9 The Riemann Mapping Theorem

References for this section:

- J. Shurman, The Riemann Mapping Theorem, available at
  http://people.reed.edu/~jerry/311/rmt.pdf
- R.B. Ash and W.P Novinger, Complex Analysis, (Dover).

The Riemann Mapping Theorem is one of the highlights of complex analysis, and is a really surprising result. It says that all simply connected proper open subsets of the plane are conformally equivalent.

**Theorem 51. (Riemann Mapping Theorem)** Let \( \Omega \) be a simply connected proper open subset of \( \mathbb{C} \). Then there is an analytic bijection \( f : \Omega \to \mathbb{D} \). Furthermore given any \( z_0 \in \Omega \), there is a unique such map satisfying \( f(z_0) = 0 \), and \( f'(z_0) \) is positive.

It is worthwhile thinking a little about the hypotheses here. If \( \Omega \) is a proper subset of \( \mathbb{C} \), then there is a point \( w \in \mathbb{C} \setminus \Omega \). If \( \Omega \) is also simply connected, this means that \( \mathbb{C} \setminus \Omega \) is connected, so there is a path connecting \( w \) to \( \infty \).

Let us start by looking at the question of uniqueness. Suppose that \( \phi : \mathbb{D} \to \mathbb{D} \) is an analytic bijection and that that \( f_1 : \Omega \to \mathbb{D} \) is an analytic bijection. Then \( f_2 = \phi \circ f_1 \) is an analytic bijection from \( \Omega \) to \( \mathbb{D} \) too.

**Question:** How many analytic bijections from \( \mathbb{D} \) to \( \mathbb{D} \) are there?

**Example:** Obvious ones are \( \phi(z) = e^{i\theta}z \). A more complicated example: fix \( z_0 \in \mathbb{D} \) and let

\[
\phi(z) = \frac{z - z_0}{-z_0z + 1}.
\]

This Möbius transformation maps \( z_0 \mapsto 0 \). Furthermore, it is easy to check that \( |\phi(1)| = |\phi(-1)| = |\phi(i)| = 1 \), and hence \( \phi \) maps the unit circle onto itself, and hence the open disk onto itself. It turns out that these two examples are basically all that you can do.

Our first result is Schwarz’s lemma which places limits on the growth of functions on the unit disk, as long as we know that they aren’t too big at the boundary.
Theorem 52. Suppose that $f : \mathbb{D} \to \mathbb{D}$ is analytic and that $f(0) = 0$. Then

(a) $|f(z)| \leq |z|$ when $|z| < 1$, and

(b) $|f'(0)| \leq 1$.

Furthermore:

(i) If equality holds at any non-zero point in (a), or in (b), then $f(z) = \lambda z$ for some constant $\lambda$.

(ii) If such an equality does not hold then, for each $R < 1$, there exists $\rho < 1$ such that $|f(z)| \leq \rho |z|$ for $|z| \leq R$.

Proof. $\frac{f(z)}{z}$ has a removable singularity at 0, so define

$$g(z) = \begin{cases} \frac{f(z)}{z}, & z \neq 0 \\ f'(0), & z = 0. \end{cases}$$

Suppose that $R < 1$. Since $g$ is analytic on $|z| < R$ and continuous on $|z| \leq R$, the maximum modulus occurs on the boundary, and only there unless $g$ is a constant. On $|z| = R$,

$$|g(z)| = \frac{|f(z)|}{R} < \frac{1}{R}$$

and so this bound holds for all $z$ inside this circle. But this is true for all $R < 1$ and hence we can deduce that $|g(z)| \leq 1$ for all $z \in \mathbb{D}$. This establishes (a) and (b).

If equality holds at some point in (a) then $g$ is a constant and there is a constant $\lambda$ such that $f(z) = \lambda z$. If $|f'(0)| = 1$ then it follows that $f$ is constant from a consideration of Cauchy’s formula for derivatives. (Alternatively, apply the Maximum Modulus Principle to $g(z)$.) The final assertion follows, for if it were false we would have a sequence $\{z_n\}$ with $|z_n| \leq R$ such that $|f(z_n)/z_n| \to 1$. Such a sequence must have a limit point $z$ with $|z| \leq R$ and $|f(z)| = |z|$.

Theorem 53. Every conformal map from the unit disk onto itself is of the form

$$\phi(z) = e^{i\theta} \frac{z - b}{1 - \bar{b}z}$$

where $\theta \in \mathbb{R}$ and $|b| < 1$.

71
Proof. Let \( \phi \) be such a conformal map from \( |z| \leq 1 \) onto itself. Then \( \phi^{-1} \) is well-defined and is itself conformal - remember \((\phi^{-1})'(\phi(z)) = \frac{1}{\phi'(z)}\). Let \( \phi^{-1}(0) = b \). Then \( |b| < 1 \) for otherwise \( \phi^{-1} \) is constant by the maximum modulus principle.

Consider now

\[
\phi_0(z) = \frac{z - b}{1 - b\bar{z}}.
\]

This maps the unit circle to itself and, since \( \phi_0(0) = -b \), the interior to itself. Now consider \( g = \phi \circ \phi_0^{-1}, g^{-1} = \phi_0 \circ \phi^{-1} \). Both satisfy the conditions of Schwarz’s lemma so

\[
|g(z)| \leq |z|, \quad |g^{-1}(w)| \leq |w| \quad \text{for } |z| \leq 1, |w| \leq 1.
\]

That is

\[
|g(z)| \leq |z| \leq |g(z)| \quad \text{for } |z| \leq 1.
\]

Thus \( g(z) = e^{i\theta}z \) by Schwarz’s lemma. Putting this together, \( g(\phi_0(z)) = e^{i\theta}\phi_0(z) \) or \( \phi(z) = e^{i\theta}\phi_0(z) \), as claimed. \( \blacksquare \)

This now allows to to show that there is at most one conformal map from a simply connected proper subset of \( \mathbb{C} \) onto \( \mathbb{D} \) with \( f(z_0) = 0 \) and \( f'(z_0) \) positive. Suppose that \( f_1, f_2 : \Omega \to \mathbb{D} \) are two such maps. Then \( \phi = f_2 \circ f_1^{-1} \) is a conformal bijection from \( \mathbb{D} \) to itself and hence

\[
\phi(z) = e^{i\theta} \frac{z - b}{1 - b\bar{z}}.
\]

But \( \phi(0) = 0 \) and so \( b = 0 \). That is \( \phi(z) = e^{i\theta}z \), or \( f_2(z) = e^{i\theta}f_1(z) \). Thus

\[
f_2'(z_0) = e^{i\theta}f_1'(z_0).
\]

But since both derivatives are real and positive at \( z_0 \), this says that \( e^{i\theta} \) must be 1, and so \( f_2 = f_1 \).

