Proceedings of Symposia in Pure Mathematics

What Gauss told Riemann about Abel's Theorem

Michael D. Fried

Abstract. Historian Otto Neuenschwanden studied Riemann's library record in Göttingen. He also consulted job position letters in Germany in the 3rd quarter of the 19th century. These show Riemann relied on personal discussions with Gauss (in the late 1840's) about harmonic functions. Mathematicians rejected that early approach after Riemann's death (in 1866) until near the end of the 20th century's 1st quarter.

Using Riemann's theta functions required generalizing Abel's work in two distinct ways. Abel compared functions on the universal cover of a complex torus; everything was in one place. Riemann had to compare functions from two different types of universal covers. We explain struggles over interpreting modern algebraic equations from the Gauss-Riemann approach.

Abel's work on elliptic integrals motivated approaches to algebraic equations that we don't easily connect to him through just his published papers. There is still no simple route through the beginnings of the subject. A student runs quickly into difficulties that call into question these topics.

- How may one divide the vast finite group area between those that are nilpotent and those that are simple?
- How may one meld existence results from complex variables with manipulative needs of algebraic equations?

[**Fr03**] and [**Vö96**] have background on analytic continuation. They cover many issues on Riemann's Existence Theorem toward modern applications. All detailed definitions in this discussion are in [**Fr03**]. We explain here subtleties in Riemann's approach and why there is so much modern work on problems one might have thought solved long ago.

We used the technical equation-oriented [**Fay73**] as a personal encyclopedia on *what counts* to theta afficianados. We also respond to a challenge from [**FaK01**, Chap. 2,§8]:

> ..., partial motivation for the results discussed so far is to better understand the multivariable case.

2000 *Mathematics Subject Classification.* Primary 11F32, 11G18, 11R58; Secondary 20B05, 20C25, 20D25, 20E18, 20F34.

Acknowledgement of support from NSF #DMS-0202259, and appropriate regards for the time spent as a senior research Alexander von Humboldt fellow in 1995–1996. That period showed me Gauss' influence through unpublished intellectual activities when I learned of his influence on the eclectic endeavors of his friend Humboldt.

[**FaK01**] only considers theta functions attached to elliptic curves, or one-dimensional thetas. That is because they consider only theta functions (and theta nulls) varying along modular curves, natural covers of the j-line that first arose in Abel's work (§4.2.3). Likely the term *multivariable* in [**FaK01**] means thetas attached to Siegel modular spaces. Our examples, however, note natural systems of j-line covers for which the corresponding thetas are multivariable. Discussions of Gauss and Riemann to which [**Ne81**] refers easily motivate these examples.

Few specialized territories exceed theta functions for giving the feeling its adherants must slave endlessly over intricate details. Yet, even here we suggest, one needn't know everything to claim knowing something significant.

1. Interpreting Abel's Problem

My notation for the Riemann sphere is $\mathbb{P}^1_z = \mathbb{C}_z \cup \{\infty\}$. This indicates for different copies of $\mathbb{C}$ the chosen variable may change. For a finite set $\boldsymbol{z}$ on $\mathbb{P}^1_z$, denote its complement by $U_{\boldsymbol{z}}$. We use the fundamental group $\pi_1(U_{\boldsymbol{z}}, z_0)$ and related fundamental groups. We start with a general problem.

1.1. Antiderivatives.

Let $\varphi : X \to \mathbb{P}^1_z$ be a map between compact 2-dimensional manifolds (the genus g_X of X comes later) with all but finitely many points $\boldsymbol{z} = \{z_1, \dots, z_r\}$ of $\mathbb{P}^1_z$ having exactly n points above them. Then φ gives X a complex structure. By that we mean, there is a unique complex manifold structure on X so that when we pull back the local functions on $U_{\boldsymbol{z}}$ analytic in z to $X \setminus \{\varphi^{-1}(\boldsymbol{z})\}$ we get the functions we call analytic on X. That means there is a unique way to declare the functions $\mathcal{H}(V)$ analytic in a neighborhood V of $x' \in \varphi^{-1}(\boldsymbol{z})$ to have two properties:

(1.1a) Functions of $\mathcal{H}(V)$ restricted to $V \setminus \{\varphi^{-1}(\boldsymbol{z})\}$ are locally the pullback of analytic functions in z.

(1.1b) There is a (possibly smaller) neighborhood W of x' in V to which the functions called analytic there form a ring isomorphic to the convergent power series in a disk about the origin on $\mathbb{C}_w$.

Notice in this formulation: Pullback of z interprets z around each point as a meromorphic function. It is a global meromorphic function on X.

The simplest case of this situation is given by $f : \mathbb{P}^1_w \to \mathbb{P}^1_z$ by $f(w) = w^e$ for some positive integer e. Even if we don't a' priori declare w as the variable from which we get our local analytic functions around 0 (convergent power series in w) and ∞ (convergent power series in $1/w$) on $\mathbb{P}^1_w$, it will be forced on us if we declare $f : \mathbb{P}^1_w \setminus \{0, \infty\} \to \mathbb{P}^1_z \setminus \{0, \infty\}$ analytic according to (1.1). Since these rules produce a local variable $w_{x'}$ giving the local analytic functions around $x' \in X$, we also have the idea of a local analytic differential $h_{x'}(w_{x'})\, dw_{x'}$.

Suppose ω is a global holomorphic differential on X. That means at each $x' \in X$ we give a neighborhood $V_{x'}$ and an expression $h_{x'}(w_{x'})\, dw_{x'}$ with $h_{x'}$ a holomorphic function of $w_{x'}$ subject to the following compatibility condition.

(1.2) If $\gamma : [0, 1] \to V_{x'} \cap V_{x''}$ is any (continuous) path then

$$\int_\gamma h_{x'}(w_{x'})\, dw_{x'} \stackrel{\text{def}}{=} \int_0^1 h_{x'}(w_{x'} \circ \gamma(t)) \frac{dw_{x'} \circ \gamma(t)}{dt} dt = \int_\gamma h_{x''}(w_{x''})\, dw_{x''}.$$

Anything contributing to adding up information along a path comes from integration of a tensor. The specific tensors that are holomorphic differentials correspond to information attached to potentials, like gravity and electric charge.

It is a fundamental in the situation of $\varphi : X \to \mathbb{P}^1_z$ that in addition to the global meromorphic function z, there is another global meromorphic function w on X defined by $\varphi_w : X \to \mathbb{P}^1_w$ so as to have the following property.

PROPOSITION 1.1 (Function half of Riemann's Existence Theorem). *We may express any local analytic function in a neighborhood of $x' \in X$ as $h(u)$ where u is a rational function in z and w and h is a convergent power series.*

Prop. 1.1 suggests saying that z and w together *algebraically uniformize* X. We differentiate this from uniformization in the sense that there is a map from a (simply-connected) universal covering space to X. Though the two notions of uniformization are related, there survive to this day many mysteries about this relation. Prop. 1.1 is obvious in Abel's situation, because it was easy to construct one such w since Abel started with a Galois cover of $\mathbb{P}^1_z$ with a group that is abelian ([**Fr03**, Chap. 1, §8.2] calls this the abelian form of Riemann's Existence Theorem). The rubric when φ is an abelian cover constructs w from branches of logarithm. So, it truly belongs to elementary complex variables, except that understanding this requires mastering analytic continuation, maybe not such an elementary idea [**Fr03**, Chap. 1, §7]. Deeper issues, however, about Prop. 1.1 arose from the result we call Abel's Theorem. Gauss discussed these things with Riemann. Especially we call attention to the combinatorial half of Riemann's Existence Theorem, closely related to the topic of *cuts* (§4.1).

EXAMPLE 1.2 (Finding where a differential is holomorphic). Consider the differential $dz/w = \omega$ on $\mathbb{P}^1_w$ where $f : \mathbb{P}^1_w \to \mathbb{P}^1_z$ by $f(w) = w^e$. The apparent zero in the denominator of ω cancels with material from the numerator, so ω is holomorphic at $w = 0$. In detail: w is a uniformizing parameter in a neighborhood of 0. From $z = w^e$ express ω as $ew^{e-1}\,dw/w = ew^{e-2}\,dw$. This expression actually holds globally where w is a uniformizing parameter. At ∞, however, it is $1/w = u$ that gives a uniformizing parameter. So $dz/w = udu^{-e} = -eu^{-e}, du$ has a pole at ∞. Notice: dz/w^{e+1} *is* holomorphic at ∞.

1.2. The gist of Abel's Theorem. The word *nonobvious* is inadequate to express the mystery in finding even one such w, whenever φ is not abelian. Abel's famous theorem showed how to construct *all* such global functions w, not just one, in his special case. Our discussion is about how Gauss and Riemann may have viewed Abel's Theorem and his special case, and how this view generated in Riemann the tools for its generalization to all $\varphi : X \to \mathbb{P}^1_z$. Our documentation of the conversations between Gauss and Riemann has much of heresay, and this author is not an historian. So, to the cogent remarks of [**Ne81**] we have added some observations of two types:

- That the subject is still very difficult today, even for experts.
- That the difficulty for even Gauss and Riemann sent them in directions still inadequately understood today.

PROBLEM 1.3 (Abel's First Question). What is the relation between elementary substitutions for the integration variable t in $\int_\gamma \omega$ where γ is a closed path and algebraic manipulations of functions coming from φ? Similarly, how do either of these substitutions relate to deforming the set of branch points $\boldsymbol{z}$?

Our question is much more general than Abel's because we ask it for general φ. We should understand the question had a long formulation, much more precise than given here, starting from a specific case of Euler and considered extensively by Legendre, long before Abel. Gauss may have already had a good sense of Abel's Theorem by the time he met Abel.

Abel investigated this fundamental problem in the example where X is a compactification of the algebraic set $\{(z,w) \mid w^2 = z^3 + cz + d\}$ and $\varphi : (z,w) \mapsto z$ is the standard projection. His formulation allowed the following operations as acceptable. Including iterated compositions of rational functions in z, multiplication by constants, selecting from *elementary functions* regarded as known, and

(1.3) the conceptual addition of functional inverse.

The problem of what we allow for elementary functions in t is handled by a key conceptual idea from complex variables. There really are but two elementary analytic functions using (1.3): z and $\log(z)$.

EXAMPLE 1.4. $\cos(z) = (e^{iz} + e^{-iz})/2$. Example: $z^{1/k} = e^{\log(z)/k}$.

We profitably consider integrals locally as an antiderivative in z. The problem is to investigate how far integration removes us from elementary functions.

1.3. Details on Abel's integral. Let $g(u)$ be a local right inverse to $f(z)$: $f(g(u)) \equiv u$. Apply the chain rule:

$$\frac{df}{dz}_{|_{z=g(u)}} \frac{dg}{du} = 1. \tag{1.4}$$

Therefore, $\frac{dg}{du} = 1/\frac{df}{dz}_{|_{z=g(u)}}$. This is the complex variable variant of how first year calculus computes an antiderivative of inverse trigonometric functions. Abel applied this to a (right) inverse of a branch of primitive from the following integral

$$\int_\gamma \frac{dz}{(z^3 + cz + d)^{\frac{1}{2}}} \tag{1.5}$$

with $c, d \in \mathbb{C}$.

If $g(z)$ is a branch of $\log(z)$ near z_0, so $e^{g(z)} = z$, then $e^{g(f(z))/k}$ is a branch of $f(z)^{1/k}$ on *any disk* (or on any simply connnected set) avoiding the zeros and poles of f [**Fr03**, Chap. 1, §6]. Related to this, there is a branch of $f(z)^{1/k}$ along *any path* in $U_{\boldsymbol{z}}$ ($\boldsymbol{z}$ containing the zeros and poles of f).

Consider $h(z)\,dz = \frac{dz}{(z^3+cz+d)^{\frac{1}{2}}}$ around some base point z_0: $h(z) = h_{c,d}(z)$ is a branch of $(z^3 + cz + d)^{-\frac{1}{2}}$. Let $f(z) = f_{c,d}(z)$ be a primitive for $h(z)\,dz$. Apply (1.4) to $f(g(u)) = u$:

(1.6) $\frac{dg(u)}{du} = (g(u)^3 + cg(u) + d)^{\frac{1}{2}}$, a parametrization of the algebraic curve from the equation $w^2 = z^3 + cz + d$.

Let $\boldsymbol{z} = \boldsymbol{z}_{c,d} = \{z_1, z_2, z_3, \infty\}$, the three (assumed distinct) zeros of $z^3 + cz + d$ and ∞. Denote the set of such $(c,d) \in \mathbb{C}^2$ by V^0. Denote the points on (z,w) satisfying (1.6) that are over $U_{\boldsymbol{z}}$ by $X^0_{c,d}$. It has a unique complex manifold compactification $X_{c,d}$.

LEMMA 1.5. *We may integrate $h_{c,d}\,dz = \omega_{c,d}$ along any path in $X_{c,d}$ (the point of it being a holomorphic integral).*

PROOF. Since there are places where the denominator of $h_{c,d}\,dz$ is 0, it behooves us to explain why the differential has no poles on $X_{c,d}$. For this, use the argument of Ex. 1.2 to express $h_{c,d}\,dz$ locally in a uniformizing parameter. Only the points in $\boldsymbol{z}$ cause any problems, and the point is that for each we should use $w = \sqrt{z - z_i}$, $i = 1, 2, 3$ (at ∞, use $\sqrt{1/z}$). $\square$

The notation $\omega_{c,d}$ for $h_{c,d}\,dz$ is especially fitting when we emphasize that we are restricting $h_{c,d}\,dz$ to paths in $X_{c,d}$.

1.4. Applying analytic continuation to the lattice from $h_{c,d}(z)\,dz$. Analytically continuing a primitive $f_{c,d}(z)$ of $(z^3 + cz + d)^{-\frac{1}{2}}$ along any closed $\gamma \in \pi_1(U_{\boldsymbol{z}}, z_0)$ produces a new analytic function $\mathrm{Int}(h_{c,d}(z)\,dz)_\gamma$ around z_0. This gives a collection of functions $\mathcal{A}_f(U_{\boldsymbol{z}}, z_0) = \{\mathrm{Int}(h_{c,d}(z)\,dz)_\gamma\}_{\gamma\in\pi_1(U_{\boldsymbol{z}},z_0)}$ around z_0.

1.4.1. *Analytically continuing* $\mathrm{Int}(h_{c,d}(z)\,dz)$ *everywhere in* $X_{c,d}$. The related subset is $\mathcal{A}_f(X^0_{c,d}, x_0)$: Analytic continuations along closed paths from $X^0_{c,d}$. Then,

$$u \in \mathbb{C}_u \mapsto \psi(u) \overset{\text{def}}{=} (g_{c,d}(u), \frac{dg_{c,d}}{du}(u)) \tag{1.7}$$

is one-one up to translation by elements of

$$L_{c,d} = \{s_\gamma = \int_\gamma h_{c,d}(z)\,dz, \gamma \in \pi_1(X_{c,d}, x_0)\}.$$

A primitive for $h_{c,d}\,dz$ analytically continues along every closed path in $X_{c,d}$ (§1.3). Since $h_{c,d}\,dz$ is a global holomorphic differential, the set $\mathcal{A}_f(X^0_{c,d}, x_0)$ depends only on the image of a path in $H_1(X_{c,d}, \mathbb{Z})$. There is no canonical way to choose representing paths $\gamma_1, \gamma_2 \in \pi_1(U_{\boldsymbol{z}}, z_0)$ that lift to generators of $H_1(X_{c,d}, \mathbb{Z})$. Still, it helps to know the double of any $\gamma \in \pi_1(U_{\boldsymbol{z}}, z_0)$ lifts to $\pi_1(X^0_{c,d}, x_0)$. Integrate $h_{c,d}(z)\,dz$ around such generating γ_1, γ_2 to get s_1 and s_2, so s_1, s_2 generate $L_{c,d}$. The following says $L_{c,d}$ is a lattice in $\mathbb{C}_u$.

PROPOSITION 1.6. *The elements s_1 and s_2 are linearly independent over the reals. So, $\psi(u)$ in* (1.7) *gives an analytic isomorphism between $X_{c,d}$ and $\mathbb{C}_u/L_{c,d}$.*

PROOF. If both s_1 and s_2 are 0, then the holomorphic differential $h_{c,d}\,dz$ is the differential of an analytic function f on $X_{c,d}$. This f would therefore have no poles on $X_{c,d}$. So, it would be an analytic function on a compact Riemann surface, and from the maximum principle, such a function is constant.

Now we know one of s_1 and s_2 is not 0. Assume that is s_1. Replace $h_{c,d}\,dz$ by $\frac{1}{s_1}h_{c,d}\,dz$ to assume s_1 is 1. We have only to show s_2 cannot be real. Suppose it is. Break $h_{c,d}\,dz$ into its real and imaginary parts. Then, the imaginary part has 0 periods. So, again, it is the differential of a harmonic function on a compact surface, defying the maximum principle for harmonic functions. $\square$

1.4.2. *Local exactness implies exactness on the universal covering space.* First year calculus: The inverse $g_1(u)$ of a primitive of $h_1(z) = (z^2 + cz + d)^{-\frac{1}{2}}$ is a function of $\sin(u)$. This has a unique analytic continuation everywhere in $\mathbb{C}$. We have just expressed Abel's discovery the same is true for the inverse $g(u) = g_{c,d}(u)$ of $f_{c,d}(z)$; it extends everywhere in $\mathbb{C}_u$ meromorphically.

Even to seasoned Riemann surface enthusiasts it must be a little puzzling as to which to emphasize first, differentials or functions. We suggest Gauss and Riemann discussed this, and came to the conclusion that there were ways they could be put on the par. The log-differential equivalent (Prop.2.5) of the next argument expresses the defining properties of θ functions.

Proposition 1.7. *For $\tilde\varphi : \tilde X \to X$ the universal cover of X, pullback of any holomorphic differential ω on X is an exact differential on $\tilde X$.*

Interpreting exactness. The definitions give this interpretation to the differential ω using a coordinate chart $\{(U_\alpha,\psi_\alpha)\}_{\alpha\in I}$ for the complex structure on X. As usual, $\psi_\alpha : U_\alpha \to \mathbb{C}_{z_\alpha}$ has analytic transition functions $\psi_\beta \circ \psi_{alpha}^{-1} : \psi_\alpha(U_\alpha \cap U_\beta) \to \psi_\beta(U_\alpha\cap U_\beta)$. This has the following properties [**Fr03**, §5.2 Chap. 3].

(1.8a) ω retricted to U_α is given by pullback from $dF_\alpha(z_\alpha)$ with on $F_\alpha(z_\alpha)$ analytic on $\psi(U_\alpha)$, $\alpha \in I$.

