We introduce Riemann surfaces and branched covers of the Riemann sphere, and describe the relationship to topology and group theory.

1. Complex manifolds

A (real) manifold of dimension \( n \) is a (Hausdorff, second countable) space which is locally homeomorphic to an open subset of \( \mathbb{R}^n \). If we wish to make a definition of a complex manifold, we could replace \( \mathbb{R}^n \) by \( \mathbb{C}^n \), but then we see that we have simply given the definition of an even-dimensional real manifold. As with real differentiable manifolds, the key is to add some additional structure via a well-behaved atlas.

Definition 1.1. Given a topological space \( X \), a chart for \( X \) is a homeomorphism \( \varphi : V \to U \), where \( V \) is an open subset of \( \mathbb{C}^n \) for some \( n \), and \( U \) is an open subset of \( X \). A holomorphic atlas for \( X \) is a collection of charts \( \varphi_i : V_i \to U_i \), with the \( V_i \subseteq \mathbb{C}^n \) for some fixed \( n \), such that:

1. the \( U_i \) cover \( X \);
2. for each \( i \neq j \), the transition map

\[
\varphi_{i,j} : \varphi_i^{-1}(U_i \cap U_j) \overset{\varphi_i}{\to} U_i \cap U_j \overset{\varphi_j^{-1}}{\to} \varphi_j^{-1}(U_i \cap U_j)
\]

is holomorphic.

A complex manifold of dimension \( n \) is a (Hausdorff, second countable) topological space \( X \) together with a holomorphic atlas.

Proposition 1.2. A complex manifold \( X \) is naturally oriented when considered as a topological manifold.

Sketch of proof. The complex numbers have a natural orientation, which induces one on \( \mathbb{C}^n \). Because holomorphic maps preserve this orientation, we get an induced orientation on any complex manifold. \( \square \)

Using the atlas and the fact that we have a definition of holomorphic functions on open subsets of \( \mathbb{C}^n \), we can define holomorphic functions on complex manifolds.

Definition 1.3. Let \( X \) be a complex manifold with atlas \( \{ \varphi_i : V_i \to U_i \} \), and \( U \subseteq X \) open. A function \( f : U \to \mathbb{C} \) is holomorphic if for each \( i \), the composed function

\[
f \circ \varphi_i : \varphi_i^{-1}(U \cap U_i) \to \mathbb{C}
\]

is holomorphic.

This definition makes sense whether or not we require the transition maps \( \varphi_{i,j} \) to be holomorphic. However, the additional hypothesis implies that we don’t need to look at more than one \( U_i \) if \( U \subseteq U_i \).

Exercise 1.4. If \( U \subseteq U_i \) for some \( i \), then \( f : U \to \mathbb{C} \) is holomorphic if and only if \( f \circ \varphi_i : \varphi_i^{-1}(U) \to \mathbb{C} \) is holomorphic.

Using the notion of holomorphic functions, we can then define a holomorphic map between two complex manifolds.
Definition 1.5. Given two holomorphic manifolds $X, X'$, a continuous map $\varphi : X \to X'$ is **holomorphic** if for all $U \subseteq X'$ open, and all $f : U \to \mathbb{C}$ holomorphic, the composed function

$$f \circ \varphi : \varphi^{-1}(U) \to \mathbb{C}$$

is also holomorphic.

This is equivalent to a definition which uses the atlas explicitly:

**Exercise 1.6.** Given holomorphic manifolds $X, X'$ with atlas $\{\varphi_i : V_i \to U_i\}$ and $\{\varphi'_i : V'_i \to U'_i\}$, a map $\varphi : X \to X'$ is holomorphic if and only if for all $i, j$ the composed map

$$(\varphi'_j)^{-1} \circ \varphi \circ \varphi_i : \varphi^{-1}(U'_j) \cap U_i \to V'_j$$

is holomorphic.

We also have:

**Exercise 1.7.** Given a complex manifold $X$, and $U \subseteq X$ open, a function $f : U \to \mathbb{C}$ is holomorphic if and only if it gives a holomorphic mapping when $\mathbb{C}$ is considered as a complex manifold via the trivial atlas.

**Definition 1.8.** A **Riemann surface** is a complex manifold of dimension 1.

**Example 1.9.** The Riemann sphere $\mathbb{CP}^1$ is a compact Riemann surface with underlying topological space $S^2$. If we realize as $S^2$ as the unit sphere in $\mathbb{R}^3$, let $U_1$ be the complement of the “north pole” $(0, 0, 1)$ and $U_2$ the complement of the “south pole” $(0, 0, -1)$. Both $V_1$ and $V_2$ are the complex plane. We then define the $\varphi_i$ as follows:

$$\varphi_1(x + iy) = \frac{1}{x^2 + y^2 + 1}(2x, 2y, x^2 + y^2 - 1)$$

and

$$\varphi_2(x + iy) = \frac{1}{x^2 + y^2 + 1}(2x, -2y, 1 - x^2 - y^2).$$

Then we compute

$$(\varphi_2^{-1} \circ \varphi_1)(x + iy) = \frac{x - iy}{x^2 + y^2} = \frac{1}{x + iy},$$

which is holomorphic, and similarly for $\varphi_1^{-1} \circ \varphi_2$. (Note that $\varphi_1^{-1}$ is stereographic projection from the north pole, and $\varphi_2^{-1}$ is stereographic projection from the south pole composed with complex conjugation)

**Example 1.10.** We often picture complex manifolds not in terms of a topological space with an atlas, but in terms of gluing together the pieces $V_i$ along biholomorphisms of open subsets. For instance, we picture $\mathbb{CP}^1$ as obtained from two copies of $\mathbb{C}$ by gluing along the open subsets $\mathbb{C} \setminus \{0\}$ via the map $1 \mapsto 1/z$. However, when working this way, one has to be careful about the Hausdorff condition.

