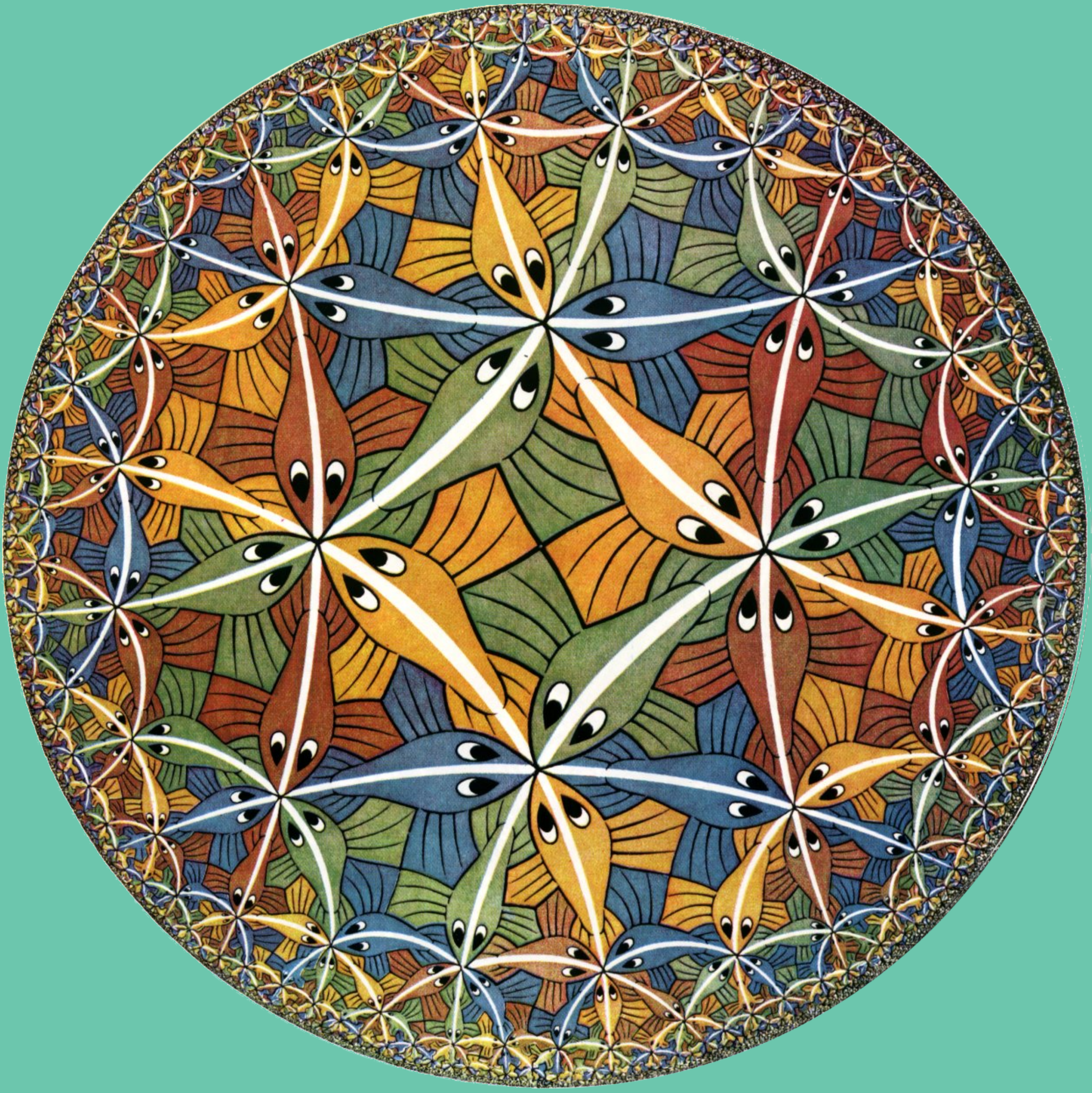
Riemann Surfaces



UT Austin, Spring 2016

M392C NOTES: RIEMANN SURFACES

ARUN DEBRAY
APRIL 4, 2016

These notes were taken in UT Austin's Math 392C (Riemann Surfaces) class in Spring 2016, taught by Tim Perutz. I live-TEXed them using `vim`, and as such there may be typos; please send questions, comments, complaints, and corrections to a.debray@math.utexas.edu. The image on the front cover is M.C. Escher's *Circle Limit III* (1959), sourced from http://www.wikiart.org/en/m-c-escher/circle-limit-iii. Thanks to Adrian Clough for finding a few typos.

Contents

Lecture 1.

Review of Complex Analysis: 1/20/16

Riemann surfaces is a subject that combines the topology of structures with complex analysis: a Riemann surface is a surface endowed with a notion of holomorphic function. This turns out to be an extremely rich idea; it's closely connected to complex analysis but also to algebraic geometry. For example, the data of a compact Riemann surface along with a projective embedding specifies a proper algebraic curve over $\mathbb{C}$, in the domain of algebraic geometry.[1] In fact, the algebraic geometry course that's currently ongoing is very relevant to this one.

The theory of Riemann surfaces ties into many other domains, some of them quite applied: number theory (via modular forms), symplectic topology (pseudo-holomorphic forms), integrable systems, group theory, and so on: so a very broad range of mathematics graduate students should find it interesting.

Moreover, by comparison with algebraic geometry or the theory of complex manifolds, there's very low overhead; we will quickly be able to write down some quite nontrivial examples and prove some deep theorems: by the middle of the semester, hopefully we will prove the analytic Riemann-Roch theorem, the fundamental theorem on compact Riemann surfaces, and use it to prove a classification theorem, called the uniformization theorem.

The course textbook is S.K. Donaldson's *Riemann Surfaces*, and the course website is at `http://www.ma.utexas.edu/users/perutz/RiemannSurfaces.html`; it currently has notes for this week's material, a rapid review of complex function theory. We will assume a small amount of complex analysis (on the level of Cauchy's theorem; much less than the complex analysis prelim) and topology (specifically, the relationship between the fundamental group and covering spaces). Some experience with calculus on manifolds will be helpful. Some real analysis will be helpful, and midway through the semester there will be a few Hilbert spaces. Thus, though this is a topics course, the demands on your knowledge will more resemble a prelim course.

Let's warm up by (quickly) reviewing basic complex analysis; the notes on the course website will delve into more detail. For the rest of this lecture, G denotes an open set in $\mathbb{C}$ (from German *gebiet*, which commonly denotes an open set).

The following definition is fundamental.

Definition. A function $f : G \to \mathbb{C}$ is *holomorphic* if for all $z \in G$, the complex derivative

$$f'(z) = \lim_{h\to 0} \frac{f(z+h) - f(z)}{h}$$

exists. The set of holomorphic functions $G \to \mathbb{C}$ is denoted $\mathscr{O}(G)$, after the Italian *funzione olomorfa*.

Even though it makes sense for this limit to be infinite, this is not allowed.

First, let's establish a few basic properties.

- If $H \subset G$ is open and $f \in \mathscr{O}(G)$, then $f|_H \in \mathscr{O}(H)$.
- The sum, product, quotient, and chain rules hold for holomorphic functions, so $\mathscr{O}(G)$ is a commutative ring (with multiplication given pointwise) and in fact a commutative $\mathbb{C}$-algebra.[2]

In other words, holomorphic functions define a *sheaf* of $\mathbb{C}$-algebras on G.

By a rephrasing of the definition, then if f is holomorphic on G, then it has a *derivative* f' on G, i.e. for all $z \in G$, one can write $f(z+h) = f(z) + f'(z)h + \varepsilon_z(h)$, where $\varepsilon_z(h) \in o(h)$ (that is, $\varepsilon_z(h)/h \to 0$ as $h \to 0$). Thus, a holomorphic function is differentiable in the real sense, as a function $G \to \mathbb{R}^2$. This means that there's an $\mathbb{R}$-linear map $D_z f : \mathbb{C} \to \mathbb{C}$ such that $f(z+h) = f(z) + (D_z f)(h) + o(h)$: here, $D_z f(h) = f'(z)h$.

However, we actually know that $D_z f$ is $\mathbb{C}$-linear. This is known as the *Cauchy-Riemann condition.* Since it's *a priori* $\mathbb{R}$-linear, saying that it's $\mathbb{C}$-linear is equivalent to it commuting with multiplication by i. $D_z f$ is represented by the Jacobian matrix

$$D_z f = \begin{bmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{bmatrix}.$$

A short calculation shows that this commutes with i iff the following equations, called the *Cauchy-Riemann equations*, hold:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \qquad \text{and} \qquad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}. \tag{1.1}$$

The content of this is exactly that $D_z f$ is complex linear.

[1]This sentence is packed with jargon you're not assumed to know yet.

[2]A $\mathbb{C}$-*algebra* is a commutative ring A with an injective map $\mathbb{C} \hookrightarrow A$, which in this case is the constant functions.

Conversely, suppose $f : G \to \mathbb{C}$ is differentiable in the real sense. Then, if it satisfies (1.1), then $D_z f$ is complex linear. But a complex linear map $\mathbb{C} \to \mathbb{C}$ must be multiplication by a complex number $f'(z)$, so f is holomorphic, with derivative f'.

Power Series. The notation $D(c, R)$ means the open disc centered at c with radius R, i.e. all points $z \in \mathbb{C}$ such that $|z - c| < R$.

Definition. Let $A(z) = \sum_{n=0}^{\infty} a_n(z-c)^n$ be a $\mathbb{C}$-valued power series centered at a $c \in \mathbb{C}$. Then, its *radius of convergence* is $R = \sup\{|z - c| : A(z) \text{ converges}\}$, which may be 0, a positive real number, or ∞.

Theorem 1.2. *Suppose $A(z) = \sum_{n \geq 0} a_n(z-c)^n$ has radius of convergence R. Then:*

(1) $R^{-1} = \limsup |a_n|^{1/n}$*;*
(2) $A(z)$ *converges absolutely on* $D(c, R)$ *to a function* $f(z)$*;*
(3) *the convergence is uniform on smaller discs* $D(c, r)$ *for* $r < R$*;*
(4) *the series* $B(z) = \sum_{n \geq 1} na_n(z-c)^{n-1}$ *has the same radius of convergence* R*, so converges on* $D(c, R)$ *to a function* $g(z)$*; and*
(5) $f \in \mathscr{O}(D(c, R))$ *and* $f' = g$*.*

These aren't extremely hard to prove: the first few rely on various series convergence tests from calculus, though the last one takes some more effort.

Paths and Cauchy's Theorem. By a *path* we mean a continuous and piecewise C^1 map $[a, b] \to \mathbb{C}$ for some real numbers $a < b$. That is, it breaks up into a finite number of chunks on which it has a continuous derivative. A *loop* is a path γ such that $\gamma(a) = \gamma(b)$.

If γ is a C^1 path in G (so its image is in G) and $f : G \to \mathbb{C}$ is continuous, we define the *integral*

$$\int_\gamma f = \int_\gamma f(z)\,\mathrm{d}z := \int_a^b f(\gamma(t))\gamma'(t)\,\mathrm{d}t.$$

This is a complex-valued function, because the rightmost integral has real and imaginary parts. This makes sense as a Riemann integral, because these real and imaginary parts are continuous. This is additive on the join of paths, so we can extend the definition to piecewise C^1 paths. Moreover, integrals behave the expected way under reparameterization, and so on.

Theorem 1.3 (Fundamental theorem of calculus)**.** *If $F \in \mathscr{O}(G)$ and $\gamma : [a, b] \to G$ is a path, then*

$$\int_\gamma F'(z)\,\mathrm{d}z = F(\gamma(b)) - F(\gamma(a)).$$

This is easy to deduce from the standard fundamental theorem of calculus. In particular, if γ is a loop, then the integral of a holomorphic function is 0.

Now, an extremely important theorem.

Definition. A *star-domain* is an open set $G \subset \mathbb{C}$ with a $z^* \in G$ such that for all $z \in G$, the line segment $[z^*, z]$ joining z^* and z is contained in G.

For example, any convex set is a star-domain.

Theorem 1.4 (Cauchy)**.** *If G is a star-domain, γ is a loop in G, and $f \in \mathscr{O}(G)$, then $\int_\gamma f = 0$. Indeed, $f = F'$, where*

$$F(z) = \int_{[z^*, z]} f.$$

The proof is in the notes, but the point is that you can check that this definition of F produces a holomorphic function whose derivative is f; then, you get the result. The idea is to compare $F(z + h)$ and $F(z)$ should be comparable, which depends on an explicit calculation of an integral of a holomorphic function around a triangle, which is not hard.

Cauchy didn't prove Cauchy's theorem this way; instead, he proved Green's theorem, using the Cauchy-Riemann equations. This is short and satisfying, but requires assuming that all holomorphic functions are C^1. This is true (which is great), but the standard (and easiest) way to show this is… Cauchy's theorem.

Lecture 2.

Review of Complex Analysis, II: 1/22/16

Today, we're going to continue not being too ambitious; next week we will begin to geometrify things. Last time, we stopped after Cauchy's theorem for a star domain G: for all f holomorphic on G and loops $\gamma \in G$, $\int_\gamma f = 0$, and in fact one can write down an antiderivative for f, and then apply the fundamental theorem of calculus.

Then one can bootstrap one's way up to a more powerful theorem; the next one is a version of the deformation theorem.

Corollary 2.1 (Deformation theorem)**.** *Let $G \subset \mathbb{C}$ be open and $\gamma_0, \gamma_1 : [a,b] \rightrightarrows G$ be C^1 loops that are C^1 homotopic through loops in G. Then, for all $f \in \mathscr{O}(G)$, $\int_{\gamma_0} f = \int_{\gamma_1} f$.*

Proof sketch. Fix a C^1 homotopy $\Gamma : [a,b] \times [0,1] \to G$ such that $\Gamma(a,s) = \Gamma(b,s)$ for all s, $\gamma_0(t) = \Gamma(t,0)$, and $\gamma_1(t) = \Gamma(t,1)$. Then, it is possible to divide $[a,b] \times [0,1]$ into a grid of rectangles fine enough such that the image of each rectangle is mapped under Γ to a subset of G contained in an open disc in $\mathbb{C}$, as in Figure 1. Now, by Cauchy's theorem in a disc, the integral does not depend on path within each disc, so we can apply

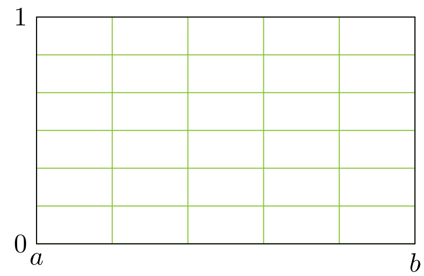


Figure 1. Subdividing $[a,b] \times [0,1]$ into rectangles.

Γ in over the rectangles from 0 to 1, showing that the two integrals are the same. ⊠

Corollary 2.2. *Cauchy's theorem holds in any simply connected open $G \subset \mathbb{C}$.*

This is considerably more general than star domains (e.g. the letter **C** is simply connected, but not a star domain). Moreover, on such a domain, any $f \in \mathscr{O}(G)$ has an antiderivative: pick some basepoint $z_0 \in G$, and let $\gamma(z_0, z)$ be a path from z_0 to z. Then,

$$F(z) = \int_{\gamma(z_0,z)} f(z)\,\mathrm{d}z$$

is well-defined, because any two choices of path differ by the integral of a holomorphic function on a loop, which is 0.

We can also use this to understand power series representations.

Proposition 2.3 (Cauchy's integral formula)**.** *, Let G be a domain in $\mathbb{C}$ containing the closed disc D. If $f \in \mathscr{O}(G)$, then*

$$f(z) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(w)}{w - z}\,\mathrm{d}w.$$

Proof idea. Suppose D is centered at z and has radius R, and let $C(z,r)$ denote the circle centered at z and with radius r. We'll also let D^* denote the punctured disc, i.e. D minus its center point. By calculating $\int_\gamma \mathrm{d}z/z = 2\pi i$, one has that

$$\frac{1}{2\pi i} \int \partial D \frac{f(w)}{z - w}\,\mathrm{d}w - f(z) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(w) - f(z)}{w - z}\,\mathrm{d}w.$$

Using Corollary 2.1, for $r \in (0, R)$,

$$= \frac{1}{2\pi i} \int_{C(z,r)} \frac{f(w) - f(z)}{w - z} \, \mathrm{d}w,$$

and as $r \to 0$, this approaches $f'(z)$, which is bounded, and the integral over smaller and smaller circles of a bounded function tends to zero. ⊠

Theorem 2.4 (Holomorphic implies analytic)**.** *If D is a disc centered at c and $f \in \mathscr{O}(D)$, then on that disc,*

$$f(z) = \sum_{n \geq 0} a_n (z - c)^n, \qquad \textit{where} \qquad a_n = \frac{1}{2\pi i} \int_{\partial D} \frac{f(z)}{(z - c)^{n+1}} \, \mathrm{d}z.$$

Proof sketch. For any $z \in D$, there's a $\delta > 0$ such that the closed disc $\overline{D}(z, \delta)$ of radius δ is contained in D. Hence, by Proposition 2.3,

$$\begin{aligned} f(z) &= \frac{1}{2\pi i} \int_{C(z,\delta)} \frac{f(w)}{w - z} \, \mathrm{d}w \\ &= \int_{C(c,R')} \frac{f(w)}{w - z} \, \mathrm{d}w \end{aligned}$$

for any $R' \in (0, \delta)$, by Corollary 2.1. We'd like to force a series on this. First, since

$$\frac{1}{w - z} = \frac{1}{(w - c) - (z - c)} = \frac{1}{w - c} \left(\frac{1}{1 - \frac{z - c}{w - c}} \right),$$

then

$$\begin{aligned} f(z) &= \frac{1}{3\pi i} \int_{C(c,R')} \frac{f(w)}{w - c} \frac{1}{1 - \frac{z - c}{w - c}} \, \mathrm{d}w \\ &= \frac{1}{2\pi i} \oint \frac{f(w)}{w - c} \sum_{n \geq 0} \frac{(z - c)^n}{(w - c)^n} \, \mathrm{d}w. \end{aligned}$$

Since $|(z - c)/(w - c)| < 1$ on $C(c, R')$, then this is well-defined, and since it's a geometric series, it has nice convergence properties, and so we can exchange the sum and integral to obtain

$$= \sum_{n \geq 0} \underbrace{\frac{1}{2\pi i} \left(\oint \frac{f(w)}{(w - c)^{n+1}} \, \mathrm{d}w \right)}_{a_n} (z - c)^n. \qquad ⊠$$

One application of this is to understand zeros of holomorphic functions. If $f \in \mathscr{O}(G)$ and $f(c) = 0$, then let $f(z) = \sum a_n (z - c)^n$ be its power series and a_m be the first nonzero coefficient. Then, in a neighborhood of c,

$$f(z) = (z - c)^m \underbrace{\sum_{n \geq m} a_n (z - c)^{n-m}}_{g(z)}.$$

This g is holomorphic and does not vanish on this neighborhood, so the takeaway is $f(z) = (z - c)^m g(z)$ near c, with g holomorphic and nonvanishing. This m is called the *multiplicity*, denoted $\operatorname{mult}(f, c)$. In particular, if $f(c) \neq 0$, then $m = 0$.

Theorem 2.5. *If G is a connected open set and $f \in \mathscr{O}(G)$ is not identically zero, then $f^{-1}(0)$ is discrete in $\mathbb{C}$.*

Proof. If $f(c) = 0$, then there's a disc D on which $f(z) = (z - c)^m g(z)$, where $m \geq 1$ and g is nonvanishing, so the only place f can vanish on D (i.e. near c) is at c itself. ⊠

Definition. A function $f \in \mathscr{O}(\mathbb{C})$, so holomorphic on the entire plane, is called *entire.*

Theorem 2.6 (Liouville)**.** *A bounded, entire function is constant.*

Proof sketch. We'll show that $f'(z) = 0$ everywhere. By Proposition 2.3, we know

$$f'(z) = \frac{1}{2\pi i} \int C(z,r) \frac{f(w)}{(w-z)^2}\, \mathrm{d}w,$$

and we can deform this loop to $C(0, R)$. Then, one bounds the integral, and the bound ends up being $O(1/R)$, so as $R \to \infty$, this necessarily goes to 0. ⊠

Lecture 3.

Meromorphic Functions and the Riemann Sphere: 1/25/16

We're still going to be doing classical function theory today, but we're going to begin to geometrify it. Recall that $G \subset \mathbb{C}$ denotes an open set.

We'll begin with the following theorem.

Theorem 3.1 (Morera)**.** *Let $f : G \to \mathbb{C}$ be a continuous function such that for all triangles $T \subset G$, $\int_{\partial T} f = 0$. Then, f is holomorphic.*

This is surprisingly easy to prove, given what we've done.

Proof. Since holomorphy is a local property, we may without loss of generality work on a disc $D(z_0, r) \subset G$. Then, define $F : D(z_0, r) \to \mathbb{C}$ by $F(z) = \int_{[z_0,z]} f$; using the hypothesis on triangles, $F' = f$. Thus, as we showed last time, this means $F \in \mathscr{O}(G)$, and so it's analytic, and therefore it has derivatives of all orders. Thus, $F' = f$ is holomorphic. ⊠

This is useful, e.g. one may have a function which is defined through an improper integral, or a pointwise limit of holomorphic functions. Then, Morera's theorem allows for an easier, indirect way to show holomorphy. Here's another application.

Definition. If $z_0 \in G$, a function $f \in \mathscr{O}(G \setminus \{z_0\})$ has a *removable singularity* at z_0 if f can be extended holomorphically to G.

Theorem 3.2. *Suppose $f \in \mathscr{O}(G \setminus \{z_0\})$ and $|f|$ is bounded near z_0. Then, f has a removable singularity at z_0.*

There are several ways to prove this quickly.

Proof. We can without loss of generality translate this to the origin, so assume $z_0 = 0$. If $g(z) = zf(z)$, then $g(z) \to 0$ as $z \to 0$, since $|f(z)|$ is bounded in a neighborhood of the origin. Thus, g extends continuously to all of G, with $g(0) = 0$.

Next, one should check that Morera's theorem applies to g; the only nontrivial example is a triangle around the origin. However, since g is holomorphic everywhere except at 0, the deformation theorem allows us to shrink the triangle as much as we want, and since $g \to 0$, the integral goes to 0 as well. If the triangle's edge or vertex touches the origin, one can use the deformation theorem to push it away again.

In particular, g is holomorphic on G and has a zero at 0, so by the discussion on multiplicities last time, $g(z) = z \cdot f(z)$, where f is holomorphic on all of G; this produces our desired extension of f. ⊠

Definition.

- If $z_0 \in G$ and $f \in \mathscr{O}(G \setminus \{z_0\})$, then f has a *pole* at z_0 if there's an $m \in \mathbb{N}$ such that $(z - z_0)^m f(z)$ is bounded near z_0 (and hence has a removable singularity there). The least such m is called the *order* of the pole.
- A *meromorphic* function on G is a pair (Δ, f) consisting of a discrete subset $\Delta \subset G$ and an $f \in \mathscr{O}(G \setminus \Delta)$ such that f has a pole at each $z \in \Delta$.

So, nothing worse than a pole happens for a meromorphic function. There are *essential singularities*, which are singularities which aren't poles, but we will not discuss them extensively; almost everything in sight will be meromorphic.

The Riemann Sphere. In some sense, the Riemann sphere is the most natural setting for meromorphic functions, and the first nontrivial example of a Riemann surface (still to be defined).

Definition. The *Riemann sphere* $\widehat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ is the *one-point compactification* of $\mathbb{C}$, so its topology has as its open sets (1) opens in $\mathbb{C}$, and (2) $(\mathbb{C} \setminus K) \cup \{\infty\}$, where $K \subset \mathbb{C}$ is compact.

There is a homeomorphism $\phi : \widehat{\mathbb{C}} \to S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ given by *stereographic projection*: send $\infty \mapsto (0, 0, 1)$ (the north pole), and then any other $z \in \mathbb{C}$ defines a line from z in the xy-plane to $(0, 0, 1)$ intersecting S^2 at one other point; this is $\phi(z)$. Hence, we will use $\widehat{\mathbb{C}}$ and S^2 interchangeably.

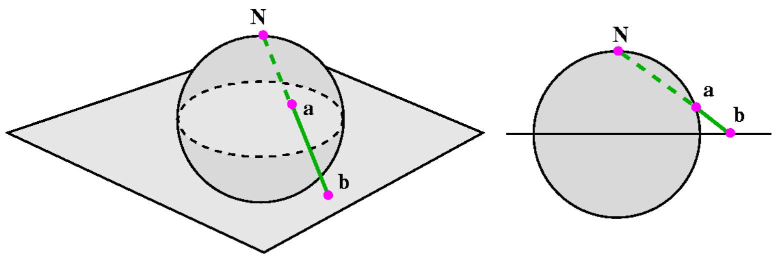


FIGURE 2. Depiction of stereographic projection, where $N = (0, 0, 1)$ is the north pole. Source: http://www.math.rutgers.edu/~greenfie/vnx/math403/diary.html.

Definition. A continuous map $f : G \to S^2$ is *holomorphic* if for all $z \in G$, either

- $f(z) \in \mathbb{C}$ (so it doesn't hit ∞) and $f : G \to \mathbb{C}$ is holomorphic, or
- if $f(z) \in \widehat{\mathbb{C}} \setminus \{0\}$, then $1/f(w) : G \to \mathbb{C}$ is holomorphic, where $1/\infty$ is understood to be 0.

If the image of f contains neither 0 not ∞, then both criteria hold, and are equivalent (since $1/z$ is holomorphic on any neighborhood not containing zero).

Proposition 3.3. *The meromorphic functions on G can be identified with the holomorphic functions $G \to S^2$.*

Proof. Suppose f is meromorphic on G, so that it has a pole of order m at z_0. Then, $f(z) = (1/(z - z_0)^m)g(z)$ for some holomorphic g with a removable singularity at z_0, and $g(z_0) \neq 0$.

By letting $1/0 = \infty$, this realizes f as a continuous map $G \to S^2$, and $1/f = (z - z_0)^m(1/g)$, which is certainly holomorphic near z_0, so f is holomorphic as a map to S^2.

The converse is quite similar, a matter of unwinding the definitions, but has been left as an exercise. ⊠

You can also define a notion of a holomorphic function coming out of S^2, not just into.

Definition. Let $G \subset S^2$ be open. A continuous $f : G \to S^2$ is *holomorphic* if one of the following is true.

- If $\infty \notin G$, then we use the same definition as above.
- If $\infty \in G$, then it's holomorphic on $G \setminus \infty$ and there's a neighborhood N of ∞ in G such that the composition
$$N^{-1} \xrightarrow{z \mapsto 1/z} N \xrightarrow{f} S^2$$
is holomorphic.

If you're used to working with manifolds, this sort of coordinate change is likely very familiar: every time we talk about ∞, we take reciprocals and talk about 0.

Example 3.4. Every rational function $p \in \mathbb{C}(z)$ is meromorphic, and extends to a holomorphic map $S^2 \to S^2$.

Now, we can talk about these geometrically: $z \mapsto z^2$ sends $e^{in\theta} \mapsto e^{2in\theta}$, so it doubles the longitude (modulo 1). In particular, it wraps the sphere twice around itself, preserving 0 and ∞, as in Figure 3.

Example 3.5. A *Möbius map* is a map
$$\mu(z) = \frac{az + b}{cz + d},$$
where $a, b, c, d \in \mathbb{R}$ and $ad - bc = 1$. This extends to a holomorphic map $S^2 \to S^2$ with a holomorphic inverse (the Möbius map associated to $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]^{-1}$). Thus, there's a homeomorphism $\mathrm{SL}_2(\mathbb{R})/\{\pm I\}$ to the group of Möbius transformations.

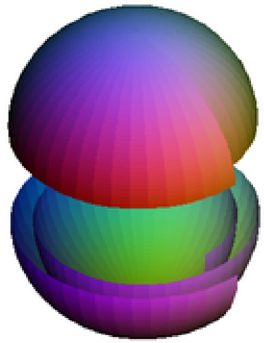


FIGURE 3. A depiction of the map $z \mapsto z^2$ on the Riemann sphere, which fixes the poles. Source: https://en.wikipedia.org/wiki/Degree_of_a_continuous_mapping.

One interesting corollary is that the point at infinity is *not* special, since there's a Möbius map sending it to any other point of S^2, and indeed they act transitively on it. So we don't really have to distinguish the point at infinity from this geometric point of view.

Theorem 3.6. *If $f : S^2 \to S^2$ is holomorphic, then it's a rational function. In particular, the Möbius maps are the only invertible holomorphic maps $S^2 \to S^2$.*

The idea is to eliminate the zeros and poles by multiplying by $(z - z_0)^m$; then, one can apply Liouville's theorem to show that the result is constant.

Lecture 4.

Analytic Continuation: 1/27/16

This corresponds to §1.1 in the textbook, and is one of the classical motivations for Riemann surfaces.

The problem is: if $G \subset \mathbb{C}$ is open and $f \in \mathscr{O}(G)$, then we would like to extend f holomorphically, or maybe meromorphically, to a larger domain $H \supset G$. Such extensions are called *analytic* (resp. *meromorphic*) *continuations* of f.[3]

Remark. If H is connected, then there exists at most one meromorphic continuation of f to H, because the difference of two continuations vanishes on the open set G, and hence vanishes everywhere.

Example 4.1. Let $f(z) = \sum_{n \geq 0} z^n$, which converges on the open unit disc, but diverges when $|z| \geq 1$. At first sight, this suggests we'll never get any farther than the disc, but this turns out to merely be an artifact of this presentation of f: we could instead write it as $f(z) = 1/(1 - z)$, which meromorphically extends f to the whole of $\mathbb{C}$ (with a single pole at $z = 1$). Thus, this power series representation is not *per se* intrinsic.

One can take this further and define analytic continuations of general functions defined by power series.

