Topics in Analysis

Prof. W.T. Gowers

Michaelmas 2005

LaTeXed by Sebastian Pancratz

These notes are based on a course of lectures given by Prof. W.T. Gowers in Part II of the Mathematical Tripos at the University of Cambridge in the academic year 2005–2006.

These notes have not been checked by Prof. W.T. Gowers and should not be regarded as official notes for the course. In particular, the responsibility for any errors is mine — please email Sebastian Pancratz (`sfp25`) with any comments or corrections.

Contents

Chapter 1

Compactness

Definition. Let X be a metric space. A subset $U \subset X$ is *open* if for all $x \in U$ there exists $\delta > 0$ such that $B_\delta \subset U$.

Definition. An *open cover* of X is a collection $\{U_\gamma : \gamma \in \Gamma\}$ of open sets U_γ such that $X \subset \bigcup_{\gamma \in \Gamma} U_\gamma$.

Definition. X is *compact* if every open cover has a finite subcover.

Remark. If you want to cover $[0,1]^2$, rational points will not do as midpoints for balls. In fact, the Lebeque measure of the union of such balls can be made arbitrarily small.

Definition. X is *sequentially compact* if every sequence $(x_n)_{n=1}^\infty$ in X has a convergent subsequence.

Recall the Heinel–Borel theorem, stating that $[0,1]$ is compact, and the Bolzano–Weierstrass theorem, stating that $[0,1]$ is sequentially compact.

Example. $U \subset \mathbb{R}^n$ is compact and sequentially compact if and only if it is closed and bounded. This is not true for the infinite-dimensional Hilbert space l_2, e.g. take the vectors e_i for $1 \leq i < \infty$. This sequence has no convergent subsequence.

Fact. (i) $[0,1]$ is compact, and sequentially compact.

(ii) Any product of compact sets is compact (in any sensible metric on the product set, i.e. one that gives rise to the product topology).

(iii) A subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

(iv) A closed subset of a compact set is compact.

(v) A compact subset of a metric space is closed.

(vi) A subset of a metric space is compact if and only if it is sequentially compact.

(vii) Let X be a compact metric space and let $f : X \to \mathbb{R}$ be continuous. Then f is bounded and attains its bounds.

(viii) A continuous image of a compact set is compact.

(ix) Let X be a metric space, let $K \subset X$ be compact. Let $F \subset X$ be closed and suppose that $K \subset F = \emptyset$. Then there exists $\delta > 0$ such that $d(x,y) \geq 0$ for all $x \in K, y \in F$.

1.1 A proof of the intermediate value theorem via a compactness argument

Lemma 1.1 (Intermediate value theorem, discrete version)**.** Let $f : \{0, 1, 2, \ldots, n\} \to \mathbb{R}$ be a function such that $f(0) < 0$ and $f(n) \geq 0$. Then there exists m such that $f(m-1) < 0$ and $f(m) \geq 0$.

Proof 1. Let m be minimal such that $f(m) \geq 0$. Then $f(m-1) < 0$. □

Proof 2. Define a function $g : \{0, 1, \ldots, n\} \to \mathbb{R}$ by

$$g(m) = \begin{cases} 1 & \text{if } f(m) \geq 0 \\ 0 & \text{if } f(m) < 0 \end{cases}.$$

Then

$$1 = g(n) - g(0) = \sum_{m=1}^{n} g(m) - g(m-1).$$

Hence there exists m such that $g(m) - g(m-1) > 0$, so it equals 1, so $g(m) = 1$, $g(m-1) = 0$. □

Theorem 1.2 (Intermediate value theorem)**.** Let $a < b$ and let $F : [a, b] \to \mathbb{R}$ be continuous with $F(a) < 0$ and $F(b) > 0$. Then there exists $c \in (a, b)$ such that $F(c) = 0$.

Proof. For each n, define $f_n : \{0, 1, \ldots, n\} \to \mathbb{R}$ by

$$f_n(m) = F(a + \tfrac{m}{n}(b-a)).$$

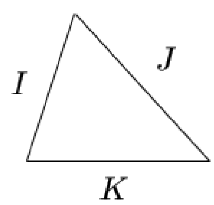
Then by Lemma 1.1 we can find some x_n (of the form $a + m/n(b-a)$) such that $F(x_n) \geq 0$ and $F(x_n - 1/n(b-a)) < 0$. Applying Bolzano–Weierstrass gives a subsequence $(x_{n_k})_{k=1}^{\infty}$ that converges to c, say.

Then $x_{n_k} - (b-a)/n_k \to c$ as well. Then $F(x_{n_k}) \to F(c)$ as F is continuous. So does $F(x_{n_k} - (b-a)/n_k)$. Since $F(x_{n_k}) \geq 0$ and $F(x_{n_k} - (b-a)/n_k) < 0$, $F(c) \geq 0$ and $F(c) \leq 0$. □

Theorem 1.3 (Brouwer's fixed-point theorem)**.** Let D be the closed unit disc in $\mathbb{R}^2$ and let $f : D \to D$ be continuous. Then there exists $x \in D$ such that $f(x) = x$.

Theorem 1.4 (No continuous retractions)**.** Let ∂D be the boundary of D. Then there is no continuous function $f : D \to \partial D$ such that $f(x) = x$.

Theorem 1.5. Let T be a triangle with sides I, J, K. Let A, B, C be closed sets such that $I \subset A$, $J \subset B$, $K \subset C$ and $T \subset A \cup B \cup C$. Then $A \cap B \cap C \neq \emptyset$.

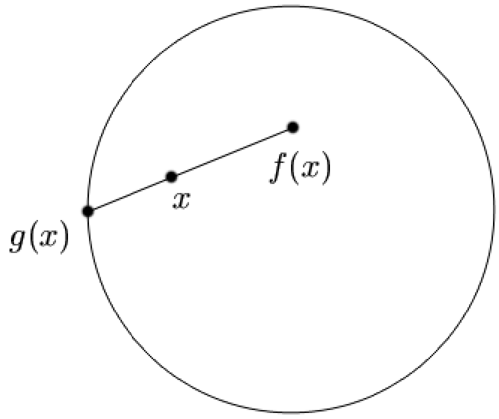


Theorem 1.6 (No continuous quasi-retractions)**.** There is no continuous function $f : T \to \partial T$ such that for all $f(I) \subset I$, $f(J) \subset J$ and $f(K) \subset K$.

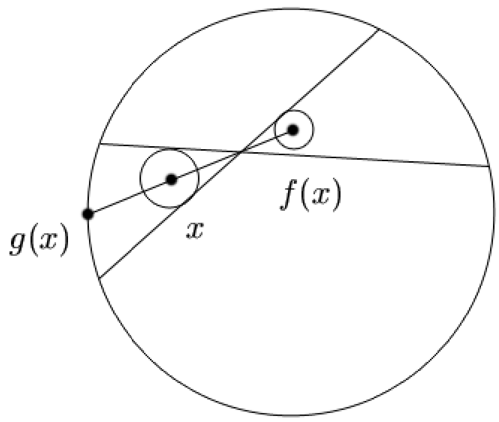
Now we shall prove that theorems 1.3, 1.4, 1.5, 1.6 are equivalent.

Proof. [1.3 $\implies$ 1.4] Suppose that $f : D \to \partial D$ is a continuous retraction. Then let ρ be a non-trivial rotation. Then, for all $x \in D$, $\rho \circ f(x) \neq x$. Also $\rho \circ f$ is continuous as ρ and f are.

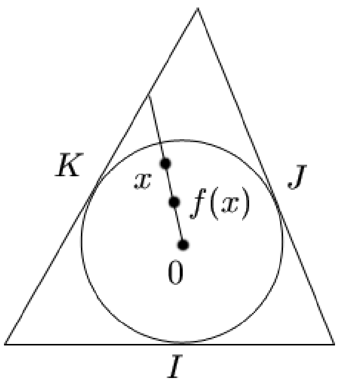
[1.4 $\implies$ 1.3] Let $f : D \to D$ be a continuous map with no fixed point. For each $x \in D$ let $g(x)$ be the point where the line from $f(x)$ to x hits the boundary.



The line is uniquely defined because $f(x) \neq x$. Then if $x \in \partial D$, $g(x) = x$. (Sketch argument to show that g is continuous: Let $\varepsilon > 0$, pick an "interval" on ∂D of radius $\varepsilon > 0$ around $g(x)$. Draw a cone with a vertex on the line segment from x to $f(x)$. Choose $\eta > 0$ such that $B_\eta(f(x))$ is inside the cone, choose $\delta > 0$ such that $f(B_\delta(x)) \subset B_\eta(f(x))$ and $B_\delta(x)$ inside the cone. Then $g(B_\delta(x)) \subset B_\varepsilon(g(x))$. See Example Sheet 1.)



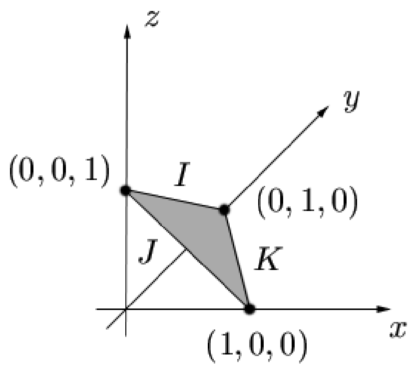
[1.3 $\implies$ 1.6] Suppose Theorem 1.6 is false.



(Draw the line from 0 through x to ∂T, take the ratio of the full line to the line segment from 0 to ∂D, multiply $\vec{0x}$ by this ratio.) Let $\phi : D \to T$ be a homeomorphism such that for every point $x \in D$, the line from the centre 0 of D to x is the same as the line from 0 to $\phi(x)$ and such that $\phi(\partial D) = \partial T$. Now let $g : D \to D$ be the map $\phi^{-1} f \phi$. If $x \in \phi^{-1}(I)$ (which is a 120° arc of ∂D) then $\phi(x) \in I$. Therefore, $f\phi(x) \in I$ (by property of f), hence $g(x) = \phi^{-1} f \phi(x) \in \phi^{-1}(I)$. So $g : \phi^{-1}(I) \to \phi^{-1}(I)$. Similarly for J and K. Also, for every $x \in D$, $\phi(x) \in T$, hence $f\phi(x) \in \partial T$, so $\phi^{-1} f \phi(x) \in \partial D$. Also, g is continuous. Let π be a 180° rotation. Then $\pi \circ g$ has no fixed point, contradicting Theorem 1.3.

[1.5 $\implies$ 1.6] Let $f : T \to \partial T$ be continuous functionsuch that $f(I) \subset I$, $f(J \subset J$, $f(K) \subset K$. Let $A = f^{-1}(I)$, $B = f^{-1}(J)$, $C = f^{-1}(K)$. Since I, J, K are closed and f is continuous, A, B, C are closed. For all $x \in T$, $f(x) \in \partial T = I \cup J \cup K$, so $A \cup B \cup C = T$. Also, since $I \cap J \cap K = \emptyset$, $A \cap B \cap C = \emptyset$, contradicting Theorem 1.5.

[1.6 $\implies$ 1.5] Let us regard T as the set $\{(x, y, z) \in \mathbb{R}^3 : x, y, z \geq 0 \ \wedge \ x + y + z = 1\}$.



Let I be the edge $x = 0$, J the edge $y = 0$, K the edge $z = 0$. (Thus, a typical point on I looks like $(0, \lambda, 1 - \lambda)$, $0 \leq \lambda \leq 1$.) Assume that A, B, C are closed sets, $T \subset A \cup B \cup C$, $I \subset A$, $J \subset B$, $K \subset C$, $A \cap B \cap C = \emptyset$. Given $x = (x, y, z) \in T$, let $f(x)$ be the point

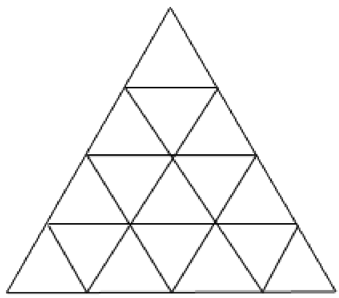
$$\frac{(d(x, A), d(x, B), d(x, C))}{d(x, A) + d(x, B) + d(x, C)}.$$

We must be sure that $d(x, A) + d(x, B) + d(x, C) \neq 0$. For each x, there is at least one of A, B, C to which is does not belong. Since A, B, C are closed, this implies that one of $d(x, A), d(x, B), d(x, C)$ is positive. Also, they are all non-negative. We therefore have that $f(x)$ is a well-defined point of ∂T. f is a quotient of continuous functions, the denominator of which is never 0, so f is continuous. If $x \in I$ then $x \in A$, so $d(x, A) = 0$, so $f(x) \in I$. Similarly for J, K, so we have contradicted Theorem 1.6.

[1.6 $\implies$ 1.4] Suppose that $f : D \to \partial D$ is a continuous retraction. Let $g : T \to \partial T$, $g = \phi f \phi^{-1}$, where $\phi : D \to T$ is the homeomorphism discussed earlier. g is a composition of continuous functions, so continuous. If $x \in T$ then $f\phi^{-1}(x) \in \partial D$, so $g(x) = \phi f \phi^{-1}(x) \in \partial T$. If $x \in \partial T$ then $\phi^{-1}(x) \in \partial D$, so $f\phi^{-1}(x) = \phi^{-1}(x)$, hence $g(x) = \phi\phi^{-1}(x) = x$. So g is a continuous retraction from T to ∂T. In particular, $g(I) \subset I$, $g(J) \subset J$, $g(K) \subset K$, contradicting Theorem 1.6. □

1.2 The discrete version of Theorem 1.5

Let T be a triangle. The nth triangular subdivision of T, or nth triangular grid, is the decomposition of T into smaller triangles, each an nth of the size of T (in each direction, some upside down).



The 4th triangular subdivision.

Theorem 1.7. Let T be a triangle with edges I, J and K and let it be divided up into a triangular grid. Let the vertices of T be coloured with three colours R, B, G (red, blue, green) such that every vertex in I is red or blue, every vertex in J is blue or green, and every vertex in K is green or red. Then there is a triangle of the grid such that all its vertices have different colours.

Proof. Given an edge of the grid, and a direction on that edge, we assign a value as follows.

Colour of u	Colour of v	Value of uv
R	R	0
R	B	1
R	G	-1
B	B	0
B	G	1
B	R	-1
G	G	0
G	R	1
G	B	-1

Suppose that a triangle is not multicoloured. Then the sum of the values of the edges going round anticlockwise is 0 (if all colours same, you get $0+0+0$; and if two same and one different, you get $0+1-1$). If it is multicoloured, then this sum is 3 or -3. Call this sum the value of the little triangle.

But we can write out the sum of the values of the triangles.

The contribution to this sum from any internal edge of the grid is 0, since it belongs to two triangles and is given opposite directions by them.

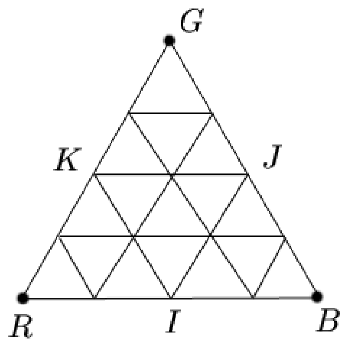
The sum of the contributions of the edges along a big edge of T is 1.

So the total sum for T equals 3. In particular, not all triangles have value 0, so there is a multicoloured triangle. □

Proof of Theorem 1.5. For each n, take the nth triangular subdivision of T. For each vertex x of the triangular grid, colour it in such a way that if x is coloured red then $x \in A$, if blue then $x \in B$, if green then $x \in C$. Do this in such a way that all vertices in I are red or blue, in J are blue or green, and in J are green or red, such that the conditions of the discrete theorem are satisfied.

Then we have a multicoloured triangle. Let its vertices be x_n, y_n, z_n. By Bolzano–Weierstrass, we can find a convergent subsequence (x_{n_k}) of (x_n). Suppose that $x_{n_k} \to x$. Since the radii of the triangles in the grid tend to 0 as $n \to \infty$, $y_{n_k} \to x$, $z_{n_k} \to x$. But $x_{n_k} \in A$, $y_{n_k} \in B$, $z_{n_k} \in C$, A, B, C closed, so $x \in A \cap B \cap C$. □

We will soon give a proof of the retraction version using the idea of winding numbers.



