

# TEMPLATE FOR REU PAPERS

NAME

ABSTRACT. This is a sample latex document with emphasis on using math mode and equation environments. You should use it as a template for your paper. Remember that a first draft must be submitted to mentors at the very latest by August 14, with no exceptions. The completed paper must be submitted by August 28, unless permission for a later date has been obtained from the director of the program.

## CONTENTS

0. Introduction	2
0.1. Motivation	2
0.2. Statement and roadmap	2
0.3. Prerequisites and conventions	3
0.4. Historical remark	3
1. Riemann surfaces: definitions and first properties	4
1.1. Holomorphic functions on domains in $\mathbb{C}$	4
1.2. Topological surfaces and Riemann surfaces	5
1.3. Holomorphic maps between Riemann surfaces	5
1.4. Examples	6
1.5. Simply connected Riemann surfaces	7
1.6. The three model simply connected surfaces	7
2. Normal families and compactness principles	9
2.1. Locally uniform convergence	9
2.2. Normal families	11
2.3. Montel's theorem	11
2.4. Hurwitz's theorem and its consequences	12
2.5. Maximum principles revisited	13
3. The hyperbolic case: extremal maps into the unit disk	14
3.1. The extremal problem	14
3.2. Existence of an extremal map into $\mathbb{D}$	16
3.3. The extremal map has no critical points	18
3.4. Covering behavior of the extremal map	22
3.5. Surjectivity onto the disk	22
3.6. Conclusion of the hyperbolic case	23
4. The parabolic and elliptic cases	24
4.1. Analytic continuation and monodromy	24
4.2. The parabolic case: the complex plane	27
4.3. The elliptic case: compact simply connected surfaces	30
4.4. Completion of the proof of uniformization	32
Appendix A. Classical complex analysis toolkit	33

Appendix B. Covering spaces	35
Appendix C. A topological input: compact simply connected surfaces	37

## 0. INTRODUCTION

**0.1. Motivation.** The uniformization theorem is one of the central classification results in the theory of Riemann surfaces. Roughly speaking, it asserts that there are only three simply connected conformal geometries in complex dimension one: the spherical, Euclidean, and hyperbolic geometries. Concretely, if  $X$  is a simply connected Riemann surface, then  $X$  is biholomorphic to exactly one of the three model surfaces

$$\widehat{\mathbb{C}}, \quad \mathbb{C}, \quad D,$$

and these three are pairwise non-biholomorphic.

This trichotomy has immediate conceptual consequences. For an arbitrary connected Riemann surface  $Y$ , the universal cover  $\widetilde{Y}$  is simply connected, so the theorem identifies  $\widetilde{Y}$  with one of the models above; many analytic and geometric questions on  $Y$  can then be studied by lifting them to  $\widetilde{Y}$  and using the much more rigid structure of its automorphism group. In this way, uniformization provides a global “normal form” for Riemann surfaces: up to biholomorphism, each surface is a quotient of one of three universal models by a discrete group of automorphisms.

There are several proof strategies in the literature. Some proofs proceed through potential theory and the construction of Green’s functions; others use the theory of algebraic curves (via Riemann–Roch and genus) or differential geometry (via curvature and PDE). The aim of this paper is to present a proof that is analytic in the classical one-variable sense: the main inputs are normal family compactness, extremal problems for holomorphic maps, and analytic continuation organized via monodromy, together with elementary covering space theory. We do not attempt to be completely self-contained at the level of the classical theorems we cite, but we do aim to keep the logical structure transparent and the uniformization-specific arguments fully detailed.

**0.2. Statement and roadmap.** We work throughout with connected Riemann surfaces. Our main result is the uniformization theorem for simply connected surfaces.

**Theorem 0.1** (Uniformization for simply connected Riemann surfaces). *Let  $X$  be a simply connected Riemann surface. Then  $X$  is biholomorphic to exactly one of the following three model surfaces:*

- (i) *the unit disk  $\mathbb{D}$ ;*
- (ii) *the complex plane  $\mathbb{C}$ ;*
- (iii) *the Riemann sphere  $\widehat{\mathbb{C}}$ .*

*Moreover, these three surfaces are pairwise non-biholomorphic, so the conformal type of  $X$  is uniquely determined.*

**Roadmap.** Our strategy separates the proof into three mutually exclusive cases and treats each by a method suited to it.

Section 1 reviews the basic language of Riemann surfaces and holomorphic maps, and introduces the three model simply connected surfaces together with their automorphism groups. This provides the targets for classification and the basic rigidity features (in particular transitivity of automorphisms) used later.

Section 2 develops the compactness tools that drive the extremal arguments: we discuss local uniform convergence on Riemann surfaces, normal families, Montel's theorem, and Hurwitz's theorem, and we collect a short list of other classical facts from one-variable complex analysis (Schwarz–Pick and the classification of isolated singularities) that will be invoked without proof.

Section 3 proves the *hyperbolic case*. Assuming  $X$  admits a nonconstant holomorphic map to the disk  $D$ , we set up an extremal problem that maximizes the derivative at a base point among all normalized disk-valued holomorphic maps. Normal-family compactness yields existence of an extremal map; a deformation argument then forces the extremal map to have no critical points, hence to be a covering map onto its image. A final step shows the image is all of  $D$ , so simple connectedness implies  $X \cong D$ .

Section 4 treats the remaining two cases by analytic continuation and monodromy. In the *parabolic case*, where  $X$  is noncompact but admits no nonconstant bounded holomorphic functions, we use monodromy to extend suitable local charts along paths to obtain global holomorphic coordinates, leading to a biholomorphism  $X \cong \mathbb{C}$ . In the compact simply connected case, we combine a standard topological input (a compact, connected, simply connected surface is homeomorphic to  $S^2$ ) with the parabolic classification of punctured surfaces and the classification of isolated singularities to conclude  $X \cong \hat{\mathbb{C}}$ . The final subsection assembles the three cases to complete the proof of Theorem 0.1.

**0.3. Prerequisites and conventions.** We assume the reader is comfortable with:

- real analysis at the level of uniform convergence, compactness, and basic metric-space arguments;
- a first course in complex analysis (holomorphic and meromorphic functions, Cauchy's theorem and Cauchy estimates, Laurent series, residues);
- basic topology and covering spaces (fundamental group, existence and uniqueness of universal covers under the standard hypotheses).

