Part IID Riemann surfaces

Dr Alexei Kovalev

Notes Michaelmas 2007

These notes are a bit terse at some points (and are not intended to be a replacement for the notes you take in lectures). Some proofs are not included, but can be chased via the given page references to textbooks. The paragraphs in small print are intended for interest only and are not examinable. I thank Owen Jones for the feedback on an earlier version of these notes. Corrections or suggestions for improvements are welcome at any time (my e-mail address is `a.g.kovalev@dpmms.cam.ac.uk`*).*

I. Holomorphic functions

Recall that a function is called *holomorphic* (I prefer this term to *analytic*—sorry) if it is complex differentiable. Furthermore, given a function, f say, from an open subset U of $\mathbb{C}$ to $\mathbb{C}$, the following are equivalent:

(i) f is *complex differentiable*: for each point z_0 in U

$$f'(z_0) = \lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} \qquad \text{exists;}$$

(ii) f admits *power series expansions*: if a disc $D(a,r) = \{|z-a| < r\}$ is contained in U then

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n, \text{ valid for all } z \in D(a,r)$$

(the coefficients can be calculated by the formula $c_n = f^{(n)}(a)/n!$);

(iii) *Cauchy's Integral Formula* holds for f: if a closed disc $\overline{D(a,r)}$ is contained in U then

$$f(z) = \frac{1}{2\pi i}\int_{\gamma(a,r)} \frac{f(w)}{w-z}dw, \quad \text{if } |z-a|<r,$$

where $\gamma(a,r)$ is a circle with centre a and radius r, bounding the disc.

Thus each of the above three properties can be used as a definition of holomorphic (analytic) function on U.

If f is holomorphic on a punctured disc $D^*(a,r) = \{0 < |z-a| < r\}$ about a, we say that f has an *isolated singularity* at a. Then f expands, in a unique way, in *Laurent series at* a, $f(z) = \sum_{n=-\infty}^{\infty} c_n(z-a)^n$, $0 < |z-a| < r$, where $c_n = \dfrac{1}{2\pi i}\displaystyle\int_{\gamma(a,r)} \frac{f(z)}{(z-a)^{n+1}}dz$. The

$\sum_{n=-\infty}^{-1} c_n(z-a)^n$ is called the *principal part* of Laurent series. Further, f has at a a *removable singularity* if $c_n = 0$ for all $n < 0$, a *pole of order* $m > 0$ if $c_{-m} \neq 0$, $c_n = 0$ for all $n < -m$, and an *essential singularity* otherwise (i.e. if for any $n \in \mathbb{Z}$ there is $n' < n$ with $c_{n'} \neq 0$).

Throughout this chapter, U denotes a connected open subset of the complex plane. For an arbitrary subset S of $\mathbb{C}$, the statement 'a function f is holomorphic on S' means 'a function f is holomorphic on some *open set* containing S'.

1 Zeros and poles.

This section gives a very rapid summary of some results from IB Complex Analysis which will be needed later.

The Identity Theorem essentially says that zeros of a holomorphic function $f : U \to \mathbb{C}$ are isolated in U unless $f \equiv 0$: more explicitly, if $f(a) = 0$, but $f \not\equiv 0$, then there is $\varepsilon > 0$ such that a is the *only* zero of f in the disc $D(a, \varepsilon)$.

Definition. A holomorphic function f has a **zero of order** $m > 0$ at a if

$$f(z) = c_m(z-a)^m + c_{m+1}(z-a)^{m+1} + \ldots, \qquad c_m \neq 0,$$

for $|z-a| < r$ and some $r > 0$.

An isolated zero of holomorphic function always has finite order.

If $f : U \to \mathbb{C}$ is holomorphic and K is a compact subset of U, then f can have only *finitely many* zeros in K.

Let $z_1, \ldots, z_N$ be the list of *all* the zeros of f in K, $m_1, \ldots, m_N$, respectively, their orders. Then

$$f(z) = (z-z_1)^{m_1} \cdot \ldots \cdot (z-z_N)^{m_N} g(z), \qquad z \in K,$$

where $g(z) \neq 0$ and g is holomorphic on K.

Proposition 1.1. *If f is a holomorphic function on a punctured disc $D^*(a,r)$ satisfying $|f(z)| \leq M|z-a|^{-m}$ for some $M > 0$ and a positive integer m, then f has at $z = a$ a pole of order at most m.*

By contrast, if f has essential singularity at a then $|f(z)|$ cannot be estimated near a by any power of $|z-a|$.

Theorem 1.2 (Weierstrass–Casorati)**.** *Let $f : D^*(a,r) \to \mathbb{C}$ be a holomorphic function having essential singularity at a. Then for any given $w \in \mathbb{C}$ there exists a sequence z_n (depending on w) converging to a such that $f(z_n) \to w$, as $n \to \infty$.*

As we shall see, poles are nicer than essential singularities.

Definition. A function $f : U \to \mathbb{C}$ is said to be **meromorphic** on U if it is holomorphic on U apart from the points where it has poles,
i.e. $f : U \setminus \{a_1, a_2 \ldots\} \to \mathbb{C}$ is holomorphic and each of a_j is an isolated non-essential singularity of f. Yet another way to say the same thing is that 'at any point of U the principal part of Laurent series for f has only finitely many terms'.
We shall adopt the formal notation $f(a) = \infty$ when f has a pole of a.

Poles are isolated by definition, therefore if $f : U \to \mathbb{C}$ is meromorphic and K is a compact subset of U, then f can have only *finitely many* poles in K.

Proposition 1.3 (Partial fractions). *Assume that f is meromorphic on U and $K \subset U$ is compact, and let $z_1, \dots, z_N$ be all the poles of f in K, with $m_1, \dots, m_N$, respectively, their orders. Then*

$$f(z) = P_1(\frac{1}{z - z_1}) + \ldots + P_N(\frac{1}{z - z_N}) + g(z), \qquad \text{for } z \in K,$$

where each P_j is a polynomial of degree m_j, with no constant term, and $g(z)$ is holomorphic on K.

Proposition 1.4 (Factorization principle for zeros and poles). *Let $f : U \to \mathbb{C}$ be meromorphic and K a compact subset of U, and $a_1, \dots, a_N$ all the poles of f in K of orders $k_1, \dots, k_N$ respectively, $b_1, \dots, b_M$ all the zeros of f in K, $\ell_1, \dots, \ell_M$ their orders. Then*

$$f(z) = \frac{(z - b_1)^{\ell_1} \cdot \ldots \cdot (z - b_M)^{\ell_M}}{(z - a_1)^{k_1} \cdot \ldots \cdot (z - a_N)^{k_N}} \tilde{g}(z) \qquad \text{for } z \in K,$$

for a uniquely determined holomorphic function $\tilde{g}$ on K.

In the next result, γ is assumed to be a simple closed curve[1], such that the winding number $n(\gamma, a)$ is 0 or 1 for any a not on γ. We say that a is *inside* γ when $n(\gamma, a) = 1$.

Theorem 1.5 (Argument principle). *Let U and γ be such that γ and any points inside γ are contained in U. Suppose that $f : U \to \mathbb{C}$ is meromorphic and has no zeros or poles on γ. Let $z_1, \dots, z_m$ be all the zeros of f inside γ, with $k_1, \dots, k_m$ their orders, and $w_1, \dots, w_n$ all the poles of f inside γ, with $l_1, \dots, l_n$ their orders. Then*

$$k_1 + \ldots + k_m - l_1 - \ldots - l_n = n(f \circ \gamma, 0). \tag{1.6}$$

2 The 'local structure' of holomorphic maps

Definition. We say that a holomorphic function f takes the value w_0 with **multiplicity** m at a point z_0, or that the **valency** of f at z_0 is m, if $f(z) - w_0$ has at z_0 a zero of multiplicity (order) m, i.e.

$$f(z) = w_0 + c(z - z_0)^m + (\text{higher power terms in } (z - z_0) \,), \quad c \neq 0.$$

We then write $v_f(z_0) = m$.

Theorem 2.7 (Local Mapping Theorem). *Let f be a holomorphic non-constant function on U and $v_f(z_0) = m$. Then there exist $\varepsilon > 0$ and $\delta > 0$, such that if $0 < |w - f(z_0)| < \delta$ then the equation $f(z) = w$ has precisely m distinct solutions z in the punctured disc $0 < |z - z_0| < \varepsilon$. The valency of f is 1 at each of these solutions.*

Notice that the Local Mapping Theorem implies its *converse*: given that for any $w \in D^*(f(z_0), \delta)$, the equation $f(z) = w$ has precisely m solutions $z \in D^*(z_0, \varepsilon)$, one has that the valency $v_f(z_0)$ is m — because this is the only value it could be!

[1] i.e. no self-intersections other than the end-points

Definition. $f : D \to D' = f(D)$ (here D, D' are open in $\mathbb{C}$) is called a **biholomorphic map**, or **conformal equivalence**, if f is holomorphic and has a holomorphic inverse $f^{-1} : D' \to D$.

Corollary 2.8 (Inverse Function Theorem)**.** *If $f : D \to \mathbb{C}$ is holomorphic and $f'(z_0) \neq 0$, for some $z_0 \in D$, then there exists $\varepsilon > 0$ such that $f : D(z_0, \varepsilon) \to D' = f(D(z_0, \varepsilon))$ is a conformal equivalence.*

A conformal equivalence $z \in U \to z' = u(z) \in U'$ can be thought of as a 'change of holomorphic coordinates'. There is no need to distinguish z as a 'canonical' coordinate on a given connected open set in $\mathbb{C}$—any biholomorphic function of z can be used just as well to 'parameterize' the same set. So, if a non-constant function $f : U \to V$ is holomorphic and $u : U \to U'$, $v : V \to V'$ are biholomorphic maps, then another holomorphic function $\tilde{f}$ is defined by the relation $v \circ f = \tilde{f} \circ u$, conveniently expressed by a commutative diagram

$$\begin{array}{ccc} U & \xrightarrow{f} & V \\ {\scriptstyle u}\downarrow & & \downarrow{\scriptstyle v} \\ U' & \xrightarrow{\tilde{f}} & V'. \end{array}$$

In this case, f and $\tilde{f} = v \circ f \circ u^{-1}$ can be regarded as *different ways to express the same holomorphic map.*

With this shift of view, the natural questions are: what properties of a holomorphic map f remain *invariant*, i.e. independent of the choice of complex coordinate? And a kind of converse: can one choose a coordinate to put a given f in a (simple) 'standard form'?

Valency is an invariant, as can be deduced from the Local Mapping Theorem. Furthermore, valency at z_0 determines the local structure of a holomorphic map near z_0.

Theorem 2.9 (The local structure of a holomorphic map)**.** *Let f be holomorphic on a disc D centred at z_0, $f(z_0) = w_0$ and $v_f(z_0) = n$. Then there are conformal equivalences u, v defined respectively near z_0, w_0, with $u(z_0) = 0$, $v(w_0) = 0$ and such that*

$$\tilde{f}(z') = v \circ f \circ u^{-1}(z') = z'^{\,n}, \quad \textit{for any } z' \textit{ near 0.}$$

Fig. 1. The local structure of f near z_0 when $v_f(z_0) = 3$.

Some page references for Chapter I

All to L.V. Ahlfors, "Complex analysis", 3rd edition.

Argument Principle: p. 131 Theorem 10 and p. 152 Theorem 18. Local Mapping Theorem: Theorem 11 p. 131. A formula for the inverse function: pp. 153-154. The 'local structure' of holomorphic maps: pp. 130–135.

IIa. Holomorphic maps on the Riemann sphere

Recall that $S^2 \subset \mathbb{R}^3$ is given by the equation $X^2 + Y^2 + Z^2 = 1$. Call $N = (0, 0, 1)$ the North Pole and $S = (0, 0, -1)$ the South Pole and put $S_0 = S^2 \setminus \{N\}$, $S_\infty = S^2 \setminus \{S\}$. The open subsets of S^2 are defined as intersections of S^2 with the open subsets of $\mathbb{R}^3$. In particular, S_0 and S_∞ are open in S^2. Consider the following bijective maps

$$\varphi : (X, Y, Z) \in S_0 \to z = \frac{X + iY}{1 - Z} \in \mathbb{C},$$

the stereographic projection from the North Pole, and

$$\psi : (X, Y, Z) \in S_\infty \to z = \frac{X - iY}{1 + Z} \in \mathbb{C},$$

the stereographic projection from the South Pole combined with complex conjugation. The maps φ and ψ are homeomorphisms (i.e. continuous with continuous inverse), indeed it can be calculated that e.g. $\varphi^{-1}(z) = \left(\frac{z + \bar{z}}{|z|^2 + 1}, \frac{z - \bar{z}}{i(|z|^2 + 1)}, \frac{|z|^2 - 1}{|z|^2 + 1} \right)$.

Note that with the above definitions $\varphi \circ \psi^{-1}(z) = 1/z$ is a *holomorphic* function on $\mathbb{C} \setminus \{0\}$.

Definition. Let U be an open subset of S^2 and $F : U \to S^2$ a continuous map. We say that F is a **holomorphic map** when all the functions $\varphi \circ F \circ \varphi^{-1}(z)$, $\varphi \circ F \circ \psi^{-1}(z)$, $\psi \circ F \circ \varphi^{-1}(z)$, $\psi \circ F \circ \psi^{-1}(z)$ are holomorphic where defined, that is for z in $\varphi(U \cap S_0 \cap F^{-1}(S_0))$, $\psi(U \cap S_\infty \cap F^{-1}(S_0))$, $\varphi(U \cap S_0 \cap F^{-1}(S_\infty))$, $\psi(U \cap S_\infty \cap F^{-1}(S_\infty))$, respectively.

If $f(z) = \varphi \circ F \circ \varphi^{-1}(z)$ then one has $\varphi \circ F \circ \psi^{-1}(z) = (\varphi \circ F \circ \varphi^{-1}) \circ (\varphi \circ \psi^{-1})(z) = f(1/z)$ when both $f(z)$ and $f(1/z)$ are defined.

Exercise 2.10. Verify by similar arguments that $\psi \circ F \circ \varphi^{-1}(z) = 1/f(z)$ and $\psi \circ F \circ \psi^{-1}(z) = 1/f(1/z)$. Thus it does make sense to require all four functions in the definition to be holomorphic.

Proposition 2.11. *If $F : S^2 \to S^2 \setminus \{N\}$ (so F misses the North Pole) and is holomorphic then F is constant.*

Proposition 2.12. *Suppose that $U \subset S^2$ is connected, open, and U does not contain the North Pole N, and $F : U \to S^2$ is a continuous map, but F is not identically equal to N. Then F is a holomorphic map if and only if $f = \varphi \circ F \circ \varphi^{-1}$ is a meromorphic function on $\varphi(U)$ with poles at $\varphi(F^{-1}(N))$.*

So, the composition with φ induces a one-to-one correspondence between holomorphic maps $U \to S^2$ (excluding the constant map to the North Pole) and meromorphic functions on U, in such a way that $N \in S^2$ and ∞ correspond. (Recall the convention that that if z_0 is a pole of f we write $f(z_0) = \infty$.)

Proof. Let z be a point in $\varphi(U)$. Put $g(z) = \psi \circ F \circ \varphi^{-1}(z)$. Then the following are equivalent: (i) $f(z)$ is **not** defined (as a complex number); (ii) $F(\varphi^{-1}(z)) = N$; (iii) $g(z) = 0$. As F is not identically N by the hypothesis, one has $g \not\equiv 0$. Recall also that $f(z) = 1/g(z)$ (from Exercise 2.10).

Now since $N \notin U$ the map F is determined by $F \circ \varphi^{-1}$. Then it follows that F is holomorphic if and only if $f(z)$ and $g(z)$ are holomorphic where defined. The point is that the two other possible compositions, in the definition, are defined where at least one of f, g is defined (and again, see Exercise 2.10). We have just shown that the singularities of f in $\varphi(U)$ are isolated and correspond one-to-one to the zeros of g. So, the singularities of f must be poles and f is meromorphic on $\varphi(U)$. □

Proposition 2.13. *Let V be a connected open subset of S^2 containing the North Pole but not the South Pole, i.e. $N \in V \subset S^2 \setminus \{S\}$. Suppose that $F : V \to S^2$ is a continuous non-constant map. Then F is holomorphic on V if and only if the following two conditions hold:*
(i) F is a holomorphic map on $V \setminus \{N\}$, i.e. $f = \varphi \circ F \circ \varphi^{-1}$ is a meromorphic *function on $V \setminus \{N\}$; and*
(ii) $\quad f(z) = \sum_{k=-\infty}^{n} c_k z^k \quad (n \in \mathbb{Z},\ c_n \neq 0)$, *for $|z| > R$, with some large R.*

Fig. 2. A neighbourhood of the North Pole and a 'neighbourhood of ∞'.

Proof. We shall prove that $F : V \to S^2$ holomorphic implies (i) and (ii). The converse is obtained by going over the same arguments backwards and is left as an exercise.

