Analysis

Dr J.M.E. Hyland[1]

Michælmas 1996

[1]Comments and corrections to soc-archim-notes@lists.cam.ac.uk

```
Revision: 1.5
Date: 1998-10-27 15:54:56+00
```

The following people have maintained these notes.

initial typing	Richard Cameron
– date	Paul Metcalfe

Contents

Introduction

These notes are based on the course "Analysis" given by Dr. J.M.E. Hyland in Cambridge in the Michælmas Term 1996. These typeset notes are totally unconnected with Dr. Hyland.

Other sets of notes are available for different courses. At the time of typing these courses were:

Probability	Discrete Mathematics
Analysis	Further Analysis
Methods	Quantum Mechanics
Fluid Dynamics 1	Quadratic Mathematics
Geometry	Dynamics of D.E.'s
Foundations of QM	Electrodynamics
Methods of Math. Phys	Fluid Dynamics 2
Waves (etc.)	Statistical Physics
General Relativity	Dynamical Systems
Physiological Fluid Dynamics	Bifurcations in Nonlinear Convection
Slow Viscous Flows	Turbulence and Self-Similarity
Acoustics	Non-Newtonian Fluids
Seismic Waves	

They may be downloaded from

`http://home.arachsys.com/~pdm/maths/` or
`http://www.cam.ac.uk/CambUniv/Societies/archim/notes.htm`

or you can email `soc-archim-notes@lists.cam.ac.uk` to get a copy of the sets you require.

Chapter 1

Real Numbers

1.1 Ordered Fields

Definition 1.1. *A* field *is a set* $\mathbb{F}$ *equipped with:*

- *an element* $0 \in \mathbb{F}$ *and a binary operation* $+ \colon \mathbb{F} \mapsto \mathbb{F}$*, making* $\mathbb{F}$ *an abelian group; we write* $-a$ *for the additive inverse of* $a \in \mathbb{F}$*;*

- *an element* $1 \in \mathbb{F}$ *and a binary operation* $\cdot \colon \mathbb{F} \mapsto \mathbb{F}$ *such*

  – *multiplication distributes over addition, that is:* $a \cdot 0 = 0$ *and* $a \cdot (b + c) = a \cdot b + a \cdot c$

  – $1 \neq 0$*, multiplication restricts to* $\mathbb{F}^{\times} = \mathbb{F} \backslash \{0\}$*, and* $\mathbb{F}^{\times}$ *is an abelian group under multiplication; we write* $a^{-1} = 1/a$ *for the multiplicative inverse of* $a \in \mathbb{F}^{\times}$

Examples: $\mathbb{Q}$ (rational numbers); $\mathbb{R}$ (real numbers); $\mathbb{C}$ (complex numbers).

Definition 1.2. *A relation* $<$ *on a set* $\mathbb{F}$ *is a* strict total order *when we have* $a \not< a$*,* $a < b$ *and* $b < c \Rightarrow a < c$*,* $a < b$ *or* $a = b$ *or* $b > a$ *for all* a, b *and* c *in* $\mathbb{F}$*. We write* $a \leq b$ *for* $a < b$ *or* $a = b$*, and note that in a total order* $a \leq b \Leftrightarrow b \not< a$*.*

Familiar ordered fields are $\mathbb{Q}$ and $\mathbb{R}$, but not $\mathbb{C}$.

1.2 Convergence of Sequences

Definition 1.3. *In an ordered field we define the* absolute value $|a|$ *of* a *as:*

$$|a| = \begin{cases} a & a > 0 \\ -a & a < 0 \\ 0 & a = 0 \end{cases}$$

and then we have the distance $d(a, b) = |a - b|$ *between* a *and* b*.*

In an ordered field the distance $d(a,b)$ satisfies

$$\begin{gathered} d(a,b) \geq 0 \quad \text{and} \quad d(a,b)=0 \text{ iff } a=b \\ d(a,b)=d(b,a) \\ d(a,c) \leq d(a,b)+d(b,c). \end{gathered}$$

Proof. Proof of this is easy. Start from

$$\begin{aligned} -|x| \quad &\leq x \leq \quad |x| \\ -|y| \quad &\leq y \leq \quad |y|\,. \end{aligned}$$

Add these to get

$$\begin{gathered} -(|x|+|y|) \leq x+y \leq |x|+|y| \\ |x+y| \leq |x|+|y|\,. \end{gathered}$$

Put $x=a-b, y=b-c$ for result. □

In general the distance takes values in the field in question; but in the case of $\mathbb{Q}$ and $\mathbb{R}$, the distance is real valued, so we have a *metric*.

Example 1.4. *Any ordered field has a copy of $\mathbb{Q}$ as an ordered subfield.*

Proof. We set

$$n = \underbrace{1+1+\ldots+1+1}_{n \quad times}$$

and so get $-n$, and so get r/s, $r \in \mathbb{Z}$, $s>0$ in $\mathbb{Z}$, all ordered correctly. □

Definition 1.5. *A sequence a_n* converges to a limit *a, or a_n tends to a in an ordered field $\mathbb{F}$, just when for all $\epsilon > 0$ in $\mathbb{F}$, there exists $N \in \mathbb{N}$ with $|a_n - a| < \epsilon$ for all $n \geq N$.*

We write $\lim_{n\to\infty} a_n = a$ or $a_n \to a$ as $n \to \infty$ or just $a_n \to a$, when a_n converges to a limit a. So we have

$$a_n \to a \quad \Leftrightarrow \quad \forall \epsilon > 0 \quad \exists N \quad \forall n \geq N \quad |a_n - a| < \epsilon$$

Example 1.6.

1. *$a_n \to a$ iff $|a_n - a| \to 0$*

2. *$b_n \geq 0, b_n \to 0, 0 \leq c_n \leq b_n$, then $c_n \to 0$*

3. *Suppose we have $N, k \in \mathbb{N}$ such that $b_n = a_{n+k}$ for all $n \geq N$, then $a_n \to a$ iff $b_n \to a$.*

4. *The sequence $a_n = n$ for $n = 0, 1, 2, \ldots$ does not converge.*

Proof. Suppose $a_n = n \to \alpha$, say.

Taking $\epsilon = 1/2$, we can find N such that $|a_n - \alpha| < 1/2$ for all $n \geq N$. Then

$$1 = |a_{n+1} - a_n| \leq |a_{n+1} - \alpha| + |a_n - \alpha| < 1/2 + 1/2 = 1.$$

This is a contradiction and so a_n does not converge.[1] □

Lemma 1.7 (Uniqueness of limit). *If $a_n \to a$ and $a_n \to a'$ then $a = a'$.*

Proof. Given $\epsilon > 0$ there exists N such that $n \geq N$ implies $|a_n - a| < \epsilon$ and K such that $n \geq K$ implies $|a_n - a'| < \epsilon$. Let L be the greater of N and K. Now

$$\begin{aligned} |a - a'| &= |a - a_n + a_n - a'| \\ &\leq |a - a_n| + |a_n - a'| \\ &\leq \epsilon + \epsilon = 2\epsilon. \end{aligned}$$

But $2\epsilon > 0$ is arbitrary, so $|a - a'| = 0$ and $a = a'$. □

Observation 1.8. *Suppose $a_n \to a$ and $a_n \leq \alpha$ for all (sufficiently large) n. Then $a \leq \alpha$.*

Proof. Suppose $\alpha < a$, so that $\epsilon = a - \alpha > 0$. We can find N such that $|a_n - a| < \epsilon$ for all $n \geq N$.

Consider

$$a_N - \alpha = (a_N - a) + (a - \alpha) = \epsilon + (a_N - a) \geq \epsilon - |a_n - a| > \epsilon - \epsilon = 0.$$

So $a_N > \alpha$ — a contradiction. We deduce $a \leq \alpha$. □

Example 1.9. *We "know" that $1/n \to 0$ in $\mathbb{R}$.* WHY? *There are ordered fields in which $1/n \not\to 0$ (e.g. $\mathbb{Q}(t)$, field of rational functions, ordered so that t is "infinite")*

(Easy to see that $1/n \to 0$ in $\mathbb{Q}$).

Proposition 1.10. *Suppose that $a_n \to a$ and $b_n \to b$. Then*

1. $a_n + b_n \to a + b$
2. $\lambda a_n \to \lambda a$
3. $a_n b_n \to ab$.

Proof of 1 and 2 are both trivial and are left to the reader.

Proof of 3. Given $\epsilon > 0$ take N such that $|a_n - a| < \epsilon$ for all $n \geq N$ and M such that $|b_n - b| < \min\{\epsilon, 1\}$ for all $n \geq M$. Let $K = \max\{M, N\}$. Now

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

for all $n \geq K$. Now $\epsilon(1 + |b| + |a|)$ can be made arbitrarily small and the result is proved. □

[1] This is a rigorous form of the thought—if $n \to \alpha$ we can't have both $n, n+1$ within $1/2$ of α.

1.3 Completeness of $\mathbb{R}$: Bounded monotonic sequences

Definition 1.11. *A sequence a_n is (monotonic)* increasing *just when $a_n \leq a_{n+1}$ for all n; it is (monotonic)* decreasing *just when $a_n \geq a_{n+1}$ for all n. To cover either case we say the sequence is* monotonic.

N.B. a_n is increasing iff $(-a_n)$ is decreasing.

A sequence a_n is bounded above *when there is B with $a_n \leq B$ for all n; it is* bounded below *when there is A with $a_n \geq A$ for all n; it is* bounded *when it is bounded above and below.*

Axiom (Completeness Axiom). *The real numbers $\mathbb{R}$ form an ordered field and every bounded monotonic sequence of reals has a limit (ie converges).*

Remarks.

- *This can be justified on further conditions, but here we take it as an axiom.*
- *It is enough to say an increasing sequence bounded above converges.*
- *In fact, this characterizes $\mathbb{R}$ as the completion of $\mathbb{Q}$.*

From now on, we consider only the complete ordered field $\mathbb{R}$, and occasionally its (incomplete) ordered subfield $\mathbb{Q}$.

Proposition 1.12 (Archimedean Property).

1. *For any real x, there is $N \in \mathbb{N}$ with $N > x$.*
2. *For any $\epsilon > 0$ there is $N \in N$ with $0 < \frac{1}{N} < \epsilon$.*
3. *The sequence $\frac{1}{n} \to 0$.*

Proof.

1. Recall that $a_n = n$ is an increasing non-convergent sequence. Hence it is not bounded above and so for any $x \in \mathbb{R}$ there is N with $x < N$.
2. If $\epsilon > 0$, then consider $\epsilon^{-1}(> 0)$ and take $N \in \mathbb{N}$ with $\epsilon^{-1} < N$. Then $0 < 1/N < \epsilon$
3. Given $\epsilon > 0$ we can find N with $0 < \frac{1}{N} < \epsilon$. Now if $n \geq N$,
$$0 < 1/n \leq 1/N < \epsilon$$
and the result is proved.

$\square$

Definition 1.13. *If a_n is a sequence and we have $n(k)$ for $k \in \mathbb{N}$, with*

$$n(k) < n(k+1)$$

then $(a_{n(k)})_{k\in\mathbb{N}}$ is a subsequence *of a_n.*

Observation 1.14. *Suppose $a_n \to a$ has a subsequence $(a_{n(k)})_{k\in\mathbb{N}}$. Then $a_{n(k)} \to a$ as $k \to \infty$.*

Theorem 1.15 (The Bolzano-Weierstrass Theorem). *Any bounded sequence of reals has a convergent subsequence.*

Cheap proof. Let a_n be a bounded sequence. Say that $m \in \mathbb{N}$ is a 'peak number' iff $a_m \geq a_k$ for all $k \geq m$.

Either there are infinitely many peak numbers, in which case we enumerate them $p(1) < p(2) < p(3) < \ldots$ in order. Then $a_{p(k)} \geq a_{p(k+1)}$ and so $a_{p(k)}$ is a bounded decreasing subsequence of a_n, so converges.

Or there are finitely many peak numbers. Let M be the greatest. Then for every $n > M$, n is not a peak number and so we can find $g(n) > n$: the least $r > n$ with $a_r > a_n$.

Define $q(k)$ inductively by $q(1) = M + 1$, $q(k + 1) = g(q(k))$.

By definition $q(k) < q(k + 1)$ for all k, and $a_{q(k)} < a_{q(k+1)}$ for all k, so $a_{q(k)}$ is a bounded, (strictly) increasing subsequence of a_n and so converges. $\square$

This basis of this proof is that any sequence in a total order has a monotonic subsequence.

1.4 Completeness of $\mathbb{R}$: Least Upper Bound Principle

Definition 1.16. *Let $(\emptyset \neq)S \subseteq \mathbb{R}$ be a (non-empty) set of reals.*

- b *is an* upper bound *for* S *iff* $s \leq b$ *for all* $s \in S$ *and if* S *has such, S is* bounded above.
- a *is a* lower bound *for* S *iff* $a \leq s$ *for all* $s \in S$*, and if* S *has such, S is* bounded below.
- S *is* bounded *iff* S *is bounded above and below, ie if* $S \subseteq [a, b]$ *for some* a, b*.*

b *is the* least upper bound *of* S *or the* supremum *of* S *iff*

- b *is an upper bound*
- *If* $c < b$ *then* $c < s$ *for some* $s \in S$ *(ie* c *is not an upper bound for* S*)*

Similarly, a *is the* greatest lower bound *of* S *or the* infimum *of* S *iff*

- a *is a lower bound*
- *If* $a < c$ *then* $s < c$ *for some* $s \in S$ *(ie* c *is not a lower bound).*[2]

Notation: $b = \operatorname{lub} S = \sup S; a = \operatorname{glb} S = \inf S$.

Theorem 1.17 (Least Upper Bound Principle). *A non-empty set* S *of reals which is bounded above has a least upper bound.*

Proof. Suppose $S \neq \emptyset$ and bounded above. Take b an upper bound and a (in S say) so that $[a, b] \cap S \neq \emptyset$.

Set $a_0 = a$, $b_0 = b$ so that $a_0 \leq b_0$ and define $a_n \leq b_n$ inductively as follows:

Suppose a_n, b_n given, then a_{n+1}, b_{n+1} are defined by stipulating:-

[2] Aside: If b, b' are both least upper bounds of S, then can't have $b < b'$ and can't have $b' < b$ and so $b = b'$.

- If $\left[\frac{a_n+b_n}{2}, b_n\right] \cap S \neq \emptyset$ then $a_{n+1} = \frac{a_n+b_n}{2}$, $b_{n+1} = b_n$.
- If otherwise, then $a_{n+1} = a_n, b_{n+1} = \frac{a_n+b_n}{2}$.

We can see inductively that:

1. $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$ for all n.
2. $(b_{n+1} - a_{n+1}) = \frac{1}{2}(b_n - a_n)$ for all n.
3. $[a_n, b_n] \cap S \neq \emptyset$ for all n.[3]
4. b_n is an upper bound of S for every n.[4]

By 1. b_n is decreasing, bounded below by a so $b_n \to \beta$ say; a_n is increasing, bounded above by b so $a_n \to \alpha$

By 2. $(b_n - a_n) = \frac{1}{2^n}(b_0 - a_0) \to 0$ as $n \to \infty$. But $b_n - a_n \to \beta - \alpha$ and so $\beta = \alpha$.

Claim: $\alpha = \beta$ is $\sup S$.

- Each b_n is an upper bound of S and so $\beta = \lim_{n\to\infty} b_n$ is an upper bound — for if $s \in S$ we have $s \leq b_n$ all n and so $s \leq \lim_{n\to\infty} b_n$
- Take $\gamma < \beta = \alpha = \lim_{n\to\infty} a_n$. We can take N such that $a_n > \gamma$ for all $n \geq N$.[5]

  But then $[a_N, b_N] \cap S \neq \emptyset$ and so there is $s \in S$ such that $s \geq a_n > \gamma$.

This shows that β is the least upper bound.

□

Observation 1.18. *We can deduce the completeness axiom from the LUB principle.*

Proof. If a_n is increasing and bounded above then $S = \{a_n : n \in \mathbb{N}\}$ is non-empty and bounded above and so we can set $a = \sup S$

Suppose $\epsilon > 0$ given. Now $a - \epsilon < a$ and so there is N with $a_N > a - \epsilon$ but then for $n \geq N$, $a - \epsilon < a_N \leq a_n \leq a$ and so $|a_n - a| < \epsilon$. □

[3]True for $n = 0$, and inductively, certainly true for $n + 1$ in first alternative, and in the 2nd alternative since

$$\left[\frac{a_n + b_n}{2}, b_n\right] \cap S = \emptyset$$

$$\left[a_n, \frac{a_n + b_n}{2}\right] \cap S = [a_n, b_n] \cap S \neq \emptyset$$

by induction hypothesis

[4]True for $n = 0$ and inductively, trivial in first case and in the second, clear as

$$[b_{n+1}, b_n] \cap S = \emptyset$$

[5]Let $\epsilon = \beta - \gamma > 0$. We can find N such that $|a_n - \beta| < \epsilon$ and thus $a_n > \gamma$.

1.5 Completeness of $\mathbb{R}$: General Principle of Convergence

Definition 1.19. *A real sequence a_n is a* Cauchy Sequence *if and only if for all $\epsilon > 0$ there exists N with*

$$|a_n - a_m| < \epsilon \quad \forall n, m \geq N.$$

That is a_n is Cauchy iff

$$\forall \epsilon > 0 \quad \exists N \quad \forall n, m \geq N \quad |a_n - a_m| < \epsilon$$

Observation 1.20. *A Cauchy sequence is bounded, For if a_n is Cauchy, take N such that $|a_n - a_m| < 1$ for all $n, m \geq N$. Then a_n is bounded by*

$$\pm \max(|a_1|, |a_2|, \ldots, |a_N + 1|)$$

Lemma 1.21. *Suppose a_n is Cauchy and has a convergent subsequence $a_{n(k)} \to a$ as $k \to \infty$. Then $a_n \to a$ as $n \to \infty$.*

Proof. Given $\epsilon > 0$, take N such that $|a_n - a_m| < \epsilon$ for all $m, n \geq N$, and take K with $n(K) \geq N$ (easy enough to require $K \geq N$) such that $\left|a_{n(k)} - a\right| < \epsilon$ for all $k \geq K$.

Then if $n \geq M = n(K)$

$$|a_n - a| \leq \left|a_n - a_{n(k)}\right| + \left|a_{n(k)} - a\right| < \epsilon + \epsilon = 2\epsilon.$$

But $2\epsilon > 0$ can be made arbitrarily small, so $a_n \to a$. □

Theorem 1.22 (The General Principle of Convergence). *A real sequence converges if and only if it is Cauchy.*

Proof. ($\Rightarrow$) Suppose $a_n \to a$. Given $\epsilon > 0$ take N such that $|a_n - a| < \epsilon$ for all $n \geq N$.