No prior familiarity with differential geometry and algebraic geometry is assumed.

We will freely use several standard theorems from one-variable complex analysis without proof. We will indicate precisely where each of these results is used; proofs can be found in any standard textbook on complex analysis. On the topological side, we use two standard inputs from covering space theory and surface topology: the existence of universal covering spaces for Riemann surfaces, and the fact that any compact, connected, simply connected surface is homeomorphic to  $S^2$ .

All remaining arguments specific to uniformization—including the extremal map construction in the hyperbolic case, the monodromy-based analytic continuation machinery used to build global coordinates in the parabolic case, and the final assembly of the three cases—are proved in detail in the text. In particular, once the classical complex-analytic tools listed above are granted, the main classification steps are carried out from first principles.

**0.4. Historical remark.** The uniformization theorem originated in the late nineteenth and early twentieth century, with foundational contributions of Riemann

and the definitive proofs given independently by Poincaré and Koebe. Over time, several proof paradigms emerged. One family of arguments proceeds through potential theory (Green’s functions and harmonic measure), another through differential geometry and PDE (constant curvature metrics), and another through algebraic geometry (Riemann–Roch and genus).

The route taken in this paper is deliberately “classical analytic” in the sense of one-variable complex analysis: normal-family compactness drives the existence of extremal holomorphic maps, and analytic continuation is organized globally via monodromy, with only elementary covering space theory in the background. Our goal is not to optimize prerequisites, but to keep the logical structure transparent and the case-splitting (hyperbolic/parabolic/elliptic) conceptually clean.

## 1. RIEMANN SURFACES: DEFINITIONS AND FIRST PROPERTIES

In this section we recall some basic facts about holomorphic functions on domains in  $\mathbb{C}$  and then introduce Riemann surfaces and holomorphic maps between them. We emphasize a formulation that is as close as possible to the classical theory of complex functions on planar domains.

**1.1. Holomorphic functions on domains in  $\mathbb{C}$ .** We begin by fixing notation and recalling standard terminology from elementary complex analysis.

**Definition 1.1.** A *domain* in  $\mathbb{C}$  is a nonempty open connected subset  $U \subset \mathbb{C}$ .

**Definition 1.2.** Let  $U \subset \mathbb{C}$  be open and let  $f : U \rightarrow \mathbb{C}$ . We say that  $f$  is *holomorphic* (or *analytic*) on  $U$  if for every  $a \in U$  the limit

$$f'(a) := \lim_{z \rightarrow a} \frac{f(z) - f(a)}{z - a}$$

exists.

The usual theory developed in a first course in complex variables shows that holomorphic functions are automatically very regular. We record some of the basic properties that will be used freely later on.

**Theorem 1.3** (Cauchy’s integral theorem and formula). *Let  $U \subset \mathbb{C}$  be a domain, let  $f : U \rightarrow \mathbb{C}$  be holomorphic and let  $\gamma$  be a closed rectifiable curve in  $U$  which is null-homotopic in  $U$ . Then*

$$\int_{\gamma} f(z) dz = 0.$$

*If furthermore  $\gamma$  is the positively oriented boundary of a disk  $\overline{D(a, r)} \subset U$ , then for all  $w \in D(a, r)$  we have*

$$f(w) = \frac{1}{2\pi i} \int_{\gamma} \frac{f(z)}{z - w} dz,$$

*and for every integer  $n \geq 1$ ,*

$$f^{(n)}(w) = \frac{n!}{2\pi i} \int_{\gamma} \frac{f(z)}{(z - w)^{n+1}} dz.$$

**1.2. Topological surfaces and Riemann surfaces.** We now introduce the basic objects of study.

**Definition 1.4.** A *topological surface* is a second countable Hausdorff topological space  $X$  such that every point  $p \in X$  has an open neighborhood  $U$  and a homeomorphism

$$\varphi : U \rightarrow V \subset \mathbb{R}^2$$

onto an open subset  $V$  of  $\mathbb{R}^2$ .

In practice we do not need the real two-dimensional description explicitly; instead we will work directly with complex coordinates.

**Definition 1.5.** Let  $X$  be a topological surface. A *chart* on  $X$  is a pair  $(U, z)$  where  $U \subset X$  is open and  $z : U \rightarrow V$  is a homeomorphism onto an open subset  $V \subset \mathbb{C}$ . An *atlas* on  $X$  is a collection of charts  $\{(U_\alpha, z_\alpha)\}_{\alpha \in A}$  such that  $\{U_\alpha\}_{\alpha \in A}$  covers  $X$ .

**Definition 1.6.** Two charts  $(U, z)$  and  $(V, w)$  on  $X$  are said to be *holomorphically compatible* if either  $U \cap V = \emptyset$ , or else the transition map

$$w \circ z^{-1} : z(U \cap V) \rightarrow w(U \cap V)$$

is a biholomorphic map between open subsets of  $\mathbb{C}$ .

**Definition 1.7.** An atlas  $\{(U_\alpha, z_\alpha)\}$  on  $X$  is called a *holomorphic atlas* if every pair of charts in the atlas is holomorphically compatible.

**Definition 1.8.** A *Riemann surface* is a pair  $(X, \mathcal{A})$ , where  $X$  is a topological surface and  $\mathcal{A}$  is a maximal holomorphic atlas on  $X$ . Often we simply write  $X$  and leave the atlas understood.

The maximality of the atlas means that if a chart is holomorphically compatible with every chart in  $\mathcal{A}$ , then it already belongs to  $\mathcal{A}$ . Given any holomorphic atlas on a topological surface, it is contained in a unique maximal holomorphic atlas, so we lose no generality by working with maximal atlases.

*Remark 1.9.* The requirement that the transition maps be biholomorphic endows  $X$  with the structure of a 1-dimensional complex manifold. In particular, a Riemann surface carries a canonical orientation.

**1.3. Holomorphic maps between Riemann surfaces.** We next define what it means for a map between Riemann surfaces to be holomorphic. The definition is modeled on the usual one for functions on domains in  $\mathbb{C}$ , but it is expressed in terms of local coordinates.

**Definition 1.10.** Let  $X$  and  $Y$  be Riemann surfaces. A continuous map  $f : X \rightarrow Y$  is *holomorphic* if for every point  $p \in X$  there exist charts  $(U, z)$  on  $X$  with  $p \in U$  and  $(V, w)$  on  $Y$  with  $f(p) \in V$  and  $f(U) \subset V$  such that the coordinate expression

$$w \circ f \circ z^{-1} : z(U) \rightarrow w(V)$$

is a holomorphic function between open subsets of  $\mathbb{C}$  in the usual sense.