Example 4.2. This example is more sophisticated, and will take longer; it reflects a common theme in this subject, that the examples are nontrivial and are worth taking seriously. Define the Γ-*function*

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, \mathrm{d}t$$

on the open set $\operatorname{Re} z > 0$. This integral is doubly improper, since there's a singularity at 0 and it's unbounded on the right, so we really should rewrite it as

$$\Gamma(z) = \lim_{\varepsilon \to 0^+} \int_\varepsilon^1 t^{z-1} e^{-t} \, \mathrm{d}t + \lim_{T \to \infty} \int_1^T t^{z-1} e^{-t} \, \mathrm{d}t.$$

Let $H_a = \{z \mid \operatorname{Re} z > a\}$. We're going to show that Γ extends to the entire plane, but first we need to show that it's holomorphic on the right half-plane.

Proposition 4.3. $\Gamma \in \mathscr{O}(H_0)$.

[3]Though "holomorphic continuation" would make more sense, tradition gives us the term "analytic continuation."

Proof sketch. Since we need to realize $\Gamma(z)$ as a limit, let

$$g_n(z) = \int_{1/n}^{n} t^{z-1}e^{-t}\,\mathrm{d}t.$$

This is an integral of a holomorphic function, so $g_n \in \mathscr{O}(\mathbb{C})$ and

$$g_n'(z) = \int_{1/n}^{n} \frac{\partial}{\partial z}\bigl(t^{z-1}e^{-t}\bigr)\,\mathrm{d}t = (z-1)g_n(z-1).$$

If $a > 0$, then g converges uniformly on the strip $a < \operatorname{Re} z < b$ — the goal is to show that g_n is uniformly Cauchy on this strip (the details of which are left to the reader) by comparing to the integral of $e^{-t/2}$ for $t \gg 0$, the point being that $e^{x-1}e^{-t} \le e^{-t/2}$ for t sufficiently large. For $t < 1$, one should compare it to the integral of t^{x-1}. Then, we need to use the following theorem.

Theorem 4.4. *If $f_n \in \mathscr{O}(G)$ and $f_n(z) \to f(z)$ locally uniformly, then $f \in \mathscr{O}(G)$.*

The proof uses Morera's theorem (Theorem 3.1) and can be found in the review notes (or Rudin, etc.). In any case, this means $\Gamma = \lim_{n\to\infty} g_n$ is holomorphic on the right half-plane. ⊠

Now, we can talk about extending Γ.

Theorem 4.5. *Γ has a meromorphic continuation to $\mathbb{C}$, whose only poles are simple poles*[4] *at 0, -1, -2, and so on.*

Proof. Since the gamma function is given by an integral, let Γ_0 be that integral from 0 to 1, and Γ_∞ be the integral from 1 to ∞. Then, the argument above shows that $\Gamma_\infty \in \mathscr{O}(\mathbb{C})$, so the only extension that we actually need to make is of

$$\Gamma_0(z) = \int_0^1 t^{z-1}e^{-t}\,\mathrm{d}t.$$

The cunning idea is that we're going to look at the n^{th}-order Taylor polynomial for e^{-t}, which provides an integral we can actually do, and then treat everything else separately. Specifically, let

$$e_n(t) = \sum_{j=0}^{n-1} \frac{(-t)^j}{j!},$$

so that

$$\begin{aligned}\Gamma_0(z) &= \underbrace{\int_0^1 t^{z-1}\bigl(e^{-t} - e_n(t)\bigr)\,\mathrm{d}t}_{\Gamma_n(z)} + \int_0^1 t^{z-1}e_n(t)\,\mathrm{d}t.\\ &= \Gamma_n(z) + \sum_{j=0}^{n-1} \frac{(-1)^j}{j!(z+j)}.\end{aligned}$$

The $(z+j)$ in the denominator on the right gives us simple poles at $0, -1, -2, \ldots, -n+1$. But $e^{-t} - e_n(t)$ has a zero of order n at $t = 0$, so

$$\int_0^1 t^{z-1}\bigl(e^{-t} - e_n(t)\bigr)\,\mathrm{d}t$$

exists on H_{-n}, so $\Gamma_n \in \mathscr{O}(H_{-n})$. Thus, we can extend Γ meromorphically to all of $\mathbb{C}$, because any $z \in \mathbb{C}$ is in some H_{-n}, so we can work this with Γ_n. ⊠

It goes without saying that Γ is one of the most prominent functions in analytic number theory.

These two successful examples of meromorphic continuation are in some sense atypical; in general, there is a problem of multi-valuedness or monodromy.

[4] A pole is *simple* if it's degree 1.

Example 4.6. For an algebraic example of this problem, consider

$$f(z) = \sum_{n \geq 0} \binom{1/2}{n} z^n,$$

where

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}.$$

By a generalized binomial theorem (or checking that it satisfies the right differential equation), one can show that f converges on $D(0,1)$ to a branch of $\sqrt{1+z}$. We can extend holomorphically to the *cut plane* $\mathbb{C} \setminus (-\infty, -1]$ by writing $f(z) = \exp((1/2)\log(1+z))$, where we can choose a branch of $\log(1+z)$ in this cut plane, such as $\log(re^{i\theta}) = \log r + i\theta$, with $\theta \in (-\pi, \pi)$.

There's nothing particularly special about this branch cut. Plenty of other branch cuts (paths from -1 to $-\infty$ whose complements are simply connected) work just as fine — but we cannot extend further, because as we go around a loop around -1, $f(z)$ flips $-f(z)$ (the other branch of $\sqrt{1+z}$), since the logarithm changes by $2\pi i$. This is a little unsatisfactory, since we can't go further.

A similar story holds for just about any algebraic function, since one has to take a branch cut to resolve the ambiguity of multiple roots.

The Riemann surfaces way to approach this is instead of making arbitrary branch cuts, it's more canonical instead to study the equation $w^2 - (1-z) = 0$, which implicitly defines w as a square root of $1+z$. Then, we consider the set

$$X = \{(z,w) \in \mathbb{C}^2 \mid P(z,w) = 0\},$$

where $P(z,w) = w^2 - (1+z)$. Soon, we will see that this X is a Riemann surface. We can play exactly the same game with any $P(z,w) : \mathbb{C}^2 \to \mathbb{C}$ that is holomorphic in each variable separately, includin any polynomial in z and w. This defines for us its zero set $X = \{P(z,w) = 0\}$.

Then, we have an implicit function theorem, which is a major classical motivation for the theory of Riemann surfaces, just as the implicit function theorem on $\mathbb{R}^n$ is a major classical motivation for defining abstract manifolds.

Theorem 4.7 (Implicit function theorem)**.** *If $(z_0, w_0) \in X$ and $\frac{\partial P}{\partial w}(z_0, w_0) \neq 0$, then there's a disc $D_1 \subset \mathbb{C}$ centered at $z_0 \in \mathbb{C}$ and a disc $D_2 \subset \mathbb{C}$ centered at w_0, and a holomorphic $\phi : D_1 \to D_2$ such that $\phi(z_0) = w_0$ and $X \cap (D_1 \times D_2)$ is the graph of ϕ, i.e. $\{(z, \phi(z)) \mid z \in D_1\}$.*

An analogue of this function holds for C^1 real functions (or C^∞ ones), and this version can be extracted from that, but it has a simpler, direct proof.

Proof. This proof hinges on a theorem called the *argument principle*, that if $f \in \mathscr{O}(G)$ and $\overline{D}$ is a closed disc in G with $f(z) \neq 0$ on $\partial\overline{D}$, then

$$\frac{1}{2\pi i} \int_{\partial \overline{D}} \frac{f'(w)}{f(w)} \, \mathrm{d}w = \sum_{\substack{z \in D \\ f(z)=0}} \mathrm{mult}(f; z). \tag{4.8}$$

That is, integrating the logarithmic derivative counts the zeros inside D, with multiplicity. There's also the related formula

$$\frac{1}{2\pi i} \int_{\partial \overline{D}} \frac{w f'(w)}{f(w)} \, \mathrm{d}w = \sum_{\substack{z \in D \\ f(z)=0}} z \, \mathrm{mult}(f; z). \tag{4.9}$$

These are nice exercises in residue calculus.

Returning to the implicit function theorem, let $f_z = P(z, \cdot)$, so $f_{z_0}(w_0) = 0$, but $f'_{z_0}(w_0) \neq 0$. Thus, $\mathrm{mult}(f_{z_0}; w_0) = 1$, and therefore by isolation of zeros, there's a disc D_2 centered at w_0 such that w_0 is the only zero of f_{z_0} in $\overline{D}_2$. Hence, by (4.8),

$$\frac{1}{2\pi i} \int_{\partial D_2} \frac{f'_{z_0}}{f_{z_0}} = 1.$$

Since $f_{z_0} \neq 0$ on the boundary, then there's a $\delta > 0$ such that $|f_{z_0}| > 2\delta > 0$ on ∂D_2. Thus, there's a disc D_1 centered at z_0 such that for all $z \in D_1$, $|f_z| > \delta$ on ∂D_2 because P is continuous. Hence,

$$\frac{1}{2\pi i}\int_{\partial D_2} \frac{f_z'}{f_z} = 1,$$

or, by (4.8), there's a unique solution $w = \phi(z)$ to $P(z,w) = 0$ with $z \in D_1$ and $w \in D_2$. Thus, we need only to show that ϕ is holomorphic. By (4.9),

$$\phi(z) = \frac{1}{2\pi i}\int_{\partial D_2} \frac{w f_z'(w)}{f_z(w)}\,\mathrm{d}w = \frac{1}{2\pi i}\int_{\partial D_2} \frac{w}{P(z,w)}\frac{\partial P}{\partial w}(z,w)\,\mathrm{d}w.$$

Hence, ϕ is holomorphic in z (since its derivative is given by differentiating under the integral sign). ⊠

Thus, even working just with zero sets of algebraic functions, Riemann surfaces show up very nicely.

Lecture 5.

Analytic Continuation Along Paths: 1/29/16

Today, we're going to talk about analytic continuation along paths and the interesting things that result. There's also a more classical Weierstrass way to look at this.

Definition. If ϕ is a holomorphic function defined on a neighborhood U of a $z_0 \in \mathbb{C}$ and $\gamma : [0,1] \to \mathbb{C}$ is a path with $\gamma(0) = z_0$, then an *analytic continuation* of ϕ along γ consists of a pair (U_t, ϕ_t) for all $t \in [0,1]$, where U_t is a neighborhood of $\gamma(t)$ and $\phi_t \in \mathscr{O}(U_t)$ such that:

- $\phi_0 = \phi$ on $U_0 \cap U$, and
- the different ϕ_t should agree, in the sense that for all $s \in [0,1]$, there's a $\delta > 0$ such that if $|t-s| < \delta$, then ϕ_s and ϕ_t agree on $U_s \cap U_t$.

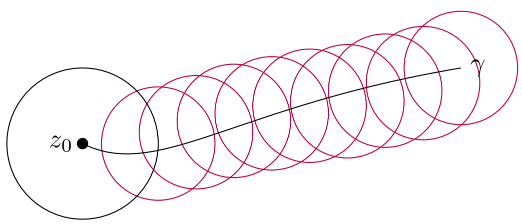


FIGURE 4. Analytic continuation along a path; on sufficiently close circles, the extensions must agree.

Note, however, that if γ intersects itself, then there's no requirement for the extensions to agree on those overlaps (if δ is sufficiently small, for example). Weierstrass said this is how one should think of complex analytic functions, and this confused a lot of people, but did lead to Weyl's work that we'll discuss in a few lectures.

Example 5.1. The logarithm is a very good example. Start with a branch of log defined on some open set U_0, so $\log(re^{i\theta}) = \log r + i\theta$, or $\log z = \log|z| + i\arg z$, for some continuous, real-valued $\arg : U_0 \to \mathbb{R}/2\pi\mathbb{Z}$.

Then, for any $\gamma : [0,1] \to \mathbb{C}^*$ with $\gamma(0) = z_0 \in U_0$, we can uniqly lift $\arg\circ\gamma : [0,1] \to \mathbb{R}/2\pi\mathbb{Z}$ to $\mathbb{R}$ consistent with $\arg(z_0)$; this lift will be called $\arg_\gamma$.[5] Then, define $\log_{\gamma_t}(z) = \log|z| + \arg_{\gamma(t)}(z)$, which defines a continuation of the logarithm around γ.

Example 5.2. For a more algebraic example, let

$$\phi(z) = \sum_{j\geq 0}\binom{1/2}{j} z^j$$

on the unit disc $D(0,1)$, so $\phi(z)^2 = z+1$. Then, one can continue along any γ with image in $\mathbb{C}\setminus\{-1\}$ by setting $\phi_t(z) = \exp((1/2)\log_{\gamma_t}(1+z))$. However, if $\gamma(t) = -1 + e^{2\pi i t}$, then γ winds around -1, and when it returns to a point, the extension of ϕ has a different value!

[5]One can think of this in terms of the theory of covering spaces, which is one reason this function lifts.

Example 5.3. Analytic continuation along paths works particularly well with differential equations: let p and q be meromorphic functions. Then, we want to find a $u(z)$ such that $u'' + p(z)u' + q(z)u = 0$, which we'll call $\boxed{p, q}$.[6] If you think differential equations are boring, questions like these are still motivated by study of $\mathcal{D}$-modules and the like in algebraic geometry.

Let's work near a point z_0 where p and q are holomorphic, so z_0 is a *regular point*, and without loss of generality make $z_0 = 0$. We're going to look for *series solutions*: set $p(z) = \sum_{n\geq 0} p_n z^n$ and $q(z) = \sum q_n z^n$ on $D(0, R)$ for some R, and we want to find $u(z) = \sum_{n\geq 0} u_n z^n$. Equating the coefficients of z_n in $\boxed{p, q}$, one obtains the recurrence relation

$$(n+1)(n+2)u_{n+1} + \sum_{i=0}^{n}(n+i-1)p_i u_{n+1-i} + \sum_{j=0}^{n} q_j u_{n-j} = 0.$$

By induction, one shows that all of the u_j are determined by a choice of $(u_0, u_1) \in \mathbb{C}^2$.

Proposition 5.4. *$\sum u_n z^n$ converges in the same disc $D(0, R)$.*

The detailed proof is a homework assignment, and depends on the following lemma, due to an idea of Cauchy.

Lemma 5.5 (Majorization)**.** *Say $|p_n| \leq P_n$ and $|q_n| \leq Q_n$. Then, let $P(z) = \sum P_n z^n$ and $Q(z) = \sum Q_n z^n$. If $u = \sum u_n z^n$ is a solution to $\boxed{p, q}$ and $U_n = \sum U_n z^n$ is a solution to $\boxed{P, Q}$, and if $U_0 = |u_0|$ and $U_1 = |u_1|$, then $|u_n| \leq |U_n|$.*

The proof involves some straightforward estimates after the recurrence formula.

Proof sketch of Proposition 5.4. Let's work on $\overline{D(0, r)}$ where $r < R$. Then, we have estimates like $|p_n| \leq M/r^n$ and $|q_n| \leq M/r^n$, where $M = \sup_{z\in\overline{D(0,r)}}\{|p(z)|, |q(z)|\}$, which follows from Cauchy's estimates (which themselves are corollaries of the Cauchy integral formula, Proposition 2.3).

Now, using the majorization lemma, we can compare $\boxed{p, q}$ to

$$\boxed{\sum_{n\geq 0}|p_n|z^n, \sum_{n\geq 0}|q_n|z^n} \qquad \text{and} \qquad \boxed{\sum \frac{M}{r^n}z^n, \sum \frac{M}{r^n}z^n}.$$

It makes sense to compare this to $\boxed{M/(1 - z/r), M/(1 - z/r)^2}$, i.e. the equation

$$u'' + \frac{Mu'}{1 - z/r} + \frac{Mu}{(1 - z/r)^2} = 0.$$

This last equation has an explicit solution $\mu/(1 - z/r)$ for some μ, and its Taylor series converges on $D(0, r)$; now, using the majorization lemma, the coefficients of our original series are smaller, and therefore it converges. ⊠

Thus, we have a 2-dimensional $\mathbb{C}$-vector space V of solutions near z_0. The tie-in to the rest of lecture is the following proposition/exercise.

Exercise. Show that if $p, q \in \mathscr{O}(G)$ and $\gamma : [0, 1] \to G$, then any solution to $\boxed{p, q}$ has a solution along γ through solutions to $\boxed{p, q}$.

Monodromy. If γ is now a loop in G, so $\gamma(0) = \gamma(1) = z_0$, then analytic continuation around γ defines a linear map $M_\gamma : V \to V$ called the *monodromy map*: you go around and end up not where you started, and it's easy to see that this dependence is linear.

Exercise. M_γ depends only on the homotopy class of γ (relative to basepoints).

Thus, this is only interesting if G isn't simply connected, so in general we get interesting examples of monodromy by going around poles of p and q. In particular, there's the oxymoronic-sounding notion of regular singular points. The prototype is the following, simpler equation:

$$u'' + \frac{A}{z}u' + \frac{B}{z^2}u = 0, \tag{5.6}$$

[6]"If you're typing notes, feel free to call it something else, like $L_{p.q}$."

where $A, B \in \mathbb{C}$ are just constants. We seek solutions of the form $u(z) = z^\alpha$, where $\alpha \in \mathbb{C}$; this is defined initially near 1, and then analytically continued along paths in $\mathbb{C}^*$. If you write down the left-hand side, you end up getting

$$u'' + \frac{A}{z}u' + \frac{B}{z^2}u = \underbrace{(\alpha(\alpha-1) + A\alpha + B)}_{I(\alpha)} z^{\alpha-2}.$$

In other words, to get a solution, we need $I(\alpha) = 0$; this is called the *indicial equation*. Since it's a quadratic, then there's one or two roots: if the roots α_1 and α_2 are distinct, then $(z^{\alpha_1}, z^{\alpha_2})$ is a basis for V (the solutions near 1), and if γ is the unit circle, then the monodromy matrix in this basis is

$$M_\gamma = \begin{bmatrix} e^{2\pi i\alpha_1} & 0 \\ 0 & e^{2\pi i\alpha_2} \end{bmatrix}. \tag{5.7}$$

If α is a related root, the basis we get is $(z^\alpha, z^\alpha \log z)$, and the monodromy matrix is a nontrivial Jordan block:

$$M_\gamma = e^{2\pi i\alpha} \begin{bmatrix} 1 & 2\pi i \\ 0 & 1 \end{bmatrix}.$$

One takeaway is that even an equation as simple as (5.6) has monodromy.

This generalizes quite naturally.

Definition. A $z_0 \in \mathbb{C}$ is a *regular singular point* of $u'' + pu' + qu = 0$ if p has a pole of order at most 1 and q has a pole of order at most 2 at z_0.

One seeks solutions via the *Frobenius method*: since p has a simple pole and q has a double pole, then there are $\widetilde{p}, \widetilde{q}$ holomorphic in a neighborhood of 0 such that $p(z) = A/z + \widetilde{p}(z)$ and $q(z) = B/z^2 + C/z + \widetilde{q}(z)$. Thus, the indicial equation is $\alpha(\alpha-1) + A\alpha + B = 0$.

Proposition 5.8. *If there are indicial roots α_1 and α_2 such that $\alpha_1 - \alpha_2 \notin \mathbb{Z}$, then there are solutions $u_1, u_2 \in V$ such that $u_1 = z^{\alpha_1}w_1$ and $u_2 = z^{\alpha_2}w_2$, and the monodromy matrix is as in* (5.7).

Lecture 6.

Riemann Surfaces and Holomorphic Maps: 2/3/16

Today, we'll begin with section 3.1 of the book, defining Riemann surfaces properly. This may be very routine to you or far from it; in any case, the notion of a manifold is central to mathematics, and now's as good a time as any to see it.

Definition. A *Riemann surface* (abbreviated R.S.) is the data of

- a Hausdorff topological space X, along with
- an *atlas*; that is, a collection $\{(U_\alpha, \phi_\alpha) \mid \alpha \in A\}$ where the $U_\alpha \subset X$ are open, $\bigcup_{\alpha\in A} U_\alpha = X$, and $\phi_\alpha : U_\alpha \to \mathbb{C}$ is a homeomorphism onto its image;[7] we require that for all $\alpha, \beta \in A$, the transition map $\tau_{\alpha\beta} = \phi_\beta \circ \phi_\alpha^{-1} : \phi_\alpha(U_\alpha \cap U_\beta) \to \mathbb{C}$ is holomorphic.

We deem two atlases on X to define the same surface if their union is also an atlas satisfying the above conditions.

The rest of this week will be devoted to examples of Riemann surfaces, and unwinding this definition.

Remark.

- If $p \in U_\alpha$, we can think of ϕ_α as defining a holomorphic coordinate z neat p on X. The definition forces us to work only with notions that are independent of the particular coordinate we chose; for example, it does make sense to ask for a holomorphic function's order of vanishing at p.
- There are many variants of this definition of a Riemann surface given by replacing holomorphicity with something else. If one instead requires the maps to be smooth, the resulting definition is for a *smooth surface*; if we require smoothness with positive Jacobian, it's a *smooth oriented surface*; and many more.

[7] Each pair (U_α, ϕ_α) is called a *chart*.

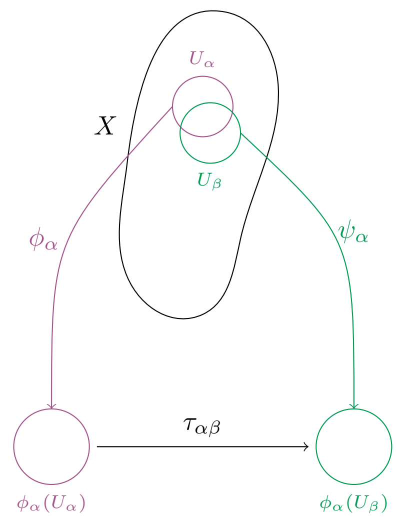


FIGURE 5. A transition map between charts for a Riemann surface X.

- A $\mathbb{C}$-linear map $\mathbb{C} \to \mathbb{C}$ has non-negative Jacobian, because the map $\mathbb{C} \to \mathbb{C}$ sending $z \mapsto az$ acts on $\mathbb{R}^2$ by $\left[\begin{smallmatrix} \operatorname{Re} a & -\operatorname{Im} a \\ \operatorname{Im} a & \operatorname{Re} a \end{smallmatrix}\right]$, which has det $|a|^2$. In particular, a Riemann surface is also a smooth oriented surface. Thus, by the classification of compact, smooth, oriented surfaces, any connected, compact Riemann surface is equivalent to a standard genus-g surface.
- In the conventional definition of smooth surfaces, one generally assumes that a smooth surface has a countable atlas; equivalently, one may take the space to be paracompact or second-countable. There are tricky counterexamples if you don't include this (e.g. they do not admit partitions of unity, and hence Riemannian metrics). However, this isn't necessary in the world of Riemann surfaces.

Theorem 6.1 (Radó)**.** *Any connected Riemann surface has a countable holomorphic atlas.*

Thus, unlike in differential geometry, where we care only about nicer surfaces, here we get that our surfaces are nice already.[8]

Now, we want to know not just what these are, but also how to map between them.

Definition. Let $(X, \{(U_\alpha, \phi_\alpha)\})$ and $(Y, \{(V_\beta, \psi_\beta)\})$ be Riemann surfaces. Then, a *holomorphic map* $f : X \to Y$ is a continuous map such that for all charts ϕ_α and ψ_β, $\psi_\beta \circ f \circ \phi_\beta^{-1}$ is holomorphic. An invertible holomorphic map is called *biholomorphic*.

Courtesy of the implicit function theorem for holomorphic functions, Theorem 4.7, the inverse of a biholomorphic function is also holomorphic.

Example 6.2.

(1) Any open set $\Omega \subset \mathbb{C}$ is a Riemann surface with just one chart $\phi : \Omega \to \mathbb{C}$ given by inclusion (the only translation functions are the identity, which is holomorphic).
(2) The Riemann sphere $S^2 = \widehat{\mathbb{C}}$ is a Riemann surface with an atlas of two charts: the copy of $\mathbb{C}$ inside $\widehat{\mathbb{C}}$ is sent to $\mathbb{C}$ by the identity, and $\mathbb{C}^* \cup \{\infty\}$ is sent to $\mathbb{C}$ by $z \mapsto 1/z$; the transition map is $z \mapsto 1/z$ on $\mathbb{C}^*$, which is holomorphic. The Möbius maps $\mu : S^2 \to S^2$ given by $\mu(z) = (az + b)/(cz + d)$, where $ad - bc = 1$, are biholomorphic.
(3) $\mathbb{D}$ will denote the *unit disc* $D(0, 1)$, and $\mathbb{H} = \{z \in \mathbb{C} \mid \operatorname{Im}(z) > 0\}$, the *upper half-plane*. The Möbius function $\mu : \mathbb{H} \to \mathbb{D}$ sending $\mu(z) = (z - i)/(z + i)$ is biholomorphic (since it's the restriction of a Möbius map on S^2), and it sends $0 \mapsto -1$, $\infty \mapsto 1$, and $1 \mapsto i$, so $\mu(\mathbb{R} \cup \infty)$ is the unit circle., Then, $\mu(i) = 0$, so $\mu(\mathbb{H}) = \mathbb{D}$ (it has to be either the inside or the outside of $\mathbb{D}$, since μ is continuous).

[8]Later in the class, we'll prove the uniformization theorem, which says that every connected Riemann surface is equivalent to a quotient of $\mathbb{C}$, the sphere, or the hyperbolic plane by a group action. This implies Radó's theorem, but is now how Radó originally proved it.

(4) Let $P(z,w)$ be holomorphic in both $z \in \mathbb{C}$ and $w \in \mathbb{C}$. Then, $X = \{(z,w) \in \mathbb{C}^2 \mid P(z,w) = 0\}$; if we know that there's no $(z,w) \in X$ where both partial derivatives of P vanish, then X is naturally a Riemann surface.

We proved that if $\frac{\partial P}{\partial w}(w_0, w_0) \neq 0$, then there are discs $D_1 \subset \mathbb{C}$ and $D_2 \subset \mathbb{C}$ centered at z_0 and w_0, respectively, and a holomorphic $\phi : D_1 \to D_2$ with $\phi(z_0) = w_0$ and $U = X \cap (D_1 \times D_2) = \mathrm{graph}(\phi)$. We can use this to get a chart $\psi : U \to D_1$ given by projection onto the first factor. Alternatively, if $\frac{\partial P}{\partial w}$ vanishes at (z_0, w_0), then $\frac{\partial P}{\partial z}$ doesn't, so we can do the same thing, but with $\phi : D_2 \to D_1$. Then, our map is projection onto the second coordinate.

Now, let's look at the change-of-charts maps. If both charts have the same type, the transition map is just the identity on some open set in $\mathbb{C}$, so let's look at what happens on a chart map between the two types, where X is locally the graph of an $f : D_1 \to D_2$. Then, ϕ^{-1} sends $z \mapsto (z, f(z))$ and ψ sends $(z,w) \mapsto w$, so the transition map is $z \mapsto f(z)$, which by construction was holomorphic; hence, X is a Riemann surface.

Lecture 7.

More Sources of Riemann Surfaces: 2/5/16

Today we'll talk about three more sources of Riemann surfaces, with more to come on Monday.

Covering Spaces. The first, easiest example is covering spaces. Recall that a continuous map $\pi : Y \to X$ is a *covering map* if X is the union of open sets U such that $\pi^{-1}(U) \to U$ is equivalent to $U \times D \to U$, where D has the discrete topology.

If X is a Riemann surface, then Y acquires a Riemann surface structure that makes π holomorphic. The idea is that the charts $X \to \mathbb{C}$ lift to several disjoint copies in Y. Each of these maps homeomorphically onto the chart, and then composing with the chart map gives a chart structure on Y. There's something to be fleshed out here, but it's straightforward; in fact, the requirement that π is holomorphic pretty much forces one's hand.

Suppose X is path-connected, with a basepoint x_0. Then, we can construct a *universal cover* $\pi : \widetilde{X} \to X$, with $\widetilde{X}$ simply connected. We'll see this again, so it's useful to remember the construction: $\widetilde{X} = \{(x,\gamma) \mid x \in X, \gamma : [0,1] \to X, \gamma(0) = x_0, \gamma(x) = x\}$ modulo the equivalence relation $(x,\gamma) \sim (x',\gamma')$ if $x = x'$ and γ and γ' are homotopic. The topology on $\widetilde{X}$ is chosen to make it a covering map.

The fundamental group of X, $\pi_1(X, x_0)$, acts on X by *deck transformations*, maps $g : \widetilde{X} \to \widetilde{X}$ that commute with the projection to X; if X and Y are Riemann surfaces, the deck transformations are biholomorphic. Moreover, any connected and path-connected covering space of X takes the form $\widetilde{X}/G$, where $G \leq \pi_1(X, x_0)$.

In summary, there's nothing new caused by making X and Y Riemann surfaces; the whole theory maps nicely into the category of Riemann surfaces and holomorphic maps.

The Riemann Surface of a Holomorphic Function. The idea is that we'll construct a "maximal analytic continuation" of a prescribed holomorphic function. The domain will be a Riemann surface, and not always an open set in the plane. It's the realization of Weierstrass' idea of considering all possible branches of a holomorphic function.

The input data will be an open $U \subset \mathbb{C}$, a $z_0 \in U$, and an $f \in \mathscr{O}(U)$. Then, an *abstract analytic continuation* (AAC) of f is $\mathscr{X} = (X, x_0, \pi, F)$, where X is a Riemann surface, $x_0 \in X$, $\pi : X \to \mathbb{C}$ sends $x_0 \mapsto z_0$ and is a *local homeomorphism* (meaning $\pi'(z)$ never vanishes). We require that if σ is a local right inverse to π near z_0 (so $\pi \circ \sigma = \mathrm{id}$), then we require that $F \circ \sigma = f$ in a neighborhood of z_0.[9]

There's a natural notion of a morphism between two abstract analytic continuations $\mathscr{X}$ and $\mathscr{X}'$ of (U, f) (respectively given by (X, x_0, π, F) and (X', x_0', π', F')): a holomorphic map $\phi : X \to X'$ respecting all the structure, i.e. it intertwines π and π', as well as F and F'. In particular, the AACs are a category C_f.