For instance, if we instead glue two copies of $\mathbb{C}$ along the open subsets $\mathbb{C} \setminus \{0\}$ via the identity map, we obtain the “plane with a doubled origin”, which is locally biholomorphic to open subsets of $\mathbb{C}$, but is not Hausdorff, as the two origins do not have disjoint open neighborhoods.

2. Branched covers

Although branched covers are interesting more generally, we will restrict our attention to branched covers of the Riemann sphere, which already have a rich and elegant structure.

**Definition 2.1.** A **branched cover** (of the Riemann sphere) is a pair $(C, f)$ where $C$ is a compact, connected Riemann surface, and $f : C \to \mathbb{CP}^1$ is a non-constant holomorphic map.
Definition 2.2. We say two branched covers \((C, f)\) and \((C', f')\) are **equivalent** if there is a biholomorphism \(g : C \to C'\) such that \(f = f' \circ g\).

The following is a basic fact from complex analysis:

**Proposition 2.3.** Suppose \((C, f)\) is a branched cover, and \(P \in C\). Then there are open neighborhoods \(U, U'\) of \(P\) and \(f(P)\) respectively, and open neighborhoods \(V, V'\) of 0 in \(C\), and biholomorphisms \(g : U \to V, g' : U' \to V'\) sending \(P\) and \(f(P)\) respectively to 0, such that the map \(g' \circ f \circ g^{-1} : V \to V'\) is equal to \(z \mapsto z^{e_P}\) for some \(e_P \geq 1\).

**Definition 2.4.** If the \(e_P\) from Proposition 2.3 is strictly greater than 1, we say that \(P\) is a **ramification point** of \(f\), with **ramification index** \(e_P\). In this case, \(f(P)\) is a **branch point** of \(f\). A neighborhood \(U\) of \(P\) such that there exist \(U', V, V', g, g'\) as in Proposition 2.3 is a **standard neighborhood** of \(P\).

It is clear that any neighborhood contained in a standard neighborhood is again a standard neighborhood.

**Corollary 2.5.** Let \((C, f)\) be a branched cover. Then there are only finitely many ramification and branch points of \(f\).

**Proof.** Given a neighborhood \(U\) of a point \(P \in C\) as in Proposition 2.3, we see that on \(U\), the map \(f\) is locally one-to-one everywhere except possibly at \(P\), where it is locally \(e_P\)-to-one. By compactness of \(C\), there is a finite open cover of \(C\) by such neighborhoods, so we conclude that \(f\) is locally one-to-one at all but finitely many points, and hence can be ramified at only finitely many points. The statement on branch points follows. \(\square\)

A stronger version of Proposition 2.3, which makes use of the compactness of the cover, is:

**Lemma 2.6.** Given a branched cover \((C, f)\), and \(Q \in \mathbb{CP}^1\), the fiber \(f^{-1}(Q)\) is finite, and there exists a connected neighborhood \(U'\) of \(Q\) such that every connected component of \(f^{-1}(U')\) contains a unique point \(P \in f^{-1}(Q)\), surjects onto \(U'\), and is a standard neighborhood of \(P\).

**Proof.** \(C\) is covered by finitely many standard neighborhoods by compactness, and on every standard neighborhood, \(f\) is finite-to-one, so we conclude the finiteness of \(f^{-1}(Q)\).

We next observe that \(f\) is a closed map, because \(C\) is compact and \(\mathbb{CP}^1\) is Hausdorff, and also an open map, by the open mapping theorem in complex analysis (or more concretely, because the map \(z \mapsto z^{e_P}\) is visibly open for any \(e\)). Then for any open subset \(U' \subseteq \mathbb{CP}^1\), the induced map \(f^{-1}(U') \to U'\) is also open and closed. It follows that if \(W \subseteq f^{-1}(U')\) is a connected component, its image is open and closed in \(U'\), and hence is all of \(U'\) if \(U'\) is connected.

Let \(P_1, \ldots, P_m\) be the preimages of \(Q\), and \(U_1, \ldots, U_m\) be standard neighborhoods of \(P_1, \ldots, P_m\). We may further suppose that they are chosen to be disjoint from one another. Each boundary \(\partial U_i\) is then a closed set with \(f(\partial U_i)\) closed and not containing \(Q\). We can then choose \(U'\) a connected neighborhood of \(Q\) disjoint from all the \(f(\partial U_i)\). Let \(U\) be a connected component of \(f^{-1}(U')\); by the above, \(U\) surjects onto \(U'\), so contains (at least) one of the \(P_i\). To see that \(U\) is a standard neighborhood and does not contain \(P_j\) for \(j \neq i\), it is enough to prove \(U \subseteq U_i\). We have \(U_i \cap f^{-1}(U')\) and \(\overline{U_i} \cap f^{-1}(U')\) open and closed respectively in \(f^{-1}(U')\) by definition. But \(\partial U_i \cap f^{-1}(U') = \emptyset\) by construction, so \(U_i \cap f^{-1}(U') = \overline{U_i} \cap f^{-1}(U')\) is also closed. Since \(U \cap U_i \neq \emptyset\) and \(U\) is connected, we conclude \(U \subseteq U_i\), as desired. \(\square\)

We now conclude:

**Corollary 2.7.** Let \((C, f)\) be a branched cover. Then \(f\) induces a (finite-degree) topological covering map after removing the branch points and their preimages.
Proof. After removing the ramification points, it is clear that $f$ is a local homeomorphism, and more specifically, a homeomorphism onto its image for every standard neighborhood. The corollary thus follows immediately from Lemma 2.6.