It is straightforward to verify that this definition does not depend on the choice of charts.

**Lemma 1.11.** *Let  $X$  and  $Y$  be Riemann surfaces and let  $f : X \rightarrow Y$  be a continuous map. If for one choice of charts  $(U, z)$  around  $p \in X$  and  $(V, w)$  around  $f(p)$  the coordinate expression  $w \circ f \circ z^{-1}$  is holomorphic on  $z(U)$ , then the same holds for any other choice of charts around  $p$  and  $f(p)$  with sufficiently small domains.*

*Proof.* Let  $(U, z)$  and  $(U', z')$  be charts on  $X$  around  $p$  and  $(V, w)$ ,  $(V', w')$  charts on  $Y$  around  $f(p)$ , with  $f(U \cap U') \subset V \cap V'$ . On the overlap we have

$$w' \circ f \circ (z')^{-1} = (w' \circ w^{-1}) \circ (w \circ f \circ z^{-1}) \circ (z \circ (z')^{-1}).$$

The transition maps  $w' \circ w^{-1}$  and  $z \circ (z')^{-1}$  are biholomorphic (by compatibility of charts), and compositions of holomorphic maps are holomorphic. Hence  $w' \circ f \circ (z')^{-1}$  is holomorphic wherever it is defined.  $\square$

**Definition 1.12.** A bijective holomorphic map  $f : X \rightarrow Y$  whose inverse  $f^{-1} : Y \rightarrow X$  is also holomorphic is called a *biholomorphism* (or a *conformal equivalence*) between  $X$  and  $Y$ . In this case we say that  $X$  and  $Y$  are *conformally equivalent*.

When we classify Riemann surfaces up to conformal equivalence, we are asking for a description of the biholomorphism classes of Riemann surfaces.

**1.4. Examples.** We record a few basic examples that will be used throughout the paper.

**Example 1.13** (Complex plane). The complex plane  $\mathbb{C}$ , equipped with the topology inherited from  $\mathbb{R}^2$  and the single global chart  $(\mathbb{C}, \text{id})$ , is a Riemann surface. Holomorphic maps  $\mathbb{C} \rightarrow \mathbb{C}$  in this sense are precisely the usual entire functions.

**Example 1.14** (Open subsets of  $\mathbb{C}$ ). Let  $\Omega \subset \mathbb{C}$  be an open set. With the subspace topology and the chart  $(\Omega, \text{id}|_{\Omega})$ ,  $\Omega$  is a Riemann surface. Holomorphic maps  $\Omega \rightarrow \mathbb{C}$  are exactly the usual holomorphic functions on the domain  $\Omega$ .

**Example 1.15** (Unit disk). The *unit disk*

$$\mathbb{D} := \{z \in \mathbb{C} : |z| < 1\}$$

is a Riemann surface as an open subset of  $\mathbb{C}$ . Many of our extremal constructions will use holomorphic maps from a given Riemann surface into  $\mathbb{D}$ .

**Example 1.16** (Riemann sphere). The *Riemann sphere*  $\widehat{\mathbb{C}}$  is the topological 2-sphere  $S^2$ , often described as  $\mathbb{C} \cup \{\infty\}$ . We equip  $\widehat{\mathbb{C}}$  with two charts:

- (i)  $(U_1, z_1)$  with  $U_1 = \widehat{\mathbb{C}} \setminus \{\infty\}$  and  $z_1$  the identity map  $z_1(z) = z$  onto  $\mathbb{C}$ ;
- (ii)  $(U_2, z_2)$  with  $U_2 = \widehat{\mathbb{C}} \setminus \{0\}$  and  $z_2(z) = 1/z$  for  $z \in \mathbb{C}^\times$  and  $z_2(\infty) = 0$ .

On the overlap  $U_1 \cap U_2 = \mathbb{C}^\times$  the transition maps are

$$z_2 \circ z_1^{-1}(z) = 1/z, \quad z_1 \circ z_2^{-1}(w) = 1/w,$$

which are biholomorphic on  $\mathbb{C}^\times$ . Thus,  $\widehat{\mathbb{C}}$  is a Riemann surface. Holomorphic functions  $f : \widehat{\mathbb{C}} \rightarrow \widehat{\mathbb{C}}$  are precisely the rational functions on  $\mathbb{C}$ , together with the value  $f(\infty)$  defined appropriately.

More complicated examples, such as tori  $\mathbb{C}/\Lambda$  and Riemann surfaces of algebraic functions, will appear only implicitly; they are not needed for the proof of the uniformization theorem.

**1.5. Simply connected Riemann surfaces.** Finally we recall the notion of simple connectedness in the topological sense.

**Definition 1.17.** A topological space  $X$  is *path-connected* if any two points  $p, q \in X$  can be joined by a continuous map  $\gamma : [0, 1] \rightarrow X$  with  $\gamma(0) = p$  and  $\gamma(1) = q$ .

A path-connected space  $X$  is *simply connected* if every continuous loop  $\gamma : [0, 1] \rightarrow X$  with  $\gamma(0) = \gamma(1)$  can be continuously deformed to a constant loop. Equivalently, the fundamental group  $\pi_1(X, x_0)$  is trivial for any (and hence every) base point  $x_0 \in X$ .

Since Riemann surfaces are locally homeomorphic to open subsets of  $\mathbb{R}^2$ , they are locally path-connected and semilocally simply connected. It follows from general topology that each Riemann surface admits a universal covering space. We will return to this in a later section.

For now we record the following examples.

**Example 1.18.** The Riemann surfaces  $\mathbb{C}$  and  $\mathbb{D}$  are simply connected. This follows from the fact that both are homeomorphic to open disks in  $\mathbb{R}^2$ , which are contractible.

**Example 1.19.** The Riemann sphere  $\widehat{\mathbb{C}}$  is simply connected. For instance, by identifying  $\widehat{\mathbb{C}}$  with the unit sphere  $S^2 \subset \mathbb{R}^3$  via stereographic projection, we may appeal to basic algebraic topology, which shows that  $\pi_1(S^2)$  is trivial.