Definition. A *terminal object* in a category C is an $X \in \mathsf{C}$ such that any $X' \in \mathsf{C}$ maps to X in a unique way.

Proposition 7.1. C_f *has a terminal object* $\mathscr{X}_f = (X_f, x_0, \pi, F)$.

[9] Up to making the neighborhood smaller, σ is unique anyways; thus, it's unique as the germ of a function.

You should think of this as a sort of maximal object. More concretely, a path $\gamma : [0,1] \to \mathbb{C}$ with $\gamma(0) = x_0$ lifts to a $\widetilde{\gamma} : [0,1] \to X$ (meaning $\pi \circ \widetilde{\gamma} = \gamma$) iff f has an analytic continuation along the path γ.

$\mathscr{X}_f$ is called the *Riemann surface of the function* f. It can be given a more concrete construction: the set of paths $\gamma : [0,1] \to \mathbb{C}$ with $\gamma(0) = z_0$ and f admits an analytic continuation f_γ along γ, modulo an equivalence relation, where $\gamma \sim \gamma'$ if $\gamma(1) = \gamma'(1)$ and $f_\gamma(\gamma(1)) = f_{\gamma'}(\gamma'(1))$. That is, since there's at most one analytic continuation along any path (up to some equivalence about the size of the neighborhoods, which is irrelevant), we identify the "same" analytic continuations: in particular, homotopic paths are identified. Thus, X_f is a quotient of the universal cover of an open set in $\mathbb{C}$, so it is itself a cover: we have a covering map $\pi : X_f \to \mathbb{C}$ sending $[\gamma] \to \gamma(1)$. The basepoint $x_0 \in X_f$ is the class of the constant path at z_0, and the map $F : X_f \to \mathbb{C}$ sends $\gamma \mapsto f_\gamma(\gamma(1))$.[10]

This seems a little abstract, but working through it is probably helpful. As an example, though, suppose f is a branch of $\sqrt{z}$ on some open $U \subset \mathbb{C}$. We know that $X = \{w^2 - z = 0\} \subset \mathbb{C}^2$ is a Riemann surface (some partial derivative-checking should be done here); then, X_f will be the subset of X where $z \neq 0$. Then, $X_f \twoheadrightarrow \mathbb{C}^*$ by $(z,w) \mapsto z$, which is a double cover.

Historically, this is one of the important examples for constructing Riemann surfaces, though "terminal object in a category" isn't the language one would have heard/

Plane Projective Algebraic Curves. This is also an extremely important class of examples.

Recall that *complex projective space*, $\mathbb{C}P^n$, is the set of one-dimensional (complex) vector subspaces in $\mathbb{C}^{n+1}$. These are given by points in $\mathbb{C}^{n+1}$ modulo the action of $\mathbb{C}^*$ acting by scaling (which doesn't change the line through a point). That is, $\mathbb{C}P^n = (\mathbb{C}^{n+1} \setminus 0)/\mathbb{C}^*$, and carries the quotient topology, making it a topological space, and it's compact (since it can also be realized as the quotient topology on $S^{2n+1}/\mathrm{U}(1)$, by scaling each vector in $\mathbb{C}^{n+1} \setminus 0$ to a unit vector).

Points in $\mathbb{C}P^n$ are usually written in *homogeneous coordinates* $[z_0 : z_1 : \cdots : z_n]$, which represents the equivalence class (modulo scaling) of $(z_0, \dots, z_n) \in \mathbb{C}^{n+1} \setminus 0$.

We can write $\mathbb{C}P^n = U_0 \cup U_1 \cup \cdots \cup U_n$, where U_j is the set of classes of points where $z_j \neq 0$, so after rescaling, $[z_0 : \cdots : z_{j-1} : 1 : z_{j+1} : \cdots : z_n]$. Thus, looking at all the other coordinates, it's identified with $\mathbb{C}^n$, and this places a (complex) manifold structure on $\mathbb{C}P^n$.

One can pass back and forth between polynomials $p \in \mathbb{C}[z_1, z_2]$ and homogeneous polynomials $P(z_0, z_1, z_2)$ in three variables: if

$$p(z_1, z_2) = \sum_{i,j} a_{ij} z_1^i z_2^j,$$

and d is the largest degree $(i+j))$ in p, then it corresponds to

$$P(Z_0, Z_1, Z_2) = \sum_{i,j} a_{ij} Z_0^{d-i-j} Z_1^i Z_2^j,$$

which is homogeneous of degree d. For example $z_1^2 + z_2^3 + 1$ is homogenized to $Z_0 Z_1^2 - Z_2^3 + Z_0^3$. Thus, we can think of $X = \{(z_1, z_2) \mid p(z_1, z_2) = 0\} \subset \mathbb{C}^2$ as a subset in $U_0 \subset \mathbb{C}^2$; then, the closure of X in $\mathbb{C}P^2$ is the compact space $\overline{X} = \{(Z_0, Z_1, Z_2) \mid P(Z_0, Z_1, Z_2) = 0\}$. Next time, we'll prove the following proposition.

Proposition 7.2. *Suppose that for all $q \in \overline{X}$, $\frac{\partial P}{\partial Z_j}$ is nonvanishing for some j. Then, the Riemann structure on $X \subset U_0$ extends to a Riemann surface structure on $\overline{X} \subset \mathbb{C}P^2$.*

Lecture 8.

Projective Surfaces and Quotients: 2/8/16

Today, we'll discuss two fundamental and important examples of Riemann surfaces: plane projective curves and quotients.

[10] In fact, another way to define X_f is as a quotient of the universal cover, subject to some conditions, but then it's less apparent that it's unique.

Plane Projective Curves. We discussed this a little bit last time, but we have a correspondence between polynomials $p(z_1, z_2) \in \mathbb{C}[z_1, z_2]$ and homogeneous polynomials $P(Z_0, Z_1, Z_2) \in \mathbb{C}[Z_0, Z_1, Z_2]$: $z_1^3 - z_2 z_1 + 1$ is sent to $Z_1^3 - Z_0 Z_1 Z_2 + Z_0^3$. Then, if $X = V(p) = \{(z_1, z_2) \mid p(z_1, z_2) = 0\} \subset \mathbb{C}^2$, then since $\mathbb{C}^2 \hookrightarrow \mathbb{C}P^2$ by $(z, w) \mapsto [1 : z_1 : z_2]$, then $X \subset \overline{X}$, which is $\{[Z_0 : Z_1 : Z_2] \mid P(Z_0, Z_1, Z_2) = 0\}$ (where p and P are identified by the above correspondence). Since $\mathbb{C}P^2$ is compact and $\overline{X}$ is closed, then $\overline{X}$ is compact.

We'd like to make $\overline{X}$ into a Riemann surface with $X \hookrightarrow \overline{X}$ holomorphic (in other words, extending the Riemann surface structure on X), which is the content of Proposition 7.2.

Proof of Proposition 7.2. Take $q = [Z_0 : Z_1 : Z_2] \in \overline{X}$. One of the Z_j is nonzero, so by the cyclic symmetry of this problem, we can assume $Z_0 \neq 0$, and scale to set $Z_0 = 1$. *Euler's identity on homogeneous polynomials* tells us that if $P(Z_1, \dots, Z_n)$ is a homogeneous polynomial (or, more generally, a homogeneous function), then

$$\sum Z_j \frac{\partial P}{\partial Z_j} = \deg P \cdot P(q).$$

The proof is but two lines, coming down to the chain rule, but is left as an exercise.

When we apply this to our choice of q, the takeaway is that

$$-\frac{\partial P}{\partial Z_0} = Z_1 \frac{\partial P}{\partial Z_1} + Z_2 \frac{\partial P}{\partial Z_2}.$$

In particular, one of $\frac{\partial P}{\partial Z_1}$ or $\frac{\partial P}{\partial Z_2}$ does not vanish. Since $p(z_1, z_2) = P(1, z_1, z_2)$, then this defines a Riemann surface $X_0 \subset U_0 = \mathbb{C}^2$, as we showed last time. In the same way, we can define Riemann surfaces $X_1 \subset \overline{X}$, where $Z_1 \neq 0$ and $Z_2 \subset \overline{X}$, where $Z_2 \neq 0$, and we know that $\overline{X} = X_0 \cup X_1 \cup X_2$.

Finally, one needs to check that the transition functions between these three components are holomorphic, which has been left as an exercise. ⊠

In a sense, $\overline{X}$ is just X along with the "points at infinity," which are the points in $\overline{X}$ where $Z_0 = 0$. If you start out with $X = V(p)$, where

$$p(z_1, z_2) = \sum_{i,j} a_{ij} z_1^i z_2^j, \qquad \text{so that} \qquad P(Z_0, Z_1, Z_2) = \sum_{i,j=0}^{d} a_{ij} Z_0^{d-i-j} Z_1^i Z_2^j,$$

and we can explicitly see what these points at infinity are:

$$P(0, Z_1, Z_2) = \sum_{i+j=d} a_{ij} Z_0^i Z_1^j.$$

In some sense, we keep only the terms of maximal degree. For instance, if $p(z_1, z_2) = z_1^2 - z_2(z_2 + 1)(z_2 - 1)$, which is a classic example of an elliptic curve, then $P(0, Z_1, Z_2) = -Z_2^3$, so the only point at infinity is $[0 : 1 : 0]$ (since $Z_2^3 = 0$). These points correspond to asymptotic behavior of your original polynomial (since the highest-degree terms dominate), which makes the name of "points at infinity" make sense.

This is a very geometric construction, depending on how you embed your Riemann surface into $\mathbb{C}$; as such, it doesn't have a whole lot of categorical significance.

Quotients. In enough detail, one could really spend an entire semester on quotients of the upper half-plane; many interesting Riemann surfaces can be realized as quotients of other Riemann surfaces by groups of biholomorphic maps.[11] The general construction is kind of hairy, but the idea can be conveyed well through a few examples. First, though, recall the following facts from complex analysis.

(1) $\mathrm{Aut}(S^2) = \mathrm{PSL}_2(\mathbb{C})$, which is also the group of Möbius maps. This is ultimately because any holomorphic map $S^2 \to S^2$ is a rational function.
(2) $\mathrm{Aut}\,\mathbb{C}$ is the set of maps in $\mathrm{PSL}_2(\mathbb{C})$ that send $\infty \mapsto \infty$, and these are therefore the maps $z \mapsto az + b$ with $a \in \mathbb{C}^*$ and $b \in \mathbb{C}$.[12]

[11] This is kind of a lame statement, since they *all* arise as quotients of S^2, $\mathbb{C}$, or $\mathbb{H}$, courtesy of the uniformization theorem.
[12] You can prove this using the Casorati-Weierstrass theorem on essential singularities, or think of this as holomorphic rational functions.

(3) Then, $\mathrm{Aut}(\mathbb{H}) = \{\mu \in \mathrm{PSL}_2(\mathbb{C}) \mid \mu(\mathbb{H}) = \mathbb{H}\}$, which is also $\mathrm{PSL}_2(\mathbb{R})$. Thus, understanding Riemann surfaces often really boils down to understanding free subgroups of this group. This also subsumes $\mathrm{Aut}(\mathbb{D})$, which is the same, because $\mathbb{H} \cong \mathbb{D}$ under a suitable Möbius map, so $\mathrm{Aut}(\mathbb{D})$ and $\mathrm{Aut}(\mathbb{H})$ are conjugate in $\mathrm{PSL}_2(\mathbb{C})$. In fact, $\mathrm{Aut}(\mathbb{D}) = \mathrm{PSU}(1,1)$; here, $\mathrm{SU}(1,1)$ is the group of unitary matrices preserving a signature-$(1,1)$ quadratic form. That is, if $\langle \underline{z}, \underline{w}\rangle = z_1\overline{w}_1 - z_2\overline{w}_2$, for $z, w \in \mathbb{C}^2$, then $\mathrm{SU}(1,1) = \{A \in \mathrm{SL}_2(\mathbb{C}) \mid \langle A\underline{z}, A\underline{w}\rangle = \langle \underline{z}, \underline{w}\rangle\}$, or more explicitly,
$$\mathrm{SU}(1,1) = \left\{ \begin{bmatrix} \alpha & \overline{\beta} \\ \beta & \overline{\alpha} \end{bmatrix} \mid |\alpha|^2 - |\beta|^2 = 1 \right\}.$$
Then, $\mathrm{PSU}(1,1)$ is the image of this in $\mathrm{PGL}_2(\mathbb{C})$, i.e. $\mathrm{SU}(1,1)/\{\pm I\}$.[13]

Now, what quotients do we get? If $\mu \neq \mathrm{id}$ is in $\mathrm{Aut}(S^2)$, then it fixes exactly two points on S^2, so no nontrivial subgroup of $\mathrm{Aut}(S^2)$ acts freely.

So instead, let's look at $\mathrm{Aut}(\mathbb{C})$. For example, $\mathbb{Z} \hookrightarrow \mathrm{Aut}(\mathbb{C})$ by $n \cdot z = z + n\lambda$, for a fixed $\lambda \in \mathbb{C}^*$. That is, $\mathbb{Z}$ acts by translation, scaled by λ. If $\Gamma < \mathrm{Aut}(\mathbb{C})$ is this subgroup, then $X = \mathbb{C}/\Gamma$, meaning the orbit space, is a Riemann surface. This looks like a cylinder, as small subsets of $\mathbb{C}$ project homeomorphically onto X, so we can create a chart structure by passing such images up to $\mathbb{C}$ and taking charts for them.

For example, if $r = |\lambda|/3$ and $z \in \mathbb{C}$, let $D_z = D(z,r)$; then, the quotient π maps D_z homeomorphically onto its image in X (since any two points whose image in the quotient is the same are at least $|\lambda|$ apart from each other). Then, the charts are $\pi(D_z) \to D_z$, since $\pi(D_z) \subset \mathbb{C}$, and the change-of-charts maps are just translations by $h\lambda$, which are smooth.

More generally, suppose $\Lambda \subset \mathbb{C}$ is a *lattice*: $\Lambda = \mathbb{Z} \oplus \mathbb{Z} \cdot \lambda$, where $\mathrm{Im}\,\lambda > 0$, as in Figure 6. This again acts on $\mathbb{C}$ by translation, and the same construction gives the quotient $X = \mathbb{C}/\Lambda$. Once again, taking $r = 1/3 \cdot \min(1, \mathrm{Im}\,\lambda))$ and $D_z = D(z,r)$ as the images of charts gives X a Riemann surface structure. Topologically, X looks like a torus.

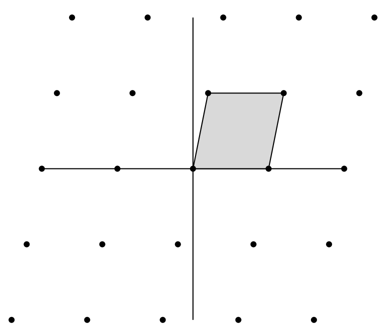


FIGURE 6. A lattice $\Lambda \subset \mathbb{C}$ and a fundamental domain for the quotient, which is a Riemann surface.

Lecture 9.

Fuchsian Groups: 2/10/16

"There's an elephant in the room, and it is hyperbolic geometry."

Today, we'll consider a specific example of quotient Riemann surfaces, quotients by the actions of Fuchsian groups. These are subgroups of $\mathrm{PSL}_2(\mathbb{R}) = \mathrm{Aut}(\mathbb{H})$, or, equivalently, $\mathrm{PSU}(1,1) = \mathrm{Aut}(\mathbb{D})$, as we established an isomorphism between these two groups, and in fact a conjugacy inside $\mathrm{PSL}_2(\mathbb{C})$.[14]

These groups have a natural topology to them. First, $\mathrm{SL}_2(\mathbb{C}) \subset \mathbb{C}^4$ (since it's a group of 2×2 matrices), so it has the subspace topology. Thus, $\mathrm{PSL}_2(\mathbb{C})$ has the quotient topology, and as subspaces, $\mathrm{Aut}(\mathbb{D})$ and $\mathrm{Aut}(\mathbb{H})$ gain the subspace topology.

[13]Usually, the story runs in reverse: using the Schwarz lemma, one discovers that all of the automorphisms of the disc are Möbius transformations, and then uses this to obtain $\mathrm{Aut}(\mathbb{H})$, $\mathrm{Aut}(S^2)$, and $\mathrm{Aut}(\mathbb{C})$.

[14]$\mathrm{SL}_2(\mathbb{C})$ is a four(-complex)-dimensional, complex Lie group, and $\mathrm{Aut}(\mathbb{D})$ and $\mathrm{Aut}(\mathbb{H})$ are 3-dimensional, noncompact, real Lie groups.

Definition. A *Fuchsian group*[15] Γ is a discrete subgroup of $\mathrm{Aut}(\mathbb{H})$.

The conjugacy of $\mathrm{Aut}(\mathbb{H})$ with $\mathrm{Aut}(\mathbb{D})$ means that this is equivalent to defining the conjugate subgroup $\Gamma' \leq \mathrm{Aut}(\mathbb{D})$.

We'd like to study quotients $\mathbb{H}/\Gamma$, or $\mathbb{D}/\Gamma'$; these turn out to all be nice Riemann surfaces. In general, we should use the hyperbolic structure on $\mathbb{H}$ or on $\mathbb{D}$ when talking about these quotients (remarkably, the biholomorphic functions exactly correspond with the isometries with respect to these hyperbolic structures). However, since we haven't done any differential geometry in this class, we'll adopt this perhaps more pedestrian, but more understandable approach.

To understand a $\gamma \in \mathrm{Aut}(\mathbb{H})$, we can think of it as a map $S^2 \to S^2$, and think about its fixed points. We'd like none of them to be in $\mathbb{H}$, so that the action is free and its quotient is a Riemann surface. γ is a fractional linear transformation $\gamma(z) = (az+b)/(cz+d)$, where $ad - bc = 1$.

- First, it's easy to check that $\gamma(\infty) = \infty$ iff $c = 0$.
- If $z \in \mathbb{C}$ is fixed, then $z = (az+b)/(cz+d)$, so $cz^2 + (d-a)z - b = 0$. If $c \neq 0$ and $\gamma \neq \mathrm{id}$, then after some case-checking, one sees that there's at most one fixed point in $\mathbb{C}$, which is actually in $\mathbb{R}$.
- The more interesting case is where $c \neq 0$, so the discriminant is $\Delta = (d-a)^2 + 4bc = (\operatorname{tr} A)^2 - 4$, where $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Thus, there are three cases:
  (1) γ is an *elliptic element* if $\Delta < 0$, i.e. $\mathrm{tr}^2(A) < 4$. In this case, $c \neq 0$, and there are two fixed points, one in $\mathbb{H}$ and its conjugate in $\overline{\mathbb{H}}$ (that is, the lower half-plane). By conjugating into $\gamma' \,\mathrm{Aut}(\mathbb{D})$, there's a unique fixed point in $\mathbb{D}$, and after a conjugation this is 0. But this means γ' is a rotation of the disc (e.g. by the Schwarz lemma); there's not quite such a simple description of $\gamma \in \mathrm{Aut}(\mathbb{H})$, but the point is that elliptic elements are conjugates of rotations. In particular, they may have finite or infinite order.
  (2) γ is a *parabolic element* if $\Delta = 0$, i.e. $\mathrm{tr}^2(A) = 4$. If $c \neq 0$, then γ has one fixed point which is in $\mathbb{R}$. If $c = 0$, then $A = \left[\begin{smallmatrix} 1 & b \\ 0 & 1 \end{smallmatrix}\right]$, so ∞ is fixed. Thus, in either case, there's one fixed point, and it's in $\mathbb{R} \cup \{\infty\}$, and γ is conjugate in $\mathrm{PSL}_2(\mathbb{R})$ to a μ fixing infinity and sending 0 to ± 1 (i.e. $\mu(z) = z \pm 1$). In particular, parabolic elements have infinite order.
  (3) γ is a *hyperbolic element* if $\Delta > 0$, so $\mathrm{tr}^2(A) > 4$. In this case, either $c \neq 0$, so there are two distinct fixed points in $\mathbb{R}$, or $c = 0$ and $a \neq d$, so there's one fixed point in $\mathbb{R}$, and ∞ is also fixed. Thus, in either case, there are two distinct fixed points in $\mathbb{R} \cup \{\infty\}$; such a γ is conjugate in $\mathrm{PSL}_2(\mathbb{R})$ to a μ fixing both 0 and ∞. Thus, $\mu(z) = \lambda z$, where $\lambda > 0$ and $\lambda \neq 1$. Thus, γ has infinite order.

This is somewhat elementary, but a complete description, and we can use it to talk about Fuchsian groups.

Example 9.1. Let p be a prime number, and let $\widetilde{\Gamma}_p = \{A \in \mathrm{SL}_2(\mathbb{Z}) \mid A \equiv \pm 1 \bmod p\} \subset \mathrm{SL}_2(\mathbb{R})$. Then, let $\Gamma_p = \widetilde{\Gamma}_p / \pm I$, which is a subgroup of $\mathrm{PSL}_2(\mathbb{R})$. An element $\gamma \in \widetilde{\Gamma}_p$ has the form

$$\gamma = \pm \begin{bmatrix} ap+1 & * \\ * & bp+1 \end{bmatrix},$$

where $a, b \in \mathbb{Z}$ and we don't know what the off-diagonal entries are. Then, $\operatorname{tr}\gamma = \pm((a+b)p + 2)$, which is generally not in $(-2, 2)$. In fact, if $p \geq 5$, then $|\operatorname{tr}\gamma| \geq 2$.[16] Thus, all elements of Γ_p are elliptic or hyperbolic, and therefore have no fixed points in $\mathbb{H}$. Hence, Γ_p acts freely on $\mathbb{H}$.

Proposition 9.2. *Let $\Gamma \subset \mathrm{Aut}(\mathbb{H})$ be a Fuchsian group.*

(1) *For all $z \in \mathbb{H}$, there's a neighborhood $N \subset \mathbb{H}$ of z such that if $q_1, q_2 \in N$ and $\gamma \in \Gamma$ satisfy $\gamma(q_1) = q_2$, then $\gamma(z) = z$ (i.e. $\gamma \in \mathrm{stab}_\Gamma(z)$).*
(2) *For all $q_1, q_2 \in \mathbb{H}$ such that $q_2 \notin \Gamma \cdot q_1$, there exist neighborhoods N_1 of q_1 and N_2 of q_2 such that $N_2 \cap \Gamma \cdot N_1 = 0$.*

Note. Part (2) says that the quotient is Hausdorff; if further Γ acts freely, then the condition from last lecture holds, and implies that the quotient is a Riemann surface. So we're always at least Hausdorff, and often a Riemann surface.

[15] These were named by Poincaré, not Fuchs, though Fuchs did study them.
[16] In fact, there are no elliptic elements if $p = 2$ or $p = 3$, and this requires a small but different argument.

Proof of Proposition 9.2, part (1). A really satisfying proof of this proposition would employ hyperbolic geometry, but we can give a hands-on proof of its first part.

We can work in $\mathbb{D}$ and without loss of generality assume $z = 0$ (since we can always conjugate by an element moving $z \mapsto 0$). Now, let $D = D(0, \varepsilon)$ (the disc of radius ε) and suppose $q, \gamma(q) \in D$ for some $\gamma \in \Gamma$.

$\gamma(z) = \alpha z + \beta/(\overline{\beta}z + \overline{\alpha})$, where $|\alpha|^2 - |\beta|^2 = 1$, so since $|q| < \varepsilon$ and $|\gamma(q)| < \varepsilon$, then

$$|\alpha q + \beta| \leq \varepsilon|\overline{\beta}q + \overline{\alpha}| \leq \varepsilon(|\beta|\varepsilon + |\alpha|).$$

We can use this to bound $|\beta|$, again by the triangle inequality: $|\beta| \leq \varepsilon(|\beta|\varepsilon + |\alpha|) + |\alpha|\varepsilon$, or $|\beta| \leq 2\varepsilon|\alpha|/(1-\varepsilon^2)$. But since $2\varepsilon/(1-\varepsilon^2) \to 0$ as $\varepsilon \to 0$, meaning we can make $|\beta|$ arbiitrarily small relative to $|\alpha|$.

The hypothesis we haven't used yet is that Γ is discrete, so suppose there is a sequence $\gamma_n \in \Gamma$ such that $|\beta_n|/|\alpha_n| \to 0$ as $n \to \infty$. Since $|\alpha_n|^2 - |\beta_n|^2 = 1$, then $|\alpha_n| \to 1$ and $|\beta_n| \to 0$. But since Γ is discrete, then eventually $|\alpha_n| = 1$ and $\beta_n = 0$. In other words, there's a $k \ll 1$ such that $|\beta| \leq k|\alpha|$: so $\beta = 0$ (e.g. take ε such that $2\varepsilon/(1-\varepsilon^) < k$), and thus $\gamma(z) = (\alpha/\overline{\alpha})z$, so it's a rotation about 0, and therefore fixes 0. ⊠

The proof of the second part can be found in the textbook.

The next surprising thing is that for *any* Fuchsian group, acting freely or not, $\mathbb{H}/\Gamma$ has the structure of a Riemann surface making the projection holomorphic. By part (2) of Proposition 9.2, we know it's Hausdorff, and we know how to make charts where the action is free, so we have to address the case where $\mathrm{stab}_\Gamma(z) \neq 1$.

The model case will be where $\Gamma_n = \mathbb{Z}/n$, which acts on $D(0;r)$ by $k \cdot z = e^{2\pi k/n}z$. Hence, $D(0;r)/\Gamma_n \cong D(0;r^n)$, by sending $[z] \mapsto z^n$ (this defines a well-defined homomorphism). This will be compatible with charts near the fixed point 0, so this quotient is a Riemann surface.

In general, if Γ is any Fuchsian group, then $\mathrm{stab}_\Gamma(0) = \Gamma_n$: it has to be a finite group of rotations (since it's a discrete subgroup of $\mathrm{U}(1)$), and for a general $z \in \mathbb{D}$, one can move it to 0 by conjugating to get a chart for it.

Lecture 10.

Properties of Holomorphic Maps: 2/15/16

First: there's no lecture Wednesday, and there may be lecture Friday. Also, read §5.1 of the textbook; it reviews calculus on manifolds: tangent vectors, cotangent vectors, and two-forms. We'll bootstrap it into calculus on Riemann surfaces later in this class.

Today, though, we're going to talk about holomorphic maps. We've seen a lot of ways in which Riemann surfaces arise (and in fact constructed all of them, thanks to the uniformization theorem). Today, the main focus will be on the structure of proper holomorphic maps.

Some properties from complex analysis generalize straightforwardly.

Lemma 10.1. *Let $f : X \to Y$ be a holomorphic map between Riemann surfaces, and $x \in X$. Then, the following are equivalent.*

(1) *f maps a neighborhood U of x homeomorphically to its image $V = f(U)$, and the inverse $f^{-1} : V \to U$ is holomorphic.*
(2) *In local coordinates near x and $f(x)$, $f'(x) \neq 0$.*[17]

(1) $\implies$ (2) by the chain rule: $f^{-1} \circ f = \mathrm{id}$, and then use the chain rule to show that $f'(z)$ is also invertible, so nonzero. (2) $\implies$ (1) relates to the inverse function theorem, and is proven using the argument principle in the same way as Theorem 4.7.

We have another lemma about the local behavior of holomorphic maps.

Lemma 10.2. *Again let $f : X \to Y$ be a holomorphic map and $x \in X$. If ψ is a holomorphic chart near $f(x)$, then there's a holomorphic chart ϕ near x such that $\widetilde{f} = \psi \circ f \circ \phi^{-1}$ takes the form $\widetilde{f}(z) = z^k$, for some integer $k \geq 0$ independent of the chart.*

So holomorphic maps look very simple, at least locally. This is a much stronger constraint than for smooth maps on smooth surfaces; Lemma 10.1 has an analogue for real manifolds, but Lemma 10.2 doesn't.

[17]There is a way to state this in a way that's independent of local coordinates, using the tangent bundle, and we'll get there in a few weeks.

Proof of Lemma 10.2. If $f'(x) \neq 0$, this reduces to Lemma 10.1: the homeomorphism defines charts for which $\widetilde{f}(z) = z$.

Otherwise, fix coordinates ψ near $f(x)$, and fix an initial choice of coordinates around x; by translation, we assume $x = 0$. In these charts,

$$\widetilde{f}(z) = \sum_{n \geq k} a_n z^n = a_k z^k \underbrace{\sum_{m \geq 0} \left(\frac{a_{m+k}}{a_k} \right) z^m}_{g(z)},$$

where $k > 1$ and $a_k \neq 0$, so $\widetilde{f}(0) = \widetilde{f}'(0) = 0$. Thus, $g(z)$ is holomorphic and $g(0) = 1$. Thus, we can define $h(z)$ to be a k^{th} root of $g(z)$, which is continuous and satisfies $h(0) = 1$, so if $a_k^{1/k}$ is any k^{th} root of a_k, then let $\phi(z) = a_k^{1/k} z h(z)$, so $\phi(z)^k = f(z)$, $\phi(0) = 0$, and $\phi'(0) = a_k^{1/k} \neq 0$, so by Lemma 10.2, ϕ is the desired coordinate chart.

We need to show this is independent of k, but this follows because $k = \min\{\ell \geq 1 \mid f^{(\ell)}(x) \neq 0\}$; this is invariant, so we're good. ⊠

This lemma is really part of complex analysis, but generalizes quite readily to Riemann surfaces.

Definition. If $x \in X$ is such that this $k \neq 1$, then x is called a *critical point* of f. The set of critical points is called $\operatorname{crit}(f)$.

These are exactly the same as the places where f' vanishes, and as the critical points of f regarded as a map between smooth manifolds.

Definition. A *critical value* of Y is a point in the image of $\operatorname{crit}(f)$. A *regular value* of f is a $y \in Y$ that's not a critical value. The set of regular values is denoted Y_0.

In differential topology, there's also the useful notion of the degree of a map; we'll find it useful and actually be able to reprove it in the holomorphic setting.

Definition. A continuous map $f : X \to Y$ of topological spaces is *proper* if whenever $K \subset Y$ is compact, $f^{-1}(K)$ is compact.

