Metric Spaces and Complex Analysis

Richard Earl

Michaelmas Term 2014
(Notes for lectures 1–16, and last year's notes for lectures 17 – 32)

SYLLABUS

Metric Spaces (10 lectures)

Basic definitions: metric spaces, isometries, continuous functions ($\varepsilon - \delta$ definition), homeomorphisms, open sets, closed sets. Examples of metric spaces, including metrics derived from a norm on a real vector space, particularly l^1, l^2, l^∞ norms on $\mathbb{R}^n$, the sup norm on the bounded real-valued functions on a set, and on the bounded continuous real-valued functions on a metric space. The characterization of continuity in terms of the pre-image of open sets or closed sets. The limit of a sequence of points in a metric space. A subset of a metric space inherits a metric. Discussion of open and closed sets in subspaces. The closure of a subset of a metric space. [3.5]

Completeness (but not completion). Completeness of the space of bounded real-valued functions on a set, equipped with the norm, and the completeness of the space of bounded continuous real-valued functions on a metric space, equipped with the metric. Lipschitz maps and contractions. Contraction Mapping Theorem. [2]

Connected metric spaces, path-connectedness. Closure of a connected space is connected, union of connected sets is connected if there is a non-empty intersection, continuous image of a connected space is connected. Path-connectedness implies connectedness. Connected open subset of a normed vector space is path-connected. [2]

Compactness. Heine-Borel theorem. The image of a compact set under a continuous map between metric spaces is compact. The equivalence of continuity and uniform continuity for functions on a compact metric space. Compact metric spaces are sequentially compact. Statement (but no proof) that sequentially compact metric spaces are compact. Compact metric spaces are complete. [2.5]

Reading

1. W. A. Sutherland, *Introduction to Metric and Topological Spaces* (Second Edition, OUP, 2009).

1. Metric Spaces. Convergence and Continuity.

Topology, loosely speaking, is the study of continuity. Consequently the objects of interest of topology are those properties that remain invariant under continuous deformations: so angle and area are out (being more rigid geometrical notions), but ideas of shape (e.g. connectedness = being in one piece) are in.

Quite what the rigorous definition of a continuous function $f\colon \mathbb{R} \to \mathbb{R}$ should be wasn't properly understood until the 19th century, with the work of Weierstrass and others, but the importance of having calculus and analysis on a rigorous footing was becoming very clear. Generalizing the definition of continuity to maps $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ was relatively straightforward but it would take much effort to fully generalize these definitions to metric and topological spaces (around the start of the 20th century), and to properly appreciate which properties of a "closed, bounded interval" had led to continuous functions being bounded and achieving those bounds, or to the Intermediate Value Theorem.

We first revisit the definition of convergence and continuity in Euclidean space.

Definition 1 *Given $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ the **distance** between $\mathbf{x}$ and $\mathbf{y}$ is*

$$|\mathbf{x}-\mathbf{y}| = \sqrt{\sum_{i=1}^{n} (x_i - y_i)^2}.$$

Now, with this notion of distance in $\mathbb{R}^n$, we can easily generalize the notions of convergence and continuity that we met in the first year. Namely:

Definition 2 *Let $(\mathbf{x}_k)$ be a sequence in $\mathbb{R}^n$ and $\mathbf{L} \in \mathbb{R}^n$. We say that $(\mathbf{x}_k)$ **converges to** $\mathbf{L}$ if*

$$|\mathbf{x}_k - \mathbf{L}| \to 0 \quad \text{as} \quad k \to \infty$$

or equivalently if

$$\forall \varepsilon > 0 \quad \exists K \in \mathbb{N} \quad \forall k \geqslant K \quad |\mathbf{x}_k - \mathbf{L}| < \varepsilon.$$

If $\mathbf{x}_k = \left(x_1^{(k)}, x_2^{(k)}, \ldots, x_n^{(k)}\right)$ and $\mathbf{L} = (L_1, L_2, \ldots, L_n)$ then $(\mathbf{x}_k)$ converging to $\mathbf{L}$ is equivalent to each sequence of coordinates $x_i^{(k)}$ converging to L_i as $k \to \infty$.

We can similarly generalize our definition of continuity:

Definition 3 *A function $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ is said to be **continuous at** $\mathbf{x} \in \mathbb{R}^n$ if*

$$\lim_{\mathbf{y}\to\mathbf{x}} \mathbf{f}(\mathbf{y}) = \mathbf{f}(\mathbf{x})$$

or equivalently if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{such that whenever } |\mathbf{y}-\mathbf{x}| < \delta \quad \text{then} \quad |\mathbf{f}(\mathbf{y}) - \mathbf{f}(\mathbf{x})| < \varepsilon.$$

From a topological point of view we will identify $\mathbb{C}$ with $\mathbb{R}^2$. (Though of course there are other important ways in which the two are different, for example when it comes to their algebra or geometry.) We start with the following simple fact:

Proposition 4 *$f\colon \mathbb{C} \to \mathbb{C}$ is continuous at z if and only if $\operatorname{Re} f$ and $\operatorname{Im} f$ are continuous at z.*

Proof Suppose that f is continuous at $z = x + iy$ and let $\varepsilon > 0$. Then there exists $\delta > 0$ such that

$$|z - w| < \delta \Longrightarrow |f(z) - f(w)| < \varepsilon.$$

Noting $|\operatorname{Re}\zeta| \leqslant |\zeta|$ and $|\operatorname{Im}\zeta| \leqslant |\zeta|$ for any $\zeta \in \mathbb{C}$, we have

$$|z - w| < \delta \Longrightarrow |(\operatorname{Re} f)(z) - (\operatorname{Re} f)(w)| = |\operatorname{Re}(f(z) - f(w))| < \varepsilon,$$

with a similar inequality for Im and hence $\operatorname{Re} f$ and $\operatorname{Im} f$ are both continuous at z. Conversely suppose that $\operatorname{Re} f$ and $\operatorname{Im} f$ are both continuous at z and let $\varepsilon > 0$. Then there exist positive δ_R and δ_I such that

$$|z - w| < \delta_R \implies |\operatorname{Re} f(z) - \operatorname{Re} f(w)| < \varepsilon/2; \qquad |z - w| < \delta_I \implies |\operatorname{Im} f(z) - \operatorname{Im} f(w)| < \varepsilon/2.$$

If we set $\delta = \min\{\delta_R, \delta_I\}$ then we have by the Triangle Inequality

$$\begin{aligned} |z - w| < \delta \implies |f(z) - f(w)| &= |(\operatorname{Re} f(z) - \operatorname{Re} f(w)) + i(\operatorname{Im} f(z) - \operatorname{Im} f(w))| \\ &\leqslant |\operatorname{Re} f(z) - \operatorname{Re} f(w)| + |\operatorname{Im} f(z) - \operatorname{Im} f(w)| \leqslant \varepsilon/2 + \varepsilon/2 = \varepsilon \end{aligned}$$

and hence f is continuous at z. ■

Here's a result from first year analysis that we see generalizes:

Proposition 5 ***(Uniqueness of Limits)*** *Let $(\mathbf{x}_k)$ be a sequence in $\mathbb{R}^n$ which converges to $\mathbf{L}$ and also to $\mathbf{M}$. Then $\mathbf{L} = \mathbf{M}$.*

Proof Suppose for a contradiction that $\mathbf{L} \neq \mathbf{M}$. Then we may set $\varepsilon = |\mathbf{L} - \mathbf{M}|/2 > 0$.

$$\begin{array}{rcl} \text{As } \mathbf{x}_k & \to & \mathbf{L} \text{ then there exists } K_1 \text{ such that } |\mathbf{x}_k - \mathbf{L}| < \varepsilon \text{ when } k \geqslant K_1; \\ \text{as } \mathbf{x}_k & \to & \mathbf{M} \text{ then there exists } K_2 \text{ such that } |\mathbf{x}_k - \mathbf{M}| < \varepsilon \text{ when } k \geqslant K_2. \end{array}$$

Hence for $k \geqslant \max\{K_1, K_2\}$ we have

$$\begin{aligned} |\mathbf{L} - \mathbf{M}| &= |(\mathbf{L} - \mathbf{x}_k) + (\mathbf{x}_k - \mathbf{M})| \leqslant |\mathbf{L} - \mathbf{x}_k| + |\mathbf{x}_k - \mathbf{M}| && (1.1) \\ &= |\mathbf{L} - \mathbf{x}_k| + |\mathbf{M} - \mathbf{x}_k| < \varepsilon + \varepsilon = |\mathbf{L} - \mathbf{M}|, && (1.2) \end{aligned}$$

which is our required contradiction. ■

In the hope of generalizing the definitions of continuity and convergence still further, we might ask ourselves this question: what underlying properties of distance are important? Whatever the setting, it seems important that limits should still be unique. We might not expect an "algebra of limits" or "a sandwich lemma"; depending on context these might not make sense. If we are interested in the question "does the Magnetic North Pole move continuously with time?" then algebra makes no sense as we can't reasonably add points on a sphere. But the above proposition *is* more fundamental topologically.

What properties of distance does the proof use? The following stand out:

- *Distances are nonnegative and distinct points are a positive distance apart.* We use this immediately in the first line.
- *Symmetry.* We make use of $|\mathbf{x}_k - \mathbf{M}| = |\mathbf{M} - \mathbf{x}_k|$ in (1.2).
- *Triangle Inequality.* We make use of this to get the first inequality in (1.1).

In light of this we define:

Definition 6 *Let M be a set. A **metric** on M is a map $d\colon M \times M \to [0, \infty)$ such that*

(M1) $d(m_1, m_2) = 0$ *if and only if* $m_1 = m_2$.

(M2) $d(m_1, m_2) = d(m_2, m_1)$ *for any* $m_1, m_2 \in M$.

(M3) $d(m_1, m_3) \leqslant d(m_1, m_2) + d(m_2, m_3)$ *for any* $m_1, m_2, m_3 \in M$.

In generalizing the distance properties to $\mathbb{R}^n$ we have met our first example of a metric space.

Example 7 *(**Euclidean Space**) Let $M = \mathbb{R}^n$ and define*

$$d_2(\mathbf{x}, \mathbf{y}) = \sqrt{(x_1 - y_1)^2 + \cdots + (x_n - y_n)^2}.$$

Example 8 *Let $M = \mathbb{R}^n$ and*

$$d_1(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + \cdots + |x_n - y_n|.$$

Example 9 *Let $M = \mathbb{R}^n$ and*

$$d_\infty(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, \ldots, |x_n - y_n|\}.$$

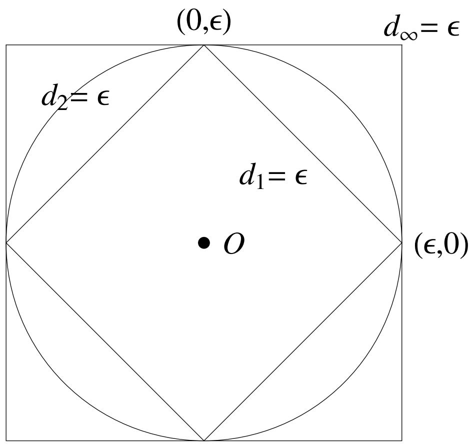
Proposition 10 *Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Show that*

$$d_2(\mathbf{x}, \mathbf{y}) \leqslant d_1(\mathbf{x}, \mathbf{y}) \leqslant \sqrt{n} d_2(\mathbf{x}, \mathbf{y}); \qquad \frac{1}{\sqrt{n}} d_2(\mathbf{x}, \mathbf{y}) \leqslant d_\infty(\mathbf{x}, \mathbf{y}) \leqslant d_2(\mathbf{x}, \mathbf{y}).$$

Proof This is seen most easily diagrammatically. Let $\varepsilon > 0$ and consider the sets

$$B_1 = \{\mathbf{x} \in \mathbb{R}^n : d_1(\mathbf{x}, \mathbf{0}) = \varepsilon\}; \quad B_2 = \{\mathbf{x} \in \mathbb{R}^n : d_2(\mathbf{x}, \mathbf{0}) = \varepsilon\}; \quad B_\infty = \{\mathbf{x} \in \mathbb{R}^n : d_\infty(\mathbf{x}, \mathbf{0}) = \varepsilon\}.$$

These sets are sketched below for $n = 2$.



Measuring with d_2, we see that the farthest B_1 is from $\mathbf{0}$ is ε at $(\varepsilon, 0, \ldots, 0)$ (say) and the nearest is $\varepsilon/\sqrt{n}$ at $(\varepsilon/n, \varepsilon/n, \ldots, \varepsilon/n)$ (say). Again measuring with d_2, we see that the farthest B_∞ is from $\mathbf{0}$ is $\sqrt{n}\varepsilon$ at $(\varepsilon, \varepsilon, \ldots, \varepsilon)$ (say) and the nearest is ε at $(\varepsilon, 0, \ldots, 0)$ (say). The result follows as the geometry of these sets is not affected by translation – i.e. we can assume $\mathbf{y} = \mathbf{0}$ without any loss of generality. ■

Example 11 *(**Discrete Metric**) Let M be any set and define d on M^2 by*

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

*Then d is a metric on M, referred to as the **discrete metric**. M1 and M2 are easily seen to be met and M3 can be verified on a case-by-case basis*

Case	$d(x,z)$	$d(x,y)+d(y,z)$	*Case*	$d(x,z)$	$d(x,y)+d(y,z)$
$x=y=z$	0	0	$x=z\neq y$	0	2
$x=y\neq z$	1	1	$x\neq y\neq z\neq x$	1	2
$x\neq y=z$	1	1			

Example 12 *Let V be a real inner product space. Distance is defined in V by*

$$d(x, y) = ||x - y||$$

*where $||v||^2 = \langle v, v \rangle$ for any $v \in V$ and this gives rise to a metric. We refer to $||v||$ as the **norm** of v.*

Example 13 *Let $C[a, b]$ denote the space of continuous functions $f \colon [a, b] \to \mathbb{R}$. As*

$$\langle f, g \rangle = \int_a^b f(t)g(t)\,\mathrm{d}t$$

defines an inner product on $C[a, b]$ then we have an associated metric

$$d_2(f, g) = \sqrt{\int_a^b (f(t) - g(t))^2\,\mathrm{d}t}.$$

Example 14 *A second metric on $C[a, b]$ is given by*

$$d_\infty(f, g) = \sup\{|f(x) - g(x)| : a \leqslant x \leqslant b\}.$$

M1 and M2 are clear. To show M3, note that for any $x \in [a, b]$ we have

$$\begin{aligned} |f(x) - h(x)| &\leqslant |f(x) - g(x)| + |g(x) - h(x)| \\ &\leqslant \sup_{a \leqslant x \leqslant b} |f(x) - g(x)| + \sup_{a \leqslant x \leqslant b} |g(x) - h(x)|. \end{aligned}$$

As sup *respects weak inequalities then*

$$\sup_{a \leqslant x \leqslant b} |f(x) - h(x)| \leqslant \sup_{a \leqslant x \leqslant b} |f(x) - g(x)| + \sup_{a \leqslant x \leqslant b} |g(x) - h(x)|$$

and so M3 follows for d_∞. More generally d_∞ defines a metric on the space of bounded real-valued functions on $[a, b]$.

Definition 15 *Let V be a real vector space. A* ***norm*** *on V is a map $|| \quad || : V \to [0, \infty)$ such that*

(N1) $||v|| = 0$ *if and only if* $v = 0_V$.

(N2) $||\lambda v|| = |\lambda| \times ||v||$ *for* $v \in V$ *and* $\lambda \in \mathbb{R}$.

(N3) $||x + y|| \leqslant ||x|| + ||y||$ *for* $x, y \in V$.

A vector space with a norm is called a ***normed vector space****. As the norm associated with an inner product satisfies N1, N2, N3 then inner product spaces are normed vector spaces.*

For any normed vector space, $d(x, y) = ||x - y||$ *defines a metric on* V.

Example 16 *(**Sequence Spaces**) We define now the following spaces of real sequences.*

$$\begin{aligned} l^1 &= \left\{ (x_n) : \sum |x_n| \ \text{ converges} \right\}, & ||(x_n)||_1 &= \sum |x_n| \, . \\ l^2 &= \left\{ (x_n) : \sum |x_n|^2 \ \text{ converges} \right\}, & ||(x_n)||_2 &= \sqrt{\sum |x_n|^2}. \\ l^\infty &= \{ (x_n) : |x_n| \ \text{ is bounded} \}, & ||(x_n)||_\infty &= \sup_n |x_n| \, . \end{aligned}$$

These spaces are all normed vector spaces. In fact l^2 is an inner product space with its norm coming from the inner product

$$\langle (x_n), (y_n) \rangle = \sum x_n y_n.$$

Example 17 *(**Subspace metrics**) Let (M, d) be a metric space and $A \subseteq M$. Then d naturally induces a metric d_A on A by setting*

$$d_A(a, b) = d(a, b) \qquad \text{for } a, b \in A.$$

Recall that our aim here was to put our notions of continuity and convergence on a more general footing. We can see that Definitions 2 and 3 can be naturally generalized as follows:

Definition 18 *Let (x_k) be a sequence in a metric space (M, d) and $x \in M$. We say that (x_k)* ***converges to*** *x if*

$$d(x_k, x) \to 0 \quad \text{as} \quad k \to \infty$$

or equivalently if

$$\forall \varepsilon > 0 \quad \exists K \in \mathbb{N} \quad \forall k \geqslant K \quad d(x_k, x) < \varepsilon.$$

Definition 19 *A function $f \colon M \to N$ between metric spaces (M, d_M) and (N, d_N) is said to be* ***continuous at*** *$x \in M$ if*

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{such that whenever } d_M(x, y) < \delta \quad \text{then} \quad d_N(f(x), f(y)) < \varepsilon.$$

Example 20 *Consider the sequence of functions $f_n(x) = x^n$ in $C[0, 1]$. Show that the sequence (f_n) converges in $(C[0, 1], d_2)$ but not in $(C[0, 1], d_\infty)$.*

Solution (i) Pointwise $f_n(x)$ converges to

$$f(x) = \begin{cases} 0 & 0 \leqslant x < 1 \\ 1 & x = 1 \end{cases}$$

but this is not continuous. However $g = 0$ is continuous and differs from f at only one point ($x = 1$). Note that

$$d_2(f_n, g)^2 = \int_0^1 (f_n(x) - g(x))^2 \, \mathrm{d}x = \int_0^1 x^{2n} \, \mathrm{d}x = \frac{1}{2n+1} \to 0$$

demonstrating that the sequence does indeed converge to g in the d_2-metric.

(ii) On the other hand suppose that there were a continuous function h such that f_n converges to h in the d_∞ metric. Then

$$\sup_{0 \leqslant x \leqslant 1} |f_n(x) - h(x)| \to 0 \quad \text{as } n \to \infty.$$

In particular this means that for any fixed x we have $|f_n(x) - h(x)| \to 0$ as $n \to \infty$. So $h(x) = \lim f_n(x) = f(x)$. However as f is not continuous we see that (f_n) is not a convergent sequence in $(C[0,1], d_\infty)$. ∎

Remark 21 *You may note that a sequence (g_n) converges to g in $(C[0,1], d_\infty)$ if and only if g_n converges uniformly to g.*

Example 22 *Show that the map $f \colon l^\infty \to l^2$ given by $f((x_n)) = (x_n/n)$ is well-defined and continuous.*

Solution If $(x_n) \in l^\infty$ then x_n is bounded, by M say. So

$$||(x_n/n)||_2 = \sqrt{\sum \frac{|x_n|^2}{n^2}} \leqslant M \sqrt{\sum \frac{1}{n^2}} = \frac{M\pi}{\sqrt{6}} < \infty$$

and $(x_n/n) \in l^2$. Now let $\varepsilon > 0$ and set $\delta = \sqrt{6}\varepsilon/\pi$. If $||(y_n) - (x_n)||_\infty < \delta$ then $|y_n - x_n| < \delta$ for each n. Hence

$$||f((y_n)) - f((x_n))||_2 = ||(y_n/n) - (x_n/n)||_2 = \sqrt{\sum \frac{|y_n - x_n|^2}{n^2}} < \delta \sqrt{\sum \frac{1}{n^2}} = \frac{\delta\pi}{\sqrt{6}} = \varepsilon.$$

∎

Proposition 23 *Let (M, d) be a metric space and $a \in M$. Define $f \colon M \to \mathbb{R}$ by $f(x) = d(x, a)$. Then f is continuous.*

Proof Let $\varepsilon > 0$ and $x \in M$. Set $\delta = \varepsilon$ and take $y \in M$ such that $d(x, y) < \delta$. Then

$$|f(y) - f(x)| = |d(y, a) - d(x, a)| \leqslant d(x, y) < \delta = \varepsilon$$

by the triangle inequality. ∎

We finish by generalizing a result of first year analysis.

Theorem 24 *Let $f\colon M \to N$ be a map between metric spaces. Then f is continuous at $x \in M$ if and only if whenever $x_k \to x$ in M then $f(x_k) \to f(x)$ in N.*

Proof Suppose firstly that f is continuous. Let $\varepsilon > 0$. By continuity there exists $\delta > 0$ such that

$$d_M(y, x) < \delta \Longrightarrow d_N(f(y), f(x)) < \varepsilon.$$

As $x_k \to x$ then there exists K such that $d_M(x_k, x) < \delta$ when $k \geqslant K$. So $d_N((f(x_k), f(x)) < \varepsilon$ when $k \geqslant K$ and hence $f(x_k) \to f(x)$ in N.

Conversely, suppose that f is not continuous at $x \in M$. Then there exists $\varepsilon > 0$ such that for every k there is $x_k \in M$ with

$$d_M(x_k, x) < 1/k \quad \text{and yet} \quad d_N(f(x_k), f(x)) \geqslant \varepsilon.$$

In this case $x_k \to x$ in M and yet $f(x_k) \nrightarrow f(x)$ in N. ■

2. Open and Closed Sets

Throughout the following (M, d) will be taken to be a metric space. We introduce here the notion of open balls – and, in due course, open sets – which will allow us to rephrase our definition of continuity.

Notation 25 *Let $x \in M$ and $\varepsilon > 0$. Then the* ***open ball*** *$B(x, \varepsilon)$ and* ***closed ball*** *$\bar{B}(x, \varepsilon)$ are defined as*

$$B(x, \varepsilon) = \{y \in M : d(x, y) < \varepsilon\}, \qquad \bar{B}(x, \varepsilon) = \{y \in M : d(x, y) \leqslant \varepsilon\}.$$

If we wish to stress that we are working in M we will write $B_M(x, \varepsilon)$ and $\bar{B}_M(x, \varepsilon)$.

We can see that Definition 19, regarding the continuity of a function $f: M \to N$ between metric spaces can be rephrased as: f is continuous at x if

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad f(B_M(x, \delta)) \subseteq B_N(f(x), \varepsilon). \tag{2.1}$$

This rephrasing will become still more streamlined and elegant once we have introduced the more general idea of open sets.

Definition 26 *$U \subseteq M$ is said to be* ***open in*** *M (or simply* ***open****) if*

$$\forall x \in U \quad \exists \varepsilon > 0 \quad B(x, \varepsilon) \subseteq U.$$

We say that x is an ***interior point*** *of U if there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$, so saying U is open is equivalent to saying all its points are interior points.*

Remark 27 *We will refer to such U as open sets but really it would be better to think of U as an* ***open subset of*** *M. Because the ambient space M will usually be obvious it's usually safe to suppress the "in M" bit, but we will see instances where the same set is open as a subset of one ambient set, but not open as a subset of another. So take some care.*

So (2.1) can be rephrased as: $f: M \to N$ is continuous at x if

whenever $f(x)$ is an interior point of $U \subseteq N$ then x is an interior point of $f^{-1}(U) \subseteq M$

or can be rephrased still more cleanly when we are talking about a function which is continuous everywhere on M as:

Theorem 28 *Let $f: M \to N$ be a map between metric spaces (M, d_M) and (N, d_N). Then f is continuous if and only if whenever $U \subseteq N$ is open then $f^{-1}(U) \subseteq M$ is open.*

Remark 29 *Hopefully this helps motivate the importance of open sets as, with this result, we see they are sufficient to determine which functions are continuous.*

Proof Suppose that f is continuous and that U is an open subset of N. Let $x \in f^{-1}(U)$ so that $f(x)$ is an interior point of U (as U is open). So there exists $\varepsilon > 0$ such that

$$B_N(f(x), \varepsilon) \subseteq U.$$

As f is continuous at x there exists $\delta > 0$ such that

$$f(B_M(x, \delta)) \subseteq B_N(f(x), \varepsilon) \subseteq U$$

and hence $B_M(x, \delta) \subseteq f^{-1}(U)$, showing that x is an interior point of $f^{-1}(U)$. As x was an arbitrary point, then $f^{-1}(U)$ is open.

Conversely suppose that the pre-image of every open set in N is open in M. Let $\varepsilon > 0$ and $x \in M$. As $B_N(f(x), \varepsilon)$ is open in N, by assumption x is an interior point of $f^{-1}(B_N(f(x), \varepsilon))$ and so there exists $\delta > 0$ such that

$$B_M(x,\delta) \subseteq f^{-1}(B_N(f(x), \varepsilon)) \quad \text{and hence} \quad f(B_M(x,\delta)) \subseteq B_N(f(x), \varepsilon),$$

which is just another way of saying the f is continuous at x. ■

Definition 30 *We refer to the open sets of M as the **topology** of M, or the topology induced by the metric d.*

Proposition 31 *(a) The intersection of finitely many open sets is open.*

(b) An arbitrary union of open sets is open.

Proof (a) Let $U_1, \ldots, U_m$ be (finitely many) open subsets of M. Let $x \in \bigcap_{i=1}^{m} U_i$. Then, for each i, there exists $\varepsilon_i > 0$ such that $B(x, \varepsilon_i) \subseteq U_i$. Let $\varepsilon = \min\{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_m\} > 0$ and then $B(x, \varepsilon) \subseteq \bigcap_{i=1}^{m} U_i$.

(b) For $i \in I$ (a not-necessarily finite indexing set), let U_i be an open subset. Let $x \in \bigcup_{i \in I} U_i$. Then there exists $j \in I$ such that $x \in U_j$ and $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U_j$. Then

$$B(x, \varepsilon) \subseteq U_j \subseteq \bigcup_{i \in I} U_i.$$

■

Remark 32 *(Off-syllabus) We saw in Theorem 28 that knowledge of the open sets is sufficient to determine which functions are continuous. So the open sets are more fundamental to continuity than the actual metric (many different metrics will lead to the same family of open sets). In line with this thinking a **topological space** is defined as follows.*

A topological space $(X, \mathcal{T})$ is a set X together with a family $\mathcal{T}$ of subsets of X with the following properties:

(i) $\varnothing \in \mathcal{T}$ and $X \in \mathcal{T}$;

(ii) $\mathcal{T}$ is closed under finite intersections;

(iii) $\mathcal{T}$ is closed under arbitrary unions.

$\mathcal{T}$ is called the ***topology*** *and U is said to be* ***open*** *when $U \in \mathcal{T}$.*

So all metric spaces are topological spaces though not all topologies are induced by metrics. However a topological space (which is a much more general notion) is sufficient to describe many aspects of continuity and indeed the proofs are often neater and more natural in that context without the unnecessary and messy inequalities associated with a metric. Those interested in this should continue with the Hilary Term long option in Topology.

Example 33 *An infinite collection of open sets need not have an open intersection. For example, consider $U_i = (-i, i)$ for $i > 0$ which have intersection $\{0\}$ which is not open.*

Definition 34 *$C \subseteq M$ is said to be* ***closed in*** *M (or simply* ***closed****) if $M \backslash C$ is open.*

Remark 35 *Note that open and closed are not opposites of one another! Sets may be open, closed, neither or both, as we shall see.*

Example 36 *(a) $\varnothing$ and M are both open and closed in M. That $\varnothing$ is open is "vacuously true" and M is clearly open by definition.*

(b) The interval $[0, 1)$ is neither open nor closed in $\mathbb{R}$. To see this, note it is not open as 0 is not an interior point of $[0, 1)$ and that the set is not closed as 1 is not an interior point of the complement.

(c) The subset $[0, \infty) \times (0, \infty) \subseteq \mathbb{R}^2$ is neither open nor closed. The point $(0, 1)$ is not interior to the subset and $(1, 0)$ is not interior to its complement.

(d) The subset $\mathbb{Z}^2$ is closed but not open in $\mathbb{R}^2$.

(e) The set $\{(x, y) : x^2 + xy + y^2 < 1\}$ is open. To see this we can note that $f : \mathbb{R}^2 \to \mathbb{R}$ given by $f(x, y) = x^2 + xy + y^2$ is continuous and then the set in question is $f^{-1}((-\infty, 1))$ which is the preimage by a continuous function of an open set (Theorem 28).

(f) Let $M = [0, 1] \cup [2, 3]$. Then $[0, 1]$ is both open and closed in M. Likewise $[2, 3]$. (We will comment more on such examples in the chapter Subspaces.)

Proposition 37 *Let $x \in M$ and $\varepsilon > 0$. Then $B(x, \varepsilon)$ is open and $\bar{B}(x, \varepsilon)$ is closed.*

Proof Firstly if $y \in B(x, \varepsilon)$ then let $\delta = \frac{1}{2}(\varepsilon - d(y, x)) > 0$. By the triangle inequality, if $z \in B(y, \delta)$ then

$$d(z, x) \leqslant d(z, y) + d(y, x) < \delta + d(y, x) = \frac{\varepsilon + d(y, x)}{2} < \frac{\varepsilon + \varepsilon}{2} = \varepsilon.$$

Hence $B(y, \delta) \subseteq B(x, \varepsilon)$ and $B(x, \varepsilon)$ is open.

To prove the second result, take $y \in M \backslash \bar{B}(x, \varepsilon)$, Then $d(x, y) > \varepsilon$ and set $\delta = \frac{1}{2}(d(x, y) - \varepsilon) > 0$. If $d(z, y) < \delta$ then by the triangle inequality

$$d(z, x) \geqslant d(x, y) - d(z, y) > d(x, y) - \delta = \varepsilon + \delta > \varepsilon$$

and hence $z \in M \backslash \bar{B}(x, \varepsilon)$. That is $M \backslash \bar{B}(x, \varepsilon)$ is open and $\bar{B}(x, \varepsilon)$ is closed. ■

Proposition 38 *Singleton sets are closed.*

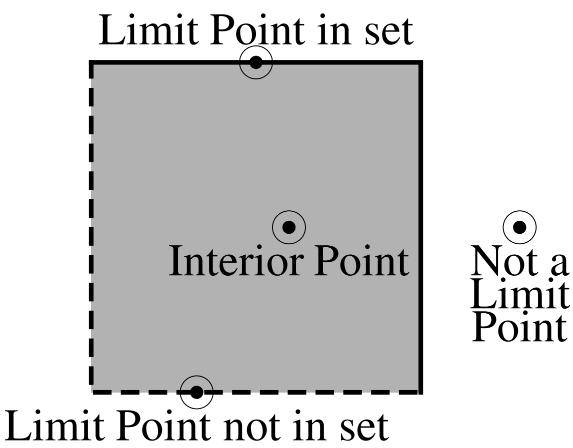
Proof Let $x \in M$. If $y \neq x$ we may set $\varepsilon = d(x, y) > 0$. Then $x \notin B(y, \varepsilon)$ and hence y is interior to $M \backslash \{x\}$. ∎

Example 39 *Note, when $M = (0, 1) \cup \{2\}$, that $B_M(2, 1) = \bar{B}_M(2, 1) = \{2\}$. In this case this open ball is also closed.*

Definition 40 *Let $S \subseteq M$ and $x \in M$. We say that x is a **limit point** of S or an **accumulation point** of S if for any $\varepsilon > 0$*

$$(B(x, \varepsilon) \cap S) \backslash \{x\} \neq \varnothing.$$

i.e. there are points of S arbitrarily close to x other than x, which may or may not be in S. We will denote the set of limit points of S as S'.



Interior and limit points of a set

Example 41 *(a) Let $M = \mathbb{R}$ and $S = [0, 1)$. Then $S' = [0, 1]$. But if $M = S = [0, 1)$, then $S' = [0, 1)$.*

(b) Let $M = \mathbb{R}$ and $S = (-\infty, 1) \cup \{2\}$. Then $S' = (-\infty, 1]$. So a point of S need not be a point of S'.

Proposition 42 *$C \subseteq M$ is closed if and only if it contains all its limit points.*

Proof Note that x is an interior point of $M \backslash C$ if and only if x is *not* a limit point of C. So

$$\begin{aligned} C \text{ is closed} &\iff M \backslash C \text{ is open} \\ &\iff \text{every point of } M \backslash C \text{ is an interior point} \\ &\iff \text{no point of } M \backslash C \text{ is a limit point of } C \\ &\iff C \text{ contains all its limit points.} \end{aligned}$$

∎

The results corresponding to Proposition 31 for closed sets are:

Proposition 43 *(a) The union of finitely many closed sets is closed.*
(b) An arbitrary intersection of closed sets is closed.

Proof These simply follow from applying De Morgan's laws to the equivalent properties of open subsets shown in Proposition 31. ■

And we also have:

Corollary 44 *(To Theorem 28) Let $f\colon M \to N$ be a map between metric spaces M and N. Then f is continuous if and only if whenever C is closed in N then $f^{-1}(C)$ is closed in M.*

Proof Noting that for any $S \subseteq N$ we have $f^{-1}(N\backslash S) = M\backslash f^{-1}(S)$, we have

$$\begin{aligned} f \text{ is continuous} &\iff f^{-1}(U) \text{ is open in } M \text{ whenever } U \text{ is open in } N \\ &\iff f^{-1}(N\backslash C) \text{ is open in } M \text{ whenever } C \text{ is closed in } N \\ &\iff M\backslash f^{-1}(C) \text{ is open in } M \text{ whenever } C \text{ is closed in } N \\ &\iff f^{-1}(C) \text{ is closed in } M \text{ whenever } C \text{ is closed in } N. \end{aligned}$$

■

Corollary 45 *Open balls are open and closed balls are closed.*

Proof We showed this in Proposition 37. However this is now an easy application of Theorem 28 and Corollary 44 to the continuous function $f(x) = d(x, a)$ from Proposition 23 as

$$B(a,\varepsilon) = f^{-1}((-\infty,\varepsilon)) \qquad \text{and} \qquad \bar{B}(a,\varepsilon) = f^{-1}((-\infty,\varepsilon]).$$

■

Example 46 *The sphere $x^2 + y^2 + z^2 = a^2$ is closed in $\mathbb{R}^3$. This is because $\{a^2\}$ is closed in $\mathbb{R}$ and $f\colon \mathbb{R}^3 \to \mathbb{R}$ is continuous where*

$$f(x, y, z) = x^2 + y^2 + z^2.$$

The sphere is then $f^{-1}(\{a^2\})$.

Example 47 *Let c denote the subset of l^∞ consisting of all convergent sequences. Then c is a closed subset of l^∞.*

Solution Say that $\left(x_n^{(k)}\right)$ is a sequence in c which converges to (X_n) in l^∞ as $k \to \infty$. We aim to show that (X_n) is in c; we will do this by showing (X_n) is Cauchy and so convergent.

Let $\varepsilon > 0$. As $\left(x_n^{(k)}\right) \to (X_n)$ in l^∞ then

$$\left\| \left(x_n^{(k)}\right) - (X_n) \right\|_\infty = \sup_n \left| x_n^{(k)} - X_n \right| \to 0 \quad \text{as} \quad k \to \infty.$$

So there exists K such that

$$\left|x_n^{(k)} - X_n\right| < \varepsilon/3 \quad \text{for } k \geqslant K \text{ and all } n.$$

As $\left(x_n^{(K)}\right)$ is convergent then it is Cauchy. So there exists N such that

$$\left|x_n^{(K)} - x_m^{(K)}\right| < \varepsilon/3 \qquad \text{for } n, m \geqslant N.$$

Thus for $m, n \geqslant N$ we have

$$|X_m - X_n| \leqslant \left|X_m - x_m^{(K)}\right| + \left|x_m^{(K)} - x_n^{(K)}\right| + \left|x_n^{(K)} - X_n\right| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon.$$

■

Example 48 *Given a metric space M, $x \in M$ and $\varepsilon > 0$, the limit points of $B(x, \varepsilon)$ need not include $\bar{B}(x, \varepsilon)$.*

Solution Take the discrete metric on a set M and $x \in M$. Then $B(x, 1) = \{x\}$ whilst $\bar{B}(x, 1) = M$. However given any $y \neq x$ we see that $x \notin B(y, 1)$ and so y is not a limit point of $B(x, 1)$. In fact there are no limit points of $B(x, 1)$. ■

3. Subspaces. Homeomorphisms. Isometries.

When describing any type of structure in mathematics, it is natural to discuss the possible substructures and isomorphisms.

Again throughout this (M, d) will be a metric space. As we have already noted (Example 17) the metric d on M induces a metric d_A on any subset A of M so that (A, d_A) is a metric space in its own right.

Recall the following Definitions 26 and 34.

- Say $S \subseteq T \subseteq M$. We say S is **open in** T if for every s in S there exists $\varepsilon > 0$ such that

$$B_T(s, \varepsilon) = \{x \in T : d_T(x, s) < \varepsilon\} \subseteq S.$$

  But as $d_T(x, s) = d(x, s)$ for $x, s \in T$ then

$$B_T(s, \varepsilon) = \{x \in T : d(x, s) < \varepsilon\} = B_M(s, \varepsilon) \cap T.$$

- We say that S is **closed in** T if $T \backslash S$ is open in T.

Definition 49 *We refer to (T, d_T) as a **subspace** of (M, d) and refer to the open sets of T as the **subspace topology** of T.*

As (T, d_T) is a metric space in its own right, it remains the case that:

- Given an arbitrary collection of open subsets of T, their union is open in T.
- Given finitely many open subsets of T, their intersection is open in T.
- Given an arbitrary collection of closed subsets of T, their intersection is closed in T.
- Given finitely many closed subsets of T, their union is closed in T.

Example 50 *(a) $[0, 1)$ is open in $[0, 1]$. Note that, for suitably small $\varepsilon > 0$, we have*

$$B_{[0,1]}(x, \varepsilon) = (x - \varepsilon, x + \varepsilon)$$

for $0 < x < 1$ and

$$B_{[0,1]}(0, \varepsilon) = [0, \varepsilon).$$

So given any x in the range $0 \leqslant x < 1$, we can find $\varepsilon > 0$ such that $B_{[0,1]}(x, \varepsilon) \subseteq [0, 1)$

(b) Given any $X \subseteq M$ then X is open and closed in itself. So the empty set is also an open and closed subset of X.

(c) Every subset of $\mathbb{Z}$ is open and closed in $\mathbb{Z}$. Note that for any integer n we have $B_{\mathbb{Z}}(n, 1/2) = \{n\}$. As singleton points are open in $\mathbb{Z}$, and an arbitrary union of open sets is open, then any subset is open in $\mathbb{Z}$.

(d) Consequently any map $f : \mathbb{Z} \to M$, to any metric space M, is continuous.

Example 51 *Let $X = [0,1] \cup [2,3]$. The map $f\colon X \to \mathbb{Z}$ defined by*

$$f(x) = \begin{cases} 0 & 0 \leqslant x \leqslant 1 \\ 1 & 2 \leqslant x \leqslant 3 \end{cases}$$

is continuous. Given $S \subseteq \mathbb{Z}$ then $f^{-1}(S)$ equals

$$\varnothing, \qquad [0,1], \qquad [2,3], \qquad X,$$

depending on whether neither, 0, 1 or both lie in S. All these possible pre-images are open in X.

Proposition 52 *Let $S \subseteq T \subseteq M$.*
(a) S is open in T if and only if there is an open set U (in M) such that $S = T \cap U$.
(b) S is closed in T if and only if there is a closed set C (in M) such that $S = T \cap C$.

Proof (a) Suppose that S is open in T. Then for each $s \in S$ there exists $\varepsilon_s > 0$ such that $B_T(s, \varepsilon_s) \subseteq S$. So

$$S = \bigcup_s B_T(s, \varepsilon_s) = \bigcup_s (B(s, \varepsilon_s) \cap T) = \left(\bigcup_s B(s, \varepsilon_s) \right) \cap T$$

and we set $U = \bigcup\limits_s B(s, \varepsilon_s)$. Conversely if $S = T \cap U$ for some open U then for any $s \in S \subseteq U$ there exists $\varepsilon > 0$ such that $B(s, \varepsilon) \subseteq U$ and then $B_T(s, \varepsilon) \subseteq U \cap T = S$.

