RIEMANN SURFACES

ALESSANDRO GIACCHETTO
ETH ZÜRICH – SPRING SEMESTER 2024

Date: October 22, 2024.

Contents

The course will be a first introduction to Riemann surfaces. These are beautiful objects that sit at the intersection of algebraic geometry, differential geometry, and analysis. We will aim to cover the theorems of Riemann–Hurwitz and Riemann–Roch, as well as the basics of Hurwitz theory.

References:

• E. Arbarello, M. Cornalba, P. A. Griffiths, J. Harris. *Geometry of Algebraic Curves, Volume 1*. Springer-Verlag, 1985

• R. Cavalieri, E. Miles. *Riemann Surfaces and Algebraic Curves*. Cambridge University Press, 2016

• O. Forster. *Lectures on Riemann Surfaces*. Springer-Verlag, 1981

1. Background

Lecture 1
Feb 22nd, 2024

1.1. Complex analysis. How do you calculate functions like e^x, $\sin(x)$, or $\cos(x)$ at any given x_0 value? This is a very old and classic question that dates back to the 14th century. The basic idea is to approximate a smooth function near an input value by the value of the function and its derivatives:

$$f(x) \approx \sum_{n=0}^{N} \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \,. \tag{1.1}$$

The right-hand side, known as the Taylor polynomial, is a polynomial that well approximates the function f around x_0. The higher the degree of the Taylor polynomial, the more derivative we incorporate into it, yielding a better and better approximation of the original function (see figure 1). Eventually, in the limit, we get what is known as the Taylor series of f around x_0.

For functions such as the exponential near $x_0 = 0$, the Taylor series does not just approximate the function, but actually equals it for all values of x for which the Taylor series converges. And for the exponential, it just so happens that the Taylor series converges for all possible values of x, meaning e^x precisely equals its Taylor series for all values of x both near and far away from zero:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \,. \tag{1.2}$$

While it is true that the Taylor polynomial will approximate the behaviour of e^x near zero to some extent, and will further improve as we use higher degree polynomials, there is no reason to assume beforehand that the approximation will improve to arbitrary precision as we keep going and ultimately reach perfect equality in the limit. The reason why not is actually simple: these polynomial approximations were constructed only using *local* information about the function at x_0, and yet somehow, it is enough to reproduce the entire function everywhere else.

Being equal to its own Taylor series, even in a small neighbourhood of the expansion point, is a peculiar feature of the function e^x. There are smooth functions whose Taylor expansions around a point are not equal to the function on any neighbourhood surrounding the expansion point, no

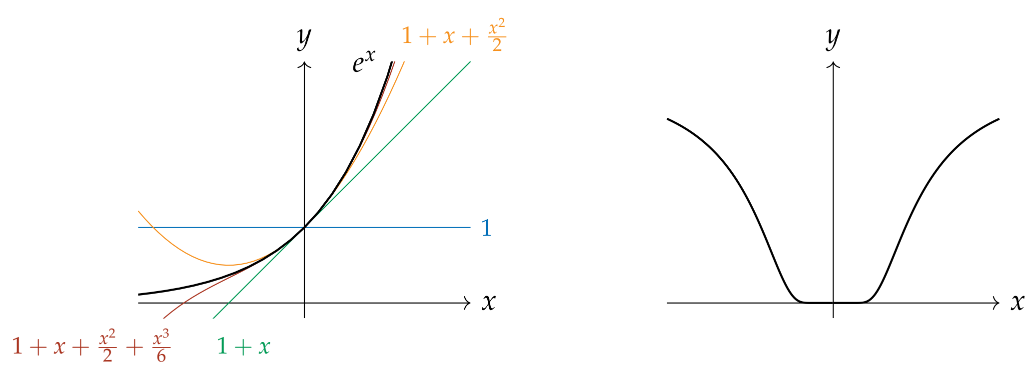


Figure 1. The exponential function e^x and its first few Taylor series (left), and the piecewise function $g(x)$ (right).

matter how small. A classic example is the piecewise function given by

$$g(x) := \begin{cases} e^{-\frac{1}{x^2}} & \text{for } x \neq 0\,, \\ 0 & \text{for } x = 0\,. \end{cases} \tag{1.3}$$

This function and all its derivatives vanish at $x_0 = 0$, implying that its Taylor polynomials and its Taylor series at $x_0 = 0$ are simply the constant function zero. However, for any non-zero value of x, the function g is non-zero, i.e. it does not match its Taylor expansion anywhere on $\mathbb{R} \setminus \{ 0 \}$ (see figure 1).

The essential element to understand the difference between these these two cases, the exponential function and the function g, is Taylor's theorem. Taylor's theorem expresses the error committed in approximating a function by its Taylor polynomial of a given order n:

$$\underbrace{f(x)}_{\substack{\text{original}\\\text{function}}} - \underbrace{\sum_{n=0}^{N} \frac{f^{(n)}(x_0)}{n!}\,(x-x_0)^n}_{\text{Taylor polynomial}} = \underbrace{\frac{f^{(N+1)}(\xi)}{(N+1)!}(x-x_0)^{N+1}}_{\text{error term}} \tag{1.4}$$

for some $\xi \in [x_0, x]$.

Let us go back to the exponential function, with expansion point $x_0 = 0$. In this case, we have $f^{(N+1)}(\xi) = e^{\xi}$ for some $\xi \in [x_0, x]$, which can be bounded independently of n. Thus, as we increase the order of the Taylor approximation, the entire remainder approaches zero since the $(N+1)!$ in the denominator will overpower the geometric factor of $(x - x_0)^{N+1}$ in the numerator. The presence of e^t cannot prevent the convergence, as it is confined in $[1, e^x]$. We thus conclude that the Taylor series converges exactly to the original function e^x for all values of x.

The key to why the exponential function, and also functions like sine and cosine, equals its Taylor series is because their N-th order derivatives grow much slower than $N!$. In fact, their higher order derivatives do not grow at all: they stay bounded as n increases. This is not the case with our pathological example $g(x)$, whose derivatives near $x_0 = 0$ grow faster than factorial speed as the derivative order increases. Now functions like e^x, $\sin(x)$ and $\cos(x)$ that equal their Taylor

n	0	1	2	3	4	5	6
$g^{(n)}(\frac{1}{4})$	$1.12 \cdot 10^{-7}$	$1.44 \cdot 10^{-5}$	$1.67 \cdot 10^{-3}$	$1.72 \cdot 10^{-1}$	15.37	1128.01	61284.06
$n!$	1	1	2	6	24	120	720

expansions near the expansion point have a special name: they are called *analytic functions*. This is a rare and special property, yet almost all functions you are likely to work with on a regular basis are analytic. Part of the reason is that any arithmetic combination or function composition of analytic functions results in another analytic function. By contrast, the piece-wise function $g(x)$ is not analytic, albeit being smooth (that is, even though it has derivatives of all orders at all points in its domain). This implies that smoothness does not imply analyticity. However, there is a world where non-analytic smooth functions can never appear: the world of complex numbers.

1.1.1. *Holomorphic functions.* To start with, notice that the concept of series makes perfect sense in the complex setting. What about the concept of differentiable function? In the complex setting, the concept of differentiable function deserves a new name: holomorphicity.

Definition 1.1. Let $U \subseteq \mathbb{C}$ be an open set. A function $f\colon U \to \mathbb{C}$ is called holomorphic at $z \in U$ if the following limit exists:

$$\lim_{h\to 0} \frac{f(z+h)-f(z)}{h}\,. \tag{1.5}$$

Here $h \in \mathbb{C}$ (and such that $z+h \in U$). We write $f'(z)$ for its value. The function f is said to be holomorphic in U if it is holomorphic at every point $z \in U$.

This is similar to the definition of differentiable functions in the real setting. However the condition of being holomorphic is much stronger. Indeed, the following result holds. As we will see later, holomorphic functions are automatically smooth and analytic. This is one example of the great irony that complex numbers make everything simpler.

In order to prove the analyticity property, let us recall some other useful facts about holomorphic functions. Notice that, in the limit defining $f'(z)$, we can choose the displacement h to be purely real or purely imaginary. This not only proves that the real and imaginary components of f are differentiable, but also implies that a specific set of partial differential equations hold: the Cauchy–Riemann equations.

Theorem 1.2 (Cauchy–Riemann equations)**.** *Let f be a holomorphic function on U. Write $f(z) = u(x,y) + \mathrm{i}v(x,y)$ for $z = x + \mathrm{i}y$, considered as a real differentiable function on $\mathbb{R}^2$. Then*

$$u_x = v_y\,, \qquad\qquad v_x = -u_y\,. \tag{1.6}$$

Exercise 1.1. *Prove the Cauchy–Riemann equations.*

Corollary 1.3. *Let f be a non-constant holomorphic function. As a function from the real plane to itself, f is orientation-preserving.*

Sketch of the proof. Intuitively, an orientation of the plane amounts to specifying the notions of clockwise and counter-clockwise. A function is orientation-preserving if it preserves such a notion. A way of characterizing orientation-preserving functions is by looking at the sign of the determinant of the Jacobian matrix. The function is orientation-preserving if the determinant is positive on a dense open set. For holomorphic functions, by Cauchy–Riemann, we find

$$\det\left(\mathrm{Jac}_f\right) = \det\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = u_x^2 + v_x^2\,, \tag{1.7}$$

which is indeed positive on an open dense set (otherwise, f would be constant). □

Another important property of non-constant holomorphic functions is that they send open sets to open sets.

Theorem 1.4 (Open mapping theorem)**.** *Let f be a non-constant holomorphic function. Then f is an open map.*

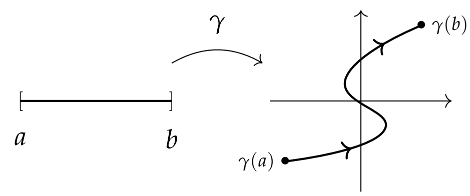


FIGURE 2. Example of a path in $\mathbb{C}$.

1.1.2. *Cauchy's formula and applications.* Given a smooth path $\gamma\colon [a,b] \to U \subseteq \mathbb{C}$ and a holomorphic function $f\colon U \to \mathbb{C}$, we define the integral of f along γ as

$$\int_\gamma f(z)\, dz := \int_a^b f(\gamma(t))\gamma'(t)\, dt\,. \tag{1.8}$$

Exercise 1.2. *Let γ_r be the counter-clock wise oriented circle of radius r centred at the origin. Prove that, for $n \in \mathbb{Z}$, $\frac{1}{2\pi \mathrm{i}} \oint_{\gamma_r} z^{n-1}\, dz = \delta_{n,0}$.*

The remarkable property of holomorphic functions is the invariance of the integral along deformations of the path.

Theorem 1.5. *Let $\gamma_0, \gamma_1\colon [a,b] \to U$ be two paths related by a smooth deformation keeping the end-points fixed. That is, there exists a smooth function $H\colon [a,b] \times [0,1] \to U$ such that $H(\cdot,0) = \gamma_0$, $H(\cdot,1) = \gamma_1$, and $H(a,\cdot) = z_a$, $H(b,\cdot) = z_b$. Then*

$$\int_{\gamma_0} f(z)\, dz = \int_{\gamma_1} f(z)\, dz\,. \tag{1.9}$$

Proof. For $s \in [0,1]$, let γ_s be the path defined by $H(\cdot,s)$. Set $I(s) := \int_{\gamma_s} f(z) dz$. A simple computation shows that

$$\frac{dI(s)}{ds} = \left[f(H(t,s)) \frac{\partial H(t,s)}{\partial s} \right]_{t=a}^{b}. \tag{1.10}$$

Since $H(a,\cdot)$ and $H(b,\cdot)$ are constant, $\frac{dI(s)}{ds}$ is constantly zero. In particular, $I(0) = I(1)$. □

A simple consequence of the above theorem is that the integral of homomorphic functions along a loop (i.e. a closed path) within a simply connected domain of holomorphicity vanishes. This is because, in this case, we can deform the loop to a point, for which the integral vanishes. Another fundamental consequence is the Cauchy's integral formula.

Theorem 1.6 (Cauchy's integral formula)**.** *Let γ be a counter-clock wise oriented loop around $z \in \mathbb{C}$, and f a holomorphic function on a neighbourhood U of γ. Then*

$$f(z) = \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{f(\zeta)}{\zeta - z} d\zeta\,. \tag{1.11}$$

Sketch of the proof. Let $g(\zeta, z) = \frac{f(\zeta)-f(z)}{\zeta - z}$. As a function of ζ, g is holomorphic in U. Indeed, the only problematic point is $\zeta = z$, for which the value is given by $f'(z)$. Thus, we conclude that $\frac{1}{2\pi \mathrm{i}} \oint_\gamma g(\zeta, z)\, d\zeta = 0$. On the other hand, exercise 1.2 implies that $\frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{f(z)}{\zeta - z} d\zeta = f(z)$. All together, we find the thesis. □

As promised, we can finally prove that holomorphic functions are analytic:

$$
\begin{aligned}
f(z) &= \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{f(\zeta)}{\zeta - z_0 + z_0 - z} d\zeta = \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{f(\zeta)}{\zeta - z_0} \frac{1}{1 - \frac{z - z_0}{\zeta - z_0}} d\zeta \\
&= \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{f(\zeta)}{\zeta - z_0} \left(\sum_{n=0}^{\infty} \left(\frac{z - z_0}{\zeta - z_0} \right)^n \right) d\zeta = \sum_{n=0}^{\infty} \left(\frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{f(\zeta)}{(\zeta - z_0)^{n+1}} d\zeta \right) (z - z_0)^n . \quad (1.12)
\end{aligned}
$$

Cauchy's integral formula not only gives a proof of the analyticity of holomorphic functions, but also provides an integral formula for their n-th derivative:

$$
f^{(n)}(z) = \frac{n!}{2\pi \mathrm{i}} \oint_\gamma \frac{f(\zeta)}{(\zeta - z)^{n+1}} d\zeta . \quad (1.13)
$$

Notice that the reason why Cauchy's formula produces something non-trivial is that we introduced a singularity at $\zeta = z$ in the integrand. This motivates the study of singularities of functions in the complex plane.

Definition 1.7. Given a positive integer k, we say that a complex function f has a pole of order k at the point $z_0 \in \mathbb{C}$ if $(z - z_0)^k f(z)$ is holomorphic at z_0 but $(z - z_0)^{k-1} f(z)$ is not. A function is called meromorphic on an open set $U \subseteq \mathbb{C}$ if for every point $z_0 \in U$, f is either holomorphic or it has a pole at z_0.

A simple consequence of the analyticity result is that a function f with a pole of order k at the point $z_0 \in \mathbb{C}$ admits an expansion of the form

$$
f(z) = \sum_{n=-k}^{\infty} a_n (z - z_0)^n \quad (1.14)
$$

with $a_{-k} \neq 0$. Series of this form are called Laurent series. An important role is played by the coefficient of $(z - z_0)^{-1}$.

Definition 1.8. Let f be a meromorphic function in U. Define the residue at $z_0 \in U$ as the coefficient of $(z - z_0)^{-1}$ in the Laurent expansion of f at z_0. We denote it as $\mathrm{Res}_{z=z_0} f(z)$.

Exercise 1.3. *Show that if f has a pole of order 1 at z_0 (also called a simple pole), then the residue can be computed as*

$$
\operatorname*{Res}_{z=z_0} f(z) = \lim_{z \to z_0} (z - z_0) f(z) . \quad (1.15)
$$

Exercise 1.4 (Residue theorem)**.** *Let γ be a counter-clock wise oriented loop in U that does not cross itself and contains the points $z_1, \ldots, z_N$. Let f be a holomorphic function in $U \setminus \{z_1, \ldots, z_N\}$ with poles*

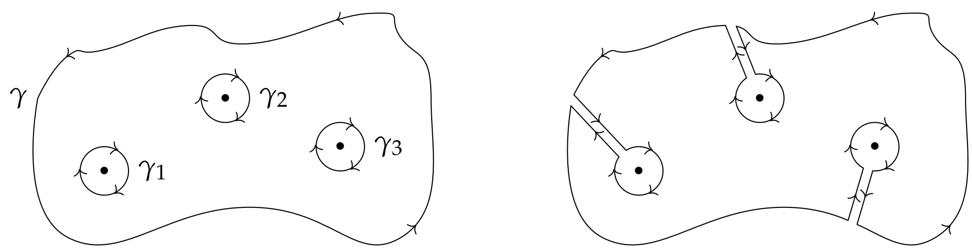


FIGURE 3. Contour deformation in the proof of the residue theorem.

at $z_1, \ldots, z_N$. Then

$$\frac{1}{2\pi \mathrm{i}} \oint_\gamma f(z)\, dz = \sum_{i=1}^{N} \operatorname*{Res}_{z=z_i} f(z)\,. \tag{1.16}$$

Hint. *Consider (small enough) counter-clock wise oriented loops around the poles, $\gamma_1, \ldots, \gamma_N$, and the deformation of the contour $\gamma \cup \gamma_1 \cup \cdots \cup \gamma_N$ depicted in figure 3.*

Exercise 1.5 (Basel problem). *Prove the identity $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{2}$. The left-hand side is the Riemann zeta function $\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$ evaluated at $s = 2$. Can you use the same strategy to prove the following equation?*

$$\zeta(2k) = (-1)^{k+1} \frac{B_{2k}(2\pi)^{2k}}{2(2k)!}\,, \tag{1.17}$$

where B_m is the m-th Bernoulli number. What is wrong with the odd values of the zeta function?

Hint. *Consider the function $f(z) = \frac{\pi}{z^2} \cot(\pi z)$ and the residue theorem.*

Lecture 2
Feb 29th, 2024

Another important application of Cauchy's formula is the inverse mapping theorem, which asserts that holomorphic functions with non-vanishing derivative are locally invertible.

Theorem 1.9 (Inverse mapping theorem)**.** *Let f be a holomorphic function in U, and $z_0 \in U$ such that $f'(z_0) \neq 0$. There exists a neighbourhood V of $f(z_0)$ and a holomorphic function g in V such that $g \circ f = \mathrm{id}$ on $g(V)$.*

Proof. Without loss of generality, we suppose that U is small enough such that f is injective and f' is non-vanishing in U. Since holomorphic functions are open, there exists a ball B centred at $f(z_0)$ and contained in $f(U)$. We take $V = B$, so that f restricted to $f^{-1}(V)$ is invertible (in the set-theoretic sense). To conclude the prove, we need to show that inverse function is holomorphic in V. Let γ be the counter-clock wise oriented loop parametrising the perimeter of B. We claim that

$$g(w) := \frac{1}{2\pi \mathrm{i}} \oint_\gamma \frac{\zeta f'(\zeta)}{f(\zeta) - w} d\zeta \tag{1.18}$$

is the desired inverse. This would conclude the proof, since the right-hand side is holomorphic. In order to prove the claim, let $z \in f^{-1}(V)$, with $w = f(z)$. The integrand in the definition of g

has a simple pole at $\zeta = z$. Thanks to the residue theorem (cf. exercise 1.3), we find

$$g(w) = \operatorname*{Res}_{\zeta=z} \frac{\zeta f'(\zeta)}{f(\zeta)-w} = \lim_{\zeta\to z} (\zeta - z)\frac{\zeta f'(\zeta)}{f(\zeta)-w}\,. \tag{1.19}$$

As $\lim_{\zeta\to z} \frac{\zeta - z}{f(\zeta)-w} = \frac{1}{f'(z)}$, we find the thesis. □

1.1.3. *Towards Riemann surfaces: the root function.* Consider the function $f(z) = z^2$. The function is not injective since both z and $-z$ have the same image. However, according to the inverse mapping theorem, the function f is locally invertible around every point z_0. Such a local inverse function is called a branch of the square-root function $w^{1/2}$. A natural question is: how large can the domain of such a branch be?

Intuitively, the most we can hope for is the complex plane minus the origin (i.e. all points where f' is non-vanishing). However, this is not possible, as can be easily seen by the following. Take $z_0 = 1$, and let $g(w)$ be the local inverse of $f(z)$ around z_0. Writing a complex number in polar coordinates as $w = re^{2\pi i\theta}$, we see that an expression for g is given by $g(w) = \sqrt{r}e^{\pi i\theta}$. If we suppose for a moment that g is well-defined on the whole $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$, we would have that a full circle around the origin would bring us back to the same point. For instance, taking $\gamma\colon [0,1) \to \mathbb{C}$ with $\gamma(t) = e^{2\pi i t}$, the unit circle, we would have

$$\lim_{t\to 1} g\big(\gamma(t)\big) = g\big(\gamma(0)\big)\,. \tag{1.20}$$

But this is not the case, since the left-hand side is -1, while the right-hand side is $+1$. Hence, a branch of the square-root function may be extended (as a holomorphic function) to any domain $U \subset \mathbb{C}^\times$ such that it is not possible to walk around the origin in U. Typical examples of maximal domains of definitions consist of the complex plane minus a real half-line stemming from the origin. We can actually extend the domain of the root function beyond the complex plane minus a real half-line by glueing two copies of such maximal domain on top of each other, as shown in figure 4.

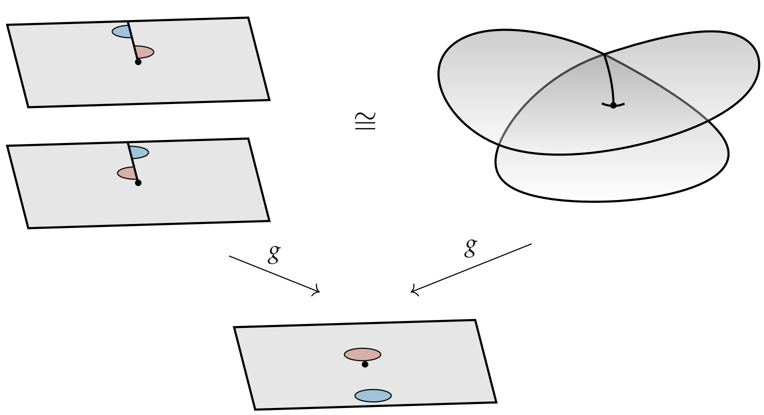


FIGURE 4. The domain of the square-root function.

A more formal approach, due to to Riemann, consists in taking as domain of $w^{1/2}$ the graph of the function z^2:

$$X := \{ (z,w) \in \mathbb{C}^\times \times \mathbb{C}^\times \mid w = z^2 \} \,. \tag{1.21}$$

For any given point $x = (z,w) \in X$, we can use w as a local coordinate for X, and composing with the projection onto the second factor gives a branch of the square-root.

Before Riemann, the square-root function (an more generally, the "inverse" of any power $z \mapsto z^r$) was traditionally viewed as a multi-valued function, implying the existence of several distinct values of z for any given $w \neq 0$. Riemann, however, brought about a shift in perspective by taking objects such as X as the domain. The space X, which captures all possible branches of the square-root without making any choice of domain restriction, is called the Riemann surface of the square-root.

To conclude this section, we show how root functions describe the local geometry of holomorphic functions.

Lemma 1.10 (Local form of holomorphic function)**.** *Let $w = f(z)$ be a holomorphic function on U and $z_0 \in U$ such that $f'(z_0) = f''(z_0) = \cdots = f^{k-1}(z_0) = 0$, and $f^{(k)}(z_0) \neq 0$. Then there exist holomorphic changes of variable $(z,w) \mapsto (\tilde{z},\tilde{w})$ such that in the new variables f is given by $\tilde{w} = \tilde{z}^k$.*

Proof. The Taylor series of f at z_0 reads

$$f(z) - f(z_0) = \sum_{n=k}^{\infty} a_n (z - z_0)^n \tag{1.22}$$

with $a_k \neq 0$. Hence, the function $g(z) := \sum_{n=k}^{\infty} a_n (z-z_0)^{n-k}$ is holomorphic with $g(z_0) \neq 0$. In particular, it admits a branch of the k-th root, that we denote as $\sqrt[k]{g(z)}$. Setting $\tilde{z} = (z - z_0)\sqrt[k]{g(z)}$, it is easy to check that $\tilde{z}$ has non-vanishing derivative at z_0, thus locally invertible by the inverse mapping theorem. In other words, $z \mapsto \tilde{z}$ is a legitimate change of variable. Furthermore, setting $\tilde{w} = w - f(z_0)$ (which is again a legitimate change of variable), we find that $\tilde{w} = \tilde{z}^k$. ☐

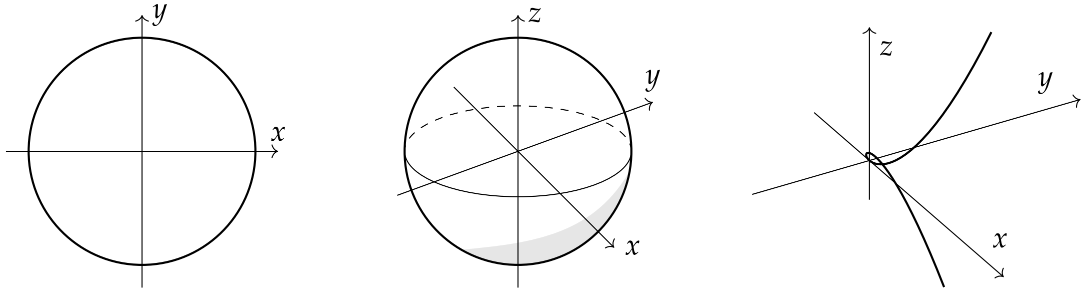
1.2. **Manifold theory.** In the previous section we understood how the domain of the square-root function is a complicated object obtained by patching together well-understood geometries, namely open subsets of $\mathbb{C}$. This concept extends to the idea of a manifold: a space whose overall geometry may be intricate, yet the local geometry around each point remains familiar.

1.2.1. *Manifolds: definition and examples.* The concept of a manifold is one of the fundamental insights of 19th century mathematics. Before that, geometric objects existed mostly only extrinsically, i.e., as subsets of the affine space $\mathbb{R}^N$. Examples include:

- The circle:
$$S^1 := \{ (x,y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \} \subset \mathbb{R}^2 \,. \tag{1.23}$$
- The sphere:
$$S^2 := \{ (x,y,z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^3 = 1 \} \subset \mathbb{R}^3 \,. \tag{1.24}$$

- Parametric curves, such as:

$$C := \{ (t, t^2, t^3) \in \mathbb{R}^3 \mid t \in \mathbb{R} \} \subset \mathbb{R}^3 . \tag{1.25}$$



(A) The circle $S^1 \subset \mathbb{R}^2$. (B) The sphere $S^2 \subset \mathbb{R}^3$. (C) The curve $C \subset \mathbb{R}^3$.

FIGURE 5. Examples of manifolds embedded in their ambient space.

However, Riemann realized that geometric objects can exist intrinsically, that is, without the need for an ambient space. Informally, a smooth manifold is a topological space that locally looks like the affine space $\mathbb{R}^n$ (see figure 6).

Definition 1.11. A real smooth manifold of dimension n is a second-countable and Hausdorff topological space X together with an open cover $X = \bigcup_{i \in I} U_i$ such that

- there exist homeomorphisms, called local charts, $\varphi_i \colon U_i \to V_i$ with $V_i \subset \mathbb{R}^n$ open;
- for all U_i, U_j with $U_i \cap U_j \neq \varnothing$, the transition map

$$\varphi_{i,j} := \varphi_j \circ \varphi_i^{-1} \colon \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j) \tag{1.26}$$

  is smooth.

The collection $\{ (U_i, \varphi_i) \}_{i \in I}$ is called an atlas.

Second countable and Hausdorff are point-set conditions that exclude pathological examples. Second countable rules out spaces considered ’too large’, like “the long line”, while Hausdorff eliminates spaces exhibiting peculiarities such as “the line with two origins”.

Example 1.12. Consider $X = S^1 = \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1 \}$, which is naturally equipped with a second-countable Hausdorff topology. Let

$$\begin{aligned} U_{\mathrm{N}} &:= \left\{ (x,y) \in S^1 \,\middle|\, y > 0 \right\}, & U_{\mathrm{S}} &:= \left\{ (x,y) \in S^1 \,\middle|\, y < 0 \right\}, \\ U_{\mathrm{E}} &:= \left\{ (x,y) \in S^1 \,\middle|\, x > 0 \right\}, & U_{\mathrm{W}} &:= \left\{ (x,y) \in S^1 \,\middle|\, x < 0 \right\}, \end{aligned} \tag{1.27}$$

which form an open cover of S^1.

Take $V := (-1, 1) \subset \mathbb{R}$, and define $\varphi_{\mathrm{N}} \colon U_{\mathrm{N}} \to V$ and $\varphi_{\mathrm{S}} \colon U_{\mathrm{S}} \to V$ as the projections to the x-axis, and likewise $\varphi_{\mathrm{E}} \colon U_{\mathrm{E}} \to V$ and $\varphi_{\mathrm{W}} \colon U_{\mathrm{W}} \to V$ as the projections to the y-axis (see figure 7). Notice

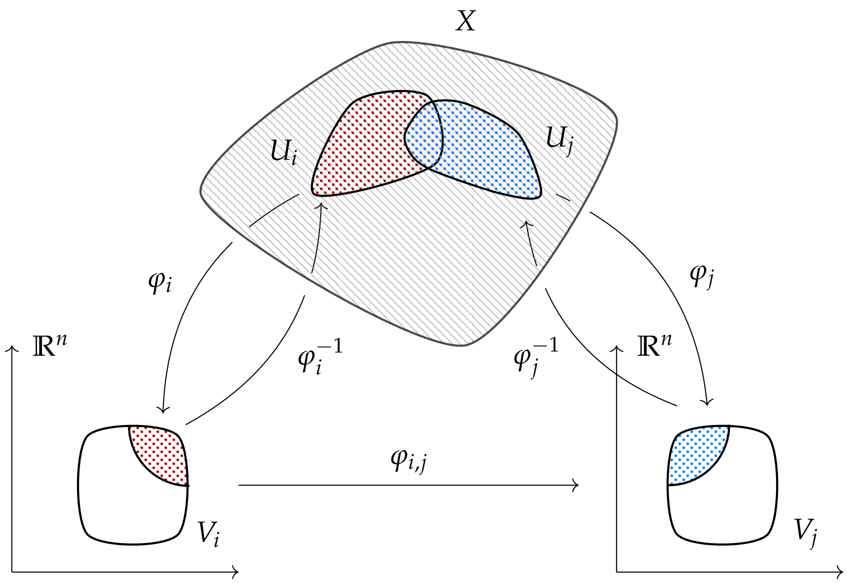


FIGURE 6. Charts and transition functions on a manifold.

that φ_* is a homeomorphism for all $* \in \{\mathrm{N}, \mathrm{S}, \mathrm{E}, \mathrm{W}\}$. For instance, $\varphi_{\mathrm{N}}^{-1}(x) = (x, \sqrt{1-x^2})$. Furthermore, the transition maps are smooth. For instance,

$$U_{\mathrm{NE}} := U_{\mathrm{N}} \cap U_{\mathrm{E}} = \left\{ (x,y) \in S^1 \;\middle|\; x > 0\,,\ y > 0 \right\}. \tag{1.28}$$

Then $\varphi_{\mathrm{NE}} \colon (0,1) \to (0,1)$ is the function $x \mapsto \sqrt{1-x^2}$, which is indeed smooth.

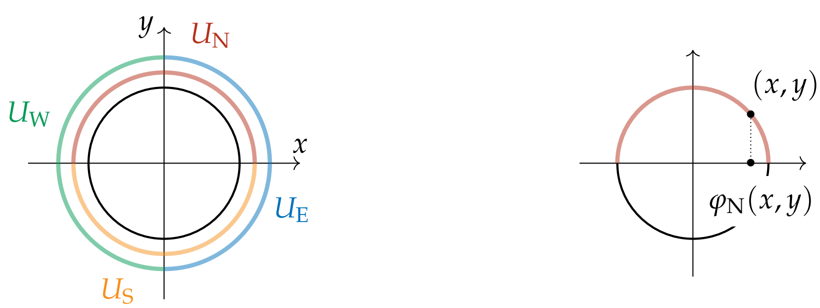


FIGURE 7. The atlas for the circle provided by the axes projections.

A priori, we could have chosen a different atlas for the circle. An example is the atlas determined by the stereographic projection.

Exercise 1.6. *On the circle S^1, let $\mathrm{N} := (0,1)$ and $\mathrm{S} := (0,-1)$ be the north and south poles respectively. Define $U'_{\mathrm{N}} := S^1 \setminus \mathrm{N}$ and the map $\varphi'_{\mathrm{N}} \colon U'_{\mathrm{N}} \to \mathbb{R}$ by declaring that $\varphi'_{\mathrm{N}}(x,y)$ is the unique intersection between the x-axis and the line passing through (x,y) and N (see figure 8). Similarly for $U'_{\mathrm{S}} := S^1 \setminus \mathrm{S}$ and the map $\varphi'_{\mathrm{S}} \colon U'_{\mathrm{S}} \to \mathbb{R}$.*

- *Find a formula for φ'_{N} and φ'_{S}.*

- *Prove that φ'_{N} and φ'_{S} are homeomorphisms.*
- *Prove that transition function φ'_{NS} is smooth.*

FIGURE 8. The atlas for the circle provided by the stereographic projections.

To convey the concept that different atlases can yield the same manifold structure, we define compatibility between two atlases $\mathcal{A} = \{ (U_i, \varphi_i) \}$ and $\mathcal{A}' = \{ (U'_j, \varphi'_j) \}$ of X if their union, $\mathcal{A} \cup \mathcal{A}'$, forms an atlas for X. Compatibility is an equivalence relation among atlases of X, and a (real smooth, n-dimensional) manifold structure is an equivalence class of compatible atlases. An example of compatible atlases is the axes projections and the stereographic projections on the circle.

An example of 'intrinsic manifold' is given by the projective space, which formalise the idea of parametrising lines in space.

Example 1.13. The real projective space of dimension n is defined as

$$P^n(\mathbb{R}) := \Big\{ (x_0, \dots, x_n) \in \mathbb{R}^{n+1} \setminus \{0\} \Big\} \Big/ \sim \tag{1.29}$$

where $(x_0, \dots, x_n) \sim (x'_0, \dots, x'_n)$ if and only if $\exists \lambda \in \mathbb{R}^\times$ such that $(x_0, \dots, x_n) = \lambda(x'_0, \dots, x'_n)$. Denote equivalence classes of points as $[x_0 : \cdots : x_n]$. The space $P^n(\mathbb{R})$ is naturally equipped with a second-countable and Hausdorff topology[1]. What are the charts of $P^n(\mathbb{R})$?

Consider the subsets $U_i \subset P^n(\mathbb{R})$ and the maps $\varphi_i \colon U_i \to \mathbb{R}^n$ defined as

$$U_i := \{ [x_0 : \cdots : x_n] \in P^n(\mathbb{R}) \mid x_i \neq 0 \} , \qquad \varphi_i \colon [x_0 : \cdots : x_n] \longmapsto \frac{1}{x_i}(x_0, \dots, \widehat{x_i}, \dots, x_n) , \tag{1.30}$$

from U_i to the affine space $\mathbb{R}^n$ with coordinates $(y_{i,0}, \dots, \widehat{y_{i,i}}, \dots, y_{i,n})$. In other words, the charts put on each U_i the coordinate $y_{i,a} = \frac{x_a}{x_i}$. The U_i's form an open cover of $P^n(\mathbb{R})$, and the charts are homeomorphisms from U_i to $\mathbb{R}^n$. Furthermore, at the intersection

$$U_i \cap U_j = \big\{ [x_0 : \cdots : x_n] \in P^n(\mathbb{R}) \bigm| x_i \neq 0 \text{ and } x_j \neq 0 \big\} , \tag{1.31}$$

[1] $P^n(\mathbb{R})$ is equipped with the quotient topology defined through the projection map $\pi \colon \mathbb{R}^{n+1} \setminus \{0\} \to P^n(\mathbb{R})$. That is, a set $U \subseteq P^n(\mathbb{R})$ is open if and only if $\pi^{-1}(U) \subseteq \mathbb{R}^{n+1} \setminus \{0\}$ is open.

we have coordinates

$$\begin{aligned}\varphi_i(U_i \cap U_j) &= \left\{ \, (y_{i,0}, \ldots, \widehat{y_{i,i}}, \ldots, y_{i,n}) \in \mathbb{R}^n \;\middle|\; y_{i,j} \neq 0 \, \right\} , \\ \varphi_j(U_i \cap U_j) &= \left\{ \, (y_{j,i}, \ldots, \widehat{y_{j,j}}, \ldots, y_{j,n}) \in \mathbb{R}^n \;\middle|\; y_{j,i} \neq 0 \, \right\} ,\end{aligned} \tag{1.32}$$

and transition maps $\varphi_{i,j}$ given by

$$\varphi_{i,j} \colon (y_{i,0}, \ldots, \widehat{y_{i,i}}, \ldots, y_{i,n}) \longmapsto (y_{j,0}, \ldots, \widehat{y_{j,j}}, \ldots, y_{j,n}) = \frac{1}{y_{i,j}} (y_{i,0}, \ldots, \widehat{y_{i,i}}, \ldots, y_{i,n}) \,. \tag{1.33}$$

This is a consequence of the relation $y_{j,a} = \frac{x_a}{x_i} = \frac{x_a}{x_i} \cdot \frac{x_i}{x_j} = \frac{y_{i,a}}{y_{i,j}}$. In other words, the transition map is the multiplication by $1/y_{i,j}$, which is a smooth function in the domain of definition.

Exercise 1.7. *Consider an open set $U \subseteq \mathbb{R}^n$ and a smooth function $f \colon U \to \mathbb{R}^m$. Show that the graph of f, that is*

$$\Gamma_f = \{ \, (x, y) \in U \times \mathbb{R}^m \mid y = f(x) \, \} \tag{1.34}$$

is a smooth manifold of dimension n.