Fact. Let S and T be smooth, oriented surfaces, $f : S \to T$ be a proper smooth map, and $y \in Y$ be a regular value of f. For any $x \in f^{-1}(y)$, let

$$\varepsilon_x = \begin{cases} 1, & \text{if } f \text{ preserves orientation near } x \text{, and} \\ 0, & \text{otherwise.} \end{cases}$$

Then, we define $\deg_y(f) = \sum_{x \in f^{-1}(y)} \varepsilon_x$. The cool fact is that this is independent of y, and is denoted $\deg(f)$.

Returning to the world of Riemann surfaces, if $f : X \to Y$ is a proper, holomorphic map of Riemann surfaces. Since $\operatorname{crit}(f)$ is the zero set of the holomorphic f', then it's discrete (assuming f is nonzero). Thus, $\Delta = f(\operatorname{crit} f)$, the critical values, is also discrete.[18] Thus, if $y \in Y \setminus \Delta$ is a critical value, then for all $x \in f^{-1}(y)$, $f'(x) \neq 0$, so f is a local homeomorphism near x preserving orientation, so $\deg f = \sum_{x \in f^{-1}(y)} = |f^{-1}(y)|$: it just counts points in the preimage!

We can also understand this as follows: there's a fact from topology that any proper local homeomorphism is a covering map, so if $Y_0 = Y \setminus \Delta$, $X_0 = f^{-1}(Y_0)$, and $f_0 = f|_{X_0}$, then f_0 is a proper local homeomorphism, so a covering map with $\deg(f)$ sheets. Now, if $y \in Y$ is arbitrary (perhaps not regular), then $\deg(f) = \sum_{x \in f^{-1}(y)} k_x$, where k_x is the value for x coming from Lemma 10.2. At this point, the proof of this is very simple: since f locally looks like $z \mapsto z^{k_x}$ near x, then its degree there is just k_x, which is the sum of the preimages, or k_x^{th} roots, of a $y \neq 0$. Again, notice how much cleaner this is than for smooth functions.

As a consequence, we have the following theorem.

Theorem 10.3. *Let X be a compact, connected Riemann surface and $f : X \to S^2$ be a holomorphic function with just one pole[19] which is simple, then f is biholomorphic.*

[18] A theorem from general topology shows that the image of a discrete set under a proper map must also be discrete. This is thus considerably stronger than Sard's theorem.

[19] Recall that a *pole* of a function into S^2 is a point in the preimage of ∞.

Proof. The hypotheses imply that $\deg f = \deg f_\infty = 1$, so for any $y \in Y$, a positively weighted counts in $f^{-1}(y)$, so $f^{-1}(y)$ is always a single point. Thus, f is bijective, so the result follows from Lemma 10.1. ⊠

Remark. One can state this for the preimage of any $y \in S^2$, and even replace S^2 with any other compact Riemann surface, but S^2 is the place in which this is the most useful, because it corresponds to meromorphic functions on Riemann surfaces.

With f as before, $f_0 : X_0 \to Y_0 = Y \setminus \Delta$ is a covering map with $d = \deg f$ sheets. Then, lifting of paths is a group homomorphism called *monodromy*, $\mathrm{mon} : \pi_1(Y_0, y_0) \to \operatorname{Aut} F$, where $F = f^{-1}(y_0)$; if X is connected, then this action is transitive on F. Hence, the data from a proper holomorphic map $f : X \to Y$ consists of a discrete set of critical values and a transitive homomorphism on the permutations of the fiber, which is a pretty nice topological thing to extract. Next time, we'll prove Riemann's existence theorem (Theorem 11.2), which provides a converse to this.

Lecture 11.

The Riemann Existence Theorem: 2/22/16

There will be an additional, optional lecture regarding calculus on surfaces; alternatively, you can read §5.1 of the textbook.

Today, we wil, cover the Riemann existence theorem and discuss normalization of algebraic curves (a way of removing singularities).

Covering Spaces and Monodromy. Let Z be a path-connected space that admits a universal cover $p : \widetilde{Z} \to Z$, e.g. a smooth surface. Fix a basepoint $b \in Z$ and let $\pi = \pi_1(Z, b)$. We usually regard connected covering spaces of Z as arising from subgroups $H \le \pi$: we have $p : \widetilde{Z}/H \to Z$, where H acts on $\widetilde{Z}$ by deck transformations.

There is a variant viewpoint which may be more useful for understanding maps of covering spaces. Given a connected covering space $\pi : Y \to Z$, it has a typical fiber $F = \pi^{-1}(b)$, and monodromy $\mathrm{mon} : \pi \to \operatorname{Aut}(F)$, a group homomorphism from the fundamental group of the base to permutations of the fiber. This is defined by path lifting: the unique lift of a path moves points in the fiber around. This action is transitive (since Y is connected) making F a *transitive π-set.*[20] Then, if we choose an $f \in F$, let $H = \mathrm{stab}_\pi(f) \le \pi$; then, we can recover $Y = \widetilde{Z} \times_\pi F$. This *fiber product* is $\widetilde{Z} \times F$ modulo the equivalence relation $(z, f) \sim (z \cdot g^{-1}, \mathrm{mon}(g) \cdot f)$ for all $g \in \pi$.

Theorem 11.1. *In fact, this identification is an equivalence of categories between*

1. *the category of connected covering spaces of Z and*
2. *the category of* canonical π-orbits*, i.e. π-sets of the form π/H, with $H \le \pi$, and with morphisms given by π-equivariant maps.*

A reference for this is May's *A Concise Course in Algebraic Topology*, which is ideal more as a second course for covering spaces than a first one.

The Riemann Existence Theorem. Last time, we established that if X and Y are connected Riemann surfaces and $F : X \to Y$ is a proper holomorphic map of degree d, then we can extract

- a discrete set $\Delta = F(\operatorname{crit} F) \subset Y$ of critical values, and
- for any basepoint $y_0 \in Y \setminus \Delta$, a monodromy homomorphism $\pi(Y \setminus \Delta, y_0) \to \operatorname{Aut}(\pi^{-1}(y_0))$.

Moreover, on $Y \setminus \Delta$, F is a covering map, and near a critical point $p \in \operatorname{crit} F$, F has the form $z \mapsto z^k$ for $k > 1$ in some coordinates. These critical points are called *branch points*, and F is called a *branched covering map*. This is a little bit tangled, but can be summed up as: if you encounter a branched cover of Riemann surfaces, you're really looking at a proper holomorphic map.

That's technically the converse, but it's still valid, and is called the Riemann existence theorem. Here, S_d denotes the symmetric group on d elements.

[20] A transitive G-set is just a set X with a transitive action of the group G on it.

Theorem 11.2 (Riemann existence theorem)**.** *Let Y be a connected Riemann surface, $\Delta \subset Y$ be a discrete subset, $y_0 \in Y$, and $\rho : \pi_1(Y \setminus \Delta) \to S_d$ be a transitive*[21] *group homomorphism. Then, there exists a connected Riemann surface X and a degree-d holomorphic map $F : X \to Y$ with $\Delta = \operatorname{crit} F$ and monodromy ρ.*

Remark. In fact, there's a category of proper holomorphic maps to Y with critical values Δ, and this category is equivalent to the category of finite canonical $\pi_1(Y \setminus \Delta, y_0)$-orbits.

Proof of Theorem 11.2. By the theory of covering spaces, there is a d-sheeted covering $F_0 : X_0 \to Y_0 = Y \setminus \Delta$ with monodromy ρ; our job is to fill in the missing points of X which will map to Δ.

The story will be same over every $\delta \in \Delta$, so let's just focus on one. Let γ be a loop starting and ending at y_0, and circling δ once, and set $\sigma = \rho(\gamma) \in S_d$. We may write σ as a product of disjoint cycles, $\sigma = c_1 \circ \cdots \circ c_r$, where the c_j are disjoint cycles, and let ℓ_j be the length of c_j.

Let D be a small coordinate disc in Y centered at δ, so $D^* = D \setminus \{0\}$ is a chart for Y_0. Then, we have a covering map $F_0 : F_0^{-1}(D^*) \to D^*$ whose sheets come together via monodromy. If D is sufficiently small (and we can shrink it if it isn't), $F_0^{-1}(D^*)$ is a disjoint union of $E_1, \dots, E_r$, where the maps $F_0 : E_j \to D^*$ is equivalent to the map $m_j : \mathbb{D}^* \to \mathbb{D}^*$ sending $z \mapsto z^{\ell_j}$. This equivalence follows from the classification of coverings of a punctured disc, so let's make this identification for each j.

Now, let e_j be a copy of $\mathbb{D}$ for each j, and define

$$X = \left(X_0 \cup \bigcup_{j=1}^{r} e_j \right) / \sim,$$

where $x \in \mathbb{D}^*$ is identified with the corresponding point in $\mathbb{D}$ for all $x \in E_j = \mathbb{D}^*$. All we've done is add the center of each disc, gluing to all the sheets of the cover; all the other points are identified with points that already existed. Thus, X is Hausdorff, because we only have to check the new points, and these are easy to separate from other points. Moreover, near each new point, there's a chart $\mathbb{D} \hookrightarrow X$, and so X is a Riemann surface.

Once we do this to all $\delta \in \Delta$, we obtain a holomorphic $F : X \to Y$ extending F_0, and on each E_j, this is an extension of $m_j : \mathbb{D}^* \to \mathbb{D}^*$ to $m_j : \mathbb{D} \to \mathbb{D}$. ⊠

Normalization of Algebraic Curves. One application of this is a canonical way to "de-singularize" plane algebraic curves, called normalization. It extends throughout algebraic geometry to any algebraic variety, which won't remove all singularities, but will push them into codimension 2.

Let $p(z, w) \in \mathbb{C}[x, y]$ be irreducible and $X = \{p = 0\} \subset \mathbb{C}^2$ be its zero set. Sitting $\mathbb{C}^2 \subset \mathbb{C}P^2$, we have a compactification $\overline{X}$, the zero set of the homogenization of p. In general, X and $\overline{X}$ are singular; let $S = \{(z, w) \in X \mid \frac{\partial p}{\partial w} = \frac{\partial p}{\partial z} = 0\}$ be the set of possible singularities of X, and the possible singularities of $\overline{X}$ are S and the points at infinity. For example, if $p(z, w) = w^2 - z^3 = 0$, we have a single singularity at the origin, and if $p(z, w) = w^2 - z^2(z + 1)$, the zero set self-intersects itself at the origin, called a *node*; there are two distinct tangent spaces. In both cases, $X \setminus S$ is a Riemann surface.

We'll prove this result next time.

Theorem 11.3. *There's a canonical construction leading to*

- *a compact Riemann surface X^*,*
- *an inclusion map $i : X \setminus S \hookrightarrow X^*$ embedding X as a dense, open subset, and*
- *a holomorphic map $\nu : X^* \to \mathbb{C}P^2$ extending the inclusion $X \setminus S \hookrightarrow \mathbb{C}P^2$, with $\operatorname{Im}(\nu) = \overline{X}$.*

(X^*, ν) is called the *normalization* of X in $\mathbb{C}P^2$. The idea is to replace things such as self-intersections with branches of a branched, 1-sheeted cover.

Lecture 12.

Normalizing Plane Algebraic Curves: 2/24/16

Today, we'll continue (and formalize) our discussion of normalization of plane algebraic curves. Let $p(z, w) \in \mathbb{C}[z, w]$ be irreducible and $P(Z_0, Z_1, Z_2)$ be its homogenization. Then, we have an algebraic curve $X = \{p = 0\} \subset \mathbb{C}^2$ and its projective closure $\overline{X} = \{P = 0\} \subset \mathbb{C}P^2$. Then, there are possible singularities,

[21] A group homomorphism $\varphi : G \to H$ is *transitive* if for all $h_1, h_2 \in H$, there's a $g \in G$ such that $\varphi(g)h_1 = h_2$; that is, the action of G on H through φ is transitive. For example, $\mathbb{Z}/n$ acts transitively on S_n by sending $i \mapsto i + 1 \bmod n$.

which are the set S where both partials of p vanish, and these sit inside $\overline{S} \subset \overline{X}$, where all three partials of P vanish.

We know that $\overline{X} \setminus \overline{S}$ is naturally a Riemann surface, but in general, if X has singularities, it's not a Riemann surface, and not even a topological surface. The typical example is a curve which intersects itself (it's hard to gain intuition about this through pictures, since they only capture the real part). For example, if $p(z, w) = w^2 - z^2(1 - z)$, the real curve intersects itself at the origin. If z is small, this factors as $(w - z\sqrt{1 - z})(w + z\sqrt{1 - z})$, which makes sense when $|z| < 1$. The square root is holomorphic, but we do need to choose a branch for it. Let $x = w - z\sqrt{1 - z}$ and $y = w + z\sqrt{1 - z}$.

In the local holomorphic coordinates (x, y) near $(0, 0)$, $X = \{xy = 0\}$, so every small, punctured neighborhood of the origin is disconnected, homeomorphic to the disjoint union of two punctured discs. Thus, X is not a surface of any kind.

We will write down a recipe to construct the normalization of $\overline{X}$, which is a compact X^* along with a surjective, continuous map $\nu : X^* \twoheadrightarrow \overline{X}$, such that $\nu : \nu^{-1}(\overline{X} \setminus \overline{S}) \to \overline{X} \setminus \overline{S}$ is a biholomorphism, so we have a commutative diagram

$$\begin{tikzcd}
\mu^{-1}(\overline{X} \setminus \overline{S}) \arrow[r, hook] \arrow[d, "\nu", "\wr"'] & X^* \arrow[d, "\nu"] \arrow[dr, "\text{incl.}"] & \\
\overline{X} \setminus \overline{S} \arrow[r, hook] & \overline{X} \arrow[r, "\text{incl.}"] & \mathbb{C}P^2.
\end{tikzcd}$$

There are a few different ways to go about this; in algebraic geometry, it has to do with integrality of rings of integers, but in this class we will see a more geometric construction. The idea is to consider projection $\pi : X \to |C$ given by $(z, w) \mapsto z$. We'd like to say that π is proper, and hence a branched covering describable in terms of its critical values and monodromy data. Then, we can use the Riemann existence theorem, Theorem 11.2, to extend this to a branched covering of S^2.

This is not going to work as stated, because π may not be proper. But we can throw out some "bad" points to make this work, and this is broadly how the construction is going to go.

First, let's get rid of a degenerate case: if $\frac{\partial p}{\partial w} = 0$ everywhere, then $p = p(z)$ is an irreducible polynomial in one variable. By the fundamental theorem of algebra, this means it's linear. Hence, $\overline{X}$ is already a Riemann surface, so we can let $X^* = \overline{X}$ and $\nu = \mathrm{id}$, which satisfies the normalization property.

Having dealt with this trivial case, we can assume $\frac{\partial p}{\partial w} \neq 0$ at least somewhere, i.e. the set $T = \{x \in X \mid \frac{\partial p}{\partial w}(x) = 0\}$ isn't all of X. This keeps track of singular points S along with vertical tangencies, which are the critical points of π.

Fact. T is finite.

The proof of this fact involves the Riemann-Roch theorem, which we haven't covered yet; it's in chapter 11 of the textbook. But the point is, S is finite too, so we can ask whether $\pi : X \setminus S \to \mathbb{C}$ is proper.

Unfortunately, this is not always the case, such as when $p(z, w) = zw + z - 1$. Then, the zero set is when $w = 1/z - 1$, so as $z \to 0$, $w \to \infty$; this map is not proper.

Nonetheless, we can write

$$p(z, w) = \sum_{j=0}^{d} a_j(z) w^j,$$

where $a_j \in \mathbb{C}[z]$ and a_d isn't identically zero, so let $F = \{z \in \mathbb{C} \mid a_d(z) = 0\}$. These are the points that actually cause us trouble.

Lemma 12.1. *$\pi : X \setminus \pi^{-1}(\pi(S) \setminus F) \to \mathbb{C} \setminus (\pi(S) \cup F)$ is proper.*

Proof idea. The point is that over a compact $K \subset \mathbb{C} \setminus (\pi(S) \cup F)$, we can divide $p(z, w)$ by $a_d(z$ to obtain something of the form

$$\frac{p(z, w)}{a_d(z)} = w^d + \sum_{j=0}^{d-1} b_j(z) w^j = 0.$$

Each b_j is holomorphic, and so on K, $|b_j(z)|$ is bounded. This implies that the solutions w must be bounded as well, which is essentially what it means to be proper.

We will write $S^+ = \pi^{-1}(\pi(S) \cup F)$, $E = \pi(S) \cup F \cup \{\infty\}$, and $\pi^+ : X \setminus S^+ \to S^2 \setminus E$ be the restriction of π. If $p(z, w) = w^2 - z^2(1 - z)$, we can work through this explicitly. Here, $F = \varnothing$ and $S = \{(0, 0)\}$, so $S^+ = S$ and $E = \{0, \infty\}$.

Now, we can apply the Riemann existence theorem to S^2, the base. There will be a discrete subset $\Delta = E \cup \pi^+(\operatorname{crit} \pi^+)$, and a monodromy map around each point of Δ which is the monodromy of π^+. The result is a compact Riemann surface X^* and a degree-d map $\pi^* : X^* \to S^2$, where π^* is branched over $E \cup \pi^+(\operatorname{crit} \pi^+)$.

In our example, the only critical value is 1, where there is a vertical tangency, so we get a $\pi^* : X^* \to S^2$, which is a degree-2 branched cover, branched over 0, 1, and ∞. This restricts to a $2 : 1$ covering map of $\mathbb{C} \setminus \{0, 1\}$. What's the monodromy? Near $z = 0$, there are two distinct branches for w, given by $w = \pm z\sqrt{1 - z}$. Thus, the monodromy here is trivial, so there's actually no branching there; we can put 0 back in, and still have a $2 : 1$ cover. On the other hand, there is nontrivial monodromy around 1 (for reasons of time, we won't check this, but it's not hard). Thus, the cover is connected, or nontrivial.

We can identify this cover with the covering $t \mapsto (1 - t^2)$, which is branched over 1 (since there's a multiple root $t = 0$ to $1 - t^2 = 1$; otherwise, there are two roots). Sending $\infty \mapsto \infty$ makes this into a map $X^* \to S^2$ branched at 1.

So far, for a general p this is an abstract construction, but we'd like to know that $X \setminus S^+$ includes into X^*, but the uniqueness of proper maps with prescribed critical values and monodromy (the uniqueness clause in the Riemann existence theorem) informs us that we have a diagram

$$\begin{tikzcd}
X \setminus S^+ \arrow[r, hook] \arrow[d] & X^* \arrow[d, "\pi^*"] \\
S^2 \setminus E \arrow[r, hook] & S^2.
\end{tikzcd}$$

Here, the green arrow is an open, dense inclusion.

The last step is to construct $\nu : X^* \twoheadrightarrow \overline{X}$, which extends holomorphically over π in the affine case. The proof can be found in the textbook.

In our example, set $z = 1 - t^2 = \pi^*(t)$; for $(z, w) \in X$, $w^2 = z^2(1 - z^2) = t^1(1 - t^2)$, so we can set $w = (1 - t^2)t$. Then, ν is fairly clear: we extend to $\mathbb{C}P^2$ by defining $\nu : S^2 = \mathbb{C}P^1 \to \mathbb{C}P^2$ by $\nu[z_0, z_1] = [1, 1 - (z_0/z_1)^2, z_0/z_1 - (z_0/z_1)^3]$; after clearing denominators, this is the same thing as $[Z_1^3 : Z_1^3 - Z_0^2 Z_1 : Z_0 Z_1^2 - Z_0^3]$, which is exactly the homogenization.

Lecture 13.

The de Rham Cohomology of Surfaces: 2/26/16

Today, we're going to cover de Rham cohomology on surfaces, corresponding to §5.2 of the textbook. After the differential topology treatment, this may be review to a lot of people, so we'll give a sketchy overview and hopefully a few new things. A good reference is Bott-Tu, *Differential Forms in Algebraic Topology*.

For the rest of this lecture, let S be a smooth surface, which need not have a Riemann surface structure. We have an $\mathbb{R}$-algebra of smooth functions $C^\infty(S) = \{f : S \to \mathbb{R} \mid f \text{ is smooth}\}$. Then, we can define three $C^\infty(S)$-modules.

- $\Omega^0(S)$, the 0-*forms*, are just $C^\infty(S)$ regarded as a module of itself.
- $\Omega^1(S)$ is the 1-*forms*, the functions which smoothly assign to each $p \in S$ an $\alpha_p \in T_p^* S$ (so a section of the cotangent bundle); in local coordinates (x, y), this is given by $\alpha = f(x, y)\, \mathrm{d}x + g(x, y)\, \mathrm{d}y$ for smooth $f, g : S \to \mathbb{R}$.
- Then, $\Omega^2(S)$ is the 2-*forms*, the functions which smoothly assign to each $p \in S$ an element $\omega_p \in \Lambda^2 T_p^* S$, which is skew-symmetric bilinear maps $T_p S \times T_p S \to \mathbb{R}$. In local coordinates, this takes on the form $\omega = f(x, y)\, \mathrm{d}x \wedge \mathrm{d}y$, where f is smooth, so 2-forms are essentially functions locally. Globally, though, S can parameterize a nontrivial family of one-dimensional vector spaces $\Lambda^2 T_p^* S$, and hence the global behavior may be more interesting.

We also have the *exterior derivative* operators d^0 and d^1:

$$\Omega^0(S) \xrightarrow{\mathrm{d}^0} \Omega^1(S) \xrightarrow{\mathrm{d}^1} \Omega^2(S). \tag{13.1}$$

It takes a little bit of work to define these rigorously, but these have nice properties. For one, they're local: if $\alpha, \beta \in \Omega^j(S)$ are such that $\alpha|_U = \beta|_U$, then $\mathrm{d}^j\alpha|_U = \mathrm{d}^j\beta|_U$. If f,, α, and β are smooth functions on S, then in local coordinates,

$$\mathrm{d}^0 f = \frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y \qquad \text{and} \qquad \mathrm{d}^1(\alpha\,\mathrm{d}x + \beta\,\mathrm{d}y) = \left(\frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y}\right)\mathrm{d}x \wedge \mathrm{d}y.$$

In particular, $\mathrm{d}^1 \circ \mathrm{d}^0 = 0$. If $F : S_1 \to S_2$ is a smooth map, we have *pullback* operators, which are linear maps $F^* : \Omega^j(S_2) \to \Omega^j(S_1)$; for example, on Ω^0, $f \mapsto f \circ F$. The fundamental fact is that $F^* \circ \mathrm{d}^j = \mathrm{d}^j \circ F^*$: certainly, plenty of operators satisfy $\mathrm{d}^2 = 0$, but there are fairly nice uniqueness results about operators commuting with all pullbacks. This is particularly nice, because it means the local structure doesn't depend on the chart.

Finally, there's a *Leibniz rule*: $\mathrm{d}^0(fg) = \mathrm{d}^0 f + f\mathrm{d}^0 g$ and $\mathrm{d}^1(f\alpha) = \mathrm{d}f \wedge \alpha + f\,\mathrm{d}\alpha$, where $f, g \in C^\infty(S)$ and $\alpha \in \Omega^1(S)$.

Definition. A differential form $\alpha \in \Omega^j(S)$ has *compact support* if it's zero outside of some compact $K \subset S$. The $C^\infty(S)$-module of j-forms with compact support is written $\Omega^j_c(S)$.

Since d^j is local, it preserves the property of being compactly supported; in particular, we have a variant of (13.1).

$$\Omega^0_c(S) \xrightarrow{\mathrm{d}^0} \Omega^1_c(S) \xrightarrow{\mathrm{d}^1} \Omega^2_c(S) \tag{13.2}$$

We can define the *de Rham cohomology* to be the cohomology of the complex in (13.1).

- We define a vector space $H^0(S) = \ker(\mathrm{d}^0)$.
- $H^1(S) = \ker(\mathrm{d}^1)/\operatorname{Im}(\mathrm{d}^0$: since $\mathrm{d}^1 \circ \mathrm{d}^0 = 0$, this is well-defined.
- $H^2(S) = \Omega^2(S)/\operatorname{Im}(\mathrm{d}^1)$.

If we use (13.2) instead of (13.1), the same definitions lead to *compactly supported cohomology* $H^j_c(S)$. Both of these are diffeomorphism invariants of S: a diffeomorphism $S \to S'$ induces an isomorphism on all of these groups.

Today, we'd like to understand a little about these spaces. First, if S is compact, we can see that $H^j_c(S) = H^j(S)$ simply through their constructions.

Lemma 13.3.

- *$H^0(S)$ is the space of locally constant functions; in particular, if S is connected, $H^0(S) \cong \mathbb{R}$.*
- *Likewise, $H^0_c(S)$ is the space of locally constant, compactly supported functions. If S is connected, $H^0_c(S)$ is $\mathbb{R}$ if S is compact, and 0 otherwise.*

This is just a matter of the definitions.

Integration. Suppose S is a smooth, *oriented* surface, meaning that it has an atlas where all change-of-charts maps have a positive Jacobian. In this case, one can define an *integration map* $\int_S : \Omega^2_c(S) \to \mathbb{R}$ characterized by the properties that

(1) it's $\mathbb{R}$-linear, and
(2) suppose $U \subset S$ is open and $\phi : U \to \widetilde{U} \subset \mathbb{R}^2$ is an oriented chart. If $\omega \in \Omega^2_c(S)$ has its support inside U and $\omega|_U = \phi^*(f(x,y)\,\mathrm{d}x \wedge \mathrm{d}y)$, then

$$\int_S \omega = \int_{\widetilde{U}} f(x,y)\,\mathrm{d}x\,\mathrm{d}y.$$

Proving that this actually defines something takes a bit of work, but remarkably, one result is that if $F : S_1 \to S_2$ is an orientation-preserving, smooth map and $\omega \in \Omega^2_c(S)$, then

$$\int_{S_1} F^*\omega = \int_{S_2} \omega,$$

so integration is completely independent of coordinates! This is very unlike ordinary integration, and is one of the reasons forms pop in: the Jacobian is absorbed into the pullback, and makes this coordinate-free and canonical. It's definitely worth saying more about this, but that's for the differential topology prelim course.

Theorem 13.4 (Stokes)**.** *If $\alpha \in \Omega^1_c(S)$, then $\int_S \mathrm{d}\alpha = 0$.*

As a corollary, this means that integration factors through the quotient by $\mathrm{Im}(\mathrm{d}^1)$ to define an integration map $\int_S : H_c^2(S) \to \mathbb{R}$.

This was all in the prelim course, but next we'll prove something that's generally not included in the prelim.

Proposition 13.5. *$\int_{\mathbb{R}^2} : H_c^2(\mathbb{R}^2) \to \mathbb{R}$ is an isomorphism.*

Proof. The analogous fact on $\mathbb{R}$ is that if $f : \mathbb{R} \to \mathbb{R}$ is compactly supported and is such that $\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t$, then f is the derivative of a compactly supported $F \in C_c^\infty(\mathbb{R})$; in fact,

$$F(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t,$$

and this is never an improper integral, since f is compactly supported. Then, $F' = f$, certainly, and F is compactly supported, because if x is below $\mathrm{supp}(x)$, we're just integrating the zero function, and if x is above it, then we've integrated all of f, so by assumption this is just 0 again. We're going to mimic this argument in two variables.

Now, suppose $\omega \in \Omega_c^2(\mathbb{R}^2)$; since $\mathbb{R}^2$ is its own single chart, then $\omega = f(x,y)\,\mathrm{d}x \wedge \mathrm{d}y$, where f is compactly supported. Let $\omega \in C_c^\infty(\mathbb{R})$ be such that $\int_{-\infty}^{\infty} \omega(t)\,\mathrm{d}t = 1$, and let $I_y = \int_{-\infty}^{\infty} f(x,y)\,\mathrm{d}x$ and $\widetilde{f}(x,y) = f(x,y) - I_y\phi(x)$. Hence,

$$\int_{-\infty}^{\infty} \widetilde{f}(x,y)\,\mathrm{d}x = I_y - I_y \int_{-\infty}^{\infty} \phi(x)\,\mathrm{d}x = 0.$$

Thus, from the single-variable case, $\widetilde{f}(x,y) = \frac{\partial}{\partial x}P(x,y)$ for some smooth, compactly supported P. In particular, $f(x,y) = \frac{\partial P}{\partial x} + I_y\phi(x)$, and

$$\int_{-\infty}^{\infty} I_y\,\mathrm{d}y = \int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0,$$

since f is compactly supported, and therefore $I_y \cdot \phi(x) = \frac{\partial Q}{\partial y}$ for some compactly supported Q. This means $\omega = f(x,y)\,\mathrm{d}x \wedge \mathrm{d}y = \mathrm{d}(P\,\mathrm{d}y + Q\,\mathrm{d}x)$, so $\omega = 0$ in $H_c^2(\mathbb{R})$. ⊠

This generalizes to the following.

Theorem 13.6. *Let S be a connected, oriented surface. Then, $\int_S : H_c^2(S) \to \mathbb{R}$ is an isomorphism.*

Proof. It's clear that $\int_S$ is surjective. Let $\rho \in \Omega_c^2(S)$ be such that $\int_S \rho = 0$. We'd like to show that $\rho = \mathrm{d}\alpha$ for an $\alpha \in \Omega_c^1(S)$. Then, ρ has support in a compact set K, which we can take to be connected.[22] In particular, we can cover K by n coordinate charts, and proceed by induction on n. The result for $n = 1$ follows from Proposition 13.5 (and the diffeomorphism-invariance of Ω_c^*).

If instead $n > 1$, let $U_1, \dots, U_n$ be our cover, and set $U = U_1$ and $V = U_2 \cup \cdots \cup U_n$. If $K \subset U$ or $K \subset V$, we're done by induction, but if not, then $K \cap U \cap V \neq \varnothing$, so take a $p \in U \cap V$ and let $\beta \in \Omega_c^2(U \cap V)$ be such that $\int_{U\cap V} \beta = 1$ (e.g. by using a bump function).