(b) Note that

$$\begin{aligned} S \text{ is closed in } T &\iff T \backslash S \text{ is open in } T \\ &\iff T \backslash S = T \cap U \text{ for some open } U \text{ in } M \\ &\iff S = T \cap (M \backslash U) \text{ for some open } U \text{ in } M \\ &\iff S = T \cap C \text{ for some closed } C \text{ in } M. \end{aligned}$$

■

Example 53 *(a) $(a,b) \cap \mathbb{Q}$ is open in $\mathbb{Q}$ for all choices of a, b and is closed in $\mathbb{Q}$ if and only a and b are irrational.*

By the above $(a,b) \cap \mathbb{Q}$ is always open in $\mathbb{Q}$. Further if a and b are irrational then

$$(a,b) \cap \mathbb{Q} = [a,b] \cap \mathbb{Q}$$

is closed in $\mathbb{Q}$. If, say, a is rational then a is not an interior point of the complement and the set is not closed in $\mathbb{Q}$.

(b) Let $S = (0,1] \cup [2,3]$. Then $(0,1]$ and $[2,3]$ are open and closed subsets of S.

$$(0,1] = (0,2) \cap S = [0,1] \cap S; \qquad [2,3] = (1,4) \cap S = [2,3] \cap S.$$

Example 54 *Let $A \subseteq B \subseteq C \subseteq M$.*

(a) Show that if A is open in B and B is open in C, then A is open in C.

(b) Show that if A is closed in B and B is closed in C, then A is closed in C.

Solution (a) As B is open in C then there exists open U such that $B = U \cap C$. As A is open in B then there exists open V such that $A = V \cap B$. Hence

$$A = V \cap B = (V \cap U) \cap C$$

and $V \cap U$ is open. Part (b) follows similarly. ■

Proposition 55 *Let $A \subseteq M$ and $f\colon M \to N$ be a continuous map between metric spaces M and N. Then the restriction $f|_A$ of f to A is continuous.*

Proof Let U be an open subset of N. Then

$$(f|_A)^{-1}(U) = f^{-1}(U) \cap A$$

which is open in A as $f^{-1}(U)$ is open in M. ■

Definition 56 *Given $A \subseteq M$ the* ***closure*** *of A, written $\overline{A}$, is the smallest closed subset of M which contains A. This is well-defined as it equals*

$$\overline{A} = \bigcap \{C : C \text{ closed in } M \text{ and } A \subseteq C\}.$$

We saw in Sheet 1, Exercise 5(ii), that $\overline{A} = A \cup A'$ where A' is the set of limit points of A.

Proposition 57 *Let $A \subseteq B \subseteq M$. Let $\overline{A}^M$ and $\overline{A}^B$ respectively denote the closures of A in M and of A in B. Then*

$$\overline{A}^B = \overline{A}^M \cap B.$$

Proof We have that

$$\begin{aligned} \overline{A}^B &= \bigcap \{C : C \text{ closed in } B \text{ and } A \subseteq C\} \\ &= \bigcap \{D \cap B : D \text{ closed in } M \text{ and } A \subseteq D\} \\ &= \bigcap \{D : D \text{ closed in } M \text{ and } A \subseteq D\} \cap B \\ &= \overline{A}^M \cap B. \end{aligned}$$

■

The notion of isomorphism for metric spaces is that of an *isometry.*

Definition 58 *An* ***isometry*** *between metric spaces (M, d_M) and (N, d_N) is a bijection $f\colon M \to N$ such that*

$$d_N(f(x), f(y)) = d_M(x, y) \qquad \text{for all } x, y \in M. \tag{3.1}$$

In such a case we would say that M and N are ***isometric****. Clearly f^{-1} is also an isometry.*

Remark 59 *Sometimes a map $f\colon M \to N$ will be termed an isometry if (3.1) holds. This is sufficient to guarantee that f is 1-1 but such an "isometry" need not be onto.*

Proposition 60 *(a) An isometry is continuous.*

(b) The isometries of a metric space form a group under composition.

(c) The isometry group of $\mathbb{R}^n$ is the so-called ***Euclidean group****. This consists of all maps of the form $\mathbf{x} \mapsto A\mathbf{x} + \mathbf{b}$ where $A \in O(n)$ and $\mathbf{b} \in \mathbb{R}^n$.*

(d) The isometry group of the unit sphere $S^{n-1} \subseteq \mathbb{R}^n$ is $O(n)$.

Proof (a) and (b) are trivial. (c) and (d) are results from (or using) the first year geometry course. ■

Example 61 *Let (M, d_M) be a metric space, N be a set and $f\colon M \to N$ be a bijection. We can then use f to induce a metric d_N on N in such a way that M and N are isometric. We simply define*

$$d_N(x, y) = d_M(f^{-1}(x), f^{-1}(y)).$$

For example, this immediately means that

$$d(x, y) = |\ln x - \ln y|$$

is a metric on $(0, \infty)$ as this is the metric induced from the bijection $\exp\colon \mathbb{R} \to (0, \infty)$.

From a point of view of continuity, though, isometries really ask too much. We are aware of non-isometric metrics on the same space leading to the same topology (i.e. the same family of open sets) and so to the same family of continuous functions. What we really need is the idea of a bijection between two spaces that induces a bijection between the topologies. This is the idea of a *homeomorphism*.

Definition 62 *Let M and N be metric spaces. A* ***homeomorphism*** *$f\colon M \to N$ is a bijection f such that f and f^{-1} are continuous. Note, in particular, that this induces a bijection between the two spaces' topologies. (i.e. if U is open in M then $f(U)$ is open in N and if V is open in N then $f^{-1}(V)$ is open in M).*

Example 63 *An isometry is a homeomorphism.*

Example 64 $(0, 1)$ *and* $\mathbb{R}$ *are homeomorphic (and not isometric).*

Solution The map $f\colon\ (0, 1) \to \mathbb{R}$ given by

$$f(x) = \frac{2x - 1}{x(1 - x)}$$

is a homeomorphism. (This is clear from a sketch of its graph.) The two spaces are clearly not isometric as 0 and 2 are distance 2 apart in $\mathbb{R}$ and no such points exist in $(0, 1)$. ■

Example 65 $\mathbb{N}$ *and* $\mathbb{Z}$ *are homeomorphic. Both spaces have the discrete topology (all subsets are open) and so any bijection between the two is automatically a homeomorphism.*

Example 66 *The homeomorphisms of* $\mathbb{R}$ *are the continuous, strictly monotone maps which are neither bounded above nor below.*

Example 67 $\mathbb{R}$ *and* $\mathbb{R}^2$ *are not homeomorphic.*

How would one show this last result? The simple answer is that we can't prove this just yet, but posing the question motivates some of our later concepts. If we were faced with a similar question in group theory – for example: show $\mathbb{Z}$ and $\mathbb{Q}$ are not isomorphic groups – then we would note that $\mathbb{Z}$ is cyclic and that $\mathbb{Q}$ is not. "Being cyclic" is an algebraic invariant preserved by isomorphisms. What we need are similar *topological invariants* to help us distinguish between these two spaces.

4. Completeness

Throughout the following (M, d) is a metric space.

Definition 68 *We say that a sequence (x_n) in M is **Cauchy** if $d(x_m, x_n) \to 0$ as $m, n \to \infty$ or equivalently*

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

Proposition 69 *A convergent sequence is Cauchy.*

Proof Say $x_n \to x$ in M. Let $\varepsilon > 0$. Then there exists N such that $d(x_n, x) < \varepsilon/2$ for all $n \geqslant N$. By the triangle inequality

$$d(x_n, x_m) \leqslant d(x_n, x) + d(x, x_m) < \varepsilon/2 + \varepsilon/2 = \varepsilon \quad \text{for all } m, n \geqslant N.$$

■

Example 70 *In $\mathbb{R}$ and $\mathbb{C}$ every Cauchy sequence is convergent. In $\mathbb{Q}$ this is not the case, for example*

$$3, \quad 3.1, \quad 3.14, \quad 3.141, \quad 3.1415, \ldots$$

is a Cauchy sequence in $\mathbb{Q}$ which does not converge in $\mathbb{Q}$. Note $1/n$ is a Cauchy sequence in $(0,1)$ which does not converge in $(0,1)$.

Definition 71 *A metric space M is said to be **complete** if every Cauchy sequence in M converges in M.*

Proposition 72 *Let (X, d) be a complete metric space and $Y \subseteq X$. Then Y is complete if and only if Y is closed in X.*

Proof Suppose that Y is closed in X and let (y_n) be a Cauchy sequence in Y. Then (y_n) is also Cauchy in X and so convergent to $x \in X$ as X is complete. As Y is closed then $x \in Y$ and we see that Y is complete. Conversely suppose that Y is complete and $y \in X$ is a limit point of Y. Then there is a sequence (y_n) in Y converging to y. But convergent sequences are Cauchy and so $y \in Y$ as Y is complete. Hence Y contains all its limit points and so is closed in X. ■

Example 73 *Note that $(0,1)$ and $\mathbb{R}$ (with the usual topologies) are homeomorphic and yet $\mathbb{R}$ is complete and $(0,1)$ is not. This shows that completeness is not a topological invariant (i.e. is not preserved by homeomorphisms). Completeness **is** preserved by isometries but the equivalent topological invariant is that of **complete metrizability**.*

For example, the map $f\colon (0,1) \to \mathbb{R}$ given by $f(x) = \tan(\pi(x - 1/2))$ is a homeomorphism. So if we define the metric d on $(0,1)$ by $d(x,y) = |f(x) - f(y)|$ then $((0,1), d)$ and $\mathbb{R}$ (with the usual metric) are isometric. In particular $((0,1), d)$ is complete whilst still having the usual topology.

Example 74 *The spaces l^1, l^2, l^∞ are all complete.*

Solution We shall prove this only for l^1. Let $\left(x_n^{(1)}\right), \left(x_n^{(2)}\right), \left(x_n^{(3)}\right), \ldots$ be a Cauchy sequence in l^1. This means that given any $\varepsilon > 0$ there exists K such that

$$\left\| \left(x_n^{(k)}\right) - \left(x_n^{(l)}\right) \right\|_1 = \sum_{n=1}^{\infty} \left| x_n^{(k)} - x_n^{(l)} \right| < \varepsilon \quad \text{for } k, l \geqslant K \tag{4.1}$$

In particular this means that $\left| x_n^{(k)} - x_n^{(l)} \right| < \varepsilon$ for $k, l \geqslant K$ and so $x_n^{(k)}$ is a Cauchy sequence which converges to some X_n as $k \to \infty$. Note that for any N and $k \geqslant K$ we have

$$\sum_{n=1}^{N} |X_n| \leqslant \sum_{n=1}^{N} \left| X_n - x_n^{(k)} \right| + \sum_{n=1}^{N} \left| x_n^{(k)} \right| \leqslant \varepsilon + \left\| \left(x_n^{(k)}\right) \right\|_1$$

and so (X_n) is absolutely summable. Further letting $l \to \infty$ in (4.1) we have

$$\left\| \left(x_n^{(k)}\right) - (X_n) \right\|_1 < \varepsilon \quad \text{for } k \geqslant K$$

so that $\left(x_n^{(k)}\right)$ converges to (X_n) as $k \to \infty$. ∎

Theorem 75 *Let X be a set and let $\mathcal{B}(X)$ denote the set of all bounded real-valued functions on X. Then*

$$\delta(f, g) = \sup \{ |f(x) - g(x)| : x \in X \}$$

defines a metric on $\mathcal{B}(X)$. Further $\mathcal{B}(X)$ is complete.

Proof Clearly $\delta(f, g) \geqslant 0$ and if $\delta(f, g) = 0$ then we have $|f(x) - g(x)| = 0$ for all x and hence $f = g$. The symmetry of δ is evident. Finally if $f, g, h \in \mathcal{B}(X)$ and $x \in X$ then:

$$|f(x) - h(x)| \leqslant |f(x) - g(x)| + |g(x) - h(x)| \leqslant \delta(f, g) + \delta(g, h).$$

If we take the supremum of the LHS over all $x \in X$ then we have the required triangle inequality. Finally let (f_n) be a Cauchy sequence in $\mathcal{B}(X)$. This means

$$\sup \{ |f_n(x) - f_m(x)| : x \in X \} \to 0 \quad \text{as} \quad m, n \to \infty.$$

For a particular $x_0 \in X$, it follows that $|f_n(x_0) - f_m(x_0)| \to 0$ as $m, n \to \infty$ and so $(f_n(x_0))$ is a real Cauchy sequence. By completeness, this converges which we'll denote the limit $f(x_0)$. Let $\varepsilon > 0$. As (f_n) is a Cauchy sequence, then there exists N such that

$$\sup \{ |f_n(x) - f_m(x)| : x \in X \} < \varepsilon \quad \text{when} \quad m, n \geqslant N$$

and so

$$\delta(f_n, f) = \sup \{ |f_n(x) - f(x)| : x \in X \} \leqslant \varepsilon \quad \text{when} \quad n \geqslant N$$

as the modulus is continuous and sup respects weak inequalities. As f_n is bounded then so is f and further f_n converges to f. ∎

If X now has a metric d, then we can discuss the bounded continuous real-valued functions on X and we have the following result.

Theorem 76 *Let (X,d) be a metric space and let $\mathcal{C}(X)$ denote the subset of $\mathcal{B}(X)$ consisting of all continuous real-valued functions. Then $\mathcal{C}(X)$ is complete.*

Proof Let (f_n) be a Cauchy sequence in $\mathcal{C}(X)$; this is then a Cauchy sequence in $\mathcal{B}(X)$ and so converges to some bounded function f by the previous theorem. It remains to show that f is continuous. Let $\varepsilon > 0$. As (f_n) is Cauchy there exists N such that

$$\sup\{|f_n(x) - f_m(x)| : x \in X\} < \varepsilon/3 \quad \text{when} \quad m, n \geqslant N$$

and hence $|f_N(x) - f(x)| < \varepsilon/3$ for all $x \in X$. As f_N is continuous at x then there exists $\delta > 0$ such that

$$|f_N(x) - f_N(y)| < \varepsilon/3 \quad \text{when} \quad d(x,y) < \delta.$$

Hence when $d(x,y) < \delta$ then

$$|f(x) - f(y)| \leqslant |f(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f(y)| < \varepsilon/3 + \varepsilon/3 + \varepsilon/3 = \varepsilon$$

and the result follows. ■

Remark 77 *Note that we could have equally shown that $\mathcal{C}(X)$ is a closed subset of $\mathcal{B}(X)$ and used Proposition 72. This approach is tantamount to showing that the uniform limit of continuous functions is continuous, a result from first year analysis.*

Definition 78 *A **contraction** (or a **contraction mapping**) on a metric space M is a map $f \colon M \to M$ such that there is some $K < 1$ with*

$$d(f(x), f(y)) \leqslant K d(x,y) \quad \textit{for all } x, y \in M.$$

*Contractions are special cases of Lipschitz maps. A map $f \colon (M, d_M) \to (N, d_N)$ is said to be **Lipschitz** if there exist $K > 0$ such that*

$$d_N(f(x), f(y)) \leqslant K d_M(x,y) \quad \textit{for all } x, y \in M.$$

Proposition 79 *Lipschitz maps (and hence contractions) are uniformly continuous.*

Proof Let $\varepsilon > 0$ and set $\delta = \varepsilon/K$. Then for any x, y and with $d_M(y,x) < \delta$ we have

$$d_N(f(x), f(y)) \leqslant K d_M(x,y) < K\delta = \varepsilon.$$

■

Theorem 80 ***Contraction Mapping Theorem*** *(Banach 1922) Let $f \colon X \to X$ be a contraction on a complete non-empty metric space (X,d). Then there is a unique fixed point $x \in X$ such that $f(x) = x$.*

Proof Let $0 \leqslant K < 1$ be such that $d(f(x), f(x')) \leqslant Kd(x, x')$ for $x, x' \in X$..Let $x_0 \in X$ and define the sequence (x_n) by $x_{n+1} = f(x_n)$ for $n \geqslant 0$. We will first show that (x_n) is Cauchy. If $n > m > 0$ then

$$d(x_n, x_{n-1}) = d(f(x_{n-1}), f(x_{n-2})) \leqslant Kd(x_{n-1}, x_{n-2}) \leqslant \cdots \leqslant K^{n-1}d(x_1, x_0)$$

and

$$\begin{aligned} d(x_n, x_m) &\leqslant d\left(x_n, x_{n-1}\right) + d(x_{n-1}, x_{n-2}) + \cdots + d(x_{m+1}, x_m) \qquad \text{[triangle inequality]} \\ &\leqslant (K^{n-1} + K^{n-2} + \cdots + K^{m-1})d(x_1, x_0) \\ &= \frac{K^m - K^n}{1-K} d(x_1, x_0) \to 0 \quad \text{as } m, n \to \infty. \end{aligned}$$

Hence, as X is complete, the sequence (x_n) is convergent, and set $x = \lim x_n$. For $n \geqslant 0$ we have $f(x_n) = x_{n+1}$ and if we let $n \to \infty$ then

$$x = \lim x_{n+1} = \lim f(x_n) = f(\lim x_n) = f(x)$$

as f is a contraction and so continuous. Finally if x and x' are two fixed points of f we have

$$d(x, x') = d(f(x), f(x')) \leqslant Kd(x, x')$$

which is a contradiction unless $d(x, x') = 0$ and $x = x'$. ■

Example 81 *Prove that there exists a unique continuous function $f\colon [0,1] \to \mathbb{R}$ such that*

$$f(x) = x + \frac{1}{2}\int_0^1 f(y)\sin(xy)\,\mathrm{d}y \quad \textit{for all } x \in [0,1].$$

Proof We know from Theorem 76 that $C[0,1]$, the space of all real-valued continuous functions on $[0,1]$ with the supremum metric

$$d(f, g) = \sup\{|f(x) - g(x)| : x \in [0,1]\}$$

is a complete metric space. Consider $F\colon C[0,1] \to C[0,1]$ defined by

$$(Ff)(x) = x + \frac{1}{2}\int_0^1 f(y)\sin(xy)\,\mathrm{d}y.$$

For $f, g \in C[0,1]$ we have

$$\begin{aligned} d(Ff, Fg) &= \sup\{|Ff(x) - Fg(x)| : x \in [0,1]\} \\ &= \sup\left\{\left|\frac{1}{2}\int_0^1 (f(y) - g(y))\sin(xy)\,\mathrm{d}y\right| : x \in [0,1]\right\} \\ &\leqslant \sup\{\frac{1}{2}\int_0^1 |(f(y) - g(y))\sin(xy)|\,\mathrm{d}y : x \in [0,1]\} \\ &\leqslant \sup\{\frac{1}{2}\int_0^1 |(f(y) - g(y))|\,\mathrm{d}y : x \in [0,1]\} \\ &= \frac{1}{2}\int_0^1 |(f(y) - g(y))|\,\mathrm{d}y \\ &\leqslant \frac{1}{2}\sup\{|f(y) - g(y)| : y \in [0,1]\} = \frac{1}{2}d(f, g). \end{aligned}$$

So F is a contraction on the complete space $C[0,1]$ and hence has a unique fixed point. ■

5. Connectedness

We meet now our first topological invariant, *connectedness*, which rigorously gives definition to the idea of a space being "in one piece".

Definition 82 *A metric space M is* ***disconnected*** *if there exist A, B, disjoint, nonempty, open subsets of M such that $M = A \cup B$. We say that S is* ***connected*** *if it is not disconnected.*

Remark 83 *Note connectedness is a property that a space itself has, or hasn't. "Being connected" has nothing to do with being connected "in" a larger space.*

Proposition 84 *The following three statements are equivalent definitions for a space M being connected.*

(a) There is no partition of M into nonempty, disjoint open subsets of M.

(b) The only open and closed subsets of M are $\varnothing$ and M.

(c) Any continuous function $f\colon M \to \mathbb{Z}$ is constant.

Proof *(a) $\iff$ (c) This is left to Exercise 6(i) on Sheet 2.*

$\neg(a) \implies \neg(b)$ If $M = A \cup B$ is a partition of M into non-empty open subsets, then A is open but also closed as $A = M \backslash B$.

$\neg(b) \implies \neg(a)$ If A is a non-empty proper open and closed subset of M then $M = A \cup (M \backslash A)$ is a partition of M into non-empty open subsets. ■

Remark 85 *Many authors use the word* ***clopen*** *for "open and closed", but the term will not be used in these notes.*

Example 86 *(a) $\mathbb{Q}$ is disconnected as*

$$\mathbb{Q} = ((-\infty, \pi) \cap \mathbb{Q}) \cup ((\pi, \infty) \cap \mathbb{Q}).$$

The two subsets on the right hand side are open and closed in $\mathbb{Q}$ as

$$(-\infty, \pi) \cap \mathbb{Q} = (-\infty, \pi] \cap \mathbb{Q} \qquad \text{and} \qquad (\pi, \infty) \cap \mathbb{Q} = [\pi, \infty) \cap \mathbb{Q}.$$

(b) $X = (0, 1] \cup (2, 3)$ is disconnected as $(0, 1]$ and $(2, 3)$ are both open and closed in X. To see this note

$$(0, 1] = [0, 1] \cap X = (0, 2) \cap X; \qquad (2, 3) = (2, 3) \cap X = [2, 3] \cap X.$$

Proposition 87 *Let $a, b \in \mathbb{R}$ with $a \leqslant b$. Then $[a, b]$ is connected.*

Proof Let C be an open and closed subset of $[a, b]$. Without any loss of generality we may assume $a \in C$; if not we could work with $[a, b] \backslash C$. Set

$$W = \{x \in [a, b] : [a, x] \subseteq C\} \text{ and } c = \sup W$$

which is well-defined as $a \in W \neq \varnothing$.

(a) Let $\varepsilon > 0$. By the Approximation Property there exists $x \in W$ such that $c - \varepsilon < x \leqslant c$ and in particular $[a, c - \varepsilon] \subseteq C$. Hence

$$\bigcup_{\varepsilon > 0} [a, c - \varepsilon] = [a, c) \subseteq C.$$

As C is closed then $[a, c] \subseteq C$ and hence that $c \in W$.

(b) Now suppose that $x \in W$ and that $x < b$. Then $[a, x] \subseteq C$, which is open, and so there exists $\varepsilon > 0$ such that $(x - 2\varepsilon, x + 2\varepsilon) \subseteq C$. Thus

$$[a, x] \cup (x - 2\varepsilon, x + 2\varepsilon) = [a, x + 2\varepsilon) \subseteq C$$

and hence $x + \varepsilon \in W$.

(c) Combining (a) and (b), if $c < b$ we would have $c + \delta \in W$ for some $\delta > 0$ contradicting the fact that $c = \sup W$. Hence $b = c \in W$ and $[a, b] = C$. ∎

Corollary 88 ***(Intermediate Value Theorem)*** *Let $f\colon\ [a, b] \to \mathbb{R}$ be continuous with*

$$f(a) < 0 < f(b).$$

Then there exists $c \in (a, b)$ such that $f(c) = 0$.

Proof Suppose for a contradiction that $f(x) \neq 0$ for all $x \in (a, b)$. Then

$$A = \{x \in [a, b] : f(x) > 0\} = f^{-1}(0, \infty) \qquad \text{and} \qquad B = \{x \in [a, b] : f(x) < 0\} = f^{-1}(-\infty, 0)$$

partition $[a, b]$ into disjoint, non-empty open subsets, contradicting its connectedness. ∎

Remark 89 *So we see that it is the connectedness of $[a, b]$ that is the driving force behind the Intermediate Value Theorem. Indeed we see more generally from this that a continuous real-valued function on any connected space which takes positive and negative values must have a root. It is just as easy to see that this isn't generally the case for continuous functions on disconnected subsets.*

Proposition 90 *The connected subsets of $\mathbb{R}$ are the intervals.*

Proof The proof of the previous proposition can be adapted to show that any interval, be it open, closed, half-open/half-closed, bounded or unbounded, is connected. Conversely, suppose for a contradiction that C is a connected subset of $\mathbb{R}$, with $a < c < b$, $a, b \in C$, $c \notin C$. Then

$$C = [(-\infty, c) \cap C] \cup [(c, \infty) \cap C]$$

is a partition of C into disjoint, non-empty open subsets of C. ∎

Example 91 *Let $X = S \cup T$ be connected where S and T are closed subsets of X. Show that if $S \cap T$ is connected then S and T are also connected.*

Proof Note that S and T are not disjoint or otherwise they would form a partition for the connected $S \cup T$ into disjoint closed subsets. Take $x \in S \cap T$ and let C be a non-empty open and closed subset of S containing x; we will aim to show $C = S$. Now $C \cap T$ is nonempty (it contains x) and is an open and closed subset of $S \cap T$; as $S \cap T$ is connected then $C \cap T = S \cap T$ or equivalently $S \cap T \subseteq C$. Note then that $S \backslash C$ and $C \cup T$ partition $S \cup T$ into disjoint closed subsets. As $S \cup T$ is connected, and as $C \cup T \neq \varnothing$ then it follows that $S \backslash C = \varnothing$ and so $C = S$. ■

Proposition 92 *If $f\colon M \to N$ is continuous and C is a connected subset of M, then $f(C)$ is connected.*

Proof Suppose that A and B provide a partition of $f(C)$ into non-empty, disjoint sets which are open in $f(C)$. Then, as f is continuous, the preimages $f^{-1}(A)$ and $f^{-1}(B)$ provide a partition of C into non-empty, disjoint sets which are open in C. As C is connected then one of these preimages is empty, say $f^{-1}(A) = \varnothing$. As f maps onto $f(C)$ then $A = ff^{-1}(A) = \varnothing$ showing $f(C)$ is connected. ■

Corollary 93 *Connectedness is a topological invariant. i.e. it is preserved under homeomorphisms.*

Proposition 94 *Let M and N be metric spaces. Let $M \times N$ be their product space with metric*

$$d((m_1, n_1), (m_2, n_2)) = d_M(m_1, m_2) + d_N(n_1, n_2).$$

Then $M \times N$ is connected if and only if M and N are connected.

Proof The projection maps

$$\pi_1 : M \times N \to M \qquad \text{and} \qquad \pi_2 : M \times N \to N$$

are continuous. (Check!) If $M \times N$ is connected then $M = \pi_1(M \times N)$ is connected and likewise N. Conversely say that M and N are connected and $f\colon M \times N \to \mathbb{Z}$ is continuous.

$$\begin{aligned} \text{For any } m &\in M, \quad \{m\} \times N \text{ is connected and so } f \text{ is constant on } \{m\} \times N. \\ \text{For any } n &\in N, \quad M \times \{n\} \text{ is connected and so } f \text{ is constant on } M \times \{n\}. \end{aligned}$$

So for any (m_1, n_1) and (m_2, n_2) in $M \times N$ we have by constancy on $\{m_1\} \times N$ and on $M \times \{n_2\}$

$$f(m_1, n_1) = f(m_1, n_2) = f(m_2, n_2).$$

Hence f is constant on $M \times N$ which we see to be connected. ■

Definition 95 *A set $S \subseteq \mathbb{R}^n$ is said to be* ***path-connected*** *if given any $a, b \in S$ there exists a continuous map $\gamma\colon [0, 1] \to S$ such that $\gamma(0) = a$ and $\gamma(1) = b$.*

Example 96 *$\mathbb{R}^n$, l^1, l^2, l^∞ are all path-connected. In fact they are all* ***convex****. That is given points v, w in any one of these spaces, the line segment connecting v and w is also in that space. The straight line path connecting v and w is given by*

$$\gamma(t) = tw + (1-t)v \qquad 0 \leqslant t \leqslant 1.$$

Proposition 97 *A path-connected set is connected.*

Proof Let U be a path-connected set and $f\colon U \to \mathbb{Z}$ be continuous. If $a, b \in U$ then there exists a continuous map $\gamma\colon\ [0,1] \to U$ connecting a to b. Then $f \circ \gamma\colon\ [0,1] \to \mathbb{Z}$ is a continuous, integer-valued map on the connected set $[0,1]$ and so constant. In particular,

$$f(a) = f(\gamma(0)) = f(\gamma(1)) = f(b).$$

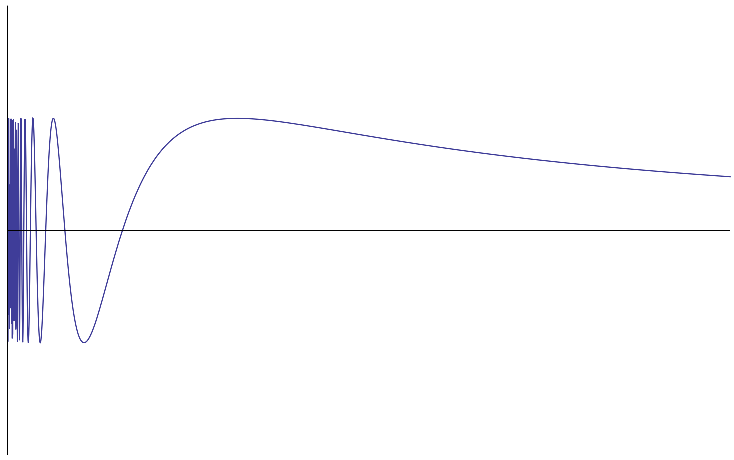
As a and b were arbitrary then f is constant on U and hence (by Proposition 84) U is connected. ∎

Proposition 98 *An open connected subset of* $\mathbb{R}^n$ *is path-connected.*

Proof Let U be an open connected subset of $\mathbb{R}^n$ and let $\mathbf{x} \in U$. Let X denote the path component of $\mathbf{x}$, that is all those points of U that can be connected to $\mathbf{x}$ by a continuous path. If $\mathbf{u} \in X$ then there is a continuous path γ connecting $\mathbf{x}$ to $\mathbf{u}$. Further, as U is open, there is $\varepsilon > 0$ such that $B(\mathbf{u},\varepsilon) \subseteq U$. Clearly $B(\mathbf{u},\varepsilon) \subseteq X$ as the path γ can be extended along a radius of the ball to any point of $B(\mathbf{u},\varepsilon)$. In particular, X is open.

We have shown that the path component X is open. Similarly any other path components are open and hence so is their union. As X is the complement of this open union then X is also closed. As X is open and closed and nonempty, and as U is connected, then $X = U$ and so U is path-connected. ∎

Example 99 ***(Topologist's Sine Curve)***



Consider the set

$$X = A \cup B \qquad \textit{where} \qquad A = \{(x, \sin(1/x)) : x > 0\} \qquad \textit{and} \qquad B = \{(0,y) : y \in \mathbb{R}\} \subseteq \mathbb{R}^2.$$

*We shall see that X is connected but **not** path-connected. Firstly connectedness: consider an open and closed subset C of X which contains the origin. Then $C \cap B$ is a non-empty open and closed subset of B. As B is connected then $C \cap B = B$ and so $B \subseteq C$. But as C is open then the origin is an interior point of C and C also includes some points from A. Again $C \cap A$ is a non-empty open and closed subset of A, and as A is connected (being a continuous image of $(0,\infty)$) then $C \cap A = A$ and $A \subseteq C$. Hence $C = A \cup B = X$ and X is connected.*

However X is not path-connected. Demonstrating this is a little fiddly, but a proof appears in Sutherland pp.101–102. Unsurprisingly the path components of X are A and B.

6. Compactness and Sequential Compactness

Definition 100 *An* ***open cover*** $\mathcal{U} = \{U_i : i \in I\}$ *for a space* M *is a collection of sets* U_i*, which are open in* M*, and such that*

$$M = \bigcup_{i \in I} U_i.$$

A ***subcover*** *of* $\mathcal{U}$ *is a collection* $\{U_i : i \in J\}$ *where* $J \subseteq I$ *such that*

$$M = \bigcup_{i \in J} U_i$$

and we say this subcover is ***finite*** *if* J *is finite.*

Definition 101 *A space* M *is said to be* ***compact*** *if every open cover of* M *has a finite subcover.*

Remark 102 *Let* $A \subseteq M$ *and let* $\mathcal{U} = \{U_i : i \in I\}$ *be an open cover of* A*. When determining the compactness or not of* A*, we might question whether it matters whether the* U_i *are open in* A *or open in* M*. In fact, it does not matter whether the* U_i *are open in* A *and*

$$A = \bigcup_{i \in I} U_i$$

or whether the U_i *are open in* M *and*

$$A \subseteq \bigcup_{i \in I} U_i.$$

Remark 103 *At this point this may seem a very odd definition and will perhaps seem moreso when we see the compact subsets of* $\mathbb{R}^n$ *are simply those that are closed and bounded. Compactness is certainly important and we shall see that:*

- *On a compact space a continuous real-valued function is bounded and attains its bounds.*
- *On a compact space a continuous function is uniformly continuous.*
- *A space is compact if and only if the Bolzano-Weierstrass Theorem holds – i.e. a sequence has a convergent subsequence.*

So why not simply focus on the closed and bounded subsets; why this extra definition? More generally, for metric spaces other than $\mathbb{R}^n$*, compact subsets and closed-and-bounded subsets differ and the three bullet points above all remain true of compact subsets of metric spaces, and *aren't* generally true of closed-and-bounded subsets.*

Further when we generalize to topological spaces – so that we have only a notion of open sets, no notion of a metric and so no notion of bounded – then the above definition of compactness, being solely in terms of open subsets, generalizes to topological spaces with most of the theory of compact sets remaining true.

Example 104 $\mathbb{R}$ *and* $(0,1]$ *are not compact.*

Solution $\{(-r,r) : r>0\}$ is an open cover of $\mathbb{R}$ which has no finite subcover. Any finite collection would be of the form $\{(-r_i,r_i) : i=1,\ldots,n\}$ whose union would be $(-R,R)$ where $R=\max r_i$.

$\{(1/k,1) : k=1,2,3,\ldots\}$ is an open cover of $(0,1]$ with no finite subcover. Any finite collection would be of the form $\{(1/k_i,1) : k=1,\ldots,n\}$ which would only cover $(1/\max k_i,1]$. ■

Example 105 *Let M be a set and d the discrete metric on M. Then $A\subseteq M$ is compact if and only if A is finite. Note that all subsets, though, are closed and bounded.*

Proposition 106 *The closed interval* $[a,b]$ *is compact.*

Proof Let $\mathcal{U}$ be an open cover of $[a,b]$. Define W to be the set

$$W=\{x\in[a,b]: \text{ a finite subcover from } \mathcal{U} \text{ for } [a,x] \text{ exists}\} \text{ and let } c=\sup(W).$$

Note that c is well-defined as $a\in W\neq\varnothing$.

(a) $c\in W$ as follows: As a is in some open subset in $\mathcal{U}$ then $c>a$. If $0<\delta<c-a$ then, by the Approximation Property, there exists $w\in W$ with $c-\delta<w$ and so $c-\delta\in W$. Say $c\in U\in\mathcal{U}$. As U is open then $(c-2\delta,c+2\delta)\subseteq U$ for some $\delta>0$ and so a finite subset of $\mathcal{U}$ covers $[a,c-\delta]\cup(c-2\delta,c+2\delta)\supseteq[a,c]$. In particular, $c\in W$.

(b) $c=b$ as follows: Say $x\in W$ and $x<b$. There exists $V\in\mathcal{U}$ such that $x\in V$ and $\delta>0$ such that $(x-2\delta,x+2\delta)\subseteq V$. So a finite subset of $\mathcal{U}$ covers

$$[a,x]\cup(x-2\delta,x+2\delta)\supseteq[a,x+\delta]$$

and hence $x+\delta\in W$. This certainly means $x\neq\sup W=c$ and so $c=b$ remains the only possibility. This shows that $[a,b]$ is compact. ■

Proposition 107 *A closed "hypercuboid"* $[a_1,b_1]\times[a_2,b_2]\times\cdots\times[a_n,b_n]$ *is compact.*

Proof This is an adaptation of the previous proof. We will prove this first for closed rectangles in $\mathbb{R}^2$. Let $\mathcal{U}=\{U_i : i\in I\}$ be an open cover of $X=[a,b]\times[c,d]$. Let

$$W=\{x\in[a,b] : \text{a finite subcover from } \mathcal{U} \text{ for } [a,x]\times[c,d] \text{ exists}\}.$$

(a) We firstly note that $a\in W$ as $\{a\}\times[c,d]$ is compact given the previous result.

(b) Now define $e=\sup W$ which is well-defined as $a\in W\neq\varnothing$. For each $y\in[c,d]$ there is an open set $U_y\in\mathcal{U}$ containing (e,y) and so some $\delta_y>0$ such that the square

$$(e-\delta_y,e+\delta_y)\times(y-\delta_y,y+\delta_y)\subseteq U_y.$$

The intervals $(y-\delta_y,y+\delta_y)$ form an open cover of $[c,d]$ which is compact; hence, there are finite $y_1,\ldots,y_n$ such that the $(y_i-\delta_{y_i},y_i+\delta_{y_i})$ cover $[c,d]$. Let $\delta=\min\delta_{y_i}>0$. Then

$$\{U_{y_1},\ldots,U_{y_n}\}$$

is an open cover for $(e-\delta, e+\delta) \times [c,d]$. Recall that $e = \sup W$ and so by the Approximation Property there exists $w \in W$ with $e - \delta < w \leqslant e$. This means there is a finite subcover $\mathcal{V}$ of $\mathcal{U}$ for $[a,w] \times [c,d]$. Hence

$$\mathcal{V} \cup \{U_{y_1}, \ldots, U_{y_n}\}$$

is a finite subcover of $\mathcal{U}$ for $[a, e+\delta) \times [c,d]$. In particular this means that $e \in W$.

(c) In fact, part (b) shows more than $e \in W$. It shows that if $x \in W$ and $x < b$, then $x + \delta/2 \in W$ for some $\delta > 0$. As $e \in W$ we would have a contradiction unless $e = b$. Hence X has a finite subcover from $\mathcal{U}$ and we see that X is compact.

