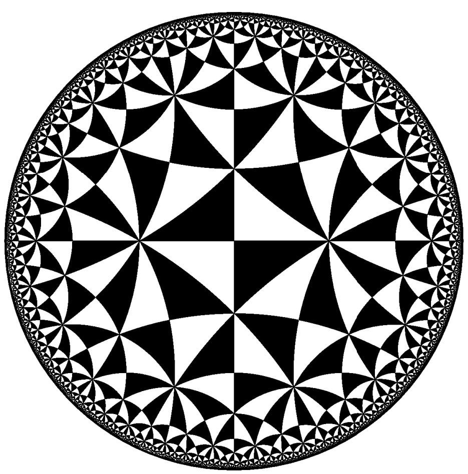
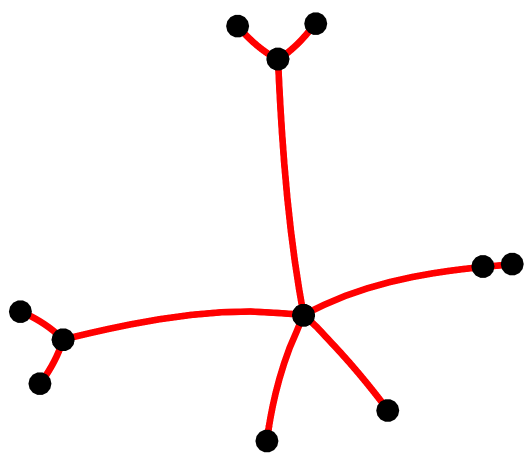
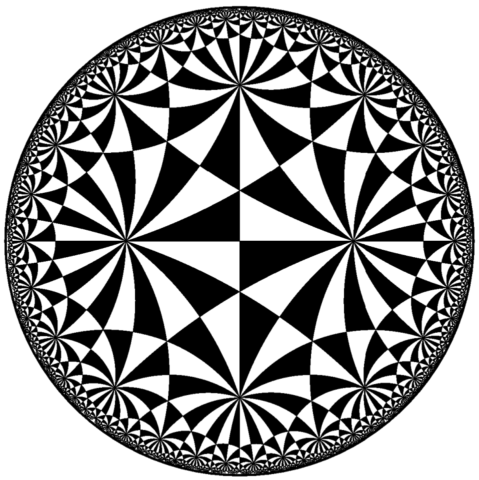
MAT 670, Fall 2023, Stony Brook University

TOPICS IN COMPLEX ANALYSIS
DESSINS AND DYNAMICS
Part I: Uniformization and Belyi's theorem

Christopher Bishop, Stony Brook







This is a course about constructing polynomials and entire functions using quasiconformal folding; a method that give good geometric control of the function and precise placement of the critical values singular values.

There are connections to the theory of *dessins d'enfant*, and we will briefly describe Belyi's theorem, although mostly we will deal with more analytic aspects of dessins,

The main tool will be quasiconformal mappings, and we will review the basics definitions and facts, including the measurable Riemann mapping theorem, which is central to our approach.

Defn: A topological surface X is a Hausdorff topological space provided with a collection homeomorphisms $\{\varphi_j\}$ (called charts) from open subsets of X to to open subsets of $\mathbb{C}$ such that:

- the union of the charts covers X,
- when two charts U_j, U_k intersect the the transition function $\varphi_j \circ \varphi_k^{-1}$ is a homeomorphism.

Defn: A Riemann surface is a connected topological surface such that the transit on functions are holomorphic.

- the plane, any planar domain (= open, connected).

- the 2-sphere

- the projective line = non-zero $\{(x, y) \in \mathbb{C}^2$ such that $(x, y) \sim (w, z)$ if $(x, y) = \lambda(w, z)$ for some $\lambda \neq 0$. This is same as the 2-sphere. The two charts are x/y if $y \neq 0$ and y/x if $n \neq 0$.

- Torus $= \mathbb{C}/\Gamma$ where $\Gamma = \{n\omega_1 + m\omega_2\}$ is a lattice

Different generators can give the same Riemann surface. Can always assume one generator is 1 and other is in upper half-plane. Different tori correspond to points in fundamental region of group $PSL(2, \mathbb{R})$.

Uniformization theorem: Every Riemann surface except sphere, plane, punctured plane and tori equals the unit disk (or upper half plane) modulo a Fuchsian group.

Fuchsian group = a discrete group of Möbius transformations mapping unit disk to itself (or upper half-plane to itself)..

Möbius transformation = linear fractional transformation

$$\frac{az+b}{cz+d}$$

.

Preserves disk if

$$e^{i\theta}\frac{z-a}{1-\overline{a}z}$$

Discrete = identity is isolated, i.e., no no-trivial sequence approaches it.

Relatively short proof of uniformization theorem given by Don Marshall.

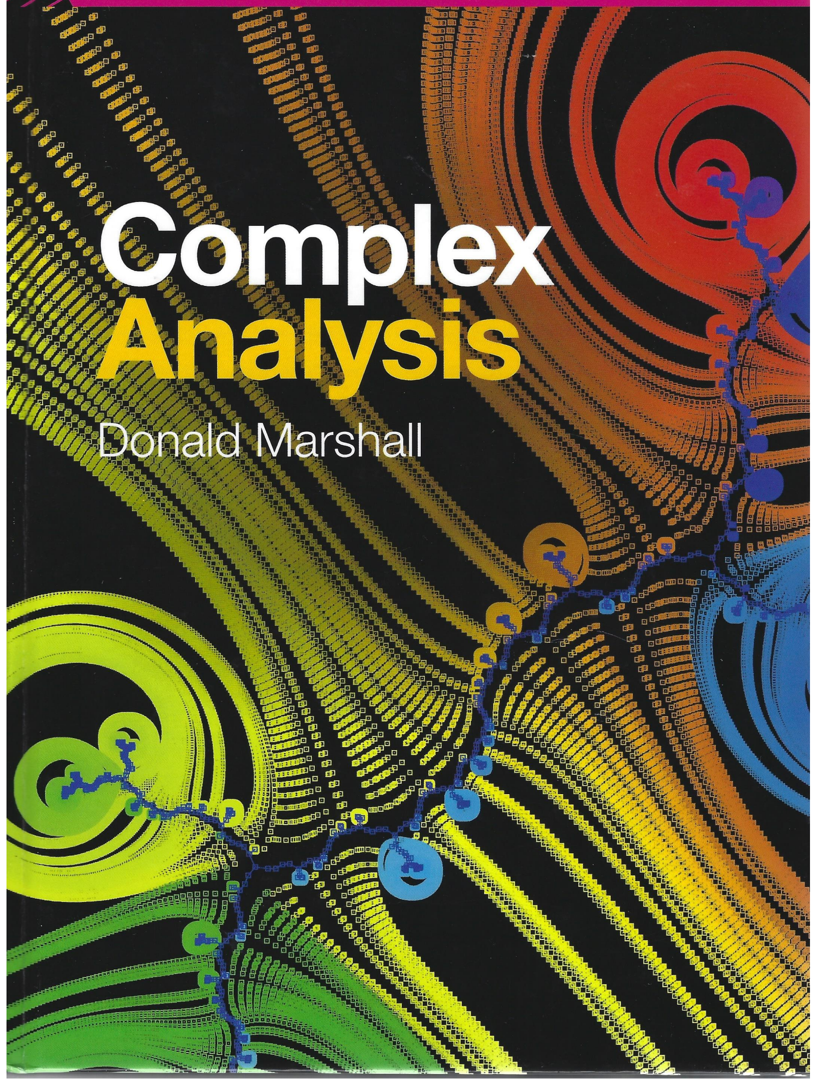
By “topology” s ait suffices to find all simply connected Riemann surfaces.

simply connected = every closed curve in is homotopic to a point.

CAMBRIDGE MATHEMATICAL TEXTBOOKS

Complex Analysis

Donald Marshall



Subharmonic functions::

A continuous function $u : \Omega \to [-\infty, \infty)$ is subharmonic if for all $z \in \Omega$ there is a $s > 0$ (depending on z) so that

$$v(z) \leq \frac{1}{2\pi} \int_0^{2\pi} v(z + re^{it} dt,$$

for all $r < s$.

Example: $\log |f|$ if f is holomorphic.

Sum and maximum of two subharmonic functions is subharmonic.

If a subharmonic funtion attains its maximum, then it is constant.

Lindelöf's Maximum principle:

Suppose Ω is a region and $\{\zeta_k\}_1^n \subset \partial\Omega$ but are not all of $\partial\Omega$. If u is subharmonic on Ω with $u \leq M < \infty$ and

$$\limsup_{z \to \zeta_k} u(z) \leq m$$

for each each $k = 1, \ldots n$, then $u \leq m$ on Ω.

Harnack's Theorem:

Suppose u is positive and harmonic on the unit disk. Then if $r = |z|$, we have

$$\left(\frac{1-r}{1+r}\right) u(0) \leq u(z) \leq q\left(\frac{1+r}{1-r}\right) u(0)$$

General version of Harnack's Theorem:

If u is positive and harmonic on a connected region Ω and $w, z \in \Omega$, there is a constant so that

$$u(w) \leq Cu(z) \leq C^2 u(w).$$

If v is subharmonic on Ω and $D \subset \Omega$ is a closed disk, then replacing v inside D by its Poisson extension from ∂D to D gives a subarhmonic function V_D with $v \leq V_D$.

Perron family: A family $\mathcal{F}$ of subharmonic functions on Ω is called a perron family if

(1) $v, u \in \mathcal{F}$ implies $\max(u, v) \in \mathcal{F}$.

(2) if $v \in \mathcal{F}$ and $D \subset \Omega$, then $v_D \in \mathcal{F}$.

(3) for each $z \in \Omega$, $v(z) > -\infty$ for some $v \in \mathcal{F}$.

Theorem: Given an Perron family $\mathcal{F}$,

$$u_{\mathcal{F}}(z) = \sup_{v \in \mathcal{F}} v(z)$$

is either harmonic on Ω or is identically $+\infty$.

If $\partial\Omega$ is compact in $\mathbb{C}$ and f is real-valued and continuous on $\partial\Omega$, the Perron family $\mathcal{F}_f$ for f consists of all subharmonic functions v on Ω so that

$$\limsup_{z\to w} v(z) \le f(w)$$

for all $w \in \partial\Omega$.

Then the sup over this family is a finite harmonic function that should be the solution of the Dirichet problem with boundary values f.

Barriers:

If $w \in \partial\Omega$, a **barrier** at w for Ω is a function b so that
(1) b is subharmonic on $\Omega \cap D(w, r)$ for some $r > 0$
(2) $b(z) < 0$ for $z \in \Omega \cap D(x, r)$
(3) $\lim_{z \to w} b(z) = 0$.

We say w is regular for Ω if a barrier exists.

If w belong to a non-trival connected component of $\partial\Omega$, then w is regular (for plane; not higher dimensions).

Theorem: The Perron solution is continous at every regular point of the boundary.

Green's function:

Suppose W is a Riemann surface. Fix $p_0 \in W$ and let $z : U \to D$ be a coordinate function such that $z(p_0) = 0$.

Let $\mathcal{F}_0$ be the collection of subharmonic functions on $W \setminus p_0$ so that

(a) $v = 0$ off some proper compact $K \subset W$,

(b) $\limsup_{p \to p_0}(v(p) + \log|z(p)|) < \infty$.

Set

$$g_W(p, p_0) = \sup\{v(p) : v \in \mathcal{F}_0\}.$$

One of the following two cases holds by Harnack's Theorem:

Case 1: $g_W(p, p_0)$is harmonic in $W \setminus p_0$, or

Case 2: $g_W(p, p_)$ $= +\infty$ for all $p \in W \setminus p_0$.