At $r = 0$, go around 0 zero times. At $r = 1$ go around 0 once. Somewhere there must be a jump from zero times to once. This is a contradiction if we show that this is a continuous process.

Lemma 1.8. Let X be a compact metric space, Y a metric space, $f : X \to Y$ continuous. Then f is uniformly continuous.

Proof 1. (Using sequential compactness.) If not, then

$$\exists \varepsilon > 0 \quad \forall \delta > 0 \quad \exists x, y \quad d(x, y) > \delta \;\wedge\; d(f(x), f(y)) \geq \varepsilon.$$

Fix such an ε. Then

$$\forall n \in \mathbb{N} \quad \exists x_n, y_n \quad d(x_n, y_n) < \tfrac{1}{n} \;\wedge\; d(f(x_n), f(y_n)) \geq \varepsilon.$$

Then the sequence (x_n) has a convergent subsequence $(x_{n_k})_{k=1}^{\infty}$. Let $x_{n_k} \to x$. Then $y_{n_k} \to x$ (since $d(x_{n_k}, y_{n_k}) < 1/n_k$). Since f is continuous, $f(x_{n_k}) \to f(x)$, $f(y_{n_k}) \to f(x)$, contradicting the fact that $d(f(x_{n_k}), f(y_{n_k})) \geq \varepsilon$ for all $k \in \mathbb{N}$. □

Proof 2. (Using compactness.) We know that

$$\forall \varepsilon > 0 \quad \forall x \quad \exists \delta_x > 0 \quad \forall y \quad d(x,y) < 2\delta_x \implies d(f(x), f(y)) < \tfrac{\varepsilon}{2}.$$

Let $\varepsilon > 0$. Then the balls $B_{\delta_x}(x)$ form an open cover of X. Since X is compact, we can find $x_1, \ldots, x_n$ such that $X = \bigcup_{i=1}^{n} B_{\delta_{x_i}}(x_i)$. Write δ_i for δ_{x_i}. Let $\delta = \min_{1 \leq i \leq n} \delta_i$.

Suppose that $d(x,y) < \delta$. We know that $x \in B_{\delta_i}(x_i)$ for some i. Since $d(x,y) < \delta \leq \delta_i$ we have $y \in B_{2\delta_i}(x_i)$ by the triangle inequality. It follows by our choice of δ_i that

$$d(f(x), f(x_i)) < \tfrac{\varepsilon}{2} \;\wedge\; d(f(y), f(x_i)) < \tfrac{\varepsilon}{2}.$$

So by the triangle inequality again $d(f(x), f(y)) < \varepsilon$. Since for every $\varepsilon > 0$ we found $\delta > 0$ with

$$d(x,y) < \delta \implies d(f(x), f(y)) < \varepsilon,$$

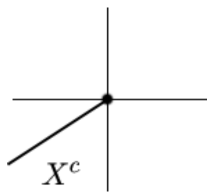
we have shown that f is uniformly continuous. □

1.3 The degree of a closed curve about a point

For convenience we shall identify $\mathbb{R}^2$ with $\mathbb{C}$ in the obvious way. However, we shall not use any ideas from *Complex Analysis*.

Let $z \in \mathbb{C}, z \neq 0$. A value of $\arg(z)$ is a real number θ such that $z = |z|e^{i\theta}$. Note that if θ is a value of $\arg(z)$ then ϕ is a value of $\arg(z)$ if and only if $(\phi - \theta)/(2\pi)$ is an integer.

Let $a \in \mathbb{R}$, let $X \subset \mathbb{C}$ be the set of all z of the form $re^{i\theta}$ with $r > 0$, $a - \pi < \theta < a + \pi$. For each $z \in X$ there is exactly one way of writing $z = re^{i\theta}$ with $a - \pi < \theta < a + \pi$. Write $\arg_a(z)$ for this θ, i.e. $\arg_a(z)$ is the unique value of $\arg(z)$ within π of a.



Lemma 1.9. The function $\arg_a$ is continuous on X.

Proof. Let $re^{i\theta}$ and $se^{i\phi}$ be two points in X. (We do not make assumptions about θ and ϕ yet.)

$$\begin{aligned} |re^{i\theta} - se^{i\phi}|^2 &= (re^{i\theta} - se^{i\phi})(re^{-i\theta} - se^{-i\phi}) \\ &= r^2 + s^2 - rs(e^{i(\theta-\phi)} + e^{-i(\theta-\phi)}) \\ &= (r-s)^2 + 2rs(1 - \cos(\theta - \phi). \end{aligned}$$

Hence if $|re^{i\theta} - se^{i\phi}|$ is sufficiently small, r and s must be approximately equal (so in particular $s > r/2$) and that $1 - \cos(\theta - \phi)$ is small, so $\theta - \phi$ is approximately a multiple of 2π. Therefore, if $\theta, \phi \in (a-\pi, a+\pi)$ we must have $\theta \approx \phi$. (Note that we first fix $re^{i\theta} \notin X^c$, then pick $\varepsilon > 0$ such that $B_\varepsilon(re^{i\theta})$ is strictly on one side of the branch cut, then $se^{i\phi}$ cannot be on the other of the branch cut, hence we cannot have $\theta \approx a - \pi$, $\phi \approx a + \pi$, so we must have $\theta \approx \phi$ instead.) □

Corollary 1.10. Let $f : [0,1] \to X$ be a continuous function. Then there is a continuous function $\theta : [0,1] \to \mathbb{R}$ such that, for every $t \in [0,1]$, $\theta(t)$ is a value of $\arg f(t)$.

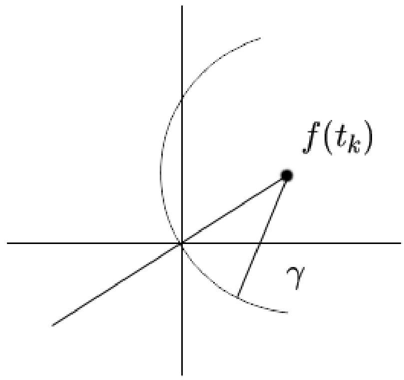
Proof. Define $\theta(t)$ to be $\arg_a(f(t))$. This is a composition of continuous functions, so it is continuous. □

Theorem 1.11. Let $f : [0,1] \to \mathbb{C} \setminus \{0\}$ be a continuous function. Then f has a continuous choice of argument. That is, there exists a continuous function $\theta : [0,1] \to \mathbb{R}$ such that $f(t) = |f(t)|e^{i\theta(t)}$ for every $t \in [0,1]$. Moreover, if θ, ϕ are any two such functions, then there exists $n \in \mathbb{Z}$ such that $\theta(t) - \phi(t) = 2n\pi \ \forall t \in [0,1]$.

Proof. Since $[0,1]$ is compact and f is continuous, $f([0,1])$ is compact, so it is closed. It does not contain 0, so there must exist $\gamma > 0$ such that $|f(t)| > \gamma \ \forall t \in [0,1]$. We also have that f is uniformly continuous, so we can find $\delta > 0$ such that

$$|s - t| < \delta \implies |f(s) - f(t)| < \gamma.$$

Let $0 = t_0 < t_1 < \cdots < t_n = 1$ be such that $t_i - t_{i-1} < \delta$ for all $i = 1, 2, \ldots, n$. We shall prove inductively that for every k there is a continuous choice of argument for the restriction of f to $[t_0, t_k]$. If $k = 0$, the result is trivial. Let $\theta(t_0)$ be any value of $\arg(f(0))$. Now suppose that θ is defined (and continuous etc.) up to t_k. Let $a = \theta(t_k)$, so $f(t_k) = re^{ia}$ for some $r > 0$ ($r = |f(t_k)|$). For every $t \in [t_k, t_{k+1}]$, we have $|f(t) - f(t_k)| < \gamma < |f(t_k)|$ by our choices of δ and γ.

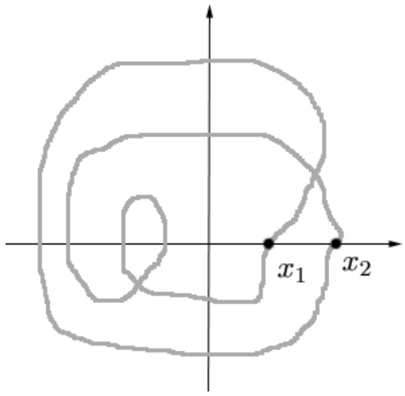


In particular, if we let X be the set of all numbers of the form $se^{i\phi}$ such that $s > 0$ and $a - pi < \phi < a + \pi$, then $f(1) \in X$ for every $t \in [t_k, t_{k+1}]$. Corollary 1.10 tells us that $\arg_a$ is a continuous choice of argument everywhere on $[t_k, t_{k+1}]$. Also, $\arg_a(f(t_k)) = a = \theta(t_k)$, so this continuous choice extends θ up to t_{k+1}, completing the induction.

(Uniqueness) If ϕ is another such function, then $\theta - \phi$ is continuous and $(\theta(t) - \phi(t))/(2\pi)$ is always an integer, since both $\theta(t)$ and $\phi(t)$ are values of $\arg f(t)$. Hence, by the intermediate value theorem, $\theta - \phi$ is constant (and takes value $2\pi n$ for some n). □

Definition. Let $f : [0,1] \to \mathbb{C} \setminus \{0\}$ be a continuous closed path, that is, $f(0) = f(1)$. Then the *winding number*, or *degree*, of f about 0 is $(\theta(1) - \theta(0))/2\pi$ for any continuous choice of θ of $\arg f(t)$. (The uniqueness part of Theorem 1.11 implies that this is well-defined.) We shall denote this by $\omega(f, 0)$. If z is some other point, then $\omega(f, z)$ can be defined as $\omega(f - z, 0)$.

Remark. This can be used as follows.



If we can rigorously prove what the points are at which the positive x-axis is crossed, then define a continuous choice of argument, e.g. from 0 to 2π between x_1 and x_2 and 2π to 4π between x_2 and x_1. Then rigorously $\omega(f, 0) = 2$.

Lemma 1.12. Let $f, h : [0,1] \to \mathbb{C} \setminus \{0\}$ be closed paths. Then $\omega(fh, 0) = \omega(f, 0) + \omega(h, 0)$.

Proof. Let θ and ϕ be continuous choices of argument for f and h, respectively. For each $t \in [0,1]$,

$$fh(t) = f(t)h(t) = |f(t)h(t)|e^{i\theta(h)}e^{i\phi(t)},$$

so $\theta + \phi$ is a continuous choice of argument for fh. Thus,

$$\begin{aligned}\omega(fh, 0) &= \frac{\theta(1) + \phi(1) - (\theta(0) + \phi(0))}{2\pi} \\ &= \frac{\theta(1) - \theta(0)}{2\pi} + \frac{\phi(1) - \phi(0)}{2\pi} \\ &= \omega(f, 0) + \omega(h, 0).\end{aligned}$$

□

Corollary 1.13 (Dog-walking lemma)**.** Let $f : [0,1] \to \mathbb{C} \setminus \{0\}$ be a closed path and let $g : [0,1] \to \mathbb{C}$ be a continuous function such that $g(0) = g(1)$ and $|g(t)| < |f(t)|$ for all $t \in [0,1]$. Then $\omega(f + g, 0) = \omega(f, 0)$.

Proof. Note

$$(f + g)(t) = f(t)\left(1 + \frac{g(t)}{f(t)}\right)$$

Let

$$h(t) = 1 + \frac{g(t)}{f(t)}$$

Then $\Re h(t) > 0$, since $|g(t)| < |f(t)|$. Therefore, $t \mapsto \arg_0(h(t))$ is a continuous choice of argument for h. But $\arg_0(h(1)) - \arg_0(h(0)) = 0$ since $h(1) = h(0)$, so $\omega(h, 0) = 0$. Therefore,

$$\omega(f+g,0) = \omega(fh,0) = \omega(f,0) + \omega(h,0) = \omega(f,0). \qquad \square$$

Definition. Let $f_0, f_1 : [0,1] \to \mathbb{C} \setminus \{0\}$ be two closed paths. A *homotopy* between f_0 and f_1 in $\mathbb{C}\setminus\{0\}$ is a continuous function $F : [0,1]^2 \to \mathbb{C}\setminus\{0\}$ such that $F(s,0) = F(s,1)$ for all s, and $F(0,t) = f_0(t)$ and $F(1,t) = f_1(t)$ for all t. It is sometimes convenient to write $f_s(t)$ for $F(s,t)$, which is an intermediate closed path. If such an F exists, then f_0 and f_1 are said to be *homotopic*.

Lemma 1.14. Let $f_0, f_1 : [0,1] \to \mathbb{C} \setminus \{0\}$ be closed paths. If they are homotopic then $\omega(f_0, 0) = \omega(f_1, 0)$.

Proof. Let $F : [0,1]^2 \to \mathbb{C} \setminus \{0\}$ be a homotopy between f_0 and f_1. Since $[0,1]^2$ is compact, $F([0,1]^2)$ is compact so closed. Hence there exists a $\gamma > 0$ such that $|F(s,t)| > \gamma$ for all s, t. Also, F is uniformly continuous. Therefore, there is some $\delta > 0$ such that

$$((s-s')^2 + (t-t')^2)^{1/2} < \delta \implies |F(s,t) - F(s',t')| < \gamma.$$

In particular, if $|s - s'| < \delta$ then $|F(s,t) - F(s',t)| < \gamma$. Write $f_s(t)$ for $F(s,t)$. Now let $0 = s_0 < \cdots < s_n = 1$ be such that $s_i - s_{i-1} < \delta$ for each i. Then, for each i and each t, $f_{s_i}(t) = f_{s_{i-1}}(t) + (f_{s_i}(t) - f_{s_{i-1}}(t))$, and $|f_{s_{i-1}}(t)| > \gamma$, $|f_{s_i}(t) - f_{s_{i-1}}(t)| < \gamma$, so by Corollary 1.13, $\omega(f_{s_i}, 0) = \omega(f_{s_{i-1}}, 0)$. Therefore, $\omega(f_0, 0) = \omega(f_1, 0)$. $\square$

Corollary 1.15. Let $f : D \to \mathbb{C}$ and let $g : [0,1] \to \mathbb{C}$ be defined by $g(t) = f(e^{2\pi i t})$. Then if $\omega(g, 0) \neq 0$, there must exist some $z \in D$ such that $f(z) = 0$.

Proof. Suppose that $f(z)$ is never 0. Define $F(s,t)$ to be $f(se^{2\pi i t})$ for every $(s,t) \in [0,1]^2$. Then $F(s,0) = F(s,1)$ for every s. F is continuous as the composition of continuous functions, $F(1,t) = g(t)$, $F(0,t)$ is constant at $f(0) \neq 0$. The winding number of a constant function is 0, since any value of arg of the constant function is a continuous choice. This contradicts Lemma 1.14. $\square$

Corollary 1.16. There is no continuous retraction from D to ∂D.

Proof. Let f be such a function, and let g be as in Corollary 1.15. Then $\omega(g, 0) = 1$, so there is some z such that $f(z) = 0$, contradicting the supposition that f is a retraction. $\square$

Corollary 1.17 (The Fundamental Theorem of Algebra)**.** Let $p : \mathbb{C} \to \mathbb{C}$ be a non-constant polynomial. Then there exists $z \in \mathbb{C}$ such that $p(z) = 0$.

