Riemann Surfaces

UT Austin, Spring 2016
These notes were taken in UT Austin’s M392C (Riemann Surfaces) class in Spring 2016, taught by Tim Perutz. I live-Texed them using vim, so 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

1. Review of Complex Analysis: 1/20/16 3
2. Review of Complex Analysis, II: 1/22/16 5
3. Meromorphic Functions and the Riemann Sphere: 1/25/16 7
4. Analytic Continuation: 1/27/16 9
5. Analytic Continuation Along Paths: 1/29/16 12
6. Riemann Surfaces and Holomorphic Maps: 2/3/16 14
7. More Sources of Riemann Surfaces: 2/5/16 16
8. Projective Surfaces and Quotients: 2/8/16 18
9. Fuchsian Groups: 2/10/16 19
10. Properties of Holomorphic Maps: 2/15/16 21
11. The Riemann Existence Theorem: 2/22/16 23
12. Normalizing Plane Algebraic Curves: 2/24/16 24
13. The de Rham Cohomology of Surfaces: 2/26/16 26
14. de Rham Cohomology in Degree 1: 2/29/16 28
15. Poincaré Duality: 3/2/16 31
16. Holomorphic 1-Forms: 3/4/16 32
17. Hyperelliptic Riemann Surfaces: 3/7/16 34
18. Elliptic Curves and Elliptic Functions: 3/9/16 36
19. The j-invariant and Moduli of Elliptic Curves: 3/11/16 38
20. The Euler Characteristic and the Riemann-Hurwitz Formula: 3/21/16 40
22. Modular Curves for Principal Congruence Subgroups: 3/25/16 45
23. Euler Characteristics and Meromorphic 1-Forms: 3/28/16 47
24. The Riemann-Roch Theorem: Background and Statement: 3/30/16 50
25. Euler Characteristics and Meromorphic 1-Forms: 3/30/16 52
26. Modular Forms: 4/4/16 52
27. Klein’s Quartic Curve: 4/6/16 54
29. The Riemann-Roch Theorem: Exact Sequences of Sheaves: 4/11/16 58
30. Serre Duality: 4/13/16 61
31. Inverting the Laplacian: 4/15/16 62
32. Inverting the Laplacian, II: 4/18/16 65
33. Weyl’s Lemma: 4/20/16 67
34. Uniformization: 4/22/16 69
35. Uniformization, II: 4/25/16 71
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 1.1.** 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 \( \mathcal{O}(G) \), 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 \( H \subseteq G \) is open and \( f \in \mathcal{O}(G) \), then \( f|_H \in \mathcal{O}(H) \).
- The sum, product, quotient, and chain rules hold for holomorphic functions, so \( \mathcal{O}(G) \) 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 \( h \to 0 \)).

---

\(^1\)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 \( \mathbb{C} \to A \), which in this case is the constant functions.
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:

\[
(1.2) \quad \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad \text{and} \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}.
\]

The content of this is exactly that \( D_z f \) is complex linear.

Conversely, suppose \( f : G \to \mathbb{C} \) is differentiable in the real sense. Then, if it satisfies (1.2), 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 1.3.** Let \( A(z) = \sum_{n=0}^{\infty} a_n(z - c)^n \) be a \( \mathbb{C} \)-valued power series centered at \( a \in \mathbb{C} \). Then, its radius of convergence is \( R = \sup \{ |z - c| : A(z) \text{ converges} \} \), which may be 0, a positive real number, or \( \infty \).

**Theorem 1.4.** Suppose \( A(z) = \sum_{n=0}^{\infty} a_n(z - c)^n \) has radius of convergence \( R \). Then:

1. \( R^{-1} = \lim \sup |a_n|^{1/n} \);
2. \( A(z) \) converges absolutely on \( D(c, R) \) to a function \( f(z) \);
3. \( f \) is continuous on \( D(c, R) \);
4. \( A(z) \) converges uniformly on smaller discs \( D(c, r) \) for \( r < R \);
5. \( A(z) \) converges to a function \( g(z) \) on \( D(c, R) \);
6. \( f \in \mathcal{O}(D(c, R)) \) 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 \( \gamma \) such that \( \gamma(a) = \gamma(b) \).

If \( \gamma \) 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

\[
\int_\gamma f = \int_a^b f(\gamma(t)) \gamma'(t) \, dt.
\]

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 1.5** (Fundamental theorem of calculus). If \( F \in \mathcal{C}(G) \) and \( \gamma : [a, b] \to G \) is a path, then

\[
\int_\gamma F'(z) \, dz = F(\gamma(b)) - F(\gamma(a)).
\]

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

Now, an extremely important theorem.

**Definition 1.6.** 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 1.7 (Cauchy). If $G$ is a star-domain, $\gamma$ is a loop in $G$, and $f \in \mathcal{O}(G)$, then $\int_{\gamma} f = 0$. Indeed, $f = F'$, where

$$F(z) = \int_{[z_0, z]} f.$$  

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.

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 $\gamma_0, \gamma_1 : [a, b] \Rightarrow G$ be $C^1$ loops that are $C^1$ homotopic through loops in $G$. Then, for all $f \in \mathcal{O}(G)$, $\int_{\gamma_0} f = \int_{\gamma_1} f$.

Proof sketch. Fix a $C^1$ homotopy $\Gamma : [a, b] \times [0, 1] \rightarrow G$ 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 $\Gamma$ 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 $\Gamma$ 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 $f \in \mathcal{O}(G)$ 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,

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

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 \( f \in \mathcal{O}(G) \), then

\[
f(z) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(w)}{w-z} \, dw.
\]

**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 \( \int_{C(z,\delta)} dz/z = 2\pi i \), one has that

\[
\frac{1}{2\pi i} \int_{\partial D} \frac{f(w)}{z-w} \, dw - f(z) = \frac{1}{2\pi i} \int_{\partial D} \frac{f(w) - f(z)}{w-z} \, dw.
\]

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

\[
= \frac{1}{2\pi i} \int_{C(z,r)} \frac{f(w) - f(z)}{w-z} \, dw,
\]

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. \( \square \)

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

\[
f(z) = \sum_{n \geq 0} a_n(z-c)^n, \quad \text{where} \quad a_n = \frac{1}{2\pi i} \int_{\partial D} \frac{f(z)}{(z-c)^{n+1}} \, dz.
\]

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

\[
f(z) = \frac{1}{2\pi i} \int_{C(z,\delta)} f(w) \, dw
\]


\[
= \int_{C(c,R')} \frac{f(w)}{w-z} \, dw
\]

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

\[
f(z) = \frac{1}{3\pi i} \int_{C(c,R')} \frac{f(w)}{w-c} \frac{1}{1 - \frac{z-c}{w-c}} \, dw
\]

\[
= \frac{1}{2\pi i} \int_{C(c,R')} \frac{f(w)}{w-c} \sum_{n \geq 0} \frac{(z-c)^n}{(w-c)^n} \, dw.
\]

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

\[
= \sum_{n \geq 0} \frac{1}{2\pi i} \left( f \left( \frac{f(w)}{(w-c)^{n+1}} \right) \right) (z-c)^n.
\]  \( \square \)

One application of this is to understand zeros of holomorphic functions. If \( f \in \mathcal{O}(G) \) 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 \( c \),

\[
f(z) = (z-c)^m \sum_{n \geq m} a_n(z-c)^{n-m}.
\]

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 \( \text{mult}(f,c) \). In particular, if \( f(c) \neq 0 \), then \( m = 0 \).
Theorem 2.5. If $G$ is a connected open set and $f \in \mathcal{O}(G)$ 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 $f(z) = (z - c)^mg(z)$, 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. \hfill \Box

Definition 2.6. A function $f \in \mathcal{O}(\mathbb{C})$, so holomorphic on the entire plane, is called \textit{entire}.

Theorem 2.7 (Liouville). A bounded, \textit{entire} function is \textit{constant}.

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

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

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. \hfill \Box

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 $F \in \mathcal{O}(G)$, and so it’s analytic, and therefore it has derivatives of all orders. Thus, $F' = f$ is holomorphic. \hfill \Box

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

Theorem 3.3. Suppose $f \in \mathcal{O}(G \setminus \{z_0\})$ 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 $f$. \hfill \Box

Definition 3.4.

- If $z_0 \in G$ and $f \in \mathcal{O}(G \setminus \{z_0\})$, then $f$ has a \textit{pole} at $z_0$ if there’s an $m \in \mathbb{N}$ such that $(z - z_0)^mf(z)$ is bounded near $z_0$ (and hence has a removable singularity there). The least such $m$ is called the \textit{order} of the pole.
A meromorphic function on $G$ is a pair $(\Delta, f)$ consisting of a discrete subset $\Delta \subset G$ and an $f \in \mathcal{O}(G \setminus \Delta)$ 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 3.5.** The Riemann sphere $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ 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 $\phi : \hat{\mathbb{C}} \to S^2 = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = 1\}$ 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 $\hat{\mathbb{C}}$ 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.](http://www.math.rutgers.edu/~greenfie/vnx/math403/diary.html)

**Definition 3.6.** A continuous map $f : G \to S^2$ is holomorphic if for all $z \in G$, either

- $f(z) \notin \mathbb{C}$ (so it doesn’t hit $\infty$) and $f : G \to \mathbb{C}$ is holomorphic, or
- if $f(z) \in \hat{\mathbb{C}} \setminus \{0\}$, 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 $0$ nor $\infty$, then both criteria hold, and are equivalent (since $1/z$ is holomorphic on any neighborhood not containing zero).

**Proposition 3.7.** 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. \(\Box\)

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

**Definition 3.8.** 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 $\infty$ 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 $\infty$, we take reciprocals and talk about $0$.

**Example 3.9.** Every rational function $p \in \mathbb{C}(z)$ is meromorphic, and extends to a holomorphic map $S^2 \to S^2$. 
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.

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 $\infty$, as in Figure 3.

**Example 3.10.** 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 $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1}$). Thus, there’s a homeomorphism $\text{SL}_2(\mathbb{R})/\{\pm I\}$ to the group of Möbius transformations.

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 3.11.** 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 $f \in \mathcal{O}(G)$, 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 $f$.

**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 $\Gamma$-function

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

Though “holomorphic continuation” would make more sense, tradition gives us the term “analytic continuation.”
on the open set $\text{Re} \, z > 0$. 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

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

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

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

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

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

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

$$
g_n'(z) = \int_{1/n}^{n} \frac{\partial}{\partial z} \left( t^{z-1} e^{-t} \right) \, dt = (z-1)g_n(z-1).
$$

If $a > 0$, then $g$ converges uniformly on the strip $a < \text{Re} \, z < b$ — 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 $e^{x-1}e^{-t} \leq e^{-t/2}$ for $t$ sufficiently large. For $t < 1$, one should compare it to the integral of $t^{z-1}$. Then, we need to use the following theorem.

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

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 $\Gamma$.

**Theorem 4.5.** $\Gamma$ 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 $\Gamma_0$ be that integral from 0 to 1, and $\Gamma_\infty$ be the integral from 1 to $\infty$. Then, the argument above shows that $\Gamma_\infty \in \mathcal{O}(\mathbb{C})$, so the only extension that we actually need to make is of

$$
\Gamma_0(z) = \int_{0}^{1} t^{z-1} e^{-t} \, dt.
$$

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

$$
\Gamma_0(z) = \underbrace{\int_{0}^{1} t^{z-1} (e^{-t} - e_n(t)) \, dt} + \int_{0}^{1} t^{z-1} e_n(t) \, dt.
$$

$$
= \Gamma_n(z) + \sum_{j=0}^{n-1} \frac{(-1)^j}{j!(z+j)}.
$$

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

$$
\int_{0}^{1} t^{z-1} (e^{-t} - e_n(t)) \, dt
$$

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

\(^4\) A pole is simple if it’s degree 1.
It goes without saying that $\Gamma$ 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.

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, including 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 $z_0 \subset \mathbb{C}$ 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 $\phi$, i.e. $\{(z, \phi(z)) \mid z \in D_1\}$.

An analogue of this function holds for $C^1$ real functions (or $C^\infty$ 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 $f \in \mathcal{O}(G)$ and $\overline{D}$ is a closed disc in $G$ with $f(z) \neq 0$ on $\partial \overline{D}$, then

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

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

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

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, \( \text{mult}(f_{z_0}; w_0) = 1 \), 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 \( \partial D_2 \). Thus, there’s a disc \( D_1 \) centered at \( z_0 \) such that for all \( z \in D_1 \), \( |f_z| > \delta \) on \( \partial D_2 \) because \( P \) is continuous. Hence,
\[
\frac{1}{2\pi i} \int_{\partial D_2} \frac{f'_{z_0}}{f_{z_0}} = 1,
\]
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 \( \phi \) is holomorphic. By (4.9),
\[
\phi(z) = \frac{1}{2\pi i} \int_{\partial D_2} \frac{w f'_z(w)}{f_z(w)} \, dw = \frac{1}{2\pi i} \int_{\partial D_2} \frac{w}{P(z, w)} \frac{\partial P}{\partial z}(z, w) \, dw.
\]
Hence, \( \phi \) 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 5.1.** If \( \phi \) 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 \( \phi \) along \( \gamma \) consists of a pair \((U_t, \phi_t)\) for all \( t \in [0, 1] \), where \( U_t \) is a neighborhood of \( \gamma(t) \) and \( \phi_t \in \mathcal{O}(U_t) \) such that:
- \( \phi_0 = \phi \) on \( U_0 \cap U \), and
- the different \( \phi_t \) should agree, in the sense that for all \( s \in [0, 1] \), there’s a \( \delta > 0 \) such that if \( |t - s| < \delta \), then \( \phi_s \) and \( \phi_t \) agree on \( U_s \cap U_t \).

![Figure 4. Analytic continuation along a path: on sufficiently close circles, the extensions must agree.](image)

Note, however, that if \( \gamma \) intersects itself, then there’s no requirement for the extensions to agree on those overlaps (if \( \delta \) 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 5.2.** The logarithm is a very good example. Start with a branch of \( \log \) defined on some open set \( U_0 \), so \( \log(r e^{i\theta}) = \log r + i\theta \), 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 uniquely 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 \( \arg_{\gamma} \). Then, define \( \log_{\gamma}(z) = \log|z| + \arg_{\gamma(t)}(z) \), which defines a continuation of the logarithm around \( \gamma \).

---

\(^5\)One can think of this in terms of the theory of covering spaces, which is one reason this function lifts.
Example 5.3. 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 \( \gamma \) with image in \( \mathbb{C} \setminus \{ -1 \} \) by setting \( \phi_\gamma(z) = \exp((1/2) \log_n(1 + z)) \). However, if \( \gamma(t) = -1 + e^{2\pi i t} \), then \( \gamma \) winds around \(-1\), and when it returns to a point, the extension of \( \phi \) has a different value!

Example 5.4. 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 \( [p,q] \). 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 \([p,q] \), 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 5.5. \( \sum u_n z^n \) 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 5.6 (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 \([p,q] \) and \( U_n = \sum U_n z^n \) is a solution to \([P,Q] \) 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 5.5. 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 \([p,q] \) to
\[
\sum_{n \geq 0} |p_n| z^n, \sum_{n \geq 0} |q_n| z^n \quad \text{and} \quad \sum_{n \geq 0} \frac{M}{r^n} z^n, \sum_{n \geq 0} \frac{M}{r^n} z^n.
\]
It makes sense to compare this to \( M/(1-z/r), M/(1-z/r)^2 \) i.e. the equation
\[
u'' + Mu'/ \frac{1-z}{r} + \frac{Mu}{(1-z/r)^2} = 0.
\]
This last equation has an explicit solution \( \mu/(1-z/r) \) for some \( \mu \), 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 5.7. Show that if \( p,q \in \mathcal{O}(G) \) and \( \gamma : [0,1] \to G \), then any solution to \([p,q] \) has a solution along \( \gamma \) through solutions to \([p,q] \).

\footnote{If you’re typing notes, feel free to call it something else, like \( L_{p,q} \).}
Monodromy. If $\gamma$ is now a loop in $G$, so $\gamma(0) = \gamma(1) = z_0$, then analytic continuation around $\gamma$ 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 5.8.** $M_{\gamma}$ depends only on the homotopy class of $\gamma$ (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 oxymoronically-sounding notion of regular singular points. The prototype is the following, simpler equation:

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

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 $\alpha_1$ and $\alpha_2$ are distinct, then $(z^{\alpha_1}, z^{\alpha_2})$ is a basis for $V$ (the solutions near $1$), and if $\gamma$ is the unit circle, then the monodromy matrix in this basis is

$$(5.10) \quad M_{\gamma} = \begin{bmatrix} e^{2\pi i \alpha_1} & 0 \\ 0 & e^{2\pi i \alpha_2} \end{bmatrix}.$$  

If $\alpha$ 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 $(5.9)$ has monodromy.

This generalizes quite naturally.

**Definition 5.11.** 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 $\tilde{p}, \tilde{q}$ holomorphic in a neighborhood of $0$ such that $p(z) = A/z + \tilde{p}(z)$ and $q(z) = B/z^2 + C/z + \tilde{q}(z)$. Thus, the indicial equation is $\alpha(\alpha - 1) + A\alpha + B = 0$.

**Proposition 5.12.** If there are indicial roots $\alpha_1$ and $\alpha_2$ 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 $(5.10)$.

---

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 6.1.** 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.

\(^7\)Each pair $(U_\alpha, \phi_\alpha)$ is called a chart.
Figure 5. A transition map between charts for a Riemann surface $X$.

Remark.

- If $p \in U_\alpha$, we can think of $\phi_\alpha$ 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.

- 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 $\begin{bmatrix} \text{Re}a & -\text{Im}a \\ \text{Im}a & \text{Re}a \end{bmatrix}$, 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 6.2 (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 these are, but also how to map between them.

Definition 6.3. 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 $\phi_\alpha$ and $\psi_\beta$, $\psi_\beta \circ f \circ \phi_\alpha^{-1}$ 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 6.4.

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).

\(^8\)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.
(2) The Riemann sphere $S^2 = \hat{\mathbb{C}}$ is a Riemann surface with an atlas of two charts: the copy of $\mathbb{C}$ inside $\hat{\mathbb{C}}$ 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 $ad - be = 1$, are biholomorphic.

(3) $\mathbb{D}$ will denote the unit disc $D(0, 1)$, and $\mathbb{H} = \{z \in \mathbb{C} | \text{Im}(z) > 0\}$, 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 $\mu$ is continuous).

(4) Let $P(z, w)$ be holomorphic in both $z \in \mathbb{C}$ and $w \in \mathbb{C}$. Then, $X = \{(z, w) \in \mathbb{C}^2 | P(z, w) = 0\}$; 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 $w_0$ and $w_0$, respectively, and a holomorphic $\phi : D_1 \to D_2$ with $\phi(z_0) = w_0$ and $U = X \cap (D_1 \times D_2) = \text{graph}(\phi)$.

   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, $\phi^{-1}$ sends $z \mapsto (z, f(z))$ and $\psi$ 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 $\pi^{-1}(U) \to U$, where $D$ has the discrete topology.

If $X$ is a Riemann surface, then $Y$ acquires a Riemann surface structure that makes $\pi$ 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 $\pi$ 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 $\pi : \tilde{X} \to X$, with $X$ simply connected. We’ll see this again, so it’s useful to remember the construction: $\tilde{X} = \{ (x, \gamma) | x \in X, \gamma : [0, 1] \to X, \gamma(0) = x_0, \gamma(x) = x \}$ modulo the equivalence relation $(x, \gamma) \sim (x', \gamma')$ if $x = x'$ and $\gamma$ and $\gamma'$ are homotopic. The topology on $X$ 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 $g : \tilde{X} \to \tilde{X}$ 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 $\tilde{X}/G$, 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 $f \in O(U)$. Then, an abstract analytic continuation (AAC) of $f$ is $\mathcal{F} = (X, x_0, \pi, F)$, 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 $\sigma$ is a local right inverse to $\pi$ near $z_0$ (so $\pi \circ \sigma = \text{id}$), then we require that $F \circ \sigma = f$ in a neighborhood of $z_0$. \footnote{Up to making the neighborhood smaller, $\sigma$ is unique anyways; thus, it’s unique as the germ of a function.}
There’s a natural notion of a morphism between two abstract analytic continuations \( \mathcal{X} \) and \( \mathcal{X}' \) of \((U, f)\) (respectively given by \((X, x_0, \pi, F)\) and \((X', x'_0, \pi', F')\)): a holomorphic map \( \phi : X \to X' \) respecting all the structure, i.e. it intertwines \( \pi \) and \( \pi' \), as well as \( F \) and \( F' \). In particular, the test functions and \( \mathcal{X} \) are a category \( \mathcal{C}_f \).

**Definition 7.1.** A terminal object in a category \( \mathcal{C} \) is an \( X \in \mathcal{C} \) such that any \( X' \in \mathcal{C} \) maps to \( X \) in a unique way.

**Proposition 7.2.** \( \mathcal{C}_f \) has a terminal object \( \mathcal{X}_f = (X_f, x_0, \pi, F) \).

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 \( \tilde{\gamma} : [0, 1] \to X \) (meaning \( \pi \circ \tilde{\gamma} = \gamma \)) iff \( f \) has an analytic continuation along the path \( \gamma \).

\( \mathcal{X}_f \) 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 \( f_{\gamma} \) along \( \gamma \), 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 \( \gamma \to f_{\gamma}(\gamma(1)) \).

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, \( X_f \to \mathbb{C}^* \) 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, \( \mathbb{CP}^n \), 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, \( \mathbb{CP}^n = (\mathbb{C}^{n+1} \setminus 0)/\mathbb{C}^* \), 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 \( S^{2n+1}/U(1) \), by scaling each vector in \( \mathbb{C}^{n+1} \setminus 0 \) to a unit vector).

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

We can write \( \mathbb{CP}^n = U_0 \cup U_1 \cup \cdots \cup U_n \), 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 \( \mathbb{CP}^n \).

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

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

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 \( X = \{(z_1, z_2) \mid p(z_1, z_2) = 0\} \subset \mathbb{CP}^2 \) as a subset in \( U_0 \subset \mathbb{C}^2 \); then, the closure of \( X \) in \( \mathbb{CP}^2 \) 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 7.3.** 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 \( \overline{X} \subset \mathbb{CP}^2 \).

\(^{10}\)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.
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.

**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 \( \mathbb{C}^2 \to \mathbb{C}P^2 \) by \((z, w) \to [1 : z_1 : z_2]\), then \( X \subset \bar{X} \), 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 \( \bar{X} \) is closed, then \( \bar{X} \) is compact.

We’d like to make \( \bar{X} \) into a Riemann surface with \( X \hookrightarrow \bar{X} \) holomorphic (in other words, extending the Riemann surface structure on \( X \)), which is the content of Proposition 7.3.

*Proof of Proposition 7.3.* Take \( q = [Z_0 : Z_1 : Z_2] \in \bar{X} \). One of the \( Z_i \) 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 \( P(Z_1, \ldots, Z_n) \) 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 \( X_1 \subset \bar{X} \), where \( Z_1 \neq 0 \) and \( Z_2 \subset \bar{X} \), where \( Z_2 \neq 0 \), and we know that \( \bar{X} = X_0 \cup X_1 \cup X_2 \).

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, \( \bar{X} \) is just \( X \) along with the “points at infinity,” which are the points in \( \bar{X} \) where \( Z_0 = 0 \). If you start out with \( X = V(p) \), where

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

so that

\[P(Z_0, Z_1, Z_2) = \sum_{i,j=0}^d a_{ij} Z_0^{d-i-j} Z_1^i Z_2^j,
\]

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

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

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. \( \text{Aut}(S^2) = \text{PSL}_2(\mathbb{C}) \), which is also the group of Möbius maps. This is ultimately because any holomorphic map \( S^2 \to S^2 \) is a rational function.

\(^{11}\)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.
(2) \( \text{Aut } \mathbb{C} \) is the set of maps in \( \text{PSL}_2(\mathbb{C}) \) 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

(3) Then, \( \text{Aut}(\mathbb{H}) = \{ \mu \in \text{PSL}_2(\mathbb{C}) \mid \mu(\mathbb{H}) = \mathbb{H} \} \), which is also \( \text{PSL}_2(\mathbb{R}) \). Thus, understanding Riemann surfaces often really boils down to understanding free subgroups of this group. This also subsumes \( \text{Aut}(\mathbb{D}) \), which is the same, because \( \mathbb{H} \cong \mathbb{D} \) under a suitable Möbius map, so \( \text{Aut}(\mathbb{D}) \) and \( \text{Aut}(\mathbb{H}) \) are conjugate in \( \text{PSL}_2(\mathbb{C}) \). In fact, \( \text{Aut}(\mathbb{D}) = \text{PSU}(1, 1) \); here, \( \text{SU}(1, 1) \) is the group of unitary matrices preserving a signature-(1, 1) quadratic form. That is, if \( \langle z, w \rangle = z_1w_1 - z_2w_2 \), for \( z, w \in \mathbb{C}^2 \), then \( \text{SU}(1, 1) = \{ A \in \text{SL}_2(\mathbb{C}) \mid \langle A_z, Aw \rangle = \langle z, w \rangle \} \), or more explicitly,

\[
\text{SU}(1, 1) = \left\{ \begin{pmatrix} \alpha & \beta \\ \beta^* & \alpha^* \end{pmatrix} \mid |\alpha|^2 - |\beta|^2 = 1 \right\}.
\]

Then, \( \text{PSU}(1, 1) \) is the image of this in \( \text{PGL}_2(\mathbb{C}) \), i.e. \( \text{SU}(1, 1)/\{ \pm 1 \} \).13

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

So instead, let’s look at \( \text{Aut}(\mathbb{C}) \). For example, \( \mathbb{Z} \mapsto \text{Aut}(\mathbb{C}) \) by \( n \cdot z = z + n\lambda \), for a fixed \( \lambda \in \mathbb{C}^* \). That is, \( \mathbb{Z} \) acts by translation, scaled by \( \lambda \). If \( \Gamma < \text{Aut}(\mathbb{C}) \) is this subgroup, then \( X = \mathbb{C}/\Gamma \) 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 \( \pi \) 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 \( \text{Im } \lambda > 0 \), 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 \( r = 1/3 \cdot \text{min}(1, \text{Im } \lambda) \) 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.](image)

---

### 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 \( \text{PSL}_2(\mathbb{R}) = \text{Aut}(\mathbb{H}) \), or, equivalently, \( \text{PSL}(1, 1) = \text{Aut}(\mathbb{D}) \), as we established an isomorphism between these two groups, and in fact a conjugacy inside \( \text{PSL}_2(\mathbb{C}) \).14

12You can prove this using the Casorati-Weierstrass theorem on essential singularities, or think of this as holomorphic rational functions.

13Usually, 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 \( \text{Aut}(\mathbb{H}) \), \( \text{Aut}(S^2) \), and \( \text{Aut}(\mathbb{C}) \).

14\( \text{SL}_2(\mathbb{C}) \) is a four(-complex)-dimensional, complex Lie group, and \( \text{Aut}(\mathbb{D}) \) and \( \text{Aut}(\mathbb{H}) \) are 3-dimensional, noncompact, real Lie groups.
These groups have a natural topology to them. First, $\text{SL}_2(\mathbb{C}) \subset \mathbb{C}^4$ (since it’s a group of $2 \times 2$ matrices), so it has the subspace topology. Thus, $\text{PSL}_2(\mathbb{C})$ has the quotient topology, and as subspaces, $\text{Aut}(\mathbb{D})$ and $\text{Aut}(\mathbb{H})$ gain the subspace topology.

**Definition 9.1.** A Fuchsian group $^{15} \Gamma$ is a discrete subgroup of $\text{Aut}(\mathbb{H})$.

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

We’d like to study quotients $\mathbb{H}/\Gamma$, or $\mathbb{D}/\Gamma^*$; these turn out to be all nice Riemann surfaces. In general, we should use the hyperbolic structure on $2\mathbb{M}$ (Riemann Surfaces) Lecture Notes

**Proposition 9.3.**

**Definition 9.1.** Let $\Gamma \leq \text{Aut}(\mathbb{D})$ and therefore have no fixed points in $\mathbb{D}$; generally not in $\mathbb{H}$.

This means that either $\Gamma$ is trivial, or $\Gamma$ is a Fuchsian group.