Then if $m, n \geq N$,

$$|a_n - a_m| \leq |a_n - a| + |a_m - a| \leq \epsilon + \epsilon = 2\epsilon.$$

As $2\epsilon > 0$ can be made arbitrarily small, a_n is Cauchy.

($\Leftarrow$) Suppose a_n is Cauchy.[6] Then a_n is bounded and so we can apply Bolzano-Weierstrass to obtain a convergent subsequence $a_{n(k)} \to a$ as $k \to \infty$. By lemma 1.21, $a_n \to a$.

□

Alternative Proof. Suppose a_n is Cauchy. Then it is bounded, say $a_n \in [\alpha, \beta]$

Consider

$$S = \{s : a_n \geq s \text{ for infinitely many } n\}.$$

First, $\alpha \in S$ and so $S \neq \emptyset$. S is bounded above by $\beta + 1$ (in fact by β). By the LUB principle we can take $a = \sup S$.

[6] This second direction contains the completeness information.

Given $\epsilon > 0$, $a-\epsilon < a$ and so there is $s \in S$ with $a-\epsilon < s$. Then there are infinitely many n with $a_n \geq s > a - \epsilon$. $a + \epsilon > a$, so $a + \epsilon \notin S$ and so there are only finitely many n with $a_n \geq a + \epsilon$. Thus there are infinitely many n with $a_n \in (a - \epsilon, a + \epsilon)$.

Take N such that $|a_n - a_m| < \epsilon$ for all $m, n \geq N$. We can find $m \geq N$ with $a_m \in (a - \epsilon, a + \epsilon)$ ie $|a_m - a| < \epsilon$. Then if $n \geq N$,

$$|a_n - a| \leq |a_n - a_m| + |a_m - a| < \epsilon + \epsilon = 2\epsilon$$

As 2ϵ can be made arbitrarily small this shows $a_n \to a$. □

Remarks.

- *This second proof can be modified to give a proof of Bolzano-Weierstrass from the LUB principle.*
- *In the proof by bisection of the LUB principle, we could have used GPC (general principle of convergence) instead of Completeness Axiom.*
- *We can prove GPC directly from completeness axiom as follows:*

  Given a_n Cauchy, define

  $$b_n = \inf\{a_m : m \geq n\}$$

  b_n is increasing, so $b_n \to b$ $(= \liminf a_n)$. Then show $a_n \to b$.
- *The Completeness Axiom, LUB principle, and the GPC are equivalent expressions of the completeness of $\mathbb{R}$.*

Chapter 2

Euclidean Space

2.1 The Euclidean Metric

Recall that $\mathbb{R}^n$ is a vector space with coordinate-wise addition and scalar multiplication.

Definition 2.1. *The* Euclidean norm[1] $\|\cdot\| : \mathbb{R}^n \mapsto \mathbb{R}$ *is defined by*

$$\|x\| = \|(x_1, \ldots, x_n)\| = +\sqrt{\sum_{i=1}^{n} x_i^2}$$

and the Euclidean distance $d(x, y)$ *between* x *and* y *is* $d(x, y) = \|x - y\|$.

Observation 2.2. *The norm satisfies*

$$\begin{gathered}\|x\| \geq 0, \qquad \|x\| = 0 \Leftrightarrow x = 0 \in \mathbb{R}^n \\ \|\lambda x\| = |\lambda| \ \|x\| \\ \|x + y\| \leq \|x\| + \|y\|\end{gathered}$$

and the distance satisfies

$$\begin{gathered}d(x, y) \geq 0, \qquad d(x, y) = 0 \Leftrightarrow x = y \\ d(x, y) = d(y, x) \\ d(x, z) \leq d(x, y) + d(y, z).\end{gathered}$$

2.2 Sequences in Euclidean Space

We can write $x^{(n)}$ or $x(n)$ for a sequence of points in $\mathbb{R}^p$. Then

$$x_i^{(n)} = x_i(n) \qquad 1 \leq i \leq p$$

for the i^{th} coordinate of the n^{th} number of the sequence.

[1] The norm arises from the standard inner product

$$< x, y >= \sum_{i=1}^{n} x_i y_i$$

Definition 2.3. *A sequence $x^{(n)}$ converges to x in $\mathbb{R}^p$ when for any $\epsilon > 0$ there exists N such that* [2]

$$\left\|x^{(n)} - x\right\| < \epsilon \quad \textit{for all} \quad n \geq N$$

In symbols:

$$x^{(n)} \to x \Leftrightarrow \forall \epsilon > 0 \quad \exists N \quad \forall n \geq N \quad \left\|x^{(n)} - x\right\| < \epsilon$$

Proposition 2.4. *$x^{(n)} \to x$ in $\mathbb{R}^p$ iff $x_i^{(n)} \to x$ in $\mathbb{R}$ for $1 \leq i \leq p$.*

Proof. Note that

$$0 < \left|x_i^{(n)} - x_i\right| \leq \left\|x^{(n)} - x\right\| \to 0$$

and

$$0 \leq \left\|x^{(n)} - x\right\| \leq \sum_{i=1}^{p} \left|x_i^{(n)} - x_i\right| \to 0.$$

□

Definition 2.5. *A sequence $x^{(n)} \in \mathbb{R}^p$ is bounded if and only if there exists R such that $\left\|x^{(n)}\right\| \leq R$ for all n.*

Theorem 2.6 (Bolzano-Weierstrass Theorem for $\mathbb{R}^p$). *Any bounded sequence in $\mathbb{R}^p$ has a convergent subsequence.*

Proof (Version 1). Suppose $x^{(n)}$ is bounded by R. Then all the coordinates $x_i^{(n)}$ are bounded by R. By Bolzano-Weierstrass in $\mathbb{R}$ we can take a subsequence such that the 1st coordinates converge; now by Bolzano-Weierstrass we can take a subsequence of this sequence such that the 2nd coordinates converge. Continuing in this way (in p steps) we get a subsequence all of whose coordinates converge. But then this converges in $\mathbb{R}^p$. □

Version 2. By induction on p. The result is known for $p = 1$ (Bolzano-Weierstrass in $\mathbb{R}$) and is trivial for $p = 0$. Suppose result is true for p.

Take x^n a bounded subsequence in $\mathbb{R}^p$ and write each $x^{(n)}$ as $x^{(n)} = (y^{(n)}, x_{p+1}^{(n)})$ where $y^{(n)} \in \mathbb{R}^p$ and $x_{p+1}^{(n)} \in \mathbb{R}$ is the $(p+1)^{\text{th}}$ coordinate.

Now $y^{(n)}$ and $x_{p+1}^{(n)}$ are both bounded, so we can apply Bolzano-Weierstrass in $\mathbb{R}^p$ to get a subsequence $y^{(n(k))} \to y$. Apply Bolzano-Weierstrass in $\mathbb{R}$ to get $x_{p+1}^{(n(k(j)))} \to x$. Then

$$x^{(n(k(j)))} \to (y, x) \text{ as } j \to \infty.$$

□

[2] $x^n \to x$ in $\mathbb{R}^p$ iff $\left\|x^{(n)} - x\right\| \to 0$ in $\mathbb{R}$.

Definition 2.7. *A sequence $x^{(n)} \in \mathbb{R}^p$ is a* Cauchy sequence *iff for any $\epsilon > 0$ there is N with $\left\|x^{(n)} - x^{(m)}\right\| < \epsilon$ for $n, m \geq N$. In symbols this is*

$$\forall \epsilon > 0 \quad \exists N \quad \forall n, m \geq N \qquad \left\|x^{(n)} - x^{(m)}\right\| < \epsilon.$$

Observation 2.8. *$x^{(n)}$ is Cauchy in $\mathbb{R}^p$ iff each $x_i^{(n)} = x_i(n)$ is Cauchy in $\mathbb{R}$ for $1 \leq i \leq p$.*

Proof. Suppose $x^{(n)}$ is Cauchy. Take $1 \leq i \leq p$. Given $\epsilon > 0$, we can find N such that $\left\|x^{(n)} - x^{(m)}\right\| < \epsilon$ for all $n, m \geq N$. But then for $n, m \geq N$,

$$|x_i(n) - x_i(m)| \leq \left\|x^{(n)} - x^{(m)}\right\| < \epsilon$$

so as $\epsilon > 0$ is arbitrary, $x_i(n)$ is Cauchy.

Conversely, suppose each $x_i(n)$ is Cauchy for $1 \leq i \leq p$. Given $\epsilon > 0$, we can find $N_1, \ldots, N_p$ such that

$$|x_i(n) - x_i(m)| < \epsilon \quad for \quad n, m \geq N_i \qquad (1 \leq i \leq p)$$

Now if $n, m \geq N = \max\{N_1, \ldots, N_p\}$ then

$$\left\|x^{(n)} - x^{(m)}\right\| \leq \sum_{i=1}^{p} \left|x_i^{(n)} - x_i^{(m)}\right| < p\epsilon$$

As $p\epsilon$ can be made arbitrarily small, $x^{(n)}$ is Cauchy. □

Theorem 2.9 (General Principle of Convergence in $\mathbb{R}^p$). *A sequence $x^{(n)}$ in $\mathbb{R}^p$ is convergent if and only if $x^{(n)}$ is Cauchy.*

Proof. $x^{(n)}$ converges in $\mathbb{R}^p$
iff $x_i(n)$ converges in $\mathbb{R}$ $(1 \leq i \leq p)$
iff $x_i(n)$ is Cauchy in $\mathbb{R}$ $(1 \leq i \leq p)$
iff $x^{(n)}$ is Cauchy in $\mathbb{R}^p$. □

2.3 The Topology of Euclidean Space

For $a \in \mathbb{R}^p$ and $r \geq 0$ we have the *open ball* $B(a, r) = O(a, r)$, defined by

$$B(a, r) = O(a, r) = \{x : \|x - a\| < r\}$$

Also we have the *closed ball* $C(a, r)$ defined by

$$C(a, r) = \{x : \|x - a\| \leq r\}$$

Also we shall sometimes need the *"punctured" open ball*

$$\{x : 0 < \|x - a\| < r\}$$

Definition 2.10. *A subset $U \subseteq \mathbb{R}^p$ is* open *if and only if for all $a \in U$ there exists $\epsilon > 0$ such that*

$$\|x - a\| < \epsilon \Rightarrow x \in U$$

[That is: U is open iff for all $a \in U$ there exists $\epsilon > 0$ with $B(a, \epsilon) \subseteq U$].

The empty set $\emptyset$ is trivially open.

Example 2.11.

- *$O(a, r)$ is open, for if $b \in O(a, r)$, then $\|b - a\| < r$, setting*

$$\epsilon = r - \|b - a\| > 0$$

we see $O(b, \epsilon) \subseteq O(a, r)$.

- *Similarly $\{x : 0 < \|x - a\| < r\}$ is open.*
- *But $C(a, r)$ is not open for any $r \geq 0$.*

Definition 2.12. *A subset $A \subseteq \mathbb{R}^p$ is* closed *iff whenever a_n is a sequence in A and $a_n \to a$, then $a \in A$. In symbols this is*

$$a_n \to a, a_n \in A \Rightarrow a \in A$$

Example 2.13.

- *$C(a, r)$ is closed, for suppose $b_n \to b$ and $b_n \in C(a, r)$ then $\|b_n - a\| \leq r$ for all n. Now*

$$\|b - a\| \leq \|b_n - b\| + \|b_n - a\| \leq r + \|b_n - b\|$$

As $b_n \to b$, $\|b_n - b\| \to 0$, and so $r + \|b_n - b\| \to r$ as $n \to \infty$. Therefore $\|b - a\| \leq r$.

- *A product $[a_1, b_1] \times \ldots \times [a_p, b_p] \subseteq \mathbb{R}^p$ of closed intervals is closed. For if $c^{(n)} \to c$ and*

$$c^{(n)} \in [\,] \times \ldots \times [\,]$$

then each $c_i^{(n)} \to c_i$ with $c_i^{(n)} \in [a_i, b_i]$ so that $c_i \in [a_i, b_i]$. Therefore

$$c \in [\,] \times \ldots \times [\,].$$

- *But $O(a, r)$ is not closed unless $r = 0$.*

Proposition 2.14. *A set $U \subseteq \mathbb{R}^p$ is open (in $\mathbb{R}^p$) iff its complement $\mathbb{R}^p \setminus U$ is closed in $\mathbb{R}^p$. A set $U \subseteq \mathbb{R}^p$ is closed (in $\mathbb{R}^p$) iff its complement $\mathbb{R}^p \setminus U$ is open in $\mathbb{R}^p$.*[3]

Proof. Exercise. □

2.4 Continuity of Functions

We consider functions $f : E \mapsto \mathbb{R}^m$ defined on some $E \subseteq \mathbb{R}^n$. For now imagine that E is a simple open or closed set as in §2.3.

[3]Warning: Sets need not be either open or closed: the half open interval $(a, b]$ is neither open nor closed in $\mathbb{R}$.

Definition 2.15. *Suppose $f\colon E \mapsto \mathbb{R}^m$ (with $E \subseteq \mathbb{R}^n$) Then f is* continuous at a *iff for any $\epsilon > 0$ there exists*[4] $\delta > 0$ *such that*

$$\|x-a\| < \delta \to \|f(x) - f(a)\| < \epsilon \quad \textit{for all} \quad x \in E.$$

In symbols:

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in E \quad \|x-a\| < \delta \Rightarrow \|f(x) - f(a)\| < \epsilon.$$

f is continuous *iff f is continuous at every point.*

This can be reformulated in terms of limit notation as follows:

Definition 2.16. *Suppose $f\colon E \mapsto \mathbb{R}^n$. Then $f(x) \to b$ as $x \to a$ in E*[5] *if an only if for any $\epsilon > 0$ there exists $\delta > 0$ such that*

$$0 < \|x-a\| < \delta \Rightarrow \|f(x) - b)\| < \epsilon \quad \textit{for all} \quad x \in E.$$

Remarks.

- *We typically use this when E is open and some punctured ball*

$$\{x : 0 < \|x-a\| < r\}$$

  is contained in E. Then the limit notion is independent of E.

- *If $f(x) \to b$ as $x \to a$, then defining $f(a) = b$ extends f to a function continuous at a.*

Proposition 2.17. *Suppose $f\colon E \mapsto \mathbb{R}^m$*

- *f is continuous (in E) if and only if whenever $a_n \to a$ in E, then $f(a_n) \to f(a)$. This is known as* sequential continuity.

- *f is continuous (in E) if and only if for any open subset $V \subseteq \mathbb{R}^m$:*

$$F^{-1}(V) = \{x \in E : f(x) \in V\}$$

  is open in E.

Proof. We will only prove the first part for now. The proof of the second part is given in theorem 5.16 in a more general form.

Assume f is continuous at a and take a convergent sequence $a_n \to a$ in E. Suppose $\epsilon > 0$ given. By continuity of f, there exists $\delta > 0$ such that

$$\|x-a\| < \delta \Rightarrow \|f(x) - f(a)\| < \epsilon.$$

As $a_n \to a$ take N such that $\|a_n - a\| < \delta$ for all $n \geq N$.

Now if $n \geq N$, $\|f(a_n) - f(a)\| < \epsilon$. Since $\epsilon > 0$ can be made arbitrarily small, $f(a_n) \to f(a)$.

The converse is clear. □

[4]The continuity of f at a depends only on the behavior of f in an open ball $B(a,r)$, $r > 0$.

[5]Then f is continuous at a iff $f(x) \to f(a)$ as $x \to a$ in E.

Remark. *$f(x) \to b$ as $x \to a$ iff $\|f(x) - b\| \to 0$ as $x \to a$.*

Observation 2.18.

- *Any linear map $\alpha\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ is continuous.*

  Proof. If α has matrix $A = (a_{ij})$ with respect to the standard basis then

  $$\alpha(x) = \alpha(x_1, \ldots, x_n) = \left(\sum_{j=1}^{n} a_{ij}x_j, \ldots, \sum_{j=1}^{n} a_{mj}x_j\right)$$

  and so

  $$\|\alpha(x)\| \leq \sum_{ij} |a_{ij}|\,|x_j| \leq \underbrace{\left(\sum_{i,j} |a_{ij}|\right)}_{K} \|x\|\,.$$

  Fix $a \in \mathbb{R}^n$. Given $\epsilon > 0$ we note that if $\|x - a\| < \epsilon$ then

  $$\|\alpha(x) - \alpha(a)\| = \|\alpha(x - a)\| \leq K\,\|x - a\| < K\epsilon$$

  As $K\epsilon$ can be made arbitrarily small, f is continuous at a. But $a \in \mathbb{R}^n$ arbitrary, so f is continuous. □

- *If $f\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ is continuous at a, and $g\colon \mathbb{R}^m \mapsto \mathbb{R}^p$ is continuous at $f(a)$, then $g \circ f\colon \mathbb{R}^n \mapsto \mathbb{R}^p$ is continuous at a.*

  Proof. Given $\epsilon > 0$ take $\eta > 0$ such that

  $$\|y - f(a)\| < \eta \Rightarrow \|g(y) - g(f(a))\| < \epsilon.$$

  Take $\delta > 0$ such that $\|x - a\| < \delta \Rightarrow \|f(x) - f(a)\| < \eta$.
  Then $\|x - a\| < \delta \Rightarrow \|g(f(x)) - g(f(a))\| < \epsilon$. □

Proposition 2.19. *Suppose $f, g\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ are continuous at a. Then*

1. *$f + g$ is continuous at a.*
2. *λf is continuous at a, any $\lambda \in \mathbb{R}$.*
3. *If $m = 1$, $f \cdot g$ is continuous at a.*

Proof. Proof is trivial. Just apply propositions 1.10 and 2.17. □

Suppose $f\colon \mathbb{R}^n \mapsto \mathbb{R}^m$. Then we can write:

$$f(x) = (f_1(x), \ldots, f_m(x))$$

where $f_j\colon \mathbb{R}^n \mapsto \mathbb{R}$ is f composed with the j^{th} projection or coordinate function.

Then f is continuous if and only if each $f_1, \ldots, f_m$ is continuous.

Theorem 2.20. *Suppose that $f\colon E \mapsto \mathbb{R}$ is continuous on E, a closed and bounded subset of $\mathbb{R}^n$. Then f is bounded and (so long as $E \neq \emptyset$) attains its bounds.*

Proof. Suppose f not bounded. Then we can take $a_n \in E$ with $|f(a_n)| > n$. By Bolzano-Weierstrass we can take a convergent subsequence $a_{n(k)} \to a$ as $k \to \infty$ and as E is closed, $a \in E$.

By the continuity of f, $f(a_{n(k)}) \to f(a)$ as $k \to \infty$. But $f(a_{n(k)})$ is unbounded — a contradiction and so f is bounded.