**Definition 2.8.** The **degree** of a branched cover is the degree of the induced covering map after removing the branch points and their preimages.

We also conclude the following result, which says that we can think of a branched cover as having degree $d$ even over the branch points, if we count with appropriate multiplicities.

**Corollary 2.9.** If $(C, f)$ is a branched cover of degree $d$, then for any $Q \in \mathbb{CP}^1$, we have

$$\sum_{P \in f^{-1}(Q)} e_P = d.$$  

Proof. Take $Q' \neq Q$ a point which lies in the neighborhood $U'$ provided by Lemma 2.6 (this is then necessarily not a branch point). Given $P \in f^{-1}(Q)$, if $U$ is the connected component of $f^{-1}(U')$ containing $P$, then Proposition 2.3 implies that $Q'$ has $e_P$ preimages in $U$. We conclude that the number of points in $f^{-1}(Q')$ is equal to $\sum_{P \in f^{-1}(Q)} e_P$. On the other hand, by Proposition 2.7, the number of points in $f^{-1}(Q')$ is equal to $d$, giving the desired identity.

This motivates the following definition:

**Definition 2.10.** Given a branched cover $(C, f)$ of degree $d$, and $Q$ a branch point, the **branch type** of $f$ at $Q$ is the partition of $d$ given by $\{e_P : P \in f^{-1}(Q)\}$.

Thus, a branched cover comes with the combinatorial data of its degree $d$ and a tuple of partitions of $d$ determined by its branching. We often assume that the branch points are marked in an ordered fashion $Q_1, \ldots, Q_r$; in this case, we can place the combinatorial data into a tuple $(d; T_1, \ldots, T_r)$ where each $T_i$ is the branch type at $Q_i$; this is called the **type** of the cover $(C, f)$.

We now return to the discussion of the topological covering map induced by $(C, f)$, according to Corollary 2.7. Note that removing a finite set of points from a connected surface leaves it connected, so the covering space coming from Corollary 2.7 is also connected. Remarkably, we have the following converse result, known as the Riemann existence theorem:

**Theorem 2.11** (Riemann existence). Given distinct points $Q_1, \ldots, Q_r \in \mathbb{CP}^1$, write $\hat{\mathbb{CP}}^1 := \mathbb{CP}^1 \setminus \{Q_1, \ldots, Q_r\}$, and choose also a point $Q \in \hat{\mathbb{CP}}^1$. Suppose we have a branched cover $(C, f)$ branched only at the $Q_i$, with branch
type \( T_i \) at each \( Q_i \). Let \( \hat{\gamma} : \hat{C} \to \hat{\mathbb{C}P^1} \) be the topological covering space of degree \( d \) we obtain from Corollary 2.7. Then if \( \gamma \) is a loop in \( \hat{\mathbb{C}P^1} \) based at \( Q \), and \( P \in f^{-1}(Q) \), it is a standard result of topology that there is a unique lift \( \tilde{\gamma}_P \) of \( \gamma \) to a path in \( \hat{C} \) which starts at \( P \). Then \( \tilde{\gamma}_P \) necessarily ends in \( f^{-1}(Q) \), but not necessarily at \( P \). Denote the endpoint by \( \mu(\gamma)(P) \). Thus, we have a function

\[
\mu(\gamma) : f^{-1}(P) \to f^{-1}(P).
\]

We see immediately that this is invertible, since if \( \gamma^{-1} \) is the same loop as \( \gamma \) with its direction reversed, we clearly have \( \mu(\gamma^{-1}) = \mu(\gamma)^{-1} \). Thus, \( \mu(\gamma) \) is a permutation of \( f^{-1}(Q) \). This is invariant under homotopy, and also compatible with composition. If we choose a labeling of \( f^{-1}(Q) \), we get an isomorphism \( \text{Sym}(f^{-1}(Q)) \sim S_d \), so we conclude:

**Proposition 3.1.** Given a branched cover \((C,f)\) with branch points \( Q_i \), and a labeling of \( f^{-1}(Q) \), then lifting of loops based at \( Q \) induces a homomorphism

\[
\mu : \pi_1(\hat{\mathbb{C}P^1},Q) \to S_d.
\]

**Definition 3.2.** The homomorphism \( \mu \) is the **monodromy map**, and \( \mu(\gamma) \) is the **monodromy** of \((C,f)\) around \( \gamma \).

The next step involves a good understanding of the fundamental group of \( \hat{\mathbb{C}P^1} \). A standard calculation from topology gives the following:

**Proposition 3.3.** There exist loops \( \gamma_1, \ldots, \gamma_r \subseteq \hat{\mathbb{C}P^1} \) based at \( Q \), satisfying:

(i) The \( \gamma_i \) generate \( \pi_1(\hat{\mathbb{C}P^1},Q) \).

(ii) The only relation among the \( \gamma_i \) in \( \pi_1(\hat{\mathbb{C}P^1},Q) \) is that \( \gamma_1 \cdots \gamma_r = 1 \).

(iii) Each \( \gamma_i \) is homotopic to a small loop around \( Q_i \).

Note that (ii) implies that \( \pi_1(\hat{\mathbb{C}P^1},Q) \) is the free group generated by any \( r - 1 \) of the \( \gamma_i \).

**Remark 3.4.** Choice of such \( \gamma_i \) is not unique. Indeed, this non-uniqueness will be very important to us when we discuss braid actions and connected components of Hurwitz spaces.