To understand $\gamma \in \text{Aut}(\mathbb{H})$, 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. $\gamma$ 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 $\gamma \neq \text{id}$, then after some case-checking, one sees that 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 $\Delta = (d - a)^2 + 4bc = (tr A)^2 - 4$, where $A = [a b; c d]$. Thus, there are three cases:

  1. $\gamma$ is an elliptic element if $\Delta < 0$, i.e. $\text{tr}^2(A) < 4$. In this case, $c \neq 0$, and there are two fixed points, one in $\mathbb{H}$ and its conjugate in $\mathbb{H}$ (that is, the lower half-plane). By conjugating into $\gamma' \text{Aut}(\mathbb{D})$, there’s a unique fixed point in $\mathbb{D}$, and after a conjugation this is 0. But this means $\gamma'$ is a rotation of the disc (e.g. by the Schwarz lemma); there’s not quite such a simple description of $\gamma \in \text{Aut}(\mathbb{H})$, but the point is that elliptic elements are conjugates of rotations. In particular, they may have finite or infinite order.

  2. $\gamma$ is a parabolic element if $\Delta = 0$, i.e. $\text{tr}^2(A) = 4$. If $c = 0$, then $\gamma$ has only one fixed point which is in $\mathbb{R}$. If $c \neq 0$, then $A = [1 0; c 1]$, so $\infty$ is fixed. Thus, in either case, there’s one fixed point, and it’s in $\mathbb{R} \cup \{\infty\}$, and $\gamma$ is conjugate in $\text{PSL}_2(\mathbb{R})$ to a $\mu$ fixing infinity and sending 0 to $\pm 1$ (i.e. $\mu(z) = z \pm 1$). In particular, parabolic elements have infinite order.

  3. $\gamma$ is a hyperbolic element if $\Delta > 0$, so $\text{tr}^2(A) > 4$. In this case, either $c \neq 0$, so there are two distinct fixed points in $\mathbb{R}$, or $c = 0$ and $a \neq 0$, so there’s one fixed point in $\mathbb{R}$, and $\infty$ is also fixed. Thus, in either case, there are two distinct fixed points in $\mathbb{R} \cup \{\infty\}$; such a $\gamma$ is conjugate in $\text{PSL}_2(\mathbb{R})$ to a $\mu$ fixing both 0 and $\infty$. Thus, $\mu(z) = \lambda z$, where $\lambda > 0$ and $\lambda \neq 1$. Thus, $\gamma$ has infinite order.

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

**Example 9.2.** Let $p$ be a prime number, and let $\Gamma_p = \{A \in \text{SL}_2(\mathbb{Z}) \mid A \equiv \pm 1 \mod p\} \subset \text{SL}_2(\mathbb{R})$. Then, let $\Gamma_p = \Gamma_p/\pm I$, which is a subgroup of $\text{PSL}_2(\mathbb{R})$. An element $\gamma \in \Gamma_p$ has the form

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

where $a, b \in \mathbb{Z}$ and we don’t know what the off-diagonal entries are. Then, $\text{tr} \gamma = \pm((a + b)p + 2)$, which is generally not in $(-2, 2)$. In fact, if $p \geq 5$, then $|\text{tr} \gamma| \geq 2.1^{16}$ Thus, all elements of $\Gamma_p$ are elliptic or hyperbolic, and therefore have no fixed points in $\mathbb{H}$. Hence, $\Gamma_p$ acts freely on $\mathbb{H}$.

**Proposition 9.3.** Let $\Gamma \subset \text{Aut}(\mathbb{H})$ 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. $\gamma \in \text{stab}_p(z)$).
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 $N_2 \cap \Gamma \cdot N_1 = \emptyset$.

---

$^{15}$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.
Note. Part (2) says that the quotient is Hausdorff; if further $\Gamma$ 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.

Proof of Proposition 9.3, 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 $D$ 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 $\varepsilon$) and suppose $q, \gamma(q) \in D$ for some $\gamma \in \Gamma$.

$\gamma(z) = az + \beta/(\overline{\beta}z + \overline{\alpha})$, 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 $|\beta|$ arbitrarily small relative to $|\alpha|$.

The hypothesis we haven’t used yet is that $\Gamma$ 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 $\Gamma$ 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 $\varepsilon$ such that $2\varepsilon/(1-\varepsilon^2) < k$), 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 9.3, 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 $\text{stab}_\Gamma(z) \neq 1$.

The model case will be where $\Gamma_n = \mathbb{Z}/n$, which acts on $D(0;r)$ by $k \cdot z = e^{2\pi ki/n}z$. Hence, $D(0;r)/\Gamma_n \cong D(0;rn)$, by sending $|z| \mapsto zn$ (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 $\Gamma$ is any Fuchsian group, then $\text{stab}_\Gamma(0) = \Gamma_n$: it has to be a finite group of rotations (since it’s a discrete subgroup of $U(1)$), and for a general $z \in D$, 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 $f(x)$, $f'(x) \neq 0$.

$(1) \implies (2)$ by the chain rule: $f^{-1} \circ f = \text{id}$, 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 $\psi$ is a holomorphic chart near $f(x)$, then there’s a holomorphic chart $\phi$ near $x$ such that $f = \psi \circ f \circ \phi^{-1}$ takes the form $f(z) = z^k$, for some integer $k \geq 0$ independent of the chart.

---

17There 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.
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.

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

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

\[
\tilde{f}(z) = \sum_{n \geq k} a_n z^n = a_k z^k \sum_{m \geq 0} \left( \frac{a_{m+k}}{a_k} \right) z^m,
\]

where \( k > 1 \) and \( a_k \neq 0 \), so \( \tilde{f}(0) = f'(0) = 0 \). 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, \( \phi \) 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. \( \rlap{\checkmark} \)

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

Definition 10.3. 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 \( \text{crit}(f) \).

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 10.4. A critical value of \( Y \) is a point in the image of \( \text{crit}(f) \). 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 10.5. 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 \( \text{crit}(f) \) is the zero set of the holomorphic \( f' \), then it’s discrete (assuming \( f \) is nonzero). Thus, \( \Delta = f(\text{crit} f) \), 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 \( \deg f = \sum_{x \in f^{-1}(y)} |f'(x)| \); 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 \( k_x^{1/k} \) 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.

\(^{18}\)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.
Theorem 10.6. 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.

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. \(\Box\)

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, $\text{mon} : \pi_1(Y_0, y_0) \to \text{Aut } F$, 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 will 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 $p : \tilde{Z} \to Z$, 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 $H \leq \pi$: we have $p : \tilde{Z}/H \to Z$, where $H$ acts on $\tilde{Z}$ 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 $\text{mon} : \pi \to \text{Aut}(F)$, 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 $\pi$-set.\(^{20}\) Then, if we choose an $f \in F$, let $H = \text{stab}_\pi(f) \leq \pi$; then, we can recover $Y = \tilde{Z} \times_\pi F$. This fiber product is $\tilde{Z} \times F$ modulo the equivalence relation $(z, f) \sim (z \cdot g^{-1}, \text{mon}(g) \cdot f)$ 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 $\pi$-orbits, i.e. $\pi$-sets of the form $\pi/H$, with $H \leq \pi$, and with morphisms given by $\pi$-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 $\Delta = F(\text{crit } F) \subset Y$ of critical values, and
- for any basepoint $y_0 \in Y \setminus \Delta$, a monodromy homomorphism $\pi(Y \setminus \Delta, y_0) \to \text{Aut}(\pi^{-1}(y_0))$.

Moreover, on $Y \setminus \Delta$, $F$ is a covering map, and near a critical point $p \in \text{crit } F$, $F$ 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.

\(^{19}\)Recall that a pole of a function into $S^2$ is a point in the preimage of $\infty$.

\(^{20}\)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 \( \Delta = \text{crit } F \) and monodromy \( \rho \).

Remark. In fact, there’s a category of proper holomorphic maps to \( Y \) with critical values \( \Delta \), 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 \( \rho \); our job is to fill in the missing points of \( X \) which will map to \( \Delta \).

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

Let \( D \) be a small coordinate disc in \( \Delta \) centered at \( \delta \), 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 \( E_1, \ldots, E_r \), where the maps \( F_0 : E_j \to D^* \) is equivalent to the map \( m_j : D^* \to \mathbb{D}^* \) 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 \( m_j : D^* \to \mathbb{D}^* \) to \( m_j : \mathbb{D} \to \mathbb{D} \).

**Normalizing 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 \( \mathbb{C}^2 \subset \mathbb{CP}^2 \), we have a compactification \( \overline{X} \), the zero set of the homogenization of \( p \). In general, \( X \) and \( \overline{X} \) are singular; let \( S = \{(z, w) \in X \mid \frac{\partial p}{\partial z} = \frac{\partial p}{\partial w} = 0\} \) 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 \( \nu : X^* \to \mathbb{CP}^2 \) extending the inclusion \( X \setminus S \hookrightarrow \mathbb{CP}^2 \), with \( \text{Im}(\nu) = \overline{X} \).

\((X^*, \nu)\) is called the normalization of \( X \) in \( \mathbb{CP}^2 \). 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

---

\(^{21}\)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 \( \varphi \) is transitive. For example, \( \mathbb{Z}/n \) acts transitively on \( S_n \) by sending \( i \mapsto i + 1 \mod n \).
which are the set $S$ where both partials of $p$ vanish, and these sit inside $\mathcal{S} \subset \mathcal{X}$, where all three partials of $P$ vanish.

We know that $\mathcal{X} \setminus \mathcal{S}$ 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 $\mathcal{X}$, which is a compact $X^*$ along with a surjective, continuous map $\nu: X^* \to \mathcal{X}$, such that $\nu: \nu^{-1}(\mathcal{X} \setminus \mathcal{S}) \to \mathcal{X} \setminus \mathcal{S}$ is a biholomorphism, so we have a commutative diagram

\[
\begin{array}{ccc}
\mu^{-1}(\mathcal{X} \setminus \mathcal{S}) & \xrightarrow{\nu} & X^* \\
\downarrow & & \downarrow \\
\mathcal{X} \setminus \mathcal{S} & \xrightarrow{\nu \text{ incl.}} & \mathcal{X} \hookrightarrow \mathbb{C}P^2.
\end{array}
\]

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 $\pi: X \to \mathbb{C}$ given by $(z, w) \mapsto z$. We'd like to say that $\pi$ 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 $\mathbb{C}P^2$.

This is not going to work as stated, because $\pi$ 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 $\frac{\partial \pi}{\partial w} = 0$ everywhere, then $p = p(z)$ is an irreducible polynomial in one variable. By the fundamental theorem of algebra, this means it’s linear. Hence, $\mathcal{X}$ is already a Riemann surface, so we can let $X^* = \mathcal{X}$ and $\nu = \text{id}$, which satisfies the normalization property.

Having dealt with this trivial case, we can assume $\frac{\partial p}{\partial w}(x) = 0$ 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 $\pi$.

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

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

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 $\pi$. If $p(z, w) = w^2 - z^2(1 - z)$, we can work through this explicitly. Here, $F = \emptyset$ 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 $\Delta = E \cup \pi^+(\text{crit } \pi^+)$, and a monodromy map around each point of $\Delta$ which is the monodromy of $\pi^+$. The result is a compact Riemann surface $X^*$ and a degree-$d$ map $\pi^* : X^* \to S^2$, where $\pi^*$ is branched over $E \cup \pi^+(\text{crit } \pi^+)$.

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 $\infty$. 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 $\pi : \text{crit } \pi^* \to S$, which extends holomorphically over $\pi$ 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 $w = (1 - t^2)$. Then, $\nu$ 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 $[Z_1^3 : Z_0^2 - Z_1^2 : Z_0^3]$, which is exactly the homogenization.

---

**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 $\alpha = f(x, y) \, dx + g(x, y) \, dy$ 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 $\omega = f(x, y) \, dx \wedge dy$, 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 $d^1$:

$$
\begin{align*}
\Omega^0(S) & \xrightarrow{d^0} \Omega^1(S) & \xrightarrow{d^1} & \Omega^2(S).
\end{align*}
$$

---
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 \( d^j \alpha|_U = d^j \beta|_U \). If \( f, \alpha, \) and \( \beta \) are smooth functions on \( S \), then in local coordinates,

\[
\begin{align*}
d^0 f &= \frac{\partial f}{\partial x} \, dx + \frac{\partial f}{\partial y} \, dy, \\
d^1 (\alpha \, dx + \beta \, dy) &= \left( \frac{\partial \beta}{\partial x} - \frac{\partial \alpha}{\partial y} \right) \, dx \wedge dy.
\end{align*}
\]

In particular, \( d^1 \circ d^0 = 0 \). 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 \( \Omega^0, f \mapsto f \circ F \). The fundamental fact is that \( F^* \circ d^j = d^j \circ F^* \): certainly, plenty of operators satisfy \( d^2 = 0 \), 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: \( d^0 (fg) = d^0 f + f d^0 g \) and \( d^1 (f \alpha) = df \wedge \alpha + f \, d\alpha \), where \( f, g \in C^\infty(S) \) and \( \alpha \in \Omega^j(S) \).

**Definition 13.2.** 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).

\[
\Omega^0_c(S) \xrightarrow{d^0} \Omega^1_c(S) \xrightarrow{d^1} \Omega^2_c(S)
\]

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

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

If we use (13.3) 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 13.4.**

- \( H^0(S) \) is the space of locally constant functions; in particular, if \( S \) is connected, \( H^0(S) \cong \mathbb{R} \).
- Likewise, \( H^1_c(S) \) is the space of locally constant, compactly supported functions. If \( S \) is connected, \( H^1_c(S) \) 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 \( \phi : U \to \tilde{U} \subset \mathbb{R}^2 \) is an oriented chart. If \( \omega \in \Omega^2_c(S) \) has its support inside \( U \) and \( \omega|_U = \phi^*(f(x,y) \, dx \wedge dy) \), then

\[
\int_S \omega = \int_{\tilde{U}} f(x,y) \, dx \wedge dy.
\]

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 13.5 (Stokes).** If \( \alpha \in \Omega^1_c(S) \), then \( \int_S d\alpha = 0 \).
As a corollary, this means that integration factors through the quotient by \( \text{Im}(d^1) \) to define an integration map \( \int_S : H^2_c(S) \to \mathbb{R} \).

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

**Proposition 13.6.** \( \int_{\mathbb{R}^2} : H^2_c(\mathbb{R}^2) \to \mathbb{R} \) 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 \( f(x, y) \) is compactly supported, then \( f \) is the derivative of a compactly supported \( F \in C^\infty_c(\mathbb{R}) \); in fact,

\[
F(x) = \int_x^\infty f(t) \, dt,
\]

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 \( \text{supp}(x) \), we're just integrating the zero function, and if \( x \) is above it, then \( \int_g \) is integrated all of \( f \); by assumption this is just \( \alpha \). We're going to mimic this argument in two variables.

Now, suppose \( \omega \in \Omega^2_c(\mathbb{R}^2) \); since \( \mathbb{R}^2 \) is its own single chart, then \( \omega = f(x, y) \, dx \wedge dy \), where \( f \) is compactly supported. Let \( \omega \in C^\infty_c(\mathbb{R}) \) be such that \( f(x, y) \) is compactly supported. Let \( \omega \in \Omega^2_c(\mathbb{R}^2) \) be such that \( f(x, y) \) is compactly supported. Let \( I_y = \int f(x, y) \, dx \) and \( \tilde{f}(x, y) = f(x, y) - I_y \phi(x) \). Hence,

\[
\int_{-\infty}^{\infty} \tilde{f}(x, y) \, dx = I_y - I_y \int_{-\infty}^{\infty} \phi(x) \, dx = 0.
\]

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

\[
\int_{-\infty}^{\infty} I_y \, dy = \int_{-\infty}^{\infty} f(x, y) \, dx \, dy = 0,
\]

since \( f \) is compactly supported, and therefore \( I_y \cdot \phi(x) = \frac{\partial}{\partial y} Q \) for some compactly supported \( Q \). This means \( \omega = f(x, y) \, dx \wedge dy = d(\rho \, dy + Q \, dx) \), so \( \omega = 0 \) in \( H^2_c(\mathbb{R}) \).

This generalizes to the following.

**Theorem 13.7.** Let \( S \) be a connected, oriented surface. Then, \( \int_S : H^2_c(S) \to \mathbb{R} \) is an isomorphism.

**Proof.** It's clear that \( \int_S \) is surjective. Let \( \rho \in \Omega^2_c(S) \) be such that \( \int_S \rho = 0 \). We'd like to show that \( \rho = 0 \) for an \( \alpha \in \Omega^1_c(S) \). Then, \( \rho \) 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 13.6 (and the diffeomorphism-invariance of \( \Omega^2_c \)).

If instead \( n > 1 \), let \( U_1, \ldots, U_n \) 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 \( K \cap U \cap V \neq \emptyset \), so take a \( p \in U \cap V \) and let \( \beta \in \Omega^2_c(U \cap V) \) 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,

\[
f_1 \rho - \left( \int_{U \cup V} f_1 \rho \right) \beta
\]

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

---

\(^{22}\)This was left as an exercise in lecture.
Last time, we proved Theorem 13.7, that integration provides an isomorphism between $H^2(S) \to \mathbb{R}$ for all smooth, oriented surfaces $S$. Today, we’re going to investigate $H^1(S)$: what does “closed 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 $\alpha$. 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 $\mathcal{U}$ of $S$ such that every $U \in \mathcal{U}$ is diffeomorphic to $\mathbb{R}^2$, and hence $H^1(U) = 0$. Hence, if $\alpha \in \Omega^1(S)$ and $d\alpha = 0$, 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 $d(f_i - f_j) = 0$ on $U_{ij}$, and hence $c_{ij} = f_i - f_j$ is locally constant.

**Definition 14.2.**

- A Čech 1-cochain $\zeta$ for $(S, \mathcal{U})$ is an assignment of a locally constant $\zeta_{ij}$ on $U_{ij} = U_i \cap U_j$ for all $U_i, U_j \in \mathcal{U}$. These form a vector space, denoted $\check{C}^1(S; \mathcal{U})$ or $\check{C}^1$.
- A Čech 1-cochain is a 1-cocycle if for all triples $U_i, U_j, U_k \in \mathcal{U}$, 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-coycle is a 1-coboundary if there exist locally constant functions $f_i \in \Omega^0(U_i)$ for each $U_i \in \mathcal{U}$ 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 $\check{H}^1(S; \mathcal{U}) = \check{Z}^1/\check{B}^1$.

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

**Definition 14.3.** A cover $\mathcal{U}$ of a surface $S$ is acyclic if it’s locally finite and for all nonempty intersections $U_{i_1, i_2 \ldots i_N} = U_{i_1} \cap \ldots \cap U_{i_N}$, we have $H^1(U_{i_1, i_2 \ldots i_N}) = 0$.

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),

**Theorem 14.4.** If $\mathcal{U}$ 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 $\mathcal{U}$, so long as it satisfies the hypotheses.

**Proof of Theorem 14.4.** 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, $d(f_i - a_i) = \alpha$ 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, $dg = \alpha$.

For surjectivity, let $\{\rho_i\}$ be a partition of unity for $\mathcal{U}$, so $\rho_i \in \Omega^0(S)$, 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 $U_{ij}$. This is because the $f_i$ don’t have to be locally constant; in particular, we define

$$f_i = \sum_{U_j \in \mathcal{U}} p_j \zeta_{ij}.$$  

---

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.
A priori, this is only defined on $U_{ij}$, since $\zeta_{ij}$ is only defined there, but since $\rho_j$ 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 $U_{ij}$,

$$f_i - f_j = \sum_k \rho_k(\zeta_{ik} - \zeta_{jk})$$

$$= \sum_k \rho_k(\zeta_{ik} + \zeta_{kj})$$

by the cocycle condition. Then, since all the $\rho_k$ sum to 1,

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

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

The injectivity of $\tilde{c}$ 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. $H^1(S) = \text{Hom}_{\mathbb{Z}}(H_1(S), \mathbb{R})$. That is, for any loop $\gamma : S^1 \to S$ and a closed 1-form $\alpha$, 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 $f(t) \, dt$ for some $f$, and hence we can explicitly integrate.

$I$ is linear, and if $\alpha = dg$ is exact, then

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

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

Moreover, suppose we have a homotopy $\Gamma : S^1 \times [0, 1] \to S$ between two loops $\gamma_0$ and $\gamma_1$. In this case,

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

but since $\alpha$ is closed, then this is 0.\footnote{This calculation follows from Stokes’ theorem on surfaces, in case it seems confusing.} In particular, $I(\gamma, \alpha)$ depends only on the free homotopy class of $\gamma$, which can be expressed as stating that $I$ defines a bilinear map $I : H_1(S) \times H^1(S) \to \mathbb{R}$.\footnote{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 $\pi_1(S, \ast)$.} In other words, we have a homomorphism of abelian groups $H^1(S) \to \text{Hom}_{\mathbb{R}}(H_1(S), \mathbb{R})$ (i.e. to homomorphisms of abelian groups), and since $H_1(S)$ is the abelianization of $\pi_1(S, \ast)$, this is also $\text{Hom}_{\text{Grp}}(\pi_1(S, \ast), \mathbb{R})$. (Here, $\ast$ is any point in $S$, regarded as the basepoint for the homotopy group.)

**Theorem 14.5.** The integration map $I : H^1(S) \to \text{Hom}_{\mathbb{R}}(H_1(S), \mathbb{R})$ 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 $\text{Hom}(H_1(S), \mathbb{R})$, the first singular cohomology. The construction used in the proof of Theorem 14.4 can be adapted to show that there’s an isomorphism $H^1(S; \mathbb{R}) \to \check{H}^1(S; \mathbb{R})$; then, one checks that the composition of that isomorphism with the one for de Rham cohomology is given by $I$.\footnote{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 $\pi_1(S, \ast)$.}
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 $\star : H^1(S) \hookrightarrow \text{Hom}(H^1_c(S), \mathbb{R})$ 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 nondegenerate 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 $\alpha$ and a closed form $\beta$ with compact support (so $d\beta = 0$). Then,

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

by Stokes’ theorem. The analogous proof works for $\langle \alpha, d\beta \rangle$.

**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 $\alpha$ 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 $\gamma$ is an embedding; there are several ways to argue this, e.g. homotope $\gamma$ 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 $\alpha$ 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 $\gamma$, we’ll construct a Thom form $\tau_\gamma$, a closed 1-form supported in an annular neighborhood of $\gamma$ such that for all closed 1-forms $\alpha$ on $S$,

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

That is, it will be a sort of dual to $\gamma$ itself.

First, let’s fix an annular neighborhood (that is, a tubular neighborhood) $U$ for $\gamma$. 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 $\Gamma$ to preserve orientation. It suffices to construct $\tau_\gamma$ 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 $\tau_\gamma$ 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 $\phi$ won’t matter beyond this description.

Now, we can define $\tau = \phi^* ds = \phi'(s) ds$. If $\alpha \in \Omega^1(C)$ is such that $d\alpha = 0$, 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 $\alpha - p^* i^* \alpha = df$ for some $f \in C^\infty(C)$.

We can check that $\tau_\gamma$ is closed:

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

so by Stokes’ theorem, since $\tau_\gamma$ is compactly supported,

$$\int_C \tau_\gamma \wedge df = \int_C d(\tau_\gamma \wedge f) = 0.$$
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:

\[
\int_{C} \tau_{\gamma} \wedge \alpha' = \iint_{(-1,1) \times S^1} \phi'(s)g(t) \, ds \, dt
\]

\[
= \int_{-1}^{1} \phi'(s) \, ds \int_{S^1} g(t) \, dt
\]

\[
= \int_{S^1} i^* \alpha = \int_{S^1} \gamma^* \alpha.
\]

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

**Corollary 15.3.** If \( \gamma_1 \) and \( \gamma_2 \) are embedded loops in \( S \), then \( \int_{S} \tau_{\gamma_1} \cdot \tau_{\gamma_2} \) is the signed intersection number of \( \gamma_1 \) and \( \gamma_2 \).

**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 \( \gamma_1 \) is traveling rightwards, assign \(+1\) if \( \gamma_2 \) 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 \( \Sigma_g \) be a compact, oriented surface of genus \( g \). In this case, \( H_1(\Sigma_g) \cong \mathbb{Z}^{2g} \), spanned by loops \( a_1, \ldots, a_g, b_1, \ldots, b_g \), 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 \( a_i \cdot b_j = \delta_{ij} \).

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 \( J_x : T_x X \to T_x X \) such that \( J_x^2 = -\text{id} \) 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 \( e_1, e_2 \in T_x X \), 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 \( \text{SO}(2) \times \mathbb{R}_+ \), and a complex structure is the reduction of the structure group to \( \text{GL}_1(\mathbb{C}) = \mathbb{C}^* \). In higher dimensions, the isomorphism of these groups no longer holds.

---

**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 \( 90^\circ \) (multiplication by \( i \)) on each tangent space. In particular, above a point \( p \in X \), \( J_p \) is linear, and \( J_p^2 = -\text{id} \). 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 $T^*_p X = \text{Hom}(T_p X, \mathbb{R})$. 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 $\text{Hom}_\mathbb{R}(T_p X, \mathbb{C})$. Since $(J^*)^2 = -\text{id}$, 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, $T^*_p X' = \text{Hom}_\mathbb{C}(T_p X, \mathbb{C})$, 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 $f \, dz + g \, dy$, 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 $\alpha$ splits as $\alpha = \alpha^{1,0} + \alpha^{0,1}$, where if $v$ is a tangent vector we have

$$\alpha^{1,0}(v) = \frac{1}{2} (\alpha(v) - i \alpha(Jv)),$$

$$\alpha^{0,1}(v) = \frac{1}{2} (\alpha(v) + i \alpha(Jv)).$$

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 $dz = dx + i \, dy$ and $d\bar{z} = dx - i \, dy$, then a $(1,0)$-form locally looks like $f(x, y) \, dz$ and a $(0,1)$-form locally looks like $g(x, y) \, d\bar{z}$. In particular, if $f$ is a holomorphic function on $X$, then $df \in \Omega^{1,0}(X)$: $df = \frac{i}{2} \partial f \, dz$, 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

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

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

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

$$df = \partial f + \overline{\partial} f = \frac{1}{2} (f_x - i f_y) \, dz + \frac{1}{2} (f_x + i f_y) \, d\bar{z} = \frac{\partial f}{\partial z} \, dz + \frac{\partial f}{\partial \bar{z}} \, d\bar{z}.$$ 

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

$$\partial(A \, d\bar{z}) = \frac{\partial A}{\partial z} \, dz \wedge d\bar{z} = 2i \frac{\partial A}{\partial z} \, dx \wedge dy.$$ 

Correspondingly, if $B \, dz \in \Omega^{1,0}(X)$, then

$$\overline{\partial}(B \, dz) = \frac{\partial B}{\partial z} \, d\bar{z} \wedge dz = -2i \frac{\partial B}{\partial z} \, dx \wedge dy.$$ 

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

**Definition 16.1.** A **holomorphic 1-form** on $X$ is an $\alpha \in \Omega^{1,0}(X)$ such that $\overline{\partial} \alpha = 0$. 

**Proof.**
That is, it’s closed and in $\Omega^{1,0}$. Such a form locally looks like $\alpha = A \, dz$, 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

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

In local coordinates, $\Delta f = -(f_{xx} + f_{yy}) \, dx \wedge dy$.\(^{20}\)

**Definition 16.2.** A smooth $\phi$ is harmonic if $\Delta \phi = 0$.

**Lemma 16.3.** If $\phi$ is harmonic, then locally, $\phi = \text{Re}(f)$, 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 $A = 2 \text{Re}(i\bar{\partial}\phi) = i\bar{\partial}\phi + \bar{i\partial}\phi$. Thus, $dA = 2 \text{Re}(i\partial\bar{\partial}\phi) = 2 \text{Re}(i\partial\bar{\partial}\phi) = 0$, since $\phi$ is harmonic.

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

This $\psi$ is sometimes called the harmonic conjugate of $\phi$. 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 16.4 (Maximum principle).** If a harmonic function $\phi$ has a local maximum near $p$, then it’s constant near $p$.

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

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

**Hyperelliptic surfaces.**

**Definition 16.5.** 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 $\infty$ 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 composition 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 $\infty$, and so $X^* = X \cup \{P\}$. If $n$ is even, we need to add two points over $\infty$, and so it’s not a branch point.