(1.8b) Restricting $\tilde\varphi$ to $\tilde\varphi^{-1}(U_\alpha)$ is one-one to U_α on each component, $\alpha \in I$.

Consider any (piecewise differentiable) path $\gamma : [0,1] \to X$ with $\gamma(0) \in U_{\alpha_0}$. A primitive for ω with initial value $F_{\alpha_0}(\psi(\gamma(t))$ along the path is a continuous $F : [0,1] \to \mathbb{C}$ so for any $t_0 \in [0,1]$ the following holds. If $\gamma(t) \in U_\beta$ for t close to t_0, then $F(t) = F_\beta(\psi_\beta(\gamma(t))) + c_\beta$ with c_β constant (in t). This assures for t close to 1, $F(t)$ is the restriction of an analytic function in a neighborhood of $\gamma(1)$. We have analytic continuations of a meromorphic function along every path. So, the Monodromy Theorem [**Fr03**, Chap. 3 Prop. 6.11] gives an analytic $\tilde F$ on $\tilde X$ with

$$\tilde F(\tilde x) - F_\alpha(\psi_\alpha(\tilde\varphi(\tilde x))) \text{ constant on any component of } \tilde\varphi^{-1}(U_\alpha). \tag{1.9}$$

Then, the differential of $\tilde F$ is the pullback of ω to $\tilde X$. That is the meaning of exactness in the proposition. □

2. Abel's Theorem

Before we launch into our main points, we note the formula for the theta function on a complex torus is in [**Ah79**, Chap. 7]. It differs, however, in that we don't start with a doubly periodic function, but a holomorphic differential on a Riemann surface. Secondly: We will see modular curves arise, Abel being the first to produce them, and Galois one of the first to apply them. Thirdly: We will ask questions about the nature of the functions described by Abel that influenced Gauss and Riemann. These have no appearance in [**Ah79**] because they give nonabelian covers. Thoughout Riemann's generalization, holomorphic differentials are the starting point. Modern treatments of the topic tend just to give the σ, eschewing our discussion on properties of the sought for object. In the forty years, however, between Abel's work and Riemann's generalization there was time for such contemplation.

2.1. Substitutions versus field operations.

A substitution in a variable corresponds to a composition of functions. Suppose we substitute $w(z)$ for z to get $f_{a,b}(w(z))$ from $f_{a,b}$. Consider values of w where $w(z)$ has a local inverse, which we express as $w^{-1}(z)$. Then, rewrite $f \circ g(u) = u$ as $f \circ w \circ w^{-1} \circ g(u) = u$.

Investigating the substitution $f \circ w(z)$ is equivalent to considering the composition $w^{-1} \circ g(u)$. An important case is when w^{-1} is a rational function. The two functions $g_{c,d}(u)$ and its derivative in u generate a field over $\mathbb{C}$. Denote this $M_{c,d}$. As $g_{c,d}$ has up to translation a unique pole of order 2 at $u = 0$, and no residue, we can guarantee $g(u) - g(-u)$ is 0 at the origin, bounded everywhere, and so is indentically 0: $g(u)$ is *even*. From (1.6) this is closed under taking derivatives in u. Abel rephrased his investigation.

Problem 2.1. For what pairs (c,d) and (c',d') is $g_{c',d'}$ an element of $M_{c,d}$?

While technically this doesn't pin down all allowable substitutions, the answer shows the whole story. For each (c,d), $g_{c,d}(u)$ is to the exponential function e^u as $f_{c,d}(z)$ in (1.5) is to a branch of $\log(z)$.

Abel described every element of $M_{c,d}$. So, describing elements of $M_{c,d}$ is the same as describing analytic maps $\varphi : \mathbb{C}/L_{c,d} \to \mathbb{P}^1_w$ [**Fr03**, Chap. 4, Prop. 2.10]. Such a map has as many zeros $D_0(\varphi) = \{a_1, \dots, a_n\}$ (with multiplicity) as it has poles $D_\infty(\varphi) = \{b_1, \dots, b_n\}$ [**Fr03**, Chap. 4, Lem. 2.1].

Further, describing the function fields $M_{c',d'} \subset Mc,d$ is the same as describing the analytic maps $\psi_{(c,d),(c',d')} : \mathbb{C}/L_{c,d} \to \mathbb{C}/L_{c',d'}$. Such a map has a degree, and Abel described a valuable equivalence class of such maps $\psi_{(c,d),(c',d')}$ of each prime degree p. Modern notation calls that equivalence the *modular curve* $X_0(p)$: It relates the *j-invariant* $j(\boldsymbol{z}_{c,d})$ of $\boldsymbol{z}_{c,d}$ to that of $\boldsymbol{z}_{c',d'}$. §4.2.3 describes this in detail.

2.2. Analytic maps $\mathbb{P}^1_u \to \mathbb{P}^1_z$. Let $\varphi : X \to \mathbb{P}^1_z$ denote a meromorphic function on a compact Riemann surface. Use $\mathbb{C}(X)$ for the complete collection of these. They form a field, using addition and multiplication of any two functions φ_1 and φ_2: $\varphi_1 \cdot \varphi_2(x) = \varphi_1(x)\varphi_2(x)$ with the understanding you resolve the meaning of the product when $\varphi_1(x) = 0$ and $\varphi_2(x) = 0$ by expressing the functions in a local analytic parameter. Up to a multiplicative constant $D_0(\varphi)$ and $D_\infty(\varphi)$ determine φ. Conversely, suppose $\varphi : \mathbb{P}^1_u \to \mathbb{P}^1_z$. Up to multiplication by $\mathbb{C}^*$,

$$\varphi(u) = \frac{\prod_{i=1}^n (u - a_i)}{\prod_{i=1}^n (u - b_i)}.$$

Replace $u - a_i$ or $u - b_i$ by 1, if either a_i or b_i is ∞. So, u is an odd function (exactly one zero of multiplicity one, at $u = 0$) whose translations give the local behavior of $\varphi(u)$. From it we craft the desired function $\varphi(u)$ having the right zeros and poles.

Given maps between $\mathbb{P}^1_z$ and $\mathbb{P}^1_u$, it is to our advantage to have the expression of functions on $\mathbb{P}^1_u$ pulled back to $\mathbb{P}^1_z$ be compatible with that on $\mathbb{P}^1_z$. The oddness of u is a nice normalizing condition, leaving the only ambiguity in the choice of u multiplication by an element of $\mathbb{C}^*$. It is forced on us by this condition: We want a globally defined function with exactly one zero of multiplicity one at $u = 0$. With u already attached to $\mathbb{P}^1_u$, it is part of our naming rubric to have normalized u.

2.3. Log-differentials and Imitating the genus 0 case. Suppose there is a nonconstant analytic map $\varphi : \mathbb{C}/L_{c,d} = X_{c,d} \to \mathbb{P}^1_w$ with $D^0 = \varphi^{-1}(0)$ and $D^\infty = \varphi^{-1}(\infty)$. Denote its branch points by $\boldsymbol{w} = \{w_1, \dots, w_{r'}\}$. Tricky notation: Denote the subset of $X_{c,d}$ over $U_{\boldsymbol{w}}$ by $X^{w,0}_{c,d}$. Take a path γ from 0 to ∞ on $U_{\boldsymbol{w}}$.

2.3.1. *Abel's necessary condition.* Let γ_i be the unique lift to $X^{w,0}_{c,d}$ of γ starting at a_i (it will end in D^∞). If $0 \in \boldsymbol{w}$ or $\infty \in \boldsymbol{w}$ (D^0 or D^∞ have points with multiplicity), lift γ without its endpoints. Then take the closures of these paths.

Proposition 2.2 (Abel's necessary condition). *Then,* $\sum_{i=1}^n \int_{\gamma_i} h_{c,d}\, dz = 0$. *So, existance for φ requires there are paths $\{\gamma'_i\}_{i=1}^\infty$ on $X_{c,d}$ with initial points D^0, end points D^∞ and*

$$\sum_{i=1}^n \int_{\gamma'_i} h_{c,d}\, dz = 0.$$

Comments. Here are brief details from the argument of [**Fr03**, Chap. 4, §2.6.1] (a rougher version is in [**Spr57**, Thm. 10-22]). Suppose $\varphi : X \to \mathbb{P}^1_w$ is a finite map of analytic manifolds. For any differential ω (not necessarily holomorphic) on X

there is differential $\mathrm{t}(\omega)$ (the *trace*) on $\mathbb{P}^1_w$ so that $\sum_{i=1}^n \int_{\gamma_i} \omega = \int_\gamma \mathrm{t}(\omega)$. Further, $\mathrm{t}(\omega)$ is holomorphic if ω is. So, in that case, $\mathrm{t}(\omega) = 0$ since there are no global holomorphic differentials on $\mathbb{P}^1_w$. As above we need γ to avoid the branch points of φ, *except* at its end points. So, this does allow zeros and poles to appear with multiplicity. To handle that case, take the lifts of γ with its endpoints, and then just take the closures of the resulting paths on X. □

2.3.2. *The log-differential property.* Here and for general X replacing $X_{c,d}$, we emulate aspects of the sphere case. Find a function $\sigma(u)$ with a zero of multiplicity one whose translates craft a given function through its zeros and poles in the sense of expression (2.1). (Significantly, we will find we are able to normalize it to be odd.) There is no such function on $\mathbb{C}/L_{c,d}$: Or else that function would give a one-one onto map of the complex torus to the sphere. Use this notation: $u \in \mathbb{C}_u \mapsto [u] \in \mathbb{C}/L_{c,d}$. The function u on $\mathbb{P}^1_u$ has only one zero, though it also has one pole (∞).

Definition 2.3. A meromorphic differential ω on a Riemann surface X is a *log differential* if $\omega = dv/v$ for some meromorphic v on X. It is *local log* if locally around each point it is a log differential.

Any local log differential ω has its poles of multiplicity 1, and its periods along paths bounding a disc about a pole are integer multiples of $2\pi i$. On a compact X, the sum of the residues of ω is 0: It defines a degree 0 *polar divisor* D_ω.

Example 2.4. On $\mathbb{P}^1_u$, When $\varphi(u) = u$ the log differential has poles at 0 and ∞ with residue respectively $+1$ and -1 according to Ex. 1.2.

Here is the analog for local log differentials of Prop. 1.7.

Proposition 2.5. *For $\tilde\varphi : \tilde X \to X$ the universal cover of X, the pullback of any local log differential ω on X is a log differential on $\tilde X$.*

Proof. Use notation of the proof of Prop. 1.7. For each $\alpha \in I$, denote restriction of ω to U_α by ω_α which by hypothesis expresses as $dv_\alpha(z_\alpha)/v_\alpha(z_\alpha)$ on $\psi_\alpha(U_\alpha)$. The general principle is that this is a differential equation for $v_\alpha(\psi_\alpha(x)$, $x \in U_\alpha$. So, analytic continuations of v_α along any path $\gamma : [0,1] \to X$ are uniquely well-defined. As a reminder of this, it comes to considering the case of continuing to $U_\alpha \cup U_\beta$ from its solutions $v_\alpha(\psi_\alpha(x))$, $x \in U_\alpha$, and $v_\beta(\psi_\beta(x))$, $x \in U_\beta$, with $U_\alpha \cap U_\beta$ connected. Further, on the overlap the two functions differ only by multiplication by a constant $c_{\beta,\alpha}$. To extend the solution requires only to multiply $v_\beta(\psi_\beta(x))$ by $1/c_{\beta,\alpha}$. It now matches $v_\alpha(\psi_\alpha(x))$ on the overlap.

Again, apply the Monodromy Theorem to conclude the existence of $\tilde v$ on $\tilde X$ so the pullback $\tilde\omega$ of ω has the form $d\tilde v/\tilde v$. □

2.4. Exactness on $\tilde X$ and the θ property. The proof of Prop. 2.5 produces from ω a 1-cocycle $c_\omega = \{c_{\beta,\alpha}\}_{\alpha,\beta \in I \times I}$. The co-cycle property $c_{\gamma,\beta}c_{\beta,\alpha} = c_{\gamma,\alpha}$ for $U_\alpha \cap U_\beta \cap U_\gamma$ nonempty, follows immediately from the uniqueness up to multiplicative constant statement in the proof.

2.4.1. *A locally constant sheaf from a local log differential.* In modern parlance: The collection $\{v_\alpha\}_{\alpha \in I}$ is a meromorphic global section of the locally constant fiber bundle (sheaf) defined by c_ω. In turn this cocycle gives an element of $\mathrm{Hom}(\pi_1(X, x_0), \mathbb{C}^*)$. If $f : X \to \mathbb{P}^1_z$ is any meromorphic function on X, then $\{fv_\alpha\}_{\alpha \in I}$ is another meromorphic section of this sheaf. The local log differential attached to this meromorphic section is $\omega + df/f$. This has divisor of poles the divisor of poles of ω plus the divisor of f. We summarise.

PROPOSITION 2.6. *The divisor of poles D_ω of any log differential ω on a compact X defines a (degree 0) divisor class $[D_\omega]$ and a locally constant sheaf $\mathcal{L}_\omega$. Its global meromorphic sections are $\{fv_\alpha\}_{\alpha\in I}$ running over all $f\in\mathbb{C}(X)$.*

Suppose ω' is another local log divisor for which $D_\omega = D_{\omega'}$. Then, the difference $\omega-\omega'$ is a holomorphic differential on X. If $\omega-\omega'\ne 0$, $\mathcal{L}_\omega$ is holomorphically isomorphic to $\mathcal{L}_{\omega'}$ through a locally nonconstant isomorphism.

PROOF. Only the last statement requires an argument. For ω (resp. ω') the proof of Prop. 2.5 gives functions $\{v_\alpha\}_{\alpha\in I}$ (resp. $\{v'_\alpha\}_{\alpha\in I}$). The ratio $v_\alpha/v'_\alpha = w_\alpha$ expresses the local exactness of $\omega-\omega'$ as the differential of a branch of log of w_α (which has no zeros or poles). Then, the collection of maps $\mathcal{O}_{U_\alpha}\xrightarrow{\text{mult by} w_\alpha}\mathcal{O}_{U_\alpha}$ twines between $\mathcal{L}_\omega$ and $\mathcal{L}_{\omega'}$. □

REMARK 2.7 (Unitary bundle). Consider the compact Riemann surface X and ω as in Prop. 2.5. Periods for any paths around disks bounding the poles of ω are imaginary since locally the differential is mdz/z for some integer m. So, it makes sense to ask about differentials ω with polar divisor D for which $\int_\gamma\omega$ is pure imaginary. Suppose there is a holomorphic differential μ with $\omega-\mu=\omega'$ having all its periods pure imaginary. Then, $\mathcal{L}_{\omega'}$ corresponds to a representation of $\pi_1(X)$ into the circle group of absolute value 1 imaginary numbers (a *unitary representation*).

PROPOSITION 2.8. *Suppose ω is a local log differential on $X_{c,d}$ with divisor of poles not on γ_1 or γ_2. Then, there is a unique a_1 (resp. a_2) so $\int_{\gamma_1}\omega-a_1\omega_{c,d}=0$ (resp. $\int_{\gamma_i}\omega-a_2\omega_{c,d}$ is pure imaginary for $i=1,2$).*

So any degree 0 divisor D on $X_{c,d}$ gives a unique locally constant unitary bundle. Also, we can analytically continue a primitive W_D for $\omega-a_2\omega_{c,d}$ along any path avoiding the support of D. Then, $R_D(x;x_0)=\Re(W_D(x)-W_D(x_0))$ ($x_0,x\notin D$), a harmonic function of x, is independent of the path between x_0 and x.

PROOF. Multiply $\omega_{c,d}$ by a nonzero element of $\mathbb{C}$ to assume (as in Prop. 3.4) $s_1=2\pi i$ and $s_2=\tau'$ with $\Re(\tau')<0$. So, pick a_1 so $2\pi i a_1=\int_{\gamma_1}\omega$. Subtract $\Re(\int_{\gamma_1}\omega)\,\omega_{c,d}$ to go from ω with a 0 period along γ_1 to a differential with the same polar divisor and pure imaginary periods. □

REMARK 2.9 (Green's functions). The harmonic expression for existence of local log differents is the existance of Green's functions $R_D(x,x_0)$. Its uniqueness is from the maximum modulus principle on the compact surface, and it having local log behavior at the support of D.

2.4.2. *Universally constructing all local log differentials.* Legendre had written a whole book on the topic of differentials on $X_{c,d}$ (though not expressed in that language). To whit: One algebraically commands the nature of all meromorphic differentials on $X_{c,d}$ (from his perspective, of rational functions in z and $\sqrt{z^3+cz+d}$) from a primitive $f_{c,d}(z)$. For $X_{c,d}$, by the early 1800s, it was clear there are local log differentials with any degree 0 polar divisor.

The hypothesis of Prop. 2.10 that we can construct all log differentials on $\mathbb{C}_u$ from the desired $\sigma(u)$ gives the rubric for generalizing the case $\mathbb{P}^1_z$. Refer to this as $\sigma(u)$ has the *log-differential property*.

PROPOSITION 2.10. *Suppose $\sigma(u)$ is holomorphic on $\mathbb{C}_u$ with zeros of multiplicity one at each element of $L_{c,d}$. Assume also, for any n and $a_1,\dots,a_n,b_1,\dots,b_n\in$*

$\mathbb{C}_u$, translates by elements of $L_{c,d}$ leave invariant the log differential of

$$\varphi(u) = \frac{\prod_{i=1}^n \sigma(u-a_i)}{\prod_{i=1}^n \sigma(u-b_i)} \tag{2.1}$$

(2.2) *Then, $\sigma(u+s) = e^{k_s u + l_s}\sigma(u)$ for each $s \in L_{c,d}$: Exactly the log-differential property.*

For $\sigma(u)$ with these properties, every degree 0 divisor D on $X_{c,d}$ has a local log differential with D as polar divisor, and with $e^{au}\sigma(u)$, running over $a \in \mathbb{C}$, substituting for σ, log differentials of $\varphi(u)$ give all log differentials on $X_{c,d}$.

Translations from $L_{c,d}$ leave (2.1) *invariant (giving a meromorphic function on $X_{c,d}$) if and only if $\sum_{i=1}^n [a_i] - \sum_{i=1}^n [b_i] \in \mathbb{C}/L_{c,d}$ is 0. Equivalently, $\frac{d\varphi(u)}{du}/\varphi(u)\,du$ is a logarithmic differential on $X_{c,d}$.*

There exist $k, l \in \mathbb{C}$ with $e^{ku+l}\sigma(u)$ an odd function of u. It, too, has the log-differential property. This determines k, and it determines l modulo $\pi i\mathbb{Z}$.

