Topics in Analysis
In a Time of Covid

T. W. Körner

January 17, 2021

Small print In normal times these notes would be a supplement to the actual lectures with the statements of theorems in one portion of the notes and the proofs in a separate portion. (If you would prefer such a system then the old notes are available from my home page.) Instead I suspect the actual lectures will be a supplement to these notes. I strongly advise writing out your own notes at some time.

I suggest that you print out these notes and refer to them when watching the lectures. I shall do my best to keep to the same order and use the same notation as is used in the notes.

This course is not intended as a foundation for further courses, so the lecturer is allowed substantial flexibility. Exam questions will only cover topics in these notes but unless specifically stated otherwise, anything lectured in the course will be considered examinable. I will not talk about topological spaces but confine myself to ordinary Euclidean spaces $\mathbb{R}^n$, $\mathbb{C}$ and complete metric spaces.

I can provide some notes on the exercises for supervisors by e-mail. The exercises themselves form sections 18 to 21 of these notes.

These notes are written in $\LaTeX\,2_\varepsilon$ and should be available in pdf and tex format from my home page

http://www.dpmms.cam.ac.uk/~twk/

Under present circumstances corrections are *particularly welcome* (email me at twk@dpmms.cam.ac.uk).

Contents

1 Metric spaces

This section is devoted to fairly formal preliminaries. Things get more interesting in the next section and the course gets fully under way in the third. Both those students who find the early material worryingly familiar and those who find it worryingly unfamiliar are asked to suspend judgement until then.

Most Part II students will be familiar with the notion of a metric space.

Definition 1.1. *Suppose that X is a non-empty set and $d : X^2 \to \mathbb{R}$ is a function which obeys the following rules.*

(i) $d(x, y) \geq 0$ for all x, $y \in X$.

(ii) $d(x, y) = 0$ if and only if $x = y$.

(iii) $d(x, y) = d(y, x)$ for all x, $y \in X$.

(iv) $d(x, y) + d(y, z) \geq d(x, z)$ for all x, y, $z \in X$.

Then we say that d is a metric on X and that (X, d) is a metric space.

For most of the course we shall be concerned with metrics which you already know well.

Lemma 1.2. *(i) Consider $\mathbb{R}^n$. If we take d to be ordinary Euclidean distance*

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \left(\sum_{j=1}^{n} |x_j - y_j|^2 \right)^{1/2},$$

then $(\mathbb{R}^n, d)$ is a metric space. We refer to this space as Euclidean space.

(ii) Consider $\mathbb{C}$. If we take $d(z, w) = |z - w|$, then $(\mathbb{C}, d)$ is a metric space.

Proof. Proved in previous courses (and set as Exercise 18.1). □

The next definitions work in any metric space, but you can concentrate on what they mean for ordinary Euclidean space.

Definition 1.3. *If (X, d) is a metric space $x_n \in X$, $x \in X$ and $d(x_n, x) \to 0$, then we say that $x_n \underset{d}{\to} x$ as $n \to \infty$.*

Definition 1.4. *If (X, d) is a metric space $x \in X$ and $r > 0$, then we write*

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

and call $B(x, r)$ the open ball of radius r with centre x.

Definition 1.5. *Let* (X, d) *be a metric space.*

(i) We say that a subset E *of* X *is* closed *if, whenever* $x_n \in E$ *and* $x_n \underset{d}{\to} x$*, it follows that* $x \in E$.

(ii) We say that a subset U *of* X *is open if, whenever* $u \in U$*, we can find a* $\delta > 0$ *such that* $B(u, \delta) \subseteq U$.

Exercise 1.6. *(i) If* $x \in X$ *and* $r > 0$*, then* $B(x, r)$ *is open in the sense of Definition 1.5.*

(ii) If $x \in X$ *and* $r > 0$*, then the set*

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

is closed. (Naturally enough, we call $\bar{B}(x, r)$ *a closed ball.)*

(iii) If E *is closed, then* $X \setminus E$ *is open.*

(iv) If E *is open then* $X \setminus E$ *is closed.*

Solution. (i) If $y \in B(x, r)$ then $\delta = r - d(x, y) > 0$. Now observe that, if $z \in B(y, \delta)$, then

$$d(x, z) \leq d(x, y) + d(y, z) < d(y, x) + \delta < r.$$

(ii) If $y_n \in \bar{B}(x, r)$ and $y_n \underset{d}{\to} y$, then

$$d(x, y) \leq d(x, y_n) + d(y_n, y) \leq r + d(y_n, y) \to r,$$

so $d(x, y) \leq r$ and $y \in \bar{B}(x, r)$.

(iii) Suppose that $X \setminus E$ is not open. Then there is a point $y \notin E$ such that $B(y, r) \cap E \neq \varnothing$ whenever $r > 0$. Choose $y_n \in B(y, 1/n) \cap E$. We have $y_n \in E$, $y_n \underset{d}{\to} y$ and yet $y \notin E$. Thus E is not closed.

(iii) Suppose that $X \setminus E$ is not closed. Then there is a sequence $y_n \notin E$ with $y_n \underset{d}{\to} y$ and yet $y \in E$. Thus $B(y, r) \nsubseteq E$ for all $r > 0$ and E is not open. □

We recall (without proof) the following important results from 1A.

Theorem 1.7. [Cauchy criterion] *A sequence* $a_n \in \mathbb{R}$ *converges if and only if, given* $\epsilon > 0$*, we can find an* $N(\epsilon)$ *such that* $|a_n - a_m| < \epsilon$ *for all* n, $m \geq N(\epsilon)$

Generalisation leads us to the following definitions.

Definition 1.8. *If* (X, d) *is a metric space, then a sequence* (a_n) *with* $a_n \in X$ *is called a* Cauchy sequence *if, given* $\epsilon > 0$*, we can find an* $N(\epsilon)$ *such that* $d(a_n, a_m) < \epsilon$ *for all* n, $m \geq N(\epsilon)$.

Definition 1.9. *We say that a metric space* (X, d) *is* complete *if every Cauchy sequence converges.*

Exercise 1.10. *Show that if* (X, d) *is a metric space (complete or not), then every convergent sequence is Cauchy.*

Solution. Suppose $x_n \underset{d}{\to} x$. Let $\epsilon > 0$. We can find an N such that $d(x_n, x) < \epsilon/2$ for all $n \geq N$. It follows that

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

for all n, $m \geq N$. □

We note the following very useful remarks.

Lemma 1.11. *Let* (X, d) *be a metric space.*
(i) If a Cauchy sequence x_n *in* (X, d) *has a convergent subsequence with limit* x*, then* $x_n \to x$.
(ii) Let $\epsilon_n > 0$ *and* $\epsilon_n \to 0$ *as* $n \to \infty$*. If* (X, d) *has the property that, whenever* $d(x_n, x_{n+1}) < \epsilon_n$ *for* $n \geq 1$*, it follows that the sequence* x_n *converges, then* (X, d) *is complete.*

Proof. (i) Let $\epsilon > 0$. We can find an N such that $d(x_n, x_m) < \epsilon/2$ for m, $n \geq N$. We can now find a J such that $n(J) \geq N$ and $d(x_{n(J)}, x) < \epsilon/2$. We now observe that, if $m \geq N$, we get

$$d(x_m, x) \leq d(x_m, x_{n(J)}) + d(x_{n(J)}, x) < \epsilon/2 + \epsilon/2 = \epsilon.$$

(ii) If x_n is Cauchy we can find a strictly increasing sequence $n(j)$ with

$$d(x_n, x_m) < \epsilon(j)$$

for all n, $m \geq n(j)$. By hypothesis, $x_{n(j)}$ converges as $j \to \infty$. Part (i) now tells us that the sequence x_n converges. □

Lemma 1.11 (ii) is most useful when we have $\sum_{n=1}^{\infty} \epsilon_n$ convergent, for example if $\epsilon_n = 2^{-n}$.

The next exercise is simply a matter of disentangling notation.

Exercise 1.12. *Suppose that* (X, d) *is a metric space and* Y *is a non-empty subset of* X.
(i) Show that, if $d_Y(a, b) = d(a, b)$ *for all* a, $b \in Y$*, then* (Y, d_Y) *is a metric space.*
(ii) Show that, if (X, d) *is complete and* Y *is closed in* (X, d)*, then* (Y, d_Y) *is complete.*
(iii) Show that, if (Y, d_Y) *is complete, then (whether* (X, d) *is complete or not)* Y *is closed in* (X, d).

Solution. (i) Observe that, whenever $x,\ y,\ z \in Y$,

$$d_Y(x,y) = d(x,y) \geq 0,$$
$$d_Y(x,y) = 0 \Leftrightarrow d(x,y) = 0 \Leftrightarrow x = y,$$
$$d_Y(x,y) = d(x,y) = d(y,x) = d_Y(y,x),$$
$$d_Y(x,y) + d_Y(y,z) = d(x,y) + d(y,z) \geq d(x,z) = d_Y(x,z).$$

(ii) Suppose the sequence x_n is Cauchy in (Y, d_Y). Then the sequence x_n is Cauchy in (X, d), so $x_n \to x$ for some $x \in X$. But Y is closed, so $x \in Y$ and $x_n \to x$ in (Y, d_Y).

(iii) If $y_n \in Y$ and $y_n \to y$ in (X, d), then y_n is Cauchy in (X, d), so Cauchy in (Y, d_Y), so $y_n \to z$ in (Y, d_y) for some $z \in Y$. It follows that $y_n \to z$ in (X, d) so, by the uniqueness of limits, $y = z \in Y$. Thus Y is closed. □

We now come to our first real theorem.

Theorem 1.13. *The Euclidean space $\mathbb{R}^n$ with the usual metric is complete.*

We shall usually prove such theorems in the case $n = 2$ and remark that the general case is similar.

Proof Theorem 1.13 for $n = 2$. We prove the case when $n = 2$. Suppose that $\mathbf{x}_n = (x_n, y_n)$ is Cauchy in $\mathbb{R}^2$. Since

$$|x_n - x_m| \leq \|\mathbf{x}_n - \mathbf{x}_m\|,$$

x_n is Cauchy in $\mathbb{R}$ and by our 1A theorem (Theorem 1.7) converges to a limit x. Similarly y_n converges to a limit y in $\mathbb{R}$. If we set $\mathbf{x} = (x, y)$, then

$$\|\mathbf{x}_n - \mathbf{x}\| \leq |x_n - x| + |y_n - y| \to 0$$

as $n \to \infty$. □

2 Compact sets in Euclidean Space

In Part 1A we showed that any bounded sequence had a convergent subsequence. This result generalises to n dimensions.

Theorem 2.1. *(Bolzano–Weierstrass theorem). If $K > 0$ and $\mathbf{x}_r \in \mathbb{R}^m$ satisfies $\|\mathbf{x}_r\| \leq K$ for all r, then we can find an $\mathbf{x} \in \mathbb{R}^m$ and $r(k) \to \infty$ such that $\mathbf{x}_{r(k)} \to \mathbf{x}$ as $k \to \infty$.*

Proof. We prove the case $m = 2$. Write $\mathbf{x}_n = (x_n, y_n)$. We have that x_n is a bounded sequence in $\mathbb{R}$ and so (by the 1A result) there exists an $x \in \mathbb{R}$ and a sequence $n(j) \to \infty$ such that $x_{n(j)} \to x$ as $j \to \infty$. Now $y_{n(j)}$ is a bounded sequence in $\mathbb{R}$ and so there exists a $y \in \mathbb{R}$ and a sequence $j(k) \to \infty$ such that $y_{n(j(k))} \to y$ as $k \to \infty$. Now set $r(k) = n(j(k))$ and $\mathbf{x} = (x, y)$ to obtain $r(k) \to \infty$ and $\mathbf{x}_{r(k)} \to \mathbf{x}$ as $k \to \infty$. □

We now prove a very useful theorem.

Theorem 2.2. *(i) If E is closed bounded subset of $\mathbb{R}^m$, then any sequence $\mathbf{x}_r \in E$ has a subsequence with a limit in E.*

(ii) Conversely, if E is a subset of $\mathbb{R}^m$ with the property that any sequence $\mathbf{x}_r \in E$ has a subsequence with a limit in E, then E is closed and bounded.

Theorem 2.2. (i) Since $\mathbf{x}_r \in E$, we know that the $\mathbf{x}_r$ form a bounded sequence and so have a convergent subsequence $\mathbf{x}_{r(k)} \to \mathbf{x}$. Since E is closed, $\mathbf{x} \in E$.

(ii) If E is not bounded, we can find $\mathbf{x}_r \in E$ with $\|\mathbf{x}_{r+1}\| \geq \|\mathbf{x}_r\| + 1$. If $r > s$

$$\|\mathbf{x}_r - \mathbf{x}_s\| \geq \|\mathbf{x}_r\| - \|\mathbf{x}_s\| \geq 1,$$

so no subsequence can be Cauchy and so no subsequence can converge.

If E is not closed, we can find $\mathbf{x}_r \in E$ and $\mathbf{x} \notin E$ such that $\mathbf{x}_r \to \mathbf{x}$. Any subsequence of $\mathbf{x}_r$ will still converge to $\mathbf{x} \notin E$. □

We shall refer to the property described in (i) as the Bolzano–Weierstrass property.

If you cannot see how to prove a result in $\mathbb{R}^m$ using the Bolzano–Weierstrass property then the 1A proof for $\mathbb{R}$ will often provide a hint.

We often refer to closed bounded subsets of $\mathbb{R}^m$ as compact sets. (The reader is warned that, in general metric spaces, ‘closed and bounded’ does not mean the same thing as ‘compact’ (see, for example, Exercise 19.5). If we deal with the even more general case of topological spaces we have to distinguish between compactness and sequential compactness[1]. We shall only talk about compact sets in $\mathbb{R}^m$.)

The reader will be familiar with definitions of the following type.

Definition 2.3. *If (X, d) and (Y, ρ) are metric spaces and $E \subseteq X$, we say that a function $f : E \to Y$ is continuous if, given $\epsilon > 0$ and $x \in E$, we can find a $\delta(x, \epsilon) > 0$ such that, whenever $z \in E$ and $d(z, x) < \delta(x, \epsilon)$, we have $\rho(f(z), f(x)) < \epsilon$.*

[1]If you intend to climb Everest you need your own oxygen supply. If you intend to climb the Gog Magogs you do not.

Exercise 2.4. *Let (X, d) and (Y, ρ) be metric spaces and let $f : X \to Y$ be a function.*

(i) f is continuous if and only if $f^{-1}(U)$ is open whenever U is.

(ii) f is continuous if and only if $f^{-1}(E)$ is closed whenever E is.

Solution. (i) Suppose $f^{-1}(U)$ is open whenever U is. If $x \in X$, $\epsilon > 0$, we know that $B(f(x), \epsilon)$ is an open subset of Y, so $f^{-1}\big(B(f(x), \epsilon)\big)$ is an open subset of X containing x. Thus we can find a $\delta > 0$ with $B(x, \delta) \subseteq f^{-1}\big(B(f(x), \epsilon)\big)$. In other words,

$$z \in B(x, \delta) \Rightarrow f(z) \in B(f(x), \epsilon).$$

Thus f is continuous.

Conversely, if f is continuous and U open in Y, then, given $x \in X$ with $f(x) \in U$, we can find a $\delta > 0$ such that $B(f(x), \delta) \subseteq U$ and an $\epsilon > 0$ such that

$$z \in B(x, \delta) \Rightarrow f(z) \in B(f(x), \epsilon).$$

Thus $B(x, \epsilon) \subseteq f^{-1}(U)$. We have shown that $f^{-1}(U)$ is open.

(ii) Complementation. If $f^{-1}(F)$ is closed for all F closed then

$$U \text{ open} \Rightarrow Y \setminus U \text{ closed} \Rightarrow X \setminus f^{-1}(U) = f^{-1}(Y \setminus U) \text{ closed} \Rightarrow f^{-1}(U) \text{ open},$$

so f is continuous.

The converse is proved similarly. □

Even if the reader has not met the general metric space definition, she will probably have a good idea of the properties of continuous functions $f : E \to \mathbb{R}^n$ when E is a subset of $\mathbb{R}^m$.

The following idea will be used several times during the course.

Lemma 2.5. *If (X, d) is a metric space and A is a non-empty closed subset of X let us write*

$$d(x, A) = \inf\{d(x, a) \,:\, a \in A\}.$$

Then $d(x, A) = 0$ implies $x \in A$. Further the map $x \mapsto d(x, A)$ is continuous.

Proof. If $d(x, A) = 0$, then we can find $x_n \in A$ such that $d(x_n, x) \leq 1/n$, so $x_n \to x$. But A is closed, so $x \in A$.

Let x, $y \in X$. Given $\epsilon > 0$, we can find $a \in A$ such that $d(x, a) \leq d(x, A) + \epsilon$. Now

$$d(y, A) \leq d(y, a) \leq d(x, y) + d(x, a) \leq d(x, y) + d(x, A) + \epsilon.$$

Since ϵ was arbitrary,

$$d(y, A) \le d(x, y) + d(x, A).$$

The same argument shows that $d(x, A) \le d(x, y) + d(y, A)$ so

$$|d(x, A) - d(y, A)| \le d(x, y).$$

This shows that the map $x \mapsto d(x, A)$ is continuous. □

We now prove that the continuous image of a compact set is compact.

Theorem 2.6. *If E is a compact subset of $\mathbb{R}^m$ and $f : E \to \mathbb{R}^n$ is continuous, then $f(E)$ is a compact subset of $\mathbb{R}^n$.*

Proof. Suppose that $y_n \in f(E)$. Then $y_n = f(x_n)$ for some $x_n \in E$. By the Bolzano–Weierstrass property, we can find $n(j) \to \infty$ and $x \in E$ such that $x_{n(j)} \to x$ as $j \to \infty$. Now, by continuity,

$$y_{n(j)} = f(x_{n(j)}) \to f(x) \in f(E)$$

so we are done. □

At first sight Theorem 2.6 seems too abstract to be useful, but it has an immediate corollary.

Theorem 2.7. *If E is a non-empty compact subset of $\mathbb{R}^m$ and $f : E \to \mathbb{R}$ is continuous, then there exist $\mathbf{a}$, $\mathbf{b} \in E$ such that*

$$f(\mathbf{a}) \ge f(\mathbf{x}) \ge f(\mathbf{b})$$

for all $\mathbf{x} \in E$.

Proof. By Theorem 2.6, $f(E)$ is closed and bounded. Since $f(E)$ is non-empty, it has a supremum (see 1A), α, say. By the definition of the supremum, we can find $\mathbf{a}_n \in E$ such that

$$\alpha - 1/n \le f(\mathbf{a}_n) \le \alpha.$$

By the Bolzano–Weierstrass property, we can find $n(j) \to \infty$ and $\mathbf{a} \in E$ such that $\mathbf{a}_{n(j)} \to \mathbf{a}$ as $j \to \infty$. We have $f(\mathbf{a}_{n(j)}) \to f(\mathbf{a})$, so $f(\mathbf{a}) = \alpha$. Thus

$$f(\mathbf{a}) \ge f(\mathbf{x})$$

for all $\mathbf{x} \in E$. We find $\mathbf{b}$ in a similar manner. □

Thus a continuous real valued function on a compact set is bounded and attains its bounds.

Exercise 2.8. *Deduce the theorem in 1A which states that if $f : [c, d] \to \mathbb{R}$ is continuous, then f is bounded and attains its bounds.*

Theorem 2.7 gives a neat proof of the fundamental theorem of algebra (which, contrary to its name, is a theorem of analysis).

Theorem 2.9. [Fundamental Theorem of Algebra] *If we work in the complex numbers, every non-constant polynomial has a root.*

Proof. Let $n \geq 1$. Suppose $P(z) = \sum_{j=0}^{n} a_j z^j$ where, without loss of generality, we take $a_n = 1$.

If $R \geq 2(2 + \sum_{j=0}^{n-1} |a_j|)$, then, whenever $|z| \geq R$, we have

$$\begin{aligned} |P(z)| &\geq |z|^n - \sum_{j=0}^{n-1} |a_j||z|^j \\ &= |z|^n \left(1 - \sum_{j=0}^{n-1} |a_j||z|^{j-n}\right) \\ &\geq |z|^n/2 > |a_0|. \end{aligned}$$

Since $\bar{D}_R = \{z \in \mathbb{C} : |z| \leq R\}$ is closed and bounded (that is to say compact) and the map $z \mapsto |P(z)|$ is continuous, $|P|$ attains a minimum on $\bar{D}_R$ at a point z_0, say. By the previous paragraph, $|z_0| < R$ (since $|P(z_0)| \leq |P(0)|$) and so we can find a $\delta > 0$ such that $|P(z)| \geq |P(z_0)|$ for all $|z - z_0| < \delta$.

By replacing $P(z)$ by $P(z - z_0)$, we may assume that $z_0 = 0$ so that $|P(z)| \geq |P(0)|$ for all $|z| < \delta$. If $a_0 = 0$, then we have $P(0) = 0$ and we are done.

We show that the assumption that $a_0 \neq 0$ leads to a contradiction. Observe that

$$P(z) = \sum_{j=m}^{n} a_j z^j + a_0 = a_0 \left(1 - \sum_{j=m}^{n} b_j z^j\right)$$

with $a_m \neq 0$ and so $b_m \neq 0$. Choose θ so that $b_m \exp(im\theta)$ is real and positive. Then

$$|P(\eta \exp i\theta)| \leq |a_0| - |b_m|\eta^m + |a_0|\eta^{m+1} \sum_{j=m}^{n} |b_j| \leq |a_0| - |b_m|\eta^m/2 < |P(0)|$$

when η is strictly positive and sufficiently small. We have the required contradiction. □

The reader will probably have seen, but may well have forgotten, the contents of the next exercise.

Exercise 2.10. *We work in the complex numbers.*

(i) Use induction on n to show that, if P is a polynomial of degree n and $a \in \mathbb{C}$, then there exists a polynomial Q of degree $n-1$ and an $r \in \mathbb{C}$ such that

$$P(z) = (z-a)Q(z) + r.$$

(ii) Deduce that, if P is a polynomial of degree n with root a, then there exists a polynomial Q of degree $n-1$ such that

$$P(z) = (z-a)Q(z).$$

(iii) Use induction and the fundamental theorem of algebra to show that every polynomial P of degree n can be written in the form

$$A\prod_{j=1}^{n}(z-a_j)$$

with A, a_1, a_2, …, $a_n \in \mathbb{C}$ and $A \neq 0$.

(iv) Show that, if P is a polynomial of degree at most n which vanishes at $n+1$ points, then P is the zero polynomial.

Solution. (i) Let $S(m)$ be the statement that, if P is a polynomial of degree n with $n \leq m$ and $a \in \mathbb{C}$, then there exists a polynomial Q of degree $n-1$ and an $r \in \mathbb{C}$ such that

$$P(z) = (z-a)Q(z) + r.$$

Suppose that $S(m)$ is true and P is a polynomial of degree $m+1$. Then $P(z) = Az^{m+1} + Q(z)$ where $A \neq 0$ and Q is a polynomial of degree at most m. We have

$$P(z) = A(z-a)z^m + q(z)$$

where $q(z) = Q(z) + az^m$, so q is a polynomial of degree at most m and, by the inductive hypothesis,

$$q(z) = (z-a)u(z) + r$$

with u a polynomial of degree at most $m-1$. Thus $P(z) = (z-a)Q(z) + r$ with $Q(z) = Az^m + u(z)$. We have shown that $S(m+1)$ is true.

Now $S(1)$ is true, since $cz + d = c(z-a) + (d - ca)$, so the required result follows by induction.

(ii) We have $P(z) = (z-a)Q(z) + r$ by (i). Setting $z = a$, we have $0 = P(a) = r$ so $r = 0$ and the result follows.

(iii) If P_n has degree $n \geq 1$, then the fundamental theorem of algebra tells us that P_n has a root a_n. By (ii), there exists a polynomial P_{n-1} of degree $n-1$ such that

$$P(z) = (z - a_n)P_{n-1}(z).$$

Using induction, we deduce that

$$P_n(z) = P_0(z)\prod_{j=1}^{n}(z - a_j),$$

where $P_0(z)$ is a polynomial of degree 0, that is to say, $P_0(z) = A$ with A a constant.

(iv) If P is not the zero polynomial, then (iii) tells we can find $m \leq n$ such that

$$P(z) = A\prod_{j=1}^{m}(z - a_j)$$

with A, a_1, a_2, ..., $a_m \in \mathbb{C}$ and $A \neq 0$. Now $P(z) = 0$ if and only if $z = a_j$ for some $1 \leq j \leq m$. The result follows. □

3 Laplace's equation

We need a preliminary definition.

Definition 3.1. *Let (X, d) be a metric space and E a subset of X.*

(i) The interior of E is the set of all points x such that there exists a $\delta > 0$ (depending on x) such that $B(x, \delta) \subseteq E$. We write $\operatorname{Int} E$ *for the interior of E.*

(ii) The closure of E is the set of points x in X such that we can find $e_n \in E$ with $e_n \underset{d}{\rightarrow} x$. We write $\operatorname{Cl} E$ *for the closure of E.*

(iii) The boundary of E is the set $\partial E = \operatorname{Cl} E \setminus \operatorname{Int} E$.

Exercise 3.2. *(i) Show that* $\operatorname{Int} E$ *is open. Show also that, if V is open and $V \subseteq E$, then $V \subseteq \operatorname{Int} E$.*

(ii) Show that $\operatorname{Cl} E$ *is closed. Show also that, if F is closed and $F \supseteq E$, then $F \supseteq \operatorname{Cl} E$.*

(Thus $\operatorname{Int} E$ *is the largest open set contained in E and* $\operatorname{Cl} E$ *is the smallest closed set containing E.)*

(iii) Show that ∂E is closed.

(iv) Suppose that we work in $\mathbb{R}^m$ with the usual metric. Show that if E is bounded, then so is $\operatorname{Cl} E$.

Solution. There are a wide variety of ways of doing this exercise. Any way that works is fine.

(i) If $x \in \operatorname{Int} E$, we can find a $\delta > 0$ such that $B(x, 2\delta) \subseteq E$. If $y \in B(x, \delta)$, then, by the triangle inequality,

$$z \in B(y, \delta) \Rightarrow z \in B(x, 2\delta) \subseteq E.$$

Thus $\operatorname{Int} E$ is open.

If V is open and $V \subseteq E$, then, if $v \in V$, there exists a $\delta > 0$ with $B(v, \delta) \subseteq V \subseteq E$. Thus $V \subseteq \operatorname{Int} E$.

(ii) If $x_n \in \operatorname{Cl} E$ and $x_n \to x$, then we can find $y_n \in E$ such that $d(y_n, x_n) < 1/n$. By the triangle inequality,

$$d(y_n, x) \leq d(y_n, x_n) + d(x_n, x) \to 0 + 0 = 0$$

so $x \in \operatorname{Cl} E$.

If F is closed and $F \supseteq E$, then, whenever $x_n \in E$ and $x_n \to x$, we have $x_n \in F$, so $x \in F$. Thus $F \supseteq \operatorname{Cl} E$.

(iii) The complement of an open set is closed and the intersection of two closed sets is closed, so

$$\partial E = \operatorname{Cl} E \cap (\operatorname{Int} E)^c$$

is closed.

(iv) If E is closed, then we can find an $R > 0$ such that $E \subseteq \bar{B}(0, R)$. Since $\bar{B}(0, R)$ is closed, $\operatorname{Cl} E \subseteq \bar{B}(0, R)$. □

Recall that, if ϕ is a real valued function in $\mathbb{R}^m$ with sufficiently many derivatives, then we write

$$\nabla^2 \phi = \sum_{j=1}^{m} \frac{\partial^2 \phi}{\partial x_j^2}.$$

In this section we look at solutions of Laplace's equation

$$\nabla^2 \phi = 0.$$

Our first collection of results lead up to a proof of uniqueness.

Lemma 3.3. *Let Ω be a bounded open subset of $\mathbb{R}^m$. Suppose $\phi : \operatorname{Cl} \Omega \to \mathbb{R}$ is continuous and satisfies*

$$\nabla^2 \phi > 0$$

on Ω. Then ϕ cannot attain its maximum on Ω.

Proof. We prove the result for $m = 2$. Since $\operatorname{Cl}\Omega$ is compact, we know that ϕ attains a maximum at some point $(x_0, y_0) \in \operatorname{Cl}\Omega$. We need to show that it is impossible that $(x_0, y_0) \in \Omega$.

Suppose, if possible, that $(x_0, y_0) \in \Omega$. Since Ω is open, we can find a $\delta > 0$ such that $B\big((x_0, y_0), \delta\big) \subseteq \Omega$. Consider the function $f(y) = \phi(x_0, y)$ defined for $y \in (y_0 - \delta, y_0 + \delta)$. We have f twice differentiable with a maximum at y_0. Thus, by 1A analysis, $f''(y_0) \leq 0$. It follows that

$$\frac{\partial^2 \phi}{\partial y^2}(x_0, y_0) \leq 0.$$

The same argument applies for the partial derivatives with respect to x, so

$$\nabla^2 \phi(x_0, y_0) = \frac{\partial^2 \phi}{\partial x^2}(x_0, y_0) + \frac{\partial^2 \phi}{\partial y^2}(x_0, y_0) \leq 0$$

contradicting our hypotheses. □

Theorem 3.4. *Let Ω be a bounded open subset of $\mathbb{R}^m$. Suppose $\phi : \operatorname{Cl}\Omega \to \mathbb{R}$ is continuous on $\operatorname{Cl}\Omega$ and satisfies*

$$\nabla^2 \phi = 0$$

on Ω. Then ϕ attains its maximum on $\partial\Omega$.

Proof. Again we prove the result for $m = 2$. Let $\psi(x, y) = x^2 + y^2$. Since ψ is continuous and $\operatorname{Cl}\Omega$ is compact, we know that there exists a M with $M \geq \psi(x, y)$ for all $(x, y) \in \operatorname{Cl}\Omega$. By direct calculation, $\nabla^2 \psi = 4$ everywhere.

Set $\phi_n = \phi + n^{-1}\psi$. Then ϕ_n satisfies the conditions of Lemma 3.3 with $\epsilon = 4/n$. It follows that there is an $\mathbf{x}_n = (x_n, y_n) \in \partial\Omega$ with

$$\phi_n(\mathbf{x}_n) \geq \phi_n(\mathbf{t})$$

for all $\mathbf{t} \in \operatorname{Cl}\Omega$. Automatically,

$$\phi(\mathbf{x}_n) \geq \phi(\mathbf{t}) - 8M/n.$$

Since $\partial\Omega$ is compact, we can find an $\mathbf{x} \in \partial\Omega$ and $n(j) \to \infty$ such that $\mathbf{x}_{n(j)} \to \mathbf{x}$. By continuity

$$\phi(\mathbf{x}) \geq \phi(\mathbf{t})$$

for all $\mathbf{t} \in \operatorname{Cl}\Omega$. □

Exercise 3.5. *Let Ω be a bounded open subset of $\mathbb{C}$. Suppose that*

$$f : \operatorname{Cl}\Omega \to \mathbb{C}$$

is continuous and f is analytic on Ω. Recall that the real and imaginary parts of f satisfy Laplace's equation on $\operatorname{Int}\Omega$. Show that $|f|$ attains its maximum on $\partial\Omega$.
[Hint: Why can we assume that $\Re f(z_0) = |f(z_0)|$ at any particular point z_0?]

Solution. The map $z \mapsto |f(z)|$ is continuous so, by compactness, there exists a $z_0 = x_0 + iy_0 \in \operatorname{Cl}\Omega$ with $|f(z_0)| \geq |f(z)|$ for all $z \in \operatorname{Cl}\Omega$. By replacing $f(z)$ by $e^{i\theta} f(z)$, we may assume that $f(z_0)$ is real and positive.

Write $f(x + iy) = u(x, y) + iv(x, y)$ with u and v real. We have

$$u(x_0, y_0) = |f(z_0)| \geq |f(x + iy)| \geq u(x, y)$$

and u satisfies Laplace's equation. Thus there exists a $x_1 + iy_1 = z_1 \in \partial\Omega$ such that $u(x_1, y_1) = u(x_0, y_0)$ and so $|f(z_1)| \geq |f(z)|$ for all $z \in \operatorname{Cl}\Omega$. □

Theorem 3.6. *Let Ω be a bounded open subset of $\mathbb{R}^m$. Suppose that the functions $\phi, \psi : \operatorname{Cl}\Omega \to \mathbb{R}$ are continuous and satisfy*

$$\nabla^2\phi = 0, \ \nabla^2\psi = 0$$

on Ω. Then, if $\phi = \psi$ on $\partial\Omega$, it follows that $\phi = \psi$ on $\operatorname{Cl}\Omega$.

Proof. Observe that, if $\tau = \phi - \psi$, then τ satisfies the conditions of Theorem 3.4 and so attains its maximum on $\partial\Omega$. But $\tau = 0$ on $\partial\Omega$. Thus $\tau(\mathbf{x}) \leq 0$ for $\mathbf{x} = \operatorname{Cl}\Omega$. The same argument applied to $-\tau$ shows that $-\tau(\mathbf{x}) \leq 0$ for $\mathbf{x} = \operatorname{Cl}\Omega$. Thus $\tau = 0$ on $\operatorname{Cl}\Omega$ and we are done. □

You proved a version of Theorem 3.6 in Part 1A but only for 'nice boundaries' and functions that behaved 'nicely' near the boundary.

You have met the kind of arguments used above when you proved Rolle's theorem in 1A. Another use of this argument is given in Exercise 18.13 which provides a nice revision for this section.

In Part 1A you assumed that you could always solve Laplace's equation. The next exercise (which is part of the course) shows that this is not the case.

Exercise 3.7. *(i) Let*

$$\Omega = \{\mathbf{x} \in \mathbb{R}^2 \, : \, 0 < \|\mathbf{x}\| < 1\}.$$

Show that Ω *is open, that*

$$Cl\Omega = \{\mathbf{x} \in \mathbb{R}^2 \,:\, \|\mathbf{x}\| \leq 1\}$$

and

$$\partial\Omega = \{\mathbf{x} \in \mathbb{R}^2 \,:\, \|\mathbf{x}\| = 1\} \cup \{\mathbf{0}\}.$$

(ii) Suppose that $\phi : \operatorname{Cl}\Omega \to \mathbb{R}$ *is continuous that* ϕ *is twice differentiable on* Ω *and satisfies*

$$\nabla^2\phi = 0$$

on Ω *together with the boundary conditions*

$$\phi(\mathbf{x}) = \begin{cases} 0 & \textit{if } \|\mathbf{x}\| = 1, \\ 1 & \textit{if } \mathbf{x} = \mathbf{0}. \end{cases}$$

Use the uniqueness of solutions of Laplace's equation to show that ϕ *must be radially symmetric in the sense that*

$$\phi(\mathbf{x}) = f(\|\mathbf{x}\|)$$

for some function $f : [0,1] \to \mathbb{R}$*.*

(iii) Show that

$$\frac{d}{dr}\big(rf(r)\big) = 0$$

for $0 < r < 1$ *and deduce that* $f(r) = A + B\log r$ $[0 < r < 1]$ *for some constants* A *and* B*.*

(iv) Conclude that the function ϕ *described in (ii) can not exist.*

Solution. (i) If $\mathbf{x} \in \Omega$, then, setting

$$\delta = \min\{\|\mathbf{x}\|, 1 - \|\mathbf{x}\|\},$$

we have $\delta > 0$ and $B(\mathbf{x}, \delta) \subseteq \Omega$. Thus Ω is open.