(d) With an inductive proof based on the above we can see that closed bounded hypercuboids in $\mathbb{R}^n$ are compact. ■

Proposition 108 *Let M be a metric space and $A \subseteq M$. If A is compact then A is closed (in M).*

Proof Let $x \in M \backslash A$. For any $y \in A$ there exists $\varepsilon_y = \frac{1}{2} d(x,y) > 0$ such that

$$B(y, \varepsilon_y) \cap B(x, \varepsilon_y) = \varnothing.$$

Now $\mathcal{U} = \{B_A(y, \varepsilon_y) : y \in A\}$ is an open cover of A and so, by compactness, there exist finitely many $y_1, y_2, \ldots, y_n$ such that

$$A = B_A(y_1, \varepsilon_{y_1}) \cup \cdots \cup B_A(y_n, \varepsilon_{y_n}).$$

Let $\varepsilon = \min \varepsilon_{y_i} > 0$ and then we see

$$B(x, \varepsilon) = \bigcap B(x, \varepsilon_{y_i}) \subseteq \left(\bigcup B(y_i, \varepsilon_{y_i}) \right)^c \subseteq A^c$$

and hence A^c is open, i.e.A is closed. ■

Proposition 109 *Let M be a metric space and $A \subseteq M$. If A is compact then A is bounded.*

Proof Note that for any $a \in A$,

$$\mathcal{U} = \{B_A(a, n) : n \in \mathbb{N}\}$$

is an open cover of A. As A is compact then there exist $n_1 < n_2 < \cdots < n_k$ such that

$$A \subseteq \bigcup_{i=1}^{k} B_A(a, n_i) = B_A(a, n_k).$$

Hence A is bounded. ■

Proposition 110 *A closed subset of a compact space is compact.*

Proof Let M be a compact space and $A \subseteq M$ closed. We aim to show that A is compact (in its own right as a metric space). Let

$$\mathcal{U} = \{U_i : i \in I\}$$

be an open cover of A. As A is closed then $M \backslash A$ is open. So

$$\{U_i : i \in I\} \cup \{M \backslash A\}$$

is open cover of M. By the compactness of M, there is a finite subcover for M. Say

$$M = U_{i_1} \cup \cdots \cup U_{i_n} \cup \{M \backslash A\}.$$

Then

$$A = U_{i_1} \cup \cdots \cup U_{i_n}$$

and so there is a finite subcover of $\mathcal{U}$ for A. ■

Theorem 111 *(**Heine-Borel**) Let $C \subseteq \mathbb{R}^n$. Then C is compact if and only if C is closed and bounded.*

Proof We have shown generally in metric spaces that compact subsets are closed and bounded. Conversely let C be a closed and bounded subset of $\mathbb{R}^n$. As C is bounded there exist real a_i, b_i with

$$C \subseteq [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n].$$

This hypercuboid is compact and C is a closed subset of it, so by the previous proposition C is compact. ■

Proposition 112 *If $f : M \to N$ is continuous and C is a compact subset of M then $f(C)$ is compact.*

Proof Let $\{U_i : i \in I\}$ be an open cover of $f(C)$. Then, as f is continuous, $\{f^{-1}(U_i) : i \in I\}$ is an open cover of C. As C is compact then there is a finite subcover⩾

$$\left\{f^{-1}(U_{i_1}), f^{-1}(U_{i_2}), \ldots, f^{-1}(U_{i_n})\right\}$$

but, in that case, $\{U_{i_1}, U_{i_2}, \ldots, U_{i_n}\}$ is a finite subcover of $\{U_i : i \in I\}$. ■

Corollary 113 *Compactness is a topological invariant – i.e. it is preserved by homeomorphisms.*

Corollary 114 *A continuous real-valued function $f : C \to \mathbb{R}$ on a compact subset C of $\mathbb{R}^n$ is bounded and attains its bounds.*

Proof By the previous proposition $f(C)$ is a compact subset of $\mathbb{R}$. So by the Heine-Borel Theorem $f(C)$ is bounded (i.e. f is a bounded function) and $f(C)$ is closed, so f attains its bounds as the supremum and infimum of a set are either in the set or are limit points thereof. ■

Definition 115 *We say that a map $f\colon M \to N$ between metric spaces is* ***uniformly continuous*** *if*

$$\forall \varepsilon > 0 \quad \exists \delta_\varepsilon > 0 \quad \forall m \in M \quad d_M(m, m') < \delta \Longrightarrow d_N(f(m), f(m')) < \varepsilon.$$

Note how this differs from continuity

$$\forall \varepsilon > 0 \quad \forall m \in M \quad \exists \delta_{\varepsilon,m} > 0 \quad d_M(m, m') < \delta \Longrightarrow d_N(f(m), f(m')) < \varepsilon.$$

Example 116 *The function $f(x) = (1-x)^{-1}$ is not uniformly continuous on $(0,1)$. Let $\varepsilon = 1$. For $0 < a < 1$ set $b = (2-a)^{-1} > a$. Note that $f(b) - f(a) = 1 = \varepsilon$ and yet*

$$\delta = b - a = \frac{1}{2-a} - a = \frac{(1-a)^2}{2-a} \to 0 \quad \textit{as } a \to 1.$$

The function $f(x) = x^2$ is uniformly continuous on $(0,1)$; in fact, it's Lipschitz as by the Mean Value Theorem we have

$$|f(b) - f(a)| = |f'(c)|\,|b-a| = 2c\,|b-a| < 2\,|b-a|\,.$$

Theorem 117 *Let $f\colon M \to N$ be a continuous map between metric spaces. If M is compact then f is uniformly continuous.*

Proof Let $\varepsilon > 0$. As f is continuous then for every $x \in M$ there exists $\delta_x > 0$ such that $f(B(x,\delta_x)) \subseteq B(f(x), \varepsilon/2)$. The collection

$$\mathcal{U} = \left\{ B(x, \frac{\delta_x}{2}) : x \in M \right\}$$

form an open cover for M and so, by compactness, this has a finite subcover

$$\{B(x_i, \delta_{x_i}/2) : i = 1.\ldots, n\}\,.$$

We set

$$\delta = \min_i \left(\frac{\delta_{x_i}}{2} \right) > 0.$$

Take $x \in M$. Then $x \in B(x_k, \delta_{x_k}/2)$ for some k as the sets form a subcover. If $d_M(x,y) < \delta$ then we have

$$d_M(x_k, y) \leqslant d_M(x_k, x) + d_M(x, y) < \frac{\delta_{x_k}}{2} + \delta \leqslant \frac{\delta_{x_k}}{2} + \frac{\delta_{x_k}}{2} = \delta_{x_k}.$$

Hence x and y both lie in $B(x_k, \delta_{x_k})$ and

$$d_N(f(x), f(y)) \leqslant d_N(f(x), f(x_k)) + d_N(f(x_k), f(y)) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

■

Definition 118 *Let M be a metric space. We say that M is* ***sequentially compact*** *if every sequence in M has a convergent subsequence.*

Example 119 $[a, b]$ *is sequentially compact. This is just a restating of the Bolzano-Weierstrass Theorem.*

$(0, 1)$ *is not sequentially compact as the sequence* $1/n$ *has no convergent subsequence.*

Likewise $\mathbb{Q}$ *is not sequentially compact as the terminating decimal expansions of* π *do not have a convergent subsequence.*

Theorem 120 *Compact spaces are sequentially compact.*

Proof Let (x_k) be a sequence in the compact metric space M. Let

$$X_n = \overline{\{x_k : k \geqslant n\}} \qquad \text{and} \qquad U_n = M \backslash X_n.$$

As X_n is a decreasing sequence of closed sets then U_n is an increasing sequence of open sets. Suppose for a contradiction that $\bigcap X_n = \varnothing$ so that $\bigcup U_n = M$. That is, the U_n form an open cover for M. As M is compact then there is a finite subcover so that there are $i_1 < i_2 < \cdots < i_n$ such that

$$M = U_{i_1} \cup U_{i_2} \cup \cdots \cup U_{i_n} = U_{i_n}$$

which means that $X_{i_n} = \varnothing$, a contradiction. Hence there exists $x \in \bigcap X_n$. This means that

$$B(x, 1/n) \cap \{x_k : k \geqslant n\} \neq \varnothing \quad \text{for each } n$$

so that we can pick a subsequence (x_{k_n}) which converges to x. ■

Remark 121 *The converse – that sequentially compact metric spaces are compact – is also true. This fact is on the A2 syllabus but its proof is beyond this course.*

Lemma 122 *If a Cauchy sequence has a convergent subsequence then the sequence is convergent.*

Proof Let (x_n) be a Cauchy sequence in a metric space M and suppose that the subsequence (x_{n_k}) converges to x. Let $\varepsilon > 0$. As (x_n) is Cauchy there exists N such that

$$d(x_n, x_m) < \varepsilon/2 \quad \text{for } m, n \geqslant N.$$

As x_{n_k} converges to x there exists K such that

$$d\left(x_{n_k}, x\right) < \varepsilon/2 \quad \text{for } k \geqslant K.$$

So for $n \geqslant N$ and $k \geqslant K$ such that $n_k \geqslant N$,

$$d\left(x_n, x\right) \leqslant d(x_n, x_{n_k}) + d(x_{n_k}, x) < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

Hence (x_n) converges to x. ■

Corollary 123 *Compact spaces are complete.*

Proof Let (x_n) be a Cauchy sequence in a compact metric space M. By Theorem 120 (x_n) has a convergent subsequence converging to a limit x. By the previous lemma, the entire sequence converges to x. ■

Example 124 *The unit cube in* l^∞

$$C = \{(x_k) : |x_k| \leqslant 1\}$$

is not compact.

Solution The sequence $e_1 = (1, 0, \ldots), e_2 = (0, 1, 0, \ldots), e_3 = (0, 0, 1, 0, \ldots)$ has no convergent subsequence as $||e_i - e_j||_\infty = 1$ for $i \neq j$. ■

Remark 125 *Here then is an example of a closed and bounded metric space which is not compact.*

Remark 126 *(Off-syllabus) A metric space is compact if and only if it complete and* ***totally bounded****. A space* M *is said to be totally bounded if for every* $\varepsilon > 0$ *there exist finite points* $x_1, \ldots, x_n$ *such that*

$$M \subseteq \bigcup_{i=1}^{n} B(x_i, \varepsilon).$$

SYLLABUS

Complex Analysis (22 lectures)

Basic geometry and topology of the complex plane, including the equations of lines and circles. [1]

Complex differentiation. Holomorphic functions. Cauchy-Riemann equations (including $z, \bar{z}$ version). Real and imaginary parts of a holomorphic function are harmonic. [2]

Recap on power series and differentiation of power series. Exponential function and logarithm function. Fractional powers — examples of multifunctions. The use of cuts as method of defining a branch of a multifunction. [3]

Path integration. Cauchy's Theorem. (Sketch of proof only — students referred to various texts for proof.) Fundamental Theorem of Calculus in the path integral/holomorphic situation. [3]

Cauchy's Integral formulae. Taylor expansion. Liouville's Theorem. Identity Theorem. Morera's Theorem. [3]

Laurent's expansion. Classification of isolated singularities. Calculation of principal parts, particularly residues. [2]

Residue Theorem. Evaluation of integrals by the method of residues (straightforward examples only but to include the use of Jordan's Lemma and simple poles on contour of integration). [3]

Extended complex plane, Riemann sphere, stereographic projection. Möbius transformations acting on the extended complex plane. Möbius transformations take circlines to circlines. [2]

Conformal mappings. Riemann mapping theorem (no proof): Möbius transformations, exponential functions, fractional powers; mapping regions (not Christoffel transformations or Joukowski's transformation). [3]

Reading

1. H. A. Priestley, *Introduction to Complex Analysis* (second edition, OUP, 2003).

Further Reading

1. L. Ahlfors, *Complex Analysis* (McGraw-Hill, 1979).
2. Reinhold Remmert, *Theory of Complex Functions* (Springer, 1989) (Graduate Texts in Mathematics 122).

7. Complex Algebra and Geometry

In the first year course *Introduction to Complex Numbers* the algebra and geometry of the complex numbers were defined. Further, in the Analysis I course the complex exponential and trigonometric functions were defined. We introduce some examples here by way of a reminder, and also move on the prove Apollonius' Theorem.

Example 127 *Express in the form $f(z) = a\bar{z} + b$ reflection in the line $x + y = 1$.*

Remark 128 *The isometries of $\mathbb{C}$ are the maps $f(z) = az + b$ and $f(z) = a\bar{z} + b$ where $|a| = 1$. The former are orientation-preserving and rotations. The latter are orientation-reversing.*

Solution Knowing from the remark that the reflection has the form $f(z) = a\bar{z} + b$ we can find a and b by considering where two points go to. As 1 and i both lie on the line of reflection then they are both fixed. So

$$a1 + b = a\bar{1} + b = 1, \qquad -ai + b = a\bar{i} + b = i.$$

Substituting $b = 1 - a$ into the second equation we find

$$a = \frac{1-i}{1+i} = -i,$$

and $b = 1 + i$. Hence $f(z) = -i\bar{z} + 1 + i$. ■

Example 129 *Prove the* ***Parallelogram Law*** *which states that:*

"A quadrilateral is a parallelogram if and only if the sum of the squares of the diagonals equals the sum of the squares of the sides."

Solution If a quadrilateral has vertices 0, a, b, c then we have

$$\begin{aligned}
& \text{parallelogram law holds} \\
\iff & |b|^2 + |a-c|^2 = |a|^2 + |c|^2 + |a-b|^2 + |b-c|^2 \\
\iff & b\bar{b} + a\bar{a} + c\bar{c} - a\bar{c} - \bar{a}c = a\bar{a} + c\bar{c} + a\bar{a} - a\bar{b} - \bar{a}b + b\bar{b} + b\bar{b} - b\bar{c} - \bar{b}c + c\bar{c} \\
\iff & a\bar{a} + a\bar{c} + \bar{a}c - a\bar{b} - \bar{a}b + b\bar{b} - b\bar{c} - \bar{b}c + c\bar{c} = 0 \\
\iff & (a + c - b)\left(\bar{a} + \bar{c} - \bar{b}\right) = 0 \\
\iff & |b - a - c|^2 = 0 \\
\iff & b = a + c \\
\iff & 0abc \text{ is a parallelogram.}
\end{aligned}$$

■

Example 130 *Find the image of each given region $A_i \subseteq \mathbb{C}$ under the given map $f_i \colon A_i \to \mathbb{C}$.*

- A_1 is the line $\operatorname{Re} z = 1$ and $f_1(z) = \sin z$;

A general point on the line has the form $z = 1 + it$ and we have

$$f_1(z) = \sin(1 + it) = \sin 1 \cos it + \cos 1 \sin it = \sin 1 \cosh t + i \cos 1 \sinh t.$$

So the image is given parametrically by

$$x(t) = \sin 1 \cosh t, \qquad y(t) = \cos 1 \sinh t.$$

(Note in particular that $\sin z$ is unbounded here!) Eliminating t by means of the identity $\cosh^2 t - \sinh^2 t = 1$ we see the image is one branch of the hyperbola

$$\frac{x^2}{\sin^2 1} - \frac{y^2}{\cos^2 1} = 1.$$

- A_2 is the region $\operatorname{Im} z > \operatorname{Re} z > 0$ and $f_2(z) = z^2$;

A general point in A_2 can be written in the form $z = re^{i\theta}$ where $\pi/4 < \theta < \pi/2$ and $r > 0$. As $z^2 = r^2 e^{i2\theta}$ has double the argument of z then

$$f_2(A_2) = \{z \in \mathbb{C} : \pi/2 < \arg z < \pi\}.$$

- A_3 is the unit disc $|z| < 1$ and $f_3(z) = (1 + z)/(1 - z)$.

The map f_3 is a bijection from the set $\mathbb{C} \setminus \{1\}$ to $\mathbb{C} \setminus \{-1\}$, with $f_3^{-1}(z) = (z - 1)/(z + 1)$. To see this note

$$w = \frac{1 + z}{1 - z} \iff w - wz = 1 + z \iff z(1 + w) = w - 1 \iff z = \frac{w - 1}{1 + w}.$$

So

$$z \in f_3(A_3) \iff f_3^{-1}(z) \in A_3 \iff \left|\frac{z - 1}{z + 1}\right| < 1 \iff |z - 1| < |z + 1|$$

which is the half-plane of points closer to 1 than -1, or equivalently the half-plane $\operatorname{Re} z > 0$.

Example 131 *Let ABC be a triangle. Prove the cosine rule*

$$|BC|^2 = |AB|^2 + |AC|^2 - 2|AB||AC|\cos \hat{A}. \tag{7.1}$$

Solution We can choose complex coordinates in the plane so that A is at the origin and B is at 1. Let C be at the point z. So in terms of our co-ordinates:

$$|AB| = 1, \qquad |BC| = |z - 1|, \qquad |AC| = |z|, \qquad \hat{A} = \arg z.$$

So

$$\begin{aligned}
\text{RHS of (7.1)} &= |z|^2 + 1 - 2|z|\cos\arg z = z\bar{z} + 1 - 2|z| \times \frac{\operatorname{Re} z}{|z|} \\
&= z\bar{z} + 1 - 2 \times \frac{(z + \bar{z})}{2} = z\bar{z} + 1 - z - \bar{z} \\
&= (z - 1)(\bar{z} - 1) = |z - 1|^2 = \text{LHS of (7.1)}.
\end{aligned}$$

■

Example 132 *Let $\omega = e^{2\pi i/3}$. Show that the triangle abc in $\mathbb{C}$ (with the vertices taken in anticlockwise order) is equilateral if and only if*

$$a + \omega b + \omega^2 c = 0$$

Solution Firstly note that $1 + \omega + \omega^2 = 0$, as $0 = \omega^3 - 1 = (\omega - 1)(\omega^2 + \omega + 1)$. The triangle abc is equilateral if and only if $c - b$ is the side $b - a$ rotated through $2\pi/3$ anticlockwise — i.e. if and only if

$$\begin{aligned} & c - b = \omega(b - a) \\ \iff \quad & \omega a + (-1 - \omega) b + c = 0 \\ \iff \quad & \omega a + \omega^2 b + c = 0 \\ \iff \quad & a + \omega b + \omega^2 c = 0. \end{aligned}$$

■

Proposition 133 *The equation*

$$Az\bar{z} + B\bar{z} + \bar{B}z + C = 0, \tag{7.2}$$

where $A, C \in \mathbb{R}$ and $B \in \mathbb{C}$ represents:-

(i) a line when $A = 0$;

(ii) a circle if $A \neq 0$ and $|B|^2 \geqslant AC$ with centre $-B/A$ and radius $\frac{1}{|A|}\sqrt{|B|^2 - AC}$; or otherwise has no solutions. All circles and lines can be represented in this way.

Proof If $A \neq 0$ then we can rewrite (7.2) as

$$z\bar{z} + \frac{B}{A}\bar{z} + \frac{\bar{B}}{A}z + \frac{C}{A} = 0, \left| z + \frac{B}{A} \right|^2 = \frac{|B|^2 - AC}{A^2}$$

which is the given circle described in (ii) unless $|B|^2 < AC$ in which case there are no solutions. If $A = 0$ then our equation reads

$$B\bar{z} + \bar{B}z + C = 0 \iff 2ux + 2vy + C = 0$$

where $B = u + iv$ and $z = x + iy$ and this is clearly a line. Clearly any line is of this form and likewise any circle $|z - a|^2 = r^2$ can be put into the form of (7.2). ■

Apollonius of Perga (c. 260-190 BC) wrote a book *Conics,* which was considered second only to the Elements in its importance. We meet here a description of circles due to him. All circles can be represented in this way. In this form a circle is often referred to as a *Circle of Apollonius.*

Theorem 134 ***(Apollonius' Theorem)*** *Let k be positive, with $k \neq 1$, and let α, β be distinct complex numbers. Then the locus of points satisfying the equation*

$$|z - \alpha| = k\,|z - \beta|$$

is a circle with centre c and radius r where

$$c = \frac{k^2\beta - \alpha}{k^2 - 1}, \qquad r = \frac{k\,|\alpha - \beta|}{|k^2 - 1|}.$$

Further α and β are inverse points — i.e. α and β are collinear with c and $|c - \alpha|\,|c - \beta| = r^2$.

Proof Squaring up the equation we have

$$(z-\alpha)(\bar{z}-\bar{\alpha}) = k^2 (z-\beta)\left(\bar{z}-\bar{\beta}\right)$$

and rearranging this becomes

$$\begin{aligned}
\left(k^2-1\right) z\bar{z} + \left(\bar{\alpha}-k^2\bar{\beta}\right) z + \left(\alpha - k^2\beta\right)\bar{z} &= \alpha\bar{\alpha} - k^2\beta\bar{\beta} \\
\iff \left| z - \left(\frac{k^2\beta-\alpha}{k^2-1}\right)\right|^2 &= \frac{\alpha\bar{\alpha}-k^2\beta\bar{\beta}}{k^2-1} + \frac{\left|k^2\beta-\alpha\right|^2}{\left(k^2-1\right)^2} \\
\iff \left| z - \left(\frac{k^2\beta-\alpha}{k^2-1}\right)\right|^2 &= \frac{\left(\alpha\bar{\alpha}-k^2\beta\bar{\beta}\right)\left(k^2-1\right) + \left(k^2\beta-\alpha\right)\left(k^2\bar{\beta}-\bar{\alpha}\right)}{\left(k^2-1\right)^2} \\
\iff \left| z - \left(\frac{k^2\beta-\alpha}{k^2-1}\right)\right|^2 &= \frac{k^2\alpha\bar{\alpha} + k^2\beta\bar{\beta} - k^2\alpha\bar{\beta} - k^2\bar{\alpha}\beta}{\left(k^2-1\right)^2} \\
\iff \left| z - \left(\frac{k^2\beta-\alpha}{k^2-1}\right)\right|^2 &= \frac{k^2\left|\alpha-\beta\right|^2}{\left(k^2-1\right)^2},
\end{aligned}$$

which is a circle with centre c and radius r as given above. Note that c is collinear with α and β as

$$c = \alpha + \frac{k^2}{k^2-1}(\beta-\alpha) = \beta + \frac{1}{k^2-1}(\beta-\alpha)$$

and from these formulae we see that

$$|c-\alpha|\,|c-\beta| = \frac{k^2}{\left(k^2-1\right)^2}\,|\beta-\alpha|^2 = r^2.$$

■

8. Holomorphic Functions

We move on from algebra and geometry now to look at what it means for a complex function to be differentiable. It will turn out that being differentiable on the complex plane is a stronger requirement than it is on the real line. Consequently we will be able to prove a much richer analytical theory than is possible on the real line, whilst still considering a class of functions general enough to include all the important everyday functions of mathematics.

Notation 135 *In line with Priestley's notation, for $a \in \mathbb{C}$ and $r > 0$, we will write*

$$\begin{array}{rcl} \textit{open disc} & : & D(a,r) = \{z \in \mathbb{C} : |z-a| < r\}\,; \\ \textit{closed disc} & : & \bar{D}(a,r) = \{z \in \mathbb{C} : |z-a| \leqslant r\}\,; \\ \textit{punctured disc} & : & D'(a,r) = \{z \in \mathbb{C} : 0 < |z-a| < r\}\,. \end{array}$$

Definition 136 *Let U be an open subset of $\mathbb{C}$ and $f\colon U \to \mathbb{C}$. Let $z_0 \in U$. Then:*
(a) f is said to be ***differentiable*** *at z_0 if*

$$\lim_{h\to 0} \frac{f(z_0+h)-f(z_0)}{h}$$

exists, in which case this limit is denoted $f'(z_0)$. Note that, precisely, this means that

$$\exists L \in \mathbb{C} \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall h \in D'(0,\delta) \qquad \left| \frac{f(z_0+h)-f(z_0)}{h} - L \right| < \varepsilon.$$

Note, in particular, that this is sufficient to guarantee f is continuous at z_0 as

$$f(z_0+h) = f(z_0) + hf'(z_0) + o(h) \to f(z_0) \quad \textit{as } h \to 0.$$

(b) f is said to be ***holomorphic*** *(or* ***analytic****) on U if it is differentiable at every point of U.*

Remark 137 *Note for the limit in (a) to exist, h must be able to approach 0 from any direction; this is a stronger requirement than the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ existing which would just require that the limit to exist as h approaches 0 along either axis.*

Example 138 *(a) z^2 is holomorphic on $\mathbb{C}$*
(b) $\bar{z}$ is not differentiable anywhere in $\mathbb{C}$.
(c) z^{-1} is holomorphic on $\mathbb{C}\backslash\{0\}$.

Solution (a) For any $z, h \in \mathbb{C}$, $h \neq 0$,

$$\frac{(z+h)^2 - z^2}{h} = 2z + h \to 2z \quad \text{as } h \to 0.$$

(b) For any $z, h \in \mathbb{C}$, $h \neq 0$,

$$\frac{\overline{z+h} - \bar{z}}{h} = \frac{\bar{h}}{h} \to \begin{cases} 1 & \text{as } h \to 0 \text{ along positive real axis} \\ -1 & \text{as } h \to 0 \text{ along positive imaginary axis} \end{cases}$$

so that the required limit exists nowhere.

(c) For any $z, h \neq 0$,

$$\frac{1}{h}\left(\frac{1}{z+h}-\frac{1}{z}\right)=\frac{-1}{z\,(z+h)}\to\frac{-1}{z^2}\quad \text{as } h\to 0$$

by the Algebra of Limits, provided $z \neq 0$. ■

Remark 139 ***(Algebra of Differentiation)*** *The addition, product, quotient and chain rules of real calculus all apply in a similar fashion when doing complex differentiation. This is entirely down to the fact that the Algebra of Limits for real sequences and functions carries over to complex sequences and functions.*

It is important to note that differentiability, in the sense of Definition 136, of a function $f(x+iy)$ requires more than just the partial derivatives $\partial f/\partial x$ and $\partial f/\partial y$ existing. We see this straight away now in the Cauchy-Riemann equations.

Theorem 140 ***(Cauchy-Riemann Equations)*** *Let $f\colon U \to \mathbb{C}$ be differentiable at $z_0 \in U$. For $z \in U$ let $x = \operatorname{Re} z$ and $y = \operatorname{Im} z$, and let $u = \operatorname{Re} f$, $v = \operatorname{Im} f$. Then, at z_0,*

$$u_x = v_y, \qquad u_y = -v_x.$$

Proof The proof follows by approaching z_0 horizontally and vertically and equating the limits. Let $z_0 = x_0 + iy_0 \in \mathbb{C}$. Then

$$\begin{aligned}
f'(z_0) &= \lim_{h\to 0,\, h\in\mathbb{R}} \frac{f(z_0+h)-f(z_0)}{h}\\
&= \lim_{h\to 0} \frac{[u(x_0+h,y_0)+iv(x_0+h,y_0)]-[u(x_0,y_0)+iv(x_0,y_0)]}{h}\\
&= \lim_{h\to 0}\left\{\left(\frac{u(x_0+h,y_0)-u(x_0,y_0)}{h}\right)+i\left(\frac{v(x_0+h,y_0)-v(x_0,y_0)}{h}\right)\right\}\\
&= u_x(x_0,y_0)+iv_x(x_0,y_0).
\end{aligned}$$

Likewise

$$\begin{aligned}
f'(z_0) &= \lim_{h\to 0,\, h\in\mathbb{R}} \frac{f(z_0+ih)-f(z_0)}{ih}\\
&= \lim_{h\to 0} \frac{[u(x_0,y_0+h)+iv(x_0,y_0+h)]-[u(x_0,y_0)+iv(x_0,y_0)]}{ih}\\
&= \lim_{h\to 0}\left\{\frac{1}{i}\left(\frac{u(x_0,y_0+h)-u(x_0,y_0)}{h}\right)+\left(\frac{v(x_0,y_0+h)-v(x_0,y_0)}{h}\right)\right\}\\
&= -iu_y(x_0,y_0)+v_y(x_0,y_0).
\end{aligned}$$

Hence, comparing real and imaginary parts, $u_x = v_y$ and $u_y = -v_x$. ■

Example 141 *Show directly that the real and imaginary parts of z^3 satisfy the CREs.*

Solution As $z^3 = (x+iy)^3 = (x^3 - 3xy^2) + i\,(3x^2y - y^3)$ then we see

$$\begin{aligned} u_x &= \left(x^3 - 3xy^2\right)_x = 3x^2 - 3y^2; \qquad v_y = \left(3x^2y - y^3\right)_y = 3x^2 - 3y^2. \\ u_y &= \left(x^3 - 3xy^2\right)_y = -6xy; \qquad -v_x = \left(y^3 - 3x^2y\right)_x = -6xy. \end{aligned}$$

■

Example 142 *Show that* $\operatorname{Re} z$ *and* $\operatorname{Im} z$ *are differentiable nowhere in* $\mathbb{C}$.

Solution For $\operatorname{Re} z$ we have $u = x, v = 0$ and so $u_x = 1 \neq 0 = v_y$. For $\operatorname{Im} z$ we have $u = y, v = 0$ and so $u_y = 1 \neq 0 = -v_x$. (Note that we could also have demonstrated these facts directly from the definition of differentiability.) ■

Another way of thinking about differentiability is via the Wirtinger derivatives. These are notationally convenient, but also reinforce the idea that if f is differentiable then it is a function of z alone and independent of $\bar{z}$. Any function of $x = \operatorname{Re} z$ and $y = \operatorname{Im} z$ can instead be written in terms of z and $\bar{z}$ via the equations

$$x = \frac{z+\bar{z}}{2}, \qquad y = \frac{z-\bar{z}}{2i}.$$

The Cauchy-Riemann equations are then equivalent to $\partial f/\partial \bar{z} = 0$.

Definition 143 *The* ***Wirtinger derivatives****, for a complex valued function on* $\mathbb{C}$*, are defined as*

$$\frac{\partial}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right); \qquad \frac{\partial}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right).$$

Remark 144 *Note that these derivatives are* ***formally*** *what one expect from the chain rule as we'd have, for example,*

$$\frac{\partial}{\partial z} = \frac{\partial x}{\partial z}\frac{\partial}{\partial x} + \frac{\partial y}{\partial z}\frac{\partial}{\partial y} = \frac{1}{2}\frac{\partial}{\partial x} + \frac{1}{2i}\frac{\partial y}{\partial z}\frac{\partial}{\partial y} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right),$$

but the expression $\partial x/\partial z$ *has no natural meaning except via the Wirtinger derivatives.*

Corollary 145 *Let* $f \colon U \to \mathbb{C}$ *be a complex-valued function on an open set* $U \subseteq \mathbb{C}$*. The Cauchy-Riemann equations are equivalent to*

$$\frac{\partial f}{\partial \bar{z}} = 0,$$

and if f *is differentiable then*

$$\frac{\mathrm{d}f}{\mathrm{d}z} = \frac{\partial f}{\partial z}.$$

Proof If $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ then

$$\frac{\partial f}{\partial \bar{z}} = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(u(x,y) + iv(x,y)) = \frac{1}{2}\left[(u_x - v_y) + i\,(u_y + v_x)\right].$$

Comparing real and imaginary parts, we see $\partial f/\partial \bar{z} = 0$ if and only if $u_x = v_y$ and $u_y = -v_x$. If $\partial f/\partial \bar{z} = 0$ then

$$\frac{\partial f}{\partial z} = \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right)(u + iv) = \frac{1}{2}\left[(u_x + v_y) + i\,(-u_y + v_x)\right] = u_x + iv_x = \frac{\mathrm{d}f}{\mathrm{d}z}.$$

■

Example 146 *Note for the non-differentiable functions we have met so far*

$$\frac{\partial \bar{z}}{\partial \bar{z}} = 1; \qquad \frac{\partial\,(\operatorname{Re} z)}{\partial \bar{z}} = \frac{1}{2}; \qquad \frac{\partial\,(\operatorname{Im} z)}{\partial \bar{z}} = \frac{i}{2}.$$

Example 147 *Show that the function*

$$f(z) = \begin{cases} z^5/\left|z\right|^4 & z \neq 0 \\ 0 & 0 \end{cases}$$

is continuous at 0*, satisfies the Cauchy-Riemann equations at* $z = 0$*, but is not differentiable at* 0*. (So the Cauchy-Riemann equations holding at a point are not sufficient to guarantee differentiability at that point.)*

Solution Continuity at 0 follows from the fact that $|f(z)| = |z|$ for all z. Secondly

$$\frac{f(0+h) - f(0)}{h} = \frac{h^5/\left|h\right|^4}{h} = \frac{h^4}{\left|h\right|^4}$$

has no limit as $h \to 0$. To see this we can approach 0 along the ray $\arg z = \pm\pi/8$ and get the different limits of $\pm i$. Finally for $z \neq 0$,

$$u = \operatorname{Re} f = \frac{x^5 - 10x^3y^2 + 5xy^4}{\left(x^2 + y^2\right)^2}; \qquad v = \operatorname{Im} f = \frac{5x^4y - 10x^2y^3 + y^5}{\left(x^2 + y^2\right)^2}.$$

So at $(0,0)$ we have

$$u_x = \lim_{\varepsilon \to 0} \frac{u(\varepsilon,0) - u(0,0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{\varepsilon - 0}{\varepsilon} = 1; \quad u_y = \lim_{\varepsilon \to 0} \frac{u(0,\varepsilon) - u(0,0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{0 - 0}{\varepsilon} = 0;$$

$$v_x = \lim_{\varepsilon \to 0} \frac{v(\varepsilon,0) - v(0,0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{0 - 0}{\varepsilon} = 0; \quad v_y = \lim_{\varepsilon \to 0} \frac{v(0,\varepsilon) - v(0,0)}{\varepsilon} = \lim_{\varepsilon \to 0} \frac{\varepsilon - 0}{\varepsilon} = 1.$$

In fact there are more pathological examples – the function

$$f(z) = \begin{cases} \exp(-z^{-4}) & z \neq 0 \\ 0 & z = 0 \end{cases}$$

satisfies the Cauchy-Riemann equations everywhere on $\mathbb{C}$ but is not even continuous at 0. ■

But by way of a partial converse to the Cauchy-Riemann equations, we can prove the following result.

Proposition 148 *(Goursat) Given $f\colon U \to \mathbb{C}$, let $u = \operatorname{Re} f$, $v = \operatorname{Im} f$. Suppose that u_x, u_y, v_x, v_y exist, are continuous and satisfy the Cauchy-Riemann Equations at $z \in U$. Then f is differentiable at z.*

Proof Let $z = x + iy \in U$ and $\varepsilon > 0$ be such that $D(z,\varepsilon) \subseteq U$. Let $h = p + iq$ where $|h| < \varepsilon$. Then $(f(z+h) - f(z))/h$ equals

$$\begin{aligned}
& \frac{p}{h}\left(\frac{u\,(x+p,y+q) - u\,(x,y+q) + iv\,(x+p,y+q) - iv(x,y+q)}{p}\right) \\
& + \frac{q}{h}\left(\frac{u\,(x,y+q) - u\,(x,y) + iv\,(x,y+q) - iv(x,y)}{q}\right) \\
= \ & \frac{p}{h}\left(u_x\,(x+\theta_1 p, y+q) + iv_x\,(x+\theta_2 p, y+q)\right) + \frac{q}{h}\left(u_y\,(x, y+\theta_3 q) + iv_y\,(x, y+\theta_4 q)\right)
\end{aligned}$$

by the Mean-value Theorem and where each $\theta_i \in (0,1)$. Using the Cauchy-Riemann Equations, the above can be rewritten as

$$\frac{[pu_x\,(x+\theta_1 p, y+q) + iqu_x\,(x, y+\theta_4 q)]}{p+iq} + \frac{i\,[pv_x\,(x+\theta_2 p, y+q) + iqv_x\,(x, y+\theta_3 q)]}{p+iq}.$$

Finally, using the continuity of u_x and v_x we see that

$$f'(z) = \lim_{h \to 0}\left(\frac{f(z+h) - f(z)}{h}\right) = u_x(x,y) + iv_x(x,y)$$

exists. ■

Remark 149 *There is a stronger result and something closer to a full converse; this is the* ***Looman-Menchoff Theorem*** *(1923) which states that if f is continuous on an open set U and satisfies the Cauchy-Riemann equations on all of U, then f is holomorphic. But this theorem is beyond the scope of this introductory course.*

Remark 150 ***(Local behaviour of a holomorphic map)*** *Consider a holomorphic map f defined on an open set U with $z_0 = x_0 + iy_0 \in U$. We will consider a nearby point $z_0 + h$ where $|h| < \varepsilon$ The existence of $f'(z_0)$ means that*

$$f(z_0 + h) = f(z_0) + hf'(z_0) + o(h).$$

So, viewing h as a coordinate on $D(a,\varepsilon)$, we see that the linear approximation of f on $D(z_0,\varepsilon)$ is multiplication by $f'(z_0)$, which geometrically represents scaling by $|f'(z_0)|$ and rotation by $\arg f'(z_0)$, followed by translation by $f(z_0)$. If we instead use two real coordinates, say $h = \alpha + i\beta$, and write $u = \operatorname{Re} f$, $v = \operatorname{Im} f$, then we see

$$\begin{aligned}
\begin{pmatrix} u\,(x_0+\alpha, y_0+\beta) \\ v(x_0+\alpha, y_0+\beta) \end{pmatrix} & \approx \begin{pmatrix} u(x_0,y_0) + u_x(x_0,y_0)\alpha + u_y(x_0,y_0)\beta \\ v(x_0,y_0) + v_x(x_0,y_0)\alpha + v_y(x_0,y_0)\beta \end{pmatrix} \\
& = \begin{pmatrix} u(x_0,y_0) \\ v(x_0,y_0) \end{pmatrix} + \begin{pmatrix} u_x(x_0,y_0) & u_y(x_0,y_0) \\ v_x(x_0,y_0) & v_y(x_0,y_0) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \\
& = \begin{pmatrix} u(x_0,y_0) \\ v(x_0,y_0) \end{pmatrix} + \begin{pmatrix} u_x(x_0,y_0) & u_y(x_0,y_0) \\ -u_y(x_0,y_0) & u_x(x_0,y_0) \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} \qquad \text{[by CREqs]} \\
& = \begin{pmatrix} u(x_0,y_0) \\ v(x_0,y_0) \end{pmatrix} + \lambda \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\ \theta \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix}
\end{aligned}$$

where $\lambda = \sqrt{u_x^2 + u_y^2} = |f'(z_0)|$ *and* $\theta = \arg f'(z_0)$. *Again we see that the matrix*

$$\lambda \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\ \theta \end{pmatrix}$$

denotes a scaling and rotation. Those who go on to do Multivariable Calculus will recognize this as the derivative of f *at* z_0 *(essentially the Jacobian).*

We shall see, in due course, that the derivative of a holomorphic function is itself holomorphic; however for now it seems sensible just to assume this so that we can state the following.

Corollary 151 *Let* f *be holomorphic on an open set* U. *Then* $u = \operatorname{Re} f$ *and* $v = \operatorname{Im} f$ *are harmonic functions (i.e. they satisfy Laplace's equation).*

Proof As f is holomorphic then we have $u_x = v_y$ and $u_y = -v_x$ in the usual way. But, with our assumption, the real and imaginary parts of

$$f' = u_x + iv_x$$

also satisfy the Cauchy-Riemann equations. Hence

$$(u_x)_x = (v_x)_y\,; \qquad (u_x)_y = -(v_x)_x$$

with all these second-derivatives themselves being continuous. So

$$\begin{aligned} u_{xx} &= (v_x)_y = (-u_y)_y = -u_{yy} \implies u_{xx} + u_{yy} = 0; \\ v_{xx} &= -(u_x)_y = -(v_y)_y = -v_{yy} \implies v_{xx} + v_{yy} = 0. \end{aligned}$$

■

Example 152 *Show directly that the following functions* $u(x,y)$ *and* $U(x,y)$ *are harmonic on* $\mathbb{C}$.

$$u(x,y) = x^3 - 3xy^2; \qquad U(x,y) = \sin x \cosh y$$

Find holomorphic functions f *and* F *such that* $u = \operatorname{Re} f$ *and* $U = \operatorname{Re} F$.

Solution Firstly

$$\nabla^2 u = (6x) + (-6x) = 0; \qquad \nabla^2 U = (-\sin x \cosh y) + (\sin x \cosh y) = 0.$$

By inspection we note that

$$\begin{aligned} \operatorname{Re}(x+iy)^3 &= \operatorname{Re}\left(x^3 + 3ix^2y - 3xy^2 - iy^3\right) = x^3 - 3xy^2 = u; \\ \operatorname{Re}\sin(x+iy) &= \operatorname{Re}(\sin x \cosh y + i\cos x \sinh y) = \sin x \cosh y = U. \end{aligned}$$

■

Definition 153 *If* u, v *are harmonic functions on the open set* $U \subseteq \mathbb{C}$ *such that* $u + iv$ *is holomorphic on* U, *then* v *is said to be a* ***harmonic conjugate*** *of* u..

9. Power Series. Complex Exponential.

Recall the following from the first year course.

Definition 154 *Let $a \in \mathbb{C}$. A* ***power series centred on*** *a is a series of the form*

$$\sum_{n=0}^{\infty} a_n (z-a)^n \tag{9.1}$$

where (a_n) is a complex sequence and $z \in \mathbb{C}$. We here treat (a_n) as given and z as a complex variable so that the series, where it converges, determines a function in z.

Theorem 155 *Given a power series $\sum_{n=0}^{\infty} a_n (z-a)^n$, there exists a unique R in the range $0 \leqslant R \leqslant \infty$ called the* ***radius of convergence*** *of the series such that:*

$$\begin{aligned} \text{(9.1) converges absolutely when } |z-a| &< R, \\ \text{(9.1) diverges when } |z-a| &> R. \end{aligned}$$

Further, within the ***disc of convergence*** *$|z-a| < R$, the power series $\sum_{n=0}^{\infty} a_n (z-a)^n$ defines a differentiable function $f(z)$ and*

$$f'(z) = \sum_{n=1}^{\infty} n a_n (z-a)^{n-1}.$$

(i.e. term-by-term differentiation is valid within the disc of convergence.)

Proposition 156 *Given a power series $\sum_{n=0}^{\infty} a_n (z-a)^n$, if*

$$L = \lim_{n\to\infty} \frac{a_n}{a_{n+1}}$$

exists then $R = |L|$.

Example 157 *(a) Using the previous proposition, the radius of convergence of*

$$\sum_{n=0}^{\infty} z^n$$

is 1 and so the series defines a holomorphic function on $D(0,1)$. The function it defines is $(1-z)^{-1}$. The series converges nowhere on the boundary $|z| = 1$.

(b) By differentiating term by term in $D(0,1)$ we get

$$(1-z)^{-2} = \sum_{n=1}^{\infty} n z^{n-1} = \sum_{n=0}^{\infty} (n+1) z^n.$$

We can also arrive at this by multiplying series; so on $D(0,1)$ we can validly write

$$\begin{aligned}(1-z)^{-2} &= (1-z)^{-1}(1-z)^{-1} = \left(\sum_{k=0}^{\infty} z^k\right)\left(\sum_{l=0}^{\infty} z^l\right) \\ &= \sum_{k=0}^{\infty}\sum_{l=0}^{\infty} z^{k+l} \\ &= \sum_{n=0}^{\infty} z^n \sum_{\substack{k+l=n \\ k,l\geqslant 0}} 1 \\ &= \sum_{n=0}^{\infty} (n+1)\, z^n.\end{aligned}$$

(c) We can find a power series for the function $(1-z)^{-1}$ centred at any $a \neq 1$. For example, with $a = i$, we see

$$\begin{aligned}(1-z)^{-1} &= (1-i-(z-i))^{-1} \\ &= (1-i)^{-1}\left(1-\left(\frac{z-i}{1-i}\right)\right)^{-1} \\ &= \sum_{n=0}^{\infty} \frac{(z-i)^n}{(1-i)^{n+1}}\end{aligned}$$

and this has radius of convergence $\sqrt{2}$.