Proof. Let $p(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$. Choose $R > 1$ such that $|a_n| R > |a_{n-1}| + \cdots |a_0|$. Then, if $|z| = R$, we have

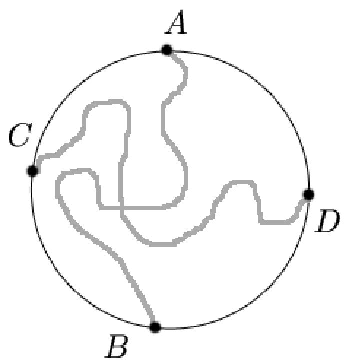
$$\begin{aligned} |a_n z^n| &= |a_n| R^n \\ &> (|a_{n-1}| + \cdots + |a_0|) R^{n-1} \\ &\geq |a_{n-1}| R^{n-1} + |a_{n-2}| R^{n-2} + \cdots + |a_1| R + |a_0| \\ &= |a_{n-1} z^{n-1}| + \cdots + |a_1 z| + |a_0| \end{aligned}$$

$$\geq |a_{n-1}z^{n-1} + \cdots + a_1 z + a_0|.$$

Now let $f : D \to \mathbb{C}$ be defined by $f(z) = p(Rz)$. Let $g(z) = a_n(Rz)^n$, $h(z) = a_{n-1}(Rz)^{n-1} + \cdots + a_1(Rz) + a_0$. Then $f(z) = g(z) + h(z)$ for every z, and $|g(z)| > |h(z)|$ whenever $|z| = 1$.

Let $u, v, w : [0,1] \to \mathbb{C}$ be defined by $u(t) = f(e^{2\pi it})$, $v(t) = g(e^{2\pi it})$, $w(t) = h(e^{2\pi it})$. Then by Corollary 1.13, $\omega(u, 0) = \omega(v, 0)$, since $|w(t)| < |v(t)|$ for all t. But $v(t) = a_n(Re^{2\pi it})^n = a_n R^n e^{2\pi int}$. Let $a_n = qe^{i\alpha}$. Then $v(t) = qR^n e^{i(2\pi nt+\alpha)}$, so $\theta(t) = 2\pi nt + \alpha$ is a continuous choice of argument for v. Therefore $\omega(v, 0) = n$. Hence, $\omega(u, 0) = n$, so by Corollary 1.15, there is some z such that $f(z) = 0$, so $p(Rz) = 0$. □

We also have another corollary. Let $f : [0,1] \to D$ from A to B, $g : [0,1] \to D$ from C to D be as in the picture. Then they have to cross somewhere.



If $f(s) \neq g(t)$ for all s, t then we can define $u : [0,1]^2 \to \partial D$ by setting $u(s,t)$ to be the point where the line from $f(s)$ to $g(t)$ hits ∂D.

Hence u has a zero in $[0,1]^2$. Therefore, $\omega(u|_{(\partial[0,1]^2)}, 0) = 1$, contradiction.

Chapter 2

Polynomial approximation

Our first target is Weierstrass's approximation theorem, which says that every continuous function in a closed bounded interval can be uniformly approximated by polynomials. The proof will use elementary probability, so we start by reviewing a few simple facts.

Lemma 2.1. Let X be a random variable with finite variance. Then $\operatorname{Var} X = \mathbb{E}\, X^2 - (\mathbb{E}\, X)^2$.

Proof. The variance of X is defined as $\mathbb{E}(X - \mathbb{E}\, X)^2$. But

$$\begin{aligned}\mathbb{E}(X - \mathbb{E}\, X)^2 &= \mathbb{E}(X^2 - 2X\,\mathbb{E}\, X + (\mathbb{E}\, X)^2) \\ &= \mathbb{E}\, X^2 - 2\,\mathbb{E}\, X\,\mathbb{E}\, X + (\mathbb{E}\, X)^2 \\ &= \mathbb{E}\, X^2 - (\mathbb{E}\, X)^2.\end{aligned}$$ □

Lemma 2.2. Let $X_1, \dots, X_n$ be pairwise independent random variables with finite variance and let $X = X_1 + \cdots + X_n$. Then

$$\operatorname{Var} X = \operatorname{Var} X_1 + \cdots + \operatorname{Var} X_n.$$

Proof.

$$\begin{aligned}\mathbb{E}\, X^2 &= \mathbb{E}(X_1 + \cdots + X_n)^2 \\ &= \sum_{i=1}^{n} \mathbb{E}\, X_i^2 + 2\sum_{i<j} \mathbb{E}(X_i X_j).\end{aligned}$$

Since $i < j$ implies X_i, X_j are independent,

$$2\sum_{i<j} \mathbb{E}(X_i X_j) = 2\sum_{i<j} \mathbb{E}\, X_i\,\mathbb{E}\, X_j.$$

Also

$$(\mathbb{E}\, X)^2 = (\mathbb{E}\, X_1 + \cdots + \mathbb{E}\, X_n)^2 = \sum_{i=1}^{n} (\mathbb{E}\, X_i)^2 + 2\sum_{i<j} \mathbb{E}\, X_i\,\mathbb{E}\, X_j$$

So

$$\mathbb{E}\, X^2 - (\mathbb{E}\, X)^2 = \sum_{i=1}^{n} (\mathbb{E}\, X_i^2 - (\mathbb{E}\, X_i)^2) = \sum_{i=1}^{n} \operatorname{Var} X_i.$$ □

Now let $X_1, \dots, X_n$ be independent Bernoulli variables with mean p. That is,

$$X_i = \begin{cases} 1 & \text{with probability } p \\ 0 & \text{with probability } 1-p \end{cases}.$$

Corollary 2.3. Let $X_1, \dots, X_n$ be as above, $X = X_1 + \dots + X_n$. Then $\mathbb{E}\, X = pn$ and $\operatorname{Var} X = p(1-p)n$.

Proof.

$$\begin{aligned} \mathbb{E}\, X &= \sum_{i=1}^{n} \mathbb{E}\, X_i = pn \\ \operatorname{Var} X &= \sum_{i=1}^{n} \operatorname{Var} X_i = n \operatorname{Var} X_1 \\ &= n(\mathbb{E}\, X_1^2 - (\mathbb{E}\, X_1)^2) \\ &= n(p - p^2) \\ &= np(1-p). \end{aligned}$$

□

Lemma 2.4 (Chebyshev's inequality)**.** Let X be a random variable with mean μ and variance σ^2. Then for every $C > 1$

$$\mathbb{P}[|X - \mu| \geq C\sigma] \leq \frac{1}{C^2}.$$

Proof.

$$\sigma^2 = \mathbb{E}(X-\mu)^2 \geq (C\sigma)^2\, \mathbb{P}[|X-\mu| \geq C\sigma]$$

e.g. with indicator functions, $|X - \mu| \geq C\sigma I_{|X-\mu| \geq C\sigma}$. Hence the result follows. □

Corollary 2.5. Let $0 \leq t \leq 1$, let $X_1, \dots, X_n$ be independent Bernoulli variables of mean t and let $Y_t = 1/n(X_1 + \dots + X_n)$. Then for every $\delta > 0$ we have

$$\mathbb{P}[|Y_t - t| \geq \delta] \leq \frac{1}{n\delta^2}.$$

Proof. The mean of Y_t is t and the variance is $(t(1-t))/n < 1/n$. Since $\delta = \sqrt{n}\delta/\sqrt{n}$, we can apply Chebyshev's inequality with $C = \sqrt{n}\delta$ which gives the result. □

Theorem 2.6 (Weierstrass's approximation theorem)**.** Let $f : [0,1] \to \mathbb{R}$ be continuous and let $\varepsilon > 0$. Then there exists a polynomial p such that $|p(t) - f(t)| < \varepsilon$ for every $t \in [0,1]$.

Proof. Let n be positive integer and let Y_t be as in Corollary 2.5. Then, as we shall show, $\mathbb{E}(f(Y_t))$ is a polynomial in t, and, if n is large enough,

$$|\mathbb{E}(f(Y_t)) - f(t)| < \varepsilon$$

for every $t \in [0,1]$.

To see that $\mathbb{E}(f(Y_t))$ is a polynomial, write it out. Notice that Y_t takes values in the set $\{0, \frac{1}{n}, \frac{2}{n}, \ldots, \frac{n-1}{n}, 1\}$. Hence

$$\begin{aligned}\mathbb{E}(f(Y_t)) &= \sum_{k=0}^{n} f(\tfrac{k}{n})\,\mathbb{P}(Y_t = \tfrac{k}{n}) \\ &= \sum_{k=0}^{n} f(\tfrac{k}{n})\binom{n}{k} t^k (1-t)^{n-k}\end{aligned}$$

which is a polynomial of degree n in t.

Why should we expect $\mathbb{E}(f(Y_t))$ to be close to $f(t)$? If n is large, then Corollary 2.5 implies that Y_t is close to t with high probability. Since f is continuous, this tells us that $f(Y_t)$ is close to $f(t)$ with high probability. To make this rigorous, we need to use compactness. As $[0,1]$ is compact and f is continuous, f is bounded and uniformly continuous. Hence, we can find M such that $|f(t)| \le M$ for every $t \in [0,1]$, and for given $\varepsilon > 0$ there exists $\delta > 0$ such that $|f(s) - f(t)| < \varepsilon/2$ whenever $|s-t| < \delta$.

If $|Y_t - t| < \delta$ then $|f(Y_t) - f(t)| < \varepsilon/2$ by our choice of δ. Therefore,

$$\mathbb{P}[|f(Y_t) - f(t)| \ge \tfrac{\varepsilon}{2}] \le \frac{1}{n\delta^2}.$$

Let us choose n such that $\frac{1}{n\delta^2} < \frac{\varepsilon}{4M}$. Then

$$\begin{aligned}|\mathbb{E}\, f(Y_t) - f(t)| &= |\mathbb{E}(f(Y_t)) - f(t)| \\ &\le \mathbb{E}|f(Y_t) - f(t)| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{4M} 2M = \varepsilon.\end{aligned}$$ □

Remark. There is a much more general version, the Stone–Weierstrass Theorem. Our proof does not carry over to general metric spaces and separating algebras of functions, though.

Corollary 2.7. Let $f : [0,1] \to \mathbb{R}$ be continuous. Then there is a sequence of polynomials $(p_n)_{n=1}^{\infty}$ such that $p_n \to f$ uniformly.

For the next part of the course a useful book is *Fourier Analysis* by T.W. Körner.

Lemma 2.8. Let $f : [a,b] \to \mathbb{R}$ be a continuous function and let p be a polynomial of degree at most n. Suppose there exist $a \le a_0 < a_1 < \cdots < a_{n+1} \le b$ such that, writing $r = \sup_{t\in[a,b]}|f(t) - p(t)|$, we have

$$\begin{aligned} & f(a_k) - p(a_k) = (-1)^k r && \text{for every } k \in \{0,1,\ldots,n+1\} \\ \text{or}\quad & f(a_k) - p(a_k) = (-1)^{k+1} r && \text{for every } k \in \{0,1,\ldots,n+1\}.\end{aligned}$$

Then p is the polynomial of degree at most n that minimises the distance $\sup_{t\in[a,b]}|f(t) - p(t)|$, i.e. it gives the best uniform approximation among all polynomials of degree at most n.

If p satisfies this condition, then it is said to satisfy the *equal-ripple criterion*.

Proof. Suppose that p has the property and that q is another polynomial of degree at most n that gives a strictly better approximation, i.e. $\sup_{t\in[a,b]}|f(t)-q(t)| < r$. Let us suppose that $f(a_k) - p(a_k) = (-1)^k r$ (the other case is similar). Then we know that for k even, $f(a_k) - p(a_k) = r$, so $f(a_k) - q(a_k) < r$ which implies that $p(a_k) < q(a_k)$. For k odd we have $f(a_k) - p(a_k) = -r$, so $f(a_k) - q(a_k) > -r$, so $q(a_k) < p(a_k)$. So between any a_k and a_{k+1}, $p(t) - q(t)$ changes sign. Hence, $p - q$ has a root in every interval (a_k, a_{k+1}), $k = 0, 1, \ldots, n$. So it has at least $n+1$ roots but degree at most n. So $p = q$, contradiction. □

Now let us try to find the polynomial of degree less than n that best approximates x^n on the interval $[-1, 1]$.

Lemma 2.9. $\cos n\theta$ is a polynomial of degree n in $\cos\theta$, with leading term 2^{n-1}.

Proof.

$$\begin{aligned}\cos n\theta &= \Re(e^{in\theta}) = \Re((\cos\theta + i\sin\theta)^n)\\ &= \cos^n\theta - \binom{n}{2}\cos^{n-2}\theta\sin^2\theta + \binom{n}{4}\cos^{n-4}\theta\sin^4\theta - \cdots\\ &= \cos^n\theta - \binom{n}{2}\cos^{n-2}\theta(1-\cos^2\theta)\\ &\quad + \binom{n}{4}\cos^{n-4}\theta(1-\cos^2\theta)^2 + \cdots\end{aligned}$$

which is a polynomial of degree n in $\cos\theta$.

The leading term is $1 + \binom{n}{2} + \binom{n}{4} + \binom{n}{6} + \cdots = 2^{n-1}$. For example, this can be seen as follows.

$$\begin{aligned}(1+1)^n &= 1 + \binom{n}{1} + \binom{n}{2} + \binom{n}{3} + \cdots\\ (1-1)^n &= 1 - \binom{n-1}{+}\binom{n}{2} - \binom{n}{3} + \cdots\\ 2^n &= 2(1 + \binom{n}{2} + \binom{n}{4} + \cdots).\end{aligned}$$

□

Definition. Let us write $T_n(\cos\theta) = \cos n\theta$. T_n is the nth *Chebyshev polynomial.*

From the definition we can read off a number of facts about the behaviour of $T_n(x)$ for $-1 \le x \le 1$.

For example, if $-1 \le x \le 1$ then $|T_n(x)| \le 1$. (Find θ such that $x = \cos\theta$. Then $T_n(x) = \cos n\theta$.)

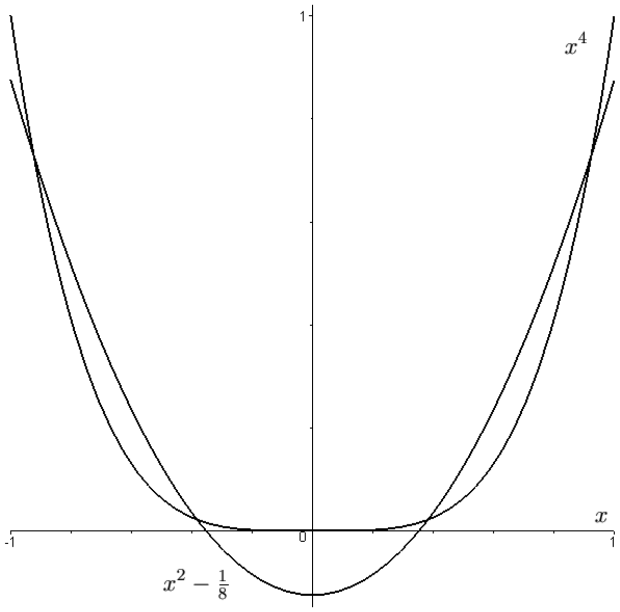
As θ increases from 0 to π, $\cos\theta$ decreases from 1 to -1. If $\theta = k\pi/n$ then $\cos n\theta = \cos k\pi = (-1)^k$. For $k = 0, 1, \ldots, n$ let us set $a_k = \cos(k\pi/n)$. Then $1 = a_0 > a_1 > \cdots > a_n = -1$. Also, $T_n(a_k) = (-1)^k$.

Now let us write $T_n(x) = 2^{n-1}x^n - S_n(x)$. Then S_n is a polynomial of degree less than n. Now S_n is the best uniform approximatio to $2^{n-1}x^n$ in the interval $[-1, 1]$, since $|2^{n-1}x^n - S_n(x)| \le 1$ for every $x \in [-1, 1]$ and we have just verified the equal-ripple criterion. Dividing through by 2^{n-1}, we find that $2^{-(n-1)}S_n$ is always within $2^{-(n-1)}$ of x^n in $[-1, 1]$, and that this is the best possible uniform approximation.

Example. Consider the case $n = 4$.

$$\cos 4\theta = \cos^4\theta - 6\cos^2\theta(1-\cos^2\theta) + (1-\cos^2\theta)^2$$
$$= 8\cos^4\theta - 8\cos^2\theta + 1$$

So, by the above general argument, the best uniform approximation to x^4 on $[-1,1]$ by a polynomial of degree less than 4 is $x^2 - \frac{1}{8}$.



Note that

$$x^4 - (x^2 - \tfrac{1}{8}) = \begin{cases} \frac{1}{8} & x = -1, 0, 1 \\ -\frac{1}{8} & x = \pm\frac{1}{\sqrt{2}} \end{cases},$$

and note $1/\sqrt{2}$ is "half way" between 0 and 1 in the $\cos\theta$ sense.