The harder part of the proof of the Riemann Mapping Theorem is to show that any suitable bijection exists at all! Even Riemann fudged when he first ‘proved’ this (in 1851). It took until 1912 before Carathéodory produced what is now considered a valid proof.

Fix then the simply connected proper open subset \( \Omega \) and the point \( z_0 \in \Omega \). Let

\[
\mathcal{F} = \{ f : \Omega \to \mathbb{D} : f \text{ is analytic, } 1-1 \text{ and } f(z_0) = 0 \}.
\]

Note that we don’t require that elements of \( \mathcal{F} \) are onto.
Lemma 54. Suppose that \( f : \Omega \rightarrow \mathbb{C} \) is an analytic 1–1 function. Then \( f(\Omega) \) is homeomorphic to \( \Omega \) and hence \( f(\Omega) \) is simply connected.

Proof. The Inverse Function Theorem implies that \( f \) is a homeomorphism and hence \( f \) preserves simple connectedness.  

If \( \Omega \) is bounded, then it is clear that \( \mathcal{F} \) is nonempty: \( f(z) = (z - z_0)/M \) will work for large enough \( M \).

Lemma 55. \( \mathcal{F} \) is nonempty.

Proof. Since \( \Omega \) is a proper subset of \( \mathbb{C} \) we can choose \( a \notin \Omega \). Since \( \Omega \) is simply connected, its complement is connected as a subset of the Riemann sphere. Thus, there is some continuous path in \( \Omega^c \) connecting \( a \) to \( \infty \). This means (think back to Assignment 1!) that we can choose a branch of the square root function

\[
f : \Omega \rightarrow \mathbb{C}, \quad r(z) = \sqrt{z - a}
\]

which is analytic on \( \Omega \).

Since each point \( w \in \mathbb{C} \) has a unique square, the function \( r \) can only have one of \( w \) or \( -w \) in its range. Pick some \( w_0 \in r(\Omega) \). By the Open Mapping Theorem, \( r(\Omega) \) is open, so for some \( \epsilon > 0 \), \( B(w_0, \epsilon) \subset r(\Omega) \). But this means that \( r(\Omega) \) contains none of the elements of \( B(-w_0, \epsilon) \).

Now let’s start composing maps. The range of the map \( z \mapsto r(z) + w_0 \) now contains none of the points of \( B(0, \epsilon) \), and so

\[
g(z) = \frac{2}{\epsilon}(r(z) + w_0)
\]

completely misses the unit disk. Thus \( h : \Omega \rightarrow \mathbb{D} \)

\[
h(z) = \frac{1}{g(z)}
\]

is well-defined, analytic and 1–1. We can now further compose with an automorphism of the disk to produce a suitable function that sends \( z_0 \) to 0, and so we have produced an element of \( \mathcal{F} \).  

The next lemma is the clever part. For notational convenience, for \( w \in \mathbb{D} \), lets

\[
T_w(z) = \frac{z - w}{1 - \overline{w}z}
\]
denote the natural automorphism of \( D \) which maps \( w \) to 0. This has inverse
\[
T_w^{-1}(z) = \frac{z + w}{1 + wz}.
\]

**Lemma 56.** Suppose that \( f \in F \) satisfies
\[
|f'(z_0)| \geq |g'(z_0)|, \quad \text{for all } g \in F.
\]
Then \( f \) is a bijection from \( \Omega \) to \( D \).

**Proof.** [Remember that we are assuming that elements of \( F \) are 1–1.]

Suppose that \( f \in F \) is not onto. That is, there exists \( w \in D \) which is not in \( f(\Omega) \). Our aim now is to construct a function \( g \in F \) with \( |g'(z_0)| > |f'(z_0)| \).

The map \( \Omega \to D, z \mapsto T_w(f(z)) \) has a simply connected image which does not contain 0. As in the proof of Lemma 55, we can define an analytic branch of the square root function \( r : T_w(f(\Omega)) \to D \). Let \( w' = r(-w) \), and let
\[
g = T_{w'} \circ r \circ T_w \circ f
\]
which is an analytic 1–1 map from \( \Omega \) to \( D \). That is, \( g \in F \).

Let \( s : D \to D \) be
\[
s(z) = T_w^{-1} \left( (T_{w'}^{-1}(z))^2 \right).
\]
Then
\[
s(0) = T_w^{-1} ((w')^2) = T_w^{-1}(-w) = 0.
\]
By Schwarz’s Lemma, \( |s'(0)| \leq 1 \), and if \( |s'(0)| = 1 \) then \( s(z) = \lambda z \) for some \( \lambda \). But \( s \) is not 1–1, so \( s \) is not of this form. Therefore we must have that \( |s'(0)| < 1 \).

Note now that \( f = s \circ g \). Using the chain rule
\[
|f'(z_0)| = |s'(g(z_0))||g'(z_0)| = |s'(0)||g'(z_0)| < |g'(z_0)|
\]
as required. \( \blacksquare \)

To complete the proof of the Riemann Mapping Theorem, we need to show that there is a function \( f \in F \) which has a maximal derivative at \( z_0 \). This is where the more delicate analysis lies. One way of attacking such a problem would be to define the map \( \Psi : F \to \mathbb{R} \), \( \Psi(f) = |f'(z_0)| \). If you could show that \( F \) were compact and \( \Psi \) continuous, then the fact
that continuous images of compact sets are compact would ensure the existence of a suitable function in \( F \). But what topology should one choose to make this work?

Let \( \mathcal{A}(\Omega) \) denote the set of all analytic functions defined on \( \Omega \). This is a vector space, but not (naturally) a normed or metric space, since the functions are not necessarily bounded. The usual topology that we use here is that of uniform convergence on compact subsets of \( \Omega \). Since \( F \subseteq \mathcal{A}(\Omega) \), it would be nice to know if it were compact, or at least precompact.

The way we will progress is slightly less abstract, but you should be aware that what we are doing is really some sort of compactness argument. What we will actually show is

1. If \( \{f_n\} \) is a sequence in \( F \) such that \( \lim_{n \to \infty} |f_n(z_0)| = \sup_{f \in F} |f'(z_0)| \), then this sequence admits a subsequence which converges uniformly on compact subsets of \( \Omega \), to some function \( f : \Omega \to \mathbb{D} \).

2. \( f \) is analytic, 1–1 and \( |f'(z_0)| \) is maximal, and so by Lemma 56 it is also onto.

The result we need here is called Montel’s Theorem and is the one part of the proof of the Riemann Mapping Theorem we won’t prove.

**Definition:** A subset \( S \subseteq \mathcal{A}(\Omega) \) is **locally bounded** if for every compact subset \( K \subseteq \Omega \),

\[
\sup_{f \in S} \left\{ \sup_{z \in K} |f(z)| \right\}
\]

is bounded.

**Theorem 57.** (Montel’s Theorem\(^2\)) Let \( S \) be a locally bounded subset of \( \mathcal{A}(\Omega) \). Then every sequence \( \{f_n\} \subseteq S \) has a subsequence which converges uniformly on compact subsets of \( \Omega \) to an analytic function.