Example 158 *Determine the power series, centred at* 0, *for* $(1+z+z^2)^{-1}$.

Solution Note that, for $|z| < 1$, we have

$$\begin{aligned}\frac{1}{1+z+z^2} &= \frac{1-z}{1-z^3} \\ &= (1-z)\left(1+z^3+z^6+\cdots\right) \\ &= 1-z+z^3-z^4+z^6-z^7+\cdots\end{aligned}$$

■

Remark 159 *We can start now to appreciate the radius of convergence of a power series as something naturally linked to the function it defines. In due course we will prove Taylor's Theorem that a holomorphic function on $D(a,r)$ has a convergent power series $\sum_0^\infty a_n(z-a)^n$. A function $f(z)$ which is holomorphic at a has a power series, and the radius of convergence is R where R is the distance from a to the nearest* ***singularity****. If f is holomorphic on $D(a,R)$ then by Taylor's Theorem f has a convergent power series on $D(a,R)$; on the other hand if the power series converged on $D(a,S)$ where $S > R$ then f would be holomorphic at the singularity as power series define differentiable functions.*

Remark 160 *Consider the set of power series with disc of convergence $D(0,R)$ where R is finite. It is still an open problem as to which subsets of the boundary are sets of convergence for some power series. It was shown by Piranian and Herzog (1949,1953) that any countable union of closed subsets of the boundary is the set of convergence for some power series.*

However there is a theorem of Abel's which goes a little way to describing behaviour of a power series on the boundary of convergence.

Theorem 161 *(Abel) Assume that we have, for fixed θ,*

$$f(r) = \sum_{n=0}^{\infty} a_n (re^{i\theta})^n \qquad \text{converges for } -R < r < R.$$

If the series also converges at $r = R$ then the limit $\lim_{r \to R_-} f(r)$ *exists and*

$$\lim_{r \to R_-} f(r) = \sum_{n=0}^{\infty} a_n (Re^{i\theta})^n.$$

Remark 162 *This theorem is off-syllabus but we shall make use of it in Exercise Sheet 4, Questions 7 and 8. For a proof see, for example, Apostol's Mathematical Analysis.*

Further it was shown in Analysis I that

Theorem 163 *The following power series define the complex exponential and trigonometric functions on the entire complex plane.*

$$\exp z = \sum_{n=0}^{\infty} \frac{z^n}{n!}; \qquad \sin z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!}; \qquad \cos z = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!}.$$

Further these functions have the following properties, that for any $z, w \in \mathbb{C}$:

(a) $\exp' z = \exp z$, $\sin' z = \cos z$ *and* $\cos' z = -\sin z$.

(b) $\exp(z+w) = \exp z \times \exp w$.

(c) $\exp(iz) = \cos z + i \sin z$.

(d)
$$\cos z = \frac{\exp(iz) + \exp(-iz)}{2}; \qquad \sin z = \frac{\exp(iz) - \exp(-iz)}{2i}.$$

(e) $\sin^2 z + \cos^2 z = 1$.

(f) $\cos z$ *and* $\sin z$ *have period* 2π *and no smaller period.*

(g) $\exp z$ *has period* $2\pi i$ *and no smaller period.*

(h) $\exp r = e^r$ *for real values of* r *(and so we will often write* e^z *for* $\exp z$*).*

Example 164 *Show that e^z takes all non-zero values.*

Solution Let $z = x + iy$ so that

$$e^z = e^{x+iy} = e^x e^{iy}.$$

We see that e^x is the modulus of e^z and can take any positive value. y is the argument of e^z and so e^z can take any argument (for infinitely many y in fact) and hence the image of exp is $\mathbb{C}\backslash\{0\}$. ■

Example 165 *Show that the only solutions of* $e^z = 1$ *are* $2n\pi i$ *where* $n \in \mathbb{Z}$. *Find all the solutions of* $\sin z = 1$.

Solution If $e^{x+iy} = 1$, where x and y are real, then equating moduli we see that $e^x = 1$ and $x = 0$. Then $\cos y + i \sin y = 1$ and we see that $\cos y = 1$, $\sin y = 0$ so that $y = 2n\pi$.

We also have

$$\begin{aligned} & \sin z = \frac{e^{iz} - e^{-iz}}{2i} = 1 \\ \iff & e^{2iz} - 2ie^{iz} - 1 = 0 \\ \iff & \left(e^{iz} - i\right)^2 = 0 \\ \iff & e^{iz} = i. \end{aligned}$$

If we set $z = x + iy$ we then have

$$\begin{aligned} e^{ix}e^{-y} = i \iff & \quad y = 0 \quad \text{and} \quad \cos x = 0, \sin x = 1 \\ \iff & \quad y = 0 \quad \text{and} \quad x = \frac{\pi}{2} + 2n\pi \quad \text{for some } n \in \mathbb{Z}. \end{aligned}$$

(So $\sin z = 1$ has no further solutions beyond the real ones that we are accustomed to.) ■

Example 166 *Let* $z = x + iy$. *Then*

$$\sin(x + iy) = \sin x \cosh y + i \cos x \sinh y.$$

Solution Note that

$$\begin{aligned} \text{RHS} &= \left(\frac{e^{ix} - e^{-ix}}{2i}\right)\left(\frac{e^y + e^{-y}}{2}\right) + i\left(\frac{e^{ix} + e^{-ix}}{2}\right)\left(\frac{e^y - e^{-y}}{2}\right) \\ &= \left(\frac{e^{ix} - e^{-ix}}{2i}\right)\left(\frac{e^y + e^{-y}}{2}\right) - \left(\frac{e^{ix} + e^{-ix}}{2i}\right)\left(\frac{e^y - e^{-y}}{2}\right) \\ &= \frac{2e^{ix-y} - 2e^{-ix+y}}{4i} = \text{LHS}. \end{aligned}$$

■

When it comes to the complex logarithm and other multifunctions (or multivalued functions) defined in terms of the complex logarithm (such as square root) then some considerable care needs to be taken. Certainly reckless assumptions that rules previously true of positive numbers remain so for complex or negative numbers leads quickly to issues.

Example 167 *Clearly*

$$\frac{-1}{1} = \frac{1}{-1}.$$

Taking square roots we find

$$\sqrt{\frac{-1}{1}} = \sqrt{\frac{1}{-1}}$$

and as $\sqrt{a/b} = \sqrt{a}/\sqrt{b}$ *then*

$$\frac{i}{1} = \frac{\sqrt{-1}}{\sqrt{1}} = \frac{\sqrt{1}}{\sqrt{-1}} = \frac{1}{i}.$$

This rearranges to give $i^2 = 1$*, which is plainly false as* $i^2 = -1$*. The error of course is in assuming that the rule* $\sqrt{a/b} = \sqrt{a}/\sqrt{b}$ *still holds.*

Proposition 168 *(a) Any* $z \in \mathbb{C}\backslash(-\infty, 0]$ *can be written as* $z = re^{i\theta}$ *where* $r > 0, \theta \in (-\pi, \pi)$ *in a unique fashion.*
(b) The function $L\colon \mathbb{C}\backslash(-\infty, 0] \to \mathbb{C}$ *given by*

$$L(z) = \log r + i\theta$$

satisfies $\exp(L(z)) = z$ *and is holomorphic with* $L'(z) = 1/z$*.*

Proof (a) follows from choosing $r = |z|$ and $\theta = \arg z$, which takes a unique principal value in the given range. For (b) we firstly note that

$$\exp(L(z)) = e^{\log r} e^{i\theta} = re^{i\theta} = z$$

and that

$$u(x, y) = \log \sqrt{x^2 + y^2} = \frac{1}{2}\log(x^2 + y^2); \qquad v(x, y) = \tan^{-1}(y/x)$$

(at least when $x > 0$). Then

$$\begin{aligned} u_x &= \frac{x}{x^2 + y^2}; \qquad v_y = \frac{\frac{1}{x}}{1 + \frac{y^2}{x^2}} = \frac{x}{x^2 + y^2}; \\ u_y &= \frac{y}{x^2 + y^2}; \qquad v_x = \frac{-\frac{y}{x^2}}{1 + \frac{y^2}{x^2}} = \frac{-y}{x^2 + y^2}. \end{aligned}$$

So the Cauchy-Riemann equations are satisfied, and u_x, u_y, v_x, v_y are clearly continuous, so by Proposition 148 we have that L is holomorphic. Further

$$L'(z) = u_x + iv_x = \frac{x - iy}{x^2 + y^2} = \frac{1}{x + iy} = \frac{1}{z}.$$

(Some care should really be taken with different formulas for argument separately for $y > 0$ and $y < 0$ but all the above follows in a near-identical fashion.) ■

Corollary 169 *Let* $\alpha \in \mathbb{C}$*. Then*

$$z^\alpha := \exp(\alpha L(z)) \tag{9.2}$$

defines a holomorphic function on $\mathbb{C}\backslash(-\infty, 0]$ *and*

$$\frac{\mathrm{d}}{\mathrm{d}z}(z^\alpha) = \alpha z^{\alpha - 1}.$$

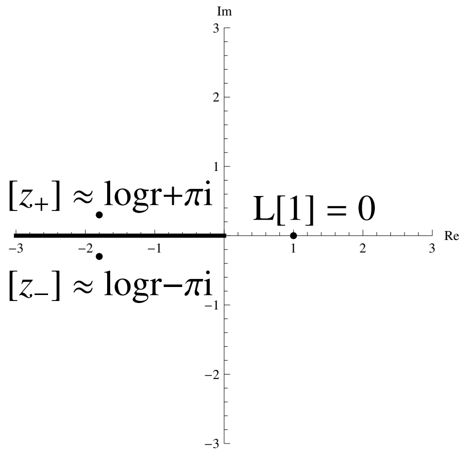
Proof As L and exp are holomorphic then so is z^α. Then by the chain rule

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}z}(z^\alpha) &= \frac{\mathrm{d}}{\mathrm{d}z}\exp(\alpha L(z)) = \alpha L'(z)\exp(\alpha L(z)) \\ &= \frac{\alpha}{z}\exp(\alpha L(z)) \\ &= \alpha\exp(\alpha L(z))\exp(-L(z)) \\ &= \alpha\exp((\alpha-1)L(z)) \\ &= \alpha z^{\alpha-1}.\end{aligned}$$

■

Remark 170 *The function $L(z)$ defined on $\mathbb{C}\backslash(-\infty,0]$ is a* ***holomorphic branch*** *(or just* ***branch****) of the complex logarithm. The other holomorphic branches of the logarithm on this cut plane are $L(z)+2n\pi i$. (See Exercise Sheet 4, Question 5(i)).*

We consider now $L(z)$ near the cut $(-\infty,0]$ and do similarly for $z^{1/2}$ and $z^{1/3}$ as defined in the above corollary.



If we were to take points z_+ and z_-, respectively just above and below the cut $(-\infty,0]$ then we would have

$$z_+ = re^{i\theta_+} \quad \text{where } \theta_+ \approx \pi; \qquad z_- = re^{i\theta_-} \quad \text{where } \theta_- \approx -\pi.$$

So

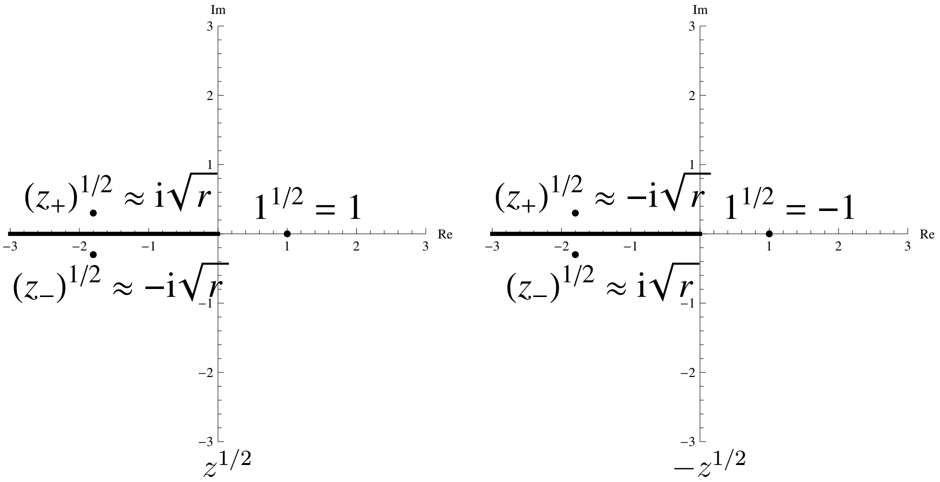
$$L(z_+) \approx \log r + i\pi; \qquad L(z_-) \approx \log r - i\pi,$$

hence there is a discontinuity of $2\pi i$ across the cut. If we consider similarly $z^{1/2}$ as defined in (9.2) we see

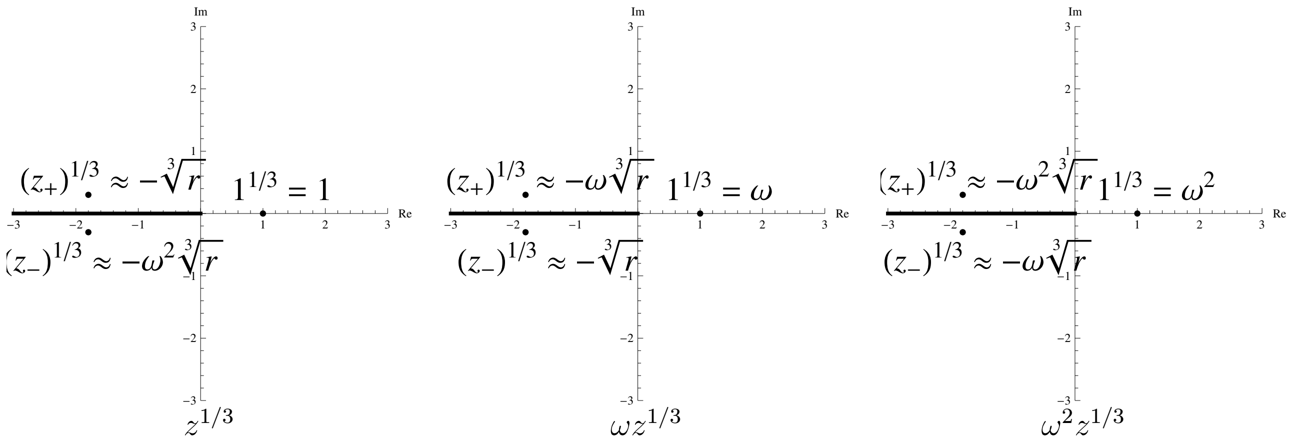
$$(z_+)^{1/2} \approx \sqrt{r}e^{i\pi/2} = i\sqrt{r}; \qquad (z_-)^{1/2} \approx \sqrt{r}e^{-i\pi/2} = -i\sqrt{r}.$$

We see this time that there is a sign change as we cross the cut.

The only other holomorphic function on $\mathbb{C}\backslash(-\infty,0]$ which satisfies $w^2=z$ is $w=-z^{1/2}$ (in the sense of (9.2)) and these two functions, $z^{1/2}$ and $-z^{1/2}$ are the two holomorphic branches *of $\sqrt{z}$. We see that as we cross the cut we move from one branch to the other.*



Something similar happens with $z^{1/3}$ but now there are three holomorphic branches on $\mathbb{C}\backslash(-\infty,0]$ namely $z^{1/3}$, $\omega z^{1/3}$ and $\omega^2 z^{1/3}$ where $\omega = e^{2\pi i/3}$ is a cube root of unity. Again we see that as we cross the cut we move from one branch to another. Specifically if we cross the cut in a downward direction then we move from $z^{1/3}$ to $\omega z^{1/3}$, from $\omega z^{1/3}$ to $\omega^2 z^{1/3}$, from $\omega^2 z^{1/3}$ to $z^{1/3}$.



Remark 171 *If we had wanted to define a holomorphic branch of the logarithm on a different cut plane, for example on $X = \mathbb{C}\backslash\{$negative imaginary axis$\}$, then this could have been done similarly. Any $z \in X$ can be uniquely written as $z = re^{i\theta}$ where $\theta \in (-\pi/2, 3\pi/2)$ and we would then set $\log z = \log r + i\theta$ as before. We could again prove that this function is holomorphic or we could simply note that this choice of logarithm is simply $L(z/i) + i\pi/2$.*

Given any simple (i.e. non-intersecting) curve C which runs from 0 to ∞ there are infinitely many branches of the logarithm that can be defined on $\mathbb{C}\backslash C$ which differ from each other by a constant which is an integer multiple of $2\pi i$. If the curve C wraps around the origin then such branches may share values with many or indeed all of the branches $L(z) + 2n\pi i$ that we defined for the cut plane $\mathbb{C}\backslash(-\infty,0]$.

Example 172 *Note that*

$$L(zw) \neq L(z) + L(w) \qquad z, w \in \mathbb{C}\backslash(-\infty,0]$$

in general. In fact it might even be the case that $L(zw)$ is not defined (e.g. $z = w = i$) Or we may find cases like $z = w = e^{3\pi i/4}$ where

$$L(zw) = -\pi i/2 \neq 3\pi i/2 = L(z) + L(w).$$

Of course what has happened is that zw has in effect moved into the domain of a different holomorphic branch.

Example 173 *Show that the power series*

$$\sum_{1}^{\infty} \frac{z^n}{n} \tag{9.3}$$

converges to $-L(1-z)$ in $D(0,1)$. Hence evaluate the two series

$$\frac{1}{2\times 2^2} - \frac{1}{4\times 2^4} + \frac{1}{6\times 2^6} - \frac{1}{8\times 2^8} + \cdots, \qquad \frac{1}{1\times 2} - \frac{1}{3\times 2^3} + \frac{1}{5\times 2^5} - \frac{1}{7\times 2^7} + \cdots.$$

Solution The power series is easily seen to have radius of convergence 1; denote by $f(z)$ the function it defines on $D(0,1)$. Differentiating term by term we see that

$$f'(z) = \sum_{1}^{\infty} z^{n-1} = \sum_{0}^{\infty} z^n = (1-z)^{-1} = -\frac{\mathrm{d}}{\mathrm{d}z} L(1-z).$$

A holomorphic function with zero derivative on a path-connected subset is constant. (We haven't proven this general result yet but will do in the next chapter.) So f and $-L(1-z)$ differ by a constant; as they agree at $z = 0$ then this constant is 0 and $f(z) = -L(1-z)$ on $D(0,1)$.

If we put $z = i/2$ into the series we get

$$\sum_{1}^{\infty} \frac{i^n}{n 2^n} = -L(1 - i/2) = -\left(\log \frac{\sqrt{5}}{2} - i \tan^{-1} \frac{1}{2}\right).$$

Taking the negative real part of both sides we get

$$\frac{1}{2\times 2^2} - \frac{1}{4\times 2^4} + \frac{1}{6\times 2^6} - \frac{1}{8\times 2^8} + \cdots = \frac{1}{2}\log 5 - \log 2$$

and taking the imaginary part we get

$$\frac{1}{1\times 2} - \frac{1}{3\times 2^3} + \frac{1}{5\times 2^5} - \frac{1}{7\times 2^7} + \cdots = \tan^{-1} \frac{1}{2}.$$

■

Remark 174 *The series in (9.3) has radius of convergence $R = 1$. It diverges at $z = 1$ (as this is just the harmonic series) but converges for all other z with $|z| = 1$. See Exercise Sheet 4, Question 7(ii).*

Instead of focussing on a principal set of values for a multivalued function we might instead seek to treat all values at once. Following notation introduced in Priestley we will write:

Notation 175 *Given $z \in \mathbb{C}$, $z \neq 0$, $a \in \mathbb{C}$ we define (using the notation of Priestley)*

$$\begin{aligned} [\arg z] &= \{\theta \in \mathbb{R} : z = |z| e^{i\theta}\}; \\ [\log z] &= \{w \in \mathbb{C} : e^w = z\}; \\ [z^a] &= \exp(a[\log z]). \end{aligned}$$

Proposition 176 *If $e^{w_0} = z$, show that*

$$[\log z] = w_0 + 2\pi i\mathbb{Z} \qquad \textit{and that} \qquad [\log z] = \log|z| + i[\arg z].$$

Proof Note that

$$\begin{aligned} w \in [\log z] &\iff e^w = z \\ &\iff e^w = e^{w_0} \\ &\iff e^{w-w_0} = 1 \\ &\iff w - w_0 \in 2\pi i\mathbb{Z} \\ &\iff w \in w_0 + 2\pi i\mathbb{Z}. \end{aligned}$$

Also

$$\begin{aligned} \theta \in [\arg z] &\iff e^{i\theta} = z/|z| \\ &\iff e^{i\theta + \log|z|} = z \\ &\iff \log|z| + i\theta \in [\log z]. \end{aligned}$$

■

Proposition 177 *For $z, w, a, b \in \mathbb{C}$ with $z, w \neq 0$*
(a) $[\log z] + [\log w] = [\log(zw)]$;
(b) $[z^a] \cdot [w^a] = [(zw)^a]$
(c) in general $[z^a] \cdot [z^b]$ *does not equal* $[z^{a+b}]$.

Proof (a) If $w_0 \in [\log w]$ and $z_0 \in [\log z]$ then

$$e^{w_0+z_0} = e^{w_0} e^{z_0} = wz \implies w_0 + z_0 \in [\log(zw)].$$

So

$$\begin{aligned} [\log z] + [\log w] &= (z_0 + 2\pi i\mathbb{Z}) + (w_0 + 2\pi i\mathbb{Z}) \\ &= z_0 + w_0 + (2\pi i\mathbb{Z} + 2\pi i\mathbb{Z}) \\ &= z_0 + w_0 + 2\pi i\mathbb{Z} \\ &= [\log(zw)]. \end{aligned}$$

(b)

$$\begin{aligned}
[z^a] \cdot [w^a] &= \exp(a[\log z]) \exp(a[\log w]) \\
&= \exp\left(a\left[\log z\right] + a\left[\log w\right]\right) \\
&= \exp\left(a\left(\left[\log z\right] + \left[\log w\right]\right)\right) \\
&= \exp\left(a\left[\log(zw)\right]\right) \\
&= \left[(zw)^a\right].
\end{aligned}$$

(c) But if we take $z = w = 1$ and $a = b = 1/2$ then we see that

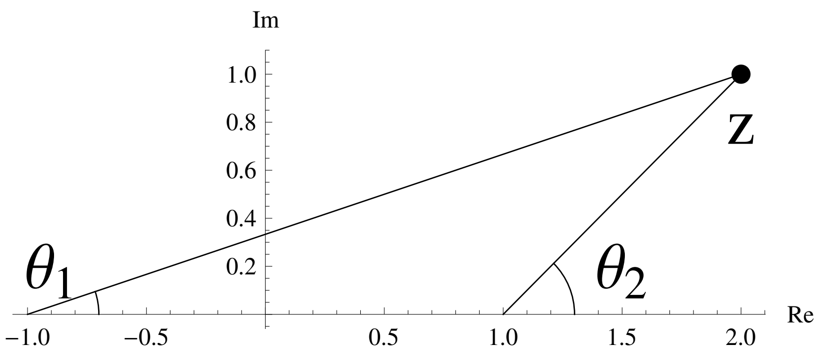
$$\left[1^{1/2}\right] \cdot \left[1^{1/2}\right] = \{1, -1\} \cdot \{1, -1\} = \{1, -1\} \neq \{1\} = \left[1^1\right].$$

■

Example 178 *For z in the cut plane $\mathbb{C}\backslash(-\infty, 1]$ we will let*

$$\begin{array}{llll}
\theta_1 & \textit{denote the value of } \arg(z+1) & \textit{in the range} & (-\pi, \pi), \\
\theta_2 & \textit{denote the value of } \arg(z-1) & \textit{in the range} & (-\pi, \pi),
\end{array}$$

as in the diagram below.



So we have

$$(z+1)(z-1) = |z+1|\, e^{i\theta_1}\, |z-1|\, e^{i\theta_2}$$

and

$$w = \sqrt{|z+1|\,|z-1|}e^{i(\theta_1+\theta_2)/2}$$

is a holomorphic function on $\mathbb{C}\backslash(-\infty, 1]$ which satisfies

$$w^2 = z^2 - 1.$$

What about the continuity, or otherwise, of w over the cut? Firstly let r be a real number in the range $-1 < r < 1$ and let r_+ and r_- be complex numbers just above and just below r in the complex plane. Then

$$\begin{array}{lll}
\textit{for } r_+ \textit{ we have } \theta_1 & \approx & 0 \textit{ and } \theta_2 \approx \pi; \\
\textit{for } r_+ \textit{ we have } \theta_1 & \approx & 0 \textit{ and } \theta_2 \approx -\pi.
\end{array}$$

So

$$\begin{aligned} w_+ &\approx \sqrt{1-r^2}e^{i(0+\pi)/2} = i\sqrt{1-r^2}; \\ w_- &\approx \sqrt{1-r^2}e^{i(0-\pi)/2} = -i\sqrt{1-r^2}. \end{aligned}$$

So we see that we have a sign discontinuity across $(-1,1)$.

However if we take r *be a real number in the range* $r<-1$ *and let* r_+ *and* r_- *be complex numbers just above and just below* r *in the complex plane. Then*

$$\begin{aligned} \text{for } r_+ \text{ we have } \theta_1 &\approx \pi \text{ and } \theta_2 \approx \pi; \\ \text{for } r_+ \text{ we have } \theta_1 &\approx -\pi \text{ and } \theta_2 \approx -\pi. \end{aligned}$$

So

$$\begin{aligned} w_+ &\approx \sqrt{r^2-1}e^{i(\pi+\pi)/2} = -\sqrt{r^2-1}; \\ w_- &\approx \sqrt{r^2-1}e^{i(-\pi-\pi)/2} = -\sqrt{r^2-1}. \end{aligned}$$

We see that w *is actually continuous across* $(-\infty,-1)$ *and we can in fact extend* w *to a holomorphic function on all of* $\mathbb{C}\backslash[-1,1]$.

Note the behaviour of w *near the points* -1 *and* 1. *If* $z\approx-1$ *then* $w\approx\sqrt{2}i\sqrt{z+1}$ *where* $\sqrt{z+1}$ *is a standard branch of* $\sqrt{z+1}$ *on the cut plane* $\mathbb{C}\backslash[-1,\infty)$. *If* $z\approx 1$ *then* $w\approx\sqrt{2}\sqrt{z-1}$ *where* $\sqrt{z-1}$ *is a standard branch of* $\sqrt{z-1}$ *on the cut plane* $\mathbb{C}\backslash(-\infty,1]$.

Remark 179 *To properly consider the* ***multifunction*** $\sqrt{z^2-1}$ *(or any similar multi-valued function) it helps to consider its* ***Riemann Surface****. In this case the Riemann surface is the set of points*

$$\Sigma=\left\{(z,\zeta)\in\mathbb{C}^2:\zeta^2=z^2-1\right\}.$$

Firstly consider the situation in $\mathbb{R}^2$. *The curve* $y^2=x^2-1$ *is a hyperbola. Above* $(1,\infty)$ *and* $(-\infty,-1)$ *sit branches* $y=\pm\sqrt{x^2-1}$ *and these two branches meet at* $(\pm1,0)$. *So most of the curve is in one or other of the sets*

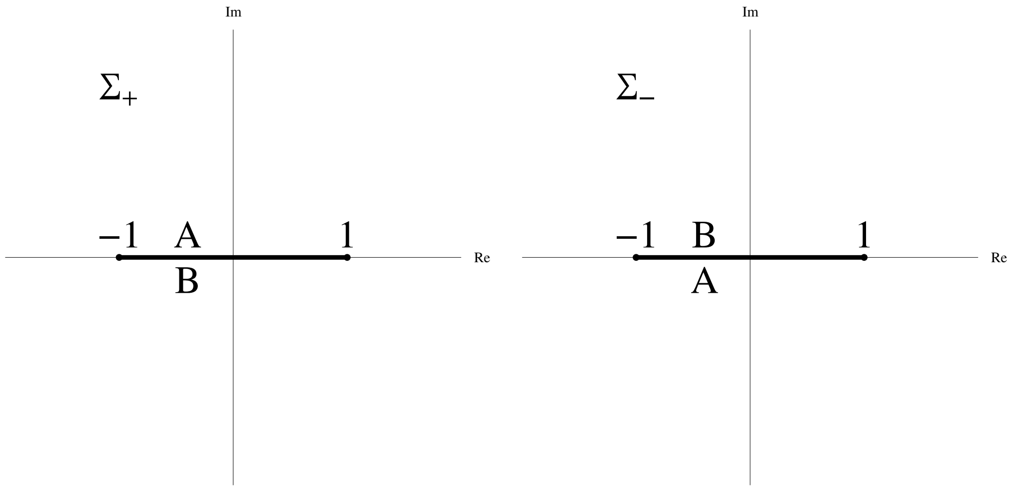
$$C_+=\left\{\left(x,\sqrt{x^2-1}\right):|x|>1\right\};\qquad C_-=\left\{\left(x,-\sqrt{x^2-1}\right):|x|>1\right\}.$$

In fact $C_+\cup C_-$ *excludes only the branch points* $(\pm1,0)$ *and we also see that as we cross the branch points we move from* C_+ *to* C_- *(or vice versa).*

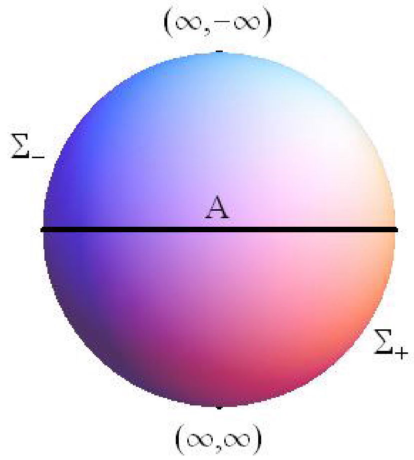
In the complex case, for $z\notin[-1,1]$ *there are two values of* ζ, *namely* $\pm w$. *For* $z=\pm1$ *the only value of* ζ *is* 0. *The points* (z,w) *and* $(z,-w)$ *have already been described as two different branches of* $\sqrt{z^2-1}$ *but we need to take some care to see how these branches fit together as subset of* Σ. *If we set as above*

$$\Sigma_+=\{(z,w):z\notin[-1,1]\}\qquad\text{and}\qquad\Sigma_-=\{(z,-w):z\notin[-1,1]\}.$$

Then $\Sigma_+\cup\Sigma_-$ *is most of* Σ *missing only those points associated with* $z\in[-1,1]$. *We can note, as with previous branches, that as* z *crosses the cut* $[-1,1]$ *then* (z,w) *moves continuously to* *the other branch Σ_- and likewise $(z, -w)$ moves continuously to the other branch Σ_+.*



So Σ_+ and Σ_- fit together on Σ by gluing the either side of $[-1, 1]$ as shown in the diagram above. We can then see that topologically Σ is a cylinder in $\mathbb{C}^2$.



However it often makes sense to include a point at infinity to the complex numbers, something we will meet in more detail later in the course. Likewise it makes sense to introduce some points at infinity to Σ. As z becomes large then the points (z, w) and $(z, -w)$ move to the different ends of the cylinder. So it is fairly natural to include points at infinity at either end of the cylinder to reflect this behaviour. Topologically, with these points included, Σ is a sphere (in what is called complex projective space).

The theory of Riemann surfaces is a particularly rich one involving much algebra, topology and complex analysis with the topology of the surface (primarily) being determined by the degree of the defining polynomial.

Intermission

In the second half of this course we shall begin to see how very different Complex Analysis is from last year's Real Analysis. One immediate difference we have, when working in the complex plane, is that extra dimension and this really makes a difference when it comes to integration. Rather than just integrating over an interval in $\mathbb{R}$ we can now consider **path integrals** in $\mathbb{C}$. It proves (as some of you will have already seen in Multivariable Calculus) natural to focus on simple closed curves (deformed circles) and our first major result is

- **Cauchy's Theorem:** the integral around a simple closed curve, of a function which is holomorphic in and on that curve, is zero.

This theorem has natural links with Green's Theorem, but we will be able to build up a much richer theory because of the complex element of the course.

As we have seen with the Cauchy-Riemann equations, a function $f\colon \mathbb{C} \to \mathbb{C}$ being differentiable in the complex sense is more demanding than being differentiable $\mathbb{R}^2 \to \mathbb{R}^2$. Holomorphic functions, it turns out, are not just once differentiable but infinitely so; in fact, more than this, holomorphic functions are analytic and can be defined locally by a power series. Whilst being holomorphic then is somewhat more restrictive, it is still inclusive enough that most functions that we might be interested in are holomorphic or at least approachable within this theory. But we will also see some rather alien results such as **Liouville's Theorem** which states that bounded holomorphic functions on $\mathbb{C}$ are constant.

Because of Cauchy's Theorem, path integrals around simple closed curves have a robust topological nature. Deforming a path integral to include a region where the integrand is holomorphic does not affect the integral. Consequently the only contributing elements come from the integrand's singularities. We shall see that if a function $f(z)$ has an **isolated singularity** at a point $a \in \mathbb{C}$ (the function is locally holomorphic except at a) then **Laurent's Theorem** applies and we can write $f(z)$ as a doubly infinite series involving positive and negative powers of $z - a$. In fact, we shall see that most of those powers do not contribute to the integral. In fact, if the path only includes the singularity a we have

$$\int f(z)\,\mathrm{d}z = \int \sum_{n=-\infty}^{\infty} c_n (z-a)^n \,\mathrm{d}z = 2\pi i c_{-1}.$$

The coefficient c_{-1} is called the **residue** of f at a.

Given that path integrals are complex numbers you might wonder what physical meaning they might possibly have. But their real or imaginary parts *are* real integrals and we will meet an array of techniques to calculate such integrals and infinite sums such as

$$\int_0^\infty \frac{\cos^2 x}{1+x^2}\,\mathrm{d}x = \frac{\pi}{4}(1+e^{-2}), \qquad \int_0^{2\pi} \frac{\cos^2 x}{2+\sin x}\,\mathrm{d}x = \frac{2\pi}{2+\sqrt{3}}, \qquad \sum_{n=-\infty}^{\infty} \frac{1}{n^2+a^2} = \frac{\pi}{a\tanh(\pi a)}.$$

10. Cauchy's Theorem

Definition 180 *A subset $U \subseteq \mathbb{C}$ is said to be a **domain** if U is non-empty, open and connected. (Note that Priestley refers to such subsets as "regions".)*

Definition 181 *A **path** or **contour** γ is a continuous and piecewise continuously differentiable function $\gamma\colon [a,b] \to \mathbb{C}$.*

(Here "piecewise continuously differentiable" means that γ is continuously differentiable except at finitely many points). We shall also use the same notation γ to denote the image of γ in $\mathbb{C}$ and also refer to that image as a path or contour.

*A path $\gamma\colon [a,b] \to \mathbb{C}$ is said to be **simple** if γ is 1–1 with the possible exception that $\gamma(a)$ may equal $\gamma(b)$ and is said to be **closed** if $\gamma(a) = \gamma(b)$.*

We will need to assume in this course an important theorem. However intuitively obvious the result might be, its proof is subtle and well beyond the course, usually being the subject of algebraic topology courses at fourth year or graduate level.

Theorem 182 ***(Jordan Curve Theorem,*** *Jordan 1887) Let γ be the image of a simple closed path in $\mathbb{C}$. The complement $\mathbb{C}\backslash\gamma$ has exactly two connected components. One of these, known as the **interior**, is bounded and the other, known as the **exterior**, is unbounded.*

Example 183 *Given $a \in \mathbb{C}$ and $r > 0$, we shall denote by*

$$\gamma(a,r)$$

*the circle with centre a and radius r with a **positive orientation** (i.e. oriented anticlockwise) parametrized by*

$$z = a + re^{i\theta} \qquad 0 \leqslant \theta \leqslant 2\pi.$$

We shall denote by $\gamma^+(a,r)$ the upper semicircle of $\gamma(a,r)$ and shall denote by $\gamma^-(a,r)$ the lower semicircle of $\gamma(a,r)$, both positively oriented.

$\gamma(a,r)$ is simple and closed; its interior is $D(a,r)$.

Example 184 *For $w_1, w_2 \in \mathbb{C}$ we shall denote by $[w_1, w_2]$ the line segment in $\mathbb{C}$ from w_1 to w_2 with parametrization*

$$z = w_1 + t(w_2 - w_1) \qquad 0 \leqslant t \leqslant 1.$$

Definition 185 *Given a path $\gamma\colon [a,b] \to \mathbb{C}$ a **reparametrization** of γ is a second map $\Gamma\colon [c,d] \to \mathbb{C}$ such that*

$$\Gamma = \gamma \circ \psi$$

where $\psi\colon [c,d] \to [a,b]$ is a bijection with positive continuous derivative. ψ should be viewed as a change of coordinates taking a point's Γ-coordinate to its γ-coordinate.

Definition 186 *Let $\gamma\colon [a,b] \to \mathbb{C}$ be a path. The length $\mathcal{L}(\gamma)$ of γ is defined to be*

$$\mathcal{L}(\gamma) = \int_a^b |\gamma'(t)| \, \mathrm{d}t.$$

Proposition 187 *The length of a path γ is invariant under reparametrization.*

Proof Let $\gamma_1\colon\ [a_1,b_1]\to\mathbb{C}$ and $\gamma_2\colon\ [a_2,b_2]\to\mathbb{C}$ be two parametrizations of the same curve γ. Let $\alpha\colon\ [a_1,b_1]\to[a_2,b_2]$ be the change of coordinate map so that $\gamma_2(\alpha(t))=\gamma_1(t)$.

As $\alpha(a_1)=a_2$ and $\alpha(b_1)=b_2$ then

$$\begin{aligned}\int_{a_1}^{b_1}|\gamma_1'(t)|\ \mathrm{d}t &= \int_{a_1}^{b_1}|\gamma_2'(\alpha(t))|\ |\alpha'(t)|\ \mathrm{d}t\\ &= \int_{a_1}^{b_1}|\gamma_2'(\alpha(t))|\ \alpha'(t)\,\mathrm{d}t=\int_{a_2}^{b_2}|\gamma_2'(u)|\ \mathrm{d}u.\end{aligned}$$

■

Definition 188 *Let $f\colon U\to\mathbb{C}$ be a continuous function defined on a domain U and let $\gamma\colon\ [a,b]\to U$ be a path in U. We define the* ***path integral***

$$\int_\gamma f(z)\,\mathrm{d}z=\int_a^b f(\gamma(t))\,\gamma'(t)\,\mathrm{d}t.$$

Remark 189 *Note that we are integrating here a complex function over a real interval $[a,b]$ but the definition of the such integrals is very much what you would expect. Given continuous real functions $x(t)$ and $y(t)$ on $[a,b]$ we defined*

$$\int_a^b(x(t)+iy(t))\ \mathrm{d}t=\int_a^b x(t)\,\mathrm{d}t+i\int_a^b y(t)\,\mathrm{d}t.$$

It is an easy check that integration remains linear: that is, for complex scalars λ,μ and complex integrands $w(t)$ and $z(t)$ we have

$$\int_a^b(\lambda w(t)+\mu z(t))\ \mathrm{d}t=\lambda\int_a^b w(t)\,\mathrm{d}t+\mu\int_a^b z(t)\,\mathrm{d}t.$$

Remark 190 *Given two paths $\gamma_1\colon\ [a_1,b_1]\to\mathbb{C}$ and $\gamma_2\colon\ [a_2,b_2]\to\mathbb{C}$ with $\gamma_1(b_1)=\gamma_2(a_2)$ we define their* ***join*** *$\gamma_1\cup\gamma_2\colon\ [a_1,b_1+b_2-a_2]\to\mathbb{C}$ by*

$$\gamma_1\cup\gamma_2(t)=\begin{cases}\gamma_1(t) & a_1\leqslant t\leqslant b_1\\ \gamma_2(t-b_1+a_2) & b_1\leqslant t\leqslant b_1+b_2-a_2\end{cases}$$

In case it is not clear then $\gamma_1\cup\gamma_2$ follows traces out γ_1 and then traces out γ_2. Unsurprisingly for any continuous f defined on both paths

$$\int_{\gamma_1\cup\gamma_2}f(z)\,\mathrm{d}z=\int_{\gamma_1}f(z)\,\mathrm{d}z+\int_{\gamma_2}f(z)\,\mathrm{d}z.$$

In a similar fashion for a path $\gamma\colon\ [a,b]\to\mathbb{C}$ we can also define the reverse path $-\gamma\colon\ [a,b]\to\mathbb{C}$ by

$$-\gamma(t)=\gamma(b+a-t).$$

Unsurprisingly

$$\int_{-\gamma}f(z)\,\mathrm{d}z=-\int_\gamma f(z)\,\mathrm{d}z.$$

Proposition 191 *The path integral $\int_\gamma f(z)\,\mathrm{d}z$ is invariant under reparametrization.*

Proof Let $\gamma_i\colon\ [a_i,b_i]\to\mathbb{C}$ for $i=1,2$ be two parametrizations of γ and let $\alpha\colon\ [a_1,b_1]\to[a_2,b_2]$ be the change of coordinate map so that $\gamma_2(\alpha(t))=\gamma_1(t)$. Then

$$\int_{a_2}^{b_2} f(\gamma_2(t))\ \gamma_2'(t)\,\mathrm{d}t=\int_{a_1}^{b_1} f(\gamma_2(\alpha(t))\ \gamma_2'(\alpha(t))\,\alpha'(t)\,\mathrm{d}t=\int_{a_1}^{b_1} f(\gamma_1(t))\ \gamma_1'(t)\,\mathrm{d}t$$

by the chain rule. ∎

Example 192 *Rewrite the integral*

$$\int_0^{2\pi}\frac{\mathrm{d}\theta}{(3+\cos\theta)^2}$$

as a path integral around $\gamma(0,1)$.

