Class Notes for MATH 255.

by
S. W. Drury

Contents

Prologue

LimSup and LimInf

Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. We can then define

$$y_m = \sup_{n \geq m} x_n$$

with the convention that $y_m = \infty$ if $(x_n)_{n=1}^{\infty}$ is unbounded above. The key point about y_m is that $y_\ell \geq y_m$ for $\ell \leq m$. This is because as m increases, the supremum is taken over a smaller set. Since $(y_m)_{m=1}^{\infty}$ is a decreasing sequence (in the wide sense) it is either unbounded below or it tends to a finite limit. This allows us to define

$$\limsup_{n\to\infty} x_n = \begin{cases} \infty & \text{if } (x_n)_{n=1}^{\infty} \text{ is unbounded above,} \\ -\infty & \text{if } (y_m)_{m=1}^{\infty} \text{ is unbounded below,} \\ \lim_{m\to\infty} y_m & \text{otherwise.} \end{cases}$$

The advantage of the limsup over the usual limit is that it always exists, but with the drawback that it may possibly take the values ∞ or $-\infty$. In the same way, we define $z_m = \inf_{n \geq m} x_n$ and

$$\liminf_{n\to\infty} x_n = \begin{cases} -\infty & \text{if } (x_n)_{n=1}^{\infty} \text{ is unbounded below,} \\ \infty & \text{if } (z_m)_{m=1}^{\infty} \text{ is unbounded above,} \\ \lim_{m\to\infty} z_m & \text{otherwise.} \end{cases}$$

The following Theorem lays out the basic properties of limsup and liminf.

THEOREM 1 *Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. Then*

1. $\liminf_{n\to\infty} x_n \leq \limsup_{n\to\infty} x_n$.
2. *If* $\lim_{n\to\infty} x_n$ *exists, then*

$$\liminf_{n\to\infty} x_n = \lim_{n\to\infty} x_n = \limsup_{n\to\infty} x_n. \tag{0.1}$$

3. *If* $\liminf_{n\to\infty} x_n = \limsup_{n\to\infty} x_n$ *is a finite quantity, then* $\lim_{n\to\infty} x_n$ *exists and* (0.1) *holds.*

Proof. For the first statement, we first get rid of the infinite cases. If $(x_n)_{n=1}^{\infty}$ is unbounded above, then $\limsup_{n\to\infty} x_n = \infty$ and there is nothing to show. Similarly if $(x_n)_{n=1}^{\infty}$ is unbounded below, there is nothing to show. So, we can assume that (x_n) is bounded and we need only point out that $z_n \leq x_n \leq y_n$. Passing to the limit in $z_n \leq y_n$ gives the desired result.

For the second statement, suppose that $\lim_{n\to\infty} x_n = x$. Let $\epsilon > 0$. Then there exists $N \in \mathbb{N}$ such that

$$n \geq N \Longrightarrow |x - x_n| < \epsilon$$

and it follows that

$$n \geq N \Longrightarrow |x - y_n| \leq \epsilon \text{ and } |x - z_n| \leq \epsilon.$$

Passing to the limit yields

$$|x - \limsup_{n\to\infty} x_n| \leq \epsilon \text{ and } |x - \liminf_{n\to\infty} x_n| \leq \epsilon,$$

and the desired conclusion follows.

For the last statement, it suffices to observe again that $z_n \leq x_n \leq y_n$ and apply the Squeeze Lemma. ■

Some examples would be a good idea.

EXAMPLE Let

$$x_n = \begin{cases} n & \text{if } n \text{ is even,} \\ -n^{-1} & \text{if } n \text{ is odd.} \end{cases}$$

In other words (x_n) is the sequence $-1, 2, -1/3, 4, -1/5, 6, -1/7, 8, \ldots$ and we find that (y_n) is identically infinite and (z_n) the sequence $-1, -1/3, -1/3, -1/5, -1/5, -1/7, -1/7, -1/9, \ldots$ which converges to 0. Of course (x_n) does not converge. □

EXAMPLE Let

$$x_n = \begin{cases} 2 + n^{-1} & \text{if } n \text{ is even,} \\ -n^{-1} & \text{if } n \text{ is odd.} \end{cases}$$

Then (x_n) is the sequence $-1, 5/2, -1/3, 9/4, -1/5, 13/6, -1/7, 17/8, \ldots$ and we find that (y_n) is the sequence $5/2$, $5/2$, $9/4$, $9/4$, $13/6$, $13/6$, $17/8$, $17/8, \ldots$ which converges to 2 and (z_n) the sequence -1, $-1/3$, $-1/3$, $-1/5$, $-1/5$, $-1/7$, $-1/7$, $-1/9, \ldots$ which converges to 0. Again (x_n) does not converge. □

EXAMPLE Let

$$x_n = \begin{cases} 2 - n^{-1} & \text{if } n \text{ is even,} \\ -n^{-1} & \text{if } n \text{ is odd.} \end{cases}$$

Very similar to the previous example, except that now (y_n) is the constant sequence equal to 2. Again (x_n) does not converge. □

EXAMPLE Finally let $x_n = -n^{-1}$. Now (y_n) is the constant sequence equal to 0, (z_n) is the same as (x_n) and we do have convergence of (x_n) to zero. □

There are other ways of understanding the limsup and liminf, but perhaps the next question to answer is why do we need it? The answer is that it allows us to write certain types of proof very succinctly. If it was not available, we would have to jump through hoops to express ourselves. Here is an example.

LEMMA 2 *Let $a_n > 0$ and suppose that $\lim_{n \longrightarrow \infty} \frac{a_{n+1}}{a_n} = \rho$. Show that*

$$\lim_{n \longrightarrow \infty} \Big(a_n\Big)^{\frac{1}{n}} = \rho.$$

Proof. Let $\epsilon > 0$. Then, by hypothesis, there exists $N \in \mathbb{N}$ such that $n \geq N$ implies that $\left| \frac{a_{n+1}}{a_n} - \rho \right| < \epsilon$. This gives $\frac{a_{n+1}}{a_n} < \rho + \epsilon$ and a straightforward induction argument shows that $\frac{a_n}{a_N} \leq (\rho + \epsilon)^{n-N}$ for $n \geq N$. So

$$a_n \leq \frac{a_N}{(\rho + \epsilon)^N} (\rho + \epsilon)^n,$$

for $n \geq N$. Thus, taking nth roots, we find that

$$(a_n)^{\frac{1}{n}} \leq (\rho + \epsilon) \left(\frac{a_N}{(\rho + \epsilon)^N} \right)^{\frac{1}{n}},$$

again for $n \geq N$. Taking the $\limsup$ of both sides now readily gives

$$\limsup_{n \to \infty} (a_n)^{\frac{1}{n}} \leq (\rho + \epsilon) \limsup_{n \to \infty} \left(\frac{a_N}{(\rho + \epsilon)^N} \right)^{\frac{1}{n}} = \rho + \epsilon.$$

Now, suppose that $\rho > 0$. We will choose ϵ such that $\rho > \epsilon > 0$. Then an entirely similar argument shows that

$$\liminf_{n \to \infty} (a_n)^{\frac{1}{n}} \geq (\rho - \epsilon) \liminf_{n \to \infty} \left(\frac{a_N}{(\rho - \epsilon)^N} \right)^{\frac{1}{n}} = \rho - \epsilon.$$

On the other hand, if $\rho = 0$ then since $a_n > 0$ we have

$$\liminf_{n \to \infty} (a_n)^{\frac{1}{n}} \geq 0.$$

Combining the above results gives

$$\rho - \epsilon \leq \liminf_{n \to \infty} (a_n)^{\frac{1}{n}} \leq \limsup_{n \to \infty} (a_n)^{\frac{1}{n}} \leq \rho + \epsilon.$$

Since ϵ is a positive number that can be taken as small as we please, we are able to conclude that

$$\liminf_{n \to \infty} (a_n)^{\frac{1}{n}} = \limsup_{n \to \infty} (a_n)^{\frac{1}{n}} = \rho,$$

and the result follows. ■

There are two other useful ways of understanding the limsup and liminf.

LEMMA 3 *We have for any sequence (x_n) of real numbers.*

- $\limsup_{n \to \infty} x_n = \inf\{t; \{n; x_n > t\}$ *is finite*$\}$.
- $\liminf_{n \to \infty} x_n = \sup\{t; \{n; x_n < t\}$ *is finite*$\}$.

Here, we need to have some understanding of the conventions to be used in special cases. In the first statement, if $\{n; x_n > t\}$ is infinite for all t, then the infimum (of any empty set) is interpreted as ∞. If $\{n; x_n > t\}$ is finite for all t, then the infimum (of $\mathbb{R}$) is interpreted as $-\infty$. Similar conventions apply also to $\{n; x_n < t\}$.

Proof. We just prove the first statement. The second is similar. If (x_n) is unbounded above, the result is evident. Let $t = \limsup_{n\to\infty} x_n$, $A = \{t; \{n; x_n > t\} \text{ is finite}\}$ and as before $y_n = \sup_{m\geq n} x_m$. Keep in mind that $t = -\infty$ is a possible case. Now $y_n \downarrow t$, so given $s > t$, $\exists N$ such that $y_N < s$. So $\{n; x_n > s\} \subseteq \{1, 2, \ldots, N\}$ and it follows that $s \in A$. Since s is an arbitrary number with $s > t$ it follows that $\inf A \leq t$. Conversely, if $\inf A < t$ we may find s with $\inf A < s < t$ and $\{n; x_n > s\}$ finite. But then there exists $N \in \mathbb{N}$ such that $n \geq N \Longrightarrow x_n \leq s$. It follows that $y_N \leq s$. But $s < t$ and (y_n) decreases to t, a contradiction. ■

The final way of thinking about limsup and liminf is by means of the limit set L. This idea works only for bounded sequences. For any bounded sequence (x_n) of real numbers, we say that x is a ***limit point*** if and only if there exists a natural subsequence (n_k) such that $x_{n_k} \underset{k\to\infty}{\longrightarrow} x$. The set L is the set of all limit points.

LEMMA 4 *Then we have* $\limsup_{n\to\infty} x_n = \sup L$ *and* $\liminf_{n\to\infty} x_n = \inf L$.

Proof. Again, we prove only the first statement. Let $t = \limsup_{n\to\infty} x_n$ and let s be arbitrary with $s > t$. Then as seen above, $\{n; x_n > s\}$ is finite and it follows that s is an upper bound for L. This shows that $\sup L \leq t$.

In the opposite direction, let $s < t$. Then $y_n = \sup_{m\geq n} x_m > s$ for all n. Since $y_1 > s$, there exists $n_1 \geq 1$ such that $x_{n_1} > s$. Since $y_{n_1+1} > s$, there exists $n_2 > n_1$ such that $x_{n_2} > s$. Since $y_{n_2+1} > s$, there exists $n_3 > n_2$ such that $x_{n_3} > s \ldots$

We have found a subsequence (x_{n_k}) with $x_{n_k} > s$ for all k. Since this subsequence is bounded, we can extract from it a further subsequence which actually converges using the Bolzano–Weierstrass Theorem. The limiting value is necessarily $\geq s$. So, there exists $u \in L$ with $u \geq s$. Hence $\sup L \geq t$. ■

1

Metric Spaces and Analysis in Several Variables

1.1 Metric Spaces

In this section we introduce the concept of a ***metric space*** . A metric space is simply a set together with a distance function which measures the distance between any two points of the space. Starting from the distance function it is possible to introduce all the concepts we dealt with last semester to do with convergent sequences, continuity and limits etc. Thus in order to have the concept of convergence in a certain set of objects (5×5 real matrices for example), it suffices to have a concept of distance between any two such objects.

Our objective here is not to exhaustively study metric spaces: that is covered in Analysis III. We just want to introduce the basic ideas without going too deeply into the subject.

DEFINITION *A **metric space** (X, d) is a set X together with a **distance function** or **metric** $d : X \times X \longrightarrow \mathbb{R}^+$ satisfying the following properties.*

- $d(x, x) = 0 \qquad \forall x \in X.$
- $x, y \in X, d(x, y) = 0 \quad \Rightarrow \quad x = y.$
- $d(x, y) = d(y, x) \qquad \forall x, y \in X.$
- $d(x, z) \leq d(x, y) + d(y, z) \qquad \forall x, y, z \in X.$

The fourth axiom for a distance function is called the ***triangle inequality*** . It is easy to derive the ***extended triangle inequality***

$$d(x_1, x_n) \leq d(x_1, x_2) + d(x_2, x_3,) + \cdots + d(x_{n-1}, x_n) \quad \forall x_1, \ldots, x_n \in X \quad (1.1)$$

directly from the axioms.

Sometimes we will abuse notation and say that X is a metric space when the intended distance function is understood.

The real line $\mathbb{R}$ is a metric space with the distance function $d(x, y) = |x - y|$.

A simple construction allows us to build new metric spaces out of existing ones. Let X be a metric space and let $Y \subseteq X$. Then the restriction of the distance function of X to the subset $Y \times Y$ of $X \times X$ is a distance function on Y. Sometimes this is called the ***restriction metric*** or the ***relative metric***. If the four axioms listed above hold for all points of X then *a fortiori* they hold for all points of Y. Thus every subset of a metric space is again a metric space in its own right.

We can construct more interesting examples from vector spaces.

1.2 Normed Spaces

We start by introducing the concept of a ***norm***. This generalization of the absolute value on $\mathbb{R}$ (or $\mathbb{C}$) to the framework of vector spaces is central to modern analysis.

The zero element of a vector space V (over $\mathbb{R}$ or $\mathbb{C}$) will be denoted 0_V. For an element v of the vector space V the norm of v (denoted $\|v\|$) is to be thought of as the distance from 0_V to v, or as the “size” or “length” of v. In the case of the absolute value on the field of scalars, there is really only one possible candidate, but in vector spaces of more than one dimension a wealth of possibilities arises.

DEFINITION *A* ***norm*** *on a vector space* V *over* $\mathbb{R}$ *or* $\mathbb{C}$ *is a mapping*

$$v \longrightarrow \|v\|$$

from V *to* $\mathbb{R}^+$ *with the following properties.*

- $\|0_V\| = 0$.
- $v \in V, \|v\| = 0 \Rightarrow v = 0_V$.
- $\|tv\| = |t|\|v\| \qquad \forall t$ *a scalar,* $v \in V$.
- $\|v_1 + v_2\| \leq \|v_1\| + \|v_2\| \qquad \forall v_1, v_2 \in V$.

The last of these conditions is called the ***subadditivity inequality***. There are really two definitions here, that of a ***real norm*** applicable to real vector spaces and that of a ***complex norm*** applicable to complex vector spaces. However, every complex vector space can also be considered as a real vector space — one simply "forgets" how to multiply vectors by complex scalars that are not real scalars. This process is called ***realification*** . In such a situation, the two definitions are different. For instance,

$$\|x+iy\| = \max(|x|, 2|y|) \qquad (x, y \in \mathbb{R})$$

defines a perfectly good real norm on $\mathbb{C}$ considered as a real vector space. On the other hand, the only complex norms on $\mathbb{C}$ have the form

$$\|x+iy\| = t(x^2+y^2)^{\frac{1}{2}}$$

for some $t > 0$.

The inequality

$$\|t_1v_1 + t_2v_2 + \cdots + t_nv_n\| \le |t_1|\|v_1\| + |t_2|\|v_2\| + \cdots + |t_n|\|v_n\|$$

holds for scalars $t_1, \ldots, t_n$ and elements $v_1, \ldots, v_n$ of V. It is an immediate consequence of the definition.

If $\| \quad \|$ is a norm on V and $t > 0$ then

$$|||v||| = t\|v\|$$

defines a new norm $||| \quad |||$ on V. We note that in the case of a norm there is often no natural way to normalize it. On the other hand, an absolute value is normalized so that $|1| = 1$, possible since the field of scalars contains a distinguished element 1.

1.3 Some Norms on Euclidean Space

Because of the central role of $\mathbb{R}^n$ as a vector space it is worth looking at some of the norms that are commonly defined on this space.

EXAMPLE On $\mathbb{R}^n$ we may define a norm by

$$\|(x_1, \ldots, x_n)\|_\infty = \max_{j=1}^{n} |x_j|. \tag{1.2}$$

□

EXAMPLE Another norm on $\mathbb{R}^n$ is given by

$$\|(x_1, \ldots, x_n)\|_1 = \sum_{j=1}^{n} |x_j|.$$

□

EXAMPLE The ***Euclidean norm*** on $\mathbb{R}^n$ is given by

$$\|(x_1, \ldots, x_n)\|_2 = \left(\sum_{j=1}^{n} |x_j|^2 \right)^{\frac{1}{2}}.$$

This is the standard norm, representing the standard Euclidean distance to $\mathbf{0}$. The symbol $\mathbf{0}$ will be used to denote the zero vector of $\mathbb{R}^n$ or $\mathbb{C}^n$. □

These examples can be generalized by defining in case $1 \le p < \infty$

$$\|(x_1, \ldots, x_n)\|_p = \left(\sum_{j=1}^{n} |x_j|^p \right)^{\frac{1}{p}}.$$

In case that $p = \infty$ we use (1.2) to define $\| \ \|_\infty$. It is true that $\| \ \|_p$ is a norm on $\mathbb{R}^n$, but we will not prove this fact here.

1.4 Inner Product Spaces

Inner product spaces play a very central role in analysis. They have many applications. For example the physics of Quantum Mechanics is based on inner product spaces. In this section we only scratch the surface of the subject.

DEFINITION *A **real inner product space** is a real vector space V together with an inner product. An **inner product** is a mapping from $V \times V$ to $\mathbb{R}$ denoted by*

$$(v_1, v_2) \longrightarrow \langle v_1, v_2 \rangle$$

and satisfying the following properties

- $\langle w, t_1 v_1 + t_2 v_2 \rangle = t_1 \langle w, v_1 \rangle + t_2 \langle w, v_2 \rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{R}.$
- $\langle v_1, v_2 \rangle = \langle v_2, v_1 \rangle \qquad \forall v_1, v_2 \in V.$

- $\langle v, v\rangle \geq 0 \qquad \forall v \in V.$
- *If* $v \in V$ *and* $\langle v, v\rangle = 0$, *then* $v = 0_V$.

The symmetry and the linearity in the second variable implies that the inner product is also linear in the first variable.

$$\langle t_1v_1 + t_2v_2, w\rangle = t_1\langle v_1, w\rangle + t_2\langle v_2, w\rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{R}.$$

EXAMPLE The ***standard inner product*** on $\mathbb{R}^n$ is given by

$$\langle x, y\rangle = \sum_{j=1}^{n} x_j y_j$$

□

The most general inner product on $\mathbb{R}^n$ is given by

$$\langle x, y\rangle = \sum_{j=1}^{n}\sum_{k=1}^{n} p_{j,k}x_j y_k$$

where the $n \times n$ real matrix $P = (p_{j,k})$ is a ***positive definite*** matrix. This means that

- P is a symmetric matrix.
- We have
$$\sum_{j=1}^{n}\sum_{k=1}^{n} p_{j,k}x_j x_k \geq 0$$
for every vector $(x_1, \ldots, x_n)$ of $\mathbb{R}^n$.
- The circumstance
$$\sum_{j=1}^{n}\sum_{k=1}^{n} p_{j,k}x_j x_k = 0$$
only occurs when $x_1 = 0, \ldots, x_n = 0$.

In the complex case, the definition is slightly more complicated.

DEFINITION *A **complex inner product space** is a complex vector space V together with a **complex inner product** , that is a mapping from $V \times V$ to $\mathbb{C}$ denoted*

$$(v_1, v_2) \longrightarrow \langle v_1, v_2 \rangle$$

and satisfying the following properties

- $\langle w, t_1 v_1 + t_2 v_2 \rangle = t_1 \langle w, v_1 \rangle + t_2 \langle w, v_2 \rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{C}.$
- $\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle} \qquad \forall v_1, v_2 \in V.$
- $\langle v, v \rangle \geq 0 \qquad \forall v \in V.$
- *If $v \in V$ and $\langle v, v \rangle = 0$, then $v = 0_V$.*

It will be noted that a complex inner product is linear in its second variable and conjugate linear in its first variable.

$$\langle t_1 v_1 + t_2 v_2, w \rangle = \overline{t_1} \langle v_1, w \rangle + \overline{t_2} \langle v_2, w \rangle \qquad \forall w, v_1, v_2 \in V, t_1, t_2 \in \mathbb{C}.$$

EXAMPLE The standard inner product on $\mathbb{C}^n$ is given by

$$\langle x, y \rangle = \sum_{j=1}^{n} \overline{x_j} y_j$$

□

The most general inner product on $\mathbb{C}^n$ is given by

$$\langle x, y \rangle = \sum_{j=1}^{n} \sum_{k=1}^{n} p_{j,k} \overline{x_j} y_k$$

where the $n \times n$ complex matrix $P = (p_{j,k})$ is a ***positive definite*** matrix. This means that

- P is a hermitian matrix, in other words $p_{jk} = \overline{p_{kj}}$.
- We have
$$\sum_{j=1}^{n} \sum_{k=1}^{n} p_{j,k} \overline{x_j} x_k \geq 0$$
for every vector $(x_1, \ldots, x_n)$ of $\mathbb{C}^n$.

- The circumstance
$$\sum_{j=1}^{n}\sum_{k=1}^{n} p_{j,k}\overline{x_j}x_k = 0$$
only occurs when $x_1 = 0, \ldots, x_n = 0$.

DEFINITION *Let V be an inner product space. Then we define*

$$\|v\| = (\langle v, v\rangle)^{\frac{1}{2}} \tag{1.3}$$

the ***associated norm*** *.*

It is not immediately clear from the definition that the associated norm satisfies the subadditivity condition. Towards this, we establish the abstract Cauchy-Schwarz inequality.

PROPOSITION 5 (CAUCHY-SCHWARZ INEQUALITY) *Let V be an inner product space and $u, v \in V$. Then*

$$|\langle u, v\rangle| \leq \|u\|\|v\| \tag{1.4}$$

holds.

Proof of the Cauchy-Schwarz Inequality. We give the proof in the complex case. The proof in the real case is slightly easier. If $v = 0_V$ then the inequality is evident. We therefore assume that $\|v\| > 0$. Similarly, we may assume that $\|u\| > 0$.

Let $t \in \mathbb{C}$. Then we have

$$\begin{aligned} 0 \leq \|u + tv\|^2 &= \langle u + tv, u + tv\rangle \\ &= \langle u, u\rangle + \bar{t}\langle v, u\rangle + t\langle u, v\rangle + t\bar{t}\langle v, v\rangle \\ &= \|u\|^2 + 2\Re t\langle u, v\rangle + |t|^2\|v\|^2. \end{aligned} \tag{1.5}$$

Now choose t such that

$$t\langle u, v\rangle \text{ is real and } \leq 0 \tag{1.6}$$

and

$$|t| = \frac{\|u\|}{\|v\|}. \tag{1.7}$$

Here, (1.7) designates the absolute value of t and (1.6) specifies its argument. Substituting back into (1.5) we obtain

$$2\frac{\|u\|}{\|v\|}|\langle u, v\rangle| \leq \|u\|^2 + \left(\frac{\|u\|}{\|v\|}\right)^2 \|v\|^2$$

which simplifies to the desired inequality (1.4). ■

PROPOSITION 6 *In an inner product space* (1.3) *defines a norm.*

Proof. We verify the subadditivity of $v \longrightarrow \|v\|$. The other requirements of a norm are straightforward to establish. We have

$$\begin{aligned}
\|u+v\|^2 &= \langle u+v, u+v\rangle \\
&= \|u\|^2 + \langle v, u\rangle + \langle u, v\rangle + \|v\|^2 \\
&= \|u\|^2 + 2\Re\langle u, v\rangle + \|v\|^2 \\
&\leq \|u\|^2 + 2|\Re\langle u, v\rangle| + \|v\|^2 \\
&\leq \|u\|^2 + 2|\langle u, v\rangle| + \|v\|^2 \\
&\leq \|u\|^2 + 2\|u\|\|v\| + \|v\|^2 \\
&= (\|u\| + \|v\|)^2 \qquad (1.8)
\end{aligned}$$

using the Cauchy-Schwarz Inequality (1.4). Taking square roots yields

$$\|u+v\| \leq \|u\| + \|v\|$$

as required. ■

1.5 Geometry of Norms

It is possible to understand the concept of norm from the geometrical point of view. Towards this we associate with each norm a geometrical object — its unit ball.

DEFINITION *Let V be a normed vector space. Then the* ***unit ball*** *B of V is defined by*

$$B = \{v; v \in V, \|v\| \leq 1\}.$$

DEFINITION *Let V be a vector space and let $B \subseteq V$. We say that B is* ***convex*** *iff*

$$t_1 v_1 + t_2 v_2 \in B \qquad \forall v_1, v_2 \in B,\ \forall t_1, t_2 \geq 0 \text{ such that } t_1 + t_2 = 1.$$

In other words, a set B is convex iff whenever we take two points of B, the line segment joining them lies entirely in B.

DEFINITION *Let V be a vector space and let $B \subseteq V$. We say that B satisfies the* ***line condition*** *iff for every $v \in V \setminus \{0_V\}$, there exists a constant $a \in]0, \infty[$ such that*

$$tv \in B \quad \Leftrightarrow \quad |t| \leq a.$$

The line condition says that the intersection of B with every one-dimensional subspace R of V is the unit ball for some norm on R. The line condition involves a multitude of considerations. It implies that the set B is ***symmetric*** about the zero element. The fact that $a > 0$ is sometimes expressed by saying that B is ***absorbing*** . This expresses the fact that every point v of V lies in some (large) multiple of B. Finally the fact that $a < \infty$ is a ***boundedness condition*** .

The following theorem gives a geometrical way of understanding norms. We will not prove this Theorem.

THEOREM 7 *Let V be a vector space and let $B \subseteq V$. Then the following two statements are equivalent.*

- *There is a norm on V for which B is the unit ball.*
- *B is convex and satisfies the line condition.*

EXAMPLE Let us define the a subset B of $\mathbb{R}^2$ by

$$(x, y) \in B \text{ if } \begin{cases} x^2 + y^2 \leq 1 & \text{in case } x \geq 0 \text{ and } y \geq 0, \\ \max(-x, y) \leq 1 & \text{in case } x \leq 0 \text{ and } y \geq 0, \\ x^2 + y^2 \leq 1 & \text{in case } x \leq 0 \text{ and } y \leq 0, \\ \max(x, -y) \leq 1 & \text{in case } x \geq 0 \text{ and } y \leq 0. \end{cases}$$

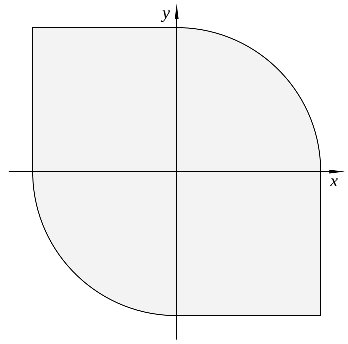

Figure 1.1: The unit ball for a norm on $\mathbb{R}^2$.

It is geometrically obvious that B is a convex subset of $\mathbb{R}^2$ and satisfies the line condition — see Figure 1.1. Therefore it defines a norm. Clearly this norm is given by

$$\|(x,y)\| = \begin{cases} (x^2+y^2)^{\frac{1}{2}} & \text{if } x \geq 0 \text{ and } y \geq 0, \\ \max(|x|,|y|) & \text{if } x \leq 0 \text{ and } y \geq 0, \\ (x^2+y^2)^{\frac{1}{2}} & \text{if } x \leq 0 \text{ and } y \leq 0, \\ \max(|x|,|y|) & \text{if } x \geq 0 \text{ and } y \leq 0. \end{cases}$$

□

1.6 Examples of Metric Spaces

In the previous section we discussed the concept of the norm of a vector. In a normed vector space, the expression $\|u - v\|$ represents the size of the difference $u - v$ of two vectors u and v. It can be thought of as the distance between u and v. Just as a vector space may have many possible norms, there can be many possible concepts of distance.

EXAMPLE Let V be a normed vector space with norm $\|\ \|$. Then V is a metric space with the distance function

$$d(u,v) = \|u - v\|.$$

We check that the triangle inequality is a consequence of the subadditivity of the norm.

$$d(u,w) = \|u-w\| = \|(u-v)+(v-w)\| \le \|u-v\|+\|v-w\| = d(u,v)+d(v,w)$$

□

EXAMPLE It follows that every subset X of a normed vector space is a metric space in the distance function induced from the norm. □

1.7 Neighbourhoods and Open Sets

It is customary to refer to the elements of a metric space as ***points*** . In the remainder of this chapter we will develop the ***point-set topology*** of metric spaces. This is done through concepts such as ***neighbourhoods*** , ***open sets*** , ***closed sets*** and ***sequences*** . Any of these concepts can be used to define more advanced concepts such as the continuity of mappings from one metric space to another. They are, as it were, languages for the further development of the subject. We study them all and most particularly the relationships between them.

DEFINITION *Let (X,d) be a metric space. For $t > 0$ and $x \in X$, we define*

$$U(x,t) = \{y; y \in X, d(x,y) < t\}$$

and

$$B(x,t) = \{y; y \in X, d(x,y) \le t\}.$$

the ***open ball*** *$U(x,t)$ centred at x of radius t and the corresponding* ***closed ball*** *$B(x,t)$.*

DEFINITION *Let V be a subset of a metric space X and let $x \in V$. Then we say that V is a* ***neighbourhood*** *of x or x is an* ***interior point*** *of V iff there exists $t > 0$ such that $U(x,t) \subseteq V$.*

Thus V is a neighbourhood of x iff all points sufficiently close to x lie in V.

PROPOSITION 8

- *If V is a neighbourbood of x and $V \subseteq W \subseteq X$. Then W is a neighbourhood of x.*

- *If $V_1, V_2, \ldots, V_n$ are finitely many neighbourhoods of x, then $\bigcap_{j=1}^n V_j$ is also a neighbourhood of x.*

Proof. For the first statement, since V is a neighbourhood of x, there exists t with $t > 0$ such that $U(x,t) \subseteq V$. But $V \subseteq W$, so $U(x,t) \subseteq W$. Hence W is a neighbourhood of x. For the second, applying the definition, we may find $t_1, t_2, \ldots, t_n > 0$ such that $U(x, t_j) \subseteq V_j$. It follows that

$$\bigcap_{j=1}^{n} U(x, t_j) \subseteq \bigcap_{j=1}^{n} V_j. \tag{1.9}$$

But the left-hand side of (1.9) is just $U(x,t)$ where $t = \min t_j > 0$. It now follows that $\bigcap_{j=1}^n V_j$ is a neighbourhood of x. ■

Neighbourhoods are a local concept. We now introduce the corresponding global concept.

DEFINITION *Let (X, d) be a metric space and let $V \subseteq X$. Then V is an* ***open subset*** *of X iff V is a neighbourhood of every point x that lies in V.*

EXAMPLE For all $t > 0$, the open ball $U(x,t)$ is an open set. To see this, let $y \in U(x,t)$, that is $d(x,y) < t$. We must show that $U(x,t)$ is a neighbourhood of y. Let $s = t - d(x,y) > 0$. We claim that $U(y,s) \subseteq U(x,t)$. To prove the claim, let $z \in U(y,s)$. Then $d(y,z) < s$. We now find that

$$d(x,z) \leq d(x,y) + d(y,z) < d(x,y) + s = t,$$

so that $z \in U(x,t)$ as required. □

EXAMPLE An almost identical argument to that used in the previous example shows that for all $t > 0$, $\{y; y \in X, d(x,y) > t\}$ is an open set. □

EXAMPLE In $\mathbb{R}$ every interval of the form $]a,b[$ is an open set. Here, a and b are real and satisfy $a < b$. We also allow the possibilities $a = -\infty$ and $b = \infty$. □

THEOREM 9 *In a metric space (X, d) we have*

- *X is an open subset of X.*

- *$\emptyset$ is an open subset of X.*

- *If V_α is open for every α in some index set I, then $\cup_{\alpha\in I}V_\alpha$ is again open.*

- *If V_j is open for $j = 1, \ldots, n$, then the finite intersection $\cap_{j=1}^{n} V_j$ is again open.*

Proof. For every $x \in X$ and any $t > 0$, we have $U(x, t) \subseteq X$, so X is open. On the other hand, $\emptyset$ is open because it does not have any points. Thus the condition to be checked is vacuous.

To check the third statement, let $x \in \cup_{\alpha\in I}V_\alpha$. Then there exists $\alpha \in I$ such that $x \in V_\alpha$. Since V_α is open, V_α is a neighbourhood of x. The result now follows from the first part of Proposition 8.

Finally let $x \in \cap_{j=1}^{n} V_j$. Then since V_j is open, it is a neighbourhood of x for $j = 1, \ldots, n$. Now apply the second part of Proposition 8. ∎

DEFINITION *Let X be a set. Let $\mathcal{V}$ be a "family of open sets" satisfying the four conditions of Theorem 9. Then $\mathcal{V}$ is a* ***topology*** *on X and $(X, \mathcal{V})$ is a* ***topological space*** *.*

Not every topology arises from a metric. In these notes we are *not* concerned with topological spaces. This is an advanced concept.

1.8 The Open subsets of $\mathbb{R}$

It is worth recording here that there is a complete description of the open subsets of $\mathbb{R}$. A subset V of $\mathbb{R}$ is open iff it is a disjoint union of open intervals (possibly of infinite length). Furthermore, such a union is necessarily countable. In order to discuss the proof properly, we need to review the concept of an equivalence relation.

DEFINITION *Let X be a set. An* ***equivalence relation*** *on X is a relation $\sim$ that enjoys the following properties*

- *$x \sim x$ for all $x \in X$. (This is called reflexivity).*

- *If $x \sim y$, then $y \sim x$. (This is called symmetry).*

- *If $x \sim y$ and $y \sim z$, then $x \sim z$. (This is called transitivity).*

For $x \sim y$ read x is equivalent to y. The simplest example of an equivalence relation is equality, i.e. $x \sim y$ if and only if $x = y$. Another way of making equivalence relations is to consider an "attribute" of the elements of the set X. Then we say that two elements are equivalent if they have the same attribute. For an example of this, let X be the set of all students in the class and let the attribute be the colour of the student's shirt. So, in this example, two students are "equivalent" if they are wearing the same colour shirt. We introduce equivalence relations when we can decide when two elements have equal attributes (i.e. are equivalent) but we do not yet have a handle on the attribute itself. In the example at hand, we can decide when two students have the same colour shirt, but the concept "colour of shirt" is something that we have not yet defined.

The following theorem says informally that every equivalence relation can be defined in terms of an attribute.

THEOREM 10 *Let X be a set and let $\sim$ be an equivalence relation on X. Then there is a set Q and a surjective mapping $\pi : X \longrightarrow Q$ such that $x \sim y$ if and only if $\pi(x) = \pi(y)$.*

The mapping π is called the ***canonical projection*** . So again in our example, the set Q is the set of all colours of shirts of students in the class and the mapping π is the mapping which maps a student to the colour of his/her shirt. For a given $q \in Q$ we can also define $\pi^{-1}(\{q\})$ which is the subset of all elements x of X which get mapped to q. This is an ***equivalence class*** . So the elements of Q are in one-to-one correspondence with the equivalence classes.

If we have two equivalence classes, then they are either equal or disjoint. This is just the statement that if $q_1, q_2 \in Q$, then either $q_1 = q_2$ in which case $\pi^{-1}(\{q_1\}) = \pi^{-1}(\{q_2\})$ or else $q_1 \neq q_2$ which gives $\pi^{-1}(\{q_1\}) \cap \pi^{-1}(\{q_2\}) = \emptyset$. Also, every point of X is in some equivalence class, so effectively, the equivalence classes partition the space X.

Proof of Theorem 10. We define the equivalence classes from the equivalence relation. Let ρ be the mapping from X to $\mathcal{P}X$ (the power set of X) given by

$$\rho(x) = \{y; y \in X, y \sim x\}.$$

Note that since $x \sim x$, $x \in \rho(x)$. So, in fact, $\rho(x)$ is the equivalence class to which x belongs. Now let $Q = \rho(X)$ and let π be the mapping $\pi : X \longrightarrow Q$ given by $\pi(x) = \rho(x)$. Now we still have to show that $\pi(x) = \pi(y)$ if and only if $x \sim y$.

If $\pi(x) = \pi(y)$, then since $x \in \pi(x)$, $x \in \pi(y)$ and therefore $x \sim y$ by definition of $\pi(y)$. Conversely suppose that $x \sim y$. If $z \in \pi(x)$ then $z \sim x$. So by the transitivity axiom, we have $z \sim y$ or equivalently $z \in \pi(y)$. We have shown $\pi(x) \subseteq \pi(y)$. But by symmetry we will also have $y \sim x$ and this will lead to $\pi(y) \subseteq \pi(x)$ by the same argument. Hence $\pi(x) = \pi(y)$ as required. ■

We can now attempt to prove our theorem about open subsets of $\mathbb{R}$.

THEOREM 11 *Every open subset U of $\mathbb{R}$ is a disjoint countable union of open intervals.*

Proof. For $a, b \in U$ define

$$L(a,b) = \begin{cases} [a,b] & \text{if } a < b, \\ [b,a] & \text{if } b < a, \\ \{a\} & \text{if } a = b. \end{cases}$$

The equivalence relation on U that we now introduce is $a \sim b$ if and only if $L(a,b) \subseteq U$. Since $a \in U$ we have $a \sim a$. Since $L(a,b) = L(b,a)$ we have symmetry. Transitivity is trickier. Let $a \sim b$ and $b \sim c$. Then $L(a,b) \subseteq U$ and $L(b,c) \subseteq U$. It is easy, but long to show that $L(a,c) \subseteq L(a,b) \bigcup L(b,c)$ and this yields the transitivity. Let Q be the set of equivalence classes. Then the sets $\pi^{-1}\{q\}$ are disjoint as q runs over Q and their union is U. We show that these sets are open intervals.

So, fix $q \in Q$ and let $a, b \in \pi^{-1}\{q\}$. Now let $c \in L(a,b)$. Then, since $a \sim b$, $L(a,b) \subseteq U$. Hence $L(a,c) \subseteq L(a,b) \subseteq U$ and $a \sim c$. Hence $c \in \pi^{-1}\{q\}$. This shows that $\pi^{-1}\{q\}$ is an interval. Next, we show that it is open. Let $a \in \pi^{-1}\{q\}$. Then since U is open, there exists $\delta > 0$ such that $[a-\delta, a+\delta] \subseteq U$. But then, $b \in [a-\delta, a+\delta]$ implies that $L(a,b) \subseteq U$ and hence that $b \in \pi^{-1}\{q\}$. So a is an interior point of $\pi^{-1}\{q\}$ and we conclude that $\pi^{-1}\{q\}$ is open.

Finally, we need to show that Q is countable. Let $q \in Q$. Since $\pi^{-1}\{q\}$ is an open nonempty interval, it contains a rational number r_q. Different $q \in Q$ yield different r_q because the corresponding $\pi^{-1}\{q\}$ are disjoint. Hence we can map Q into the set of rational numbers. Thus Q is countable.

1.9 Convergent Sequences

A sequence $x_1, x_2, x_3, \ldots$ of points of a set X is really a mapping from $\mathbb{N}$ to X. Normally, we denote such a sequence by (x_n). For $x \in X$ the sequence given by $x_n = x$ is called the ***constant sequence*** with value x.

DEFINITION *Let X be a metric space. Let (x_n) be a sequence in X. Then (x_n)* ***converges to*** *$x \in X$ iff for every $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ for all $n > N$. In this case, we write $x_n \longrightarrow x$ or*

$$x_n \overset{n\to\infty}{\longrightarrow} x.$$

Sometimes, we say that x is the ***limit*** *of (x_n). Proposition 12 below justifies the use of the indefinite article. To say that (x_n) is a* ***convergent sequence*** *is to say that there exists some $x \in X$ such that (x_n) converges to x.*

EXAMPLE Perhaps the most familiar example of a convergent sequence is the sequence

$$x_n = \frac{1}{n}$$

in $\mathbb{R}$. This sequence converges to 0. To see this, let $\epsilon > 0$ be given. Then choose a natural number N so large that $N > \epsilon^{-1}$. It is easy to see that

$$n > N \qquad \Rightarrow \qquad \left|\frac{1}{n}\right| < \epsilon$$

Hence $x_n \longrightarrow 0$. □

PROPOSITION 12 *Let (x_n) be a convergent sequence in X. Then the limit is unique.*

Proof. Suppose that x and y are both limits of the sequence (x_n). We will show that $x = y$. If not, then $d(x, y) > 0$. Let us choose $\epsilon = \frac{1}{2}d(x, y)$. Then there exist natural numbers N_x and N_y such that

$$\begin{aligned} n > N_x \qquad &\Rightarrow \qquad d(x_n, x) < \epsilon, \\ n > N_y \qquad &\Rightarrow \qquad d(x_n, y) < \epsilon. \end{aligned}$$

Choose now $n = \max(N_x, N_y) + 1$ so that both $n > N_x$ and $n > N_y$. It now follows that

$$2\epsilon = d(x, y) \leq d(x, x_n) + d(x_n, y) < \epsilon + \epsilon$$

a contradiction. ■

PROPOSITION 13 *Let X be a metric space and let (x_n) be a sequence in X. Let $x \in X$. The following conditions are equivalent to the convergence of (x_n) to x.*

- *For every neighbourhood V of x in X, there exists $N \in \mathbb{N}$ such that*

$$n > N \qquad \Rightarrow \qquad x_n \in V. \tag{1.10}$$

- *The sequence $(d(x_n, x))$ converges to 0 in $\mathbb{R}$.*

Proof. Suppose that $x_n \longrightarrow x$. For the first statement, since V is a neighbourhood of x, there exists $\epsilon > 0$ such that $U(x, \epsilon) \subseteq V$. Now applying this ϵ in the definition of convergence, we find the existence of $N \in \mathbb{N}$ such that $n > N$ implies that $d(x_n, x) < \epsilon$ or equivalently that $x_n \in U(x, \epsilon)$. Hence $n > N$ implies that $x_n \in V$. For the second statement, we see that since $d(x_n, x) \geq 0$ (distances are always nonnegative), we have $|d(x_n, x) - 0| = d(x_n, x)$. So, given $\epsilon > 0$ we have the existence of $N \in \mathbb{N}$ such that

$$n > N \quad \Longrightarrow \quad d(x_n, x) < \epsilon \quad \Longrightarrow \quad |d(x_n, x) - 0| < \epsilon.$$

This shows that $d(x_n, x) \longrightarrow 0$.