Proof. Assume $\sigma(u)$ is holomorphic in $\mathbb{C}_u$, and it has a zero of multiplicity one at each period, $\sigma(u+s)/\sigma(u)$ has no zeros or poles. So, there is a well- defined branch $h_s(u)$ of log on all of $\mathbb{C}_u$. Conclude the ratio has the form $e^{h_s(u)}$. The difference between the log differential of $\varphi(u)$ and $\varphi(u+s)$ in (2.1) is $\sum_{i=1}^n \frac{dh_s(u-a_i)}{du} - \sum_{i=1}^n \frac{dh_s(u-b_i)}{du}$. Our hypothesis is that this is 0 for all $\boldsymbol{a}, \boldsymbol{b}$. So, $\frac{dh_s(u)}{du}$ is constant and $h_s(u) = k_s u + l_s$ for constants k_s and l_s. Clearly this produces a local log differential for any $[a_1], \dots, [a_n], [b_1], \dots, [b_n] \in \mathbb{C}_u/L_{c,d}$.

Now consider when the $d\varphi(u)/\varphi(u)$ is a logarithmic differential dv/v on $X_{c,d}$. From Abel's necessary condition, with no loss assume $\sum_{i=1}^n a_i - \sum_{i=1}^n b_i = 0$. Check translation of $s \in L_{c,d}$ leaves (2.1) invariant:

$$\varphi(u+s)/\varphi(u) = e^{k_s(\sum_{i=1}^n a_i - \sum_{i=1}^n b_i)} = 1,\ j = 1, 2.$$

Multiplication by -1 on $\mathbb{C}_u$ induces a corresponding automorphism on $\mathbb{C}_u/L_{c,d}$, which in turn induces an isomorphism on all meromorphic differentials. Therefore, $\sigma(-u)$ has the defining property of the paragraph above. As $\sigma(-u)$ has the same zeros (with multiplicity 1) as does $\sigma(u)$, conclude that $\sigma(-u) = e^{h(u)}\sigma(u)$ for some $h(u)$ holomorphic on $\mathbb{C}_u$. Taking the log differential of (2.1) with $\sigma(-u)$ replacing $\sigma(u)$, conclude from the above computation:

$$\sum_{i=1}^n \frac{dh(u-a_i)}{du} - \sum_{i=1}^n \frac{dh(u-b_i)}{du} \text{ is 0 for all } \boldsymbol{a}, \boldsymbol{b}.$$

Again, the conclusion is $h(u) = k'u + l'$ for some $k', l' \in \mathbb{C}$. It is automatic that $e^{k^*u+l^*}\sigma(u)$ has the log-differential property for any $k^*, l^* \in \mathbb{C}$. Iterate applying the automorphism induced by -1 on $\mathbb{C}_u$. Then, $\sigma(u) = e^{h(-u)}e^{h(u)}\sigma(u)$, and so $e^{-h(-u)} = e^{h(u)}$. Conclude: $2l' \in 2\pi i\mathbb{Z}$. Let $\sigma^*(u) = e^{k'u/2}\sigma(u)$. Compute:

$$\sigma^*(-u) = e^{-k'u/2}\sigma(-u) = e^{k'u + l' - k'u/2}\sigma(u) = e^{l'}\sigma^*(u).$$

Since σ^* has a zero of multiplicity one at 0, from local behavior around 0 conclude $e^{l'}$ is -1 (l' can be taken as πi). □

3. Implications from an odd σ with the log-differential property

The rubric of §2.3.2 goes far toward generalizing Abel's Theorem. We see that from natural puzzles likely occurring in Gauss and Riemann discussions (§3.2).

3.1. The unique odd θ with zero at the origin. In the last step, we actually find this odd $\sigma(u)$. When we do, we have what practioners often call the dimension one (or genus one) odd θ (it is a θ function, though tradition calls this Weierstrass version σ). Yet, there is unfinished business even then (§6.4).

PROPOSITION 3.1. *Up to multiplication by a constant, there is a unique odd homolomorphic function $\sigma_{c,d}(u)$ on $\mathbb{C}_u$ with a unique multiplicity one zero (modulo $L_{c,d}$) at $u = 0$, for which*

$$\sigma_{c,d}(u+s) = e^{k_s u + l_s}\sigma_{c,d}(u) = \sigma(u) \tag{3.1}$$

for some $k_s \in \mathbb{C}^*$ and $l_s \in \pi i\mathbb{Z}$, $s \in L_{c,d}$. The derivative in u of $\frac{d\sigma(u)}{du}$ is a translate of the even function $g_{c,d}(u)$.

COMMENTS ON FINDING σ. We know from Prop. 2.10, the log-differential property forces the conditions (3.1). Given (3.1), remove the *cocycle factor* $e^{k_s u + l_s}$ by forming the logarithmic derivative

$$\frac{d(\frac{d\sigma}{du})}{\sigma}(u+s) = k_s + \frac{d(\frac{d\sigma}{du})}{\sigma}(u). \tag{3.2}$$

One more derivative in u gives $g(u)$ invariant under translations by $L_{c,d}$.

Since $\sigma(u)$ has (modulo $L_{c,d}$) but one zero of multiplicity one (at 0), we know $g(u)$ has (modulo $L_{c,d}$) precisely one pole of multiplicy two (at 0). Such a function is $g_{c,d}$; $u = 0$ corresponds to the point on $X_{c,d}$ lying over $\infty \in \mathbb{P}^1_z$. That the sum of the residues is 0 determines $g_{c,d}$ up to a change $z \mapsto az+b$, with $a, b \in \mathbb{C}$. Inverting the process of taking the derivative of the log derivative gives σ. Consequently, from §1.4, k_s is $\int_{\gamma_s} g_{c,d}\,dz$. For the two generating periods s_i, $i = 1, 2$, it is traditional to use η_i for k_{s_i}, $i = 1, 2$. We have only to figure what is e^{l_s} to conclude the uniqueness of σ up to constant multiple. For that, the oddness of σ is crucial. For s_i, $\sigma(u - s_i)_{u=s_i/2} = -\sigma(s_i/2)$ and it also equals $e^{-\eta_i s_i/2 - l_i}\sigma(s_i/2)$. That determines e^{l_i}, $i = 1, 2$.

Apply (3.1) to $\sigma(u + s + s')$ to see $k_{s+s'} = k_s + k_{s'}$.

The main point left is to consider for $a_1, \dots, a_n, b_1, \dots, b_n \in \mathbb{C}_u$, what happens if $\frac{d\varphi(u)}{du}/\varphi(u)\,du = dv/v$ for some meromorphic function v on $X_{c,d}$. Pull v back to $\mathbb{C}_u$, and conclude from the differential equation that $me^{yu+n}v = \varphi(u)$ for some $m, y, n \in \mathbb{C}$. From this equation, v has the divisor $\sum_{i=1}^n [a_i] - \sum_{i=1}^n [b_i]$, and so by the above, $\varphi(u)$ defines a function on $X_{c,d}$. $\square$

Denote $L_{a,b} \setminus \{0\}$ by $L^*_{a,b}$:

$$\sigma_{c,d}(u) = \sigma(u) = u \prod_{s \in L^*_{a,b}} \left(1 - \frac{u}{s}\right) e^{u/s + \frac{1}{2}(u/s)^2}. \tag{3.3}$$

Clearly, $\sigma(u)$ is an odd function. Also, $\sigma(u+s_1) = -\sigma(u)e^{\eta_1(u+s_1/2)}$, etc. As in the proof Prop. 3.1, the oddness of $\sigma(u)$ gives a special role to the $u = s_i/2$, $i = 1, 2$. The construction of $\sigma = \sigma_{c,d}$ gives a holomorphic function (as in §1.3):

$$\sigma : ((c,d), u) \in V^0 \times \mathbb{C}_u \mapsto \sigma_{c,d}(u). \tag{3.4}$$

PROPOSITION 3.2. *We may craft all local log differentials (and so all meromorphic functions) on all complex 1-dimensional torii and all compact curves of form $X_{c,d}$ from (3.4).*

3.2. Puzzles from Abel's Theorem. Here are puzzles Riemann handled to describe all functions on a compact Riemann surface.

(3.5a) For a general $\varphi : \mathbb{C}/L_{c,d} \to \mathbb{P}^1_w$, there is no branch of log description of φ (Galois' discovery). So, how to picture such a cover?

(3.5b) How to relate the beginning and end points of allowable paths lifting γ from D^0 to D^∞ in Abel's Prop. 2.2 condition (from describing φ)?

(3.5c) For any odd n, how to describe equivalence classes V^0_n from corresponding (c, d) and (c', d') if there is a degree n analytic map $X_{c,d} \to X_{c',d'}$?

(3.5d) What conditions govern normalization of the function $\sigma_{c,d}$ in Prop. 2.10?

If n is composite, several different types of maps have degree n in (3.5c), though the question still takes a good shape. When $n = 1$, the phrasing would be to describe V^0_1, equivalence classes of the many pairs $(c, d) \in V^0$ corresponding to the same isomorphism class of complex torus.

The function $\sigma_{c,d}$ in Prop. 2.10 is actually Weierstrass' version of a θ function, not Riemann's. The likely Gauss-Riemann conversation helps considerably in figuring the different possibilities for normalization.

3.3. Normalizing the θ s. We have already put one condition into our normalization that we must relax to appreciate the long history of θ functions: We've kept the origin in $\mathbb{C}_u$ as a zero (of the function $\sigma_{c,d}(u)$).

3.3.1. *Even and odd thetas.* Any translation $u + u_0$ of the variable u also gives a θ function $\sigma_{c,d}(u + u_0)$ according to Prop. 2.10: Use the functional equation $\sigma(u + s) = e^{k_s u + l_s}\sigma(u)$ for each $s \in L_{c,d}$. Further, a necessary condition that $\sigma(u)$ be either even or odd is that multiplication by -1 preserves its zero set.

LEMMA 3.3. *Among the functions $e^{au}\sigma_{c,d}(u + u_0)$, those that are either odd or even correspond to values of u_0 for which $2u_0 \in L_{c,d}$. For these there is a unique $a = a_{u_0}$ for which it is odd or even. More precisely, $\sigma_{c,d}(u)$ is odd, and $e^{a_{u_0}}\sigma_{c,d}(u + u_0)$ with $u_0 = s_1 + s_2$ or $u_0 = s_i/2$, $i = 1, 2$, is even.*

PROOF. Except for the precise values for evenness and oddness, this is in Prop. 2.10. Since $\sigma_{c,d}(u)$ has exactly one zero (mod $L_{c,d}$), for u_0 a nontrivial 2-division point, $\sigma_{c,d}(u+u_0)$ is not zero, and so it is §3.1 for the precise changes. □

Riemann's choice of the fundamental θ function is that with $u_0 = s_1 + s_2$. The other three complete the list of *θ s with half- integer characteristics*. Denote any one of the functions $e^{au}\sigma(u+u_0)$ by $\theta(u)$. An equivalent condition to it satisfying (2.2) appears in the proof of Prop. 2.10. For each $s \in L_{c,d}$, $d\theta(u+s)/\theta(u+s)-d\theta(u)/\theta(u)$ is a constant k_s times du, the holomorphic differential on $\mathbb{C}_u/L_{c,d}$.

The following properties are explicit from $\sigma_{c,d}$. Still, they restate the general technique given in Prop. 2.8.

PROPOSITION 3.4 (A-period normalization). *Replace u by cu, $c \in \mathbb{C}^*$, so $s_1 = 2\pi i$. Denote $s_2 = \tau'$. With no loss $\Re(\tau') < 0$. So, with $\tau \stackrel{\text{def}}{=} \tau'/2\pi i$, $\Im(\tau) > 0$. Given any θ, multiplying it by e^{au} for some a produces a θ with $\theta(u + 2\pi i) = \theta(u)$. Such a θ has a Fourier series expansion in the variable u.*

3.3.2. *The Fourier expansion and θ-characteristics.* There is no aspect of θ functions more used than this Fourier expansion. Euler, Jacobi and many others used it (as explained and generalized in [**Sie29**, Chap. 1-2] and [**FaK01**, Chap. 3] titled *Function theory for modular group and its subgroups*). Once we see the significance of using τ' (or τ) as an essential parameter for the complex torii $\mathbb{C}_u/L_{c,d}$, a whole subject takes off. Yet, §6.3 shows there are problems applying this to natural equations not officially from the study of theta functions.

Suppose $\theta(u)$ is odd *and* has a Fourier expansion. Evaluate $\theta(-u+2\pi i) = -\theta(u)$ at $u = \pi i$ to conclude $\theta(\pi i) = 0$. This is contrary to our assumption of only one zero mod $L_{c,d}$. So, we can't get the Fourier expansion and oddness simultaneously. Having, however, the Fourier expansion was so valuable, Riemann followed Jacobi to choose an even function to express it [**Fay73**, p. 1]:

$$\theta(u,\tau_{c,d}) = \theta\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}(u,\tau_{c,d}) = \sum_{m\in\mathbb{Z}} e^{m\pi i\tau_{c,d}m+mu} = \sum_{m\in\mathbb{Z}} e^{\pi i\tau_{c,d}m^2+mu}. \tag{3.6}$$

§3.3.3 suggests why we expect such an expression for all $\tau_{c,d}$. Of course, $\tau_{c,d}$ is an entirely mysterious function of (c,d), almost the whole point of Abel's investigations. Further, it depends on a choice: The basis of $L_{c,d}$. The other three even and odd thetas get a similar look from this. Following Fay, use ϵ and δ to indicate real multiples of the periods: Each point of $\mathbb{C}_u$ has the form $\boldsymbol{e} = 2\pi i(\epsilon+\delta\tau_{c,d})$. For an integer n, the n-division points $\mathbb{C}_u/L_{c,d}$ have representatives by taking $\delta,\epsilon \in \frac{1}{n}\mathbb{Z}$. The theta with characteristics (δ,ϵ):

$$\begin{aligned}\theta\begin{bmatrix}\delta\\ \epsilon\end{bmatrix}(u,\tau_{c,d}) = {}& e^{\delta\pi i\tau_{c,d}\delta+(u+2\pi i\epsilon)\delta}\,\theta\begin{bmatrix}\mathbf{0}\\ \mathbf{0}\end{bmatrix}(u+\boldsymbol{e},\tau_{c,d})\\ &= \textstyle\sum_{m\in\mathbb{Z}} e^{(m+\delta)\pi i\tau_{c,d}(m+\delta)+(u+2\pi i\epsilon)(m+\delta))}.\end{aligned} \tag{3.7}$$

LEMMA 3.5. *With $\epsilon,\delta \in \frac{1}{2}\mathbb{Z}$, (3.7) gives even and odd functions. The one odd function goes with $\epsilon = \delta = \frac{1}{2}$.*

For $(\delta,\epsilon) \in \frac{1}{k}\mathbb{Z}$, the θs have $\frac{1}{k}$-*characteristics*. The notational difference between [**Fay73**] and [**FaK01**] appears in changing e^u in the former to $e^{2\pi iu}$ in the latter. This has a slight effect on the expression for the *heat equation* given by comparing the effect of $\frac{\partial^2}{\partial u^2}$ with $\frac{\partial}{\partial\tau}$ on (3.6). As theta function topics go, this equation is reasonably memorable.

(3.8a) (3.6) is invariant under $2\pi i\tau \mapsto 2\pi i\tau + 2\cdot 2\pi i$: A Fourier series in $2\cdot\tau$.

(3.8b) The odd θ, $\theta\begin{bmatrix}\frac{1}{2}\\ \frac{1}{2}\end{bmatrix}(u,\tau_{c,d})$ is invariant under $z \mapsto z+2\cdot 2\pi i$, close to a Fourier series in z.

(3.8c) (3.6) makes sense in g complex variables: $u \mapsto \boldsymbol{u} \in \mathbb{C}_{(u_1,\dots,u_g)} \stackrel{\text{def}}{=} \mathbb{C}_{\boldsymbol{u}}$; replace τ by any symmetric $g\times g$ matrix Π.

(3.8d) For $k \in \mathbb{N}^+$, the collection $\{\theta\begin{bmatrix}\frac{m}{k}\\ \frac{m'}{k}\end{bmatrix}(u,\tau)\}_{\{ {m,m'\in\mathbb{Z}| \atop (m,k)=(m',k)=1} \}}$ evaluated at 0 (or use derivatives in u) are analytic functions on τ-space.

Reality check on (3.8c): Convergence in z for a particular Π requires its real part be negative definite. Also, replace $(m+\delta)\pi i\tau_{c,d}(m+\delta)$ by $(\boldsymbol{m}+\boldsymbol{\delta})\pi i\Pi(\boldsymbol{m}+\boldsymbol{\delta})^{\mathrm{t}}$ with $\boldsymbol{m} \in \mathbb{Z}^g$ and $\boldsymbol{\delta} \in \mathbb{R}^g$.

3.3.3. *The jump to τ-space.* Consider why we expect in (3.6) a function with a natural analytic continuation in $\tau = \tau_{c,d}$. Given a set of branch points $\boldsymbol{z}$, and a choice of γ_1,γ_2 (as in §1.4), you may fix γ_1,γ_2 even as you wiggle the $\boldsymbol{z}$ a little, and therefore continue $\tau = \tau_{\boldsymbol{z}}$ uniquely. Further, in a continuous wiggle of $\boldsymbol{z}$, you uniquely determine $\frac{2\pi i+\tau'}{2}$ the place of the zero mod L of the θ function, among the discrete set of points of order 2 on $\mathbb{C}/L_{\boldsymbol{z}}$. So, now we know $\theta(u,\tau_{\boldsymbol{z}})$ uniquely from $\mathbb{C}/L_{\boldsymbol{z}}$, under this small wiggle hypothesis, up to multiplication by a (nonzero) constant. With no loss assume for each $\tau_{\boldsymbol{z}}$ the constant coefficient of the Fourier series expansion in e^u is 1. Then, this function is extensible over the whole τ plane.

That (3.6) is a theta function is now an elementary check. Example: $u \mapsto 2\pi i\tau$ has the effect $e^{\pi i\tau m^2+mu} \mapsto e^{-u-\pi i\tau}e^{\pi i\tau(m+1)^2+(m+1)u}$ for each m.

Suppose $\Phi : \mathcal{T} \to \mathcal{H}$ is a connected family of genus 1 curves. A section $\mu : \mathcal{H} \to \mathcal{T}$ to Φ allows us to regard Φ as a family of complex torii: The point $\mu(\boldsymbol{p})$ gives a canonical isomorphism: $x \in \mathcal{T}_{\boldsymbol{p}} \mapsto x - \mu(\boldsymbol{p}) \in \mathrm{Pic}^{(0)}_{\boldsymbol{p}}$ from $\mathcal{T}_{\boldsymbol{p}}$ to divisor classes of degree 0. We assume known that $\mathrm{Pic}^{(0)}_{\boldsymbol{p}}$ has a structure of complex torus through a holomorphic differential on $\mathcal{T}_{\boldsymbol{p}}$ (§6.3).

