Riemann Surfaces

Henry Wilton*

Michaelmas 2020

Contents

*Comments and corrections are always welcome. Please send them to h.wilton@maths.cam.ac.uk.

Lecture 1: Analytic and meromorphic functions

This course picks up where *IB Complex Analysis* leaves off. The IB course contains many classical results about complex-differentiable functions of one variable. However, it also raises some natural questions, such as:

> 'What do we really mean by multi-valued functions, such as the complex logarithm or mth roots?'

The reader may like to keep this question in mind as this course begins. We will see that these questions naturally lead us to define and study a whole new class of mathematical objects, the *Riemann surfaces* of the course title. As we shall see, Riemann surfaces exhibit a beautiful interplay between analysis and geometry.

Since the course leans heavily on some of the results of *IB Complex Analysis*, we will start by recalling some of the definitions and results from that course.

1.1 Analytic functions and their zeroes

The functions we study will be defined on domains. A *domain* is an open, connected subset of the complex plane $\mathbb{C}$. Two of the most important kinds of domains are the *(open) disc*

$$D(z_0, r) := \{z \in \mathbb{C} \mid |z - z_0| < r\}$$

and the *(open) punctured disc*

$$D_*(z_0, r) := \{z \in \mathbb{C} \mid 0 < |z - z_0| < r\}\,.$$

More generally, a *neighbourhood* U of a point z_0 in a domain D is any subdomain $U \subseteq D$ containing z_0, and a *punctured neighbourhood of* z_0 is any open set of the form $U \smallsetminus \{z_0\}$, where U is a neighbourhood of z_0.

Definition 1.1. Let $D \subseteq \mathbb{C}$ be a domain. A function $f : D \to \mathbb{C}$ is called *holomorphic* or *analytic* if either of the following two equivalent definitions are satisfied:

(i) f is $\mathbb{C}$-differentiable at every $z_0 \in D$; or

(ii) for any $z_0 \in D$, there is $r > 0$ such that f has a power-series expansion

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

for any $z \in D(z_0, r) \subseteq D$.

More precisely, Definition 1.1(i) is the definition of a holomorphic function, and Definition 1.1(ii) is the definition of an analytic function. It is a theorem of *IB Complex Analysis* that the two definitions coincide. As a consequence, analytic functions are surprisingly rigid: many behaviours that we are used to from real analysis are impossible in the complex setting.

Proposition 1.2 (Principle of isolated zeroes)**.** *Let $f : D \to \mathbb{C}$ be an analytic function on a domain $D \subseteq \mathbb{C}$. If $f(z_0) = 0$, then either f is identically zero in a neighbourhood of z_0, or f is non-zero on a punctured neighbourhood of z_0.*

Proof. Unless $f \equiv 0$ in a neighbourhood of z_0, there is a minimal $m \geq 0$ such that $a_m \neq 0$. Hence,

$$f(z) = (z - z_0)^m g(z)$$

for some analytic function g, defined on an open disc $D(z_0, \rho)$, with $g(z_0) \neq 0$. By continuity of g, there is $r > 0$ with $g(z) \neq 0$ for all $z \in D(z_0, r)$, and the result follows. □

It follows that, if two analytic functions on a domain agree fairly often, then they are equal. Here, 'fairly often' means, precisely, on a non-discrete subset. A subset A of $\mathbb{C}$ is called *discrete* if the subspace topology induced on A is the discrete topology. Equivalently, A is discrete unless some $a \in A$ is a limit of a sequence in its complement $A \setminus \{a\}$.

Corollary 1.3 (Identity principle)**.** *Let f, g be analytic functions defined on a domain D in $\mathbb{C}$. Unless the set*

$$\{z \in D \mid f(z) = g(z)\}$$

is discrete, $f \equiv g$ on D.

Proof. The proof is an easy application of the fact that D is connected. Let A be the set of points $z \in D$ on which f and g agree in a punctured neighbourhood of z: that is,

$$A := \{z \in U \mid \exists r > 0, \ \forall z \in D(z, r), f(z) = g(z)\}.$$

Likewise, let D be the set of points on which f and g disagree in some punctured neighbourhood: that is,

$$B := \{z \in D \mid \exists r > 0, \ \forall z \in D_*(z, r), f(z) \neq g(z)\}.$$

The sets A and B are open and disjoint by definition. More surprisingly, the principle of isolated zeros applied to the analytic function $f - g$ implies that A and B together cover D. Therefore, either A or B must be empty, because D is connected. If f and g agree on a non-discrete subset, then it follows that B does not cover D, so $A = D$ and the result follows. □

1.2 Meromorphic functions and singularities

Singularities arise when analytic functions are defined on punctured discs.

Definition 1.4. An analytic function $f : D_*(z_0, r) \to \mathbb{C}$ is said to have an *isolated singularity* at z_0.

Just as holomorphic functions have Taylor series, so you saw in *IB Complex Analysis* that analytic functions have Laurent series at their singularities.

Proposition 1.5 (Laurent series)**.** *If an analytic function f has an isolated singularity at z_0, then*

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$$

on $D_*(z_0, r)$ for some $r > 0$.

The coefficients a_n lead to a natural classification of singularities.

Definition 1.6 (Classification of singularities)**.** Suppose that f has an isolated singularity at z_0, so

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$$

on a punctured neighbourhood of z_0.

(i) If $a_n = 0$ for $n < 0$ then z_0 is a *removable singularity*. In this case, f can be extended to an analytic function $g(z)$ defined on a neighbourhood of z_0.

(ii) If there is $m > 0$ such that $a_m = 0$ for all $n < -m$ but $a_{-m} \neq 0$, then f is said to have a *pole of order* m at z_0. In this case,

$$f(z) = (z - z_0)^{-m} g(z)$$

on a neighbourhood of z_0, for some analytic function g with $g(z_0) \neq 0$.

(iii) Otherwise (that is, if $a_n \neq 0$ for infinitely many $n < 0$), then f is said to have an *essential singularity* at z_0.

Fortunately, we don't need to classify singularities by calculating Laurent series. The following two important results are also proved in *IB Complex Analysis*.

Theorem 1.7 (Removable singularities)**.** *An analytic function f has a removable singularity at z_0 if and only if f is bounded on some punctured disc $D_*(z_0, r)$.*

The corresponding theorem for essential singularities plays a very important role in this course.

Theorem 1.8 (Casorati–Weierstrass)**.** *An analytic function f on a domain D has an essential singularity at z_0 if and only if $f(D_*(z_0, r))$ is dense in $\mathbb{C}$, for any $r > 0$ such that $D_*(z_0, r) \subseteq D$.*

Proof. For the 'only if' direction, notice that $f(D_*(z_0, r))$ is not dense if z_0 is removable or a pole. Indeed, if z_0 is removable, then some $f(D_*(z_0, r))$ is bounded, and in particular not dense, by Theorem 1.7. If z_0 is a pole of order m, then

$$f(z) = (z - z_0)^{-m} g(z)$$

where g is analytic and non-zero in a neighbourhood of z_0. There is an $\epsilon > 0$ so that

$$|g(z)| \geq \epsilon > 0$$

on some punctured disc $D_*(z_0, r)$, and therefore, on that disc,

$$|f(z)| \geq \frac{|g(z)|}{|z - z_0|^m} > \frac{\epsilon}{r^m} .$$

So f is bounded away from 0 on $D_*(z_0, r)$, and not dense.

For the other direction, if $f(D_*(z_0, r))$ is not dense, then there is some open disc $D(w_0, \epsilon)$ disjoint from $f(D_*(z_0, r))$. Consider the function

$$h(z) = \frac{1}{f(z) - w_0}$$

defined on $D_*(z_0, r)$. Since $|f(z) - w_0| \geq \epsilon$ whenever $z \in D_*(z_0, r)$, it follows that $|h(z)| \leq 1/\epsilon$. By Theorem 1.7, h has a removable singularity at z_0, and can be extended across z_0. Writing

$$f(z) = \frac{1}{h(z)} + w_0 ,$$

it follows that f has a removable singularity at z_0, unless $h(z_0) = 0$, in which case f has a pole. □

Poles are almost as nice a removable singularities, so it is useful to have terminology for functions with poles.

Definition 1.9. Let D be a domain in $\mathbb{C}$. If there is a discrete subset A of D, and f is a holomorphic function on $D \setminus A$ with poles at the points of A, then f is said to be a *meromorphic function on D*.

We close with an example.

Example 1.10. Let

$$f(z) = \frac{1}{e^{1/z} - 1},$$

which is analytic on the complement of $A = \{0\} \cup \{1/2\pi i n \mid n \in \mathbb{Z} \setminus \{0\}\}$. From the power series expansion of the exponential function, we see that

$$g(w) = e^w - 1$$

has a simple zero at $w = 1$, and hence, by periodocity, at every point of $2\pi i\mathbb{Z}$. Therefore, $h(z) = e^{1/z} - 1$ has a simple zero at $1/2\pi i n$, for each non-zero integer n, and so f has simple poles at those points.

However, 0 must be an essential singularity of f. Indeed, $h(z) = 0$ at a sequence of points tending to 0 so, by the principle of isolated zeroes, if 0 were a removable singularity then h would be 0 in a neighbourhood of 0, which is absurd. Furthermore, if h has a pole at 0, then $|h(z)| > 0$ in a neighbourhood of 0, which is also a contradiction. Therefore, h has an essential singularity at 0, and f does too.

1.3 Analytic continuation

The identity principle leads to a very surprising rigidity for analytic functions. Given analytic f defined on a small open disc $D_1 = D(z_1, r_1)$, and another small open disc D_2 that intersects D_1, there is at most one way to extend f analytically across D_2! This move from one disc to a neighbouring one is called *analytic continuation.*

Definition 1.11. Let D be a domain in $\mathbb{C}$. A *function element* on D is a pair (f, U), where U is a subdomain in D and f is an analytic function on U. If (g, V) is another function element on D, then

$$(f, U) \sim (g, V)$$

means that $U \cap V \neq \varnothing$ and $f|_{U\cap V} = g|_{U\cap V}$. In this case, (g, V) is said to be a *direct analytic continuation* of (f, U). If there is a finite sequence of direct analytic continuations

$$(f, U) = (f_1, U_1) \sim \cdots \sim (f_{n-1}, U_{n-1}) \sim (f_n, U_n) = (g, V)$$

then (g, V) is said to be an *analytic continuation of* (f, U), and we write $(f, U) \approx (g, V)$.

Note that $\approx$ is an equivalence relation.

Definition 1.12. A $\approx$-equivalence class $\mathcal{F}$ of function elements on a domain D is called a *complete analytic function* on D.

1.4 The complex logarithm

If $(f, U) \sim (g, V)$ then, by the identity principle, g is determined by f. It is tempting to think that this observation extends from direct analytic continuations to all analytic continuations. However, that is not quite correct, as we shall illustrate with one of the most famous motivating examples in complex analysis. The notation $\mathbb{C}_*$ denotes the punctured complex plane $\mathbb{C} \setminus \{0\}$.

Example 1.13 (The complex logarithm). The complex logarithm log arises from attempting to invert the exponential function $\exp : \mathbb{C} \to \mathbb{C}_*$. This isn't possible globally, because exp is not injective: for instance,

$$\exp(0) = 1 = \exp(2\pi i)\,.$$

Therefore, to define log as a function, we need to restrict attention to a subdomain of $\mathbb{C}$ on which exp is injective. In the language of *IB Complex Analysis*, we could do this explicitly using *branch cuts*. In the language of this course, any definition of log on a subdomain U of $\mathbb{C}$ defines a function element. In total, although log isn't strictly speaking a function, this collection of function elements defines log as a complete analytic function.

To be concrete, let's exhibit some explicit function elements that enable us to make sense of all possible choices of log, everywhere on $\mathbb{C}_*$.

Given any $(\alpha, \beta) \subseteq \mathbb{R}$ with $|\alpha - \beta| < 2\pi$, define

$$U_{(\alpha,\beta)} = \{re^{i\theta} \mid r > 0,\ \alpha < \theta < \beta\}\,.$$

For instance, $U_{(0,2\pi)} = \mathbb{C} \setminus \mathbb{R}_{\geq 0}$. For $z \in U_{(\alpha,\beta)}$, which we think of as $z = re^{i\theta}$ with $\theta \in (\alpha, \beta)$, define

$$f_{(\alpha,\beta)}(z) = \log r + i\theta\,,$$

where log is the real logarithm. It is not hard to check that $f_{(\alpha,\beta)}$ is analytic: for instance, noting that the real part is independent of θ and the imaginary part is independent of r, the Cauchy–Riemann equations in polar coordinates reduce to

$$\frac{\partial u}{\partial r} = \frac{1}{r}\frac{\partial v}{\partial \theta}$$

which is true in this case. Therefore, together, these define function elements $F_{(\alpha,\beta)} = (f_{(\alpha,\beta)}, U_{(\alpha,\beta)})$.

To understand the complex logarithm, we just need to choose a collection of open intervals that cover $\mathbb{R}$. For instance, we can take the intervals

$$I(n) = ((n-1)\pi/2, (n+1)\pi/2)$$

across all $n \in \mathbb{Z}$. Next, we need to analyse what happens when these intervals overlap, to determine when $F_{I(m)} \sim F_{I(n)}$.

There are four cases to consider, depending on $m - n$ modulo 4.

(i) If $m \equiv n$ modulo 4 the $U_{I(m)} = U_{I(n)}$, but $I(m)$ and $I(n)$ are disjoint unless $m = n$. Therefore, $F_{I(m)} \sim F_{I(n)}$ if and only if $m = n$.

(ii) If $m \equiv n + 1$ modulo 4 then $U_{I(m)} \cap U_{I(n)}$ is a quadrant of $\mathbb{C}$, but $I(m)$ and $I(n)$ are disjoint unless $m = n + 1$. In this last case, $f_{I(m)}$ and $f_{I(n)}$ agree on $U_{I(m)} \cap U_{I(n)}$. Therefore, $F_{I(m)} \sim F_{I(n)}$ if and only if $m = n + 1$.

(iii) If $m \equiv n+2$ modulo 4 then $U_{I(m)}$ and $U_{I(n)}$ are disjoint, so $F_{I(m)} \nsim F_{I(n)}$.

(iv) If $m \equiv n + 3$ modulo 4 then, as in the case (ii), $F_{I(m)} \sim F_{I(n)}$ if and only if $m = n - 1$.

In conclusion, $F_{I(m)} \sim F_{I(n)}$ if and only if $|m - n| \le 1$.

Applying these direct analytic continuations iteratively, all of the function elements $F_{I(n)}$ are in the same $\approx$-equivalence class, so they all define the same complete analytic function. This is the complex logarithm.

This example shows us that analytic continuation may not be unique. Indeed, in the notation of Example 1.13, $U_{I(0)} = U_{I(4)}$, but $f_{I(0)}(1) = 0$ whereas $f_{I(4)}(1) = 2\pi i$. This is of course the familiar phenomenon that, as we follow a clockwise loop around the origin, the value of the logarithm changes by $2\pi i$. To make this precise, we introduce the idea of analytically continuing along a curve.

Definition 1.14 (Analytic continuation along a path)**.** Let (f, U) be a function element in the domain D, and consider an analytic continuation $(f, U) \approx (g, V)$, exhibited by a sequence of direct analytic continuations as in the definition.

$$(f, U) = (f_1, U_1) \sim \ldots \sim (f_{n-1}, U_{n-1}) \sim (f_n, U_n) = (g, V);$$

Let $\gamma : [0,1] \to D$ be a continuous path. If there is a dissection

$$0 = t_0 < t_1 < \ldots < t_{n-1} < t_n = 1$$

such that $\gamma([t_{i-1}, t_i]) \subseteq U_i$ for each $1 \le i \le n$, then (g,V) is an *analytic continuation of* (f,U) *along* γ, and we write $(f,U) \approx_\gamma (g,V)$.

Remark 1.15. If $(f,U) \approx (g,V)$ then $(f,U) \approx_\gamma (g,V)$ for some path γ.

Although we saw above that analytic continuations are not always unique, it turns out that they are determined the path γ: if $(f,U) \approx_\gamma (g,V)$, then $g(\gamma(1))$ only depends on (f,U) and γ. This can be proved directly, although it also follows easily from later results in the course. Indeed, in due course we shall see that much more is true.

Lecture 2: Natural boundary, gluing constructions and roots

2.1 Natural boundary

In the last lecture we saw that, when we try to analytically continue, we may build very large complete analytic functions, taking infinitely many values at any point. On the other hand, sometimes we simply cannot analytically continue very far. This phenomenon is the subject of the next subsection.

Let's introduce some new notation for some of our favourite subsets of $\mathbb{C}$: $\mathbb{D}$ is the open unit disc $D(0,1)$, and $\mathbb{T}$ is its boundary, the unit circle.

Consider a power series

$$f(z) = \sum_{n \ge 0} a_n z^n$$

with radius of convergence 1 (without loss of generality). In particular, the series converges absolutely and uniformly on any closed disc $\overline{D} \subseteq \mathbb{D}$.

Definition 2.1. A point $z_0 \in \mathbb{T}$ is called *regular for* f if there is an open neighbourhood U of z_0 and an analytic function g on U such that $g \equiv f$ on $U \cap \mathbb{D}$. Otherwise, z_0 is called *singular for* f.

The point is, of course, that we may analytically continue to the regular points, but not to the singular points.

Remark 2.2. The set of regular points in ∂T is open by definition; hence, the set of singular points is closed.

Beware! It is not the case that z_0 is a regular point if and only if $\sum_{n\geq 0} a_n z_0^n$ converges; indeed, neither implication is true.

Example 2.3. Consider the power series

$$f(z) = \frac{1}{1-z} = \sum_{n\geq 0} z^n ;$$

evidently, every point of $\mathbb{T} \setminus \{1\}$ is regular. However, the power series does not converge at $z_0 = -1$.

Example 2.4. Consider the power series

$$g(z) = \sum_{n\geq 2} \frac{z^n}{n(n-1)} .$$

The series

$$g(1) = \sum_{n\geq 2} \frac{1}{n(n-1)}$$

is convergent, but if 1 were a regular point for g, then it would follow that it is also regular for $g'' = f$, which is absurd since 1 is a pole of f.

There is always at least one singular point.

Proposition 2.5. *If a power series*

$$f(z) = \sum_{n\geq 0} a_n z^n$$

has radius of convergence 1, then some point of $\mathbb{T}$ *is singular for* f.

Proof. If not, for each $z \in \mathbb{T}$ there is $\epsilon_z > 0$ such that f extends analytically over $D(z, \epsilon_z)$. Finitely many of these discs cover $\mathbb{T}$ by compactness, so f extends analytically over some $D(0, 1+\delta)$ with $\delta > 0$. But this implies that the radius of convergence is at least $1+\delta$, which is a contradiction. □

The concept of natural boundary refers to the extreme case where the number of singular points is as large as possible.

Definition 2.6. If every point of $\mathbb{T}$ is singular, then $\mathbb{T}$ is said to be a *natural boundary* for f.

This can happen in practice, as the following example shows.

Example 2.7. Let

$$f(z) = \sum_{n \geq 0} z^{n!} ,$$

which has radius of convergence 1 by the ratio test, and consider a qth root of unity $\omega = \exp(2\pi i p/q)$. We will show that ω is singular, and conclude that f has natural boundary $\mathbb{T}$ by Remark 2.2, since roots of unity are dense in $\mathbb{T}$.

Whenever $0 < r < 1$, we have

$$f(r\omega) = \sum_{n=0}^{q-1} r^{n!}\omega^{n!} + \sum_{n \geq q} r^{n!} .$$

The second term diverges to ∞ as $r \to 1$: indeed,

$$\lim_{r \to 1} \sum_{n=q}^{q+M} r^{n!} = M + 1 ,$$

for any real number M, so for r large enough,

$$\sum_{n \geq q} r^{n!} > \sum_{n=q}^{q+M} r^{n!} > M .$$

However, if ω were a regular value then $f(r\omega)$ would converge to $f(\omega)$ as $r \to 1$, so ω must be singular, as claimed.

There are easy extensions of the notion of natural boundary to other curves in the plane; for instance, it is not difficult to adapt Example 2.7 to define a function with natural boundary equal to $\mathbb{R}$. However, in general, the natural boundary of an analytic function can be extremely complicated, and a complete treatment is beyond the scope of this course.

2.2 A gluing construction

In the previous lecture, we constructed the complex logarithm as a collection of function elements. An important idea in this course is that these collections can be glued together to give geometric objects, whose geometry encodes the function we are interested in. In this subsection, we will briefly outline that construction for the complex logarithm.