In the opposite direction, assume that the first statement holds. Let $\epsilon > 0$ and take $V = U(x, \epsilon)$ a neighbourhood of x. Then, there exists $N \in \mathbb{N}$ such that

$$n > N \quad \Longrightarrow \quad x_n \in V = U(x, \epsilon) \quad \Longrightarrow \quad d(x_n, x) < \epsilon.$$

Now assume that the second statement holds. Let $\epsilon > 0$. Then, there exists $N \in \mathbb{N}$ such that $|d(x_n, x) - 0| < \epsilon$ for $n > N$. So we have

$$n > N \quad \Longrightarrow \quad |d(x_n, x) - 0| < \epsilon \quad \Longrightarrow \quad d(x_n, x) < \epsilon.$$

■

The first item here is significant because it leads to the concept of the ***tail*** of a sequence. The sequence (t_n) defined by $t_k = x_{N+k}$ is called the Nth tail sequence of (x_n). The set of points $T_N = \{x_n; n > N\}$ is the Nth tail set. The condition (1.10) can be rewritten as $T_N \subseteq V$.

DEFINITION *Let A be a subset of a metric space X. Then A is* ***bounded*** *if either $A = \emptyset$ or if $\{d(a, x); a \in A\}$ is bounded above in $\mathbb{R}$ for some element x of X.*

The boundedness of $\{d(a, x); a \in A\}$ does not depend on the choice of x because if x' is some other element of X we always have $d(a, x') \leq d(a, x) + d(x, x')$ and $d(x, x')$ does not depend on a. In a normed vector space, we usually take the special element x to be the zero vector, so that boundedness of A is equivalent to the boundedness of $\{\|a\|, a \in A\}$ in $\mathbb{R}$.

PROPOSITION 14 *If (x_n) is a convergent sequence in a metric space X, then the underlying set $\{x_n; n \in \mathbb{N}\}$ is bounded in X.*

Proof. Let $x \in X$ be the limit point of (x_n). Then $d(x_n, x) \longrightarrow 0$ in $\mathbb{R}$, so $(d(x_n, x))_{n=1}^{\infty}$ is a convergent sequence of real numbers and hence bounded. It follows that $\{x_n; n \in \mathbb{N}\}$ is also bounded. ■

Sequences provide one of the key tools for understanding metric spaces. They lead naturally to the concept of ***closed subsets*** of a metric space.

DEFINITION *Let X be a metric space. Then a subset $A \subseteq X$ is said to be* ***closed*** *iff whenever (x_n) is a sequence in A (that is $x_n \in A \quad \forall n \in \mathbb{N}$) converging to a limit x in X, then $x \in A$.*

The link between closed subsets and open subsets is contained in the following result.

THEOREM 15 *In a metric space X, a subset A is closed if and only if $X \setminus A$ is open.*

It follows from this Theorem that U is open in X iff $X \setminus U$ is closed.

Proof. First suppose that A is closed. We must show that $X \setminus A$ is open. Towards this, let $x \in X \setminus A$. We claim that there exists $\epsilon > 0$ such that $U(x, \epsilon) \subseteq X \setminus A$. Suppose not. Then taking for each $n \in \mathbb{N}$, $\epsilon_n = \frac{1}{n}$ we find that there exists $x_n \in A \cap U(x, \frac{1}{n})$. But now (x_n) is a sequence of elements of A converging to x. Since A is closed $x \in A$. But this is a contradiction.

For the converse assertion, suppose that $X \setminus A$ is open. We will show that A is closed. Let (x_n) be a sequence in A converging to some $x \in X$. If $x \in X \setminus A$ then since $X \setminus A$ is open, there exists $\epsilon > 0$ such that

$$U(x, \epsilon) \subseteq X \setminus A. \tag{1.11}$$

But since (x_n) converges to x, there exists $N \in \mathbb{N}$ such that $x_n \in U(x, \epsilon)$ for $n > N$. Choose $n = N + 1$. Then we find that $x_n \in A \cap U(x, \epsilon)$ which contradicts (1.11). ■

Combining now Theorems 9 and 15 we have the following corollary.

COROLLARY 16 *In a metric space (X, d) we have*

- *X is an closed subset of X.*
- *$\emptyset$ is an closed subset of X.*
- *If A_α is closed for every α in some index set I, then $\cap_{\alpha \in I} A_\alpha$ is again closed.*
- *If A_j is closed for $j = 1, \ldots, n$, then the finite union $\cup_{j=1}^{n} A_j$ is again closed.*

Notice that nothing prevents a subset of a metric space from being both open and closed at the same time. In fact, the empty set and the whole space always have this property. For $\mathbb{R}$ (with the standard metric) these are the only two sets that are both open and closed. In more general metric spaces there may be others. This issue is related to the connectedness of the metric space.

EXAMPLE In a metric space every singleton is closed. To see this we remark that a sequence in a singleton is necessarily a constant sequence and hence convergent to its constant value. □

EXAMPLE Combining the previous example with the last assertion of Corollary 16, we see that in a metric space, every finite subset is closed. □

EXAMPLE Let (x_n) be a sequence converging to x. Then the set

$$\{x_n; n \in \mathbb{N}\} \cup \{x\}$$

is a closed subset. □

EXAMPLE In $\mathbb{R}$, the intervals $[a, b]$, $[a, \infty[$ and $]-\infty, b]$ are closed subsets. □

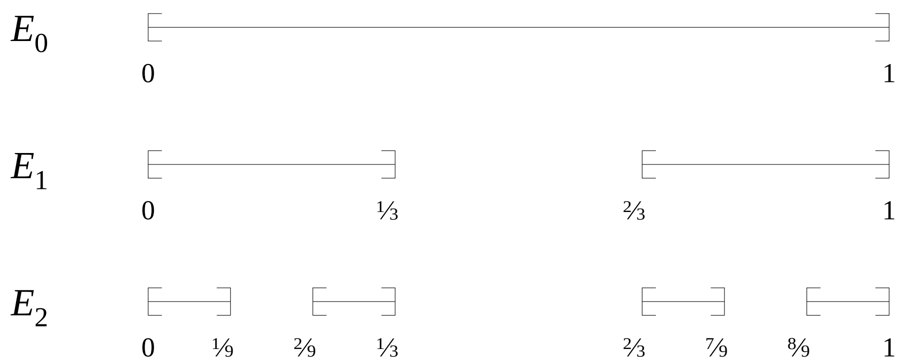

Figure 1.2: The sets E_0, E_1 and E_2.

EXAMPLE A more complicated example of a closed subset of $\mathbb{R}$ is the ***Cantor set.*** Let $E_0 = [0, 1]$. To obtain E_1 from E_0 we remove the middle third of E_0. Thus $E_1 = [0, \frac{1}{3}] \cup [\frac{2}{3}, 1]$. To obtain E_2 from E_1 we remove the middle thirds from both the constituent intervals of E_1. Thus

$$E_2 = [0, \tfrac{1}{9}] \cup [\tfrac{2}{9}, \tfrac{1}{3}] \cup [\tfrac{2}{3}, \tfrac{7}{9}] \cup [\tfrac{8}{9}, 1].$$

Continuing in this way, we find that E_k is a union of 2^k closed intervals of length 3^{-k}. The Cantor set E is now defined as

$$E = \bigcap_{k=0}^{\infty} E_k.$$

By Corollary 16 it is clear that E is a closed subset of $\mathbb{R}$. However, E does not contain any interval of positive length. Let $x, y \in E$ with $x < y$ we will show that $[x, y] \not\subseteq E$. Towards this, find $k \in \mathbb{Z}^+$ such that

$$3^{-(k+1)} \leq y - x < 3^{-k}.$$

Now, E_k consists of intervals separated by a distance of at least 3^{-k}. Since $x, y \in E \subseteq E_k$, it must be the case that x and y lie in the same constituent interval J of E_k. If x lies in the lower third and y in the upper third of J, then already $[x, y] \not\subseteq E_{k+1}$. So, since $3^{-(k+1)} \leq y - x$, x and y must be the extremities of either the lower third of J or the upper third of J. Now it is clear that $[x, y] \not\subseteq E_{k+2}$.

The sculptor Rodin once said that to make a sculpture one starts with a block of marble and removes everything that is unimportant. This is the approach that we have just taken in building the Cantor set. There is a second way of constructing the Cantor set which works by building the set from the inside out. In fact, we have

$$E = \{\sum_{k=1}^{\infty} \omega_k 3^{-k}; \omega_k \in \{0, 2\}, k = 1, 2, \ldots\}. \tag{1.12}$$

Another way of saying this is that E consists of all numbers in [0,1] with a ternary (i.e. base 3) expansion in which only the "tergits" 0 and 2 occur. This is why Cantor's set is sometimes called the ternary set. The proof of (1.12) is not too difficult, but we do not give it here. □

1.10 Continuity

The primary purpose of the preceding sections is to define the concept of ***continuity*** of mappings. This concept is the mainspring of mathematical analysis.

DEFINITION *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Let $x \in X$. Then f is* ***continuous at*** *x iff for all $\epsilon > 0$, there exists $\delta > 0$ such that*

$$z \in U(x, \delta) \qquad \Rightarrow \qquad f(z) \in U(f(x), \epsilon). \tag{1.13}$$

The $\forall \ldots \exists \ldots$ combination suggests the role of the "devil's advocate" type of argument. Let us illustrate this with an example.

EXAMPLE The mapping $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by $f(x) = x^2$ is continuous at $x = 1$. To prove this, we suppose that the devil's advocate provides us with a number $\epsilon > 0$ chosen cunningly small. We have to "reply" with a number $\delta > 0$ (depending on ϵ) such that (1.13) holds. In the present context, we choose

$$\delta = \min(\tfrac{1}{4}\epsilon, 1)$$

so that for $|x - 1| < \delta$ we have

$$|x^2 - 1| \le |x - 1||x + 1| < (\tfrac{1}{4}\epsilon)(3) < \epsilon$$

since $|x - 1| < \delta$ and $|x + 1| = |(x - 1) + 2| \le |x - 1| + 2 < 3$. □

EXAMPLE Continuity at a point — a single point that is, does not have much strength. Consider the function $f : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$f(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q}, \\ x & \text{if } x \in \mathbb{Q}. \end{cases}$$

This function is continuous at 0 but at no other point of $\mathbb{R}$. □

EXAMPLE An interesting contrast is provided by the function $g : \mathbb{R} \longrightarrow \mathbb{R}$ given by

$$g(x) = \begin{cases} 0 & \text{if } x \in \mathbb{R} \setminus \mathbb{Q} \text{ or if } x = 0, \\ \frac{1}{q} & \text{if } x = \frac{p}{q} \text{ where } p \in \mathbb{Z} \setminus \{0\},\, q \in \mathbb{N} \text{ are coprime.} \end{cases}$$

The function g is continuous at x iff x is zero or irrational. To see this, we first observe that if $x \in \mathbb{Q} \setminus \{0\}$, then $g(x) \neq 0$ but there are irrational numbers z as close as we like to x which satisfy $g(z) = 0$. Thus g is not continuous at the points of $\mathbb{Q} \setminus \{0\}$. On the other hand, if $x \in \mathbb{R} \setminus \mathbb{Q}$ or $x = 0$, we can establish continuity of g at x by an epsilon delta argument. We agree that whatever $\epsilon > 0$ we will always choose $\delta < 1$. Then the number of points z in the interval $]x - \delta, x + \delta[$ where $|g(z)| \geq \epsilon$ is finite because such a z is necessarily a rational number that can be expressed in the form $\frac{p}{q}$ where $1 \leq q < \epsilon^{-1}$. With only finitely many points to avoid, it is now easy to find $\delta > 0$ such that

$$|z - x| < \delta \qquad \Longrightarrow \qquad |g(z) - g(x)| = |g(z)| < \epsilon.$$

□

There are various other ways of formulating continuity at a point.

THEOREM 17 *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Let $x \in X$. Then the following statements are equivalent.*

- *f is continuous at x.*
- *For every neighbourhood V of $f(x)$ in Y, $f^{-1}(V)$ is a neighbourhood of x in X.*
- *For every sequence (x_n) in X converging to x, the sequence $(f(x_n))$ converges to $f(x)$ in Y.*

Proof. We show that the first statement implies the second. Let f be continuous at x and suppose that V is a neighbourhood of $f(x)$ in Y. Then there exists $\epsilon > 0$ such that $U(f(x), \epsilon) \subseteq V$ in Y. By definition of continuity at a point, there exists $\delta > 0$ such that

$$\begin{aligned} z \in U(x, \delta) \qquad &\Rightarrow \qquad f(z) \in U(f(x), \epsilon) \\ &\Rightarrow \qquad f(z) \in V \\ &\Rightarrow \qquad z \in f^{-1}(V). \end{aligned}$$

Hence $f^{-1}(V)$ is a neighbourhood of x in X.

Next, we assume the second statement and establish the third. Let (x_n) be a sequence in X converging to x. Let $\epsilon > 0$. Then $U(f(x), \epsilon)$ is a neighbourhood of $f(x)$ in Y. By hypothesis, $f^{-1}(U(f(x), \epsilon))$ is a neighbourhood of x in X. By the first part of Proposition 13 there exists $N \in \mathbb{N}$ such that

$$n > N \qquad \Rightarrow \qquad x_n \in f^{-1}(U(f(x), \epsilon)).$$

But this is equivalent to

$$n > N \qquad \Rightarrow \qquad f(x_n) \in U(f(x), \epsilon).$$

Thus $(f(x_n))$ converges to $f(x)$ in Y.

Finally we show that the third statement implies the first. We argue by contradiction. Suppose that f is not continuous at x. Then there exists $\epsilon > 0$ such that for all $\delta > 0$, there exists $z \in X$ with $d(x, z) < \delta$, but $d(f(x), f(z)) \geq \epsilon$. We take choice $\delta = \frac{1}{n}$ for $n = 1, 2, \ldots$ in sequence. We find that there exist x_n in X with $d(x, x_n) < \frac{1}{n}$, but $d(f(x), f(x_n)) \geq \epsilon$. But now, the sequence (x_n) converges to x in X while the sequence $(f(x_n))$ does not converge to $f(x)$ in Y. ■

We next build the global version of continuity from the concept of continuity at a point.

DEFINITION *Let X and Y be metric spaces and let $f : X \longrightarrow Y$. Then the mapping f is* ***continuous*** *iff f is continuous at every point x of X.*

There are also many possible reformulations of global continuity.

THEOREM 18 *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Then the following statements are equivalent to the continuity of f.*

- *For every open set U in Y, $f^{-1}(U)$ is open in X.*

- *For every closed set A in Y, $f^{-1}(A)$ is closed in X.*
- *For every convergent sequence (x_n) in X with limit x, the sequence $(f(x_n))$ converges to $f(x)$ in Y.*

Proof. Let f be continuous. We check that the first statement holds. Let $x \in f^{-1}(U)$. Then $f(x) \in U$. Since U is open in Y, U is a neighbourhood of $f(x)$. Hence, by Theorem 17 $f^{-1}(U)$ is a neighbourhood of x. We have just shown that $f^{-1}(U)$ is a neighbourhood of each of its points. Hence $f^{-1}(U)$ is open in X. For the converse, we assume that the first statement holds. Let x be an arbitrary point of X. We must show that f is continuous at x. Again we plan to use Theorem 17. Let V be a neighbourhood of $f(x)$ in Y. Then, there exists $t > 0$ such that $U(f(x), t) \subseteq V$. It is shown on page 17 that $U(f(x), t)$ is an open subset of Y. Hence using the hypothesis, $f^{-1}(U(f(x), t))$ is open in X. Since $x \in f^{-1}(U(f(x), t))$, this set is a neighbourhood of x, and it follows that so is the larger subset $f^{-1}(V)$.

The second statement is clearly equivalent to the first. For instance if A is closed in Y, then $Y \setminus A$ is an open subset. Then

$$X \setminus f^{-1}(A) = f^{-1}(Y \setminus A)$$

is open in X and it follows that $f^{-1}(A)$ is closed in X. The converse entirely similar.

The third statement is equivalent directly from the definition. ■

One very useful condition that implies continuity is the Lipschitz condition.

DEFINITION *Let X and Y be metric spaces. Let $f : X \longrightarrow Y$. Then f is a* ***Lipschitz map*** *iff there is a constant C with $0 < C < \infty$ such that*

$$d_Y(f(x_1), f(x_2)) \leq C d_X(x_1, x_2) \qquad \forall x_1, x_2 \in X.$$

In the special case that $C = 1$ we say that f is a ***nonexpansive mapping*** *. In the even more restricted case that*

$$d_Y(f(x_1), f(x_2)) = d_X(x_1, x_2) \qquad \forall x_1, x_2 \in X,$$

we say that f is an ***isometry*** *.*

PROPOSITION 19 *Every Lipschitz map is continuous.*

Proof. We work directly. Let $\epsilon > 0$. The set $\delta = C^{-1}\epsilon$. Then $d_X(z, x) < \delta$ implies that

$$d_Y(f(z), f(x)) \le C d_X(z, x) \le C\delta = \epsilon.$$

as required. ■

1.11 Compositions of Functions

DEFINITION *Let X, Y and Z be sets. Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be mappings. Then we can make a new mapping $h : X \longrightarrow Z$ by $h(x) = g(f(x))$. In other words, to map by h we first map by f from X to Y and then by g from Y to Z. The mapping h is called the* ***composition*** *or* ***composed mapping*** *of f and g. It is usually denoted by $h = g \circ f$.*

Composition occurs in very many situations in mathematics. It is the primary tool for building new mappings out of old.

THEOREM 20 *Let X, Y and Z be metric spaces. Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be continuous mappings. Then the composition $g \circ f$ is a continuous mapping from X to Z.*

THEOREM 21 *Let X, Y and Z be metric spaces. Let $f : X \longrightarrow Y$ and $g : Y \longrightarrow Z$ be mappings. Suppose that $x \in X$, that f is continuous at x and that g is continuous at $f(x)$. Then the composition $g \circ f$ is a continuous at x.*

Proof of Theorems 20 and 21. There are many possible ways of proving these results using the tools from Theorem 18 and 17. It is even relatively easy to work directly from the definition.

Let us use sequences. In the local case, we take x as a fixed point of X whereas in the global case we take x to be a generic point of X.

Let (x_n) be a sequence in X convergent to x. Then since f is continuous at x, $(f(x_n))$ converges to $f(x)$. But, then using the fact that g is continuous at $f(x)$, we find that $(g(f(x_n)))$ converges to $g(f(x))$. This says that $(g \circ f(x_n))$ converges to $g \circ f(x)$. Since this holds for every sequence (x_n) convergent to x, it follows that $g \circ f$ is continuous (respectively continuous at x). ■

1.12 Interior and Closure

We return to discuss subsets and sequences in metric spaces in greater detail. Let X be a metric space and let A be an arbitrary subset of X. Then $\emptyset$ is an open subset of X contained in A, so we can define the ***interior*** $\mathrm{int}(A)$ of A by

$$\mathrm{int}(A) = \bigcup_{U \text{ open } \subseteq A} U. \tag{1.14}$$

By Theorem 9 (page 18), we see that $\mathrm{int}(A)$ is itself an open subset of X contained in A. Thus $\mathrm{int}(A)$ is the unique open subset of X contained in A which in turn contains all open subsets of X contained in A. There is a simple characterization of $\mathrm{int}(A)$ in terms of interior points (page 16).

PROPOSITION 22 *Let X be a metric space and let $A \subseteq X$. Then*

$$\mathrm{int}(A) = \{x; x \textit{ is an interior point of } A\}.$$

Proof. Let $x \in \mathrm{int}(A)$. Then since $\mathrm{int}(A)$ is open, it is a neighbourhood of x. But then the (possibly) larger set A is also a neighbourhood of x. This just says that x is an interior point of A.

For the converse, let x be an interior point of A. Then by definition, there exists $t > 0$ such that $U(x,t) \subseteq A$. But it is shown on page 17, that $U(x,t)$ is open. Thus $U = U(x,t)$ figures in the union in (1.14), and since $x \in U(x,t)$ it follows that $x \in \mathrm{int}(A)$. ■

EXAMPLE The interior of the closed interval $[a,b]$ of $\mathbb{R}$ is just $]a,b[$. □

EXAMPLE The Cantor set E has empty interior in $\mathbb{R}$. Suppose not. Let x be an interior point of E. Then there exist $\epsilon > 0$ such that $U(x,\epsilon) \subseteq E$. Choose now n so large that $3^{-n} < \epsilon$. Then we also have $U(x,\epsilon) \subseteq E_n$. For the notation see page 25. This says that E_n contains an open interval of length $2(3^{-n})$ which is clearly not the case. □

By passing to the complement and using Theorem 15 (page 23) we see that there is a unique closed subset of X containing A which is contained in every closed subset of X which contains A. The formal definition is

$$\mathrm{cl}(A) = \bigcap_{E \text{ closed } \supseteq A} E. \tag{1.15}$$

The set $\mathrm{cl}(A)$ is called the ***closure*** of A. We would like to have a simple characterization of the closure.

PROPOSITION 23 *Let X be a metric space and let $A \subseteq X$. Let $x \in X$. Then $x \in \mathrm{cl}(A)$ is equivalent to the existence of a sequence of points (x_n) in A converging to x.*

Proof. Let $x \in \mathrm{cl}(A)$. Then x is not in $\mathrm{int}(X \setminus A)$. Then by Proposition 22, x is not an interior point of $X \setminus A$. Then, for each $n \in \mathbb{N}$, there must be a point $x_n \in A \cap U(x, \frac{1}{n})$. But now, $x_n \in A$ and (x_n) converges to x.

For the converse, let (x_n) be a sequence of points of A converging to x. Then $x_n \in \mathrm{cl}(A)$ and since $\mathrm{cl}(A)$ is closed, it follows from the definition of a closed set that $x \in \mathrm{cl}(A)$. ■

While Proposition 23 is perfectly satisfactory for many purposes, there is a subtle variant that is sometimes necessary.

DEFINITION *Let X be a metric space and let $A \subseteq X$. Let $x \in X$. Then x is an **accumulation point** or a **limit point** of A iff $x \in \mathrm{cl}(A \setminus \{x\})$.*

PROPOSITION 24 *Let X be a metric space and let $A \subseteq X$. Let $x \in X$. Then the following statements are equivalent.*

- *$x \in \mathrm{cl}(A)$.*
- *$x \in A$ or x is an accumulation point of A.*

Proof. That the second statement implies the first follows easily from Proposition 23. We establish the converse. Let $x \in \mathrm{cl}(A)$. We may suppose that $x \notin A$, for else we are done. Now apply the argument of Proposition 23 again. For each $n \in \mathbb{N}$, there is a point $x_n \in A \cap U(x, \frac{1}{n})$. Since $x \notin A$, we have $A = A \setminus \{x\}$. Thus we have found $x_n \in A \setminus \{x\}$ with (x_n) converging to x. ■

DEFINITION *Let X be a metric space and let $A \subseteq X$. Let $x \in A$. Then x is an **isolated point** of A iff there exists $t > 0$ such that $A \cap U(x, t) = \{x\}$.*

We leave the reader to check that a point of A is an isolated point of A if and only if it is not an accumulation point of A.

A very important concept related to closure is the concept of density.

DEFINITION *Let X be a metric space and let $A \subseteq X$. Then A is said to be **dense** in X if $\mathrm{cl}(A) = X$.*

If A is dense in X, then by definition, for every $x \in X$ there exists a sequence (x_n) in A converging to x.

PROPOSITION 25 *Let f and g be continuous mappings from X to Y. Suppose that A is a dense subset of X and that $f(x) = g(x)$ for all $x \in A$. Then $f(x) = g(x)$ for all $x \in X$.*

Proof. Let $x \in X$ and let (x_n) be a sequence in A converging to x. Then $f(x_n) = g(x_n)$ for all $n \in \mathbb{N}$. So the sequences $(f(x_n))$ and $(g(x_n))$ which converge to $f(x)$ and $g(x)$ respectively, are in fact identical. By the uniqueness of the limit, Proposition 12 (page 21), it follows that $f(x) = g(x)$. This holds for all $x \in X$ so that $f = g$. ∎

1.13 Limits in Metric Spaces

DEFINITION *Let X be a metric space and let $t > 0$. Then for $x \in X$ the **deleted open ball** $U'(x,t)$ is defined by*

$$U'(x,t) = \{z; z \in X, 0 < d(x,z) < t\} = U(x,t) \setminus \{x\}.$$

Let A be a subset of X then it is routine to check that x is an accumulation point of A if and only if for all $t > 0$, $U'(x,t) \cap A \neq \emptyset$. Deleted open balls are also used to define the concept of a ***limit*** .

DEFINITION *Let X and Y be metric spaces. Let x be an accumulation point of X. Let $f : X \setminus \{x\} \longrightarrow Y$. Then $f(z)$ has **limit** y **as** z **tends to** x **in** X, in symbols*

$$\lim_{z \to x} f(z) = y \tag{1.16}$$

if and only if for all $\epsilon > 0$ there exists $\delta > 0$ such that

$$z \in U'(x,\delta) \Longrightarrow f(z) \in U(y,\epsilon).$$

In the same way one also defines $f(z)$ has ***a limit as*** z ***tends to*** x ***in*** X*, which simply means that* (1.16) *holds for some $y \in Y$.*

Note that in the above definition, the quantity $f(x)$ is undefined. The purpose of taking the limit is to "attach a value" to $f(x)$. The following Lemma connects this idea with the concept of continuity at a point. We leave the proof to the reader.

LEMMA 26 *Let X and Y be metric spaces. Let x be an accumulation point of X. Let $f : X \setminus \{x\} \longrightarrow Y$. Suppose that* (1.16) *holds for some $y \in Y$. Now define $\tilde{f} : X \longrightarrow Y$ by*

$$\tilde{f}(z) = \begin{cases} f(z) & \text{if } z \in X \setminus \{x\}, \\ y & \text{if } z = x. \end{cases}$$

Then $\tilde{f}$ is continuous at x.

1.14 Uniform Continuity

For many purposes, continuity of mappings is not enough. The following strong form of continuity is often needed.

DEFINITION *Let X and Y be metric spaces and let $f : X \longrightarrow Y$. Then we say that f is* ***uniformly continuous*** *iff for all $\epsilon > 0$ there exists $\delta > 0$ such that*

$$x_1, x_2 \in X, d_X(x_1, x_2) < \delta \qquad \Rightarrow \qquad d_Y(f(x_1), f(x_2)) < \epsilon. \tag{1.17}$$

In the definition of continuity, the number δ is allowed to depend on the point x_1 as well as ϵ.

EXAMPLE The function $f(x) = x^2$ is continuous, but not uniformly continuous as a mapping $f : \mathbb{R} \longrightarrow \mathbb{R}$. Certainly the identity mapping $x \longrightarrow x$ is continuous because it is an isometry. So f, which is the pointwise product of the identity mapping with itself is also continuous. We now show that f is not uniformly continuous. Let us take $\epsilon = 1$. Then, we must show that for all $\delta > 0$ there exist points x_1 and x_2 with $|x_1 - x_2| < \delta$, but $|x_1^2 - x_2^2| \geq 1$. Let us take $x_2 = x - \frac{1}{4}\delta$ and $x_1 = x + \frac{1}{4}\delta$. Then

$$x_1^2 - x_2^2 = (x_1 - x_2)(x_1 + x_2) = x\delta.$$

It remains to choose $x = \delta^{-1}$ to complete the argument. □

EXAMPLE Any function satisfying a Lipschitz condition (page 29) is uniformly continuous. Let X and Y be metric spaces. Let $f : X \longrightarrow Y$ with constant C. Then

$$d_Y(f(x_1), f(x_2)) \le C d_X(x_1, x_2) \qquad \forall x_1, x_2 \in X.$$

Given $\epsilon > 0$ it suffices to choose $\delta = C^{-1}\epsilon > 0$ in order for $d_X(x_1, x_2) < \delta$ to imply $d_Y(f(x_1), f(x_2)) < \epsilon$. □

It should be noted that one cannot determine (in general) if a mapping is uniformly continuous from a knowledge only of the open subsets of X and Y. Thus, uniform continuity is not a topological property. It depends upon other aspects of the metrics involved.

In order to clarify the concept of uniform continuity and for other purposes, one introduces the ***modulus of continuity*** ω_f of a function f. Suppose that $f : X \longrightarrow Y$. Then $\omega_f(t)$ is defined for $t \ge 0$ by

$$\omega_f(t) = \sup\{d_Y(f(x_1), f(x_2)); x_1, x_2 \in X, d_X(x_1, x_2) \le t\}. \tag{1.18}$$

It is easy to see that the uniform continuity of f is equivalent to

$$\forall \epsilon > 0,\ \exists \delta > 0 \text{ such that } 0 < t < \delta \quad \Rightarrow \quad \omega_f(t) < \epsilon.$$

We observe that $\omega_f(0) = 0$ and regard $\omega_f : \mathbb{R}^+ \longrightarrow \mathbb{R}^+$. Then the uniform continuity of f is also equivalent to the continuity of ω_f at 0.

1.15 Subsequences and Sequential Compactness

Subsequences are important because they are used to approach the topics of sequential compactness and later completeness. These ideas are used to establish the existence of a limit when the actual limiting value is not known explicitly. To show that $x_n \longrightarrow x$ in a metric space, we show that $d(x_n, x) \longrightarrow 0$ in $\mathbb{R}$. This supposes that the limit x is known in advance. We need ways of showing that sequences are convergent when the limit is not known in advance.

DEFINITION *A sequence (n_k) of natural numbers is called a* ***natural subsequence*** *if $n_k < n_{k+1}$ for all $k \in \mathbb{N}$.*

Since $n_1 \ge 1$, a straightforward induction argument yields that $n_k \ge k$ for all $k \in \mathbb{N}$.

DEFINITION *Let (x_n) be a sequence of elements of a set X. A* ***subsequence*** *of (x_n) is a sequence (y_k) of elements of X given by*

$$y_k = x_{n_k}$$

where (n_k) is a natural subsequence.

The key result about subsequences is the following.

LEMMA 27 *Let (x_n) be a sequence in a metric space X converging to an element $x \in X$. Then any subsequence (x_{n_k}) also converges to x.*

Proof. Since (x_n) converges to x, given $\epsilon > 0$ there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \epsilon$ whenever $n \geq N$. But now, $k \geq N$ implies that $n_k \geq k \geq N$ and therefore that $d(x_{n_k}, x) < \epsilon$. ■

DEFINITION *Let X be a metric space and A a subset of X. Then A is* ***sequentially compact*** *iff every sequence (a_n) in A possesses a subsequence which converges to some element of A.*

LEMMA 28 *A sequentially compact subset is both closed and bounded.*

Proof. Let A be a sequentially compact subset of a metric space X. Suppose that A is not closed. Then there is a point $x \in X \setminus A$ and a sequence (a_n) with $a_n \longrightarrow x$. But then every subsequence of (a_n) will also converge to x and hence not to an element of A since limits are unique. To show that A is bounded, again, suppose not. If A is empty then we are done. If not, let a be a reference point of A. Then there is an element a_n of A such that $d(a_n, a) > n$ for otherwise, every element of A would be within distance n of a. But for any subsequence a_{n_k} we will have $d(a_{n_k}, a) \geq n_k \geq k$ and the sequence a_{n_k} cannot converge because it is also unbounded. ■

In the real line, it is an easy consequence of the Bolzano–Weierstrass Theorem that every closed bounded subset is sequentially compact. This statement is not true in general metric spaces.

One of the key results about sequentially compact spaces is the following.

THEOREM 29 *Let A be a sequentially compact subset of a metric space X. Let $f : X \longrightarrow \mathbb{R}$ be a continuous mapping. Then $f(A)$ is bounded above and the supremum* $\sup f(A)$ *is attained. Similarly $f(A)$ is bounded below and* $\inf f(A)$ *is attained.*

Proof. First we show that $f(A)$ is bounded above. If not, then there exists $a_n \in A$ such that $f(a_n) > n$. Since A is sequentially compact, there is a subsequence (a_{n_k}) of (a_n) and an element $a \in A$ such that $a_{n_k} \longrightarrow a$ as $k \longrightarrow \infty$. But then $f(a_{n_k}) \longrightarrow f(a)$ as $k \longrightarrow \infty$ and $(f(a_{n_k}))_{k=1}^{\infty}$ is a bounded sequence. Clearly this contradicts $f(a_{n_k}) > n_k \geq k$.

We show that the sup is attained. Let (ϵ_n) be a sequence of positive numbers converging to 0. Then, there exists $a_n \in A$ such that $f(a_n) > \sup f(A) - \epsilon_n$. Since A is sequentially compact, there is a subsequence (a_{n_k}) of (a_n) and an element $a \in A$ such that $a_{n_k} \longrightarrow a$ as $k \longrightarrow \infty$. Since $f(a_{n_k}) > \sup f(A) - \epsilon_{n_k}$, f is continuous and $\epsilon_{n_k} \longrightarrow 0$, we find that $f(a) \geq \sup f(A)$. But obviously, since $a \in A$ we also have $f(a) \leq \sup f(A)$. Therefore $f(a) = \sup f(A)$. ■

Another important result concerns uniform continuity.

THEOREM 30 *Let X be a sequentially compact metric space and let Y be any metric space. Let $f : X \longrightarrow Y$ be a continuous function. Then f is uniformly continuous.*

Proof. Suppose not. Then there exists $\epsilon > 0$ such that the uniform continuity condition fails. This means that for any $\delta > 0$ there will exist $a, b \in X$ such that $d_X(a, b) < \delta$ but $d_Y(f(a), f(b)) \geq \epsilon$. So, choose a sequence of positive numbers (δ_n) converging to zero and applying this with $\delta = \delta_n$ we find sequences (a_n) and (b_n) in X such that $d_X(a_n, b_n) < \delta_n$ and $d_Y(f(a_n), f(b_n)) \geq \epsilon$. We now use the sequential compactness of X to find a point x of X where things impact. So, there is a subsequence (a_{n_k}) of (a_n) and a point $x \in X$ such that $a_{n_k} \longrightarrow x$ as $k \longrightarrow \infty$. Since

$$d_X(x, b_{n_k}) \leq d_X(x, a_{n_k}) + d_X(a_{n_k}, b_{n_k}) < d_X(x, a_{n_k}) + \delta_{n_k}$$

we find that also $b_{n_k} \longrightarrow x$ as $k \longrightarrow \infty$.

Now apply the definition of continuity at x with the "epsilon" replaced by $\frac{\epsilon}{3}$. We find that there exists $\delta > 0$ such that

$$d_X(x, y) < \delta \quad \Longrightarrow \quad d_Y(f(x), f(y)) < \frac{\epsilon}{3}.$$

But, for k large enough, we will have both $d_X(x, a_{n_k}) < \delta$ and $d_X(x, b_{n_k}) < \delta$, so that both $d_Y(f(x), f(a_{n_k})) < \frac{\epsilon}{3}$ and $d_Y(f(x), f(b_{n_k})) < \frac{\epsilon}{3}$. Now, by the triangle inequality we have $d_Y(f(a_{n_k}), f(b_{n_k})) < \frac{2\epsilon}{3}$ which contradicts the statement $d_Y(f(a_n), f(b_n)) \geq \epsilon$. This contradiction shows that f must be uniformly continuous. ■

1.16 Sequential Compactness in Normed Vector Spaces

We now turn our attention to normed vector spaces. We start by considering $\mathbb{R}^d$ with the Euclidean norm.

LEMMA 31 *A sequence converging coordinatewise in Euclidean space also converges in norm.*

Proof. Let $(x_n)_{n=1}^{\infty}$ be a sequence in $\mathbb{R}^d$ such that $x_{n,k} \longrightarrow \xi_k$ as $n \longrightarrow \infty$ for each $k = 1, 2, \ldots, d$. Then we will show that $x_n \longrightarrow x$ as $n \longrightarrow \infty$ where $x = (\xi_1, \xi_2, \ldots, \xi_d)$.

Let $\epsilon > 0$. Then, for each $k = 1, 2, \ldots, d$, there exists $N_k \in \mathbb{N}$ such that

$$n > N_k \quad \Longrightarrow \quad |x_{n,k} - \xi_k| < \frac{\epsilon}{d}$$

It follows that

$$\begin{aligned} n > \max(N_1, \ldots, N_k) \quad &\Longrightarrow \quad \sum_{k=1}^{d} |x_{n,k} - \xi_k|^2 < d\left(\frac{\epsilon}{d}\right)^2 \leq \epsilon^2 \\ &\Longrightarrow \quad \|x_n - x\| < \epsilon \end{aligned}$$

■

THEOREM 32 *Every closed bounded subset of Euclidean space is sequentially compact.*

Proof. Let (v_n) be a bounded sequence in $\mathbb{R}^d$ for the Euclidean norm. Let $v_{n,k}$ be the coordinates of v_n for $k = 1, 2, \ldots, d$. Then, for each k we have a coordinate sequence $(v_{n,k})_{n=1}^{\infty}$ which is bounded sequence in $\mathbb{R}$. From the first coordinate sequence, $(v_{n,1})_{n=1}^{\infty}$ we can extract a convergent subsequence $(v(n_\ell, 1))$ converging say to u_1. Then, from the corresponding subsequence $(v(n_\ell, 2))$ of the second coordinate sequence extract a further subsequence $(v(n_{\ell_m}, 2))_{m=1}^{\infty}$ converging say to u_2. But now, $(v(n_{\ell_m}, 1))_{m=1}^{\infty}$ is still converging to u_1 because it is a subsequence of $(v(n_\ell, 1))$. So, if $d = 2$, we are done because we have found a subsequence which is converging coordinatewise to (u_1, u_2). If $d > 2$ then we have to repeat the argument and take further subsequences in the remaining coordinates. Details are left to the reader. ■

Now we may consider another norm. So on $\mathbb{R}^d$ we will denote by $\| \ \|$ the Euclidean norm and $|\!|\!| \ |\!|\!|$ some other norm.

LEMMA 33 *There is a constant C such that*

$$|\!|\!|x|\!|\!| \le C\|x\| \qquad \forall x \in \mathbb{R}^d.$$

Proof. Let e_k denote the standard coordinate vectors in $\mathbb{R}^d$. We can write

$$x = x_1 e_1 + x_2 e_2 + \cdots + x_d e_d$$

and so

$$\begin{aligned} |\!|\!|x|\!|\!| &\le |x_1| |\!|\!|e_1|\!|\!| + |x_2| |\!|\!|e_2|\!|\!| + \cdots + |x_d| |\!|\!|e_d|\!|\!| \\ &\le \left\{x_1^2 + x_2^2 + \cdots + x_d^2\right\}^{\frac{1}{2}} \left\{|\!|\!|e_1|\!|\!|^2 + |\!|\!|e_2|\!|\!|^2 + \cdots + |\!|\!|e_d|\!|\!|^2\right\}^{\frac{1}{2}} \\ &= \|x\| \left\{|\!|\!|e_1|\!|\!|^2 + |\!|\!|e_2|\!|\!|^2 + \cdots + |\!|\!|e_d|\!|\!|^2\right\}^{\frac{1}{2}} \end{aligned}$$

using the Cauchy–Schwarz Inequality. Much more remarkable is the following result.

THEOREM 34 *There is a constant C' such that*

$$\|x\| \le C' |\!|\!|x|\!|\!| \qquad \forall x \in \mathbb{R}^d.$$

Proof. Consider the mapping $f : \mathbb{R}^d \longrightarrow \mathbb{R}$ given by $f(x) = |||x|||$. This mapping is continuous for the Euclidean norm on $\mathbb{R}^d$ by Lemma 33 since

$$|f(x) - f(y)| = ||||x||| - |||y|||| \leq |||x - y||| \leq C\|x - y\|$$

shows that f is Lipschitz. Now consider the (Euclidean) unit sphere S^{d-1} in $\mathbb{R}^d$. This is a closed bounded subset of $\mathbb{R}^d$ for the Euclidean metric. So, by Theorem 29 the infimum

$$\alpha = \inf_{x \in S^{d-1}} |||x|||$$

is attained. Now, since norms are nonnegative, we have $\alpha \geq 0$. If $\alpha = 0$ it follows that there exists $x \in S^{d-1}$ such that $|||x||| = 0$. This is impossible because $|||x||| = 0$ implies that $x = \mathbf{0}$ and $\mathbf{0} \notin S^{d-1}$. Therefore we must have $\alpha > 0$. Now let $x \in \mathbb{R}^d$ with $x \neq \mathbf{0}$. Then $\|x\|^{-1}x \in S^{d-1}$ and we must then have

$$\alpha \leq |||\|x\|^{-1}x||| = \|x\|^{-1}|||x|||,$$

or equivalently

$$\|x\| \leq \alpha^{-1}|||x|||. \tag{1.19}$$

But (1.19) is also true if $x = \mathbf{0}$ and the result is proved. ■

The consequences of Lemma 33 and Theorem 34 are:

- On a finite dimensional vector space over $\mathbb{R}$ or $\mathbb{C}$, all norms are equivalent.
- On a finite dimensional normed vector space over $\mathbb{R}$ or $\mathbb{C}$, all linear functions are continuous.

It goes without saying that both these statements are false in the infinite dimensional case.

1.17 Cauchy Sequences and Completeness

We will assume that the reader is familiar with the completeness of $\mathbb{R}$. Usually $\mathbb{R}$ is defined as the unique order-complete totally ordered field. The order completeness postulate is that every subset B of $\mathbb{R}$ which is bounded above possesses a least upper bound (or supremum). From this the metric completeness of $\mathbb{R}$ is deduced. Metric completeness is formulated in terms of the convergence of Cauchy sequences.

DEFINITION *Let X be a metric space. Let (x_n) be a sequence in X. Then (x_n) is a* ***Cauchy sequence*** *iff for every number $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that*

$$p, q > N \qquad \Rightarrow \qquad d(x_p, x_q) < \epsilon.$$

LEMMA 35 *Every convergent sequence is Cauchy.*

Proof. Let X be a metric space. Let (x_n) be a sequence in X converging to $x \in X$. Then given $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_n, x) < \frac{1}{2}\epsilon$ for $n > N$. Thus for $p, q > N$ the triangle inequality gives

$$d(x_p, x_q) \leq d(x_p, x) + d(x, x_q) < \tfrac{1}{2}\epsilon + \tfrac{1}{2}\epsilon = \epsilon.$$

Hence (x_n) is Cauchy. ■

The Cauchy condition on a sequence says that the diameters of the successive tails of the sequence converge to zero. One feels that this is almost equivalent to convergence except that no limit is explicitly mentioned. Sometimes, Cauchy sequences fail to converge because the "would be limit" is not in the space. It is the existence of such "gaps" in the space that prevent it from being complete.

Note that it is also true that every Cauchy sequence is bounded.