A natural approach to creating interesting manifolds within subsets of Euclidean space involves examining level sets of functions. For instance, considering the function $f(x, y) = x^2 + y^2$, we can view the unit circle as the level set $f^{-1}(1)$. Another illustration is provided by the function $g(x, y) = xy$ and its level set $g^{-1}(1)$, which forms a hyperbola. However, the level set $g^{-1}(0)$ is the union of the x- and y-axes, and it cannot be endowed with the structure of a smooth manifold. This is because no neighbourhood of the origin can be homeomorphic to an open set in $\mathbb{R}$.

The question of which level sets exhibit the structure of a manifold finds its solution in this classical theorem from analysis.

Theorem 1.14 (Implicit function theorem)**.** *Let $F \colon \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function, and $x \in \mathbb{R}^n$ such that the Jacobian matrix $\mathrm{Jac}_F(x)$ is a surjective linear function. Say $F(x) = c$. Then there exist:*

- *a neighbourhood $A_x \subseteq \mathbb{R}^n$ of x,*
- *an open set $V_x \subseteq \mathbb{R}^{n-k}$ and a smooth function $f_x \colon V_x \to \mathbb{R}^m$,*

such that $F^{-1}(c)$ is locally in A_x the graph Γ_{f_x} of the function f_x. In other words, $F^{-1}(c) \cap A_x = \Gamma_{f_x}$.

The implicit function theorem conveys a fundamentally intuitive idea: at a point $x \in \mathbb{R}^n$, if there exist m coordinates where the determinant of the matrix of corresponding partial derivatives is non-zero, then locally around x you can select the remaining $n - m$ coordinates as local coordinates for the level set of F passing through x. The projection onto these coordinates establishes a local chart for the level set in the proximity of x. Keeping this in mind, the subsequent result appears inherently natural.

Theorem 1.15. *Let $F \colon \mathbb{R}^n \to \mathbb{R}^m$ be a smooth function and $c \in \mathbb{R}^m$ a regular value for F. That is, for every point $x \in F^{-1}(c)$, the Jacobian matrix $\mathrm{Jac}_F(x)$ is a surjective linear function. Then $F^{-1}(c)$ is a real smooth manifold of dimension $n - m$.*

Sketch of the proof. Firstly, observe that $X := F^{-1}(c)$ is second-countable and Hausdorff since it is a subset of $\mathbb{R}^n$ with the induced topology.

Secondly, applying the implicit function theorem, we conclude that X is locally $\mathbb{R}^{n-m}$. More precisely, we define an atlas as follows: for each point $x \in X$, we find a neighbourhood $U_x = X \cap A_x$ of x in the level set. We also define the map $\varphi_x \colon U_x \to V_x$ as follows. By the implicit function theorem, every point in U_x is of the form $(p, f_x(p))$ for some $p \in V_x$; we set $\varphi_x(p, f(p)) = p$. In other words, φ_x is the projection $\pi \colon \mathbb{R}^n \to \mathbb{R}^{n-m}$. Then, φ_x is a homeomorphism with inverse $\varphi_x^{-1}(p) = (p, f_x(p))$.

To conclude, we must show that the transition maps are smooth. Given two points $x, x' \in X$ such that $U_x \cap U_{x'} \neq \varnothing$, the transition function is simply $\pi \circ f_{x'}$, which is indeed smooth. $\square$

1.2.2. *Complex analytic manifolds.* In our course we will not be concerned with smooth manifolds but, rather, with complex manifolds. These are spaces that locally look like the complex affine space $\mathbb{C}^n$ and the transition functions are holomorphic functions.

Definition 1.16. A ~~real smooth~~ *complex analytic* manifold of dimension n is a second-countable and Hausdorff topological space X together with an open cover $X = \bigcup_{i \in I} U_i$ such that

- there exist homeomorphisms, called charts, $\varphi_i \colon U_i \to V_i$ with $V_i \subset \cancel{\mathbb{R}^n}\ \mathbb{C}^n$ open;
- for all U_i, U_j with $U_i \cap U_j \neq \varnothing$, the transition map
$$\varphi_{i,j} := \varphi_j \circ \varphi_i^{-1} \colon \varphi_i(U_i \cap U_j) \longrightarrow \varphi_j(U_i \cap U_j) \tag{1.35}$$
is ~~smooth~~ *holomorphic*[2].

Through the identification $\mathbb{C}^n \cong \mathbb{R}^{2n}$ and the fact that every holomorphic function between open subsets of $\mathbb{C}^n$ is a smooth function between open subsets of $\mathbb{R}^{2n}$, we deduce that every n-dimensional complex analytic manifold is also a $2n$-dimensional real smooth manifold.

Another important remark is that all methods of construction of real smooth manifolds discussed in the previous section still holds when substituting real spaces by complex spaces and smooth functions by holomorphic functions. In particular:

- graphs of holomorphic functions,
- level sets of regular values,

[2] A function $f \colon U \to \mathbb{C}^m$, with $U \subseteq \mathbb{C}^n$ open, is called holomorphic at a point $z \in U$ if there exists a a complex linear map $L \colon \mathbb{C}^n \to \mathbb{C}^m$ such that
$$f(z+h) = f(z) + L(h) + \mathrm{o}(h)\,.$$
As a consequence, f is holomorphic as a function of each individual variables at z. Conversely, Hartog's theorem states that if f is holomorphic in each variable separately, then f is holomorphic. Hartog's result has no real smooth counterpart: it's a feature of multivariable complex analysis.

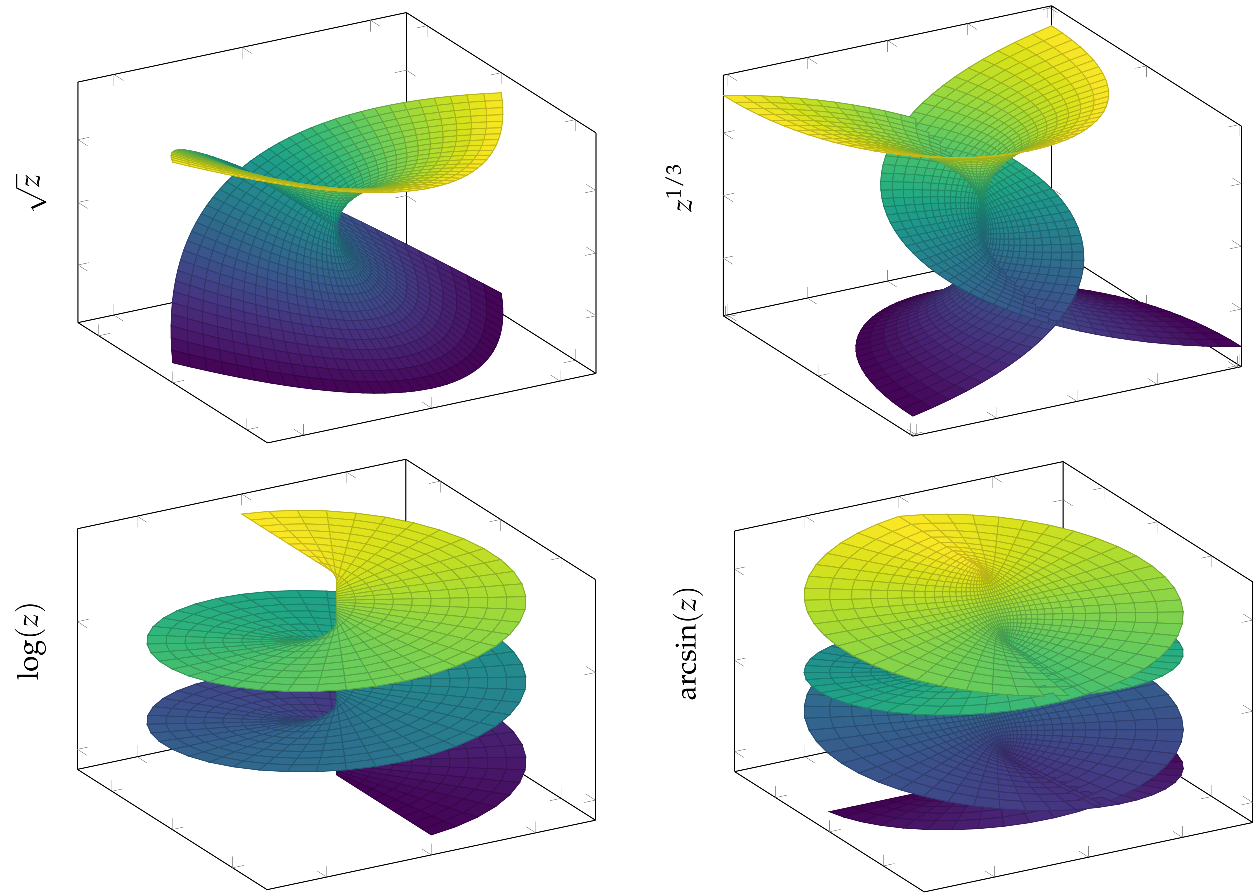


FIGURE 9. The Riemann surfaces of the square-root, the third-root, the logarithm, and the arcsine.

are examples of complex analytic manifolds. For instance, the graph of the function $z \mapsto z^2$ discussed in relation to the square-root function is an example of a one-dimensional complex analytic manifold. More examples stemming from multi-valued function are shown in figure 9.

Another example that generalise to the complex analytic world is that of projective spaces.

Example 1.17. Define the complex projective space of dimension n as

$$P^n(\mathbb{C}) := \left\{ (x_0, \ldots, x_n) \in \mathbb{C}^{n+1} \setminus \{0\} \right\} \Big/ \sim \tag{1.36}$$

where $(x_0, \ldots, x_n) \sim (x_0', \ldots, x_n')$ if and only if $\exists \lambda \in \mathbb{C}^\times$ such that $(x_0, \ldots, x_n) = \lambda(x_0', \ldots, x_n')$. Mutatis mutandis, one can show that $P^n(\mathbb{C})$ is a complex analytic space of dimension n.

Exercise 1.8 (Riemann sphere)**.** *Prove that any point $[z_0 : z_1] \in P^1(\mathbb{C})$ can be realised as $[x_1 : x_2 + \mathrm{i}x_3]$, with x_i real, $x_1^2 + x_2^2 + x_3^2 = 1$, and $x_1 \geq 0$. Deduce that $P^1(\mathbb{C})$, as a 2-dimensional real smooth manifold, is the 2-dimensional sphere S^2. For this reason, $P^1(\mathbb{C})$ is also called the Riemann sphere. Furthermore, we will identify $z \in \mathbb{C}$ with the point $[z : 1] \in P^1(\mathbb{C})$, and the new point $\boldsymbol{\infty} = [1 : 0] \in P^1(\mathbb{C})$ will be called the point at infinity.*

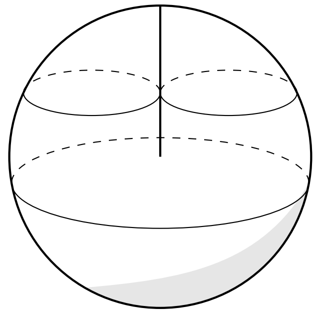


FIGURE 10. The cross-cap.

Exercise 1.9. *Repeat the same argument for $P^2(\mathbb{R})$, and deduce that the real projective plane can be identified with a hemisphere with boundary glued along the antipodal map (known as the cross-cap, see figure 10).*

Finally, we can introduce the concept of Riemann surfaces, which represent the most basic instances of complex analytic manifolds.

Definition 1.18. A Riemann surface is a connected, complex analytic, 1-dimensional manifold.

Remark 1.19. The connectedness assumption is not essential, and different authors use different conventions. For instance, some authors require Riemann surfaces to be compact. We will restrict ourselves to compact Riemann surfaces in the next chapters, but for the time being we allow non-compact objects.

Georg Friedrich Bernhard Riemann (1826–1866)

The name "Riemann surface" is attributed to Riemann, who introduced the concept in his thesis while investigating the problem of multi-valued functions. The term "surface" denotes that every Riemann surface is indeed a surface, a real 2-dimensional manifold, but it encompasses additional structure, specifically a complex structure. Perhaps a more fitting name would be "complex curve", but the terminology "Riemann surface" is now firmly established.

Examples of Riemann surfaces are the graphs of $z \mapsto z^k$, $z \mapsto e^z$, and $z \mapsto \sin(z)$, which correspond to the Riemann surface of the k-th root function, the logarithm, and the arcsine (figure 9). Other (compact) examples include $P^1(\mathbb{C})$ and complex tori.

Exercise 1.10. *Let ω_1 and ω_2 be two complex numbers which are linearly independent over $\mathbb{R}$ (that is, they do not lie on the same real line through 0 in $\mathbb{C}$). The set of all integral linear combinations of ω_1 and ω_2, that is*

$$\Lambda := \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2 = \{\, n_1\omega_1 + n_2\omega_2 \in \mathbb{C} \mid n_1,\ n_2 \in \mathbb{Z} \,\}\,, \tag{1.37}$$

is called a lattice. Define $T := \mathbb{C}/\Lambda$, equipped with the quotient topology induced by the projection map $\pi\colon \mathbb{C} \to T$ (see figure 11).

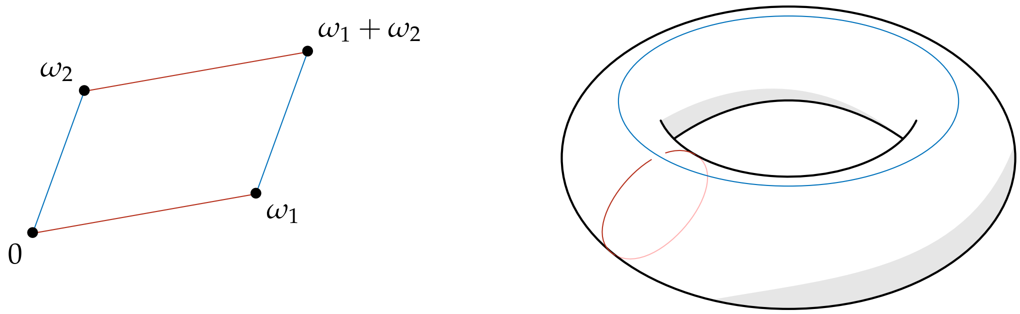


Figure 11. Identification of the torus with the gluing of a polygon.

- *Consider the closed polygon $P \subset \mathbb{C}$ with vertices 0, ω_1, ω_2, $\omega_1 + \omega_2$. Show that for any $z \in \mathbb{C}$, there exists $z_0 \in P$ such that $z - z_0 \in \Lambda$. Thus, $\pi|_P \colon P \to T$ is surjective. What happens at the sides of P?*
- *Deduce that every point in T has a neighbourhood homeomorphic to a disc in $\mathbb{C}$, and that the transitions maps are translations.*

Since translations are holomorphic, we deduce that T is a Riemann surface.

Lecture 3
Mar 7th, 2024

1.2.3. *The geometer's dream.* Up to this point, we have explored two distinct notions of manifolds: smooth real and complex analytic. However, there are numerous other categories of manifolds. A real manifold can be 'smooth', 'topological', or 'piece-wise linear' (PL). The latter are defined by requiring that transition maps should be homeomorphisms or PL maps between open subsets of real space.

The primary aspiration of a geometer, and also an algebraist, is to classify objects within the chosen category. To achieve this goal, it is crucial not only to possess a precise definition of the objects in question but also to establish the concept of maps (or morphisms) between the given objects.

Definition 1.20. Let X and Y be smooth real manifolds with atlases $\{(U_i, \varphi_i)\}$ of dimensions n on X and $\{(V_j, \psi_j)\}$of dimension m on Y. A map $f \colon X \to Y$ is described locally as follows. For any point $x \in X$, let $x \in U_i$ and $f(x) \in V_j$. Then we get a map

$$\psi_j \circ f \circ \varphi_i^{-1} \colon \varphi_i(U_i) \longrightarrow \psi_j(V_j) \tag{1.38}$$

between open subsets of $\mathbb{R}^n$ and $\mathbb{R}^m$. The map f is called smooth if all its coordinate representations are smooth.

Clearly, 'smooth' can be substituted with 'continuous' or 'PL' if the manifolds in question are topological or PL. Alternatively, in the complex analytic setting, it can be replaced by 'holomorphic'. The crucial point is this: we now have a well-defined notion of morphism between the selected objects, forming a category for topological manifolds (**Top**), PL manifolds (**PL**), smooth manifolds (**Diff**), analytic manifolds (**Hol**), and so on. The classification problem stands out as a key challenge in manifold theory. In its strongest form, it involves classifying all objects within the chosen category up to isomorphism. Isomorphism in the different categories are sometimes called by different names: homeomorphisms in **Top**, PL homeomorphisms in **PL**, diffeomorphisms in **Diff**, and biholomorphisms in **Hol**.

Changing the category drastically changes the landscape of possibilities. For example, the validity of the generalised Poincaré conjecture[3] depends on the specific setting:

- **Top**: true in all dimensions (Michael Freedman received the Fields Medal in 1986 for solving the $n = 4$ case).
- **PL**: true in all dimensions other than 4; unknown in dimension 4, where it is equivalent to the smooth version (Stephen Smale received the Fields Medal for his work in 1966).
- **Diff**: false generally, the first known counterexample is in dimension 7 (John Milnor received the Fields Medal in 1962 for the counterexample). True in some dimensions including 1, 2, 3, 5, 6, 12, 56 and 61 (Grigori Perelman received the Fields Medal in 2006 and the Millennium Prize in 2010, both declined, for his solution of the smooth 3-dimensional Poincaré conjecture).

Another example involves the classification of compact orientable topological surfaces compared to compact Riemann surfaces, a topic that will be partially explored in the upcoming section. While the former are categorised by a single discrete invariant, the genus, the latter are distinguished by the discrete genus parameter along with a continuous set of parameters forming a moduli space.

In general, a fundamental tool in the classification problem is the concept of *invariant*: an object (such as a number, a group, a polynomial, etc.) assigned to each object in the chosen category that remains unchanged under isomorphism. Though the classification problem varies across different categories, invariants prove to be useful throughout. For instance, a topological invariant automatically becomes a smooth invariant. For this reason, gaining an understanding of topological surfaces will be beneficial in comprehending Riemann surfaces, as discussed in the next section.

1.2.4. *Topology of compact surfaces.* In this section, we refer to a real topological, compact, connected surface simply as 'surface'. The assumption of connectedness is relatively innocuous, as the same arguments would apply to each connected component. Conversely, the assumption of compactness is more drastic and can be regarded as a finiteness condition.

We already discussed some examples of surfaces in the previous section: S^2, the sphere, and $P^2(\mathbb{R})$, the real projective plane. Another example is the topological torus $T := S^1 \times S^1$. A remarkable theorem from algebraic topology states that these three examples are the building blocks of all surfaces. Before stating the theorem, we have to introduce the concept of connected sum.

[3] Every homotopy n-sphere (a closed n-manifold which is homotopy equivalent to the n-sphere) in the chosen category (i.e. topological manifolds, PL manifolds, or smooth manifolds) is isomorphic in the chosen category (i.e. homeomorphic, PL-homeomorphic, or diffeomorphic) to the standard n-sphere.

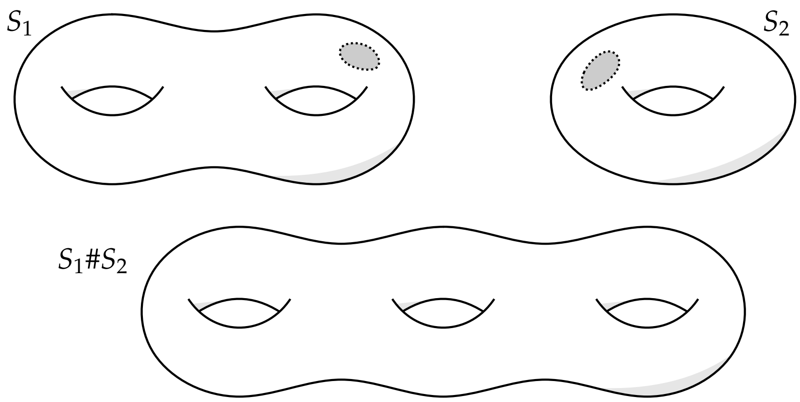


FIGURE 12. Connected sum of surfaces.

Definition 1.21. Consider two surfaces S_1 and S_2. The connected sum $S_1\#S_2$ is the surface obtained by removing a small open disc from each of the two surfaces and identifying the two boundaries via a homeomorphism.

Figure 12 illustrates the connected sum operation. Naturally, it is necessary to prove that the operation is well-defined, meaning that the outcome remains consistent regardless of the choice of discs and the choice of homeomorphism used to connect their boundaries. In other words, it is a surface well-defined up to homeomorphism. One can prove that the connected sum is associative, commutative, with identity being the sphere. More importantly, the connected sum allows to produce a complete list of surfaces, solving the classification problem in this specific case.

Theorem 1.22 (Classification of surfaces)**.** *Every connected, compact surface can be homeomorphically identified with precisely one surface from the following list:*

- *orientable surfaces: $T^{\#g}$, the connected sum of g tori for $g \geq 0$ (here $T^{\#0} = S^2$ is the sphere); g is called the genus (see figure 13);*
- *non-orientable surfaces: $P^2(\mathbb{R})^{\#m}$, the connected sum of m projective planes for $m \geq 1$; m is called the demigenus.*

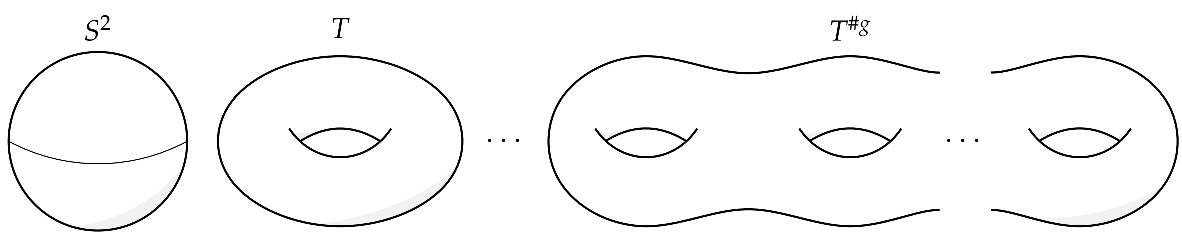
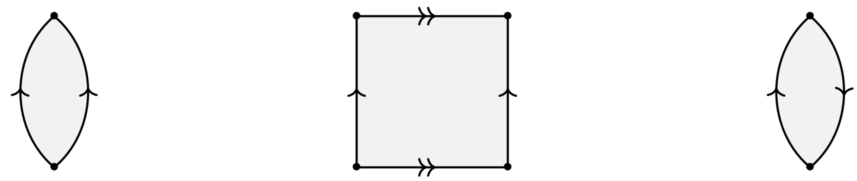


FIGURE 13. Classification of compact, orientable surfaces.

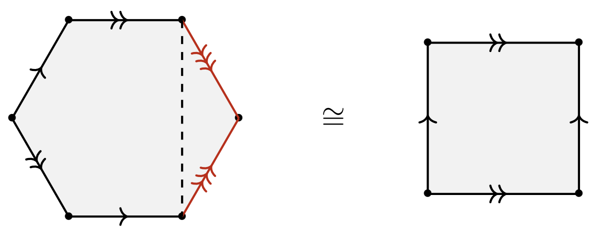
A brief sketch of the proof is as follows: to establish the above list, one demonstrates that any surface can be represented by an identification polygon. Secondly, it is shown that all identification polygons correspond to surfaces in the list. To prove that different elements of the list are non-homeomorphic, two topological invariants–orientability and Euler characteristic–are introduced. One can prove that no two different elements in the list share the same values for both invariants. The rest of the section introduces these concepts.

Identification polygons. Let us start with some intuitive examples: the sphere (genus 0), the torus (genus 1), and the real projective plane as the cross-cap (demigenus 1):

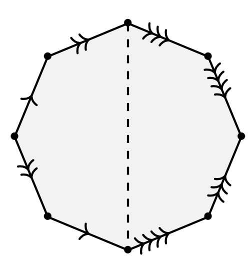


Sphere Torus Cross-cup

But how can we obtain connected sums? A simple example would be the following, realising $T\#S^2 \cong T$. The dashed line represents the connected sum.



The isomorphism is realised by "shrinking" the red part, the sphere, down to a single point. A less trivial example is $T\#T$, the 2-holed torus (that is, the surface S_1 in figure 12).



A more systematic way of constructing glueings is given as follows. We define a set A of n letters as an alphabet, and the set $A \cup \bar{A}$, formed by repeating each letter with a bar above it, is called a doubled n-letter alphabet. Each pair $(a, \bar{a})$ is referred to as a pair of conjugate letters.

Definition 1.23. An identification polygon with $2n$ sides is a word w constructed from a doubled n-letter alphabet such that, for each pair of conjugate letters, w contains exactly two letters from that pair (repetitions allowed). In particular, the word w must have exactly $2n$ letters.

As example, take $A = \{a, b\}$, so that $A \cup \bar{A} = \{a, b, \bar{a}, \bar{b}\}$. Examples of identification polygons with 4 sides are

$$w_1 = a\bar{a}b\bar{b}\,, \qquad w_2 = aab\bar{b}\,, \qquad w_3 = ab\bar{a}\bar{b}\,, \qquad w_4 = aba\bar{b}\,. \tag{1.39}$$

Non-examples are, for instance: $aaaa$ (there is no letter from the pair $(b, \bar{b})$) or $a\bar{a}ab$ (there is only one letter from the pair $(b, \bar{b})$).

An identification polygon w gives rise to a compact surface as follows. Consider a regular $2n$-gon and label its sides in a counter-clockwise manner to spell out the word w. Assign a counter-clockwise orientation to each side for normal letters and a clockwise orientation for barred letters. Subsequently, for every pair of equal or conjugate letters in the alphabet, align and identify the two sides based on the specified orientation. A surface S that is homeomorphic to the resulting surface is described as being represented by the identification polygon w.

Exercise 1.11. *Which surfaces are represented by the identification polygons in equation (1.39)?*

Connected sums of surfaces represented by identification polygons is easily realised as concatenation of words.

Lemma 1.24. *If S_1, S_2 are surfaces represented by the polygons w_1 and w_2, then the surface $S_1 \# S_2$ is represented by the polygon $w_1 w_2$ (we are assuming that the two words use letters from different alphabets).*

Exercise 1.12. *Convince yourself that $P^2(\mathbb{R})\#P^2(\mathbb{R})$ is the Klein bottle, and prove using polygon identification that $P^2(\mathbb{R})\#T \cong P^2(\mathbb{R})^{\#3}$.*

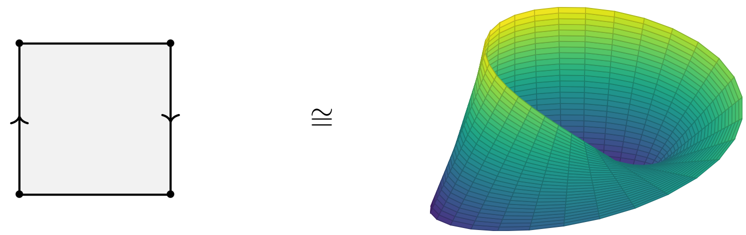
Exercise 1.13. *Come up with a list of identification polygons that produce the list in theorem 1.22. Convince yourself that any identification polygon gives a surface homeomorphic to one in the list.*

Orientability and Euler characteristic. We are now armed to define two invariants that completely characterise: orientability and Euler characteristic. Orientability is a general concept than can be defined for any topological manifold. The easiest definition requires the manifold to be C^1, but an alternative definition can be given in the topological case using homology[4].

Definition 1.25. A smooth manifold X is orientable if and only if it admits an atlas such that all transition functions are orientation-preserving (i.e. the determinant of the Jacobian matrix is positive in an open dense).

[4]In the topological setting, we say that a connected n-dimensional manifold X is orientable if and only if its n-th singular homology group $H_n(X, \mathbb{Z})$ is isomorphic to the integers $\mathbb{Z}$. Intuitively, the n-th homology group of an n-dimensional manifold measures 'how many sides' the manifold has. The integers are generated by two elements, ± 1, which reflects the intuitive idea that an orientable manifolds has two sides.

An example of non-orientable surface is the projective plane $P^2(\mathbb{R})$. Another example is the Klein bottle $P^2(\mathbb{R})\#P^2(\mathbb{R})$. One can show that a surface is non-orientable if and only if it contains an open subset homeomorphic to a Möbius strip. The Möbius strip can be realised as the following 'non-compact' identification polygon.



As such, one can deduce that all surfaces of the form $P^2(\mathbb{R})^{\# m}$ are non-orientable. On the contrary, the sphere and the connected sum of g tori are orientable. In other words, orientability distinguishes elements from the first and the second item of the list in the classification theorem 1.22.

The second invariant, the Euler characteristic, can be defined for smooth manifolds (or more generally, manifolds of finite CW-complex type). In the surface case, the definition reduces to the historical definition by Euler, which implicitly was known already to Descartes around 1620.

Definition 1.26. A good graph on a surface S is a graph Γ on S such that:

- $S \setminus \Gamma$ is homeomorphic to a disjoint union of open discs,
- edges only cross at vertices,
- no edge ends without a vertex.

For a given good graph Γ on S, the Euler characteristic is defined as

$$\chi_S := |V_\Gamma| - |E_\Gamma| + |F_\Gamma|\,, \tag{1.40}$$

where V_Γ, E_Γ, F_Γ are the sets of vertices, edges and faces of Γ.

It can be shown that the Euler characteristic is independent of the choice of good graph.

Exercise 1.14. *Prove that the Euler characteristic of $T^{\# g}$ is $2-2g$. Prove that the Euler characteristic of $P^2(\mathbb{R})^{\# m}$ is $2-m$. Conclude that the genus and the demigenus uniquely characterise the surface within the orientable and the non-orientable classes.*

To sum up: isomorphism classes of compact connected topological surfaces are divided into two classes (orientable/non-orientable), and both classes are labelled by an integer (genus/demigenus); furthermore, each isomorphism class consists of a single element.

2. Riemann surfaces: general theory

Lecture 4
Mar 14th, 2024

2.1. **Examples.** In the previous chapter, we have introduce the concept of Riemann surfaces and discussed a handful of examples:

- the complex plane $\mathbb{C}$, and more generally any open subset $U \subseteq \mathbb{C}$,
- the Riemann sphere $P^1(\mathbb{C})$,
- the tori $\mathbb{C}/\Lambda$ (a more detailed discussion will follow),
- graphs of functions, such as the Riemann surface of the square root,
- level sets of regular values of holomorphic functions of the form $F\colon \mathbb{C}^m \to \mathbb{C}^{m-1}$ (also known as affine curves if F is a polynomial).

We also understood how, in the compact case, every topological surfaces is uniquely characterised by the orientability/non-orientability and the Euler characteristic. Since every Riemann surfaces is a topological surface, it is natural to ask ourself: can we endow every compact topological surface with a Riemannian structure?

As established in corollary 1.3, a holomorphic function preserves orientation when regarded as a differentiable function from the real plane to itself. Given that all transition functions of a Riemann surface are holomorphic, it follows that every Riemann surface is orientable. Consequently, the real projective plane, the Klein bottle, and, more generally, every non-orientable compact topological surface cannot be Riemann surfaces. Conversely, we have previously seen that $P^1(\mathbb{C})$ is topologically a sphere, while $\mathbb{C}/\Lambda$ is topologically a torus. It can be shown that all orientable compact topological surfaces can be endowed with a Riemannian structure.

On the other hand, not every bicontinuous function is biholomorphic. For example, every pair of complex tori $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ are homeomorphic, since the homeomorphism class of a torus consists of a single element. The natural question is: are $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ biholomorphic too? For concreteness, consider the example of

$$\Lambda = \mathbb{Z} + \tau\mathbb{Z} \qquad \text{and} \qquad \Lambda' = \mathbb{Z} + \bar{\tau}\mathbb{Z}\,. \tag{2.1}$$

The map $z \mapsto \bar{z}$ is a homeomorphism between $\mathbb{C}/\Lambda$ and $\mathbb{C}/\Lambda'$ (geometrically, it is the reflection along the real axis), but it is not a holomorphic map! Complex tori given by different lattices are not necessarily biholomorphic, but a characterisation can be given in accordance with the following theorem.

Theorem 2.1 (The modular curve)**.** *The set of equivalence classes of complex tori is in one-to-one correspondence with the upper-half plane* $\mathbb{H} := \{\tau \in \mathbb{C} \mid \Im(\tau) > 0\}$*, quotient by the action of* $\mathrm{SL}(2,\mathbb{Z})$*:*

$$\mathbb{H}/\mathrm{SL}(2,\mathbb{Z}) \xrightarrow{1:1} \{\,\mathbb{C}/\Lambda\,\}\,/\textit{biholomorphism}\,, \qquad [\tau] \longmapsto [\mathbb{C}/(\mathbb{Z}+\tau\mathbb{Z})]\,. \tag{2.2}$$

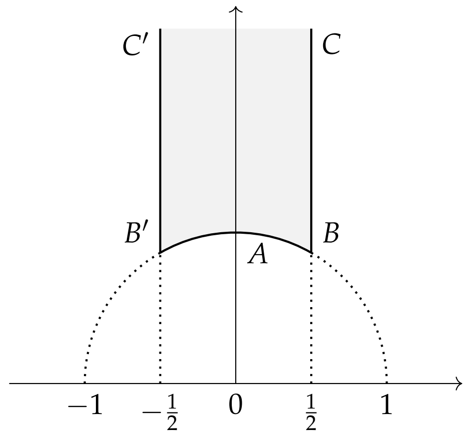


FIGURE 14. A fundamental domain for the quotient $\mathbb{H}/\mathrm{SL}(2,\mathbb{Z})$. The arcs AB and AB' and the half-lines BC and $B'C'$ are identified, where $A = \mathrm{i}$, $B = \frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$, $B' = -\frac{1}{2} + \mathrm{i}\frac{\sqrt{3}}{2}$, $C = \frac{1}{2} + \mathrm{i}\infty$, $C' = -\frac{1}{2} + \mathrm{i}\infty$.

The group $\mathrm{SL}(2,\mathbb{Z})$*, called the modular group, is the group of* 2×2 *matrices with integer coefficients and determinant* 1*. Its action on the upper-half plane, called the modular action, is defined as*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} . \tau := \frac{a\tau + b}{c\tau + d} \,. \tag{2.3}$$

The set $\mathbb{H}/\mathrm{SL}(2,\mathbb{Z})$*, called the modular curve, is pictured in figure 14.*

Exercise 2.1. *The proof is left as a guided exercise.*

- *Let* $\Lambda, \Lambda' \subset \mathbb{C}$ *be two lattices. Suppose* $\exists \alpha \in \mathbb{C}^{\times}$ *such that* $\alpha\Lambda \subseteq \Lambda'$*. Show that the map* $\mathbb{C} \to \mathbb{C}$*,* $z \mapsto \alpha z$ *induces a holomorphic map* $\mathbb{C}/\Lambda \to \mathbb{C}/\Lambda'$*, which is biholomorphic if and only if* $\alpha\Lambda = \Lambda'$*.*
- *Show that every torus* $\mathbb{C}/\Lambda$ *is isomorphic to a torus of the form* $T(\tau) := \mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$*, where* $\tau \in \mathbb{H}$*.*
- *Suppose* $\gamma \in \mathrm{SL}(2,\mathbb{C})$ *and* $\tau \in \mathbb{H}$*. Let* $\tau' := \gamma.\tau$*, according to the action defined in equation (2.3). Prove that* $T(\tau)$ *and* $T(\tau')$ *are isomorphic.*

Remark 2.2. The one-to-one correspondence given in theorem 2.1 can be extended from a set correspondence to a correspondence of topological spaces. Furthermore, both sides of the correspondence can be endowed with a natural structure of 1-dimensional complex orbifolds, a generalisation of manifolds where points exhibit symmetry. This elucidates the term 'curve' in 'modular curve': the set that parametrises complex tori forms a complex (orbi)curve. Additionally, it resolves the classification problem for complex tori by characterising their equivalence classes in terms of a well-understood geometric object. This case is much more difficult than the topological case, where the set of homeomorphism classes of tori is a just single point. Here, instead, we find a whole continuum of biholomorphism classes of complex tori.

The space $\mathbb{H}/\mathrm{SL}(2,\mathbb{Z})$ is an example of a *moduli space*: a space that parametrises specific geometric objects, while the space itself is naturally endowed with an intrinsic geometry. Another familiar example is the projective space, which parametrises directions in a given real/complex vector space and is naturally endowed with a real-smooth/complex-analytic manifold structure. Part of the classification problem of compact Riemann surfaces (beyond the genus 1 case) involves understanding the spaces that classify their equivalence classes, that is, understand the moduli space of genus g Riemann surfaces.

Other interesting examples of Riemann surfaces are affine plane curves, that is level sets of polynomials in two variables.

Example 2.3 (Plane affine curves)**.** For a given polynomial $F \in \mathbb{C}[z, w]$ with 0 as a regular value, we denote the associated level set at $Z(F) := F^{-1}(0)$, also called smooth affine plane curve.'' Notice that the regular value assumption is equivalent to $\{p \in \mathbb{C}^2 \mid F(p) = \frac{\partial F}{\partial z}(p) = \frac{\partial F}{\partial w}(p) = 0\} = \varnothing$.

