# Complex Analysis on Riemann Surfaces

Math 213b — Harvard University
C. McMullen

## Contents



## 1 Introduction

**Scope of relations to other fields.**

1. Topology: genus, manifolds. Algebraic topology, intersection form on $H^1(X,\mathbb{Z})$, $\int \alpha \wedge \beta$.

2. 3-manifolds. (a) Knot theory of singularities. (b) Isometries of $\mathbb{H}^3$ and $\operatorname{Aut}\widehat{\mathbb{C}}$. (c) Deformations of $M^3$ and $\partial M^3$.

3. 4-manifolds. $(M,\omega)$ studied by introducing $J$ and then pseudo-holomorphic curves.

4. Differential geometry: every Riemann surface carries a conformal metric of constant curvature. Einstein metrics, uniformization in higher dimensions. String theory.

5. Complex geometry: Sheaf theory; several complex variables; Hodge theory.

6. Algebraic geometry: compact Riemann surfaces are the same as algebraic curves. Intrinsic point of view: $x^2+y^2=1$, $x=1$, $y^2=x^2(x+1)$ are all 'the same' curve. Moduli of curves. $\pi_1(\mathcal{M}_g)$ is the mapping class group.

7. Arithmetic geometry: Genus $g \geq 2$ implies $X(\mathbb{Q})$ is finite. Other extreme: solutions of polynomials; $\mathbb{C}$ is an algebraically closed field.

8. Lie groups and homogeneous spaces. We can write $X = \mathbb{H}/\Gamma$, and $\mathcal{M}_1 = \mathbb{H}/\operatorname{SL}_2(\mathbb{Z})$. The Jacobian $\operatorname{Jac}(X) \cong \mathbb{C}^g/\Lambda$ determines an element of $\mathfrak{H}_g/\operatorname{Sp}_{2g}(\mathbb{Z})$: arithmetic quotients of bounded domains.

9. Number theory. Automorphic forms and theta functions. Let $q = \exp(\pi i z)$ on $\mathbb{H}$. Then $f(q) = \sum q^{n^2}$ is an automorphic form and $f(q)^k = \sum a_n(k) q^n$ where $a_n(k)$ is the number of ways to represent $n$ as a sum of $k$ ordered squares.

10. Dynamics. Unimodal maps exceedingly rich, can be studied by complexification: Mandelbrot set, Feigenbaum constant, etc. Billiards can be studied via Riemann surfaces. The geodesic flow on a hyperbolic surface is an excellent concrete example of a chaotic (ergodic, mixing) dynamical system.

**Definition and examples of Riemann surfaces.** A *Riemann surface* is a connected, Hausdorff topological space $X$ equipped with an open covering $U_i$ and a collection of homeomorphisms $f_i : U_i \to \mathbb{C}$ such that there exist analytic maps $g_{ij}$ satisfying

$$f_i = g_{ij} \circ f_j$$

on $U_{ij} = U_i \cap U_j$.

We put the definition into this form because it suggest that $(g_{ij})$ is a sort of 1-coboundary of the 0-chain $(f_i)$. This equation will recur in the definition of sheaf cohomology.

**The Riemann sphere.** The simplest compact Riemann surface is $\widehat{\mathbb{C}} = \mathbb{C} \cup \infty$ with charts $U_1 = \mathbb{C}$ and $U_2 = \widehat{\mathbb{C}} - \{0\}$ with $f_1(z) = z$ and $f_2(z) = 1/z$.

Alternatively, $\widehat{\mathbb{C}} \cong \mathbb{P}^1$ is the space of lines in $\mathbb{C}^2$:

$$\mathbb{P}^1 = (\mathbb{C}^2 - 0)/\mathbb{C}^*.$$

The isomorphism is given by $[Z_0 : Z_1] \mapsto z = Z_1/Z_0$.

Note that we have many natural (inverse) charts $\mathbb{C} \to \mathbb{P}^1$ given by $f(t) = [ct+d, at+b]$. The image omits $a/c$. Transitions between these charts are given by Möbius transformations.

In fact the full automorphism group of $\widehat{\mathbb{C}}$ is the quotient of $G = \mathrm{SL}_2(\mathbb{C})$ by $\pm I$. This means every automorphism lifts to a linear map on $\mathbb{C}^2$ of determinant 1, well-defined up to sign. One can also regard the full automorphism group as $\mathrm{PGL}_2(\mathbb{C}) = \mathrm{GL}_2(\mathbb{C})/\mathbb{C}^*$.

Other types of geometry on the Riemann sphere and its homogeneous subdomains correspond to subgroups $H \subset G$ as below.

**What is a circle?** Recall the classification of Hermitian forms, i.e. nondegenerate bilinear forms on $\mathbb{C}^n$ satisfying $z \cdot w = \overline{w \cdot z}$. Any such form is equivalent, by an element of $\mathrm{GL}_n(\mathbb{C})$, to the standard form of signature $(p, q)$, given by

$$z \cdot z = |z_1|^2 + \cdots + |z_p|^2 - |z_{p+1}|^2 - \cdots - |z_{p+q}|^2$$

In the case of $\mathbb{C}^2$, a form of signature $(1, 1)$ (an *indefinite* form) has a *light cone* of vectors with $z \cdot z = 0$. The lines in this light cone give a circle on $\mathbb{P}^1$, and vice-versa.

The *positive vectors* in $\mathbb{C}^2$ pick out a distinguished component of the complement of a circle. Consequently the *space of oriented circles* on $\widehat{\mathbb{C}}$ is isomorphic to $C = \mathrm{SL}_2(\mathbb{C})/\mathrm{SU}(1, 1)$.

The space $C' = C/(\mathbb{Z}/2)$ of unoriented circles can be thought of as the complement of a ball $B \subset \mathbb{RP}^3$, with the boundary of the ball identified with $\widehat{\mathbb{C}} \cong S^2$, and the horizon as seen from $p \notin B$ the circle attached to $p$. We have $\pi_1(C') \cong \mathbb{Z}/2$, and its double cover is the space of oriented circles. Thus $\mathrm{SL}_2(\mathbb{C})/\mathrm{SU}(1, 1)$ is simply–connected.

**The round sphere.** The round metric on $\widehat{\mathbb{C}}$ is given by $2|dz|/(1 + |z|^2)$ and preserved exactly by

$$\mathrm{SU}(2) = \left\{ \begin{pmatrix} a & b \\ -\bar{b} & \overline{a} \end{pmatrix} : |a|^2 + |b|^2 = 1 \right\} \cong S^3.$$

Alternatively, if we represent points $z_i$ in $\mathbb{P}^1$ by unit vectors $v_i$ in $\mathbb{C}^2$, then the spherical distance satisfies:

$$\cos^2(d(z_1, z_2)/2) = |\langle v_1, v_2 \rangle|^2.$$

For example, $d(1, e^{i\theta}) = \theta$, and if we set $v_1 = (1, 1)/\sqrt{2}$, $v_2 = (1, e^{i\theta})/\sqrt{2}$, then

$$|\langle v_1, v_2 \rangle|^2 = |1 + e^{i\theta}|^2/4 = (1 + \cos\theta)/2 = \cos^2(\theta/2).$$

The factor of 1/2 is not unexpected: for real projective space, the map

$$S^1 \subset \mathbb{R}^2 - \{0\} \to \mathbb{RP}^1 = S^1/(\pm 1)$$

doubles angles as well.

Alternatively, we can compute that in the round metric

$$d(0,r) = \int_0^r \frac{2\,ds}{1+s^2} = 2\tan^{-1}(r),$$

so $r = \tan(d/2)$. (This is the familiar transformation from calculus that turns any rational function of $\sin(d)$ and $\cos(d)$ into the integral of an ordinary rational function.) Then $v_1 = (1,0)$, $v_2 = (1,r)/\sqrt{1+r^2}$, and hence

$$\langle v_1, v_2\rangle^2 = \frac{1}{1+r^2} = \frac{1}{1+\tan^2(d/2)} = \cos^2(d/2),$$

using the identity $1 + \tan^2\theta = \sec^2\theta$.

**The Euclidean plane.** The other two simply-connected Riemann surfaces are $\mathbb{C}$ and $\mathbb{H}$, with automorphism groups $B = AN$ and $\mathrm{SL}_2(\mathbb{R})$. These are simply their stabilizers in $\mathrm{SL}_2(\mathbb{C})$.

Note that $B$ is the same as the stabilizer of $\infty \in \widehat{\mathbb{C}}$, and hence $\widehat{\mathbb{C}} \cong G/B$; $B$ is called a *Borel subgroup* because the associated homogeneous space is a projective variety.

**The hyperbolic plane.** Note that $\mathrm{SL}_2(\mathbb{R})$ preserves not only the compactified real axis $\widehat{\mathbb{R}}$, but also its *orientation*. (Note that $z \mapsto 1/z$ is *not* in $\mathrm{SL}_2(\mathbb{R}.)$

Miraculously, the full automorphism of $\mathbb{H}$ preserves the hyperbolic metric $|dz|/y$. On the unit disk $\Delta \cong \mathbb{H}$ the automorphism group becomes

$$\mathrm{SU}(1,1) = \left\{ \begin{pmatrix} a & b \\ \overline{b} & \overline{a} \end{pmatrix} \,:\, |a|^2 - |b|^2 = 1 \right\},$$

and the metric becomes $2|dz|/(1-|z|^2)$.

As a 3-manifold, $\mathrm{SU}(1,1) \cong \mathrm{SL}_2(\mathbb{R})$ is homeomorphic to a solid torus, since it double covers the unit tangent bundle of $\Delta$. Thus $\mathrm{SL}_2(\mathbb{R})$ retracts onto $\mathrm{SU}(2)$ (as is well-known) but not onto $SU(1,1)$, since they have different homotopy types. The reason is that the Gram–Schmidt process fails for indefinite metrics, because a null vector cannot be normalized.

**Hyperbolic distance in $\boldsymbol{\Delta}$.** The unit disk $\Delta$ can be interpreted as the space of *positive lines* in $\mathbb{C}^2$ for the Hermitian form $\|Z\|^2 = |Z_0|^2 - |Z_1|^2$.

As in the case of the sphere, we can use the underlying inner product to compute distances in the hyperbolic metric on the disk. We find:

$$\cosh^2(d(z_1, z_2)/2) = |\langle v_1, v_2\rangle|^2.$$

Here it is useful to recall that $t \mapsto (\cosh t, \sinh t)$ sweeps out (one sheet of) the unit 'circle' in $\mathbb{R}^{1,1}$ defined by $x^2 - y^2 = 1$, just as $(\cos t, \sin t)$ sweeps out the unit circle in $\mathbb{R}^2$.

For example, $d(0, r) = \int_0^r 2\,dx/(1 - x^2) = 2\tanh^{-1}(r)$, and hence

$$\cosh^2(d(0, r)/2) = (1 - r^2)^{-1},$$

which agrees with the inner product squared between the unit vectors $v_0 = (1, 0)$ and $v_r = (1, r)/\sqrt{1 - r^2}$.

**Aside: $\mathbb{H}$ as the moduli space of triangles.** We can think of $\mathbb{H}$ as the space of *marked* triangles (with ordered vertices) up to similarity (allowing reversal of orientation). The triangle attached to $\tau \in \mathbb{H}$ is the one with vertices $(0, 1, \tau)$.

The 'modular group' $S_3$ acts by reflections, in the line $x = 1/2$ and the circles of radius 1 centered at 0 and 1. It is generated by $\tau \mapsto 1/\overline{\tau}$ and $\tau \mapsto 1 - \overline{\tau}$. Note that the fixed loci of these maps are geodesics, and that the special point where they come together corresponds to an equilateral triangle, and that each line corresponds to isosceles triangles.

One can also consider barycentric subdivision. One of the 6 subdivided triangles is given by replacing $\tau$ with the barycenter, i.e. $\beta(\tau) = (1 + \tau)/3$.

As an exercise, one can show that $S_3$ and $\beta$ generate a dense subgroup of $\operatorname{Isom}\mathbb{H}$. Thus any triangle can be approximated by one in $T_n$, the $n$th barycentric subdivision of an equilateral triangle. A more subtle exercise is to show that almost all of the triangles in $T_n$ are long and thin.

**Aside: the Schwarz lemma and dynamics.** A fundamental property of the hyperbolic metric: if $f : \mathbb{H} \to \mathbb{H}$ is holomorphic, then either $f$ is an isometry, or $\|f'(z)\| < 1$ for every $z \in \mathbb{H}$, where the norm is measured in the hyperbolic metric.

Sample use: if $f(z) = z^2 + c$ and $p$ is an attracting periodic point for $f$, then $p$ attracts $z = 0$. (The immediate basin $B$ of $p$ is a bounded disk, by the maximum principle; and this disk must contain a critical point of $f^n$, where $n$ is the period of $p$.)

Sample use: let $p(z)$ be a polynomial. If Newton's method works for the point of inflection of $p$, then it works for almost every point on the sphere.

Idea of the proof: the critical points of Newton's method $N(z) = z - p(z)/p'(z)$ come exactly from the zeros of $p(z)p''(z)$.

**Description of all Riemann surfaces by covering spaces.** Once all the universal covering spaces have been identified, we can easily see:

**Theorem 1.1** *Every Riemann surface $X$ is isomorphic to*

$$\widehat{\mathbb{C}}, \mathbb{C}, \mathbb{C}^*, \mathbb{C}/\Lambda \quad \text{or} \quad \mathbb{H}/\Gamma,$$

*for some lattice $\Lambda \subset \mathbb{C}$ or torsion–free discrete subgroup $\Gamma \subset \mathrm{SL}_2(\mathbb{R})$.*

The torsion–free condition insures that $\Gamma$ acts *freely* on $\mathbb{H}$.

**Warning.** It is not always true that the quotient of a planar domain by a group acting freely with discrete orbits is a Hausdorff surface! (Consider the acting $(x, y) \mapsto (2x, y/2)$ acting on $\mathbb{R}^2 - \{0, 0\}$.) However: if $\Gamma$ is a discrete group of isometries acting freely on a connected Riemannian manifold $Y$, then the action is properly discontinuous and $Y$ covers the quotient manifold $Y/\Gamma$.

**Doubly-connected Riemann surfaces.** Applying the result above we can easily show that the Riemann surfaces with $\pi_1(X) \cong \mathbb{Z}$ are given by $\mathbb{C}^*$, $\Delta^*$ and

$$A(r) = \{z \,:\, r < |z| < 1\},$$

up to isomorphism.

In terms of covering spaces, we have $\mathbb{C}^* \cong \mathbb{C}/\mathbb{Z}$, $\Delta^* \cong \mathbb{H}/\mathbb{Z}$, and

$$A(r) \cong \mathbb{H}/\lambda^{\mathbb{Z}},$$

where $\lambda > 1$ and $r = \exp(-2\pi^2/\log\lambda)$.

The covering map $f : \mathbb{H} \to A(r)$ can be factored as $z \mapsto \log(z)$, $z \mapsto \alpha z$ and then $z \mapsto e^z$, where $\alpha = 2\pi i/\log\lambda$. In other words, $f(z) = z^\alpha$ with $\alpha$ purely imaginary.

The fact that $f(z) = z^\alpha = \exp(i\beta\log z)$ can be guessed from the fact that $f$ should send $\mathbb{R}_+$ to the unit circle. For this map to have the right deck group, we need

$$f(\lambda) = 1 = \exp(i\beta\log\lambda) = \exp(2\pi i),$$

which gives $\beta = 2\pi/L$, where $L = \log\lambda$ is the length of the closed hyperbolic geodesic on $A(r)$. Then $r = f(-1) = \exp(i\beta\pi i) = \exp(-2\pi^2/L)$. Thus $r \to 0$ as $L \to \infty$.

**Triply–connected Riemann surfaces, etc.** Any finitely–connected planar surface is equivalent to $\widehat{\mathbb{C}}$ with a finite number of points and round disks removed. In particular, such a surface depends on only finitely many *moduli*.

The double of a surface with $g+1$ circular boundaries is a closed Riemann surface of genus $g$, with a *real symmetry*. In fact the space of such planar regions is isomorphic to the space of Riemann surfaces defined over $\mathbb{R}$ with the maximal number of real components.

Another canonical form for a planar region, say containing a neighborhood of infinity, is the complement of finitely many disjoint horizontal segments. This canonical conformal representation can be found by maximizing $\operatorname{Re} b_1$, where $f : U \to \mathbb{C}$ is given by $f(z) = z + b_1/z + O(1/z^2)$. See e.g. [Gol].

**The triply–punctured sphere.** The group $\Gamma(2) \subset \mathrm{SL}_2(\mathbb{Z})$; isometric for hyperbolic metric; quotient is triply-punctured sphere. The map

$$\lambda : \mathbb{H} \to \widehat{\mathbb{C}} - \{0, 1, \infty\}$$

can be interpreted as the elliptic modular function. That is, given $\tau \in \mathbb{H}$ we set $\Lambda = \mathbb{Z} \oplus \mathbb{Z}\tau$ and form $E = \mathbb{C}/\Lambda$; then $E$ admits a degree two map to $\widehat{\mathbb{C}}$ branched over $0, 1, \infty$ and $\lambda(\tau)$. The ordering of the 4 branch points requires a marking of $E[2]$, which explains that passage to $\Gamma(2)$.

**More constructions of Riemann surfaces.**

1. Any connected, oriented Riemannian 2-manifold $(X, g)$ has a unique compatible complex structure, such that $g$ is a conformal metric. (This result requires isothermal coordinates or harmonic functions.)

2. Algebraic curves. A curve $C \subset \mathbb{C}^2$ defined by $f(x, y) = 0$ is *smooth* $df \neq 0$ along $C$. Example: $f(x, y) = y^2 - p(x)$ is smooth iff $p$ has no multiple roots. Non-example: $y^2 = x^3$, $y^2 = x^2(x + 1)$. Double zeros give nodes, triple zeros give cusps.

3. Polygons in $\mathbb{C}$, $\widehat{\mathbb{C}}$, $\mathbb{H}$, glued together by isometries, naturally form Riemannian surfaces. Polyhedra in $\mathbb{R}^3$, such as the octahedron and the icosahedron. Triangulated surfaces. (Relation to billiards.)

4. If $X$ is a Riemann surface and $\pi : Y \to X$ is a covering map, with $Y$ connected, then there is a unique complex structure on $Y$ such that $\pi$ is holomorphic and $Y$ is a Riemann surface in its own right.

5. Finite branched coverings. Let $X$ be a compact Riemann surface and $E \subset X$ a finite set, and let $X^* = X - E$ and $\rho : \pi_1(X^*) \to G$ is a map to a finite group. Then the corresponding $G$-covering space $Y^* \to X^*$ completes to a give a holomorphic map $\pi : Y \to X$, branched over $E$.

6. Finite quotients. If $G \subset \text{Aut}(Y)$ is a finite group of automorphisms, then $X = Y/G$ is a topological surface, and $X$ carries a unique complex structure such that $\pi : Y \to X = Y/G$ is holomorphic.

   The Weierstrass $\wp$-function gives the map $E \mapsto E/G \cong \widehat{\mathbb{C}}$ where $G = \mathbb{Z}/2$ acts by $z \mapsto -z$ in the group law on an elliptic curve $E$.

**The octagon, elaborated.** To illustrate connections between these constructions more fully, we give 3 different descriptions of the same Riemann surface.

1. Let $Z$ be the regular octahedron, and $B \subset Z$ its 6 vertices. Then $Z$ has the structure of a Riemann surface, and we can take the canonical 2-fold covering $X \to Z$ branched over $B$. The result is a surface of genus two that can be assembled out of 16 equilateral triangles, meeting 8 to a vertex.

2. Let $X$ be the curve defined by $y^2 = x(x^4 - 1)$. This Riemann surface can be regarded as the degree 2 cover of $\mathbb{P}^1$ branched over the 6 points $(0, \infty, \sqrt[4]{1})$.

3. Let $X$ be a regular Euclidean octagon with opposite sides identified.

To see the first two surfaces are the same, we use the uniformization theorem to see $Z \cong \widehat{\mathbb{C}}$. Now $Z$ has symmetry group $S_4$, and the obvious rotation of order 4 means we can take $B$ equal to 0, $\infty$ and the 4th roots of unity. Then the uniqueness of branched coverings shows the result is isomorphic to the algebraic curve $y^2 = x(x^4 - 1)$.

This isomorphism reveals that $\text{Aut}(X)$ has a subgroup $G$ of order 48, the 'double cover' of $S_4$, isomorphic to the preimage of $S_4$ in $\text{SU}(2)$. In fact $y \mapsto -y$ gives the center $\mathbb{Z}/2$ of $G$. In particular, $X$ has an automorphism of order 8 given by $(x, y) \mapsto (ix, \sqrt{i}y)$. (We will eventually prove, in the study of hyperelliptic Riemann surfaces, that $G = \text{Aut}(X)$.)

The regular octagon gives the same surface of genus two. Indeed, the regular octagon $P$ also has an automorphism of order 8, which gives rise to an automorphism $r : X \to X$ of order 8 since the gluing instructions are respected. Note that $r$ has 6 fixed points: 1 coming from the center of the octagon, 1 coming from the 8 vertices (which are all identified), and 4 coming from the edge midpoints (which are identified in pairs).

We can construct an explicit map of degree 2 from $X = P/\sim$ to $\widehat{\mathbb{C}}$ branched over $(0, 1, \sqrt[4]{1})$. More precisely, we will construct an isomorphism of $Y = X/\langle r^4 \rangle$ with $\widehat{\mathbb{C}}$ such that $X \to Y \cong \widehat{\mathbb{C}}$ gives the desired branched cover. Note that its six fixed points (the Weierstrass points) correspond to the vertices and edge-midpoints of the octagon, and its center.

To describe $Y$, first observe that the quotient of the octagon by $\langle r^4 \rangle$ is a square $Q$. The edge identification on the square *fold* each edge at its midpoint, identifying the two halves, and $Y \cong Q/\sim$.

Topologically, the quotient of the boundary of the square by these identifications gives a plus sign. To make an isomorphism $Y \cong \widehat{\mathbb{C}}$, map $Q$ conformally to complement in $\widehat{\mathbb{C}}$ of the line segments from the 4th roots of unity to the origin, sending the center of the square to infinity, and sending the *midpoint* of each edge of the square to one of the 4th roots of unity. This map folds the edges of the square and identifies them. Because of its folding, this map respects the identifications on the sides of the square and the octagon, and gives an explicit degree two map of the octagon surface to $\widehat{\mathbb{C}}$, branched over the vertices of a octahedron.

## 2 Maps between Riemann surfaces

The arrows in the category of Riemann surfaces are as important as the objects themselves. In this section we discuss holomorphic maps $f : X \to Y$ between Riemann surfaces. Particular types of maps with good features that we will discuss are proper maps and branched coverings, especially regular (Galois) branched coverings.

As we will explain, every compact Riemann surface $X$ can be presented as a branched covering of $\widehat{\mathbb{C}}$, and if $X$ is defined over a number field then we can arrange that the covering is branched over just 3 points.

**Local analysis and multiplicity.** Let $f : X \to Y$ be holomorphic, with $f(p) = q$. Then we can find charts $(U, p) \cong (\Delta, 0)$ and $(V, q) \cong (\Delta, 0)$ so that $f$ goes over to an analytic map

$$F : (\Delta, 0) \to (\Delta, 0).$$

**Theorem 2.1** *The charts can be chosen so that $F(z) = 0$ or $F(z) = z^d$, $d > 0$.*

**Proof.** First choose arbitrary charts, and assume $F(z) \neq 0$; then $F(z) = z^d g(z)$ where $g(0) \neq 0$. The function $g(z)^{1/d}$ can therefore be defined for $|z|$ sufficiently small, yielding a factorization $f(z) = h(z)^d$ defined near 0, where $h'(0) \neq 0$. Since $h$ is invertible near 0, it can be absorbed into the choice of chart; then after suitably rescaling in domain and range, we obtain $F(z) = z^d$ on $\Delta$. ∎

We let $\mathrm{mult}(f,p) = d$ when $F(z) = z^d$, and $\mathrm{mult}(f,p) = \infty$ when $F(z)$ is constant. Note that the multiplicity can be described topologically: in the first case, $f$ is locally $d$-to-1 near (but not at) $p$, while in the second case $f$ is locally constant.

By the description above, the set where $\mathrm{mult}(f,p) = \infty$ is open *and* the set where $\mathrm{mult}(f,p)$ is finite is open. Since $X$ is connected, one of these must be empty. This shows:

*If $f$ is not globally constant, then* $\mathrm{mult}_p(f) < \infty$ *for all $p$.*

In particular, one $f$ is constant on a nontrivial open set, it is globally constant. (This is intuitively reasonable by analytic continuation.)

*From now on in this section we assume $f$ is not constant.*

**Critical points and branching.** Let $C(f) = \{p \in X : \mathrm{mult}(f,p) > 1\}$ denote the discrete set of *critical point* of $f$. The *branch locus* $B(f) \subset Y$ is its image, $B(f) = f(C(f))$. The branch locus obeys the useful cocycle formula:

$$B(f \circ g) = B(f) \cup f(B(g)).$$

Here are some general properties of nonconstant analytic maps.

1. The map $f$ is *open* and *discrete*. That is, when $U$ is open so is $f(U)$, and $f^{-1}(q)$ is a discrete subset of $X$ for all $q \in Y$.

2. If $X$ is compact, then $Y$ is compact and $f(X) = Y$, since $f(X)$ is both open and closed. Moreover the fibers of $f$ are finite and $C(f)$ and $B(f)$ are finite.

3. An analytic function $f : X \to \mathbb{C}$ on a compact Riemann surface is constant.

4. To prove $\mathbb{C}$ is algebraically closed, observe that a polynomial with no zero would define an analytic map $f : \widehat{\mathbb{C}} \to Y = \widehat{\mathbb{C}} - \{0\}$, and $Y$ is not compact.

**Proper maps.** While compactness of $X$ and $Y$ are useful assumptions, many results follow by imposing a compactness property on $f$. We say $f : X \to Y$ is *proper* of $f^{-1}$ sends compact sets to compact sets; that is, whenever $K \subset Y$ is compact, so is $f^{-1}(K)$.

Assume $f$ is analytic and proper. Then:

1. $f$ is closed: i.e. $E$ closed implies $f(E)$ closed.

2. $f$ is surjective. (Since $f(X)$ is open and closed.)

3. If $D \subset X$ is discrete, so is $f(D)$.

4. In particular, the branch locus $B(f) \subset Y$ is discrete.

5. $f^{-1}(q)$ is finite for all $q \in Y$. (Since $f^{-1}(q)$ is compact and discrete.)

6. For any neighborhood $U$ of $f^{-1}(q)$ there exists a neighborhood $V$ of $q$ whose preimage is contained in $U$. (Since $f(X-U)$ is closed and does not contain $q$.)

7. If $f$ is a proper local homeomorphism, then it is a covering map.

   Indeed, a local homeomorphism is discrete, so given $q \in Y$ we can choose neighborhoods $U_i$ of its preimages $p_1, \ldots, p_n$ such that $f_i : U_i \to V_i$ is a homeomorphism, where $f_i = f|U_i$. Let $V$ be a disk neighborhood of $q$ such that $V \subset \bigcap f(U_i)$ and $f^{-1}(V)$ is contained in $\bigcup U_i$. Then it is easy to check that $f^{-1}(V)$ has a components $W_i \subset U_i$ and $f_i : W_i \to V$ is a covering map. ■

8. For any $q \in Y$ with $f^{-1}(q) = \{p_1, \ldots, p_n\}$, there exists a disk $V$ containing $q$ such that $f^{-1}(V) = \bigcup U_i$ with $U_i$ a disk, $p_i \in U_i$ and $f_i = f|U_i$ satisfies $f_i : (U_i, p_i) \to (V, q)$ is conjugate to $z \mapsto z^{d_i}$ on $\Delta$.

   (Here one can use the Riemann mapping theorem to prove that if $g : \Delta \to \Delta$ is given by $g(z) = z^m$, and $V \subset \Delta$ is a simply-connected neighborhood of $z = 0$, then the pullback of $g$ to $V$ is also conjugate to $z \mapsto z^m$.)

9. The function $d(q) = \sum_{f(p)=q} m(f,p)$ is independent of $q$. It is called the *degree* of $f$. In fact, the condition that $d(q)$ is finite and constant is equivalent to the condition that $f$ is proper.

**Finite branched coverings.** Let $f : X \to Y$ be a proper map of degree $d = \deg(f)$. Let $Y^* = Y - B(f)$ and $X^* = f^{-1}(Y^*)$. Then:

$$f : X^* \to Y^*$$

is a proper local homeomorphism, hence a covering map.

Conversely, one can easily show: given a discrete set $B \subset Y$ and a finite covering space $f : X^* \to Y^* = Y - B$, there is a unique way to extend $X^*$ and $f$ to obtain a proper map

$$f : X \to Y.$$

As a basic example, if $Y = \Delta$, $B = \{0\}$, then $Y^* = \Delta^*$ with $\pi_1(Y^*) \cong \mathbb{Z}$. Thus a covering space is determined by its degree $d$. But $f_d(z) = z^d : \Delta^* \to \Delta^*$ gives an explicitly model for $X^*/Y^*$, and it clearly extends unique to a proper map on $\Delta$.

The general case follows from this example. Here is a related result on filling in punctures. Let $X^* = X - D$ be the complement of a discrete set on a topological surface $X$. Suppose $X^*$ is given the structure of a Riemann surface. Then:

**Theorem 2.2** *The complex structure extends from $X^*$ to $X$ iff every point $x \in D$ has a neighborhood $U$ such that $U \cap X^* \cong \Delta^*$.*

Here we use the removability of isolated singularities for bounded functions. As an example, this result applies to the one-point compactification of $\mathbb{C}$; but *not* to the one-point compactification of $\mathbb{C}^*$. (For a fancy proof, observe that if $X = \Delta \cup \{\infty\}$ were a compact Riemann surface then it would have to be isomorphic to $\widehat{\mathbb{C}}$, and hence the complement of a point in $\widehat{\mathbb{C}}$ would be isomorphic to $\Delta$, contradicting Liouville's theorem.

**Regular branched covers.** We define the Galois group of a branched covering (= proper) map $f : X \to Y$ by

$$\operatorname{Gal}(X/Y) = \{g \in \operatorname{Aut}(X) \ : \ f \circ g = f\}.$$

If $|\operatorname{Gal}(X/Y)| = \deg(f)$, we say $X/Y$ is a *regular* (or Galois) branched cover.

Conversely, if $G \subset \operatorname{Aut}(X)$ is a finite group, the quotient $Y = X/G$ can be given the structure of a Riemann surface is a unique way such that the quotient map

$$f : X \to Y = X/G$$

is a proper holomorphic map. We have

$$C(f) = \{x \in X \ : \ |G^x| > 1\},$$

where $G^x$ denotes the stabilizer of $x$.

**Examples.**

1. Any analytic map between compact Riemann surfaces is proper.

2. Thus every rational map $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ is proper, with $B(f) = \{0, \infty\}$. The map $f(z) = z^d$ is regular, with $G \cong \mathbb{Z}/d$ generated by $z \mapsto \zeta_d z$.

3. The rational map $f(z) = z + 1/z$ is regular with Galois group $\mathbb{Z}/2$ generated by $z \mapsto 1/z$ and $B(f) = \{-2, 2\}$. If we restrict this to a map
$$f : \mathbb{C}^* \to \mathbb{C},$$
we see the covering map from an infinite cylinder to the infinite pillowcase.

4. The rational map $f(z) = z^n + 1/z^n$ has the same branch locus and Galois group the dihedral group $D_{2n}$.

5. There are similar rational maps for $A_4$, $S_4$ and $A_5$ of degrees 12, 24 and 60. See Klein's *Lectures on the Icosahedron*.

6. Any finite covering map between Riemann surfaces is proper.

7. For a complex torus $E = \mathbb{C}/\Lambda$, we obtain a finite covering map $f_\alpha : E \to E$ for every element of
$$\mathrm{End}(E) = \{\alpha \in \mathbb{C}^* \ : \ \alpha\Lambda \subset \Lambda\}.$$
We have $\deg(f_\alpha) = |\alpha|^2$. Note that many such $\alpha$ exist when e.g. $\Lambda = \mathbb{Z}[\sqrt{-d}]$ is a ring (more generally, when $\Lambda \otimes \mathbb{Q}$ is a quadratic field.)

8. The Weierstrass $\wp$ function gives a regular proper map $f : E \to \widehat{\mathbb{C}}$ of degree 2 with $|B(f)| = 4$.

9. Let $E = \mathbb{C}/\mathbb{Z}[i]$ and let $G = \mathbb{Z}/4$ acting by rotations. Then $E/G \cong \widehat{\mathbb{C}}$ and the natural map $f : E/G \to \widehat{\mathbb{C}}$ has degree 4 and $|B(f)| = 3$. Indeed, $C(f)$ consists of the 4 points represented by the center, vertex, and edge midpoints of the square, and the two types of edge midpoints are identified by the action of $G$.

   The quotient can, if desired, be constructed as the composition of a degree two map $\wp : E \to \widehat{\mathbb{C}}$ branched over $0, \infty, \pm - 1$, with $s(z) = z^2$, resulting in $B(f) = \{0, 1, \infty\}$ with $f = s \circ \wp$. This is a good illustration of the principle
$$B(s \circ \wp) = B(s) \cup s(B(\wp)).$$

Geometrically, $E/G$ is the double of an isosceles right triangle. (It is called the $(2,2,4)$–orbifold.)

10. An entire function is proper iff it is a polynomial.

11. Any proper map of the disk to itself is given by a Blaschke product,

$$f(z) = \exp(i\theta) \prod_1^d \frac{z - a_i}{1 - \overline{a}_i z}.$$

(Proof: $f$ has have a least one zero. Taking a quotient with a Blaschke product, the result is either a constant of modulus one or a proper map $g : \Delta \to \Delta$ with no zeros. The latter is impossible.)

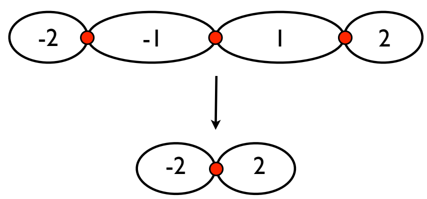



Figure 1. Topology of a cubic polynomial.

12. A basic example of an irregular branched cover is given by $f : \mathbb{C} \to \mathbb{C}$ the cubic polynomial
$$f(z) = z^3 - 3z.$$
(See Figure 1). This map has critical points $C(f) = \{-1, 1\}$ and branch values $B(f) = \{\pm 2\}$; we have $\widetilde{B} = f^{-1}(B) = \{\pm 1, \pm 2\}$, and $f : \mathbb{C} - \widetilde{B} \to \mathbb{C} - B$ is an irregular degree three cover. It corresponds to the *abbaab* triple cover of a bouquet of circles with fundamental group $F_2 = \langle a, b \rangle$.

We remark that *every* cubic polynomial $p(z)$ is equivalent $f(z)$ or $c(z) = z^3$, in the sense that it is equal to $f$ or $c$ after an affine change of coordinates in domain and range.

13. The map $f(z) = z^4 : \mathbb{H} \to \mathbb{C}^*$ is not proper, even though it is a local homeomorphism. Indeed, the fibers of $f$ have cardinality 2 for most points in $\mathbb{C}^*$, but the fiber over $z = 1$ consists of the single point $i$ with multiplicity 1. Thus the 'degree' of $f$ is not constant.

14. Let $B \subset \widehat{\mathbb{C}}$ be a finite set with $|B| = 2n$, and let $Y^* = \widehat{\mathbb{C}} - B$. Then $\pi_1(Y^*) \cong F_{2n-1}$, and there is a unique map to $\mathbb{Z}/2$ which sends each peripheral loop to 1. The resulting 2-fold covering space can be completed to give a compact Riemann surface $f : X \to \widehat{\mathbb{C}}$ with $B(f) = E$, $\deg(f) = 2$ and the genus of $X$ is $n - 1$. If $B$ coincides with the zero set of a polynomial $p(x)$ of degree $2n$, then $X$ is nothing more than the (completion of) the algebraic curve $y^2 = p(x)$.

**Infinite branched coverings.** Another class of well–behaved maps between Riemann surfaces are the covering maps, possibly of infinite degree. We would like to allow these maps to be branched as well. To this end, we say $f : X \to Y$ is a *branched covering* if

1. Each $q \in Y$ has a neighborhood $V$ such that $f_i : U_i \to V$ is proper, when $U_i$ are the components of $f^{-1}(V)$ and $f_i = f|U_i$.
2. The branch locus $B(f) \subset Y$ is discrete.

(Some authors omit the second requirement).

The Galois group is defined as before. As before, we can set $Y^* = Y - B(f)$, $X^* = f^{-1}(Y^*)$ and obtain a covering map

$$f : X^* \to Y^*.$$

This data uniquely determines $(X, f)$ up to isomorphism over $Y$. Conversely, for any discrete set $B \subset Y$ and any covering space $p : X^* \to Y^*$ as above, *with* the property that whenever $U \subset Y$ is a disk and $|U \cap B| \leq 1$, each component of $p^{-1}(U^*)$ maps to $U$ with finite degree, there is an associated branched covering $f : X \to Y$.

Regular branched coverings are also easily constructed: given any group $G \subset \mathrm{Aut}(X)$ acting *properly discontinuously* on $X$, there exists a Riemann surface $Y \cong X/G$ and a regular branched covering

$$f : X \to Y = X/G$$

with Galois group $G$.

**Examples.**

1. The map $f : \mathbb{C} \to \mathbb{C}^*$ given by $f(z) = e^{iz}$ is a regular covering map, with degree group $G = 2\pi\mathbb{Z}$.

2. The map $\cos : \mathbb{C} \to \mathbb{C}$ is a regular *branched* covering map, with $B(\cos) = \{\pm 1\}$. This gives the universal covering map for the infinite pillowcase. Its Galois group is $2\pi\mathbb{Z} \ltimes \mathbb{Z}/2$.

   The map $\cos(z) = (e^{i\theta} + 1/e^{i\theta})/2$ can be factored as $\mathbb{C} \to \mathbb{C}^* \to \mathbb{C}$, where the first map is a covering map and the second has degree two.

3. Given a lattice $\Lambda$, the map $\wp : \mathbb{C} \to \widehat{\mathbb{C}}$ is a regular branched covering, with Galois group the dihedral group $\Lambda \ltimes \mathbb{Z}/2$.

4. The map $f(z) = \tan(z) : \mathbb{C} \to f(\mathbb{C}) \subset \widehat{\mathbb{C}}$ is a *covering map.* In fact $\tan(z) = g(e^{2iz})$ where $g(z) = -i(z-1)/(z+1)$. Thus $f(\mathbb{C}) = \widehat{\mathbb{C}} - \{\pm i\}$.

**The Riemann-Hurwitz formula.** We now return our attention to compact Riemann surfaces, or more generally Riemann surfaces of *finite type.*

The Euler characteristic of a compact Riemann surface $X$ of genus $g$ is given by

$$\chi(X) = 2 - 2g.$$

If we remove $n$ points from $X$ to obtain an open Riemann surface $X^*$, then

$$\chi(X^*) = 2 - 2g - n.$$

We refer to $X^*$ as a Riemann surface of *finite type.*

**Theorem 2.3** *If $f : X \to Y$ is a branched covering map between Riemann surfaces of finite type, then*

$$\chi(X) = d\chi(Y) - \sum_{x \in C(f)} (\operatorname{mult}(f, x) - 1).$$

**Examples.**

1. Let $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ be a rational map of degree $d$, and let $C$ be the number of critical points of $f$, counted with multiplicities. Then $2 = 2d - C$ and thus $C = 2d - 2$.

2. Let $\wp : E \to \widehat{\mathbb{C}}$ be the Weierstrass $\wp$-function. Then $\deg(\wp) = 2$, so all its branched points are simple, and from

   $$0 = 2 \cdot 2 - |B(\wp)|$$

   we conclude that $|B(\wp)| = 4$.

3. Let $f : E \to \widehat{\mathbb{C}}$ give the quotient of the square torus by a 90 degree rotation, and let $C$ denote the number of critical points of $f$ counted with multiplicities. Then

$$0 = 4 \cdot 2 - C$$

so $C = 8$. In fact, $f$ has critical points of multiplicity 3 at the center and vertices of a fundamental square, and simple critical points at the edge midpoints.

4. If $X$ is compact and $f : X \to X$ has degree $d > 1$, then $\chi(X) \geq 0$. Thus the sphere and the torus are the only compact Riemann surfaces admit self–maps of higher degree. Moreover, if $X = E$ is a complex torus, then $f$ is simply a covering map.

5. Let $f : X \to \widehat{\mathbb{C}}$ be a map of degree 2 branched over $2n$ points, and let $g$ be the genus of $X$. Then we have:

$$2 - 2g = 2 \cdot 2 - 2n,$$

and thus $g = n - 1$ as previously discussed.

**Branched covers in higher dimensions.** It is a special feature of the topology in dimension two that the branched cover of a surface is a surface, and the quotient $X/G$ is a surface when $G$ preserved orientation.

A nice example in higher dimensions is furnished by the three spaces:

1. The quotient $\mathbb{C}^2/G$ where $G$ acts by $(x, y) \mapsto (-x, -y)$;
2. The singular quadric $z^2 = xy$ in $\mathbb{C}^3$; and
3. The 2-fold cover of $\mathbb{C}^2$ branched over the pair of lines $x = 0$ and $y = 0$.

These 3 spaces are isomorphic and they are all singular complex surfaces. The topology of the isolated singularity is seen most clearly in the case of (1): the link of the singularity is $S^3/G \cong \mathbb{RP}^3$. The reason this type of example does not exist in complex dimension 1 is that $S^1/G \cong S^1$ when $G$ is a finite group of rotations.

To see the isomorphic between (1) and (2), observe that $(u, v, w) = (x^2, y^2, xy)$ are invariant under $G$, and $w^2 = uv$.

As for (3), one should observe that the locus $xy = 0$ meets $S^3$ in the Hopf link $L \subset S^3$, and the 2-fold cover of $S^3$ branched over $L$ is $\mathbb{RP}^3$.

For a related example, $y^2 = x^3$ meets $S^3$ in the trefoil knot, and $y^p = x^q$ meets $S^3$ in the $(p, q)$ torus knot.

**Belyi's Theorem.** The Riemann surfaces that arise as coverings of $\widehat{\mathbb{C}}$ branched over $B = \{0, 1, \infty\}$ are especially interesting. In this case the configuration $B$ has no moduli, so $X$ is determined by purely combinatorial data.

**Theorem 2.4** *A compact Riemann surface $X$ is defined over a number field iff it can be presented as a branched cover of $\widehat{\mathbb{C}}$, branched over just 3 points.*

We first need to explain what it means for $X$ to be defined over a number field. For our purposes this means there exists a branched covering map $f : X \to \widehat{\mathbb{C}}$ with $B(f)$ consisting of algebraic numbers. Alternatively, $X$ can be described as the completion of a curve in $\mathbb{C}^2$ defined by an equation $f(x, y) = 0$ with algebraic coefficients.