We now suppose we have fixed \( Q_1, \ldots, Q_r \) and \( Q \), but not necessarily any branched cover.

**Proposition 3.5.** Suppose \((C,f)\) is a branched cover with branch points \( Q_1, \ldots, Q_r \), and branch type \( T_i \) at \( Q_i \) for \( i = 1, \ldots, r \), and with a labeling of \( f^{-1}(Q) \). Set \( \sigma_i = \mu(\gamma_i) \in S_d \) for \( i = 1, \ldots, r \). Then we have:

(i) The subgroup of \( S_d \) generated by the \( \sigma_i \) is transitive.

(ii) \( \sigma_1 \cdots \sigma_r = 1 \).

(iii) The cycle decomposition of \( \sigma_i \) agrees with \( T_i \), for \( i = 1, \ldots, r \).

Recall that a subgroup \( G \subseteq S_d \) is transitive if for all \( i, j \in \{1, \ldots, d\} \), there is some \( \sigma \in G \) with \( \sigma(i) = j \). For \( \sigma \in S_d \), the decomposition of \( \sigma \) into disjoint cycles yields a partition of \( d \) by considering the lengths of the cycles (if we consider all fixed points of \( \sigma \) to lie in cycles of length 1), and thus it makes sense to compare to the \( T_i \).

**Proof.** (i) By Proposition 3.5 (i), this is equivalent to showing that \( \mu(\pi_1(\hat{\mathbb{C}P^1},Q)) \) is transitive. This follows from the connectedness of \( C \): first observe that \( \hat{C} := C \setminus (\bigcup_i f^{-1}(Q_i)) \) is still connected, and indeed path connected, so given \( P, P' \in f^{-1}(Q) \), there is a path \( \hat{\gamma} \) in \( \hat{C} \) connecting them. Then \( f(\hat{\gamma}) \) is a loop in \( \hat{\mathbb{C}P^1} \) based at \( Q \), and by definition \( \hat{\gamma} \) is its unique lift starting at \( P \). Thus \( \mu(f(\hat{\gamma})) \) sends \( P \) to \( P' \), and since \( P, P' \) were arbitrary in \( f^{-1}(Q) \), we conclude the desired transitivity.
(ii) This follows immediately from Proposition 3.5 (ii), and the fact that $\mu$ is a homomorphism.

(iii) By Proposition 3.5 (iii), it is enough to consider the case that $Q$ lies on such a small loop, say $\gamma$. By Lemma 2.6, we see that $f^{-1}(\gamma)$ is a disjoint union of paths, each contained in a standard neighborhood of some $P \in f^{-1}(Q_i)$. Fix such a standard neighborhood, so that we are simply looking at lifts of a loop around 0 under the map $z \mapsto z^e P$. Now, $Q$ has $e P$ preimages under this map, differing by powers of the $e P$th root of unity $\zeta_{e P} := e^{2\pi i / e P}$, and we see that a lift of $\gamma$ starting at one point $P'$ will end at $\zeta_{e P} P'$. Thus $\mu(\gamma)$ cyclically permutes the $e P$ points lying over $Q$ in the standard neighborhood of $P$, and the desired statement follows. 

This motivates:

**Definition 3.6.** A tuple $(\sigma_1, \ldots, \sigma_r) \in (S_d)^r$ is a **Hurwitz factorization** of type $(d; T_1, \ldots, T_r)$ if it satisfies (i)-(iii) of Proposition 3.5.

Thus, the proposition says that a branched cover of type $(d; T_1, \ldots, T_r)$, together with a labeling of $f^{-1}(Q)$, yields a Hurwitz factorization of the same type. What happens if we choose a different labeling? We change all the $\sigma_i$ by a fixed relabeling of $\{1, \ldots, d\}$, or equivalently, conjugate them all by a fixed element of $S_d$.

**Definition 3.7.** Two Hurwitz factorizations $(\sigma_1, \ldots, \sigma_r)$ and $(\sigma'_1, \ldots, \sigma'_r)$ are **equivalent** if there exists a relabeling of $\{1, \ldots, d\}$ sending $\sigma_i$ to $\sigma'_i$ for all $i$, or equivalently, if there exists $\tau \in S_d$ such that $\sigma_i = \tau \sigma'_i \tau^{-1}$ for all $i$.

We thus have that a branched cover yields a well-defined equivalence class of Hurwitz factorizations. On the other hand, given a Hurwitz factorization, by Proposition 3.5 (i) and (ii) we obtain a homomorphism $\pi_1(\mathbb{C}P^1, Q) \to S_d$ with transitive image, and a basic theorem of topology relating the fundamental group to covering maps then implies that we obtain a corresponding topological covering map $\hat{C} \to \mathbb{C}P^1$ of degree $d$, with $\hat{C}$ connected. Applying the Riemann existence theorem, we conclude:

**Theorem 3.8.** The monodromy map induces a bijection between equivalence classes of branched covers of type $(d; T_1, \ldots, T_r)$ with branch points $Q_1, \ldots, Q_r$, and equivalence classes of Hurwitz factorizations of the same type.

### 4. Hurwitz Theory

We immediately conclude the following from Theorem 3.8:

**Corollary 4.1.** If we fix $Q_1, \ldots, Q_r \in \mathbb{C}P^1$, and a type $(d; T_1, \ldots, T_r)$, then there are only finitely many equivalence classes of branched covers of type $(d; T_1, \ldots, T_r)$ with branch points $Q_1, \ldots, Q_r$. Moreover, this number does not depend on the $Q_i$.