Next time, we’ll construct a holomorphic 1-form $\alpha$ on $X$ and see that it extends meromorphically to $X^*$; if $n \leq 4$, 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 them, 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

\(^{20}\)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.
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 $\infty$. Concretely, if $n = \deg(f)$ is odd, the $X^* = X \cup \{P_1, P_2\}$; we know the inclusion $X \hookrightarrow \mathbb{CP}^2$ sending $(z, w) \mapsto [z : w : 1]$ extends to a holomorphic map $X^* \to \mathbb{CP}^2$, so we just need to know what exists over $[1 : 0 : 1]$. If $n$ is odd, $P \in \text{crit}(F)$, and $\infty$ is a critical value; if $n$ is even, them $P_{\pm} \notin \text{crit} F$, and $\infty$ 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 $180^\circ$ 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 $f : \mathbb{D} \to \mathbb{D}$ sending $z \mapsto z^2$. If $\mathbb{Z}/2$ acts on $\mathbb{D}$ by sending $z \mapsto -z$, then $f$ is $\mathbb{Z}/2$-equivariant, and $\overline{\mathbb{D}} = \mathbb{D}/(\mathbb{Z}/2)$. 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.](image)

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 $\alpha = dz/w$ is holomorphic for certain values of $n$.

A priori this looks meromorphic, rather than holomorphic, but on $X$, $2w \, dw = f'(z) \, dz$ as complex 1-forms, so $dz/w = z/f'(z) \, dw$. 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 $\alpha$ is holomorphic is to say that $\partial_\alpha = 0$, i.e. locally, $\alpha = \varphi(z) \, dz$, where $\varphi$ is holomorphic. This is not a problem except maybe at $\infty$. 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

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

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

The point is, $\alpha = dz/w$, so

$$\alpha = \frac{dz}{w} = \frac{\tau^n}{\sqrt{1 + a_1 \tau^2 + \cdots}} \left( -\frac{2}{\tau^3} \right) \, d\tau = \frac{-2\tau^{n-3}}{\sqrt{1 + a_1 \tau^2 + \cdots}} \, d\tau.$$

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

- $\alpha$ has a double pole if $n = 1$;
- $\alpha$ is holomorphic and nonvanishing if $n = 3$; and
- $\alpha$ 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, $\alpha$ 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 $\alpha$ 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 $\alpha_1$ and $\alpha_2$ 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 17.1.**

- 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 $p_1 \mapsto p_2$.

**Theorem 17.2.** 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 $\text{Aut}(E)$, 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 $q : \mathbb{C} \to \mathbb{C}/\Lambda$ is the projection map and $\beta$ is the holomorphic 1-form on $\mathbb{C}/\Lambda$ such that $q^*\beta = dz$, 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 $\Lambda$ be the period lattice of $\alpha$, 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 $x \to \int_p^x \alpha \bmod \Lambda$ (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 17.2; specifically, we’ll prove (1) implies (2), and then prove that every elliptic curve embeds as a Weierstrass cubic into $\mathbb{C}^2$, which uses elliptic functions, which we’ll also discuss.

---

27This 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 \(\alpha\) 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 \(\alpha\) as above,

1. \(\Lambda_{E, \alpha}\) is a lattice in \(\mathbb{C}\);
2. there is a unique biholomorphic map \(F : E \rightarrow \mathbb{C}/\Lambda_{E, \alpha}\) sending \(p \mapsto [0]\) and such that \(F^*(dz) = \alpha\); and
3. if \(\Lambda_1\) and \(\Lambda_2\) are lattices in \(\mathbb{C}\) and \(\phi : \mathbb{C}/\Lambda_1 \rightarrow \mathbb{C}/\Lambda_2\) 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 \(\alpha\) locally takes the form \(\alpha = f(z)dz\), then \(\beta\) locally takes the form \(\beta = f\,dz + \overline{f}\,d\overline{z}\) (so their zeros would be the same).

The Poincaré-Hopf theorem\(^2\) states that the Euler characteristic is given by the number of times \(\alpha\) 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, \(\Lambda = \mathbb{Z} \int_0^\alpha + \mathbb{Z} \int_b \alpha\). Thus, to show \(\Lambda\) is a lattice, it suffices to show that \(\int_0^\alpha\) and \(\int_b \alpha\) are \(\mathbb{R}\)-linearly independent. So suppose they’re linearly dependent, so after rescaling \(\alpha, \Lambda \subset \mathbb{R}\). Then, we can define \(\phi : E \rightarrow \mathbb{R}\) by \(\phi(x) = \Im \int_0^x \alpha\), which is independent of path. And \(\phi\) is harmonic, because locally, it’s the imaginary part of a holomorphic function, so by the maximum principle, \(\phi\) is constant! Since \(\phi(p) = 0\), then \(\phi = 0\). However, \(d\phi = \Im(\alpha)\), which by the open mapping theorem is not identically zero, which is a contradiction. Hence, \(\Lambda\) is a lattice, proving (1).

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

\[
x \mapsto \int_p^x \alpha \text{ mod } \Lambda.
\]

Clearly, \(F(p) = 0\), and by construction, \(F^*(dz) = \alpha\). Now, we would like to show that \(F\) is a biholomorphism. Since \(dF = \alpha\) 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 \(\alpha\) is sent to \(dz\), which is almost exactly what we wanted — but we don’t know that \(E\) is the period lattice yet. However, the period lattice of \((\mathbb{C}/L, dz)\) is exactly \(L\), so the period lattice of \(E\) has to be \(L\), i.e. \(L = \Lambda\).

For (3), suppose we have such a \(\phi\). Then, \(\phi^*(dz) = c\,dz\) 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 \(\mathbb{CP}^2\) 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 18.2.** 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 18.3.** A holomorphic elliptic function is constant.

\(^2\)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.\footnote{It’s so important that it gets its own typeface: it’s an old German p, but not the Gothic p! In LaTeX, you should use \wp to typeset $\wp$.}

**Definition 18.4.** 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 $\Lambda$.
- $\wp$ is an even function: $\wp(-z) = \wp(z)$.

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

**Theorem 18.5.** Any elliptic function for $\Lambda$ 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
\[
\frac{1}{(z - \lambda)^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2} \left( \frac{1}{(1 - z/\lambda)^2} - 1 \right) = \sum_{n \geq 1} (n + 1) \frac{z^n}{\lambda^{n+2}}.
\]
Hence, if we take $|z| \leq \min\{|\lambda| \mid \lambda \in \Lambda\}$, then
\[
\wp(z) = \frac{1}{z^2} + \sum_{\lambda \in \Lambda \setminus 0} \sum_{n \geq 1} (n + 1) \frac{z^n}{\lambda^{n+2}}
= \frac{1}{z^2} + \sum_{n \geq 1} \left( (n + 1) \sum_{\lambda \in \Lambda \setminus 0} \lambda^{-n-2} \right) z^n.
\]
Since $\wp$ is even, it only has even powers in its Laurent series, so
\[
(18.6) \quad \wp(z) = \frac{1}{z^2} + \sum_{n \geq 1} (2n + 1) G_{2n+2}(\Lambda) z^{2n}.
\]
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 $\Lambda$-invariant meromorphic function $\varphi : \mathbb{C} \to S^2$ (or $\mathbb{C}/\Lambda \to S^3$) defined by

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

We also calculated its Laurent series in (18.6) (recall that $G_{2k}(\Lambda) = \sum_{\lambda} \lambda^{-2k}$). We’ll now derive a differential equation for $\varphi$. From (18.6),

$$\varphi'(z) = -\frac{2}{z^3} + \sum_{n \geq 1} 2n(2n + 1)G_{2n+2}z^{2n-1}$$

$$= -\frac{2}{z^3} + 6G_4z + 20G_6z^3 + O(z^5).$$

Therefore

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

Compare this to

$$\varphi(z) = \frac{1}{z^2} + 3G_4z^2 + 5G_6z^4 + O(z^6)$$

$$\varphi(z)^3 = \frac{1}{z^6} + 9G_4z^2 + 15G_4 + O(z^6).$$

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 18.3. By inspection of the first few terms, we should take

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

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 $g_3 = 140G_6$, $\varphi$ satisfies the equation

$$(\varphi')^2 = 4\varphi^3 - g_2\varphi - g_3.$$  

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 $[z] \mapsto [\varphi(z) : \varphi'(z) : 1]$ (for $z \not\in \Lambda$). Near $z = 0$, $\theta(z) = [z^3 \varphi(z) : z^5 \varphi'(z) : z^7]$, and therefore approaches $[0 : 1 : 0]$. Hence, we let $\theta(0) = [0 : 1 : 0]$. This means that the image of $\theta$ is $\text{Im}(\theta) = \overline{X} = \{y^2z = 4x^3 - g_2xz^2 - g_3z^3\} \subset \mathbb{C}P^2$.

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

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

In particular, $\overline{X}$ is a Riemann surface. $\theta$ 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 $\Lambda_1$ 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 \( \mathfrak{X}_1 \to \mathfrak{X}_2 \) 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 \( \delta \) 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 19.5.** 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 19.6.** \( j \) 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 \( \Lambda \) 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 \( \Delta = -16(-g_2^3/4^2 + (27/4)g_3^3) \). In particular, the \( j \)-invariant of the lattice \( \Lambda \) is

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

Any \( \Lambda \) can be rescaled to one of the form \( \Lambda = \mathbb{Z} \cdot 1 + \mathbb{Z} \cdot \tau \), where \( \tau \in \mathbb{H} \); \( \tau \) isn’t uniquely determined by \( \Lambda \), but only up to the action of the modular group \( \text{PSL}_2(\mathbb{Z}) \) acting on \( \mathbb{H} \) by Möbius maps. In other words, this lattice doesn’t have a unique \( \mathbb{Z} \)-basis, but we can define \( J : \mathbb{H}/(\text{PSL}_2(\mathbb{Z})) \to \mathbb{C} \) by \( J([\tau]) = j(\Lambda_{1 \in \mathbb{Z} \cdot \tau}) \). In particular, \( \mathbb{H}/\text{PSL}_2(\mathbb{Z}) \) 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.

---

\(^{30}\)A coarse moduli space is one that does not keep track of automorphisms of a given object.
The Riemann-Hurwitz Formula. To talk about the Euler characteristic we’ll need to exclude a few pathological examples.

Definition 20.1. If $S$ is a surface, it has finite (topological) type if $\dim H^j(S)$ is finite for $j = 0, 1, 2$.

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 20.2. 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 20.3 (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.\textsuperscript{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 20.4. 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 20.5. 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 20.6 (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.\textsuperscript{32} Then,

\begin{equation}
\chi(X) = \deg(f)\chi(Y) + R_f,
\end{equation}

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

\begin{equation}
R_f = \sum_{x \in X} (k_x - 1),
\end{equation}

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 (20.8) 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 (20.7) are hard to remember, but if you can do the calculation below, you’ll be fine.

Example 20.9. 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 20.6: $\chi(X) = 2 - 2g$, 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 20.10. The total ramification $R_f$ is even.

\textsuperscript{31} M. Victoris 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!

\textsuperscript{32} If $X$ and $Y$ are connected, then all this means is that $f$ is nonconstant.
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 20.11.** 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 $Y$.

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 (20.7) is $-\deg f : \chi(Y) + R_f \geq -\chi(Y)$, 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 20.6.** Let $\Delta \subset Y$ be the set of critical values; since $Y$ is compact, this is finite, so let $b = |\Delta|$. Let $N_\Delta$ be a small neighborhood of $\Delta$ 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 20.5, $\chi(X_0) = d\chi(Y_0)$. Now, we’ll relate $\chi(X)$ and $\chi(Y)$ to this relation.

We can use the Mayer-Vietoris property, Proposition 20.3, 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 $\delta$ 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 $\delta$ is

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

Thus,

$$\chi(X) = \chi(X_0) + \sum_{\delta \in \Delta} \left( d - \sum_{x \in f^{-1}(\delta)} (k_x - 1) \right)$$

$$= \chi(X_0) + db - \sum_{x \in \text{crit}(f)} (k_x - 1).$$

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 20.12.** 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 $\text{Aut}(X)$ is also finite, so one can take $G = \text{Aut}(X)$. 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 $\text{Aut}(X)$, and the induced map $\text{Aut}(X) \to S_N$ 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 $|\text{Aut } X| = 168 = 2 \cdot 84$, so this bound is satisfied. ($\text{Aut}(X)$ is actually the simple group of order 168, $\text{PSL}_2(\mathbb{F}_7)$.) It’s known that the bound is sharp for some infinite sequence of genera.

**Proof of Theorem 20.12.** 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 $\text{stab}_G(x)$ is a group of automorphisms of $D$ fixing 0, and therefore a group of rotations. Thus, $\text{stab}_G(x) = \mu_n$: 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

\[
R_\pi = \sum_{y \in G \cdot x} \left( |G| - |\text{stab}_G(x)| \right),
\]

which therefore implies that

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

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

\[
2g(Y) - 2 + \sum_{y \in G \cdot x} \left( 1 - \frac{1}{|\text{orb}_G(x)|} \right) > 0.
\]

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

\[
2g - 2 + \sum_{i=1}^N \left( 1 - \frac{1}{e_i} \right)
\]

is 42.\(^3\) This, together with (20.13), shows that

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

which proves the theorem.

So why 42? To minimize (20.14), we should definitely set \( g = 0 \). Let \( E = \sum_{i=1}^N (1 - 1/e_i) \), which needs to be greater than 2; since \( e_1 \geq 2 \), then \( 1 > 1 - 1/e_1 \geq 1/2 \), 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 \( \Gamma = \text{PSL}_2(\mathbb{Z}) = \text{SL}_2(\mathbb{Z})/\{\pm 1\} \). 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. \( \Lambda \sim \Lambda c \) 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 \( \Gamma \) acting on \( \mathbb{H} \).

\(^3\)It’s the answer to the Ultimate Question of Life, the Universe, and Everything!
• Talk about modular curves $Y_N = \mathbb{H}/\Gamma_N$, where $\Gamma_N$ is a principal congruence subgroup of $\Gamma$; these are the elements that reduce mod $N$ to the identity. We’ll also discuss their compactifications. Through the Riemann-Hurwitz formula (Theorem 20.6), we’ll calculate their genus.

We begin with two group-theoretic facts.

Fact.

• If

$$S = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad T = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix},$$

then $\Gamma$ 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 21.1. A fundamental region for $\Gamma$ acting on $\mathbb{H}$ is a set $X$ such that every $\Gamma$-orbit intersects the closure $\overline{X}$, and no two points of $X$ lie in the same $\Gamma$-orbit.

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

Theorem 21.2. If $\Omega = \{ z \in \mathbb{H} \mid |z| > 1, |\Re z| < 1/2 \}$, then $\Omega$ is a fundamental region for $\Gamma$ acting on $\mathbb{H}$.

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

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure8.png}
\caption{A fundamental region for $\Gamma$ acting on $\mathbb{H}$, as in Theorem 21.2.}
\end{figure}

Proof of Theorem 21.2. Pick an $A \in \text{SL}_2(\mathbb{Z})$, so $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$. If $c = 0$, then $ad = 1$, so $A = \begin{bmatrix} \pm 1 & b \\ 0 & \pm 1 \end{bmatrix}$; 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

$$\text{Im}(A(\tau)) = \frac{\text{Im}(\tau)}{|c\tau + d|^2}. \quad (21.3)$$

Now, we’re going to march through a few lemmas. The overarching idea is that, within some $\Gamma$-orbit, we’d like to find the point in $\overline{\Omega}$ that maximizes $\text{Im}(\tau)$; we will also show that given any $\Gamma$-orbit, we do have some such point, and then we’re done.

Lemma 21.4. If $\tau \in \Omega$ and $c \neq 0$, then $\text{Im}(A(\tau)) < \text{Im}(\tau)$.

This is a quick check using (21.3).

Lemma 21.5. If $\tau, A\tau \in \Omega$, then $A = \pm I$, so $A\tau = \tau$.

Proof. If $c \neq 0$, then by Lemma 21.4, $\text{Im}(A\tau) < \text{Im}(\tau) = \text{Im}(A^{-1}A\tau) < \text{Im}(A\tau)$, which is a contradiction. Hence, $c = 0$, so $A = \pm T^m$ is a translation. However, $T^m(\Omega) \cap \Omega = \emptyset$ unless $m = 0$, so $A = \pm I$ as desired. \(\Box\)

\(^{34}\)For a reference, see T. Apostol, Modular Functions and Dirichlet Series in Number Theory.
We've now shown that no two points of \( \Omega \) lie in the same \( \Gamma \)-orbit, which is half of the proof. But we still need to show that every \( \Gamma \)-orbit is represented.

**Lemma 21.6.** Suppose \( \tau^* \) maximizes \( \text{Im}(\tau) \) within its \( \Gamma \)-orbit; then, \( |\tau^*| \geq 1 \).

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

So in this case, you're almost in \( \Omega \) — 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 \( \gamma \mapsto \text{Im}(\gamma \cdot \tau_0) \).

**Lemma 21.7.** \( h \) is bounded above.

**Proof.** One could approach this with hyperbolic geometry, but it’s just as true if you prove it directly: \( \text{Im}(A\tau_0) = \text{Im}(\tau_0)/|c\tau_0 + d|^2 \) by (21.3). However, \( c\tau_0 + d \) is a nonzero point in the lattice \( \Lambda_{\tau_0} \), and therefore it cannot be arbitrarily close to 0, and therefore its modulus is bounded.

One lemma remaining.

**Lemma 21.8.** \( h \) 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 21.7. Thus, \( \text{Im}(\tau_n) \) is also bounded above. Using translations \( T^n \), we may assume \( |\text{Re}(\tau_n)| \leq 1/2 \), 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 \( \tau^* \). We’d like this to lie in the orbit of \( \tau_0 \), but this is not immediate: we do know that \( \tau^* \in \overline{\Gamma \cdot \tau_0} \). However, \( \Gamma \) 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 9.3.) Thus, \( \tau_{n_j} \) must be constant for large \( j \), and in particular \( \tau^* \in \Gamma \cdot \tau_0 \). Thus, \( \tau^* \) maximizes \( h \).

In conclusion, let \( \tau^* \) maximize \( h \), and through translations, we may assume \( \text{Re}(\tau^*) \leq 1/2 \), and by Lemma 21.6, \( |\tau^*| \geq 1 \). Thus, \( \tau^* \in \overline{\Omega} \), so \( \overline{\Omega} \) intersects all \( \Gamma \)-orbits.

On \( Y = \mathbb{H}/\Gamma \), we have to identify the two left and right sides of \( \Omega \) 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 \( \Omega \) is actually promoted to a full cuicircle through \( z \mapsto z^2 \). Similarly, the map \( z \mapsto z^3 \) makes a neighborhood of \( \rho \) 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).

---

**Lecture 22.**

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

Last time, we proved Theorem 21.2, which characterized a fundamental region (almost, but not quite, a fundamental domain) for the modular group \( \Gamma = \text{PSL}_2(\mathbb{Z}) \). This region \( \Omega \) is the set of \( z \in \mathbb{H} \) with \( |\text{Re} z| < 1/2 \) and \( |z| > 1 \).

Today, we’re going to talk about the subgroups \( Q_N = \text{PSL}_2(\mathbb{Z}/N) \) 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 \( \Gamma_p \) for \( p \) prime, as in Example 9.2, but the primality is actually irrelevant: the action of \( \Gamma_N \) 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 \( \pi_N \) denote the projection \( Y_N \to Y_N/Q_N \cong \mathbb{H}/\Gamma = Y \) (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 22.1.** A level \( N \) structure on a lattice \( \Lambda \) 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 22.2.** 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 $\Gamma_N$ on $\mathbb{H}$? Since $\Gamma_N$ 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 $\Gamma_N$. Then, let $\Omega_N$ be the interior of $g_1\Omega \cup \cdots \cup g_q\Omega$; this is a fundamental region for $\Gamma_N$. 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 $g_i\Omega$, and hence is still closed and bounded in $\mathbb{H}$, hence compact.

**Example 22.3.** Let’s see what happens when $n = 2$. Then, the first column of a matrix in $\text{PSL}_2(\mathbb{Z}/2)$ 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 identify them up to multiplication in $(\mathbb{Z}/2)^\times$, but this just contains the identity; hence, we get 6 elements in $\text{PSL}_2(\mathbb{Z}/2)$.

Let $S$ and $T$ be the generators we defined last lecture; then, we will take the cosets of $\Gamma_2$ 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 $\Omega_2$. 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 $\rho^2$ (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 $\pi_2 : Y_2 \to Y$ is a degree-6 map branched over two critical points, $\rho$ and $i$ (since these are the two non-free $\Gamma$-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 \( \infty \).
- Using the monodromy at \( \infty \) for \( \pi_N \), we get a recipe for extending \( \pi_N : Y_N \rightarrow Y \) to a branched covering \( \pi_N : X_N \rightarrow S^2 \); 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 \( \pi_2 \) has branching order 2 at each cusp.

We’re going to use the Riemann-Hurwitz formula (Theorem 20.6) 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 \( \Omega_N \) by taking the set of cosets \( I, T, T^2, \ldots, T^{n-1} \), and some others (many others, though we won’t write them down). Thus, \( \Omega_N \) consists of a bunch of translates of \( \Omega \) by \( T \), plus some extra stuff given by the action of \( S \). The action of \( T^N \) identifies the ends of \( P = \Omega \cup T\Omega \cup \cdots \cup T^{n-1}\Omega \), 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, \( \pi_N \) 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 20.12 implies that

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

For \( Q_N \) acting on \( X_N \), the nonfree orbits are those of \( \rho \), whose stabilizer has size 3; \( i \), whose stabilizer has size 2; and \( \infty \), whose stabilizer has size \( N \). Plugging this into (22.4),

\[
2g(X_N) - 2 = q_n \left( -2 + \left( 1 - \frac{1}{3} \right) + \left( 1 - \frac{1}{2} \right) + \left( 1 - \frac{1}{N} \right) \right)
= q_N \left( \frac{1}{6} - 1N \right) = \frac{q_N(N - 6)}{6N}.
\]

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

**Proposition 22.5.** 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 \( \text{GL}_2(\mathbb{Z}/p) \rightarrow (\mathbb{Z}/p)^* \) 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 22.3, \( g(X_2) = 0 \).
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 20.12; it has the largest possible symmetry group.

These \( X_N \) are where modular forms live; but that’s another 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 \( \alpha \), 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 \( \gamma \) 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).

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

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 $\theta$ be a holomorphic 1-form on $X$. Assume $\theta$ isn’t identically $0$ on any connected component of $X$ (which automatically implies its zeros are isolated). Then,

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

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

**Proof.** Let $\alpha = \text{Re}(\theta) = (1/2)(\theta + \overline{\theta})$, so the zeros of $\alpha$ are the zeros of $\theta$. Near such a zero, $z = x + yi$, so if $\theta = f(z) \, dz$, then $\alpha = (\text{Re} \, f) \, dx + (\text{Im} \, f) \, dy$, and so the winding number of $(\text{Re} \, f, \text{Im} \, f)$ around $z = re^{i\theta}$ 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 $\sum k_p(\theta) \leq 0$ for any such $\theta$.

We can also define $k_p(\theta)$ when $p$ is a pole of a meromorphic 1-form: if $\theta = f(z) \, dz$ 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 $\theta$ be a meromorphic 1-form on $X$ that isn’t identically $0$ on any connected component of $X$. Then,

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

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

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

We’d like to obtain an ordinary 1-form out of $\theta$, 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 $0 \leq t \leq 1$, $\psi(t) = 1$.
- If $t \geq 2$, then $\psi(t) = 1/t$.
- On $(1, 2)$, $\psi$ smoothly interpolates between these two curves.

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

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

Thus, $\overline{\theta}$ is a complex 1-form, whose zeros are the zeros and poles of $\theta$, and so we can repeat the argument that proved Proposition 23.2, checking that the multiplicities for $\alpha = \text{Re}(\theta)$ 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 20.6. 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 (20.8).

Proof of Theorem 20.6 using Theorem 23.4. This proof is contingent on the existence of a nonvanishing, meromorphic 1-form $\theta$ 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, $\theta = g(w) \, dw$, so $f^*\theta = g(z^k) \, d(z^m) = mg(z^m)z^{m-1} \, dz$; 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 $\ell_y = k_0(y)$. Thus,

$$
\sum_{x \in f^{-1}(y)} \frac{1}{k_f(x)} = \sum_{x \in f^{-1}(y)} (m_x\ell_y + m_x - 1) = (\deg f)\ell_y + \sum_{x \in f^{-1}(y)} (m_x - 1).
$$

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

$$
-\chi(X) = -(\deg f)\chi(X) + \sum_{x \in X} (m_x - 1).
$$

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 $\infty$, 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

$$
dp = \frac{\partial p}{\partial z} \, dz + \frac{\partial p}{\partial w} \, dw = 0,
$$

so if we let

$$
\theta = \frac{dz}{\partial p/\partial w} = -\frac{dw}{\partial p/\partial z},
$$

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

\(^{35}\)This is a special case of an adjunction formula characterizing the genus of a compact Riemann surface embedded in more general manifolds, and which admits an entirely topological proof.
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.)

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, though.

**Definition 24.1.** 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 $\text{Div}(X)$. 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 24.2.** 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 $\text{PDiv}(X) \subseteq \text{Div}(X)$. We’ll soon see that not every divisor is principal.

**Definition 24.3.** Two divisors $D_1$ and $D_2$ are linearly equivalent if $D_1 - D_2 \in \text{PDiv}(X)$.

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 $\text{Cl}(X) = \text{Div}(X)/\text{PDiv}(X)$.\(^{36}\)

**Definition 24.4.** 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 $\text{Div} X$.

This partial order does not pass through to $\text{Cl}(X)$.

**Definition 24.5.** The degree of a divisor $D = \sum a_x \cdot x$ is

$$\deg D = \sum_{x \in X} a_x \in \mathbb{Z}.$$

Thus, $\deg : \text{Div} X \to \mathbb{Z}$ is a homomorphism of groups.

**Lemma 24.6.** If $D \in \text{PDiv}(X)$, then $\deg D = 0$.

\(^{36}\)A fact which we will not use: there is an isomorphism $\text{Cl}(X) \cong \text{Pic}(X)$, where the latter is the Picard group of $X$. 
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, \( \text{deg}((f)) = 0 \) 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,

\[
\text{deg}((f)) = \sum_{i=1}^N \int_{\partial \Delta_i} \frac{df}{f}.
\]

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,

\[
\sum_{i=1}^N \int_{\partial \Delta_i} \frac{df}{f} = -\int_{\partial X_0} \frac{df}{f} = -\int_{X_0} \left( \frac{df}{f} \right),
\]

because \( df/f \) is nonsingular on \( X_0 \). However, since \( f \) is holomorphic on \( X_0 \), then \( d(df/f) = 0 \): in local coordinates, \( df = f'(z) \, dz \), and so the only interesting thing can come from \( d\bar{z} \), but applying this to a holomorphic function gives you \( 0 \). Thus, \( \text{deg}((f)) = 0 \). \( \square \)

The point is that we can use Stokes’ theorem to bring the argument principle to general Riemann surfaces.\(^\text{37} \)

As a consequence, we get a degree homomorphism \( \text{deg} : \mathcal{Ct}(X) \to \mathbb{Z} \).

**Definition 24.7.** Let \( D \) be a divisor; then, we will let \( \mathcal{O}_X(D) \) denote the *sheaf of effective divisors*, the sheaf of functions defined as follows: if \( U \subset X \) is open, then \( \mathcal{O}_X(D)(U) \) is the set of \( f : X \to S^2 \) meromorphic on \( U \) and such that \( \text{deg}((f)) \geq 0 \) 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 \( \text{supp} \, D \), 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 24.8.** Let \( H^0(\mathcal{O}_X(D)) = \mathcal{O}_X(D)(X) = \{ f \in \mathcal{M}_X \mid f = 0 \text{ or } f \in \mathcal{M}_X^\times \text{ and } \text{deg}((f)) \geq 0 \}. \) 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 \( h^0 : \mathcal{Ct}(X) \to \mathbb{Z}_{\geq 0} \).

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 \( \text{deg} \, D < 0 \), then \( H^0(D) = 0 \), because if \( D + (f) \geq 0 \), then \( \text{deg}(D + (f)) \geq 0 \), but this was \( \text{deg} \, D \), 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 \( p \notin \text{supp}(D') \), 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. \( af + bg \in H^0(D) \). Thus, \( \dim(h^0(D + p)/h^0(D)) \leq 1 \), and we can repeat this argument for the rest of \( \text{supp}(D) \setminus \text{supp}(D') \).

3. Combining these two, if \( D \) is linearly equivalent to an effective divisor, \( h^0(D) \leq \text{deg} \, D + 1 \), 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

\[
h^0(D) \leq \max(0, \text{deg}(D) + 1).
\]

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 24.9 (Riemann-Roch inequality\(^\text{38} \)).** If \( X \) is a compact, connected Riemann surface and \( D \) is a divisor on \( X \), then \( h^0(D) \geq \text{deg} \, D + 1 - g(X) \).