The idea is to take a quotient of the disjoint union of the domains of the function elements. In the notation of Example 1.13, let

$$R = \left(\coprod_{n \in \mathbb{Z}} U_{I(n)} \right) / \sim$$

where $\sim$ identifies $z_1 \in U_{I(m)}$ and $z_2 \in U_{I(n)}$ if and only if $z_1 = z_2$ in $\mathbb{C}$ and, furthermore, $f_{I(m)}(z_1) = f_{I(n)}(z_2)$; in other words, z_1 and z_2 are the same point in $U_{I(m)} \cap U_{I(n)}$, and $F_{I(n)}$ is a direct analytic continuation of $F_{I(m)}$. By giving R the quotient topology, it becomes a topological space. Furthermore, R is path-connected, because all of the function elements $F_\bullet$ are in the same complete analytic function.

Informally, this construction can be thought of as an 'infinite multi-storey car park': standing on the level $U_{I(n)}$, one can go up through level $U_{I(n+1)}$ to the level above, $U_{I(n+1)}$; alternatively, one can go down through the level $U_{I(n-1)}$ to the level below, $U_{I(n-2)}$.

As well as being an abstract space, R is equipped with two well-defined functions to $\mathbb{C}$. The first comes from the function elements $F_\bullet = (f_\bullet, U_\bullet)$. Since f_A and f_B agree wherever points of U_A and U_B are identified in R, there is a well-defined map

$$f : R \to \mathbb{C}$$

defined as follows: given $[z] \in R$, choose any representative $z \in U_A$ (where $A = I(n)$ or $A = J(n)$, for some n) and set $f([z]) = f_A(z)$. The second is defined in the same way, but using the natural inclusion maps $\pi_\bullet : U_\bullet \hookrightarrow \mathbb{C}$. Again, these agree whenever two points are identified in R, and so together give a well-defined map $\pi : R \to \mathbb{C}$.

Remark 2.8. Since each $f_\bullet$ is defined as an inverse to the exponential map, it follows that $\exp \circ f_\bullet \equiv \pi_\bullet$, and so $\exp \circ f \equiv \pi$.

The functions f and π can be used together to define a continuous map

$$\begin{aligned} \Phi : R &\to \mathbb{C}^2 \\ z &\mapsto (\pi(z), f(z)) \end{aligned}$$

which can be used to study R. Looking back at the definition of the relation that defines R, we see that

$$z_1 \sim z_2 \Leftrightarrow z_1 = z_2 \text{ and } f(z_1) = f(z_2) \Leftrightarrow \Phi(z_1) = \Phi(z_2)$$

so Φ is injective. In particular, R is homeomorphic to a subset of $\mathbb{C}^2$, so is Hausdorff. In fact, with a little thought, we can see that this map identifies R with the graph of the exponential function,

$$\{(z,w) \in \mathbb{C}^2 \mid w = \exp(z)\}\,.$$

We shall see much more of this as the course goes on!

The moral of this construction is that, as we try to analytically continue, we may build functions whose domain (the 'infinite multi-storey car park') is much larger than the original domain that we started with ($\mathbb{C}_*$).

2.3 Complex roots

Along with the complex logarithm, kth roots provide a family of important examples of multi-valued functions, arising, of course, in an attempt to invert the kth power map $p_k : z \mapsto z^k$. These also define complete analytic functions, just as we saw the complex logarithm does in §1.4. Having already handled the complex logarithm, it is easy to deal with kth roots by thinking of them as

$$\sqrt[k]{z} = \exp\left(\frac{1}{k}\log z\right)\,.$$

Example 2.9. Consider intervals $I(n) \subseteq \mathbb{R}$ and open sets $U_{I(n)}$ as in Example 1.13, with each $U_\bullet$ equipped with an analytic branch of log, $f_\bullet$. Define

$$g_\bullet(z) = \exp\left(\frac{1}{k}f_\bullet(z)\right)$$

to obtain a set of function elements $G_{I(n)} := (g_{I(n)}, U_{I(n)})$. These function elements only depend on n modulo $4k$, and so this time we may as well $n \in \mathbb{Z}/4k\mathbb{Z}$. With that modification, the analysis proceeds as before: we have

$$G_{I(m)} \sim G_{I(n)}\,,$$

if and only if $m - n = 0, \pm 1$ modulo $4k$, so the function elements $G_\bullet$ define a complete analytic function, which is the complex kth root.

Furthermore, as in §2.2, we may proceed to define a path-connected, Hausdorff topological space R_k, equipped with a pair of functions $\pi, g : R_m \to \mathbb{C}$, satisfying $\pi(z) = g(z)^k$.

Lecture 3: Riemann surfaces and analytic maps

3.1 Covering maps

In §2.2, we constructed a topological space R that fitted into a commutative diagram of maps.

$$\begin{array}{ccc} R & \xrightarrow{f} & \mathbb{C} \\ & \searrow^{\pi} & \downarrow \exp \\ & & \mathbb{C}_* \end{array}$$

If π were a bijection, then the function $f \circ \pi^{-1}$ would be a single-valued complex logarithm. In fact, π isn't a bijection, but from a topological point of view it is the next best thing.

Definition 3.1 (Covering map)**.** Let X and $\widetilde{X}$ be path-connected, Hausdorff, topological spaces. A *covering map* $\pi : \widetilde{X} \to X$ is a local homeomorphism: that is, each $\tilde{x} \in \widetilde{X}$ has an open neighbourhood $\widetilde{U}$ such that $\pi|_{\widetilde{U}}$ is a homeomorphism onto its image.

A covering map $\pi : \widetilde{X} \to X$ is *regular* if, for each $x \in X$, there is an open neighbourhood U of x and a discrete set Δ_x such that $\pi^{-1}(U)$ is homeomorphic to the direct product $U \times \Delta_x$ and the diagram

$$\begin{array}{ccc} \pi^{-1}(U) & \xrightarrow{\cong} & U \times \Delta_x \\ & \searrow^{\pi} & \downarrow \\ & & U \end{array}$$

commutes, where the map $U \times \Delta_x \to U$ is projection onto the first factor.

Remark 3.2. Beware! The regular covering maps of Definition 3.1 are just called 'covering maps' in *II Algebraic Topology.*

The definition of a regular covering map is often illustrated with a 'stack of records' picture.

Example 3.3. The map $\pi : R \to \mathbb{C}_*$ defined in §2.2 is a covering map, indeed a regular covering map. Each $z \in \mathbb{C}_*$ lives in at least one $U_{I(n)}$ – the left, right, upper or lower half-plane – and in each case the preimage is of the claimed form:

$$\pi^{-1}(U_{I(n)}) = \coprod_{m \cong n \bmod 4} U_{I(m)} \cong U_{I(n)} \times \mathbb{Z}\,.$$

This example is closely related to the exponential map itself.

Example 3.4. For each open interval $I \subseteq \mathbb{R}$, let

$$\widetilde{V}_I = \mathbb{R} + iI = \{x + iy \mid x \in \mathbb{R}, y \in I\}\,.$$

As long as I is of diameter at most 2π, the exponential map restricts to a homeomorphism $\widetilde{V}_I \to U_I$, with inverse provided by the map f_I of Example 1.13. As in the previous example, every $z \in \mathbb{C}_*$ is contained in some $U_{I(n)}$, and

$$\exp^{-1}(U_{I(n)}) = \coprod_{m \cong n \bmod 4} \widetilde{V}_{I(m)} \cong U_{I(n)} \times \mathbb{Z}\,,$$

so $\exp : \mathbb{C} \to \mathbb{C}_*$ is a regular covering map.

However, not all covering maps are regular. Easy examples of irregular covering maps can be constructed by restricting the size of the domain. For instance, the inclusion $\mathbb{D} \hookrightarrow \mathbb{C}$ is a covering map, but not regular, since the preimage of a small disc $D(1, \epsilon)$ is a proper subset of that disc. For a more interesting example see, for instance, question 7 of Example Sheet 3.

Example 3.5. The maps $\pi : R_k \to \mathbb{C}_*$ constructed in Example 2.9 and associated to kth roots are also regular covering maps, just as in Example 3.3.

3.2 Abstract Riemann surfaces

Intuitively, a Riemann surface is a space in which, at every point, there are systems of coordinates that look like the complex plane $\mathbb{C}$. The notion of systems of coordinates is made precise in the next definition. Throughout this section, R is a connected, Hausdorff, topological space.

Definition 3.6. A *chart* on R is a pair (ϕ, U), where U is an open subset of R and $\phi : U \to D$ is a homeomorphism to an open subset of $\mathbb{C}$. A set of charts $\mathcal{A}$ is called an *atlas on* R if the following two conditions hold:

(i) $\bigcup_{(\phi,U)\in\mathcal{A}} U = R$;

(ii) if $(\phi_1, U_1), (\phi_2, U_2) \in \mathcal{A}$ and $U_1 \cap U_2 \neq \varnothing$, then

$$\phi_1 \circ \phi_2^{-1} \equiv (\phi_1|_{U_1 \cap U_2}) \circ (\phi_2|_{U_1 \cap U_2})^{-1}$$

is analytic on $\phi_2(U_1 \cap U_2)$.

The composition $\phi_1 \circ \phi_2^{-1}$ is called a *transition function.*

The reader may like to compare this with the definition of an abstract surface from *IB Geometry*. The main difference is the requirement that the transition functions should be analytic. This extra hypothesis will enable us to use notions from the complex plane when we work in the coordinates provided by charts in $\mathcal{A}$.

Remark 3.7. A few remarks are in order.

(i) Transition functions are always invertible, since $(\phi_1 \circ \phi_2^{-1})^{-1} = \phi_2 \circ \phi_1^{-1}$.

(ii) The space R is connected and locally path-connected, hence path-connected.

Since Riemann surface are locally modelled on $\mathbb{C}$, it should come as no surprise that $\mathbb{C}$ is a Riemann surface.

Example 3.8. Taking

$$\mathcal{A} = \{\mathrm{id} : \mathbb{C} \to \mathbb{C}\}$$

defines an atlas on $\mathbb{C}$. Many others are possible; for instance, $\mathcal{A}' = \{z \mapsto z+1\}$ is an atlas, and so is the union $\mathcal{A} \cup \mathcal{A}'$.

Example 3.8 illustrates a problem with the definition of an atlas: the two different atlases $\mathcal{A}$ and $\mathcal{A}'$ carry the same information, because they are both contained in a common larger atlas, $\mathcal{A} \cup \mathcal{A}'$. Thus, we should only consider atlases that are as large as possible.

Definition 3.9. A *conformal structure* on R is an atlas $\mathcal{A}$ on R which is maximal in the following sense: if (ψ, V) is a chart on R such that, for any $(\phi, U) \in \mathcal{A}$, the transition function $\phi \circ \psi^{-1}$ is analytic, then $(\psi, V) \in \mathcal{A}$.

We now have all the concepts we need to define Riemann surfaces.

Definition 3.10. A *Riemann surface* is a pair $(R, \mathcal{A})$, where $\mathcal{A}$ is a conformal structure on R. By abuse of notation, the notation R will usually denote the Riemann surface $(R, \mathcal{A})$.

This definition presents a technical problem: because conformal structures are necessarily huge, it is rarely practical to exhibit an entire conformal structure. Fortunately, by the next result, it suffices to exhibit any atlas.

Lemma 3.11. *Every atlas $\mathcal{A}$ is contained in a unique conformal structure $\widehat{\mathcal{A}}$.*

Proof. Let $\widehat{\mathcal{A}}$ be the set of all charts (ψ, V) on R such that $\psi \circ \phi^{-1}$ is analytic for every $(\phi, U) \in \mathcal{A}$. This is necessarily maximal so, to prove existence, it remains to show that it is an atlas. Take $(\psi_1, V_1), (\psi_2, V_2) \in \widehat{\mathcal{A}}$, and an arbitrary point $p \in V_1 \cap V_2$. Since $\mathcal{A}$ is an atlas, there is $(\phi, U) \in \mathcal{A}$ with $p \in U$. Therefore,

$$\psi_1 \circ \psi_2^{-1} = (\psi_1 \circ \phi^{-1}) \circ (\phi \circ \psi_2^{-1})$$

is analytic at $\psi_2(p)$, since it is a composition of analytic maps. Since $\psi_2(p)$ is an arbitrary point in the domain of $\psi_1 \circ \psi_2^{-1}$, it follows that $\psi_1 \circ \psi_2^{-1}$ is analytic, so $\widehat{\mathcal{A}}$ is an atlas, and hence a conformal structure.

Uniqueness is clear: if $\mathcal{A}'$ is any atlas containing $\mathcal{A}$ then, by the definition of an atlas, every chart (ϕ', U') in $\mathcal{A}'$ has analytic transition function with every $(\phi, U) \in \mathcal{A}$, so $\mathcal{A}' \subseteq \widehat{\mathcal{A}}$. In particular, $\widehat{\mathcal{A}}$ contains any other conformal structure that contains $\mathcal{A}$, so is unique. □

Example 3.12. The atlas

$$\mathcal{A} = \{\mathrm{id} : \mathbb{C} \to \mathbb{C}\}$$

from Example 3.9 is contained in a unique conformal structure $\widehat{\mathcal{A}}$ on $\mathbb{C}$. Thus, $\mathbb{C}$ is a Riemann surface.

Beware! This is not the only Riemann surface structure on $\mathbb{C}$.

Example 3.13. The atlas

$$\overline{\mathcal{A}} = \{z \mapsto \bar{z}\}$$

on $\mathbb{C}$ is *not* contained in the conformal structure $\widehat{A}$, so defines a different conformal structure. The Riemann surface obtained by equipping $\mathbb{C}$ with this conformal structure is denoted by $\overline{\mathbb{C}}$.

Although other conformal structures on $\mathbb{C}$ exist, $\widehat{A}$ is called the *canonical* conformal structure on $\mathbb{C}$, and is the one we will always use.

Domains in $\mathbb{C}$ provide many more examples.

Example 3.14. If S is any open subset of a Riemann surface R, then restricting the charts of the conformal structure on R to S defines a conformal structure on S. In particular, every domain in $\mathbb{C}$ is a Riemann surface. Particular favourites include $\mathbb{C}_*$, the punctured plane, and $\mathbb{D}$, the unit disc.

However, the real power of the definition stems from being able to equip *larger* spaces with conformal structures. The next example is one of the most famous Riemann surfaces, and you have probably encountered it before.

Example 3.15 (Riemann sphere). Let $\mathbb{C}_\infty$ denote the 2-sphere S^2, identified with $\mathbb{C} \cup \{\infty\}$ via stereographic projection. We will define an atlas with two charts. The first chart is (ϕ, U), where $U = \mathbb{C}$ and $\phi : U \to \mathbb{C}$ is the identity. The second is (ψ, V), where $V = \mathbb{C}_\infty \smallsetminus \{0\}$ and $\psi(z) = 1/z$ (taking, of course, $\psi(\infty) = 0$). Setting

$$\mathcal{A} = \{(\phi, U), (\psi, V)\},$$

both transition functions are $z \mapsto 1/z$ on $\mathbb{C}_*$, which is analytic. Therefore, $\mathcal{A}$ is an atlas and defines a conformal structure on $\mathbb{C}_\infty$. This Riemann surface is called the *Riemann sphere*.

Remark 3.16. The very alert reader may have noticed that, in contrast to the definition of abstract surfaces in *IB Geometry*, Riemann surfaces are not assumed to be second countable. In fact, a remarkable theorem of Radó asserts that every Riemann surface is second countable!

3.3 Analytic maps

As mentioned in the previous section, the fact that the transition functions in an atlas are analytic makes it possible to import notions from the complex plane to R. Most importantly, the notion of an *analytic map* between Riemann surfaces makes sense.

Definition 3.17. Let R and S be Riemann surfaces. A continuous map

$$f : R \to S$$

is *analytic* or *holomorphic* if, for all charts (ϕ, U) on R and (ψ, V) on S, the map $\psi \circ f \circ \phi^{-1}$ is analytic on $\phi(U \cap f^{-1}V)$.

Because transition functions are analytic, it follows that this definition can be checked at each point, just like the usual notion of an analytic map on the complex plane.

Lemma 3.18. *A continuous map of Riemann surfaces $f : R \to S$ is analytic if and only if the following holds: for each $p \in R$, there is a chart (ϕ_p, U_p) on R with $p \in U_p$ and a chart (ψ_p, V_p) on S with $f(p) \in V_p$ such that the map of open subsets of $\mathbb{C}$*

$$\psi_p \circ f \circ \phi_p^{-1} : \phi_p(U_p \cap f^{-1}V_p) \to \psi_p(V_p)$$

is analytic at $\phi_p(p)$.

Proof. The 'only if' direction is immediate. For the 'if' direction, given charts (ϕ, U) on R and (ψ, V) on S, it suffices to show that $\psi \circ f \circ \phi^{-1}$ is analytic at an arbitrary point $\phi(p)$ of its domain $\phi(U \cap f^{-1}V)$. The given hypothesis provides charts (ϕ_p, U_p) on R and (ψ_p, V_p) on S such that

$$\psi_p \circ f \circ \phi_p^{-1} : \phi_p(U \cap f^{-1}V) \to \psi_p(V)$$

is analytic at $\phi_p(p)$. Hence

$$\psi \circ f \circ \phi^{-1} = (\psi \circ \psi_p^{-1}) \circ (\psi_p \circ f \circ \phi_p^{-1}) \circ (\phi_p \circ \phi^{-1})$$

is analytic at $\phi(p)$, since the transition functions $\psi \circ \psi_p^{-1}$ and $\phi_p \circ \phi^{-1}$ are analytic. □

To make use of the definition of analytic maps, we need to check that various properties of analytic maps on $\mathbb{C}$ carry over to the setting of Riemann surfaces. The proofs of these are usually easy: just work in local coordinates provided by charts, and then appeal to standard results about the complex plane. Lemma 3.18 makes it particularly easy.

Lemma 3.19. *If $f : R \to S$ and $g : S \to T$ are analytic maps of Riemann surfaces, then $g \circ f : R \to T$ is analytic.*

Proof. Let $p \in R$. Since f is analytic, there are charts (ϕ_p, U_p) about p and (ψ_p, V_p) about $f(p)$ such that $\psi_p \circ f \circ \phi_p$ is analytic at $\phi_p(p)$. Likewise, since g is analytic, there are charts $(\psi_{f(p)}, V_{f(p)})$ about $f(p)$ and $(\theta_{f(p)}, W_{f(p)})$ about $g \circ f(p)$ such that $\theta_{f(p)} \circ g \circ \psi_{f(p)}^{-1}$ is analytic at $\psi_{f(p)}(f(p))$. Therefore,

$$\theta_{f(p)} \circ (g \circ f) \circ \phi_p^{-1} = (\theta_{f(p)} \circ g \circ \psi_{f(p)}^{-1}) \circ (\psi_{f(p)} \circ \psi_p^{-1}) \circ (\psi_p \circ f \circ \phi_p^{-1})$$

is a composition of analytic maps and hence analytic at $\phi_p(p)$. Therefore, $g \circ f$ is analytic by Lemma 3.18. □

In pure mathematics, we are always interested in finding the right notion of equivalence for the objects we study.

Definition 3.20. A *conformal equivalence* or *biholomorphism* is an analytic bijection of Riemann surfaces $f : R \to S$ with an analytic inverse $f^{-1} : S \to R$. By Lemma 3.19, conformal equivalence is an equivalence relation.

Example 3.21. Complex conjugation $z \mapsto \bar{z}$ defines analytic bijections in both directions between the Riemann surfaces $\mathbb{C}$ and $\overline{\mathbb{C}}$. Thus, these two Riemann surfaces are conformally equivalent.

Lecture 4: Examples of conformal structures and analytic functions

To apply this theory to the object R we associated to the complex logarithm in §2.2, we need to find a way to put a conformal structure on R. Fortunately, the covering maps of §3.1 are a big help.

4.1 Covering maps and analyticity

Using covering maps, we can 'pull back' a conformal structure from the range to the domain.

Lemma 4.1. *If $\pi : \widetilde{R} \to R$ is a covering map and R is a Riemann surface then there is a unique conformal structure on $\widetilde{R}$ such that π is analytic.*

Proof. Each point $p \in \widetilde{R}$ has an open neighbourhood $\widetilde{N}_p$ such that the restriction of π to $\widetilde{N}_p$ is a homeomorphism onto its image. Because R is a Riemann surface, there is a chart (ϕ_p, U_p) about $\pi(p)$. Setting $\tilde{\phi}_p = \phi_p \circ \pi$ (suitably restricted) and $\widetilde{U}_p = \widetilde{N}_p \cap \pi^{-1}(U_p)$ defines a chart $(\tilde{\phi}_p, \widetilde{U}_p)$ about p.

We will show that

$$\widetilde{\mathcal{A}} = \{(\tilde{\phi}_p, \widetilde{U}_p) \mid p \in \widetilde{R}\}$$

is an atlas on $\widetilde{R}$. The sets $\widetilde{U}_p$ cover $\widetilde{R}$ by construction, so it suffices to check that the transition functions are analytic. Given two points $p, q \in \widetilde{R}$, the map π is invertible when restricted to the intersection $\widetilde{U}_p \cap \widetilde{U}_q$ by construction. Therefore, the transition function of the corresponding charts is

$$\tilde{\phi}_p \circ \tilde{\phi}_q^{-1} = \phi_p \circ \pi \circ \pi^{-1} \circ \phi_q^{-1} = \phi_p \circ \phi_q^{-1}$$

which is also a transition function of R, and so is analytic. Therefore $\widetilde{\mathcal{A}}$ is indeed an atlas, and defines a conformal structure.