DEFINITION *Let X be a metric space. Then X is* ***complete*** *iff every Cauchy sequence in X converges in X.*

EXAMPLE The real line $\mathbb{R}$ is complete. □

EXAMPLE The set $\mathbb{Q}$ of rational numbers is not complete. Consider the sequence defined inductively by

$$x_1 = 2 \quad \text{and} \quad x_{n+1} = \frac{1}{2}\left(x_n + \frac{2}{x_n}\right), \quad n = 1, 2, \ldots. \tag{1.20}$$

Then one can show that (x_n) converges to $\sqrt{2}$ in $\mathbb{R}$. It follows that (x_n) is a Cauchy sequence in $\mathbb{Q}$ which does not converge in $\mathbb{Q}$. Hence $\mathbb{Q}$ is not complete.

To fill in the details, observe first that (1.20) can also be written in both of the alternative forms

$$2x_n(x_{n+1} - \sqrt{2}) = (x_n - \sqrt{2})^2,$$
$$x_{n+1} - x_n = -\left(\frac{x_n^2 - 2}{2x_n}\right).$$

We now observe the following in succession.

- $x_n > 0$ for all $n \in \mathbb{N}$.
- $x_n > \sqrt{2}$ for all $n \in \mathbb{N}$.
- x_n is decreasing with n.
- $x_n \leq 2$ for all $n \in \mathbb{N}$.
- $|x_{n+1} - \sqrt{2}| \leq \dfrac{|x_n - \sqrt{2}|^2}{2\sqrt{2}}$ for all $n \in \mathbb{N}$.
- $|x_{n+1} - \sqrt{2}| \leq \dfrac{2 - \sqrt{2}}{2\sqrt{2}}|x_n - \sqrt{2}|$ for all $n \in \mathbb{N}$.

The convergence of (x_n) to $\sqrt{2}$ follows easily. □

Completeness is very important because in the general metric space setting, it is the only tool that we have at our disposal for proving the convergence of a sequence when we do not know what the limit is.

LEMMA 36 *$\mathbb{R}^d$ is complete with the Euclidean metric.*

Proof. Let $(x_n)_{n=1}^\infty$ be a Cauchy sequence of vectors in $\mathbb{R}^d$. Since $|x_{n,k} - x_{m,k}| \leq \|x_n - x_m\|$ we see that each of the coordinate sequences $(x_{n,k})_{n=1}^\infty$ is Cauchy $(1 \leq k \leq d)$. Hence each of the coordinate sequences converges in $\mathbb{R}$ say to ξ_k. Here we have used the fact that $\mathbb{R}$ is complete. But, now by Lemma 31, we see that $(x_n)_{n=1}^\infty$ converges to $\xi = (\xi_1, \ldots, \xi_d)$.

COROLLARY 37 *Every finite dimensional normed vector space over $\mathbb{R}$ or $\mathbb{C}$ is complete.*

Proof. Combine Lemma 36 with Lemma 33 and Theorem 34. ■

Once again the Corollary is not true in the infinite dimensional setting.

2

Numerical Series

In this chapter, we want to make sense of an infinite sum. Typically we are given real numbers a_n and we wish to attach a meaning to

$$\sum_{n=1}^{\infty} a_n$$

The way that we do this is to define the ***partial sum*** $s_N = \sum_{n=1}^{N} a_n$. This gives us a sequence $(s_N)_{N=1}^{\infty}$.

DEFINITION *We say that $\sum_{n=1}^{\infty} a_n$ exists and equals s, or simply $s = \sum_{n=1}^{\infty} a_n$ if and only if $s_N \longrightarrow s$ as $N \longrightarrow \infty$. We say that the $\sum_{n=1}^{\infty} a_n$* ***converges*** *if (s_N) converges to some limit.*

Since the limit of a sequence is uniquely determined when it exists, the sum of a series is likewise unique when the series converges.

There are some cases when the sum of a series can be found explicitly because we can find a formula for all the partial sums. Perhaps the most basic example is the ***geometric series*** .

EXAMPLE We have $\sum_{n=0}^{N} r^n = \dfrac{1 - r^{N+1}}{1 - r}$ unless $r = 1$ in which case we have $\sum_{n=0}^{N} r^n = N + 1$. It is easy to see that we have

$$\sum_{n=0}^{\infty} r^n = \frac{1}{1 - r}$$

if and only if $|r| < 1$. In case $|r| \geq 1$ the series does not converge. □

EXAMPLE Another example where all the partial sums can be computed explicitly is

$$\begin{aligned}\sum_{n=1}^{N}\frac{1}{n(n+1)} &= \sum_{n=1}^{N}\left(\frac{1}{n}-\frac{1}{n+1}\right)\\ &= \left(\frac{1}{1}-\frac{1}{2}\right)+\left(\frac{1}{2}-\frac{1}{3}\right)+\left(\frac{1}{3}-\frac{1}{4}\right)+\cdots+\left(\frac{1}{N}-\frac{1}{N+1}\right)\\ &= \frac{1}{1}-\frac{1}{N+1}.\end{aligned}$$

What happens here is that the second term of each bracket cancels off with the first term of the following bracket. The only terms that do not cancel off in this way are the first term of the first bracket and the second term of the last bracket. We call this a ***telescoping sum*** . As $N \to \infty$ we find

$$\sum_{n=1}^{\infty}\frac{1}{n(n+1)} = 1$$

□

There is one important principle which applies to all series. We see that if a series converges, i.e. $s_n \longrightarrow s$ then we also have $s_{n-1} \longrightarrow s$ and it follows that $a_n = s_n - s_{n-1} \longrightarrow s - s = 0$. So if a series converges, then the sequence of terms of the series must converge to zero. Conversely, if the sequence of terms does not converge to 0 then the series cannot converge.

From the theorem on linear combinations of sequences, we have the following result for series.

THEOREM 38 *Let* $\sum_{n=1}^{\infty} a_n$ *and* $\sum_{n=1}^{\infty} b_n$ *be convergent series. Then so is* $\sum_{n=1}^{\infty}(ta_n + sb_n)$ *and*

$$\sum_{n=1}^{\infty}(ta_n + sb_n) = t\sum_{n=1}^{\infty} a_n + s\sum_{n=1}^{\infty} b_n.$$

It is also clear that the convergence of a series remains unchanged if only finitely many terms are altered. In fact, we have

THEOREM 39 *We have*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{N} a_n + \sum_{n=N+1}^{\infty} a_n \tag{2.1}$$

in the sense that if one of the infinite series converges, then so does the other and (2.1) *holds.*

Proof. For $k > N$ we have

$$\sum_{n=1}^{k} a_n = \sum_{n=1}^{N} a_n + \sum_{n=N+1}^{k} a_n,$$

and it suffices to let $k \to \infty$. ■

2.1 Series of Positive Terms

In the case that $a_n \geq 0$ for all n, we find that s_n is increasing (since $s_n - s_{n-1} = a_n \geq 0$). Now an increasing sequence of real numbers converges if and only if it is bounded above. Furthermore, an increasing sequence which is bounded above converges to the sup of the sequence, so we have

THEOREM 40 *If* $a_n \geq 0$. *Then* $\sum_{n=1}^{\infty} a_n = \sup_N \sum_{n=1}^{N} a_n$

If the partial sums are not bounded, then we may interpret the supremum as infinite and this gives us the notation $\sum_{n=1}^{\infty} a_n = \infty$ expressing the fact that the series does not converge, some times we say that the series ***diverges*** . Likewise, we write $\sum_{n=1}^{\infty} a_n < \infty$ to express the fact that the series does converge. These notations should only be used for series of positive terms.

There is an important corollary of the last two Theorems stated which will be used extensively later.

COROLLARY 41 *If* $a_n \geq 0$ *and* $\sum_{n=1}^{\infty} a_n < \infty$, *then* $\lim_{N\to\infty} \sum_{n=N}^{\infty} a_n = 0$.

There is a collection of recipes for deciding whether a series of positive terms converges or diverges.

Comparison Test : Suppose that $\sum_{n=1}^{\infty} a_n < \infty$ and that $0 \le b_n \le a_n$ for all n. Then $\sum_{n=1}^{\infty} b_n < \infty$. Obviously, we have

$$\sum_{n=1}^{N} b_n \le \sum_{n=1}^{N} a_n$$

for all N and it follows that

$$\sum_{n=1}^{\infty} b_n = \sup_N \sum_{n=1}^{N} b_n \le \sup_N \sum_{n=1}^{N} a_n = \sum_{n=1}^{\infty} a_n < \infty$$

EXAMPLE We have $\dfrac{1}{n^2} \le \dfrac{2}{n(n+1)}$, so we find

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \le \sum_{n=1}^{\infty} \frac{2}{n(n+1)} = 2.$$

□

The comparison test can also be turned around. If $\sum_{n=1}^{\infty} b_n = \infty$ and we have $0 \le b_n \le a_n$ for all n, then $\sum_{n=1}^{\infty} a_n = \infty$.

Limit Comparison Test : This is a more sophisticated version of the comparison test, so we give the most sophisticated version.

LEMMA 42 *Let $a_n > 0$ and $b_n \ge 0$. Suppose that $\sum_{n=1}^{\infty} a_n < \infty$ and that $\limsup_{n\to\infty} \dfrac{b_n}{a_n} < \infty$. Then $\sum_{n=1}^{\infty} b_n < \infty$*

Proof. Let $c = \limsup_{n\to\infty} \dfrac{b_n}{a_n}$. Then, taking $\epsilon = 1$ in the definition of $\lim\sup$, we have the existence of N such that

$$\frac{b_n}{a_n} \le c + 1$$

for $n > N$. We now get for $k > N$

$$\begin{aligned} \sum_{n=1}^{k} b_n = \sum_{n=1}^{N} b_n + \sum_{n=N+1}^{k} b_n &\le \sum_{n=1}^{N} b_n + (c+1) \sum_{n=N+1}^{k} a_n \\ &\le \sum_{n=1}^{N} b_n + (c+1) \sum_{n=1}^{\infty} a_n \qquad (2.2) \end{aligned}$$

Since the member in (2.2) is finite and independent of k we have shown that the partial sums $\sum_{n=1}^{k} b_n$ are bounded. ■

Similarly, the limit comparison test can be turned around to show the divergence of one series from the divergence of another. Since we have a point to make, let's actually write that down explicitly.

LEMMA 43 *Let $a_n > 0$ and $b_n \geq 0$. Suppose that $\sum_{n=1}^{\infty} a_n = \infty$ and that $\liminf\limits_{n\to\infty} \dfrac{b_n}{a_n} > 0$. Then $\sum_{n=1}^{\infty} b_n = \infty$*

The point to note here is that the $\limsup$ gets changed into a $\liminf$.

Ratio Test

LEMMA 44 *Suppose that $a_n > 0$ and that $\limsup_{n\to\infty} \dfrac{a_{n+1}}{a_n} < 1$. Then $\sum_{n=1}^{\infty} a_n < \infty$.*

Proof. Again we play the sandwich game. Let $\limsup_{n\to\infty} \dfrac{a_{n+1}}{a_n} < r < 1$. Then, there exists $N \in \mathbb{N}$ such that $\dfrac{a_{n+1}}{a_n} < r$ for $n \geq N$. But now a simple induction shows that $a_{N+k} \leq a_N r^k$ for $k \in \mathbb{Z}^+$. So, for $n \geq N$ we have $a_n \leq a_N r^{n-N}$. Again the limit comparison test shows that $\sum_{n=1}^{\infty} a_n < \infty$ since we know that $\sum_{n=1}^{\infty} r^n < \infty$. ■

The converse part of the ratio test is given as follows.

LEMMA 45 *Suppose that $a_n > 0$ and that $\liminf_{n\to\infty} \dfrac{a_{n+1}}{a_n} > 1$. Then $\sum_{n=1}^{\infty} a_n = \infty$.*

Proof. Let r be such that $\liminf_{n\to\infty} \dfrac{a_{n+1}}{a_n} > r > 1$. Then, there exists $N \in \mathbb{N}$ such that $\dfrac{a_{n+1}}{a_n} > r$ for $n \geq N$. But now a simple induction shows that $a_{N+k} \geq a_N r^k$ for $k \in \mathbb{Z}^+$. So, for $n \geq N$ we have $a_n \geq a_N r^{n-N} \geq a_N > 0$. The terms of the series no not tend to zero and consequently, the series does not converge. ■

Note that if you have $\lim_{n\to\infty} \dfrac{a_{n+1}}{a_n} = 1$ then you cannot apply the ratio test.

Root Test :

LEMMA 46 *Suppose that* $a_n \geq 0$ *and that* $\limsup_{n\to\infty}(a_n)^{\frac{1}{n}} < 1$. *Then* $\sum_{n=1}^{\infty} a_n < \infty$.

Proof. We can find a number r which can be sandwiched

$$\limsup_{n\to\infty}(a_n)^{\frac{1}{n}} < r < 1.$$

Now, there exists $N \in \mathbb{N}$ such that $(a_n)^{\frac{1}{n}} < r$ for $n > N$. So, $a_n < r^n$ for $n > N$ and the limit comparison test shows that $\sum_{n=1}^{\infty} a_n < \infty$ since we know that $\sum_{n=1}^{\infty} r^n < \infty$. ■

Again we have a result in the opposite direction.

LEMMA 47 *Suppose that* $a_n \geq 0$ *and that* $\limsup_{n\to\infty}(a_n)^{\frac{1}{n}} > 1$. *Then* $\sum_{n=1}^{\infty} a_n = \infty$.

Here there is a very remarkable contrast with the limit comparison test and the ratio test. Note that for the root test it is still the $\limsup$ that figures in the converse part.

Proof. From the definition of the $\limsup$ we find a natural subsequence $(n_k)_{k=1}^{\infty}$ such that $(a_{n_k})^{\frac{1}{n_k}} > 1$. It follows that $a_{n_k} > 1$ and we cannot have that $a_n \longrightarrow 0$. It follows that $\sum_{n=1}^{\infty} a_n = \infty$.

Note that if you have $\limsup_{n\to\infty}(a_n)^{\frac{1}{n}} = 1$ then you cannot apply the root test.

Lemma 2 suggests that the root test is stronger than the ratio test and in fact this is the case. The advantage of the ratio test over the root test is that it is often considerably easier to apply.

Condensation Test :

The condensation test applies only in the case that a_n is positive and decreasing. It is fiendishly clever.

LEMMA 48 *Suppose that $a_n \geq 0$ and that $a_{n+1} \leq a_n$ for all $n \in \mathbb{N}$. Then*

$$\sum_{n=1}^{\infty} a_n < \infty \quad \Longleftrightarrow \quad \sum_{k=0}^{\infty} 2^k a_{2^k} < \infty \tag{2.3}$$

Proof. The idea is to bracket the series. We will develop this idea later. We write

$$s_7 = (a_1) + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) \leq a_1 + 2a_2 + 4a_4.$$

In each bracket, each term has been bounded above by the first term in the bracket. Obviously, the same argument can be used to show that

$$s_{2^K-1} \leq \sum_{k=0}^{K-1} 2^k a_{2^k}.$$

Thus the convergence of the series on the right of (2.3) implies that of the series on the left.

For the converse, we put the brackets in different places.

$$s_8 = (a_1) + (a_2) + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) \geq a_1 + a_2 + 2a_4 + 4a_8$$

This time, each term in a bracket is bounded below by the last term in the bracket. The generalization is

$$s_{2^K} \geq a_1 + \sum_{k=1}^{K} 2^{k-1} a_{2^k} = \frac{1}{2}\left(a_1 + \sum_{k=0}^{K} 2^k a_{2^k}\right)$$

The convergence of the series on the left of (2.3) now implies that of the series on the right. ■

EXAMPLE The Condensation test is the one which allows us to figure out which of the p-series converge. Let's suppose that $p > 0$ then obviously n^{-p} is decreasing as n increases. So $\sum_{n=1}^{\infty} n^{-p}$ converges iff $\sum_{k=0}^{\infty} 2^k 2^{-pk}$ does. But the second series is geometric and converges iff $p > 1$. Of course, if $p \leq 0$ then $\sum_{n=1}^{\infty} n^{-p}$ diverges because the terms do not tend to zero. The case $p = 1$

$$\sum_{n=1}^{\infty} \frac{1}{n} = \infty$$

is called the ***harmonic series*** . □

EXAMPLE It can also be shown that $\sum_n \frac{1}{n(\ln(n))^p}$ also converges iff $p > 1$. Applying the condensation test to this series yield a series you can compare with the previous example. In case you were wondering about the series

$$\sum_n \frac{1}{n\ln(n)(\ln(\ln(n)))^p},$$

well it also converges iff $p > 1$. □

Raabe's Test

This is a more powerful version of the ratio test.

LEMMA 49 *Suppose that $a_n > 0$ and that there exists $\alpha > 1$ and $N \in \mathbb{N}$ such that*

$$\frac{a_{n+1}}{a_n} \leq 1 - \frac{\alpha}{n}$$

for $n \geq N$. Then $\sum_{n=1}^{\infty} a_n < \infty$.

Proof. We can write the Raabe condition as

$$na_{n+1} \leq (n-1)a_n - (\alpha - 1)a_n$$

and then manipulate it into the form

$$(\alpha - 1)a_n \leq (n-1)a_n - na_{n+1}.$$

Now, for $K \geq N$ we have

$$\begin{aligned} (\alpha - 1)\sum_{n=N}^{K} a_n &\leq \sum_{n=N}^{K} \Big((n-1)a_n - na_{n+1}\Big) \qquad (2.4) \\ &= (N-1)a_N - Ka_{K+1} \\ &\leq (N-1)a_N \end{aligned}$$

The key point here is that the right-hand side of (2.4) is a telescoping sum. So, we finally get

$$s_K = s_{N-1} + \sum_{n=N}^{K} a_n \leq s_{N-1} + \frac{(N-1)a_N}{\alpha - 1}$$

and the right-hand side is independent of K. Hence the result. ■

There is also a converse version.

LEMMA 50 *Suppose that $a_n > 0$ and that*

$$\frac{a_{n+1}}{a_n} \geq 1 - \frac{1}{n}$$

for $n \geq N$. Then $\sum_{n=1}^{\infty} a_n = \infty$.

Proof. First, be sure to take $N > 1$. We rewrite the condition as $na_{n+1} \geq (n-1)a_n$ for $n \geq N$. Now a simple induction gives for $K \geq N$ that $Ka_{K+1} \geq (N-1)a_N$ or equivalently that $a_{K+1} \geq \dfrac{(N-1)a_N}{K}$ and it follows that $\sum_{n=1}^{\infty} a_n = \infty$ by limit comparison with the harmonic series. ∎

EXAMPLE Consider $\displaystyle\sum_{n=1}^{\infty} \frac{(2n)!}{(n!)^2 4^n}$. We get $\dfrac{a_{n+1}}{a_n} = \dfrac{2n+1}{2n+2}$ and it is clear that the ratio is too large. We find that $\dfrac{2n+1}{2n+2} \geq 1 - \dfrac{1}{n}$ is equivalent to $2n^2 + n \geq 2n^2 - 2$ which is always true for $n \geq 1$. So the series diverges. □

EXAMPLE Consider $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^2}$. We get $\dfrac{a_{n+1}}{a_n} = \dfrac{n^2}{(n+1)^2}$. Let us try for $\alpha = \dfrac{3}{2}$. We will need $\dfrac{n^2}{(n+1)^2} \leq 1 - \dfrac{3}{2n}$, or equivalently $2n^3 \leq (2n-3)(n+1)^2 = 2n^3 + n^2 - 4n - 3$, and this is true for $n \geq 5$ since $n^2 - 4n - 3 = (n-5)(n+1) + 2$. So the series converges. □

Notice that the ratio test would fail to give a conclusion for either of the two examples above.

Finally, despite this plethora of tests, sometimes the correct way to proceed is simply to show directly that the partial sums are bounded above.

EXAMPLE Let $(n_k)_{k=1}^{\infty}$ be the increasing enumeration of those nonnegative integers that do not have a 4 in their decimal expansion. We claim that $\displaystyle\sum_{k=1}^{\infty} \frac{1}{n_k} < \infty$. To see this we simply count the number of such integers from 10^j to $10^{j+1} - 1$ inclusive. These are the integers that have exactly $j+1$ digits in their decimal expansion. (For example, when $j = 2$, the integers from 100 to 999 are those that have a 3-digit expansion.) The first digit of such an n_k is one of $1, 2, 3, 5, 6, 7, 8, 9$ (8 choices) and the remaining digits are chosen from $0, 1, 2, 3, 5, 6, 7, 8, 9$ (9 choices). Hence, there are $8 \cdot 9^j$ integers of the form n_k in the given range. It follows that the sum of $\dfrac{1}{n_k}$ over these integers is bounded above by $\dfrac{8 \cdot 9^j}{10^j}$. Thus we have

$$\sum_{k=1}^{\infty} \frac{1}{n_k} \leq \sum_{j=0}^{\infty} 8(0.9)^j = 80 < \infty.$$

□

2.2 Signed Series

We now come to the convergence of general series of real numbers. The first line of approach is to determine if the series is absolutely convergent.

DEFINITION *A series $\sum_{n=1}^{\infty} a_n$ is* ***absolutely convergent*** *iff $\sum_{n=1}^{\infty} |a_n| < \infty$.*

THEOREM 51 *An absolutely convergent series is convergent and*

$$\left| \sum_{n=1}^{\infty} a_n \right| \leq \sum_{n=1}^{\infty} |a_n| \tag{2.5}$$

Proof. By Corollary 41, we have $\lim_{N \to \infty} \sum_{n=N}^{\infty} |a_n| = 0$. Let $\epsilon > 0$. Then, there exists N such that $\sum_{n=N}^{\infty} |a_n| < \epsilon$. (The sequence $(t_N = \sum_{n=N}^{\infty} |a_n|$ is decreasing in N, so we really only need just one term to be less than ϵ). Now, let $N \leq p \leq q$ and denote by $s_k = \sum_{n=1}^{k} a_n$. Then we have

$$|s_q - s_p| = \left| \sum_{n=p+1}^{q} a_n \right| \leq \sum_{n=p+1}^{q} |a_n| \leq \sum_{n=N}^{\infty} |a_n| < \epsilon$$

Thus, $(s_n)_{n=1}^{\infty}$ is a Cauchy sequence in $\mathbb{R}$ and hence converges.

So far, so good. Now we have to estimate the sum. We have

$$|s_N| = \left| \sum_{n=1}^{N} a_n \right| \leq \sum_{n=1}^{N} |a_n| \leq \sum_{n=1}^{\infty} |a_n| \tag{2.6}$$

But, as $N \to \infty$, $s_N \to \sum_{n=1}^{\infty} a_n$ and so $|s_N| \to |\sum_{n=1}^{\infty} a_n|$. It follows now that (2.5) holds by letting $N \to \infty$ in (2.6).

EXAMPLE The series $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n(n+1)}$ converges since the series $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ converges. □

EXAMPLE There is also a vector-valued version of Theorem 51. Let V be a complete normed space. Let v_n be vectors in V such that $\sum_{n=1}^{\infty} \|v_n\| < \infty$, then $\sum_{n=1}^{\infty} v_n$ converges in V and we have the inequality

$$\left\| \sum_{n=1}^{\infty} v_n \right\| \leq \sum_{n=1}^{\infty} \|v_n\|$$

You cannot dispense with the completeness of V in this result. □

There are some series that converge but do not converge absolutely. Such series are called ***conditionally convergent*** and their convergence depends upon cancellation of terms. Before we discuss this, it should be made clear that we usually check for absolute convergence first. If absolute convergence fails, then we may be able to conclude immediately that the series does not converge. The ratio test and the root test establish divergence by showing that the terms do not tend to zero and if $|a_n| \underset{n\to\infty}{\longrightarrow} 0$ fails, then $a_n \underset{n\to\infty}{\longrightarrow} 0$ also fails and the original (signed) series cannot converge.

EXAMPLE Consider $\sum_{n=1}^{\infty} (-1)^n \frac{(2n)!}{(n!)^2 3^n}$. We get $\frac{|a_{n+1}|}{|a_n|} = \frac{2(2n+1)}{3(n+1)} \to \frac{4}{3}$. The ratio test shows that $\sum_{n=1}^{\infty} \frac{(2n)!}{(n!)^2 3^n} = \infty$ by showing that $\frac{(2n)!}{(n!)^2 3^n} \underset{n\to\infty}{\longrightarrow} 0$ fails. It follows that the terms of the original signed series do not tend to zero and so the signed series must also fail to converge. □

2.3 Alternating Series

There is a quite remarkable result which goes by the name of the ***alternating series test*** . We will write an alternating series in the form

$$\sum_{n=1}^{\infty} (-)^{n-1} a_n = a_1 - a_2 + a_3 - a_4 + a_5 - a_6 + \cdots \tag{2.7}$$

The notation $(-)^n$ which bothers some people means $+$ if n is even and $-$ if n is odd.

THEOREM 52 *Suppose that the series in* (2.7) *satisfies*

- *The series is alternating, i.e.* $a_n > 0$ *for all* $n \in \mathbb{N}$.
- *The terms are decreasing in absolute value, i.e.* $a_{n+1} \le a_n$ *for all* $n \in \mathbb{N}$.
- *We have* $\lim_{n\to\infty} a_n = 0$.

Then the series $\sum_{n=1}^{\infty}(-)^{n-1}a_n$ *converges.*

Proof.

Let $s_k = \sum_{n=1}^{k}(-)^{n-1}a_n$. The proof is based on the following three observations

- If k is odd, then $s_{k+1} < s_k$ since $s_{k+1} = s_k - a_{k+1}$.
- The odd partial sums are decreasing. For k odd, $s_{k+2} = s_k - (a_{k+1} - a_{k+2})$.
- The even partial sums are increasing. For k even, $s_{k+2} = s_k + (a_{k+1} - a_{k+2})$.

From these we deduce that if k is odd, then $s_k > s_{k+1} \ge s_2$, the right-hand inequality is because the even partial sums are increasing. Similarly, if k is even, then $s_k < s_{k-1} \le s_1$. The right-hand inequality is because the odd partial sums are decreasing. So the subsequence of odd partial sums is decreasing and bounded below by s_2 and the subsequence of even partial sums is increasing and bounded above by s_1. We can show this symbolically by

$$s_2 < s_4 < s_6 < s_8 < \cdots < s_9 < s_7 < s_5 < s_3 < s_1$$

So both subsequences converge say to s_{odd} and s_{even} respectively. But, $s_{2k+1} - s_{2k} = a_{2k+1} \to 0$, so that $s_{\text{odd}} = s_{\text{even}}$. It follows that s_k converges to the common value. ■

EXAMPLE The series $\displaystyle\sum_{n=2}^{\infty}(-)^n \frac{1}{\ln n}$ converges. Note that this series converges very slowly. □

EXAMPLE We show that e is irrational. Let us suppose that $e = \dfrac{p}{q}$ where $p, q \in \mathbb{N}$. Then $e^{-1} = \dfrac{q}{p}$. Choose $N \in \mathbb{N}$ with $2N \geq p$. We will use the alternating series

$$e^{-1} = \sum_{n=2}^{\infty} (-)^n \frac{1}{n!}$$

We have $s_{2N-1} < e^{-1} = \dfrac{q}{p} < s_{2N}$, for $N \geq 2$ and where we have denoted $s_M = \displaystyle\sum_{n=2}^{M} (-)^n \frac{1}{n!}$. This is a bit confusing. Because the series starts with a positive term at $n = 2$, it is the odd partial sums that are small and the even ones which are large! We get

$$(2N)!\, s_{2N-1} < (2N)!\, \frac{q}{p} < (2N)!\, s_{2N}. \tag{2.8}$$

All three terms in (2.8) are integers and $(2N)!\, s_{2N} = (2N)!\, s_{2N-1} + 1$. This is a contradiction. □

2.4 Bracketting Series

Obviously, the alternating series test has limited applicability. Not all signed series will have their signs neatly alternating. A bracketted series has brackets placed around groups of terms of the original series. Then each sum within a bracket becomes a term in the bracketted series. Thus, starting from

$$\sum_{n=1}^{\infty} a_n = a_1 + a_2 + a_3 + \cdots$$

we insert brackets to obtain

$$(a_1 + a_2 + \cdots + a_{n_1}) + (a_{n_1+1} + a_{n_1+2} + \cdots + a_{n_2}) + \cdots$$

Thus, using the convention $n_0 = 0$ we have the formula $b_k = \sum_{n=n_{k-1}+1}^{n_k} a_n$ for the kth term of the bracketted series. Let us denote $t_\ell = \sum_{k=1}^{\ell} b_k$, then it is clear that $t_\ell = s_{n_\ell}$. Expressed in words this says that the ℓth partial sum of the bracketted series is the n_ℓth partial sum of the original series, (t_ℓ) is a subsequence of (s_n). Hence, if the original series converges, so does the bracketted series.

Usually, we want to go in the opposite direction. Suppose that we have shown that the bracketted series converges, say to t. We want to be able to deduce that the original series converges. This is not automatic, it requires an additional condition. Let us suppose that $n_k + 1 \leq n \leq n_{k+1}$, then we define $\alpha_n = \sum_{m=n_k+1}^{n} a_m = s_n - t_k$. The additional condition needed is that $\alpha_n \longrightarrow 0$ as $n \longrightarrow \infty$. As $n \longrightarrow \infty$, we get that $k \longrightarrow \infty$ (since each bracket contains at least one term) and so $s_n = t_k + \alpha_n \longrightarrow t + 0 = t$. The good thing about α_n is that it involves terms which live only in a single bracket.

EXAMPLE The example we look at here is

$$\frac{1}{1} - \frac{1}{2} - \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} - \frac{1}{7} - \frac{1}{8} - \frac{1}{9} - \frac{1}{10} + \frac{1}{11} + \cdots$$

where each block of signs increases in length by one at each step. This series is not absolutely convergent because the harmonic series diverges. The signs are not alternating, so the alternating series test cannot be applied at least directly. One possible approach might be to bracket the terms to make an alternating series as in

$$\left(\frac{1}{1}\right) - \left(\frac{1}{2} + \frac{1}{3}\right) + \left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6}\right) - \left(\frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \frac{1}{10}\right) + \cdots \tag{2.9}$$

but the rather stringent conditions of the alternating series test do not make this appealing. We therefore take our brackets as

$$\left(\frac{1}{1} - \frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{4} + \frac{1}{5} + \frac{1}{6} - \frac{1}{7} - \frac{1}{8} - \frac{1}{9} - \frac{1}{10}\right) + \cdots$$

and actually, this makes more sense because the idea of bracketting is usually to capture cancellation of terms within a bracket. So usually, there should be terms of both signs within each bracket.

Some fairly horrible calculations now give

$$b_k = \sum_{n=2k^2-3k+2}^{2k^2-k} \frac{1}{n} - \sum_{n=2k^2-k+1}^{2k^2+k} \frac{1}{n} \tag{2.10}$$

there being $2k-1$ terms in the first sum and $2k$ terms in the second sum. We now rearrange the right hand side of (2.10) by combining each term of the first sum with the corresponding term in the second sum. Since the second sum has one more term than the first sum, the last term of the second sum remains unmatched. Combining the terms in this way captures the cancellation well enough for us to establish convergence.

$$b_k = \left\{ \sum_{n=2k^2-3k+2}^{2k^2-k} \left(\frac{1}{n} - \frac{1}{n+2k-1} \right) \right\} - \frac{1}{2k^2+k}$$

$$= \left\{ \sum_{n=2k^2-3k+2}^{2k^2-k} \frac{2k-1}{n(n+2k-1)} \right\} - \frac{1}{2k^2+k}$$

$$|b_k| \leq \left\{ \sum_{n=2k^2-3k+2}^{2k^2-k} \frac{2k-1}{n(n+2k-1)} \right\} + \frac{1}{2k^2+k}$$

$$\leq \frac{(2k-1)^2}{(2k^2-3k+2)(2k^2-k+1)} + \frac{1}{2k^2+k}$$

So $\displaystyle\sum_{k=1}^{\infty} |b_k| < \infty$ by comparison with $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$ and we find that $\sum_{k=1}^{\infty} b_k$ converges absolutely.

So far, so good. Now, if $n_k + 1 \leq n \leq n_{k+1}$, then we can bound $|\alpha_n|$ by the sum of the absolute values of the terms in the $(k+1)$st bracket.

$$|\alpha_n| \leq \sum_{n=2k^2+k+1}^{2k^2+5k+3} \frac{1}{n} \leq \frac{4k+3}{2k^2+k+1} \longrightarrow 0$$

as $n \longrightarrow \infty$. This shows that the original series converges. □

EXAMPLE Consider next

$$\frac{1}{\sqrt{1}} - \frac{1}{\sqrt{2}} - \frac{1}{\sqrt{3}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{6}} - \frac{1}{\sqrt{7}} - \frac{1}{\sqrt{8}} - \frac{1}{\sqrt{9}} - \frac{1}{\sqrt{10}} + \frac{1}{\sqrt{11}} + \cdots$$

where again each block of signs increases in length by one at each step. Put the brackets as

$$(1) - \left(\frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} \right) + \left(\frac{1}{\sqrt{4}} + \frac{1}{\sqrt{5}} + \frac{1}{\sqrt{6}} \right) - \left(\frac{1}{\sqrt{7}} + \frac{1}{\sqrt{8}} + \frac{1}{\sqrt{9}} + \frac{1}{\sqrt{10}} \right) + \cdots$$

Now, estimate the absolute value of b_k from below by using the last term in each bracket

$$|b_k| \geq \frac{k}{\sqrt{\frac{k(k+1)}{2}}} = \sqrt{\frac{2k}{k+1}} \geq 1$$

and we see that the $|b_k|$ are bounded away from zero. So the bracketted series does not converge and hence neither does the original one. □

2.5 Summation by Parts

Another approach to conditionally convergent series is given by the summation by parts formula. Let's start by deriving that formula. We wish to study the series

$$\sum_{n=1}^{\infty} a_n b_n$$

where we have a good grip on the partial sums $s_N = \sum_{n=1}^{N} a_n$ of the series $\sum_{n=1}^{\infty} a_n$. We will denote by t_N the partial sums of the series we are studying

$$t_N = \sum_{n=1}^{N} a_n b_n.$$

Now, for $M > N$ we get

$$\begin{aligned} t_M - t_N &= \sum_{n=N+1}^{M} a_n b_n \\ &= \sum_{n=N+1}^{M} (s_n - s_{n-1}) b_n && (2.11) \\ &= \sum_{n=N+1}^{M} s_n b_n - \sum_{n=N+1}^{M} s_{n-1} b_n && (2.12) \\ &= \sum_{n=N+1}^{M} s_n b_n - \sum_{n=N}^{M-1} s_n b_{n+1} && (2.13) \end{aligned}$$

$$= s_M b_M - s_N b_{N+1} + \sum_{n=N+1}^{M-1} s_n(b_n - b_{n+1}) \tag{2.14}$$

In (2.11) we have replaced a_n by $s_n - s_{n-1}$. In (2.12) we have multiplied out in (2.11) and distributed the sum. In (2.13) we have left the first summation alone and changed the summation variable from n to $n+1$ in the second summation. This is reflected both in the variable change $n \to n+1$ and in the change in the limits of summation. Finally in (2.14) we have written down first the terms corresponding to either $n = N$ or $n = M$ and then written all the remaining terms (in the range $N+1 \le n \le M-1$) as a combined summation.

THEOREM 53 *Suppose that*

- $\sum_{n=1}^{\infty} s_n(b_n - b_{n+1})$ *converges.*
- $s_n b_n \longrightarrow 0$ *as* $n \longrightarrow \infty$.

Then $\sum_{n=1}^{\infty} a_n b_n$ *converges and equals* $\sum_{n=1}^{\infty} s_n(b_n - b_{n+1})$.

Proof. Putting $N = 0$ into (2.14) we get

$$\sum_{n=1}^{M} a_n b_n = s_M b_M + \sum_{n=1}^{M-1} s_n(b_n - b_{n+1})$$

since $s_0 = 0$. We now let $M \longrightarrow \infty$. ■

As with bracketted series, the key idea of summation by parts is to capture cancellation. We typically choose the a_n to have terms of both signs and also so that we have a precise formula for s_n. This is how the cancellation is captured. On the other hand, the b_n should vary only slightly, so that the quantities $|b_n - b_{n+1}|$ may be relatively small.

EXAMPLE A very celebrated series[1] is

$$\sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)t$$

[1] What does not come out in this discussion is that the sum of the series is quite simply $\frac{\pi}{4} \operatorname{sgn}(\sin(t))$.

for $t \in \mathbb{R}$. If t is an integer multiple of π, then the series vanishes identically and convergence is trivial. If not, then the series is not absolutely convergent. This is not quite obvious, but certainly the estimate $|\sin(2n-1)t| \leq 1$ does not yield absolute convergence.

Let us suppose that t is not an integer multiple of π so that $\sin t \neq 0$. We choose to take $a_n = \sin(2n-1)t$ and $b_n = \dfrac{1}{2n-1}$. We have

$$\begin{aligned} 2\sin t\sin(2n-1)t &= \cos((2n-1)t-t) - \cos((2n-1)t+t) \\ &= \cos((2n-2)t) - \cos(2nt). \end{aligned}$$

We get a telescoping sum

$$2\sin t\sum_{n=1}^{N}\sin(2n-1)t = \sum_{n=1}^{N}\Big(\cos((2n-2)t) - \cos(2nt)\Big) = 1-\cos(2Nt)$$

and it follows that

$$s_N = \sum_{n=1}^{N}\sin(2n-1)t = \frac{1-\cos(2Nt)}{2\sin t}.$$

So, we have for all n that $|s_n| \leq \dfrac{1}{|\sin t|}$. Since $b_n \longrightarrow 0$ we have that $s_n b_n \longrightarrow 0$ as $n \longrightarrow \infty$. On the other hand, we have

$$\sum_{n=1}^{\infty} s_n(b_n - b_{n+1}) = \sum_{n=1}^{\infty}\frac{1-\cos 2nt}{(2n-1)(2n-3)\sin t}$$

and the right hand side is absolutely convergent. Hence the series actually converges for all real t. □

EXAMPLE As a second example, let's rework an example for which we used the bracketting method.

$$\frac{1}{1}-\frac{1}{2}-\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}-\frac{1}{7}-\frac{1}{8}-\frac{1}{9}-\frac{1}{10}+\frac{1}{11}+\cdots$$

Let's denote this as $\displaystyle\sum_{n=1}^{\infty}\omega_n\frac{1}{n}$ where $\omega_n = \pm 1$ is the sign of the term. Let us define $k(n)$ to be the unique positive integer k such that $\dfrac{k(k-1)}{2} < n \leq \dfrac{k(k+1)}{2}$.

This means that the term $\omega_n \frac{1}{n}$ is in the kth bracket of (2.9). Now we write $a_n = \omega_n \frac{1}{k(n)}$ and $b_n = \frac{k(n)}{n}$. Then it is easy to see that $0 \leq s_n \leq 1$ for all n. It is also clear that $b_n \longrightarrow 0$ as $n \longrightarrow 0$. So, it remains to show that $\sum_{n=1}^{\infty} |b_n - b_{n+1}| < \infty$. Now if n is not a triangular number we get

$$b_n - b_{n+1} = \frac{k(n)}{n} - \frac{k(n)}{n+1} = \frac{k(n)}{n(n+1)} \sim n^{-\frac{3}{2}}.$$

On the other hand, if n is a triangular number i.e. $n = \frac{k(n)(k(n)+1)}{2}$, then

$$b_n - b_{n+1} = \frac{k(n)}{n} - \frac{k(n)+1}{n+1} = -\frac{n-k(n)}{n(n+1)} \sim k(n)^{-2}.$$

It follows that $\sum_{n=1}^{\infty} |b_n - b_{n+1}| \leq C \left(\sum_{n=1}^{\infty} n^{-\frac{3}{2}} + \sum_{k=1}^{\infty} k^{-2} \right) < \infty$ for some suitably chosen constant C. $\square$

2.6 Rearrangements

We start this section with an example.

EXAMPLE Consider the following two series.

$$\overbrace{\frac{1}{1} - \frac{1}{2} - \frac{1}{2}} + \overbrace{\frac{1}{2} - \frac{1}{4} - \frac{1}{4}} + \overbrace{\frac{1}{2} - \frac{1}{4} - \frac{1}{4}}$$
$$+ \overbrace{\frac{1}{4} - \frac{1}{8} - \frac{1}{8}} + \overbrace{\frac{1}{4} - \frac{1}{8} - \frac{1}{8}} + \overbrace{\frac{1}{4} - \frac{1}{8} - \frac{1}{8}} + \overbrace{\frac{1}{4} - \frac{1}{8} - \frac{1}{8}}$$
$$+ \cdots \tag{2.15}$$

and

$$1 \overbrace{-\frac{1}{2} + \frac{1}{2}} \overbrace{-\frac{1}{2} + \frac{1}{2}} \overbrace{-\frac{1}{4} + \frac{1}{4}} \overbrace{-\frac{1}{4} + \frac{1}{4}} \overbrace{-\frac{1}{4} + \frac{1}{4}} \overbrace{-\frac{1}{4} + \frac{1}{4}}$$
$$\overbrace{-\frac{1}{8} + \frac{1}{8}} \overbrace{-\frac{1}{8} + \frac{1}{8}} \overbrace{-\frac{1}{8} + \frac{1}{8}} \overbrace{-\frac{1}{8} + \frac{1}{8}} \overbrace{-\frac{1}{8} + \frac{1}{8}} \overbrace{-\frac{1}{8} + \frac{1}{8}}$$
$$\overbrace{-\frac{1}{8} + \frac{1}{8}} \overbrace{-\frac{1}{8} + \frac{1}{8}} - \cdots \tag{2.16}$$

In (2.15) the terms are grouped in threes and the number of occurrences of a given group doubles at each step. In (2.16), apart from the first term, the terms are grouped in pairs and again, the number of occurrences of each group doubles at each step. By placing brackets around the groups, it is easy to see that both series converge, (2.15) converges to 0 and (2.16) converges to 1. The big deal here is that the two series have exactly the same terms, but they are jumbled up. So, for instance the term $+\dfrac{1}{4}$ occurs in both series exactly 4 times etc. We could with some trouble define a bijection $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ such that if the first series is denoted $\sum_{n=1}^{\infty} a_n$ then the second one is $\sum_{n=1}^{\infty} a_{\sigma(n)}$. In this case, we call the second series a ***rearrangement*** of the first. What this example shows is that rearranged series do not always have the same sum. Even worse, we could define a third series

$$1 \overbrace{-\frac{1}{2}-\frac{1}{2}} \overbrace{+\frac{1}{2}+\frac{1}{2}} \overbrace{-\frac{1}{4}-\frac{1}{4}-\frac{1}{4}-\frac{1}{4}} \overbrace{+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}+\frac{1}{4}} - \cdots \tag{2.17}$$

which is a rearrangement of both (2.15) and (2.16) and does not converge at all! □

So, what can be shown about rearrangements? We have the following three results.