Now suppose $\beta = \sup\{f(x)\colon x \in E\}$. We can take $c_n \in E$ with

$$|f(c_n) - \beta| < \frac{1}{n}.$$

By Bolzano-Weierstrass we can take a convergent subsequence $c_{n(k)} \to c$. As E is closed, $c \in E$. By continuity of f, $f(c_{n(k)}) \to f(c)$, but by construction $f(c_{n(k)}) \to \beta$ as $k \to \infty$. So $f(c) = \beta$. □

Essentially the same argument shows the more general fact. If $f\colon E \mapsto \mathbb{R}^n$ is continuous in E, closed and bounded, then the image $f(E)$ is closed and bounded. N.B. *compactness*.

2.5 Uniform Continuity

Definition 2.21. *Suppose $f\colon E \mapsto \mathbb{R}^m$ where $E \subseteq \mathbb{R}^n$. f is* uniformly continuous *on E iff for any $\epsilon > 0$ there exists $\delta > 0$ such that*

$$\|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon \quad \textit{for all} \quad x, y \in E.$$

In symbols:

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x, y \in E \quad \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon$$

Compare this with the definition of continuity of f at all points $x \in E$:

$$\forall x \in E \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall y \in E \quad \|x - y\| < \delta \Rightarrow \|f(x) - f(y)\| < \epsilon$$

The difference is that for continuity, the $\delta > 0$ to be found depends on both x and $\epsilon > 0$; for uniform continuity the $\delta > 0$ depends only on $\epsilon > 0$ and is independent of x.

Example 2.22. $x \mapsto x^{-1} : (0, 1] \mapsto [1, \infty)$ *is continuous but* not *uniformly continuous.*

Consider

$$\left|\frac{1}{x} - \frac{1}{y}\right| = \left|\frac{y - x}{xy}\right|$$

Take $x = \eta$, $y = 2\eta$. Then

$$\left|\frac{1}{x} - \frac{1}{y}\right| = \left|\frac{1}{2\eta}\right|$$

while $|x - y| = \eta$.

Theorem 2.23. *Suppose $f\colon E \mapsto \mathbb{R}^m$ is continuous on E, a closed and bounded subset of $\mathbb{R}^n$. Then f is uniformly continuous on E.*

Proof. Suppose f continuous but not uniformly continuous on E.

Then there is some $\epsilon > 0$ such that for no $\delta > 0$ is it the case that

$$\|x-y\| < \delta \Rightarrow \|f(x)-f(y)\| < \epsilon \quad \forall\, x, y \in E.$$

Therefore[6] for every $\delta > 0$ there exist $x, y \in E$ with $\|x-y\| < \delta$ and

$$\|f(x)-f(y)\| \geq \epsilon.$$

Now for every $n \geq 1$ we can take $x_n, y_n \in E$ with $\|x_n - y_n\| < \frac{1}{n}$ and

$$\|f(x_n)-f(y_n)\| \geq \epsilon.$$

By Bolzano-Weierstrass, we can take a convergent subsequence $x_{n(k)} \to x$ as $k \to \infty$. $x \in E$ since E is closed.

Now

$$\left\|y_{n(k)} - x\right\| \leq \left\|y_{n(k)} - x_{n(k)}\right\| + \left\|x_{n(k)} - x\right\| \to 0 \text{ as } k \to 0.$$

Hence $y_{n(k)} \to x$. So $x_{n(k)} - y_{n(k)} \to 0$ as $k \to \infty$ and so $f(x_{n(k)}) - f(y_{n(k)}) \to 0$ (by continuity of f). So

$$\underbrace{\left\|f(x_{n(k)}) - f(y_{n(k)})\right\|}_{\geq \epsilon} \to 0 \text{ as } k \to \infty.$$

This is a contradiction and it follows that f must be uniformly continuous. $\square$

[6]We want the "opposite" of

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x, y \in E \quad \|x-y\| < \delta \Rightarrow \|f(x)-f(y)\| < \epsilon.$$

It is:

$$\exists \epsilon > 0 \quad \forall \delta > 0 \quad \exists x, y \in E \quad \|x-y\| < \delta \quad \text{and} \quad \|f(x)-f(y)\| \geq \epsilon.$$

Chapter 3

Differentiation

3.1 The Derivative

Definition 3.1. *Let $f\colon E \mapsto \mathbb{R}^m$ be defined on E, an open subset of $\mathbb{R}^n$. Then f is* differentiable *at $a \in E$ with derivative $Df_a \equiv f'(a) \in L(\mathbb{R}^n, \mathbb{R}^m)$ when*

$$\frac{\|f(a+h) - f(a) - f'(a)h\|}{\|h\|} \to 0 \text{ as } h \to 0.$$

The idea is that the best linear approximation to f at $a \in E$ is the affine map $x \mapsto a + f'(a)(x-a)$.

Observation 3.2 (Uniqueness of derivative). *If f is differentiable at a then its derivative is unique.*

Proof. Suppose $Df_a, \hat{Df_a}$ are both derivatives of f at a. Then

$$\frac{\left\|Df_a(h) - \hat{Df}_a(h)\right\|}{\|h\|} \leq$$
$$\frac{\|f(a+h) - f(a) - Df_a(h)\|}{\|h\|} + \frac{\left\|f(a+h) - f(a) - \hat{Df}_a(h)\right\|}{\|h\|} \to 0.$$

Thus

$$LHS = \left\|(Df_a - \hat{D}f_a)\left(\frac{h}{\|h\|}\right)\right\| \to 0 \text{ as } h \to 0.$$

This shows that $Df_a - \hat{D}f_a$ is zero on all unit vectors, and so $Df_a \equiv \hat{D}f_a$. □

Proposition 3.3. *If $f\colon E \mapsto \mathbb{R}^m$ is differentiable at a, then f is continuous at a.*

Proof. Now

$$\|f(x) - f(a)\| \leq \|f(x) - f(a) - Df_a(x-a)\| + \|Df_a(x-a)\|$$

But $\|f(x) - f(a) - Df_a(x-a)\| \to 0$ as $x \to a$ and $\|Df_a(x-a)\| \to 0$ as $x \to a$ and so the result is proved. $\square$

Proposition 3.4 (Differentiation as a linear operator). *Suppose that $f, g\colon E \mapsto \mathbb{R}^n$ are differentiable at $a \in E$. Then*

1. *$f+g$ is differentiable at a with $(f+g)'(a) = f'(a) + g'(a)$;*

2. *λf is differentiable at a with $(\lambda f)'(a) = \lambda f'(a)$ for all $\lambda \in \mathbb{R}$.*

Proof. Exercise. $\square$

Observation 3.5 (Derivative of a linear map). *If $\alpha\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ is a linear map, then α is differentiable at every $a \in \mathbb{R}^n$ with $\alpha'(a) = \alpha(a)$.*

Proof is simple, note that

$$\frac{\|\alpha(a+h) - \alpha(a) - \alpha(h)\|}{\|h\|} \equiv 0.$$

Observation 3.6 (Derivative of a bilinear map). *If $\beta\colon \mathbb{R}^n \times \mathbb{R}^m \mapsto \mathbb{R}^p$ is bilinear then β is differentiable at each $(a,b) \in \mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ with*

$$\beta'(a,b)(h,k) = \beta(h,b) + \beta(a,k)$$

Proof.

$$\frac{\|\beta(a+h, b+k) - \beta(a,b) - \beta(h,b) - \beta(a,k)\|}{\|(h,k)\|} = \frac{\|\beta(h,k)\|}{\|(h,k)\|}$$

If β is bilinear then there is (b_k^{ij}) such that

$$\beta(h,k) = \left(\sum_{i=1,j=1}^{n,m} b_1^{ij} h_i k_j, \ldots, \sum_{i=1,j=1}^{n,m} b_p^{ij} h_i k_j \right)$$

$$\|\beta(h,k)\| \leq \sum_{i,j,k} |b_k^{ij}|\, |h_i|\, |k_j| \leq \underbrace{\sum_{i,j,k} |b_n^{ij}|}_{=K} \|h\|\, \|k\| \leq \frac{K}{2}(\|h\|^2 + \|k\|^2)$$

So $\frac{\|\beta(h,k)\|}{\|(h,k)\|} \leq \left(\frac{K}{2}\right) \|(h,k)\|$ and so $\to 0$ as $(h,k) \to 0$. $\square$

Example 3.7. *The simplest bilinear map is multiplication $m\colon \mathbb{R} \times \mathbb{R} \mapsto \mathbb{R}$ and we have*

$$m'(a,b)(h,k) = hb + ak (= bh + ak).$$

3.2 Partial Derivatives

Example 3.8 (Derivative of a function $\mathbb{R} \mapsto \mathbb{R}$). *Suppose $f\colon E \mapsto \mathbb{R}$ with $E \subseteq \mathbb{R}$ open is differentiable at $a \in E$. Then the derivative map $f'(a) \in L(\mathbb{R}, \mathbb{R})$ and any such is multiplication by a scalar, also called the derivative $f'(a) = \frac{\mathrm{d}f}{\mathrm{d}x}\Big|_a$.*

Now

$$\frac{|f(a+h) - f(a) - f'(a)h|}{|h|} = \left| \frac{f(a+h) - f(a)}{h} - f'(a) \right| \to 0 \quad as \quad h \to 0$$

we see that

$$f'(a) = \lim_{h\to 0} \frac{f(a+h) - f(a)}{h}.$$

WARNING: This limit formula only makes sense in this case

Definition 3.9 (Partial derivatives). *Suppose $f\colon E \mapsto \mathbb{R}$ with $E \subseteq \mathbb{R}^n$ open. Take $a \in E$. For each $1 \le i \le n$ we can consider the function*

$$x_i \mapsto f(a_1, \ldots, a_{i-1}, x_i, a_{i+1}, \ldots, a_n)$$

which is a real-valued function defined at least on some open interval containing a_i.

If this is differentiable at a_i we write

$$D_i f(a) = \frac{\partial f}{\partial x_i}\bigg|_a$$

for its derivative—the i^{th} partial derivative of f at a.

Now suppose f is differentiable at a with derivative $f'(a) \in L(\mathbb{R}^n, \mathbb{R})$. From linear maths, any such linear map is uniquely of the form

$$(h_1, \ldots, h_n) \mapsto \sum_{i=1}^n t_i h_i$$

for $t_1, \ldots, t_n \in \mathbb{R}$ (the coefficients w.r.t. the standard basis). Therefore

$$\frac{|f(a+h) - f(a) - \sum t_i h_i|}{\|h\|} \to 0$$

as $h \to 0$. Specialize to $h = (0, \ldots, 0, h_i, 0, \ldots, 0)$. We get

$$\frac{|f(a_1, \ldots, a_{i-1}, a_i + h, a_{i+1}, \ldots, a_n) - f(a_1, \ldots, a_n) - t_i h_i|}{|h_i|} \to 0 \text{ as } h_i \to 0.$$

It follows that $t_i = D_i f(a) \equiv \frac{\partial f}{\partial x_i}\Big|_a$ and thus the coefficients of $f'(a)$ are the partial derivatives.

Example 3.10. *$m(x, y) = xy$. Then*

$$\frac{\partial m}{\partial x} = y, \frac{\partial m}{\partial y} = x$$

and

$$m'(x, y)(h, k) = \frac{\partial m}{\partial x} h + \frac{\partial m}{\partial y} k = yh + xk$$

Proposition 3.11. *Suppose $f\colon E \mapsto \mathbb{R}^m$ with $E \subseteq \mathbb{R}^n$ open. We can write*

$$f = (f_1, \dots, f_m),$$

where $f_j\colon E \mapsto \mathbb{R}$, $1 \leq j \leq m$. Then f is differentiable at $a \in E$ if and only if all the f_i are differentiable at $a \in E$. Then

$$f'(a) = (f_1'(x), \dots, f_m'(x)) \in L(\mathbb{R}^n, \mathbb{R}^m)$$

Proof. If f is differentiable with $f'(a) = ((f'(a))_1, \dots, (f'(a))_m)$ then

$$\frac{|f_j(a+h) - f_j(a) - (f'(a))_j(h)|}{\|h\|} \leq \frac{\|f(a+h) - f(a) - f'(a)(h)\|}{\|h\|} \to 0.$$

So $(f'(a))_j$ is the derivative $f_j'(a)$ at a. Conversely, if all the f_j's are differentiable, then

$$\frac{\|f(a+h) - f(a) - (f_1'(a)(h), \dots, f_m'(a)(h))\|}{\|h\|} \\ \leq \sum_{j=1}^{m} \frac{\left|f_j(a+h) - f_j(a) - f_j'(a)(h)\right|}{\|h\|} \to 0 \text{ as } h \to 0.$$

Therefore f is differentiable with the required derivative. □

It follows that if f is differentiable at a, then $f'(a)$ has the matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

all evaluated at a with respect to the standard basis.

Remark. *If the $\frac{\partial f_j}{\partial x_i}$ are continuous at a then $f'(a)$ exists.*

3.3 The Chain Rule

Theorem 3.12 (The Chain Rule). *Suppose $f\colon \mathbb{R}^n \mapsto \mathbb{R}^m$ is differentiable at $a \in \mathbb{R}^n$ and $g\colon \mathbb{R}^m \mapsto \mathbb{R}^p$ is differentiable at $b = f(a) \in \mathbb{R}^m$, then $g \circ f\colon \mathbb{R}^n \to \mathbb{R}^p$ is differentiable at $a \in \mathbb{R}^n$ and $(g \circ f)'(a) = g'(f(a)) \circ f'(a)$.*

Proof. Let $f(a+h) = f(a) + f'(a)(h) + R(a,h)$, where

$$\frac{\|R(a,h)\|}{\|h\|} \to 0 \text{ as } h \to 0.$$

We also have $g(b+k) = g(b) + g'(b)(k) + S(b,k)$, where $S(b,k) \to 0$ in the same manner.

We can now define $\sigma(b,k) = \frac{S(b,k)}{\|k\|}$ for $k \neq 0$, and $\sigma(b,k) = 0$ otherwise, so that $\sigma(b,k)$ is continuous at $k = 0$.

Now

$$\begin{aligned} g(f(a+h)) &= g(f(a) + f'(a)(h) + R(a,h)) \\ &= g(f(a)) + g'(f(a))(f'(a)(h) + R(a,h)) \\ &\quad + \sigma(f(a), f'(a)(h) + R(a,h)) \, \|f'(a)(h) + R(a,h)\| \\ &= g(f(a)) + g'(f(a))(f'(a)(h)) + g'(f(a))(R(a,h)) + Y \end{aligned}$$

as $g'(f(a))$ is linear. So it remains to show that

$$\frac{g'(f(a))(R(a,h)) + Y}{\|h\|} \to 0 \text{ as } h \to 0.$$

1.

$$\frac{g'(f(a))(R(a,h))}{\|h\|} = g'(f(a)) \left(\frac{R(a,h)}{\|h\|} \right)$$

but as $h \to 0$, $\frac{R(a,h)}{\|h\|} \to 0$, and since $g'(f(a))$ is continuous

$$\frac{g'(f(a))(R(a,h))}{\|h\|} \to 0 \text{ as } h \to 0.$$

2.

$$\frac{\|f'(a)(h)\|}{\|h\|} \le K \frac{\|h\|}{\|h\|} = K$$

as $f'(a)$ is linear (and K is the sum of the norms of the entries in the matrix $f'(a)$). Also

$$\frac{\|R(a,h)\|}{\|h\|} \to 0 \text{ as } h \to 0$$

so we can find $\delta > 0$ such that $0 < \|h\| < \delta \Rightarrow \frac{\|R(a,h)\|}{\|h\|} < 1$. Therefore, if $0 \le \|h\| < \delta$ then

$$\frac{\|f'(a)(h) + R(a,h)\|}{\|h\|} < K + 1$$

Hence $f'(a)(h) + R(a,h) \to 0$ as $h \to 0$ and so $\sigma(f(a), f'(a)(h) + R(a,h)) \to 0$ as $h \to 0$. Thus $\frac{Y}{\|h\|} \to 0$ as $h \to 0$.

□

Remark. *In the 1-D case it is tempting to write $f(a) = b, f(a+h) = b + k$ and then consider*

$$\lim_{h \to 0} \frac{g(f(a+h)) - g(f(a))}{h} = \lim_{h \to 0} \frac{g(b+k) - g(b)}{k} \frac{f(a+h) - f(a))}{h}.$$

But k could be zero. The introduction of σ is for the analogous problem in many variables.

Application. *Suppose $f, g \colon \mathbb{R}^n \mapsto \mathbb{R}$ are differentiable at a. Then their product $(f \cdot g) \colon \mathbb{R}^n \mapsto \mathbb{R}$ is differentiable at a, with derivative*[1]

$$(f \cdot g)'(a)(h) = g(a) \cdot f'(a)(h) + f(a) \cdot g'(a)(h)$$

[1] multiplication in $\mathbb{R}$ is commutative!

3.4 Mean Value Theorems

Suppose $f\colon [a,b] \mapsto \mathbb{R}$ is continuous on the (closed, bounded) interval $[a,b]$ and differentiable on (a,b). Then we have both Rolle's theorem and the mean value theorem.

Theorem 3.13 (Rolle's Theorem). *If $f(a) = f(b)$ then there exists $c \in (a,b)$ with $f'(c) = 0$.*

Proof. Either f is constant and the result is then trivial, or else without loss of generality, f takes values greater than $f(a) = f(b)$. Then there exists $c \in (a,b)$ such that $f(c) = \sup\{f(t) : t \in [a,b]\}$. Thus $f'(c) = 0$.[2] □

Theorem 3.14 (Mean Value Theorem). *Suppose $f\colon [a,b] \mapsto \mathbb{R}(a < b)$ is continuous and differentiable on (a,b). Then there exists $c \in (a,b)$ with*

$$\frac{f(b) - f(a)}{b - a} = f'(c).$$

Proof. Set $g(x) = f(x) - \frac{x-a}{b-a}(f(b) - f(a))$. Then $g(a) = f(a) = g(b)$ so we can apply Rolle's theorem to get $c \in (a,b)$ with $g'(c) = 0$. This c does the trick. □

Theorem 3.15. *Suppose that $f\colon E \mapsto \mathbb{R}^m$ (E open in $\mathbb{R}^n$) is such that the partial derivatives*

$$D_i f_j(x) = \frac{\partial f_j}{\partial x_i}$$

evaluated at x (exist and) are continuous in E. Then f is differentiable in E.

Proof. Note that since f is differentiable iff each f_j is differentiable $(1 \leq j \leq m)$, it is sufficient to consider the case $f\colon E \mapsto \mathbb{R}$. Take $a = (a_1, \dots, a_n) \in E$.