At an arbitrary point $p \in \widetilde{R}$, we may take the charts $(\tilde{\phi}_p, \widetilde{U}_p)$ about p and (ϕ_p, U_p) about $\pi(p)$, and then

$$\phi_p \circ \pi \circ \tilde{\phi}_p^{-1} = \phi_p \circ \pi \circ \pi^{-1} \circ \phi_p^{-1} = \mathrm{id}$$

which is certainly analytic at p, so this conformal structure does indeed make π into an analytic map.

For uniqueness, suppose that $\widetilde{\mathcal{B}}$ is some conformal structure that makes π into an analytic map, let $p \in \widetilde{R}$ be arbitrary, and let $(\psi, V) \in \widetilde{\mathcal{B}}$ be any chart about p. Since π is analytic, the composition

$$\tilde{\phi}_p \circ \psi^{-1} = \phi_p \circ \pi \circ \psi^{-1}$$

is analytic, so $\tilde{\phi}_p$ has analytic transition function with every chart in $\widetilde{\mathcal{B}}$. Hence $\widetilde{\mathcal{A}} = \widetilde{\mathcal{B}}$ by Lemma 3.11. □

This makes it easy to put a conformal structure on the space we associated to the complex logarithm in §2.2.

Example 4.2. Let R be as in §2.2. In Example 3.3 we saw that natural map $\pi : R \to \mathbb{C}_*$ is a covering map. Hence, by Lemma 4.1, there is a unique conformal structure on R that makes π analytic; in particular, R is a Riemann surface.

In fact, more is true. Recall that R and π fitted into a commutative diagram.

$$\begin{array}{ccc} R & \xrightarrow{f} & \mathbb{C} \\ & \searrow^{\pi} & \downarrow \exp \\ & & \mathbb{C}_* \end{array}$$

On any given domain $U_{I(n)} \subseteq R$, in the standard charts, f can be written as

$$f_{I(n)} : U_{I(n)} \to \mathbb{C}$$

which is analytic, so this conformal structure on R also makes f into an analytic map.

Furthermore, recall from Example 3.4 that $f_{I(n)} : U_{I(n)} \to \widetilde{V}_{I(n)}$ has an analytic inverse $f_{I(n)}^{-1} = \exp|_{\widetilde{V}_{I(n)}}$, for each $n \in \mathbb{Z}$. These inverses agree on the intersections $\widetilde{V}_{I(m)} \cap \widetilde{V}_{I(n)}$, and so piece together to give a globally defined analytic inverse $f^{-1} : \mathbb{C} \to R$. In particular, f is a conformal equivalence.

A similar analysis applies to kth roots.

Example 4.3. Let R_k be as in Example 2.9, which fits into a commutative diagram

$$\begin{array}{ccc} R_k & \xrightarrow{g} & \mathbb{C}_* \\ & \searrow^{\pi} & \downarrow p_k \\ & & \mathbb{C}_* \end{array}$$

where p_k is the power map $z \mapsto z^k$. In Example 3.5, we saw that π is a covering map so, by Lemma 4.1, there is a unique conformal structure on R_k that makes π analytic. As in Example 4.2, g takes the form

$$g_{I(n)} : U_{I(n)} \to \mathbb{C}_*$$

on any open set $U_{I(n)} \subseteq R_k$, so g is also analytic. Again, the functions $g_{I(n)}$ have analytic inverses $g_{I(n)}^{-1}$ (defined by suitable restrictions of the power maps), which piece together to define a global inverse $g^{-1} : \mathbb{C}_* \to R_k$, so g is a conformal equivalance.

Unlike the complex logarithm, the kth root has an additional nice property. The space R_k can be embedded in a compact Riemann surface $\widehat{R}_k$, and g and π extended, to obtain a commutative diagram of compact Riemann surfaces.

$$\begin{array}{ccc} \widehat{R}_k & \xrightarrow{\hat{g}} & \mathbb{C}_\infty \\ & \searrow^{\hat{\pi}} & \downarrow{\scriptstyle p_k} \\ & & \mathbb{C}_\infty \end{array}$$

Compactifying loses the nice property that π is a covering map, since p_k fails to be a local homeomorphism. However, it is a huge advantage to work with compact spaces, so this sacrifice is usually worth making. We will learn a lot more about compact Riemann surfaces later in the course.

4.2 Analytic functions

An important way to study a Riemann surface R is via analytic maps from R to well-understood Riemann surfaces. Since $\mathbb{C}$ is the prototypical Riemann surface, maps to $\mathbb{C}$ are especially important.

Definition 4.4. An *analytic function* on a Riemann surface R is an analytic map $R \to \mathbb{C}$.

To check that a function $f : R \to \mathbb{C}$ is analytic, we only need to find a chart (ϕ, U) about each point $p \in R$ such that $f \circ \phi$ is analytic. Recall the statement of the inverse function theorem from *IB Complex Analysis*.

Theorem 4.5 (Inverse function theorem)**.** *Let f be an analytic function on a domain $D \subseteq \mathbb{C}$. If $f'(z_0) \neq 0$ for $z_0 \in D$, then there are open neighbourhoods U of z_0 and V of $f(z_0)$ such that f restricts to a biholomorphism $U \to V$.*

One advantage of working with Riemann surfaces is that, in a suitable chart, any analytic function can be put into a very simple local form. The proof of this uses the inverse function theorem.

Proposition 4.6. *Let f be a non-constant analytic function on a Riemann surface R and let $p \in R$ be a zero of f. There is a chart (ϕ, U) about p with $\phi(p) = 0$ such that*

$$f \circ \phi^{-1}(z) = z^m$$

for some integer $m > 0$.

Proof. Choose a chart (ψ, V) about p; without loss of generality, $\psi(p) = 0$. As in the proof of the principle of isolated zeroes, there is an analytic function g with $g(0) \neq 0$ such that

$$f \circ \psi^{-1}(z) = z^m g(z)$$

on a neighbourhood of 0, for some natural number m. By the identity principle for Riemann surfaces (question 10 of Example Sheet 1), $m > 0$ since f is non-constant.

Since $g(0) \neq 0$ and is continuous, there is $\delta > 0$ such that $g(D(0,\delta)) \subseteq D(g(0), |g(0)|)$. Since there is an analytic mth root defined on $D(g(0), |g(0)|)$ (for instance, one could use one of the function elements from Example 2.9), there is an analytic function $\sqrt[m]{g(z)}$ defined in a neighbourhood of 0. If $h(z) = z\sqrt[m]{g(z)}$, then $f \circ \psi^{-1}(z) = (h(z))^m$. Furthermore,

$$h'(0) = \sqrt[m]{g(0)} \neq 0\,,$$

so h has an analytic inverse defined on some $D(0,\epsilon)$, by the inverse function theorem. Setting $\phi = h \circ \psi$ and $U = \phi^{-1}(D(0,\epsilon))$ gives the required chart, because

$$f \circ \phi^{-1}(z) = f \circ \psi^{-1} \circ h^{-1}(z) = (h \circ h^{-1}(z))^m = z^m$$

as required. □

Lecture 5: Complex tori and the open mapping theorem

5.1 Complex tori

So far, the Riemann sphere is our only example of a compact Riemann surface. In fact, the collection of Riemann surfaces is rich and diverse. In this section, we construct a much more interesting family of examples.

Example 5.1 (Complex tori). Let τ_1, τ_2 be complex numbers that are linearly independent over $\mathbb{R}$ – that is, $\tau_i \in \mathbb{C}_*$ and $\tau_2/\tau_1 \notin \mathbb{R}$. Let $\Lambda = \langle \tau_1, \tau_2 \rangle$, the additive subgroup generated by them, let T be the quotient group $\mathbb{C}/\Lambda$, and let $\pi : \mathbb{C} \to T$ be the quotient map.

As a topological space, T is equipped with the quotient topology. This can be studied via the *fundamental parallelogram* $P \subseteq \mathbb{C}$, the parallelogram with corners $\{0, \tau_1, \tau_2, \tau_1 + \tau_2\}$: T is obtained by identifying opposite sides of P. From this description, it is not hard to prove that T is homeomorphic to the torus $S^1 \times S^1$; in particular, T is compact and Hausdorff.

The map π is a regular covering map. Indeed, if

$$\epsilon < \min\{|\lambda| \mid \lambda \in \Lambda \smallsetminus \{0\}\}/2$$

then, for any $z \in \mathbb{C}$, $D(z, \epsilon) \cap \{z + \Lambda\} = \{z\}$. If $p = \pi(z_0) \in T$, it follows that $U = \pi(D(z_0, \epsilon))$ is open with preimage

$$\pi^{-1}(U) = \coprod_{\lambda \in \Lambda} D(z_0, \epsilon) + \lambda \cong D(z_0, \epsilon) \times \Lambda$$

as required.

Finally, for T to be a Riemann surface, we need to endow it with an atlas. For any $z_0 \in \mathbb{C}$, let $U = \pi(D(z_0, \epsilon))$, an open set in T as above. Then π restricts to a homeomorphism $D(z_0, \epsilon) \to U$, and so we may set $\phi = \pi|^{-1}_{D(z_0,\epsilon)}$ to define a chart (ϕ, U) on T. Since z_0 was arbitrary, the set of all such charts covers T. Furthermore, if (ψ, V) is another such chart with $\psi = \pi|^{-1}_{D(z_1,\epsilon)}$, then the transition function $\phi \circ \psi^{-1}$ is translation by some element of Λ. Since translation is analytic, this defines a conformal structure on T.

Remark 5.2. The construction of Example 5.1 is the reverse of the construction of Lemma 4.1. Rather than 'pulling back' the complex structure along the covering map, the complex structure of $\mathbb{C}$ is 'pushed forward' along π to define a complex structure on T. In general, it is more difficult to make this sort of 'push forward' construction work than the 'pull back' construction of Lemma 4.1. In this case, the construction works because T is a quotient of $\mathbb{C}$ by a group (in this case, Λ), acting by analytic maps on $\mathbb{C}$.

Topologically, these complex tori are all the same: they are all homeomorphic to $S^1 \times S^1$. However, in question 5 of Example Sheet 2, we shall see that different lattices Λ often give rise to complex tori that are not conformally equivalent. In particular, they will provide a source of infinitely many different conformal equivalence classes of compact Riemann surfaces.

5.2 The open mapping theorem

As mentioned in §4.2, we often want to study Riemann surfaces via their analytic functions. However, we shall see in this subsection that the analytic functions on compact Riemann surfaces are not very interesting. This is an easy consequence of the open mapping theorem for Riemann surfaces which, like many other theorems about analytic functions, follows easily from the open mapping theorem on the complex plane from *IB Complex Analysis*.

Theorem 5.3 (Open mapping theorem for Riemann surfaces)**.** *Any non-constant, analytic map of Riemann surfaces $f : R \to S$ is an open map.*

Proof. Consider an open set $W \subseteq R$, and let $p \in R$ be arbitrary. Take charts (ϕ, U) about p and (ψ, V) about $f(p)$. By the identity principle for Riemann surfaces (question 10 of Example Sheet 1), $\psi \circ f \circ \phi^{-1}$ is a non-constant function from $\phi(U \cap W \cap f^{-1}V)$ to $\psi(V)$, so $\psi \circ f(U \cap W \cap f^{-1}V)$ is open, by the open mapping theorem for domains in $\mathbb{C}$. Because ψ is a homeomorphism, $f(U \cap W \cap f^{-1}V)$ is an open neighbourhood of $f(p)$ in $f(W)$, and since p was arbitrary, it follows that $f(W)$ is open, as required. □

This has especially profound consequences for analytic maps of compact Riemann surfaces, such as $\mathbb{C}_\infty$ or the complex tori of Example 5.1.

Corollary 5.4. *Let $f : R \to S$ be a non-constant, analytic map of Riemann surfaces. If R is compact, then f is surjective, and in particular, S is also compact.*

Proof. By the open mapping theorem, $f(R)$ is open. Since R is compact, $f(R)$ is also compact and so closed. The result follows since S is connected. □

This immediately rules out any non-constant analytic function on a compact Riemann surface, since $\mathbb{C}$ is not compact.

Corollary 5.5. *Every analytic function on a compact Riemann surface is constant.*

5.3 Harmonic functions

This subsection is a brief digression into *real* functions on Riemann surfaces. Of course, it doesn't make sense to ask for a real-valued function to be complex-differentiable. Since we need to work with real coordinates, we write $z = x + iy$ for $x, y \in \mathbb{R}$.

Definition 5.6. Let $D \subseteq \mathbb{C}$ be a domain. A smooth function $u : D \to \mathbb{R}$ is called *harmonic* if

$$\Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} \equiv 0$$

on D. (Recall that Δu is the *Laplacian* of u.)

The next lemma explains the connection between analytic and harmonic functions.

Lemma 5.7. *Consider a disc $D \subseteq \mathbb{C}$. A function $u : D \to \mathbb{C}$ is harmonic if and only if*

$$u = \mathrm{Re}(f)$$

for some analytic function f.

Proof. If $f = u + iv$ is analytic, then it satisfies the Cauchy–Riemann equations:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.$$

Therefore,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = \frac{\partial^2 v}{\partial x \partial y} - \frac{\partial^2 v}{\partial y \partial x} = 0$$

so u is harmonic.

For the other direction, see question 11 of Example Sheet 1. □

With this lemma in hand, we can make sense of harmonic functions on Riemann surfaces.

Definition 5.8. Let R be a Riemann surface. A function $u : R \to \mathbb{R}$ is *harmonic* if, for any chart (ϕ, U) on R, the composition

$$u \circ \phi^{-1} : U \to \mathbb{R}$$

is harmonic.

Remark 5.9. Lemma 5.7 implies that Definition 5.8 can be checked on any atlas. If $u \circ \phi^{-1}$ is harmonic then, by the lemma,

$$u \circ \phi^{-1} = \mathbb{R}(f)\,.$$

For any overlapping chart (ψ, V), we therefore have

$$u \circ \psi^{-1} = \mathrm{Re}(f \circ \phi \circ \psi^{-1})\,,$$

which is the real part of a composition of analytic functions, and so harmonic.

Lemma 5.7 makes it quite easy to extend results about analytic functions on Riemann surfaces to harmonic functions. For instance, they satisfy an identity principle.

Proposition 5.10 (Identity principle for harmonic functions)**.** *Let u and v be harmonic functions on a Riemann surface R. Either $u \equiv v$ on R, or the set where they coincide*

$$\{p \in R \mid u(p) = v(p)\}$$

has empty interior.

Proof. See Example Sheet 1, question 12. □

Moreover, harmonic functions also satisfy an open mapping theorem.

Theorem 5.11 (Open mapping theorem for harmonic functions)**.** *Any non-constant harmonic function u on a Riemann surface R is an open map.*

Proof. Suppose $W \subseteq R$ is open. Let $p \in W$, and consider a chart (ϕ, U) about p. By Lemma 5.7, after shrinking U if necessary, there is an analytic function $f : \phi(U) \to \mathbb{C}$ such that $u \circ \phi^{-1} = \mathrm{Re}(f)$. If f is constant on $\phi(U)$ then u is constant on U, and hence on R by Proposition 5.10. Therefore f is non-constant and so, by the open mapping theorem, $f \circ \phi(U)$ is open.

Let $f \circ \phi(p) = a + ib$. Since the topology on $\mathbb{C}$ is identified with the product topology on $\mathbb{R}^2$, $f \circ \phi(U)$ contains an open set of the form

$$(a - \delta, a + \delta) + i(b - \epsilon, b + \epsilon)\,.$$

But $a = u(p)$, so $u(W)$ contains $(u(p) - \epsilon, u(p) + \epsilon)$. Since $p \in W$ was arbitrary, $u(W)$ is open, as required. □

Just as in the analytic case, it follows that all harmonic functions on compact Riemann surfaces are constant.

Corollary 5.12. *If R is a compact Riemann surface, all harmonic functions on R are constant.*

Harmonic functions don't play much of a role in the rest of the course, but the results of this section do at least indicate a connection with Laplace's equation. This connection can be exploited in both directions. Corollary 5.12 shows that Laplace's equation doesn't have interesting solutions on compact Riemann surfaces. In the other direction, Lemma 5.7 indicates that we could aim to construct analytic functions by solving Laplace's equation.

Lecture 6: Meromorphic functions and a worked example

6.1 Meromorphic functions

Let R be a compact Riemann surface. Corollary 5.5 shows us that the analytic functions on R will not yield much information. However, Corollary 5.4 also points to a solution: the problem is that $\mathbb{C}$ is non-compact, and so we can obtain a more interesting theory by compactifying $\mathbb{C}$. This motivates the next definition.

Definition 6.1. A *meromorphic function* on a Riemann surface R is an analytic map $f : R \to \mathbb{C}_\infty$, where $\mathbb{C}_\infty$ is the Riemann sphere, which is not identically ∞.

This is an elegant definition, but to avoid confusion we need to check that it coincides with the definition of a meromorphic function on the complex plane that we already use!

Proposition 6.2. *Let $D \subseteq \mathbb{C}$ be a domain. A function $f : D \to \mathbb{C}$ is meromorphic if and only if there is a discrete subset $A \subseteq D$ such that $f : D \smallsetminus A \to \mathbb{C}$ is analytic, and f has a pole at each $a \in A$.*

Proof. Let's start with the 'only if' direction. Let $A = f^{-1}(\infty)$. By the identity principle for Riemann surfaces A is discrete, since otherwise f would be constant and equal to ∞, contrary to hypothesis. It remains to prove that each $a \in A$ is a pole. Considering the standard chart on $\mathbb{C}_\infty$ about ∞, any $a \in A$ has a neighbourhood on which $1/f(z)$ is analytic and so, as in the proof of the identity principle, takes the form

$$1/f(z) = (z-a)^m g(z)$$

for some $m \geq 1$ and some analytic g with $g(a) \neq 0$. Therefore, on some possibly smaller neighbourhood of a,

$$f(z) = (z-a)^{-m} h(z)$$

for $h(z) = 1/g(z)$ analytic. Thus a is a pole of f, as required.

For the 'if' direction, suppose that f has a pole of order m at $a \in A$, meaning that, in a neighbourhood of a,

$$f(z) = (z-a)^{-m} h(z)$$

for some $m \geq 1$ and some analytic function h with $h(a) \neq 0$. Passing to a smaller neighbourhood on which h is never zero,

$$1/f(z) = (z-a)^m g(z)$$

for $g(z) = 1/h(z)$ analytic. Therefore, f extends at each $a \in A$ to an analytic map to $\mathbb{C}_\infty$. □

6.2 A worked example

Our next topic will give a theoretical proof that analytically continuing any function element defines a Riemann surface. However, before seeing the theoretical construction, we provide a detailed worked example.

We have already seen this in action in a few cases: Example 1.13 associated a Riemann surface to the complex logarithm, while Example 2.9 did the same for kth roots. In this subsection, we will look at a function element for

$$w = \sqrt{z^3 - z}\,.$$

There are two approaches to constructing this Riemann surface. One is to put a conformal structure on the graph

$$\{(w,z) \in \mathbb{C}^2 \mid w^2 = z^3 - z\}\,.$$

However, here we will construct this Riemann surface by gluing together function elements.

Example 6.3. Consider $f(z) = z^3 - z = z(z-1)(z+1)$. Since the complex square root fails to be locally defined at 0 and ∞, we are most concerned with the points $z = 0, 1, -1, \infty$, where $\sqrt{f(z)}$ is not locally defined. This motivates us to define a domain D by joining these points with branch cuts.

$$D = \mathbb{C} \setminus ([-1, 0] \cup [0, \infty))$$

Our goal is to construct analytic functions $g_\pm(z)$ on D that satisfy $g_\pm(z)^2 = f(z)$. Pick any base point $z_0 \in D$, and let $g_\pm(z_0)$ be the two values such that $g_\pm(z_0)^2 = f(z_0)$. For any $z \in D$, choose a path γ in U from z_0 to z. The function $g_\pm(z)$ is then defined by a path integral.

$$g_\pm(z) \coloneqq g_\pm(z_0) \exp\left(\frac{1}{2}\int_\gamma \frac{f'(\zeta)}{f(\zeta)} d\zeta\right)$$

We claim that this definition is independent of the choice of path. To prove this, it suffices to show that

$$\exp \frac{1}{2} \oint_\gamma \frac{f'(\zeta)}{f(\zeta)} d\zeta = 1$$

for any closed curve γ. By the argument principle from *IB Complex analysis*,

$$\begin{aligned} \oint_\gamma \frac{f'(\zeta)}{f(\zeta)} d\zeta &= 2\pi i \left(\sum_{\text{zeroes of } f} n(\gamma, Z) - \sum_{\text{poles of } f} n(\gamma, P) \right) \\ &= 2\pi i (n(\gamma, 1) + n(\gamma, 0) + n(\gamma, -1)), \end{aligned}$$

where $n(\gamma, z)$ denotes winding numbers. In question 1 of Example Sheet 1, we saw that $n(\gamma, 1) = 0$ and $n(\gamma, 0) = n(\gamma, -1)$. Therefore,

$$\frac{1}{2} \oint_\gamma \frac{f'(\zeta)}{f(\zeta)} d\zeta \in 2\pi i \mathbb{Z}$$

and so the claim follows and $g_\pm(z)$ is well-defined.