THEOREM 54 *Let $a_n \geq 0$ for all $n \in \mathbb{N}$ and let $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ be a bijection. Then*

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma(n)} \tag{2.18}$$

in the sense that if one series converges, then so does the other and the values of the sum are equal.

Proof. Let $N \in \mathbb{N}$. Then define $M = \max\{\sigma(n); n = 1 \ldots N\}$. Then clearly, $\{\sigma(1), \sigma(2), \ldots, \sigma(N)\} \subseteq \{1, 2, \ldots, M\}$. It follows that

$$\sum_{n=1}^{N} a_{\sigma(n)} \leq \sum_{m=1}^{M} a_m$$

because every term on the left hand side also appears on the right and the terms that appear on the right but not on the left are nonnegative. Thus,

$$\sum_{n=1}^{\infty} a_{\sigma(n)} = \sup_N \sum_{n=1}^{N} a_{\sigma(n)} \leq \sup_M \sum_{m=1}^{M} a_m = \sum_{m=1}^{\infty} a_m.$$

So, if the series on the left of (2.18) converges, then so does the series on the right. If the series on the right of (2.18) diverges, then so does the series on the left. But defining $b_n = a_{\sigma(n)}$ we have that $a_n = b_{\sigma^{-1}(n)}$ and σ^{-1} is also a bijection. Thus the original series $\sum_{n=1}^{\infty} a_n$ is also a rearrangement of $\sum_{n=1}^{\infty} a_{\sigma(n)}$ and the reverse inequality

$$\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_{\sigma^{-1}(n)} \leq \sum_{m=1}^{\infty} b_m = \sum_{m=1}^{\infty} a_{\sigma(m)}$$

also holds. Hence we have equality in (2.18). ∎

THEOREM 55 *Let $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ be a bijection. Then if the series $\sum_{n=1}^{\infty} a_n$ converges absolutely, then so does $\sum_{n=1}^{\infty} a_{\sigma(n)}$ and $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma(n)}$ holds.*

Proof. The statement about absolute convergence follows immediately from Theorem 54. It remains to show that the sums are the same. Towards this, let $\epsilon > 0$. Then, since $\sum_{n=1}^{\infty} |a_{\sigma(n)}| < \infty$ there exists N such that $\sum_{n>N} |a_{\sigma(n)}| < \epsilon$. We also then have

$$\left| \sum_{n=1}^{\infty} a_{\sigma(n)} - \sum_{n=1}^{N} a_{\sigma(n)} \right| = \left| \sum_{n>N} a_{\sigma(n)} \right| \leq \sum_{n>N} |a_{\sigma(n)}| < \epsilon. \tag{2.19}$$

Now, let $M \geq \max\{\sigma(n); n = 1 \ldots N\}$, so that $\{\sigma(1), \sigma(2), \ldots, \sigma(N)\} \subseteq \{1, 2, \ldots, M\}$. We find that

$$\sum_{m=1}^{M} a_m - \sum_{n=1}^{N} a_{\sigma(n)} = \sum_{m \in Z} a_m = \sum_{n \in \sigma^{-1}(Z)} a_{\sigma(n)}$$

where $Z = \{1, 2, \ldots, M\} \setminus \{\sigma(1), \sigma(2), \ldots, \sigma(N)\}$. But the finite set $\sigma^{-1}(Z)$ is contained in $\{N+1, N+2, \ldots\}$ and it follows that

$$\left| \sum_{m=1}^{M} a_m - \sum_{n=1}^{N} a_{\sigma(n)} \right| \leq \sum_{n \in \sigma^{-1}(Z)} |a_{\sigma(n)}| \leq \sum_{n>N} |a_{\sigma(n)}| < \epsilon \tag{2.20}$$

Combining (2.19) and (2.20) with the triangle inequality we find that

$$\left|\sum_{m=1}^{M} a_m - \sum_{n=1}^{\infty} a_{\sigma(n)}\right| < 2\epsilon$$

for all M such that $M \geq \max\{\sigma(n); n = 1 \dots N\}$. Letting $M \to \infty$, we finally obtain

$$\left|\sum_{m=1}^{\infty} a_m - \sum_{n=1}^{\infty} a_{\sigma(n)}\right| \leq 2\epsilon$$

and since $\epsilon > 0$ can be as small as we wish, the two sums must be equal. ■

The final theorem in this group is interesting, but we do not know of any practical applications, so we omit the proof.

THEOREM 56 *Let $\sum_{n=1}^{\infty} a_n$ be a conditionally convergent series of real numbers. (Specifically, this means that it is convergent, but not absolutely convergent). Let s be any real number. Then, there is a bijection $\sigma : \mathbb{N} \longrightarrow \mathbb{N}$ such that the series*

$$\sum_{n=1}^{\infty} a_{\sigma(n)}$$

converges and has sum s.

•2.7 Unconditional Summation

The material in this section is not normally included in analysis texts. Infinite sums depend on the order of the terms. We may feel that this is unnatural and ask if it would be possible to define an infinite sum that does not depend on the ordering of terms. The answer is yes, but the properties are rather disappointing.

So in this section, X is an index set and for each $x \in X$ we have a real number a_x (so that a is really a function $a : X \longrightarrow \mathbb{R}$). There is no ordering or structure of any kind on X. We want to define $\sum_{x\in X} a_x$. Let us denote $s_F = \sum_{x\in F} a_x$ for every finite subset F of X. This does have a valid meaning.

DEFINITION *We say that $\sum_{x\in X} a_x$ **converges unconditionally** to a number s if for every $\epsilon > 0$, there exists a finite subset F of X such that for every finite subset G of X with $G \supseteq F$, we have $|s - s_G| < \epsilon$.*

The sum s is unique, for if there is another sum s', then, for any $\epsilon > 0$ we can find finite sets F, F' such that

$$F \subseteq G \text{ finite} \subseteq X \Longrightarrow |s_G - s| < \epsilon$$
$$F' \subseteq G \text{ finite} \subseteq X \Longrightarrow |s_G - s'| < \epsilon$$

It now suffices to take $G = F \bigcup F'$ to deduce that $|s - s'| < 2\epsilon$ and, since ϵ is an arbitrary positive number, we must have $s = s'$.

LEMMA 57 *Suppose that $\sum_{x \in X} a_x$ is unconditionally convergent to s. Then*

- *There is a countable subset C of X such that $a_x = 0$ for all $x \in X \setminus C$.*
- *For any enumeration of C (i.e. bijective mapping) $\varphi : \mathbb{N} \longrightarrow C$, the series $\sum_{n=1}^{\infty} a_{\varphi(n)}$ is absolutely convergent and converges to s.*

Proof. Let (ϵ_j) be a sequence of positive numbers tending to zero. Then there exist finite subsets F_j of X such that

$$F_j \subseteq G \text{ finite} \subseteq X \Longrightarrow |s_G - s| < \epsilon_j.$$

Let us put $C = \bigcup_{j=1}^{\infty} F_j$ a countable subset of X. Let us suppose that there exists $x \in X \setminus C$ such that $a_x \neq 0$. Then choose j such that $2\epsilon_j < |a_x|$. Take for G the sets F_j and $F_j \bigcup \{x\}$. Then we get

$$2\epsilon_j < |a_x| = |s_{F_j \cup \{x\}} - s_{F_j}| \leq |s_{F_j \cup \{x\}} - s| + |s - s_{F_j}| < 2\epsilon_j$$

a contradiction.

Next we show the absolute convergence. Let us take $\epsilon = 1$, then there is a finite subset F such that

$$F \subseteq G \text{ finite} \subseteq X \Longrightarrow |s_G - s| < 1.$$

We will show that for every finite G we have $\sum_{x \in G} |a_x| \leq 2 + \sum_{x \in F} |a_x|$. Let $C_+ = \{x; a_x > 0\}$ and $C_- = \{x; a_x < 0\}$ so that $C = C_+ \bigcup C_-$ and $\emptyset = C_+ \bigcap C_-$. Let $H = G \setminus F$ and $H_\pm = H \bigcap C_\pm$. Then we have

$$|s_{F \cup H_+} - s| < 1 \text{ and } |s_{F \cup H_-} - s| < 1.$$

Subtracting off gives

$$\sum_{x\in H}|a_x| = s_{H_+} - s_{H_-} = s_{F\cup H_+} - s_{F\cup H_-} = (s_{F\cup H_+} - s) - (s_{F\cup H_-} - s) < 2$$

Since $G \subseteq F \bigcup H$ we finally get

$$\sum_{x\in G}|a_x| \leq \sum_{x\in H}|a_x| + \sum_{x\in F}|a_x| \leq 2 + \sum_{x\in F}|a_x|$$

as required. It follows that the series $\sum_{n=1}^{\infty}|a_{\varphi(n)}| < \infty$.

The final step is to show that $s = \sum_{n=1}^{\infty} a_{\varphi(n)}$. Let $\epsilon > 0$. Find $N \in \mathbb{N}$ such that $\sum_{n>N}|a_{\varphi(n)}| < \epsilon$ and a finite set F such that

$$F \subseteq G \text{ finite} \subseteq X \Longrightarrow |s_G - s| < \epsilon.$$

Now, let $G = F \bigcup \{\varphi(1), \varphi(2), \ldots, \varphi(N)\}$. Then $s_G = \sum_{n=1}^{N} a_{\varphi(n)} + \sum_{n\in T} a_{\varphi(n)}$ where T is a finite subset of $\{N+1, N+2, \ldots\}$. So we find

$$|s_G - \sum_{n=1}^{N} a_{\varphi(n)}| \leq \sum_{n\in T}|a_{\varphi(n)}| < \epsilon$$

Thus

$$|s - \sum_{n=1}^{N} a_{\varphi(n)}| < 2\epsilon$$

Finally, since $|\sum_{n=1}^{\infty} a_{\varphi(n)} - \sum_{n=1}^{N} a_{\varphi(n)}| \leq \sum_{n>N}|a_{\varphi(n)}| < \epsilon$, we get

$$|s - \sum_{n=1}^{\infty} a_{\varphi(n)}| < 3\epsilon.$$

But ϵ is an arbitrary positive number and we have our result. ■

2.8 Double Summation

We are familiar with the idea of a spreadsheet, really a matrix of real numbers. If we add down the columns and then add together all the column sums, we should get the same answer as we get from computing the row sums and then totalling them. For a finite matrix this works fine, but not for an infinite one as the following simple example shows.

$$\begin{array}{cccccc} 1 & -1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & -1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & -1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & -1 & \\ \vdots & \vdots & \vdots & & \ddots & \ddots \end{array}$$

All the row sums are zero as are the column sums, except the first which is 1. So, adding down the columns first gives the answer 1 and adding along the rows first gives the answer 0. In other words, in general it is false that

$$\sum_{p=1}^{\infty}\sum_{q=1}^{\infty} a_{p,q} = \sum_{q=1}^{\infty}\sum_{p=1}^{\infty} a_{p,q} \tag{2.21}$$

THEOREM 58 *If $a_{p,q} \geq 0$, then (2.21) holds in the sense that if one side of the equation is finite, so is the other and they are equal.*

Proof. Let us suppose that the right hand side is finite. Then each of the quantities

$$\alpha_q = \sum_{p=1}^{\infty} a_{p,q}$$

is finite and $\sum_{q=1}^{\infty} \alpha_q$ is also finite. Since $a_{p,q} \leq \alpha_q$ we see that for fixed p the series $\sum_{q=1}^{\infty} a_{p,q}$ converges. Now we have

$$\sum_{p=1}^{P}\sum_{q=1}^{\infty} a_{p,q} = \sum_{q=1}^{\infty}\sum_{p=1}^{P} a_{p,q} \leq \sum_{q=1}^{\infty}\sum_{p=1}^{\infty} a_{p,q} < \infty$$

since we know that convergent series are linear (use Theorem 38 and induction). Since the partial sums of the outer series on the left are bounded, we have convergence and

$$\sum_{p=1}^{\infty}\sum_{q=1}^{\infty} a_{p,q} \leq \sum_{q=1}^{\infty}\sum_{p=1}^{\infty} a_{p,q} < \infty$$

But, now we know that the left hand side is finite we can repeat the same argument to show

$$\sum_{q=1}^{\infty}\sum_{p=1}^{\infty} a_{p,q} \leq \sum_{p=1}^{\infty}\sum_{q=1}^{\infty} a_{p,q} < \infty$$

and (2.21) holds. Clearly arguing by contradiction, if one side is infinite, then so is the other. ■

For signed series an additional condition is needed. The following is a special case of a theorem due to Fubini.

THEOREM 59 *If* $\sum_{p=1}^{\infty}\sum_{q=1}^{\infty} |a_{p,q}| < \infty$, *then* (2.21) *holds.*

While this result is stated for real series, it is also true for complex series and with the same proof. There is also a version for complete normed vector spaces.

Proof. It is clear from the previous result that both left and right hand side of (2.21) are absolutely convergent. Let $\epsilon > 0$. Let $\alpha_q = \sum_{p=1}^{\infty} |a_{p,q}|$. Then, since $\sum_{q=1}^{\infty} \alpha_q < \infty$, there exists $Q \in \mathbb{N}$ such that $\sum_{q>Q} \alpha_q < \epsilon$. Also, for each q with $1 \leq q \leq Q$, there exists P_q such that

$$\sum_{p>P_q} |a_{p,q}| < \frac{\epsilon}{Q}.$$

Let P be any integer with $P \geq \max(P_1, P_2, \ldots, P_Q)$. Then we find

$$\begin{aligned}
\left|\sum_{q=1}^{\infty}\sum_{p=1}^{\infty} a_{p,q} - \sum_{q=1}^{\infty}\sum_{p=1}^{P} a_{p,q}\right| &= \left|\sum_{q=1}^{\infty}\sum_{p>P} a_{p,q}\right| \\
&\leq \sum_{q=1}^{\infty}\left|\sum_{p>P} a_{p,q}\right| \\
&\leq \sum_{q=1}^{\infty}\sum_{p>P} |a_{p,q}| \\
&\leq \sum_{q=1}^{Q}\sum_{p>P} |a_{p,q}| + \sum_{q>Q}\sum_{p>P} |a_{p,q}| \\
&\leq \sum_{q=1}^{Q}\frac{\epsilon}{Q} + \sum_{q>Q}\sum_{p=1}^{\infty} |a_{p,q}|
\end{aligned}$$

$$\leq \epsilon + \sum_{q>Q} \alpha_q < 2\epsilon$$

But, once again

$$\sum_{p=1}^{P} \sum_{q=1}^{\infty} a_{p,q} = \sum_{q=1}^{\infty} \sum_{p=1}^{P} a_{p,q}$$

and from this it follows that

$$\left| \sum_{q=1}^{\infty} \sum_{p=1}^{\infty} a_{p,q} - \sum_{p=1}^{P} \sum_{q=1}^{\infty} a_{p,q} \right| < 2\epsilon$$

whenever $P \geq \max(P_1, P_2, \ldots, P_Q)$. This shows (again) that the series on the right of (2.21) converges but (more to the point) that it converges to the left hand side of (2.21). ■

2.9 Infinite Products

We define inner products in much the same way as we define infinite sums. The partial products are defined by

$$p_N = \prod_{n=1}^{N} a_n$$

and the existence of the infinite product is equivalent to the convergence of the sequence (p_N). If we have $p_n \longrightarrow p$ as $n \longrightarrow \infty$ we write

$$p = \prod_{n=1}^{\infty} a_n \tag{2.22}$$

While in general, a_n can be arbitrary real numbers, usually they can be taken as positive. If the a_n are not eventually nonnegative, then the partial products will change signs infinitely often and the only possible limit will be zero. If just one of the a_n is zero, then the partial products will vanish eventually. So the only case of real interest is $a_n > 0$ for all n and we can then instead study the series

$$\sum_{n=1}^{\infty} \ln(a_n),$$

which is essentially equivalent to (2.22). Note however that we find $\prod_{n=1}^{\infty} a_n = 0$ if the partial sums of $\sum_{n=1}^{\infty} \ln(a_n)$ diverge properly to $-\infty$.

THEOREM 60 *Let $0 < a_n < 1$ for all $n \in \mathbb{N}$ Then $\prod_{n=1}^{\infty}(1 - a_n) > 0$ if and only if $\sum_{n=1}^{\infty} a_n < \infty$.*

Proof. So, converting the product to a sum, we must show that $\sum_{n=1}^{\infty} -\ln(1 - a_n) < \infty$ is equivalent to $\sum_{n=1}^{\infty} a_n < \infty$. For $0 < x < 1$ we have $x \leq -\ln(1-x)$, so the convergence of the product implies the convergence of the sum. However, if the sum converges, then the terms a_n must converge to zero and so eventually, (i.e. for n large enough) we have $0 < a_n < \frac{1}{2}$. In this range, $-\ln(1-x) \leq x\ln(4)$ and so convergence of the sum implies convergence of the product. See figure 2.1 for a graphical representation of the underlying inequalities. They are easily established using differential calculus. ∎

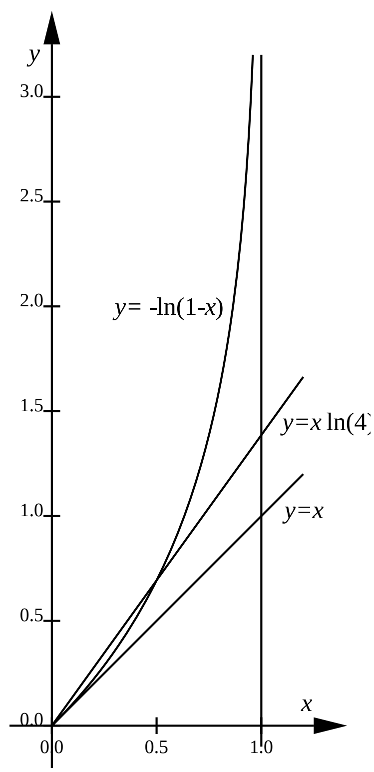

Figure 2.1: Comparison of $y = x$, $y = x\ln(4)$ and $y = -\ln(1 - x)$.

EXAMPLE The case $a_n = \dfrac{1}{n+1}$ leads to a telescoping product. Indeed, it is clear that $\displaystyle\prod_{n=1}^{\infty} \frac{n}{n+1} = 0$, so $\displaystyle\sum_{n=1}^{n} \frac{1}{n+1} = \infty$. This establishes the divergence of the harmonic series once again. □

EXAMPLE The series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{(2n)^2} < \infty$, so we have $\displaystyle\prod_{n=1}^{\infty} \left(1 - \frac{1}{(2n)^2}\right) > 0$. The actual value of this product is $\dfrac{2}{\pi}$ as we will see later (4.14). □

EXAMPLE Let $p_1 = 2$, $p_2 = 3$, $p_3 = 5$, $p_4 = 7$, $p_5 = 11$,... be the increasing enumeration of the prime numbers. We will show that

$$\sum_{k=1}^{\infty} \frac{1}{p_k} = \infty. \tag{2.23}$$

Let A be a large positive number. Then, since the harmonic series diverges, there exists a natural number N such that $\displaystyle\sum_{n=1}^{N} \frac{1}{n} \geq A$. Now choose a natural number K so large that every integer n with $1 \leq n \leq N$ has a (unique) factorization

$$n = \prod_{k=1}^{K} p_k^{\alpha_k}$$

where α_k are integers with $0 \leq \alpha_k \leq K$. In fact, we can take $K = N$. Then we have

$$\prod_{k=1}^{K} \left(1 + \frac{1}{p_k} + \frac{1}{p_k^2} + \cdots + \frac{1}{p_k^K}\right) \geq \sum_{n=1}^{N} \frac{1}{n} \geq A \tag{2.24}$$

because the left hand side can be multiplied out to give a sum $\displaystyle\sum_{n \in S} \frac{1}{n}$ where S is a subset of $\mathbb{N}$ containing $\{1, 2, \ldots, N\}$. Note that it is the uniqueness of the prime factorization which guarantees that each term $\dfrac{1}{n}$ occurs at most once. We deduce from (2.24) that

$$\prod_{k=1}^{K} \left(\frac{p_k}{p_k - 1}\right) \geq A$$

by replacing each finite sum in parentheses on the left by the corresponding infinite sum. Therefore we have

$$\prod_{k=1}^{\infty}\left(\frac{p_k}{p_k-1}\right)=\infty$$

or equivalently that

$$\prod_{k=1}^{\infty}\left(1-\frac{1}{p_k}\right)=0.$$

Equation (2.23) is now an immediate consequence of Theorem 60. □

•2.10 Continued Fractions

Infinite continued fractions have been around since late in the sixteenth century. We start with two sequences $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ of real (or complex) numbers and define the ***approximants***

$$f_1=\frac{a_1}{b_1},\ f_2=\frac{a_1}{b_1+\dfrac{a_2}{b_2}},\ f_3=\frac{a_1}{b_1+\dfrac{a_2}{b_2+\dfrac{a_3}{b_3}}},\ f_4=\frac{a_1}{b_1+\dfrac{a_2}{b_2+\dfrac{a_3}{b_3+\dfrac{a_4}{b_4}}}},\ldots$$

This notation is too demanding of space, so we write it more compactly as

$$f_1=\frac{a_1}{b_1},\ f_2=\frac{a_1}{b_1+}\frac{a_2}{b_2},\ f_3=\frac{a_1}{b_1+}\frac{a_2}{b_2+}\frac{a_3}{b_3},\ f_4=\frac{a_1}{b_1+}\frac{a_2}{b_2+}\frac{a_3}{b_3+}\frac{a_4}{b_4},\ldots$$

The infinite partial fraction converges to a number f if and only if $f_n \longrightarrow f$ as $n \longrightarrow \infty$ and we write

$$f=\frac{a_1}{b_1+}\frac{a_2}{b_2+}\frac{a_3}{b_3+}\frac{a_4}{b_4+}\cdots$$

In an infinite summation, it is easy to how the $n+1^{\text{st}}$ partial sum is related to the n^{th} partial sum and in a infinite product, it is easy to see how the $n+1^{\text{st}}$ partial product is related to the n^{th} partial product, but it is hard to see how f_{n+1} is related to f_n. In other words, the approximants appear not to be defined incrementally. The key is that if we expand the fractions in the obvious way

$$f_1=\frac{a_1}{b_1},\quad f_2=\frac{a_1b_2}{b_1b_2+a_2},\quad f_3=\frac{a_1a_3+a_1b_1b_2}{a_2b_3+a_3b_1+b_1b_2b_3}$$

then the numerators and denominators of these fractions can be defined incrementally. To simplify the calculation, we get rid of the b's. Define $c_0 = a_1/b_1$, $c_n = a_{n+1}/(b_n b_{n+1})$ for $n = 1, 2, 3, \ldots$ and it follows that

$$f_n = \frac{c_0}{1+}\frac{c_1}{1+}\frac{c_2}{1+\cdots+}\frac{c_{n-1}}{1} = c_0\left(\frac{1}{1+}\frac{c_1}{1+}\frac{c_2}{1+\cdots+}\frac{c_{n-1}}{1}\right).$$

We see that c_0 is just a normalizing constant, so we assume that $c_0 = 1$. Now write

$$f_{n+1} = \frac{1}{1+}\frac{c_1}{1+}\frac{c_2}{1+\cdots+}\frac{c_n}{1} = \frac{p_n(c_1,\ldots,c_n)}{q_n(c_1,\ldots,c_n)}$$

along the lines indicated above. Then, we find the recurrence relations

$$p_n(c_1,\ldots,c_n) = q_{n-1}(c_2,\ldots,c_n) \tag{2.25}$$

$$\begin{aligned} q_n(c_1,\ldots,c_n) &= c_1 p_{n-1}(c_2,\ldots,c_n) + q_{n-1}(c_2,\ldots,c_n) \\ &= c_1 q_{n-2}(c_3,\ldots,c_n) + q_{n-1}(c_2,\ldots,c_n) \end{aligned} \tag{2.26}$$

say for $n = 2, 3, \ldots$ with the starting values $q_0 = 1$ and $q_1(c) = 1 + c$.

DEFINITION *A subset X of $\{1, 2, \ldots, n\}$ is* ***neighbour free*** *if whenever $k \in X$ we have $k + 1 \notin X$.*

In other words X is neighbour free if and only if for all $k = 1, 2, \ldots, (n-1)$ we do not simultaneously have $k \in X$ and $k + 1 \in X$.

LEMMA 61 *We have* $q_n(c_1,\ldots,c_n) = \sum_{X \text{ neighbour free}} c^X$ *where* $c^X = \prod_{k\in X} c_k$.

Proof. First of all, the result is correct for $n = 0$ and $n = 1$. So, let $n \geq 2$ and suppose that the result is correct for all smaller values of n. Let X be a neighbour free subset of $\{1, 2, \ldots, n\}$. Either $1 \in X$ or $1 \notin X$. In the first case $2 \notin X$ and $Y = X \setminus \{1\}$ is a neighbour free subset of $\{3, 4, \ldots, n\}$. Conversely, every such subset Y, yields a neighbour free subset $X = Y \cup \{1\}$ containing 1. In the second case, then X is a neighbour free subset of $\{2, 3, \ldots, n\}$. So

$$\sum_{\substack{X\subseteq\{1,2,\ldots,n\}\\ X \text{ neighbour free}}} c^X = c_1 \sum_{\substack{Y\subseteq\{3,4,\ldots,n\}\\ Y \text{ neighbour free}}} c^Y + \sum_{\substack{X\subseteq\{2,3,\ldots,n\}\\ X \text{ neighbour free}}} c^X,$$

$$= c_1 q_{n-2}(c_3, \ldots, c_n) + q_{n-1}(c_2, \ldots, c_n),$$

by the strong induction hypothesis

$$= q_n(c_1, \ldots, c_n).$$

by (2.26)

■

You can use the same idea to show that the number of neighbour free subsets of $\{1, 2, \ldots, n\}$ is the Fibonacci number F_{n+2}.

LEMMA 62 *For $n \geq 2$ we have*

$$q_n(c_1, \ldots, c_n) = c_n q_{n-2}(c_1, \ldots, c_{n-2}) + q_{n-1}(c_1, \ldots, c_{n-1}). \tag{2.27}$$

Proof. This is really the same proof as for Lemma 61. The result is correct for $n = 0$ and $n = 1$. So, let $n \geq 2$. Let X be a neighbour free subset of $\{1, 2, \ldots, n\}$. Either $n \in X$ or $n \notin X$. In the first case $n - 1 \notin X$ and $Y = X \setminus \{n\}$ is a neighbour free subset of $\{1, 2, \ldots, n-2\}$. Conversely, every such subset Y, yields a neighbour free subset $X = Y \cup \{n\}$ containing n. In the second case, then X is a neighbour free subset of $\{1, 2, \ldots, n-1\}$. So

$$\begin{aligned} q_n(c_1, \ldots, c_n) &= \sum_{\substack{X \subseteq \{1,2,\ldots,n\} \\ X \text{ neighbour free}}} c^X \\ &= c_n \sum_{\substack{Y \subseteq \{1,2,\ldots,n-2\} \\ Y \text{ neighbour free}}} c^Y + \sum_{\substack{X \subseteq \{1,2,\ldots,n-1\} \\ X \text{ neighbour free}}} c^X, \\ &= c_n q_{n-2}(c_1, \ldots, c_{n-2}) + q_{n-1}(c_1, \ldots, c_{n-1}). \end{aligned}$$

■

Also, we find from (2.25)

$$p_n(c_1, \ldots, c_n) = c_n p_{n-2}(c_1, \ldots, c_{n-2}) + p_{n-1}(c_1, \ldots, c_{n-1}) \tag{2.28}$$

for $n \geq 2$ with starting values $p_0 = 1$ and $p_1 = 1$. The equations (2.27) and (2.28) give the incremental definition of the fraction $\dfrac{p_n}{q_n}$. Here, we are using the symbol p_n without arguments to stand for $p_n(c_1, c_2, \ldots, c_n)$ and the same for q_n. Equation

(2.27) is the famous three term recurrence relation which has connections to many other branches of analysis.

With a little more work, we can actually obtain the approximants as the partial sums of an infinite series

$$\begin{aligned} f_n - f_{n-1} &= \frac{p_{n-1}}{q_{n-1}} - \frac{p_{n-2}}{q_{n-2}} = \frac{p_{n-1}q_{n-2} - p_{n-2}q_{n-1}}{q_{n-1}q_{n-2}} \\ &= \frac{(c_{n-1}p_{n-3} + p_{n-2})q_{n-2} - p_{n-2}(c_{n-1}q_{n-3} + q_{n-2})}{q_{n-1}q_{n-2}} \\ &= -\frac{c_{n-1}}{q_{n-1}q_{n-2}}(p_{n-2}q_{n-3} - p_{n-3}q_{n-2}) \\ &= -\frac{c_{n-1}q_{n-3}}{q_{n-1}}(f_{n-1} - f_{n-2}) \end{aligned}$$

It now follows that f_n is the n^{th} partial sum of the series

$$1 - \frac{c_1}{q_1} + \frac{c_1c_2}{q_1q_2} - \frac{c_1c_2c_3}{q_2q_3} + \frac{c_1c_2c_3c_4}{q_3q_4} - \cdots \tag{2.29}$$

EXAMPLE If $c_n > 0$ for all $n = 1, 2, \cdots$, then the signs of this series are alternating and the absolute values of the terms are decreasing. We have

$$\frac{c_1c_2\cdots c_n}{q_{n-1}q_n} < \frac{c_1c_2\cdots c_{n-1}}{q_{n-2}q_{n-1}}$$

since $c_nq_{n-2} < q_n$ which in turn follows since $q_n - c_nq_{n-2} = q_{n-1} > 0$. Thus, by the Alternating Series Test (Theorem 52), the partial fraction converges if and only if $\lim_{n\to\infty} \frac{c_1c_2\cdots c_n}{q_{n-1}q_n} = 0$. Since

$$q_n \geq \sum_{X\subseteq\{k;1\leq k\leq n,\ n-k \text{ even}\}} c^X = \prod_{\substack{1\leq k\leq n\\ n-k \text{ even}}} (1 + c_k),$$

we find

$$q_{n-1}q_n \geq \prod_{k=1}^{n}(1 + c_k)$$

from which we see that a sufficient condition for convergence is $\prod_{k=1}^{\infty} \frac{c_k}{1+c_k} = 0$.

By Theorem 60 this is equivalent to $\sum_{k=1}^{\infty} \frac{1}{1+c_k} = \infty$. □

EXAMPLE A very basic example is

$$f = \frac{a}{2b+}\frac{a}{2b+}\frac{a}{2b+}\cdots = 2b\left(\frac{\rho}{1+}\frac{\rho}{1+}\frac{\rho}{1+}\cdots\right) \tag{2.30}$$

where $\rho = \dfrac{a}{4b^2}$. If $a > 0$, then the previous example shows that the series converges and it is easy to see that $f = \dfrac{a}{2b+f}$, so that $f^2 + 2bf - a = 0$ and $f = -b \pm \sqrt{b^2+a}$. By (2.29) we see that $f > 0$ in this case and we can deduce that in fact $f = -b + \sqrt{b^2+a}$. On the other hand, if $b^2 + a < 0$, then convergence cannot be possible. What happens when $-b^2 \le a \le 0$? The approximants can be obtained by iteration. For the continued fraction in the brackets on the right of (2.30) we have $f_{n+1} = \rho/(1+f_n)$ with $f_0 = \rho$. It's easy to see that f_n converges to $-\frac{1}{2} + \sqrt{\rho + \frac{1}{4}}$ in case $-\frac{1}{4} \le \rho \le 0$ using the facts that

- f_n decreases with n.
- $f_n > -\frac{1}{2} + \sqrt{\rho + \frac{1}{4}}$ for all n.

□

EXAMPLE Another example is the case $c_n = n^2$. Leading to

$$f = \frac{1}{1+}\frac{1^2}{1+}\frac{2^2}{1+}\frac{3^2}{1+}\frac{4^2}{1+}\cdots$$

Then it is easy to see that the solution of the recurrence relation $q_n = c_n q_{n-2} + q_{n-1}$, $q_0 = 1$, $q_1 = 2$ is given by $q_n = (n+1)!$. Then (2.29) becomes

$$1 - \frac{(1!)^2}{2!} + \frac{(2!)^2}{2!\,3!} - \frac{(3!)^2}{3!\,4!} + \cdots = \sum_{k=1}^{\infty}(-1)^{k-1}\frac{1}{k}$$

which as we shall see later converges to $\ln(2)$. □

There are in fact many other interesting examples and indeed, many of the transcendental functions have neat continued fraction expansions. This material is beyond the scope of this course.

3

The Riemann Integral

In this chapter we develop the theory of Riemann Integration. In the form that we present it here this is a quick, unpolished theory which gets the job done for moderately nice functions. However, there is also a theory of integration due to Lebesgue which is more general and more powerful. As we all know from our calculus courses the integral, in its simplest form, is an attempt to measure the area bounded below by the x-axis, above by the graph of a function f and on the left and right by x-ordinates. In the Riemann theory, this area is cut up vertically so that then vertical partition corresponds to a collection of intervals in the x-axis. Since an interval has a well-defined length this poses few problems. In the Lebesgue theory, the area is cut up horizontally, the partition corresponding to a collection of intervals in the y axis. If J is one such interval, the corresponding subset of the x-axis that has to be measured is the inverse image $f^{-1}(J)$. This set is no longer necessarily an interval, in fact it can be quite complicated and we need to determine its length. The problem of achieving this in a systematic way is called measure theory and it is a necessary prerequisite to the Lebesgue Integral. Beyond this lies abstract measure theory and abstract integration. Within this more abstract framework lies the theory of probability where the so called events are subsets of the sample space which are assigned a probability which one thinks of as a kind of measure. To develop all of these ideas takes a couple of courses ...

So we come back to earth. Here we look only at the 1-dimensional Riemann theory. We shall then be attempting to integrate a bounded function f over a bounded interval $[a, b]$.

3.1 Partitions

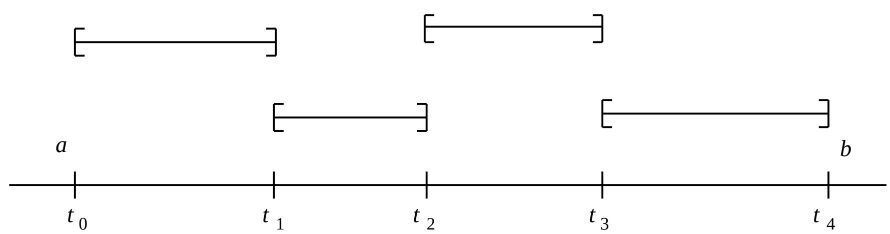

Figure 3.1: A Riemann Partition and its intervals.

DEFINITION *A **Riemann partition** P of the interval $[a, b]$ is specified by real numbers $(t_n)_{n=0}^N$ such that*

$$a = t_0 < t_1 < t_2 < \cdots < t_{N-1} < t_N = b.$$

The intervals of the partition are $[t_{n-1}, t_n]$ for $n = 1, 2, \ldots, N$. Strictly speaking, the partition P is the collection of intervals $[t_{n-1}, t_n]$ which cover $[a, b]$ and overlap only at their endpoints.

DEFINITION *Given two Riemann partitions $P = (t_n)_{n=0}^N$ and $Q = (s_k)_{k=0}^K$, we say that Q is a **refinement** of P if $\{t_n; 0 \leq n \leq N\} \subseteq \{s_k; 0 \leq k \leq K\}$. In terms of intervals, this means that each interval of P, can be decomposed as a finite union of the intervals of Q overlapping only at their endpoints.*

DEFINITION *A **tagged partition** $\mathcal{P}$ of the interval $[a, b]$ is a Riemann Partition P together with a choice of **partition points** $(\xi_n)_{n=1}^N$ with $\xi_n \in [t_{n-1}, t_n]$ for $n = 1, 2, \ldots, N$.*

DEFINITION *For every tagged partition $\mathcal{P}$, we can define the **Riemann Sum***

$$S(\mathcal{P}, f) = \sum_{n=1}^{N} f(\xi_n)(t_n - t_{n-1})$$

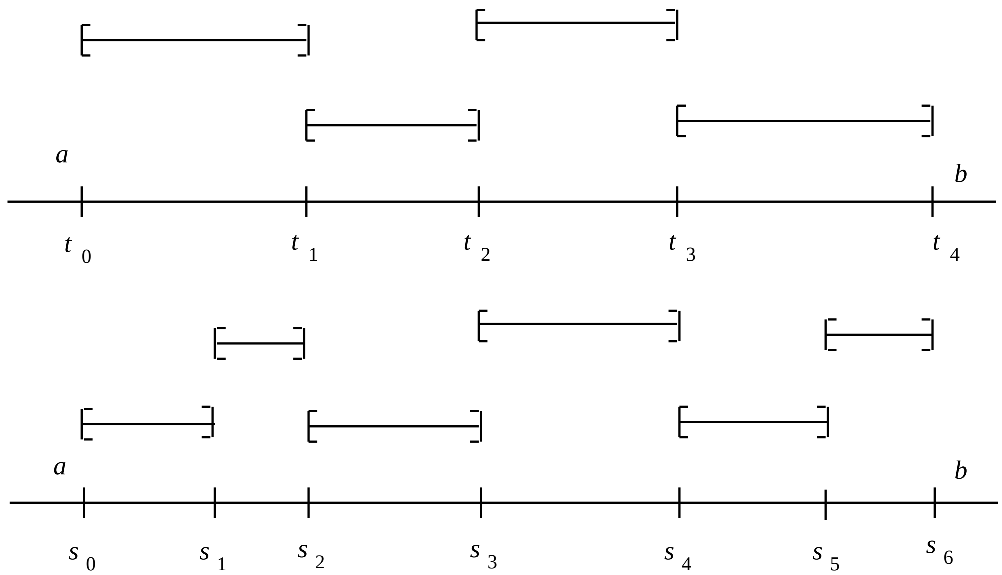

Figure 3.2: Top: A Riemann Partition and its intervals, Bottom: A refining partition and its intervals.

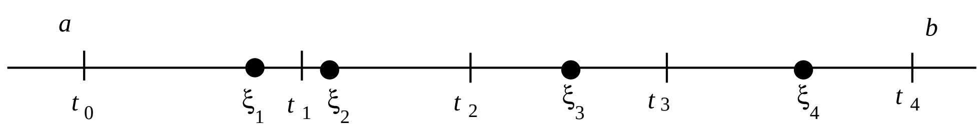

Figure 3.3: A Tagged Partition.

We can write this more succinctly as

$$S(\mathcal{P}, f) = \sum_{n=1}^{N} f(\xi_n)|J_n|$$

where $J_n = [t_{n-1}, t_n]$ *is a typical interval of the partition* P *and* $|J_n|$ *denotes its length.*

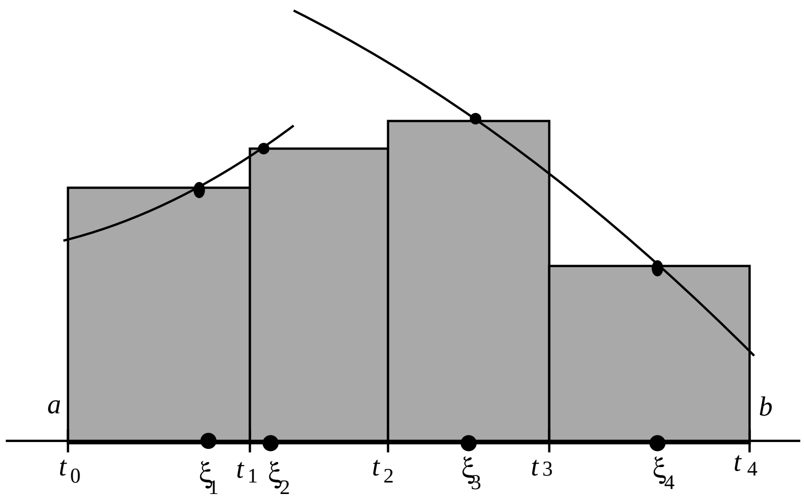

Figure 3.4: Area representing a Riemann sum.

3.2 Upper and Lower Sums and Integrals

There are two approaches to the Riemann integral. Perhaps the easiest to understand uses upper and lower sums. Let P be a Riemann partition. Then the upper sum $U(P, f)$ is the supremum of all Riemann sums $S(\mathcal{P}, f)$ as $\mathcal{P}$ runs over all tagged partitions based on P. Notice that since we are assuming that f is a bounded function, we always have

$$S(\mathcal{P}, f) \leq (b-a) \sup_{[a,b]} f$$

so the supremum is defined. Since the ξ_n are independent of one another we also can write

$$U(P, f) = \sum_{n=1}^{N} (t_n - t_{n-1}) \sup_{[t_{n-1}, t_n]} f = \sum_J |J| \sup_J f$$

The lower sum $L(P, f)$ is defined similarly as the infimum of all Riemann sums $S(\mathcal{P}, f)$ as $\mathcal{P}$ runs over all tagged partitions based on P and we find

$$L(P, f) = \sum_{n=1}^{N}(t_n - t_{n-1}) \inf_{[t_{n-1},t_n]} f = \sum_J |J| \inf_J f.$$

Obviously, we have $L(P, f) \leq U(P, f)$.

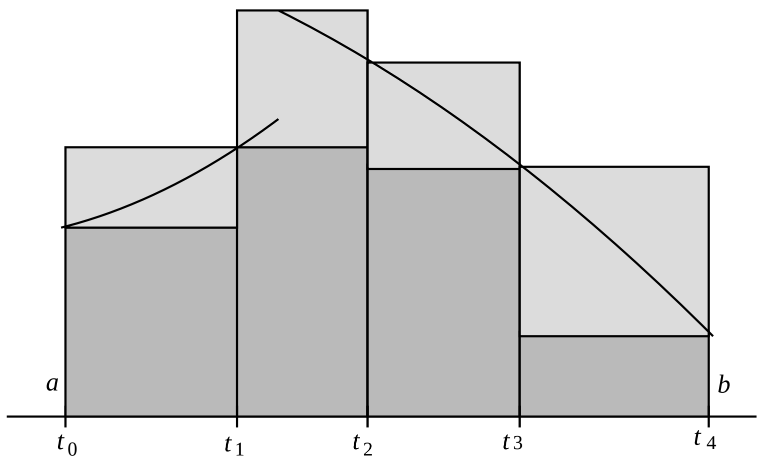

Figure 3.5: The upper sum corresponds to the area shaded in gray, the lower sum to the area shaded in the darker gray.