- Lines. For $F(z, w) = az + bw + c$ and $(a, b) \neq (0, 0)$,
$$Z(F) = \left\{ (z, w) \in \mathbb{C}^2 \;\middle|\; az + bw = c \right\} \tag{2.4}$$
is a Riemann surface biholomorphic to $\mathbb{C}$.
- The 'complex circle'. Take $F(x, y) = z^2 + w^2 - 1$, and set
$$Z(F) = \left\{ (z, w) \in \mathbb{C}^2 \;\middle|\; z^2 + w^2 = 1 \right\} . \tag{2.5}$$
Although the equation seems the same as that of a circle in the real plane, the Riemann surface $Z(F)$ is drastically different. Indeed, we can factorise the polynomial as $F(z, w) = (z + \mathrm{i}w)(z - \mathrm{i}w) - 1$ and perform the change of variables $(z, w) \mapsto (u, v) = (z + \mathrm{i}w, z - \mathrm{i}w)$. Thus, we can rewrite $Z(F)$ as $\{(u, v) \in \mathbb{C}^2 \mid uv = 1\}$, that is a 'complex hyperbola'. One can check that the map
$$\left\{ (u, v) \in \mathbb{C}^2 \;\middle|\; uv = 1 \right\} \longrightarrow \mathbb{C} \setminus \{0\} , \qquad (u, v) \longmapsto u \tag{2.6}$$
is an biholomorphism of complex manifolds. Hence, $Z(F) \cong \mathbb{C} \setminus \{0\}$ as a Riemann surface. This is an example of a smooth affine conic. A similar result holds for all smooth affine conics.

Other examples of compact Riemann surfaces arise when considering level sets of polynomials in projective spaces (rather than affine spaces). The first problem, though, is that projective spaces are quotients, so in order to obtained a well-defined level set we need to check that our definition is independent of the choice of representatives. We will only focus on the plane case, although the discussion can be easily generalised beyond the planar case.

Definition 2.4. A polynomial $F \in \mathbb{C}[z_0, z_1, z_2]$ is called homogeneous of degree d if
$$F(\lambda z_0, \lambda z_1, \lambda z_2) = \lambda^d F(z_0, z_1, z_2) \tag{2.7}$$
for all $\lambda \in \mathbb{C}^\times$.

Exercise 2.2. *Let $F \in \mathbb{C}[z_0, z_1, z_2]$. Prove that the following are equivalent.*

- *F is homogeneous of degree d.*

- *Every monomial in F has degree d.*
- *F satisfies the so-called Euler's identity: $z_0 \frac{\partial F}{\partial z_0} + z_1 \frac{\partial F}{\partial z_1} + z_2 \frac{\partial F}{\partial z_2} = d \cdot F$.*

For a homogeneous polynomial, the property of being zero or not depends only on the equivalence class of (z_0, z_1, z_2) in $P^2(\mathbb{C})$. Hence, the following definition is well-posed.

Definition 2.5. For a given homogeneous polynomial $F \in \mathbb{C}[z_0, z_1, z_2]$ of degree d, define

$$Z(F) := \left\{ [z_0 : z_1 : z_2] \in P^2(\mathbb{C}) \;\middle|\; F(z_0, z_1, z_2) = 0 \right\}. \tag{2.8}$$

The set $Z(F)$ is called a projective plane curve of degree d.

The main question is: are projective plane curves Riemann surfaces? The answer is yes, provided that a 'regular value' condition is satisfied.

Proposition 2.6. *Let $F \in \mathbb{C}[z_0, z_1, z_2]$ be a homogeneous polynomial. Suppose that*

$$\left\{ p \in \mathbb{C}^3 \setminus \{0\} \;\middle|\; F(p) = \frac{\partial F}{\partial z_0}(p) = \frac{\partial F}{\partial z_1}(p) = \frac{\partial F}{\partial z_2}(p) = 0 \right\} = \varnothing\,. \tag{2.9}$$

Then $Z(F) \subset \mathbb{P}^2(\mathbb{C})$ has a natural structure of compact Riemann surface. In this case, $Z(F)$ is called a smooth projective plane curve.

Proof. We first show that $Z(F)$ is compact by showing that $Z(F)$ is a closed subset of $P^2(\mathbb{C})$, which is a compact topological space. To this end, consider the diagram

$$\begin{array}{ccc} \mathbb{C}^3 \setminus \{0\} & \xrightarrow{F} & \mathbb{C} \\ {\scriptstyle\pi}\big\downarrow & & \\ P^2(\mathbb{C}) & & \end{array} \tag{2.10}$$

where π is the natural projection. By definition, $Z(F)$ is a closed subset of $P^2(\mathbb{C})$ if and only if $\pi^{-1}(Z(F))$ is closed in $\mathbb{C}^3 \setminus \{0\}$. As $\pi^{-1}(Z(F)) = F^{-1}(0)$ and the latter is closed (being the inverse image of $\{0\} \subset \mathbb{C}$ under the continuous function F), we deduce that $Z(F)$ is closed. Notice that we did not used the smoothness assumption here.

To prove that $Z(F)$ is a Riemann surface, it is sufficient to show that its intersection with any of the charts of $P^2(\mathbb{C})$ is a Riemann surface (think about why that is true). Consider the chart

$$U = \{ [z_0 : z_1 : z_2] \mid z_0 \neq 0 \}. \tag{2.11}$$

with coordinates $\varphi([z_0 : z_1 : z_2]) = \frac{1}{z_0}(z_1, z_2) =: (w_1, w_2)$. Then the dehomogenisation of F with respect to z_2, defined as $f(w_1, w_2) := F(1, w_1, w_2)$, satisfies

$$\varphi(Z(F) \cap U) = Z(f)\,. \tag{2.12}$$

Notice that $Z(F)$ is a subset of the complex projective plane, while $Z(f)$ is a subset of the $\mathbb{C}$. From the relations

$$\frac{\partial F}{\partial z_1}(1, w_1, w_2) = \frac{\partial f}{\partial w_1}(w_1, w_2)\,, \qquad \frac{\partial F}{\partial z_2}(1, w_1, w_2) = \frac{\partial f}{\partial w_2}(w_1, w_2)\,, \tag{2.13}$$

and Euler's identity, we conclude that equation (2.9) implies that 0 is a regular value for f. Indeed, suppose by contradiction that there exists $(w_1, w_2) \in \mathbb{C}$ such that

$$f(w_1, w_2) = \frac{\partial f}{\partial w_1}(w_1, w_2) = \frac{\partial f}{\partial w_2}(w_1, w_2) = 0\,. \tag{2.14}$$

Then $F(1, w_1, w_2) = 0$, and $\frac{\partial F}{\partial z_1}(1, w_1, w_2) = \frac{\partial F}{\partial z_2}(1, w_1, w_2) = 0$. By Euler's identity, we deduce that $\frac{\partial F}{\partial z_0}(1, w_1, w_2) = 0$ as well, which contradicts equation (2.9). Hence, $Z(f)$ is a Riemann surface, being the zero set of a polynomial at a regular value. A similar argument holds for the other local charts of $P^2(\mathbb{C})$, thus proving the proposition. $\square$

Remark 2.7 (Genus-degree formula)**.** It can be shown that smooth projective plane curves are connected. Since they are compact and orientable, by the classification theorem of surfaces they must be topologically homeomorphic to a genus g surface. The main question is: can we compute the genus in terms of the degree? The answer is yes:

$$g = \frac{(d-1)(d-2)}{2}\,. \tag{2.15}$$

We will prove the genus-degree formula as a consequence of the Riemann–Hurwitz formula.

Example 2.8 (Conic)**.** Consider the projective plane curve $C := Z(F)$ of degree 2 defined by $F(z_0, z_1, z_2) := z_0^2 + z_1^2 - z_2^2$. It is easy to check that F satisfies the smoothness condition, hence C is a compact Riemann surface. Thanks to the genus-degree formula, we see that C has genus 0.

On the other hand, dehomogenising with respect to z_2 we find the Riemann surface of the 'complex circle':

$$\left\{ (w_0, w_1) \in \mathbb{C}^2 \;\middle|\; w_0^2 + w_1^2 = 1 \right\}\,, \tag{2.16}$$

which is biholomorphic to $\mathbb{C}^\times$. Dehomogenising with respect to z_0 or z_1 gives a 'complex hyperbola', which again is biholomorphic to $\mathbb{C}^\times$. In other words, the projective curve C intersected with the three charts of the projective plane gives three copies of $\mathbb{C}^\times$. However, the Riemann surface C is compact, and intuition suggests that it is the sphere obtained by compactifying $\mathbb{C}^\times$. Indeed, this is in accordance with the genus-degree formula.

The curve $Z(F)$ is an example of a smooth conic (a conic is projective curves of degree 2). This has to do with the fact that affine plane conics are obtained as plane sections of a cone: before we identify all the points on a line through the origin, the solutions of $F(z_0, z_1, z_2) = z_0^2 + z_1^2 - z_2^2 = 0$ in $\mathbb{C}^3$ give a cone, and slicing it with different planes just amounts to restricting the projective curve $Z(F)$ to different affine charts.

Example 2.9 (Fermat curve)**.** Consider the projective plane curve of degree $n \geq 1$, called Fermat curve, defined by the homogeneous polynomial

$$F(z_0, z_1, z_2) := z_0^n + z_1^n - z_2^n\,. \tag{2.17}$$

It is easy to check that $Z(F)$ is smooth, therefore of genus $g = \frac{(n-1)(n-2)}{2}$. Fermat's Last Theorem states that there are no non-trivial integer solutions to the Fermat equation $x^n + y^n = z^n$ for $n \geq 3$; therefore, the Fermat curve has no non-trivial rational points for $n \geq 3$. Fermat's Last Theorem is one of the biggest mathematical achievements of the 20th century, proved 358 years after it was conjectured thanks to the contribution of several mathematicians such as Sophie Germain, Kummer, Shimura, Taniyama, Weil, Frey, Ribet, Taylor, and Wiles.

Example 2.10 (Elliptic curves)**.** Consider the projective plane curve $E := Z(F)$ of degree 3, called elliptic curve, defined by the homogeneous polynomial Lecture 5 Mar 21st, 2024

$$F(z_0, z_1, z_2) := z_1^2 z_2 - (z_0 - \alpha_1 z_2)(z_0 - \alpha_2 z_2)(z_0 - \alpha_3 z_2) \tag{2.18}$$

with $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{C}$. It is easy to check that, if the α_i's are distinct, E is smooth. Hence, E has genus 1. It can be shown that elliptic curves are isomorphic to complex tori; such connection is a very beautiful and classical story, and it will be partially explored in a future exercise.

Notice that complex tori have a natural addition operation induced by the complex sum in $\mathbb{C}$. That is, the operation

$$\mathbb{C}/\Lambda \times \mathbb{C}/\Lambda \longrightarrow \mathbb{C}/\Lambda\,, \qquad [z] + [w] := [z + w] \tag{2.19}$$

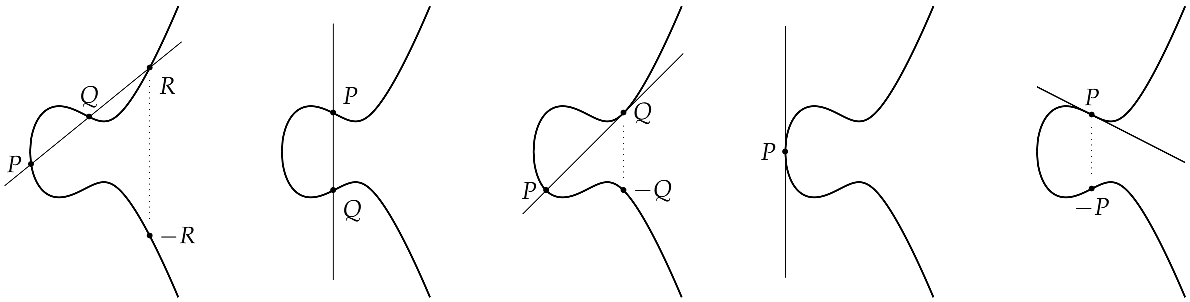
is well defined. If we believe that elliptic curves are indeed complex tori, a natural question arises: how is the addition defined for elliptic curves? The answer is very simple and geometric. To start with, let us transform the elliptic curve in its Weierstraß normal form:

$$G(Z_0, Z_1, Z_2) = Z_1^2 Z_2 - Z_0^3 - aZ_0 Z_2^2 - bZ_2^3\,. \tag{2.20}$$

This can be achieved through the linear transform $(z_0, z_1, z_2) \mapsto (Z_0, Z_1, Z_2) = (z_0 + \frac{1}{3}(\alpha_1 + \alpha_2 + \alpha_3)z_2, z_1, z_2)$. Furthermore, the smoothness condition is equivalent to $\Delta := -16(4a^3 + 27b^2) \neq 0$. For an elliptic curve in the Weierstraß form, define the group law as follows.

- The identity element is $O := [0 : 1 : 0]$.
- For a point $P = [Z_0 : Z_1 : Z_2] \in E$, define $-P := [Z_0 : -Z_1 : Z_2]$. Thanks to the symmetry of the elliptic curve in Weierstraß form, we have $-P \in E$.
- If for a line $\ell \subset P^2(\mathbb{C})$, the intersection $\ell \cap E$ consists of:
  - three distinct points P, Q, R, then $P + Q + R = O$ (generic case);
  - two distinct points P, Q, with ℓ tangent to E at Q, then $P + Q + Q = O$ (tangent point);
  - a single point P, then $P + P + P = O$ (inflection point).

The group law is summarised in the following pictures, representing the real points of the dehomogenisation with respect to Z_2. The first two drawings represent the generic case (in the second picture, the third point of intersection is O); the second two drawings represent the tangency case (in the fourth picture, the second point of intersection is O); the last picture represents the inflection case.



$P+Q+R=O$ $\quad$ $P+Q+O=O$ $\quad$ $P+Q+Q=O$ $\quad$ $P+P+O=O$ $\quad$ $P+P+P=O$

It can be shown that the group law defined above is associative, commutative, with O being the identity element.

Exercise 2.3 (The Puzzle of the Doctor of Physic ☠). *Find a positive rational solution (different from the trivial solution $\{1, 2\}$) to the equation*

$$x^3 + y^3 = 9\,. \tag{2.21}$$

This is a version of Fermat's Last Theorem, with the right-hand side equal to 9 (rather than 1).

The question is a reformulation of the 20th puzzle from The Canterbury Puzzles and Other Curious Problems *(1907) by Henry Dudeney.*

> *This Doctor, learned though he was, for "In all this world to him there was none like To speak of physic and of surgery," and "He knew the cause of every malady," yet was he not indifferent to the more material side of life. "Gold in physic is a cordial; Therefore he lovéd gold in special." The problem that the Doctor propounded to the assembled pilgrims was this. He produced two spherical phials, as shown in our illustration, and pointed out that one phial was exactly a foot in circumference, and the other two feet in circumference.*
>
> *"I do wish," said the Doctor, addressing the company, "to have the exact measures of two other phials, of a like shape but different in size, that may together contain just as much liquid as is contained by these two." To find exact dimensions in the smallest possible numbers is one of the toughest nuts I have attempted. Of course the thickness of the glass, and the neck and base, are to be ignored.*

In mathematical terms, the puzzle can be rephrased as follows. The Doctor has two spherical phials of circumference 1 and 2 feet respectively, that is radii $\frac{1}{2\pi}$ and $\frac{1}{\pi}$ respectively. Hence, the total volume contained in the two phials is $\frac{4}{3}\pi(\frac{1}{8\pi^3} + \frac{1}{\pi^3}) = \frac{3}{2\pi^2}$. The Doctor is asking for two more spherical phials with the same total volume, but different rational circumferences. That is, he is looking for a pair of positive rational numbers (x, y), different from $(1, 2)$ and $(2, 1)$ such that

$$\frac{4}{3}\pi\left(\left(\frac{x}{2\pi}\right)^3 + \left(\frac{y}{2\pi}\right)^3\right) = \frac{3}{2\pi^2}\,. \tag{2.22}$$

Simplifying, we get to find a non-trivial positive rational solution to equation (2.21).

Hint. *Notice that $E = Z(z_0^3 + z_1^3 - 9z_2^3)$ is a smooth projective plane curve of degree 3, hence an elliptic curve with its group structure. Can you define the group law (even if E is not in Weierstraß form)? And can you use the group law to find new rational solutions, starting from the trivial ones?*

2.2. **The Riemann–Hurwitz formula.** Consider a non-constant holomorphic map $f\colon X \to Y$ between Riemann surfaces. For any point $x \in X$, the local form of holomorphic functions (lemma 1.10) guarantees that there exists local coordinate (z, w) centred at $(x, f(x))$ such that the function f has the form

$$z \longmapsto z^k \tag{2.23}$$

for some integer $k > 0$. This means that there exist neighbourhoods U of x and V of $f(x)$ and homeomorphisms $\varphi\colon U \to U' \subseteq \mathbb{C}$, $\psi\colon V \to V' \subseteq \mathbb{C}$, such that $\varphi(x) = 0$, $\psi(f(x)) = 0$, and

$$(\psi \circ f \circ \varphi^{-1})(z) = z^k\,. \tag{2.24}$$

Exercise 2.4. *Prove that the integer k does not depend on the choice of local coordinates.*

Definition 2.11. With the notation above, define the multiplicity (or ramification index) of the map f at the point x to be

$$\mu_x(f) := k\,. \tag{2.25}$$

A function f with $\mu_x(f) = 1$ is called unramified at x, and ramified if $\mu_x(f) > 1$. In the latter case, x is called a ramification point. The ramification locus Ram_f is the subset of X consisting of all ramification points. If x is a ramification point, then $f(x) \in Y$ is called a branch point. The branch locus Brnch_f is the subset of Y consisting of all branch points.

Geometrically, the multiplicity tells you how many distinct solutions there are to the equation $f(x) = y$ for fixed y. If locally the map is of the form $z \mapsto z^k$, it is clear that $w = 0$ is a bad point for the equation $z^k = w$: only $z = 0$ is a solution, while for $w \neq 0$ there are k distinct solutions. Intuitively, a ramification point is where the local number of solutions to $f(x) = y$ has suddenly dropped, and the ramification index counts exactly by how much such number has dropped.

Example 2.12. Consider the map $f\colon \mathbb{C} \to \mathbb{C}$ defined by $z \mapsto z^2$. The only ramification point is the origin, with multiplicity equal to 2.

The above map can be extended to a map between complex projective lines: $f\colon P^1(\mathbb{C}) \to P^1(\mathbb{C})$, defined by $[z_0 : z_1] \mapsto [z_0^2 : z_1^2]$. In any of the local charts U_0 or U_1 of the complex projective line, the map is simply $z \mapsto z^2$. Hence, the ramification points are $[0 : 1]$ and $[1 : 0]$, both of multiplicity equal to 2. Thinking about $P^1(\mathbb{C})$ as the one-point compactification of $\mathbb{C}$, we can say that $z \mapsto z^2$ is ramified at the origin and at infinity (see figure 15).

Example 2.13. Let E be an elliptic curve, that is $E = Z(F)$ with $F(z_0, z_1, z_2) = z_1^2 z_2 - (z_0 - \alpha_1 z_2)(z_0 - \alpha_2 z_2)(z_0 - \alpha_3 z_2)$ and distinct α_i's. Consider the map $g\colon E \to P^1(\mathbb{C})$ defined by

$$\begin{cases} [z_0 : z_1 : z_2] \longmapsto [z_0 : z_2] & \text{if } z_2 \neq 0\,, \\ [0 : 1 : 0] \longmapsto [0 : 1]\,. \end{cases} \tag{2.26}$$

The map is well-defined, as the point $[0 : 1 : 0]$ is the only point with $z_2 = 0$. The structure of the map is perhaps more transparent after dehomogenisation with respect to z_2. In the corresponding affine chart of $P^2(\mathbb{C})$, the elliptic curve is of the form

$$y^2 = (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)\,, \tag{2.27}$$

and the map g sends $(x, y) \mapsto x$. It is then clear that a generic point has multiplicity 2, since we can use $y = \sqrt{(x - \alpha_1)(x - \alpha_2)(x - \alpha_3)}$ as local coordinate and the map is then a square. It is also clear that $x = \alpha_1,\ \alpha_2,\ \alpha_3$ are ramification points of multiplicity 2. In the projective plane embedding of E, they correspond to the points $p_i := [\alpha_i : 0 : 1] \in E$. As in the previous example though, we have to be careful: dehomogenisation with respect to z_2 loses information for the point $\infty := [0 : 1 : 0]$. One can check that ∞ is indeed a ramification point of multiplicity 2. See figure 15 for a pictorial representation of the map.

Let us turn our attention to compact Riemann surfaces. In this case, holomorphic functions enjoy several useful properties.

Lemma 2.14. *Let $f\colon X \to Y$ be a non-constant holomorphic function between Riemann surfaces with X compact.*

(1) Then f is surjective. Hence, Y is compact.
(2) The fibres $f^{-1}(y)$ and Ram_f are finite sets.

Proof. Let us start with the first property. By the open mapping theorem (theorem 1.4), a non-constant holomorphic map is open. This is a local statement, so $f\colon X \to Y$ is open. Furthermore, f is closed because X is compact and Y is Hausdorff[5]. Hence, $f(X)$ is clopen in Y. The only clopen subsets of a connected space (recall that a Riemann surface is connected by definition) are the empty set and the whole space. Since $f(X)$ is non-empty, we conclude that $f(X) = Y$, that is f is surjective. We also conclude that Y is compact, since it is the image of X that is compact.

[5]Recall the *closed/proper mapping theorem*: every continuous function $f\colon X \to Y$ from a compact space X to a Hausdorff space Y is closed and proper (i.e. it sends closed sets to closed sets, and the preimage of compact sets are compact).

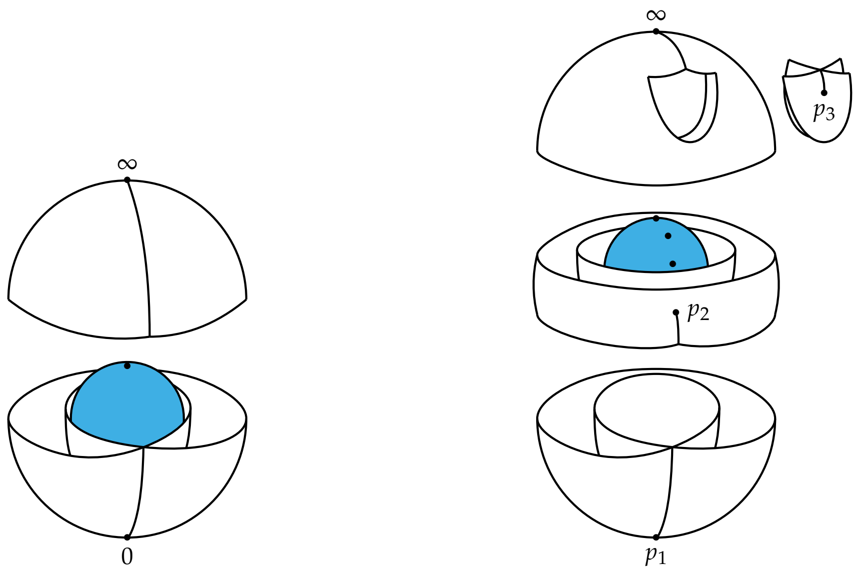


FIGURE 15. Left: the map $f\colon P^1(\mathbb{C}) \to P^1(\mathbb{C})$, extension of $z \mapsto z^2$. The domain is the outermost white sphere; the codomain is the innermost blue sphere; the map is the projection from the outermost to the innermost sphere. Notice the two ramification points 0 and ∞.
Right: the map $g\colon E \to P^1(\mathbb{C})$. Again, the domain is the outermost surface, a torus; the codomain is the innermost sphere. Notice the four ramification points p_1, p_2, p_3 and ∞.

As for the second property, consider the fibre $f^{-1}(y)$. From the local form of holomorphic functions, it is clear that all fibres are discrete[6]. Furthermore, f is proper because X is compact and Y is Hausdorff, so all fibres are compact and discrete. Since compact discrete spaces are finite, we conclude. The same argument holds for Ram_f. □

Lecture 6
Mar 28th, 2024

An immediate corollary, often referred to as Liouville's theorem, asserts that the only holomorphic functions from a compact Riemann surface X to a non-compact Riemann surface Y are the constant functions. Denoting by $\mathcal{O}(X)$ the $\mathbb{C}$-algebra of holomorphic functions on X, that is

$$\mathcal{O}(X) := \{\, f\colon X \to \mathbb{C} \mid f \text{ is holomorphic} \,\}\,, \tag{2.28}$$

we then obtain the following result.

Corollary 2.15 (Holomorphic functions on compact Riemann surfaces)**.** *The space of holomorphic forms on a compact Riemann surface X is 1-dimensional:*

$$\mathcal{O}(X) \cong \mathbb{C}\,. \tag{2.29}$$

Remark 2.16 ('How many functions are out there?')**.** We can think about the above theorem as an answer to 'how many holomorphic functions are there on a compact Riemann surface?'. The answer is: there is only one linearly independent holomorphic function, namely the constant

[6] A subset $D \subset X$ is discrete if every point in D has a neighbourhood in X that contains no other points of D.

function 1. This is in contrast with the smooth case: thanks to the existence of bump functions, it is possible to show that the vector space of real smooth functions $f\colon M \to \mathbb{R}$ on any (compact) real n-manifold is infinite dimensional. In the next sections, we will generalise the above result by relaxing the holomorphicity condition, counting meromorphic functions on compact Riemann surface. The result will still be an finite dimensional vector space, whose dimension is given by the celebrated Riemann–Roch formula.

We are now armed to define the most important invariant of holomorphic functions between compact Riemann surfaces: the degree.

Theorem 2.17 (Degree of holomorphic functions)**.** *Let $f\colon X \to Y$ be a non-constant holomorphic function between compact Riemann surfaces. The number of points in the fibre, counted with multiplicity, is constant:*

$$d := \sum_{x \in f^{-1}(y)} \mu_x(f) \qquad \text{is independent of } y \in Y. \tag{2.30}$$

The positive integer d is called the degree of f, denoted $\deg(f)$*. For constant functions, the degree is set to be zero.*

Proof. Fix a point $\bar{y} \in Y$, and let $f^{-1}(\bar{y}) = \{\bar{x}_1, \ldots, \bar{x}_n\}$. By appropriately restricting local charts and and changing variables, we can find disjoint neighbourhoods U_i of $\bar{x}_i$ and V of $\bar{y}$ such that $f^{-1}(V) \subset U_1 \sqcup \cdots \sqcup U_n$ and locally f is expressed as $z \mapsto z^{k_i}$ around $\bar{x}_i$. In this case, the degree at $\bar{y}$ is given by $d := k_1 + \cdots + k_n$.

We can now show that the function

$$Y \longrightarrow \mathbb{Z}, \qquad y \longmapsto d(y) := \sum_{x \in f^{-1}(y)} \mu_x(f) \tag{2.31}$$

is constant on V. For any $y \in V$, $f^{-1}(y) \subset U_1 \sqcup \cdots \sqcup U_n$ where we have a local description of the function as $z \mapsto z^{k_i}$ on U_i. In this case, the fibre intersected with U_i contains

- one point of multiplicity k_i if the point corresponds to 0,
- k_i point of multiplicity 1 otherwise.

In any case, the sum of points with multiplicity is $d(y) = k_1 + \cdots + k_n = d$. This proves that $d(y)$ is locally constant, and therefore globally constant due to the connectedness of Y. $\square$

The general philosophy is that, when counting properly, magic things happen. In this case 'counting properly' means counting elements in the fibres according to their multiplicity. By doing so, the total number is constant! The theorem also gives a formula for the cardinality of the fibres. Indeed, the equation $d = \sum_{x \in f^{-1}(y)} \mu_x(f)$, valid for any $y \in Y$, can be rearranged as

$$|f^{-1}(y)| = d - \sum_{x \in f^{-1}(y)} \big(\mu_x(f) - 1\big)\,. \tag{2.32}$$

The manipulation is legit, since fibres are always finite. In particular, if y is not a branch point, $|f^{-1}(y)| = d$.

Remark 2.18. If $f\colon X \to Y$ is a degree 1, non-constant holomorphic map between Riemann surfaces, then f is a biholomorphism. Indeed, f is surjective, and degree 1 means that f is also injective. A bijective holomorphic map is automatically a biholomorphism, thank to the open mapping theorem.

In subsection 1.2.4, we introduced the concept of genus, a topological invariant for any connected compact surface. Despite its topological nature, the genus is closely intertwined with the holomorphic structure of a Riemann surface, as highlighted in one of the most significant examples: the Riemann–Hurwitz formula.

Theorem 2.19 (Riemann–Hurwitz formula)**.** *Let $f\colon X \to Y$ be a holomorphic map of degree d between compact Riemann surfaces of genus g_X and g_Y respectively. Then*

$$2g_X - 2 = d(2g_Y - 2) + \sum_{x \in \mathsf{Ram}_f} \big(\mu_x(f) - 1\big)\,. \tag{2.33}$$

Adolf Hurwitz (1859–1919)

Proof. First, notice that the Riemann–Hurwitz formula can be restated in terms of Euler characteristics as

$$\chi_X = d\,\chi_Y - \sum_{x \in \mathsf{Ram}_f} \big(\mu_x(f) - 1\big)\,. \tag{2.34}$$

As the Euler characteristic is defined in terms of good graphs, a natural strategy is to choose a suitable good graph on Y and "lift" it to a good graph on X which we use to compute. A suitable choice of good graph Γ_Y in Y is such that $\mathsf{Brnch}_f \subseteq V_{\Gamma_Y}$. We define the graph Γ_X to be the pullback of Γ_Y via f, that is $\Gamma_X := f^{-1}(\Gamma_Y)$ with the natural graph structure. It is not hard to check that Γ_X is indeed a good graph in X. Furthermore, since the branch locus is contained in V_{Γ_Y}, we deduce from equation (2.32) the relations $|E_{\Gamma_X}| = d\,|E_{\Gamma_Y}|$ and $|F_{\Gamma_X}| = d\,|F_{\Gamma_Y}|$. On the other hand, if $y \in V_Y$ is a vertex we can still use equation (2.32) to get

$$|V_{\Gamma_X}| = \sum_{y \in V_{\Gamma_X}} |f^{-1}(y)| = d\,|V_{\Gamma_Y}| - \sum_{y \in V_{\Gamma_X}} \sum_{x \in f^{-1}(y)} \big(\mu_x(f) - 1\big) = d\,|V_{\Gamma_Y}| - \sum_{x \in \mathsf{Ram}_f} \big(\mu_x(f) - 1\big)\,. \tag{2.35}$$

All together, we find

$$\begin{aligned} \chi_X &= |V_{\Gamma_X}| - |E_{\Gamma_X}| + |F_{\Gamma_X}| \\ &= d\,|V_{\Gamma_Y}| - \sum_{x \in \mathsf{Ram}_f} \big(\mu_x(f) - 1\big) - d\,|E_{\Gamma_Y}| + d\,|F_{\Gamma_Y}| \\ &= d\,\chi_Y - \sum_{x \in \mathsf{Ram}_f} \big(\mu_x(f) - 1\big)\,. \end{aligned} \tag{2.36}$$

This concludes the proof. □

The Riemann–Hurwitz formula has several immediate consequences.

Exercise 2.5. *Let $f\colon X \to Y$ be a non-constant holomorphic map between compact Riemann surfaces.*

- *Show that $g_X \geq g_Y$.*
- *Suppose that both X and Y have genus 1. Conclude that f, must be unramified.*
- *Suppose that $Y = P^1(\mathbb{C})$ and f is unramified. Conclude that X is biholomorphic to $\mathbb{P}^1(\mathbb{C})$.*
- *Show that if $g_X = g_Y \geq 2$, then X and Y are biholomorphic.*

In general, the Riemann–Hurwitz formula imposes a condition on the genera of the source and target, the degree of the map, and the type of ramification points it can have.

Let us observe the formula in action with the examples considered previously.

Example 2.20. The degree of the maps

$$\begin{array}{ccc} P^1(\mathbb{C}) & \xrightarrow{f} & P^1(\mathbb{C}) \\ [z_0 : z_1] & \longmapsto & [z_0^2 : z_1^2] \end{array} \qquad \begin{array}{ccc} E & \xrightarrow{g} & P^1(\mathbb{C}) \\ [z_0 : z_1 : z_2] & \longmapsto & [z_0 : z_2] \end{array} \tag{2.37}$$

from examples 2.12 and 2.13 is 2 in both cases. As a sanity check, let us verify the Riemann–Hurwitz formula (assuming that E has genus one). The first map has two ramification points, $\mathbf{0} = [0 : 1]$ and $\boldsymbol{\infty} = [1 : 0]$, both of multiplicity 2. The second map has four ramification points, say p_1, p_2, and p_3 (corresponding to the values α_i), and $\boldsymbol{\infty} = [0 : 1 : 0]$, all of multiplicity 2. Thus, we find

$$\underbrace{2g_{P^1(\mathbb{C})} - 2}_{=-2} = \underbrace{d_f\left(2g_{P^1(\mathbb{C})} - 2\right)}_{=-4} + \underbrace{\left(\mu_{\mathbf{0}}(f) + \mu_{\boldsymbol{\infty}}(f) - 2\right)}_{=+2} \tag{2.38}$$

in the first example, and

$$\underbrace{2g_E - 2}_{=0} = \underbrace{d_g\left(2g_{P^1(\mathbb{C})} - 2\right)}_{=-4} + \underbrace{\left(\mu_{p_1}(g) + \mu_{p_2}(g) + \mu_{p_3}(g) + \mu_{\boldsymbol{\infty}}(g) - 4\right)}_{=+4} \tag{2.39}$$

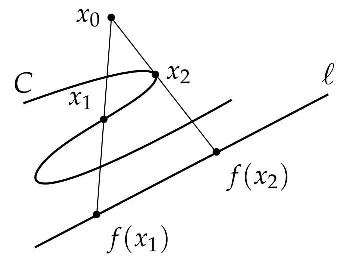
Exercise 2.6. *Consider the meromorphic function on $\mathbb{C}$ given by $f(z) = \frac{z^3}{1-z^2}$. Can you define a map $F\colon P^1(\mathbb{C}) \to P^1(\mathbb{C})$ that equals f in the affine chart $\{ [z : w] \mid w \neq 0 \}$? Show that F has degree 3, find its ramification points, and verify the Riemann–Hurwitz formula in this case.*

Exercise 2.7 (Fermat's Last Theorem for polynomials)**.** *Let $F, G, H \in \mathbb{C}[z, w]$ be non-constant, coprime, homogeneous polynomials such that $F^n + G^n = H^n$. Show that $n \leq 2$.*

As promised, we can now prove the genus-degree formula for smooth projective plane curves. We will assume the validity of Bézout's theorem, a classical theorem from algebraic geometry.

Theorem 2.21 (Bézout's theorem)**.** *Suppose that C_1 and C_2 are two plane projective curves of degree d_1 and d_2 respectively that do not have a common component (in other words, C_1 and C_2 are defined by polynomials without common divisor of positive degree). Then the total number of intersection points of C_1 and C_2, counted with their multiplicities, is equal to $d_1 \cdot d_2$.*

FIGURE 16. Albert Einstein, Adolf Hurwitz and his daughter Lisbeth Hurwitz, Zürich, August 1913. Photo credit: ETH Library.



Theorem 2.22 (Genus-degree formula)**.** *Let $C \subset P^2(\mathbb{C})$ be a smooth projective plane curve of degree d. Its genus is given by*

$$g_C = \frac{(d-1)(d-2)}{2}\,. \tag{2.40}$$

Proof. Consider a general point $x_0 \in P^2(\mathbb{C})$, a generic line $\ell \subset P^2(\mathbb{C})$, and the projection from $f\colon C \to \ell \cong P^1(\mathbb{C})$. It can be shown that f is holomorphic, so a natural question is: what is the degree of f? For $y \in \ell$, the fibre $f^{-1}(y)$ consists of all point in the intersection of C with the line passing through x_0 and y:

$$f^{-1}(y) = \mathsf{line}(x_0, y) \cap C\,. \tag{2.41}$$

As the line has degree 1 and C has degree d, the total number of points, counted with multiplicity, is equal to d. Moreover, the ramification points are exactly those $x \in C$ such that $\mathsf{line}(x_0, x)$ is the tangent line T_xC to C at x. Since x_0 is generic, one can show that every line passing through x_0 is tangent to C with multiplicity at most 2. Thus,

$$\sum_{x \in \mathsf{Ram}_f} \big(\mu_x(f) - 1\big) = |\mathsf{Ram}_f| = |\,\{\, x \in C \mid T_xC \text{ passes through } x_0 \,\}\,|\,. \tag{2.42}$$

Let F be the polynomial defining the curve C. The tangent line to C at x is given by the equation

$$T_xC = \left\{ [z_0 : z_1 : z_2] \in P^2(\mathbb{C}) \;\middle|\; \frac{\partial F}{\partial z_0}(x) z_0 + \frac{\partial F}{\partial z_1}(x) z_1 + \frac{\partial F}{\partial z_2}(x) z_2 = 0 \right\}. \tag{2.43}$$

Writing $x_0 = [a : b : c]$, we find that the conditions $x \in C$ and T_xC passing through x_0 are simply

$$\{\, x \in C \mid T_xC \text{ passes through } x_0 \,\} = \{\, F(x) = 0 \,\} \cap \left\{ \frac{\partial F}{\partial z_0}(x) a + \frac{\partial F}{\partial z_1}(x) b + \frac{\partial F}{\partial z_2}(x) c = 0 \right\}. \tag{2.44}$$

The polynomial F has degree d, while its derivative has degree $d-1$. By Bézout's theorem, the number of intersection points is then $d(d-1)$. To conclude, we apply Riemann–Hurwitz to obtain

$$2g_C - 2 = d\big(2g_\ell - 2\big) + \sum_{x \in \mathsf{Ram}_f} \big(\mu_x(f) - 1\big) = -2d + d(d-1)\,, \tag{2.45}$$

and the genus-degree formula follows. □

3. Meromorphic functions and divisors

In the first chapter, we have seen how meromorphic functions naturally show up when studying holomorphic functions. The main instance is Cauchy's integral formula: the value of a holomorphic function (and its derivatives) can be expressed as a contour integral of a meromorphic one. It is then natural to extend the study of holomorphic functions from $\mathbb{C}$ to Riemann surfaces.