The uniformization theorem will ultimately assert that every simply connected Riemann surface is conformally equivalent to exactly one of  $D$ ,  $\mathbb{C}$ , or  $\widehat{\mathbb{C}}$ . In the remainder of this section we record some basic properties of these three model surfaces; the following sections will then develop the analytic tools needed to prove this classification.

**1.6. The three model simply connected surfaces.** We now single out three basic Riemann surfaces which will serve as the models in the uniformization theorem:

$$\widehat{\mathbb{C}}, \mathbb{C}, D.$$

All three have already appeared as examples (see Examples 1.13, 1.15, 1.16). Here we recall their role as simply connected model surfaces and describe their holomorphic automorphism groups.

**1.6.1. The Riemann sphere.** We begin with the compact model.

**Definition 1.20.** The *Riemann sphere* is the Riemann surface  $\widehat{\mathbb{C}}$  constructed in Example 1.20 as the one-point compactification of  $\mathbb{C}$  by adding a point denoted  $\infty$ . We always regard it as equipped with the complex structure described there.

The holomorphic automorphism group of  $\widehat{\mathbb{C}}$  is well known.

**Proposition 1.21.** Every biholomorphism  $f : \widehat{\mathbb{C}} \rightarrow \widehat{\mathbb{C}}$  is a Möbius transformation, that is, there exist  $a, b, c, d \in \mathbb{C}$  with  $ad - bc \neq 0$  such that

$$f(z) = \frac{az + b}{cz + d}$$

for  $z \in \mathbb{C}$  with  $cz + d \neq 0$ , and  $f(\infty) = a/c$  if  $c \neq 0$ , while  $f(\infty) = \infty$  if  $c = 0$ . Conversely, every such Möbius transformation defines a biholomorphism of  $\widehat{\mathbb{C}}$ .

*Proof.* It is classical that nonconstant holomorphic self-maps of  $\widehat{\mathbb{C}}$  are rational functions of some degree  $d \geq 1$ . If  $f$  is a biholomorphism, then  $f^{-1}$  is also holomorphic, hence rational. This forces  $d = 1$ , so  $f$  must be a Möbius transformation. The converse is a direct computation.  $\square$

We denote the automorphism group of  $\widehat{\mathbb{C}}$  by

$$\text{Aut}(\widehat{\mathbb{C}}) = \left\{ z \mapsto \frac{az + b}{cz + d} : a, b, c, d \in \mathbb{C}, ad - bc \neq 0 \right\} / \mathbb{C}^\times,$$

which is naturally isomorphic to the projective linear group  $\text{PSL}(2, \mathbb{C})$ .

**1.6.2. The complex plane.** Next we consider the noncompact model of parabolic type. The automorphism group of  $\mathbb{C}$  as a Riemann surface is very simple.

**Proposition 1.22.** *Every biholomorphism  $f : \mathbb{C} \rightarrow \mathbb{C}$  is an affine map*

$$f(z) = az + b$$

*with  $a \in \mathbb{C}^\times$  and  $b \in \mathbb{C}$ . Conversely, every such affine map is a biholomorphism.*

*Proof.* Let  $f : \mathbb{C} \rightarrow \mathbb{C}$  be a biholomorphism. Since  $f$  is entire and injective, its derivative  $f'(z)$  never vanishes. By Liouville's theorem, an entire function is bounded if and only if it is constant. To see that  $f'$  is bounded, note that  $f$  is a homeomorphism of  $\mathbb{C}$  onto itself and therefore proper; in particular,  $|f(z)| \rightarrow \infty$  as  $|z| \rightarrow \infty$ . Cauchy's estimates on large circles then imply that  $f'$  is bounded on  $\mathbb{C}$ , hence constant. Thus  $f(z) = az + b$  for some  $a, b \in \mathbb{C}$  with  $a \neq 0$ . Conversely, any such affine map is a global biholomorphism of  $\mathbb{C}$ .  $\square$

We will write

$$\text{Aut}(\mathbb{C}) = \{z \mapsto az + b : a \in \mathbb{C}^\times, b \in \mathbb{C}\}.$$

**1.6.3. The unit disk.** Finally we introduce the hyperbolic model. The biholomorphisms of  $D$  are given by the classical disk automorphisms.

**Proposition 1.23.** *Every biholomorphism  $f : D \rightarrow D$  has the form*

$$f(z) = e^{i\theta} \frac{z - a}{1 - \bar{a}z}$$

*for some  $a \in D$  and real  $\theta \in \mathbb{R}$ . Conversely, every map of this form is a biholomorphism of  $D$  onto itself.*

*Proof.* This is a standard result in classical complex analysis, usually proved using the.  $\square$

We will denote the automorphism group of the unit disk by

$$\text{Aut}(D) = \left\{ z \mapsto e^{i\theta} \frac{z - a}{1 - \bar{a}z} : a \in D, \theta \in \mathbb{R} \right\}.$$

1.6.4. *Transitivity of automorphism groups.* A key feature of the three model surfaces is that their automorphism groups act transitively.

**Proposition 1.24.** *For each of the Riemann surfaces  $\widehat{\mathbb{C}}$ ,  $\mathbb{C}$  and  $D$ , the automorphism group acts transitively on the underlying surface. More precisely:*

- (i) *For any two points  $p, q \in \widehat{\mathbb{C}}$ , there exists a Möbius transformation  $\varphi \in \text{Aut}(\widehat{\mathbb{C}})$  with  $\varphi(p) = q$ .*
- (ii) *For any two points  $p, q \in \mathbb{C}$ , there exists an affine map  $\varphi \in \text{Aut}(\mathbb{C})$  with  $\varphi(p) = q$ .*
- (iii) *For any two points  $p, q \in D$ , there exists a disk automorphism  $\varphi \in \text{Aut}(D)$  with  $\varphi(p) = q$ .*

*Proof.* For (i), if  $p \neq \infty$  and  $q \neq \infty$ , the Möbius transformation

$$z \mapsto \frac{z - p}{z - q}$$

maps  $p$  to 0 and  $q$  to  $\infty$ . Composing with suitable scalings and translations gives a transformation sending  $p$  to  $q$ . The other cases are similar and standard.

For (ii), the map  $z \mapsto z - p + q$  is an automorphism sending  $p$  to  $q$ .