Observe that $(0, 1/n) \to (0,0)$, so $\mathbf{0} \in \operatorname{Cl}\Omega$. Again, if $\|\mathbf{x}\| = 1$, then $(1 - 1/n)\mathbf{x} \to \mathbf{x}$, so $\mathbf{x} \in \operatorname{Cl}\Omega$. Thus $\operatorname{Cl}\Omega \supseteq \bar{B}(\mathbf{0}, 1)$. Since $\bar{B}(\mathbf{0}, 1)$ is closed $\operatorname{Cl}\Omega = \bar{B}(\mathbf{0}, 1)$.

Finally,

$$\partial\Omega = \operatorname{Cl}\Omega \setminus \operatorname{Int}\Omega = \operatorname{Cl}\Omega \setminus \Omega = \{\mathbf{x} \in \mathbb{R}^2 \,:\, \|\mathbf{x}\| = 1\} \cup \{\mathbf{0}\}.$$

(ii) Let T be a rotation with centre the origin. If $\psi = \phi T$, then (using the chain rule if you do not know the result already from applied courses)

$$\nabla^2\psi = 0.$$

But

$$\psi(\mathbf{x}) = \begin{cases} 0 & \text{if } \|\mathbf{x}\| = 1, \\ 1 & \text{if } \mathbf{x} = \mathbf{0}. \end{cases}$$

Thus, by uniqueness, $\psi = \phi$ and so, since T was an arbitrary rotation,

$$\phi(\mathbf{x}) = f(\|\mathbf{x}\|)$$

for some function $f : [0, 1] \to \mathbb{R}$.

(iii) The chain rule gives

$$\frac{\partial \phi}{\partial x} = f'(r)\frac{x}{r} \text{ and } \frac{\partial^2 \phi}{\partial x^2} = f''(r)\frac{x^2}{r^2} + f'(r)\left(\frac{1}{r} - \frac{x^2}{r^3}\right)$$

so, using the parallel result for derivatives with respect to y,

$$\nabla^2 \phi = f''(r) + f'(r)r^{-1} = r^{-1}\frac{d}{dr}\big(rf(r)\big).$$

(Or we can just quote this result from applied courses.) Thus

$$\frac{d}{dr}\big(rf(r)\big) = 0$$

so $rf'(r) = B$ and $f(r) = A + B\log r$ for appropriate constants A and B.

(iv) We need $f(r) \to 1$ as $r \to 0+$, so $B = 0$ and $A = 1$. This gives $f(1) = 1$, contradicting the condition $\phi(\mathbf{x}) = 0$ if $\|\mathbf{x}\| = 1$. □

This result is due to Zaremba, one of the founding fathers of Polish mathematics. Later Lebesgue produced a three dimensional example (the Lebesgue thorn) which will be briefly discussed by the lecturer, but does not form part of the course.

4 Fixed points

In Part 1A we proved the intermediate value theorem.

Theorem 4.1. *If $f : [a, b] \to \mathbb{R}$ is continuous function and $f(a) \geq c \geq f(b)$, then we can find a $y \in [a, b]$ such that $f(y) = c$.*

We then used it to the following very pretty fixed point theorem.

Theorem 4.2. *If $f : [a, b] \to [a, b]$ is a continuous function, then we can find a $w \in [a, b]$ such that $f(w) = w$.*

Notice that we can reverse the implication and use Theorem 4.2 to prove Theorem 4.1. (See Exercise 4.9.)

The object of this section is to extend the fixed point theorem to two dimensions.

Theorem 4.3. *Let* $\bar{D} = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$*. If* $f : \bar{D} \to \bar{D}$ *is a continuous function, then we can find a* $\mathbf{w} \in \bar{D}$ *such that* $f(\mathbf{w}) = \mathbf{w}$*.*

Although we will keep strictly to two dimensions the reader should note that the result and many of its consequences holds in n dimensions.

Notice also that the result can be extended to (closed) squares, triangles and so on.

Lemma 4.4. *Let* $\bar{D} = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}$*. Suppose that* $g : \bar{D} \to A$ *is a bijective function with* g *and* g^{-1} *continuous. Then if* $F : A \to A$ *is a continuous function, we can find an* $a \in A$ *such that* $F(a) = a$*.*

Proof. Observe that $f = g^{-1}Fg$ is a continuous function from $\bar{D}$ to $\bar{D}$ and so, by Theorem 4.3, has a fixed point w. Set $a = g(w)$. □

From now on we shall use extensions of the type given for Lemma 4.4 without comment.

The proof of Theorem 4.3 will take us some time. It consists in showing that a number of interesting statements are equivalent. The proof thus consists of lemmas of the form $A \Rightarrow B$ or$B \Leftrightarrow C$, ... I suggest that the reader considers each of these implications individually and then steps back to see how they hang together.

We write

$$D = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| < 1\},\ \bar{D} = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| \leq 1\}, \partial D = \{\mathbf{x} \in \mathbb{R}^2 : \|\mathbf{x}\| = 1\}.$$

Theorem 4.5. *The following two statements are equivalent.*

(i) If $f : \bar{D} \to \bar{D}$ *is continuous, then we can find a* $\mathbf{w} \in \bar{D}$ *such that* $f(\mathbf{w}) = \mathbf{w}$*. (We say that every continuous function of the closed disc into itself has a fixed point.)*

(ii) There does not exist a continuous function $g : \bar{D} \to \partial D$ *with* $g(\mathbf{x}) = \mathbf{x}$ *for all* $\mathbf{x} \in \partial D$*. (We say that there is no* retraction mapping *from* $\bar{D}$ *to* ∂D*.)*

Proof. (i)⇒ (ii) Suppose, if possible, that there exists a continuous function $g : \bar{D} \to \partial D$ with $g(\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \partial D$. If T is a rotation through π about $\mathbf{0}$, then $f = T \circ g$ is a continuous function from $\bar{D}$ to itself with no fixed points, contradicting (i).

(ii)$\Rightarrow$ (i) Suppose, if possible, that $f : \bar{D} \to \bar{D}$ is a continuous function with no fixed points. If we define

$$E = \{(\mathbf{x}, \mathbf{y}) \in \bar{D}^2 \ : \ \mathbf{x} \neq \mathbf{y}\}$$

and $u : E \to \partial D$ by taking $u(\mathbf{x}, \mathbf{y})$ to be the point where the straight line joining $\mathbf{x}$ to $\mathbf{y}$ in the indicated direction cuts ∂D then u is continuous. (We shall take this as geometrically obvious. The algebraic details are messy. The really conscientious student can do Exercise 18.14.) Using the chain rule for continuous functions, we see that

$$g(\mathbf{x}) = u\big(\mathbf{x}, f(\mathbf{x})\big)$$

defines a retraction mapping from $\bar{D}$ to ∂D, contradicting (ii). □

Let $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ be unit vectors making angles of $\pm 2\pi/3$ with each other. We take T to be the closed triangle with vertices $\mathbf{a}_1$, $\mathbf{a}_2$ and $\mathbf{a}_3$ and sides I, J and K.

(For those who insist on things being spelt out

$$T = \{\lambda_1\mathbf{a}_1 + \lambda_2\mathbf{a}_2 + \lambda_3\mathbf{a}_3 \ : \ \lambda_1 + \lambda_2 + \lambda_3 = 1, \ \lambda_j \geq 0\}$$

but though such ultra precision has its place, that place is not this course.)

The next collection of equivalences is easy to prove.

Lemma 4.6. *The following three statements are equivalent.*

(i) There is no retraction mapping from $\bar{D}$ to ∂D.

(ii) Let

$$\tilde{I} = \{(\cos\theta, \sin\theta) \ : \ 0 \leq \theta \leq 2\pi/3\}, \ \tilde{J} = \{(\cos\theta, \sin\theta) \ : \ 2\pi/3 \leq \theta \leq 4\pi/3\}$$
$$\text{and } \tilde{K} = \{(\cos\theta, \sin\theta) \ : \ 4\pi/3 \leq \theta \leq 2\pi\}.$$

Then there does not exist a continuous function $\tilde{k} : \bar{D} \to \partial D$ with

$$\tilde{k}(\mathbf{x}) \in \tilde{I} \text{ for all } \mathbf{x} \in \tilde{I}, \ \tilde{k}(\mathbf{x}) \in \tilde{J} \text{ for all } \mathbf{x} \in \tilde{J}, \ \tilde{k}(\mathbf{x}) \in \tilde{K} \text{ for all } \mathbf{x} \in \tilde{K}.$$

(iii) There does not exist a continuous function $k : T \to \partial T$ with

$$k(\mathbf{x}) \in I \text{ for all } \mathbf{x} \in I, \ k(\mathbf{x}) \in J \text{ for all } \mathbf{x} \in J, \ k(\mathbf{x}) \in K \text{ for all } \mathbf{x} \in K.$$

Proof. (i)$\Rightarrow$(ii) Suppose, if possible, that $\tilde{k}$ exists with the properties stated in (ii), Then, if T is a rotation through π, about $\mathbf{0}$, we see that $f = T \circ \tilde{k}$ is a continuous map from $\bar{D}$ to $\bar{D}$ without a fixed point. By Theorem 4.5 this contradicts (i).

(ii)$\Rightarrow$(i) If $\tilde{k}$ is a continuous retract from $\bar{D}$ to ∂D, then it certainly satisfies (ii).

(iii)$\Leftrightarrow$(ii) We use an argument of the type used for Lemma 4.4. □

We now prove a slightly more difficult equivalence.

Lemma 4.7. *The following two statements are equivalent,*
(i) There does not exist a continuous function $h : T \to \partial T$ *with*

$$h(\mathbf{x}) \in I \textit{ for all } \mathbf{x} \in I,\ h(\mathbf{x}) \in J \textit{ for all } \mathbf{x} \in J,\ h(\mathbf{x}) \in K \textit{ for all } \mathbf{x} \in K.$$

(ii) If A, B *and* C *are closed subsets of* T *with* $A \supseteq I$, $B \supseteq J$ *and* $C \supseteq K$ *and* $A \cup B \cup C = T$, *then* $A \cap B \cap C \neq \varnothing$.

Lemma 4.7. (ii)⇒(i) Let $h : \bar{T} \to \partial T$ be continuous with $h(I) \subseteq I$, $h(J) \subseteq J$, $h(K) \subseteq K$. Let $A = h^{-1}(I)$, $B = h^{-1}(J)$, $C = h^{-1}(K)$. Since h is continuous A, B and C are closed. Since $I \cup J \cup K = \partial D$, $A \cup B \cup C = \bar{D}$. But

$$A \cap B \cap C = h^{-1}(I) \cap h^{-1}(J) \cap h^{-1}(K) = h^{-1}(I \cap J \cap K) = h^{-1}(\varnothing) = \varnothing$$

contradicting (ii).
(i)⇒(ii) Suppose that A, B and C are closed subsets of T with $A \supseteq I$, $B \supseteq J$, $C \supseteq K$, $A \cup B \cup C = T$, and $A \cap B \cap C = \varnothing$.

We consider T as the triangle

$$T = \{(x, y, z) \in \mathbb{R}^3 \,:\, x + y + z = 1,\ x,\ y,\ z \geq 0\}.$$

(In my school days we called these 'barycentric coordinates'.) If $\mathbf{x} \in T$, we know that $\mathbf{x}$ lies in at most two of the sets A, B and C so (by Lemma 2.5) at least one of $d(\mathbf{x}, A)$, $d(\mathbf{x}, B)$ and $d(\mathbf{x}, C)$ is non-zero. Thus

$$h(\mathbf{x}) = \frac{1}{d(\mathbf{x}, A) + d(\mathbf{x}, B) + d(\mathbf{x}, C)} \big(d(\mathbf{x}, A), d(\mathbf{x}, B), d(\mathbf{x}, C)\big)$$

defines a continuous function $h : T \to T$. If $\mathbf{x} \in I$, then $d(\mathbf{x}, A) = 0$ and so $h(\mathbf{x}) \in I$. Similarly $h(J) \subseteq J$ and $h(K) \subseteq K$ contradicting (i). □

We shall prove statement (ii) of Lemma 4.7 from which the remaining statements will then follow. The key step is Sperner's Lemma.

Lemma 4.8. *Consider a triangle* DEF *divided up into a triangular grid. If all the vertices of the grid are coloured red, green or blue and every vertex on the side* DE *of the big triangle (with the exception of* E*) are coloured red, every vertex of* EF *(with the exception of* F*) green and every vertex of* FD *(with the exception of* D*) blue then there is a triangle of the grid all of whose vertices have different colours.*

Proof. Given an edge of the grid joining vertices u and v we assign a value $E(u,v)$ to the edge by a rule which ensures that, if u and v have the same colour, $E(u,v) = 0$, if u and v, have different colours X and Y, then $E(u,v) = \zeta(X,Y)$ with $\zeta(X,Y) = -\zeta(Y,X)$ and $\zeta(X,Y) = \pm 1$.

This table which follows gives an example.

colour u	colour v	$E(u,v)$
R	R	0
R	G	1
R	B	-1
G	R	-1
G	G	0
G	B	1
B	R	1
B	G	-1
B	B	0

If uvw is a grid triangle then, by inspection, the sum of the edge values (going round anticlockwise) is zero unless all of the vertices have different colours. By considering internal cancellation, the total sum of the edge values is the sum of the edge values going round the outer edge and this is non-zero. Thus one of the grid triangles must have all three vertices of different colours. □

Proof of Lemma 4.7 (ii). Suppose that A, B and C are closed subsets of T with $A \supseteq I$, $B \supseteq J$ and $C \supseteq K$ and $A \cup B \cup C = T$.

Take a triangular grid formed by n equally spaced parallel lines for each of the three sides dividing T into a grid of congruent triangles. Colour the vertices red, blue or green so that all the red vertices lie in A, all the blue vertices lie in B and all the green vertices lie in C, making sure that the outside edges are coloured as required by Lemma 4.8.

Lemma 4.8 tells us that there is a grid triangle with vertex $\mathbf{a}_n$ red, so in A, vertex $\mathbf{b}_n \in B$ and $\mathbf{c}_n \in C$. By compactness, we can find $n(j) \to \infty$ and $\mathbf{x} \in T$ such that $\mathbf{a}_{n(j)} \to \mathbf{x}$ and so $\mathbf{b}_{n(j)} \to \mathbf{x}$, $\mathbf{c}_{n(j)} \to \mathbf{x}$. Since A, B and C are closed $\mathbf{x} \in A \cap B \cap C$, so $A \cap B \cap C \neq \varnothing$ □

We can now prove the statement (ii) of Lemma 4.7 and so of Theorem 4.3 and all its equivalent forms.

The following pair of exercises (set as Exercises 18.7 and 18.8) may be helpful in thinking about the arguments of this section.

Exercise 4.9. *The following four statements are equivalent.*

(i) If $f : [0,1] \to [0,1]$ is continuous, then we can find a $w \in [0,1]$ such that $f(w) = w$.

(ii) There does not exist a continuous function $g : [0,1] \to \{0,1\}$ with $g(0) = 0$ and $g(1) = 1$. (In topology courses we say that $[0,1]$ is connected.)

(iii) If A and B are closed subsets of $[0,1]$ with $0 \in A$, $1 \in B$ and $A \cup B = [0,1]$ then $A \cap B \neq \varnothing$.

(iv) If $h : [0,1] \to \mathbb{R}$ is continuous and $h(0) \leq c \leq h(1)$, then we can find a $y \in [0,1]$ such that $h(y) = c$.

Exercise 4.10. *Suppose that we colour the points r/n red or blue $[r = 0, 1, \dots, n]$ with 0 red and 1 blue. Show that there are a pair of neighbouring points u/n, $(u+1)/n$ of different colours. Use this result to prove statement (iii) of Exercise 4.9.*

Sperner's lemma can be extended to higher dimensions and once this is done the remainder of our proofs together with Brouwer's theorem extend with simple changes to higher dimensions.

Here is an example of the use of Brouwer's theorem.

Exercise 4.11. *Suppose that $A = (a_{ij})$ is a 3×3 matrix such that $a_{ij} \geq 0$ for all $1 \leq i, j \leq 3$ and $\sum_{i=1}^{3} a_{ij} = 1$ for all $1 \leq j \leq 3$. Let*

$$T = \{\mathbf{x} \in \mathbb{R}^3 \ : \ x_j \geq 0 \text{ for all } j \text{ and } x_1 + x_2 + x_3 = 1\}.$$

By considering the effect of A on T, show that A has an eigenvector lying in T with eigenvalue 1.

Solution. T is a closed triangle in the appropriate plane. If $\mathbf{X} \in T$ and we write $\mathbf{y} = T\mathbf{x}$, then $y_i \geq 0$ and

$$\sum_{i=1}^{3} y_i = \sum_{i=1}^{3}\sum_{j=1}^{3} a_{ij}x_j = \sum_{j=1}^{3}\sum_{i=1}^{3} a_{ij}x_j = \sum_{j=1}^{3} x_j = 1,$$

so $\mathbf{y} \in T$. Thus T is a continuous map of T into itself and has a fixed point $\mathbf{e}$. We observe that $\mathbf{e}$ is an eigenvector lying in T with eigenvalue 1. □

If you have not done the 1B Markov chains course Exercise 4.11 may appear somewhat artificial. However, if you have done that course, you will see that is not.

Exercise 4.12. *(Only for those who understand the terms used. This is not part of the course.) Use the argument of Exercise 4.11 to show that every 3 state Markov chain has an invariant measure. (Remember that in Markov chains 'the i's and j's swap places'.) What result can you obtain under the assumption that Brouwer's theorem holds in higher dimensions?*

Brouwer's theorem is rather deep. Here is a result which can be proved using it.

Exercise 4.13. *Show that if $ABCD$ is a square and γ is a continuous path joining A and C whilst τ is a continuous path joining B and D, then γ and τ intersect,*
[*See Exercise 18.11 for a more detailed statement and a description of the proof.*]

Another interesting result is given as Exercise 18.9.

5 Non-zero sum games

It is said that converting the front garden of a house into a parking place raises the value of a house, but lowers the value of the other houses in the road. Once everybody has done the conversion, the value of each house is lower than before the process started.

Let us make a simple model of such a situation involving just two people with just two choices to see what we can say about it.

Suppose that Albert has the choice of doing A_1 or A_2 and Bertha the choice of doing B_1 or B_2. If A_i and B_j occur, then Albert gets a_{ij} units and Bertha gets b_{ij} units. If you went to the 1B course on optimisation you learnt how to deal with the case when $a_{ij} = -b_{ij}$ (this is called zero-sum case since $a_{ij} + b_{ij} = 0$ and Albert's loss is Bertha's gain). Albert and Bertha agree that Albert will choose A_i with probability p_i and Bertha will choose B_j with probability q_j. The expected value of the arrangement to Albert is

$$A(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij} p_i q_j$$

and the expected value of the arrangement to Bertha

$$B(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^{2} \sum_{j=1}^{2} b_{ij} p_i q_j$$

In 1B you saw that *in the zero-sum case* there is a choice of $\mathbf{p}$ and $\mathbf{q}$ such that if Albert chooses $\mathbf{p}$ and Bertha $\mathbf{q}$ then *even if he knows Bertha's choice* Albert will not change his choice and *even if she knows Albert's choice* Bertha will not change her choice.

So far so good, but if we consider the non zero-sum case we can imagine other situations in which Albert chooses $\mathbf{p}$ and then Bertha chooses $\mathbf{q}$

but, now knowing Bertha's choice, Albert changes his choice to $\mathbf{p}'$ and then, knowing Albert's new choice, Bertha changes to $\mathbf{q}'$ and then .... The question we ask ourselves is whether there is a 'stable choice' of $\mathbf{p}$ and $\mathbf{q}$ such that neither party can do better by unilaterally choosing a new value. This question is answered by a remarkable theorem of Nash.

Theorem 5.1. *Suppose a_{ij} and b_{ij} are real numbers. Let $E = \{(p, q) : 1 \geq p, q \geq 0\}$, set $p_1 = p$, $p_2 = 1 - p$, $q_1 = q$, $q_2 = 1 - q$,*

$$A(p, q) = \sum_{i=1}^{2} \sum_{j=1}^{2} a_{ij} p_i q_j \text{ and } B(p, q) = \sum_{i=1}^{2} \sum_{j=1}^{2} b_{ij} p_i q_j.$$

Then we can find $(p^*, q^*) \in E$ such that

$$B(p^*, q^*) \geq B(p^*, q) \text{ for all } (p^*, q) \in E$$

and

$$A(p^*, q^*) \geq A(p, q^*) \text{ for all } (p, q^*) \in E.$$

Proof. Let $\tilde{E} = \{(p, 1-p, q, 1-q) : 0 \leq p, q \leq 1\}$. (Thus $\tilde{E}$ is a two dimensional square embedded in $\mathbb{R}^4$.)

Suppose $(\mathbf{p}, \mathbf{q}) \in \tilde{E}$. Write

$$u_1(\mathbf{p}, \mathbf{q}) = \max\{0, A(1, 0, \mathbf{q}) - A(\mathbf{p}, \mathbf{q})\}.$$

Thus u_1 is Albert's expected gain if, instead of choosing $\mathbf{p}$ when Bertha chooses $\mathbf{q}$, he chooses $(1, 0)$ and Bertha maintains her choice *provided this is positive* and u_1 is zero otherwise. Similarly

$$u_2(\mathbf{p}, \mathbf{q}) = \max\{0, A(0, 1, \mathbf{q}) - A(\mathbf{p}, \mathbf{q})\},$$

so u_2 is Albert's expected gain if, instead of choosing $\mathbf{p}$ when Bertha chooses $\mathbf{q}$, he chooses $(0, 1)$ and Bertha maintains her choice *provided this is positive* and u_2 is zero otherwise. In the same way, we take

$$v_1(\mathbf{p}, \mathbf{q}) = \max\{0, B(\mathbf{p}, 1, 0) - B(\mathbf{p}, \mathbf{q})\}.$$

and

$$v_2(\mathbf{p}, \mathbf{q}) = \max\{0, B(\mathbf{p}, 0, 1) - B(\mathbf{p}, \mathbf{q})\}.$$

Now define

$$g(\mathbf{p}, \mathbf{q}) = (\mathbf{p}', \mathbf{q}')$$

with

$$\mathbf{p}' = \frac{\mathbf{p} + \mathbf{u}(\mathbf{p}, \mathbf{q})}{1 + u_1(\mathbf{p}, \mathbf{q}) + u_2(\mathbf{p}, \mathbf{q})}$$

and

$$\mathbf{q}' = \frac{\mathbf{q} + \mathbf{v}(\mathbf{p}, \mathbf{q})}{1 + v_1(\mathbf{p}, \mathbf{q}) + v_2(\mathbf{p}, \mathbf{q})}.$$

We observe that g is a well defined continuous function from $\tilde{E}$ into itself and so has a fixed point $(\mathbf{p}^*, \mathbf{q}^*)$.

We claim that $(\mathbf{p}^*, \mathbf{q}^*)$ is a Nash stable point.

Suppose, if possible, that $A\big((r, 1-r), \mathbf{q}^*\big) > A\big((p^*, 1-p^*), \mathbf{q}^*\big)$. Without loss of generality, we may suppose that $r > p^*$ so that

$$A\big((1,0), \mathbf{q}^*\big) > A\big((p^*, 1-p^*), \mathbf{q}^*\big)$$

and

$$A\big((0,1), \mathbf{q}^*\big) < A\big((p^*, 1-p^*), \mathbf{q}^*\big).$$

Thus $u_1(\mathbf{p}^*, \mathbf{q}^*) > 0$ and $u_2(\mathbf{p}^*, \mathbf{q}^*) = 0$, whence $\mathbf{p}^* = (1, 0)$ and $u_1(\mathbf{p}^*, \mathbf{q}^*) = 0$ which contradicts our earlier assertion.

We have shown that

$$A\big(\mathbf{p}^*, \mathbf{q}^*\big) \geq A\big((p, 1-p), \mathbf{q}^*\big)$$

for all $1 \geq p \geq 0$. The same argument shows that

$$B\big(\mathbf{p}^*, \mathbf{q}^*\big) \geq B\big(\mathbf{p}^*, (q, 1-q)\big)$$

for all $1 \geq q \geq 0$ so we are done. □

The pair (p^*, q^*) is called a Nash equilibrium point or Nash stable point.

The interested reader should have no difficulty in convincing herself (given Brouwer's fixed point theorem in the appropriate dimension) that the result can be extended to many participants with many choices to state that there is always a choice of probabilities such that *no single participant* has an incentive to change their choice[2]. Note that the game theory you did in 1B only applies to two players.

Unfortunately the stable points need not be unique. Suppose that Albert and Bertha have to choose scissors or paper. If they both choose scissors they get £1 each. If they both choose paper they get £2 each but if they disagree they get nothing. It is clear that the points corresponding to 'both choose paper' and 'both choose scissors' are stable. The same is true if when they disagree they both get nothing but Albert gets £1 and Bertha £2 if they both choose scissors whilst, if they both choose paper the payments are reversed.

[2]However, this applies only to single participants. There may be an incentive for two or more participants (if they can agree) to change their choices jointly.

Exercise 5.2. [Chicken] *Albert and Bartholomew drive cars fast at one another. If they both swerve they both lose 1 prestige points. If one swerves and the other does not the swerver looses 5 prestige points and the non-swerver gains 10 prestige points. If neither swerves they both loose 100 points. Identify the Nash equilibrium points.*

Solution. Suppose that A swerves with probability a and B with probability b. The value of the game to A is

$$V(a,b) = -ab + 10(1-a)b - 5a(1-b) - 100(1-a)(1-b).$$

If $0 < a < 1, 0 < b < 1$

$$\frac{\partial V}{\partial a}(a,b) = 95 - 106b,$$

so by symmetry we have a Nash equilibrium point $(a,b) = (95/106, 95/106)$. However

$$V(a,0) = -5a - 100(1-a) = 95a - 100, V(a,1) = 10 - 11a$$

so, again using symmetry, $(1,0)$ and $(0,1)$ are also Nash equilibrium points. $\square$

Notice that it is genuinely easy to solve toy problems like this when they appear in exercises and examinations. First look at the interior of the square and apply elementary calculus to find stationary points. Then look at the interior of each edge and apply elementary calculus to find stationary points. Finally look at the vertices. The same idea applies when there are three participants, but now we need to examine the interior of the cube, the interior of each face, the interior of each edge and the vertices in turn. Obviously if we attack the problem in this way, we run into the curse of dimensionality — each step is easy but the number of steps increases very rapidly with the number of participants. So far as I know, there is no way of avoiding this phenomenon. (But I know of no real life situation where we would wish to solve a high dimensional problem.)

It is also clear from examples like the one that began this section that even if the stable point is unique it may be unpleasant for all concerned[3]. However this is not the concern of the mathematician.

[3] McNamara, the US Defence Secretary at the time, was of the opinion that, during the Cuban crisis, all the participants behaved in a perfectly rational manner and only good luck prevented a full scale nuclear war.

6 Dividing the pot

Faced with problems like those of the previous section, the young and tender hearted often ask 'Why not cooperate?' It is, of course, true that under certain conditions people are willing to cooperate, but, even if these conditions are met, the question remains of how to divide up the gains due to cooperation.

Exercise 6.1. *(You will be asked to solve this as Exercise 19.4.) Consider two rival firms A and B engaged in an advertising war. So long as the war continues, the additional costs of advertising mean that the larger firm A loses 3 million pounds a year and the smaller firm B loses 1 million pounds a year. If they can agree to cease hostilities then A will make 8 million a year and B will make 1 million a year. How much should A pay B per year to achieve this end[4]?*

Nash produced a striking answer to this question. There are objections to his argument but I hope the reader will agree with me that it is a notable contribution.

In order to examine his answer we need to introduce the notion of a convex set.

Definition 6.2. *A subset E of $\mathbb{R}^m$ is* convex *if, whenever $\mathbf{u}, \mathbf{v} \in E$ and $1 \geq p \geq 0$, we have*

$$p\mathbf{u} + (1-p)\mathbf{v} \in E.$$

Nash considers a situation in which m players must choose a point $\mathbf{x} \in E$ where E is a closed, bounded, convex set in $\mathbb{R}^m$. The value of the outcome to the jth participant is x_j. To see why is reasonable to take E convex suppose that the participants can choose two points $\mathbf{u}$ and $\mathbf{v}$. The participants can agree among themselves to toss a suitable coin and choose $\mathbf{u}$ with probability p and $\mathbf{v}$ with probability $1-p$. The expected value of the outcome to the jth participant is $pu_j + (1-p)v_j$, that is to say, the value of the jth component of $p\mathbf{u} + (1-p)\mathbf{v}$.

The participants also know a point $\mathbf{s} \in E$ (the status quo) which will be the result if they can not agree on any other point.

Nash argues that a best point $\mathbf{x}^*$ if it exists must have the following properties.

(1) $x_j^* \geq s_j$ for all j. (Everyone must be at least as well off as if they failed to agree.)

[4] The reader may feel that it would be very difficult for rival firms to come to an agreement in this way. In fact, it appears to be so easy that most countries have strict laws against such behaviour.

(2) (Pareto Optimality) If $\mathbf{x} \in E$ and $x_j \geq x_j^*$ for all j, then $\mathbf{x} = \mathbf{x}^*$. (If there is a choice which makes some strictly better off and nobody worse off, then the participants should take it.)

(3) (Independence of irrelevant alternatives.) Suppose E' is a closed bounded convex set with $E' \supseteq E$ and $\mathbf{x}^{**}$ is a best point for E'. Then, if $\mathbf{x}^{**} \in E$ it follows that $\mathbf{x}^{**}$ is a best point for E.

(4) If E is symmetric (that is, if whenever $\mathbf{x} \in E$ and $y_1, y_2, \dots, y_m$ is some rearrangement of $x_1, x_2, \dots, x_m$, then $\mathbf{y} \in E$) and $\mathbf{s}$ is symmetric, then $x_1^* = x_2^* = \dots = x_m^*$. This corresponds to our beliefs about 'fairness'.

(5) Our final assumption is that we must treat the poor man's penny with the same respect as the rich man's pound. Suppose that $\mathbf{x}^*$ is a best point for E. If we change coordinates and consider

$$E' = \{\mathbf{x}' : x'_j = a_j x_j + b_j \text{ for } 1 \leq j \leq m \text{ and } \mathbf{x} \in E\}$$

with $a_j > 0$, then $\mathbf{y}^*$ with $y_j^* = a_j x_j^* + b_j$ is a best point for E'.

Exercise 6.3. *Show that the E' defined above is closed bounded and convex.*

Solution. If $\mathbf{x}'$, $\mathbf{y}' \in E'$ and $0 \leq t \leq 1$, then

$$x'_j = a_j x_j + b_j, \ y'_j = a_j y_j + b_j$$

with $\mathbf{x}$, $\mathbf{y} \in E$ and so

$$t x'_j + (1-t) y'_j = a_j(t x_j + (1-t) y_j) + b_j$$

for all j. But $t\mathbf{x} + (1-t)\mathbf{y} \in E$, since E is convex, so $t\mathbf{x}' + (1-t)\mathbf{y}' \in E'$ and is convex.

We now recall Theorem 2.6 and observe that E' is the continuous image of a compact set so compact. □

There is no difficulty in remembering these conditions since they each play a particular role in the proof. If the reader prefers initially only to deal with the case $m = 2$, she will lose nothing of the argument. We need a preliminary lemma.

Lemma 6.4. *If K is a convex set in $\mathbb{R}^n$ such that $(1, 1, \dots, 1) \in K$ and $\prod_{j=1}^n x_j \leq 1$ for all $\mathbf{x} \in K$ with $x_j \geq 0$ $[1 \leq j \leq n]$, then*

$$K \subseteq \{\mathbf{x} : x_1 + x_2 + \dots x_n \leq n\}.$$

Proof. If $\mathbf{x} \in K$ and $0 \leq t \leq 1$, then, since $\mathbf{1} \in K$ and K is convex, we have

$$(1-t)\mathbf{1} + t\mathbf{x} \in K$$

so, by our hypothesis,

$$\begin{aligned} 1 &\geq \prod_{j=1}^{n} \big(tx_j + (1-t)\big) = \prod_{j=1}^{n} \big(1 + t(x_j - 1)\big) \\ &= 1 + t\sum_{j=1}^{n}(x_j - 1) + t^2 P(t) \end{aligned}$$

where P is a polynomial with coefficients depending on $\mathbf{x}$. It follows that, if $0 \leq t \leq 1$, we have

$$0 \geq \sum_{j=1}^{n}(x_j - 1) + tP(t).$$

Allowing $t \to 0+$ gives

$$0 \geq \sum_{j=1}^{n}(x_j - 1),$$

which is the desired result. □

Theorem 6.5. *Suppose that we agree to the Nash conditions. If E is closed bounded convex set in $\mathbb{R}^m$, $\mathbf{s}$ is the status quo point and the function $f : E \to \mathbb{R}$ given by*

$$f(\mathbf{x}) = \prod_{j=1}^{m}(x_j - s_j)$$

has a maximum (in E) with $x_j - s_j > 0$ at $\mathbf{x}^*$, then $\mathbf{x}^*$ is the unique best point.

Proof. The Nash conditions mean that the problem is invariant under affine transformation (i.e. transformations of the type discussed in Exercise 6.3). Thus we may assume that $\mathbf{s} = \mathbf{0}$. If the hyperboloid $\prod_{j=1}^{n} y_j = K$ touches the convex set E' at $\mathbf{y}$ (with $y_j > 0$) then the transformation $x_j = K^{-1/n} y_j / y_j^*$ gives a hyperboloid $\prod_{j=1}^{n} x_j = 1$ touching a convex set E at $(1, 1, \ldots, 1)$.

Thus we may assume that $\mathbf{s} = \mathbf{0}$ and $x_1^* = x_2^* = \cdots = x_n^* = 1$.