2.1 Legendre polynomials and numerical integeration

Lemma 2.10. There is a sequence of polynomials $p_0, p_1, \ldots$ such that for each $n \in \mathbb{N}$ p_n has degree n and

$$\int_{-1}^{1} p_m(t)p_n(t)\,dt = 0$$

whenever $m \neq n$. Moreover, the polynomials are unique up to scalar multiples, i.e. if $q_0, q_1, \ldots$ also work then there are constants $\lambda_0, \lambda_1, \ldots$ with $q_n(x) = \lambda_n p_n(x)$.

Proof. Let us write $\langle f, g\rangle$ for $\int_{-1}^{1} f(t)g(t)\,dt$, and $\|f\|$ for $\langle f, f\rangle^{1/2}$. Now let us choose inductively a sequence $p_0, p_1, \ldots$ with the required properties, together with the property that $\|p_n\| = 1$. Suppose we have chosen $p_0, p_1, \ldots, p_n$. (Note p_0 is the constant $1/\sqrt{2}$.) Let $f(t) = t^{n+1}$ and define $p'_{n+1} = f - \sum_{i=0}^{n}\langle f, p_i\rangle p_i$. Then, if $j \leq n$,

$$\langle p'_{n+1}, p_j\rangle = \langle f, p_j\rangle - \sum_{i=0}^{n}\langle f, p_i\rangle\langle p_i, p_j\rangle$$

$$= \langle f, p_j \rangle - \sum_{i=0}^{n} \langle f, p_i \rangle \delta_{ij}$$
$$= \langle f, p_j \rangle - \langle f, p_j \rangle = 0.$$

Now let
$$p_{n+1} = \frac{p'_{n+1}}{\|p'_{n+1}\|}.$$

Then $\|p_{n+1}\| = 1$.

To see that $\deg p_{n+1} = n+1$, note that the polynomials $p_0, p_1, \ldots$ are orthogonal, so linearly independent. Therefore, $p_0, \ldots, p_n$ span all polynomials of degree at most n. But p_{n+1} cannot lie in the space spanned by $p_0, \ldots, p_n$, so has degree greater than n.

For the uniqueness part, note that p_{n+1} has to satisfy $n+1$ independent linear conditions $\langle p_i, p_{n+1} \rangle = 0$ in an $(n+2)$-dimensional vector space. Therefore, it must lie in some fixed 1-dimensional subspace. □

Definition. p_n is called the nth *Legendre polynomial.*

Note that, since $\deg p_n = n$ for each n, the polynomials $p_0, \ldots, p_n$ span the space of all polynomials of degree at most n. (The relevant system of simultaneous equations gives rise to an upper triangular matrix with no zeros on the diagonal, so can be solved easily. Alternatively, note that $p_0, \ldots, p_n$ are orthogonal, so linearly independent, so they span by comparing dimensions.) As a consequence, we have the following very useful property of Legendre polynomials.
$$\int_{-1}^{1} f(t) p_n(t)\, dt = 0$$
whenever f is a polynomial of degree less than n. As we have just commented, f can be written as $\sum_{i=0}^{n-1} \lambda_i p_i$, so this follows from the fact that $\langle p_i, p_n \rangle = 0$ for $i < n$.

Now we are going to look at Gaussian quadrature, a method of approximating $\int_{-1}^{1} f(t)\, dt$ by expressions of the form $\sum_{i=1}^{n} \lambda_i f(a_i)$.

Lemma 2.11. Let $-1 \le a_1 < a_2 < \cdots < a_n \le 1$. Then there exist unique $\lambda_1, \ldots, \lambda_n$ such that
$$\int_{-1}^{1} f(t)\, dt = \sum_{i=1}^{n} \lambda_i f(a_i)$$
whenever f is a polynomial of degree less than n.

Proof. For each i let
$$r_i(t) = \frac{\prod_{j \ne i}(t - a_j)}{\prod_{j \ne i}(a_i - a_j}.$$
Then $r_i(a_j) = \delta_{ij}$ and $\deg(r_i) = n - 1$. Then

$$f(a_i) = \sum_{j=1}^{n} f(a_j) \delta_{ij}$$
$$= \sum_{j=1}^{n} f(a_j) r_j(a_i).$$

Hence, if we set

$$g(t) = \sum_{j=1}^{n} f(a_j) r_j(t),$$

we have that g is a polynomial of degree less than n, and $f(a_i) = g(a_i)$ for $i = 1, \dots, n$. So $f - g$ has n roots, so $f = g$. Therefore,

$$\begin{aligned}\int_{-1}^{1} f(t)\,dt &= \int_{-1}^{1} g(t)\,dt \\ &= \int_{-1}^{1} \sum_{j=1}^{n} f(a_j) r_j(t)\,dt \\ &= \sum_{j=1}^{n} \left(\int_{-1}^{1} r_j(t)\,dt \right) f(a_j).\end{aligned}$$

So we can take $\lambda_j = \int_{-1}^{1} r_j(t)\,dt$.

Now for the uniqueness part, let us suppose that $\lambda_1, \dots, \lambda_n$ are chosen. Then $\int_{-1}^{1} r_i(t)\,dt$ must be equal to

$$\sum_{j=1}^{n} \lambda_j r_i(a_j) = \lambda_i. \qquad \square$$

Claim. The n roots of p_n are distinct and lie in $[-1, 1]$.

Proof. Let $-1 \leq a_1 < \dots < a_k \leq 1$ be the distinct crossing roots in $[-1, 1]$, i.e. the points where the graph crosses the axis. Let $q(t) = \prod_{i=1}^{k} (t - a_i)$. Then q crosses the axis exactly where p_n does, so $q(t) p_n(t)$ never changes sign. So $\int_{-1}^{1} q(t) p_n(t)\,dt \neq 0$, a contradiction unless $k = n$. $\square$

Theorem 2.12. Let $-1 \leq a_1 < \dots < a_n \leq 1$ be the roots of the nth Legendre polynomial and let $\lambda_1, \dots, \lambda_n$ be as given by Lemma 2.11. Then

$$\int_{-1}^{1} f(t)\,dt = \sum_{i=1}^{n} \lambda_i f(a_i)$$

whenever f is a polynomial of degree less than $2n$.

Proof. If f is such a polynomial then we can write $f(t) = Q(t) p_n(t) + R(t)$ with $\deg Q, \deg R < n$. Therefore,

$$\int_{-1}^{1} f(t)\,dt = \int_{-1}^{1} Q(t) p_n(t)\,dt + \int_{-1}^{1} R(t)\,dt.$$

Since $\deg Q < n$, the first part is 0, so

$$\int_{-1}^{1} f(t)\,dt = \int_{-1}^{1} R(t)\,dt.$$

As for $\sum_{i=1}^{n} \lambda_i f(a_i)$, it is $\sum_{i=1}^{n} \lambda_i Q(a_i) p_n(a_i) + \sum_{i=1}^{n} \lambda_i R(a_i)$. But each $p_n(a_i) = 0$, so this is $\sum_{i=1}^{n} \lambda_i R(a_i)$. Since $\deg R < n$,

$$\int_{-1}^{1} R(t)\,dt = \sum_{i=1}^{n} \lambda_i R(a_i). \qquad \square$$

Theorem 2.13. Let $-1 \le a_1 < \cdots < a_n \le 1$ and $\lambda_1, \ldots, \lambda_n$ be such that

$$\int_{-1}^{1} f(t)\ dt = \sum_{i=1}^{n} \lambda_i f(a_i)$$

whenever f is a polynomial of degree less than $2n$. Then $a_1, \ldots, a_n$ are the roots of the nth Legendre polynomial p_n.

Proof. Let $p(t) = \prod_{i=1}^{n}(t - a_i)$. We would like to show that P is proportional to p_n. By the uniqueness of the construction of the p_n, it is enough to show that $\int_{-1}^{1} p(t)f(t)\ dt = 0$ whenever f is a polynomial of degree less than n.

But pf is a polynomial of degree less than $2n$, so

$$\int_{-1}^{1} p(t)f(t)\ dt = \sum_{i=1}^{n} \lambda_i p(a_i) f(a_i) = 0.$$

So the result follows. □

Lemma 2.14. Let $a_1, \ldots, a_n$ be the roots of the nth Legendre polynomial and let $\lambda_1, \ldots, \lambda_n$ be such that

$$\int_{-1}^{1} f(t)\ dt = \sum_{i=1}^{n} \lambda_i f(a_i)$$

for every polynomial f of degree less than $2n$. Then

(i) $\sum_{i=1}^{n} \lambda_i = 2$;

(ii) $\lambda_i \ge 0$ for all i.

Proof. (i) Take the polynomial $f(t) = 1$.

(ii) Let $f(t) = \prod_{j \ne i}(t - a_j)^2$. Then $f(a_j) = 0$ if $j \ne i$ and $f(t) \ge 0$ for all t. So $\int_{-1}^{1} f(t)\ dt > 0$ and $\sum_{j=1}^{n} f(a_j)\lambda_j = \lambda_i f(a_i)$. Since $f(a_i) > 0$, the result follows. □

Proposition 2.15. Let $f : [-1, 1] \to \mathbb{R}$ be a continuous function and let $\varepsilon > 0$. Then for every sufficiently large n, if $a_1, \ldots, a_n$, $\lambda_1, \ldots, \lambda_n$ are as above, then

$$\Big| \int_{-1}^{1} f(t)\ dt - \sum_{i=1}^{n} \lambda_i f(a_i) \Big| < \varepsilon.$$

Proof. By the Weierstrass approximation theorem there exists a polynomial p such that $|f(t) - p(t)| < \varepsilon/4$ for every $t \in [-1, 1]$. Let n be such that $\deg p < 2n$. Then

$$\begin{aligned} \Big| \int_{-1}^{1} f(t)\ dt - \sum_{i=1}^{n} \lambda_i f(a_i) \Big| \le{} & \Big| \int_{-1}^{1} f(t)\ dt - \int_{-1}^{1} p(t)\ dt \Big| \\ & + \Big| \int_{-1}^{1} p(t)\ dt - \sum_{i=1}^{n} \lambda_i p(a_i) \Big| \end{aligned}$$

$$
\begin{aligned}
&+\Big|\sum_{i=1}^{n}\lambda_i p(a_i)-\sum_{i=1}^{n}\lambda_i f(a_i)\Big| \\
&\le \int_{-1}^{1}|f(t)-p(t)|\,dt+0 \\
&+\Big(\sum_{i=1}^{n}|\lambda_i|\Big)\max_i|p(a_i)-f(a_i)| \\
&< 2\frac{\varepsilon}{4}+0+2\frac{\varepsilon}{4} \\
&= \varepsilon.
\end{aligned}
$$

□

Proposition 2.16. Let

$$f_n(t)=\frac{d^n}{dt^n}(1-t^2)^n.$$

Then f_n is a polynomial of degree n and if $m\neq n$ then $\int_{-1}^{1} f_m(t)f_n(t)\,dt=0$. Hence, f_n is proportional to the nth Legendre polynomial p_n.

Proof. $(1-t^2)^n$ is a polynomial of degree $2n$, which implies the first assertion. Now let us suppose that $m<n$ and calculate $\int_{-1}^{1} f_m(t)f_n(t)\,dt$ by parts.

$$
\begin{aligned}
\int_{-1}^{1} f_m(t)f_n(t)\,dt &= \int_{-1}^{1}\frac{d^m}{dt^m}(1-t^2)^m\frac{d^n}{dt^n}(1-t^2)^n\,dt \\
&= \left[\frac{d^m}{dt^m}(1-t^2)^m\frac{d^{n-1}}{dt^{n-1}}(1-t^2)^n\right]_{-1}^{1} \\
&\quad -\int_{-1}^{1}\frac{d^{m+1}}{dt^{m+1}}(1-t^2)^m\frac{d^{n-1}}{dt^{n-1}}(1-t^2)^n\,dt
\end{aligned}
$$

Since $(1-t^2)^n=(1-t)^n(1+t)^n$ has roots of multiplicity n at ±1, the $(n-1)$th derivative is zero there, as are the $(n-k)$th derivatives for all $k>0$. Therefore,

$$=-\int_{-1}^{1}\frac{d^{m+1}}{dt^{m+1}}(1-t^2)^m\frac{d^{n-1}}{dt^{n-1}}(1-t^2)^n\,dt.$$

Repeating this argument shows that for every $0\le k\le n$,

$$\int_{-1}^{1}\frac{d^m}{dt^m}(1-t^2)^m\frac{d^n}{dt^n}(1-t^2)^n\,dt=(-1)^k\int_{-1}^{1}\frac{d^{m+k}}{dt^{m+k}}(1-t^2)^m\frac{d^{n-k}}{dt^{n-k}}(1-t^2)^n\,dt.$$

Let $k=m+1$. Then

$$\frac{d^{m+k}}{dt^{m+k}}(1-t^2)^m=\frac{d^{2m+1}}{dt^{2m+1}}(1-t^2)^m=0$$

and the result follows. □

We can also work with weight functions to get Chebyshev or Hermite polynomials and others.

Chapter 3

Approximation by complex polynomials

Definition. A *path* in $\mathbb{C}$ was defined earlier as a continuous function from $[0,1]$ to $\mathbb{C}$, or more generally from $[a,b]$ to $\mathbb{C}$.

In *Complex Analysis* one restricts attention to continuous functions γ that are piecewise continuously differentiable, that is, we ask for sequence $a = a_0 < a_1 < \cdots < a_n = b$ such that for each k we ask for γ to be continuously differentiable on $[a_{k-1}, a_k]$, where the derivatives at a_{k-1} and a_k are right and left, respectively.

Definition. Now let us suppose that $\gamma : [a,b] \to \mathbb{C}$ is continuously differentiable and that $f : \mathbb{C} \to \mathbb{C}$ is a continuous function. Then the *path integral* $\int_\gamma f(z)\, dz$ is defined to be

$$\int_a^b f(\gamma(t))\gamma'(t)\, dt.$$

3.1 Why is the definition of the path integral natural?

Remark. Consider $f : \mathbb{C} \to \mathbb{C}$ and a continuously differentiable function $\gamma : [a,b] \to C$. We want the limit of the sum

$$\sum_{i=1}^{n} f(z_i)(z_i - z_{i-1}).$$

If $z_i = \gamma(t_i)$ with $a = t_0 < t_1 < \cdots < t_n = b$ and if $t_i - t_{i-1}$ is small, then

$$z_i - z_{i-1} = \gamma(t_i) - \gamma(t_{i-1}) = (t_i - t_{i-1})\gamma'(t_i)$$

since γ is continuously differentiable so γ' is approximately constant on $[t_{i-1}, t_i]$. So the sum

$$\sum_{i=1}^{n} f(z_i)(z_i - z_{i-1})$$

is roughly

$$\sum_{i=1}^{n} f(\gamma(t_i))\gamma'(t_i)(t_i - t_{i-1})$$

which we can write as

$$\sum_{i=1}^{n} f(\gamma(t_i))\gamma'(t_i)\delta t_i.$$

So in the limit we get

$$\int_a^b f(\gamma(t))\gamma'(t)\ dt.$$

Definition. A *domain* D in $\mathbb{C}$ is a connected open set of complex numbers. A function $f : D \to \mathbb{C}$ is *analytic* if it is complex-differentiable on all of D.

3.2 Cauchy's theorem

This has different versions, of different levels of generality. They all say that, under certain circumstances, if $f : D \to \mathbb{C}$ is an analytic function and C is a closed path then $\int_C f(z)\ dz = 0$. E.g. this is true if D is (i) a disc, (ii) convex, (iii) star-shaped, (iv) simply-connected. This last property means the following: every closed curve in D is homotopic to a point, or *contractible*, in D. E.g. an annulus is not simply connected.

To prove (iv) one proves the following important fact: if C_1 and C_2 are homotopic closed paths in D, and f is analytic, then

$$\int_{C_1} f(z)\ dz = \int_{C_2} f(z)\ dz,$$

i.e. the integral of f on C is a *homotopy invariant*.