\(^\text{37}\)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.

\(^\text{38}\)This is also called the weak Riemann-Roch theorem.
The proof of this will have to wait.

**Corollary 24.10.** 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 24.9. 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.

Of course, the Riemann-Roch inequality is a more precise result than this, but it’s still a nice fact to know.

**Corollary 24.11.** If $X$ is a compact, connected Riemann surface and $g(X) = 0$, then $X$ is biholomorphic 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 24.12.** Let $H^0(K_X(D))$ denote the vector space of meromorphic 1-forms on $X$ with poles at worst $D$.

This is defined in local coordinates: if $\alpha$ is a meromorphic 1-form, then in local coordinates $z$, $\alpha = f(z) \, dz$, and we require that $v_f(x) \leq m$ (meaning $z^m f(z)$ is holomorphic).

**Theorem 24.13 (Riemann-Roch).** Let $D$ be a divisor on the compact Riemann surface $X$. Then,

$$\dim H^0(O_X(D)) - \dim H^0(K_X(-D)) = \deg D + 1 - g.$$  

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: 4/4/16**

“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 $\Gamma = \text{PSL}_2(\mathbb{Z})$, the modular group, and $\Gamma_N$ be the kernel of the reduction mod $N$, $\Gamma \to \text{PSL}_2(\mathbb{Z}/N)$ (so $\Gamma_1 = \Gamma$). Thus, $\Gamma_N$ 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 $\Gamma_N$. 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.

\(^{39}\)If you, like me, didn’t recognize this colloquialism, it means “to make a [slight] mess of it.”
Definition 26.1. Let $k \in \mathbb{Z}$. A weakly modular function of weight $2k$ for the group $\Gamma_N$ is a meromorphic function $f : \mathbb{H} \to \mathbb{C}$ such that

\[(26.2) \quad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_N, \text{ we have } f(\gamma(t)) = (ct + d)^{2k}f(t).\]

This relation appears to depend on the choice of representative for $\gamma \in \text{PSL}_2(\mathbb{Z}/N)$, but the condition that $\det(\gamma) = 1$ means that $c$ and $d$ are determined up to one choice of sign, meaning that taking $(ct + d)^{2k}$ makes the ambiguity disappear. This is why we only consider even weights.

**Proposition 26.4.**

Let $\eta$ be a weakly modular function of weight $k$ for the group $\Gamma_N$. Then let $\tilde{\eta} = \eta|_\Gamma$, where $\Gamma$ is the subgroup of $\text{PSL}_2(\mathbb{Z}/N)$ consisting of matrices with determinant $1$. Then $\tilde{\eta}$ is a holomorphic function of weight $k$ on $X_N$.

**Proof.**

The idea is that if $\eta$ is a weakly modular function of weight $k$ on $X_N$, then $\tilde{\eta}$ is a holomorphic function of weight $k$ on $X$. Thus, $\tilde{\eta}$ is a holomorphic function of weight $k$ on $X_N$. Therefore, $\tilde{\eta}$ is a holomorphic function of weight $k$ on $X_N$.

\[\tilde{\eta}(\gamma z) = (cz + d)^{-k}\tilde{\eta}(z)\]

This relation is known as the equivariance property. It implies that $\tilde{\eta}$ satisfies the same transformation law as $\eta$, but now it is holomorphic. Therefore, $\tilde{\eta}$ is a weakly modular function of weight $k$.

We’d like to relate these to a geometric construction of differentials on a Riemann surface.

**Definition 26.3.** Let $X$ be a Riemann surface and $\mathcal{X}_X$ 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 $\eta$ of $\mathcal{X}_X^\otimes k$. That is, $\eta$ attaches to each $x \in X$ an element $\eta_x \in ((T^1_0)^*)^k X$ that varies holomorphically.

Locally, in a holomorphic coordinate $z$, there’s a holomorphic $f$ such that $\eta = f(z)(dz)^\otimes k$. The holomorphic $k$-differentials form a vector space $H^0(\mathcal{X}_X^\otimes k) \cong H^0(kK_X)$ 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 26.4.** There’s an isomorphism between the space of meromorphic $k$-differentials on $X_N$ and the weakly modular functions for $\Gamma_N$ with weight $2k$.

**Proof.**

The idea is that if $\eta$ is a meromorphic $k$-differential on $Y_N$ and $\pi : \mathbb{H} \to Y_N$ is the canonical projection, then let $\tilde{\eta} = \pi^*\eta$. Hence, we can write $\tilde{\eta}f(\tau)(d\tau)^\otimes k$, 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 $\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_N$. Then, $\gamma^*\tilde{\eta} = \tilde{\eta}$, so

\[\gamma^*(\tilde{\eta}) = f(\gamma(\eta))(d(\gamma(\eta)))^\otimes k\]

\[= f(\gamma(\eta))\left(\frac{d\gamma}{(ct + d)^2}\right)^\otimes k\]

\[= f(\gamma(\eta))(ct + d)^{2k}(d\gamma)^\otimes k,\]

and since this is equal to $\tilde{\eta} = f(\tau)(d\tau)^\otimes k$, then $f$ is invariant under $\gamma$.

Finally, one needs to check that one can reconstruct $\gamma$ 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 $\eta$ 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 $\Gamma$ 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 $\Gamma$, let $\tilde{P}$ be a preimage of $P$. If $z$ is a coordinate on $\mathbb{H}$ near $\tilde{P}$ and $w$ is one on $Y$ near $P$, then $z \mapsto z^3 = w$, so if $\eta = f(w)(dw)^\otimes k$, then $\tilde{\eta} = f(z^3)(dz^3)^\otimes k$. Thus, if $(\text{div } \eta)_P = m$, then $(\text{div } \tilde{\eta})_{\tilde{P}} = 2k + 3m$. 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 26.5.**

- Let $f : \mathbb{H} \to \mathbb{C}$ be meromorphic, and let $q^{1/2} = e^{\pi i/3}$, which is a holomorphic coordinate on $Y_N$ near the cusp at $i\infty$ in $\mathbb{H}$. The function $f$ is holomorphic at $i\infty$ if $q^{1/2}$ 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 \( i\infty \).  
• A modular form \( f \) for \( \Gamma_N \) of weight \( 2k \) is a weakly modular function of weight \( 2k \) that is
  - holomorphic on \( \mathbb{H} \), and
  - holomorphic at the cusps in the sense above.

The modular forms of weight \( 2k \) for \( \Gamma_N \) form a vector space, denoted \( M_{2k}(\Gamma_N) \).

**Example 26.6.** Recall that the Eisenstein series for a lattice \( \Lambda_7 = \mathbb{Z} \oplus \mathbb{Z} \tau \) is defined by
\[
G_{2k}(\tau) = \sum_{x \in \Lambda_7 \setminus 0} \frac{1}{x^{2k}}.
\]
This is a modular form for \( \Gamma \) 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 \( \eta = g(\zeta)(d\zeta)^{\otimes k} \), then when we pull back to \( \mathbb{H} \),
\[
\tilde{\eta} = \left( \frac{2\pi i}{N} \right)^k g(\zeta)(\zeta)^k(d\zeta)^k,
\]
because \( d\zeta = (2\pi i/N)(\zeta)d\zeta \).

This means \( \tilde{\eta} \) 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 \( \deg D > \deg K_N \), then \( h^0(D) = \deg D + 1 - g \). One can therefore compute that \( \deg(K_{X_N} + C) = \left\lfloor G/6 \right\rfloor \), 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
\[
\dim M_{2k}(\Gamma_N) = \deg(k(K_{X_N} + C)) + 1 - g(X_N)
\]
\[
= |G| \left( \frac{2k - 1}{12} + \frac{1}{2N} \right).
\]

A similar computation goes through for \( N = 1 \).

**Klein’s Quartic Curve: 4/6/16**

Today, we’re going to focus on a specific example of a modular form, *Klein’s quartic curve* \( Y_7 = \mathbb{H}/\Gamma_7 \).

Here, \( \Gamma_7 \) is the matrices \( A \in \text{PSL}_2(\mathbb{Z}) \) such that \( A \equiv \pm 1 \mod 7 \). We’ve computed that \( G = \text{PSL}_2(\mathbb{Z}/7) \) acts on \( X_7 = Y_7 \cup \{ \text{cusps} \} \), where the cusps form a \( G \)-orbit of size 24. We also computed that \( g(X_7) = 3 \), so \( |G| = 168 \), which is the maximal group of symmetries of a genus-3 Riemann surface. This is a good reason that one should care about it: it’s maximally symmetric, so should be interesting and nice.

We’d like to embed \( X_7 \) as a plane algebraic curve (the name suggests that it’ll be a quartic, but we don’t know that yet). Since \( h^0(K_{X_7}) = g(X_7) = 3 \), then if we can find a basis \((x,y,z)\) for \( H^0(X_{X_7}) \), we’ll get a map \( X_7 \to \mathbb{CP}^2 \) induced by \([x:y:z]\).\(^41\) We’ve also seen that a homogeneous polynomial cuts a genus-3 Riemann surface out of \( \mathbb{CP}^2 \) exactly when its degree is 4, so the image of \([x:y:z]\) will be a quartic.

This was known relatively long ago, though in a slightly different form.

**Theorem 27.1 (F. Klein).** There is a basis \((x,y,z)\) for \( H^0(X_{X_7}) \) such that \([x:y:z]\) is an embedding \( X_7 \to \mathbb{CP}^2 \) with image the zero set of \( \phi(x,y,z) = x^3y + y^3z + z^3x \).

\( ^{40}\)It suffices to take one coset representative for each coset of \( \Gamma_N \) in \( \Gamma \).

\( ^{41}\)\( K_{X_7} \) has no basepoints, as with any canonical divisor on a Riemann surface of genus at least 2.
This $\phi$ is also called Klein’s quartic.

The first thing we need is a quick computation, which has been left as an exercise.

**Lemma 27.2.** There is no $(x, y, z) \in \mathbb{C}^3 \setminus \{0\}$ such that $\phi, \phi_x, \phi_y, \text{ and } \phi_z$ all vanish; in particular, $\phi^{-1}(0)$ is a Riemann surface.

Hence, once we know $[x : y : z]$ maps into $\phi^{-1}(0)$, we’re done: it’s a holomorphic map $f$ between two genus-$3$ Riemann surfaces, so the Riemann-Hurwitz formula tells us that $4 = \deg f \cdot 4 + R_f$, and so $R_f = 0$ and $\deg f = 1$, i.e. $f$ is biholomorphic.

Since $G \subset \text{Aut}(X_7)$, then $G$ acts on the space $V = H^0(X_7)$, the holomorphic one-forms, by linear automorphisms. In other words, we have a representation $\rho : G \rightarrow \text{GL}(V)$.\(^{42}\)

**Lemma 27.3.** Let $X$ be a Riemann surface of genus $g > 1$ and $\Gamma \leq \text{Aut} \ X$ be a finite subgroup.\(^{43}\) Then, the action of $\Gamma$ on $H^0(X)$ is faithful.\(^{44}\)

**Proof.** Once again, we can use the Riemann-Hurwitz formula! Suppose $g \in \Gamma$ acts trivially on $H^0(X_7)$, and let $Y = X/\langle g \rangle$, which is a Riemann surface. Then, if $\pi : X \rightarrow Y$ is projection, then $\pi^* : H^0(X_7) \rightarrow H^0(Y)$ is injective. Since $g$ acts trivially, then every holomorphic 1-form is pulled back via $\pi^*$, so $\pi^*$ is also surjective, hence an isomorphism. Thus, $g(Y) = \dim H^0(X_7) = \dim H^0(Y) = g(X)$, so by the Riemann-Hurwitz formula, $\deg \pi = 1$, so $Y = X$ and $g$ must be the identity. \(\Box\)

$G = \text{PSL}_2(\mathbb{Z}/7)$ is a complicated group, so let’s focus on a parabolic subgroup of $G$, the upper triangular matrices

$$P = \left\{ \pmatrix{a & b \\ 0 & a^{-1}} \mid a, b \in \mathbb{Z}/7, a \neq 0 \right\} \leq G.$$  

This is a subgroup of order $7 \cdot 6/2 = 21$. We’ll use $P$ to pin down the equation of our quartic: it has to be left-invariant under a faithful action of $P$, which is easier to calculate with than all of $G$.

$P$ is generated by the matrices

$$\alpha = \begin{bmatrix} 3 & 0 \\ 0 & 5 \end{bmatrix} \quad \text{and} \quad \beta = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}.$$  

Since this is a subgroup of $\text{PSL}_2(\mathbb{Z}/7)$, $\alpha$ has been identified with $-\alpha$, and $\beta$ with $-\beta$. The relations for $P$ (which you can check) are $P = \langle \alpha, \beta \mid \alpha^3 = 1, \beta^7 = 1, \alpha \beta \alpha^{-1} = \beta^2 \rangle$. In particular, if $A = \rho(\alpha) \in \text{GL}(V)$ and $B = \rho(\beta) \in \text{GL}(V)$, then $A^3 = I$, $B^7 = I$, and $ABA^{-1} = B^2$. If you pick a basis for $V$, these will be $3 \times 3$ matrices.

Let $e$ be an eigenvector for $B$, so $Be = \lambda e$, where $\lambda$ is a $7^{\text{th}}$ root of unity. Since $B \neq I$, then we can pick $e$ and $\lambda$ such that $\lambda \neq 1$, i.e. it’s a primitive $7^{\text{th}}$ root of unity. We’d like to diagonalize $B$. We know $B^2(Ae) = AB(Ae) = \lambda^4(Ae)$

$$B(Ae) = (B^2)^4(Ae) = \lambda^4(Ae)$$

$$B^2(A^2e) = B^2(Ae) = ABAe = \lambda^4A^2e$$

$$B(A^2e) = (B^2)^4A^2e = (\lambda^4)^4A^2e = \lambda^2(A^2e).$$

This seemed a little unmotivated, but the point is that $(1, A, A^2e)$ is a basis of eigenvectors for $V$, with eigenvectors $\lambda$, $\lambda^4$, and $\lambda^2$, respectively.

This means that we can rename $e$ if necessary (to one of $Ae$ or $A^2e$) such that $\lambda = e^{\pm 2\pi i/7}$. This determines $V$ fully as a $P$-representation.

Now, let $(x, y, z)$ be a basis for $V^*$ that’s dual to $(1, A, A^2e)$. We’ll determine all of the quartics in $(x, y, z)$ that are $P$-invariant, up to scalar factors; these will be our candidates for the things that $X_7$ could cut out of $\mathbb{C}P^2$.

\(^{42}\)If you’ve never seen representation theory, we’re not going to do anything fancy; $\rho$ is simply a group homomorphism from $G$ to $\text{GL}(V)$.

\(^{43}\)As we mentioned, $\text{Aut} \ X$ is finite, but we can’t prove that yet, so the finiteness hypothesis isn’t really necessary, but we can’t get rid of it yet.

\(^{44}\)A representation $\rho : G \rightarrow \text{GL}(V)$ is faithful if $\rho$ is an injective map.
Under $B$, $x^pg^qz^r$ scales as $\lambda^w(x^pg^qz^r)$, where $w = -(p + 4q + 2r)$ (the minus sign comes about because we've dualized). Pur quartic is a sum of monomials of the same weight $w$, so we need the weights to be invariant under cyclic permutations $(x \mapsto y \mapsto z \mapsto x)$. Thus, we need $p$, $q$, and $r$ to satisfy

\begin{align*}
p + 4q + 2 &\equiv q + 4r + 2p \equiv r + 4p + 2q \mod 7 \\
p + q + 4 &\equiv 4.
\end{align*}

One can check this forces $w = 0$, and the only monomials of weight 0 are $x^3y$, $y^3z$, and $z^3x$. Thus, the only thing it could be is $\phi(x, y, z) = x^3y + y^3z + z^3x$, so if $(x, y, z)$ obeys a quartic constraint, then this is the correct polynomial. But we still have to check that.

However, $h^0(4K_X) = \deg(4K_X) + 1 - g + h^0((1 - 4)K_X)$ by Riemann-Roch; this is also equal to $4(2g - 2) + 1 - g = 14$. The number of monomials of degree 4 in $x$, $y$, and $z$ is the number of monomials in $x$ and $y$ of degree at most 4, which is $1 + 2 + \cdots + 5 = 15$. There are too many of these, so they must satisfy a linear relation, meaning $x$, $y$, and $z$ satisfy a quartic. Thus, we've proved Theorem 27.1, albeit in a pretty roundabout way.

From a modular perspective, $H^0(\mathcal{M}_X)$ is the space of modular forms of weight 2 for $\Gamma_0$ vanishing at the cusps, so-called cusp forms. Thus, $(x, y, z)$ is dual to a basis for these cusp forms. These cusp forms have been explicitly identified: if $\alpha = (-1 + \sqrt{-7})/12$, then we get $q$-series that are “modified $\Theta$-functions” in some sense:

$$x^* = \sum_{\beta \in \mathbb{Z}[\alpha]} \text{Re}(\beta)q^{\beta^2},$$

and $y^*$ is the same except $\beta \equiv 4 \mod -7$, and $z^*$ uses $1 \mod -7$.

### Lecture 28.

**The Riemann-Roch Theorem: Towards a Proof: 4/8/16**

Today, we’re going to start talking about the proof of the Riemann-Roch theorem, Theorem 24.13. We’re not going to provide the proof today, but we’ll introduce some language.

We want to prove that if $D$ is a divisor on a compact Riemann surface $X$, then $h^0(\mathcal{O}_X(D)) - h^0(\mathcal{O}_X(-D)) = \deg D + 1 - g$. If $D \geq 0$, so $D = \sum a_ip_i$ with $a_i > 0$, the basic problem is:

*To what extent can we find a meromorphic function with prescribed “tails” (principal parts) of order at most $a_i$ at each $p_i$?*

We can’t always do this: there is an obstruction to defining such functions which lives in Dolbeault cohomology $H^{0,1}(X)$, which is a vector space that measures the obstruction to solving $\bar{\partial}f = \eta$, where $\eta$ is a prescribed $(0,1)$-form.

**Remark.**

- It turns out that $\dim H^{0,1}(X) = g(X)$, which is where the genus term emerges from in the Riemann-Roch theorem.
- $\bar{\partial}f = \eta$ is always locally solvable, but may not be globally solvable. This means that there’s a sheaf $\mathcal{O}_X$ and an isomorphism of $H^{0,1}(X)$ with the sheaf cohomology $H^1(\mathcal{O}_X)$.

In order to compute $\dim H^{0,1}(X)$, we will define and set up the Serre pairing $H^0(\mathcal{M}_X) \times H^{0,1}(X) \to \mathbb{C}$, and prove that it’s nondegenerate. We’ll also need a more general form of this pairing to be nondegenerate; this is called Serre duality. The main difficulty in proving Riemann-Roch will be proving Serre duality.

There are two routes to Serre duality.

1. The first, due to Serre, uses local methods and the sheaf theory of currents. This is pretty, but uses perhaps a bit more sheaf theory than we’d like to have in this class.
2. There’s a global method which investigates the *Poisson equation* $\bar{\partial}\bar{\partial}u = p$, where $p$ is a 2-form of integral 0.

---

We’ll take the second approach, which is the one taken in the textbook. It does involve globally solving a PDE, yes, but we can leverage some theorems from analysis to help us. Moreover, very similar methods are used to prove the Hodge theorem and set up Hodge theory, which is extremely useful in complex geometry and algebraic geometry. Moreover, the same PDE appears in the proof of the uniformization theorem, so it seems like a good idea to work with it now. The idea is that we basically have to do stuff with harmonic functions, potentials, and their friends.

Today, though, we’re going to set up the language we need to state these results.

Invertible Sheaves. These are also known as line bundles.

**Definition 28.1.** A presheaf of abelian groups \( \mathcal{F} \) on a topological space \( X \) associates to each open set \( U \subseteq X \) an abelian group \( \mathcal{F}(U) \) (such that \( \mathcal{F}(\emptyset) = 0 \)), and to each inclusion \( V \subseteq U \) of open sets associates a “restriction” homomorphism \( \rho_{UV}: \mathcal{F}(U) \to \mathcal{F}(V) \), such that

1. \( \rho_{UU} = \text{id}_{\mathcal{F}(U)} \) and
2. If \( W \subseteq V \subseteq U \), then \( \rho_{VW} \circ \rho_{UV} = \rho_{UW} \).

We will write \( f|_V \) for \( \rho_{UV}(f) \).

A sheaf of abelian groups is a presheaf of abelian groups that also satisfies the following two properties whenever an open subset \( U \subseteq X \) is covered by an open cover \( \mathcal{U} = \{ U_i \mid i \in I \} \).

- **Locality:** If \( f \in \mathcal{F}(U) \) and \( f|_{U_i} = 0 \) for all \( U_i \in \mathcal{U} \), then \( f = 0 \).
- **Patching:** Given \( f_i \in \mathcal{F}(U_i) \) for each \( i \) such that \( f_i|_{U_i \cap U_j} = f_j|_{U_i \cap U_j} \) for all \( i, j \in I \), then there exists an \( f \in \mathcal{F}(U) \) such that \( f_i = f|_{U_i} \) for all \( i \).

A sheaf is like a presheaf, but we can glue elements together in a unique way, as long as they agree on overlaps.

**Example 28.2.**

1. If \( X \) is any topological space, \( C^0(X) \), the sheaf of continuous functions on \( X \) is defined by \( C^0(X)(U) = \{ f : U \to \mathbb{R} \text{ continuous} \} \), and \( \rho_{UV} \) is restriction of functions. This sheaf contains several other important sheaves inside of it.
2. If \( X \) is a smooth manifold, there’s also \( C^\infty(X) \), the sheaf of \( C^\infty \) functions, which is the same except assigning smooth functions, rather than continuous ones.
3. Using locally constant, real-valued functions instead of continuous ones defines the sheaf of locally constant functions \( \mathbb{R}(X) \) (so \( \mathbb{R}(X)(U) \) is the set of locally constant functions from \( U \to \mathbb{R} \)). However, choosing just the constant functions does not define a sheaf: on a disconnected space, one could choose a different constant function on each connected component, and then there would be no way to glue them. A sheaf has to be built out of data that can be checked locally.
4. If \( X \) is a Riemann surface, its holomorphic functions form a sheaf \( \mathcal{O}_X \) (sometimes called the structure sheaf), and its meromorphic functions form a sheaf \( \mathcal{M}_X \).
5. If \( D \) is a divisor on \( X \), there’s a sheaf \( \mathcal{O}_X(D) \) for which \( \mathcal{O}_X(D)(U) \) is the set of meromorphic functions \( f : U \to \mathbb{R} \) such that on each compact \( K \subseteq U \), either \( f = 0 \) or \( v_x(f) + D_x \geq 0 \).

All these examples were functions, but not all sheaves are sheaves of functions.

**Example 28.3.**

1. Given a divisor \( D \), the sheaf of meromorphic 1-forms with poles no worse than \( D \) is denoted \( \mathcal{K}_X(D) \)
2. Similarly, \( \mathcal{O}^{0,1}_X \) denotes the sheaf of \( C^\infty \) (0, 1)-forms.

For a sheaf \( \mathcal{F} \), one says that \( H^0(\mathcal{F}) = \mathcal{F}(X) \), the global sections of \( \mathcal{F} \).

The structure sheaf \( \mathcal{O}_X \) for a Riemann surface \( X \) is more than just a sheaf of abelian groups: in fact, it’s a sheaf of commutative \( \mathbb{C} \)-algebras, meaning that for every open \( U \subseteq X \), \( \mathcal{O}_X(U) \) is a commutative \( \mathbb{C} \)-algebra and the restriction maps are \( \mathbb{C} \)-algebra homomorphisms (since, after all, restriction commutes with pointwise multiplication).

**Definition 28.4.** A sheaf \( \mathcal{F} \) on a Riemann surface \( X \) is a sheaf of \( \mathcal{O}_X \)-modules if each \( \mathcal{F}(U) \) is an \( \mathcal{O}_X(U) \)-module, and the restriction maps \( \rho_{UV} \) are \( \mathcal{O}_X(U) \)-linear.

\( ^{46}\)Sheaves are often labeled \( \mathcal{F} \), since they are ultimately French in origin, and the French word for sheaf is *faisceau*. 
For example, \( \mathcal{O}_X \), \( \mathcal{K}_X(D) \), \( C^\infty(X, \mathbb{C}) \), and \( \Omega_X^{0,1} \) are all sheaves of \( \mathcal{O}_X \)-modules. \( \mathbb{R}_X \) is not, though.

**Definition 28.5.** If \( X \) is a Riemann surface, an **invertible sheaf of \( \mathcal{O}_X \)-modules** \( \mathcal{L} \) is an \( \mathcal{O}_X \)-module such that there’s an open cover \( \mathfrak{U} \) of \( X \) such that for each \( U \in \mathfrak{U} \), \( \mathcal{L}(U) \) is a free \( \mathcal{O}_X(U) \)-module of rank 1.

That is, locally \( \mathcal{L}(U) \cong \mathcal{O}_X(U) \), but globally it may be “more twisted;” in particular, invertible sheaves are holomorphic line bundles. Not all of the sheaves we’ve mentioned are invertible, but \( \mathcal{K}_X(D) \) and \( \mathcal{O}_X(D) \) are.

**Dolbeault cohomology.** We have two operators \( \overline{\partial} : \Omega^{0,0}(X) \to \Omega^{0,1}(X) \) and \( \partial : \Omega^{1,0}(X) \to \Omega^{1,1}(X) \) which are related to properties of holomorphic functions.

**Definition 28.6.** The **Dolbeault cohomology** of \( X \) is the collection of the following groups.

- \( H^{0,0}(X) = \ker(\overline{\partial} : \Omega^{0,0}(X) \to \Omega^{0,1}(X)) \). These are the global holomorphic functions \( H^0(\mathcal{O}_X) \), which is isomorphic to \( \mathbb{C} \) if \( X \) is compact and connected.
- \( H^{1,0}(X) = \ker(\partial : \Omega^{1,0}(X) \to \Omega^{1,1}(X)) \). These are the holomorphic 1-forms \( H^0(\mathcal{K}_X) \), which suggests why we might care about this in the proof of the Riemann-Roch theorem.
- \( H^{0,1}(X) = \text{coker}(\overline{\partial} : \Omega^{0,0}(X) \to \Omega^{0,1}(X)) = \Omega^{0,1}(X) / \text{Im}(\overline{\partial}) \).
- \( H^{1,1}(X) = \text{coker}(\partial : \Omega^{1,0}(X) \to \Omega^{1,1}(X)) = \Omega^{1,1}(X) / \text{Im}(\partial) \).

The cokernels are in general more challenging to understand than the kernels, but turn out to play a very crucial role.

One can set this theory up more generally: let \( \mathcal{F} \) be a sheaf of \( \mathcal{O}_X \)-modules. Then, we can define

\[
\mathcal{F}_\mathcal{F} = \mathcal{F} \otimes \text{id}_\mathcal{F} : (\Omega^{0,0} \otimes \mathcal{O}_X, \mathcal{F}) \to \Omega^{0,1} \otimes \mathcal{O}_X, \mathcal{F}.
\]

This is a little wonky: what’s the tensor product of sheaves of \( \mathcal{O}_X \)-modules? The answer is a little nuanced to define: it’s not just the tensor product on each open subset, since that may not satisfy the gluing axiom. Indeed, one has to use sheafification, which means defining everything on small open subsets and using gluing to define things globally. To that end, if \( \mathcal{M} \) and \( \mathcal{N} \) are \( \mathcal{O}_X \)-modules, then \( (\mathcal{M} \otimes \mathcal{O}_X, \mathcal{N})(U) = \mathcal{M}(U) \otimes \mathcal{O}_X(U), \mathcal{N}(U) \) for small opens \( U \) in a certain cover of \( X \); then, gluing data determines the rest of the sheaf.

In any case, we know \( \overline{\partial}_{\mathcal{F}}(f \otimes s) = \overline{\partial}(f) \otimes s \) at least locally, and define \( H^{0,0}(\mathcal{F}) = \ker \overline{\partial}_{\mathcal{F}} \) and \( H^{0,1}(\mathcal{F}) = \text{coker} \overline{\partial}_{\mathcal{F}}. \) One can do something similar to get \( H^{1,0}(\mathcal{F}) \) and \( H^{1,1}(\mathcal{F}) \). We’ll have to use Dolbeault cohomology of invertible sheaves in the proof of Riemann-Roch.

---

Lecture 29.