By Lemma 6.4, we have

$$K \subseteq L = \{\mathbf{x} : \sum_{j=1}^{n} x_j \leq n\},$$

and, by the independence of irrelevant alternatives, if $\mathbf{x}^*$ is best for L, it is best for K. Now L is symmetric so any best point $\mathbf{x}$ for L must lie on $x_1 = x_2 = \ldots = x_n$. But, amongst these points, only $\mathbf{x}^*$ is Pareto optimal so we are done. □

We complete our discussion by observing that a best point always exists.

Lemma 6.6. *If E is closed bounded convex set in $\mathbb{R}^m$, $\mathbf{s} \in \operatorname{Int} E$ and the function $f : E \to \mathbb{R}$ given by*

$$f(\mathbf{x}) = \prod_{j=1}^{m} (x_j - s_j),$$

then there is a unique point in E with $x_j - s_j \geq 0$ where f attains a maximum.

Lemma 6.6. By compactness, there is a point $\mathbf{x}^*$ where f attains its maximum. By translation, we may suppose $\mathbf{s} = \mathbf{0}$ and, re-scaling the axes, we may suppose $\mathbf{x}^* = \mathbf{e} = (1, 1, \ldots, 1)$.

Lemma 6.4 tells us that

$$\{\mathbf{k} \in K : k_j \geq 0 \ \forall j\} \subseteq \{\mathbf{x} \in K : x_j \geq 0 \ \forall j \text{ and } x_1 + x_2 + \ldots + x_n = n\}.$$

The uniqueness of the maximum now follows from the conditions for equality in the arithmetic geometric inequality. □

'There is no patent for immortality under the the moon' but I suspect that Nash's results will be remembered long after the last celluloid copy of *A Beautiful Life* has crumbled to dust.

The book *Games, Theory and Applications* [6] by L. C. Thomas maintains a reasonable balance between the technical and non-technical and would make a good port of first call if you wish to learn more along these lines.

7 Approximation by polynomials

It is a guiding idea of both the calculus and of numerical analysis that 'well behaved functions look like polynomials'. Like most guiding principles, it needs to be used judiciously.

If asked to justify it, we might mutter something about Taylor's Theorem, but Cauchy produced the following example to show that this is not sufficient[5]

[5] When we do 1A this result is a counterexample but, by Part II, if we need a 'partition of unity' or a 'smoothing kernel', it has become an invaluable tool.

Exercise 7.1. *Let $E : \mathbb{R} \to \mathbb{R}$ be defined by*

$$E(t) = \begin{cases} \exp(-1/t^2) & \text{if } t \neq 0, \\ 0 & \text{if } t = 0. \end{cases}$$

(i) E is infinitely differentiable, except, possibly, at 0, with

$$E^{(n)}(t) = P_n(1/t)E(t)$$

for all $t \neq 0$ for some polynomial P_n.

(ii) E is infinitely differentiable everywhere with

$$E^{(n)}(0) = 0.$$

(iii) We have

$$E(t) \neq \sum_{n=0}^{\infty} \frac{E^{(n)}(0)}{n!} t^n$$

for all $t \neq 0$.

(It is very unlikely that you have not seen this exercise before, but if you have not you should study it.)

Solution for Exercise 7.1. (i) We use induction on n to show that E is n times differentiable with

$$E^{(n)}(t) = P_n(1/t)E(t)$$

for all $t \neq 0$ and some polynomial P_n.

The result is certainly true for $n = 0$ with $P_0 = 1$. If it is true for $n = m$, then the standard rules for differentiation show that $E^{(m)}$ is differentiable with

$$E^{(m+1)}(t) = t^{-2}P_m'(1/t)E(t) - 2t^{-3}P_m(1/t)E(t) = P_{m+1}(1/t)E(t)$$

for all $t \neq 0$ and the polynomial $P_{m+1}(s) = s^2 P_m'(s) - 2s^3 P_m(s)$.

(ii) We use induction on n to show that E is n times differentiable at 0 with

$$E^{(n)}(0) = 0.$$

The result is true for $n = 0$. If it is true for $n = m$, then

$$\frac{E^{(m)}(h) - E^{(m)}(0)}{h} = h^{-1}P(h^{-1})E(h) \to 0$$

as $h \to 0$, so it is true for $n = m + 1$.

(iii) We have

$$E(t) \neq 0 = \sum_{n=0}^{\infty} \frac{0}{n!} t^n = \sum_{n=0}^{\infty} \frac{E^{(n)}(0)}{n!} t^n$$

for all $t \neq 0$, as stated. □

We could also mutter something like 'interpolation'. The reader probably knows all the facts given in the next lemma.

Lemma 7.2. *Let x_0, x_1, ..., x_n be distinct points of $[a, b]$.*

(i) If $f : [a, b] \to \mathbb{R}$ then there is at most one polynomial of degree no greater than n with $P(x_j) = f(x_j)$ for $0 \leq j \leq n$.

(ii) Write

$$e_j(t) = \prod_{k \neq j} \frac{t - x_k}{x_j - x_k}.$$

If $f : [a, b] \to \mathbb{R}$ then

$$P = \sum_{j=0}^{n} f(x_j) e_j$$

is a polynomial of degree at most n with $P(x_i) = f(x_i)$ for $0 \leq i \leq n$. (Thus we can replace 'at most' by 'exactly' in (i).)

(iii) In the language of vector spaces, the e_j form a basis for the vector space of polynomials $\mathcal{P}_n$ of degree n or less.

Proof. (i) If P and Q have degree at most n and

$$P(x_j) = Q(x_j) = f(x_j)$$

for $0 \leq j \leq n$, then $P - Q$ is a polynomial of degree at most n vanishing at at least $n + 1$ points. Thus $P - Q = 0$, by Exercise 2.10, so $P = Q$.

(ii) We observe that $e_j(x_i) = 1$ if $i = j$, but $e_j(x_i) = 0$ otherwise and that e_j is a polynomial of degree n. Thus

$$P = \sum_{j=0}^{n} f(x_j) e_j$$

is a polynomial of degree at most n with

$$P(x_i) = \sum_{j=0}^{n} f(x_j) e_j(x_i) = f(x_i)$$

for $0 \le i \le n$.

(iii) It is easy to check that $\mathcal{P}_n$ is a vector space. Part (ii) shows that the e_j span $\mathcal{P}_n$. If

$$\sum_{j=0}^{n} \lambda_j e_j = 0,$$

then

$$\lambda_i = \sum_{j=0}^{n} \lambda_j e_j(x_i) = 0$$

for each i, so the the e_j are linearly independent. $\square$

However polynomials can behave in rather odd ways.

Theorem 7.3. *There exist polynomials T_n of degree n and U_{n-1} of degree $n-1$ such that*

$$T_n(\cos\theta) = \cos n\theta$$

for all θ and

$$U_{n-1}(\cos\theta) = \frac{\sin n\theta}{\sin\theta}$$

for $\sin\theta \ne 0$. The value of $U_{n-1}(\cos\theta)$ when $\sin\theta = 0$ is given by continuity and will be $\pm n$. The roots of U_{n-1} are $\cos(r\pi/n)$ with $1 \le r \le n-1$ and the roots of T_n are $\cos\big((r+\frac{1}{2})\pi/n\big)$ with $0 \le r \le n-1$.

The coefficient of t^n in T_n is 2^{n-1} for $n \ge 1$.

Proof. By de Moivre's theorem,

$$\begin{aligned}
\cos n\theta + i\sin n\theta &= (\cos\theta + i\sin\theta)^n \\
&= \sum_{r=0}^{n} i^r \binom{n}{r} (\cos\theta)^{n-r} (\sin\theta)^r \\
&= \sum_{0\le 2r\le n} (-1)^r \binom{n}{2r} (\cos\theta)^{n-2r} (\sin\theta)^{2r} \\
&\qquad + i\sin\theta \sum_{0\le 2r\le n-1} (-1)^r \binom{n}{2r+1} (\cos\theta)^{n-1-2r} (\sin\theta)^{2r} \\
&= \sum_{0\le 2r\le n} (-1)^r \binom{n}{2r} (\cos\theta)^{n-2r} \big(1-(\cos\theta)^2\big)^r \\
&\qquad + i\sin\theta \sum_{0\le 2r\le n-1} (-1)^r \binom{n}{2r+1} (\cos\theta)^{n-1-2r} \big(1-(\cos\theta)^2\big)^r \\
&= T_n(\cos\theta) + i\sin\theta U_{n-1}(\cos\theta),
\end{aligned}$$

where T_n is a polynomial of degree at most n and U_{n-1} a polynomial of degree at most $n-1$.

Taking real and imaginary parts, we obtain

$$T_n(\cos\theta) = \cos n\theta$$

for all θ and

$$U_{n-1}(\cos\theta) = \frac{\sin n\theta}{\sin\theta}$$

for $\sin\theta \neq 0$, $U_{n-1}(1) = n$, $U_{n-1}(-1) = (-1)^{n-1}n$. The roots of U_{n-1} and T_n can be read off directly and show that the two polynomials have full degree.

The coefficient of t^n in T_n is

$$\sum_{0\leq 2r\leq n} \binom{n}{2r} = \frac{1}{2}\big((1+1)^n + (1-1)^n\big) = 2^{n-1}$$

for $n \geq 1$. □

We call T_n the Chebychev[6] polynomial of degree n. The U_n are called Chebychev polynomials of the second kind. Looking at the Chebychev polynomials of the second kind, we see that we can choose a well behaved function f which is well behaved at $n+1$ reasonably well spaced points but whose nth degree interpolating polynomial is very large at some other point. It can be shown (though this is harder to prove) that this kind of thing can happen however we choose our points of interpolation.

A little thought shows that we are not even sure what it means for one function to look like another. It is natural to interpret f looks like g as saying that f and g are close in some metric. However there are a number of 'obvious' metrics. The next exercise will be familiar to almost all my audience.

Exercise 7.4. *Show that the following define metrics on the space $C([0,1])$ of continuous functions $f : [0,1] \to \mathbb{R}$.*

(i) $\|f\|_1 = \int_0^1 |f(t)|\,dt$ defines a norm with associated distance

$$d_1(f,g) = \|f-g\|_1 = \int_0^1 |f(t)-g(t)|\,dt.$$

(ii) The equation

$$\langle f,g\rangle = \int_a^b f(t)g(t)\,dt$$

[6] Or Tchebychev, hence the T.

defines an inner product on $C([a,b])$ *with associated norm* $\| \ \|_2$ *and so a distance*

$$d_2(f,g) = \|f-g\|_2 = \left(\int_a^b (f(t)-g(t))^2\,dt\right)^{1/2}$$

for the derived norm.

(iii) The equation $\|f\|_\infty = \sup_{t\in[0,1]}|f(t)|$ *defines a norm and so a distance*

$$d_3(f,g) = \|f-g\|_\infty = \sup_{t\in[0,1]}|f(t)-g(t)|.$$

Show that

$$\|f-g\|_\infty \le \|f-g\|_1 \le \|f-g\|_2.$$

Let

$$f_n(t) = \begin{cases} (1-nt) & \text{for } 0 \le t \le 1/n, \\ 0 & \text{otherwise.} \end{cases}$$

Compute $\|f_n\|_\infty/\|f_n\|_1$ *and* $\|f_n\|_1/\|f_n\|_2$*. Comment.*

Solution. The key result that we use in (i) and (ii) is that, if $f \in C([0,1])$, $f(t) \ge 0$ for all $t \in [0,1]$ and $\int_0^1 |f(t)|\,dt = 0$, then $f(t) = 0$ for all $t \in [0,1]$.

(i) Observe that

$$\|f\|_1 = \int_0^1 |f(t)|\,dt \ge 0$$

and that, if $\|f\|_1 = 0$, then

$$\int_0^1 |f(t)|\,dt = 0,$$

so $|f(t)| = 0$ for all t, so $f(t) = 0$ for all t and $f = 0$.

Further

$$\|\lambda f\|_1 = \int_0^1 |\lambda||f(t)|\,dt = |\lambda|\int_0^1 |f(t)|\,dt = |\lambda|\|f\|_1$$

and, since $|f(t)+g(t)| \le |f(t)| + |g(t)|$, we have

$$\|f+g\|_1 = \int_0^1 |f(t)+g(t)|\,dt \le \int_0^1 |f(t)|+|g(t)|\,dt = \|f\|_1 + \|g\|_1,$$

so we have a norm.

(ii) We have

$$\langle f,f\rangle = \int_0^1 f(t)^2\,dt \ge 0.$$

If $\langle f, f\rangle = 0$, then $\int_0^1 f(t)^2\,dt = 0$, so $f(t)^2 = 0$ for all t, so $f(t) = 0$ for all t and $f = 0$.

We have

$$\langle f, g\rangle = \int_0^1 f(t)g(t)\,dt = \int_0^1 g(t)f(t)\,dt = \langle g, f\rangle$$

and

$$\begin{aligned}\langle f+g, h\rangle &= \int_0^1 (f(t)+g(t))h(t)\,dt \\ &= \int_0^1 f(t)h(t)\,dt + \int_0^1 g(t)h(t)\,dt = \langle f, h\rangle + \langle g, h\rangle\end{aligned}$$

whilst

$$\langle \lambda f, g\rangle = \int_0^1 \lambda f(t)g(t)\,dt = \lambda \int_0^1 f(t)g(t)\,dt = \lambda\langle f, g\rangle,$$

so we have an inner product.

(iii) Observe that $|f(t)| \geq 0$,so

$$\|f\|_\infty = \sup_{t\in[0,1]} |f(t)| \geq 0,$$

that

$$\|f\|_\infty = 0 \Rightarrow \sup_{t\in[0,1]} |f(t)| = 0 \Rightarrow |f(t)| = 0\ \forall t \Rightarrow f = 0,$$

that

$$\|\lambda f\|_\infty = \sup_{t\in[0,1]} |\lambda f(t)| = \sup_{t\in[0,1]} |\lambda||f(t)| = |\lambda| \sup_{t\in[0,1]} |f(t)| = \lambda\|f\|_\infty,$$

and, that, since $|f(t)+g(t)| \leq |f(t)| + |g(t)|$,

$$\begin{aligned}\|f+g\|_\infty &= \sup_{t\in[0,1]} |f(t)+g(t)| \leq \sup_{t\in[0,1]} (|f(t)| + |g(t)|) \\ &\leq \sup_{t,s\in[0,1]} (|f(t)| + |g(s)|) = \|f\|_\infty + \|g\|_\infty,\end{aligned}$$

so we are done.

If f_n is as stated, $\|f_n\|_1 = 2\int_0^{1/n} nt\,dt = 1/n$, $\|f_n\|_\infty = 1$ and

$$\|f_n\|_2 = \left(2\int_0^{1/n} (nt)^2\,dt\right)^{1/2} = \left(\frac{2}{3n}\right)^{1/2}.$$

Thus $\|f_n\|_\infty / \|f_n\|_1 = n \to \infty$ and $\|f_n\|_1/\|f_n\|_2 = (3/2)^{1/2} n^{1/2} \to \infty$ as $n \to \infty$. We have genuinely different measures of difference. □

Each of these metrics has its advantages and all are used in practice. We shall concentrate on the metric d_3. We quote the following result from a previous course (where it is known as the General Principle of Uniform Convergence).

Theorem 7.5. *If $[a, b]$ is a closed interval and $C([a, b])$ is the space of continuous functions on $[a, b]$ then the uniform metric*

$$d(f, g) = \|f - g\|_\infty$$

is complete.

We have now obtained a precise question. If f is a continuous function can we find polynomials which are arbitrarily close in the uniform norm? This question was answered in the affirmative by Weierstrass in a paper published when he was 70 years old. Since then, several different proofs have been discovered. We present one due to Bernstein based on probability theory[7]

Before that, we need a definition and theorem which the reader will have met in a simpler form earlier.

Definition 7.6. *Let (X, d) and (Y, ρ) be metric spaces. We say that a function $f : X \to Y$ is uniformly continuous if, given $\epsilon > 0$, we can find a $\delta > 0$ such that $\rho\big(f(a), f(b)\big) < \epsilon$ whenever $d(a, b) < \delta$.*

Theorem 7.7. *If E is a bounded closed set in $\mathbb{R}^m$ and $f : E \to \mathbb{R}^p$ is continuous, then f is uniformly continuous.*

Proof. Suppose that f is not uniformly continuous. Then we can find an $\epsilon > 0$ and $\mathbf{x}_n$, $\mathbf{y}_n \in E$ such that

$$\|\mathbf{x}_n - \mathbf{y}_n\| \leq 1/n \text{ and } \|f(\mathbf{x}_n) - f(\mathbf{y}_n)\| \geq \epsilon.$$

By compactness, we can find $\mathbf{e} \in E$ and $n(j) \to \infty$ such that $\mathbf{x}_{n(j)} \to \mathbf{e}$. The triangle inequality tells us that $\mathbf{y}_{n(j)} \to \mathbf{e}$ and so

$$\|f(\mathbf{x}_{n(j)}) - f(\mathbf{y}_{n(j)})\| \leq \|f(\mathbf{x}_{n(j)}) - f(\mathbf{e})\| + \|f(\mathbf{y}_{n(j)}) - f(\mathbf{e})\| \to 0 + 0 = 0.$$

We have a contradiction. □

Although we require probability theory we only need deal with the simplest case of a random variable taking a finite number of values and, if the reader wishes, she need only prove the next result in that case.

[7] Some other proofs are given in Exercises 20.6, 20.7 and 20.8. It is the author's belief that one can not have too many (insightful) proofs of Weierstrass's theorem.

Theorem 7.8. [Chebychev's inequality] *If X is a real valued bounded random variable, then, writing*

$$\sigma^2 = \operatorname{var} X = \mathbb{E}(X - \mathbb{E}X)^2,$$

we have

$$\Pr(|X - \mathbb{E}X| \geq a) \leq \frac{\sigma^2}{a^2}$$

for all $a > 0$.

Proof. By replacing X by $Y = X - \mathbb{E}X$, we may suppose that $\mathbb{E}X = 0$.

Let

$$\mathbb{I}_{\mathbb{R}\setminus(-a,a)}(t) = \begin{cases} 0 & \text{if } |t| < a, \\ 1 & \text{otherwise.} \end{cases}$$

Then

$$\frac{t^2}{a^2} \geq \mathbb{I}_{\mathbb{R}\setminus(-a,a)}(t)$$

for all t, so, automatically,

$$\frac{X^2}{a^2} \geq \mathbb{I}_{\mathbb{R}\setminus(-a,a)}(X)$$

and

$$\frac{\sigma^2}{a^2} = \mathbb{E}\frac{X^2}{a^2} \geq \mathbb{E}\mathbb{I}_{\mathbb{R}\setminus(-a,a)}(X) = \Pr(|X| \geq a).$$

□

Theorem 7.9. [Bernstein] *Suppose $f : [0,1] \to \mathbb{R}$ is continuous. Let X_1, X_2, $\dots X_n$ be independent identically distributed random variables with $\Pr(X_r = 0) = 1 - t$ and $\Pr(X_r = 1) = t$ (think of tossing a biased coin). Let*

$$Y_n(t) = \frac{X_1 + X_2 + \dots + X_n}{n}$$

and let

$$p_n(t) = \mathbb{E}f(Y_n(t)).$$

Then

(i) p_n is polynomial of degree n. Indeed,

$$p_n(t) = \sum_{j=0}^{n} \binom{n}{j} f(j/n) t^j (1-t)^{n-j}.$$

(ii) $\|p_n - f\|_\infty \to 0$ as $n \to \infty$.

Proof. (i) We have

$$\begin{aligned} p_n(t) &= \mathbb{E} f(Y_n(t)) \\ &= \sum_{j=0}^{n} f(j/n) \Pr(X_1 + X_2 + \cdots + X_n = j) \\ &= \sum_{j=0}^{n} \binom{n}{j} f(j/n) t^j (1-t)^{n-j}. \end{aligned}$$

(ii) Automatically,

$$\mathbb{E} Y_n = \mathbb{E} \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{\mathbb{E} X_1 + \mathbb{E} X_2 + \cdots + \mathbb{E} X_n}{n} = \frac{nt}{n} = t$$

and, since the X_j are independent,

$$\begin{aligned} \operatorname{var} Y_n = \operatorname{var} \frac{X_1 + X_2 + \cdots + X_n}{n} &= n^{-2} \operatorname{var}(X_1 + X_2 + \cdots + X_n) \\ &= n^{-2}(\operatorname{var} X_1 + \operatorname{var} X_2 + \ldots + \operatorname{var} X_n) \\ &= n^{-1} \operatorname{var} X_1 = n^{-1} t(1-t) \leq n^{-1}. \end{aligned}$$

Let $\epsilon > 0$. By uniform continuity we can find an $\eta > 0$ such that $|f(t) - f(s)| \leq \epsilon$ for $|t - s| \leq \eta$ and $t, s \in [0, 1]$. Thus, using Chebychev's inequality,

$$\begin{aligned} |p_n(t) - f(t)| = \left|\mathbb{E}\big(f(Y_n) - f(t)\big)\right| &\leq \mathbb{E}|f(Y_n) - f(t)| \\ &\leq \epsilon \Pr(|Y_n - t| < \eta) + 2\|f\|_\infty \Pr(|Y_n - t| \geq \eta) \\ &\leq \epsilon + 2\|f\|_\infty \Pr(|Y_n - \mathbb{E} Y_n| \geq \eta) \\ &\leq \epsilon + 2\|f\|_\infty \eta^{-2}/n \leq 3\epsilon, \end{aligned}$$

provided only that $n \geq \epsilon^{-1}(2\|f\| + 1)\eta^{-2}$. Since ϵ is arbitrary, the result follows. □

Bernstein's result differs from many proofs of Weierstrass's theorem in giving an elegant *explicit* approximating polynomial.

8 Best approximation by polynomials

Bernstein's theorem gives an explicit approximating polynomial but, except in very special circumstances, not the best approximating polynomial. (Indeed, we have not yet shown that such a polynomial exists.)

Chebychev was very interested in this problem and gave a way of telling when we do have a best approximation.

Theorem 8.1. **[The Chebychev equiripple criterion]** *Let $f : [a, b] \to \mathbb{R}$ be a continuous function and P a polynomial of degree at most $n-1$. Suppose that we can find $a \leq a_0 < a_1 < \cdots < a_n \leq b$ such that, writing $\sigma = \|f - P\|_\infty$ we have either*

$$f(a_j) - P(a_j) = (-1)^j \sigma \text{ for all } 0 \leq j \leq n$$

or

$$f(a_j) - P(a_j) = (-1)^{j+1} \sigma \text{ for all } 0 \leq j \leq n.$$

Then $\|P - f\|_\infty \leq \|Q - f\|_\infty$ for all polynomials Q of degree $n-1$ or less.

Proof. Without loss of generality, suppose that

$$f(a_j) - P(a_j) = (-1)^j \sigma \text{ for all } 0 \leq j \leq n.$$

Suppose, if possible, that Q is a polynomial of degree $n-1$ or less such that $\|P - f\|_\infty > \|Q - f\|_\infty$.

We look at $R = P - Q$. Note first that R is a polynomial of degree at most $n-1$. If j is odd,

$$\begin{aligned} R(a_j) &= \big(P(a_j) - f(a_j)\big) + \big(f(a_j) - Q(a_j)\big) \\ &= |P(a_j) - f(a_j)| + \big(f(a_j) - Q(a_j)\big) \\ &\geq |P(a_j) - f(a_j)| - \|Q - f\|_\infty = \|P - f\|_\infty - \|Q - f\|_\infty > 0. \end{aligned}$$

and a similar argument shows that

$$R(a_j) < 0$$

when j is even.

The intermediate value theorem now tells that R has at least n zeros, so $R = 0$ and $P = Q$, contradicting our initial assumption. □

We apply this to find the polynomial of degree $n-1$ which gives the best approximation to t^n on $[-1, 1]$.

Theorem 8.2. *Write $S_n(t) = t^n - 2^{1-n} T_n(t)$, where T_n is the Chebychev polynomial of degree n. Then (if $n \geq 1$)*

$$\sup_{t \in [-1,1]} |t^n - Q(t)| \geq \sup_{t \in [-1,1]} |t^n - S_n(t)| = 2^{1-n}$$

for all polynomials Q of degree $n-1$.

Proof. If $t = \cos\theta$, then

$$t^n - S_n(t) = 2^{1-n} T_n(t) = 2^{1-n} \cos n\theta$$

Thus

$$|t^n - S_n(t)| \leq 2^{1-n}$$

for $t \in [-1, 1]$ and

$$t^n - S_n(t) = (-1)^j 2^{1-n}$$

for $t = \cos j\pi/n$ $[0 \leq j \leq n]$.

The stated result now follows from the equiripple criterion. □

Corollary 8.3. *We work on* $[-1, 1]$.

(i) If $P(t) = \sum_{j=0}^{n} a_j t^j$ *is a polynomial of degree* n *with* $|a_n| \geq 1$ *then* $\|P\|_\infty \geq 2^{-n+1}$.

(ii) We can find $\epsilon(n) > 0$ *with the following property. If* $P(t) = \sum_{j=0}^{n} a_j t^j$ *is a polynomial of degree at most* n *and* $|a_k| \geq 1$ *for some* $n \geq k \geq 0$ *then* $\|P\|_\infty \geq \epsilon(n)$.

Proof. (i) This is just a restatement of Theorem 8.2.

(ii) Let $\Gamma(n)$ be the statement given in (ii) with the extra condition $\epsilon_n \leq 1$. $\Gamma(0)$ is true with $\epsilon_0 = 1$ by inspection.

Suppose that Γ_n is true, that $P(t) = \sum_{j=0}^{n+1} a_j t^j$ is a polynomial of degree at most $n+1$, and that $|a_k| \geq 1$ for some $n+1 \geq k \geq 0$. If $|a_{n+1}| \leq \epsilon_n/2$, then

$$P(t) = a_{n+1} t^{n+1} + Q(t)$$

where $Q(t) = \sum_{j=0}^{n} a_j t^j$ is a polynomial of degree at most $n+1$ and $|a_k| \geq 1$ for some $n \geq k \geq 0$. Thus

$$\|P\|_\infty \geq \|Q\|_\infty - |a_{n+1}| \geq \epsilon_n/2.$$

On the other hand, if $|a_{n+1}| \geq \epsilon_n/2$, then part (i) tells us that

$$\|P\|_\infty \geq 2^{-n+1}\epsilon_n/2 = 2^{-n}\epsilon_n.$$

Thus, whatever the value of a_{n+1},

$$\|P\|_\infty \geq 2^{-n-1}\epsilon_n$$

and $\Gamma(n+1)$ holds with $\epsilon_{n+1} = 2^{-n-1}\epsilon_n$.

The required result holds by induction. □

We can now use a compactness argument to prove that there does exist a best approximation.

Theorem 8.4. *If $f : [a, b] \to \mathbb{R}$ is a continuous function, then there exists a polynomial P of degree at most n such that $\|P - f\|_\infty \leq \|Q - f\|_\infty$ for all polynomials Q of degree n or less.*

Proof. By rescaling and translation, we may suppose that $[a, b] = [-1, 1]$. Consider the map $F : \mathbb{R}^{n+1} \to \mathbb{R}$ given by $F(\mathbf{a}) = \|f - Q\|_\infty$ where

$$Q(t) = \sum_{j=0}^{n} a_j t^j.$$

Recalling the inequality $\big||d(f, g)| - |d(f, h)|\big| \leq d(g, h)$, we have

$$|F(\mathbf{a}) - F(\mathbf{b})| \leq \sup_{t \in [-1,1]} \left| \sum_{j=0}^{n} a_j t^j - \sum_{j=0}^{n} b_j t^j \right| \leq \sum_{j=0}^{n} |a_j - b_j| \leq (n + 1)\|\mathbf{a} - \mathbf{b}\|,$$

so F is continuous. Also

$$F(\mathbf{a}) \geq \sup_{t \in [-1,1]} \left| \sum_{j=0}^{n} a_j t^j \right| - \|f\|_\infty$$

so, by Corollary 8.3 (ii), we can find a $K > 0$ such that

$$\mathbf{a} \notin [-K, K]^{n+1} \Rightarrow F(\mathbf{a}) \geq F(\mathbf{0}).$$

By compactness, F attains a minimum at some point $\mathbf{p} \in [-K, K]^{n+1}$ and

$$P(t) = \sum_{j=0}^{n} p_j t^j$$

is the required polynomial. □

We could also have proved Corollary 8.3 (ii) directly by a compactness argument without using Chebchev's result.

We have only shown that the Chebychev criterion is a sufficient condition. However, it can be shown that it is also a necessary one. The proof is given in Exercise 20.10 but is not part of the course.

9 Gaussian quadrature

How should we attempt to estimate $\int_a^b f(x)\, dx$ if we only know f at certain points. One, rather naive, approach is to find the interpolating polynomial for those points and integrate that. This leads rapidly, via Lemma 7.2, to the following result.

Lemma 9.1. *Let x_0, x_1, ..., x_n be distinct points of $[a,b]$. Then there are unique real numbers A_0, A_1, ..., A_n with the property that*

$$\int_a^b P(x)\,dx = \sum_{j=0}^{n} A_j P(x_j)$$

for all polynomials of degree n or less.

Proof. As in Lemma 7.2, we take

$$e_k(x) = \prod_{j\neq k} \frac{x - x_j}{x_k - x_j}.$$

If

$$\int_a^b P(x)\,dx = \sum_{j=0}^{n} A_j P(x_j)$$

for all polynomials of degree n or less, then, setting $P = e_k$, gives us

$$A_k = \int_a^b e_k(x)\,dx,$$

proving uniqueness.

On the other hand, if P has degree n or less,

$$Q = P - \sum_{j=0}^{n} P(x_j)e_j$$

has degree n or less but vanishes at the $n+1$ points x_j. Thus $Q = 0$ and

$$P = \sum_{j=0}^{n} P(x_j)e_j,$$

whence

$$\int_a^b P(x)\,dx = \sum_{j=0}^{n} A_j P(x_j)$$

with

$$A_j = \int_a^b e_j(x)\,dx.$$

□

However our previous remarks about interpolating polynomials suggest, and experience confirms, that it may not always be wise to use the approximation

$$\int_a^b f(x)\,dx \approx \sum_{j=0}^{n} A_j f(x_j)$$

even when f well behaved. In the particular case when the interpolation points are equally spaced, computation suggests that as the number of points used increases the A_j begin to vary in sign and become large in absolute value. It can be shown that this is actually the case and that this means that the approximation can actually get worse as the number of points increase.

It is rather surprising that there is a choice of points which avoids this problem. Earlier (in Exercise 7.4 (ii)) we observed that the definition

$$\langle f, g\rangle = \int_{-1}^{1} f(t)g(t)\,dt$$

gives an inner product on the vector space $C([-1,1])$. Let us recall some results from vector space theory.

Lemma 9.2. [Gramm–Schmidt] *Let V be a vector space with an inner product. Suppose that $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_n$ are orthonormal and $\mathbf{f}$ is not in their linear span. If we set*

$$\mathbf{v} = \mathbf{f} - \sum_{j=1}^{n} \langle \mathbf{f}, \mathbf{e}_j\rangle \mathbf{e}_j$$

we know that $\mathbf{v} \neq \mathbf{0}$ and that, setting $\mathbf{e}_{n+1} = \|\mathbf{v}\|^{-1}\mathbf{v}$ the vectors $\mathbf{e}_1$, $\mathbf{e}_2$, …, $\mathbf{e}_{n+1}$ are orthonormal.

Lemma 9.2. Linear independence shows that $\mathbf{v} \neq \mathbf{0}$. We have $\|\mathbf{e}_{n+1}\| = \|\mathbf{v}\|^{-1}\|\mathbf{v}\| = 1$. Now

$$\begin{aligned}\langle \mathbf{v}, \mathbf{e}_k\rangle &= \left\langle \mathbf{f} - \sum_{j=1}^{n} \langle \mathbf{f}, \mathbf{e}_j\rangle \mathbf{e}_j, \mathbf{e}_k \right\rangle\\ &= \langle \mathbf{f}, \mathbf{e}_k\rangle - \sum_{j=1}^{n} \langle \mathbf{f}, \mathbf{e}_j\rangle\langle \mathbf{e}_k, \mathbf{e}_j\rangle\\ &= \langle \mathbf{f}, \mathbf{e}_k\rangle - \langle \mathbf{f}, \mathbf{e}_k\rangle = 0\end{aligned}$$

so $\langle \mathbf{e}_{n+1}, \mathbf{e}_k\rangle = 0$ for all $1 \le k \le n$. □

Lemma 9.2 enables us to make the following definition.

Definition 9.3. *The Legendre polynomials p_n are the the polynomials given by the following conditions*[8].

(i) p_n is a polynomial of degree n with positive leading coefficient.

(ii) $$\int_{-1}^{1} p_n(t)p_m(t)\,dt = \delta_{nm} = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases}$$

Lemma 9.4. *The nth Legendre polynomial p_n has n distinct roots all lying in $(-1, 1)$.*

Proof. Suppose that p_n has k roots α_j of *odd* order (that is to say the polynomial changes sign at the root) on $(-1, 1)$. If we set $Q(t) = \prod_{j=1}^{k}(t - \alpha_j)$, then $p_n(t)Q(t)$ is a continuous single signed not everywhere zero function so

$$\int_{-1}^{1} Q(t)p_n(t)\,dt \neq 0.$$

Thus Q has degree at least n, so $k \geq n$.

It follows that $k = n$ and all of the roots of p_n are simple lying in $(-1, 1)$. □

Gauss had the happy idea of choosing the evaluation points to be the roots of a Legendre polynomial.

Theorem 9.5. [Gaussian quadrature] *(i) If $\alpha_1, \alpha_2, \ldots \alpha_n$ are the n roots of the nth Legendre polynomial p_n and the A_j are chosen so that*

$$\int_{-1}^{1} P(x)\,dx = \sum_{j=1}^{n} A_j P(\alpha_j)$$

for every polynomial P of degree $n - 1$ or less, then, in fact

$$\int_{-1}^{1} Q(x)\,dx = \sum_{j=1}^{n} A_j Q(\alpha_j)$$

for every polynomial Q of degree $2n - 1$ or less.

(ii) If $\beta_j \in [-1, 1]$ and B_j are such that

$$\int_{-1}^{1} Q(x)\,dx = \sum_{j=1}^{n} B_j Q(\beta_j)$$

for every polynomial Q of degree $2n - 1$ or less, then the β_j are the roots of the nth Legendre polynomial.

[8] There are various other definitions but they all give polynomials of the form $b_n p_n$. The only difference is in the choice of b_n. As may be seen from Exercise 19.14 our choice is not very convenient.

Proof. Suppose that p_n has k roots α_j of *odd* order on $(-1,1)$. If we set $Q(t) = \prod_{j=1}^{k}(t-\alpha_j)$, then $p_n(t)Q(t)$ is a continuous single signed not everywhere zero function so

$$\int_{-1}^{1} Q(t)p_n(t)\,dt \neq 0.$$

Thus Q has degree at least n, so $k \geq n$.