The (i) follows from Proposition 2.12 by restricting to $V \setminus \{N\}$. To deduce (ii), notice first that $F : V \to S^2$ is determined by the map $F \circ \psi^{-1} : \psi(V) \to S^2$. Then $h(z) = \varphi \circ F \circ \psi^{-1}(z)$ is meromorphic for $z \in \psi(V)$. The proof of this latter claim follows Proposition 2, but with necessary changes in notation (compose with ψ^{-1} instead of φ^{-1}, etc.). As $0 \in \psi(V)$, $h(z)$ at $z = 0$ has Laurent expansion of the form $h(z) = \tilde{c}_m z^m + \tilde{c}_{m+1} z^{m+1} + \ldots$, valid in some punctured disc $0 < |z| < r$. Now using the relation $h(z) = f(1/z)$ for $z \in \psi(V)$, $z \neq 0$ we find $f(z) = c_n z^n + c_{n-1} z^{n-1} + \ldots$ for $|z| > R$ with $R = 1/r$, $n = -m$, $c_n = \tilde{c}_m$, $c_{n-1} = \tilde{c}_{m+1}, \ldots$, as we had to prove. □

Example. Recall that a rational function is defined as

$$f(z) = \frac{p(z)}{q(z)},$$

where p and q are polynomials with no common factor. Then f is meromorphic on $\mathbb{C}$. Furthermore, it is not difficult to verify that $h(z) = f(1/z)$ defines again a rational function, so $h(z)$ has Laurent series at $z = 0$ and $f(z)$ satisfies the condition (ii) above. Now Propositions 2.12 and 2.13 together imply that every rational function f induces a holomorphic map $F : S^2 \to S^2$, such that $f = \varphi \circ F \circ \varphi^{-1}$.

It will be convenient to adopt the following.

Definition. Let $f : \{|z| > R\} \to \mathbb{C}$ be holomorphic. Then $f(\infty) = a_0$ $(a_0 \in \mathbb{C})$ *with multiplicity* m if $h(z) = f(1/z)$ extends holomorphically over zero by putting $h(0) = a$ and h takes its value at $z = 0$ with multiplicity m (i.e. the valency $v_h(0) = m$). Similarly, define $f(\infty) = \infty$ *with multiplicity* m if $h(z)$ has at $z = 0$ a pole of order m.

The definition and the preceding results justify the thinking of S^2 as $\mathbb{C} \cup \{\infty\}$; under stereographic projection φ, the ∞ corresponds to the North Pole (the point from which the projection is performed). Then the condition (ii) of Proposition 2.13 can be read as f being 'meromorphic near ∞'.

Now if f is rational we can write

$$f(z) = \begin{cases} p(z)/q(z), & \text{if } q(z) \neq 0; \\ \infty, & \text{if } q(z) = 0; \\ \lim_{|z|\to\infty} p(z)/q(z), & \text{if } z = \infty. \end{cases}$$

Exercise 2.14. Derive a formula for $\lim_{|z|\to\infty} p(z)/q(z)$ in terms of the degrees and coefficients of p and q. (Hint: consider the rational function $h(z) = f(1/z)$ and let $z \to 0$.)

Theorem 2.15. *Any holomorphic map $F : S^2 \to S^2$ is defined, in the way outlined above, by a rational function. Furthermore, if F is non-constant then F is onto (surjective).*

Proof. We may assume without loss of generality that F is non-constant. Then $f = \varphi \circ F \circ \varphi^{-1}$ is meromorphic on $\mathbb{C}$ by Proposition 2.12. On the other hand, by Proposition 2.13, since F is holomorphic near $N \in S^2$, there is $R > 0$ such that f has Laurent series

$$f(z) = c_n z^n + c_{n-1} z^{n-1} + \dots \qquad \text{valid for } |z| > R. \qquad (*)$$

In particular, f is holomorphic and has no poles in $\{|z| > R\}$. Now $\{|z| \leq R\}$ is compact, so f can have at most a finite number of poles there, $z_1, \dots, z_N$ say. Define $P_0(z) = c_n z^n + \dots + c_1 z$, if $n > 0$, and $P_0(z) \equiv 0$ otherwise. Pull off $P_0(z)$ and the principal parts of f at z_j,

$$f(z) = P_0(z) + P_1\left(\frac{1}{z - z_1}\right) + \dots + P_N\left(\frac{1}{z - z_N}\right) + g(z).$$

Here P_0, $P_1, \dots, P_N(z)$ are polynomials with no constant terms and $g(z)$ is some function holomorphic on $\{|z| \leq R\}$. In fact, $g = f - P_0 - P_1 - \dots - P_N$ is also holomorphic on $\{|z| > R\}$, since all of the $f, P_0, P_1, \dots P_N$ are so, and thus g is holomorphic on $\mathbb{C}$.

Further, as $P_1, \dots P_N$ only contribute negative powers to the expansion of f in $(*)$, $\hat{g}(z) = g(1/z)$ has removable singularity $z = 0$, so $\hat{g}$ is holomorphic and bounded for $0 < |z| < 1/R$. But then it follows that $g(z)$ is bounded on $|z| > R$ and thus bounded on $\mathbb{C}$, so g has to be constant by Liouville's Theorem. The formula $(*)$ then implies that $f(z)$ is rational.

Now, for the second part of the theorem. We suppose

$$f(z) = \frac{p(z)}{q(z)},$$

and that the degree of at least one of p and q is positive. Consider the case that the degree d, say, of p is strictly greater than the degree ℓ of q. Now if $a \in \mathbb{C}$ we have $f(z) = a$ if and only if

$$p(z) - aq(z) = 0.$$

But this is a polynomial equation of degree d and hence has d solutions (counted with multiplicity) for $z \in \mathbb{C}$. It remains to check that f does not miss ∞ (so F does not miss $N \in S^2$). In fact, since $d > \ell$, f has a pole of order $d - \ell$ at infinity, and so $f(\infty) = \infty$ with multiplicity $d - \ell$. In addition q has ℓ zeros in $\mathbb{C}$ (counted with multiplicity), so there are d solutions of $F(P) = \infty$, $P \in S^2$, counted with multiplicity.

Exercise 2.16. Show that f is surjective in the remaining cases: when the degrees of p and q are equal, $d = \ell$, and when $d = \deg p < \ell = \deg q$. (Hint: check that now one has $f(\infty) \in \mathbb{C}$ and calculate the multiplicity of f at ∞.)

This exercise completes the proof. □

In the course of the last part of the proof we noticed that for a non-constant holomorphic map $F : S^2 \to S^2$, the number of solutions P of the equation $F(P) = Q$ is independent of the point Q in S^2 (at least if one takes account of the multiplicities). This number, which is the larger of the degrees of the polynomials p and q is called the **degree** of the rational function $f = p/q$ and also the degree of the (induced) holomorphic map F.

A notable consequence of the theorem is that $F : S^2 \to S^2$ is **biholomorphic** (has a holomorphic inverse) if and only if the degree of F is equal to 1. Biholomorphic maps $S^2 \to S^2$ are simply the **Möbius transformations**

$$T(z) = \frac{az + b}{cz + d} \quad \text{with } ad - bc = 1.$$

More precisely, T defines a map $S^2 \to S^2$ by setting $T(\infty) = a/c$ and $T(-d/c) = \infty$ when $c \neq 0$, and $T(\infty) = \infty$ when $c = 0$.

We note also the following

Theorem 2.17. *Let $P_1, \ldots, P_k$ and $Q_1, \ldots, Q_k$ be two lists of points of $\mathbb{C} \cup \{\infty\}$ (possibly with repetitions), such that $P_i \neq Q_j$ for all i, j. Then there is a rational function f such that $f(P) = 0$ if and only if P is one of the P_j's, and $f(P) = \infty$ if and only if P if one of the Q_j's. If P_j occurs n times in the list then the statement is to be interpreted as saying that f has a zero of order n at P_j, and similarly for Q_j.*

Furthermore, if g is any other rational function with the same property as f then there is a constant $c \neq 0$ such that $g(z) = cf(z)$.

In other words, there always exist rational functions with prescribed zeros and poles on $\mathbb{C} \cup \infty$, provided only that the number of zeros is equal to the number of poles.

Finally, two philosophical remarks. First, the definitions given here for the notions of holomorphic maps to and from S^2 may seem somewhat arbitrary. They fit into a general framework of holomorphic mappings between Riemann surfaces later on, however.

Secondly, note that we have combined the condition of **holomorphy** with **compactness** in the domain and range (both equal to S^2) and found that the resulting objects are **algebraic** (given by polynomials). This is the first instance of a remarkable phenomenon that turns out to be very general and very important.

IIb. Elliptic functions

3 Fourier series for periodic holomorphic functions

Let f be a periodic function with period λ, $f(z+\lambda)=f(z)$. We may assume without loss of generality that $\lambda=1$ (otherwise, consider $\tilde{f}(z)=f(\lambda z)$). Suppose further that $f(z)$ is defined on an infinite horizontal strip $S=\{z:\alpha<\operatorname{Im} z<\beta\}$. Put $\mathbf{e}(z)=\exp(2\pi iz)$, then $\mathbf{e}(z)$ has period 1 and maps S to an annulus $\mathbf{e}(S)=A=\{e^{-\beta}<|w|<e^{-\alpha}\}$.

For any $w\in A$, define $F(w)=F(e^{2\pi iz})=F(\mathbf{e}(z))=f(z)$. The periodicity of f ensures that $F(w)$ is well-defined and the Inverse Function Theorem (see Chapter I) further ensures that F is holomorphic on A. Respectively, the Laurent series for $F(w)=\sum_{n=-\infty}^{\infty}c_nw^n$ become *Fourier series* for f,

$$f(z)=\sum_{n=-\infty}^{\infty}c_ne^{2\pi inz},\qquad c_n=\int_a^{a+1}f(z)e^{-2\pi inz}dz,$$

where the formula for c_n is obtained by the change of variable $w=\mathbf{e}(z)$ in the formula for the coefficients of Laurent series.

4 Constraints on elliptic functions

Let τ be a complex number with $\operatorname{Im}\tau>0$. A **lattice Λ generated by 1 and τ** is a subset of $\mathbb{C}$ given by $\Lambda=\{n+m\tau:n,m\in\mathbb{Z}\}$ (it is actually a subgroup of the additive group $\mathbb{C}$). A choice of τ, and hence Λ, will be fixed throughout this chapter.

Definition. A function $f(z)$ is said to be **doubly-periodic** with period lattice Λ (or with periods 1 and τ) if for any z in the domain of definition of f, $f(z+1)$ and $f(z+\tau)$ are also defined and

$$f(z+1)=f(z),\quad f(z+\tau)=f(z).$$

Doubly-periodic meromorphic functions on $\mathbb{C}$ are called **elliptic functions**.

Denote by P a closed parallelogram with vertices $\xi,\xi+1,\xi+1+\tau,\xi+\tau$, where ξ is some complex number. Then a Λ-periodic function is determined by its values on P (and P is the closure of a 'fundamental region' of Λ). Think of the boundary ∂P of P as a piece-wise smooth simple closed curve oriented so that its winding number around any point not on ∂P is either 0 or 1.

Theorem 4.18 (Constraints on the elliptic functions)**.** *Let F be an elliptic function and assume that F has no zeros or poles on the boundary of P.*[2] *Suppose that $a_1,\dots,a_k$ are all the zeros of F in P repeated according to multiplicities and $b_1,\dots,b_\ell$ are all the poles of F in P repeated according to multiplicities. Then*

(i) if F is holomorphic on $\mathbb{C}$ then F is constant;

(ii) $\sum_{z\in P}\operatorname{res}_zF=0$*;*

[2]Remark: As the zeros and poles of a meromorphic function are isolated and P is compact, it is always possible to choose ξ so that there are no zeros or poles of F on the boundary of P.

(iii) the numbers of zeros and poles of F in P are equal, this number is called the **degree of elliptic function** *F, $\deg F = k = \ell$;*

(iv) $\sum_{j=1}^{k} a_j - \sum_{j=1}^{\ell} b_j \in \Lambda$.

Remarks. A function $F(z)-c$, $c \in \mathbb{C}$, is elliptic and has the same poles as $F(z)$. It follows that any complex value c of F is taken precisely $k = \deg F$ times on the parallelogram P, counting with multiplicities and assuming an appropriate choice of P, so that $F(z) \neq c$ on ∂P.

Also, it follows from (ii) (and from (iv)) that the degree of F cannot be equal to 1, thus a non-constant elliptic function must have degree at least 2.

Proof. (i) If F is holomorphic on $\mathbb{C}$ then F is continuous, hence bounded on the compact P. But then, by double-periodicity, F is a bounded holomorphic function on $\mathbb{C}$ and must be constant by Liouville's Theorem.

(ii) By the residue theorem, the sum of the residues in P is computed by

$$\frac{1}{2\pi i} \int_{\partial P} F(z) dz.$$

On the other hand, double-periodicity of F implies that this integral vanishes.

(iii) By the argument principle,

$$k - \ell = \frac{1}{2\pi i} \int_{\partial P} \frac{F'(z)}{F(z)} dz.$$

Now if F is doubly-periodic then so is F'/F. The double-periodicity of F'/F implies that the integral in the RHS vanishes.

(iv) Applying the residue theorem again and a generalized argument principle (Example sheet 1, Q2) we find

$$\frac{1}{2\pi i} \int_{\partial P} z \frac{F'(z)}{F(z)} dz = \sum_{j=1}^{k} a_j - \sum_{j=1}^{\ell} b_j.$$

On the other hand, in the same integral, calculating the contribution from the sides $[\xi, \xi+1]$ and $[\xi+1+\tau, \xi+\tau]$, one obtains

$$-\frac{\tau}{2\pi i} \int_{\xi}^{\xi+1} \frac{F'(z)}{F(z)} dz = -\frac{\tau}{2\pi i} \int_{F\circ\gamma} \frac{dw}{w} = -\tau n(F\circ\gamma, 0),$$

$$\text{where } \gamma(t) = \xi + t,\ 0 \le t \le 1, w = F(z).$$

Now $F \circ \gamma$ is a closed curve as F has period 1 and the winding number is always an integer. A similar argument shows that the contribution from the other two sides of ∂P is an integer. Thus $\sum_{j=1}^{k} a_j - \sum_{j=1}^{\ell} b_j = n + m\tau$, for some $n, m \in \mathbb{Z}$ as required. □

5 Theta-functions

Theta-functions, although not themselves elliptic, are useful for constructing the elliptic functions. As before, we denote $\mathbf{e}(z) = \exp(2\pi i z)$.

Definition. $\vartheta(z, \tau) = \sum_{n=-\infty}^{\infty} \mathbf{e}(\frac{1}{2} n^2 \tau + nz)$.

Remark. Having fixed τ, I shall often write $\vartheta(z)$ for this function. The dependence of ϑ on τ is an issue in its own right, however it is not discussed here.

The first thing is to check that $\vartheta(z)$ is well-defined and is a holomorphic function on $\mathbb{C}$ (i.e. an entire function). This is done by application of the Weierstrass M-test to show the local uniform convergence on $\mathbb{C}$ for the series for $\theta(z)$. Consider the horizontal strip $S_R = \{|\operatorname{Im} z| \leq R\}$ in $\mathbb{C}$. Write $\tau = \tau_1 + i\tau_2$ (so $\tau_2 = \operatorname{Im}\tau > 0$), $z = x + iy$. Then on S_R

$$\begin{aligned}|\mathbf{e}(\frac{1}{2}n^2\tau + nz)| &= \exp(-\pi n^2\tau_2 - 2\pi n y) \leq \exp(-\pi n^2\tau_2 + 2\pi|n|R) \\ &= \exp(-\pi\tau_2(|n| - R/\tau_2)^2 + \pi R^2/\tau_2) \leq M_R e^{-\pi\tau_2(|n|-\frac{R}{\tau_2})^2},\end{aligned}$$

where a factor M_R is independent of n. Now it is not difficult to check that the series $\sum_{n=-\infty}^{\infty} e^{-\pi\tau_2(|n|-\frac{R}{\tau_2})^2}$ converges, e.g. by comparison with the convergent geometric series $\sum_{n=0}^{\infty}(e^{-\pi\tau_2})^n$ (remember that $\tau_2 > 0$, so $0 < e^{-\pi\tau_2} < 1$). As any compact set in $\mathbb{C}$ is contained in the horizontal strip $\{|\operatorname{Im} z| \leq R\}$ for some $R > 0$, Weierstrass M-test applies to show that $\vartheta(z)$ is holomorphic on any compact subset of $\mathbb{C}$, and thus holomorphic on $\mathbb{C}$.

Exercise 5.19. Show that $\vartheta(-z) = \vartheta(z)$.