Using a technique called a *partition of unity*, one can find $f_1, f_2 \in C^\infty(U \cup V)$ such that f_1 is supported in U, f_2 is supported in V, and $f_1 + f_2 = 1$. In particular, $\rho = f_1\rho + f_2\rho$ on $U \cup V$. Hence,

$$f_j\rho - \left(\int_{U\cup V} f_j\rho\right)\beta$$

integrates to 0 on U, so it's $\mathrm{d}\alpha_j$ for some compactly supported α_j by the inductive assumption, and now $\rho = \mathrm{d}(\alpha_1 + \alpha_2)$. ⊠

[22] This was left as an exercise in lecture.

Lecture 14.

de Rham Cohomology in Degree 1: 2/29/16

Today, we're going to do some more de Rham cohomology, which on the one hand isn't specifically about Riemann surfaces, but on the other hand isn't covered in the prelim, and is going to be important in our discussion of Riemann-Roch.

Last time, we proved Theorem 13.6, that integration provides an isomorphism between $H_c^2(S) \to \mathbb{R}$ for all smooth, oriented surfaces S. Today, we're going to investigate $H^1(S)$: what does "closesd 1-forms modulo exact 1-forms" mean? There are at least three or four answers, and we'll go through two of them today.

Obstruction to Global Primitives. One answer is that it measures the obstruction to assembling local primitives f_i for a closed 1-form into a globally-defined primitive for α. Specifically, suppose $S = \bigcup_{i \in I} U_i$ and we know that $\alpha|_{U_i} = af_i$; in this case, can we stitch these together into a global primitive? We'll understand this more precisely using Čech cohomology.

The answer begins with the following lemma.

Lemma 14.1 (Poincaré)**.** $H^1(\mathbb{R}^2) = H^2(\mathbb{R}^2) = 0$.

In particular, a closed 1-form on a surface is locally exact. Thus, we can choose an open cover $\mathfrak{U}$ of S such that every $U \in \mathfrak{U}$ is diffeomorphic to $\mathbb{R}^2$, and hence $H^1(U) = 0$. Hence, if $\alpha \in \Omega^1(S)$ and $\mathrm{d}\alpha = 0$, then $\alpha|_{U_i} = af_i$, where $f_i \in \Omega^0(U_i)$. On the overlaps $U_{ij} = U_i \cap U_j$, it's not necessarily true that $f_i = f_j$. However, we do have $\mathrm{d}(f_i - f_j) = 0$ on U_{ij}, and hence $c_{ij} = f_i - f_j$ is locally constant.

Definition.

- A *Čech 1-cochain* ζ for $(S, \mathfrak{U})$ is an assignment of a locally constant ζ_{ij} on $U_{ij} = U_i \cap U_j$ for all $U_i, U_j \in \mathfrak{U}$. These form a vector space, denoted $\check{C}^1(S; \mathfrak{U})$ or $\check{C}^1$.
- A Čech 1-cochain is a 1-*cocycle* if for all triples $U_i, U_j, U_k \in \mathfrak{U}$, the *cocycle condition* $\zeta_{ij} + \zeta_{jk} + \zeta_{ki} = 0$ on $U_i \cap U_j \cap U_k$. In particular, $\zeta_{ii} = 0$ and $\zeta_{ij} + \zeta_{ji} = 0$. 1-cocycles form a subspace $\check{Z}^1 \subset \check{C}^1$.
- A Čech 1-cocycle is a 1-*coboundary* if there exist locally constant functions $f_i \in \Omega^0(U_i)$ for each $U_i \in \mathfrak{U}$ such that on U_{ij}, $\zeta_{ij} = f_i - f_j$. The 1-coboundaries form a subspace $\check{B}^1 \subset \check{Z}^1$.
- Finally, we define the *Čech cohomology* $\check{H}^1(S; \mathfrak{U}) = \check{Z}^1 / \check{B}^1$.

The procedure above that assigned to a closed 1-form α the data (c_{ij}) defines a linear map $\check{c} : H^1(S) \to \check{H}^1(S; \mathfrak{U})$, because (c_{ij}) is a cocycle, and is well-defined up to coboundaries. If $\alpha = \mathrm{d}f$, we can take $c_{ij} = f|_{U_i} - f_{U_j} = 0$.

Definition. A cover $\mathfrak{U}$ of a surface S is *acyclic* if it's locally finite and for all nonempty intersections $U_{i_1 i_2 \cdots i_N} = U_{i_1} \cap \cdots \cap U_{i_N}$, we have $H^j(U_{i_1 i_2 \cdots i_N}) = 0$.

Hence, it suffices to make each intersection diffeomorphic to $\mathbb{R}^2$. Such an acyclic cover always exists: we can embed S properly into $\mathbb{R}^N$ for some N large, and then cover S by Euclidean balls in $\mathbb{R}^N$. By making them smaller if necessary (so S looks like a linear subspace).[23]

Theorem 14.2. *If $\mathfrak{U}$ is an acyclic cover, then $\check{c}$ is an isomorphism.*

Note. This is a special case of de Rham's theorem; this formulation is due to André Weil. And as an interesting corollary, this means the Čech cohomology is independent of the cover $\mathfrak{U}$, so long as it satisfies the hypotheses.

Proof of Theorem 14.2. Though we'll write down an inverse map, let's first see why $\check{c}$ is injective. Suppose $\check{c}([\alpha]) = 0$, so that it's a coboundary. Hence, $\alpha|_{U_i} = af_i$ and $c_{ij} = f_i - f_j$ on U_{ij}, and there exist locally constant functions $a_i \in \Omega^0(U_i)$ such that $c_{ij} = a_i - a_j$. Then, $\mathrm{d}(f_i - a_i) = \alpha$ and $(f_i - a_i) - (f_j - a_j) = c_{ij} - c_{ij} = 0$ on U_{ij}, meaning there's a $g \in \Omega^0(S)$ such that $g|_{U_i} = f_i - a_i$; hence, $\mathrm{d}g = \alpha$.

For surjectivity, let $\{\rho_i\}$ be a partition of unity for $\mathfrak{U}$, so $\rho_i \in \Omega^0(S)$, $\operatorname{supp}(\rho_i) \subset U_i$, and $\sum \rho_i = 1$. Now, let $\zeta = (\zeta_{ij})$ be any 1-cocycle. We do know there are smooth functions $f_i \in \Omega^0(U_i)$ such that $f_i - f_j = \zeta_{ij}$ on

[23] If you want to see this argument worked out in detail, search for the notion of a *good cover*, which is in between acyclic and having all intersections diffeomorphic to $\mathbb{R}^2$: a good cover is one for which all intersections are contractible.

U_{ij}. This is because the f_i don't have to be locally constant; in particular, we define

$$f_i = \sum_{U_j \in \mathfrak{U}} \rho_j \zeta_{ij}.$$

A priori, this is only defined on U_{ij}, since ζ_{ij} is only defined there, but since ρ_j smoothly extends to 0 outside U_j, we can just define $\rho_j \zeta_{ij}$ to be 0 outside U_j. In any case, on U_{ij},

$$\begin{aligned} f_i - f_j &= \sum_k \rho_k(\zeta_{ik} - \zeta_{jk}) \\ &= \sum_k \rho_k(\zeta_{ik} + \zeta_{kj}) \end{aligned}$$

by the cocycle condition. Then, since all the ρ_k sum to 1,

$$= \sum_k \rho_k \zeta_{ij} = \zeta_{ij}.$$

Now, let $\alpha_i = \mathrm{d}f_i$, which is an exact 1-form on U_i, and $\alpha_i - \alpha_j = \mathrm{d}(f_i - f_j) = \mathrm{d}\zeta_{ij} = 0$ on U_{ij}, so the α_i are restrictions of a closed $\alpha \in H^1(S)$, meaning $\zeta_{ij} = \check{c}(\alpha)$. ⊠

The injectivity of $\check{c}$ is what we meant by thinking of H^1 (the de Rham cohomology) as an obstruction to having a global primitive.

Dual to First Singular Homology. Another way of understanding $H^1(S)$ is as the dual to first singular homology (with integral coefficients), i.e. $H^1(S) = \mathrm{Hom}_{\mathsf{Ab}}(H_1(S), \mathbb{R})$. That is, for any loop $\gamma : S^1 \to S$ and a closed 1-form α, we have the integral

$$I(\gamma, \alpha) = \int_{S^1} \gamma^* \alpha.$$

$\gamma^*\alpha$ is a closed 1-form on $S^1 \cong \mathbb{R}/\mathbb{Z}$, so it has the form $f(t)\,\mathrm{d}t$ for some f, and hence we can explicitly integrate.

I is linear, and if $\alpha = \mathrm{d}g$ is exact, then

$$I(\gamma, \mathrm{d}g) = \int_{S^1} \gamma^* \,\mathrm{d}g = \int_{S^1} \mathrm{d}(\gamma^* g) = 0$$

by the fundamental theorem of calculus. Thus, $I(\gamma, \alpha)$ depends only on the class of α in $H^1(S)$.

Moreover, suppose we have a homotopy $\Gamma : S^1 \times [0,1] \to S$ between two loops γ_0 and γ_1. In this case,

$$\int_{S^1} \gamma_0^* \alpha - \int_{S^1} \gamma_1^* \alpha = \int_{S^1 \times I} \Gamma^*(\mathrm{d}\alpha),$$

but since α is closed, then this is 0.[24] In particular, $I(\gamma, \alpha)$ depends only on the free homotopy class of γ, which can be expressed as stating that I defines a bilinear map $I : H_1(S) \times H^1(S) \to \mathbb{R}$.[25] In other words, we have a homomorphism of abelian groups $H^1(S) \to \mathrm{Hom}_{\mathsf{Ab}}(H_1(S), \mathbb{R})$ (i.e. to homomorphisms of abelian groups), and since $H_1(S)$ is the abelianization of $\pi_1(S, *)$, this is also $\mathrm{Hom}_{\mathsf{Grp}}(\pi_1(S, *), \mathbb{R})$. (Here, $*$ is any point in S, regarded as the basepoint for the homotopy group.)

Theorem 14.3. *The integration map* $I : H^1(S) \to \mathrm{Hom}_{\mathsf{Ab}}(H_1(S), \mathbb{R})$ *is an isomorphism.*

This is another facet of de Rham's theorem, in the form conjectured by H. Cartan, and also proved by de Rham.

Proof idea. Let $H^1(S; \mathbb{R})$ denote $\mathrm{Hom}(H_1(S), \mathbb{R})$, the *first singular cohomology*. The construction used in the proof of Theorem 14.2 can be adapted to show that there's an isomorphism $H^1(S; \mathbb{R}) \to \check{H}^1(S; \mathfrak{U})$; then, one checks that the composition of that isomorphism with the one for de Rham cohomology is given by I.

[24] This calculation follows from Stokes' theorem on surfaces, in case it seems confusing.

[25] If the 1st homology group $H_1(S)$ isn't intuitive to you, then think of it this way: if S is connected, $H_1(S)$ is the abelianization of the fundamental group $\pi_1(S, *)$.

Lecture 15.

Poincaré Duality: 3/2/16

Today, we're going to finish our discussion of algebraic topology on surfaces by covering Poincaré duality, and then discuss what the complex structure does to vectors and covectors on Riemann surfaces.

Theorem 15.1 (Poincaré duality)**.** *Let S be a connected, oriented smooth surface. Then, the bilinear map $\langle \cdot, \cdot \rangle : H^1(S) \times H^1_c(S) \to \mathbb{R}$ defined by*

$$\langle [\alpha], [\beta] \rangle = \int_S \alpha \wedge \beta$$

is nondegenerate, so $H^1_c(S)$ is dual to $H^1(S)$.

To prove this completely would be perhaps too much of a detour from our course, as we need to set up the Mayer-Vietoris sequence for both H^* and H^*_c, and analyze what happens to the pairing under these sequences. For a reference for the complete proof, see Bott and Tu's book.

We will prove a weaker statement: that there is an injection $\star : H^1(S) \hookrightarrow \mathrm{Hom}(H^1_c(S), \mathbb{R})$ sending $[\alpha] \mapsto \langle [\alpha], \cdot \rangle$. In particular, when S is compact, the pairing is $\langle \cdot, \cdot \rangle : H^1(S) \times H^1(S) \to \mathbb{R}$, and is skew-symmetric. Since it's nondegnerate on one side, it's therefore nondegenerate on the other, so the injectivity of $\star$ implies it's an isomorphism.

First, though, why is this pairing well-defined? Let's compute it on an exact form $\mathrm{d}\alpha$ and a closed form β with compact support (so $\mathrm{d}\beta = 0$). Then,

$$\int_S \mathrm{d}\alpha \wedge \beta = \int_S \mathrm{d}(\alpha \wedge \beta) = 0,$$

by Stokes' theorem. The analogous proof works for $\langle \alpha, \mathrm{d}\beta \rangle$.

Proof. Suppose $[\alpha] \neq 0$ in $H^1(S)$; we'd like to find a closed $\beta \in \Omega^1_c(S)$ such that $\int_S \alpha \wedge \beta \neq 0$.

Since α represents a nonzero class in cohomology, there must be a loop $\gamma : \mathbb{R}/\mathbb{Z} \to S$ such that $\int_{\mathbb{R}/\mathbb{Z}} \gamma^* \alpha \neq 0$. Without loss of generality, one can assume that γ is an embedding; there are several ways to argue this, e.g. homotope γ to a self-transverse immersion, thanks to the standard transversality theory package. Now, replace this immersed loop with one of its circle components (and the integral of α must be nonzero on at least one of them), so it's now a piecewise smooth embedding, and then can be smoothed out with a homotopy.

Now, given an embedded loop γ, we'll construct a *Thom form* τ_γ, a closed 1-form supported in an annular neighborhood of γ such that for all closed 1-forms α on S,

$$\int_S \tau_\gamma \wedge \alpha = \int_{\mathbb{R}/\mathbb{Z}} \gamma^* \alpha. \tag{15.2}$$

That is, it will be a sort of dual to γ itself.

First, let's fix an annular neighborhood (that is, a tubular neighborhood) U for γ. This is a smooth embedding of a cylinder $\Gamma : (-1, 1) \times S^1 \to S$ with the property that $\gamma(0, t) = \gamma(t)$; let's take Γ to preserve orientation. It suffices to construct τ_γ on $(-1, 1) \times S^1$, and then embed it in S. We want to control the integral on the left side of (15.2), so we'd like to restrict τ_γ to be compactly supported.

Fix a smooth $\phi : (-1, 1) \to \mathbb{R}$ such that

- $\phi(s) = 1/2$ for s sufficiently close to 1, and
- $\phi(s) = -1/2$ for s sufficiently close to -1.

The precise shape of ϕ won't matter beyond this description.

Now, we can define $\tau = \phi^* \,\mathrm{d}s = \phi'(s)\,\mathrm{d}s$. If $\alpha \in \Omega^1(C)$ is such that $\mathrm{d}\alpha = 0$, then we have an inclusion map $i : S^1 \to C$ sending $t \mapsto (0, t)$ and a projection map $p : C \to S^1$ sending $(s, t) \mapsto t$; the homotopy invariance of de Rham cohomology implies that $\alpha - p^* i^* \alpha = \mathrm{d}f$ for some $f \in C^\infty(C)$.

We can check that τ_γ is closed:

$$\mathrm{d}\tau_\gamma = \frac{\partial}{\partial t} \phi'(s)\,\mathrm{d}t \wedge \mathrm{d}s = 0,$$

so by Stokes' theorem, since τ_γ is compactly supported,

$$\int_C \tau_\gamma \wedge \mathrm{d}f = \int_C \mathrm{d}(\tau_\gamma \wedge f) = 0.$$

In particular, $\int_C \tau_\gamma \wedge \alpha = \int_C \tau_\gamma \wedge \alpha'$, where $\alpha' = p^* i^* \alpha$, and we can compute this directly:

$$\begin{aligned}\int_C \tau_\gamma \wedge \alpha' &= \iint\limits_{(-1,1)\times S^1} \phi'(s) g(t)\,\mathrm{d}s\,\mathrm{d}t \\ &= \int_{-1}^{1} \phi'(s)\,\mathrm{d}s \int_{S'} g(t)\,\mathrm{d}t \\ &= \int_{S^1} i^*\alpha = \int_{S^1} \gamma^*\alpha.\end{aligned}$$

⊠

Poincaré duality generalizes to higher dimensions; the proof is harder, but the Thom form generalizes very nicely.

Corollary 15.3. *If γ_1 and γ_2 are embedded loops in S, then $\int_S \tau_{\gamma_1} \cdot \tau_{\gamma_2}$ is the signed intersection number of γ_1 and γ_2.*

Proof sketch. The construction given in the previous proof doesn't depend on the choice of tubular neighborhood, so let's use this freedom to choose tubular neighborhoods such that the curves look "standard" near intersection points, i.e. like coordinate axes (which entails making the curves transverse). If γ_1 is traveling rightwards, assign $+1$ if γ_2 is traveling up and -1 if it's traveling down; then, it's quick to check that this point contributes that value to $\int_S \tau_{\gamma_1} \wedge \tau_{\gamma_2}$. ⊠

Note. Let Σ_g be a compact, oriented surface of genus g. In this case, $H_1(\Sigma_g) \cong \mathbb{Z}^{2g}$, spanned by loops $a_1, \dots, a_g, b_1, \dots, b_g$, where a_i goes around the i^{th} hole and b_j loops from that hole to the edge (if $g = 1$, a slice of the donut). In this case, the intersection form is $a_i \cdot a_j = 0$, $b_j \cdot b_j = 0$, and $a_i \cdot b_j = \partial_{ij}$.

Thus, by Poincaré duality, $H^1(\Sigma_g) \cong \mathbb{R}^{2g}$, with a basis $\alpha_i = \tau_{a_i}$ and $\beta_j = \tau_{b_j}$. This is in fact a *symplectic basis* with respect to $\langle\cdot,\cdot\rangle$: $\int_{\Sigma_g} \alpha_i \wedge \alpha_j = 0$, $\int_{\Sigma_g} \beta_i \wedge \beta_j = 0$, and $\int_{\Sigma_g} \alpha_i \wedge \beta_j = -\int_{\Sigma_g} \beta_j \wedge \alpha_i = 0$.

Now, let's specialize to Riemann surfaces. We'll do a taste now, and then turn to elliptic curves; later on, we'll need more and develop more. There is a smooth, even linear map $i : \mathbb{C} \to \mathbb{C}$ that is multiplication by i. Its derivative $j = D_0 i : T_0\mathbb{C} \to T_0\mathbb{C}$ is just i again, under the natural identification of $T_0\mathbb{C}$ and $\mathbb{C}$.

If X is a Riemann surface, it has a *complex structure*, consisting of a $\mathbb{R}$-linear map $J_x : T_xX \to T_xX$ such that $J_x^2 = -\operatorname{id}$ for each $x \in X$ and varying smoothly in x. This structure is just multiplication by i on the tangent space (the action of j) in holomorphic coordinates. This is independent of coordinates, because the change-of-coordinates map is holomorphic, so its derivative is complex linear, and hence commutes with j. That is, a holomorphic atlas induces a smooth oriented atlas with a complex structure J.

One can also view J as a *conformal structure*; holomorphic maps $\mathbb{C} \to \mathbb{C}$ with nonvanishing derivatives are *conformal* (infinitesimally preserving angles), thanks to the Cauchy-Riemann equations. This allows us to measure the angle $\theta \in S^1$ between two nonzero tangent vectors $e_1, e_2 \in T_xX$, where X is a Riemann surface, because the holomorphic change-of-charts map will preserve it. In complex-structure terms, the angle between v and Jv is necessarily a right angle. That is, *holomorphic maps with nonzero derivative are exactly the angle-preserving maps.* From this perspective, a complex structure is exactly the same as a conformal structure, which is a way of measuring angles.

Another way to see this is that an oriented conformal structure on a smooth surface is the reduction of the structure group of its tangent bundle to $\mathrm{SO}(2) \times \mathbb{R}_+$, and a complex structure is the reduction of the structure group to $\mathrm{GL}_1(\mathbb{C}) = \mathbb{C}^*$. In higher dimensions, the isomorphism of these groups no longer holds.

Lecture 16.

Holomorphic 1-Forms: 3/4/16

Today, we'll discuss 1-forms on a Riemann surface, including holomorphic 1-forms on hyperelliptic Riemann surfaces, corresponding to §5.3 and §6.1 in the textbook. This leads into calculating a distinguished, nonvanishing 1-form on certain Riemann surfaces, and to a discussion of elliptic curves.

Last time, we mentioned that a Riemann surface X comes with a complex structure $J : TX \to TX$ which is rotation by 90° (multiplication by i) on each tangent space. In particular, above a point $p \in X$, J_p is linear, and $J_p^2 = -\operatorname{id}$. We constructed it as the derivative of the map $z \mapsto iz$ in local coordinates.

There is an induced map J_p^* on the cotangent space $T_p^*X = \operatorname{Hom}(T_pX, \mathbb{R})$. Now, we can consider the complexified cotangent space $(T_p^*X) \otimes \mathbb{C}$, which is a two-dimensional complex vector space, and can be thought of as $\operatorname{Hom}_{\mathbb{R}}(T_pX, \mathbb{C})$. Since $(J^*)^2 = -\operatorname{id}$, we can decompose the complexified cotangent space into the i- and $-i$-eigenspaces for J^*: $T_p^*X \otimes \mathbb{C} = T_p^*X' \oplus T_p^*X''$, where J^* acts as $+i$ on T_p^*X' and $-i$ on T_p^*X''. Thus, $T_p^*X' = \operatorname{Hom}_{\mathbb{C}}(T_pX, \mathbb{C})$, the $\mathbb{C}$-linear maps, and T_p^*X'' is the space of $\mathbb{C}$-antilinear maps.

Similarly, we can decompose the complex-valued 1-forms (so letting p vary) $\Omega^1(X)_{\mathbb{C}} = \Omega^1(X) \otimes C^\infty(X, \mathbb{C})$ (the tensor is over $C^\infty(X, \mathbb{R})$; $\Omega^1(X)$ is a module over this ring). Hence, an $\alpha \in \Omega^1(X)_{\mathbb{C}}$ locally looks like $f\,\mathrm{d}x + g\,\mathrm{d}y$, where f and g are $\mathbb{C}$-valued smooth functions, so $\alpha_p \in T_p^*X \otimes \mathbb{C}$. Now, we apply the decomposition into $\pm i$-eigenspaces to get that $\Omega^1(X)_{\mathbb{C}} = \Omega^{1,0}(X) \oplus \Omega^{0,1}(X)$, where $\Omega^{1,0}(X)$ is valued in T_p^*X' and $\Omega^{0,1}(X)$ is valued in T_p^*X''. Explicitly, a complex-valued 1-form α splits as $\alpha = \alpha^{1,0} + \alpha^{0,1}$, where if v is a tangent vector we have

$$\begin{aligned}
\alpha^{1,0}(v) &= \frac{1}{2}(\alpha(v) - i\alpha(Jv)) \\
\alpha^{0,1}(v) &= \frac{1}{2}(\alpha(v) + i\alpha(Jv)).
\end{aligned}$$

It's quick to check that the first one is $\mathbb{C}$-linear, the second is $\mathbb{C}$-antilinear, and they sum to $\alpha(v)$.

If we let $\mathrm{d}z = \mathrm{d}x + i\,\mathrm{d}y$ and $\mathrm{d}\bar{z} = \mathrm{d}x - i\,\mathrm{d}y$, then a $(1,0)$-form locally looks like $f(x,y)\,\mathrm{d}z$ and a $(0,1)$-form locally looks like $g(x,y)\,\mathrm{d}\bar{z}$. In particular, if f is a holomorphic function on X, then $\mathrm{d}f \in \Omega^{1,0}(X)$: $\mathrm{d}f = \frac{\partial f}{\partial z}\,\mathrm{d}z$, because $\frac{\partial}{\partial \bar{z}}$ is 0 for any holomorphic function.

The exterior derivative also splits. Akin to the exterior derivative on real differential forms, we have $\Omega^0(X)_{\mathbb{C}} = C^\infty(X, \mathbb{C})$ and $\Omega^2(X)_{\mathbb{C}} = \Omega^2(X) \otimes C^\infty(X, \mathbb{C})$. Then, d extends $\mathbb{C}$-linearly, meaning our complex de Rham complex is the sequence

$$\Omega^0(X)_{\mathbb{C}} \xrightarrow{\mathrm{d}} \Omega^1(X)_{\mathbb{C}} \xrightarrow{\mathrm{d}} \Omega^2(X)_{\mathbb{C}}.$$

Now, when we split $\Omega^1(X)_{\mathbb{C}}$, each arrow also splits into two.

$$\begin{tikzcd}
 & \Omega^{1,0}(X) \arrow[dr, "\overline{\partial}"] & \\
\Omega^0(X)_{\mathbb{C}} \arrow[ur, "\partial"] \arrow[dr, "\overline{\partial}"'] & & \Omega^2(X)_{\mathbb{C}} \\
 & \Omega^{0,1}(X) \arrow[ur, "\partial"'] &
\end{tikzcd}$$

For $\mathrm{d} : \Omega^0(X)_{\mathbb{C}} \to \Omega^1(X)_{\mathbb{C}}$, the splitting is $\mathrm{d} = \partial + \overline{\partial}$, so for a smooth f, in local coordinates we have

$$\begin{aligned}
\mathrm{d}f &= \partial f + \overline{\partial} f \\
&= \frac{1}{2}(f_x - i f_y)\,\mathrm{d}z + \frac{1}{2}(f_x + i f_y)\,\mathrm{d}\bar{z} \\
&= \frac{\partial f}{\partial z}\,\mathrm{d}z + \frac{\partial f}{\partial \bar{z}}\,\mathrm{d}\bar{z}.
\end{aligned}$$

For $\mathrm{d} : \Omega^1(X)_{\mathbb{C}} \to \Omega^2(X)_{\mathbb{C}}$, we let $\overline{\partial} = \mathrm{d}|_{\Omega^{1,0}}$ and $\partial = \mathrm{d}|_{\Omega^{0,1}}$; if $A\,\mathrm{d}\bar{z} \in \Omega^{0,1}(X)$, then

$$\partial(A\,\mathrm{d}\bar{z}) = \frac{\partial A}{\partial z}\,\mathrm{d}z \wedge \mathrm{d}\bar{z} = 2i\frac{\partial A}{\partial z}\,\mathrm{d}x \wedge \mathrm{d}y.$$

Correspondingly, if $B\,\mathrm{d}z \in \Omega^{1,0}$, then

$$\overline{\partial}(B\,\mathrm{d}z) = \frac{\partial B}{\partial \bar{z}}\,\mathrm{d}\bar{z} \wedge \mathrm{d}z = -2i\frac{\partial B}{\partial \bar{z}}\,\mathrm{d}x \wedge \mathrm{d}y.$$

With this $\mathbb{C}$-linear algebra in mind, we have the following definition.

Definition. A *holomorphic* 1*-form* on X is an $\alpha \in \Omega^{1,0}(X)$ such that $\overline{\partial}\alpha = 0$.

That is, it's closed and in $\Omega^{1,0}$. Such a form locally looks like $\alpha = A\,\mathrm{d}z$, where A is a holomorphic function.

Using this, one can define the *Laplacian* $\Delta : \Omega^0(X) \to \Omega^2(X)$, sending real functions to real 2-forms, given by the formula

$$\Delta = 2i\overline{\partial}\partial = -2i\partial\overline{\partial}.$$

In local coordinates, $\Delta f = -(f_{xx} + f_{yy})\,\mathrm{d}x \wedge \mathrm{d}y$.[26]

Definition. A smooth ϕ is *harmonic* if $\Delta\phi = 0$.

Lemma 16.1. *If ϕ is harmonic, then locally, $\phi = \mathrm{Re}(f)$, where f is holomorphic.*

This is a classical theorem in complex analysis, albeit in a slightly different context.

Proof. You can prove this using the Poincaré lemma — suppose $\Delta\phi = 0$, and let $A = 2\,\mathrm{Re}(i\overline{\partial}\phi) = i\overline{\partial}\phi + \overline{i\overline{\partial}\phi}$. Thus, $\mathrm{d}A = 2\,\mathrm{Re}(\mathrm{d}i\overline{\partial}\phi) = 2\,\mathrm{Re}\,i(\partial\overline{\partial}\phi) = 0$, since ϕ is harmonic.

Since A is closed, then it's locally exact, so $A = \mathrm{d}\psi$ for some real-valued function ψ. Hence, $\partial\psi = A^{1,0} = -i\partial\phi$ and $\overline{\partial}\psi = A^{0,1} = i\overline{\partial}\phi$. Thus, $\overline{\partial}(\phi + i\psi) = \overline{\partial}\phi + i\overline{\partial}\psi = \overline{\partial}\phi - \overline{\partial}\phi = 0$, so this function is holomorphic. ⊠

This ψ is sometimes called the *harmonic conjugate* of ϕ. The proof was really the same as the standard construction using the Cauchy-Riemann equations, though draped in different clothes.

From this, we immediately get the maximum principle.

Corollary 16.2 (Maximum principle)**.** *If a harmonic function ϕ has a local maximum near p, then it's constant near p.*

Proof. Let $\phi = \mathrm{Re}(f)$, for f holomorphic, and apply the open mapping theorem. ⊠

Again, this is the same proof as usual, just in a different guise.

Hyperelliptic surfaces.

Definition. A compact, connected Riemann surface Z is called *hyperelliptic* if there's a degree-2 map $f : Z \to S^2$.

Since a proper map is determined by its branching data, Z is identified with the Riemann surface $X = \{(z, w) \in \mathbb{C}^2 \mid w^2 = f(z)\}$, where the roots of f (the critical values) are all distinct (hence it really is a Riemann surface: repeated roots correspond to both partial derivatives vanishing). We haven't looked at ∞ yet, which will be important, but the reason this works is that f is a branched double cover for S^2, and $w^2 = f(z)$ is a branched double cover, with the covering map projection onto z.

If C is a circle centered at the origin and large enough to contain all the roots of f in its interior, then the monodromy around C is the same as the composite of the monodromies around each root, since if we removed the roots, the two loops would be homotopic. Hence, the monodromy is trivial if there are an even number of roots, and nontrivial if there's an odd number, which allows us to understand the monodromy at infinity. Let n be the number of roots.