In the first case, g_W is called Green's function with pole at p_0.

In the second case we say that Green's function does not exist.

In probability theory the Green's function can be defined using Brownian motion.

For a measurable set E

$$T(E) = T(E) = \int_E g(z,p)\,dxdy$$

is the amount of time that a Brownian motion started at p spends in E.

If $G \equiv \infty$, then Brownian motion is recurrent and visits every positive area set E infinitely (and spends infinite time in E).

If $G < \infty$, then Brownian motion is transcient and leaves every compact set eventually.

The Uniformization Theorem (Koebe[1907]).

Suppose W is a simply connected Riemann surface.

(1) If Green's function exists for W, then there is a one-to-one analytic map of W onto the unit disk, $\mathbb{D}$.

(2) If W is compact, then there is a one-to-one analytic map of W onto the Riemann sphere $\mathbb{C}^*$.

(3) If W is not compact and if Green's function does not exist for W, then there is a one-to-one analytic map of W onto $\mathbb{C}$.

Lemma 1. *Suppose* $p_0 \in W$ *and suppose* $z : U \to \mathbb{D}$ *is a coordinate function such that* $z(p_0) = 0$. *If* g_W *exists then,*

(1) $g_W(p, p_0) > 0$ *for* $p \in W \setminus p_0$, *and*

(2) $g_W(p, p_0) + \log|z(p)|$ *extends to be harmonic in* U.

Proof: The function $v_0(p) = -\log|z(p)|$ on U and zero elsewhere is admissible (i.e., it is in $\mathcal{F}_0$). It is non-negative, so $g_W(p, p_0) \geq 0$ and $g_W(p, p_0) > 0$ if $p \in U$.

By the maximum principle $g_W > 0$ (and the fact that if a harmonic functions vanishes on a disk, it vanishes everywhere).

If v is admissible and $\epsilon > 0$, then $v + (1+\epsilon)\log|z|$ extends to be subharmonic in U , and equal to $-\infty$ at p_0. Thus

$$\sup_U (v + (1+\epsilon)\log|z|) = \sup_{\partial U} v \leq \sup_{\partial U} g_W < \infty.$$

Taking the supremum over admissible v and sending $\epsilon \to 0$, we obtain

$$g_W + \log|z| \le \sup_{\partial U} g_W = C < \infty$$

in $U \setminus p_0$. We also have that

$$g_W + \log|z| \ge v_0 + \log|z| \ge 0$$

for $p \in U \setminus p_0$. Thus p_0 is a removable singularity for the harmonic function $g_W + \log|z|$. $\blacksquare$

The Green's function for the unit disk D is given by

$$g_D(z,a) = \log |\frac{1-\overline{a}z}{z-a}|.$$

To prove this note that by the maximum principle, each candidate subharmonic function v is bounded above by $(1+\epsilon)g_D(z,a)$, for $\epsilon > 0$.

On the other hand $\max(g_D(z,a) - \epsilon, 0)$ is an admissible subharmonic function that vanishes off a compact set, and their supremum is g_D given above..

Lemma 2. *Suppose W_0 is a Riemann surface and suppose U_0 is a coordinate disk whose closure is compact in W_0. Set $W = W_0 \setminus \overline{U_0}$. Then $g_W(p, p_0)$ exists for all $p, p_0 \in W$ with $p \neq p_0$.*

Proof: Fix $p_0 \in W$ and let U be a coordinate disk as above with coordinate function x taking p_0 to 0.

We have to show the family of admissible subharmonic functions $\mathcal{F}_0$ is bounded above.

Fix $0 < r < 1$, and set $rU = \{p \in W : |z(p)| < r\}$.

If v is admissible and $\epsilon > 0$, then by the maximum principle

$$v(p) + (1+\epsilon)\log|z(p)| \leq \max_{\partial U}(v(q) + (1+\epsilon)\log|z(q)|) = \max_{\partial U} v(q),$$

for all $p \in U$.

Letting $\epsilon \to 0$, we obtain

$$\max_{\partial rU} v(p) + \log r \leq \max_{\partial U} v(p). \tag{1}$$

Let $\omega(p) = \omega(p, \partial rU, W \setminus rU)$ be the Perron solution to the Dirichlet problem on $W \setminus rU = W_0 \setminus \{U_0 \cup rU\}$ with boundary data 1 on ∂rU and 0 on ∂U_0.

More precisely, let $\mathcal{F}$ denote the collection of functions u which are subharmonic on $W \setminus rU$ with $u = 0$ on $W \setminus K$ for some compact set K, depending on u, and such that

$$\limsup_{p \to \zeta} u(p) \leq 1$$

for $\zeta \in \partial rU$. By definition $\omega(p) = \sup\{u(p) : u \in \mathcal{F}\}$.

By the Perron process, ω is harmonic in $W \setminus U$.

Regularity for the Dirichlet problem is a local question. We can define a local barrier at each point of the boundary of rU and at each point of the boundary of U_0.

Thus the harmonic function ω extends to be continuous at each point of ∂U_0 and each point of ∂rU so that $\omega(p) = 0$ on ∂U_0 and $\omega(p) = 1$ on ∂rU

This implies ω is not constant and $0 < \omega(p) < 1$ for $p \in W \setminus rU$.

By the maximum principle we have that

$$v(p) \leq \left(\max_{\partial rU} v \right) \cdot \omega(p)$$

for $p \in W \setminus rU$ since $v = 0$ off a compact subset of W. So

$$\max_{\partial U} v \leq \left(\max_{\partial rU} v \right) \max_{\partial U} \omega(p) \leq \left(\max_{\partial rU} v \right) (1 - \delta) \tag{2}$$

for some $\delta > 0$.

Note $1 - \delta = \max_{\partial U} \omega$. Since ω is continuous on the compact set ∂U it attains a maximum and this is strictly less than 1 since $0 < \omega < 1$.

Recall inequalities (1) and (2):

$$\max_{\partial rU} v(p) + \log r \leq \max_{\partial U} v(p)$$

$$\max_{\partial U} v \leq \left(\max_{\partial rU} v\right) \max_{\partial U} \omega(p) \leq \left(\max_{\partial rU} v\right)(1-\delta)$$

Combining these gives

$$\max_{\partial rU} v(p) \leq \log \frac{1}{r} + \left(\max_{\partial rU} v\right)(1-\delta)$$

or equivalently

$$\delta \max_{\partial rU} v \leq \log \frac{1}{r}$$

for every admissible v and δ independent of v.

This implies that Case 2 does not hold and hence Green's function exists.

$\blacksquare$

Lemma 3. *Suppose W is a Riemann surface with Green's function. Let W^* be a simply connected universal covering surface of W and let π be the universal covering map. Then g_{W^*} exists and satisfies*

$$g_W(\pi(p), \pi(p_0)) = \sum_{q:\pi(q)=\pi(p_0)} g_{W^*}g(p, q).$$

The terms of the series are non-negative, so the supremum over all finite sub-sums is either finite or $+\infty$.

Proof:. Suppose $q_1, \ldots q_n$ are distinct points in W^* that all project to p_0. Let p_0^* be one of these points projecting to p_0.

Suppose v_j is admissible for the Perron family used to construct G_{W^*} with pole at q_j. So v_j is zero off a compact set K_j of W^* and

$$\limsup_{p \to q_j} \left(v(p) + \log |z(\pi(p))| \right) < \infty,$$

where z is a coordinate chart on W with $z(p_0) = 0$.

Recall that $g_W(p, p_0) + \log|z(p)|$ extends to be finite and continuous at p_0, and hence

$$\lim_{p \to q} g_W(\pi(p), p_0) + \log|z(\pi(p))|$$

exists and is finite where $\pi(q_j) = p_0$.

Thus for $\epsilon > 0$,

$$\left(\sum_{j=1}^{n} v_j(p) \right) - (1+\epsilon) g_W(\pi(p), p_0)$$

extends to be subharmonic and equal to $-\infty$ at q_j for $j = 1, \ldots, n$, and less than or equal to 0 off $\cup_j K_j$.

By the maximum principle, this function is bounded above by 0. By letting $\epsilon \to 0$ and taking the supremum over all such v we conclude that $g_{W^*}(p, q_j)$ exists and

$$\sum_{j=1}^{n} g_{W^*}(p, q_j) \leq g_W(\pi(p), q).$$

Taking the supremum over all such finite sums (letting $n \nearrow \infty$) we have

$$S(p) \equiv \sum_{q:\pi(q)=p_0} g_{W^*}(p, q) \leq g_W(\pi(p), p_0).$$

This is half the desired equality. Next we prove the other direction.

Since $S(p) + \log|z(\pi(p))|$ is a supremum of finite sums of positive harmonic functions, it is harmonic in a neighborhood of each q_j by Harnack's Theorem.

Now take v in the Perron family used to construct $g_W(p, p_0)$.

Let U^* be a coordinate disk containing p_0^* such that $z \circ \pi$ is a coordinate function mapping U^* onto $\mathbb{D}$.

We claim that
$$v(\pi(p)) - (1 + \epsilon)S(p) \leq 0$$
for $p \in U^*$ and $\epsilon > 0$.

Since S is invariant under the group of deck transfomations, it is well defined on W. It has a logarithmic pole at p_0 so
$$v(\pi(p)) - (1 + \epsilon)S(p) \to -\infty$$
as $p \to p_0$.

On the other hand v is zero off some compact set K of W. Thus $v = 0$ on ∂K and $S > 0$ on ∂K since S is a sum of Green's functions that are positive everywhere. Thus

$$v(\pi(p)) - (1+\epsilon)S(p) \leq 0$$

for $p \in \partial H$ and $\epsilon > 0$.

By the maximum princple

$$v(\pi(p)) - (1+\epsilon)S(p) \leq 0$$

everwhere on $K \setminus \{p_0\}$ and hence on U and U^*, as claimed.

Taking $\epsilon \to 0$ gives

$$v(\pi(p)) \leq S(p)$$

$\in U^*$.

Finally, taking the supremum over all admissible v, we obtain

$$g_W(\pi(p), p_0) \leq S(p)$$

which proves the lemma. ■

We will use the following standard results without proof.

Riemann Mapping Theorem: *Any proper, simply connected planar domain can be conformally mapped to* $\mathbb{D}$*.*

Monodromy Theorem: *Suppose* Ω *is simply connected and suppose* f_0 *is defined and analytic in a neighborhood of* $b \in \Omega$*. If* f_0 *can be analytically continued along all curves in* Ω *beginning at* b*, then there is an analytic function* f *on* Ω *so that* $f = f_0$ *in a neighborhood of* b*.*

Theorem (Uniformization, Part I): If W is a simply connected Riemann surface then the following are equivalent:

(1) $g_W(p, p_0)$ exists for some $p_0 \in W$.

(2) $g_W(p, p_0)$ exists for all $p_0 \in W$.

(3) There is a one-to-one analytic map φ from W onto $\mathbb{D}$.

Moreover if g_W exists, then

$$g_W(p_1, p_0) = g_W(p_0, p_1),$$

and $g_W(p, p_0) = -\log|\varphi(p)|$, where $\varphi(p_0) = 0$.

Proof: Suppose there is a one-to-one, onto analytic map $\varphi : W \to D$ and let $p_0 \in W$. By composing this map with a Möbius transformation, we can assume that p_0 maps to 0.

If v is admissible for g_W and if $\epsilon > 0$, then by definition,

$$v + (1 + \epsilon) \log |\varphi|$$

is subharmonic in W and equal to $-\infty$ at p_0. By the maximum principle, since $v = 0$ off a compact set $K \subset W$,

$$v + (1 + \epsilon) \log |\varphi| \leq 0$$

on W .