For (iii), Proposition 1.23 shows that for each  $a \in D$  there is an automorphism

$$\varphi_a(z) = \frac{z - a}{1 - \bar{a}z}$$

sending  $a$  to 0. Composing two such maps shows that for any  $p, q \in D$  there exists an automorphism sending  $p$  to  $q$ .  $\square$

Transitivity will be important later when we discuss the uniqueness aspect of uniformization: once a simply connected Riemann surface is identified with one of the three models, the identification is unique up to the action of the corresponding automorphism group.

## 2. NORMAL FAMILIES AND COMPACTNESS PRINCIPLES

In this section we introduce the notion of a normal family of holomorphic maps and record the basic compactness results that will be used in the extremal construction of the uniformizing map. The presentation is modeled on the classical theory for holomorphic functions on domains in  $\mathbb{C}$ , but phrased for maps between Riemann surfaces.

Throughout,  $X$  will denote a Riemann surface and  $\Omega \subset X$  a nonempty open subset.

**2.1. Locally uniform convergence.** We start with the notion of locally uniform convergence.

**Definition 2.1.** Let  $Y$  be a metric space and let  $(f_n)$  be a sequence of maps  $f_n : \Omega \rightarrow Y$ . We say that  $f_n$  converges *locally uniformly* to a map  $f : \Omega \rightarrow Y$  if for every compact set  $K \subset \Omega$  the sequence  $f_n|_K$  converges uniformly to  $f|_K$  with respect to the metric on  $Y$ .

In the situations that concern us,  $Y$  will be  $\mathbb{C}$ , the unit disk  $\mathbb{D}$ , or the Riemann sphere  $\widehat{\mathbb{C}}$ , all endowed with their standard metrics.

When  $Y$  is a Riemann surface, the above definition is interpreted using local charts.

**Lemma 2.2.** *Let  $Y$  be a Riemann surface and let  $f_n : \Omega \rightarrow Y$  be a sequence of continuous maps. The following are equivalent:*

- (i)  $f_n$  converges locally uniformly to  $f : \Omega \rightarrow Y$  (with respect to some compatible metric on  $Y$ );
- (ii) for every chart  $(V, w)$  on  $Y$  and every compact  $K \subset f^{-1}(V)$ , the coordinate expressions  $w \circ f_n|_K$  converge uniformly to  $w \circ f|_K$ .

*Proof.* Fix once and for all a metric  $d_Y$  on  $Y$  that induces its given topology.

(i)  $\Rightarrow$  (ii). Let  $(V, w)$  be a chart on  $Y$  and let  $K \subset f^{-1}(V)$  be compact. Then  $f(K)$  is a compact subset of the open set  $V$ , so there exists an open set  $W$  with

$$f(K) \subset W \subset \overline{W} \subset V.$$

Since  $f_n \rightarrow f$  locally uniformly, there exists  $N$  such that

$$\sup_{x \in K} d_Y(f_n(x), f(x)) < \text{dist}(f(K), Y \setminus W)$$

for all  $n \geq N$ . In particular, for  $n \geq N$  we have  $f_n(K) \subset W \subset V$ , so  $w \circ f_n$  is defined on  $K$ .

On the compact set  $\overline{W}$ , the chart map

$$w : \overline{W} \rightarrow w(\overline{W}) \subset \mathbb{C}$$

is a homeomorphism between compact metric spaces, hence uniformly continuous. Thus the uniform convergence  $f_n|_K \rightarrow f|_K$  in the metric  $d_Y$  implies uniform convergence in coordinates:

$$\sup_{x \in K} |w(f_n(x)) - w(f(x))| \rightarrow 0.$$

This is exactly (ii).

(ii)  $\Rightarrow$  (i). Let  $K \subset \Omega$  be compact. The compact set  $f(K) \subset Y$  is covered by charts  $\{(V_\alpha, w_\alpha)\}_{\alpha \in A}$ . Since  $f(K)$  is compact, there exist finitely many indices  $\alpha_1, \dots, \alpha_m$  such that

$$f(K) \subset \bigcup_{j=1}^m V_{\alpha_j}.$$

Set

$$K_j := K \cap f^{-1}(V_{\alpha_j}), \quad j = 1, \dots, m.$$

Each  $K_j$  is compact and  $K = \bigcup_{j=1}^m K_j$ .

Fix  $j$ . By assumption (ii), there exists  $N_j$  such that for all  $n \geq N_j$ , the points  $f_n(K_j)$  lie in  $V_{\alpha_j}$  and

$$w_{\alpha_j} \circ f_n|_{K_j} \rightarrow w_{\alpha_j} \circ f|_{K_j}$$

uniformly on  $K_j$ .

Since  $f_n(K_j)$  and  $f(K_j)$  are contained in  $V_{\alpha_j}$  for  $n$  large, we can restrict  $w_{\alpha_j}$  to a compact neighborhood  $W_j \subset V_{\alpha_j}$  of  $f(K_j)$ . On  $W_j$ , the map

$$w_{\alpha_j} : W_j \rightarrow w_{\alpha_j}(W_j)$$

is a homeomorphism between compact metric spaces, hence both  $w_{\alpha_j}$  and its inverse are uniformly continuous. Thus the uniform convergence in coordinates implies that

$$\sup_{x \in K_j} d_Y(f_n(x), f(x)) \rightarrow 0$$

as  $n \rightarrow \infty$ .

Now let  $\varepsilon > 0$ . For each  $j$  choose  $N_j$  so that

$$\sup_{x \in K_j} d_Y(f_n(x), f(x)) < \varepsilon$$

for all  $n \geq N_j$ . Let  $N := \max\{N_1, \dots, N_m\}$ . Then for all  $n \geq N$ ,

$$\sup_{x \in K} d_Y(f_n(x), f(x)) = \max_{1 \leq j \leq m} \sup_{x \in K_j} d_Y(f_n(x), f(x)) < \varepsilon.$$

Since  $K$  was an arbitrary compact subset of  $\Omega$ , this is precisely local uniform convergence in the sense of Definition 2.1.  $\square$

In particular, when  $Y$  is  $\mathbb{C}$  or  $\mathbb{D}$  equipped with the Euclidean metric, locally uniform convergence coincides with locally uniform convergence in the usual sense on planar domains.

If  $(f_n)$  is a sequence of holomorphic maps  $f_n : \Omega \rightarrow Y$  that converges locally uniformly to a map  $f : \Omega \rightarrow Y$ , then by passing to local coordinates one sees that  $f$  is again holomorphic. Thus the class of holomorphic maps between Riemann surfaces is closed under locally uniform limits.