THEOREM 63 *If Q refines P then $U(Q, f) \leq U(P, f)$ and $L(Q, f) \geq L(P, f)$.*

Proof. Let us work with points. Let $P = (t_n)_{n=0}^N$ and $Q = (s_k)_{k=0}^K$. Since Q refines P every $t_n = s_{k(n)}$ for a suitable $k(n)$. Clearly $0 = k(0) < k(1) < \cdots < k(N) = K$.

$$\begin{aligned} U(Q, f) &= \sum_{k=1}^{K}(s_k - s_{k-1}) \sup_{[s_{k-1},s_k]} f \\ &= \sum_{n=1}^{N} \sum_{k=k(n-1)+1}^{k(n)} (s_k - s_{k-1}) \sup_{[s_{k-1},s_k]} f \end{aligned}$$

$$
\begin{aligned}
&\le \sum_{n=1}^{N} \left\{ \sup_{[t_{n-1},t_n]} f \right\} \sum_{k=k(n-1)+1}^{k(n)} (s_k - s_{k-1}) \\
&= \sum_{n=1}^{N} \left\{ \sup_{[t_{n-1},t_n]} f \right\} (t_n - t_{n-1}) \\
&= U(P, f)
\end{aligned}
$$

since for $k(n-1) < k \le k(n)$, we have that $\sup_{[s_{k-1},s_k]} f \le \sup_{[t_{n-1},t_n]} f$ since $[s_{k-1}, s_k] \subseteq [t_{n-1}, t_n]$. The argument for the lower sums is similar. ■

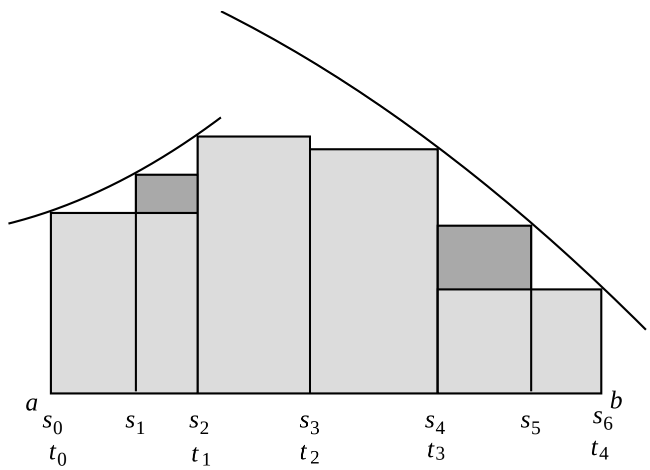

Figure 3.6: Areas representing $L(P, f)$ and $L(Q, f)$. The area corresponding to $L(Q, f) - L(P, f)$ is shown in the darker shade of gray.

DEFINITION *We now set*

$$
\overline{\int_a^b} f(x)dx = \inf_P U(P, f) \qquad \textit{and} \qquad \underline{\int_a^b} f(x)dx = \sup_P L(P, f)
$$

the sup *and* inf *being taken over all Riemann partitions P of $[a, b]$. These expressions are called the* ***upper integral*** *and* ***lower integral*** *respectively. They are well-defined for all bounded functions on $[a, b]$.*

Of course, we would like to know that these definitions imply that

$$\underline{\int}_a^b f(x)dx \leq \overline{\int}_a^b f(x)dx$$

To see this, suppose not. Then $\underline{\int}_a^b f(x)dx > \overline{\int}_a^b f(x)dx$ and we can find Riemann partitions Q and R such that $L(Q, f) > U(R, f)$. Now let P be a partition that refines both Q and R. To make P it suffices to take the union of the endpoint sets as the endpoint set of P. We have $L(P, f) \geq L(Q, f) > U(R, f) \geq U(P, f) \geq L(P, f)$, a contradiction.

DEFINITION *We say that the function f is* ***Riemann integrable*** *over the interval $[a, b]$ iff*

$$\overline{\int}_a^b f(x)dx = \underline{\int}_a^b f(x)dx.$$

In this case the value of the integral is the common value and it is denoted $\int_a^b f(x)dx$.

Note that the above definition is for the case $a < b$. We may also define

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

and indeed

$$\int_a^a f(x)dx = 0.$$

EXAMPLE Consider the function

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q}, \\ 0 & \text{otherwise.} \end{cases}$$

Now each interval J of strictly positive length contains both rational and irrational numbers. So we will find that $\sup_J f = 1$ and $\inf_J f = 0$. It follows that for any Riemann partition of $[0, 1]$ we have $U(P, f) = 1$ and $L(P, f) = 0$. So we get

$$\overline{\int}_0^1 f(x)dx = 1 > 0 \underline{\int}_0^1 f(x)dx$$

and f is not Riemann integrable.

This is in stark contrast to the Lebesgue theory which would determine that the set $\mathbb{Q} \bigcap [0,1]$ has zero length (because it can be covered by a countable union of open intervals of total length as small as we please) and decide that the value of the integral should be 0. □

THEOREM 64 *The following condition is equivalent to the Riemann integrability of f on $[a,b]$. For all $\epsilon > 0$ there exists a Riemann partition P such that $U(P,f) - L(P,f) < \epsilon$.*

Proof. First suppose that f is Riemann integrable. Then there exists a Riemann partition Q such that $U(Q,f) < \int_a^b f(x)dx + \frac{\epsilon}{2}$. This is because $\int_a^b f(x)dx + \frac{\epsilon}{2}$ is not a lower bound for $\overline{\int}_a^b f(x)dx = \inf_Q U(Q,f)$. In the same way, we have a Riemann partition R such that $L(R,f) > \int_a^b f(x)dx - \frac{\epsilon}{2}$. Now let P be a partition that refines both Q and R. Then we get

$$\int_a^b f(x)dx - \frac{\epsilon}{2} < L(R,f) \le L(P,f) \le U(P,f) \le U(Q,f) < \int_a^b f(x)dx + \frac{\epsilon}{2}.$$

and it follows that $U(P,f) - L(P,f) < \epsilon$.

Now for the opposite direction, we have simply that

$$L(P,f) \le \underline{\int}_a^b f(x)dx \le \overline{\int}_a^b f(x)dx \le U(P,f) < L(P,f) + \epsilon.$$

Therefore,

$$0 \le \overline{\int}_a^b f(x)dx - \underline{\int}_a^b f(x)dx < \epsilon$$

and since ϵ can be as small as we please, we must have $\overline{\int}_a^b f(x)dx = \underline{\int}_a^b f(x)dx$.

The condition in Theorem 64 is called ***Riemann's condition*** and it is useful to express it in a different way. Note that

$$U(P,f) - L(P,f) = \sum_J |J| \sup_J f - \sum_J |J| \inf_J f = \sum_J |J| \operatorname{osc}_J f$$

where $\operatorname{osc}_J f = \sup_J f - \inf_J f$, the ***oscillation*** of f on the interval J. We can also characterize $\operatorname{osc}_J f = \sup_{x,x' \in J} |f(x) - f(x')|$. The alternate form of Riemann's condition is that for every $\epsilon > 0$ there exists a Riemann partition P such that

$$\sum_J |J| \operatorname*{osc}_J f < \epsilon \tag{3.1}$$

where the sum is taken over the intervals of P.

3.3 Conditions for Riemann Integrability

There are two big theorems here.

THEOREM 65 *Let $f : [a, b] \longrightarrow \mathbb{R}$ be continuous. Then f is Riemann integrable on $[a, b]$.*

Proof. If J is an interval of length at most $\delta > 0$ then $\operatorname{osc}_J f \leq \omega_f(\delta)$ using the modulus of continuity notation ω_f. So if P is a Riemann partition in which the intervals have length at most δ we get

$$\sum_J |J| \operatorname*{osc}_J f \leq \sum_J |J| \omega_f(\delta) = (b - a)\omega_f(\delta).$$

Since f is continuous on the sequentially compact set $[a, b]$, it is also uniformly continuous on this set and therefore, given $\epsilon > 0$ we can find $\delta > 0$ such that $\omega_f(\delta) < \dfrac{\epsilon}{b - a}$. It is then easy to construct a suitable P and it follows that

$$\sum_J |J| \operatorname*{osc}_J f < \epsilon.$$

■

THEOREM 66 *Let $f : [a, b] \longrightarrow \mathbb{R}$ be monotone and bounded. Then f is Riemann integrable on $[a, b]$.*

Proof. Let's suppose without loss of generality that f is increasing. Then we have $\operatorname{osc}_J f = f(\beta_J) - f(\alpha_J)$ where $J = [\alpha_J, \beta_J]$. Let us suppose again that P is a Riemann partition in which the intervals have length at most δ. Then

$$\sum_J |J| \operatorname*{osc}_J f \leq \sum_J \delta \operatorname*{osc}_J f = \delta \sum_J \Big(f(\beta_J) - f(\alpha_J) \Big) = \delta(f(b) - f(a))$$

because the sum $\sum_J \Big(f(\beta_J) - f(\alpha_J)\Big)$ telescopes. We need only choose $\delta = \dfrac{\epsilon}{1 + f(b) - f(a)}$ in order to satisfy Riemann's condition. ■

3.4 Properties of the Riemann Integral

There are a number of fairly routine properties that need to be verified.

THEOREM 67 *If f and g are Riemann integrable on the interval $[a, b]$, then so is the linear combination $tf + sg$ for $t, s \in \mathbb{R}$. Furthermore*

$$\int_a^b (tf + sg)(x)dx = t\int_a^b f(x)dx + s\int_a^b g(x)dx.$$

Proof. We divide this result up into two separate parts. First scalar multiples. If f is Riemann integrable, so is tf. This is more or less obvious, but some care is needed because the case $t < 0$ flips the upper and lower sums and integrals. We leave this and the identity

$$\int_a^b (tf)(x)dx = t\int_a^b f(x)dx$$

as an exercise for the reader.

It remains to deal with the sum. Obviously, for a tagged partition we have

$$S(\mathcal{P}, f + g) = S(\mathcal{P}, f) + S(\mathcal{P}, g).$$

When we take the supremum (over all tagged partitions $\mathcal{P}$ based on P), we get

$$U(P, f + g) \leq U(P, f) + U(P, g).$$

We cannot dispense with the inequality in this case[1]. Similarly we get

$$L(P, f + g) \geq L(P, f) + L(P, g).$$

Now, let $\epsilon > 0$. Then there are partitions Q and R such that $U(Q, f) - L(Q, f) < \epsilon$ and $U(R, g) - L(R, g) < \epsilon$. Let P be a partition that refines both Q and R.

[1] To see this, take $a = 0$, $b = 1$, P the partition with just one interval, $f(x) = x$, $g(x) = 1 - x$. Then $U(P, f + g) = U(P, f) = U(P, g) = 1$.

Then we get $U(P,f) - L(P,f) < \epsilon$ and $U(P,g) - L(P,g) < \epsilon$. So we have a sandwich

$$\begin{aligned} L(P,f) + L(P,g) &\le \int_a^b f(x)dx + \int_a^b g(x)dx \\ &\le U(P,f) + U(P,g) \\ &\le L(P,f) + \epsilon + L(P,g) + \epsilon \\ &= L(P,f) + L(P,g) + 2\epsilon \end{aligned}$$

and at the same time another sandwich

$$\begin{aligned} L(P,f) + L(P,g) &\le L(P, f+g) \\ &\le \underline{\int}_a^b \Big(f(x) + g(x)\Big)dx \\ &\le \overline{\int}_a^b \Big(f(x) + g(x)\Big)dx \\ &\le U(P, f+g) \\ &\le U(P,f) + U(P,g) \\ &\le L(P,f) + L(P,g) + 2\epsilon \end{aligned}$$

So, the quantities $\underline{\int}_a^b \Big(f(x)+g(x)\Big)dx$ and $\overline{\int}_a^b \Big(f(x)+g(x)\Big)dx$ lie in an interval of length 2ϵ and ϵ is an arbitrary positive number. So the two quantities must be equal. This shows that $f+g$ is Riemann Integrable. But then $\int_a^b \Big(f(x)+g(x)\Big)dx$ and $\int_a^b f(x)dx + \int_a^b g(x)dx$ also lie within the same interval of length 2ϵ and we conclude that

$$\int_a^b \Big(f(x) + g(x)\Big)dx = \int_a^b f(x)dx + \int_a^b g(x)dx,$$

as required. ■

THEOREM 68 *If f is Riemann integrable on $[a,b]$ and $f(x) \ge 0$ for $a \le x \le b$, then $\int_a^b f(x)dx \ge 0$.*

Proof. Since $S(\mathcal{P}, f) \ge 0$ we have $L(P,f) \ge 0$ for all P and the result follows.■

COROLLARY 69 *If f and g are Riemann integrable on $[a, b]$ and if $f(x) \geq g(x)$ for $a \leq x \leq b$, then $\int_a^b f(x)dx \geq \int_a^b g(x)dx$.*

THEOREM 70 *Let $a < b < c$. Suppose that f is Riemann integrable on $[a, b]$ and on $[b, c]$, then it is integrable on $[a, c]$ and*

$$\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx.$$

Proof. Let $\epsilon > 0$. Let Q be a Riemann partition of $[a, b]$ and R a Riemann partition of $[b, c]$ such that $U(Q, f) - L(Q, f) < \epsilon$ and $U(R, f) - L(R, f) < \epsilon$. We form the join P of these two partitions. The intervals of P are the intervals of Q together with the intervals of R. In fact, we get $U(P, f) = U(Q, f) + U(R, f)$ and $L(P, f) = L(Q, f) + L(R, f)$. We build two sandwiches

$$\begin{array}{ccccccc} L(P,f) & \leq & \underline{\int}_a^c f(x)dx & \leq & \overline{\int}_a^c f(x)dx & \leq & U(P,f) \\ \| & & & & & & \| \\ L(Q,f)+L(R,f) & \leq & & \int_a^b f(x)dx + \int_b^c f(x)dx & & \leq & U(Q,f)+U(R,f) \end{array}$$

Since $U(P, f) \leq L(P, f) + 2\epsilon$ we find that $\underline{\int}_a^c f(x)dx \leq \overline{\int}_a^c f(x)dx$ showing that f is Riemann integrable on $[a, c]$. Furthermore we also get $\int_a^c f(x)dx = \int_a^b f(x)dx + \int_b^c f(x)dx$. ■

THEOREM 71 *If a function f is Riemann integrable on a closed bounded interval I then it is also integrable on every closed subinterval J of I.*

The proof is left to the reader.

3.5 Another Approach to the Riemann Integral

The upper and lower sum approach to defining the Riemann integral has some disadvantages. Perhaps the most serious of these is that it does not apply to vector valued functions. You cannot take infs and sups of vectors. Here is an equivalent way of defining the integral. It uses the partial ordering of the set of all Riemann partitions.

THEOREM 72 *Let $f : [a,b] \longrightarrow \mathbb{R}$ be a bounded function. The following condition is equivalent to the existence of the Riemann integral $\int_a^b f(x)dx$ and its equality with the real number s.*

For every positive number ϵ, there exists a Riemann partition P of $[a,b]$ such that for every Riemann partition Q refining P and every tagged partition $\mathcal{Q}$ based on Q we have

$$|s - S(\mathcal{Q}, f)| < \epsilon. \tag{3.2}$$

Proof. First suppose that the Riemann integral $\int_a^b f(x)dx$ exists and equals s. Let $\epsilon > 0$. Then, by Riemann's criterion, there is a Riemann partition P such that

$$L(P, f) \le s \le U(P, f) < L(P, f) + \epsilon.$$

On the other hand, if Q is a Riemann partition refining P and $\mathcal{Q}$ is a tagged partition based on Q we have

$$L(P, f) \le L(Q, f) \le S(\mathcal{Q}, f) \le U(Q, f) \le U(P, f).$$

Thus, both s and $S(\mathcal{Q}, f)$ lie in the interval $[L(P, f), L(P, f) + \epsilon[$, and the required conclusion follows.

In the opposite direction, suppose that (3.2) holds. Then we have, taking just $Q = P$ that

$$U(P, f) = \sup S(\mathcal{P}, f) \le s + \epsilon.$$

Rewriting this and combining with the similar statement for the lower sum, we get

$$U(P, f) - \epsilon \le s \le L(P, f) + \epsilon$$

which forces $U(P, f) - L(P, f) \le 2\epsilon$. So Riemann's condition is satisfied. This shows that the integral exists, but we still need its equality with s. Towards this, we have

$$L(P, f) \le \int_a^b f(x)dx \le U(P, f)$$

which yields

$$\left| s - \int_a^b f(x)dx \right| \le \epsilon.$$

Since this holds for all positive ϵ we have our result. ■

So, in (3.2) it appears that the refinement procedure is unnecessary. It would be just as good to write:

For every positive number ϵ, there exists a Riemann partition P of $[a, b]$ such that for every tagged partition $\mathcal{P}$ based on P we have

$$|s - S(\mathcal{P}, f)| < \epsilon. \tag{3.3}$$

So why don't we? Well if you are going to use (3.3) as the definition of the integral, it is far from being immediately clear that s is unique. Taking (3.2) as the definition, the uniqueness is almost immediate. If s and s' are both possible integrals, then according to (3.2) there are partitions P and P' such that

$$|s - S(\mathcal{Q}, f)| < \epsilon \text{ for all } \mathcal{Q} \text{ based on a refinement of } P$$

and

$$|s' - S(\mathcal{Q}, f)| < \epsilon \text{ for all } \mathcal{Q} \text{ based on a refinement of } P'$$

Choose Q to be a common refinement of P and P' and $\mathcal{Q}$ any tagged partition based on Q and we deduce that $|s - s'| < 2\epsilon$. Since ϵ is an arbitrary positive number, we must have $s = s'$.

If we wish to deal with Riemann integrals of functions taking values in a complete normed vector space, then upper and lower sums cannot be used. This means that we have to take (3.2) as the *definition* of the Riemann integral and the entire theory has to reworked within this framework. We shall not do this here.

•3.6 Lebesgue's Theorem and other Thorny Issues

There is a theorem due to Lebesgue that completely characterizes which functions are Riemann integrable. We will need the following definition.

DEFINITION *A subset N of $\mathbb{R}$ has* ***zero length*** *if for every positive number ϵ, there exists a countable collection of open intervals $(J_j)_{j=1}^{\infty}$ such that*

$$N \subseteq \bigcup_{j=1}^{\infty} J_j \qquad \text{and} \qquad \sum_{j=1}^{\infty} |J_j| < \epsilon.$$

THEOREM 73 *Let $[a, b]$ be a closed bounded interval and let $f : [a, b] \longrightarrow \mathbb{R}$ be a bounded function. Then f is Riemann integrable on $[a, b]$ if and only if the set of points where f fails to be continuous has zero length.*

It is important to read the condition in Theorem 73 carefully. It is completely different for example to say that there is a subset N of zero length such that the restriction of f to $[a, b] \setminus N$ is continuous. For example let $f = \mathbb{1}_{\mathbb{Q}}$, the indicator function of the set of rational numbers. Then f is discontinuous everywhere. Nevertheless, the restriction of the function to the set of irrational numbers (whose complement, namely $\mathbb{Q}$ is a set of zero length) is identically zero and hence continuous.

We do not have the tools to prove Theorem 73 in this course. However, we do have the tools to prove one of its corollaries.

COROLLARY 74 *Let $[a, b]$ be a closed bounded interval and let $f : [a, b] \longrightarrow [c, d]$ be a bounded function which is Riemann integrable on $[a, b]$. Let $\varphi : [c, d] \longrightarrow \mathbb{R}$ be continuous. Then $\varphi \circ f$ is Riemann integrable on $[a, b]$.*

Proof. Since φ is continuous on a closed bounded interval, it is uniformly continuous. Also, it is bounded, so let $\varphi([c, d])$ be contained in an interval of length $L > 0$. Let $\epsilon > 0$. Then, by the uniform continuity of φ there exists $\delta > 0$ such that $|x - x'| \leq \delta$ implies that $|\varphi(x) - \varphi(x')| < \dfrac{\epsilon}{2(b-a)}$. If J is a subinterval of $[a, b]$ such that $\operatorname{osc}_f(J) \leq \delta$, then we have $\operatorname{osc}_{\varphi\circ f}(J) \leq \dfrac{\epsilon}{2(b-a)}$. Since f is Riemann integrable on $[a, b]$, there is a Riemann partition P such that

$$\sum_J |J| \operatorname*{osc}_f(J) < \frac{\epsilon\delta}{2L} \tag{3.4}$$

where the sum is taken over the intervals of P. These intervals are divided into two types, the red intervals and the black intervals. A red interval is one such that $\operatorname{osc}_f(J) \leq \delta$ and the black intervals satisfy $\operatorname{osc}_f(J) > \delta$. It follows from (3.4) that the total length of the black intervals is less than $\dfrac{\epsilon}{2L}$, but we can say very little about $\operatorname{osc}_{\varphi\circ f}(J)$, only that it is bounded by L. On the other hand, for each red interval we have $\operatorname{osc}_{\varphi\circ f}(J) \leq \dfrac{\epsilon}{2(b-a)}$, but we can say very little about the total length of the red intervals, in fact, only that their total length is less than $b-a$. So we find

$$\begin{aligned}\sum_{J}|J|\operatorname*{osc}_{\varphi\circ f}(J) &\le \sum_{J\text{ red}}|J|\operatorname*{osc}_{\varphi\circ f}(J)+\sum_{J\text{ black}}|J|\operatorname*{osc}_{\varphi\circ f}(J)\\ &\le \sum_{J\text{ red}}|J|\frac{\epsilon}{2(b-a)}+\sum_{J\text{ black}}|J|L\\ &\le (b-a)\frac{\epsilon}{2(b-a)}+\frac{\epsilon}{2L}L=\epsilon\end{aligned}$$

Since ϵ is an arbitrary positive number, we see that $\varphi\circ f$ satisfies Riemann's condition and is therefore Riemann integrable on $[a,b]$. ■

COROLLARY 75 *Let $[a,b]$ be a closed bounded interval and let $f,g:[a,b]\longrightarrow[c,d]$ be bounded functions, Riemann integrable on $[a,b]$. Then so is $f\cdot g$.*

Proof. We have

$$f\cdot g=\frac{1}{4}\left((f+g)^2-(f-g)^2\right)$$

and can apply the previous corollary with $\varphi(x)=x^2$. ■

Another very important inequality is contained in the following corollary.

COROLLARY 76 *Let $[a,b]$ be a closed bounded interval and let $f:[a,b]\longrightarrow\mathbb{R}$ be a bounded function, Riemann integrable on $[a,b]$. Then so is $|f|$ and*

$$\left|\int_a^b f(x)dx\right|\le\int_a^b|f(x)|dx$$

Proof. The composition theorem yields that $|f|$ is Riemann integrable on $[a,b]$. (In fact this is easy since $\operatorname{osc}_{|f|}(J)\le\operatorname{osc}_f(J)$ for any interval J). Then using the fact that $|f|\mp f$ is a nonnegative function, we get $\displaystyle\int_a^b\Big(|f(x)|\mp f(x)\Big)dx\ge 0$ and hence that

$$\pm\int_a^b f(x)dx\le\int_a^b|f(x)|dx$$

which is equivalent to the desired conclusion. ■

Traditionally, the Riemann integral is defined by convergence through the partially ordered set of all Riemann partitions. It is interesting to ask if it could also be done through the step of the partition. The answer is yes, but the proof of this fact is not quite trivial.

DEFINITION *Let P be a Riemann partition. The* ***step*** *of P is the length of the longest interval of P.*

THEOREM 77 *Let $[a, b]$ be a closed bounded interval and let $f : [a, b] \longrightarrow \mathbb{R}$ be a bounded function which is Riemann integrable on $[a, b]$. Then, given $\epsilon > 0$ there exists $\delta > 0$ such that for every Riemann partition P of step less than δ and every tagged partition $\mathcal{P}$ based on P we have*

$$\left| \int_a^b f(x)dx - S(\mathcal{P}, f) \right| < \epsilon$$

Proof. We start out by showing that if I_1, I_2 and J are three intervals with $J \subseteq I_1 \cup I_2$, then

$$\operatorname*{osc}_J f \leq \operatorname*{osc}_{I_1} f + \operatorname*{osc}_{I_2} f.$$

The cases where $J \subseteq I_1$ or $J \subseteq I_2$ are trivial, so we can assume that J meets both I_1 and I_2 and hence $I_1 \cap I_2 \neq \emptyset$. Let $x_3 \in I_1 \cap I_2$. Let $x_1, x_2 \in J$. We want to show

$$|f(x_1) - f(x_2)| \leq \operatorname*{osc}_{I_1} f + \operatorname*{osc}_{I_2} f. \tag{3.5}$$

If $x_1, x_2 \in I_1$ this is trivial, and similarly if both points are in I_2. So, we can assume without loss of generality that $x_1 \in I_1$ and $x_2 \in I_2$. Now we have

$$|f(x_1) - f(x_2)| \leq |f(x_1) - f(x_3)| + |f(x_3) - f(x_2)| \leq \operatorname*{osc}_{I_1} f + \operatorname*{osc}_{I_2} f,$$

as required. Taking the supremum in (3.5) over x_1 and x_2 establishes the claim.

Now, to the main issue. Let $\epsilon > 0$ and find a Riemann partition Q with intervals $I_1, \ldots, I_K$ such that

$$\sum_{k=1}^{K} |I_k| \operatorname*{osc}_{I_k} f < \frac{1}{4}\epsilon.$$

Let $\delta = \min_{k=1}^{K} |I_k| > 0$. Now, let P be another Riemann partition having all of its intervals $J_1, \ldots, J_L$ of length less than δ. Then, of course the intervals J_ℓ may be contained in a single I_k or in the union $I_k \cup I_{k+1}$ of two consecutive intervals of $\mathcal{Q}$, but they cannot extend over more than two of these intervals. Thus, using the claim, we have

$$\sum_{\ell=1}^{L} |J_\ell| \operatorname*{osc}_{J_\ell} f \leq \sum_{\ell=1}^{L} |J_\ell| \sum_{I_k \cap J_\ell \neq \emptyset} \operatorname*{osc}_{I_k} f$$

the inner sum being over those k such that $I_k \cap J_\ell \neq \emptyset$

$$= \sum_{k=1}^{K} \operatorname*{osc}_{I_k} f \sum_{I_k \cap J_\ell \neq \emptyset} |J_\ell|$$

the inner sum now being over those ℓ such that $I_k \cap J_\ell \neq \emptyset$

$$\leq \sum_{k=1}^{K} 3|I_k| \operatorname*{osc}_{I_k} f < \epsilon$$

because the total length of the intervals J_ℓ meeting a fixed I_k cannot exceed 3 times the length of I_k. To see this point, observe that the J_ℓ are disjoint and contained in $I_k^\star$, the interval with the same centre as I_k but three times the length. Hence the Riemann condition is satisfied.

From this, we need to get to the statement regarding tagged partitions. Let $\mathcal{P}$ be a tagged partition based on P. Then we have $U(P, f) - L(P, f) < \epsilon$ and also the inequality chains

$$L(P, f) \leq \int_a^b f(x)dx \leq U(P, f),$$
$$L(P, f) \leq S(\mathcal{P}, f) \leq U(P, f).$$

When combined, these inequalities yield the desired conclusion. ■

One of the consequences of Theorem 77 is the following corollary.

COROLLARY 78 *Let $[a, b]$ be a closed bounded interval and let $f : [a, b] \longrightarrow \mathbb{R}$ be a bounded function which is Riemann integrable on $[a, b]$. Then*

$$\lim_{N\to\infty} \frac{b-a}{N} \sum_{n=1}^{N} f\left(a + \frac{n}{N}(b-a)\right) = \int_a^b f(x)dx.$$

We can use this Corollary to evaluate certain limits.

EXAMPLE

$$\lim_{n\to\infty}\sum_{k=1}^{n}\frac{n}{n^2+k^2}=\lim_{n\to\infty}\frac{1}{n}\sum_{k=1}^{n}\frac{1}{1+\left(\frac{k}{n}\right)^2}=\int_0^1\frac{1}{1+x^2}dx=\frac{\pi}{4}.$$

□

EXAMPLE

$$\begin{aligned}\lim_{n\to\infty}\left\{\prod_{k=n+1}^{2n}\frac{k}{n}\right\}^{\frac{1}{n}} &= \exp\left(\lim_{n\to\infty}\frac{1}{n}\sum_{k=n+1}^{2n}\ln\left(\frac{k}{n}\right)\right)\\ &= \exp\left(\int_1^2\ln(x)dx\right)=\exp\Big(2\ln(2)-1\Big)=\frac{4}{e}.\end{aligned}$$

□

3.7 The Fundamental Theorem of Calculus

There are actually two versions of the fundamental theorem.

THEOREM 79 *Let $f:[a,b]\longrightarrow\mathbb{R}$ be Riemann integrable on $[a,b]$. Let $F:[a,b]\longrightarrow\mathbb{R}$ be continuous on $[a,b]$. Suppose that $F'(x)$ exists and equals $f(x)$ for all $x\in\,]a,b[$. Then*

$$\int_a^b f(x)dx=F(b)-F(a) \tag{3.6}$$

Proof. Let $\epsilon>0$. Then there exists a Riemann partition P of $[a,b]$ such that

$$\left|S(\mathcal{P},f)-\int_a^b f(x)dx\right|<\epsilon \tag{3.7}$$

for every tagged partition $\mathcal{P}$ based on P. Let us suppose that the intervals of P are $[t_{j-1},t_j]$ as j runs from 1 to n. Then, we apply the Mean Value Theorem to F on each of these intervals. This establishes the existence of a point ξ_j of $]t_{j-1},t_j[$ for

$j = 1, \ldots, n$ such that $F(t_j) - F(t_{j-1}) = F'(\xi_j)(t_j - t_{j-1}) = f(\xi_j)(t_j - t_{j-1})$. Adding up these equalities now gives

$$F(b) - F(a) = \sum_{j=1}^{n} \Big(F(t_j) - F(t_{j-1}) \Big)$$

$$= \sum_{j=1}^{n} f(\xi_j)(t_j - t_{j-1}) = S(\mathcal{P}, f) \tag{3.8}$$

for a certain tagged partition $\mathcal{P}$ based on P. Combining (3.7) and (3.8) we get

$$\left| \Big(F(b) - F(a) \Big) - \int_a^b f(x)dx \right| < \epsilon$$

and all dependence on the partition P has disappeared. We are therefore free to take ϵ as small as we wish. Equation (3.6) follows. ■

We tend to think of the fundamental theorem with a constant and b varying, that is, the form

$$\int_a^x f(x)dx = F(x) - F(a)$$

is more common. The flaw in the Theorem 79 is that the differentiability of F is an assumption rather than a conclusion. But, if we want this, then we have to pay extra.

THEOREM 80 *Let $f :]a, b[\longrightarrow \mathbb{R}$ be continuous. Let $c \in]a, b[$ be fixed and define*

$$F(x) = \int_c^x f(t)dt$$

for $x \in]a, b[$. Then F is differentiable at every point of $]a, b[$ and $F'(x) = f(x)$ for all $x \in]a, b[$.

Proof. Let $x \in]a, b[$ be fixed and suppose that $|h|$ is positive, but so small that $x + h \in]a, b[$ also. Note that we have to allow h to be negative here, so some caution is needed. We find

$$F(x+h) - F(x) = \int_x^{x+h} f(t)dt$$

and indeed,

$$\frac{F(x+h)-F(x)}{h} - f(x) = \frac{1}{h}\int_x^{x+h} \Big(f(t)-f(x)\Big)dt$$

Let $\epsilon > 0$. Then, since f is continuous at x there exists $\delta > 0$ such that $|t-x| < \delta \implies |f(t)-f(x)| < \epsilon$. Therefore, $0 < |h| < \delta$ implies that

$$\left|\frac{F(x+h)-F(x)}{h} - f(x)\right| \leq \frac{1}{h}\int_x^{x+h} |f(t)-f(x)|dt \leq \epsilon,$$

so that $F'(x)$ exists and equals $f(x)$. ■

The Fundamental Theorem of Calculus gives us the ability to make substitutions in integrals.

THEOREM 81 (CHANGE OF VARIABLES THEOREM) *Let $\varphi :]a,b[\longrightarrow]\alpha,\beta[$ be a differentiable mapping with continuous derivative. Let $c \in]a,b[$ and $\gamma \in]\alpha,\beta[$ be basepoints such that $\varphi(c) = \gamma$. Let $f :]\alpha,\beta[\longrightarrow \mathbb{R}$ be a continuous mapping. Then for $u \in]a,b[$, we have*

$$\int_\gamma^{\varphi(u)} f(t)dt = \int_c^u f(\varphi(s))\varphi'(s)ds.$$

Proof. We define for $v \in]\alpha,\beta[$,

$$g(v) = \int_\gamma^v f(t)dt.$$

Then according to the Fundamental Theorem of Calculus (Theorem 80), g is differentiable on $]\alpha,\beta[$ and

$$g'(v) = f(v).$$

Then, by the Chain Rule, for $u \in]a,b[$ we have

$$(g\circ\varphi)'(u) = (g'\circ\varphi)(u)\varphi'(u) = (f\circ\varphi)(u)\varphi'(u).$$

Since

$$u \longrightarrow (f\circ\varphi)(u)\varphi'(u)$$

is a continuous mapping, the Fundamental Theorem of Calculus can be applied again to show that if $h :]a, b[\longrightarrow \mathbb{R}$ is defined by

$$h(u) = \int_c^u f(\varphi(s))\varphi'(s)ds,$$

then $h'(u) = (f \circ \varphi)(u)\varphi'(u) = (g \circ \varphi)'(u)$. The Mean Value Theorem now shows that $h(u) - g(\varphi(u))$ is constant. Substituting $u = c$ shows that the constant is zero. Hence $h(u) = g(\varphi(u))$ for all $u \in]a, b[$. This is exactly what was to be proved. ■

The second objective of this section is to be able to differentiate under the integral sign.

THEOREM 82 *Let $\alpha < \beta$ and $a < b$. Suppose that*

$$f, g : [a, b] \times [\alpha, \beta] \longrightarrow \mathbb{R}$$

are continuous mappings such that $\dfrac{\partial g}{\partial t}(t, s)$ exists and equals $f(t, s)$ for all (t, s) in $]a, b[\times [\alpha, \beta]$. Let us define a new function $G : [a, b] \to \mathbb{R}$ by

$$G(t) = \int_\alpha^\beta g(t, s)ds \qquad (a \le t \le b).$$

Then $G'(t)$ exists for $a < t < b$ and

$$G'(t) = \int_\alpha^\beta f(t, s)ds \qquad (a < t < b). \tag{3.9}$$

Proof. For shortness of notation, let us define

$$F(t) = \int_\alpha^\beta f(t, s)ds \qquad (a < t < b).$$

Then, we have for $a < t < b$ and small enough h that

$$\begin{aligned} G(t + h) - G(t) - hF(t) &= \int_\alpha^\beta (g(t + h, s) - g(t, s) - hf(t, s))ds \\ &= \int_\alpha^\beta \left\{ \int_t^{t+h} \Big(f(u, s) - f(t, s) \Big) du \right\} ds \end{aligned}$$

where we have used the Fundamental Theorem of Calculus (Theorem 80) in the last step. Since the points (u, s) and (t, s) are separated by a distance of at most $|h|$, the inner integral satisfies

$$\left| \int_t^{t+h} \Big(f(u,s) - f(t,s) \Big) du \right| \leq |h| \omega_f(|h|).$$

It follows that

$$|G(t+h) - G(t) - hF(t)| \leq (\beta - \alpha)|h|\omega_f(|h|).$$

or equivalently

$$\left| \frac{G(t+h) - G(t)}{h} - F(t) \right| \leq (\beta - \alpha)\omega_f(|h|).$$

Since f is a continuous function on the compact space $[a, b] \times [\alpha, \beta]$, it follows that f is uniformly continuous, so as $h \longrightarrow 0$, we find that $(\beta - \alpha)\omega_f(|h|) \longrightarrow 0$ and so

$$\lim_{h \to 0} \frac{G(t+h) - G(t)}{h} = F(t),$$

showing that G is differentiable at $t \in]a, b[$ with derivative $F(t)$. ■

In Theorem 82 we have avoided the issues of one-sided derivatives by establishing (3.9) only on the open interval $]a, b[$. This is for the sake of simplicity only. If we impose, for example, the additional condition that $\frac{\partial g}{\partial t}(a, s)$ exists as a right-hand derivative and equals $f(a, s)$ for all s in $[\alpha, \beta]$, then the same argument (with $t = a$ and $h > 0$) will show that $G'(a)$ exists as a right-hand derivative and that $G'(a) = \int_\alpha^\beta f(a, s) ds$.

There is also a version of Theorem 82 in which α and β are allowed to vary with t. This is called ***Leibnitz' formula***

$$\frac{d}{dt} \int_{\alpha(t)}^{\beta(t)} g(t, x) dx = g(t, \beta(t))\beta'(t) - g(t, \alpha(t))\alpha'(t) + \int_{\alpha(t)}^{\beta(t)} \frac{\partial g(t, x)}{\partial t} dx.$$

This is best proved (with appropriate hypotheses) as a consequence of Theorem 82 using the several variable chain rule. Since this version of the chain rule is not available to us, we will omit the proof. It can also be proved directly.

EXAMPLE Consider the following two functions defined for $x \geq 0$.

$$f(x) = \left(\int_0^x e^{-t^2} dt \right)^2 \qquad g(x) = \int_0^1 \frac{e^{-x^2(1+t^2)}}{t^2+1} dt$$

Using the Fundamental Theorem of Calculus (Theorem 80), we have

$$f'(x) = 2e^{-x^2} \int_0^x e^{-t^2} dt = 2 \int_0^1 e^{-x^2 - s^2 x^2} x ds$$

after making the substitution $t = sx$. On the other hand, to differentiate g we use Theorem 82 to get

$$g'(x) = \int_0^1 -2x(1+t^2) \frac{e^{-x^2(1+t^2)}}{t^2+1} dt = -2 \int_0^1 e^{-x^2 - s^2 x^2} x ds$$

after simplifying and replacing the t by s. Clearly we have $f'(x) + g'(x) = 0$ for all $x > 0$ and since $f(0) + g(0) = \int_0^1 \frac{1}{1+t^2} dt = \frac{\pi}{4}$. We deduce that

$$\left(\int_0^x e^{-t^2} dt \right)^2 + \int_0^1 \frac{e^{-x^2(1+t^2)}}{t^2+1} dt = \frac{\pi}{4} \tag{3.10}$$

for all $x \geq 0$. We clearly have

$$0 \leq \int_0^1 \frac{e^{-x^2(1+t^2)}}{t^2+1} dt \leq e^{-x^2},$$

and can therefore deduce from (3.10) that

$$0 \leq \frac{\pi}{4} - \left(\int_0^x e^{-t^2} dt \right)^2 \leq e^{-x^2}$$

and it follows that

$$\lim_{x \to \infty} \int_0^x e^{-t^2} dt = \frac{\sqrt{\pi}}{2}.$$

□

3.8 Improper Integrals and the Integral Test

Improper integrals are an attempt to extend the Riemann integral to the case where the interval over which the integral is taken is of infinite length or the function being integrated is unbounded. It should be said that the Lebesgue theory does not suffer from the restriction of bounded functions or intervals of finite length and it handles these cases automatically. However, the results are not the same. Sometimes the improper Riemann integral can exist when the Lebesgue integral does not.

The improper integral is defined using limits. Thus by $\int_a^\infty f(x)dx$ we mean

$$\lim_{b\longrightarrow\infty}\int_a^b f(x)dx.$$

EXAMPLE We evaluate $\int_1^\infty x^{-2}dx$. By the Fundamental Theorem of Calculus (Theorem 80), we get $\int_1^b x^{-2}dx = 1 - b^{-1}$, so letting $b \longrightarrow \infty$ we have $\int_1^\infty x^{-2}dx = 1$. □

EXAMPLE We saw in the last section that $\displaystyle\int_0^\infty e^{-t^2}dt = \lim_{x\to\infty}\int_0^x e^{-t^2}dt = \frac{\sqrt{\pi}}{2}$. □

These examples do not involve cancellation, the integrand was positive. In this case, the limit is an increasing one and we have a situation similar to the convergence of series of positive terms.

LEMMA 83 *If f is a nonnegative function on $[0,\infty[$ which is Riemann integrable on every interval $[0,b]$ for all $b > 0$ and if $\int_0^\infty f(x)dx < \infty$, then, given $\epsilon > 0$ there exists $c > 0$ such that $c \le a \le b$ implies*

$$\int_a^b f(x)dx \le \epsilon \tag{3.11}$$

and, indeed

$$\int_c^\infty f(x) \le \epsilon. \tag{3.12}$$

Proof. Indeed, by Theorem 70 we have

$$\int_0^b f(x)dx = \int_0^c f(x)dx + \int_c^b f(x)dx. \tag{3.13}$$

Now letting $b \longrightarrow \infty$, the left-hand side of (3.13) tends to a limit. Thus, so does the second term on the right in (3.13) and we have

$$\int_0^\infty f(x)dx = \int_0^c f(x)dx + \int_c^\infty f(x)dx. \tag{3.14}$$

But now as $c \longrightarrow \infty$, the first term on the right in (3.14) tends to the left-hand side of (3.14). Therefore $\lim\limits_{c\to\infty} \displaystyle\int_c^\infty f(x)dx = 0$. So, given $\epsilon > 0$, there exists c such that (3.12) holds. A fortiori, (3.11) also holds. ■

LEMMA 84 *If f is a function on $[0, \infty[$ which is Riemann integrable on every interval $[0, b]$ for all $b > 0$ and if $\int_0^\infty |f(x)|dx < \infty$, then, $\int_0^\infty f(x)dx$ exists and*

$$\left|\int_0^\infty f(x)dx\right| \leq \int_0^\infty |f(x)|dx \tag{3.15}$$

Proof. This is a messy proof and follows the same line as the proof of extension by uniform continuity. We start with a sequence $a_n > 0$ tending to ∞. Let

$$I_n = \int_0^{a_n} f(x)dx.$$

We will show that (I_n) is a Cauchy sequence. So, let $\epsilon > 0$. Then, according to Lemma 83, there exists $c > 0$ such that (3.11) holds for $|f|$. Then, choose N such that $n > N$ implies $a_n > c$. We have, for $p \geq q > N$ that

$$|I_p - I_q| = \left|\int_0^{a_p} f(x)dx - \int_0^{a_q} f(x)dx\right| = \left|\int_{a_q}^{a_p} f(x)dx\right| \leq \int_{a_q}^{a_p} |f(x)|dx \leq \epsilon$$

This shows that (I_n) is a Cauchy sequence and therefore converges to some limit which we will call I. We claim that $\lim_{b\to\infty} \int_0^b f(x)dx = I$. If not, then there is another sequence (b_n) such that $\lim_{n\to\infty} \int_0^{b_n} f(x)dx$ does not converge to I. Now combine the sequences (a_n) and (b_n) with one as the even subsequence and the other as the odd subsequence and reapply the original argument. We will get the required contradiction. ■

This deals with absolutely convergent improper integrals, but some also exist because of cancellation.