Taking the supremum over all such v and letting $\epsilon \to 0$, shows that $g_W(p, p_0) < \infty$ and therefore the Green's function exists for all p_0. Trivially, this implies it exists for some p_0.

Now suppose we have existence for some pole p_0. (8) holds. By Lemma 1 the first lemma,

$$\mathrm{Re}(f(p)) = g_W(p, p_0) + \log|z(p)|,$$

is harmonic in U, so there is an analytic function f defined on a coordinate disk U containing p_0 so that

$$\mathrm{Re}(f(p)) = g_W(p, p_0) + \log|z(p)|,$$

for $p \in U$.

Hence the function

$$\varphi(p) = \exp(-f(p))$$

is analytic in U and satisfies

$$|\varphi(p)| = \exp(g_W(p, p_0))$$

$$\varphi(p_0) = 0.$$

On any coordinate disk U_α not containing p_0, $g_W(p, p_0)$ is the real part of an analytic function. Thus by the monodromy theorem, there is a function φ, analytic on W, such that

$$|\varphi(p)| = \exp(-g_W(p, p_0)) < 1.$$

We claim that φ is one-to-one.

If $\varphi(p) = \varphi(p_0) = 0$, then clearly $p = p_0$. Let $p_1 \in W$, with $p_1 \neq p_0$. Then since Green's function is positive at p_1, we have $|\varphi(p_1)| < 1$ and define

$$\varphi_1 \frac{\varphi - \varphi(p_1)}{1 - \overline{\varphi(p_1)}\varphi}$$

is analytic on W with absolute value bounded by 1. If $v \in \mathcal{F}_{p_1}$, then by the maximum principle

$$v + (1 + \epsilon) \log |\varphi_1| \leq 0.$$

Taking the supremum over all such v and sending $\epsilon \to 0$, we see that $g_W(p, p_1)$ exists and that

$$g_W(p, p_1) + \log |\varphi_1| \leq 0. \tag{3}$$

Setting $p = p_0$ above gives

$$g_W(p_0, p_1) \leq -\log|\varphi_1(p_0)| = -\log|\varphi(p_1)| = g_W(p_1, p_0).$$

Switching the roles of p_0 and p_1 gives the symmetry of Green's function

$$g_W(p_0, p_1) = g_W(p_1, p_0).$$

Moreover equality holds in (3) at $p = p_0$ so that

$$g_W(p, p_1) = -\log|\varphi_1(p)|$$

for all $p \in W \setminus \{p_1\}$. Now if $\varphi(p_2) = \varphi(p_1)$, then by the definition $\varphi_1(p_2) = 0$ and thus $gW(p_2, p_1) = \infty$ and so $p_2 = p_1$. Therefore φ is one-o-one, as claimed.

The image $\varphi(W) \subset D$ is simply connected, so if $\varphi(W) \neq D$ then by the Riemann Mapping Theorem we can find a one-to-one analytic map ψ of $\varphi(W)$ onto $\mathbb{D}$ with $\psi(0) = 0$. The map $\psi \circ \varphi$ is then a one-to-one analytic map of W onto $\mathbb{D}$, with $\psi \circ \varphi(p_0) = 0$, proving the theorem. ■

We have already seen the symmetry of Green's function when it exists on a simply connected Riemann surface. The proof above implies it is symmetric on any surface that has a Green's function.

Corollary: Suppose W is a Riemann surface for which Green's function exists, for some pole $q \in W$. Then $g_W(p,q)$ exists for all $p \neq q \in W$ and $g_W(p,q) = g_W(q,p)$.

Not every Riemann surface has a Green's function, but every such surface has a dipole Green's function that has two poles: one negative and one positive.

Lemma 4. *Suppose W is a Riemann surface and for $j = 1, 2$, suppose that $z_j : U_j \to \mathbb{D}$ are coordinate functions with disjoint coordinate disks, mapping p_1, p_2 to zero respectively.*

Then there is a function $G(p) \equiv G(p, p_1, p_2)$, that is harmonic in $W \setminus \{p_1, p|2\}$ such that

(1) $G + \log |z_1|$ extends to be harmonic in U_1, and

(2) $G - \log |z_2|$ extends to be harmonic in U_2.

Proof of the Uniformization Theorem, Case 2 (assuming Lemma 4, dipole Green's function exists:

By Part I, we may suppose that $g_W(p, p_1)$ does not exist for all $p, p_1 \in W$.

By the monodromy theorem and the lemma, there is a meromorphic function φ defined on W such that

$$|\varphi(p)| = \exp(-G(p, p_1, p_2)).$$

φ has a simple zero at p_1, a simple pole at p_2 and no other zeros or poles.

We claim φ is one-to-one. Take $p_0 \in W \setminus \{p_1, p_2\}$. Let φ_1 be the meromorphic function on W such that

$$|\varphi_1(p)| = \exp(-G(p, p_0, p_2)$$

and consider the function

$$H(p) = \frac{\varphi(p) - \varphi(p_0)}{\varphi_1(p)}.$$

Then H is analytic on W because its poles at p_2 cancel and because φ_1 has a simple zero at p_0.

By the lemma and the analyticity of H, $|H|$ is bounded on W.a But if v is in the Perron family for $g_W(p, p_1)$, and $\epsilon > 0$, then by the maximum principle

$$v(p) + (1+\epsilon)\log|\frac{H(p) - H(p_1)}{2\sup_W |H|}| \leq 0.$$

Since the Green's function of W does not exist, $\sup v(p) = \infty$, and therefore

$$\log|\frac{H(p) - H(p_1)}{2\sup_W |H|}| \equiv -\infty$$

or

$$H(p) \equiv H(p_1) = \frac{-\varphi(p_0)}{\varphi_1(p_1)} \neq 0, \infty.$$

From the definition of H, if $\varphi(p)$ is finite and not zero, then $\varphi(p) \neq \varphi(p_0)$since H is a non-zero constant.

If $\varphi_1(p) = 0$ then $p = p_0$ from the definition of φ_1.

Finally, φ_1 has a pole only at p_2. But φ also has a pole at p_2, and only at p_2, and thus $\varphi(p_2) \neq \varphi(p_0)$.

Thus $\varphi(p) = \varphi(p_0)$ only if $p = p_0$.

Since p_0 is arbitrary, this proves that φ is one-to-one.

Therefore φ is a one-to-one analytic map from W to a simply connected region $\varphi(W) \subset \mathbb{C}^*$. If the difference contains two or more points, then by the Riemann mapping Theorem, there is a one-to-one analytic map of this region, and hence of W, onto $\mathbb{D}$.

Since we assumed that g_W does not exist, this contradicts Part I of the uniformization theorem. Thus $\varphi(W)$ is either the sphere or the sphere minus one point.

If W is compact, so is the image, so it must be the sphere. If it is not compact, the image must be the plane (after moving the the omitted point to ∞ by a Möbius transformation.

This proves the uniformization theorem (except for Lemma 4). $\blacksquare$

Proof of Lemma 4:

Suppose z_0 is a coordinate function with coordinate chart U_0 that is disjoint from U_1 and U_2. Let p_0 be the point so that $z_0(p_0) = 0$. Set

$$tU_0 = \{p \in W : |z_0(p)| < t\}$$

and set

$$W_t = W \setminus tU_0.$$

By Lemma 2 and Theorem 4, the Green's function for W_t exists for all $p, p_1 \in W_t$ with $p \neq p_1$. Fix $0 < r < 1$, and set

$$rU_1 = \{p \in W : |z_1(p)| < r\}.$$

By the maximum principle

$$g_{W_t}(p, p_1) \leq M_1(t) \equiv \max_{q \in \partial U_1} g_{W_t}(q, p_1) \tag{4}$$

for all $p \in W_t \setminus rU_1$, because the same bound holds for all candidates in the Perron family defining g_{W_t}.

By (1)

$$M_1(t) \leq \max \partial U_1 g_{W_t}(p, p_1) + \log \frac{1}{r}. \tag{5}$$

By (4), $u_t(p) \equiv M_1(t) - g_{W_t}(p, p_1)$ is a positive harmonic function in $W_t \setminus rU_1$ and by (5) there exists $q \in \partial U_1$ with $u_t(q) \leq \log 1/r$.

Riemann surfaces are pathwise connected so that if K is a compact subset of $W_1 \setminus rU_1$ containing $\{p_2\} \cup \partial rU_1$, then by Harnack's inequality there is a constant C depending on K and r but not on t, so that for all $p \in K$ $0 \leq u_t(p) \leq C$, and

$$|g_{W_t}(p, p_1) - g_{W_t}(p_2, p_1)| = |u_t(p_2) - u_t(p)| \leq 2C.$$

Likewise, if K' is a compact subset of $W_1 \setminus \{|z_2| < r\}$ containing $\{p_1\} \cup \partial rU_2$, there is a constant C so that

$$|g_{W_t}(p, p_2) - g_{W_t}(p_1, p_2)| \leq C$$

for all $p \in K'$.

By the symmetry of Green's function, $g_{W_t}(p_1, p_2) = g_{W_t}(p_2, p_1)$ so the function

$$
\begin{aligned}
G_t(p, p_1, p_2) &\equiv g_{W_t}(p, p_1) - g_{W_t}(p, p_2) \\
&= [g_{W_t}(p, p_1) - g_{W_t}(p_2, p_1)] - [g_{W_t}(p, p_2) - g_{W_t}(p_1, p_2)]
\end{aligned}
$$

is harmonic in $W_t \setminus \{p_1, p_2\}$ and is bounded by C for all $p \in K \cap K'$, for some finite C independent of t.

We may suppose, for instance, that $K \cap K'$ contains $\partial U_1 \cup \partial U_2$. If v is in the Perron family for $g_{W_t}(p, p_1)$, then since $v = 0$ off a compact subset of W_t and since $g_{W_t} > 0$, by the maximum principle

$$\begin{aligned}
\sup_{W_t \backslash U_1} v(p) - g_{W_t}(p, p_2) &\leq \max(0, \sup_{\partial U_1} v(p) - g_{W_t}(p, p_2)) \\
&\leq \max(0, \sup_{\partial U_1} g_{W_t}(p, p_1) - g_{W_t}(p, p_2)) \\
&\leq C,
\end{aligned}$$

and taking the supremum over all such v yields

$$\sup_{W_t \backslash U_1} G(p, p_1, p_2) \leq C.$$

Similarly,

$$\sup_{W_t \setminus U_2} G(p, p_1, p_2) = - \sup_{W_t \setminus U_2} -G(p, p_1, p_2) \geq -C$$

and

$$|G_t(p, p_1, p_2)| \leq C$$

for all $p \in W_t \setminus \{U_1 \cup U_2\}$.

The function $G_t + \log|z_1|$ extends to be harmonic in U_1, so by the maximum principle, we have that

$$\sup_{U_1} |G_t + \log|z_1|| = \sup_{\partial U_1} |G_t + \log|z_1|| = \sup_{\partial U_1} |G_t| \leq C$$

and

$$\sup_{U_2} |G_t + \log|z_2|| = \sup_{\partial U_2} |G_t + \log|z_2|| = \sup_{\partial U_2} |G_t| \leq C$$

By normal families, there exists a sequence $t_n \to 0$ so that G_{t_n} converges uniformly on compact subsets of $W \setminus \{p_0, p_1, p_2\}$ to a function $G(p, p_1, p_2)$ satisfying the conclusions of the lemma. The function $G(p, p_1, p_2)$ extends to be harmonic at p_0 because it is bounded in a punctured neighborhood of p_0. ■

This completes the proof of the uniformization theorem.

Paul Koebe

In terms of Brownian motion on W,

$$\int_E g(x, y)dx$$

is the expected amount of time that a Brownian motion started at y will spend in the set E (over all time).

Green's function is finite iff Brownian motion is transient, i.e., it leaves every compact set eventually.