**2.2. Normal families.** We now define the central compactness notion.

**Definition 2.3.** Let  $Y$  be a Riemann surface and  $\Omega \subset X$  open. A family  $\mathcal{F}$  of holomorphic maps  $f : \Omega \rightarrow Y$  is called *normal* if every sequence  $(f_n)$  in  $\mathcal{F}$  admits a subsequence that converges locally uniformly on  $\Omega$  to a holomorphic map  $f : \Omega \rightarrow Y$ .

In the classical literature one often allows convergence to the constant map with value  $\infty$  when  $Y = \widehat{\mathbb{C}}$ ; we will not need this variant here and will always take  $Y$  to be either  $\mathbb{C}$  or  $\mathbb{D}$  in applications.

Normal families provide a substitute for compactness in infinite-dimensional spaces of holomorphic maps: although the space of all holomorphic maps  $\Omega \rightarrow Y$  is not compact in any reasonable sense, normality ensures that bounded sequences admit convergent subsequences on compact subsets.

*Remark 2.4.* The definition is independent of the choice of metric on  $Y$ . Indeed, any two metrics that induce the same topology on a compact set are uniformly equivalent, and locally uniform convergence is a local notion.

**2.3. Montel's theorem.** We now state the basic compactness criterion due to Montel. We first recall the classical version for holomorphic functions on planar domains.

**Theorem 2.5** (Montel's theorem on domains in  $\mathbb{C}$ ). *Let  $D \subset \mathbb{C}$  be a domain and let  $\mathcal{F}$  be a family of holomorphic functions  $f : D \rightarrow \mathbb{C}$  which is locally bounded in the following sense: for every compact set  $K \subset D$  there exists  $M_K > 0$  such that*

$$|f(z)| \leq M_K \quad \text{for all } z \in K \text{ and all } f \in \mathcal{F}.$$

*Then  $\mathcal{F}$  is a normal family on  $D$ .*

A proof of this theorem can be found in any standard text on complex analysis, which is based on the Arzelà–Ascoli theorem and Cauchy's integral estimates.

We next extend Montel's theorem to holomorphic maps defined on open subsets of arbitrary Riemann surfaces and taking values in  $\mathbb{C}$  or  $\mathbb{D}$ .

**Theorem 2.6** (Montel's theorem on Riemann surfaces). *Let  $X$  be a Riemann surface,  $\Omega \subset X$  a nonempty open set, and let  $\mathcal{F}$  be a family of holomorphic maps  $f : \Omega \rightarrow \mathbb{C}$  which is locally bounded: for every compact  $K \subset \Omega$  there exists  $M_K > 0$  such that*

$$|f(z)| \leq M_K \quad \text{for all } z \in K \text{ and all } f \in \mathcal{F}.$$

*Then  $\mathcal{F}$  is a normal family on  $\Omega$ .*

*In particular, any family of holomorphic maps  $f : \Omega \rightarrow \mathbb{D}$  is normal.*

*Proof.* Cover  $\Omega$  by a countable collection of coordinate disks  $(U_j, z_j)$  with  $U_j \subset \Omega$  and  $z_j(U_j)$  a disk in  $\mathbb{C}$ . Let  $(f_n)$  be any sequence in  $\mathcal{F}$ . Fix  $j$  and consider the coordinate expressions

$$g_{n,j} := f_n \circ z_j^{-1} : z_j(U_j) \rightarrow \mathbb{C}.$$

The local boundedness hypothesis on  $\mathcal{F}$  implies that for each compact  $K \subset U_j$  there is a bound on  $|f(z)|$  for  $z \in K$  and  $f \in \mathcal{F}$ ; equivalently, the family  $\{g_{n,j}\}_n$  is locally bounded on the planar domain  $z_j(U_j)$ . By Theorem 2.5, for each fixed  $j$  there is a subsequence of  $(f_n)$ , which we still denote by  $(f_n)$  for simplicity, such that  $(g_{n,j})$  converges locally uniformly on  $z_j(U_j)$  to a holomorphic function  $g_j$ .

By a standard diagonal argument, passing to a subsequence of  $(f_n)$  we may assume that for every  $j$  the sequence  $(g_{n,j})$  converges locally uniformly on  $z_j(U_j)$  to some holomorphic function  $g_j$ . The compatibility of the charts and uniqueness of limits imply that these holomorphic functions  $g_j$  patch together to define a holomorphic map  $f : \Omega \rightarrow \mathbb{C}$  such that  $f_n \rightarrow f$  locally uniformly on  $\Omega$ . Thus  $\mathcal{F}$  is normal.

If each  $f \in \mathcal{F}$  maps into  $\mathbb{D}$ , then the local boundedness assumption is automatically satisfied with  $M_K = 1$  for every  $K$ , so the conclusion follows.  $\square$

**2.4. Hurwitz's theorem and its consequences.** We now recall Hurwitz's theorem on the behavior of zeros of holomorphic functions under locally uniform convergence.

**Theorem 2.7** (Hurwitz's theorem on domains in  $\mathbb{C}$ ). *Let  $D \subset \mathbb{C}$  be a domain and let  $(f_n)$  be a sequence of nonconstant holomorphic functions  $f_n : D \rightarrow \mathbb{C}$  converging locally uniformly on  $D$  to a holomorphic function  $f$ .*

- (i) *If each  $f_n$  has no zeros in  $D$ , then either  $f$  has no zeros in  $D$  or  $f \equiv 0$ .*
- (ii) *Suppose that for some  $a \in \mathbb{C}$  each  $f_n - a$  has at most one zero in  $D$ . If  $f$  is nonconstant and  $f(z_1) = f(z_2) = a$  for some  $z_1 \neq z_2$  in  $D$ , then for all sufficiently large  $n$  the function  $f_n - a$  has at least two zeros in  $D$ , counted with multiplicity. This contradicts the hypothesis.*

*In particular, if each  $f_n$  is injective and  $f$  is nonconstant, then  $f$  is injective.*

A proof can be completed by using the argument principle.

We now formulate a version of Hurwitz's theorem that is suitable for maps between Riemann surfaces.