This leads us to the following:

**Definition 4.2.** The **Hurwitz number** $h(d; T_1, \ldots, T_r)$ is the number of equivalence classes of branched covers of type $(d; T_1, \ldots, T_r)$ with branched points $Q_1, \ldots, Q_r$, for any choice of distinct $Q_i \in \mathbb{C}P^1$. Equivalently, $h(d; T_1, \ldots, T_r)$ is the number of equivalence classes of Hurwitz factorizations of type $(d; T_1, \ldots, T_r)$.

**Remark 4.3.** We have given equivalent characterizations of branched covers in terms of complex geometry, topology, and group theory. One can also give an equivalent characterization in terms of algebraic geometry, as Chow’s theorem states that any compact Riemann surface has a (unique) structure of a projective nonsingular curve, and holomorphic maps correspond to algebraic morphisms. Using this definition, we can generalize branched covers from the complex numbers to any field.
From this point of view, independence of the choice of $Q_i$ is not at all a trivial fact: in fact, there are simple examples demonstrating that the same statement fails if we work instead over fields of positive characteristic.

**Remark 4.4.** In fact, the Hurwitz number is usually defined slightly differently: we use a weighted count of branched covers, dividing by the size of the automorphism group of each cover. The automorphisms of a cover correspond to the relabelings that leave the corresponding Hurwitz factorization unchanged, so group-theoretically, this corresponds to the number of Hurwitz factorizations divided by $d!$. We use the above definition because its motivation is clearer, but it turns out the standard definition is in many ways more natural. In any case, in many interesting cases there will not be any automorphisms of the covers in question, so the two definitions will agree.

Now we consider what happens if we allow the $Q_i$ to vary. Given a choice of $\gamma = (\gamma_1, \ldots, \gamma_r)$ satisfying the conditions of Proposition 3.5, let $\mathcal{U} = (U_1, \ldots, U_r)$ be a tuple of small open disks centered at $Q_1, \ldots, Q_r$ respectively, and with $U_i$ contained in the interior of $\gamma_i$ for each $i$. Then for any choices of $Q'_i \in U_i$, we still have $\gamma_i$ homotopic to a small loop around $Q'_i$. Denote by $\Delta_{\gamma, \mathcal{U}}(C, f)$ the set of all branched covers $(C, f')$ with labeled branch points $Q'_i \in U_i$, and such that the equivalence classes of Hurwitz factorizations associated to each by $\gamma$ are the same.

We can then define:

**Definition 4.5.** Given a type $(d; T_1, \ldots, T_r)$, let the Hurwitz space $\mathcal{H}(d; T_1, \ldots, T_r)$ of type $(d; T_1, \ldots, T_r)$ be the set of branched covers $(C, f)$ with labeled branch points and type $(d; T_1, \ldots, T_r)$, equipped with the topology with base consisting of the sets $\Delta_{\gamma, \mathcal{U}}(C, f)$ as $\gamma, \mathcal{U}$ and $(C, f)$ vary over all possibilities.

We have:

**Proposition 4.6.** The topology on $\mathcal{H}(d; T_1, \ldots, T_r)$ is well defined; that is, the $\Delta_{\gamma, \mathcal{U}}(C, f)$ satisfy the conditions for a base of a topology.

In order to examine this topology further, let $\mathcal{U}_r \subseteq (\mathbb{CP}^1)^r$ denote the open subset consisting of $r$-tuples of distinct points in $\mathbb{CP}^1$; thus, a choice of $(Q_1, \ldots, Q_r)$ corresponds to a point of $\mathcal{U}_r$. If we have a branched cover together with a labeling of its branch points, we obtain a point of $\mathcal{U}_r$, so we have a map $\mathcal{H}(d; T_1, \ldots, T_r) \to \mathcal{U}_r$.

We then have the following consequence of the Riemann existence theorem:

**Proposition 4.7.** The natural map

$$\mathcal{H}(d; T_1, \ldots, T_r) \to \mathcal{U}_r$$

is a topological covering map, of degree equal to $h(d; T_1, \ldots, T_r)$.

**Proof.** It is clear that for any $Q_1, \ldots, Q_r$, a tuple of open neighborhoods $\mathcal{U} = (U_1, \ldots, U_r)$ gives an open neighborhood of $(Q_1, \ldots, Q_r)$ inside $\mathcal{U}_r$. Moreover, if we fix any choice of $\gamma = (\gamma_1, \ldots, \gamma_r)$ as above, we see immediately that the preimage of $\mathcal{U}$ in $\mathcal{H}(d; T_1, \ldots, T_r)$ is, as a set, the disjoint union over all $(C, f)$ with branch points $Q_i$ of the sets $\Delta_{\gamma, \mathcal{U}}(C, f)$. It is moreover clear that each $\Delta_{\gamma, \mathcal{U}}(C, f)$ maps bijectively to $\mathcal{U} \subseteq \mathcal{U}_r$, so it is enough to see that $\Delta_{\gamma, \mathcal{U}}(C, f)$ maps homeomorphically to $\mathcal{U}$, and each $\Delta_{\gamma, \mathcal{U}}(C, f)$ is in fact a connected component of the preimage of $\mathcal{U}$. Since $\mathcal{U}$ is connected, the latter follows from the former together with the observation that each $\Delta_{\gamma, \mathcal{U}}(C, f)$, being an element of a base for the topology, is open in $\mathcal{H}(d; T_1, \ldots, T_r)$.