**The Riemann-Roch Theorem: Exact Sequences of Sheaves: 4/11/16**

Today, we’re going to talk about Serre duality, leading to part of the proof of the Riemann-Roch theorem. The goal will be to use the language of sheaves; if \( X \) is a compact Riemann surface and \( p \in X \), we’d like an intrinsic description of the meromorphic functions with (for example) a simple pole at \( p \).

Consider the **exact sequence of sheaves**

\[
\begin{CD}
0 @>>> \mathcal{O}_X @>>> \mathcal{O}_X(p) @>>> \mathcal{O}_X(p)/\mathcal{O}_X @>>> 0.
\end{CD}
\]

To ease notation, we’ll let \( \mathcal{O}_p = \mathcal{O}_X(p)/\mathcal{O}_X \). Here, \( \mathcal{O}_X \) is the sheaf of holomorphic functions on \( X \), and \( \mathcal{O}_X(p) \) is the sheaf of meromorphic functions on \( X \) with at worst a simple pole at \( p \).

How is the quotient \( \mathcal{O}_p \) defined? The idea is that for every \( x \in X \), there’s a small open neighborhood \( U \) of \( x \) such that \( \mathcal{O}_p(U) = \mathcal{O}_X(p)(U)/\mathcal{O}_X(U) \) as \( \mathcal{O}_X(U) \)-modules; then, for general opens, we patch these together using the sheaf axiom. In particular, \( \mathcal{O}_p(V) \cong \mathbb{C} \) if \( p \in V \), and is 0 otherwise.

That \( (29.1) \) is a (short) exact sequence is not a definition, but rather a property.

**Definition 29.2.** Let

\[
\begin{CD}
\cdots @>>> \mathcal{A} @>a>> \mathcal{B} @>b>> \mathcal{C} @>>> \cdots
\end{CD}
\]

be a sequence of sheaves of abelian groups over a space \( X \), with all arrows morphisms of sheaves.

---

47 Maybe it’s sad that the naïve approach doesn’t work and we have to do something more complicated, but “sheafification” is so much fun to say!

48 Formally, what we’re doing is taking the sheafification of the presheaf quotient, but we’re not going to use this formalism in this class.
(29.3) is exact at $\mathcal{B}$ if on a small open neighborhood of every point, $\ker(b) = \operatorname{Im}(a)$.
(29.3) is exact if it’s exact at each place in the sequence.
As with abelian groups, a short exact sequence of sheaves is an exact sequence of the form
\[ 0 \longrightarrow \mathcal{A} \overset{a}{\longrightarrow} \mathcal{B} \overset{b}{\longrightarrow} \mathcal{C} \longrightarrow 0. \]

In general, if $X$ is a compact Riemann surface and we have a short exact sequence (29.4) of $\mathcal{O}_X$-modules, then by the abstract nonsense of homological algebra we have a long exact sequence (of abelian groups) in Dolbeault cohomology:
\[ 0 \longrightarrow H^0(\mathcal{A}) \overset{a^\ast}{\longrightarrow} H^0(\mathcal{B}) \overset{b^\ast}{\longrightarrow} H^0(\mathcal{C}) \overset{\delta}{\longrightarrow} H^0,1(\mathcal{A}) \longrightarrow H^0,1(\mathcal{B}) \longrightarrow H^0,1(\mathcal{C}) \longrightarrow 0. \]

The map $\delta$ is called the connecting map and is defined in the following way.

- Choose a $\gamma \in H^0(\mathcal{C})$, so $\gamma \in (\Omega^0,0_X \otimes \mathcal{O}_X)(X)$ and $\overline{\partial}_{\mathcal{E}} \gamma = (\overline{\partial} \otimes \operatorname{id}_{\mathcal{E}})\gamma = 0$.
- Since $b$ and therefore $b^\ast$ are surjective, we can lift $\gamma$ to a $\beta \in (\Omega^0,0_X \otimes \mathcal{O}_X)(X)$, so $b(\beta) = \gamma$. Then, $b(\overline{\partial}_{\mathcal{E}} \beta) = \overline{\partial}_{\mathcal{E}}(b\beta) = \overline{\partial}_{\mathcal{E}} \gamma = 0$, so $\overline{\partial}_{\mathcal{E}} \beta \in \ker(b)$.
- Thus, by exactness, $\overline{\partial}_{\mathcal{E}} \beta \in \operatorname{Im}(a)$, so we can choose a lift $\delta \gamma \in \Omega^0,1 \otimes \mathcal{O}_X \mathcal{A}$ of $\overline{\partial}_{\mathcal{E}} \beta$. Finally, we define $\delta(\gamma)$ to be the class of $\delta \gamma$ in $H^0,1(\mathcal{A})$.

One can check that this defines a homomorphism making (29.5) exact by hand, or by invoking principles of homological algebra.

We’re going to apply this to the short exact sequence (29.1); the long exact sequence we obtain in cohomology is
\[ 0 \longrightarrow H^0(\mathcal{O}_X) \longrightarrow H^0(\mathcal{O}_X(p)) \longrightarrow H^0(\mathcal{O}_p) \overset{\delta}{\longrightarrow} H^0,1(\mathcal{O}_X) \longrightarrow H^0,1(\mathcal{O}_p) \longrightarrow 0. \]

Here, $H^0(\mathcal{O}_p) \cong \mathbb{C}$.

What does $\delta$ geometrically mean in this situation? Let $D = D(0;1)$ be a coordinate disc centered at $p$. Then $t = 1/z$ on $D$ represents a generator for $H^0(\mathcal{O}_p) \cong \mathbb{C}$, and we can lift it to $(\Omega^0,0_X \otimes \mathcal{O}_X(\mathcal{O}_p))(X)$. Let $\chi$ be a smooth cutoff function for $D$, meaning that $\chi(x) = 1$ on $(1/3)D$, but $\chi(x) = 0$ outside $(2/3)D$. Then, we have another lift
\[ \tilde{t} = \begin{cases} \chi \otimes (1/z), & \text{on } D \\ 0, & \text{outside } D. \end{cases} \]

This tensor product should be thought of as pointwise multiplication. Then, $\delta(t)$ is the lift to $\Omega^0,1 \otimes \mathcal{O}_X \mathcal{O}_X$ of $\overline{\partial}(\tilde{t})$: it’s equal to $\overline{\partial}(\chi) \otimes (1/z)$ on $D$, and is $0$ elsewhere; in particular, it’s supported in the annulus $(2/3)D \setminus (1/3)D$. Then, $\delta(t)$ is the class of $\delta(t)$ in $H^0,1(X)$. That is, $\delta$ is obtained by multiplying by a cutoff function and differentiating.

One thinks of $t$ as the “tail” of a function (meaning the principal part of $f$ is $1/z$ in this local coordinate), in the following sense.

**Proposition 29.6.** Given $X$, $p$, and $t$ as above, the following are equivalent.

1. There exists a meromorphic function with a single pole at $p$ and whose tail is $t$.
2. $\delta(t) = 0$ in $H^0,1(X)$.
3. $\delta t = \overline{\partial} f$ for some $f \in \Omega^0,0(X)$.

If such an $f$ exists, $\overline{\partial}(\chi \cdot (1/z) - f) = 0$, so $\chi \cdot (1/z) - f$ is the desired meromorphic function.

So the obstruction is a class in Dolbeault cohomology; to progress, we need to understand $H^0,1(X)$. The crucial fact here is that $\dim H^0,1(X) = g(X)$, meaning that for a compact Riemann surface $X$ of genus $0$, there’s always a solution, so there’s a meromorphic function with just a pole at $p$, and hence $X \cong S^2$.

We can repeat all of this replacing $\mathcal{O}_X$ by any invertible $\mathcal{O}_X$-module $\mathcal{L}$, especially if $\mathcal{L} = \mathcal{O}_X(D)$ for a divisor $D$. If $\mathcal{L}(p) = \mathcal{L} \otimes_{\mathcal{O}_X} \mathcal{O}_X(p)$ and $\mathcal{L}_p = \mathcal{L}(p)/\mathcal{L}$, then by the same line of reasoning there’s a short exact sequence of $\mathcal{O}_X$-modules
\[ 0 \longrightarrow \mathcal{L} \longrightarrow \mathcal{L}(p) \longrightarrow \mathcal{L}_p \longrightarrow 0. \]
Recall that \( \mathcal{L}_p \) is a skyscraper sheaf, supported entirely at \( p \) (and only sets containing it). The induced long exact sequence in cohomology is

\[
\begin{array}{ccccccc}
0 & \longrightarrow & H^0(\mathcal{L}) & \longrightarrow & H^0(\mathcal{L}(p)) & \longrightarrow & H^0(\mathcal{L}_p) \\
& & & \delta & & & \\
& & & H^0(\mathcal{L}|\mathcal{L}_p) & \longrightarrow & H^0(\mathcal{L}(p)) & \longrightarrow & H^0(\mathcal{L}_p) & \longrightarrow & 0.
\end{array}
\]

**Lemma 29.7.** \( H^{0,1}(\mathcal{L}_p) = 0. \)

The proof boils down to the fact that \( 1/z \, d\bar{z} = \overline{\partial}(z/z) \); in particular, we obtain a slightly shorter long exact sequence

\[
\begin{array}{ccccccc}
0 & \longrightarrow & H^0(\mathcal{L}) & \longrightarrow & H^0(\mathcal{L}(p)) & \longrightarrow & H^0(\mathcal{L}_p) \\
& & & & \delta & & \Longrightarrow & \\
& & & & H^0(\mathcal{L}|\mathcal{L}_p) & \longrightarrow & H^0(\mathcal{L}(p)) & \longrightarrow & H^0(\mathcal{L}_p) & \longrightarrow & 0.
\end{array}
\]

Great; so what does this have to do with the Riemann-Roch theorem? Well, there’s a lemma from linear algebra that helps.

**Lemma 29.9.** Let

\[
\begin{array}{cccccc}
0 & \longrightarrow & V_1 & \longrightarrow & V_2 & \longrightarrow & \cdots & \longrightarrow & V_N & \longrightarrow & 0
\end{array}
\]

be an exact sequence of vector spaces. If for every adjacent 3 vector spaces, two are finite-dimensional, then all of them are finite-dimensional, and the alternating sum

\[
\sum_{i=1}^N (-1)^i \dim V_i = 0.
\]

This is a consequence of the rank-nullity theorem.

Applying this lemma to the sequence (29.8), we deduce that \( H^0(\mathcal{L}) \) and \( H^{0,1}(\mathcal{L}) \) are finite-dimensional iff \( H^0(\mathcal{L}(p)) \) and \( H^{0,1}(\mathcal{L}(p)) \) are finite-dimensional. In this case, we can define \( \chi(\mathcal{L}) = \dim H^0(\mathcal{L}) - \dim H^{0,1}(\mathcal{L}), \) a sort of Euler characteristic of \( \mathcal{L}, \) and because \( H^0(\mathcal{L}_p) \cong \mathbb{C}, \ \chi(\mathcal{L}(p)) - \chi(\mathcal{L}) = 1. \)

In other words, each time you add a point, you add 1 to the Euler characteristic, and vice versa.

**Corollary 29.10.** Let \( \mathcal{L} \) be an invertible \( \mathcal{O}_X \)-module with finite-dimensional \( H^0(\mathcal{L}) \) and \( H^{0,1}(\mathcal{L}). \) If

\[
D = \sum_{i=1}^N p_i - \sum_{j=1}^M q_j,
\]

then

\[
\chi(\mathcal{L} \otimes \mathcal{O}_X(D)) - \chi(\mathcal{L}) = \deg D.
\]

Part of the statement is that these sheaves have finite-dimensional cohomology, so that their Euler characteristics make sense.

Combining this with the statement that \( h^{0,1}(X) = g, \) then we see that for all divisors \( D, \)

\[
(29.11) \quad \chi(\mathcal{O}_X(D)) = \chi(\mathcal{O}_X) + \deg D = 1 - g + \deg D.
\]

This looks very much like the Riemann-Roch theorem.

**Serre duality.** We will spend the last five minutes on Serre duality. To complete the proof of Riemann-Roch from this point, we certainly need (29.11), but we also need to know that \( h^{0,1}(\mathcal{O}_X(D)) = h^0(\mathcal{K}_X \otimes \mathcal{O}_X(-1)). \)

In fact, if we can show this, we’re done, and this is where Serre duality comes in.

Let \( \mathcal{L} \) be an invertible sheaf of \( \mathcal{O}_X \)-modules, e.g. \( \mathcal{O}_X(D), \) and \( \mathcal{L}^* = \text{Hom}_{\mathcal{O}_X}(\mathcal{L}, \mathcal{O}_X), \) which is the dual or inverse sheaf; in this case, \( \mathcal{O}_X(-D). \) If \( \Omega_X^{p,q}(\mathcal{L}^*) = (\Omega_X^{p,q} \otimes \mathcal{O}_X(\mathcal{L}^*)) \) denotes the \((p,q)\)-forms valued in \( \mathcal{L}^*, \) then there’s a bilinear pairing \( \Omega_X^{p,q}(\mathcal{L}^*) \times \Omega_X^{1-p,1-q}(\mathcal{L}) \rightarrow \Omega_X^{1,1} \) combining the wedge product with evaluation \( \mathcal{L}^* \times \mathcal{L} \rightarrow \mathcal{O}_X. \)

**Definition 29.12.** The **Serre pairings** are \( \sigma^1 : H^{1,1}(\mathcal{L}^*) \times H^0(\mathcal{L}) \rightarrow \mathbb{C} \) and \( \sigma^0 : H^{1,0}(\mathcal{L}^*) \times H^{0,1}(\mathcal{L}) \rightarrow \mathbb{C} \)

sending \( (\eta, \theta) \rightarrow \int_X (\eta \wedge \theta). \)

**Theorem 29.13** (Serre duality). **The Serre pairings are nondegenerate.**

In particular, this implies \( H^{0,1}(\mathcal{L}) \cong (H^{1,0}(\mathcal{L}^*))^* \cong (H^0(\mathcal{K}_X \otimes \mathcal{L}^*))^*. \) Thus, the dimensions agree: \( h^{0,1}(\mathcal{L}) = h^0(\mathcal{K}_X \otimes \mathcal{L}^*), \) and with \( \mathcal{L} = \mathcal{O}_X(D), \) we’ve finished the proof of Riemann-Roch! But we still need to prove Serre duality, which we’ll do next time.
Lecture 30.

Serre Duality: 4/13/16

Last time, we showed that Serre duality for $\mathcal{O}_X(D)$ and that the rank of the Dolbeault cohomology satisfying $h^{0,1}(X) = g(X)$ together imply the Riemann-Roch theorem for $\mathcal{O}_X(D)$. This was all stated somewhat quickly, and using some exact sequences, but nothing too deep.

Today, we’ll first show that Serre duality for $\mathcal{O}_X$ implies Serre duality for $\mathcal{O}_X(D)$, which is a somewhat formal argument; nothing very much will happen. Instead, all the work goes into proving Serre duality for $\mathcal{O}_X$ and showing that $h^{0,1}(X) = g(X)$, which amounts to inverting the Laplacian.

Let $\mathcal{L}$ be an invertible sheaf on a compact Riemann surface $X$. If $\mathcal{L}^*$ is the dual invertible sheaf $\mathcal{L}^* = \text{Hom}_{\mathcal{O}_X}(\mathcal{L}, \mathcal{O}_X)$, then $\mathcal{L} \otimes_{\mathcal{O}_X} \mathcal{L}^* \cong \mathcal{O}_X$, which is the sense in which $\mathcal{L}$ is invertible. Recall that the Serre pairings are $\sigma_0^\mathcal{L} : H^0(\mathcal{L}) \times H^1(\mathcal{L}^*) \to \mathbb{C}$ defined by sending $(s, [\omega]) \mapsto \int_X s(\omega)$, evaluating $\omega$ on $s$; and $\sigma_1^\mathcal{L} : H^0(\mathcal{L}) \times H^1(\mathcal{L}^*) \to \mathbb{C}$ sending $[\theta], \eta \mapsto \int_X \theta \wedge \eta$. That these are well-defined follows from Stokes’ theorem.

The statement of Serre duality for $\mathcal{L}$ (Theorem 29.13) is that that the pairings $\sigma_0^\mathcal{L}$ and $\sigma_1^\mathcal{L}$ are non-degenerate. That is, the maps induced by these pairings, $H^0(\mathcal{L}) \to H^1(\mathcal{L}^*)$ sending $s \mapsto \sigma_0^\mathcal{L}(s, \cdot)$ and $H^1(\mathcal{L}) \to H^0(\mathcal{L}^*)$ sending $\theta \mapsto \sigma_1^\mathcal{L}(\theta, \cdot)$.

Serre proved that Serre duality holds for all invertible sheaves $\mathcal{L}$. We don’t need this in full generality.

Theorem 30.1. If $D$ is a divisor on a compact Riemann surface $X$, Serre duality holds for $\mathcal{O}_X(D)$.

Remark.

- It turns out that any invertible sheaf is $\mathcal{O}_X(D)$ for some divisor $D$ (induced from a meromorphic section), but this is nontrivial and so we won’t prove it. So this is Serre duality in generality, sort of.
- Serre’s proof uses sheaf cohomology, the $\partial$-Poincaré lemma, on solutions of the form $\partial f = g$, and the theory of smoothing of currents.\footnote{Serre’s proof uses sheaf cohomology, the $\partial$-Poincaré lemma, on solutions of the form $\partial f = g$, and the theory of smoothing of currents.}

Proposition 30.2. Let $p \in X$. Then, Serre duality for $\mathcal{L}$ is equivalent to $\mathcal{L}(p) = \mathcal{L} \otimes_{\mathcal{O}_X} \mathcal{O}_X(p)$.

Proof. We have a commutative diagram

$$
\begin{array}{ccc}
H^0(\mathcal{L}) & \xrightarrow{i^*} & H^0(\mathcal{L}(p)) \\
\sigma_0^\mathcal{L} & & \sigma_0^{\mathcal{L}(p)} \\
H^0(\mathcal{L} \otimes \mathcal{L}^*) & \xrightarrow{j^*} & H^0(\mathcal{L}(p) \otimes \mathcal{L}^*(p))^*.
\end{array}
$$

Here, $i^*$ is the map induced from the inclusion $i : \mathcal{L} \to \mathcal{L}(p)$ and $j^*$ is the dual of the map induced from the inclusion $j : \mathcal{L}(p) \otimes \mathcal{L}^*(p) \to \mathcal{L}(p) \otimes \mathcal{L}^*$. Since we can identify $H^0(\mathcal{L}(p) \otimes \mathcal{L}^*(p)) = H^1(\mathcal{L}^*)$ and $H^0(\mathcal{L}(p) \otimes \mathcal{L}^*(p)) = H^1(\mathcal{L}(p)^*)$, so the vertical arrows make sense. The commutativity of the diagram is an exercise, but amounts to definition-checking.

In the same way, there’s a commutative diagram

$$
\begin{array}{ccc}
H^1(\mathcal{L}) & \xrightarrow{i^*} & H^1(\mathcal{L}(p)) \\
\sigma_1^\mathcal{L} & & \sigma_1^{\mathcal{L}(p)} \\
H^1(\mathcal{L}^*) & \xrightarrow{j^*} & H^1(\mathcal{L}(p)^*)^*.
\end{array}
$$

We’ll compare the exact sequences that arise in Dolbeault cohomology from two short exact sequences of sheaves. The first is

(30.3a) \hspace{2cm} 0 \longrightarrow \mathcal{L} \longrightarrow \mathcal{L}(p) \longrightarrow \mathcal{L}_p \longrightarrow 0,

where $\mathcal{L}_p = \mathcal{L}(p)/\mathcal{L}$ just like last time; the second is

(30.3b) \hspace{2cm} 0 \longrightarrow \mathcal{H} \otimes \mathcal{L}^*(-p) \longrightarrow \mathcal{H} \otimes \mathcal{L}^* \longrightarrow (\mathcal{H} \otimes \mathcal{L})_{-p} \longrightarrow 0,

\footnote{That is, on sufficiently small opens $U \subset X$, $\mathcal{L}^*(U) = \text{Hom}_{\mathcal{O}_X(U)}(\mathcal{L}(U), \mathcal{O}(U))$.}

\footnote{For a reference, consult Serre’s paper “Un théorème de dualité.”}
where again the third term is just the quotient. Now, we get long exact sequences in cohomology from (30.3a) and the dualized version of (30.3b), and the Serre pairings define vertical morphisms making the following big diagram commutative.

\[
\begin{array}{cccccccc}
0 & \rightarrow & H^0(\mathcal{L}) & \rightarrow & H^0(\mathcal{L}(p)) & \rightarrow & H^0(\mathcal{L}_p) & \rightarrow & H^0,1(\mathcal{L}) & \rightarrow & H^0,1(\mathcal{L}(p)) & \rightarrow & 0 \\
& & \sigma_x & \downarrow \sigma^0_{x(p)} & & \downarrow \tau & & \sigma_x & \downarrow \sigma^0_{x(p)} & & \downarrow & & \downarrow 0 \\
0 = H^0(\mathcal{H} \otimes \mathcal{L}^*) & \rightarrow & H^0,1(\mathcal{H} \otimes \mathcal{L}^*) & \rightarrow & H^0(\mathcal{H} \otimes \mathcal{L}^*) & \rightarrow & H^0(\mathcal{H} \otimes \mathcal{L}^*) & \rightarrow & H^0(\mathcal{H} \otimes \mathcal{L}^*)(p) & \rightarrow & 0.
\end{array}
\]

We can’t fill in a vertical map \( \tau \) in the middle column a priori, but we will be able to construct it. Fix a local generator \( s \) for \( \mathcal{L} \) near \( p \), so for small opens \( U \) containing \( p \), \( \mathcal{L}(U) = \mathcal{O}_X(U) \cdot s \). For a local coordinate \( z \) centered on \( p \), \( H^0(\mathcal{L}_p) = \{c/z \cdot s \mid c \in \mathbb{C} \} \) and \( H^0((\mathcal{H} \otimes \mathcal{L}^*)_p) = \{a \, dz \otimes s^* \mid a \in \mathbb{C} \} \), where \( s^* \) is the dual generator for \( \mathcal{L}^* \). One can check that the pairing \( (c/z \cdot s, a \, dz \otimes s^*) \mapsto 2\pi i \text{Res}(ac/z \, dz) = 2\pi iac \) is independent of \( s \) and defines an isomorphism \( \tau \) that makes the big diagram commute. Now, the result follows from a big diagram chase, using the five lemma a few times: if certain vertical arrows are isomorphisms, then so are the rest, which gives you the if and only if.

**Corollary 30.4.** If Serre duality holds for \( \mathcal{O}_X \), then it also holds for \( \mathcal{O}_X(D) \) for any divisor \( D \).

This is because Proposition 30.2 allows us to add and subtract points until we obtain \( D \).

This means we need to prove Serre duality for \( \mathcal{O}_X \), which no longer involves divisors, which is cool. There are two things to check: that integration \( \int_X : H^1,1(X) \rightarrow \mathbb{C} \) is an isomorphism and \( \sigma : H^0,1(X) \times H^0,1(\mathcal{H}_X) \rightarrow |C | \) defined by \( (\theta), \eta \mapsto \int_X \theta \wedge \eta \) is nondegenerate. Once we have this, \( h^0(\mathcal{H}_X) = h^0,1(X) \), and we need to show this is equal to \( g(X) \). Clearly, the integral is surjective, so we need to show that if \( \int_X \rho = 0 \), then there’s an \( \alpha \in \Omega^1,0(X) \) such that \( \rho = \partial \alpha \).

If \( \eta \in \Omega^1,0 \), then it has a complex conjugate \( \overline{\eta} \in \Omega^0,1(X) \), defined in local coordinates by \( \overline{f(z)} \, dz = f(z) \, d\overline{z} \). This defines a \( \mathbb{C} \)-linear conjugation map \( c : H^0(\mathcal{H}_X) \rightarrow H^0,1(X) \). For \( \eta \neq 0 \) in \( H^0,1(X) \), \( i\sigma^1(\eta, \overline{\eta}) = i \int \eta \wedge \overline{\eta} > 0 \), so \( c \) is injective. If we knew \( c \) were surjective, then \( \sigma^1 \) would be nondegenerate: surjectivity means that given a \( \theta \in \Omega^0,1(X) \), there’s an \( f \in \Omega^0(X) \) such that \( \overline{\theta} + \overline{\sigma^1f} = 0 \), i.e. \( \theta + \sigma^1f = 0 \), i.e. \( \partial \overline{\sigma^1f} = -\partial \theta \).

To prove this, we’ll invert the Laplacian.

**Theorem 30.5** (Invertibility of the Laplacian). Let \( \Delta = 2i\partial \overline{\partial} \), the Laplacian, \( \Omega^0(X) \) be the real differential 0-forms, and \( \Omega^2_0(X) \) be the 2-forms on \( X \) whose total integral is 0. Then, the Laplacian \( \Delta : \Omega^0(X) \rightarrow \Omega^2_0(X) \) is an invertible, \( \mathbb{R} \)-linear map.

By the maximum principle, \( \Delta u = 0 \) iff \( u \) is constant, and by Stokes’ theorem,

\[
\int_X \Delta u = 2i \int \partial(\overline{\partial} u) = 0,
\]

so \( \text{Im}(\Delta) \subset \Omega^2(X) \). The issue is surjectivity, so we want to solve \( \Delta u = \rho \). Moreover, we’ve already observed that this theorem implies Serre duality for \( X \). A similar argument will work for \( h^0,1(X) = g(X) \).

---

**Lecture 31.**

**Inverting the Laplacian:** 4/15/16

Last time, we stated Theorem 30.5 on the invertibility of the Laplacian, that the Laplacian operator \( \Delta = 2i\partial \overline{\partial} : \Omega^0(X) \rightarrow \Omega^2_0(X) \) is invertible on any compact, connected Riemann surface \( X \). We also showed that this suffices to prove Serre duality for \( \mathcal{O}_X \). We’ll also show today that it implies that \( 2g(X) = h^0(\mathcal{H}_X) + h^0,1(X) \), so that \( g(X) = h^0,1(\mathcal{H}_X) = h^0,1(X) \); these together imply the Riemann-Roch theorem.

Before we invert the Laplacian (and do some actual analysis!), let’s explain why \( 2g(X) = h^0(\mathcal{H}_X) + h^0,1(X) \), so that \( g(X) = h^0,1(\mathcal{H}_X) = h^0,1(X) \); these together imply the Riemann-Roch theorem.

Before we invert the Laplacian (and do some actual analysis!), let’s explain why \( 2g(X) = h^0(\mathcal{H}_X) + h^0,1(X) \), so that \( g(X) = h^0,1(\mathcal{H}_X) = h^0,1(X) \); these together imply the Riemann-Roch theorem.

Let \( \iota : H^0(\mathcal{H}_X) \rightarrow H^1(X) \) send \( \eta \rightarrow [\eta] \) and \( c : H^0(\mathcal{H}_X) \rightarrow H^0,1(X) \) be complex conjugation, as in last time; we discussed why Serre duality implies that \( c \) is an isomorphism. Then, using the invertibility of the Laplacian, one can check that the map \( H^0(\mathcal{H}_X) \oplus H^0,1(X) \rightarrow H^1(X) \) sending \( (\eta, [\theta]) \mapsto \iota(\eta) + \iota(c^{-1}([\theta])) \) is an isomorphism, using the Serre pairing, and therefore their dimensions satisfy \( h^0(\mathcal{H}_X) + h^0,1(X) = 2g(X) \).
Remark. One can also show $2g - 2 = (h^{0,1} - 1) + (h^0(\mathcal{X}^2) - 1)$ by taking Euler characteristics in sheaf cohomology for the exact sequence of sheaves

$$0 \longrightarrow \mathbb{C}_X \longrightarrow \mathcal{O}_X \stackrel{\partial}{\longrightarrow} \mathcal{X}_X \longrightarrow 0,$$

where $\mathbb{C}_X$ denotes the constant sheaf.

The argument for inverting the Laplacian is a prototype for understanding elliptic operators on manifolds; generalized and harder versions of these arguments are used in the Hodge theorem, the Atiyah-Singer index theorem, and more. Our proof will not set up the general machinery, and so will be somewhat different (making fewer choices) than the general Hodge-theoretic approach. We’ll also reuse some of the story in the noncompact case, as an ingredient in the proof of the uniformization theorem.