For $h = (h_1, \dots, h_n)$ write $h(r) = (h_1, \dots, h_r, 0, \dots, 0)$. Then by the MVT we can write

$$f(a + h(r)) - f(a + h(r-1)) = h_r D_r f(\xi_r)$$

where ξ_r lies in the "interval" $(a + h(r-1), a + h(r))$. Summing, we get

$$f(a + h) - f(a) = \sum_{i=1}^{n} D_i f(\xi_i) h_i.$$

Hence

$$\begin{aligned}\frac{\left|f(a+h) - f(a) - \sum_{i=1}^n D_i f(a) h_i\right|}{\|h\|} &= \frac{\left|\sum_{i=1}^n (D_i f(\xi_i) - D_i f(a)) h_i\right|}{\|h\|} \\ &\leq \sum_{i=1}^{n} \left|D_i f(\xi_i) - D_i f(a)\right|.\end{aligned}$$

As $h \to 0$, the $\xi_i \to a$ and so by the continuity of the $D_i f$'s the RHS $\to 0$ and so the LHS $\to 0$ as required. □

[2] This requires proof, which is left as an exercise.

Alternatively: Given $\epsilon > 0$, take $\delta > 0$ such that[3] for $0 < |h| < \delta$,

$$|D_i f(a+h) - D_i f(a)| < \epsilon.$$

Then if $0 < |h| < \delta$, $|\xi_i - a| < \delta$ and so LHS $\leq$ RHS $< n\epsilon$, which can be made arbitrarily small. This shows that the LHS $\to 0$ as $h \to 0$.

3.5 Double Differentiation

Suppose $f\colon E \mapsto \mathbb{R}^m$ (E open in $\mathbb{R}^n$) is differentiable. We can thus consider the function

$$f'\colon E \mapsto L(\mathbb{R}^n, \mathbb{R}^m) \text{ given by } x \mapsto f'(x).$$

Vulgarly, we can identify $L(\mathbb{R}^n, \mathbb{R}^m)$ with $\mathbb{R}^{mn}$ via matrices, and so can ask whether f' is differentiable. If it is differentiable at $a \in E$, then its derivative $f''(a)$ is a linear map $\mathbb{R}^n \mapsto L(\mathbb{R}^n, \mathbb{R}^m)$. It is better regarded as a bilinear map $\mathbb{R}^n \times \mathbb{R}^n \mapsto \mathbb{R}^m$. Thus $(f''(a)(h))(k)$ is regarded as $f''(a)(h,k)$. Similarly, if the partial derivatives $D_i f_j$ exist in E, we can ask whether the functions

$$x \mapsto D_i f_j(x), \qquad E \mapsto \mathbb{R}$$

are differentiable or even whether their partial derivatives

$$D_k D_i f_j(a) \equiv \left.\frac{\partial^2 f_j}{\partial x_k \partial x_i}\right|_a$$

exist.

Theorem 3.16. *Suppose $f\colon E \mapsto \mathbb{R}^m$ with $E \subseteq \mathbb{R}^n$ open, is such that all the partial derivatives $D_k D_i f_j(x)$ (exist and) are continuous in E. Then f is twice differentiable in E and the double derivative $f''(a)$ is a symmetric bilinear map for all $a \in E$.*

Remarks.

- *Sufficient to deal with $m = 1$.*
- *It follows from previous results that $f''(a)$ exists for all $a \in E$.*
- *It remains to show $D_i D_j f(a) = D_j D_i f(a), in E$, where $f\colon E \mapsto \mathbb{R}$.*

  For this we can keep things constant except in the i^{th} and j^{th} components.

It suffices to prove the following:

Proposition 3.17. *Suppose $f\colon E \mapsto \mathbb{R}$, $E \subseteq \mathbb{R}^2$ is such that the partial derivatives $D_1 D_2 f(x)$ and $D_2 D_1 f(x)$ are continuous. Then $D_1 D_2 f(x) = D_2 D_1 f(x)$.*

[3] $B(a,\delta) \subseteq E$ is also necessary.

Proof. Take $(a_1, a_2) \in E$. For (h_1, h_2) small enough (for $a + h \in E$) define

$$\begin{aligned} T(h_1, h_2) &= f(a_1 + h_1, a_2 + h_2) - f(a_1, a_2 + h_2) \\ &\quad - f(a_1 + h_1, a_2) + f(a_1, a_2) \end{aligned}$$

Apply the MVT to $y \mapsto f(a_1 + h, y) - f(a_1, y)$ to get $\hat{y} \in (a_2, a_2 + h_2)$ such that

$$T(h_1, h_2) = (D_2 f(a_1 + h, \hat{y}) - D_2 f(a_1, \hat{y}))h_2$$

Now apply MVT to $x \mapsto D_2 f(x, \hat{y})$ to get $\hat{x} \in (a_1, a_1 + h_1)$ with

$$T(h_1, h_2) = (D_1 D_2 f(\hat{x}, \hat{y}))h_1 h_2$$

As $h_1, h_2 \to 0$ separately, $(\hat{x}, \hat{y}) \to (a_1, a_2)$, and so, by continuity of $D_1 D_2$:

$$\lim_{h_1 \to 0, h_2 \to 0} \frac{T(h_1, h_2)}{h_1 h_2} = D_1 D_2 f(a_1, a_2)$$

Similarly

$$\lim_{h_1 \to 0, h_2 \to 0} \frac{T(h_1, h_2)}{h_1 h_2} = D_2 D_1 f(a_1, a_2).$$

The result follows by uniqueness of limits. □

3.6 Mean Value Theorems in Many Variables

Suppose first that $f \colon [a, b] \mapsto \mathbb{R}^m$ is continuous and is differentiable on (a, b). Then the derivative $f'(t) \in L(\mathbb{R}, \mathbb{R}^m)$ for $a < t < b$. We identify $L(\mathbb{R}, \mathbb{R}^m)$ with $\mathbb{R}^m$ via

$$\alpha \in L(\mathbb{R}, \mathbb{R}^m) \mapsto \alpha(1) \in \mathbb{R}^m$$

Then write $\|f'(t)\| = \|f'(t)(1)\|$.

Theorem 3.18. *With f as above, suppose $\|f'(t)\| \leq K$ for all $t \in (a, b)$. Then*

$$\|f(b) - f(a)\| \leq K\,|b - a|\,.$$

Proof. Set $e = f(b) - f(a)$ and let $\phi(t) = \langle f(t), e\rangle$, the inner product with e. By the one dimensional MVT we have $\phi(b) - \phi(a) = \phi'(c)(b - a)$ for some $c \in (a, b)$.

We can calculate $\phi'(t)$ by the chain rule as $\phi'(t) = \langle f'(t), e\rangle$. ($f'(t)$ regarded as begin a vector in $\mathbb{R}^m$). Now

$$\begin{aligned} \phi(b) - \phi(a) &= \langle f(b), e\rangle - \langle f(a), e\rangle \\ &= \langle f(b) - f(a), f(b) - f(a)\rangle \\ &= \|f(b) - f(a)\|^2\,. \end{aligned}$$

Therefore

$$\begin{aligned} \|f(b) - f(a)\|^2 &= |\langle f'(c), e\rangle|\;|b - a| \\ &\leq \|f'(c)\|\;\|f(b) - f(a)\|\;|b - a| \end{aligned}$$

and so $\|f(b) - f(a)\| \leq K\,|b - a|$. □

Finally, take the case $f\colon E \mapsto \mathbb{R}^m$ differentiable on E with E open in $\mathbb{R}^n$. For any $d \in E$, $f'(d) \in L(\mathbb{R}^n, \mathbb{R}^m)$.

For $\alpha \in L(\mathbb{R}^n, \mathbb{R}^m)$ we can define $\|\alpha\|$ by

$$\|\alpha\| = \sup_{x \neq 0} \frac{\|\alpha(x)\|}{\|x\|}$$

So $\|\alpha\|$ is least such that

$$\|\alpha(x)\| \leq \|\alpha\| \, \|x\|$$

for all x.

Theorem 3.19. *Suppose f is as above and $a, b \in E$ are such that the interval $[a,b]$ (line segment), $[a,b] = \{c(t) = tb + (1-t)a : 0 \leq t \leq 1\}$.*

Then if $\|f'(d)\| < K$ for all $d \in (a,b)$,

$$\|f(b) - f(a)\| \leq K \|b - a\| .$$

Proof. Let $g(t) = f(c(t))$, so that $g\colon [0,1] \mapsto \mathbb{R}^m$. By theorem 3.18,

$$\|f(b) - f(a)\| = \|g(1) - g(0)\| \leq L \cdot 1 = L$$

for $L \geq \|g'(t)\|$, $0 < t < 1$. But by the chain rule

$$g'(t) = f'(t) \underbrace{(b-a)}_{=c'(t)},$$

so that $\|g'(t)\| \leq \|f'(t)\| . \|b - a\| \leq K \|b - a\|$. The result follows. □

Chapter 4

Integration

4.1 The Riemann Integral

Definition 4.1. *A* dissection $\mathfrak{D}$ *of an interval* $[a,b]$ $(a < b)$*, is a sequence*

$$\mathfrak{D} = [x_0, \dots, x_n] \quad \textit{where} \quad a = x_0 < x_1 < x_2 < \dots < x_n = b.$$

A dissection $\mathfrak{D}_1$ *is* finer *than (or a* refinement *of) a dissection* $\mathfrak{D}_2$ *if and only if all the points of* $\mathfrak{D}_2$ *appear in* $\mathfrak{D}_1$*. Write* $\mathfrak{D}_1 < \mathfrak{D}_2$*.* [1]

Definition 4.2. *For* $f\colon [a,b] \mapsto \mathbb{R}$ *bounded and* $\mathfrak{D}$ *a dissection of* $[a,b]$ *we define*

$$S_{\mathfrak{D}} = \sum_{i=1}^{n} (x_i - x_{i-1}) \sup_{x_{i-1} \le x \le x_i} \{f(x)\}$$
$$s_{\mathfrak{D}} = \sum_{i=1}^{n} (x_i - x_{i-1}) \inf_{x_{i-1} \le x \le x_i} \{f(x)\}.$$

These are reasonable upper and lower estimates of the area under f. For general f we take the area below the axis to be negative.

Combinatorial Facts

Lemma 4.3. *For any* $\mathfrak{D}$*,* $s_{\mathfrak{D}} \le S_{\mathfrak{D}}$*.*

Lemma 4.4. *If* $\mathfrak{D}_1 \le \mathfrak{D}_2$*, then* $S_{\mathfrak{D}_1} \le S_{\mathfrak{D}_2}$ *and* $s_{\mathfrak{D}_1} \ge s_{\mathfrak{D}_2}$*.*

Lemma 4.5. *For any dissections* $\mathfrak{D}_1$ *and* $\mathfrak{D}_2$*,* $s_{\mathfrak{D}_1} \le S_{\mathfrak{D}_2}$*.*

Proof. Take a common refinement $\mathfrak{D}_3$, say, and

$$s_{\mathfrak{D}_1} \le s_{\mathfrak{D}_3} \le S_{\mathfrak{D}_3} \le S_{\mathfrak{D}_2}$$

It follows that the $s_{\mathfrak{D}}$ are bounded by an $S_{\mathfrak{D}_0}$, and the $S_{\mathfrak{D}}$ are bounded by any $s_{\mathfrak{D}_0}$. □

[1] The *mesh* of $\mathfrak{D} = [x_0, \dots, x_n]$ is $\max_{1 \le i \le n}\{|x_i - x_{i-1}|\}$. If $\mathfrak{D}_1 \le \mathfrak{D}_2$ then $\mathrm{mesh}(\mathfrak{D}_1) \le \mathrm{mesh}(\mathfrak{D}_2)$.

Definition 4.6. *For $f \colon [a,b] \mapsto \mathbb{R}$ bounded, define the upper Riemann integral*

$$S(f) \equiv \overline{\int_a^b} f(x)\,\mathrm{d}x = \inf_{\mathfrak{D}}\{S_{\mathfrak{D}}(f)\}$$

and the lower Riemann integral

$$s(f) \underline{\int_a^b} f(x)\,\mathrm{d}x = \sup_{\mathfrak{D}}\{s_{\mathfrak{D}}(f)\}.$$

Note that $s(f) \leq S(f)$. f is said to be Riemann integrable, *with $\int_a^b f(x)\,\mathrm{d}x = \sigma$ iff $s(f) = S(f)$.*

Example 4.7.

- $f(x) = \begin{cases} 0 & x \text{ irrational}, \\ 1 & x \text{ rational}. \end{cases} \quad x \in [0,1]$

  Then $S(f) = 1, s(f) = 0$ and so f is not Riemann integrable.

- $f(x) = \begin{cases} 0 & x \text{ irrational}, \\ \frac{1}{q} & x \text{ rational} = \frac{p}{q} \text{ in lowest terms}. \end{cases} \quad x \in [0,1]$

  is Riemann integrable with

$$\int_0^1 f(x)\,\mathrm{d}x = 0$$

Conventions

We defined $\int_a^b f(x)\,\mathrm{d}x$ for $a < b$ only. For $a = b$, $\int_a^b f(x)\,\mathrm{d}x = 0$ and for $b < a$, $\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x$.

These give a general additivity of the integral with respect to intervals, ie:

If f is Riemann integrable on the largest of the intervals,

$$[a,b], [a,c], [c,b]$$

then it is integrable on the others, with

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x.$$

This makes sense in the obvious case $a \leq c \leq b$, but also in all others, eg $b \leq a \leq c$.

Proof. Left to the reader. ☐

4.2 Riemann's Condition: A GPC for integrability

Theorem 4.8. *Suppose $f \colon [a,b] \mapsto \mathbb{R}$ is bounded. Then f is Riemann-integrable iff for all $\epsilon > 0$ there exists a dissection $\mathfrak{D}$ with $S_{\mathfrak{D}} - s_{\mathfrak{D}} < \epsilon$.*

Proof.
($\Rightarrow$) Take $\epsilon > 0$, Pick $\mathfrak{D}_1$ such that

$$S_{\mathfrak{D}_1} - \int_a^b f(x)\,\mathrm{d}x < \frac{\epsilon}{2}$$

Pick $\mathfrak{D}_2$ such that

$$\int_a^b f(x)\,\mathrm{d}x - s_{\mathfrak{D}_2} < \frac{\epsilon}{2}$$

Then if $\mathfrak{D}$ is a common refinement,

$$S_{\mathfrak{D}} - s_{\mathfrak{D}} \leq \left(S_{\mathfrak{D}_1} - \int_a^b f(x)\,\mathrm{d}x\right) + \left(\int_a^b f(x)\,\mathrm{d}x - s_{\mathfrak{D}_2}\right) < \epsilon$$

($\Leftarrow$) Generally, $S_{\mathfrak{D}} \geq S \geq s \geq s_{\mathfrak{D}}$ Riemann's condition gives $S - s < \epsilon$ for all $\epsilon > 0$. Hence $S = s$ and f is integrable. □

Remarks.

- *If σ is such that $\forall \epsilon > 0\ \exists \mathfrak{D}$ with $S_{\mathfrak{D}} - s_{\mathfrak{D}} < \epsilon$ and $S_{\mathfrak{D}} \geq \sigma \geq s_{\mathfrak{D}}$ then σ is $\int_a^b f(x)\,\mathrm{d}x$.*
- *A sum of the form*

$$\sigma_{\mathfrak{D}}(f) = \sum_{i=1}^{n} (x_i - x_{i-1}) f(\xi_i)$$

where $\xi_i \in [x_{i-1}, x_i]$, is an arbitrary Riemann sum. *Then f is Riemann integrable with*

$$\int_a^b f(x)\,\mathrm{d}x = \sigma$$

if and only if $\forall \epsilon > 0\ \exists \delta > 0\ \forall \mathfrak{D}$ with $\mathrm{mesh}(\mathfrak{D}) < \delta$ and all arbitrary sums

$$|\sigma_{\mathfrak{D}}(f) - \sigma| < \epsilon$$

Applications

A function $f \colon [a, b] \mapsto \mathbb{R}$ is

increasing if and only if

$$x \leq y \Rightarrow f(x) \leq f(y), \qquad x, y \in [a, b]$$

decreasing if and only if

$$x \leq y \Rightarrow f(x) \geq f(y), \qquad x, y \in [a, b]$$

monotonic if and only if it is either increasing or decreasing.

Proposition 4.9. *Any monotonic function is Riemann integrable on* $[a, b]$.

Proof. By symmetry, enough to consider the case when f is increasing. Dissect $[a, b]$ into n equal intervals, ie

$$\begin{aligned}\mathfrak{D} &= \left[a, a + \frac{(b-a)}{n}, a + 2\frac{(b-a)}{n}, \dots, b\right] \\ &= [x_0, x_1, \dots, x_n].\end{aligned}$$

Note that if $c < d$ then $\sup_{x \in [c,d]}\{f(x)\} = f(d)$ and $\inf_{x \in [c,d]}\{f(x)\} = f(c)$. Thus

$$\begin{aligned}S_{\mathfrak{D}} - s_{\mathfrak{D}} &= \sum_{i=1}^{n}(x_i - x_{i-1})(f(x_i) - f(x_{i-1})) \\ &= \frac{b-a}{n}\sum_{i=1}^{n}(f(x_i) - f(x_{i-1})) \\ &= \frac{b-a}{n}(f(b) - f(a))\end{aligned}$$

Now, the RHS$\to 0$ as $n \to \infty$ and so given $\epsilon > 0$ we can find n with

$$\frac{b-a}{n}(f(b) - f(a)) < \epsilon$$

and so we have $\mathfrak{D}$ with $S_{\mathfrak{D}} - s_{\mathfrak{D}} < \epsilon$. Thus f is Riemann integrable by Riemann's condition. □

Theorem 4.10. *If* $f : [a, b] \mapsto \mathbb{R}$ *is continuous, then* f *is Riemann integrable.*

Note that f is bounded on a closed interval.

Proof. We will use theorem 2.23, which states that if f is continuous on $[a, b]$, f is uniformly continuous on $[a, b]$. Therefore, given $\eta > 0$ we can find $\delta > 0$ such that for all $x, y \in [a, b]$:

$$|x - y| < \delta \Rightarrow |f(x) - f(y)| < \eta$$

Take n such that $\frac{b-a}{n} < \delta$ and consider the dissection

$$\begin{aligned}\mathfrak{D} &= \left[a, a + \frac{(b-a)}{n}, a + 2\frac{(b-a)}{n}, \dots, b\right] \\ &= [x_0, x_1, \dots, x_n].\end{aligned}$$

Now if $x, y \in [x_{i-1}, x_i]$ then $|x - y| < \delta$ and so $|f(x) - f(y)| < \eta$. Therefore

$$\sup_{x \in [x_{i-1}, x_i]}\{f(x)\} - \inf_{x \in [x_{i-1}, x_i]}\{f(x)\} \leq \eta.$$

We see that

$$S_{\mathfrak{D}} - s_{\mathfrak{D}} \leq \sum_{i-1}^{n}(x_i - x_{i-1}) \cdot \eta = (b-a)\eta$$

Now assume $\epsilon > 0$ given. Take η such that $(b-a)\eta < \epsilon$. As above, we can find $\mathfrak{D}$ with $S_{\mathfrak{D}} - s_{\mathfrak{D}} \leq (b-a)\eta < \epsilon$, so that f is Riemann integrable by Riemann's condition. □

4.3 Closure Properties

Notation. *Define* $M(f;c,d) \equiv \sup_{x\in[c,d]}\{f(x)\}$ *and* $m(f;c,d) \equiv \inf_{x\in[c,d]}\{f(x)\}$.

