Topics in Analysis

T. W. Körner

August 11, 2022

Small print This course is not intended as a foundation for further courses, so the lecturer is allowed substantial flexibility. Exam questions will be set on the course as given (so, if I do not have time to cover everything in these notes, the topics examined will be those actually lectured) and will only cover topics in these notes. Unless there is more time to spare than I expect, I will not talk about topological spaces but confine myself to ordinary Euclidean spaces $\mathbb{R}^n$, $\mathbb{C}$ and complete metric spaces.

These notes lay out the results of the course. There is a separate set of notes giving the associated proofs, but those notes are intended for home use and not as a means of following the live lectures.

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_ε and should be available in tex, ps, pdf and dvi format from my home page

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

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.

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.*

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.

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).

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.

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 $k \to \infty$.*

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.

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 $\mathbb{R}^m$.)

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

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* $\rho(f(z), f(x)) < \epsilon$*.*

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.*

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.*

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* $\mathbb{R}^n$*.*

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*

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

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

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.*

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, $\ldots$, $a_n \in \mathbb{C}$ 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.

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* $\operatorname{Cl} E$*.*

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*.*

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* $\partial \Omega$*.*

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*?]*

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* $\operatorname{Cl} \Omega$*.*

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 forms 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 $\phi : \mathrm{Cl}\,\Omega \to \mathbb{R}$ is continuous, that ϕ is twice differentiable on Ω and that ϕ 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.

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 hold 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 $F(a) = a$.*

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.)*

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

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

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

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

(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 $A \cap B \cap C \neq \varnothing$.

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.*

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* $[r = 0, 1, \ldots, n]$ *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*

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

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

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

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

and the expected value of the arrangement to Bertha

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

If you went to the 1B optimisation course, then 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

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

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.

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.*

---

[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.

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.

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.

[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.

[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.

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.)

(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 $x_1^* = x_2^* = \cdots = x_m^*$. 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 in (5) above is closed bounded and convex.*

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*

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

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.

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.

'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.)

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

Lemma 7.2. *Let $x_0, x_1, \ldots, x_n$ 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.

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 $\sin\theta \neq 0$. 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 $1 \leq r \leq n-1$ and the roots of T_n are $\cos\big((r+\frac{1}{2})\pi/n\big)$ with $0 \leq r \leq n-1$.

The coefficient of t^n in T_n is 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

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

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.

[6] Or Tchebychev, hence the T.

(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

$$\|f\|_\infty \geq \|f\|_2 \geq \|f\|_1.$$

Let

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

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

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.*

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 $a > 0$.

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*

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

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 $n \to \infty$.

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.*

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* $n-1$.

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* $\|P\|_\infty \geq \epsilon(n)$.

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.*

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.

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 increases.

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*

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

then 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 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].

[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 for most uses.

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

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

Lemma 9.4. *The* n*th Legendre polynomial* p_n *has* n *distinct roots all 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* n*th 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* n*th Legendre polynomial.*

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 $n \to \infty$.

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*

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

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.

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.

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.*

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 $E \cap F \neq \varnothing$.

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 $d(\mathbf{e}, F) = \sigma(E,F)$.*

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 $E \subseteq F$.

However, σ does obey the triangle inequality.

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

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

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 $\mathbb{R}^m$.*

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 $p \to \infty$.

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

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

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. We write $\bar{z}$ for the complex conjugate of z.

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

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

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

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

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 $w \neq 0$, 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 $z \in \Omega$.

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.

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*

$$f(z) = \frac{1}{2\pi i} \sum_{m=1}^{M} \int_{C_m} \frac{f(w)}{w - z}\, dw$$

whenever $z \in K$ and $f : \Omega \to \mathbb{C}$ is analytic.

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.*

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*

$$\left| f(z) - \sum_{n=1}^{N} \frac{A_n}{z - \alpha_n} \right| \leq \epsilon$$

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 $z \in K$.

Let us make a temporary definition.

[9] And, provided she does not twist the examiner's nose, few marks in the exam.