In our particular case, if \( K \) is a compact subset of \( \Omega \) and \( f \in F \), then \( f(K) \subseteq f(\Omega) \subseteq \mathbb{D} \) and so \( F \) is clearly locally bounded. Let \( L = \sup_{f \in F} |f'(z_0)| \). Then there exists a sequence \( \{f_n\} \subseteq F \) such that \( |f_n'(z_0)| \to L \). By Montel’s Theorem then, this sequence has a subsequence \( \{f_{n_k}\} \) which converges uniformly on compact subsets of \( F \) to an analytic function.

\(^2\)This result depends on the Arzelà-Ascoli theorem which identifies relatively compact subsets of \( C(K) \). The main step is to show that the elements of \( S \) must be ‘equicontinuous’, which is to say, that for every \( \epsilon > 0 \), there exists \( \delta > 0 \) such that for all \( f \in S \),

\[
|f_n(x) - f_n(y)| < \epsilon, \quad \text{whenever } |x - y| < \delta.
\]

You can look up the detail (involving finite subcovers of open covers and \( \frac{\epsilon}{n} \) arguments, etc) in the references given earlier.
Note that the derivatives also converge uniformly on compact subsets of $\Omega$ and so in particular
\[ f_{n_k}(z_0) = 0 \to f(z_0), \quad \text{and} \quad f'_{n_k}(z_0) \to f'(z_0). \]

**Exercise:** Prove that $f(\Omega) \subseteq \mathbb{D}$.

If we can prove that $f \in \mathcal{F}$ then we are done, since then $f$ would satisfy the hypotheses of Lemma 56, and hence it would be a bijection. The remaining property of $f$ that needs to be shown is that $f$ is 1–1.

This is not entirely straightforward. Suppose that $\Omega = \mathbb{D}$ and define $f_n(z) = z/n$. If $z_0 = 0$ then all these functions are in the set $\mathcal{F}$ for this $\Omega$ and $z_0$. The sequence $\{f_n\}$ clearly converges to $f(z) = 0$ uniformly on compact subsets of $\mathbb{D}$, but of course this $f$ is not 1–1.

The extra ingredient that we need to show that $f$ is 1–1 is that $f$ is not constant. We know this because $f'(z_0) \neq 0$ (since $|f'(z_0)| \geq |f'_1(z_0)| > 0$).

For notational sanity, let’s re-choose the sequence so that in fact $f_n \to f$ uniformly on compact subsets of $\Omega$.

Suppose that $w \in \Omega$. Let $g_n : \Omega \to \mathbb{C}$ be defined by $g_n(z) = f_n(z) - f_n(w)$ and let $g(z) = f(z) - f(w)$. Let $\Omega' = \Omega \setminus \{w\}$, which is open and connected. Now for every $n$, $g_n$ has no zeros in $\Omega'$ as $f_n$ is 1–1. Suppose however that $g$ has a zero at some point $v$ in $\Omega'$. Since $g$ is analytic and not identically zero, $v$ must be an isolated zero. That is, there exists $r > 0$ such that if $0 < |z - v| \leq r$ then $g(z) \neq 0$. (We of course choose $r$ so that $D(v, r) \subseteq \Omega'$.) It follows that $\epsilon = \min \{|g(z)| : |z - v| = r\}$ is strictly positive.

Now $g_n \to g$ uniformly on $|z - v| = r$ so there exists $N$ such that if $n \geq N$,
\[ \sup_{|z-v|=r} |g_n(z) - g(z)| < \epsilon. \]

Thus, if $n \geq N$,
\[ |g_n(z) - g(z)| < |g(z)| \quad \text{on} \quad |z - v| = r. \]

It follows from Rouché’s Theorem that for such $n$, $g_n$ and $g$ have the same number of zeros in $|z-v| < r$. Thus $g_n$ must have a zero in this disk and hence in $\Omega'$. But this is a contradiction. Thus $g$ has no zeros in $\Omega'$ and hence $f$ is 1–1 on $\Omega$.

Putting everything together now, we have shown that there exists a function $f \in \mathcal{F}$ such that $|f'(z_0)| \geq |g'(z_0)|$ for all $g \in \mathcal{F}$, and hence we have found a bijective element of $\mathcal{F}$, and this completes the proof of the Riemann Mapping Theorem.
This last bit of the proof essentially establishes Hurwitz’s Theorem, and more particularly the following corollary.

**Theorem 58.** Suppose that \( \{ f_n \} \) is a sequence of 1–1 functions on an open connected set \( \Omega \) which converges uniformly on compact subsets to the function \( f \). Then \( f \) is either 1–1 or constant.

**Remark:** An issue that arose earlier was what happens at the boundary. That is, can the bijection \( f \) be extended to a continuous map sending the the closure of \( \Omega \) to the closed unit disk. The answer turns out to depend very much on the set \( \Omega \). We saw for example, that every analytic bijection of \( \mathbb{D} \) onto itself is of a particular form, and all those maps to extend beyond the open unit disk. On the other hand, If \( \Omega = \mathbb{C} \setminus (-\infty, 0] \) then, taking all square roots in the right half-plane,

\[ f(z) = \frac{\sqrt{z} - 1}{\sqrt{z} + 1} \]

maps \( \Omega \) conformally onto \( \mathbb{D} \), but evidently cannot be extended to the closure of \( \Omega \) which is the whole complex plane.

**Definition:** A point \( w \) in the boundary of \( \Omega \) is called simple if for any sequence \( \{ z_n \} \) converging to \( w \), there is a contour \( \gamma : [0, 1] \to \Omega \cup \{ w \} \) in \( \Omega \) that contains the points \( z_n \) and which terminates at \( w \).

**Example:** Let \( \Omega = \mathbb{C} \setminus (-\infty, 0] \). Then \( -1 \) is not a simple boundary point. To see this, take \( z_n = -1 + \frac{(-1)^n i}{n} \). No path joining these points can be continuous at \( t = 1 \). On the other hand, 0 is simple.

We won’t prove the following, but include it to show that for large classes of sets \( \Omega \), a continuous extension will exist. The proof is not easy!

**Theorem 59.** Suppose that \( \Omega \) is a bounded simply connected subset of \( \mathbb{C} \) and that every boundary point of \( \Omega \) is simple. If \( f \) is an analytic bijection from \( \Omega \) onto \( \mathbb{D} \), then \( f \) extends to a homeomorphism of \( \overline{\Omega} \) onto \( \overline{\mathbb{D}} \).