Solution If we set $z=e^{i\theta}$ then z wraps positively around $\gamma(0,1)$ and we also have

$$\cos\theta=\frac{e^{i\theta}+e^{-i\theta}}{2}=\frac{z+z^{-1}}{2};\qquad \mathrm{d}z=ie^{i\theta}\,\mathrm{d}\theta=iz\,\mathrm{d}\theta.$$

So

$$\int_0^{2\pi}\frac{\mathrm{d}\theta}{(3+\cos\theta)^2}=\int_{\gamma(0,1)}\frac{\mathrm{d}z/(iz)}{(3+\frac{z+z^{-1}}{2})^2}=\frac{4}{i}\int_{\gamma(0,1)}\frac{z\,\mathrm{d}z}{(z^2+6z+1)^2}.$$

∎

Example 193 *Evaluate the following path integrals:*

$$\int_{\gamma^+(0,1)}\frac{\mathrm{d}z}{z}.\qquad \int_{[a,b]}z^2\,\mathrm{d}z,\qquad \int_{[1,-i]}\frac{\mathrm{d}z}{z}.$$

Solution Writing $z=e^{i\theta}$ where $0\leqslant\theta\leqslant\pi$ we see that

$$\int_{\gamma^+(0,1)}\frac{\mathrm{d}z}{z}=\int_0^\pi\frac{ie^{i\theta}\,\mathrm{d}\theta}{e^{i\theta}}=i\int_0^\pi\mathrm{d}\theta=\pi i.$$

Parametrizing $[a,b]$ by $z=a+t(b-a)$ with $0\leqslant t\leqslant 1$, we see

$$\begin{aligned}\int_{[a,b]}z^2\,\mathrm{d}z &= \int_0^1 (a+t(b-a))^2\ \mathrm{d}(a+t(b-a))\\ &= (b-a)\int_0^1 (a+t(b-a))^2\ \mathrm{d}t\\ &= (b-a)\left[a^2+2a(b-a)\int_0^1 t\,\mathrm{d}t+(b-a)^2\int_0^1 t^2\,\mathrm{d}t\right]\\ &= \frac{1}{3}(b-a)\left[a^2+ab+b^2\right]=\frac{b^3}{3}-\frac{a^3}{3}.\end{aligned}$$

An oddly familiar answer!!

Finally we will parametrize $[1,-i]$ by $z = 1-(1+i)t$ where $0 \leqslant t \leqslant 1$ and find

$$\begin{aligned}
\int_{[1,-i]} \frac{\mathrm{d}z}{z} &= \int_0^1 \frac{-(1+i)\,\mathrm{d}t}{1-(1+i)t} \\
&= -(1+i)\int_0^1 \frac{[(1-t)+it]\ \mathrm{d}t}{(1-t)^2+t^2} \\
&= \frac{-(1+i)}{2}\int_0^1 \frac{[(1-t)+it]\ \mathrm{d}t}{(t-1/2)^2+1/4} \\
&= \frac{-(1+i)}{2}\int_{-\pi/4}^{\pi/4} \frac{\left[\left(\frac{1}{2}-\frac{1}{2}\tan\theta\right)+i\left(\frac{1}{2}+\frac{1}{2}\tan\theta\right)\right]\ \left(\frac{1}{2}\sec^2\theta\right)\mathrm{d}\theta}{(1/4)\sec^2\theta}
\end{aligned}$$

where the last integral is arrived at by substituting $t = (1+\tan\theta)/2$. With some simplifying the above becomes

$$\begin{aligned}
& \frac{-(1+i)}{2}\int_{-\pi/4}^{\pi/4} [(1+i)+(i-1)\tan\theta]\ \mathrm{d}\theta \\
= & \frac{-(1+i)}{2}\int_{-\pi/4}^{\pi/4} (1+i)\ \mathrm{d}\theta \qquad [\text{as } \tan\theta \text{ is odd}] \\
= & \frac{-(1+i)}{2}\times(1+i)\times\frac{\pi}{2} = \frac{-\pi i}{2}.
\end{aligned}$$

■

Proposition 194 *For $a \in \mathbb{C}$, $r > 0$ and $k \in \mathbb{Z}$,*

$$\int_{\gamma(a,r)} (z-a)^k\ \mathrm{d}z = \begin{cases} 0 & k \neq -1, \\ 2\pi i & k = -1. \end{cases}$$

Proof We set $z = a + re^{i\theta}$ where $0 \leqslant \theta \leqslant 2\pi$ and then

$$\begin{aligned}
\int_{\gamma(a,r)} (z-a)^k\ \mathrm{d}z &= \int_0^{2\pi} \left(re^{i\theta}\right)^k \left(ire^{i\theta}\,\mathrm{d}\theta\right) \\
&= ir^{k+1}\int_0^{2\pi} e^{i(k+1)\theta}\,\mathrm{d}\theta \\
&= ir^{k+1}\int_0^{2\pi} [\cos(k+1)\theta + i\sin(k+1)\theta]\ \mathrm{d}\theta.
\end{aligned}$$

This last integral is clearly zero when $k \neq -1$. When $k = -1$ we have

$$\int_{\gamma(a,r)} (z-a)^{-1}\ \mathrm{d}z = ir^0\int_0^{2\pi} \mathrm{d}\theta = 2\pi i.$$

■

Remark 195 *Whilst the above calculation may look rather trifling, it will prove to be central to the theory of residues and contour integration. In fact, Preistley refers to it as the* ***Fundamental Integral*** *(though this is a phrase of her coining and not widely used). It is at least the reason why you will see the number* $2\pi i$ *appearing in a great many theorems. We shall see in due course that a function* $f(z)$ *which is holomorphic on* $D'(a,R)$ *can be written in the form*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n$$

for unique c_n*. This is Laurent's Theorem. If* $0 < r < R$ *and* ***providing convergence allows it*** *then we can see that*

$$\int_{\gamma(a,r)} f(z)\,\mathrm{d}z = \int_{\gamma(a,r)} \left(\sum_{n=-\infty}^{\infty} c_n (z-a)^n \right) \mathrm{d}z = \sum_{n=-\infty}^{\infty} c_n \left(\int_{\gamma(a,r)} (z-a)^n\,\mathrm{d}z \right) = 2\pi i c_{-1}.$$

So the function's integral depends only on the coefficient c_{-1}*, known as the* ***residue*** *of* f *at* a*. Of course, we will need to carefully demonstrate that the sum and integral can indeed be interchanged.*

Proposition 196 ***(Fundamental Theorem of Calculus for path integrals)*** *Let* f *be a holomorphic function on a domain* U*, and let* γ *be a path in* U *from* p *to* q*. Then*

$$\int_\gamma f'(z)\,\mathrm{d}z = f(q) - f(p).$$

Proof Let $\gamma\colon [a,b] \to U$ be a parametrization of γ with $\gamma(a) = p$ and $\gamma(b) = q$. Then

$$\begin{aligned} \int_\gamma f'(z)\,\mathrm{d}z &= \int_a^b f'(\gamma(t))\gamma'(t)\,\mathrm{d}t \\ &= \int_a^b \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\,\mathrm{d}t \\ &= [f(\gamma(t))]_{t=a}^{t=b} \\ &= f(\gamma(b)) - f(\gamma(a)) \\ &= f(q) - f(p) \end{aligned}$$

as required. (The application of the usual Fundamental Theorem of Calculus relies on derivatives to be continuous, but we will see with Corollary 225 that holomorphic functions have continuous derivatives.) ∎

Corollary 197 *If* $f' = 0$ *on a domain* U *then* f *is constant.*

Example 198 *If we return to the three integrals in Example 193, we see that each of them can be determined simply using the above version of the Fundamental Theorem.*

- If we take a branch of logarithm on $\mathbb{C}\backslash\{\text{negative imaginary axis}\}$ defined by

$$\log z = \log|z| + i\arg z \qquad -\pi/2 < \arg z < 3\pi/2$$

then this branch of $\log z$ is defined and holomorphic on all of $\gamma^+(0,1)$ and we see

$$\int_{\gamma^+(0,1)} \frac{\mathrm{d}z}{z} = \log(-1) - \log 1 = \pi i - 0 = \pi i.$$

- We can define $f(z) = z^3/3$ on all of $\mathbb{C}$ (and so in particular on $[a,b]$) with $f'(z) = z^2$ and so

$$\int_{[a,b]} z^2\,\mathrm{d}z = \left[\frac{z^3}{3}\right]_a^b = \frac{b^3}{3} - \frac{a^3}{3}.$$

- If we take a branch of logarithm on $\mathbb{C}\backslash(-\infty,0]$ defined by

$$\log z = \log|z| + i\arg z \qquad -\pi < \arg z < \pi,$$

then this branch of $\log z$ is defined and holomorphic on all of $[1,-i]$ and we see

$$\int_{[1,-i]} \frac{\mathrm{d}z}{z} = \log(-i) - \log 1 = -\pi i/2 = 0 = -\pi i/2.$$

Proposition 199 *(**Estimation Theorem**) Let U be a domain, $\gamma\colon [a,b] \to U$ a path in U and $f\colon U \to \mathbb{C}$. Then*

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant \int_a^b |f(\gamma(t))\gamma'(t)|\,\mathrm{d}t.$$

In particular if $|f(z)| \leqslant M$ on γ then

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant M\,\mathcal{L}(\gamma).$$

Proof The proof that

$$\left|\int_a^b z(t)\,\mathrm{d}t\right| \leqslant \int_a^b |z(t)|\,\mathrm{d}t$$

for a complex integrand follows in a similar fashion to that for real integrands (being essentially a continuous version of the triangle inequality). We then have

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| = \left|\int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}t\right| \leqslant \int_a^b |f(\gamma(t))\gamma'(t)|\,\mathrm{d}t.$$

Further if $|f(z)| \leqslant M$ on γ then

$$\left|\int_\gamma f(z)\,\mathrm{d}z\right| \leqslant \int_a^b M\,|\gamma'(t)|\,\mathrm{d}t = M\,\mathcal{L}(\gamma).$$

■

Proposition 200 *(**Uniform Convergence**) Suppose that the functions $f_n(z)$ uniformly converge to $f(z)$ on the path γ. Then*

$$\lim_{n\to\infty}\int_\gamma f_n(z)\,\mathrm{d}z = \int_\gamma f(z)\,\mathrm{d}z.$$

Proof Let $\varepsilon > 0$. There exists N such that

$$|f_n(z) - f(z)| < \frac{\varepsilon}{\mathcal{L}(\gamma)} \qquad \text{for } n \geqslant N \text{ and } z \in \gamma.$$

Hence for $n \geqslant N$ we have

$$\left|\left(\int_\gamma f_n(z)\,\mathrm{d}z\right) - \left(\int_\gamma f(z)\,\mathrm{d}z\right)\right| = \left|\int_\gamma (f_n(z) - f(z))\,\mathrm{d}z\right| < \mathcal{L}(\gamma) \times \frac{\varepsilon}{\mathcal{L}(\gamma)} = \varepsilon$$

and the result follows. ■

Example 201 *Show that*

$$\int_{\gamma^+(0,R)} \frac{\mathrm{d}z}{z^2+1} \to 0 \qquad \textit{as } R \to \infty.$$

Solution On $\gamma^+(0,R)$ we have

$$\left|z^2+1\right| = \left|z^2 - (-1)\right| \geqslant \left|z^2\right| - |-1| = R^2 - 1$$

using $|z-w| \geqslant |z| - |w|$. Hence

$$\left|\int_{\gamma^+(0,R)} \frac{\mathrm{d}z}{z^2+1}\right| \leqslant \mathcal{L}(\gamma^+(0,R)) \times \frac{1}{R^2-1} = \frac{\pi R}{R^2-1} \to 0 \qquad \text{as } R \to \infty.$$

■

We now move to address the major theorem of the course: **Cauchy's Theorem** (or perhaps more properly the **Cauchy-Goursat Theorem**). This states:

Theorem 202 *(**Cauchy 1825, Goursat 1900**) Suppose that $f(z)$ is holomorphic inside and on a closed path γ. Then*

$$\int_\gamma f(z)\,\mathrm{d}z = 0.$$

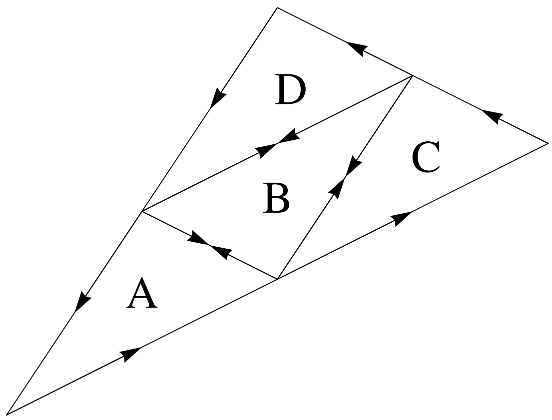
Remark 203 *Whilst Cauchy's Theorem is one of the most important in analysis, it may well also not be that surprising at this stage in your studies. Those who has met Green's and Stokes' Theorems have seen similar results before, and a weaker version of Cauchy's Theorem does indeed follow from Green's Theorem (see Exercise Sheet 5, Question 1). Further, in those cases when f is the derivative of some function F then Cauchy's Theorem just follows from the Fundamental Theorem of Calculus for a closed curve. In fact we shall in due course see that f is indeed the derivative of such F – of course we will have needed to make much use of Cauchy's Theorem to demonstrate the existence of such F.*

It is beyond the scope of this course to give a full rigorous proof of this result. We shall first prove this result for a triangle.

Theorem 204 ***(Cauchy's Theorem for a triangle)*** *Let f be holomorphic on a domain which includes a closed triangular region T. Let Δ denote the boundary of the triangle, positively oriented. Then*

$$\int_{\Delta} f(z)\,\mathrm{d}z = 0.$$

Proof STEP ONE: *Creating nested triangles and their intersection.* If we join the three mid-points of the sides of Δ to make four new triangular paths A, B, C, D, and orient them as in the diagram below,



then we see that

$$\int_{\Delta} f(z)\,\mathrm{d}z = \int_{A} f(z)\,\mathrm{d}z + \int_{B} f(z)\,\mathrm{d}z + \int_{C} f(z)\,\mathrm{d}z + \int_{D} f(z)\,\mathrm{d}z$$

as the contributions to the RHS from the interior edges cancel out. Of the four summands on the RHS, one has maximal modulus, call it Δ_1, and then by the triangle inequality

$$\left|\int_{\Delta} f(z)\,\mathrm{d}z\right| \leqslant 4\left|\int_{\Delta_1} f(z)\,\mathrm{d}z\right|.$$

We can then apply this repeatedly, producing triangle $\Delta_1, \Delta_2, \ldots$ such that

$$\left|\int_{\Delta} f(z)\,\mathrm{d}z\right| \leqslant 4^n\left|\int_{\Delta_n} f(z)\,\mathrm{d}z\right| \qquad \text{for any } n.$$

As the closed triangular regions (boundaries and interiors) are nested compact subsets then their intersection

$$\bigcap_{n=1}^{\infty} \Delta_n$$

is non-empty (this is left as a lemma at the end of the theorem), but can contain no more than one point as the triangles lie in discs whose radii tend to 0. Let ζ denote the single complex number in the intersection.

STEP TWO: *Applying the differentiability of f at ζ.* As $f'(\zeta)$ exists, then for any $\varepsilon > 0$ there exists $\delta > 0$ such that

$$\left|\frac{f(z)-f(\zeta)}{z-\zeta} - f'(\zeta)\right| < \varepsilon \quad \text{whenever} \quad |z-\zeta| < \delta$$

or equivalently that

$$f(z) = f(\zeta) + (z-\zeta)\, f'(\zeta) + (z-\zeta)w \quad \text{where} \quad |w| < \varepsilon.$$

If we choose n sufficiently large that $\Delta_n \subseteq D(\zeta, \delta)$ then we see

$$\int_{\Delta_n} f(z)\,\mathrm{d}z = f(\zeta)\int_{\Delta_n} \mathrm{d}z + f'(\zeta)\int_{\Delta_n} (z-\zeta)\,\mathrm{d}z + \int_{\Delta_n} (z-\zeta)w(z)\,\mathrm{d}z = \int_{\Delta_n} (z-\zeta)w(z)\,\mathrm{d}z$$

as the first two integrals are zero by the Fundamental Theorem of Calculus. Note, from the nature of the triangles' construction, that$\mathcal{L}(\Delta_n) = \frac{1}{2}\mathcal{L}(\Delta_{n-1}) = \frac{1}{2^n}\mathcal{L}(\Delta)$ and that for any $z \in \Delta_n$ we have $|z-\zeta| < \mathcal{L}(\Delta_n)$. So, by the Estimation Theorem,

$$\left|\int_{\Delta_n} f(z)\,\mathrm{d}z\right| = \left|\int_{\Delta_n} (z-\zeta)w(z)\,\mathrm{d}z\right| \leqslant \mathcal{L}(\Delta_n) \times \mathcal{L}(\Delta_n) \times \varepsilon = \frac{\varepsilon}{4^n}\mathcal{L}(\Delta)^2.$$

Finally, then,

$$\left|\int_{\Delta} f(z)\,\mathrm{d}z\right| \leqslant 4^n \left|\int_{\Delta_n} f(z)\,\mathrm{d}z\right| \leqslant \varepsilon\mathcal{L}(\Delta)^2$$

and as ε was arbitrarily chosen then it follows that

$$\int_{\Delta} f(z)\,\mathrm{d}z = 0.$$

■

Lemma 205 *Let M be a compact metric space and C_n a decreasing sequence of closed non-empty subsets. Then $\bigcap C_n \neq \varnothing$.*

Proof Let $U_n = M \backslash C_n$ so that the U_n form an increasing sequence of open subsets. Suppose that the intersection $\bigcap C_n$ is empty and then, by De Morgan's Law

$$\bigcup U_n = \bigcup (M\backslash C_n) = M \Big\backslash \bigcap C_n = M,$$

shows that the U_n form an open cover for M. By compactness there are $n_1 < n_2 < \cdots < n_k$ such that

$$U_{n_k} = U_{n_1} \cup \cdots \cup U_{n_k} = M$$

but this leads to the conclusion that $C_{n_k} = \varnothing$, a contradiction. ■

Recall that a domain U is called **convex** if for any $a, b \in U$ it follows that $[a,b] \subseteq U$. We now have:

Theorem 206 ***(Antiderivative Theorem)*** *Let f be a function which is holomorphic on a convex domain U. Then there exists a holomorphic function F on U such that $F'(z) = f(z)$.*

Proof Fix $a \in U$ and for any $z \in U$ let $[a, z]$ denote the oriented line segment from a to z. Define

$$F(z) = \int_{[a,z]} f(w)\,\mathrm{d}w.$$

Let $\varepsilon > 0$ be such that $D(z, \varepsilon) \subseteq U$ and take h with $|h| < \varepsilon$. By Cauchy's Theorem for triangles

$$\int_{[a,z]} f(w)\,\mathrm{d}w + \int_{[z,z+h]} f(w)\,\mathrm{d}w + \int_{[z+h,a]} f(w)\,\mathrm{d}w = 0$$

which may be rearranged to

$$F(z+h) - F(z) = \int_{[z,z+h]} f(w)\,\mathrm{d}w.$$

So

$$\left|\frac{F(z+h) - F(z)}{h} - f(z)\right| = \left|\left(\frac{1}{h}\int_{[z,z+h]} f(w)\,\mathrm{d}w\right) - f(z)\right| = \left|\frac{1}{h}\int_{[z,z+h]} [f(w) - f(z)]\,\mathrm{d}w\right|$$

as $\int_{[z,z+h]} k\,\mathrm{d}w = kh$ for any constant function of w. Finally, by the Estimation Theorem, we have

$$\begin{aligned}\left|\frac{F(z+h) - F(z)}{h} - f(z)\right| &= \left|\frac{1}{h}\int_{[z,z+h]} [f(w) - f(z)]\,\mathrm{d}w\right| \\ &\leqslant \frac{1}{|h|} \times |h| \times \sup_{w\in[z,z+h]} |f(w) - f(z)| \\ &= \sup_{w\in[z,z+h]} |f(w) - f(z)|\end{aligned}$$

which tends to zero as $h \to 0$ by the continuity of f at z, showing that $F'(z) = f(z)$ as required. ■

Corollary 207 ***(Cauchy's Theorem for a convex region)*** *Let f be holomorphic on a convex domain U. Then for any closed path γ in U we have*

$$\int_\gamma f(z)\,\mathrm{d}z = 0.$$

Proof This follows from the Antiderivative Theorem and the Fundamental Theorem of Calculus. ■

We will now proceed to state, though not prove, two more general versions of Cauchy's Theorem. For these we will need to introduce some further topological notions.

Definition 208 *Let $\gamma_1 \colon [0,1] \to U$ and $\gamma_2 \colon [0,1] \to U$ be two closed paths in a domain U. We say that γ_1 and γ_2 are* ***homotopic*** *if there is a continuous function $H \colon [0,1]^2 \to \mathbb{C}$ such that for each $u \in [0,1]$, then $H(\_, u) \colon [0,1] \to U$ is a closed path in U and for all $t \in [0,1]$,*

$$H(t,0) = \gamma_1(t), \qquad H(t,1) = \gamma_2(t).$$

We might think of the second variable u as a time parameter. The homotopy H can then be viewed as a continuous deformation from γ_1 (when $u = 0$) to γ_2 (when $u = 1$). Whilst the full rigour of the following result is somewhat difficult, it should not be a surprising result. Indeed $\gamma_1 \cup (-\gamma_2)$ might be viewed as the boundary to a domain U on which f is holomorphic and thus, by an appropriate version of Cauchy's Theorem, we might expect

$$0 = \int_{\gamma_1 \cup (-\gamma_2)} f(z)\,\mathrm{d}z = \int_{\gamma_1} f(z)\,\mathrm{d}z - \int_{\gamma_2} f(z)\,\mathrm{d}z.$$

Theorem 209 ***(Deformation Theorem)*** *Let f be holomorphic on a domain U and let γ_1 and γ_2 be homotopic closed paths in U. Then*

$$\int_{\gamma_1} f(z)\,\mathrm{d}z = \int_{\gamma_2} f(z)\,\mathrm{d}z.$$

Example 210 *Any $\gamma(a_1, r_1)$ and $\gamma(a_2, r_2)$ are homotopic in $\mathbb{C}$.*

Solution We may use

$$H(\theta, u) = ua_2 + (1-u)\,a_1 + (ur_2 + (1-u)\,r_1)\,e^{2\pi i\theta}.$$

■

Example 211 *$\gamma(0,1)$ and $\gamma(2,1)$ are not homotopic in $\mathbb{C}\backslash\{0\}$.*

Solution Let $f(z) = z^{-1}$ which is homotopic on $\mathbb{C}\backslash\{0\}$. Note that

$$\int_{\gamma(0,1)} \frac{\mathrm{d}z}{z} = 2\pi i \qquad \text{and} \qquad \int_{\gamma(2,1)} \frac{\mathrm{d}z}{z} = 0,$$

(the first by direct calculation, the second by Cauchy's Theorem). Hence, by the Deformation Theorem, $\gamma(0,1)$ and $\gamma(2,1)$ cannot be homotopic in $\mathbb{C}\backslash\{0\}$. ■

Definition 212 *We say that a domain U is* ***simply-connected*** *if every closed curve in U is homotopic to a point (i.e. to a constant path).*

Example 213 $\mathbb{C}$, *(open and closed) discs, (open and closed) half-planes, line segments and cut-planes are all simply-connected domains.*

Punctured discs, circles and $\mathbb{C}\backslash\{0\}$ are domains which are not simply-connected.

The following are then corollaries to the Deformation Theorem.

Corollary 214 *(**Cauchy's Theorem on a Simply Connected Domain**) Let f be holomorphic on a simply connected domain U and let γ be a closed path in U. Then*

$$\int_\gamma f(z)\,\mathrm{d}z = 0.$$

Proof γ is homotopic to a constant path Γ and so by the Deformation Theorem

$$\int_\gamma f(z)\,\mathrm{d}z = \int_\Gamma f(z)\,\mathrm{d}z = 0.$$

■

Corollary 215 *Let f be holomorphic on a simply-connected domain U. Let $a, b \in U$, and let γ_1 and γ_2 be two paths from a to b. Then*

$$\int_{\gamma_1} f(z)\,\mathrm{d}z = \int_{\gamma_2} f(z)\,\mathrm{d}z.$$

Proof Apply the previous corollary to $\gamma_1 \cup (-\gamma_2)$. ■

Corollary 216 *(**Antiderivative Theorem for a simply-connected domain**) Let f be holomorphic on a simply-connected domain U. Then there exists F on U such that $F' = f$.*

Proof Fix $a \in U$ and for $z \in U$ let γ be a path from a to z in U. Define

$$F(z) = \int_\gamma f(w)\,\mathrm{d}w.$$

By the previous corollary F is independent of the choice of path and so a function of z alone. By an argument identical to that for the Antiderivative Theorem in a convex domain, it follows that $F'(z) = f(z)$. ■

Corollary 217 *(**Logarithm for a simply connected domain**) Let U be a simply-connected domain not containing zero. Then there exists a holomorphic function $l(z)$ on U such that*

$$\exp(l(z)) = z.$$

Proof Let $a \in U$ and $c \in [\log a]$ so that $\exp c = a$. For $z \in U$, define

$$l(z) = c + \int_\gamma \frac{\mathrm{d}w}{w}$$

where γ is a path from a to z. By the previous corollary $l'(z) = z^{-1}$. Define

$$G(z) = \frac{\exp(l(z))}{z}.$$

Then

$$G'(z) = \frac{zl'(z)\exp(l(z)) - \exp(l(z))}{z^2} = \frac{\exp(l(z)) - \exp(l(z))}{z^2} = 0$$

and so G is constant on U by connectedness. At a we see

$$G(a) = \frac{\exp(l(a))}{a} = \frac{\exp c}{a} = \frac{a}{a} = 1$$

and hence $\exp(l(z)) = z$ as required. ■

11. Consequences of Cauchy's Theorem

Let $\gamma\colon [0,1] \to \mathbb{C}$ be a simple path in $\mathbb{C}$ so that, in particular, γ is a homeomorphism onto its image. If $\Gamma\colon [0,1] \to \mathbb{C}$ is also a homeomorphism onto the image of γ then $\gamma \circ \Gamma^{-1}$ is a homeomorphism of $[0,1]$. This means $\gamma \circ \Gamma^{-1}$ is strictly monotone and so either increasing or decreasing.

Definition 218 *We say that γ and Γ have the same* ***orientation*** *if $\gamma \circ \Gamma^{-1}$ is increasing and have* ***reverse orientations*** *if $\gamma \circ \Gamma^{-1}$ is decreasing.*

If γ and Γ have the same orientation then

$$\int_\gamma f(z)\,\mathrm{d}z = \int_\Gamma f(z)\,\mathrm{d}z$$

and if they have reverse orientations then

$$\int_\gamma f(z)\,\mathrm{d}z = -\int_\Gamma f(z)\,\mathrm{d}z.$$

Now let γ be a simple closed path in $\mathbb{C}$. By the Jordan Curve Theorem $\mathbb{C}\backslash\gamma$ has precisely two connected components, one of which, the interior, is bounded. If a is in the interior of γ, and is such that the circle $\gamma(a,r)$ is interior to γ also, then γ is homotopic to $\gamma(a,r)$ or to $-\gamma(a,r)$. If γ is homotopic to $\gamma(a,r)$ then by the Deformation Theorem

$$\int_\gamma \frac{\mathrm{d}z}{z-a} = \int_{\gamma(a,r)} \frac{\mathrm{d}z}{z-a} = 2\pi i,$$

and otherwise the integral is $-2\pi i$.

Definition 219 *Let γ be a simple closed curve and a a point in the interior of γ. We say that γ is* ***positively oriented*** *if*

$$\int_\gamma \frac{\mathrm{d}z}{z-a} = 2\pi i.$$

Remark 220 *More generally if γ is a (not necessarily simple) closed curve, and a is a point not on γ, then*

$$\frac{1}{2\pi i}\int_\gamma \frac{\mathrm{d}z}{z-a}$$

is an integer, called the ***winding number of*** *γ* ***about*** *a.*

Theorem 221 ***(Cauchy's Integral Formula)*** *Let f be holomorphic on and inside a positively oriented, simple, closed curve γ and let a be a point inside γ. Then*

$$\frac{1}{2\pi i}\int_\gamma \frac{f(w)}{w-a}\,\mathrm{d}w = f(a).$$

Proof By the Deformation Theorem we know that

$$\frac{1}{2\pi i}\int_{\gamma}\frac{f(w)}{w-a}\,\mathrm{d}w=\frac{1}{2\pi i}\int_{\gamma(a,r)}\frac{f(w)}{w-a}\,\mathrm{d}w$$

for *any* circle $\gamma(a,r)$ inside γ. From Proposition 194, we know that

$$\int_{\gamma(a,r)}\frac{\mathrm{d}w}{w-a}=2\pi i$$

so that

$$f(a)=\frac{1}{2\pi i}\int_{\gamma(a,r)}\frac{f(a)\,\mathrm{d}w}{w-a}.$$

Hence

$$\begin{aligned}\left|\left(\frac{1}{2\pi i}\int_{\gamma(a,r)}\frac{f(w)}{w-a}\,\mathrm{d}w\right)-f(a)\right| &= \left|\frac{1}{2\pi i}\int_{\gamma(a,r)}\frac{f(w)-f(a)}{w-a}\,\mathrm{d}w\right|\\ &= \frac{1}{2\pi}\left|\int_0^{2\pi}\frac{f(a+re^{i\theta})-f(a)}{re^{i\theta}}ire^{i\theta}\mathrm{d}\theta\right|\\ &= \frac{1}{2\pi}\left|\int_0^{2\pi}\left(f(a+re^{i\theta})-f(a)\right)\,\mathrm{d}\theta\right|\\ &\leqslant \frac{1}{2\pi}\times 2\pi\sup_{0\leqslant\theta\leqslant 2\pi}\left|f(a+re^{i\theta})-f(a)\right|\\ &= \sup_{0\leqslant\theta\leqslant 2\pi}\left|f(a+re^{i\theta})-f(a)\right|\to 0\text{ as }r\to 0\end{aligned}$$

by the continuity of f at a. The result follows. ■

Example 222 *Evaluate, using Cauchy's Integral Formula only,*

$$\int_{\gamma(0,1)}\frac{\mathrm{d}z}{z^2-4z+1},$$

where $\gamma(0,1)$ *denotes the positively oriented unit circle centred at 0. By means of the substitution* $z=e^{i\theta}$*, or otherwise, prove that*

$$\int_0^{2\pi}\frac{\mathrm{d}\theta}{2-\cos\theta}=\frac{2\pi}{\sqrt{3}}.$$

Solution Note that $z^2-4z+1=0$ only at α and β where

$$\alpha=2+\sqrt{3}\quad\text{and}\quad\beta=2-\sqrt{3}.$$

As $|\alpha|>1$ then $(z-\alpha)^{-1}$ is holomorphic in the disc $D(0,1)$ and so by the Integral Formula

$$\int_{\gamma(0,1)}\frac{\mathrm{d}z}{z^2-4z+1}=\int_{\gamma(0,1)}\frac{(z-\alpha)^{-1}\,\mathrm{d}z}{z-\beta}=\frac{2\pi i}{\beta-\alpha}=\frac{-\pi i}{\sqrt{3}}.$$

If we set $z = e^{i\theta}$ in the integral we see

$$\begin{aligned}\int_{\gamma(0,1)} \frac{\mathrm{d}z}{z^2 - 4z + 1} &= \int_0^{2\pi} \frac{ie^{i\theta}\,\mathrm{d}\theta}{e^{2i\theta} - 4e^{i\theta} + 1} \\ &= i\int_0^{2\pi} \frac{\mathrm{d}\theta}{e^{i\theta} + e^{-i\theta} - 4} \\ &= \frac{-i}{2}\int_0^{2\pi} \frac{\mathrm{d}\theta}{2 - \cos\theta}\end{aligned}$$

and hence

$$\int_0^{2\pi} \frac{\mathrm{d}\theta}{2 - \cos\theta} = \frac{2}{-i} \times \frac{-\pi i}{\sqrt{3}} = \frac{2\pi}{\sqrt{3}}.$$

■

Example 223 *Use Cauchy's Integral Formula to determine*

$$\int_{\gamma(0,1)} \frac{\operatorname{Im} z}{2z - 1}\,\mathrm{d}z \qquad \textit{and} \qquad \int_{\gamma(0,1)} \frac{\exp z}{4z^2 + 1}\,\mathrm{d}z.$$

Solution The function $\operatorname{Im} z$ is not holomorphic so we cannot immediately apply Cauchy's Integral Formula. However, for $z \in \gamma(0,1)$ we have

$$\operatorname{Im} z = \frac{1}{2i}(z - \bar{z}) = \frac{1}{2i}\left(z - z^{-1}\right).$$

Hence

$$\begin{aligned}\int_{\gamma(0,1)} \frac{\operatorname{Im} z}{2z - 1}\,\mathrm{d}z &= \frac{1}{2i}\int_{\gamma(0,1)} \frac{z - z^{-1}}{2z - 1}\,\mathrm{d}z \\ &= \frac{1}{2i}\int_{\gamma(0,1)} \frac{z^2 - 1}{z(2z - 1)}\,\mathrm{d}z \\ &= \frac{1}{2i}\int_{\gamma(0,1)} \left(\frac{1}{2} + \frac{1}{z} - \frac{3/4}{z - 1/2}\right)\mathrm{d}z \\ &= \frac{2\pi i}{2i}\left(1\big|_0 - \frac{3}{4}\bigg|_{1/2}\right) = \frac{\pi}{4}.\end{aligned}$$

For the second integral there are two points $z = \pm i/2$ at which the integrand is not holomorphic. However, using partial fractions we can argue as follows.

$$\begin{aligned}\int_{\gamma(0,1)} \frac{\exp z}{4z^2 + 1}\,\mathrm{d}z &= \int_{\gamma(0,1)} \frac{\exp z}{(2z + i)(2z - i)}\,\mathrm{d}z \\ &= \frac{1}{4i}\int_{\gamma(0,1)} \left[\frac{\exp z}{z - i/2} - \frac{\exp z}{z + i/2}\right]\mathrm{d}z \\ &= \frac{2\pi i}{4i}\left[\exp\left(\frac{i}{2}\right) - \exp\left(\frac{-i}{2}\right)\right] \\ &= \frac{\pi}{2} \times 2i \times \sin\frac{1}{2} = \pi i \sin\frac{1}{2}.\end{aligned}$$

■

Theorem 224 ***(Taylor's Theorem)*** *Let $a \in \mathbb{C}$, $\varepsilon > 0$ and let $f\colon D(a,\varepsilon) \to \mathbb{C}$ be a holomorphic function. Then there exist unique $c_n \in \mathbb{C}$ such that*

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n, \qquad z \in D(a,\varepsilon). \tag{11.1}$$

Further

$$c_n = \frac{f^{(n)}(a)}{n!} = \frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)}{(w-a)^{n+1}}\,\mathrm{d}w \qquad (0 < r < \varepsilon). \tag{11.2}$$

Equation (11.2) is often referred to as ***Cauchy's Formula for Derivatives****. We refer to the c_n as the* ***Taylor coefficients****.*

Proof *Existence*: Choose r such that $|z-a| < r < \varepsilon$. By Cauchy's Integral Formula we have

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)}{w-z}\,\mathrm{d}w \\ &= \frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)}{(w-a)-(z-a)}\,\mathrm{d}w \\ &= \frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)/(w-a)}{1-\left(\frac{z-a}{w-a}\right)}\,\mathrm{d}w. \end{aligned}$$

For $w \in \gamma(a,r)$ we have $|z-a| < |w-a| = r$, and so

$$\left(1-\left(\frac{z-a}{w-a}\right)\right)^{-1} = \sum_{n=0}^{\infty}\left(\frac{z-a}{w-a}\right)^n.$$

Hence

$$f(z) = \frac{1}{2\pi i}\int_{\gamma(a,r)} \left[\sum_{n=0}^{\infty} f(w)\frac{(z-a)^n}{(w-a)^{n+1}}\right]\mathrm{d}w.$$

As $\gamma(a,r)$ is compact and $f(w)$ is continuous, then there exists $M > 0$ such that $|f(w)| < M$ on $\gamma(a,r)$. So, for $w \in \gamma(a,r)$,

$$\left|f(w)\frac{(z-a)^n}{(w-a)^{n+1}}\right| < M\frac{|z-a|^n}{r^{n+1}} = \frac{M}{r}\left(\frac{|z-a|}{r}\right)^n =: M_n.$$

As $\sum M_n$ converges – being a convergent geometric series – then, by the Weierstrass M–Test,

$$\sum_{n=0}^{\infty} f(w)\frac{(z-a)^n}{(w-a)^{n+1}}$$

converges uniformly and we can interchange the series and integral to find

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\int_{\gamma(a,r)} \left[\sum_{n=0}^{\infty} f(w)\frac{(z-a)^n}{(w-a)^{n+1}}\right]\mathrm{d}w \\ &= \sum_{n=0}^{\infty} \underbrace{\left[\frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)\,\mathrm{d}w}{(w-a)^{n+1}}\right]}_{c_n}(z-a)^n. \end{aligned}$$

Uniqueness: To demonstrate uniqueness we can make use of the fact that, within the disc of convergence, the term-by-term derivative of a power series converges to the derivative of the function. Hence if (11.1) holds, then

$$f^{(k)}(z)=\sum_{n=k}^{\infty}\frac{n!c_n}{(n-k)!}\,(z-a)^{n-k}$$

and in particular $f^{(k)}(a)=k!c_k$. ■

Corollary 225 *Let f be holomorphic on a domain U. Then $f^{(n)}$ exists and is holomorphic for all $n\geqslant 0$.*

Proof Let $a\in U$ and $\varepsilon>0$ such that $D(a,\varepsilon)\subseteq U$. By Taylor's Theorem we know that $f(z)=\sum_0^\infty a_n z^n$ is defined by a power series. From first year analysis we know that a power series defines a differentiable function on its disc of convergence and that term-by-term differentiation is valid so that $f'(z)=\sum_1^\infty n a_n z^{n-1}$ By induction, f has derivatives of all orders. ■

Definition 226 *Let f be holomorphic on some domain U and let $a\in U$. Let the Taylor series of f at a be*

$$f(z)=\sum_{n=0}^{\infty}c_n\,(z-a)^n\,.$$

*We say that a is a **zero of** f if $f(a)=c_0=0$ and the **order** of the zero to be N where N is the least n such that $c_n\neq 0$.*

Example 227 *Find the orders of the following zeros.*
(a) $\sin^2 z$ *at* 0.
(b) $(\cos z+1)^3$ *at* π.

Solution (a) For $\sin^2 z$ we could either argue that

$$\sin z=z+O(z)\implies\sin^2 z=z^2+O(z^3)$$

or note that

$$\sin^2 0=0;\qquad \left.\frac{\mathrm{d}}{\mathrm{d}z}\right|_0\sin^2 z=2\sin 0\cos 0=0;\qquad \left.\frac{\mathrm{d}^2}{\mathrm{d}z^2}\right|_0\sin^2 z=2\left(\cos^2 0-\sin^2 0\right)=2.$$

Either way we see that the order is 2.