3.3 The only path integral you ever need to calculate from first principles

Let C be a circle of radius r about 0, or more formally the path $\gamma : [0,1] \to \mathbb{C}, \gamma(t) = re^{2\pi it}$. Then

$$\int_C z^n\ dz = \begin{cases} 0 & n \neq -1 \\ 2\pi i & n = -1 \end{cases}$$

where $n \in \mathbb{Z}$. Indeed,

$$\begin{aligned} \int_C z^n\ dz &= \int_0^1 r^n e^{2\pi int} 2\pi i r e^{2\pi it}\ dt \\ &= r^{n+1} 2\pi i \int_0^1 e^{2\pi i(n+1)t}\ dt \\ &= r^{n+1} 2\pi i \left[\frac{e^{2\pi i(n+1)t}}{2\pi i(n+1)} \right]_0^1 \\ &= 0 \end{aligned}$$

using $n \neq -1$ in the penultimate step. If $n = -1$ we have $2\pi i \int_0^1\ dt = 2\pi i$.

Theorem 3.1 (Cauchy's integral formula)**.** Let D be a simply-connected domain, let $z \in D$ and let C be a closed path in D that winds once around z. Let $f : D \to \mathbb{C}$ be analytic. Then

$$f(z) = \frac{1}{2\pi i} \int_C \frac{f(w)}{w - z}\, dw.$$

Proof. It can be shown that C is homotopic in $D \setminus \{z\}$ to a path that goes round a small circle about z once. Once could assume this in the statement of the integral formula, it is just as useful.

The function $f(w)/(w - z)$ is analytic in $D \setminus \{z\}$, so homotopy invariance tells us that, for suitably small r,

$$\int_C \frac{f(w)}{w - z}\, dw = \int_{C_r} \frac{f(w)}{w - z}\, dw$$

where C_r is the circle of radius r about z. Let $\varepsilon > 0$, choose r such that $|w - z| \le r \implies |f(w) - f(z)| \le \varepsilon$. Then

$$\int_{C_r} \frac{f(w)}{w - z}\, dw = \int_{C_r} \frac{f(w) - f(z)}{w - z}\, dw + f(z) \int_{C_r} \frac{1}{w - z}\, dw.$$

By our earlier calculation, $\int_{C_r} 1/(w - z)\, dw = 2\pi i$. (The path is $w = z + re^{2\pi i t}$.) But

$$\begin{aligned}
\Big|\int_{C_r} \frac{f(w) - f(z)}{w - z}\, dw\Big| &\le \sup\Big\{\frac{f(w) - f(z)}{w - z} : w \in C_r\Big\} L(C_r) \\
&< \frac{\varepsilon}{r} 2\pi r \\
&= 2\pi\varepsilon.
\end{aligned}$$

Hence

$$\Big|\frac{1}{2\pi i} \int_{C_r} \frac{f(w)}{w - z} - f(z)\Big| \le \varepsilon$$

so

$$\Big|\frac{1}{2\pi i} \int_{C} \frac{f(w)}{w - z} - f(z)\Big| \le \varepsilon$$

for all $\varepsilon > 0$, giving the result. □

Lemma 3.2. Let $\sum_{n=0}^{\infty} a_n z^n$ be a power series with radius of convergence R and let $r < R$. Then the partial sums of $\sum_{n=0}^{\infty} a_n z^n$ converge uniformly on the closed disc of radius r about 0.

Proof. Let ρ be such that $r < \rho < R$. Then the sum $\sum_{n=0}^{\infty} a_n \rho^n$ converges, so in particular the numbers $a_n \rho^n$ are bounded. Let us say that $|a_n \rho^n| \le M$ for all n. Now let z be such that $|z| \le r$. Then

$$\begin{aligned}
\Big|\sum_{n=N}^{\infty} a_n z^n\Big| &= \Big|\sum_{n=N}^{\infty} a_n \rho^n \left(\frac{z}{\rho}\right)^n\Big| \\
&\le M \sum_{n=N}^{\infty} \Big|\frac{z}{\rho}\Big|^n
\end{aligned}$$

$$\begin{aligned}
&\le M \sum_{n=N}^{\infty} \left|\frac{r}{\rho}\right|^n \\
&= \frac{M\left(\frac{r}{\rho}\right)^N}{1-\frac{r}{\rho}}.
\end{aligned}$$

This tends to 0 as $N \to \infty$ and does not mention z, so the result is proved. □

Theorem 3.3. Let D be a domain and let $f : D \to \mathbb{C}$ be analytic. Let $z_0 \in D$. Then there exist unique coefficients $(a_n)_{n=0}^{\infty}$ with the following property: whenever the open ball $B_r(z_0) = \{z : |z - z_0| < r\}$ lies entirely in D, we have $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ for every $z \in B_r(z_0)$.

Proof. Let $z_0 \in D$ and let $r > 0$ be such that $B_r(z_0) \subset D$. Let $z \in B_r(z_0)$. Pick ρ such that $|z - z_0| < \rho < r$.

Let C be the path $t \mapsto z_0 + \rho e^{2\pi i t}$, i.e. a circle of radius ρ about z_0. By Cauchy's integral formula,

$$\begin{aligned}
f(z) &= \frac{1}{2\pi i}\int_C \frac{f(w)}{w-z}\,dw \\
&= \frac{1}{2\pi i}\int_C \frac{f(w)}{(w-z_0)-(z-z_0)}\,dw \\
&= \frac{1}{2\pi i}\int_C \frac{f(w)}{(w-z_0)(1-\frac{z-z_0}{w-z_0})}\,dw.
\end{aligned}$$

Since $|(z - z_0)/(w - z_0)| < 1$, this equals

$$= \frac{1}{2\pi i}\int_C \frac{f(w)}{w-z_0}\left(1+\frac{z-z_0}{w-z_0}+\left(\frac{z-z_0}{w-z_0}\right)^2+\cdots\right)dw.$$

The partial sums of this series converge uniformly, so we can interchange sums and integrals to obtain

$$= \sum_{n=0}^{\infty} \frac{1}{2\pi i}(z-z_0)^n \int_C \frac{f(w)}{(w-z_0)^{n+1}}\,dw.$$

So we can take

$$a_n = \frac{1}{2\pi i}\int_C \frac{f(w)}{(w-z_0)^{n+1}}\,dw.$$

To prove uniqueness, recall that the nth derivative of z_0 of the function $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ is $n!a_n$. So a_n must be $f^{(n)}(z_0)/n!$. □

Note that we obtain Cauchy's integral formula for derivatives,

$$f^{(n)}(z_0) = \frac{n!}{2\pi i}\int_C \frac{f(w)}{(w-z)^{n+1}}\,dw.$$

Remark. It follows from the above results that if f is analytic on an open disc $B_r(z_0)$ then for every $\rho < r$, f can be uniformly approximated by polynomials on the closed disc $\overline{B_\rho(z_0)} = \{z : |z - z_0| \le \rho\}$. For a proof, write $f(z) = \sum_{n=0}^{\infty} a_n(z - z_0)^n$ and take partial sums.

3.4 Runge's theorem and pointwise convergence of analytic functions

Theorem 3.4 (Runge's theorem)**.** Let K be a compact subset of $\mathbb{C}$ such that the complement $\mathbb{C} \setminus K$ is connected. Let $f : \mathbb{C} \to \mathbb{C} \cup \{\infty\}$ be a rational function whose poles all lie outside K. Then f can be uniformly approximated by polynomials on K.

Remark. This result follows from follows from the previous remark if there is an open disc $B_r(z_0)$ containing K and not containing any poles of f. This is because f is analytic on $B_r(z_0)$, and since K is compact one can find $\rho < r$ such that $K \subset \overline{B_\rho(z_0)}$.

Lemma 3.5. Let K be compact and let $f, g : K \to \mathbb{C}$ ba functions that can be uniformly approximated by polynomials. Then $\lambda f + \mu g$ can be uniformly approximated, as can fg. Hence, any polynomial in f, that is, a function $F(z) = a_n f(z)^n + \cdots + a_1 f(z) + a_0$ can be as well.

Proof. If p and q are polynomials with $\sup_{z\in K}|f(z)-p(z)| < \varepsilon$, $\sup_{z\in K}|g(z)-q(z)| < \varepsilon$, then

$$\sup_{z\in K}\big|\lambda f(z) + \mu g(z) - (\lambda p(z) + \mu q(z))\big| < \varepsilon(|\lambda| + |\mu|).$$

This proves the assertion about linear combinations. Since f and g can be uniformly approximated by continuous functions, they are continuous on the compact set K, so bounded on K. Suppose $|f(z)|, |g(z)| \leq M$ for all $z \in K$. Let p, q be polynomials as above. Then

$$\begin{aligned}
|f(z)g(z) - p(z)q(z)| &= |f(z)(g(z) - q(z)) + q(z)(f(z) - p(z))| \\
&\leq |f(z)||g(z) - q(z)| + |q(z)||f(z) - p(z)| \\
&< M\varepsilon + (M + \varepsilon)\varepsilon.
\end{aligned}$$

This does the part about products and hence polynomials. $\square$

Lemma 3.6. Let K be a compact subset of $\mathbb{C}$ and let $\lambda, \mu \in \mathbb{C}$ be such that $|\lambda - \mu| < d(\lambda, K) = \inf_{z\in K}|\lambda - z| = \min_{z\in K}|\lambda - z|$ as K is compact. Suppose that the function $1/(z - \lambda)$ can be uniformly approximated on K by polynomials. Then so can $1/(z - \mu)$.

Notation. We shall write "f is UAP" for "f can be uniformly approximated on K by polynomials".

Proof.

$$\frac{1}{z-\mu} = \frac{1}{(z-\lambda)-(\mu-\lambda)} = \frac{1}{z-\lambda}\frac{1}{1-\frac{\mu-\lambda}{z-\lambda}}$$

There exists $\rho < 1$ such that $|(\mu - \lambda)/(z - \lambda)| \leq \rho$ for every $z \in K$, because $|\lambda - \mu| < d(\lambda, K)$. Therefore,

$$\sum_{n=0}^{N}\left(\frac{\mu-\lambda}{z-\lambda}\right)^n \to \frac{1}{1-\frac{\mu-\lambda}{z-\lambda}}$$

uniformly on K. Let $\varepsilon > 0$. Then we can find N such that

$$|\sum_{n=0}^{N}\left(\frac{\mu-\lambda}{z-\lambda}\right)^n - \frac{1}{1-\frac{\mu-\lambda}{z-\lambda}}| < \frac{\varepsilon}{2}$$

for every $z \in K$. Since $1/(z-\lambda)$ is UAP, Lemma 3.5 tells us that $\sum_{n=0}^{N}((\mu-\lambda)/(z-\lambda))^n$ is UAP. So we can find a polynomials P such that

$$|P(z) - \sum_{n=0}^{N} \left(\frac{\mu-\lambda}{z-\lambda}\right)^n| < \frac{\varepsilon}{2}$$

for all $z \in K$. Therefore,

$$|P(z) - \frac{1}{1-\frac{\mu-\lambda}{z-\lambda}}| < \varepsilon$$

for all $z \in K$, so

$$\frac{1}{1-\frac{\mu-\lambda}{z-\lambda}}$$

is UAP. Hence, by Lemma 3.5, so is

$$\frac{1}{z-\lambda}\frac{1}{1-\frac{\mu-\lambda}{z-\lambda}} = \frac{1}{z-\mu}.$$

□

Corollary 3.7. Let $K \subset \mathbb{C}$ be compact and suppose that $\mathbb{C} \setminus K$ is connected. Then $1/(z-\lambda)$ is UAP on K for every $\lambda \neq K$.

Proof. We have already shown that if $B_r(z_0)$ is some open disc that contains K and $\lambda \notin B_r(z_0)$ then $1/(z-\lambda)$ is UAP on K. Hence, there exists some $\lambda \notin K$ such that $1/(z-\lambda)$ is UAP. Now let $\mu \notin K$. Since $\mathbb{C} \setminus K$ is path-connected, we can find a path from λ to μ, i.e. we can find a continuous function $\gamma : [0,1] \to \mathbb{C} \setminus K$ such that $\gamma(0) = \lambda$ and $\gamma(1) = \mu$.

The function $t \mapsto d(\gamma(t), K)$ is continuous, so it attains its minimum value. Hence, there exists $\varepsilon > 0$ such that $|\gamma(t) - z| \geq \varepsilon$ for all $t \in [0,1]$ and $z \in K$. Also, γ is uniformly continuous, so there exists $\delta > 0$ such that $|s-t| < \delta \implies |\gamma(s) - \gamma(t)| < \varepsilon$. So let $0 = t_0 < t_1 < \cdots < t_n = 1$ be such that $t_i - t_{i-1} \leq \delta$ for every i. Then let $\lambda_i = \gamma(t_i)$, so $\lambda_0 = \lambda$, $\lambda_n = \mu$. We then have $|\lambda_i - \lambda_{i-1}| < \varepsilon \leq d(\lambda_{i-1}, K)$. Hence, by Lemma 3.6, if $1/(z-\lambda_{i-1})$ is UAP then so is $1/(z-\lambda_i)$. But $1/(z-\lambda_0)$ is UAP, so this implies that $1/(z-\lambda_n) = 1/(z-\mu)$ is UAP. □

Proof of Theorem 3.4. Let $f(z) = P(z)/Q(z)$ be a rational function with poles at $\lambda_1, \ldots, \lambda_k$, $\lambda_i \notin K$. That is, let $f(z)$ be a function of the form

$$\frac{P(z)}{(z-\lambda_1)^{r_1} \cdots (z-\lambda_k)^{r_k}}.$$

By Corollary 3.7, $1/(z-\lambda_i)$ is UAP for each i. Also, P is UAP since P is a polynomial. So, by Lemma 3.5, f is UAP. □

Theorem 3.8. There exists a sequence of polynomials that converges pointwise on the closed unit disc $\{z : |z| \leq 1\}$ to a discontinuous function. In particular, pointwise limits of analytic functions need not be analytic.

Proof. For each $n \in \mathbb{N}$ let

$$J_n = \left\{ z \in \mathbb{C} : \frac{1}{n} \leq |z| \leq 1 \ \wedge\ \frac{2\pi}{n} < \arg z \leq 2\pi(1 - \frac{1}{n}) \right\}.$$

Let $[0,1]$ stand for $\{z \in \mathbb{C} : \Im z = 0 \ \wedge \ 0 \le \Re z \le 1\}$, let $K_n = J_n \cup [0,1]$. Let C_n be a sensible simply closed path that winds once round every point in J_n and zero times round every point in $[0,1]$.

Then define, for $z \notin C_n$,

$$f(z) = \frac{1}{2\pi i}\int_{C_n} \frac{1}{w-z}\,dw.$$

By standard results in complex analysis, $f(z) = 1$ whenever $z \in J_n$ and $f(z) = 0$ whenever $z \in [0,1]$. (This is the formula for the winding number, which follows e.g. from Cauchy's integral formula, Cauchy's theorem, and homotopy invariance of path integrals of analytic functions.)

If $\gamma : [a,b] \to \mathbb{C}$ is a continuously differentiable function that gives C_n then

$$f(z) = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t)-z}\,dt$$

by the definition of path-integration.

For each positive integer n, let

$$g_n(z) = \frac{1}{2\pi i}\frac{b-a}{n}\sum_{r=1}^{n} \frac{\gamma'(a+\frac{r}{n}(b-a))}{\gamma(a+\frac{r}{n}(b-a))-z}.$$

It can be shown that $g_n \to f$ uniformly on K_n (see Example Sheet 3). Relevant facts: γ and γ' are piecewise continuous and hence uniformly continuous. Also, there exists $\delta > 0$ such that $|\gamma(t) - z| > \delta$ for all $t \in [a,b]$ and $z \in K_n$. We may therefore choose n such that $|f(z) - g_n(z)| < \frac{1}{2n}$ for all $z \in K_n$. But g_n is a rational function with poles in C_n, hence not in K_n. So we can find, by Runge's theorem, a polynomial P_n such that $|P_n(z) - g_n(z)| < \frac{1}{2n}$ for all $z \in K_n$. So $|P_n(z) - f(z)| < \frac{1}{n}$ for all $z \in K_n$. But $\bigcup_{n=1}^{\infty} K_n = \{z : |z| \le 1\}$. So the polynomials P_n converge pointwise on the closed disc to

$$F(z) = \begin{cases} 0 & z \in [0,1] \\ 1 & \text{otherwise} \end{cases}.$$

□

Chapter 4

Irrationality, transcendence and continued fractions

Theorem 4.1. e is irrational.