From now onwards, we denote the Riemann sphere $P^1(\mathbb{C})$ simply as $\mathbb{P}^1$. More generally, we denote the complex projective space $P^n(\mathbb{C})$ as $\mathbb{P}^n$.

3.1. **Meromorphic functions.** Recall the definition of meromorphic functions on an open set $U \subseteq \mathbb{C}$: f is meromorphic if, for every point $z_0 \in U$, there exists $k \geq 0$ such that $(z-z_0)^k f(z)$ is holomorphic at z_0 but $(z-z_0)^{k-1} f(z)$ is not. If $k = 0$, the function is holomorphic at z_0; if $k > 0$, we say that f has a pole of order k at z_0. Notice that, since being holomorphic at a point z_0 implies holomorphicity in a neighbourhood, we deduce that poles must be isolated points.

It is easy to check that the set of meromorphic functions on U, denoted $\mathcal{M}(U)$, is a field. For $f \in \mathcal{M}(U)$ not identically zero and a fixed point $z_0 \in U$, we have the Laurent series expansion

$$f(z) = \sum_{n=m}^{\infty} a_n (z-z_0)^n \,, \qquad a_m \neq 0 \,, \tag{3.1}$$

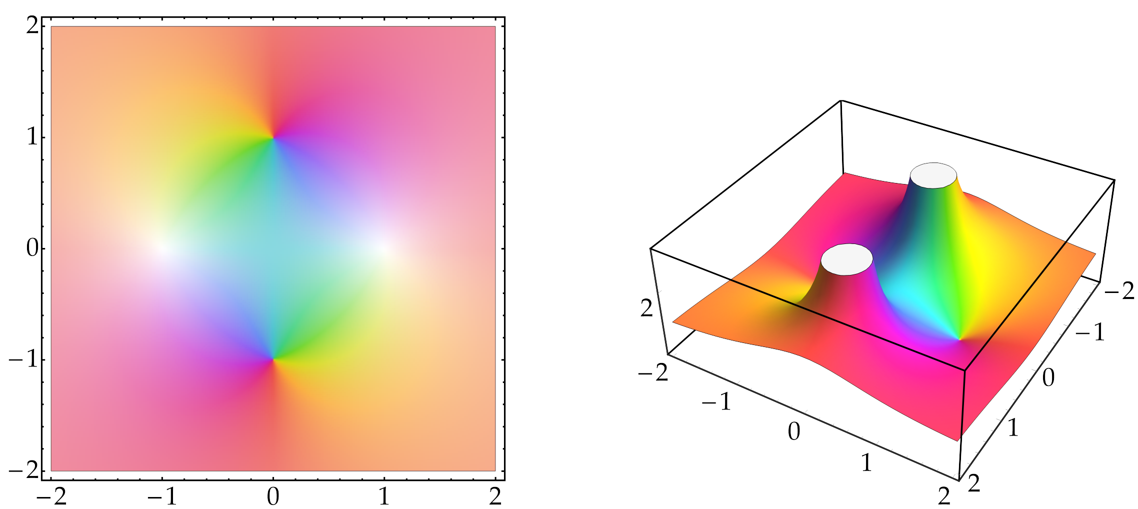


Figure 17. Meromorphic functions can be visualised by means of two techniques, 2d colouring and 3d colouring. In both cases, the argument $\arg(f(z))$ corresponds to a color on the color wheel: red for 0, lime for $\frac{\pi}{2}$, cyan for π, violet for $\frac{3\pi}{2}$. In the 2d colouring, the module $|f(z)|$ is represented via saturation: black is a zero, white is a pole. In the 3d colouring, the module $|f(z)|$ is represented on the z-axis. Above, the two representations of the function $f(z) = \frac{z^2+1}{z^2-1}$.

for some $m \in \mathbb{Z}$. Define the order of f at z_0 as

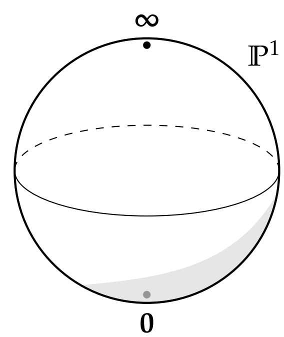
$$\mathrm{ord}_{z_0}(f) = m\,. \tag{3.2}$$

In accordance with our previous definition, we say that f has a zero of order m if $m > 0$, and it has a pole of order $-m$ if $m < 0$. We also set $\mathrm{ord}_{z_0}(f) = \infty$ for $f \equiv 0$. As a consequence, we get a map

$$\mathrm{ord}_{z_0} \colon \mathcal{M}(U) \longrightarrow \mathbb{Z} \cup \{\infty\}\,. \tag{3.3}$$

Exercise 3.1. *Prove that* $\mathrm{ord}_{z_0} \colon \mathcal{M}(U) \to \mathbb{Z} \cup \{\infty\}$ *is a discrete valuation, that is:*

- $\mathrm{ord}_{z_0}(f \cdot g) = \mathrm{ord}_{z_0}(f) + \mathrm{ord}_{z_0}(g)$*,*
- $\mathrm{ord}_{z_0}(f + g) \geq \min\{\, \mathrm{ord}_{z_0}(f), \mathrm{ord}_{z_0}(g)\,\}$*,*
- $\mathrm{ord}_{z_0}(f) = \infty$ *if and only if* $f = 0$*.*



Notice that a meromorphic function f on U, sometimes denoted as $f \colon U \dashrightarrow \mathbb{C}$, is not strictly speaking a function. Near a pole, a meromorphic function is unbounded, 'taking' the value ∞ at pole. Thinking about $\mathbb{C} \cup \{\infty\}$ as the Riemann sphere $\mathbb{P}^1$, it is then tempting to consider a meromorphic function on U as a holomorphic function to $\mathbb{P}^1$. This is indeed the case. To simplify the notation, set

$$\mathbf{0} = [0:1] \in \mathbb{P}^1 \qquad \text{and} \qquad \boldsymbol{\infty} = [1:0] \in \mathbb{P}^1 \tag{3.4}$$

for the point at zero and infinity on the Riemann sphere.

Lemma 3.1. *There is a one-to-one correspondence between meromorphic functions* $f \colon U \dashrightarrow \mathbb{C}$ *and holomorphic functions* $F \colon U \to \mathbb{P}^1$ *that are not constantly* $\boldsymbol{\infty}$*. Furthermore:*

- *zeros of f corresponds to $F^{-1}(\mathbf{0})$, poles of f corresponds to $F^{-1}(\boldsymbol{\infty})$,*
- $\mathrm{ord}_{z_0}(f) = \mu_{z_0}(F)$ *if* $F(z_0) = \mathbf{0}$*, and* $\mathrm{ord}_{z_0}(f) = -\mu_{z_0}(F)$ *if* $F(z_0) = \boldsymbol{\infty}$*.*

Proof. We restrict ourselves to the case finitely many poles. This would correspond to finitely many points on the fibre over $\boldsymbol{\infty}$, which is automatic when we are going to replace U by a compact Riemann surface. This said, denoted by $z_1, \ldots, z_n$ the poles of f, with orders $k_1, \ldots, k_n$. Then

$$g(z) = \left(\prod_{i=1}^{n} (z - z_i)^{k_i} \right) f(z) \tag{3.5}$$

is holomorphic and non-vanishing at the z_i's. Moreover, the zeros of f correspond to the zeros of g with the same multiplicity. Define

$$F \colon U \longrightarrow \mathbb{P}^1\,, \qquad z \longmapsto \left[g(z) : \prod_{i=1}^{n} (z - z_i)^{k_i} \right]\,. \tag{3.6}$$

The function F is holomorphic (being $g(z)$ and $\prod_{i=1}^{n} (z - z_i)^{k_i}$ on the two charts of $\mathbb{P}^1$). Moreover:

- points on the fibre over $\mathbf{0}$ correspond to zeros of g, which in turn correspond to zeros of f,
- points on the fibre over ∞ correspond to zeros of $\prod_{i=1}^{n}(z-z_i)^{k_i}$, that is to poles of f.

The correspondence of orders/multiplicity also follows. Conversely, given a holomorphic function $F\colon U \to \mathbb{P}^1$ with a finite fibre over ∞, we must have $F(z) = [g_1(z) : g_2(z)\prod_{i=1}^{n}(z-z_i)^{k_i}]$ with g_1 holomorphic and non-vanishing at the z_i's, and g_2 holomorphic and nowhere vanishing. Setting $f(z) = \frac{g_1(z)}{g_2(z)\prod_{i=1}^{n}(z-z_i)^{k_i}}$, we have the thesis. □ Happy Easter Break!

Lecture 7 Apr 11th, 2024

3.1.1. *Meromorphic functions on Riemann surfaces.* Thanks to the above characterisation, we can now define meromorphic functions on arbitrary Riemann surfaces.

Definition 3.2. A meromorphic function on a Riemann surface X is a holomorphic function $f\colon X \to \mathbb{P}^1$ that is not constantly ∞. Denote the field of meromorphic functions on X by $\mathcal{M}(X)$, and the subalgebra of holomorphic functions by $\mathcal{O}(X)$.

For a meromorphic function not identically zero, set

$$\operatorname{ord}_x(f) := \begin{cases} \mu_x(f) & \text{if } f(x) = \mathbf{0}\,, \\ -\mu_x(f) & \text{if } f(x) = \infty\,, \\ 0 & \text{otherwise.} \end{cases} \tag{3.7}$$

We say that x is a zero of order k if $\operatorname{ord}_x(f) = k > 0$, a pole of order k if $\operatorname{ord}_x(f) = -k < 0$. A zero or pole is called simple, double, triple, etc. if it is of order 1, 2, 3, etc. Again, the map $\operatorname{ord}_x\colon \mathcal{M}(X) \to \mathbb{Z}\cup\{\infty\}$ is a discrete valuation (setting $\operatorname{ord}_x(f) := \infty$ for $f \equiv 0$).

The next result states that for a meromorphic function on a compact Riemann surface the number of zeros equals the numbers of poles (if counted with multiplicity).

Lemma 3.3 (Functions: #zeros = #poles)**.** *Let X be a compact Riemann surface, f a non-constant meromorphic function. Then*

$$\sum_{x\in X}\operatorname{ord}_x(f) = 0\,. \tag{3.8}$$

Moreover, f has at least one zero and at least one pole.

Proof. First, the sum is well-defined since the only non-trivial contributions come from the fibres over $\mathbf{0}$ and ∞, which are finite as X is compact. Moreover,

$$\begin{aligned}\sum_{x\in X}\operatorname{ord}_x(f) &= \sum_{x\in f^{-1}(\mathbf{0})}\operatorname{ord}_x(f) + \sum_{x\in f^{-1}(\infty)}\operatorname{ord}_x(f) \\ &= \sum_{x\in f^{-1}(\mathbf{0})}\mu_x(f) - \sum_{x\in f^{-1}(\infty)}\mu_x(f)\end{aligned} \tag{3.9}$$

and the sum is zero since both summands equal $\deg(f)$. To conclude, recall that a non-constant holomorphic function with compact source is surjective. Hence, there exists at least one zero and at least one pole. □

3.1.2. *Meromorphic functions on the Riemann sphere and the torus.* Let us explore the case of meromorphic functions on our two main examples of compact Riemann surfaces: the Riemann sphere and the complex tori. For the Riemann sphere, meromorphic functions can be easily classified.

Exercise 3.2. *Prove that all meromorphic on $\mathbb{P}^1$ are of the form $[z : w] \mapsto [F(z,w) : G(z,w)]$ where $F, G \in \mathbb{C}[z,w]$ are homogeneous polynomials of the same degree and no common factors.*

As for the torus, the classification is more complicated but extremely beautiful. We start by noticing that a meromorphic function on a torus $\mathbb{C}/\Lambda$ is the same as a Λ-periodic meromorphic function on $\mathbb{C}$. That is, a holomorphic function $f \colon \mathbb{C} \to \mathbb{P}^1$ such that

$$f(z) = f(z+\omega_1) = f(z+\omega_2) \qquad \forall z \in \mathbb{C}\,, \tag{3.10}$$

where ω_1, ω_2 are the generators of the lattice.

This said, since meromorphic functions on compact Riemann surfaces have at least a pole, we can start by considering the case of simple poles.

Exercise 3.3 (Weierstraß gap theorem for the torus)**.** *Assume $f \colon \mathbb{C} \to \mathbb{P}^1$ is a Λ-periodic function with no poles on the boundary of the closed polygon $P \subset \mathbb{C}$ with vertices 0, ω_1, ω_2, $\omega_1+\omega_2$ (the fundamental domain). Show that the sum of the residues of all poles of f inside P is zero. Conclude that no such f can have a unique simple pole in P.*

Since meromorphic functions on tori with a single simple pole cannot exist, let us consider the case of a single double pole. The simplest Λ-periodic function with a single double pole is the Weierstraß $\wp$-function, defined for $z \notin \Lambda$ as

$$\wp(z) := \frac{1}{z^2} + \sum_{\omega\in\Lambda^\times} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right). \tag{3.11}$$

Here $\Lambda^\times := \Lambda \setminus \{0\}$. Notice that the Weierstraß $\wp$-function depends on the lattice Λ. For this reason, some authors denote it as $\wp(z,\Lambda)$.

Karl Theodor Wilhelm Weierstraß (1815–1897)

Proposition 3.4. *The Weierstraß $\wp$-function defines a meromorphic function on the complex torus $\mathbb{C}/\Lambda$ with a single double pole at $[0]$. Up to translation, scaling and adding constant, $\wp$ is the unique meromorphic function on $\mathbb{C}/\Lambda$ with a single double pole. It satisfies the differential equation*

$$(\wp')^2 = 4\,\wp^3 - g_2\,\wp - g_3\,, \tag{3.12}$$

where g_2 and g_3 are the coefficients of the Laurent expansion of $\wp$ at zero:

$$\wp(z) = \frac{1}{z^2} + \frac{1}{20} g_2\, z^2 + \frac{1}{28} g_3\, z^4 + \mathrm{O}(z^6)\,. \tag{3.13}$$

The proof is left as a guided exercise.

Exercise 3.4 (Weierstraß $\wp$-function). *Fix a lattice $\Lambda = \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2$, and consider*

$$\wp(z) = \frac{1}{z^2} + \sum_{\omega \in \Lambda^\times} \left(\frac{1}{(z-\omega)^2} - \frac{1}{\omega^2} \right), \tag{3.14}$$

the Weierstraß $\wp$-function associated to Λ.

(i) Show that the series defining $\wp$ converges absolutely for every $z \in \mathbb{C} \setminus \Lambda$ and uniformly on compact subsets of $\mathbb{C} \setminus \Lambda$.

Hint. *Use Weierstraß's M-test.*

(ii) Show that the derivative $\wp'$ is Λ-periodic and odd and that $\wp$ is even. Use this to show that $\wp$ is Λ-periodic. Conclude that $\wp$ is a meromorphic function on $\mathbb{C}/\Lambda$ with a single double pole at $[0]$.

(iii) Show that $\wp'$ has simple zeroes exactly at the points $[\frac{\omega_1}{2}], [\frac{\omega_1}{2}], [\frac{\omega_1+\omega_2}{2}] \in \mathbb{C}/\Lambda$.

Hint. *Use that $\wp'$ satisfies $\sum_{[z]} \operatorname{ord}_{[z]}(\wp') = 0$.*

(iv) Prove that every Λ-periodic function f with double poles exactly at the elements of Λ is of the form $f(z) = a\wp(z) + b$ for $a \in \mathbb{C}^\times, b \in \mathbb{C}$. Hence, up to translation, scaling and adding constant functions, $\wp$ is the unique meromorphic function on $\mathbb{C}/\Lambda$ with a single double pole.

Hint. *Use Weierstraß's gap theorem.*

(v) Let the Laurent expansion of (the even function) $\wp$ around $z = 0$ be given by

$$\wp(z) = \frac{1}{z^2} + c + \frac{1}{20} g_2 z^2 + \frac{1}{28} g_3 z^4 + \mathrm{O}(z^6). \tag{3.15}$$

Show that $c = 0$ and that $\wp$ satisfies the differential equation

$$(\wp')^2 = 4\wp^3 - g_2 \wp - g_3. \tag{3.16}$$

The differential equation satisfied by the Weierstraß $\wp$-function is reminiscent of the elliptic curve defined as $E_\Lambda := Z(z_1^2 z_2 - z_0^3 + \frac{g_2}{4} z_0 z_2^2 + \frac{g_3}{4} z_0^3) \subset \mathbb{P}^2$. Indeed, the differential equation implies that the following holomorphic map

$$\begin{aligned} \mathbb{C}/\Lambda &\xrightarrow{\quad \phi \quad} E_\Lambda \\ [z] &\longmapsto [\wp(z) : 2\wp'(z) : 1] \end{aligned} \tag{3.17}$$

is well-defined[7]. Here we set $\phi([0]) = [0:1:0]$.

Exercise 3.5 (Complex tori and elliptic curves). *Show that E_Λ is a smooth projective plane curve if and only if $\Delta := g_2^3 - 27g_3^2 \neq 0$. Conclude that ϕ is a biholomorphism.*

Since every smooth elliptic curve can be brought to the Weierstraß form $Z(z_1^2 z_2 - z_0^3 - az_0 z_2^2 - bz_0^3)$ with $\Delta = -16(4a^3 + 27b^2) \neq 0$, we deduce that every elliptic curve is isomorphic to a complex torus (and vice versa).

Moreover, one can show that ϕ is a group homomorphism. Let us check it in the easiest case: consider two points P and Q on the elliptic curve different form the point at infinity and not the inverse of each other. Say that $P = \phi([z])$ and $Q = \phi([w])$. We want to check that $-(P+Q) =$

[7]You can think of ϕ as the $2d$ version of the isomorphism $\mathbb{R}/2\pi\mathbb{Z} \to S^1$, $[t] \mapsto (\sin(t), \sin'(t)) = (\sin(t), \cos(t))$.

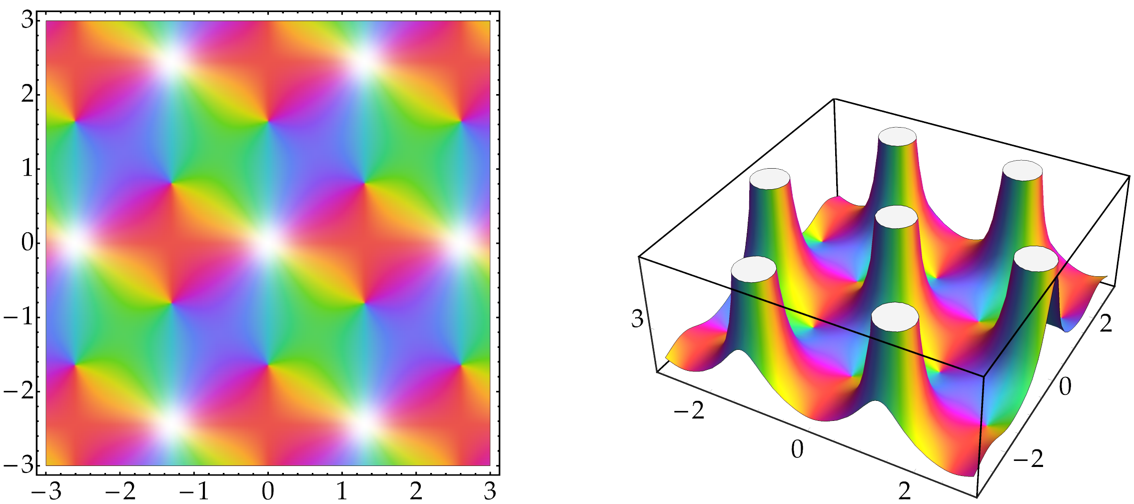


FIGURE 18. The Weierstraß $\wp$-function on the hexagonal lattice.

$\phi(-[z+w])$, and by definition $-(P+Q)$ is the only other point on the elliptic curve that is collinear to both P and Q. The collinearity condition is equivalent to show that the area formed by the triangle defined by the points $(\wp(z), 2\wp'(z))$, $(\wp(w), 2\wp'(w))$, and $(\wp(-z-w), 2\wp'(-z-w))$ is zero, which in turn is equivalent to

$$\det\begin{pmatrix} \wp(z) & 2\wp'(z) & 1 \\ \wp(w) & 2\wp'(w) & 1 \\ \wp(-z-w) & 2\wp'(-z-w) & 1 \end{pmatrix} = 0\,. \tag{3.18}$$

This is a consequence of a more general result that holds for meromorphic functions of complex tori, that can be interpreted as a special case of the Abel's theorem.

Proposition 3.5 (Abel's theorem for the torus)**.** *Let $f\colon \mathbb{C}/\Lambda \to \mathbb{P}^1$ be a non-constant meromorphic function. Then*

$$\sum_{x\in\mathbb{C}/\Lambda} \mathrm{ord}_x(f)\cdot x = [0]\,. \tag{3.19}$$

Here the sum is intended according to the group law on $\mathbb{C}/\Lambda$.

Proof. Without loss of generality, we can assume that the zeros and poles are located away from the boundary ∂P of fundamental domain (if not, we can simply shift). Consider now the meromorphic function $z\,\frac{f'(z)}{f(z)}$ defined in a neighbourhood of P. By the residue theorem, we have

$$\frac{1}{2\pi \mathrm{i}}\oint_{\partial P} z\,\frac{f'(z)}{f(z)}dz = \sum_{z_0 \text{ pole in } P} \operatorname*{Res}_{z=z_0}\, z\,\frac{f'(z)}{f(z)}\,. \tag{3.20}$$

On the other hand, the function $z\,\frac{f'(z)}{f(z)}$ is designed so that it has poles at the zeros and poles of f. Indeed, if f has a zero or a pole of order k at z_0, then the Laurent expansion of f and f' reads

$$\begin{aligned} f(z) &= a_k(z-z_0)^k + a_{k+1}(z-z_0)^{k+1} + \mathrm{O}\big((z-z_0)^{k+2}\big)\,, \\ f'(z) &= k\,a_k(z-z_0)^{k-1} + (k-1)\,a_{k+1}(z-z_0)^k + \mathrm{O}\big((z-z_0)^{k+1}\big)\,, \end{aligned} \tag{3.21}$$

with $a_k \neq 0$. Writing $z = z_0 + (z - z_0)$, we find

$$z \frac{f'(z)}{f(z)} = k\, z_0 \frac{1}{z - z_0} + \mathrm{O}(1)\,. \tag{3.22}$$

We conclude that the residue of $z \frac{f'(z)}{f(z)}$ at a zero or pole z_0 of f in P is $\mathrm{ord}_{z_0}(f) \cdot z_0$. Hence, the right-hand side of the residue computation is $\sum_z \mathrm{ord}_z(f) \cdot z$, where the sum is over all zeros and poles of f in P. On the other hand, the integral along ∂P can be broken into four integrals along the sides of ∂P. Using the double periodicity of f and f', and the fact that $\frac{f'(z)}{f(z)} = \frac{d \log f(z)}{dz}$, we find

$$\begin{aligned}
\frac{1}{2\pi \mathrm{i}} \oint_{\partial P} z \frac{f'(z)}{f(z)} dz &= \frac{1}{2\pi \mathrm{i}} \left[\int_{[0,\omega_1]-[\omega_2,\omega_1+\omega_2]} z \frac{f'(z)}{f(z)} dz - \int_{[0,\omega_2]-[\omega_1,\omega_1+\omega_2]} z \frac{f'(z)}{f(z)} dz \right] \\
&= -\frac{1}{2\pi \mathrm{i}} \left[\omega_2 \int_{[0,\omega_1]} \frac{f'(z)}{f(z)} dz - \omega_1 \int_{[0,\omega_2]} \frac{f'(z)}{f(z)} dz \right] \\
&= -\frac{1}{2\pi \mathrm{i}} \Big[\omega_2 \big(\log f(\omega_1) - \log f(0) \big) - \omega_1 \big(\log f(\omega_2) - \log f(0) \big) \Big]\,.
\end{aligned} \tag{3.23}$$

The properties of the complex logarithm imply that $\log f(\omega_1) - \log f(0) = -2\pi \mathrm{i} n_2$ and $\log f(\omega_2) - \log f(0) = 2\pi \mathrm{i} n_1$ for some integers n_1, n_2, so the contour integral equals $n_1\omega_1 + n_2\omega_2$. All together, the result of the residue theorem modulo Λ gives the thesis. $\square$

Let us go back to our problem. Consider the meromorphic function on the torus defined by

$$f \colon [u] \longmapsto \det \begin{pmatrix} \wp(z) & 2\wp'(z) & 1 \\ \wp(w) & 2\wp'(w) & 1 \\ \wp(u) & 2\wp'(u) & 1 \end{pmatrix}. \tag{3.24}$$

As a function of $[u] \in \mathbb{C}/\Lambda$, it has a pole of order 3 at $[0]$ and no other poles (since the only poles of $\wp$ and $\wp'$ are at $[0]$, of order 2 and 3 respectively), and it has simple zeros at $[u] = [z]$ and $[u] = [w]$. From the above proposition, we deduce that f has a third simple zero at $[u] = [-z-w]$, since zeros and poles must satisfy $\sum_{x \in \mathbb{C}/\Lambda} \mathrm{ord}_x(f) \cdot x = [0]$. Hence, we find that $f([-z-w]) = 0$, which is the thesis.

Bonus section (not required for the exam)

3.1.3. *More on elliptic functions.* There are many more interesting facts about the Weierstraß $\wp$-function, and more generally about meromorphic functions on the torus.

Elliptic integrals. The Weierstraß $\wp$-function is the inverse function of

$$z \longmapsto \int_z^\infty \frac{ds}{\sqrt{4s^3 - g_2 s - g_3}}\,. \tag{3.25}$$

Integrals of this form are called 'elliptic integrals', originally arising in connection with the problem of finding the arc length of an ellipse. This also explains the origin of the term 'elliptic curve'.

Field of meromorphic functions. The field of meromorphic functions on $\mathbb{C}/\Lambda$ with poles only at the identity $[0]$ is isomorphic to

$$\mathbb{C}(x)[y]/(y^2 - 4x^3 - g_2 x - g_3)\,, \tag{3.26}$$

with $x \mapsto \wp$ and $y \mapsto \wp'$.

Eisenstein series. As functions of the lattice Λ, the invariants g_2 and g_3 take the form

$$g_2(\Lambda) = -60 \sum_{\omega \in \Lambda^\times} \frac{1}{\omega^4}\,, \qquad g_3(\Lambda) = -140 \sum_{\omega \in \Lambda^\times} \frac{1}{\omega^6}\,. \tag{3.27}$$

They are part of a bigger family, called Eisenstein series:

$$G_k(\Lambda) \coloneqq \sum_{\omega \in \Lambda^\times} \frac{1}{\omega^k}\,. \tag{3.28}$$

Here we require $k > 2$ for convergence, and for k odd is it easy to see that G_k is identically zero. Hence, $g_2 = -60 G_4$ and $g_3 = -140 G_6$ are the first non-trivial Eisenstein series. They naturally appear as Laurent series coefficients of the Weierstraß function:

$$\wp(z) = \frac{1}{z^2} + \sum_{k \geq 1} (2k+1) G_{2k+2}\, z^{2k}\,. \tag{3.29}$$

Modular forms. For $\Lambda = \mathbb{Z} + \tau\mathbb{Z}$, we can interpret Eisenstein series as holomorphic functions on the upper-half plane $\mathbb{H}$. As such, they satisfy several amazing properties.

- Eisenstein series behave well under the action of the modular group:
$$G_k\left(\frac{a\tau + b}{c\tau + d}\right) = (c\tau + d)^k\, G_k(\tau)\,. \tag{3.30}$$
A holomorphic function on the upper-half plane which stays bounded as $\tau \to \mathrm{i}\infty$ and satisfies such property is called a modular form of weight k.
- Writing $q = e^{2\pi\mathrm{i}\tau}$, called the nome, we can consider the Taylor expansion around $q = 0$ (that is, a Fourier series):
$$G_{2k}(\tau) = 2\zeta(2k)\left(1 + \frac{2}{\zeta(1-2k)} \sum_{n=1}^{\infty} \sigma_{2k-1}(n)\, q^n\right), \tag{3.31}$$
where $\zeta(s)$ is the Riemann zeta function and $\sigma_m(n) = \sum_{d|n} d^m$ is the divisor sum function, the sum of the m-th powers of the divisors of n (see figure 19). The q-series can be resummed as a Lambert series, that is
$$\sum_{n=1}^{\infty} \sigma_{2k-1}(n)\, q^n = \sum_{n=1}^{\infty} n^{2k-1} \frac{q^n}{1 - q^n}\,. \tag{3.32}$$
It is customary to denote $E_{2k} \coloneqq \frac{G_{2k}}{2\zeta(2k)}$.
- The set M_k of modular forms of weight k is a complex vector space of dimension
$$\dim_{\mathbb{C}} M_k = \begin{cases} 0 & \text{if } k \text{ is odd,} \\ \lfloor k/12 \rfloor & \text{if is even and } k \equiv 2 \pmod{12}, \\ \lfloor k/12 \rfloor + 1 & \text{otherwise.} \end{cases} \tag{3.33}$$
The graded ring $M \coloneqq \bigoplus_k M_k$ is isomorphic, as an algebra over $\mathbb{C}$, to the polynomial ring in E_4 and E_6. Since the space of modular forms of weight $2k$ is one-dimensional for $2k =$

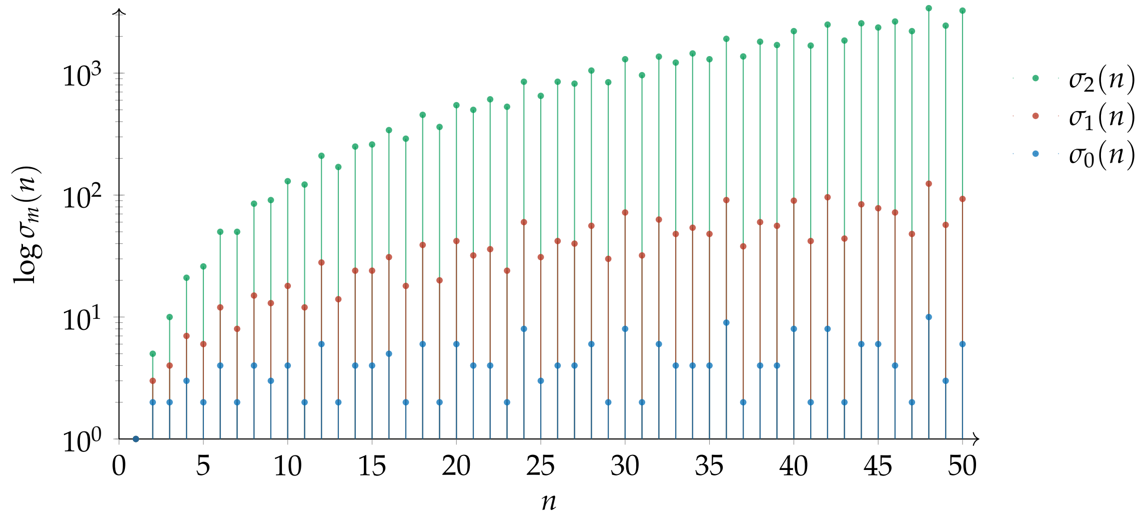


FIGURE 19. Log-linear plots of divisor sum functions σ_m for $m = 0, 1, 2$.

4, 6, 8, 10, 14, different products of Eisenstein series having those weights have to be equal up to a scalar multiple. In fact, we obtain the identities:

$$E_4^2 = E_8\,, \qquad E_4E_6 = E_{10}\,, \qquad E_4E_{10} = E_{14}\,, \qquad E_6E_8 = E_{14}\,. \tag{3.34}$$

Using the q-expansions of the Eisenstein series, they may be restated as identities involving the sums of powers of divisors. For instance, the first relation reads

$$\sigma_7(n) = \sigma_3(n) + 120 \sum_{m=1}^{n-1} \sigma_3(m)\sigma_3(n-m)\,. \tag{3.35}$$

Modular discriminant, Dedekind eta function, j-invariant. Dropping the holomorphicity condition, we obtain modular functions. There are meromorphic functions in the upper-half plane which behave well under the action of the modular group. The discriminant $\Delta = g_2^3 - 27g_3^2$ is a modular function of weight 12. It can be expressed as the infinite product

$$\Delta(\tau) = (2\pi \mathrm{i})^{12}\, q \prod_{k=1}^{\infty}(1-q^k)^{24}\,. \tag{3.36}$$

The presence of 24 is related to the fact that the celebrated Leech lattice has 24 dimensions. The function $\eta(\tau) := q^{1/24}\prod_{k\geq 1}(1-q^k)$ is called the Dedekind eta function, which is a modular function of weight $1/2$. The Fourier coefficients of Δ, that is

$$\Delta(\tau) = (2\pi \mathrm{i})^{12} \sum_{n\geq 1} \boldsymbol{\tau}(n)\, q^n\,, \tag{3.37}$$

defines a function $\boldsymbol{\tau}\colon \mathbb{N} \to \mathbb{Z}$ (not to be confused with the variable $\tau = \frac{1}{2\pi \mathrm{i}}\log(q) \in \mathbb{H}$) called the Ramanujan tau function (see figure 20). It satisfies some remarkable arithmetic properties, firstly conjectured by Ramanujan in 1916:

- $\boldsymbol{\tau}(mn) = \boldsymbol{\tau}(m)\boldsymbol{\tau}(n)$ if $\gcd(m,n) = 1$,
- $\boldsymbol{\tau}(p^{r+1}) = \boldsymbol{\tau}(p)\boldsymbol{\tau}(p^r) - p^{11}\boldsymbol{\tau}(p^{r-1})$ for p prime and $r > 0$,

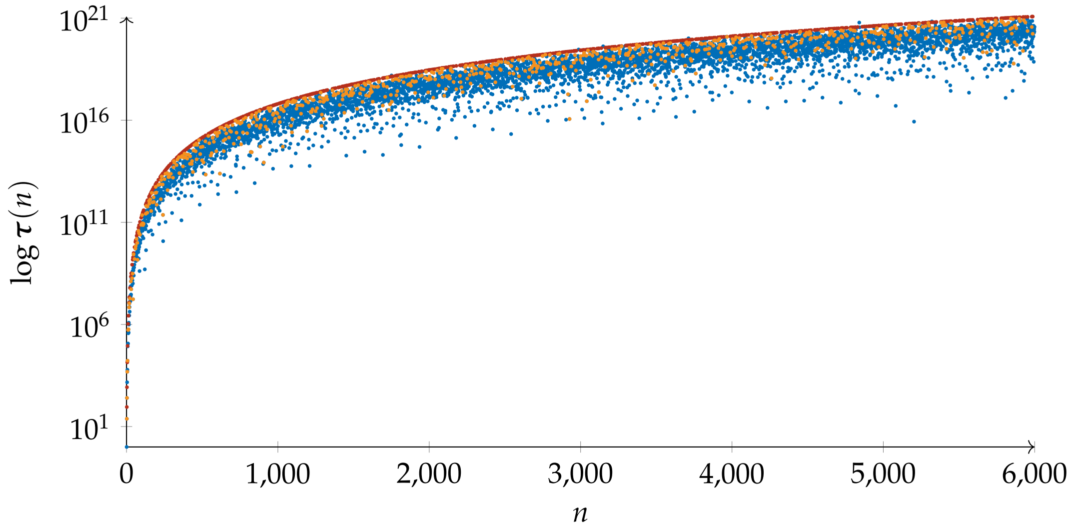


FIGURE 20. The log-linear plot of the absolute value of Ramanujan's tau function (in blue), with its values at prime numbers highlighted in orange and the values $2p^{11/2}$ for p prime in red. Ramanujan's conjecture state that all orange dots lays below the red ones.

- $|\boldsymbol{\tau}(p)| \leq 2p^{11/2}$ for all primes p.

The first two properties where proved by Mordell in 1917. The last property, known as Ramanujan's conjecture, was proved by Pierre Deligne in 1974 as a consequence of his proof of the Weil conjectures, a finite-field analogue of Riemann's hypothesis (Deligne received the Fields medal in 1978 for his proof).

Another interesting function is Klein's j-invariant, a modular function of weight 0:

$$j(\tau) := 1728 \, \frac{E_4(\tau)^3}{E_4(\tau)^3 - E_6(\tau)^2} \, . \tag{3.38}$$

The j-invariant defines a biholomorphism between the one-point compactification of the modular curve $\mathbb{H}/\mathrm{SL}(2,\mathbb{Z})$ (adding the point $\mathrm{i}\infty$) and the Riemann sphere $\mathbb{P}^1$. Apart from its relation with elliptic curves, it has surprising connections to the symmetries of the monster group, the largest sporadic simple group of order $2^{46} \cdot 3^{20} \cdot 5^9 \cdot 7^6 \cdot 11^2 \cdot 13^3 \cdot 17 \cdot 19 \cdot 23 \cdot 29 \cdot 31 \cdot 41 \cdot 47 \cdot 59 \cdot 71 \approx 8 \cdot 10^{53}$. This connection is referred to as monstrous moonshine (Richard Borcherds received the Fields medal in 1998 for his work on the subject).