(b) Let $w=z-\pi$. Then

$$\begin{aligned}(\cos z+1)^3&=(\cos(w+\pi)+1)^3\\&=(1-\cos w)^3\\&=\left(1-\left(1-\frac{w^2}{2}+O(w^3)\right)\right)^3\\&=\left(\frac{w^2}{2}+O(w^3)\right)^3\\&=\frac{w^6}{8}+O(w^7).\end{aligned}$$

Hence the order of the zero is 6. ■

Theorem 228 ***(Liouville's Theorem)*** *Let $f\colon \mathbb{C} \to \mathbb{C}$ be a holomorphic function which is bounded. Then f is constant.*

Proof As f is bounded there exists M such that $|f(z)| < M$ for all $z \in \mathbb{C}$. By Taylor's Theorem we have $f(z) = \sum c_n z^n$ for $z \in \mathbb{C}$. So, for $n \geqslant 1$ and $r > 0$ we have

$$\begin{aligned} |c_n| &= \left| \frac{1}{2\pi i} \int_{\gamma(0,r)} \frac{f(w)}{w^{n+1}} \,\mathrm{d}w \right| \\ &\leqslant \frac{1}{2\pi} \times 2\pi r \times \sup_{|w|=r} \left| \frac{f(w)}{w^{n+1}} \right| \\ &\leqslant \frac{1}{2\pi} \times 2\pi r \times \frac{M}{r^{n+1}} \\ &= Mr^{-n} \to 0 \quad \text{as } r \to \infty. \end{aligned}$$

Hence $c_n = 0$ for $n \geqslant 1$ and $f(z) = c_0$ is constant. ■

Corollary 229 *Let f be holomorphic on $\mathbb{C}$ and non-constant. Then $f(\mathbb{C})$ is dense in $\mathbb{C}$ – that is $\overline{f(\mathbb{C})} = \mathbb{C}$.*

Proof Take $a \in \mathbb{C}$ and $\delta > 0$ and suppose for a contradiction that $D(a, \delta) \subseteq \mathbb{C} \backslash f(\mathbb{C})$. Then for all $z \in \mathbb{C}$

$$|f(z) - a| \geqslant \delta$$

and so

$$\frac{1}{|f(z) - a|} \leqslant \frac{1}{\delta}.$$

Hence $(f(z) - a)^{-1}$ is a bounded holomorphic function on $\mathbb{C}$ and so, by Liouville, constant. Hence $f(z)$ is constant, a contradiction. ■

Remark 230 *Liouville's Theorem is considerably weaker than* ***Picard's Little Theorem*** *which is off-syllabus, but which states that a non-constant holomorphic function on $\mathbb{C}$ has image either $\mathbb{C}$ or $\mathbb{C} \backslash \{\text{a single value}\}$.*

Theorem 231 ***(Fundamental Theorem of Algebra)*** *(Gauss 1799) Let p be a non-constant polynomial with complex coefficients. Then there exists $\alpha \in \mathbb{C}$ such that $p(\alpha) = 0$.*

Proof Say that $p(z) = a_n z^n + \cdots + a_0$ where $n \geqslant 1$, $a_i \in \mathbb{C}$ and $a_n \neq 0$. As $p(z)/z^n \to a_n$ as $z \to \infty$ then there exists $R > 0$ such that

$$\left| \frac{p(z)}{z^n} \right| > \frac{|a_n|}{2} \qquad \text{for } |z| > R.$$

Suppose for a contradiction that p has no roots so that $1/p$ is holomorphic. Then, by Cauchy's Integral Formula, with $r > R$, we have

$$\begin{aligned} 0 &\neq \left|\frac{1}{p(0)}\right| = \left|\frac{1}{2\pi i}\int_{\gamma(0,r)} \frac{\mathrm{d}w}{wp(w)}\right| \\ &\leqslant \frac{1}{2\pi} \times 2\pi r \times \sup_{|w|=r}\left|\frac{1}{wp(w)}\right| \\ &\leqslant \frac{1}{2\pi} \times 2\pi r \times \frac{2}{|a_n|\, r^{n+1}} \\ &= \frac{2}{|a_n|\, r^n} \to 0 \qquad \text{as } r \to \infty, \end{aligned}$$

which is the required contradiction. ■

Theorem 232 ***(Morera's Theorem)*** *Let $f\colon U \to \mathbb{C}$ be a continuous function on a domain such that*

$$\int_\gamma f(z)\,\mathrm{d}z = 0$$

for any closed path γ. Then f is holomorphic.

Proof Let $z_0 \in U$. As U is open and connected then it is path-connected and so for any $z \in U$ there is a path $\gamma(z)$ connecting z_0 to z. We will then define

$$F(z) = \int_{\gamma(z)} f(w)\,\mathrm{d}w.$$

Note that if γ_1 and γ_2 are two such paths then $\gamma_1 \cup (-\gamma_2)$ is a closed path and hence by hypothesis

$$0 = \int_{\gamma_1\gamma_2^{-1}} f(w)\,\mathrm{d}w = \int_{\gamma_1} f(w)\,\mathrm{d}w - \int_{\gamma_2} f(w)\,\mathrm{d}w$$

and we see that $F(z)$ is well-defined.

Now take $z_0 \in U$ and $r > 0$ such that $D(z_0, r) \subseteq U$ and $h \in D(z_0, r)$. Then

$$F(z_0 + h) = F(z_0) + \int_{[z_0, z_0+h]} f(w)\,\mathrm{d}w.$$

Hence

$$\begin{aligned} \left|\frac{F(z_0+h) - F(z_0)}{h} - f(z_0)\right| &= \left|\left(\frac{1}{h}\int_{[z_0,z_0+h]} f(w)\,\mathrm{d}w\right) - f(z_0)\right| \\ &= \left|\frac{1}{h}\int_{[z_0,z_0+h]} \left(f(w) - f(z_0)\right)\,\mathrm{d}w\right| \\ &\leqslant \frac{1}{|h|}\,|h| \sup_{[z_0,z_0+h]} |f(w) - f(z_0)| \qquad \text{[by the Estimation Theorem]} \\ &= \sup_{[z_0,z_0+h]} |f(w) - f(z_0)| \to 0 \text{ as } h \to 0 \text{ by the continuity of } f \text{ at } z_0. \end{aligned}$$

Hence F is holomorphic and $F' = f$. By Corollary 225 f is holomorphic. ■

Theorem 233 *(Identity Theorem) Let f be holomorphic on a domain U. Then the following are equivalent:*

(a) $f(z) = 0$ for all $z \in U$.

(b) The zero set $f^{-1}(0)$ has a limit point in U.

(c) There exists $a \in U$ such that $f^{(k)}(a) = 0$ for all $k \geqslant 0$.

Proof We shall demonstrate (a) $\implies$ (c). $\implies$ (b) $\implies$ (a)

- (a) $\implies$ (c): This implication is obvious.
- (c). $\implies$ (b): As U is open then there exists $\varepsilon > 0$ such that $D(a, \varepsilon) \subseteq U$. By Taylor's Theorem, for $z \in D(a, \varepsilon)$ we have

$$f(z) = \sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!}(z-a)^k = 0.$$

  Then $D(a, \varepsilon) \subseteq f^{-1}(0)$ and so $f^{-1}(0)$ has a limit point in U.

- (b) $\implies$ (a): Let a be a limit point of $f^{-1}(0)$ in U. Let $\varepsilon > 0$ be such that $D(a, \varepsilon) \subseteq U$ and let

$$f(z) = \sum_{k=0}^{\infty} c_k (z-a)^k$$

  be the Taylor expansion of f centred about a. Suppose for a contradiction that not all c_k are zero and let K be the smallest k such that $c_K \neq 0$. Then

$$f(z) = (z-a)^K g(z) \qquad \text{where} \qquad g(z) = \sum_{k=K}^{\infty} c_k (z-a)^{k-K}.$$

  Note that $g(z)$ is holomorphic on $D(a, \varepsilon)$ and $g(a) = c_K \neq 0$. By continuity, there exists $\delta > 0$ such that $g(z) \neq 0$ on $D(a, \delta)$. Thus, in $D(a, \delta)$, we see $f(z) = 0$ only holds at a, contradicting the fact that a is a limit point of $f^{-1}(0)$. Hence $c_k = 0$ for all $k \geqslant 0$, and so $f(z) = 0$ on $D(a, \varepsilon)$.

  Now let S denote the set of limit points of $f^{-1}(0)$. By assumption $S \neq \varnothing$. As f is continuous, then $f^{-1}(0)$ is closed and so $S \subseteq f^{-1}(0)$. By the previous part of the argument if $a \in S$ then $D(a, \varepsilon) \subseteq S$ for some $\varepsilon > 0$ and hence S is open. Finally if $z \in U \backslash S$ then z is not a limit point of $f^{-1}(0)$ and so there exists $r > 0$ such that $D(z, r) \subseteq U \backslash S$ and we see that $U \backslash S$ is open and S is closed. As S is open and closed, and as U is connected, then $U = S \subseteq f^{-1}(0)$.

■

Corollary 234 *Let f and g be holomorphic on a domain U. Then the following are equivalent:*

(a) $f(z) = g(z)$ for all $z \in U$.

(b) $f(z) = g(z)$ for all $z \in S$ where $S \subseteq U$ has a limit point in U.

(c) There exists $a \in U$ such that $f^{(k)}(a) = g^{(k)}(a)$ for all $k \geqslant 0$.

Proof Apply the Identity Theorem to $f - g$. ∎

Remark 235 *The Identity Theorem is far from being true on the real line for differentiable, even infinitely differentiable, functions. The function*

$$f(x) = \begin{cases} \exp\left(\frac{-1}{x^2}\right) & x \neq 0, \\ 0 & x = 0, \end{cases}$$

is a commonly given example to show that an infinitely differentiable function need not have a locally convergent Taylor series. In this case $f^{(n)}(0) = 0$ for all n and yet the function is clearly zero only at zero. This contradicts the Identity Theorem. Another example is

$$g(x) = \begin{cases} x^2 \sin\left(\frac{\pi}{x}\right) & x \neq 0, \\ 0 & x = 0, \end{cases}$$

which is differentiable everywhere. The zero set includes $\left\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\right\}$ which has a limit point in $\mathbb{R}$ (namely 0) but the function isn't identically zero. The Identity Theorem gives a sense of how rigid holomorphic functions are compared with even infinitely differentiable real functions.

Example 236 *Find all the functions $f(z)$ which are holomorphic on $D(0,1)$ and which satisfy*

$$f(1/n) = n^2 f(1/n)^3 \qquad n = 2, 3, 4, \ldots$$

Solution We can rewrite this as

$$f\left(\frac{1}{n}\right)\left(f\left(\frac{1}{n}\right) - \frac{1}{n}\right)\left(f\left(\frac{1}{n}\right) + \frac{1}{n}\right) = 0.$$

So, at each n, one of the following holds

$$f\left(\frac{1}{n}\right) = 0, \qquad f\left(\frac{1}{n}\right) = \frac{1}{n}, \qquad f\left(\frac{1}{n}\right) = \frac{-1}{n}.$$

At least one of these three equations must hold for infinitely many n, say $n_1, n_2, \ldots$ This means that one of

$$f(z) = 0, \qquad f(z) = z, \qquad f(z) = -z$$

holds on $S = \{1/n_1, 1/n_2, \ldots\}$. Applying the Identity Theorem we see that $f(z) = 0$ or z or $-z$ on all of $D(0,1)$. ∎

Proposition 237 ***(Counting Zeros)*** *Let f be holomorphic inside and on a positively oriented closed path; assume further than f is non-zero on γ. Then*

$$\text{number of zeros of } f \text{ in } \gamma \text{ (counting multiplicities)} = \frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}\,\mathrm{d}z$$

Proof Let the zeros of f be $a_1, \ldots, a_k$ with multiplicities $m_1, \ldots, m_k$. Then $f'(z)/f(z)$ is holomorphic inside γ except at the a_i. By Taylor's Theorem we know that in an open disc $D(a_i, r_i)$ around a_i we can write

$$f(z) = \sum_{k=m_i}^{\infty} c_k \left(z - a_i\right)^k = \left(z - a_i\right)^{m_i} \underbrace{\sum_{k=m_i}^{\infty} c_k \left(z - a_i\right)^{k-m_i}}_{g(z)} = \left(z - a_i\right)^{m_i} g(z).$$

where $g(z)$ is holomorphic and $g(a_i) \neq 0$. So in $D'(a_i, r_i)$ we have

$$\frac{f'(z)}{f(z)} = \frac{m_i \left(z - a_i\right)^{m_i - 1} g(z) + \left(z - a_i\right)^{m_i} g'(z)}{\left(z - a_i\right)^{m_i} g(z)} = \frac{m_i}{(z - a_i)} + \frac{g'(z)}{g(z)}.$$

In particular we see that

$$\frac{f'(z)}{f(z)} - \frac{m_i}{(z - a_i)} = \frac{g'(z)}{g(z)}$$

is holomorphic in $D(a, r)$. So we similarly see that

$$F(z) = \frac{f'(z)}{f(z)} - \sum_{i=1}^{k} \frac{m_i}{(z - a_i)}$$

is holomorphic inside and on γ, having been suitably adjusted at each zero a_i. By Cauchy's Theorem

$$0 = \int_\gamma F(z)\,\mathrm{d}z = \int_\gamma \frac{f'(z)}{f(z)}\,\mathrm{d}z - \sum_{i=1}^{k} \left(\int_\gamma \frac{m_i}{(z - a_i)} \right) \mathrm{d}z = \int_\gamma \frac{f'(z)}{f(z)}\,\mathrm{d}z - 2\pi i \sum_{i=1}^{k} m_i$$

and the result follows. ■

12. Laurent's Theorem

Theorem 238 ***(Laurent's Theorem)*** *Let f be holomorphic on the annulus*

$$A = \{z \in \mathbb{C} : R < |z - a| < S\}.$$

Then there exist unique c_k $(k \in \mathbb{Z})$ such that

$$f(z) = \sum_{k=-\infty}^{\infty} c_k (z-a)^k \qquad (z \in A)$$

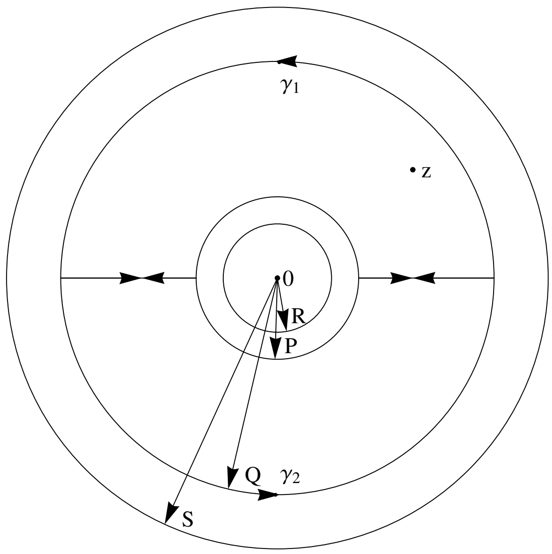
where

$$c_k = \frac{1}{2\pi i} \int_{\gamma(a,r)} \frac{f(w)}{(w-a)^{k+1}} \, \mathrm{d}w \qquad (R < r < S).$$

Proof *Existence*: Let $z \in A$ and choose P, Q such that

$$R < P < |z - a| < Q < S.$$

Let γ_1 and γ_2 be as given in the diagram below.



By Cauchy's Integral Formula and Cauchy's Theorem respectively, we know that

$$f(z) = \frac{1}{2\pi i} \int_{\gamma_1} \frac{f(w)}{w - z} \, \mathrm{d}w, \qquad 0 = \frac{1}{2\pi i} \int_{\gamma_2} \frac{f(w)}{w - z} \, \mathrm{d}w.$$

As the path integrals along the internal line segments cancel then we have

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_{\gamma_1} \frac{f(w)}{w - z} \, \mathrm{d}w + \frac{1}{2\pi i} \int_{\gamma_2} \frac{f(w)}{w - z} \, \mathrm{d}w \\ &= \frac{1}{2\pi i} \int_{\gamma(a,Q)} \frac{f(w)}{w - z} \, \mathrm{d}w - \frac{1}{2\pi i} \int_{\gamma(a,P)} \frac{f(w)}{w - z} \, \mathrm{d}w. \end{aligned}$$

For $w \in \gamma(a,Q)$ note $|z-a| < |w-a|$ and for $w \in \gamma(a,P)$ note $|z-a| > |w-a|$. Hence

$$\begin{aligned} f(z) &= \frac{1}{2\pi i}\int_{\gamma(a,Q)} \frac{f(w)/(w-a)}{1-\frac{(z-a)}{(w-a)}}\,\mathrm{d}w + \frac{1}{2\pi i}\int_{\gamma(a,P)} \frac{f(w)/(z-a)}{1-\frac{(w-a)}{(z-a)}}\,\mathrm{d}w \\ &= \frac{1}{2\pi i}\int_{\gamma(a,Q)} \sum_{k=0}^{\infty} \frac{f(w)(z-a)^k}{(w-a)^{k+1}}\,\mathrm{d}w + \frac{1}{2\pi i}\int_{\gamma(a,P)} \sum_{k=0}^{\infty} \frac{f(w)(w-a)^k}{(z-a)^{k+1}}\,\mathrm{d}w. \end{aligned}$$

Arguing as in Taylor's Theorem with the Weierstrass M–Test, we may show that these sums converge uniformly. Hence we may change the order of the integral and summation to obtain

$$\begin{aligned} f(z) &= \sum_{k=0}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma(a,Q)} \frac{f(w)}{(w-a)^{k+1}}\,\mathrm{d}w\right)(z-a)^k + \sum_{k=0}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma(a,P)} f(w)(w-a)^k\,\mathrm{d}w\right)(z-a)^{-k-1} \\ &= \sum_{k=0}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)}{(w-a)^{k+1}}\,\mathrm{d}w\right)(z-a)^k + \sum_{k=0}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma(a,r)} f(w)\,(w-a)^k\,\mathrm{d}w\right)(z-a)^{-k-1} \\ &= \sum_{k=-\infty}^{\infty}\left(\frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)}{(w-a)^{k+1}}\,\mathrm{d}w\right)(z-a)^k \end{aligned}$$

as required.

Uniqueness: Suppose now that

$$f(z) = \sum_{k=-\infty}^{\infty} d_k (z-a)^k \qquad (z \in A).$$

For $R < r < S$ we have

$$\begin{aligned} 2\pi i c_n &= \int_{\gamma(a,r)} \frac{f(w)}{(w-a)^{n+1}}\,\mathrm{d}w \\ &= \int_{\gamma(a,r)} \sum_{k=-\infty}^{\infty} d_k\,(w-a)^{k-n-1}\,\mathrm{d}w \\ &= \left(\int_{\gamma(a,r)} \sum_{k=0}^{\infty} d_k\,(w-a)^{k-n-1}\,\mathrm{d}w\right) + \left(\int_{\gamma(a,r)} \sum_{k=1}^{\infty} d_{-k}\,(w-a)^{-k-n-1}\,\mathrm{d}w\right). \end{aligned}$$

As the separate power series converge uniformly on $\gamma(a,r)$ then we may exchange integration and summation and find that

$$2\pi i c_n = \sum_{k=-\infty}^{\infty}\left(\int_{\gamma(a,r)} d_k\,(w-a)^{k-n-1}\,\mathrm{d}w\right) = 2\pi i d_n.$$

■

Remark 239 *We will almost never use this integral expression for c_n in order to determine c_n. Rather we will instead use standard power series for familiar functions to determine Laurent series.*

Remark 240 *If f is holomorphic at a then, by uniqueness, the Laurent series of f centred at a is simply the Taylor series of f centred at a.*

Example 241 *Find the Laurent series of:*
(a) $(1-z)^{-1}$ *about* $a=1$.
(b) $(1+z^2)^{-1}\exp z$ *about* $a=i$.

Solution (a) We just note

$$\frac{1}{1-z}=\frac{-1}{z-1}$$

so that $c_{-1}=-1$ and $c_n=0$ for $n\neq -1$. This is obviously convergent on $\mathbb{C}\backslash\{1\}$.

(b) For ease of notation write $w=z-i$. Then

$$\frac{\exp z}{z^2+1}=\frac{\exp(w+i)}{w(w+2i)}=\frac{e^w e^i}{2iw}\left(1+\frac{w}{2i}\right)^{-1}=\frac{e^i}{2i}\frac{1}{w}\left(\sum_{k=0}^{\infty}\frac{w^k}{k!}\right)\left(\sum_{l=0}^{\infty}\frac{(-w)^l}{(2i)^l}\right).$$

Hence $c_n=0$ for $n<-1$ and for $n\geqslant -1$

$$c_n=\frac{e^i}{2i}\sum_{k+l=n+1}\frac{(-1)^l}{k!\,(2i)^l}=\frac{e^i}{2i}\sum_{k+l=n+1}\frac{(i/2)^l}{k!}.$$

This is convergent on $D'(i,2)$. ■

Example 242 *Find c_n, where $n\leqslant 3$, for $f(z)=\cot z$ and $a=0$.*

Solution Using the standard series for sine and cosine we see, for suitably small z,

$$\begin{aligned}
\cot z &= \frac{\cos z}{\sin z}=\frac{1-\frac{1}{2}z^2+\frac{1}{24}z^4-\cdots}{z-\frac{1}{6}z^3+\frac{1}{120}z^5-\cdots}\\
&= \frac{1}{z}\left(1-\frac{z^2}{2}+\frac{z^4}{24}-\cdots\right)\left(1-\frac{z^2}{6}+\frac{z^4}{120}-\cdots\right)^{-1}\\
&= \frac{1}{z}\left(1-\frac{z^2}{2}+\frac{z^4}{24}-\cdots\right)\left(1+\left(\frac{z^2}{6}-\frac{z^4}{120}+\cdots\right)+\left(\frac{z^2}{6}-\frac{z^4}{120}+\cdots\right)^2+\cdots\right)\\
&= \frac{1}{z}\left(1-\frac{z^2}{2}+\frac{z^4}{24}-\cdots\right)\left(1+\frac{z^2}{6}+\left(\frac{1}{36}-\frac{1}{120}\right)z^4+\cdots\right)\\
&= \frac{1}{z}\left(1+\left(\frac{1}{6}-\frac{1}{2}\right)z^2+\left(\frac{1}{24}-\frac{1}{12}+\frac{1}{36}-\frac{1}{120}\right)z^4+\cdots\right)\\
&= \frac{1}{z}\left(1-\frac{1}{3}z^2-\frac{1}{45}z^4+\cdots\right).
\end{aligned}$$

So $c_n=0$ for $n\leqslant 3$ except

$$c_{-1}=1,\qquad c_1=\frac{-1}{3},\qquad c_3=\frac{-1}{45}.$$

■

Example 243 *Find the Laurent expansion of* $(z\,(z-1))^{-1}$ *in the following annuli:*

$$A_1 = \{z \in \mathbb{C} : 0 < |z| < 1\}; \qquad A_2 = \{z \in \mathbb{C} : 1 < |z-2| < 2\}; \qquad A_3 = \left\{z \in \mathbb{C} : |z-i| > \sqrt{2}\right\}.$$

Solution In $A_1 = \{z \in \mathbb{C} : 0 < |z| < 1\}$ we have

$$\frac{1}{z\,(z-1)} = \frac{-1}{z}(1-z)^{-1} = \frac{-1}{z}\left(1+z+z^2+\cdots\right) = -\frac{1}{z} - 1 - z - z^2 - \cdots.$$

In $A_2 = \{z \in \mathbb{C} : 1 < |z-2| < 2\}$, writing $w = z-2$, we have

$$\begin{aligned}
\frac{1}{z\,(z-1)} &= \frac{1}{(1+w)\,(2+w)} = \frac{1}{1+w} - \frac{1}{2+w} \\
&= \frac{1/w}{1+1/w} + \frac{1/2}{1+w/2} \qquad \text{[noting } \left|w^{-1}\right| < 1 \text{ and } |w/2| < 1 \text{ in } A_2] \\
&= \frac{1}{w}\left(1 - \frac{1}{w} + \frac{1}{w^2} - \cdots\right) + \frac{1}{2}\left(1 - \frac{w}{2} + \frac{w^2}{4} - \cdots\right) \\
&= \sum_{-\infty}^{\infty} c_n w^n \quad \text{where} \quad c_n = \begin{cases} (-1)^n\,2^{-n-1} & n \geqslant 0 \\ (-1)^{n+1} & n < 0 \end{cases}
\end{aligned}$$

In $A_3 = \left\{z \in \mathbb{C} : |z-i| > \sqrt{2}\right\}$, writing $w = z-i$, we have

$$\begin{aligned}
\frac{1}{z\,(z-1)} &= \frac{1}{(w+i)\,(w-1+i)} = \frac{-1}{w+i} + \frac{1}{w-1+i} \\
&= \frac{-1/w}{1+i/w} + \frac{1/w}{1+(i-1)/w} \qquad \text{[noting } \left|w^{-1}\right| < |(i-1)/w| < 1 \text{ in } A_3] \\
&= \frac{-1}{w}\left(1 - \frac{i}{w} + \frac{i^2}{w^2} - \frac{i^3}{w^3} + \cdots\right) + \frac{1}{w}\left(1 + \frac{i-1}{w} + \left(\frac{i-1}{w}\right)^2 + \cdots\right) \\
&= \sum_{-\infty}^{\infty} c_n w^n \quad \text{where} \quad c_n = \begin{cases} 0 & n \geqslant 0 \\ (i-1)^{n-1} - (-i)^{n-1} & n < 0 \end{cases}
\end{aligned}$$

■

Definition 244 *Let* $f : U \to \mathbb{C}$ *be defined on a domain* U.

(a) We say that $a \in U$ *is a* ***regular point*** *if* f *is holomorphic at* a.

(b) We say that $a \in U$ *is a* ***singularity*** *if* f *is not holomorphic at* a *but* a *is a limit point of regular points.*

(c) We say that a singularity $a \in U$ *is* ***isolated*** *if* f *is holomorphic on some* $D'(a,r) \subseteq U$.

Definition 245 *Let* a *be an isolated singularity of* f. *By Laurent's Theorem there exist unique* c_n *such that*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n \qquad z \in D'(a,\varepsilon).$$

(i) The ***principal part*** *of f at a is*

$$\sum_{n=-\infty}^{-1} c_n (z-a)^n.$$

(ii) The ***residue*** *of f at a is c_{-1}.*
(iii) a is said to be a ***removable singularity*** *of f if $c_n = 0$ for all $n < 0$.*
(iv) a is said to be a ***pole of order*** *k if $c_{-k} \neq 0$ and $c_n = 0$ for all $n < -k$. Poles of order $1, 2, 3$ are respectively referred to as* ***simple, double*** *and* ***triple poles****.*
(v) a is said to be an ***essential singularity*** *if $c_n \neq 0$ for infinitely many negative n.*
(vi) We will simply classify a singularity which is not isolated as ***non-isolated****.*

Proposition 246 *Let $f, g \colon U \to \mathbb{C}$ be holomorphic and $a \in U$. Suppose that f has a zero of order m at a and g has a zero of order n at a. Then*

$$f/g \text{ has } \begin{cases} \text{a pole of order } n-m \text{ at } a & \text{if } m < n, \\ \text{a removable singularity at } a & \text{if } m \geqslant n. \end{cases}$$

Proof Given the hypotheses, $f(z) = (z-a)^m F(z)$ and $g(z) = (z-a)^n G(z)$ where F and G are holomorphic on U and $F(a) \neq 0 \neq G(a)$. Hence

$$\frac{f(z)}{g(z)} = (z-a)^{m-n} \frac{F(z)}{G(z)} \qquad (z \neq a)$$

where F/G is holomorphic about a and non-zero. The result follows. Note that if $m \geqslant n$ then there is still a singularity at a as the function is undefined, but that if this singularity were removed then f/g would have a zero of order $m-n$ at a. ■

Example 247 *Locate and classify the singularities of the following functions:*

$$f(z) = \frac{z+1}{z^3+1}, \qquad g(z) = \frac{1}{z} - \frac{1}{\sin z}, \qquad h(z) = \frac{1}{e^{1/z}+1}$$

Solution As $z^3 + 1 = 0$ at $z = -1, e^{\pi i/3}, e^{-\pi i/3}$ then these are the singularities of f and they are clearly isolated. Further when $z^3 + 1 \neq 0$ we see

$$f(z) = \frac{1}{(z - e^{\pi i/3})(z - e^{-\pi i/3})}.$$

So f has a removable singularity at -1 and simple poles at $e^{\pi i/3}$ and $e^{-\pi i/3}$.

Note that

$$g(z) = \frac{\sin z - z}{z \sin z}.$$

This has singularities at $z = n\pi$ where $n \in \mathbb{Z}$. At $z = 0$ we note

$$\sin z - z = -\frac{z^3}{3!} + \cdots, \qquad z \sin z = z^2 + \cdots.$$

So $\sin z - z$ has a triple zero and $z \sin z$ has a double zero, so that $g(z)$ has a removable singularity at $z = 0$. At $z = n\pi$ where $n \neq 0$ then $z \sin z$ has a single zero whilst $\sin z - z \neq 0$. Hence $g(z)$ has a simple pole.

We see that $h(z)$ has a singularity when $z = 0$ and when $e^{1/z} = -1$, i.e. at $z_n = (2n\pi i)^{-1}$. As $z_n \to 0$ as $n \to \infty$, then $z = 0$ is a non-isolated singularity. At z_n, looking at the denominator, we have

$$\left.\frac{\mathrm{d}}{\mathrm{d}z}\right|_{z=z_n} \left(e^{1/z} - 1\right) = \frac{-1}{z_n^2} e^{1/z_n} = -\left(2n\pi i\right)^2 (-1) = -4n^2\pi^2 \neq 0.$$

Thus $e^{1/z} - 1$ has a simple zero at $z = z_n$ and hence $h(z)$ has a simple pole at $z = z_n$. ■

Proposition 248 ***(Residues at Simple Poles)*** *Suppose that $f(z)$ has a simple pole at a. Then*

$$\operatorname{res}(f(z); a) = \lim_{z \to a} (z - a) f(z).$$

If $f(z) = g(z)/h(z)$, where g and h are holomorphic at a and h has a simple zero at a, then

$$\operatorname{res}(f(z); a) = \frac{g(a)}{h'(a)}.$$

Proof Let $R = \operatorname{res}(f(z); a)$. Then there exists a holomorphic function h such that

$$f(z) = \frac{R}{z - a} + h(z) \qquad \text{on some } D'(a, \varepsilon).$$

So

$$(z - a)\, f(z) = R + (z - a)\, h(z) \to R \quad \text{as} \quad z \to a.$$

If $f(z) = g(z)/h(z)$ as given then by Taylor's Theorem,

$$g(z) = g(a) + O(z - a), \qquad h(z) = h'(a)\,(z - a) + O((z - a)^2),$$

so that

$$\begin{aligned}
\lim_{z \to a} (z - a) f(z) &= \lim_{z \to a} \frac{(z - a) g(z)}{h(z)} \\
&= \lim_{z \to a} \frac{(z - a)\,[g(a) + O(z - a)]}{h'(a)\,(z - a) + O((z - a)^2)} \\
&= \lim_{z \to a} \frac{g(a) + O(z - a)}{h'(a) + O(z - a)} \\
&= \frac{g(a)}{h'(a)}.
\end{aligned}$$

■

Proposition 249 ***(Residues at Overt Multiple Poles)*** *Suppose that*

$$f(z) = \frac{g(z)}{(z - a)^n} \qquad \textit{where } g \textit{ is holomorphic at } a \textit{ and } g(a) \neq 0,$$

(so that g has an overt pole of order n at a). Then

$$\operatorname{res}(f(z); a) = \frac{g^{(n-1)}(a)}{(n - 1)!}.$$

Proof By Taylor's Theorem we can write, on some $D(a,\varepsilon)$,

$$g(z) = \sum_{k=0}^{\infty} c_k (z-a)^k$$

where $c_k = g^{(k)}(a)/k!$. So on $D'(a,\varepsilon)$ we have

$$f(z) = \sum_{k=0}^{\infty} c_k (z-a)^{k-n}$$

and we see that $\operatorname{res}(f(z);a) = c_{n-1}$ as required. ■

Example 250 *Determine the following residues.*

$$\operatorname{res}((z^2+1)^{-1};i), \qquad \operatorname{res}((z^2-1)^{-3};1), \qquad \operatorname{res}((z-\sin z)^{-1};0), \qquad \operatorname{res}((z-\sin z)^{-2};0).$$

Solution The first example is a (relatively overt) simple pole and we see

$$\operatorname{res}((z^2+1)^{-1};i) = \operatorname{res}\left(\frac{1}{(z-i)(z+i)};i\right) = \left.\frac{1}{z+i}\right|_{z=i} = \frac{1}{2i}.$$

The second example is a (relatively overt) triple pole. We can write

$$\frac{1}{(z^2-1)^3} = \frac{(z+1)^{-3}}{(z-1)^3}$$

and then see

$$\operatorname{res}((z^2-1)^{-3};1) = \frac{1}{2!}\left.\frac{\mathrm{d}^2}{\mathrm{d}z^2}\right|_{z=1}\left((z+1)^{-3}\right) = \frac{(-3)\,(-4)}{2!}(1+1)^{-5} = \frac{12}{2\times 32} = \frac{3}{16}.$$

The third example is a covert triple pole and we have no alternative but to calculate the principal part of the Laurent series. We have a Taylor expansion at 0 of

$$z - \sin z = z - \left(z - \frac{z^3}{6} + \frac{z^5}{120} - \cdots\right) = \frac{z^3}{6} - \frac{z^5}{120} + \cdots.$$

So, by the binomial theorem, and for suitably small z, we have

$$\begin{aligned}
(z-\sin z)^{-1} &= \left(\frac{z^3}{6} - \frac{z^5}{120} + \cdots.\right)^{-1} \\
&= \frac{6}{z^3}\left(1 - \frac{z^2}{20} + \cdots.\right)^{-1} \\
&= \frac{6}{z^3}\left(1 + \left(\frac{z^2}{20} + \cdots\right) + \left(\frac{z^2}{20} + \cdots\right)^2 + \cdots\right) \\
&= \frac{6}{z^3} + \frac{3}{10z} + \cdots
\end{aligned}$$

so that

$$\operatorname{res}((z-\sin z)^{-1};0) = \frac{3}{10}.$$

The fourth example is a covert pole of order six. At first glance calculating the residue looks intimidating, but note that the function is even (about 0) and so its Laurent series will only involve even powers. Hence the residue is zero. ■

13. Residue Theorem & Applications.

Theorem 251 ***(Cauchy's Residue Theorem)*** *Let f be holomorphic inside and on a simple closed, positively oriented path γ except at points $a_1, \ldots, a_n$ inside γ. Then*

$$\int_\gamma f(z)\,\mathrm{d}z = 2\pi i \sum_{k=1}^{n} \operatorname{res}(f(z); a_k).$$

Proof On a disc $D(a_k, \varepsilon_k)$ about a_k we can expand $f(z)$ as a Laurent expansion

$$f(z) = \underbrace{\sum_{n=0}^{\infty} c_n^{(k)}(z-a_k)^n}_{g_k(z)} + \underbrace{\sum_{n=-\infty}^{-1} c_n^{(k)}(z-a_k)^n}_{h_k(z)} = g_k(z) + h_k(z),$$

so that $h_k(z)$ is the principal part of f about a_i. Note also that the sum defining $h_k(z)$ converges except at a_k. Hence

$$F(z) = f(z) - \sum_{k=1}^{n} h_k(z)$$

is holomorphic in γ except for removable singularities at $a_1, \ldots, a_n$. As we have seen each $F(a_k)$ may be assigned a value so as to extend F to a holomorphic function in γ. By Cauchy's Theorem and the Deformation Theorem we have

$$\int_\gamma f(z)\,\mathrm{d}z = \sum_{k=1}^{n} \int_\gamma h_k(z)\,\mathrm{d}z = \sum_{k=1}^{n} \int_{\gamma(a_k, r_k)} h_k(z)\,\mathrm{d}z$$

where $0 < r_k < \varepsilon_k$. As we saw in the proof of Laurent's Theorem, the sum defining $h_k(z)$ converges uniformly on $\gamma(a_k, r_k)$ and hence

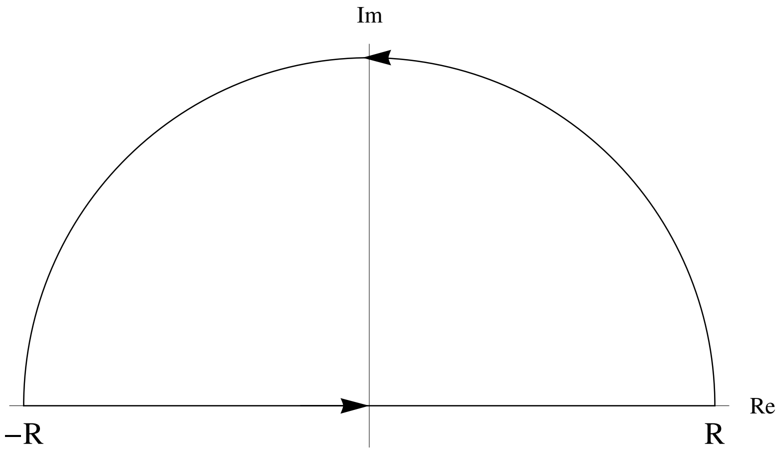
$$\int_{\gamma(a_k,r_k)} h_k(z)\,\mathrm{d}z = \int_{\gamma(a_k,r_k)} \sum_{n=-\infty}^{-1} c_n^{(k)}(z-a_k)^n\,\mathrm{d}z = \sum_{n=-\infty}^{-1} c_n^{(k)} \int_{\gamma(a_k,r_k)} (z-a_k)^n\,\mathrm{d}z = 2\pi i c_{-1}^{(k)}$$

and the result follows. ■

Remark 252 ***Standard Contours.*** *As we shall see, the Residue Theorem is a powerful means of calculating certain real integrals. There are three standard contours that we will use and we shall also introduce certain embellishments here and there when slightly more finesse is needed. If ever a contour is needed other than the standard ones below then you will typically be advised what contour to use.*

- ***Integrals between*** 0 ***and*** 2π. *We have already seen integrals of this type: as θ moves from 0 to 2π then $z = e^{i\theta}$ moves positively around the unit circle $\gamma(0,1)$.*

- ***Integrals from* 0 *to* ∞ *or* $-\infty$ *to* ∞.** *We will standardly use a semicircular contour in the upper half-plane made up of* $\gamma^+(0,R)$ *and the line segment* $[-R,R]$, *as pictured below,*



with the aim of showing that as $R\to\infty$ *the contribution from* $\gamma^+(0,R)$ *tends to* 0. *The contribution from* $[-R,R]$ *then tends to the integral on* $(-\infty,\infty)$ *which we are seeking. If the integrand is even then the integral on* $(0,\infty)$ *will be half that on* $(-\infty,\infty)$.

- ***Infinite sums.*** *We will use a square contour* Γ_N *with vertices at* $(N+1/2)(\pm 1\pm i)$ *where* N *is a positive integer and aim to show that the integral around the contour tends to* 0 *as* $N\to\infty$.

  Full details follow in examples below.

13.1 Integrals on $\mathbb{R}$

Example 253 *Evaluate*

$$\int_{-\infty}^{\infty}\frac{\mathrm{d}x}{1+x^2+x^4}.$$

Solution We will consider $f(z)=(1+z^2+z^4)^{-1}$. As

$$\left(1+z^2+z^4\right)\left(z^2-1\right)=z^6-1$$

then $f(z)$ has simple poles at $e^{\pi i/3}$, $e^{2\pi i/3}$, $e^{4\pi i/3}$, $e^{5\pi i/3}$. If we use the contour Γ which consists of the positively-oriented contour consisting of the interval $[-R,R]$ and the upper semicircle $\gamma^+(0,R)$ then only the simple poles $\alpha=e^{\pi i/3}$ and $\alpha^2=e^{2\pi i/3}$ lie inside Γ and we have

$$\operatorname{res}(f(z);\alpha)=\left.\frac{1}{4z^3+2z}\right|_{z=\alpha}=\frac{1}{4\alpha^3+2\alpha}=\frac{1}{-4+2\alpha};\qquad \operatorname{res}(f(z);\alpha^2)=\frac{1}{4\alpha^6+2\alpha^2}=\frac{1}{4+2\alpha^2}.$$

By Cauchy's Residue Theorem, and noting $\alpha^2+\alpha=i\sqrt{3}$, $\alpha^2-\alpha=-1$ we see

$$\int_\Gamma\frac{1}{1+z^2+z^4}\,\mathrm{d}z=2\pi i\left(\frac{1}{2\alpha-4}+\frac{1}{2\alpha^2+4}\right)=\frac{2\pi i\,(2\alpha^2+2\alpha)}{4(-1-2\alpha^2+2\alpha-4}=\frac{\pi}{\sqrt{3}}.\tag{13.1}$$

Note that if $|z| = R$ then

$$\left|1+z^2+z^4\right| \geqslant \left|z^4\right| - \left|z^2+1\right| \geqslant \left|z^4\right| - \left|z^2\right| - |1| = R^4 - R^2 - 1,$$

so that, by the Estimation Theorem, we have

$$\left|\int_{\gamma^+(0,R)} \frac{\mathrm{d}z}{1+z^2+z^4}\right| \leqslant \frac{\pi R}{R^4-R^2-1} = O\left(\frac{1}{R^3}\right) \to 0 \quad \text{as } R \to \infty.$$

Finally, letting $R \to \infty$ in (13.1) we have

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}x}{1+x^2+x^4} = \frac{\pi}{\sqrt{3}}.$$

■

Example 254 *Let $a \geqslant 0$. Evaluate*

$$\int_0^{\infty} \frac{\cos ax}{1+x^2}\,\mathrm{d}x.$$

What is the value of the integral when $a < 0$?