Proof.

$$e = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \cdots$$

If $e = p/q$, then $q!e$ is an integer. But

$$\begin{aligned} q!e &= q! + q! + \frac{q!}{2!} + \cdots + \frac{q!}{q!} + \frac{q!}{(q+1)!} + \cdots \\ &= M + \frac{1}{q+1} + \frac{1}{(q+1)(q+2)} + \cdots \end{aligned}$$

for some integer M. So

$$M < q!e < M + \frac{1}{q+1} + \frac{1}{(q+1)^2} + \frac{1}{(q+1)^3} + \cdots = M + \frac{1}{q},$$

contradiction. □

Theorem 4.2. π is irrational.

Proof. Suppose that $\pi = p/q$, $p, q \in \mathbb{N}$. For each $n \in \mathbb{N}$, let

$$f_n(x) = \tfrac{1}{n!} q^n x^n (\pi - x)^n.$$

Since $\pi = p/q$, this can also be written as

$$f_n(x) = \tfrac{1}{n!} x^n (p - qx)^n.$$

We shall look at $\int_0^\pi f_n(x) \sin x \, dx$ in two different ways.

Since $n!f_n$ is a polynomial with integer coefficients, and has a factor x^n, $n!f_n^{(m)}(0) = 0$ if $m < n$ and is a multiple of $n!$ when $m \geq n$. Also, f_n has degree $2n$, so $f_n^{(m)}(0) = 0$ if $m > 2n$.

But $f_n(x) = f_n(\pi - x)$, so we can say the same for $f_n^{(m)}(\pi)$. Therefore, if we repeatedly integrate $\int_0^\pi f_n(x) \sin x \, dx$ by parts, we obtain integrals such as $\pm \int_0^\pi f_n^{(m)}(x) \sin x \, dx$ or $\pm \int_0^\pi f_n^{(m)}(x) \cos x \, dx$ and square-bracket parts of the form $\pm \big[f_n^{(m)}(x) \sin x \big]_0^\pi$ or

$\pm\left[f_n^{(m)}(x)\cos x\right]_0^\pi$. Since $f^{(m)}$, $\cos x$ and $\sin x$ are all integers at 0 and π, and $f_n^{(2n+1)}(x) = 0$, this shows that $\int_0^\pi f_n(x)\sin x\,dx$ is an integer.

However, $x(\pi - x)$ is maximised at $x = \pi/2$, so when $0 \leq x \leq \pi$, $0 \leq f_n(x) \leq q^n\left(\frac{\pi}{2}\right)^{2n}/n!$. Hence, for n large enough, we have $0 \leq f_n(x) \leq \frac{1}{2\pi}$ for every $0 \leq x \leq \pi$. Also, $\sin x \geq 0$ when $0 \leq x \leq \pi$ with equality only at 0 and π, so

$$0 < \int_0^\pi f_n(x)\sin x\,dx \leq 1\frac{1}{2\pi}\pi = \frac{1}{2},$$

contradiction. □

4.1 Continued fractions

Let t be a positivee real number. Then to work out its *continued-fraction expansion* you do the following: First write $t = t_0 = a_0 + s_0$ with $0 \leq s_0 < 1$, so $a_0 = \lfloor t \rfloor$. If $s_0 = 0$ then stop. Otherwise, $0 < s_0 < 1$, and we can write $t_0 = a_0 + 1/t_1$, where $t_1 = 1/s_0$. Now repeat the process for t_1, obtaining $t_0 = a_0 + 1/a_1$ or

$$t_0 = a_0 + \cfrac{1}{a_1 + \cfrac{1}{t_2}}.$$

Continue until some $s_n = 0$, or else, if this does not happen, go on for ever.

Remark. A nice book for this is *Strange Curves, Counting Rabbits and Other Mathematical Explorations* by Keith Ball, Princeton University Press.

Proposition 4.3. The continued-fraction expansion of a real number t terminates, i.e. is finite, if and only t is rational.

Proof. If the continued-fraction expansion terminates, then we have

$$t = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}$$

which is built out of the integers $a_0, a_1, \ldots, a_n$ by a mixture of addition and division. Since $\mathbb{Q}$ is closed under addition and division, $t \in \mathbb{Q}$.

Now suppose that $t = p/q$. Write $p = aq + b$ with $0 \leq b < q$. Then the first stage of the calculation is

$$\frac{p}{q} = a + \frac{b}{q} = a + \frac{1}{\frac{q}{b}},$$

so $t_0 = p/q$, $t_1 = q/b$. Notice that b, the denominator of t_1, is smaller than q, the denominator of t_0. Therefore, the process terminates as you cannot have an infinite decreasing sequence of denominators. □

Example.

$$\begin{aligned}\frac{79}{33} &= 2+\frac{13}{33}\\ &= 2+\cfrac{1}{\frac{33}{13}} = 2+\cfrac{1}{2+\frac{7}{13}}\\ &= 2+\cfrac{1}{2+\cfrac{1}{\frac{13}{7}}} = 2+\cfrac{1}{2+\cfrac{1}{1+\frac{6}{7}}}\\ &= 2+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{\frac{7}{6}}}} = 2+\cfrac{1}{2+\cfrac{1}{1+\cfrac{1}{1+\frac{1}{6}}}}\end{aligned}$$

Compare this with

$$\begin{aligned}79 &= 2\times 33+13\\ 33 &= 2\times 13+7\\ 13 &= 1\times 7+6\\ 7 &= 1\times 6+1\\ 6 &= 6\times 1\end{aligned}$$

Corollary 4.4. $\sqrt{2}$ is irrational.

Proof 1.

$$\begin{aligned}\sqrt{2} = 1+\sqrt{2}-1 = 1+\frac{1}{\sqrt{2}+1} &= 1+\frac{1}{2+\sqrt{2}-1}\\ &= 1+\cfrac{1}{2+\cfrac{1}{\sqrt{2}+1}}\end{aligned}$$

Since we have repeated the term $\sqrt{2}+1$, this process must go on for ever. Therefore, by Proposition 4.3, $\sqrt{2}$ is irrational. □

Proof 2. Suppose $\sqrt{2}$ is rational.

$$\sqrt{2} = \frac{p}{q} \implies 2 = \frac{p^2}{q^2} \implies \frac{2q}{p} = \frac{p}{q}.$$

Hence

$$\frac{2qa+pb}{pa+qb} = \frac{p}{q}$$

for all $a,b \in \mathbb{Z}$ such that $pa+qb \neq 0$. Suppose p/q is in its lowest terms. Then $q < p < 2q$, so $0 < p-q < q$. So

$$\frac{2q-p}{p-q}$$

is another expression for $\sqrt{2}$ with smaller denominator, contradiction. □

Proof 3. Yet another version of this argument: if $\sqrt{2}$ is rational, then write $\sqrt{2} = p/q$ with $\gcd(p,q) = 1$. Then find $a, b \in \mathbb{Z}$ such that $ap + bq = 1$ by Euclid's algorithm. Then

$$\sqrt{2} = \frac{p}{q} = \frac{2qa + pb}{ap + qb} = 2qa + pb \in \mathbb{Z},$$

contradiction. □

4.2 Continued fractions and matrices

Consider the expression

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}$$

To work it out, we start with a_n, take the reciprocal, add a_{n-1}, take the reciprocal etc. We get

$$a_n \to \frac{1}{a_n} \to a_{n-1} + \frac{1}{a_n} = \frac{a_{n-1}a_n + 1}{a_n}$$
$$\to \frac{a_n}{a_{n-1}a_n + 1} \to a_{n-2} + \frac{a_n}{a_{n-1}a_n + 1} = \frac{a_{n-2}a_{n-1}a_n + a_{n-2} + a_n}{a_{n-1}a_n + 1}$$

Let us write

$$\frac{r_k}{s_k} = a_k + \cfrac{1}{a_{k+1} + \cfrac{1}{a_{k+2} + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}$$

Then

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

So

$$\begin{pmatrix} r_k \\ s_k \end{pmatrix} = \begin{pmatrix} a_k r_{k+1} + s_{k+1} \\ r_{k+1} \end{pmatrix} = \begin{pmatrix} a_k & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_{k+1} \\ s_{k+1} \end{pmatrix}$$

Hence,

$$\begin{pmatrix} r_0 \\ s_0 \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ 1 \end{pmatrix}$$

Let us write

$$\frac{p_m}{q_m} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_m}}}}$$

Then

$$\begin{pmatrix} p_{n-1} \\ q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & 1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-2} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_{n-1} \\ 1 \end{pmatrix}$$

But

$$\begin{pmatrix} a_{n-1} \\ 1 \end{pmatrix} = \begin{pmatrix} a_{n-1} & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Therefore

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

Lemma 4.5. Let the fractions p_n/q_n be as above. Then

$$p_n q_{n-1} - q_n p_{n-1} = (-1)^{n+1}.$$

Proof. This follows immediately from the product rule for determinants. □

If we devide through by $q_{n-1}q_n$, we obtain

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n+1}}{q_{n-1}q_n}.$$

But whenever a/q_n, b/q_{n-1} are different fractions, we know that

$$\left|\frac{a}{q_n} - \frac{b}{q_{n-1}}\right| = \left|\frac{aq_{n-1} - bq_n}{q_{n-1}q_n}\right| \geq \frac{1}{q_{n-1}q_n}.$$

So p_{n-1}/q_{n-1} is approximating p_n/q_n as well as it can, given that its denominator is at most q_{n-1}.

Theorem 4.6. Let α be a positive real number, and suppose that

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots}}}$$

Let p_n, q_n be as above. Then p_n/q_n is closer to α than any other fraction with denominator at most q_n.

Proof. It follows easily from the recurrence relation

$$\begin{pmatrix} p_{n+1} & p_n \\ q_{n+1} & q_n \end{pmatrix} = \begin{pmatrix} p_n & p_{n-1} \\ q_n & q_{n-1} \end{pmatrix} \begin{pmatrix} a_{n+1} & 1 \\ 1 & 0 \end{pmatrix}$$

that p_n and q_n form increasing sequences. Therefore

$$\frac{p_n}{q_n} - \frac{p_{n-1}}{q_{n-1}} = \frac{(-1)^{n+1}}{q_{n-1}q_n}$$

is an alternating sequence with decreasing magnitude. Therefore, $p_1/q_1 > p_3/q_3 > p_5/q_5 > \cdots$ and $p_2/q_2 < p_4/q_4 < \cdots$, and $p_n/q_n > \alpha$ when n is odd, $p_n/q_n < \alpha$ when n is even. Also, since $(-1)^{n+1}/(q_n q_{n-1}) \to 0$, $p_n/q_n \to \alpha$.

Notice that p_n/q_n is closer to p_{n+1}/q_{n+1} than any other fraction of denominator at most q_n, since

$$\left|\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}\right| = \frac{1}{q_n q_{n+1}},$$

while

$$\left|\frac{a}{q} - \frac{b}{q_{n+1}}\right| \geq \frac{1}{q q_{n+1}}$$

whenever a/q, b/q_{n+1} are distinct fractions. Note $\gcd(p_{n+1}, q_{n+1}) = 1$, since $p_{n+1}q_n - q_{n+1}p_n = (-1)^n$. But any number that is closer to α than p_n/q_n is also closer to p_{n+1}/q_{n+1} than p_n/q_n since α is between p_n/q_n and p_{n+1}/q_{n+1}. Therefore, the result follows. $\square$

Example. The expansion of $\sqrt{2}$ starts

$$1 + \cfrac{1}{2 + \cfrac{1}{2 + \cfrac{1}{2}}} = 1 + \cfrac{1}{2 + \cfrac{1}{\frac{5}{2}}} = 1 + \cfrac{1}{\frac{12}{5}} = \frac{17}{12}$$

and $17^2 = 289$ and $2 \times 12^2 = 288$.

4.3 An expansion for $\tan x$

We shall consider more general expressions, of the form

$$a_0 + \cfrac{b_0}{a_1 + \cfrac{b_1}{a_2 + \cfrac{b_2}{\cdots + \cfrac{b_{n-1}}{a_n}}}}$$

Sometimes, if all the b_i are 1, we call the expression a *simple continued fraction*. We need to generalise the matrix formulation. So let us write

$$\frac{p_m}{q_m} = a_0 + \cfrac{b_0}{a_1 + \cfrac{b_1}{\dots + \cfrac{b_{m-1}}{a_m}}} \qquad \frac{r_k}{s_k} = a_k + \cfrac{b_k}{a_{k+1} + \cfrac{b_{k+1}}{\dots + \cfrac{b_{n-1}}{a_n}}}$$

Then

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

So

$$\begin{pmatrix} r_k \\ s_k \end{pmatrix} = \begin{pmatrix} a_k & b_k \\ 1 & 0 \end{pmatrix} \begin{pmatrix} r_{k+1} \\ s_{k+1} \end{pmatrix}$$

Therefore,

$$\begin{pmatrix} p_n \\ q_n \end{pmatrix} = \begin{pmatrix} r_0 \\ s_0 \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & b_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n \\ 1 \end{pmatrix}$$

So

$$\begin{aligned} \begin{pmatrix} p_{n-1} \\ q_{n-1} \end{pmatrix} &= \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} \\ 1 \end{pmatrix} \\ &= \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & b_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \end{aligned}$$

Hence,

$$\begin{pmatrix} p_n & b_n p_{n-1} \\ q_n & b_n q_{n-1} \end{pmatrix} = \begin{pmatrix} a_0 & b_0 \\ 1 & 0 \end{pmatrix} \cdots \begin{pmatrix} a_{n-1} & b_{n-1} \\ 1 & 0 \end{pmatrix} \begin{pmatrix} a_n & b_n \\ 1 & 0 \end{pmatrix}.$$

In particular,

$$\begin{pmatrix} p_n & b_n p_{n-1} \\ q_n & b_n q_{n-1} \end{pmatrix} = \begin{pmatrix} p_{n-1} & b_{n-1} p_{n-2} \\ q_{n-1} & b_{n-1} q_{n-2} \end{pmatrix} \begin{pmatrix} a_n & b_n \\ 1 & 0 \end{pmatrix},$$

giving

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

We shall now show that

$$\tan x = \cfrac{x}{1 - \cfrac{x^2}{3 - \cfrac{x^2}{5 - \cfrac{x^2}{7 - \cfrac{x^2}{\dots}}}}}$$

To do this, let

$$\frac{P_n(x)}{Q_n(x)} = \cfrac{x}{1 - \cfrac{x^2}{3 - \cfrac{x^2}{\dots}}}$$

at least for $|x| < \frac{\pi}{2}$. So, for example,

$$\begin{aligned}
\frac{P_1(x)}{Q_1(x)} &= \frac{x}{1} \\
\frac{P_2(x)}{Q_2(x)} &= \cfrac{x}{1 - \cfrac{x^2}{3}} = \frac{3x}{3 - x^2} \\
\frac{P_3(x)}{Q_3(x)} &= \cfrac{x}{1 - \cfrac{x^2}{3 - \cfrac{x^2}{5}}} = \cfrac{x}{1 - \cfrac{5x^2}{15 - x^2}} = \frac{15x - x^3}{15 - 6x^2}
\end{aligned}$$

In general,

$$\begin{aligned}
P_n(x) &= (2n-1)P_{n-1}(x) - x^2 P_{n-2}(x) \\
Q_n(x) &= (2n-1)Q_{n-1}(x) - x^2 Q_{n-2}(x)
\end{aligned}$$

by the recurrence formula for generalised continued fractions. We would like to prove that

$$\frac{P_n(x)}{Q_n(x)} \to \tan x$$

pointwise as $n \to \infty$. Now

$$\tan x - \frac{P_n(x)}{Q_n(x)} = \frac{Q_n(x)\sin x - P_n(x)\cos x}{Q_n(x)\cos x}.$$