A Fuchsian group is called divergence type if the quotient Riemann surface has no Green's function.

A Riemann surface is sometimes called parabolic if there is no Green's function and hyperbolic if there is one.

Confusing since "hyperbolic" also used to mean universal cover is the disk. Must ask or deduce meaning from context.

Divergence/Convegence terminology is clearer.

Compact surfaces have no Green's functions.

Some infinite surfaces have no Green's function (cyclic cover of compact surface)

The following are all equivalent to divergence type (no Green's function):

(1) Brownian motion is recurrent.

(2) Geodesic flow on the unit tangent bundle of W is ergodic.

(3) Poincare series of covering group Γ diverges.

(4) Γ has the Mostow rigidity property (cojugating circle homeomorphisms are Möbius or singular).

(5) Γ has the Bowen's property.

(6) Almost every geodesic ray is recurrent. Equivalently, the set of escaping geodesic rays from a point $p \in W$ has zero (visual) measure.

Graduate Texts in Mathematics

Alan F. Beardon

The Geometry of Discrete Groups

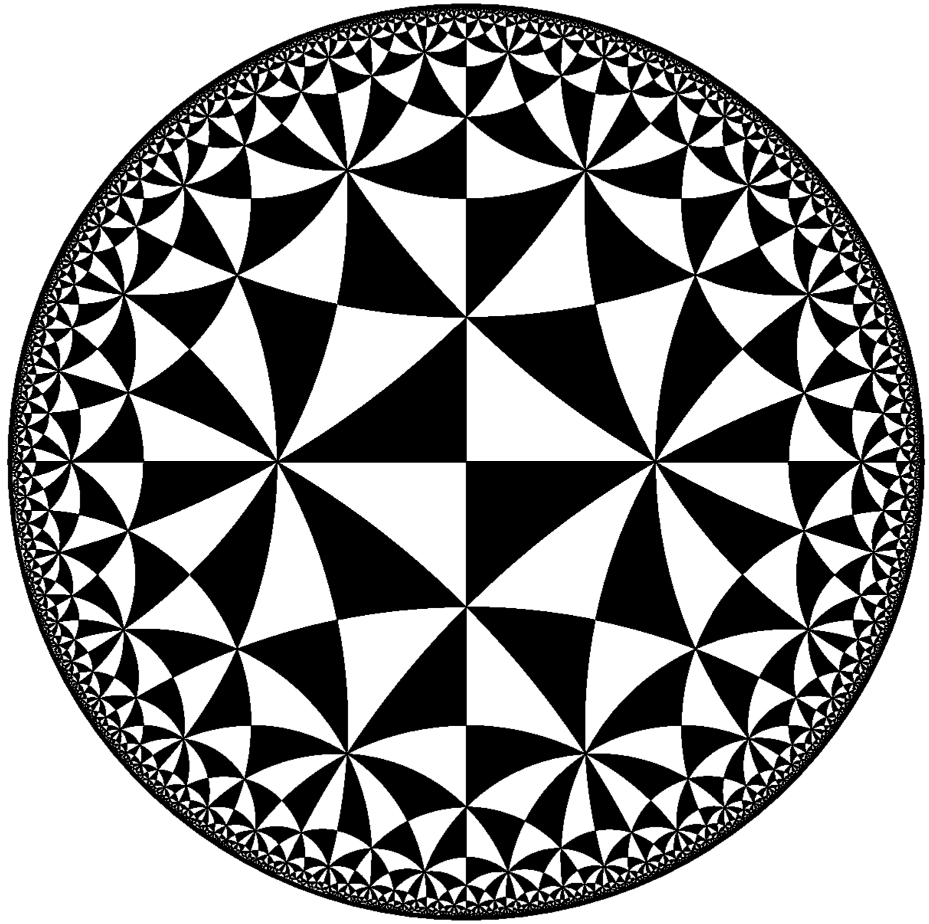
Springer-Verlag
New York Heidelberg Berlin

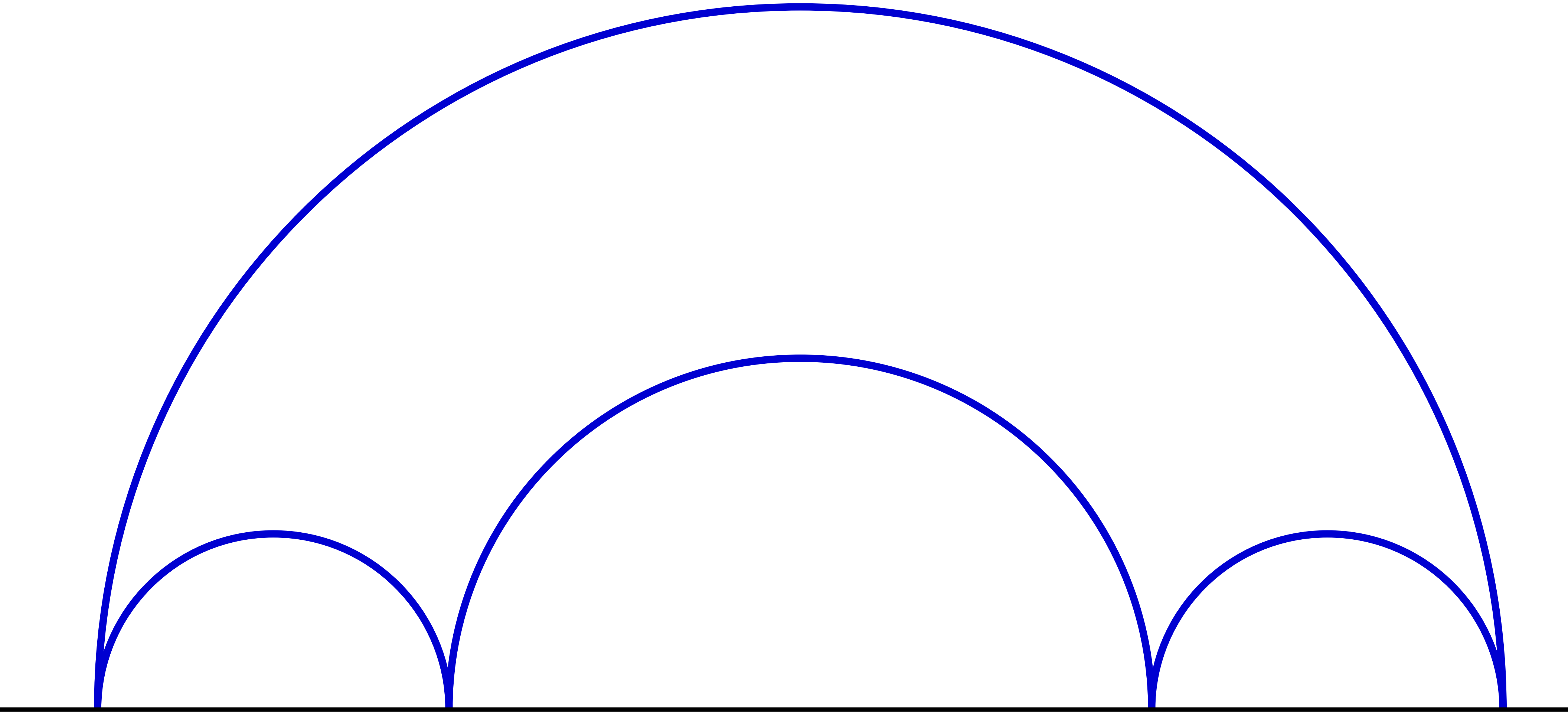
Suppose Γ is a Fuchsian group acting on the hyperbolic disk or half-space.

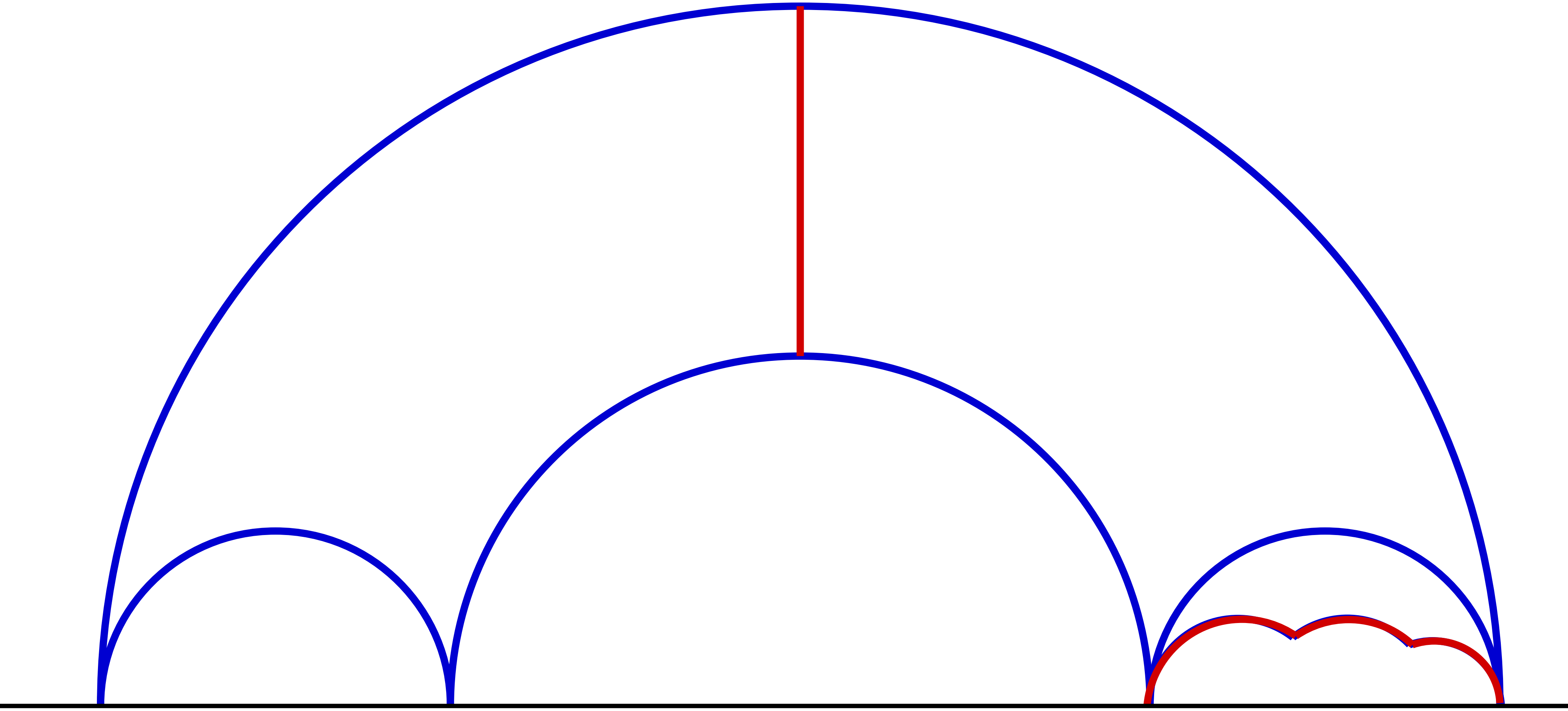
A **fundamental set** is a set that contains exactly one point from each orbit.