Grothendieck wrote that this was the most striking theorem he had heard since at age 10, in a concentration camp, he learned the definition of a circle as the locus of points equidistant from a given center.

**Proof.** I. Suppose $X$ is defined over a number field, and $f : X \to \widehat{\mathbb{C}}$ is given with $B(f)$ algebraic. We will show there exists a polynomial $p : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ such that $B(p) \subset \{0, 1, \infty\}$ and $p(B(f)) \subset \{0, 1, \infty\}$. This suffices, since

$$B(p \circ f) = B(p) \cup p(B(f)).$$

We may assume $\infty \notin B(f)$. Let $\deg(B)$ denote the maximal degree over $\mathbb{Q}$ of the points of $B$, and let $z \in B(f)$ have degree $\deg(B)$. Then there exists a polynomial $p \in \mathbb{Q}[z]$ of degree $d$ such that $p(z) = 0$. Moreover $\deg(B(p)) \le d - 1$ since the critical values of $B$ are the images of the zeros of $p'(z)$. Thus

$$B' = B(p) \cup p(B)$$

has fewer points of degree $\deg(B)$ over $\mathbb{Q}$. Iterating this process, we can reduce to the case where $B \subset \mathbb{Q}$.

Now comes a second beautiful trick. Consider the polynomial $p(z) = Cz^a(1 - z)^b$. This polynomial has critical points at 0, 1 and $w = a/(a + b)$. By choosing the value of $C \in \mathbb{Q}$ correctly, we can arrange that $p(w) = 1$ and hence $B(p) \subset \{0, 1, \infty\}$. Thus if $x \in B$ is rational, we can first apply a Möbius transform so $0 < x < 1$; then, write $x = a/(a + b)$, and choose $p$ as above so that $B' = p(B) \cup B(p)$ has fewer points outside $\{0, 1, \infty\}$. Thus we can eventually eliminate all such points.

II. Let $X$ be a Riemann surface presented as a branched covering $f : X \to \widehat{\mathbb{C}}$ with $B(f) = \{0, 1, \infty\}$. We need a nontrivial fact: there exists a second $g \in \mathcal{M}(X)$ and a cofinite set $X^* \subset X$ such that $(f, g) : X^* \to \mathbb{C}^2$ is an immersion with image the zero locus $Z(p)$ of a polynomial $p(x, y) \in \mathbb{C}[x, y]$.

This polynomial has the property that the projection of (the normalization of) $Z(p)$ to the first coordinate — which is just $f$ — is branched over just 0, 1 and $\infty$. The set of all such polynomials of given degree is an algebraic subset $W$ of the space of coefficients $\mathbb{C}^N$ for some large $N$. This locus $W$ is defined by rational equations. Thus the component of $W$ containing our given $p$ also contains a polynomial $q$ with coefficients in a number field. But $Z(q)$ and $Z(p)$ are isomorphic, since they have the same branch locus and the same covering data under projection to the first coordinate. Thus $X^* \cong Z(p) \cong Z(q)$ is defined over a number field. ∎

**Corollary 2.5** *$X$ is defined over a number field iff $X$ can be built by gluing together finitely many unit equilateral triangles.*

**Proof.** If $X$ can be built in this way, one can 2-color the barycentric subdivision and hence present $X$ as a branched cover of the double of a 30-60-90 triangle. The converse is clear, since the sphere can be regarded as the double of an equilateral triangle. ∎

Example: The square torus can be built by gluing together 8 isosceles right triangles. But these can also be taken to be equilateral triangles, with the same result! This is because the double of *any* two triangles is the same, as a Riemann surface with 3 marked points.

**Trees and Belyi polynomials.** Let us say $p : \mathbb{C} \to \mathbb{C}$ is a *Belyi polynomial* if $p(0) = 0$, $p(1) = 1$ and the critical values of $p$ are contained in $\{0, 1\}$. Then $p^{-1}[0, 1]$ is a topological tree $T \subset \mathbb{C}$. (This is because $p^{-1}(\mathbb{C} - T)$ maps to $\mathbb{C} - T$ by a covering map of degree $\deg(p)$.)

We adopt the convention that the inverse images of 0 or 1 coincide with the vertices of $T$, and that the vertices of $T$ *at* 0 and 1 are labeled. The images of the vertices of the $T$ then alternate between 0 and 1.

A tree of this type encodes a branched covering of $\widehat{\mathbb{C}}$ of genus 0, which by the uniformization theorem is simply a copy of $\widehat{\mathbb{C}}$ again. Thus:

**Theorem 2.6** *Any finite topological tree $T \subset \mathbb{C}$ with 2 vertices at 0 and 1 determines a unique Belyi polynomial.*

**Examples.**

1. A star with $d$ endpoints, and its center labeled zero, corresponds to $p(z) = z^d$.

2. A segment with 2 internal vertices marked 0 and 1 corresponds to the cubic polynomial $p(z) = 3z^2 - 2z^3$. To see this, we note that $p'(z) = az(1-z)$ since the two vertices are internal, and use the normalization $p(0) = 0$ and $p(1) = 1$ to solve for $a$ and then for $p(z)$.

**Describing compact Riemann surfaces and their meromorphic functions.** Let $X$ be a compact Riemann surface, and let $f : X \to \widehat{\mathbb{C}}$ be a meromorphic function of degree $d \geq 1$. Then $(X, f)$ determines the following data:

- A finite set $B = B(f) \subset \widehat{\mathbb{C}}$; and
- A *monodromy map* $\phi : \pi_1(\widehat{\mathbb{C}} - B, y) \to S_d$.

To obtain the latter, choose $y \in Y^* = \widehat{\mathbb{C}} - B$ and label the points in $f^{-1}(y)$ as $x_1, \ldots, x_d$. Then path lifting gives a natural action of $\pi_1(Y^*, y)$ on the fiber over $y$, and hence a representation $\phi : \pi_1(Y^*, y) \to S_d$. Explicitly, $\rho(\gamma)(i) = j$ if a lift of $\gamma$ starting at $x_i$ terminates at $x_j$.

Let $X^* = f^{-1}(Y^*)$. Then $f : X^* \to Y^*$ is a covering map, and clearly

$$\pi_1(X^*, x_i) \cong \{\gamma \in \pi_1(Y^*, y) \,:\, \rho(\gamma)(i) = i\}.$$

To be even more explicit, let us choose loops $\gamma_1, \ldots, \gamma_n$ based at $y$ and encircling the points of $B = (b_1, \ldots, b_n)$ in order. Then $\pi_1(Y^*, y)$ is generated by these elements, with the single relation

$$\gamma_1 \cdots \gamma_n = e.$$

Thus $\phi$ is uniquely determined by the permutations $\sigma_i = \phi(\gamma_i)$, which must satisfy two properties:

- The product $\sigma_1 \cdots \sigma_n = e$ in $S_d$; and
- The group $G = \langle \sigma_1, \ldots, \sigma_n \rangle$ is a transitive subgroup of $S_d$.

The conjugacy class of $\sigma_i$ in $S_d$ determines a partition $p_i$ of $d$, which is independent of choices. This partition records how the sheets of $X^*$ come together over $b_i$; it can be alternatively defined by

$$p_i = \{\mathrm{mult}(f, x) \,:\, f(x) = b_i\}.$$

**Theorem 2.7** *The pair $(X, f)$ is uniquely determined by the data $(B, \sigma_1, \ldots, \sigma_n)$ above, and any data satisfying the product and transitivity conditions arises from some $(X, f)$.*

**Proof.** The covering space $f : X^* \to Y^*$ is determined, up to isomorphism over $Y$, by the preimage of $G \cap S_{d-1}$ in $\pi_1(Y^*, y)$; filling in the branch points, we obtain the desired compact Riemann surface $X$ and meromorphic function $f$. ■

**Example: Belyi trees.** For the case of Belyi polynomials, with $B = \{0, 1, \infty\}$, we can take $z = 1/2$ as the basepoint of $\widehat{\mathbb{C}} - B$ and use loops $\gamma_0$ and $\gamma_1$ around $z = 0$ and $z = 1$ as generators of the fundamental group. The permutation representation, $\sigma_i = \rho(\gamma_i)$, is easily read off from a picture of the tree formed by the preimage of $[0, 1]$. An example is shown in Figure 2.

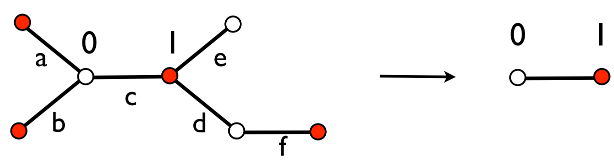



Figure 2. This Belyi tree corresponds to $\sigma_0 = (abc)(df)$, $\sigma_1 = (cde)$.

**Moduli of Riemann surfaces.** Note that if we fix the genus of $X$ and the degree of $f$, then the cardinality of $B$ is also bounded. On the other hand, one can show that $X$ admits a meromorphic whose degree is bounded by $g + 1$. Thus Riemann surfaces of genus $g$ depend on only finitely many parameters.

**Hurwitz problem.** Here is an unsolved problem. Let $f : X \to \widehat{\mathbb{C}}$ be a branched covering of degree $d$ with critical values $B = (b_i)_1^n$. Let $p_i$ be the partition of $d$ corresponding to the fiber over $b_i$. What partitions $(p_1, \ldots, p_n)$ can be so realized?

This is really a problem in topology or group theory. We have to lift each $p_i$ to an element $g_i \in S_d$ in the conjugacy class specified by $p_i$, in such a way that $g_1 \cdots g_n = e$ and $\langle g_1, \ldots, g_n \rangle$ acts transitively on $(1, \ldots, d)$.

**Examples.**

1. For $n = 3$ and $d = 4$, the partitions $2+2, 2+2, 3+1$ cannot be realized; since any 2 involutions in $A_4$ commute, $\sigma_1\sigma_2$ has order 1 or 2, not 3.

2. If $\sigma_1, \ldots, \sigma_{n-1}$ lie in $A_n$, then so must $\sigma_n$. So for example, the partitions of $d = 5$ given by $2 + 2 + 1$, $3 + 1 + 1$ and $2 + 1 + 1 + 1$ cannot coexist.

In fact, this $A_n$ condition just insures that $\deg C(f)$ is even — a necessary condition, by the Riemann–Hurwitz formula. Note that the Euler characteristic of $X$ can be calculated from the partitions $(p_i)$. For more details, see [EKS].

**Have we seen all compact Riemann surfaces?** Yes! A basic result, the *Riemann Existence Theorem*, to be proved later (for compact $X$), states:

**Theorem 2.8** *Every Riemann surface admits a nonconstant meromorphic function. In fact, the elements of $\mathcal{M}(X)$ separate points and provide local charts for $X$.*

In more detail, this means for every $p \neq q$ in $X$ there is an $f : X \to \widehat{\mathbb{C}}$ such that $f(p) \neq f(q)$ and $df(p) \neq 0$. In classical terminology, this means $f$ also separates *infinitely near* points.

# 3 Sheaves and analytic continuation

Sheaves play an essential role in geometry, algebra and analysis. They allow one to cleanly separate the local and global features of a construction, and to move information between analytical and topological regimes.

For example, it is trivial to locally construct a meromorphic function on a compact Riemann surface $X$. We will use sheaf theory to investigate the global obstructions. It will turn out these obstructions are finite dimensional and thus they can be surmounted.

**Presheaves and sheaves.** A *presheaf of abelian groups* on $X$ is a functor $\mathcal{F}(U)$ from the category of open sets in $X$, with inclusions, to the category of abelian groups, with homomorphisms. (This functor is covariant, i.e. it reverses the directions of arrows.)

It is a *sheaf* if

(I) elements $f \in \mathcal{F}(U)$ are determined by their restrictions to an open cover $U_i$, and
(II) any collection $f_i \in \mathcal{F}(U_i)$ with $f_i = f_j$ on $U_{ij}$ for all $i, j$ comes from an $f \in \mathcal{F}(U)$.

These axioms say we have an exact sequence

$$0 \to \mathcal{F}(U) \to \prod \mathcal{F}(U_i) \to \prod \mathcal{F}(U_{ij}).$$

Here $U = \bigcup U_i$ is an open cover, the index $ij$ is considered to be *ordered*, $U_{ij} = U_i \cap U_j$, and the maps are given by $f \mapsto f|U_i$ and $(f_i) \mapsto (f_i|U_{ij}) - (f_j|U_{ij})$.

Since $\mathcal{F}(U)$ is simply the kernel of the consistency map, a sheaf is *uniquely determined* once we know $\mathcal{F}(U)$ for all $U$ sufficiently small.

**Cohomology.** The exact sequence above hints at a cohomology theory, and indeed it can be interpreted as:

$$0 \to Z^0(\mathcal{F}) \to C^0(\mathcal{F}) \stackrel{d}{\to} C^1(\mathcal{F}),$$

where the chain groups are taken relative to the covering $(U_i)$. There are no 0-coboundaries, so in fact

$$Z^0(\mathcal{F}) \cong H^0(\mathcal{F}) \cong \mathcal{F}(U).$$

**Examples.**

1. The sheaves $\mathcal{C}$, $\mathcal{C}^\infty$, $\mathcal{O}$ and $\mathcal{M}$, of rings of continuous, smooth, holomorphic and meromorphic functions. The multiplicative group sheaves $\mathcal{O}^*$ and $\mathcal{M}^*$.

2. Applying (I) to the empty cover of the empty set, we see $\mathcal{F}(\emptyset) = (0)$. Thus the presheaf that assigns a fixed, nontrivial group $G$ to every open set is not a sheaf.

3. If $G$ is a nontrivial abelian group, $\mathcal{F}(U) = G$ for $U$ nonempty and $\mathcal{F}(\emptyset) = (0)$ is a presheaf. But it is not a sheaf: for a pair of disjoint nonempty open sets, we have $G = \mathcal{F}(U_1 \sqcup U_2) \neq G \oplus G$.

4. To rectify this, we define $\mathcal{F}(U)$ to be the additive group of *locally constant maps* $f : U \to G$. This is now a sheaf; it is often denoted simply by $G$. (E.g. $\mathbb{R}$, $\mathbb{C}$, $\mathbb{Q}$, $S^1$, $\mathbb{Z}$.)

5. The presheaf $\mathcal{F}(U) = \mathcal{O}(U)/\mathbb{C}(U)$ is not a sheaf. For example, local logarithms do not assemble. This can be thought of as the presheaf of exact holomorphic 1-forms. The exactness condition is not local.

**Pushforward.** Sheaves naturally push *forward* under continuous maps, i.e. we can define

$$\pi_*(\mathcal{F})(U) = \mathcal{F}(\pi^{-1}(U)).$$

In particular, homeomorphisms between spaces give rise to isomorphisms of sheaves.

We can also pull back a sheaf under a *local homeomorphism.* This means we define $(\pi^*(\mathcal{F}))(U) = \mathcal{F}(U)$ when $U$ is small enough that $\pi$ is a local homeomorphism, and then extend using the sheaf axiom.

**Structure.** From a modern perspective, a Riemann surface can be defined as a pair $(X, \mathcal{O}_X)$ consisting of a connected Hausdorff topological space and a sheaf that is locally isomorphic to $(\mathbb{C}, \mathcal{O}_\mathbb{C})$. Similarly definitions work for smooth manifolds. For example: a homeomorphism $f : X \to Y$ between Riemann surfaces is holomorphic iff it sends $\mathcal{O}_X$ to $\mathcal{O}_Y$.

**Stalks.** Let $\mathcal{F}$ be a *presheaf.*

The *stalk* $\mathcal{F}_x$ is the direct limit of $\mathcal{F}(U)$ over the (directed) system of open sets containing $x \in X$. It can be described directly as the disjoint union of these groups modulo $f_1 \sim f_2$ if they have a common restriction near $x$. (Alternatively, $\mathcal{F}_x = (\oplus \mathcal{F}(U))/N$ where $N$ is generated by elements of the form $(f|U_1) - (f|U_2)$, $f \in \mathcal{F}(U_1 \cup U_2)$.)

There is a natural map $\mathcal{F}(U) \to \mathcal{F}_x$ for any neighborhood $U$ of $x$. We let $f_x$ denote the image of $f \in \mathcal{F}(U)$ under this map. We have $f_x = 0$ iff there is a neighborhood $V$ of $x$ such that $f|V = 0$.

**Example: The stalk as a local ring.** For any point $a \in X$ a Riemann surface, we have $\mathcal{O}_a \cong \mathbb{C}\{\{z_a\}\}$, the ring of convergent power series, for any local chart (uniformizer) $z_a : U \to \mathbb{C}$, $z_a(a) = 0$. Note that $\mathcal{O}_a$ is a *local ring*, with maximal ideal $m_a = (z_a)$ and *residue field* $\mathbb{C} = \mathcal{O}_a / m_a$. The natural map $\mathcal{O}_a \to \mathbb{C}$ is given by evaluation at $a$, and the associated discrete valuation on $\mathcal{O}_a$ measure the order of vanishing of the germ of a function.

**Theorem 3.1** *Let $\mathcal{F}$ be a sheaf. Then $f \in \mathcal{F}(U)$ is zero iff $f_x = 0$ for all $x \in U$.*

**Proof.** If $f_x = 0$ for each $x \in U$ then there is a covering of $U$ by open sets $U_i$ such that $f|U_i = 0$; then $f = 0$ by the sheaf axioms. The converse is obvious. ∎

**Espace étalé.** The *espace étalé* $|\mathcal{F}|$ of a presheaf is the disjoint union of the stalks $\mathcal{F}_x$, with a base for the topology given by sets of the form

$[U, f] = \{f_x \,:\, x \in U\}$. It comes equipped with a natural projection

$$p : |\mathcal{F}| \to X$$

which is a local homeomorphism. Every $f \in \mathcal{F}(U)$ gives a continuous section $s : U \to |\mathcal{F}|$ by $s(x) = f_x$.

**Example.** For a (discrete) group $G$, we have $|G| = G \times X$ with the product topology.

We say $\mathcal{F}$ satisfies the *identity theorem* if whenever $U$ is open and connected, $f, g \in \mathcal{F}(U)$ and $f_x = g_x$ for some $x \in U$, then $f = g$. Examples: $\mathcal{O}$, $\mathcal{M}$ and $G$.

**Theorem 3.2** *Assume $X$ is Hausdorff and locally connected, and $\mathcal{F}$ satisfies the identity theorem. Then $|\mathcal{F}|$ is Hausdorff.*

**Proof.** We need to find disjoint neighborhoods centered at a pair of distinct germs $f_x$ and $g_y$. This is easy if $x \neq y$, since $X$ is Hausdorff. Otherwise, choose a connected open set $U$ containing $x = y$, and small enough that both $f_x$ and $g_x$ are represented by $f, g \in \mathcal{F}(U)$. Then the identity theorem implies the neighborhoods $[U, f]$ and $[U, g]$ are disjoint, as required. ■

**Structure of $|\mathcal{O}|$.** Since analytic functions satisfy the identity theorem, $|\mathcal{O}|$ is Hausdorff. There is a unique complex structure on $|\mathcal{O}|$ such that $p : |\mathcal{O}| \to X$ is an *analytic* local homeomorphism.

There is also a natural map $F : |\mathcal{O}| \to \mathbb{C}$ given by $F(f_x) = f(x)$. This map is analytic. However $|\mathcal{O}|$ is never connected, since the stalks $\mathcal{O}_x$ are uncountable.

Thus an analytic function $f : U \to \mathbb{C}$ has *three* manifestations:

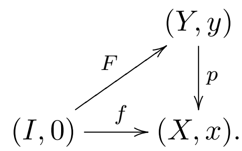
(i) as an element of an abelian group, $f \in \mathcal{O}(U)$;
(ii) as a complex number at each point, $f(a) \in \mathbb{C}$; and
(iii) as the germ of an analytic function, $f_a \in \mathcal{O}_a$.

**Values of functions.** The construction of the function $F$ on $|\mathcal{O}|$ appears *ad hoc*. It can be made more understandable by recalling that the stalks $\mathcal{O}_x$ are *local rings* with *residue field* $k_x = \mathcal{O}_x / m_x \cong \mathbb{C}$. Then $F(f_x)$ gives the value of $f_x$ in $k_x$.

For $k_x$ to be *canonically* identified with $\mathbb{C}$, one must regard $\mathcal{O}$ as a sheaf of $\mathbb{C}$-algebras; otherwise, complex conjugation might intervene.

**Path lifting.** We now recall some results from the theory of covering spaces.

Let $p : (Y, y) \to (X, x)$ be a local homeomorphism between Hausdorff spaces, and let $f : (I, 0) \to (X, x)$ be a path. A *lift* of $f$ is a path $F : (I, 0) \to (Y, y)$, satisfying $p \circ F = f$:




**Theorem 3.3 (Uniqueness of lifts)** *A lift of $f$ is unique if it exists.*

**Proof.** If we have two liftings, $F_1$ and $F_2$, the set of $t \in I$ such that $F_1(t) = F_2(t)$ is open since $p$ is a local homeomorphism, closed since $X$ is Hausdorff, and nonempty since $F_1(0) = F_2(0) = x$. By connectedness it is the whole interval. ∎

Now let $f_s(t)$ be a homotopy of paths parameterized by $s \in [0, 1]$, such that $f_s(0) = y_0$ and $f_s(1) = y_1$ are constant. Suppose for every $s$ there is a lift $F_s$ of $f_s$ based at $x_0$. We then have:

**Theorem 3.4 (Monodromy Theorem)** *The terminus $F_s(1) = x_1$ is independent of $s$. Moreover $F_s(t)$ is a lift of $f_s(t)$ as a function on $I \times I$. In particular, $F_0(t)$ and $F_1(t)$ are homotopic* rel *their endpoints.*

**Structure of $|C^k|$.** There are plenty of natural sheaves such that $|\mathcal{F}|$ is not Hausdorff and path lifting is not unique.

A good example is provide by the sheaf $C^k$ of $k$-times continuously differentiable functions on $X = \mathbb{R}$. You can think of $|C^k|$ as the space of all graphs of $C^k$ functions, with overlap graphs separated like the strands in a knot projection. However graphs that coincide over an open set are glued together. A point know which graph it is on, at least locally.

Now any $C^k$ function $f : \mathbb{R} \to \mathbb{R}$ defines a section $s(x) = f_x$, and hence a path in $|C^k|$. Consider, for example, $k = 0$ and the sections $s_1$ and $s_2$ defined by $f_1(x) = x$ and $f_2(x) = |x|$. These agree for $x > 0$ but not for $x \leq 0$. Thus path–lifting is not unique, and the germs of $x$ and $|x|$ at $x = 0$ cannot be separated by disjoint open sets in $|C^k|$, so the space is not Hausdorff.

Note that we are at least guaranteed that when the lift of a path goes wayward, the two germs at least have the same value at the point of departure (here $x = 0$).

In $|C^k|$ the first $k$ derivatives must coincide at the point of departure. In $|C^\infty|$, the graphs must make contact of infinite order. Finally in the space of real–analytic functions, the graphs must coincide and the space becomes Hausdorff.

**Structure of Spec $\mathbb{Z}$.** There is a similar natural sheaf defined on the spectrum of any commutative ring, e.g.

$$X = \operatorname{Spec} \mathbb{Z} = \{0, 2, 3, 5, 7, 11, \ldots\}$$

with the *Zariski topology.* Here the primes $p > 0$ are closed points, but the closure of the generic point 0 is the whole space. Over a prime $p$ we have a stalk

$$\mathcal{O}_p \cong \mathbb{Z}[1/q \,:\, \gcd(p, q) = 1],$$

with residue field $\mathbb{F}_p$, and over the generic point the stalk is $\mathbb{Q}$. We can define a map $F : |\mathcal{O}| \to \sqcup \mathbb{F}_p$ as before; then an integer $n \in \mathcal{O}(X) = \mathbb{Z}$ gives a section $n : X \to \sqcup \mathbb{F}_p$ with $n(p) = n \bmod p$.

**Completions; affine curves.** In the analytic study of a Riemann surface, all local rings are isomorphic. But from the algebraic point of view, this is not the case. That is, if we start with $X = \operatorname{Spec} \mathbb{C}[x, y]/p(x, y)$ a smooth affine curve, with the Zariski topology and the algebraic structure sheaf, then the field of fractions of $\mathcal{O}_a$ is $K(X)$. The stalks remember too much. To rectify this, it is standard to *complete* these local rings and consider instead

$$\varprojlim \mathcal{O}_a / m_a^n \cong \mathbb{C}[[z_a]].$$

This gives the ring of formal power series, which is now sufficiently localized that it is the same ring at every point.

**Analytic continuation.** Let $\mathcal{O}$ be the sheaf of analytic functions on $X = \mathbb{C}$. Let $\gamma : [0, 1] \to \mathbb{C}$ be path with $a = \gamma(0)$ and $b = \gamma(1)$. Let $f_a$ be the germ of an analytic function at $z = a$, i.e. let $f \in \mathcal{O}_a$.

We say $f_b \in \mathcal{O}_b$ is obtained from $f_a$ by *analytic continuation along* $\gamma$ if there are analytic functions $g_i$ on balls $B_i$ centered at $\gamma(t_i)$, $i = 0, \ldots, n$ such that:

1. $0 = t_0 < t_1 < \cdots < t_n = 1$;
2. $B_i$ meets $B_{i+1}$, $g_i$ and $g_{i+1}$ agree there, and $\gamma[t_i, t_{i+1}] \subset B_i \cap B_{i+1}$; and
3. $g_0 = f_a$ and $g_n = f_b$.

**Modern translation.** Analytic continuation along $\gamma$ is equivalent to path-lifting to the space $p : |\mathcal{O}| \to \mathbb{C}$, with the lift defined by $\widetilde{\gamma}(t) = [g_i] \in \mathcal{O}_{\gamma(t)}$ for $t \in [t_i, t_{i+1}]$. Thus we have, from the previous results on path–lifting:

**Corollary 3.5** *An analytic continuation of $f_a$ along a given path $\gamma$ is unique if it exists.*

This Corollary is also obvious from the definitions, by applying the identity theorem on the connected sets $B_i \cap B_{i+1}$ and induction on $i$.

**Corollary 3.6 (Monodromy theorem)** *If analytic continuation from $a$ to $b$ is possible along a family of paths $\gamma_s$, $s \in I$, then the result $f_b$ is always the same.*

**Maximal analytic continuation (spreads).** To keep track of the potential multi-valuedness, we define an *analytic continuation* of $f_a \in \mathcal{O}_a$, $a \in \mathbb{C}$, to be a pointed Riemann surface $(Y, b)$ endowed with an analytic local homeomorphism $p : (Y, b) \to (\mathbb{C}, a)$ and an $F \in \mathcal{O}(Y)$ such that $p_* F_b = f F_a$.

We say an analytic continuation $(Y, b)$ is *maximal* (universal might be a better terminology) if, for any other analytic continuation $(Y', b')$, there is an (analytic) local homeomorphism $j : (Y', b') \to (Y, b)$ making the natural diagrams commute.

The map $j$ is unique if it exists, by our assumption on basepoints and uniqueness of path lifting. Thus any 2 maximal analytic continuations are isomorphic.

However $j$ need not be an embedding! For example, the 2-sheeted map $p : (\mathbb{C}^*, 1) \to (\mathbb{C}, 1)$ with $F(z) = p(z) = z^2$ gives an analytic continuation of $f(z) = z$, even though the unique maximal analytic continuation is 1-sheeted.

**Theorem 3.7** *Every germ of an analytic function $f_a$ has a unique maximal analytic continuation, obtained by taking $Y$ to be the connected component of $|\mathcal{O}|$ containing $f_a$, and restricting the natural maps $F : |\mathcal{O}| \to \mathbb{C}$ and $p : |\mathcal{O}| \to \mathbb{C}$ to $Y$.*

**Proof.** The map $y \mapsto p_*(F_y)$ gives a local homeomorphism of any analytic continuation into $|\mathcal{O}|$, which is then tautologically contained in the maximal one described above. ∎

**Examples.**

1. The maximal domain of $f(z) = \sum z^{n!}$ is the unit disk.

2. For $f(z) = \sum z^n$, the maximal analytic continuation in $\mathbb{C}$ is $p : \mathbb{C} - \{1\} \to \mathbb{C}$ with $p(z) = z$ and $F(z) = 1/(1-z)$. Evidentally $f(p(z)) = 1/(1-z) = F(z)$.

3. For $f(z) = \sum z^{n+1}/(n+1)$ — which gives a branch of $\log(1/(1-z))$ on the unit disk — the maximal analytic continuation is given by the universal covering map $p : \mathbb{C} \to \mathbb{C} - \{1\}$ with $p(z) = 1 - e^z$ and $F(z) = -z$. Check: $f(p(z)) = \log(1/e^z) = -z$.

4. Let $f(z) = \sqrt{q(z)}$ where $q(z) = 4z^3 + az + b$. Then its maximal analytic continuation is given by $Y = E - E[2]$ with $E = \mathbb{C}/\Lambda$, with $p : Y \to \mathbb{C}$ given $p(z) = \wp(z)$, and with $F(z) = \wp'(z)$.

**Remark.** One can replace the base $\mathbb{C}$ of analytic continuation with any other Riemann surface $X$. Note however that the 'maximal' analytic continuation may become larger under an inclusion $X \hookrightarrow X'$. An example follows.

**Counterintuitive extensions.** Let $f : \Delta \to \mathbb{C}$ be an analytic function such that $Z(f) \subset \Delta$ accumulates on every point of $S^1$. Then $f$ admits no analytic continuation in the sense above.

However, if may be that $(\Delta, f)$ can be extended as an abstract function on a Riemann surface. That is, there may be a proper inclusion $j : \Delta \hookrightarrow X$ and a map $F : X \to \mathbb{C}$ that extends $f$. For this somewhat surprising phenomenon to occur, it must certain be true that $j(Z(f))$ is a discrete subsets of $X$.

To construct explicit examples, it is useful to know that there exist Riemann mappings $j : \Delta \to U \subset \mathbb{C}$ such that the radial limit of $j(z)$ is infinity at every point. (The Fatou set of $J(\lambda e^z)$ for $\lambda$ small provides an explicit example of such a $U$.) Then one can construct a set $B \subset U$ which is discrete in $\mathbb{C}$, such that $j^{-1}(B)$ accumulates everywhere on $S^1$. Now consider an entire function $F : \mathbb{C} \to \mathbb{C}$ with zero set $B$, and let $f(z) = F \circ j(z)$ on the disk. Then $f$ has zeros accumulating everywhere on $S^1$, but $f$ extends from $\Delta$ to $\mathbb{C}$ under the embedding $j$.

# 4 Algebraic functions

We will develop two main results. Our main case of interest will be compact Riemann surfaces, but we will formulate them in general.

**Theorem 4.1** *Let $\pi : X \to Y$ be a holomorphic proper map of degree $d$. Then $\mathcal{M}(X)/\mathcal{M}(Y)$ is an algebraic field extension of degree $d$.*

To see the degree is exactly $d$ we need to know that $\mathcal{M}(X)$ separates the points of $X$, i.e. we will appeal to the Riemann Existence Theorem.

**Theorem 4.2** *Let $K/\mathcal{M}(Y)$ be a field extension of degree $d$. Then there is a unique degree $d$ branched covering $\pi : X \to Y$ such that $K \cong \mathcal{M}(X)$ over $\mathcal{M}(Y)$.*

*Moreover $\mathcal{M}(X)/\mathcal{M}(Y)$ is Galois iff $X/Y$ is Galois, in which case there is a natural isomorphism*

$$\mathrm{Gal}(\mathcal{M}(X)/\mathcal{M}(Y)) \cong \mathrm{Deck}(X/Y).$$

Putting these results together, we find for example that $X \mapsto \mathcal{M}(X)$ establishes an equivalence between (i) the category of finite-sheeted branched coverings of $\widehat{\mathbb{C}}$, and (ii) the category of finite field extensions of $\mathbb{C}(x)$.

**Symmetric functions.** The proof will implicitly use the idea of *symmetric functions.* Recall that the 'elementary symmetric functions' $s_i$ of $f_1, \ldots, f_d$ are related to the coefficients $c_i \in \mathbb{Z}[f_1, \ldots, f_d]$ of the polynomial

$$P(T) = \prod_1^d (T - f_i) = T^d + c_1 T^{d-1} + \cdots + c_d \tag{4.1}$$

by $s_i = (-1)^i c_i$. (Thus $s_1 = \sum f_i$, $s_2 = \sum_{i<j} f_i f_j$ and $s_d = \prod f_i$.) Clearly these polynomials $s_i$ are invariant under permutation of the $f_i$, and in fact the ring of invariant polynomials is given by:

$$\mathbb{Z}[f_i]^{S_d} = \mathbb{Z}[s_i];$$

see [La, §IV.6].

**Separation of points.** We also observe the following consequence of the fact that $\mathcal{M}(X)$ separates points. This should be compared to the theorem of the primitive element.

**Proposition 4.3** *Let $A \subset X$ be a finite subset of a Riemann surface. Then there exists a meromorphic function $f : X \to \widehat{\mathbb{C}}$ such that $f|A$ is injective.*

**Proof.** Let $V \subset \mathcal{M}(X)$ be a finite-dimensional space of functions that separate the points of $A = \{a_1, \ldots, a_n\}$. Then for any $i \neq j$, the subspace $V_{ij} \subset V$ where $f(a_i) \neq f(a_j)$ has codimension at least one. It follows that $V \neq \bigcup V_{ij}$, and hence a typical $f \in V$ is injective on $A$. ∎

**Proof of Theorem 4.1.** Let $\pi : X^* \to Y^*$ be the unbranched part of the degree $d$ branched covering $\pi : X \to Y$, and consider $f \in \mathcal{M}(X)$. By deleting more points if necessary, we can assume $f \in \mathcal{O}(X^*)$. We will show $\deg(f/\mathcal{M}(Y)) \le d$.

Given $y \in Y^*$, define a degree $d$ polynomial in $T$ by

$$P(T, y) = \prod_{\pi(x)=y} (T - f(x)) = \sum c_i(y) T^{d-i}.$$

Clearly $P(T, y)$ is *globally* well-defined on $Y^*$.

We now check its *local* properties. Since $\pi$ is a covering map over $Y^*$, we can choose local sections $q_i : U \to X^*$ such that $P(T, y) = \prod (T - f(q_i(y)))$ on $U$. This shows that the coefficients of $P(T, y)$ are analytic functions of $y$. On the other hand, the definition of $P(T, y)$ makes no reference to these local sections, so we get a globally well-defined polynomial $P \in \mathcal{O}(Y^*)[T]$ satisfying $P(f) = 0$.

Near any puncture of $Y$ over which $f$ takes finite values, the coefficients $c_i(y)$ of $P(T, y)$ are bounded, so they extend across the puncture. At any puncture where $f$ has a pole, we can choose a local coordinate $z$ on $Y$ and $n > 0$ such that $z^n f$ has no poles. Then there is a monic polynomial $Q(T)$ of degree $d$, with holomorphic coefficients, such that $Q(z^n f) = 0$ locally, constructed just as above. On the other hand, $P(T) = z^{-nd} Q(z^n T)$, since both sides are monic polynomials with the same zeros. Thus the coefficients of $P(T)$ extend to these punctures as well, yielding meromorphic functions on all of $Y$.

In other words, we have shown there is a degree $d$ polynomial $P \in \mathcal{M}(Y)[T]$ such that $P(f) = 0$. It follows that $\deg(\mathcal{M}(Y)/\mathcal{M}(X))$ is at most $d$ (e.g. by the theory of the primitive element).

To see the degree is exactly $d$, choose a function $f \in \mathcal{M}(Y)$ which separates the $d$ points in a fiber over $y_0 \in Y^*$. Then $P(f) = 0$ for a polynomial of degree $d$ as above. If $P$ factors $Q(T)R(T)$, then $Q(f)R(f) = 0$ in $\mathcal{M}(Y)$. Since $X$ is connected, $\mathcal{M}(X)$ is a field and one of these terms vanishes identically — say $Q(f) = 0$. Then the values of $f$ on the fiber over $y_0$ are zeros of the polynomial $Q_{y_0}(T)$. Since $\deg Q < d$, this is a contradiction. ∎

**Basic point: pushforwards.** The basic point in the preceding argument is that if $f \in \mathcal{M}(X)$ and $\pi : X \to Y$ is proper, then the *pushforward* $\pi_*(f)$ can be defined and lies in $\mathcal{M}(Y)$. The pushfoward is just the sum of $f$ over the fibers of $\pi$, i.e. it is the 'trace', with multiplicities taken into account:

$$\pi_*(f)(y) = \sum_{\pi(x)=y} \operatorname{mult}(\pi, x) f(x).$$

Then $\pi_*(f)$ gives the first symmetric function of $f$, and all the symmetric functions can be expressed in terms of $\pi_*(f^k)$, $k = 1, 2, \ldots, d$.

**From a field extension to a branched cover.** We now turn to the proof of Theorem 4.2. That is, starting from a finite extension field $K/\mathcal{M}(Y)$ we will construct a Riemann surface and a proper map $\pi : X \to Y$, such that $\mathcal{M}(X) \cong K$ over $\mathcal{M}(Y)$. The pair $(X, \pi)$ will be uniquely determined, up to isomorphism over $Y$.

To see there must be something interesting in this construction, we can ask: what will the branch locus $B = B(\pi)$ will be? The field $K$ must implicitly know the answer to this question.

Our first approach to the construction of $X$ will be related to analytic continuation and sheaves.

**Local algebraic functions.** We first work locally. Let

$$P(T, z) = T^d + c_1(z)T^{d-1} + \cdots + c_d(z)$$

be a monic polynomial with coefficients in $\mathcal{M}(Y)$. Then for each $a \in Y$ outside the poles of the $c_i$, we get two objects:

$$P_a(T) \in \mathcal{O}_a[T] \quad \text{and} \quad P(T, a) \in \mathbb{C}[T].$$

We will see the behavior of the second influences the factorization of the first.

**Theorem 4.4** *If $P(T, a)$ has simple zeros, then there exist $f_i \in \mathcal{O}_a$ such that $P(T) = \prod_1^d (T - f_i)$.*

**Proof.** This is simply the statement that the zeros of $P(T, z)$ vary holomorphically as its coefficients do. To make this precise, let $w_1, \ldots, w_d$ be the zeros of $P(T, a)$, and let $\gamma_1, \ldots, \gamma_d$ be the boundaries of disjoint disks in $\mathbb{C}$, centered at these zeros. By continuity, there is a connected neighborhood $U$ of $a$ such that $P$ has no zeros of $\bigcup \gamma_i$. For $z \in U$ we define

$$N_i(z) = \frac{1}{2\pi i} \int_{\gamma_i} \frac{P'(t, z)\, dt}{P(t, z)}$$

and

$$f_i(z) = \frac{1}{2\pi i} \int_{\gamma_i} \frac{tP'(t,z)\,dt}{P(t,z)},$$

where $P' = dP/dt$. Then $N_i(z)$ is the number of zeros of $P(T,z)$ enclosed by $\gamma_i$. But $N_i(a) = 1$ and $N_i(z)$ is continuous, so $N_i(z) = 1$ for all $z \in U$. It follows that $f_i(a) = w_i$ and $f_i(z)$ is the unique zero of $P(T,z)$ inside $\gamma_i$. Since the loops $\gamma_i$ enclosed disjoint disks, the values $(f_1(z), \ldots, f_d(z))$ are distinct, and hence $P(T) = \prod(T - f_i)$. ■

**Resultant and discriminant.** The preceding results shows the zeros of $P(T)$ behave well when they are all simple. To analyze this condition, it is useful to recall the *resultant* $R(f,g)$.

Let $K$ be a field, and let $f, g \in K[x]$ be nonzero polynomials with $\deg(f) = d$ and $\deg(g) = e$. Recall that $K[x]$ is a PID and hence a UFD. We wish to determine if $f$ and $g$ have a common factor, say $h$, of degree 1 or more. In this case $f = hf_1$, $g = hg_1$ and hence $g_1 f - f_1 g = 0$. Conversely, if we can find nonzero $r$ and $s$ with $\deg(r) < e$ and $\deg(s) < d$ such that $rf + sg = 0$, then $f$ and $g$ have a common factor.

The existence of such $r$ and $s$ is the same as a linear relation among the elements $(f, xf, \ldots, x^{e-1}f, g, xg, \ldots, x^{d-1}g)$, and hence it can be written as a determinant $R(f,g)$ which is simply a polynomial in the coefficients of $f$ and $g$. We have $R(f,g) = 0$ iff $f$ and $g$ have a common factor.

In general, $\gcd(f,g) = h$ can be found by the Euclidean division algorithm, and $(f,g) = (h)$ as an ideal in $K[x]$. We have $R(f,g) = 0$ iff $(f,g)$ is a proper ideal, rather than all of $K[x]$.

The *discriminant* $D(f) = R(f, f')$ is nonzero iff $f$ has simple zeros.

**Example: the cubic discriminant.** Let $f(x) = x^3 + ax + b$, so $f'(x) = 3x^2 + a$. Then $D(f) = R(f, f')$ is given by:

$$\det \begin{pmatrix} 1 & 0 & a & b & 0 \\ 0 & 1 & 0 & a & b \\ 3 & 0 & a & 0 & 0 \\ 0 & 3 & 0 & a & 0 \\ 0 & 0 & 3 & 0 & a \end{pmatrix} = 4a^3 + 27b^2.$$

The resultant $R(f,g)$ and discriminant $D(f)$ are, from this perspective, only well-defined up to sign. The usual more precise definition, $D(f) = \prod_{i<j}(x_i - x_j)^2$, gives $D(f) = -4a^3 - 27b^2$. With this sign, $K[x]/f(x)$ is Galois $\iff$ $D(f)$ is a square in $K$ iff the Galois group of the splitting field

is $\mathbb{Z}/3$ (rather than $S_3$). In general the discriminant begin a square implies the Galois group is contained in $A_n$. This is because an even permutation preserves the sign of $\prod_{i-j}(x_i - x_j) = \sqrt{D(f)}$.

**Proof of Theorem 4.2.** Let $K$ be a degree $d$ field extension of $\mathcal{M}(Y)$. By the theorem of the primitive element, there exists a monic irreducible polynomial $P(T)$ with coefficients $c_i \in \mathcal{M}(Y)$ such that $K \cong \mathcal{M}(Y)[T]/(P(T))$.

By irreducibility, the discriminant $D(P) \in \mathcal{M}(Y)$ is not identically zero. (Otherwise we could compute the gcd and $P$ and $P'$ to obtain a factor of $P$.) Let $Y^* \subset Y$ be the complement of the zeros and poles of $D(P)$, and of the poles of the coefficients $c_i$.

Let $X^* \subset |\mathcal{O}_{Y^*}|$ be the set of germs $f_a$ such that $P(f_a, a) = 0$. By the local analysis in Theorem 4.4, $X^*$ is a degree $d$ covering space of $Y^*$. The tautological map $f : X^* \to \mathbb{C}$ sending $f_a$ to $f_a(a)$ is analytic. We can complete $X^*$ to a branched covering $\pi : X \to Y$, and extend $f$ to a meromorphic function on $X$ using Riemann's removable singularities theorem.