It follows that $k = n$ and all of the roots of p_n are simple lying in $(-1,1)$. □

Theorem 9.5 looks impressive but it is the next result which really shows how good Gauss's idea is.

Theorem 9.6. *We continue with the notation of Theorem 9.5.*

(i) $A_j \geq 0$ *for each* $1 \leq j \leq n$.

(ii) $\sum_{j=1}^{n} A_j = 2$.

(iii) If $f : [-1,1] \to \mathbb{R}$ *is continuous and* P *is any polynomial of degree at most* $2n-1$*, then*

$$\left| \int_{-1}^{1} f(x)\,dx - \sum_{j=1}^{n} A_j f(\alpha_j) \right| \leq 4\|f-P\|_\infty.$$

(iv) Write $G_n f$ *for the estimate of* $\int_{-1}^{1} f(t)\,dt$ *obtained using Gauss's idea with the* n*th Legendre polynomial. Then, if* f *is continuous on* $[-1,1]$,

$$G_n f \to \int_{-1}^{1} f(t)\,dt$$

as $n \to \infty$.

Proof. (i) Let

$$P_k(x) = \left(\prod_{j \neq k} \frac{x - x_j}{x_k - x_j} \right)^2.$$

Then P_k has degree $2n-2$, so

$$0 < \int_{-1}^{1} P_k(x)\,dx = \sum_{j=1}^{n} A_j P_k(\alpha_j) = A_k.$$

(ii) Taking $P = 1$ in the formula, we obtain

$$2 = \int_{-1}^{1} 1\,dx = \sum_{j=1}^{n} A_j.$$

(iii) We have

$$\begin{aligned}
\left|\int_{-1}^{1} f(x)\,dx - \sum_{j=1}^{n} A_j f(\alpha_j)\right| \\
&= \left|\int_{-1}^{1} \big(f(x)-P(x)\big)\,dx - \sum_{j=1}^{n} A_j\big(f(\alpha_j)-P(\alpha_j)\big)\right| \\
&\leq \left|\int_{-1}^{1} \big(f(x)-P(x)\big)\,dx\right| + \left|\sum_{j=1}^{n} A_j\big(f(\alpha_j)-P(\alpha_j)\big)\right| \\
&\leq \int_{-1}^{1} |f(x)-P(x)|\,dx + \sum_{j=1}^{n} A_j|f(\alpha_j)-P(\alpha_j)| \\
&\leq 2\|f-P\|_\infty + \sum_{j=1}^{n} A_j\|f-P\|_\infty \leq 4\|f-P\|_\infty.
\end{aligned}$$

(iv) Let $\epsilon > 0$. By Weierstrass's theorem, we can find a polynomial P such that $\|f-P\|_\infty \leq \epsilon/4$. Then, if n is greater than the degree of P, part (iii) tells us that

$$\left|\int_{-1}^{1} f(x)\,dx - G_n f\right| \leq 4\|f-P\|_\infty \leq \epsilon.$$

□

10 Distance and compact sets

This section could come almost anywhere in the notes, but provides some helpful background to the section on Runge's theorem. We start by strengthening Lemma 2.5.

Lemma 10.1. *If E is a non-empty compact set in $\mathbb{R}^m$ and $\mathbf{a} \in \mathbb{R}^m$, then there is a point $\mathbf{e} \in E$ such that*

$$\|\mathbf{a}-\mathbf{e}\| = \inf_{\mathbf{x}\in E} \|\mathbf{a}-\mathbf{x}\|.$$

Proof. It is a standard observation about metric spaces (X, d) that, since $d(x, y) + d(y, z) \geq d(x, z)$, we have $d(y, z) \geq d(x, z) - d(x, y)$ and similarly $d(y, z) = d(z, y) \geq d(x, y) - d(x, z)$, so that

$$d(y, z) \geq |d(x, z) - d(x, y)|.$$

Thus, if we write $f(\mathbf{x}) = \|\mathbf{a} - \mathbf{x}\|$, we have

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq \|\mathbf{x} - \mathbf{y}\|,$$

so f is continuous and attains its minimum on the compact set E. $\square$

As before we write $d(\mathbf{a}, E) = \inf_{\mathbf{x}\in E} \|\mathbf{a} - \mathbf{x}\|$.

Exercise 10.2. *(i) Give an example to show that the point* $\mathbf{e}$ *in Lemma 10.1 need not be unique.*

(ii) Show that, if E is convex, $\mathbf{e}$ *is unique.*

Solution. (i) Consider $n = 2$,

$$E = \{(x, y) \,:\, x^2 + y^2 = 1\} \text{ and } \mathbf{e} = \mathbf{0}.$$

Any point of E will do.

(ii) Suppose that E is convex, $\mathbf{e},\ \mathbf{f} \in E$ and $\|\mathbf{a} - \mathbf{e}\| = \|\mathbf{a} - \mathbf{f}\|$. Then

$$\frac{\mathbf{e} + \mathbf{f}}{2} \in E$$

but the parallelogram law tells us that

$$\begin{aligned} 4\|\mathbf{a} - \mathbf{e}\|^2 &= 2\|\mathbf{a} - \mathbf{e}\|^2 + \|\mathbf{a} - \mathbf{f}\|^2 \\ &= \|(\mathbf{a} - \mathbf{e}) + (\mathbf{a} - \mathbf{f})\|^2 + \|(\mathbf{a} - \mathbf{e}) - (\mathbf{a} - \mathbf{f})\|^2 \\ &= 4\left\|\mathbf{a} - \frac{\mathbf{e} + \mathbf{f}}{2}\right\|^2 + \|\mathbf{e} - \mathbf{f}\|^2 \end{aligned}$$

and so

$$\left\|\mathbf{a} - \frac{\mathbf{e} + \mathbf{f}}{2}\right\| \leq \|\mathbf{a} - \mathbf{e}|$$

with equality only if $\mathbf{e} = \mathbf{f}$. (Alternatively draw a diagram and use a little school geometry to obtain the same result.) $\square$

Lemma 10.3. *(i) If E and F are non-empty compact sets in $\mathbb{R}^m$, then there exist points* $\mathbf{e} \in E$ *and* $\mathbf{f} \in F$ *such that*

$$\|\mathbf{e} - \mathbf{f}\| = \inf_{\mathbf{y}\in E} d(\mathbf{y}, F).$$

(ii) The result in (i) remains true when F is compact and non-empty and E is closed and non-empty.

(iii) The result in (i) may fail when E and F are closed and non-empty.

Proof. (i) Recall that, if $\mathbf{u} \in \mathbb{R}^n$, then we can find $\mathbf{v} \in F$ such that $\|\mathbf{u}-\mathbf{v}\| = d(\mathbf{u}, F)$. If $\mathbf{u}' \in \mathbb{R}^n$, then

$$d(\mathbf{u}', F) \leq \|\mathbf{u}' - \mathbf{v}\| \leq \|\mathbf{u}' - \mathbf{u}\| + \|\mathbf{u} - \mathbf{v}\| = \|\mathbf{u}' - \mathbf{u}\| + d(\mathbf{u}, F).$$

The same argument shows that $d(\mathbf{u}, F) \leq \|\mathbf{u}' - \mathbf{u}\| + d(\mathbf{u}', F)$. Thus

$$|d(\mathbf{u}, F) - d(\mathbf{u}', F)| \leq \|\mathbf{u}' - \mathbf{u}\|.$$

and the map $\mathbf{u} \mapsto d(\mathbf{u}, F)$ is continuous. By compactness, it attains its minimum on E and this is the required result.

(ii) Chose $\mathbf{u} \in E$. Since F is bounded we can find an R such that $B(\boldsymbol{u}, R) \supseteq F$. Let

$$E^* = \bar{B}(\mathbf{u}, 2R + 1) \cap E.$$

If $\mathbf{e} \in E \setminus E^*$, then $d(\mathbf{e}, F) \geq d(\mathbf{u}, F) + 1$.

Since E^* is compact, part (i) tells us that there exist a $\mathbf{e} \in E^*$ and a $\mathbf{f} \in F$ such that

$$\|\mathbf{e} - \mathbf{f}\| = \inf_{\mathbf{y} \in E^*} d(\mathbf{y}, F)$$

and so by the previous paragraph

$$\|\mathbf{e} - \mathbf{f}\| = \inf_{\mathbf{y} \in E} d(\mathbf{y}, F)$$

(iii) Let $n = 1$, $E = \{r + 1/r \,:\, r \in \mathbb{Z},\ r \geq 2\}$ and $F = \{r \,:\, r \in \mathbb{Z},\ r \geq 2\}$. We have E and F closed and $\tau(E, F) = 0$, but $|e - f| > 0$ for all $e \in E$, $f \in F$. □

Let us write $\tau(E, F) = \inf_{\mathbf{y} \in E} d(\mathbf{y}, F)$.

Exercise 10.4. *Give an example to show that the points* $\mathbf{e}$ *and* $\mathbf{f}$ *in Lemma 10.3 need not be unique.*

Solution. Repeat the counter-example of Exercise 10.2. Take $n = 2$,

$$E = \{(x, y) \,:\, x^2 + y^2 = 1\},\ F = \{\mathbf{0}\}.$$

□

The statement of the next exercise requires us to recall the definition of a metric.

Definition 1.1. *Suppose that X is a non-empty set and $d : X^2 \to \mathbb{R}$ is a function such that*

(i) $d(x, y) \geq 0$ for all x, $y \in X$.

(ii) $d(x, y) = 0$ if and only if $x = y$.

(iii) $d(x, y) = d(y, x)$ for all x, $y \in X$.

(iv) $d(x, y) + d(y, z) \geq d(x, z)$ for all x, y, $z \in X$.

Then we say that d is a metric on X and that (X, d) is a metric space.

Exercise 10.5. *Show that, if we consider the space $\mathcal{K}$ of non-empty compact sets in $\mathbb{R}^m$, then τ obeys conditions (i) and (iii) for a metric but not conditions (ii) and (iv).*

Show that, if $E, F \in \mathcal{K}$, then $\tau(E, F) = 0$ if and only if $E \cap F \neq \varnothing$.

Solution. (i) follows directly from the definition. (Alternatively take the $\mathbf{e}$ and $\mathbf{f}$ of Lemma 10.3 (i) and observe that

$$\tau(E, f) = \|\mathbf{e} - \mathbf{f}\| \geq 0.)$$

Lemma 10.3 (i) also shows that

$$\tau(F, E) \leq \|\mathbf{e} - \mathbf{f}\| = \tau(E, F).$$

Interchanging E and F, yields $\tau(E, F) \leq \tau(F, E)$ so $\tau(E, F) = \tau(F, E)$.

If we work with $n = 1$, setting $E = \{0\}$, $F = \{0, 1\}$, gives $\tau(E, F) = 0$, but $E \neq F$.

If we work with $n = 1$, then setting $E = \{0\}$, $F = \{0, 1\}$, $G = \{1\}$ gives $\tau(E, F) + \tau(F, G) = 0 + 0 = 0$, but $\tau(E, g) = 1$. □

Since τ does not provide a satisfactory metric on $\mathcal{K}$, we try some thing else. If E and F are compact sets in $\mathbb{R}^m$, let us set $\sigma(E, F) = \sup_{\mathbf{y} \in E} d(\mathbf{y}, F)$.

Exercise 10.6. *Suppose that E and F are non-empty compact sets. Show that there exists an $\mathbf{e} \in E$ such that $d(\mathbf{e}, F) = \sigma(E, F)$.*

Solution. In our proof of Lemma 10.3 (i) we showed that $\mathbf{u} \mapsto d(\mathbf{u}, F)$ is continuous. It follows that it attains its maximum on the compact set E. □

Exercise 10.7. *Show that, if we consider the space $\mathcal{K}$ of non-empty compact sets in $\mathbb{R}^m$, then σ obeys condition (i) for a metric but not conditions (ii) and (iii).*

Show that $\sigma(E, F) = 0$ if and only if $E \subseteq F$.

Solution. Since $d(\mathbf{e}, F) \geq 0$ for all $\mathbf{e}$, we have $\sigma(E, F) \geq 0$.

If $n = 1$, $E = \{0\}$, $F = [0, 1]$, then $\sigma(E, F) = 0$, but $\sigma(F, E) = 1$, so conditions (ii) and (iii) fail.

$$\sigma(E, F) = 0 \Leftrightarrow d(\mathbf{e}, F) = 0 \ \forall \mathbf{e} \in E \Leftrightarrow \mathbf{e} \in F \ \forall \mathbf{e} \in E \Leftrightarrow E \subseteq F.$$

□

However, σ does obey the triangle inequality.

Lemma 10.8. *If E, F and G are non-empty compact sets then*

$$\sigma(E, G) \leq \sigma(E, F) + \sigma(F, G).$$

Lemma 10.8. Given $\mathbf{e} \in E$, we can find $\mathbf{f} \in F$ such that $\|\mathbf{e} - \mathbf{f}\| = d(\mathbf{e}, F)$. If $\mathbf{g} \in G$, then

$$\begin{aligned} d(\mathbf{e}, G) &\leq \|\mathbf{e} - \mathbf{g}\| \leq \|\mathbf{e} - \mathbf{f}\| + \|\mathbf{f} - \mathbf{g}\| \\ &= d(\mathbf{e}, F) + \|\mathbf{f} - \mathbf{g}.\| \end{aligned}$$

Since $\mathbf{g} \in G$ was arbitrary,

$$d(\mathbf{e}, G) \leq d(\mathbf{e}, F) + d(\mathbf{f}, G) \leq \sigma(E, F) + \sigma(F, G)$$

and so

$$\sigma(E, G) \leq \sigma(E, F) + \sigma(F, G).$$

□

This enables us to define the Hausdorff metric ρ.

Definition 10.9. *If E and F are non-empty compact subsets of $\mathbb{R}^m$, we set*

$$\rho(E, F) = \sigma(E, F) + \sigma(F, E),$$

that is to say,

$$\rho(E, F) = \sup_{\mathbf{e} \in E} \inf_{\mathbf{f} \in F} \|\mathbf{e} - \mathbf{f}\| + \sup_{\mathbf{f} \in F} \inf_{\mathbf{e} \in E} \|\mathbf{e} - \mathbf{f}\|.$$

Theorem 10.10. *The Hausdorff metric ρ is indeed a metric on the space $\mathcal{K}$ of non-empty compact subsets of $\mathbb{R}^m$.*

Proof. Observe that

$$\begin{aligned}
\rho(E,F) &= \sigma(E,F)+\sigma(F,E)\geq 0\\
\rho(E,F)=0 &\Leftrightarrow \sigma(E,F)=\sigma(F,E)=0\Leftrightarrow E\subseteq F,\ F\subseteq E\Leftrightarrow E=F\\
\rho(E,F) &= \sigma(E,F)+\sigma(F,E)=\sigma(F,E)+\sigma(E,F)=\rho(F,E)\\
\rho(E,F)+\rho(F,G) &= \sigma(E,F)+\sigma(F,G)+\sigma(G,F)+\sigma(F,E)\\
&\geq \sigma(E,G)+\sigma(G,E)=\rho(E,G),
\end{aligned}$$

as desired. □

Indeed, we can say something even stronger which will come in useful when we give examples of the use of Baire's theorem in Section 13

Theorem 10.11. *The Hausdorff metric ρ is a complete metric on the space $\mathcal{K}$ of non-empty compact subsets of $\mathbb{R}^m$.*

Our proof of Theorem 10.11 makes use of two observations.

Theorem 10.12. *(i) Suppose that we have a sequence of non-empty compact sets in $\mathbb{R}^m$ such that*

$$K_1\supseteq K_2\supseteq K_3\supseteq\dots.$$

Then $K=\bigcap_{p=1}^{\infty}K_p$ is a non-empty compact set.

(ii) Further, $K_p \underset{\rho}{\to} K$ as $p\to\infty$.

Proof. (i) This part may be familiar from 1B. (Indeed the reader may well be able to supply a more sophisticated proof.) Since the intersection of closed sets is closed and the intersection of bounded sets is bounded we only have to show that K is non-empty.

Choose $\mathbf{x}_n\in K_n$. Since K_1 is compact and $\mathbf{x}_n\in K_1$ for every n we can find an $\mathbf{x}\in K_1$ and $n(j)\geq j$ such that $\mathbf{x}_{n(j)}\to\mathbf{x}$ (in the Euclidean metric) as $j\to\infty$.

Automatically,

$$\mathbf{x}_{n(j)}\in K_{n(j)}\subseteq K_j\subseteq K_p$$

for all $j\geq p$, so, since K_p is closed, $\mathbf{x}\in K_p$ for all $p\geq 1$. It follows that $\mathbf{x}\in K$ and K is non-empty.

(ii) Since $K\subseteq K_p$ it follows that

$$\rho(K,K_p)=\sup_{\mathbf{e}\in K_p}\inf_{\mathbf{k}\in K}\|\mathbf{e}-\mathbf{k}\|$$

and, in particular that $\rho(K,K_p)$ is a decreasing positive sequence.

Thus if $K_p \underset{\rho}{\nrightarrow} K$ there must exist an $\eta > 0$ with

$$\rho(K, K_p) \geq 2\eta$$

and there must exist $\mathbf{k}_p \in K_p$ with

$$\|\mathbf{k}_p - \mathbf{k}\| \geq \eta$$

for all $\mathbf{k} \in K$.

Since K_1 is compact and $\mathbf{k}_p \in K_1$ for every p we can find an $\mathbf{x} \in K_1$ and $p(j) \geq j$ such that $\mathbf{k}_{p(j)} \to \mathbf{x}$ (in the Euclidean metric) as $j \to \infty$. As in part (i), we know that $\mathbf{x} \in K$ so

$$\|\mathbf{k}_{p(j)} - \mathbf{x}\| \geq \eta.$$

for all j giving us a contradiction.

Part (ii) follows by reductio ad absurdum. □

Lemma 10.13. *If K is compact in $\mathbb{R}^m$ so is*

$$K + \bar{B}(0, r) = \{\mathbf{x} + \mathbf{y} \,:\, \mathbf{x} \in K,\ \|\mathbf{y}\| \leq r\}.$$

Proof. If $\mathbf{z}_n \in K + \bar{B}(0, r)$, then $\mathbf{z}_n = \mathbf{x}_n + \mathbf{y}_n$ with $\mathbf{x}_n \in K$, $\|\mathbf{y}_n\| \leq r$. By compactness, we can first extract a convergent subsequence $\mathbf{x}_{n(j)} \in K$ and then a convergent subsequence $\mathbf{y}_{n(j(k))} \in \bar{B}(0, r)$. It follows that $\mathbf{z}_{n(j(k))} = \mathbf{x}_{n(j(k))} + \mathbf{y}_{n(j(k))}$ converges to a point in $K + \bar{B}(0, r)$ so we are done. □

Proof of Theorem 10.11. By Lemma 1.11, it suffices to show that, if we have sequence of non-empty compact sets with $\rho(E_n, E_{n+1}) < 8^{-n}$ for $n \geq 1$, then the sequence converges. Set

$$K_n = E_n + \bar{B}(0, 6 \times 8^{-n}).$$

Then K_n is compact and $\rho(E_n, K_n) = 6 \times 8^{-n}$ so it is sufficient to show that K_n converges.

To do this, we observe that $K_{n+1} \subseteq K_n$ and so we may apply Theorem 10.12. □

11 Runge's theorem

The existence of two different introductory courses in complex variable is one of many mad things in the Cambridge system. The contents of this section should be accessible to anyone who has gone to *either*. As I shall emphasise from time to time, the reader will need to know some of the results from those courses but will not be required to prove them.

Weierstrass's theorem tells us that every continuous real valued function on $[a, b]$ can be uniformly approximated by polynomials. Does a similar theorem hold for complex variable?

Cauchy's theorem enables us to answer with a resounding no.

Example 11.1. *Let* $\bar{D} = \{z : |z| \leq 1\}$ *and define* $f : \bar{D} \to \mathbb{C}$ *by*

$$f(z) = \bar{z}.$$

If P *is any polynomial, then*

$$\sup_{z \in \bar{D}} |f(z) - p(z)| \geq 1.$$

Proof. Observe that, taking C to be the contour $z = e^{i\theta}$ as θ runs from 0 to 2π, we have

$$\begin{aligned}
\sup_{z \in \bar{D}} |f(z) - p(z)| &\geq \frac{1}{2\pi} \left| \int_C f(z) - p(z)\, dz \right| \\
&= \frac{1}{2\pi} \left| \int_C f(z)\, dz \right| = \frac{1}{2\pi} \left| \int_0^{2\pi} e^{-i\theta} i e^{i\theta}\, d\theta \right| \\
&= \frac{1}{2\pi} \left| \int_0^{2\pi} i\, d\theta \right| = 1.
\end{aligned}$$

□

After looking at this example the reader may recall the following theorem (whose proof does not form part of this course).

Theorem 11.2. *If* Ω *is an open subset of* $\mathbb{C}$ *and* $f_n : \Omega \to \mathbb{C}$ *is analytic, then if* $f_n \to f$ *uniformly on* Ω *(or, more generally, if* $f_n \to f$ *uniformly on each compact subset of* Ω*) then* f *is analytic.*

We might now conjecture that every analytic function on a well behaved set can be uniformly approximated by polynomials. Cauchy's theorem again shows that the matter is not straightforward.

Example 11.3. *Let* $T = \{z : 1/2 \leq |z| \leq 2\}$ *and define* $f : T \to \mathbb{C}$ *by*

$$f(z) = \frac{1}{z}.$$

If P *is any polynomial then*

$$\sup_{z \in \bar{T}} |f(z) - p(z)| \geq 1.$$

Proof. By exactly the same computations as in Example 11.1,

$$\sup_{z\in T}|f(z)-p(z)| \geq \frac{1}{2\pi}\left|\int_C f(z)-p(z)\,dz\right| = 1.$$

□

Thus the best we can hope for is a theorem that tells us that every analytic function on a suitable set 'without holes' can be uniformly approximated by polynomials.

We shall see that the following definition gives a suitable 'no holes' condition.

Definition 11.4. *An open set $U \subseteq \mathbb{C}$ is* path connected *if, given z_0, $z_1 \in U$, we can find a continuous map $\gamma : [0,1] \to U$ with $\gamma(0) = z_0$ and $\gamma(1) = z_1$.*

We will obtain results for a bounded sets whose *complement* is path connected.

The reader may ask why we could not simply use Taylor's theorem in complex variables. To see that this would not work we recall various earlier results. (As I said earlier the proofs are not part of the course, but you are expected to know the results.)

Lemma 11.5. *If $a_j \in \mathbb{C}$, then there exists an $R \in [0,\infty]$ (with suitable conventions when $R = \infty$) such that $\sum_{j=0}^\infty a_j z^j$ converges for $|z| < R$ and diverges for $|z| > R$.*

We call R the radius of convergence of $\sum_{j=0}^\infty a_j z^j$.

Lemma 11.6. *If $\sum_{j=0}^\infty a_j z^j$ has radius of convergence R and $R' < R$ then $\sum_{j=0}^\infty a_j z^j$ converges uniformly for $|z| \leq R'$.*

Lemma 11.7. *Suppose that $\sum_{j=0}^\infty a_j z^j$ has radius of convergence R and that $\sum_{j=0}^\infty b_j z^j$ has radius of convergence R'. If there exists an R'' with $0 < R'' \leq R, R'$ such that*

$$\sum_{j=0}^\infty a_j z^j = \sum_{j=0}^\infty b_j z^j$$

for all $|z| < R''$, then $a_j = b_j$ for all j.

By a careful use of Taylor's theorem Lemma 11.7 can be used to give the following extension. (The proof is not part of the course but, again, you are expected to know the result.)

Lemma 11.8. *Suppose that $f, g : B(w, r) \to \mathbb{C}$ are analytic. If there is a non-empty open subset U of $B(w, r)$ such that $f(z) = g(z)$ for all $z \in U$, it follows that $g = f$.*

Exercise 11.9. *(i) If $w \neq 0$,w show that we can find a power series $\sum_{j=0}^{\infty} a_j(z-w)^j$ with radius of convergence $|w|$ such that*

$$z^{-1} = \sum_{j=0}^{\infty} a_j(z-w)^j$$

for all $|z - w| < |w|$.

(ii) Let

$$\Omega = \{z : 10^{-2} < |z| < 1\} \setminus \{x : x \in \mathbb{R},\ x \leq 0\}$$

Show that Ω is open, path connected and bounded and $f(z) = 1/z$ defines a bounded analytic function on Ω, but we can not find z_0 and b_j such that

$$z^{-1} = \sum_{j=0}^{\infty} b_j(z-z_0)^j$$

for all $z \in \Omega$.

Proof. (i) Just observe that

$$\frac{1}{z} = \frac{1}{w + (z-w)} = \frac{1}{w(1 + (z-w)/w)} \sum_{j=0}^{\infty} \frac{(z-w)^j}{w^{j+1}}$$

for $|(z - w)/w| < 1$.

(ii) It is easy to check that Ω is open and bounded. To see that Ω is connected, suppose that w_1, $w_2 \in \Omega$. Then we can find r_k and θ_k with $10^{-2} < r_k < 1$ and $-\pi < \theta < \pi$ such that $w_k = r_k e^{i\theta_k}$ $[k = 1, 2]$. If we define $\gamma : [0, 1] \to \Omega$ by

$$\gamma(t) = \big((1-t)r_1 + tr_2\big) \exp\big(i(1-t)\theta_1 + it\theta_2\big),$$

then γ gives a path from w_1 to w_2.

Suppose, if possible, that

$$z^{-1} = \sum_{j=0}^{\infty} b_j(z-z_0)^j$$

for all $z \in \Omega$. Then the power series converges on some open disc D centre z_0 with $D \supseteq \Omega$. Thus $D \supseteq \{z : |z| < 1\}$. By Lemma 11.8,

$$z^{-1} = \sum_{j=0}^{\infty} b_j (z - z_0)^j$$

for all z with $0 < |z| < 1$. Allowing $z \to 0$, gives a contradiction. □

Let us see what Taylor's theorem actually says.

Theorem 11.10. [Taylor's Theorem] *Suppose that Ω is an open set in $\mathbb{C}$ and $f : \Omega \to \mathbb{C}$ is analytic. If the open disc*

$$B(z_0, \delta) = \{z : |z - z_0| < \delta\}$$

lies in Ω, then we can find $a_j \in \mathbb{C}$ such that $\sum_{j=0}^{\infty} a_j z^j$ has radius of convergence at least δ and

$$\sum_{j=0}^{\infty} a_j (z - z_0)^j = f(z)$$

for all $z \in B(z_0, \delta)$.

Thus Taylor's theorem for analytic functions says (among other things) that an analytic function can be *locally* approximated uniformly by polynomials. Runge's theorem asserts that (under certain conditions) an analytic function can be *globally* approximated uniformly by polynomials.

Theorem 11.11. [Runge's theorem] *Suppose that Ω is an open set and $f : \Omega \to \mathbb{C}$ is analytic. Suppose that K is a compact set with $K \subseteq \Omega$ and $\mathbb{C} \setminus K$ path connected. Then given any $\epsilon > 0$, we can find a polynomial P with*

$$\sup_{z \in K} |f(z) - P(z)| < \epsilon.$$

I shall make a number of remarks before moving on to the proof. The first is that (as might be expected) Theorem 11.11 is the simplest of a family of results which go by the name of Runge's theorem. However, I think that it is fair to say that, once the proof of this simplest case is understood, both the proofs and the meanings of the more general theorems are not hard to grasp.

The second remark is that the reader will lose very little understanding[9] if she concentrates on the example of Runge's theorem for geometrically simple K and Ω (like rectangles and triangles).

Our proof of Runge's theorem splits into several steps.

[9]And, provided she does not twist the examiner's nose, few marks in the exam.

Lemma 11.12. *Suppose that K is a compact set with $K \subseteq \Omega$. Then we can find a finite set of piece-wise linear contours C_m lying entirely within $\Omega \setminus K$ such that*

$$f(z) = \sum_{m=1}^{M} \int_{C_m} \frac{f(w)}{w-z}\, dw$$

whenever $z \in K$ and $f : \Omega \to \mathbb{C}$ is analytic.

Proof. We may suppose K non-empty. Since K is compact, $\mathbb{C} \setminus \Omega$ closed and the two sets are disjoint, it follows that $\eta = \tau(K, \mathbb{C} \setminus \Omega)/8 > 0$ (i.e. $|k - w| > 8\eta$ for all $k \in K$, $w \notin \Omega$).

Consider a grid of squares side η. We consider the collection Γ of closed squares S lying entirely within Ω with boundary contours $C(S)$. Observe that

$$\bigcup_{S\in\Gamma} C(S) \supseteq \{k + u \,:\, k \in K,\ |u| \leq 2\eta\}$$

By Cauchy's theorem

$$f(z) = \frac{1}{2\pi i} \sum_{S\in\Gamma} \int_{C(S)} \frac{f(w)}{w-z}\, dw$$

for all $z \in K$ such that z does not lie on the boundary of some S. By cancelling internal sides,

$$f(z) = \sum_{m=1}^{M} \int_{C_m} \frac{f(w)}{w-z}\, dw \qquad \bigstar$$

with the piece-wise linear contours C_m $[1 \leq m \leq M]$ lying entirely within $\Omega \setminus K$.

We deal with the case when z lies on the boundary of some S by erasing any sides through z and repeating the argument with the new (non-regular) grid. (Alternatively, we could observe that both sides of equation $\bigstar$ are continuous.) □

It is worth making the following observation explicit.

Lemma 11.13. *With the notation and conditions of Lemma 11.12, we can find a $\delta > 0$ such that $|z - w| \geq \delta$ whenever $z \in K$ and w is a point of one of the contours C_m.*

Lemma 11.13. Observe that K and $\bigcup_{m=1}^{M} C_m$ are compact and disjoint. □

We use Lemma 11.12 to prove the following result which takes us closer to our goal.

Lemma 11.14. *Suppose that K is a compact set with $K \subseteq \Omega$. Then given any analytic $f : \Omega \to \mathbb{C}$ and any $\epsilon > 0$ we can find an integer N, complex numbers $A_1, A_2, \ldots, A_N$ and $\alpha_1, \alpha_2, \ldots, \alpha_N \in \Omega \setminus K$ such that*

$$\left| f(z) - \sum_{n=1}^{N} \frac{A_n}{z - \alpha_n} \right| < \epsilon$$

for all $z \in K$.

Proof. Consider the system in Lemma 11.12. Write

$$g_z(w) = \frac{1}{z - w}.$$

By Lemma 11.13, we can find an M_1 such that $|g_k(w)| \leq M_1$ for all $k \in K$, $w \in \bigcup_{m=1}^{M} C_m$. Further

$$|g_k(w) - g_k(w')| = \frac{|w - w'|}{|z - w||z - w'|} \leq M_1^2 |w - w'|.$$

Since f is continuous on the compact set $\bigcup_{m=1}^{M} C_m$, we can find an M_2 such that $|f(w)| \leq M_2$ for all $w \in \bigcup_{m=1}^{M} C_m$ and, given ϵ, we can find a $\delta > 0$ such that $|f(w) - f(w')| \leq \epsilon$ whenever $|w - w'| \leq \delta$ and $w,\ w' \in \bigcup_{m=1}^{M} C_m$.

If $k \in K$, $w,\ w' \in \bigcup_{m=1}^{M} C_m$ and $|w - w'| \leq \min\{\delta, \epsilon\}$ then

$$\begin{aligned}
&|f(w)g_k(w) - f(w')g_k(w')| \\
&\qquad \leq |f(w)g_k(w) - f(w)g_k(w')| + |f(w)g_k(w') - f(w')g_k(w')| \\
&\qquad = |f(w)||g_k(w) - g_k(w')| + |f(w) - f(w')||g_k(w')| \\
&\qquad \leq M_2 M_1^2 \epsilon + M_1 \epsilon.
\end{aligned}$$

Let L be the total length of the contours C_m. Since ϵ is arbitrary, we have shown that, given $\epsilon > 0$, there exists a $\beta > 0$ such that

$$\left| \frac{f(w)}{w - z} - \frac{f(w')}{w' - z} \right| \leq \frac{\epsilon}{8L}$$

for all $w,\ w' \in \bigcup_{m=1}^{M} C_m$ with $|w - w'| \leq \beta$ and $z \in K$. Thus, if we pick $\alpha_{1,m},\ \alpha_{2,m},\ \ldots,\ \alpha_{N(m),m},\ \alpha_{N(m)+1,m} = \alpha_{1,m}$ as points in anticlockwise order round C_m with $|\alpha_{r,m} - \alpha_{r+1,m}| \leq \beta$ $[1 \leq r \leq N(m)]$, we will have

$$\sum_{m=1}^{M} \left| \sum_{r=1}^{N(m)} \frac{f(\alpha_{r,m})}{z - \alpha_{r,m}} - \int_{C_m} \frac{f(w)}{w - z}\, dw \right| < \epsilon$$

and so

$$\left|f(z) - \sum_{m=1}^{M} \sum_{r=1}^{N(m)} \frac{f(\alpha_{r,m})}{z - \alpha_{r,m}}\right| < \epsilon$$

for all $z \in K$. □

Thus Runge's theorems follows at once from the following special case.

Lemma 11.15. *Suppose that K is a compact set and $\mathbb{C} \setminus K$ path connected. Then, given any $\alpha \notin K$ and any $\epsilon > 0$, we can find a polynomial P with*

$$\left|P(z) - \frac{1}{z - \alpha}\right| < \epsilon$$

for all $z \in K$.

Proof of Theorem 11.11 from Lemma 11.15. We use the result and notation of Lemma 11.14. Choose polynomials P_n such that

$$\left|P_n(z) - \frac{1}{z - \alpha_n}\right| \leq \frac{\epsilon}{(N+1)(|A_n| + 1)}$$

for all $z \in K$. Then, if

$$P(z) = \sum_{n=1}^{N} A_n P_n(z),$$

P is a polynomial and

$$\begin{aligned} |f(z) - P(z)| &\leq \left|f(z) - \sum_{n=1}^{N} \frac{A_n}{z - \alpha_n}\right| + \sum_{n=1}^{N} |A_n| \left|P_n(z) - \frac{1}{z - \alpha_n}\right| \\ &\leq \epsilon + N\frac{\epsilon}{N+1} \leq 2\epsilon. \end{aligned}$$

Since ϵ was arbitrary the result follows. □

Let us make a temporary definition.

Definition 11.16. *Let K be a compact set in $\mathbb{C}$. We write $\Lambda(K)$ for the set of points $\alpha \notin K$ such that, given any $\epsilon > 0$, we can find a polynomial P with*

$$\left|P(z) - \frac{1}{z - \alpha}\right| < \epsilon$$

for all $z \in K$.