**Definition 31.1.** Let $X$ be any (not necessarily compact) Riemann surface and $\alpha \in \Omega^1_c(X)$. Then, we define the **Dirichlet norm** $\|\alpha\|$ of $\alpha$ to be the nonnegative real number such that

$$\|\alpha\|^2 = \int_X i\alpha \wedge \overline{\alpha}.$$ 

This will be the key, and it’s choice-free. Locally, $z = x + yi$ and $\alpha = p(z) \, dz$, so $dz \wedge d\overline{z} = -2i \, dx \wedge dy$. Thus,

$$i\alpha \wedge \overline{\alpha} = i|p|^2 \, dz \wedge d\overline{z} = 2|p|^2 \, dx \wedge dy,$$

and therefore $\|\alpha\|^2 \geq 0$, and $\|\alpha\| = 0$ iff $\alpha = 0$.

But we care about real $1$-forms $\alpha \in \Omega^1_c(X)$; there’s an isomorphism $\Omega^1_c(X) \rightarrow \Omega^{1,0}_c(X)$ sending $a \mapsto a^{1,0}$, so we also define the Dirichlet norm for $\alpha \in \Omega^1_c(X)$ by

$$\|\alpha\| = \|\alpha^{1,0}\| = -i \int_X \alpha^{1,0} \wedge \alpha^{0,1}.$$ 

This norm arises from an inner product $\langle \cdot, \cdot \rangle$ on $\Omega^1_c(X)$: if $\alpha, \beta \in \Omega^1_c(X)$, then

$$\langle a, b \rangle = -i \int_X a^{1,0} \wedge \beta^{0,1}.$$ 

Thus, $\langle a, b \rangle = \langle b, a \rangle \in \mathbb{R}$, and $\|a\|^2 = \langle a, a \rangle$, which means that $\|\cdot\|$ really is a norm.

**Lemma 31.2** (Cauchy-Schwarz).

$$\int |a \wedge b| \leq \|a\| \cdot \|b\|.$$

The proof is similar to the proof of other Cauchy-Schwarz inequalities for inner product spaces; see p. 79 of the textbook.

**Definition 31.3.** This allows us to define the **Dirichlet inner product** for smooth functions as

$$\langle f, g \rangle_D = \langle df, dg \rangle = 2i \int_X \partial f \wedge \overline{g}.$$ 

This makes sense as long as at least one of $f$ and $g$ is compactly supported. This allows us to define the **Dirichlet norm** on functions $f \in \Omega^0_c(X)$ by

$$\langle f \rangle_D^2 = \langle f, f \rangle_D = 2i \int_X \partial f \wedge \overline{f}.$$ 

This is intrinsic, and doesn’t change if you add a constant to a compactly supported function. The first hint that the Dirichlet norm relates to the Laplacian is that

$$\langle f, g \rangle_D = \int g \cdot \Delta f = \int f \cdot \Delta g.$$
Indeed,

\[
\langle f, g \rangle_D = 2i \int_X \partial f \wedge \overline{\partial} g
= 2i \int_X \partial(f \overline{\partial} g) - f \overline{\partial} \partial g
= 2i \int d(f \overline{\partial} g) - f \overline{\partial} \partial g
= \int f \cdot \Delta g,
\]

since the integral of \( d \) of something vanishes by Stokes’ theorem, proving (31.4).

**Proof of Theorem 30.5.** To invert the Laplacian, we want to solve \( \Delta u = \rho \) given a \( \rho \in \Omega^0_0(X) \). We know that \( \Delta u = \rho \) iff \( \int v \Delta u = \int \rho \) for all \( v \in \Omega^0_0(X) \), i.e. \( \langle u, v \rangle_D = \int_X \rho \) for all \( v \). This is the key insight that will allow us to use Hilbert space techniques.

At this point, we must assume \( X \) is compact and connected. We’ll proceed in a few stages, which are absolutely typical for PDE problems involving elliptic PDEs.

1. First, we’ll find a weak solution \( u \) to \( \Delta u = \rho \), meaning \( u \) is an element of the Hilbert space completion \( H \) of \( \Omega^0(X)/\mathbb{R} \) (functions modulo constants), completed with respect to the Dirichlet norm, satisfying \( \langle u, v \rangle_D = \int_X \rho \) for all \( v \). If \( u \) were smooth, we’d be done.
2. The second step is to show that \( u \) is smooth, which requires two steps.
   a. First, we’ll show that \( u \), which is formally a Cauchy sequence of functions, can be represented by a **locally** \( L^2 \) function on \( X \) (that is, \( X \) can be covered by coordinate charts on which \( u \) is an honest \( L^2 \) function).
   b. In a process called **elliptic regularity**, we’ll show that \( L^2_{\text{loc}} \) weak solutions to \( \Delta u = \rho \) are necessarily smooth.

These together imply the theorem, and form a strategy that’s common to many other problems in the same type; however, the intrinsic definition of the Dirichlet inner product makes this particularly clean.

For part 1, let’s find a weak solution. \( \langle \cdot, \cdot \rangle_D \) is a nondegenerate inner product on \( \Omega^0(X)/\mathbb{R} \), because the only degeneracy on \( \Omega^0(X) \) was due to functions \( f \) with \( df = 0 \), and we’ve just killed those. Let \( H \) be the real Hilbert space completion of \( (\Omega^0(X)/\mathbb{R}, \langle \cdot, \cdot \rangle_D) \), i.e. the vector space of Cauchy sequences \( (f_n) \in \Omega^0(X)/\mathbb{R} \) quotiented by the subspace of sequences converging to 0 in the Dirichlet norm.

\( H \) inherits the Dirichlet inner product:

\[
\langle (f_n), (g_n) \rangle_D = \lim_{n \to \infty} \langle f_n, g_n \rangle_D \in \mathbb{R},
\]

and is a Hilbert space, so Cauchy sequences converge. Now, we can use a very standard tool, the Riesz representation theorem.

**Theorem 31.5** (Riesz representation theorem). Let \( H \) be a Hilbert space and \( L : H \to \mathbb{R} \) be a bounded linear functional; then, there’s a \( u \in H \) such that \( L(v) = \langle u, v \rangle \) for all \( v \in H \).

So we need to prove the following theorem.

**Theorem 31.6** (Boundedness). \( \hat{\rho} : \Omega^0(X)/\mathbb{R} \to \mathbb{R} \) defined by sending \( f \mapsto \int f \rho \) is bounded in the Dirichlet norm, meaning \( \|\hat{\rho}(f)\|_D \leq C\|f\|_D \) for some constant \( C \in \mathbb{R} \).

Assuming these theorems, \( \hat{\rho} \) extends to a bounded functional \( H \to \mathbb{R} \) and gives a solution \( u \in H \) to \( \langle u, v \rangle_D = \int \rho \) for all \( v \), which is part (1).

First off, it suffices to prove that \( \hat{\rho} \) is bounded when \( \rho \) is compactly supported inside some coordinate chart \( U \): by the compactness of \( X \), we can choose a finite open cover \( U_1, \ldots, U_n \), and since \( \int \rho = 0 \), then \( \rho = 0 \) in \( H^2(X) \), so \( \rho = d\theta \) for a 1-form \( \theta \). Let \( \{\chi_i\} \) be a partition of unity subordinate to \( \{U_i\} \). Then, we can use the partition of unity to write \( \theta \) as a sum of forms supported on coordinate charts, and therefore their derivatives sum to \( \rho \).

\[\text{This is well-defined, because if one adds a constant to } f, \text{ that’s multiplied by the integral of } \rho, \text{ which is 0}.\]
Recall that we’re in the middle of inverting the Laplacian: if $X$ is a compact Riemann surface and $\rho$ is a given 2-form such that $\int \rho = 0$, we’d like to solve the Poisson equation $\Delta u = \rho$. This is the same equation that governs the behavior of a point of charge in an electrical field. In the plane, this has an explicit solution, given by convolution with the function that would arise if $\rho$ were the delta function at $0$; this explicit solution will show up again, but not today.

Recall that $H$ denotes the Hilbert space completion of $\Omega^0(X)/\mathbb{R}$ (i.e. modulo constants), and we defined the Dirichlet norm on it by

$$\langle f, g \rangle_D = 2i \int_X \partial f \wedge \partial g.$$ 

Right now, we’re looking for a weak solution $u$ to $\Delta u = \rho$, so a $u \in H$ such that

$$\langle u, f \rangle_D = \hat{\rho}(f) = \int_X f \rho.$$ 

By the Riesz representation theorem, it suffices to show $\hat{\rho}$ is bounded, meaning $|\hat{\rho}(f)| \leq C_\rho \|f\|_D$ for some $C_\rho \in \mathbb{R}$ (in particular, it’s independent of $f$). Finally, we showed that it suffices to treat the case where $\text{supp } \rho$ is contained in a coordinate neighborhood.

**Proposition 32.1.** If $\Omega \subset \mathbb{C}$ is a bounded, convex open set, $U$ is a neighborhood of $\Omega$, and $\psi : U \to \mathbb{R}$ is a $C^\infty$ function with average $\bar{\psi} = \frac{1}{\mu(\Omega)} \int_\Omega \psi(z) \, d\mu$,

where $\mu$ denotes the usual Lebesgue measure, then there exists a constant $c$ such that

$$\int_\Omega |\psi - \bar{\psi}|^2 \, d\mu \leq c \int_\Omega |\nabla \psi|^2 \, d\mu.$$ 

This proposition implies that $\hat{\rho}$ is bounded, because if $\psi$ and $\hat{\psi}$ are as in Proposition 32.1, then

$$\|\psi - \bar{\psi}\|_{L^2(\Omega)}^2 = \int_\Omega |\psi - \bar{\psi}|^2 \, d\mu \leq c \int_\Omega |\nabla \psi|^2 \, d\mu = c \int_\Omega (\psi \Delta \psi) \, d\mu$$

$$\leq c \int_X \psi \cdot \Delta \psi = c \|\psi\|_D.$$ 

If $\rho$ is supported in $\Omega$, then since $\int_X \rho = 0$, then

$$\hat{\rho}(\psi) = \int_X \psi \rho = \int_X (\psi - \bar{\psi}) \rho = \int_\Omega (\psi - \bar{\psi}) \rho.$$ 

Hence, by the Cauchy-Schwarz inequality,

$$|\hat{\rho}(\psi)| \leq \|\rho\|_{L^2(\Omega)} \|\psi - \bar{\psi}\|_{L^2(\Omega)}$$

$$\leq \sqrt{c}\|\rho\|_{L^2(\Omega)} \|\psi\|_D,$$ 

which shows that $\hat{\rho}$ is bounded assuming Proposition 32.1. To prove this, we’ll need the following lemma, whose proof can be found on page 122 of the book, and is pretty elementary, depending on polar coordinates and double integrals.

**Lemma 32.2.** Let $\Omega$ and $\psi$ be as in Proposition 32.1 and $d = \text{diam}(\Omega)$. Then,

$$|\psi(z) - \bar{\psi}| \leq \frac{d^2}{2\mu(\Omega)} \int_{\Omega} \frac{1}{|z - w|} |\nabla \psi(w)| \, d\mu_w.$$ 

The intuition, familiar from the mean value theorem, is that the deviation of a function from its average is controlled by its gradient.

---

52The definition was given incorrectly last time; I’ve corrected it in my notes, but check the book to be sure.
Proof of Proposition 32.1. Recall that if $f$ and $g$ are integrable functions on $\mathbb{R}^2$, their convolution is the function $f * g : \mathbb{R}^2 \to \mathbb{R}$ given by

$$(f * g)(z) = \int_{\mathbb{R}^2} f(w)g(z-w) \, d\mu_w.$$

If we define

$$g(z) = \begin{cases} |\nabla \psi(z)|, & z \in \Omega \\ 0, & z \not\in \Omega \end{cases} \quad \text{and} \quad K(z) = \begin{cases} \frac{d^2}{2\mu(\Omega)} \frac{1}{|z|}, & |z| \leq d \\ 0, & |z| > d, \end{cases}$$

then Lemma 32.2 implies that $|\psi(z) - \tilde{\psi}| \leq (K*g)(z)$. We’d like $\|\psi - \tilde{\psi}\|_{L^2(\Omega)}^2 \leq C\|g\|_{L^2(\Omega)}^2$ for some constant $C$; at this point things follow almost formally.

Recall that if $\nu$ is a reasonable norm on a space of functions (the $L^2$ norm certainly satisfies this), then

$$\nu\left(\int f(\cdot, w) \, d\mu_w\right) \leq \int \nu(f(\cdot, w)) \, d\mu_w.$$

If $\nu$ is also translation-invariant, meaning $\nu(f(\cdot + c)) = \nu(f)$, then this provides an estimate for the norm of the convolution:

$$\nu(f_1 * f_2) = \nu\left(\int f_1(w)f_2(\cdot - w) \, d\mu_w\right) \leq \int \nu(f_1(w)f_2(\cdot - w)) \, d\mu_w \leq \int |f_1(w)|\nu(f_2(\cdot - w)) \, d\mu_w \leq \int |f_1(w)|\nu(f_2) \, d\mu_w = \|f_1\|_{L^1, \nu} \nu(f_2).$$

In our case, with $\nu = \|\cdot\|_{L^2(\Omega)}$,

$$\|f_1 * f_2\|_{L^2(\Omega)} \leq \|f_1\|_{L^1(\Omega)}\|f_2\|_{L^2(\Omega)},$$

and therefore

$$\|\psi - \tilde{\psi}\|_{L^2(\Omega)}^2 \leq \|K*g\|_{L^2(\Omega)}^2 \leq \|K\|_{L^1(\Omega)}^2 \|g\|_{L^2(\Omega)}^2,$$

and $\|K\|_{L^1(\Omega)} = \pi d^3/\mu(\Omega)$, so we win. $\Box$

Essentially, the boundedness of $\rho$ comes from a pointwise inequality which, by general properties of convolutions, is strong enough to imply an integral inequality.

**Regularity: weak solutions are solutions.** The second half of the theorem involves showing that all weak solutions are honest solutions. This is really an argument in the land of Sobolev spaces; we have control over the derivative of $u$, and so we could think of this as an argument in $L^2$, as more general approaches to these kinds of equations do; however, we’re not going to use this formalism, since we don’t need all of it to approach this problem.

By definition, $u$ is represented by a Cauchy sequence $(u_n)$ in $\Omega^0(X)/\mathbb{R}$, so that for all $\chi \in \Omega^0(X)$,

$$\int_X \Delta u_n \cdot \chi = \int_X u_n \cdot \Delta \chi \to \bar{\rho}(\chi).$$

We’ll upgrade the abstract solution $u$ to a concrete one, which will be a function that’s locally $L^2$; this isn’t quite what we’re looking for, but is easier to work with than a Cauchy sequence, and will help us place it back in $\Omega^0(X)/\mathbb{R}$. Specifically, we’ll find a function $u : X \to \mathbb{R}$ and an atlas $\mathcal{U}$ of small coordinate discs such that for each $U \in \mathcal{U}$, $u|_U \in L^2(U)$ (so it’s square-integrable in local coordinates) and $u_n|_U \to u|_U$ in $L^2(U)$.53

Fix an open set in $X$ identified with a bounded, convex neighborhood $\Omega \subset \mathbb{C}$, and let $\pi_n$ denote the average of $u_n|_\Omega$ by adding constants to $u_n$, we may assume $\pi_n = 0$ (though only on $\Omega$). Now, we’ll show that these $u_n$ are Cauchy in $L^2(\Omega)$: because we’ve forced the averages to be 0,

$$\|u_n - u_m\|_{L^2(\Omega)}^2 = \int_\Omega |u_n - u_m|^2 \, d\mu = \int_\Omega |(u_n - u_m) - (u_n - u_m)|^2 \, d\mu.$$

53If we put a metric on $X$, this argument can be rephrased globally, but the textbook takes this approach and so will we.
By Proposition 32.1, there’s a constant $C(\Omega)$ depending only on $\Omega$ such that
\[
\leq C(\Omega) \int_{\Omega} |\nabla (u_n - u_m)|^2 \, d\mu
\leq C(\Omega) \|u_n - u_m\|_D^2 \to 0.
\]
Thus, $u_n|\Omega$ is Cauchy in the $L^2$-norm, and this norm is complete, so there’s some limit $u_\Omega \in L^2(\Omega)$. We need to extend this to all of $X$, so let $A$ denote the largest open set containing $\Omega$ on which there exists a function $u \in L^2(A)$ such that $u_n \to u$ in the $L^2(A)$-norm. $A$ is by definition a nonempty open set, and since $X$ is connected, we’d like to show $A$ is closed.

Suppose that $A$ isn’t closed, so we can pick a $y \in \overline{A} \setminus A$ and a small coordinate neighborhood $\Omega' \subset X$ containing $y$. Just as on $\Omega$, there’s a sequence $(c'_n)$ of real numbers such that $u_n - c'_n$ converges in $L^2(\Omega')$. Thus, for any $x \in A \cap \Omega'$, both $u_n$ and $c'_n$ converge in $L^2$ on a neighborhood of $x$, so $c'_n$ converges to some $c \in \mathbb{R}$, and therefore $u_n$ converges in $L^2(\Omega')$, which is a contradiction. Hence, $A$ is closed and open, and hence $A = X$.

Next time, we’ll prove Weyl’s lemma, that this solution must be smooth.

---

**Lecture 33.**

---

**Weyl’s Lemma: 4/20/16**

“This whole proof is full of what seem to me, as a non-analyst, somewhat devious maneuvers.”

Today, we will finally finish off the proof of the Riemann-Roch theorem; we’ve proven everything contingent on the assumption that we can invert the Laplacian; that is, if $\rho \in \Omega^2(X)$ is such that $\int_X \rho = 0$, we would like to find a solution $u \in \Omega^0(X)$ such that $\Delta u = \rho$.

So far, we have showed that there’s a weak solution, specifically a Cauchy sequence $(u_n) \subset \Omega^0(X)$, so that $\|u_m - u_n\|_D \to 0$ as $m, n \to \infty$ and $(u_n, \chi)_D \to \int_X \chi \rho$ for all $\chi \in \Omega^0(X)$. After adding constants to the $u_n$, we may assume there’s a $u : X \to \mathbb{R}$ such that $u_n \to u$ in $L^2$ on small coordinate discs $U \subset X$. Equivalently, if $\text{supp}(\chi) \subset U$, then $\int_U u_n \cdot \chi$ converges to $\int_U u \chi$, and therefore $\int_U u_n \cdot \Delta \chi \to \int_U u_n \cdot \Delta \chi$. Globally, this means that for every $f \in \Omega^0(X)$,
\[
\int_X u_n \cdot \Delta f \to \int_X u \cdot \Delta f = \int_X f \rho.
\]
That is, $u$ is a weak $L^2$ solution to $\Delta u = \rho$; if $u$ were smooth, we would be done. In other words, we need to prove the following.

**Lemma 33.1** (Weyl). Let $\Omega \subset \mathbb{C}$ be a bounded open set and $\rho$ a smooth 2-form on $\Omega$. Suppose $u \in L^2(\Omega)$ is such that for all compactly supported $\chi$ on $\Omega$,
\[
\int_{\Omega} u \Delta \chi = \int_{\Omega} \chi \rho.
\]
Then, $u$ is $C^\infty$ and $\Delta u = \rho$.

Since smoothness is a local condition, this suffices to show that we have a solution on all of $X$.

One might do this systematically, using Sobolev spaces, elliptic inequalities, etc., but we’ll follow the text’s more hands-on approach.

**Remark.** If $\rho \in \Omega^2(\mathbb{C})$ is compactly supported, one can find an explicit solution to $\Delta u = \rho$. The *Newton kernel* is
\[
K(z) = \frac{1}{2\pi} \log|z|,
\]
which is locally integrable and $C^\infty$ everywhere except 0; then, if $f \in \Omega^0(\mathbb{C})$, consider the convolution
\[
K \ast f = \int_{\mathbb{C}} K(y)f(\bullet - y) \, d\mu_y = \lim_{r \to 0^+} \int_{\{ |y| \geq r \}} K(y)f(\bullet - y) \, d\mu_y
\]

54 This hypothesis is only needed so that the convolution is well-defined; you can also relax this to a suitable integrability condition.
is an improper integral, which is why we evaluate it with a limit. By differentiating under the integral sign, \((K * f)' = K * (f')\), and in particular, \(K * f\) must be \(C^\infty\). Moreover, if \(\rho\) and \(\sigma\) are compactly supported, \(\Delta(K * \rho) = \rho\) and \(K * (\Delta\sigma) = \sigma\); this can be proven very classically, and boils down to the fact that \(\Delta K\) is the Dirac delta at the origin.\(^{55}\)

It’s possible to write down explicit solutions on the torus, using Fourier analysis, and the textbook does this, though we won’t. In general, explicit solutions are quite difficult to find.

**Proposition 33.2.** Weyl’s lemma for \(\rho = 0\) (and all domains \(\Omega\)) implies Weyl’s lemma for all \(\rho\).

That is, if you can show that weakly harmonic functions are smooth, then the rest of Weyl’s lemma holds.

**Proof.** We want to show that \(\Delta u = \rho\) weakly in \(L^2\) implies \(u\) is smooth on \(\Omega\), where we assume that \(\Delta u = 0\) on \(L^2\) implies \(u\) is smooth. To prove this, it suffices to prove smoothness on coordinate neighborhoods \(\Omega' \subset \Omega\), where \(\Omega\) contains an \(\varepsilon\)-neighborhood of \(\Omega'\) (where \(\varepsilon\) can be arbitrarily small); since \(\Omega\) can be exhausted by such sets, this suffices.

In particular, we can find a compactly supported \(\rho'\) on \(\Omega\) such that \(\rho' = \rho\) on the \(\varepsilon/2\)-neighborhood of \(\Omega'\). Then, the equation \(\Delta u' = \rho'\) on \(\Omega\) has the solution \(u' = K * \rho'\), which is smooth, by the remark. We’d like \(u\) to be smooth, which is equivalent to \(u - u'\) being smooth, but \(\Delta(u - u') = \rho - \rho' = 0\) weakly in \(L^2(\Omega)\), so the assumption shows \(u - u'\) is smooth, and therefore that \(u\) is.

Thus, we should focus on weakly harmonic functions \(u\), such that \(\int u \Delta \chi = 0\) for all \(\chi \in C^\infty_\varepsilon(\Omega)\). To show that \(u\) is smooth, we’ll show that \(u = B * u\) for some smooth kernel \(B\).

Let \(\beta : [0, \infty) \to \mathbb{R}\) be a smooth cutoff function, so \(\beta\) is constant in a neighborhood of 0, is 0 when \(x > \varepsilon\) (for some chosen \(\varepsilon > 0\)), and is normalized such that

\[
\int_0^\varepsilon r \beta(r) \, dr = \frac{1}{2\pi}. \tag{33.3}
\]

**Lemma 33.4.** If \(\phi\) is \(C^\infty\) and \(\Delta \phi = 0\), then \(B * \phi = 0\), where \(B(z) = \beta(|z|)\).

**Proof.** This follows from the mean-value property for harmonic functions: since \(\phi\) is harmonic, then if \(r > 0\), \(\phi(z)\) is equal the average of \(\phi\) on a circle of radius \(r\) around \(z\). That is,

\[
\phi(z) = \frac{1}{2\pi} \int_0^{2\pi} \phi(z + re^{i\theta}) \, d\theta. \tag{33.5}
\]

By translation-invariance, it suffices to show that \((B * \phi)(0) = \phi(0)\), since \(\phi\) is harmonic on a neighborhood of a closed disc centered at the origin. Here we may use calculus:

\[
B * \phi(0) = \int_C B(-z) \phi(z) \, dm_z = \int_C \beta(|z|) \phi(z) \, dm_z.
\]

Switching to polar coordinates,

\[
= \int_0^\infty \int_0^{2\pi} r \beta(r) \phi(r e^{i\theta}) \, d\theta \, dr.
\]

By (33.5), this is

\[
= \int_0^\infty 2\pi r \beta(r) \phi(0) \, dr = \phi(0),
\]

since we fixed this integral in (33.3).

As a corollary, if \(\phi\) is \(C^\infty\) and \(\Delta \phi\) is compactly supported inside \(J \subset \mathbb{C}\), then \(B * \phi = \phi\) outside the \(\varepsilon\)-neighborhood of \(J\).

\(^{55}\)\(K * \rho\) is not compactly supported; the convolution smears it out; but taking the Laplacian gets us back to \(\rho\), which is.
Proof of Lemma 33.1. By Proposition 33.2, it suffices to show that if \( \Delta u = 0 \) weakly on \( \Omega \), then \( u \) is smooth. In particular, by the proof of that proposition, it suffices to prove smoothness on \( \Omega' \). We’ll show that \( u = B \ast u \) on \( \Omega' \), which proves \( u \) is smooth because \( B \) is; this is equivalent to \( \langle \chi, u - B \ast u \rangle = 0 \) for all \( \chi \in C^\infty_c(\Omega') \) (here, \( \langle \cdot, \cdot \rangle \) is the \( L^2 \) inner product). Since \( \langle f, g \ast h \rangle = \langle g \ast f, h \rangle \) (by checking that these actually are the same integral), then

\[
\langle \chi, u - B \ast u \rangle = \langle \chi, u \rangle - \langle B \ast \chi, u \rangle = \langle \chi - B \ast \chi, u \rangle.
\]

Thus, we want to show that \( \langle \chi - B \ast \chi, u \rangle \) is identically zero, given that \( \langle \Delta \phi, u \rangle = 0 \) for all test functions \( \phi \).

Since \( \Delta(K \ast \chi) = \chi \), then \( \Delta(K \ast \chi) = 0 \) outside \( \text{supp}(\chi) \subset \Omega' \), and therefore \( B \ast (K \ast \chi) = K \ast \chi \) outside \( \Omega' \). If \( h = K \ast \chi - B \ast K \ast \chi = K \ast (\chi - B \ast \chi) \) (since convolution is commutative), then \( h \) is compactly supported in \( \Omega \), so we may use it as a test function. That is, \( \langle u, \Delta h \rangle = 0 \), i.e. \( \langle u, \Delta K \ast (\chi - B \ast \chi) \rangle = 0 \), but this is \( \langle u, \chi - B \ast \chi \rangle \), i.e. \( u = B \ast u \).

Next time, we will march on towards uniformization.

Lecture 34.

**Uniformization: 4/22/16**

Today, we’re going to begin talking about the uniformization theorem, a strong, yet amazingly simple, classification theorem. It’s very rare to find such a clean classification theorem, especially for noncompact spaces.

**Theorem 34.1 (Uniformization).** Any connected, simply connected Riemann surface is isomorphic to one of \( S^2 \), \( \mathbb{C} \), or \( \mathbb{H} \).

The case of bounded domains in the plane is the Riemann mapping theorem.

**Corollary 34.2.** Any connected Riemann surface is isomorphic to one of the following:

1. the Riemann sphere \( S^2 \);
2. \( \mathbb{C} \) or \( \mathbb{C}^* \cong \mathbb{C}/\mathbb{Z} \);
3. \( \mathbb{C}/\Lambda \), where \( \Lambda \) is a lattice; or
4. \( \mathbb{H}/\Gamma \), where \( \Gamma \) is a Fuchsian group acting freely.

The term “uniformization theorem” is used to describe either or both of Theorem 34.1 and Corollary 34.2, and the two aren’t usually distinguished.

Corollary 34.2 follows from Theorem 34.1 by taking the universal cover \( \tilde{X} \) of a connected Riemann surface \( X \), so \( X = \tilde{X}/G \), where \( G \leq \text{Aut } \tilde{X} \) is discrete and acts freely on \( \text{Aut } \tilde{X} \). Then, we know \( \text{Aut}(S^2) = \text{PSL}_2(\mathbb{C}) \) has no nontrivial freely-acting subgroups, since all Möbius transformations fix points, and \( \text{Aut } \mathbb{C} \) is the set of maps \( z \mapsto az + b \) for \( a, b \in \mathbb{C} \) and \( a \neq 0 \). These are groups of translations \( \mathbb{Z} \) or lattices.

**History.** In 1882, Poincaré and Felix Klein simultaneously announced uniformization for algebraic curves, a somewhat simpler case; the name “uniformization” meant that this result allowed one to describe a surface in terms of a fundamental domain, and in particular in terms of one coordinate.

They proofs they furnished were rooted in the idea that the number of parameters needed to specify an algebraic curve of a given genus is the same as the number of parameters describing \( \mathbb{H}/\Gamma \) with \( \Gamma \) acting freely. This was more or less what Klein had, so work continued over the next 25 years to put this (and other results, such as the Riemann mapping theorem, which Riemann didn’t get quite right) on sound footing. This work was done by Poincaré, Schwarz, Harnack, and Hilbert.