Although this theorem deals with most of the cases of interest, it is far from giving a characterization of sets where an extension to the boundary exists.
10 Riemann Surfaces

10.1 Complex manifolds

In differential geometry you study $d$-dimensional manifolds, which are sets which are locally like bits of $\mathbb{R}^d$. In particular, a topological space $M$ is a 2-dimensional manifold if there is an open covering $\mathcal{U} = \{U_\alpha\}$ of $M$ and corresponding homeomorphisms (or charts) $h_{U_\alpha}$ between $U_\alpha$ and an open subset of $\mathbb{R}^2$. These need to join nicely: if $U, V \in \mathcal{U}$ and $U \cap V$ is nonempty, then the transition map $T_{U,V} = h_U \circ h_V^{-1}$ which maps an open subset of $\mathbb{R}^2$ to another open subset of $\mathbb{R}^2$ should have suitable properties. For example, we say $M$ is a differentiable if the transition maps are all differentiable.

**Example:** Take $M = \overline{\mathbb{C}}$. Let $\mathcal{U}$ consist of the two open sets $D = \{z : |z| < 2\}$ and $U = \{z : |z| > 1\} \cup \{\infty\}$, where we are writing subsets of $\mathbb{R}^2$ in complex notation. Let $h_D(z) = z$ and $h_U(z) = \frac{1}{z}$, which has image $|z| < 1$. In the open region

$$\Omega_1 = h_U(D \cap U) = \{z : \frac{1}{2} < |z| < 1\}$$

we have the transition map

$$T_{D,U}(z) = \frac{1}{z}$$

which gives a homeomorphism of $\Omega_1$ onto

$$\Omega_2 = \{z : 1 < |z| < 2\}.$$

(The other transition map is just the inverse mapping!)

All this allows one to unambiguously decide that a function $f : M \to \mathbb{C}$ is differentiable if $f \circ h_U^{-1}$ is always a differentiable function on the set $h_U(U) \subseteq \mathbb{R}^2$.

As in the example, topologically, $\mathbb{C}$ is identical to $\mathbb{R}^2$. In the complex realm however we want to be able to have the induced composition maps $f \circ h_U^{-1}$ being analytic. For this we need that the transition maps $T_{U,V}$ are analytic functions. If this is the case we’ll say that $M$ is a complex manifold. Note that for such spaces, the Riemann Mapping Theorem ensures that we could always choose things so that $h_U$ is a homeomorphism onto the unit disk $\mathbb{D}$.

In the example above, the transition maps are analytic on their domains, but it would be easy to construct one that weren’t!

**Definition:** Suppose then that $M, N$ are complex manifold, with charts $\{h_U\}$ and $\{g_V\}$ respectively. A function $f : M \to N$ is **holomorphic** if for every pair of charts $h_U$ and $g_V$, ...
the composition $F = g_V \circ f \circ h_U^{-1}$ is analytic wherever it is defined. (That is, at those $z$ in the domain of $h_U^{-1}$ such that $f(h_U^{-1}(z)) \in V$.)

The charts let us transfer a complex analytic structure from $\mathbb{C}$ onto the manifolds. I have deliberately used the word holomorphic here just so that we can easily distinguish between this definition and what we had earlier.

**Example:** Let $M = N = \mathbb{C}$. Recall that with our earlier definition the function $f(w) = w$ on $\mathbb{C}$ is not analytic at $\infty$ since $f(1/w) = 1/w$ is not analytic at 0.

To check whether $f$ is holomorphic we need to look at the compositions with the charts:

$$g_D, h_D \colon z \mapsto z, \quad D = \{|z| < 2\}$$

$$g_U, h_U \colon z \mapsto \frac{1}{z}, \quad U = \{|z| > 1\}$$

For example:

$$g_U \circ f \circ h_U^{-1}(z) = z, \quad |z| < 1,$$

which is certainly analytic. Or

$$F(z) = g_U \circ f \circ h_D^{-1}(z) = \frac{1}{z}$$

for those $z$ in $|z| < 2$ for which $f \circ h_D^{-1}(z) = z$ is in $U = \{|z| > 1\}$. That is here, the domain of $F$ is the annulus $\{1 < |z| < 2\}$ and on that set $F$ is analytic. In any case, checking all the possibilities shows that $f$ is holomorphic.

Basically, meromorphic functions on $\mathbb{C}$ are holomorphic functions from $\mathbb{C}$ to $\mathbb{C}$ in this new context. Many references will talk about automorphisms of the Riemann sphere as being biholomorphic functions: holomorphic bijections with holomorphic inverses.

The chart basically lays a piece of $\mathbb{C}$ onto the surface $M$ is a way that one can then do complex analysis. At this point, the theory can disappear off into very abstract settings, so we’ll just look at something a bit more concrete.

### 10.2 Riemann surfaces for functions

Let $M = \{(r, \theta) : r > 0, \ 0 \leq \theta < 4\pi\}$. If we join the boundary $\theta = 0$ to $\theta = 4\pi$ we get what is geometrically an infinite cylinder and so we have a natural topology on $M$.

One can map elements of $M$ onto the complex plane by $\Phi(r, \theta) = re^{i\theta}$. This map is locally invertible, but of course it is 2–1 not 1–1 on $M$. Indeed it is often better to think
of $M$ as being two copies of the punctured plane, but joined together in a way that if you start at $z = 1$ on one plane and go around the origin once you end up not back where you started, but on the other copy of the plane.

It is easy to construct some charts on $M$ to make it a complex manifold: for $n = 0, 1, 4$, let $U_n = \{(r, \theta) : n\pi < \theta < (n + 2)\pi\}$, where the angles are all calculated modulo $4\pi$. The corresponding homeomorphisms are all of the form $h_n(r, \theta) = re^{i\theta}$.

**Question:** What do the transition maps do?

Consider the function $f : M \to \mathbb{C}$ defined by $f(r, \theta) = \sqrt{re^{i\theta}}$. This function is now a continuous bijection from $M$ to $N = \mathbb{C} \setminus \{0\}$. We can take the identity chart on $\mathbb{C}$, so $f$ is holomorphic if

$$F_n = f \circ h_n^{-1}$$

is analytic for each $n$. Each of these maps is just an analytic branch of the square root function.

The usual map $z \mapsto z^{1/2}$ is slightly problematic as it is multiple valued. The principal value version requires that we take a branch cut along which the function is not continuous, and which only maps onto half the complex plane. This ‘Riemann surface’ version allows us to have a single-valued holomorphic square root, essentially by allowing copies of the domain space, joined together in a nice way.

Thought of another way, one could start with $s$, the principal branch of the square root function and try to analytically continue this. For example, at the point $-10 + i$, you can write a Taylor series $\tilde{s}$ for $\sqrt{e^{i\theta}}$ which converges on a disk (of radius $\sqrt{101}$ that crosses the negative real axis. Clearly $\tilde{s}(-10 - i) \neq s(-10 - i)$, so you want to consider $\tilde{s}$ as being defined on a different ‘copy’ of the plane. However, once one continues around the origin twice, you find that the continuation gets back to the original function $s$. Thus the Riemann surface $M$ is the natural surface for the square root function.

One can do essentially the same thing with other multivalued functions, except that you might not end up back where you started!