Further, $\vartheta(z)$ is clearly periodic with period 1, $\vartheta(z+1) = \vartheta(z)$, as it is given by Fourier series in z. It is not periodic in τ direction, however

$$\begin{aligned}\vartheta(z+\tau) &= \sum_{n=-\infty}^{\infty} \mathbf{e}(\frac{1}{2}n^2\tau + nz + n\tau) = \sum_{n=-\infty}^{\infty} \mathbf{e}(\frac{1}{2}(n+1)^2\tau - \tau/2 + nz) \\ &= \sum_{n=-\infty}^{\infty} \mathbf{e}(\frac{1}{2}(n+1)^2\tau + (n+1)z - \tau/2 - z) = \mathbf{e}(-\tau/2 - z)\vartheta(z).\end{aligned} \tag{5.20}$$

Proposition 5.21. *With ξ chosen so that $\vartheta(z) \neq 0$ on ∂P, $\vartheta(z)$ has precisely one zero in P and this is a simple zero.*

Proof. As $\vartheta(z)$ is holomorphic on $\mathbb{C}$, the integral $(2\pi i)^{-1}\int_{\partial P}(\vartheta'(z)/\vartheta(z))dz$ gives the number of zeros in P, by the argument principle. As $\vartheta(z)$ has period 1, the contributions from the sloping sides $[\xi+\tau, \xi]$ and $[\xi+1, \xi+1+\tau]$ cancel each other. As calculation gives

$$\frac{\vartheta'(z+\tau)}{\vartheta(z+\tau)} = -2\pi i + \frac{\vartheta'(z)}{\vartheta(z)}, \tag{5.22}$$

it follows that the contribution along the pair of horizontal sides is 1. □

Proposition 5.23. *The unique zero of $\vartheta(z)$ in P is $(1+\tau)/2$ (here P is chosen with $\xi = 0$).*

Proof. Put $f(z) = \vartheta(z - 1/2 - \tau/2)$. Then, reading the formula (5.20) backwards for $z - 1/2 - \tau/2$,

$$\begin{aligned}f(z)\mathbf{e}(-z+1/2) &= \vartheta(z - 1/2 + \tau/2) \\ &= \vartheta(z + 1/2 + \tau/2) \qquad \text{as } \vartheta \text{ has period 1} \\ &= \vartheta(-z - 1/2 - \tau/2) \qquad \text{as } \vartheta \text{ is even} \\ &= f(-z).\end{aligned}$$

With $z = 0$, so $\mathbf{e}(-z+1/2) = -1$, the above argument gives $f(0) = -f(0)$ and therefore $f(0) = 0 = \vartheta(1/2 + \tau/2)$. □

Theorem 5.24 (Existence of elliptic functions)**.** *Let a_j and b_j, $j = 1, \dots, k$, be two lists of points (possibly with repetitions) in P, but not on the boundary ∂P, such that $\sum_{j=1}^k a_j - \sum_{j=1}^k b_j \in \Lambda$ and $a_i \neq b_j$ for all i, j. Then there exists an elliptic function with zeros in P precisely at a_j and poles in P precisely at b_j, as in Theorem 4.18. Furthermore, such a function is unique up to multiplication by a constant.*

Proof. To prove the uniqueness claim, suppose that the elliptic functions F_1 and F_2 both have the required properties. Then F_1/F_2 is doubly periodic and extends holomorphically to all of $\mathbb{C}$ — hence F_1/F_2 is constant.

For the existence, consider

$$F(z) = \frac{f(z - a_1) \dots f(z - a_k)}{f(z - b_1) \dots f(z - b_k)},$$

where again $f(z) = \vartheta(z - 1/2 - \tau/2)$. It is clear that $F(z + 1) = F(z)$. Also, $F(z + \tau) = \mathbf{e}(a_1 + \ldots + a_k - b_1 - \ldots - b_k)F(z) = \mathbf{e}(m\tau)F(z)$ for some $m \in \mathbb{Z}$, by Theorem 4.18(iv). Hence $\mathbf{e}(-mz)F(z)$ gives the required elliptic function. □

6 Weierstrass $\wp$-function

Definition. Weierstrass $\wp$-function is a meromorphic doubly-periodic function such that
(i) $\wp$ has degree 2 and a double pole at $z = 0$; and
(ii) the Laurent expansion of $\wp$ at $z = 0$ has coefficient 1 at the leading term z^{-2} and no constant term.

It follows that $\wp(z)$, if it exists, is uniquely determined and must be an even function. To see the uniqueness of $\wp$ assume that there are two functions, $\wp_1$ and $\wp_2$, satisfying the definition. Then the function $\wp_1 - \wp_2$ is elliptic and may only have singularities at the points of Λ. The condition (ii) holds for both $\wp_1$ and $\wp_2$, so $\wp_1 - \wp_2$ may have at worst a simple pole at $z = 0$. Thus the degree of $\wp_1 - \wp_2$ is less than 2 and so, by the remark following Theorem 4.18 this function is constant. Then $\wp_1 - \wp_2$ must be zero, as its Laurent series at $z = 0$ has no constant term.

A similar argument applies to the difference $\wp(z) - \wp(-z)$ to show that it must be identically zero, and therefore $\wp$ must be an even function. Note, in particular, that the principal part of Laurent series for $\wp(z)$ at $z = 0$ has to be z^{-2} (i.e. no z^{-1} term).

The existence of $\wp$ is ensured by the following.

Theorem 6.25. *Weierstrass $\wp$-function can be expressed as*

$$\wp(z) = A - \frac{d^2}{dz^2} \log f(z) = A - \frac{d}{dz}\left(\frac{f'(z)}{f(z)}\right) \tag{6.26}$$

and also as

$$\wp(z) = C + D \cdot \mathbf{e}(z)\left(\frac{f(z + 1/2 + \tau/2))}{f(z)}\right)^2 = C + D \cdot \mathbf{e}(z)\left(\frac{\vartheta(z)}{\vartheta(z - 1/2 - \tau/2)}\right)^2 \tag{6.27}$$

for some constants $A, C, D \in \mathbb{C}$. Here $f(z) = \vartheta(z - 1/2 - \tau/2)$.

Proof. It is not difficult to see that both (6.26) and (6.27) give meromorphic functions.

Formula (6.26). We know that $f(z)$ has a simple zero at $z = 0$, so $f(z) = zg(z)$, $g(0) \neq 0$. Then

$$-\frac{d}{dz}\left(\frac{f'}{f}\right) = -\frac{d}{dz}\left(\frac{1}{z} + \frac{g'}{g}\right) = \frac{1}{z^2} - \left(\frac{g'}{g}\right)',$$

so the principal part at 0 is as required. Further, the choice $A = (g'/g)'(0)$ kills the constant term and thus the condition (ii) in the definition is satisfied.

Exercise 6.28. Verify the double-periodicity of the RHS in (6.26). (Hint: use the previously calculated (5.22).)

Formula (6.27). For any constants C and D the function given in (ii) is elliptic, by application of the argument in Theorem 5.24. We have (using the same notation) $a_1 = a_2 = -(1+\tau)/2$ and $b_1 = b_2 = 0$, so $a_1 + a_2 - b_1 - b_2 = -1 - \tau$ and $m = -1$. Thus the RHS of (6.27) is an elliptic function satisfying the condition (i) in the definition of the $\wp$-function. Choose the constant D so as to have coefficient 1 at z^{-2} in the Laurent expansion at $z = 0$. Choose the constant C to get rid of the constant term and now (ii) is satisfied too. □

The derivative $\wp'$ is an odd elliptic function of degree 3 with triple poles at the points of Λ.

Theorem 6.29. *The function $\wp'$ has simple zeros at $1/2$, $\tau/2$ and $(1+\tau)/2$.*

As $\deg \wp' = 3$, these are *all* the zeros of $\wp'$ up to addition of elements of the lattice of periods Λ.

Corollary 6.30. *If z is not a pole of the Weierstrass $\wp$-function then the valency of $\wp$ at z is given by*

$$v_\wp(z) = \begin{cases} 2, & \text{if } z - 1/2 \in \Lambda \text{ or } z - \tau/2 \in \Lambda \text{ or } z - (1+\tau)/2 \in \Lambda, \\ 1, & \text{otherwise.} \end{cases}$$

As $\wp$ is even, its Laurent expansion at $z = 0$ has only even powers of z and one can write

$$\wp(z) = \frac{1}{z^2} + 3E_4 z^2 + 5E_6 z^4 + \ldots = \frac{1}{z^2} + \sum_{n=2}^{\infty} (2k-1) E_{2k} z^{2n-2} \qquad (6.31)$$

for some complex constants E_{2k}.

In fact, it can be shown that the constants E_{2k} are the so-called ‘Eisenstein series’ (for the lattice Λ) given by

$$E_{2k} = \sum_{\substack{n,m\in\mathbb{Z} \\ (n,m)\neq(0,0)}} \frac{1}{(n+m\tau)^{2k}} = \sum_{\omega\in\Lambda\setminus\{0\}} \frac{1}{\omega^{2k}}.$$

(details can be found in Jones & Singerman). However, the calculation of E_{2k} will *not* be used here.

Theorem 6.32 (Differential equation for $\wp$)**.**

$$\wp'(z)^2 = 4\wp(z)^3 - 60E_4\wp(z) - 140E_6.$$

Proof. In the proof, notation $f_j(z)$, $j = 1, 2, \ldots$, will be used to denote some holomorphic function near $z = 0$ given by convergent power series. Then from the Laurent series (6.31) we have

$$\wp(z) = \frac{1}{z^2} + 3E_4 z^2 + 5E_6 z^4 + z^6 f_1(z).$$

Differentiating the above, we obtain

$$\wp'(z) = \frac{-2}{z^3} + 6E_4 z + 20E_6 z^3 + z^5 f_2(z),$$

and further

$$\wp'(z)^2 = \frac{4}{z^6} - \frac{24E_4}{z^2} - 80E_6 + z^2 f_3(z),$$

also,

$$4\wp(z)^3 = \frac{4}{z^6} + \frac{36E_4}{z^2} + 60E_6 + z^2 f_4(z).$$

Then, by examination of the coefficients at the non-positive powers of z in the above series, we obtain

$$\wp'(z)^2 - 4\wp(z)^3 + 60E_4\wp(z) + 140E_6 = z^2 f_5(z).$$

The LHS is an elliptic function, being a polynomial expression in the elliptic functions $\wp$ and $\wp'$. Further, the singularities of the function on the LHS may only occur at the points in Λ.

On the other hand, the RHS shows that the same function must have holomorphic extension over $z = 0$, hence, by periodicity, this function must extend holomorphically over any point of Λ. So $\wp'(z)^2 - 4\wp(z)^3 + 60E_4\wp(z) + 140E_6$ defines a holomorphic on $\mathbb{C}$ elliptic function and must be constant. Evaluating the RHS at $z = 0$ shows that this constant is zero. □

One consequence of Theorem 6.32 is that the inverse of the $\wp$-function may be expressed by an *elliptic integral*

$$\wp^{-1}(z) - \wp^{-1}(z_0) = \int_{z_0}^{z} \frac{dz}{\sqrt{4z^3 - 60E_4 z - 140E_6}}.$$

Note that $\wp^{-1}(z)$ is not uniquely defined and, respectively, the value of integral in the RHS depends on the choice of path joining z_0 and z.

7 The field of elliptic functions

If f, g are two meromorphic functions then $f \pm g$, fg, and (if $g \not\equiv 0$) f/g are again meromorphic. It is easy to see that meromorphic functions on $\mathbb{C}$ (more generally, on some fixed open domain) form a field. Elliptic functions having the same lattice of periods form a subfield of this field.

Theorem 7.33. *(i) If f is an even elliptic function then $f(z) = R(\wp(z))$, for some rational function R.*

(ii) If f is any elliptic function then $f(z) = \wp'(z)R_1(\wp(z)) + R_2(\wp(z))$, for some rational functions R_1, R_2.

III. Riemann surfaces and holomorphic maps.

8 The definitions and examples

For a general theory of Riemann surfaces, we need to recall some topological definitions.

A set X is called a **topological space** if there is a class of **open subsets** defined for X with the following properties: X, $\emptyset$ are open; the intersection of any two open sets is open; the union of any collection of open sets is open. A topological space X is **Hausdorff** if any two distinct points $x_1, x_2 \in X$ possess open neighbourhoods, $x_1 \in U_1,\ x_2 \in U_2$, such that $U_1 \cap U_2 = \emptyset$.

A map f between two topological spaces is **continuous** if the inverse image $f^{-1}(V)$ of any open set V is open. A **homeomorphism** is a continuous map having a continuous inverse.

8.1 Riemann surfaces

A **topological surface** X is a Hausdorff topological space such that any point p of X has an open neighbourhood $U = U_p$ homeomorphic to an open disc in $\mathbb{C}$.

The corresponding homeomorphisms $\varphi : U \subseteq X \to \varphi(U) \subseteq \mathbb{C}$ are called **(coordinate) charts**, and the domains U are called **coordinate neighbourhoods** in X, and $z \in \varphi(U)$ is a local coordinate on X.

The maps $\tau_{ij} = \varphi_j \circ \varphi_i^{-1}$ are called **transition functions**. Clearly, one has $\tau_{ji} = \tau_{ij}^{-1}$. Notice that the τ_{ij} are complex-valued functions on open domains in $\mathbb{C}$.

Fig. 3. Transition between two charts.

Definition. A **complex structure** on X is a system of charts with domains covering X and such that all the transition functions are *holomorphic* (then necessarily biholomorphic).
A **Riemann surface** is a topological surface endowed with a complex structure.

Examples:

- Any **connected open set in** $\mathbb{C}$ is in a obvious way a Riemann surface (with just one chart, the identity map).
  Likewise, it is not difficult to see that any connected open subset of a Riemann surface is again a Riemann surface. (Just take the charts for the original Riemann surface and intersect the coordinate neighbourhoods with that connected open subset...)

- The **Riemann sphere** S^2 with charts given by the two stereographic projections φ and ψ. We have already met this example in the Lectures.

The next example is closely related to the doubly-periodic meromorphic, i.e. elliptic, functions.

- **Elliptic curves**

Let $\Lambda = \{n + m\tau : n, m \in \mathbb{Z}\}$, $\operatorname{Im}\tau > 0$, be a lattice in the complex plane. An **elliptic curve** $E = \mathbb{C}/\Lambda$ is defined as a space of all 'cosets of Λ' in $\mathbb{C}$, where a coset of Λ is a subset of $\mathbb{C}$ of the form $z + \Lambda := \{z + \lambda : \lambda \in \Lambda\}$, for some $z \in \mathbb{C}$. Of course, different z's may define the same cosets of Λ. More precisely, there is an "obvious" projection, sometimes called the **quotient map**, $\pi : \mathbb{C} \to \mathbb{C}/\Lambda$ assigning to each z the coset $z + \Lambda$. Then $\pi(z) = \pi(z')$, i.e. $z + \Lambda = z' + \Lambda$, if and only if $z - z' \in \Lambda$.

Define a subset W of $E = \mathbb{C}/\Lambda$ to be open when its inverse image under π, $\pi^{-1}(W) = \{z \in \mathbb{C} : \pi(z) \in W\}$ is open in $\mathbb{C}$. With this definition, E becomes a topological space, moreover a Hausdorff space (I leave it as an optional exercise in topology to verify the latter claim), and π a continuous map.

> The topology we defined on $\mathbb{C}/\Lambda$ is sometimes called the *quotient topology*. It is the strongest topology (informally, the maximal possible class of open sets) on $\mathbb{C}/\Lambda$ so that the quotient map π is a continuous map.

Consider a family of (open) discs D_i covering the complex plane, $\cup_i D_i = \mathbb{C}$. Assume that the diameter of each disc D_i is less than $\min\{\frac{1}{2}, \operatorname{Im}\frac{\tau}{2}\}$. Then D_i intersects any coset of Λ in at most one point, that is, if both z and $z + \lambda$ are in D_i, for some $\lambda \in \Lambda$, then $\lambda = 0$. Therefore, π maps D_i one-to-one to an open subset $U_i = \pi(D_i)$ in E. Define $\varphi_i : U_i \to D_i$ to be the inverse map, so $\pi \circ \varphi_i = \mathrm{id}_{U_i}$. It follows directly from the construction of φ_i that φ_i is a homeomorphism. Then, given two such maps, say $\varphi_i : U_i \to D_i$ and $\varphi_j : U_j \to D_j$, with nonempty $U_i \cap U_j$, the transition map is $\varphi_i \circ \varphi_j^{-1}(z) = z + \lambda_{ij}$, for some $\lambda_{ij} \in \Lambda$, certainly holomorphic.

To see that indeed $\varphi_i \circ \varphi_j^{-1}(z) = z + \lambda_{ij}$ note that we must have $\varphi_i \circ \varphi_j^{-1}(z) - z \in \Lambda$ for every z, from the definition of the maps φ_i. Now $\varphi_i \circ \varphi_j^{-1}(z) - z$ is a continuous function on a connected set $\varphi_i(U_i \cap U_j)$ and taking values in a **discrete** set Λ, so $\varphi_i \circ \varphi_j^{-1}(z) - z = \lambda_{ij}$ must be constant, independent of z. (The upper bound on the diameters of D_i,D_j ensures that $\varphi_i(U_i \cap U_j) \subseteq D_i \subset \mathbb{C}$ is connected for all i, j.)