Proposition 4.11. *If* $f, g\colon [a,b] \mapsto \mathbb{R}$ *are Riemann integrable, so are*

1. $f+g\colon [a,b] \mapsto \mathbb{R}$, *with* $\int_a^b (f+g)\,\mathrm{d}x = \int_a^b f\,\mathrm{d}x + \int_a^b g\,\mathrm{d}x$.
2. $\lambda f\colon [a,b] \mapsto \mathbb{R}$ *(*$\lambda \in \mathbb{R}$*) with* $\int_a^b \lambda f\,\mathrm{d}x = \lambda \int_a^b f\,\mathrm{d}x$.

Proof of 1. Given $\epsilon > 0$. Take a dissection $\mathfrak{D}_1$ with $S_{\mathfrak{D}_1}(f) - s_{\mathfrak{D}_1}(f) < \frac{\epsilon}{2}$ and a dissection $\mathfrak{D}_2$ with $S_{\mathfrak{D}_2}(g) - s_{\mathfrak{D}_2}(g) < \frac{\epsilon}{2}$. Let $\mathfrak{D}$ be a common refinement. Note that

$$\begin{aligned} M(f+g;c,d) &\le M(f;c,d) + M(g;c,d) \\ m(f+g;c,d) &\ge m(f;c,d) + m(g;c,d) \end{aligned}$$

Hence

$$s_{\mathfrak{D}}(f) + s_{\mathfrak{D}}(g) \le s_{\mathfrak{D}}(f+g) \le S_{\mathfrak{D}}(f+g) \le S_{\mathfrak{D}}(f) + S_{\mathfrak{D}}(g)$$

and so $S_{\mathfrak{D}}(f+g) - s_{\mathfrak{D}}(f+g) < \epsilon$. Thus $f+g$ is Riemann integrable (by Riemann's condition). Further, given $\epsilon > 0$ we have a dissection $\mathfrak{D}$ with

$$\begin{aligned} S_{\mathfrak{D}}(f) - s_{\mathfrak{D}}(f) &< \frac{\epsilon}{2} \\ S_{\mathfrak{D}}(g) - s_{\mathfrak{D}}(g) &< \frac{\epsilon}{2}. \end{aligned}$$

Then

$$\begin{aligned} s_{\mathfrak{D}}(f) + s_{\mathfrak{D}}(g) &\le s_{\mathfrak{D}}(f+g) \\ &\le \int_a^b (f+g)\,\mathrm{d}x \\ &\le S_{\mathfrak{D}}(f+g) \\ &\le S_{\mathfrak{D}}(f) + S_{\mathfrak{D}}(g) \end{aligned}$$

and so

$$\begin{aligned} \left(\int_a^b f\,\mathrm{d}x - \frac{\epsilon}{2}\right) + \left(\int_a^b g\,\mathrm{d}x - \frac{\epsilon}{2}\right) &< \int_a^b (f+g)\,\mathrm{d}x \\ &< \left(\int_a^b f\,\mathrm{d}x + \frac{\epsilon}{2}\right) + \left(\int_a^b g\,\mathrm{d}x + \frac{\epsilon}{2}\right) \end{aligned}$$

Since $\epsilon > 0$ arbitrarily small, we have:

$$\int_a^b (f+g)\,\mathrm{d}x = \int_a^b f\,\mathrm{d}x + \int_a^b g\,\mathrm{d}x$$

□

Proof of 2 is left as an exercise.

Proposition 4.12. *Suppose $f, g\colon [a,b] \mapsto \mathbb{R}$ are bounded and Riemann integrable. Then $|f|$, f^2 and fg are all Riemann integrable.*

Proof. Note that

$$M(|f|\,; c, d) - m(|f|\,; c, d) \le M(f; c, d) - m(f; c, d),$$

and so, given $\epsilon > 0$, we can find a dissection $\mathfrak{D}$ with $S_{\mathfrak{D}}(f) - s_{\mathfrak{D}}(f) < \epsilon$ and then

$$S_{\mathfrak{D}}(|f|) - s_{\mathfrak{D}}(|f|) \le S_{\mathfrak{D}}(f) - s_{\mathfrak{D}}(f) < \epsilon.$$

Therefore $|f|$ is Riemann-integrable.
As for f^2, note that

$$\begin{aligned} &M(f^2; c, d) - m(f^2; c, d) \\ &\qquad = [M(|f|\,; c, d) + m(|f|\,; c, d)] \times [M(|f|\,; c, d) - m(|f|\,; c, d)] \\ &\qquad\qquad \le 2K\,(M(|f|\,; c, d) - m(|f|\,; c, d)) \end{aligned}$$

where K is some bound for $|f|$.
Given $\epsilon > 0$, take a dissection $\mathfrak{D}$ with $S_{\mathfrak{D}}(|f|) - s_{\mathfrak{D}}(|f|) < \frac{\epsilon}{2K}$. Then

$$S_{\mathfrak{D}}(f^2) - s_{\mathfrak{D}}(f^2) \le 2K(S_{\mathfrak{D}}(|f|) - s_{\mathfrak{D}}(|f|)) < \epsilon.$$

Therefore f^2 is Riemann-integrable.
The integrability of fg follows at once, since

$$fg = \frac{1}{2}\left((f+g)^2 - f^2 - g^2\right).$$

□

Estimates on Integrals

1. Suppose $F\colon [a,b] \mapsto \mathbb{R}$ is Riemann-integrable, $a < b$. If we take $\mathfrak{D} = [a,b]$ then we see that

$$(b-a)m(f; a, b) \le \int_a^b f(x)\,\mathrm{d}x \le (b-a)M(f; a, b).$$

   It follows that if $|f| \le K$ then

$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \le K\,|b-a|\,.$$

   This is true even if $a \ge b$.

2. Suppose $f\colon [a,b] \mapsto \mathbb{R}$ is Riemann-integrable, $a < b$. Then $S_{\mathfrak{D}}\,|f| \ge S_{\mathfrak{D}}(f)$ and so

$$\int_a^b |f| \ge \int_a^b f\,\mathrm{d}x.$$

Also $S_{\mathfrak{D}} |f| \geq S_{\mathfrak{D}}(-f)$. and so

$$\int_a^b |f| \geq -\int_a^b f \, \mathrm{d}x$$

Thus[2]

$$\left| \int_a^b f \, \mathrm{d}x \right| \leq \int_a^b |f| \, \mathrm{d}x.$$

4.4 The Fundamental Theorem of Calculus

If $f \colon [a, b] \mapsto \mathbb{R}$ is Riemann-integrable, then for any $[c, d] \subseteq [a, b]$, f is Riemann integrable on $[c, d]$.[3] Hence for $c \in [a, b]$ we can define a function

$$F(x) = \int_c^x f(t) \, \mathrm{d}t$$

on $[a, b]$.

Observation 4.13.

$$F(x) = \int_c^x f(t) \, \mathrm{d}t$$

is continuous on $[a, b]$ *if* f *is bounded.*

Proof. Note that

$$|F(x + h) - F(x)| = \int_x^{x+h} f(t) \, \mathrm{d}t \leq |h| K$$

where K is an upper bound for $|f|$. Now $|h| K \to 0$ as $h \to 0$, so F is continuous. □

Theorem 4.14 (The Fundamental Theorem of Calculus). *Suppose* $f \colon [a, b] \mapsto \mathbb{R}$ *is Riemann integrable. Take* $c, d \in [a, b]$ *and define*

$$F(x) = \int_c^x f(t) \, \mathrm{d}t.$$

If f *is continuous at* d*, then* F *is differentiable at* d *with* $F'(d) = f(d)$.[4]

---

[2] For general a, b;

$$\left| \int_a^b f \, \mathrm{d}x \right| \leq \left| \int_a^b |f| \, \mathrm{d}x \right|$$

[3] For if $\mathfrak{D}$ is a dissection of $[a, b]$ such that $S_{\mathfrak{D}}(f) - s_{\mathfrak{D}}(f) < \epsilon$ then $\mathfrak{D}$ restricts to $\mathfrak{D}'$, a dissection of $[c, d]$ with $S_{\mathfrak{D}'}(f) - s_{\mathfrak{D}'}(f) < \epsilon$.

[4] In the case d is a or b ($a < b$), we have right and left derivatives. We ignore these cases (result just as easy) and concentrate on $d \in (a, b)$.

Proof. Suppose $\epsilon > 0$ is given. By the continuity of f at d we can take $\delta > 0$ such that $(d-\delta, d+\delta) \subset (a,b)$ and

$$|k| < \delta \Rightarrow |f(k+d) - f(d)| < \epsilon.$$

If $0 < |h| < \delta$ then

$$\begin{aligned}
\left|\frac{F(d+h)-F(d)}{h} - f(d)\right| &= \left|\frac{1}{h}\int_d^{d+h} (f(t)-f(d))\,\mathrm{d}t\right| \\
&\le \frac{1}{|h|}\epsilon\,|h| \\
&< 2\epsilon.
\end{aligned}$$

□

Corollary 4.15 (Integration is anti-differentiation). *If $f = g'$ is continuous on $[a,b]$ then*

$$\int_a^b f(t)\,\mathrm{d}t = g(b) - g(a).$$

Proof. Set $F(x) = \int_a^x f(t)\,\mathrm{d}t$. Then

$$\frac{\mathrm{d}}{\mathrm{d}x}(F(x) - g(x)) = F'(x) - g'(x) = f(x) - f(x) = 0$$

and so $F(x) - g(x) = k$ is constant. Therefore

$$\int_a^b f(t)\,\mathrm{d}t = F(b) - F(a) = g(b) - g(a).$$

□

Corollary 4.16 (Integration by parts). *Suppose f, g are differentiable on (a,b) and f', g' continuous on $[a,b]$. Then*

$$\int_a^b f(t)g'(t)\,\mathrm{d}t = [f(t)g(t)]_a^b - \int_a^b f'(t)g(t)\,\mathrm{d}t.$$

Proof. Note that

$$\frac{\mathrm{d}}{\mathrm{d}x}(f(x)g(x)) = f(x)g'(x) + f'(x)g(x),$$

and so

$$\begin{aligned}
[f(t)g(t)]_a^b &= f(b)g(b) - f(a)g(a) \\
&= \int_a^b (fg)'(t)\,\mathrm{d}t \\
&= \int_a^b f'(t)g(t)\,\mathrm{d}t + \int_a^b f(t)g'(t)\,\mathrm{d}t.
\end{aligned}$$

□

Corollary 4.17 (Integration by Substitution). *Take $g\colon [a,b] \mapsto [c,d]$ with g' is continuous in $[a,b]$ and $f\colon [c,d] \mapsto \mathbb{R}$ continuous. Then*

$$\int_{g(a)}^{g(b)} f(t)\,\mathrm{d}t = \int_a^b f(g(s))g'(s)\,\mathrm{d}s.$$

Proof. Set $F(x) = \int_c^x f(t)\,\mathrm{d}t$. Now

$$\begin{aligned}\int_{g(a)}^{g(b)} f(t)\,\mathrm{d}t &= F(g(b)) - F(g(a)) \\ &= \int_a^b (F \circ g)'(s)\,\mathrm{d}s \\ &= \int_a^b F'(g(s))g'(s)\,\mathrm{d}s \quad \text{by Chain Rule} \\ &= \int_a^b f(g(s))g'(s)\,\mathrm{d}s.\end{aligned}$$

□

4.5 Differentiating Through the Integral

Suppose $g\colon \mathbb{R} \times [a,b] \mapsto \mathbb{R}$ is continuous. Then we can define

$$G(x) = \int_a^b g(x,t)\,\mathrm{d}t.$$

Proposition 4.18. *G is continuous as a function of x.*

Proof. Fix $x \in \mathbb{R}$ and suppose $\epsilon > 0$ is given. Now g is continuous and so is uniformly continuous on the closed bounded set $E = [x-1, x+1] \times [a,b]$. Hence we can take $\delta \in (0,1)$ such that for $u, v \in E$,

$$\|u - v\| < \delta \Rightarrow |g(u_x, u_t) - g(v_x, v_t)| < \epsilon.$$

So if $|h| < \delta$ then $\|(x+h,t) - (x,t)\| = |h| < \delta$ and so

$$|g(x+h,t) - g(x,t)| < \epsilon.$$

Therefore $|G(x+h) - G(x)| \leq |b-a|\,\epsilon < 2\,|b-a|\,\epsilon$, and as $2\,|b-a|\,\epsilon$ can be made arbitrarily small $G(x+h) \to G(x)$ as $h \to 0$. □

Now suppose also that $D_1 g(x,t) = \frac{\partial g}{\partial x}$ exists and is continuous throughout $\mathbb{R} \times [a,b]$.

Theorem 4.19. *Then G is differentiable with*

$$G'(x) = \int_a^b D_1 g(x,t)\,\mathrm{d}t$$

Proof. Fix $x \in \mathbb{R}$ and suppose $\epsilon > 0$ is given.

Now $D_1 g$ is continuous and so uniformly continuous on the closed and bounded set $E = [x-1, x+1] \times [a, b]$. We can therefore take $\delta > \in (0, 1)$ such that for $u, v \in E$,

$$\|u - v\| < \delta \Rightarrow |D_1 g(a) - D_1 g(x, t)| < \epsilon.$$

Now

$$\left| \frac{G(x+h) - G(x)}{h} - \int_a^b D_1 g(x, t)\,\mathrm{d}t \right| \\ = \frac{1}{|h|} \left| \int_a^b g(x+h, t) - g(x, t) - h D_1 g(x, t)\,\mathrm{d}t \right|.$$

But

$$g(x+h, t) - g(x, t) - h D_1 g(x, t) = h(D_1 g(\xi, t) - D_1 g(x, t))$$

for some $\xi \in (x, x+h)$ by the MVT.

Now if $0 < |h| < \delta$ we have $\|(\xi, t) - (x, t)\| < \delta$ and so

$$|g(x+h, t) - g(x, t) - h D_1 g(x, t)| < |h|\,\epsilon.$$

Hence

$$\begin{aligned} \left| \frac{G(x+h) - G(x)}{h} - \int_a^b D_1 g(x, t)\,\mathrm{d}t \right| &\leq \frac{1}{|h|}\,|b-a|\;|h|\,\epsilon \\ &< 2\,|b-a|\,\epsilon. \end{aligned}$$

But $2\,|b-a|\,\epsilon$ can be made arbitrarily small, so that

$$G'(x) = \int_a^b D_1 g(x, t)\,\mathrm{d}t.$$

□

4.6 Miscellaneous Topics

Improper Integrals

1. Case $f : [a, b] \mapsto \mathbb{R}$ but is unbounded (and possibly undefined at a finite number of places). Set

$$f_{N,M}(x) = \begin{cases} N & f(x) > N \\ f(x) & -M \leq f(x) \leq N \\ -M & f(x) < -M. \end{cases}$$

If

$$\int_a^b f_{N,M}(x)\,\mathrm{d}x \to \text{limit}$$

as $N, M \to \infty$ (separately), then the limit is the improper integral

$$\int_a^b f(x)\,\mathrm{d}x.$$

2. Case $f\colon (-\infty, \infty) \mapsto \mathbb{R}$ say.
   Then if

$$\int_{-x}^{+y} f(t)\,\mathrm{d}t \to \text{limit}$$

as $x, y \to \infty$ then the limit is the improper integral

$$\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t.$$

Integration of Functions $f\colon [a,b] \mapsto \mathbb{R}^n$

It is enough to integrate the coordinate functions separately so that

$$\int_a^b f(t)\,\mathrm{d}t = \left(\int_{a_1}^{b_1} f_1(t)\,\mathrm{d}t, \dots, \int_{a_n}^{b_n} f_n(t)\,\mathrm{d}t\right),$$

but there is a more intrinsic way of defining this.

Chapter 5

Metric Spaces

5.1 Definition and Examples

Definition 5.1. *A* metric space (X, d) *consists of a set* X *(the set of* points *of the space) and a function* $d\colon X \times X \mapsto \mathbb{R}$ *(the* metric *or* distance*) such that*

- $d(a, b) \geq 0$ *and* $d(a, b) = 0$ *iff* $a = b$*,*
- $d(a, b) = d(b, a)$*,*
- $d(a, c) \leq d(a, b) + d(b, c)$ $\forall a, b, c \in X$.

Examples

1. $\mathbb{R}^n$ with the Euclidean metric
$$d(x, y) = +\sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$
2. $\mathbb{R}^n$ with the $\sup$ metric
$$d(x, y) = \sup_{1 \leq i \leq n} \{|x_i - y_i|\}$$
3. $\mathbb{R}^n$ with the "grid" metric
$$d(x, y) = \sum_{i=1}^{n} |x_i - y_i|$$
4. $C[a, b]$ with the $\sup$ metric[1][2]
$$d(f, g) = \sup_{t \in [a,b]} \{|f(t) - g(t)|\}$$

[1]Define
$$C[a, b] = \{f\colon [a, b] \mapsto \mathbb{R} : f \text{ is continuous}\}$$

[2]This is the standard metric on $C[a, b]$. It's the one meant unless we say otherwise.

5. $C[a, b]$ with the L^1-metric
$$d(f, g) = \int_a^b |f(t) - g(t)| \; \mathrm{d}t$$
6. $C[a, b]$ with the L^2-metric
$$d(f, g) = \left(\int_a^b |f(t) - g(t)|^2 \; \mathrm{d}t \right)^{\frac{1}{2}}$$
analogous to the Euclidean metric.
7. Spherical Geometry: Consider $S^2 = \{x \in \mathbb{R}^3 : \|x\| = 1\}$. We can consider continuously differentiable paths $\gamma\colon [0, 1] \mapsto S^2$ and define the length of such a path as
$$L(\gamma) = \int_0^1 \|\gamma'(t)\| \; \mathrm{d}t.$$
The spherical distance is
$$S(x, y) = \inf_{\gamma \text{ a path from } x \text{ to } y \text{ in } S^2} \{L(\gamma)\}.$$
This distance is realized along great circles.
8. Hyperbolic geometry: Similarly for $\mathcal{D}$: the unit disc in $\mathbb{C}$. Take $\gamma\colon [0, 1] \mapsto \mathcal{D}$ and
$$L(\gamma) = \int_0^1 \frac{2\,|\gamma'(t)|}{1 + |\gamma(t)|^2} \, \mathrm{d}t.$$
Then
$$h(z, w) = \inf_{\gamma \text{ a path from } z \text{ to } w \text{ in } S^2} \{L(\gamma)\}$$
is realized on circles through z, w meeting $\partial\mathcal{D} = S'$ (boundary of $\mathcal{D}$) at right angles.
9. The discrete metric: Take any set X and define
$$d(x, y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$
10. The “British Rail Metric”: On $\mathbb{R}^2$ set
$$d(x, y) = \begin{cases} |x| + |y| & x \neq y \\ 0 & x = y \end{cases}$$

Definition 5.2. *Suppose (X, d) is a metric space and $Y \subseteq X$. Then d restricts to a map $d|_{Y \times Y} \mapsto \mathbb{R}$ which is a metric in Y. (Y, d) is a* (metric) subspace *of (X, d), d on Y is the* induced metric.