We must show that $X$ is connected. But suppose instead $X = X_1 \sqcup X_2$. Then $f|X_i$ satisfies a unique monic polynomial $Q \in \mathcal{M}(Y)[T]$, and $P = Q_1 Q_2$. This contradicts irreducibility of $P$.

Thus $\mathcal{M}(X)$ is a field, and by construction, $P(f) = 0$. Thus $\mathcal{M}(X)$ contains a copy of $K$. But $\deg(\pi) = d$, so $\deg(\mathcal{M}(X)/\mathcal{M}(Y)) \le d$. Thus $\mathcal{M}(X) \cong K$.

It remains to consider the Galois group. Suppose $X/Y$ is Galois as a branched covering. Since $P_a(T)$ has distinct zeros for some $a \in Y$, the group $\mathrm{Deck}(X/Y)$ maps injectively into $\mathrm{Gal}(\mathcal{M}(X)/\mathcal{M}(Y))$; indeed, only the identity stabilizes the function $f$. By degree considerations, this map is surjective as well. Similarly, if $\mathcal{M}(X)/\mathcal{M}(Y)$ is Galois, then $G = \mathrm{Gal}(\mathcal{M}(X)/\mathcal{M}(Y))$ permutes the roots of the polynomial $P(T)$, and hence the germs in the factorization of $P_a(T)$. These correspond to the sheets of $X$, so we get a map $G \to \mathrm{Deck}(X/Y)$ which is an isomorphism again by degree considerations.
■

The Riemann surface $X$ can be regarded as a completion of the maximal analytic continuation of $f_a$, for any germ $f_a \in \mathcal{O}_a(Y)$ satisfying $P(f_a, a) = 0$ at a point where the discriminant of $P(T)$ is not zero.

**Construction as a curve.** An alternative to the construction of $\pi : X \to Y$ is the following. Fix $P(T) \in \mathcal{M}(Y)[T]$, irreducible, and as before let $Y^*$ be the locus where the discriminant is nonzero and where the coefficients of $P(T)$ are holomorphic. Then define

$$X^* = \{(x, y) \,:\, P(x, y) = 0\} \subset \mathbb{C} \times Y^*.$$

Let $(F, \pi)$ be projections of $X^*$ to its two coordinates. Then $\pi : X^* \to Y^*$ is a covering map, and $F : X^* \to \mathbb{C}$ is an analytic function satisfying $P(F) = 0$; and the remainder of the construction carries through as before.

**Examples: quadratic extensions.** For any polynomial $p(z) \in \mathbb{C}[z]$ which is not a square (i.e. which at least one root of odd order), we can form the Riemann surface $\pi : X \to \widehat{\mathbb{C}}$ corresponding to adjoining $f(z) = \sqrt{p(z)}$ to $K = \mathbb{C}(z) = \mathcal{M}(\widehat{\mathbb{C}})$. When $p$ has $2n$ or $2n-1$ simple zeros, the map $\pi$ is branched over $2n$ points (including infinity in the latter case), and hence $X$ has genus $n-1$. Note that the polynomial $P(T) = T^2 - p(z)$ has discriminant $\mathrm{Res}(P, P') = -4p(z)$.

**Remark: Compositums of fields.** It is well-known in field theory that if $K_1, K_2$ are two finite extensions of $k$, then there exists an extension $L$ containing both – the *compositum* of $K_1$ and $K_2$. An analogous fact holds for Riemann surfaces.

**Theorem 4.5** *Let $f_i : X_i \to \widehat{\mathbb{C}}$, $i = 1, 2$ be a pair of Riemann surfaces presented as branched coverings of $\widehat{\mathbb{C}}$. Then there is a third Riemann surface $Y$ dominating them both.*

*More precisely, there is a Riemann surface $Y$ and a pair of holomorphic maps $g_i : Y \to X_i$ such that $f_1 \circ g_1 = f_2 \circ g_2$.*

**Proof.** Let $B = B(f_1) \cup B(f_2)$. Then we can regard $X_i$ as the branched covering of $\widehat{\mathbb{C}}$ determined by a subgroup of finite index $H_i \subset \pi_1(\widehat{\mathbb{C}} - B)$. Now let $Y$ be the branched covering determined by $H_1 \cap H_2$. ■

The field $\mathcal{M}(Y)$ is a compositum of $\mathcal{M}(X_1)$ and $\mathcal{M}(X_2)$.

Note that $Y$ is not quite canonical, because we must choose basepoints to make the construction. However $Y$ can be regarded as an irreducible component of the fiber product

$$X_1 \times_{\widehat{\mathbb{C}}} X_2 = \{(x_1, x_2) \in X_1 \times X_2 \;:\; f_1(x_1) = f_2(x_2)\},$$

which is canonical.

**Puiseux series.**

*Avec les series de Puiseux,*
*Je marche comme sur des oeufs.*
*Il s'ensuit que je les fuis*
*Comme un poltron que je suis.*
—A. Douady, 1996

(Variante: *Et me refugie dans la nuit.*)

Let $\mathcal{M}_p$ be the local field of a point $p$ on a Riemann surface; it is isomorphic to the field of convergent Laurent series $\sum_{-n}^{\infty} a_i z^i$ in a local parameter $z \in \mathcal{O}_p$ with $z(p) = 0$.

**Theorem 4.6** *Every algebraic extension of $\mathcal{M}_p$ of degree $d$ is of the form $K = \mathcal{M}_p[\zeta]$, where $\zeta^d = z$.*

**Proof.** We may assume $p = 0$ in $\mathbb{C}$. Let $P(T) \in \mathcal{M}_p[T]$ be the degree $d$ irreducible polynomial for a primitive element $f \in K$ (so $K = \mathcal{M}_0(f)$).) Then we can find an $r > 0$ such that the coefficients $c_i$ of $P(T)$ are well-defined on $\Delta(r)$, only have poles at $z = 0$, and the discriminant $D(P)|Y$ vanishes only at $z = 0$ as well.

Note that $P(T)$ remains irreducible as a polynomial in $\mathcal{M}(Y)[T]$, $Y = \Delta(r)$ Thus it defines a degree $d$ branched covering $X \to Y$, branched only over the origin $z = 0$. But there is also an obvious branched covering $\Delta(r^{1/d}) \to Y$ given by $\zeta \mapsto \zeta^d = z$. Since $\pi_1(Y^*) \cong \mathbb{Z}$ has a unique subgroup of index $d$, we have $X \cong \Delta(r^{1/d})$ over $Y$, and thus $\mathcal{M}(X) = \mathcal{M}(Y)[\zeta]$. It follows that $K = \mathcal{M}_0[\zeta]$. ∎

**Corollary 4.7** *Any irreducible, degree $d$ polynomial $P(T) = 0$ with coefficients in $\mathcal{M}_p$ has a solution of the form*

$$f(z) = \sum_{-n}^{\infty} a_i z^{i/d}.$$

In particular, any polynomial is locally 'solvable by radicals', i.e. its roots can be expressed in the form $f_i(z) = g_i(z^{1/d})$, where $g_i(z)$ is analytic. Equivalent, after the change of variables $z = \zeta^d$, the polynomial $P(T)$ factors into linear terms.

Example. For $P(T) = T^3 + T^2 - z = 0$, we have a solution

$$f(z) = z^{1/2} - \frac{z}{2} + \frac{5z^{3/2}}{8} - z^2 + \frac{231 z^{5/2}}{128} - \cdots$$

Why is this only of degree 1/2? Because the original polynomial has two distinct roots when $z = 0$: it is reducible over $\mathcal{M}_0$!

This calculation can be carried out recursively by setting $\zeta = z^{1/2}$ and observing that

$$f = \zeta\sqrt{1 - f^3/\zeta^2} = \zeta\left(1 - \frac{f^3}{2\zeta^2} - \frac{f^6}{8\zeta^4} - \cdots\right).$$

Note: it is a general fact that only the primes dividing $d$ appear in the denominators of the power series for $(1-x)^{1/d}$.

| **Riemann surfaces** | **Number fields** |
|---|---|
| $K = \mathcal{M}(\widehat{\mathbb{C}}) = \mathbb{C}(z)$ | $K = \mathbb{Q}$ |
| $A = \mathbb{C}[z]$ | $A = \mathbb{Z}$ |
| $p \in \mathbb{C}$ | $p\mathbb{Z}$ prime ideal |
| uniformizer $z_p = (z-p) \in \mathbb{C}[z]$ | uniformizer $p \in \mathbb{Z}$ |
| order of vanishing of $f(z)$ at p | power of $p$ dividing $n \in \mathbb{Z}$ |
| $A_p = \mathcal{O}_p = \{\sum_0^\infty a_n z_p^n\}$ | $A_p = \mathbb{Z}_p = \{\sum a_n p^n\}$ |
| $m_p = z_p \mathcal{O}_p$ | $m_p = p\mathbb{Z}_p$ |
| residue field $k = A_p/m_p = \mathbb{C}$ | residue field $k = A_p/m_p = \mathbb{F}_p$ |
| value $f(p)$, $f \in \mathbb{C}[z]$ | value $n \bmod p$, $n \in \mathbb{Z}$ |
| $L = \mathcal{M}(X), \pi : X \to \widehat{\mathbb{C}}$ | extension field $L/\mathbb{Q}$ |
| $B = \{f$ holomorphic on $X_0 = \pi^{-1}(\mathbb{C})\}$ | $B =$ integral closure of $A$ in $L$ |
| $P \in X_0 \ : \ \pi(P) = p$ | prime $P$ lying over $p$ |
| $B_P \cdot m_p = m_P^e$; $e =$ ramification index | |
| $k' = B_P/m_P = \mathbb{C}$ | $k' = B_P/m_P = \mathbb{F}_{p^f}$; $f =$ residue degree |

Table 3. A brief dictionary.

**Valuations on $\widehat{\mathbb{C}}$.** How does $\mathcal{M}(X)$ know about the points of $X$? The answer can be formulated in terms of valuations.

A *discrete valuation* on a field $K$ is a surjective homomorphism $v : K^* \to \mathbb{Z}$ such that

$$v(f+g) \geq \min v(f), v(g).$$

The homomorphism property means

$$v(fg) = v(f) + v(g).$$

It is also convenient to define $v(0) = +\infty$.

**Examples.**

1. For $K = \mathbb{C}(z) = \mathcal{M}(\widehat{\mathbb{C}})$ and $p \in \widehat{\mathbb{C}}$, there is a unique *point valuation* $v_p$ such that $v_p(f) = n > 0$ iff $f$ has a zero of order $n$ at $p$. It is called the *point valuation* $v_p(f)$.

2. For $K = \mathbb{Q}$ and $p$ a prime, there is a unique *p-adic valuation* $v_p$ such that $v_p(p^k m) = k$ when $\gcd(p, m) = 1$.

3. We might try to define $v_{10}(n)$ as the number of 0's in the base 10 expansion of an integer $n$. This works well additively, but not multiplicatively: $v_{10}(2 \cdot 5) = 1 \neq v_{10}(2)v_{10}(5)$.

   This is an argument for work with 'decimals' in a prime base – the most natural such base being $b = 2$.

**Case of equality.** Perhaps the most important property of a valuation is that:

$$v(f+g) = \min(v(f), v(g)) \quad \text{provided } v(f) \neq v(g). \tag{4.2}$$

For example, adding a constant cannot destroy a pole, while it always destroys a zero. To see (4.2), just note that if $v(f) > v(g)$ then:

$$v(g) = v(f+g-f) = \min(v(f+g), v(f)) = v(f+g) \geq \min(v(f), v(g)) = v(g).$$

(Here the first minimum cannot be $v(f)$ because $v(f) > v(g)$).

**Theorem 4.8** *Every valuation on $\mathcal{M}(\widehat{\mathbb{C}})$ is a point valuation.*

**Proof.** Let $v : K^* \to \mathbb{Z}$ be a valuation. We claim $v$ vanishes on the constant subfield $\mathbb{C} \subset K$. Indeed, if $f \in K^*$ has arbitrarily large roots $f^{1/n}$, then $v(f^{1/n}) = v(f)/n \to 0$ and thus $v(f) = 0$.

Next suppose $v(z-a) > 0$ for some $a$. Then, $v(z-a) > v(a-b)$ for all $b \neq a$. Thus 4.2 implies $v(z-b) = 0$ for all $b \neq 0$. Since a rational function is a product of linear terms and their reciprocals, this shows $v$ is a multiple of the point valuation at $a$. But $v$ maps surjectively to $\mathbb{Z}$, so we must have $v(z-a) = 1$ and hence $v = v_a$.

Now suppose $v(a) \leq 0$ for all $a$. Since $v$ is surjective, $v(z-a) < 0$ for some $a$. But then $v(z-b) = v(z-a)$ for all $b$, by (4.2) again. It follows that $v(z-a) = -1$ for all $a$ and hence $v = v_\infty$ gives the order of vanishing at infinity. ∎

**Valuations on Riemann surfaces.** Using the Riemann existence theorem, it can be shown that the valuations on $\mathcal{M}(X)$ correspond to the points of $X$, for any compact Riemann surface $X$. This provides a natural way to associate $X$ canonically to the abstract field ($\mathbb{C}$-algebra) $\mathcal{M}(X)$. By pushing valuations forward, we can then associate to a finite field extension $\mathcal{M}(Y) \subset \mathcal{M}(X)$ a canonical map $f : X \to Y$, which turns out to be holomorphic. Of course we have already seen how to do this directly, by writing $\mathcal{M}(X) = \mathcal{M}(Y)[T]/P(T)$, looking at the discriminant of $P(T)$, etc.

The general theory of valuations on a field $K$ is facilities by the following observations:

1. The elements with $v(x) \geq 0$ form an integral domain $A \subset K$.
2. The domain $A$ is integrally closed. Indeed, if $f \in K$ and $f^n + a_1 f^{n-1} + \cdots + a_{n-1} f = a_n$ with $a_i \in A$, and $v(f) < 0$; then $v(f^n)$ has the smallest valuation in this sum, and thus $v(f^n) = v(a_n) < 0$, a contradiction.
3. The elements with $v(x) \geq 1$ form a prime ideal $I \subset A$.

Now in the case of a compact Riemann surface $X \to \widehat{\mathbb{C}}$, a valuation $v$ on $\mathcal{M}(X)$ restricts to one on $\mathbb{C}(x)$, so it lies over a point evaluation $v_p$ for $\widehat{\mathbb{C}}$; and then the corresponding prime ideal must come from one of the points on $X$ lying over $p$.

**Local rings and fields.** For a summary of connection, see Table 3. Example: Let $L = \mathbb{Q}(\sqrt{D})$, where $D \in \mathbb{Z}$ is square-free. Then (forgetting the prime 2) we have *ramification* over the primes $p|D$. At these primes we have $B_P \cong \mathbb{Z}_p[p^{1/2}]$, just as for Puiseux series.

You might think there are '2 points' lying above primes $p$ which do not divide $D$. In fact, for such primes, either:

(i) $T^2 - D$ is irreducible $\bmod p$, and there is a unique $P$ over $p$;
or
(ii) $D$ is a square mod $p$, and there are two primes $P_1$ and $P_2$ over $p$.

In case (i) the prime $p$ behaves like a circle (closed string?) rather than a point, and $P$ like a double cover of $p$. These cases are distinguished by the Legendre symbol and each occurs half the time on average.

In general when $P/p$ has ramification index $e$ and residue degree $f$, it can be thought of roughly as modeled on the map $(z, w) \mapsto (z^e, w^f)$ of $\Delta \times S^1$ to itself.

## 5 Holomorphic and harmonic forms

In this section we discuss differential forms on Riemann surfaces, from the perspective of sheaves and stalks. A special role is played by the holomorphic 1-forms, and more generally the harmonic 1-forms, which can exist in abundance even on a compact Riemann surface (where harmonic functions do not).

**Local rings again.** Let $p \in X$ be a point on a Riemann surface, and let $m_p \subset \mathcal{C}_p^\infty$ be the ideal of smooth, complex-valued functions vanishing at $p$. Let $m_p^2$ be the ideal generated by products of element in $m_p$.

**Theorem 5.1** *The ideal $m_p^2$ consists exactly of the smooth functions all of whose derivatives vanish at $p$.*

**Proof.** (Cf. [Hel, p.10]). Clearly the derivatives of $fg$, $f, g \in m_p$, all vanish.

For the converse, let us work locally at $p = 0 \in \mathbb{C}$, and suppose $f \in m_0$. Let $f_1$ and $f_2$ denote the partial derivatives of $f$ with respect to $x$ and $y$. Then we have:

$$\begin{aligned} f(x,y) &= \int_0^1 \frac{d}{dt} f(tx,ty)\,dt = x\int_0^1 f_1(tx,ty)\,dt + y\int_0^1 f_2(tx,ty)\,dt \\ &= xg_1(x,y) + yg_2(x,y) \end{aligned}$$

where $g_i(0,0) = f_i(0,0)$. Thus $g_1, g_2 \in m_p$ if both derivatives of $f$ vanish, which implies $f \in m_p^2$. ∎

**Corollary 5.2** *The vector space $T_p^{(1)} = m_p/m_p^2$ is isomorphic to $\mathbb{C}^2$.*

**Comparison between $C^\infty$ and $\mathcal{O}$.** Forster skirts this point (p.60) by *defining* $m_p^2$ to consist of the functions vanishing to order 2. Note that for the local ring $\mathcal{O}_p$, it is much easier to analyze $m_p$ because it is principle — $m_p = (z_p)$, $m_p^2 = (z_p^2)$, etc. For $C_p^\infty$, on the other hand, $m_p^2$ is not principal, because the *zero loci* $Z(f)$ for smooth $f \in m_p^2$ are not all the same!

**Tangent and cotangent spaces.** We refer to $T_p^{(1)}$ as the *complexified cotangent space* of the real surface $X$ at $p$. The exterior differential of a function can then be defined as the map

$$(df)_p = [f(z) - f(p)] \in m_p/m_p^2.$$

The subspace $d\mathcal{O}_p$ is $T_p^{(1,0)} = T_p^* X \cong \mathbb{C}$, the *complex* (or holomorphic) tangent space. Its complex conjugate is $T^{(0,1)}$. In a local holomorphic coordinate $z = x + iy$, we have:

$$T_p^{(1)} = \mathbb{C} \cdot dx \oplus \mathbb{C} \cdot dy = \mathbb{C} \cdot dz \oplus \mathbb{C} \cdot d\overline{z} = T_p^{(1,0)} \oplus T_p^{(0,1)}.$$

The dual of $T_p^*$ is $T_p$, the complex tangent space to $X$ at $p$.

Note that $TX$ and $T^*X$ are *complex* line bundles which naturally isomorphic to the 2-dimensional *real* vector bundles given by the tangent and cotangent bundles. In other words, the complex structure on $X$ turns these real vector spaces into complex vector spaces, and these smooth vector bundles into holomorphic line bundles.

As for $T^{(1)}X$, it is isomorphic to $T^*X \otimes_{\mathbb{R}} \mathbb{C}$.

**The exterior algebra of forms.** A *smooth* 1-form $\alpha$ on a Riemann surface is a smooth section of the vector bundle $T^{(1)}(X)$. Each 1-form splits naturally into its $(1,0)$ and $(0,1)$ parts, given locally by

$$\alpha = f(z)\, dz + g(z)\, d\overline{z},$$

where $f$ and $g$ are smooth functions.

One we have exterior $d$ on functions, we can form the deRham complex by taking exterior products of forms, and extending $d$ so $d^2 = 0$ and the Leibniz rule holds:

$$d(\sum f_i d\alpha_i) = \sum (df_i) d\alpha_i,$$

and more generally

$$d(\alpha\beta) = (d\alpha)\beta + (-1)^{\deg \alpha} \alpha)\, d\beta.$$

The 2-forms on $X$ are locally of the form

$$\beta = f(x,y) dz\, d\overline{z},$$

i.e. they are $(1,1)$ forms. Note that $dx\, dy = (i/2) dz\, d\overline{z} > 0$, in the sense that the integral of this form over any region is positive.

Stokes' theorem gives that for any compact, smoothly bounded region $\overline{\Omega} \subset X$, and any smooth 1-form $\alpha$, we have

$$\int_\Omega d\alpha = \int_{\partial\Omega} \alpha.$$

For a general smooth manifold, we can now define the *deRham cohomology* groups by

$$H^k_{DR}(M) = (\text{closed } k\text{-forms}) / d(\text{arbitrary } (k-1) \text{ forms}) .$$

**Integrals of metrics.** We will also meet metrics of the form $\rho(z)|dz|$, and their associated area forms $\rho^2 = \rho(z)^2|dz|^2$. These can also be integrated along paths and regions. If $\gamma : [0,1] \to X$ is a smooth path, then

$$\int_\gamma \rho = \int_0^1 \rho(\gamma(t))|\gamma'(t)|\,dt,$$

and if $\Omega \subset X$ is a region with coordinate $z$, then

$$\int_\Omega \rho^2 = \int_\Omega \rho(z)^2\,dx\,dy = \frac{i}{2}\int_\Omega \rho(z)^2\,dz \wedge d\overline{z}.$$

Note that a metric transforms by

$$f^*(\rho) = \rho(f(z))\,|f'(z)|\,|dz|.$$

**The *d*-bar operator.** On a Riemann surface the exterior derivative has a natural splitting $d = \partial + \overline{\partial}$.

On functions, this gives the splitting of $df$ into its $(1,0)$ and $(0,1)$ parts:

$$df = \partial f + \overline{\partial} f = \frac{df}{dz}\,dz + \frac{df}{d\overline{z}}\,d\overline{z}.$$

Here, if $z = x + iy$, we have

$$\frac{df}{dz} = \frac{1}{2}\left(\frac{df}{dx} - i\frac{df}{dx}\right) \quad \text{and} \quad \frac{df}{d\overline{z}} = \frac{1}{2}\left(\frac{df}{dx} + i\frac{df}{dx}\right).$$

Note that $df/d\overline{z} = 0$ iff $f$ is holomorphic (by the Cauchy–Riemann equations), in which case $f'(z) = df/dz$.

This splitting of $df$ is related to the factor that a real-linear map $T : \mathbb{C} \to \mathbb{C}$ can be written canonically in the form

$$T(z) = az + b\overline{z},$$

with $a, b \in \mathbb{C}$. For $T = Df$, we have $a = df/dz$ and $b = df/d\overline{z}$. Intrinsically, we have a real-linear map $df : T_pX \to \mathbb{C}$, and

$$df(v) = (\partial f)(v) + (\overline{\partial} f)(v)$$

where these two pieces are complex–linear and conjugate–linear respectively.

Note that

$$d(f\,dz) = (\overline{\partial} f + \overline{\partial} f)\,dz = (\overline{\partial} f)\,dz,$$

since $dz \wedge dz = 0$. Similarly, a 1-form $\alpha = f\,dz + g\,d\overline{z}$ satisfies

$$d\alpha = (\overline{\partial} f)\,dz + (\partial g)\,d\overline{z} = \left(-\frac{df}{d\overline{z}} + \frac{dg}{dz}\right) dz\,d\overline{z}.$$

**Holomorphic functions.** A function $f : X \to \mathbb{C}$ is *holomorphic* if $\overline{\partial} f = 0$. Equivalently, $df = \partial f$.

**Holomorphic forms.** A 1-form $\alpha$ on $X$ is *holomorphic* if it has type $(1, 0)$, and the following equivalent conditions hold:

1. $\overline{\partial}\alpha = 0$.
2. $d\alpha = 0$; i.e. $\alpha$ is closed.
3. Locally $\alpha = df$ with $f$ holomorphic.
4. Locally $\alpha = \alpha(z)\,dz$ with $\alpha(z)$ holomorphic.

The space of all such forms on $X$ is denoted by $\Omega(X)$.

Examples: $\Omega(\widehat{\mathbb{C}}) = 0$; $\Omega(\mathbb{C}/\Lambda) = \mathbb{C} \cdot dz$; $dz/z \in \Omega(\mathbb{C}^*)$.

**Cauchy's theorem.** Since a holomorphic form is *closed*, it integrates to zero over any contractible loop in $X$. In particular we have Cauchy's formula in the plane:

$$\int_\gamma f(z)\,dz = 0,$$

if $\gamma$ is contractible in the domain of $f$. Note that this really a theorem about the *1-form* $f(z)\,dz$.

**Harmonic functions.** A complex-valued function $u : X \to \mathbb{C}$ is *harmonic* if the following equivalent conditions hold:

1. $\Delta u = d^2u/dx^2 + d^2u/dy^2 = 0$ in local coordinates $z = x + iy$.
2. $\overline{\partial}\partial u = 0$.
3. $\partial u$ is a holomorphic 1-form.

If $u$ is real-valued, then we can add the condition:

Locally $u = \operatorname{Re} f$ for some holomorphic function $f$.

Indeed, $f$ can be found by integrating $\partial u$ to obtain a holomorphic function with $\partial f = \partial u$; then $\overline{\partial} u = \overline{\partial f}$ (since $u$ is real), and hence $du = d(f + \overline{f})$ which implies $u = 2\operatorname{Re}(f)$ up to a constant.

**Harmonic 1-forms.** A 1-form $\omega$ on $X$ is *harmonic* if the following equivalent conditions hold:

1. Locally $\omega = du$ with $u$ harmonic;
2. $\partial\omega = \overline{\partial}\omega = 0$ (in particular, $\omega$ is closed);
3. There exist $\alpha, \beta \in \Omega(X)$ such that $\omega = \alpha + \overline{\beta}$.

Example: $dx$ and $dy$ are harmonic 1-forms on $X = \mathbb{C}/\Lambda$.

In the last condition, $\alpha$ and $\overline{\beta}$ are simply the $(1,0)$ and $(0,1)$ parts of $\omega$. In particular, every holomorphic 1-form is harmonic. Denoting the space of all holomorphic 1-forms by $\mathcal{H}^1(X)$, this gives:

**Theorem 5.3** *On any Riemann surface the space of harmonic forms splits as:*

$$\mathcal{H}^1(X) = \Omega(X) \oplus \overline{\Omega}(X).$$

**Harmonic conjugates.** It is a classical fact that if $u$ is a real–valued harmonic function, then locally we can find a *conjugate* harmonic function $v$ such that $u + iv$ is analytic. In fact, since $\partial u$ is a holomorphic 1-form, we can find an analytic function $f$ such that

$$\partial u = df,$$

and then, since $u$ is real,

$$du = \partial u + \overline{\partial} u = d(f + \overline{f}),$$

and so $u = 2\,\mathrm{Re}(f)$ and $v = 2\,\mathrm{Im}(f)$ will work, up to an additive constant.

**Periods.** On a compact Riemann surface, all these spaces are finite–dimensional. This is a corollary of general facts about elliptic differential operators such as the Laplacian. To see it explicitly in this case at hand, we define the *period map* by associating to each 1-form $\omega$ the homomorphism

$$\phi : \pi_1(X) \to \mathbb{C}$$

given by $\gamma \mapsto \int_\gamma \omega$ on $\pi_1(X)$.

**Theorem 5.4** *If $X$ is compact, then the period map*

$$\mathcal{H}^1(X) \to \mathrm{Hom}(\pi_1(X), \mathbb{C}) \cong H^1(X, \mathbb{C}) \cong \mathbb{C}^{2g}$$

*is injective.*

**Proof.** Any form in the kernel can be expressed globally as $\omega = du$ where $u$ is harmonic — and hence constant — on the compact Riemann surface $X$. ∎

**Corollary 5.5** *If $X$ has genus $g$, then* $\dim \Omega(X) \leq g$ *and* $\dim \mathcal{H}^1(X) \leq 2g$.

In fact equality holds, but to see this will require more work.

**Example: the torus.** For $X = \mathbb{C}/\Lambda$ we have $\Omega(X) = \mathbb{C}\omega$ with $\omega = dz$, and the period map sends $\pi_1(X) \cong \mathbb{Z}^2$ to $\Lambda$.

**Example: the regular octagon.** Here is concrete example of a compact Riemann surface $X$ and a holomorphic 1-form $\omega \in \Omega(X)$. Namely let $X = Q/\equiv$ where $Q$ is a regular octagon in the plane, with vertices at the 8th roots of unity, and $\equiv$ identifies opposite edges. Since translations preserve $dz$, the form $\omega = dz/\equiv$ is well-defined on $X$. It has a zero of order two at the single point $p \in X$ coming from the vertex. The edges of $Q$ form a system of generators for $\pi_1(X)$, and the periods $\int_\alpha \omega$ are given by $\zeta^{i+1} - \zeta^i$ where $\zeta$ is a primitive 8th root of unity.

**Meromorphic forms and the residue theorem.** A *meromorphic 1-form* $\omega$ is a $(1,0)$-form with meromorphic coefficients. That is, locally we have a meromorphic function $\omega(z)$ such that $\omega = \omega(z)\,dz$.

Suppose we locally have $\omega = \sum a_n z^n \, dz$. We then define the *order* of $\omega$ at $p$ by:

$$\operatorname{ord}_p \omega = \inf\{n \,:\, a_n \neq 0\},$$

and the *residue* by

$$\operatorname{Res}_p(\omega) = a_{-1}.$$

The order of pole or zero of $\omega$ at a point $p \in X$ is easily seen to be independent of the choice of coordinate system. The invariance of the residue follows from the following equivalent definition:

$$\operatorname{Res}_p(\omega) = \int_\gamma \omega,$$

where $\gamma$ is a small loop around $p$ (inside of which the only pole of $\omega$ is at $p$). On a compact Riemann surface, the *degree* of $\omega$ is defined by

$$\deg(\omega) = \sum_{p \in X} \operatorname{ord}_p(\omega).$$

**Theorem 5.6** *If $X$ is compact, then* $\sum_p \operatorname{Res}_p(\alpha) = 0$.

By considering $df/f = d\log f$, this gives another proof of:

**Corollary 5.7** *If $f : X \to \widehat{\mathbb{C}}$ is a meromorphic function on a compact Riemann surface, then*

$$\sum_p \mathrm{ord}_p(f) = 0.$$

*In other words, $f$ has the same number of zeros as poles (counted with multiplicity).*

(The first proof was that $\sum_{f(x)=y} \mathrm{mult}(f,x) = \deg(f)$ by general considerations of proper mappings.)

Since the ratio $\omega_1/\omega_2$ of any two meromorphic 1-forms is a meromorphic function (so long as $\omega_2 \neq 0$), we have:

**Theorem 5.8** *The 'degree' of a meromorphic 1-form on a compact Riemann surface — that is, the difference between the number of zeros and the number of poles — is independent of the form.*

**Meromorphic forms on $\widehat{\mathbb{C}}$.** Meromorphic 1-forms on $\widehat{\mathbb{C}}$ have degree $-2$, i.e. they have 2 more poles than zeros. To check this, note that $dz$ has a double order pole at $z = \infty$, since $d(1/z) = -dz/z^2$. More generally, we have:

**Theorem 5.9** *Let $X$ be a compact Riemann surface of genus $g$. Then every nonzero meromorphic form on $X$ has degree $2g - 2$.*

**Proof.** This is an easy consequence of the Riemann–Hurwitz theorem. Let $f : X \to \widehat{\mathbb{C}}$ be a meromorphic function of degree $d > 0$ whose branch locus does not include $\infty$. Let $\omega = f^*(dz)$. Then over $\infty$, $\omega$ has a total of $2d$ poles. Each critical point of $f$, on the other hand, creates a simple zero. Thus

$$\deg(\omega) = \sum_{p \in X} (\mathrm{mult}_p(f) - 1) - 2\deg(f) = -\chi(X) = 2g - 2.$$

■

**Example: hyperelliptic Riemann surfaces.** We will eventually show that $\dim \Omega(X) = g$. Let us first check this for the important case of 2-fold branched covers of the sphere.

**Theorem 5.10** *Let $B \subset \mathbb{C}$ be a finite set of cardinality $2n$, and let $p(T) = \prod_B (T-b)$. Let $\pi : X \to \widehat{\mathbb{C}}$ be the unique 2-fold covering of $\widehat{\mathbb{C}}$ branched over $B$. Then $X$ has genus $g = n-1$, and the forms*

$$\omega_i = \frac{z^i\, dz}{\sqrt{p(z)}}, \quad i = 0, 1, \ldots, n-2$$

*form a basis for $\Omega(X)$. In particular* $\dim \Omega(X) = g$.

**Proof.** To check this, we begin by investigating $\omega_0$. Note that if $z = w^2$ then $dz = 2w\, dw$. Thus the pullback of $dz$ to $X$ has simple zeros at the branch points of $\pi$, which lie over the zeros of $p(z)$. Also $\sqrt{p(z)}$, as a meromorphic function on $X$, has simple zeros in the same locations. These zeros cancel when we form the quotient $dz/\sqrt{p(z)}$, and thus $\omega_0$ is holomorphic except possibly over the two unbranched points $p_1, p_2 \in X$ lying over $z = \infty$. But now $dz$ has a pole of order 2 at $z = \infty$, while $1/\sqrt{p(z)}$ has a zero of order $n$ at $z = \infty$. We conclude:

> *The form $\omega_0$ is holomorphic on $X$, with zeros of order $n-2$ at $p_1$ and $p_2$ and nowhere else.*

See Table 4. In particular, the degree of a meromorphic 1-form on $X$ is $2n - 4 = 2g - 2$. Since $z^i$ has a pole of order $i$ at $z = \infty$, it follows that the $g$ forms $\omega_i$ on $X$ are holomorphic for $i \leq n-2$. ∎

| | $z^k$ | $dz$ | $1/\sqrt{p}$ |
|---|---|---|---|
| $\widetilde{B}$ | $\geq 0$ | $+1$ | $-1$ |
| $\infty$ | $-k$ | $-2$ | $+n$ |

Table 4. Singularities on a hyperelliptic surface.

**Note: periods and hyperelliptic integrals.** When $B = \{r_1, \ldots, r_{2n}\} \subset \mathbb{R}$, the periods of $\omega_0$ can be expressed in terms of the integrals

$$\int_{r_i}^{r_{i+1}} \frac{dx}{\sqrt{(x-r_1)\cdots(x-r_{2n})}}.$$

We can also allow one point of $B$ to become infinity. An example of a period that can be determined explicitly comes from the square torus:

$$\int_0^1 \frac{dx}{\sqrt{x(1-x^2)}} = \frac{2\sqrt{\pi}\,\Gamma(5/4)}{\Gamma(3/4)}.$$

A more general type of period is:

$$\zeta(3) = \sum n^{-3} = \int_{0<x<y<z<1} \frac{dx\,dy\,dz}{(1-x)yz}.$$

See [KZ] for much more on periods.

**Hodge theory on Riemann surfaces.** The Laplacian and the notion of harmonic functions and forms are actually elements of the differential geometry of Riemannian manifolds. The part of discussion which survives on a Riemann surface (without a chosen metric) is the *Hodge star* operator on 1-forms, $\omega \mapsto *\omega$; and the associated Hermitian inner product.

**Hodge star.** The Hodge $*$-operator is a real–linear map

$$* : T_p^{(1)}X \to T_p^{(1)}X$$

for each $p \in X$, satisfying

$$*^2 = -1.$$

In local coordinates $z = x + iy$, it is given by

$$*dx = dy \quad \text{and} \quad *\,dy = -dx.$$

In other words, $*\omega$ 'rotates' $\omega$ by 90 degrees.

We adopt the convention (usual for complex manifolds) that $*$ is conjugate linear, meaning

$$*(a\omega) = \overline{a}(*\omega).$$

Thus an equivalent local definition of the Hodge star operator is

$$*dz = i\,d\overline{z} \quad \text{and} \quad *\,d\overline{z} = -i\,dz.$$

By applying the Hodge $*$ pointwise, we obtain an operator on 1-forms. If we write a general 1-form as

$$\omega = \alpha + \overline{\beta},$$

where $\alpha, \beta$ are $(1,0)$ forms, then

$$*\omega = *(\alpha + \overline{\beta}) = i(\overline{\alpha} - \beta).$$

One can check that

$$\omega_1 \wedge *\omega_2 = \overline{\omega_2 \wedge *\omega_1}.$$

In other words, the product is Hermitian.

**Positivity.** It is useful to measure the size of a 1-form by the natural 2-form

$$|\omega|^2_{\mathbb{R}} = \omega \wedge *\omega.$$

This 'real norm' satisfies

$$|a\,dx + b\,dy|^2_{\mathbb{R}} = (|a|^2 + |b|^2)\,dx\,dy.$$

This normalization is slightly unnatural for complex analysis, since

$$|dz|^2_{\mathbb{R}} = 2\,dx\,dy.$$

As an alternative, we have the 'complex norm', defined by the measure

$$|a\,dz + \,b\,d\overline{z}|^2_{\mathbb{C}} = (|a|^2 + |b|^2)\,|dz|^2.$$

The measure $|dz|^2$ coincides with that induced by $dx\,dy$ together with the canonical orientation of $X$. With this identification, we have

$$|\omega|^2_{\mathbb{R}} = 2|\omega|^2_{\mathbb{C}}.$$

Both these (real) 2-forms are *positive*, in the sense that they are non-negative multiples of the orientation form on $X$, or equivalently that they integrate to non-negative quantities.

There is a natural *real linear* isomorphism between the real cotangent space, $T^*_p$, and the complex space $T^{(1,0)}_p$, sending $dx$ to $dz$ and $dy$ to $-i\,dz$. This map (whose inverse is $\alpha \mapsto \mathrm{Re}(\alpha)$) is an *isometry* from the real norm to the complex norm.

**Hodge norm and inner product.** Now assume $X$ is compact. We can then define a Hermitian inner product on the space of smooth 1-forms by

$$\langle \omega_1, \omega_2 \rangle = \int_X \omega_1 \wedge *\omega_2.$$

The associated $L^2$-norm is

$$\|\omega\|^2 = \langle \omega, \omega \rangle = \int_X |\omega|^2_{\mathbb{R}}.$$

Note that $\Omega(X)$ and $\overline{\Omega}(X)$ are *orthogonal* with respect to the Hodge inner product. On the space $\Omega(X)$ there is a second commonly used inner product,

$$\{\alpha, \beta\} = \frac{i}{2} \int_X \alpha \wedge \overline{\beta}.$$

It satisfies

$$\{\alpha, \alpha\} = \int_X |\alpha|^2_{\mathbb{C}} = \frac{1}{2}\langle \alpha, \alpha \rangle.$$

**Harmonic 1-forms.** We can now give a more intrinsic definition of a harmonic 1-form, that generalizes to arbitrary Riemannian manifolds.

From the formulas above, it is evident that $d * \omega$ is a linear combination of $\partial\overline{\omega}$ and $\overline{\partial}\overline{\omega}$, linear independent from $d\omega$, and thus

$$(d\omega = d * \omega = 0) \iff (\partial\omega = \overline{\partial}\omega = 0) \iff (\omega \text{ is harmonic}).$$

Now suppose $X$ is compact and we have a deRham cohomology class $[\omega] \in H^1(X)$. We wish to choose the representative $\omega$ to *minimize* $\|\omega\|^2$ in its cohomology class. Formally, this means $\omega$ must be perpendicular to the exact forms, and thus

$$\langle df, \omega \rangle = \int_X df \wedge *\omega = 0$$

for all smooth functions $f$. Integrating by parts, we see this is equivalent to the condition that $\int f d * \omega = 0$ for all $f$, which just says that $d * \omega = 0$ — the form $\omega$ is co-closed. This shows:

**Theorem 5.11** *A 1-form on a compact Riemann surface is harmonic iff it is closed and (formally) norm-minimizing in its cohomology class.*

The general theory of elliptic operators insures that this minimum is actual achieved at a smooth form, and thus $\mathcal{H}^1(X) \cong H^1(X, \mathbb{C})$. This is the first instance of the *Hodge theorem.* We will soon prove it, for Riemann surfaces, using sheaf cohomology.

**The Laplacian.** The Hodge star operator also provides any Riemann surface with a *natural* Laplacian

$$\Delta : (\text{function on } X) \to (\text{2-forms on } X).$$

This is consistent with electromagnetism, where $\mu = \Delta\phi$ should be a *measure* giving the charge density for the static electric field potential $\phi$. This Laplacian is defined by

$$\Delta f = d * df;$$

it satisfies

$$\Delta f = \left( \frac{d^2 f}{dx^2} + \frac{d^2 f}{dy^2} \right) dx\, dy$$

in local coordinates $z = x + iy$. Integrating by parts, we find the basic identity

$$\int_X f\,\Delta f = \int_X f(d * df) = -\int_X df \wedge *df = -\int_X \|df\|^2_{\mathbb{R}}.$$

Note that in local coordinates we have

$$|df|^2_{\mathbb{R}} = \left(\left(\frac{df}{dx}\right)^2 + \left(\frac{df}{dy}\right)^2\right) dx\,dy.$$

Thus we obtain an intrinsic form of the familiar formula for compactly supported functions in Euclidean space,

$$\int_{\mathbb{R}^n} f\,\Delta f = \int_{\mathbb{R}^n} |\nabla f|^2.$$

**Complex Hessian.** The Laplacian from functions to 2-forms can also be thought of naturally as giving the complex Hessian of $f$: namely, we have

$$\Delta f = 2i\partial\overline{\partial} f = 4\frac{d^2 f}{dz\,d\overline{z}} \cdot \frac{i}{2} dz\,d\overline{z}.$$

**Curvature.** We remark that, given a conformal metric $\rho = \rho(z)\,|dz|$, the 2-form $R(\rho) = \Delta \log \rho$ is naturally the Ricci curvature of $\rho$, and its ratio to the volume form of $\rho$ itself gives the Gaussian curvature:

$$K(\rho) = -\frac{\Delta \log \rho}{\rho^2}\cdot$$

For example: the curvature of the metric on $\mathbb{C}^*$ given by

$$\rho_\alpha = |z|^\alpha\,|dz|$$

is zero for all $\alpha \in \mathbb{R}$. Note that $\rho_{-1}$ makes $\mathbb{C}^*$ into a (complete) cylinder; other values near $-1$ make it into a (incomplete) cone.

To see this metric is flat, just note that $\log|z| = \operatorname{Re}\log(z)$ is harmonic. In fact, we can locally change coordinates so that $\rho_\alpha$ becomes $|dz|$: for $f(z) = z^{\alpha+1}/(\alpha+1)$, we have

$$f^*(|dz|) = |f'(z)|\,|dz| = \rho_\alpha.$$

For $\alpha = -1$ we can take $f(z) = \log z$.