A series of observations about $\Lambda(K)$ brings the proof of Runge's theorem to a close.

Lemma 11.17. *Let K be a compact set in $\mathbb{C}$. Then there exists an R such that $|\alpha| > R$ implies $\alpha \in \Lambda(K)$.*

Proof. Since K is compact, it is bounded and we can find an $R > 0$ such that $|z| < R/2$ whenever $z \in K$. The standard geometric series result shows that if $|\alpha| > R$

$$\frac{-1}{\alpha}\sum_{r=0}^{n}\frac{z^r}{\alpha^r} \to \frac{-1}{\alpha} \times \frac{1}{1-(z/\alpha)} = \frac{1}{z-\alpha}$$

uniformly for $|z| \leq R/2$ and so for $z \in K$. □

Lemma 11.18. *Let K be a compact set in $\mathbb{C}$. If $\alpha \in \Lambda(K)$ and $|\alpha - \beta| < d(\alpha, K)$ then $\beta \in \Lambda(K)$.*

Proof. Since $\alpha \in \Lambda(K)$ we know that there exists a sequence of polynomials P_n such that

$$P_n(z) \to \frac{1}{z-\alpha}$$

uniformly on K. Moreover, since (by compactness) $z \mapsto (z-\alpha)^{-1}$ is bounded on K, the P_n are uniformly bounded.

On the other hand,

$$\frac{1}{z-\beta} = \frac{1}{z-\alpha-(\beta-\alpha)} = \frac{-1}{z-\alpha} \times \frac{1-(\beta-\alpha)}{z-\alpha}.$$

Since

$$\left|\frac{\beta-\alpha}{z-\alpha}\right| \leq \frac{|\beta-\alpha|}{d(\alpha,K)} < 1$$

for all $z \in K$, we know that, given $\epsilon > 0$, there exists an N with

$$\left|\frac{1}{z-\beta} - \sum_{j=0}^{N}\frac{(\beta-\alpha)^j}{(z-\alpha)^{j+1}}\right| < \epsilon/2$$

for all $z \in K$. By the first paragraph, we can find an M such that

$$\left|\frac{(\beta-\alpha)^j}{(z-\alpha)^{j+1}} - (\beta-\alpha)^j P_M(z)^j\right| < \epsilon/(2N+4)$$

for each $0 \leq j \leq N$ and so

$$\left|\frac{1}{z-\beta} - \sum_{j=0}^{N}(\beta-\alpha)^j P_M(z)^j\right| < \epsilon$$

for all $z \in K$. We have shown that $\beta \in \Lambda(K)$. □

Lemma 11.19. *Suppose that K is a compact set in $\mathbb{C}$ and $\mathbb{C} \setminus K$ is path connected. Then $\Lambda(K) = \mathbb{C} \setminus K$.*

Proof. Let $a \in \mathbb{C} \setminus K$. By Lemma 11.17, $\Lambda(K)$ is non-empty so we may choose a $b \in \Lambda(K)$. Since $\mathbb{C} \setminus K$ is path connected we can find a continuous $\gamma : [0,1] \to \mathbb{C} \setminus K$ with $\gamma(0) = b$, $\gamma(1) = a$. The continuous image of a compact set is compact and $\gamma([0,1]) \cap K = \varnothing$ so (see Lemma 10.3) there exists a $\delta > 0$ such that $|\gamma(t) - k| > \delta$ for all $k \in K$ and all $t \in [0,1]$.

By uniform continuity, we can find an N such that

$$|s - t| \leq 1/N \Rightarrow |\gamma(t) - \gamma(s)| < \delta/2.$$

Writing $x_r = \gamma(r/N)$, we see that $x_0 = b \in \Lambda(K)$ and, applying Lemma 11.18,

$$x_{r-1} \in \Lambda(K) \Rightarrow x_r \in \Lambda(K)$$

for $1 \leq r \leq N$. Thus $a = x_N \in \Lambda(K)$ and we are done. □

Since Lemma 11.19 is equivalent to Lemma 11.15, this completes the proof of our version of Runge's theorem.

It is natural to ask if the condition of uniform convergence can be dropped in Theorem 11.2. We can use Runge's theorem to show that it can not.

Example 11.20. *Let $D = \{z : |z| < 1\}$ and define $f : D \to \mathbb{C}$ by*

$$f(re^{i\theta}) = r^{3/2}e^{3i\theta/2}$$

for $r \geq 0$ and $0 < \theta \leq 2\pi$ (so that f is not even continuous). Then we can find a sequence of polynomials P_n such that $P_n(z) \to f(z)$ as $n \to \infty$ for all $z \in D$.

Proof. Let

$$\Omega_n = \mathbb{C} \setminus \{x : x \in \mathbb{R}, x \geq 2^{-n}\}.$$

Then Ω_n is open and the function g_n given by

$$g_n\big(2^{-n} + r\exp(i\theta)\big) = r^{3/2}\exp(3i\theta/2)$$

with $r > 0$, $2\pi > \theta > 0$ is a well defined analytic function on Ω_n.

If

$$K_n = \{z \in \mathbb{C} : |z| \leq 1\} \setminus \{(2^{-n-1}\exp(2^{-n}i) + r\exp(i\theta) : r > 0,\ 2^{-n} - 2^{-4n} > \theta > 2^{-n} + 2^{-4n}\},$$

then K_n is compact and the function

$$f_n(z) = g_n(\exp(2^{-n}i)z)$$

is analytic on an open set containing K_n. Thus, by Runge's theorem, we can find a polynomial P_n with

$$|P_n(z) - f_n(z)| < 2^{-n}$$

for all $z \in K_n$.

Now, if $z \in D$, we can find an N such that $z \in K_n$ for all $n \geq N$. Thus

$$|P_n(z) - f_n(z)| \to 0$$

as $n \to \infty$. Similarly $f_n(z) \to f(z)$ so $P_n(z) \to f(z)$ for all $z \in D$ and we are done. □

Thus the pointwise limit of analytic functions need not be analytic.

12 Odd numbers

According to Von Neumann[10] 'In mathematics you don't understand things. You just get used to them.' The real line is one of the most extraordinary objects in mathematics[11]. A single apparently innocuous axiom ('every increasing bounded sequence has a limit' or some equivalent formulation) calls into being an indescribably complicated object.

We know from 1A that $\mathbb{R}$ is uncountable (a different proof of this fact will be given later in Corollary 13.8). But, if we have a finite alphabet of n symbols (including punctuation), then we can only describe at most n^m real numbers in phrases exactly m symbols long. Thus the collection of describable real numbers is the countable union of finite (so countable) sets so (quoting 1A again) countable! We find ourselves echoing Newton.

> I do not know what I may appear to the world, but to myself I seem to have been only like a boy playing on the sea-shore, and diverting myself in now and then finding a smoother pebble or a prettier shell than ordinary, whilst the great ocean of truth lay all undiscovered before me. [*Memoirs of the Life, Writings, and Discoveries of Sir Isaac Newton* Brewster (Volume II. Ch. 27)]

[10]Dr Bloom suggests a a footnote in the 'The Dancing Wu Li Masters' by Gary Zukav as a source.

[11]I am being modest on behalf of analysis, I suspect the real line is the most extraordinary object in mathematics.

Let us look at some of the prettier shells.

Theorem 12.1. *The number e is irrational.*

Proof. This is probably familiar from 1A.

Suppose, if possible, that $e = p/q$ with p and q integers and $q \geq 2$. Then

$$q!e \text{ and } q!\sum_{r=0}^{q}\frac{1}{r!}$$

are integers so

$$M = q!\sum_{r=q+1}^{\infty}\frac{1}{r!} = q!e - q!\sum_{r=0}^{q}\frac{1}{r!}$$

is an integer. But

$$0 < M = q!\sum_{r=q+1}^{\infty}\frac{1}{r!} \leq \sum_{r=q+1}^{\infty}\frac{1}{q^{r-q}} = \frac{1}{q} \times \frac{1}{1-q^{-1}} = \frac{1}{q-1} < 1$$

and there is no integer strictly between 0 and 1. Our assumption has led to a contradiction so $e$must be irrational. □

Theorem 12.2. *The number π is irrational.*

Our proof of Theorem 12.2 depends on the following lemma.

Lemma 12.3. *If we write $f_n(x) = x^n(\pi - x)^n$ then*

$$\int_0^{\pi} f_n(x)\sin x\,dx = n!\sum_{j=0}^{n} a_j\pi^j$$

with a_j an integer.

Proof. Observe that

$$f_n(x) = \sum_{s=0}^{n}\binom{n}{s}\pi^{n-s}x^{n+s}$$

Thus

$$f^{(r)}(0) = 0$$

if $0 \leq r \leq n-1$ or $2n+1 \leq r$ and

$$f^{(n+r)}(0) = (n+r)!\binom{n}{r}\pi^{n-r}$$

for $0 \leq r \leq n$. By symmetry about $\pi/2$,

$$f^{(r)}(\pi) = (-1)^r f^{(r)}(0).$$

Thus $f^{(r)}(0)$ and $f^{(r)}(\pi)$ always take the form of $M \times n! \times \pi^k$ where M is an integer and k is an integer with $0 \leq k \leq n$.

Now integration by parts gives

$$\begin{aligned}\int_0^\pi f^{(m)}(x) \cos x \, dx &= \left[f^{(m)}(x) \sin x\right]_0^\pi - \int_0^\pi f^{(m+1)}(x) \sin x \, dx \\ &= -\int_0^\pi f^{(m+1)}(x) \sin x \, dx\end{aligned}$$

and

$$\begin{aligned}\int_0^\pi f^{(m)}(x) \sin x \, dx &= -\left[f^{(m)}(x) \cos x\right]_0^\pi + \int_0^\pi f^{(m+1)}(x) \cos x \, dx \\ &= (f^{(m)}(\pi) - f^{(m)}(0)) + \int_0^\pi f^{(m+1)}(x) \cos x \, dx.\end{aligned}$$

Thus integration by parts $2n+1$ times gives

$$\int_0^\pi f(x) \sin x \, dx = n! U(\pi)$$

where U is a polynomial of degree at most n with integer coefficients. □

Proof of Theorem 12.2 from Lemma 12.3. Suppose that $\pi = p/q$ with p and q integers and $q \geq 1$. It follows from Lemma 12.3 that

$$\frac{q^n}{n!} \int_0^\pi f_n(x) \sin x \, dx = q^n \sum_{j=0}^n a_j \pi^j = \sum_{j=0}^n a_j q^{n-j} p^j \in \mathbb{Z}.$$

But (by school calculus or completing the square or the AM-GM inequality) $x(\pi - x)$ takes its maximum when $x = \pi/2$ so

$$0 \leq f_n(x) \leq (\pi/2)^{2n}$$

and, since $f_n(x) \sin x$ is strictly positive for $0 < x < \pi$,

$$0 < \int_0^\pi f_n(x) \sin x \, dx \leq \int_0^\pi (\pi/2)^{2n} \, dx = \pi^{2n+1} 2^{-2n}.$$

Thus

$$0 < \frac{q^n}{n!} \int_0^\pi f_n(x) \sin x \, dx \leq \frac{1}{n!} \pi^{2n+1} 2^{-2n} q^n < 1$$

for n sufficiently large.

However there is no integer strictly between 0 and 1. Our assumption has led to a contradiction. Thus π is irrational. □

Faced with a proof like that of Theorem 12.2 the reader may cry 'How did you think of looking at $f_n(x)$?' The first, though not very helpful, answer is 'I did not, I learnt it from someone else[12]'. The second is that, even admitting that we could not have thought of it in a thousand years, once we are presented with the argument we can see a path (though not, I suspect, the actual one) which it might have been thought of. We are all familiar with the evaluation of $\int_0^\pi x^n \sin x \, dx$ and the fact that this takes the form $P(\pi)$ where P is polynomial of degree at most n with integer coefficients. It follows that if Q is a polynomial of degree n with integer coefficients then $\int_0^\pi Q(x) \sin x \, dx$ takes the form $U(\pi)$ where u is polynomial of degree at most n with integer coefficients. If $\pi = p/q$ then $q^n U(\pi)$ is an integer. We now experiment, trying to make $\int_0^\pi Q(x) \sin x \, dx$ lie between 0 and 1 in the manner of our proof that e was irrational.

For what it is worth, I think the restriction 'candidates will not be required to quote elaborate formula from memory' ought to mean that the examiners remind you of the formula for f_n in a question that requires it. However, it is also my opinion that examiners, like umpires, are always right.

It may be worth remembering that, after 300 years we still do not know if Euler's constant

$$\gamma = \lim_{N\to\infty} \left(\sum_{n=1}^{N} \frac{1}{n} - \log N \right)$$

is irrational or not.

If we think of the rationals as 'the best understood numbers' then the algebraic numbers can be thought of as 'the next best understood numbers'.

Definition 12.4. *We say that a real number α is algebraic if it is a zero of a polynomial with integer coefficients. Real numbers which are not algebraic are called transcendental.*

Exercise 12.5. *Show that a real number α is algebraic if and only if it is a zero of a polynomial with rational coefficients.*

Proof. Only if is trivial since integers are rational numbers.

To see *if*, observe that, if α satisfies

$$\sum_{j=0}^{N} \frac{p_j}{q_j} \alpha^j = 0$$

with p_j, q_j integers, $q_j \neq 0$ for all j, $p_N \neq 0$, $N \geq 1$, then

$$\sum_{j=0}^{N} p_j \prod_{i \neq j} q_i \alpha^j = 0.$$

[12]Professor Gowers in this case.

□

Lemma 12.6. *The algebraic numbers are countable.*

Proof. This was done in 1A. There are only finitely many polynomials of the form

$$\sum_{j=0}^{N} a_j x^j$$

with $n \geq N \geq 1$, $a_N \neq 0$ and all a_j integers with $|a_j| \leq n$. A polynomial has only finitely many roots, so the set E_n of roots of such polynomials is finite so countable. Thus $E = \bigcup_{n=1}^{\infty} E_n$ is the countable union of countable sets so countable. But E is the set of algebraic numbers, so we are done. □

Since the reals are uncountable, this shows that transcendental numbers exist.

The argument just given (which you saw in 1A) is due to Cantor. It is beautiful but non-constructive. It tells us that transcendental numbers exist (indeed that uncountably many transcendental numbers exist) without showing us any.

The first proof that transcendentals exist is due to Liouville. It is longer but actually produces particular examples.

Theorem 12.7. [Liouville] *Suppose α is an irrational root of the equation*

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0 = 0$$

where $a_j \in \mathbb{Z}$ $[0 \leq j \leq n]$, $n \geq 1$ and $a_n \neq 0$. Then there is a constant $c > 0$ (depending on the a_j) such that

$$\left| \alpha - \frac{p}{q} \right| \geq \frac{c}{q^n}$$

for all p, $q \in \mathbb{Z}$ with $q \neq 0$

Proof. Let

$$P(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_0.$$

Since a polynomial has only finitely many roots, we can find an $R \geq 1$ such that all the roots of P lie in $[-R+1, R-1]$. If we take $0 < c \leq 1$, the required result will be automatic for $p/q \notin [-R, R]$.

Now P' is continuous, so, by compactness, there exists an $M > 1$ such that $|P'(t)| \leq M$ for $t \in [-R, R]$. (We could also prove this directly.) If α is an irrational root, $p, q \in \mathbb{Z}$ with $q \neq 0$, $p/q \in [-R, R]$ and $P(p/q) \neq 0$, then the mean value theorem yields

$$|P(\alpha) - P(p/q)| \leq M \left| \alpha - \frac{p}{q} \right|$$

so, since $P(\alpha) = 0$,

$$|P(p/q)| \leq M \left| \alpha - \frac{p}{q} \right|.$$

Now $q^n P(p/q)$ is a non-zero integer, so $|q^n P(p/q)| \geq 1$ and

$$q^{-n} \leq M \left| \alpha - \frac{p}{q} \right|,$$

that is to say

$$M^{-1} q^{-n} \leq \left| \alpha - \frac{p}{q} \right|.$$

Since there are only a finite number of roots and so only a finite number of irrational roots, we know that there is a $c' > 0$ such that

$$\left| \alpha - \frac{p}{q} \right| \geq c' q^{-n}$$

whenever α is an irrational root, $p, q \in \mathbb{Z}$ with $q \neq 0$ and $P(p/q) = 0$.

Taking $c = \min\{M^{-1}, c', 1\}$, we have the required result. □

We say that 'irrational algebraic numbers are not well approximated by rationals'.

Theorem 12.8. *The number*

$$\sum_{j=0}^{\infty} \frac{1}{10^{j!}}$$

is transcendental.

Proof. Let

$$L = \sum_{n=0}^{\infty} \frac{1}{10^{n!}}.$$

We observe that L is irrational since its decimal expansion is not recurring. If $q_m = 10^{m!}$ and

$$p_m = q_m \sum_{n=0}^{m} \frac{1}{10^{n!}},$$

then p_m and q_m are integers with $q_m \neq 0$.

We observe that

$$\left|L - \frac{p_m}{q_m}\right| = \sum_{j=m+1}^{\infty} \frac{1}{10^{j!}} \leq \frac{1}{10^{(m+1)!}} \sum_{j=0}^{\infty} \frac{1}{10^j} \leq \frac{2}{10^{(m+1)!}}$$

and, given any $c > 0$ and any integer $n \geq 1$, we can find an m such that

$$\left|L - \frac{p_m}{q_m}\right| \leq \frac{2}{10^{(m+1)!}} < \frac{c}{q^n}.$$

Thus Theorem 12.7 tells us that L is transcendental. □

Exercise 12.9. *By considering*

$$\sum_{n=0}^{\infty} \frac{b_n}{10^{n!}}$$

with $b_j \in \{1, 2\}$, give another proof that the set of transcendental numbers is uncountable.

Solution. Essentially the same argument as for Theorem 12.8 tells us that

$$\sum_{n=0}^{\infty} \frac{b_j}{10^{n!}}$$

with $b_j \in \{1, 2\}$ is transcendental.

The map

$$\theta : \{0, 1\}^{\mathbb{N}} \to \mathbb{R}$$

given by

$$\theta(\boldsymbol{\zeta}) = \sum_{n=0}^{\infty} \frac{\zeta(j) + 1}{10^{n!}}$$

is injective and its image (as we have just seen) consists of transcendental numbers. Since, as we saw in 1A, the set $\{0, 1\}^{\mathbb{N}}$ is uncountable and θ is injective, $\theta(\mathbb{N})$ is uncountable and we are done. □

As might be expected, it turned out to be very hard to show that particular numbers are transcendental. Hermite proved that e is transcendental and Lindemann adapted Hermite's method to show that π is transcendental (and so the circle can not be squared). Alan Baker contributed greatly to this field, and his book *Transcendental number theory* [2] contains accessible[13] proofs of the transcendence of e and π.

[13]But miraculous.

13 The Baire category theorem

The following theorem turns out to be much more useful than its somewhat abstract formulation makes it appear.

Theorem 13.1. [The Baire category theorem] *If (X,d) is a complete non-empty metric space and U_1, U_2, ... are open sets whose complements have empty interior, then*

$$\bigcap_{j=1}^{\infty} U_j \neq \varnothing.$$

Proof. We construct $x_j \in X$ and $\delta_j > 0$ inductively as follows. Choose any $x_0 \in X$ and set $\delta_0 = 1$.

Suppose that x_j and δ_j have been found. Since $X \setminus U_{j+1}$ has empty interior, we can find an $x_{j+1} \in U_{j+1}$ with $d(x_{j+1}, x_j) \leq \delta_j/4$. Since U_{j+1} is open we can find a $\delta_{j+1} > 0$ with $\delta_{j+1} \leq \delta_j/4$ such that $B(x_{j+1}, \delta_{j+1}) \subseteq U_{j+1}$.

By induction, $\delta_{j+k} \leq 4^{-k}\delta_j$ for $j, k \geq 0$, so, if $m \geq n \geq 0$,

$$\begin{aligned} d(x_n, x_m) &\leq \sum_{r=0}^{m-n-1} d(x_{n+r}, x_{n+r+1}) \\ &\leq \sum_{r=0}^{m-n-1} \delta_{n+r}/4 \leq \sum_{r=0}^{m-n-1} \delta_n 4^{-r-1} \leq \delta_n/2, \end{aligned}$$

so the sequence x_n is Cauchy and so converges to some point a.

We observe that

$$d(x_n, a) \leq d(x_n, x_m) + d(x_m, a) \leq \delta_n/2 + d(x_m, a) \to \delta_n/2$$

as $m \to \infty$. Thus

$$d(x_n, a) \leq d(x_n, x_m) + d(x_m, a) \leq \delta_n/2 + d(x_m, a) \to \delta_n/2$$

as $m \to \infty$. Thus

$$a \in B(x_j, \delta_j) \subseteq U_j$$

for all $j \geq 1$ and

$$a \in \bigcap_{j=1}^{\infty} U_j.$$

The result is proved □

Taking complements gives the following equivalent form.

Theorem 13.2. *If (X, d) is a complete non-empty metric space and F_1, F_2, ... are closed sets with empty interior, then*

$$\bigcup_{j=1}^{\infty} F_j \neq X.$$

Proof of the equivalence of Theorems 13.1 and 13.2. Set $F_j = X \setminus U_j$. □

I think of Baire's theorem in yet another equivalent form.

Theorem 13.3. *Let (X, d) be a non-empty complete metric space. Suppose that P_j is a property such that:-*

(i) The property of being P_j is stable *in the sense that, given $x \in X$ which has property P_j, we can find an $\epsilon > 0$ such that whenever $d(x, y) < \epsilon$ the point y has the property P_j.*

(ii) The property of not being P_j is unstable *in the sense that, given $x \in X$ and $\epsilon > 0$, we can find a $y \in X$ with $d(x, y) < \epsilon$ which has the property P_j.*

Then there is an $x_0 \in X$ which has all of the of the properties P_1, P_2, ...

Proof of the equivalence of Theorems 13.1 and 13.3. Let x have the property P_j if and only if $x \notin U_j$. □

Baire's theorem has given rise to the following standard definitions[14].

Definition 13.4. *A set in a metric space is said to be* nowhere dense *if its closure has empty interior. A set in a metric space is said to be of* first category *if it is a subset of a countable union of nowhere dense closed sets. Any set which is not of first category is said to be of* second category.

Your lecturer will try never to use the words second category but always to talk about 'not first category'. If all points outside a set of first category have a property P, I shall say that *quasi-all* points have property P.

Two key facts about first countable sets are stated in the next lemma.

Lemma 13.5. *(i) If (X, d) is a non-empty complete metric space and E is a subset of first category, then $E \neq X$.*

(ii) The countable union of sets of first category is of first category.

Proof. (i) This is just a restatement of Theorem 13.2.

(ii) The countable union of countable sets is countable. □

[14] Your lecturer thinks that, whilst the concepts defined are very useful, the nomenclature is particularly unfortunate.

We need one more definition.

Definition 13.6. *(i) If (X, d) is a metric space, we say that a point $x \in X$ is* isolated *if we can find a $\delta > 0$ such that $B(x, \delta) = \{x\}$.*

(ii) If (X, d) is a metric space, we say that a subset E of X contains no isolated points if, whenever $x \in E$ and $\delta > 0$, we have $B(x, \delta) \cap E \neq \{x\}$.

Theorem 13.7. *A non-empty complete metric space without isolated points is uncountable.*

Proof. Suppose that (E, d) is a non-empty countable complete space with no isolated points. Then each $\{e\}$ with $e \in E$ is closed (since singletons are always closed in metric spaces). However, since e not isolated, $B(x, \delta) \not\subseteq \{e\}$ for all $\delta > 0$, so $\{e\}$ is not open and $\{e\}$ has empty interior. Thus E is the countable union of closed sets $\{e\}$ with empty interior contradicting Theorem 13.2. The required result follows by reductio ad absurdum. □

Corollary 13.8. *The real numbers are uncountable.*

Proof. Observe that $\mathbb{R}$ with the usual metric is complete without isolated points. Theorem 13.7 now tells us that $\mathbb{R}$ is uncountable. □

The proof we have given here is much closer to Cantor's original proof than that given in 1A. It avoids the use of extraneous concepts like decimal representation.

Banach was a master of using the Baire category theorem. Here is one of his results.

Theorem 13.9. *Consider $C([0, 1])$ with the uniform norm. The set of anywhere differentiable functions is a set of the first category. Thus continuous nowhere differentiable functions exist.*

Proof. Banach's clever idea is to consider the set E_m consisting of all those $f \in C([0, 1])$ such that there exists an $x \in [0, 1]$ with the property

$$|f(x) - f(y)| \leq m|x - y|$$

for all $y \in [0, 1]$. Our proof falls into several parts.

(a) We show that, if f is differentiable at some point $x \in [0, 1]$, then there exists a positive integer m such that $f \in E_m$. It will then follow that any $g \in C([0, 1]) \setminus \bigcup_{m=1}^{\infty} E_m$ is nowhere differentiable.

To this end, suppose that f is differentiable at x. We can find an $\epsilon > 0$ such that

$$\left| \frac{f(x + h) - f(x)}{h} - f'(x) \right| \leq 1$$

for all $0 < |h| < \epsilon$, when $x + h \in [0, 1]$. Thus

$$|f(x+h) - f(x)| \leq (|f'(x)| + 1)|h|$$

for all $0 < |h| < \epsilon$ when $x + h \in [0, 1]$. We thus have

$$|f(x) - f(y)| \leq (|f'(x)| + 1)|x - y|$$

for all $y \in [0, 1]$ such that $|y - x| < \epsilon$. If we choose m with $m \geq |f'(x)| + 1$ and $m \geq K\epsilon^{-1}$, we will have $f \in E_m$.

(b) We now show that E_m is closed.

Suppose that $f_n \in E_m$ and $\|f_n - f\|_\infty \to 0$. By definition, there exists an $x_n \in [0, 1]$ with the property

$$|f_n(x_n) - f(y)| \leq m|x_n - y|$$

for all $y \in [0, 1]$. By the Bolzano–Weierstrass property, we can find $x \in [0, 1]$ and $n(r) \to \infty$ such that $x_{n(r)} \to x$ as $r \to \infty$.

Let $y \in [0, 1]$. We have

$$\begin{aligned}
|f(x) - f(y)| &\leq |f(x) - f(x_{n(r)})| + |f(x_{n(r)}) - f_{n(r)}(x_{n(r)})| \\
&\qquad + |f_{n(r)}(x_{n(r)}) - f_{n(r)}(y)| + |f_{n(r)}(y) - f(y)| \\
&\leq 2\|f - f_{n(r)}\|_\infty + |f(x) - f(x_{n(r)})| + m|x_{n(r)} - y| \\
&\to 0 + 0 + m|x - y| = m|x - y|.
\end{aligned}$$

Since y was arbitrary, $f \in E_m$.

(c) Next we show that E_m has a dense complement.

Suppose that $f \in C([0, 1])$ and $\epsilon > 0$. By Weierstrass's theorem on polynomial approximation (see Theorem 7.9), we can find a polynomial P such that

$$\|f - P\|_\infty \leq \epsilon/3.$$

Since P is continuously differentiable, there is a K such that $|P'(t)| \leq K$ for all $t \in [0, 1]$. By the mean value theorem, it follows that

$$|P(x) - P(y)| \leq K|x - y|$$

for all x, $y \in [0, 1]$.

Let $g(t) = P(t) + (\epsilon/3)\cos 2\pi N t$. Automatically,

$$\|g - f\|_\infty \leq \|f - P\|_\infty + \epsilon/3 \leq 2\epsilon/3 < \epsilon.$$

We claim that, provided only that N is large enough, $g \notin E_m$.

To see this choose r an integer with $0 \leq r \leq N-1$ such that $0 \leq x - r/N \leq 1/N$. We have

$$\begin{aligned}\max\{|g(r/N) - g(x)|, |g((r+1)/N) - g(x)|\}\\ &\geq \frac{|g(r/N) - g(x)| + |g((r+1)/N) - g(x)|}{2}\\ &\geq \frac{|g(r/N) - g((r+1)/N)|}{2}\\ &\geq \frac{2\epsilon/3 - |P(r/N) - P((r+1)/N)|}{2}\\ &\geq \epsilon/3 - K/N \geq \epsilon/6 \geq 4m/N.\end{aligned}$$

Thus at least one of the statements

$$|g(r/N) - g(x)| > m|r/n - x| \text{ or } |g((r+1)/N) - g(x)| > m|(r+1)/n - x|$$

is true for N sufficiently large (with N not depending on the choice of x).

(d) Thus $\bigcup_{m=1}^{\infty} E_m$ is a set of first category and we are done. □

Here is another corollary of Theorem 13.7.

Corollary 13.10. *A non-empty closed subset of* $\mathbb{R}$ *without isolated points is uncountable.*

Proof. Observe that a closed subset of a complete metric space is complete under the inherited metric and that $\mathbb{R}$ is complete under the standard metric. □

Do there exist nowhere dense closed subsets of $\mathbb{R}$ with no isolated points[15]? We shall answer this question by applying Baire's theorem in the context of the Hausdorff metric.

Lemma 13.11. *Consider the space* $\mathcal{K}$ *of non-empty compact subsets of* $[0,1]$ *with the Hausdorff metric* ρ. *Let* $\mathcal{E}_k$ *be the collection of compact sets* E *such that there exists an* $x \in E$ *with* $B(x, 1/k) \cap E = \{x\}$.

(i) The set $\mathcal{E}_k$ *is closed in the Hausdorff metric.*

(ii) The set $\mathcal{E}_k$ *is nowhere dense in the Hausdorff metric.*

(iii) The set $\mathcal{E}$ *of compact sets with an isolated point is of first category with respect to the Hausdorff metric.*

[15] Such sets are called perfect. If they make pretty pictures they are called fractals. You do not have to remember either name.

Proof. (i) Suppose that $E_n \in \mathcal{E}_k$ and $E_n \to E$ in the Hausdorff metric. By definition we can find $x_n \in E_n$ with $B(x_n, 1/k) \cap E = \{x_n\}$. By Bolzano–Weierstrass, we can find $n(j) \to \infty$ and $x \in [0,1]$ such that $|x_n(j) - x| \to 0$. We observe that $x \in E$.

Suppose, if possible, that $B(x, 1/k) \cap E \neq \{x\}$. Then we can find a $y \in E$ such that $|x - y| < 1/k$. Set $\delta = (1/k - |x - y|)/2$. Since $E_n \to E$ and $n(j) \to \infty$, we can find a J such that the Hausdorff distance $\rho(E_{n(J)}, E) < \delta$ and so there exists a $y' \in E_{n(J)}$ with $|y' - y| < \delta$ and so with $|x_{n(J)} - y'| < 1/k$, contrary to our hypothesis.

Thus $E \in \mathcal{E}_k$ and $\mathcal{E}_k$ is closed.

(ii) Let $G \in \mathcal{K}$ and let $\epsilon > 0$. Choose an integer $N > 5(\epsilon^{-1} + k + 1)$. Let

$$F = \{r/N : |r/N - x| \leq 4/N \text{ for some } x \in G\}.$$

By construction,

$$\sigma(G, F) = \sup_{\mathbf{y} \in G} d(y, F) \leq 1/N \text{ and } \sigma(F, G) = \sup_{\mathbf{y} \in F} d(y, G) \leq 4/N$$

so

$$\rho(G, F) = \sigma(G, F) + \sigma(F, G) \leq 5/N < \epsilon$$

But, if $y \in G$, then either $y + 1/N$ or $y - 1/N$ (or both) lies in G, so $G \notin \mathcal{E}_k$.

(iii) Observe that $\mathcal{E} = \cup_{k=1}^{\infty} \mathcal{E}_k$. □

Lemma 13.12. *Consider the space $\mathcal{K}$ of non-empty compact subsets of $[0,1]$ with the Hausdorff metric ρ. Let $\mathcal{F}_{j,k}$ be the collection of compact sets F such that $F \supseteq [j/k, (j+1)/k]$ is closed and nowhere dense $[0 \leq j \leq k,\ 1 \leq k]$.*

(i) The set $\mathcal{F}_{j,k}$ is closed in the Hausdorff metric.

(ii) The set $\mathcal{F}_{j,k}$ is nowhere dense in the Hausdorff metric.

(ii) The set $\mathcal{F}$ of compact sets with non-empty interior is of first category.

Lemma 13.12. (i) Suppose that $F_n \in \mathcal{F}_{j,k}$ and $F_n \to F$ in the Hausdorff metric. By definition, $F_n \supseteq [j/k, (j+1)/k]$, so $F \supseteq [j/k, (j+1)/k]$.

Thus $\mathcal{F}_{j,k}$ is closed

(ii) Let $G \in \mathcal{K}$ and let $\epsilon > 0$. Choose an integer $N > 2\epsilon^{-1} + 1$. Let

$$E = \{r/N : |r/N - x| \leq 1/N \text{ for some } x \in G\}.$$

By construction,

$$\sigma(G, E) = \sup_{\mathbf{y} \in G} d(y, E) \leq 1/N \text{ and } \sigma(E, G) = \sup_{\mathbf{y} \in E} d(y, G) \leq 1/N$$

so

$$\rho(G, E) = \sigma(G, E) + \sigma(E, G) \leq 2/N < \epsilon$$

but $E \notin \mathcal{F}_{j,k}$.

(iii) Observe that $\mathcal{F} = \bigcup_{k=1}^{\infty} \bigcup_{j=0}^{k} \mathcal{F}_{j,k}$, so $\mathcal{F}$ is the countable union of closed nowhere dense sets. □

Theorem 13.13. *The set $\mathcal{C}$ of non-empty compact sets with empty interior and no isolated points is the complement of a set of first category in the space $\mathcal{K}$ of non-empty compact subsets of $[0,1]$ with the Hausdorff metric ρ.*

Proof. With the notation of Lemmas 13.11 and 13.12,

$$\mathcal{K} \setminus \mathcal{C} = \mathcal{E} \cup \mathcal{F}.$$

Since the union of two first category sets is of first category, $\mathcal{K} \setminus \mathcal{C}$ is of the first category. □

Since $\mathcal{K}$ with the Hausdorff metric is complete it follows that non-empty compact sets with empty interior and no isolated points exist.

The following example provides a background to our next use of Baire category.

Exercise 13.14. *(i) Show that we can find continuous functions $g_n : [0,1] \to \mathbb{R}$ such that $g_n(x) \to 0$ for each $x \in [0,1]$ but*

$$\sup_{t \in [0,1]} g_n(t) \to \infty$$

as $n \to \infty$.
[*Hint: Witch's hat.*]

(ii) Show that we can find continuous functions $f_n : [0,1] \to \mathbb{R}$ such that $f_n(x) \to 0$ for each $x \in [0,1]$ but

$$\sup_{t \in [2^{-r-1}, 2^{-r}]} f_n(t) \to \infty$$

as $n \to \infty$ for each integer $r \geq 0$.