Subtly different is considering on each fiber $\mathcal{T}_{\boldsymbol{p}}$ the set of divisor classes of degree 0 whose squares are the trivial divisor class. Denote this $\mathcal{D}^0_{\boldsymbol{p}}$. The subtlety is that we won't easily recognize such divisor classes without an explicit torus structure. Unless there are special conditions, there won't usually be specific divisors representating them, varying analytically with $\boldsymbol{p}$. This is a $\frac{1}{2}$-*canonical divisor class*.

The definition works for general genus. In (4.7), $\sum_{i=1}^r \mathrm{ind}(\mathrm{C}_i)-2n$ is the degree of the differential $d\varphi$ of $\varphi : X \to \mathbb{P}^1_z$. This defines $g(X)$ (independent of φ) because all differentials on X have the same degree. So $\frac{1}{2}$-canonical divisor classes have degree $g(X) - 1$.

DEFINITION 3.6. Suppose $(d\varphi) = 2D_\varphi$ for some divisor D_φ. Call D_φ an exact $\frac{1}{2}$-canonical divisor.

We use a technical lemma later [**BFr02**, §B.1].

LEMMA 3.7. *Such a D_φ exists if and only if all elements of* $\mathbf{C}$ *have odd order. For any $\alpha \in \mathrm{PGL}_2(\mathbb{C})$, $D_{\alpha\circ\varphi}$ is linearly equivalent to D_φ.*

THE EXPLICIT LINEAR EQUIVALENCE. With no loss represent α by $\frac{az+b}{cz+d}$ with $ad - bc = 1$. Then,

$$d\alpha \circ \varphi = \frac{d\alpha(w)}{dw}_{|w=\varphi(z)} d\varphi(z) = \frac{1}{(c\varphi(z)+d)^2} d\varphi(z).$$

So $D_{\alpha\circ\varphi} = D_\varphi - (c\varphi(z) + d)$. □

It would be wonderful if one could summarize Riemann's work by saying (3.8c) works for general compact Riemann surfaces, and it's a great memory device, so done. It isn't quite that easy.

Still, we profit by turning to questions that arose from *cuts* have a modern cast and many modern applications. The unique odd theta function when the genus is 1 is in (3.4): One function works simultaneously to describe all functions on all compact Riemann surfaces appearing as complex torii. That is very useful, if solving a problem about functions depends only on zeros and poles, rather than on the branch points and if the coordinates given by the variable τ are appropriate for the problem. We can phrase why that isn't so by looking at (3.5a) and analyzing the signficance of finding out that all analytic maps φ don't come from (3.5c). The conversations between Gauss and Riemann, on whose outcome the rest of this paper concentrates, are appropriate for seeing this even through the case $g = 1$.

Abel's Theorem is perfect for forming abelian covers of a complex torus similar to using branches of log to describe abelian covers of $\mathbb{P}^1_u$. Knowing only the generators of a function field like $M_{c,d}$ gives surprisingly little help in understanding properties of elements in the field. Abel's Theorem is valuable for many questions, and yet it too leaves us powerless against those in Ex. 6.5.

4. Compact surfaces from cuts and the puzzle (3.5a)

To this point we interpreted the compact Riemann surface $X_{c,d}$ as coming from a branch of $\sqrt{z^3+cz+d}$. The value was that the differential $h_{c,d}(z)\,dz$ interprets as a holomorphic differential on $X_{c,d}$. Its integrals around all closed paths on $X_{c,d}$ produce a lattice $L_{c,d}$. The goal (as in (3.5c)) is to glean how this integral changes as (c,d) varies. The phrasing of that goal was how to put order in relating (c,d) and (c',d') when $g_{c',d'} \in M_{c,d}$. The production, however, of all functions on the compact Riemann surface $X_{c,d}$ raises the question of how objects on $X_{c,d}$ appear from the view of $\mathbb{P}^1_w$ given by $\varphi : X_{c,d} \to \mathbb{P}^1_w$.

For example, is a general φ really given by one of those $g_{c',d'}$ s? If not, how would holomorphic (and meromorphic) differentials appear as a function of w? Our discussion documents that the negative answer to the former and how to consider the latter must have occurred in discussions between Gauss and Riemann.

4.1. Data for cuts and Nielsen classes. The treatment of [**Fr03**, Chap. 4, §2.4] emphasizes unramified covers as locally constant structures (and has complete details). We here go for the simplier goal of clarifying one look at Riemann's Existence Theorem. Here is the starting data for connected degree n covers.

(4.1a) $r+1$ distinct points z_0 and $\boldsymbol{z} = \{z_1, \dots, z_r\}$ on the sphere.
(4.1b) Semi-simplicial paths $\bar{\gamma}_i$, (with range) from z_0 to z_i, $i = 1, \dots, r$, meeting only at their beginning points, that emanate clockwise from z_0.
(4.1c) a collection of elements $g_1, \dots, g_r \in S_n$ satisfying two conditions:
- Generation: The group $G(\boldsymbol{g})$ they generate is transitive.
- Product-one: The product $g_1 \cdots g_r \overset{\text{def}}{=} \Pi(\boldsymbol{g})$ (in that order) is 1.

4.1.1. *Covers from cut data.* We show how the data $C = C(\bar{\boldsymbol{\gamma}}, \boldsymbol{g})$ (4.1) canonically produces a new compact Riemann surface cover $\varphi_C : X_C \to \mathbb{P}^1_z$. Equivalences between two such covers are important, though we suppress that here. Given that there are cuts as in (4.1b), call the elements $\boldsymbol{g}$ from (4.1c) a *branch-cycle description*.

Let $\mathbb{P}^1_i$, $i = 1, \dots, r$, be copies of $\mathbb{P}^1_z$, and on each remove the points labeled $z_0, z_1, \dots, z_r$. Call the result $\mathbb{P}_j$. Form a pre-manifold $\mathbb{P}^{\pm}_j$ (not Hausdorff) from $\mathbb{P}_j$ by replacing each point z along any one of the γ_i s by two points: z^+ and z^-. We form a manifold from an equivalence on the union of the $\mathbb{P}^{\pm}_j$, $j = 1, \dots, n$, using the expected neighborhoods of all points except z^+ and z^-. For neighborhoods of these, we use the following sets. Let $D_{i,z}$ be a disk around z. Write this as as a union of two sets: $D^+_{i,z}$ (resp. $D^-_{i,z}$), all points on and to the left (resp. right) of γ_i.

PROPOSITION 4.1. *Form a manifold from an equivalence relation (in the proof) on $\cup_{i=1}^n \mathbb{P}^{\pm}_j$ based on using the r-tuple $\boldsymbol{g}$. Running over all n and product-one r-tuples $\boldsymbol{g}$ (even with the cuts fixed), forming the compactification gives all possible compact Riemann surfaces mapping to $\mathbb{P}^1_z$ ramified over $\boldsymbol{z}$.*

PROOF. If g_i maps k to l, then identify $z^- \in \mathbb{P}^{\pm}_k$ in the g_i cut with $z^+ \in \mathbb{P}^{\pm}_l$. In the resulting set, put on a topology where the neighborhood of such a z^- is $D^+_{l,z} \cup D^-_{k,z}$ identified along the part of γ_i running through z. □

4.1.2. *Changing the cuts and s-equivalences.* There are many ways to consider the cuts changing. Eventually one must consider that in great generality. The simplest change, however is just to change the $\bar{\gamma}_1, \dots, \bar{\gamma}_r$, noting this canonically changes the $\boldsymbol{g}$ if we keep the analytic map $\varphi_C : X_C \to \mathbb{P}^1_z$ fixed [**Fr03**, §2.4.3].

LEMMA 4.2. *Suppose we fix $\boldsymbol{z}$, but change the cuts (leaving z_0 fixed) and possibly changing the order of $(z_1, \dots, z_r)$ to correspond to condition* (4.1b)*, the cuts emanate clockwise from z_0. Then, the corresponding cover $\varphi_C : X_C \to \mathbb{P}^1_z$ has a new (canonical) branch cycle description $\boldsymbol{g}' = (g'_1, \dots, g'_r)$ satisfying these conditions.*

(4.2a) $G(\boldsymbol{g}) = G(\boldsymbol{g}')$ *(and $\Pi(\boldsymbol{g}') = 1$).*
(4.2b) *For $\pi \in S_r$ corresponding to the change in order of $\boldsymbol{z}$ ($z_i \mapsto z_{(i)\pi}$, $i = 1, \dots, r$), g'_i is conjugate in $G(\boldsymbol{g})$ to $g_{(i)\pi}$.*

Record the set of conjugacy classes in $G = G(\boldsymbol{g})$, repeating them with multiplicity though without regard to order, as $\mathbf{C} = \mathbf{C}_{\boldsymbol{g}}$. We say $\boldsymbol{g}, \boldsymbol{g}' \in \mathbf{C}$.

DEFINITION 4.3. We call the collection $\mathrm{Ni}(G, \mathbf{C})$ of $\boldsymbol{g}'$ satisfying (4.2) the *Nielsen class* of $(G, \mathbf{C})$.

There are many possible further equivalences on Nielsen classes. All come from modding out by the action of certain groups on $\mathrm{Ni}(G, \mathbf{C})$. Describing the cuts starts with a permutation representation of $G(\boldsymbol{g})$. Let $N_{S_n}(\mathbf{C})$ denote the subgroup of S_n that normalizes G and permutes the conjugacy classes $\mathbf{C}$. It makes sense to conjugate any r-tuple $\boldsymbol{g} \in \mathrm{Ni}(G, \mathbf{C})$ by elements of $N_{S_n}(\mathbf{C})$. Two equivalences apply for all values of r:

(4.3a) Inner Nielsen classes: This is the set $\mathrm{Ni}(G, \mathbf{C})/G \stackrel{\mathrm{def}}{=} \mathrm{Ni}(G, \mathbf{C})^{\mathrm{in}}$.
(4.3b) Absolute Nielsen classes: $\mathrm{Ni}(G, \mathbf{C})/N_{S_n}(\mathbf{C}) \stackrel{\mathrm{def}}{=} \mathrm{Ni}(G, \mathbf{C})^{\mathrm{abs}}$ with an understanding this requires giving a permutation representation.

We say an element $\boldsymbol{g} \in \mathrm{Ni}(G, \mathbf{C})$ represents the Nielsen class. It also represents an (absolute or inner) s-equivalence class, so giving an element of $\mathrm{Ni}^{\mathrm{abs}}$ or $\mathrm{Ni}^{\mathrm{in}}$.

A natural set of operators, which we designate as $q_1, \dots, q_{r-1}$, acts on any of the s-equivalence classes of a Nielsen class. For $\boldsymbol{g} \in \mathrm{Ni}^{\mathrm{in}}$ (or $\mathrm{Ni}^{\mathrm{abs}}$), q_i sends $(g_1, \dots, g_r) = \boldsymbol{g}$ (in order) to the new r-tuple of $G(\boldsymbol{g})$ generators

$$(g_1, \dots, g_{i-1}, g_i g_{i+1} g_i^{-1}, g_i, g_{i+2}, \dots, g_r),\ i = 1, \dots, r-1. \tag{4.4}$$

There is an actual group, the *Hurwitz monodromy group* H_r that these operators generate. It has a presentation by generators and relations. We list these.

(4.5a) Braid relations: $q_i q_j = q_j q_i$, $1 \le i \le j \le r-1$; $j \ne i-1$ or $i+1$, and $q_i q_{i+1} q_i = q_{i+1} q_i q_{i+1}$, $i = 1, \dots, r-2$.
(4.5b) Hurwitz relation: $q(r) = q_1 q_2 \cdots q_{r-1} q_{r-1} \cdots q_2 q_1$.

Let $U_r = \mathbb{P}^r \setminus D_r$ be the space of r distinct unordered points in $\mathbb{P}^1$, the image of $(\mathbb{P}^1)^r \setminus \Delta_r = U^r$. Thus, $\Psi_r : U^r \to U_r$ is an unramified Galois cover with group S_r. [**BFr02**, §2.1] reminds how giving a set of cuts identifies H_r with $\pi_1(U_r, \boldsymbol{z})$.

It has great value to be precise about what happens as we change the cuts. While this was not apparantly explicit in the discussions of Gauss and Riemann, it will be a tool for our discussion.

PROPOSITION 4.4. *The Hurwitz relation $q(r)$ acts as $\boldsymbol{g} \mapsto g_1(\boldsymbol{g}) q_1^{-1}$ on Nielsen class elements, so it acts on inner (or absolute) Nielsen classes* [**Fr03**, §3.1.2]*. The action of H_r commutes with the action of $N_{S_n}(\mathbf{C})$.*

Keep the hypotheses on the cuts as in Lem. 4.2 (like fix $\boldsymbol{z}$, etc.). Then, the action of H_r on s-equivalence classes gives all the possible changes in the branch cycles from changing the cuts.

RELATING TO THE FUNDAMENTAL GROUP OF U_r. . Suppose $\gamma : [0,1] \to U_r$ is any closed path with beginning and endpoint $\boldsymbol{z} \in U_r$. Since $\Psi_r : U^r \to U_r$ is unramified, any such path lifts uniquely to $\gamma^* : [0,1] \to U^r$ starting at $(z_1, \dots, z_r)$ and ending at some $(z_{(1)\pi}, \dots, z_{(r)\pi})$. So, each coordinate γ_i^*, gives a path on $\mathbb{P}^1_z$. Then, given $\bar\gamma_1, \dots, \bar\gamma_r$ giving cuts, we may form a new r-tuple of paths $(\bar\gamma_1 \cdot \gamma_1^*, \dots, \bar\gamma_r \cdot \gamma_r^*)$. These can serve as the basis for a new set of cuts, so long as the hypotheses for cuts hold. [**Fr03**, Chap. 5] (or [**Fri77**, §4]) lists γ s by which we calculate explicitly the effect of the $q_1, \dots, q_{r-1}$. It shows for allowable such paths, the result only depends on the homotopy class of γ, and that the effect of the $q_1, \dots, q_{r-1}$ generate all such changes. The proof comes to computing representatives of $\pi_1(U_r, \boldsymbol{z})$ modulo providing such representatives satisfying a few constraints. □

4.1.3. *Cuts and r-equivalences.* There are no new equivalences for $r \ge 5$. We tend to ignore $r = 2$, whenever we can, and the other equivalences apply to $r = 3$ and $r = 4$. For $r = 4$ these come from the action of a group $\mathcal{Q}''$ that acts *through* a Klein 4-group $(\mathbb{Z}/2 \times \mathbb{Z}/2)$ on any Nielsen classes. First consider the shift $\mathbf{sh} = q_1q_2q_3$: $(g_1, \dots, g_4) \in \mathrm{Ni}(G, \mathbf{C}) \mapsto (g_2, g_3, g_4, g_1)$. Then, consider $q_1 q_3^{-1}$ which acts as $(g_1, \dots, g_4) \in \mathrm{Ni}(G, \mathbf{C}) \mapsto (g_1 g_2 g_1^{-1}, g_2, g_3, g_3^{-1} g_4 g_3^{-1})$.

Two new equivalences for $r = 4$ [**BFr02**, §2]:

(4.6a) Reduced inner Nielsen classes: $\mathrm{Ni}(G, \mathbf{C})/\langle G, \mathcal{Q}''\rangle \stackrel{\text{def}}{=} \mathrm{Ni}(G, \mathbf{C})^{\text{in,rd}}$.

(4.6b) Reduced absolute Nielsen classes: $\mathrm{Ni}(G, \mathbf{C})/N_{S_n}(\mathbf{C}) \stackrel{\text{def}}{=} \mathrm{Ni}(G, \mathbf{C})^{\text{abs,rd}}$.

Reduced equivalence classes for $r = 3$ replace the group $\mathcal{Q}''$, by the group $\langle q_1q_2, q_1\rangle$. This comes from setting $q_3 = 1$ in $\mathcal{Q}''$, though that is misleading, for this group acts through S_3, and is not a quotient of $\mathcal{Q}''$. The operators $q_1, \dots, q_{r-1}$ are specific to a particular value of r. We simplify in dropping that notation, though we must be careful for there is no natural homomorphism from H_r to H_{r-1}.

4.2. Source of the cuts and modular curves. B. Riemann (1826–1866) from his thesis 1851 and his 1857 articles on abelian functions, used the Cauchy-Riemann equations exclusively. He based many of his proofs on potential theory.

4.2.1. *Where the cuts came from and a map through the rest of the presentation.* [**Ne81**, p. 89]: It was Gauss' (1777–1855) writings the young Riemann studied wth special zeal. From these he drew significant inspirations for his [doctoral] thesis. He wrote his father how he found these papers. What he especially appreciated was Gauss' contributions to conformal mapping using essentially a Dirichlet principle.

According to Betti, Riemann said he got the idea of *cuts* from conversations with Gauss [**Ne81**, p. 90]. Letters of Klein and Schering attest to Gauss' influence on Riemann's theory of hypergeometric series. Though this influence came partly from Gauss' papers, it is striking to consider, possibly in 1849, the over 70 year old Gauss sketching plans for such an etherial construction to the very young Riemann.

We show how modular curves arise from the cuts. Then, we give examples parallel to modular curves that show what Gauss and Riemann might have been considering. The group generated by the product-one $\boldsymbol{g}$ is the (monodromy) group $G(\boldsymbol{g})$ of the Galois closure of the cut map $\varphi : X \to \mathbb{P}^1_z$. §4.2.3 uses this group to characterize the answer to (3.5c). Euler's classification of compact 2-dimensional manifolds raises the question of finding which complex manifolds have given presentations (§6.3.1).

4.2.2. *Appearance of the j-line.* Linear fractional transformations, $z \mapsto \frac{az+b}{cz+d}$ with $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}_2(\mathbb{C})$ form a group acting on $\mathbb{P}^1_z$. Identify the elements of projective r-space, $\mathbb{P}^r$ as nonzero monic polynomials in a variable z of degree *at most* r. For example, if $(a_0, a_1, \dots, a_r)$ represents a point of $\mathbb{P}^r$, and $z_0 \neq 0$, by scaling it by $\frac{1}{z_0}$ assume with no loss $z_0 = 1$. Then, take the polynomial associated to this point as $z^r + \sum_{i=0}^{r-1}(-i)^{r-i}a_{r-i}z^i$. There is a natural permutation action of $\pi \in S_r$ on the entries of $(\mathbb{P}^1_z)^r$: $\pi : (z_1, \dots, z_r) \mapsto (z_{((1)\pi}, \dots, z_{(r)\pi})$.