Example 5.3. *Any $E \subseteq \mathbb{R}^n$ is a metric subspace of $\mathbb{R}^n$ with the metric induced from the Euclidean metric.*[3]

[3] For instance, the Euclidean metric on S^2 is
$$(x, y) \mapsto 2 \sin\left(\frac{1}{2} S(x, y)\right).$$

5.2 Continuity and Uniform Continuity

Definition 5.4. *Let (X, d) and (Y, c) be metric spaces. A map $f\colon X \mapsto Y$ is continuous at $x \in X$ if and only if*

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x' \in X \quad d(x, x') < \delta \Rightarrow c(f(x), f(x')) < \epsilon.$$

Then $f\colon (X, d) \mapsto (Y, c)$ is continuous iff f is continuous at all $x \in X$.
Finally $f\colon (X, d) \mapsto (Y, c)$ is uniformly continuous iff

$$\forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x, x' \in X \quad d(x, x') < \delta \Rightarrow c(f(x), f(x')) < \epsilon.$$

A bijective continuous map $f\colon (X, d) \mapsto (Y, c)$ with continuous inverse is a *homeomorphism*.

A bijective uniformly continuous map $f\colon (X, d) \mapsto (Y, c)$ with uniformly continuous inverse is a *uniform homeomorphism*.

1. There are continuous bijections whose inverse is not continuous. For instance
   (a) Let d_1 be the discrete metric on $\mathbb{R}$ and d_2 the Euclidean metric. Then the identity map $id\colon (\mathbb{R}, d_1) \mapsto (\mathbb{R}, d_2)$ is a continuous bijection; its inverse is not.
   (b) (Geometric Example) Consider the map
   $$\begin{aligned}[0, 1) &\mapsto S^1 = \{z \in \mathbb{C} : |z| = 1\},\\ \theta &\mapsto e^{2\pi i\theta}\end{aligned}$$
   with the usual metrics. This map is continuous and bijective but its inverse is not continuous at $z = 1$.
2. Recall that a continuous map $f\colon E \mapsto \mathbb{R}^m$ where E is closed and bounded in $\mathbb{R}^n$ is uniformly continuous. Usually there are lots of continuous not uniformly continuous maps: For example
   $$\tan\colon \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \mapsto \mathbb{R}$$
   is continuous but not uniformly continuous, essentially because
   $$\tan'(x) \to \infty \quad \text{as} \quad x \to \frac{\pi}{2}.$$

Definition 5.5. *Let d_1, d_2 be two metrics on X. d_1 and d_2 are* equivalent *if and only if $id\colon (X, d_1) \mapsto (X, d_2)$ is a homeomorphism. In symbols, this becomes*

$$\begin{aligned}&\forall x \in X \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall y \in X \quad d_1(y, x) < \delta \Rightarrow d_2(y, x) < \epsilon \qquad \textit{and}\\ &\forall x \in X \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall y \in X \quad d_2(y, x) < \delta \Rightarrow d_1(y, x) < \epsilon.\end{aligned}$$

Notation. *Define $O(x, r) \equiv N(x, r) \equiv N_r(x) \equiv \{y\colon d(x, y) < r\}$.*

Then d_1 and d_2 are equivalent if and only if

1. $\forall x \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad N^1_\delta(x) \subseteq N^2_\epsilon(x)$.
2. $\forall x \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad N^2_\delta(x) \subseteq N^1_\epsilon(x)$.

Definition 5.6. *d_1 and d_2 are* uniformly equivalent *if and only if*

$$id\colon (X, d_1) \mapsto (X, d_2)$$

is a uniform homeomorphism. In symbols this is

$$\begin{array}{ll} \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in X \quad N^1_\delta(x) \subseteq N^2_\epsilon(x) & \textit{and} \\ \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall x \in X \quad N^2_\delta(x) \subseteq N^1_\epsilon(x) & \end{array}$$

The point of the definitions emerges from the following observation.

Observation 5.7.

1. *$id\colon (X, d) \mapsto (X, d)$ is (uniformly) continuous.*
2. *If $f\colon (X, d) \mapsto (Y, c)$ and $g\colon (Y, c) \mapsto (Z, e)$ are (uniformly) continuous then so is their composite.*

   Hence

   (a) *for topological considerations an equivalent metric works just as well;*

   (b) *for uniform considerations a uniformly equivalent metric works as well.*

Example 5.8. *On $\mathbb{R}^n$, the Euclidean, $\sup$, and grid metrics are uniformly equivalent.*

Proof. **Euclidean and** $\sup$

$$N^{\text{Euc}}_\epsilon(x) \subseteq N^{\sup}_\epsilon(x) \quad \text{and} \quad N^{\sup}_{\frac{\epsilon}{\sqrt{n}}} \subseteq N^{\text{Euc}}_\epsilon(x)$$

(A circle contained in a square; and a square contained in a circle).

Euclidean and Grid

$$N^{\text{grid}}_\epsilon(x) \subseteq N^{\text{Euc}}_\epsilon(x) \quad \text{and} \quad N^{\text{Euc}}_{\frac{\epsilon}{\sqrt{n}}} \subseteq N^{\text{grid}}_\epsilon(x).$$

□

Compare this with work in chapters 2 and 3.

5.3 Limits of sequences

Definition 5.9. *Let x_n be a sequence in a metric space (X, d). Then x_n converges to x as $n \to \infty$ if and only if $\forall \epsilon > 0\ \exists N\ \forall n \geq N d(x_n, x) < \epsilon$. Clearly $x_n \to x$ iff $d(x_n, x) \to 0$ as $n \to \infty$.*

Note that the limit of a sequence is unique. Proof is as in lemma 1.7.

Theorem 5.10. *Suppose (X, d_X) and (Y, d_Y) are metric spaces. A map*

$$f\colon (X, d_X) \mapsto (Y, d_Y)$$

is continuous if and only if whenever $x_n \to x$ in X then $f(x_n) \to f(x)$ in Y.

Proof.

$\Rightarrow$ Assume f continuous and take $x_n \to x$ in X. Suppose $\epsilon > 0$ given. By the continuity of f, we can take $\delta > 0$ such that

$$d(x, x') < \delta \Rightarrow d(f(x), f(x')) < \epsilon$$

As $x_n \to x$ we can take N such that, for all $n \geq N$, $d(x_n, x) < \delta$. Now if $n \geq N$, $d(f(x_n), f(x)) < \epsilon$. But since $\epsilon > 0$ was arbitrary $f(x_n) \to f(x)$.

$\Leftarrow$ Suppose f is not continuous at $x \in X$. Then there exists $\epsilon > 0$ such that for any $\delta > 0$ there is $x \in N_\delta(x')$ with $d(f(x), f(x')) \geq \epsilon$.

Fix such an $\epsilon > 0$. For each $n \geq 1$ pick x_n with $d(x_n, x) < n^{-1}$ and $d(f(x_n), f(x)) \geq \epsilon$. Then $x_n \to x$ but $f(x_n) \not\to f(x)$.

□

Definition 5.11. *A sequence x_n in a metric space (X, d) is* Cauchy *if and only if*

$$\forall \epsilon > 0 \quad \exists N \quad \forall n, m \geq N \quad d(x_n, x_m) < \epsilon.$$

Observation 5.12. *If $f\colon (X, d_X) \mapsto (Y, d_Y)$ is uniformly continuous, then x_n Cauchy in $X \Rightarrow f(x_n)$ Cauchy in Y.*

Proof. Take x_n Cauchy in X and suppose $\epsilon > 0$ is given. By uniform continuity we can pick $\delta > 0$ such that

$$\forall x, x' \in X \quad d_X(x, x') < \delta \Rightarrow d_Y(f(x), f(x')) < \epsilon.$$

Now pick N such that $\forall n, m \geq N d_X(x_n, x_m) < \epsilon$. Then $d_Y(f(x_n), f(x_m)) < \delta$ for all $m, n \geq N$. Since $\epsilon > 0$ arbitrary, $f(x_n)$ is Cauchy in Y. □

Definition 5.13. *A metric space (X, d) is* complete *if and only if every Cauchy sequence in X converges in X.*

A metric space (X, d) is compact *if and only if every sequence in X has a convergent subsequence.*

Remarks.

1. *$[0, 1]$ or any closed bounded set $E \subseteq \mathbb{R}^n$ is both complete and compact.*

   $(0, 1]$ is neither complete nor compact.

   Indeed if $E \subseteq \mathbb{R}^n$ is compact it must be closed and bounded and if E is complete and bounded, it is compact.

2. *Compactness $\Rightarrow$ completeness:*

   Proof. Take a Cauchy sequence x_n in a compact metric space. Then there is a convergent subsequence $x_{n(k)} \to x$ as $k \to \infty$. Therefore $x_n \to x$ as $n \to \infty$. □

However $C[a,b]$ with the $\sup$ metric is complete but not compact.

What is more, given $f \in C[a,b]$ and $r > 0$, the set $\{g : d(g,f) \leq r\}$ is closed and bounded — but not compact.

3. *Compactness is a "topological property". If (X, d_X) and (Y, d_Y) are homeomorphic, then X compact implies Y compact.*

   However, this isn't true for completeness: $(0,1]$ is homeomorphic to $[1,\infty)$ via $x \mapsto 1/x$ but $(0,1]$ is not complete while $[1,\infty)$ is.

   However if (X, d_Y) and (Y, d_Y) are uniformly homeomorphic, then X complete implies Y complete.

5.4 Open and Closed Sets in Metric Spaces

Definition 5.14. *Let (X,d) be a metric space. A subset $U \subseteq X$ is open iff whenever $x \in U$ there is $\epsilon > 0$ with $d(x',x) < \epsilon \Rightarrow x' \in U$ or*[4] *$N_\epsilon \subseteq U$.*

Observation 5.15. *$N_\epsilon(x)$ is itself open in (X,d).*

Proof. If $x' \in N_\epsilon(x)$ then $d(x',x) < \epsilon$ so that $\delta = \epsilon - d(x,x') > 0$. Therefore $N_\delta(x') \subseteq N_\epsilon(x)$. □

Theorem 5.16. *Let (X, d_X) and (Y, d_Y) be metric spaces. Then*

$$f\colon (X, d_X) \mapsto (Y, d_Y)$$

is continuous if and only if $f^{-1}(V)$[5] *is open in X whenever V is open in Y.*

Proof.

$\Rightarrow$ Assume f is continuous. Take V open in Y and $x \in f^{-1}(V)$. As V is open we can take $\epsilon > 0$ such that $N_\epsilon(f(x)) \subseteq V$. By continuity of f at x we can take $\delta > 0$ such that $d(x,x') < \delta \Rightarrow d(f(x'), f(x)) < \epsilon$, or alternatively $x' \in N_\delta(x) \Rightarrow f(x') \in N_\epsilon(f(x))$ so that $x' \in N_\delta(x) \Rightarrow f(x') \in V$. Therefore $x' \in f^{-1}(V)$ and so $N_\delta(x) \subseteq f^{-1}(V)$ and $f^{-1}(V)$ is open.

$\Leftarrow$ Conversely, assume $f^{-1}(V)$ is open in X whenever V is open in Y. Take $x \in X$ and suppose $\epsilon > 0$ is given. Then $N_\epsilon(f(x))$ is open in Y and so by assumption $f^{-1}(N_\epsilon(f(x)))$ is open in X. But $x \in f^{-1}(N_\epsilon(f(x)))$ and so we can take $\delta > 0$ such that $N_\delta(x) \subseteq f^{-1}(N_\epsilon(f(x)))$. Therefore $d(x',x) < \delta \Rightarrow d(f(x'), f(x)) < \epsilon$ and as $\epsilon > 0$ is arbitrary, f is continuous at x. As x is arbitrary, f is continuous.

□

Corollary 5.17. *Two metrics d_1, d_2 on X are equivalent if and only if they induce the same notion of open set. This is because d_1 and d_2 are equivalent iff*

- *For all V d_2-open, $id^{-1}(V) = V$ is d_1-open.*
- *For all U d_1-open, $id^{-1}(U) = U$ is d_2-open.*

[4] Recall that in a metric space (X,d): $N_\epsilon(x) = \{x' : d(x,x') < \epsilon\}$.

[5] Where $f^{-1}(V) = \{x \in X : f(x) \in V\}$.

Definition 5.18. *Suppose (X, d) is a metric space and $A \subseteq X$. A is closed if and only if $x_n \to x$ and $x_n \in A$ for all n implies $x \in A$.*

Proposition 5.19. *Let (X, d) be a metric space.*

1. *U is open in X if and only if $X \setminus U$ is closed in X.*
2. *A is closed in X if and only if $X \setminus A$ is open in X.*

Proof. We only need to show 1.

$\Rightarrow$ Suppose U is open in X. Take $x_n \to x$ with $x \in U$. As U is open we can take $\epsilon > 0$ with $N_\epsilon(x) \subseteq U$. As $x_n \to x$, we can take N such that

$$\forall n \geq N \quad x_n \in N_\epsilon(x).$$

So $x_n \in X$ for all $n \geq N$. Then if $x_n \to x$ and $x_n \in X \setminus U$ then $x \notin U$, which is the same as $x \in X \setminus U$. Therefore $X \setminus U$ is closed.

$\Leftarrow$ Suppose $X \setminus U$ is closed in X. Take $x \in U$. Suppose that for no $\epsilon > 0$ do we have $N_\epsilon(x) \subseteq U$. Then for $n \geq 1$ we can pick $x_n \in N_{\frac{1}{n}}(x) \setminus U$. Then $x_n \to x$ and so as $X \setminus U$ is closed, $x \in X \setminus U$. But $x \in U$, giving a contradiction. Thus the supposition is false, and there exists $\epsilon > 0$ with $N_\epsilon(x) \subseteq U$. As $x \in U$ is arbitrary, this shows U is open.

$\square$

Corollary 5.20. *A map $f\colon (X, d_X) \mapsto (Y, d_Y)$ is continuous iff $f^{-1}(B)$ is closed in X for all B closed in Y.*[6]

5.5 Compactness

If (X, d) is a metric space and $a \in X$ is fixed then the function $x \mapsto d(x, a)$ is (uniformly) continuous. This is because $|d(x, a) - d(y, a)| \leq d(x, y)$, so that if $d(x, y) < \epsilon$ then $|d(x, a) - d(y, a)| < \epsilon$.

Recall. *A metric space (X, d) is compact if and only if every sequence in (X, d) has a convergent subsequence.*

If $A \subseteq X$ with (X, d) a metric space we say that A is compact iff the induced subspace (A, d_A) is compact.[7]

Observation 5.21. *A subset/subspace $E \subseteq \mathbb{R}^n$ is compact if and only if it is closed and bounded.*

Proof.

$\Rightarrow$ This is essentially Bolzano-Weierstrass. Let x_n be a sequence in E. As E is bounded, x_n is bounded, so by Bolzano-Weierstrass x_n has a convergent subsequence. But as E is closed the limit of this subsequence is in E.

[6] Because $f^{-1}(Y \setminus B) = X \setminus f^{-1}(B)$.
[7] $x_n \in A$ implies x_n has a convergent subsequence.

$\Leftarrow$ Suppose E is compact. If E is not bounded then we can pick a sequence $x_n \in E$ with $\|x_n\| > n$ for all $n \geq 1$. Then x_n has no convergent subsequence. For if $x_{n(k)} \to x$ as $k \to \infty$, then

$$\left\|x_{n(k)}\right\| \to \|x\| \text{ as } k \to \infty,$$

but clearly

$$\left\|x_{n(k)}\right\| \to \infty \text{ as } k \to \infty.$$

This shows that E is bounded.

If E is not closed, then there is $x_n \in E$ with $x_n \to x \notin E$. But any subsequence $x_{n(k)} \to x \notin E$ and so $x_{n(k)} \not\to y \in E$ as limits of sequences are unique—a contradiction.

This shows that E is closed.

□

Thus, quite generally, if E is compact in a metric space (X, d), then E is closed and E is bounded in the sense that there exists $a \in E, r \in \mathbb{R}$ such that

$$E \subseteq \{x : d(x, a) < r\}$$

This is not enough for compactness. For instance, take

$$l^\infty = \{(x_n) : x_n \text{ is a bounded sequence in } \mathbb{R}\}$$

with $d((x_n), (y_n)) = \sup_n |x_n - y_n|$. Then consider the points

$$e^{(n)} = (0, \ldots, 0, \overbrace{1}^{n^{\text{th}} \text{ position}}, 0, \ldots), \text{ or } \left(e^{(n)}\right)_r = \delta_{nr}$$

Then $d(e^{(n)}, e^{(m)}) = 1$ for all $n \neq m$. So $E = \{e^{(n)}\}$ is closed and bounded: $E \subseteq \{(x_n) : d(x_n, 0) \leq 1\}$ But $\left(e^{(n)}\right)$ has no convergent subsequence.

Theorem 5.22. *Suppose $f\colon (X, d_X) \mapsto (Y, d_Y)$ is continuous and surjective. Then (X, d_X) compact implies (Y, d_Y) compact.*

Proof. Take y_n a sequence in Y. Since f is surjective, for each n pick x_n with $f(x_n) = y_n$. Then x_n is a sequence in X and so has a convergent subsequence $x_{n(k)} \to x$ as $k \to \infty$. As f is continuous, $f(x_{n(k)}) \to f(x)$ as $k \to \infty$, or $y_{n(k)} \to y = f(x)$ as $k \to \infty$.

Therefore y_n has a convergent subsequence and so Y is compact. □

Application. *Suppose $f\colon E \mapsto \mathbb{R}^n, E \subseteq \mathbb{R}^n$ closed and bounded. Then the image $f(E) \in \mathbb{R}^n$ is closed and bounded. In particular when $f\colon E \mapsto \mathbb{R}$ we have $f(E) \subseteq \mathbb{R}$ closed and bounded. But if $F \subseteq \mathbb{R}$ is closed and bounded then* $\inf F, \sup F \subseteq F$. *Therefore f is bounded and attains its bounds.*

Theorem 5.23. *If $f\colon (X, d_X) \mapsto (Y, d_Y)$ is continuous with (X, d_X) compact then f is uniformly continuous.*

Proof. As in theorem 2.23. □

Lemma 5.24. *Let (X, d) be a compact metric space. If $A \subseteq X$ is closed then A is compact.*

Proof. Take a sequence x_n in A. As (X, d) is compact, x_n has a convergent subsequence $x_{n(k)} \to x$ as $k \to \infty$. As A is closed, $x \in A$ and so $x_{n(k)} \to x \in A$. This shows A is compact. $\square$

Note that if $A \subseteq X$ is a compact subspace of a metric space (X, d) then A is closed.