In 1907, Poincaré and Koebe announced proofs for the uniformization theorem; Koebe devoted much of his life to this theorem, and produced several different proofs. The proof we’ll give is one of his.

**Consequences of uniformization.** This general classification result has some nice corollaries. This first one is somewhat minor.

**Corollary 34.3 (Rado’s theorem).** Every connected Riemann surface has a countable atlas.

Pass to the universal cover, which has a countable atlas, and use this to define charts on the quotient. In particular, this means that our proof shouldn’t invoke second-countability in any way!
This next corollary is more important. A Riemannian metric on a Riemann surface $X$ is compatible with the complex structure $J : TX \to TX$ if in the angles defined on each tangent space by the metric, $J$ is always a rotation by $\pi/2$.

**Corollary 34.4.** Every Riemann surface admits a constant-curvature Riemannian metric compatible with the complex structure.

The curvature is $+1$ for $S^2$, $0$ for $\mathbb{C}$, $\mathbb{C}^*$, and $\mathbb{C}/\mathbb{Z}$, and $-1$ for $\mathbb{H}/\Gamma$. The main idea, which occurred to Riemann as he got on a bus, is that $\text{Aut} \mathbb{H} = \text{PSL}_2(\mathbb{R})$ is also the isometry group of the hyperbolic metric on $\mathbb{H}$.

Thurston's geometrization conjecture for compact 3-manifolds (proved by Hamilton-Perelman) is analogous, determining to what extent a 3-manifold admits a constant-curvature Riemannian metric. Both this and uniformization have the spirit that most of the manifolds they classify are hyperbolic; the geometrization theorem is one of the major developments in geometry and topology over the past quarter-century, and uniformization is widely regarded as the culmination of the theory of one complex variable.

**First steps.** As such, it will take us a while to prove it. Much like Riemann-Roch, we'll prove it from the top down, reducing it to the analytic problem of solving Poisson's equation on a non-compact surface.

**Definition 34.5.** If $X$ is a Riemann surface, a slit mapping is a meromorphic map $f : X \to S^2$ which defines a biholomorphic equivalence with $S^2 \setminus I$, where $I \subset \mathbb{R} \subset \mathbb{C} \subset S^2$ is either a point or a closed, bounded interval.

The idea is that the map hits everything except a slit in the sphere (which could have length 0).

**Lemma 34.6.** If $X$ admits a slit mapping, then $X \cong \mathbb{C}$ or $X \cong \mathbb{D}$ (and so also $\mathbb{H}$).

**Proof.** If $I$ is a single point $\{a\}$, then $X \cong S^2 \setminus pt \cong \mathbb{C}$.

If $I = [a,b]$ is an interval, then $X \cong S^2 \setminus [a,b]$, so by some Möbius map, $X \cong S^2 \setminus [-1,1]$. Define $g : S^2 \to S^2$ by $g(z) = (1/2)(z + z^{-1})$; then, $g$ maps $\mathbb{D}$ isometrically onto $S^2 \setminus [-1,1]$. Since $\deg g = 2$, then $g$ is surjective, and its fiber at $g(z)$ is $\{z, z^{-1}\}$. Thus, it's injective on $\mathbb{D}$, and $g(\mathbb{D}) = S^2 \setminus g(\partial \mathbb{D})$, and $g(e^{i\theta}) = \cos \theta \in [-1,1]$, and thus $g(\mathbb{D}) = S^2 \setminus [-1,1]$.

One big step towards uniformization is the following, which we will prove.

**Theorem 34.7.** Any non-compact, simply connected Riemann surface admits a slit mapping.

This implies uniformization for noncompact surfaces. If $X$ is simply connected and compact, then $H^1(X) = \text{Hom}(\pi_1(X), \mathbb{R}) = 0$, so $g(X) = 0$ (as we proved). Thus, by the weak Riemann-Roch theorem, $X \cong S^2$. So the hard part is the noncompact case.

**Definition 34.8.** Let $X$ be a Riemann surface and $f : X \to S^2$ be meromorphic. One says that $\text{Im}(f) \to 0$ at $\infty$ in $X$ if for all $\varepsilon > 0$, there exists a compact $K \subset X$ such that $|\text{Im}(f)| \leq \varepsilon$ on $X \setminus K$.

**Proposition 34.9.** Let $X$ be a noncompact, simply connected Riemann surface and $f : X \to S^2$ be meromorphic, with a single, simple pole such that $\text{Im}(f(z)) \to 0$ at $\infty$ on $X$. Then, $f$ is a slit mapping.

**Proof.** First, let $H_+$ denote the open northern hemisphere in $S^2$ and $H_-$ denote the open southern hemisphere in $S^2$, so that $H_+ \cong \mathbb{H}$ and $H_- \cong \overline{\mathbb{H}}$. Let $X_\pm = f^{-1}(H_\pm)$; by the open mapping theorem, $X_+ \cup X_-$ is dense in $X$. Let $f_\pm = f|_{X_\pm}$.

**Claim.** $f_+$ and $f_-$ are proper maps.

**Proof.** Suppose $B \subset H_+$ is compact; thus, there exists an $\varepsilon > 0$ such that $\text{Im} z > \varepsilon$ for all $z \in B$, and there's a compact $K \subset X$ such that $|\text{Im}(f(z))| < \varepsilon$ on $X \setminus K$. In particular, $f_+^{-1}(B) \subset K$, and therefore $f_+^{-1}(B) = f^{-1}(B)$ is a closed subset of $X$ contained in the compact set $K$, and therefore is compact. The same argument works for $f_-$.\footnote{This notation might be confusing: this is the conjugate of $\mathbb{H}$, not the closure.}
Since \( f_+ \) and \( f_- \) are proper, they have degrees. The next step is to prove that \( \deg(f_+) = \deg(f_-) = 1 \); the key to this is that \( f \) has a simple pole at some \( p \in X \), and no other poles.

Since \( f \) has a single, simple pole at \( p \), it maps a neighborhood \( U \) of \( p \) diffeomorphically to its image. This in particular implies \( X_+ \) and \( X_- \) are nonempty.

Let \( K \subset X \) be compact such that \( \Im f \) is a contradiction. First, suppose \( \deg(f_+) \geq 2 \), so there exist sequence \( (x_n) \) and \( (x'_n) \) in \( H_+ \) such that \( f(x_n) = f(x'_n) = \infty \) and either \( x_n \neq x_n \) or \( f'(x_n) = 0 \). Hence, \( \Im(f(x_n)) = \Im(f(x'_n)) = n \geq 1 \), so \( x_n, x'_n \in K \) for all \( n \); since \( K \) is compact, there exist convergent subsequences of each; replace each sequence by a convergent subsequence, so we can assume \( x_n \to x_\infty \) and \( x'_n \to x'_\infty \). In particular, \( f(x_\infty) = f(x'_\infty) = \infty \), so for \( n \gg 0 \), \( x_n, x'_n \in f(U) \), since \( f(U) \) is a neighborhood of \( \infty \). On the other hand, \( f \) is injective on \( U \) with nonvanishing derivative, which is a contradiction to our construction.

Thus, \( \deg f_+ = 1 \), and similarly for \( f_- \); we’ll finish the rest of the proof next time.

---

**Uniformization, II: 4/25/16**

Recall that we’re in the middle of proving the uniformization theorem, that all connected Riemann surfaces are quotients of \( \mathbb{C} \) or \( \mathbb{H} \) by groups acting freely.

Keep in mind that it’s quite easy to create very complicated Riemann surfaces: closed sets in \( \mathbb{C} \) can be very complicated, including Cantor or Julia sets, and their complements are open subsets of \( \mathbb{C} \) and hence Riemann surfaces (typically, these are quotients of \( \mathbb{H} \)).

Riemann-Roch addressed the compact case, so we’d like to show that a connected, simply-connected, noncompact Riemann surface \( X \) is isomorphic to either \( \mathbb{C} \) or \( \mathbb{H} \). To that end, we were showing that if \( f : X \to S^2 \) is meromorphic with a single pole \( p \) that’s simple, then \( f \) is a slit mapping, i.e. an isomorphism \( X \cong S^2 \setminus I \), where \( I \) is an interval or a point inside \( \mathbb{R} \subset \mathbb{C} \subset S^2 \).

**Continuation of Proof of Proposition 34.9.** We defined \( X_\pm = f^{-1}(H_\pm) \), where \( H_+ = \mathbb{H} \subset S^2 \) and \( H_- = \mathbb{H} \), and let \( f_\pm : X_\pm \to H_\pm \) be the restriction of \( f \) to \( H_+ \) or \( H_- \), respectively. Importantly, we proved \( f_+ \) and \( f_- \) are proper of degree \( 1 \), and hence are biholomorphic maps \( X_\pm \to H_\pm \).

Next, we can show \( f \) is injective: if not, there exist \( x_1, x_2 \in X \) such that \( f(x_1) = f(x_2) \). The image \( y = f(x_1) \) must be inside \( \mathbb{R} \subset S^2 \), since \( f_+ \) and \( f_- \) are injective; let \( D_1 \) and \( D_2 \) be disjoint open neighborhoods of \( x_1 \) and \( x_2 \), respectively. By the open mapping theorem, \( f(D_1) \) and \( f(D_2) \) are open neighborhoods of \( y \). Let \( N = f(D_1) \cap f(D_2) \), which is also a neighborhood of \( y \) open in \( S^2 \), so there’s a \( z \in N \cap H_+ \), so there are \( x'_1 \in D_1 \cap X_+ \) and \( x'_2 \in D_2 \cap X_- \) such that \( z = f(x'_1) = f(x'_2) \), which is a contradiction; hence, \( f \) is injective.

Thus, we find that \( f : X \to S^2 \setminus I \) is a biholomorphism for some closed \( I \subset \mathbb{R} \).\(^{57}\) We’d like to show that \( I \) is path-connected; since \( X \) is simply connected, then so must \( S^2 \setminus I \).

Intuitively, this makes sense: if \( I \) isn’t path-connected, then there’s a loop around one path component but not the others isn’t contractible. Formally, if \( I \) isn’t path-connected, there exist \( x_1, x_2 \in I \) and a \( y \in \mathbb{R} \setminus I \) such that \( y \) lies between \( x_1 \) and \( x_2 \) (without loss of generality, we can assume \( y \neq \infty \)). Let \( \gamma \) be a loop through \( y \) and \( \infty \); since \( S^2 \setminus I \) must be simply connected, \( \gamma \) is contractible in it, and therefore must be contractible in \( S^2 \setminus \{x_1, x_2\} \). However, a loop around a point \( p_1 \) generates the first homology of \( H_1(S^2 \setminus \{p_1, p_2\}) \cong \mathbb{Z} \), which is a contradiction. Thus, \( I \) must be path-connected.

Moreover, we know that \( I \) must be closed in \( \mathbb{R} \cup \infty \), and hence is a compact interval, either a point or \([a, b] \).

Now the analysis will begin creeping in. The next theorem is arguably the main theorem; at least 80% of the work will go into it.

**Theorem 35.1 (Main theorem).** Let \( X \) be a noncompact, simply connected Riemann surface and \( \rho \in \Omega^2_c(X) \) be such that \( \int_X \rho = 0 \). Then, there’s a smooth solution \( u \) to \( \Delta u = \rho \) such that \( u \to 0 \) at \( \infty \) in \( X \).

Here’s the reason this theorem is important on our way to uniformization.

**Proposition 35.2.** Theorem 35.1 implies the existence of a meromorphic function \( f \) with one pole that’s simple such that \( \Im f \to 0 \) at \( \infty \) in \( X \).

\(^{57}\)Notice that we haven’t yet used that \( X \) is simply connected.
That is, Theorem 35.1 implies the uniformization theorem, at least after the reductions that we’ve already done. This proof strategy should feel very similar to the proof of Riemann-Roch, with an additional decay at infinity that is part of the noncompact data, and isn’t tracked by sheaf theory, since it’s not local data.

**Proof of Proposition 35.2.** Choose a holomorphic coordinate $z$ near a point $p \in X$, and let $D \subset X$ be the unit disc in this coordinate. Let $\chi : [0, \infty) \to [0, 1]$ be a smooth cutoff function, so $\chi(x) = 1$ for $x < 1/3$ and $\chi(x) = 0$ for $x > 2/3$, and let

$$A = \bar{\partial} \left( \chi(|z|) \cdot \frac{1}{z} \right)$$

on $D$. Since $A$ is only supported in the annulus $1/3 \leq |z| \leq 2/3$, then we can extend it by 0 to all of $X$, and let $\rho = \partial A = dA$; by Stokes’ theorem, this means $\int \rho = 0$.

Apply Theorem 35.1 to $\text{Re} \rho$ and $\text{Im} \rho$ separately to obtain a $C^\infty u : X \to \mathbb{C}$ such that $\partial \bar{\partial} u = \rho$ and $u \to 0$ at $\infty$. The $(0, 1)$-form $A - \bar{\partial} u$ is $\partial$-closed.

There’s a bijection from the space of real $1$-forms to the space of $(0, 1)$-forms sending $b \mapsto b^{0,1}$, and the inverse maps sends $\alpha \mapsto 2 \text{Re}(\alpha)$. Hence, let $a = 2 \text{Re}(A - \bar{\partial} u)$, so $da = 2 \text{Re} d(A - \bar{\partial} u) = 2 \text{Re} \partial (A - \bar{\partial} u) = 0$, meaning $a$ is closed. Since $X$ is simply connected, then $\pi_1(X) = 1$, and therefore $H^1(X) = \text{Hom}(\pi_1(X), \mathbb{R}) = 0$.

Thus, $a$ must be exact: there exists a $\psi : X \to \mathbb{R}$ such that $a = d\psi$, so $\psi$ is a real 0-form.

Now, we can compute that

$$A - \bar{\partial} u = 2 \left( \text{Re}(A - \bar{\partial} u) \right)^{0,1} = a^{0,1} = (d \psi)^{0,1} = \bar{\partial} \psi,$$

or $A = \bar{\partial}(u + \psi)$, and in particular, if

$$f = \begin{cases} u + \psi - \chi(|z|) \frac{1}{z}, & z \in D \\ u + \psi, & z \not\in D, \end{cases}$$

then $f$ is holomorphic away from $p$ and has a simple pole at $p$. Off of $D$, $\text{Im}(f) = \text{Im}(u)$, which goes to 0 at $\infty$.

**Example 35.3.** On $\mathbb{C}$, we can write down a solution to $\Delta u = \delta_{z_0}$ for a $z_0 \in C$. Sure, this isn’t a smooth function, and the $u$ we obtain won’t be $C^\infty$, but far away at $\infty$, there’s no way to tell. There’s a way to formalize this with smooth approximations to $\delta_{z_0}$, but we’re not going to discuss this in depth. Anyways, we obtain the Newton kernel

$$u(z) = \frac{1}{2\pi} \log|z - z_0|.$$ 

This grows at infinity, which is weird! You might think it has something to do with the singularity, but by smooth approximations this isn’t the case. Instead, it has to do with the “fact” (in some sense that can be made rigorous)

$$\int_{\mathbb{C}} \delta_{z_0} d\mu_z = 1.$$ 

This isn’t covered by the theorem. Instead, we want to solve $\Delta u = \delta_{z_0} - \delta_{z_1}$, which does have total integral 0 and gives us something that decays at $\infty$:

$$2\pi u(z) = \log|z - z_0| - \log|z - z_1| = \log \left| \frac{z - z_0}{z - z_1} \right| \sim \frac{C}{|z|},$$

for some constant $C$. So it does decay, albeit not particularly fast.

We’ve already done a lot of analysis in the compact case; let’s refresh what we’ve already obtained. We can use a partition of unity to reduce to the case where $\text{supp}(\rho)$ is contained in some coordinate disc $D \subset X$.

**Proposition 35.4.** If $\rho$ is such that $\text{supp}(\rho) \subset D \subset X$ and $\int_X \rho = 0$, then there exists a smooth solution $u$ to $\Delta u = \rho$ which can be approximated by a sequence $(u_n)$ of functions in $C^\infty(X)$ such that:

1. There’s a sequence $(c_n)$ of real numbers such that $u_n + c_n$ is compactly supported for each $n$.
2. For all $\alpha \in \Omega^1_c(X)$,

$$\int_X |u - u_n|^2 |\alpha|^2 \to 0.$$ 

---

\footnote{Here, $|\alpha| = 2\alpha^{1,0} \wedge \alpha^{0,1}$.}
Remark. That is, will secretly be the last step is to show this is a strong solution.

The weak solution in the Hilbert space completion, and the second statement is that the solution is locally \(L^2\). Suppose there exists a \(u \in \Omega^2_c(X)\) such that \(\int_X \rho = 0\), then we want to show there's a \(u \in C^\infty(X)\) such that \(\Delta u = \rho\) and \(u \to 0\) at \(\infty\) in \(X\).

We'll work with the inner product space \((\Omega^1_c(X), \langle \cdot, \cdot \rangle_D)\), where the inner product is the Dirichlet product

\[
\langle f, g \rangle_D = 2 \int_X \partial f \land \overline{\partial} g
\]

as before. If \(\|f\|_D^2 = 0\), then \(\int |df|^2 = 0\) (where \(|df| = 2i\partial f \land \overline{\partial} f\)); this means \(df = 0\), so \(f\) is constant, and therefore 0 (since it goes to 0 at \(\infty\)). Thus, the Dirichlet product is nondegenerate already; unlike the proof in the compact case, we don’t need to mod out by constants.

Let \(H\) denote the Hilbert space completion of \(\Omega^2_0(X)\); then, \(\Delta u = \rho\) is equivalent to \(\langle u, \chi \rangle_D = \hat{\rho}(\chi)\), for all \(\chi \in \Omega^2_c(X)\), where \(\hat{\rho}(\chi)\) is defined to be \(\int_X \chi \rho\).

As in the compact case, we showed that \(\hat{\rho} : \Omega^1_c(X) \to \mathbb{R}\) is a bounded operator, which used the fact that \(\int_X \rho = 0\). By the Riesz representation theorem, there must be some \(u \in H\) such that \(\langle u, \chi \rangle_D = \hat{\rho}(\chi)\) for all \(\chi\).

That is, \(u\) is represented by a Cauchy sequence \((u_n)\) in \(\Omega^2_0(X)\), so that \(\|u_m - u_n\|_D \to 0\) as \(m, n \to \infty\) and \(\langle u_n, \chi \rangle_D \to \hat{\rho}(\chi)\) for all \(\chi \in \Omega^2_c(X)\).

The next step is to show that the limit \(u\) is actually a function, as in the compact case. The goal is to show that there exist constants \(c_n \in \mathbb{R}\) such that if \(v_n = u_n - c_n\), then there's a \(v : X \to \mathbb{R}\) such that \(v_n \to v\) in \(L^2\) on small coordinate discs. That is, in a fixed coordinate disc \(\mathbb{D} \subset X\) we've estimated that if \(C_n\) is “the average” of \(u_n\) on \(\mathbb{D}\), in the sense that \(v_n = u_n - c_n\) has integral 0 on \(\mathbb{D}\), then

\[
\|v_n - v_m\|_{L^2(\mathbb{D})} \leq C\|v_n - v_m\|_{D(\mathbb{D})} \leq C\|v_n - v_m\|_{D(X)} = C\|u_n - u_m\|_{D(X)} \to 0,
\]

so \((v_n)\) is Cauchy on \(L^2(\mathbb{D})\), meaning there’s a limit \(v_0\) such that \(v_n \to v_0\) in \(L^2(\mathbb{D})\), because \(L^2(\mathbb{D})\) is complete!

The new step will be to show that \(v_n\) converges in \(L^2\) throughout \(X\); let \(A\) be the set of \(x \in X\) where \(v_n\) converges in \(L^2\) on a small coordinate disc centered at \(x\). By construction, \(A\) is open, and is nonempty, since it contains \(\mathbb{D}\). Suppose there exists a \(y \in \overline{A} \setminus A\), so there exist \(c'_n\) such that \(v_n + c'_n\) converges in \(L^2\) on some disc around \(y\), but then \(c'_n\) converges to some \(c\), which implies that \((v_n)\) converges in \(L^2\) in a neighborhood of \(y\). Hence, \(A\) is closed, so \(A = X\). We now have a \(v : X \to \mathbb{R}\) that is approximated by \(v_n = c_n\).
First, we can show that \( v \) is a weak solution to \( \Delta v = \rho \); for any \( \chi \in \Omega^0_c(X) \), since \( \chi \) is compactly supported, then

\[
\langle v, \chi \rangle_D = \int_X v \Delta \chi = \lim_{n \to \infty} \int_X v_n \Delta \chi = \lim_{n \to \infty} \langle v_n, \chi \rangle_D = \lim_{n \to \infty} \langle u_n, \chi \rangle_D = \hat{\rho}(\chi)
\]

by the definition of \( v_n \).

The next step was to show that \( v \) is smooth. This follows from Weyl’s lemma (Lemma 33.1) again, because smoothness is a local condition. That is, we have everything except for the asymptotic behavior of \( v \).

There’s an infinite-dimensional space of harmonic functions on \( X \) (e.g. if \( X = \mathbb{C} \) there are harmonic polynomials of arbitrary degree), and therefore infinitely many solutions to \( \Delta u = \rho \). However, there’s only one solution \( u \) such that \( u \to 0 \) at \( \infty \). Did we bet on the right one? It’s a reasonable guess — it’s a limit of compactly supported functions, which is reassuring, but the Dirichlet norm knows nothing about constants; we may have to add a constant to \( v \) to get it to go to \( 0 \) at \( \infty \), and there’s no guarantee yet that this will work.

**Proposition 36.1.** For all \( \alpha \in \Omega^1_c(X) \),

\[
(36.2) \quad \int_X |v - v_n|^2 |\alpha|^2 \to 0
\]

as \( n \to \infty \), and \( \|v - v_n\|_D \to 0 \) as \( n \to \infty \).

**Proof.** The first part, (36.2), is just a restatement of locally \( L^2 \) convergence. For the second part, one can check that \( \langle v - v_n, v_m \rangle_D = \langle u - u_n, u_m \rangle_D \) and therefore \( \langle v - v_n, v \rangle_D = \langle u - u_n, u \rangle_D \), and therefore we can deduce that \( \langle v - v_n \rangle_D^2 = \langle u - u_n \rangle_D^2 \to 0 \).

This was a technical detail; necessary details.

Our task now is to show that \( v \) goes to a constant at \( \infty \); then, we’ll subtract that constant to obtain a solution that goes to \( 0 \) at \( \infty \). We won’t get to the heart of the argument today; before we do, we should set up some preliminary lemmas.

First, you might recall that we assumed \( X \) to be simply connected, but haven’t used that yet. This shows up in the following lemma from topology.

**Lemma 36.3.** If \( S \) is a noncompact, simply connected smooth surface, then it’s connected at infinity, i.e. for all compact \( K \subset S \), \( S \setminus K \) has a single connected component that’s noncompact.

For example, an annulus separates \( \mathbb{C} \) into two parts, but only one is unbounded. By contrast, on a cylinder \( S^1 \times \mathbb{R} \), removing a subcylinder in the middle produces two noncompact pieces (though of course the cylinder isn’t simply connected).

**Proof.** For the sake of convenience, we’ll rely on a proof of Poincaré duality that we haven’t proven; it simplifies one’s understanding of the proof, but if you don’t like that, there’s a workaround in the textbook. The statement is that \( H_1(S) \otimes \mathbb{R} \cong H^1_c(S) \). Since \( \pi_1(S) \) is trivial, then so is \( H_1(S) = \pi_1(S)^{ab} \), and therefore \( H^1_c(S) = 0 \).

Suppose \( K \subset X \) is compact and \( S \setminus K \) has two noncompact components \( U_0 \) and \( U_1 \). We can enlarge \( K \) if necessary so that \( \overline{U_0} \cap \overline{U_1} = \emptyset \). Thus, there exists a smooth \( f \in C^\infty(S) \) such that \( f|_{U_0} = 0 \) and \( f|_{U_1} = 1 \). Then, \( df \) is a closed, compactly supported 1-form, so defines a class \([df] \in H^1_c(S) = 0\), and therefore is exact. Hence, \( \theta = dg \), where \( g \in C^\infty(S) \), and in particular \( d(f - g) = 0 \). Thus, \( g \) and \( f \) differ by a constant, \( f \) is supported on all of \( U_1 \), and \( g \) is compactly supported; this is a contradiction, because there’s no constant one can add to \( f \) to give it compact support.

There’s one more lemma we can get out of the way today.

**Lemma 36.4.** Let \( X \) be a noncompact topological space and \( \phi : X \to \mathbb{R} \) be continuous. Then, one of the following holds.

- \( \phi \to c \in \mathbb{R} \) at \( \infty \) in \( X \).
• \( \phi \to \infty \) at \( \infty \).
• \( \phi \to -\infty \) at \( \infty \).
• There exist \( a, b \in \mathbb{R} \) such that \( \phi^{-1}((-\infty, a]) \) and \( \phi^{-1}([b, \infty)) \) are both noncompact.

This isn’t an incredibly strong conclusion, but perhaps that’s why this is a lemma and not a theorem. The last possibility is the most confusing; it encodes things such as oscillation at \( \infty \).

**Proof.** Let \( A^- = \{ a \in \mathbb{R} \mid \phi^{-1}((-\infty, a]) \) is compact \}, and \( A^+ = \{ b \in \mathbb{R} \mid \phi^{-1}([b, \infty)) \) is compact \}.

\( \phi \to \infty \) is by definition saying that for all \( a \in \mathbb{R} \), there’s a compact \( K \) such that \( \phi|_{X \setminus K} \geq a \), which is equivalent to \( \phi^{-1}((-\infty, a]) \) being compact for all \( a \in \mathbb{R} \), or that \( A^- = \mathbb{R} \) or \( A^- = \emptyset \). Similarly, \( \phi \to -\infty \) iff \( A^- = \emptyset \).

Thus, we’re left considering the case where both \( A^- \) and \( A^+ \) are nonempty; let \( a^- = \sup A^- \) and \( a^+ = \inf A^+ \); since \( X \) is noncompact, \( a^- \leq a^+ \); if they’re equal, then \( \phi \to a^- \) at \( \infty \), and if they’re not, then we’ve met the last condition.

This is mostly a definition-unwinding argument; it’s pretty elementary.

---

**Lecture 37.**

**Uniformization, IV: 4/29/16**

---

**Lecture 38.**

**Uniformization, V: 5/2/16**

Today, we finish the proof of Theorem 35.1, the theorem that if \( X \) is a noncompact, simply connected Riemann surface and \( \rho \in \Omega^2_c(X) \) is such that \( \int_X \rho = 0 \), then there’s a smooth solution \( u \) such that \( \Delta u = \rho \) and \( u \to \infty \) in \( X \). This last clause is what remains.

So far, we’ve found a smooth solution \( u \) to \( \Delta u = \rho \), and that there exist \( a, b \in \mathbb{R} \) and closed, noncompact, and path-connected sets \( Z_a, Z_b \subset X \) such that \( u_n|_{Z_a} \leq a \) and \( u|_{Z_b} \geq b \), with \( a \leq b \). We also figured out that there are sequences \( (u_n) \) in \( \Omega^0(X) \) and \( (c_n) \) in \( \mathbb{R} \) such that \( u_n + c_n \) is compactly supported for every \( n \) and \( \|u - u_n\|_D \to 0 \) as \( n \to \infty \). We know that \( u \) is locally \( L^2 \) (though we’ll no longer use that) and that \( X \) is connected at \( \infty \).

Thus far, this story is not terribly different from the solution of Poisson’s equation in the compact case, but here something genuinely different happens. We want to show that \( u \) goes to a constant; we’ve showed that it can’t converge to \( \infty \) or \( -\infty \), and have put conditions on whether it can oscillate. What we need to show is that \( a = b \).

Roughly speaking, our strategy will use the scenario with \( Z_a, Z_b \), and \( u_n \) to concoct some loops \( \gamma_n \) in \( X \), which bound discs \( K_n \). Assuming \( a \) is strictly less than \( b \), we can show that \( (u, u_n) \) maps \( K_n \) to a “big” region in \( \mathbb{R}^2 \), meaning with area above a \( \delta > 0 \) independent of \( n \). We’ll then show this contradicts \( \|u - u_n\|_D \to 0 \).

**Lemma 38.1.** Let \( S \) be an oriented smooth surface (which need not have a countable atlas) and \( f : S \to \mathbb{R}^2 \) be \( C^\infty \). Then, for all \( K \subset X \) compact,

\[
\mu(f(K)) \leq \int_S |f^*(dx \wedge dy)|,
\]

where \( \mu \) denotes the Lesbegue measure on \( \mathbb{R}^2 \).