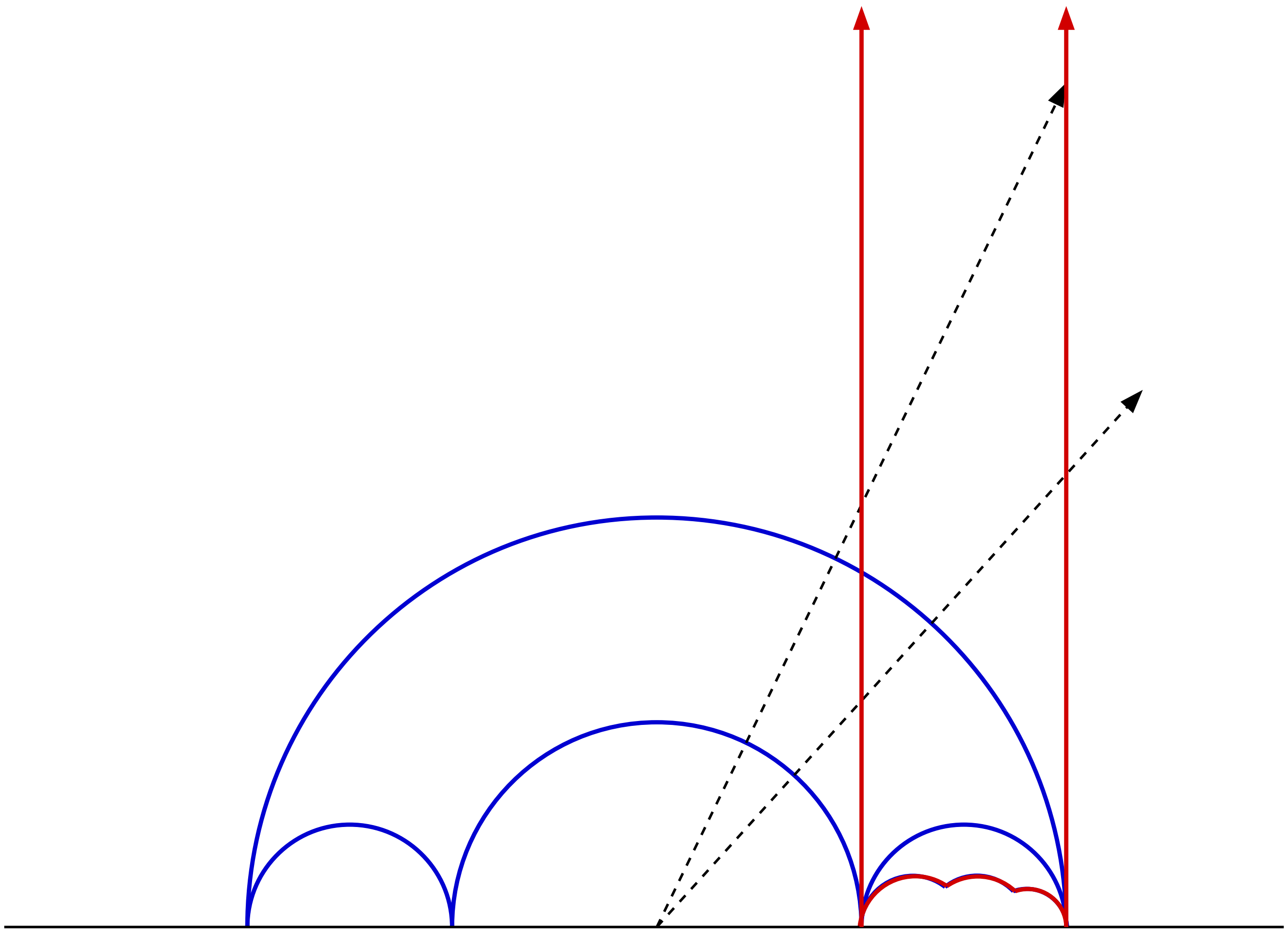
A **fundamental domain** is an open, connected set D so that there is a fundamental set F with $D \subset F \subset \overline{D}$ and so that ∂D has zero area.

A fundamental domain D is called **locally finite** if every compact set meets only finitely many images of $\overline{D}$.









Theorem (Thm 9.2.7, Beardon) If D is a locally finite fundamental region for a Fuchsian group Γ, then Γ is generated by the elements $g \in \Gamma$ so that $g(\overline{D}) \cap \overline{D} \neq \emptyset$.

If Γ is a Fuchsian group, then P is a convex fundamental polygon for Γ if it is a convex, locally finite fundamental domain.

A side of P is a geodesic segment of positive length of the form $g(\overline{D}) \cap \overline{D}$ for some $g \neq \mathrm{Id}$.

A vertex of P is a single point of the form $g(\overline{D}) \cap h(\overline{D}) \cap \overline{D}$ for some $g \neq h$, neither the identity.

Convex polygons exist:

For a point w the Dirichlet polygon $D(w)$ consists of a all points strictly closer to w than to any $g(w)$, $g \in \Gamma$, $g \neq \mathrm{Id}$.

This is an intersection of open half-planes, so is convex.

Poincaré's Theorem (Thm 9.8.4 Beardon):

Suppose P is a hyperbolic polygon and Φ is a set of side-pairing transformations.

Assume that for each vertex x of P there are vertices $x_0 = x, x_1, \ldots, x_n$ of P and elements $f_0 = \mathrm{ID}, \ldots f_n$ of Γ such that the sets $f_j(N_j)$ are non-overlapping sets whose union is $D(x, \epsilon)$ and such that each f_{j+1} is of the form $f_j g_s$ for some s.

Also assume the ϵ above can be chosen independent of x.

For a polygon P with a side-pairing Φ satisfying these conditions Γ is discrete and P is a fundamental polygon of Γ.

Theorem (Thm 10.1.2 Beardon): If Γ is a non-elementary Fuchsian group, then TFAE:

(1) Γ is finitely generated.
(2) Γ has a finite sided, convex fundamental polygon.
(3) Every convex fundamental polygon of Γ is finite sided.

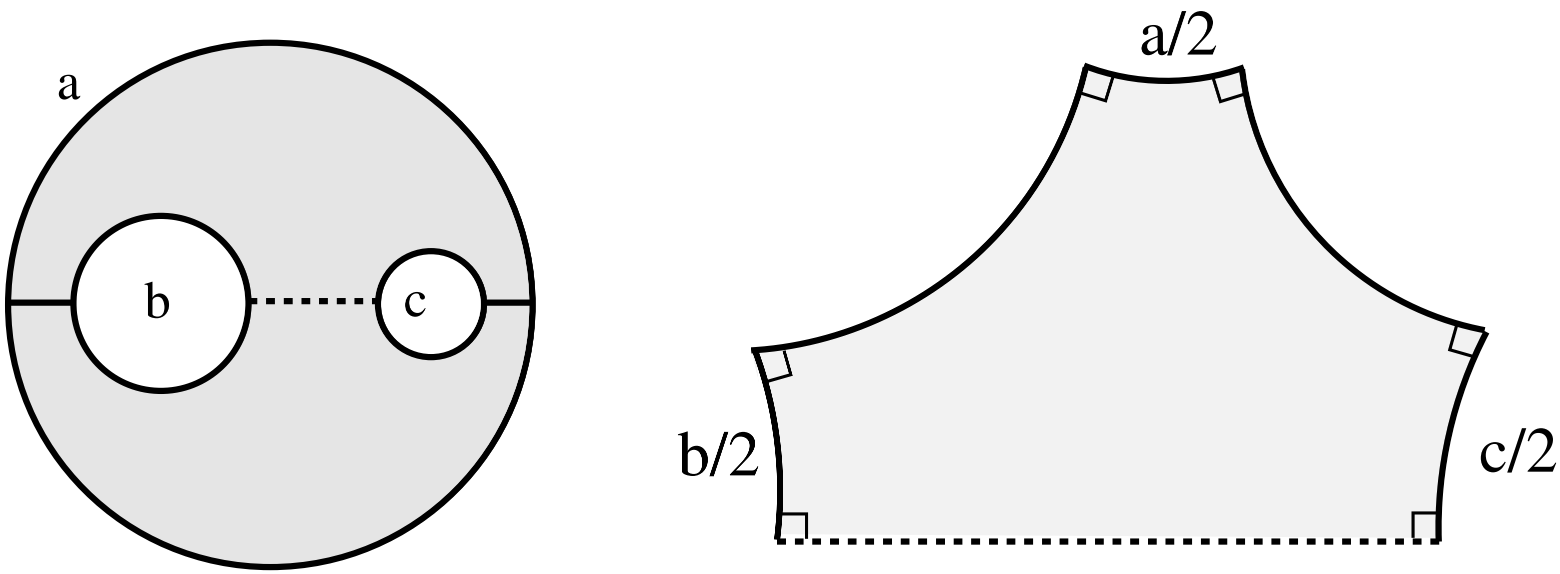
Not true for Kleinian groups acting on hyperbolic 3-space.

A Kleinian group is called geometrically finite if it has a finite sided fundamental polyhedron.

Y-pieces: A Y-piece is a bordered Riemann surface which is topologically a sphere with three disks removed and in which each of the three boundary components is a hyperbolic geodesic.

A generalized Y-piece is similarly defined, except that we also allow boundaries of length zero, i.e., instead of removing a disk we may remove a point.

A Y-piece can always be realized as a hyperbolic octagon with three side pairings. Moreover, there is a line of symmetry which divides the octagon into two isometric right hexagons. The alternate sides of these hexagons are given by $a_1 = a/2$, $a_2 = b/2$ and $a_3 = c/3$.



A right hexagon is determined by three alternating sides, so a Y piece is uniquely determined by its three side lengths.

Theorem: every compact Riemann surface is a finite union of Y-pieces.

Theorem: every finite area Riemann surface is a finite union of generalized Y-pieces (zero length boundaries allowed).

Theorem (Álvarez and Rogríguez, JLMS, 2004)

Every hyperbolic Riemann surface except for $\mathbb{D} \setminus \{0\}$ is the union (with pairwise disjoint interiors) of funnels, half-disks and a set G which can be exhausted by geodesic domains. Furthermore, if the surface is not $\mathbb{D}$ or an annulus, the set G appears always in the decomposition.

Half-disks are sometimes needed, e.g., $X = \mathbb{D} \setminus \{x_n\}$ where $0 < x_n \nearrow 1$.

Lemma 5. *If* Γ *is a Fuchsian group and* $z \in \mathbb{D}$ *then*

$$\sum_{\gamma \in \Gamma} |\gamma'(z)|^2 < \infty.$$

Lemma 6. *If* Γ *is a Fuchsian group and* $a \in \mathbb{D}$ *then*

$$Q_a(z) = \sum_{\gamma \in \Gamma} \frac{(\gamma'(z))^2}{\gamma(z) - a}$$

defines a meromorphic function on $\mathbb{D}$ *such that*

$$Q_a(\gamma(z)) = Q_a(z)/(\gamma'(z))^2$$

for all $\gamma \in \Gamma$*.*

Lemma 7. *If* Γ *is a Fuchsian group and* $a, b \in \mathbb{D}$ *then* $Q_a(z)/Q_b(z)$ *is a meromorphic,* Γ*-invariant function that has a simple pole at* a *and a simple zero at* b*.*

Corollary 8. *For any compact Riemann surface* X *and* $p \in X$ *there is a meromorphic function with a simple zero at* p*.*

Corollary 9. *For any compact Riemann surface X and $\{p_1, \ldots, p_n\} \subset X$ there is a meromorphic function taking any given n values $a_1, \ldots, a_n$ at these points. In particular, there is are functions taking n distinct values at these points.*

Equilateral triangulations and Belyi functions

Let T be a closed equilateral triangle. Starting from either a finite even number or a countably infinite number of copies of T, glue these triangles together by identifying every edge with exactly one edge of another triangle, in such a way that the identification map is the restriction of an orientation-reversing symmetry of T.

Assume furthermore that the resulting space E is connected, and that any vertex is identified with only finitely many other vertices. Then E is an orientable topological surface, which is compact if and only if the number of triangles we started with was finite.

We say that E is an *equilateral surface*.

Every equilateral surface comes equipped with a Riemann surface structure: On the interior of a face or of an edge, the complex structure is inherited from T.