Theorem 5.25. *Suppose $f\colon (X, d_X) \mapsto (Y, d_Y)$ is a continuous bijection. Then if (X, d_X) is compact, then (so is (Y, d_Y) and) f is a homeomorphism.*

Proof. Write $g\colon (Y, d_Y) \mapsto (X, d_X)$ for the inverse of f. We want this to be continuous. Take A closed in X. By lemma 5.24, A is compact, and so as f is continuous, $f(A)$ is compact in Y. Therefore $f(A)$ is closed in Y.

But as f is a bijection, $f(A) = g^{-1}(A)$. Thus A closed in X implies $g^{-1}(A)$ closed in Y and so g is continuous. $\square$

5.6 Completeness

Recall that a metric space (X, d) is complete if and only if every Cauchy sequence in X converges. If $A \subseteq X$ then A is complete if and only if the induced metric space (A, d_A) is complete. That is: A is complete iff every Cauchy sequence in A converges to a point of A.

Observation 5.26. *$E \subseteq \mathbb{R}^n$ is complete if and only if E is closed.*

Proof.

$\Leftarrow$ This is essentially the GPC. If x_n is Cauchy in E, then $x_n \to x$ in $\mathbb{R}^n$ by the GPC. But E is closed so that $x \in E$ and so $x_n \to x \in E$.

$\Rightarrow$ If E is not closed then there is a sequence $x_n \in E$ with $x_n \to x \notin E$. But x_n is Cauchy and by the uniqueness of limits $x_n \not\to y \in E$ for any $y \in E$. So E is not complete.

$\square$

Examples.

1. *$[1, \infty)$ is complete but $(0, 1]$ is not complete.*
2. *Any set X with the discrete metric is complete.*
3. *$\{1, 2, .., n\}$ with*

$$d(n, m) = \left| \frac{1}{n} - \frac{1}{m} \right|$$

is not complete.

Consider the space $B(X, \mathbb{R})$ of bounded real-valued functions $f\colon X \mapsto \mathbb{R}$ on a set $X \neq \emptyset$; with

$$d(f, g) = \sup_{x \in X} |f(x) - g(x)|,$$

the $\sup$ *metric*.

Proposition 5.27. *The space* $B(X, \mathbb{R})$ *with the* $\sup$ *metric is complete.*

Proof. Take f_n a Cauchy sequence in $B(X, \mathbb{R})$. Fix $x \in X$. Given $\epsilon > 0$ we can take N such that

$$\forall n, m \geq N \quad d(f_n, f_m) < \epsilon$$

Then

$$\forall n, m \geq N \quad d(f_n(x), f_m(x)) < \epsilon.$$

This shows that $f_n(x)$ is a Cauchy sequence in $\mathbb{R}$ and so has a limit, say $f(x)$. As $x \in X$ arbitrary, this defines a function $x \mapsto f(x)$ from X to $\mathbb{R}$.

Claim: $f_n \to f$. Suppose $\epsilon > 0$ given. Take N such that

$$\forall n, m \geq N \quad d(f_m, f_n) < \epsilon.$$

Then for any $x \in X$

$$\forall n, m \geq N \quad |f_n(x) - f_m(x)| < \epsilon.$$

Letting $m \to \infty$ we deduce that $|f_n(x) - f(x)| \leq \epsilon$ for any $x \in X$.

Thus $d(f_n, f) \leq \epsilon < 2\epsilon$ for all $n \geq N$. But $2\epsilon > 0$ is arbitrary, so this shows $f_n \to f$. $\square$

Chapter 6

Uniform Convergence

6.1 Motivation and Definition

Consider the binomial expansion

$$(1+x)^{\alpha} = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

for $|x| < 1$. This is quite easy to show via some form of Taylor's Theorem. Thus

$$\lim_{N\to\infty} \sum_{n=0}^{N} \binom{\alpha}{n} x^n = (1+x)^{\alpha}$$

As it stands this is for each individual x such that $|x| < 1$. It is pointwise convergence.

For functions $f_n, f \colon X \mapsto \mathbb{R}$, we say that $f_n \to f$ *pointwise* iff

$$\forall x \in X \quad f_n(x) \to f(x).$$

This notion is "useless". It does not preserve any important properties of f_n.

Examples.

- *A pointwise limit of continuous functions need not be continuous.*

$$f_n(x) = \begin{cases} 0 & x \le 0 \\ 1 & x \ge \frac{1}{n} \\ nx & 0 < x < \frac{1}{n} \end{cases}$$

is a sequence of continuous functions which converge pointwise to

$$f(x) = \begin{cases} 0 & x \le 0 \\ 1 & x > 0 \end{cases}$$

which is discontinuous.

- *The integral of a pointwise limit need not be the limit of the integrals.*

$$f_n(x) = \begin{cases} 0 & x \leq 0 \text{ or } x \geq \frac{2}{n} \\ xn^2 & 0 \leq n \leq \frac{1}{n} \\ n - n^2(x - \frac{1}{n}) & \frac{1}{n} \leq x \leq \frac{2}{n} \end{cases}$$

has

$$\int_0^2 f_n(x)\,\mathrm{d}x = 1$$

for all $n \geq 1$, but f_n converges pointwise to $f(x) = 0$ which has

$$\int_0^2 f(x)\,\mathrm{d}x = 0.$$

We focus on real valued functions but everything goes through for complex valued or vector valued functions.

We will often tacitly assume that sets X (metric spaces (X, d)) are non-empty.

Definition 6.1. *Let f_n, f be real valued functions on a set X. Then $f_n \to f$ uniformly if and only if given $\epsilon > 0$ there is N such that for all $x \in X$*

$$|f_n(x) - f(x)| < \epsilon$$

all $n \geq N$. In symbols:

$$\forall \epsilon > 0 \quad \exists N \quad \forall x \in X \quad \forall n \geq N \quad |f_n(x) - f(x)| < \epsilon.$$

This is equivalent to

Definition 6.2. *Let $f_n, f \in B(X, \mathbb{R})$. Then $f_n \to f$ uniformly iff $f_n \to f$ in the sup metric.*

The connection is as follows:

- If $f_n, f \in B(X, \mathbb{R})$, then these definitions are equivalent. (There's a bit of $< \epsilon$ vs $\leq \epsilon$ at issue).
- Suppose $f_n \to f$ in the sense of the first definition. There will be N such that

  $$\forall x \in X \quad |f_n(x) - f(x)| < 1$$

  for all $n \geq N$. Then $(f_n - f)_{n \geq N} \to 0$ uniformly in the sense of the second definition.

Theorem 6.3 (The General Principle of Convergence). *Suppose $f_n \colon X \mapsto \mathbb{R}$ such that*

Either

$$\forall \epsilon > 0 \quad \exists N \quad \forall x \in X \quad \forall n, m \geq N \quad |f_n(x) - f_m(x)| < \epsilon$$

or *Suppose $f_n \in B(X, \mathbb{R})$ is a Cauchy sequence. Then there is $f \colon X \mapsto \mathbb{R}$ with $f_n \to f$ uniformly.*

Proof. $B(X, \mathbb{R})$ is complete. ∎

6.2 The space $C(X)$

Definition 6.4. *Let (X, d) be a metric space. $C(X) \equiv C(X, \mathbb{R})$ is the space of bounded continuous functions from X to $\mathbb{R}$ with the $\sup$ metric.*

This notation is usually used when X is compact, when all continuous functions are bounded.

Proposition 6.5. *Suppose (X, d) is a metric space, that f_n is a sequence of continuous real-valued functions and that $f_n \to f$ uniformly on X. Then f is continuous.*

Proof. Fix $x \in X$ and suppose $\epsilon > 0$ given. Take N such that for all $y \in X$

$$\forall n \geq N \quad |f_n(y) - f(y)| < \epsilon.$$

As f_N is continuous at x we can take $\delta > 0$ such that

$$d(y, x) < \delta \Rightarrow |f_N(y) - f_N(x)| < \epsilon.$$

Then if $d(y, x) < \delta$,

$$\begin{aligned} |f(y) - f(x)| &\leq |f_N(y) - f(y)| + |f_N(x) - f(x)| + |f_N(y) - f_n(x)| \\ &< 3\epsilon. \end{aligned}$$

But 3ϵ can be made arbitrarily small and so f is continuous at x. But $x \in X$ is arbitrary, so f is continuous. □

Theorem 6.6. *The space $C(X)$ (with the $\sup$ metric) is complete.*

Proof. We know that $B(X, \mathbb{R})$ is complete, and the proposition says that $C(X)$ is closed in $B(X, \mathbb{R})$. □

Sketch of Direct Proof. Take f_n Cauchy in $C(X)$.

- For each $x \in X$, $f_n(x)$ is Cauchy, and so converges to a limit $f(x)$.
- f_n converges to f uniformly.
- f is continuous by the above argument.

□

Theorem 6.7 (Weierstrass Approximation Theorem). *If $f \in C[a, b]$, then f is the uniform limit of a sequence of polynomials.*

Proof. Omitted. □

6.3 The Integral as a Continuous Function

Restrict attention to $C[a,b]$, the space of continuous functions on the closed interval $[a,b]$.

Proposition 6.8. *Suppose $f_n \to f$ in $C[a,b]$. Then*

$$\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x \quad \textit{in } \mathbb{R}.$$

Proof. Suppose $\epsilon > 0$. Take N such that $\forall n \geq N d(f_n, f) < \epsilon$. Then if $c < d$ in $[a,b]$

$$m(f_n; c, d) - \epsilon \leq m(f; c, d) \leq M(f; c, d) \leq M(f_n; c, d) + \epsilon$$

for all $n \geq N$. So for any dissection $\mathfrak{D}$,

$$s_{\mathfrak{D}}(f_n) - \epsilon(b-a) \leq s_{\mathfrak{D}}(f) \leq S_{\mathfrak{D}}(f) \leq S_{\mathfrak{D}}(f_n) + \epsilon(b-a)$$

for all $n \geq N$.

Taking $\sup$s and $\inf$s, it follows that

$$\int_a^b f_n(x)\,\mathrm{d}x - \epsilon(b-a) \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b f_n(x)\,\mathrm{d}x + \epsilon(b-a)$$

for all $n \geq N$.

Then as $\epsilon(b-a) > 0$ can be made arbitrarily small,

$$\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x.$$

$\square$

We can make the superficial generalization: If $f \in C[a,b]$ then so is

$$x \mapsto \int_a^x f(t)\,\mathrm{d}t.$$

So

$$\int_a^x : C[a,b] \mapsto C[a,b].$$

Theorem 6.9. *The map*

$$\int_a^x : C[a,b] \mapsto C[a,b]$$

is continuous with respect to the $\sup$ *metric.*

That is, if $f_n \to f$ (uniformly), then

$$\int_a^x f_n(t)\,\mathrm{d}t \to \int_a^x f(t)\,\mathrm{d}t$$

(uniformly in x).

Proof. We see from the previous proof that if N is such that for all $y \in [a, b]$,

$$\forall n \geq N \quad |f_n(y) - f(y)| < \epsilon$$

then

$$\left| \int_a^x f_n(t)\,\mathrm{d}t - \int_a^x f(t)\,\mathrm{d}t \right| \leq \epsilon(x-a) \leq \epsilon(b-a) < 2\epsilon(b-a).$$

As $2\epsilon(b-a)$ is arbitrarily small (and independent of x), this shows

$$\int_a^x f_n(t)\,\mathrm{d}t \to \int_a^x f(t)\,\mathrm{d}t$$

uniformly in x. □

Uniform convergence controls integration, but *not* differentiation, for example the functions

$$f_n(x) = \frac{1}{n}\sin nx$$

converge uniformly to zero as $n \to \infty$, but the derivatives $\cos nx$ converge only at exceptional values.

Warning. *There are sequences of infinitely differentiable functions (polynomials even) which converge uniformly to functions which are necessarily continuous but nowhere differentiable. However, if we have uniform convergence of derivatives, all is well.*

Theorem 6.10. *Suppose $f_n \colon [a, b] \mapsto \mathbb{R}$ is a sequence of functions such that*

1. *the derivatives f_n' exist and are continuous on $[a, b]$*
2. *$f_n' \to g(x)$ uniformly on $[a, b]$*
3. *for some $c \in [a, b]$, $f_n(c)$ converges to a limit, d, say.*

Then $f_n(x)$ converges uniformly to a function $f(x)$, with $f'(x)$ (continuous and) equal to $g(x)$.[1]

Proof. By the FTC,

$$f_n(x) = f_n(c) + \int_c^x f_n'(t)\,\mathrm{d}t.$$

Using the lemma that if $f_n \to f$ uniformly and $g_n \to g$ uniformly then $f_n + g_n \to f + g$ uniformly[2], we see that

$$f_n(x) \to d + \int_c^x g(t)\,\mathrm{d}t$$

uniformly in X (by theorem 6.9). Thus

$$f_n(x) \to f(x) = d + \int_c^x g(t)\,\mathrm{d}t$$

, and $f(x)$ has continuous derivative $f'(x) = g(x)$ by FTC. □

[1] In these cases we do have

$$\frac{\mathrm{d}}{\mathrm{d}x}\left(\lim_{n\to\infty} f_n(x)\right) = \lim_{n\to\infty}\left(\frac{\mathrm{d}}{\mathrm{d}x} f_n(x)\right)$$

[2] This lemma is not actually part of the original lecture notes

6.4 Application to Power Series

For $M \geq N$, and $|z| \leq r$,

$$\begin{aligned}\left|\sum_{N+1}^{M} a_n z^n\right| &\leq \sum_{N+1}^{M} |a_n z^n| \\ &= \sum_{N+1}^{M} |a_n z_0^n| \left|\frac{z}{z_0}\right|^n \\ &\leq \sum_{N+1}^{M} k \left(\frac{r}{|z_0|}\right)^n \\ &< k \left(\frac{r}{|z_0|}\right)^{N+1} \frac{1}{1-\frac{r}{|z_0|}}\end{aligned}$$

which tends to zero as $N \to \infty$. This shows that the power series is absolutely convergent, uniformly in z for $|z| \leq r$. Whence, not only do power series $\sum a_n z^n$ have a radius of convergence $R \in [0, \infty]$ but also if $r < R$, then they converge uniformly in $\{z : |z| \leq r\}$.

Also, if $\sum a_n z_0^n$ converges, so that $|a_n z_0^n| < k$ say, we have the following for $r < |z_0|$. Choose s with $r < s < |z_0|$. Then for $|z| \leq r$ and $M \geq N$ we have

$$\begin{aligned}\left|\sum_{N+1}^{M} n a_n z^{n-1}\right| &\leq \sum_{N+1}^{M} \left|n a_n z^{n-1}\right| \\ &\leq \sum_{N+1}^{M} \left|a_n z_0^{n-1}\right| n \left(\frac{|z|}{s}\right)^{n-1} \left(\frac{s}{|z_0|}\right)^{n-1} \\ &\leq \sum_{N+1}^{M} k' n \left(\frac{r}{s}\right)^{n-1} \left(\frac{s}{|z_0|}\right)^{n-1} \qquad \text{where } \left|a_n z_0^{n-1}\right| \leq k'.\end{aligned}$$

For $n \geq N_0$, $n\left(\frac{r}{s}\right)^{n-1} \leq 1$ and so for $N \geq N_0$,

$$\begin{aligned}\left|\sum_{N+1}^{M} n a_n z^{n-1}\right| &\leq \sum_{N+1}^{M} k' \left(\frac{s}{|z_0|}\right)^{n-1} \\ &\leq k \left(\frac{s}{|z_0|}\right)^{N} \frac{1}{1-\frac{s}{|z_0|}} \to 0 \quad \text{as } N \to \infty.\end{aligned}$$

This shows that the series $\sum_{n \geq 1} n a_n z^{n-1}$ converges uniformly inside the radius of convergence. So what we've done, in the *real* case[3] is to deduce that

$$\sum_{n \geq 1} n a_n z^{n-1}$$

is the derivative of

$$\sum_{n \geq 1} a_n z^n$$

within the radius of convergence.

[3] And with more work, in the complex case.

6.5 Application to Fourier Series

Proposition 6.11 (Simplest Version). *Suppose a_n is a sequence such that*

$$\sum_{n\geq 1} n\,|a_n|$$

converges. Then

$$\sum_{n\geq 1} a_n \cos nt$$

converges uniformly and has a derivative

$$\sum_{n\geq 1} -na_n \sin nt$$

which is uniformly convergent to a continuous function.

Proof. Let $S_N(t)$ be the partial sum

$$S_N(t) = \sum_{n=1}^{N} a_n \cos nt. \quad \text{Then} \quad S'_N(t) = \sum_{n=1}^{N} -na_n \sin nt$$

is a sequence of continuous functions. Now for $M \geq N$

$$\begin{aligned}
|S_M(t) - S_N(t)| &= \left| \sum_{N+1}^{M} a_n \cos nt \right| \\
&\leq \sum_{N+1}^{M} |a_n \cos nt| \\
&\leq \sum_{N+1}^{M} |a_n| \\
&\leq \sum_{N+1}^{M} n\,|a_n| \to 0 \quad \text{as } N \to \infty. \\
\text{Also, } |S'_M(t) - S'_N(t)| &= \left| \sum_{N+1}^{M} -na_n \sin nt \right| \\
&\leq \sum_{N+1}^{M} |-na_n \sin nt| \\
&\leq \sum_{N+1}^{M} n\,|a_n| \to 0 \quad \text{as } N \to \infty.
\end{aligned}$$

So both $S_N(t)$ and $S'_N(t)$ are uniformly convergent and we deduce that

$$\frac{\mathrm{d}}{\mathrm{d}t} \sum_{n\geq 1} a_n \cos nt = \sum_{n\geq 1} -na_n \sin nt.$$

□

The right context for Fourier series is the L^2 norm arising from the inner product

$$\langle f, g \rangle = \frac{1}{\pi} \int_0^{2\pi} f(t) g(t) \, \mathrm{d}t$$

on functions on $[0, 2\pi]$. We take Fourier coefficients of a function $f(x)$

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos nt \, \mathrm{d}t \quad n \geq 0$$
$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(t) \sin nt \, \mathrm{d}t \quad n \geq 1$$

and hope that

$$f(x) = \frac{1}{2} a_0 + \sum_{n \geq 1} a_n \cos nx + b_n \sin nx.$$

This works for smooth functions; and much more generally in the L^2-sense; so that for example, for continuous functions we have Parseval's Identity:

$$\int_0^{2\pi} |f(x)|^2 \, \mathrm{d}x = \frac{a_0^2}{2} + \sum_{n \geq 1} \left(a_n^2 + b_n^2 \right).$$

Chapter 7

The Contraction Mapping Theorem

7.1 Statement and Proof

Definition 7.1. *A map $T\colon (X,d) \mapsto (X,d)$ on a metric space (X,d) is a contraction if and only if for some k, $0 \leq k < 1$*

$$\forall x, y \in X \quad d(Tx, Ty) \leq k d(x,y)$$

Theorem 7.2 (Contraction Mapping Theorem). *Suppose that $T\colon (X,d) \mapsto (X,d)$ is a contraction on a (non-empty) complete metric space (X,d). Then T has a unique fixed point.*

That is, there is a unique $a \in X$ with $Ta = a$.[1]

Proof. Pick a point $x_0 \in X$ and define inductively $x_{n+1} = Tx_n$ so that $x_n = T^n x_0$. For any $n, p \geq 0$ we have

$$\begin{aligned}
d(x_n, x_{n+p}) &= d(T^n x_0, T^n x_p) \\
&\leq k^n d(x_0, x_p \\
&\leq k^n [d(x_0, x_1) + d(x_1, x_2) + \ldots + d(x_{p-1}, x_p)] \\
&\leq k^n d(x_0, x_1)[1 + k + k^2 + \ldots + k^{p-1}] \\
&\leq \frac{k^n}{1-k} d(x_0, x_1).
\end{aligned}$$

Now

$$\frac{k^n}{1-k} d(x_0, x_1) \to 0 \quad \text{as } n \to \infty,$$

and so x_n is a Cauchy sequence. As (X,d) is complete, $x_n \to a \in X$. We now claim that a is a fixed point of T.