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 $\alpha \in \Lambda(K)$.*

Lemma 11.18. *Let K be a compact set in $\mathbb{C}$. If $\alpha \in \Lambda(K)$ and $|\alpha - \beta| < d(\alpha, K)$ then $\beta \in \Lambda(K)$.*

Lemma 11.19. *Suppose that K is a compact set in $\mathbb{C}$ and $\mathbb{C} \setminus K$ is path connected. Then $\Lambda(K) = \mathbb{C} \setminus K$.*

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 $z \in D$.

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.

[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.

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)]

Let us look at some of the prettier shells.

Theorem 12.1. *The number e is 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.

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.

[12]Professor Gowers in this case.

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.*

Lemma 12.6. *The algebraic numbers are countable.*

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 was given earlier by Liouville[13]. It is longer but actually produces particular examples.

Theorem 12.7. [Liouville] *Suppose α is an irrational root of the equation*

$$a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0 = 0$$

[13] Once when lecturing to a class [Kelvin] used the word 'mathematician,' and then interrupting himself asked his class: 'Do you know what a mathematician is?' Stepping to the blackboard he wrote upon it:—

$$\int_{-\infty}^{\infty} e^{-x^2/2}\,dx = \sqrt{2\pi}.$$

Then putting his finger on what he had written, he turned to his class and said: 'A mathematician is one to whom that is as obvious as that twice two makes four is to you. Liouville was a mathematician.' [S. P. Thompson, *Life of Lord Kelvin*] *where* $a_j \in \mathbb{Z}$ $[0 \le j \le n]$, $n \ge 1$ *and* $a_n \ne 0$. *Then there is a constant* $c > 0$ *(depending on the* a_j*) such that*

$$\left|\alpha - \frac{p}{q}\right| \ge \frac{c}{q^n}$$

for all $p,\ q \in \mathbb{Z}$ *with* $q \ne 0$

We can now exhibit a transcendental number.

Theorem 12.8. *The number*

$$\sum_{j=0}^{\infty} \frac{1}{10^{j!}}$$

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.*

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[14] proofs of the transcendence of e and π.

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* $U_1,\ U_2,\ \dots$ *are open sets whose complements have empty interior, then*

$$\bigcap_{j=1}^{\infty} U_j \ne \varnothing.$$

[14] But miraculous.

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*

$$\bigcup_{j=1}^{\infty} F_j \neq X.$$

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, ...

Baire's theorem has given rise to the following standard definitions[15].

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.

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 $B(x, \delta) \cap E \neq \{x\}$.

[15]Your lecturer thinks that, whilst the concepts defined are very useful, the nomenclature is particularly unfortunate.

Theorem 13.7. *A non-empty complete metric space without isolated points is uncountable.*

Corollary 13.8. *The real numbers are 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.*

Here is another corollary of Theorem 13.7.

Corollary 13.10. *A non-empty closed subset of* $\mathbb{R}$ *without isolated points is uncountable.*

Do there exist nowhere dense closed subsets of $\mathbb{R}$ with no isolated points[16]? 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.*

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* $F \supseteq [j/k, (j+1)/k]$ $[0 \leq j \leq k,\ 1 \leq k]$.

(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.*

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.

[16] Such sets are called perfect. If they make pretty pictures they are called fractals. You do not have to remember either name.

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* $r \geq 0$.

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*

$$|f_n(t)| \leq M$$

for all $t \in (a,b)$ *and all* $n \geq 1$.

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 $1 > T(x) \geq 0$. Thus

$$Nx = \left[\frac{1}{x}\right], \; Tx = \frac{1}{x} - \left[\frac{1}{x}\right]$$

. 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[17] 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 $N(x)$, $NT(x)$, $NT^2(x)$, $\dots$. In the next section we shall show first that the continued fraction can be assigned a numerical meaning and then that the assigned meaning is, as we might hope, x.]

[17] As opposed to typesetters; this sort of thing turned their hair prematurely grey.

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.*

Once we have done a couple of examples it is clear that we are simply echoing Euclid's algorithm[18].