Our main task will be to show that

$$s_n(x) = Q_n(x)\sin x - P_n(x)\cos x \to 0$$

as $n \to \infty$. To do this, we shall consider the following integral

$$I_n(x) = \frac{1}{2^n n!}\int_0^x (x^2 - t^2)^n \cos t \, dt.$$

If $n = 0$ we get $\int_0^x \cos t \, dt = \sin x$. We shall show that $I_n(x) = s_n(x)$ for $n \geq 1$. Our strategy will be to evaluate $I_n(x)$ for small n and to prove that $I_n(x) = (2n-1)I_{n-1}(x) - x^2 I_{n-2}(x)$.

$$\begin{aligned}
I_1(x) &= \tfrac{1}{2}\int_0^x (x^2 - t^2)\cos t \, dt \\
&= \tfrac{1}{2}\left[(x^2 - t^2)\sin t\right]_0^x - \tfrac{1}{2}\int_0^x (-2t)\sin t \, dt \\
&= \int_0^x t \sin t \, dt = [-t\cos t]_0^x + \int_0^x \cos t \, dt
\end{aligned}$$

$$= -x\cos x + \sin x = s_1(x)$$

For $n \geq 2$,

$$\begin{aligned}
I_n(x) &= \frac{1}{2^n n!}\int_0^x (x^2-t^2)^n \cos t\, dt \\
&= \frac{1}{2^n n!}\left[(x^2-t^2)^n \sin t\right]_0^x - \frac{1}{2^n n!}\int_0^x -2tn(x^2-t^2)^{n-1}\sin t\, dt \\
&= \frac{1}{2^{n-1}(n-1)!}\int_0^x t(x^2-t^2)^{n-1}\sin t\, dt \\
&= \frac{-1}{2^{n-1}(n-1)!}\left[t(x^2-t^2)^{n-1}\cos t\right]_0^x \\
&\quad + \frac{1}{2^{n-1}(n-1)!}\int_0^x \left((x^2-t^2)^{n-1} - 2t^2(n-1)(x^2-t^2)^{n-2}\right)\cos t\, dt \\
&= \frac{1}{2^{n-1}(n-1)!}\int_0^x (x^2-t^2)^{n-1}\cos t\, dt \\
&\quad - \frac{1}{2^{n-2}(n-2)!}\int_0^x t^2(x^2-t^2)^{n-2}\cos t\, dt \\
&= I_{n-1}(x) - x^2 I_{n-2}(x) + \frac{1}{2^{n-2}(n-2)!}\int_0^x (x^2-t^2)(x^2-t^2)^{n-2}\cos t\, dt \\
&= I_{n-1}(x) - x^2 I_{n-2}(x) + 2(n-1)I_{n-1}(x) \\
&= (2n-1)I_{n-1}(x) - x^2 I_{n-2}(x).
\end{aligned}$$

Therefore, $I_2(x) = 3I_1(x) - x^2 I_0(x) = 3(\sin x - x\cos x) - x^2\sin x = s_2(x)$. So $I_1(x) = s_1(x)$, $I_2(x) = s_2(x)$, $I_n(x) = (2n-1)I_{n-1}(x) - x^2 I_{n-2}(x)$ and $s_n(x) = (2n-1)s_{n-1} - x^2 s_{n-2}(x)$ for all $n \geq 3$. Hence $I_n(x) = s_n(x)$ for all $n \geq 1$. So

$$\begin{aligned}
|s_n(x)| &= \left|\frac{1}{2^n n!}\int_0^x (x^2-t^2)^n \cos t\, dt\right| \\
&\leq \frac{1}{2^n n!} x x^{2n} = \frac{x^{2n+1}}{2^n n!} \to 0
\end{aligned}$$

as $n \to \infty$. It remains to show that $Q_n(x)$ does not tend to 0 and is in fact bounded below. Let us write $r_n(x) = Q_n(x)/Q_{n-1}(x)$. Then $Q_n(x) = (2n-1)Q_{n-1}(x) - x^2 Q_{n-1}(x)$, so

$$\frac{Q_n(x)}{Q_{n-1}(x)} = 2n - 1 - x^2\frac{Q_{n-2}(x)}{Q_{n-1}(x)}$$

So

$$r_n(x) = 2n - 1 - \frac{x^2}{r_{n-1}(x)}.$$

If $|x| \leq 1$ and $|r_{n-1}(x)| \geq 1$, then $r_n(x) \geq 2n - 1 - 1 \geq 2$ when $n \geq 2$. Therefore, if ever $|r_{n-1}(x)| \geq 1$ with $n \geq 2$, we get $Q_n(x) \to \infty$ or $Q_n(x) \to -\infty$ as $n \to \infty$ for $|x| \leq 1$.

But if $|x| \leq 1$ then $Q_2(x)/Q_1(x) \geq 2$, so in this case we are done. With a little more care, one can show $Q_n(x) \to \pm\infty$ for every x. It follows that

$$\tan x = \cfrac{x}{1 - \cfrac{x^2}{3 - \cfrac{x^2}{5 - \cfrac{x^2}{\cdots}}}}$$

when $|x| \leq 1$, and with a bit more effort the argument shows that the expressionis valid for all x when $\cos x \neq 0$.

Theorem 4.7 (Liouville)**.** Let α be an algebraic number of degree $n > 1$. That is, n is the smallest integer such that $P(\alpha) = 0$ for some polynomial P of degree n with integer coefficients. Then there is $c > 0$, which may depend on α, such that

$$|\alpha - \frac{p}{q}| \geq \frac{c}{q^n}$$

for every pair of integers p, q with $q \neq 0$.

Proof. Let $P(x) = a_n x^n + \cdots + a_1 x + a_0$ be a polynomial of degree $n > 1$ with integer coefficients such that $P(\alpha) = 0$. Then $q^n P(p/q) = a_n p^n + a_{n-1}p^{n-1}q + \cdots + a_1 p q^{n-1} + a_0 q^n$ is a non-zero integer, unless p/q is a root of P. Let C_1 be the minimum difference between α and any rational root of P. (Note that there are at most n, but in fact none.) If p/q is not a root, we conclude that $|P(p/q)| \geq 1/q^n$. Therefore, $|P(\alpha) - P(p/q)| \geq 1/q^n$. But

$$\left|\alpha^k - \left(\frac{p}{q}\right)^k\right| \leq \left|\alpha - \frac{p}{q}\right|\left|\alpha^{k-1} + \alpha^{k-2}\frac{p}{q} + \cdots + \alpha\left(\frac{p}{q}\right)^{k-2} + \left(\frac{p}{q}\right)^{k-1}\right|.$$

If $|p/q| \leq 2|\alpha|$ then the second bracket is at most $2^{k-1}k\alpha^{k-1}$. Hence

$$\begin{aligned} \left|P(\alpha) - P(\frac{p}{q})\right| &\leq \left|\alpha - \frac{p}{q}\right|\Big(|a_1| + |a_2 \cdot 2 \cdot 2 \cdot \alpha| + |a_3 \cdot 4 \cdot 3 \cdot \alpha^2| \\ &\quad + |a_4 \cdot 8 \cdot 4 \cdot \alpha^3| + \cdots + |a_n \cdot 2^{n-1} \cdot n\alpha^{n-1}|\Big) \\ &=: |\alpha - \frac{p}{q}|A \end{aligned}$$

when $|p/q| \leq 2|\alpha|$ and p/q not a root of P. Therefore, $|\alpha - p/q|$ is always at least as big as one of c_1, $|\alpha|$ or c_2/q^n where $c_2 = 1/A$. So let $c = \min\{c_1, c_2, |\alpha|\}$. ☐

Proposition 4.8. For every irrational number α there is an infinite sequence of fractions p_n/q_n such that

$$|\alpha - \frac{p_n}{q_n}| \leq \frac{1}{q_n^2}.$$

Proof 1. Let

$$\alpha = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots}}}$$

and let

$$\frac{p_n}{q_n} = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{\cdots + \cfrac{1}{a_n}}}}$$

Then we have shown that α lies between p_n/q_n and p_{n+1}/q_{n+1}, and also that

$$|\frac{p_n}{q_n} - \frac{p_{n+1}}{q_{n+1}}| = \frac{1}{q_n q_{n+1}} \leq \frac{1}{q_n^2}.$$

Hence,

$$|\frac{p_n}{q_n} - \alpha| \leq \frac{1}{q_n^2}.$$

□

Proof 2. Fix a positive integer m. Then, of the numbers $0, \alpha, 2\alpha, \dots, m\alpha$, there must be two that have fractional part within $1/m$ of each other, since all fractional parts lie in $[0, 1)$. Suppose that these are $r\alpha$ and $s\alpha$. Then we can find positive integers a and b such that $|r\alpha - a - (s\alpha - b)| \leq 1/m$, so $|(r - s)\alpha - (a - b)| \leq 1/m$. Therefore,

$$|\alpha - \frac{a - b}{r - s}| \leq \frac{1}{m|r - s|} \leq \frac{1}{(r - s)^2}.$$

This proves the existence of one fraction of the required kind. But as $m \to \infty$ you cannot use the same fraction infinitely many times, using the irrationality of α, so we get an infinite sequence of them. □

Theorem 4.9. The number $\sum_{n=1}^{\infty} 10^{-n!}$ is transcendental.

Proof. Let $\alpha = \sum_{n=1}^{\infty} 10^{-n!}$. Suppose that α is algebraic of degree k. Then, by Theorem 4.7, there is some $c > 0$ such that $|\alpha - p/q| \geq c/q^k$ for all p, q. The number $\sum_{n=1}^{N} 10^{-n!}$ can be written as p/q with $q = 10^{N!}$. Also

$$\begin{aligned} |\alpha - \sum_{n=1}^{N} 10^{-n!}| &= \sum_{n=N+1}^{\infty} 10^{-n!} \\ &\leq 2 \cdot 10^{-(N+1)!} = \frac{2}{q^{N+1}}. \end{aligned}$$

This is a contradiction if $2/q^{N+1} < c/q^k$, so choose N large enough that $q^{N+1-k} > 2/c$, this is possible as $q = 10^{n!}$. □

Remark (Roth's theorem)**.** For all $\varepsilon > 0$ and for every algebraic number α there exists c such that

$$|\alpha - \frac{p}{q}| \geq \frac{c}{q^{2+\varepsilon}}$$

for all p, q, so Theorem 4.7 is not best possible.

Chapter 5

Properties of compact Hausdorff topological spaces

Definition. Let X be a set. A *topology* τ on X is a collection of subsets of X satisfying the following axioms,

(i) $\emptyset \in \tau$ and $X \in \tau$;

(ii) if $U_1, \ldots, U_n \in \tau$ then $\bigcap_{i=1}^n U_i \in \tau$;

(iii) if $U_\gamma \in \tau \ \forall \gamma \in \Gamma$ then $\bigcup_{\gamma \in \Gamma} U_\gamma \in \tau$.

The sets in τ are called *open*. A set $F \subset X$ is *closed* if $X \setminus F$ is open.

Example. (i) Any metric space with τ the collection of all open sets (in the metric space sense).

(ii) The *discrete topology* on X is the collection of all subsets of X.

(iii) The *indiscrete topology* on X is $\tau = \{\emptyset, X\}$.

(iv) The *cofinite topology* on X consists of all sets $Y \subset X$ such that $X \subset Y$ is finite or $Y = \emptyset$. (The co-countable topology is defined similarly.)

Definition. Let X, Y be topological spaces and let $f : X \to Y$. Then f is *continuous* if $f^{-1}(U)$ is open in X whenever U is open in Y. That is, the inverse image of any open set is open. If $x \in X$ then we say that f is *continuous at x* if for every open set V containing $f(x)$ there exists an open set U containing x such that $f(U) \subset V$.

Proposition 5.1. Let X and Y be topological spaces and let $f : X \to Y$. Then f is continuous if and only if f is continuous at every $x \in X$.

Proof. Suppose that f is continuous. Let $x \in X$ and let V be an open set containing $f(x)$. Then $f^{-1}(V)$ is open, and contains x, and maps into V.

Now suppose that f is continuous at every $x \in X$ and let $V \subset Y$ be open. We must show that $f^{-1}(V)$ is open. Let $x \in f^{-1}(V)$. Then $f(x) \in V$. Since f is continuous at x, we can find an open set U_x containing x such that $f(U_x) \subset V$. But then $\bigcup_{x \in f^{-1}(V)} U_x \subset f^{-1}(V)$ (since each $U_x \subset f^{-1}(V)$), $\bigcup_{x \in f^{-1}(V)} U_x \supset f^{-1}(V)$ (since it contains each $x \in f^{-1}(V)$) and is open (since each U_x is open). Therefore, $f^{-1}(V)$ is open. □

Definition. As with metric spaces, a topological space is *compact* if every open cover has a finite subcover. That is, X is compact if, whenever $X = \bigcup_{\gamma \in \Gamma} U_\gamma$ with each U_γ open, one can find $\gamma_1, \ldots, \gamma_n \in \Gamma$ such that $X \subset \bigcup_{i=1}^n U_{\gamma_i}$.

Definition. X is *Hausdorff* if, whenever $x, y \in X$, $x \neq y$, there exist disjoint open sets U, V with $x \in U$ and $y \in Y$.

Lemma 5.2. A closed subset of a compact topological space is compact.

Proof. Let X be a compact topological space and let $F \subset X$ be closed. Suppose that $F \subset \bigcup_{\gamma \in \Gamma} U_\gamma$ with the U_γ open. Then $X = (X \setminus F) \cup \bigcup_{\gamma \in \Gamma} U_\gamma$. Since X is compact, we can find a finite subcover of this cover. So we can find $\gamma_1, \dots, \gamma_n$ such that $X = (X \setminus F) \cup U_{\gamma_1} \cup \cdots \cup U_{\gamma_n}$, where, if the finite subcover did not include the open set $X \setminus F$, we put it in. But since $F \cap (X \setminus F) = \emptyset$, we have $F \subset U_{\gamma_1} \cup \cdots \cup U_{\gamma_n}$. □

Lemma 5.3. A continuous image of a compact topological space is compact.

Proof. Let X be compact and let $f : X \to Y$ be continuous. Suppose that $f(X) \subset \bigcup_{\gamma \in \Gamma} U_\gamma$ with U_γ open in Y. Then $X \subset \bigcup_{\gamma \in \Gamma} f^{-1}(U_\gamma)$, and each $f^{-1}(U_\gamma)$ is open, since f is continuous. Since X is compact, we can find $\gamma_1, \dots, \gamma_n$ such that $X = f^{-1}(U_{\gamma_1}) \cup \cdots \cup f^{-1}(U_{\gamma_n})$. But then $f(X) \subset U_{\gamma_1} \cup \cdots \cup U_{\gamma_n}$. □

Lemma 5.4. A compact subset of a Hausdorff topological space is closed.

Proof. Let X be Hausdorff and let $K \subset X$ be compact. Let $x \in X \setminus K$, noting there is nothing to prove if $X \setminus K = \emptyset$. For every $y \in K$ we can find open sets U_y, V_y such that $x \in U_y$, $y \in V_y$, $U_y \cap V_y = \emptyset$. The sets V_y form an open cover of K, so we can find $y_1, \dots, y_n$ such that $K \subset V_{y_1} \cup \cdots V_{y_n}$. But $(V_{y_1} \cup \cdots \cup V_{y_n}) \cap (U_{y_1} \cap \cdots \cap U_{y_n}) = \emptyset$, so $U_{y_1} \cap \cdots \cap U_{y_n}$ is an open set containing x disjoint from K. Since we can do this for every $x \notin K$, $X \setminus K$ is open. □

Theorem 5.5. Let X, Y be topological spaces with X compact and Y Hausdorff. Let $f : X \to Y$ be a continuous bijection. Then f^{-1} is continuous.

Proof. We need to show that if U is open in X then $f(U)$ is open in Y. But as U is open, $X \setminus U$ is closed, so by Lemma 5.2 $X \setminus U$ is compact. Hence, by Lemma 5.3 $f(X \setminus U)$ is compact. Therefore, Lemma 5.4 implies $f(X \setminus U)$ is closed. But $f(X \setminus U) = Y \setminus f(U)$ as f is a bijection. So $f(U)$ is open. □

Theorem 5.6. Let X be a compact Hausdorff topological space. Then no finer topology on X is compact and no coarser topology is Hausdorff. (If σ, τ are topologies on X such that $\tau \subset \sigma$, we say that σ is *finer* and τ is *coarser*.)