If $\varphi : X \to \mathbb{P}^1_z$ represents an s-equivalence class of covers in a given Nielsen class Ni, then the collection $\{\alpha \circ \varphi : X \to \mathbb{P}^1_z\}_{\alpha \in \mathrm{PGL}_2(\mathbb{C})}$ gives the set of covers r-equivalent to φ. The cover $\alpha \circ \varphi$ has as branch points, α applied to the branch points of φ. Denote the equivalence classes for the action of $\mathrm{PGL}_2(\mathbb{C})$ on U_r (resp. U^r; §4.1.2) by J_r (resp. Λ_r).

As a quotient of an affine space by a reductive group, J_r is also an affine space. Then, Λ_4 identifies with $\mathbb{P}^1_\lambda \setminus \{0, 1\infty\} = U_{0,1,\infty}$:

$$(z_1, z_2, z_3, z_4) \mapsto \lambda_{\boldsymbol{z}} = \frac{(z_2 - z_3)}{(z_2 - z_1)} \frac{(z_4 - z_1)}{(z_4 - z_3)} \in \Lambda_4,$$

where z_4 goes under the fractional transformation taking (z_1, z_2, z_3) to $(0, 1, \infty)$. Notice: Cycling the order to (z_2, z_3, z_4, z_1) has the effect $\lambda_{\boldsymbol{z}} \mapsto 1/\lambda_{\boldsymbol{z}}$, so the square of the cycle fixes $\lambda_{\boldsymbol{z}}$. This is true for all conjugates in S_4 of the cycle. So, S_4 acts through $S_3 = S_4/K_4$, with K_4 a Klein 4-group on the collection of (z_1, z_2, z_3, z_4) over a given $\{z_1, z_2, z_3, z_4\}$. Let C_2 and C_3 be the respective conjugacy classes of order 2 and 3 in S_3. Let $F : \mathbb{P}^1_\lambda \to \mathbb{P}^1_z$ by the Galois cover with group S_3 ramified at $0, 1, \infty$ (we normalize F to use 1 rather than the number theorists' 1728) with conjugacy classes $\mathbf{C}_\lambda = (C_3, C_2, C_2)$ in order, and having $\{0, 1, \infty\}$ lying over ∞. This identifies J_4 with [**Ah79**, p. 282]:

$$\{z_1, \dots, z_4\} \mapsto j_{\boldsymbol{z}} = F(\lambda_{\boldsymbol{z}}) \in J_4.$$

It is a worthy exercise — to those new to Nielsen classes — to show $\mathrm{Ni}(S_3, \mathbf{C}_\lambda)^{\mathrm{in}}$ has but three elements. So, with the normalizations above this determines F.

4.2.3. *Modular curves.* We first consider how to describe the situation of (3.5c) as coming from cuts. Covers from which we may recover everything in (3.5c) come from covers $\varphi : X \to \mathbb{P}^1_w$ with $r = 4$ and monodromy group $G(\boldsymbol{g})$ a dihedral group when the elements of $\boldsymbol{g}$ are involutions. These are *dihedral involution* covers. Following that, we ask if $X_{c,d}$ is a cover of $\mathbb{P}^1_w$ in an entirely different way than given by Abel's situation (3.5c).

The field $\mathbb{C}(g_{c,d}(u)) = R_{c,d}$ identifies as the fixed field of even functions in $\mathbb{C}(g_{c,d}(u), \frac{dg_{c,d}}{du}) = \hat{R}_{c,d}$ of the automorphism induced by $u \mapsto -u$. The inclusion of $g_{c',d'}$ in $\mathbb{C}(g_{c,d}(u), \frac{dg_{c,d}}{du})$ induces a field extension $\mathbb{C}(g_{c,d})/\mathbb{C}(g_{c',d'}) = R_{c,d}/R_{c',d'}$. Let H_2 be the group $\{\pm 1\}$. Below, if A is an abelian group we regard H_2 acting on it by multiplications by ± 1. For a positive integer m, we consider the semidirect product action $\mathbb{Z}/m \times \mathbb{Z}/m \times^s H_2$ and various of its subgroups. If V is a subgroup of $\mathbb{Z}/m \times \mathbb{Z}/m$, then there is an induced group $\mathbb{Z}/m \times \mathbb{Z}/m/V \times^s H_2$. If the quotient $\mathbb{Z}/m \times \mathbb{Z}/m/V$ is cyclic of order n, then $\mathbb{Z}/m \times \mathbb{Z}/m/V \times^s H_2$ is the dihedral group D_n (having order $2n$; of degree n).

PROPOSITION 4.5. *An extension $R_{c,d}/R_{c',d'}$ (degree m) corresponds to a cover*

$$\varphi = \varphi_{(c,d),(c',d')} : \mathbb{P}^1_w \to \mathbb{P}^1_z : \text{ for } u \in \mathbb{C}_u,\ w = g_{c,d}(u) \mapsto z = g_{c',d'}(u).$$

The Galois closure of this cover corresponds to the extension $\hat{R}_{c,d} \to R_{c'd'}$. Its group G_φ is a quotient of $\mathbb{Z}/m \times \mathbb{Z}/m \times^s H_2$. If G_φ is D_n with n odd, then there is a branch cycle description for φ in the Nielsen class $\mathrm{Ni}(D_n, \mathbf{C}_{2^4})^{\mathrm{abs}}$ *and conversely.*

The collection of absolute reduced (resp. inner reduced) equivalence classes of covers in the Nielsen class $\mathrm{Ni}(D_n, \mathbf{C}_{2^4})$ *identifies with the open subset of the modular curve $X_0(n)$ (resp. $X_1(n)$) over $\mathbb{P}^1_j \setminus \{\infty\}$.*

CONSTRUCTING MODULAR CURVE EQUATIONS. Assume m is odd to simplify. On any complex torus, $X_{c',d'}$, we may multiply by an integer m. Denote this map m^*. The Galois closure $X_{c,d} \to \mathbb{P}^1_{g_{c',d'}}$ factors through $X_{c',d'}$, also degree m. All such covers fit between $m^* : X_{c',d'} \to X_{c',d'}$, which is Galois with group $\mathbb{Z}/m \times /bZ/m$, so identifying $m^* : X_{c',d'} \to \mathbb{P}^1_{g_{c',d'}}$ with $\mathbb{Z}/m \times \mathbb{Z}/m \times^s H_2$, and the group of $X_{c,d} \to \mathbb{P}^1_{g_{c',d'}}$ with a quotient by the subgroup of the Galois cover $X_{c',d'} \to X_{c,d}$. For m odd, the generator of H_2 gives a unique conjugacy class of involutions in $\mathbb{Z}/m \times \mathbb{Z}/m \times^s H_2$. This is the only conjugacy class fixing points on $X_{c',d'}$. The fixed points of H_2 are exactly the 2-division points on $X_{c,d}$.

So, $\varphi = \varphi_{(c,d),(c',d')} : \mathbb{P}^1_w \to \mathbb{P}^1_z$ has precisely four branch points $z_1, \ldots, z_4$, and above each z_i precisely one point w_i does not ramify, and the cover is in the Nielsen class $\mathrm{Ni}(D_n, \mathbf{C}_{2^4})$. In the other direction, suppose given $j_{\boldsymbol{z}}$ and a cover $\varphi : X \to \mathbb{P}^1_z$ with branch points $\boldsymbol{z}$ and in the Nielsen class $\mathrm{Ni}(D_n, \mathbf{C}_{2^4})$. This relates $j_{\boldsymbol{w}}$ and $j_{\boldsymbol{z}}$ (notation from §4.2.2).

We can see $j_{\boldsymbol{w}}$ is an algebraic function of $j_{\boldsymbol{z}}$ by using analytic continuation. Avoid $j \in \{0, 1, \infty\}$. For any particular $\boldsymbol{z}$ lying over $j_{\boldsymbol{z}}$, consider all the covers D_n covers of $\mathbb{P}^1_{g_{c',d'}}$ that factor through $X_{c',d'} \xrightarrow{n^*} X_{c',d'} \to \mathbb{P}^1_{g_{c',d'}}$. The case with n general comes by taking fiber products from the case n is prime-power (as in §5.1). So, if $n = p^e$ (p odd), there are as many such D_{p^e} covers as there $\mathbb{Z}/p^e$ quotients of $\mathbb{Z}/p^e \times \mathbb{Z}/p^e$. A **sh**-incidence argument in §5.2.4 illustrates a quick way to see this count is correct for the Nielsen classes for $\mathrm{Ni}(D_n, \mathbf{C}_{2^4})$. This assures there is no cover $\varphi : X \to \mathbb{P}^1_z$ in the Nielsen class not included (up to equivalence) by Abel's considerations.

A different argument brings up a key point for the Gauss-Riemann discussion. Elements h in a conjugacy class C of S_n have an index ind(C) given by n minus the number of orbits of h. The genus $g_{\mathrm{Ni}} = g(X)$ of X from a cover φ in a Nielsen class Ni come from the Riemann-Hurwitz formula:

$$2(n + g_{\mathrm{Ni}} - 1) = \sum_{i=1}^{r} \mathrm{ind}(\mathrm{C}_i). \tag{4.7}$$

(Fig. 1 shows how cuts produce a triangulation from which we compute the genus of the covering curve as an Euler characteristic.) So, without the count argument we need that $\varphi : X \to \mathbb{P}^1_z$ in $\mathrm{Ni}(D_n, \mathbf{C}_{2^4})$, with X having genus 0, identifies with $\mathbb{P}^1_w$. This then associates values $w_1, \ldots, w_4$ to $z_1, \ldots, z_4$. □

The proof of Prop. 4.5 required knowing a cover $X \to \mathbb{P}^1_z$ with $g(X) = 0$ assures X is analytically isomorphic to $\mathbb{P}^1_w$. It used the j-invariant attached to four points on $\mathbb{P}^1_w$ to give coordinates to modular curves. We have a respectable argument for that built from the cuts. Still, we are about to come upon genus 1 surfaces where it is difficult to find such an argument. The Riemann-Roch Theorem shows all genus 0 Riemann surfaces are analytically isomorphic, though we are relating this to a discussion before that theorem. If we knew this about X, then we could construct the degree two genus 1 cover from the $\boldsymbol{w}$ set, and from §1.3 to have a complex torus on it. Conversely, given a genus 1 compact surface Y, a nontrivial holomorphic differential ω_X identifies the universal covering space of X with $\mathbb{C}_u$ and gives Abel's situation.

5. Modular curve generalizations

Let any finite group H act on any lattice L or on any finitely generated free group F. We *include* the case L or F is trivial. We may replace $\mathbb{Z}$ by L or F and H_2 by H in the discussion of modular curves. Also, let $\mathbf{C}$ be some generating conjugacy classes for H. In the next discussion avoid all primes p dividing the order of any element in $\mathbf{C}$. For the most serious results we add that the finite quotient groups are *p-perfect*: Have no $\mathbb{Z}/p$ quotient. This is the setup for Modular Towers (as in [**BFr02**]). Here we illustrate possibilities for the Gauss-Riemann discussion, which have modern counterparts (say, in the Inverse Galois Problem and cryptology).

5.1. Universal p-Frattini cover. For any prime p, consider the pro-p completion ${}_pF$ of F (or L if that is the case). A pro-p group $\tilde{P}$ has a Frattini subgroup $\Phi(\tilde{P})$ generated by its pth powers and commutators. For p not dividing $|H|$, the H action extends to ${}_pF$, and gives ${}_pF \times^s H$ ([**BFr02**, Rem. 5.2] or [**FJ86**, Chap. 21]). This is the universal p-Frattini extension of ${}_pF/\Phi({}_pF) \times^s H = G$ (this is G_0 below).

For any finite group G and each prime p, $p \,|\, |G|$, there is a universal p-Frattini cover $\psi_p : {}_p\tilde{G} \to G$ with these properties [**Fr95a**, Part II].

(5.1a) $\tilde{G}$ is the fiber product of ${}_p\tilde{G}$ (over G) over p primes dividing $|G|$.
(5.1b) Both $\ker(\psi_p)$ and a p-Sylow of ${}_p\tilde{G}$ are pro-free pro-p groups, and ${}_p\tilde{G}$ is the minimal profinite cover of G with this property.
(5.1c) ${}_p\tilde{G}$ has a characteristic sequence of finite quotients $\{G_k\}_{k=0}^{\infty}$.
(5.1d) Each p'-conjugacy class of G lifts uniquely to a p' class of ${}_p\tilde{G}$.
(5.1e) If $G^* \le G_k$ has image in G all of G, then $G^* = G_k$.

Denote $\ker(\psi_p)$ by $\ker_0$. For any pro-p group, the Frattini subgroup is the closed subgroup that commutators and pth powers generate. Let $\ker_1$ be the Frattini subgroup of $\ker_0$. Continue inductively to form $\ker_k$ as the Frattini subgroup of $\ker_{k-1}$. Then, $G_k = {}_p\tilde{G}/\ker_k$. To simplify notation, suppress the appearance of p in forming the characteristic sequence $\{G_k\}_{k=0}^{\infty}$. Use of modular representation theory throughout this paper is from the action of G_k on $\ker_k / \ker_{k+1} \stackrel{\text{def}}{=} M_k$, a natural $\mathbb{Z}/p[G_k]$ module.

Suppose p does not divide H, but P is a p-group, and $\tilde{P}$ is minimal pro-free pro-p cover of P. Then, the universal p- Frattini cover of $P \times^s H$ is $\tilde{P} \times^s H$ from the H action extending to $\tilde{P}$. Applications must consider the general case where $p||H|$. Then, [**Fr02**, Prop. 2.8] gives the rank of the p-Frattini kernel. The key information comes from a two step process going from the normalizer N of a p-Sylow of G_0, which is a split case from which we compute $\ker_0 / \ker_1$, which is an N module. The correct rank is that of a natural indecomposable module in the G_0 module induced from N. We give one case of it here, for contrast with our main example.

EXAMPLE 5.1 ($G_0 = A_5$). F is trivial, and $H = G_0$ is A_5. We use the prime $p = 2$ and $\mathbf{C} = \mathbf{C}_{3^4}$, four conjugacy classes of elements of order 3. For absolute equivalence for $\mathrm{Ni}(G_0, \mathbf{C}_{3^4})$ use the cosets of A_4. Call this permutation representation T_{A_4}. For larger values of k, [**BFr02**] uses several different coset representations extending the standard one, for good reason as we will see.

5.2. Using $D_\infty = \mathbb{Z} \times^s H_2$ as a model. We give an analog of the modular curve situation based on using four 3-cycles.

5.2.1. *Finite quotients of the group $F_2 \times^s H_3$.* Let $H = H_3 = \mathbb{Z}/3$ act on a free group F_2 with two generators $\boldsymbol{v}_1, \boldsymbol{v}_2$: $\langle \mu \rangle \stackrel{\text{def}}{=} \mathbb{Z}/3$ acts as $(\boldsymbol{v}_1, bv_2) \mapsto (\boldsymbol{v}_2^{-1}, \boldsymbol{v}_1\boldsymbol{v}_2^{-1})$. We use the conjugacy classes $\mathbf{C}_{\pm 3^2}$: Four conjugacy classes of elements of order 3, two mapping to $\mu \in \mathbb{Z}/3$ and two mapping to $-\mu$. That is As in (5.1d), regard $\mathbf{C}_{\pm 3^2}$ as conjugacy classes in all quotients of ${}_pF^2 \times^s H_3 = {}_p\tilde{G}$. So, we avoid only the prime 3 (akin to 2 for the modular curve case). For any other prime p, G_k in our notation above is $G_k((\mathbb{Z}/p)^2) \times^s H_3$. Use a copy of H_3 in $G_k((\mathbb{Z}/p)^2) \times^s H_3$ for each k ($p \neq 3$) to define absolute classes, denoting the corresponding permutation representation by T_{H_3}. As in Ex. 5.1, this may not always be the best choice.

The following statements are done in great generality in [**FV91**]. The collection of conjugacy classes in both examples is a *rational union*. In our illustrating examples we use that all spaces formed from a Nielsen class $\mathrm{Ni}(G, \mathbf{C})$ where $\mathbf{C}$ is a rational union have equations over $\mathbb{Q}$. They, may, however, have components not defined over $\mathbb{Q}$. Further, we use that inner spaces have unique total families over $\mathbb{Q}$ if there is also no center, and the same holds for absolute spaces if the image of G under the permutation representation $T : G \to S_N$ has no centralizer in S_N.

PROPOSITION 5.2. *The Nielsen class* $\mathrm{Ni}(G_k((\mathbb{Z}/p)^2) \times^s H_3, \mathbf{C}_{\pm 3^2}) = \mathrm{Ni}$ *is nonempty. Covers in the inner classes form a space analogous to $X_1(p^{k+1})$; in the absolute classes analogous to $X_0(p^{k+1})$.*

FORMING NONEMPTY SPACES. For each k we show there are Harbater- Mumford (H-M) reps.: $(g_1 g_1^{-1}, g_2, g_2^{-1})$. Since G_k is a Frattini cover of G_0, we can lift any elements g_1, g_2 having order 3 to G_k. The Frattini property says they automatically generated G_k, and so produce an H-M rep. at level k. So, it suffices to find two order 3 generating elements $g_1, g_2 \in (\mathbb{Z}/p)^2) \times^s H_3$ mapping to μ. For the action of H on $(\mathbb{Z}/p)^2$ there are no invariant subspaces. Take $g_1 = \mu$ and $g_2 = (v^\mu - v, \mu)$ for any v not commuting with μ. When $r = 4$,

$$\gamma_0 = q_1 q_2,\ \gamma_1 = q_1 q_2 q_1,\ \gamma_\infty = q_2 \tag{5.2}$$

acting on $\mathrm{Ni}(G, \mathbf{C})^{\mathrm{rd}}$ give a branch cycle description of this space as a cover of the j-line. Just apply this to $\mathrm{Ni}(G_k, \mathbf{C}_{\pm 3^2})^{\mathrm{rd}}$. This is a major point: [**BFr02**, Prop. 4.4] or [**DF99**, Prop. 6.5]. □

5.2.2. *Family from Ex. 5.1.* Each $\boldsymbol{p} \in \mathcal{H}(G_0, \mathbf{C}_{3^4})^{\mathrm{abs,rd}}$ corresponds to a cover $\varphi_{\boldsymbol{p}} : X_{\boldsymbol{p}} \to \mathbb{P}^1_z$ up to reduced equivalence. If $\boldsymbol{z}(\boldsymbol{p})$ are the four unordered branch points of the cover, we see the following.