Jacobi theta functions. Another fundamental function associated to the lattice $\mathbb{Z} + \tau\mathbb{Z}$ is the Jacobi theta function, defined as

$$\vartheta(z;\tau) := \sum_{n=-\infty}^{\infty} e^{\pi \mathrm{i}(n^2\tau + nz)} \, . \tag{3.39}$$

The above series converges uniformly on every compact of $\mathbb{C} \times \mathbb{H}$. In particular, for any fixed $\tau \in \mathbb{H}$, $\vartheta(z;\tau)$ is holomorphic in z. It can be shown that its zeros are all simple, and located at

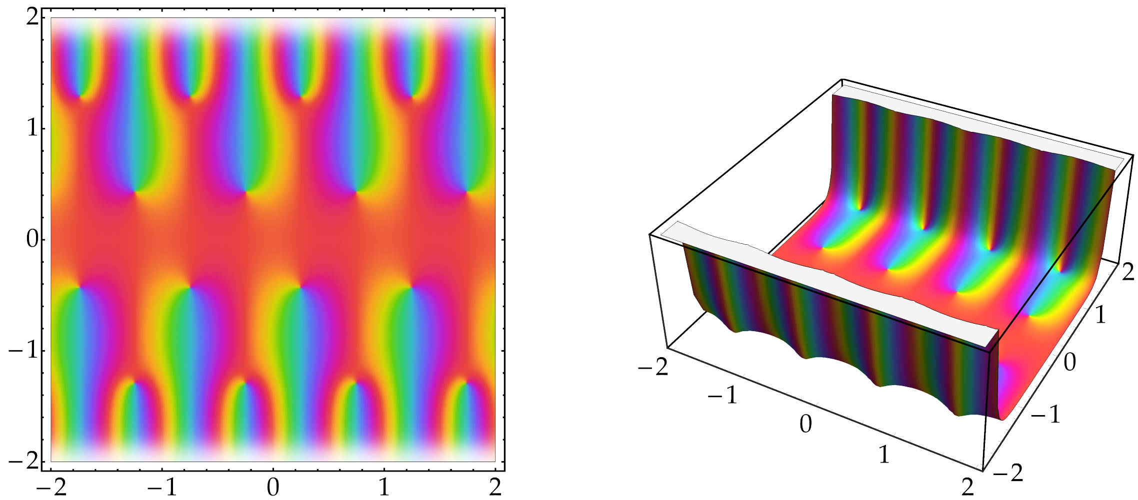


FIGURE 21. The Jacobi theta function on the square lattice.

$z = n\frac{1}{2} + m\frac{\tau}{2}$ for $n, m \in \mathbb{Z}$. However, ϑ does not quite define a function on the torus $\mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$: this is because it is quasi-periodic, meaning that

$$\vartheta(z+1;\tau) = \vartheta(z;\tau) \qquad \text{and} \qquad \vartheta(z+\tau;\tau) = e^{-\pi \mathrm{i}(\tau+2z)}\,\vartheta(z;\tau)\,. \tag{3.40}$$

The location of the zeros and the quasi-periodicity can be used to construct arbitrary meromorphic functions on the torus. Indeed, for any $x \in \mathbb{C}$, consider the translated theta function $\vartheta_{(x)}(z;\tau) := \vartheta(z - \frac{1}{2} - \frac{\tau}{2} - x;\tau)$. For fixed integers $m_1, \ldots, m_M, n_1, \ldots, n_N \geq 1$ and distinct points $x_1, \ldots, x_M, y_1, \ldots, y_N \in \mathbb{C}$ such that $\sum_i m_i x_i - \sum_j n_j y_j \in \mathbb{Z}$, it is easy to see that

$$\frac{\prod_{i=1}^{M} \vartheta_{(x_i)}(z;\tau)^{m_i}}{\prod_{j=1}^{N} \vartheta_{(y_j)}(z;\tau)^{n_j}} \tag{3.41}$$

defines a $(\mathbb{Z} + \tau\mathbb{Z})$-periodic meromorphic function with zeros at $x_1, \ldots, x_M$ of order $m_1, \ldots, m_M$ and poles at $y_1, \ldots, y_N$ of order $n_1, \ldots, n_N$. In particular, it defines a meromorphic function on the torus with prescribed zeros and poles.

Theta functions are intimately related to modular functions and the general theory of q-series. For instance, ϑ can be expressed as an infinite triple product as

$$\vartheta(z;\tau) = \prod_{n=1}^{\infty} (1 - q^{2n})(1 + e^{2\pi \mathrm{i} z} q^{2n-1})(1 + e^{-2\pi \mathrm{i} z} q^{2n-1})\,, \tag{3.42}$$

or more compactly as $\vartheta(z;\tau) = (q^2;q^2)_\infty\,(-e^{2\pi \mathrm{i} z} q;q^2)_\infty\,(-e^{-2\pi \mathrm{i} z} q;q^2)_\infty$ in terms of the q-Pochhammer symbol $(a;q)_\infty := \prod_{n\geq 1}(1 - aq^n)$.

3.2. **Divisors.** On complex tori, we have a natural notion of summation which has proven to be very fruitful in the study of meromorphic functions. Although arbitrary Riemann surfaces are not endowed with a natural group law structure, we can still define a formal group. Lecture 8 Apr 18th, 2024

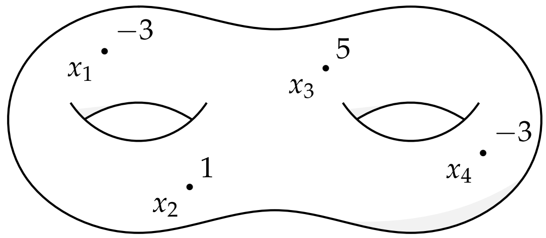


FIGURE 22. A degree-zero divisor on a genus 2 Riemann surface.

Definition 3.6. Let X be a Riemann surface. Define $\mathrm{Div}(X)$ as the free abelian group of points on X:

$$\mathrm{Div}(X) := \left\{ \sum_{x \in X} n_x[x] \;\middle|\; n_x \in \mathbb{Z},\ n_x = 0 \text{ for all but finitely many } x \right\}. \tag{3.43}$$

For $D = \sum_{x\in X} n_x[x]$, the (finite) set of points x such that $n_x \neq 0$ is called the support of D. Define the degree of $D = \sum_{x\in X} n_x[x]$ as $\deg(D) := \sum_{x\in X} n_x$. In particular, $\deg\colon \mathrm{Div}(X) \to \mathbb{Z}$ defines a group homeomorphism, and we denote by $\mathrm{Div}_0(X)$ its kernel, the subgroup of degree-zero divisors.

In other words, a divisor is simply a finite collection of points on a Riemann surface weighted by integers (see figure 22). The degree is simply the sum of all such integral weights.

3.2.1. *Principal divisors, linear equivalence, partial ordering.* Inspired again by the result for tori, it is natural to assign a divisor to any meromorphic function.

Definition 3.7. For a non-zero meromorphic function $f \in \mathcal{M}(X)$ on a compact X, define its associated divisor, divisor of zeros, and divisor of poles as

$$\begin{aligned} \mathrm{div}(f) &:= \sum_{x\in X} \mathrm{ord}_x(f)\,[x]\,, \\ \mathrm{div}_0(f) &:= + \sum_{\substack{x\in X \\ \mathrm{ord}_x(f)>0}} \mathrm{ord}_x(f)\,[x]\,, \\ \mathrm{div}_\infty(f) &:= - \sum_{\substack{x\in X \\ \mathrm{ord}_x(f)<0}} \mathrm{ord}_x(f)\,[x]\,. \end{aligned} \tag{3.44}$$

Clearly, $\mathrm{div}(f) = \mathrm{div}_0(f) - \mathrm{div}_\infty(f)$. We also set $\mathrm{div}(f) = 0$ for $f \equiv 0$. Every divisor of the form $D = \mathrm{div}(f)$ for a meromorphic function f is called principal. The set of principal divisors on X is denoted by $\mathrm{PDiv}(X)$. Notice that, as $\sum_x \mathrm{ord}_x(f) = 0$ for meromorphic functions on compact Riemann surfaces, we have that $\deg(D) = 0$ for all D principal divisors.

Remark 3.8. Do not confuse the degree of a meromorphic function $f\colon X \to \mathbb{P}^1$ with the degree of its divisor $\mathrm{div}(f)$. The latter is always zero. The degree of f is equal to the degree of its divisor of zeros and divisor of poles: $\deg(f) = \deg(\mathrm{div}_0(f)) = \deg(\mathrm{div}_\infty(f))$.

As $\mathrm{ord}_x\colon \mathcal{M}(X) \to \mathbb{Z} \cup \{\infty\}$ is a discrete valuation, The following result immediately follows.

Lemma 3.9. *Let f, g be meromorphic functions on a compact Riemann surface X.*

- $\operatorname{div}(f \cdot g) = \operatorname{div}(f) + \operatorname{div}(g)$,
- $\operatorname{div}(1/f) = -\operatorname{div}(f)$,
- $\operatorname{div}(f/g) = \operatorname{div}(f) - \operatorname{div}(g)$.

In particular, $\operatorname{PDiv}(X)$ *is a subgroup of* $\operatorname{Div}(X)$. *In fact, it is a subgroup of* $\operatorname{Div}_0(X)$, *the group of degree zero divisors.*

Definition 3.10. For a compact Riemann surface X, define the Picard group[8] as the quotient $\operatorname{Pic}(X) := \operatorname{Div}(X)/\operatorname{PDiv}(X)$. In other words, two divisors D and E define the same class element if and only if there exists a meromorphic function on X such that

$$D - E = \operatorname{div}(f)\,. \tag{3.45}$$

In this case, we say that D and E are linearly equivalent, and write $D \sim E$.

Intuitively, the Picard group measures 'how far' a divisor is from being the divisor of a meromorphic function. A description of the Picard group is provided by the Abel–Jacobi theory presented in section 5.

Example 3.11 (Picard group of $\mathbb{P}^1$)**.** For $\mathbb{P}^1$, every degree-zero divisor is principal. Indeed, any such divisor is of the form

$$D = \sum_{i=1}^{M} m_i[x_i] - \sum_{j=1}^{N} n_j[y_j] \tag{3.46}$$

for some integers $m_1, \ldots, m_M, n_1, \ldots, n_N \geq 1$ such that $\sum_i m_i = \sum_j n_j$ and distinct points $x_1, \ldots, x_M$, $y_1, \ldots, y_N \in \mathbb{P}^1$. Denote $x_i = [a_i : b_i]$ and $y_j = [c_j : d_j]$. Setting

$$\begin{cases} F(z,w) = \prod_{i=1}^{M}(b_i z - a_i w)^{m_i}\,, \\ G(z,w) = \prod_{j=1}^{N}(d_j z - c_j w)^{n_j}\,, \end{cases} \qquad \text{and} \qquad \begin{array}{rcl} \mathbb{P}^1 & \xrightarrow{\quad f \quad} & \mathbb{P}^1 \\ [z:w] & \longmapsto & [F(z,w) : G(z,w)] \end{array}\,, \tag{3.47}$$

we find that f is a meromorphic function on $\mathbb{P}^1$ with $\operatorname{div}(f) = D$. On the other hand, the degree homomorphism $\deg\colon \operatorname{Div}(\mathbb{P}^1) \to \mathbb{Z}$ is clearly surjective (this is true for all compact Riemann surfaces), and the above observation shows that $\ker(\deg) = \operatorname{PDiv}(\mathbb{P}^1)$. Thus, we find that

$$\deg\colon \operatorname{Pic}(\mathbb{P}^1) \longrightarrow \mathbb{Z} \tag{3.48}$$

is a group isomorphism: $\operatorname{Pic}(\mathbb{P}^1) \cong \mathbb{Z}$.

Exercise 3.6 (Abel's theorem for the torus)**.** *On a torus $T = \mathbb{C}/\Lambda$, consider the map*

$$\mathrm{A}\colon \operatorname{Div}(T) \longrightarrow T\,, \qquad \sum_{x \in T} n_x[x] \longmapsto \sum_{x\in T} n_x\, x\,. \tag{3.49}$$

[8]In modern algebraic geometry, for a scheme X, the term 'Picard group' is usually reserved for the group of isomorphism classes of invertible sheaves. If X is nice enough (that is, Noetherian, integral, separated and regular in codimension 1) then the Picard group is isomorphic to the group of divisors modulo principal divisors. For Riemann surfaces the two notions coincide.

The sum on the right-hand side is intended according to the group law on T. Show that D is principal if and only if $\deg(D) = 0$ *and* $\mathrm{A}(D) = [0]$. *Conclude that two divisors* $D, E \in \mathrm{Div}(X)$ *are linearly equivalent if and only if* $\deg(D) = \deg(E)$ *and* $\mathrm{A}(D) = \mathrm{A}(E)$.

Hint. *Use Abel's theorem for the torus (proposition 3.5) and the existence of theta functions.*

Notice that, in general, meromorphic functions are determined by their associated divisor up to scalar. Indeed, if $f, g \in \mathcal{M}(X)$ are such that $\mathrm{div}(f) = \mathrm{div}(g)$, then $\mathrm{div}(f/g) = 0$, as ord_x is a discrete valuation. As a consequence, f/g has no zeros and no poles, which means that it must be a non-zero constant.

A useful property of the divisors of zeros and poles of a meromorphic function is that all coefficients are non-negative. It is then natural to generalise such notion as follows.

Definition 3.12. For a divisor $D = \sum_{x \in X} n_x[x]$ on a Riemann surface X, we write:

- $D \geq 0$ if and only if $n_x \geq 0$ for all $x \in X$,
- $D > 0$ if $D \geq 0$ and $D \neq 0$.

For two divisors D and E, we write $D \geq E$ if and only if $D - E \geq 0$ (and likewise for $>$). This defines a partial ordering on the group of divisors.

Notice that every divisor can be uniquely written as $D = D_+ - D_-$, where $D_+, D_- \geq 0$ and with disjoint support. They are called the positive and negative parts of D. For instance, if $D = \mathrm{div}(f)$ is principal, then $D_+ = \mathrm{div}_0(f)$ and $D_- = \mathrm{div}_\infty(f)$.

Exercise 3.7 (Pullback of divisors)**.** *Let* $\varphi\colon X \to Y$ *be a holomorphic map between compact Riemann surfaces. For any point* $y \in Y$, *define the pullback*

$$\varphi^*[y] := \sum_{x \in \varphi^{-1}(y)} \mu_x(\varphi) \cdot [x]\,, \tag{3.50}$$

and extend it by linearity to a group morphism $\varphi^*\colon \mathrm{Div}(Y) \to \mathrm{Div}(X)$. *Prove that* $\deg(\varphi^* D) = \deg(\varphi) \cdot \deg(D)$. *What happens when you pullback principal divisors?*

3.2.2. *Space of functions associated to a divisor.* We have seen how the language of divisors can be effectively used to translate geometric properties of meromorphic functions on Riemann surfaces in more algebraic terms. Along these lines, we introduce the following space of meromorphic functions.

Definition 3.13. For a divisor D on a compact Riemann surface X, define the space

$$\mathcal{L}(D) := \left\{ f \in \mathcal{M}(X)^\times \;\middle|\; \mathrm{div}(f) \geq -D \right\} \cup \{ 0 \}\,. \tag{3.51}$$

The space $\mathcal{L}(D)$ has a nice geometric interpretation: it is the set meromorphic functions whose zeros and poles are bounded by D. More precisely $f \in \mathcal{L}(D)$ if and only if for every $x \in X$:

- if the coefficient n_x in D is negative, f has a zero of order at least $-n_x$ at x,
- if the coefficient n_x in D is positive, f can have a pole of order at most n_x at x,
- if the coefficient n_x in D is zero, f is holomorphic at x.

Alternatively, we can think about the above conditions in terms of Laurent series. For any point $x \in X$, choose a local coordinate z centred at x. Then $f \in \mathcal{L}(D)$ if and only if its local Laurent expansion at x has no terms lower than z^{-n_x}.

For instance, $\mathcal{L}(0)$ is precisely the space of holomorphic functions on X.

Example 3.14. Let $X = \mathbb{P}^1$, $D = d[\infty]$, where $\infty = [1 : 0]$ is the point at infinity on the Riemann sphere and $d \in \mathbb{Z}$ is a fixed integer. By definition, a meromorphic function $f \in \mathcal{M}(X)^\times$ belongs to $\mathcal{L}(d[\infty])$ if and only if $\operatorname{div}(f) + d[\infty] \geq 0$. Write $\operatorname{div}(f) = \sum_{i=1}^{M} m_i[x_i] - \sum_{j=1}^{N} n_j[y_j]$ with $n_i, m_j > 0$. We have two cases.

- $d < 0$. In this case we need f to have a zero of order at least $-d$ at infinity (in order to cancel $d[\infty]$) and no pole. A function with no poles must be constant. Since it must have at least a zero, it must be identically zero.
- $d \geq 0$. In this case we need f to have poles only at ∞ of order at most d. Such meromorphic functions are parametrised as
$$f([z : w]) = [F(z, w) : w^d] \tag{3.52}$$
with F a homogeneous degree d polynomial.

In other words,
$$\mathcal{L}(d[\infty]) = \begin{cases} \{ 0 \} & \text{if } d < 0\,, \\ \mathbb{C}[z, w]_d & \text{if } d \geq 0\,, \end{cases} \tag{3.53}$$
where $\mathbb{C}[z, w]_d$ denotes the vector space of degree d homogeneous polynomials.

We collect here some properties of $\mathcal{L}(D)$.

Proposition 3.15. *Let X be a compact Riemann surface, $D, E \in \operatorname{Div}(X)$.*

(1) $\mathcal{L}(D)$ is a complex vector space.
(2) If $D \geq E$, then $\mathcal{L}(D) \supseteq \mathcal{L}(E)$.
(3) If $D \sim E$, there is a canonical isomorphism $\mathcal{L}(D) \cong \mathcal{L}(E)$.
(4) If $D \sim 0$ (that is, D is principal), then $\mathcal{L}(D) = \mathcal{O}(X) \cong \mathbb{C}$.
(5) If $\deg(D) < 0$, then $\mathcal{L}(D) = \{ 0 \}$.

Proof. The first property follows from ord_x being a discrete valuation. The second property follows immediately from the definition. As for the third property, let $g \in \mathcal{M}(X)^\times$ such that $D - E = \operatorname{div}(g)$. Set
$$\mu_g \colon \mathcal{L}(D) \longrightarrow \mathcal{L}(E)\,, \qquad f \longmapsto f \cdot g\,. \tag{3.54}$$

It is easy to check that the above map is well-defined (that is $f \cdot g \in \mathcal{L}(E)$ for any $f \in \mathcal{L}(D)$), it is linear, and its inverse is the multiplication by $1/g$. For the fourth property, $D \sim 0$ implies that $\mathcal{L}(D)$ is the space of holomorphic functions on X, which are simply the constant functions. For the last property, suppose that $\deg(D) < 0$ and assume there exists a non-zero $f \in \mathcal{L}(D)$. By definition, $\operatorname{div}(f) + D \geq 0$, but taking the degree we find

$$0 \leq \underbrace{\deg(\operatorname{div}(f))}_{=0} + \deg(D) = \deg(D)\,, \tag{3.55}$$

which is a contradiction. □

We can now prove a crucial property of $\mathcal{L}(D)$: a bound on its dimension. Let us start with a preliminary result comparing D and $D - [x]$.

Lemma 3.16 (Codimension 1 lemma). *Let X be a compact Riemann surface, $x \in X$ a point, and $D \in \operatorname{Div}(X)$ a divisor. Then either $\mathcal{L}(D - [x]) = \mathcal{L}(D)$, or $\mathcal{L}(D - [x])$ has codimension 1 inside $\mathcal{L}(D)$. In other words,*

$$\dim\big(\mathcal{L}(D)/\mathcal{L}(D - [x])\big) \leq 1\,. \tag{3.56}$$

Proof. Choose a local coordinate z centred at x, and let n be the coefficient of $[x]$ in D. For any $f \in \mathcal{L}(D)$, its local Laurent expansion at x is of the form $c_f z^{-n}$ plus higher order term, with $c_f \in \mathbb{C}$ (possibly zero). Consider the map

$$\phi \colon \mathcal{L}(D) \longrightarrow \mathbb{C}\,, \qquad f \longmapsto c_f\,. \tag{3.57}$$

Then ϕ is linear, with kernel being $\mathcal{L}(D - [x])$. If ϕ is identically zero, then $\mathcal{L}(D - [x]) = \mathcal{L}(D)$. Otherwise, ϕ is onto (being a non-trivial linear map to $\mathbb{C}$), hence $\mathcal{L}(D - [x])$ has codimension 1 inside $\mathcal{L}(D)$. □

Remark 3.17. It is important to notice that the value c_f from above does depend on the choice of local coordinates. Indeed, if $w = \tau(z)$ is a different local coordinate centred at x, then $\tau(z) = \lambda z + \mathrm{O}(z^2)$ for some $\lambda \neq 0$. Thus, the Laurent expansion of f in the w coordinate will be

$$c_f\big(\lambda z + \mathrm{O}(z^2)\big)^{-n} + \mathrm{O}(z^{-n-1}) = (\lambda \cdot c_f)z^{-n} + \mathrm{O}(z^{-n-1})\,. \tag{3.58}$$

In other words, the coefficient we are interested in gets rescaled by a non-zero scalar. Multiplication by a non-zero scalar is just a linear automorphism of $\mathbb{C}$, and it does not affect the kernel of the map ϕ. Hence, the above argument holds.

Lecture 9
Apr 25th, 2024

Theorem 3.18 (Finiteness theorem). *Let X be a compact Riemann surface and $D \in \operatorname{Div}(X)$ a divisor. Then $\mathcal{L}(D)$ is a finite-dimensional vector space. More precisely, setting $\ell(D) := \dim \mathcal{L}(D)$, we have*

$$\ell(D) \leq 1 + \deg(D_+). \tag{3.59}$$

where D_+ is the positive part of D.

Proof. We proceed by induction on the degree of D_+. If $\deg(D_+) = 0$, then $D_+ = 0$, so $\ell(D_+) = 1$ from property (4) of proposition 3.15. Since $D \le D_+$, we find $\ell(D) \le \ell(D_+) = 1 = 1 + \deg(D_+)$.

Suppose now that $\deg(D_+) > 0$ and that the bound holds for all divisors whose positive part has degree $< \deg(D_+)$. Let x be a point on the support of D_+, and consider $D - [x]$. Then the positive part of $D - [x]$ has degree $\deg(D_+) - 1$, so

$$\ell(D - [x]) \le 1 + (\deg(D_+) - 1) = \deg(D_+)\,. \tag{3.60}$$

Thanks to lemma 3.16, we have $\ell(D) \le 1 + \ell(D - [x])$, which concludes the proof. ☐

Exercise 3.8 (Complete linear systems)**.** *Let X be a compact Riemann surface, $D \in \mathrm{Div}(X)$. Define the complete linear system associated to D as*

$$|D| := \{\, E \in \mathrm{Div}(X) \mid E \ge 0 \text{ and } E \sim D \,\}\ . \tag{3.61}$$

Prove that the map

$$\mathbb{P}(\mathcal{L}(D)) \longrightarrow |D|\,, \qquad [f] \longmapsto \mathrm{div}(f) + D \tag{3.62}$$

is well-defined and is an isomorphism. Compute $|D|$ for:

- *$X = \mathbb{P}^1$ and $D = d[\infty]$;*
- *$X = \mathbb{C}/\Lambda$ and $D = [x]$.*

Exercise 3.9 (Complete linear systems and projective embeddings)**.** *Given a divisor D on a compact Riemann surface X, define the map*

$$\varphi_D \colon X \longrightarrow \mathbb{P}^n\,, \qquad x \longmapsto [f_0(x) : f_1(x) : \cdots : f_n(x)]\,, \tag{3.63}$$

where $f_0, \ldots, f_n$ form a basis[9] of $\mathcal{L}(D)$. A basic question is: when is φ_D and embedding?

- *We say that the complete linear system $|D|$ is free iff $\ell(D) - l(D - [x_0]) = 1$ for every x_0 in the support of D. Fix a point $x \in X$. Show that if $|D|$ is free, then there exists a basis $f_0, f_1, \ldots, f_n$ for $\mathcal{L}(D)$ such that $\mathrm{ord}_x(f_0) = -n_x$ and $\mathrm{ord}_x(f_i) > -n_x$ for $i \ge 1$.*
- *Let $|D|$ be free. Show that $\varphi_D(x) = \varphi_D(y)$ iff $\mathcal{L}(D - [x] - [y]) = \mathcal{L}(D - [x]) = \mathcal{L}(D - [y])$. Conclude that φ_D is injective iff $\ell(D - [x] - [y]) = \ell(D) - 2$ for all distinct $x, y \in X$.*
- *Conclude that, for $|D|$ be free, φ_D is an embedding iff $\ell(D - [x] - [y]) = \ell(D) - 2$ for all $x, y \in X$ (including coinciding points). In this case, $|D|$ is called very ample.*

[9] If the negative part of D is non-trivial, then the functions f_i have common zeros of certain orders. In this case, the common factors can be simplified, as coordinates on the projective space are well-defined up to global scaling. Thus, the definition of φ_D is well-posed.

4. Differential forms and Riemann–Roch theorem

In the previous chapter we have seen that, for a fixed compact Riemann surface X and a divisor D on it, the space $\mathcal{L}(D)$ of meromorphic functions with zero and poles bounded by D is a finite-dimensional complex vector space. Hence, a natural question arises:

How can we compute $\ell(D)$?

In other words, how many linearly independent meromorphic functions can we find with a specific zero/pole structure? The answer is given by the celebrated Riemann–Roch theorem.

Before discussing it, we introduce another useful object on Riemann surfaces: meromorphic differential forms. Again, the main motivation is Cauchy's integral formula: the value of a holomorphic function is determined by a contour integration of a meromorphic form.

4.1. **Meromorphic differential forms.**

Definition 4.1. A meromorphic (differential 1-) form on an open set $U \subseteq \mathbb{C}$ is an expression of the form

$$\omega = f(z)\, dz \tag{4.1}$$

where f is a meromorphic function. We say that ω is meromorphic in the coordinate z. We say that ω is a holomorphic form if f is holomorphic.

Remark 4.2. Notice that, if ω, η are meromorphic forms on U, so do $\omega + \eta$ and the product $f \cdot \omega$ for any meromorphic function f on U. In other words meromorphic forms on U constitute an $\mathcal{M}(U)$-module. Similarly, holomorphic forms on U constitute an $\mathcal{O}(U)$-module.

Suppose we have two meromorphic form $\omega_i = f_i(z_i)\, dz_i$ in $U_i \subseteq \mathbb{C}$ in the coordinate z_i ($i = 1, 2$), and a holomorphic map $\tau \colon U_1 \to U_2$, $z_2 = \tau(z_1)$. We say that ω_1 transforms into ω_2 under τ if

$$f_1(z_1) = f_2\big(\tau(z_1)\big)\, \tau'(z_1) \tag{4.2}$$

The definition is formulated to be compatible with the chain rule for integration, where $dz_2 = \tau'(z_1) dz_1$. Notice that, if τ is invertible, then ω_2 transforms into ω_1 under τ^{-1}.

4.1.1. *Meromorphic forms on Riemann surfaces.*

We are now ready to define meromorphic forms on Riemann surfaces.

Definition 4.3. A meromorphic form on a Riemann surface X is a collection of meromorphic forms $\{\omega_i\}$, one for each chart $\varphi_i \colon U_i \to V_i$ in the coordinate of V_i, such that for every pair of overlapping charts $U_i \cap U_j \neq \varnothing$, the form ω_i transforms into ω_j under the transition map $\varphi_{i,j} = \varphi_j \circ \varphi_i^{-1}$. We denote the $\mathcal{M}(X)$-module of meromorphic forms on X by $\mathcal{K}(X)$, and the $\mathcal{O}(X)$-module of holomorphic forms by $\Omega(X)$.

A simple way of constructing meromorphic forms goes as follows. Consider a meromorphic function $f \in \mathcal{M}(X)$. This is noting but a collection of meromorphic functions $\{ f_i(z_i) \}$, one for each chart (U_i, φ_i) of X. It is easy to check that the collection of meromorphic form $\{ f'(z_i)\, dz_i \}$ defines a meromorphic form on X, denoted df. Meromorphic forms obtained this way are called exact. It is also easy to check that the following Leibniz rule holds:

$$d(f \cdot g) = f \cdot dg + g \cdot df\,. \tag{4.3}$$

Exercise 4.1. *Prove that the only holomorphic form on $\mathbb{P}^1$ is the trivial one (identically zero). Prove that the only holomorphic form on a torus $\mathbb{C}/\Lambda$ is the one induced by scalar multiples of dz.*

Given a meromorphic form ω on a compact Riemann surface X and a point $x \in X$, choose local coordinates centred at x, so that locally around x the differential form is expressed as $f(z)dz$. We define the order of ω at x, denoted $\mathrm{ord}_x(\omega)$ as the order of f at $z = 0$. We say that ω has a zero of order m at x if $m = \mathrm{ord}_x(\omega) > 0$, and a pole of order m if $-m = \mathrm{ord}_x(\omega) < 0$.

The set of zeros and poles are discrete, and finite if X is compact. It is then natural to give the following definition.

Definition 4.4. For a meromorphic form ω on a compact Riemann surface X, define the divisor

$$\mathrm{div}(\omega) := \sum_{x \in X} \mathrm{ord}_x(\omega)\,[x]\,. \tag{4.4}$$

We also have the notion of $\mathrm{div}_0(\omega)$ and $\mathrm{div}_\infty(\omega)$ defined accordingly. Every divisor of the form $D = \mathrm{div}(\omega)$ for a non-trivial meromorphic form ω is called canonical. The set of canonical divisors on X is denoted by $\mathrm{KDiv}(X)$.

Notice that, for any meromorphic function f and meromorphic form ω on X, we have

$$\mathrm{div}(f \cdot \omega) = \mathrm{div}(f) + \mathrm{div}(\omega)\,. \tag{4.5}$$

In other words, if we add a principal divisor to a canonical divisor, the result is still canonical. In algebraic terms, $\mathrm{PDiv}(X) + \mathrm{div}(\omega) \subseteq \mathrm{KDiv}(X)$. The other inclusion is also true.

Lemma 4.5. *Let ω_1 and ω_2 be two meromorphic forms on a Riemann surface X, with ω_1 not identically zero. Then there exists a meromorphic function f such that $\omega_2 = f \cdot \omega_1$.*

Proof. For a local chart (U, φ) of X, write $\omega_i = g_i(z)\, dz$. We have that g_1 is not identically zero, since ω_1 is not. Define the meromorphic function f on X as g_2/g_1 on the local chart (U, φ). It is easy to check that f is well-defined (that is, it is independent of the choice of local coordinates), and $\omega_2 = f \cdot \omega_1$. □

Corollary 4.6. *The set of canonical divisor is a coset of the subgroup of principal divisors:* $\mathrm{KDiv}(X) = \mathrm{PDiv}(X) + \mathrm{div}(\omega)$. *In other words, any two canonical divisors are equivalent.*

Recall that a meromorphic function on a compact Riemann surface has vanishing degree. This is a consequence of the '# zeros = # poles' result for meromorphic functions. What about meromorphic forms? It is easy to find an example where the degree is non-zero.

Exercise 4.2. *Consider the meromorphic function* $f\colon \mathbb{P}^1 \to \mathbb{P}^1$, $[z : w] \mapsto [z : w]$. *Let* $\omega = df$. *Show that the degree of* $\operatorname{div}(\omega)$ *is* -2.

The general answer is given by the following result.

Lemma 4.7 (Forms: #zeros − #poles = $2g - 2$)**.** *Let* X *be a compact Riemann surface,* ω *a non-trivial meromorphic form. Then* $\deg(\operatorname{div}(\omega)) = 2g - 2$, *where* g *is the genus of* X.

Proof. In light of the fact that any two canonical divisors are equivalent, any meromorphic form would work. Assume for a moment that there exists a non-constant meromorphic function f on X, that is a holomorphic map $f\colon X \to \mathbb{P}^1$, and set $\omega = df$. Our goal is to compute $\deg(\operatorname{div}(\omega))$.

Let $d = \deg(f)$. We can assume without loss of generality that f is unramified at the pre-images of ∞ (if not, we can compose it with an automorphism of $\mathbb{P}^1$). As a consequence, f has exactly d simple poles, say $x_1, \dots, x_d$, which are double poles for ω (as $d(1/z) = -1/z^2 dz$). On the other hand, the zeros of ω are exactly the ramification points of f and $\mu_x(f) - 1 = \operatorname{ord}_x(\omega)$ (as $d(z^k) = kz^{k-1}dz$). As a consequence,

$$\operatorname{div}(\omega) = \sum_{x \in \mathsf{Ram}_f} (\mu_x(f) - 1)[x] - 2\big([x_1] + \cdots + [x_d]\big) . \tag{4.6}$$

In particular, the degree is $\deg(\operatorname{div}(\omega)) = \sum_{x \in \mathsf{Ram}_f} (\mu_x(f) - 1) - 2d$. On the other hand, the Riemann–Hurwitz formula implies that

$$2g_X - 2 = -2d + \sum_{x \in \mathsf{Ram}_f} (\mu_x(f) - 1) , \tag{4.7}$$

hence the thesis. □

The above proof relies on the assumption that, for any compact Riemann surface, there exists a non-constant meromorphic function f on X. Notice that the theorem is false if we substitute 'meromorphic' with 'holomorphic', so the assertion is highly non-trivial! Such result is known as Riemann's existence theorem, whose proof is omitted.

Theorem 4.8 (Riemann's existence theorem)**.** *Every compact Riemann surface admits a non-constant meromorphic function.*

Exercise 4.3 (Riemann–Hurwitz formula on divisors)**.** *Let* $f\colon X \to Y$ *be a holomorphic map between compact Riemann surfaces. Let* K_X *and* K_Y *be any canonical divisor on* X *and* Y *respectively, and define the ramification divisor as*

$$R_f := \sum_{x \in \mathsf{Ram}_f} (\mu_x(f) - 1)[x] \in \operatorname{Div}(X) . \tag{4.8}$$

Prove that

$$K_X \sim f^* K_Y + R_f \tag{4.9}$$

and deduce the Riemann–Hurwitz formula by taking degrees.

4.1.2. *Space of forms associated to a divisor.* As for meromorphic functions, it is natural to use divisors to bound the zeros and poles of a meromorphic form.

Definition 4.9. For a divisor D on a Riemann surface X, define the space

$$\mathcal{I}(D) := \left\{ \omega \in \mathcal{K}(X)^{\times} \;\middle|\; \operatorname{div}(\omega) \geq -D \right\} \cup \{ 0 \} . \tag{4.10}$$

A simple consequence of the fact that any two canonical divisors are equivalent is an isomorphism between $\mathcal{I}(D)$ and $\mathcal{L}(D+K)$, where $K \in \operatorname{KDiv}(X)$ is any canonical divisor. This is a special case of the celebrated Serre duality (Jean-Pierre Serre received the Fields Medal in 1954 for his work on sheaf theory).

Proposition 4.10 (Serre duality for Riemann surfaces)**.** *There is a natural isomorphism of vector spaces*

$$\mu_{\omega} \colon \mathcal{L}(D+K) \longrightarrow \mathcal{I}(D) , \qquad f \longmapsto f \cdot \omega , \tag{4.11}$$

where $K = \operatorname{div}(\omega) \in \operatorname{KDiv}(X)$ *is any canonical divisor. In particular,* $\mathcal{I}(D)$ *is finite dimensional.*

For the space $\mathcal{L}(D)$, we have a good understanding in the case of $D \sim 0$ principal: $\mathcal{L}(0) = \mathcal{O}(X)$ is the space of holomorphic functions, and $\mathcal{O}(X) \cong \mathbb{C}$ since the only holomorphic functions on a compact Riemann surface are the constant one. As for $\mathcal{I}(D)$, we can ask the same question: can we characterise $\mathcal{I}(0)$? From the definition, it is clear that

$$\mathcal{I}(0) = \Omega(X) \tag{4.12}$$

is the space of holomorphic differentials on X. For compact X, this must be finite dimensional. But contrary to the case of functions, we do not have yet a clear understanding of $\Omega(X)$. It is easy to see that we can have non-constant holomorphic forms: on the torus, dz is such an example. It turns out that we can completely characterise the space of holomorphic forms in terms of the genus of X.

Theorem 4.11 (Holomorphic forms on compact Riemann surfaces)**.** *The space of holomorphic forms on a compact Riemann surface X is g-dimensional:*

$$\Omega(X) \cong \mathbb{C}^{g} , \tag{4.13}$$

where g is the genus of X.

The proof is not required for the exam.

Sketch of the proof (for the reader familiar with differential calculus). We omit the most technical parts of the proof but provide a general idea for readers familiar with differential topology.

Step 1: harmonic forms. We start by considering the space of smooth real-valued forms on X (considered as a smooth 2-dimensional real manifold), that is expressions that locally (for $z = x + \mathrm{i}y$) look like

$$\nu \overset{\text{loc}}{=} p\, dx + q\, dy = \frac{p - \mathrm{i}q}{2} dz + \frac{p + \mathrm{i}q}{2} d\bar{z} \tag{4.14}$$

for $p = p(x,y)$ and $q = q(x,y)$ smooth real-valued functions. Here $\stackrel{\text{loc}}{=}$ means that the equality holds in local coordinates. On the space of smooth real-valued forms on X, we have the involution defined by the Hodge star:

$$\star\nu \stackrel{\text{loc}}{=} -q\,dx + p\,dy = -\mathrm{i}\frac{p-\mathrm{i}q}{2}dz + \mathrm{i}\frac{p+\mathrm{i}q}{2}d\bar{z}\,. \tag{4.15}$$

In turn, it defines a norm (the L^2-norm) as

$$\|\nu\|^2 = \int_X \nu \wedge \star\nu\,. \tag{4.16}$$

Here $\nu \wedge \star\nu \stackrel{\text{loc}}{=} (p^2+q^2)dxdy$. It can be checked that all definitions do not depend on the choice of local coordinates.