Next, let's check that $g_\pm$ is continuous. Indeed, since f'/f is continuous at any $z \in D$, there is δ such that

$$\left| \frac{f'(z)}{f(z)} - \frac{f'(\tau)}{f(\tau)} \right| \le 1$$

for any $\tau \in D(z,\delta)$. Therefore,

$$\left|\int_{z_0}^{\tau} \frac{f'(\zeta)}{f(\zeta)} d\zeta - \int_{z_0}^{z} \frac{f'(\zeta)}{f(\zeta)} d\zeta\right| = \left|\int_{z}^{\tau} \frac{f'(\zeta)}{f(\zeta)} d\zeta\right| \leq M|z-\tau| \to 0$$

as $\tau \to z$, where $M = |f'(z)/f(z)| + 1$. This shows that $\int_{z_0}^{z} \frac{f'(\zeta)}{f(\zeta)} d\zeta$ is a continuous function of z, and so $g_\pm$ is also continuous.

Given continuity, $g_\pm(z) = \sqrt{f(z)}$ locally, for a suitable choice of square root, and so $g_\pm$ is analytic.

This construction gives two function elements, denoted by

$$(g_+, D_+) \text{ and } (g_-, D_-)$$

where $D_\pm = D$. The Riemann surface will be defined by gluing these two function elements together. Before we do this, let's discuss the topology of D.

The domain D is obtained by deleting two closed intervals from the Riemann sphere $\mathbb{C}_\infty$, which lie on the equator $\mathbb{R} \cup \{\infty\}$. From this, it follows that D is topologically an open annulus $S^1 \times \mathbb{R}$.

For any z_0 in the intervals $(-1,0) \cup (0,\infty)$, let $z \to z_0^+$ denote the approaching z_0 from the upper half-plane, and let $z \to z_0^-$ denote approaching z_0 from the lower half-plane. It is not hard to see that

$$\lim_{z \to z_0^-} g_+(z_0) = \lim_{z \to z_0^+} g_-(z_0)$$

and

$$\lim_{z \to z_0^+} g_+(z_0) = \lim_{z \to z_0^-} g_-(z_0)\,,$$

for any z_0 in the interiors of the branch cuts. Thus, we can glue the two function elements together along the open intervals $(-1,0) \cup (0,\infty)$ to obtain a Riemann surface R and a well-defined function

$$g : R \to \mathbb{C}\,.$$

Since R is obtained by gluing two annuli along four open intervals, we can see that R is homeomorphic to the torus $S^1 \times S^1$ with four points deleted. The Riemann surface R is also equipped with another function $\pi : R \to \mathbb{C}$, defined by the inclusions of $D_\pm$ into $\mathbb{C}$. By construction two functions together satisfy the equation

$$g(p)^2 = f \circ \pi(p)$$

at every $p \in R$. It is in this sense that g solves the equation $w^2 = f(z)$.

To be fully rigorous about this gluing, we need to describe the complex structure at the glued points $z_0 \in (-1,0) \cup (0,\infty)$. This can be done by considering function elements on small discs $D(z_0,\epsilon)$. Alternatively, one can make the opposite choice of branch cuts

$$E = \mathbb{C} \setminus ((-\infty,-1] \cup [0,1])$$

define function elements $(h_\pm, E_\pm)$ similarly to above, and then construct R by gluing together $D_\pm$ and $E_\pm$ wherever they overlap and the function elements agree. However, we will usually suppress this detail.

Lecture 7: Covering-space theory

This section introduces some tools from topology that will be have many applications in this course, especially when we come to give a complete treatment of analytic continuation.

7.1 Covering-space theory

Lifts are the fundamental objects of study here.

Definition 7.1. Suppose $\pi : \widetilde{X} \to X$ is a covering map and $\gamma : [0,1] \to X$ is a path. A *lift of* γ *along* π is a path $\tilde{\gamma} : [0,1] \to \widetilde{X}$ such that $\pi \circ \tilde{\gamma} = \gamma$.

Example 7.2. In Example 3.4 we saw that $\exp : \mathbb{C} \to \mathbb{C}_*$ is a covering map. Consider the anticlockwise loop around the unit circle

$$\gamma(t) = e^{2\pi i t}.$$

Then

$$\tilde{\gamma}_1(t) = 2\pi i t$$

and

$$\tilde{\gamma}_2(t) = 2\pi i(1+t)$$

are both lifts of γ.

As this example shows, lifts are not unique. However, the two different lifts given start in different places, and it turns out that this extra information – the start point – is all that's needed to determine a lift.

Proposition 7.3 (Uniqueness of lifts). *Suppose $\tilde{\gamma}_1, \tilde{\gamma}_2$ are both lifts of γ along a covering map π. If $\tilde{\gamma}_1(0) = \tilde{\gamma}_2(0)$ then $\tilde{\gamma}_1 \equiv \tilde{\gamma}_2$.*

Proof. The result follows from the claim that the subset $I = \{t \in I \mid \tilde{\gamma}_1(t) = \tilde{\gamma}_2(t)\}$ is both open and closed in $[0,1]$. Since I is hypothesised to contain 0, it follows that $I = [0,1]$ by connectedness.

The fact that I is closed follows from very general facts. Because X is Hausdorff, the diagonal X is a closed subset of the product $X \times X$, endowed with the product topology. Since the map

$$\gamma_1 \times \gamma_2 : I \to X \times X$$

sending $t \mapsto (\gamma_1(t), \gamma_2(t))$ is continuous, it follows that I, the preimage of the diagonal, is also closed.

To prove that I is open, consider any $t \in I$. Since π is a covering, $\tilde{\gamma}_1(t) = \tilde{\gamma}_2(t)$ has an open neighbourhood $\widetilde{U}$ such that $\pi|_{\widetilde{U}}$ is a a homeomorphism onto its image U. By continuity, there is $\delta > 0$ such that $\tilde{\gamma}_1(t-\delta, t+\delta)$ and $\tilde{\gamma}_2(t-\delta, t+\delta)$ are both contained in $\widetilde{U}$. For all $t - \delta < s < t + \delta$, we therefore have

$$\tilde{\gamma}_1(s) = \pi|_{\widetilde{U}}^{-1} \circ \gamma(s) = \tilde{\gamma}_2(s)\,,$$

so $s \in I$ as required. □

Mathematicians like to complement uniqueness results with existence results. However, for arbitrary coverings, lifts need not exist, even if the covering map is surjective.

Example 7.4. Let $\widetilde{X} = \{z \in \mathbb{C} \mid -\pi < \mathrm{Im}(z) < 2\pi\}$, and consider the surjective covering map

$$\pi : \widetilde{X} \to \mathbb{C}_*$$

which is the restriction of the exponential map. Let $\gamma(t) = e^{2\pi i t}$, as in Example 7.2. Since $\pi^{-1}(1) = \{0\}$, any lift $\tilde{\gamma}$ of γ must start at 0. But such a lift $\tilde{\gamma}$ is also a lift of γ along the exponential map so, by uniqueness of lifts, it must coincide with $\tilde{\gamma}_1$ from Example 7.2. However, $\tilde{\gamma}_1(1) = 2\pi i \notin \widetilde{X}$, so this is impossible. Hence, there is no lift of γ along π.

This is where *regular* coverings come into their own. The proof of the next lemma follows a similar strategy to the proof of uniqueness of lifts.

Proposition 7.5 (Path-lifting lemma). *Let $\pi : \widetilde{X} \to X$ be a regular covering map. Let $\gamma : [0,1] \to X$ be a path, and suppose that $\pi(\tilde{x}) = \gamma(0)$. Then there is a (unique) lift $\tilde{\gamma}$ of γ such that $\tilde{\gamma}(0) = \tilde{x}$.*

Proof. Let I be the set of $t \in [0,1]$ such that there exists a lift $\tilde{\gamma} : [0,t] \to \widetilde{X}$ of $\gamma|_{[0,t]}$ along π with $\tilde{\gamma}(0) = \tilde{x}$. Again, the result follows from the claim that I is both open and closed, since $0 \in I$ by hypothesis.

To show that I is closed, consider a sequence $t_n \in I$ converging to some $\tau \in [0,1]$. Since π is a regular covering map and X is locally path-connected, $\gamma(\tau)$ has a path-connected open neighbourhood U such that

$$\pi^{-1}(U) = \coprod_{\delta \in \Delta} U_\delta \cong U \times \Delta$$

for some discrete set Δ. Let N be large enough that $\gamma(t_n) \in U$ for all $n \geq N$. It follows that $\tilde{\gamma}(t_n)$ are in the same path component U_δ of $\pi^{-1}(U)$ for all $n \geq N$. Therefore, setting

$$\tilde{\gamma}(\tau) = (\pi|_{U_\delta})^{-1} \circ \gamma(\tau)$$

extends $\tilde{\gamma}$ to a continuous lift at τ, so $\tau \in I$ as required.

To show that I is open, consider $\tau \in I$, and again, let U be a path-connected open neighbourhood of $\gamma(\tau)$ such that

$$\pi^{-1}(U) = \coprod_{\delta \in \Delta} U_\delta \,.$$

There is a unique δ such that $\tilde{\gamma}(\tau) \in U_\delta$. Let $\epsilon > 0$ be such that $\gamma(t) \in U$ whenever $|t - \tau| < \epsilon$. By uniqueness of lifts,

$$\tilde{\gamma}(t) = (\pi|_{U_\delta})^{-1} \circ \gamma(t)$$

for every $t \in (\tau - \epsilon, \tau + \epsilon) \cap I$. This shows that $\tilde{\gamma}$ can be extended to an open neighbourhood of τ, as required. □

Rather than just being interested in individual curves, we often want to deform one into another.

Definition 7.6. Let X be a a topological space and $\alpha, \beta : [0,1] \to X$ paths with $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$. This pair of paths is *homotopic* (conveniently written $\alpha \simeq \beta$) if there exists a family of paths $(\alpha_s)_{s \in [0,1]}$ such that:

(i) $\alpha_0 \equiv \alpha$ and $\alpha_1 \equiv \beta$;

(ii) $\alpha_s(0) = \alpha(0)$ and $\alpha_s(1) = \alpha(1)$ for all s;

(iii) the map $(t,s) \mapsto \alpha_s(t)$ is a continuous map $[0,1]^2 \to X$.

In particular, this enables us to make rigorous sense of the notion of a 'space with no holes'.

Definition 7.7. A topological space X is *simply connected* if:

(i) X is path-connected; and

(ii) every pair of paths $\alpha, \beta : [0,1] \to X$ with the same endpoints (i.e. $\alpha(0) = \beta(0)$ and $\alpha(1) = \beta(1)$) is homotopic.

Remark 7.8. Let $D \subseteq \mathbb{C}$ be a convex domain. The formula

$$\alpha_s(t) = (1-s)\alpha(t) + s\beta(t)$$

defines a homotopy between any two paths α, β in D with equal endpoints, so D is simply connected.

The next result is an important statement about the existence of homotopies. In due course, it will enable us to prove uniqueness results about analytic continuation.

Theorem 7.9 (Monodromy theorem)**.** *Let $\pi : \widetilde{X} \to X$ be a covering map and let α, β be paths in X. Suppose that:*

(i) $\alpha \simeq \beta$ in X;

(ii) there are lifts $\tilde{\alpha}$ of α and $\tilde{\beta}$ of β with $\tilde{\alpha}(0) = \tilde{\beta}(0)$;

(iii) every path γ in X with $\gamma(0) = \alpha(0) = \beta(0)$ has a lift to $\widetilde{X}$ with $\tilde{\gamma(0)} = \tilde{\alpha}(0) = \tilde{\beta}(0)$.

Then $\tilde{\alpha} \simeq \tilde{\beta}$; in particular, $\tilde{\alpha}(1) = \tilde{\beta}(1)$.

This result is not proved here. In *II Algebraic topology*, an essentially identical result is called the *homotopy lifting lemma*, and a proof is given.

Remark 7.10. Note that hypotheses (ii) and (iii) of Theorem 7.9 are automatically satisfied if π is a regular covering map.

Lecture 8: The monodromy group and the space of germs

8.1 The monodromy group

In this section, we shall see that the monodromy theorem can be used to define an invariant of regular covering maps – the *monodromy group*. It's an important object, corresponding to the Galois group in Galois theory and to the deck group in algebraic topology. We won't discuss its properties much in this course, but you will be asked to compute it on example sheets.

Let $\pi : \widetilde{X} \to X$ be a regular covering map, and let's pick a base point $x_0 \in X$. Now consider a loop $\gamma : [0,1] \to X$ based at x_0; that is, a path with $\gamma(0) = \gamma(1) = x_0$. Path lifting enables us to associate to γ a self-map σ_γ of the preimage $\pi^{-1}(x_0)$.

Definition 8.1. Let $\tilde{x} \in \pi^{-1}(x_0)$, and let $\tilde{\gamma}_{\tilde{x}}$ be the unique lift of γ starting at $\tilde{x}$. Since it is a lift,

$$\pi(\tilde{\gamma}_{\tilde{x}}(1)) = \gamma(1) = x_0$$

so $\tilde{\gamma}_{\tilde{x}}(1) \in \pi^{-1}(x_0)$. Therefore, we may define $\sigma_\gamma : \pi^{-1}(x_0) \to \pi^{-1}(x_0)$ by

$$\sigma_\gamma(\tilde{x}) := \tilde{\gamma}_{\tilde{x}}(1)$$

for any $\tilde{x} \in \pi^{-1}(x_0)$.

The next remarks collect together a number of nice properties of this definition.

Remark 8.2. Let $\pi : \widetilde{X} \to X$, x_0 and γ be as above.

(i) For the constant loop $\iota : t \mapsto x_0$, the corresponding map σ_ι is the identity.

(ii) Let $\bar{\gamma}$ be the loop

$$\bar{\gamma}(t) := \gamma(1-t)\,.$$

Using uniqueness of lifts, it is easy to see that $\sigma_{\bar{\gamma}} = \sigma_\gamma^{-1}$. In particular, σ_γ is always a bijection, i.e. a permutation.

(iii) If α and β are both loops based at x_0, we can define the *concatenation*

$$\alpha \cdot \beta(t) := \begin{cases} \alpha(2t) & 0 \le t \le 1/2 \\ \beta(2t-1) & 1/2 \le t \le 1 \end{cases}$$

which is also a loop based at x_0. If $\tilde{x}_1 \in \pi^{-1}(x_0)$ and the lift $\tilde{\alpha}_{\tilde{x}_1}$ ends at $\tilde{x}_2$ then,

$$(\widetilde{\alpha \cdot \beta})_{\tilde{x}_1} = \tilde{\alpha}_{\tilde{x}_1} \cdot \tilde{\beta}_{\tilde{x}_2}$$

by uniqueness of lifts. In particular,

$$\sigma_{\alpha \cdot \beta})(\tilde{x}_1) = \sigma_\beta(\tilde{x}_2) = \sigma_\beta \circ \sigma_\alpha(\tilde{x}_1)\,.$$

Taking all of these properties together, we see that the set of all permutations σ_γ of $\pi^{-1}(x_0)$ corresponding to loops based at x_0 form a subgroup H_{x_0} of the symmetric group $\mathrm{Sym}(\pi^{-1}(x_0))$. This is the *monodromy group* of the regular covering map π. The definition appears to also depend on the choice of base point, but in fact this dependence is an illusion.

Remark 8.3. Let $\pi : \widetilde{X} \to X$ be as above, with X path connected.

(i) Let x_0 and x_1 be different choices of base point in X. Concatenation with α can be used to transform a loop γ based at x_1 into a loop base at x_0, via the map

$$\gamma \mapsto \alpha \cdot \gamma \cdot \bar{\alpha}\,.$$

This map defines a homomorphism of monodromy groups

$$\theta_\alpha : H_{x_1} \to H_{x_0}$$

and, by uniqueness of lifts again, it is not hard to see that $\theta_{\bar{\alpha}} = \theta_\alpha^{-1}$. In particular, the isomorphism class of the monodromy group is independent of the choice of base point.

(ii) Computation of the monodromy group seems daunting, since there are usually uncountably many loops based at a point. However, the monodromy theorem implies that $\sigma_\alpha = \sigma_\beta$ if $\alpha \simeq \beta$; that is, σ_α only depends on the homotopy class of α. This makes computation much more tractable.

This section finishes with the easiest example.

Example 8.4. The power map $p_n(z) = z^n$ defines a regular covering map $\mathbb{C}_* \to \mathbb{C}_*$, for any natural number n. Let $\gamma(t) = e^{2\pi i t}$, the standard clockwise loop around the unit circle in $\mathbb{C}_*$. Let ζ_n be a primitive nth root of unity, and let $\tilde{\gamma}_k$ be the unique lift of γ at ζ_n^k. By uniqueness of lifts,

$$\tilde{\gamma}_k(t) = \exp\left(2\pi i(k+t)/n\right)\,;$$

in particular, $\tilde{\gamma}_k(1) = \zeta_n^{k+1}$. Therefore, γ naturally defines a permutation $\sigma_\gamma \in \mathrm{Sym}(n)$ via

$$\tilde{\gamma}_k(1) = \zeta_n^{\sigma_\gamma(k)}.$$

Any loop in $\mathbb{C}_*$ starting and ending at 1 is homotopic to γ^n for some $n \in \mathbb{Z}$. Therefore, the regular covering map p_n and the choice of base point 1 defines a subgroup of $\mathrm{Sym}(n)$, namely $\langle \sigma_\gamma \rangle \cong C_n$.

See question 9 of Example Sheet 2 for an example of a similar computation.

8.2 The space of germs

The main construction of this section associates a space to any domain – the *space of germs*. The path components of the space of germs are Riemann surfaces, which can be thought of as the result of all possible analytic continuations. The construction makes sense for any domain $D \subseteq \mathbb{C}$.

Definition 8.5. Let (f, U) and (g, V) be function elements on D. For any $z \in D \cap E$, write

$$(f, U) \equiv_z (g, V)$$

if f and g agree on a neighbourhood of z.

Note that $\equiv_z$ is an equivalence relation on the set of function elements (f, U) such that $z \in U$.

Definition 8.6 (Germ)**.** Let (f, U) be a function element and $z \in U$. The equivalence class of (f, U) under $\equiv_z$ is called the *germ* of f at z, and is denoted by $[f]_z$. In summary, two germs $[f]_z$ and $[g]_w$ are equal if and only if $z = w$ and $f = g$ on a neighbourhood of $z = w$.

Our main task is to give the collection of all possible germs the structure of a nice topological space – ultimately, a disjoint union of Riemann surfaces.

Definition 8.7 (The space of germs)**.** The *space of germs over D* is

$$\mathcal{G} := \{[f]_z \mid z \in D,\ (f, U) \text{ a function element with } z \in U\}$$

as a set.

To endow $\mathcal{G}$ with a topology, it is convenient to write

$$[f]_U := \{[f]_z \mid z \in U\}$$

for any function element U. The open sets of the topology are all unions of all sets of the form $[f]_U$, for all function elements (f, U) on D. Let us check that this is indeed a topology.

Lemma 8.8. *Unions of sets of the form $[f]_U$ define a topology on $\mathcal{G}$.*

Proof. Taking the empty union shows that $\varnothing$ is open. By definition, each point $[f]_z \in [f]_U$ for some U, whence $\mathcal{G}$ is also open. The topology is closed under taking unions by definition, so it only remains to check that it is closed under finite intersections.

Since the intersection of two unions is the union of the intersections, to check that the topology is closed under finite intersections, it suffices to check that a set of the form $[f]_U \cap [g]_V$ is open. Consider any germ $[h]_z \in [f]_U \cap [g]_V$. This means that $z \in U \cap V$ and that h agrees with both f and g on a neighbourhood W of z. Thus $[h]_W$ provides an open neighbourhood of $[h]_z$ in the intersection, so the intersection is open claimed. □

If the components of $\mathcal{G}$ are to be Riemann surfaces, they must in particular be Hausdorff.

Lemma 8.9. *The space of germs $\mathcal{G}$ is Hausdorff.*

Proof. Consider two distinct germs $[f]_z \neq [g]_w$, and choose representative function elements (f, U) and (g, V). If $z \neq w$ then, by shrinking U and V, we may assume that U and V are disjoint, whence $[f]_U \cap [g]_V = \varnothing$, as required.