Solution. (i) (This may be familiar from 1B.) Set

$$g_n(x) = \begin{cases} 2^{2n}x & \text{if } 0 \leq x \leq 2^{-n-1}, \\ 2^{2n}(2^{-n} - x) & \text{if } 2^{-n-1} \leq x \leq 2^{-n}, \\ 0 & \text{otherwise.} \end{cases}$$

If $x \neq 0$, then $x \geq 2^{-m}$ for some m and so $g_n(x) = 0$ for $n \geq m$. Since $g_n(0) = 0$ for all n, we have $g_n(x) \to 0$ as $n \to \infty$ for all x.

However,

$$\sup_{t\in[0,1]} g_n(t) = g(2^{-n-1}) = 2^{n-1} \to \infty$$

as $n \to \infty$.

(ii) Extend g_n to a function on $\mathbb{R}$ by setting $g_n(t) = 0$ for $t \notin [0,1]$. Set $f_n(t) = \sum_{j=1}^{\infty} 2^{-j} g_n\big(2^j(t-2^{-j})\big)$ and use (i). □

In spite of the previous example we have the following remarkable theorem.

Theorem 13.15. *Suppose that we have a sequence of continuous functions $f_n : [0,1] \to \mathbb{R}$ such that $f_n(x) \to 0$ for each $x \in [0,1]$ as $n \to \infty$. Then we can find a non-empty interval $(a,b) \subseteq [0,1]$ and an $M > 0$ such that*

$$|f_n(t)| \le M$$

for all $t \in (a,b)$ and all $n \ge 1$.

Proof. Observe that

$$E_{n,m} = \{x : |f_n(x)| \le m\} = f_n^{-1}([-m,m])$$

is closed (since f_n is continuous), so

$$E_m = \bigcap_{n=1}^{\infty} E_{n,m}$$

is.

If we fix $x \in [0,1]$ for the moment, we know that $f_n(x) \to 0$ as $n \to \infty$. In particular, we can find an $N(x)$ such that $|f_n(x)| \le 1$ for all $n \ge N(x)$. Thus

$$|f_n(x)| \le \max\{1, \max_{1\le j\le N(x)} |f_n(x)|\}$$

and so $x \in E_{m(x)}$ for some integer $m(x)$.

The previous paragraph shows that

$$[0,1] = \bigcup_{m=1}^{\infty} E_m,$$

but Baire's category theorem tells us that $[0,1]$ cannot be the countable union of closed sets with empty interior. Thus there must exist an M such that E_M has non-empty interior, so $E_M \supseteq (a,b)$ for some non-empty interval (a,b). □

A slightly stronger version of the result is given as Exercise 20.11.

14 Continued fractions

We are used to writing real numbers as decimals, but there are other ways of specifying real numbers which may be more convenient. The oldest of these is the method of continued fractions. Suppose that x is irrational and $1 \geq x > 0$. We know that there is a strictly positive integer $N(x)$ such that

$$\frac{1}{N(x)} \geq x > \frac{1}{N(x)+1}$$

so we can write

$$x = \frac{1}{N(x)+T(x)}$$

where $T(x)$ is irrational and $1 > T(x) \geq 0$. We can do the same things to $T(x)$ as we did to x, obtaining

$$T(x) = \frac{1}{N(T(x))+T(T(x))}$$

and so, using the standard notation for composition of functions,

$$x = \cfrac{1}{N(x) + \cfrac{1}{NT(x)+T^2(x)}}.$$

The reader[16] will have no difficulty in proceeding to the next step and obtaining

$$x = \cfrac{1}{N(x) + \cfrac{1}{NT(x) + \cfrac{1}{NT^2(x)+T^3(x)}}},$$

and so on indefinitely. We call

$$\cfrac{1}{N(x) + \cfrac{1}{NT(x) + \cfrac{1}{NT^2(x) + \cfrac{1}{NT^3(x)+\dots}}}}$$

the *continued fraction expansion* of x. [Note that, for the moment, this is simply a pretty way of writing the infinite sequence $N(x)$, $NT(x)$, $NT^2(x)$,

[16]As opposed to typesetters; this sort of thing turned their hair prematurely grey.

.... In the next section we shall show first that the continued fraction can be asigned a numerical meaning and then that the asigned meaning is, as we might hope, x.]

If y is a general irrational number, we call

$$[y] + \cfrac{1}{N(x) + \cfrac{1}{NT(x) + \cfrac{1}{NT^2(x) + \cfrac{1}{NT^3(x) + \dots}}}}.$$

the *continued fraction expansion* of y.

We can do the same thing if y is rational, but we must allow for the possibility that the process does not continue indefinitely. It is instructive to carry out the process in a particular case.

Exercise 14.1. *Carry out the process outlined above for* 100/37. *Carry out the process for the rational of your choice.*

Solution.

$$\begin{aligned} \frac{100}{37} &= 2 + \frac{26}{37}, \\ \frac{37}{26} &= 1 + \frac{11}{26}, \\ \frac{26}{11} &= 2 + \frac{4}{11}, \\ \frac{11}{4} &= 2 + \frac{3}{4}, \\ \frac{4}{3} &= 1 + \frac{1}{3}. \end{aligned}$$

Thus

$$\frac{100}{37} = 2 + \cfrac{1}{1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{1 + \frac{1}{3}}}}}.$$

□

Once we have done a couple of examples it is clear that we are simply echoing Euclid's algorithm[17].

Lemma 14.2. *(i) Suppose that* $0 < x < 1$, $x = a_1/a_0$ *with* a_1 *and* a_0 *coprime strictly positive integers and* $a_0 \geq 2$. *Then*

$$x = \cfrac{1}{k_1 + \cfrac{a_2}{a_1}}$$

with

$$a_0 = k_1 a_1 + a_2$$

for some positive integer k_1. *Thus the pair* a_1, a_2 *is obtained from* (a_0, a_1) *by applying one step of the Euclidean algorithm.*

(ii) If we use the notation of (i) a_1 *and* a_2 *are coprime strictly positive integers with* $0 < a_2/a_1 < 1$.

(iii) If y *is a rational number, its continued fraction expansion (obtained by the method described above) terminates.*

Lemma 14.2. (i) This is immediate.

(ii) This is a result from 1A. (The highest common factor of each pair in the Euclidean algorithm is the same.)

(iii) We saw in 1A that the Euclidean algorithm terminates. (Or we could repeat the 1A proof by observing that the elements of the pairs are strictly decreasing.) □

Exercise 14.3. *Show that* $\sqrt{2}$ *has continued fraction expansion*

$$1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \dots}}}}.$$

Deduce that $\sqrt{2}$ *is irrational.*

Solution. We know that $1 < \sqrt{2} < 2$, so

$$\sqrt{2} = 1 + \alpha$$

[17] I think historians would reverse the order and say that continued fractions gave rise to Euclid's algorithm

with $0 < \alpha = \sqrt{2} - 1 < 1$.

Now

$$\frac{1}{\alpha} = \frac{1}{\sqrt{2}-1} = \sqrt{2} + 1 = 2 + \alpha$$

so $N(\alpha) = 2$ and $T(\alpha) = 2+\alpha$. Thus $\sqrt{2}$ has the *non-terminating* continued fraction

$$1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2 + \dots}}}}$$

and cannot be rational. $\square$

If we look at a random variable with the uniform distribution on $[0, 1]$ then the successive terms in the decimal expansion of X will be independent and will take the value j with probability $1/10$ $[0 \leq j \leq 9]$.

Exercise 14.4. *(This easy exercise formalises the remark just made.) If $x \in [0, 1)$ let us write $Dx = 10x - [10x]$ (in other words, Dx is the fractional part of $10x$) and $Nx = [10x]$). Show that*

$$x = 10^{-1}(Dx + Nx) = 10^{-1}Nx + 10^{-2}NDx + 10^{-2}D^2x = \dots$$

and write down the next term in the chain of equalities explicitly.

If X is a random variable with uniform distribution on $[0, 1]$ show that NX, NDX, ND^2X, ... are independent and

$$\Pr(ND^kX = j) = 1/10$$

for $0 \leq j \leq 9$.

Solution. We have

$$Dx = 10x - [10x] = 10x - Nx$$

so

$$x = 10^{-1}(Dx + Nx).$$

Since $Dx \in [0, 1)$, we have

$$\begin{aligned} x &= 10^{-1}(Dx + Nx) = 10^{-1}\big(10^{-1}\big(D(Dx) + N(Dx)\big) + Nx\big) \\ &= 10^{-1}Nx + 10^{-2}NDx + 10^{-2}D^2x \\ &= 10^{-1}Nx + 10^{-2}NDx + 10^{-3}ND^2x + 10^{-3}D^3x \\ &= \dots. \end{aligned}$$

We have

$$\begin{aligned}\Pr(ND^rX = k_r \text{ for } 1 \leq r \leq n) \\ = \Pr\left(\sum_{r=1}^{n} k_r 10^{-r} \leq X < 10^{-n} + \sum_{r=1}^{n} k_r 10^{-r}\right) = 10^{-n},\end{aligned}$$

so

$$\Pr(ND^kX = j) = 1/10$$

for $0 \leq j \leq 9$ and

$$\Pr(ND^rX = k_r \text{ for } 1 \leq r \leq n\} = 10^{-n} = \prod_{r=1}^{n} \Pr(ND^rX = k_r),$$

showing that NX, NDX, ND^2X, ... are independent □

Gauss made the following observation.

Lemma 14.5. *Suppose that X is a random variable on $[0, 1)$ with density function*

$$f(x) = \left(\frac{1}{\log 2}\right) \frac{1}{1+x}.$$

Then TX is a random variable with the same density function.

Proof. Observe that

$$\begin{aligned}
\Pr(TX \le a) &= \Pr(n \le X^{-1} \le n + a \text{ for some integer } n \ge 1) \\
&= \Pr\big((n+a)^{-1} \le X \le n^{-1} \text{ for some integer } n \ge 1\big) \\
&= \sum_{n=1}^{\infty} \Pr\big((n+a)^{-1} \le X \le n^{-1}\big) \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \int_{(n+a)^{-1}}^{n^{-1}} \frac{1}{1+x}\, dx \\
&= \sum_{n=1}^{\infty} \frac{1}{\log 2} \Big[\log(1+x) \Big]_{(n+a)^{-1}}^{n^{-1}} \\
&= \frac{1}{\log 2} \sum_{n=1}^{\infty} \big(\log(1+n^{-1}) - \log(1+(n+a)^{-1}) \big) \\
&= \frac{1}{\log 2} \sum_{n=1}^{\infty} \Big(\big(\log(n+1) - \log n \big) - \big(\log(1+n+a) - \log(n+a) \big) \Big) \\
&= \frac{1}{\log 2} \lim_{N\to\infty} \big(\log(N+1) - \log(1+N+a) + \log(1+a) \big) \\
&= \frac{1}{\log 2} \left(\log(1+a) + \lim_{N\to\infty} \log \frac{N+1}{1+N+a} \right) \\
&= \frac{\log(1+a)}{\log 2} = \Pr(X \le a).
\end{aligned}$$

□

Corollary 14.6. *Suppose that X is a random variable on $[0,1]$ with density function*

$$f(x) = \left(\frac{1}{\log 2} \right) \frac{1}{1+x}.$$

Then

$$\Pr(NT^m X = j) = \frac{1}{\log 2} \int_{1/(j+1)}^{1/j} \frac{dx}{1+x} = \frac{1}{\log 2} \log \left(\frac{(j+1)^2}{j(j+2)} \right).$$

Proof. By Lemma 14.5,

$$\begin{aligned}\Pr(NT^m X = j) &= \Pr(NX = j)\\ &= \frac{1}{\log 2}\int_{j^{-1}}^{(j+1)^{-1}} \frac{1}{1+x}\,dx\\ &= \frac{1}{\log 2}\Big[\log(1+x)\Big]_{(j+1)^{-1}}^{j^{-1}}\\ &= \frac{1}{\log 2}\big(\log(j+2) - \log(j+1)\big) = \frac{1}{\log 2}\log\frac{j+2}{j+1}.\end{aligned}$$

□

With a little extra work (which we shall not do) we can show that, if X has the density suggested by Gauss, then NX, NTX, NT^2X, ... are independent random variables all with the same probability distribution. It is also not hard to guess, and not very hard to prove, that if Y is uniformly distributed on $[0,1]$, then

$$\Pr(NT^m Y = j) \to \frac{1}{\log 2}\log\left(\frac{(j+1)^2}{j(j+2)}\right).$$

as $m \to \infty$, but we shall not take the matter further.

15 Continued fractions (continued)

In the previous section we showed how to compute the continued fraction associated with a real number x, but we did not really consider what exact meaning was to be assigned to the result. In this section we show that continued fractions really do what we might hope they do.

Definition 15.1. *If a_1, a_2, ... is a sequence of strictly positive integers and a_0 is a positive integer we call*

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \dots}}}}$$

the associated continued fraction.

Lemma 15.2. *(i) We use the notation of Definition 15.1. If we take*

$$r_n = a_n, \;\; s_n = 1$$

and define r_k and s_k in terms of r_{k+1} and s_{k+1}

$$\begin{pmatrix} r_k \\ s_k \end{pmatrix} = \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_{k+1} \\ s_{k+1} \end{pmatrix},$$

then

$$\frac{r_k}{s_k} = a_k + \cfrac{1}{a_{k+1} + \cfrac{1}{a_{k+2} + \cfrac{1}{a_{k+3} + \cfrac{1}{a_{k+4} + \cfrac{1}{\ddots \, \cfrac{}{a_{n-1} + \cfrac{1}{a_n}}}}}}}.$$

(ii) If we set

$$\begin{pmatrix} r_0 \\ s_0 \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \dots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ 1 \end{pmatrix},$$

then

$$\frac{r_0}{s_0} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cfrac{1}{\ddots \, \cfrac{}{a_{n-1} + \cfrac{1}{a_n}}}}}}}$$

Proof. (i) We use backwards induction on k. Since

$$\frac{r_n}{s_n} = a_n$$

the result is true for $k = n$.

Suppose the result is true for $m+1$ with $0 \leq m \leq n-1$. Then, by definition,

$$\begin{pmatrix} r_m \\ s_m \end{pmatrix} = \begin{pmatrix} a_m & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_{m+1} \\ s_{m+1} \end{pmatrix} = \begin{pmatrix} a_m r_{m+1} + s_{m+1} \\ r_{m+1} \end{pmatrix},$$

and, by the inductive hypothesis,

$$\begin{aligned} a_m + &\cfrac{1}{a_{m+1} + \cfrac{1}{a_{m+2} + \cfrac{1}{a_{m+3} + \cfrac{1}{a_{m+4} + \cfrac{1}{\ddots \cfrac{}{a_{n-1} + \cfrac{1}{a_n}}}}}}} \\ &= a_m + \frac{1}{r_{m+1}/s_{m+1}} \\ &= a_m + \frac{s_{m+1}}{r_{m+1}} \\ &= \frac{a_m r_{m+1} + s_{m+1}}{r_{m+1}} = \frac{r_m}{s_m}. \end{aligned}$$

The required result now follows.

(ii) Apply (i) repeatedly. □

We now read everything off in the opposite direction

Lemma 15.3. *(i) If we set*

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \dots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ 1 \end{pmatrix},$$

then

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cfrac{1}{\ddots \cfrac{}{a_{n-1} + \cfrac{1}{a_n}}}}}}}$$

(ii) Further

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix},$$

Lemma 15.3. (i) This is just a restatement of Lemma 15.2 (ii).

(ii) We have

$$\begin{pmatrix} p_{n-1} \\ q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-2} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ 1 \end{pmatrix}$$

and

$$\begin{pmatrix} a_{n-1} \\ 1 \end{pmatrix} = \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

□

The pay-off for our work in recasting matters in matricial form comes in the next theorem.

Theorem 15.4. *Choose p_j and q_j as in Lemma 15.3.*

(i) $p_k q_{k-1} - q_k p_{k-1} = (-1)^{k+1}$ for all k.

(ii) $q_k = a_k q_{k-1} + q_{k-2}$ and $p_k = a_k p_{k-1} + p_{k-2}$ for all $k \geq 2$.

(iii) p_k and q_k are coprime for all k.

(iv) We have

$$\frac{p_{2k}}{q_{2k}} > \frac{p_{2k-2}}{q_{2k-2}}, \ \frac{p_{2k-1}}{q_{2k-1}} > \frac{p_{2k+1}}{q_{2k+1}}$$

and

$$\left| \frac{p_k}{q_k} - \frac{p_{k+1}}{q_{k+1}} \right| = \frac{1}{q_k q_{k+1}}.$$

(v) Suppose a_j is a sequence of strictly positive integers for $j \geq 1$ and a_0 is a positive integer. Then there exists an $\alpha \in \mathbb{R}$ such that

$$\frac{p_n}{q_n} \to \alpha.$$

Further

$$\left| \frac{p_n}{q_n} - \alpha \right| \leq \frac{1}{q_n q_{n+1}}.$$

Proof. (i) Using Lemma 15.3 (ii), we have

$$\begin{aligned} p_k q_{k-1} - q_k p_{k-1} &= \det \begin{pmatrix} p_k & p_{k-1} \\ q_k & q_{k-1} \end{pmatrix} \\ &= \det \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} \\ &= \det \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \det \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \det \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} \\ &= (-1)^{k+1}. \end{aligned}$$

(ii) Either use the matricial formula

$$\begin{pmatrix} p_k & p_{k-1} \\ q_k & q_{k-1} \end{pmatrix} = \begin{pmatrix} p_{k-1} & p_{k-2} \\ q_{k-1} & q_{k-2} \end{pmatrix} \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix}$$

or direct computation.

(iii) Follows from the formula of (i).

(iv) By (ii) (or direct observation), the q_k form a strictly increasing sequence of strictly positive integers. Thus the $q_{k-1}q_k$ form a strictly increasing sequence of strictly positive integers.

The formula of (i) gives

$$\frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}} = (-1)^{k+1} \frac{1}{q_k q_{k-1}},$$

so the remark of the previous paragraph shows that

$$\frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}}$$

is an alternating sequence with decreasing magnitude. Thus

$$\frac{p_{2k}}{q_{2k}} < \frac{p_{2k-2}}{q_{2k-2}}, \quad \frac{p_{2k-1}}{q_{2k-1}} < \frac{p_{2k+1}}{q_{2k+1}}.$$

We also have

$$\left| \frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}} \right| = \frac{1}{q_k q_{k-1}} \to 0.$$

(v) A decreasing sequence bounded below tends to a limit, so

$$\frac{p_{2k+1}}{q_{2k+1}} \to \alpha$$

as $k \to \infty$ for some α. Since

$$\left| \frac{p_k}{q_k} - \frac{p_{k-1}}{q_{k-1}} \right| \to 0$$

this tells us that

$$\frac{p_{2k}}{q_{2k}} \to \alpha$$

as $k \to \infty$. Thus

$$\frac{p_n}{q_n} \to \alpha$$

and

$$\frac{p_{2k+1}}{q_{2k+1}} > \alpha > \frac{p_{2k}}{q_{2k}}.$$

□

Naturally we call α the value of the associated continued fraction.

Exercise 15.5. *Suppose we have an irrational $x \in (0,1]$ and we form a continued fraction (with $a_0 = 0$)*

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \dots}}}}$$

is the manner of Section 14. Show that

$$\frac{p_{2k}}{q_{2k}} < x < \frac{p_{2k-1}}{q_{2k-1}}$$

for all k and deduce that $x = \alpha$ where α is the value of the associated continued fraction.

Solution. Observe that if $a > 0$

$$s > t > 0 \Rightarrow \frac{1}{a+t} > \frac{1}{a+s}.$$

Thus (using a formal induction if more details are required)

$$\frac{p_{2k}}{q_{2k}} < x < \frac{p_{2k-1}}{q_{2k-1}}.$$

We know from Theorem 15.4 (v) that

$$\frac{p_n}{q_n} \to \alpha,$$

so $\alpha = x$. □

Theorem 15.6. *Continuing with the ideas and notation of Theorem 15.4, p_n/q_n is closer to α than any other fraction with denominator no larger than q_n. In other words,*

$$\left|\frac{p_n}{q_n} - \alpha\right| \leq \left|\frac{p}{q} - \alpha\right|$$

whenever p and q are integers with $1 \leq q \leq q_n$.

Proof. Observe that if q and u are positive integers with $q \leq q_n$, then

$$\left|\frac{u}{q} - \frac{p_{n+1}}{q_{n+1}}\right| \geq \frac{1}{qq_{n+1}},$$

with equality only if $q = q_n$ and, in this case, only if $u = p_n$. Thus p_n/q_n is the closest fraction of the form u/q (with $q \leq q_n$) to p_{n+1}/q_{n+1}. But α lies between p_n/q_n and p_{n+1}/q_{n+1}, so p_n/q_n is also the closest fraction of the form u/q (with $q \leq q_n$) to α. □

Theorem 15.7. *If x is irrational, we can find u_n and v_n integers with $v_n \to \infty$ such that*

$$\left|\frac{u_n}{v_n} - x\right| < \frac{1}{v_n^2}.$$

Proof. We may assume that $0 < x < 1$ without loss of generality. Using the notation of this section we observe that x lies between p_n/q_n and p_{n+1}/q_{n+1} so

$$\left|\frac{p_n}{q_n} - x\right| < \left|\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}\right| = \frac{1}{q_n q_{n+1}} < \frac{1}{q_n^2}.$$

□

This result should be compared with Theorem 12.7. We give another proof of Theorem 15.7 in Exercise 21.5. We give a slight improvement in Exercise 21.9.

Exercise 15.8. *Which earlier result tells us that, if α is the irrational root of a quadratic with integer coefficients, then there exists a C (depending on α) such that*

$$\left|\frac{u}{v} - \alpha\right| \geq \frac{C}{v^2}$$

whenever u and v are integers with $v \geq 1$?

Solution. Theorem 12.7 with $n = 2$. □

We can treat 'terminating continued fractions' and rationals in the same way.

Speaking rather vaguely, we see that the occurrence of large a_j's in a continued fraction expansion gives rise to large q_m's and associated good approximations. It is reasonable to look at the most extreme opposite case.

Exercise 15.9. *(i) Show that, if we write*

$$\sigma = \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \dots}}}},$$

then

$$\sigma = \frac{-1 + \sqrt{5}}{2}.$$

(ii) Show that, if we form p_n and q_n in the usual way for the continued fraction above, then $p_n = F_n$, $q_n = F_{n+1}$ where F_m is the mth Fibonacci number given by $F_0 = 0$, $F_1 = 1$ and

$$F_{m+1} = F_m + F_{m-1}.$$

(iii) Show that

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n.$$

Solution. (i) Observe that

$$\sigma = \frac{1}{1 + \sigma}$$

so that

$$\sigma^2 + \sigma - 1 = 0$$

and

$$\sigma = \frac{-1 \pm \sqrt{5}}{2}.$$

Since $\sigma > 0$, we must have

$$\sigma = \frac{-1 + \sqrt{5}}{2}.$$

(ii) By Theorem 15.4 (iii), $q_k = a_k q_{k-1} + q_{k-2}$ and $p_k = a_k p_{k-1} + p_{k-2}$ with $a_k = 1$. Thus

$$q_k = q_{k-1} + q_{k-2} \text{ and } p_k = p_{k-1} + p_{k-2}.$$

Now $q_0 = 1 = F_1$, $q_1 = 1 = F_2$, $p_0 = 0 = F_0$, $p_1 = 1 = F_1$, so by an inductive argument (or general knowledge of recurrence relations),

$$q_n = F_{n+1} \text{ and } p_n = F_n.$$

(iii) By Theorem 15.4 (i),

$$F_{n+1}F_{n-1} - F_n^2 = -(p_{n-1}q_n - q_{n-1}p_n) = (-1)^{n+1}.$$

□

Exercise 15.10. *Use the continued fraction expansion of σ and Theorem 15.6 to show that there exists an $m > 0$ such that*

$$\left|\frac{p}{q} - \sigma\right| > \frac{m}{q^2}$$

whenever p and q are integers with $q \geq 1$.

Exercise 21.10 gives a more general version of this idea.

Solution. By Theorem 15.6 F_n/F_{n+1} is closer to σ than any other fraction with denominator no larger than F_{n+1}. Thus

$$\begin{aligned}\left|\frac{p}{q} - \sigma\right| &\geq \max\left\{\left|\frac{F_n}{F_{n+1}} - \sigma\right|, \left|\frac{F_{n+1}}{F_{n+2}} - \sigma\right|\right\} \geq \frac{1}{2}\left(\left|\frac{F_n}{F_{n+1}} - \sigma\right| + \left|\frac{F_{n+1}}{F_{n+2}} - \sigma\right|\right)\\ &= \frac{1}{2}\left|\frac{F_n}{F_{n+1}} - \frac{F_{n+1}}{F_{n+2}}\right| = \frac{1}{2F_{n+1}F_{n+2}}\end{aligned}$$

whenever $q \leq F_n$ and so whenever $F_{n-1} \leq q \leq F_n$.

Now $F_r \leq 2F_{r-1}$ so

$$\left|\frac{p}{q} - \sigma\right| \geq \frac{1}{2F_{n+1}F_{n+2}} \geq \frac{1}{64F_{n-1}^2} \geq \frac{1}{64q^2}$$

whenever $F_{n-1} \leq q \leq F_n$ for all $n \geq 2$ and the result follows. □

Exercise 15.11. *In one of Lewis Carroll's favourite puzzles an 8×8 square is reassembled to form a 13×5 rectangle as shown in Figure 1.*

What is the connection with Exercise 15.9? Can you design the next puzzle in the sequence?

Solution. Observe that $F_5 = 5$, $F_6 = 8$, $F_7 = 13$ so the difference in areas is 1 so we only need to hide one unit of area.

Identify the Fibonacci numbers in the diagram and generalise. □

Hardy and Wright's *An Introduction to the Theory of Numbers* [5] contains a chapter on approximation by rationals in which they show, among other things, that σ is indeed particularly resistant to being so approximated by rationals. If I was asked to nominate a book to be taken away by some one leaving *professional* mathematics, but wishing to keep up an interest in the subject, this book would be my first choice.

Figure 1: Carroll's puzzle

16 A nice formula (not lectured in 2020/2021)

This section is non-examinable and will not be lectured on this year.

The notion of a continued fraction can be extended in many ways.

We are used to the idea of approximating functions f by polynomials P. Sometimes it may be more useful to approximate f by a rational function P/Q where P and Q are polynomials. If we approximate by polynomials we are led to look at Taylor series. If we approximate by rational functions it might be worth looking at some generalisation of continued fractions.

Here is a very pretty formula along these lines.

$$\tan x = \cfrac{x}{1 - \cfrac{x^2}{3 - \cfrac{x^2}{5 - \cfrac{x^2}{7 - \dots}}}}.$$

The following theorem of Lambert makes the statement precise.

Theorem 16.1. *If we write*

$$R_n(x) = \cfrac{x}{1 - \cfrac{x^2}{3 - \cfrac{x^2}{\ddots - \cfrac{x^2}{2n-3 - \cfrac{x^2}{2n-1}}}}},$$

then $R_n(x) \to \tan x$ *as* $n \to \infty$ *for all real* x *with* $|x| \leq 1$.

In order to attack this we start by generalising an earlier result

Exercise 16.2. *Suppose that* a_j *and* b_j $[j = 0, 1, 2, \dots]$ *are chosen so that we never divide by zero (for example all strictly positive). Show that if*

$$\begin{pmatrix} p_n & b_n p_{n-1} \\ q_n & b_n q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & b_n \\ 1 & 0 \end{pmatrix}.$$

then

$$\frac{p_n}{q_n} = a_0 + \cfrac{b_0}{a_1 + \cfrac{b_1}{a_2 + \cfrac{b_2}{a_3 + \cfrac{b_3}{a_4 + \cfrac{b_4}{\ddots \cfrac{}{a_{n-1} + \cfrac{b_{n-1}}{a_n}}}}}}}$$

Show that

$$\begin{aligned} p_n &= a_n p_{n-1} + b_{n-1} p_{n-2} \\ q_n &= a_n q_{n-1} + b_{n-1} q_{n-2}. \end{aligned}$$

Solution. (i) The proof follows the lines of that of Lemma 15.2.

Observe that, if s, $r \neq 0$,

$$\begin{pmatrix} r' \\ s' \end{pmatrix} = \begin{pmatrix} a & b \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix} = \begin{pmatrix} ar + bs \\ r \end{pmatrix},$$

and

$$\frac{ar + bs}{r} = a + \frac{b}{r/s}.$$

Thus, by induction, if

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & b_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ 1 \end{pmatrix},$$

then

$$\frac{p_n}{q_n} = a_0 + \cfrac{b_0}{a_1 + \cfrac{b_1}{a_2 + \cfrac{b_2}{a_3 + \cfrac{b_3}{a_4 + \cfrac{b_4}{\ddots \, \cfrac{}{a_{n-1} + \cfrac{b_{n-1}}{a_n}}}}}}}.$$

Observe that

$$\begin{aligned} &\begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & b_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &\qquad = \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-2} & b_{n-2} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ 1 \end{pmatrix} = \begin{pmatrix} p_{n-1} \\ q_{n-1} \end{pmatrix} \end{aligned}$$

and thus

$$\begin{pmatrix} p_n & b_n p_{n-1} \\ q_n & b_n q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_n & b_n \\ 1 & 0 \end{pmatrix}.$$

We deduce that

$$\begin{pmatrix} p_n & b_n p_{n-1} \\ q_n & b_n q_{n-1} \end{pmatrix} = \begin{pmatrix} p_{n-1} & b_{n-1} p_{n-2} \\ q_{n-1} & b_{n-1} q_{n-2} \end{pmatrix} \begin{pmatrix} a_n & b_n \\ 1 & 0 \end{pmatrix}.$$

Looking at the first column gives

$$\begin{aligned} p_n &= a_n p_{n-1} + b_{n-1} p_{n-2}, \\ q_n &= a_n q_{n-1} + b_{n-1} q_{n-2}, \end{aligned}$$

as required. □

We now use the following result which is clearly related to the manipulations used in Lemma 12.3.

Lemma 16.3. *Let us write*

$$S_n(x) = \frac{1}{2^n n!} \int_0^x (x^2 - t^2)^n \cos t \, dt.$$

Then $S_n(x) = q_n(x) \sin x - p_n(x) \cos x$ *where* p_n *and* q_n *satisfy the recurrence relations*

$$\begin{aligned} p_n(x) &= (2n-1)p_{n-1}(x) - x^2 p_{n-2}(x), \\ q_n(x) &= (2n-1)q_{n-1}(x) - x^2 q_{n-2}(x) \end{aligned}$$

for $n \geq 2$ *and* $p_0(x) = 0$, $q_0(x) = 1$, $p_1(x) = x$, $q_1(x) = 1$.

Proof. We have

$$S_0(x) = \int_0^x \cos t \, dt = \sin x,$$

so $p_0(x) = 0$, $q_0(x) = 1$. Integration by parts gives

$$\begin{aligned} S_1(x) &= \frac{1}{2} \int_0^x (x^2 - t^2) \cos t \, dt \\ &= \left[\frac{1}{2}(x^2 - t^2) \sin t \right]_0^x + \int_0^x t \sin t \, dt = \int_0^x t \sin t \, dt \\ &= [-t \cos t]_0^x + \int_0^x t \cos t \, dt = -x \cos x + \sin x \end{aligned}$$

so $p_1(x) = x$, $q_1(x) = 1$.

If $n \geq 2$, a similar repeated integration by parts gives

$$\begin{aligned}
S_n(x) &= \frac{1}{2^n n!} \int_0^x (x^2 - t^2)^n \cos t \, dt \\
&= \left[\frac{1}{2^n n!} (x^2 - t^2)^n \sin t \right]_0^x + \frac{1}{2^{n-1}(n-1)!} \int_0^x t(x^2 - t^2)^{n-1} \sin t \, dt \\
&= \frac{1}{2^{n-1}(n-1)!} \int_0^x t(x^2 - t^2)^{n-1} \sin t \, dt \\
&= \frac{1}{2^{n-1}(n-1)!} \left[-t(x^2 - t^2)^{n-1} \cos t \right]_0^x \\
&\quad + \frac{1}{2^{n-1}(n-1)!} \int_0^x \left((x^2 - t^2)^{n-1} - 2(n-1)t^2(x^2 - t^2)^{n-2} \right) \cos t \, dt \\
&= S_{n-1}(x) - \frac{1}{2^{n-2}(n-2)!} \int_0^x \left(t^2(x^2 - t^2)^{n-2} \right) \cos t \, dt \\
&= S_{n-1}(x) + \frac{1}{2^{n-2}(n-2)!} \int_0^x \left((x^2 - t^2)t^2(x^2 - t^2)^{n-2} \right) \cos t \, dt \\
&\quad - x^2 \int_0^x (x^2 - t^2)^{n-2} \cos t \, dt \\
&= (2n-1)S_{n-1}(x) - x^2 S_{n-2}(x).
\end{aligned}$$

Thus

$$\begin{aligned}
p_n(x) &= (2n-1)p_{n-1}(x) - x^2 p_{n-2}(x), \\
q_n(x) &= (2n-1)q_{n-1}(x) - x^2 q_{n-2}(x)
\end{aligned}$$

and we are done. □

Theorem 16.1. Since

$$S_n(x) = q_n(x) \sin x - p_n(x) \cos x,$$

rearrangement gives

$$\tan x = \frac{p_n(x)}{q_n(x)} + \frac{S_n(x)}{q_n(x) \cos x}$$

so we need to show that

$$\frac{S_n(x)}{q_n(x) \cos x} \to 0.$$

It is easy to see that

$$\begin{aligned}
|S_n(x)| &\leq \left| \frac{1}{2^n n!} \int_0^x (x^2 - t^2)^n \cos t \, dt \right| \\
&\leq \frac{1}{2^n n!} |x||x|^{2n} \to 0
\end{aligned}$$

as $n \to \infty$ for all x.

We shall show that if $|x| \leq 1$ then $q_n(x) \to \infty$. (Actually it can be shown that $|q_n(x)| \to \infty$ for all x.) Observe that

$$q_n(x) = (2n-1)q_{n-1}(x) - x^2 q_{n-2}(x)$$

so, if $|x| \leq 1$,

$$q_n(x) \geq (2n-1)q_{n-1}(x) - q_{n-2}(x) \geq 3q_{n-1}(x) - q_{n-2}(x)$$

for $n \geq 2$. Since $q_0(x) = q_1(x) = 1$ we have $q_2(x) \geq 2$ and a simple induction gives $q_n(x) \geq 2^{n-1}$ for $n \geq 1$ and $|x| \leq 1$.