EXAMPLE Consider $\int_0^\infty \frac{\sin x}{x} dx$. Strictly speaking there is a problem at both 0 and at ∞, but if we believe in the fact

$$\lim_{x \to 0} \frac{\sin x}{x} = 1$$

then we see that the integrand can be extended continuously to the left-hand endpoint 0. We will assume that this has been done. Splitting the range of integration and integrating by parts, we get

$$\begin{aligned}
\int_0^b \frac{\sin x}{x} dx &= \int_0^{2\pi} \frac{\sin x}{x} dx + \int_{2\pi}^b \frac{\sin x}{x} dx \\
&= \int_0^{2\pi} \frac{\sin x}{x} dx + \left[\frac{1 - \cos x}{x} \right]_{2\pi}^b + \int_{2\pi}^b \frac{1 - \cos x}{x^2} dx \\
&= \int_0^{2\pi} \frac{\sin x}{x} dx + \frac{1 - \cos b}{b} + \int_{2\pi}^b \frac{1 - \cos x}{x^2} dx
\end{aligned}$$

and it is clear that the limit exists as $b \longrightarrow \infty$ since

$$\int_{2\pi}^b \frac{1 - \cos x}{x^2} dx \le \int_{2\pi}^b 2x^{-2} dx \le \frac{1}{\pi} < \infty.$$

In fact, $\int_0^\infty \frac{\sin x}{x} dx = \frac{\pi}{2}$. □

There are other kinds of improper integral and we leave the details to the reader's imagination. They can be two-sided as in

$$\int_{-\infty}^\infty e^{-\frac{1}{2}x^2} dx$$

or they can with an unbounded integrand on a bounded interval as in

$$\int_0^1 x^{-\frac{1}{2}} dx = \lim_{t \to 0+} \int_t^1 x^{-\frac{1}{2}} dx = \lim_{t \to 0+} 2(1 - t^{\frac{1}{2}}) = 2.$$

One of the major applications of improper integrals is to a test for convergence of series. This is significant, because sometimes integrals can be computed explicitly. In today's world, we are better at figuring out integrals than we are at summing series.

THEOREM 85 *Let f be a positive bounded decreasing continuous function on $[0, \infty[$. Then*

$$\int_1^\infty f(x)dx \leq \sum_{n=1}^\infty f(n) \leq \int_0^\infty f(x)dx.$$

In particular, the series $\sum_{n=1}^\infty f(n)$ and the integral $\int_0^\infty f(x)dx$ converge or diverge together.

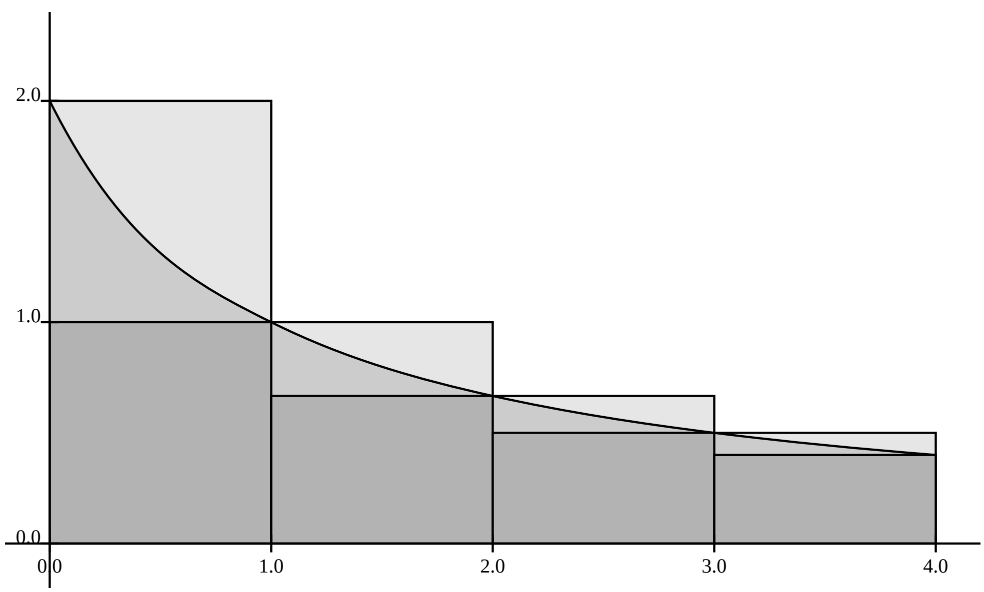

Figure 3.7: Areas involved the the Integral Test. Well, not quite! What is offered is a pictorial proof of $\sum_{n=1}^4 f(n) \leq \int_0^4 f(x)dx \leq \sum_{n=0}^3 f(n)$. The quantity $\sum_{n=1}^4 f(n)$ is represented as the area shaded with the darkest shade of gray, $\int_0^4 f(x)dx$ corresponds to the area shaded with the darkest shade of gray and the middle gray and finally $\sum_{n=0}^3 f(n)$ is represented as the area shaded in any gray. This shows how to bound an integral above and below by sums. We have stated the integral test the other way around, because usually you know the integral and want to estimate the sum.

Proof. Let $g(x) = \sum_{n=1}^\infty f(n)\mathbb{1}_{[n,n+1[}(x)$. Then $f(x) \leq g(x)$ for all $x \geq 1$. This is because, if $1 \leq n \leq x < n+1$, then $g(x) = f(n) \geq f(x)$. So

$$\int_1^{n+1} f(x)dx \leq \int_1^{n+1} g(x)dx = \int_1^{n+1} \sum_{k=1}^{n} f(k)\mathbb{1}_{[k,k+1[}(x)dx = \sum_{k=1}^{n} f(k)$$

Letting $n \longrightarrow \infty$ gives the left-hand inequality. For the other direction, we take $h(x) \sum_{n=1}^{\infty} f(n)\mathbb{1}_{]n-1,n]}(x)$. Then, for $x > 0$ we have $h(x) \leq f(x)$. This is because if $n-1 < x \leq n$, then $h(x) = f(n) \leq f(x)$. We get

$$\int_0^{n} f(x)dx \geq \int_0^{n} h(x)dx = \int_0^{n} \sum_{k=1}^{n} f(k)\mathbb{1}_{]k-1,k]}(x)dx = \sum_{k=1}^{n} f(k)$$

and the result follows on letting $n \longrightarrow \infty$. ■

EXAMPLE

$$\sum_{n=1}^{\infty} \frac{1}{n^2} \leq 1 + \sum_{n=2}^{\infty} \frac{1}{n^2} \leq 1 + \int_1^{\infty} x^{-2}dx = 2$$

□

EXAMPLE We can get precise information about the divergence of the harmonic series. We let

$$\begin{aligned}\gamma_n &= -\ln n + \sum_{k=1}^{n} \frac{1}{k} \\ &= 1 - \int_1^n \frac{1}{x}dx + \sum_{k=2}^{n} \frac{1}{k} \\ &= 1 - \int_1^n \left(\frac{1}{x} - \frac{1}{\lceil x \rceil}\right) dx\end{aligned}$$

and see that γ_n is decreasing with n because the integrand $\dfrac{1}{x} - \dfrac{1}{\lceil x \rceil}$ is nonnegative and the interval of integration is increasing with n. Also, its clear that

$$\int_1^{\infty} \left(\frac{1}{x} - \frac{1}{\lceil x \rceil}\right) dx \leq \int_1^{\infty} \frac{\lceil x \rceil - x}{x\lceil x \rceil}dx \leq \int_1^{\infty} \frac{1}{x^2}dx = 1,$$

so (γ_n) is bounded below by 0 and must therefore tend to a limit γ between 0 and 1. This number is called Euler's constant. □

3.9 Taylor's Theorem

Before we can tackle Taylor's Theorem, we need to extend the Mean-Value Theorem.

THEOREM 86 (EXTENDED MEAN VALUE THEOREM) *Let a and b be real numbers such that $a < b$. Let $g, h : [a, b] \longrightarrow \mathbb{R}$ be continuous maps. Suppose that g and h are differentiable at every point of $]a, b[$. Then there exists ξ such that $a < \xi < b$ and*

$$(g(b) - g(a))h'(\xi) = g'(\xi)(h(b) - h(a)).$$

Proof. Let us define

$$f(x) = g(x)(h(b) - h(a)) - (g(b) - g(a))h(x).$$

Then routine calculations show that

$$f(a) = g(a)h(b) - g(b)h(a) = f(b).$$

Since f is continuous on $[a, b]$ and differentiable on $]a, b[$, we can apply Rolle's Theorem to establish the existence of $\xi \in \,]a, b[$ such that $f'(\xi) = 0$, a statement equivalent to the desired conclusion. ■

DEFINITION *Let f be a function $f :]a, b[\longrightarrow \mathbb{R}$ which is n times differentiable. Formally this means that the successive derivatives f', f'',. . . ,$f^{(n)}$ exist on $]a, b[$. Let $c \in \,]a, b[$ be a basepoint. Then we can construct the* ***Taylor Polynomial*** *$T_{n,c}f$ of order n at c by*

$$T_{n,c}f(x) = \sum_{k=0}^{n} \frac{1}{k!} f^{(k)}(c)(x - c)^k.$$

To make the notations clear we point out that $f^{(0)} = f$, that $0! = 1$ and that $(x - c)^0 = 1$. In fact even $0^0 = 1$ because it is viewed as an “empty product”.

THEOREM 87 (TAYLOR'S THEOREM) *Let f be a function $f\ :\ ]a,b[\ \longrightarrow \mathbb{R}$ which is $n+1$ times differentiable. Let $c \in\ ]a,b[$ be a basepoint. Then there exists a point ξ between c and x, such that*

$$f(x) = T_{n,c}f(x) + \frac{1}{(n+1)!}f^{(n+1)}(\xi)(x-c)^{n+1}. \tag{3.16}$$

The statement ξ is between c and x means that

$$\begin{cases} c < \xi < x & \text{if } c < x, \\ c = \xi = x & \text{if } c = x, \\ x < \xi < c & \text{if } c > x. \end{cases}$$

The second term on the right of (3.16) is called the ***remainder term*** and in fact this specific form of the remainder is called the ***Lagrange remainder*** . It is the most common form. When we look at (3.16), we think of writing the function f as a polynomial plus an error term (the remainder). Of course, there is no guarantee that the remainder term is small.

All this presupposes that f is a function of x and indeed this is the obvious point of view when we are applying Taylor's Theorem. However for the proof, we take the other point of view and regard x as the constant and c as the variable.

Proof. First of all, if $x = c$ there is nothing to prove. We can therefore assume that $x \neq c$. We regard x as fixed and let c vary in $]a,b[$. We define

$$g(c) = f(x) - T_{n,c}f(x) \qquad \text{and} \qquad h(c) = (x-c)^{n+1}.$$

On differentiating g with respect to c we obtain a telescoping sum which yields

$$g'(\xi) = -\frac{1}{n!}f^{(n+1)}(\xi)(x-\xi)^n. \tag{3.17}$$

On the other hand we have, differentiating h with respect to c,

$$h'(\xi) = -(n+1)(x-\xi)^n.$$

Applying now the extended Mean-Value Theorem, we obtain

$$(g(c) - g(x))h'(\xi) = g'(\xi)(h(c) - h(x)),$$

where ξ is between c and x. Since both $g(x) = 0$ and $h(x) = 0$ (remember g and h are viewed as functions of c, so here we are substituting $c = x$), this is equivalent to

$$(f(x) - T_{n,c}f(x))(-(n+1)(x-\xi)^n) = (-\frac{1}{n!}f^{(n+1)}(\xi)(x-\xi)^n)(x-c)^{n+1},$$

Since $x \neq c$, we have that $x \neq \xi$ and we may divide by $(x - \xi)^n$ and obtain the desired conclusion. ■

In many situations, we can use estimates on the Lagrange remainder to establish the validity of power series expansion

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(c)(x - c)^k,$$

for x in some open interval around c. Such estimates are sometimes fraught with difficulties because all that one knows about ξ is that it lies between c and x. Because the precise location of ξ is not known, some information may have been lost irrecoverably. Usually, there are better ways of establishing the validity of power series expansions. These will be investigated later in this course.

However, there is a way of obtaining estimates of the Taylor remainder in which no information is sacrificed.

THEOREM 88 (INTEGRAL REMAINDER THEOREM) *Suppose that* $f : \,]a, b[\longrightarrow \mathbb{R}$ *is* $n + 1$ *times differentiable and that* $f^{(n+1)}$ *is continuous. Let* $c \in \,]a, b[$ *be a basepoint. Then we have for* $x \in \,]a, b[$

$$f(x) = T_{n,c}f(x) + \frac{1}{n!} \int_{\xi=c}^{x} (x - \xi)^n \, f^{(n+1)}(\xi) \, d\xi, \tag{3.18}$$

or equivalently by change of variables

$$f(x) = T_{n,c}f(x) + \frac{(x - c)^{n+1}}{n!} \int_{t=0}^{1} (1 - t)^n \, f^{(n+1)}\Big((1 - t)c + tx\Big) \, dt. \tag{3.19}$$

This Theorem provides an explicit formula for the remainder term which involves an integral. Note that in order to define the integral it is supposed that f is slightly more regular than is the case with the Lagrange form of the remainder.

Proof. Again we tackle (3.18) by viewing x as the constant and c as the variable. Equation (3.18) follows immediately from the Fundamental Theorem of Calculus (Theorem 80) and (3.17). The second formulation (3.19) follows by the Change of Variables Theorem, using the substitution $\xi = (1 - t)c + tx$. ■

EXAMPLE Let $\alpha > 0$ and consider

$$f(x) = (1-x)^{-\alpha}$$

for $-1 < x < 1$. The Taylor series of this function is

$$f(x) = 1 + \alpha x + \frac{\alpha(\alpha+1)}{2!}x^2 + \dots$$

actually valid for $-1 < x < 1$. If we try to obtain this result using the Lagrange form of the remainder

$$\frac{\alpha(\alpha+1)\dots(\alpha+n)}{(n+1)!}(1-\xi)^{-\alpha-n-1}x^{n+1}$$

we are able to show that the remainder tends to zero as n tends to infinity provided that

$$\sup\left|\frac{x}{1-\xi}\right| < 1,$$

where the sup is taken over all ξ between 0 and x. If $x > 0$ the worst case is when ξ is very close to x. Convergence of the Lagrange remainder to zero is guaranteed only if $0 < x < \frac{1}{2}$. On the other hand, if $x < 0$ then the worst location of ξ is $\xi = 0$. Convergence of the Lagrange remainder is then guaranteed for $-1 < x < 0$. Combining the two cases, we see that the Lagrange remainder can be controlled only for $-1 < x < \frac{1}{2}$.

For the same function, the integral form of the remainder is

$$\frac{\alpha(\alpha+1)\dots(\alpha+n)}{n!}\int_0^x (1-\xi)^{-\alpha-n-1}(x-\xi)^n d\xi.$$

For ξ between 0 and x we have

$$\left|\frac{x-\xi}{1-\xi}\right| \le |x|,$$

for $-1 < x < 1$. This estimate allows us to show that the remainder tends to zero over the full range $-1 < x < 1$. □

4

Sequences of Functions

In this chapter we look at the convergence of sequences of functions. In chapter 1 we spent a lot of time introducing the metric space concept to deal with convergence, so it comes as something a a nuisance that metric spaces are not ideally suited to describing the situation here.

4.1 Pointwise Convergence

The simplest type of convergence is in the pointwise sense.

DEFINITION *Let X be a set and let $f_n : X \longrightarrow \mathbb{R}$ for $n \in \mathbb{N}$. Then (f_n) is a sequence of real-valued functions on X. We say that (f_n) converges to a function $f : X \longrightarrow \mathbb{R}$ iff for every $x \in X$, we have $f_n(x) \longrightarrow f(x)$ as $n \longrightarrow \infty$.*

In other words, pointwise convergence is convergence at every point of the domain. We could in this definition replace $\mathbb{R}$ by a general metric space. Unless X is finite, you cannot find a metric on the space of all real-valued function on X for which the metric space convergence agrees with pointwise convergence.

EXAMPLE Let us consider the following sequence of functions defined on the interval $[0, 1]$.

$$f_n(x) = \begin{cases} 1 - nx & \text{if } 0 \le x \le \dfrac{1}{n}, \\ 0 & \text{if } \dfrac{1}{n} \le x \le 1. \end{cases}$$

The two cases agree on their overlap, i.e. when $x = \dfrac{1}{n}$. Now, if $x = 0$ we have $f_n(0) = 1$ for all n. So $(f_n(0))$ is a constant sequence and it converges to its constant value 1. On the other hand, if $0 < x \leq 1$, then as soon as $n \geq \left\lceil \frac{1}{x} \right\rceil$, we have $f_n(x) = 0$, so eventually the sequence vanishes. Hence the limit in this case is 0. We have shown that $f_n \longrightarrow f$ pointwise on $[0, 1]$ where f is given by

$$f(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

One thing that we learn from this example is that we should not expect the pointwise limit of continuous functions to be continuous. □

4.2 Uniform Convergence

The definition of uniform convergence is similar to that of uniform continuity in that the definition requires one of the quantities to be chosen independent of another.

DEFINITION *Let X be a set and let (f_n) be a sequence of real-valued functions on X. We say that (f_n) converges to a function $f : X \longrightarrow \mathbb{R}$* ***uniformly*** *iff for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that*

$$n \geq N \qquad \Longrightarrow \qquad |f(x) - f_n(x)| < \epsilon.$$

So, here it is the N which has to be chosen to be independent of the $x \in X$. If N were allowed to depend on x we would have exactly the definition of pointwise convergence. Obviously then $f_n \longrightarrow f$ uniformly implies $f_n \longrightarrow f$ pointwise.

Again, we could replace $\mathbb{R}$ by a general metric space and the definition would still make sense.

In order to stress the independence on x we can embody that in a supremum. It is easy to check that the above definition is equivalent to the following. The sequence (f_n) converges uniformly to f if for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$n \geq N \qquad \Longrightarrow \qquad \sup_{x \in X} |f(x) - f_n(x)| < \epsilon.$$

Now this looks like a norm and so we may think of uniform convergence as convergence in a normed space. But it does not work quite perfectly. We denote by $B(X)$ the space of all bounded real-valued function on X. Then it can be shown that

$$\|f\| = \sup_{x \in X} |f(x)|$$

defines a norm on $B(X)$ and furthermore, convergence in this norm is exactly uniform convergence. However doing this unfortunately restricts uniform convergence to bounded functions.

EXAMPLE Consider the case $X = \mathbb{R}$ and let $f_n(x) = x + \frac{1}{n}\sin(x)$. Let also $f(x) = x$. Then

$$\sup_{x \in \mathbb{R}} |f_n(x) - f(x)| = \sup_{x \in \mathbb{R}} \left| \frac{1}{n} \sin(x) \right| = \frac{1}{n} \longrightarrow 0,$$

so that convergence is uniform. However the functions f_n and f are not bounded. □

EXAMPLE How do we show that a sequence of functions does not converge uniformly? Consider $f_n(x) = \dfrac{nx}{n^2x^2+1}$ on $[0, \infty[$. The first thing to do is to see if the sequence converges pointwise. If it doesn't converge pointwise, then it certainly doesn't converge uniformly. But, if it does converge pointwise, we still get useful information. In the example in question, it is easy to see that $f_n(x) \longrightarrow 0$ as $n \longrightarrow \infty$ for every $x \in [0, \infty[$. This means that *if* (f_n) converges uniformly, then it converges uniformly to the zero function. The pointwise convergence identifies the function that would have to be the limit. So, now we know that we must have $f \equiv 0$, we can calculate

$$\sup_{x \geq 0} |f_n(x) - f(x)| = \sup_{x \geq 0} \left| \frac{nx}{n^2x^2+1} \right| = \frac{1}{2} > 0$$

for all n, the sup being taken at $x = \frac{1}{n}$. Since $\sup_{x \geq 0} |f_n(x) - f(x)|$ is bounded away from zero independent of n, convergence is not uniform. □

Perhaps the most important thing about uniform convergence is that it preserves continuity.

THEOREM 89 *If X is a metric space, f_n and f are real-valued functions on X, f_n is continuous for each $n \in \mathbb{N}$ and $f_n \longrightarrow f$ uniformly on X, then f is also continuous.*

Proof. The result is actually true for continuity at a point. Let $x_0 \in X$. We will show that f is continuous at x_0. Let $\epsilon > 0$. Then, there exists $N \in \mathbb{N}$ such that

$$n \geq N \qquad \Longrightarrow \qquad \sup_{x \in X} |f(x) - f_n(x)| < \frac{\epsilon}{3}. \tag{4.1}$$

Now we use the fact that f_N is continuous at x_0. There exists $\delta > 0$ such that

$$|x - x_0| < \delta \qquad \Longrightarrow \qquad |f_N(x) - f_N(x_0)| < \frac{\epsilon}{3}.$$

We apply (4.1) with $n = N$ and at both x and x_0 to get

$$\begin{aligned} |f(x) - f(x_0)| &\leq |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| \\ &< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

for $|x - x_0| < \delta$. This shows that f is continuous at x_0. Since x_0 is an arbitrary point of X we deduce that f is continuous on X. ∎

In case you were wondering, the following theorem is also true and the proof is so similar to the previous one that we omit it.

THEOREM 90 *If X is a metric space, f_n and f are real-valued functions on X, f_n is uniformly continuous for each $n \in \mathbb{N}$ and $f_n \longrightarrow f$ uniformly on X, then f is also uniformly continuous.*

EXAMPLE Let $f : [0,1] \times [0,1] \longrightarrow \mathbb{R}$ be a continuous mapping of the unit square in the plane. We will use the Euclidean metric on $[0,1] \times [0,1]$. We can now define mappings $g_n : [0,1] \longrightarrow \mathbb{R}$ and $g : [0,1] \longrightarrow \mathbb{R}$ by

$$g_n(x) = f\left(x, \frac{1}{n}\right) \text{ and } g(x) = f(x,0)$$

effectively horizontal slices of the function f. Then $g_n \longrightarrow g$ uniformly as $n \longrightarrow \infty$. To prove this, we observe first that $[0,1] \times [0,1]$ is a bounded closed subset of $\mathbb{R}^2$ and is hence sequentially compact. Therefore f is uniformly continuous. So

$$|g_n(x) - g(x)| = \left| f\left(x, \frac{1}{n}\right) - f(x,0) \right| \leq \omega_f(\|(x, n^{-1}) - (x,0)\|) = \omega_f(n^{-1})$$

So, we find

$$\sup_{0 \leq x \leq 1} |g_n(x) - g(x)| \leq \omega_f(n^{-1})$$

and $\omega_f(n^{-1}) \longrightarrow 0$ as $n \longrightarrow \infty$. □

EXAMPLE The result of the previous example fails if the whole x-axis is used in place of $[0,1]$. The function $f(x,y) = \sin(xy)$ is certainly continuous on $\mathbb{R}\times[0,1]$ and the corresponding functions are

$$g_n(x) = \sin\left(\frac{x}{n}\right) \text{ and } g(x) = 0$$

defined on the whole of $\mathbb{R}$. But $\sup_{x\in\mathbb{R}} |g_n(x) - g(x)| = 1$, so convergence is not uniform here. Pointwise convergence does hold. □

So uniform continuity is an important tool that is often vital to establish uniform convergence. We can for instance study approximation by piecewise linear functions. Let $f : [0,1] \longrightarrow \mathbb{R}$ be a given continuous function. Then the nth piecewise linear approximation $P_n(f,\cdot)$ is the function that agrees with f at the $n+1$ points $\frac{k}{n}$ $(k = 0, 1, \ldots, n)$ and is linear on each of the intervals $\left[\frac{k-1}{n}, \frac{k}{n}\right]$ for $k = 1, 2, \ldots, n$. There is a succinct way of writing down $P_n(f,\cdot)$. Let

$$\Delta(x) = \begin{cases} 1+x & \text{if } -1 \le x \le 0, \\ 1-x & \text{if } 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Then we have simply $P_n(f,x) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) \Delta(nx-k)$.

THEOREM 91 *Let $f : [0,1] \longrightarrow \mathbb{R}$ be a given continuous function. Then $(P_n(f,\cdot))$ converges uniformly to f on $[0,1]$.*

Proof. Note that for $x \in [0,1]$ we have $\sum_{k=0}^{n} \Delta(nx-k) = 1$. Hence we can write

$$f(x) - P_n(f,x) = \sum_{k=0}^{n} \left(f(x) - f\left(\frac{k}{n}\right)\right) \Delta(nx-k)$$

and therefore

$$|f(x) - P_n(f,x)| \le \sum_{k=0}^{n} \left|f(x) - f\left(\frac{k}{n}\right)\right| \Delta(nx-k)$$

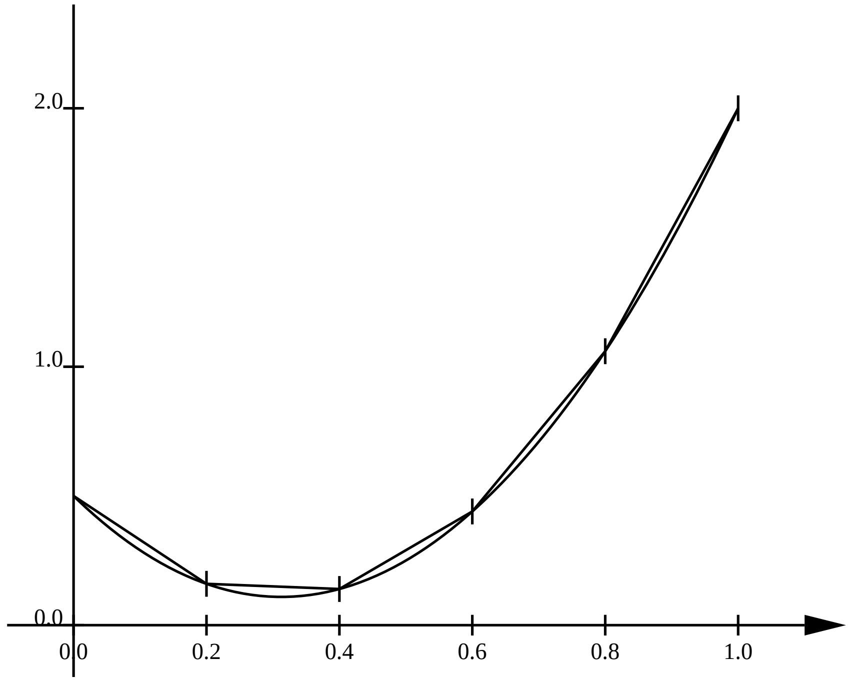

Figure 4.1: A function f and the corresponding $P_5(f,\cdot)$

Now, for each term, either we have $\Delta(nx-k)=0$ or we have $\left|x-\dfrac{k}{n}\right|<\dfrac{1}{n}$, so we can write

$$|f(x)-P_n(f,x)| \le \sum_{k=0}^{n} \omega_f\left(\frac{1}{n}\right)\Delta(nx-k) = \omega_f\left(\frac{1}{n}\right). \tag{4.2}$$

The right-hand side of (4.2) is independent of x, so we can say

$$\sup_{x\in[0,1]} |f(x)-P_n(f,x)| \le \omega_f\left(\frac{1}{n}\right)$$

and since f is uniformly continuous (being continuous on a sequentially compact set) we find that $\omega_f\left(\dfrac{1}{n}\right) \longrightarrow 0$ as $n \longrightarrow \infty$. ∎

We can prove a similar theorem for approximation by polynomials, but it is substantially harder.

THEOREM 92 (BERNSTEIN APPROXIMATION THEOREM) *Let* $f : [0,1] \longrightarrow \mathbb{R}$ *be a continuous function. Define the* n*th* ***Bernstein polynomial*** *by*

$$B_n(f,x) = \sum_{k=0}^{n} {}^nC_k\, f\left(\frac{k}{n}\right) x^k(1-x)^{n-k}.$$

Then $(B_n(f,\cdot))$ *converges uniformly to* f *on* $[0,1]$.

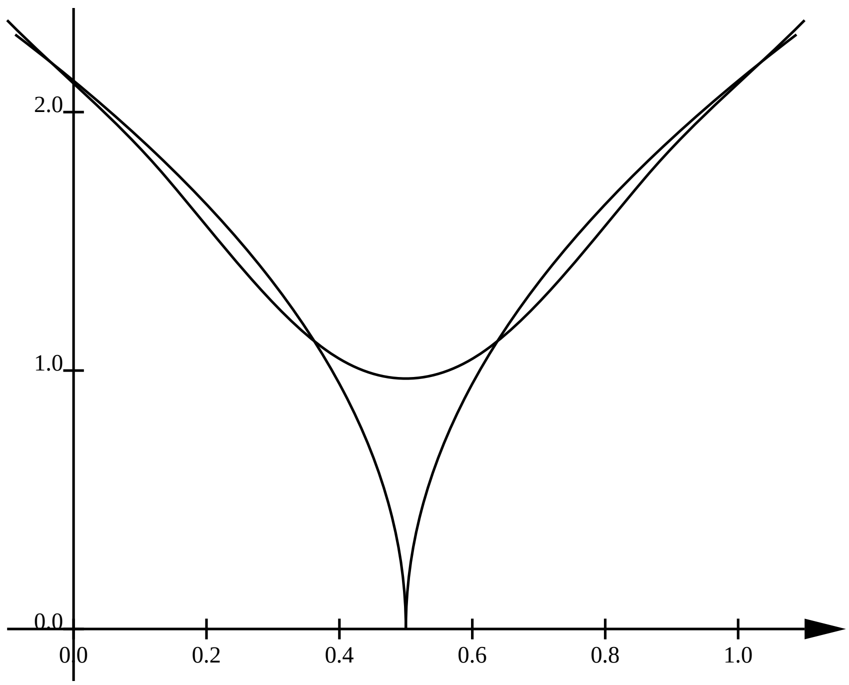

Figure 4.2: The function $f(x) = 3\sqrt{\left|x - \frac{1}{2}\right|}$ and the corresponding $B_6(f,\cdot)$

Before tackling the proof, we need to do some fairly horrible calculations.

LEMMA 93 *We have*

$$\sum_{k=0}^{n} \left(x - \frac{k}{n}\right)^2 \; {}^nC_k\, x^k(1-x)^{n-k} = \frac{1}{n}x(1-x). \tag{4.3}$$

Proof. We start with the Binomial Theorem

$$(x+y)^n = \sum_{k=0}^{n} {}^nC_k\, x^k y^{n-k} \tag{4.4}$$

which we differentiate twice partially with respect to x to get

$$n(x+y)^{n-1} = \sum_{k=0}^{n} k\, {}^nC_k\, x^{k-1} y^{n-k} \tag{4.5}$$

and

$$n(n-1)(x+y)^{n-2} = \sum_{k=0}^{n} k(k-1)\, {}^nC_k\, x^{k-2} y^{n-k} \tag{4.6}$$

Multiplying (4.5) by x, (4.6) by x^2 and then substituting $y = 1 - x$ into (4.4), (4.5) and (4.6) we get

$$1 = \sum_{k=0}^{n} {}^nC_k\, x^k (1-x)^{n-k}, \tag{4.7}$$

$$nx = \sum_{k=0}^{n} k\ {}^nC_k\, x^k (1-x)^{n-k}, \tag{4.8}$$

$$n(n-1)x^2 = \sum_{k=0}^{n} k(k-1)\ {}^nC_k\, x^k (1-x)^{n-k}. \tag{4.9}$$

Then it is easy to see that

$$\begin{aligned}
&\sum_{k=0}^{n} \left(x - \frac{k}{n}\right)^2\ {}^nC_k\, x^k (1-x)^{n-k} \\
&= \sum_{k=0}^{n} \left(x^2 - \frac{2k}{n}x + \frac{k^2-k}{n^2} + \frac{k}{n^2}\right)\ {}^nC_k\, x^k (1-x)^{n-k}, \\
&= x^2 - 2x^2 + \frac{n(n-1)}{n^2}x^2 + \frac{n}{n^2}x, \\
&= \frac{1}{n}x(1-x).
\end{aligned}$$

by applying (4.7), (4.8) and (4.9). ∎

Proof of the Bernstein Approximation Theorem. From (4.3) we have for $\delta > 0$ the Chebyshev inequality

$$\sum_{|x-\frac{k}{n}|>\delta} \delta^2 \ {}^nC_k\, x^k(1-x)^{n-k} \le \sum_{|x-\frac{k}{n}|>\delta} \left(x - \frac{k}{n}\right)^2 \ {}^nC_k\, x^k(1-x)^{n-k}$$
$$\le \frac{1}{n}x(1-x).$$

We are now ready to study the approximation. Since f is continuous on the compact set $[0,1]$ it is also uniformly continuous and bounded. Thus we have

$$\begin{aligned} f(x) - B_n(f,x) &= f(x) - \sum_{k=0}^{n} {}^nC_k\, f\left(\frac{k}{n}\right) x^k(1-x)^{n-k}, \\ &= \sum_{k=0}^{n} {}^nC_k \left(f(x) - f\left(\frac{k}{n}\right)\right) x^k(1-x)^{n-k}, \end{aligned}$$

and

$$\begin{aligned} |f(x) - B_n(f,x)| &\le \sum_{k=0}^{n} {}^nC_k \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k(1-x)^{n-k}, \\ &\le E_1 + E_2, \end{aligned} \tag{4.10}$$

where

$$\begin{aligned} E_1 &= \sum_{|x-\frac{k}{n}|>\delta} {}^nC_k \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k(1-x)^{n-k}, \\ &\le 2\|f\|_\infty \delta^{-2}\frac{1}{n}x(1-x), \\ &\le \frac{1}{2n}\|f\|_\infty \delta^{-2}, \end{aligned} \tag{4.11}$$

and

$$\begin{aligned} E_2 &= \sum_{|x-\frac{k}{n}|\le\delta} {}^nC_k \left|f(x) - f\left(\frac{k}{n}\right)\right| x^k(1-x)^{n-k}, \\ &\le \sum_{k=0}^{n} {}^nC_k\, \omega_f(\delta) x^k(1-x)^{n-k}, \\ &= \omega_f(\delta). \end{aligned} \tag{4.12}$$

Suppose now that ϵ is a strictly positive number. Then, using the uniform continuity of f, choose $\delta > 0$ so small that $\omega_f(\delta) < \frac{1}{2}\epsilon$. Then, with δ now fixed, select N so large that $\frac{1}{2N}\|f\|_\infty \delta^{-2} < \frac{1}{2}\epsilon$. It follows by combining (4.10), (4.11) and (4.12) that

$$\sup_{0 \le x \le 1} |f(x) - B_n(f, x)| \le \epsilon \qquad \forall n \ge N,$$

as required for uniform convergence of the Bernstein polynomials to f. ∎

This proof does not address the question of motivation. Where do the Bernstein polynomials come from? To answer this question, we need to assume that the reader has a rudimentary knowledge of probability theory. Let X be a random variable taking values in $\{0, 1\}$, often called a Bernoulli random variable. Assume that it takes the value 1 with probability x and the value 0 with probability $1 - x$. Now assume that we have n independent random variables $X_1, \ldots, X_n$ all with the same distribution as X. Let

$$S_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$$

Then it is an easy calculation to see that

$$P\left(S_n = \frac{k}{n}\right) = {}^nC_k\, x^k (1 - x)^{n-k}$$

where $P(E)$ stands for the probability of the event E. It follows that

$$\mathbb{E}(f(S_n)) = \sum_{k=0}^{n} f\left(\frac{k}{n}\right) P\left(S_n = \frac{k}{n}\right) = B_n(f, x)$$

where $\mathbb{E}(Y)$ stands for the expectation of the random variable Y. By the law of averages, we should expect S_n to "converge to" x as n converges to ∞. Hence $\mathbb{E}(f(S_n)) = B_n(f, x)$ should converge to $f(x)$ as n tends to ∞. The above argument is imprecise, but it is possible to give a rigorous proof of the Bernstein Approximation Theorem using the Law of Large Numbers.

4.3 Uniform on Compacta Convergence

DEFINITION *Let X be a metric space and let (f_n) be a sequence of real-valued functions on X. Let $f : X \longrightarrow \mathbb{R}$. We say that (f_n) converges uniformly on compacta to f if for every (sequentially) compact subset K of X we have*

$$f_n|_K \longrightarrow f|_K$$

uniformly.

EXAMPLE We looked at the example

$$g_n(x) = \sin\left(\frac{x}{n}\right) \text{ and } g(x) = 0$$

and decided that (g_n) does not converge to g uniformly. But convergence is uniform on compacta. Every (sequentially) compact subset of $\mathbb{R}$ is bounded, so we need only show uniform convergence on every symmetric interval $[-a, a]$ for $a > 0$. But

$$\sup_{|x| \leq a} \left|\sin\left(\frac{x}{n}\right)\right| \leq \frac{a}{n} \longrightarrow 0$$

as $n \longrightarrow \infty$. □

EXAMPLE Consider again the example $f_n(x) = \dfrac{nx}{n^2x^2 + 1}$ but this time on on $]0, 1]$. It can be shown that every sequentially compact subset K of $]0, 1]$ is contained in $[\delta, 1]$ for some $\delta > 0$, for if not, there would be a sequence in K decreasing strictly to 0 and such a sequence could not have a subsequence converging to anything in $]0, 1]$.

If we *wait* until $n\delta > 1$, then f_n is decreasing on $[\delta, 1]$ and so

$$\sup_{]\delta,1]} |f_n(x)| = f_n(\delta) \longrightarrow 0$$

as $n \longrightarrow \infty$. So uniform on compacta convergence holds. □

4.4 Convergence under the Integral Sign

Let f_n be Riemann integrable functions on $[a, b]$ and suppose that f_n converges to a Riemann integrable function f on $[a, b]$. Do we have

$$\lim_{n\to\infty} \int_a^b f_n(x)dx = \int_a^b f(x)dx? \tag{4.13}$$

The first thing to say is that this is not true in general.

EXAMPLE Let

$$f_n(x) = \begin{cases} n^2 x & \text{if } 0 \le x \le \dfrac{1}{n}, \\ 2n - n^2 x & \text{if } \dfrac{1}{n} \le x \le \dfrac{2}{n}, \\ 0 & \text{if } \dfrac{2}{n} \le x \le 2. \end{cases}$$

be a sequence of functions on $[0, 2]$. It's easy to see that $f_n(x) \longrightarrow 0$ as $n \longrightarrow \infty$ for each $x \in [0, 2]$. But we also have

$$\begin{aligned} \int_0^2 f_n(x)dx &= \int_0^{n^{-1}} n^2 x dx + \int_{n^{-1}}^{2n^{-1}} (2n - n^2 x)dx + \int_{2n^{-1}}^2 0dx \\ &= \frac{1}{2} + \frac{1}{2} = 1. \end{aligned}$$

On the other hand

$$\int_0^2 f(x)dx = 0.$$

□

It is true if the convergence is uniform as the following theorem shows.

THEOREM 94 *Let f_n be Riemann integrable functions on $[a, b]$ and suppose that f_n converges uniformly to a function f on $[a, b]$. Then f is Riemann integrable on $[a, b]$ and (4.13) holds.*

Proof. First we show that f is Riemann integrable on $[a, b]$. Let $\epsilon > 0$ we will show that f satisfies Riemann's integrability condition. First, we find n such that

$$\sup_{a \le x \le b} |f(x) - f_n(x)| \le \frac{\epsilon}{3(b - a)}$$

Now f_n is integrable by hypothesis, so there is a Riemann partition P such that

$$\sum_J |J| \operatorname*{osc}_J f_n < \frac{\epsilon}{3}.$$

But

$$
\begin{aligned}
\operatorname*{osc}_J f &= \sup_{x,x'\in J} |f(x)-f(x')| \\
&\le \sup_{x,x'\in J} \Big(|f_n(x)-f_n(x')| + |f(x)-f_n(x)| + |f(x')-f_n(x')| \Big) \\
&\le \operatorname*{osc}_J f_n + \frac{2\epsilon}{3(b-a)}.
\end{aligned}
$$

Therefore

$$\sum_J |J| \operatorname*{osc}_J f \le \sum_J |J| \operatorname*{osc}_J f_n + \frac{2\epsilon}{3(b-a)} \sum_J |J| < \epsilon.$$

It follows that f is Riemann integrable.

Now for the convergence issue. We have

$$
\begin{aligned}
\left| \int_a^b f_n(x)dx - \int_a^b f(x)dx \right| &\le \int_a^b |f_n(x)-f(x)|dx \\
&\le \int_a^b \sup_{a\le t\le b} |f_n(t)-f(t)|dx \\
&\le (b-a) \sup_{a\le t\le b} |f_n(t)-f(t)|
\end{aligned}
$$

■

That was easy. Too easy. In fact, much more is true, but usually such results are proved along with the Lebesgue theory. It's unfortunate that the proofs use the Lebesgue theory in an essential way and are not accessible to us with our present knowledge. So, if you want to prove convergence under the integral sign and convergence is not uniform, you need to investigate on a case by case basis.

•4.5 The Wallis Product and Sterling's Formula

In the following saga we will assume the properties of the trig functions.

LEMMA 95 *We have*

$$\lim_{n\to\infty} \prod_{k=1}^{n} \frac{4k^2}{4k^2-1} = \frac{\pi}{2} \tag{4.14}$$

and

$$\lim_{n\to\infty} \sqrt{n} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos x)^n dx = \sqrt{2\pi}. \tag{4.15}$$

Proof. From integration by parts, one obtains for $n \geq 2$

$$I_n = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos x)^n dx = \frac{n-1}{n} \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos x)^{n-2} dx.$$

This in turn leads to the formulæ

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos x)^{2n} dx = \pi \, \frac{3\cdot 5\cdot 7\cdots(2n-1)}{2\cdot 4\cdot 6\cdots 2n}$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (\cos x)^{2n+1} dx = 2 \, \frac{2\cdot 4\cdot 6\cdots 2n}{3\cdot 5\cdot 7\cdots(2n+1)}$$

Now since cos is nonnegative and bounded above by 1 on $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, we get $I_{2n+2} \leq I_{2n+1} \leq I_{2n}$ and hence

$$\frac{2n+1}{2n+2} = \frac{I_{2n+2}}{I_{2n}} \leq \frac{I_{2n+1}}{I_{2n}} \leq 1$$

and an application of the Squeeze Lemma shows that

$$\lim_{n\to\infty} \frac{I_{2n+1}}{I_{2n}} = 1.$$

Equally well,

$$\frac{2n+1}{2n+2} = \frac{I_{2n+2}}{I_{2n}} \leq \frac{I_{2n+2}}{I_{2n+1}} \leq 1$$

and an application of the Squeeze Lemma shows that

$$\lim_{n\to\infty} \frac{I_{2n+2}}{I_{2n+1}} = 1.$$

So, actually $\lim_{n\to\infty} \frac{I_{n+1}}{I_n} = 1$.