Therefore, the system of open subsets $U_i \subset E$ and the maps (the charts) φ_i is a *complex structure*, making E into a *Riemann surface*.

The underlying topological surface is a torus.

8.2 Holomorphic maps between Riemann surfaces

Let S, S' be two Riemann surfaces with complex structures given respectively by the charts $\varphi_i : U_i \to \mathbb{C}$ and $\psi_\alpha : W_\alpha \to \mathbb{C}$.

Definition. A continuous map $f : S \to S'$ is said to be **holomorphic** when $\psi_\alpha \circ f \circ \varphi_i^{-1}(z)$ are holomorphic functions of z, whenever the composition is well-defined (i.e. whenever the set $\varphi_i(U_i \cap f^{-1}(W_\alpha))$ is non-empty).

Examples:

- Holomorphic functions on open sets in $\mathbb{C}$
- Meromorphic functions on open sets $U \subset \mathbb{C}$ are holomorphic maps $U \to S^2$ to the Riemann sphere (as we have seen).

- We have also seen that the rational functions on $\mathbb{C} \cup \{\infty\}$ are precisely the holomorphic maps from the Riemann sphere to itself.
- Let Λ be a lattice in $\mathbb{C}$. The projection $\pi : z \in \mathbb{C} \to z + \Lambda \in \mathbb{C}/\Lambda$ is a holomorphic map. This is an immediate corollary of the construction of the complex structure on $\mathbb{C}/\Lambda$.
- It follows from the discussion of elliptic curves that the Λ-periodic meromorphic functions descend, with the help of projection π, to give holomorphic maps $\mathbb{C}/\Lambda \to S^2$.

Definition. A map between two Riemann surfaces is called **biholomorphic**, or **conformal equivalence**, if it is holomorphic and has a holomorphic inverse.

Notice that with this definition, any chart φ on a Riemann surface is a biholomorphic map of a coordinate neighbourhood U onto complex plane domain $\varphi(U) \subset \mathbb{C}$.

Remark. 'Biholomorphic' is a much more rigid kind of equivalence than 'homeomorphic' (see e.g. Example sheet 2 question 7).

9 A closer look at holomorphic maps

Throughout this section, let R and S be Riemann surfaces such that both are *connected* and let $f : R \to S$ be a non-constant holomorphic map. A number of the most important properties of holomorphic functions on open sets in $\mathbb{C}$ carries over to this more general setting. In order to see this, the following notation will be useful. Suppose p is a point of R and q is a point of S and $f(p) = q$. We can choose charts φ and ψ 'centred' at p and q, this means, respectively, $\varphi(p) = 0$ and $\psi(q) = 0$. Then $\hat{f} = \psi \circ f \circ \varphi^{-1}$ is a holomorphic function defined near $z = 0$, with $\hat{f}(0) = 0$. Of course, $\hat{f}$ depends on the choice of centred charts φ and ψ. However, a different choice gives a function, $\hat{f}_1$ say, such that $\hat{f}_1(z) = v(\hat{f}(u^{-1}(z)))$, where u and v are conformal equivalences (biholomorphic maps) near $z = 0$, with $u(0) = 0$, $v(0) = 0$.

'Isolated values principle'. If f is not constant and R is connected then $f^{-1}(q)$ is a discrete subset of R, for any $q \in S$.

Proof. Form two subsets X, Y of R:

> $x \in X$ precisely if there exists an open neighbourhood U_x of x such that f is constant on U_x;
>
> $y \in Y$ precisely if there is an open neighbourhood V_y of y such that $f(y) \neq f(y')$ for all $y' \in V_y$ with $y \neq y'$.

From the definitions, X and Y are disjoint and X is open. We claim that Y is open and $X \cup Y = R$. To see this, let $p \in R$, $f(p) = q$, with φ and ψ centred charts as above, and $\hat{f} = \psi \circ f \circ \varphi^{-1}$ the corresponding holomorphic function defined near 0. Since $\hat{f}$ is holomorphic, there are just two possibilities: either $\hat{f}$ is constant near 0 or 0 is an isolated zero of $\hat{f}$: $\hat{f}(z) \neq 0$ for $z \neq 0$. In the first case f is identically constant on a neighbourhood of p and $p \in X$; in the second case, similarly, p satisfies the conditions to be a point of Y.

Also in the second case, as $\hat{f}$ is non-constant, the zeros of any function of the form $\hat{f}(z) - \hat{f}(z_0)$ are isolated for z_0 in some neighbourhood of 0 (in the domain of $\hat{f}$). Therefore, by applying the chart φ, we obtain a neighbourhood of y contained in Y and so Y is open.

Since R is connected, we have either $X = R$ or $Y = R$. Hence f is either constant or for any $q \in S$ the points of $f^{-1}(q)$ are isolated in R. $\square$

The branching order. Suppose that $f(p) = q$ and let φ, ψ and $\hat{f}$ be as above. Assume that $\hat{f}$ is non-constant. Since $\hat{f}(0) = 0$, there is an integer n, the order of this zero. This integer is called the **branching order** of f at p and will be denoted by $v_f(p)$. If $v_f(p) > 1$ then p is called a **ramification point** of f and q is called a **branch point**. Note that $v_f(p)$ is also the valency of $\hat{f}$ at $z = 0$ and is therefore unchanged if $\hat{f}$ is replaced by $v \circ \hat{f} \circ u^{-1}$, for any conformal equivalences u, v. That is, $v_f(p) = v_{\psi \circ f \circ \varphi^{-1}}(z)$, $z = \varphi(p)$, for any charts φ, ψ near p, q. Thus $v_f(p)$ is well-defined and generalizes the valency of holomorphic function.

The set of ramification points of a non-constant holomorphic f on a connected R is a **discrete** *subset of R.* To see this, suppose that p is a ramification point of f. By definition this means that the corresponding map $\hat{f}$ has a multiple zero at $z = 0$, so $\hat{f}'(0) = 0$. Since $\hat{f}$ is non-constant, the zeros of $\hat{f}'$ are isolated, so there is a neighbourhood D of 0 such that $\hat{f}'(z) \neq 0$ if $z \neq 0$, $z \in D$. Now if $U = \varphi^{-1}(D)$ then U is open in R and the previous statement translates into the statement that p is the only ramification point of f in U. Thus any ramification point p is isolated, as claimed.

The local structure of holomorphic maps. If the branching order of f at p is equal to n then there exist a neighbourhood V of q and a neighbourhood U of p such that for all $q' \in V$, such that $q' \neq q$, the set $f^{-1}(q') \cap U$ consists of precisely n points.

This was proved for holomorphic functions on open subsets of $\mathbb{C}$. For the general result, apply this theorem to the holomorphic function $\hat{f}$ and then transfer back to R and S using the charts (which are bijections). Further, the following consequences are proved similarly to the case of maps between domains in $\mathbb{C}$.

Open Mapping Theorem. A non-constant holomorphic map between connected Riemann surfaces takes open sets to open sets. (In particular, $f(R)$ is open in S.)

Inverse Mapping Theorem. Suppose that the branching order of f at p is equal to 1. Then there are neighbourhoods U and V, of p and q, respectively, and a holomorphic map $g : V \to U$, such that $g \circ f = \mathrm{id}_U$ and $f \circ g = \mathrm{id}_V$. Such g is called a **local inverse** of f.

The degree of holomorphic map. Let f, R, S be as above (thus f is non-constant and R, S are connected) and assume also that R and S are **compact**. If q is any point of S define

$$k(q) = \sum_{p \in f^{-1}(q)} v_f(p).$$

The sum is finite because $f^{-1}(q)$ is discrete (by the isolated values principle above) and closed in the compact R, thus itself compact, and any discrete compact set is finite. $k(q)$ is the number of solutions of the equation $f(p) = q$, where the solutions are counted according to multiplicity. We claim that $k(q)$ is independent of q.

Proof. For any p such that $f(p) = q$, considering the local structure of holomorphic f near p, we can choose neighbourhoods N_p of p and V_p of q, with the property that the equation $f(x) = q'$ has precisely $v_f(p)$ solutions x in N_p if $q' \neq q$ and $q' \in V_p$. It also follows that $v_f(x) = 1$ for any of these solutions in N_p. Put

$$V = \cap_{p \in f^{-1}(q)} V_p, \qquad U_p = f^{-1}(V) \cap N_p.$$

By shrinking V if necessary, we may assume that the U_p are disjoint subsets of R. Further shrinking V we can find $\tilde{V}$ such that $q \in \tilde{V} \subseteq V$ and $f^{-1}(\tilde{V})$ is the union of $\tilde{U}_p$, for all $p \in f^{-1}(q)$, where $\tilde{U}_p = f^{-1}(\tilde{V}) \cap N_p$. The existence of such $\tilde{V}$ is a consequence of the compactness of R and the continuity of f.

Indeed, assume that there is no such $\tilde{V}$. This implies the existence of two sequences: $\{q_n\}$, $q_n \neq q$, with $\lim_{n\to\infty} q_n = q$, and $\{p_n\}$, $f(p_n) = q_n$, such that, for any n, p_n is not in any of the U_p's. But since R is compact, the sequence p_n must have a converging subsequence. The limit of that subsequence must be some $p \in f^{-1}(q)$, by the property of continuous functions

$$f(\lim_{m\to\infty} p_m) = \lim_{m\to\infty} f(p_m).$$

This yields a contradiction with the assumption that $p_n \notin U_p$ for any n and any $p \in f^{-1}(q)$.

Now, for the pair of sets U_p and V we still have that for any $q' \neq q$, $q' \in V$, there are precisely $v_f(p)$ solutions of $f(x) = q'$ in U_p. In light of this and since U_p are disjoint, we conclude that in the whole of R there are

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

solutions of $f(x) = q'$. This shows that the map $y \mapsto k(y)$ is constant near $y = q$ (remember that, by the above arrangements, $v_f(x) = 1$ for any x in any U_p if $f(x) \neq q$). In particular, the map $k : S \to \mathbb{Z}$ is continuous near each $q \in S$. But S is connected, so k must be constant on all of S and the claim is proved. □

The integer $k = k(q)$ (any $q \in R$) is called the **degree** of the map f.

In the case of holomorphic maps $S^2 \to S^2$ and $\mathbb{C}/\Lambda \to S^2$ this definition gives the (previously considered) degree of, respectively, the rational functions and the elliptic functions.

10 Triangulations and the Euler characteristic

Let S be a compact connected Riemann surface. In this section, we temporarily disregard a complex structure and consider S as a **topological** surface. This topological surface is compact, connected, and is also necessarily orientable (informally: 'has two sides').

For a precise definition of orientability, start from a real vector space, W say. The orientation of W is an equivalence class of bases of W, where two bases are equivalent (have the same orientation) if and only if the matrix expressing the transition between these bases has positive determinant. A surface is locally modelled on open neighbourhoods in $\mathbb{R}^2$ with two different local models related by a smooth map $U \subset \mathbb{R}^2 \to \mathbb{R}^2$ on the overlap U. This map is, of course, non-linear in general but has at each point a 'good linear approximation', the *derivative*. The derivative is given by a 2×2 matrix of partial derivatives and acts on the vectors in $\mathbb{R}^2$. Then, for our surface to be *orientable*, we require that it has an open cover by local coordinate patches, so that for all the corresponding changes of local coordinates, the 2×2 matrices of partial derivatives 'preserve the orientation'—i.e. have positive determinant at every point where these are defined.

In the case of a Riemann surface, we are therefore looking at linearizations of holomorphic functions, i.e. the maps $z \in \mathbb{C} \to \tau'(z_0)\, z \in \mathbb{C}$ (for all z_0 in local coordinate patches and all biholomorphic τ's expressing transitions between local coordinates); the corresponding 2×2 real matrices *always* have positive determinant.

Compact connected orientable surfaces are **classified up to a homeomorphism** by their Euler characteristic χ (or, equivalently, by the genus). In order to define the Euler characteristic, we shall need a concept of **triangulation**. I shall only include here an outline (not aiming for maximum generality), assuming some statements without proofs.

Informally, a triangulation T is cutting a surface S into a finite number of 'polygonal' regions, called faces, by smooth non-self-intersecting arcs, called edges, joined at vertices (so that a triangulated surface looks like a 'topological polyhedron'). More precisely, an **edge** of T is a homeomorphic image in S of the interval $[0,1] \subset \mathbb{R}$ and the images of 0 and 1 are **vertices** of T. The complement in S of the edges of T consists of (finitely many) connected components; each one is required to be homeomorphic to an open disc. The **faces** of T are the closures of these components. In addition, one requires the following properties:

- any two faces share only one edge, if at all; each edge belongs to the boundaries of exactly two faces;
- two edges meet only in one common end-point (vertex), if at all;
- any vertex has a neighbourhood homeomorphic to an open disc with edges corresponding to rays from the centre to the boundary of that disc. Any two distinct sectors cut out by these rays correspond to distinct faces of T. (Consequently, at least 3 edges meet at each vertex.)

Remark. In the literature, there are some variations on what is allowed or disallowed in a triangulation. However, these variations will not be important for us as they lead to the same Euler characteristic.

Let $V(T), E(T), F(T)$ denote the number of vertices, edges, and faces of T. A remarkable fact is that the quantity
$$\chi(S,T) = V(T) - E(T) + F(T)$$
is independent of the choice of a triangulation of S. Therefore it can be written as $\chi(S)$ and is called the **Euler characteristic** of S. The **genus** $g(S)$ is related to this by $\chi(S) = 2 - 2g(S)$ and $g(S)$ can be visualized as 'the number of handles that one needs to attach to the sphere in order to obtain S'. In particular, $g(S) \geq 0$ and so $\chi(S) \leq 2$.

For example, $\chi(S^2) = 2$ and $g(S^2) = 0$; the torus has $\chi = 0$ and $g = 1$.

Fig. 4. Triangulating a 'topological pretzel': $\chi = 24 - 44 + 18 = -2, \;\; g = 2$.

You may notice from the above that $\chi(S)$ is even. Remember though that we are only considering *orientable* surfaces here. An example of topological surface with an odd χ is the quotient $S^2/\pm 1$ by the antipodal map (it has $\chi = 1$), but it is not orientable (and cannot arise as a Riemann surface).

We shall assume without proof the following topological result.

Theorem 10.34. Every topological surface underlying a compact Riemann surface has a triangulation.

What really matters here is whether a surface S is a so-called 'second countable' topological space. By definition, second countable means that there is a countable family of open neighbourhoods $U_n \subset S$, $n = 1, 2, ...$, so that any open subset of S can be obtained as a union of some U_n's.

The plane $\mathbb{R}^2$ (and any $\mathbb{R}^n$) is second countable as one can choose the family of all the open discs with rational radii and with centres whose coordinates are rational numbers.

A connected surface has a triangulation if and only if it is 'second countable'. Prüfer gave an example of a connected surface which is not second countable (and so cannot be triangulated). Such examples are necessarily quite pathological and it is common to exclude them by requiring the second countable property. But it can be proved that a *connected* Riemann surface is always second countable (which is why this condition was not included in the definition). The proof is not obvious, see, for example, G.Springer 'Introduction to Riemann surfaces'; you will also find Prüfer's example there.

11 The Riemann–Hurwitz formula

We have seen that for any holomorphic map between two compact connected Riemann surfaces one can consider

- an **analytic** quantity 'the branching order at a point' (defined using power series);
- an **algebraic** quantity 'the degree' (generalizes the degree of a polynomial);
- **topological** quantities: 'Euler characteristic' or 'genus' (depend only on the topological surface underlying a given Riemann surface).

The following result gives a very useful relation between the above three quantities.

Theorem 11.35 (Riemann–Hurwitz formula)**.** *Let R and S be compact connected Riemann surfaces and $f : R \to S$ a non-constant holomorphic map of degree k. Then*

$$\chi(R) = k\,\chi(S) - \sum_{p\in R}(v_f(p) - 1),$$

$$g(R) - 1 = k(g(S) - 1) + \frac{1}{2}\sum_{p\in R}(v_f(p) - 1).$$

Proof (assuming Theorem 10.34). The idea is to consider a triangulation of S and 'pull it back' to R. Suppose first that there is no branching. This means that every point $q \in S$ has a neighbourhood V_q such that $f^{-1}(q) = U_q$ is a disjoint union of open sets, each of which is mapped biholomorphically onto V_q by f. Choose a triangulation T of S so fine that every face lies in some one of the V_q.

(A triangulation can be refined by performing sufficiently many times a 'barycentric subdivision' procedure: mark a point, a new vertex, in the interior of a face and further new vertices in the middle of each edge of this face. Join the new vertex in the interior by new edges with those in the middle of the edges and also with all the 'old' vertices on the boundary of this face.)

Fig. 5. Barycentric subdivision.