Notice the rapidity of convergence of the continued fraction in this case. □

The results of this section and other interesting topics are discussed in a book [3] which is a model of how a high level recreational mathematics text should be put together.

17 Winding numbers

We all know that complex analysis has a lot to say about 'the number of times a curves goes round a point'. In this final section we make the notion precise.

Theorem 17.1. *Let*

$$\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}.$$

If $g : [0,1] \to \mathbb{T}$ is continuous with $g(0) = e^{i\theta_0}$, then there is a unique continuous function $\theta : [0,1] \to \mathbb{R}$ with $\theta(0) = \theta_0$ such that $g(t) = e^{i\theta(t)}$ for all $t \in [0,1]$.

Proof. We prove existence, leaving uniqueness to follow from Exercise 17.2. Let E be the set of $u \in [0,1]$ for which there exists a continuous function $\theta : [0,u] \to \mathbb{R}$ with $\theta(0) = \theta_0$ such that $g(t) = e^{i\theta(t)}$ for all $t \in [0,u]$. Since $0 \in E$ (just take $\theta(0) = \theta_0$) and E is bounded, E must have an upper bound w.

Suppose that $w \in (0,1)$. Since g is continuous, we can find a $\delta > 0$ such that $(w-2\delta, w+2\delta) \subseteq [0,1]$ and $|g(t) - g(w)| < 1/2$ for $t \in (w-2\delta, w+2\delta)$ We know that there is a unique continuous function

$$\phi : \{z : |z-1| < 1/2, \ |z| = 1\} \to [-\pi/2, \pi/2]$$

such that

$$z = e^{i\phi(z)} \text{ for all } |z-1| < 1/2,\ |z| = 1.$$

Thus, if we choose θ_1 such that $e^{i\theta_1} = g(w)$ and define $\tilde{\theta} : (w-2\delta, w+2\delta) \to [-\pi/2, \pi/2]$ by

$$\tilde{\theta}(t) = \theta_1 + \phi\big(g(t)/g(w)\big)$$

we will have $\tilde{\theta}$ continuous and

$$g(t) = e^{i\tilde{\theta}(t)}$$

for $t \in (w-2\delta, w+2\delta)$.

By the definition of an upper bound, we can find $u \in (w-\delta, w]$ and a continuous function $\psi : [0,u] \to \mathbb{R}$ with $\psi(0) = \theta_0$ such that $g(t) = e^{i\psi(t)}$ for all $t \in [0,u]$. Since

$$e^{i\tilde{\theta}(u)} = g(u) = e^{i\psi(u)}$$

we must have $\tilde{\theta}(u) = \psi(u) + 2N\pi$ for some integer N. Taking

$$\theta(t) = \begin{cases} \psi(t) & \text{for } t \in [0,u] \\ \tilde{\theta}(t) - 2N\pi & \text{for } t \in [u, u+\delta] \end{cases}$$

we see that $\theta : [0, u+\delta] \to \mathbb{R}$ is a continuous function with $\theta(0) = \theta_0$ such that $g(t) = e^{i\theta(t)}$ for all $t \in [0, u+\delta]$. Thus $u+\delta \in E$ and $u+\delta > w$, contradicting our assumption that u is an upper bound.

A similar argument show that 0 is not an upper bound. Thus $\sup E = 1$ and much the same argument as above shows that $1 \in E$, so we are done. □

The uniqueness part of Theorem 17.1 follows from the next exercise.

Exercise 17.2. *Suppose ψ, $\phi : [0,1] \to \mathbb{R}$ are continuous with $e^{i\psi(t)} = e^{i\phi(t)}$ for all $t \in [0,1]$. Show that there exists an integer n such that $\psi(t) = \phi(t) + 2n\pi$ for all $t \in [0,1]$.*

Solution. Observe that $f : [0,1] \to \mathbb{R}$ defined by

$$f(t) = \frac{\psi(t) - \phi(t)}{2\pi}$$

is an integer valued continuous function on $[0,1]$ and so must be constant. (Or quote the intermediate value theorem directly.) □

Corollary 17.3. *If $\gamma : [0,1] \to \mathbb{C} \setminus \{0\}$ is continuous with $\gamma(0) = |\gamma(0)|e^{i\theta_0}$, then there is a unique continuous function $\theta : [0,1] \to \mathbb{R}$ with $\theta(0) = \theta_0$ such that $\gamma(t) = |\gamma(t)|e^{i\theta(t)}$.*

Proof. Set $g(t) = \gamma(t)/|\gamma(t)|$ and use Theorem 17.1. □

Definition 17.4. *If $\gamma : [0,1] \to \mathbb{C} \setminus \{0\}$ and $\theta : [0,1] \to \mathbb{R}$ are continuous with $\gamma(t) = |\gamma(t)|e^{i\theta(t)}$ then we define*

$$w(\gamma, 0) = \frac{\theta(1) - \theta(0)}{2\pi}.$$

Exercise 17.2 shows that $w(\gamma, 0)$ does not depend on the choice of θ.

Exercise 17.5. *(i) If $\gamma : [0,1] \to \mathbb{C} \setminus \{0\}$ is continuous and $\gamma(0) = \gamma(1)$ (that is to say, the path is* closed*) show that $w(\gamma, 0)$ is an integer.*

(ii) Give an example to show that, under the conditions of (i), $w(\gamma, 0)$ can take any integer value.

Solution. (i) Since $\gamma(0) \neq 0$ and

$$|\gamma(0)|e^{i\theta(0)} = \gamma(0) = \gamma(1) = |\gamma(1)|e^{i\theta(1)} = |\gamma(0)|e^{i\theta(1)},$$

we have $e^{i(\theta(1)-\theta(0))} = 1$, so $\theta(1) - \theta(0)$ is an integer multiple of 2π.

(ii) If we take $\gamma(t) = \exp(irt)$, we get $w(\gamma, 0) = r$ [$r \in \mathbb{Z}$] □

We are only interested in the winding number of *closed* curves.

If $a \in \mathbb{C}$, it is natural to define the winding number round a of a curve given by a continuous map

$$\gamma : [0,1] \to \mathbb{C} \setminus \{a\}$$

to be

$$w(\gamma, a) = w(\gamma - a, 0),$$

but we shall not use this slight extension.

Lemma 17.6. *If γ_1, $\gamma_2 : [0,1] \to \mathbb{C} \setminus \{0\}$ then the product $\gamma_1\gamma_2$ satisfies*

$$w(\gamma_1\gamma_2, 0) = w(\gamma_1, 0) + w(\gamma_2, 0).$$

Proof. By Corollary 17.3, we can write

$$\gamma_j(t) = |\gamma_j(t)| \exp\big(i\theta_j(t)\big)$$

with $\theta_j : [0,1] \to \mathbb{R}$ continuous. We now have

$$\begin{aligned}\gamma_1(t)\gamma_2(t) &= |\gamma_1(t)| \exp\big(i\theta_1(t)\big)|\gamma_2(t)| \exp\big(i\theta_2(t)\big)\\ &= |\gamma_1(t)\gamma_2(t)| \exp\big(i(\theta_1(t) + \theta_2(t))\big),\end{aligned}$$

so that

$$\begin{aligned}w(\gamma_1\gamma_2,0)&=\frac{1}{2\pi}\Big(\big(\theta_1(1)+\theta_2(1)\big)-\big(\theta_1(0)+\theta_2(0)\big)\Big)\\&=\frac{1}{2\pi}\big(\theta_1(1)-\theta_1(0)\big)+\frac{1}{2\pi}\big(\theta_2(1)-\theta_2(0)\big)=w(\gamma_1,0)+w(\gamma_2,0).\end{aligned}$$

□

Lemma 17.7. [Dog walking lemma] *If γ_1, $\gamma_2 : [0,1]\to\mathbb{C}\setminus\{0\}$ are continuous, $\gamma_1(0)=\gamma_1(1)$, $\gamma_2(0)=\gamma_2(1)$ and*

$$|\gamma_2(t)|<|\gamma_1(t)|$$

for all $t\in[0,1]$, then $\gamma_1+\gamma_2$ never takes the value 0 and $w(\gamma_1+\gamma_2,0)=w(\gamma_1,0)$.

Proof. This argument may be familiar from 1B complex variable.

Write $\gamma(t)=\big(1+\gamma_2(t)/\gamma_1(t)\big)$. By Lemma 17.6,

$$w(\gamma_1+\gamma_2,0)=w(\gamma_1\gamma,0)=w(\gamma_1,0)+w(\gamma,0),$$

so it suffices to prove that $w(\gamma,0)=0$. We shall do this by noting that $|\gamma_2(t)/\gamma_1(t)|<1$ and so

$$\Re\gamma(t)>0$$

for all $t\in[0,1]$.

By Corollary 17.3, we can write

$$\gamma(t)=|\gamma(t)|\exp\big(i\theta(t)\big)$$

with $\theta:[0,1]\to\mathbb{R}$ continuous and $\theta(0)\in(-\pi/2,\pi/2)$. If $|\theta(t)|\geq\pi/2$ for any $t\in[0,1]$, the intermediate value theorem tells us that there is an $s\in[0,t]$ such that $|\theta(s)|=\pi/2$ and so $\Re\gamma(s)=0$, which is impossible. Thus $|\theta(t)|<\pi/2$ for all $t\in[0,1]$.

In particular $|\theta(0)|,|\theta(1)|<\pi/2$, so $|\theta(1)-\theta(0)|<\pi$. It follows that $w(\gamma,0)$ is an integer with $|w(\gamma,0)|<1/2$ and so $w(\gamma,0)=0$. □

Many interesting results in 'applied complex analysis' are obtained by 'deforming contours'. The idea of 'continuously deforming curves' can be made precise in a rather clever manner.

Definition 17.8. *Suppose that γ_0, γ_1 are closed paths not passing through 0 (so we have, $\gamma_j:[0,1]\to\mathbb{C}\setminus\{0\}$). Then we say that γ_0 is* homotopic *to γ_1 by* *closed curves not passing through zero if we can find a continuous function* $\Gamma : [0,1]^2 \to \mathbb{C} \setminus \{0\}$ *such that*

$$\begin{aligned} \Gamma(s,0) &= \Gamma(s,1) && \textit{for all } s \in [0,1], \\ \Gamma(0,t) &= \gamma_0(t) && \textit{for all } t \in [0,1], \\ \Gamma(1,t) &= \gamma_1(t) && \textit{for all } t \in [0,1]. \end{aligned}$$

We often write $\gamma_s(t) = \Gamma(s,t)$.

Exercise 17.9. *If* γ_0 *and* γ_1 *satisfy the conditions of Definition 17.8, we write* $\gamma_0 \simeq \gamma_1$. *Show that* $\simeq$ *is an equivalence relation on closed curves not passing through zero.*

Solution. Setting

$$\Gamma(s,t) = \gamma(t),$$

we see that $\gamma \simeq \gamma$.

If $\gamma_0 \simeq \gamma_1$, then we can find a continuous function $\Gamma : [0,1]^2 \to \mathbb{C} \setminus \{0\}$ such that

$$\begin{aligned} \Gamma(s,0) &= \Gamma(s,1) && \text{for all } s \in [0,1], \\ \Gamma(0,t) &= \gamma_0(t) && \text{for all } t \in [0,1], \\ \Gamma(1,t) &= \gamma_1(t) && \text{for all } t \in [0,1]. \end{aligned}$$

If we set $\tilde{\Gamma}(s,t) = \Gamma(1-s,t)$, then $\tilde{\Gamma} : [0,1]^2 \to \mathbb{C} \setminus \{0\}$ is a continuous function such that

$$\begin{aligned} \tilde{\Gamma}(s,0) &= \tilde{\Gamma}(s,1) && \text{for all } s \in [0,1], \\ \tilde{\Gamma}(0,t) &= \gamma_1(t) && \text{for all } t \in [0,1], \\ \tilde{\Gamma}(1,t) &= \gamma_0(t) && \text{for all } t \in [0,1], \end{aligned}$$

and so $\gamma_1 \simeq \gamma_0$.

If $\gamma_0 \simeq \gamma_1$ and $\gamma_1 \simeq \gamma_2$, then we can find a continuous functions $\Gamma_j : [0,1]^2 \to \mathbb{C} \setminus \{0\}$ such that

$$\begin{aligned} \Gamma_j(s,0) &= \Gamma_j(s,1) && \text{for all } s \in [0,1], \\ \Gamma_j(0,t) &= \gamma_{0+j}(t) && \text{for all } t \in [0,1], \\ \Gamma_j(1,t) &= \gamma_{1+j}(t) && \text{for all } t \in [0,1] \end{aligned}$$

for $j = 0,\ 1$.

If we set

$$\Gamma(s,t) = \begin{cases} \Gamma_0(2s,t) & \text{for all } s \in [0,1/2],\ t \in [0,1] \\ \Gamma_1(2s-1,t) & \text{for all } s \in (1/2,1],\ t \in [0,1] \end{cases}$$

then $\Gamma : [0,1]^2 \to \mathbb{C} \setminus \{0\}$ is a continuous function such that

$$\begin{aligned} \tilde{\Gamma}(s,0) &= \tilde{\Gamma}(s,1) && \text{for all } s \in [0,1], \\ \tilde{\Gamma}(0,t) &= \gamma_0(t) && \text{for all } t \in [0,1], \\ \tilde{\Gamma}(1,t) &= \gamma_2(t) && \text{for all } t \in [0,1], \end{aligned}$$

and so $\gamma_0 \simeq \gamma_2$. □

The proof of the next theorem illustrates the utility of Definition 17.8. The proof itself is sometimes referred to as 'dog walking along a canal'.

Theorem 17.10. *If γ_0 and γ_1 satisfy the conditions of Definition 17.8, then $w(\gamma_0, 0) = w(\gamma_1, 0)$.*

Proof. Let Γ be as in Definition 17.8. The map $(s,t) \mapsto |\Gamma(s,t)|$ is continuous so, by compactness, $|\Gamma(s,t)|$ attains a minimum m on the compact set $[0,1]^2$. Since Γ is never zero, we must have $m > 0$.

By compactness (see Theorem 7.7 if necessary), Γ is uniformly continuous and so we can find a strictly positive integer N such that

$$|s - s'|, |t - t'| < 2/N \Rightarrow |\Gamma(s,t) - \Gamma(s',t')| < m/2.$$

If $0 \le r \le N$ let us define

$$\beta_r(t) = \Gamma(r/N, t)$$

for $t \in [0,1]$. We observe that

$$\begin{aligned} |\beta_r(t)| &= |\Gamma(r/N, t)| \ge m > m/2 \\ &> \big|\Gamma\big(r/N, t\big) - \Gamma\big((r+1)/N, t\big)\big| = |\beta_r(t) - \beta_{r+1}(t)| \end{aligned}$$

for all $t \in [0,1]$, so by the dog walking lemma (Lemma 17.7),

$$w(\beta_r, 0) = w(\beta_{r+1}, 0)$$

for all $0 \le r \le N-1$. It follows that

$$w(\gamma_0, 0) = w(\beta_0, 0) = w(\beta_N, 0) = w(\gamma_1, 0).$$

□

As before, let us write

$$\begin{aligned} \bar{D} &= \{z \in \mathbb{C} : |z| \le 1\}, \\ D &= \{z \in \mathbb{C} : |z| < 1\}, \\ \partial D &= \{z \in \mathbb{C} : |z| = 1\}. \end{aligned}$$

Corollary 17.11. *Suppose $f : \bar{D} \to \mathbb{C}$ is continuous, $f(z) \neq 0$ for $z \in \partial D$, and we define $\gamma : [0,1] \to \mathbb{C}$ by*

$$\gamma(t) = f(e^{2\pi i t})$$

for all $t \in [0,1]$. If $w(\gamma, 0) \neq 0$, then there must exist a $z \in D$ with $f(z) = 0$.

Corollary 17.11. Suppose, if possible, that $f(z) \neq 0$ for $z \in D$. The nowhere-zero function $G : [0,1]^2 \mapsto \mathbb{C}$ given by

$$G(s,t) = f(se^{2\pi i t})$$

is continuous with $G(s,0) = G(s,1)$ for all $s \in [0,1]$, $G(1,t) = \gamma(t)$ and $G(0,t) = \gamma_0(t)$ where $\gamma_0(t) = f(0)$ for all $t \in [0,1]$. Thus γ and γ_0 are homotopic closed curves not passing through 0. By Theorem 17.10, $w(\gamma, 0) = w(\gamma_0, 0) = 0$ contradicting our hypothesis. □

This gives us another proof of the Fundamental Theorem of Algebra (Theorem 2.9).

Corollary 17.12. *If we work in the complex numbers, every non-trivial polynomial has a root.*

Proof. It is sufficient to consider polynomials P of the form

$$P(z) = z^n + Q(z)$$

with $Q(z) = \sum_{j=0}^{n-1} a_j z^j$. If we set $R = 1 + \sum_{j=0}^{n-1} |a_j|$ and consider $p(z) = R^{-n} p(Rz)$ we see that P has a root if p has root and that

$$p(z) = z^n + q(z)$$

with $|q(z)| < 1$ when $|z| = 1$.

By the dog walking lemma, the map $t \mapsto p(e^{2\pi i t})$ for $t \in [0,1]$ has the same winding number as $t \mapsto (e^{2\pi i t})^n = e^{2\pi i n t}$, that is to say, n. By Corollary 17.11, there must exist a $z \in D$ with $p(z) = 0$, so we are done. □

We also obtain a second proof of Brouwer's theorem in two dimensions in the 'no retraction' form of Theorem 4.5.)

Corollary 17.13. *There does not exist a continuous function $f : \bar{D} \to \partial D$ with $f(z) = z$ for all $z \in \partial D$.*

Proof. Suppose such a function existed. The continuous map $G : [0,1]^2 \to \partial D$ given by

$$G(s,t) = f(se^{2\pi it})$$

gives a homotopy between γ_0 defined by $\gamma_0(t) = f(0)$ and γ_1 defined by $\gamma_1(t) = f(t) = e^{2\pi t}$ using closed curves not passing through 0. By Theorem 17.10, this gives

$$1 = w(\gamma_1, 0) = w(\gamma_0, 0) = 0,$$

which is absurd.

The required result follows by reductio ad absurdum. □

The earlier combinatorial proof that we gave requires less technology to extend to higher dimensions.

The contents of this section show that parts of complex analysis are really just special cases of general 'topological theorems'. On the other hand, other parts (such as Taylor's theorem and Cauchy's theorem itself) depend crucially on the the fact that we are dealing with the very restricted class of functions which satisfy the Cauchy-Riemann equations.

In traditional courses on complex analysis, this fact appears, if it appears at all, rather late in the day. Beardon's *Complex Analysis* [4] shows that it is possible to do things differently and is well worth a look[18].

References

[1] C. Adams, *Zombies and Calculus*, Princeton University Press, 2014. (I have not read this, but according to one reviewer, the book 'shows how calculus can be used to understand many different real-world phenomena.')

[2] A. Baker, *Transcendental number theory*, CUP, 1975.

[3] K. Ball, *Strange Curves, Counting Rabbits and Other Mathematical Explorations*, Princeton University Press, 2003.

[4] A. F. Beardon *Complex Analysis*, Wiley, 1979.

[5] G. H. Hardy and E. M. Wright *An Introduction to the Theory of Numbers* OUP, 1937.

[18] Though, like other good texts, it will not please those 'who want from books, plain cooking made still plainer by plain cooks.'

[6] L. C. Thomas *Games, Theory and Applications*, Dover reprint of a book first published by Wiley in 1984.

‘You are old, Father William’ the young man said,
And your hair has become very white;
And yet you incessantly stand on your head
Do you think, at your age, it is right?’

‘In my youth’ Father William replied to his son,
‘I feared it might injure the brain;
But, now that I’m perfectly sure I have none,
Why, I do it again and again.’

‘You are old’ said the youth ‘as I mentioned before,
And have grown most uncommonly fat;
Yet you turned a back-somersault in at the door
Pray, what is the reason of that?’

‘In my youth’ said the sage, as he shook his grey locks,
‘I kept all my limbs very supple
By the use of this ointment one shilling the box
Allow me to sell you a couple?’

‘You are old’ said the youth ‘and your jaws are too weak
For anything tougher than suet;
Yet you finished the goose, with the bones and the beak
Pray how did you manage to do it?’

‘In my youth’ said his father ‘I took to the law,
And argued each case with my wife;
And the muscular strength, which it gave to my jaw,
Has lasted the rest of my life.’

‘You are old’ said the youth ‘one would hardly suppose
That your eye was as steady as ever;
Yet you balanced an eel on the end of your nose
What made you so awfully clever?’

‘I have answered three questions, and that is enough’
Said his father; ‘don’t give yourself airs!
Do you think I can listen all day to such stuff?
Be off, or I’ll kick you down stairs!’

Lewis Carroll

18 Question sheet 1

Note The Pro-Vice-Chancellor for Education has determined that the amount of work required by a student to understand a course should be the same as that required by any other student. Traditionally, teachers and students in the Faculty of Mathematics have interpreted this as meaning that each example sheet should have exactly twelve questions. The number of questions in the example sheets for this course may be reduced to twelve by omitting any marked with a ★.

Exercise 18.1. (i) Consider $\mathbb{R}^n$. If we take d to be ordinary Euclidean distance

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = \left(\sum_{j=1}^{n} |x_j - y_j|^2\right)^{1/2},$$

show that $(\mathbb{R}^n, d)$ is a metric space.
[Hint: Use inner products.]

(ii) Consider $\mathbb{C}$. If we take $d(z, w) = |z - w|$, show that $(\mathbb{C}, d)$ is a metric space.

(iii) Let X be a non-empty set. If we write

$$d(x, y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y, \end{cases}$$

show that (X, d) is a metric space (d is called the discrete metric).

(iv) Consider $X = \{0, 1\}^n$. If we take

$$d(\mathbf{x}, \mathbf{y}) = \sum_{j=1}^{n} |x_j - y_j|$$

then (X, d) is a metric space.

Exercise 18.2. (i) Give an example of a continuous bijection $f : \mathbb{R} \to \mathbb{R}$ with no fixed points.

(ii) If A is an infinite countable set, show that there exists a bijection $f : A \to A$ with no fixed points. What can you say if A is finite non-empty set? (Be careful to cover all possible cases.)

[Remark: We can replace 'infinite countable' by 'infinite' provided we accept the appropriate set theoretic axioms.]

Exercise 18.3. (i) Show that a subset E of $\mathbb{R}^m$ (with the usual metric) is compact if every continuous function $f : E \to \mathbb{R}$ is bounded.

(ii) Show that a subset E of $\mathbb{R}^m$ (with the usual metric) is compact if every bounded continuous function $f : E \to \mathbb{R}$ attains its bounds.

Exercise 18.4. Suppose that d is the usual metric on $\mathbb{R}^m$, X is a compact set in $\mathbb{R}^m$ and $f : X \to X$ is a continuous *distance increasing* map. In other words,

$$d(f(x), f(y)) \geq d(x, y)$$

for all $x, y \in X$. The object of this question is to show that f must be a surjection.

(i) Let $f^0(x) = x$ $f^n(x) = f(f^{n-1}(x))$. Explain why $X_\infty = \bigcap_{n=0}^\infty f^n(X)$ is compact and why the map $g : X_\infty \to X_\infty$ is well defined by $f\big(g(y)\big) = y$.

(ii) If $z \in X$, consider the sequence $f^n(z)$. By using compactness and part (i) show that given $\epsilon > 0$ we can find an $w \in X_\infty$ such that $d(w, z) < \epsilon$. Deduce that f is surjective.

(iii) Let $X = \{x \in \mathbb{R} : x \geq 0\}$. Find a continuous function $f : X \to X$ such that $|f(x) - f(y)| \geq 2|x - y|$ for all $x, y \in X$ but $f(X) \neq X$. Why does this not contradict (ii)?

(iv) We work in $\mathbb{C}$. Let α be irrational and let $\omega = \exp(2\pi\alpha i)$. If

$$X = \{\omega^n : n \geq 1\},$$

find a continuous function $f : X \to X$ such that $|f(w) - f(z)| = |w - z|$ for all $w, z \in X$ but $f(X) \neq X$. Why does this not contradict (ii)?

Exercise 18.5. A function $f : \mathbb{R}^n \to \mathbb{R}$ is called upper semi-continuous if, given $\mathbf{x} \in \mathbb{R}^n$ and $\epsilon > 0$, we can find a $\delta > 0$ such that

$$\|\mathbf{x} - \mathbf{y}\| < \delta \Rightarrow f(\mathbf{y}) \leq f(\mathbf{x}) + \epsilon.$$

(i) If E is a subset of $\mathbb{R}^n$ and we define the indicator function $\mathbb{I}_E$ by

$$\mathbb{I}_E(\mathbf{x}) = \begin{cases} 1 & \text{if } \mathbf{x} \in E, \\ 0 & \text{otherwise,} \end{cases}$$

show that $\mathbb{I}_E$ is upper semi-continuous if and only if E is closed.

(ii) State and prove necessary and sufficient conditions for $-\mathbb{I}_E$ to be upper semi-continuous.

(iii) If $f : \mathbb{R}^n \to \mathbb{R}$ is upper semi-continuous and K is compact show that there exists a $\mathbf{z} \in K$ such that $f(\mathbf{z}) \geq f(\mathbf{k})$ for all $\mathbf{k} \in K$. (In other words f attains a maximum on K.)

(iv) If $g : \mathbb{R} \to \mathbb{R}$ is defined by

$$g(x) = \begin{cases} -1/|x| & \text{if } x \neq 0, \\ 0 & \text{if } x = 0, \end{cases}$$

show that g is upper semi-continuous but that g is unbounded on $[-1, 1]$.

Exercise 18.6. Let

$$\begin{aligned}\mathbb{H} &= \{z \in \mathbb{C} : \Re z > 0\},\\ \bar{\mathbb{H}} &= \{z \in \mathbb{C} : \Re z \geq 0\},\\ \mathbb{D} &= \{z \in \mathbb{C} : |z| < 1\},\\ \bar{\mathbb{D}} &= \{z \in \mathbb{C} : |z| \leq 1\}.\end{aligned}$$

(i) Does every bijective continuous map $f : \mathbb{C} \to \mathbb{C}$ have a fixed point?
(ii) Does every bijective continuous map $f : \mathbb{H} \to \mathbb{H}$ have a fixed point?
(iii) Does every bijective continuous map $f : \bar{\mathbb{H}} \to \bar{\mathbb{H}}$ have a fixed point?
(iv) Does every bijective continuous map $f : \mathbb{D} \to \mathbb{D}$ have a fixed point?
(v) Does every bijective continuous map $f : \bar{\mathbb{D}} \to \bar{\mathbb{D}}$ have a fixed point?
Give reasons for your answers.

Exercise 18.7. (Exercise 4.9) Show that the following four statements are equivalent.

(i) If $f : [0, 1] \to [0, 1]$ is continuous, then we can find a $w \in [0, 1]$ such that $f(w) = w$.

(ii) There does not exist a continuous function $g : [0, 1] \to \{0, 1\}$ with $g(0) = 0$ and $g(1) = 1$. (In topology courses we say that $[0, 1]$ is connected.)

(iii) If A and B are closed subsets of $[0, 1]$ with $0 \in A$, $1 \in B$ and $A \cup B = [0, 1]$ then $A \cap B \neq \varnothing$.

(iv) If $h : [0, 1] \to \mathbb{R}$ is continuous and $h(0) \leq c \leq h(1)$, then we can find a $y \in [0, 1]$ such that $h(y) = c$.

Exercise 18.8. (Exercise 4.10) Suppose that we colour the points r/n red or blue $[r = 0, 1, \dots, n]$ with 0 red and 1 blue. Show that there are a pair of neighbouring points u/n, $(u + 1)/n$ of different colours. Use this result to prove statement (iii) of Exercise 4.9.

Exercise 18.9. Suppose that $g : \mathbb{R}^2 \to \mathbb{R}^2$ is a continuous function such that there exists a $K > 0$ with $\|g(\mathbf{x}) - \mathbf{x}\| \leq K$.

(i) By constructing a function $f : \mathbb{R}^2 \to \mathbb{R}^2$, taking a disc into itself, and such that

$$f(\mathbf{t}) = \mathbf{t} \Rightarrow g(\mathbf{t}) = \mathbf{0}$$

show that $\mathbf{0}$ lies in the image of g.

(ii) Show that, in fact, g is surjective.

(iii) Is it necessarily true that g has a fixed point? Give reasons.

(iv) Is g necessarily injective? Give reasons.

Exercise 18.10. Use the Brouwer fixed point theorem to show that there is a complex number z with $|z| \leq 1$ and

$$z^4 - z^3 + 8z^2 + 11z + 1 = 0.$$

Exercise 18.11. Consider the square $S = [-1,1]^2$. Suppose that $\beta, \gamma : [-1,1] \to S$ are continuous with $\beta(-1) = (-1,-1)$, $\beta(1) = (1,1)$, $\gamma(-1) = (-1,1)$, $\gamma(1) = (1,-1)$. The object of this question is to show that there exist $(s_0,t_0) \in [-1,1]^2$ such that $\beta(s_0) = \gamma(t_0)$. (Note that this is just a formal version of Exercise 4.13.)

Our proof will be by contradiction, so assume that no such (s_0,t_0) exists. We write

$$\beta(s) = \big(\beta_1(s), \beta_2(s)\big),\ \gamma(t) = \big(\gamma_1(t), \gamma_2(t)\big)$$

(i) Show carefully that the function $F : S \to S$ given by

$$F(s,t) = \frac{-1}{\max\{|\beta_1(s) - \gamma_1(t)|, |\beta_2(s) - \gamma_2(t)|\}}\big(\beta_1(s) - \gamma_1(t), \beta_2(s) - \gamma_2(t)\big)$$

is well defined and continuous.

(ii) Show by considering the possible values of the fixed points of F or otherwise that F has no fixed points, Brouwer's fixed point theorem now gives a contradiction.

Exercise 18.12. Here is a variation on Lemma 4.7 (ii). It can be proved in the same way.

Suppose that T, I, J, K are as in Lemma 4.7 and that A, B and C are closed subsets of T with

$$A \cup B \cup C = T,$$
$$A \cup B \supseteq I,\ B \cup C \supseteq J,\ C \cup A \supseteq K,$$
$$A \supseteq K \cap I,\ B \supseteq I \cap J,\ C \supseteq J \cap K.$$

Show that $A \cap B \cap C \neq \varnothing$.

Exercise 18.13. ★ Cantor started the researches which led him to his studies of infinite sets by looking at work of Riemann on trigonometric series. He needed to show that if $F : [a,b] \to \mathbb{R}$ is continuous and

$$\frac{F(x+h) - 2F(x) + F(x-h)}{h^2} \to 0$$

as $h \to 0$ for all $x \in (a,b)$ then F is linear. (Note that there are no differentiability conditions on F.) Schwarz was able to supply a proof.

(i) Suppose that $F : [a,b] \to \mathbb{R}$ is continuous, $F(a) = F(b)$ and there exists an $\epsilon > 0$ such that

$$\limsup_{h\to 0} \frac{F(x+h) - 2F(x) + F(x-h)}{h^2} \geq \epsilon$$

for all $x \in (a,b)$. Show that F cannot attain a maximum at any $x \in (a,b)$. Deduce that

$$F(x) \leq F(a)$$

for all $x \in [a,b]$.

(ii) Suppose that $F : [a,b] \to \mathbb{R}$ is continuous, $F(a) = F(b)$ and

$$\limsup_{h \to 0} \frac{F(x+h) - 2F(x) + F(x-h)}{h^2} \geq 0$$

for all $x \in (a,b)$. Let $c = (a+b)/2$. By considering $G(x) = F(x) + \epsilon(x-c)^2/4$ or otherwise, show that

$$F(x) \leq F(a)$$

for all $x \in [a,b]$.

(iii) Suppose that $F : [a,b] \to \mathbb{R}$ is continuous, $F(a) = F(b)$ and

$$\lim_{h \to 0} \frac{F(x+h) - 2F(x) + F(x-h)}{h^2} = 0$$

for all $x \in (a,b)$. By considering F and $-F$ show that

$$F(x) = F(a)$$

for all $x \in [a,b]$.

(iv) Show that if $F : [a,b] \to \mathbb{R}$ is continuous and

$$\frac{F(x+h) - 2F(x) + F(x-h)}{h} \to 0$$

as $h \to 0$ for all $x \in (a,b)$ then F is linear.

Exercise 18.14. ★ As usual $\bar{D}$ is the closed unit disc in $\mathbb{R}^2$ and ∂D its boundary. Let us write

$$\Delta = \{(\mathbf{x}, \mathbf{y}) \in \bar{D}^2 \, : \, \mathbf{x} \neq \mathbf{y}\}$$

and consider Δ as a subset of $\mathbb{R}^4$ with the usual metric. We define $F : \Delta \to \partial D$ as follows.

Given $(\mathbf{x}, \mathbf{y}) \in \Gamma$, take the line from $\mathbf{x}$ to $\mathbf{y}$ and extend it (in the $\mathbf{x}$ to $\mathbf{y}$ direction) until it first hits the boundary at $\mathbf{z}$. We write $F(\mathbf{x}, \mathbf{y}) = \mathbf{z}$.

In the proof of Theorem 4.5 I claimed that it was obvious that F was continuous. Suppose, if possible, that $g : \bar{D} \to \bar{D}$ is a continuous map with $g(\mathbf{x}) \neq \mathbf{x}$ for all $\mathbf{x} \in \bar{D}$. Explain why if my claim is true that the map

$$\mathbf{x} \mapsto F\big(\mathbf{x}, g(\mathbf{x})\big)$$

is a continuous map.

The claimed result is obvious (in some sense) and you may take it as obvious in the exam. However, if we can not prove the obvious it ceases to be obvious. This question outlines one method of proof, but, frankly, the reader may find it easier to find their own method. Any correct method counts as a solution.

(i) Suppose that $0 < y_0 < 1$, $x_0 > 0$ and $x_0^2 + y_0^2 = 1$. Show that, given $\epsilon > 0$ we can find an $\eta > 0$ such that if $x^2 + y^2 = 1$, $x > 0$ and $|y - y_0| < \eta$ implies $\|(x, y) - (x_0, y_0)\| < \epsilon$.

(ii) Suppose that $(x_1, y_0), (x_2, y_0) \in \bar{D}$, $y_0 \geq 0$ and $x_1 \neq x_2$. By using (i), or otherwise, show that, given any $\epsilon > 0$, we can find a $\delta > 0$ such that, whenever

$$\|(x_1', y_1') - (x_1, y)\|,\ \|(x_2', y_2') - (x_2, y)\| < \delta$$

and (x_1', y_1'), $(x_2', y_2') \in \bar{D}$, we have $(x_1', y_1') \neq (x_2', y_2')$ and

$$F\big((x_1', y_1'), (x_2', y_2')\big) = (u, v) \text{ with } |v - y| < \epsilon.$$

(iii) Hence show that $F : \Delta \to \bar{D}$ is continuous.