A smooth real-valued form ν is called harmonic if it is closed ($d\nu = 0$) and co-closed ($d\star\nu = 0$), which locally translates into

$$d\nu \stackrel{\text{loc}}{=} (-p_y + q_x)\,dxdy = 0\,, \qquad d\star\nu \stackrel{\text{loc}}{=} (q_y + p_x)\,dxdy = 0\,. \tag{4.17}$$

The Cauchy–Riemann equations implies that if ν is harmonic, then $\mathrm{i}\nu - \star\nu$ is holomorphic; conversely, if ω is holomorphic, then $\Im\omega$ is harmonic. A simple computations shows that such maps are inverse of each other. In other words, there is an isomorphism

$$\begin{array}{ccc} \mathcal{H}(X) & \rightleftarrows & \Omega(X) \\ \nu & \longmapsto & \mathrm{i}\nu - \star\nu \\ \Im\omega & \longleftarrow\!\shortmid & \omega \end{array} \tag{4.18}$$

as real vector spaces. Hence, we simply have to prove that $\mathcal{H}(X)$ has real dimension $2g$.

Step 2: homology. Consider the real vector space $H_1(X,\mathbb{R})$ generated by homology cycles. That is, the space generated by piecewise smooth oriented contours (i.e. closed paths), up to equivalence defined as $\gamma_1 \sim \gamma_2$ if and only if there exists a 2-cell c in X with $\partial c = \gamma_1 - \gamma_2$. The opposite of a contour is the contour oriented in the opposite direction. A classical result from algebraic topology states that $H_1(X,\mathbb{R})$ is a real vector space of dimension $2g$.

Step 3: Dirichlet's principle and Weyl's lemma. It can be shown that, for a fixed basis $\gamma = (\gamma_1,\dots,\gamma_{2g})$ of $H_1(X,\mathbb{R})$, the map

$$\mathcal{H}(X) \longrightarrow \mathbb{R}^{2g}\,, \qquad \nu \longmapsto \left(\oint_{\gamma_1}\nu,\dots,\oint_{\gamma_{2g}}\nu\right), \tag{4.19}$$

is a linear isomorphism. The proof goes as follows.

For the injectivity, suppose that all cycle integrals of ν vanish. Then ν is exact: $\nu = df$, with $f(x) = \int_{x_0}^x \nu$ (the definition does not depend on the choice of base-point $x_0 \in X$ nor on path of integration, since all cycle integrals of ν vanish). In this case, a simple computation shows that $\nu \wedge \star\nu = -d(f\cdot df)$, and applying Stokes theorem we obtain $\|\nu\|^2 = 0$, hence $\nu = 0$.

As for the surjectivity, fix $(\epsilon_1, \dots, \epsilon_{2g}) \in \mathbb{R}^{2g}$. One can show that the space of smooth forms satisfying $\oint_{\gamma_i} \nu = \epsilon_i$ is a non-empty affine space of the form $\eta_{(\gamma,\epsilon)} + E(X)$, where $\eta_{(\gamma,\epsilon)}$ is an (explicit, (γ, ϵ)-dependent) closed form and $E(X)$ is the space of exact forms on X. Through the general machinery of Hilbert spaces, one shows that the minimiser of the L^2-norm on such affine space exists. This is known as the Dirichlet's principle. Weyl's lemma states that such minimiser is then harmonic. $\square$

The above ideas led to a more general set of concepts known as Hodge theory, which provides an understanding of cohomology groups of complex projective varieties in terms of harmonic forms. One of the Millennium Prize Problems, the Hodge conjecture, revolves around harmonic forms and algebraic subvarieties.

Remark 4.12. The above strategy can be slightly modified to prove the following statement: given a compact Riemann surface X and two distinct points $x_1, x_2 \in X$, there exists a meromorphic form ω whose only poles are simple, located at x_1 with residue $+1$ and at x_2 with residue -1 (the residue of ω at x is the coefficient of $\frac{1}{z}dz$ in any local coordinate z centred at x).

With this result, we can give a simple proof of Riemann's existence theorem. On a Riemann surface X, fix three distinct points x_1, x_2, x_3. Let ω_1 be the meromorphic form with poles at x_1 and x_2, and residues $+1$ and -1. Let ω_2 be the meromorphic form with poles at x_2 and x_3, and residues $+1$ and -1. The $f = \frac{\omega_1}{\omega_2}$ is a meromorphic form on X with a simple pole at x_1 and a simple zero at x_3.

Exercise 4.4 (Holomorphic forms on projective curves)**.** *Let $X = Z(F) \subset \mathbb{P}^2$ be a smooth projective plane curve defined by a homogeneous polynomial $F(x, y, z)$ of degree $d \geq 3$. Let $f(u, v) = F(u, v, 1)$ define the associated affine curve.*

- *Prove that du and dv define meromorphic forms on X.*
- *Show that $f_u \, du = -f_v \, dv$ as meromorphic forms on X (here $f_u = \frac{\partial f}{\partial u}$ and $f_v = \frac{\partial f}{\partial v}$).*
- *Show that for any polynomial $p(u, v)$ of degree $\leq d - 3$, the form*

$$\omega_p = p(u, v) \frac{du}{f_v} \tag{4.20}$$

is holomorphic. How many polynomials of degree $\leq d - 3$ in two variables there are? How does it relate to the genus of X?

Gustav Adolph Roch (1839–1866)

Lecture 10
May 2nd, 2024

4.2. **The Riemann–Roch theorem.** We can finally state the Riemann–Roch theorem, arguably the most important result in the theory of Riemann surfaces.

Theorem 4.13 (Riemann–Roch)**.** *Let X be a compact Riemann surface, $D \in \mathrm{Div}(X)$ a divisor. Then*

$$\ell(D) - \ell(K - D) = \deg(D) + 1 - g\,, \tag{4.21}$$

where K is a canonical divisor and g is the genus of X.

The inequality $\ell(D) \geq \deg(D) + 1 - g$, known as Riemann's inequality, was originally proved by Riemann in 1857; the error term $\ell(K - D)$ was provided by his student Gustav Roch in 1865. The theorem was later generalised by Weil for vector bundles on Riemann surfaces, by Noether on complex surfaces, and by Hirzebruch for an arbitrary holomorphic vector bundle E on a compact complex manifold X of dimension n:

$$\chi(X, E) = \int_X \mathrm{ch}(E)\,\mathrm{td}(X)\,. \tag{4.22}$$

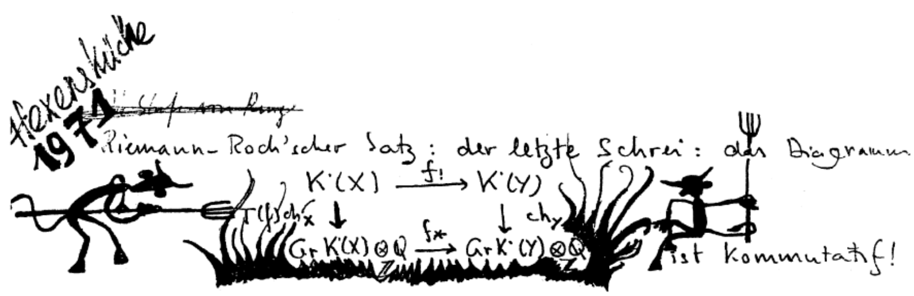
Here $\chi(X, E)$ is the Euler characteristic of E, $\mathrm{ch}(E)$ is its Chern character, and $\mathrm{td}(X)$ is the Todd class of X. If X is a compact Riemann surface and $E = \mathcal{O}(D)$ is a certain line bundle associated to a divisor D, the the left-hand side of Hirzebruch–Riemann–Roch's formula reduces to

$$\chi(X, \mathcal{O}(D)) = \ell(D) - \ell(K - D)\,, \tag{4.23}$$

while the right-hand side simplifies as

$$\int_X \mathrm{ch}(\mathcal{O}(D))\,\mathrm{td}(X) = \deg(D) + 1 - g\,. \tag{4.24}$$

Hirzebruch's result paved the way for the Grothendieck–Riemann–Roch theorem, a far-reaching result on coherent cohomology. The main change of perspective in Grothendieck's formulation is that the result is best formulated in terms of a map from one variety to another, $f\colon X \to Y$. Hirzebruch's version is retrieved by choosing $Y = \{\,*\,\}$. Grothendieck received the Fields Medal in 1966 for introducing the idea of K-theory, a key concept in the statement of his generalise Riemann–Roch theorem. His result was also fundamental in the development of the Atiyah–Singer index theorem, another generalisation of Riemann–Roch related to elliptic differential operator on compact manifolds (Atiyah received the Fields Medal alongside Grothendieck for his work). Below, Grothendieck's notes on his version of the theorem.



Witches Kitchen 1971. Riemann–Roch Theorem: The final cry: The diagram

$$\begin{tikzcd} K^{\cdot}(X) \arrow[r, "f_!"] \arrow[d, "\mathrm{T}(f)\,\mathrm{ch}_X"'] & K^{\cdot}(Y) \arrow[d, "\mathrm{ch}_Y"] \\ \mathrm{Gr}K^{\cdot}(X) \otimes_{\mathbb{Z}} \mathbb{Q} \arrow[r, "f_*"] & \mathrm{Gr}K^{\cdot}(Y) \otimes_{\mathbb{Z}} \mathbb{Q} \end{tikzcd}$$

is commutative!

4.2.1. *Consequences of Riemann–Roch.* Before giving a proof of Riemann–Roch's theorem, let us explore some consequences.

1) Riemann's existence theorem. Every compact Riemann surface admits a non-constant meromorphic function.

Proof. Choose a divisor D of degree $g+1$. Then $\ell(D) \geq 2$, which implies the existence of a non-constant meromorphic function. $\square$

2) Classification of genus 0 Riemann surfaces. If X has genus 0, it is biholomorphic to $\mathbb{P}^1$.

Proof. Fix a point $x \in X$, and consider the point divisor $D = [x]$. Then $\ell(D) \geq 2$, which means that there exists a non-constant function $f\colon X \to \mathbb{P}^1$ with a simple pole at x and no other poles. Hence, f has degree 1 (the degree of the divisor of poles), which implies that f is a biholomorphism. $\square$

3) Dimension of $\ell(k[x])$. Let $k \geq 0$ be an integer, $x \in X$ a point, and consider the sequence of dimensions $\ell_k := \ell(k[x])$. We already know that $\ell_0 = 1$. Riemann–Roch implies that $\ell_k = k+1-g$ for $k > 2g-2$, since $\ell(K-D) = 0$ in this range. In other words:

$$\begin{array}{c|cccccccc} k & 0 & 1 & \cdots & 2g-2 & 2g-1 & 2g & \to \\ \ell_k & 1 & ? & & ? & g & g+1 & \end{array} \tag{4.25}$$

4) Weierstrass's gap theorem. Let X be a compact Riemann surface of genus $g > 0$, and let $x \in X$ be a point. There are precisely g integers (depending on x) of the form

$$1 \leq k_1 < k_2 < \cdots < k_g < 2g \tag{4.26}$$

such that there exists no meromorphic function on X with a single pole of order k_i at x.

We can immediately check Weierstraß's gap theorem for complex tori: the result states that there is no meromorphic function on a complex torus with a single simple pole. This was precisely the statement of exercise 3.3.

Proof. We already know, thanks to Riemann–Roch, that $\ell(0) = 1$ and $\ell((2g-1)[x]) = g$. On the other hand, the finiteness theorem 3.18 states that

$$\ell(k[x]) - \ell((k-1)[x]) \leq 1\,. \tag{4.27}$$

Hence, $\ell_k := \ell(k[x])$ is a monotone increasing sequence of positive integers, which increases by at most 1 at each step, with $\ell_0 = 1$ and $\ell_{2g-1} = g$. Therefore, there must be a total of g integers in the range $1 \leq k < 2g$, say $1 \leq k_1 < k_2 < \cdots < k_g < 2g$ for which $\ell(k[x]) = \ell((k-1)[x])$. If k is such an integer, then the quotient space

$$\mathcal{L}(k[x])/\mathcal{L}((k-1)[x]) \tag{4.28}$$

is trivial. But this space is exactly the space of meromorphic functions with a single pole at x of order k. This proves the theorem. $\square$

Weierstraß's gap can be regarded as an obstruction result to the existence of meromorphic functions with specific pole structures. It is a special case of Noether's gap theorem, whose proof is omitted.

5) Noether's gap theorem. Let X be a compact Riemann surface of genus $g > 0$, and let $x_1, x_2, \ldots$ be a sequence of points in X. Define recursively the following sequence of divisors:

$$D_0 = 0\,, \qquad D_k = [x_k] + D_{k-1}\,. \tag{4.29}$$

We now pose the following sequence of questions:

Q_k: Is there a non-zero meromorphic function in $\mathcal{L}(D_k)/\mathcal{L}(D_{k-1})$?

Noether's gap theorem states that there are precisely g integers (depending on the sequence of points) of the form $1 \leq k_1 < k_2 < \cdots < k_g < 2g$ for which the answer to Q_k is no if and only if k is one of the integers in the list.

6) Compact Riemann surfaces are projective. Every compact Riemann surface X admits an embedding in projective space.

Proof. Let D be a divisor of degree $\deg(D) \geq 2g - 1$. Then $\deg(K - D) < 0$, so Riemann–Roch implies that $\ell(D) = \deg(D) + 1 - g$. As a consequence:

- if $\deg(D) \geq 2g$, then $|D|$ is free;
- if $\deg(D) \geq 2g + 1$, then $|D|$ is very ample.

Thus, every divisor D of degree $\deg(D) \geq 2g + 1$ induces an embedding $\varphi_D \colon X \hookrightarrow \mathbb{P}^{\ell(D)+1}$ (see exercise 3.9). $\square$

The above argument can be easily improved to show that every genus 1 Riemann surface is isomorphic to an elliptic curve. Since elliptic curves are isomorphic to complex tori, we deduce that the modular curve $\mathbb{H}/\mathrm{SL}(2, \mathbb{Z})$ from theorem 2.1 classifies all genus 1 Riemann surfaces.

7) Holomorphic q-differentials. For $q \in \mathbb{Z}$, a holomorphic differential of order q (or simply q-differential) on an open set $U \subset \mathbb{C}$ is an expression of the form

$$f(z)\, dz^q\,. \tag{4.30}$$

As in the case of holomorphic functions or differentials (corresponding to the cases $q = 0$ and $q = 1$ respectively), there is a natural notion of transformation under a holomorphic change of coordinate $w = \tau(z)$: it is the one induced by $dw^q = (\tau'(z))^q\, dz^q$. Hence, we have a well-defined notion of holomorphic q-differentials on a Riemann surface. It is easy to see that, for a compact

Riemann surface X with canonical divisor K, the space $\mathcal{L}(qK)$ is isomorphic to the space of q-differential on X. Riemann–Roch allows us to compute its dimension, in accordance to the following table.

Genus	Order	Dimension
$g=0$	$q \leq 0$	$1-2q$
	$q>0$	0
$g=1$	q	1
$g \geq 2$	$q<0$	0
	$q=0$	1
	$q=1$	g
	$q>1$	$(2q-1)(g-1)$

Proof. We want to compute $\ell(qK)$. Notice that $\deg(qK) = q(2g-2)$. Hence, we immediately conclude that $\ell(qK) = 0$ for

$$g = 0 \ \text{ and } \ q > 0 \qquad \text{or} \qquad g \geq 2 \ \text{ and } \ q < 0 \tag{4.31}$$

because the degree is negative in these cases. On the other hand, Riemann–Roch implies that

$$\ell(qK) = q(2g-2) + 1 - g + \ell((1-q)K) = (2q-1)(g-1) + \ell((1-q)K)\,. \tag{4.32}$$

Hence, we deduce the dimension for $g = 0$ and $q > 0$ or $g \geq 2$ and $q > 1$. The case $g \geq 2$ and $q = 0$ or $q = 1$ are known already: they correspond to holomorphic functions and holomorphic forms respectively.

For $g = 1$, Riemann–Roch gives no useful information. However, we can argue as follows. If $\omega \in \Omega(X)$ is a non-zero holomorphic form, then it cannot have zeros (since it has no poles and its degree is 0). Thus multiplication by ω gives an isomorphism between the space of q-differential and the space of $(q+1)$-differentials, while multiplication by $1/\omega$ gives an isomorphism between the space of q-differential and the space of $(q-1)$-differentials. Since the space of holomorphic 0-differentials (i.e. holomorphic functions) is one-dimensional, we conclude. □

8) Hyperelliptic curves. A Riemann surface X is called hyperelliptic if and only if it admits a degree two map to $\mathbb{P}^1$. These are generalisations of elliptic curves, where the degree two map to the Riemann sphere is the map given in example 2.13. It can be shown that every genus 0, genus 1, and genus 2 Riemann surface is hyperelliptic.

Proof. Since genus 0 and genus 1 curves are always biholomorphic to the Riemann sphere or an elliptic curve, in these cases the statement follows from examples 2.12 and 2.13 respectively. As for genus 2: the canonical divisor K_X of a genus 2 curve is such that $\ell(K_X) = 2$, which in turn gives a map $\varphi\colon X \to \mathbb{P}^1$ thanks to exercise 3.9. From the definition of the map and the fact that the degree of K_X is $2g-2=2$, one concludes that φ has degree two. □

4.2.2. *Proof of Riemann–Roch.* We propose a proof of the Riemann–Roch theorem by dividing it into three parts:

(1) Show that Riemann–Roch holds for divisors D satisfying $D \geq 0$.
(2) Prove the Riemann–Roch inequality $\ell(D) - \ell(K - D) \geq \deg(D) + 1 - g$ for all D.
(3) Deduce the Riemann–Roch theorem in full generality.

Before proceeding with the proof, we need to discuss the concept of residue on Riemann surfaces. First, notice that forms can be integrated along paths on Riemann surfaces:

$$\int_\gamma \omega \tag{4.33}$$

for $\gamma\colon [0,1] \to X$ a path and ω holomorphic along γ. This makes sense locally (as ω is of the form $f(z)dz$ for a holomorphic function f) and the integration is additive (the integral from a to b is equal to the integral from a to c plus the integral from c to b). So the above integral can be defined by breaking the path into pieces, each contained in local chart, perform each integral in local coordinates, and sum the contributions up. The integral in local coordinates does not depend on the choice thanks to the chain rule for integrals.

Now, if ω is a meromorphic form on a Riemann surface X, we define its residue at a pole $x \in X$ to be

$$\operatorname*{Res}_x \, \omega := \frac{1}{2\pi \mathrm{i}} \oint_{\gamma_x} \omega \,, \tag{4.34}$$

where γ_x is a small counter-clock wise oriented circle around x that contains x and no other poles. The definition does not depend on the choice of contour as in the case of $\mathbb{C}$. Notice that, on arbitrary Riemann surfaces, the residue is well-defined for meromorphic forms only (it is not well-defined for meromorphic functions). Another possibility is to define the residue as the coefficient of $z^{-1}dz$ in the Laurent expansion of ω in any local coordinate z centred at x. It can be shown that the definition does not depend on the choice of local coordinate. For instance, consider the case of a simple pole: $\omega = f(z)dz$ with $f(z) = c\,z^{-1} + \mathrm{O}(1)$, and take $w = \tau(z) = \lambda\, z + \mathrm{O}(z^2)$ a holomorphic change of coordinates centred at the origin (hence, $\lambda \neq 0$). Then

$$f(\tau(z))\tau'(z) = \left(\frac{c}{z(\lambda + \mathrm{O}(z))} + \mathrm{O}(1) \right) \big(\lambda + \mathrm{O}(z)\big) = c\, z^{-1} + \mathrm{O}(1) \,, \tag{4.35}$$

so the definition does not depend on the choice of coordinates. The proof can be generalised to poles of arbitrary order.

A remarkable property of compact Riemann surfaces, which is a special case of the residue theorem, is that the sum of all residues of a meromorphic form is equal to zero.

Lemma 4.14 (Residue theorem on compact Riemann surfaces)**.** *Let ω be a meromorphic form on a compact Riemann surface X with poles at $x_1, \dots, x_N$. Then*

$$\sum_{i=1}^{N} \operatorname*{Res}_{x_i} \, \omega = 0 \,. \tag{4.36}$$

The proof employs a simple trick: it interprets the sum of the residues across the entire compact Riemann surface as a Cauchy integral along a contour oriented in the opposite direction, resulting in zero because it encircles no poles.

We are now ready to prove the Riemann–Roch theorem.

Part 1. If $D = 0$, then Riemann–Roch follows from $\mathcal{L}(0) \cong \mathbb{C}$ and $\mathcal{L}(K) \cong \mathbb{C}^g$. Take then $D > 0$, that is $D = \sum_{i=1}^N n_i[x_i]$ with $n_i > 0$ for all $i = 1, \ldots, N$. Define the complex vector space V spanned by N-tuples $(f_1, \ldots, f_N)$ of the form

$$f_i = \frac{c_{i,n_i}}{z^{n_i}} + \cdots + \frac{c_{i,1}}{z} \tag{4.37}$$

for some constants $c_{i,n_i}, \ldots, c_{i,1} \in \mathbb{C}$. Clearly, $\dim V = \deg(D)$. Define the linear map (the 'principal part map')

$$\pi \colon \mathcal{L}(D) \longrightarrow V\,, \tag{4.38}$$

that sends a meromorphic function f to the N-tuple $(f_1, \ldots, f_N)$ of negative-power portion of the Laurent series at $(x_1, \ldots, x_N)$, where the Laurent series is taken with respect to some local coordinate[10] around x_i. The function f_i is called the principal part of f at x_i (with respect to the fixed local coordinates). Notice that, for any $f \in \mathcal{L}(D)$, its Laurent series at x_i is a $\mathrm{O}(z^{-n_i})$; hence, the definition is well-posed. By the rank-nullity theorem from linear algebra,

$$\ell(D) = \dim \mathcal{L}(D) = \dim \ker(\pi) + \dim \operatorname{im}(\pi)\,. \tag{4.39}$$

Notice that $f \in \ker(\pi)$ if and only if f has no poles, hence it is constant. Thus, $\ker(\pi) = \mathbb{C}$ and $\ell(D) = 1 + \dim \operatorname{im}(\pi)$. Our goal is then to understand $W := \operatorname{im}(\pi)$, and in particular to show that $\dim W = \deg(D) - g + \ell(K - D)$. Geometrically, we are looking for a meromorphic function f with prescribed principal parts $(f_1, \ldots, f_N)$ at $(x_1, \ldots, x_N)$ and no other poles. This is known as the Mittag-Leffler's problem. Solutions to the Mittag-Leffler's problem can be characterised as follows.

Magnus Gösta Mittag-Leffler (1846–1927)

Lemma 4.15. *The following are equivalent:*

- *There exists a meromorphic function f with prescribed principal parts $(f_1, \ldots, f_N)$ at $(x_1, \ldots, x_N)$ and no other poles.*
- $\sum_{i=1}^N \operatorname{Res}_{x_i} f_i \cdot \omega = 0$ *for all holomorphic form ω in X.*

Sketch of the proof. Let f be a solution to Mittag-Leffler's problem. According to the residue theorem, the sum of the residues, $\sum_{i=1}^N \operatorname{Res}_{x_i}(f \cdot \omega)$, equals zero for any holomorphic form ω on X.

[10] It is important to notice that Laurent expansion coefficients of a meromorphic function depend on the choice of local coordinates. Here we fix once and for all local coordinates centred around each x_i, and consider all expansions with respect to such choice. However, different choices would produce the same result. For instance, a different choice would give a new map $\pi' \colon \mathcal{L}(D) \to V$ that is related to π by a linear automorphism of V. In particular, the quantities we are interested in (such as $\dim \ker(\pi)$ and $\dim \operatorname{im}(\pi)$) do not depend on such choice.

Since f only has poles at the points $(x_1, \ldots, x_N)$, and the principal part at each x_i is f_i, the residue of $f \cdot \omega$ at x_i equals the residue of $f_i \cdot \omega$ at the same point. This concludes one implication. The converse is an application of Serre duality and is omitted. □

According to the lemma above, $W = \ker(\rho)$, where ρ is the linear map (the 'residue' map)

$$\rho\colon V \longrightarrow \Omega(X)^*\,, \qquad \rho(f_1, \ldots, f_N)(\omega) = \sum_{i=1}^{N} \operatorname*{Res}_{x_i}\ f_i \cdot \omega\,. \tag{4.40}$$

Again by the rank-nullity theorem, $\dim V = \dim\ker(\rho) + \dim\operatorname{im}(\rho)$. We already know that $\dim V = \deg(D)$, so we only have to prove that $\dim\operatorname{im}(\rho) = g - \ell(K-D)$. Since the dimension of $\Omega(X)$ is g, this is equivalent to proving that $\dim\operatorname{coker}(\rho) = \ell(K-D)$. In order to compute the cokernel of ρ, we can use one more linear algebra trick: $\operatorname{coker}(\rho) \cong \ker(\rho^*)^*$, where ρ^* is the adjoint (or transpose) map:

$$\rho^*\colon \Omega(X) \longrightarrow V^*\,, \qquad \rho^*(\omega)(f_1, \ldots, f_N) = \sum_{i=1}^{N} \operatorname*{Res}_{x_i}\ f_i \cdot \omega\,. \tag{4.41}$$

Since $f_i = O(z^{-n_i})$, we deduce that $\rho^*(\omega) = 0$ if and only if ω has a zero of order $\geq n_i$ at x_i for all $i = 1, \ldots, N$, which is the same as asking that $\operatorname{div}(\omega) \geq D$. This proves that $\ker(\rho^*) = \mathcal{I}(-D)$. By Serre duality, the dimension of such space is precisely $\ell(K-D)$.

To summarise, we have constructed the exact sequence of complex vector spaces:

$$\begin{tikzcd} 0 \arrow[r] & \ker(\pi) \arrow[r] \arrow[d, equal] & \mathcal{L}(D) \arrow[r, "\pi"] & V \arrow[r, "\rho"] & \Omega(X)^* \arrow[r] & \operatorname{coker}(\rho) \arrow[r] \arrow[d, equal] & 0 \\ & \mathcal{O}(X) & & & & \mathcal{I}(-D)^* & \end{tikzcd} \tag{4.42}$$

where π is the 'principal part' map and ρ is the 'residue' map. In turn, the exact sequence gives the following dimension formula:

$$\dim\mathcal{O}(X) - \dim\mathcal{L}(D) + \dim V - \dim\Omega(X) + \dim\mathcal{I}(-D) = 0\,. \tag{4.43}$$

The values $\dim\mathcal{O}(X) = 1$, $\dim\mathcal{L}(D) = \ell(D)$, $\dim V = \deg(D)$, $\dim\Omega(X) = g$, $\dim\mathcal{I}(-D) = \ell(K-D)$ conclude the proof.

Part II. Consider now an arbitrary divisor D. We shall prove the Riemann–Roch inequality by induction on $\deg(D_-)$. If $\deg(D_-) = 0$, then $D \geq 0$ and this is simply Part I. Suppose now that $\deg(D_-) > 0$ and proceed by induction: take a point x in the support of D_-. The induction hypothesis implies that the Riemann–Roch inequality holds for $D + [x]$:

$$\ell(D+[x]) - \ell(K-D-[x]) \geq \underbrace{\deg(D+[x])}_{=\deg(D)+1} + 1 - g = \deg(D) + 2 - g\,. \tag{4.44}$$

By lemma 3.16, the following inequalities hold:

$$\ell(D) \geq \ell(D+[x]) - 1 \qquad \text{and} \qquad \ell(K-D) \leq \ell(K-D-[x]) + 1\,. \tag{4.45}$$

So the only scenario in which the inequality would not hold for D is when $\ell(D) \geq \ell(D+[x]) - 1$ and $\ell(K-D) \leq \ell(K-D-[x]) + 1$. Let us proof that that this is not possible.

By contradiction, suppose that this is the case. The two equalities imply that there exist a non-trivial meromorphic function f and a non-trivial meromorphic form ω such that

$$f \in \mathcal{L}(D+[x]) \setminus \mathcal{L}(D) \qquad \text{and} \qquad \omega \in \mathcal{I}(-D) \setminus \mathcal{I}(-D-[x])\,. \tag{4.46}$$

Denoting by $-n$ the coefficient of $[x]$ in D (with $n > 0$), this is equivalent to say that f has a zero of order exactly $n-1$ at x, and ω has a pole of order exactly n at x. Hence, the meromorphic form $\eta := f \cdot \omega$ has a pole of order exactly 1 at x. In particular, $\mathrm{Res}_x\, \eta \neq 0$. On the the other hand, η does not have other poles: for any point $y \neq x$, denoting by n_y the coefficient of $[y]$ in D,

$$\mathrm{ord}_y(\eta) \geq \underbrace{\mathrm{ord}_y(f)}_{\geq -n_y} + \underbrace{\mathrm{ord}_y(\omega)}_{\geq n_y} \geq 0\,. \tag{4.47}$$

We deduce that the sum of all residues of η equals $\mathrm{Res}_x\;\eta \neq 0$, in contradiction with the residue theorem.

Part iii. Finally, for the last part of the proof we use a clever trick to obtain the desired equality for an arbitrary divisor D: We substitute $K-D$ for D in the inequality proved in Part ii. Since $\ell(K-D) = \ell(-D)$ by Serre duality, we deduce that

$$\ell(K-D) - \ell(D) \geq \underbrace{\deg(K-D)}_{=2g-2-\deg(D)} + 1 - g = g - 1 - \deg(D)\,. \tag{4.48}$$

Reversing the inequality, we find $\ell(D) - \ell(K-D) \leq \deg(D) + 1 - g$. Together with the inequality from Part ii, we find the Riemann–Roch equality.

5. The Abel–Jacobi theory

Lecture 11
May 16th, 2024

In the previous chapters, we introduced the concept of the divisor group of a compact Riemann surface X, which provides a natural framework to filter meromorphic functions into finite-dimensional vector spaces. The dimension of these spaces is computed by the Riemann–Roch.

A related 'smaller' group is that of divisors of degree zero up to linear equivalence, the degree zero Picard group $\mathrm{Pic}_0(X)$. Geometrically, it measure 'how many' degree zero divisors are not principal. A natural question arises:

How can we compute $\mathrm{Pic}_0(X)$*?*

The answer is provided by the Abel–Jacobi theorem. A central role in the Abel–Jacobi theory is played by the integration of differential forms over cycles, which also played a crucial role in the proof of the Riemann–Roch theorem.

5.1. **Periods and Riemann bilinear relations.** In the previous section we have sketched the idea of integrations of differential forms along paths on Riemann surfaces:

$$\int_\gamma \omega \tag{5.1}$$

for $\gamma\colon [0,1] \to X$ a path and ω a differential form that is smooth along γ. This makes sense locally thanks to the chain rule for integration and thanks to the additivity of the integral. As in the local case, integration of differential forms is invariant under deformation.

We can also talk about integrals of wedge products of differential forms: if locally $\omega \overset{\text{loc}}{=} p(x,y)dx + q(x,y)dy$ and $\eta \overset{\text{loc}}{=} u(x,y)dx + v(x,y)dy$, then $\omega \wedge \eta \overset{\text{loc}}{=} (p \cdot v - q \cdot u)dx\,dy$, which can be integrated over the Riemann surface:

$$\int_X \omega \wedge \eta\,. \tag{5.2}$$

Notice that, if both ω and η are holomorphic (under the identification $z = x + \mathrm{i}y$), then $\omega \wedge \eta = 0$. However, it we get a non-trivial result by integrating $\omega \wedge \bar{\eta}$. Integration over the whole Riemann surface and integration along contours are intimately linked to each other. This is known are Riemann's bilinear identity. In order to state it, we need to recall some basic facts about contours on compact Riemann surfaces.

Let X be a genus g Riemann surface. Topologically, it is identified with the identification polygon

$$a_1 b_1 \bar{a}_1 \bar{b}_1 \cdots a_g b_g \bar{a}_g \bar{b}_g\,. \tag{5.3}$$

Indeed, each word $a_i b_i \bar{a}_i \bar{b}_i$ is a topological torus, and the concatenation of g such words is precisely the connected sum of g tori. Throughout this section, we will work on the fixed identification polygon and perform most of the computations on the fundamental domain P associated to it. A classical result from algebraic topology shows that shows that the homology group

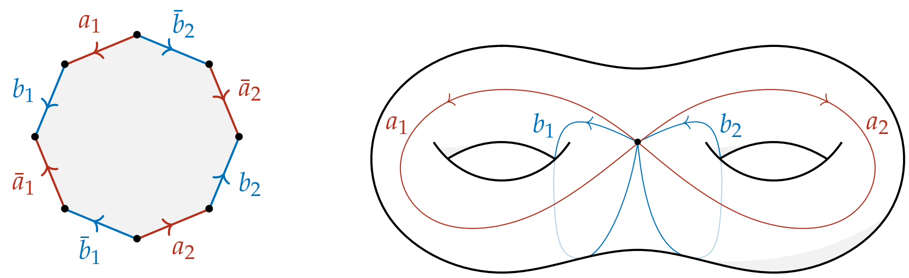


FIGURE 23. A surface of genus 2 with its identification polygon and the induced cycles on the surface.

$H_1(X, \mathbb{Z})$ is generated by the loops $\{ a_i, b_i \}_{i=1}^{g}$. For any differential form η, its integrals

$$A_i(\eta) := \oint_{a_i} \eta\,, \qquad\qquad B_i(\eta) := \oint_{b_i} \eta\,, \tag{5.4}$$

for $i = 1, \ldots, g$ are called the A- and B-periods of η (with respect to the chosen basis $\{ a_i, b_i \}_{i=1}^{g}$).

Theorem 5.1 (Riemann bilinear identity). *Let ω and η be two closed*[11] *differential forms on X. Then*

$$\int_X \omega \wedge \eta = \sum_{i=1}^{g} \big(A_i(\omega) B_i(\eta) - A_i(\eta) B_i(\omega) \big)\,. \tag{5.5}$$

Proof (for the reader familiar with differential calculus). We first observe that the integral over X can be written as an integral over the fundamental domain P of the identification polygon. Moreover, it is easy to check that, sine both ω and η are closed, we can write $\eta \wedge \omega$ inside P as the differential of the 1-form $f_\omega \cdot \eta$, where f_ω is a primitive of ω:

$$f_\omega(x) := \int_{x_0}^{x} \omega \tag{5.6}$$

for a path inside P from an arbitrary base-point x_0 to x. It is important to remark that f_ω is single-valued on P, but not on X (in other words, it depends on the choice of the identification polygon). Since P is simply connected, there is only one homotopy class of such path, hence f_ω is a well–defined smooth function.

If $x_1, x_2 \in \partial P$ are points on the boundary of the domain P that correspond to the same point in X, then there is a unique homotopy class of contour $\gamma \subset P$ joining them; this contour correspond to a homotopy class in X that we denote with the same symbol, and

$$\oint_\gamma \omega = f_\omega(x_2) - f_\omega(x_1)\,. \tag{5.7}$$

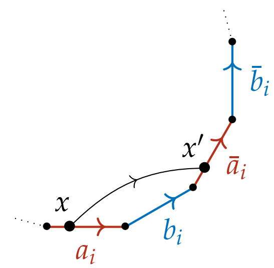
We can now apply Stokes' theorem to compute the integral over P:

$$\int_X \omega \wedge \eta = \oint_{\partial P} f_\omega \cdot \eta = \sum_{i=1}^{g} \left(\int_{a_i} + \int_{\bar{a}_i} \right) f_\omega \cdot \eta + \sum_{i=1}^{g} \left(\int_{b_i} + \int_{\bar{b}_i} \right) f_\omega \cdot \eta\,. \tag{5.8}$$

---

[11]Recall that a differential form is closed if its exterior derivative is zero. The differential of a 1-form locally defined as $v \overset{\text{loc}}{=} u\,dx + v\,dy$ is simply defined as $dv \overset{\text{loc}}{=} (-u_y + v_x) dx\,dy$.

Consider the i-th summand in the first sum. For any pair of points $x \in a_i$ and $x' \in \bar{a}_i$, we have $f_\omega(x') - f_\omega(x) = \oint_{b_i} \omega = B_i(\omega)$:



(5.9)

Moreover, the differential form η is single valued on X. Hence, the i-th summand in the first sum can be rewritten as

$$\begin{aligned} \int_{a_i} f_\omega(x) \cdot \eta(x) + \int_{\bar{a}_i} f_\omega(x') \cdot \eta(x') &= \int_{a_i} f_\omega(x) \cdot \eta(x) - \int_{a_i} \big(f_\omega(x) + B_i(\omega)\big) \cdot \eta(x) \\ &= -B_i(\omega) \int_{a_i} \eta = -A_i(\eta) B_i(\omega) \,. \end{aligned} \tag{5.10}$$

The minus sign in the second equality is due to the fact that $\bar{a}_i$ is the cycle with opposite direction to a_i in X. A similar computation holds for the second sum. All together, we get the thesis. $\square$

The Riemann bilinear identity gives some nice consequence when applied to holomorphic forms.

Proposition 5.2 (Riemann bilinear inequality)**.** *For any holomorphic form $\omega \in \Omega(X)$,*

$$\Im\left(\sum_{i=1}^{g} A_i(\omega)\overline{B_i(\omega)}\right) \le 0 \,, \tag{5.11}$$

with the equality being valid if and only if ω is identically zero.

Proof. Consider the Riemann bilinear identity for with $\eta = \bar{\omega}$. If $\omega \overset{\text{loc}}{=} f(z)dz$, it is easy to check that $\omega \wedge \bar{\omega} \overset{\text{loc}}{=} -2\mathrm{i}|f(z)|^2\, dx\, dy$. Thus, $\frac{1}{2\mathrm{i}} \int_X \omega \wedge \bar{\omega} \le 0$, with the equality being true if and only if ω is identically zero. On the other hand, the Riemann bilinear identity implies that

$$\frac{1}{2\mathrm{i}} \int_X \omega \wedge \bar{\omega} = \frac{1}{2\mathrm{i}} \sum_{i=1}^{g} \big(A_i(\omega)B_i(\bar{\omega}) - A_i(\bar{\omega})B_i(\omega)\big) \,. \tag{5.12}$$

Since $A_i(\bar{\omega}) = \overline{A_i(\omega)}$ (ans similarly for B-periods), we find the thesis. $\square$

As simple consequence of the above result is a criterion for establishing the triviality of holomorphic forms.