Then f lifts back to a triangulation T' of R which has k times as much of everything that T has. So $\chi(R) = k\chi(S)$ in this case.

Fig. 6. No branching ($k = 2$).

Next let f have just one ramification point, say p with $v_f(p) = r \leq k$ and $f(p) = q$. Make sure q is a vertex of the triangulation T of S (perform a subdivision if necessary). Then the map f near q looks like $z \mapsto z^r$ near $z = 0$, which is $r : 1$ for $z \neq 0$. Therefore, lifting the triangulation back to R, as before, will yield k times of everything—except for the vertex q, which has only $k - (r - 1)$ preimages in R. So now one has $\chi(R) = k\chi(S) - (r - 1)$.

Fig. 7. Near a ramification point[3] ($k = 4$, $r = 4$).

In general, if there is more than one ramification point, then treat them in turn as above to obtain the Riemann–Hurwitz formula as claimed. □

[3]Here I only showed some relevant edges of T in S and $f^{-1}(T)$ in R.

IV. Algebraic curves

12 Curves in $\mathbb{C}^2$

Definition. Let $P(s,t)$ be a non-constant complex polynomial in two variables. A (complex) **algebraic curve** C in $\mathbb{C}^2$ is the zero set of P, $C = \{(s,t) \in \mathbb{C}^2 : P(s,t) = 0\}$.

An algebraic curve C is **non-singular** if $(\partial P/\partial s(s_0,t_0), \partial P/\partial t(s_0,t_0)) \neq (0,0)$, for any point (s_0,t_0) in C.

Theorem 12.36. *A non-singular algebraic curve is, in a natural way, a Riemann surface.*

The proof uses the following result from complex analysis.

Implicit Function Theorem. *Suppose that $P(s,t) = \sum P_{m,n} s^m t^n$ is a complex polynomial in two variables $(s,t) \in \mathbb{C}^2$, with $P(s_0,t_0) = 0$ and $\partial P/\partial t(s_0,t_0) \neq 0$.*

Then there is a unique holomorphic function $\varphi : D(s_0,\varepsilon) \to D(t_0,\delta)$ between discs with small radii ε, δ and such that $\varphi(s_0) = t_0$ and $P(s,\varphi(s)) \equiv 0$.

In fact, the proof of the Implicit Function Theorem only requires that $P(s,t)$ be holomorphic in each of the variables s,t with the other variable held fixed.

Proof of Theorem 12.36 assuming the Implicit Function Theorem. As $C \subset \mathbb{C}^2$ is a metric space it is Hausdorff. We need to show that C is a topological surface and find a complex structure on C.

Define $f_1, f_2 : C \to \mathbb{C}$ to be the restrictions to C of the projection maps $(s,t) \to s$ and $(s,t) \to t$ respectively.

Consider an arbitrary point (s_0,t_0) in C. We can assume, without loss of generality, that $\partial P/\partial t(s_0,t_0) \neq 0$. Then the Implicit Function Theorem implies that, for small positive ε, δ the set $U = C \cap \{(s,t) \in \mathbb{C}^2 : |s - s_0| < \varepsilon, |t - t_0| < \delta\}$ is a graph of a uniquely determined holomorphic function $\varphi(s)$ such that $\varphi(s_0) = t_0$ and $P(s,\varphi(s)) \equiv 0$. That is,

$$U = \{(s,\varphi(s)) : |s - s_0| < \varepsilon\}$$

and so f_1 projects U homeomorphically onto the disc $\{|s - s_0| < \varepsilon\}$ in $\mathbb{C}$. Thus $f_1|_U$ is a coordinate chart for C near (s_0,t_0). Proceeding in this fashion, we find near each point a chart given by the first or second projection. Thus C fits the definition of a topological surface.

Furthermore, with the coordinate charts on C as defined above, the transition function for a pair of charts given by projection to the same variable is an identity map. Otherwise, the transition function has the form

$$s \overset{f_1^{-1}}{\mapsto} (s,\varphi(s)) \overset{f_2}{\mapsto} \varphi(s)$$

and φ is holomorphic, by the Implicit Function Theorem. □

Thus the curve C has a natural complex structure. Note that with this complex structure the projections $f_1, f_2 : C \to \mathbb{C}$ become holomorphic maps.

An algebraic curve $\{P(s,t) = 0\}$ is said to be **irreducible** when the polynomial P cannot be written as a product $P(s,t) = p(s,t)q(s,t)$ of two non-constant polynomials p,q. It can be shown that any irreducible algebraic curve is *connected*.

13 Projective curves

There is a way to extend a complex vector space by adding points 'at infinity' corresponding to complex 1-dimensional subspaces (lines). This leads to a concept of projective space and, further, to the concept of projective algebraic curves. An advantage of projective curves, over curves in $\mathbb{C}^2$ (alias affine curves) is that projective curves have better topological properties, in particular are always *compact*.

Definition. The **projective space** $\mathbb{P}^n$ is the space of all the complex lines in $\mathbb{C}^{n+1}$ passing through the origin (i.e. 1-dimensional subspaces of $\mathbb{C}^{n+1}$). The points of $\mathbb{P}^n$ are denoted by $z_0 : z_1 : \ldots : z_n$, where not all z_i are zero. The z_i's are called the **homogeneous coordinates** and $z_0 : z_1 : \ldots : z_n = \lambda z_0 : \lambda z_1 : \ldots : \lambda z_n$ are identified, for any complex $\lambda \neq 0$.

It will be convenient to introduce the projection

$$\Pi : (z_0, z_1, \ldots, z_n) \in \mathbb{C}^{n+1} \setminus \{0\} \to z_0 : z_1 : \ldots : z_n \in \mathbb{P}^n.$$

We use on $\mathbb{P}^n$ the *quotient topology*: $W \subset \mathbb{P}^n$ is open if an only if $\Pi^{-1}(W) \subset \mathbb{C}^{n_1} \setminus \{0\}$ is so. With this topology, the map Π is continuous. Furthermore, it is not difficult to check that Π maps the affine space $A_j = \{(z_0, z_1, \ldots, z_n)\mathbb{C}^{n_1} | z_j = 1\}$, for each $i = 0, 1, 2, \ldots$, homeomorphically onto its image. Each A_j is homeomorphic to $\mathbb{C}^n$ and $\mathbb{P}^n - A_j$ is given by $z_j = 0$, thus identified as a copy of $\mathbb{P}^{n-1}$. We find that $\mathbb{P}^n$ decomposes as $\mathbb{C}^n \cup \mathbb{P}^{n-1}$.

We also find that the projective space $\mathbb{P}^n$ is *Hausdorff* and *compact*. The Hausdorff property follows by exploiting that each A_j is Hausdorff. For the compactness, notice that $\mathbb{P}^n = \Pi\big(\{\sum_{i=0}^n |z_n|^2 = 1\}\big)$ and the unit sphere $\{\sum_{i=0}^n |z_n|^2 = 1\}$ is compact in $\mathbb{C}^{n+1}$.

In these notes, we shall only consider the projective plane $\mathbb{P}^2$ and the projective line $\mathbb{P}^1$.

Definition. Let $P(X, Y, Z)$ be a non-constant homogeneous complex polynomial in three variables. A **projective curve** C in $\mathbb{P}^2$ is the zero set of P, $C = \{X : Y : Z \in \mathbb{P}^2 \mid P(X, Y, Z) = 0\}$.

A projective curve C is **non-singular** if $(\partial P/\partial X(a, b, c), \partial P/\partial Y(a, b, c), \partial P/\partial Z(a, b, c)) \neq (0, 0, 0)$, for any point $a : b : c$ in C.

Theorem 13.37. *A non-singular projective curve in $\mathbb{P}^2$ is, in a natural way, a Riemann surface. Moreover, this Riemann surface is* compact*.*

Proof (gist). The basic idea is to consider 'affine pieces' of $C = \{P(X, Y, Z) = 0\}$, i.e. the algebraic curves C_j, $j = 0, 1, 2$, in $\mathbb{C}^2 \subset \mathbb{P}^2$ defined by $C_j = \{X_0 : X_1 : X_2 \in C \mid X_j \neq 0\}$. Thus e.g. $C_0 = \{P(1, s, t) = 0\}$ (put $s = Y/X$, $t = Z/X$) and $C_0 \subset C$ is an open subset.

The key property is that if $a : b : c \in C$ is a non-singular point, with $a \neq 0$ say, then $(b/a, c/a)$ is a non-singular point in C_0 (and similarly for points in C_1 and C_2). So if $a \neq 0$ then $\partial P/\partial Y(a, b, c)$ or $\partial P/\partial Z(a, b, c)$ is non-zero and if say $\partial P/\partial Z(a, b, c) \neq 0$ then $X : Y : Z \in C \to Y/X$ is well-defined near $a : b : c$ and may be used for a chart (it corresponds to a first projection chart on the C_0). Thus we put together all the charts defined in the proof of Theorem 12.36, for each $C_j \subset C$. It remains to check, by inspection, that for the charts defined in this way on C the transition functions are always holomorphic.

As $\mathbb{C}^3 \setminus P^{-1}(0)$ is open it follows that $C \subset \mathbb{P}^2$ is closed, hence compact. □

Proposition 13.38. *The projective line $\mathbb{P}^1$ is a Riemann surface biholomorphic to the Riemann sphere.*

Remarks. A projective curve C may be thought of as a *compactification* of its affine piece, say $C_0 \subset \mathbb{C}^2$. (Note that an algebraic curve in $\mathbb{C}^2$ is never compact, Ex.sheet 3 Q1.) Conversely, given an algebraic curve C_0 in $\mathbb{C}^2$ it is easy to determine the equation of its 'projective version' C, so that C_0 is identified with $C - \{\text{finite set of points}\}$. However, C need not automatically be a Riemann surface as the added points may be singular.

With some further work in algebra, one can show that a non-singular projective curve $C = \{P(X, Y, Z) = 0\}$ in $\mathbb{P}^2$ is always irreducible (i.e. $P(X, Y, Z) \neq p(X, Y, Z)q(X, Y, Z)$ for any non-constant homogeneous polynomials p, q) and hence *connected*. This is another topological advantage of working with projective (rather than affine) curves. We shall assume the latter result without proof. (Details can be found in F.Kirwan 'Complex algebraic curves'.)

14 Branched covers

Definition. Let R and S be Riemann surfaces and $F : R \to S$ a non-constant holomorphic map. The set $B \subset S$ of all the branch points of F is called the **branch locus** of F and one says that R is a **cover of S branched over** B.

The concept of a branched cover is useful e.g. when a surface S is easier to understand than R. In these notes we shall mostly look at branched covers of open subsets of $\mathbb{C}$.

If R and S are *compact* then the properties of the degree of F imply that the restriction of F to $R \setminus F^{-1}(B)$ is a covering map as defined in Algebraic Topology. On the other hand, F near any point $x \in F^{-1}(B)$ (in particular, near any ramification point), is expressed in suitable local coordinates by a simple formula $z \mapsto z^n$, where $n = v_F(x)$.

The following result gives a systematic way to find ramification points and branch locus for the restriction of the first projection to an algebraic curve in $\mathbb{C}^2$.

Proposition 14.39. *Let $C = \{P(s, t) = 0\}$ be a non-singular complex algebraic curve in $\mathbb{C}^2$. Let f be the restriction to C of the first projection $(s, t) \to s$.*

If $(s_0, t_0) \in C$ then $v_f(s_0, t_0) > 1$ if and only if $(\partial P/\partial t)(s_0, t_0) = 0$. Moreover, $v_f(s_0, t_0) = n$ if and only if

$$\frac{\partial P}{\partial t}(s_0, t_0) = \ldots = \frac{\partial^{n-1} P}{\partial t^{n-1}}(s_0, t_0) = 0, \quad \frac{\partial^n P}{\partial t^n}(s_0, t_0) \neq 0.$$

The proof uses the charts defined on an algebraic curve as in the proof of Theorem 12.36. Since $\partial P/\partial t = 0$ at (s_0, t_0), the chart is given by the restriction of the second projection $(s, t) \to t$.

One interesting application is the following.

Theorem 14.40. *There exist compact Riemann surfaces of any genus g.*

Proof (gist). For the proof we consider a non-singular algebraic curve C in $\mathbb{C}^2$ defined by $t^2 - h(s) = 0$, where h is a polynomial of degree $2g + 2$ ($g \geq 0$) with no repeated roots. We then define another curve $Y = \{(z, w) \in \mathbb{C}^2 : w^2 - k(z) = 0\}$, with $k(z) = z^{2g+2}h(1/z)$. Let X be the disjoint union of C and Y, where each point $(s, t) \in C$ with $s \neq 0$ is identified with $(z, w) = (1/s, t/s^{g+1}) \in Y$. Informally, X is the 'gluing' of C and Y according to the above relation. We may think of C and Y as subsets of X and then each $X \setminus C$ is a set of one or two points (depending on whether $h(0) = 0$) and $X \setminus Y$ is a set of exactly two points.

We make X into a topological space by defining $U \subset X$ to be open precisely if both $U \cap C$ and $U \cap Y$ are open. Furthermore, from the property that points of C and Y are identified via a biholomorphic map we deduce that the standard charts of C and Y together give a complex structure on X, so X is a Riemann surface.

The Riemann surface X is *compact*, being a union of two compact sets

$$X = \{(s,t) \in C : |s| \leq 1\} \cup \{(z,w) \in Y : |z| \leq 1\}.$$

Indeed $\{(s,t) \in C : |s| \leq 1\}$ is closed as it is the inverse image of the closed unit disc under the continuous first projection $C \to \mathbb{C}$. This latter set is also bounded as $|t|^2 = |h(s)|$ is bounded on a compact set $|s| \leq 1$ and thus a compact subset in $\mathbb{C}^2$. Similarly, $\{(z,w) \in Y : |z| \leq 1\}$ is compact. It turns out that X is also *connected* as both C and Y are so; we shall assume the latter claim without proof.

The first projections $C_0 \to \mathbb{C}$, $Y_0 \to \mathbb{C}$ induce a well-defined holomorphic map f from X to the Riemann sphere $\mathbb{P}^1$. The map f has degree 2 because there are infinitely many points in $\mathbb{P}^1$ with precisely two pre-images, but X and $\mathbb{P}^1$ are compact, so f has only finitely many branch points with the number of pre-images strictly less than $\deg f$. We further find that f has $2g+2$ ramification points, all with branching order 2. Then the Riemann-Hurwitz formula computes the genus $g(X) = g$. □

To show that C is connected is not difficult, but is largely a matter of topology. Let (s_1, t_1) be a point in C. Choose some root s_0 of the polynomial h. We shall check that there is a continuous path in C connecting (s_1, t_1) to $(s_0, 0)$. Since (s_1, t_1) was chosen arbitrarily this will show that C is connected. Firstly, $(s_0, 0)$ has a neighbourhood V in C homeomorphic to a disc, via a chart. The first projection $p : C \to \mathbb{C}$ is holomorphic, so $p(V)$ is a connected open set in $\mathbb{C}$; we may assume that $p(V)$ does not contain any roots of h. Let $s_* \in p(V)$ and $s_* \neq s_0$; it suffices to connect (s_1, t_1) to some $(s_*, t_*) \in V \subset C$.

Consider a path $\alpha(x)$ in $\mathbb{C}$ $(0 \leq x \leq 1)$ with $\alpha(0) = s_1$ and $\alpha(1) = s_*$, and such that α does not pass through any roots of h. Then (i) for every x, $p^{-1}(\alpha(x))$ is exactly two points in C and (ii) p gives a chart of C near each point in $p^{-1}(\alpha(x))$. It follows that for each $\alpha(x)$, there is a disc $D_x = D(\alpha(x), \varepsilon_x)$ such that $p^{-1}(D_x)$ is the disjoint union of two open sets D_x^+, D_x^- in C each homeomorphic to D_x. The discs D_x form an open cover of a compact path $\alpha([0,1])$, so there is a finite subcover $D_1, \ldots, D_N$, $\alpha(0) \in D_1$. We can now construct the desired path in C: for each D_i use a homeomorphism to either D_i^+ or D_i^-. We shall be done after N steps. Thus C is connected.
That Y is connected is checked similarly.

15 Meromorphic differentials

Recall that holomorphic maps from a Riemann surface, S say, to the Riemann sphere may be identified with the meromorphic functions on S (cf. Propn. 2.12 and Sec. 8.2). More explicitly, if we use the $\mathbb{P}^1$ model of the Riemann sphere (Propn. 13.38) then the meromorphic function corresponding to a non-constant holomorphic map $F : x \in S \to F_0(x) : F_1(x) \in \mathbb{P}^1$ is just $f(x) = F_0(x)/F_1(x)$. For each chart $\varphi_\alpha : V_\alpha \to \mathbb{C}$ on S, the composition $f \circ \varphi_\alpha^{-1}$ defines a meromorphic function on $\varphi_\alpha(V_\alpha) \subset \mathbb{C}$ with poles precisely at $F_1^{-1}(0) \cap \varphi_\alpha(V_\alpha)$. As any such f takes values in $\mathbb{C}$ the addition and multiplication is well-defined and the meromorphic functions on S form a field.