The case $z = w$ is all that remains, in which case we may take representative function elements (f, U) and (g, U) for U connected. Suppose that $[h]_x \in [f]_U \cap [g]_U$. This implies that x has a neighbourhood W in U on which f and g both agree with h, and hence with each other. By the identity principle, f and g agree on U, so $[f]_U = [g]_U]$ whence $[f]_z = [g]_z$, a contradiction. □

To define a conformal structure on each component of $\mathcal{G}$, we will use a useful extra piece of structure – the natural 'forgetful' map to D.

Definition 8.10. Let $\mathcal{G}$ be the space of germs over a domain D. The *forgetful map* $\pi : \mathcal{G} \to D$ is defined by

$$\pi([f]_z) = z\,.$$

Lemma 8.11. *For each component $G \subseteq \mathcal{G}$, the restriction of the forgetful map*

$$\pi : G \to D$$

is a covering map

Proof. Let $U \subseteq D$ be open. The preimage of U under π is

$$\pi^{-1}(U) = \bigcup_{V \subseteq U} [f]_V$$

where the union ranges over all function elements on U. In particular, π is continuous.

The restriction of π to any open set of the form $[f]_U$ has an inverse, namely $z \mapsto [f]_z$. Furthermore, this inverse is continuous, since the preimage of an open set $[f]_V$ is the open set $V \cap U$. Since the sets $[f]_U$ cover $\mathcal{G}$, π is indeed a local homeomorphism, and so its restriction to any component is a covering map. □

By Lemma 4.1, each component of $\mathcal{G}$ has a unique conformal structure that makes π into an analytic map. An explicit atlas is easy to write down, with each chart of the form $(\pi|_U, [f]_U)$.

As well as the forgetful map, the space of germs carries another naturally defined map.

Definition 8.12. Let $\mathcal{G}$ be the space of germs on a domain D. The *evaluation map* $\mathcal{E} : \mathcal{G} \to \mathbb{C}$ is defined by

$$\mathcal{E}([f]_z) = f(z)\,.$$

Remark 8.13. With respect to a standard chart $(\pi|_U, [f]_U)$, the evaluation map takes the form

$$\mathcal{E} \circ (\pi|_U)^{-1}(z) = f(z)$$

which is analytic, since f is. Hence $\mathcal{E}$ (or, more precisely, its restriction to each component of $\mathcal{G}$) is analytic.

Lecture 9: Uniqueness of analytic continuation and gluing

9.1 Analytic continuation revisited

When combined with the theory of covering spaces, the space of germs constructed in the last lecture gives a very clean account of analytic continuation.

Theorem 9.1. *Let (f,U) and (g,V) be function elements on a domain $D \subseteq \mathbb{C}$, and let $\gamma : [0,1] \to D$ be a path starting in U and ending in V. Then $(f,U) \approx_\gamma (g,V)$ if and only if the lift $\tilde{\gamma}$ to (a component of) $\mathcal{G}$ starting at $[f]_{\gamma(0)}$ exists, and ends at $[g]_{\tilde{\gamma}(1)}$.*

Proof. For the 'only if' direction, suppose that $(f,U) \approx_\gamma (g,V)$. That is, there is a sequence of direct analytic continuations as in the definition.

$$(f,U) = (f_1,U_1) \sim \ldots \sim (f_{n-1},U_{n-1}) \sim (f_n,U_n) = (g,V)\,,$$

a continuous path $\gamma : [0,1] \to D$, and a dissection

$$0 = t_0 < t_1 < \ldots < t_{n-1} < t_n = 1$$

such that $\gamma([t_{i-1},t_i]) \subseteq U_i$ for each $1 \le i \le n$. Define a lift by

$$\tilde{\gamma}(t) = [f_i]_{\gamma(t)}$$

whenever $t \in [t_{i-1},t_i]$, which is well-defined since $[f_i]_{\gamma(t_i)} = [f_{i+1}]_{\gamma(t_i)}$ for each $0 < i < n$. For continuity, note that

$$\tilde{\gamma}(t) = (\pi|_{[f_i]})^{-1} \circ \gamma(t)$$

for $t \in [t_{i-1},t_i]$, so $\tilde{\gamma}$ is continuous on each interval in the dissection and hence continuous on $[0,1]$.

For the 'if' direction, suppose that there is a lift $\tilde{\gamma}$ of γ to $\mathcal{G}$ such that $\tilde{\gamma}(0) = [f]_{\gamma(0)}$ and $\tilde{\gamma}(1) = [g]_{\gamma(1)}$. By the compactness of $[0,1]$, there is a finite sequence of function elements (f_i,U_i) for $1 \le i \le n$ and a dissection

$$0 = t_0 < t_1 < \ldots < t_{n-1} < t_n = 1$$

such that $\tilde{\gamma}([t_{i-1},t_i]) \subseteq [f_i]_{U_i}$ for each $1 \le i \le n$. Indeed, we may assume that each U_i is an open disc in $\mathbb{C}$. Applying the forgetful map π, it follows that

$\gamma([t_{i-1}, t_i]) \subseteq U_i$, so it remains only to prove that $(f_{i-1}, U_{i-1}) \sim (f_i, U_i)$ for $1 < i \leq n$. For each such i,

$$[f_{i-1}]_{\gamma(t_{i-1})} = \tilde{\gamma}(t_{i-1}) = [f_i]_{\gamma(t_{i-1})}$$

so f_i and f_{i-1} agree on a neighbourhood of $\gamma(t_{i-1}) \in U_{i-1} \cap U_i$. Since the U_i are discs, $U_{i-1} \cap U_i$ is connected, so it follows by the identity principle that f_{i-1} and f_i agree on the whole intersection $U_{i-1} \cap U_i$. Therefore $(f_{i-1}, U_{i-1}) \sim (f_i, U_i)$ as required. □

Theorem 9.1 immediately implies that we can give a much more concrete description of complete analytic functions.

Corollary 9.2. *Let $\mathcal{F}$ be a complete analytic function on a domain $D \subseteq \mathbb{C}$. Then*

$$\mathcal{G}_{\mathcal{F}} := \bigcup_{(f,U) \in \mathcal{F}} [f]_U$$

is a path component of $\mathcal{G}$.

Since the space of germs is equipped with the evaluation map $\mathcal{E}$, this result tells us that a complete analytic function on a domain D is essentially the same thing as a Riemann surface R equipped with a covering map $\pi : R \to D$ and an analytic function $R \to \mathbb{C}$.

Definition 9.3. The component $\mathcal{G}_{\mathcal{F}}$ is the *Riemann surface associated to $\mathcal{F}$*.

9.2 The classical monodromy theorem

With these tools in hand, it becomes easy to prove the uniqueness results we want about analytic continuation. Specifically, analytic continuation along a path only depends on the homotopy class of the path.

Theorem 9.4 (Classical monodromy theorem)**.** *Let $D \subseteq \mathbb{C}$ be a domain. Suppose that (f, U) is a function element in D and can be continued along any path in D starting in U. If $(f, U) \approx_\alpha (g_1, V)$ and $(f, U) \approx_\beta (g_2, V)$ and $\alpha \simeq \beta$ then $g_1 \equiv g_2$ on V.*

Proof. Let $\tilde{\alpha}, \tilde{\beta}$ be the lifts to $\mathcal{G}$ of α, β respectively, starting at $[f]_{\alpha(0)} = [f]_\beta(0)$. By Theorem 9.1 and the monodromy theorem,

$$[g_1]_{\alpha(1)} = \tilde{\alpha}(1) = \tilde{\beta}(1) = [g_2]_{\beta(1)}$$

so g_1 and g_2 agree on a neighbourhood of $\alpha(1) = \beta(1)$. Therefore, g_1 and g_2 agree on V by the identity principle. □

In particular, analytic continuation is unique on any simply connected domain.

Corollary 9.5. *Let D be a simply connected domain and (f,U) a function element on D. If (f,U) can be analytically continued along every path in D starting in U then (f,U) extends to an analytic function $f : D \to \mathbb{C}$.*

9.3 Gluing Riemann surfaces

In Example 4.3, we analysed kth roots on $\mathbb{C}$ by constructing the associated Riemann surface R_k, together with a covering map $\pi : R_k \to \mathbb{C}_*$ and an analytic function $g : R_k \to \mathbb{C}_*$. We also noted that R_k can be embedded in a *compact* Riemann surface $\widehat{R}_k$. But how can we do this in practice? The answer is by *gluing*.

Definition 9.6. Let X, Y be topological spaces. Suppose we have subspaces $X' \subseteq X$ and $Y' \subseteq Y$, and a homeomorphism $\Phi : X' \to Y'$. The quotient space

$$Z := (X \sqcup Y)/\sim$$

where $\sim$ is the finest equivalence relation such that $x \sim \Phi(x)$ for all $x \in X'$, is called the result of *gluing X and Y along* Φ. It may also sometimes be denoted by $X \cup_\Phi Y$, or even by $X \cup_{X'} Y$ if the map Φ is implicit.

In the context of this course, we particularly want to know when the result of gluing produces a Riemann surface.

Proposition 9.7. *Let R_1, R_2 be Riemann surfaces. Suppose that $S_j \subseteq R_j$ (for $j = 1,2$) are non-empty, connected, open subsets, and $\Phi : S_1 \to S_2$ is a conformal equivalence of Riemann surfaces. There is a unique conformal structure on*

$$R = R_1 \cup_\Phi R_2$$

such that the inclusion maps $i_j : R_j \hookrightarrow R$ are analytic. In particular, if the resulting gluing R is Hausdorff then it is a Riemann surface.

Proof. For $j = 1,2$, every chart (ϕ_j, U_j) on R_j gives rise to a chart $(\phi_j \circ i_j^{-1}, i_j(U))$ on R. By construction these charts cover R. The transition functions between two charts arising from R_j are just transition functions of R_j, hence analytic. If (ϕ_1, U_1) and (ϕ_2, U_2) are charts on R_1 and R_2 respectively, giving rise to charts $(\phi_1 \circ i_1^{-1}, i_1(U))$ and $(\phi_2 \circ i_2^{-1}, i_2(U))$, then the resulting transition function is

$$\phi_2 \circ i_2^{-1} \circ i_1 \circ \phi_1^{-1} = \phi_2 \circ \Phi \circ \phi_1^{-1} ,$$

which is analytic because Φ is a conformal equivalence.

To prove the uniqueness statement, suppose that $\mathcal{A}$ is any conformal structure on R that makes i_j analytic. If (ϕ_j, U_j) is a chart on R_j and $(\psi, V) \in \mathcal{A}$, then

$$\psi \circ i_j \circ \phi_j^{-1}$$

is analytic. Thus $(\phi_j \circ i_j^{-1}, i_j(U))$ has analytic transition functions with every chart in $\mathcal{A}$, so is contained in $\mathcal{A}$ by maximality. This proves uniqueness.

Finally, note that, since R_1 and R_2 are path-connected and the S_i are non-empty, the gluing R is also path-connected. Since we have assumed that R is Hausdorff, it follows that it is a Riemann surface. □

It is an irritation that non-Hausdorff spaces are quite easy to construct by gluing, as the following example shows.

Example 9.8. Let $R_1 = R_2 = \mathbb{C}$, let $S_1 = S_2 = \mathbb{C}_*$, and let $\Phi : S_1 \to S_2$ be the identity map. The gluing

$$R = \mathbb{C} \cup_{\mathbb{C}_*} \mathbb{C}$$

is not Hausdorff, since the two copies of 0 do not have disjoint open neighbourhoods.

Therefore, we always have to check Hausdorffness of a gluing by hand.

Gluing is a convenient way to construct compactifications. For instance, we can use gluing to compactify the complex plane to the Riemann sphere.

Example 9.9. Let $R_1 = R_2 = \mathbb{C}$, let $S_1 = S_2 = \mathbb{C}_*$, and let $\Phi : S_1 \to S_2$ be the inversion map $z \mapsto 1/z$. Every pair of points in the gluing

$$R = \mathbb{C} \cup_\Phi \mathbb{C}$$

is contained in either R_1 or R_2, except for the pair $\{i_1(0), i_2(0)\}$. Therefore, to check that R is Hausdorff, we only need to check that this pair have disjoint open neighbourhoods. Indeed, $i_1(\mathbb{D})$ and $i_2(\mathbb{D})$ is a disjoint pair of open neighbourhoods, so R is Hausdorff. By Proposition 9.9, R is therefore a Riemann surface, and indeed it is easily seen to be the Riemann sphere $\mathbb{C}_\infty$. Furthermore, this approach makes it clear that R is compact: $R = i_1(\overline{\mathbb{D}}) \cup i_2(\overline{\mathbb{D}})$, so is the continuous image of two closed discs.

Lecture 10: More gluing and branching

10.1 A more detailed gluing example

In this section, we give a more involved example of compactification by gluing. In Example 6.3, we constructed the Riemann surface R associated to the complete analytic function $\sqrt{z^3 - z}$, and noticed that the result was a torus with four points removed. We would like to be able to compactify this Riemann surface to a compact torus.

In Example Sheet 1, question 14, you saw that the graph

$$R_1 = \{(z, w) \in \mathbb{C}^2 \mid w^2 = z^3 - z\},$$

is conformally equivalent to R with three points added.[1] To complete the compactification of R, it remains to 'add a point at infinity'.

Example 10.1. First, we perform a change of coordinates to send the point at infinity to a finite point, as in the example of the Riemann sphere. Setting $u = 1/z$ and $v = z/w$, we can rearrange $w^2 = z^3 - z$ to obtain the equation $u = v^2(1 - u^2)$.[2] Therefore, consider the graph

$$R_2 = \{(u, v) \in \mathbb{C}^2 \mid u = v^2(1 - u^2)\}.$$

We next need to check that R_2 is indeed a Riemann surface and carries a conformal structure. As in question 14 of Example Sheet 1, the idea is to define an atlas on R_2 using the two coordinate projections, $\pi(u, v) = u$ and $\tau(u, v) = v$.

Noting that $u \neq \pm 1$, the defining equation of R_2 can be rewritten as

$$v = v(u) = \pm\sqrt{\frac{u}{1 - u^2}}.$$

The projection π therefore has local inverses defined by

$$\pi^{-1}(u) = (u, v(u))$$

[1] In the example sheet, the graph is denoted by R, and R_0 is the result of deleting three points.

[2] We won't discuss how we found this change of coordinates further in this course, so you will have to think of it as found by trial and error. In questions, changes of coordinates will usually be given. If you take *Part II Algebraic Geometry*, you will learn more about where these kinds of changes of coordinates come from.

for suitable choices of branches of the square root. These branches can be chosen unless $u = 0$, in which case $v = 0$ too. Thus, restrictions of π provide charts on the whole of $R_2 \setminus \{(0,0)\}$.

Thus, it remains to provide a chart about $(0,0)$, and for this we use the other coordinate projection, τ. The defining equation of the graph can also be rewritten as

$$v^2u^2 + u - v^2 = 0\,,$$

so τ has local inverses

$$\tau^{-1}(v) = (u(v), v)$$

where

$$u(v) = \frac{-1 \pm \sqrt{1+4v^4}}{2v^2}$$

for a suitable choice of branch of square root. This is not defined at $v = 0$, but has a removable singularity there if we make the correct choice of branch. Indeed, about $v = 0$, taking the positive branch of the square root, we have the power series expansion

$$u(v) = \frac{1 + 4v^4/2 + o(v^8) - 1}{2v^2} = v^2 + o(v^6)\,.$$

Thus, τ has an inverse in a neighbourhood of $\{(0,0)\}$, and so defines a chart there. Taken together with restrictions of π, we have charts about every point of R_2, and the transition functions for this collection of charts, apart from the identity, are the analytic maps $u(v)$ and $v(u)$. Therefore, they define an atlas, and hence a conformal structure. Furthermore, R_2 is a subspace of $\mathbb{C}^2$, hence is Hausdorff. Below, we shall see that R_2 is obtained from R_1 by deleting one point and adding one point; from this it follows that R_2 is connected. Hence, R_2 is a Riemann surface.

The map Φ from a subset of R_1 to a subset of R_2 is given by the change of coordinates, which is to say that

$$\Phi(z,w) = (1/z, z/w) \text{ and } \Phi^{-1}(u,v) = (1/u, 1/uv)\,.$$

The domain of Φ is

$$S_1 = \{(z,w) \in R_1 \mid z, w \neq 0\} = R_1 \setminus \{(0,0), (1,0), (-1,0)\}\,,$$

and the domain of Φ^{-1} is

$$S_2 = \{(u,v) \in R_2 \mid u, v \neq 0\} = R_2 \setminus \{(0,0)\}\,.$$

Since the atlases on R_1 and R_2 are given by coordinate projections, Φ and Φ^{-1} are given in local charts by their coordinate formulae, which are analytic functions, so Φ is a conformal equivalence.

Using all of the above, we may define the gluing

$$\widehat{R} = R_1 \cup_\Phi R_2 \,.$$

We will shortly check that this is Hausdorff, but first, notice that the meromorphic functions $\hat{\pi}_1(z,w) = z$ on R_1 and $\hat{\pi}_2(u,v) = 1/u$ on R_2 satisfy

$$\hat{\pi}_2 \circ \Phi = \hat{\pi}_1$$

on S_1, and so descend to a meromorphic function $\hat{\pi} : \widehat{R} \to \mathbb{C}_\infty$. This function makes it easy to prove that $\widehat{R}$ is Hausdorff. Indeed, since S_1 and S_2 are identified in $\widehat{R}$, it suffices to find a disjoint pair of open neighbourhoods of $i_1(0,0), i_1(\pm 1, 0)$ and $i_2(0,0)$, and $\hat{\pi}^{-1}(\{|z| < 2\})$ and $\hat{\pi}^{-1}(\{|z| > 2\} \cup \{\infty\})$ are such a pair.

Therefore, by Proposition 9.7, $\widehat{R}$ is a Riemann surface. Let's prove that it is (sequentially) compact. For any sequence of points $(p_n)_{n \geq 0}$ of R_1, if $\hat{\pi}(p_n)$ remains bounded then, after passing to a subsequence, we may assume that $\hat{\pi}(p_n)$ converges to some limit $z_0 \in \mathbb{C}$. But there are at most two points in R_1 with z-coordinate equal to z_0 so, after passing to a further subsequence, it follows that (p_n) converges in R_1. On the other hand, if $\hat{\pi}(p_n)$ is unbounded then, after passing to a subsequence, $\hat{\pi}(p_n) \to \infty$, so p_n converges to $i_2(0,0)$, which is the unique point p of $\widehat{R}$ with $\hat{\pi}(p) = \infty$.

In summary, we have compactified R_1 to a Riemann surface $\widehat{R}$, in such a way that π extends to a meromorphic function on $\widehat{R}$. Similarly, the function g on R_1 also extends to a meromorphic function on $\widehat{R}$; the details are left to the reader.

In the case of Example 10.1, we can understand the topology of $\widehat{R}$ directly, since we visualised it via a gluing construction in Example 6.3: it is homeomorphic a torus. In general, however, these kinds of computations are difficult to perform. It turns out that the key to determining the topology of a Riemann surface $\widehat{R}$ is to understand the points where a meromorphic function like $\hat{\pi}$ fails to be a local homeomorphism. This phenomenon is called *branching*, and is the topic of our next section.

10.2 Branching

Compactifying Riemann surfaces often leads to meromorphic functions that are no longer covering maps. For instance, although the power map $p_k(z) = z^k$ is a covering map $\mathbb{C}_* \to \mathbb{C}_*$, its natural extension to the Riemann sphere $\hat{p}_k : \mathbb{C}_\infty \to \mathbb{C}_\infty$ fails to be a local homeomorphism at 0 and ∞ if $k \geq 2$.

Definition 10.2. Let $f : R \to S$ be an analytic map of Riemann surfaces and let $p \in R$. By Proposition 4.6, there are choices of charts (ϕ, U) about p and (ψ, V) about $f(p)$, with $\phi(p) = 0$, such that

$$\psi \circ f \circ \phi^{-1}(z) = z^{m_f(p)}$$

for some integer $m_f(p) \geq 0$. Note that the integer $m_f(p)$ is equal to the number of preimages in a sufficiently small neighbourhood of p of any point in a sufficiently small punctured neighbourhood of $f(p)$, and so is independent of the choice of charts. This is the *multiplicity* of f at p.

Unless f is constant, 'most' points have ramification index equal to 1. The remaining points are especially interesting, and so we introduce special terminology for them.

Definition 10.3. If $m_f(p) > 1$ then p is called a *ramification point* and $f(p)$ is called a *branch point* of f. In this case, the multiplicity $m_f(p)$ is also sometimes called the *ramification index* of p.

Example 10.4. Let $\hat{p}_k : \mathbb{C}_\infty \to \mathbb{C}_\infty$ be the power map $z \mapsto z^k$ for $k \geq 2$, thought of as meromorphic function on the Riemann sphere. The only points with multiplicity greater than 1 are 0 and ∞, so these are the ramification points, each of which has ramification index equal to k. The branch points are their images, which are also 0 and ∞.

Arbitrary polynomials behave similarly.