**Example:** Let $\ell(z) = \log z = \ln |z| + i \arg z$. The natural surface here would be

$$M = \{(r, \theta) : r > 0, \ \theta \in \mathbb{R}\}.$$ 

We can impose a complex analytic structure on $M$ using a now infinite family of charts $(U_n, h_n)$ defined just as in square root example. Let $f : M \to \mathbb{C} \setminus \{0\}$ be $f(r, \theta) = \ln r + i\theta$. 

80
Again this is now a holomorphic bijection of $M$ onto $N = \mathbb{C} \setminus \{0\}$. Now we have infinitely many surfaces sitting above each point in the complex plane, one for each value of $\log z$, joined up like some infinite screw thread.

**Example:** Let $a(z) = \arcsin z$. Again this is a multivalued function. The set of solutions of $\sin w = z$ is rather more complicated. One wants to place each of these points on a different sheet of the Riemann surface. Nonetheless, you can imagine moving about the plane in a continuous way, patching together locally analytic branches of the arcsin function to create a rather complicated surface.

Note that we can look at all these things backwards. If $M$ is the double sheeted surface we first defined associated with the square root function, and $N$ is the punctured plane $\mathbb{C} \setminus \{0\}$, then we can think of the squaring function $S(re^{i\theta}) = (r^2, 2\theta)$ as being a version of $z \mapsto z^2$ mapping $N$ holomorphically onto $M$.

### 10.3 Compact Riemann surfaces

The Riemann surfaces looked at so far are not compact. On the other hand, it is possible to put a complex analytic structure on a torus via a suitable selection of charts.

The cylinder $M$ arose from a periodicity in the coordinate $\theta$ in the associated version of the square root function: $f(r, \theta + 4\pi) = f(r, \theta) = \sqrt{r}e^{i\theta/2}$. A torus can arise when the function has periodicity in two ‘coordinates’. In particular, if the function has the property that there exist ‘independent’ $\omega_1, \omega_2 \in \mathbb{C}$

$$f(z + k\omega_1 + \ell\omega_2) = f(z)$$

for all $k, \ell \in \mathbb{Z}$, then the associated Riemann surface will be a torus.

**Question:** Do such functions exist?

This leads on to the topic of elliptic functions!
11 Elliptic functions

A classical problem, going back to Fagnano and Euler in the 18th century, is to calculate the arc length of an ellipse. This involves calculating integrals of the form

$$\int_c^x R(t, \sqrt{\phi(t)}) \, dt$$

where $\phi$ is a polynomial of degree 3 or 4, and $R$ is a rational function. This led to a more general theory of what are now called elliptic integrals.

In this theory we now allow $\phi(z)$ to be a polynomial in $z$ of degree 3 or 4 with complex coefficients, and $R(z, w)$ to be a rational function in $z$ and $w$. An elliptic integral is an antiderivative of the form

$$F(z) = \int R(z, \sqrt{\phi(z)}) \, dz.$$ 

Any elliptic integral may be expressed by a suitable change of variables as a sum of elementary functions and elliptic integrals of the following 3 kinds:

$$F(z; k) = \int \frac{dz}{\sqrt{(1 - z^2)(1 - k^2z^2)}} \quad (1^{st} \text{ kind})$$

$$E(z; k) = \int \sqrt{\frac{1 - k^2z^2}{1 - z^2}} \, dz \quad (2^{nd} \text{ kind})$$

$$\Pi(z; a, k) = \int \frac{dz}{(1 - a^2z^2) \sqrt{(1 - z^2)(1 - k^2z^2)}} \quad (3^{rd} \text{ kind})$$

The elliptic functions were first introduced as the inverse functions of elliptic integrals. This is analogous to defining the sin function as being the inverse of

$$s(x) = \int \frac{dx}{\sqrt{1 - x^2}}.$$ 

However, since it has been realized that elliptic functions can be characterized as functions with double periodicity, it is now customary to define them as doubly periodic functions.

**Definition:** A non-zero number $w$ is a period of $f$ if $f(z + w) = f(z)$ for all $z$ (for example, $e^z$ has period $2\pi i$).

It turns out that there are only two interesting cases – either $f$ has essentially one period or two periods, $w_1$, $w_2$ with $\frac{w_1}{w_2}$ not real. Let us make this precise.

We say $f$ has *arbitrary small periods* if there is a sequence of periods $\{w_n\}$ with $|w_n| \to 0$. 

82
Lemma 60. If $f$ is meromorphic in $\mathbb{C}$ and has arbitrary small periods then $f$ is constant.

Proof. Whenever $f'(z)$ exists we have
\[
\lim_{n \to \infty} \frac{f(z + w_n) - f(z)}{w_n} = f'(z) = 0.
\]
Since poles are isolated (being isolated singularities) this implies $f$ is constant. ■

Lemma 61. If $f$ has periods $w_1, w_2, \ldots, w_k$ then for any integers $n_1, n_2, \ldots, n_k$
\[
w = n_1w_1 + n_2w_2 + \cdots + n_kw_k
\]
is a period (if it is not zero).

Proof. This is obvious! ■

Let $\Lambda_f$ denote the set of all periods of $f$, plus the origin. A consequence of Lemma 61 is that if $f$ is not constant, then the set $\Lambda$ forms a (point) lattice\(^3\) in the plane, that is a discrete subgroup of the additive group of points in the plane. The next result is really about the structure of lattices in the plane, rather than being about periodic functions.

Theorem 62. Suppose $f$ is periodic but not constant. Then $\Lambda_f$ is of one the following two forms:

1) $\{nw : n \in \mathbb{Z}\}$ for some $w \neq 0$, or
2) $\{n_1w_1 + n_2w_2 : n_1, n_2 \in \mathbb{Z}\}$ where $\frac{w_1}{w_2}$ is not real.

We break the proof up as a chain of lemmas and remarks. The first step uses a lemma of Kronecker:

Lemma 63. If $\theta$ is irrational, $\epsilon > 0$, there exist integers $p, q$ such that
\[
|\frac{p}{q} - \theta| < \frac{\epsilon}{q}.
\]

\(^3\)Some authors require that a lattice is of full dimension, that is, that it doesn’t sit within a proper subspace.
Proof. Choose \( N > \frac{1}{\epsilon} \) and consider the fractional parts of \( \theta, 2\theta, \ldots, N\theta, (N + 1)\theta \). There are \( N + 1 \) numbers to be fitted into the \( N \)-intervals \( \left[ 0, \frac{1}{N} \right], \left[ \frac{1}{N}, \frac{2}{N} \right], \ldots, \left[ \frac{N-1}{N}, 1 \right] \) which divide \([0, 1] - \) and no number coincides with an end point. Consequently there must be two of our numbers in one of these subdivisions, that is, \(|k\theta - l\theta - m| < \frac{1}{N} \) for integers \( k, l, m \) with \( 1 \leq k < l \leq N + 1 \) (This is an application of the so called “pigeon-hole principle”). We may rewrite this as

\[ |q\theta - p| < \epsilon \quad \text{where} \quad 1 \leq q \leq N. \]

Returning to the theorem, suppose that \( w_1, w_2 \) are periods of \( f \), and hence that all nonzero numbers in the set \( P = \{ n_1w_1 + n_2w_2 : n_1, n_2 \in \mathbb{Z} \} \) are also periods. We first want to show that if \( \alpha = \frac{w_1}{w_2} \) is real then it must be rational, and from this deduce that that (1) holds. If \( \alpha \) is real, then the elements of \( P \) lie on a line through the origin.