Meromorphic functions on a Riemann surface S are supposed to be generalizations of the meromorphic functions on open domains in $\mathbb{C}$. Can we make sense of the derivative of a meromorphic function f on a general Riemann surface?

For any coordinate chart φ_α on a neighbourhood of $x \in S$ the derivative of $f_\alpha(z) = f(\varphi_\alpha^{-1}(z))$ is defined. However, this depends on the choice of a coordinate chart! For any two local coordinate expressions $f_\alpha(z), f_\beta(z)$, we can write $f_\alpha = f_\beta \circ \tau_{\alpha\beta}$, where $\tau_{\alpha\beta} = \varphi_\beta \circ \varphi_\alpha^{-1}$ is the transition function. The respective derivatives are then related via

$$f'_\alpha(z) = f'_\beta(\tau_{\alpha\beta}(z)) \cdot \tau'_{\alpha\beta}(z) \tag{15.41}$$

by the chain rule.

We find that the derivative of f is **not** a well-defined meromorphic function on S but is an example of a different, new object.

Definition. A **meromorphic differential** (or just differential) η on a Riemann surface S is a collection of meromorphic functions η_α on $\varphi_\alpha(V_\alpha) \subset \mathbb{C}$ assigned to the coordinate charts $\varphi_\alpha : V_\alpha \to \mathbb{C}$ on S and satisfying

$$\eta_\alpha(z) = \eta_\beta(\tau_{\alpha\beta}(z)) \cdot \tau'_{\alpha\beta}(z), \quad z \in \varphi_\alpha(V_\alpha \cap V_\beta), \tag{15.42}$$

for any pair $\alpha, \beta \in A$.

We say that $\eta_\alpha(z)dz$, for $z \in \varphi_\alpha(V_\alpha)$, is the **local expression** of η with respect to a chart φ_α.

Remark. In the local expression of a differential we choose not to assign any meaning to the 'dz' part, so it serves as just a *notation.* This notation is at least convenient though: e.g. the (15.42) becomes a rather intuitive $\eta_\beta(w)dw = \eta_\beta(w)\frac{dw}{dz}dz = \eta_\beta(\tau_{\alpha\beta}(z))\tau'_{\alpha\beta}(z)dz$, where we used z and w to denote the local coordinate on, respectively, V_α and V_β.

In view of the above, the derivative of a meromorphic function f on S is a well-defined differential. We shall denote this differential as df; the local expressions of df are $f'_\alpha(z)dz$, $z \in \varphi_\alpha(V_\alpha)$. The (15.41) then may be written as $f'_\alpha(z)dz = f'_\beta(w)dw$, with z and w as before.

More generally, given two meromorphic functions, say f, g, on S, we can construct a differential gdf whose local expressions are $g_\alpha(z)f'_\alpha(z)dz$. It is straightforward to verify that (15.42) holds and thus gdf is well-defined. Notice also that if $f'_\alpha(z_0) \neq 0$, for some $z_0 \in V_\alpha$, then by the Inverse Function Theorem f is biholomorphic near $\varphi_\alpha^{-1}(z_0)$. So the restriction of f to a neighbourhood of $\varphi_\alpha^{-1}(z_0)$ is a well-defined chart. Then f may be understood as a valid *local coordinate* near $\varphi_\alpha^{-1}(z_0)$ and, respectively gdf also makes sense as the local expression of a differential, with respect to this latter coordinate.

Remark. In the case of a complex plane domain $U \subseteq \mathbb{C}$ there is an obvious preferred choice of a complex coordinate z. Then, for any meromorphic function f on U, the differential df has a preferred 'local expression' $f'dz$ or just f'. Then df is identified with a meromorphic function f'. But such an identification is no longer possible on a general Riemann surface S as we no longer have a 'global' complex coordinate defined on all of S.

If $\eta = \{\eta_\alpha\}$ and $\zeta = \{\zeta_\alpha\}$ are two meromorphic differentials on S and $\zeta_\alpha \neq 0$ then (15.42) implies

$$\frac{\eta_\alpha}{\zeta_\alpha} = \frac{\eta_\beta}{\zeta_\beta}$$

on $V_\alpha \cap V_\beta$. This defined a meromorphic extension of a function η_α/ζ_α on V_α to a $V_\alpha \cup V_\beta$ (assuming that a non-empty overlap) and the function is further extended in this way over all of S. Thus we proved

Proposition 15.43. *Given a non-zero meromorphic differential ζ on a connected Riemann surface S, any other meromorphic differential η on S is expressed as $\eta = f\zeta$, for some meromorphic function f.*

Conversely, if ζ is a meromorphic differential and g is a meromorphic function then $g\zeta$ is a meromorphic differential; its local expressions are $g_\alpha \zeta_\alpha dz$. (The gdf above is an example of this.)

Thus the existence of meromorphic differentials on S is related to the existence of meromorphic functions. If S a non-singular algebraic curve in $\mathbb{C}^2$ then an example of meromorphic function on S is provided by the first (or second) projection to $\mathbb{C}$. If $S \subset \mathbb{P}^2$ is a non-singular projective curve then S is a connected Riemann surface (remark on page 25). Assume that S is *not* a projective line $\{\alpha X + \beta Y = 0\}$, for any $(\alpha, \beta) \neq (0, 0)$, and that $0 : 0 : 1$ in not in S. (We may assume this without loss.) Then $X : Y : Z \in S \to X : Y \in \mathbb{P}^1$ is an example of non-constant meromorphic function on S.

In fact, more is true.

Theorem 15.44. *Every compact Riemann surface carries a non-constant meromorphic function.*

The proof requires serious technical work in analysis, dealing with the Laplace equation on surface domains and studying the solutions with prescribed singularities. Theorem 15.44 is a crucial ingredient in the proof of another important result.

Theorem 15.45. *Every compact connected Riemann surface is* algebraic*, i.e. is biholomorphic to an algebraic curve defined by a system of polynomial equations on $\mathbb{P}^n$, for some n.*

This was conjectured by Riemann around 1850 and proved over the next five decades.

Here is some evidence for the significance of Theorem 15.44 for this latter result. Assume, for the moment, that there are 'sufficiently many' meromorphic functions on a given compact Riemann surface, R say. A system of meromorphic functions defines a map of R into a projective space, $x \in R \to 1 : f_1(x) : \ldots : f_n(x) \in \mathbb{P}^n$. Slightly generalizing Example sheet 2 Question 11, one can show that the f_i's are algebraically related and R is mapped to the zero locus of a system of homogeneous polynomials. If the map of R is a bijection onto this zero locus and the image of R is non-singular then we obtain a projective embedding of R. (I.e. R is biholomorphic to a curve in $\mathbb{P}^n$.) You will find an example of projective embedding in $\mathbb{P}^2$ in Example sheet 3 Question 5.

Definition. The point $x \in S$ is a **pole of order** m of a meromorphic differential η if $\varphi_\alpha(x)$ is a pole of order m of some meromorphic function η_α, such that $x \in V_\alpha$ (φ is a chart of S defined on V_α). Respectively, x is a **zero of order** m of η if $\varphi_\alpha(x)$ is a zero of order m of some η_α.

The orders of zeros and poles of η are well-defined independent of the choice of coordinate chart: note (15.42) and recall that the derivative of a biholomorphic function $\tau_{\alpha\beta}$ is never zero.

A meromorphic differential is **holomorphic** if it has no poles.

Remark. A compact Riemann surface has no non-constant holomorphic functions but you will see in Example sheet 3 that there are non-trivial holomorphic differentials e.g. on elliptic curves.

16 Divisors and the Riemann–Roch theorem

Throughout this section, S is a compact connected Riemann surface. We assume without proof that S admits non-constant meromorphic functions.

Definition. A **divisor** on S is a formal sum $D = \sum_{P_i \in S} n_i P_i$, where $n_i \in \mathbb{Z}$ and only finitely many n_i are non-zero.

It is easy to see that the set of all divisors on a given Riemann surface forms a group by addition.

The **degree** of a divisor $D = \sum n_i P_i$ is $\deg D = \sum n_i$. Notice that deg defines a homomorphism from the group of all divisors on S to $\mathbb{Z}$.

If $D = \sum n_j P_j$ is a divisor then we say that D is **effective**, and write $D \geq 0$, if $n_j \geq 0$, for all j.

The divisor (f) of a meromorphic function f which is not identically zero on S is defined as

$$(f) = \sum_i k_i A_i - \sum_j \ell_j B_j,$$

where A_i are all the zeros of f, k_i respectively their orders and B_j are all the poles of f, ℓ_j their orders. We say that D is a **principal divisor** if $D = (f)$ for some meromorphic f.

Two divisors D and D' are **linearly equivalent** if $D - D'$ is a principal divisor.

As the zeros and poles (and their orders) of a meromorphic differential ω are well-defined and are isolated, we can define the divisor (ω) of any $\omega \not\equiv 0$ in just the same way as a divisor of a meromorphic function. The divisor of a (not identically zero) meromorphic differential is called a **canonical divisor**.

The following are easy consequences of the above definitions:

(1) Any principal divisor has degree zero (as the number of zeros and the number of poles of a meromorphic function of a *compact* Riemann surface are the same, equal to the degree of f).

(2) $(fg) = (f) + (g)$ and $(f/g) = (f) - (g)$ for any meromorphic functions f, g, which are not identically zero.

(3) Any two canonical divisors (ω), $(\tilde{\omega})$ are linearly equivalent (as $\tilde{\omega} = f\omega$, for some meromorphic function f). It follows that the degree of a canonical divisor is independent of the choice of a meromorphic differential ω.

For any divisor D we can consider the effective divisors which are linearly equivalent to it. In other words, we can consider the vector space

$$\mathcal{L}(D) = \{\text{meromorphic functions } f \text{ on } S : D + (f) \geq 0\} \cup \{0\}.$$

The condition $D + (f) \geq 0$ means that the poles of f may only occur at those points $P_i \in S$ which appear in D with positive coefficients m_i and may have orders at most m_i. Choosing a chart at each P_i we find that the possible non-zero coefficients of the principle parts of Laurent series for f at all P_i's form a *finite-dimensional* vector space. It follows that $\mathcal{L}(D)$ is always finite-dimensional; denote $\ell(D) = \dim \mathcal{L}(D)$.

The following properties are more-or-less evident

(i) If K_S is a canonical divisor on S then $\mathcal{L}(K_S)$ is the vector space of all the *holomorphic differentials* on S.

(ii) If $\deg D < 0$ then any divisor linearly equivalent to D satisfies $\deg(D + (f)) < 0$ and thus $\ell(D) = 0$

(iii) If $D' = D + (h)$ is a divisor linearly equivalent to D then a meromorphic function f is in $\mathcal{L}(D')$ if and only if $fh \in \mathcal{L}(D)$, so the two vector spaces are isomorphic and $\ell(D) = \ell(D')$.

A way to calculate $\ell(D)$ is given by the following fundamental result.

Riemann–Roch Theorem. *Let S be a compact connected Riemann surface. Then*

$$\ell(D) = 1 - g(S) + \deg D + \ell(K_S - D),$$

where $g(S)$ denotes the genus and K_S is a canonical divisor on S.

In these lectures, we assume Riemann–Roch Theorem without proof.

Here are some immediate consequences.

- A special case when $D = 0$ yields $g(S) = \ell(K_S)$, an alternative characterization of the genus of a Riemann surface as the dimension of the space of holomorphic differentials.

  The number of linearly independent (over $\mathbb{C}$) holomorphic differentials S is called the **geometric genus** of G and denoted $p_g(S)$. Both the Euler characteristic $\chi(S)$ and the geometric genus generalize to higher dimensions and need not be equivalent then — but this is another, more advanced story.

- Putting $D = K_S$ we obtain that $\deg K_S = 2g(S) - 2 = -\chi(S)$.
- If $\deg D > 2g(S) - 2$ then the Riemann–Roch becomes $\ell(D) = 1 - g(S) + \deg D$.

The Riemann–Roch theorem has a number of applications in the theory of algebraic curves and Riemann surfaces. We mention some of these.

Theorem 16.46. *A Riemann surface of genus 0 is biholomorphic to the Riemann sphere.*

Proof. For $D = P$ a point in S, the Riemann–Roch gives $\ell(P) = 2$, so there must be a meromorphic function $S \to \mathbb{P}^1$ of degree 1. Then S must be biholomorphic to $\mathbb{P}^1$ (why?). □

Theorem 16.47. *Every compact Riemann surface of genus 2 is hyperelliptic (i.e. carries a meromorphic function of degree 2).*

Proof (gist). As $\ell(K_S) > 0$ there is a non-zero holomorphic differential, ω say, on S. For any $P, Q \in S$, Riemann–Roch gives $\ell(K_S - P - Q) = \ell(P + Q) - 1$. Now the number of zeros of ω is $\deg(\omega) = 2$. Let $P, Q \in S$ be these zeros ($Q = P$ in the case of a double zero). Then $\ell(K_S - P - Q) = \ell((\omega) - P - Q) = \ell(0) = 1$, so $\ell(P + Q) = 2$ and S must admit a non-constant meromorphic function with at most two (hence exactly two) poles. □

It is easy to see that Riemann surfaces of genus 0 or 1 are also hyperelliptic.

By contrast, a 'generic' compact Riemann surface of genus > 2 is *not* hyperelliptic. An example of non-hyperelliptic Riemann surface if the quartic projective curve $X^4+Y^4 = Z^4$ (this can be proved by studying the subfield of meromorphic functions on S which are obtainable as the quotients of two *holomorphic differentials*).

The group law on a genus-one curve. Suppose now that $g(S) = 1$. Let $Cl^0(S)$ denote the set of linear equivalence classes of the divisors of degree zero on S. It is easy to see that $Cl^0(S)$ is an additive group. The group structure can be induced from $Cl^0(S)$ to S via a bijective map defined as follows.

Fix a point $P_0 \in S$. If $\deg D = 0$ then by the Riemann–Roch $\ell(D + P_0) = 1$ which means that there is a unique *effective* divisor linearly equivalent to $D + P_0$. But effective divisor of degree 1 is a point, P say. Thus $D \in Cl^0(S) \to P \in S$ defines the required bijection. E.g. the zero of $Cl^0(S)$ is mapped to P_0.

If $S = \mathbb{C}/\Lambda$ and $P_0 = 0 + \Lambda$, for a lattice Λ, then the above construction recovers the addition of complex numbers modulo Λ. This is a consequence of the constraints on zeros and poles of the elliptic functions (Theorem 4.18).

The next theorem is part of Question 10 in the Example sheet 4.

Theorem 16.48. *If S is a non-singular projective curve of genus 1 then S is biholomorphic to a cubic curve in the generalized Weierstrass normal form $C = \{X : Y : Z \in \mathbb{P}^2 \mid Y^2Z - a_2XYZ - a_4YZ^2 = X^3 + a_1X^2Z + a_3XZ^2 + a_5Z^3\}$, for some $a_i \in \mathbb{C}$.*

Some references for Chapter IV

all to F. Kirwan 'Complex algebraic curves'

Affine and projective curves, projective spaces Ch. 2.
Implicit function theorem Appendix B.
Meromorphic differentials §6.1.
Divisors and the Riemann-Roch Theorem §6.3.

There is also a very good book R. Miranda 'Algebraic curves and Riemann surfaces' (existence of compact Riemann surface of any genus is based on Lemma 1.7 in Ch. III there). However, it is more advanced than Kirwan's book and has quite a few additional topics.

V. Analytic continuation and the covering surfaces

17 The complete analytic functions

Definition. (i) A **function element** is a pair (f, D) where f is a holomorphic function defined on an open connected set D.

(ii) A function element (g, E) is called a **direct analytic continuation** of (f, D), denoted $(g, E) \approx (f, D)$, if $D \cap E \neq \emptyset$ and $f|_{D\cap E} = g|_{D\cap E}$.

(iii) A function element (g, E) is called an **analytic continuation** of (f, D), denoted $(g, E) \sim (f, D)$, if there is a finite set of function elements (f_i, D_i), so that

$$(f, D) = (f_0, D_0) \approx (f_1, D_1) \approx \ldots \approx (f_N, D_N) = (g, E).$$

It is easy to check that $\sim$ is an equivalence relation (but $\approx$ is not).

Definition. A **complete analytic function** (in the sense of Weierstrass) is the $\sim$ equivalence class of a function element.

It may happen that a holomorphic function f on $D \neq \mathbb{C}$ does not have any direct analytic continuation at all to any open set not contained in D [4]. More precisely, a point $c \in \partial D$ in the boundary of D is called **regular** if there exists a function element $(g, \{|z - c| < \varepsilon\}) \approx (f, D)$, for some $\varepsilon > 0$. c is called **singular** otherwise. If the boundary of D contains no regular points then ∂D is called the **natural boundary** of the function element (f, D).