Example 10.5. Let $f : \mathbb{C}_\infty \to \mathbb{C}_\infty$ be any polynomial

$$f(z) = a_d z^d + \ldots + a_1 z + a_0$$

with $a_d \neq 0$. If $w = \phi(z) = 1/z$ is the chart about infinity then

$$\begin{aligned} \phi \circ f \circ \phi^{-1}(w) &= 1/(a_d w^{-d} + a_{d-1} w^{1-d} + \ldots + a_0) \\ &= w^d/(a_d + a_{d-1} w + \ldots + a_0 w^d) \end{aligned}$$

which shows that $m_f(\infty) = d$, since $1/(a_d + \ldots + a_0 w^d)$ is analytic and non-zero in a neighbourhood of 0. That is, for any polynomial, the multiplicity of ∞ is the degree d.

For more complicated maps it may not be easy to immediately see which points are ramification points. However, differentiation gives a convenient method for analytic functions.

Remark 10.6. Let f be a non-constant analytic function on a Riemann surface R, let $p \in R$ and let (ϕ, U) be any chart about p with $\phi(p) = z_0$. Then

$$F = f \circ \phi^{-1}(z) = (z - z_0)^{m_f(p)} g(z)$$

for some analytic function g with $g(z_0) \neq 0$. Hence,

$$F'(z) = (m_f(p) g(z) + (z - z_0) g'(z))(z - z_0)^{m_f(p)-1}$$

by the product rule. Therefore, we have

$$F'(z_0) = g(z_0) \neq 0$$

if $m_f(p) = 1$, whereas we have $F'(z_0) = 0$ if $m_f(p) > 1$. In summary, the ramification points of an analytic function are exactly the points where the derivative vanishes, in any choice of local coordinates.

The next lemma tells us that multiplicity behaves as you might hope under composition of maps.

Lemma 10.7. *If $f : R \to S$ and $g : S \to T$ are analytic functions of Riemann surfaces then*

$$m_{g \circ f}(p) = m_g(f(p)) m_f(p)$$

for any point $p \in R$.

Proof. Fix any chart (θ, W) about $g \circ f(p)$, with $\theta \circ g \circ f(p) = 0$. By Proposition 4.6, there is a chart (ψ, V) about $f(p)$ such that

$$\theta \circ g \circ \psi^{-1}(z) = z^{m_g(f(p))}$$

on a neighbourhood of 0. Likewise, there is a chart (ϕ, U) about p such that

$$\psi \circ f \circ \phi^{-1}(z) = z^{m_f(p)}$$

on a neighbourhood of 0, by the same proposition. Therefore,

$$\theta \circ (g \circ f) \circ \phi^{-1}(z) = (\theta \circ g \circ \psi^{-1}) \circ (\psi \circ f \circ \phi^{-1})(z) = (z^{m_f(p)})^{m_g(f(p))}$$

and the result follows. $\square$

Lecture 11: The valency and Riemann–Hurwitz theorems

11.1 The valency theorem

The next theorem, the main result of this section, is our first step towards using analytic maps to determine the topology of surfaces. It says that any non-constant map f between compact Riemann surfaces is n-to-1, for some well-defined integer n, as long as we count with multiplicities.

Theorem 11.1 (Valency theorem)**.** *Suppose that $f : R \to S$ is a non-constant, analytic map between compact Riemann surfaces. The function $n : S \to \mathbb{N}$ defined by*

$$n(q) := \sum_{p \in f^{-1}(q)} m_f(p)$$

is constant on S.

Proof. Since R is compact, each $q \in S$ only has finitely many pre-images in R, by the identity principle, so n is indeed a well-defined map to $\mathbb{N}$.

Since R is connected, it suffices to prove that n is locally constant. Therefore, taking an arbitrary $q_0 \in S$ with $n(q_0) = n_0$, it suffices to find an open neighbourhood of q_0 on which $n(q) = n_0$. Let

$$f^{-1}(q_0) = \{p_1, \ldots, p_k\},$$

and fix a chart (ψ, V) about q_0. By Proposition 4.6, there is a chart (ϕ_i, U_i) about each p_i, such that

$$\psi \circ f \circ \phi_i^{-1}(z) = z^{m_f(p_i)}$$

on U_i. Passing to smaller charts if necessary, we may furthermore assume that the $\{U_i\}$ are disjoint. Now $R \setminus \bigcup_i U_i$ is closed, hence compact so the image $K = f(R \setminus \bigcup_i U_i)$ is also compact, hence closed. Therefore, there is a connected open neighbourhood $V' \subseteq V$ of q_0 that is disjoint from K. Thus

$$f^{-1}(V') \subseteq R \setminus f^{-1}(K) \subseteq R \setminus (R \setminus \bigcup_i U_i) = \bigcup_i U_i .$$

Taking $U_i' = f^{-1}(V') \cap U_i$, we obtain charts (ϕ_i, U_i') about p_i and (ψ, V') about q such that the local form of f is a power map everywhere on the preimage of V'. In particular, $n(q) = n$ for all $q \in V'$, as required. □

Evidently the constant n is an important invariant of f.

Definition 11.2. The constant n that Theorem 11.1 associates to a non-constant analytic map f of compact Riemann surfaces is called the *degree* or *valency* of f, and denoted by $\deg(f)$.

Note that, by Example 10.5, this notion of degree coincides with the usual one when f is a polynomial. In any course on complex analysis It is traditional to give a proof of the fundamental theorem of algebra. We could have done so several times already, but this is a particularly good moment, since it is now an easy corollary of the valency theorem.

Corollary 11.3 (Fundamental theorem of algebra)**.** *Any polynomial f of degree d has exactly d zeroes, counted with multiplicity.*

Proof. By Example 10.5, the valency of f is d. The result now follows from the valency theorem. □

11.2 Euler characteristic

In this section, we build on the valency theorem, and see how to use the branching data of analytic maps to determine the topology of compact Riemann surfaces. First, we have to briefly review the topological classification of compact surfaces. The crucial invariant is the *Euler characteristic*.

Definition 11.4. Let S be a compact Riemann surface. A *topological triangle* in S is a continuous embedding $\Delta \hookrightarrow S$, where Δ is a closed triangle in the plane $\mathbb{R}^2$. A *triangulation* of S is a finite collection of topological triangles $\{\Delta_i\}$ satisfying the following conditions.

(i) The union $\bigcup_i \Delta_i$ is the whole of S.

(ii) Unless $i = j$, the intersection $\Delta_i \cap \Delta_j$ is either empty, a vertex or an edge.

(iii) Each edge is an edge of exactly two triangles.

The *Euler characteristic* of the triangulation $\{\Delta_i\}$ is defined to be

$$\chi = V - E + F$$

where F is the number of triangles Δ_i, E is the number of edges, and V is the number of vertices.

These ideas can be applied to Riemann surfaces, because of two important facts, which we do not prove in this course. They are summarised in the following theorem.

Theorem 11.5. *Every compact Riemann surface S admits a triangulation, and the Euler characteristic χ is independent of the choice of triangulation.*

The second assertion – that χ is independent of the choice of triangulation – is proved in *Part II Algebraic Topology*. In light of the theorem, we may write $\chi(S)$ for the Euler characteristic of S.

Example 11.6. The Riemann sphere is homeomorphic to a tetrahedron. Therefore

$$\chi(\mathbb{C}_\infty) = 4 - 6 + 4 = 2\,.$$

Example 11.7. Subdivide the square into nine squares, then divide each of these into two triangles. Identifying opposite sides of the square in the usual way, this exhibits a triangulation of the torus $S^1 \times S^1$. Counting the number of faces, edges and vertices, it follows that

$$\chi(\mathbb{C}/\Lambda) = 9 - 18 + 9 = 0\,,$$

for any complex torus $\mathbb{C}/\Lambda$.

The sphere and the torus are, in a sense, the simplest compact Riemann surfaces. The number of 'holes' is called the *genus*: the sphere is the surface of genus 0, and the torus is the surface of genus 1. A third important fact, which again will remain unproved, is that every compact Riemann surface is homeomorphic to some surface of genus g, for some g.

Theorem 11.8 (Topological classification of Riemann surfaces)**.** *Every compact Riemann surface S is homeomorphic to the surface of genus g, for some g. The Euler characteristic of S is then*

$$\chi(S) = 2 - 2g\,.$$

In particular, the Euler characteristic determines the topological type of any compact Riemann surface. That is, if R and S are Riemann surfaces such that $\chi(R) = \chi(S)$, then R and S are homeomorphic.

11.3 The Riemann–Hurwitz theorem

In this section we will state, and sketch the proof of, the Riemann–Hurwitz theorem. This is the key tool that enables us to deduce topological information from the branching data of an analytic map.

Theorem 11.9 (Riemann–Hurwitz). *Let $f : R \to S$ be any non-constant analytic map of compact Riemann surfaces. Then*

$$\chi(R) = \deg(f)\chi(S) - \sum_{p\in R}(m_f(p) - 1)\,.$$

Note that $m_f(p) - 1 > 0$ only if p is a ramification point of f. Since R is compact and ramification points are isolated, there are only finitely many such points, and so the sum in the theorem is finite.

Sketch proof. As in the proof of the valency theorem, each $q \in S$ has a neighbourhood U such that f takes the form of a power map on each component of the preimage of U. These neighbourhoods form an open cover of S so, by compactness, there is a finite subcover

$$\{U_1, \dots, U_k\}\,,$$

where U_i is the neighbourhood associated the point q_i. The only point of U_i that can be a branch point is q_i itself, so there are only finitely many branch points.

Take a triangulation of S. Each triangle contains at most finitely many branch points, so after subdividing further, we may assume that each triangle contains at most one branch point. Subdividing further still, we may assume that each branch point is a vertex of the triangulation.

Now refine the triangulation further still, until each triangle is contained in some V_i. In particular, the preimages of the triangles in R together form a triangulation of R. To relate the Euler characteristics of these triangulations, let's introduce some notation: $n = \deg(f)$; V_S, E_S, F_S respectively denote the number of vertices, edges and triangles in S; likewise, V_R, E_R, F_R respectively denote the number of vertices, edges and triangles in R.

We now count preimages in R.

(i) Each triangle in S has exactly n preimages in R, so $F_R = nF_S$.

(ii) Each edge in S has exactly n preimages in R, so $E_R = nE_S$.

(iii) Each vertex q in S has exactly

$$n - \sum_{p \in f^{-1}(q)} (m_f(p) - 1)$$

preimages in R. (Note that the sum is zero unless q is a branch point.) Therefore,

$$V_R = nV_S - \sum_{q \in S} \sum_{p \in f^{-1}(q)} (m_f(p) - 1) = nV_S - \sum_{p \in R} (m_f(p) - 1)\,.$$

Putting these together completes the proof:

$$\begin{aligned} \chi(R) &= F_R - E_R + V_R \\ &= nF_S - nE_S + nV_S - \sum_{p \in R} (m_f(p) - 1) \\ &= n\chi(S) - \sum_{p \in R} (m_f(p) - 1) \end{aligned}$$

as required. □

Lecture 12: Applications of Riemann–Hurwitz

12.1 Immediate consequences

Last time, we sketched the proof of the Riemann–Hurwitz theorem. Writing g_R and g_S for the genera of R and S respectively, and $n = \deg(f)$, the Riemann–Hurwitz equation can be rewritten as

$$2g_R - 2 = n(2g_S - 2) + \sum_{p \in R} (m_f(p) - 1)\,,$$

which is how we will usually use it.

The most important application is to the computation of the topological types of compact Riemann surfaces. Let's complete the story of our running example by using Riemann–Hurwitz to compute its genus.

Example 12.1. In Example 10.1, we compactified the Riemann surface associated to the complete analytic function $\sqrt{z^3 - z}$, building a compact Riemann surface $\widehat{R}$ equipped with a meromorphic function $\hat{\pi}$. Furthermore, $\hat{\pi}$ has four branch points – $0, \pm 1, \infty$ – each 'totally ramified', meaning that their preimages each consist of exactly one ramification point, of index equal to $\deg(\hat{\pi})$. Since $\deg(\hat{\pi}) = 2$, Riemann–Hurwitz therefore gives

$$2g_{\widehat{R}} - 2 = 2 \times -2 + 4 \times (2-1)\,,$$

which rearranges to give

$$g_{\widehat{R}} = -2 + 2 \times (2-1) + 1 = 1\,.$$

This is consistent with our observation from the gluing construction that $\widehat{R}$ is a torus.

Usually, it will be much easier to apply Riemann–Hurwitz than a specific gluing construction! Riemann–Hurwitz also has many general consequences, which are worth noting. Here is one.

Remark 12.2. The correction term $\sum_{p \in R} m_f(p) - 1$ is even.

This observation is especially useful when looking at meromorphic functions of degree 2. Let's take one last look at our favourite example.

Example 12.3. Let

$$\widehat{R} = R_1 \cup_\Phi R_2$$

be as in Example 10.1, but let's *not* use any information about R_2 – that is, let's not use any explicit information about $\hat{\pi}^{-1}(\infty)$. Knowing that $\hat{\pi}$ is a map of degree two, with three totally ramified branch points in $\mathbb{C}$, the correction term in Riemann–Hurwitz is

$$3 \times (2-1) + C\,,$$

where $C = \sum_{p \in \hat{\pi}^{-1}(\infty)} (m_{\hat{\pi}}(p) - 1)$. In particular, C is odd. Since $\deg(\hat{\pi}) = 2$, there are only two possibilities for $\hat{\pi}^{-1}(\infty)$.

(i) If $\#\hat{\pi}^{-1}(\infty) = 2$ then both preimages are unramified, so ∞ is not a branch point and $C = 0$.

(ii) If $\#\hat{\pi}^{-1}(\infty) = 1$ then ∞ is totally ramified and $C = 1$.

Knowing that C must be even, only the second case is possible, and the computation that $g_{\widehat{R}} = 1$ follows as before.

The moral here is that, for meromorphic functions of degree two, we do not need to explicitly construct the compactification at ∞ to deduce the branching data there – it is enough to know that the compactification exists.

Riemannn–Hurwitz also places restrictions on the kinds of covering maps between Riemann surfaces that can occur.

Remark 12.4. If f is a covering map (also called *unramified*) then the correction term coming from the branching data is zero, so

$$g_R - 1 = n(g_S - 1)\,.$$

(i) If $g_S = 0$ then we must have $n = 1$ and $g_R = 0$. In particular, n is a conformal equivalence.

(ii) If $g_S = 1$ then we also have $g_R = 1$, and there is no constraint on n.

(iii) If $g_S > 1$ then either $n = 1$, in which case f is a conformal equivalence, or $g_R > g_S$.

We do not have the tools to construct examples realising the third item, but our understanding of complex tori makes the second item easy to realise.

Example 12.5. For any integer $n \geq 1$, consider the lattice $\Lambda_n = \langle n, i \rangle$ in $\mathbb{C}$. The natural inclusion $\Lambda_1 \hookrightarrow \Lambda_n$ induces a quotient map of complex tori

$$\mathbb{C}/\Lambda_n \to \mathbb{C}/\Lambda_1$$

which is a covering map of degree n.

12.2 Higher-genus Riemann surfaces

We haven't yet seen any example of Riemann surfaces of genus greater than 1. We'll do this as with the previous example – by writing down an explicit graph and then compactifying.

Example 12.6. Consider the *Fermat curve* of degree d, namely the graph

$$F'_d \coloneqq \{(x, y) \in \mathbb{C}^2 \mid x^d + y^d = 1\}$$

so-called because rational points on this graph correspond to integer solutions to the famous Fermat equation $x^d + y^d = z^d$.

The projection $\pi_x : (x, y) \mapsto x$ has local inverse

$$\pi_x^{-1}(x) = \sqrt[d]{1 - x^d}$$

unless x is a dth root of unity ζ_d^i, in which case $y = 0$. Symmetrically, the projection π_y to the y coordinate is a local homeomorphism except at $(0, \zeta_d^i)$. This collection of charts covers F_d', and the transition functions are either the identity or $x \mapsto \sqrt[d]{1 - x^d}$, hence analytic. Therefore, this defines an atlas, and hence a conformal structure. It is Hausdorff as a subset of $\mathbb{C}^2$.

To check path-connectedness, define a simply-connected domain D in the x-plane by making branch cuts joining the dth roots of unity to ∞. That is:

$$D = \mathbb{C} \smallsetminus \bigcup_{i=1}^{d} \{t\zeta_d^i \mid t \geq 1\}\,.$$

Since D is simply connected, there are well-defined branches of

$$y(x) = \sqrt[d]{1 - x^d}$$

defined on D, which extend continuously to the branch points ζ_d^i. Let $(x_0, y_0) \in F_d'$ with $x_0 \in D$, and choose the branch of $y(x)$ so that $y(x_0) = y_0$. Since D is path-connected, there is a continuous path γ in D from x_0 to 1, so the path

$$\tilde{\gamma}(t) := (\gamma(t), y(\gamma(t)))$$

joins $(1, 0)$ to some (x_0, y_0).

In summary, every point $(x_0, y_0) \in F_d'$ is in the same path component of $(1, 0)$ unless $x_0 \notin D$. If $x_0 \notin D$, the same argument on a small disc about x_0 shows that a short path joins (x_0, y_0) to some $(x_1, y_1) \in F_d'$ such that $x_1 \in D$. Therefore, F_d' is path connected.

On question 13 of Example Sheet 3, you show that F_d' can be compactified to a Riemann surface F_d by adding d points, and π_x extends to a meromorphic function $\hat{\pi}_x$, in such a way that such that

$$\hat{\pi}_x^{-1}(\infty) = F_d \smallsetminus F_d'\,.$$

For a generic choice of $x_0 \in \mathbb{C}$, there are p points in F_d' satisfying $\pi_x(p) = x_0$, so $\deg(\hat{\pi}_x) = d$. Thus, ∞ is not a branch point of $\hat{\pi}_x$, and our computations above showed that the only branch points of π_x are the roots of unity ζ_d^i, which are totally ramified. In summary, $\hat{\pi}_x$ has d ramification points, each of multiplicity d.

From Riemann–Hurwitz applied to $\hat{\pi}_x$ we now obtain the formula

$$2g_d - 2 = d \times -2 + d(d-1)$$

for the genus g_d of F_d. Rearranging this gives

$$g_d = (d-1)(d-2)/2\,.$$

Thus, this construction gives us compact Riemann surfaces of arbitrarily large (but not all) genera.

Lecture 13: Rational and periodic functions

Corollary 9.2 tells us that understanding analytic continuation is equivalent to understanding certain functions on Riemann surfaces. In general this is an impossible task, but there is hope when the Riemann surface is compact. In this lecture, we will classify the meromorphic functions on the Riemann sphere, and then move on try to understand the meromorphic functions on complex tori.

13.1 Rational functions

Riemann–Hurwitz gives us a much better understanding of compact Riemann surfaces, by enabling us to compute their topology. As well as compact Riemann surfaces themselves, we are also interested in analytic and meromorphic functions on those surfaces. Corollary 5.5 tells us that analytic functions will be constant, but the set of meromorphic functions on a compact Riemann surface R can still be very interesting, and we would like to classify them if possible. In this section, we will do that for the simplest possible compact Riemann surface – the Riemann sphere, $\mathbb{C}_\infty$.

Proposition 13.1. *Every meromorphic function on the Riemann sphere $f : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is a rational function. That is, it is of the form*

$$f(z) = c\frac{(z-a_1)\dots(z-a_m)}{(z-b_1)\dots(z-b_n)}$$

for some integers $m, n \geq 0$ and constants $a_i, b_j, c \in \mathbb{C}$.

Proof. Clearly we may assume that f is non-constant. After replacing f by $1/f$ if necessary, we may also assume that $f(\infty) \in \mathbb{C}$. Now $f^{-1}(\infty)$ is a finite set of poles $b_1, \ldots, b_{n'} \in \mathbb{C}$, and f takes the form

$$f(z) = \sum_{l=-k_j}^{\infty} c_{j,l}(z - b_j)^l$$

in a punctured neighbourhood of each pole b_j. Defining Q_j to be the principal part

$$Q_j(z) = \sum_{l=-k_j}^{-1} c_{j,l}(z - b_j)^l \,,$$

we see that all the singularities of the function

$$g(z) = f(z) - \sum_{j=1}^{n'} Q_j(z)$$

are removable. Hence $g(z)$ is non-surjective self-map from $\mathbb{C}_\infty$, and hence constant. Since the Q_j are all rational functions, the result follows. □

Of course, we should assume that the representation

$$f(z) = c\frac{(z - a_1)\ldots(z - a_m)}{(z - b_1)\ldots(z - b_n)}$$

is chosen so that m and n are minimal, or equivalently that the a_i are all distinct from the b_j. The hypothesis that $f(\infty) \in \mathbb{C}$ in the proof of Proposition 13.1 is equivalent to assuming that $m \leq n$, and in this case the proof shows that $\deg(f) = n$.

Remark 13.2. For a rational function

$$f(z) = c\frac{(z - a_1)\ldots(z - a_m)}{(z - b_1)\ldots(z - b_n)}$$

as above, we have $\deg(f) = \max\{m, n\}$.