We thus prove that $\Delta_{\gamma, \mathcal{U}}(C, f)$ maps homeomorphically to $\mathcal{U}$. If we consider all $\mathcal{U}' = (U'_1, \ldots, U'_r)$ with each $U'_i$ an open disk contained in $U_i$, then the $\mathcal{U}'$ give a base for the topology of $\mathcal{U}$, and their preimages in $\Delta_{\gamma, \mathcal{U}}(C, f)$ are precisely $\Delta_{\gamma, \mathcal{U}'}(C, f)$, so we conclude that the map $\Delta_{\gamma, \mathcal{U}'}(C, f)$ is continuous. The argument that it is open involves comparing different choices of $\gamma$, and is similar to the proof of Proposition 4.6.
Finally, the degree is by definition given by \( h(d; T_1,\ldots,T_r). \)

Since \( \mathcal{U}_r \) is naturally a complex manifold of (complex) dimension \( r \), we immediately conclude:

**Corollary 4.8.** The Hurwitz space \( \mathcal{H}(d; T_1,\ldots,T_r) \) is a (complex) manifold of dimension \( r \).

A basic question one can ask is the following:

**Question 4.9.** How can one characterize the connected components of \( \mathcal{H}(d; T_1,\ldots,T_r) \)? In particular, for which \((d; T_1,\ldots,T_r)\) is \( \mathcal{H}(d; T_1,\ldots,T_r) \) connected?

Just as with the Hurwitz number, the connected components of Hurwitz spaces can be interpreted purely in terms of group theory. In order to give this interpretation, we observe that there is a simple way to move from one Hurwitz factorization to others.

**Definition 4.10.** Given a Hurwitz factorization \( (\sigma_1,\ldots,\sigma_r) \) of type \((d; T_1,\ldots,T_r)\), and \( i \) with \( 1 \leq i \leq r-1 \), let \( B_i(\sigma_1,\ldots,\sigma_r) \) be the Hurwitz factorization of type \((d; T_1,\ldots,T_{i-1},T_{i+1},T_i,T_{i+2},\ldots,T_r)\) given by

\[
(\sigma_1,\ldots,\sigma_{i-1},\sigma_{i+1},\sigma_i^{-1}\sigma_i,\sigma_{i+1},\sigma_{i+2},\ldots,\sigma_r).
\]

Let the free group \( F_{r-1} \) act on the set of Hurwitz factorizations by the \( i \)th generator \( g_i \) acting as \( B_i \); in general, this permutes the indices in the type. If we define a homomorphism \( F_{r-1} \to S_r \) by sending \( g_i \) to the transposition \((i, i + 1)\), then the kernel \( P_r \) of this homomorphisms does not permute the indices of the type when it acts on Hurwitz factorizations, so we obtain an action of \( P_r \) on Hurwitz factorizations of type \((d; T_1,\ldots,T_r)\). This is called the pure braid action.

**Example 4.11.** If we simply repeat \( B_i \) twice for any given \( i \), we obtain an action on Hurwitz factorizations of fixed type \((d; T_1,\ldots,T_r)\) which is in general non-trivial.

We observe that the action of \( P_r \) is well-defined on equivalence classes. The main fact is then the following:

**Proposition 4.12.** Two branched covers of type \((d; T_1,\ldots,T_r)\), with fixed branch points \( Q_1,\ldots,Q_r \), lie in the same connected component of \( \mathcal{H}(d; T_1,\ldots,T_r) \) if and only if for some (equivalently all) choice of \( \gamma_i \), their associated equivalence classes of Hurwitz factorizations lie in the same pure braid orbit.

**Sketch of proof.** We first give a geometric interpretation of \( B_i \). If we fix a choice of the \( \gamma_i \), then the action of \( B_i \) is precisely what is obtained by switching \( Q_i \) with \( Q_{i+1} \), while having the \( \gamma_i \) follows the points as they move, without letting them pass through any of the \( Q_j \). It follows that for any combination of the \( B_i \) giving a pure braid action, we obtain a path in \( \mathcal{H}(d; T_1,\ldots,T_r) \) connecting each cover to its image under \( B_i \), and thus the pure braid orbits are contained in the same connected component. Arguing as in Proposition 3.5 (i), proving the converse amounts to showing that the loops in \( \mathcal{U}_r \) obtained in this way from the pure braid group in fact generate the fundamental group of \( \mathcal{U}_r \). \( \square \)

5. Maps of branched covers

As one might deduce from the definition of equivalence of branched covers, we can also define a notion of a map of branched covers as follows:

**Definition 5.1.** Suppose \((C, f)\) and \((C', f')\) are branched covers. A map \((C, f) \to (C', f')\) is a holomorphic map \( \varphi : C \to C' \) such that \( f = f' \circ \varphi \).

At first glance, this may seem like a mild condition to impose, but in fact it means there are very few maps between branched covers; in fact, we have that \((C, f)\) has a map to another branched cover \((C', f')\) if and only if it factors through \((C', f')\). Many branched covers don’t factor through
any other non-trivial branched covers, and thus have no non-trivial maps. In general, we have that a map \( \varphi \) is open because it is holomorphic, and closed because \( C \) is compact, so because we have assumed that \( C \) and \( C' \) are connected, we conclude that \( \varphi \) is surjective, and thus that the degree of a cover can only go down under maps between covers.

We see immediately that \((C,f)\) is equivalent to \((C',f')\) if and only if there is an invertible map \((C,f) \rightarrow (C',f')\). Riemann’s existence theorem asserted that there is a unique branched cover, up to equivalence, corresponding to a given Hurwitz factorization. We give a sharper version of this uniqueness, and sketch the proof.