Suppose then that \( \alpha \) is irrational. Fix \( \epsilon > 0 \). By the lemma there exist integers \( p, q \) such that

\[ \frac{|p - \alpha q|}{q} \leq \frac{\epsilon}{q|w_2|} \]

(applying the lemma with \( \epsilon/|w_2| \) rather than \( \epsilon! \)). Then, using the fact that \( w_1 - \alpha w_2 = 0 \),

\[
|qw_1 - pw_2| = q \left| w_1 - \frac{p}{q} w_2 \right|
= q \left| (w_1 - \frac{p}{q} w_2) - (w_1 - \alpha w_2) \right|
= q \left| \frac{p}{q} - \alpha \right| |w_2|
< \epsilon.
\]

This implies \( f \) has arbitrary small periods, contradicting the hypothesis on \( f \). Therefore \( \alpha \) must be rational, say \( \alpha = \frac{p}{q} \).

In this case

\[
P = \left\{ \frac{n_1p + n_2q}{q} w_2 : n_1, n_2 \in \mathbb{Z} \right\}
\]

consists of certain integer multiples of \( \frac{w_2}{q} \), and so there is a nonzero element of \( P \), say \( w \), closest to the origin. Write

\[ w_1 = nw + \rho, \quad 0 \leq |\rho| < |w|. \]
If \( \rho \neq 0 \) then \( \rho \) is a period contradicting the choice of \( w \). So \( w_1 \) is an integer multiple of \( w \) as is any period on the line through \( w_1, w_2 \).

**Exercise:** Use this to show that either case (1) of the theorem holds or else there exist 2 periods \( w_1, w_2 \) of \( f \) such that \( \frac{w_1}{w_2} \) is not real.

Suppose then that the elements of \( \Lambda_f \) don’t all lie in a line through the origin. Again, if \( f \) does not admit arbitrarily small periods, then there must be a positive lower bound on the distance between the elements of \( \Lambda_f \). And this lower bound must be achieved (why?), and at just a finite number of points (why?).

Choose \( w_1 \) such that \( |w_1| \) is a minimum and such that \( w_1 \) has the least nonnegative argument amongst such points. Now choose \( w_2 \) on the same circle if possible having the next smallest argument. Otherwise pass to the next smallest circle containing a period and choose \( w_2 \) with least nonnegative argument not equal to that of \( w_1 \).

Note that \( |w_1| \leq |w_2| \), and (since \( w_1 \pm w_2 \) are periods) \( |w_1 + w_2| \geq |w_2| \) and \( |w_1 - w_2| \geq |w_2| \).

It follows from the construction that there are no periods in the triangle with vertices \( 0, w_1, w_2 \), that is, no periods of the form

\[
  w = \alpha w_1 + \beta w_2 \quad \text{with} \quad 0 < \alpha < 1, \quad 0 < \beta < 1.
\]

Now let \( w_3 \) be any period. Since \( w_1, w_2 \) are linearly independent vectors in \( \mathbb{C} \) we can write \( w_3 = aw_1 + bw_2 \) where \( a, b \) are real numbers and \( w_3 = mw_1 + nw_2 + \alpha w_1 + \beta w_2 \) where \( m, n \), are integers and \( 0 \leq \alpha < 1, \quad 0 \leq \beta < 1 \). Since \( w_3 - mw_1 - nw_2 \) is a period we must have \( \alpha = \beta = 0 \). Hence \( w_3 = mw_1 + nw_2 \) and the theorem is proved.

**Definition:** The parallelogram \( 0, w_1, w_2, w_1 + w_2 \) is called the *fundamental period parallelogram*. The set \( \Lambda_f \) is called the *period lattice* for \( f \).
Definition: An elliptic function is a doubly periodic meromorphic function. That is, a meromorphic function with two linearly independent periods.

Thus, once one know the values of an elliptic function \( f \) on the fundamental period parallelogram, one know it everywhere. In fact, what one really needs is just to know the values of \( f \) on any translate of the fundamental period parallelogram. Such parallelograms are called cells. From our point of view, one might consider the cell as a torus, and an elliptic function as a holomorphic function from this Riemann surface to the Riemann sphere. We won’t however pursue this very far.

Since poles are isolated singularities any bounded region in the plane contains only a finite number of poles. It is often inconvenient to have the poles on a boundary so we choose a cell so that all the poles lie in the interior of the cell.

Lemma 64. If \( f \) is an elliptic function then in any cell the number of poles of \( f \) equals the number of zeros of \( f \).

Proof. Let \( C \) be the boundary contour of a cell. By periodicity \( \int_C \frac{f'}{f} = 0 \). But \( \int_C \frac{f'}{f} = 2\pi i (Z - P) \) where \( Z = \) the number of zeros and \( P = \) the number of poles. ■

Definition: The order of an elliptic function is the number of poles per cell, counted according to multiplicity.

Suppose now that \( \Lambda \) is a point lattice in the plane with basis \( w_1 \) and \( w_2 \). One can look at the set

\[ \mathcal{E}(w_1, w_2) = \{ f : \mathbb{C} \to \mathbb{C} : f \text{ is an elliptic function with period lattice } \Lambda \} \]

The set \( \mathcal{E}(w_1, w_2) \) is of course non-empty since the constant functions are elements. The following result is easy.

Theorem 65. \( \mathcal{E}(w_1, w_2) \) is a subfield of the field of meromorphic functions. Moreover, if \( f \in \mathcal{E}(w_1, w_2) \) then \( f' \in \mathcal{E}(w_1, w_2) \) too.

The problem remains as to whether there are any non-constant elliptic functions!
12 Construction of elliptic functions

We shall begin by constructing Weierstrass’s \( \wp \) function, which is a nontrivial elliptic function with a single pole (in each cell) of order two.

Let \( \Lambda = \Lambda(w_1, w_2) \) be our period lattice, and let \( \Lambda' \) denote the nonzero elements of this lattice. One way to construct a periodic function would be to try to define

\[
f(z) = \sum_{w \in \Lambda} \frac{1}{(z - w)^2}.
\]

**Example:** Let’s take \( \Lambda = \Lambda(1, i) \). Then we could try

\[
f(z) = \sum_{n,m \in \mathbb{Z}} \frac{1}{(z - n - mi)^2}.
\]

Unfortunately, as we shall see, there is a problem with convergence!