13.2 Simply periodic functions

Having classified the meromorphic functions on $\mathbb{C}_\infty$, we would like to carry out something similar for other Riemann surfaces. The fact that many Riemann surfaces can be constructed as quotients is useful here, because it enables us to give an alternative description of the meromorphic functions as *periodic* functions on domains in $\mathbb{C}$.

Definition 13.3. Let $f : \mathbb{C} \to \mathbb{C}_\infty$ be meromorphic. A *period* of f is a complex number $\omega \in \mathbb{C}$ such that $f(z+\omega) = f(z)$ for all $z \in \mathbb{C}$.

Note that the set of periods Ω of any meromorphic function f on $\mathbb{C}$ is an additive subgroup of $\mathbb{C}$. We may therefore ask what the subgroup of periods can be, and it turns out that there are not too many possibilities.

Lemma 13.4. *Let Ω be the set of periods of a meromorphic function f on $\mathbb{C}$. One of the following holds:*

(i) $\Omega = \{0\}$*;*

(ii) $\Omega = \langle\omega\rangle \cong \mathbb{Z}$ *for some* $\omega \neq 0$*;*

(iii) $\Omega = \langle\omega_1\rangle \oplus \langle\omega_2\rangle \cong \mathbb{Z}^2$ *for some* ω_1, ω_2 *linearly independent over* $\mathbb{R}$*;*

(iv) $\Omega = \mathbb{C}$*.*

Proof. See question 1 of Example Sheet 3. □

The case of $\Omega = \{0\}$ is of course the 'generic' case, and such f are too complicated to classify. By contrast, the case of $\Omega = \mathbb{C}$ is very simple: it occurs exactly when f is constant. Before moving on to the case of $\Omega \cong \mathbb{Z}^2$, let's briefly consider the case $\Omega \cong \mathbb{Z}$.

Definition 13.5. A meromorphic function f on $\mathbb{C}$ for which the group of periods is isomorphic to $\mathbb{Z}$ is called *simply periodic*.

Some of our favourite examples are simply periodic.

Example 13.6. The exponential function has periods $\langle 2\pi i\rangle$.

Another way to think about the exponential function is as a regular covering map $\exp : \mathbb{C} \to \mathbb{C}_*$, as in Example 3.4. This suggests an alternative interpretation of simply periodic functions.

Proposition 13.7. *If f is a meromorphic function on $\mathbb{C}$ and the periods of f contain an infinite cyclic subgroup $\langle\omega\rangle$, then there is a unique meromorphic function $\bar{f}$ on $\mathbb{C}_*$ such that*

$$f(z) = \bar{f} \circ \exp((2\pi i/\omega)z)$$

for all $z \in \mathbb{C}$.

Proof. On a small open neighbourhood of any point in $\mathbb{C}_*$, choose a branch of the complex logarithm and define

$$\bar{f}(w) = f((\omega/2\pi i)\log(w)),$$

which is evidently a locally defined analytic function that satisfies

$$\bar{f} \circ \exp((2\pi i/\omega)z) = f((\omega/2\pi i)\log(\exp((2\pi i/\omega)z))) = f(z)$$

as required. If $\hat{f}$ is defined in the same way by choosing a different choice of branch, differing from the original choice by $2\pi in$ for some $n \in \mathbb{Z}$, then

$$\begin{aligned} \hat{f}(w) &= f((\omega/2\pi i)(\log(w) + 2\pi in)) \\ &= f((\omega/2\pi i)\log(w) + n\omega)) \\ &= f((\omega/2\pi i)\log(w)) \\ &= \bar{f}(w) \end{aligned}$$

since $n\omega \in \langle\omega\rangle$ is a period of f. Therefore the definition of $\bar{f}$ is independent of the choice of branch of logarithm, which proves the result. □

The moral of Proposition 13.7 is that simply periodic functions are essentially the same things are functions on $\mathbb{C}_*$. Since $\mathbb{C}_*$ is non-compact, classifying meromorphic functions on $\mathbb{C}_*$ is still too much to ask for. However, the idea that functions on a Riemann surface are equivalent to periodic functions on a covering space is very useful, as we shall see in the next section.

13.3 Doubly periodic functions

Definition 13.8. A meromorphic function f on $\mathbb{C}$ with periods $\Omega = \langle\omega_1\rangle \oplus \langle\omega_2\rangle \cong \mathbb{Z}^2$ is said to be *doubly periodic* or *elliptic*.

In this case the period group is a *lattice* Λ, as in Example 5.1, where we constructed the complex tori $\mathbb{C}/\Lambda$. Indeed, replacing exp by the quotient map $\pi : \mathbb{C} \to \mathbb{C}/\Lambda$, there is a precise analogue to Proposition 13.7.

Proposition 13.9. *If f is a meromorphic function on $\mathbb{C}$ and the periods of f contain a lattice Λ, then there is a unique meromorphic function $\bar{f}$ on the complex torus such that*

$$f(z) = \bar{f} \circ \pi$$

for all $z \in \mathbb{C}$.

Proof. The proof is identical to the proof of Proposition 13.7, using π instead of the exponential function and local inverses to π instead of branches of the logarithm. □

Thus, we can study functions on complex tori by studying doubly periodic functions on $\mathbb{C}$. We will often abuse notation by identifying a doubly periodic function f with the induced function $\bar{f}$ on a complex torus. The next result follows immediately from Proposition 13.9 and Corollary 5.5.

Corollary 13.10. *Any analytic function f on $\mathbb{C}$ that is doubly periodic is constant.*

Here's another example of an application of the theory of compact Riemann surfaces, this time using Riemann–Hurwitz. If f is doubly periodic, we define $\deg(f)$ to be the degree of the associated function $\bar{f}$ on a complex torus.

Corollary 13.11. *If f is a doubly periodic, non-constant, meromorphic function then* $\deg(f) \geq 2$.

Proof. If $\deg(f) = 1$ then in particular f has no ramification points, and hence $m_f(p) = 1$ for all p. The Riemann–Hurwitz theorem then gives

$$(2 \times 1 - 2) = 1 \times (2 \times 0 - 2)$$

which is absurd. □

A doubly periodic function f is determined by its values on any *period parallelogram*

$$\mathcal{P} := \{z_0 + t_1\omega_1 + t_2\omega_2\}$$

where $t_1, t_2 \in [0,1]$. (We will usually take $z_0 = 0$, but sometimes it will be convenient to perturb $\mathcal{P}$ slightly.) This point of view offers a more concrete approach to studying doubly periodic functions. To illustrate this, we give another proof of the last result.

Alternative proof of Corollary 13.11. Let $\mathcal{P}$ be a period parallelogram, chosen so that no zeroes or poles lie on the boundary $\partial\mathcal{P}$. (We can perturb $\mathcal{P}$ to make this true, since there are only $\deg(f)$ zeroes or poles by the valency theorem.) The residue theorem now gives

$$\sum_{z \in \mathrm{poles}(f) \cap \mathcal{P}} \mathrm{res}_z(f) = \frac{1}{2\pi i} \oint_{\partial\mathcal{P}} f(z)\,dz\,.$$

If the path α is the straight line from z_0 to $z_0+\omega_1$ and β is the straight line from $z_0+\omega_1$ to $z_0+\omega_1+\omega_2$ then the boundary $\partial\mathcal{P}$ decomposes as the concatenation

$$\alpha\cdot\beta\cdot(\bar{\alpha}+\omega_2)\cdot(\bar{\beta}-\omega_1)\,.$$

Now

$$\int_\alpha f(z)dz+\int_{\bar{\alpha}+\omega_2} f(z)dz=\int_\alpha f(z)dz+\int_{\bar{\alpha}} f(z)dz=0$$

since ω_2 is a period of f. The same holds to the sides parallel to β so we conclude that the sum of the residues is zero. This is only possible if f has at least two poles in $\mathcal{P}$, counted with multiplicity. □

Lecture 14: The Weierstrass $\wp$-function

14.1 The definition

Let $\Lambda=\langle\omega_1,\omega_2\rangle$ be a lattice in $\mathbb{C}$. Can we construct any non-constant functions on a complex torus $\mathbb{C}/\Lambda$, or equivalently, any doubly periodic functions with periods Λ? By Corollary 5.5 such a function must have a pole, and Corollary 13.11 tells us that it must have either two poles or a pole of degree 2. This suggests that we could try to construct a function that is approximately $1/z^2$ in a small disc about 0. A few minutes' thought about the constraints we have imposed on ourselves suggests the following definition.

Definition 14.1. Let Λ be a lattice in $\mathbb{C}^2$. The associated *Weierstrass $\wp$-function* is defined by

$$\wp_\Lambda(z):=\frac{1}{z^2}+\sum_{\omega\in\Lambda\smallsetminus\{0\}}\left(\frac{1}{(z-\omega)^2}-\frac{1}{\omega^2}\right).$$

We will usually write $\wp\equiv\wp_\Lambda$ when the lattice Λ is understood.

We now need to check that this series really does converge. To do this, we use the following lemma.

Lemma 14.2. *Let $\Lambda=\langle\omega_1,\omega_2\rangle$ be a lattice in $\mathbb{C}$ and $t\in\mathbb{R}$. The sum*

$$\sum_{\omega\in\Lambda\smallsetminus\{0\}}\frac{1}{|\omega|^t}$$

converges if and only if $t>2$.

Proof. The set $\{(t_1,t_2) \in \mathbb{R}^2 \mid |t_1| + |t_2| = 1\}$ is compact, so the function

$$(t_1,t_2) \mapsto |t_1\omega_1 + t_2\omega_2|$$

achieves its maximum M and minimum m, which are both non-negative numbers. If $m = 0$ then

$$t_1\omega_1 + t_2\omega_2 = 0$$

for some non-zero t_1 and t_2, which contradicts the hypothesis that ω_1 and ω_2 are linearly independent over $\mathbb{R}$. In summary,

$$0 < m \le |t_1\omega_1 + t_2\omega_2| \le M < \infty$$

whenever $|t_1| + |t_2| = 1$. Now let $(k,l) \in \mathbb{Z}^2$. Setting $t_1 = k/(|k| + |l|)$ and $t_2 = l/(|k| + |l|)$ gives

$$m(|k| + |l|) \le |k\omega_1 + l\omega_2| \le M(|k| + |l|)\,.$$

Therefore the sum we're interested in, namely

$$\sum_{(k,l)\in\mathbb{Z}^2\smallsetminus\{0\}} \frac{1}{|k\omega_1 + l\omega_2|^t}\,,$$

is bounded above and below by constant multiples of the series

$$\sum_{(k,l)\in\mathbb{Z}^2\smallsetminus\{0\}} \frac{1}{(|k| + |l|)^t}\,.$$

Setting $n = |k| + |l|$ and noting that there are precisely $4n$ pairs (k,l) with $|k| + |l| = n > 0$, this sum can be rewritten as

$$\sum_{(k,l)\in\mathbb{Z}^2\smallsetminus\{0\}} \frac{1}{(|k| + |l|)^t} = \sum_{n=1}^{\infty} \frac{4n}{n^t} = 4\sum_{n=1}^{\infty} \frac{1}{n^{t-1}}\,,$$

so the sum converges if and only if $t > 2$, as required. □

Using this result, we can prove that our definition of the $\wp$-function is well defined.

Theorem 14.3. *The function $\wp_\Lambda$ is a well-defined elliptic function with periods Λ. Moreover, $\wp_\Lambda$ is even and of degree 2.*

Proof. First, let's prove convergence of the sum for a fixed z. Start by estimating the summands:

$$\begin{aligned}
\left|\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2}\right| &= \left|\frac{\omega^2-(z-\omega)^2}{\omega^2(z-\omega)^2}\right| \\
&= \left|\frac{z(2\omega-z)}{\omega^2(z-\omega)^2}\right| \\
&= \left|\frac{z}{\omega^2}\right|\left|\frac{(2\omega-z)}{(z-\omega)^2}\right| \\
&\le \frac{|z|}{|\omega|^2}\cdot\frac{2|\omega-z|+|z|}{|z-\omega|^2} \\
&= \frac{|z|}{|\omega|^2}\left(\frac{2}{|z-\omega|}+\frac{|z|}{|z-\omega|^2}\right)
\end{aligned}$$

Fix some $R \ge |z|$. For all but finitely many $\omega \in \Lambda$, $|\omega| \ge 2R$, whence $|\omega - z| \ge |\omega|/2 \ge R$. These summands are therefore bounded above by

$$\frac{R}{|\omega|^2}\left(\frac{4}{|\omega|}+\frac{R}{R|\omega|/2}\right) = \frac{6R}{|\omega|^3}$$

The exponent of 3 implies that the sum defining $\wp_\Lambda$ converges absolutely and uniformly on compact subsets, by Lemma 14.2.

The definition implies immediately that $\wp_\Lambda$ is even.

To prove that $\wp_\Lambda$ is elliptic, we need to prove that each $\omega_0 \in \Lambda$ is a period of $\wp_\Lambda$. Let's start by considering the derivative:

$$\wp'_\Lambda(z) = \sum_{\omega\in\Lambda} \frac{-2}{(z-\omega)^3}\,.$$

Note that his expression as a sum implies immediately that ω_0 is a period of $\wp'_\Lambda$. Therefore, the function

$$\wp_\Lambda(z+\omega_0) - \wp_\Lambda(z) = c$$

for some constant c, because its derivative is zero. Setting $z = -\omega_0/2$ and using the fact that $\wp_\Lambda$ is even gives

$$\wp_\Lambda(\omega_0/2) = c + \wp_\Lambda(-\omega_0/2) = c + \wp_\Lambda(\omega/2)\,,$$

so $c = 0$, and ω_0 is indeed a period of $\wp_\Lambda$ as required.

Finally, note that, by Lemma 14.2, the only poles of $\wp'_\Lambda$ are at the lattice points Λ, so the periods are precisely Λ. In particular, $\wp_\Lambda$ has a unique pole of order 2 on $\mathbb{C}/\Lambda$, so $\deg(\wp_\Lambda) = 2$ as claimed. □

Remark 14.4. The proof shows that $\wp_\Lambda$ has the following properties:

(i) $\wp_\Lambda$ is meromorphic with periods Λ;

(ii) $\wp_\Lambda$ has poles only at Λ;

(iii) $\lim_{z\to 0}(\wp_\Lambda(z) - 1/z^2) = 0$.

Indeed, these properties uniquely characterise $\wp_\Lambda$. If $f(z)$ is any other function satisfying (i) and (ii), then $f(z) - \wp_\Lambda(z)$ is an elliptic function on $\mathbb{C}$ with poles only at the lattice points Λ. But item (iii) implies that $f(z) - \wp_\Lambda(z) \to 0$ as $z \to 0$, so the poles are in fact removable singularities, and $f(z) - \wp_\Lambda(z)$ is constant by Corollary 13.10. This constant is 0 by item (iii) again, so $f = \wp_\Lambda$.

14.2 Branching properties of $\wp_\Lambda$

Our goal for this section is to investigate the branching behaviour of $\wp_\Lambda$. To do this, Remark 10.6 tells us that we should investigate the derivative $\wp_\Lambda$.

Remark 14.5. Since

$$\wp'_\Lambda(z) = \sum_{\omega\in\Lambda} \frac{-2}{(z-\omega)^3}$$

we see that $\wp'_\Lambda$ has poles precisely at the lattice points Λ, and these poles are of order 3. Hence the periods are precisely Λ and $\deg(\wp'_\Lambda) = 3$. It also follows immediately that $\wp'_\Lambda$ is odd.

Therefore, for any $\omega \in \Lambda$,

$$\wp'_\Lambda(\omega/2) = -\wp'_\Lambda(-\omega/2) = -\wp'_\Lambda(\omega/2),$$

by oddness and periodicity, so $\wp'_\Lambda(\omega/2) = 0$. Thus, $\wp'_\Lambda$ has zeroes at the three half-lattice points $\omega_1/2$, $\omega_2/2$ and $(\omega_1 + \omega_2)/2$ in the fundamental parallelogram. By the valency theorem, these are all the zeroes and each is simple, since $\deg(\wp'_\Lambda) = 3$.

This information about the zeroes of $\wp'_\Lambda$ translates into the information we want about the branching of $\wp_\Lambda$.

Remark 14.6. By Remark 10.6, aside from its poles, the ramification points of $\wp_\Lambda$ are the points where the derivative vanishes. Therefore, the ramification points of $\wp_\Lambda$ in the fundamental parallelogram are 0 and the three half-lattice points, and each has multiplicity 2. The corresponding branch points are ∞ and three finite values $\wp_\Lambda(\omega_1/2)$, $\wp_\Lambda(\omega_2/2)$ and $\wp_\Lambda((\omega_1+\omega_2)/2)$. Note that these three values are distinct, by the valency theorem. In the future, we shall let e_1, e_2 and e_3 denote them.

Finally, note that this is consistent with our expectations: Riemann–Hurwitz predicts that, for any elliptic function f of degree 2,

$$0 = 2 \times -2 + \sum_{p \in \mathbb{C}/\Lambda} (m_f(p) - 1)$$

which in turn implies that there must be exactly four ramification points.

14.3 An algebraic relation

Although the we derived $\wp'_\Lambda$ from $\wp_\Lambda$ by an analytic procedure – differentiation – it turns out that there is an algebraic relation between the two functions.

Proposition 14.7. *There are constants $g_1, g_2 \in \mathbb{C}$, depending only on Λ, such that*

$$(\wp'_\Lambda)^2 \equiv 4\wp_\Lambda^3 - g_2\wp_\Lambda - g_3\,.$$

Proof. Since $\wp$ is even, every term in its Laurent expansion about 0 has an even exponent. By item (iii) of Remark 14.4 the constant term is 0, and so

$$\wp(z) = 1/z^2 + az^2 + o(z^4)$$

for some $a \in \mathbb{C}$. Cubing this gives

$$(\wp(z))^3 = 1/z^6 + f(z)$$

where f is analytic in a neighbourhood of 0. Differentiating $\wp$ gives

$$\wp'(z) = -2/z^3 + 2az + o(z^3)$$

and squaring this expression gives

$$(\wp'(z))^2 = 4/z^6 - 8a/z^2 + g(z)$$

where g is another analytic function in a neighbourhood of 0. Therefore,

$$(\wp'(z))^2 - 4(\wp(z))^3 = -8a/z^2 - h(z)$$

for another analytic function h. Setting $g_2 = 8a$, we get that

$$(\wp'(z))^2 - 4(\wp(z))^3 + g_2\wp(z)$$

is doubly periodic, with the only possible poles at Λ, but analytic in a neighbourhood of 0. Hence it is equal to some constant $-g_3$ and the result follows. □

In fact, the constants g_2 and g_3 are related to the branch points e_1, e_2 and e_3 of $\wp$.

Remark 14.8. Recall that e_1, e_2 and e_3 are the images of the half-lattice points under $\wp$, which are the zeroes of $\wp' = 0$. Therefore, the cubic equation

$$4x^3 - g_2x - g_3 = 0$$

has three distinct zeroes, namely e_1, e_2 and e_3, and so we can factor it:

$$4x^3 - g_2x - g_3 = 4(x - e_1)(x - e_2)(x - e_3)\,.$$

Thus, the relation from Proposition 14.7 can be rewritten as follows.

$$(\wp')^2 \equiv 4(\wp - e_1)(\wp - e_2)(\wp - e_3)$$

Since the cubic equation had zero coefficient in x^2, this also tells us that $e_1 + e_2 + e_3 = 0$.

Lecture 15: More $\wp$-function and quotients

15.1 An elliptic curve

In this course, we have seen two constructions of complex tori. On the one hand, we constructed them as quotients $\mathbb{C}/\Lambda$. On the other hand, complex tori have also arisen as compactifications of certain Riemann surfaces defined by graphs, as in Example 10.1. It is reasonable to ask whether these constructions are connected in any way. The next result, remarkably, shows us that every torus constructed as a quotient also arises as a compactification of a graph.

Corollary 15.1. *Let $\mathbb{C}/\Lambda$ be a complex torus. There are constants g_2, g_3 such that $\mathbb{C}/\Lambda$ is biholomorphic to a one-point compactification of the graph*

$$X' := \{(x,y) \in \mathbb{C}^2 \mid y^2 = 4x^3 - g_2 x - g_3\}\,.$$

Sketch proof. Take g_2, g_3 as in Proposition 14.7. As in previous examples, such as Question 14 of Example Sheet 1 or Example 10.1, the x- and y-coordinate projections turn out to define an atlas on X'. Furthermore, as in Example 10.1 there is a compactification $X' \subseteq X$ with a single point at infinity, and the coordinate projections extend to meromorphic functions.

Because the conformal structure on X includes the coordinate projections, the map $F : \mathbb{C} \to X$ defined by

$$F(z) := (\wp(z), \wp'(z))$$

is analytic. Since Λ is the group of periods of $\wp$ and $\wp'$, F descends to a well-defined analytic map

$$\Phi : \mathbb{C}/\Lambda \to X\,,$$

sending the coset $0 + \Lambda$ to ∞.