**Spectrum of the Laplacian.** When we have a metric $\rho$ on $X$, we can divide by the volume form $\rho^2$ to get a well-defined Laplace operator $\Delta_\rho$ from functions to functions. It is usually normalized so its spectrum is positive; then, it is defined locally by:

$$\Delta_\rho(f) = -\frac{\Delta f}{\rho^2}.$$

For example, the function $f(y) = y^\alpha$ on $\mathbb{H}$ is an *eigenfunction* of the hyperbolic Laplacian (for the metric $\rho = |dz|/y$), since

$$\Delta_\rho(y^\alpha) = \frac{\alpha(1-\alpha)y^{\alpha-2}}{y^{-2}} = \alpha(1-\alpha)y^\alpha.$$

Note that the curvature is simple the $\rho$-Laplacian of $\log \rho$:

$$K(\rho) = \Delta_\rho(\log \rho).$$

**1-forms and foliations.** Geometrically, a (locally nonzero) 1-form defines a foliation $\mathcal{F}$ tangent to $\operatorname{Ker} \alpha$, and a measure on transversals given by $m(\tau) = \int_\tau \alpha$. We have $d\alpha = 0$ iff $m$ is a transverse *invariant* measure, meaning $m(\tau)$ is unchanged if we slide it along the leaves of $\mathcal{F}$. The transverse measure on the orthogonal foliation, defined by $*\alpha$, is *also* invariant iff $\alpha$ is harmonic.

**Example.** The level sets of $\operatorname{Re} f(z)$ and $\operatorname{Im} f(z)$ give the foliations associated to the form $\alpha = du$, $u = \operatorname{Re} f(z)$. The case $f(z) = z + 1/z$ in particular gives foliations by confocal ellipses and hyperboli, with foci $\pm 1$ coming from the critical points of $f$.

**Appendix: The Hodge star operator on Riemannian manifolds.** Let $V$ be an $n$-dimensional real vector space with an inner product $\langle v_1, v_2 \rangle$. Choose an orthonormal basis $e_1, \ldots, e_n$ for $V$. Then the wedge products $e_I = e_{i_1} \wedge \cdots e_{i_k}$ provide an orthonormal basis for $\bigwedge^k V$. The *Hodge star* operator $* : \bigwedge^k V \to \bigwedge^{n-k} V$ is the unique linear map satisfying

$$*e_I = e_J \quad \text{where} \quad e_I \wedge *e_I = e_1 \wedge e_2 \wedge \cdots e_n.$$

Here $J$ are the indices not occurring in $I$, ordered so the second equation holds. More generally we have:

$$v \wedge *w = \langle v, w \rangle \, e_1 \wedge e_2 \wedge \cdots e_n,$$

and thus $v \wedge *w = w \wedge *v$. Since $v \wedge *v = (-1)^{k(n-k)} *v \wedge v$, we have $*^2 = (-1)^{k(n-k)}$ on $\bigwedge^k V$. Equivalently, $*^2 = (-1)^k$ when $n$ is even, and $*^2 = 1$ when $n$ is odd.

Now let $(M, g)$ be a compact Riemannian manifold. We can then try to represent each cohomology class by a closed form minimizing $\int_M \langle \alpha, \alpha \rangle$. Formally this minimization property implies:

$$d\alpha = d * \alpha = 0, \tag{5.1}$$

using the fact that a minimizer satisfies

$$\int_M \langle d\beta, \alpha \rangle = \int_M (d\beta) \wedge *\alpha = -\int_M \beta \wedge d * \alpha = 0$$

for all smooth $(k-1)$-forms $\beta$. Thus we call $\alpha$ *harmonic* if (5.1) holds, and let $\mathcal{H}^k(M)$ denote the space of all harmonic $k$-forms.

**Theorem 5.12 (Hodge)** *There is a natural isomorphism $\mathcal{H}^k(M) \cong H^k_{DR}(M)$.*

**Adjoints.** The *adjoint differential* $d^* : \mathcal{E}^k(M) \to \mathcal{E}^{k-1}(M)$ is defined so that:

$$\langle d\alpha, \beta \rangle = \langle \alpha, d^*\beta \rangle,$$

where $\langle \alpha, \beta \rangle = \int_M \alpha \wedge *\beta$. It is given by

$$d^*(\alpha) = \pm * d * \alpha$$

for a suitable choice of sign, since:

$$\langle d\alpha, \beta \rangle = \int d\alpha \wedge *\beta = -\int \alpha \wedge d * \beta = \pm \int \alpha \wedge *(*d * \beta).$$

Since $*^2 = 1$ on an odd-dimensional manifold, in that case we have $d^* = - * d *$. For a $k$-form $\beta$ on an even dimensional manifold, we have instead:

$$d^*\beta = (-1)^k * d * \beta.$$

Here we have used the fact that $*^2 = (-1)^{k-1}$ on $n-(k-1)$ forms such as $d * \beta$.

**Generalized Hodge theorem.** Once the adjoint $d^*$ is in play, the arguments of the Hodge theorem give a complete picture of *all* smooth $k$-forms on $M$.

**Theorem 5.13** *The space of smooth $k$-forms has an orthogonal splitting:*

$$\mathcal{E}^k(M) = d\mathcal{E}^{k-1}(M) \oplus \mathcal{H}^k(M) \oplus d^*\mathcal{E}^{k+1}(M).$$

**The Laplacian.** Once we have a metric we can combine $d$ and $d^*$ to obtain the *Hodge Laplacian*

$$\Delta : \mathcal{E}^k(M) \to \mathcal{E}^k(M),$$

defined by

$$\Delta\alpha = (dd^* + d^*d)\alpha.$$

**Theorem 5.14** *A form $\alpha$ on a compact manifold is harmonic iff $\Delta\alpha = 0$.*

**Proof.** Clearly $d\alpha = d*\alpha = 0$ implies $\Delta\alpha = 0$. Conversely if $\Delta\alpha = 0$ then:

$$\begin{aligned} 0 &= \int_M \langle \Delta\alpha, \alpha \rangle \\ &= \int_M \langle d\alpha, d\alpha \rangle + \langle d^*\alpha, d^*\alpha \rangle \end{aligned}$$

and so $d\alpha = d*\alpha = 0$ as well. ∎

Note that on functions these definitions give

$$\Delta f = d^*df = - * d * df,$$

independent of the dimension of $M$. This satisfies $\int \langle f, \Delta f \rangle \geq 0$, but differs by a sign from the usual Euclidean Laplacian. (For example on $S^1$ we have $\int f f'' = -\int |f'|^2 \leq 0$ for the usual Laplacian.)

**Back to Riemann surfaces.** Now suppose $M$ has even dimension $n = 2k$. The Hodge star on the middle-dimensional $k$-forms is then *conformally invariant*. Thus it makes sense to talk about harmonic $k$-forms when only a conformal structure is present. Similarly, on a 4-manifold one can talk about self-dual 2-forms, satisfying $*\alpha = \alpha$. These play an important role in Yang–Mills theory.

## 6 Cohomology of sheaves

In this section we introduce sheaf cohomology, an algebraic structure that measures the global obstruction to solving equations for which local solutions are available.

**Maps of sheaves; exact sequences.** A map between sheaves is always specified at the level of open sets, by a family of compatible morphisms $\mathcal{F}(U) \to \mathcal{G}(U)$. A map of sheaves *induces* maps $\mathcal{F}_p \to \mathcal{G}_p$ between stalks. We say $\mathcal{F} \to \mathcal{G}$ is injective, surjective, an isomorphism, etc. iff $\mathcal{F}_p \to \mathcal{G}_p$ has the same property for each point $p$.

We say a sequence of sheaves $\mathcal{A} \to \mathcal{B} \to \mathcal{C}$ is *exact* at $\mathcal{B}$ if the sequence of groups

$$\mathcal{A}_p \to \mathcal{B}_p \to \mathcal{C}_p$$

is exact, for every $p$.

**The exponential sequence.** As a prime example: on any Riemann surface $X$, the sequence of sheaves

$$0 \to \mathbb{Z} \to \mathcal{O} \to \mathcal{O}^* \to 0$$

is exact. But it is only exact on the level of stalks! For every open set the sequence

$$0 \to \mathbb{Z}(U) \to \mathcal{O}(U) \to \mathcal{O}^*(U)$$

is exact, but the final arrow need not be surjective. (Consider $f(z) = z \in \mathcal{O}^*(\mathbb{C}^*)$; it cannot be written in the form $f(z) = \exp(g(z))$ with $g \in \mathcal{O}(\mathbb{C}^*)$.)

More generally, we have:

**Theorem 6.1** *The global section functor is left exact. That is, for any short exact sequence of sheaves,* $0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$, *the sequence of global sections*

$$0 \to \mathcal{A}(U) \to \mathcal{B}(U) \to \mathcal{C}(U)$$

*is also exact.*

**Proof.** We have a commutative diagram

$$\begin{array}{ccccccc} 0 & \longrightarrow & \mathcal{A}(U) & \longrightarrow & \mathcal{B}(U) & \longrightarrow & \mathcal{C}(U) \\ & & \downarrow & & \downarrow & & \downarrow \\ 0 & \longrightarrow & \prod \mathcal{A}_p & \longrightarrow & \prod \mathcal{B}_p & \longrightarrow & \prod \mathcal{C}_p \end{array}$$

where the bottom row is exact and the vertical arrows are injective. Exactness at $\mathcal{A}(U)$ now follows easily: going down and across (from $\mathcal{A}(U)$) is 1-1, so going across and down must be too.

For exactness at $\mathcal{B}(U)$, we must show that if $b \in \mathcal{B}(U)$ maps to zero in $\mathcal{C}(U)$, then it is in the image of $\mathcal{A}(U)$. This is true on the level of stalks: for every $p \in U$ there is a neighborhood $U_p$ of $p$ and an $a_p \in \mathcal{A}(U_p)$ such that $a_p|U_p \mapsto b|U_p$. Now by the first part of the proof, $\mathcal{A}(U_p) \to \mathcal{B}(U_p)$ is $1-1$, so $a_p$ is unique. But then $a_p - a_q = 0$ on $U_{pq}$, so these local solutions piece together to give an element $a \in \mathcal{A}(U)$ that maps to $b$. ∎

Sheaf cohomology is the derived functor which measures the failure of exactness to hold on the right.

**Čech theory: the nerve of a covering.** A precursor to sheaf cohomology is Čech cohomology. The idea here is that any open covering $\mathfrak{U} = (U_i)$ of $X$ has an associated simplicial complex $\Sigma(\mathfrak{U})$ that is an approximation to the topology of $X$. The simplices in this complex are simply ordered finite sequence of indices $I$ such that $\bigcap_I U_i \neq \emptyset$.

This works especially well if we require that all the multiple intersections are *connected*. Note that this is equivalent to requiring that $\mathbb{Z}(U_I) \cong \mathbb{Z}$ whenever $U_I \neq \emptyset$. Then the ordinary simplicial cohomology $H^*(\Sigma(\mathfrak{U}), \mathbb{Z})$ turns out to agree with the sheaf cohomology $H^*(\mathfrak{U}, \mathbb{Z})$.

**Cochains, cocycles and coboundaries.** When we have a general sheaf $\mathcal{F}$ in play, we can modify Čech cohomology by allowing the 'weights' on the simplex $U_I$ to take values in the group $\mathcal{F}(U_I)$. (For the usual Čech cohomology, we would take $\mathcal{F} = \mathbb{Z}$.)

With these 'weights on simplices', we define the space of $q$-cochains by:

$$C^q(\mathfrak{U}, \mathcal{F}) = \prod_{|I|=q+1} \mathcal{F}(U_I).$$

Here $I$ ranges over *ordered* sets of indices $(i_0, \ldots, i_q)$, and

$$U_I = U_{i_0} \cap \cdots \cap U_{i_q}.$$

Examples: a 0-cochain is the data $f_i \in \mathcal{F}(U_i)$; a 0-cochain is the data $g_{ij} \in \mathcal{F}(U_i \cap U_j)$; etc.

Next we define a boundary operator

$$\delta : C^q(\mathfrak{U}, \mathcal{F}) \to C^{q+1}(\mathfrak{U}, \mathcal{F})$$

by setting $\delta f = g$ where, for $q = 0$:

$$g_{ij} = f_j - f_i;$$

for $q = 1$:

$$g_{ijk} = f_{jk} - f_{ik} + f_{ij},$$

and more generally

$$g_I = \sum_0^q (-1)^j f_{I_j}$$

where $I_j = (i_0, i_1, \ldots, \widehat{i_j}, \ldots, i_{q+1})$. When two indices are eliminated, they come with opposite sign, so $\delta^2 = 0$.

The kernel of $\delta$ is the group of *cocycles* $Z^q(\mathfrak{U}, \mathcal{F})$, its image is the group of *coboundaries* $B^q(\mathfrak{U}, \mathcal{F})$, and the $q$th cohomology group of $\mathcal{F}$ relative to the covering $\mathfrak{U}$ is defined by:

$$H^q(\mathfrak{U}, \mathcal{F}) = Z^q(\mathfrak{U}, \mathcal{F})/B^q(\mathfrak{U}, \mathcal{F}).$$

**Example: $\boldsymbol{H^0}$.** A 0-cocycle $(f_i)$ is a coboundary iff $f_j - f_i = g_{ij} = 0$ for all $i$ and $j$. By the sheaf axioms, this happens iff $f_i = f|U_i$, and thus:

$$H^0(\mathfrak{U}, \mathcal{F}) = \mathcal{F}(X).$$

**Example: $\boldsymbol{H^1}$.** A 1-cocycle $g_{ij}$ satisfies $g_{ii} = 0$, $g_{ij} = -g_{ji}$ and

$$g_{ij} + g_{jk} = g_{ik}$$

on $U_{ijk}$. It is a coboundary if it can be written in the form $g_{ij} = f_i - f_j$.

**Example on $\boldsymbol{S^1}$.** Let $\mathcal{F} = \mathbb{Z}$. Suppose we cover $S^1$ with $n \geq 3$ intervals $U_1, \ldots, U_n$ such that $U_{ij}$ is an interval when $i$ and $j$ are consecutive, and otherwise $U_{ij}$ is empty. Then the space of 1-cocycles in simply $\mathbb{Z}^n$, but $g_{ij} = f_i - f_j$ iff $\sum g_{i,i+1} = 0$. Thus $H^1(\mathfrak{U}, \mathbb{Z}) = \mathbb{Z}$.

**Refinement.** Whenever $\mathfrak{V} = (V_i)$ is a finer covering than $\mathfrak{U} = (U_i)$, we can choose a refinement map on indices such that $V_i \subset U_{\rho i}$. Once $\rho$ is specified, it determines maps $\mathcal{F}(U_{\rho I}) \to \mathcal{F}(V_I)$, and hence chain maps giving rise to a homomorphism

$$H^q(\mathfrak{U}, \mathcal{F}) \to H^q(\mathfrak{V}, \mathcal{F}),$$

which sends the chain $(f_I|U_I)$ to the chain $(f_{\rho I}|V_I)$.

**Theorem 6.2** *The refinement map $H^q(\mathfrak{U}, \mathcal{F}) \to H^q(\mathfrak{V}, \mathcal{F})$ is independent of $\rho$.*

**Definition.** We define the cohomology of $X$ with coefficients in $\mathcal{F}$ by:

$$H^q(X, \mathcal{F}) = \lim_{\longrightarrow} H^q(\mathfrak{U}, \mathcal{F}),$$

where the limit is taken over the system of all open coverings, directed by refinement.

**Theorem 6.3** *The refinement map $H^1(\mathfrak{U}, \mathcal{F}) \to H^1(\mathfrak{V}, \mathcal{F})$ is injective.*

**Proof.** Suppose we are given coverings $(U_i)$ and $(V_i)$ with $V_i \subset U_{\rho i}$. Let $g_{ij}$ be a 1-cocycle for the covering $(U_i)$ that becomes trivial for $(V_i)$. That means there exist $f_i \in \mathcal{F}(V_i)$ such that

$$g_{\rho i, \rho j} = f_i - f_j$$

on $V_{ij}$.

Our goal is to find $h_k \in \mathcal{F}(U_k)$ so $g_{km} = h_k - h_m$. The first attempt is to set

$$h_k^i = f_i$$

on $U_k \cap V_i$, and hope that these patch together. But in fact we have

$$h_k^i - h_k^j = f_i - f_j = g_{\rho i, \rho j} = g_{\rho i, k} - g_{\rho j, k}$$

on $U_k \cap V_i \cap V_j$. The clever part is the last expression, which follows from the cocycle condition on $g_{ij}$. It suggests correcting our definition of $h_k$ as follows:

$$h_k = f_i - g_{\rho i, k}$$

on $U_k \cap V_i$. Now this definition agrees on overlaps, so by the sheaf axioms it gives a well–defined section $h_k \in \mathcal{F}(U_k)$.

The rest of the proof is now straightforward: by the cocycle condition again, on $U_{km} \cap V_i$, we have

$$h_k - h_m = f_i - g_{\rho i, k} - f_i + g_{\rho i, m} = g_{km}$$

as desired. ∎

**Corollary 6.4** *We have* $H^1(X, \mathcal{F}) = \bigcup H^1(X, \mathcal{U}; \mathcal{F})$.

**Intuition for $q = 1$.** To get a nontrivial element of $H^1(X, \mathcal{U}; \mathbb{Z})$, it suffices to find a covering of $X$ by *connected* open sets $U_1, \ldots, U_n$ such that $U_i$ meets $U_{i+1}$ (and $U_1$ meets $U_n$) and all other pairwise intersections are trivial. This shows $X$ is 'coarsely' a circle, and this property persists under refinements because the $U_i$ are connected.

**Failure for $q \geq 2$.** Refinement is not injective for $q \geq 2$ in general. For example, let $X$ be the 1-skeleton of a 3-simplex and let $\mathcal{U} = (U_1, \ldots, U_4)$ an open covering obtained by thickening each of its triangular 'faces'. Then the nerve of $\mathcal{U}$ is a simplicial 2-sphere and hence $H^2(X, \mathcal{U}; \mathbb{Z}) = \mathbb{Z}$ even though $H^2(X; \mathbb{Z}) = 0$.

Note that $\mathcal{U}$ has a refinement $\mathcal{V}$ with no 3-fold intersections.

**Leray coverings.** An open set $U$ is *acyclic* (for $\mathcal{F}$) if $H^q(U,\mathcal{F}) = 0$ for all $q > 0$. We say $\mathfrak{U}$ is a *Leray* covering of $(X,\mathcal{F})$ if $U_I$ is acyclic for all indices $I$.

**Theorem 6.5 (Leray)** *If $\mathfrak{U}$ is a Leray covering, then*

$$H^q(X,\mathcal{F}) \cong H^q(\mathfrak{U},\mathcal{F}).$$

*If just $H^1(U_i,\mathcal{F}) = 0$ for every $i$, $\mathfrak{U}$ can still be used to compute $H^1$.*

(Note: Forster works only with Leray covers of 'first order', i.e. those in the second part of the theorem.)

**Vanishing theorem by dimension.** Using the existence of fine coverings where the $(n+2)$-fold intersections are all empty, we have:

**Theorem 6.6** *For any $n$-dimensional space $X$ and any sheaf $\mathcal{F}$, $H^p(X,\mathcal{F}) = 0$ for all $p > n$.*

**Vanishing theorems for smooth functions, forms, etc.** Let $\mathcal{F}$ be the sheaf of $C^\infty$ functions on a (paracompact) manifold $X$, or more generally a sheaf of modules over $C^\infty$. We then have:

**Theorem 6.7** *The cohomology groups $H^q(X,\mathcal{F}) = 0$ for all $q > 0$.*

**Proof for $\boldsymbol{q = 1}$.** To indicate the argument, we will show $H^1(\mathfrak{U},\mathcal{F}) = 0$ for any open covering $\mathfrak{U} = (U_i)$. We will use the fact that there exists a partition of unity $\rho_i \in C^\infty(X)$ subordinate to $U_i$: that is, a set of functions with $K_i = \operatorname{supp}\rho_i \subset U_i$, such that $K_i$ forms a locally finite covering of $X$ and $\sum \rho_i(x) = 1$ for all $x \in X$.

Let $g_{ij} \in Z^1(\mathfrak{U},\mathcal{F})$ be a 1-cocycle. Then $g_{ii} = 0$, $g_{ij} = -g_{ji}$ and $g_{ij} + g_{jk} = g_{ik}$. Our goal is to write $g_{ij} = f_j - f_i$ (or $f_i - f_j$).

How will we ever get from $g_{ij}$, which is only defined on $U_{ij}$, a function $f_i$ define on all of $U_i$? The central observation is that:

> $\rho_j g_{ij}$*, extended by* $0$*, is smooth on* $U_i$*.*

This is because $\rho_j$ vanishes on $U_i - U_j$. More precisely, $\rho_j g_{ij}$ is smooth wherever $g_{ij}$ is defined or $\rho_j$ is zero. That is, it is smooth on

$$(U_i - U_j) \cup (U_{ij}) = U_i.$$

(Some care should be taken at the points of $\partial U_j \cap U_i$.) Thus we can define:

$$f_i = \sum_k \rho_k g_{ik};$$

and then:

$$f_i - f_j = \sum_k \rho_k (g_{ik} - g_{jk}) = \sum_k \rho_k (g_{ik} + g_{kj}) = \sum_k \rho_k g_{ij} = g_{ij}.$$

■

**Proof for $q = 2$.** Let $f_{ab} = \sum_x \rho_x g_{abx}$. Then

$$\delta(f)_{abc} = f_{bc} - f_{ac} + f_{ab} = \sum_x \rho_x (g_{bcx} - g_{acx} + g_{abx}) = \sum_x \rho_x g_{abc} = g_{abc}.$$

■

**The exact cohomology sequence; deRham cohomology.** We can now explain how sheaf cohomology is used to capture global aspects of analytic problems that can be solved locally.

Let $\mathcal{E}^p$ denote the sheaf of smooth $p$-forms on a manifold $X$. Suppose $\alpha \in \mathcal{E}^1(X)$ is closed; then locally $\alpha = df$ for $f \in \mathcal{E}^0(X)$. When we can we find a global primitive for $\alpha$?

To solve this problem, let $\mathfrak{U} = (U_i)$ be an open covering of $X$ by disks. Then we can write $\alpha = df_i$ on $U_i$. On the overlaps, $g_{ij} = f_i - f_j$ satisfies $dg_{ij} = 0$, i.e. it is a constant function. Moreover we obviously have $g_{ij} + g_{jk} = g_{ik}$, i.e. $g_{ij}$ is an element of $Z^1(\mathfrak{U}, \mathbb{C})$.

Now we may have chosen our $f_i$ wrong to fit together, since $f_i$ is not uniquely determined by the condition $df_i = \alpha_i$; we can always add a constant function $c_i$. But if we replace $f_i$ by $f_i + c_i$, then $g_{ij}$ will change by the coboundary $c_i - c_j$. Thus we can conclude:

$$\alpha = df \text{ iff } [g_{ij}] = 0 \text{ in } H^1(X, \mathbb{C}) .$$

**The exact cohomology sequence.** The conceptual theorem underlying the preceding discussion is the following:

**Theorem 6.8** *Any short exact sequence of sheaves on $X$,*

$$0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0,$$

*gives rise to a long exact sequence*

$$\begin{aligned} 0 \quad \to \quad & H^0(X,\mathcal{A}) \to H^0(X,\mathcal{B}) \to H^0(X,\mathcal{C}) \to \\ & H^1(X,\mathcal{A}) \to H^1(X,\mathcal{B}) \to H^1(X,\mathcal{C}) \to \\ & H^2(X,\mathcal{A}) \to H^2(X,\mathcal{B}) \to H^2(X,\mathcal{C}) \to \cdots \end{aligned}$$

*on the level of cohomology.*

Note: for *any* open set $U$, we get an exact sequence

$$0 \to \mathcal{A}(U) \to \mathcal{B}(U) \to \mathcal{C}(U), \tag{6.1}$$

as can be checked using the sheaf axioms. Surjectivity of the maps $\mathcal{B}_x \to \mathcal{C}_x$ implies that for any $c \in \mathcal{C}(X)$, there is an open covering $(U_i)$ and $b_i \in \mathcal{B}(U_i)$ mapping to $c$.

To obtain the *connecting homomorphism*

$$\xi : H^0(X,\mathcal{C}) \to H^1(X,\mathcal{A}),$$

we use the exactness of (6.1) to write $b_i - b_j$ as the image of $a_{ij}$. The resulting cocycle $[a_{ij}] \in H^1(X,\mathcal{A})$ is the image $\xi(c)$.

**Sheaves and deRham cohomology.** Let $\mathcal{Z}^p$ denote the sheaf of closed $p$-forms. Recall that the deRham cohomology groups of $X$ are given by:

$$H^p_{DR}(X) = \mathcal{Z}^p(X)/d\mathcal{E}^{p-1}(X).$$

The *Poincaré Lemma* asserts that

$$H^p_{DR}(\mathbb{R}^n) = 0$$

for all $p$ and $n$.

**Chain homotopy.** The proof is by induction on $n$. We have natural maps $\pi : \mathbb{R}^{n+1} \to \mathbb{R}^n$ and $s : \mathbb{R}^n \to \mathbb{R}^{n+1}$, such that $\pi \circ s = \mathrm{id}$. These maps in fact give a homotopy equivalence — no surprise, since both spaces are contractible. We wish to show they induce isomorphisms on deRham cohomology, which amount so showing that the map on forms given by

$$\phi = I - \pi^* \circ s^*$$

induces the identity on $H^*_{DR}(\mathbb{R}^{n+1})$.

To this end there is a basic tool. Note that $\phi$ is a chain map — it commutes with $d$. To show $\phi$ is trivial on cohomology, it suffices to construct a *chain homotopy* $K$, which lowers the degree of forms by one, such that

$$\phi = dK + Kd.$$

(Note that $d(dK + Kd) = dKd = (dK + Kd)d$). If $[\alpha]$ represents a cohomology class, then $d\alpha = 0$, so

$$\phi(\alpha) = (dK + Kd)(\alpha) = dK(\alpha)$$

is cohomologous to zero, and we are done. Here are the details.

**Proof.** We can regard $\mathbb{R}^{n+1}$ as a bundle over $\mathbb{R}^n$ with projection $\pi$ and zero section $s$. We wish to construct $K$ such that

$$I - \pi^* \circ s^* = Kd + dK.$$

Let $(t, x)$ denote coordinates on $\mathbb{R} \times \mathbb{R}^n$, and let $\alpha$ denote a product of some subset of $dx_1, \ldots, dx_n$. (So $d\alpha = 0$). Then we define $K(f(t, x)\alpha) = 0$, and

$$K(f(t, x)\, dt\, \alpha) = F(t, x)\alpha,$$

where $F(t, x) = \int_0^t f(t, x)\, dt$.

The verification that $K$ gives a chain homotopy is straightforward. For example, $dK + Kd$ should be the identity on a form like $f\, dt\alpha$, since $s^*(dt) = 0$. And indeed,

$$dK(f\, dt, \alpha) = d(F\alpha) = f\, dt\, \alpha + (d_x F)\alpha$$

while

$$Kd(f\, dt\alpha) = K((d_x f)\, dt\alpha) = -(d_x F)\alpha.$$

The sign enters because we must move $dt$ in front of $d_x f$.

On a form like $f\alpha$ we should get $(f(t, x) - f(0, x))\alpha$, and indeed $K$ kills this form while

$$Kd(f\alpha) = K((df/ft)\, dt\alpha + (d_x f)\alpha) = (f(t, x) - f(0, x))\alpha.$$

∎

**The deRham theorem.** Since every manifold looks locally like $\mathbb{R}^n$, the Poincaré Lemma implies we have an exact sequence of sheaves:

$$0 \to \mathcal{Z}^{p-1} \to \mathcal{E}^{p-1} \xrightarrow{d} \mathcal{Z}^p \to 0.$$

Now let $p = 1$. Then $\mathcal{Z}^{p-1} = \mathbb{C}$. By examining the associate long exact sequence we find:

**Theorem 6.9** *For any manifold $X$, we have $H^1_{DR}(X) \cong H^1(X, \mathbb{C})$.*

More generally, using all the terms in the exact sequence and all values of $p$, we find:

**Theorem 6.10** *For any manifold $X$, we have*

$$H^p_{DR}(X) \cong H^1(X, \mathcal{Z}^{p-1}) \cong H^2(X, \mathcal{Z}^{p-2}) \cong \cdots H^p(X, \mathcal{Z}^0) = H^p(X, \mathbb{C}).$$

A key point in this discussion is that:

*The sheaf $\mathcal{Z}^p$, unlike $\mathcal{E}^p$, is not a module over $C^\infty$.*

This is particularly clear for $\mathcal{Z}^0 \cong \mathbb{C}$, and it explains why the cohomology of $\mathcal{Z}^p$ need not vanish.

**Corollary 6.11** *The deRham cohomology groups of homeomorphic smooth manifolds are isomorphic.*

(In fact one can replace 'homeomorphic' by 'homotopy equivalent' here.)

**Finiteness.** Now it is easy to prove that $H^p_{DR}(\mathbb{R}^n) = 0$ for all $p > 0$. We thus have, by taking a Leray covering:

**Theorem 6.12** *For any compact manifold, the cohomology groups $H^p(X, \mathbb{C})$ are finite.*

**Periods revisited.** Finally we mention the fundamental group:

**Theorem 6.13** *For any connected manifold $X$, we have*

$$H^1(X, \mathbb{C}) \cong H^1_{DR}(X) \cong \mathrm{Hom}(\pi_1(X), \mathbb{C}).$$

**Proof.** We have $\int_\gamma df = 0$ for all closed loops $\gamma$, so the period map is well-defined; if $\alpha$ has zero periods then $f(q) = \int_p^q \alpha$ is also well-defined and satisfies $df = \alpha$. To prove surjectivity, take a (Leray) covering of $X$ by geodesicly convex sets, and observe that (i) every element of $\pi_1(X)$ is represented by a 1-chain and (ii) every 1-boundary is a product of commutators. ∎

**Warning: exotic topological spaces** . One can define the period map $H^1(X,\mathbb{C}) \to \text{Hom}(\pi_1(X),\mathbb{C})$ directly, using $\gamma : S^1 \to X$ to obtain from $\xi \in H^1(X,\mathbb{C})$ a class $\phi(\gamma) \in H^1(S^1,\mathbb{C}) \cong \mathbb{C}$. For more exotic topological spaces, however, this map need not be an isomorphism: e.g. the 'topologist's sine curve' is a compact, connected space $X$ with $\pi_1(X) = 0$ but $H^1(X,\mathbb{Z}) = \mathbb{Z}$.

# 7 Cohomology on a Riemann surface

On a Riemann surface we have the notion of *holomorphic* functions and forms. Thus in addition to the sheaves $\mathcal{E}^p$ we have the important sheaves:

$\mathcal{O}$ — the sheaf of holomorphic functions; and
$\Omega$ — the sheaf of holomorphic 1-forms.

Let $h^i(\mathcal{F}) = \dim H^i(X,\mathcal{F})$. We will show that a compact Riemann surface satisfies:

$$h^0(\mathcal{O}) = 1, \quad h^0(\Omega) = g,$$
$$h^1(\mathcal{O}) = g, \quad h^1(\Omega) = 1.$$

The symmetry of this table is not accidental: it is rather the first instance of Serre duality, which we will also prove.

We remark that a higher cohomology group, like $H^1(X,\mathcal{O})$, is a sort of place holder, like paper money. It has no actual worth until it is exchanged for something else. Serre duality makes such an exchange possible.

**The Dolbeault Lemma.** Just as the closed forms can be regarded as the subsheaf $\text{Ker}\, d \subset \mathcal{E}^p$, the holomorphic functions can be regarded as the subsheaf $\text{Ker}\,\overline{\partial} \subset \mathcal{C}^\infty = \mathcal{E}^0$. So we must begin by studying the $\overline{\partial}$ operator.

We begin by studying the equation $df/d\overline{z} = g \in L^1(\mathbb{C})$. An example is given for each $r > 0$ by:

$$f_r(z) = \begin{cases} 1/z & \text{if } |z| > r, \\ \overline{z}/r^2 & \text{if } |z| \le r, \end{cases}$$

which satisfies $g_r = df/d\overline{z} = (1/r^2)\chi_{B(0,r)}(z)$. In particular $\int g_r = \pi$ is independent of $r$, which suggests the distributional equation:

$$\frac{d}{d\overline{z}}\frac{1}{z} = \pi\delta_0.$$

Using this fundamental solution (and the fact that $dx\,dy = (i/2)dz \wedge d\overline{z}$), we obtain:

**Theorem 7.1** *For any $g \in C_c^\infty(\mathbb{C})$, a solution to the equation $df/d\overline{z} = g$ is given by:*

$$f(z) = g * \frac{1}{\pi z} = \frac{i}{2\pi} \int_{\mathbb{C}} \frac{g(w)}{z-w}\, dw \wedge d\overline{w}.$$

**Proof.** It is enough to check $df/d\overline{z}$ at $z=0$. Using that fact that $df/d\overline{z} = (dg/d\overline{z}) * \frac{1}{\pi z}$, we find:

$$\begin{aligned} -2\pi i \frac{df}{dz}(0) &= \int_{\mathbb{C}} \frac{dg}{d\overline{z}} \left(\frac{-1}{z}\right) dz\, d\overline{z} = \int_{\mathbb{C}} d\left(\frac{g(z)\, dz}{z}\right) \\ &= \lim_{r\to 0} \int_{\mathbb{C}-B(0,r)} d\left(\frac{g(z)\, dz}{z}\right) = -\lim_{r\to 0} \int_{S^1(0,r)} g(z)\frac{dz}{z} = -2\pi i\, g(0). \end{aligned}$$

The penultimate minus sign comes from the back that $S^1(0,r)$ acquires a negative orientation as $\partial(\mathbb{C} - B(0,r))$. ■

**Theorem 7.2** *For any $g \in \mathcal{C}^\infty(\Delta)$, there is an $f \in C^\infty(\Delta)$ with $df/d\overline{z} = g$.*

**Proof.** Write $g = \sum g_n$ where each $g_n$ is smooth and compactly supported outside the disk $D_n$ of radius $1 - 1/n$. Solve $df_n/d\overline{z} = g_n$. Then $f_n$ is holomorphic on $D_n$. Expanding it in power series, we can find holomorphic functions $h_n$ on the disk such that $|f_n - h_n| < 2^{-n}$ n $D_n$. Then $f = \sum(f_n - h_n) = \lim F_i$ converges uniformly, and $F - F_i$ is holomorphic on $D_n$ for all $i > n$, so the convergence is also $C^\infty$. ■

**Corollary 7.3** *For any $g \in \mathcal{C}^\infty(\Delta)$, there is an $f \in C^\infty(\Delta)$ with $\Delta f = g$.*

**Proof.** First solve $dh/dz = g$, then $df/d\overline{z} = h$. Then $\Delta(f/4) = d^2 f/dz\, d\overline{z} = g$. ■

**Remarks.** The same results hold with $\Delta$ replaced by $\mathbb{C}$. It is easy to solve the equation $df/d\overline{z} = g$ on $\mathbb{C}$ when $g(z, \overline{z})$ is a polynomial: simply formally integrate with respect to the $\overline{z}$ variable.

In higher dimensions, $\omega = \overline{\partial} g$ should be thought of as a $(0,1)$-form, and this equation can be solved for $g$ only when $\overline{\partial}\omega = 0$.

**Functions with small $\overline{\partial}$ derivative.** Here is an easy consequence of the Dolbeault lemma above. Suppose $f(z)$ is smooth on the disk $2\Delta$ of radius 2. Then on the unit disk $\Delta$, we can write

$$f(z) = h(z) + g(z)$$

where $h(z)$ is holomorphic and $g$ is small on $\Delta$ if $\overline{\partial} f$ is small on $2\Delta$. Here smallness can be measured, e.g. in the $C^k$ norm. Informally,

$$f = \text{holomorphic} + O(\overline{\partial} f).$$

To see this, we simply solve $\overline{\partial} g = \overline{\partial} \rho f$ by convolution with $1/z$, where $\rho$ is a smooth cutoff function $= 1$ on $\Delta$ and $= 0$ outside $2\Delta$.

**Dolbeault cohomology.** Let us define, for any Riemann surface $X$,

$$H^{0,1}(X) = \mathcal{E}^{0,1}(X) \big/ \overline{\partial} \mathcal{E}^0(X).$$

The preceding results show the sequence of sheaves:

$$0 \to \mathcal{O} \to \mathcal{E}^0 \xrightarrow{\overline{\partial}} \mathcal{E}^{0,1} \to 0$$

is exact. Consequently we have:

**Theorem 7.4** *On any Riemann surface $X$, we have $H^1(X, \mathcal{O}) \cong H^{0,1}(X)$.*

Thus the Dolbeault lemma can be reformulated as:

**Theorem 7.5** *The unit disk satisfies $H^1(\Delta, \mathcal{O}) \cong H^{0,1}(\Delta) = 0$. The same is true for the complex plane.*

**Corollary 7.6** *We have $H^1(\widehat{\mathbb{C}}, \mathcal{O}) = 0$.*

**Proof.** Let $U_1 \cup U_2$ be the usual covering by $U_1 = \mathbb{C}$ and $U_2 = \widehat{\mathbb{C}} - \{0\}$. By the preceding result, $H^1(U_i, \mathcal{O}) = 0$. Thus by Leray's theorem, this covering is sufficient for computing $H^1$: $H^1(\widehat{\mathbb{C}}, \mathcal{O}) = H^1(\mathfrak{U}, \mathcal{O})$. Suppose $g_{12} \in \mathcal{O}(U_{12}) = \mathcal{O}(\mathbb{C}^*)$ is given. Then $g_{12}(z) = \sum_{-\infty}^{\infty} a_n z^n$. Splitting this Laurent series into its positive and negative parts, we obtain $f_i \in \mathcal{O}(U_i)$ such that $g_{12} = f_2 - f_1$. ∎

Similarly, we define

$$H^{1,1}(X) = \mathcal{E}^{1,1}(X) \big/ \overline{\partial} \mathcal{E}^{1,0}(X).$$

**Theorem 7.7** *On any Riemann surface $X$ we have $H^1(X, \Omega) \cong H^{1,1}(X)$.*

**Proof.** Consider the exact sequence of sheaves

$$0 \to \Omega \to \mathcal{E}^{1,0} \xrightarrow{\overline{\partial}} \mathcal{E}^{1,1} \to 0.$$

∎

**Corollary 7.8** *The dimension of $H^1(X, \Omega)$ is $\geq 1$.*

**Proof.** The map $\alpha \mapsto \int_X \alpha$ on $H^{1,1}(X)$ has image $\mathbb{C}$. ∎

We will later see that equality holds.

**The Dolbeault isomorphism.** Using the fact that $\overline{\partial}^2 = 0$ one can defined the Dolbeault cohomology groups for general complex manifolds, and prove using sheaf theory the following variant of the deRham theorem:

**Theorem 7.9** *For any compact complex manifold $X$, we have $H^{p,q}_{\overline{\partial}}(X) \cong H^q(X, \Omega^p)$.*

Here $\Omega^q$ is the sheaf of holomorphic $(q, 0)$-forms.

**Finiteness.** Recall that we already know $\dim \Omega(X) \leq g$, by period considerations; in particular, $\Omega(X)$ is finite-dimensional. We remark that a more robust proof of this finite-dimensionality can be given by endowing $\Omega(X)$ with a reasonable normal — e.g. $\|\alpha\|^2 = \int_X |\alpha|^2$ — and then observing that the unit ball is compact. (The same proof applies to show $\dim \mathcal{O}(X) < \infty$, without using the maximum principle. More generally the space of holomorphic sections of a complex line bundle over a compact space is finite-dimensional.)

A proof of the *finiteness* of $\dim H^1(X, \mathcal{O})$, based on norms as in the discussion, is given in Forster.

**Serre duality, special case.** We will approach finiteness instead through elliptic regularity. We begin with a special case of Serre duality, which allows one to relate $H^1$ and $H^0$. The proof below is based on *Weyl's lemma*, which states that a distribution $f$ is represented by a holomorphic function if $\overline{\partial} f = 0$.

**Theorem 7.10** *On any compact Riemann surface $X$, we have $H^1(X, \mathcal{O})^* \cong \Omega(X)$. In particular, $H^1(X, \mathcal{O})$ is finite-dimensional.*

We let $g_a = \dim \Omega(X)$, the *arithmetic genus* of $X$. We will eventually see that $g_a = g =$ the topological genus.

**Proof.** By Dolbeault we have

$$H^1(X, \mathcal{O}) \cong H^{0,1}(X) = \mathcal{E}^{(0,1)}(X)/\overline{\partial}\mathcal{E}^0(X).$$

We claim $\overline{\partial}\mathcal{E}^0(X)$ is closed in $\mathcal{E}^{(0,1)}(X)$ in the $C^\infty$ topology. If not, there is a sequence $f_n \in \mathcal{E}^0(X)$ with $\overline{\partial} f_n \to \omega$ in the $C^\infty$ topology, such that

$f_n$ has no convergent subsequence in $\mathcal{E}^0(X)$. Since $\overline{\partial}$ annihilates constants, we may normalize so that $f_n(p) = 0$ for some $p$ and all $n$. Since bounded sets in $\mathcal{E}^0(X)$ are compact, the latter condition implies for some $k \geq 0$, $\|f_n\|_{C^k} \to \infty$. From this we will obtain a contradiction.

Dividing through by the $C^k$-norm, we can arrange that $\|f_n\|_{C^k} = 1$ in $\mathcal{E}^0$, and that $\overline{\partial} f_n \to 0$ in the $C^\infty$ topology. Solving the $\overline{\partial}$-equation locally, this means we can write locally write

$$f_n = h_n + g_n$$

where $h_n$ is holomorphic and $g_n \to 0$ in $C^\infty$. But then $h_n$ is bounded, so we can pass to a subsequence so $h_n$ converges smoothly. Then $f_n \to F$ smoothly, so $\overline{\partial} F = 0$, so $F$ is constant. Indeed, $F = 0$ since $f_n(p) = 0$. On the other hand, our normalization that $\|f_n\|_{C_k} = 1$ implies $\|F\|_{C^k} = 1$, a contradiction.

Since $\overline{\partial}\mathcal{E}^0(X)$ is closed, we have

$$H^1(X, \mathcal{O})^* = W \cong (\overline{\partial}\mathcal{E}^0(X))^\perp \subset (\mathcal{E}^{0,1}(X))^* = \mathcal{D}^{1,0}(X).$$

But any $(1,0)$-current $\omega \in W$ satisfies $\int \omega \wedge (\overline{\partial} f) = 0$ for all smooth $f$, and thus $\overline{\partial}\omega = 0$, which implies $\omega$ is holomorphic and thus $W = \Omega(X)$. ∎

**Corollary 7.11** *We have natural isomorphisms* $H^{1,0}(X) \cong \Omega(X)$ *and* $H^{0,1}(X) \cong \overline{\Omega}(X)$.

**Proof.** By definition, $H^{1,0}(X)$ consists of $\overline{\partial}$-closed $(1,0)$ forms, which means it is isomorphic to $\Omega(X)$.