**Lemma 66.** For real \( \alpha \), \( \sum_{w \in \Lambda'} \frac{1}{w^\alpha} \) converges absolutely if and only if \( \alpha > 2 \).

**Proof.** Let’s just do this for the lattice in the example. The general case is pretty similar.

We sum in concentric “annular” rings.

Straightening the picture, look at the diagram above. In the first “ring” there are 8 vertices, that is periods, \( w \) with \( 1 \leq |w| \leq \sqrt{2} \). In the next ring there are \( 2 \times 8 \) periods with \( 2 \leq |w| \leq 2\sqrt{2} \), and, in general, \( 8n \) with \( n \leq |w| \leq n\sqrt{2} \).

Now we have

\[
\sum_{n=1}^{\infty} \frac{8n}{n^\alpha 2^{\alpha/2}} \leq \sum_{w \in \Omega} \frac{1}{|w|^\alpha} \leq \sum_{n=1}^{\infty} \frac{8n}{n^\alpha}
\]

which gives the result.
Lemma 67. If $\alpha > 2$, $R > 0$ then \[
\sum_{w \in \Lambda, \ |w| > R} \frac{1}{|z-w|^\alpha} \]
converges uniformly and absolutely on $|z| \leq R$.

Proof. Fix $\alpha, R$. We will show
\[
\frac{1}{|z-w|^\alpha} \leq M \frac{1}{|w|^\alpha}
\]
so that we will have the result by the previous lemma. In fact there is a $w$ of least modulus with $|w| > R$ so we have some $\rho$ with
\[
\left| \frac{z}{w} \right|^{\alpha} \leq \rho < 1 \quad \text{for all } |z| \leq R
\]

\[
\left| \frac{w}{z-w} \right|^\alpha = \frac{1}{|z/w-1|^\alpha} \leq (1-\rho)^{-\alpha}, \quad \text{so take } M = (1-\rho)^{-\alpha}.
\]

Definition: Weierstrass’s $\wp$ function: $\wp(z) = \frac{1}{z^2} + \sum_{w \in \Lambda, w \neq 0} \left\{ \frac{1}{(z-w)^2} - \frac{1}{w^2} \right\}$.

We should of course talk about the Weierstrass elliptic function associated to the period lattice $\Lambda$, but we shall suppress this dependence in what follows.

Theorem 68. $\wp$ has periods $w_1, w_2$ and is analytic except at $w \in \Lambda$ where it has double poles. Further, $\wp$ is an even function.

Proof. We first check that the series appropriately converges. Consider any disk $D_R = \{|z| \leq R\}$. For any $z \in D_R$ consider the deleted series
\[
\sum_{w \in \Lambda \setminus D_R} \left( \frac{1}{(z-w)^2} - \frac{1}{w^2} \right)
\]

The series converges uniformly and absolutely because
\[
\left| \frac{1}{(z-w)^2} - \frac{1}{w^2} \right| \leq \frac{z(2w-z)}{|z-w|^2 w^2} \leq \frac{R(2|w|+R)}{|z-w|^2 |w|^2} \leq \frac{R(2+\frac{R}{|w|})}{|z-w|^2 |w|},
\]

so that
\[
\left| \frac{1}{(z-w)^2} - \frac{1}{w^2} \right| \leq 3MR \frac{1}{|w|^3},
\]

This sum is therefore analytic for $z \in D_R$. Adding back the missing terms shows that $\wp$ is analytic except at $w \in \Lambda$ where it clearly has double poles.
The evenness of $\wp$ follows once we are convinced that
\[
\frac{1}{z^2} + \sum_{w \in \Omega, w \neq 0} \left( \frac{1}{(z-w)^2} - \frac{1}{w^2} \right) = \frac{1}{z^2} + \sum_{w \in \Lambda, w \neq 0} \left( \frac{1}{(z-w)^2} - \frac{1}{w^2} \right).
\]

The point is that \(\frac{1}{z^2} + \sum_{w \in \Omega, w \neq 0} \left( \frac{1}{(z-w)^2} - \frac{1}{w^2} \right)\) runs through the same values as \(\frac{1}{(z-w)^2} - \frac{1}{w^2}\) as $\Lambda$ is symmetric, and we can rearrange because of *absolute convergence*. It remains to show that the function has periods $w_1, w_2$. Direct verification is relatively cumbersome. By termwise differentiation

\[
\wp'(z) = \sum_{w \in \Omega} \frac{-2}{(z-w)^3}
\]

which obviously converges uniformly and absolutely in every disk $|z| \leq R$ (provided we omit a finite number of terms). However

\[
\wp'(z + w_1) = \sum_{w \in \Omega} \frac{-2}{(z + w_1 - w)^3} = \sum_{w \in \Omega} \frac{-2}{(z-w)^3}
\]
(again we cannot have any problems as the series converges absolutely) so $\wp'$ has periods $w_1$ and $w_2$. So $\wp(z + w_1) - \wp(z)$ is constant. Let $z = -\frac{1}{2}w_1$ to see by evenness that the constant is zero. ■

For $n \geq 3$ define the constants $G_n = \sum_{w \in \Lambda} \frac{1}{w^n}$ (which converges by Lemma 66). The next result gives the Laurent series for $\wp$.

**Lemma 69.** Let $r = \min\{|w| : w \in \Lambda, w \neq 0\}$. Then, for $0 < |z| < r$,
\[
\wp(z) = \frac{1}{z^2} + \sum_{n=1}^{\infty} (2n+1) G_{2n+2} z^{2n}.
\]

**Proof.** For $0 < |z| < r$, \(\left| \frac{z}{w} \right| < 1\) and
\[
\frac{1}{(z-w)^2} - \frac{1}{w^2} = \frac{1}{w^2} \left( \left( 1 - \frac{z}{w} \right)^{-2} - 1 \right) = \sum_{j=1}^{\infty} (j+1) \frac{z^j}{w^{j+2}}
\]

Absolute convergence allows us to interchange the order of the sums, so
\[
\frac{1}{z^2} + \sum_{w \neq 0} \left\{ \frac{1}{(z-w)^2} - \frac{1}{w^2} \right\} = \frac{1}{z^2} + \sum_{j=1}^{\infty} (j+1) \left( \sum_{w \neq 0} \frac{1}{w^{j+2}} \right) z^j = \frac{1}{z^2} + \sum_{j=1}^{\infty} (j+1) G_{j+2} z^j
\]

But $\wp$ is even so the odd terms vanish and the result is proved, by setting $j = 2n$. ■

**Notation:** It is customary to set $g_2 = 60G_4$, $g_3 = 140G_6$.

Weierstrass’s $\wp$ function appears in the theory of differential equations as the solution of a non-linear first order ODE.
**Theorem 70.** \((\wp'(z))^2 = 4\wp^3(z) - g_2\wp(z) - g_3.\)