Lemma 14.2. *(i) Suppose that r_k, s_k are coprime positive integers with $r_k < s_k$; that r_{k+1}, s_k are coprime strictly positive integers with $r_{k+1} < s_{k+1}$. and that a_k is a strictly positive integer. If*

$$\frac{r_k}{s_k} = \cfrac{1}{a_k + \cfrac{r_{k+1}}{s_{k+1}}},$$

then $s_{k+1} = r_k$

$$s_k = a_k s_{k+1} + r_{k+1}$$

for some positive integer k_1. Thus the pair (r_{k+1}, s_{k+1}) is obtained from (r_k, s_k) by applying one step of the Euclidean algorithm.

(ii) If y is a rational number, its continued fraction expansion (obtained by the method described above) terminates.

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.

[18]I think historians would reverse the order and say that continued fractions gave rise to Euclid's algorithm

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 $0 \leq j \leq 9$.

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.

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

$$\Pr(NT^mX = j) = \frac{1}{\log 2}\int_{1/(j+1)}^{1/j}\frac{dx}{1+x} = \frac{1}{\log 2}\log\left(\frac{(j+1)^2}{j(j+2)}\right).$$

Proof. By Lemma 14.5,

$$\begin{aligned}
\Pr(NT^mX = j) &= \Pr(NX = j)\\
&= \frac{1}{\log 2}\int_{j^{-1}}^{(j+1)^{-1}}\frac{1}{1+x}\,dx\\
&= \frac{1}{\log 2}\Big[\log(1+x)\Big]_{(j+1)^{-1}}^{j^{-1}}\\
&= \frac{1}{\log 2}\big(\log(j+2) - \log(j+1)\big) = \frac{1}{\log 2}\log\frac{j+2}{j+1}.
\end{aligned}$$

□

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

$$\Pr(NT^mY = j) \to \frac{1}{\log 2}\log\left(\frac{(j+1)^2}{j(j+2)}\right).$$

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 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 continued fraction associated with $\mathbf{a} = (a_0, a_1, a_2, \dots)$.

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
$$\frac{r_0}{s_0} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \cfrac{1}{\ddots \, \cfrac{}{a_{n-1} + \cfrac{1}{a_n}}}}}}}$$

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

$$\begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n & 1 \\ 1 & 0 \end{pmatrix},$$

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

$$\left| \frac{p_n}{q_n} - \alpha \right| \leq \frac{1}{q_n q_{n+1}}.$$

Exercise 15.5. *Suppose we have an irrational $x \in (0, 1]$ and we form a continued fraction (with $a_0 = 0$)*

$$\cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cfrac{1}{a_4 + \dots}}}}$$

in 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.

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 $1 \leq q \leq q_n$.

Theorem 15.7. *If x is irrational, we can find u_n and v_n integers with $v_n \to \infty$ such that*

$$\left|\frac{u_n}{v_n} - x\right| < \frac{1}{v_n^2}.$$

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 $v \geq 1$?

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

$$F_{n+1}F_{n-1} - F_n^2 = (-1)^n.$$

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. Exercise 21.11 (vii) suggests a better estimate for m.

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.*

Figure 1: Carroll's puzzle

What is the connection with Exercise 15.9? Can you design the next puzzle in the sequence?

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.

16 A nice, but starred, formula

This section is non-examinable

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

$$\begin{aligned} p_n &= a_n p_{n-1} + b_{n-1} p_{n-2} \\ q_n &= a_n q_{n-1} + b_{n-1} q_{n-2}. \end{aligned}$$

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* $p_0(x) = 0$, $q_0(x) = 1$, $p_1(x) = x$, $q_1(x) = 1$.

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 $t \in [0,1]$.

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 $t \in [0,1]$.*

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 $\gamma(t) = |\gamma(t)|e^{i\theta(t)}$.*

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.

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*
$$w(\gamma_1\gamma_2, 0) = w(\gamma_1, 0) + w(\gamma_2, 0).$$

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 $w(\gamma_1 + \gamma_2, 0) = w(\gamma_1, 0)$.