It is easy to see that each vertex is conformally a puncture, and therefore the complex structure extends to all of E; indeed, local charts can be defined by using appropriate power maps.

We say that a Riemann surface is *equilaterally triangulable* if it is conformally equivalent to an equilateral surface;

There are only countably many ways to glue finitely many triangles together. So there are only countably many compact equilateral surfaces; therefore most compact Riemann surfaces can *not* be equilaterally triangulated.

Let $\mathcal{T}$ be a triangulation and let Δ be the Euclidean equilateral triangle inscribed in the unit circle, with a vertex at 1. For each topological triangle $T \in \mathcal{T}$, let ϕ_T denote a biholomoprhic isomorphism that takes T to Δ, mapping vertices to vertices. Observe that ϕ_T is unique up to post-composition by a rotational symmetry of Δ.

Defn: The triangulation $\mathcal{T}$ is *equilateral* if, on every edge e with two adjacent triangles T and $\tilde{T}$, the maps ϕ_T and $\phi_{\tilde{T}}$ agree up to a reflection symmetry of Δ.

If such a triangulation exists, we say that X is *equilaterally triangulable.*

It is elementary to see that this agrees with the definition given before, except that earlier we allowed allowed two triangles to intersect in more than one edge.

Given such an equilateral triangulation, we can perform a barycentric subdivision of all triangles, and get a triangulation of the same surface with no triangle glued to itself or to multiple edges of a distinct triangle.

Lemma 10 (Equilateral triangulations and reflections)**.** *A triangulation of X is equilateral if and only if the two triangles adjacent to a given edge are related by reflection.*

That is, suppose that the triangles T and $\tilde{T}$ are both adjacent to an edge e. Then there exists an anti-holomorphic homeomorphism $\iota \colon T \to \tilde{T}$ that fixes e pointwise and maps the third vertex of T to the corresponding vertex of $\tilde{T}$.

Proof:

Let e, T and $\tilde{T}$ be as in the statement, and let ϕ_T and $\phi_{\tilde{T}}$ be as defined above. Suppose that $\phi_{\tilde{T}}|_e = R \circ \phi_T|_e$, where R is a reflection symmetry of Δ. Then

$$\iota := \phi_{\tilde{T}}^{-1} \circ R \circ \phi_T$$

is an anti-holomorphic bijection as in the statement of the observation.

Conversely, suppose ι is such a bijection. Then $R := \phi_{\tilde{T}} \circ \iota \circ \phi_T^{-1}$ is an anti-holomorphic automorphism of the triangle Δ, mapping vertices to vertices. Thus R is a reflection symmetry of Δ, as required. ■

Definition 1. *Let X be a (compact or non-compact) Riemann surface. A meromorphic function $f\colon X \to \mathbb{C}^*$ is a* Belyi function *if f is a branched covering whose branched points lie only over -1, 1 and ∞.*

Proposition 11 (Triangulations and Belyi functions)**.** *A Riemann surface X is equilaterally trianguable if and only if there is a Belyi function on X.*

Proof. First suppose that $f\colon X \to \mathbb{C}^*$ is a Belyi function. Consider the generalized triangulation of the sphere into two triangles corresponding to the upper and lower half-plane, with vertices at 1, -1 and ∞. By the Schwarz reflection principle and our previous remarks, this triangulation is equilateral.

Since the critical values of f are at the vertices of the triangulation, we may lift it to X, to obtain a generalized equilateral triangulation. As discussed above, a barycentric subdivision leads to a triangulation in the stricter sense, and the proof of the "if" direction is complete.

Now suppose that an equilateral triangulation of the surface X is given. Let $\mathcal{T}$ be the corresponding collection of topological triangles, with conformal maps $\phi_T\colon T \to \Delta$ for $T \in \mathcal{T}$, as above. Let $\psi\colon \Delta \to \mathbb{D}$ be the conformal isomorphism that fixes 0 and 1, and consider the function

$$f\colon X \to \mathbb{C}^*; : z \mapsto F_3(\psi(\phi_T(z))) \qquad (z \in T),$$

where F_3 is the degree 6 rational map

$$F_3(z) := \frac{1}{2}(z^3 + z^{-3}).$$

Let ρ denote rotation by 60° around 0, and let σ denote complex conjugation.

Observe that ψ commutes with both operations, and that $F_3 \circ \rho = F_3 \circ \sigma = F_3$ on $\partial\mathbb{D}$. The group of symmetries of Δ is generated by ρ and σ, and thus f is indeed a well-defined holomorphic function on X.

Clearly f is a branched covering with no critical values outside of -1, 1 and ∞; so f is a Belyi function. $\square$

Defn: A smooth affine algebraic curve is
$$X = \{(x, y) \in \mathbb{C}^2 : f(x, y) = 0\}$$
where f is a polynomial such at each point $p \in X$ either
$$\frac{\partial f}{\partial x}(p) \neq 0 \quad \text{or} \quad \frac{\partial f}{\partial y}(p) \neq 0,$$

Implict function theorem covers X by charts where either x or y are the maps to complexes.

For example, hyperelliptic curves

$$y^2 = (x - a_1)\dots(x - a_n).$$

Fermat curves:

$$x^n + y^n = 1$$

Defn: Complex projective space is

$$\{(x, y, z) \in \mathbb{C}^3 : (x, y, z) \neq 0\}$$

with $(x, y, z) = (\lambda x, \lambda y, \lambda z)$ for $\lambda \neq 0$.

Is a compact, 4-dimensional (real) manifold.

Defn: A smooth projective algebraic curve

$$X = \{[x, y, z] \in \mathbb{P}^2(\mathbb{C}) : f(x, y, z) = 0\}$$

where f is a homogeneous polynomial and charts are x/z or y/z

Such curves are compact (closed subset of compact space).

Fermat curve: $x^n + y^n + z^n = 1.$

Hyperelliptic curve: $y^2 z^{n-2} = \prod_{j=1}^{n}(x - a_j z)$

Defn: The Euler characteristic of a compact surface is $\chi = V - E + F$ where V, E, F are the number of vertices, edges and faces of a triangulation.

Is independent of the triangulation.

For sphere, $\chi = 2$

For torus, $\chi = 0$

For genus g surface, $\chi = 2 - 2g$

Riemann-Hurwitz formula: If $f : X \to Y$ is non-constant holomorphic map between Riemann surfaces, then

$$2g(X) - 2 = (\deg f)(2g(Y) - 2) + \sum_{p \in X} (\text{mult}_p f - 1)$$

$\deg(f)$ is size of preimage of generic point (non-critical value).

For polynomial on sphere, this says (including ∞)

$$2(\deg(f) - 1) = \#(\text{criticalpoints}).$$

Suppose X is a Riemann surface.

a divisor $=$ is a finite linear combination of points of the surface with integer coefficients.

$\deg(D) =$ the sum of the coefficients occurring in D.

Every meromorphic function defines a divisor as set of zeros and poles. Coefficient is a a zero of order a and $-a$ at pole of order a.

Any such divisor is called a principle divisor.

Two divisors are linearly equivalent if difference is principle.

A divisor of a global meromorphic 1-form is called the canonical divisor, denoted K.

Any two meromorphic 1-forms will yield linearly equivalent divisors, so the canonical divisor is uniquely determined up to linear equivalence.

Riemann-Roch Thm *For any divisor D on a Riemann surface X,*

$$\ell(D) - \ell(K - D) = \deg(D) - g + 1.$$

where

g is genus of surface X.

$\deg(D) =$ the sum of the coefficients occurring in D D.

$\ell(D)$ is the dimension over $\mathbb{C}$ of the vector space of meromorphic functions h on X, such that all the coefficients of $(h) + D$ are non-negative.

This is the space of all meromorphic functions so that if the coefficient in D at z is negative, then h has a zero of at least that multiplicity at z, and if the coefficient in D is positive, then h can have a pole of at most that order.

Theorem 12. *There is an equivalence between the categories of compact Riemann surfaces and smooth complex projective curves.*

More informally, every compact Riemann surface is of the form $\{p(x, y) = 0\}$ for some polynomial p, and every meromorphic function on X is the restriction of a rational function $R(x, y)$ to this variety.

Springer Monographs in Mathematics

Gareth A. Jones
Jürgen Wolfart

Dessins d'Enfants on Riemann Surfaces

Springer

Sketch of a Proof (following Jones and Wolfart, Section 1.2.5):

Projective algebraic curves are compact Riemann surfaces, so need only prove converse.

Idea is to find a pair of meromorphic functions f and g on X that are algebraically dependent, i.e., that $p(f, g) \equiv 0$ for some non-zero polynomial p. Then show X is equivalent to $p(x, y) = 0$.

We do this in eleven steps $1, 2, \ldots 11$.

1. If X is any compact Riemann surface then there is a non-constant meromorphic function f on X.

This follows from definition of Q_a/Q_b earlier (or Riemann-Roch theorem).

Recall that for any $p \in X$ there is a meromorphic function having a simple zero at p.

2. Let $n = \deg(f)$.

Let g be any other meromorphic function on X.

Let F be the finite subset of the sphere consisting of ∞, critical values of f, and the images under f of the poles of g $(= f(g^{-1})(\infty))$.

Then for each $q \in X \setminus F$, $f^{-1}(q) = \{p_1, \ldots, p_n\}$ consists of n distinct points.

Later we will want

(1) f and g to have non-overlapping critical points.

(2) g to have distinct values at $f^{-1}(q)$ for some complex q.

3. Since g is finite at these points we can define the elementary symmetric functions

$$S_1(q) = \sum_j g(p_j)$$

$$S_2(q) = \sum_{i<j} g(p_i)g(p_j)$$

$$\vdots$$

$$S_n(q) = \prod_j g(p_j)$$

where

$$\{p_j\} = f^{-1}(q).$$

Locally these are sums and products of holomorphic inverses of f, and independent of ordering of preimages. By construction, these are single-valued analytic functions on the sphere minus F.

4. We claim each S_k is a rational function.

It is holomorphic except at the finite set F, so it is enough to check it is finite or has poles at these points (no essential singularities).

But $S_k(q)$ is sum and product of values in $\{g(p_j)\}$ and these limit on $f^{-1}(q_0)$ as $q \to q_0$. So S_k has a limit (possibly infinite) at each point of F. Hence no essential singularity

Alternate proof from textbook.

Around each point $q \notin F$ we can use $z = q - q_0$ as a local coordinate (or $z = 1/q$ at ∞),

Each S_j, is represented near each q_0 as a Laurent series in $x^{1/k}$ for some finite k.

Since S_j is single-valued, only integer powers of z can appear in this series.

Since g is meromorphic, only finitely many negative powers can appear.

Thus each S_j is meromorphic on the Riemann sphere, hence rational.

5. For each j, the composition of

$$f : X \to \mathbb{C}^{\infty}$$

and

$$S_j : \mathbb{C}^{\infty} \to \mathbb{C}^{\infty}$$

defines

$$s_j = S_j \circ f : X \to \mathbb{C}^{\infty}$$

that is a meromorphic function on X which is a rational function of f.

6. There is a well-known relationship between the coefficients of a polynomial and the symmetric functions of its roots: $(-1)^r S_r$ is the coefficient of t^{n-r} in the polynomial

$$A(t) = \prod_{j=1}^{n} (t - g(p_j))).$$

Thus

$$A(t) = t^n - S_1 t^{n-1} + \cdots + (-1)^n S_n.$$

If $p \in X \setminus f^{-1}(F)$, let $q = f(p) \notin F$.

Obviously $p = p_j \in f^{-1}(q)$ for some j.

For any $p \in X \setminus f^{-1}(F)$, plug $g(p)$ into

$$A(t) = \prod_{j=1}^{n} (t - g(p_j))).$$

$$A(g(p)) = \prod_{j=1}^{n} (g(p) - g(p_j))) = 0.$$

since $g(p) = g(p_j)$ for some j.