(5.3) From R-H: $g(X_{\boldsymbol{p}}) = 0$ and the unique point $x_i \in X_{\boldsymbol{p}}$ over $z_i \in \boldsymbol{z}$ corresponding to the 3-cycle ramification ramification, gives an unordered 4-tuple $\boldsymbol{x}$ on $X_{\boldsymbol{p}}$ associated to $\boldsymbol{p}$.

PROPOSITION 5.3. *The map* $\boldsymbol{p} \in \mathcal{H}(G_0, \mathbf{C}_{3^4})^{\mathrm{abs,rd}} \mapsto (j_{\boldsymbol{z}(\boldsymbol{p})}, j_{\boldsymbol{x}(\boldsymbol{p})}) = \psi(\boldsymbol{p})$ *is generically one-one. The projection* $\psi(\boldsymbol{p}) \mapsto j_{\boldsymbol{z}(\boldsymbol{p})}$ *has degree 9 and monodromy* *group A_9. So, this presentation of $\mathcal{H}(G_0,\mathbf{C}_{3^4})^{\mathrm{abs,rd}}$ resembles the description of modular curves given by Prop. 4.5, yet it is not a modular curve.*

USING THE **sh**-INCIDENCE MATRIX. This is a brief review of [**BFr02**, §2.9]. For a general reduced Nielsen class, list γ_∞ orbits as $O_1,\dots,O_n$. The **sh**-incidence matrix $A(G,\mathbf{C})$ has (i,j) term $|(O_i)\mathbf{sh}\cap O_j|$. Since **sh** has order two on reduced Nielsen classes, this is a symmetric matrix when $r=4$. Equivalence $n\times n$ matrices A and $TA{}^{\mathrm{t}}T$ running over permutation matrices T (${}^{\mathrm{t}}T$ is its transpose) associated to elements of S_n. List γ_∞ orbits as

$$O_{1,1},\dots,O_{1,t_1},O_{2,1},\dots,O_{2,t_2},\dots,O_{u,1},\dots,O_{u,t_u}$$

corresponding to $\bar{M}_4$ orbits. Choose T to assume $A(G,\mathbf{C})$ is arranged in blocks along the diagonal. The blocks correspond to connnected components of $\mathcal{H}(G,\mathbf{C})^{\mathrm{rd}}$.

5.2.3. **sh**-*incidence matrix of* $\mathrm{Ni}(A_5,\mathbf{C}_{3^4})^{\mathrm{in,rd}}=\mathrm{Ni}_0^{\mathrm{in,rd}}$. Denote γ_∞ orbits of

$$\boldsymbol{g}_1=((1\,2\,3),(1\,3\,2),(1\,4\,5),(1\,5\,4)) \text{ and } \boldsymbol{g}_2=((1\,2\,3),(1\,3\,2),(1\,5\,4),(1\,4\,5))$$

by $O(5,5;1)$ and $O(5,5;2)$; γ_∞ orbits of

$$((5\,1\,3),(2\,4\,5),(1\,5\,4),(1\,2\,3)) \text{ and } ((3\,2\,4),(5\,1\,3),(1\,5\,4),(1\,2\,3))$$

by $O(3,3;1)$ and $O(3,3;2)$; and of $(\boldsymbol{g}_1)\mathbf{sh}$ by $O(1,2)$. Add conjugation by a 2-cycle

TABLE 1. **sh**-Incidence Matrix for Ni_0

Orbit	$O(5,5;1)$	$O(5,5;2)$	$O(3,3;1)$	$O(3,3;2)$	$O(1,2)$
$O(5,5;1)$	0	2	1	1	1
$O(5,5;2)$	2	0	1	1	1
$O(3,3;1)$	1	1	0	1	0
$O(3,3;2)$	1	1	1	0	0
$O(1,2)$	1	1	0	0	0

to get this for $\mathrm{Ni}_0^{\mathrm{abs,rd}}$, which gives exactly 3 γ_∞ orbits of respective widths 5, 3 and 1. So the monodromy group is a primitive subgroup of A_9 containing a 3-cycle, and so it is A_9. ☐

5.2.4. *Modular Curve example.* How **sh**-incidence matrix shows there is one component. Here p is odd.

Describe Nielsen classes $\mathrm{Ni}(D_p,\mathbf{C}_{2^4})^{\mathrm{in}}$ as 4-tuples (b_1,b_2,b_3,b_4) with $b_1,\dots,b_4\in\mathbb{Z}/p^{k+1}$, where the differences b_j-b_{j+1}, $j=1,2,3$, generate, $b_1-b_2+b_3-b_4=0$, modulo translation by elements of $\mathbb{Z}/p^{k+1}$ and $(b_1,b_2,b_3,b_4)\mapsto-(b_1,b_2,b_3,b_4)$.

For absolute Nielsen classes $\mathrm{Ni}(D_{p^{k+1}},\mathbf{C}_{2^4})^{\mathrm{abs}}$, allow all affine actions, including $(b_1,b_2,b_3,b_4)\mapsto a(b_1,b_2,b_3,b_4)$, $a\in(\mathbb{Z}/p^{k+1})^*$. Further, a γ_∞ orbit consists of the set (b_1,b_2+k,b_3+k,b_4) with $0\le k\le \mathrm{ord}(b_2-b_3)$.

The length one orbits for γ_∞ have representatives $(0,b,b,0)$ $(b\ne0)$, so there is just one absolute width one absolute orbit, and $\frac{p-1}{2}$ such inner orbits. In all other cases the orbits have width p, and so there is just absolute orbit (with representative $(0,0,1,1)$, an H-M representative) and $\frac{p-1}{2}$ such inner orbits. Let T_n be the $n\times n$ matrix with 1's in each position.

PROPOSITION 5.4. *Assume $k = 0$. Reduced Nielsen classes are the same as Nielsen classes because the action of $\mathcal{Q}''$ is trivial. The* **sh***-incidence matrix for* $\mathrm{Ni}(D_p, \mathbf{C}_{2^4})^{\mathrm{abs,rd}}$ *is* $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. *That for* $\mathrm{Ni}(D_p, \mathbf{C}_{2^4})^{\mathrm{in,rd}}$ *is* $\begin{pmatrix} 0_n & I_n \\ I_n & T_n \end{pmatrix}$.

PROOF. Label the inner γ_∞ width one orbits as $O_{1,b}$, as given by the representatives $(0, b, b, 0)$, $b = 1, \dots, p-1$. Similarly, Label the inner γ_∞ width p orbits as $O_{p,b}$, as given by γ_∞ orbit representatives $(0, 0, b, b)$, $b = 1, \dots, p-1$.

Now check $|O_{1,b} \cap (O_{p,b'})\mathbf{sh}| = 1$: The one element of intersection is $(0, b, b, 0)$ occurring if and only if $b = b'$. Further $|O_{1,b} \cap (O_{1,b'})\mathbf{sh}| = \delta_{b,b'}$. That leaves us tto show $|O_{p,b} \cap (O_{p,b'})\mathbf{sh}| = 1$ for each pair of nonzero entries $b, b' \in (\mathbb{Z}/p)^*$.

Inner representatives of $O_{p,b}$ are the elements $\{(0, a, a+b, b\}_{a \in \mathbb{Z}/p}$. Applying **sh** to these gives these as inner representatives: $\{(0, b, b-a, -a)\}_{a \in \mathbb{Z}/p}$. This has exactly one representative in $\{(0, a', a'+b', b'\}_{a' \in \mathbb{Z}/p}$, when $b = a'$ and $-a = b'$. □

The proposition of §5.2 is not really far from a Gauss-Riemann discussion. We'll do the first case to see that. Take $k = 0$, $p = 2$ and absolute classes. We are looking at $\mathcal{H}(G_0((\mathbb{Z}/2)^2) \times^s H_3, \mathbf{C}_{\pm 3})^{\mathrm{abs,rd}} = \mathcal{H}_{0,2}^{\mathrm{abs,rd}}$. Each $\boldsymbol{p} \in \mathcal{H}_{0,2}^{\mathrm{abs,rd}}$ corresponds to a cover $\varphi_{\boldsymbol{p}} : X_{\boldsymbol{p}} \to \mathbb{P}^1_z$ with the genus of $g(X_{\boldsymbol{p}}) = 1$.

The end of §4.2.3 discusses the symbiotic relation between genus 0 and genus 1, and identifies the practical point of recognizing genus 1 surfaces as complex torii. It poses if the genus 1 curves in this space relate to those given by modular curves. This relates to (3.5a): Do the functions and differentials described by Abel's Theorem relate to these $X_{\boldsymbol{p}}$ s. If we could answer yes to this, then the nature of the cut construction guarantees there is a map from $\mathcal{H}_{0,2}^{\mathrm{abs,rd}}$ to the j-line by mapping $\boldsymbol{p}$ to the j-invariant of $X_{\boldsymbol{p}}$. Then we would know which of the $X_{c,d}$ are appearing in this family? Even it you already know about elliptic curves, isn't it possible that only one is?

PROPOSITION 5.5. *The space $\mathcal{H}_{0,2}^{\mathrm{abs,rd}}$ has two components.*

6. Riemann's formulation of the generalization

A collection of results about the complex 1-dimensional torus case that require generalization. We start with a compact Riemann surface cover $\varphi : X \to \mathbb{P}^1_z$. We understand others might choose a different starting point, for example not having φ at all, just X. Our points are around this φ. Use the notation $X_\varphi^0 = X \setminus \varphi^{-1}(\boldsymbol{z})$.

(6.1a) Some algebraic function $\varphi : X \to \mathbb{P}^1_w$ separates points of X_{z_0}, $z_0 \notin \boldsymbol{z}$.
(6.1b) Any degree 0 divisor on X is the polar divisor of a local log differential.
(6.1c) The universal covers of $\tilde{X}$ and $\tilde{X}_\varphi^0$ are analytic open sets in $\mathbb{P}^1_w$.

These properties have subtle relations, exposed from different approaches to filling in Riemann's legacy. We get back to the Gauss-Riemann conversation with some of these. Example: If any unitary line bundle gives a divisor and associated Green's function, that would invert the relation between local log differentials and such bundles, answering (6.1b). It is an alternate way to say existence of a theta function.

6.1. Using the existence theorem to uniformize compact surface minus points. We start by showing topologically what is the universal covering space of a compact surface minus a finite set of points.

Lemma 6.1. *For a fixed $g \geq 0$, there are ∞-ly many n with 3 branch point $\varphi : X \to \mathbb{P}^1$, $g(X) = g$ and monodromy group S_n. In both cases in* (6.1c)*, the universal covering space is topologically a subset of the plane. Further, any genus g compact Riemann surface with r punctures ($r \geq 3$ if $g = 0$, $r \geq 1$, if $g = 1$), has the disk as a universal cover.*

Proof. We do just the compact case using the cuts and the function $\lambda(\tau)$. Fix g and consider the following branch cycle descriptions giving 3 branch point $\varphi : X \to \mathbb{P}^1$, $g(X) = g$, monodromy S_n.

Take $n = m_1 + \cdots + m_s$. Modify $S_{n,0}$ covers:

$$g_1 = (1 \dots m_1) \cdots (m_1 + \cdots + m_{s-1} + 1 \dots n).$$

Genus 0: Take

$$\begin{aligned} g_2 = \quad & (m_1 - 1 \dots 1)(m_1 + m_2 - 1 \dots m_1 + 1) \cdots \\ & (n-1 \dots n - m_s + 1)(m_1\, m_1 + m_2 \dots n). \end{aligned}$$

Then, $\mathrm{ind}(g_1) = n - s$ and $\mathrm{ind}(g_2) = n - s - 1$. Compute $g_1 g_2 = g_3$:

$$(m_1\, m_1 - 1\, m_1 + m_2\, m_1 + m_2 - 1 \dots n\, n-1).$$

So, $\mathrm{ind}(g_3) = 2s - 1$. RET gives genus 0 cover. For S_n, select $m_1, \dots, m_s$ accordingly.

For $g = 1$ covers, switch 1 and 2 in g_1, but not in g_2. Changes nothing from conclusions, except adding 2 to index of g_3:

$$(m_1\, 2\, 1\, m_1 - 1\, m_1 + m_2\, m_1 + m_2 - 1 \dots n\, n-1).$$

$\square$

6.1.1. *Complex spaces, topologically a subspace of $\mathbb{P}^1_z$.* What we want to know is that the universal covering spaces are *analytically* a subset of the plane. From there we have the Riemann mapping Theorem, which shows that we have a disk or the complex plane analytically. All we need is a one-one analytic map to $\mathbb{P}^1_z$. [**Ah79**, p. 248–251] The dirichlet problem on a disk or something conformal to a disk, extended by using Perron's method of exhaustion. [**Ah79**, p. 257–259] Green's functions: $g(z, z_0) = G(z, z_0) - \log|z - z_0|$, G is symmetric and harmonic in each variable, with boundary values $\log|z - z_0|$, so g has boundary values 0.

It is interesting to consider [**H49**] because it comes long after the topic of Riemann's resurrection and an initial concern about algebra versus function theory, and long before modern decisions about that significance. Starts by referring to the modern treatment of the Dirichlet Problem by Perron and Carathéodory, subject of a treatment by Ahlfors and Beurling. Our case is one — as explained — where we already know the space is topologically a simply-connected subset of the plane, but we do not know if its complex structure comes from the plane. All we need is a one-one analytic map to the plane, and then the Riemann mapping theorem in the plane tells us what we want. It is really quite simple to conclude the existence of a Green's function $g_n(\cdot, q)$, with a pole at $q \in T_n$, for $p \in T_n$ on a set T_n with compact closure, so the sets T_n are increasing, have compact closure and exhaust the open surface $\tilde{X}$.

It is well-known $\{g_n(p, q)\}$ is monotone non-decreasing for each p, and its limit $g(p, q)$ is either $+\infty$ on $\tilde{X}$ (parabolic case) or harmonic and nonnegative on X save at q where it has a logarithmic singularity (hyperbolic case). That is, $g(p, q) \equiv -\log(r_z) + h_z(p)$ were z is a local parameter in a neighborhood of q, $r_z = |z(p) - z(q)|$

and $h_z(p)$ is harmonic at q. The two cases don't depend on the sequence or the exhaustion.

The monodromy principle says there is $f(p,q)$ analytic on $\tilde{X}$ with $\log|f(p,q)| \equiv -g(p,q)$, determining f up to constant factor of modulus 1, and f has a simple zero at q and modulus less than one on $\tilde{X}$.

If we knew (for q fixed) that g tends to zero as p goes to the ideal boundary of $\tilde{X}$, then $|f|$ would tend to one as p goes to the ideal boundary and so each point $|w|<1$ would be covered equally often and the origin precisely once. The rest is a detour around this information not being available.

Let $q_1 \ne q_2 \in X$, and consider $f(p,q_1)$ and $f(p,q_2)$ for comparison. Let

$$\varphi(p) = \frac{f(p,q_1) - f(q_2,q_1)}{1 - \bar{f}(q_2,q_1) f(p,q_1)}.$$

This is analytic, modulus less than 1 on $\tilde{X}$ and vanishes at q_2. By the maximum principle, for large n, $\log|\varphi(p)| \le -g_n(p,q_2), p \in T_n$. So, $\log|\varphi(p)| \le -g(p,q_2) = \log|f(p,q_2)|$, $p \in \tilde{X}$.

For $p = q_1$, these imply $|f(q_2,q_1)| \le |f(q_1,q_2)|$ and by symmetry equality. By the maximum principle this implies $\varphi(p)$ is identically a constant of modulus 1 times $f(p,q_2)$. So, $\varphi(p)$ vanishes only at q_2 and just once. Thus, $f(p,q_1) = f(q_2,q_1)$ implies $p = q_2$ and since q_2 is arbitrary in $\tilde{X}$, $f(p,q_1)$ is one-one in p. He then does the parabolic case.

6.2. Comparing with modular curves. Suppose $\Phi^{\rm rd} : \mathcal{H}^{\rm rd} \to \mathbb{P}^1_j$ is the compactification of $\mathbb{H}/\Gamma$ with $\Gamma \le \mathrm{SL}_2(\mathbb{Z})$. Let N_Γ be the least common multiple of the cusp widths. Equivalently: N_Γ is the least common multiple of the ramification orders of points of $\mathcal{H}^{\rm rd}$ over $j=\infty$; or the order of γ_∞ on reduced Nielsen classes. Wohlfahrt's Theorem [**Wo64**] says Γ is congruence if and only if Γ contains $\Gamma(N_\Gamma)$. We are most interested in considering situations that have a modular curve-like aspect, though they are not modular curves. So, a necessary numerical check for $\Phi^{\rm rd}$ to be modular, starts by computing γ_∞ orbits on $\mathrm{Ni}^{\rm rd}$, and then checking their distribution among $\bar{M}_4 = \langle \gamma_\infty, \mathbf{sh}\rangle$ orbits, which correspond to $\mathcal{H}^{\rm rd}$ components. For each component $\mathcal{H}'$ of $\mathcal{H}^{\rm rd}$ check separately the lcm of length of γ_∞ orbits to compute N'. Then, there is a final check if there is a permutation representation of $\Gamma(N')$ that could produce the cover $\Phi' : \mathcal{H}' \to \mathbb{P}^1_j$, and the type of cusps now computed.

6.2.1. *A nonmodular family of elliptic curves.* We can select an example to illustrate all aspects of this prescription. It is a family of genus 1 curves. Take as Nielsen classes in A_4, $\mathbf{C}_{\pm 3}$: Two pairs of 3-cycles in each of the two conjugacy classes having order 3. The total Nielsen class $\mathrm{Ni}(A_3, C_{\pm 3^2})$ contains exactly six elements corresponding to the six possible arrangements of the conjugacy classes. Since A_3 is abelian, the inner classes are the same. Also, the outer automorphism of A_n ($n=3$ or 4) from conjugation by $(1\,2) \in S_n$ restricts to A_3 to send a conjugacy class arrangement to its complement. Here is a convenient list of the arrangements, and their complements:

$$\begin{array}{lll} [1]\,+-+- & [2]\,++-- & [3]\,+--+ \\ [4]\,-+-+ & [5]\,--++ & [6]\,-++-. \end{array}$$

The group $\mathcal{Q}'' = \langle q_1 q_3^{-1}, \mathbf{sh}^2\rangle$ equates elements in this list and their complements. So, absolute and inner reduced classes are the same. Conclude: $\mathcal{H}(A_3, \mathbf{C}_{\pm 3^2})^{\rm in,rd} \to$

$\mathbb{P}^1_j$ is a degree three cover with branch cycles $(\gamma_0^*, \gamma_1^*, \gamma_\infty^*) = ((1\,3\,2), (2\,3), (1\,2))$. It is an easy check that if $(g_1, \dots, g_4)$ maps to [1], and (with no loss) $g_1 = (1\,2\,3)$, then either this is $\boldsymbol{g}_{1,1}$ or g_1g_2 has order 2. Listing the four elements of order 2 gives a total of five elements in the reduced Nielsen class $\mathrm{Ni}(A_4, \mathbf{C}_{\pm 3^2})^{\mathrm{in,rd}}$ lying over [1].