Corollary 5.3. *Let $\omega \in \Omega(X)$ be a holomorphic form on X. Then ω is identically zero if an only if all its A-periods (or all its B–periods) vanish.*

Since the space of holomorphic forms is g-dimensional, we can collect the values of $A_i(\omega)$ and $B_i(\omega)$ into two $g \times g$ matrices. That is, fix a basis $\{\, \omega_1, \dots, \omega_g \,\}$ of $\Omega(X)$ and set

$$A := \Big(A_i(\omega_j)\Big)_{1 \le i,j \le g}, \qquad B := \Big(B_i(\omega_j)\Big)_{1 \le i,j \le g}. \tag{5.13}$$

These matrices are called the period matrices with respect the given basis of $\Omega(X)$ and $H_1(X, Z)$. The results above can be re-stated purely in terms of period matrices, and are collectively known are Riemann bilinear relations.

Corollary 5.4 (Riemann bilinear relations)**.** *The matrices A and B are invertible and are such that:*

(1) $A^{\mathsf{T}} B = B^{\mathsf{T}} A$;
(2) the matrix $\tau := A^{-1} B$ is symmetric and its imaginary part is positive definite.

Here M^{T} denotes the matrix transpose of M.

Proof. The invertibility is just a restatement of corollary 5.3. The identity $A^{\mathsf{T}} B = B^{\mathsf{T}} A$ is the Riemann bilinear identity applied to $\omega_i \wedge \omega_j = 0$. As for τ, the symmetry is again a consequence of the Riemann bilinear identity, while the imaginary part being positive definite is a restatement of the Riemann bilinear inequality. ☐

Exercise 5.1. *For a fixed identification polygon decomposition, prove that there exists a choice of basis of holomorphic differentials $(\omega_1, \dots, \omega_g)$ such that*

$$\int_{a_i} \omega_j = \delta_{i,j}\,. \tag{5.14}$$

In this case, it is customary to denote the matrix of B-periods as $\tau = (\tau_{i,j})_{1 \le i,j \le g}$, where $\tau_{i,j} = \oint_{b_i} \omega_j$. Prove that the matrix τ is symmetric and its imaginary part is positive-definite.

The space of $g \times g$ symmetric matrices over the complex numbers whose imaginary part is positive definite is called the Siegel upper-half space of degree g, and it generalises the upper-half space $\mathbb{H}$ (corresponding to $g = 1$) to higher dimensions. Lie-theoretically, it is the symmetric space associated to the symplectic group $\mathrm{Sp}(2g, \mathbb{R})$.

We have just seen how A- and B-periods play a special role in the theory of integration on Riemann surfaces. Since the cycles $\{\, a_i, b_i \,\}_{i=1}^{g}$ for a basis of the first homology group of X, it is natural to introduce the group homomorphism

$$H_1(\mathrm{X}, \mathbb{Z}) \longrightarrow \Omega(\mathrm{X})^*, \qquad \gamma \longmapsto \oint_\gamma\,. \tag{5.15}$$

Recall that the first homology group is the group generated by piecewise smooth oriented contours (i.e. closed paths), up to equivalence defined as $\gamma_1 \sim \gamma_2$ if and only if there exists a 2-cell c in X with $\partial c = \gamma_1 - \gamma_2$. The opposite of a contour is the contour oriented in the opposite direction.

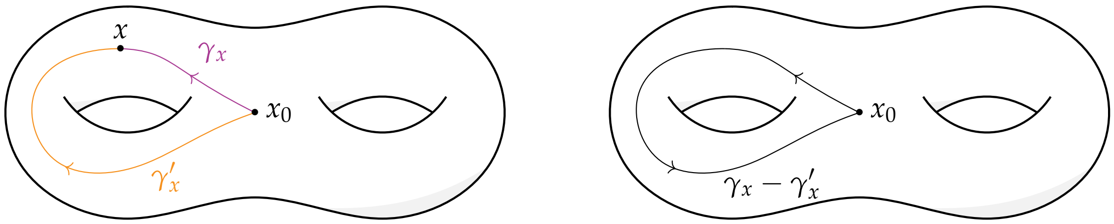


FIGURE 24. Two choices of paths γ_x and γ'_x from x to x_0 which differ by a closed path.

Definition 5.5. A period on a compact Riemann surface X is a linear map $\lambda \in \Omega(X)^*$ of the form $\lambda = \oint_\gamma$ for some $\gamma \in H_1(X, \mathbb{Z})$.

It is not hard to show that the space of periods on X, denoted $\Lambda(X)$ is a lattice in $\Omega(X)^*$. In other words, there exists $\lambda_1, \ldots, \lambda_{2g} \in \Lambda(X)$ that are linearly independent over $\mathbb{R}$ such that

$$\Lambda(X) = \{ \mu_1\lambda_1 + \cdots \mu_{2g}\lambda_{2g} \mid \mu_i \in \mathbb{Z} \} . \tag{5.16}$$

For a fixed basis $\{ a_i, b_i \}_{i=1}^g$ as above, we can choose $\lambda_i = \oint_{a_i}$ and $\lambda_{g+i} = \oint_{B_i}$ for all $i = 1, \ldots, g$.

Definition 5.6. Define the Jacobian of X as the group $\Omega(X)^*$ modulo its subgroup $\Lambda(X)$ of periods:

$$J(X) \coloneqq \Omega(X)^* / \Lambda(X) . \tag{5.17}$$

As a group, it is isomorphic to the complex torus $\mathbb{C}^g/\mathbb{Z}^{2g}$, which is isomorphic to $(S^1)^{2g}$.

Closed paths on a compact Riemann surface naturally leads to the definition of the Jacobian. We can then ask ourselves: what happens when we consider non-closed paths? To this end, fix a base-point $x_0 \in X$, and for any point $x \in X$ choose a path γ_x from x_0 to x. Thus, we get a map

$$\mathrm{AJ}\colon X \longrightarrow \Omega(X)^* , \qquad \mathrm{AJ}(x)(\omega) = \int_{\gamma_x} \omega , \tag{5.18}$$

which is a group homomorphism. The map depends on the choice of base-point x_0 and on the choice of path γ_x from x to x_0. However, for any two choices γ_x and γ'_x of paths from x_0 to x, the difference $\gamma_x - \gamma'_x$ is a closed path (cf. figure 24). Hence, $\oint_{\gamma_x - \gamma'_x} \omega$ is period, and the map AJ descends to the Jacobian:

$$\mathrm{AJ}\colon X \longrightarrow J(X) , \qquad \mathrm{AJ}(x)(\omega) = \left[\int_{\gamma_x} \omega \right] . \tag{5.19}$$

Abusing notation, we denote it with the same symbol. It is called the Abel–Jacobi map, and it does not depend on the choice of path from x_0 to x. However, it still depends on the choice of x_0.

We can now extend the map by linearity to the divisor group, and denote by AJ_0 its restriction to the subgroup $\mathrm{Div}_0(X)$ of degree-zero divisors.

Lemma 5.7. *The Abel–Jacobi map*

$$\mathrm{AJ}_0 \colon \mathrm{Div}_0(X) \longrightarrow J(X) \tag{5.20}$$

does not depend on the choice of base-point x_0.

Proof. Let x_0 and x_0' be two base-points and denote the associated Abel–Jacobi maps as AJ_0 and AJ_0' respectively. Choose a path γ from x_0 to x_0'. Then

$$\mathrm{AJ}_0(x)(\omega) - \mathrm{AJ}_0'(x)(\omega) = \left[\int_\gamma \omega\right] \tag{5.21}$$

and denote the right-hand side as $j \in \mathrm{Jac}(X)$. As a consequence, we find that for all $D = \sum_x n_x\,[x] \in \mathrm{Div}(X)$:

$$\mathrm{AJ}_0(D)(\omega) - \mathrm{AJ}_0'(D)(\omega) = \sum_x n_x\, j = \deg(D)\, j\,. \tag{5.22}$$

We conclude that $\mathrm{AJ}_0 \equiv \mathrm{AJ}_0'$ on $\mathrm{Div}_0(X)$. □

5.2. **Abel's and Jacobi's theorems.** The Abel–Jacobi map $\mathrm{AJ}_0 \colon \mathrm{Div}_0(X) \to J(X)$ is a well-defined group homomorphism that does not depend on any choice. Abel's and Jacobi's theorem describe its kernel and cokernel respectively.

The proofs of Abel's and Jacobi's theorems are not required for the exam (but the content of their theorems is).

Theorem 5.8 (Abel's theorem)**.** *Let $D \in \mathrm{Div}_0(X)$ be a degree-zero divisor. Then D is principal if and only of $\mathrm{AJ}_0(D) = 0$. In other words,* $\ker(\mathrm{AJ}_0) = \mathrm{PDiv}(X)$.

Niels Henrik Abel
(1802–1829)

We are going to prove the two inclusions separately, starting with $\mathrm{PDiv}(X) \subseteq \ker(\mathrm{AJ}_0)$. To this end, we introduce the concept of pushforward of differential forms and pullback of paths.

Fix a holomorphic map $f \colon X \to Y$, and let ω be a meromorphic form on X. We define a meromorphic form $f_*\omega$ on Y (called the pushforward of ω along f) as follows. Fix a point $y \in Y$ with local coordinates around w centred at y and z_i centred at $x_1, \dots, x_m$ for all x_i in the fibre $f^{-1}(y)$. We can assume that the function f locally looks like $z_i \mapsto w = z_i^{k_i}$, and that $\omega \overset{\mathrm{loc}}{=} g_i(z_i)\, dz_i$. We then set locally around y

$$f_*\omega \overset{\mathrm{loc}}{:=} \sum_{i=1}^m \sum_{j=0}^{k_i-1} \frac{g_i(\zeta_i^j z_i)}{k_i(\zeta_i^j z_i)^{k_i-1}} dw\,. \tag{5.23}$$

Here $\zeta_i = e^{\frac{2\pi \mathrm{i}}{k_i}}$ is a k_i-th root of unity. Intuitively, $f_*\omega$ is summing the differential forms ω over at the preimages, according to the local form of f. In terms of Laurent series, a simple computations shows that for $\omega = \sum_n c_n z^n\, dz$ we find

$$f_*\omega = \sum_{i=1}^m \sum_{n_i} c_{n_i k_i - 1} w^{n_i}\, dw\,. \tag{5.24}$$

Notice that, if f at the preimages of y (i.e. $k_i = 1$ for all $i = 1, \dots, m$), then $f_*\omega$ is simply the sum of the differential forms at the preimages.

Another important tool is the pullback of paths. Given f as before and γ a path in Y, we can pull it back to d paths in X by using the path-lifting properties of covers, as long as the path does not go through any of the branch points. Even if it does, we can delete those points, lift the path in X, and take the closure. We denote the pullback of a path γ defined in this way by $f^*\gamma$.

The relation between pushforward of holomorphic forms and pullback of paths is given by the following result.

Lemma 5.9. *Let $f\colon X \to Y$ be a non-constant holomorphic map, ω a holomorphic form on X, and γ a path on Y. Then*

$$\int_{f^*\gamma} \omega = \int_\gamma f_*\omega\,. \tag{5.25}$$

Sketch of the proof. Since the ramification points are finite, they have measure zero and thus do not affect the integral. We may assume that the path does not pass through the branch points. In this case, the left-hand side is just an integral along the d lifts of γ. Recall that the pushforward of a differential form around an unramified point is simply the sum of the differential forms at the preimages. This means the right-hand side is also the sum of the integrals of ω along the lifted paths of γ. □

We can now prove one of the directions of Abel's theorem.

Proof of Abel's theorem: $\mathrm{PDiv}(X) \subseteq \ker(\mathrm{AJ}_0)$. Let $D = \mathrm{div}(f)$ be a principal divisor associated to a holomorphic function $f\colon X \to \mathbb{P}^1$ of, say, degree d. Denote the set of zeros and poles of f as $\{ z_1, \dots, z_d \}$ and $\{ p_1, \dots, p_d \}$, where points are counted with multiplicity (i.e. repetitions are allowed). Thus $\mathrm{div}(f) = \sum_{k=1}^d (z_k - p_k)$. Given a base-point $x_0 \in X$, choose a path α_k from x_0 to z_k and a path β_k from x_0 to p_k, so that the Abel–Jacobi map reads

$$\mathrm{AJ}_0(D)(\omega) = \left[\sum_{k=1}^d \left(\int_{\alpha_k} - \int_{\beta_k} \right) \omega \right]. \tag{5.26}$$

Consider now a path γ in $\mathbb{P}^1$ from $\mathbf{0}$ to $\boldsymbol{\infty}$. Without loss of generality, we can suppose that γ does not pass through any of the branch points of f. Then we can write $f^*\gamma = \gamma_1 + \cdots + \gamma_d$, where γ_k is a path on X from z_k to p_k. Since $\alpha_k - \beta_k + \gamma_k$ is closed (see figure 25), we find that

$$\mathrm{AJ}_0(D)(\omega) = -\left[\sum_{k=1}^d \int_{\gamma_k} \omega \right] = -\left[\int_{f^*\gamma} \omega \right]. \tag{5.27}$$

Thanks to lemma 5.9, we find $\mathrm{AJ}_0(D)(\omega) = -[\int_\gamma f_*\omega]$. As ω is holomorphic, so is $f_*\omega$. But the only holomorphic form on $\mathbb{P}^1$ is the identically-zero form, hence $\mathrm{AJ}_0(D)(\omega) = 0$. □

The converse statement is more involved.

Proof of Abel's theorem: $\ker(\mathrm{AJ}_0) \subseteq \mathrm{PDiv}(X)$. Let $D = \sum_{k=1}^d n_k[x_k]$ be a degree 0 divisor such that $\mathrm{AJ}_0(D) = 0$. We want to show that there exists a meromorphic function f such that $D = \mathrm{div}(f)$.

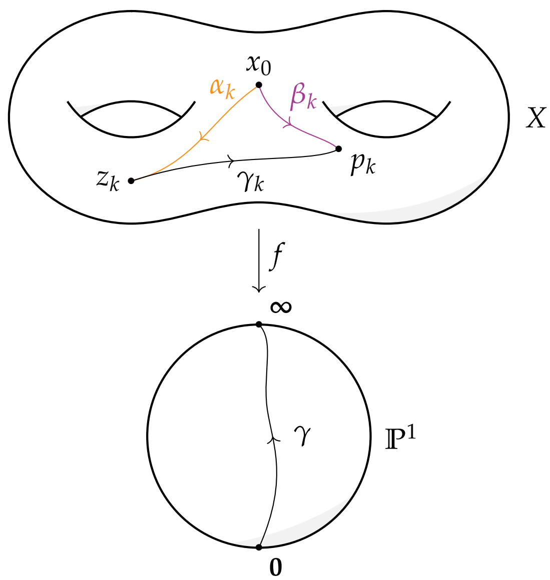


FIGURE 25. The closed path $\alpha_k - \beta_k + \gamma_k$ in the 'ker $\subseteq$ PDiv' part of Abel's theorem.

We proceed as follows. We first construct a meromorphic form ω such that:

(1) ω has simple poles at x_k for all $k = 1, \dots, d$, and is holomorphic at any other point;
(2) $\operatorname{Res}_{x_k} \omega = n_k$ for all $k = 1, \dots, d$;
(3) the periods $A_i(\omega)$ and $B_i(\omega)$ are integral multiples of $2\pi\mathrm{i}$ for all $i = 1, \dots, g$.

Before proving the existence of such a form, let us show how it implies the thesis. We define the meromorphic function

$$f(x) = \exp\left(\int_{\gamma_x} \omega\right), \tag{5.28}$$

where γ_x is any path from the base-point x_0 to x. Since the periods are integral multiples of $2\pi\mathrm{i}$, the function does not depend on the choice of path. Hence, we only need to show that f is meromorphic and $\operatorname{div}(f) = D$. Clearly, f is holomorphic and non-vanishing where ω is holomorphic. Consider then a pole x_k of ω. Since ω has a simple pole with residue n_k, we deduce that locally around x_k we have $\omega = (\frac{n_k}{z} + g(z))dz$ for some holomorphic function $g(z)$. Thus, locally around x_k the function f will be of the form $f = z^{n_k} e^{h(z)}$ for some holomorphic function $h(z)$. The computation shows precisely that $\operatorname{div}(f) = D$.

We are then left with proving the existence of a meromorphic form ω satisfying (1-2-3). The existence of a meromorphic form η satisfying (1-2) is a consequence of Riemann–Roch (and simply uses the fact that D has degree zero). Notice that for any such η and any holomorphic η', then $\eta - \eta'$ still satisfies (1-2). In other words, writing $\eta' = \sum_{i=1}^{g} c_i \omega_i$, we are reduced to show that we can properly choose $c = (c_1, \dots, c_g)$ such that $\omega := \eta - \sum_{i=1}^{g} c_i \omega_i$ satisfies (3) as well.

First, we can assume without loss of generality that the curves a_i and b_i do not pass through the poles of η. Define the holomorphic functions

$$f_{\omega_j}(x) := \int_{x_0}^{x} \omega_j\,, \qquad j = 1, \dots, g \tag{5.29}$$

where the integral is over a path from x_0 to x inside the fundamental domain P. Moreover, define the scalars

$$\rho_j := \sum_{x \in X} \operatorname*{Res}_{x}\, f_{\omega_j}\, \eta\,, \qquad j = 1, \dots, g\,. \tag{5.30}$$

Thanks to the residue theorem, Stokes' theorem and the Riemann bilinear identity, the scalar ρ_i can be expressed as

$$\rho_j = \frac{1}{2\pi \mathrm{i}} \oint_{\partial P} f_{\omega_j}\, \eta = \frac{1}{2\pi \mathrm{i}} \int_X \omega_j \wedge \eta = \frac{1}{2\pi \mathrm{i}} \sum_{i=1}^{g} \big(A_i(\omega_j) B_i(\eta) - A_i(\eta) B_i(\omega_j) \big)\,. \tag{5.31}$$

On the other hand, we can directly compute the residue thanks to the known properties of f_{ω_j} and η. Indeed, f_{ω_j} is holomorphic and η has only simple poles at x_k with residue n_k. Thus:

$$\rho_j = \sum_{k=1}^{d} n_k\, f_{\omega_j}(x_k) = \sum_{k=1}^{d} n_k \int_{x_0}^{x} \omega_j\,. \tag{5.32}$$

Modulo the period lattice, this is precisely the $\mathrm{AJ}_0(D)(\omega_i)$. Since by hypothesis $\mathrm{AJ}_0(D) = 0$, we deduce that there exists γ such that the functional $\sum_{k=1}^{d} n_k \int_{x_0}^{x}$ is precisely $\oint_\gamma$. We can express γ in the basis $(a_i, b_i)_{i=1}^{g}$ as $\gamma = \sum_{i=1}^{g} (\mu_i a_i - \nu_i b_i)$ for some $\mu_i, \nu_i \in \mathbb{Z}$, so that

$$\rho_j = \sum_{i=1}^{g} \big(\mu_i A_i(\omega_j) - \nu_i B_i(\omega_j) \big)\,. \tag{5.33}$$

All together, we deduce that

$$\sum_{i=1}^{g} A_i(\omega_j)(B_i(\eta) - 2\pi\mathrm{i}\,\mu_i) - \sum_{i=1}^{g} B_i(\omega_j)(A_i(\eta) - 2\pi\mathrm{i}\,\nu_i) = 0\,. \tag{5.34}$$

In matrix form, the above collection of equation reads $A^{\mathsf{T}}\beta - B^{\mathsf{T}}\alpha = 0$, where β is the column vector whose i-th entry is $B_i(\eta) - 2\pi\mathrm{i}\mu_i$ and similarly for α.

Consider now the linear maps

$$\begin{array}{ccc} \mathbb{C}^g & \xrightarrow{\ \varphi\ } & \mathbb{C}^{2g} \\ v & \longmapsto & \begin{pmatrix} Av \\ Bv \end{pmatrix} \end{array} \qquad\qquad \begin{array}{ccc} \mathbb{C}^{2g} & \xrightarrow{\ \psi\ } & \mathbb{C}^g \\ \begin{pmatrix} v \\ w \end{pmatrix} & \longmapsto & A^{\mathsf{T}}v - B^{\mathsf{T}}w \end{array} \tag{5.35}$$

which by corollary 5.4 satisfy $\psi \circ \varphi = 0$. Moreover, since both A and B are full rank, we deduce that $\ker(\psi) = \operatorname{im}(\varphi)$. As the vector composed by β and α is in $\ker(\psi)$, we find that it is realised as the image via ψ of some $c \in \mathbb{C}^g$. It is then easy to show that c is the desired vector of coefficients: setting $\omega := \eta - \sum_{i=1}^{g} c_i\, \omega_i$, then

$$A_i(\omega) = 2\pi\mathrm{i}\,\nu_i \qquad \text{and} \qquad B_i(\omega) = 2\pi\mathrm{i}\,\mu_i \tag{5.36}$$

are integral multiples of 2πi. □

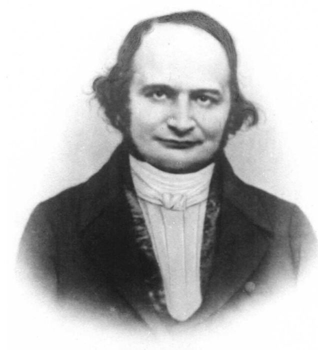
Geometrically, Abel's theorem answers a natural question regarding meromorphic functions on compact Riemann surface. That is, given a degree zero divisor D on X, is there a meromorphic function f satisfying $D = \operatorname{div}(f)$? The answer is affirmative if and only if $\mathrm{AJ}_0(D)(\omega) = 0$ for all holomorphic forms ω.

Exercise 5.2. *Prove the Weierstraß gap theorem on the torus $\mathbb{C}/(\mathbb{Z} + \tau\mathbb{Z})$ using Abel's theorem.*

An immediate consequence of Abel's theorem is that the Abel–Jacobi map descends to an injective group homomorphism from the Picard group of degree-zero divisors up to principal divisors to the Jacobian:

$$\mathrm{AJ}_0 \colon \mathrm{Pic}_0(X) \longrightarrow J(X)\,. \tag{5.37}$$

Again, abusing notation, we denote it with the same symbol. Jacobi's theorem asserts that the map is also surjective.



Carl Gustav Jacob Jacobi (1804–1851)

Theorem 5.10 (Jacobi's inversion theorem)**.** *The Abel–Jacobi map*

$$\mathrm{AJ}_0 \colon \mathrm{Pic}_0(X) \longrightarrow J(X) \tag{5.38}$$

is surjective.

Proof. As the map AJ_0 is a group morphism, it suffices to show that the image of AJ_0 contains an open set. This is because on a connected topological group G (such as $J(X)$), every non-trivial open subgroup H is automatically the whole group: $H = G$.

We proceed as follows. Pick a base point $x_0 \in X$ and consider the map $I \colon X^g \to J(X)$ defined as

$$I(x_1, \ldots, x_g) = \mathrm{AJ}_0\left(\sum_{i=1}^{g}(x_i - x_0)\right). \tag{5.39}$$

In order to prove that the image of AJ_0 contains an open set, it suffices to show that the image of I contains an open set. Because both X^g and $J(X)$ are g-dimensional complex manifolds, one approach would be to show that there exists a point $(x_1, \ldots, x_g) \in X^g$ where the Jacobian matrix of I is invertible. The inverse function theorem will then tell us that the image contains an open set. To this end, fix a basis $(\omega_1, \ldots, \omega_g)$ of $\Omega(X)$, which gives an identification of $\Omega(X)^*$ with $\mathbb{C}^g$. The map I then reads

$$I(x_1, \ldots, x_g) = \begin{pmatrix} \sum_{i=1}^{g} \int_{x_0}^{x_i} \omega_1 \\ \vdots \\ \sum_{i=1}^{g} \int_{x_0}^{x_i} \omega_g \end{pmatrix}, \tag{5.40}$$

where the path is an arbitrary path from x_0 to x_i. For ease of notation, denote the j-th row of I as I_j. Let us choose local coordinates z_i centred at x_i; then the Jacobian matrix of I in local coordinates reads

$$\mathrm{Jac}_I(x_1,\dots,x_g) = \begin{pmatrix} \frac{\omega_1(x_1)}{dz_1} & \dots & \frac{\omega_1(x_g)}{dz_g} \\ \vdots & \ddots & \vdots \\ \frac{\omega_g(x_1)}{dz_1} & \dots & \frac{\omega_g(x_g)}{dz_g} \end{pmatrix}. \tag{5.41}$$

The above expression needs an explanation. In coordinates, I_j is given by $\sum_{i=1}^{g} \int_{x_0}^{x_i} \omega_j$, and if we differentiate it with respect to the x_k coordinate, all terms except for the k-th summand vanish since they do not depend on x_k. The derivative of the k-th summand with respect to the coordinate x_k can be computed by writing ω_j locally in the coordinate z_k as $\omega_j = h_{j,k}(z_k)dz_k$. In this case, the expression for the derivative is computed as

$$\frac{\partial I_j}{\partial x_k} = \frac{\partial}{\partial x_k}\int_{x_0}^{x_k} h_{j,k}(z_k)dz_k\,, \tag{5.42}$$

which from the fundamental theorem of calculus equals $h_{j,k}(z_k)$. This motivates the notation $\frac{\omega_j(x_k)}{dz_k}$ in place of $h_{j,k}(z_k)$.

We now have to show that there exists a point $(x_1,\dots,x_g) \in X^g$ where the above matrix in invertible. We start by choosing $x_1 \in X$ as a point where the form ω_1 does not vanish. We then modify the basis of $\Omega(X)$ as follows: we replace ω_j for $j \geq 2$ with

$$\omega_j' := \omega_j - \frac{\omega_j(x_1)}{\omega_1(x_1)\omega_1}\,. \tag{5.43}$$

The definition is well-posed, since $\omega_1(x_1) \neq 0$. With this choice, we deduce that $\frac{\omega_j(x_1)}{dz_1} = 0$ for all $j \geq 2$, and is different from zero for $j = 1$. In other words, the Jacobian is of the form

$$\mathrm{Jac}_I(x_1,\dots,x_g) = \begin{pmatrix} \frac{\omega_1(x_1)}{dz_1} & \frac{\omega_1(x_2)}{dz_2} & \dots & \frac{\omega_1(x_g)}{dz_g} \\ 0 & \frac{\omega_2(x_2)}{dz_2} & \dots & \frac{\omega_2(x_g)}{dz_g} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \frac{\omega_g(x_2)}{dz_2} & \dots & \frac{\omega_g(x_g)}{dz_g} \end{pmatrix} \qquad \text{with} \qquad \frac{\omega_1(x_1)}{dz_1} \neq 0\,. \tag{5.44}$$

We then repeat this process again, by choosing $x_2 \in X$ as a point where ω_2' does not vanish and modifying the basis elements ω_j' for $j \geq 3$ accordingly. By repeating this possess g times, we find g points $(x_1,\dots,x_g) \in X^g$ and a basis of holomorphic differentials such that the Jacobian matrix of I at $(x_1,\dots,x_g)$ is upper-triangular with non-vanishing elements along the diagonal, showing that it is indeed invertible. □

We can collect Abel's and Jacobi's results in a unique statement:

Theorem 5.11 (Abel–Jacobi theorem)**.** *The Abel–Jacobi map is a group isomorphism:*

$$\mathrm{Pic}_0(X) \cong J(X)\,. \tag{5.45}$$

The Abel–Jacobi theorem completely characterise the Picard group of degree zero divisors modulo principal divisors, i.e. 'how far' degree zero divisors are from being principal: it is a complex

g-dimensional torus. A natural question would be to characterise the whole Picard group. A simple computation shows that, after choosing a base-point $x_0 \in X$, the map

$$\mathrm{Div}(X) \longrightarrow \mathrm{Div}_0(X) \times \mathbb{Z}\,, \qquad D \longmapsto \big(D - \deg(D)[x_0], \deg(D)\big) \tag{5.46}$$

is a group isomorphism that descends to the Picard group: $\mathrm{Pic}(X) \cong \mathrm{Pic}_0(X) \times \mathbb{Z}$. As a consequence, we deduce the following characterisation of the full Picard group:

$$\mathrm{Pic}(X) \cong J(X) \times \mathbb{Z}\,. \tag{5.47}$$

An interesting problem related to the Abel–Jacobi map, known as Schottky's problem, asks for necessary and sufficient conditions for a g-dimensional complex torus to be the Jacobian of a compact genus g Riemann surface. Interestingly, the problem has deep relations with the Kadomtsev–Petviashvili equation, a partial differential equation which describes non-linear wave motion.

A simple example of interaction between wave equations and Riemann surfaces is provided by the following exercise.

Exercise 5.3. *Consider the Korteweg–de Vries (KdV) equation*

$$u_t = 6uu_x - u_{xxx}\,, \tag{5.48}$$

which describes the motion of a wave $u = u(x,t)$ on shallow water surfaces. Prove that any periodic solution of KdV in the form of a travelling wave $u = u(x-ct)$ has the form

$$u = 2\wp(x - ct - x_0; \Lambda) - \frac{c}{6} \tag{5.49}$$

for an arbitrary lattice Λ and $x_0 \in \mathbb{C}$.

6. Hurwitz theory

Lecture 12
May 23rd, 2024

We now introduce the counting problem for maps of Riemann surfaces: fixing a compact Riemann surface Y and a finite number of points $y_1, \ldots, y_m \in Y$, how many maps to Y have a specified ramification behaviour over the chosen branch points and are unramified elsewhere? Natural questions that arise are:

(1) Is the number of such maps finite?
(2) Does it depend on the Riemann surface Y?
(3) Does it depend on the configuration of the points y_i?

Fortunately, the answers are quite favourable: the number is always finite, it depends only on the genus of Y; and it also depends on the choice of ramification over the points y_i, but not on the positions of the points. A key reason for these favourable answers is that maps of Riemann surfaces are essentially "controlled by topology".

6.1. Hurwitz numbers.

To start with, we introduce a combinatorial gadget which will be used to control the ramification behaviour over the chosen branch points.

Definition 6.1. Let $d > 0$ be an integer. A partition of d is an ordered tuple of positive integers $\mu = (\mu_1, \ldots, \mu_\ell)$ such that $\mu_1 \geq \cdots \geq \mu_\ell > 0$ and $\sum_{i=1}^{\ell} \mu_i = d$. The elements μ_i are called the parts of μ. The sum d is called the size and denoted $|\mu|$. The number ℓ of parts in μ is called the length and denoted $\ell(\mu)$. We write $\mu \vdash d$ to denote a partition of d.

Example 6.2. There are five distinct partitions of size 4: (4), $(3,1)$, $(2,2)$, $(2,1,1)$ and $(1,1,1,1)$. Their lengths are respectively one, two, two, three, and four.

Let $f \colon X \to Y$ be a holomorphic map between compact Riemann surfaces of degree d and fix a point $y \in Y$. For all points $x \in f^{-1}(y)$, we have a multiplicity $\mu_x(f)$ and we can arrange all such integers in a partition μ_y. This is called the ramification profile of f over y. Since the sum of the multiplicities is the degree, we have that $\mu_y \vdash d$. Notice that if y is not a branch point, then $\mu_y = (1, \ldots, 1)$, i.e. the partition consisting of 'd ones'.

Example 6.3. Consider the degree d map $[z_0 : z_1] \mapsto [z_0^d : z_1^d]$ from $\mathbb{P}^1$ to itself. The branch points are $\mathbf{0}$ and $\boldsymbol{\infty}$, whose preimage are $\mathbf{0}$ and $\boldsymbol{\infty}$ respectively both with multiplicity d. Thus, in both cases the ramification profile is the partition (d).

Another important feature of Hurwitz numbers is that we are interested in counting maps up to automorphism.

Definition 6.4. Two holomorphic maps of $f \colon X \to Y$ and $f' \colon X' \to Y$ between compact Riemann surfaces are called isomorphic if there is a biholomorphism $\varphi \colon X \to X'$ such that $f = f' \circ \varphi$. In other words, the following diagram commutes.

$$\begin{tikzcd} X \arrow[rr, "\varphi"] \arrow[dr, "f"'] & & X' \arrow[dl, "f'"] \\ & Y & \end{tikzcd} \tag{6.1}$$

An automorphism of $f\colon X \to Y$ is a biholomorphism $\varphi\colon X \to X$ such that $f = f \circ \varphi$. The group of automorphisms of f is denoted $\mathrm{Aut}(f)$.

We are now ready to define Hurwitz numbers.

Definition 6.5 (Hurwitz numbers)**.** Let Y be a compact Riemann surface of genus h. Fix points $y_1, \ldots, y_m \in Y$ and let $\mu_1, \ldots, \mu_m$ be partitions of a positive integer d. We define the Hurwitz number as

$$H_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_m) := \sum_{[f]} \frac{1}{|\mathrm{Aut}(f)|} \tag{6.2}$$

where the sum runs over all isomorphism class of $f\colon X \to Y$ where:

(1) X is a compact Riemann surface of genus g,
(2) $f\colon X \to Y$ is a homomorphic map of degree d,
(3) the points $y_1, \ldots, y_m$ are the branch locus of f,
(4) the ramification profile of f over y_i is μ_i.

A map f satisfying (1–4) is called a Hurwitz cover for the discrete data $(g, h, d, \mu_1, \ldots, \mu_m)$.

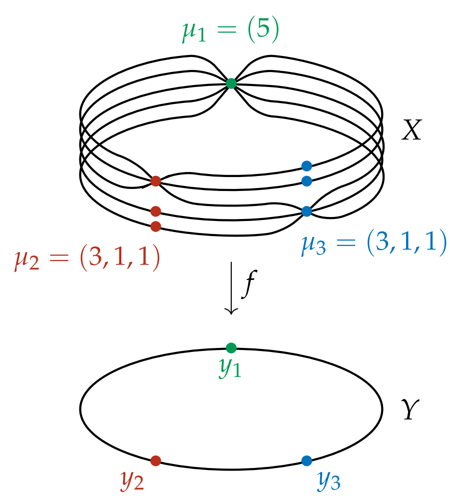


Figure 26. A schematic picture of a Hurwitz cover of degree 5 and three branch points with ramification profiles (5), $(3, 1, 1)$ and $(3, 1, 1)$.

Notice that, for Hurwitz covers to exist, the discrete data must satisfy the Riemann–Hurwitz formula (theorem 2.19):

$$2g - 2 = d(2h-2) + \sum_{i=1}^{m} \sum_{j=1}^{\ell(\mu_i)} (\mu_{i,j} - 1)\,, \tag{6.3}$$

where we set $\mu_i = (\mu_{i,1}, \mu_{i,2}, \dots)$ for the elements of the i-th partition. In particular, the notation is redundant: the degree d can be read out of the sizes of all partitions $\mu_1, \dots, \mu_m$, and the genus g of the cover can be deduced from the Riemann–Hurwitz formula.

We also point out that the above definition assumes various claims:

- the automorphism group $\mathrm{Aut}(f)$ is finite,
- there are finitely many isomorphism classes Hurwitz covers for a given discrete set of data,
- Hurwitz numbers only depend on the genus of the base Y and the ramification profiles, but not on the position of the branch points (hence the notation on the left-hand side).

Finiteness of the automorphism group can be deduced from the Riemann–Hurwitz formula. The other properties will be deduced a posteriori from the monodromy representation.

Exercise 6.1. *Prove that*

$$H_{0 \xrightarrow{d} 0}\big((d),(d)\big) = \frac{1}{d}\,. \tag{6.4}$$

Hint. *Choose $y_1 = \mathbf{0}$ and $y_2 = \boldsymbol{\infty}$ on the target $\mathbb{P}^1$. Write down all possible maps $f\colon \mathbb{P}^1 \to \mathbb{P}^1$ branched over $\mathbf{0}$ and $\boldsymbol{\infty}$, and show that they are all isomorphic. Can you write down the automorphism group of f?*

It will prove useful to allow the source Riemann surface to be disconnected. We call the corresponding count a disconnected Hurwitz number and denote it by $H^{\bullet}_{g \xrightarrow{d} h}(\mu_1, \dots, \mu_m)$.

Before proceeding further, let us pause to consider the fact that we have not yet defined the genus of a disconnected Riemann surface. If you are inclined to assume that the genus should be the sum of the genera of the connected components, recall that there is another fundamental topological invariant for surfaces, the Euler characteristic, which is naturally additive under disjoint unions if defined by the usual formula (1.40). We then define the genus of a disconnected surface as the value g such that the formula $\chi = 2 - 2g$. still holds. It is easy to show that, if the connected components of a disconnected surface X are $X_1, \dots, X_n$, then

$$g = g_1 + \cdots + g_n + 1 - n\,, \tag{6.5}$$

where g_i is the genus of X_i.

6.2. **Monodromy representation: from geometry to group theory.** The natural question we now try to answer is:

How can we compute Hurwitz numbers?

As mentioned earlier, there is a natural topological approach to Hurwitz covers, which (following Cavalieri–Miles) we might call the *IKEA approach*.

Suppose you ordered your preferred Hurwitz cover $f\colon X \to Y$ online, associated with the discrete data $(g, h, d, \mu_1, \ldots, \mu_m)$. To save on shipping costs, the warehouse cuts Y along appropriate segments to obtain a single polygon P_0 (i.e. a topological disc). If the branch locus Brnch_f is contained within the cuts, then cutting X along the inverse image of the cuts produces d disjoint identical copies of P_0, which, with Scandinavian precision, would be labelled $P_1, \ldots, P_d$. What you receive is a box containing these $d+1$ discs and an assembly instruction manual.