In particular, if X^* denotes the *compactification* of the algebraic curve X (i.e. the normalization of the projective closure of X in $\mathbb{C}P^2$), then if n is odd, then we just need to ad one branch point lying over ∞, and so $X^* = X \cup \{P\}$. If n is even, we need to add two points over ∞, and so it's not a branch point.

Next time, we'll construct a holomorphic 1-form α on X and see that it extends meromorphically to X^*; if $n \le 4$, this extension is holomorphic. This will lead to a discussion of Riemann surfaces that have a nowhere-vanishing holomorphic one-form.

Lecture 17.

Hyperelliptic Riemann Surfaces: 3/7/16

Today, we're going to discuss hyperelliptic Riemann surfaces and a special case of then, elliptic curves. This corresponds to §6.1 of the book.

Recall that a hyperelliptic surface is a compact, connected Riemann surface Z that admits a degree-2 map $F : Z \to S^2$. Equivalently, there is a holomorphic action of the group $\mathbb{Z}/2$ on Z, and $Z/(\mathbb{Z}/2)$ is biholomorphic

[26] This is the so-called "geometer's Laplacian;" the "analyst's Laplacian" has no minus sign. In Riemannian geometry, this choice of the Laplacian is very natural.

to S^2 (which is nice because this is a branched double cover; the branched points are the fixed points of the action).

One concrete realization of such a hyperelliptic surface is as the compactification X^* of an algebraic curve $X = \{(z, w) \mid w^2 = f(z)\}$, where f has no repeated roots; then, the map F sends $(z, w) \mapsto z$, so the critical values are the roots of f, and possibly also ∞. Concretely, if $n = \deg(f)$ is odd, the $X^* = X \cup \{P\}$, and if n is even, $X^* = X \cup \{P_1, P_2\}$: we know the inclusion $X \hookrightarrow \mathbb{C}P^2$ sending $(z, w) \mapsto [z : w : 1]$ extends to a holomorphic map $X^* \to \mathbb{C}P^2$, so we just need to know what exists over $[1 : 0] = \infty$. If n is odd, $P \in \mathrm{crit}(F)$, and ∞ is a critical value; if n is even, them $P_\pm \notin \operatorname{crit} F$, and ∞ isn't a critical value. Hence, there's always an even number of critical values of F.

Topologically, imagine that you have a genus-g surface, and skewer it through its holes so that the skewer intersects the surface at $2g - 2$ points. Then, rotating by $180°$ is a $\mathbb{Z}/2$-action and these intersections are the fixed points of that action. This is in fact that model for the $\mathbb{Z}/2$-action on a hyperelliptic surface, and the quotient is S^2, but this is not easy to see. One way to think about it is to consider the map $f : \overline{\mathbb{D}} \to \overline{\mathbb{D}}$ sending $z \mapsto z^2$. If $\mathbb{Z}/2$ acts on $\overline{\mathbb{D}}$ by sending $z \mapsto -z$, then f is $\mathbb{Z}/2$-equivariant, and $\overline{\mathbb{D}} = \overline{\mathbb{D}}/(\mathbb{Z}/2)$. Then, if we excise a disc in the codomain, we have to remove two discs in the domain, as in Figure 7.

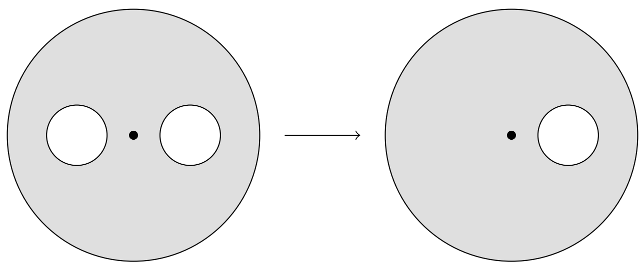


FIGURE 7. Visualizing the $\mathbb{Z}/2$-action of a hyperelliptic surface.

Now, we can redraw this as a pair of pants mapping to the cylinder by a quotient, and the $\mathbb{Z}/2$-action becomes rotating around the skewer in the middle. Now, if we stick several of these back-to-back, the more general case of $2g - 2$ branch points can be realized in this way, and capping off the cylinder gives you S^2.

We'd like to relate these to holomorphic 1-forms, and the goal will be to show that if there are at most 4 branch points, there exists a nowhere-vanishing holomorphic 1-form. Specifically, on $X = \{(z, w) \mid w^2 = f(z)\} \subset \mathbb{C}^2$, the form $\alpha = \mathrm{d}z/w$ is holomorphic for certain values of n.

A priori this looks meromorphic, rather than holomorphic, but on X, $2w\,\mathrm{d}w = f'(z)\,\mathrm{d}z$ as complex 1-forms, so $\mathrm{d}z/w = z/f'(z)\,\mathrm{d}w$. When $w = 0$, we know $f(z) = 0$, but therefore $f'(z) \neq 0$ (since the roots of f are distinct), and therefore there's not actually a pole at $w = 0$.

To say that α is holomorphic is to say that $\overline{\partial}\alpha = 0$, i.e. locally, $\alpha = \varphi(z)\,\mathrm{d}z$, where φ is holomorphic. This is not a problem except maybe at ∞. If n is odd, then there is a holomorphic coordinate near $P \in X^*$ such that F takes the form $F(\tau) = \tau^2$, given by $\tau^2 = 1/z$. Hence, we can directly substitute

$$f(z) = z^n + a_1 z^{n-1} + \cdots + a_n = \tau^{-2n}\bigl(1 + a_1\tau^2 + \cdots + a_n\tau^{2n}\bigr)$$

and $w = \pm\sqrt{f(z)} = \pm\tau^{-n}\sqrt{1 + a_1\tau^2 + \cdots}$, and this square root is well-deifned and holomorphic for τ in a neighborhood of 0.

The point is, $\alpha = \mathrm{d}z/w$, so

$$\begin{aligned}\alpha = \frac{\mathrm{d}z}{w} &= \frac{\tau^n}{\sqrt{1 + \alpha_1\tau^2 + \cdots}}\left(-\frac{2}{\tau^3}\right)\mathrm{d}\tau \\ &= \frac{-2\tau^{n-3}}{\sqrt{1 + a_1\tau^2 + \cdots}}\,\mathrm{d}\tau.\end{aligned}$$

Hence, at P, where $\tau = 0$:

- α has a double pole if $n = 1$;
- α is holomorphic and nonvanihsing if $n = 3$; and
- α has a 0 of order $n - 3$ if $n > 3$.

For $n = 1$, if you defined a "meromorphic 1-form" to be one that locally looks like a meromorphic function, then that's what we get; for $n = 3$, the most interesting case, α is a *holomorphic volume form* on X^*, i.e. a nowhere-vanishing holomorphic 1-form.

A similar analysis for n even shows that when $n = 4$, we once again get that α is a holomorphic volume form.

Note. Like regular volume forms, holomorphic volume forms are unique up to scaling: if Z is a compact Riemann surface and α_1 and α_2 are holomorphic volume forms, then there's a $c \in \mathbb{C}^*$ such that $\alpha_2 = c\alpha_1$; this is because $\alpha_2 = f\alpha_1$ for some holomorphic $f : Z \to \mathbb{C}^*$, but since Z is compact, then f must be constant (by the maximum modulus theorem).

Elliptic Curves.

Definition.

- An *elliptic curve* (E, p) is a compact, connected Riemann surface E together with a point $p \in E$ such that E admits a holomorphic volume form.
- An *isomorphism of elliptic curves* $f : (E_1, p_1) \to (E_2, p_2)$ is a biholomorphic map $f : E_1 \to E_2$ that maps $p_1 \mapsto p_2$.[27]

Theorem 17.1. *If E is a compact, connected Riemann surface and $p \in E$, then the following are equivalent.*

(1) *(E, p) is an elliptic curve.*
(2) *$(E, p) \cong (\mathbb{C}/\Lambda, [0])$ (meaning p is the equivalence class of 0) for a lattice $\Lambda \subset \mathbb{C}$.*
(3) *E embeds into $\mathbb{C}P^2$ as a* nonsingular Weierstrass cubic*; that is, it's the projective closure of an equation $w^2 = (z - a_1)(z - a_2)(z - a_3)$ for distinct a_1, a_2, a_3, where p is the point at infinity.*
(4) *There exists a branched cover $F : E \to S^2$ of degree 2 branched over exactly four points, where p is one of the critical points.*

Remark.

- Condition (2) means that $\mathrm{Aut}(E)$, the biholomorphic maps $E \to E$, acts transitively on E, and are induced from translations of $\mathbb{C}$. Hence, all choices of p are equally good.
- Nonetheless, the choice of p will be useful when we classify elliptic curves, and later, using the Riemann-Roch theorem, we'll introduce a group law on an elliptic curve, and it will be helpful to have p around.
- Later, there will be another equivalent condition, that E has genus 1 (so topologically a torus); right now, we can see that (2) implies this, and later will show that it implies (3).

We've proven (4) implies (1) by the discussion earlier this lecture, so our challenge is to run the rest of the proof. The idea for (1) implying (2) is that we need to construct a lattice given an elliptic curve; we'd like to do this so that if $q : \mathbb{C} \twoheadrightarrow \mathbb{C}/\Lambda$ is the projection map and β is the holomorphic 1-form on $\mathbb{C}/\Lambda$ such that $q^*\beta = \mathrm{d}z$, then we'd like the isomorphism $F : E \to \mathbb{C}/\Lambda$ to set $F^*\beta = \alpha$.

If we knew $H_1(E) = \mathbb{Z}^2 = \langle a, b\rangle$, we could do this more explicitly: let Λ be the *period lattice* of α, i.e. the points

$$\Lambda = \left\{ \int_\gamma \alpha \mid \gamma \text{ is a 1-cycle in } E \right\}.$$

Since a and b generate $H_1(E)$, this means $\Lambda = \mathbb{Z}\int_a \alpha + \mathbb{Z}\int_b \alpha$, so we really get a lattice; then, we can define $F : E \to \mathbb{C}/\Lambda$ by sending $x \mapsto \int_p^x \alpha \bmod \Lambda$ (this is not well-defined, but does define a unique point in $\mathbb{C}/\Lambda$).

Since we don't yet know $H_1(E) = \mathbb{Z}^2$ (even though *a posteriori* we will discover this), we will have to work a little harder to make an analogous construction.

Lecture 18.

Elliptic Curves and Elliptic Functions: 3/9/16

Today, we're going to make some headway on Theorem 17.1; specifically, we'll prove (1) implies (2), and then prove that every elliptic curve embeds as a Weierstrass cubuc into $\mathbb{C}^2$, which uses elliptic functions, which we'll also discuss.

[27] This is not required to do anything special to the volume form; this means that the different scaled versions of a different volume give you isomorphic elliptic curves.

Let (E, p) be an elliptic curve and α be a choice of holomorphic volume form. We have the abelian group of *periods*

$$\Lambda_{E,\alpha} = \left\{ \int_\gamma \alpha \mid \gamma \text{ is a 1-cycle in } E \right\}.$$

This is a subgroup of $\mathbb{C}$, and we can say a lot about this.

Theorem 18.1. *With (E, p) and α as above,*

(1) *$\Lambda_{E,\alpha}$ is a lattice in $\mathbb{C}$;*
(2) *there is a unique biholomorphic map $F : E \to \mathbb{C}/\Lambda_{E,\alpha}$ sending $p \mapsto [0]$ and such that $F^*(\mathrm{d}z) = \alpha$; and*
(3) *if Λ_1 and Λ_2 are lattices in $\mathbb{C}$ and $\phi : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$ is a biholomorphic map such that $\phi^*[0] = [0]$, then $\phi^*[z] = [cz]$, where $c \in \mathbb{C}^*$ is such that $c(\Lambda_1) = \Lambda_2$.*

Proving this inevitably involves working with some technical issues; the textbook tries to deal with these minimally to avoid muddling the proof idea, but it's not too hard to deal with them head-on.

Proof. First, we'll attack (1). We must prove that $\Lambda = \Lambda_{E,\alpha}$ is a lattice. We have a real, nowhere-vanishing 1-form given by $\beta = \alpha + \overline{\alpha}$; since α locally takes the form $\alpha = f(z)\,\mathrm{d}z$, then β locally takes the form $\beta = f\,\mathrm{d}z + \overline{f}\,\mathrm{d}\bar{z}$ (so their zeros would be the same).

The Poincaré-Hopf theorem[28] states that the Euler characteristic is given by the number of times α vanishes, so $\chi(E) = 2 - 2g = 0$, and since $2g = \dim H^1(E)$, then $g = 1$, and $H_1(E; \mathbb{Z}) = \mathbb{Z}^2 = \mathbb{Z} \cdot a \oplus \mathbb{Z} \cdot b$.

Thus, $\Lambda = \mathbb{Z} \int_a \alpha + \mathbb{Z} \int_b \alpha$. Thus, to show Λ is a lattice, it suffices to show that $\int_a \alpha$ and $\int_b \alpha$ are $\mathbb{R}$-linearly independent. So suppose they're linearly dependent, so after rescaling α, $\Lambda \subset \mathbb{R}$. Then, we can define $\phi : E \to \mathbb{R}$ by $\phi(x) = \mathrm{Im} \int_p^x \alpha$, which is independent of path. And ϕ is harmonic, because locally, it's the imaginary part of a holomorphic function, so by the maximum principle, ϕ is constant! Since $\phi(p) = 0$, then $\phi = 0$. However, $\mathrm{d}\phi = \mathrm{Im}(\alpha)$, which by the open mapping theorem is not identically zero, which is a contradiction. Hence, Λ is a lattice, proving (1).

Next, for (2), define the *period map* or *Albanese map* $F : E \to \mathbb{C}/\Lambda$ by

$$x \longmapsto \int_p^x \alpha \bmod \Lambda.$$

Clearly, $F(p) = 0$, and by construction, $F^*(\mathrm{d}z) = \alpha$. Now, we would like to show that F is a biholomorphism. Since $\mathrm{d}F = \alpha$ is nonvanishing, then F is locally biholomorphic. It's also proper, because E is compact, and as such, it's a covering map.

All finite coverings of $\mathbb{C}/\Lambda$ take the form $\mathbb{C}/L$ where $L \subset \Lambda$ is a lattice, so (E, p) is identified with $(\mathbb{C}/L, [0])$, and α is sent to $\mathrm{d}z$, which is almost exactly what we wanted — but we don't know that E is the period lattice yet. However, the period lattice of $(\mathbb{C}/L, \mathrm{d}z)$ is exactly L, so the period lattice of E has to be L, i.e. $L = \Lambda$.

For (3), suppose we have such a ϕ. Then, $\phi^*(\mathrm{d}z) = c\,\mathrm{d}z$ for a $c \in \mathbb{C}^*$, because these are the only nonvanishing holomorphic 1-forms (so any two differ by a constant). In particular, $\phi'(z) = c$ and $\phi(0) = 0$, so $\phi(z) = cz$. ⊠

Now, we will talk about elliptic functions. The goal is to show that $\mathbb{C}/\Lambda$ is (biholomorphic to) the projective closure in $\mathbb{C}P^2$ of $\{w^2 = z^3 + az + b\} \subset \mathbb{C}^2$, where the cubic has distinct roots and the biholomorpism sends 0 to the unique point at infinity.

Definition. An *elliptic function* for a lattice $\Lambda \subset \mathbb{C}$ is a meromorphic f on $\mathbb{C}$ with the property that $f(z + \lambda) = f(z)$ for all $\lambda \in \Lambda$ and $z \in \mathbb{C}$.

In other words, this function is translation-invariant by the lattice, so it descends to a meromorphic function on the quotient torus $\mathbb{C}/\Lambda$.

Lemma 18.2. *A holomorphic elliptic function is constant.*

[28] We'll discuss this theorem, and the Euler characteristic in general, in a few lectures; its proof does not depend on anything we do today.

This is by Liouville's theorem (it's defined just on the compact torus, and hence bounded and holomorphic); indeed, this is quite possibly why Liouville proved this theorem. Alternatively, you can use the fact that any holomorphic function on a compact Riemann surface is constant, thanks to the maximum principle.

The most important elliptic function is the Weierstrass $\wp$-function.[29]

Definition. The *Weierstrass $\wp$-function* $\wp = \wp_\Lambda$ is defined by

$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda\in\Lambda\setminus 0}\left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2}\right).$$

One also defines the *Eisenstein series*

$$G_{2k}(\Lambda) = \sum_{\lambda\in\Lambda\setminus 0}\lambda^{-2k}.$$

We have a bunch of facts about these, which are somewhat straightforward estimates that are left as exercises. Alternatively, consult any text on elliptic functions.

Fact.

- The series defining G_{2k} converges absolutely for $k > 1$.
- The series defining $\wp$ converges absolutely and uniformly on compact subsets of $\mathbb{C}/\Lambda$, so $\wp$ is meromorphic with double poles at each $\lambda \in \Lambda$, and is holomorphic on $\mathbb{C} \setminus \Lambda$.
- $\wp$ is an elliptic function for Λ.
- $\wp$ is an even function: $\wp(-z) = \wp(z)$.

The reason we care about $\wp$ is the following theorem.

Theorem 18.3. *Any elliptic function for Λ is a rational function in $\wp$ and $\wp'$.*

We can also write down the Laurent expansion for $\wp$ around $z = 0$. If $|z| < |\lambda|$, then

$$\begin{aligned}\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} &= \frac{1}{\lambda^2}\left(\frac{1}{(1-z/\lambda)^2} - 1\right)\\ &= \sum_{n\geq 1}(n+1)\frac{z^n}{\lambda^{n+2}}.\end{aligned}$$

Hence, if we take $|z| \leq \min\{|\lambda| \mid \lambda \in \Lambda\}$, then

$$\begin{aligned}\wp(z) &= \frac{1}{z^2} + \sum_{\lambda\in\Lambda\setminus 0}\sum_{n\geq 1}(n+1)\frac{z^n}{\lambda^{n+2}}\\ &= \frac{1}{z^2} + \sum_{n\geq 1}\left((n+1)\sum_{\lambda\in\Lambda\setminus 0}\lambda^{-n-2}\right)z^n.\end{aligned}$$

Since $\wp$ is even, it only has even powers in its Laurent series, so

$$\wp(z) = \frac{1}{z^2} + \sum_{n\geq 1}(2n+1)G_{2n+2}(\Lambda)z^{2n}. \tag{18.4}$$

Lecture 19.

The *j*-invariant and Moduli of Elliptic Curves: 3/11/16

"There are going to be some constants in this picture, and you're going to think they look stupid. But I promise you they're not."

[29]It's so important that it gets its own typeface: it's an old German p, but not the Gothic $\mathfrak{p}$! In LaTeX, you should use `\wp` to typeset $\wp$.

Today we're going to finish the proof that $\mathbb{C}$ mod a lattice is a cubic curve, and then discuss the moduli space of elliptic curves and the j-invariant; this could be a whole subject, and makes contact with the world of modular forms.

Recall that last time, we took a lattice $\Lambda \subset \mathbb{C}$ and produced a Λ-invariant meromorphic function $\wp : \mathbb{C} \to S^2$ (or $\mathbb{C}/\Lambda \to S^2$) defined by

$$\wp(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda \setminus 0} \left(\frac{1}{(z-\lambda)^2} - \frac{1}{\lambda^2} \right).$$

We also calculated its Laurent series in (18.4) (recall that $G_{2k}(\Lambda) = \sum_\lambda \lambda^{-2k}$). We'll now derive a differential equation for $\wp$. From (18.4),

$$\begin{aligned} \wp'(z) &= -\frac{2}{z^3} + \sum_{n \geq 1} 2n(2n+1) G_{2n+2} z^{2n-1} \\ &= -\frac{2}{z^3} + 6G_4 z + 20 G_6 z^3 + O(z^5). \end{aligned}$$

Therefore

$$\wp'(z)^2 = \frac{4}{z^6} - \frac{24G_4}{z^2} - 80G_6 + O(z^2). \tag{19.1}$$

Compare this to

$$\wp(z) = \frac{1}{z^2} + 3G_4 z^2 + 5G_6 z^4 + O(z^6) \tag{19.2}$$

$$\wp(z)^3 = \frac{1}{z^6} + \frac{9G_4}{z^2} + 15G_4 + O(z^6). \tag{19.3}$$

What we would like to do is combine (19.1), (19.2), and (19.3) somehow in a way that eliminates all of their poles; since these are elliptic functions, such a combination would necessarily be constant, by Lemma 18.2. By inspection of the first few terms, we should take

$$\wp'(z)^2 - 4\wp(z)^3 + 60G_4\wp(z) + 140G_4 = O(z^2).$$

The left-hand side is elliptic, holomorphic on $\mathbb{C}$, and vanishes at 0, so it must be identically 0. Thus, if we let $g_2 = 60G_4$ and $g_3 = 140G_6$, $\wp$ satisfies the equation

$$(\wp')^2 = 4\wp^3 - g_2\wp - g_3. \tag{19.4}$$

All told, it really could have been worse; these estimates weren't so bad.

We'll use this to map $\mathbb{C}/\Lambda$ to a cubic curve by the map $\theta : \mathbb{C}/\Lambda \to \mathbb{C}P^2$ that sends $[z] \mapsto [\wp(z) : \wp'(z) : 1]$ (for $z \notin \Lambda$). Near $z = 0$, $\theta(z) = [z^3\wp(z) : z^3\wp'(z) : z^3]$, and therefore approaches $[0 : 1 : 0]$. Hence, we let $\theta([0]) = [0 : 1 : 0]$. This means that the image of θ is $\mathrm{Im}(\theta) = \overline{X} = \{y^2 z = 4x^3 - g_2xz^2 - g_3z^3\} \subset \mathbb{C}P^2$.

Claim. The polynomial $p(x) = 4x^2 - g_2x - g_3$ has three distinct roots, and hence $\overline{X}$ is a Riemann surface and the closure of $\{y^2 = p(x)\}$.

The last part does require smoothness at $[0 : 1 : 0]$, but we've already checked that.

Proof. Returning to the differential equation (19.4), if $4\wp(u)^3 - g_2\wp(u) - \wp(u) = 0$, then $\wp'(u) = 0$, which seems like something we can work out. In fact, $\wp'$ is odd, so if $\lambda \in (1/2)\Lambda$, then $\wp'(\lambda) = -\wp'(-\lambda) = -\wp'(\lambda)$, so it must be either 0 or ∞. We know $(\wp')^{-1}(\infty) \in \Lambda$, so all of the zeroes of $\wp'$ are in $(1/2)\Lambda \setminus \Lambda$, and in the fundamental domain there are three strict half-lattice points. In particular, if λ_1 and λ_2 are independent lattice vectors, the zeros of $\wp'$ are $\wp(\lambda_1/2)$, $\wp(\lambda_2/2)$, and $\wp((\lambda_1+\lambda_2)/2)$.

Are these still distinct after we've hit them with $\wp$? Yes, because $\wp - \wp(\lambda_1/2)$ is a degree-2, even function with a double zero at $\lambda_1/2<$ so it has no other zeroes. Thus, $\wp(\lambda_1/2) \neq \wp(\lambda_2/2)$ and is also different from $\wp((\lambda_1+\lambda_2)/2)$; the same argument works for $\lambda_2/2$. ⊠

In particular, $\overline{X}$ is a Riemann surface. θ is non-constant, hence surjective by the open mapping theorem; since $\theta^{-1}([0 : 1 : 0]) = \{[0]\}$, meaning it has multiplicity 1, then $\deg\theta = 1$, so it's a biholomorphism. Thus, $\mathbb{C}/\Lambda \cong \overline{X}$.

What does this to do an isomorphism $f : \mathbb{C}/\Lambda_1 \to \mathbb{C}/\Lambda_2$? Last time, we saw this is of the form $f([z]) = [cz]$, where $c \in \mathbb{C}^*$ and $c\Lambda_1 = \Lambda_2$. Passing from Λ_1 to $\Lambda_2 = c\Lambda_1$ is akin to rescaling the coefficients of the cubic:

$g_2(c\Lambda_1) = c^{-4}g_2(\Lambda_1)$ and $g_3(c\Lambda_1) = c^{-6}g_3(\Lambda_1)$, and so we get an isomorphism $\overline{X}_1 \to \overline{X}_2$ by a change of variables $x' = c^{-2}x$ and $y' = c^{-3}y$.

This concludes our proof of the equivalences of the various realizations of elliptic curves.

Classifying Elliptic Curves. We would like to set up and study a *coarse*[30] *moduli space* $\mathcal{M}$ parameterizing elliptic curves up to isomorphism. The various models we have of elliptic curves all work to do this.

Let's start with the Weierstrass cubic, for which an elliptic curve is represented as (the projective closure of) $y^2 = x^3 + ax + b$. The right-hand side of this has distinct roots, so its discriminant δ is nonzero; if the roots are $\{e_1, e_2, e_3\}$, then $\delta = (e_1 - e_2)^2(e_2 - e_3)^2(e_3 - e_1)^2 = 4a^3 + 27b^2$, which comes to us from Galois theory.

First, let $\Delta = -16\delta$. Then, our moduli space is $\mathcal{M} = (\mathbb{C}^2 \setminus \Delta^{-1}(0))/\mathbb{C}^*$, where $\mathbb{C}^*$ acts by $t \cdot (a, b) = (t^{-4}a, t^{-6}b)$. Thus, $\mathcal{M}$ is in bijection with isomorphism classes of elliptic curves, by the preceding discussion. This is a one-dimensional complex manifold, or in other words, a Riemann surface!

Definition. Define the *j-function* (or *j-invariant*) $j : \mathcal{M} \to \mathbb{C}$ by $j(\{y^2 = x^3 + ax + b\}) = -12^3a^3/\Delta(a, b)$.

We can divide out by $\Delta(a, b)$ because it never vanishes. Thus, this is a pretty transparent construction of the j-invariant, if puzzling: why -1728?

Proposition 19.5. *j is a biholomorphic map.*

Well, that's cool. Elliptic curves, up to isomorphism, are parameterized (coarsely) by $\mathbb{C}$! We won't prove this, but it's a direct check of injectivity and surjectivity: you can use the $\mathbb{C}^*$-action to work less hard.

But we can also express j in terms of lattices: if Λ is a lattice, $\mathbb{C}/\Lambda \cong \{y^2 = x^3 - g_2(\Lambda)x/4 - g_3(\Lambda)/4\}$, so $a^3 = -g_2^3/4^3$ and $\Delta = -16(-g_2^3/4^2 + (27/4)g_3^2)$. In particular, the j-invariant of the lattice Λ is

$$j(\Lambda) = 27\left(\frac{g_2^3}{27g_3^2 - g_3^2}\right).$$

Any Λ can be rescaled to one of the form $\Lambda = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau$, where $\tau \in \mathbb{H}$; τ isn't uniquely determined by Λ, but only up to the action of the modular group $\mathrm{PSL}_2(\mathbb{Z})$ acting on $\mathbb{H}$ by Möbius maps. In other words, this lattice doesn't have a unique $\mathbb{Z}$-basis, but we can define $J : \mathbb{H}/(\mathrm{PSL}_2(\mathbb{Z})) \to \mathbb{C}$ by $J([\tau]) = j(\Lambda_{\mathbb{Z}\cdot 1 \oplus \mathbb{Z}\cdot\tau})$. In particular, $\mathbb{H}/\mathrm{PSL}_2(\mathbb{Z})$ is the space of lattices, and J is also biholomorphic. There's a lot one could say about this function as well: it's invariant under $\tau \mapsto \tau + 1$, and it has a Fourier expansion in powers of $q = e^{2\pi i\tau}$. It turns out that this has the form

$$J(\tau) = \frac{1}{q} + 744 + \sum_{n \geq 1} c_n q^n,$$

where c_n are positive integers that can be made explicit. On the one hand, this explains why we normalized the j-invariant we defined, but the 744?

Even more amazingly, Conway and Norton's "monstrous moonshine" conjecture (later a Fields Medal-winning proof) realizes these c_n as dimensions of a graded algebra that the monster simple group acts on!

Lecture 20.

The Euler Characteristic and the Riemann-Hurwitz Formula: 3/21/16

This week, we're going to talk about the Euler characteristic, corresponding to Chapter 7 of the textbook. We'll prove a bunch of interesting or cute facts, none of which are particularly deep; but after this, we move into the second part of the textbook, which has deeper results: Dolbeault cohomology, the Riemann-Roch theorem, and the uniformization theorem.

The Riemann-Hurwitz Formula. To talk about the Euler characteristic we'll need to exclude a few pathological examples.

Definition. If S is a surface, it has *finite (topological) type* if $\dim H^j(S)$ is finite for $j = 0, 1, 2$.

[30] A coarse moduli space is one that does not keep track of automorphisms of a given object.

For example, all compact surfaces have finite type. A counterexample would be an infinite disjoint union of spheres (for which H^0 is infinite) or a surface with infinite genus (which has infinite-dimensional H^1).

The Euler characteristic is an essentially topological definition.

Definition. If S is a surface of finite type, then its *Euler characteristic* is $\chi(S) = \dim H^0(S) - \dim H^1(S) + \dim H^2(S)$.

Proposition 20.1 (Mayer-Vietoris property)**.** *Suppose $S = S_0 \cup S_1$ for open subsets $S_0, S_1 \subset S$. If S_0, S_1, and $S_0 \cap S_1$ have finite type, then so does S, and $\chi(S) = \chi(S_0) + \chi(S_1) - \chi(S_0 \cap S_1)$.*

The proof is immediate from the Mayer-Vietoris sequence for cohomology.[31]

The fact that singular cohomology with coefficients in $\mathbb{R}$ agrees with de Rham cohomology gives us the following property, which is often taken as a more elementary definition of the Euler characteristic.

Proposition 20.2. *If S is a surface with a finite triangulation, $\chi(S)$ is the number of vertices minus the number of edges plus the number of faces.*

If $f : S \to T$ is a finite-degree covering map, one can pull a triangulation on T back to a triangulation on S; there are a number of different geometric or algebraic proofs of this, including one with spectral sequences! But the point is the following corollary.