Now we have

$$(2n+1) I_{2n} I_{2n+1} = 2\pi$$

and one usually deduces from this that

$$\prod_{k=1}^{n} \frac{4k^2}{4k^2-1} = \frac{2}{1}\cdot\frac{2}{3}\cdot\frac{4}{3}\cdot\frac{4}{5}\cdot\frac{6}{5}\cdots\frac{2n}{2n-1}\cdot\frac{2n}{2n+1}$$
$$= \left(\frac{2}{1}\cdot\frac{4}{3}\cdot\frac{6}{5}\cdots\frac{2n}{2n-1}\right)\cdot\left(\frac{2}{3}\cdot\frac{4}{5}\cdot\frac{6}{7}\cdots\frac{2n}{2n+1}\right)$$
$$= \left(\frac{\pi}{I_{2n}}\right)\cdot\left(\frac{I_{2n+1}}{2}\right) \underset{n\to\infty}{\longrightarrow} \frac{\pi}{2},$$

which is the famous Wallis product and equivalent to (4.14).

However, our interest is the fact that (4.15) holds. ■

Now suppose that $r = \sup_{a\le x\le b} |f(x)| < 1$. Then it is clear that

$$\lim_{n\to\infty} \sqrt{n} \int_a^b \Big(f(x)\Big)^n dx = 0. \tag{4.16}$$

This is just a consequence of $\lim_{n\to\infty} \sqrt{n}\, r^n = 0$. Combining this with (4.15) we see that for every $0 < \delta < \dfrac{\pi}{2}$ we have

$$\lim_{n\to\infty} \sqrt{n} \int_{-\delta}^{\delta} (\cos x)^n dx = \sqrt{2\pi}.$$

PROPOSITION 96 *Let f be a twice continuously differentiable function defined on $[-1,1]$ with $f(0) = 1$ and $|f(x)| < 1$ on $[-1,1]\setminus\{0\}$. Since f has a local maximum at 0 it will necessarily be the case that $f'(0) = 0$. Assume also that $f''(0) = -1$. Then*

$$\lim_{n\to\infty} \sqrt{n} \int_{-1}^{1} \Big(f(x)\Big)^n dx = \sqrt{2\pi}. \tag{4.17}$$

For instance, $f(x) = \cos x$ satisfies the properties listed in Proposition 96. The idea of the proof is to transfer the known properties of cos to the function f.

Proof. Let $\lambda > 1 > \mu$. We claim that there exists $\delta > 0$ such that

$$\cos(\lambda x) \le f(x) \le \cos(\mu x) \tag{4.18}$$

for $|x| < \delta$. To prove this consider $g(x) = \dfrac{f(x)}{\cos(\lambda x)}$ and show that $g(0) = 1$, $g'(0) = 0$ and $g''(0) = \lambda - 1 > 0$. Since g'' is continuous, it follows that there exists $\delta_1 > 0$ such that $g''(\xi) > 0$ for $|\xi| < \delta_1$. But now we apply Taylor's Theorem with the Lagrange remainder to get $g(x) = 1 + \dfrac{1}{2}g''(\xi)x^2 > 1$ whenever $|x| < \delta_1$ (and hence $|\xi| < \delta_1$). Similarly, there exists $\delta_2 > 0$ such that $\dfrac{f(x)}{\cos(\mu x)} < 1$ for $|x| < \delta_2$.

From (4.18), we get

$$\begin{aligned}
\sqrt{\frac{2\pi}{\lambda}} &= \liminf_{n\to\infty} \sqrt{n} \int_{-\delta}^{\delta} \Big(\cos(\lambda x)\Big)^n dx \\
&\le \liminf_{n\to\infty} \sqrt{n} \int_{-\delta}^{\delta} \Big(f(x)\Big)^n dx \\
&\le \limsup_{n\to\infty} \sqrt{n} \int_{-\delta}^{\delta} \Big(f(x)\Big)^n dx \\
&\le \limsup_{n\to\infty} \sqrt{n} \int_{-\delta}^{\delta} \Big(\cos(\mu x)\Big)^n dx \\
&= \sqrt{\frac{2\pi}{\mu}}
\end{aligned}$$

Combining this with (4.16), we deduce that

$$\sqrt{\frac{2\pi}{\lambda}} \le \liminf_{n\to\infty} \sqrt{n} \int_{-1}^{1} \Big(f(x)\Big)^n dx \le \limsup_{n\to\infty} \sqrt{n} \int_{-1}^{1} \Big(f(x)\Big)^n dx \le \sqrt{\frac{2\pi}{\mu}}$$

and again, since λ and μ are arbitrary satisfying $\lambda > 1 > \mu$ that the claim (4.17) holds. ■

THEOREM 97 (STERLING'S FORMULA)

$$\lim_{n\to\infty} \frac{n!}{n^{n+\frac{1}{2}}e^{-n}} = \sqrt{2\pi}$$

A much more precise statement will be given later (6.5).

Proof. Now, we have by an easy induction

$$\begin{aligned} n! &= \int_0^\infty x^n e^{-x} dx \\ &= \int_0^\infty (nx)^n e^{-nx} n dx \\ &= n^{n+1} \int_0^\infty \left(x e^{-x} \right)^n dx \\ &= n^{n+1} \int_{-1}^\infty \left((x+1) e^{-(x+1)} \right)^n dx \\ &= n^{n+1} e^{-n} \int_{-1}^\infty \left((x+1) e^{-x} \right)^n dx \\ &= n^{n+\frac{1}{2}} e^{-n} \left(\sqrt{n} \int_{-1}^\infty \left((x+1) e^{-x} \right)^n dx \right) \end{aligned}$$

Now the function $f(x) = (x+1)e^{-x}$ has the properties of Proposition 96. We also remark that

$$f(x) = (x+1)e^{-x} \le 2e^{-\frac{1+x}{2}} \qquad \text{for } x \ge 1$$

so that

$$\sqrt{n} \int_1^\infty \left((x+1)e^{-x} \right)^n dx \le \sqrt{n} \int_1^\infty \left(2e^{-\frac{1+x}{2}} \right)^n dx = \frac{2}{\sqrt{n}} \left(\frac{2}{e} \right)^n \longrightarrow 0$$

as $n \longrightarrow \infty$. so we find

$$\lim_{n\to\infty} \frac{n!}{n^{n+\frac{1}{2}} e^{-n}} = \lim_{n\to\infty} \sqrt{n} \int_{-1}^1 \left((x+1)e^{-x} \right)^n dx = \sqrt{2\pi}$$

as required. ■

EXAMPLE Consider again the series $\displaystyle\sum_{n=1}^\infty \frac{(2n)!}{(n!)^2 4^n}$ which we originally handled using Raabe's test. We now have

$$\frac{(2n)!}{(n!)^2 4^n} \sim \frac{\sqrt{2\pi}\, (2n)^{2n+\frac{1}{2}}\, e^{-2n}}{2\pi\, n^{2n+1}\, e^{-2n}\, 4^n} \sim \frac{1}{\sqrt{\pi n}}$$

and the series diverges in comparison with $\sum_{n=1}^\infty n^{-\frac{1}{2}}$. With Sterling's formula we see exactly how big the terms really are. □

4.6 Uniform Convergence and the Cauchy Condition

PROPOSITION 98 *The normed vector space $B(X)$ of all bounded real-valued functions on a set X with the supremum norm is complete.*

Proof. The pattern of most completeness proofs is the same. Take a Cauchy sequence. Use some existing completeness information to deduce that the sequence converges in some weak sense. Use the Cauchy condition again to establish that the sequence converges in the metric sense.

Let (f_n) be a Cauchy sequence in $B(X)$. Then, for each $x \in X$, it is straightforward to check that $(f_n(x))$ is a Cauchy sequence in $\mathbb{R}$ and hence converges to some element of $\mathbb{R}$. This can be viewed as a rule for assigning an element of $\mathbb{R}$ to every element of X — in other words, a function f from X to $\mathbb{R}$. We have just shown that (f_n) converges to f pointwise.

Now let $\epsilon > 0$. Then for each $x \in X$ there exists $N_x \in \mathbb{N}$ such that

$$q > N_x \qquad \Rightarrow \qquad |f_q(x) - f(x)| < \tfrac{1}{3}\epsilon. \tag{4.19}$$

Now we reuse the Cauchy condition — there exists $N \in \mathbb{N}$ such that

$$p, q > N \qquad \Rightarrow \qquad \sup_{x \in X} |f_p(x) - f_q(x)| < \tfrac{1}{3}\epsilon. \tag{4.20}$$

Now, combining (4.19) and (4.20) with the triangle inequality and choosing q explicitly as $q = \max(N, N_x) + 1$, we find that

$$p > N \qquad \Rightarrow \qquad |f_p(x) - f(x)| < \tfrac{2}{3}\epsilon \quad \forall x \in X. \tag{4.21}$$

We emphasize the crucial point that N depends only on ϵ. It does not depend on x. Thus we may deduce

$$p > N \qquad \Rightarrow \qquad \sup_{x \in X} |f_p(x) - f(x)| < \epsilon. \tag{4.22}$$

from (4.21).

This would be the end of the proof, if it were not for the fact that we still do not know that $f \in B(X)$. For this, choose an explicit value of ϵ, say $\epsilon = 1$. Then, using the corresponding specialization of (4.22), we see that there exists $r \in \mathbb{N}$ such that

$$\sup_{x \in X} |f_r(x) - f(x)| < 1. \tag{4.23}$$

Now, use (4.23) to obtain

$$\sup_{x \in X} |f(x)| \leq \sup_{x \in X} |f_r(x)| + \sup_{x \in X} |f_r(x) - f(x)|$$

It now follows that since f_r is bounded, so is f. Finally, with the knowledge that $f \in B(X)$ we see that (f_n) converges to f in $B(X)$ by (4.22). ■

There is an alternative way of deducing (4.22) from (4.20) which worth mentioning. Conceptually it is simpler than the argument presented above, but perhaps less rigorous. We write (4.20) in the form

$$p, q > N \qquad \Rightarrow \qquad |f_p(x) - f_q(x)| < \tfrac{1}{3}\epsilon. \tag{4.24}$$

where x is a general point of X. The vital key is that N depends only on ϵ and not on x. Now, letting $q \longrightarrow \infty$ in (4.24) we find

$$p > N \qquad \Rightarrow \qquad |f_p(x) - f(x)| \leq \tfrac{1}{3}\epsilon. \tag{4.25}$$

because $f_q(x)$ converges pointwise to $f(x)$. Here we are using the fact that $[0, \frac{1}{3}\epsilon]$ is a closed subset of $\mathbb{R}$. Since N depends only on ϵ we can then deduce (4.22) from (4.25).

The result also extends in an informal sense to real-valued functions that are not bounded.

COROLLARY 99 *Let (f_n) be a sequence of functions on X that satisfies the Cauchy condition that for all $\epsilon > 0$, there exists N such that*

$$p, q > N \qquad \Rightarrow \qquad \sup_{x \in X} |f_p(x) - f_q(x)| < \epsilon.$$

Then there is a real-valued function f on X such that (f_n) converges to f uniformly on X.

Proof. Choose $\epsilon = 1$. Then there exists M such that

$$p, q \geq M \qquad \Rightarrow \qquad \sup_{x \in X} |f_p(x) - f_q(x)| < 1.$$

Now define $g_n = f_n - f_M$. Then g_n is a bounded function for $n \geq M$ and moreover $(g_n)_{n=M}^{\infty}$ is a Cauchy sequence in $B(X)$. So, g_n must converge to some function g in $B(X)$. It now follows easily that (f_n) converges uniformly to $f_M + g$. ■

There is a canned version of this result that applies to series.

THEOREM 100 (WEIERSTRASS M-TEST) *Let $M_n = \sup_{x \in X} |a_n(x)|$. Suppose that $\sum_{n=1}^{\infty} M_n < \infty$, then the series of functions $\sum_{n=1}^{\infty} a_n(x)$ converges uniformly on X.*

Proof. Let $s_p(x) = \sum_{n=1}^{p} a_n(x)$. Then for $p \geq q$ we have

$$\begin{aligned}
\sup_{x \in X} |s_p(x) - s_q(x)| &= \sup_{x \in X} \left| \sum_{n=q+1}^{p} a_n(x) \right| \\
&\leq \sup_{x \in X} \sum_{n=q+1}^{p} |a_n(x)| \\
&\leq \sum_{n=q+1}^{p} \sup_{x \in X} |a_n(x)| \\
&= \sum_{n=q+1}^{p} M_n
\end{aligned}$$

It follows from $\sum_{n=1}^{\infty} M_n < \infty$ that (s_n) is a uniform cauchy sequence on X. Hence (s_n) converges uniformly to some limit. But, we know that the limit is $\sum_{n=1}^{\infty} a_n(x)$, because after all this series converges (absolutely) pointwise. So the series also converges uniformly. ■

It is important to realize that the M-test is a sufficient condition for a series to converge uniformly. It is not a necessary condition.

EXAMPLE Recall the series

$$\sum_{n=1}^{\infty} \frac{1}{2n-1} \sin(2n-1)t$$

that was discussed in the section on summation by parts. The proof given there can be extended to show that convergence is uniform on the compact subsets of $]0, \pi[$. Let us just use the fact that it is uniform on $\left[\frac{\pi}{3}, \frac{2\pi}{3}\right]$. Now, for each $n \in \mathbb{N}$, $\sin(2n-1)t = \pm 1$ for some $t \in \left[\frac{\pi}{3}, \frac{2\pi}{3}\right]$, in fact for $t = \frac{\pi}{2}$. It follows that

$$M_n = \sup_{\frac{\pi}{3} \leq t \leq \frac{2\pi}{3}} \left| \frac{1}{2n-1} \sin(2n-1)t \right| = \frac{1}{2n-1},$$

and the M-test fails since $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2n-1} = \infty$. □

4.7 Differentiation and Uniform Convergence

We now prove the theorem that links together uniform convergence and derivatives. The proof of this theorem is really very subtle. This theorem is vital for understanding power series.

THEOREM 101 *Let $-\infty < a < c < b < \infty$ and let f_n be a sequence of differentiable functions on $]a, b[$. We suppose that*

- *$(f_n(c))$ is a convergent sequence of real numbers.*
- *(f'_n) converges uniformly to a function g on $]a, b[$.*

Then (f_n) converges uniformly to a function f on $]a, b[$. Furthermore f is differentiable on $]a, b[$ and $f' = g$.

Proof. The first step is to apply the Mean Value Theorem to the function $f_p - f_q$. This gives

$$\Big(f_p(x) - f_q(x)\Big) - \Big(f_p(c) - f_q(c)\Big) = \Big(f'_p(\xi) - f'_q(\xi)\Big)(x - c)$$

and so

$$\sup_{a<x<b} |f_p(x) - f_q(x)| \le |f_p(c) - f_q(c)| + (b - a) \sup_{a<\xi<b} \left|f'_p(\xi) - f'_q(\xi)\right|$$

The hypotheses show that (f_n) is a uniform Cauchy sequence on $]a, b[$, and therefore it converges to some function f.

Now for the tricky part. Let $\epsilon > 0$. Then there exists N such that $n \ge N$ implies that

$$\sup_{a<x<b} |f'_n(x) - g(x)| < \epsilon \tag{4.26}$$

Now since f_N is differentiable at x, there exists $\delta > 0$ (depending on ϵ and x) such that $0 < |h| < \delta$ forces both $a < x + h < b$ and

$$\left|\frac{f_N(x+h) - f_N(x)}{h} - f'_N(x)\right| < \epsilon.$$

Let $p, q \geq N$ and apply once again the Mean Value Theorem to $f_p - f_q$. We get

$$\frac{(f_p - f_q)(x+h) - (f_p - f_q)(x)}{h} = (f_p' - f_q')(\xi)$$

so that

$$\left| \frac{(f_p - f_q)(x+h) - (f_p - f_q)(x)}{h} \right| < 2\epsilon$$

by (4.26). Now, let $p \longrightarrow \infty$ and put $q = N$. We find

$$\left| \frac{(f - f_N)(x+h) - (f - f_N)(x)}{h} \right| \leq 2\epsilon$$

Now we get

$$\begin{aligned} &\left| \frac{f(x+h) - f(x)}{h} - g(x) \right| \\ &\leq \left| \frac{f_N(x+h) - f_N(x)}{h} - f_N'(x) \right| \\ &\quad + \left| \frac{(f - f_N)(x+h) - (f - f_N)(x)}{h} \right| + |f_N'(x) - g(x)| \\ &< \epsilon + 2\epsilon + \epsilon = 4\epsilon \end{aligned}$$

always provided $0 < |h| < \delta$. This shows that $f'(x)$ exists and equals $g(x)$ for every x in $]a, b[$. ■

EXAMPLE Consider the sequence of functions $(f_n)_{n=1}^{\infty}$ is defined on $[-1, 3]$ by

$$f_n(x) = \begin{cases} 0 & \text{if } -1 \leq x \leq 0, \\ n^2 x^3 & \text{if } 0 \leq x \leq \dfrac{1}{n}, \\ -2x + 4nx^2 - n^2 x^3 & \text{if } \dfrac{1}{n} \leq x \leq \dfrac{2}{n}, \\ 2x & \text{if } \dfrac{2}{n} \leq x \leq 3. \end{cases}$$

Let also

$$f(x) = \begin{cases} 0 & \text{if } -1 \leq x \leq 0, \\ 2x & \text{if } 0 \leq x \leq 3. \end{cases}$$

We can write $f(x) - f_n(x) = xg(nx)$ where in fact g is a bounded continuous function on $\mathbb{R}$ which vanishes on $]-\infty, 0] \bigcup [2, \infty[$. Also $f(x) - f_n(x) = 0$ unless $0 \le x \le \dfrac{2}{n}$. Hence

$$\sup\{|f(x) - f_n(x)|; -1 \le x \le 3\} \le \frac{2}{n} \sup\{|g(u)|; u \in \mathbb{R}\}$$

and it follows that (f_n) converges to f uniformly on $[-1, 3]$. Now we check that f_n is differentiable on $[-1, 3]$. We do this by elementary calculus except at the points $x = 0$, $x = \dfrac{1}{n}$ and $x = \dfrac{2}{n}$. For the exceptional points, it's enough to check that the left and right derivatives coincide. We get

$$f_n'(x) = \begin{cases} 0 & \text{if } -1 \le x \le 0, \\ 3n^2x^2 & \text{if } 0 \le x \le \dfrac{1}{n}, \\ -2 + 8nx - 3n^2x^2 & \text{if } \dfrac{1}{n} \le x \le \dfrac{2}{n}, \\ 2 & \text{if } \dfrac{2}{n} \le x \le 3. \end{cases}$$

For $-1 \le x \le 0$, we have $f_n'(x) = 0$ for all n so $\lim_{n\to\infty} f_n'(x) = 0$. Now let $x > 0$, then for n sufficiently large, we are in the fourth option. Thus $f_n'(x) = 2$ for $n \ge \left\lceil \dfrac{2}{x} \right\rceil$. We find $\lim_{n\to\infty} f_n'(x) = 2$ for $x > 0$. Thus, the derivatives f_n' converge pointwise on $[-1, 3]$. It's easy to check that the derivatives f_n' are continuous on $[-1, 3]$, but the pointwise limit of (f_n') is not. So, convergence cannot be uniform on $[-1, 3]$. The limit function f is not differentiable on $[-1, 3]$, showing that in Theorem 101, you cannot replace the uniform of convergence of the (f_n') with pointwise convergence. □

5

Power Series

In this chapter, we study power series, that is, series of the form

$$a_0 + a_1(x-\alpha) + a_2(x-\alpha)^2 + a_3(x-\alpha)^3 + \cdots$$

This is described as a power series about $x = \alpha$, or a series in powers of $x - \alpha$. Usually, it is no restriction to take $\alpha = 0$ and we will usually work with this case. Power series are important because they allow us to define the standard transcendental functions. They are also a powerful tool in numerical analysis.

For those with an algebraic bent of mind, the set of all ***formal power series*** of the form

$$a_0 + a_1x + a_2x^2 + a_3x^3 + \cdots$$

where $a_0, a_1, a_2, a_3, \ldots$ are *any* real numbers form a ring which is usually denoted $\mathbb{R}[[x]]$. The addition and multiplication in the ring are defined *formally*, this means that if

$$(c_0 + c_1x + c_2x^2 + c_3x^3 + \cdots) = (a_0 + a_1x + a_2x^2 + \cdots)(b_0 + b_1x + b_2x^2 + \cdots)$$

then each coefficient c_n is obtained by a finite calculation from the a's and b's. For example

$$c_2 = a_0b_2 + a_1b_1 + a_2b_0,$$

the convergence of the series is not considered and not needed. Similarly, one can formally differentiate a power series, or in case $a_0 \neq 0$, formally invert it. The use of the word *formal* in the sequel means that the operation is to be carried out in this sense.

5.1 Convergence of Power Series

So, now we are back to the world of analysis, we need to know where a power series converges. The answer is supplied by the following theorem.

THEOREM 102 *There is a "number" $\rho \in [0, \infty]$ such that the series*

$$a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \tag{5.1}$$

converges if $|x| < \rho$ and does not converge if $|x| > \rho$.

There are two extreme cases. In the case $\rho = 0$, the series converges only if $x = 0$. In this case, the series is for all intents and purposes useless. The other extreme case is when $\rho = \infty$ and then the series converges for all real x. This theorem is so important that we give two proofs, the first a simple-minded one and the second uses the root test and actually produces a formula for ρ. The number ρ is called the ***radius of convergence*** .

First proof.
The series always converges for $x = 0$, so let us define

$$\rho = \sup\{|x|; \text{Series (5.1) converges}\},$$

with the understanding that $\rho = \infty$ in case the set is unbounded above. Then, by definition, $|x| > \rho$ implies that (5.1) does not converge. It remains to show that $|x| < \rho$ implies that (5.1) converges. In this case, there exists t with $|x| < |t|$ such that $a_0 + a_1 t + a_2 t^2 + a_3 t^3 + \cdots$ converges, for otherwise $|x|$ would be an upper bound for the set over which the sup was taken. Therefore $a_n t^n \longrightarrow 0$ as $n \longrightarrow \infty$. So, there is a constant C such that $|a_n t^n| \le C$ for all $n \in \mathbb{Z}^+$. But, now $\sum_{n=1}^{\infty} a_n x^n$ converges absolutely by comparison with a geometric series

$$\sum_{n=0}^{\infty} |a_n x^n| \le \sum_{n=1}^{\infty} |a_n t^n| \left|\frac{x}{t}\right|^n \le \sum_{n=0}^{\infty} C \left|\frac{x}{t}\right|^n < \infty$$

■

Note that the proof actually shows that (5.1) converges *absolutely* for $|x| < \rho$.

Second proof. We apply the root test. Series (5.1) converges absolutely if $\limsup_{n\to\infty} |a_n x^n|^{\frac{1}{n}} < 1$. and does not converge if $\limsup_{n\to\infty} |a_n x^n|^{\frac{1}{n}} > 1$, because the *proof* of the root test shows that the terms do not tend to zero. This gives the formula

$$\rho = \liminf_{n\to\infty} |a_n|^{-\frac{1}{n}}$$

which has to be interpreted by taking $|a_n|^{-\frac{1}{n}} = \infty$ if $a_n = 0$ and $\rho = \infty$ if $a_n = 0$ eventually, or if $\inf_{n \geq N} |a_n|^{-\frac{1}{n}}$ tends properly to ∞ as $N \longrightarrow \infty$.

EXAMPLE Some fairly simple examples show that anything can happen at $\pm\rho$. The following series all have radius 1. We find that $\sum_{n=0}^{\infty} x^n$ does not converge at either 1 or -1. The series $\sum_{n=0}^{\infty} \frac{1}{n} x^n$ converges at -1, but not at 1. The series $\sum_{n=0}^{\infty} (-1)^n \frac{1}{n} x^n$ converges at 1, but not at -1. The series $\sum_{n=0}^{\infty} \frac{1}{n^2} x^n$ converges at both -1 and 1. □

EXAMPLE Consider $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n!} x^n$. Then, applying the ratio test, the series absolutely converges if

$$\lim_{n\to\infty} \left| \frac{n!}{(n+1)!} \frac{x^{n+1}}{x^n} \right| < 1.$$

Since the limit is zero for all x, the radius of convergence is infinite. Of course, this series defines the exponential function. □

EXAMPLE Consider $\displaystyle\sum_{n=0}^{\infty} n! x^n$. Then, it is clear that the terms of the series do not tend to zero unless $x = 0$. This series has zero radius of convergence and is totally useless. □

EXAMPLE The series $1 + 2x + 2x^2 + 2x^3 + 2x^4 + \cdots$ gives $\displaystyle\lim_{n\to\infty} \left| \frac{a_{n+1} x^{n+1}}{a_n x^n} \right| = |x|$, so that the series converges if $|x| < 1$ and diverges if $|x| > 1$. Hence the radius of convergence is 1. Actually, after the first term this series is geometric and it therefore converges to

$$1 + 2x \cdot \frac{1}{1-x} = \frac{1+x}{1-x}$$

for $|x| < 1$. □

PROPOSITION 103 *Let (5.1) have radius $\rho > 0$. Then (5.1) converges uniformly on the compact subsets of* $]-\rho, \rho[$.

Proof. Let $0 < r < \rho$. Then, as in either proof of Theorem 102, $\sum_{n=0}^{\infty} |a_n| r^n < \infty$. So $\sum_{n=0}^{\infty} \left\{ \sup_{|x| \leq r} |a_n x^n| \right\} < \infty$, and $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on

$[-r, r]$ by the M-test (Theorem 100). ■

COROLLARY 104 *Let $f(x) = \sum_{n=0}^{\infty} a_n x^n$. Then f is continuous on $]-\rho, \rho[$.*

Proof. Let $0 < r < \rho$. Then the series converges uniformly on $[-r, r]$ and hence f is continuous on $[-r, r]$. Therefore f is continuous on $]-\rho, \rho[$. ■

It is a remarkable fact that if a series converges at ρ (respectively at $-\rho$) then it converges uniformly on $[0, \rho]$ (respectively $[-\rho, 0]$). Showing this boils down to the following theorem, due to Abel.

THEOREM 105 *Let $\sum_{n=0}^{\infty} a_n$ be a convergent series. Then $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[0, 1]$.*

Proof. The idea is to use summation by parts. However, the summation by parts formula that we proved earlier does not cut the mustard. The trick is to devise a formula that uses the *tail sums* rather than the partial sums. Let

$$r_N = \sum_{n=N}^{\infty} a_n.$$

Then,

$$\begin{aligned}
\sum_{n=p}^{q} a_n x^n &= \sum_{n=p}^{q} (r_n - r_{n+1}) x^n \\
&= \sum_{n=p}^{q} r_n x^n - \sum_{n=p}^{q} r_{n+1} x^n \\
&= \sum_{n=p}^{q} r_n x^n - \sum_{n=p+1}^{q+1} r_n x^{n-1} \\
&= r_p x^p - r_{q+1} x^q + \sum_{n=p+1}^{q} r_n x^n - \sum_{n=p+1}^{q} r_n x^{n-1} \\
&= r_p x^p - r_{q+1} x^q + \sum_{n=p+1}^{q} r_n (x^n - x^{n-1})
\end{aligned}$$

so that,

$$\left|\sum_{n=p}^{q} a_n x^n\right| \le |r_p x^p| + |r_{q+1} x^q| + \left|\sum_{n=p+1}^{q} r_n (x^n - x^{n-1})\right|$$

$$\le |r_p| + |r_{q+1}| + \sum_{n=p+1}^{q} |r_n| (x^{n-1} - x^n)$$

because $x^{n-1} - x^n \ge 0$ for $0 \le x \le 1$,

$$\le \left\{\sup_{n \ge p} |r_n|\right\} \left\{1 + 1 + \sum_{n=p+1}^{q} (x^{n-1} - x^n)\right\}$$

$$\le 3 \sup_{n \ge p} |r_n|,$$

since the series $\sum_{n=p+1}^{q} (x^{n-1} - x^n)$ telescopes to $x^p - x^q \le 1$ for $0 \le x \le 1$. But since $r_n \longrightarrow 0$ as $n \longrightarrow \infty$, we see that $s_p(x) = \sum_{n=0}^{p} a_n x^n$ is a uniform Cauchy sequence on $[0, 1]$. It follows that $\sum_{n=0}^{\infty} a_n x^n$ converges uniformly on $[0, 1]$. ■

There's actually something else that one can prove here with the same idea.

THEOREM 106 *Let $\sum_{n=0}^{\infty} a_n$ be a convergent series and (x_n) a positive decreasing sequence. Then $\sum_{n=0}^{\infty} a_n x_n$ converges.*

We leave the proof to the reader.

5.2 Manipulation of Power Series

In this section we will assume that

$$f(x) = a_0 + a_1 x + a_2 x^2 + a_3 x^3 + \cdots \tag{5.2}$$

with radius r and for $|x| < r$ and that

$$g(x) = b_0 + b_1 x + b_2 x^2 + b_3 x^3 + \cdots$$

with radius s and for $|x| < s$. We already know from general principles the following result.

PROPOSITION 107 *The series $\sum_{n=0}^{\infty} (\lambda a_n + \mu b_n) x^n$ has radius at least $\min(r, s)$ and it converges to $\lambda f(x) + \mu g(x)$ for $|x| < \min(r, s)$.*

It is easy to find examples where the radius of $\sum_{n=0}^{\infty} (\lambda a_n + \mu b_n) x^n$ is strictly larger than $\min(r, s)$.

Perhaps the most important result concerns differentiation.

THEOREM 108 *The function f is differentiable on $]-r, r[$. Further the formally differentiated series*

$$\sum_{n=0}^{\infty} a_n n x^{n-1} = \sum_{n=1}^{\infty} a_n n x^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} x^n$$

has radius r and converges to $f'(x)$ for $|x| < r$.

Proof. We start with the formally differentiated series. Multiplying this by x does not affect the radius of convergence, so let us find the radius of

$$\sum_{n=0}^{\infty} a_n n x^n.$$

By the root test this is

$$\liminf_{n\to\infty} |n a_n|^{-\frac{1}{n}}.$$

But $\lim_{n\to\infty} n^{-\frac{1}{n}} = 1$, so we have

$$\liminf_{n\to\infty} |n a_n|^{-\frac{1}{n}} = \liminf_{n\to\infty} |a_n|^{-\frac{1}{n}} = r.$$

This shows that the formally differentiated series has radius r.

Now we apply Theorem 101 to the interval $]-\rho, \rho[$ for $\rho < r$. The original series converges at 0. The differentiated series converges uniformly on $]-\rho, \rho[$. So, f is differentiable on $]-\rho, \rho[$ and the sum of the formally differentiated series is $f'(x)$ for $|x| < \rho$. But this is true for every $\rho < r$. Hence the result. ■

This result is very important because it can be iterated. So in fact, f is infinitely differentiable on $]-r, r[$ and we have a power series expansion for the kth derivative.

$$f^{(k)}(x) = \sum_{n=0}^{\infty} (n+k)(n+k-1)\cdots(n+1) a_{n+k} x^n = k! \sum_{n=0}^{\infty} {}^{n+k}C_n \, a_{n+k} x^n$$

Substituting $x = 0$ into this formula gives $f^{(k)}(0) = k! a_k$ or

$$a_k = \frac{f^{(k)}(0)}{k!}. \tag{5.3}$$

This is very important because it shows that the function f determines the coefficients a_k for all $k = 0, 1, 2, \ldots$ You cannot have two different functions with the same power series.

COROLLARY 109 *If the power series $\sum_{n=0}^{\infty} a_n x^n$ converges to zero for $|x| < r$ and $r > 0$ then $a_n = 0$ for all $n \in \mathbb{Z}^+$.*

COROLLARY 110 *If the power series $\sum_{n=0}^{\infty} a_n x^n$ converges to $f(x)$ for $|x| < r$ and $r > 0$ then f is an infinitely differentiable function on $]-r, r[$.*

EXAMPLE The converse of Corollary 110 is false. There are infinitely differentiable functions which do not have a power series expansion with strictly positive radius. Consider

$$f(x) = \begin{cases} \exp\left(-x^{-2}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

Then it can be shown by induction on n that there is a polynomial function p_n such that

$$f^{(n)}(x) = \begin{cases} p_n\left(x^{-1}\right) \exp\left(-x^{-2}\right) & \text{if } x \neq 0, \\ 0 & \text{if } x = 0. \end{cases}$$

so that, in fact, f is infinitely differentiable. Since the derivatives of f of all order vanish at 0, it follows from (5.3) that if $f(x) = \sum_{n=0}^{\infty} a_n x^n$ for in a neighbourhood of 0, then $a_n = 0$ for all $n = 0, 1, 2, \ldots$. But that would mean that f is identically zero which is not the case. □

It's also true that we can formally integrate power series.

THEOREM 111 *The formally integrated series*

$$\sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n} x^n \tag{5.4}$$

has radius r and converges to $\int_0^x f(t)dt$ for $|x| < r$.

Proof. Well of course, formally differentiating the formally integrated series (5.4) gets us back to the original series. So, by the proof of Theorem 108 we see that they have the same radius of convergence. So, (5.4) has radius r. Now, fix a with $0 < a < r$. Then the series (5.2) converges uniformly on $[-a, a]$. It follows from the Theorem 4.13 that

$$\lim_{N\to\infty} \int_0^x \sum_{n=0}^{N} a_n t^n dt = \int_0^x \lim_{N\to\infty} \sum_{n=0}^{N} a_n t^n dt$$

for $|x| \le a$. But this says that

$$\sum_{n=0}^{\infty} \frac{a_n}{n+1} x^{n+1} = \int_0^x \sum_{n=0}^{\infty} a_n t^n dt = \int_0^x f(t) dt$$

for $|x| \le a$. But since we can a as close as we like to r, we have that (5.4) holds for all x such that $|x| < r$. ■

THEOREM 112 *The formal product series $\sum_{n=0}^{\infty} c_n x^n$ has radius of convergence at least $\min(r,s)$ and it converges to $f(x)g(x)$ for $|x| < \min(r,s)$. Explicit formulæ for c_n are given by*

$$c_n = \sum_{p=0}^{n} a_p b_{n-p} = \sum_{q=0}^{n} a_{n-q} b_q,$$

and furthermore

$$\sum_{n=0}^{\infty} |c_n| t^n \le \left\{ \sum_{p=0}^{\infty} |a_p| t^p \right\} \left\{ \sum_{q=0}^{\infty} |b_q| t^q \right\} \tag{5.5}$$

for $0 \le t < \min(r,s)$.

Proof. Let $0 \le |x| \le t < \min(r,s)$ and $\epsilon > 0$. We denote by

$$f_N(x) = \sum_{p=0}^{N} a_p x^p$$

$$g_N(x) = \sum_{q=0}^{N} b_q x^q$$

$$h_N(x) = \sum_{n=0}^{N} c_n x^n$$

then, a tricky calculation shows that

$$f_N(x) g_N(x) - h_N(x) = \sum_{0 \le p,q \le N} a_p b_q x^{p+q} - \sum_{\substack{0 \le p,q \\ p+q \le N}} a_p b_q x^{p+q} = \sum_{\substack{0 \le p,q \le N \\ p+q > N}} a_p b_q x^{p+q}.$$

This gives

$$|f_N(x)g_N(x) - h_N(x)| \leq \sum_{\substack{0\leq p,q\leq N\\ p+q>N}} |a_p||b_q|t^{p+q}$$

and, since $p+q>N$ implies $p>N/2$ or $q>N/2$

$$\begin{aligned}
&\leq \sum_{\substack{0\leq p\leq N\\ N/2<q\leq N}} |a_p||b_q|t^{p+q} + \sum_{\substack{0\leq q\leq N\\ N/2<p\leq N}} |a_p||b_q|t^{p+q}\\
&\leq \left(\sum_{p=0}^{N} |a_p|t^p\right) \sum_{q=\lceil \frac{N}{2}\rceil}^{N} |b_q|t^q + \left(\sum_{q=0}^{N} |b_q|t^q\right) \sum_{p=\lceil \frac{N}{2}\rceil}^{N} |a_p|t^p\\
&\leq \left(\sum_{p=0}^{\infty} |a_p|t^p\right) \sum_{q=\lceil \frac{N}{2}\rceil}^{\infty} |b_q|t^q + \left(\sum_{q=0}^{\infty} |b_q|t^q\right) \sum_{p=\lceil \frac{N}{2}\rceil}^{\infty} |a_p|t^p\\
&< \epsilon
\end{aligned}$$

if N is large enough. But, on the other hand, we also have

$$|f(x)g(x) - f_N(x)g_N(x)| < \epsilon$$

if N is large enough. Therefore, for N large enough

$$|f(x)g(x) - h_N(x)| < 2\epsilon.$$

Since ϵ is an arbitrary positive number this shows that the partial sums (h_N) of the formal product series converge to the product of the sums of the given series. This holds for all x with $|x| < t$, but since t may approach $\min(r, s)$, it holds for all x with $|x| < \min(r, s)$. So the radius of convergence of the formal product series is at least $\min(r, s)$.

To show (5.5) we remark that

$$\sum_{n=0}^{N} |c_n|t^n \leq \sum_{\substack{0\leq p,q\\ p+q\leq N}} |a_p||b_q|t^{p+q} \leq \sum_{0\leq p,q\leq N} |a_p||b_q|t^{p+q} = \left\{\sum_{p=0}^{N} |a_p|t^p\right\}\left\{\sum_{q=0}^{N} |b_q|t^q\right\}$$

and let N tend to ∞. ∎

EXAMPLE The radius of convergence of a product series can exceed the minimum of the individual radii. To see this, take

$$\frac{1+x}{1-x} = 1 + 2x + 2x^2 + 2x^3 + 2x^4 + \cdots$$

and

$$\frac{1-x}{1+x} = 1 - 2x + 2x^2 - 2x^3 + 2x^4 + \cdots$$

both of which have radius 1. But, not surprisingly, the product series is

$$1 = 1 + 0x + 0x^2 + 0x^3 + 0x^4 + \cdots$$

which has infinite radius. □

COROLLARY 113 *Let* $K \in \mathbb{Z}^+$. *The formal* K*-fold product series* $\sum_{n=0}^{\infty} c_{K,n} x^n$ *of* $\sum_{n=0}^{\infty} a_n x^n$ *has radius at least* r *and converges to* $\Big(f(x)\Big)^K$. *Furthermore*

$$\sum_{n=0}^{\infty} |c_{K,n}| t^n \le \left\{ \sum_{n=0}^{\infty} |a_n| t^n \right\}^K$$

for $0 \le t < r$.

EXAMPLE Again, the product series may have larger radius of convergence than the original. Consider

$$(1+2x)^{\frac{1}{2}} = 1 + x - \frac{1}{2!}x^2 + \frac{1 \cdot 3}{3!}x^3 - \frac{1 \cdot 3 \cdot 5}{4!}x^4 + \cdots$$

which has radius $\frac{1}{2}$. However, the square of this series is just $1 + 2x$ which has infinite radius. □

Now we come to compositions. Usually, the formal composition does not make sense. The formal K-fold product series has constant term $C_{K,0} = a_0^K$ and so the constant term of $g \circ f$ would be $\sum_{K=0}^{\infty} b_K a_0^K$ which is an infinite sum. So, in general, composition of power series is not a formal operation. However, if we suppose that $a_0 = 0$, then it does become a formal operation. In this special case, the series $\{\sum_{n=0}^{\infty} a_n x^n\}^K$ starts with the term in x^K (or later if $a_1 = 0$). Hence, in computing the coefficient of x^n in $g \circ f$ we need only consider $K = 0, 1, \ldots, n$. So each coefficient of $g \circ f$ is in fact a finite sum.

In practice, it is convenient to break up the discussion of general compositions of power series into two separate operations. One of these is the special type of composition with $a_0 = 0$ which is a formal operation and the other is the recentering of power series which is not a formal operation. We will deal with recentering later.

THEOREM 114 *Suppose that $a_0 = 0$. Then the formally composed series of $g \circ f$ has strictly positive radius and converges to $g(f(x))$ for x in some neighbourhood of 0. In most situations, one can say nothing about the radius of convergence, except that it is strictly positive. However, if s is infinite, then the formally composed series has radius of convergence at least r.*

Proof. Let

$$\varphi(t) = \sum_{n=1}^{\infty} |a_n| t^n.$$

The series has radius $r > 0$ and so φ is continuous at 0. Hence, there is a number $\rho > 0$ such that

$$0 \le t < \rho \qquad \Longrightarrow \qquad |\varphi(t)| < s.$$

Now we have for $|x| < \rho$, $|f(x)| = |\sum_{n=1}^{\infty} a_n x^n| \le \varphi(|x|) < s$, so that

$$g \circ f(x) = \sum_{K=0}^{\infty} b_K \Big(f(x) \Big)^K.$$

Note that if $s = \infty$, then we may take $\rho = r$ if r is finite or ρ to be any positive number (as large as we please) if $r = \infty$. Now, using the Corollary 113 we get

$$g \circ f(x) = \sum_{K=0}^{\infty} b_K \sum_{n=K}^{\infty} c_{K,n} x^n. \tag{5.6}$$

The inner sum could be taken from $n = 0$ to infinity, but $c_{K,n} = 0$ for $0 \le n < K$. What we would like to do is to interchange the order of summation in (5.6). This would yield

$$g \circ f(x) = \sum_{n=0}^{\infty} \left\{ \sum_{K=0}^{n} b_K c_{K,n} \right\} x^n.$$

and indeed, $\sum_{K=0}^{n} b_K c_{K,n}$ is the coefficient of x^n in the formal powers series for $g \circ f$. To justify this interchange, we must apply Theorem 59. We need to show

$$\sum_{K=0}^{\infty} |b_K| \sum_{n=K}^{\infty} |c_{K,n}| t^n < \infty. \tag{5.7}$$

But, according to Corollary 113,

$$\sum_{n=K}^{\infty} |c_{K,n}| t^n \le \Big(\varphi(t)\Big)^K$$

and (5.7) holds since $\varphi(t) < s$. The radius of convergence of the formally composed series is then at least ρ. In the special case $s = \infty$, (5.7) holds for $\rho = r$ if r is finite, or for every finite $\rho > 0$ if r is infinite. The radius of convergence of the series is therefore at least r. ∎

COROLLARY 115 *Suppose that $a_0 \neq 0$. Then*

$$\frac{1}{f(x)} = d_0 + d_1 x + d_2 x^2 + d_3 x^3 + \cdots$$

with strictly positive radius. In fact, the coefficients $d_0, d_1, \ldots$ can be obtained by successively solving the recurrence relations

$$\begin{aligned} 1 &= a_0 d_0 \\ 0 &= a_1 d_0 + a_0 d_1 \\ 0 &= a_2 d_0 + a_1 d_1 + a_0 d_2 \\ 0 &= a_3 d_0 + a_2 d_1 + a_1 d_2 + a_0 d_3 \end{aligned}$$

et cetera.