We can write this as

$$a(g) = g^n - s_1 g^{n-1} + \cdots + (-1)^n s_n = 0.$$

where coefficients depend on p, but not g. Note $s_k = S_k \circ f$ as above.

Thus g satisfies a polynomial with coefficients that are rational functions of f, except on a finite set. Hence it satisfies it on all of X.

(7) We claim the polynomial

$$a(t) = t^n - s_1 t^{n-1} + \cdots + (-1)^n s_n.$$

is irreducible in $\mathbb{C}(f)[t]$.

Suppose that it factorises as $a(t) = b(t)c(t)$ where $b, c \in \mathbb{C}(f)[t]$.

If a is reducible, we can choose b, c with strictly smaller degrees.

Meromorphic functions can be chosen to separate pairs of points (that is, to have distinct values at any given pair of points), and hence, one can show they separate any finite set of points (Corollary 9).

Choose some q_0 and then choose the meromorphic function g above so it takes distinct finite values at $\{p_1, \ldots, p_n\} = f^{-1}(q_0)$.

Near p_1, g has a power series $s(z)$ in local coordinates and either $b(s(z)) \equiv 0$ or $c(s(z)) \equiv 0$. Assume the former.

By analytic continuation along paths connecting the p_j, we see $b(g) = 0$ along these paths.

Thus $b(g) = 0$ at all the points $\{p_j\} = f^{-1}(q)$. But at these points b has the same coefficients (since these are polynomials in f and f takes the value q at all these points).

Thus at the points $\{p_j\}$ the polynomial $b = b(f, t)$ is independent of f. and vanishes at the n distinct values $\{g(p_j)\}$.

Hence the degreee of b (in t) is at least $n = \deg(a)$, so a is irreducible.

8. For $p \in X$, let $z = f(p)$ and $w = g(p)$. We have

$$a(w) = w^n - s_1 w^{n-1} + \cdots + (-1)^n s_n$$

where each s_1 is a rational function of z.

Multiplying be the LCM of the denominators, we get a polynomial and w and z

$$P(z, w) = 0$$

$$P(f(p), g(p)) = 0$$

This gives an affine model for X, which can made projective by adding powers of z to make it homogeneous.

9. We are not quite done, since models might not be smooth, i.e., both partials might vanish at some point.

We would be done if we can choose g with distinct critical points from f.

Alternatively, we know that given $z \in X$ there is a meromorphic function g with a simple zero at z.

We can replace g by a finite set of meromorphic functions $\{g_j\}$ so that at least one is un-ramified at each ramification point of f.

The resulting set of algebraic equations yields a nonsingular projective model of X in $\mathbb{P}^M(\mathbb{C})$ for $M < \infty$.

10. This shows the equivalence between compact Riemann surfaces and smooth complex projective algebraic curves.

We also need the equivalence of the morphisms in these categories.

More precisely, any holomorphic map $f : X \to Y$ between compact Riemann surfaces gives a rational map on the corresponding algebraic curves. This is easy if we consider X and Y as algebraic curves in $\mathbb{P}^N(\mathbb{C})$ and $\mathbb{P}^M(\mathbb{C})$.

The graph of f is

$$G_f = \{(p, f(p)) \in X \times Y : p \in X\}.$$

is an algebraic curve in $\mathbb{P}^N(\mathbb{C}) \times \mathbb{P}^M(\mathbb{C})$.

G_f is isomorphic to X via the first coordinate projection. Then $\pi_2 \circ \pi_1^{-1}$ is a rational map from X to Y.

11. The proof shows that every meromorphic g on X is a root of a polynomial of degree at most $n = \deg(f)$ with coefficients in $\mathbb{C}(f)$.

The field of meromorphic functions on X can be identified with the quotient of the polynomial ring $\mathbb{C}(f)[w]$ by the ideal generated by $a(w)$, which is a maximal ideal by the irreducibility of $a(w)$.

If X is given by $F(x, y) = 0$, then $p = (x, y) \to x$ and $p \to y$ are meromorphic functions.

Every rational function $R(x, y)$ gives meromorphic function on X.

Field of meromorphic functions on X is $\mathbb{C}(x, y)$ (rational functions in x and y) modulo the ideal generated by F.

There is an equivalence between:

- Compact Riemann surfaces
- Finite extensions of $\mathbb{C}(x)$ (function fields in one variable)
- Irreducible algebraic curves $F(x, y) = 0$.

There is a functor establishing isomorphism of categories.

How to tell if two surfaces are isomorphic, algebraically

Suppose $X = \{F(x, y) = 0\}$ and $Y = \{G(x, y) = 0\}$.

Defining a holomorphic map $f : X \to Y$ is equivalent to giving a pair of rational maps $R_1 = P_1/Q_1$, and $R_2 = P_2/Q_2$, so that

$$A_1^n Q_2^m G(R_1, R_2) = H \cdot F$$

where $n = \deg_x(G)$, $m = \deg_y(F)$, $H = \mathbb{C}[x, y]$

$f : X \to Y$ is an isomorphism if there is an h so $h \circ f$ is identity. h is given by rational functions $W_1 = U_1/V_1$, $W_2 = U_2/V_2$ so

$$V_1^s V_2^t F(W_1, W_2) = T \cdot G.$$

Doing some algebra shows this occurs iff there are rational functions R_1, R_2 so that

$$Q_1^d Q_2^k (U_1(R_1, R_2) - X V_1(R_1, R_2)) = H_1 F$$
$$Q_1^d Q_2^k (U_2(R_1, R_2) - Y V_1(R_1, R_2)) = H_2 F$$

X and Y are isomorphic iff there are polynomials F_j, Q_j, U_j, V_j, H_j, H and T so that three equalities above hold.

Belyi's Theorem: *A compact Riemann surface S is defined over $\overline{\mathbb{Q}}$ if and only if it supports a Belyi function,i.e., a meromorphic function with at most three critical values.*

Belyi's Theorem: *A compact Riemann surface S is defined over $\overline{\mathbb{Q}}$ if and only if it supports a Belyi function,i.e., a meromorphic function with at most three critical values.*

$\overline{\mathbb{Q}}$ = algebraic closure of rational

Defined over $\overline{\mathbb{Q}}$ means S can be represented at $P(w, z) = 0$ where all coefficients are in $\overline{\mathbb{Q}}$.

Since P has finitely many terms, this happens iff all coefficients are in a single finite extension of $\mathbb{Q}$

Belyi function $\Rightarrow$ Defined over $\overline{\mathbb{Q}}$

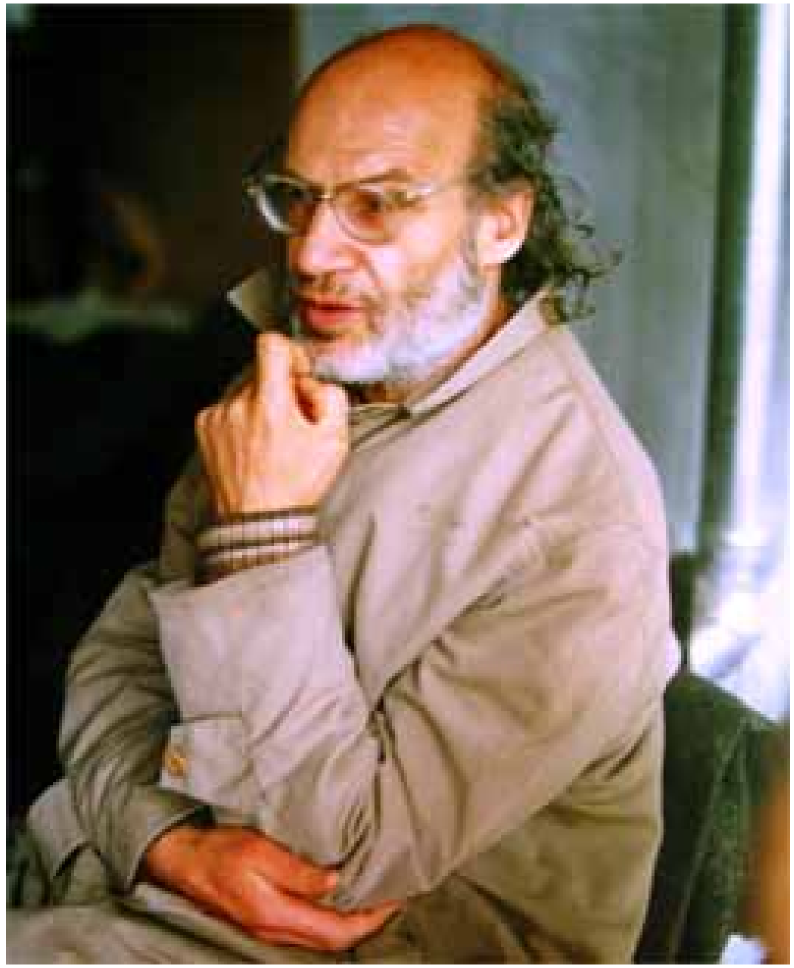
Known to Grothendieck.

Only briefly sketched in Belyi's original paper.

Detailed proof published 25 years later.

We will outline this direction later.

G.V. Belyi



Alexander Grothendieck

Other direction is easier: Defined over $\overline{\mathbb{Q}} \Rightarrow$ Belyi function.

Basic idea:

1. if S is compact and defined over $\overline{\mathbb{Q}}$ then there is a meromorphic f with finitely many critical values, all algebraic integers.

2.Post-compose f with a rational map so all critical values are rational.

3. Post-compose with other rational maps to collapse finite set of rational critical values into a smaller set of rational critical values.

4. Continue until only three left.

Introduction to Compact Riemann Surfaces and Dessins d'Enfants

ERNESTO GIRONDO
GABINO GONZÁLEZ-DIEZ

London Mathematical Society
Student Texts **79**

The London Mathematical Society

CAMBRIDGE

We follow book of Girondo and González-Diez, Section 3.1.

Suppose f is meromorphic on S with only rational critical values (or ∞).

By composing with a Möbius transformation can assume all critical values are $\{0, 1, \infty, \lambda_1, \ldots, \lambda_n\}$ and $\lambda_1 \in (0, 1)$.

Suppose $\lambda_1 = m/(m+n)$, $m, n > 0$ integers. Define

$$P_{m,n}(x) = \frac{(m+1)^{m+n}}{M^m n^n} x^m (1-x)^n.$$

This has critical values only at 0, 1, ∞ and $\lambda_1 = m/(m+n)$ and maps

$$0 \to 0, \quad ,1 \to 0, \quad \infty \to \infty, \quad \lambda \to 1$$

Therefore $P_{m,n} \circ f$ has critical values $0, 1, \infty$ and the at most $n-1$ images of $\lambda_2, \ldots, \lambda_n$.

Continue until only $0, 1, \infty$ remain.

Thus if suffices to prove there is a meromorphic function with only rational critical values.

Suppose S is given by a polynomial

$$F(x, y) = p_0(x)y^n + \ldots + p_n(x)$$

were all coefficients are in $\overline{\mathbb{Q}}$.

Consider the map $(x, y) \to x$. This is a meromorphic function on S.

We claim its critical values are algebraic. Need some results from algebra.

$K = \overline{\mathbb{Q}}$, $K[x]$ = polynomials, $K(x)$ = rational functions.

Weak Bezout's Theorem (Th, 1.84 in GG-D): If $F(z, y)$ and $G(x, y)$ are relatively prime, then $\{F = 0\} \cap \{G = 0\}$ is finite set, with both coordinates in K.

Proof: We may regard F, G as elements of $K(x)[y]$ (polynomials in y with coefficients that are rational functions of x).