Proposition 20.3. *Let T be a surface of finite type and $f : S \to T$ is a degree-d covering map (d must be finite), then $\chi(S) = d\chi(T)$.*

Now, we can state the Riemann-Hurwitz formula. Note, though, that there are typos in the textbook; the version given in class is more accurate.

Theorem 20.4 (Riemann-Hurwitz formula)**.** *Let X and Y be compact Riemann surfaces and $f : X \to Y$ be a surjective holomorphic map that's nowhere locally constant.*[32] *Then,*

$$-\chi(X) = -\deg(f)\chi(Y) + R_f, \tag{20.5}$$

where R_f, the total ramification*, is given by*

$$R_f = \sum_{x \in X} (k_x - 1), \tag{20.6}$$

where k_x is the local multiplicity of f at x (i.e. there are coordinates centered at x in which $f(z) = z^{k_x}$).

The sum in (20.6) is finite because $k_x = 1$ unless x is a critical point, and because X and Y are compact, there are only finitely many critical points. The theorem might be true for noncompact X and Y with certain proper maps f, but we're only going to need it in the compact case.

The signs in the formula (20.5) are hard to remember, but if you can do the calculation below, you'll be fine.

Example 20.7. We know that if X is a genus-g hyperelliptic Riemann surface, then X is branched over b points in S^2, where $b = 2g + 2$ (this was the "skewering" that we constructed a few lectures back). This is a consequence of Theorem 20.4: $\chi(X) = 2 - 2g$, so $-\chi(X) = 2g - 2$, and the right-hand side is $2(-2) + b$ (each critical point contributes 1 to the total ramification), and so $2g - 2 = b - 4$, or $b = 2g + 2$.

Trying this with a genus-1 curve over the 2-sphere can help you remember the sign rule in a pinch.

We'll prove this momentarily; first, we present some corollaries.

Corollary 20.8. *The total ramification R_f is even.*

This is because $\chi(X)$ and $\chi(Y)$ are both even: a hyperelliptic surface can't be branched at three points over the 2-sphere.

Corollary 20.9. *Let X and Y be compact, connected Riemann surfaces and $f : X \to Y$ be a nonconstant holomorphic map. Then, the genus of X is at least the genus of Y.*

---

[31] Vietoris did work on this in the 1920s, but lived until 2002! He was Austria's oldest citizen for some time, which is also surprising because the oldest citizen is usually a woman. How different the subject must have looked at the end of his life from what he grew up studying!

[32] If X and Y are connected, then all this means is that f is nonconstant.

For example, you can't map $S^2 \to T^2$ (or any higher-genus surface) except by the constant map.

Proof. Let $g(S)$ denote the genus of a surface S. If $g(Y) = 0$, there's nothing to prove, so suppose $g(Y) > 0$, or $-\chi(Y) \geq 0$, and therefore the right-hand side of (20.5) is $-\deg f \cdot \chi(Y) + R_f \geq -\chi(Y)$, because $-\deg f$ is nonnegative and R_f is nonnegative. Thus, $2g(X) - 2 \geq 2g(Y) - 2$. ⊠

This is interesting because it's a restriction on what kinds of holomorphic maps can exist based on a purely topological input. The converse is not true, however: there are Riemann surfaces of the same genus (e.g. different elliptic curves) with no nonconstant maps between them.

Proof of Theorem 20.4. Let $\Delta \subset Y$ be the set of critical values; since Y is compact, this is finite, so let $b = |\Delta|$. Let N_Δ be a small neighborhood of Δ consisting of a small closed disc around each $\delta \in \Delta$.

Let $Y_0 = Y \setminus N_\Delta$ and $X_0 = f^{-1}(Y_0)$, so $f : X_0 \to Y_0$ is a proper map with no critical points, hence a covering map of degree $d = \deg(f)$. Thus, by Proposition 20.3, $\chi(X_0) = d\chi(Y_0)$. Now, we'll relate $\chi(X)$ and $\chi(Y)$ to this relation.

We can use the Mayer-Vietoris property, Proposition 20.1, to show that $\chi(Y) = \chi(Y_0) + b\chi(D) - b\chi(A)$, where D is a disc and A is an annulus. Since D is contractible, $\chi(D) = 1$, and since $A \simeq S^1$, then $\chi(A) = 0$. Thus, $\chi(Y) = \chi(Y_0) + b$.

Next, let's look at X. Over each $\delta \in \Delta$, we're gluing a number of discs to X_0, where the fibers over points near δ are grouped by monodromy cycles. Thus, for each critical point $x \in f^{-1}(\delta)$, we're gluing k_x discs (e.g. if the monodromy is trivial, we glue in one disc), so the number of discs over preimages over δ is

$$d - \sum_{x \in f^{-1}(\delta)} (k_x - 1).$$

Thus,

$$\begin{aligned} \chi(X) &= \chi(X_0) + \sum_{\delta \in \Delta} \left(d - \sum_{x \in f^{-1}(\delta)} (k_x - 1) \right) \\ &= \chi(X_0) + db - \sum_{x \in \mathrm{crit}(f)} (k_x - 1). \end{aligned}$$

Combining this with the formula for $\chi(Y)$, we get that $\chi(X) + R_f = d\chi(Y)$. ⊠

The proof ultimately comes from the way we counted the discs that we've glued in.

Here's another nice application of the Riemann-Hurwitz formula.

Theorem 20.10. *Let X be a compact, connected Riemann surface of genus $g \geq 2$ and G be a finite group of automorphisms acting on X. Then, $|G| \leq 84(g - 1)$.*

Remark.

(1) It turns out that $\mathrm{Aut}(X)$ is also finite, so one can take $G = \mathrm{Aut}(X)$. This is a stronger theorem, and doesn't follow from the Riemann-Hurwitz theorem. The proof emerges from the Riemann-Roch formula; one has a finite set of points, called *Weierstrass points* permuted by $\mathrm{Aut}(X)$, and the induced map $\mathrm{Aut}(X) \to S_N$ will be injective. Of course, the automorphism group of a genus-0 or 1 surface isn't finite.

(2) *Klein's quartic curve* X has genus $g = 3$ and $|\mathrm{Aut}\, X| = 168 = 2 \cdot 84$, so this bound is satisfied. ($\mathrm{Aut}(X)$ is actually the simple group of order 168, $\mathrm{PSL}_2(\mathbb{F}_7)$.) It's known that the bound is sharp for some infinite sequence of genera.

Proof of Theorem 20.10. It turns out that $Y = X/G$ has the structure of a Riemann surface. This is not obvious, since points in X may have nontrivial stabilizers, but the proof is similar to that for Fuchsian groups. Near an $x \in X$ with nontrivial stabilizer, find a G-invariant coordinate disc D, so that $\mathrm{stab}_G(x)$ is a group of automorphisms of D fixing 0, and therefore a group of rotations. Thus, $\mathrm{stab}_G(x) = \mu_n$: it's the n^{th} roots of unity for some n. That is, the projection $X \to Y$ is the map $z \mapsto z^n$, which we know gives the structure of a smooth manifold near x.

Now, let's apply Riemann-Hurwitz to $\pi : X \to Y = X/G$. The degree is $|G|$, so $2g(x) - 2 = |G|(2g(Y) - 2) + R_\pi$. In the same way as the proof of the Riemann-Hurwitz formula, one can show (yes, this is an exercise) that

$$R_\pi = \sum_{\substack{\text{non-free orbits} \\ y = G\cdot x}} (|G| - |\mathrm{stab}_G(x)|),$$

which therefore implies that

$$\frac{2g(X) - 2}{|G|} = 2g(Y) - 2 + \sum_y \left(1 - \frac{1}{|\mathrm{orb}_G(x)|}\right).$$

If $g(X) \geq 2$, then the left-hand side of this equation is nonzero, so

$$2g(Y) - 2 + \sum_{y=G\cdot x} \left(1 - \frac{1}{|\mathrm{orb}_G(x)|}\right) > 0. \tag{20.11}$$

The somewhat magical fact is that if we minimize over $g(Y) \in \mathbb{Z}_{\geq 0}$ and $e_1, \dots, e_N \geq 2$, then the minimum positive value of

$$2g - 2 + \sum_{i=1}^N \left(1 - \frac{1}{e_i}\right) \tag{20.12}$$

is 42.[33] This, together with (20.11), shows that

$$\frac{2g(X) - 2}{|G|} \geq \frac{1}{42},$$

which proves the theorem.

So why 42? To minimize (20.12), we should definitely set $g = 0$. Let $E = \sum_{i=1}^N (1 - 1/e_i)$, which needs to be greater than 2; since $e_i \geq 2$, then $1 > 1 - 1/e_i \geq 1/2$, so we need $N > 2$ to get anything interesting. If $N \geq 5$, then $E \geq 5 \cdot (1/2) = 2 + (1/2)$, so we don't do better than $1/2$. If $N = 4$, we can exceed 2 by $1/2 + 1/2 + 1/2 + 2/3$, so we can get $1/6$. If $N = 3$, we can use $2/3 + 2/3 + 3/4$, giving us $1/12$, which is better. We can take $e_1 = 2$, so $e_2, e_3 \geq 3$, and if they're both at least 4, we get $E \geq 1/2 + 3/4 + 4/5 = 1/20$, which is another lower bound… but we actually achieve the minimum with $e_1 = 2$ and $e_2 = 3$, so $E = 1/2 + 2/3 + 6/7$, giving us $2 + 1/42$. ⊠

This is essentially an exercise in Hartshorne's algebraic geometry textbook.

Lecture 21.

Modular Curves: 3/23/16

Today, we're going to talk about modular curves, which are covered in §6.3.2 and §7.2.4 of the textbook. One beautiful aspect of Riemann surfaces is how many interesting and nontrivial things you can do with just examples (e.g. elliptic curves, which we're already briefly discussed). Modular curves are where modular forms live, and these are loved by number theorists and representation theorists; hopefully, we can return to these.

We've already talked about the *modular group* $\Gamma = \mathrm{PSL}_2(\mathbb{Z}) = \mathrm{SL}_2(\mathbb{Z})/\{\pm 1\}$. This acts on the upper half-plane $\mathbb{H}$ by Möbius maps, so the modular group is a Fuchsian group; the quotient $Y = \mathbb{H}/\Gamma$ is the Riemann surface of lattice $\Lambda \subset \mathbb{C}$ up to *homothety* (i.e. $\Lambda \sim c\Lambda$ for all $c \in \mathbb{C}^*$), where $\tau \in \mathbb{H} \mapsto \Lambda_\tau = \mathbb{Z}\cdot 1 \oplus \mathbb{Z}\cdot\tau$. This Y is an interesting space, because it is also the moduli space of elliptic curves up to isomorphism. Through this identification, we found that the j-invariant defines a biholomorphic map $J : Y \to \mathbb{C}$.

Today (and for part of Friday's lecture), we're going to do the following.

- Establish a fundamental domain for Γ acting on $\mathbb{H}$.
- Talk about modular curves $Y_N = \mathbb{H}/\Gamma_N$, where Γ_N is a principal congruence subgroup of Γ; these are the elements that reduce mod N to the identity. We'll also discuss their compactifications. Through the Riemann-Hurwitz formula (Theorem 20.4), we'll calculate their genus.

We begin with two group-theoretic facts.

[33] It's the answer to the Ultimate Question of Life, the Universe, and Everything!

Fact.

- If
$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \qquad \text{and} \qquad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$
then Γ is generated by S and T. Geometrically, if $\tau \in \mathbb{H}$, $S(\tau) = 1/\tau$, so S is inversion, and $T(\tau) = \tau + 1$, meaning T is translation to the right.[34]
- One can see that $S^2 = I$ and $(ST)^3 = I$, and these are in fact all the relations: $\Gamma = \langle S, T \mid S^2 = (ST)^3 = I\rangle$.

Definition. A *fundamental region* for Γ acting on $\mathbb{H}$ is a set X such that every Γ-orbit intersects the closure $\overline{X}$, and no two points of X lie in the same Γ-orbit.

Since we have to use the closure, this is not the same as a fundamental domain.

Theorem 21.1. *If $\Omega = \{z \in \mathbb{H} \mid |z| > 1, |\mathrm{Re}\, z| < 1/2\}$, then Ω is a fundamental region for Γ acting on $\mathbb{H}$.*

See Figure 8. Geometrically, this makes sense: T marches to the right by 1, so $T(\Omega)$ and Ω are distinct, and every orbit intersects $\overline{\Omega}$, and S inverts, so the inverses of things in Ω are inside the unit circle. This means that Y looks like Ω: the two points of Ω are i and $\rho = e^{i\pi/3}$, corresponding to the square lattice $\Lambda_i \in Y$, which is invariant under multiplication by i (rotation by 90°), and the hexagonal lattice $\Lambda_\rho \in Y$, which is invariant under multiplication by ρ (which is rotation by 60°).

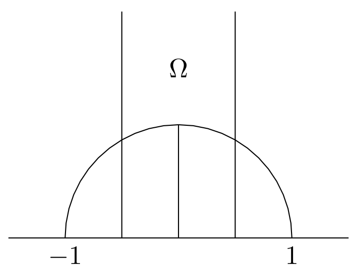


FIGURE 8. A fundamental region for Γ acting on $\mathbb{H}$, as in Theorem 21.1.

Proof of Theorem 21.1. Pick an $A \in \mathrm{SL}_2(\mathbb{Z})$, so $A = \left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. If $c = 0$, then $ad = 1$, so $A = \left[\begin{smallmatrix} \pm 1 & b \\ 0 & \pm 1 \end{smallmatrix}\right]$; in particular, in this case, $A = \pm T^m$ is a translation. Correspondingly, if A is a translation, then $c = 0$.

In general, we can directly compute a useful formula: if $\tau \in \mathbb{H}$, then
$$\mathrm{Im}(A(\tau)) = \frac{\mathrm{Im}(\tau)}{|c\tau + d|^2}. \tag{21.2}$$
Now, we're going to march through a few lemmas. The overarching idea is that, within some Γ-orbit, we'd like to find the point in $\overline{\Omega}$ that maximizes $\mathrm{Im}(\tau)$; we will also show that given any Γ-orbit, we do have some such point, and then we're done.

Lemma 21.3. *If $\tau \in \Omega$ and $c \neq 0$, then $\mathrm{Im}(A(\tau)) < \mathrm{Im}(\tau)$.*

This is a quick check using (21.2).

Lemma 21.4. *If $\tau, A\tau \in \Omega$, then $A = \pm I$, so $A\tau = \tau$.*

Proof. If $c \neq 0$, then by Lemma 21.3, $\mathrm{Im}(A\tau) < \mathrm{Im}(\tau) = \mathrm{Im}(A^{-1}A\tau) < \mathrm{Im}(A\tau)$, which is a contradiction. Hence, $c = 0$, so $A = \pm T^m$ is a translation. However, $T^m(\Omega) \cap \Omega = \varnothing$ unless $m = 0$, so $A = \pm I$ as desired. ⊠

We've now shown that no two points of Ω lie in the same Γ-orbit, which is half of the proof. But we still need to show that every Γ-orbit is represented.

Lemma 21.5. *Suppose τ^* maximizes $\mathrm{Im}(\tau)$ within its Γ-orbit; then, $|\tau^*| \geq 1$.*

[34] For a reference, see T. Apostol, *Modular Functions and Dirichlet Series in Number Theory.*

Proof. If $|\tau| < 1$, then $\operatorname{Im}(S\tau) = \operatorname{Im}(-1/\tau) = \operatorname{Im}(\tau)/|\tau|^2 > \operatorname{Im}(\tau)$. ⊠

So in this case, you're almost in Ω — but might be in a translate of it.

Fix a $\tau_0 \in \mathbb{H}$ and let $h = h_{\tau_0} : \Gamma \to \mathbb{R}$ send $\gamma \mapsto \operatorname{Im}(\gamma \cdot \tau_0)$.

Lemma 21.6. *h is bounded above.*

Proof. One could approach this with hyperbolic geometry, but it's just as true if you prove it directly: $\operatorname{Im}(A\tau_0) = \operatorname{Im}(\tau_0)/|c\tau_0 + d|^2$ by (21.2). However, $c\tau_0 + d$ is a nonzero point in the lattice Λ_{τ_0}, and therefore it cannot be arbitrarily close to 0, and therefore its modulus is bounded. ⊠

One lemma remaining.

Lemma 21.7. *h attains its maximum.*

Proof. Take a sequence $\tau_n = \gamma_n \cdot \tau_0$ such that $h(\tau_0) \to \sup h$, which is finite by Lemma 21.6. Thus, $\operatorname{Im}(\tau_n)$ is also bounded above. Using translations T^m, we may assume $|\operatorname{Re}(\tau_n)| \leq 1/2$, so $\tau_n \in [-1/2, 1/2] \times [0, \sup h]$. This is compact, so there's a convergent sequence $\{\tau_{n_j}\}$ converging to some τ^*. We'd like this to lie in the orbit of τ_0, but this is not immediate: we do know that $\tau^* \in \overline{\Gamma \cdot \tau_0}$. However, Γ is a Fuchsian group, so its action on the upper half-plane is properly discontinuous. That means that every point $p \in \mathbb{H}$ has a neighborhood N in which $N \cap (\Gamma \cdot p) = \{p\}$: we can separate the points in an orbit. (This is something we proved, in Proposition 9.2.) Thus, τ_{n_j} must be constant for large j, and in particular $\tau^* \in \Gamma \cdot \tau_0$. Thus, τ^* maximizes h. ⊠

In conclusion, let τ^* maximize h, and through translations, we may assume $\operatorname{Re}(\tau^*) \leq 1/2$, and by Lemma 21.5, $|\tau^*| \geq 1$. Thus, $\tau^* \in \overline{\Omega}$, so $\overline{\Omega}$ intersects all Γ-orbits. ⊠

On $Y = \mathbb{H}/\Gamma$, we have to identify the two left and right sides of Ω because of T, and the two sides of the arc on the bottom, thanks to S. This makes the Riemann surface structure on Y look a little strange: the semicircle around i in Ω is actually promoted to a full cuircle through $z \mapsto z^2$. Simialarly, the map $z \mapsto z^3$ makes a neighborhood of ρ into a disc neighborhood.

Another consequence is that one can tile $\mathbb{H}$ by translations of $\overline{\Omega}$ (as with any fundamental domain or region).

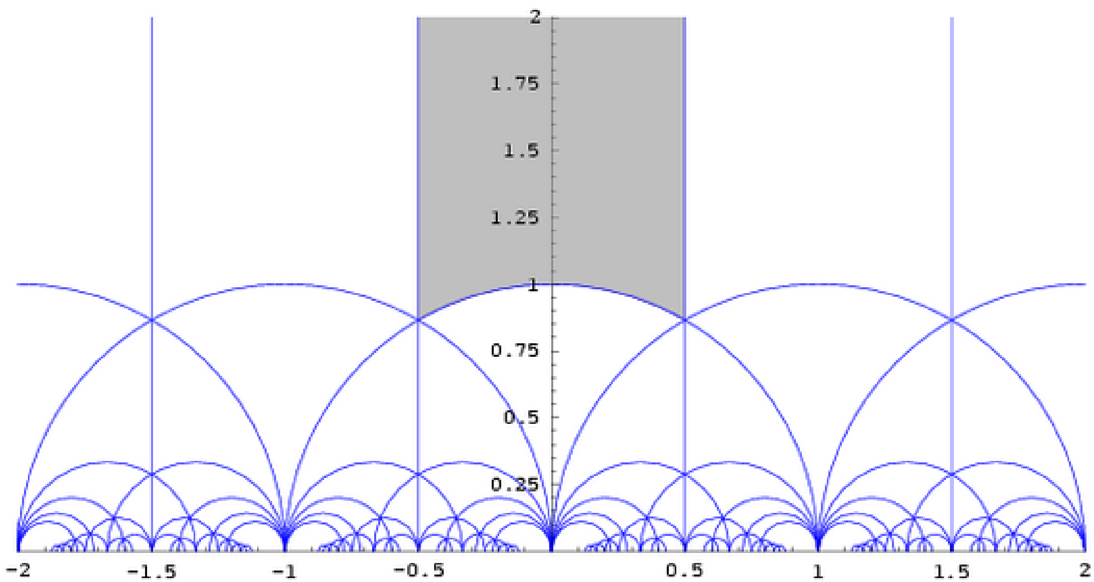


Figure 9. Tiling the upper-half plane with translations of $\overline{\Omega}$ by the actions of S and T. Source: http://www.math.mcgill.ca/goren/ModularForms07-08/ModularForms.html.

Lecture 22.

Modular Curves for Principal Congruence Subgroups: 3/25/16

Last time, we proved Theorem 21.1, which characterized a fundamental region (almost, but not quite, a fundamental domain) for the modular group $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$. This region Ω is the set of $z \in \mathbb{H}$ with $|\mathrm{Re}\, z| < 1/2$ and $|z| > 1$.

Today, we're going to talk about the subgroups $Q_N = \mathrm{PSL}_2(\mathbb{Z}/N)$ for $N = 1, 2, 3, \ldots$. It's a standard fact in modular arithmetic that reducing mod N defines a map $\varphi_N : \Gamma \to Q_N$ which is surjective! Then, let $\Gamma_N = \ker(\varphi_N)$. We've previously studied Γ_p for p prime, as in Example 9.1, but the primality is actually irrelevant: the action of Γ_N on $\mathbb{H}$ is free for $N \geq 2$.

Define $Y_N = \mathbb{H}/\Gamma_N$, which carries a residual action by Q_N; let π_N denote the projection $Y_N \twoheadrightarrow Y_N/Q_N \cong \mathbb{H}/\Gamma = Y$ (by the third isomorphism theorem, more or less). These Y_N are quotients by Fuchsian groups, hence Riemann surfaces; they will be our principal object of study today.

Y parameterizes lattices $\Lambda \subset \mathbb{C}$, but since we've quotiented by a smaller group, we also remember some extra structure: Y_N parameterizes lattices along with something called a level N structure modulo homothety (where $\Lambda \sim c\Lambda$ as usual, and which does something sensible to level structures). Since $\mathbb{C}/\Lambda$ is an abelian group, we can let $(\mathbb{C}/\Lambda)[N]$ denote its N-torsion subgroup.

Definition. A *level N structure* on a lattice Λ is an isomorphism $(\mathbb{C}/\Lambda)[N] \cong (\mathbb{Z}/N) \oplus (\mathbb{Z}/N)$; in other words, it's a $\mathbb{Z}/N$-basis for the N-torsion.

Similarly, since Y is a moduli space for elliptic curves, Y_N should be too, and should carry some extra information, also called a level N structure. Using the abelian group structure of the elliptic curve, one could also frame the level N structure in terms of torsion subgroups, but we'll give a different definition.

Definition. A *level N structure* on an elliptic curve E is an isomorphism $H_1(E; \mathbb{Z}/N) \cong (\mathbb{Z}/N) \oplus (\mathbb{Z}/N)$.

The main reason we care about these things is that they're the home of modular forms for these particular principal congruence subgroups; we'll hopefully get to that by the end.

First, though, can we come up with a fundamental region for the action of Γ_N on $\mathbb{H}$? Since Γ_N is smaller, we should get a larger region. Let $q = q_N = |Q_N|$, and let $g_1, \ldots, g_q \in \Gamma$ be representatives for the cosets of Γ_N. Then, let Ω_N be the interior of $g_1\overline{\Omega} \cup \cdots \cup g_q\overline{\Omega}$; this is a fundamental region for Γ_N. Depending on the representatives of cosets you picked, this might not be connected.

One corollary of this is that the projection map $\pi_N : Y_N \to Y$ is proper; the preimage of a compact set is contained in finitely many $g_i\overline{\Omega}$, and hence is still closed and bounded in $\mathbb{H}$, hence compact.

Example 22.1. Let's see what happens when $n = 2$. Then, the first column of a matrix in $\mathrm{PSL}_2(\mathbb{Z}/2)$ can be anything except all zeroes, and the second column has the same, but can't be a multiple of the first; thus, this gives us $(4-1)(4-2) = 6$ elements. We idtneify them up to multiplication in $(\mathbb{Z}/2)^\times$, but this just contains the identity; hence, we get 6 elements in $\mathrm{PSL}_2(\mathbb{Z}/2)$.

Let S and T be the generators we defined last lecture; then, we will take the cosets of Γ_2 to be those given by 1, T^{-1}, S, $T^{-1}S$, STS, and ST. We choose these because they produce a nice connected region of $\mathbb{H}$, bounded by a hyperbolic hexagon. This includes a point at infinity, a "end" of Ω_2. If $\rho = e^{i\pi/3}$, then one can check that $ST(\rho^2) = \rho^2$, so ST is a rotation of order 3 around ρ^2 (the center of the hexagon). These give us the three "ends at infinity," either going to infinity in the imaginary direction or touching the x-axis. $T^2 \in \Gamma_2$, and it identifies the two vertical boundaries going to infinity.

The point is that Y_2 is this hexagon with three pairs of adjacent edges identified. Making these identifications pinches them off, and so Y_2 is biholomorphic to the thrice-punctured sphere, and $\pi_2 : Y_2 \twoheadrightarrow Y$ is a degree-6 map branched over two critical points, ρ and i (since these are the two non-free Γ-orbits); for these points, the stabilizers have size 3 and 2, respectively.

This example should aid visualization (the pictures were drawn on the blackboard; I couldn't reproduce them quickly, but I encourage you to draw them, or look up a program that generates such fundamental domains).

So for Y_2, we can compactify it by putting in the three points at infinity, to obtain S^2. In general, we can compactify Y_N; we'll call the compactification X_N.

- First, compactify Y to get S^2, by adding a point at ∞.

- Using the monodromy at ∞ for π_N, we get a recipe for extending $\pi_N : Y_N \to Y$ to a branched covering $\pi_n : X_N \to S^2$; this is analogous to the construction we used in things such as the Riemann existence theorem earlier in the course. Thus, X_N is Y_N along with a number of "cusps" over $\infty \in S^2$.

Q_N still acts on X_N, where the cusps form exactly one Q_N-orbit; for $N = 2$, there are 3 cusps, and π_2 has branching order 2 at each cusp.

We're going to use the Riemann-Hurwitz formula (Theorem 20.4) to compute the genus $g(X_N)$, thus characterizing Y_N as a surface of a specified genus minus a known number of points. To do this, we'll define Ω_N by taking the set of cosets $I, T, T^2, \dots, T^{n-1}$, and some others (many others, though we won't write them down). Thus, Ω_N consists of a bunch of translates of Ω by T, plus some extra stuff given by the action of S. The action of T^N identifies the ends of $P = \overline{\Omega} \cup T\overline{\Omega} \cup \dots \cup T^{n-1}\overline{\Omega}$, which wraps the ends to a cylinder, giving one cusp at infinity, but doesn't affect whatever is going on with S. And in particular, it contains the neighborhood of exactly one cusp, and at this cusp, π_N is branched to order N (around this cusp, it looks like the cyclic map $z \mapsto z^N$).

In general, for a finite group action G on a compact Riemann surface X, the proof of Theorem 20.10 implies that

$$-\chi(X) = |G|\left(-\chi(X/G) + \sum_{\text{orbits}} \left(1 - \frac{1}{|\text{stab}_G\, x|}\right)\right). \tag{22.2}$$

For Q_N acting on X_N, the nonfree orbits are those of ρ, whose stabilizer has size 3; i, whose stabilizer has size 2; and ∞, whose stabilizer has size N. Plugging this into (22.2),

$$\begin{aligned} 2g(X_N) - 2 &= q_n\left(-2 + \left(1 - \frac{1}{3}\right) + \left(1 - \frac{1}{2}\right) + \left(1 - \frac{1}{N}\right)\right) \\ &= q_N\left(\frac{1}{6} - 1N\right) = \frac{q_N(N-6)}{6N}. \end{aligned}$$

That is, we've derived the following genus formula.

Proposition 22.3. *With X_N as in the above discussion,*

$$g(X_N) = 1 + \frac{q_N(N-6)}{12N},$$

where $q_N = |\text{PSL}_2(\mathbb{Z}/N)|$.

Let's make this more explicit: if $N = p$ is an odd prime, $|\text{GL}_2(\mathbb{Z}/p)| = (p^2-1)(p^2-p)$ (since the first column can be anything nonzero, and the second column can be anything other than a multiple of the first), and the kernel of the quotient $\text{GL}_2(\mathbb{Z}/p) \to (\mathbb{Z}/p)^*$ is $\text{SL}_2(\mathbb{Z}/p)$, so we've divided by $|(\mathbb{Z}/p)^*| = p - 1$, and therefore $|\text{SL}_2(\mathbb{Z}/p)| = p(p^2-1)$. Thus, $|Q_p| = |\text{PSL}_2(\mathbb{Z}/p)| = p(p^2-1)/2$ for an odd prime p. That is, $g(X_p) = (1/24)(p+2)(p-3)(p-5)$.

(1) By Example 22.1, $g(X_2) = 0$.
(2) $g(X_3) = 0$ and $g(X_5) = 0$, and 2, 3, and 5 are the only examples with genus 0.
(3) $g(X_7) = 3$. $q_3 = |\text{PSL}_2(\mathbb{Z}/7)| = 168$. This is a famous example: $\text{PSL}_2(\mathbb{Z}/7)$ is a nonabelian simple group of order 168, and its action on X_7 realizes the bound we found in Theorem 20.10: it has the largest possible symmetry group.

These X_N are where modular forms live; but that's a story for another day.

Lecture 23.

Euler Characteristics and Meromorphic 1-Forms: 3/28/16

Today, we're going to talk about the Poincaré-Hopf theorem.

Let S be a compact, oriented surface and $\alpha \in \Omega^1(S)$ be a 1-form whose zeros are isolated. If p is a zero of α, we'd like to assign it a *multiplicity* $m_\alpha(p) \in \mathbb{Z}$, which we do as follows. Suppose (x, y) are oriented coordinates centered at p, so that in these coordinates, $\alpha = \alpha_1\, dx + \alpha_2\, dy$. If γ is a small circle around p, so that $\gamma(t) = (r\cos(2\pi t), r\sin(2\pi t))$, for $t \in [0, 1]$, then $\alpha \circ \gamma : \mathbb{R}/\mathbb{Z} \to \mathbb{R}^2 \setminus \{(0,0)\}$; then, we defined $m_\alpha(p)$ to be the winding number of $\alpha \circ \gamma$. It's easy to show this is independent of the coordinates we chose.