As for $H^{0,1}(X)$, we have a natural map $\overline{\Omega}(X) \to H^{0,1}(X)$. On this subspace, the pairing with $\Omega(X)$ given by $\int \alpha \wedge \overline{\beta}$ is obviously perfect, since $(i/2)\int \alpha \wedge \overline{\alpha} > 0$. Thus $\overline{\Omega}$ is isomorphic to $H^{0,1}(X)$. ∎

**Elliptic regularity.** Here is a brief sketch of the proof of elliptic regularity. We will show that if $f$ is a distribution on $\mathbb{R}^n$ and $\Delta f = 0$, then $f$ is represented by a smooth harmonic function.

The first point to observe is that morally, if $f$ is a compactly supported distribution and

$$\Delta f = u$$

with $u \in C^k$, then $f \in C^{k+2}$. Here $k$ is allowed to be negative. This can be made precise in terms of Fourier transforms and Sobolev spaces; it comes from the fact that the symbol of $\Delta$, namely $\sum \xi_i^2$, has quadratic growth.

Now suppose we have a distribution with $\Delta f = 0$. Let $\rho$ be a $C^\infty$, compactly supported cutoff function. Then

$$\Delta(\rho f) = \rho \Delta f + u = u$$

where $u$ involves only the first derivatives of $f$ (and $\rho$). Thus if $f$ is in $C^k$, then $u \in C^{k-1}$ and hence $\rho f$ is in $C^{k+1}$. With some finesse to get around the fact that $\rho f$ is not quite the same as $f$, this shows $f \in C^\infty$.

For more details, see e.g. Rudin, *Functional Analysis*.

**Taking stock.** In preparation for Riemann-Roch, we now know $H^1(X, \mathcal{O})^* \cong \Omega(X)$, $g_a = \dim \Omega(X) \leq g$, and $\dim H^1(X, \Omega) \geq 1$ because of the residue map.

**The space $\boldsymbol{H^1(X, \Omega)}$.** We can mimic the proof that $H^1(X, \mathcal{O}) \cong \Omega(X)$ to show

$$H^{1,1}(X) \cong H^0(X, \mathcal{O})^* \cong \mathbb{C}.$$

The main point is that

$$H^{1,1}(X)^* \cong (\overline{\partial} \mathcal{E}^{1,0}(X))^\perp \subset \mathcal{D}^0(X),$$

and we have

$$\int f \overline{\partial} \alpha = 0$$

for all $\alpha \in \mathcal{E}^{(1,0)}(X)$ iff $\overline{\partial} f = 0$ iff $f$ is holomorphic.

Another argument will be given below.

# 8 Riemann-Roch

One of the most basic questions about a compact Riemann surface $X$ is: does there exist a nonconstant holomorphic map $f : X \to \widehat{\mathbb{C}}$? But as we have seen already in the discussion of the field $\mathcal{M}(X)$, it is desirable to ask more: for example, do the meromorphic functions on $X$ separate points? Even better, does there exist a holomorphic embedding $X \to \mathbb{P}^n$ for some $n$?

To answer these questions we aim to determine the *dimension* of the space of meromorphic functions with *controlled* zeros and poles. This is the Riemann-Roch problem.

**Divisors.** Let $X$ be a compact Riemann surface. The group of *divisors* $\mathrm{Div}(X)$ is the free abelian group generated by the points of $X$. A divisor is sum $D = \sum a_P P$, where $a_P \in \mathbb{Z}$ and $a_P = 0$ for all but finitely many $P \in X$.

A divisor is *effective* if $D \geq 0$, meaning $a_P \geq 0$ for all $P$. Any divisor can be written uniquely as a difference of effective divisors, $D = D_+ - D_-$. We write $E \geq D$ if $E - D \geq 0$.

The *degree* of a divisor, $\deg(D) = \sum a_P$, defines a homomorphism $\deg : \mathrm{Div}(X) \to \mathbb{Z}$. We let $\mathrm{Div}_0(X)$ denote the subgroup of divisors of degree zero.

**The sheaf of divisors.** The quotient $\mathcal{M}^*/\mathcal{O}^*$ is not officially a sheaf, but it can be made into a sheaf in a standard way (take sections of the espace étalé). A section of $\mathcal{M}^*/\mathcal{O}^*$ over $U$ is given by a covering $U_i$ and elements $f_i \in \mathcal{M}^*(U_i)$ such that $f_i/f_j \in \mathcal{O}^*(U_{ij})$. Note that $\deg_P(f_i)$ gives the same value for any $i$ with $P \in U_i$.

One can also define a *sheaf* by $\mathrm{Div}(U) = \{\sum a_P P\}$ where the sum is *locally* finite. We then have a natural isomorphism of sheaves,

$$\mathcal{M}^*/\mathcal{O}^* \cong \mathrm{Div},$$

send $(f_i)$ to $\sum \deg_P(f_i)$.

**Principal divisors.** The divisor of a *global* meromorphic function $f \in \mathcal{M}^*(X)$ is defined by

$$(f) = \sum_P \mathrm{ord}(f, P) \cdot P.$$

The divisors that arise in this way are said to be *principal.* Note that:

$$(fg) = (f) + (g) \quad \text{and} \quad \deg((f)) = 0,$$

so $f \mapsto (f)$ defines a homomorphism from

$$\mathcal{M}^*(X) \to \mathrm{Div}_0(X).$$

(Its cokernel is a beautiful complex torus, the *Jacobian* of $X$, to be discussed later.)

Note that the *topological* degree of $f : X \to \widehat{\mathbb{C}}$ is given by $\deg((f)_+)$, which gives the number of zeros of $f$ counted with multiplicities.

The map $(f) \mapsto \mathrm{Div}(f)$ can be thought of as part of the long exact sequence associated to

$$0 \to \mathcal{O}^* \to \mathcal{M}^* \to \mathcal{M}^*/\mathcal{O}^* \to 0,$$

which is short exact by the definition of the quotient.

**Constrained zeros and poles** We defined $\mathcal{O}_D$ as the sheaf of complex vector spaces given by meromorphic functions $f$ such that $(f) + D \geq 0$. For example, $\mathcal{O}_{nP}(X)$ is the vector space of meromorphic functions on $X$ with poles of order at most $n$ at $p$. In general $\mathcal{O}_D$ keeps the poles of $f$ from being to big and forces some zeros at the points where $a_P < 0$.

If $D - E = (f)$ is principal, we say $D$ and $E$ are *linearly equivalent.* Then the map $h \mapsto hf$ gives an isomorphism of sheaves:

$$\mathcal{O}_D = \mathcal{O}_{E+(f)} \cong \mathcal{O}_E,$$

because

$$(hf) + E = (h) + (f) + E = (h) + D.$$

**Constrained forms.** We let $\mathcal{M}^1(X)$ denote the space of *meromorphic* 1-forms on $X$. It is a 1-dimensional vector space over the field $\mathcal{M}(X)$. The divisor $K = (\omega)$ of a nonzero meromorphic 1-form $\omega$ is defined just as for a meromorphic function, in terms of the zeros and poles and $\omega$.

Once we have such a *canonical* divisor, we get an isomorphism

$$\mathcal{O}_K \cong \Omega$$

by $h \mapsto h\omega$, because $h\omega$ is holomorphic iff

$$(h\omega) = (h) + K \geq 0$$

iff $h \in \mathcal{O}_K$.

Similarly, for any divisor $D$ we let

$$\Omega_D(X) = \{\omega \in \mathcal{M}^1(X) \; : \; (\omega) + D \geq 0\}.$$

The sheaf $\Omega_D$ is defined similarly, and it satisfies

$$\Omega_D \cong \mathcal{O}_{K+D}\,.$$

Global sections of this sheaf correspond to meromorphic forms with constrained zeros and poles.

**Riemann-Roch Problem.** The *Riemann-Roch problem* is to calculate or estimate, for a given divisor $D$, the number

$$h^0(D) = \dim H^0(X, \mathcal{O}_D) = \dim \mathcal{O}_D(X).$$

Example: if $X$ is a complex torus, we have $h^0(nP) = 0, 1, 1, 2, 3, \ldots$ for $n = 0, 1, 2, 3, 4, \ldots$. This can be explained by the fact that we must have $\mathrm{Res}_P(f\,dz) = 0$.

It is a general principle that Euler characteristics are more stable than individual cohomology groups, and for any sheaf of complex vector spaces $\mathcal{F}$ on $X$ we define:

$$\chi(\mathcal{F}) = \sum (-1)^q h^q(\mathcal{F}),$$

assuming these spaces are finite–dimensional. (Of course $h^q(\mathcal{F}) = 0$ for all $q > 2$ by topological considerations.)

Example: Using the fact that $h^1(\mathcal{O}) = h^0(\Omega) = g_a$, we have

$$\chi(\mathcal{O}) = 1 - g_a.$$

We may now state:

**Theorem 8.1 (Riemann-Roch, Euler characteristic version)** *For any divisor $D$, we have*

$$\chi(\mathcal{O}_D) = \chi(\mathcal{O}) + \deg(D).$$

Equivalently, we have

$$h^0(\mathcal{O}_D) - h^1(\mathcal{O}_D) = \deg D - g_a + 1.$$

We will soon show that $g_a = g$.

For the proof we will use:

**Theorem 8.2** *Let $0 \to V_1 \to \cdots \to V_n \to 0$ be an exact sequence of finite–dimensional vector spaces. Then $\sum (-1)^i \dim(V_i) = 0$.*

**Proof.** Let $\phi_i : V_i \to V_{i+1}$. Then $\dim(V_i) = \dim \operatorname{Im} \phi_i + \dim \operatorname{Ker} \phi_i$, while exactness gives $\dim \operatorname{Im} \phi_i = \dim \operatorname{Ker} \phi_{i+1}$. Thus

$$\sum (-1)^i \dim V_i = \sum (-1)^i (\dim \operatorname{Ker} \phi_i + \dim \operatorname{Ker} \phi_{i+1}) = 0.$$

■

**Corollary 8.3** *If $0 \to \mathcal{A} \to \mathcal{B} \to \mathcal{C} \to 0$ is an exact sequence of sheaves, and two of the three have well-defined Euler characteristics, then so does the third, and we have:*

$$\chi(B) = \chi(A) + \chi(C).$$

**Proof.** Apply the previous result to the long exact sequence in cohomology.

■

**Skyscrapers.** The *skyscraper sheaf* $\mathbb{C}_P$ is given by $\mathcal{O}(U) = \mathbb{C}$ if $P \in U$, and $\mathcal{O}(U) = 0$ otherwise. For any divisor $D$, we have the exact sequence:

$$0 \to \mathcal{O}_D \to \mathcal{O}_{D+P} \to \mathbb{C}_P \to 0, \tag{8.1}$$

where the final map records the leading coefficient of the polar part of $f$ at $P$. It is easy to see (e.g. by taking fine enough coverings, without $P$ in any multiple intersections):

**Theorem 8.4** *We have $H^p(X, \mathbb{C}_P) = 0$ for all $p > 0$. In particular $\chi(\mathbb{C}_P) = h^0(\mathbb{C}_P) = 1$.*

**Theorem 8.5** *The cohomology groups $H^p(X, \mathcal{O}_D)$ are finite-dimensional for $p = 0, 1$ and vanish for all $p \geq 2$.*

**Proof.** We have already seen the result is true for $D = 0$: the space $H^1(X, \mathcal{O}) \cong \Omega(X)^*$ is finite-dimensional, and using the Dolbeault sequence, one can show $H^p(X, \mathcal{O}) = 0$ for all $p \geq 2$. The result for general $D$ then follows induction, using the skyscraper sheaf. ∎

**Proof of Riemann-Roch.** Since $h^1(\mathcal{O}) = \dim \Omega(X) = g_a$, the formula is correct for the trivial divisor. Using (8.1) and the additivity of the Euler characteristic for short exact sequences, we find:

$$\chi(\mathcal{O}_{D+P}) = \chi(\mathcal{O}_D) + \chi(\mathbb{C}_P) = \chi(\mathcal{O}_D) + 1,$$

which implies Riemann-Roch for an arbitrary divisor $D$. ∎

**Existence of meromorphic functions and forms.** A more general form of Serre duality will lead to a more useful formulation of Riemann-Roch, but we can already deduce several useful consequences.

**Theorem 8.6** *Any compact Riemann surface admits a nonconstant map $f : X \to \mathbb{P}^1$ with $\deg(f) \leq g_a + 1$.*

**Proof.** We have $\dim H^0(X, \mathcal{O}_D) \geq \deg D - g_a + 1$. Once $\deg D > g_a$ this gives $\dim H^0(\mathcal{O}_D) > 1 = \dim \mathbb{C}$. ∎

(Note: this bound on $\deg(f)$ is not optimal. The optimal bound, coming from Brill–Noether theory, is given by $(g+3)/2$ rounded down. We will later see this bound explicitly for genus $g \leq 3$.)

**Corollary 8.7** *Any surface with $g_a = 0$ is isomorphic to $\mathbb{P}^1$.*

**Proof.** It admits a map to $\mathbb{P}^1$ of degree $g_a + 1 = 1$. ∎

**Corollary 8.8** *Any compact Riemann carries a nonzero meromorphic 1-form.*

**Proof.** Take $\omega = df$.

**Corollary 8.9** *Canonical divisors $K = (\omega)$ exist, and satisfy $\mathcal{O}_K \cong \Omega$.*

The isomorphism is given by $f \mapsto f\omega$.

**Arithmetic and topological genus; degree of canonical divisors; residues.**

**Corollary 8.10** *The degree of any canonical divisor is $2g - 2$, where $g$ is the topological genus of $X$.*

**Proof.** Apply Riemann-Hurwitz to compute $\deg(df)$. ∎

**Corollary 8.11** *The topological and arithmetic genus of $X$ agree: we have $g = g_a$; and $\dim H^1(X, \Omega) = 1$.*

**Proof.** Apply Riemann-Roch to a canonical divisor $K$: we get

$$h^0(K) - h^1(K) = g_a - h^1(K) = 1 - g_a + \deg(K) = 2g - g_a - 1.$$

Using the fact that $h^0(K) = g_a$, we can solve for $h^1(K)$:

$$h^1(K) = 2(g_a - g) + 1.$$

Now we know $g_a - g \leq 0$ and $h^1(K) \geq 1$, so we must have $g_a = g$ and $h^1(K) = 1$. ∎

**Theorem 8.12 (Hodge theorem)** *On a compact Riemann surface every class in $H^1_{DR}(X,\mathbb{C})$ is represented by a harmonic 1-form. More precisely we have*

$$H^1_{DR}(X) = \Omega(X) \oplus \overline{\Omega(X)} = H^{1,0}(X) \oplus H^{0,1}(X) = \mathcal{H}^1(X).$$

**Proof.** We already know the harmonic forms inject into deRham cohomology, by considering their periods; since the topological and arithmetic genus agree, they also surject. ■

**The space of smooth 1-forms.** On a compact Riemann surface, the full Hodge theorem

$$\mathcal{E}^1(X) = d\mathcal{E}^0(X) \oplus \mathcal{H}^1(X) \oplus d^*\mathcal{E}^2(X)$$

becomes the statement:

$$\mathcal{E}^1(X) = (\partial + \overline{\partial})\mathcal{E}^0(X) \oplus (\Omega(X) \oplus \overline{\Omega}(X)) \oplus (\partial - \overline{\partial})\mathcal{E}^0(X).$$

We have now proved this statement. Indeed, it suffices to show $\partial\mathcal{E}^0(X) \oplus \overline{\partial}\mathcal{E}^0(X)$ spans the complement of the harmonic forms. But the isomorphism $H^{0,1}(X) \cong \Omega(X)^* \cong \overline{\Omega}(X)$ gives $\mathcal{E}^{0,1} \cong \overline{\Omega}(X) \oplus \overline{\partial}\mathcal{E}^0(X)$, and similarly for $\mathcal{E}^{0,1}$.

**Theorem 8.13** *A smooth* $(1,1)$ *form* $\alpha$ *lies in the image of the Laplacian iff* $\int_X \alpha = 0$.

**Proof.** We must show $\alpha = \partial\overline{\partial}f$. Since the residue map on $H^{1,1}(X)$ is an isomorphism, we can write $\alpha = \overline{\partial}\beta$ where $\beta$ is a $(1,0)$-form. Since $\overline{\Omega}(X) \cong H^{0,1}(X)$, we can find a holomorphic 1-form $\omega$ such that $\overline{\beta} - \overline{\omega} = \overline{\partial f}$ for some smooth function $f$. Then $\partial f = \beta - \omega$, and hence $\overline{\partial}\partial f = \overline{\partial}\beta = \alpha$. ■

**Remark: isothermal coordinates.** The argument we have just given also proves the Hodge theorem for any compact, oriented Riemannian 2-manifold $(X,g)$. To see this, however, we need to know that every Riemannian metric is locally conformally flat; i.e. that one can introduce 'isothermal coordinates' to make $X$ into a Riemann surface with $g$ a conformal metric.

## 9 The Mittag–Leffler problems

In this section we show that the residue theorem is the only obstruction to the construction of a meromorphic 1-form with prescribed principal parts. The residue map on $H^1(X, \Omega)$ will also play a useful role in the discussion of Serre duality.

We then solve the traditional Mittag-Leffler theorem for meromorphic functions, and use it to prove the Riemann–Roch theorem for *effective* divisors.

**Mittag–Leffler for 1-forms.** The Mittag-Leffler problem for 1-forms is to construct a meromorphic 1-form $\omega \in \mathcal{M}^1(X)$ with prescribed polar parts. The input data is specified by meromorphic forms

$$\beta_i \in \mathcal{M}^1(U_i)$$

such that

$$\alpha_{ij} = \beta_i - \beta_j \in \Omega(U_i)$$

for all $i$. A *solution* to the Mittag-Leffler problem is given by a meromorphic form $\omega$ with

$$\omega = \beta_i + \gamma_i, \quad \gamma_i \in \Omega(U_i).$$

This just says that $\omega$ is a global meromorphic form whose principal parts are those given by $\beta_i$.

In terms of sheaves, we have

$$0 \to \Omega \to \mathcal{M}^1 \to \mathcal{M}^1/\Omega \to 0,$$

where as before the quotient on the right must be converted to a sheaf in the natural way. The Mittag-Leffler data $(\beta_i)$ determines an element of $H^0(\mathcal{M}^1/\Omega)$, and it has a solution if this element is in the image of $\mathcal{M}^1)(X)$. Thus the obstruct to finding a solution is given by the *Mittag–Leffler coboundary*

$$[\alpha_{ij}] \in H^1(\Omega).$$

**The residue map.** Now recall that the exact sequence

$$0 \to \Omega \to \mathcal{E}^{1,0} \xrightarrow{\overline{\partial}} \mathcal{E}^{1,1} \to 0$$

determines a natural isomorphism

$$H^1(X, \Omega) \cong H^{1,1}(X) \cong \mathbb{C},$$

and that the *residue map*, defined by

$$\mathrm{Res}(\omega) = \frac{1}{2\pi i}\int_X \omega,$$

determines a natural isomorphism

$$H^{1,1}(X) \cong \mathbb{C}.$$

The terminology is justified by the following result:

**Theorem 9.1** *On a Mittag-Leffler coboundary, $\alpha_{ij} = \beta_i - \beta_j$, we have*

$$\mathrm{Res}([\alpha_{ij}]) = \sum_p \mathrm{Res}_p(\beta_i).$$

Here we choose any $U_i$ such that $p \in U_i$ to compute the residue.

**Proof.** We must find explicitly the form $\omega \in \mathcal{E}^{1,1}$ representing the cocycle $\alpha_{ij}$. To do this we first write

$$\alpha_{ij} = \xi_i - \xi_j$$

for smooth $(1,0)$ forms $\xi_i$. Then we observe that, since $\alpha_{ij}$ is holomorphic, the $(1,1)$-forms $\overline{\partial}\xi_i$ fits together over patches to give a smooth $(1,1)$-form $\omega$. Now we also have $\alpha_{ij} = \beta_i - \beta_j$, and thus

$$\eta_i = \beta_i - \xi_i = \beta_j - \xi_j$$

on $U_{ij}$. So these forms piece together to give a global form $\eta$ such that

$$\overline{\partial}\eta = -\overline{\partial}\xi_i = -\omega.$$

The only nuance is that $\eta \in \mathcal{E}^{1,0} + \mathcal{M}^1$, i.e. it is locally the sum of a smooth form and a form with poles. However, $\eta$ is smooth on $X^*$, the result of puncturing $X$ at all the poles of the data $(\beta_i)$. If we replace $X^*$ by a surface with boundary circles around each of these poles, then we get

$$2\pi i\,\mathrm{Res}([\alpha_{ij}]) = \int_X \omega \approx \int_{X^*} -\overline{\partial}\eta = \int_{-\partial X^*} \eta \approx \sum 2\pi i\,\mathrm{Res}_p(\beta_i).$$

Letting the boundary of $X^*$ shrink to zero yields the Theorem. ∎

**Corollary 9.2** *There exists a meromorphic 1-form with prescribed principal parts iff the sum of its residues is equal to zero.*

**Corollary 9.3** *Given any pair of distinct points $p_1, p_2 \in X$, there is a meromorphic 1-form $\omega$ with simple poles of residues $(-1)^i$ at $p_i$ and no other singularities.*

This 'elementary differential of the third kind' is unique up to the addition of a global holomorphic differential. Example: the form $dz/z$ works for $0, \infty \in \widehat{\mathbb{C}}$.

**Corollary 9.4** *For any $p \in X$ and $n \geq 2$ there exists a meromorphic 1-form $\omega$ with a pole of order $n$ at $p$ (but vanishing residue) and no other singularities.*

This is an 'elementary differential of the second kind'.

**Residues and currents.** Underlying the calculation above (and the $\overline{\partial}$ equation) is a simple calculation the distributional derivative *on forms*:

$$\int_{\Delta} \overline{\partial}\left(\frac{dz}{z}\right) = \int_{S^1} \frac{dz}{z} = 2\pi i.$$

Using distributions, we get a more direct proof of the calculation above. Namely, we have an exact sequence

$$0 \to \Omega \to \mathcal{D}^{1,0} \xrightarrow{\overline{\partial}} \mathcal{D}^{1,1} \to 0,$$

and then if $\alpha_{ij} = \beta_i - \beta_j$, the *distributions* $\eta_i = \overline{\partial}\beta_i$ already agree on overlaps and give a global $(1,1)$-current $\eta$, satisfying, by the calculation above,

$$\operatorname{Res}(\eta) = \frac{1}{2\pi i} \int_X \eta = \sum \operatorname{Res}_p(\beta_i).$$

**Mittag-Leffler for functions.** Given a finite set of points $p_i \in X$, and the Laurent tails

$$f_i(z) = \frac{b_n}{z^n} + \cdots + \frac{b_1}{z}$$

of meromorphic functions $f_i$ in local coordinates near $P_i$, when can we find a global meromorphic function $f$ on $X$ with the given principal parts?

In terms of sheaves, we are now studying the short exact sequence

$$0 \to \mathcal{O} \to \mathcal{M} \to \mathcal{M}/\mathcal{O} \to 0.$$

Given given $f_i \in \mathcal{M}(U_i)$, we want to determine when $[f_i - f_j] \in H^1(X, \mathcal{O})$ is a coboundary.

Now by the case of Serre duality we have already proven, $H^1(X, \mathcal{O})$ is isomorphic to $\Omega(X)^*$, so there is a natural pairing between $(\delta f_i) \in H^1(X, \mathcal{O})$ and $\omega \in \Omega(X)$. In fact, by Theorem 9.1, this pairing is given by

$$\langle \delta f_i, \omega \rangle = \mathrm{Res}(\omega f_i).$$

A class in $H^1(X, \mathcal{O})$ vanishes iff it pairs trivially with all $\omega \in \Omega(X)$. Thus we have:

**Theorem 9.5** *The Mittag-Leffler problem specified by* $(f_i)$ *has a solution iff*

$$\sum \mathrm{Res}_p(f_i \omega) = 0$$

*for every* $\omega \in \Omega(X)$.

**Geometric Riemann–Roch.** Using the solution to the Mittag–Leffler problem, we get a geometric and natural version of the Riemann-Roch theorem for effective divisors. We will soon see that the same result holds without the assumption that $D \geq 0$.

**Theorem 9.6** *Suppose* $D \geq 0$. *Then*

$$h^0(D) = 1 - g + \deg(D) + h^0(K - D).$$

**Proof.** Let $T \cong \mathbb{C}^n$, $n = \deg D$, be the vector space of all Laurent tails associated to $D$. (More functorially, $T$ is the space of global sections of $\mathcal{O}_D \,/\, \mathcal{O}$; the latter is a generalized skyscraper sheaf, with the same support as $D$.) Taking into account the constant functions, we then have

$$H^0(X, \mathcal{O}_D)/\mathbb{C} \cong \Omega(X)^\perp \subset T,$$

using the residue pairing above. But the map $\Omega(X) \to T^*$ has a kernel, namely $\Omega_D(X)$, the space of holomorphic 1-forms that vanish to such higher order that they automatically cancel all the principal parts in $T$. This shows

$$\begin{aligned} h^0(D) - 1 &= \dim \Omega(X)^\perp = \dim T - \dim \Omega(X) + \dim \Omega_D(X) \\ &= \deg D - g + h^0(K - D), \end{aligned}$$

which gives the formula above. ∎

**Remarks.** The proof is quite effective; e.g. if we know the principal parts of a basis for $\Omega(X)$ at $P \in X$, then we can compute explicitly the Laurent tails of all meromorphic function in $H^0(X, \mathcal{O}_{nP})$.

The computation above can be summarized more functorially: for $D \geq 0$ we have a short exact sequence of sheaves,

$$0 \to \mathcal{O} \to \mathcal{O}_D \to \mathcal{O}_D/\mathcal{O} \to 0,$$

which gives rise to a long exact sequence, using duality:

$$0 \to \mathbb{C} \to H^0(\mathcal{O}_D) \to H^0(\mathcal{O}_D/\mathcal{O}) \to H^1(\mathcal{O}) \cong \Omega(X)^* \to \Omega_{-D}(X)^* \to 0.$$

Note that $H^0(\mathcal{O}_D\,/\mathcal{O})$ is exactly the space of principal parts supported on $D$; its dimension is $\deg D$.

To carry out this computation we needed the inclusion $\mathcal{O}_D \subset \mathcal{O}$, which is equivalent to the condition that $D \geq 0$. We remark that effective divisors are 'infinitely rare' among all divisors.

## 10 Serre duality

The goal of this section is to present:

**Theorem 10.1 (Serre duality)** *For any divisor $D$, we have a canonical isomorphism*

$$H^1(X, \mathcal{O}_{-D})^* \cong \Omega_D(X).$$

The natural isomorphism comes from the pairing

$$H^0(X, \Omega_D) \otimes H^1(X, \mathcal{O}_{-D}) \to H^1(X, \Omega) \cong \mathbb{C}.$$

This result allows us to eliminate $h^1$ entirely from the statement of Riemann-Roch, without the requirement that $D$ is effective, and obtain:

**Theorem 10.2 (Riemann-Roch, final version)** *For any divisor $D$ on a compact Riemann surface $X$, we have*

$$h^0(D) = \deg D - g + 1 + h^0(K - D).$$

**Proof of Serre duality: analysis.** A proof of Serre duality can be given using the elliptic regularity, just as we did for the case $\Omega(X) \cong H^1(X, \mathcal{O})^*$. For this one should either consider forms that are locally (smooth) + (meromorphic), *or* consider smooth sections of the line bundle defined by a divisor, which we will elaborate later. ∎

**Proof of Serre duality: dimension counts.** We first show

$$\Omega_{-D}(X) \to H^1(X, \mathcal{O}_D)^*$$

is injective. For this, given $\omega \in \Omega_{-D}(X)$ choose a small neighborhood $U_0$ centered at point $P$ with $U_0$ outside the support of $D$ and $(\omega)$, on which we have a local coordinate where $\omega = dz$. Then set $f_0 = 1/z$ on $U_0$ and $f_1 = 0$ on Then set $U_1 = X - \{P\}$. We then find $\delta f_i \in H^1(X, \mathcal{O}_D)^*$, and

$$\mathrm{Res}(f_i \omega) = 1,$$

so $\omega$ is nontrivial in the dual space.

This shows $h^0(K - D) \le h^1(D)$, or $h^0(K - D) - h^1(D) \le 0$. Now we sum the following two applications of Riemann-Roch:

$$\begin{aligned} h^0(K-D) - h^1(K-D) &= 1 - g + \deg(K - D) \\ h^0(D) - h^1(D) &= 1 - g + \deg(D) \end{aligned}$$

to get

$$(h^0(D) - h^1(K-D)) + (h^0(K-D) - h^1(D)) = 0$$

(using the fact that $\deg(K) = 2g - 2$). Since both terms on the left are $\le 0$, they must vanish, and this gives Serre duality. ■ (I am grateful to Levent

Alpoge for pointing out this last step in the argument.)

**Remark on injectivity.** It seems strange at first sight that, in our proof that

$$\Omega_{-D}(X) \hookrightarrow H^1(X, \mathcal{O}_D)^*,$$

almost no use was made of the divisor $D$. Why doesn't the same argument prove that $\Omega_{-E}(X) \hookrightarrow H^1(X, \mathcal{O}_D)^*$ for any $D$ and $E$?

The point is that we need the inclusion of $\Omega_{-D}(X)$ into $H^1(X, \mathcal{O}_D)$ to *be defined*, and for this it is essential that $\omega$ pairs trivially with *coboundaries* in the target. To get this fact we are really using $D$, so that the product goes into $H^1(X, \Omega) \cong \mathbb{C}$.

**Proof of Serre duality: using just finiteness.** Finally, here is an argument that just uses *finiteness* of $h^1(\mathcal{O})$. In particular it shows that finiteness implies $H^1(X, \mathcal{O}) \cong \Omega(X)^*$. It is based loosely on Forster.

We begin with some useful qualitative dimension counts that follow immediately from the fact that $h^0(D) = 0$ if $\deg(D) < 0$, $h^1(D) \ge 0$ and the fact that $\Omega \cong \mathcal{O}_K$.

**Theorem 10.3** *For* $\deg(D) > 0$, *we have*

$$\begin{aligned}\dim \mathcal{O}_D(X) &\geq \deg(D) + O(1), \\ \dim \Omega_D(X) &\geq \deg(D) + O(1), \quad \textit{and} \\ \dim H^1(X, \mathcal{O}_{-D}) &= \deg(D) + O(1).\end{aligned}$$

**Pairings.** Next we couple the product map

$$\mathcal{O}_{-D} \otimes \Omega_D \to \Omega$$

together with the residue map $\mathrm{Res} : H^1(X, \Omega) \to \mathbb{C}$ to obtain a map

$$H^1(X, \mathcal{O}_{-D}) \otimes \Omega_D(X) \to H^1(X, \Omega) \to \mathbb{C},$$

or equivalently a natural map

$$\Omega_D(X) \to H^1(X, \mathcal{O}_{-D})^*.$$

This map explicitly sends $\omega$ to the linear functional defined by

$$\phi(\xi) = \mathrm{Res}(\xi\omega).$$

Using the long exact sequence associated to equation (8.1) we obtain:

**Theorem 10.4** *The inclusion* $\mathcal{O}_D \to \mathcal{O}_{D+P}$ *induces a surjection:*

$$H^1(X, \mathcal{O}_D) \to H^1(X, \mathcal{O}_{D+P}) \to 0.$$

**Corollary 10.5** *We have a natural surjective map:*

$$H^1(X, \mathcal{O}_D) \to H^1(X, \mathcal{O}_E) \to 0$$

*for any* $E \geq D$.

Put differently, for $E \geq D$ we get a surjective map $H^1(X, \mathcal{O}_{-E}) \to H^1(X, \mathcal{O}_{-D})$ and thus an injective map on the level of duals. For organizational convenience we take the direct limit over increasing divisors and set

$$V(X) = \lim_{D \to +\infty} H^1(X, \mathcal{O}_{-D})^*.$$

Clearly $\Omega_D(X)$ maps into $V(X)$ for every $D$, so we get a map $\mathcal{M}^1(X) \to V(X)$.

**Theorem 10.6** *The natural map $\mathcal{M}^1(X) \to V(X)$ is injective. Moreover a meromorphic form $\omega$ maps into $H^1(X, \mathcal{O}_{-D})^*$ iff $\omega \in \Omega_D(X)$.*

**Proof.** We have already shown injectivity in the previous proof. Now for the second statement. If $\omega$ is in $H^1(X, \mathcal{O}_{-D})^*$, then it must vanish on all coboundaries for this group. Suppose however $-(k+1) = \mathrm{ord}_P(\omega) < -D(P)$ for some $P$. Then $k \geq D(P)$, so in the construction above we can arrange that $\xi = 0$ in $H^1(X, \mathcal{O}_{-D})$. This is a contradiction. Thus $\mathrm{ord}_P(\omega) \geq -D(P)$ for all $P$, i.e. $\omega \in \Omega_D(X)$. ∎

**Completion of the proof of Serre duality.** Note that both $\mathcal{M}^1(X)$ and $V(X)$ are vector spaces over the field of meromorphic functions $\mathcal{M}(X)$. (Indeed the former vector space is one-dimensional, generated by any meromorphic 1-form.)

Given $\phi \in H^1(X, \mathcal{O}_{-D})^* \subset V(X)$, we must show $\phi$ is represented by a meromorphic 1-form $\omega$. (By the preceding result, $\omega$ will automatically lie in $\Omega_D(X)$.)

The proof will be by a dimension count. We note that for $n \gg 0$, we have

$$\dim H^1(X, \mathcal{O}_{-D-nP})^* = n + O(1).$$

On the other hand, this space contains $\Omega_{D+nP}(X)$ as well as $\mathcal{O}_{nP}(X) \cdot \phi$. Both of these spaces have dimension bounded below by $n+O(1)$. Thus for $n$ large enough, they meet in a nontrivial subspace. This means we can write $f\phi = \omega$ where $f$ is a nonzero meromorphic function and $\omega$ is a meromorphic form. But then $\phi = \omega/f$ is also a meromorphic form, so we are done! ∎

We can now round out the discussion by proving some results promised above.

**Theorem 10.7** *For any divisor $D$ we have*

$$H^0(X, \mathcal{O}_D) \cong H^1(X, \Omega_{-D})^*.$$

**Proof.** We have

$$H^0(X, \mathcal{O}_D) \cong H^0(X, \Omega_{D-K}) \cong H^1(X, \mathcal{O}_{K-D})^* \cong H^1(X, \Omega_{-D})^*$$

∎

**Corollary 10.8** *We have* $H^1(X, \mathcal{O}_D) = 0$ *as soon as* $\deg(D) > \deg(K) = 2g - 2$.

**Proof.** Because then $H^1(X, \mathcal{O}_D)^* \cong \Omega_{-D}(X) = 0$. ∎

**Corollary 10.9** *If* $\deg(D) > 2g - 2$*, then* $h^0(D) = 1 - g + \deg(D)$.

**Corollary 10.10** *We have* $H^1(X, \mathcal{M}) = H^1(X, \mathcal{M}^1) = 0$.

**Proof.** Any representative cocycle $(f_{ij})$ for a class in $H^1(X, \mathcal{M})$ can be regarded as a class in $H^1(X, \mathcal{O}_D)$ for some $D$ of large degree. But $H^1(X, \mathcal{O}_D) = 0$ once $\deg(D)$ is sufficiently large. Thus $(f_{ij})$ splits for the sheaf $\mathcal{O}_D$, and hence for $\mathcal{M}$. ∎

**Mittag-Leffler for 1-forms, revisited.** Here is another proof of the Mittag-Leffler theorem for 1-forms. Suppose we consider all possible principal parts with poles of order at most $n_i > 0$ at $P_i$, and let $D = \sum n_i P_i$. The dimension of the space of principal parts is then $n = \sum n_i = \deg D$. In addition, the principal part determines the solution to the Mittag-Leffler problem up to adding a holomorphic 1-form. That is, the solutions lie in the space $\Omega_D(X)$, and the map to the principal parts has $\Omega(X)$ as its kernel.

Thus the dimension of the space of principal parts that have solutions is:

$$k = \dim \Omega_D(X) - \dim \Omega(X).$$

But by Riemann-Roch we have

$$\begin{aligned} \dim \Omega_D(X) &= h^0(K + D) = h^0(-D) + \deg(K + D) - g + 1 \\ &= 2g - 2 + \deg D - g - 1 = g + \deg D - 1 = g + n - 1, \end{aligned}$$

so the solvable data has dimension $k = n - 1$. Thus there is one condition on the principal parts for solvability, and that condition is given by the residue theorem.

# 11 Maps to projective space

In this section we explain the connection between the sheaves $\mathcal{O}_D$, linear systems and maps to projective space. In the next section we relate this classical theory to the modern concept of a holomorphic line bundle. Useful references for this material include [GH, Ch. 1].

Historically, algebraic curves in projective space were studied from an extrinsic point of view, i.e. as embedded subvarieties. From this perspective, it is not at all clear when the same intrinsic curve is being presented in two different ways. The modern perspective is to fix the intrinsic curve $X$ — which we will consider as a Riemann surface — and then study its various projective manifestations. The embedding of $X$ into $\mathbb{P}^n$ then becomes visible as the family of effective divisors — the linear system — coming from intersections of $X$ with hyperplanes in $\mathbb{P}^n$.

**Projective space.** Let $V$ be an $(n+1)$-dimensional vector space over $\mathbb{C}$. The space of lines (one-dimensional subspaces) in $V$ forms the *projective space*

$$\mathbb{P}V = (V - \{0\})/\mathbb{C}^*.$$

It has the structure of a complex $n$-manifold. The subspaces $S \subset V$ give rise to planes $\mathbb{P}S \subset \mathbb{P}V$; when $S$ has codimension one, $\mathbb{P}S$ is a *hyperplane*.

The *dual* projective space $\mathbb{P}V^*$ parameterizes the hyperplanes in $\mathbb{P}V$, via the correspondence $\phi \in V^* \mapsto \mathrm{Ker}(\phi) = S \subset V$.

We can also form the quotient space $W = V/S$. Any line $L$ in $V$ that is not entirely contained in $S$ projects to a line in $W$. Thus we obtain a natural map

$$\pi : (\mathbb{P}V - \mathbb{P}S) \to \mathbb{P}(V/S).$$

All the analytic automorphisms of projective space come from linear automorphisms of the underlying vector space: that is,

$$\mathrm{Aut}(\mathbb{P}V) = \mathrm{GL}(V)/\mathbb{C}^* = \mathrm{PGL}(V).$$

We let $\mathbb{P}^n = \mathbb{P}\mathbb{C}^{n+1}$ with homogeneous coordinates $[Z] = [Z_0 : \cdots : Z_n]$. It satisfies

$$\mathrm{Aut}(\mathbb{P}^n) = \mathrm{PGL}_{n+1}(\mathbb{C}).$$

The case $n = 1$ gives the usual identification of automorphisms of $\widehat{\mathbb{C}}$ with Möbius transformations.

**The Hopf fibration.** By considering the unit sphere in $\mathbb{C}^{n+1}$, we obtain the *Hopf fibration* $\pi : S^{2n+1} \to \mathbb{CP}^n$ with fibers $S^1$. This shows that projective space is compact. Moreover, for $n = 1$ the fibers of $\pi$ are linked circles in $S^3$, and $\pi$ generates $\pi_3(S^2) \cong \mathbb{Z}$.

**Metrics on projective space.** The compact group $\mathrm{SU}_{n+1}$ also acts transitively on projective space and has a unique invariant metric, up to scale, called the Fubini–Study metric. In this metric all linear subvarieties $\mathbb{P}^k \subset \mathbb{P}^n$ have the same $(2k)$–dimensional volume. In fact $\mathrm{SU}_{n+1}$ acts transitively on these subvarieties.

More remarkably, the volume of *any* irreducible subvariety of $\mathbb{P}^n$ is determined by its dimension and its degree. In particular, all curves of degree $d$ in $\mathbb{P}^2$ have the same volume. This is a basic property of *Kähler manifolds*; it comes from the fact that the symplectic form associated to the Fubini–Study metric is *closed*.

**Projective varieties.** The zero set of a homogeneous polynomial $f(Z)$ defines an *algebraic hypersurface* $V(f) \subset \mathbb{P}^n$. A *projective algebraic variety* is the locus $V(f_1, \ldots, f_n)$ obtained as an intersection of hypersurfaces.

**Affine charts and cohomology.** The locus $\mathbb{A}^n = \mathbb{P}^n - V(Z_0)$ is isomorphic to $\mathbb{C}^n$ with coordinates $(z_1, \ldots, z_n) = Z_i/Z_0$, while $V(Z_0)$ itself is a hyperplane; thus we have

$$\mathbb{P}^n \cong \mathbb{C}^n \cup \mathbb{P}^{n-1}.$$

By permuting the coordinates, we get a covering of $\mathbb{P}^n$ by $n+1$ affine charts. It is easy to see that the transition functions are algebraic. Thus $\mathbb{P}^n$ is a complex manifold.

These affine charts shows that as a topological space, $\mathbb{P}^n$ is obtained by adding a single cell in each even dimension $0, 1, \ldots, 2n$, and thus:

$$H^{2k}(\mathbb{P}^n, \mathbb{Z}) = \mathbb{Z} \quad \text{for } k = 0, 1, \ldots, n;$$

moreover $[\mathbb{P}^k]$ gives a canonical generator in dimension $2k$. The remaining cohomology groups are zero.

**The degree of a subvariety.** If $Z^k \subset \mathbb{P}^n$ is a variety of dimension $k$, then $Z^k$ is a topological cycle and its degree is defined so that:

$$[Z^k] = \deg(Z) \cdot [\mathbb{P}^k] \subset H^{2k}(\mathbb{P}^n, \mathbb{Z}).$$

Alternatively, $\deg(Z^k) = |\mathbb{P}^{n-k} \cap Z|$ is the number of intersections with a generic linear subspace of complementary dimension. In particular, if $Z = V(f)$ is a hypersurface and $f$ is separable (a product of distinct irreducibles), then

$$\deg(Z) = \deg(f).$$

**The space of hypersurfaces.** The space of homogeneous polynomials of degree $d$ in $\mathbb{P}^n$ has dimension given by

$$\dim P_d(\mathbb{C}^{n+1}) = \binom{d+n}{n}.$$

This can be seen by inserting $n$ markers into a list of $d$ symbols, and turning all the symbols up to the first marker into $Z_0$'s, then the next stretch into $Z_1$'s, etc.

Thus the space of curves of degree $d$ in $\mathbb{P}^2$ is itself a projective space $\mathbb{P}^N$, with $N = d(d+3)/2$. E.g. there is a 5-dimensional space of conics, a 9-dimensional space of cubics and a 14-dimensional space of quartics. Note that these spaces include reducible elements, e.g. the union of $d$ lines is an example of a (reducible) plane curve of degree $d$.