6.2.2. *Effect of γ_∞ on the Nielsen class.* Start with an H-M rep over [1] in A_3:

$$\boldsymbol{g}_{1,1} = ((1\,2\,3), (1\,3\,2), (1\,3\,4), (1\,4\,3)) \in \mathrm{Ni}(A_4, \mathbf{C}_{\pm 3^2}).$$

The middle twist squared on this conjugates the middle two by $(1\,4)(2\,3)$ to give

$$\boldsymbol{g}_{1,2} = ((1\,2\,3), (4\,2\,3), (4\,2\,1), (1\,4\,3)).$$

The result is a γ_∞ orbit of length 4. The middle twist squared on

$$\boldsymbol{g}_{1,3} = ((1\,2\,3), (1\,2\,4), (1\,4\,2), (1\,3\,2))$$

leaves it fixed, giving a γ_∞ orbit of length 2. Similarly, the square of the middle twist on $\boldsymbol{g}_{1,4} = ((1\,2\,3), (1\,2\,4), (1\,2\,3), (1\,2\,4))$ conjugates the middle pair by $(1\,3)(2\,4)$ producing $\boldsymbol{g}_{1,5} = ((1\,2\,3), (1\,2\,4), (2\,4\,3), (1\,4\,3))$. Again the middle twist gives an element of order 4 on reduced Nielsen classes.

Let $\boldsymbol{g}_{3,1} = ((1\,2\,3), (1\,3\,2), (1\,4\,3), (1\,3\,4)) \in \mathrm{Ni}(A_4, \mathbf{C}_{\pm 3^2})$. This maps to [3] in A_3 and it is an H-M rep. Applying γ_∞ gives $\boldsymbol{g}_{3,2} = ((1\,2\,3), (1\,2\,4), (1\,3\,2), (1\,3\,4))$, which is the same as conjugating on the middle two by $(2\,4\,3)$. The result is a γ_∞ orbit of length 3.

On Nielsen class representatives over [3], γ_∞ has one orbit of length 3 and two of length one. See this by listing the second and third positions (leaving $(1\,2\,3)$ as the first). Label these as

$$\begin{aligned} &1' = ((1\,3\,2), (1\,4\,3)), 2' = ((1\,2\,4), (1\,3\,2)), 3' = ((1\,2\,4), (2\,3\,4)),\\ &4' = ((1\,2\,4), (1\,2\,4)), 5' = ((1\,2\,4), (1\,4\,3)). \end{aligned}$$

Proposition 6.2. *Then, γ_∞ fixes $4'$ and $5'$ and cycles $1' \to 2' \to 3'$. So there are two $\bar{M}_4$ orbits on $\mathrm{Ni}(A_4, \mathbf{C}_{\pm 3^2})^{\mathrm{in,rd}}$, Ni_0^+ and Ni_0^-, having respective degrees 9 and 6 and respective lifting invariants to $\hat{A}_4$ of $+1$ and -1. The first, containing all H-M reps., has orbit widths 2,4 and 3. The second has orbit widths 1,1 and 4. Neither defines a modular curve cover of $\mathbb{P}^1_j$.*

Denote the corresponding H_4 orbits on $\mathrm{Ni}(A_4, \mathbf{C}_{\pm 3^2})^{\mathrm{in}}$ by $\mathrm{Ni}_0^{\mathrm{in},+}$ and $\mathrm{Ni}_0^{\mathrm{in},-}$. The $\mathcal{Q}''$ orbits on both have length 2.

Proof. As γ_∞ fixes $4'$ and it maps $5'$ to $((1\,2\,3), (2\,3\,4), (1\,2\,4), (3\,1\,2))$ (conjugate by $(1\,2\,3)$ to $5'$), these computations establish the orbit lengths:

$$\begin{aligned} (g_{1,1})\gamma_\infty &= ((1\,2\,3), (1\,4\,2), (1\,3\,2), (1\,4\,3)) = (3')\mathbf{sh},\\ (g_{1,3})\gamma_\infty &= ((1\,2\,3), (1\,4\,2), (1\,2\,4), (1\,3\,2)) = (1')\mathbf{sh}. \end{aligned}$$

They put the H-M rep. in the $\bar{M}_4$ orbit with γ_∞ orbits of length 2,3 and 4 (in the orbit of the $1' \to 2' \to 3'$ cycle). Use Ni_0^+ for the Nielsen reps. in this $\bar{M}_4$ orbit.

Apply $\mathbf{sh}$ to $4'$ to see $g_{1,4}, g_{1,5}, 4', 5'$ all lie in one $\bar{M}_4$ orbit. Any H-M rep. has lifting invariant $+1$, and since it is a $\bar{M}_4$ invariant, all elements in Ni_0^+ have lifting invariant $+1$. For the other orbit, we have only to check the lifting invariant on $4'$, with its full expression given by $((1\,2\,3), ((1\,2\,4), (1\,2\,4), (4\,3\,2)) = (g_1, \dots, g_4)$. Compute the lifting invariant as $\hat{g}_1\hat{g}_2\hat{g}_3\hat{g}_4)$. Since $g_2 = g_3$ (and their lifts are the same), the invariant is $\hat{g}_1\hat{g}_2^2\hat{g}_4$. Apply Serre's formula (not necessary, though illuminating). The hypotheses of genus zero for a degree 4 cover hold for $((1\,2\,3), (1\,4\,2), (4\,3\,2))$ and the lifting invariant is $(-1)^{3\cdot(3^2-1)/8} = -1$.

From Wohlfahrt's Theorem [**Wo64**], if the degree nine orbit is modular, the group is a quotient of $\mathrm{PSL}_2(12)$. If the degree 6 orbit is modular, the group is a quotient of $\mathrm{PSL}_2(4)$. Since $\mathrm{PSL}_2(\mathbb{Z}/4)$ has the λ-line as a quotient, with 2,2,2 as the cusp lengths, these cusp lengths are wrong for the second orbit to correspond to the λ-line. Similarly, for the longer orbit, as $\mathrm{PSL}_2(\mathbb{Z}/12)$ has both $\mathrm{PSL}_2(\mathbb{Z}/4)$ and $\mathrm{PSL}_2(\mathbb{Z}/3)$ as a quotient, and so the cusp lengths are wrong.

We can check the length of a $\mathcal{Q}''$ orbit on $\mathrm{Ni}_0^{\mathrm{in},+}$ and $\mathrm{Ni}_0^{\mathrm{in},-}$ by checking the length of the orbit of any particular element. If an orbit has an H-M rep. like $\boldsymbol{g}_{1,1}$ it is always convenient to check that: $(\boldsymbol{g}_{1,1})\mathbf{sh}^2 = (1\,3)(2\,4)\boldsymbol{g}_{1,1}(1\,3)(2\,4)$ and $(\boldsymbol{g}_{1,1})q_1q_3^{-1} = (1\,3)\boldsymbol{g}_{1,1}(1\,3)$. So, in $\mathrm{Ni}_0^{\mathrm{in},+}$, $\mathbf{sh}^2$ fixes $\boldsymbol{g}_{1,1}$, but $q_1q_3^{-1}$ does not. For $\mathrm{Ni}_0^{\mathrm{in},-}$, $\boldsymbol{g}_{1,4}$ is transparently fixed by $\mathbf{sh}^2$, and $(\boldsymbol{g}_{1,4})q_1q_3^{-1} = (3\,4)\boldsymbol{g}_{1,4}(3\,4)$. Conclude the orbit length of $\mathcal{Q}''$ on both $\mathrm{Ni}_0^{\mathrm{in},+}$ and $\mathrm{Ni}_0^{\mathrm{in},-}$ is 2. $\square$

The **sh**-incidence matrix of Ni_0^+ comes from the following data. Elements $\boldsymbol{g}_{1,1},\boldsymbol{g}_{1,2},\boldsymbol{g}_{1,3}$ over [1] map by γ_∞ respectively to $\boldsymbol{g}_{2,1},\boldsymbol{g}_{2,2},\boldsymbol{g}_{2,3}$ over [2], and these map respectively to $\boldsymbol{g}_{1,2},\boldsymbol{g}_{1,1},\boldsymbol{g}_{1,3}$, while $\boldsymbol{g}_{3,1},\boldsymbol{g}_{3,2},\boldsymbol{g}_{3,3}$ cycle among each other under γ_∞. So, there are three γ_∞ orbits, $O_{1,1}$, $O_{1,3}$ and $O_{3,1}$ on Ni_0^+ named for the subscripts of a representing element.

The data in the proof of the proposition shows

$$|O_{1,1} \cap (O_{3,1})\mathbf{sh}| = 2,\ |O_{1,3} \cap (O_{3,1})\mathbf{sh}| = 1.$$

Compute: **sh** applied to $\boldsymbol{g}_{1,3}$ is $g_{1,1}$ so $|O_{1,1} \cap (O_{1,3})\mathbf{sh}| = 1$. The rest are determined by symmetry and having elements in a row or column add up to the total number elements in the set labeling that row or column.

TABLE 2. **sh**-Incidence Matrix for Ni_0^+

Orbit	$O_{1,1}$	$O_{1,3}$	$O_{3,1}$
$O_{1,1}$	1	1	2
$O_{1,3}$	1	0	1
$O_{3,1}$	2	1	0

Similarly, the **sh**-incidence matrix of Ni_0^- comes from the following data. Elements $\boldsymbol{g}_{1,4},\boldsymbol{g}_{1,5}$ over [1] map by γ_∞ respectively to $\boldsymbol{g}_{2,4},\boldsymbol{g}_{2,5}$ over [2], and these map respectively to $\boldsymbol{g}_{1,5},\boldsymbol{g}_{1,4}$, while γ_∞ fixes both $\boldsymbol{g}_{3,4},\boldsymbol{g}_{3,5}$. So, there are three γ_∞ orbits, $O_{1,4}$, $O_{3,4}$ and $O_{3,5}$ on Ni_0^-.

TABLE 3. **sh**-Incidence Matrix for Ni_0^-

Orbit	$O_{1,4}$	$O_{3,4}$	$O_{3,5}$
$O_{1,4}$	2	1	1
$O_{3,4}$	1	0	0
$O_{3,5}$	1	0	0

6.3. From whence the genus one covers.

6.3.1. *Genus greater than one and the universal cover.* Cuts gave Riemann clues the universal covering space of X is analytically a disk if $g_X > 1$.

Can get $g_X = 1$ (from Riemann Hurwitz; say by taking n 3-cycles), though we can't assure such a map will go with a specific (c, d).

Gauss' approach to the hypergeometric equation (known to us through the Schwarz-Cristoffel transformation; see §7.2) gave many examples in imitation of the complex torus case. That is, by integration and forming an inverse of an anti-derivative, they gave a clear simply connected fundamental domain on the upper half plane for the Riemann surface. This gave examples showing the universal covering space is analytically a disk. This must have given Riemann confidence in the result that was eventually in contention.

PROPOSITION 6.3. *The universal covering space for a compact Riemann surface X, when $g_X > 1$, would always be a disk (*[**Spr57**, p. 219–225]*, proof adapted from* [**Wey55**]*, though due to Koebe).*

6.4. Completing the genus one case. For some —especially complex geometers —it is reassuring to see the particular θ ($\sigma(u)$ in §3.1) function on $\mathbb{C}_u$. For others —especially algebraists —enhancing characterizations of it is more important. The two views come at loggerheads over places where curves meet abelian varieties. The now very old problem of characterizing (the Jacobian variety of) curves among all Abelian varieties is one such topic. The topic of those moduli spaces called Shimura varieties is close to pure abelian variety territory. Other moduli spaces use algebraic varieties characterized as families of curves that do not include all curves of a given genus. Here the look of expressions for θ functions helps much less.

6.5. Other uses of the cuts. Any holomorphic differential ω on X from a cut construction $\varphi : X \to \mathbb{P}^1_z$ has a description around $z_0 \in U_{\boldsymbol{z}}$ as $h(z)\,dz$. We can then integrate it along any $\gamma \in \pi_1(U_{\boldsymbol{z}}, z_0)$ or in $\pi_1(X, x_0)$. The generalization of Abel's condition comes by taking a basis $\mathcal{B} = (\omega_1, \dots, \omega_g)$ of holomorphic differentials (g is the genus, though that isn't obvious) and integrals along paths on X. Define L_X to be $\{\int_\gamma \mathcal{B} \mid \gamma \in H_1(X, \mathbb{Z})\}$.

PROPOSITION 6.4. *If $D^0 = \varphi^{-1}(0)$ and $D^\infty = \varphi^{-1}(\infty)$, for an n-tuple of lifts $\gamma_1, \dots, \gamma_n$ of $\gamma \in \pi_1(U_{\boldsymbol{z}}, z_0)$,*

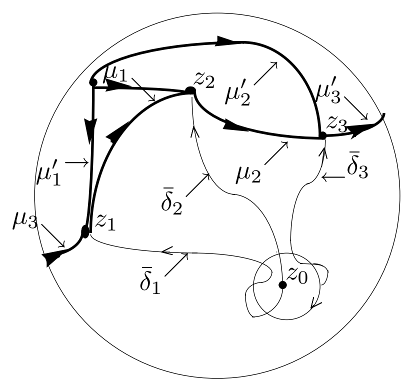
$$(\sum_{i=1}^{n} \int_{\gamma_i} \omega_1, \dots, \sum_{i=1}^{n} \int_{\gamma_i} \omega_g) = \int_{\boldsymbol{\gamma}} \mathcal{B} = \mathbf{0}. \tag{6.2}$$

First part of Riemann's Theorem: Having paths $\gamma_1, \dots, \gamma_n$ starting at D^0 and ending at D^∞ satisfying (6.2) *is sufficient for existence of φ.*

EXAMPLE 6.5 (Illustrate (3.5a)). Suppose (c, d) in (3.5) is general. In what ways are there $\varphi : X_{c,d} \to \mathbb{P}^1_w$ so the cut description of φ has branch cycles $\boldsymbol{g}$ consisting of 3-cycles? For every degree $n \geq 4$ there are two distinct such presentations for general (c, d). The generalization of this special case for a general curve of genus g is an active research topic right now.

The relation of this one case with Riemann's use of half canonical classes appears in [**Fr95a**], [**Ser90a**] and [**Ser90b**]. [**Fr96**] shows there are exactly two

FIGURE 1. Cuts for a triangulation of X_c when $r = 3$



continuous families of such covers separated by the evenness and oddness of characteristics attached to the covers. [**AP03**], [**FMV03**] and [**SS02**] have different versions of showing the general genus g surface has such 3-cycle covers.

Further, for other types of branch cycle — a *Nielsen class* given by specifying the conjugacy classes of entries of $\boldsymbol{g}$ in $G(\boldsymbol{g})$ — there is no general method. When $r = 4$ and the genus is 1, the investigation of this question appears repeatedly in explicit versions of the abc-conjecture.

6.6. Higher genus θ s. As in the genus one case, we may take the logarithmic differential of the function φ, $d\varphi/\varphi$ we have a meromorphic differential on X with residue $\frac{1}{2\pi i}$ at the a_i s and residue - $\frac{1}{2\pi i}$ at the b_i s, and no other residues. Riemann's uniformization by a disk showed we don't actually need φ to construct such a differential: It always exists, manifesting that the sum of the residues is 0. Coordinates on $\mathbb{C}^g$ come from integrating $\mathcal{B}$ along paths going from an $x_0 \in X$ to some general points $x_1, \dots, x_g$, though to express a function on X we would integrate along from some point to x.

This is what we expect of φ where we imitate what happened in Abel's case. We want an odd function θ on $\mathbb{C}^g$ whose translates allow us to craft any $\varphi : X \to \mathbb{P}^1_w$. That would be tantamount to an expression like this:

$$\varphi(x) = \prod_{i=1}^{n} \theta(\int_{a_i}^{x} \mathcal{B}) / \prod_{i=1}^{n} \theta(\int_{b_i}^{x} \mathcal{B}). \tag{6.3}$$

In θ you see g coordinates; the *ith* entry is $\int_{a_i}^{x} \omega_i$. Each holds an integral over one basis element from $\mathcal{B}$. It is totally well-defined if the integrations are on a space that is to X as $\mathbb{C}$ was to $\mathbb{C}/L_{c,d}$ in Abel's problem.

Here you see the differential equation defining θ functions. In the expression for φ, replace the integrand in each $\int_{a_i}^{x} \mathcal{B}$ and $\int_{b_i}^{x} \mathcal{B}$ by a vector $\boldsymbol{w}$ in $\mathbb{C}^g$. Form the logarithmic differential of it: $d\theta(\boldsymbol{w})/\theta(\boldsymbol{w})$. Translations by periods will change it by addition of a constant. With ∇ the gradient in $\boldsymbol{w}$, $\nabla(\nabla\theta(\boldsymbol{w})/\theta(\boldsymbol{w}))$ is invariant under the lattice of periods.

7. Competition between the algebraic and analytic approaches

Many of these comments are from [**Ne81**]. Many who use Riemann's theory haveopinions on what Riemann might have been thinking. For one, many nontrivial issues going back a long way still arise. Galois has been the biggest victim of this kind of surmising. It is good to have a historian investigate the record.

Though there are other sources, too. For example, Alexander von Humboldt was good friend of Gauss'. His natural history volumes from his South American expeditions are rife with isothermal coordinates laid out on mountains and level curves of magnetic flow lines. Gauss was his consultant in the early 1800's, and so probably influenced natural history profoundly.

7.1. Riemann's early education. During his time in Berlin (1847–1849) P.G. Dirichlet (1805–1859), G. Eisenstein (1823–1852) and C.G.J. Jacobi (1804–1851) especially influenced Riemann. He attended Dirichlet's lectures on partial differential equations, and Eisenstein and Jacobi lectures on elliptic integrals. Riemann read Cauchy and Legendre on elliptic functions. [**Ne81**, p. 91]:

> Riemann was suitable, as no other German mathematician then was, to effect the first synthesis of the "French" and "German" approaches in general complex function theory.