If a function element does admit non-trivial analytic continuations then it is quite possible to 'return to the same point with a different value', i.e. have $(f, D) \sim (g, E)$ with $D = E$ but $f \neq g$.

The remarkable idea of Riemann was that the so-called 'multivalued functions' are just considered on a wrong domain: the natural domain is a *surface* covering $U \subset \mathbb{C}$ several (possibly infinitely many) times.

Definition. Let $p : X \to Y$ be a continuous map between topological surfaces. We shall say that a continuous map $q : V \to X$ is a **local section** of p over $V \subset Y$ if $p \circ q = \mathrm{id}_V$.

It follows that p maps $q(V)$ homeomorphically onto V. Informally, the image $q(V)$ is a 'layer' over V in the covering surface X. Note that X need not be homeomorphic to the product of Y and some discrete space and so it may not be possible, in general, to find a section of p defined on all of Y.

Theorem 17.49. *Any complete analytic function $\mathcal{F}$ determines, in a natural way, a Riemann surface $S(\mathcal{F})$ endowed with two holomorphic maps $\pi : S(\mathcal{F}) \to \mathbb{C}$ and $u : S(\mathcal{F}) \to \mathbb{C}$, such that if (f, D) is a function element of $\mathcal{F}$ then $f(z) = u \circ q(z)$, for some holomorphic section $q : D \to S(\mathcal{F})$ of π.*

Remark. As $\pi(S(\mathcal{F}))$ is open, thus non-compact, $S(\mathcal{F})$ is never compact (though $S(\mathcal{F})$ may admit a 'natural' compactification in some examples).

Theorem 17.49 makes precise the intuitive idea of choosing over $D \subset \mathbb{C}$ a single-valued 'holomorphic branch' of a multivalued function.

The basic idea of the proof is to construct $S(\mathcal{F})$ by 'gluing' together the function elements in $\mathcal{F}$ according to the direct analytic continuations. To this end, one uses the following refinement of the equivalence relation $\sim$.

[4] even if D is bounded, e.g. a unit disc, and even if f extends continuously to the boundary ∂D

Definition. A germ $[f, z]$ of a function element (f, D) is the equivalence class of function elements $(g, E) \underset{z}{\sim} (f, D)$, where the notation $\underset{z}{\sim}$ means that $z \in D \cap E$ and $f = g$ on some open neighbourhood of z.

Proof of Theorem 17.49 (gist). The $S(\mathcal{F})$ is defined as the set of all germs $[f, z]$, $z \in D$, of all the functional elements (f, D) in $\mathcal{F}$. Then the maps $\pi([f, z]) := z$ and $u([f, z]) := f(z)$ are well-defined.

A Hausdorff topology on $S(\mathcal{F})$ is induced by a system of 'basic neighbourhoods' $[f, D] = \cup_{z\in D}[f, z]$, for all $(f, D) \in \mathcal{F}$. More precisely, $U \subset S(\mathcal{F})$ is open if for each point $[f, z] \in U$ there is a function element $(f, D) \in \mathcal{F}$, so that $[f, z] \in [f, D] \subset U$. It readily follows that the union of any collection of open subset of $S(\mathcal{F})$ is open, in particular $\emptyset$ and $S(\mathcal{F})$ are open. To deal with the intersections, it suffices to check the intersection of two basic neighbourhoods. Note that if $[f_1, D_1] \cap [f_2, D_2]$ is non-empty then $(f_1, D_1) \approx (f_2, D_2)$, so $(f_1, D_1 \cap D_2)$ is a valid function element of $\mathcal{F}$ and we find that $[f_1, D_1] \cap [f_2, D_2] = [f_1, D_1 \cap D_2]$ is open.

A slightly subtle point in checking the Hausdorff condition is that it is possible to have $[f_1, z_1] \neq [f_2, z_2]$ with $z_1 = z_2$ (and even $f_1(z_1) = f_2(z_2)$). But then $f_1 - f_2$ is never-zero on some *punctured disc* $D^*(z_1, \varepsilon)$ and so $[f_1, D^*(z_1, \varepsilon)] \cap [f_2, D^*(z_1, \varepsilon)] = \emptyset$.

The charts on $S(\mathcal{F})$ are the 'obvious' maps $[f, z] \in [f, D] \to z \in D \subset \mathbb{C}$ corresponding to the function elements (f, D) of $\mathcal{F}$. We assume that each D is an *open disc* in $\mathbb{C}$. With these arrangements, two coordinate neighbourhoods $[f, D]$ and $[g, E]$ overlap if and only if the function elements (f, D) and (g, E) are direct analytic continuations of each other. The transition function between local coordinates (on $[f, D]$ and $[g, E]$) is then just $\mathrm{id}_{D\cap E}$.

Finally, it is straightforward to verify that π and u defined above are holomorphic (local expressions for u will be given by $f(z)$ for the function elements (f, D) of $\mathcal{F}$). □

The Riemann surfaces of some multivalued functions

Here are two simple examples for Theorem 17.49.

1. The power series $z - z^2/2 + z^3/3 - \ldots + (-1)^{n+1}z^n/n \ldots$ defines a holomorphic function $h(z)$, for $|z - 1| < 1$, such that $\exp(h(z)) = 1 + z$. The function element $(h, \{|z - 1| < 1\})$ generates a complete analytic function which we shall denote by Log. The function elements of Log are local holomorphic inverses $\tilde{h}(z)$ of $\exp(z)$ and any such $\tilde{h}(z)$, on a neighbourhood of z_0 say, is uniquely determined by the value $w_0 = \tilde{h}(z_0)$. It follows that we may identify the germs in $S(\mathrm{Log})$ as pairs $\{(z, w) : z - \exp(w) = 0\}$ and the two maps defined in Theorem 17.49 are given by $\pi(z, w) = z$ and $u(z, w) = w$. We find that, in this example, the map u is *invertible* (and biholomorphic), with $u^{-1}(w) = (\exp(w), w)$ and $\pi(u^{-1}(w)) = \exp(w)$. Thus the Riemann surface of the logarithm is identified with the graph of the function exp and is biholomorphic to $\mathbb{C}$. Note that the identification with $\mathbb{C}$ goes via u, whereas via the map π the Riemann surface $S(\mathrm{Log})$ 'spirals' over $\mathbb{C}$ infinitely many times.

2. The situation with $\sqrt[n]{z}$ is quite similar to the above example. The corresponding complete analytic function contains local holomorphic inverses of w^n (i.e. holomorphic $h_n(z)$ satisfying $(h_n(z))^n = z$). Such holomorphic functions $h_n(z)$ may be defined on any open disc $D \subset (\mathbb{C} \setminus \{0\})$. Note that although $\sqrt[n]{0}$ is defined there is no function element of $\sqrt[n]{z}$ whose domain is an open neighbourhood of 0 (otherwise the real-valued function $\sqrt[n]{t}$ would be differentiable at $t = 0$, but it is not!).

The Riemann surface of $\sqrt[n]{z}$ may be identified with an open subset $C_0 = \{(z, w) \in C : z \neq 0\}$ of an *algebraic curve* $C = \{(z, w) \in C : z - w^n = 0\}$ in $\mathbb{C}^2$. In this example, we again have $\pi(z,w) = z$ and $u(z,w) = w$. Notice that the charts of $S(\sqrt[n]{z})$ naturally correspond to the charts of C_0 given by the *first projection.* Note also that the first projection does *not* define a chart near $(0,0) \in C$.

Once again u is invertible with $u^{-1}(w) = (w^n, w)$ and $\pi(u^{-1}(w)) = w^n$. The Riemann surface $S(\sqrt[n]{z}) = C_0$ is identified with the graph of the function $w \mapsto z = w^n$, $w \neq 0$, and so C_0 is biholomorphic to the punctured complex plane $\mathbb{C}^*$.

Remark. It is tempting to assume from the last example that the Riemann surface of a function $w = w(z)$ defined implicitly by the vanishing of a polynomial $P(z,w)$ in two variables is biholomorphic to an open subset of an algebraic curve $\{P(z,w) = 0\}$ obtained by deleting the ramifications points of the first projection $(z,w) \in \{P(z,w) = 0\} \to z \in \mathbb{C}$. However, the situation is not quite so simple as in general the 'obvious' algebraic curve may have singular points...

18 A topological digression: the covering surfaces

To better understand the Riemann surface $S(\mathcal{F})$ of a complete analytic function, we shall need some facts from topology.

Definition. Let R be a topological surface. A **'covering' surface** for R (in the sense of complex analysis) is a topological surface $\tilde{R}$ endowed with a continuous surjective map $p : \tilde{R} \to R$, such that any $x \in \tilde{R}$ has a neighbourhood $\tilde{U}$ so that the restriction $p|_{\tilde{U}}$ is a homeomorphism onto its image in R.

The Riemann surface $S(\mathcal{F})$ of a complete analytic function $\mathcal{F}$, together with the map π, is the 'covering' surface for the image $\pi(S(\mathcal{F})) \subseteq \mathbb{C}$. This follows directly from the way we constructed the charts on $S(\mathcal{F})$ in the proof of Theorem 17.49. (Note also that $\pi(S(\mathcal{F}))$ is an open subset in $\mathbb{C}$ as π is holomorphic.)

Remark. In topology, it is customary to use a stronger definition of a covering $p : \tilde{X} \to X$ when each $x \in X$ must have an open neighbourhood V with $p^{-1}(V)$ a disjoint union of open sets in $\tilde{X}$, so that p maps each of the latter open sets homeomorphically onto V. Also $\tilde{X}$ is often required to be path-connected. In particular, this is the definition used in IID Algebraic Topology. However, there are examples of Riemann surfaces of complete analytic functions, obtainable by Theorem 17.49, which do *not* fit the topological definition. (See example 18.52 below.) This is why a weaker definition had to be adopted here.

We shall soon identify a sufficient condition when the two definitions of a covering (one in the sense of complex analysis and one in the sense of algebraic topology) will agree.

Proposition 18.50. *If R is a Riemann surface and $p : \tilde{R} \to R$ is a 'covering' surface for R then $\tilde{R}$ has a structure of a Riemann surface (i.e. a complex structure), such that p is a holomorphic map.*

Proof (gist). The charts of the required complex structure of $\tilde{R}$ are obtained as the compositions $\varphi \circ p$ (on small neighbourhoods), for every chart φ of the complex structure of R. □

Definition. Let $p : \tilde{R} \to R$ be a 'covering' of a surface R, and $\gamma(t) : [0,1] \to R$ a continuous path, and w_0 a point in $\tilde{R}$ with $p(w_0) = \gamma(0)$. Then **a lift of $\gamma(t)$ from w_0** is a continuous path $\Gamma(t) : [0,1] \to \tilde{R}$ such that $\Gamma(0) = w_0$ and $p(\Gamma(t)) = \gamma(t)$ for all $0 \leq t \leq 1$.

In the context of a 'covering' $\pi : S(\mathcal{F}) \to \pi(S(\mathcal{F})) \subseteq \mathbb{C}$, as in Theorem 17.49, examples of the lifts are obtained as follows. Let (f, D) be a function element in $\mathcal{F}$. If γ is a path in D, $z = \gamma(0)$, then the path of germs $\Gamma(t) = [f, \gamma(t)]$ (determined by (f, D)) defines a lift of γ from the germ $w_0 = [f, z] \in S(\mathcal{F})$.

If a lift of $\gamma(t)$ from w_0 exists then it is unique. However, there is in general no guarantee that a lift (of a given path from a given point) exists.

Definition. A 'covering' $p : \tilde{R} \to R$ is called **regular** if any continuous path $\gamma(t) : [0, 1] \to R$ has a lift to $\tilde{R}$ from any $w_0 \in \tilde{R}$ such that $p(w_0) = \gamma(0)$.

It turns out that a regular 'covering' is a covering in the (stronger) sense of algebraic topology; see the remark after Corollary 19.55.

The following topological result, which we assume here without proof, gives a useful sufficient condition for a 'covering' to be regular.

Theorem 18.51. *A 'covering' $\pi : \tilde{R} \to R$ will be regular if each $z \in R$ has a neighbourhood K such that every connected component of $\pi^{-1}(\overline{K})$ is compact, where $\overline{K}$ denotes the closure of K.*

Example 18.52. Consider the Riemann surface $S(\mathcal{F})$ of the complete analytic function $\mathcal{F}(z) = \sqrt{1 + \sqrt{z}}$. The holomorphic map $\pi : S(\mathcal{F}) \to \mathbb{C}$ is an example of non-regular 'covering': the path $\gamma(t) = 1 - \varepsilon/2 + \varepsilon t,\ \ 0 \le t \le 1$, with $0 < \varepsilon < 1/2$ does not have a lift from the germ $[g(1 - h(z)), \gamma(0)]$, where g and h are holomorphic, $g(z)^2 = h(z)^2 = z$ and $h(1) = 1$. (Example sheet 4 Question 4.)

Examples of regular coverings are

$$z \in \mathbb{C} \to z + \Lambda \in \mathbb{C}/\Lambda, \qquad \text{for a lattice } \Lambda \text{ in } \mathbb{C},$$
$$z \in \mathbb{C}^* \to z^n \in \mathbb{C}^*, \qquad \text{for a positive integer } n,$$

Notice that in the latter example $0 \in \mathbb{C}$ must be excluded.

Definition. Let $\gamma, \sigma : [a, b] \to Y$ be two continuous paths in a topological space Y, with common end points $\gamma(a) = \sigma(a)$, $\gamma(b) = \sigma(b)$. A **homotopy between paths** γ **and** σ is a continuous map $H(s, t) : [0, 1] \times [a, b] \to Y$ such that

$$H(s, a) = \gamma(a) = \sigma(a), \quad H(s, b) = \gamma(b) = \sigma(b), \qquad \text{for all } 0 \le s \le 1,$$
$$H(0, t) = \gamma(t), \qquad H(1, t) = \sigma(t), \qquad \text{for all } a \le t \le b,$$

Remark. If we set $X = [a, b]$ and $A = \{a, b\} \subset X$ then the above definition becomes a special case of the general definition of homotopy between (continuous) maps $X \to Y$ relative to A, used in Algebraic Topology.

Definition. A topological space Y if called **simply-connected** if it is path-connected and any closed path in Y is homotopic to a point.

For example, an open disc in $\mathbb{C}$ and the Riemann sphere are simply connected, whereas an annulus $\{1 < |z| < 2\}$, an elliptic curve (torus) and, more generally, compact Riemann surfaces of genus $g \ge 1$ are not simply-connected.

N.B. The terms 'connected' and 'path-connected' are equivalent for topological surfaces (but not for general topological spaces).

We shall need the following result proved in topology.

Theorem 18.53. *Every path-connected topological surface R has a regular covering by a simply-connected topological surface, $\tilde{R}$ say. The surface $\tilde{R}$ then is called a* **universal regular cover** *of R.*

19 The Monodromy Theorem and its applications

The Monodromy Theorem. *If $p : \tilde{R} \to R$ is a regular covering then any homotopy between paths in R can be lifted to a homotopy in $\tilde{R}$.*

More precisely, if $H(s,t) : [0,1] \times [a,b] \to R$ is a homotopy between paths $\gamma(t)$ and $\sigma(t)$ in R and if $w_0 \in \tilde{R}$ is such that $p(w_0) = \gamma(a) = \sigma(a)$, then the lifts $\tilde{\gamma}(t)$ and $\tilde{\sigma}(t)$, of γ and σ, from w_0 are homotopic in $\tilde{R}$. In particular, $\tilde{\gamma}(b) = \tilde{\sigma}(b)$.

We assume the Monodromy Theorem without proof. Here are two applications.

Recall that any function element (f, D) determines a complete analytic function $\mathcal{F}$ whose Riemann surface $S(\mathcal{F})$ covers an open subset of $R \subseteq \mathbb{C}$, denote the covering map $\pi : S(\mathcal{F}) \to R$. An **analytic continuation of the germ of** (f, D) **along a path** $\gamma : [a, b]$, $\gamma(a) = z \in D$ is the lift of γ from the germ $[f, z]$.

Theorem 19.54. *Suppose that $R \subset \mathbb{C}$ is an open connected subset and a germ $[f, z]$ can be analytically continued along any path contained in R. Then the analytic continuations along any two homotopic paths $\gamma, \sigma : [a, b] \to R$ starting at z give the same germ at $w = \gamma(b) = \sigma(b)$.*

The hypothesis of Theorem 19.54 ensures that $\pi : S(\mathcal{F}) \to R$ is a regular cover, so the Monodromy Theorem applies.