We can either use continuity of distance:

[1]As a preliminary remark, we see that as T is a contraction, it is certainly uniformly continuous

$$\begin{aligned}
d(Ta,a) &= d\left(Ta, \lim_{n\to\infty} x_n\right) \\
&= \lim_{n\to\infty} d(Ta, x_n) \\
&= \lim_{n\to\infty} d(Ta, Tx_{n-1}) \\
&\leq \lim_{n\to\infty} d(a, x_{n-1}) \\
&= d\left(a, \lim_{n\to\infty} x_{n-1}\right) \\
&= d(a,a) \\
&= 0,
\end{aligned}$$

and so $d(Ta,a) = 0$. Or we can use the (uniform) continuity of T.

$$\begin{aligned}
Ta &= T\left(\lim_{n\to\infty} x_n\right) \\
&= \lim_{n\to\infty} Tx_n \\
&= \lim_{n\to\infty} x_{n+1} \\
&= a.
\end{aligned}$$

As for uniqueness, suppose a, b are fixed points of T. Then

$$d(a,b) = d(Ta,Tb) \leq kd(a,b)$$

and since $0 \leq k < 1$, $d(a,b) = 0$ and so $a = b$. □

Corollary 7.3. *Suppose that* $T\colon (X,d) \mapsto (X,d)$ *is a map on a complete metric space* (X,d) *such that for some* $m \geq 1$, T^m *is a contraction, ie*

$$d(T^m x, T^m y) \leq kT(x,y).$$

Then T *has a unique fixed point.*

Proof. By the contraction mapping theorem, T^m has a unique fixed point a. Consider

$$\begin{aligned}
d(Ta,a) &= d(T^{m+1}a, T^m a) \\
&= d(T^m(Ta), T^m a) \\
&\leq kd(Ta,a).
\end{aligned}$$

So $d(Ta,a) = 0$ and thus a is a fixed point of T. If a, b are fixed points of T, they are fixed points of T^m and so $a = b$. □

Example 7.4. *Suppose we wish to solve* $x^2+2x-1 = 0$. *(The solutions are* $-1\pm\sqrt{2}$*.)*

We write this as

$$x = \frac{1}{2}(1-x^2)$$

and seek a fixed point of the map

$$T : x \mapsto \frac{1}{2}(1-x^2)$$

So we seek an interval $[a,b]$ *with* $T\colon [a,b] \mapsto [a,b]$ *and* T *a contraction on* $[a,b]$. *Now*

$$\begin{aligned} |Tx - Ty| &= \left|\frac{1}{2}x^2 - \frac{1}{2}y^2\right| \\ &= \frac{1}{2}\,|x+y|\,\,|x-y|\,. \end{aligned}$$

So if $|x|\,,|y| \le \frac{3}{4}$ *then*

$$|Tx - Ty| \le \frac{1}{2}(|x| + |y|)\,|x-y| \le \frac{3}{4}\,|x-y|$$

and so T *is a contraction on* $[-3/4, 3/4]$. *Actually*

$$T\colon \left[-\frac{3}{4}, \frac{3}{4}\right] \mapsto \left[0, \frac{1}{2}\right]$$

and so certainly

$$T\colon \left[-\frac{3}{4}, \frac{3}{4}\right] \mapsto \left[-\frac{3}{4}, \frac{3}{4}\right]$$

is a contraction.

So there is a unique fixed point of T *in* $[-3/4, 3/4]$. *The contraction mapping principle even gives a way of approximating it as closely as we want.*

7.2 Application to Differential Equations

Consider a differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = F(x,y) \tag{7.1}$$

subject to $y = y_0$ when $x = x_0$. We assume

$$F\colon [a,b] \times \mathbb{R} \mapsto \mathbb{R}$$

is continuous, $x_0 \in [a,b]$ and $y_0 \in \mathbb{R}$.

Observation 7.5. $g\colon [a,b] \mapsto \mathbb{R}$ *is a solution of (7.1) ie* g *is continuous,* $g'(x) = F(x, g(x))$ *for* $x \in (a,b)$ *and* $g(x_0) = y_0$*, iff* g *satisfies the Volterra integral equation*

$$g(x) = y_0 + \int_{x_0}^{x} F(t, g(t))\,\mathrm{d}t$$

on $[a,b]$.

Proof. Essentially the FTC.[2] ☐

[2] If g satisfies the differential equation, as $F(x, g(x))$ will be continuous we can integrate to get the integral equation and vice-versa.

Theorem 7.6. *Suppose $x_0 \in [a,b]$ closed interval, $y_0 \in \mathbb{R}$,*

$$F\colon [a,b] \times \mathbb{R} \mapsto \mathbb{R}$$

is continuous and satisfies a Lipschitz condition; ie there is K such that for all $x \in [a,b]$

$$|F(x,y_1) - F(x,y_2)| \leq K\,|y_1 - y_2|\,.$$

Then the differential equation (7.1) subject to the initial condition $y(x_0) = y_0$ has a unique solution in $C[a,b]$.

Proof. We consider the map $T\colon C[a,b] \mapsto C[a,b]$ defined by

$$Tf(x) = y_0 + \int_{x_0}^{x} F(t, f(t))\,\mathrm{d}t.$$

We claim that for all n,

$$|T^n f_1(x) - T^n f_2(x)| \leq \frac{K^n\,|x - x_0|}{n!} d(f_1, f_2)$$

The proof is by induction on n. The case $n = 0$ is trivial (and $n = 1$ is already done). The induction step is as follows:

$$\begin{aligned}
\left|T^{n+1} f_1(x) - T^{n+1} f_2(x)\right| &= \left|\int_{x_0}^{x} F(t, T^n f_1(t)) - F(t, T^n f_2(t))\,\mathrm{d}t\right| \\
&\leq \left|\int_{x_0}^{x} K\,|T^n f_1(t) - T^n f_2(t)|\,\mathrm{d}t\right| \\
&\leq \left|\int_{x_0}^{x} \frac{K.K^n\,|t - x_0|^n}{n!} d(f_1, f_2)\,\mathrm{d}t\right| \\
&= \frac{K^{n+1}\,|x - x_0|^{n+1}}{(n+1)!} d(f_1, f_2)
\end{aligned}$$

But

$$\frac{K^{n+1}\,|x - x_0|^{n+1}}{(n+1)!} d(f_1, f_2) \leq \frac{K^{n+1}\,|b - a|^{n+1}}{(n+1)!} d(f_1, f_2) \to 0$$

as $n \to \infty$. So for n sufficiently large,

$$\left(\frac{k^{n+1}\,|b - a|^{n+1}}{(n+1)!} < 1\right)$$

and so T^n is a contraction on $C[a,b]$.

Thus T has a unique fixed point in $C[a,b]$, which gives a unique solution to the differential equation. □

Example 7.7. *Solve $y' = y'$ with $y = 1$ at $x = 0$. Here $F(x,y) = y$ and the Lipschitz condition is trivial. So we have a unique solution on any closed interval $[a,b]$ with $0 \in [a,b]$. Thus we have a unique solution on $(-\infty, \infty)$.*

In fact[3] *we can do better than this and construct a solution by iterating T starting from $f_0 = 0$.*

$$\begin{aligned}
f_0(x) &= 0,\\
f_1(x) &= 1 + \int_0^x 0 \,\mathrm{d}t,\\
f_2(x) &= 1 + \int_0^x \mathrm{d}t = 1 + x,\\
f_3(x) &= 1 + x + \frac{x^2}{2!}\\
&\vdots
\end{aligned}$$

and so on. So (of course we knew this), the series for $\exp(x)$ *converges uniformly on bounded closed intervals.*

We can make a trivial generalization to higher dimensions.

Suppose $[a, b]$ is closed interval with $x_0 \in [a, b]$, $y_0 \in \mathbb{R}^n$ and $F\colon [a, b] \times \mathbb{R}^n \mapsto \mathbb{R}^n$ continuous and satisfying a Lipschitz condition: $\exists K$ such that

$$\|F(x, y_1) - F(x, y_2)\| \le K \|y_1 - y_2\| .$$

Then the differential equation

$$\frac{\mathrm{d}y}{\mathrm{d}x} = F(x, y)$$

with $y(x_0) = y_0$ has a unique solution in $C([a, b], \mathbb{R}^n)$. The proof is the same, but with $\|\cdot\|$s instead of $|\cdot|$s.

This kind of generalization is good for higher order differential equations. For example if we have

$$\frac{\mathrm{d}^2y}{\mathrm{d}x^2} = F\left(x, y, \frac{\mathrm{d}y}{\mathrm{d}x}\right)$$

with $y = y_0$, $dy/\mathrm{d}x = v_0$ at $x = x_0$ we can set $v = \frac{\mathrm{d}y}{\mathrm{d}x}$ and rewrite the equation as

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} v \\ F(x, y, v) \end{pmatrix}$$

with

$$\begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} y_0 \\ v_0 \end{pmatrix}$$

at $x = x_0$.

With a suitable Lipschitz condition we are home.

[3] This is not a general phenomenon!

7.3 Differential Equations: pathologies

The problem is that the Lipschitz condition seldom holds outright.

Trivial way Failure happens as $x \to$ something. The typical case is $x \to \infty$ but we can always consider bounded intervals and then expand them.

OK way Failure happens as $y \to \infty$.

Example 7.8.

$$\frac{\mathrm{d}y}{\mathrm{d}x} = 1 + y^2$$

with $y(0) = 0$. *Here* $F(x, y) = 1 + y^2$ *and so*

$$|F(x, y_1) - F(x, y_2)| = |y_1 + y_2|\ |y_1 - y_2|\,,$$

which is large for y *large.*

So F *as a map* $[a, b] \times \mathbb{R} \mapsto \mathbb{R}$ *does not satisfy a Lipschitz condition.*

Theorem 7.9. *Suppose* $x_0 \in (a, b), y_0 \in (c, d)$*, and*

$$F\colon [a, b] \times [c, d] \mapsto \mathbb{R}$$

is continuous and satisfies a Lipschitz condition: there is k *with*

$$|F(x, y_1) - F(x, y_2)| \leq k\,|y_1 - y_2|$$

in $[a, b] \times [c, d]$ *then there is* $\delta > 0$ *such that*

$$\frac{\mathrm{d}y}{\mathrm{d}x} = F(x, y)$$

with $y(0) = x_0$*, has a unique solution in* $[x_0 - \delta, x_0 + \delta]$.

Proof. Suppose that L is a bound for F on $[a, b] \times [c, d]$.[4]

Take $\eta > 0$ such that

$$[y_0 - \eta, y_0 + \eta] \subseteq [c, d]$$

Observe that if $|x - x_0| < \delta$ then

$$|Tf - y_0| = \left| \int_{x_0}^{x} F(t, f(t))\,\mathrm{d}t \right| \leq \delta L$$

so long as $f \in C$. So set $\delta = L^{-1}$.

[4]We aim to find a closed and so complete subspace

$$C \subseteq C[x_0 - \delta, x_0 + \delta]$$

of the form

$$C = \{f \colon C[x_0 - \delta, x_0 + \delta] : |f(x) - y_0| \leq \eta\}$$

for $\eta > 0$ with T mapping C to C.

Then C as above is complete, the map

$$T\colon f \mapsto y_0 + \int_{x_0}^{x} F(t, f(t))\,\mathrm{d}t$$

maps C to C, and by the argument of §7.2, T^n is a contraction for n sufficiently large.

Hence T has a unique fixed point and so the differential equation has a unique solution on $[x_0 - \delta, x_0 + \delta]$. □

Now we have a value $f(x_0+\delta)$ at $x_0+\delta$, so we can solve $\frac{\mathrm{d}y}{\mathrm{d}x} = F(x,y)$ with $y = f(x_0+\delta)$ at $x = x_0+\delta$, and so we extend the solution uniquely. This goes on until the solution goes off to $\pm\infty$. In this example we get $y = \tan x$.

Really bad case "Lipschitz fails at finite values of y." For example, consider $\frac{\mathrm{d}y}{\mathrm{d}x} = 2y^{\frac{1}{2}}$ with $y(0) = 0$.

Now $F(x,y) = 2y^{\frac{1}{2}}$ in $(-\infty, +\infty) \times [0, \infty)$ and

$$|F(x,y_1) - F(x,y_2)| = \frac{|y_1 - y_2|}{y_1^{1/2} + y_2^{1/2}},$$

which has problems as $y_1, y_2 \to 0$. We lose uniqueness of solutions.

7.4 Boundary Value Problems: Green's functions

Consider the second order linear ODE

$$\mathcal{L}y = \frac{\mathrm{d}^2 y}{\mathrm{d}x^2} + p(x)\frac{\mathrm{d}y}{\mathrm{d}x} + q(x)y = r(x)$$

subject to $y(a) = y(b) = 0$. (Here $p, q, r \in C[a,b]$).

The problem is that solutions are not always unique.

Example 7.10.

$$\frac{\mathrm{d}^2 y}{\mathrm{d}x^2} = -y$$

with $y(0) = y(\pi) = 0$ *has solutions* $A \sin x$ *for all* A.

Write $C^2[a,b]$ for the twice continuously differentiable functions on $[a,b]$, so that

$$\mathcal{L}\colon C^2[a,b] \mapsto C[a,b].$$

Write

$$C_0^2[a,b] = \{f \in C^2[a,b] : f(a) = f(b) = 0\}$$

and

$$\mathcal{L}_0\colon C_0^2[a,b] \mapsto C[a,b]$$

for the restricted map. Either $\ker \mathcal{L}_0 = \{0\}$ then a solution (if it exists) is unique, or $\ker \mathcal{L}_0 \neq \{0\}$, when we lose uniqueness. Note that because p, q, r have no y or $\frac{\mathrm{d}y}{\mathrm{d}x}$ dependence the Lipschitz condition for

$$\mathcal{L}y = \begin{cases} 0 \\ r \end{cases}$$

in the 2-dimensional form

$$\frac{\mathrm{d}}{\mathrm{d}x}\begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} v \\ -pv - qy + r \end{pmatrix}$$

is easy and so initial value problems always have unique solutions.

Assume $\ker \mathcal{L}_0 = \{0\}$. Now take $g_a(x)$, a solution to $\mathcal{L}y = 0$ with $y(a) = 1$, $y'(a) = 0$. $g_a(x) \not\equiv 0$ as $g_a'(a) = 1$. If $g_a(b) = 0$, $g_a \in C_0^2[a, b]$, contradicting $\ker \mathcal{L}_0 = \{0\}$ and so $g_a(b) \neq 0$.

We can similarly take $g_b(x)$, a solution to $\mathcal{L}y = 0$ with $y(b) = 0, y'(b) = 1$ and we have $g_b(a) \neq 0$. Now if h is a solution of $\mathcal{L}y = r$, then

$$f(x) = h(x) - \frac{h(a)}{g_b(a)} g_b(x) - \frac{h(b)}{g_a(b)} g_a(x)$$

is a solution to the boundary value problem. In fact this solution has an integral form:

$$f(x) = \int_a^b G(x, t) r(t) \, \mathrm{d}t.$$

We take the Wronskian

$$\mathcal{W}(x) = \begin{vmatrix} g_a(x) & g_b(x) \\ g_a'(x) & g_b'(x) \end{vmatrix}$$

and note that

$$\mathcal{W}'(x) + p(x)\mathcal{W}(x) = 0$$

and so

$$\mathcal{W}(x) = C \exp\left[-\int_a^x p(t)\, \mathrm{d}t\right]$$

$\mathcal{W}(a)$ and $\mathcal{W}(b) \neq 0$ so $C \neq 0$, so $\mathcal{W}(x) \neq 0$. Then we define

$$G(x, t) = \begin{cases} \frac{1}{\mathcal{W}(t)} g_b(x) g_a(t) & t \leq x \\ \frac{1}{\mathcal{W}(t)} g_b(t) g_a(x) & x \leq t \end{cases}$$

and check directly that

$$\int_a^b G(x, t) r(t) \, \mathrm{d}t$$

solves the initial value problem.

7.5 **The Inverse Function Theorem**

This is a theorem you should be aware of. Proof is omitted.

Theorem 7.11. *Suppose $f\colon E \mapsto \mathbb{R}^n$, $E \subseteq \mathbb{R}^n$ is open and continuously differentiable and that $f'(a)$ is invertible at some point $a \in E$. Then there are open U, V with $a \in U$, $b = f(a) \in V$ with $f\colon U \mapsto V$ bijective and the inverse of f, g say, continuously differentiable.*

References

- R. Haggarty, *Fundamentals of Modern Analysis*, Addison-Wesley, 1993.

  A new and well-presented book on the basics of real analysis. The exposition is excellent, but there's not enough content and you will rapidly outgrow this book. Worth a read but probably not worth buying.

- W. Rudin, *Principles of Mathematical Analysis*, Third ed., McGraw-Hill, 1976.

  This is a good book for this course. It is rather hard though.

- W.A. Sutherland, *Introduction to Metric and Topological Spaces*, OUP, 1975.

  This book is good on the metric space part of the course. It's also good for Further Analysis.