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.*

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 $w(\gamma_0, 0) = w(\gamma_1, 0)$.*

As before, let us write

$$\begin{aligned} \bar{D} &= \{z \in \mathbb{C} : |z| \leq 1\}, \\ D &= \{z \in \mathbb{C} : |z| < 1\}, \\ \partial D &= \{z \in \mathbb{C} : |z| = 1\}. \end{aligned}$$

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 $f(z) = 0$.

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.*

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* $z \in \partial D$.

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[19].

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.

[6] L. C. Thomas *Games, Theory and Applications*, Dover reprint of a book first published by Wiley in 1984.

[19] Though, like other good texts, it will not please those 'who want from books, plain cooking made still plainer by plain cooks.'

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. Check that, if we write

$$d(x,y) = \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y, \end{cases}$$

then (X, d) is a metric space (d is called the discrete metric).

(iv) Consider $X = \{0,1\}^n$. If we take

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

show that (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 the first two parts 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

$$\mathbb{H} = \{z \in \mathbb{C} : \Re z > 0\},$$
$$\bar{\mathbb{H}} = \{z \in \mathbb{C} : \Re z \geq 0\},$$
$$\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\},$$
$$\bar{\mathbb{D}} = \{z \in \mathbb{C} : |z| \leq 1\}.$$

(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 $[r = 0, 1, \ldots, n]$ 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 $\|g(\mathbf{x}) - \mathbf{x}\| \leq K$ for all $\mathbf{x} \in \mathbb{R}^2$.

(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

$$\begin{gathered} A \cup B \cup C = T, \\ A \cup B \supseteq I,\ B \cup C \supseteq J,\ C \cup A \supseteq K, \\ A \supseteq K \cap I,\ B \supseteq I \cap J,\ C \supseteq J \cap K. \end{gathered}$$

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

$$\frac{F(x+h) - 2F(x) + F(x-h)}{h^2} \to 0$$

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, 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[20], 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 $\mathbf{b} \in \mathbb{R}^m$. Show that

$$\{\mathbf{b} + \alpha\mathbf{x} \, : \, \mathbf{x} \in E\}$$

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

[20] This has nothing to do with the question but I cannot resist passing on the information that TfL employs hawks to clear its larger stations of pigeons. The hawks actually catch pigeons, but are not allowed to eat them, because London pigeons have too many diseases.

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 at most $2n$.

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.

(iv) If $f : \mathbb{R} \to \mathbb{R}$ is continuous, does there always exist a sequence of polynomials p_n with $p_n(x) \to f(x)$ for each x as $n \to \infty$.

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 Q_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 $n = 0, 1, 2 \dots, 4$ and the Chebychev polynomials U_{n-1} of the second kind for $n = 1, 2 \dots, 4$.

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 $A_n p_n$ (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 $\epsilon > 0$, 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 a 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 saw 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 $q \leq n$, 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

$$|f_n(t)| \le \epsilon$$

for all $t \in (a,b)$ and all $n \ge N(\epsilon)$.
Hint Consider the sets

$$E_N = \{x \in [0,1] : |f_n(x)| \le \epsilon, \text{for all } n \ge N\}.$$

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 $N \ge 1$ and

$$0 \le x_0 < x_1 < x_2 < \ldots < x_N \le 1$$

we have

$$\sum_{j=1}^{N} |f(x_{j-1}) - f(x_j)| \le M.$$

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 $M \ge 1$, we can find $N \ge 1$ and

$$0 \le x_0 < x_1 < x_2 < \ldots < x_N \le 1$$

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 $0 \le a < b \le 1$ and any $M \ge 1$, we can find $N \ge 1$ and

$$a \le x_0 < x_1 < x_2 < \ldots < x_N \le b$$

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 $\sqrt{3}$ 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?

(vii) What is the relation between τ and the σ of Exercise 15.9. Use the result of part (i) to obtain an estimate for

$$\frac{F_n}{F_{n+1}} - \sigma$$

correct to within a constant multiple of σ^{4n}.

Exercise 21.12. Let $p(z) = z^2 - 4z + 3$ and let $\gamma : [0,1] \to \mathbb{C}$ be given by $\gamma(t) = p(2e^{2\pi i t})$. 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