19 Question sheet 2

Exercise 19.1. In the two player game of Hawks and Doves, player i chooses a probability p_i which announce publicly. Players may change their mind before the game begins but must stick to their last announced decision.

Once the game begins, player i becomes a hawk with probability p_i and a dove with probability $1 - p_i$. Two doves divide food so that each gets $V/2$. A hawk scares off a dove so the hawk gets V and the dove 0. Two hawks fight, the winner gets $V - D$ and the looser $-D$ (D is the damage). The probability of winning such an encounter is $1/2$ for each bird.

If $V > 2D$ show that there is only one Nash equilibrium point. Give a simple explanation of this fact.

If $V < 2D$ show that there are three equilibrium points and identify them.

What happens if $V = 2D$?

Exercise 19.2. (i) Suppose that E is a compact convex set in $\mathbb{R}^n$, that $\alpha : \mathbb{R}^n \to \mathbb{R}^m$ is linear and $\mathbf{b} \in \mathbb{R}^m$, then

$$\{\mathbf{b} + \alpha\mathbf{x} \,:\, \mathbf{x} \in E\}$$

is compact and convex.

(ii) Suppose that E is a compact convex set in $\mathbb{R}^n$, that $f : \mathbb{R}^n \to \mathbb{R}^m$ is continuous and $\mathbf{b} \in \mathbb{R}^m$. Set

$$E' = \{\mathbf{b} + f(\mathbf{x}) \,:\, \mathbf{x} \in E\}$$

(a) Is E' necessarily convex if $n = 1$?
(b) Is E' necessarily convex if $m = 1$?
(c) E' necessarily convex for general m and n?
Give reasons.

Exercise 19.3. (If you have done the 1B optimisation course.) We use the notation of Theorem 5.1. Suppose that $a_{ij} = -b_{ij}$, that is to say that Albert's gain is Bertha's loss. Explain why the 1B game theoretic solution will always be a Nash equilibrium point and vice versa.

Exercise 19.4. (This is Exercise 6.1) Consider two rival firms A and B engaged in an advertising war. So long as the war continues, the additional costs of advertising mean that the larger firm A loses 3 million pounds a year and the smaller firm B loses 1 million pounds a year. If they can agree to cease hostilities then A will make 8 million a year and B will make 1 million a year. How much does Nash say should A pay B per year to achieve this end.

[One way of doing this is to apply an affine transformation.]

Exercise 19.5. Consider the continuous functions on $[0,1]$ with the uniform norm. Show that the unit ball

$$\{f \in C([0,1]) : \|f\|_\infty \leq 1\}$$

is a closed bounded subset of the complete space $(C([0,1]), \| \ \|_\infty)$, but is not compact.

Exercise 19.6. (i) Let $f : [0,1] \to \mathbb{R}$ be a continuous function which is not a polynomial. If p_n is a polynomial of degree d_n and $p_n \to f$ uniformly on $[0,1]$, show that $d_n \to \infty$.

[Hint. Look at Corollary 8.3.]

(ii) If q_n is a polynomials of degree e_n with $e_n \to \infty$ and $q_n \to g$ uniformly on $[0,1]$, does it follow that g is not a polynomial? Give reasons.

Exercise 19.7. Show that no formula of the form

$$\int_{-1}^{1} f(t)\, dt = \sum_{j=1}^{n} A_j f(x_j)$$

(with $x_j, A_j \in \mathbb{R}$) can hold for polynomials f of degree $2n$.

Exercise 19.8. Let $f : [0,1] \to \mathbb{R}$ and $g : [-1,1] \to \mathbb{R}$ be continuous.

(i) By using the Weierstrass approximation theorem, show that

$$\int_0^1 x^n f(x)\, dx = 0 \text{ for all } n \geq 0 \Rightarrow f \text{ is the zero function.}$$

(ii) Show that

$$\int_0^1 x^{2n} f(x)\, dx = 0 \text{ for all } n \geq 0 \Rightarrow f \text{ is the zero function.}$$

(iii) Is it true that if $\int_0^1 x^{2n+1} f(x)\, dx = 0$ for all $n \geq 0$, then f must be the zero-function? Give reasons.

(iv) Is it true that, if $\int_{-1}^1 x^{2n} g(x)\, dx{=}0$ for all $n \geq 0$, then g must be the zero-function? Give reasons.

Exercise 19.9. (i) (This just to remind you that discontinuous functions come in many shapes and sizes.) Let $f : \mathbb{R} \to \mathbb{R}$ be given by $f(x) = \sin 1/x$ for $x \neq 0$ and $f(0) = a$. Show that, whatever the choice of a, f is discontinuous.

(ii) Does there exist a discontinuous function $g : [0,1] \to \mathbb{R}$ which can be approximated uniformly by polynomials? Why?

(iii) Does there exist a smooth function $h : \mathbb{R} \to \mathbb{R}$ which cannot be approximated uniformly by polynomials? Prove your answer.

Exercise 19.10. We say that a function $f:\mathbb{R}\to\mathbb{R}$ has the intermediate value property if whenever $a,b\in\mathbb{R}$ and $f(a)\geq c\geq f(b)$ we can find a t in the closed interval with end points a and b such that $f(t)=c$.

(i) Give an example of a function satisfying the intermediate value property which is not continuous.

(ii) Show that if f has the intermediate value property and in addition $f^{-1}(\alpha)$ is closed for every α in a dense subset of of $\mathbb{R}$ then f is continuous.

Exercise 19.11. ★ Are the following statements true or false? Give reasons.

(i) If $f:(0,1)\to\mathbb{R}$ is continuous, we can find a sequence of polynomials P_n converging uniformly to f on every compact subset of $(0,1)$.

(ii) If $g:(0,1)\to\mathbb{R}$ is continuously differentiable we can find a sequence of polynomials Q_n with Q_n' converging uniformly to g' and R_n converging uniformly to g on every compact subset of $(0,1)$.

(iii) If $h:(0,1)\to\mathbb{R}$ is continuous and bounded we can find a sequence of polynomials R_n with

$$\sup_{t\in(0,1)}|R_n(t)|\leq\sup_{t\in(0,1)}|h(t)|$$

converging uniformly to h on every compact subset of $(0,1)$.

Exercise 19.12. ★ Compute the Chebychev polynomials T_n of the first kind for $n=0,\,1,\,2\,\ldots,\,4$ and the Chebychev polynomials U_{n-1} of the second kind for $n=1,\,2\,\ldots,\,4$.

Recall that we say that a function $f;[-1,1]\to\mathbb{R}$ is *even* if $f(x)=f(-x)$ for all x and *odd* if $f(x)=-f(-x)$ for all x.

Explain why we know, without calculation, that the Chebychev polynomials T_n are even when n is even and odd when n is odd. What can you say about the Chebychev polynomials U_n of the second kind?

Exercise 19.13. The Chebychev polynomials are orthogonal with respect to a certain non-zero positive weight function w. In other words,

$$\int_{-1}^{1}T_m(x)T_n(x)w(x)\,dx=0$$

for all $m\neq n$. Use a change of variables to find a suitable w.

Exercise 19.14. ★ (i) Use the Gramm–Schmidt method (see Lemma 9.2) to compute the Legendre polynomials p_n for $n=0,\,1,\,2,\,3,\,4$. You may leave your answers in the form A_np_n (i.e. ignore normalisation).

(ii) Explain why we know, without calculation, that the Legendre polynomials p_n are even when n is even and odd when n is odd.

(iii) Explain why

$$\frac{d^m}{dx^m}(1-x)^n(1+x)^n$$

vanishes when $x = 1$ or $x = -1$ whenever $m < n$.

Suppose that

$$P_n(x) = \frac{d^n}{dx^n}(1-x^2)^n.$$

Use integration by parts to show that

$$\int_{-1}^{1} P_n(x)P_m(x)\,dx = 0$$

for $m \neq n$. Conclude that the P_n are scalar multiple of the Legendre polynomials p_n.

(iv) Compute P_n for $n = 0, 1, 2, 3, 4$ and check that these verify the last sentence of (iii).

(iv) Let $u_n(x) = x^n$. Find the choice of v which minimises

$$\|u_n - v\|_2 = \left(\int_{-1}^{1} |x^n - v(x)|^2\,dx\right)^{1/2}$$

for v a polynomial of degree at most $n-1$

Exercise 19.15. Are the following statements true or false. Give reasons.

(i) For all $n \geq 1$, there exists a polynomial P_n of degree at most n such that

$$P_n(\cosh t) = \cosh nt.$$

(ii) For all $n \geq 1$ there exists a polynomial Q_n of degree at most n such that

$$Q_n(\cosh t) = \sinh nt.$$

(iii) For all $n \geq 1$ there exists a polynomial R_n of degree at most n such that

$$R_n(\sin t) = \sin nt.$$

Exercise 19.16. ★ (i) Suppose $f : [a, b] \to \mathbb{R}$ is continuous and $\epsilon > 0$. Why can we find an infinitely differentiable function $g : [a, b] \to \mathbb{R}$ such that $\|f - g\|_\infty < \epsilon$.

(ii) By using Chebychev polynomials and Weierstrass's approximation theorem, show that given any continuous $f : [0, \pi] \to \mathbb{R}$ and any $\epsilon > 0$ we can find N and $a_j \in \mathbb{R}$ $0 \leq j \leq N$ such that

$$\left| f(s) - \sum_{j=0}^{N} a_j \cos js \right| < \epsilon$$

for all $s \in [0, \pi]$.

(iii) Let $\epsilon > 0$. If $f : [0, \pi] \to \mathbb{R}$ is continuous with $f(0) = 0$ show that we can find N and $b_j \in \mathbb{R}$ $0 \leq j \leq N$ such that

$$\left| f(s) - b_0 s - \sum_{j=0}^{N} b_j \sin js \right| < \epsilon$$

for all $s \in [0, \pi]$.

(iv) Let $\epsilon > 0$. If $f : [0, \pi] \to \mathbb{R}$ is continuous with $f(0) = f(\pi) = 0$ show that we can find N and $b_j \in \mathbb{R}$ $0 \leq j \leq N$ such that

$$\left| f(s) - \sum_{j=0}^{N} b_j \sin js \right| < \epsilon$$

for all $s \in [0, \pi]$.

(v) Hence show that given any continuous $f : [-\pi, \pi] \to \mathbb{R}$ with $f(-\pi) = f(\pi)$ and any $\epsilon > 0$ we can find N and $\alpha_j, \beta_j \in \mathbb{R}$ such that

$$\left| f(t) - \sum_{j=0}^{N} b_j \cos jt - \sum_{j=1}^{N} c_j \sin jt \right| < \epsilon$$

for all $t \in [-\pi, \pi]$.

20 Question Sheet 3

Exercise 20.1. Let $f : [-1, 1] \to \mathbb{R}$ be a function and let $M > 0$. Show that there exists at most one polynomial of degree n such that

$$|f(x) - P(x)| \leq M|x|^{n+1}$$

for all $x \in [-1, 1]$.

Must there always exist such a P if f is everywhere infinitely differentiable and we choose M sufficiently large?

Exercise 20.2. Let T_j be the jth Chebychev polynomial. Suppose that γ_j is a sequence of non-negative numbers such that $\sum_{j=1}^{\infty} \gamma_j$ converges. Explain why $\sum_{j=1}^{\infty} \gamma_j T_{3^j}$ converges uniformly on $[-1, 1]$ to a continuous function f.

Let us write $P_n = \sum_{j=1}^{n} \gamma_j T_{3^j}$. Show that we can find points

$$-1 \leq x_0 < x_1 < \ldots < x_{3^{n+1}} \leq 1$$

such that

$$f(x_k) - P_n(x_k) = (-1)^{k+1} \sum_{j=n+1}^{\infty} \gamma_j.$$

Exercise 20.3. Use Exercise 20.2 to show that, given any decreasing sequence $\delta_n \to 0$, we can find a continuous function $f : [-1, 1] \to \mathbb{R}$ such that (writing $\| \ \|_\infty$ for the uniform norm on $[-1, 1]$)

$$\inf\{\|f - P\|_\infty : P \text{ a polynomial of degree at most } n\} \geq \delta_n\}.$$

Why does this not contradict the Weierstrass approximation theorem?

Exercise 20.4. Use the ideas of Theorem 7.9 to show that, if $f : [0, 1]^2 \to \mathbb{R}$ is continuous, then, given $\epsilon > 0$, we can find a polynomial P in two variables such that

$$|f(x, y) - P(x, y)| < \epsilon$$

for all $x, y \in [0, 1]$.

Exercise 20.5. (Not very much to do with the course but a nice question which you should have met at least once in your life.) Suppose $f : [-1, 1]^2 \to \mathbb{R}$ is a bounded function such that the map $x \mapsto f(x, y)$ is continuous for each fixed y and the map $y \mapsto f(x, y)$ is continuous for each fixed x. By means of a proof or counterexample establish whether f is necessarily continuous.

The next three questions give alternative proofs of Weierstrass's theorem. Each involves some heavy lifting but each introduces ideas which are very useful in a variety of circumstances. If you are finding the course heavy going or your busy social schedule limits the time you can spend thinking to an absolute minimum you can skip them. If you want to do any sort of analysis in the future they are highly recommended.

Exercise 20.6. Here is an alternative proof of Bernstein's theorem using a different set of ideas.

(i) Let $f \in C([0,1])$. Show that given $\epsilon > 0$ we can find an $A > 0$ such that

$$f(x) + A(t-x)^2 + \epsilon/2 \geq f(t) \geq f(x) - A(t-x)^2 - \epsilon/2$$

for all t, $x \in [0,1]$.

(ii) Now show that we can find an N such that, writing

$$h_r(t) = f(r/N) + A(t - r/N)^2, \; g_r(t) = f(r/N) - A(t - r/N)^2,$$

we have

$$g_r(t) + \epsilon \geq f(t) \geq h_r(t) - \epsilon$$

for $|t - r/N| \leq 1/N$. (You may find it helpful to draw diagrams here and in (iii).)

(iii) We say that a linear map $S : C([0,1]) \to C([0,1])$ is positive if $F(t) \geq 0$ for all $t \in [0,1]$ implies $SF(t) \geq 0$ for all $t \in [0,1]$. Suppose that S is such a positive linear operator. Show that if $F(t) \geq G(t)$ for all $t \in [0,1]$, then $(SF)(t) \geq (SG)(t)$ for all $t \in [0,1]$ $[F, G \in C([0,1])]$. Show also that if, $F \in C([0,1])$, then $\|SF\|_\infty \leq \|S1\|_\infty \|F\|_\infty$.

(iv) Write $e_r(t) = t^r$. Suppose that S_n is a sequence of positive linear functions such that $\|S_n e_r - e_r\|_\infty \to 0$ as $n \to \infty$ for $r = 0$, $r = 1$ and $r = 2$. Show, using (ii), or otherwise, that $\|S_n f - f\|_\infty \to 0$ as $n \to \infty$ for all $f \in C([0,1])$.

(v) Let

$$(S_n f)(t) = \sum_{j=0}^{n} \binom{n}{j} f(j/n) t^j (1-t)^{n-j}.$$

Verify that S_n satisfies the hypotheses of part (iv) and deduce Bernstein's theorem.

Exercise 20.7. ★ Here is another proof of Weierstrass's theorem which is closer to his original proof. We wish to show show that any continuous function function $f : [-1/2, 1/2] \to \mathbb{R}$ can be uniformly approximated by polynomials on $[-1/2, 1/2]$. To do this we show that any continuous function $g : \mathbb{R} \to \mathbb{R}$ with $g(x) = 0$ for $|x| \geq 1$ can be uniformly approximated by polynomials on $[-1/2, 1/2]$. Why does this give the desired result?

Let

$$L_n(x) = \begin{cases} (4 - x^2)^n & \text{for } |x| \leq 2, \\ 0 & \text{otherwise,} \end{cases}$$

let

$$A_n = \int_{-\infty}^{\infty} L_n(x)\, dx$$

and let $K_n(x) = A_n^{-1} L_n(x)$.

(i) Show that

$$P_n(x) = K_n * g(x) = \int_{-\infty}^{\infty} K_n(x-t) g(t)\, dt$$

is a polynomial in x on the interval $[-1/2, 1/2]$ It may be helpful to recall that $f * g = g * f$.)

(ii) Let $\delta > 0$ be fixed. Show that $K_n(x) \to 0$ uniformly for $|x| \geq \delta$ and

$$\int_{-\delta}^{\delta} K_n(x)\, dx \to 1$$

as $n \to \infty$.

(iii) Use the fact that g is bounded and uniformly continuous together with the formula

$$P_n(x) = \int_{-\delta}^{\delta} K_n(t) g(x-t)\, dt + \int_{t \notin (-\delta, \delta)} K_n(t) g(x-t)\, dt$$

to show that $P_n(x) \to g(x)$ uniformly on $[-1/2, 1/2]$.

Exercise 20.8. Here is another proof of Weierstrass's theorem, this time due to Lebesgue.

(i) If $a < b$ sketch the graph of $|x-a| - |x-b|$.

(ii) Show that if $g : [0,1] \to \mathbb{R}$ is piece-wise linear, then we can find $n \geq 1$, $\lambda_j \in \mathbb{R}$ and $a_j \in [0,1]$ such that

$$g(t) = \lambda_0 + \sum_{j=1}^{n} \lambda_j |t - a_j|.$$

Deduce that, given $f : [0,1] \to \mathbb{R}$ and $\epsilon > 0$, then we can find $n \geq 1$, $\lambda_j \in \mathbb{R}$ and $a_j \in [0,1]$ such that

$$\left| f(t) - \lambda_0 - \sum_{j=1}^{n} \lambda_j |t - a_j| \right| < \epsilon.$$

(iii) Let

$$u_n(t) = 3\sqrt{\left(1 + \frac{1}{n}\right) - \left(1 - \frac{t^2}{9}\right)}.$$

Explain using results on the general binomial expansion (which you need not prove) why u_n can be uniformly approximated by polynomials on $[-2, 2]$.

Explain why $u_n(t) \to |t|$ uniformly on $[-2, 2]$ as $n \to \infty$. Deduce that there exist polynomials q_r with $q_r(t) \to |t|$ uniformly on $[-1, 1]$ as $r \to \infty$.

(iv) Use (ii) and (iii) to prove the Weierstrass approximation theorem. [Lebesgue's idea provides the basis for the proof of the more general Stone–Weierstrass theorem.]

Exercise 20.9. (This will look less odd if you have done the previous exercise.)

(i) Let the sequence of distinct x_n form a dense subset of $[0, 1]$ with $x_0 = 0$, $x_1 = 1$. If $f \in C([0, 1])$, define $f_n : [0, 1] \to \mathbb{R}$ to be the simplest piece-wise linear function with $f_n(x_j) = f(x_j)$ for $0 \leq j \leq n$. Show that $f_n \to f$ uniformly.

(ii) Use (i) to show that there exists a sequence of continuous functions ϕ_n such that, for each $f \in C([0, 1])$ there exists a unique sequence a_n such that

$$\sum_{j=0}^{n} a_j \phi_j \to f$$

uniformly on $[0, 1]$.

[In practice the sequence x_j is usually taken to be 0, 1, 1/2, 1/4, 3/4 1/8, 3/8, 5/8, 7/8,1/16, 3/16, ...]

Exercise 20.10. ★ In Theorem 8.4 we say that, if $f : [a, b] \to \mathbb{R}$ is a continuous function there exists a polynomial P, of degree at most $n - 1$, such that $\|P - f\|_\infty \leq \|Q - f\|_\infty$ for all polynomials Q degree n or less. The object of this question is to show that the polynomial P satisfies the equiripple criterion.

We claim that we can find $a \leq a_0 \leq a_1 \leq \cdots \leq a_n \leq b$ such that, writing $\sigma = \|f - P\|_\infty$ we have either

$$f(a_j) - P(a_j) = (-1)^j \sigma \text{ for all } 0 \leq j \leq n$$

or

$$f(a_j) - P(a_j) = (-1)^{j+1} \sigma \text{ for all } 0 \leq j \leq n.$$

Our proof will be by reductio ad absurdum.

We assume without loss of generality that $[a, b] = [0, 1]$ and $\sigma = 1$.

(i) Write $g = f - P$. Explain why we can find an integer $N \geq 1$ such that if, $1 \leq r \leq N$, at least one of the following statements must be true

$$g(x) \geq 1/2 \text{ for all } x \in [(r - 1)/N, r/N],$$

or

$$g(x) \leq -1/2 \text{ for all } x \in [(r - 1)/N, r/N],$$

or

$$|g(x)| \leq 3/4 \text{ for all } x \in [(r-1)/N, r/N].$$

(ii) Using the result of (i) show that, if our putative theorem is false, we can find an integer $q \leq n$ and integers

$$0 = u(1) < v(1) < u(2) < v(2) < \cdots < u(q) < v(q) = N$$

and $w \in \{0, 1\}$ such that

$$\begin{aligned}(-1)^{w+j} g(x) > -1 &\text{ for all } x \in [u(j)/N, v(j)/N]\\ |g(x)| < 1 &\text{ for all } x \in [v(j)/N, u(j+1)/N].\end{aligned}$$

Without loss of generality, we take $w = 0$.

(iii) Explain why we can find an $\eta > 0$ with

$$\begin{aligned}(-1)^{j} g(x) > -1 + \eta &\text{ for all } x \in [u(j)/N, v(j)/N]\\ |g(x)| < 1 - \eta &\text{ for all } x \in [v(j)/N, u(j+1)/N],\end{aligned}$$

for all j. We may take $\eta < 1/8$ and will do so.

(iv) Explain how to find a polynomial R of degree n or less with $\|R\|_\infty = 1$ such that

$$(-1)^j R(x) > 0 \text{ for all } x \in [u(j)/N, v(j)/N]$$

and $j = 1, 2, \ldots, q$.

(v) Show that

$$|g(x) - (\eta/2)R(x)| < 1 - \eta/2$$

for all $x \in [0, 1]$. Hence obtain a contradiction.

Exercise 20.11. Suppose that we have a sequence of continuous functions $f_n : [0,1] \to \mathbb{R}$ such that $f_n(x) \to 0$ for each $x \in [0,1]$ as $n \to \infty$. Then, given $\epsilon > 0$, we can find a non-empty interval $(a,b) \subseteq [0,1]$ and an $N(\epsilon)$ such that

$$|f_n(t)| \leq \epsilon$$

for all $t \in (a,b)$ and all $n \geq N(\epsilon)$.
Hint Consider the sets

$$E_N = \{x \in [0,1] : |f_n(x)| < \epsilon, \text{for all } n \geq N\}.$$

Exercise 20.12. Suppose that $f : [1,\infty) \to \mathbb{R}$ is a continuous function and $f(nx) \to 0$ as $n \to \infty$ for each $x \in [1,\infty)$. Show that $f(x) \to 0$ as $x \to \infty$.

Exercise 20.13. ★ (i) Consider $C([0,1])$ with the uniform norm. If M is strictly positive integer, let $\mathcal{E}_M$ be the set of $f \in C([0,1])$ such that whenever $N \geq 1$ and

$$0 \leq x_0 < x_1 < x_2 < \ldots < x_N \leq 1$$

we have

$$\sum_{j=1}^{N} |f(x_{j-1}) - f(x_j)| \leq M.$$

Show that $\mathcal{E}_M$ is closed with dense complement. Deduce that there is a set $\mathcal{G}$ which is the complement of a set of first category such that given any $f \in \mathcal{G}$ and any $M \geq 1$ we can find $N \geq 1$ and

$$0 \leq x_0 < x_1 < x_2 < \ldots < x_N \leq 1$$

with

$$\sum_{j=1}^{N} |f(x_{j-1}) - f(x_j)| > M.$$

(ii) Show that there is a set $\mathcal{H}$ which is the complement of a set of first category such that given any $f \in \mathcal{H}$, any a and b with $0 \leq a < b \leq 1$ and any $M \geq 1$ we can find $N \geq 1$ and

$$a \leq x_0 < x_1 < x_2 < \ldots < x_N \leq b$$

with

$$\sum_{j=1}^{N} |f(x_{j-1}) - f(x_j)| > M.$$

Exercise 20.14. Let $h : [0,1] \to \mathbb{R}$ be a continuous strictly increasing function with $h(0) = 0$. We say that a compact set E is thin if, given $\epsilon > 0$, we can find a finite collection of intervals I_j of length l_j [$N \geq j \geq 1$ such that

$$E \subseteq \bigcup_{j=1}^{N} I_j, \text{ but } \sum_{j=1}^{N} h(l_j) < \epsilon.$$

Show that the set $\mathcal{C}$ of thin compact sets is the complement of a set of first category in the space $\mathcal{K}$ of compact subsets of $[0,1]$ with the Hausdorff metric ρ.

Exercise 20.15. ★ Let $A = \{z \in \mathbb{C} : 1/2 < |z| < 1\}$ and let $D = \{z \in \mathbb{C} : |z| < 1\}$. Suppose that $f : A \to \mathbb{C}$ is analytic and we can find polynomials p_n with $p_n(z) \to f(z)$ uniformly on A. Show that we can find an analytic function $g : D \to \mathbb{C}$ with $f(z) = g(z)$ for all $z \in A$.
[Hint: Use the maximum modulus principle and the general principle of uniform convergence.]

Exercise 20.16. (i) We work in $\mathbb{C}$. Show that there exists a sequence of polynomials P_n such that

$$P_n(z) \to \begin{cases} 1 & \text{if } |z| < 1 \text{ and } \Re z \geq 0 \\ 0 & \text{if } |z| < 1 \text{ and } \Re z < 0 \end{cases}$$

as $n \to \infty$.
[Hint: Recall that, if Ω_1 and Ω_2 are disjoint open sets and $f(z) = 0$ for $z \in \Omega_1$ and $f(z) = 1$ for $z \in \Omega_2$, then f is analytic on $\Omega_1 \cup \Omega_2$.]

(ii) Show that there exists a sequence of polynomials Q_n such that

$$Q_n(z) \to \begin{cases} 1 & \text{if } \Re z \geq 0 \\ 0 & \text{if } \Re z < 0 \end{cases}$$

as $n \to \infty$.

21 Question sheet 4

Exercise 21.1. By quoting the appropriate theorems show that, if Ω is an open set in $\mathbb{C}$, then $f : \Omega \to \mathbb{C}$ is analytic if and only if, whenever K is a compact subset of Ω with path-connected complement and $\epsilon > 0$, we can find a polynomial P with $|f(z) - P(z)| < \epsilon$ for all $z \in K$.

Exercise 21.2. In this exercise we suppose that K is a bounded compact subset of $\mathbb{C}$ and E is a non-empty bounded connected component of $\mathbb{C} \setminus K$. Give a simple example of such a K and E. Our object is to show that if $a \in E$ the function $f(z) = (z-a)^{-1}$ is not uniformly approximable on K by polynomials.

Suppose P is a polynomial with $|p(z) - (z-a)^{-1}| \leq \epsilon$ or all $z \in K$. By observing that the boundary ∂E of E lies in K and using the maximum modulus principle deduce that $|p(w)(w-a) - 1| \leq \epsilon \sup_{z \in K} |z - a|$. By choosing w appropriately deduce that $\epsilon \geq (\sup_{z\in K} |z-a|)^{-1}$.

Exercise 21.3. Show that $\cos 1$ is irrational. Show more generally that $\cos 1/n$ is irrational whenever n is a non zero integer.

Exercise 21.4. Use the idea of Louiville's theorem to write down a continued fraction whose value is transcendental. Justify your answer.

Exercise 21.5. Let us write $\langle y \rangle = y - [y]$ so that $\langle y \rangle$ is the fractional part of y

Suppose that x is irrational. If m is strictly positive integer consider the $m + 1$ points

$$0 = \langle 0x \rangle,\ \langle 1x \rangle,\ \ldots,\ \langle kx \rangle,\ \ldots,\ \langle mx \rangle$$

and explain why there must exist integers r and s with $0 \leq s < r \leq m$ and

$$|\langle rx \rangle - \langle sx \rangle| \leq 1/m.$$

Deduce that we can find an integer v with $1 \leq v \leq m$ and and integer u with

$$|vx - u| \leq 1/m$$

and so with

$$\left| x - \frac{u}{v} \right| \leq \frac{1}{mv} \leq \frac{1}{v^2}.$$

Deduce that we can find u_n, v_n integers with $v_n \to \infty$ such that

$$\left| x - \frac{u_n}{v_n} \right| \leq \frac{1}{v_n^2}.$$

Exercise 21.6. Determine the continued fraction expansion of $71/49$ and use your result to find the rational number with denominator no greater than 10 which best approximates $71/49$.

Exercise 21.7. (i) Determine the continued fraction expansions of $\sqrt{3}$.

(ii) Explain why the form of the continued fraction shows that for $\sqrt{3}$ is irrational.

(iii) Let p_n/q_n be the nth convergent for $\sqrt{3}$. Compute $(p_n/q_n)^2$ for $n = 1, 2, 3, 4, 5$.

Exercise 21.8. Let a and b be strictly positive integers. If

$$x = \cfrac{1}{a + \cfrac{1}{b + \cfrac{1}{a + \cfrac{1}{b + \dots}}}},$$

show that $ax^2 + abx - b = 0$

Exercise 21.9. If x is irrational, we can find u_n and v_n show that we can find integers with $v_n \to \infty$ such that

$$\left|\frac{u_n}{v_n} - x\right| < \frac{1}{2v_n^2}.$$

[Hint: Show that, in fact, at least one of the convergents p_n/q_n or p_{n+1}/q_{n+1} must satisfy the required inequality.]

Exercise 21.10. Show that if all the integers a_n in the continued fraction

$$x = \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \dots}}}},$$

are bounded, then there exists an $M > 0$ such that

$$\left|x - \frac{p}{q}\right| > \frac{M}{q^2}$$

for all integers p and q with $q \neq 0$.

Exercise 21.11. The Fibonacci sequence has many interesting aspects. (It is, so far as I know the only series with its own Fanzine — *The Fibonacci Quarterly.*)

(i) Find the general solution of the difference equation

$$u_{n+1} = u_n + u_{n-1}.$$

The Fibonacci series is the particular solution $F_n = u_n$ with $u_0 = 0$, $u_1 = 1$. Write F_n in the appropriate form.

(ii) Show, by using (i), or otherwise, that if $n \geq 1$, F_n is the closest integer to

$$\frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n.$$

We call

$$\tau = \frac{1+\sqrt{5}}{2}$$

the golden ratio.

(iii) Prove the two identities

$$\begin{aligned} F_{2n+1} &= F_n^2 + F_{n+1}^2 \\ F_{2n} &= F_n(F_{n-1} + F_{n+1}) \end{aligned}$$

by using the result of (i).

(iv) Explain why

$$\begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix} = A^n$$

where

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Use the result $A^{n+m} = A^n A^m$ to deduce that

$$\begin{aligned} F_{n+m+1} &= F_{n+1}F_{m+1} + F_n F_m \\ F_{n+m} &= F_n F_{m+1} + F_{n-1} F_m. \end{aligned}$$

Obtain (iii) as a special case.

(v) Let $x_n = F_{n+1}/F_n$. Use (iii) to express x_{2n} as a rational function of x_n.

(vi) Suppose now we take $y_k = x_{2^k}$. Write down y_{n+1} as a rational function of y_n. Use (i) to show that y_k converges very rapidly to τ. Can you link this with the Newton–Raphson method for finding a root of a particular function?

Exercise 21.12. Let $p(z) = z^2 - 4z + 3$ and let $\gamma : [0,1] \to \mathbb{C}$ be given by $\gamma(t) = p(2e^{2\pi it})$. Show that closed path associated with γ does not pass through 0.

Compute $w(\gamma, 0)$

(i) Non-rigorously direct from the definition by obtaining enough information about γ, (You could write the real and imaginary parts of $\gamma(t)$ in terms of $\cos t$ and $\sin t$.)

(ii) by factoring, and

(iii) by the dog walking lemma.

Exercise 21.13. ★ Suppose that $\gamma : [0,1] \to \mathbb{C} \setminus \{0\}$ is a continuously differentiable function with $\gamma(0) = \gamma(1)$.

If we define $r : [0,1] \to \mathbb{C}$ by

$$r(t) = \exp\left(\int_0^t \frac{\gamma'(s)}{\gamma(s)}\, ds\right)$$

compute the derivative of $r(t)/\gamma(t)$ and deduce that

$$w(\gamma, 0) = \frac{1}{2\pi i}\int_0^1 \frac{\gamma'(s)}{\gamma(s)}\, ds.$$

Use this result and the residue theorem to compute $w(\gamma, 0)$ in Exercise 21.12.

[The example used in the last two questions has been chosen to make things easy. However, if you are prepared to work hard, it is possible to obtain enough information about γ to find the winding number of closed curves even in quite complicated cases. If many winding numbers are required (as may be the case when studying stability in an engineering context then we can use numerical methods (this question suggests a possibility, though not necessarily a good one) together with the knowledge that the winding number is an integer to obtain winding numbers on an industrial scale.]

Exercise 21.14. ★ Take your electronic calculator out of your old school satchel (or use the expensive piece of equipment on which you play games) and find the first few terms of the continued fraction for π (or, more strictly for the rational number that your calculator gives when you ask it for π.) Compute first few associated convergents (what we have called $3 + p_n/q_n$).

Verify that 355/113 is an extremely good approximation for π and explain why this is so. Apparently the approximation was first discovered by the astronomer Tsu Ch'ung-Chih in the fifth century A.D.

The entries a_n in the continued fraction expansion for π look, so far as anyone knows, just like those you would expect from a random real number (in a sense made precise in Corollary 14.6).

I would be inclined to say that this was precisely what one should expect if there was not a beautiful expansion (using a generalisation of the kind of continued fraction discussed in the course) found by Lord Brouncker in 1654.

$$\frac{\pi}{4} = 1 + \cfrac{1^2}{1 + \cfrac{3^2}{2 + \cfrac{5^2}{2 + \cfrac{7^2}{2 + \dots}}}}.$$

You may easily verify that the first convergents are

$$1, \ , 1 - \frac{1}{3}, \ 1 - \frac{1}{3} + \frac{1}{5} -, \ \dots$$

and, if your name is Euler, that the nth convergent is

$$\sum_{j=0}^{n} \frac{(-1)^j}{2j+1}$$

and then, if your name is Leibniz, you will prove the result

$$\sum_{j=0}^{n} \frac{(-1)^j}{2j+1} \to \frac{\pi}{4}.$$

The convergence is, however, terribly slow and it is no wonder that Huygen's initially disbelieved Brouncker's result.

Index