**Proposition 5.2.** Fix points \( Q_1, \ldots, Q_r \in \mathbb{C}P^1 \), an additional base point \( Q \), and loops \( \gamma_1, \ldots, \gamma_r \) based at \( Q \) satisfying the conditions of Proposition 3.3. Suppose \((C,f)\) and \((C',f')\) are branched covers of degrees \( d \) and \( d' \), with branch points (contained in the) \( Q_i \), and chosen labelings of \( f^{-1}(Q) \) and \( f'^{-1}(Q') \). Let \((\sigma_1, \ldots, \sigma_r)\) and \((\sigma'_1, \ldots, \sigma'_r)\) be the associated Hurwitz factorizations.

Then maps \( \varphi : (C,f) \rightarrow (C',f') \) correspond to set maps \( \psi : \{1, \ldots, d\} \rightarrow \{1, \ldots, d'\} \) satisfying the condition that
\[
\psi \circ \sigma_i = \sigma'_i \circ \psi
\]
for each \( i \).

**Sketch of proof.** A map \( \varphi \) yields a \( \psi : \{1, \ldots, d\} \rightarrow \{1, \ldots, d'\} \) by restricting to \( f^{-1}(Q) \) and using the labelings of the fibers over \( Q \). We have that \( \varphi \) maps \( f^{-1}(Q) \) to \((f')^{-1}(Q)\) because \( f = f' \circ \varphi \). Moreover, \( \psi \) satisfies the desired condition because \( \varphi \) must map lifts of \( \gamma_i \) to lifts of \( \gamma_i \) for each \( i \). The key then is to check that the correspondence is invertible. Away from the branch points, this follows from standard topological results, as we can take paths starting at a point of \( f^{-1}(Q) \), take their image in \( \mathbb{C}P^1 \), and then lift them to paths in \( C' \) starting at the point of \((f')^{-1}(Q)\) determined by \( \psi \). That such maps extend to holomorphic maps even over the branch points involves a local calculation. \( \square \)

A closely related definition from group theory is the following:

**Definition 5.3.** A transitive permutation group \( G \subseteq S_d \) is **imprimitive** if there exists a non-trivial decomposition of \( \{1, \ldots, d\} \) into blocks on which the action of \( G \) is well-defined. If \( G \) is not imprimitive, we say \( G \) is **primitive**.

The essence of the definition is that \( G \) is imprimitive if it is induced from a smaller-degree permutation group.

**Example 5.4.** Consider the cyclic subgroup of \( S_4 \) generated by the 4-cycle \( \sigma = (1,2,3,4) \). This is imprimitive, because we can decompose \( \{1,2,3,4\} \) into the blocks \( \{1,3\} \) and \( \{2,4\} \), and \( \sigma \) then acts as a transposition on these two blocks.

The relationship to branched covers is the following:

**Proposition 5.5.** A branched cover \((C,f)\) of degree \( d \) has a map to some branched cover \((C',f')\) of degree \( d' \), with \( 1 < d' < d \), if and only if its monodromy group is imprimitive.

**Proof.** If we have such a map, it is clear that the fibers of the corresponding \( \psi : \{1, \ldots, d\} \rightarrow \{1, \ldots, d'\} \) give a decomposition of \( \{1, \ldots, d\} \), and one sees that the action of the \( \sigma_i \) is necessarily well-defined on each block. Conversely, given a decomposition we use it to define a smaller-degree permutation action of the \( \sigma_i \), giving a Hurwitz factorization of smaller degree and a \( \psi \) satisfying the necessary condition. \( \square \)

**Remark 5.6.** If we relax our definition of branched cover to allow the base of the cover to be an arbitrary (compact, connected) Riemann surface, then a map of branched covers \((C,f) \rightarrow (C',f')\) is itself a branched cover of \( C' \) by \( C \). Thus, in this more general setting the primitive covers are
precisely the ones that cannot be written as a composition of smaller-degree covers, and many questions can be reduced to studying primitive covers.

### 6. Automorphisms of Branched Covers

We have alluded before to a more common definition of Hurwitz number which involved weighting the count of branched covers by dividing by the number of automorphisms. We now explore this more fully.

As one might expect, we have the following definition:

**Definition 6.1.** An **automorphism** of a branched cover \((C, f)\) is a biholomorphism \(\varphi : C \overset{\sim}{\rightarrow} C\) such that \(f = f \circ \varphi\).

We have:

**Proposition 6.2.** Any map \((C, f) \rightarrow (C, f)\) is an automorphism.

If \((C, f)\) has monodromy group \(G \subseteq S_d\), then the automorphism group of \((C, f)\) is isomorphic to the centralizer of \(G\) in \(S_d\) – that is, the subgroup of \(S_d\) consisting of elements which commute with every element of \(G\).

**Proof.** We have seen that a map \((C, f) \rightarrow (C, f)\) is necessarily surjective, meaning that it induces a bijection on \(f^{-1}(Q)\); the inverse bijection then induces the inverse map.

For the second assertion, we see from Proposition 5.2 that automorphisms are in bijection with permutations \(\tau \in S_d\) such that \(\tau \sigma_i = \sigma_i \tau\) for all \(i\); since the \(\sigma_i\) generate \(G\), this is precisely equivalent to requiring that \(\tau\) commutes with every element of \(G\).

The following terminology isn’t standard:

**Definition 6.3.** The **stacky Hurwitz number** \(h_{\text{st}}(d; T_1, \ldots, T_r)\) is the number of branched covers of type \((d; T_1, \ldots, T_r)\), with each cover \((C, f)\) counted with weight \(1/|\text{Aut}(C, f)|\).