Theorem 23.1 (Poincaré-Hopf theorem for surfaces)**.**

$$\chi(S) = -\sum_{p\colon \alpha(p)=0} m_\alpha(p).$$

There's a more general version for manifolds; moreover, there's a version that uses vector fields instead of 1-forms, in which case the minus sign disappears. This can be found in many textbooks on differential topology (though, curiously, not Guillemin and Pollack's); the proof generally involves relating both sides to the self-intersection number of the diagonal $\Delta_S \subset S \times S$, which is equal to the Euler characteristic (this is a theorem, since we didn't define the Euler characteristic in this way).

Assuming Theorem 23.1, we derive the following corollary.

Proposition 23.2. *Let X be a compact Riemann surface and θ be a holomorphic 1-form on X. Assume θ isn't identically 0 on any connected component of X (which automatically implies its zeros are isolated). Then,*

$$\chi(X) = -\sum_{p\colon \theta(p)=0} k_p(\theta),$$

where in local coordinates, $\theta = f(z)\,\mathrm{d}z$, so we define $k_p(\theta)$ to be the order of vanishing of f at p.

Proof. Let $\alpha = \mathrm{Re}(\theta) = (1/2)(\theta + \overline{\theta})$, so the zeros of α are the zeros of θ. Near such a zero, $z = x + yi$, so if $\theta = f(z)\,\mathrm{d}z$, then $\alpha = (\mathrm{Re}\, f)\,\mathrm{d}x + (\mathrm{Im}\, f)\,\mathrm{d}y$, and so the winding number of $(\mathrm{Re}\, f, \mathrm{Im}\, f)$ around $z = re^{i\vartheta}$ is the order of vanishing of f at 0, by Theorem 23.1. ⊠

Corollary 23.3. *S^2 has no nonzero holomorphic 1-forms.*

This is because $\chi(S^2) = 2$, so $\sum k_p(\theta) \le 0$ for any such θ.

We can also define $k_p(\theta)$ when p is a pole of a meromorphic 1-form: if $\theta = f(z)\,\mathrm{d}z$ and f has a pole of order n at p, then we set $k_p(\theta) = -n$. For example, a simple pole contributes -1 to the sum.

Theorem 23.4. *Let X be a compact Riemann surface and θ be a meromorphic 1-form on X that isn't identically 0 on any connected component of X. Then,*

$$\chi(X) = -\sum_{p\colon \text{poles and zeros}} k_p(\theta).$$

This is the version of the Poincaré-Hopf theorem that we will find the most useful.

Proof. We know what happens at zeros of θ, so fix a coordinate disc D_p around each pole p. By shrinking D_p if needed, we can assume that the D_ps are disjoint, and that θ doesn't vanish on any D_p. Thus, each D_p is biholomorphic to the closed unit disc $\overline{\mathbb{D}}$.

We'd like to obtain an ordinary 1-form out of θ, but thanks to the poles this will be a little harder. Let $\psi : [0, \infty) \to [0, \infty)$ be a smooth function defined such that:

- If $0 \le t \le 1$, $\psi(t) = 1$.
- If $t \ge 2$, then $\psi(t) = 1/t$.
- On $(1, 2)$, ψ smoothly interpolates between these two curves.

Now, for a pole p, regarding D_p as $\overline{\mathbb{D}}$, we can write $\theta(z) = cf(z)\,\mathrm{d}z$, where $c \in \mathbb{C}^\times$ and f are chosen such that $|f| \le 1$ on the annulus $\{z \mid 1/2 \le z \le 1\}$.

Now, on D_p, set $\widetilde{\theta}_p(z) = \psi(|f(z)|^2)cf(z)\,\mathrm{d}z$. For $1/2 \le |z| \le 1$, $\widetilde{\theta}_p(z) = \theta(z)$, so we can extend to a 1-form $\widetilde{\theta}$, which is defined to be θ outside of all D_p, and to be $\widetilde{\theta}_p$ on D_p. This is a smooth 1-form (and has no poles), but isn't holomorphic! Near $p = 0 \in D_p$, $\widetilde{\theta}_p(z) = cf(z)/|f(z)|^2\,\mathrm{d}z = c/\overline{f(z)}\,\mathrm{d}z$, and therefore $\widetilde{\theta}_p(0) = 0$.

Thus, $\widetilde{\theta}$ is a complex 1-form, whose zeros are the zeros and poles of θ, and so we can repeat the argument that proved Proposition 23.2, checking that the multiplicities for $\alpha = \mathrm{Re}(\widetilde{\theta})$ are exactly the values of $k_p(\theta)$; the trick with the poles is that the complex conjugate we had in a neighborhood of p switches the sign to the correct one. ⊠

There might be ways to prove this that lie entirely within algebraic geometry.

One consequence of Theorem 23.4 is that it allows a new proof of the Riemann-Hurwitz formula, Theorem 20.4. Recall that this says if X and Y are compact Riemann surfaces and $f : X \to Y$ is smooth, then $-\chi(X) = -(\deg f)\chi(Y) + R_f$, where R_f is the total ramification of f, defined in (20.6).

Proof of Theorem 20.4 using Theorem 23.4. This proof is contingent on the existence of a nonvanishing, meromorphic 1-form θ on Y; this exists by the Riemann-Roch theorem, which we will prove later. Thus, $f^*\theta$ is a meromorphic 1-form on X; we'd like to apply our formula to it.

Let $p \in X$ be a pole of $f^*\theta$. Let z be a holomorphic local coordinate for X centered at p, and w be a holomorphic local coordinate for Y centered at $f(p)$. We can (and do) choose z and w such that $f(z) = w = z^m$, where k is the multiplicity of f at p.

In w-coordinates, $\theta = g(w)\,\mathrm{d}w$, so $f^*\theta = g(z^k)\,\mathrm{d}(z^m) = mg(z^m)z^{m-1}\,\mathrm{d}z$; hence, if $m_{f(p)}(\theta) = \ell$, then $f^*\theta$ has multiplicity $m\ell + m - 1$ at $f(p)$.

For $x \in X$, let m_x denote the multiplicity of f at x, and for all $y \in Y$, let $\ell_y = k_\theta(y)$. Thus,

$$\begin{aligned}\sum_{x \in f^{-1}(y)} k_{f^*\theta}(x) &= \sum_{x \in f^{-1}(y)} (m_x \ell_y + m_x - 1)\\ &= (\deg f)\ell_y + \sum_{x \in f^{-1}(y)} (m_x - 1).\end{aligned}$$

Summing over $y \in Y$, we get that

$$-\chi(X) = -(\deg f)\chi(X) + \underbrace{\sum_{x \in X} (m_x - 1)}_{R_f}. \qquad \boxtimes$$

This is useful for generalizing the Riemann-Hurwitz formula to branched coverings of higher-dimensional complex manifolds.

Another useful consequence of Theorem 23.4 is to calculate the genus of plane algebraic curves.

Theorem 23.5. *Let P be a degree-d homogeneous polynomial in $\mathbb{C}[z_0, z_1, z_2]$ and $X = \{P = 0\} \subset \mathbb{C}P^2$. Suppose that the $\frac{\partial P}{\partial z_i}$ don't all vanish at any point, so X is a Riemann surface. Then, the genus of X is*

$$g(X) = \frac{1}{2}(d-1)(d-2).$$

This is quite powerful for characterizing genera of algebraic curves: a degree-1 or 2 curve must have genus 0, a cubic must be degree 1, and in general we only get triangular numbers.[35]

Proof. Fix a projective line $L \subset \mathbb{C}P^2$ which intersects X at d distinct points; this is a generic condition, so we can always find such an L. We can make a change of coordinates of $\mathbb{C}P^2$ such that L is the line at ∞, the points of the form $[0 : z_1 : z_2]$. Thus, on $\mathbb{C}P^2 \setminus L$, we have coordinates $(z, w) = (Z_1/Z_0, Z_2/Z_0)$; set $p(z, w) = P(1, z, w)$. On X, this means that

$$\mathrm{d}p = \frac{\partial p}{\partial z}\,\mathrm{d}z + \frac{\partial p}{\partial w}\,\mathrm{d}w = 0,$$

so if we let

$$\theta = \frac{\mathrm{d}z}{\partial p/\partial w} = -\frac{\mathrm{d}w}{\partial p/\partial z},$$

then this is a holomorphic, nowhere-vanishing 1-form on X. Then, computing the multiplicities of the zeros of θ calculates $\chi(X)$ and therefore also $g(X)$ (see the textbook for more details). $\boxtimes$

Lecture 24.

The Riemann-Roch Theorem: Background and Statement: 3/30/16

"For a while, everything is going to be formal, until it won't be."

Today, we're going to start part III of the textbook, which centers on the Riemann-Roch theorem, which is really the most fundamental theorem of compact Riemann surfaces. (Uniformization would be the fundamental theorem for noncompact surfaces.)

[35] This is a special case of a *adjunction formula* characterizing the genus of a compact Riemann surface embedded in more general manifolds, and which admits an entirely topological proof.

Today, we'll discuss the background, and over the next two lectures, we'll discuss some applications. After that, we'll provide the proof, which begins geometrically and ends with the very analytic story of inverting a Laplacian.

For the rest of this lecture, X will denote a Riemann surface.

Divisors on a Riemann surface. Let $\mathcal{M}_X$ denote the field of meromorphic functions $X \to S^2$ and $\mathcal{M}_X^\times = \mathcal{M}_X \setminus \{0\}$. For any $x \in X$, we can define a valuation $v_x : \mathcal{M}_X^\times \to \mathbb{Z}$ as follows: in a local coordinate z centered at x, write $f \in \mathcal{M}_X^\times$ as its Laurent series: $f(z) = cz^m + O(z^{m+1})$. Then, we define $v_x(f) = m$.

Here are some basic properties of this valuation.

- $v_x(f) \geq 0$ iff f is holomorphic near x, and $v_x(f) > 0$ iff f is holomorphic near x and $f(x) = 0$.
- If $f, g \in \mathcal{M}_X^\times$, $v_x(fg) = v_x(f) + v_x(g)$ and $v_x(f+g) \geq \min(v_x(f), v_x(g))$.

It may seem a bit curious to make this an $\mathcal{M}_X^\times$-valued function, rather than an X-valued function, but these nice properties are why we do this; some divisors arise when we think of this as an X-valued function, though.

Definition. A *divisor* on X is a function $D : X \to \mathbb{Z}$ sending $x \mapsto D_x$ with finite support.

That is, $D_x = 0$ for all but finitely many x. A divisor is often written as the sum

$$D = \sum_{x \in X} a_x \cdot x,$$

where $a_x = D_x \in \mathbb{Z}$. Since $\mathbb{Z}$ is an abelian group, one can add divisors pointwise to get an abelian group $\mathrm{Div}(X)$. Given a $p \in X$, one example of a divisor is the one $D = p$, whose value at p is 1 and whose values everywhere else is 0.

Now, assume X is compact. Then, any nonzero meromorphic function $f \in \mathcal{M}_X^\times$ defines a divisor

$$(f) = \sum_{x \in X} v_x(f) \cdot x,$$

or $v_x(f)$ regarded as a function of x.

Definition. If $D = (f)$ for some $f \in \mathcal{M}_X^\times$, then D is a *principal divisor*.

The properties of v_x imply that $(fg) = (f) + (g)$ and $(1/f) = (-f)$; hence, the principal divisors form a subgroup $\mathrm{PDiv}(X) \leq \mathrm{Div}(X)$. We'll soon see that not every divisor is principal.

Definition. Two divisors D_1 and D_2 are *linearly equivalent* if $D_1 - D_2 \in \mathrm{PDiv}(X)$.

That is, they differ by the principal divisor associated to a nonzero meromorphic function. The group of linear equivalence classes is called the *class group* $\mathrm{C}\ell(X) = \mathrm{Div}(X)/\mathrm{PDiv}(X)$.[36]

Definition. A divisor D is *effective* if its coefficients are all nonnegative; this is denoted $D \geq 0$. We write $D_1 \geq D_2$ if $D_1 - D_2 \geq 0$, which defines a partial order on $\mathrm{Div}\, X$.

This partial order does not pass through to $\mathrm{C}\ell(X)$.

Definition. The *degree* of a divisor $D = \sum a_x \cdot x$ is

$$\deg D = \sum_{x \in X} a_x \in \mathbb{Z}.$$

Thus, $\deg : \mathrm{Div}\, X \to \mathbb{Z}$ is a homomorphism of groups.

Lemma 24.1. *If* $D \in \mathrm{PDiv}(X)$*, then* $\deg D = 0$.

Proof. It may come as no surprise that we'll use the argument principle, (4.8). Choose an $f \in \mathcal{M}_X^\times$; thus, $\deg((f)) = 0$ is equivalent to the number of poles and zeros of f, counted with multiplicity, being equal.

Choose disjoint closed discs $\Delta_1, \ldots, \Delta_n$, each containing a unique zero or pole of f. By the argument principle,

$$\deg((f)) = \sum_{i=1}^{N} \int_{\partial \Delta_i} \frac{\mathrm{d}f}{f}.$$

[36] A fact which we will not use: there is an isomorphism $\mathrm{C}\ell(X) \cong \mathrm{Pic}(X)$, where the latter is the *Picard group* of X.

Let X_0 be X minus the interiors of all these discs, so X_0 is a compact surface with boundary. Then, by Stokes' theorem,

$$\sum_{i=1}^N \int_{\partial\Delta_i} \frac{\mathrm{d}f}{f} = -\int_{\partial X_0} \frac{\mathrm{d}f}{f} = -\int_{X_0} \mathrm{d}\left(\frac{\mathrm{d}f}{f}\right),$$

because $\mathrm{d}f/f$ is nonsingular on X_0. However, since f is holomorphic on X_0, then $\mathrm{d}(\mathrm{d}f/f) = 0$: in local coordinates, $\mathrm{d}f = f'(z)\,\mathrm{d}z$, and so the only interesting thing can come from $\mathrm{d}\bar{z}$, but applying this to a holomorphic function gives you 0. Thus, $\deg((f)) = 0$. ⊠

The point is that we can use Stokes' theorem to bring the argument principle to general Riemann surfaces.[37] As a consequence, we get a degree homomorphism $\deg : \mathrm{C}\ell(X) \to \mathbb{Z}$.

Definition. Let D be a divisor; then, we will let $\mathscr{O}_X(D)$ denote the *sheaf of effective divisors*, the sheaf of functions defined as follows: if $U \subset X$ is open, then $\mathscr{O}_X(D)(U)$ is the set of $f : X \to S^2$ meromorphic on U and such that $D((f)) \geq 0$ on U.

These are the "functions with poles at worst D," meaning that if $D = \sum a_x x$ is effective, f is allowed to have poles only on $\operatorname{supp} D$, and the order of the allowed poles at x is bounded by a_x.

The key definition is this one, whose notation is motivated by sheaf cohomology.

Definition. Let $H^0(\mathscr{O}_X(D)) = \mathscr{O}_X(D)(X) = \{f \in \mathcal{M}_X \mid f = 0 \text{ or } f \in \mathcal{M}_X^\times \text{ and } D((f)) \geq 0\}$. This will sometimes be abbreviated $H^0(D)$.

For example, if P and Q are distinct divisors on X, then $H^0(2P - Q)$ is the space of meromorphic functions with at worst a double pole at P, and vanishing at Q.

In particular, $H^0(D)$ is a vector space, and if $g \in \mathcal{M}_X^\times$, $H^0(D) \cong H^0(D + (g))$ through the map $f \mapsto f/g$. We will let $h^0(D) = \dim H^0(D)$, so that we have a function $h^0 : \mathrm{C}\ell(X) \to \mathbb{Z}_{\geq 0}$.

The *Riemann-Roch problem* is to compute $h^0(D)$. It's worth putting this in context: *a priori*, for a general compact Riemann surface X, we don't know if there are any nonconstant meromorphic functions!

We can make a few observations quickly, though.

(1) If $\deg D < 0$, then $H^0(D) = 0$, because if $D + (f) \geq 0$, then $\deg(D + (f)) \geq 0$, but this was $\deg D$, which is a contradiction.
(2) If $D_1 \leq D_2$, then $D_2 = D_1 + E$, where E is an effective divisor. In this case, we have $H^0(D_1) \subset H^0(D_2)$: if $D = D' + mp$, where $p \notin \operatorname{supp}(D')$, then choose $f, g \in H^0(D + p) = H^0(D' + (m+1)p)$. Then, some linear combination $af + bg$ has valuation at most m at p (intuitively, we're canceling out large multiple poles), i.e. $af + bf \in H^0(D)$. Thus, $\dim(H^0(D + p)/H^0(D)) \leq 1$, and we can repeat this argument for the rest of $\operatorname{supp}(D) \setminus \operatorname{supp}(D')$.
(3) Combining these two, if D is linearly equivalent to an effective divisor, $h^0(D) \leq \deg D + 1$, since h^0 is invariant under linear equivalence.

On the other hand, D is linearly equivalent to an effective divisor iff $H^0(D) \neq 0$, so in general our upper bound is

$$h^0(D) \leq \max(0, \deg(D) + 1).$$

We still haven't said anything that implies there are nontrivial meromorphic functions on X yet. That's where the Riemann-Roch inequality comes in.

Theorem 24.2 (Riemann-Roch inequality[38])**.** *If X is a compact, connected Riemann surface and D is a divisor on X, then $h^0(D) \geq \deg D + 1 - g(X)$.*

The proof of this will have to wait.

Corollary 24.3. *Every compact Riemann surface admits nonconstant meromorphic functions.*

Proof. It suffices to prove this on a connected component of a compact Riemann surface X, so we can use Theorem 24.2. Let D be effective of degree D, so $h^0(D) \geq \deg D + 1 - g(X) > 1$, so $H^0(D)$ contains more than just constant functions. ⊠

[37] A related theorem which is distinct, but has a very similar proof, is the residue theorem for compact Riemann surfaces, which shows that the sum of the residues and orders of zeros of a meromorphic function on a compact Riemann surface is 0.

[38] This is also called the weak Riemann-Roch theorem.

Of course, the Riemann-Roch inequality is a more precise result than this, but it's still a nice fact to know.

Corollary 24.4. *If X is a compact, connected Riemann surface and $g(X) = 0$, then X is biholomorphic to S^2.*

Before this, all we knew was that X is diffeomorphic to S^2.

Proof. Take a $p \in X$, and regard it as a divisor. Then, $h^0(P) \geq 1 + 1 - 0 = 2$, so there is a nonconstant $f \in H^0(p)$, which has a single, simple pole. In particular, $f : X \to S^2$ has exactly one pole, so it's a degree-1 map, and therefore is biholomorphic. ⊠

There's also a strong (or full) Riemann-Roch theorem, which measures the discrepancy between $h^0(D)$ and $\deg D + 1 - g(X)$.

Definition. Let $H^0(\mathscr{K}_X(D))$ denote the vector space of meromorphic 1-forms on X with poles at worst D.

This is defined in local coordinates: if α is a meromorphic 1-form, then in local coordinates z, $\alpha = f(z)\,\mathrm{d}z$, and we require that $v_f(x) \leq m$ (meaning $z^m f(z)$ is holomorphic).

Theorem 24.5 (Riemann-Roch)**.** *Let D be a divisor on the compact Riemann surface X. Then,*

$$\dim H^0(\mathscr{O}_X(D)) - \dim H^0(\mathscr{K}_X(-D)) = \deg D + 1 - g.$$

This means that we're comparing the dimensions of the meromorphic functions with poles at worst D and meromorphic 1-forms with zeros at worst D.

Lecture 25.

: 4/1/16

Lecture 26.

Modular Forms: 4/3/16

"We're not going to prove Fermat's last theorem today; maybe next week."

Last time, we discussed some applications of the Riemann-Roch theorem motivated towards algebraic geometry, e.g. embedding surfaces in projective space.

Today, we're going to discuss modular forms, a related but different application. There are many references for this material:

- Diamond and Shurman, *A First Course in Modular Forms*, though they seem to make a meal of it.[39]
- Of course, Serre's *A Course in Arithmetic* is beautiful.
- Milne's online modular forms course notes are also a great reference.

Modular forms touch on a great deal of current research in representation theory and number theory, most famously in the modularity theorem of Wiles et. al., which led to the proof of Fermat's last theorem. But today, we're going to do some basics.

Throughout today, let $\Gamma = \mathrm{PSL}_2(\mathbb{Z})$, the modular group, and Γ_N be the kernel of the reduction mod N, $\Gamma \to \mathrm{PSL}_2(\mathbb{Z}/N)$ (so $\Gamma_1 = \Gamma$). Thus, Γ_N acts on the upper half-plane $\mathbb{H}$, and $Y_N = \mathbb{H}/\Gamma_N$. The compactification $X_N = Y_N \cup \{\text{cusps}\}$, which are a single orbit of Γ_N. The number of cusps is $|\Gamma_N|/N$. Through the j-invariant, we proved that $Y_1 = \mathbb{H}/\Gamma_1 \cong \mathbb{C}$, and $X_1 \supset Y_1$ is biholomorphic to $S^2 = \mathbb{C} \cup \{\infty\}$: there's one cusp.

Modular forms are special examples of weakly modular functions.

Definition. Let $k \in \mathbb{Z}$. A *weakly modular function* of *weight* $2k$ for the group Γ_N is a meromorphic function $f : \mathbb{H} \to S^2$ such that

$$\text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_N, \text{ we have } f(\gamma(t)) = (c\tau + d)^{2k} f(\tau). \tag{26.1}$$

[39] If you, like me, didn't recognize this colloquialism, it means "to make a [slight] mess of it."

This relation appears to depend on the choice of representative for $\gamma \in \mathrm{PSL}_2(\mathbb{Z}/N)$, but the condition that $\det(\gamma) = 1$ means that c and d are determined up to one choice of sign, meaning that taking $(c\tau+d)^{2k}$ makes the ambiguity disappear. This is why we only consider even weights.

(26.1) implies that $f(\tau+N) = f(\tau)$, since $\tau \mapsto \tau + N \in \Gamma_N$, so it's periodic, and therefore has a Fourier expansion: let $q^{1/N} = e^{2\pi i\tau/N}$. In that case, there's a meromorphic $\widetilde{f} : \mathbb{D}^* \to S^2$ such that $f(\tau) = \widetilde{f}(q^{1/N}$.

When $N = 1$ (meaning for Γ itself), (26.1) can be restated:

$$(26.1) \iff \begin{cases} f(\tau+1) = f(\tau) \\ f(-1/\tau) = \tau^{2k} f(\tau). \end{cases}$$

We'd like to relate these to a geometric construction of differentials on a Riemann surface.

Definition. Let X be a Riemann surface and $\mathscr{K}_X$ be its sheaf of holomorphic 1-forms (on each open $U \subset X$, it's the holomorphic 1-forms on U). If $k \geq 0$, a *holomorphic k-differential* is a holomorphic section η of $\mathscr{K}_X^{\otimes k}$. That is, η attaches to each $x \in X$ an element $\eta_x \in (T_x^{1,0}X)^{\otimes k}$ that varies holomorphically.

Locally, in a holomorphic coordinate z, there's a holomorphic f such that $\eta = f(z)(\mathrm{d}z)^{\otimes k}$. The holomorphic k-differentials form a vector space $H^0(\mathscr{K}_X^{\otimes k}) \cong H^0(kK_X)$ for a canonical divisor K_X, meaning it's the kind of vector space the Riemann-Roch theorem is good at analyzing.

In the same way, one can define *meromorphic k-differentials* to be those valued in the sheaf of meromorphic functions, or in local coordinates given by a meromorphic f.

We've talked about two different things today; it turns out they're closely related.

Proposition 26.2. *There's an isomorphism between the space of meromorphic k-differentials on X_N and the weakly modular functions for Γ_N with weight $2k$.*

Proof. The idea is that if η is a meromorphic k-differential on Y_N and $\pi : \mathbb{H} \to Y_N$ is the canonical projection, then let $\widetilde{\eta} = \pi^*\eta$. Hence, we can write $\widetilde{\eta} f(\tau)(\mathrm{d}\tau)^{\otimes k}$, so f is meromorphic. Out bijection will send $\eta \mapsto f$; by definition, this is linear and injective, but why is f weakly modular?

Suppose $\gamma = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_N$. Then, $\gamma^*\widetilde{\eta} = \widetilde{\eta}$, so

$$\begin{aligned} \gamma^*(\widetilde{\eta}) &= f(\gamma \circ \tau)(\mathrm{d}(\gamma \circ \tau))^{\otimes k} \\ &= f(\gamma \circ \tau)\left(\frac{\mathrm{d}\tau}{(c\tau+d)^2}\right)^{\otimes k} \\ &= \frac{f(\gamma \circ \tau)}{(c\tau+d)^{2k}}(\mathrm{d}\tau)^{\otimes k}, \end{aligned}$$

and since this is equal to $\widetilde{\eta} = f(\tau)(\mathrm{d}\tau)^{\otimes k}$, then f is invariant under γ.

Finally, one needs to check that one can reconstruct γ from f, but this is easy to do. ⊠

One interesting nuance of this is that in the case $N = 1$, it's worth thinking through what happens to the poles of η versus f at the *elliptic points* $P = [e^{\pi i/3}] \in Y_1 = \mathbb{H}/\Gamma$ and $Q = [i]$. These are the orbits of points for which Γ has nontrivial stabilizer: for P, it's of order 3, and for Q, it's of order 2.

If $Y = Y_1$ and $\pi : \mathbb{H} \to Y$ is the quotient by Γ, let $\widetilde{P}$ be a preimage of P. If z is a coordinate on $\mathbb{H}$ near $\widetilde{P}$ and w is one on Y near P, then $z \mapsto z^3 = w$, so if $\eta = f(w)(\mathrm{d}w)^{\otimes k}$, then $\widetilde{\eta} = f(z^3)\,\mathrm{d}(z^3)^{\otimes k}$. Thus, if $(\operatorname{div}\eta)_P = m$, then $(\operatorname{div}\widetilde{\eta})_{\widetilde{P}} = 2k + 3m$. Perhaps this is not what you expected, but at least this is only weird for $N = 1$. In the same way, for a preimage of Q, we get $k + 2m$.

One can also ask what happens at infinity, which leads to the notion of a modular form.

Definition.

- Let $f : \mathbb{H} \to S^2$ be meromorphic, and let $q^{1/N} = e^{2\pi i\tau/N}$, which is a holomorphic coordinate on Y_N near the cusp at $i\infty$ in $\mathbb{D}^*$. The function f is *holomorphic at $i\infty$* if $q^{1/N}$ has a removable singularity at 0 (so it can be "filled in").
- f is *holomorphic at the cusps* if for all $\gamma \in \Gamma$, $f \circ \gamma$ is holomorphic at $i\infty$.[40]
- A *modular form* f for Γ_N of *weight* $2k$ is a weakly modular function of weight $2k$ that is
  - holomorphic on $\mathbb{H}$, and

[40] It suffices to take one coset representative for each coset of Γ_N in Γ.

– holomorphic at the cusps in the sense above.

The modular forms of weight $2k$ for Γ_N form a vector space, denoted $M_{2k}(\Gamma_N)$.

Example 26.3. Recall that the Eisenstein series for a lattice $\Lambda_\tau = \mathbb{Z} \oplus \mathbb{Z}\tau$ is defined by

$$G_{2k}(\tau) = \sum_{\lambda \in \Lambda_\tau \setminus 0} \frac{1}{\lambda^{2k}}.$$

This is a modular form for Γ of weight $2k$.

We can interpret modular forms as holomorphic k-differentials on X_N with certain poles. In particular, we need to pay attention to the cusps. Near $i\infty$, there's a model for $\mathbb{H} \to Y_N$ given in local coordinates by $z \mapsto \zeta = e^{2\pi i z/N}$, and, as above, we can let $\zeta = q^{1/N} \in \mathbb{D}^*$. If $\eta = g(\zeta)(\mathrm{d}\zeta)^{\otimes k}$, then when we pull back to $\mathbb{H}$,

$$\widetilde{\eta} = \left(\frac{2\pi i}{N}\right)^k g(\zeta(z))\zeta(z)^k(\mathrm{d}z)^k,$$

because $\mathrm{d}\zeta = (2\pi i/N)\zeta\,\mathrm{d}z$.

This means $\widetilde{\eta}$ is holomorphic at $z = 0$ iff g has a pole of order at most k at $\zeta = 0$. Thus, we can allow k-differentials on X_N which have poles of order at most k at cusps, because they'll pull back to modular forms.

That is, if K_{X_N} is a canonical divisor for X_B and $N \geq 2$, $M_{2k}(\Gamma_N) = H^0(k(K_{X_N} + C))$, where C is the divisor that's the formal sum of the cusps of X_N. When $N = 1$, we have the slightly more complicated result

$$M_{2k}(\Gamma_1) = H^0\left(k(K_{X_1} + C) + \left\lfloor\frac{2k}{3}\right\rfloor P + \left\lfloor\frac{k}{2}\right\rfloor Q\right).$$

This gives a geometric interpretation of modular forms; now, Riemann-Roch computes dimensions: it tells us that if $g > 0$ and $\deg D > \deg K_X <$ then $h^0(D) = \deg D + 1 - g$. One can therefore compute that $\deg(K_{X_N} + C) = |G|/6 >)$, and $\deg(k(K_{X_N} + C)) > \deg K_{X_N}$ for $k > 0$.

Hence, for $N \geq 2$, Riemann-Roch applies, and therefore we can use some computations we've already done to conclude

$$\begin{aligned} \dim M_{2k}(\Gamma_N) &= \deg(k(K_{X_N} + C)) + 1 - g(X_N) \\ &= |G|\left(\frac{2k-1}{12} + \frac{1}{2N}\right). \end{aligned}$$

A similar computation goes through for $N = 1$.