Corollary 19.55. *Assume the hypotheses of Theorem 19.54. Let $U \subseteq R$ be a* simply-connected *domain and (f, D) a function element in $\mathcal{F}$.*

Then the analytic continuations of (f, D) along paths in U give a single-valued *holomorphic function $f : U \to \mathbb{C}$.*

A basic example is given by 'holomorphic branches' of $\log z$ and $\sqrt[n]{z}$; for any simply-connected $U \subset \mathbb{C} \setminus \{0\}$, there are well-defined (single-valued) holomorphic functions $h(z), h_n(z) : U \to \mathbb{C}$ so that $\exp(h(z)) = z$ and $\big(h_n(z)\big)^n = z$. E.g. a suitable choice of U is a cut plane $\mathbb{C} \setminus \{z \in \mathbb{R} : z \leq 0\}$.

Remark. One consequence of Corollary 19.55 is that (f, U) is a well-defined function element of $\mathcal{F}$ and the germs of (f, U) define an open set $\tilde{U} \subseteq S(\mathcal{F})$ biholomorphic to U.

It follows that each $z \in R = \pi(S(\mathcal{F}))$ has an open neighbourhood V, such that $p^{-1}(V)$ a disjoint union of open sets in $\tilde{R}$, so that p maps each of the latter open sets homeomorphically (moreover, biholomorphically) onto V. This also verifies, as promised in the previous section, that a regular cover is the same as a cover in the (stronger) sense of algebraic topology.

The second application of the Monodromy Theorem is to the elliptic curves.

Theorem 19.56. *Two elliptic curves $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$ are conformally equivalent if and only if $\Lambda_1 = \alpha\Lambda_2$, for some $\alpha \neq 0$.*

Proof (gist). The non-trivial part is to show that any biholomorphic map f of $\mathbb{C}/\Lambda_1$ onto $\mathbb{C}/\Lambda_2$ induces a biholomorphic automorphism F of $\mathbb{C}$. To construct F, we choose $w_0 \in \mathbb{C}$ such that $f(\pi_1(0)) = \pi_2(w_0)$. (Here $\pi_i : \mathbb{C} \to \mathbb{C}/\Lambda_i$ are the quotient maps.)

Given a $z \in \mathbb{C}$, connect 0 to z by a continuous path $\gamma(t)$, $0 \leq t \leq 1$. Let $\gamma_2(t)$ be the unique lift of $f(\pi_1(\gamma(t)))$ from w_0. The Monodromy Theorem ensures that the end-point $\gamma_2(1)$ is independent of the choice of the path γ to z and so we obtain a well defined function $F : z \in \mathbb{C} \to F(z) := \gamma_2(1) \in \mathbb{C}$. This function is bijective (swap the roles of $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$ to get F^{-1}). It is also holomorphic as on a neighbourhood of z we can write $F = \varphi \circ f \circ \pi_1$, where φ is a chart of $\mathbb{C}/\Lambda_2$ near $f(\pi_1(z))$ given by a local inverse of π_2. A similar argument, with $\mathbb{C}/\Lambda_1$ and $\mathbb{C}/\Lambda_2$ swapped, shows that F^{-1} is holomorphic.

Since F is a biholomorphic map of $\mathbb{C}$ onto itself it must have the form $F(z) = az + b$, $a \neq 0$ (Question 6(ii) from Example sheet 1). Then $b = w_0$. If $\omega \in \Lambda_1$ then $\pi_2(F(\omega)) = f(\pi_1(\omega)) = f(\pi_1(0)) = \pi_2(w_0)$ whence $F(\omega) - w_0 = \omega' \in \Lambda_2$. That is $\omega' = F(\omega) - F(0) = a\omega$ and $a\Lambda_1 = \Lambda_2$, hence $\Lambda_1 = \alpha\Lambda_2$ with $\alpha = a^{-1}$. □

Corollary 19.57. *There are uncountably many elliptic curves not conformally equivalent to each other.*

Every elliptic curve is biholomorphic to one defined by a lattice $\mathbb{Z} + \tau\mathbb{Z}$, for some $\tau > 0$. However, two different choices of τ may well define biholomorphic elliptic curves.

The problem of identifying a 'parameter space' for the conformal equivalence classes of elliptic curves is equivalent to describing the set $\mathcal{M}$ of lattices modulo rescaling. For the latter issue, it is useful to consider the group

$$PSL(2,\mathbb{Z}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : a, b, c, d \in \mathbb{Z},\ ad - bc = 1 \right\} / \pm 1.$$

$PSL(2,\mathbb{Z})$ acts on the upper half-plane $\mathbb{H} = \{\operatorname{Im}\tau > 0\}$ by the Möbius transformations,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \tau = \frac{a\tau + b}{c\tau + d}.$$

Proposition 19.58. $\mathcal{M}$ *is bijective to the orbits of the action of* $PSL(2,\mathbb{Z})$ *on* $\mathbb{H}$.

The main point of the proof is that if two lattices $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$ and $\tilde{\Lambda} = \mathbb{Z}\tilde{\omega}_1 + \mathbb{Z}\tilde{\omega}_2$ are related by rescaling then ω_1/ω_2 and $\tilde{\omega}_1/\tilde{\omega}_2$ are related by a Möbius transformation in $PSL(2,\mathbb{Z})$.

Given a group G acting on a space X, a **fundamental region** D of this action is defined as a subset D of X such that every orbit of G intersects D in precisely one point.

Theorem 19.59. *A fundamental region of* $PSL(2,\mathbb{Z})$ *in* $\mathbb{H}$ *may be given by* $D = \{\tau \in \mathbb{H} \mid -\frac{1}{2} < \operatorname{Re}\tau \leq \frac{1}{2},\ |\tau| \geq 1,$ *and if* $\operatorname{Re}\tau < 0$ *then* $|\tau| > 1\}$.

One can show that any $\tau \in \mathbb{H}$ can be mapped into D by alternating (finitely many times) the following two $PSL(2,\mathbb{Z})$ transformations: $\left(\begin{smallmatrix} 1 & m \\ 0 & 1 \end{smallmatrix}\right)$, mapping each τ to $\tau + m$ $(m \in \mathbb{Z})$, and $\left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right)$, mapping each τ to $-1/\tau$. (These transformations actually generate the group $PSL(2,\mathbb{Z})$.) Thus any $PSL(2,\mathbb{Z})$-orbit contains a point in D.

Further, it turns out that any $PSL(2,\mathbb{Z})$-orbit meets D in only one point; a (rather lengthy) proof can be found in McKean and Moll 'Elliptic durves', Chapter 4.3.

The space $\mathcal{M} \cong (\mathbb{H}/PSL(2,\mathbb{Z}))$ of the conformal equivalence classes of the elliptic curves is homeomorphic to $\mathbb{C}$ via the function $J(\tau) : \mathbb{H} \to \mathbb{C}$ defined in Example sheet 4 Question 7.

Some page references for Chapter V

to A.F. Beardon, "A primer on Riemann surfaces", CUP 1984.

analytic continuation pp. 95–100;
covering surfaces and regular coverings pp. 101–104;
applications of the Monodromy theorem pp. 100–101,105,108–109.

VI. The Uniformization Theorem

We introduced Riemann surfaces as the most general domain for the holomorphic functions. We saw that Riemann surfaces arise naturally in many ways and investigated a number of examples illustrating this (e.g. algebraic curves, analytic continuation). Is there a way to describe *all* of the Riemann surfaces thus giving a complete classification? As we shall now see, the answer is yes.

Recall that we successfully studied elliptic functions on the genus 1 Riemann surfaces by presenting the surface as a quotient $\mathbb{C}/\Lambda$. One is then led to ask if we could study other surfaces by viewing them as quotients of $\mathbb{C}$ and lifting the meromorphic functions to $\mathbb{C}$. It turns out that by taking a quotient of $\mathbb{C}$ we can obtain very few examples, but more generally the idea of a *quotient* is still a winner and leads to classification of all the Riemann surfaces.

We have seen that any Riemann surface S has universal (simply-connected) regular cover $\tilde{S}$ (Theorem 18.53) and this cover is itself a Riemann surface, so that the covering map is holomorphic (Proposition 18.50). If we can classify all the *simply-connected* Riemann surfaces and show that S as above is obtainable from its universal cover by a quotient construction $S = \tilde{S}/G$ similar to that of elliptic curves $\mathbb{C}/\Lambda$ then the classification of Riemann surfaces will be complete. This is precisely what is achieved by the Uniformization Theorem.

Definition 19.60. A group G of homeomorphisms of a topological space X acts **properly discontinuously** when every $x \in X$ has a neighbourhood U, so that the images $g(U)$, for $g \in G$, are pair-wise disjoint.

The Uniformization Theorem.

1. (difficult part) *Every simply-connected Riemann surface is conformally equivalent to precisely one of:*
(a) the Riemann sphere S^2;
(b) the complex plane $\mathbb{C}$;
(c) the unit disc Δ.

2. (easy part) *Every connected Riemann surface R is conformally equivalent to a quotient of a simply-connected Riemann surface by a subgroup of biholomorphic automorphisms acting properly discontinuously.*

We also say that the latter simply-connected surface **uniformizes** R.

The Uniformization Theorem reduces the classification of Riemann surfaces to the classification of subgroups G of the groups of Möbius transformations,

$$\operatorname{Aut} S^2 \cong SL(2,C)/\pm 1, \qquad \operatorname{Aut}\mathbb{C} = \{f(z) = az + b, a \neq 0\}, \qquad \operatorname{Aut}\Delta \cong SU(1,1)/\pm 1,$$

such that G acts properly discontinuously on the respective Riemann surface.

There are no non-trivial subgroups of $\operatorname{Aut} S^2$ acting properly discontinuously and the only non-trivial such subgroups of $\operatorname{Aut}\mathbb{C}$ are the groups of *translations* $\mathbb{Z}\omega$ ($\omega \in \mathbb{C}^*$) and $\mathbb{Z}\lambda + \mathbb{Z}\mu$ ($\lambda/\mu \notin \mathbb{R}$). This accounts for $S^2 \cong \mathbb{P}^1$, $\mathbb{C}$, $\mathbb{C}^*$ [5] and the elliptic curves $\mathbb{C}/\Lambda$. All the other Riemann surfaces are quotients Δ/Γ of the unit disc.

Remark. Recall from IB Geometry that the unit disc Δ is a model for the *hyperbolic plane* and the Möbius transformations in $SU(1,1)/\pm 1$ are isometries of the hyprebolic metric.

[5] The Riemann surface $\mathbb{C}^*$ is biholomorphic to the cylinder (an exercise).

Here is one immediate application of the Uniformization Theorem.

Corollary 19.61. *Any compact simply-connected Riemann surface is biholomorphic to the Riemann sphere S^2.*

The proof of the first part of the Uniformization Theorem includes, as a special case, the result whose applications were important in IB Complex Methods.

Riemann Mapping Theorem. *Any simply-connected region $D \subset \mathbb{C}$ which is not $\mathbb{C}$ itself is conformally equivalent to the open unit disc Δ.*

In general, the proof of first part includes some hard work in analysis constructing on a given simply-connected surface a holomorphic (or meromorphic) function with special properties. We shall assume the first part of the Uniformization Theorem without proof.

The following extension of the Riemann Mapping Theorem is not part of the Uniformization Theorem but is significant in its own right.

Theorem 19.62 (Boundary behaviour)**.** *Suppose D is a simply-connected open subset of $\mathbb{C}$ and the boundary ∂D contains an open interval of $I = \{a < x < b\} \subset \mathbb{R} \subset \mathbb{C}$. Suppose also that for any $z \in I$, there is an open disc $D(z, \varepsilon)$ such that*

(i) $\partial D \cap D(z, \varepsilon) = I \cap D(z, \varepsilon)$ ('the rest of the boundary stays away from I'); and

(ii) only one of the two half-discs of $D(z, \varepsilon)$ with the diameter $I \cap D(z, \varepsilon)$ is contained in D.

Then a conformal equivalence f from D onto Δ extends to a homeomorphism of $D \cup I$ onto the union of Δ and an arc of the unit circle.

Theorem 19.62 can be generalized to the situation when I is an arc of a circle.

Review of the proof of the second part of the Uniformization Theorem. (non-examinable)

Let S be a connected Riemann surface and $p : \tilde{S} \to S$ a universal regular cover of S, so $\tilde{S}$ is simply-connected. By Proposition 18.50, $\tilde{S}$ then is a Riemann surface and p is a holomorphic map.

We shall say that a homeomorphism g of $\tilde{S}$ onto itself is a **cover transformation** if $p(g(\tilde{z})) = p(\tilde{z})$, for every $\tilde{z} \in \tilde{S}$.

The proof of part 2 of the Uniformization Theorem hinges on four propositions and turns out to be essentially topological; it exploits the theory of covering maps.

Proposition 19.63. *Any cover transformation g of $\tilde{S}$ is biholomorphic, $g \in \operatorname{Aut} \tilde{S}$.*

Proof (gist). Let $z \in \tilde{S}$ and g a cover transformation. If φ is a complex coordinate chart at $p(z) \in S$ then $\varphi \circ p|_{V_1}$ and $\varphi \circ p|_{V_2}$ are charts at respectively z and $g(z)$ in $\tilde{S}$, for suitable choices of neighbourhoods V_1 of z and V_2 of $g(z)$. The local expression of g is complex coordinates is then $(\varphi \circ p|_{V_2}) \circ g \circ (\varphi \circ p|_{V_1})^{-1}(z) = \varphi \circ (p|_{V_2} \circ g) \circ (\varphi \circ p|_{V_1})^{-1}(z) = \varphi \circ p|_{V_2} \circ (p|_{V_1})^{-1} \circ \varphi^{-1}(z) = z$, a holomorphic function (on $(\varphi \circ p)(V_1 \cap V_2)$). □

Proposition 19.64. *The set of all cover transformations form a subgroup $G \subset \operatorname{Aut} \tilde{S}$ and G acts properly discontinuously on $\tilde{S}$.*

Proof (gist). To check the properly discontinuous property, suppose, for contradiction, that $z \in \tilde{S}$ and for any neighbourhood V of z we have $V \cap g(V) \neq \emptyset$, for some cover transformation g. Choose V so that $p|_V$ is a homeomorphism onto image. If there is a $z_0 \in V$ with $g(z_0) \in V$ then we must have $g(z_0) = z_0$. But then, joining z_0 to an arbitrary $w \in \tilde{S}$ by a continuous path $\gamma(t)$ say, we obtain that $g(w) = w$ by the uniqueness of lift of the path $p(\gamma(t))$ from S with the starting point $z \in \tilde{S}$. So g is the identity map of $\tilde{S}$. □

Proposition 19.65. *For any two points $w_1, w_2 \in \tilde{S}$, such that $p(w_1) = p(w_2)$, there is a unique cover transformation g of $\tilde{S}$, such that $g(w_1) = w_2$.*

Proof (gist). Consider an arbitrary $w \in \tilde{S}$. Join w_1 to w by a continuous path $\tilde{\gamma}(t)$. Project $\tilde{\gamma}$ to a path $\gamma = p \circ \tilde{\gamma}$ in S. This γ starts at $p(w_1)$; lift γ to $\tilde{S}$ from w_2, denote the lift $\tilde{\gamma}_2$. Define $g(w)$ to be the other end-point of $\tilde{\gamma}_2$. By construction, $g \circ p = p$ and g is a homeomorphism (g^{-1} is obtained by swapping w_1 and w_2). Also g is uniquely determined because all the choices of $\tilde{\gamma}$ are homotopic (as $\tilde{S}$ is simply-connected), so by the Monodromy Theorem $g(z)$ is independent of $\tilde{\gamma}$. □

Thus there is a bijection $q : z \in S \to p^{-1}(z) \in \tilde{S}/G$, where we denoted by $\tilde{S}/G$ the space of *orbits* of G. Let $\pi : \tilde{S} \to \tilde{S}/G$ denote the quotient map (assigning to each x the orbit Gx). By definition, a set $U \subseteq \tilde{S}/G$ is open when $\pi^{-1}(U)$ is open. (Compare with elliptic curves.) With these arrangements, q becomes a homeomorphism (and, in particular, we see that $\tilde{S}/G$ is a Hausdorff space). Finally, and again by analogy with elliptic curves, $\tilde{S}/G$ has a complex structure such that the quotient map is holomorphic. To obtain a chart near $Gx \in \tilde{S}/G$, take a neighbourhood U of $x \in \tilde{S}$ so that the sets $g(U)$, for $g \in G$, are mutually disjoint (possible by Definition 19.60). Then π maps U homeomorphically onto a neighbourhood of Gx and a chart of $\tilde{S}/G$near Gx is the composision of $(\pi|_U)^{-1}$ with some chart of $\tilde{S}$ near x. The properly discontinuous property is crucial for checking that $\tilde{S}/G$ is a well-defined Riemann surface.

Proposition 19.66. *The map $q : z \in S \to p^{-1}(z) \in \tilde{S}/G$ is a conformal equivalence between S and the quotient Riemann surface $\tilde{S}/G$.*

This is a consequence of the construction of the complex structure on $\tilde{S}/G$ and the property that $p : \tilde{S} \to S$ is a holomorphic regular covering.

You can find a complete account of the proof of the Uniformization Theorem in A.F. Beardon, "A primer on Riemann surfaces", Ch.9.