This is a close relative of the change-of-variables formula, which states that (38.2) is an equality when \( f \) is a diffeomorphism. As such, it has a proof along the same lines.

**Proof.** Since \( K \) is compact, we can cover it by finitely many coordinate discs \( D_1, \ldots, D_n \), and replace \( S \) by \( \bigcup_{i=1}^n D_i \), which only decreases the total value of an integral on \( S \) of a nonnegative function. Hence, we may assume \( S \) has a finite atlas.

By Sard’s theorem, \( \mu(f(\text{crit } f)) = 0 \), and therefore \( \mu(f(K)) = \mu(f(K \setminus \text{crit } f)) \), so it’s enough to work with the regular points of \( f \), or equivalently to prove (38.2) when \( f \) has no critical points, \( K \subset S \) is closed (but maybe not compact), and \( S \) has a countable atlas.

By the inverse function theorem, there is a countable, locally finite atlas \( \mathcal{U} \) such that on each \( U \in \mathcal{U} \), \( f \) is a diffeomorphism onto its image. We can use this to “tile” \( S \) into closed subsets with disjoint interiors, and it
suffices to prove (38.2) on each such closed subset: there may be overlaps, but since we’re proving an upper bound that’s acceptable.

Finally, on each closed tile, equality holds in (38.2) by the change-of-variables formula, since \( f \) restricts to a diffeomorphism onto its image there.

Now, fix an \( x_a \in Z_a \) and an \( x_b \in Z_b \), and let \( \Gamma \) be a path from \( x_a \) to \( x_b \). The \( u_n \) we cooked up in our partial proof of the main theorem can be adjusted to be equal to \( u \) on any compact subset of \( X \) (since we know \( u \) is smooth, and care about its behavior at infinity), and in particular we choose them to be equal on \( \text{Im}(\Gamma) \). Parameterize \( S^1 \) by \([0, 1]\), where 0 \( \sim \) 1, and let \( S_n = \text{supp}(u_n + c_n) \), which is a compact subset of \( X \).

Now, we choose loops \( \gamma_n : S^1 \to X \) according to the following constraints:

- On \([0, 1/4]\), \( \gamma_n \) traces out \( \Gamma \) from \( x_a \) to \( x_b \).
- Choose a \( y_b^{(n)} \in Z_b \setminus S_n \), and on \([1/4, 1/2]\) \( \gamma \) traces a path from \( x_b \) to \( y_b^{(n)} \) contained entirely in \( Z_b \).
- Choose a \( y_a^{(n)} \in Z_a \setminus S_n \), and on \([1/2, 3/4]\), \( \gamma \) traces a path from \( y_a^{(n)} \) to \( y_a^{(n)} \) contained entirely in \( X \setminus S_n \).
- On \([3/4, 1]\), \( \gamma \) traces a path from \( y_a^{(n)} \) to \( x_a \) contained entirely in \( Z_a \).

This is all reasonable, except in the third step we need to verify why we can choose a path in \( X \setminus S_n \). Since \( X \) is connected at \( \infty \), then \( \overline{X \setminus S_n} \) has just one noncompact connected component, but \( y_a \) and \( y_b \) necessarily lie in noncompact components of \( \overline{X \setminus S_n} \), because \( S_n \) is closed, and therefore must be in the same path component. Since \( X \) is simply connected, each \( \gamma_n \) can be “filled in,” in that for each \( n \) there’s a compact \( K_n \subset X \) containing \( \gamma_n \) such that \( \gamma_n \) is contractible in \( K_n \).

Let \( F_n = (u, u_n) : X \to \mathbb{R}^2 \) and \( g_n = F_n \circ \gamma_n : S^1 \to \mathbb{R}^2 \), and give \( \mathbb{R}^2 \) the usual \((x, y)\)-coordinates. What does \( g_n \) do?

- On \([0, 1/4]\), \( \gamma_n(t) \in \text{Im}(\Gamma) \), and therefore \( u(\gamma(t)) = u_n(\gamma(t)) \). Thus, \( g_n(t) \) is in the diagonal \( \{(x, x)\} \subset \mathbb{R}^2 \).
- On \([1/4, 1/2]\), \( \gamma_n(t) \in Z_b \), so \( u(\gamma_n(t)) \geq b \), so \( g_n(t) \) is to the right of the line \( x = b \).
- On \([1/2, 3/4]\), we’re outside of \( S_n \), so \( u_n = -c_n \), and therefore \( u_n(\gamma_n(t)) = -c_n \). Thus, \( g_n(t) \) is contained in the line \( y = -c_n \).
- On \([3/4, 1]\), \( \gamma_n(t) \in Z_a \), so \( g_n(t) \) is to the left of the line \( x = a \).

This sharply constrains \( g_n \) to one of only a few possible pictures, depending on the size order of \( a, b, \) and \( -c_n \). In particular, the region enclosed by \( \gamma_n \) necessarily contains the region \( Z_n = \{(x, y) \mid x \in [a, b], \min(x, -c_n) < y < \max(x, -c_n)\} \), and (checking on a case-by-case basis) \( \mu(Z_n) \geq (b - a)^2/4 \).

In each case, \( g_n \) encloses \( Z_n \), meaning that for all \( p \in Z_n \), the winding number of \( g_n \) around \( p \) is \( \pm 1 \): after a homotopy, \( \text{Im}(g_n) = \partial Z_n \), with one or the other orientation, by “pulling it taut” at 0, \( 1/4, 1/2, \) and \( 3/4 \). This homotopy never passes through the interior of \( Z_n \), so it preserves the winding number. In particular, \( Z_n \subset F_n(K_n) \): if not, then there would be a \( p \in Z_n \subset F_n(K_n) \), so we could contract \( g_n \) in \( \mathbb{R}^2 \setminus p \), but we know it has a nonzero winding number, which is a contradiction.

Now, \( F_n \) maps to an area of the plane whose area is at least \( \delta = (b - a)^2/4 \), so we may finish by applying Lemma 38.1 to \( F_n \):

\[
0 < \delta \leq \mu(Z_n) \leq \mu(F_n(K_n)) \leq \int_X |F_n^*(dx \wedge dy)| = \int_X |du \wedge du_n| = \int_X |d(u - u_n) \wedge du_n| \leq \|u - u_n\|_D \cdot \|u_n\|_D.
\]

This is a Cauchy sequence, but it doesn’t go to 0 unless \( b = a \), so we’ve finally finished the proof of uniformization.
Lecture 39.

The Abel-Jacobi Map: 5/4/16

For the last two lectures, we’ll discuss the Abel-Jacobi map, which is an interesting subject tying together a lot of interesting things from this quarter: divisors, Serre duality, and so forth. There will be some things that will necessarily be glossed over, and some language of complex manifolds or Čech cohomology that we haven’t discussed, but it should for the most part form a good conclusion. For this and the next lecture, \( X \) will denote a compact Riemann surface.

The geometry of the divisors \( X \), considered collectively, is encoded in the topology and geometry of higher-dimensional complex manifolds and maps between them.

Definition 39.1. Let \( d \geq 0 \).

- The **symmetric power** \( \text{Sym}^d(X) \) is the space of effective\(^59\) divisors of degree \( d \) on \( X \). Equivalently, it is the quotient \( X^d/S_d \), where the symmetric group \( S_d \) acts on \( X^d \) by permuting the coordinates. Intuitively, this is the space of unordered \( n \)-uples of elements of \( X \).
- The **Picard torus** \( \text{Pic}^d(X) \) is the space of isomorphism classes of invertible sheaves of \( \mathcal{O}_X \)-modules of degree \( d \).
- The **Jacobian torus** is the space \( \text{Jac}(X) \) defined as \( H^0(\mathcal{K}_X)^* \) modulo periods.
- The **Abel-Jacobi maps** are the maps \( \text{AJ}_d : \text{Sym}^d(X) \to \text{Pic}^d(X) \) sending \( D \mapsto [\mathcal{O}_X(D)] \).

\( \text{Pic}^0(X) \) and \( \text{Jac}(X) \) are both complex Lie groups of dimension \( g = g(X) \); it turns out that \( \text{Pic}^0(X) \cong \text{Jac}(X) \). Hence, we can and will reinterpret one Abel-Jacobi map as a map to \( \text{Jac}(X) \). We’ll show that \( \text{AJ}_d : \text{Pic}^d(X) \to \text{Pic}^d(X) \) is surjective and generically one-to-one.

First, let’s talk about Picard tori. Recall that an invertible sheaf of \( \mathcal{O}_X \)-modules \( \mathcal{L} \) is a sheaf such that locally (on an open cover of \( X \)) we have \( \mathcal{L}(U) = \mathcal{O}_X(U) \cdot s_U \), so it’s a free module on \( s_U \). The isomorphism classes of invertible sheaves form an abelian group under \( \otimes \), called the **Picard group**; the inverse of a sheaf is its dual sheaf: \( \mathcal{L}^{-1} = \mathcal{L}^* = \text{Hom}_{\mathcal{O}_X}(\mathcal{L}, \mathcal{O}_X) \).

If \( \mathcal{U} \) is such an open cover of \( X \), meaning that \( \mathcal{L}(U_i) \) is free of rank 1 on \( \mathcal{O}_X(U_i) \) for all \( U_i \in \mathcal{U} \), then we have transition functions \( \tau_{ij} \in \mathcal{O}_X(U_{ij}) \) (where \( U_{ij} = U_i \cap U_j \)) given by \( \tau_{ij} = s_{ij}/s_{ij} \). These obey a cocycle condition \( \tau_{ij} \tau_{jk} \tau_{ki} = 1 \), and therefore define a Čech cohomology class \( \{[\tau_{ij}]\} \in H^1(X, \mathcal{O}_X^*) \).

Proposition 39.2. This class completely determines \( \mathcal{L} \); that is, \( \text{Pic}(X) \cong \check{H}^1(X, \mathcal{O}_X^*) \) as abelian groups.

This is because any cocycle \( \{\tau_{ij}\} \) defines an invertible sheaf trivial over each \( U_i \in \mathcal{U} \) and with transition functions \( \tau_{ij} \), and if this cocycle is the zero class in cohomology, then by definition \( \tau_{ij} = \sigma_i/\sigma_j \), where \( \sigma_j \in \mathcal{O}_X^*(U_i) \). Then, we can tweak our generators \( s_i \) to \( s_i' = s_i/\sigma_i \), which agree on overlaps (the transition functions are trivial) and therefore define a global section of \( \mathcal{L} \), trivializing it. This sets up the isomorphism as sets; then, one checks that tensor products are sent to sums.

The **exponential exact sequence** of sheaves is the sequence

\[
\begin{array}{cccccc}
0 & \longrightarrow & \mathbb{Z}_X & \longrightarrow & \mathcal{O}_X & \xrightarrow{\exp} & \mathcal{O}_X^* & \longrightarrow & 0, \\
& & \text{(39.3)} & & \\
& & \cdots & \longrightarrow & \check{H}^1(X; \mathcal{O}_X) & \longrightarrow & \hat{H}^1(X; \mathcal{O}_X) & \delta & \longrightarrow & \check{H}^2(X; \mathbb{Z}_X). \\
\end{array}
\]

One can check\(^60\) that \( \check{H}^k(X; \mathbb{Z}_X) \) is isomorphic to the integer lattice \( H^k(X; \mathbb{Z}) \) inside the \( k \)-th de Rham cohomology (de Rham classes with integer coefficients), so is just \( \mathbb{Z} \) when \( k = 2 \). Thus, this sequence is actually

\[
\begin{array}{cccccc}
\cdots & \longrightarrow & \mathbb{Z} & \longrightarrow & \check{H}^1(X; \mathcal{O}_X) & \longrightarrow & \text{Pic}(X) & \delta & \longrightarrow & \mathbb{Z}. \\
& & & & \text{(39.3)} & & \\
\end{array}
\]

We’ve obtained a map \( \delta : \text{Pic} X \to \mathbb{Z} \); let \( \text{Pic}^0(X) = \ker(\delta) \).

---

\(^{59}\)Recall that a divisor is effective if all of its coefficients are nonnegative.

\(^{60}\)No pun intended
Fact. $\delta(\mathcal{O}_X(D)) = \deg D$.

This is cool, but we’re not going to prove it. In any case, it implies that $\text{Pic}^d(X) = \delta^{-1}(d)$ is the a coset of $\text{Pic}^0(X)$.

The exact sequence (39.3) induces a group isomorphism $\text{Pic}^0(X) \cong \hat{H}^1(X; \mathcal{O}_X)/\hat{H}^1(X; \mathcal{O}_X)$.

**Proposition 39.4.** There is an isomorphism of Čech and Dolbeault cohomology $\hat{H}^1(X; \mathcal{O}_X) \cong H^{0,1}(X)$.

We’re not going to prove this; it follows from something called the $\dbar$-Poincaré lemma and a strategy of proof similar to de Rham’s theorem. The isomorphism sends an $\alpha \in \Omega^{0,1}(X)$ to the cocycle $\{f_i - f_j\}$, where $\alpha|_{U_i} = \dbar f_i$.

Proposition 39.4 and the identification $\mathbb{H}^1(X; \mathbb{Z}) \cong H^1(X; \mathbb{Z})$ together imply that $\text{Pic}^0(X)$ is isomorphic to $\mathbb{C}^g$ modulo the image of the lattice $H^1(X; \mathbb{Z})$ in $\hat{H}^1(X; \mathcal{O}_X)$.

**Proposition 39.5.** $\hat{H}^1(X; \mathbb{Z})$ is a lattice in the vector space $\hat{H}^1(X; \mathcal{O}_X)$.

That is, we want it to be discrete and of full rank $2g$.

**Proof sketch.** Locally constant integer-valued functions include into $\mathbb{R}_X$ (since they’re real-valued), and then into $\mathbb{C}_X$, and then into $\mathcal{O}_X$, so we have a commutative diagram

$$
\begin{array}{ccccc}
\hat{H}^1(X; \mathbb{Z}) & \longrightarrow & \hat{H}^1(X; \mathbb{R}) & \longrightarrow & \hat{H}^1(X; \mathbb{C}) \\
\uparrow & & \uparrow & & \uparrow \\
H^1(X; \mathbb{Z}) & \longrightarrow & H^1(X; \mathbb{R}) & \rightarrow & H^1(X; \mathbb{C}) \cong \mathbb{C}^{\gamma} \\
\uparrow & & \uparrow & & \uparrow \\
& & H^1(X; \mathbb{R}) \otimes \mathbb{C} \cong H^{0,1}(X) \\
\end{array}
$$

We emphasize that the rows are not exact.

The vertical arrows are all isomorphisms, the rightmost one by Proposition 39.4, the middle two by de Rham’s theorem, and the leftmost one as we already discussed. We know $H^1(X; \mathbb{Z})$ is a lattice in $H^1(X; \mathbb{R}) \cong \mathbb{R}^{2g}$, and the composite $H^1(X) \rightarrow H^1(X; \mathbb{R}) \otimes \mathbb{C} \rightarrow H^{0,1}(X)$ is an $\mathbb{R}$-linear isomorphism, which follows from the decomposition $H^1(X; \mathbb{R}) \otimes \mathbb{C} = H^0(X; \mathcal{O}_X) \oplus H^{0,1}(X)$, which we established from Serre duality, so $H^1(X; \mathbb{Z})$ is a lattice inside $H^{0,1}(X) \cong \hat{H}^1(X; \mathcal{O}_X)$.

Consequently, we know $\text{Pic}^0(X)$ is a complex, $g$-dimensional vector space modulo a lattice, so it is a $g$-dimensional complex manifold and a complex Lie group: multiplication and inversion in its group structure are holomorphic functions. Moreover, it’s $C^\infty$ diffeomorphic to $H^1(X; \mathbb{R}) \cong (S^1)^{2g}$. Therefore $\text{Pic}^d(X)$ is also a complex manifold. A better way to phrase this is that $X$ is a complex Lie group, with components $\text{Pic}^d(X)$ for each $d$.

**The Abel-Jacobi map.** Let’s return to the symmetric power $\text{Sym}^d(X) = X^d/S_d$. *A priori* a topological space, for Riemann surfaces specifically it has a $d$-dimensional complex manifold structure.\(^{\text{61}}\) We will return to this in the future, but assume it for now.

The Abel-Jacobi map $\text{AJ}_d : \text{Sym}^d(X) \rightarrow \text{Pic}^d(X)$ was defined to send $D \mapsto [\mathcal{O}_X(D)]$; right now we can see only that it’s a map of spaces. Its fibers are projective spaces: $\text{AJ}_d^{-1}(\mathcal{L})$ is the space of divisors $D$ such that $\mathcal{O}_X(D) \cong \mathcal{L}$, or the divisors of nonzero holomorphic sections $s \in H^0(\mathcal{L})$; that is, $\text{AJ}_d^{-1}(\mathcal{L}) = \mathbb{P}H^0(\mathcal{L})$, and is empty if $H^0(\mathcal{L}) = 0$.

Next time, we’ll show that $\text{Sym}^d$ is a $d$-dimensional complex manifold and $\text{AJ}_d$ is holomorphic. We’ll compute its derivative (which we’ve already seen in another guise), and show that $\text{AJ}_d$ is surjective, along with a few consequences, which also uses Serre duality.

---

\(^{\text{61}}\)There is no analogue of this for, e.g. complex surfaces.
The basic version of the Abel-Jacobi map for a compact Riemann surface $X$ is $\text{AJ}^1 : X \to \text{Pic}^1(X)$ sending $p \mapsto \mathcal{O}_X(p)$, where $\text{Pic}^1(X)$ is the degree-1 invertible sheaves. We showed last time that $\text{Pic} X$ has the structure of a complex Lie group, with a $\mathbb{Z}$ worth of connected components. Through the exponential sequence, we have $\text{Pic}^0(X) \cong \check{H}^1(X; \mathcal{O}_X)/\check{H}^1(X; \mathbb{Z}_X)$, which is the quotient of $\mathbb{C}^g$ by a lattice. This gives us a complex structure, but it’s not the best way to realize the complex structure on $\text{Pic} X$.

**Fact.** There’s a nicer way to characterize the complex structure on $\text{Pic} X$: for a complex manifold $M$, a smooth map $\phi : M \to \text{Pic} X$ is holomorphic iff $\phi$ arises from an invertible sheaf $\mathcal{L}$ on $M \times X$ such that for all $p \in X$, $\mathcal{L}|_{(p) \times X}$ represents $\phi(p)$. It’s not too hard to check this agrees with the complex structure we defined first, but it represents a moduli-theoretic point of view, in which $\text{Pic} X$ is the moduli of invertible line bundles on $X$.

**Lemma 40.1.** $\text{AJ}^1$ is a holomorphic map, and its derivative $D_p\text{AJ}^1$ makes the following diagram commute.

$$
\begin{array}{ccc}
T_p M & \longrightarrow & \check{H}^1(X, \mathcal{O}_X) \\
\downarrow a & & \downarrow 2\pi i \delta \\
H^0(\mathcal{O}_X(p)/\mathcal{O}_X) & & \\
\end{array}
$$

where $a : \frac{\partial}{\partial z} \mapsto 1/z$, where $z$ is a local coordinate near $p$, and $\delta$ is the connecting morphism in the cohomology long exact sequence coming from the short exact sequence of sheaves

$$0 \longrightarrow \mathcal{O}_X \longrightarrow \mathcal{O}_X(p) \longrightarrow \mathcal{O}_X(p)/\mathcal{O}_X \longrightarrow 0.$$

**Proof idea.** The relevant invertible sheaf on $X \times X$ is $\mathcal{O}_{X \times X}(\Delta)$, where $\Delta$ is the diagonal submanifold of $X \times X$, because this restricts along $\{(p) \times X \to \mathcal{O}_X(p)$. One can make this explicit by writing down transition functions for $\mathcal{O}_X(p')$ for $p'$ near $p$ and then computing the derivative.  

This is the major black box of this lecture.

The next player in this story is the Jacobian $\text{Jac}(X) = \check{H}^0(\mathcal{X}_X)^*/\Lambda$, where $\Lambda$ is the period group, the image of the map $H_1(X) \to \check{H}^0(\mathcal{X}_X)^*$ sending

$$[\gamma] \mapsto \left( \omega \mapsto \int_\gamma \omega \right)$$

for holomorphic 1-forms $\omega$.

The Dolbeault isomorphism and Serre duality together provide an isomorphism $\sigma : \check{H}^1(\mathcal{O}_X) \to \check{H}^0(X) \to \check{H}^0(\mathcal{X}_X)^*$ such that $\sigma(\check{H}^1(X; \mathbb{Z}_X)) = \Lambda$, so there is an isomorphism of abelian groups $\sigma : \text{Pic}^0 X \to \text{Jac} X$.

Fix a basepoint $q \in X$; we’ll define a map $u : X \to \text{Jac} X$ which will secretly be the Abel-Jacobi map again. This sends a point $p$ to the function $\omega \mapsto \int_q^p \omega$; these integrals are not independent of path, but the differences are periods, hence lie in $\Lambda$. Thus, this is well-defined as a map into $\text{Jac} X$.

**Lemma 40.2.** $u$ is a holomorphic map, and $D_p u : T_p X \to \check{H}^0(\mathcal{X}_X)^*$ is the dual evaluation map $\text{ev}^*$, where $\text{ev} : \check{H}^0(\mathcal{X}_X) \to T^*_p X$. That is, $(D_p u)(v)$ is the function $\omega \mapsto \omega(v)$.

Compared to Lemma 40.1, this is much more hands-on, even if it’ll be the same thing in the end.

**Proof.** Let $\Omega \subset \mathbb{C}$ be a convex open set containing the origin and $\phi : \Omega \to X$ send $0 \mapsto 0$. Then,

$$u(\phi(z))(\omega) = u(p)(\omega) \int_0^z \omega,$$

which is clearly holomorphic, and which does what we said it did.  

$u$ depends on the choice of basepoint $q$, but merely by a translation along a standard path between the two basepoints.

Now, we’d like $\text{AJ}^1$ to be “the same as” $u$. More precisely, given a $q \in X$, let $T_{-q} : \text{Pic}^1(X) \to \text{Pic}^0(X)$ send $\mathcal{L} \mapsto \mathcal{L} \otimes \mathcal{O}_X(-q)$, which we’ll think of as a translation (after all, it’s biholomorphic).
Theorem 40.3. The following diagram commutes.

\[
\begin{array}{ccc}
X & \xrightarrow{\text{AJ}^1} & \text{Pic}^1(X) \\
& \searrow & \Downarrow_{T_{-q} \cong} \\
& & \text{Pic}^0(X) \\
& \swarrow & \\
& & \text{Jac} X
\end{array}
\]

\(u\) and \(T_{-q}\) are the only things in this diagram that depend on basepoint.

Proof. Let \(\widetilde{\text{AJ}} = \sigma \circ T_{-q} \circ \text{AJ}^1 : X \to \text{Jac} X\). Since \(\widetilde{\text{AJ}}(q) = u(q) = 0\), then it suffices to show that they have the same derivative everywhere (equivalently, since the codomain is a complex Lie group, we can take the difference of these functions and show it vanishes). By Lemma 40.1, the derivative of \(T_{-q} \circ \text{AJ}^1\) is \(2\pi i\delta\), since \(T_{-q}X = H^0(\Theta_X(p)/\Theta_X)\), and by Lemma 40.2, the derivative of \(u\) is \(ev^*\). But when we proved Serre duality for \(\Theta_X(D)\) is equivalent to that for \(\Theta_X(D + p)\), we saw that \(2\pi i\delta\) is “Serre dual” to \(ev^*\).

This illustrates how close Abel and Jacobi’s theory is to the theory of divisors and Serre duality.

Now, we can pass to symmetric products, defining \(\text{AJ}^d : \text{Sym}^d X \to \text{Pic}^d X\) to send \(D = \sum p_i\) to \([\Theta_X(D)] = \sum \text{AJ}^1(p_i)\). We expect a lot of things to be very similar, since this is built out of \(\text{AJ}^1\).

First, \(u\) extends to the abelian sums map \(u_d : \text{Sym}^d X \to \text{Jac} X\) sending

\[D = \sum p_i \mapsto \left( \omega \mapsto \sum_i \int_{p_i}^{q} \omega \right),\]

or sending \(D\) to the function \(\omega \mapsto u(p_1) + u(p_2) + \cdots + u(p_m)\).

The diagram in Theorem 40.3 is preserved by the addition of divisors (in a sense, it’s a monoid homomorphism).

Theorem 40.4. The following diagram commutes.

\[
\begin{array}{ccc}
\text{Sym}^d X & \xrightarrow{\text{AJ}^d} & \text{Pic}^d(X) \\
& \searrow & \Downarrow_{T_{-q} \cong} \\
& & \text{Pic}^0(X) \\
& \swarrow & \\
& & \text{Jac} X
\end{array}
\]

Everything is in place, except we only know \(\text{Sym}^d X\) as a topological space, rather than a complex manifold. This is standard material and so we will only sketch it, but the starting point is a homeomorphism \(\mathbb{C}^d \to \text{Sym}^d \mathbb{C}\) sending \((a_1, \ldots, a_d)\) to the set of roots of \(z^d + \sum a_i z^{d-i}\) with multiplicity, and the inverse map uses the elementary symmetric functions as the coordinates on \(\mathbb{C}^d\). More generally, let \(\pi : X^d \to \text{Sym}^d X\) be projection. Then, we define the complex structure on \(\text{Sym}^d X\) by saying that if \(M\) is a complex manifold, a map \(F : \text{Sym}^d X \to M\) is holomorphic iff the composite \(F \circ \pi\).

One curious result is that there’s no canonical smooth structure on the \(d^{th}\) symmetric power on a surface; to get one, pick a complex structure, take this complex structure on \(\text{Sym}^d X\), and then forget the complex structure!

In any case, the definitions of \(\text{AJ}^d\) and \(u_d\) are both holomorphic in this complex structure.

 Abel’s theorem says we can learn about the abelian sums map using the Abel-Jacobi map.

Theorem 40.5 (Abel). The fibers of \(u\) are projective spaces: \(u^{-1}(u(D)) = |D|\) is the set of divisors linearly equivalent to \(D\), i.e. \(\mathbb{P}(H^0(\Theta_X(D)))\).

The proof is that this is clear for \(\text{AJ}^d\).

Corollary 40.6 (Jacobi inversion). There’s an open dense set \(U \subset \text{Pic}^0 X\) consisting of points in the image of \(\text{AJ}^0\) which are regular values for \(\text{AJ}^0\). Hence, \(\text{AJ}^d\) is surjective and \(\deg(\text{AJ}^0) = 1\) (the zero-dimensional fibers are zero-dimensional projective spaces).

Proof. We’ll prove it using \(u_g : \text{Sym}^g X \to \text{Jac} X\). Let \(\omega_1, \ldots, \omega_g\) be a basis for \(H^0(K_X)\) and \(p_1 \in X\) be a point where \(\omega_1(p_1) \neq 0\); the set of possible choices is a set \(U_1\), which is a cofinite open subset of \(X\). Then, we can adjust \(\omega_2, \ldots, \omega_g\) to vanish at our chosen \(p_1\), and repeat the process with \(p_2\), and so on.
Thus, the possible choices for \((p_1, \ldots, p_g)\) fit inside \(U_1 \times \cdots \times U_g\), and each \(U_i\) is cofinite in \(X\). We specified \(\omega_i(p_i) \neq 0\) and \(\omega_i(p_j) = 0\) if \(i < j\). Since we can choose these \(p_i\) to be distinct, this gives us an isomorphism \(D\pi : T_{(p_1, \ldots, p_g)}X^d \to T_{(p_1, \ldots, p_g)} \operatorname{Sym}^d X\) is an isomorphism.

Thus, \(D\pi \circ Du_g\) is a lower triangular matrix with nonzero diagonal entries, and hence it too is nonsingular. The desired open dense set is hence the one where \(u_g\) is a local diffeomorphism. 

The point is, we get some cool corollaries.

**Corollary 40.7.**
- Every degree-\(g\) invertible sheaf has a holomorphic section \(\mathcal{L} \cong \mathcal{O}_X(D)\), because \(\operatorname{AJ}^g\) is surjective.
- Every invertible sheaf has a meromorphic section.

**Corollary 40.8.** For a “generic” effective divisor \(D\), we have

\[
h^0(D) = \begin{cases} 
1, & \deg D \leq g \\
\deg D + 1 - g, & \deg D > g.
\end{cases}
\]

In other words, the Riemann-Roch bound is very often an equality! The genericity comes from the open dense subset that appeared in Jacobi inversion.

This opens the door to a whole lot more about the geometry about generic and special divisors.