**Meromorphic (rational) functions on $\mathbb{P}^n$.** The ratio

$$F(Z) = f_1(Z)/f_0(Z) = [f_0(Z) : f_1(Z)]$$

of two homogeneous polynomials of the same degree defines a meromorphic 'function'

$$F : \mathbb{P}^n \dashrightarrow \mathbb{P}^1.$$

Its values are undetermined on the subvariety $V(f_0, f_1)$, which has codimension two if $f_0$ and $f_1$ are relatively prime. Away from this subvariety, $F(Z)$ gives a holomorphic map to $\mathbb{P}^1$. The field of all rational functions is

$$K(\mathbb{P}^n) \cong \mathbb{C}(z_1, \ldots, z_n).$$

**Projectivization.** Any ordinary polynomial $p(z_i)$ has a unique homogeneous version $P(Z_i)$ of the same maximal degree, such that $p(z_i) = P(1, z_1, \ldots, z_n)$. Thus any affine variety $V(p_1, \ldots, p_m)$ has a natural completion $V(P_1, \ldots, P_m) \subset \mathbb{P}^m$. This variety is *smooth* if it is a smooth submanifold in each affine chart.

**Examples: Curves in $\mathbb{P}^2$.**

1. The affine curve $x^2 - y^2 = 1$ meets the line at infinity in two points corresponding to its two asymptotes; its homogenization $X^2 - Y^2 = Z^2$ is smooth in every chart. In $(y, z)$ coordinates it becomes $1 - y^2 = z^2$, which means the line at infinity $(z = 0)$ in two points.

2. The affine curve defined by $p(x, y) = y - x^3 = 0$ is smooth in $\mathbb{C}^2$, but its homogenization $P(X, Y, Z) = YZ^2 - X^3$ defines the curve $z^2 = x^3$ in the affine chart where $Y \neq 0$, which has a cusp.

**The normalization of a singular curve.** Every *irreducible* homogeneous polynomial $f$ on $\mathbb{P}^2$ determines a compact Riemann surface $X$ together with a generically injective map $\nu : X \to V(f)$. The Riemann surface $X$ is called the *normalization* of the (possibly singular) curve $V(f)$.

To construct $X$, use projection from a typical point $P \in \mathbb{P}^2 - V(f)$ to obtain a surjective map $\pi : V(f) \to \mathbb{P}^1$. After deleting a finite set from domain and range, including all the singular points of $V(f)$, we obtain an *open Riemann surface* $X^* = V(f) - C$ and a degree $d$ *covering map*

$$\pi : X^* \to (\mathbb{P}^1 - B).$$

As we have seen, there is a canonical way to complete $X^*$ and $\pi$ to a compact Riemann surface $X$ and a branched covering $\pi : X \to \mathbb{P}^1$. It is then easy to see that the isomorphism

$$\nu : X^* \to V(f) - C \subset \mathbb{P}^2$$

extends to a holomorphic map $\nu : X \to \mathbb{P}^2$ with image $V(f)$.

**Examples: cusps and nodes.** The cuspidal cubic $y^2 = x^3$ is normalized by $\nu : \mathbb{P}^1 \to \mathbb{P}^2$ given by $\nu(t) = (t^2, t^3)$.

The nodal cubic $y^2 = x^2(1-x)$ is also interesting to uniformize. The idea is to project from the node at $(x, y) = (0, 0)$, i.e. let us use the parameter $t = y/x$ or equivalently set $y = tx$. Then we find

$$t^2 x^2 = x^2(1 - x)$$

and so $(x, y) = (1 - t^2, t(1 - t^2)$ give a parameterization of this curve.

**Maps to $\mathbb{P}^n$.** Now let $X$ be a compact Riemann surface. Generalizing the study of meromorphic functions $f : X \to \mathbb{P}^1$, we wish to describe all holomorphic maps

$$\phi : X \to \mathbb{P}^n.$$

There are at least two ways to approach the description of $\phi$: using $\mathcal{M}(X)$, and using $\mathrm{Div}(X)$.

**Maps from meromorphic functions.** Let $S \subset \mathcal{M}(X)$ be a linear subspace of meromorphic functions of dimension $n + 1$. We have a *natural* map

$$\phi : X \to \mathbb{P}S^*$$

given by $\phi(x)(f) = f(x)$. More concretely, if we choose a basis for $S$ then the map is given by

$$\phi(x) = [f_0(x) : \cdots : f_n(x)].$$

The linear functional $x \mapsto f(x)$ in $S^*$ makes sense at most points of $X$; the only potential problems come from zeros and poles. To handle these, we simply choose a local coordinate $z$ near $x$ with $z(x) = 0$, and define

$$\phi(x)(f) = (z^p f(z))|_{z=0},$$

where $p$ is chosen so all values are finite and so at least one value is nonzero.

Note that the map $\phi$ is *nondegenerate*: its image is contained in no hyperplane.

Note also that if we replace $S$ with $fS$, where $f \in \mathcal{M}^*(X)$, then the map $\phi$ remains the same. Thus we can always normalize so that $1 \in S$.

**Linear systems of divisors.** We now turn to the approach via divisors. Recall that divisors $D, E$ are *linearly equivalent* if $D - E = (f)$ is principal; then $\mathcal{O}_D \cong \mathcal{O}_E$.

The *complete linear system* $|D|$ determined by $D$ is the collection of *effective* divisors $E$ linearly equivalent to $D$. These are exactly the divisors of the form $E = (f) + D$, where $f \in H^0(X, \mathcal{O}_D)$. The divisor only depends on the line determined by $f$. We have a natural bijection

$$|D| \cong \mathbb{P}\mathcal{O}_D(X).$$

In particular, we have

$$\dim |D| = h^0(D) - 1.$$

Technically $|D|$ is a subset of the space of all divisors on $X$. Equivalent divisors determine the *same* linear system, i.e.

$$|E| = |D| \subset \mathrm{Div}_d^+(X) \cong X^{(d)}.$$

We emphasize that all divisors in $|D|$ are effective and of the same degree.

A general linear system is simply a subspace $L \cong \mathbb{P}^r \subset |D|$. It corresponds to a subspace $S \subset H^0(X, \mathcal{O}_D)$.

**Base locus and geometric linear systems.** The *base locus* of a linear system is the largest divisor $B \geq 0$ such that $E \geq B$ for all $E \in L$. A linear system is *basepoint free* if $B = 0$.

Let $X \subset \mathbb{P}^n$ be a smooth nondegenerate curve, and let $H$ be a hyperplane. Then $D = H \cap X$ is an effective divisor called a *hyperplane section*. Any two such divisors are *linearly equivalent*, because $H_1 - H_2 = (f)$ where $f : \mathbb{P}^n \dashrightarrow \mathbb{P}^1$ is a rational map of degree one (a ratio of linear forms), and $f|X$ is a meromorphic function.

Thus the set of hyperplane sections determines a linear system $L \subset |D|$. It is *basepoint free* because for any $P \in X$ there is an $H$ disjoint from $P$.

On the other hand, if we take only those hyperplanes passing through $P \in X$, then we get a smaller linear system whose base locus is $P$ itself.

**Different perspectives on maps to projective space.** The following data, up to suitable equivalences, are in natural bijection:

1. Nondegenerate holomorphic maps $\phi : X \to \mathbb{P}^n$ of degree $d$; up to $\mathrm{Aut}(\mathbb{P}^n)$;

2. Subspaces $S \subset \mathcal{M}(X)$ of dimension $(n+1)$ with $\deg(f/g) = d$ for generic $f, g \in S$; up to $S \mapsto fS$;

3. Basepoint-free linear systems $L \subset \mathrm{Div}(X)$ of degree $d$ and dimension $n$.

Here are some of the relations between them.

$1 \implies 2$. Let $S$ be spanned by the meromorphic function $f_i = Z_i/Z_0$, restricted to $X$. Note that $f/g$ is the projection of $\phi(X)$ to $\mathbb{P}^1$ from a generic $\mathbb{P}^{n-2}$, which will have degree $d$ so long as the $\mathbb{P}^{n-2}$ is disjoint from $\phi(X)$.

$1 \implies 3$. We let $D = X \cap H$ and let $L \subset |D|$ be the collection of all hyperplane sections.

$2 \implies 1$. We have seen that $S$ canonically determines a map to $\mathbb{P}S^*$, and clearly $fS$ determines an equivalent map.

$2 \implies 3$. Let $D$ be the smallest divisor such that $D + (f)$ is effective for all (nonzero) $f \in S$. This is given, concretely, by $D(x) = -\inf_S \mathrm{ord}_f(x)$. Then we consider the linear system of divisors of the form $E = (f) + D$ such that $f \in S$.

$3 \implies 2$. Picking any $D \in L$, we let $S = \{f : (f) = E - D, E \in L\}$.

$3 \implies 1$. For each $x \in X$ we get a hyperplane in $L$ by considering the set of $E$ containing $x$. This gives a natural map $X \to \mathbb{P}^n \cong L^*$.

**Embeddings and smoothness for complete linear systems.** We now study the complete linear system $|D|$, and the associated map

$$\phi_D : X \to \mathbb{P}H^0(X, \mathcal{O}_D)^*$$

in more detail.

Recall that $\mathcal{O}_D$ is a $\mathcal{O}$-module. We say $\mathcal{O}_D$ is *generated by global sections* if for each $x \in X$ we have a global section $f \in H^0(X, \mathcal{O}_D)$ such that the stalk $\mathcal{O}_{D,x}$, which is an $\mathcal{O}_x$-module, is generated by $f$; that is, $\mathcal{O}_{D,x} = \mathcal{O}_x \cdot f$.

**Theorem 11.1** *The following are equivalent.*

1. $|D|$ *is basepoint–free.*

2. $h^0(D - P) = h^0(D) - 1$ *for all* $P \in X$.

3. $\mathcal{O}_D$ *is generated by global sections.*

4. *For any* $x \in X$ *there is an* $f \in \mathcal{O}_D(X)$ *such that* $\mathrm{ord}_x(f) = -D(x)$.

Example. Note that $h^0(nP) = 1$ for $n = 0$ and $= g$ for $n = 2g - 1$. Thus (for large genus) there must be values of $n$ such that $h^0(nP) = h^0(nP - P)$. In this case, $D = nP$ is *not* globally generated.

**Theorem 11.2** *Let $\phi : X \to \mathbb{P}^n$ be a map of $X$ to projective space with the image not contained in a hyperplane. Then $\phi = \pi \circ \phi_D$, where $D$ is the divisor of a hyperplane section, $|D|$ is a base-point free linear system, and where $\pi : \mathbb{P}^N \dashrightarrow \mathbb{P}^n$ is projection from a linear subspace $\mathbb{P}S$ disjoint from $\phi_D(X)$.*

**Proof.** Since hyperplanes can be moved, the linear system $|D|$ is base-point free, and all hyperplane sections are linearly equivalent to $D$. Thus $\phi$ can be regarded as the natural map from $X$ to $\mathbb{P}S^*$, where $S$ is a subspace of $H^0(X, \mathcal{O}_D)$, so $\phi$ can be factored through the map $\phi_D$ to $\mathbb{P}H^0(X, \mathcal{O}_D)^*$. ∎

**Theorem 11.3** *Let $|D|$ be base-point free. Then for any effective divisor $P = P_1 + \ldots + P_n$ on $X$, the linear system*

$$P + |D - P| \subset |D|$$

*consists exactly of the hyperplane sections $E = \phi_D^{-1}(H)$ passing through $(P_1, \ldots, P_n)$. In particular, the dimension of the space of such hyperplanes is*

$$\dim |D - P| = h^0(D - P) - 1.$$

**Proof.** A hyperplane section $E \in |D|$ passes through $P$ iff $E - P \geq 0$ iff $E = P + E'$ where $E' \in |D - P|$. ∎

**Theorem 11.4** *If $h^0(D - P - Q) = h^0(D) - 2$ for any $P, Q \in X$, then $|D|$ provides a smooth embedding of $X$ into projective space.*

**Proof.** This condition says exactly that the set of hyperplanes passing through $\phi_D(P)$ and $\phi_D(Q)$ has dimension 2 less than the set of all hyperplanes. Thus $\phi_D(P) \neq \phi_D(Q)$, so $\phi_D$ is $1-1$. The condition on $\phi(D - 2P)$ says that the set of hyperplanes containing $\phi_D(P)$ and $\phi'_D(P)$ is also 2 dimensions less, and thus $\phi_D$ is an immersion. Thus $\phi_D$ is a smooth embedding. ∎

**Theorem 11.5** *The linear system $|D|$ is base-point free if $\deg D \geq 2g$, and gives an embedding into projective space if $\deg D \geq 2g+1$.*

**Proof.** Use the fact that if $\deg D > \deg K = 2g-2$, then we have $h^0(D) = \deg D - g + 1$, which is linear in the degree. ∎

**Corollary 11.6** *Every compact Riemann surface embeds in projective space.*

**Remark.** By projecting we get an embedding of $X$ into $\mathbb{P}^3$ and an immersion into $\mathbb{P}^2$.

Not every Riemann surface can be embedded into the plane! In fact a smooth curve of degree $d$ has genus $g = (d-1)(d-2)/2$, so for example there are no curves of genus 2 embedded in $\mathbb{P}^2$.

**Examples of linear systems.**

1. *Genus 0.* On $\mathbb{P}^1$, we have $\mathcal{O}_D \cong \mathcal{O}_E$ if $d = \deg(D) = \deg(E)$. This sheaf is usually referred to as $\mathcal{O}(d)$; it is well-defined up to isomorphism, $|D|$ consists of all effective divisors of degree $d$. Note that $\mathcal{O}_{d\infty}(\mathbb{P}^1)$ is the $d+1$-dimensional space of polynomials of degree $d$. The corresponding map $\phi_D : \mathbb{P}^1 \to \mathbb{P}^d$ is given in affine coordinates by $\phi_D(t) = (t, t^2, \ldots, t^d)$. Its image is the *rational normal curve* of degree $d$. Particular cases are smooth conics and the twisted cubic.

2. *Genus 1.* On $X = \mathbb{C}/\Lambda$ with $P = 0$, the linear system $|P|$ is *not* base-point free, but $|2P|$ is, and $|3P|$ gives an embedding into the plane, via the map $z \mapsto (\wp(z), \wp'(z))$.

   Recall that $D = \sum m_i P_i$ is a principal divisor on $X$ iff $\deg(D) = 0$ and $e(D) = \sum m_i P_i$ in $\mathbb{C}/\Lambda$ is zero. Thus $3Q \in |3P|$ iff $e(3(P-Q)) = 0$. This shows:

   > *A smooth cubic curve $X$ in $\mathbb{P}^2$ has 9 flex points, corresponding to the points of order 3 in the group law on $X$.*

# 12 The canonical map

Note that the embedding of a curve $X$ of genus 1 into $\mathbb{P}^2$ by the linear series $|3P|$ *breaks* the symmetry group of the curve: since $\mathrm{Aut}(X)$ acts transitively, the 9 flexes of $Y = \phi_{3P}(X)$ are not intrinsically special.

For genus $g \geq 2$ on the other hand there is a more natural embedding — one which does not break the symmetries of $X$ — given by the canonical linear system $|K|$.

The linear system $|K|$ corresponds to the map

$$\phi_K : X \to \mathbb{P}\Omega(X)^* \cong \mathbb{P}^{g-1}$$

given by $\phi(x) = [\omega_1(x), \ldots, \omega_g(x)]$, where $\omega_i$ is a basis for $\Omega(X)$.

**Theorem 12.1** *(1) The linear system $|K|$ is base-point free for $g > 0$.*

*(2) For $g \geq 1$, either $|K|$ gives an embedding of $X$ into $\mathbb{P}^{g-1}$, or $X$ is hyperelliptic.*

**Proof.** Let $P_k \geq 0$ be an effective divisor of degree $k$ on $X$. Then

$$h^0(K - P_k) = 1 - g + (2g - 2) - k + h^0(P_k).$$

We have $1 \leq h^0(P_k)$ for all $k$. If equality holds then we have

$$h^0(K - P_k) = h^0(K) - k,$$

which is what we want to get a basepoint free linear system ($k = 1$) and an embedding ($k = 2$). Thus $|K|$ is basepoint free iff $h^0(P_1) = 1$ iff $g > 0$, and $|K|$ gives an embedding iff $h^0(P_2) = 1$ iff $X$ is not hyperelliptic. ∎

**The hyperelliptic case.** Let us analyze the canonical map in the hyperelliptic case. Then there is a degree two holomorphic map $f : X \to \mathbb{P}^1$ branched over the zeros of a polynomial $p(z)$ of degree $2g + 2$. A basis for the holomorphic 1-forms on $X$ is given by

$$\omega_i = \frac{z^i\, dz}{\sqrt{p(z)}}$$

for $i = 0, \ldots, g - 1$. That is, $\omega_i(x) = f(x)^i \omega_0$. It follows that the canonical map $\phi : X \to \mathbb{P}^{g-1}$ is given by

$$f(x) = [\omega_i(x)] = [\omega_0(x) f(x)^i] = [f(x)^i].$$

In other words, the canonical map factors as $\phi = \psi \circ f$, where $\psi : \mathbb{P}^1 \to \mathbb{P}^{g-1}$ is the rational normal curve of degree $g - 1$. ∎

**Geometry of canonical divisors.** With the canonical map in hand, it is easy to visualize at once all the effective canonical divisors $K$ on $X$, and hence all the elements of $\Omega(X)$. There are just two cases:

1. The curve $X \subset \mathbb{P}^{g-1}$ is embedded by the canonical map as a curve of degree $2g-2$. Then the hyperplane sections $K = X \cap H$ are exactly the canonical divisors, i.e. these are all the elements of the complete linear system $|K|$.

2. The canonical maps factors through degree two map $\pi : X \to \mathbb{P}^1$. Then the canonical divisors are the preimages $K = \pi^{-1}(E)$ of divisors of degree $g-1$ on $\mathbb{P}^1$.

In the second, hyperelliptic case, the image of the canonical map is the rational normal curve $Z \subset \mathbb{P}^{g-1}$, which has degree $(g-1)$, and the canonical map is the composition

$$\phi : X \xrightarrow{\pi} \mathbb{P}^1 \cong Z \subset \mathbb{P}^{g-1}.$$

Note that any set of $g-1$ points on $Z$ span a unique hyperplane $H$, so even in the hyperelliptic case we can regard the canonical divisors as simple the divisors of the form $\phi^{-1}(H)$.

**Geometric picture for $h^0(K-D)$.** Let $D = \sum_1^n P_i$ be an effective divisor on $X$. The canonical curve proves a clear picture of the 'hard term' $h^0(K-D)$ in Riemann–Roch, in terms of the space of hyperplanes containing all these points $P_i$; namely,

$$\begin{aligned} h^0(K-D)-1 &= \dim|K-D| \\ &= \dim\{H \subset \mathbb{P}^{g-1} \ : \ \phi(P_i) \in H, \ 1 \le i \le n\}. \end{aligned}$$

**Special divisors.** An effective divisor $D = \sum P_i$ is *special* if $h^0(K-D) > 0$, i.e. if there is a holomorphic 1-form $\omega \neq 0$ vanishing at the points $(P_i)$, or equivalent if the points $(P_i)$ lie on a hyperplane in $\mathbb{P}^{g-1}$.

Of course divisors of degree $\le g-1$ are all special. The first interesting case is degree $g$. There exist plenty of such divisors – note that $|H \cap \phi(X)| = 2g-2$, so a given hyperplane determines many such divisors. On the other hand, $g$ typical points on $\phi(X)$ do *not* span a hyperplane, so these divisors really are *special*.

**Theorem 12.2** *An effective divisor $D$ of degree $g$ is special if and only if there is a nonconstant meromorphic function on $X$ with $(f) + D \ge 0$.*

**Proof.** By Riemann-Roch, we have

$$h^0(D) = 1 - g + \deg D + h^0(K - D) = 1 + h^0(K - D) > 1$$

iff $h^0(K - D) > 0$ iff $D$ is special. ∎

Since special advisors do exist, we have:

**Corollary 12.3** *If $g \geq 2$ then $X$ admits a meromorphic function of degree $\leq g$.*

**Example: genus 3.** A curve of genus 3 either admits a map to $\mathbb{P}^1$ of degree two, or it embeds as a curve $X \subset \mathbb{P}^2$ of degree 4. In the latter case, projection to $\mathbb{P}^1$ from any $P \in X$ gives a map of degree $g = 3$.

**Explicit construction of meromorphic functions.** The map $f : X \to \mathbb{P}^1$ determined by a special divisor $D$ of degree $g$ can be constructed geometrically as follows. Let $H$ be a hyperplane containing $D$, and let $H \cap X = D + E$. Then $\deg(E) = g - 2$, so $E$ lies in linear subspace $J \cong \mathbb{P}^{g-3}$. Then projection of $\mathbb{P}^{g-1}$ from $J$ to $\mathbb{P}^1$ gives a map $f : X \to \mathbb{P}^1$ with $D$ as one of its fibers.

**Weierstrass points.** Now we focus on divisors of the form $D = gP$. We say $P$ is a *Weierstrass point* if $gP$ is special.

**Proposition 12.4** *The following are equivalent:*

1. *$P$ is a Weierstrass point.*
2. *There is a hyperplane $H$ such that $H \cap \phi(X)$ has multiplicity $g$ at $\phi(P)$.*
3. *There is an $\omega \in \Omega(X)$ vanishing to order $g$ at $P$.*
4. *There is a meromorphic function $f$ on $X$ with a pole just at $P$ and with $1 \leq \deg(f) \leq g$.*

and otherwise holomorphic.

Example: there are no Weierstrass points on a Riemann surface of genus 1. The branch points of every hyperelliptic surface of genus $g \geq 2$ are Weierstrass points.

The map $f$ can be constructed as follow: choose a subspace $J \cong \mathbb{P}^{g-3}$ in $\mathbb{P}^{g-1}$ passing through the $g - 2$ points of $H \cap \phi(X)$ other than $P$. Then let $f$ be the projection of $X$ from $J$ to $\mathbb{P}^1$, normalized so $H \cap X$ maps to $\infty$. Then $f$ has a pole of order $\geq g$ at $P$ and nowhere else.

**The Wronskian.** To have $H \cap \phi(X) = D = gP$, the hyperplane $H$ should contain not just $P$ but the appropriate set of tangent directions at $P$, namely

$$(\phi(P), \phi'(P), \phi''(P), \ldots, \phi^{(g-1)}(P)).$$

For these tangent $(g-1)$ tangent directions to span a $(g-2)$-dimensional plane $H$ through $\phi(P)$, there must be a linear relation among them; that is, the *Wronskian determinant* $W(P)$ must vanish.

In terms of a basis for $\Omega$ and a local coordinate $z$ at $P$, the Wronskian is given by

$$W(z) = \det \left( \frac{d^j \omega_i}{dz^j} \right),$$

where $j = 0, \ldots, g-1$ and $i = 1, \ldots, g$.

We can see directly the vanishing of the Wronskian is equivalent to $gP$ being special.

**Theorem 12.5** *The Wronskian vanishes at $P$ iff there is a holomorphic 1-form $\omega \neq 0$ with a zero at $P$ of order at least $g$.*

**Proof.** The determinant vanishes iff there is a linear combination of the basis elements $\omega_i$ whose derivatives through order $(g-1)$ vanish at $P$. ∎

The quantity $W = W(z)\,dz^N$ turns out to be independent of the choice of coordinate, where $N = 1 + 2 + \ldots + g = g(g+1)/2$. This value of $N$ arises because the $j$th derivative of a 1-form behaves like $dz^{j+1}$.

Thus $W(z)$ is a section of $\mathcal{O}_{NK}$, so its number of zeros is $\deg NK = N(2g-2) = (g-1)g(g+1)$. This shows:

**Theorem 12.6** *Any Riemann surface of genus $g$ has $(g-1)g(g+1)$ Weierstrass points, counted with multiplicity.*

**Weierstrass points of a hyperelliptic curve.** These correspond to the branch points of the hyperelliptic map $\pi : X \to \mathbb{P}^1$, since the projective normal curve $\mathbb{P}^1 \to \mathbb{P}^{g-1}$ has no flexes.

**Canonical curves of genus two.** We will now describe more geometrically the canonical curves of genus two, three and four.

**Theorem 12.7** *Any Riemann surface $X$ of genus 2 is hyperelliptic, and any degree two map of $X$ to $\mathbb{P}^1$ agrees with the canonical map (up to $\operatorname{Aut} \mathbb{P}^1$).*

**Proof.** In genus 2, we have $\deg K = 2g - 2 = 2$, so the canonical map $\phi : X \to \mathbb{P}^{g-1} = \mathbb{P}^1$ already presents $X$ as a hyperelliptic curve. If $f : X \to \mathbb{P}^1$ is another such map, with polar divisor $P + Q$, then we have $h^0(P + Q) = 2 = h^0(K - P - Q) + 2 - 2 + 1$; thus there exists an $\omega$ with zeros just at $P, Q$ and therefore $P + Q$ is a canonical divisor. ∎

**Corollary 12.8** *The moduli space of curves of genus two is isomorphic to the 3-dimensional space $\mathcal{M}_{0,6}$ of isomorphism classes of unordered 6-tuples of points on $\mathbb{P}^1$. Thus $\mathcal{M}_2$ is finitely covered by $\mathbb{C}^3 - D$, where $D$ consists of the hyperplanes $x_i = 0$, $x_i = 1$ and $x_i = x_j$.*

**Remark.** Here is a topological fact, related to the fact that every curve of genus two is hyperelliptic: if $S$ has genus two, then the center of the mapping-class group $\operatorname{Mod}(S)$ is $\mathbb{Z}/2$, generated by any hyperelliptic involution. (In higher genus the center of $\operatorname{Mod}(S)$ is trivial.)

**Canonical curves of genus 3.** Let $X$ be a curve of genus 3. Then $X$ is either hyperelliptic, or its canonical map realizes it as a smooth plane quartic. We will later see that, conversely, any smooth quartic is a canonical curve (we already know it has genus 3). This shows:

**Theorem 12.9** *The moduli space of curves of genus 3 is the union of the 5-dimensional space $\mathcal{M}_{0,8}$ and the 6-dimensional moduli space of smooth quartics, $(\mathbb{P}^{14} - D)/\operatorname{PGL}_3(\mathbb{C})$.*

Note: a smooth quartic curve that degenerates to a hyperelliptic one becomes a double conic. The eight hyperelliptic branch points can be thought of as the intersection of this conic with an infinitely near quartic curve.

**Weierstrass points in genus 3.** The Weierstrass points on a smooth plane quartic correspond to flexes; there are $2 \cdot 3 \cdot 4 = 24$ of them in general.

At the flexes we have $h^0(3P) = 2$. How can one go from a flex $P$ to a degree 3 branched covering $f : X \to \mathbb{P}^1$? We can try projection $f_P$ from $P$, but in general the line $L$ tangent to $X$ at $P$ will meet $X$ in a fourth point $Q$. Thus $f_P$ will have a double pole at $P$ and a simple pole at $Q$.

Instead, we project from $Q$! Then the line $L$ through $P$ and $Q$ has multiplicity 3 at $P$, giving a triple order pole there.

**The Fermat quartic.** The *Fermat quartic*, defined by

$$X^4 + Y^4 + Z^4 = 0,$$

is somewhat similar; its symmetry group is $\mathrm{PSL}_2(\mathbb{Z}/8)$, which has order 96; the quotient is the $(2, 3, 8)$ orbifold, of Euler characteristic $-1/24$, and this curve is tiled by 12 octagons. In this case the full symmetry group is easily visible: there is an $S^3$ coming from permutations of the coordinates, and a $(\mathbb{Z}/4)^2$ action coming from multiplication of $X$ and $Y$ by powers of $\sqrt{-1}$.

On the Fermat curve, given in affine coordinates by $x^4 + y^4 = 1$, there are 12 flexes altogether, each of multiplicity 2. Of these, 8 lie in the affine plane, and arise when one coordinate vanishes and the other is a 4th root of unity.

These are examples of Weierstrass points of *order two*; a line tangent to one of these flexes meets the curve in a *unique point* $P$. In this case, projection from $P$ itself really does give a rational map $f : X \to \mathbb{P}^1$ with a triple pole just at $P$.

**Flexes of plane curves.** In general, if $C$ is defined by $F(X, Y, Z) = 0$, then the flexes of $C$ are the locus where both $F$ and the Hessian $H = \det(d^2F/dZ_i\, dZ_j)$ of $F$ vanish. For the Fermat curve, we have $F(X, Y, Z) = X^4 + Y^4 + Z^4$ and $H = 1728(XYZ)^2$. On a smooth curve of degree $d$ the number of flexes is $3d(d - 2)$.

**The Klein quartic.** The most symmetric curve of genus 3 is the *Klein quartic*, given by

$$X^3Y + Y^3Z + Z^3X = 0.$$

Its symmetry group $G \cong \mathrm{PSL}_2(\mathbb{Z}/7)$ has order $168 = 7 \cdot 24$ and gives as quotient the $(2, 3, 7)$-orbifold. A subgroup of order 21 is easily visible in the form above: we can cyclically permute the coordinates and also send $(X, Y, Z)$ to $(\zeta^2X, \zeta Y, \zeta^4Z)$ where $\zeta^7 = 1$. More geometrically, the Klein quartic is tiled by 24 regular 7-gons. It is easily shown that $-1/42 = \chi(S^2(2, 3, 7))$ is the largest possible Euler characteristic for a hyperbolic orbifold, which implies:

**Theorem 12.10 (Hurwitz)** *For any compact Riemann surface of genus $g$ we have $|\mathrm{Aut}(X)| \le 84(g - 1)$.*

Equality is achieved in the case above.

**Canonical curves of genus 4.** Now let $X \subset \mathbb{P}^3$ be a canonical curve of genus 4 (in the non-hyperelliptic case). Then $X$ has degree 6.

**Theorem 12.11** *$X$ is the intersection of an irreducible quadric and cubic hypersurface in $\mathbb{P}^3$.*

**Proof.** The proof is by dimension counting again. There is a natural linear map from $\mathrm{Sym}^2(\Omega(X))$ into $H^0(X, \mathcal{O}_{2K})$. Since $\dim \Omega(X) = 4$, the first space has dimension $\binom{3+2}{2} = 10$, while the second has dimension $3g - 3 = 9$ by Riemann-Roch. Thus there is a nontrivial quadratic equation $Q(\omega_1, \ldots, \omega_4) = 0$ satisfied by the holomorphic 1-forms on $X$; equivalent $X$ lies on a quadric. The quadric is irreducible because $X$ does *not* lie on a hyperplane. Moreover, it is unique: if $X$ lies on 2 quadrics, then it is a curve of degree 4, not 6.

Carrying out a similar calculation for degree 3, we find $\dim \mathrm{Sym}^3(\Omega(X)) = \binom{3+3}{3} = 20$ while $\dim H^0(X, \mathcal{O}_{3K}) = 5g - 5 = 15$. Thus there is a 5-dimensional space of cubic relations satisfied by the $(\omega_i)$. In this space, a 4-dimensional subspace is accounted for by the product of $Q$ with an arbitrary linear equation. Thus there must be, in addition, an irreducible cubic surface containing $X$. ∎

We will later see that the converse also holds.

**Dimension counts for linear systems.** Here is another perspective on the preceding proof. Intersection of surfaces of degree $d$ in $\mathbb{P}^3$ with $X$ gives a birational map between projective spaces,

$$|dH| \to |dK|.$$

Now note that in general a linear map $\phi : A \to B$ between vector spaces gives a birational map

$$\Phi : \mathbb{P}A \dashrightarrow \mathbb{P}B$$

which is projection from $\mathbb{P}C$ where $C = \mathrm{Ker}\,\phi$. Then $\dim A - \dim B \le \dim C$ and thus

$$\dim \mathbb{P}C \ge \dim \mathbb{P}A - \dim \mathbb{P}B - 1.$$

In the case at hand $\mathbb{P}C$ corresponds to the linear system $S_d(X)$ of surfaces of degree $d$ containing $X$. Thus we get:

$$\dim S_d(X) \ge \dim |dH| - \dim |dK| - 1.$$

For $d = 2$ this gives

$$\dim S_2(X) \ge 9 - (3g - 4) - 1 = 9 - 8 - 1 = 0$$

which shows there is a unique quadric $Q$ containing $X$. For $d = 3$ we get

$$\dim S_3(X) \ge 19 - (5g - 6) - 1 = 19 - 14 - 1 = 4.$$

Within $S_3(X)$ we have $Q + |H|$ which is 3-dimensional, and thus there must be a cubic surface $C$ not containing $Q$ in $S_3(X)$ as well.

This cubic is *not* unique, since we can move it in concert with $Q + H$.

**The dimension of moduli space $\mathcal{M}_g$ : Riemann's count.** What is the dimension of $\mathcal{M}_g$? We know the dimension is 0, 1 and 3 for genus $g = 0$, 1 and 2 (using 6 points on $\mathbb{P}^1$ for the last computation).

Here is Riemann's heuristic. Take a large degree $d \gg g$, and consider the bundle $\mathcal{F}_d \to \mathcal{M}_g$ whose fibers are meromorphic functions $f : X \to \mathbb{P}^1$ of degree $d$. Now for a fixed $X$, we can describe $f \in \mathcal{F}_d(X)$ by first giving its polar divisor $D \geq 0$; then $f$ is a typical element of $H^0(X, \mathcal{O}_D)$. (The parameters determining $f$ are its principal parts on $D$.) Altogether with find

$$\dim \mathcal{F}_d(X) = d + h^0(D) = 2d - g + 1.$$

On the other hand, $f$ has $b$ critical points, where

$$\chi(X) = 2 - 2g = 2d - b,$$

so $b = 2d + 2g - 2$. Assuming the critical values are distinct, they can be continuously deformed to determine new branched covers $(X', f')$. Thus the dimension of the total space is given by

$$b = \dim \mathcal{F}_d = \dim \mathcal{F}_d(X) + \dim \mathcal{M}_g = 2d + 2g - 2 = 2d - g + 1 + \dim \mathcal{M}_g,$$

and thus $\dim \mathcal{M}_g = 3g - 3$. This dimension is in fact correct.

On the other hand, the space of hyperelliptic Riemann surfaces clearly satisfies

$$\dim \mathcal{H}_g = 2g + 2 - \dim \operatorname{Aut} \mathbb{P}^1 = 2g - 1,$$

since such a surface is branched over $2g+2$ points. Thus for $g > 2$ a typical Riemann surface is not hyperelliptic.

**Tangent space to $\mathcal{M}_g$.** As an alternative to Riemann's count, we note that the tangent space to the deformations of $X$ is $H^1(X, \Theta)$, where $\Theta \cong \Omega^*$ is the sheaf of holomorphic vector fields on $X$. By Serre duality and Riemann-Roch, we have

$$\dim H^1(X, \Theta) = \dim H^1(X, \mathcal{O}_{-K}) = h^0(2K) = 4g - 4 - g + 1 = 3g - 3.$$

Serre duality also shows $H^1(X, \Theta)$ is naturally dual to the space of holomorphic quadratic differentials $Q(X)$.

**Plane curves again.** The space of homogeneous polynomials on $\mathbb{C}^{n+1}$ of degree $d$ has dimension $N = \binom{n+d}{n}$. Thus the space of plane curves of degree $d$, up to automorphisms of $\mathbb{P}^2$, has dimension

$$N_d = \binom{2+d}{d} - 9.$$

We find

$$N_d = \begin{cases} -3 = \dim \operatorname{Aut} \mathbb{P}^1 & \text{for } d = 2, \\ 1 = \dim \mathcal{M}_1 & \text{for } d = 3, \\ 6 = \dim \mathcal{M}_3 & \text{for } d = 4, \\ 12 < \dim \mathcal{M}_6 = 15 & \text{for } d = 5. \end{cases}$$

Thus most curves of genus 6 *cannot* be realized as plane curves. In a sense made precise by the Theorem below, there is no way to simply parameterize the moduli space of curves of high genus:

**Theorem 12.12 (Harris-Mumford)** *For $g$ sufficiently large, $\mathcal{M}_g$ is of general type.*

In fact $g \geq 24$ will do.

# 13 Line bundles

Let $X$ be a *complex manifold.* A *line bundle* $\pi : L \to X$ is a 1-dimensional holomorphic vector bundle.

This means there exists a collection of trivializations of $L$ over charts $U_i$ on $X$, say $L_i \cong U_i \times \mathbb{C}$. Viewing $L$ in two different charts, we obtain clutching data $g_{ij} : U_i \cap U_j \to \mathbb{C}^*$ such that $(x, y_j) \in L_j$ is equivalent to $(x, y_i) \in L_i$ iff $g_{ij}(x) y_j = y_i$. This data satisfies the cocycle condition $g_{ij} g_{jk} = g_{ik}$.

In terms of charts, a *holomorphic section* $s : X \to L$ is encoded by holomorphic functions $s_i = y_i \circ s(x)$ on $U_i$, such that

$$s_i(x) = g_{ij}(x) s_j(x).$$

**Examples:** the trivial bundle $X \times \mathbb{C}$; the canonical bundle $\wedge^n \mathrm{T}^* X$. Here the transition functions are $g_{ij} = 1$ for the trivial bundle and $g_{ij} = 1/\det D(\phi_i \circ \phi_j^{-1})$ for the canonical bundle, with charts $\phi_i : U_i \to \mathbb{C}^n$.

In detail, on a Riemann surface $X$, with local coordinates $z_i : U_i \to \mathbb{C}$, a section of the canonical bundle is locally given by $\omega_i = s_i(z)\, dz_i$; it must satisfy $s_i(z)\, dz_i = s_j(z)\, dz_j$, so $s_i = (dz_j/dz_i) s_j$.

**Tensor powers.** From $L$ we can form the line bundle $L^* = L^{-1}$, and more generally $L^d$, with transition functions $g_{ij}^d$.

A line bundle is *trivial* if it admits a nowhere-vanishing holomorphic section (which then provides an isomorphism between $L$ and $X \times \mathbb{C}$). Such a section exists iff there are $s_i \in \mathcal{O}^*(U_i)$ such that $s_i/s_j = g_{ij}$, i.e. iff $g_{ij}$ is a coboundary.

Thus line bundles up to isomorphism over $X$ are classified by the cohomology group $H^1(X, \mathcal{O}^*)$.

**Sections and divisors.** Now consider a divisor $D$ on a Riemann surface $X$. Then we can locally find functions $s_i \in \mathcal{M}(U_i)$ with $(s_i) = D$. From this data we construction a line bundle $L_D$ with transition functions $g_{ij} = s_i/s_j$. These transitions functions are chosen so that $s_i$ is automatically a *meromorphic section* of $L_D$; indeed, a holomorphic section if $D$ is effective.

**Theorem 13.1** *The sheaf of holomorphic sections $\mathcal{L}$ of $L = L_D$ is isomorphic to $\mathcal{O}_D$.*

**Proof.** Choose a meromorphic section $s : X \to L$ with $(s) = D$. (On a compact Riemann surface, $s$ is well-defined up to a constant multiple.) Then a local section $t : U \to L$ is holomorphic if and only if the meromorphic function $f = t/s$ satisfies

$$(t) = (fs) = (f) + D \geq 0$$

on $U$, which is exactly the condition that $f \in \mathcal{O}_D$. Thus the map $f \mapsto fs$ gives an isomorphism between $\mathcal{O}_D(U)$ and $\mathcal{L}(U)$. ∎

**Line bundles on Riemann surfaces.** Conversely, it can be shown that every line bundle $L$ on a Riemann surface admits a non-constant meromorphic section, and hence $L = L_D$ for some $D$. More precisely, if $\mathcal{L}$ is the sheaf of sections of $L$ one can show (see e.g. Forster, Ch. 29):

**Theorem 13.2** *The group $H^1(X, \mathcal{L})$ is finite-dimensional.*

**Corollary 13.3** *Given any $P \in X$, there exists a meromorphic section $s : X \to L$ with a pole of degree $\leq 1 + h^1(\mathcal{L})$ at $P$ and otherwise holomorphic.*

**Corollary 13.4** *Every line bundle has the form $L \cong L_D$ for some divisor $D$.*

From the point of view of sheaf theory, we have

$$0 \to \mathcal{O}^* \to \mathcal{M}^* \to \mathrm{Div} \to H^1(X, \mathcal{O}^*) \to H^1(X, \mathcal{M}^*) \to 0,$$

and since every line bundle is represented by a divisor, we find:

**Corollary 13.5** *The group $H^1(X, \mathcal{M}^*) = 0$.*

**Divisors and line bundles in higher dimensions.** On complex manifolds of higher dimension, we can similarly construct line bundles from divisors. First, a divisor is simply an element of $H^0(\mathcal{M}^*/\mathcal{O}^*)$; this means it is locally a formal sum of analytic hypersurfaces, $D = \sum(f_i)$. Then the associated transition functions are $g_{ij} = f_i/f_j$ as before, and we find:

**Theorem 13.6** *Any divisor $D$ on a complex manifold $X$ determines a line bundle $L_D \to X$ and a meromorphic section $s : X \to L$ with $(s) = D$.*

**Failure of every line bundle to admit a nonzero section.** However in general not every line bundle arises in this way. For example, there exist complex 2-tori $M = \mathbb{C}^2/\Lambda$ with no divisors but with plenty of line bundles (coming from characters $\chi : \pi_1(M) \to S^1$).

**Degree.** The degree of a line bundle, $\deg(L)$, is the degree of the divisor of any meromorphic section.

In terms of cohomology, the degree is associated to the exponential sequence: we have

$$\ldots H^1(X,\mathbb{Z}) \to H^1(X,\mathcal{O}) \to H^1(X,\mathcal{O}^*) \to H^2(X,\mathbb{Z}) \cong \mathbb{Z}.$$

This allows one to define the degree or *first Chern class*, $c_1(L) \in H^2(X,\mathbb{Z})$, for a line bundle on any complex manifold.

**Projective space.** For projective space, we have $H^1(\mathbb{P}^n,\mathbb{C}) = 0$, and in fact $H^1(\mathbb{P}^n,\mathcal{O}) = H^{0,1}(\mathbb{P}^n) = 0$. (This follows from the Hodge theorem, which implies that

$$H^n(X,\mathbb{C}) = \oplus_{p+q=n} H^{p,q}(X).$$

This is nontrivial for the case of $\mathbb{P}^n$, $n \geq 2$, because elements of $H^{0,1}$ are represented by $\overline{\partial}$-closed forms, rather than simply *all* $(0,1)$-forms as in the case of a Riemann surface).

It follows that line bundles on projective space are *classified by their degree*: we have

$$0 \to H^1(\mathbb{P}^n,\mathcal{O}^*) \to H^2(\mathbb{P}^n,\mathbb{Z}) \cong \mathbb{Z}.$$

We let $\mathcal{O}(d)$ denote the (sheaf of sections) of the unique line bundle of degree $d$. It has the property that for any meromorphic section $s \in H^0(\mathbb{P}^n,\mathcal{O}(d))$, the divisor $D = (s)$ represents $d[H] \in H^2(\mathbb{P}^n,\mathbb{Z})$, where $H \cong \mathbb{P}^{n-1}$ is a hyperplane.