One way to provide assembly instructions is to specify, for every loop on Y, its lifts to X. For instance, suppose that a loop γ exits P_0 at a point x and re-enters it at another point x'. Suppose you are also informed that when you lift γ starting from polygon P_1, you end up in polygon P_3. This information tells you that you should glue the points x and x' together, and that you should glue the point corresponding to x in P_1 to the point corresponding to x' in P_3. It is easy to imagine that if you know such information for every possible loop, you could eventually glue back all sides of P_0 to reconstruct Y and all sides of the various P_k's to obtain X. While at first this seems like an overwhelming amount of information to manage, because the endpoints of lifts of loops are invariant under homotopy, all such information is contained in a group homomorphism

$$\rho\colon \pi_1(Y^\times, y_0) \longrightarrow S_d \tag{6.6}$$

called the monodromy representation of the cover f. Here $Y^\times \coloneqq Y \setminus \{ y_1, \ldots, y_m \}$ is the surface Y removed of all the branch points, $y_0 \in Y^\times$ is an arbitrary base-point, and S_d is the symmetric group of d elements.

To properly define the above map (in the case of possibly disconnected domains), proceed as follows. Denote the preimages of y_0 as $f^{-1}(y_0) = \{ x_1, \ldots, x_d \}$. There are exactly d distinct preimages, since f is unramified at every point of $Y^\times$. For any loop $\alpha \in \pi_1(Y^\times, y_0)$, there exists a unique lift[12] $\tilde{\alpha}_k$ in X that starts at x_k. The end-point of $\tilde{\alpha}_k$ must be in the fibre over y_0, so the index k must be permuted to another index in $\{ 1, \ldots, d \}$. In other words, there exists a permutation $\sigma_\alpha \in S_d$ such that $\tilde{\alpha}_k(1) = x_{\sigma(k)}$ (see figure 27 for an example). The monodromy representation is precisely defined as

$$\rho(\alpha) \coloneqq \sigma_\alpha \,. \tag{6.7}$$

One can check that the map so defined is a group homomorphism. Moreover, for every loop γ_i around y_i one can show that the cycle type of the permutation $\rho(\gamma_i)$ is exactly μ_i. That is, $\rho(\gamma_i)$ decomposes into disjoint cycles of length $(\mu_{i,1}, \mu_{i,2}, \ldots)$.

It should be noted that, in order to define ρ, we fixed a base-point y_0 and a labelling $\{ x_1, \ldots, x_d \}$ of the preimage of y_0. A Hurwitz cover with such a choice is called a a y_0-labelled Hurwitz cover. Two y_0-labelled Hurwitz cover $f\colon X \to Y$ and $f'\colon X' \to Y$ with $f^{-1}(y_0) = \{ x_1, \ldots, x_d \}$

[12]Recall the *path lifting lemma*. If $p\colon E \to B$ is a covering space, $\alpha\colon [0,1] \to B$ is any path, and $e_0 \in p^{-1}(\alpha(0))$, then there exists a unique path (called the lift) $\tilde{\alpha}\colon [0,1] \to E$ such that $p \circ \tilde{\alpha} = \alpha$ and $\tilde{\alpha}(0) = e_0$.

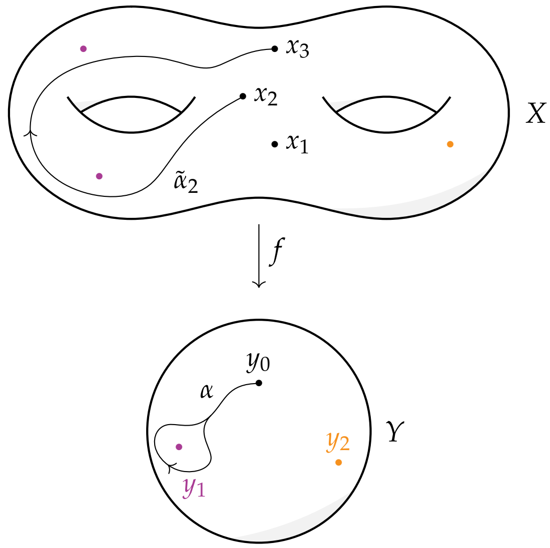


FIGURE 27. The lift of a path α on the base generates a permutation σ_α. In this example, we have $\sigma_\alpha(2) = 3$.

and $f'^{-1}(y_0) = \{ x'_1, \dots, x'_d \}$ are called isomorphic if there exists a biholomorphism $\varphi\colon X \to X'$ such that $f' = f \circ \varphi$ and $\varphi(x_k) = x'_k$.

The above construction shows that there exists a map of sets, called the monodromy map

$$\left\{ \begin{array}{c} \text{isomorphism classes of } y_0\text{-labelled Hurwitz covers} \\ f\colon X \to Y \text{ of degree } d \\ \text{with } X \text{ possibly disconnected,} \\ f \text{ branched over } (y_1, \dots, y_m) \\ \text{with ramification profiles } (\mu_1, \dots, \mu_m) \end{array} \right\} \xrightarrow{\mathsf{M}} \left\{ \begin{array}{c} \text{group homomorphism } \rho\colon \pi_1(Y^\times, y_0) \to S_d \\ \rho(\gamma_i) \text{ has cycle type } \mu_i \\ \text{for } \gamma_i \text{ a loop around } y_i \end{array} \right\}. \tag{6.8}$$

We call elements on the domain set: y_0-labelled isomorphism classes of possibly disconnected covers of type $(g, h, d, \mu_1, \dots, \mu_m)$; elements in the codomain set are called: possibly disconnected monodromy representations of type $(g, h, d, \mu_1, \dots, \mu_m)$.

Lecture 13
May 30th, 2024

We can actually do much better: we can invert the monodromy map by gluing together polygons following the instructions provided by a monodromy representation! In other words, to given a monodromy representations ρ of type $(g, h, d, \mu_1, \dots, \mu_m)$ we can associate to it an isomorphism class $[f]$ of y_0-cover of type $(g, h, d, \mu_1, \dots, \mu_m)$ such that $\mathsf{M}([f]) = \rho$. This is essentially a refined version of Riemann's existence theorem.

Theorem 6.6. *There is a one-to-one correspondence of sets*

$$\left\{ \begin{array}{c} \textit{isomorphism classes of } y_0\textit{-labelled Hurwitz covers} \\ f\colon X \to Y \textit{ of degree } d \\ \textit{with } X \textit{ is possibly disconnected,} \\ f \textit{ branched over } (y_1, \dots, y_m) \\ \textit{with ramification profiles } (\mu_1, \dots, \mu_m) \end{array} \right\} \overset{1:1}{\longleftrightarrow} \left\{ \begin{array}{c} \textit{group homomorphism } \rho\colon \pi_1(Y^\times, y_0) \to S_d \\ \rho(\gamma_i) \textit{ has cycle type } \mu_i \\ \textit{for } \gamma_i \textit{ a loop around } y_i \end{array} \right\}. \tag{6.9}$$

One question still stands: how do we compute the original Hurwitz numbers? The main difference between Hurwitz covers (what we want to count) and y_0-labelled Hurwitz cover (what is in bijection with monodromy representations) is precisely in the labelling. It is not hard to prove that the counting of Hurwitz covers weighted by their automorphisms is precisely the number of y_0-labelled Hurwitz cover, divided by the total number of labelling: $d!$. In other words, we have the following result.

Corollary 6.7 (Hurwitz numbers by counting monodromy representations)**.** *Let $M^{\bullet}_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_m)$ be the set of possibly disconnected monodromy representations of type $(g, h, d, \mu_1, \ldots, \mu_m)$. Then*

$$H^{\bullet}_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_m) = \frac{1}{d!} \left| M^{\bullet}_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_m) \right| . \tag{6.10}$$

In particular, we see that the definition of Hurwitz numbers is well-posed: the right-hand side is finite (as the fundamental group $\pi_1(Y^{\times}, y_0)$ is finitely generated and the symmetric group is a finite group) and it only depends on the discrete data $(g, h, d, \mu_1, \ldots, \mu_m)$.

6.3. **Burnside formula: from group theory to representation theory.** At first sight, the above corollary might seem a complication rather than a simplification. The reason why counting monodromy representations is much easier than counting Hurwitz covers is that the structure of the fundamental group of a surface is fairly well-understood. Namely, if Y is a genus h curve, then the fundamental group of the punctured $Y^{\times} = Y \setminus \{ y_1, \ldots, y_m \}$ can be presented in a standard way as a set of generators corresponding to loops winding around the handles, a set of loops winding around the branch points, satisfying a simple relation.

To simplify the discussion, consider the case of a genus 0 base curve: $Y = \mathbb{P}^1$. In this case the fundamental group of $\mathbb{P}^1 \setminus \{ y_1, \ldots, y_m \}$ is presented in terms of m generators and a single relation:

$$\pi_1(\mathbb{P}^1 \setminus \{ y_1, \ldots, y_m \} , y_0) \cong \langle \gamma_1, \ldots, \gamma_m \,|\, \gamma_1 \cdots \gamma_m \rangle , \tag{6.11}$$

where γ_i is a small loop around y_i. The product is simply the concatenation of loops.

As a consequence, we deduce that a monodromy representation of type $(g, 0, d, \mu_1, \ldots, \mu_m)$ is simply a choice of permutations $\sigma_1, \ldots, \sigma_m$ multiplying to the identity and having cycle type $(\mu_1, \ldots, \mu_m)$:

$$H^{\bullet}_{g \xrightarrow{d} 0}(\mu_1, \ldots, \mu_m) = \frac{1}{d!} \left| \left\{ \sigma_1, \ldots, \sigma_m \in S_d \;\middle|\; \substack{\sigma_1 \cdots \sigma_m = \mathrm{id}\,, \\ \sigma_i \text{ has cycle type } \mu_i} \right\} \right| . \tag{6.12}$$

In other words, we have reduced the counting of Hurwitz covers to the counting of permutations in the symmetric group satisfying certain properties!

The main trick to compute such numbers is the following. Consider the so-called group algebra of the symmetric group:

$$\mathbb{C}[S_d] \coloneqq \left\{ \sum_{\sigma \in S_d} a_\sigma \, \sigma \;\middle|\; a_\sigma \in \mathbb{C} \right\} . \tag{6.13}$$

It has an algebra structure, where the vector space structure is defined as

$$\lambda \left(\sum_{\sigma \in S_d} a_\sigma \, \sigma \right) + \mu \left(\sum_{\sigma \in S_d} b_\sigma \, \sigma \right) = \sum_{\sigma \in S_d} (\lambda a_\sigma + \mu b_\sigma) \sigma \qquad \forall \lambda, \mu \in \mathbb{C} , \tag{6.14}$$

and the (non-commutative) product is defined as the $\mathbb{C}$-linear extension of the usual product on S_d. For any $\mu \vdash d$, define

$$\mathcal{C}_\mu := \{ \sigma \in S_d \mid \sigma \text{ has cycle type } \mu \} , \qquad C_\mu := \sum_{\sigma \in \mathcal{C}_\mu} \sigma . \tag{6.15}$$

The element $C_\mu \in \mathbb{C}[S_d]$ collects in a single term all permutations of cycle type μ, and is called the conjugacy class element of type μ.

Exercise 6.2. *Compute $x + y$ and $x \cdot y$ in the group algebra $\mathbb{C}[S_3]$, where $x = 3(1\,2) + 5(1\,2\,3)$ and $y = 4(1\,3) - 6(1\,2\,3)$. Compute C_μ for $\mu = (2, 1)$.*

With this language, it is now easy to describe the above counting.

Theorem 6.8 (Hurwitz numbers by counting permutations)**.** *The following equation holds:*

$$H^\bullet_{g \xrightarrow{d} 0}(\mu_1, \ldots, \mu_m) = \frac{1}{d!} [\mathrm{id}] C_{\mu_1} \cdots C_{\mu_m} , \tag{6.16}$$

where $[\mathrm{id}] \colon \mathbb{C}[S_d] \to \mathbb{C}$ is the operator extracting the coefficient of the identity in the expansion of the product $C_{\mu_1} \cdots C_{\mu_m}$.

As the next example and exercise show, the above formula is extremely easy to implement: it reduces the counting of Hurwitz covers to a simple expansion of monomials and product in the symmetric group. The formula can even be implemented on a computer!

Example 6.9. Consider in $\mathbb{C}[S_2]$

$$[\mathrm{id}] C_{(2)}^2 = [\mathrm{id}] (1\,2)^2 = 1 . \tag{6.17}$$

The computation gives $H^\bullet_{0 \xrightarrow{2} 0}\big((2), (2)\big) = \frac{1}{2}$, which geometrically is the cover

$$\begin{aligned} \mathbb{P}^1 &\xrightarrow{\quad f \quad} \mathbb{P}^1 \\ [z_0 : z_1] &\longmapsto [z_0^2 : z_1^2] \end{aligned} \tag{6.18}$$

from example 2.12 (with the non-trivial automorphism $[z_0 : z_1] \mapsto [-z_0 : z_1]$).

A similar example, again in $\mathbb{C}[S_2]$, is

$$[\mathrm{id}] C_{(2)}^4 = [\mathrm{id}] (1\,2)^4 = 1 . \tag{6.19}$$

This gives $H^\bullet_{1 \xrightarrow{2} 0}\big((2), (2), (2), (2)\big) = \frac{1}{2}$, which geometrically is the projectivisation of the cover

$$\begin{aligned} E &\xrightarrow{\quad f \quad} \mathbb{P}^1 \\ [z_0 : z_1 : z_2] &\longmapsto [z_0 : z_2] \end{aligned} \tag{6.20}$$

for $E = Z\big(z_1^2 z_2 - (z_0 - \alpha_1 z_2)(z_0 - \alpha_2 z_2)(z_0 - \alpha_3 z_2)\big)$ the elliptic curve of example 2.13 (with the non-trivial automorphism $[z_0 : z_1 : z_2] \mapsto [z_0 : -z_1 : z_2]$).

Exercise 6.3. *Compute $H^\bullet_{0 \xrightarrow{3} 0}\big((3), (3)\big)$ by counting permutation. Can you generalise the computation to $H^\bullet_{0 \xrightarrow{d} 0}\big((d), (d)\big)$? And can you motivate why $H^\bullet_{0 \xrightarrow{d} 0}\big((d), (d)\big) = H_{0 \xrightarrow{d} 0}\big((d), (d)\big)$ in this case?*

Exercise 6.4. *Show that* $H^\bullet_{0\xrightarrow{4}0}\big((3,1),(2,2),(2,2)\big) = 0$ *by counting permutation.*

The previous exercise demonstrates that, even if the discrete data satisfies the Riemann–Hurwitz formula, it is still possible that no Hurwitz covers exist for that data. Remarkably, determining the necessary and sufficient conditions for a Hurwitz number to be non-zero remains an open problem, known as the Hurwitz existence problem.

Remark 6.10. Theorem 6.8 can be generalised to a target Riemann surface Y of arbitrary genus h as follows. First, the fundamental group of $Y^\times = Y \setminus \{ y_1, \dots, y_m \}$ admits the presentation

$$\pi_1(Y^\times, y_0) \cong \langle \alpha_1, \beta_1, \dots, \alpha_h, \beta_h, \gamma_1, \dots, \gamma_m \mid [\alpha_1, \beta_1] \cdots [\alpha_h, \beta_h]\gamma_1 \cdots \gamma_m \rangle \,, \tag{6.21}$$

where α_j, β_j are the loops defined by the sides of the fundamental polygon of Y and γ_i is a small loop around y_i. The relation can be easily seen from the fundamental polygon. Secondly, the relation can be encoded in the following element of the group algebra:

$$\mathfrak{K} := d! \sum_{\mu \vdash d} \frac{1}{|\mathcal{C}_\mu|} C_\mu^2 \in \mathbb{C}[S_d] \,, \tag{6.22}$$

so that the Hurwitz number is computed as

$$H^\bullet_{g\xrightarrow{d}h}(\mu_1, \dots, \mu_m) = \frac{1}{d!} \, [\mathrm{id}] \mathfrak{K}^h C_{\mu_1} \cdots C_{\mu_m} \,. \tag{6.23}$$

Bonus paragraph (not required for the exam)

For the reader familiar with the representation theory of the symmetric group, we conclude with one more formula for Hurwitz numbers known as Burnside formula. First, notice that the conjugacy class elements belong to a subalgebra of $\mathbb{C}[S_d]$: its centre

$$Z\mathbb{C}[S_d] := \{ x \in \mathbb{C}[S_d] \mid xy = yx \text{ for all } y \in \mathbb{C}[S_d] \} \,. \tag{6.24}$$

Thus, all computations regarding Hurwitz numbers can performed on the (much nicer) subalgebra $Z\mathbb{C}[S_d]$ of $\mathbb{C}[S_d]$. Secondly, recall[13] that the irreducible representations and the conjugacy classes of element in the symmetric group S_d and the are in one-to-one correspondence with partitions of d. In particular, it makes sense to consider the character $\chi_\lambda(\mu)$, that is the trace of any element of conjugacy type μ in the irreducible representation labelled by λ. We also denote by $\dim(\lambda)$ the dimension of the irreducible representation labelled by λ. The link between the centre of the group algebra of S_d and its irreducible characters is provided by Maschke's theorem.

Lemma 6.11 (Maschke's theorem)**.** *The conjugacy class elements form a basis of* $Z\mathbb{C}[S_d]$ *as a complex vector space:*

$$Z\mathbb{C}[S_d] \cong \bigoplus_{\mu \vdash d} \mathbb{C} C_\mu \,. \tag{6.25}$$

[13]Here is a crash course on representation theory. Let G be a group; a (complex) representation of G is the data of a finite-dimensional complex vector space V together with a group homomorphism $\rho \colon G \to \mathrm{GL}(V)$. A representation is is called irreducible if it contains no non-trivial subspaces that are invariant under the group action. The character of a representation ρ is the map $\chi_\rho \colon G \to \mathbb{C}$, $g \mapsto \mathrm{tr}(\rho(g))$. Notice that the character only depends on the conjugacy class of an element g. If the irreducible representations of G are indexed as $\{ (V_\lambda, \rho_\lambda) \}_\lambda$ and the conjugacy classes are indexed as $\{ C_\mu \}_\mu$, we denote by $\chi_\lambda(\mu)$ the character of any element in C_μ for the representation $(V_\lambda, \rho_\lambda)$.

Further, as an algebra, it is semisimple. This means that there exists a basis $(e_\lambda)_{\lambda \vdash d}$, called the idempotent basis, such that $e_\lambda \cdot e_{\lambda'} = \delta_{\lambda,\lambda'} e_\lambda$. The change of basis $(C_\mu)_{\mu \vdash d} \leftrightarrow (e_\lambda)_{\lambda \vdash d}$ is explicitly given as

$$C_\mu = |\mathcal{C}_\mu| \sum_{\lambda \vdash d} \frac{\chi_\lambda(\mu)}{\dim(\lambda)} e_\lambda \,, \qquad\qquad e_\lambda = \frac{\dim(\lambda)}{d!} \sum_{\mu \vdash d} \chi_\lambda(\mu)\, C_\mu \,. \tag{6.26}$$

Thanks to Maschke's theorem, we can easily deduce Burnside's formula, expressing Hurwitz numbers in terms of characters of the symmetric group.

Theorem 6.12 (Burnside character formula)**.** *The following equation holds:*

$$H^{\bullet}_{g \xrightarrow{d} 0}(\mu_1, \dots, \mu_m) = \sum_{\lambda \vdash d} \left(\frac{\dim(\lambda)}{d!} \right)^2 \prod_{i=1}^{m} |\mathcal{C}_{\mu_i}| \, \frac{\chi_\lambda(\mu_i)}{\dim(\lambda)} \,. \tag{6.27}$$

Proof. Start from equation (6.16), expressing Hurwitz numbers in terms of conjugacy class elements, and let us perform the change of basis writing the C_{μ_i}'s in the idempotent basis. Since the basis is idempotent, the product simplifies considerably:

$$\begin{aligned} H^{\bullet}_{g \xrightarrow{d} 0}(\mu_1, \dots, \mu_m) &= \frac{1}{d!} \, [\mathrm{id}] C_{\mu_1} \cdots C_{\mu_m} \\ &= \frac{1}{d!} \, [\mathrm{id}] \sum_{\lambda_1, \dots, \lambda_m \vdash d} \left(\prod_{i=1}^{m} |\mathcal{C}_{\mu_i}| \frac{\chi_{\lambda_i}(\mu_i)}{\dim(\lambda_i)} \right) e_{\lambda_1} \cdots e_{\lambda_m} \\ &= \frac{1}{d!} \, [\mathrm{id}] \sum_{\lambda \vdash d} \left(\prod_{i=1}^{m} |\mathcal{C}_{\mu_i}| \frac{\chi_\lambda(\mu_i)}{\dim(\lambda)} \right) e_\lambda \,. \end{aligned} \tag{6.28}$$

On the other hand, Maschke's theorem gives an explicit description of the coefficient of the identity in e_λ:

$$e_\lambda = \frac{\dim(\lambda)^2}{d!} \mathrm{id} + \cdots \,, \tag{6.29}$$

where the dots stand for elements not proportional to the identity. The equation is due to the fact that $\dim(\lambda) = \chi_\lambda((1, \dots, 1))$, $C_{(1,\dots,1)} = \mathrm{id}$, and the identity does not appear in any other conjugacy class element. As a consequence, we find

$$\begin{aligned} H^{\bullet}_{g \xrightarrow{d} 0}(\mu_1, \dots, \mu_m) &= \frac{1}{d!} \, [\mathrm{id}] \sum_{\lambda \vdash d} \left(\prod_{i=1}^{m} |\mathcal{C}_{\mu_i}| \frac{\chi_\lambda(\mu_i)}{\dim(\lambda)} \right) \left(\frac{\dim(\lambda)^2}{d!} \mathrm{id} + \cdots \right) \\ &= \sum_{\lambda \vdash d} \left(\frac{\dim(\lambda)}{d!} \right)^2 \prod_{i=1}^{m} |\mathcal{C}_{\mu_i}| \, \frac{\chi_\lambda(\mu_i)}{\dim(\lambda)} \,. \end{aligned} \tag{6.30}$$

This concludes the proof. ☐

Remark 6.13. Burnside's formula also generalises to targets of arbitrary genus h as

$$H^{\bullet}_{g \xrightarrow{d} h}(\mu_1, \dots, \mu_m) = \sum_{\lambda \vdash d} \left(\frac{\dim(\lambda)}{d!} \right)^{2-2h} \prod_{i=1}^{m} |\mathcal{C}_{\mu_i}| \, \frac{\chi_\lambda(\mu_i)}{\dim(\lambda)} \,. \tag{6.31}$$

Although perhaps 'less practical' compared to the expression in terms of permutations, Burnside's formula is a formidable tool. It links Hurwitz numbers to the elegant representation theory of the symmetric group. Surprisingly, this theory is closely related to that of soliton waves and to the topological string theory of 1-dimensional space-times. Burnside's formula is a crucial step in establishing these connections.

7. What next?

Bonus chapter (not required for the exam)

7.1. **Algebraic geometry.** One of the main consequences of Riemann–Roch is that all compact Riemann surfaces are algebraic. It turns out that many of the desirable properties of compact Riemann surfaces are preserved when dealing with projective varieties, which are subsets of some projective space defined as the zero-locus of a finite collection of homogeneous polynomials.

There are several advantages to this point of view. Firstly, most of the theory works seamlessly when considering polynomial equations over an arbitrary algebraically closed field k. Secondly, the theory naturally accommodates potentially singular objects. For example, the affine curve defined by the equation $xy = 0$ is singular (it is the union of two crossing lines), but the singularity is mild enough (being defined by a simple polynomial equation) to not pose significant problems. Thirdly, one can use powerful tools from commutative algebra rather than analysis to prove the desired results.

Here is a list of concepts and results introduced in this course that naturally generalise to an arbitrary n-dimensional projective variety X (or schemes, stack, or further generalisations).

- The ring of *holomorphic functions* is generalised by that of regular functions on X, which is better described as a sheaf $\mathcal{O}_X$ of rings on X. For a variety X over a field k we have $\mathcal{O}_X(X) = k$, which generalises Liouville's theorem for compact Riemann surfaces.
- The field of *meromorphic functions* is generalised by the so-called function field of X, usually denoted $k(X)$. Elements of $k(X)$ are locally ratios of regular functions, in the same way that meromorphic functions are locally ratios of holomorphic functions.
- The concept of a *divisor* generalises to that of Weil divisors, which are formal sums of codimension 1 subvarieties. The concepts of principal divisors and the Picard group also extend without too many modifications (although the correct terminology in this context would be the Weil divisor class group).
- The *spaces of meromorphic functions and forms*, $\mathcal{L}(D)$ and $\mathcal{I}(D)$, generalise considerably within the framework of sheaf theory. They are instances of the 0-th and 1-st cohomology groups of the sheaf $\mathcal{O}(D)$ associated with the divisor D. Sheaf theory generalises most of the results we proved for the spaces of meromorphic functions and meromorphic forms, with some complications arising from the higher dimension of the variety. These include:
  - *Finiteness theorem*. For a coherent sheaf $\mathcal{F}$, the cohomology groups $H^p(X, \mathcal{F})$ are finite-dimensional vector spaces which are trivial for all $p > \dim X$.
  - *Serre duality*. For smooth X, there exists a sheaf, called the canonical sheaf ω_X and constructed through differential forms on X, such that for any coherent sheaf $\mathcal{F}$ the following duality holds:

$$H^p(X, \mathcal{F}) \cong H^{n-p}(X, \mathcal{F}^* \otimes \omega_X)^* . \tag{7.1}$$

  Albeit not, strictly speaking, a generalisation of Poincaré duality from differential topology, Serre duality serves a similar role in algebraic geometry.

– *Riemann–Roch formula*. As mentioned before, Riemann–Roch generalises to a relative statement due to Grothendieck regarding a proper map $f\colon X \to Y$ between smooth projective varieties. It relates the pushforward map in K-theory

$$f_! \coloneqq \sum_{p=0}^{n} (-1)^p \, R^p f_* \colon K_0(X) \longrightarrow K_0(Y)\,, \tag{7.2}$$

which can be thought of as a generalisation of the Euler characteristic, to the pushforward map in cohomology

$$f_*\colon H(X) \longrightarrow H(Y) \tag{7.3}$$

by the formula

$$\mathrm{ch}(f_!\mathcal{F}) = f_*\big(\mathrm{ch}(\mathcal{F})\mathrm{td}(T_f)\big)\,. \tag{7.4}$$

Here $\mathrm{ch}\colon K_0(X) \to H(X)$ is the Chern character, $\mathrm{td}(T_f)$ is the Todd genus of the relative tangent sheaf of f, and $\mathcal{F}$ is an arbitrary bounded complex of coherent sheaves. If X is a Riemann surface, $Y = \{\,*\,\}$ is a point, and $\mathcal{F} = \mathcal{O}(D)$ for a divisor D, the formula reduces to Riemann–Roch as

$$\mathrm{ch}(f_!\mathcal{F}) = \ell(D) - \ell(K-D)\,, \qquad f_*\big(\mathrm{ch}(\mathcal{F})\mathrm{td}(T_f)\big) = \deg(D) + 1 - g\,. \tag{7.5}$$

- There are various generalisations of the *Abel–Jacobi theory*, although not as rich as in the curve case. One example is the intermediate Jacobians:

$$J^{k+1}(X) \coloneqq H^{2k+1}(X,\mathbb{R})/H^{2k+1}(X,\mathbb{Z})\,, \qquad k = 0,\dots,n-1\,, \tag{7.6}$$

which reduces to the ordinary Jacobian for $\dim(X) = 1$ and $k = 0$. Other generalisations include the Picard variety and the Albanese variety. An interesting concept that generalises well is that of Riemann bilinear relations, which can be formalised in the theory of Hodge structures.

Other key concepts relevant to Riemann surfaces, however, do not generalise to higher dimensions. This is the case for most 'topological' notions and statements such as the Riemann–Hurwitz formula and Hurwitz theory. The main reason behind this difference is that compact orientable topological surfaces are uniquely characterised by their genus. In higher dimensions, this is not the case, making all related concepts much more challenging. For instance, one can define the geometric genus p_g and the arithmetic genus p_a of a projective variety as

$$p_g \coloneqq \dim H^0(X,\omega_X) \qquad \text{and} \qquad p_a \coloneqq (-1)^n\big(\chi(\mathcal{O}_X) - 1\big)\,. \tag{7.7}$$

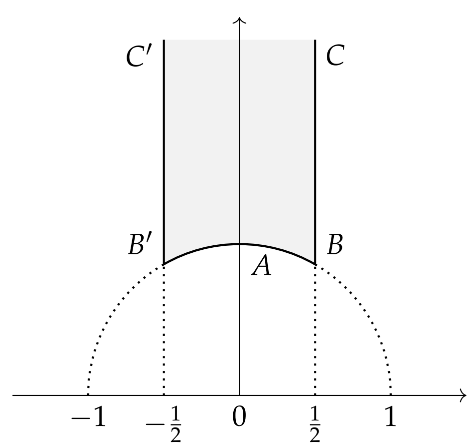
For a smooth variety. X, one finds $p_g = p_a$, and they both coincide with the topological genus if X is a Riemann surface (since $H^0(X,\omega_X) = \Omega(X)$ and $\chi(\mathcal{O}_X) = 1 - g$). Albeit more complicated, the classification of varieties based on invariants such as the geometric and arithmetic genus is an important field of algebraic geometry known as birational geometry.

7.2. **Moduli spaces, string theory, and all that.** A different direction that, in a sense, remains within the realm of Riemann surfaces is the study of their moduli spaces, which aims at understanding the various ways we can regard a genus g topological surface as a Riemann surface. One of the consequences of Riemann–Roch is that for $g = 0$ there is a unique way to endow a sphere with a complex structure (this is why it is sensible to refer to *the* Riemann sphere). However, for $g > 0$, there are infinitely many distinct complex structures. Thus, it is meaningful to consider a space parametrising all genus g Riemann surfaces: a *moduli space*.

A moduli space $\mathcal{M}$ is, first and foremost, a topological space. What distinguishes a moduli space from merely a topological space is that the underlying set of points corresponds bijectively with certain interesting geometric objects. More precisely, each point of $\mathcal{M}$ corresponds to an isomorphism class of some object we wish to study, and the intuitive idea is that two points in $\mathcal{M}$ are 'close' to each other if the corresponding isomorphism classes of geometric objects are 'close' to each other. A classical example is the projective space $\mathbb{P}^n$ parametrising lines in $\mathbb{C}^{n+1}$. Each point of $\mathbb{P}^n$ corresponds to a line, and moving continuously from one point to a nearby point amounts to continuously adjusting the corresponding lines.

Similarly, we have the moduli space $\mathcal{M}_g$, whose points correspond to isomorphism classes of genus g Riemann surfaces. In other words, two Riemann surfaces X and X' represent the same point in $\mathcal{M}_g$ if they are isomorphic. The topology on $\mathcal{M}_g$ is difficult to describe formally, but it satisfies some intuitive properties. For instance, if $X = Z(F)$ for a certain homogeneous polynomial F, then slightly altering the coefficients of F will result in a new Riemann surface X' that we perceive as close to X inside $\mathcal{M}_g$. The example of genus 1 surfaces was illustrated by the so-called modular curve, where closely related lattices correspond to closely related tori.

$$\left\{\begin{matrix}\text{genus } 1\\ \text{Riemann surfaces}\end{matrix}\right\}\Big/\!\sim \;\cong\; \mathbb{H}\big/\mathrm{SL}(2,\mathbb{Z}) \;= \tag{7.8}$$



Understanding the geometry of the moduli space $\mathcal{M}_g$ is equivalent to understanding all possible complex structures that can be imposed on a genus g surface. Hurwitz theory emerged as a method to address some of these questions. For example, it can be shown using Hurwitz theory that $\mathcal{M}_g$ is connected. Geometrically, this means that any complex structure can be continuously deformed into any other. Another interesting application of Hurwitz theory provides the dimension of $\mathcal{M}_g$: for $g \geq 2$,

$$\dim \mathcal{M}_g = 3g - 3\,. \tag{7.9}$$

This result was already known to Riemann himself, who also coined the term 'moduli space' (from the Latin word *modus*, meaning measure):

> Diese Bestimmung der Anzahl der Moduln einer Klasse $\overline{2p+1}$ fach zusammenhängender algebraischer Functionen gilt jedoch nur unter der Voraussetzung, dass es $2\mu - p + 1$ Verzweigungswerthe giebt, welche von einander unabhängige Functionen der willkürlichen Constanten in der Function ζ sind. Diese Voraussetzung trifft nur zu, wenn $p > 1$, und die Anzahl der Moduln ist nur dann $= 3p - 3$, für $p = 1$ aber $= 1$. Die directe Untersuchung derselben wird indess schwierig durch die Art und Weise, wie die willkürlichen Constanten in ζ enthalten sind. Man führe deshalb in einem Systeme gleichverzweigter $\overline{2p+1}$ fach zusammenhängender Functionen, um die Anzahl der Moduln zu bestimmen, als unabhängig veränderliche Grösse nicht eine dieser Functionen, sondern ein allenthalben endliches Integral einer solchen Function ein.

Riemann's argument combines (what are now called) the Riemann–Roch and Riemann–Hurwitz formulae. He starts by considering the moduli space of pairs (X, f), where X is a genus g Riemann surface and f is a degree d holomorphic map from X to $\mathbb{P}^1$ (i.e. a meromorphic function on X). Such a space is sometimes referred to as the Hurwitz space, denoted $\mathcal{H}_{g,d}$. Riemann computes its dimension in two different ways.

Method 1. On the one hand, the dimension of $\mathcal{H}_{g,d}$ is simply the dimension of $\mathcal{M}_g$, counting the 'number of deformation parameters' of the Riemann surface X, plus the 'number of deformation parameters' of the function f. A generic f has d simple poles that can be fixed at $x_1, \ldots, x_d$, providing d degrees of freedom. Once the poles have been fixed, the number of degree d functions $f \colon X \to \mathbb{P}^1$ with simple poles at $x_1, \ldots, x_d$ is computed by Riemann–Roch: for $D = x_1 + \cdots + x_d$ we have $\ell(D) = d + 1 - g$, assuming d is large enough so that $\ell(K - D) = 0$. To sum up, we find

$$\dim \mathcal{H}_{g,d} = \underbrace{\dim \mathcal{M}_g}_{\text{moving } X} + \underbrace{d}_{\substack{\text{fixing} \\ d \text{ poles}}} + \underbrace{d + 1 - g}_{\substack{\text{moving } f \\ \text{with fixed poles}}} = \dim \mathcal{M}_g + 2d + 1 - g\,. \tag{7.10}$$

Method 2. On the other hand, a generic degree d holomorphic map $f \colon X \to \mathbb{P}^1$ has only simple ramification points that can be arbitrary, so $\dim \mathcal{H}_{g,d} = \#\mathsf{Ram}_f$. This number is provided by Riemann–Hurwitz: if f has only simple ramification points,

$$2g - 2 = -2d + \#\mathsf{Ram}_f\,, \tag{7.11}$$

so that $\dim \mathcal{H}_{g,d} = 2d + 2g - 2$.

The two computations imply the claimed formula: $\dim \mathcal{M}_g = 3g - 3$. The argument can be made rigorous by saying that the map

$$p \colon \mathcal{H}_{g,d} \longrightarrow \mathcal{M}_g\,, \qquad (X, f) \longmapsto X \tag{7.12}$$

which 'forgets' the function f is (the algebraic-geometric equivalent of) a fibration, so that by the fibre-dimension theorem

$$\dim \mathcal{H}_{g,d} = \dim \mathcal{M}_g + \dim F \tag{7.13}$$

for F a generic fibre of p. Riemann's argument described above is essentially the computation of $\dim F = 2d + 1 - g$ using Riemann–Roch and $\dim \mathcal{H}_{g,d} = 2d + 2g - 2$ using Riemann–Hurwitz.

It is also worth mentioning an exciting application of Hurwitz numbers in relation to the moduli space of Riemann surfaces: the ELSV formula, named for its discoverers Ekedhal, Lando, Shapiro, and Vainshtein. The ELSV formula expresses certain integrals over the moduli space of Riemann surfaces in terms of Hurwitz numbers. The integrals involved are fundamentally important but relatively intractable, while we have seen that Hurwitz numbers are combinatorial and computable. This connection is a key step, in conjunction with Burnside's formula, in Okounkov and Pandharipande's solution of the topological string theory of one-dimensional space-times.

At this point, one can ask: why would physics have anything to do with Riemann surface? It turns out that physics, and in particular *string theory*, does indeed have much to say Riemann surfaces and their moduli space. String theory is a branch of theoretical physics where elementary particles are replaced by strings. As a string travels through space-time it traces out a Riemann surface, that is the worldsheet of the string (see figure 28). These are a stringy versions of Feynman diagrams. The path integrals of the theory are mathematically described as integrals over the moduli spaces of Riemann surfaces mapping to the space-time.

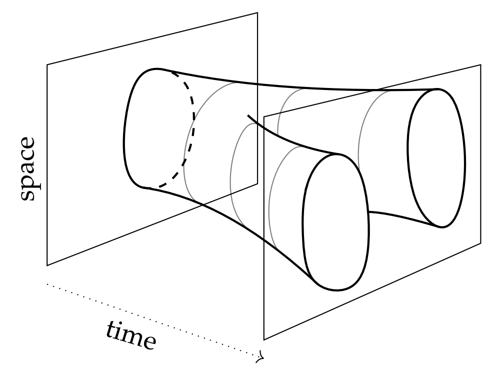


FIGURE 28. A worldsheet in space-time.

One of the most useful tools at the disposal of physicists is physical dualities. Roughly speaking, physicists aim to construct mathematical models that explain the universe, but occasionally, multiple mathematical models yield the same observable quantities. When this occurs, both models are considered valid physical descriptions from a physics standpoint. We say that these models are dual. However, dualities have an unexpected consequence. Dual physical models may be constructed using completely different mathematical structures. For instance, a given physical model may be formulated in the language of algebraic geometry, while a dual model may involve objects from number theory. Consequently, the physical duality implies a connection between certain objects defined in seemingly unrelated areas of mathematics!