Remark 255 *At first glance this looks no more complicated than the previous example. However note that*

$$\left|e^{iz}\right| = e^{-\operatorname{Im} z} \qquad \text{and} \qquad \cos z = \frac{e^{iz}+e^{-iz}}{2}.$$

So on $\gamma^+(0,R)$ we see that e^{-iz} and hence $\cos z$ are very large – we would, in due course, have found that the integral around $\gamma^+(0,R)$ did not converge. So rather than integrating $\cos az/(1+z^2)$ we will use $e^{iaz}/(1+z^2)$, noting that e^{iaz} is small on $\gamma^+(0,R)$. When done we can note $\cos ax = \operatorname{Re} e^{iax}$ for real x.

Solution We will consider the function

$$f(z) = \frac{e^{iaz}}{1+z^2},$$

which has simple poles at i and $-i$, and use the contour Γ which denotes the positively-oriented contour consisting of the interval $[-R, R]$ and the upper semicircle $\gamma^+(0,R)$. Only the simple pole i lies inside Γ and we have

$$\operatorname{res}(f(z); i) = \left.\frac{e^{iaz}}{z+i}\right|_{z=i} = \frac{e^{-a}}{2i}.$$

By Cauchy's Residue Theorem

$$\int_{\Gamma} \frac{e^{iaz}}{1+z^2}\,\mathrm{d}z = 2\pi i\left(\frac{e^{-a}}{2i}\right) = \pi e^{-a}.$$

Note that

$$\text{if } \ z = Re^{i\theta} \ \text{ then } \ \left|e^{iaz}\right| = e^{-a\,\mathrm{Re}\,z} = e^{-aR\sin\theta} \leqslant 1;$$
$$\text{if } \ |z| = R \ \text{ then } \ \left|z^2+1\right| = \left|z^2-(-1)\right| \geqslant \left|z^2\right| - |-1| = R^2-1,$$

so that, by the Estimation Theorem, we have

$$\left|\int_{\gamma^+(0,R)} \frac{e^{iaz}}{1+z^2}\,\mathrm{d}z\right| \leqslant \frac{\pi R}{R^2-1} = O\left(\frac{1}{R}\right) \to 0 \ \text{ as } R \to \infty.$$

(Note that this part of the argument relies on a being positive so that $e^{-aR\sin\theta} \leqslant 1$.) Now, letting $R \to \infty$ in the identity

$$\int_{-R}^{R} \frac{e^{iaz}}{1+z^2}\,\mathrm{d}z + \int_{\gamma^+(0,R)} \frac{e^{iaz}}{1+z^2}\,\mathrm{d}z = \pi e^{-a}$$

we arrive at

$$\int_{-\infty}^{\infty} \frac{e^{iax}}{1+x^2}\,\mathrm{d}x = \pi e^{-a}.$$

Comparing real parts we have

$$\int_{-\infty}^{\infty} \frac{\cos ax}{1+x^2}\,\mathrm{d}x = \pi e^{-a}.$$

Using the evenness of the integrand we arrive at

$$\int_{0}^{\infty} \frac{\cos ax}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2e^a}.$$

We clearly couldn't have argued quite this way if a had been negative. One alternative would have been to use the lower semicircle $\gamma^-(0,R)$ and repeat a similar calculation, but we can be smarter than that and note, because of the evenness of $\cos ax$ as a function of a that

$$\int_{0}^{\infty} \frac{\cos(-a)x}{1+x^2}\,\mathrm{d}x = \int_{0}^{\infty} \frac{\cos ax}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2e^a}$$

and so for general $a \in \mathbb{R}$ we have

$$\int_{0}^{\infty} \frac{\cos ax}{1+x^2}\,\mathrm{d}x = \frac{\pi}{2}e^{-|a|}.$$

■

13.2 Infinite Series

Example 256 *Determine*

$$\sum_{n=1}^{\infty} \frac{1}{1+n^2} \qquad \textit{and} \qquad \sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2}.$$

Before we start this example we will need to introduce some relevant theory. The contour we shall use is the positively-oriented square contour Γ_N with vertices $\left(N+\frac{1}{2}\right)(\pm 1 \pm i)$ and we shall use the integrands

$$f(z) = \frac{\pi}{(1+z^2)\tan \pi z}, \qquad g(z) = \frac{\pi}{(1+z^2)\sin \pi z}.$$

Quite why becomes apparent with the next two lemmas.

Lemma 257 *Suppose that the function $\phi(z)$ is holomorphic at $z = n \in \mathbb{Z}$ with $\phi(n) \neq 0$. Then*
(a) $\pi\phi(z)\cot \pi z$ has a simple pole at n with residue $\phi(n)$;
(b) $\pi\phi(z)\csc \pi z$ has a simple pole at n with residue $(-1)^n\phi(n)$.

Proof Note that $\tan \pi z$ and $\sin \pi z$ have simple zeros at $z = n$ and hence $\pi\phi(z)\cot \pi z$ and $\pi\phi(z)\csc \pi z$ have simple poles there. So

$$\operatorname{res}\left(\frac{\pi\phi(z)}{\tan \pi z}; n\right) = \frac{\pi\phi(n)}{\pi\sec^2 \pi n} = \phi(n); \qquad \operatorname{res}\left(\frac{\pi\phi(z)}{\sin \pi z}; n\right) = \frac{\pi\phi(n)}{\pi\cos \pi n} = (-1)^n\phi(n).$$

■

Lemma 258 *There exists a $C > 0$ such that for all N and for all $z \in \Gamma_N$*

$$\left|\frac{\pi}{\tan \pi z}\right| \leqslant C \qquad \text{and} \qquad \left|\frac{\pi}{\sin \pi z}\right| \leqslant C.$$

Proof For the square's top and bottom we note

$$\begin{aligned}
\left|\frac{\pi}{\tan \pi z}\right| &= \pi\left|\frac{\cos(\pi x \pm i\pi(N+1/2))}{\sin(\pi x \pm i\pi(N+1/2))}\right| \\
&= \pi\left|\frac{\exp(i\pi x \mp \pi(N+1/2)) + \exp(-i\pi x \pm \pi(N+1/2))}{\exp(i\pi x \mp \pi(N+1/2)) - \exp(-i\pi x \pm \pi(N+1/2))}\right| \\
&\leqslant \pi\frac{\exp(\mp\pi(N+1/2)) + \exp(\pm\pi(N+1/2))}{|\exp(\mp\pi(N+1/2)) - \exp(\pm\pi(N+1/2))|} \\
&= \pi\coth((N+1/2)\pi) \\
&\leqslant \pi\coth(3\pi/2).
\end{aligned}$$

On the sides we have (using the periodicity of tan)

$$\left|\frac{\pi}{\tan \pi z}\right| = \left|\frac{\pi}{\tan(\pm\pi/2 + \pi i y)}\right| = \left|\frac{\pi}{\cot \pi i y}\right| = \pi\left|\tanh \pi y\right| \leqslant \pi.$$

Similarly

$$\begin{aligned}
\left|\frac{\pi}{\sin \pi z}\right| &= \frac{\pi}{|\sin(\pi x \pm i\pi(N+1/2))|} \\
&= \frac{2\pi}{|\exp(i\pi x \mp \pi(N+1/2)) - \exp(-i\pi x \pm \pi(N+1/2))|} \\
&\leqslant \frac{2\pi}{|\exp(\mp\pi(N+1/2)) - \exp(\pm\pi(N+1/2))|} \\
&= \frac{2\pi}{\sinh((N+1/2)\pi)} \\
&\leqslant \frac{2\pi}{\sinh(3\pi/2)}
\end{aligned}$$

and

$$\left|\frac{\pi}{\sin \pi z}\right| = \left|\frac{\pi}{\sin(\pm\pi/2 + \pi i y)}\right| = \left|\frac{\pi}{\cos \pi i y}\right| = \pi\,|\mathrm{sech} \pi y| \leqslant \pi.$$

Set C to be the maximum of these three numbers. ■

Solution (To Example 256) Consider the positively-oriented square contour Γ_N with vertices $\left(N+\frac{1}{2}\right)(\pm 1 \pm i)$ and the functions

$$f(z) = \frac{\pi}{(1+z^2)\tan \pi z}, \qquad g(z) = \frac{\pi}{(1+z^2)\sin \pi z}.$$

We know from Lemma 257 that f and g have simple poles at $n \in \mathbb{Z}$ and that

$$\operatorname{res}(f(z); n) = \frac{1}{1+n^2}, \qquad \operatorname{res}(g(z); n) = \frac{(-1)^n}{1+n^2}.$$

Further there are simple poles at $\pm i$ with

$$\begin{aligned} \operatorname{res}(f(z); i) &= \operatorname{res}(f(z); -i) = \frac{\pi}{2i \tan \pi i} = \frac{-\pi}{2\tanh \pi}; \\ \operatorname{res}(g(z); i) &= \operatorname{res}(g(z); -i) = \frac{\pi}{2i \sin \pi i} = \frac{-\pi}{2\sinh \pi}. \end{aligned}$$

Hence, by Cauchy's Residue Theorem, we have

$$\begin{aligned} \frac{1}{2\pi i}\int_{\Gamma_N} f(z)\,\mathrm{d}z &= \sum_{n=-N}^{N} \frac{1}{1+n^2} - \frac{\pi}{\tanh \pi} = 2\sum_{n=1}^{N} \frac{1}{1+n^2} + 1 - \frac{\pi}{\tanh \pi}; \\ \frac{1}{2\pi i}\int_{\Gamma_N} g(z)\,\mathrm{d}z &= \sum_{n=-N}^{N} \frac{(-1)^n}{1+n^2} - \frac{\pi}{\sinh \pi} = 2\sum_{n=1}^{N} \frac{(-1)^n}{1+n^2} + 1 - \frac{\pi}{\sinh \pi}. \end{aligned}$$

Now by the Estimation Theorem and Lemma 258 there exists $C > 0$ such that

$$\begin{aligned} \left|\int_{\Gamma_N} f(z)\,\mathrm{d}z\right| &\leqslant 4(2N+1)\,C \sup_{z\in\Gamma_N}\left|\frac{1}{1+z^2}\right| \leqslant \frac{4(2N+1)\,C}{\left(N+\frac{1}{2}\right)^2 - 1} = O\left(\frac{1}{N}\right) \to 0 \text{ as } N \to \infty. \\ \left|\int_{\Gamma_N} g(z)\,\mathrm{d}z\right| &\leqslant 4(2N+1)\,C \sup_{z\in\Gamma_N}\left|\frac{1}{1+z^2}\right| \leqslant \frac{4(2N+1)\,C}{\left(N+\frac{1}{2}\right)^2 - 1} = O\left(\frac{1}{N}\right) \to 0 \text{ as } N \to \infty. \end{aligned}$$

Letting $N \to \infty$ we then have

$$\begin{aligned} 2\sum_{n=1}^{\infty} \frac{1}{1+n^2} + 1 - \frac{\pi}{\tanh \pi} &= 0 \implies \sum_{n=1}^{\infty} \frac{1}{1+n^2} = \frac{\pi \coth \pi - 1}{2}. \\ 2\sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2} + 1 - \frac{\pi}{\sinh \pi} &= 0 \implies \sum_{n=1}^{\infty} \frac{(-1)^n}{1+n^2} = \frac{\pi \,\mathrm{csch}\, \pi - 1}{2}. \end{aligned}$$

■

Remark 259 *Should we wish to sum*

$$\sum_{n=1}^{\infty} \frac{1}{n^4}, \qquad \sum_{n=2}^{\infty} \frac{(-1)^n}{n^2-1},$$

we would need to note that our chosen ϕ *would have poles that mingle with the poles of* $\cot \pi z$ *and* $\csc \pi z$*. This is not in itself problematic save that it means such poles need to be treated separately.*

13.3 Refinements

Remark 260 *For integrals such as*

$$\int_{-\infty}^{\infty} \frac{x \sin x}{1+x^2+x^4}\, \mathrm{d}x \qquad \int_{-\infty}^{\infty} \frac{\cos x}{1+x^6}\, \mathrm{d}x$$

we would have no great trouble arguing along the lines of Example 254. These integrals converge relatively quickly and we would see that the contribution from $\gamma^{+}(0,R)$ *was* $O(R^{-2})$ *and* $O(R^{-5})$ *respectively. For integrals such as*

$$\int_{-\infty}^{\infty} \frac{\sin x}{x}\, \mathrm{d}x, \qquad \int_{0}^{\infty} \frac{x \sin x}{1+x^2}\, \mathrm{d}x$$

without subtler reasoning we would estimate the contribution from $\gamma^{+}(0,R)$ *as* $O(1)$*. The problem is that these integrals do not converge fast enough. (To be technical these integrals in fact only exist as improper integrals.) We need the following inequality due to Jordan to strengthen our estimate.*

Lemma 261 ***(Jordan's Lemma)***

$$\frac{2}{\pi} < \frac{\sin \theta}{\theta} < 1 \qquad \textit{for } \ 0 < \theta < \pi/2.$$

Proof Let $f(\theta) = \sin\theta/\theta$ on $(0,\pi/2)$. Note that

$$f'(\theta) = \frac{\theta \cos\theta - \sin\theta}{\theta^2}.$$

Now, for $0 < \theta < \pi/2$ let

$$g(\theta) = \theta\cos\theta - \sin\theta,$$

so that

$$g'(\theta) = -\theta\sin\theta + \cos\theta - \cos\theta = -\theta\sin\theta < 0,$$

so that g is decreasing on $(0,\pi/2)$. As $g(0) = 0$ then $g(\theta) < 0$ for $0 < \theta < \pi/2$. So

$$f'(\theta) = \frac{g(\theta)}{\theta^2} < 0 \qquad \text{for } \ 0 < \theta < \pi/2$$

so that f is decreasing on $(0,\pi/2)$. As $f(\theta) \to 1$ as $\theta \to 0$ then

$$\frac{2}{\pi} < f(\theta) < 1 \qquad \text{for } \ 0 < \theta < \pi/2$$

■

Example 262 *Determine*

$$\int_0^\infty \frac{x \sin x}{1+x^2}\,\mathrm{d}x.$$

Solution We will consider the function

$$f(z) = \frac{ze^{iz}}{1+z^2},$$

which has simple poles at i and $-i$, and use the contour Γ which denotes the positively-oriented contour consisting of the interval $[-R, R]$ and the upper semicircle $\gamma^+(0, R)$. Only the simple pole i lies inside Γ and we have

$$\operatorname{res}(f(z); i) = \left.\frac{ze^{iz}}{z+i}\right|_{z=i} = \frac{ie^{-1}}{2i} = \frac{1}{2e}.$$

By Cauchy's Residue Theorem

$$\int_\Gamma \frac{ze^{iz}}{1+z^2}\,\mathrm{d}z = \frac{2\pi i}{2e} = \frac{\pi i}{e}. \tag{13.2}$$

Note that

$$\begin{aligned} \text{if } \ z = Re^{i\theta} \ \text{ then } \left|e^{iz}\right| &= e^{-\operatorname{Re} z} = e^{-R\sin\theta}; \\ \text{if } \ |z| = R \ \text{ then } \left|z^2+1\right| &= \left|z^2 - (-1)\right| \geqslant \left|z^2\right| - |-1| = R^2 - 1, \end{aligned}$$

so that, by the Estimation Theorem, we have

$$\begin{aligned} \left|\int_{\gamma^+(0,R)} \frac{ze^{iz}}{1+z^2}\,\mathrm{d}z\right| &= \left|\int_0^\pi \frac{Re^{i\theta}e^{iRe^{i\theta}}}{1+R^2e^{2i\theta}}\, iRe^{i\theta}\mathrm{d}\theta\right| \\ &\leqslant \int_0^\pi \frac{Re^{-R\sin\theta}}{R^2-1}\, R\,\mathrm{d}\theta \\ &= \frac{2R^2}{R^2-1}\int_0^{\pi/2} e^{-R\sin\theta}\,\mathrm{d}\theta \\ &\leqslant \frac{2R^2}{R^2-1}\int_0^{\pi/2} e^{-2R\theta/\pi}\,\mathrm{d}\theta \qquad \text{[by Jordan's Lemma]} \\ &= \frac{2R^2}{R^2-1} \times \left(\frac{-\pi}{2R}\right) \times \left[e^{-2R\theta/\pi}\right]_0^{\pi/2} \\ &= \frac{\pi R^2\left(1-e^{-R}\right)}{R(R^2-1)} = O\left(\frac{1}{R}\right) \to 0 \ \text{ as } R \to \infty. \end{aligned}$$

Letting $R \to \infty$ in (13.2) we obtain

$$\int_{-\infty}^\infty \frac{xe^{ix}}{1+x^2}\,\mathrm{d}x = \frac{\pi i}{e}.$$

Taking imaginary parts we find

$$\int_{-\infty}^{\infty} \frac{x \sin x}{1+x^2}\, \mathrm{d}x = \frac{\pi}{e}.$$

Finally using the even-ness of the integrand we obtain

$$\int_{0}^{\infty} \frac{x \sin x}{1+x^2}\, \mathrm{d}x = \frac{\pi}{2e}.$$

■

Remark 263 *With some integrals the natural integrand to take may leave us with a singularity on the contour. This isn't a problem when that singularity is a simple pole as we may* ***indent*** *at such singularities (see lemma below).* ***It is impossible to sensibly indent at other singularities*** *and have any hope of convergence; it is only possible to indent at simple poles because the contour has length $O(\varepsilon)$ and the integrand grows as $O(\varepsilon^{-1})$.*

Lemma 264 *Let f be a holomorphic function on $D'(a, \varepsilon)$ with a simple pole at a. Let $0 < r < \varepsilon$, and let γ_r be the positively oriented arc*

$$\gamma_r(\theta) = a + re^{i\theta} \qquad \alpha < \theta < \beta.$$

Then

$$\lim_{r \to 0} \int_{\gamma_r} f(z)\, \mathrm{d}z = (\beta - \alpha)\, i \times \operatorname{res}(f(z); a).$$

Proof Let $R = \operatorname{res}(f(z); a)$. There is a holomorphic function h on $D'(a, \varepsilon)$ such that

$$f(z) = \frac{R}{z-a} + h(z).$$

Now

$$\int_{\gamma_r} \frac{\mathrm{d}z}{z-a} = \int_{\alpha}^{\beta} \frac{ire^{i\theta}\, \mathrm{d}\theta}{re^{i\theta}} = i \int_{\alpha}^{\beta} \mathrm{d}\theta = i\,(\beta - \alpha)\,.$$

Further, as $\bar{D}(a, \varepsilon/2)$ is compact and h is continuous, then h is bounded on $\bar{D}(a, \varepsilon/2)$, and so on every γ_r with $r < \varepsilon/2$, say by M. So, by the Estimation Theorem

$$\left| \int_{\gamma_r} h(z)\, \mathrm{d}z \right| \leqslant \mathcal{L}(\gamma_r) \times M = M\,(\beta - \alpha)\, r \to 0 \qquad \text{as } r \to 0.$$

Finally then

$$\lim_{r \to 0} \int_{\gamma_r} f(z)\, \mathrm{d}z = \lim_{r \to 0} \int_{\gamma_r} \frac{R\, \mathrm{d}z}{z-a} + \lim_{r \to 0} \int_{\gamma_r} h(z)\, \mathrm{d}z = iR\,(\beta - \alpha) + 0 = iR\,(\beta - \alpha)$$

as required. ■

Example 265 *Determine*

$$\int_{0}^{\infty} \frac{\sin^2 x}{x^2}\, \mathrm{d}x.$$

Solution Note that
$$\frac{\sin^2 x}{x^2} = \frac{1-\cos 2x}{2x^2}$$
and so for an integrand we will use $(1-e^{2iz})\,/\,(2z^2)$. This has a simple pole at 0 as

$$\frac{1-e^{2iz}}{2z^2} = \frac{1-\left(1+2iz+\frac{(2iz)^2}{2!}+\cdots\right)}{2z^2} = \frac{-i}{z}+2+\cdots$$

So we shall indent our usual contour Γ so that we now have a contour made up of line segments $[-R,-\varepsilon]$ and $[\varepsilon, R]$, a positively oriented semicircle $\gamma^+(0,R)$ and a negatively oriented indent $\gamma^+(0,\varepsilon)$. As the integrand is holomorphic inside the contour Γ then, by Cauchy's Theorem, we have
$$\int_\Gamma \frac{1-e^{2iz}}{2z^2}\,\mathrm{d}z = 0.$$
On the large semicircle
$$\left|\frac{1-e^{2iz}}{2z^2}\right| = \frac{1-e^{-2R\sin\theta}}{2R^2} \leqslant \frac{1}{2R^2}$$
so that, by the estimation theorem,

$$\left|\int_{\gamma^+(0,R)} \frac{1-e^{2iz}}{2z^2}\,\mathrm{d}z\right| \leqslant \frac{\pi R}{2R^2} = O\left(\frac{1}{R}\right) \to 0 \quad \text{as} \quad R\to\infty.$$

We also have from the previous lemma that

$$-\int_{\gamma^+(0,\varepsilon)} \frac{1-e^{2iz}}{2z^2}\,\mathrm{d}z = -\pi i \times \operatorname{res}\left(\frac{1-e^{2iz}}{2z^2};0\right) = (-\pi i)\times(-i) = -\pi.$$

We have the identity

$$\int_{\gamma^+(0,R)} \frac{1-e^{2iz}}{2z^2}\,\mathrm{d}z + \int_{-R}^{-\varepsilon} \frac{1-e^{2ix}}{2x^2}\,\mathrm{d}x - \int_{\gamma^+(0,\varepsilon)} \frac{1-e^{2iz}}{2z^2}\,\mathrm{d}z + \int_{\varepsilon}^{R} \frac{1-e^{2ix}}{2x^2}\,\mathrm{d}x = 0$$

in R and ε. If we let $R\to\infty$, $\varepsilon\to 0$ and take real parts then we obtain

$$-\pi + 2\lim_{\substack{R\to\infty\\ \varepsilon\to 0}} \int_\varepsilon^R \frac{1-\cos 2x}{2x^2}\,\mathrm{d}z = 0.$$

Hence
$$\int_0^\infty \frac{\sin^2 x}{x^2}\,\mathrm{d}x = \frac{\pi}{2}.$$

■

13.4 Further Contours

Example 266 *Let $-1 < a < 2$. Use a keyhole contour to determine*

$$\int_0^\infty \frac{x^a}{1+x^3}\,\mathrm{d}x.$$

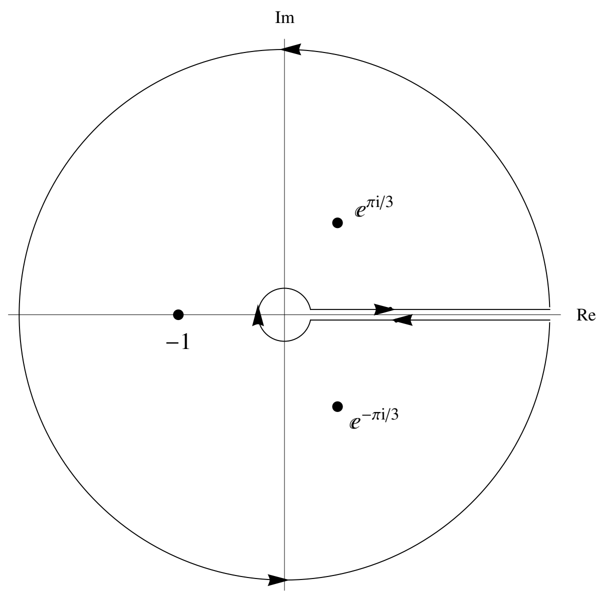
Solution We will consider the function

$$f(z) = \frac{z^a}{1+z^3}.$$

We need to define a branch of z^a on the cut-plane $\mathbb{C}\backslash[0,\infty)$. Any z in the cut-plane can be written $z = re^{i\theta}$ where $0 < \theta < 2\pi$; we then set

$$z^a = r^a e^{ia\theta}.$$

Our keyhole contour Γ comprises two line segments just above and just below $[0, R]$. Alternatively, we can consider these two line segments both as $[0, R]$ but with one on the lower branch with $\theta = 0$ and one of the upper branch with $\theta = 2\pi$. (Cauchy's Residue Theory is valid on Riemann surfaces so this viewpoint is perfectly valid.) We also have as part of our keyhole contour a small cut circle around the origin but as the integrand is $O(\varepsilon^a)$ and the contour has length $O(\varepsilon)$ then the contribution from this is zero in the limit.



Note that f is holomorphic inside Γ except at -1, $e^{\pi i/3}$ and $e^{-\pi i/3}$. Note also that we should write $e^{-\pi i/3} = e^{5\pi i/3}$ for the purpose of the branch's definition. So

$$\begin{aligned}
\operatorname{res}(f;-1) &= \left.\frac{z^a}{3z^2}\right|_{z=-1} = \frac{1}{3}e^{\pi i a};\\
\operatorname{res}(f;e^{\pi i/3}) &= \left.\frac{z^a}{3z^2}\right|_{z=e^{\pi i/3}} = \frac{1}{3}e^{\pi i(a-2)/3};\\
\operatorname{res}(f;e^{-\pi i/3}) &= \left.\frac{z^a}{3z^2}\right|_{z=e^{5\pi i/3}} = \frac{1}{3}e^{5\pi i(a-2)/3}.
\end{aligned}$$

By the Residue Theorem we have

$$\int_{\Gamma} \frac{z^a \, \mathrm{d}z}{1+z^3} = \frac{2\pi i}{3} \left(e^{\pi i a} + e^{\pi i (a-2)/3} + e^{5\pi i (a-2)/3} \right).$$

The contribution from the outside circle $\gamma(0,R)$ satisfies

$$\left| \int_{\gamma(0,R)} \frac{z^a \mathrm{d}z}{1+z^3} \right| \leqslant \frac{2\pi R^{1+a}}{R^3} = O(R^{a-2}) \to 0 \quad \text{as} \quad R \to \infty.$$

Hence we have, letting $R \to \infty$, that

$$\int_0^\infty \frac{x^a \, \mathrm{d}x}{1+x^3} - e^{2\pi i a} \int_0^\infty \frac{x^a \, \mathrm{d}x}{1+x^3} = \frac{2\pi i}{3} \left(e^{\pi i a} + e^{\pi i (a-2)/3} + e^{5\pi i (a-2)/3} \right).$$

Rearranging we have

$$\begin{aligned}
\int_0^\infty \frac{x^a \, \mathrm{d}x}{1+x^3} &= \frac{2\pi i}{3} \times \left(\frac{e^{\pi i a} + e^{\pi i (a-2)/3} + e^{5\pi i (a-2)/3}}{1 - e^{2\pi i a}} \right) \\
&= \frac{2\pi i}{3} \times \left(\frac{1 + e^{\pi i (-2a-2)/3} + e^{\pi i (2a+2)/3}}{e^{-\pi i a} - e^{\pi i a}} \right) \\
&= \frac{2\pi i}{3} \times \left(\frac{e^{\pi i (2(a+1))/3} + 1 + e^{\pi i (-2(a+1))/3}}{e^{\pi i (a+1)} - e^{-\pi i (a+1)}} \right) \\
&= \frac{2\pi i}{3} \times \left(\frac{e^{\pi i (2(a+1))/3} + 1 + e^{\pi i (-2(a+1))/3}}{e^{\pi i (a+1)} - e^{-\pi i (a+1)}} \right) \\
&= \frac{2\pi i}{3} \times \left(\frac{1}{e^{\pi i (a+1)/3} - e^{-\pi i (a+1)/3}} \right) \\
&= \frac{\pi}{3} \csc \left(\frac{(a+1)\pi}{3} \right).
\end{aligned}$$

The cancellation to the penultimate line uses the identity $(x^2+xy+y^2)/(x^3-y^3) = 1/(x-y)$.
■

Remark 267 *The above integral could also have been done by using the contour in the next example which goes around a sector.*

Example 268 *Evaluate*

$$\int_0^\infty \frac{\mathrm{d}x}{1+x^3+x^6}.$$

Solution The denominator has roots at

$$e^{2\pi i/9}, \quad e^{4\pi i/9}, \quad e^{8\pi i/9}, \quad e^{10\pi i/9}, \quad e^{14\pi i/9}, \quad e^{16\pi i/9}$$

as $1+x^3+x^6 = (x^9-1)/(x^3-1)$. Let $R > 1$ and let $\alpha = e^{2\pi i/9}$. We will consider the positively oriented contour Γ made up of the line segment from 0 to R, the circular arc centred at 0 connecting R to $R\alpha^3$ and the line segment from $R\alpha^3$ to 0. Let $f(z) = (1+z^3+z^6)^{-1}$.

Then $f(z)$ has six simple poles at the above roots with only α and α^2 being inside Γ. By Cauchy's Residue Theorem

$$\begin{aligned}
\int_\Gamma \frac{\mathrm{d}z}{1+z^3+z^6} &= 2\pi i\left[\operatorname{res}\left(\frac{1}{1+z^3+z^6};\alpha\right)+\operatorname{res}\left(\frac{1}{1+z^3+z^6};\alpha^2\right)\right] \\
&= \frac{2\pi i}{3}\left[\frac{1}{\alpha^2+2\alpha^5}+\frac{1}{\alpha^4+2\alpha}\right] \\
&= \frac{2\pi i}{3}\left[\frac{(2\alpha^2+2\alpha^{-2}+\alpha+\alpha^{-1})}{(1+2\alpha^3)(\alpha^3+2)}\right].
\end{aligned}$$

If we write $\beta = 2\pi/9$ then this simplifies as

$$\frac{2\pi i}{3\alpha^3}\left[\frac{4\cos(2\beta)+2\cos\beta}{(\alpha^{-3}+2)(\alpha^3+2)}\right] = \frac{2\pi i}{3\alpha^3}\left[\frac{4\cos(2\beta)+2\cos\beta}{|2+\alpha^3|^2}\right] = \frac{4\pi i}{9\alpha^3}\left[2\cos(2\beta)+\cos\beta\right].$$

The contribution to this integral from the circular arc satisfies

$$\left|\int_{\text{arc}} \frac{\mathrm{d}z}{1+z^3+z^6}\right| \leqslant \frac{2\pi R}{3}\sup_{\text{arc}}\left|\frac{1}{1+z^3+z^6}\right| \leqslant \frac{2\pi R}{3(R^6-R^3-1)} = O\left(R^{-5}\right)\to 0 \quad \text{as } R\to\infty.$$

Hence, letting $R\to\infty$, we have

$$\int_0^\infty \frac{\mathrm{d}x}{1+x^3+x^6}+\int_\infty^0 \frac{\alpha^3\,\mathrm{d}x}{1+x^3+x^6} = \frac{4\pi i}{9\alpha^3}\left[2\cos(2\beta)+\cos\beta\right],$$

giving

$$\begin{aligned}
\int_0^\infty \frac{\mathrm{d}x}{1+x^3+x^6} &= \frac{4\pi i}{9}\frac{\left[2\cos(2\beta)+\cos\beta\right]}{\alpha^3-\alpha^6} \\
&= \frac{4\pi}{9\sqrt{3}}\left[2\cos(2\beta)+\cos\beta\right]
\end{aligned}$$

■

Example 269 *Show that*

$$\int_0^\infty \frac{\sqrt[3]{x}\log x}{1+x^2}\,\mathrm{d}x = \frac{\pi^2}{6} \qquad \textit{and} \qquad \int_0^\infty \frac{\sqrt[3]{x}}{1+x^2}\,\mathrm{d}x = \frac{\pi}{\sqrt{3}}.$$

Remark 270 *Note that it's not necessary to use a keyhole contour here as along the negative real axis our integrand at $x=-t<0$ becomes*

$$\frac{\sqrt[3]{-t}\log(-t)}{1+(-t)^2} = \frac{e^{\pi i/3}\sqrt[3]{t}\,(\log t+\pi i)}{1+t^2}$$

which contributes, in its real and imaginary parts, towards both integrals.

Solution Let $R > 1$ and let U be the cut complex plane with the negative imaginary axis removed. Consider the positively oriented contour Γ comprising $[-R, R]$ and $\gamma^+(0, R)$ and the function

$$f(z) = \frac{\sqrt[3]{z}\log z}{1+z^2} \qquad (z \in U)$$

where

$$\log z = \log|z| + i\arg z, \qquad \sqrt[3]{z} = \exp\left(\frac{1}{3}\log z\right)$$

and $-\frac{\pi}{2} < \arg z < \frac{3\pi}{2}$. By these definitions $\sqrt[3]{i} = e^{\pi i/6}$ and $\log i = \pi i/2$. So

$$\int_\Gamma \frac{\sqrt[3]{z}\log z}{1+z^2}\,\mathrm{d}z = 2\pi i \operatorname{res}(f; i) = 2\pi i\left(\frac{\sqrt[3]{i}\log i}{2i}\right) = \frac{\pi^2 i}{2}e^{\pi i/6}.$$

Now

$$\left|\int_{\gamma^+(0,R)} \frac{\sqrt[3]{z}\log z}{1+z^2}\,\mathrm{d}z\right| \leqslant \pi R \sup_{z\in\gamma^+(0,R)}\left|\frac{\sqrt[3]{z}\log z}{1+z^2}\right| \leqslant \frac{\pi R^{4/3}(\log R + \pi)}{R^2-1} \to 0 \quad \text{as } R \to \infty.$$

For $z = -x$ where $x > 0$ we have $\sqrt[3]{-x} = \sqrt[3]{x}e^{\pi i/3}$ and $\log(-x) = \log x + i\pi$. Hence, letting $R \to \infty$ we have

$$\int_0^\infty \frac{\sqrt[3]{x}e^{\pi i/3}(\log x + i\pi)}{1+x^2}\,\mathrm{d}x + \int_0^\infty \frac{\sqrt[3]{x}\log x}{1+x^2}\,\mathrm{d}x = \frac{\pi^2 i}{2}e^{\pi i/6}$$

which simplifies to

$$i\pi e^{\pi i/3}\int_0^\infty \frac{\sqrt[3]{x}}{1+x^2}\,\mathrm{d}x + \left(1+e^{\pi i/3}\right)\int_0^\infty \frac{\sqrt[3]{x}\log x}{1+x^2}\,\mathrm{d}x = \frac{\pi^2 i}{2}e^{\pi i/6}.$$

As $e^{\pi i/6}/\left(1+e^{\pi i/3}\right) = 1/\left(e^{-\pi i/6}+e^{\pi i/6}\right) = 1/(2\cos\pi/6) = 1/\sqrt{3}$ then

$$\frac{i\pi e^{\pi i/6}}{\sqrt{3}}\int_0^\infty \frac{\sqrt[3]{x}}{1+x^2}\,\mathrm{d}x + \int_0^\infty \frac{\sqrt[3]{x}\log x}{1+x^2}\,\mathrm{d}x = \frac{\pi^2 i}{2\sqrt{3}}. \tag{13.3}$$

Taking imaginary parts we find

$$\frac{\pi}{2}\int_0^\infty \frac{\sqrt[3]{x}}{1+x^2}\,\mathrm{d}x = \frac{\pi^2}{2\sqrt{3}} \implies \int_0^\infty \frac{\sqrt[3]{x}}{1+x^2}\,\mathrm{d}x = \frac{\pi}{\sqrt{3}}$$

and then taking real parts in (13.3) we find

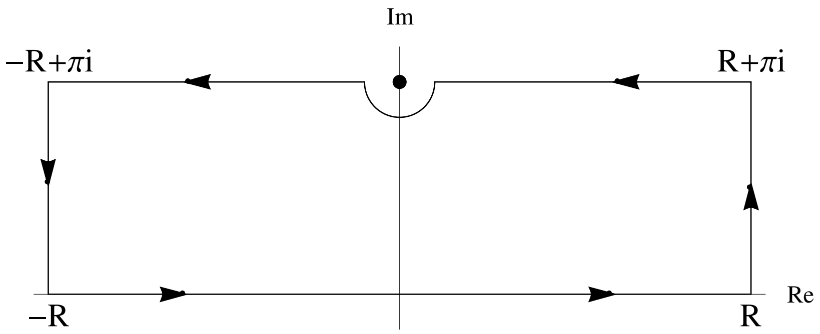
$$\int_0^\infty \frac{\sqrt[3]{x}\log x}{1+x^2}\,\mathrm{d}x = \operatorname{Re}\left(-\frac{i\pi e^{\pi i/6}}{\sqrt{3}}\right)\int_0^\infty \frac{\sqrt[3]{x}}{1+x^2}\,\mathrm{d}x = \left(\frac{\pi}{2\sqrt{3}}\right)\left(\frac{\pi}{\sqrt{3}}\right) = \frac{\pi^2}{6}.$$

■

Example 271 *By considering a rectangular contour Γ with corners at R, $R+i\pi$, $-R+i\pi$, $-R$, and with an appropriate indentation, determine*

$$\int_0^\infty \frac{x}{\sinh x}\mathrm{d}x.$$

Solution We will consider the function $f(z) = z/\sinh z$ on the suggested contour. Then $f(z)$ has a removable singularity at 0 (which we can remove, and so treat the function as holomorphic there) and also a simple pole at $z = \pi i$ which we need to indent around.