His introductory lectures started with these topics: the Cauchy integral formulae; operations on infinite series; the Laurent series; and analytic continuation by power series. [**Ne81**, p. 92] includes a photocopy of a famous picture on analytic continuation from Riemann's own hand. Picard and Lefschetz both used this picture (from Riemann's collected works) in autobiographies of what influenced critical theorems of theirs. Riemann also lectured on the argument principle, the product represention of an entire fuction with arbitrarily prescribed zeros and the evaluation of definite integrals by residues. His most advanced lectures were from his published papers solving the Jacobi inversion problem.

7.2. Competition between Riemann and Weierstrass. [**Ne81**, p. 93]: K. Weierstrass (1815–1897) himself stressed above all the great influence of N.H. Abel (1802–1829) on him. At first Weierstrass was an unknown. Only after his 1856 paper on abelian functions did he get his position in Berlin. It was in 1856 that the competition between Riemann and Weierstrass became intense, around the solution of the Jacobi Inversion problem.

[**Ne81**, p. 93]: May 18 and July 2, 1857, Riemann submitted his two part solution to Jacobi's general inversion problem with these carefully measured words:

> Jacobi's inversion problem, which is settled here, has already been solved for the hyperelliptic integrals in several ways through the persistent and regally successful work of Weierstrass, of which a survey has been communicated in Vol. 47 of the Journ für Math. (p. 289). Until now, however, only a part of these investigations has been fully worked out and published (vol. 52, p. 285), namely the part that was outlined in §1 and §2 of the earlier paper and in the first half of §3, pertaining to elliptic functions. Only after the full publication of the promised paper shall we be

> able to tell to what extent the later parts of the presentation agree with my article not only in results but also in the methods leading to them.

Weierstrass consequently withdrew the 3rd installment of his investigations, which he had in the meantime finished and submitted to the Berlin Academy. He explained this (much later) in his collected works as follows.

> I withdrew [the 1857 manuscript] for, a few weeks later, Riemann published an article on the same topic, [...] on entirely different foundations from mine and did not make immediately clear that it agreed completely with mine in its results. The proof for it entailed some investigations of chiefly an algebraic nature, whose execution was not altotether easy for me ... But after this difficulty was overcome ...a thorough going overhaul of my paper was necessary. ... I could only toward the end of 1869 give to the solution of the general inversion problem that form in which I have treated it from then on in my lectures.

7.3. Soon after Riemann died. Publicly they seemed to have gotten along [**Ne81**, p. 95]. Professionally the mutual influence was unquestionably great. It would be entirely conceivable that the general systematic construction of the Weierstrassian function theory, achieved around 1860, could have been inspired by the works of Riemann perstaining to the same set of ideas.

[**Ne81**, p. 96]: After Riemann's death, Weierstrass attacked his methods quite often and even openly. July 14, 1870 was when he read his now famous critique on the Dirichlet Principle before the Royal Academy in Berlin. Weierstrass showed there did not always exist a function among those admitted [in variation problems] whose expression in question attained the lower bound, as Riemann had assumed. A letter to H. A. Schwarz on Oct. 3, 1875 says:

> The more I think about the principles of function theory, the firmer becomes my conviction this must be based on the foundation of algebraic truths, and that, consequently, it is not the right way if instead of building on simple and fundamental algebraic theorems, one appeals to the "transcendental" [by which Riemann has discovered so many of the most important properties of aglebraic functions].

During its heydey (1870–1890), the Weierstrassian school took over nearly every position in Germany. For instance, Schwarz was at Göttingen.

[**Ne81**, p. 98] asserts it was the Goursat part of Cauchy's theorem that renovated Riemann's approach, starting around 1900. [**Ah79**, p. 111] with no precise citation, refers to Goursat's contribution as,

> This beautiful proof, which could hardly be simpler is due to É. Goursat, who discovered that the classical hypothesis of a continuous $f'(z)$ is redundant.

Curiously, there is exactly one reference in *all* of [**Ah79**], a footnote:

> Without use of integration R. L. Plunkett proved the continuity of the derivative (BAMS **65**, 1959). E. H. Connell and P. Porcelli proved the existence of all derivatives (BAMS

> **67**, 1961). Both proofs lean on a topological theorem due to G. T. Whyburn [**Ah79**, p. 121].

That unique quote suggests Ahlfors supports the significance of Goursat to Riemann's renovation. Yet, there is a complication in analyzing Neuenschwanden's thesis. How would one document that this event turned mathematicians to the geometric/analytic view of Riemann? Historically it seems sensible to investigate the span from [**AG1895**] to [**Wey55**] as a shift from genus 1 to higher genus. Yet, that period is clearly insufficient to deal with an aspect of the true shift, from moduli of genus 1 curves (including modular curves) to general moduli. Theories toward the latter include *Teichmüller theory* (analytic) and *geometric invariant theory* (algebraic) or expedient precursors of the Hurwitz space approach like the Schiffer-Spencer deformation theories of varying the complex structure around a single point of a Riemann surface.

I suspect Goursat's theorem is a simple explanation that first year graduate students can follow. Likely, however, serious applications and resonant questions required understanding the variation of structures on a Riemann surfaces with the variation of the surface itself. My experiences are that not only do these issues confound graduate students, often specialists in complex variables struggle with these. Both technically and conceptually handling the hidden monodromy considerations (see [**Fr03**, Chap. 5]) is a tough topic. The only tool flexible enough to handle the complexity of the structure was that of Riemann. If that is right, then it is the documentation of these applications and questions that would illuminate on the story of the resurrection of Riemann's work. This makes it all look like slow continual progress. When, however, we come to Galois, the story has a different nature. We see it through modular curves which still to this day herald those works that accrue the most prestige.

8. Using Riemann to vary algebraic equations algebraically

8.1. The Picard components. There are three geometric ingredients in Riemann's theory: $J(\bar{X}_f)$, $\bar{X}_f$ and the zero (Θ) divisor of the function $\theta = \theta_{\bar{X}_f}$ (§8.3). The first identifies with divisor classes $\mathrm{Pic}^0(\bar{X}_f) = \mathrm{Pic}^0_f$ of degree 0 on $\bar{X}_f$. The second embeds naturally (algebraically) in Pic^1_f, divisor classes of degree 1 on $\bar{X}_f$. Then, Θ_f is the dimension $g-1$ variety of positive divisor classes in Pic^{g-1}_f.

Further, Pic^g_f interprets the Riemann-Roch Theorem and the *Jacobi Inversion Problem* geometrically. It takes its group structure from adding two positive divisors of degree g together modulo linear equivalence. Weil used this for an algebraic construction of Pic^0_f years after his thesis. His principle: The nearly well defined addition on positive divisors produced a unique complete algebraic group on the homogeneous space of divisor classes. Therefore Pic^0_f is almost the symmetric product of $\bar{X}_f$ taken g times. Riemann's theory was an inspiration to Weil's 1928 thesis (§9.1). Still, Weil was not certain until later that $\mathrm{Pic}^0(X)$ and X have the same field of definition. This reminds that what now looks obvious is the result of many mathematical stories.

8.2. Half-canonical classes. All Picard components Pic^k_f are pair wise analytically isomorphic. Yet, finding an isomorphism analytic in the Hurwitz space coordinates may require moving to a cover of the Hurwitz space.

Applying Riemann's theory directly requires having X and the Θ_X divisor on Pic^0_X. For example, suppose there is an analytic assignment of a divisor class of degree $g-1$ on each curve $\bar{X}_f$ in the Hurwitz family. Then, translation of Θ_X by this divisor class puts it in Pic^0_X. Here it would be available to construct the θ function. Convenient for this might be a *half-canonical* class: two times gives divisors for meromorphic differentials.

Riemann was even less algebraic in relating X and its Jacobian.

8.3. Relating $\tilde{X}$ and $\mathbb{C}^g$. The analytic isomorphism class of $\tilde{X}$ depends on the genus g of X. If $g=0$ it is the sphere, if $g=1$ it is $\mathbb{C}$ and it is *the upper half plane* $\mathbb{H}$ (or disk) if $g \geq 2$. As with $U_{\boldsymbol{z}}$, suppose we accept that $\tilde{X}$ is an analytic subspace of the Riemann sphere. Then, this comes from the Riemann mapping theorem. Still, it is not the uniformizing space we would expect. That would be $\tilde{X}^{\text{ab}}$, the quotient of $\tilde{X}$ by the subgroup of $\pi_1(X)$ generated by commutators. This is the maximal quotient of $\tilde{X}$ that is an abelian cover of X.

Mathematics rarely looks directly at $\tilde{X}^{\text{ab}}$. It embeds in $\mathbb{C}^g$. It is on $\mathbb{C}_g$ that θ_X lives with its zeros, the Θ divisor, meeting $\tilde{X}^{\text{ab}}$ transversally. Periods of differentials on X translate $\tilde{X}^{\text{ab}}$ into itself. Yet, it is sufficiently complicated there seems to be no device for picturing it.

There are two models for picturing this. A standard picture shows the complex structure on a complex torus (like the Jacobian). It is of a fundamental domain (parallelpiped) in $\mathbb{C}^g$. Then, $2g$ vectors representing generators of the lattice defining the complex torus give the sides of the parallelpiped. Inside this sits the pullback of X. The geometry for this picture uses geodesics (straight lines) from the flat (Euclidean) metric defining distances on the complex torus.

Assume the genus of X is at least 2. Then, the universal covering $\tilde{X}$ of X is the upper half plane $\tilde{X}$. A standard picture for X appears by grace of this having the structure of a negatively curved space. Geodesics here provide a polygonal outline of a set representing points of X. Since $\tilde{X} \to \tilde{X}^{\text{ab}}$ is unramified, $\tilde{X}^{\text{ab}}$ inherits a metric tensor with constant negative curvature. Yet, it sits snuggly in a flat space. Every finite abelian (unramified) cover Y of $\bar{X}_f$ is a quotient of $\tilde{X}^{\text{ab}}$; it is a minimal cover of $\tilde{X}$ with that property.

9. The impact of Riemann's Theorem

I list two big theorems that used theta functions in imaginative ways.

9.1. Siegel-Weil use of distributions. Look at (6.3) again. Weil's thesis constructed an analog of it: $(h(x)) \equiv \prod_{i=1}^{u} \theta^{w}_{x_i^0}(x) / \prod_{i=1}^{u} \theta^{w}_{x_i^\infty}(x)$. Here is its meaning. Both sides are fractional ideals in the ring of integers $\mathcal{O}_K$ of a number field K. The $\equiv$ sign means the left and right are equal up to a bounded fractional ideal. The left side is the principal fractional ideal that $h(x)$ generates. Most important, of course, are the functions $\theta^w_{x'}$: $x \mapsto \theta_{x'}(x)$ maps K points x into integral ideals. This function is defined only up to $\equiv$. Weil's distribution theorem allowed him (and Siegel [**Sie29**]) to perform diophantine magic.

[**Sie29**] starts with a crude set of reductions by going to a finite extension of K. Doing this point-by-point along a Hurwitz space would be a disaster. Canonical heights avoid this. Here is a related allusion to the odd half-canonical classes.

9.2. Mumford's coordinates on families of abelian varieties.

10. Final anecdote

Not long ago I was at Max Planck Institute to give talks on my favorite invention: Modular Towers. My wife and I have a German friend who lives in Bonn, who we knew in the US, a literature scholar. She is a very bright woman, full of energy. Yewt,that energy sometimes expresses a humanities approach to science that confounds scientists. Martina loves discussions on black holes, quantum mechanics, relativity theory, string theory. I'm sufficiently versed on these to often feel exasperated.

If you discuss the model for energy interchange in quantum mechanics, state spaces, spectra, such words will set Martina off with visions of other models she finds equally compelling. I tried to explain to her and her friend what this approach signals to physical scientists.

> By not understooding the constraints, you have diminished the significance of the one hard fought model that supports the evidence.

Often, and definitely at this place, she raised her vision of Einstein who in modern parlance "thought out of the box."

"Martina," I exhorted, "You see Einstein as totally without precedent. You suggest, we should follow humanities people out of our dull intellectual ruts." I explained that Einstein was far from without precedent; that we mathematicians had geniuses with at least his imaginative. My example was Riemann: I called him the man who formed the equation that gave Einstein his scalar curvature criterion gravity. "Mike," she said, "You're just making that up! Who is Riemann?"

In the land that has a Zehn Deutsche Mark with Gauss' picture on it, a woman of intellectual energy and eclectic temperament, who hadn't heard of Riemann! I made her an offer.

Would she change her mind in future discussions if I could convince her Riemann was at least the ilk of Einstein? "How can you do that?"was her retort. My response: "The evidence is in this room."

Books lined her walls. I was certain I could find something among the philosophy books – though they were in German – that would serve my purpose. She took the challenge, and I, expecting a long haul, went about my work. To start I took the *R* book in her encyclopedia series. Opening to Riemann, I found this in the first paragraph: Bernhard Riemann was one of the most profound geniuses of modern times. Notable among his discoveries were the equations that Einstein later applied to general relativity theory.

Consider the context: Riemann spent very little of his life on Riemannian geometry and much of his life on theta functions and their consequences. This paper attempts to explain that his efforts inspired many and have still yet to completely fulfill their significance. We mathematicians write equations revealing what our senses might never directly feel. That our senses feel it indirectly is our constraint.

Riemann's based his subject on harmonic functions, deriving from the distribution of electronic charge along a capacitor. You couldn't see electrons then. You can't see them too well now. Only with quantum mechanics had we a serious model that produced electrons. Yet, we believe they exist as surely do computers and light bulbs. Riemann produced equations based on spin long before the existance of half-spin comparable to that of Fermions (in electrons) long before Stern-Gerlach experiment.

Many of us love what Riemann did, and we wish more of the world knew of his ineffable contributions.

References

[AG1895] P. Appell and É. Goursat, *Théorie des Fonctions Algébriques et de leurs Intégrales*, Gauthier-Villars et C[ie], Paris, 1895, with a revised edition by P. Fatou, 1929.

[Ah79] L. Ahlfors, *Introduction to the Theory of Analytic Functions of One*, 3rd edition, Inter. Series in Pure and Applied Math., McGraw-Hill Complex Variable, 1979.

[AP03] M. Artebani and G.P. Pirola, *Algebraic functions with even monodromy*, preprint Feb. 2003.

[BFr02] P. Bailey and M. Fried, *Hurwitz monodromy, spin separation and higher levels of a Modular Tower*, in Proceedings of Symposia in Pure Mathematics **70** (2002) editors M. Fried and Y. Ihara, 1999 von Neumann Conference on Arithmetic Fundamental Groups and Noncommutative Algebra, August 16-27, 1999 MSRI, 79–221.

[Ben91] D.J. Benson, *I: Basic representation theory of finite groups and associative algebras*, Cambridge Studies in advanced math., vol. 30, Cambridge U. Press, Cambridge, 1991.

[DF99] P. Dèbes and M. Fried, *Integral specialization of families of rational functions*, PJM **190** (1999) 45–85 (typos corrected at www.math.uci/˜mfried/#math).

[FaK01] H. Farkas and I. Kra, *Theta Constants, Riemann Surfaces and the Modular Group*, AMS graduate text series **37**, 2001.

[Fay73] J. Fay, *Theta Functions on Riemann Surfaces*, Lecture notes in Mathematics **352**, Springer Verlag, Heidelberg, 1973.

[FMV03] G. Frey, K. Maagard and H. Völklein, *The monodromy group of a function on a general curve*, preprint as of 10/24/02.

[Fri77] M. Fried, *Fields of definition of function fields and Hurwitz families and groups as Galois groups*, Communications in Algebra **5** (1977), 17–82.

[Fri78] M. Fried, *Galois groups and Complex Multiplication*, Trans.A.M.S. **235** (1978) 141–162.

[FJ86] M. Fried and M. Jarden, *Field arithmetic*, Ergebnisse der Mathematik III, vol. 11, Springer Verlag, Heidelberg, 1986.

[FV91] M. Fried and H. Völklein, *The inverse Galois problem and rational points on moduli spaces* , Math. Annalen **290** (1991), 771–800.

[FKK01] M. Fried, E. Klassen, Y. Kopeliovic, *Realizing alternating groups as monodromy groups of genus one covers*, PAMS **129** (2000), 111–119.

[Fr96] M. Fried, *Alternating groups and lifting invariants*, Preprint 07/01/96 (1996), 1–34.

[Fr95a] M. Fried, *Introduction to Modular Towers: Generalizing the relation between dihedral groups and modular curves*, Proceedings AMS-NSF Summer Conf., **186**, 1995, Cont. Math series, Recent Developments in the Inverse Galois Problem, pp. 111–171.

[Fr02] M. Fried, *Moduli of relatively nilpotent extensions*, Institute of Mathematical Science Analysis **1267**, June 2002, Comm. in Arithmetic Fundamental Groups, 70–94.

[Fr03] M. Fried, *Riemann's Existence Theorem: Elementary approach to moduli*, www.math.uci.edu/ mfried/#ret, especially Chaps. 2 and 4. Chap. 1 is an overview of the whole five chapters.

[H49] M. Heins, *The conformal mapping of simply-connected Riemann surfaces*, Annals of Math. **50** (1949), 686–690.

[Ne81] E. Neuenschwanden, *Studies in the history of complex function theory II: Interactions among the French school, Riemann and Weierstrass*, BAMS **5** (1981), 87–105.

[SS02] S. Schröer, *Curves With Only Triple Ramification*, preprint arXiv:math.AG/0206091 10 June 2002.

[Ser90a] J.-P. Serre, *Relêvements dans* $\tilde{A}_n$, C. R. Acad. Sci. Paris **311** (1990), 477–482.

[Ser84] J.-P. Serre, *L'invariant de Witt del la forme* $\mathrm{Tr}(x^2)$, Comm. Math. Helv. **59** (1984), 651–676.

[Ser90b] J.-P. Serre, *Revêtements a ramification impaire et thêta-caractéristiques*, C. R. Acad. Sci. Paris **311** (1990), 547–552.

[Sie29] C.L. Siegel, *Über einige Anwendungen diophantischer Approximationen*, Abh. Pr. Akad. Wiss. **1** (1929), 41–69.

[Spr57] G. Springer, *Introduction to Riemann Surfaces*, 1st ed., Addison-Wesley Mathematics Series, vol. 11, Addison-Wesley, Reading, Mass., 1957.

[Vö96] H. Völklein, *Groups as Galois Groups* **53**, Cambridge Studies in Advanced Mathematics, Camb. U. Press, Camb. England, 1996.

[Wey55] H. Weyl, Die Idee der Riemannschen Fläche, Third Edition, B. G. Teubner, Stuttgart, 1955.

[Wo64] K. Wohlfahrt, *An extension of F. Klein's level concept*, Ill. J. Math. **8** (1964), 529–535.

UC Irvine, Irvine, CA 92697, USA
E-mail address: `mfried@math.uci.edu`