From the proposition we conclude:

**Corollary 6.4.** The stacky Hurwitz number \(h_{\text{st}}(d; T_1, \ldots, T_r)\) is equal to the number of Hurwitz factorizations divided by \(d!\).

**Proof.** A branched cover \((C, f)\) corresponds to a relabeling equivalence class of Hurwitz factorizations. On the other hand, by Proposition 6.2, the number of automorphisms of \((C, f)\) is precisely equal to the number of relabelings which fix a given Hurwitz factorization in the equivalence class. Since the total number of relabelings is \(d!\), the contribution of \((C, f)\) to the stacky Hurwitz number is the number of Hurwitz factorizations in the relabeling equivalence class, divided by \(d!\). Summing over all equivalence classes gives the desired identity.

Not surprisingly, the stacky Hurwitz number often leads to cleaner formulas. From a modern point of view (motivated by the theory of stacks), it is arguably the more natural definition. However, in many cases of interest, there are no non-trivial automorphisms of the branched covers, and the two definitions agree.

**Example 6.5.** We give an example where the stacky Hurwitz number and the Hurwitz number differ. Suppose that \(d = 2\), so that every branch point is necessarily simply branched, with the corresponding permutation being a transposition. Since \(S_2\) is abelian, the automorphism group always has order 2, so the stacky Hurwitz number is always exactly half the Hurwitz number.

From one point of view, branched covers with automorphisms are very common:
Proposition 6.6. Let \((C, f)\) be a branched cover, with monodromy group \(G\). Then there exists a (unique) branched cover \((\tilde{C}, \tilde{f})\) of degree \(|G|\) which maps to \((C, f)\), has the same branch points at \((C, f)\), has monodromy group isomorphic to \(G\), and which also has automorphism group isomorphic to \(G\).

We omit the proof, but the basic idea is to use the regular representation of \(G\). The cover \((\tilde{C}, \tilde{f})\) is called the Galois closure of \((C, f)\).

This shows that automorphisms are ubiquitous among covers with imprimitive monodromy groups. However, the covers we study are frequently primitive, and in this case automorphisms are quite rare. The following is an elementary result from the theory of permutation groups:

Theorem 6.7. Let \(G \subseteq S_d\) be a transitive permutation group. If \(G\) is abelian, then the centralizer of \(G\) in \(S_d\) is equal to \(G\). If \(G\) is non-abelian and primitive, the centralizer of \(G\) is \(S_d\) trivial.

7. Riemann’s work and moduli spaces

Although Riemann’s approach to Riemann surfaces is often described in terms of branched covers, he also gave an abstract definition quite close to the modern one, and his idea was to use branched covers to study the abstract notion. Moreover, in his study of Riemann surfaces, Riemann also introduced the concept of a moduli space. More specifically, we know that a compact, connected Riemann surface is determined topologically by its genus \(g\), a non-negative integer. Riemann realized that for any \(g\), the set of compact, connected Riemann surfaces naturally forms a topological space \(\mathcal{M}_g\), which has a complex structure and therefore a complex dimension.

We will not attempt to define these additional structures formally, but will rather settle for sketching some of Riemann’s related ideas, as well as the later elaboration on these ideas by Hurwitz and others. The basic idea is that for any type \((d; T_1, \ldots, T_r)\), we know that every cover \((C, f)\) has genus \(g\) determined by the type via the Riemann-Hurwitz formula. We thus obtain a (continuous) map \(\mathcal{H}(d; T_1, \ldots, T_r) \rightarrow \mathcal{M}_g\) by the “forgetful map” sending a point corresponding to the cover \((C, f)\) to the point corresponding to the Riemann surface \(C\). A consequence of the Riemann-Roch theorem is the following:

Theorem 7.1. Let \(C\) be a compact, connected Riemann surface of genus \(g\), and \(d \geq 2g\). Then there is a branched cover \((C, f)\) of degree \(d\), and if \(g \geq 2\), the space of all such branched covers has dimension \(2d + 1 - g\). If \(g = 0\), the space has dimension \(2d - 2\), and if \(g = 1\) it has dimension \(2d - 1\).

Moreover, for a given \(C\), “most” of the branched covers \((C, f)\) of degree \(d\) will have only simple branching. Thus, we find that for \(d \geq 2g\), if \(T_i = (2, 1, \ldots, 1)\) for all \(i\), the map

\[
\mathcal{H}(d; T_1, \ldots, T_r) \rightarrow \mathcal{M}_g
\]

is surjective, with (in the case \(g \geq 2\)) fibers of dimension \(2d + 1 - g\). Now, we know that the dimension of \(\mathcal{H}(d; T_1, \ldots, T_r)\) is equal to \(r\), and we can compute \(r\) from the Riemann-Hurwitz formula as \(2d - 2 + 2g\). We thus find that the dimension of \(\mathcal{M}_g\) is equal to

\[
2d - 2 + 2g - (2d + 1 - g) = 3g - 3.
\]

Similarly, for \(g = 0\) we get dimension 0, and for \(g = 1\) we get dimension 1.

Somewhat later, Severi was interested in studying the geometry of \(\mathcal{M}_g\) further, and he observed that the same argument above, together with the Clebsch-Hurwitz result on connectedness of Hurwitz spaces of simply branched covers, also proves that \(\mathcal{M}_g\) is connected, since a continuous image of a connected space is connected. This is an example of a surprisingly common phenomenon in mathematics: although the Hurwitz space appears more complicated than \(\mathcal{M}_g\) because it involves considering the additional structure of the covering map as well as the Riemann surface, in fact it
may be easier to study. Introducing additional structure to a problem in order to make it more tractable has emerged as a common theme in modern work on moduli spaces.