**Proposition 2.8** (Hurwitz's theorem on Riemann surfaces). *Let  $X$  and  $Y$  be Riemann surfaces and let  $f_n : X \rightarrow Y$  be a sequence of holomorphic maps converging locally uniformly to a holomorphic map  $f : X \rightarrow Y$ .*

- (i) *If each  $f_n$  is nonconstant and has no critical points (i.e. its differential is nowhere zero), then either  $f$  is constant or  $f$  has no critical points.*
- (ii) *If each  $f_n$  is injective and  $f$  is nonconstant, then  $f$  is injective.*

*Proof.* Both statements are local on  $X$  and  $Y$ . Fix a point  $p \in X$  and choose charts  $(U, z)$  on  $X$  around  $p$  and  $(V, w)$  on  $Y$  around  $f(p)$  such that  $f(U) \subset V$  and  $f_n(U) \subset V$  for all  $n$  sufficiently large. Consider the coordinate expressions

$$g_n := w \circ f_n \circ z^{-1} : z(U) \rightarrow \mathbb{C}, \quad g := w \circ f \circ z^{-1} : z(U) \rightarrow \mathbb{C}.$$

By Lemma 2.2,  $g_n$  converges locally uniformly to  $g$  on the planar domain  $z(U)$ .

For (i), the assumption that  $f_n$  has no critical points on  $X$  means that  $g'_n$  never vanishes on  $z(U)$ . Applying Theorem 2.7(i) to the sequence of holomorphic functions  $g'_n$  on  $z(U)$  shows that either  $g'$  has no zeros or  $g' \equiv 0$  on  $z(U)$ . In the latter case  $g$  is constant on  $z(U)$ , hence  $f$  is constant on  $U$ . Since  $p$  was arbitrary, either  $f$  is constant on  $X$  or else  $f$  has no critical points.

For (ii), suppose  $f$  is nonconstant and not injective. Then there exist distinct points  $p_1, p_2 \in X$  with  $f(p_1) = f(p_2)$ . Choose disjoint charts  $(U_j, z_j)$  around  $p_j$  and a chart  $(V, w)$  around the common value  $f(p_1) = f(p_2)$ , with  $f(U_j) \subset V$  for  $j = 1, 2$ . For  $n$  large enough we also have  $f_n(U_j) \subset V$ . Consider the holomorphic functions

$$g_{n,1} := w \circ f_n \circ z_1^{-1},$$

and similarly

$$g_1 := w \circ f \circ z_1^{-1}, \quad g_2 := w \circ f \circ z_2^{-1}.$$

The functions  $g_{n,j}$  converge locally uniformly to  $g_j$  for  $j = 1, 2$ , and  $g_1$  and  $g_2$  have the same value at  $z_1(p_1)$  and  $z_2(p_2)$ , respectively. A standard form of Hurwitz's theorem applied to the difference  $g_{n,1} - g_{n,2}$  on suitable small disks around these points shows that, for  $n$  large enough, there exist  $x \in U_1$ ,  $y \in U_2$ ,  $x \neq y$ , with  $f_n(x) = f_n(y)$ , contradicting injectivity of  $f_n$ . Thus  $f$  must be injective.  $\square$

The main consequence that we will use later is the following.

**Corollary 2.9.** *Let  $X$  and  $Y$  be Riemann surfaces and let  $(f_n)$  be a sequence of holomorphic maps  $f_n : X \rightarrow Y$  such that each  $f_n$  is injective. Suppose that  $f_n$  converges locally uniformly to a holomorphic map  $f : X \rightarrow Y$ . Then either  $f$  is constant or  $f$  is injective and has no critical points.*

*Proof.* Apply Proposition 2.8(ii) to obtain injectivity of  $f$  in the nonconstant case, and Proposition 2.8(i) to deduce that  $f$  then has no critical points.  $\square$

**2.5. Maximum principles revisited.** Finally we record a version of the maximum modulus principle for holomorphic maps defined on Riemann surfaces. This is essentially a restatement of the classical theorem in local coordinates.

**Proposition 2.10** (Maximum modulus principle on Riemann surfaces). *Let  $X$  be a Riemann surface,  $\Omega \subset X$  a domain, and  $f : \Omega \rightarrow \mathbb{C}$  a holomorphic function. If there exists  $a \in \Omega$  such that*

$$|f(a)| \geq |f(z)| \quad \text{for all } z \in \Omega,$$

*then  $f$  is constant.*

*Proof.* Choose a chart  $(U, z)$  around  $a$  with  $U \subset \Omega$  and consider  $g := f \circ z^{-1}$ , which is holomorphic on the planar domain  $z(U)$ . The hypothesis implies that  $|g|$  attains a local maximum at  $z(a)$ . By the classical maximum modulus principle on domains in  $\mathbb{C}$ , the function  $g$  is constant on  $z(U)$ , hence  $f$  is constant on  $U$ . By the identity theorem on  $\Omega$ ,  $f$  is constant on all of  $\Omega$ .  $\square$

We will mainly apply this principle to holomorphic maps with values in the unit disk.

**Corollary 2.11.** *Let  $X$  be a Riemann surface,  $\Omega \subset X$  a domain, and  $f : \Omega \rightarrow \mathbb{D}$  a holomorphic map. Then for every relatively compact open subset  $U \subset \Omega$  we have*

$$\sup_{z \in U} |f(z)| < 1.$$

*Proof.* Suppose on the contrary that there exists a relatively compact open set  $U \subset \Omega$  such that  $\sup_{z \in U} |f(z)| = 1$ . Then there exists a sequence  $(z_n)$  in  $U$  with  $|f(z_n)| \rightarrow 1$ . Since  $\overline{U}$  is compact, after passing to a subsequence we may assume  $z_n \rightarrow a \in \overline{U}$ . Because  $U$  is relatively compact in  $\Omega$ , we have  $a \in \Omega$ . By continuity of  $f$  we obtain  $|f(a)| = 1$ , contradicting the assumption that  $f(\Omega) \subset \mathbb{D}$ . Thus  $\sup_{z \in U} |f(z)| < 1$ .  $\square$

### 3. THE HYPERBOLIC CASE: EXTREMAL MAPS INTO THE UNIT DISK

In this section we begin the analytic part of the proof of the uniformization theorem. Let  $X$  be a simply connected Riemann surface and fix a base point  $p \in X$ . Our goal is to construct a holomorphic map from  $X$  into one of the model surfaces  $D$  or  $C$  which is extremal in a suitable sense at the point  $p$ .

We first treat the case in which there exists a nonconstant holomorphic map  $X \rightarrow D$ . The parabolic case, in which no such bounded nonconstant map exists, will be discussed later.