**Proof.** By Theorem 65, both sides of the equation are elliptic functions and the only possible pole of the difference in a cell round 0 is at 0. Let’s look at the Laurent series of both sides of this equation:

\[
\wp'(z) = -\frac{2}{z^3} + 6G_4z + 20G_6z^3 + \ldots
\]

so

\[
\wp'(z)^2 = \frac{4}{z^6} - 24\frac{G_4}{z^2} - 80G_6 + \ldots.
\]

On the other hand

\[
4\wp^3(z) - g_2\wp(z) - g_3 = \left(\frac{4}{z^6} + 36\frac{G_4}{z^2} + 60G_6 + \ldots\right) - 60G_4\left(\frac{1}{z^2} + \text{analytic}\right) - 140G_6
\]

\[
= \frac{4}{z^6} - 24\frac{G_4}{z^2} + \text{analytic}.
\]

Subtracting these shows that

\[(\wp'(z))^2 - (4\wp^3(z) - g_2\wp(z) - g_3)\]

is an analytic elliptic function (after removing the singularity) and hence it is constant. Indeed with a little more care one can check that the constant terms of these Laurent expansions match and hence the two side of the equation must be equal. \(\blacksquare\)

We could rewrite this differential equation as

\[
\frac{dw}{dz} = \sqrt{4w^3 - g_2w - g_3}
\]

(where we are writing \(w\) for \(\wp(z)\)) or

\[
\frac{dw}{\sqrt{4w^3 - g_2w - g_3}} = dz.
\]

Written more formally, suppose that \(C\) is a path in \(\mathbb{C}\) from \(z_0\) to \(z\) avoids the zeros and poles of \(\wp'(z)\), and where the sign of the square root must be chosen so that it actually equals \(\wp'(z)\). Then this says that

\[
\int_{\wp(z_0)}^{\wp(z)} \frac{dw}{\sqrt{4w^3 - g_2w - g_3}} = \int_{z_0}^{z} dz = z - z_0.
\]

Thus, \(\wp(z)\) is the inverse of an elliptic integral.

The centrality of the Weierstrass \(\wp\) function to the study of elliptic functions is underscored by the following remarkable result.
Theorem 71. Any elliptic function $f$ with given period lattice, $\Lambda$, can be expressed in the form

$$f(z) = R_1\{\wp(z)\} + \wp'(z)R_2\{\wp(z)\}$$

where $R_1$ and $R_2$ are rational functions.

Proof. First we assume that $f$ is an even function and is analytic and nonzero at the lattice points. If $a_1$ is a zero of $f$ in the fundamental period–parallelogram then a point in this parallelogram congruent to $-a_1$ is also a zero. We can, therefore, choose $n$ zeros $a_1, a_2, \ldots, a_n$ in the fundamental parallelogram, each multiple zero being repeated according to its multiplicity, in such a way that they, together with the points in the parallelogram congruent to $-a_1, \ldots, -a_n$ are all the zeros in the parallelogram. Similarly we can choose $n$ poles $b_1, \ldots, b_n$ such that they, together with the points congruent to $-b_1, \ldots, -b_n$ in the parallelogram, are all the poles in the fundamental parallelogram. The function

$$F(z) = \prod_{\nu=1}^{n} \frac{\wp(z) - \wp(a_{\nu})}{\wp(z) - \wp(b_{\nu})},$$

where $\wp(z)$ has the same primitive periods as $f(z)$, is an elliptic function having the same zeros and poles as $f(z)$. [Why?] Hence the ratio $\frac{f(z)}{F(z)}$ is an elliptic function with no poles in the fundamental period–parallelogram and so is constant by Liouville's theorem. Thus

$$f(z) = A \prod_{\nu=1}^{n} \frac{\wp(z) - \wp(a_{\nu})}{\wp(z) - \wp(b_{\nu})} \quad (*)$$

If $f$ has a pole or a zero at the origin (and hence at all lattice points) such a pole or zero must be of even order (we are still assuming that $f$ is even). By choosing the integer $m$ suitably the function $f(z)\wp^m(z)$ is an even elliptic function which is analytic and non–zero at all lattice points and is, therefore, expressible in the form $(*)$ (the integer $m$ is allowed to be negative).

If $f$ is an odd function then $\frac{f(z)}{\wp'(z)}$ is even and therefore rationally expressible in terms of $\wp(z)$. Finally we observe that any function can be written in the form

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

The first term on the right hand side of this equation is even, the second term is odd. Hence $f(z)$ can be put into the form

$$f(z) = R_1\{\wp(z)\} + \wp'(z)R_2\{\wp(z)\},$$
where $R_1, R_2$ are rational functions of their argument $\varphi(z)$. This is the result stated in the theorem.

Much, much more can be said about elliptic functions. A very important family of elliptic functions was defined by Jacobi, which mirror many of the properties of the trigonometric functions. These functions satisfy different differential equations; ones of the form

$$(y')^2 = Ay^4 + By^3 + Cy^2 + Dy.$$ 

and

$$y' = Ay^3 + By.$$ 

The equation

$$Y^2 = 4X^3 - g_2X - g_3$$

is an example of an elliptic curve, considered here with complex variables $X$ and $Y$. In algebraic geometry and differential geometry, an important concept is the complex projective plane $\mathbb{CP}^2 = \mathbb{C}^3 \setminus \{(0, 0, 0)\}$ modulo the equivalence relation $(z_1, z_2, z_3) \sim (\lambda z_1, \lambda z_2, \lambda z_3)$ for all $\lambda \neq 0$. Recall that the Weierstrass $\wp$ function is naturally defined on the torus $T$ determined by joining the appropriate edges of a cell, and that the cell/torus is a group under addition modulo the lattice $\Lambda$. The function $\wp$ gives an embedding of the torus into the complex projective plane via

$$z \mapsto (1, \varphi(z), \varphi'(z))$$

which turns out to be a group homomorphism from $T$ to $\mathbb{CP}^2$, and an isomorphism of Riemann surfaces. What this means is that topologically all complex elliptic curves are a torus, and they all have a group structure that is matched to their algebra.

From here we could head into a discussion of modular forms: an analytic function $f : H = \{\text{Im } z > 0\} \to \mathbb{C}$ such that for all $a, b, c, d \in \mathbb{Z}$ such that $ad - bc = 1$,

$$f \left( \frac{az + b}{cz + d} \right) = (cz + d)^k f(z), \quad z \in H,$$

and such that $f$ is analytic at $\infty$. Actually it is sufficient to check that $f(-1/z) = z^k f(z)$ and $f(z + 1) = f(z)$.

There is a close link between modular forms and elliptic curves over $\mathbb{Q}$. This was an important part of Wiles’ proof of Fermat’s Last Theorem$^5$.

---

$^4$It is worth noting that $\mathbb{CP}^1$, the complex projective line is just the Riemann sphere.

$^5$Alas, I have run out of room to include the proof.