As $f(z)$ is holomorphic in the rectangle then

$$\int_\Gamma f(z)\,\mathrm{d}z = 0$$

by Cauchy's Theorem. The contribution from the rectangle's sides satisfy

$$\left|\int_0^\pi \frac{\pm R + iy}{\sinh(\pm R + iy)}\,i\,\mathrm{d}y\right| = \left|\int_0^\pi \frac{2\,(\pm R + iy)}{e^{\pm R+iy} - e^{\mp R-iy}}\,\mathrm{d}y\right| \leqslant \left|\int_0^\pi \frac{2\,(\pm R + iy)}{|e^{\pm R} - e^{\mp R}|}\,\mathrm{d}y\right| \leqslant \frac{R\pi + \pi^2/2}{\sinh R} \to 0$$

as $R \to \infty$. The residue at πi is

$$\mathrm{res}(f(z); \pi i) = \frac{\pi i}{\cosh \pi i} = -\pi i.$$

Also note along the top of the rectangle that the function is

$$f(x + \pi i) = \frac{x + \pi i}{\sinh(x + \pi i)} = -\frac{x + \pi i}{\sinh x}.$$

(This identity, coming from the periodicity of $\sinh x$, is the very reason for choosing the rectangular contour.) Hence letting $R \to \infty$ and shrinking the indentation to zero, we see

$$\int_{-\infty}^\infty \frac{x\,\mathrm{d}x}{\sinh x} + \lim_{\varepsilon\to 0}\left\{\int_\infty^\varepsilon \frac{-(x+\pi i)\,\mathrm{d}x}{\sinh x} + \int_{-\varepsilon}^{-\infty} \frac{-(x+\pi i)\,\mathrm{d}x}{\sinh x}\right\} + (-\pi) \times \iota \times (-\pi i) = 0.$$

Comparing real parts we find

$$\int_{-\infty}^\infty \frac{x\,\mathrm{d}x}{\sinh x} = \frac{\pi^2}{2}.$$

■

Remark 272 *We have not discussed in any detail precisely what integration theory we are working within when calculating these integrals. In Prelims Analysis III, all integrals discussed* *were on closed bounded intervals. For those who go on to take Part A Integration they will see that many of the integrals such as*

$$\int_{-\infty}^{\infty} \frac{x\,\mathrm{d}x}{\sinh x}, \qquad \int_{0}^{\infty} \frac{\mathrm{d}x}{1+x^3+x^6}, \qquad \int_{0}^{\infty} \frac{\cos ax}{1+x^2}\,\mathrm{d}x$$

exist as ***proper Lebesgue integrals****. However integrals (usually those that require Jordan's Lemma to evaluate them) such as*

$$\int_{-\infty}^{\infty} \frac{\cos \pi x}{2x-1}\,\mathrm{d}x, \qquad \int_{0}^{\infty} \frac{x \sin x}{1+x^2}\,\mathrm{d}x$$

only exist as ***improper Lebesgue integrals*** *– that is the limits*

$$\lim_{R\to\infty} \int_{-R}^{R} \frac{\cos \pi x}{2x-1}\,\mathrm{d}x, \qquad \lim_{R\to\infty} \int_{0}^{R} \frac{x \sin x}{1+x^2}\,\mathrm{d}x$$

but the integrals are not proper Lebesgue integrals. This is closely related to the difference between absolutely summable and conditionally summable series. The limit

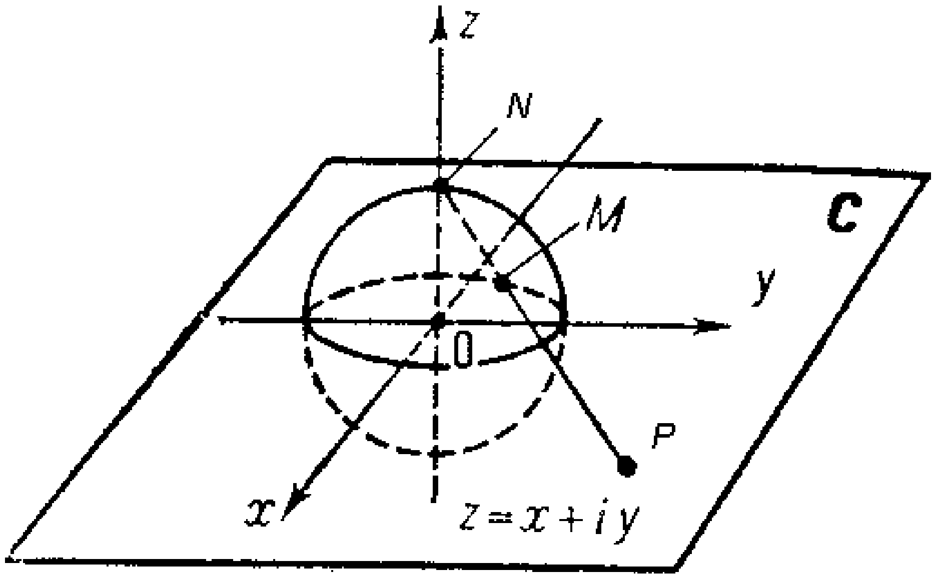
$$\lim_{R\to\infty} \int_{0}^{R} \frac{x\,|\sin x|}{1+x^2}\,\mathrm{d}x$$

does not exist and if the signed area in these improper integrals were calculated in a different order to the usual one then different answers could be determined. The values we have assigned to such an improper Lebesgue integral is called the ***principal value integral.***

14. Riemann Sphere. Möbius Transformations

Definition 273 *Let S^2 denote the unit sphere $x^2 + y^2 + z^2 = 1$ in $\mathbb{R}^3$ and let $N = (0, 0, 1)$ denote the "north pole" of S^2. Given a point $M \in S^2$, other than N, then the line connecting N and M intersects the xy-plane at a point P. If with identify the xy-plane with $\mathbb{C}$ in the natural way, then* ***stereographic projection*** *is the map*

$$\pi \colon S^2 \backslash \{N\} \to \mathbb{C} \quad \textit{defined by} \quad M \mapsto P.$$



Proposition 274 *The map π is given by*

$$\pi(a, b, c) = \frac{a + ib}{1 - c}.$$

The inverse map is given by

$$\pi^{-1}(x + iy) = \frac{(2x, 2y, x^2 + y^2 - 1)}{1 + x^2 + y^2}.$$

Proof Say $M = (a, b, c)$. Then the line connecting M and N can be written parametrically as

$$\mathbf{r}(t) = (0, 0, 1) + t\,(a, b, c - 1)\,.$$

This intersects the xy-plane when $1 + t\,(z - 1) = 0$, i.e. when $t = (1 - z)^{-1}$. Hence

$$P = \mathbf{r}\left(\frac{1}{1 - c}\right) = \left(\frac{a}{1 - c}, \frac{b}{1 - c}\right)$$

which is identified with

$$\frac{a + ib}{1 - c} \in \mathbb{C}.$$

On the other hand, if $\pi(a,b,c) = x+iy$ then

$$\frac{a+ib}{1-c} = x+iy \quad \text{and} \quad a^2+b^2+c^2 = 1.$$

Hence $(a-ib)/(1-c) = x-iy$ and so

$$x^2+y^2 = \left(\frac{a+ib}{1-c}\right)\left(\frac{a-ib}{1-c}\right) = \frac{a^2+b^2}{(1-c)^2} = \frac{1-c^2}{(1-c)^2} = \frac{1+c}{1-c} = -1+\frac{2}{1-c},$$

giving

$$\frac{2}{1+x^2+y^2} = 1-c$$

and

$$c = \frac{x^2+y^2-1}{x^2+y^2+1}.$$

Then

$$a+ib = \frac{2\,(x+iy)}{x^2+y^2+1}$$

and we may compare real and imaginary parts for the result. ■

Definition 275 *If we identify, via stereographic projection, points in $S^2\backslash\{N\}$ with points in the complex plane we see that points near N correspond to distant points in all directions. It makes a certain sense then to include a single point ∞ which is "out there" in all directions. The North Pole N then corresponds with ∞ under stereographic projection.*

Definition 276 *The* ***extended complex plane*** *is the set $\mathbb{C}\cup\{\infty\}$ and we denote this by $\tilde{\mathbb{C}}$.*

Definition 277 *When we identify S^2 with $\tilde{\mathbb{C}}$ using stereographic projection, S^2 is known as the* ***Riemann sphere****.*

Corollary 278 *If M corresponds to $z\in\tilde{\mathbb{C}}$ then the antipodal point $-M$ corresponds to $-1/\bar{z}$.*

Proof Say $M=(a,b,c)$ which corresponds to $z=(a+ib)\,/\,(1-c)$. Then $-M$ corresponds to

$$w = \frac{-a-ib}{1+c},$$

and

$$w\bar{z} = \frac{(a-ib)\,(-a-ib)}{(1-c)\,(1+c)} = \frac{-a^2-b^2}{1-c^2} = \frac{c^2-1}{1-c^2} = -1.$$

■

Theorem 279 *Circles and lines in the complex plane correspond to circles on the Riemann sphere and vice-versa.*

Remark 280 *Note that lines in $\tilde{\mathbb{C}}$ simply correspond to circles on S^2 that pass through N.*

Proof Consider the plane Π with equation $Aa + Bb + Cc = D$. This plane will intersect with S^2 in a circle if $A^2 + B^2 + C^2 > D^2$. Recall that the point corresponding to $z = x + iy$ is

$$(a, b, c) = \frac{(2x, 2y, x^2 + y^2 - 1)}{1 + x^2 + y^2}$$

which lies in the plane $Aa + Bb + Cc = D$ if and only if

$$2Ax + 2By + C\left(x^2 + y^2 - 1\right) = D\left(1 + x^2 + y^2\right).$$

This can be rewritten as

$$(C - D)\left(x^2 + y^2\right) + 2Ax + 2By + (-C - D) = 0.$$

This is the equation of a circle in $\mathbb{C}$ if $C \neq D$. The centre is $(A/(D-C), B/(D-C))$ and the radius is

$$\frac{\sqrt{A^2 + B^2 + C^2 - D^2}}{C - D}.$$

Furthermore all circles can be written in this form — we can see this by setting $C - D = 1$ and letting $A, B, C + D$ vary arbitrarily. On the other hand if $C = D$ then we have the equation

$$Ax + By = C$$

which is the equation of a line and moreover any line can be written in this form. Note that $C = D$ if and only if $N = (0, 0, 1)$ lies in the plane. So under stereographic projection lines in the complex plane correspond to circles on S^2 which pass through the north pole. ■

Definition 281 *We will use the term* ***circline*** *to denote any subset which is a circle or a line in the extended complex plane.*

Proposition 282 *Stereographic projection is conformal (i.e. angle-preserving).*

Proof Without loss of generality we can consider the angle defined by the real axis and an arbitrary line meeting it at the point $p \in \mathbb{R}$ and making an angle θ. So points on the two lines can be parametrized as

$$z = p + t, \qquad z = p + te^{i\theta},$$

where t is real. These points map onto the sphere as

$$\mathbf{r}(t) = \frac{\left(2(p+t), 0, (p+t)^2 - 1\right)}{1 + (p+t)^2}, \qquad \mathbf{s}(t) = \frac{\left(2(p + t\cos\theta), 2t\sin\theta, (p + t\cos\theta)^2 + t^2\sin^2\theta - 1\right)}{1 + (p + t\cos\theta)^2 + t^2\sin^2\theta}.$$

Then

$$\mathbf{r}'(0) = \frac{(2(p^2 - 1), 0, 4p)}{(1 + p^2)^2}, \qquad \mathbf{s}'(0) = \frac{(2(p^2 - 1)\cos\theta, 2(1 + p^2)\sin\theta, 4p\cos\theta)}{(p^2 + 1)^2}.$$

So the angle ϕ between these tangent vectors is given by

$$\begin{aligned}\cos\phi &= \frac{\left(4\left(p^2-1\right)^2\cos\theta+0+16p^2\cos\theta\right)}{\sqrt{4\left(p^2-1\right)^2+16p^2}\sqrt{4\left(p^2-1\right)^2\cos^2\theta+4\left(1+p^2\right)^2\sin^2\theta+16p^2\cos^2\theta}} \\ &= \frac{4\left(p^2+1\right)^2\cos\theta}{\left\{2\left(p^2+1\right)\right\}\left\{2\left(p^2+1\right)\right\}} \\ &= \cos\theta.\end{aligned}$$

Hence stereographic projection is conformal as required. ■

Definition 283 *A **Möbius transformation** is a map $f\colon \tilde{\mathbb{C}} \to \tilde{\mathbb{C}}$ of the form*

$$f(z) = \frac{az+b}{cz+d} \qquad \text{where } ad \neq bc.$$

We define

$$f(\infty) = \begin{cases} \frac{a}{c} & \text{if } c \neq 0, \\ \infty & \text{if } c = 0, \end{cases}$$

and also $f(-d/c) = \infty$ if $c \neq 0$.

Remark 284 *The Möbius transformations clearly have a "matrix" look to them and there are good reasons for this, though the full reason is beyond this course and more a matter for Trinity's Projective Geometry course. The extended complex plane is the **complex projective line** $P(\mathbb{C}^2)$ and the Möbius transformations are its **projective transformations** $PGL(2,\mathbb{C})$.*

The complex projective line is as follows. Suppose that we consider the equivalence classes of non-zero complex pairs $\mathbb{C}^2\backslash\{(0,0)\}$ under the equivalence relation

$$(z_1, w_1) \sim (z_2, w_2) \text{ if and only there is a non-zero } \lambda \in \mathbb{C}\backslash\{0\} \text{ such that } z_1 = \lambda w_1, z_2 = \lambda w_2.$$

Then each equivalence class has a representative of the form $(z,1)$ where $z \in \mathbb{C}$ except $(1,0)$. The former can be thought of as z and the latter as ∞.

Matrices don't act on these pairs in a well-defined way, but equivalence classes of invertible matrices do. If we identify 2×2 complex matrices under the equivalence relation

$$M_1 \sim M_2 \text{ if and only if there is a non-zero } \lambda \in \mathbb{C}\backslash\{0\}\text{, such that } M_1 = \lambda M_2$$

*then the Möbius transformations can be identified with the equivalence classes of invertible matrices which is denoted $PGL(2,\mathbb{C})$, the **P**rojective **G**eneral **L**inear group. The Möbius transformation*

$$\frac{az+b}{cz+d} \text{ is then identified with the equivalence class of } \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Proposition 285 *Möbius transformations form the group of transformations $\tilde{\mathbb{C}} \to \tilde{\mathbb{C}}$ generated (under composition) by:*

- ***translations*** *— maps of the form* $z \mapsto z + k$ *where* $k \in \mathbb{C}$*;*
- ***scalings*** *or* ***dilations*** *— maps of the form* $z \mapsto kz$ *where* $k \in \mathbb{C}\setminus\{0\}$*;*
- ***inversion*** *— the map* $z \mapsto 1/z$*. (Note this map is not an actual inversion in the sense of inverting in a circle.)*

Proof Note that if $c \neq 0$ then

$$\frac{az+b}{cz+d} = \frac{a}{c} - \frac{bc-ad}{c^2 z + dc}$$

is a composition of various translations, scalings and inversion. If $c = 0$, then $d \neq 0$ and clearly $z \mapsto (a/d)\, z + (b/d)$ is a composition of a scaling and a translation.

This shows that the set of Möbius transformations are a subset of the group generated by translations, scalings and inversion. It is also clear that translations, scalings and inversion are all special types of Möbius transformations. Further if $f(z) = (az+b)/(cz+d)$ is a Möbius transformation then

$$\begin{aligned} f(z+k) &= \frac{az+(ak+b)}{cz+(ck+d)} \text{ where } a\,(ck+d) - c\,(ak+b) = ad - bc \neq 0; \\ f(kz) &= \frac{akz+b}{ckz+d} \text{ where } (ak)\,d - b\,(ck) = k\,(ad-bc) \neq 0 \text{ if } k \neq 0; \\ f\left(\frac{1}{z}\right) &= \frac{bz+a}{dz+c} \text{ where } bc - ad \neq 0. \end{aligned}$$

Hence Möbius transformations composed with these generators yield further Möbius transformations and the result follows. ■

Proposition 286 *The Möbius transformations are bijections from* $\tilde{\mathbb{C}}$ *to* $\tilde{\mathbb{C}}$*.*

Proof

- The translation $z \mapsto z+k, \infty \mapsto \infty$ is clearly a bijection with inverse $z \mapsto z-k, \infty \mapsto \infty$.
- The scaling $z \mapsto kz, \infty \mapsto \infty$ is clearly a bijection with inverse $z \mapsto z/k, \infty \mapsto \infty$.
- And inversion $z \mapsto 1/z, 0 \mapsto \infty, \infty \mapsto 0$ is a bijection which is its own inverse.

  Hence any composition of these maps, i.e. the Möbius transformations, are all bijections.

■

Remark 287 *In fact, though we shall not prove this, the Möbius transformations correspond to the conformal bijections of* S^2*. Given any bijection* f *of* S^2 *there is a corresponding bijection* $\pi f \pi^{-1}$ *of* $\tilde{\mathbb{C}}$*. Identified this way, the conformal bijections of* S^2 *correspond to the Möbius transformations.*

Proposition 288 *Given two triples of distinct points* $z_1, z_2, z_3 \in \tilde{\mathbb{C}}$ *and* $w_1, w_2, w_3 \in \tilde{\mathbb{C}}$ *then there is a unique Möbius transformation* f *such that* $f(z_i) = w_i$ *for* $i = 1, 2, 3$*.*

Proof Note that the map

$$f(z) = \frac{(z_2 - z_3)(z - z_1)}{(z_2 - z_1)(z - z_3)}$$

is a Möbius transformation which maps z_1, z_2, z_3 to $0, 1, \infty$. There is a similar Möbius transformation g which sends w_1, w_2, w_3 to $0, 1, \infty$. By Proposition 285 $g^{-1}f$ is a Möbius transformation which maps each z_i to w_i.

To show uniqueness suppose h is another such map. Then ghf^{-1} is a Möbius transformation which maps $0, 1, \infty$ to $0, 1, \infty$. If we write

$$ghf^{-1}(z) = \frac{az + b}{cz + d}$$

then $0 \mapsto 0$ means $b = 0$, $1 \mapsto 1$ means $a + b = c + d$ and $\infty \mapsto \infty$ means $c = 0$. Hence $ghf^{-1}(z) = z$ for all z and $h = g^{-1}f$ as required. ■

Remark 289 *Recall from Proposition 133 that the circlines are the solutions sets of the equations*

$$Az\bar{z} + \bar{B}z + B\bar{z} + C = 0 \tag{14.1}$$

where $A, C \in \mathbb{R}$ and $B \in \mathbb{C}$.

Proposition 290 *The Möbius transformations map circlines to circlines.*

Proof Inversion $z \mapsto 1/z$ maps the circline (14.1) to the circline with equation

$$Cz\bar{z} + \bar{B}\bar{z} + Bz + A = 0,$$

and it is clear that translations also map circlines to circlines. Also the image of (14.1) under the scaling $z \mapsto kz$ has equation

$$Az\bar{z} + \overline{Bk}z + Bk\bar{z} + |k|^2 C = 0$$

which is another circline. Hence a Möbius transformation, which can be written as a composition of these maps, also maps circlines to circlines. ■

Example 291 *Show that the maps*

$$f(z) = \frac{e^{i\theta}(z - a)}{1 - \bar{a}z} \quad \textit{where } \theta \in \mathbb{R} \textit{ and } |a| < 1$$

maps the unit circle $|z| = 1$ to itself and a to 0.

Solution Clearly a maps to 0. Suppose now that $|z| = 1$. Then

$$|f(z)| = \left|\frac{e^{i\theta}(z - a)}{1 - \bar{a}z}\right| = \frac{|z - a|\,|\bar{z}|}{|1 - \bar{a}z|} = \frac{|1 - a\bar{z}|}{|1 - \bar{a}z|} = \left|\frac{1 - a\bar{z}}{1 - \bar{a}z}\right| = 1$$

as the numerator is the conjugate of the denominator. ■

Example 292 *Find a Möbius transformation which maps the unit circle* $|z| = 1$ *to the real axis and the real axis to the imaginary axis.*

Solution Initially it might well seem that the two diagrams — a circle divided by a line and two perpendicular lines meeting at the origin — are very different. In one, the circle meets the real axis at right angles at two points; in the second diagram this intersection is still there but "hidden" at infinity. So if we take one of these intersections, say 1, to ∞ and the other, -1, to the origin, the diagrams will look much more similar. Let's start then with

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

As required f takes -1 to 0 and 1 to ∞. Also $f(0) = -1$ and so f takes the real axis to a circline containing $0, -1, \infty$, that is the real axis (rather than the imaginary axis as desired). It takes the unit circle to a circle containing $0, \infty$ and $f(i) = (i+1)/(i-1) = -i$, that is the imaginary axis (rather than the real axis as desired). To unmuddle our images we can rotate by $\pi/2$ about the origin — hence we see the map

$$z \mapsto i\left(\frac{z+1}{z-1}\right)$$

solves the given problem. ■

15. Conformal Maps

Definition 293 *A holomorphic map $f\colon U \to \mathbb{C}$ is said to be **conformal** if $f'(z) \neq 0$ for all $z \in U$.*

Proposition 294 *A conformal map is angle-preserving and sense-preserving.*

Proof Let $f\colon U \to \mathbb{C}$ be a holomorphic map on an open set U. Let $z_0 \in U$ and let $\gamma_1\colon\ [-1,1] \to U$ and $\gamma_2\colon\ [-1,1] \to U$ be two paths which meet at $z_0 = \gamma_1(0) = \gamma_2(0)$. The original curves meet at z_0 in the (signed) angle

$$\theta = \arg \gamma_2'(0) - \arg \gamma_1'(0) = \arg \frac{\gamma_2'(0)}{\gamma_1'(0)}.$$

The images of the curves $f(\gamma_1)$ and $f(\gamma_2)$ meet at $f(z_0)$ at angle

$$\begin{aligned}
\phi &= \arg (f\gamma_2)'(0) - \arg (f\gamma_1)'(0) \\
&= \arg \frac{(f\gamma_2)'(0)}{(f\gamma_1)'(0)} \\
&= \arg \frac{f'(\gamma_2(0))\,\gamma_2'(0)}{f'(\gamma_1(0))\,\gamma_1'(0)} \qquad \text{[by the chain rule]} \\
&= \arg \frac{f'(z_0)\,\gamma_2'(0)}{f'(z_0)\,\gamma_1'(0)} \qquad [\text{as } \gamma_2(0) = \gamma_1(0) = z_0] \\
&= \arg \frac{\gamma_2'(0)}{\gamma_1'(0)} = \theta.
\end{aligned}$$

■

For various reasons, relating to areas as diverse as harmonic analysis and fluid dynamics, it is important to know what domains can be mapped to other domains by conformal bijections. Such domains would then be said to be **conformally equivalent.** We meet now our main examples of conformal maps.

Example 295 *There are three main examples of conformal maps which we shall use.*

- ***Möbius Transformations.***

  The Möbius transformations are conformal. These are maps of the form

  $$f(z) = \frac{az+b}{cz+d} \qquad (ad \neq bc)$$

  so that

  $$f'(z) = \frac{a(cz+d) - c(az+b)}{(cz+d)^2} = \frac{ad-bc}{(cz+d)^2} \neq 0.$$

  In fact, the Möbius transformations are precisely the group of conformal bijections of the Riemann Sphere. These maps are particularly useful for conformally mapping regions which are bounded by circlines.

- ***Power Maps.*** *Maps of the form*

$$z \longrightarrow z^{\alpha}$$

are conformal except when $z = 0$. These are very useful for changing the angle at the origin – for example, the map $z \longrightarrow z^2$ maps the quadrant $\{z : \operatorname{Re} z > 0, \operatorname{Im} z > 0\}$ to the upper half-plane. This is not a cheat as the origin is on the boundary of the quadrant but crucially not in it.

- ***Exponential.*** *This is particularly useful for mapping semi-infinite and infinite bars, see the example below.*

Remark 296 *Recall that:*
(a) Möbius transformations map the set of circlines (circles and lines) to the set of circlines.
(b) A circline is uniquely determined by three points.
(c) The Möbius transformations are bijections on $\tilde{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

Example 297 *Determine the image of the semi-infinite rectangles*

$$A = \{z : 0 < \operatorname{Im} z < \pi, \operatorname{Re} z < 0\}, \qquad B = \{z : 0 < \operatorname{Im} z < \pi, \operatorname{Re} z > 0\}$$

under the maps

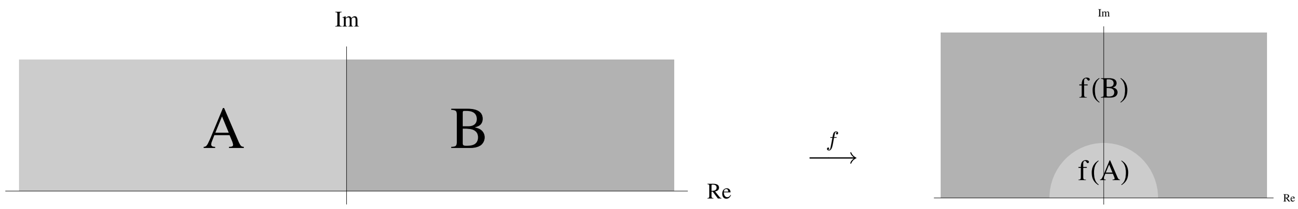
$$f(z) = \exp z \qquad \textit{and} \qquad g(z) = \exp(iz).$$

Solution An arbitrary element of A is of the form $z = x + iy$ where $y \in (0, \pi)$ and $x < 0$. So

$$f(z) = e^{x+iy} = e^x e^{iy}$$

is a general element of the half-disc $D(0,1) \cap H^+$ as $e^x < 1$ and $\operatorname{Im} e^{iy} > 0$. We likewise see that the image of B under f is $H^+ \backslash \bar{D}(0,1)$.

We can similarly note that when $y \in (0, \pi)$ and $x < 0$ then $g(z) = e^{-y}e^{ix}$ is an arbitrary complex number in the annulus $\{z : e^{-\pi} < |z| < 1\}$. Note that g is not bijective on A repeatedly wrapping A around the annulus. Again g maps B to this annulus repeatedly.



■

Remark 298 *From the Exercise sheets we know that we could use* sin *to take*

$$\{z : -\pi/2 < \operatorname{Re} z < \pi/2, \operatorname{Im} z > 0\}$$

to the upper half-plane. We could suitably adjust sin *and instead use*

$$\sin(-iz - \pi/2) \qquad \textit{or} \qquad \sin(iz + \pi/2)$$

to respectively take A and B to the upper half-plane.

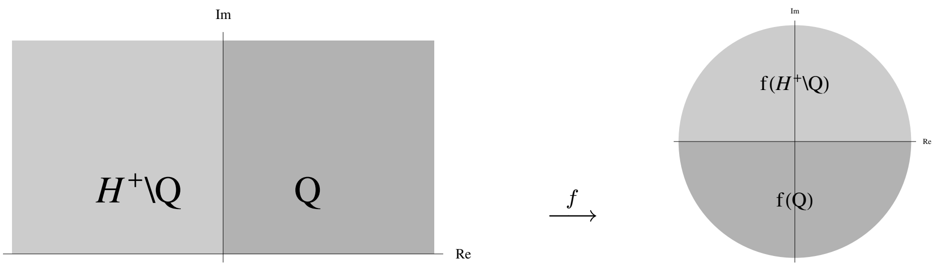
Example 299 *Determine the image of the quadrant $Q = \{z : \operatorname{Re} z > 0, \operatorname{Im} z > 0\}$ and the half-plane $H^+ = \{z : \operatorname{Re} z > 0\}$ under the Möbius transformation*

$$f(z) = \frac{z-i}{z+i}.$$

Solution Note that both Q and H^+ are bounded by circlines (lines in fact) and so their images will be bounded by circlines also. Note further that

$$f(0) = -1, \quad f(1) = -i, \quad f(\infty) = 1, \quad f(i) = 0, \quad f(1+i) = \frac{1-2i}{5}, \quad f(-1+i) = \frac{-1-2i}{5}.$$

Hence the positive imaginary axis maps to the interval $(-1, 1)$ and the positive real axis maps to the lower unit semicircle $\gamma^-(0, 1)$ (the circline connecting -1, $-i$ and 1). As $1+i$ maps into the lower unit semi-disc that then that is $f(Q)$. The image of $H^+ \backslash Q$ makes the remainder of the unit disc as shown in the diagram below.



■

Remark 300 *It is in fact quite easy to see why the map f above takes the upper half-plane onto the unit disc. The upper half-plane can be characterized as those points closer to i than to $-i$. Hence*

$$\operatorname{Im} z > 0 \iff |z-i| < |z+i| \iff \left|\frac{z-i}{z+i}\right| < 1.$$

We can see that other maps $(z-a)/(z-\bar{a})$ would also work for any $a \in H^+$ and that more generally any half-plane can be mapped to the unit disc by a map $(z-\alpha)/(z-\beta)$ where α is a point in the half-plane and β is the mirror image of α across the bounding line.

Example 301 *Find a conformal map which maps the half-disc*

$$U = \{z \in \mathbb{C} : |z| < 1, \operatorname{Im} z > 0\}$$

bijectively onto $D(0, 1)$. Explain why it is not possible to do this using only Möbius transformations.

Remark 302 ***An important tip.*** *If faced with a domain U bounded by two circline arcs or segments which meet at α and β then the best thing to do first, almost without any further consideration, is to apply the map $(z-\alpha)/(z-\beta)$. This map sends α to 0 and β to ∞ and so takes the two circline arcs or segments to half-lines meeting at the origin.*

Solution In light of the above remark, the domain U is best thought of as being bounded by two circlines, namely $\gamma(0,1)$ and the real axis, which meet at ± 1, and so we will begin with

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

As

$$f(1) = 0, \qquad f(-1) = \infty, \qquad f(i) = i, \qquad f(0) = -1, \qquad f(i/2) = \frac{i-2}{i+2} = \frac{-3+4i}{5}$$

then $\gamma(0,1)$ maps to the imaginary axis and the real axis maps to the real axis. Given the position of $f(i/2)$ then we know $f(U)$ to be the second quadrant. As Mobius maps are bijections on the extended complex plane then f is a bijection onto its image.

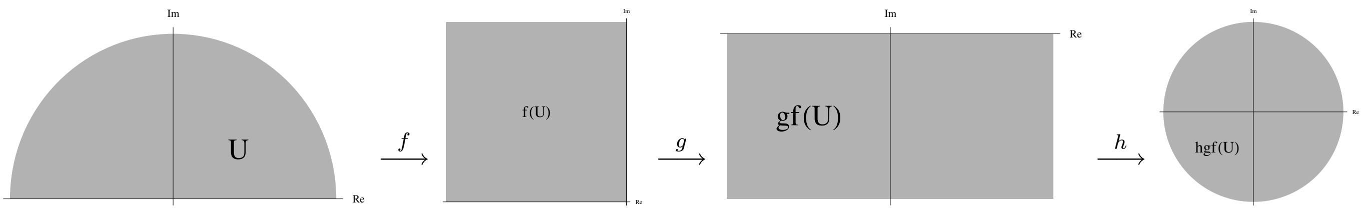
The map $g(z) = z^2$ is conformal (except at the origin) and bijective on $f(U)$ as it includes no point z and its negative $-z$. We have $g(f(U))$ is the lower half-plane.

Finally the lower half-plane can be identified as precisely those points closer to $-i$ than i (mathematically $|z+i| < |z-i|$) and hence

$$h(z) = \frac{z+i}{z-i}$$

maps the lower half-plane bijectively onto the unit disc. A map that then meets the requirements of the question is

$$\begin{aligned} h(g(f(z))) &= \frac{\left(\frac{z-1}{z+1}\right)^2 + i}{\left(\frac{z-1}{z+1}\right)^2 - i} \\ &= \frac{(z-1)^2 + i(z+1)^2}{(z-1)^2 - i(z+1)^2} \\ &= \frac{(1+i)(z^2+1) + 2z(i-1)}{(1-i)(z^2+1) - 2z(1+i)} \\ &= i\left(\frac{z^2+2iz+1}{z^2-2iz+1}\right). \end{aligned}$$



That there is no Möbius transformation which performs this is clear because any such map would keep the two circlines bounding U as two distinct circlines, whereas $D(0,1)$ is bounded by just one circline. ■

Example 303 *Find a conformal bijection mapping $U = \{z \in \mathbb{C} : |z| < 1, \operatorname{Im} z > 0\}$ to $D(0,1)$ and such that $i/2$ maps to* 0.

Solution In our previous solution, we see that $hgf(i/2) = -i/7$. However we don't need to reinvent the wheel here as we know that for any a in the lower half-plane then the map $(z-a)/(z-\bar{a})$ will suffice as a replacement for h. So note

$$gf(i/2) = \left(\frac{i/2-1}{i/2+1}\right)^2 = \left(\frac{-2+i}{2+i}\right)^2 = \left(\frac{-3+4i}{5}\right)^2 = \frac{-7-24i}{25}.$$

If we then instead use

$$\tilde{h}(z) = \frac{25z + (7+24i)}{25z + (7-24i)}$$

then $\tilde{h}gf$ will do the job. ■

Example 304 *Show that $D(0,1)$ and $\mathbb{C}$ are not conformally equivalent.*

Solution Let $f\colon \mathbb{C} \to D(0,1)$ be a holomorphic map. In particular it is bounded and so, by Liouville's Theorem, f is constant and so certainly not a bijection. ■

One might ask what domains are conformally equivalent to $D(0,1)$. As conformal equivalence is a particular type of topological isomorphism (the technical term for a topological isomorphism being *homeomorphism*) then any such domain would have to be simply-connected. But there is a wide variety of such domains, many rather nasty boundaries, and so the following result may come as striking.

Theorem 305 ***(Riemann Mapping Theorem)*** *Let U be a simply-connected domain with $U \neq \mathbb{C}$. Then U is conformally equivalent to $D(0,1)$. In the case the the boundary of U is smooth then the conformal equivalence can be extended between $U \cup \partial U$ and $D(0,1)$.*

Proof The proof of this is well beyond this course. Interested readers might try Conway *Functions of One Complex Variable, §7.4.* ■

An important use of conformal maps is their use in harmonic analysis and begins with the following result. A conformal equivalence between domains U and V induces an isomorphism between the space of harmonic functions on each domain. So if we are seeking (for example) to solve Laplace's Equation in a simply-connected proper domain of $\mathbb{C}$ with a smooth boundary then the Riemann Mapping Theorem tells us (in principle at least) that this is no harder than solving the problem on the unit disc or upper half-plane.

Proposition 306 *Let U and V be open subsets of $\mathbb{C}$, let $f\colon U \to V$ be holomorphic and let $\phi\colon V \to \mathbb{R}$ be harmonic. Then $\phi \circ f$ is harmonic.*

Proof Let $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ and we will also write $\psi = \phi \circ f$. If $z = x+iy$ then we can write ψ as

$$\psi(x+iy) = \phi(u(x,y), v(x,y)).$$

By the chain rule we have

$$\psi_x = \phi_u u_x + \phi_v v_x, \qquad \psi_y = \phi_u u_y + \phi_v v_y$$

and applying it again we get

$$\begin{aligned}
\psi_{xx} &= (\phi_{uu}u_x + \phi_{uv}v_x)\,u_x + \phi_u u_{xx} + (\phi_{vu}u_x + \phi_{vv}v_x)\,v_x + \phi_v v_{xx} \\
&= \phi_u u_{xx} + \phi_v v_{xx} + \phi_{uu}u_x^2 + 2\phi_{uv}u_x v_x + \phi_{vv}v_x^2 \\
\psi_{yy} &= (\phi_{uu}u_y + \phi_{uv}v_y)\,u_y + \phi_u u_{yy} + (\phi_{vu}u_y + \phi_{vv}v_y)\,v_y + \phi_v v_{yy} \\
&= \phi_u u_{yy} + \phi_v v_{yy} + \phi_{uu}u_y^2 + 2\phi_{uv}u_y v_y + \phi_{vv}v_y^2
\end{aligned}$$

Remember that as ϕ is harmonic then $\phi_{xx} + \phi_{yy} = 0$. Also, by the Cauchy-Riemann equations, we have

$$u_x = v_y, \quad u_y = -v_x, \quad u_{xx} + u_{yy} = 0, \quad v_{xx} + v_{yy} = 0.$$

Hence

$$\begin{aligned}
\psi_{xx} + \psi_{yy} &= \phi_u(u_{xx} + u_{yy}) + \phi_v(v_{xx} + v_{yy}) + \phi_{uu}(u_x^2 + u_y^2) + 2\phi_{uv}(u_x v_x + u_y v_y) + \phi_{vv}(v_x^2 + v_y^2) \\
&= \phi_{uu}(u_x^2 + u_y^2) + \phi_{vv}(v_x^2 + v_y^2) \qquad \text{[as consequences of the CREqs].}
\end{aligned}$$

Now, also by the Cauchy-Riemann equations, we have

$$|f'(z)|^2 = u_x^2 + u_y^2 = v_x^2 + v_y^2.$$

So

$$\psi_{xx} + \psi_{yy} = |f'(z)|^2\,(\phi_{uu} + \phi_{vv}) = 0$$

as ϕ is harmonic. Hence ψ is harmonic also. ■

Remark 307 *You may recall from the first year meeting parabolic coordinates*

$$u = \frac{x^2 - y^2}{2}, \qquad v = xy.$$

If $g(u,v) = f(x,y)$ then it is the case that $\nabla^2 f = 0$ if and only if $\nabla^2 g = 0$. This does not seem surprising now in light of the above result with $f = u + iv = z^2/2$.

Definition 308 *Given an open subset U of $\mathbb{C}$, we will denote by $H(U)$ the vector space of harmonic functions on U.*

Corollary 309 *Let U and V be conformally equivalent open subsets of $\mathbb{C}$. Then $H(U)$ and $H(V)$ are isomorphic as vector spaces.*

Proof Let $f\colon U \to V$ be a conformal equivalence. This then induces map

$$\begin{aligned}
H(V) &\to H(U) \text{ given by } \phi \mapsto \phi \circ f, \\
H(U) &\to H(V) \text{ given by } \phi \mapsto \phi \circ f^{-1},
\end{aligned}$$

which are clearly linear (in ϕ) and inverses of one another. ■

This introduces powerful techniques for handling the **Dirichlet Problem** (after the German mathematician Lejeune Dirichlet 1805-1859). The Dirichlet Problem is typically stated as follows:

- Given a domain U and a function g defined on the boundary ∂U, is there a unique function f such that $f = g$ on ∂U and $\nabla^2 f = 0$ in U.

The Dirichlet problem was rigorously demonstrated by Hilbert in 1900 for suitably nice boundaries (after Weierstrass has shown Dirichlet's own proof lacked rigour). The problem has the following solution in the upper half-plane.

Proposition 310 ***(Dirichlet Problem in the Half-Plane)*** *(Off syllabus) The function*

$$u(x,y) = \frac{y}{\pi}\int_{-\infty}^{\infty} \frac{f(t)}{(t-x)^2 + y^2}\,\mathrm{d}t$$

is harmonic in the upper half-plane and satisfies $u(x,0_+) = f(x)$.

Proof See Priestley 23.14 and 23.16. ■

Example 311 *Use the above proposition to solve Dirichlet problem in the half-plane when*

$$f(t) = \begin{cases} 1 & t < 0 \\ 0 & t > 0 \end{cases}$$

Solution We have, for $y > 0$

$$\begin{aligned} u(x,y) &= \frac{y}{\pi}\int_{-\infty}^{0} \frac{\mathrm{d}t}{(t-x)^2 + y^2} \\ &= \frac{y}{\pi}\int_{-\infty}^{-x} \frac{\mathrm{d}t}{t^2 + y^2} \\ &= \frac{1}{\pi}\left[\tan^{-1}\left(\frac{t}{y}\right)\right]_{-\infty}^{-x} \\ &= \frac{1}{\pi}\left(-\tan^{-1}\frac{x}{y} + \frac{\pi}{2}\right) \\ &= \frac{1}{\pi}\tan^{-1}\frac{y}{x} \\ &= \frac{\arg(x+iy)}{\pi} \end{aligned}$$

with $\arg(x+iy)$ taken in the range $(0,\pi)$. ■

Example 312 *By means of conformal maps, solve the Dirichlet problem in:*
(a) the quadrant $\{z : \operatorname{Re} z > 0, \operatorname{Im} z > 0\}$ *subject to the boundary conditions*

$$u(x,0) = 0 \quad (x > 0), \qquad u(0,y) = 1 \quad (y > 0).$$

(b) the half-disc $\{z : \operatorname{Im} z > 0, |z| < 1\}$ *subject to the boundary conditions*

$$u(z) = 0 \quad (|z| = 1), \qquad u(x,0) = 1 \quad (|x| < 1),$$

Solution (a) We know $z \to z^2$ is a conformal equivalence between the quadrant and the upper half-plane which matches up the boundary conditions correctly. So, by Proposition 306, we see that we should assign to the value z (in the quadrant) the value of the solution at z^2 (in the half-plane). Hence our solution in the quadrant is

$$u(x,y) = \frac{\arg\left((x+iy)^2\right)}{\pi} = \frac{2\arg(x+iy)}{\pi}.$$

(b) We saw in Example 301 that the upper half-disc is mapped to the upper half-plane by

$$w(z) = -\left(\frac{z-1}{z+1}\right)^2,$$

and that this map takes the bottom of the half-disc to the negative real axis and the semicircular arc to the positive real axis. So, again by Proposition 306, we need to assign to z the solution at w. Hence our solution is

$$u(z) = \frac{1}{\pi}\arg\left[-\left(\frac{z-1}{z+1}\right)^2\right]$$

with arg taking values in the range $(0,\pi)$. ■

Example 313 *(Off-syllabus)* ***(Fluid Flow in the Plane)*** *Let $\omega = \phi + i\psi$ be a holomorphic function on a domain U. Such a function can be considered as a **complex potential** of a steady, irrotatonal, incompressible, inviscid fluid with **velocity potential** ϕ and **streamfunction** ψ. The velocity of the flow is then*

$$\mathbf{u} = (\phi_x, \phi_y).$$

That the flow is incompressible means $\operatorname{div}\mathbf{u} = 0$*; this is true here as*

$$\operatorname{div}\mathbf{u} = \phi_{xx} + \phi_{yy} = 0$$

and we know $\phi = \operatorname{Re}\omega$ to be harmonic. That the flow is irrotational means $\operatorname{curl}\mathbf{u} = \mathbf{0}$ *which holds true as*

$$\operatorname{curl}\mathbf{u} = \left(\phi_{yx} - \phi_{xy}\right)\mathbf{k} = \mathbf{0}.$$

*The **streamlines** of the flow are the curves $\psi = c$ (a constant).*

Example 314 *An example of a uniform flow in half-plane $H = \{z : \operatorname{Im} z > 0\}$ is $\mathbf{u} = (U, 0)$ where U is a (real) constant, and we can see that this corresponds to the complex potential*

$$\omega = Uz = Ux + iUy.$$

The streamlines are (unsurprisingly) the curves $y = c$ (a positive constant).

We know from the exercise sheets that sine is a conformal equivalence between

$$U = \{z : \operatorname{Im} z > 0, -\pi/2 < \operatorname{Re} z < \pi/2\}$$

and H. Any complex potential $\omega(z)$ on H corresponds to a complex potential $\omega(\sin z)$ on U (and vice versa). Recall that

$$\sin(x+iy) = \sin x \cosh y + \cos x \sinh y.$$

The above uniform flow in H corresponds to a flow in U with streamlines

$$\cos x \sinh y = c \qquad \textit{(where } c > 0\textit{).}$$

These streamlines are sketched below.