Since F, G are co-prime in $K[x, y]$ they are also coprime in $K(x)[y]$. Thus there are $A, B \in \mathbb{C}(x)$ so that

$$1 = AF + BG$$

Clearing denominators, we get $q = (qA)F + (qB)G$

If f, G have infinitely many common roots, these are all roots of the non-zero polynomial q, a contradiction.

If $p = (x, y)$ is a common root, then $q(x) = 0$, so $x \in K$ since K is algebraically closed.

Fixing such an x, gives y as root of $F(x, y) = 0$, a polynomial with coefficients in K, so $y \in K$. $\blacksquare$

Theorem (Thm 1.86 in GG-D): Let

$$F(x, y) = p_0(x)y^n + \cdots + p_n(x).$$

be an irreducible polynomial. The branch points of $(x, y) \to x$ lie in the finite set $S \setminus \{F(x, y) = 0, F_y(x, y) \neq 0, p_0(x) \neq 0\}$.

Theorem (Thm 1.86 in GG-D): Let

$$F(x,y) = p_0(x)y^n + \cdots + p_n(x).$$

be an irreducible polynomial. The branch points of $(x,y) \to x$ lie in the finite set $S \setminus \{F(x,y) = 0, F_y(x,y) \neq 0, p_0(x) \neq 0\}$.

Proof: At other points, implicit function theorem says x is 1-to-1

Since F is irreducible, F and F_y are coprime (otherwise common factor divides F). Hence $\{F = 0\} \cap \{F_y = 0\}$ is finite and these points have coordinates in $\overline{\mathbb{Q}}$ by weak Bezout theorem. ■

This proves our critical values are algebraic.

If critical values are all rational, we are done.

Otherwise let $B_0 = \{\mu_1, \ldots, \mu_s\}$ be critical values.

Let $m_1(t)$ be the minimal polynomial of the finite branch values, i.e., the lowest degree monic polynomial vanishing at all these points.

Equivalently, m_1 is the product of the minimal polynomials of all these algebraic numbers μ_j, avoiding repetition of factors.

Let m_1' be dervative of m_1.

Let $\{\beta_1, \ldots, \beta_d\}$ be roots of m_1'.

Let $p(t)$ the corresponding minimal polynomial.

By definition $\deg(p) \leq \deg(m_1') < \deg(m_1)$.

General fact (chain rule): the critical values of $g \circ f$ are contained in the critical values of g and the g-images of the critical values of f.

By definition of m_1, it maps all finite critical values of $(x, y) \to x$ to 0.

So $(x, y) \to m_1(x)$ has critical values $0, \infty$ and critical values of m_1, denoted $\{\beta_1, \ldots, \beta_d\}$

If this set, denoted B_2, is in $\mathbb{Q} \cup \{\infty\}$ we are done.

Otherwise, let m_2 be the minimal polynomial of the critical points of m_1, i.e., of of $\{m_1(\beta_1), \ldots, m_2(\beta_d)\}$. The dimension of the field extension generated by these points is at most the dimension generated by $\{\beta_1, \ldots, \beta_d\}$.

Thus

$$\deg(m_2) \leq \deg(p) \leq \deg(m_1') < \deg(m_1).$$

we continue this way until all the finite critical values are in $\mathbb{Q}$. This must happen since the degree of the minimal polynomial of the non-rational value decreases by one at each step, so eventually becomes linear. $\blacksquare$

The other direction (Belyi function implies algebraic):

Let Gal($\mathbb{C}$) denote all automorphisms of $\mathbb{C}$.

Any such automorphism acts on polynomials by acting on the coefficients.

Polynomials define surfaces and morphisms between them, so automorphism gives action on compact surfaces and meromorphic functions.

The action on a pair (S, f) preserves the genus of S, the degree of f and applies the automorphism to the critical values.

Theorem (Criterion 3.29 of GG-D): A compact Riemann surface S is defined over $\overline{\mathbb{Q}}$ iff the orbit of S under $\mathrm{Gal}(\mathbb{C})$ is finite.

One direction is easy (if $S = \{(x, y) : P(x, y) = 0\}$ is algebraic then the orbit is finite.

The harder direction easily implies the converse of Belyi's theorem.

Converse of Belyi's theorem, given criterion:

f is a Belyi function on S and $\sigma \in \mathrm{Gal}(\mathbb{C})$, then f^σ is a Belyi function on S^σ of the same degree n.

Thus both S and S^σ are obtained by gluing together $2n$ equilateral triangles. There are only a finite number of distinct ways to to this.

By the criterion above, S is defined over $\overline{\mathbb{Q}}$. $\blacksquare$

Idea of proof of Criterion:

A finite set $\{\pi_1, \dots, \pi_d\}$ is algebraically independent over a field k if

$$p(x_1, \dots, x_d) \to p(\pi_1, \dots, \pi_d)$$

is 1-1 from $k[x_1, \dots, x_d]$ into $\mathbb{C}$.

A specialization of $(\pi_1, \dots \pi_d)$ is a choice of a complex d-tuple $(q_1, \dots q_d)$. The map $\pi_j \to q_j$ defines a k-algebra s homomorphism

$$s : \mathbb{Q}(\pi_1, \dots, \pi_d) \to \mathbb{C}.$$

This distance of a specialization is $\max_j |\pi_j - q_j|$.

We will be interested in cases when distance is small and s maps into $\overline{\mathbb{Q}}$.

A field extension is pure transcendental if it is generated by a set of algebraically independent elements.

General finitely generated extension is $\mathbb{Q}(\pi_1, \ldots, \pi_d; u)$ where u is algebraic over $\mathbb{Q}(\pi_1, \ldots, \pi_d)$.

Suppose $X = \{F(x, y) = 0\}$.

We want to take extension K of $\mathbb{Q}$ generated by coefficients of F.

If this is inside $\overline{\mathbb{Q}}$, X is algebraic, as desired.

Otherwise, we replace transcendental generators by nearby algebraic numbers (a specialization). Use finiteness of orbit to show X is isomorphic to some $Y = \{G(x, y) = 0\}$ with algebraic coefficients.

Let m_u be minimal polynomial of u over $\mathbb{Q}(\pi_1, \ldots, \pi_d)$.

Let m_u^s be image of m_u under specialization s.

Lemma 3.31 of GG-D: *A specialization of* $\mathbb{Q}(\pi_1, \ldots, \pi_d)$ *can be extended to* $\mathbb{Q}(\pi_1, \ldots, \pi_d; u)$ *iff* u *is mapped to a root of* m_u^s.

Lemma 3.32 of GG-D: *Let* $\{u_1, \ldots, u_n\}$ *be roots of* m_u *(minimal poly for u). For any* $\delta > 0$ *there is an* $\epsilon > 0$*, so that if the distance of the specialization is* $< \epsilon$ *then each root of* m_u *is approximated within* δ *by a root of* m_u^s*.*

Sketch: coefficients of m_u are polynomial combinations of field generators, and m_u^s coefficients are same combinations of nearby elements. So $m_u^s \approx m_u$ in coefficients. This implies roots also approximate. ■

Let $X = \{F(x,y) = 0\}$ and let $K_1 = \mathbb{Q}(\pi_1, \dots, \pi_d; v)$ be field generated by coefficients of F.

Let m_v be minimal polynomial of v over $\mathbb{Q}(\pi_1, \dots, \pi_d)$.

For any automorphism σ of $\mathbb{C}$ $K_2 = \mathbb{Q}(\sigma(\pi_1), \dots, \sigma(\pi_d), \sigma(v))$ is field generated by coefficients of F^σ.

Now consider automorphisms σ such that

$$\Sigma_2 = \{\pi_1, \ldots, \pi_d, \sigma(\pi_1), \ldots, \sigma(\pi_d), \}$$

algebraically independent elements.

There are infinitely many such automorphisms σ. Since there are only finitely many possible surfaces X^σ, some coincide, i.e., X^β is isometric to X^τ for some $\beta \neq \tau$.

Then X is isometric to X^σ for $\sigma = \tau^{-1} \circ \beta$.

Recall this isometry is equivalent to certain polynomials existing.

Enlarge Σ_2 by adding coefficients of the polynomials P_j, Q_j, U_j, V_j, T, H_j, H that express the isometry $X \simeq X^\sigma$.

We get larger field

$$K_3 = \mathbb{Q}(\pi_1, \ldots, \pi_n; u)$$

with $n \geq 2d$,, and u algebraic over $\mathbb{Q}(\pi_1, \ldots, \pi_n)$.

Specialize the coordinates $> d$ with elements of $\mathbb{Q}(\sqrt{-1})$ with small distance (enough to apply Lemma 3.32 of GG-D above about approximating roots).

Let s to the associated homomorphism.

Recall $\mathbb{Q}(\pi_1, \ldots, \pi_n; u)$ is ring of rational linear combinations of generators.

$\mathbb{Q}[\pi_1, \ldots, \pi_n; u]$ is field of fractions of these i.e.,

$$z = \frac{A(\pi_1, \ldots, \pi_n, u)}{B(\pi_1, \ldots, \pi_n, u)}.$$

Let $\mathbb{Q}[\pi_1, \ldots, \pi_n; u]_s$ be sub-ring of field where image of denominator is non-zero under s.

In other words $z = A/B$ is in $\mathbb{Q}[\pi_1, \ldots, \pi_n; u]_s$ if

$$s(B) = B(\pi_1, \ldots, \pi_d, q_{d+1}, \ldots, q_n, u_s) \neq 0.$$

Then s extends to a homomorphism of $\mathbb{Q}[\pi_1, \ldots, \pi_n; u]_s$ into $\mathbb{C}$.

Thus s extends to a homomorphism

$$\mathbb{Q}[\pi_1, \ldots, \pi_n; u]_s(x, y) \to \mathbb{C}(x, y)).$$

Here is the main point.

If the elements q_j are chosen sufficiently close to π_j, then all the elements of the finite set consisting of the coefficients of the polynomials $P_j, Q_j, U_j, V_j, T, H_j, H$ along with the element $v \in K_1$ all lie in $\mathbb{Q}[\pi_1, \ldots, \pi_n; u]_s$.

(A fixed non-zero element maps to a non-zero element if the distance of the specialization is small enough. We only need this for finitely many elements.)

We can therefore apply specialization to our polynomials that verify X is isometric to X^σ, i.e.,

$$Q_1^n Q_2^m G(R_1, R_2) = H \cdot F$$

becomes

$$(Q_1^n)^s (Q_2^m)^s G^s(R_1^s, R_2^s) = H^s \cdot F^s$$

The latter defines a morphism X_{F^s} to $X_{(F^\sigma)^s}$.

Doing the same for the other relations defines an isomorphism between X_{F^s} and $X_{(F^\sigma)^s}$.

Now, by construction, the coefficients of the polynomial $(F^{\sigma})^s$ are in field generated by elements of $\mathbb{Q}(\sqrt{-1})$ plus an element $s(\sigma(v))$ that is algebraic over this field.

Thus $(F^{\sigma})^s$ is algebraic.

Since X_{F^s} and $X_{(F^{\sigma})^s}$ are isomorphic, and the latter is algebraic, X_{F^s} is also algebraic.

We are done if $F = F^s$.

Specialization leaves $\pi_1, \ldots, \pi_d$ fixed, so it suffices to show $s(v) = v$.

Note that $m_v^s = m_v(v)$, again since π_j don't change.

By taking the distance of the specialization small enough, $s(v)$ a root of m_v as close to v as we wish, hence equals v.

This proves a Riemann surface with finite orbit is algebraic.

Completes proof of Belyi's theorem.

Argument using specialization to prove Belyi's theorem also shows that Belyi functions are defined over $\overline{\mathbb{Q}}$.

See Proposition 3.34 of Girondo and González-Diez.