**Example: the tautological bundle.** Let $\mathbb{P}^n$ be the projective space of $\mathbb{C}^{n+1}$ with coordinates $Z = (Z_0,\ldots,Z_n)$. The *tautological bundle* $\tau \to \mathbb{P}^n$

has, as its fiber over $p = [Z]$, the line $\tau_p = \mathbb{C} \cdot Z \subset \mathbb{C}^{n+1}$. Its total space is $\mathbb{C}^{n+1}$ with the origin blown up.

To describe $\tau$ in terms of transition functions, let $U_i = (Z_i \neq 0) \subset \mathbb{P}^n$. Then we can use the coordinate $Z_i$ itself to trivialize $\tau|U_i$; in other words, we can map $\tau|U_i$ to $U_i \times \mathbb{C}$ the map

$$(Z_0, \ldots, Z_n) \mapsto ([Z_0 : \cdots : Z_n], Z_i).$$

(Here the origin must be blown up.) Then clearly the transition functions are given simply by $g_{ij} = Z_i/Z_j$, since they satisfy

$$Z_i = g_{ij} Z_j.$$

**Theorem 13.7** *There is no nonzero holomorphic section of the tautological bundle.*

**Proof.** A section gives a map $s : \mathbb{P}^n \to \tau \to \mathbb{C}^{n+1}$ which would have to be constant because $\mathbb{P}^n$ is compact. But then the constant must be zero, since this is the only point in $\mathbb{C}^{n+1}$ that lies on every line through the origin. ■

As a typical *meromorphic* section, we can define $s(p)$ to be the intersection of $\tau_p$ with the hyperplane $Z_0 = 1$. In other words,

$$s([Z_0 : Z_1 : \cdots : Z_n]) = (1, Z_1/Z_0, \ldots, Z_n/Z_0).$$

Then $s_i = Z_i/Z_0$. Notice that this section is nowhere vanishing (since $Z_i$ has no zero on $U_i$), but it has a pole along the divisor $H_0 = (Z_0)$.

Thus we have $\tau \cong \mathcal{O}(-1)$. Similarly, $\tau^* = \mathcal{O}(1)$.

**Homogeneous polynomials.** Note that the coordinates $Z_i$ are sections of $\mathcal{O}(1)$. Indeed, any element in $V^*$ naturally determines a function on the tautological bundle over $\mathbb{P}V$, linear on the fibers, and hence a section of the dual bundle. Similarly we have:

**Theorem 13.8** *The space of global sections of $\mathcal{O}(d)$ over $\mathbb{P}^n$ can be naturally identified with the homogeneous polynomials of degree $d$ on $\mathbb{C}^{n+1}$.*

One can appeal to Hartog's extension theorem to show *all* global sections have this form.

**Corollary 13.9** *The hypersurfaces of degree $d$ in projective space are exactly the zeros of holomorphic sections of $\mathcal{O}(d)$.*

**The canonical bundle.** To compute the canonical bundle of project space, we use the coordinates $z_i = Z_i/Z_0$, $i = 1, \ldots, n$ to define a nonzero canonical form

$$\omega = dz_1 \cdots dz_n$$

on $U_0$. To examine this form in $U_1$, we use the coordinates $w_1 = Z_0/Z_1$, $w_i = Z_i/Z_1$, $i > 1$; then $z_1 = 1/w_1$ and $z_i = w_i/w_1$, $i > 1$, so we have

$$\omega = d(1/w_1)\, d(w_2/w_1) \cdots d(w_n/w_1) = -(dw_1 \cdots dw_n)/w_1^{n+1}.$$

Thus $(\omega) = (-n-1)H_0$ and thus the canonical bundle satisfies $K \cong \mathcal{O}(-n-1)$ on $\mathbb{P}^n$.

**The adjunction formula.**

**Theorem 13.10** *Let $X \subset Y$ be a smooth hypersurface inside a complex manifold. Then the canonical bundles satisfy*

$$K_X \cong (K_Y \otimes L_X)|X.$$

**Proof.** We have an exact sequence of vector bundles on $X$:

$$0 \to \mathrm{T}X \to \mathrm{T}Y \to \mathrm{T}Y/\,\mathrm{T}X = NX \to 0,$$

where $NX$ is the normal bundle. Now $(NX)^*$ is the sub-bundle of $\mathrm{T}^*Y|X$ spanned by 1-forms that annihilate $TX$. If $X$ is defined in charts $U_i$ by $f_i = 0$, then $g_{ij} = f_i/f_j$ defines $L_X$. On the other hand, $df_i$ is a nonzero holomorphic section of $(NX)^*$. The 1-forms $df_i|X$, however, do not fit together on overlaps to form a global section of $(NX)^*$. Rather, on $X$ we have $f_j = 0$ so

$$df_i = d(g_{ij} f_j) = g_{ij} df_j.$$

This shows $(df_i)$ gives a global, nonzero section of $(NX)^* \otimes L_X$, and hence this bundle is trivial on $X$.

On the other hand $K_Y = K_X \otimes (NX)^*$, by taking duals and determinants. Thus $K_X = K_Y \otimes NX = K_Y \otimes L_X$. ■

**Smooth plane curves.** Using the adjunction formula plus Riemann-Roch we can obtain some interesting properties of smooth plane curves $X \subset \mathbb{P}^2$ of degree $d$.

**Theorem 13.11** *Let $f : X \to \mathbb{P}^n$ be a holomorphic embedding. Then $f$ is given by a subspace of sections of the line bundle $L \to X$, where $L = f^*\mathcal{O}(1)$.*

**Proof.** The divisors of section of $\mathcal{O}(1)$ are hyperplanes. ■

**Theorem 13.12** *Every smooth plane curve of degree $d$ has genus $g = (d-1)(d-2)/2$.*

**Proof.** We have $K_X \cong K_{\mathbb{P}^2} \otimes L_X = \mathcal{O}(d-3)$. Any curve $Z$ of degree $d-3$ is the zero set of a section of $\mathcal{O}(d-3)$ and hence restricts to the zero set of a holomorphic 1-form on $X$. Thus we find $2g-2 = d(d-3)$. ∎

**Corollary 13.13** *Every smooth quartic plane curve $X$ is a canonical curve.*

**Proof.** We have $K_X \cong \mathcal{O}(d-3) = \mathcal{O}(1)$, which is the linear system that gives the original embedding of $X$ into $\mathbb{P}^2$. ∎

Next note that the genus $g(X) = (d-1)(d-2)/2$ coincides with the dimension of the space of homogeneous polynomials on $\mathbb{C}^3$ of degree $d-3$. This shows:

**Theorem 13.14** *Every effective canonical divisor on $X$ has the form $K = X \cap Y$, where $Y$ is a curve of degree $d-3$.*

**Explicit computation of $\boldsymbol{\Omega(X)}$.** If we unwind the proof the adjunction formula, we get a useful algorithm for computing $\Omega(X)$ for a plane curve. Let us work in affine coordinates $(x,y)$ on $\mathbb{C}^2 \subset \mathbb{P}^2$, and suppose $X$ has degree $d$ and is defined by $f(x,y) = 0$. Let $W(x,y)$ be a polynomial of degree $d-3$. We can then solve the equation

$$\omega \wedge df = W(x,y)\, dx\, dy$$

to obtain 1-form $\omega$ on $\mathbb{P}^2$. The pullback $\omega|TX$ gives a 1-form on $X$.

Note: $\omega$ is only well-defined up to adding a multiple of $df$; but $df|TX = 0$, so $\omega$ gives a well-defined 1-form on $X$.

Now the form $df$ does not vanish along $X$, since $X$ is smooth. Thus the zeros of $\omega \wedge df$ coincide with the zeros of $\omega|TX$, as well as with the zeros of $W(x,y)$. Since the curves defined by $W$ and $f$ meet in $d(d-3) = 2g-2$ points, the divisor $(\omega)$ must coincide with these points; hence $\omega|TX$ is a holomorphic 1-form. By varying $W(x,y)$ we get all such forms.

**Examples.** For the cubic curve $y^2 = x^3 - 1$ we have $d-3 = 0$, so $W(x,y) = a$ and we need to solve

$$\omega \wedge (2y\, dy - 3x^2\, dx) = a dx\, dy.$$

Clearly $\omega = dx/y$ does the job.

For the quartic curve $x^4 + y^4 = 1$, we have $W(x, y) = (ax + by + c)$, and we need to solve

$$\omega \wedge (4x^3\, dx + 4y^3\, dy) = (ax + by + c)\, dx\, dy.$$

Clearly $\omega = dy/x^3$ works for $W(x, y) = c$, and then $x\omega$ and $y\omega$ give the other 2 forms spanning $\Omega(X)$.

**Gonality of plane curves.** What is the minimum degree of a map of a plane curve to $\mathbb{P}^1$?

**Theorem 13.15** *Any $n + 1$ distinct points in $\mathbb{P}^2$ impose independent condition on curves of degree $n$.*

**Proof.** Choose $Y$ to be the union of $n$ random lines through the first $k \le n$ points. Then $Y$ is an example of a curve through the first $k$ points that does not pass through the $k + 1$st. This shows that adding the $k + 1$st point imposes an additional condition on $Y$. ■

**Theorem 13.16** *A smooth curve $X$ of degree $d > 1$ admits a nonconstant map to $\mathbb{P}^1$ of degree $d - 1$, but none of degree $d - 2$.*

**Proof.** For degree $d - 1$, simply projection from a point on $X$. For the second assertion, suppose $f : X \to \mathbb{P}^1$ has degree $e \le d - 2$. Let $E \subset X$ be a generic fiber of $f$. Then the $d - 2$ points $E$ impose independent conditions on the space of curves $Y$ degree $d - 3$. Consequently

$$h^0(K - E) = g - \deg E.$$

By Riemann-Roch we then have:

$$h^0(E) = 1 - g + \deg(E) - h^0(K - E) = 1,$$

so $|E|$ does not provide a map to $\mathbb{P}^1$. ■

**Hypersurfaces in products of projective spaces.** Here are two further instances of the adjunction theorem.

**Theorem 13.17** *Every smooth degree 6 intersection $X$ of a quadric $Q$ and a cubic surface $C$ is a canonical curve in $\mathbb{P}^3$.*

**Proof.** It can be shown that $Q$ and $C$ are smooth and transverse along $X$ (this is nontrivial). We then have $K_Q \cong K_{\mathbb{P}^3} \otimes L_Q$ and thus

$$K_X \cong K_Q \otimes L_C \cong K_{\mathbb{P}^3} \otimes L_Q \otimes L_C \cong \mathcal{O}(-4+2+3) = \mathcal{O}(1).$$

■

**Theorem 13.18** *Every smooth* $(d, e)$ *curve on* $Q = \mathbb{P}^1 \times \mathbb{P}^1$ *has genus* $g = (d-1)(e-1)$.

**Proof.** It is easy to see that $K_{X\times Y} = K_X \otimes K_Y$. Thus $K_Q = \mathcal{O}(-2, -2)$. Therefore $2g - 2 = C \cdot K_C$ and $K_C = \mathcal{O}(d-2, e-2)$, so $2g-2 = d(e-2) + e(d-2)$, which implies the result. ■

**Canonical curves of genus 5.** By a calculation with Riemann-Roch, one can show that any canonical curve $X \subset \mathbb{P}^4$ is an intersection of 3 quadrics. Conversely, a complete intersection of 3 quadrics is a canonical curve, because its canonical bundle is $\mathcal{O}(-5+2+2+2) = \mathcal{O}(1)$.

In general, the canonical bundle of a complete intersection $X$ of hypersurfaces of degrees $d_i$ in $\mathbb{P}^n$ is given by

$$K_X \cong \mathcal{O}(-n-1+\sum d_i).$$

**K3 surfaces.** Manifolds with trivial canonical bundle are often interesting — in higher dimensions they are called *Calabi-Yau* manifolds.

For Riemann surfaces, $K_X$ is trivial iff $X$ is a complex torus. For 2-dimensional manifolds, complex tori also have trivial canonical, but they are not the only examples. Another class is provided by the *K3 surfaces*, which by definition are simply-connected complex surfaces with $K_X$ trivial. Example:

**Theorem 13.19** *Every smooth surface of degree* 4 *in* $X = \mathbb{P}^3$ *and of degree* $(2, 2, 2)$ *in* $X = \mathbb{P}^1 \times \mathbb{P}^1 \times \mathbb{P}^1$ *is a K3 surface.*

**Proof.** Here $K_X = \mathcal{O}(-3)$ or $K_X = \mathcal{O}(-2, -2, -2)$, so the canonical bundle is trivial. By the Lefschetz hyperplane theorem, a hypersurface in a simply-connected complex 3-manifold is always itself simply-connected. ■

# 14 Curves and their Jacobians

We now turn to the important problem of classifying line bundles on a Riemann surface $X$; equivalently, of classifying divisors modulo linear equivalence.

**The Jacobian.** Recall that a holomorphic 1-form on $X$ is the same thing as a holomorphic map $f : X \to \mathbb{C}$ well-defined up to translation in $\mathbb{C}$. If the periods of $f$ happen to generate a discrete subgroup $\Lambda$ of $\mathbb{C}$, then we can regard $f$ as a map to $\mathbb{C}/\Lambda$. However the periods are almost always indiscrete. Nevertheless, we can put all the 1-forms together and obtain a map to $\mathbb{C}^g/\Lambda$.

**Theorem 14.1** *The natural map $H_1(X, \mathbb{Z}) \to \Omega(X)^*$ has as its image a lattice $\Lambda \cong \mathbb{Z}^{2g}$.*

**Proof.** If not, the image lies in a real hyperplane defined, for some nonzero $\omega \in \Omega(X)$, by the equation $\operatorname{Re} \alpha(\omega) = 0$. But then all the periods of $\operatorname{Re} \omega$ vanish, which implies the harmonic form $\operatorname{Re} \omega = 0$. ■

The *Jacobian variety* is the quotient space $\operatorname{Jac}(X) = \Omega(X)^*/H_1(X, \mathbb{Z})$, the cycles embedded via periods.

**Theorem 14.2** *Given any basepoint $P \in X$, there is a natural map $\phi_P : X \to \operatorname{Jac}(X)$ given by $\phi_P(Q) = \int_P^Q \omega$.*

We will later show this map is an embedding, and thus $\operatorname{Jac}(X)$, roughly speaking, makes $X$ into a group.

**Example: the pentagon.** Let $X$ be the hyperelliptic curve defined by $y^2 = x^5 - 1$. Geometrically, $X$ can be obtained by gluing together two regular pentagons. Clearly $X$ admits an automorphism $T : X \to X$ of order 5. Using the pentagon picture, one can easily show there is a cycle $C \in H_1(X, \mathbb{Z})$ such that its five images $T^i(C)$ span $H_1(X, \mathbb{Z}) \cong \mathbb{Z}^4$. This means $H_1(X, \mathbb{Z}) = A \cdot C$ is a free, rank one $A$-module, where $A = \mathbb{Z}[T]/(1 + T + T^2 + T^3 + T^4)$.

Now $T^*$ acts on $\Omega(X)$. Since $X/\langle T \rangle$ has genus zero, $T$ has no invariant forms. Thus we can choose a basis $(\omega_1, \omega_2)$ for $\Omega(X)$ such that

$$T^* \omega_i = \zeta_i \omega_i$$

where $\zeta_i$ is a primitive 5th root of unity.

Let us scale these $\omega_i$ so $\int_C \omega_i = 1$. Let

$$\pi : H_1(X, \mathbb{Z}) \to \Lambda \subset \mathbb{C}^2$$

be the period map, defined by

$$\pi(B) = \left( \int_B \omega_1, \int_B \omega_2 \right).$$

Since

$$\int_{T(B)} \omega = \int_B T^* \omega,$$

we have

$$\pi(TB) = \begin{pmatrix} \zeta_1 & 0 \\ 0 & \zeta_2 \end{pmatrix} \pi(B).$$

Since $H_1(X, \mathbb{Z}) = \mathbb{Z}[T] \cdot C$, we find that $\Lambda \subset \mathbb{C}^2$ is simply the image of the ring $\mathbb{Z}[T]$ under the ring homomorphism that sends $T$ to $(\zeta_1, \zeta_2)$.

Since $\Lambda$ is a lattice, we cannot have $\zeta_2 = \overline{\zeta}_1 = \zeta_1^4$, nor can we have $\zeta_2 = \zeta_1$. Thus $\zeta_2 = \zeta_1^2$ or $\zeta_1^3$. In the latter case we can interchange the eigenforms to obtain the former case. This finally shows:

**Theorem 14.3** *The Jacobian of the hyperelliptic curve $y^2 = x^5 - 1$ is isomorphic to $\mathbb{C}^2/\Lambda$, where $\Lambda$ is the ring $\mathbb{Z}[T]/(1+T+T^2+T^3+T^4)$ embedded in $\mathbb{C}^2$ by*

$$T \mapsto (\zeta, \zeta^2),$$

*and $\zeta$ is a primitive 5th root of unity.*

This is an example of a Jacobian variety with *complex multiplication.*

**The Picard group.** Let $\mathrm{Pic}(X)$ denote the group of all line bundles on $X$. Since every line bundle admits a meromorphic section, there is a natural isomorphism between $\mathrm{Pic}(X)$ and $\mathrm{Div}(X)/(\mathcal{M}^*(X))$, where $\mathcal{M}^*(X)$ is the group of nonzero meromorphic functions, mapping to principal divisors in $\mathrm{Pic}(X)$. Under this isomorphism, the degree and a divisor and of a line bundle agree.

The *Abel-Jacobi map* $\phi : \mathrm{Div}_0(X) \to \mathrm{Jac}(X)$ is defined by

$$\phi\left(\sum Q_i - P_i\right)(\omega) = \sum \int_{P_i}^{Q_i} \omega.$$

Because of the choice of path from $P_i$ to $Q_i$, the resulting linear functional is well-defined only modulo cycles on $X$.

For example: given any basepoint $P \in X$, we obtain a natural map $f : X \to \mathrm{Jac}(X)$ by $f(Q) = \phi(Q - P) = \int_P^Q \omega$.

One of the most basic results regarding the Jacobian is:

**Theorem 14.4** *The map $\phi$ establishes an isomorphism between* $\mathrm{Jac}(X)$ *and* $\mathrm{Pic}_0(X) = \mathrm{Div}_0(X)/$*(principal divisors).*

The proof has two parts: *Abel's theorem*, which asserts that $\mathrm{Ker}(\phi)$ coincides with the group of principal divisors, and the *Jacobi inversion theorem*, which asserts that $\phi$ is surjective.

**Theorem 14.5 (Abel's theorem)** *A divisor $D$ is principal iff $D = \sum Q_i - P_i$ and*

$$\sum \int_{P_i}^{Q_i} \omega = 0$$

*for all $\omega \in \Omega(X)$, for some choice of paths $\gamma_i$ joining $P_i$ to $Q_i$.*

**Proof in one direction.** Suppose $D = (f)$. We can assume after multiplying $f$ by a scalar, that none of its critical values are real. Let $\gamma = f^{-1}([0, \infty])$. Then we have

$$\sum \int_{P_i}^{Q_i} \omega = \int_\gamma \omega = \int_0^\infty f_*(\omega) = 0$$

since $f_*(\omega) = 0$, being a holomorphic 1-form on $\widehat{\mathbb{C}}$. (To see this, suppose locally $f(z) = z^d$. Then $f_*(z^i\,dz) = 0$ unless $z^i\,dz$ is invariant under rotation by the $d$th roots of unity. This first happens when $i = d-1$, in which case $z^{d-1}\,dz = (1/d)d(z^d)$, so the pushforward is proportional to $dz$.) ∎

**The curve in its Jacobian.** Before proceeding to the proof of Abel's theorem, we derive some consequences.

Given $P \in X$, define

$$\phi_P : X \to \mathrm{Jac}(X)$$

by $\phi_P(Q) = (Q - P)$. Note that with respect to a basis $\omega_i$ for $\Omega(X)$, the derivative of $\phi_P(Q)$ in local coordinates is given by:

$$D\phi_P(Q) = (\omega_1(Q), \ldots, \omega_g(Q)).$$

This shows:

**Theorem 14.6** *The canonical map $X \to \mathbb{P}\Omega(X)^*$ is the Gauss map of $\phi_P$.*

**Theorem 14.7** *For genus $g \geq 1$, the map $\phi_P : X \to \mathrm{Jac}(X)$ is a smooth embedding.*

**Proof.** If $Q - P = (f)$, then $f : X \to \mathbb{P}^1$ has degree 1 so $g = 0$. Since $|K|$ is basepoint-free, there is a nonzero-holomorphic 1-form at every point, and hence $D\phi_P \neq 0$. ∎

**Theorem 14.8 (Jacobi)** *The map* $\mathrm{Div}_0(X) \to \mathrm{Jac}(X)$ *is surjective. In fact, given* $(P_1, \ldots, P_g) \in X^g$*, the map*

$$\phi : X^g \to \mathrm{Jac}(X)$$

*given by*

$$\phi(Q_1, \ldots, Q_g) = \phi\left(\sum Q_i - P_i\right) = \left(\int_{P_i}^{Q_i} \omega_j\right)$$

*is surjective.*

**Proof.** It suffices to show that $\det D\phi \neq 0$ at some point, so the image is open. To this end, just note that $d\phi/dQ_i = (\omega_j(Q_i))$, and thus $\det D\phi(Q_1, \ldots, Q_g) = 0$ if and only if there is an $\omega$ vanishing simultaneously at all the $Q_i$, i.e. iff $(Q_i)$ lies on a hyperplane under the canonical embedding. For generic $Q_i$'s this will not be the case, and hence $\det D\phi \neq 0$ almost everywhere on $X^g$. ∎

**Corollary 14.9** *We have a natural isomorphism:*

$$\mathrm{Pic}_0(X) = \mathrm{Div}_0(X)/(\mathcal{M}^*(X)) \cong \mathrm{Jac}(X).$$

**Bergman metric.** We remark that the space $\Omega(X)$, and hence its dual, carries a natural norm given by:

$$\|\omega\|_2^2 = \int_X |\omega(z)|^2 \, |dz|^2.$$

This induces a canonical metric on $\mathrm{Jac}(X)$, and hence on $X$ itself.

This metric ultimately comes from the intersection pairing or *symplectic form* on $H_1(X, \mathbb{Z})$, satisfying $a_i \cdot b_j = \delta_{ij}$.

**Abel's theorem, proof I: The $\overline{\partial}$-equation.** (Cf. Forster.) For the converse, we proceed in two steps. First we will construct a smooth solution to $(f) = D$; then we will correct it to become holomorphic.

**Smooth solutions.** Let us say a *smooth* map $f : X \to \widehat{\mathbb{C}}$ satisfies $(f) = D$ if near $P_i$ (resp. $Q_i$) we have $f(z) = zh(z)$ (resp. $z^{-1}h(z)$) where $h$ is a smooth function with values in $\mathbb{C}^*$, and if $f$ has no other zeros or poles.

Note that for such an $f$, the *distributional* logarithmic derivative satisfies

$$\overline{\partial} \log f = \frac{\overline{\partial} f}{f} + \sum (Q_i' - P_i'),$$

where $Q'_i$ and $P'_i$ are $\delta$ functions (in fact currents), locally given by $\overline{\partial} \log z$, and $\overline{\partial} f/f$ is smooth.

Inspired by the proof in one direction already given, we first construct a smooth solution of $(f) = D$ which maps a disk neighborhood $U_i$ of $\gamma_i$ diffeomorphicly to a neighborhood $V$ of the interval $[0, \infty]$.

**Lemma 14.10** *Given any arc $\gamma$ joining $P$ to $Q$ on $X$, there exists a smooth solution to $(f) = Q - P$ satisfying*

$$\frac{1}{2\pi i} \int_X \frac{\overline{\partial} f}{f} \wedge \omega = \int_P^Q \omega$$

*for all $\omega \in \Omega(X)$, where the integral is taken along $\gamma$.*

**Proof.** First suppose $P$ and $Q$ are close enough that they belong to a single chart $U$, and $\gamma$ is almost a straight line. Then we can choose the isomorphism $f : U \to V \subset \widehat{\mathbb{C}}$ so that $f(P) = \infty$, $f(Q) = 0$ and $f(\gamma) = [0, \infty]$ (altering $\gamma$ by a small homotopy rel endpoints). This $f$ is already holomorphic on $U$, and it sends $\partial U$ to a contractible loop in $\mathbb{C}^*$. Thus we can extend $f$ to a smooth function sending $X - U$ into $\mathbb{C}^*$, which then satisfies $(f) = D$.

Now note that $z$ admits a single-valued logarithm on the region $\widehat{\mathbb{C}} - [0, \infty]$, and thus $\log f(z)$ has a single-valued branch on $Y = X - \gamma$. Thus $df/f = d \log f$ is an exact form on $Y$ (and we only need the $\overline{\partial}$ part when we wedge with $\omega$). However as one approaches $\gamma$ from different sides, the two branches of $\log f$ differ by $2\pi i$. Applying Stokes' theorem, we find:

$$\int_X (\overline{\partial} \log f) \wedge \omega = 2\pi i \int_\gamma \omega.$$

To handle the case of well-separated $P$ and $Q$, simply break $\gamma$ up into many small segments and take the product of the resulting $f$'s. ∎

Taking the product of the solutions for several pairs of points, and using additivity of the logarithmic derivative, we obtain:

**Corollary 14.11** *Given an arc $\gamma_i$ joining $P_i$ to $Q_i$ on $X$, there exists a smooth solution to $(f) = \sum Q_i - P_i$ satisfying*

$$\frac{1}{2\pi i} \int_X \frac{\overline{\partial} f}{f} \wedge \omega = \sum \int_{\gamma_i} \omega$$

*for all $\omega \in \Omega(X)$.*

**From smooth to holomorphic.** To complete the solution, it suffices to find a smooth function $g$ such that $fe^g$ is meromorphic. Equivalently, it suffices to solve the equation $\overline{\partial} g = -\overline{\partial} f/f$. (Note that $\overline{\partial} f/f$ is smooth even at the zeros and poles of $f$, since near there $f = z^n h$ where $h \neq 0$.) Since $H^{0,1}(X) \cong \Omega(X)^*$, such a $g$ exists iff

$$\frac{1}{2\pi i}\int(\overline{\partial}\log f)\wedge\omega\sum\int_{P_i}^{Q_i}\omega = 0$$

for all $\omega \in \Omega(X)$. This is exactly the hypothesis of Abel's theorem. ∎

**The symplectic form on $H^1$.** For the second proof of Abel's theorem, we need to make the connection between the Jacobian and the symplectic form on $H^1(X,\mathbb{C})$ more transparent. We now turn to some *topological* remarks that will be useful in describing the Jacobian more explicitly, and also in the second proof of Abel's theorem.

First, recall that a surface of genus $g$ admits a basis for $H_1(X,\mathbb{Z})$ of the form $(a_i, b_i)$, $i = 1, \ldots, g$, such that $a_i \cdot b_i = 1$ and all other products vanish. In other words, the intersection pairing makes $H_1(X,\mathbb{Z})$ into a *unimodular* symplectic space.

This symplectic form on $H_1(X,\mathbb{Z})$ gives rise to one on the space of periods, $H^1(X,\mathbb{C})$, defined by:

$$[\alpha,\beta] = \sum \alpha(a_i)\beta(b_i) - \alpha(b_i)\beta(a_i).$$

In matrix form, we can associate a vector of periods $(\alpha_a, \alpha_b) \in \mathbb{C}^{2g}$ to any homomorphism $\alpha$, and then we have

$$[\alpha,\beta] = (\alpha_a,\alpha_b)\begin{pmatrix} 0 & I \\ -I & 0\end{pmatrix}\begin{pmatrix}\beta_a \\ \beta_b\end{pmatrix} = \left|\begin{pmatrix}\alpha_a & \alpha_b \\ \beta_a & \beta_b\end{pmatrix}\right| = \alpha_a\cdot\beta_b - \alpha_b\cdot\beta_a.$$

**Theorem 14.12** *Under the period isomorphism between $H^1_{DR}(X)$ and $H^1(X,\mathbb{C})$, we have*

$$\int_X \alpha\wedge\beta = [\alpha,\beta].$$

**Proof.** Cut $X$ along the $(a_i, b_i)$ curves to obtain a surface $F$ with boundary, on which we can write $\alpha = df$. Then we have

$$\int_X \alpha\wedge\beta = \int_F (df)\wedge\beta = \int_{\partial F} f\,\beta.$$

If we write $\partial F = \sum a_i + b_i' - a_i' - b_i$, then

$$f|a_i' - f|a_i = \alpha(b_i) \quad \text{and} \quad f|b_i' - f|b_i = \alpha(a_i).$$

On the other hand, $\beta$ is the same along corresponding edges of $\partial F$. Therefore we find

$$\int \alpha \wedge \beta = \sum \alpha(a_i)\beta(b_i) - \alpha(b_i)\beta(a_i).$$

■

**Poincaré duality.** Note that for any cocycle $N \in H^1(X, \mathbb{Z})$, there is a dual cycle $C \in H_1(X, \mathbb{Z})$ such that

$$[N, \alpha] = \int_C \alpha$$

for all $\alpha \in H^1(X, \mathbb{C})$. In fact, since

$$[N, \alpha] = \sum N(a_i)\alpha(b_i) - N(b_i)\alpha(a_i),$$

we can simply take

$$C = \sum N(a_i) b_i - N(b_i) a_i.$$

Since the intersection form is unimodular, this construction gives a natural isomorphism

$$H^1(X, \mathbb{Z}) \cong H_1(X, \mathbb{Z}).$$

**Recognizing holomorphic 1-forms.** Our second remark uses the Riemann surface structure. Namely, we note that a class $\alpha \in H^1(X, \mathbb{C})$ can be represented by a holomorphic 1-form iff

$$[\alpha, \omega] = 0 \ \ \forall \omega \in \Omega(X).$$

This is because $H^1(X, \mathbb{C}) \cong \Omega(X) \oplus \overline{\Omega}(X)$.

**Meaning of the norm.** What is the meaning of $\|\omega\|^2$? It is simply the *area of* $X$ under the map $f : X \to \mathbb{C}$ defined *locally* by solving $df = \omega$. This map can be constructed globally on $X$ once it is slit along the $a_i$ and $b_i$ cycles.

For example, the area of a rectangle $R$ in $\mathbb{C}$ with adjacent sides $a$ and $b$, positively oriented, is

$$\operatorname{area}(R) = \operatorname{Im}(\overline{a} b) = \frac{i}{2}\left(a\overline{b} - \overline{a} b\right).$$

And if $X$ is a complex torus with a 1-form $\omega$, with periods $(a,b) = \omega(a_1), \omega(b_1)$, then

$$\int_X |\omega|^2 = (i/2)[\omega, \overline{\omega}] = (i/2)(a\overline{b} - b\overline{a}).$$

**Abel's theorem, proof II.** We are now prepared to give a second proof that a divisor with $\phi(D) = 0$ is principal (cf. Lang, *Algebraic Functions.*) To try to construct $f$ with $(f) = D$, we first construct a candidate for $\lambda = df/f$.

**Theorem 14.13** *For any divisor with* $\deg D = 0$*, there exists a meromorphic differential* $\lambda$ *with only simple poles such that* $\sum \operatorname{Res}_P(\lambda) \cdot P = D$*.*

**Proof.** By Mittag-Leffler for 1-forms, $\lambda$ exists because the sum of its residues is zero.

Alternatively, by Riemann-Roch, for any $P, Q \in X$ we have $\dim H^0(K + P + Q) > \dim H^0(K)$. Thus there exists a meromorphic 1-form $\lambda$ with a simple pole at one of $P$ or $Q$. By the residue theorem, $\lambda$ has poles at both points with opposite residues. Scaling $\lambda$ proves the Theorem for $Q - P$, and a general divisor of degree zero is a sum of divisors of this form. ■

Now if we could arrange that the periods of $\lambda$ are all in the group $2\pi i\mathbb{Z}$, then $f(z) = \exp \int \lambda$ would be a meromorphic function with $(f) = D$.

(Compare Mittag-Leffler's proof of Weierstrass's theorem on functions with prescribed zeros.)

**Periods of $\boldsymbol{\lambda}$.** As we have just seen, $D$ determines a meromorphic form $\lambda$ with simple poles and

$$\operatorname{Res}(\lambda) = \sum_P \operatorname{Res}_P(\lambda) \cdot P = D.$$

However $\lambda$ is not uniquely determined by this condition — it can be modified by a form in $\Omega(X)$. Furthermore, the periods of $\lambda$ are not well defined — when we move a path across a pole of $\lambda$, it changes by an integral multiple of $2\pi i$.

What we *do* obtain, canonically from $D$, is a class

$$[\lambda] \in H^1(X, \mathbb{C})/(2\pi i H^1(X, \mathbb{Z}) + \Omega(X)),$$

where $\Omega(X)$ is embedded by its periods. And this class is zero iff there is a meromorphic function such that $(f) = D$.

**Reciprocity.** To make everything better defined, we now choose a symplectic basis $(a_i, b_i)$ as before, and cut $X$ open along this basis to obtain a region $F$. We also arrange that the poles of $\lambda$ lie inside $F$. Then the periods of $\lambda$ become well-defined, by integrating along these specific representatives $a_i, b_i$ of a basis for $H_1(X, \mathbb{Z})$.

Now any $\omega \in \Omega(X)$ can be written as $\omega = df$ on $F$. On the other hand, $\omega \wedge \lambda = 0$ in $F$, at least if we exclude small loops around the poles of $\lambda$. Integration around these loops gives residues. Thus we find

$$\begin{aligned} 0 &= \int_F \omega \wedge \lambda = \int_{\partial F} f\lambda - 2\pi i \sum \operatorname{Res}(fP_i) + \operatorname{Res}(fQ_i) \\ &= [\omega, \lambda] - 2\pi i \sum \int_{P_i}^{Q_i} \omega, \end{aligned}$$

where the path of integration is taken in $F$.

Now suppose $\phi(D) = 0$. Then the sum above vanishes modulo the periods of $\omega$. In other words, there exists a cycle $C$ such that

$$[\omega, \lambda] = 2\pi i \int_C \omega = 2\pi i [\omega, N]$$

for some $N \in H^1(X, \mathbb{C})$. But this means $[\omega, \lambda - 2\pi i N] = 0$ for all $\omega$, and hence $\lambda - 2\pi i N$ is represented by a holomorphic 1-form. Modifying $\lambda$ by this form, we obtain $[\lambda] = 2\pi i N$, and thus the periods of the $\lambda$ are integral multiples of $2\pi i$. Then $f = \exp(\int \lambda)$ satisfies $(f) = D$. ∎

**Natural subvarieties of the Jacobian.** Just as we can embed $X$ into $\operatorname{Jac}(X)$ (or more naturally into $\operatorname{Pic}_1(X)$), we can map $X^{(k)} = X^k/S_k$ to $\operatorname{Jac}(X)$ (or into $\operatorname{Pic}_k(X)$). By what we have just seen, the fibers of this map consists of linearly equivalent divisors, and the dimensions of the fibers are predicted by Riemann-Roch. In particular, we have:

**Theorem 14.14** *The fibers of the natural map $X^{(k)} \to \operatorname{Pic}_k(X)$ are projective spaces $|D|$ corresponding to complete linear systems of degree $k$.*

The images of these maps in $\operatorname{Jac}(X)$ (which are well-defined up to translation) are usually denoted $W_k$; in particular, we obtain an important divisors $W_{g-1}$ in this way, while $W_g = \operatorname{Jac}(X)$.

**The Jacobian and $X^{(g)}$.** As a particular case, we note that the surjective map

$$X^{(g)} \to \operatorname{Pic}_g(X)$$

has, as its fibers, projective spaces corresponding to special divisors. For example, when $g = 2$ the only special divisor class is $|K|$, and $X^{(2)}$ is therefore a complex torus with one point blown up to a line ($\mathbb{P}^1 \cong |K|$).

**Mordell's Conjecture.**

**Theorem 14.15** *Suppose $X$ has genus $g \geq 2$. Then, given any finitely generated subgroup $A \subset \mathrm{Jac}(X)$, the set $X \cap A$ is finite.*

This theorem is in fact equivalent to Mordell's conjecture (Falting's theorem), which states that $X(K)$ is finite for any number field $K$.

**The exponential sequence, the Picard group and the Jacobian.** An alternative description of the Jacobian is via the exponential sequence which leads to the exact sequence

$$H^1(X, \mathbb{Z}) \to H^1(X, \mathcal{O}) \to H^1(X, \mathcal{O}^*) \to H^2(X, \mathbb{Z}) \to 0.$$

Under the isomorphisms $H^1(X, \mathbb{Z}) \cong H_1(X, \mathbb{Z})$ by cup product, $H^1(X, \mathcal{O}) \cong \Omega(X)^*$ by Serre duality, and $H^2(X, \mathbb{Z}) \cong \mathbb{Z}$ by degree, we obtain the isomorphism

$$\mathrm{Pic}_0(X) = \mathrm{Ker}(H^1(X, \mathcal{O}) \to H^2(X, \mathbb{Z})) \cong \Omega(X)^*/H_1(X, \mathbb{Z}) = \mathrm{Jac}(X).$$

**The norm map.** We note that if $f : X \to Y$ is a nonconstant holomorphic map, we have a natural map on divisors and hence an induced map $\mathrm{Jac}(X) \to \mathrm{Jac}(Y)$ and even a pushforward map on line bundles, $\mathrm{Pic}(X) \to \mathrm{Pic}(Y)$. We also have a *norm map* $N : \mathcal{M}^*(X) \to \mathcal{M}(Y)^*$, given by taking the product over the fibers. Note that:

$$D - D' = (g) \implies f(D) - f(D') = (N(g)).$$

For this to be correct, the points of $f(D)$ must be taken with multiplicity according to the branching of $f$. In particular, $\deg f(P)$ is not generally a continuous function of $P \in X$.

Recalling that a divisor is represented by a cochain $g_i \in \mathcal{M}^*(U_i)$ with $(\delta g)_{ij} \in \mathcal{O}^*(U_{ij})$, the pushforward of divisors can be defined by the local norm map as well.

The same construction works for equidimensional maps between complex manifolds of higher dimension. In this case $f(D) = 0$ if the image of $f$ is a lower-dimensional variety.

**The Siegel upper half-space.** The Siegel upper half-space is the space $\mathfrak{H}_g$ of symmetric complex $g \times g$ matrices $P$ such that $\mathrm{Im}\, P$ is positive-definite.

The space $\mathfrak{H}_g$ is the natural space for describing the $g$-dimensional complex torus $\mathrm{Jac}(X)$, just as $\mathbb{H}$ is the natural space for describing the 1-dimensional torus $\mathbb{C}/\Lambda$.

To describe the Jacobian via $\mathfrak{H}_g$, we need to choose a symplectic basis $(a_i, b_i)$ for $H_1(X, \mathbb{Z})$.

**Theorem 14.16** *There exists a unique basis* $\omega_i$ *of* $\Omega(X)$ *such that* $\omega_i(a_j) = \delta_{ij}$*.*

**Proof.** To see this we just need to show the map from $\Omega(X)$ into the space of $a$-periods is injective (since both have dimension $g$. But suppose the $a$-periods of $\omega$ vanish. Then the same is true for the $(0,1)$-form $\overline{\omega}$, which implies

$$\int |\omega(z)|^2\, dz\, d\overline{z} = \int \omega \wedge \overline{\omega} = \sum \omega(a_i)\overline{\omega}(b_i) - \omega(b_i)\overline{\omega}(a_i) = 0,$$

and thus $\omega = 0$. ∎

Definition. The *period matrix* of $X$ with respect to the symplectic basis $(a_i, b_i)$ is given by

$$\tau_{ij} = \int_{b_j} \omega_i = \omega_i(b_j).$$

**Theorem 14.17** *The matrix* $\tau$ *is symmetric and* $\operatorname{Im}\tau$ *is positive-definite.*

**Proof.** To see symmetry we use the fact that for any $i, j$:

$$0 = \int \omega_i \wedge \omega_j = \sum \omega_i(a_k)\omega_j(b_k) - \omega_i(b_k)\omega_j(a_k) = \delta_{ij}\tau_{jk} - \tau_{ik}\delta_{jk} = \tau_{ji} - \tau_{ij}.$$

Similarly, we have

$$\frac{i}{2}\int \omega \wedge \overline{\omega} = \int |\omega(z)|^2 |dz|^2 \geq 0.$$

Thus, if we let

$$\begin{aligned} Q_{ij} &= \frac{i}{2}\int \omega_i \wedge \overline{\omega}_j = \frac{i}{2}\sum \omega_i(a_k)\overline{\omega}_j(b_k) - \omega_i(b_k)\overline{\omega}_j(a_k) \\ &= \frac{i}{2}(\overline{\tau}_{ji} - \tau_{ij}) = \operatorname{Im}\tau_{ij}, \end{aligned}$$

then

$$\sum Q_{ij} c_i c_j = \int \sum |c_i|^2 |\omega_i|^2,$$

and thus $\operatorname{Im}\tau$ is positive-definite. ∎

The *symplectic group* $\mathrm{Sp}_{2g}(\mathbb{Z})$ now takes over for $\mathrm{SL}_2(\mathbb{Z})$, and two complex tori are isomorphic as *principally polarized* abelian varieties if and only if there are in the same orbit under $\mathrm{Sp}_{2g}(\mathbb{Z})$.

With respect to the basis $(b_1, \ldots, b_g, a_1, \ldots, a_g)$ the symplectic group acts on the full matrix of $a$ and $b$ periods by

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} \tau \\ I \end{pmatrix} = \begin{pmatrix} A\tau + B \\ C\tau + D \end{pmatrix}.$$

We must then change the choice of basis for $\Omega(X)$ to make the $a$-periods, $C\tau + D$ into the identity matrix; and we find

$$g(\tau) = (A\tau + B)(C\tau + D)^{-1},$$

which shows $\mathrm{Sp}_{2g}(\mathbb{Z})$ acts on $\mathfrak{H}_g$ by non-commutative fractional linear transformations.

**A non-Jacobian.** We note that for $g \geq 2$ there are period matrices which do not arise from Jacobians. The simplest example comes from the rank 2 diagonal matrices

$$\tau_{ij} = \begin{pmatrix} t_1 & 0 \\ 0 & t_2 \end{pmatrix}$$

with $t_i \in \mathbb{H}$ (which, incidentally, show there is a copy of $\mathbb{H}^g$ in $\mathfrak{H}_g$). It is essential, here, that the matrix above is with respect to a symplectic base.

To see this matrix cannot arise, suppose in fact $\tau_{ij} = \mathrm{Jac}(X)$. Note that $\mathrm{Per}(\omega_1) = \mathbb{Z} \oplus \mathbb{Z}t_1 = \Lambda_1$ and $\int_{a_2} \omega_1 = \int_{b_2} \omega_1 = 0$. Thus by integrating $\omega_1$ we get a map $f : X \to E = \mathbb{C}/\Lambda_1$ such that $f^*(dz) = \omega_1$. But then

$$\int_X |f^*(dz)|^2 = \frac{i}{2} \int_X \omega_1 \wedge \overline{\omega}_1 = (i/2)(\overline{\omega}_1(b_1) - \omega_1(b_1)) = \mathrm{Im}(b_1) = \int_E |dz|^2,$$

and hence $\deg(f) = 1$, which is impossible for a holomorphic map between Riemann surfaces of different genera.