Proof. Let σ be a finer topology on X. Then (X, σ) is Hausdorff. Suppose it is compact. Let $f : (X, \sigma) \to (X, \tau)$ be the indentity on X. Then (X, σ) is compact, (X, τ) is Hausdorff, and clearly f is continuous since $\tau \subset \sigma$. Hence, by Theorem 5.5, f^{-1} is continuous. So if U is open in σ then U is open in τ, and therefore $\sigma = \tau$. The argument for coarser topologies is similar. □

Definition. A topological space is *regular* if, given any point x and any closed set F not containing x, there exist disjoint open sets U, V such that $x \in U$ and $F \subset V$. It is *normal* if, given any two disjoint closed sets F, G there exist disjoint open sets U, V such that $F \subset U$, $G \subset V$.

Definition. A topological space is T_1 if, given any two disjoint points x, y, there is an open set U such that $x \notin U, y \in U$ (so $\{x\}$ is closed).

Theorem 5.7. Every compact Hausdorff topological space is normal (and regular).

Proof. Let X be a compact Hausdorff topological space. First we shall show that X is regular, so let $x \in X$, let $F \subset X$ be closed and let $x \notin F$. For each $y \in F$ we can find disjoint open sets U_y, Y_y such that $x \in U_y, y \in V_y$. The sets V_y form an open cover of F, so, since F is closed and therefore compact, we can find $y_1, \ldots, y_n$ such that $F \subset V_{y_1} \cup \cdots \cup V_{y_n}$. But then let $U = U_{y_1} \cup \cdots \cup U_{y_n}, V = V_{y_1} \cup \cdots \cup V_{y_n}$. The sets U, V are disjoint and open, $x \in U$, $F \subset X$. So X is regular.

Now let F, G be disjoint closed subsets of X. For every $x \in F$, the regularity of X gives us disjoint open sets U_x, V_x with $x \in U_x, G \subset V_x$. The sets U_x form an open cover of F, so, as F is compact, we can find $x_1, \ldots, x_n$ such that $F \subset U_{x_1} \cup \cdots \cup U_{x_n}$. Let $U = U_{x_1} \cup \cdots \cup U_{x_n}, V = V_{x_1} \cap \cdots \cap V_{x_n}$. Then U and V are disjoint and open, $F \subset U, G \subset V$. $\square$

Chapter 6

The Baire category theorem

Definition. Let X be a topological space and let $Y \subset X$. Then Y is *dense* if $Y \cap U \neq \emptyset$ for every non-empty open set U. If V is an open set, then Y is *dense in* V if $Y \cap U \neq$ for every non-empty open subset $U \subset V$. Y is *nowhere dense* if there is no non-empty open set V with Y dense in V. That is, Y is nowhere dense if, for every non-empty open set V there is a non-empty open subset $U \subset V$ with $Y \cap U = \emptyset$.

6.1 A typical example of a nowhere dense set: the Cantor set

The Cantor set is a subset of $[0, 1]$ defined as follows: let $X_0 = [0, 1]$, $X_1 = [0, 1/3] \cup [2/3, 1]$, $X_2 = [0, 1/9] \cup [2/9, 1/3] \cup [2/3, 7/9] \cup [8/9, 1]$ etc. with each X_n removing the middle third of each of the 2^{n-1} intervals that make up X_{n-1}. Then $\bigcap_{n=0}^{\infty} X_n$ is the *Cantor set*. The see that this is nowhere dense, notice that every open interval contains a rational with denominator 3^n for some n, and to one side of any such rational is an interval disjoint from X_n and hence from the Cantor set.

Alternatively, the Cantor set consists of all points whose ternary expansion consists of just 0s and 2s.

Theorem 6.1 (Baire category theorem)**.** Let X be a complete metric space and let $U_1, U_2, \ldots$ be dense open subsets of X. Then $\bigcap_{n=1}^{\infty} U_n \neq \emptyset$.

Proof. We shall construct sequences $x_1, x_2, \ldots$ of points in X and $B_1 \supset B_2 \supset \cdots$ of closed balls in X such that for all $n \in \mathbb{N}$ $x_n \in B_n \subset U_n$ and the radius of B_n is at most $1/n$.

First, let $x_1 \in U_1$. Since U_1 is open, we can find $\varepsilon \leq 1$ such that $B_{\varepsilon_1}(x_1) \subset U_1$. (We write $B_\varepsilon(x)$ for $\{y : d(x, y) < \varepsilon\}$.) Let $0 < \delta_1 < \varepsilon$, then let $B_1 = \overline{B_{\delta_1}(x_1)}$. (It does not matter whether we consider the closed ball or the closure of the open ball.)

Suppose now that we have constructed $x_1, \ldots, x_{n-1}$ and $B_1 \supset \cdots \supset B_{n-1}$ with $B_{n-1} = \overline{B_{\delta_{n-1}}(x_{n-1})}$. Since U_n is dense and $B_{\delta_{n-1}}(x_{n-1})$ is open and non-empty, we have $U_n \cap B_{\delta_{n-1}}(x_{n-1}) \neq \emptyset$. So pick $x_n \in U_n \cap B_{\delta_{n-1}}(x_{n-1})$. Since $U_n \cap B_{\delta_{n-1}}(x_{n-1})$ is open, we can find $\varepsilon_n \leq 1/n$ such that $B_{\varepsilon_n}(x_n) \subset U_n \cap B_{\delta_{n-1}}(x_{n-1}) \subset U_n \cap B_{n-1}$. Let $0 < \delta_n < \varepsilon_n$ and set $B_n = \overline{B_{\delta_n}(x_n)}$. Then $x_n \in B_n \subset U_n$.

So now we have the sequences claimed. Now suppose that $N \leq p, N \leq q$. Then $x_p, x_q \in B_N$. Since B_N has radius at most $1/N$, $d(x_p, x_q) \leq 2/N$. Thus the sequence

$(x_n)_{n=1}^{\infty}$ is Cauchy. Let x be its limit. Let $n \in N$. Then for all $m \geq n$, $x_m \in B_n$. Hence, as B_n is closed, $x \in B_n$. But $B_n \subset U_n$, so $x \in \bigcap_{n=1}^{\infty} U_n$. □

Definition. Let X be a topological space and let $Y \subset X$. Then the *interior* of Y, denoted Y° or $\operatorname{int} Y$, is the largest open set contained in Y, that is, $\bigcup\{U \subset Y : U \text{ is open}\}$. The *closure* of Y, denoted $\bar{Y}$ or $\operatorname{cl} Y$, is the smallest closed set containing Y, that is, $\bigcap\{F \supset Y : F \text{ is closed}\}$. (In a metric space, $\bar{Y}$ is the set of all y such that there are $y_1, y_2, \ldots \in Y$ with $y_n \to y$.)

Lemma 6.2. The closure of a nowhere dense set is nowhere dense.

Proof. Let X be a topological space and let $Y \subset X$ be nowhere dense. We must show that $\bar{Y}$ is nowhere dense. So let $U \subset X$ be a non-empty open set. Since Y is nowhere dense, there is a non-empty open subset $V \subset U$ such that $Y \cap V = \emptyset$. But then $X \setminus V$ is a closed set containing Y. But then $\bar{Y} \subset X \setminus V$, so $\bar{Y} \cap V = \emptyset$. Since U was arbitrary, we have shown that $\bar{Y}$ is nowhere dense. □

Lemma 6.3. Let X be a topological space and let $U \subset X$. Then U is open and dense if and only if $X \setminus U$ is closed and nowhere dense.

Proof. Let $Y = X \setminus U$. If U is open and dense then Y is certainly closed. If W is a non-empty open subset of X then $U \cap W$ is a non-empty (as U is dense) and open (as U, W are open) subset of W that is disjoint from Y, so Y is nowhere dense.

Conversely, if Y is closed and nowhere dense, then U is open. For every non-empty open W there is a non-empty open $V \subset W$ such that $V \cap Y = \emptyset$, so $V \subset U$. So U is dense. □

Theorem 6.4 (Baire category theorem, version 2)**.** Let X be a complete metric space. Then X is not a countable union of nowhere dense sets.

Proof. Let $Y_1, Y_2, \ldots$ be nowhere dense. Then, by Lemma 6.2, $\bar{Y}_1, \bar{Y}_2, \ldots$ are also nowhere dense, and they are closed. So, by Lemma 6.3, $X \setminus \bar{Y}_1, X \setminus \bar{Y}_2, \ldots$ are open and dense. Therefore, by Theorem 6.1, $\bigcap_{n=1}^{\infty}(X \setminus \bar{Y}_n) \neq \emptyset$. So $\bigcup_{n=1}^{\infty} \bar{Y}_n \neq X$. So $\bigcup_{n=1}^{\infty} Y_n \neq X$. □

Remark. It is an easy exercise to use Lemmas 6.2 and 6.3 to deduce Theorem 6.1 from Theorem 6.4.

Theorem 6.5 (Baire category theorem, version 3)**.** Let X be a complete metric space and let $F_1, F_2, \ldots$ be closed sets such that $\bigcup_{n=1}^{\infty} F_n = X$. Then some F_n has a non-empty interior, i.e. contains an open ball.

Proof. Let F be a closed set with empty interior. If U is any non-empty open set, then $U \not\subset F$, so $U \setminus F$ is a non-empty open subset of U, disjoint from F. So F is nowhere dense. Hence, if no F_n has non-empty interior, all F_n are nowhere dense, contradicting Theorem 6.4. □

6.2 Applications

Definition. Let $\mathcal{F}$ be a set of functions defined on a set X. Then $\mathcal{F}$ is *pointwise bounded* on X if for all $x \in X$ there exists M such that for all $f \in \mathcal{F}$ we have $|f(x)| \leq M$. $\mathcal{F}$ is *uniformly bounded* on X if there exists M such that for all $f \in \mathcal{F}$ and $x \in X$ we have $|f(x)| \leq M$.

Theorem 6.6. Let $\mathcal{F}$ be a collection of continuous functions from $[0, 1]$ to $\mathbb{R}$. Suppose that $\mathcal{F}$ is pointwise bounded. Then there exists a subinterval (a, b) of $[0, 1]$ such that $\mathcal{F}$ is uniformly bounded on (a, b).

Proof. For each M let $Y_M = \{x \in [0, 1] : \forall f \in \mathcal{F}\ |f(x)| \leq M\}$. Notice that $Y_M = \bigcap_{f\in\mathcal{F}} f^{-1}([-M, M])$. This is closed since the functions f are continuous and $[-M, M]$ is closed. Our hypothesis is that every x belongs to some Y_M. Thus $[0, 1] = \bigcup_{M=1}^{\infty} Y_M$. Since $[0, 1]$ is complete, some Y_M must have non-empty interior, by Theorem 6.4. So it contains some interval (a, b). On that (a, b), every $f \in \mathcal{F}$ is bounded in modulus by M. □

Chapter 7

Example sheet questions revisited

Question (Sheet 2 Question 2)**.** Let $p(z)$ be the quadratic polynomial $z^2 - 4z + 3$ and let $f : [0, 1] \to \mathbb{C}$ be the closed path $f(t) = p(2e^{2\pi it})$. (Thus, the image of f is the image of the restriction of p to the circle of radius 2 and centre 0.)

(i) Calculate the winding number of f about 0 directly from the definition of winding number.

(ii) Can you give another proof that uses some of the results of the course?

Answer. Do not find a continuous choice of argument explicitly.

(i) Find the number of times the graph of p crosses the real positive axis. $p(z) = (z - 2)^2 - 1$ is real if $z - 2$ is real or purely imaginary, but $p(z) < 0$ for $z - 2$ purely imaginary. Note $z = 2e^{2\pi it}$ so $f(t)$ is real and positive if and only if $t = 1/2$. Therefore, for $0 \leq t < 1/2$ let $\arg z \in [0, 2\pi]$ and for $1/2 < t < 1$ let $\arg z \in (2\pi, 4\pi)$. Hence, the winding number about 0 is 1.

(ii) $p(z) = (z-1)(z-3)$. The winding number of a product is the sum of the winding numbers of the factors. Hence the winding number is $1 + 0 = 1$.

Question (Sheet 3 Question 2)**.** Prove that there is no sequence of analytic functions f_n that converges uniformly on the unit circle to the function $1/z$ (which on the circle is the same as $\bar{z}$). Why does this not contradict Runge's theorem?

Answer. Note that $\int_{\mathbb{T}} f_n(z)\ dz = 0$ but $\int_{\mathbb{T}} \bar{z}\ dz = 2\pi i$ (as e.g. $\bar{z} = z^{-1}$ on the unit circle).

Question (Sheet 3 Question 9)**.** Prove that $\sqrt{3} + \sqrt{5}$ and e^2 are irrational.

Answer.

$$e^2 = 1 + 2 + \frac{2}{2!} + \frac{2^3}{3!} + \cdots + \frac{2^m}{m!} + \frac{2^{m+1}}{(m+1)!} + \cdots$$

Multply by $m!/2^k$ for a suitable power of 2 such that this is an integer and such that $m!/2^k(2^{m+1}/(m+1)! + \cdots) < 1$. This is an obvious generalisation from e, but one needs to be careful.

Alternatively, let $e^2 = p/q$, then $p = e^2 q$, so $qe - pe^{-1} = 0$. Therefore,

$$q\Big(1 + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \cdots\Big) - p\Big(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \cdots\Big) = 0.$$

Now multiply through by some $m!$ to show that the left-hand side is non-zero.

Question (Sheet 4 Question 1). (i) For each n let F_n be the nth Fibonacci number, with $F_1 = F_2 = 1$. Prove the identities $F_{2n+1} = F_n^2 + F_{n+1}^2$ and $F_{2n} = F_n(F_{n-1} + F_{n+1})$. Let x_n stand for the ratio F_{n+1}/F_n. Use the above identities to express x_{2n} in terms of x_n.

(ii) If we set $y_k = x_{2^k}$, then we have expressed each y_k as a simple rational function of y_{k-1}. Explain why the sequence (y_k) converges very rapidly to the golden ratio. How does this relate to another method that produces rapid convergence?

Answer. To show $F_{2n+1} = F_n^2 + F_{n+1}^2$, $F_{2n} = F_n(F_{n-1} + F_{n+1})$, show that

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}.$$

Then

$$\begin{pmatrix} F_{2n+1} & F_{2n} \\ F_{2n} & F_{2n-1} \end{pmatrix} = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}^2.$$

Next, to show $x_n = F_{n+1}/F_n$, find x_{2n} in terms of x_n.

$$x_{2n} = \frac{F_{2n+1}}{F_{2n}} = \frac{F_n^2 + F_{n+1}^2}{F_n(F_{n-1} + F_{n+1})} = \frac{F_n^2 + F_{n+1}^2}{F_n(2F_{n+1} - F_n)} = \frac{1 + x_n^2}{2x_n - 1}$$

If we define $y_0, y_1, \ldots$ with $y_{n+1} = (1 + y_n^2)/(2y_n - 1)$, then $y_n = x_{2^n}$, so we jump exponentially fast to the golden ratio. This is the same as the Newton–Raphson method with $f(x) = (x^2 + 1)/(2x - 1)$.

Question (Sheet 4 Question 9). Let $\mathbb{T}$ be the unit circle $\{z \in \mathbb{C} : |z| = 1\}$ with the obvious topology coming from $\mathbb{C}$. Defined an equivalence relation $\sim$ on $\mathbb{R}$ by $x \sim y$ if and only if $x - y$ is an integer. Prove that $\mathbb{T}$ is homeomorphic to $\mathbb{R}/\sim$ with the quotient topology.

Answer. Consider $\mathbb{T} = \{z : |z| = 1\}$ and $\mathbb{R}/\mathbb{Z}$, with the topology induced by the quotient map $q : \mathbb{R} \to \mathbb{R}/\mathbb{Z}, t \mapsto [t]$. The map $[t] \mapsto e^{2\pi i t}$ is a homeomorphism as it is a continuous bijection from a compact set to a Hausdorff set.