Proof. We can assume without loss of generality that $a_0 = 1$. Now, let us define $h(x) = \sum_{n=1}^{\infty} a_n x^n$ and $g(y) = (1+y)^{-1}$. Then, applying Theorem 114, we see that

$$\frac{1}{f(x)} = \frac{1}{1 + h(x)} = (g \circ h)(x)$$

has a power series expansion with strictly positive radius. Once we know this, then both f and $\frac{1}{f}$ have expansions with strictly positive radius and

$$f(x) \cdot \frac{1}{f(x)} = 1$$

so the Product Theorem 5.5 allows us to conclude that the coefficients are in fact obtained by formal multiplication, leading to the recurrence relations cited above. ■

5.3 Power Series Examples

EXAMPLE Consider $f_a(x) = |x|^a$. Does this function have a power series about $x = 0$ with strictly positive radius? Well, if a is an even nonnegative integer, the answer is clearly "Yes". If a takes other values then it is fairly easy to see that f_a is not infinitely differentiable and hence by Corollary 110 it cannot have a power series expansion with strictly positive radius. Let's look at this in detail. If $a < 0$ then f_a is unbounded in every neighbourhood of 0. If $a > 0$ and not an integer, let $k = \lfloor a \rfloor$, then $f_a^{(k)}(x) = c_a(\mathrm{sgn}(x))^k |x|^{a-k}$ with $c_a \neq 0$. This function is not differentiable at $x = 0$. If a is an odd integer, then $f_a^{(a-1)}(x) = c_a|x|$ with $c_a \neq 0$. Again this function is not differentiable at $x = 0$.

Similarly, $g_a(x) = \mathrm{sgn}(x)|x|^a$ has a power series with strictly positive radius about $x = 0$ if and only if a is a nonnegative odd integer. □

EXAMPLE Consider

$$f(x) = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{1}{n} x^n, \tag{5.8}$$

which has radius 1. Differentiation gives

$$f'(x) = \sum_{n=0}^{\infty} (-1)^n x^n = \frac{1}{1+x}$$

for $-1 < x < 1$. Since $f(0) = 0$ we can deduce from the Mean Value Theorem that

$$f(x) = \ln(1 + x)$$

at least for $-1 < x < 1$. This is because both sides agree at $x = 0$ and they have the same derivative.

But $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n}$ converges by the alternating series test. So, according to Abel's Theorem, The series in (5.8) converges uniformly on $[0, 1]$. Therefore f is continuous on $[0, 1]$. Thus $f(1) = \lim_{x\to 1-} f(x) = \lim_{x\to 1-} \ln(1 + x) = \ln 2$. It follows that

$$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n} = \ln 2.$$

□

EXAMPLE Another very similar example is

$$f(x) = \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{2n-1}x^{2n-1}, \tag{5.9}$$

which also has radius 1. Differentiation gives

$$f'(x) = \sum_{n=1}^{\infty}(-1)^{n-1}x^{2n-2} = \sum_{n=0}^{\infty}(-1)^{n}x^{2n} = \frac{1}{1+x^2}$$

for $-1 < x < 1$. Since $f(0) = 0$ we can deduce from the Mean Value Theorem that

$$f(x) = \arctan(x)$$

at least for $-1 < x < 1$. This is because both sides agree at $x = 0$ and they have the same derivative.

But $\displaystyle\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{2n-1}$ converges by the alternating series test. So, according to Abel's Theorem, The series in (5.9) converges uniformly on $[0, 1]$. Therefore f is continuous on $[0, 1]$. Thus $f(1) = \lim_{x\to 1-} f(x) = \lim_{x\to 1-} \arctan(x) = \dfrac{\pi}{4}$. It follows that

$$\sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{2n-1} = \frac{\pi}{4}.$$

□

EXAMPLE The example that we just looked at

$$\frac{1}{1+x^2} = 1 - x^2 + x^4 - x^6 + \cdots$$

is an interesting one. Why does it have radius 1? The function $x \mapsto \dfrac{1}{1+x^2}$ seems to be a very nice function on the whole real line and it doesn't appear to have any kind of singularity at $x = \pm 1$. So what is restricting the radius of convergence to be 1? The answer is that if a power series

$$a_0 + a_1 z + a_2 z^2 + a_3 z^3 + a_4 z^4 + \cdots$$

has radius 1, then it follows fairly straightforwardly that it converges for all *complex* values of z with $|z| < 1$. It is the singularities of the function

$$z \mapsto \frac{1}{1+z^2}$$

as a mapping on $\mathbb{C}$ that causes the problem. These singularities are at i and at $-i$. In fact, power series are inextricably linked to complex variables and complex analysis. A knowledge of complex analysis actually leads to simpler proofs of some of the theorems we have presented. There is a general theorem that tells that the radius of convergence of a rational function, that is a function of the type

$$x \mapsto \frac{p(x)}{q(x)}$$

where p and q are polynomial functions without (nonconstant) common factors, is the distance from the point we are expanding about, to the nearest zero of q. Thus, if we were to expand the function $x \mapsto \dfrac{1}{1+x^2}$ in powers of $x - \dfrac{1}{2}$, the radius of convergence would be $\dfrac{\sqrt{5}}{2} = \left|\dfrac{1}{2} \mp i\right|$. This expansion, does give information about the function near $x = 1$. □

EXAMPLE Power series are very useful numerically. They can be used to define all the standard transcendental functions and some not so standard ones. Some caution is needed in applying power series formulæ even when they clearly converge. Take for example the Bessel function of order zero. It has an expansion

$$1 - \frac{x^2}{4(1!)^2} + \frac{x^4}{4^2(2!)^2} - \frac{x^6}{4^3(3!)^2} + \frac{x^8}{4^4(4!)^2} - \cdots$$

which has infinite radius and appears to be very rapidly convergent. For small values of $|x|$ it is. But try to implement the formula on a modern computer with $|x|$ bigger than say 15 and you will discover the meaning of roundoff error. The terms are huge in comparison with the actual infinite sum and they change sign and produce a lot of cancellation. Even small relative errors in the terms overwhelm the final answer. $\square$

EXAMPLE If, all the terms in a series have the same sign, roundoff error is usually less of a problem. From the numerical viewpoint, it can be worth going the extra mile to find expansions with this property. For example, we obviously have

$$\int_0^x e^{-\frac{1}{2}x^2}dx = \sum_{n=0}^{\infty}(-)^n \frac{x^{2n+1}}{n!2^n(2n+1)}$$

because this is a formally integrated series. On the other hand, another series expansion is available

$$\begin{aligned}\int_0^x e^{-\frac{1}{2}x^2}dx &= e^{-\frac{x^2}{2}}\left(x + \frac{x^3}{3} + \frac{x^5}{3\cdot 5} + \frac{x^7}{3\cdot 5\cdot 7} + \cdots\right). \\ &= \frac{x + \frac{x^3}{3} + \frac{x^5}{3\cdot 5} + \frac{x^7}{3\cdot 5\cdot 7} + \cdots}{1 + \frac{x^2}{2} + \frac{x^4}{2\cdot 4} + \frac{x^6}{2\cdot 4\cdot 6} + \cdots}\end{aligned}$$

How do we justify this expansion? We define

$$f(x) = x + \frac{x^3}{3} + \frac{x^5}{3\cdot 5} + \frac{x^7}{3\cdot 5\cdot 7} + \cdots$$

The series has infinite radius and it is easy to check that it satisfies the differential equation $f'(x) = 1 + xf(x)$. (In practice of course one first argues backwards to find the correct equation and then the correct coefficients.) We then check that

$$\frac{d}{dx}e^{-\frac{1}{2}x^2}f(x) = e^{-\frac{1}{2}x^2}\Big(f'(x) - xf(x)\Big) = e^{-\frac{1}{2}x^2}.$$

Integrating gives the required result. $\square$

EXAMPLE Back in the section on uniform convergence, we showed that you can approximate any continuous function on the interval $[0, 1]$ uniformly by polynomials. We used the Bernstein polynomials to do this. Could we do the same thing

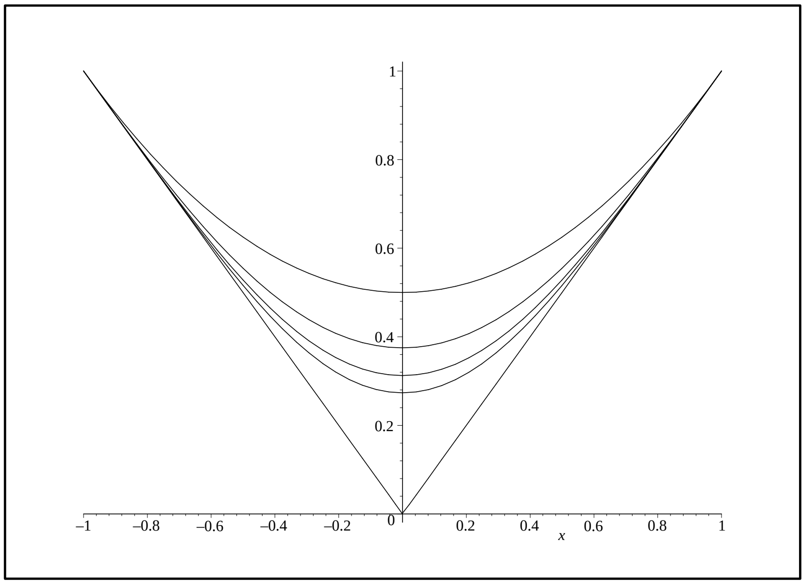

Figure 5.1: The functions $f(x) = |x|, B_2(f), B_4(f), B_6(f)$ and $B_8(f)$.

with power series, each polynomial being a partial sum of the power series. The answer is no. Because every function represented by power series is infinitely differentiable, this is not possible. Consider for example the function $f(x) = |x|$ on the interval $[-1, 1]$. After rescaling the Bernstein polynomials to $[-1, 1]$, we get

$$
\begin{aligned}
B_1(f, x) &= 1, \\
B_2(f, x) = B_3(f, x) &= 2^{-1}(1 + x^2), \\
B_4(f, x) = B_5(f, x) &= 2^{-3}(3 + 6x^2 - x^4), \\
B_6(f, x) = B_7(f, x) &= 2^{-5}(10 + 30x^2 - 10x^4 + 2x^6), \\
B_8(f, x) = B_9(f, x) &= 2^{-7}(35 + 140x^2 - 140x^4 + 28x^6 - 5x^8).
\end{aligned}
$$

Observe how the coefficients change, for example the constant coefficient is decreasing to 0. It is indeed the case that $B_n(f) = B_{n+1}(f)$ if n is even!

□

5.4 Recentering Power Series

In this section we deal with recentering power series. This is not a formal power series operation. We will start with a power series centered at 0, namely

$$a_0 + a_1 x + a_2 x^2 + \cdots$$

Let us suppose that it has radius $r > 0$. Now let $|\alpha| < r$. Then we wish to expand the same gadget about $x = \alpha$

$$b_0 + b_1(x-\alpha) + b_2(x-\alpha)^2 + b_3(x-\alpha)^3 + \cdots \tag{5.10}$$

Substituting $t = x - \alpha$ and comparing the coefficient of t^n in

$$\sum_{k=0}^{\infty} a_k (t+\alpha)^k = \sum_{n=0}^{\infty} b_n t^n \tag{5.11}$$

we find the formula

$$b_n = \sum_{k=n}^{\infty} {}^kC_n \, a_k \alpha^{k-n}. \tag{5.12}$$

So the coefficients of the recentered series are infinite sums (as opposed to finite sums) and this is why the recentering operation is not an operation on formal power series.

THEOREM 116 *Under the hypotheses given above, the series (5.12) defining b_n converges for all $n \in \mathbb{Z}^+$. The radius of convergence of (5.10) is at least $r - |\alpha|$. Finally, the identity (5.11) holds provided that $|t| < r - |\alpha|$.*

Proof. Let $|\alpha| < \rho < r$ and $\sigma = \rho - |\alpha|$. Then

$$\begin{aligned}\sum_{n=0}^{\infty} \sigma^n \sum_{k=n}^{\infty} {}^kC_n \, |a_k||\alpha|^{k-n} &= \sum_{k=0}^{\infty} |a_k| \sum_{n=0}^{k} {}^kC_n \, \sigma^n |\alpha|^{k-n} \\ &= \sum_{k=0}^{\infty} |a_k| (\sigma + |\alpha|)^k < \infty\end{aligned}$$

since the order of summation can be interchanged for series of positive terms and since $\sigma + |\alpha| = \rho < r$. In particular it follows that for each fixed n, the inner series

$\sum_{k=n}^{\infty} {}^kC_n\,|a_k||\alpha|^{k-n}$ converges and hence the series (5.12) converges absolutely for each $n \in \mathbb{Z}^+$. The same argument now shows that

$$\sum_{n=0}^{\infty} |b_n|\sigma^n \leq \sum_{n=0}^{\infty} \sigma^n \sum_{k=n}^{\infty} {}^kC_n\,|a_k||\alpha|^{k-n} < \infty$$

and so (5.10) converges absolutely whenever $|t| < r - |\alpha|$. So the radius of convergence of the recentered series is at least $r - |\alpha|$. Finally, we use Fubini's theorem to show that (5.11) holds. Effectively, since

$$\sum_{k=0}^{\infty} |a_k| \sum_{n=0}^{k} {}^kC_n\,|t|^n|\alpha|^{k-n} < \infty$$

we have

$$\begin{aligned}
\sum_{n=0}^{\infty} b_n t^n &= \sum_{n=0}^{\infty} t^n \sum_{k=n}^{\infty} {}^kC_n\,a_k\alpha^{k-n} \\
&= \sum_{k=0}^{\infty} a_k \sum_{n=0}^{k} {}^kC_n\,t^n\alpha^{k-n} \\
&= \sum_{k=0}^{\infty} a_k(t+\alpha)^k
\end{aligned}$$

by interchanging the order of summation. ■

6

The Elementary Functions

6.1 The Exponential Function

We define

$$\exp(x) = \sum_{n=0}^{\infty} \frac{1}{n!} x^n.$$

It follows from the ratio test that this power series has infinite radius of convergence, so exp is an infinitely differentiable function on the whole of $\mathbb{R}$. From differentiating the series (using Theorem 108) we get $\exp'(x) = \exp(x)$. Now, fix $a \in \mathbb{R}$ and consider

$$x \mapsto \exp(x + a) \exp(-x)$$

We find by using the chain rule that this function has derivative everywhere zero. So it must be a constant by the Mean Value Theorem. Putting $x = 0$, we find that the constant is $\exp(a)$ this gives

$$\exp(x + a) \exp(-x) = \exp(a) \qquad \text{for all } x, a \in \mathbb{R}.$$

Next substitute $a = 0$ to get $\exp(x) \exp(-x) = \exp(0) = 1$. This shows that $\exp(x) \neq 0$ for all real x. Now, $\exp(0) = 1$, so if $\exp(x) < 0$, we could find, using the continuity of exp and the Intermediate Value Theorem, a point between 0 and x where exp vanishes. Since this is impossible, we conclude that $\exp(x) > 0$.

We can also get

$$\exp(x + a) = \exp(x + a) \exp(-x) \exp(x) = \exp(a) \exp(x)$$

which is the additive-multiplicative property of the exponential.

Now, $\exp'(x) = \exp(x) > 0$ so that exp is a strictly increasing function. In particular

$$\exp(x) \begin{cases} < 1 & \text{if } x < 0, \\ = 1 & \text{if } x = 0, \\ > 1 & \text{if } x > 0. \end{cases}$$

Let $f(x) = \exp(x) - x - 1$. Then $f'(x) = \exp(x) - 1$ and it follows from the Mean Value Theorem that f is increasing for $x \geq 0$ and decreasing for $x \leq 0$. So f takes its minimum value at $x = 0$. This yields the well-known inequality

$$\exp(x) \geq 1 + x \qquad \text{for all } x \in \mathbb{R}.$$

Furthermore we get

$$\lim_{x \to \infty} \exp(x) = \infty$$

and

$$\lim_{x \to \infty} \exp(-x) = \lim_{x \to \infty} \frac{1}{\exp(x)} = 0.$$

Thus, in fact, exp takes all positive values.

6.2 The Natural Logarithm

For the definition of the natural logarithm, we will take

$$\ln(x) = \int_1^x \frac{1}{t} dt, \qquad x > 0.$$

Substituting $t = \exp(s)$, we find that $\ln(x) = y$ where y is the unique solution of the equation $x = \exp(y)$. From this it follows that both $\exp(\ln x) = x$ and, if we start from y by defining $x = \exp(y)$ that $y = \ln(\exp(y))$. So, exp and ln are inverse functions.

By the Fundamental Theorem of Calculus (Theorem 80), we have $\ln'(x) = \dfrac{1}{x}$. We can now get more information about the exponential function.

EXAMPLE We have $\lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n = \exp(x)$ for each fixed real x. To see this, we write

$$\lim_{n \to \infty} n \ln\left(1 + \frac{x}{n}\right) = \lim_{n \to \infty} \frac{\ln\left(1 + \frac{x}{n}\right) - \ln(1)}{\frac{1}{n}} = \lim_{h \to 0} \frac{\ln(1 + hx) - \ln(1)}{h} = x$$

since the derivative of the function $t \mapsto \ln(1+tx)$ at $t = 0$ is just x. Since the exponential function is continuous we now have

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n = \exp(x)$$

by applying exp to both sides and passing exp through the limit. Convergence here is not uniform since $\lim_{x\to-\infty}\exp(x) = 0$, but $\left|1+\frac{x}{n}\right|^n$ is large when x is large and negative.

Convergence is uniform on compacta. To see this, restrict x to $-a \le x \le a$ for some $a > 0$. Now compute

$$\frac{d}{dx}\left(n\ln\left(1+\frac{x}{n}\right) - x\right) = -\frac{x}{n+x}.$$

As soon as $n > a$, the only critical point is at $x = 0$. So we can deduce that

$$\left|n\ln\left(1+\frac{x}{n}\right) - x\right| \le \max\left(\left|n\ln\left(1+\frac{a}{n}\right) - a\right|, \left|n\ln\left(1+\frac{-a}{n}\right) - (-a)\right|\right)$$

So $n\ln\left(1+\frac{x}{n}\right) \longrightarrow x$ uniformly on $|x| \le a$ and we can then deduce that $\left(1+\frac{x}{n}\right)^n \longrightarrow \exp(x)$ uniformly on the same set.

Another fact about $\left(1+\frac{x}{n}\right)^n$ is that it is increasing with n when $x \ge 0$. To see this we expand by the binomial theorem

$$\begin{aligned}\left(1+\frac{x}{n}\right)^n &= \sum_{k=0}^{n} {}^nC_k \left(\frac{x}{n}\right)^k \\ &= 1 + x + \sum_{k=2}^{n}\frac{x^k}{k!}\prod_{\ell=1}^{k-1}\left(1-\frac{\ell}{n}\right)\end{aligned}$$

The expansion of $\left(1+\frac{x}{n+1}\right)^{n+1}$ is similar, but contains an extra term in x^{n+1}. This term is nonnegative. Also, the coefficient of x^k in $\left(1+\frac{x}{n+1}\right)^{n+1}$ is

$$\frac{1}{k!}\prod_{\ell=1}^{k-1}\left(1-\frac{\ell}{n+1}\right)$$

clearly larger than the corresponding coefficient in $\left(1+\frac{x}{n}\right)^n$ which is

$$\frac{1}{k!}\prod_{\ell=1}^{k-1}\left(1-\frac{\ell}{n}\right)$$

It follows that

$$\left(1+\frac{x}{n}\right)^n \le \left(1+\frac{x}{n+1}\right)^{n+1} \qquad \text{for } x \ge 0.$$

□

Next, we get the power series for the logarithm. We have

$$\begin{aligned}\ln(1+x) &= \int_1^{1+x}\frac{dt}{t} = \int_0^x \frac{ds}{1+s}\\ &= \int_0^x \sum_{n=0}^{\infty}(-)^n s^n ds\\ &= \sum_{n=0}^{\infty}(-)^n\frac{x^{n+1}}{n+1} \qquad (6.1)\end{aligned}$$

where the power series involved have radius 1. We also have similarly

$$-\ln(1-x) = \sum_{n=0}^{\infty}\frac{x^{n+1}}{n+1}$$

also with radius 1.

EXAMPLE A nice application of the power series expansion for the logarithm gives the ***Newton identities*** . For n variables $x_1, x_2, \ldots, x_n$, we define the ***elementary symmetric functions*** $e_k(x_1, x_2, \ldots, x_n)$ by

$$\prod_{j=1}^{n}(1+tx_j) = \sum_{k=0}^{n} t^k e_k(x_1, x_2, \ldots, x_n).$$

By convention $e_0(x_1, x_2, \ldots, x_n) = 1$ and we can check that

$$e_k(x_1, x_2, \ldots, x_n) = \sum_{i_1<i_2<\cdots<i_k} x_{i_1}x_{i_2}\cdots x_{i_k}$$

So, for t small, we get

$$\sum_{j=1}^{n} \ln(1+tx_j) = \ln\left(\sum_{k=0}^{n} t^k e_k(x_1, x_2, \ldots, x_n)\right).$$

and expanding the term on the left

$$\sum_{j=1}^{n}\sum_{m=0}^{\infty}(-)^m t^{m+1}\frac{x_j^{m+1}}{m+1} = \ln\left(\sum_{k=0}^{n} t^k e_k\right).$$

Now denote $\sigma_m = \sum_{j=1}^{n} x_j^m$ and manipulating the term on the left we get

$$\sum_{m=0}^{\infty}(-)^m t^{m+1}\frac{\sigma_{m+1}}{m+1} = \ln\left(\sum_{k=0}^{n} t^k e_k\right).$$

Next, get rid of the logarithm by differentiating with respect to t

$$\sum_{m=0}^{\infty}(-)^m t^m \sigma_{m+1} = \frac{\sum_{k=1}^{n} k t^{k-1} e_k}{\sum_{k=0}^{n} t^k e_k}.$$

and hence

$$\sum_{k=0}^{n-1}(k+1)t^k e_{k+1} = \left(\sum_{\ell=0}^{n} t^\ell e_\ell\right)\left(\sum_{m=0}^{\infty}(-)^m t^m \sigma_{m+1}\right).$$

Next, equate the coefficient of t^k to get

$$\sum_{m=0}^{k}(-)^m \sigma_{m+1} e_{k-m} = \begin{cases} (k+1)e_{k+1} & \text{if } 0 \le k < n, \\ 0 & \text{if } k \ge n. \end{cases}$$

These identities between symmetric polynomials in $x_1, x_2, \ldots, x_n$ are collectively called Newton's identities. □

6.3 Powers

We can use the exponential and logarithm together to define powers. For $x > 0$ and $a \in \mathbb{R}$ we simply define

$$x^a = \exp(a\ln(x)).$$

Taking logarithms we get

$$\ln(x^a) = a\ln(x)$$

There are two standard identities for powers

$$x^{a+b} = \exp((a+b)\ln(x)) = \exp(a\ln(x)) \cdot \exp(b\ln(x)) = x^a \cdot x^b$$

and

$$x^{ab} = \exp((ab)\ln(x)) = \exp(a\ln(x^b)) = (x^a)^b.$$

The definition can be extended in the following ways.

- If $x = 0$ then $x^a = 0$ if $a > 0$.
- $0^0 = 1$.
- If $x < 0$ and a is a rational number with odd denominator, it is still possible to make sense of x^a. If $a = \frac{p}{q}$ with p an integer and q an odd integer, we have $x^a = t^p$ where t is the unique real number such that $t^q = x$.

We now obtain the binomial theorem for general powers. Let $\alpha \in \mathbb{R}$ and define $c_0 = 1$, $c_1 = \alpha$, $c_2 = \frac{\alpha(\alpha+1)}{2!}$, $c_3 = \frac{\alpha(\alpha+1)(\alpha+2)}{3!}$ and so forth. Then define a power series f by

$$f(x) = \sum_{k=0}^{\infty} c_k x^k.$$

Let us find the radius of convergence of this series. If α is a nonpositive integer $0, -1, -2, \ldots$ then the series will terminate and the radius of convergence will be infinite. Otherwise, we can use the ratio test to determine the radius of convergence.

$$\left|\frac{c_{k+1}x^{k+1}}{c_k x^k}\right| = \left|\frac{\alpha+k}{k+1}\right| |x| \longrightarrow |x|$$

so the series will converge if $|x| < 1$ and diverge if $|x| > 1$. So the radius is 1.

For $-1 < x < 1$, we obtain, using known properties of power series we get

$$\begin{aligned}(1-x)f'(x) - \alpha f(x) &= \sum_{k=0}^{\infty} kc_k x^{k-1} - \sum_{k=0}^{\infty} kc_k x^k - \alpha \sum_{k=0}^{\infty} c_k x^k, \\ &= \sum_{k=0}^{\infty}(k+1)c_{k+1}x^k - \sum_{k=0}^{\infty} kc_k x^k - \alpha \sum_{k=0}^{\infty} c_k x^k, \\ &= 0,\end{aligned}$$

since $(k+1)c_{k+1} = (\alpha + k)c_k$.

Now let

$$h(x) = (1-x)^\alpha f(x) \qquad \text{for } -1 < x < 1.$$

Note that

$$\begin{aligned}\frac{d}{dx}(1-x)^\alpha &= \frac{d}{dx}\exp(\alpha\ln(1-x)) = -\frac{\alpha}{1-x}\exp(\alpha\ln(1-x)) \\ &= -\alpha(1-x)^{-1}(1-x)^\alpha = -\alpha(1-x)^{\alpha-1}\end{aligned}$$

so that

$$\begin{aligned}h'(x) &= (1-x)^\alpha f'(x) - \alpha(1-x)^{-1}(1-x)^\alpha f(x) \\ &= (1-x)^{-1}(1-x)^\alpha\Big((1-x)f'(x) - \alpha f(x)\Big) \\ &= 0,\end{aligned}$$

always for $-1 < x < 1$. So, h is constant on $]-1,1[$ and since $h(0) = f(0) = 1$, we find that

$$(1-x)^\alpha f(x) = 1$$

Hence we have

$$f(x) = (1-x)^{-\alpha} \qquad \text{for } -1 < x < 1.$$

This gives the binomial expansion for general powers

$$(1-x)^{-\alpha} = 1 + \alpha x + \frac{\alpha(\alpha+1)}{2!}x^2 + \frac{\alpha(\alpha+1)(\alpha+2)}{3!}x^3 + \cdots$$

If we wish, we can replace x by $-x$ and α by $-\alpha$ to get

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2!}x^2 + \frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3 + \cdots$$

which is the more commonly stated form. The series in both forms have radius 1.

•6.4 Stirling's Formula

We already obtained Stirling's formula by an oddball method. Now we will get a finer estimate by a more standard approach. Let

$$a_n = n!\; e^n n^{-(n+\frac{1}{2})}.$$

Then we have

$$\frac{a_n}{a_{n+1}} = e^{-1}\left(1+\frac{1}{n}\right)^{n+\frac{1}{2}}.$$

The idea is to estimate this ratio. To do this we start from two forms of (6.1)

$$\begin{aligned}\ln(1+x) &= x - \frac{1}{2}x^2 + \frac{1}{3}x^3 - \frac{1}{4}x^4 + \frac{1}{5}x^5 - \cdots \\ \ln(1-x) &= -x - \frac{1}{2}x^2 - \frac{1}{3}x^3 - \frac{1}{4}x^4 - \frac{1}{5}x^5 - \cdots\end{aligned}$$

Subtracting we get

$$\ln\left(\frac{1+x}{1-x}\right) = 2\sum_{k=0}^{\infty}\frac{x^{2k+1}}{2k+1}.$$

Now substitute $x = \dfrac{1}{2n+1}$ to get

$$\ln\left(\frac{n+1}{n}\right) = 2\sum_{k=0}^{\infty}\frac{1}{2k+1}(2n+1)^{-(2k+1)}.$$

We now get after multiplication by $n + \dfrac{1}{2}$

$$\left(n+\frac{1}{2}\right)\ln\left(1+\frac{1}{n}\right) = 1 + \sum_{k=1}^{\infty}\frac{1}{2k+1}(2n+1)^{-2k}.$$

We estimate the series on the right above and below. For the upper bound

$$\begin{aligned}\sum_{k=1}^{\infty}\frac{1}{2k+1}(2n+1)^{-2k} &< \sum_{k=1}^{\infty}\frac{1}{3}(2n+1)^{-2k} \\ &= \frac{1}{3}\cdot\frac{1}{(2n+1)^2}\cdot\left(1-\frac{1}{(2n+1)^2}\right)^{-1} \\ &= \frac{1}{3}\cdot\frac{1}{4n^2+4n} \\ &= \left(\frac{1}{12n} - \frac{1}{12(n+1)}\right).\end{aligned}$$

For the lower bound

$$\sum_{k=1}^{\infty} \frac{1}{2k+1}(2n+1)^{-2k} = \frac{1}{3} \cdot \frac{1}{(2n+1)^2} \cdot \sum_{k=1}^{\infty} \frac{3}{2k+1}(2n+1)^{-2k+2}$$

$$> \frac{1}{3} \cdot \frac{1}{(2n+1)^2} \cdot \sum_{k=1}^{\infty} \left(\frac{3}{5(2n+1)^2}\right)^{k-1}$$

since $\left(\frac{3}{5}\right)^{k-1} \leq \frac{3}{2k+1}$ for $k \geq 1$,

$$= \frac{1}{3} \cdot \frac{1}{(2n+1)^2} \cdot \left(1 - \frac{3}{5(2n+1)^2}\right)^{-1}$$

$$= \frac{5}{6} \cdot \frac{1}{10n^2 + 10n + 1}$$

$$> \frac{192}{2304n^2 + 2400n + 49}$$

for $n \geq 2$ since $5(2304n^2 + 2400n + 49) - 6 \cdot 192(10n^2 + 10n + 1) = 480n - 907$,

$$= \left(\frac{1}{12n + \frac{1}{4}} - \frac{1}{12(n+1) + \frac{1}{4}}\right).$$

So, to recap, we have

$$\left(\frac{1}{12n + \frac{1}{4}} - \frac{1}{12(n+1) + \frac{1}{4}}\right) < \left(n + \frac{1}{2}\right) \ln\left(1 + \frac{1}{n}\right) - 1$$

$$< \left(\frac{1}{12n} - \frac{1}{12(n+1)}\right). \tag{6.2}$$

Now set

$$x_n = \exp\left(-\frac{1}{12n + \frac{1}{4}}\right) \qquad \text{and} \qquad y_n = \exp\left(-\frac{1}{12n}\right)$$

then we have, exponentiating (6.2)

$$1 < \frac{x_{n+1}}{x_n} < \frac{a_n}{a_{n+1}} < \frac{y_{n+1}}{y_n} \qquad (n \geq 2). \tag{6.3}$$

There is a lot of information in (6.3). First, $(a_n)_{n=2}^{\infty}$ is a decreasing positive sequence and hence converges to a nonnegative number a. Second $(a_n x_n)_{n=2}^{\infty}$ is also decreasing and $(a_n y_n)_{n=2}^{\infty}$ is increasing. Both of these sequences must converge to a. Hence we have

$$0 < a_n y_n < a < a_n x_n \qquad (n \geq 2). \tag{6.4}$$

Now, what is the value of a? We get

$$\frac{a_n^2}{a_{2n}} = \frac{(n!)^2 e^{2n} n^{-(2n+1)}}{(2n)! e^{2n} (2n)^{-(2n+\frac{1}{2})}} = \frac{(n!)^2 4^n}{(2n)!}\sqrt{\frac{2}{n}}$$

and so

$$\frac{a_n^4}{a_{2n}^2} = \left(\frac{4^n (n!)^2 \cdot 2^n n! \cdot 2^{n-1}(n-1)!}{(2n+1)!(2n-1)!}\right)\left(\frac{2n \cdot (2n+1) \cdot 2}{2n \cdot n}\right)$$

But the first bracket on the right is just the Wallis product $\displaystyle\prod_{k=1}^{n} \frac{4k^2}{4k^2-1}$ and it follows from (4.14) that

$$\frac{a_n^4}{a_{2n}^2} \longrightarrow \frac{\pi}{2} \cdot 4 = 2\pi$$

It follows that $a = \sqrt{2\pi}$. Restating (6.4) we get

$$\sqrt{2n\pi}\; n^n \exp\left(-n + \frac{1}{12n + \frac{1}{4}}\right) < n! < \sqrt{2n\pi}\; n^n \exp\left(-n + \frac{1}{12n}\right). \tag{6.5}$$

This is a sharper form of Stirling's Theorem.

6.5 Trigonometric Functions

The functions sin and cos can also be defined by power series

$$\sin(x) = x - \frac{1}{3!}x^3 + \frac{1}{5!}x^5 - \frac{1}{7!}x^7 + \cdots$$
$$\cos(x) = 1 - \frac{1}{2!}x^2 + \frac{1}{4!}x^4 - \frac{1}{6!}x^6 + \cdots$$

which have infinite radius of convergence. We clearly have $\cos(-x) = \cos(x)$ and $\sin(-x) = -\sin(x)$. They are related to the exponential function of an imaginary argument via identities such as

$$\cos(x) = \frac{\exp(ix) + \exp(-ix)}{2} \qquad \sin(x) = \frac{\exp(ix) - \exp(-ix)}{2i}$$

and

$$\exp(ix) = \cos(x) + i\sin(x)$$

To make progress, we need to extend the additive-multiplicative property of exp to the complex setting. We have

$$\begin{aligned}
\exp(a+b) &= \sum_{n=0}^{\infty} \frac{1}{n!}(a+b)^n \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{p=0}^{n} {}^nC_p\, a^p b^{n-p} \\
&= \sum_{n=0}^{\infty} \sum_{p=0}^{n} \frac{1}{p!} \frac{1}{(n-p)!} a^p b^{n-p}
\end{aligned}$$

while

$$\begin{aligned}
\exp(a)\exp(b) &= \sum_{p=0}^{\infty} \frac{1}{p!} a^p \sum_{q=0}^{\infty} \frac{1}{q!} b^q \\
&= \sum_{p=0}^{\infty} \sum_{n=p}^{\infty} \frac{1}{p!} \frac{1}{(n-p)!} a^p b^{n-p}
\end{aligned}$$

So we need only show that the order of summation can be interchanged using Fubini's Theorem (Theorem 59). For this we need only observe that

$$\sum_{p=0}^{\infty} \sum_{n=p}^{\infty} \frac{1}{p!} \frac{1}{(n-p)!} |a|^p |b|^{n-p} = \exp(|a|)\exp(|b|) < \infty$$

Hence $\exp(a+b) = \exp(a)\exp(b)$ for all $a, b \in \mathbb{C}$.

This yields the standard addition laws for sin and cos and also

$$\begin{aligned}
1 &= \exp(0) = \exp(ix)\exp(-ix) \\
&= \Big(\cos(x) + i\sin(x)\Big)\Big(\cos(-x) + i\sin(-x)\Big) \\
&= \Big(\cos(x) + i\sin(x)\Big)\Big(\cos(x) - i\sin(x)\Big) \\
&= \Big(\cos(x)\Big)^2 + \Big(\sin(x)\Big)^2.
\end{aligned}$$

Now we turn to the power series expansion for $\cos(2)$

$$\cos(2) = \sum_{n=0}^{\infty} (-)^n \frac{4^n}{(2n)!}$$

and it is easy to see that from the second term onwards, the terms are decreasing in absolute value. Therefore, from the proof of the alternating series test we have

$$\cos(2) < 1 - \frac{4}{2} + \frac{16}{24} = -\frac{1}{3}.$$

Since cos is continuous and $\cos(0) = 1$, we may apply the Intermediate Value Theorem to show that there exists x with $0 < x < 2$ such that $\cos(x) = 0$. Now define

$$\pi = 2 \inf\{x; x \geq 0, \cos(x) = 0\}$$

Then we have $\cos\left(\frac{\pi}{2}\right) = 0$ and again by the Intermediate Value Theorem, we have

$$0 \leq x < \frac{\pi}{2} \qquad \Longrightarrow \qquad \cos(x) > 0.$$

We check easily that $\sin'(x) = \cos(x)$ so that sin is increasing on $\left[0, \frac{\pi}{2}\right]$. Since $\sin(0) = 0$ and $\sin\left(\frac{\pi}{2}\right) = \pm 1$, we must have $\sin\left(\frac{\pi}{2}\right) = 1$ and that $\sin(x) \geq 0$ on $\left[0, \frac{\pi}{2}\right]$. Since $\cos'(x) = -\sin(x)$, we now see that cos is decreasing on $\left[0, \frac{\pi}{2}\right]$.

We have $\exp\left(\frac{\pi}{2} i\right) = i$ and it follows that $\exp(\pi i) = -1$ and that $\exp(2\pi i) = 1$. So $\exp(x + 2\pi i) = \exp(x)\exp(2\pi i) = \exp(x)$ and it follows that cos and sin are periodic with period 2π.

Now we need to see that there is no shorter period. Let $0 < t < 2\pi$ and suppose that $\cos(t) = 1$ and $\sin(t) = 0$. Define $\exp\left(\frac{t}{4} i\right) = u + iv$. Then we have $1 = (u + iv)^4 = (u^4 - 6u^2v^2 + v^4) + 4uv(u^2 + v^2)i$. Since $u^2 + v^2 = 1$ we are forced to have either $u = 0$ or $v = 0$. In the first case $\cos\left(\frac{t}{4}\right) = 0$ which is impossible for $0 < \frac{t}{4} < \frac{\pi}{2}$ and in the second case we have $\sin\left(\frac{t}{4}\right) = 0$ which is also impossible in the same range.

It is now easy to establish all the standard facts about sin and cos and we leave these as an exercise. Equally well, the standard trig functions can be built out of sin and cos and their basic properties established.

•6.6 Niven's proof of the Irrationality of π

THEOREM 117 *π^2 is irrational.*

Proof. Let us suppose that π^2 is rational. Then, we can write $\pi^2 = \dfrac{a}{b}$ where $a, b \in \mathbb{N}$. Since $\displaystyle\sum_{n=0}^{\infty} \frac{a^n}{n!} = e^a < \infty$ there exists $N \in \mathbb{N}$ such that

$$\pi \frac{a^N}{N!} < 1. \tag{6.6}$$

Now define

$$f(x) = \frac{1}{N!} x^N (1-x)^N. \tag{6.7}$$

We make the following claims about f.

- $f^{(k)}(x) = 0$ for $k > 2N$ and all x.
- $f^{(k)}(0) \in \mathbb{Z}$ for $k \in \mathbb{Z}^+$.
- $f^{(k)}(1) \in \mathbb{Z}$ for $k \in \mathbb{Z}^+$.

Since f is a polynomial of degree N, the first claim is obvious. For the second claim, we expand the right hand side of (6.7) to obtain

$$f(x) = \frac{1}{N!} \sum_{n=N}^{2N} c_n x^n \tag{6.8}$$

where the $c_n \in \mathbb{Z}$. Differentiating (6.8) k times we obtain

$$f^{(k)}(x) = \frac{1}{N!} \sum_{n=N}^{2N} n(n-1)\cdots(n-k+1) c_n x^{n-k}.$$

Now, if $k = 0, 1, \ldots, N-1$, we see that $f^{(k)}(0) = 0$ and if $k = N, N+1, \ldots, 2N$, the only surviving term in the sum is the one corresponding to $n = k$ and we obtain $f^{(k)}(0) = \dfrac{k!}{N!} c_k \in \mathbb{Z}$. For the third claim, we have from (6.7) that $f(x) =$

$f(1-x)$ which, when differentiated k times yields $f^{(k)}(x) = (-)^k f^{(k)}(1-x)$. Thus $f^{(k)}(1) = (-)^k f^{(k)}(0) \in \mathbb{Z}$.

Now define

$$
\begin{aligned}
g(x) &= b^N \sum_{k=0}^{N} (-)^k \pi^{2(N-k)} f^{(2k)}(x) \\
&= \sum_{k=0}^{N} (-)^k a^{N-k} b^k f^{(2k)}(x) \qquad (6.9)
\end{aligned}
$$

since $b^N \pi^{2(N-k)} = b^N \left(\frac{a}{b}\right)^{N-k} = a^{N-k} b^k$. It follows from (6.9) that $g(0)$ and $g(1)$ are integers.

Next, let us define

$$h(x) = g'(x)\sin(\pi x) - \pi g(x)\cos(\pi x),$$

so that

$$
\begin{aligned}
h'(x) &= g''(x)\sin(\pi x) + g'(x)\pi\cos(\pi x) - \pi g'(x)\cos(\pi x) + \pi^2 g(x)\sin(\pi x) \\
&= \Big(g''(x) + \pi^2 g(x)\Big)\sin(\pi x) \\
&= b^N \sin(\pi x)\Big(\sum_{k=0}^{N} (-)^k \pi^{2(N-k)} f^{(2k+2)}(x) + \sum_{k=0}^{N} (-)^k \pi^{2(N-k)+2} f^{(2k)}(x)\Big) \\
&= b^N \sin(\pi x)\Big(\sum_{k=1}^{N+1} (-)^{k-1} \pi^{2(N-k+1)} f^{(2k)}(x) + \sum_{k=0}^{N} (-)^k \pi^{2(N-k)+2} f^{(2k)}(x)\Big) \\
&= b^N \sin(\pi x)\Big((-)^N \pi^0 f^{(2N+2)}(x) + (-)^0 \pi^{2N+2} f(x)\Big) \\
&= b^N \pi^{2N+2} f(x)\sin(\pi x) = \pi^2 a^N f(x)\sin(\pi x)
\end{aligned}
$$

by the first claim. Thus, applying the Mean Value Theorem to h we have the existence of ξ with $0 < \xi < 1$ such that

$$h(1) - h(0) = h'(\xi) = \pi^2 a^N f(\xi)\sin(\pi\xi) \qquad (6.10)$$

But on the other hand, we have

$$h(1) - h(0) = -\pi\Big(g(1) + g(0)\Big) \qquad (6.11)$$

Combining (6.10) and (6.11) we find, dividing by π that

$$-\Big(g(1) + g(0)\Big) = \pi a^N f(\xi) \sin(\pi\xi) \tag{6.12}$$

Clearly $0 < f(\xi)\sin(\pi\xi) < \dfrac{1}{N!}$ and since $\pi \dfrac{1}{N!} a^N < 1$, we find from (6.12) that

$$0 < -\Big(g(1) + g(0)\Big) < 1$$

contradicting the fact that $g(0)$ and $g(1)$ are integers. ■

Index