In fact $A = \mathbb{C}^2/\mathbb{Z}^2 \oplus \tau\mathbb{Z}^2$ can be realized as the Jacobian of a stable curve, $E_1 \vee E_2$. And *as a complex torus*, $A$ can be realized as a Jacobian — in fact, in genus two, a complex torus is a Jacobian unless it splits *symplecticly* as a product.

A dimension count shows that in genus $g \geq 4$, there are principally polarized Abelian varieties which cannot be realized by stable curves.

**The infinitesimal Torelli theorem.** A beautiful computation of Ahlfors gives the explicit derivative of the map $\mathcal{M}_g \to \mathcal{A}_g$ provided by $X \mapsto \mathrm{Jac}(X)$.

More precisely, the *coderivative* of this map sends the cotangent space at a given periodic matrix $(\tau_{ij})$ to the cotangent space to $\mathcal{M}_g$ at $X$, which is naturally the space $Q(X)$ of holomorphic quadratic differentials on $X$.

The map is given simply by

$$d\tau_{ij} = \omega_i \, \omega_j \in Q(X).$$

Now we have the following basic fact:

**Theorem 14.18 (Max Noether)** *The natural map $\Omega(X) \otimes \Omega(X) \to Q(X)$ is surjective iff $X$ is not hyperelliptic.*

**Corollary 14.19 (Local Torelli theorem)** *Away from the hyperelliptic locus, the map $\mathcal{M}_g \to \mathcal{A}_g$ is an immersion.*

Noether's theorem can be rephrased as a property of the canonical curve $X \subset \mathbb{P}^{g-1}$ — the complete linear series of quadrics in $\mathbb{P}^{g-1}$ restricts to a complete linear series on $X$. In fact, any canonical curve is projectively normal – see Griffiths and Harris.

**Ahlfors' formula: sketch of the proof.** A new complex structure on $X$ is specified by a (small) Beltrami differential $\mu = \mu(z) d\bar{z}/dz$. Then the complex cotangent space is generated at each point, not by $dz$, but by

$$dz^* = (1 + \mu) \, dz = dz + \mu(z) d\bar{z}.$$

In this way we obtain a new Riemann surface $X^*$. We let $\omega_i$ denote the basis for $\Omega(X)$ with $\omega_i(a_j) = \delta_{ij}$, and similarly for $\omega_i^*$. We have

$$\omega_i^* = \omega_i^*(z)(dz + \mu(z) d\bar{z}).$$

Let $k = \sup |\mu|$ which is assumed to be small. Since $\omega_i - \omega_i^*$ has vanishing $a$-periods, this form has self-intersection number zero in cohomology. Thus

$$0 = \frac{i}{2} \int_X (\omega_i - \omega_i^*) \wedge (\overline{\omega}_i - \overline{\omega}_i^*).$$

This yields

$$\int |\omega_i - \omega_i^*|^2 \, |dz|^2 = \int |\mu|^2 |\omega_i^*|^2 \, |dz|^2,$$

which implies

$$\omega_i^* = \omega_i + O(k).$$

A simple computation also gives

$$\tau_{ij}^* - \tau_{ij} = \int_X \omega_i \wedge \omega_i^* = \int_X \omega_i \omega_j \mu + O(k^2).$$

Thus the variation in $\tau_{ij}$ is given, in the limit, by pairing $\mu$ with $\omega_i \omega_j$. ∎

**Theta functions.** The theory of $\theta$-functions allows one to canonically attach a divisor

$$\Theta \subset A = \mathbb{C}^g/(\mathbb{Z}^g \oplus \tau\mathbb{Z}^g)$$

to the principally polarized Abelian variety $A$ determined by $\tau \in \mathfrak{H}_g$. Using the fact that $\operatorname{Im}\tau \gg 0$, we define the entire $\theta$-function $\theta : \mathbb{C}^g \to \mathbb{C}$ by

$$\theta(z) = \sum_{n\in\mathbb{Z}^g} \exp(2\pi i\langle n, z\rangle)\exp(\pi i\langle n, \tau n\rangle).$$

The zero-set of $\theta$ is $\Lambda$-invariant, and descends to a divisor $\Theta$ on $A$.

For example, when $g = 1$ and $\tau \in \mathbb{H}$, we obtain:

$$\theta(z) = \sum_{\mathbb{Z}} \exp(2\pi inz)\exp(\pi in^2\tau).$$

Clearly $\theta(z + 1) = \theta(z)$, and we have

$$\theta(z + \tau) = \theta(z)\exp(-2\pi iz - \pi i\tau).$$

A useful notation for $g = 1$ is to set $q = \exp(\pi i\tau)$ and $\zeta = \exp(2\pi iz)$; then

$$\theta(\zeta) = \sum_{n\in\mathbb{Z}} q^{n^2}\zeta^n,$$

and we have

$$\theta(q^2\eta) = (q\zeta)^{-1}\theta(\zeta).$$

**Zeros and poles.** It turns out that $\theta$ has a unique zero on $X = \mathbb{C}/\mathbb{Z}\oplus\tau\mathbb{Z} = \mathbb{C}^*/q^{2\mathbb{Z}}$. Indeed, we have

$$\theta(-q) = \sum q^{n^2+n}(-1)^n = q^{-1/4}\sum q^{(n+1/2)^2}(-1)^n = 0,$$

since the terms $q^{m^2}$ and $q^{(-m)^2}$, $m \in \mathbb{Z} + 1/2$, occur with opposite signs.

A general meromorphic function on $X$ can be expressed as a suitable product of translates of $\theta$ functions. Indeed,

$$f_a(\zeta) = \prod_1^n \theta(-q\zeta/a_i)$$

has zeros at $a_i$, and satisfies

$$f_a(q^2\zeta) = (q^2\zeta)^{-n}\prod_1^n(-a_i).$$

Thus if $\prod_1^n a_i = \prod_1^n b_i$, the function

$$F(\zeta) = \frac{f_a(\zeta)}{f_b(\zeta)}$$

is meromorphic on $X$, with zeros at $(a_i)$ and poles at $(b_i)$. The condition that the products agree guarantees that the divisor of $(F)$ maps to zero under the Abel-Jacobi map.

**Theta functions for $g \geq 2$.** A convenient notation for $\theta$ functions in higher genus is the following: we set

$$\begin{aligned}
\zeta &= (\zeta_1, \ldots, \zeta_g) = (\exp 2\pi i z_i) \in (\mathbb{C}^*)^g, \\
q &= q_{ij} = \exp(\pi i \tau_{ij}) \in \exp(\pi i \mathfrak{H}_g), \\
n &= (n_1, \ldots, n_g) \in \mathbb{Z}^g, \\
\zeta^n &= \prod_i \zeta_i^{n_i}, \quad \text{and} \\
q^{n^2} &= \prod_{i,j} q_{ij}^{n_i n_j}.
\end{aligned}$$

Then we set

$$\theta(\zeta) = \sum_{n \in \mathbb{Z}^g} q^{n^2} \zeta^n,$$

and find for all $k \in \mathbb{Z}^g$:

$$\theta(q^{2k}\zeta) = \sum q^{n^2} q^{2nk} \zeta^n = q^{-k^2} \zeta^{-k} \sum q^{(n+k)^2} \zeta^{n+k} = q^{-k^2} \zeta^{-k} \theta(\zeta).$$

Noting that $\zeta(z_i + \sum m_i + \sum \tau_{ij} k_j) = q^{2k}\zeta$, this shows that $\Theta = (\theta)$ is a divisor on $A = \mathbb{C}^g / (\mathbb{Z}^g \oplus \tau \mathbb{Z}^g)$.

**Theorem 14.20 (Riemann)** *We have $\Theta = W_{g-1} + \kappa$ for some $\kappa \in \mathrm{Jac}(X)$, where $W_{g-1} \subset \mathrm{Jac}(X)$ is the image of $X^{g-1}$ under the Abel-Jacobi map.*

In particular, for $g = 2$ the divisor $\Theta$ is isomorphic to $X$ itself. Note however that $\Theta$ makes sense even when $A$ is not a Jacobian.

**The Torelli theorem.**

**Theorem 14.21** *Suppose $(\mathrm{Jac}(X), \omega)$ and $(\mathrm{Jac}(Y), \omega')$ are isomorphic as (principally) polarized complex tori (Abelian varieties). Then $X$ is isomorphic to $Y$.*

**Sketch of the proof.** Using the polarization, we can reconstruct the divisor $W_{g-1} \subset \text{Jac}(X)$ up to translation, using $\theta$-functions.

Now the tangent space at any point to $\text{Jac}(X)$ is canonically identified with $\Omega(X)^*$, and hence tangent hyperplanes in $\text{Jac}(X)$ give hyperplanes in $\mathbb{P}\Omega(X)$, the ambient space for the canonical curve. (For convenience we will assume $X$ is not hyperelliptic.) In particular, the Gauss map on the smooth points of $W_{g-1}$ defines a natural map

$$\gamma : W_{g-1} \to \mathbb{P}\Omega(X) \cong \mathbb{P}^{g-1}.$$

The composition of $\gamma$ with the natural map $C^{g-1} \to W_{g-1}$ simply sends $g-1$ points $P_i$ on the canonical curve to the hyperplane $H(P_i)$ they (generically) span.

Clearly this map is surjective. Since $W_{g-1}$ and $\mathbb{P}^{g-1}$ have the same dimension, $\gamma$ is a local diffeomorphism at most points. The *branch locus* corresponds to the hyperplanes that are *tangent* to the canonical curve $X \subset \mathbb{P}^{g-1}$. Thus from $(\text{Jac}(X), \omega)$ we can recover the collection of hyperplanes $X^*$ tangent to $X \subset \mathbb{P}^{g-1}$.

Finally one can show geometrically that $X^* = Y^*$ implies $X = Y$. The idea is that, to each point $P \in X$ we have a $(g-3)$ dimensional family of hyperplanes $H_P \subset X^*$ containing the tangent line $T_P(X)$. These hyperplanes must all be tangent to $Y$ as well; but the only reasonable way this can happen is if $T_P(X) = T_Q(Y)$ for some $Q$ on $Y$.

Now for genus $g > 3$ one can show a tangent line meets $X$ in exactly one point; thus $Q$ is unique and we can define an isomorphism $f : X \to Y$ by $f(P) = Q$. (For example, in genus 4 the canonical curve is a sextic in $\mathbb{P}^3$. If a tangent line $L \subset \mathbb{P}^3$ were to meet $2P$ and $2Q$, then the planes through $L$ would give a complementary linear series of degree 2, so $X$ would be hyperelliptic.) For $g = 3$ there are in general a finite number of 'bitangents' to the quartic curve $X \subset \mathbb{P}^2$, and away from these points we can define $f$, then extend. (For details see Griffiths and Harris). ■

**Explicit Torelli theorem in genus 2.** In the case of genus two, the Riemann surface $X$ can be reconstructed directly from its period matrix as the zero locus of the associated theta function. As remarked above, the Gauss map then gives the canonical map from $X$ to $\mathbb{P}^1$. To give $X$ as an *algebraic curve* it suffices to determine the six branch points $B$ of the canonical map. But these are nothing more than the images under the Gauss map of the six odd $\theta$ characteristics. In brief, we find:

**Theorem 14.22** *If $g(X) = 2$, then the 6 branch points of the canonical map $X \to \mathbb{P}^1$ correspond to the values of $d\theta$ at the 6 odd theta characteristics.*

**Spin structures in genus two.** A *spin structure* on $X$ is the choice of a square-root of the canonical bundle. Equivalently it is a divisor class $[D]$ such that $2D \sim K$. The *parity* of a spin structure is defined to be the parity of $h^0(D)$.

The spin structures on $X$ lie naturally in $\mathrm{Pic}_{g-1}(X)$ and form a homogeneous space for $\mathrm{Jac}(X)[2] \cong \mathbb{F}_2^{2g}$. The parity is, in fact, determined on these points by the polarization of the Jacobian.

In the case of genus 2, it is easy to see that there are 6 odd spin structures, one for each Weierstrass point $P_i \in X$. In fact the *differences* $P_i - P_j$ of Weierstrass points span $\mathrm{Jac}(X)[2]$.

To see this more clearly, first note that $2P_i \sim 2P_j \sim K$ for any $i$ and $j$. Thus if we take any $(\epsilon_i) \in \mathbb{F}_2^6$ with $\sum \epsilon_i = 0 \bmod 2$, and lift to integers such that $\sum \epsilon_i = 0$, then

$$\phi(\epsilon) = \sum_1^6 \epsilon_i P_i \in \mathrm{Jac}(X)[2]$$

is well-defined. We also note that $D = [5P1 - \sum_2^6 P_i]$ is a principal divisor; indeed, $D = (y)$ where $y^2 = f(x)$ and $\deg f = 5$. Thus $\phi(1,1,1,1,1,1) = 0$. Taking the quotient of the space of $(\epsilon_i)$ with $\sum \epsilon_i = 0 \bmod 2$ by subspace $\langle(1,1,1,1,1,1)\rangle$, we obtain $\mathrm{Jac}(X)[2] \cong \mathbb{F}_2^4$.

# 15 Hyperbolic geometry

*(The B-side.)*

**Elements of hyperbolic geometry in the plane.** The hyperbolic metric is given by $\rho = |dz|/y$ in $\mathbb{H}$ and $\rho = 2|dz|/(1-|z|^2)$ in $\Delta$.

Thus $d(i, iy) = \log y$ in $\mathbb{H}$, and $d(0,x) = \log(x+1)/(x-1)$ in $\Delta$. Note that $(x+1)/(x-1)$ maps $(-1,1)$ to $(0,\infty)$.

An important theorem for later use gives the hyperbolic distance from the origin to a hyperbolic geodesic $\gamma$ which is an arc of a circle of radius $r$:

$$\sinh d(0,\gamma) = \frac{1}{r}\cdot$$

To see this, let $x$ be the Euclidean distance from 0 to $\gamma$. Then we have, by algebra,

$$\sinh d(0,x) = \frac{2x}{1-x^2}\cdot \qquad (15.1)$$

On the other hand, we have by a right triangle with sides 1, $r$ and $x+r$. Thus $1 + r^2 = (x+r)^2$ which implies $2x/(1-x^2) = 1/r$.

A more intrinsic statement of this theorem is that for any point $p \in \mathbb{H}$ and geodesic $\gamma \in \mathbb{H}$, we have

$$\sinh d(p, \gamma) = \cot(\theta/2),$$

where $\theta$ is the visual angle subtended by $\gamma$ as seen from $p$.

**Area of triangles and polygons.** The area of an ideal triangle is $\pi$. The area of a triangle with interior angles $(0, 0, \alpha)$ is $\pi - \alpha$. From these facts one can see the area of a general triangle is given by the angle defect:

$$T(\alpha, \beta, \gamma) = \pi - \alpha - \beta - \gamma.$$

To see this, one extends the edges of $T$ to rays reaching the vertices of an ideal triangle $I$; then we have

$$T(\alpha, \beta, \gamma) = I - T(\pi - \alpha) - T(\pi - \beta) - T(\pi - \gamma)$$

which gives $\pi - \alpha - \beta - \gamma$ for the area.

Another formulation is that the area of a triangle is the sum of its exterior angles minus $2\pi$. In this form the formula generalizes to polygons.

**Right quadrilaterals with an ideal vertex.**

**Theorem 15.1** *Let $Q$ be a quadrilateral with edges of lengths $(a, b, \infty, \infty)$ and interior angles $(\pi/2, \pi/2, \pi/2, 0)$. Then we have*

$$\sinh(a)\sinh(b) = 1.$$

**Proof.** First we make a remark in Euclidean geometry: let $Q'$ be an ideal hyperbolic quadrilateral, centered at the origin, with sides coming from circles of Euclidean radii $(r, R, r, R)$. Then $rR = 1$.

Indeed, from this picture we can construct a right triangle with right-angle vertex 0, with hypotenuse of length $r + R$, and with altitude from the right-angle vertex of 1. By basic Euclidean geometry of similar triangles, we find $rR = 1^2 = 1$.

Now cut $Q'$ into 4 triangles of the type $Q$. Then we have $\sinh(a) = 1/r$ and $\sinh(b) = 1/R$, by (15.1). Therefore $\sinh(a)\sinh(b) = 1$. ■

**Right hexagons.** For a right-angled hexagon $H$, the excess angle is $6(\pi/2) - 2\pi = \pi$, and thus $\text{area}(H) = \pi$.

**Theorem 15.2** *For any $a, b, c > 0$ there exists a unique right hexagon with alternating sides of lengths $(a, b, c)$.*

**Proof.** Equivalently, we must show there exist disjoint geodesics $\alpha, \beta, \gamma$ in $\mathbb{H}$ with $a = d(\alpha, \beta)$, $b = d(\alpha, \gamma)$ and $c = d(\beta, \gamma)$. This can be proved by continuity.

Normalize so that $\alpha$ is the imaginary axis, and choose any geodesic $\beta$ at distance $a$ from $\alpha$. Then draw the 'parallel' line $L$ of constant distance $b$ from $\alpha$, on the same side as $\beta$. This line is just a Euclidean ray in the upper half-plane. For each point $p \in L_p$ there is a unique geodesic $\gamma_p$ tangent to $L_p$ at $p$, and consequently at distance $b$ from $\alpha$.

Now consider $f(p) = d(\gamma_p, \beta)$. Then as $p$ moves away from the juncture of $\alpha$ and $L$, $f(p)$ decreases from $\infty$ to 0, with strict monotonicity since $\gamma_p \cup L$ separates $\beta$ from $\gamma_q$. Thus there is a unique $p$ such that $f(p) = c$. ■

Doubling $H$ along alternating edges, we obtain a *pair of pants* $P$. Thus $\text{area}(P) = 2\pi$.

**Corollary 15.3** *Given any triple of lengths $a, b, c > 0$, there exists a pair of pants, unique up to isometry, with boundary components of lengths $(a, b, c)$.*

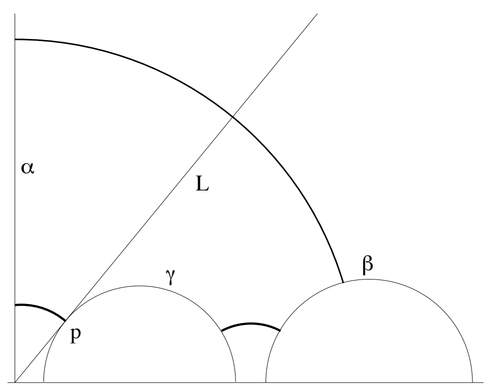



Figure 5. Right hexagons.

**Pairs of pants decomposition.**

**Theorem 15.4** *Any essential simple loop on a compact hyperbolic surface $X$ is freely homotopic to a unique simple geodesic. Any two disjoint simple loops are homotopic to disjoint simple geodesics.*

**Corollary 15.5** *Let $X$ be a compact surface of genus $g$. Then $X$ can be cut along $3g - 3$ simple geodesics into $2g - 2$ pairs of pants. In particular,* *we have*

$$\mathrm{area}(X) = 2\pi|\chi(X)|.$$

**Parallels of a geodesic.** There is a nice parameterization of the geodesic $|z| = 1$ in $\mathbb{H}$: namely

$$\delta(t) = \tanh t + i \,\mathrm{sech}\, t.$$

We have $\|\delta'(t)\| = 1$ in the hyperbolic metric.

Now given a closed simple geodesic $\gamma$ on $X$, let $C(\gamma, r)$ be a parallel curve at distance $r$ from $\gamma$. Then we have:

$$L(C(\gamma), r) = L(\gamma)\cosh(r).$$

Indeed, let $\gamma$ can be covered by the imaginary axis $i\mathbb{R}_+$ in $\mathbb{H}$. Then $C(\gamma, r)$ is a ray from 0 to $\infty$ which passes through $\delta(r)$. Thus the Euclidean slope of $C(\gamma, r)$ is the same as that of the vector $(x, y) = (\tanh t, \mathrm{sech}\, t)$. Thus projection along Euclidean horizontal lines from $C(\gamma, r)$ to $\gamma$ contracts by a factor of $y/\sqrt{x^2 + y^2} = \mathrm{sech}(t)$. Therefore $C(\gamma, r)$ is longer than $\gamma$ by a factor of $\cosh(t)$.

**The collar lemma.**

**Theorem 15.6** *Let $\alpha$ and $\beta$ be disjoint simple geodesics on a compact Riemann surface, of lengths a and b respectively. Define A and B by* $\sinh(a/2)\sinh(A) = 1$ *and* $\sinh(b/2)\sinh(B) = 1$. *Then the collars of widths $A$ and $B$ about $\alpha$ and $\beta$ are disjoint.*

**Proof.** We can assume that $\alpha$ and $\beta$ are part of a pants decomposition of $X$, which reduces the result to the case where $\alpha$ and $\beta$ are two cuffs of a pants $P$. By the Schwarz lemma, we can assume the lengths of the boundaries of $P$ are $(a, b, 0)$.

Now cut $P$ along a simple loop $\gamma$ that begins and ends at its ideal boundary component, i.e. the cuff of length zero. Then the components of $P = \gamma$ are doubles of quadrilaterals with one ideal vertex and the remaining angles $\pi/2$. The quadrilateral meeting $\alpha$ has finite sides of lengths $a/2$ and $A$ satisfying $\sinh(a/2)\sinh(A) = 1$, and similarly for the quadrilateral containing $\beta$. Thus these collars are disjoint. ∎

**Boundary of a collar.**

**Theorem 15.7** *The length of each component of the boundary of the standard collar around $\alpha$ with $L(\alpha) = a$ satisfies*

$$L(C(\alpha, r))^2 = a^2\cosh^2(r) = \frac{a^2}{1 + \sinh^{-2}(a/2)} \to 4$$

*as $a \to 0$. Thus the length of each component of the collar about $\alpha$ tends to 2 as $L(\alpha) \to 0$.*

**Proof.** Apply the preceding formulas.

Check. The limiting case is the triply-punctured sphere, the double of an ideal triangle $T \subset \mathbb{H}$ with vertices $(-1, 1, \infty)$. The collars limit to the horocycles given by the circles of radius 1 resting on $\pm 1$ together with the horizontal line segment $H$ at height 2 running from $-1 + 2i$ to $+1 + 2i$. We have $L(H) = 1$, so upon doubling we obtain a collar boundary of length 2.

**Corollary 15.8** *The collars about short geodesics on a compact hyperbolic surface cover the thin part of the surface.*

**Proof.** Suppose $x \in X$ lies in the thin part — that is, suppose there is a short essential loop $\delta$ through $x$. Then $\delta$ is homotopic to a closed geodesic $\gamma$, which is necessarily simple. But since $\delta$ is short, we see by the result above that $\delta$ must lie in the collar neighborhood of $\gamma$. ■

**Thick-thin decomposition.** There is a universal constant $r > 0$ such that any compact Riemann surface of genus $g$ can be covered by a collection of $O(g)$ balls $B(x_i, r)$ and $O(g)$ standard collars about short geodesics.

**Bers' constant.**

**Theorem 15.9** *There exists a constant $L_g$ such that $X$ admits a pants decomposition with no cuff longer than $L_g$.*

**Theorem 15.10** *We can take $L_g = O(g)$, but there exist examples requiring at least one curve of length $> C\sqrt{g} > 0$.*

**A finite-to-one map to $\mathcal{M}_g$.**

**Theorem 15.11** *For each trivalent graph of $G$ with $b_1(G) = g$, there is a finite-to-one map*

$$\phi_G : ((0, L_g] \times S^1)^{3g-3} \to \mathcal{M}_g,$$

*sending $(r_i, \theta_i)$ to the surface obtained by gluing together pants with cuffs of lengths $r_i$ and twisting by $\theta_i$, using pants and cuffs corresponding to vertices and edges of $G$.*

*The union of the images of the maps $\phi_G$ is all of $\mathcal{M}_g$.*

**Corollary 15.12 (Mumford)** *The function $L : \mathcal{M}_g \to \mathbb{R}$ sending $X$ to the length $L(X)$ of its shortest geodesic is proper.*

**The Laplacian.** Let $M$ be a Riemannian manifold. The Laplace operator $\Delta : C_0^\infty(M) \to C_0^\infty(M)$ is defined so that

$$\int_M |\nabla f|^2 = \int_M f \Delta f,$$

both integrals taken with respect to the volume element on $M$.

For example, on $\mathbb{R}$ we find $\Delta f = -d^2 f/dx^2$ by integrating by parts. Similarly on $\mathbb{R}^n$ we obtain

$$\Delta f = -\sum \frac{d^2 f}{dx_i^2}.$$

Note this is the negative of the 'traditional' Laplacian.

In terms of the Hodge star we can write

$$\int \langle \nabla f, \nabla f \rangle = \int df \wedge *df = -\int f \wedge d * df = \int f(- * d * df)\, dV,$$

and therefore we have

$$\Delta f = - * d * df.$$

For example, on a Riemann surface with a conformal metric $\rho(z)|dz|$, we have $*dx = dy$, $*dy = -dx$, and

$$\Delta f = -\rho^{-2}(z) \left( \frac{d^2 f}{dx^2} + \frac{d^2 f}{dy^2} \right).$$

As a particular case, for $\rho = |dz|/y$ on $\mathbb{H}$ we see $\Delta y^\alpha = \alpha(1-\alpha) y^\alpha$, showing that $y^\alpha$ is an eigenfunction of the hyperbolic Laplacian.

**The heat kernel.** Let $X$ be a compact hyperbolic surface. Enumerating the eigenvalues and eigenfunctions of the Laplacian, we obtain smooth functions satisfying

$$\Delta \phi_n = \lambda_n \phi_n,$$

$\lambda_n \geq 0$. The *heat kernel* $K_t(x, y)$ is defined by

$$K_t(x, y) = \sum e^{-\lambda_n t} \phi_n(x) \overline{\phi}_n(y),$$

The heat kernel is the fundamental solution to the *heat equation.* That is, for any smooth function $f$ on $X$, the solution to the *heat equation*

$$\frac{df_t}{dt} = -\Delta f_t$$

with initial data $f_0 = f$ is given by $f_t = K_t * f$. Indeed, if $f(x) = \sum a_n \phi_n(x)$ then

$$f_t(x) = K_t * f = \sum a_n e^{-\lambda_n t} \phi_n(x)$$

clearly solves the heat equation and has $f_0(x) = f(x)$.

Note also that formally, convolution with $K_t$ is the same as the operator $\exp(-\Delta)$, which acts by $\exp(-\lambda_n)$ on the $\lambda_n$-eigenspace.

**Brownian motion.** The heat kernel can also be interpreted using diffusion; namely, $K_t(x, \cdot)$ defines a probability measure on $X$ that gives the distribution of a Brownian particle $x_t$ satisfying $x_0 = x$.

For example, on the real line, the heat kernel is given by

$$K_t(x) = \frac{1}{\sqrt{4\pi t}} \exp(-x^2/(4t)).$$

Also we have $K_{s+t} = K_s * K_t$, as befits a Markov process.

To check this, note that $K_t$ solves the heat equation, and that $\int K_t = 1$ for all $t$. Thus $K_t * f \to f$ as $t \to 0$, since $K_t$ concentrates at the origin.

In terms of Brownian motion, the solution to the heat equation is given by $f_t(x) = E(f_0(x_t))$, where $x_t$ is a random path with $x_0 = x$.

**The trace.** The *trace* of the heat kernel is the function

$$\operatorname{Tr} K_t = \int_X K_t(x, x) = \sum e^{-\lambda_n t}.$$

It is easy to see that the function $\operatorname{Tr} K_t$ determines the set of eigenvalues $\lambda_n$ and their multiplicities.

**Length spectrum and eigenvalue spectrum.**

**Theorem 15.13** *The length spectrum and genus of $X$ determine the eigenvalues of the Laplacian on $X$.*

**Proof.** The proof is based on the trace of the heat kernel. Let $k_t(x, y)$ denote the heat kernel on the hyperbolic plane $\mathbb{H}$; it satisfies $k_t(x, y) = k_t(r)$ where $r = d(x, y)$. Then for $X = \mathbb{H}/\Gamma$ we have

$$K_t(x, y) = \sum_\Gamma k_t(x, \gamma y),$$

where we regard $K_t$ as an equivariant kernel on $\mathbb{H}$.

Working more intrinsically on $X$, we can consider the set of pairs $(x, \delta)$ where $\delta$ is a loop in $\pi_1(X, x)$. Let $\ell_x(\delta)$ denote the length of the geodesic representative of $\delta$ based at $x$. Then we have:

$$K_t(x, x) = \sum_\delta k_t(\ell_x(\delta)).$$

Let $\mathcal{L}(X)$ denote the space of nontrivial free homotopy classes of maps $\gamma : S^1 \to X$. For each $\gamma \in \mathcal{L}(X)$ we can build a covering space $p : X_\gamma \to X$ corresponding to $\langle \gamma \rangle \subset \pi_1(X)$.

The points of $X_\gamma$ correspond naturally to pairs $(x, \delta)$ on $X$ with $\delta$ freely homotopic to $\gamma$. Indeed, given $x'$ in $X_\gamma$, there is a unique homotopy class of loop $\delta'$ through $x'$ that is freely homotopic to $\gamma$, and we can set $(x, \delta) = (p(x), p(\delta'))$. Conversely, given $(x, \delta)$, from the free homotopy of $\delta$ to $\gamma$ we obtain a natural homotopy class of path joining $x$ to $\gamma$, which uniquely determines the lift $x'$ of $x$ to $X_\gamma$.

For $x' \in X_\gamma$, let $r(x')$ denote the length of the unique geodesic through $x'$ that is freely homotopic to $\gamma$. Then we have $\ell_x(\delta) = r(x')$. It follows that

$$\operatorname{Tr} K_t = \int_X K_t(x, x) = \int_X k_t(0) + \sum_{\mathcal{L}(X)} \int_{X_\gamma} r(x').$$

But $\int_X k_t(0) = \text{area}(X) k_t(0)$ depends only on the genus of $X$ by Gauss-Bonnet, and the remaining terms depend only on the geometry of $X_\gamma$. Since the geometry of $X_\gamma$ is determined by the length of $\gamma$, we see the length spectrum of $X$ determines the trace of the heat kernel, and hence the spectrum of the Laplacian on $X$. ∎

**Remark.** Almost nothing was used about the heat kernel in the proof. Indeed, the length spectrum of $X$ determine the trace of any kernel $K(x, y)$ on $X$ derived from a kernel $k(x, y)$ on $\mathbb{H}$ such that $k(x, y)$ depends only on $d(x, y)$.

**Remark.** In fact the genus is determined by the length spectrum.

**Isospectral Riemann surfaces.**

**Theorem 15.14** *There exist a pair of compact hyperbolic Riemann surfaces $X$ and $Y$, such that the length spectrum of $X$ and $Y$ agree (with multiplicities), but $X$ is not isomorphic to $Y$.*

**Isospectral subgroups.** Here is a related problem in group theory. Let $G$ be a finite group, and let $H_1, H_2$ be two subgroups of $G$. Suppose $|H_1 \cap C| =$

$|H_2 \cap C$ for every conjugacy class $C$ in $G$. Then are $H_1$ and $H_2$ conjugate in $G$?

The answer is *no* in general. A simple example can be given inside the group $G = S_6$. Consider the following two subgroups inside $A_6$, each isomorphic to $(\mathbb{Z}/2)^2$:

$$\begin{aligned} H_1 &= \langle e, (12)(34), (12)(56), (34)(56)\rangle, \\ H_2 &= \langle e, (12)(34), (13)(24), (14)(23)\rangle. \end{aligned}$$

Note that the second group actually sits inside $A_4$; it is related to the symmetries of a tetrahedron.

Now conjugacy classes in $S_n$ correspond to permutations of $n$, i.e. cycle structures of permutations. Clearly $|H_i \cap C| = 3$ for the cycle structure $(ab)(cd)$, and $|H_i \cap C| = 0$ for other conjugacy classes (except that of the identity). Thus $H_1$ and $H_2$ are isospectral. But they are not conjugate ('internally isomorphic'), because $H_1$ has no fixed-points while $H_2$ has two.

**Construction of isospectral manifolds.**

**Theorem 15.15 (Sunada)** *Let $X \to Z$ be a finite regular covering of compact Riemannian manifolds with deck group $G$. Let $Y_i = X/H_i$, where $H_1$ and $H_2$ are isospectral subgroups of $G$. Then $Y_1$ and $Y_2$ are also isospectral.*

**Proof.** For simplicity of notation we consider a single manifold $Y = X/G$ and assume the geodesics on $Z$ are discrete, as is case for a negatively curved manifold. Every closed geodesic on $Y$ lies over a closed geodesic on $Z$.

Fixing a closed geodesic $\alpha$ on $Y$, we will show the set of lengths $\mathcal{L}$ of geodesics on $Y$ lying over $\alpha$ depends only on the numbers $n_C = H \cap C$ for conjugacy classes $C$ in $G$.

For simplicity, assume $\alpha$ has length 1. Let $\alpha_1, \ldots, \alpha_n$ denote the components of the preimage of $\alpha$ on $X$. Let $S_i \subset G$ be the stabilizer of $\alpha_i$. The subgroups $S_i$ fill out a single conjugacy class in $G$, and we have $S_i \cong \mathbb{Z}/m$ where $nm = |G|$. Each loop $\alpha_i$ has length $m$.

Let $k$ be the index of $H \cap S_i$ in $S_i$. Then $k$ is the length of $\alpha_i/H$ in $Y$. Moreover, the number of components $\alpha_j$ in the orbit $H \cdot \alpha_i$ is $|H|/|H \cap S_i| = |H||S_i|/k$, and of course all these components descend to a single loop on $X/H$. Thus the number of times the integer $k$ occurs in $\mathcal{L}$ is exactly

$$|\mathcal{L}(k)| = \frac{kA_k}{m|H|},$$

where

$$A_k = |\{i \ : \ [S_i : S_i \cap H] = k\}|.$$

Thus to determine $\mathcal{L}$, it suffices to determine the integers $A_k$.

For example, let us compute $A_1$, the number of $i$ such that we have $S_i \subset H$. Now $H$ contains $S_i$ if and only if $H$ contains a generator $g_i$ of $S_i$. We can choose the $g_i$'s to fill out a single conjugacy class $C$, since the groups $S_i$ are all conjugate. Then the proportion of $i$'s satisfying $S_i \subset H$ is exactly $|H \cap C|/|C|$, and therefore

$$A_1 = \frac{n|H \cap C|}{|C|}.$$

An important point here: it can certainly happen that $S_i = S_j$ even when $i \neq j$. For example if $G$ is abelian, then all the groups $S_i$ are the same. But the number of $i$ such that $S_i$ is generated by a given element $g \in C$ is a constant, independent of $g$. Thus the *proportion* of $S_i$ generated by an element of $H$ is still $|H \cap C|/|C|$.

Now for $d|m$, let $C_d$ be the $d$th powers of the elements in $C$. Then the subgroups of index $d$ in the $S_i$'s are exactly the cyclic subgroups generated by elements $g \in C_d$. Again, the correspondence is not exact, but constant-to-one; the number of $i$ such that $\langle g \rangle \subset S_i$ is independent of $g \in C_d$. Thus the proportion of $S_i$'s such that $H \cap S_i$ *contains* a subgroup of index $d$ is exactly $|H \cap C_d|/|C_d|$, which implies:

$$\sum_{k|d} A_k = \frac{n|H \cap C_d|}{|C_d|}.$$

From these equations it is easy to compute $A_k$. ■

**Cayley graphs.** The spaces in Sunada's construction do not have to be Riemannian manifolds. For example, we can take $Z$ to be a bouquet of circles. Then $X$ is the Cayley group of $G$, and $Y_1$ and $Y_2$ are coset graphs on which $G$ acts. The coset graphs $Y_1$ and $Y_2$ also have the same length spectrum!

**A small example.** Let $G = (\mathbb{Z}/8)^* \ltimes \mathbb{Z}/8$ be the affine group of $A = \mathbb{Z}/8$, i.e. the group of invertible maps $f : A \to A$ of the form $f(x) = ax + b$. Let

$$\begin{aligned} H_1 &= \{x, 3x, 5x, 7x\}, \\ H_2 &= \{x, 3x + 4, 5x + 4, 7x\}. \end{aligned}$$

Then the subgroups $H_1$ and $H_2$ are isospectral.

In both cases, the coset space $Y_i = G/H_i$ can be identified with $\mathbb{Z}/8$; that is, $\mathbb{Z}/8 \times H_i = G$.

To make associated graphs, $Y_1$ and $Y_2$, we take $\langle x+1, 3x, 5x\rangle$ as generators for $G$. Note that $3x$ and $5x$ have order 2. Then the coset graph $Y_1$ is an octagon, coming from the generator $x+1$, with additional (unoriented, colored) edges joining $x$ to $3x$ and $5x$. Similarly, $Y_2$ is also an octagon, but now the colored edges join $x$ to the antipodes of $3x$ and $5x$, namely $3x+4$ and $5x+4$.

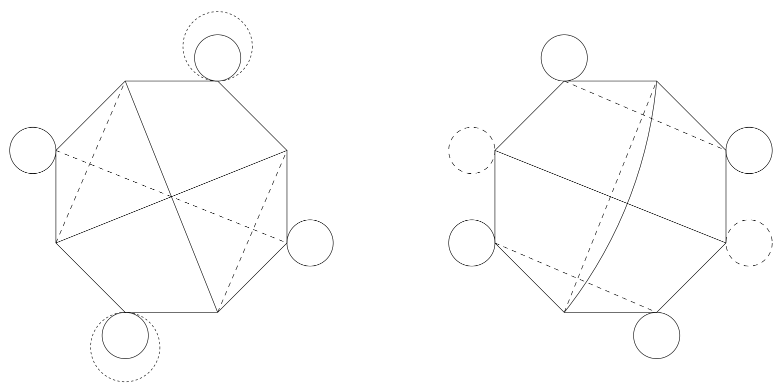


Figure 6. Isospectral graphs

These graphs are isospectral. In counting the number of loops, it is important to regard the graphs as covering spaces. For this it is best to replace each colored edge which is *not* a loop by a pair of parallel edges with opposite arrows. Each colored loop should be replaced by a single oriented edge. Then the graphs become covering spaces of the bouquet of 3 circles, and the number of loops of length $n$ is the same for both graphs.

**Not isometric.** Using short geodesics, we can arrange $Z$ such that one can reconstruct the action of $G$ on $G/H_i$ from the intrinsic geometry of $Y_i$. Then $Y_1$ and $Y_2$ are isometric iff $H_1$ and $H_2$ are conjugate. So in this way we obtain isospectral, but non-isometric, Riemann surfaces.

## 16 Uniformization

**Theorem 16.1** *Every simply connected Riemann surface is isomorphic to* $\widehat{\mathbb{C}}$, $\mathbb{C}$ *or* $\mathbb{H}$.

**Poincaré series.** We will now show that the uniformization theorem implies the Riemann existence theorem; for example, it implies that every compact Riemann surface can be presented as a branched cover of $\widehat{\mathbb{C}}$.

For surfaces of genus 0 this is clear. For genus one, uniformization implies $X = \widehat{\mathbb{C}}/\Lambda$ and then the construction of the Weierstrass $\wp$-function completes the proof.

For surfaces of higher genus we have $X = \Delta/\Gamma$. Here the idea is the following: given a meromorphic *form* $\omega = \omega(z)\,dz^k$ on the unit disk, we generate a $\Gamma$-invariant form by summation (just as one would do for the $\wp$-function):

$$\Theta(\omega) = \sum_\Gamma \gamma^*\omega.$$

More explicitly, this form is given by

$$\Theta(\omega) = \sum_\Gamma \omega(\gamma(z))(cz+d)^{-2k}\,dz^k,$$

since $\gamma'(z) = (cz+d)^{-2}$ when $\gamma(z) = (az+b)/(cz+d)$ is normalized so that $ad - bc = 1$.

There is no chance that this sum will converge when $\omega$ is a nonzero *function.* However, when $k = 2$, $|\omega|$ is naturally an area form, and hence

$$\int_X |\Theta(\omega)| \le \int_\Delta |\omega|.$$

This shows the sum converges (and defines a meromorphic form on $X$) so long as $\int |\omega| < \infty$. In fact, it is enough that $\int |\omega|$ is finite outside a compact set $K \subset \Delta$.

A second problem is that $\Theta(\omega)$ might be zero. Indeed, if $\omega = \alpha - \gamma^*\alpha$, then the sum will telescope and give zero. As in the case of the $\wp$ function, we can insure the sum is nonzero by introducing a pole. That is, if $a \in \Delta$ projects to $p \in X$, then

$$\Omega_p = \Theta\left(\frac{dz^2}{z-a}\right)$$

gives a meromorphic quadratic differential with a simple pole at $p$ and elsewhere holomorphic on $X$. It follows that for $p \neq q$,

$$f_{pq} = \Omega_p/\Omega_q$$

defines a nonconstant meromorphic function on $X$; indeed, $f_{pq}(p) = \infty$ and $f_{pq}(q) = 0$.

The *multiplicities* of these zeros and poles is not quite clear; for example, $\Omega_q$ might vanish at $p$, and then $f_{pq}$ has at least a double order pole at $p$.

But if $p$ is close enough to $q$, then $\Omega_q(p) \neq 0$, and hence $f_{pq}$ provides a local chart near $p$.

With a little more work we have established:

**Corollary 16.2** *Every Riemann surface carries a nonconstant meromorphic function, indeed, enough such functions to separate points.*

**Quasiconformal geometry.** Here is another approach to the uniformization theorem. For any $\mu$ on $\widehat{\mathbb{C}}$ with $\|\mu\|_\infty < 1$, there exists a quasiconformal map $f : \widehat{\mathbb{C}} \to \widehat{\mathbb{C}}$ with complex dilatation $\mu$.

Evidentally we can uniformize at least one Riemann surface $X_g$ of genus $g$, e.g. using a regular hyperbolic $4g$-gon. Now take any other surface $Y$ of the same genus. By topology, there is a diffeomorphism $f : X_g \to Y$. Pulling back the complex structure to $X_g$ and lifting to the universal cover, we obtain by qc conjugacy a Fuchsian group uniformizing $Y$.

# References


[EKS] A. Edmonds, R. Kulkarni, and R. E. Stong. Realizability of branched coverings of surfaces. *Trans. Amer. Math. Soc.* **282**(1984), 773–790.

[Gol] G. M. Goluzin. *Geometric Theory of Functions of a Complex Variable.* Amer. Math. Soc, 1969.

[GH] P. Griffiths and J. Harris. *Principles of Algebraic Geometry.* Wiley Interscience, 1978.

[Hel] S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces.* Academic Press, 1978.

[KZ] M. Kontsevich and D. Zagier. Periods. In *Mathematics Unlimited — 2001 and Beyond*, pages 771–808. Springer-Verlag, 2001.

[La] S. Lang. *Algebra.* Addison-Wesley, 1984.