It remains to show that Φ is a biholomorphism. As a non-constant analytic map of compact Riemann surfaces, Φ is certainly surjective.

To show injectivity, take a period parallelogram $\mathcal{P}$ centred on 0, and suppose that z and z' are $\mathcal{P}$ and that $F(z) = F(z')$. Then

$$\wp(z) = \wp(z')$$

which implies that $z' = \pm z$, because $\wp$ is even and of degree 2. But then

$$\wp'(z) = \wp'(\pm z) = \pm\wp'(z)$$

since $\wp'$ is odd. As long as $z \notin \Lambda/2$, we have that $\wp'(z) \in \mathbb{C}_*$ and so this implies that $z' = z$.

In summary, away from the ramification points of $\wp$, the map Φ has degree 1 on $\mathbb{C}/\Lambda$. Therefore, $\deg(\Phi) = 1$ by the valency theorem, and so Φ is a conformal equivalence. □

15.2 Classification of elliptic functions

Finally, let's see that we can in fact classify all meromorphic functions on a complex torus, and indeed they can all be written in terms of $\wp$. The result is phrased equivalently in terms of elliptic functions with periods Λ.

Theorem 15.2. *Let f be an elliptic function with periods Λ. There exist rational functions Q_1, Q_2 such that*

$$f(z) = Q_1(\wp(z)) + Q_2(\wp(z))\wp'(z)\,.$$

Furthermore, if f is even, then we can take $Q_2 = 0$.

Proof. First, assume f is even. Next, let's observe that we may assume that f only has simple zeroes and poles.

Indeed, let B_1 be the set of branch points of f. This is equal to the set of branch points of the induced function $\bar{f}$ on the complex torus $\mathbb{C}/\Lambda$, which is finite. Likewise, let B_2 be the set of branch points of $\wp$, which is also finite. Therefore, we may choose distinct $c, d \in \mathbb{C} \setminus (B_1 \cup B_2)$. Now the function

$$\frac{f(z) - c}{f(z) - d}$$

is still even, neither its zeroes nor its poles are ramification points of f or $\wp$. So we can and will assume that f has this property.

Since f is even, its zeroes are of the form

$$\pm a_1, \ldots, \pm a_n$$

and likewise its poles are of the form

$$\pm b_1, \ldots, \pm b_n$$

(where $2n = \deg(f)$). Therefore, the function

$$g(z) = \frac{(\wp(z) - \wp(a_1))\cdots(\wp(z) - \wp(a_n))}{(\wp(z) - \wp(b_1))\cdots(\wp(z) - \wp(b_n))}$$

is also an elliptic function with periods Λ, and with the same zeroes and poles as f. Therefore the function $f(z)/g(z)$ is a non-zero elliptic function with all singularities removable, and hence is equal to some constant c by Corollary 13.10. Therefore

$$f(z) = cg(z)\,,$$

which is a rational function as required.

If f is odd, then $f(z)/\wp'(z)$ is an even elliptic function, so equal to a rational function of $\wp(z)$ by the above. Therefore $f(z) = Q_2(\wp(z))\wp'(z)$ as required.

Finally, an arbitrary function f can be written as as sum of its even and odd parts,

$$f(z) = \left(\frac{f(z) + f(-z)}{2}\right) + \left(\frac{f(z) - f(-z)}{2}\right),$$

so the general result follows. □

15.3 Quotients of Riemann surfaces

Hopefully it is now clear that we were able to learn a great deal about complex tori $\mathbb{C}/\Lambda$ by thinking of them as quotients of $\mathbb{C}$. We would like to be able to think about all Riemann surfaces that way, but in order to make a reasonable statement, we need to make some definitions.

Definition 15.3. Let a group G act by homeomorphisms on a space X. The action is said to be *properly discontinuous* if, for every compact $K \subseteq X$, the set

$$\{g \in G \mid g(K) \cap K \neq \varnothing\}$$

is finite. If, for each $x \in X$, the stabiliser $\mathrm{Stab}_G(x)$ is trivial, then the action is said to be *free*.

This definition generalises the action of a lattice Λ on $\mathbb{C}$.

Example 15.4. If Λ is a lattice in $\mathbb{C}$ then the action of Λ on $\mathbb{C}$ by translation is properly discontinuous and free.

These slightly awkward-looking conditions turn out to be exactly what is needed to show that the quotient space is Hausdorff and the quotient map is a covering map.

Lemma 15.5. *Let G be a group acting freely and properly discontinuously by homeomorphisms on a Riemann surface R. The quotient space $G\backslash R$ is Hausdorff and the quotient map $\pi : R \to G\backslash R$ is a regular covering map.*

Proof. To prove that $G\backslash R$ is Hausdorff, let $p, q \in R$ with $\pi(p) \neq \pi(q)$. By taking charts about p and q, and using small discs in $\mathbb{C}$, we may find disjoint open sets $U \ni p$ and $V \ni q$ such that the closures $\overline{U}$ and $\overline{V}$ are compact. Therefore, setting $K = \overline{U} \cap \overline{V}$ in the definition of proper discontinuity, the set of $g \in G$ such that $g(\overline{U}) \cap \overline{V} \neq \varnothing$ is a finite set $\{g_1, \ldots, g_n\}$. Because R is Hausdorff, for each g_i there are disjoint open neighbourhoods $U_i \ni x$ and $V_i \ni g_i.y$. Now

$$U' = U \cap \bigcap_i U_i$$

and

$$V' = V \cap \bigcap_i g_i^{-1}(V_i)$$

are neighbourhoods of p and q respectively, with the property that $G.U'$ is disjoint from $G.V'$. Therefore, $\pi(U')$ and $\pi(V')$ are the required disjoint open neighbourhoods of $\pi(p)$ and $\pi(q)$.

The argument that π is a covering map is similar. For $p \in R$, as above take an open neighbourhood $U \ni p$ such that the closure $\overline{U}$ is compact. Setting $K = \overline{U}$ in the definition of proper discontinuity, the set of $g \in G$ such that $g(\overline{U}) \cap \overline{U} \neq \varnothing$ is a finite set $\{1, g_1, \ldots, g_n\}$, with each $g_i \neq 1$. Since the action is free, $g_i.x \neq x$ for each i. Therefore, for each i, there are disjoint open neighbourhoods $U_i \ni x$ and $V_i \ni g_i x$. Now

$$U' = U \cap \bigcap_i (U_i \cap g_i^{-1} V_i)$$

has the property that $g(U') \cap U' = \varnothing$ for each $g \in G \smallsetminus 1$. This can be restated as

$$\pi^{-1}(\pi(U')) = \coprod_{g \in G} g(U'),$$

which shows that π is a regular covering map. □

Just as in the case of lattices, free and properly discontinuous actions are a convenient way of constructing conformal structures on the quotients.

Proposition 15.6. *Let R be a Riemann surface, and let G be a group acting freely and properly discontinuously by conformal equivalences on R. Then the quotient $S = G \backslash R$ is a Riemann surface, and the quotient map*

$$\pi : R \to S$$

is analytic and a regular covering map.

Proof. Since R is connected, its continuous image S must be too. By Lemma 15.5, S is Hausdorff and π is a regular covering map. Finally, the construction of the conformal structure on S is exactly the same as in Example 5.1, where the result was proved for complex tori. □

As an example of the application of these ideas, we will proof a simple case of a famous theorem of Hurwitz about the automorphisms of Riemann surfaces.

Theorem 15.7 (Hurwitz theorem)**.** *Let R be a compact Riemann surface of genus $g_R \geq 2$, and suppose that a group G acts freely and properly discontinuously on R by conformal equivalences. Then G is finite, and indeed $|G| \leq g_R - 1$.*

Proof. By Proposition 15.6, the quotient $S = G\backslash R$ is a Riemann surface, and the quotient map $\pi : R \to S$ is an analytic regular covering map; in particular, it is unramified. By construction, $\deg(\pi) = |G|$. If g_S is the genus of S then the Riemann–Hurwitz theorem gives

$$2g_R - 2 = |G|(2g_S - 2)$$

so, since the left-hand side is positive, $g_S \geq 2$ and so

$$|G| \leq |G|(g_S - 1) = g_R - 1$$

as required. □

Finally, notice that the result fails if $g_R = 1$.

Example 15.8. The complex torus $\mathbb{C}/\Lambda$ is an infinite group, acting faithfully on itself by conformal equivalences. Any finite subgroup acts freely and properly discontinuously, and there are arbitrarily large finite subgroups.

Lecture 16: Uniformisation and its consequences

16.1 The uniformisation theorem

In this final lecture, we state without proof some theorems that classify all Riemann surfaces, and then deduce some consequences. The first, and most difficult, task is to address the simply connected case.

Theorem 16.1 (Uniformisation theorem)**.** *Every simply connected Riemann surface is conformally equivalent to one of:*

(i) the Riemann sphere, $\mathbb{C}_\infty$;

(ii) the complex plane $\mathbb{C}$; or

(iii) the unit disc $\mathbb{D}$.

Remark 16.2. The three Riemann surfaces listed are not conformally equivalent to each other. Indeed, $\mathbb{C}_\infty$ is compact, so not even homeomorphic to the other two. The complex plane $\mathbb{C}$ and the unit disc $\mathbb{D}$ are homeomorphic, but any analytic map $\mathbb{C} \to \mathbb{D}$ is constant by Liouville's theorem, so they are not conformally equivalent.

The proof is beyond the scope of this course. With the uniformisation theorem in hand, we can make progress on classifying all Riemann surfaces. Let's start with the case of surfaces of genus 0.

Corollary 16.3. *The conformal structure on S^2 is conformally equivalent to $\mathbb{C}_\infty$.*

Proof. Let R be a Riemann surface defined by a conformal structure on S^2. By Theorem 16.1, R is conformally equivalent to one of $\mathbb{C}_\infty$, $\mathbb{C}$ or $\mathbb{D}$; in particular, it is homeomorphic to one of these. But only $\mathbb{C}_\infty$ is compact. □

For Riemann surfaces of positive genus, we need to invoke the tools of algebraic topology.

Theorem 16.4. *Every Riemann surface R has a regular covering map $\pi : \widetilde{R} \to R$ such that $\widetilde{R}$ is simply connected. Furthermore, there is a group G acting freely and properly discontinuously by conformal equivalences on $\widetilde{R}$, and the covering map π descends to a conformal equivalence $G\backslash\widetilde{R} \cong R$.*

Sketch proof. The existence of a simply connected regular covering space $\widetilde{R}$ (the *universal cover* of R) is proved in *Part II Algebraic Topology*; furthermore, the results of that course also show that the *fundamental group* G acts freely and properly discontinuously on $\widetilde{R}$, so that π descends to a homeomorphism $G\backslash\widetilde{R} \cong R$. By Lemma 4.1, there is a unique conformal structure on $\widetilde{R}$ making π into an analytic map. In the standard local coordinates on $\widetilde{R}$ each $g \in G$ just acts as the identity, so G does indeed act by conformal equivalences. The induced homeomorphism $G\backslash\widetilde{R} \to R$ is now an analytic map of degree one, and hence a conformal equivalence. □

Combining Theorem 16.4 with the uniformisation theorem now gives a classification of all Riemann surfaces as quotients.

Corollary 16.5. *Every Riemann surface R is conformally equivalent to a quotient*

$$R \cong G\backslash\widetilde{R}\,,$$

where $\widetilde{R}$ is one of $\mathbb{C}_\infty$, $\mathbb{C}$ or $\mathbb{D}$, and G is a properly discontinuous group of conformal equivalences of $\widetilde{R}$.

We say that R is *uniformised* by $\widetilde{R}$.

Remark 16.6. From the point of view of this course, the group G can be defined as the group

$$G = \{\phi \in \mathrm{Aut}(\widetilde{R}) \mid \pi \circ \phi = \pi\},$$

where $\mathrm{Aut}(\widetilde{R})$ is the group of conformal equivalences $\widetilde{R} \to \widetilde{R}$.

16.2 Classification of Riemann surfaces

Corollary 16.5 suggests a strategy for classifying Riemann surfaces R, by studying properly discontinuous groups of conformal equivalences of $\widetilde{R} = \mathbb{C}_\infty, \mathbb{C}, \mathbb{D}$. It divides naturally into three cases, depending on $\widetilde{R}$. The simplest case is $\widetilde{R} = \mathbb{C}_\infty$.

Proposition 16.7. *If a Riemann surface R is uniformised by $\mathbb{C}_\infty$ then R is conformally equivalent to $\mathbb{C}_\infty$.*

Proof. By hypothesis, $R = G\backslash\mathbb{C}_\infty$ for some G acting properly discontinuously by conformal equivalences. In Question 7 of Example Sheet 1, you showed that the group of conformal equivalences of $\mathbb{C}_\infty$ is the group of Möbius transformations $PSL_2(\mathbb{C})$. But every Möbius transformation $g \in PSL_2(\mathbb{C})$ has at least one fixed-point on $\mathbb{C}_\infty$ (by the fundamental theorem of algebra, for instance). Therefore, no non-trivial subgroup of $PSL_2(\mathbb{C})$ can act freely on $\mathbb{C}_\infty$. So $G = 1$ and $R \cong \mathbb{C}_\infty$. □

The proof of Proposition 16.7 relies on understanding the properly discontinuous subgroups of the group of conformal automorphisms of $\widetilde{R}$. We have a similarly good understanding of this group for $\mathbb{C}$, and we can use this classify the Riemann surfaces uniformised by $\mathbb{C}$.

Proposition 16.8. *If a Riemann surface R is uniformised by $\mathbb{C}$ then one of the following holds:*

(i) $G = 1$ *and* $R \cong \mathbb{C}$;

(ii) $G \cong \mathbb{Z}$ *and* $R \cong \mathbb{C}_*$;

(iii) $G \cong \mathbb{Z}^2$ *and* $R \cong \mathbb{C}/\Lambda$ *for some lattice* Λ.

Proof. In Question 7 of Example Sheet 1 (again), you showed that the group of conformal equivalences of $\mathbb{C}$ is the group

$$\mathbb{C} \rtimes \mathbb{C}_* \cong \{z \mapsto az + b \mid a \in \mathbb{C}_*,\, b \in \mathbb{C}\}\,.$$

Again, since the action is free, non-trivial elements of G cannot have fixed points. But

$$z = az + b$$

has a solution in $\mathbb{C}$ unless $a = 1$, so G consists entirely of translations

$$z \mapsto z + b\,.$$

In Question 1 of Example Sheet 3, you prove that such a group G must be either 1 (the trivial subgroup), or be isomorphic to $\mathbb{Z}$ and generated by translation by some non-zero ω, or be isomorphic to $\mathbb{Z}^2$ and generated by a pair of translations by a pair of non-zero numbers ω_1, ω_2 with $\omega_2/\omega_1 \cong \mathbb{C}^2$,

The first case is the same as item (i) of the proposition, the second case corresponds to item (ii) by Proposition 13.7, and the final case immediately corresponds to item (iii). □

In summary, the Riemann surfaces uniformised by $\mathbb{C}_\infty$ and by $\mathbb{C}$ are some of our favourite simple invariants that we have seen in this course! Everything else is uniformised by $\mathbb{D}$. To see this, we notice that the three different uniformising spaces are mutually exclusive. This in turn relies on the following important fact.

Lemma 16.9. *Let* $f : R \to S$ *be an analytic map of Riemann surfaces. Suppose that* R *is simply connected, and let* $\pi : \widetilde{S} \to S$ *be the uniformising map of* S*. Then there is an analytic map* $F : R \to \widetilde{S}$ *such that* $f = \pi \circ F$.

Proof. See Question 3 on Example Sheet 2, where you proved the result for tori. The general proof is the same. □

Using this, we can easily show that the different uniformising surfaces are mutually exclusive.

Proposition 16.10. *A Riemann surface* R *is uniformised by at most one of* $\mathbb{C}_\infty$*,* $\mathbb{C}$ *and* $\mathbb{D}$*.*

Proof. Propositions 16.7 and 16.8 together show that no surface is uniformised by both $\mathbb{C}_\infty$ and $\mathbb{C}$. Suppose therefore that R is uniformised by both $\mathbb{D}$ and $\widetilde{R} = \mathbb{C}$ or $\mathbb{C}_\infty$, so we have uniformising maps $\pi : \mathbb{D} \to R$ and $f : \widetilde{R} \to R$. By Lemma 16.9, f lifts to an analytic map $F : S \to \mathbb{D}$. But now F is a bounded analytic function on $\mathbb{C}$ or $\mathbb{C}_\infty$, hence constant by Liouville's theorem. Therefore $f = \pi \circ F$ is also constant, which is a contradiction. □

In particular, any Riemann surface not conformally equivalent to one of the surfaces mentioned in Propositions 16.7 or 16.8 must be uniformised by $\mathbb{D}$.

Any attempt to classify the Riemann surfaces uniformised by $\mathbb{D}$ must rely on understanding the group of conformal self-equivalences of $\mathbb{D}$. The following result was proved in *IB Complex Analysis*.

Proposition 16.11. *The group of conformal equivalences of* $\mathbb{D}$ *is the group of Möbius transformations*

$$\{z \mapsto e^{i\theta}\frac{z-a}{1-\bar{a}z} \mid a \in \mathbb{C}, \theta \in \mathbb{R}\},$$

acting on $\mathbb{D}$ *in the natural way.*

Recall also that this group action is more easily understood if we use the Möbius transformation

$$\mu : z \mapsto \frac{1+iz}{1-iz}$$

to conjugate $\mathbb{D}$ to the upper half-plane $\mathbb{H} := \{z \in \mathbb{C} \mid \mathrm{Im}, z > 0\}$. The conformal equivalences of the upper half-plane are just the group $PSL_2(\mathbb{R})$ of Möbius transformations with real coefficients.

Remark 16.12. There is a connection here with *IB Geometry*: both $\mathbb{D}$ and $\mathbb{H}$ provide models of the hyperbolic plane, and in both cases the group of conformal equivalences is equal to the group of orientation-preserving isometries.

Definition 16.13. A subgroup of $PSL_2(\mathbb{R})$ that acts properly discontinuously on $\mathbb{H}$ is called a *Fuchsian group*.

In summary, the problem of classifying the Riemann surfaces uniformised by $\mathbb{D}$ involves first classifying Fuchsian groups, and then understanding all their free and properly discontinuous actions.

16.3 Consequences of uniformisation

Finally, we will give a few consequences of uniformisation.

Corollary 16.14. *If R is a compact Riemann surface and the genus of R is at least 2 then R is uniformised by $\mathbb{D}$.*

Proof. The compact Riemann surfaces uniformised by $\mathbb{C}_\infty$ or $\mathbb{C}$ have genus at most 1. Therefore, by Proposition 16.10, R is uniformised by $\mathbb{D}$. □

Even in the case of domains in $\mathbb{C}$, the uniformisation theorem is non-trivial. This instance of it is called the *Riemann mapping theorem.*

Corollary 16.15 (Riemann mapping theorem)**.** *If $D \subsetneq \mathbb{C}$ is a simply connected, proper subdomain of $\mathbb{C}$, then D is conformally equivalent to $\mathbb{D}$.*

Proof. By Theorem 16.1, it suffices to prove that D is not conformally equivalent to $\mathbb{C}_\infty$ or $\mathbb{C}$. The case of $\mathbb{C}_\infty$ is clear, because D is not compact whereas $\mathbb{C}_\infty$ is.

The case of $\mathbb{C}$ is similar to part of Question 7 from Example Sheet 1. Suppose $f : \mathbb{C} \to D$ be a conformal equivalence. If the singularity at ∞ is essential then, by the Casorati–Weierstrass theorem, $f(\{|z| > 1\})$ is dense. On the other hand, $f(\mathbb{D})$ is open by the open mapping theorem, so $f(\mathbb{D}) \cap f(\{|z| > 1\}) \neq \varnothing$, which contradicts the fact that f injective. Therefore, ∞ is a removable singularity or a pole, so f extends to an analytic map

$$\bar{f} : \mathbb{C}_\infty \to \mathbb{C}_\infty \,.$$

But $\bar{f}$ is non-constant and hence surjective, so $\mathbb{C}_\infty = D \cup \{\bar{f}(\infty)\}$. This contradicts the claim that D is a proper subdomain of $\mathbb{C}$. □

We finish where we started the course – with complex analysis – by giving a proof of Picard's theorem.

Corollary 16.16 (Picard's theorem)**.** *Any analytic function*

$$f : \mathbb{C} \to \mathbb{C} \smallsetminus \{0, 1\}$$

is constant.

Proof. In Question 9 of Example Sheet 3, you will prove that $\mathbb{C} \smallsetminus \{0, 1\}$ is uniformised by $\mathbb{D}$. Therefore, by Lemma 16.9, there is an analytic map $F : \mathbb{C} \to \mathbb{D}$ with $f = \pi \circ F$, where π is the uniformising map $\mathbb{D} \to \mathbb{C} \smallsetminus \{0, 1\}$. But F is constant by Liouville's theorem, so f is too. □