**3.1. The extremal problem.** Throughout this subsection we assume that there exists a nonconstant holomorphic map  $g : X \rightarrow D$ . By composing with an automorphism of  $D$  we may and do assume that  $g(p) = 0$  (recall from Proposition 1.24 that  $\text{Aut}(D)$  acts transitively on  $D$ ).

We begin by recording that the derivative of a holomorphic map between Riemann surfaces at a point is well defined up to multiplication by a nonzero complex number, and that its absolute value is independent of the choice of local coordinate. We will use this only in the simplest situation.

**Lemma 3.1.** *Let  $X$  and  $Y$  be Riemann surfaces, let  $f : X \rightarrow Y$  be holomorphic, and let  $p \in X$ . Choose a chart  $(U, z)$  on  $X$  with  $p \in U$  and  $z(p) = 0$ , and a chart  $(V, w)$  on  $Y$  with  $f(p) \in V$  and  $f(U) \subset V$ . Then the coordinate expression*

$$h := w \circ f \circ z^{-1}$$

*is holomorphic on  $z(U) \subset \mathbb{C}$  and we may define*

$$f'(p) := h'(0).$$

*If we replace  $(U, z)$  and  $(V, w)$  by other charts  $(U', z')$  and  $(V', w')$  around  $p$  and  $f(p)$ , respectively, then the corresponding derivative  $h'(0)$  is multiplied by a nonzero complex number depending only on the change of coordinates. In particular,  $|f'(p)|$  is well defined.*

*Proof.* First note that  $h$  is holomorphic on  $z(U)$  because it is a composition of holomorphic maps between planar domains:

$$z(U) \xrightarrow{z^{-1}} U \xrightarrow{f} V \xrightarrow{w} w(V).$$

Thus  $h'(0)$  is well defined in the usual one-variable sense.

Now suppose we use different charts  $(U', z')$  on  $X$  and  $(V', w')$  on  $Y$ , with  $p \in U'$  and  $f(U') \subset V'$ , and set

$$h' := w' \circ f \circ (z')^{-1}.$$

We must compare the derivatives of  $h$  and  $h'$  at the points corresponding to  $p$ . Since  $z(p) = 0$  and  $z'(p) = 0$ , we have  $h'(0) = (w' \circ f \circ (z')^{-1})'(0)$ .

Consider the transition maps between the two coordinate systems on  $X$  and  $Y$ :

$$\phi := z' \circ z^{-1} : z(U \cap U') \rightarrow z'(U \cap U'), \quad \psi := w' \circ w^{-1} : w(V \cap V') \rightarrow w'(V \cap V').$$

Both  $\phi$  and  $\psi$  are biholomorphisms between planar domains. On the intersection  $U \cap U'$  we have

$$z' = \phi \circ z, \quad w' = \psi \circ w,$$

so for points where both charts are defined we can write

$$h' = w' \circ f \circ (z')^{-1} = \psi \circ w \circ f \circ z^{-1} \circ \phi^{-1} = \psi \circ h \circ \phi^{-1}.$$

Now differentiate this identity at 0. Since  $z(p) = z'(p) = 0$ , we have  $\phi(0) = 0$  and  $\phi^{-1}(0) = 0$ . By the chain rule for holomorphic functions on domains in  $\mathbb{C}$ ,

$$(h')'(0) = (\psi \circ h \circ \phi^{-1})'(0) = \psi'(h(\phi^{-1}(0))) \cdot h'(\phi^{-1}(0)) \cdot (\phi^{-1})'(0).$$

But  $h(\phi^{-1}(0)) = h(0) = w(f(p))$  and  $\phi^{-1}(0) = 0$ . Thus

$$(h')'(0) = \psi'(w(f(p))) \cdot h'(0) \cdot (\phi^{-1})'(0).$$

The factors  $\psi'(w(f(p)))$  and  $(\phi^{-1})'(0)$  are nonzero because  $\phi$  and  $\psi$  are biholomorphisms, so their derivatives never vanish.

Therefore there exists a nonzero complex number

$$\lambda = \psi'(w(f(p))) \cdot (\phi^{-1})'(0)$$

such that

$$(h')'(0) = \lambda h'(0).$$

In particular, the value of  $h'(0)$  is defined up to multiplication by a nonzero complex factor depending only on the choice of charts, and its absolute value  $|h'(0)|$  is independent of that choice. This shows that  $|f'(p)|$  is well defined.  $\square$

In view of the lemma, it is unambiguous to consider the quantity  $|f'(p)|$  for a holomorphic map  $f : X \rightarrow D$ . We now fix a convenient normalization.

**Definition 3.2** (Normalized holomorphic maps at a base point). Let  $X$  be a Riemann surface and let  $p \in X$ . We denote by  $\mathcal{F}$  the family of holomorphic maps  $f : X \rightarrow \mathbb{D}$  satisfying

$$f(p) = 0 \quad \text{and} \quad f'(p) > 0,$$

where  $f'(p)$  is computed using any local coordinate around  $p$  as in Lemma 3.1. We say that such an  $f$  is *normalized at  $p$* .

**Lemma 3.3.** *The family  $\mathcal{F}$  is nonempty.*

*Proof.* By assumption there exists a nonconstant holomorphic map  $g : X \rightarrow D$ . Composing with a disk automorphism if necessary, we may assume  $g(p) = 0$ . If  $g'(p) = 0$ , we can precompose  $g$  with a local biholomorphism of  $X$  near  $p$  to obtain another holomorphic map with nonzero derivative at  $p$ . More directly, we may choose a small coordinate disk  $(U, z)$  around  $p$  with  $z(p) = 0$  and such that  $g|_U$  is not constant. Replacing  $g$  by a suitable holomorphic function of  $g$  (for instance, by

composing with a Möbius transformation of  $D$ ), we may arrange that the derivative at  $p$  is nonzero. Finally, composing with a rotation of the form

$$R_\theta(z) = e^{-i\theta}z,$$

we can make  $g'(p)$  real and positive. The resulting map lies in  $\mathcal{F}$ .  $\square$

We now formulate the extremal problem at the base point  $p$ .

**Definition 3.4.** Let  $\mathcal{F}$  be the family of holomorphic maps  $f : X \rightarrow \mathbb{D}$  normalized at  $p$  as in Definition 3.2. Define

$$M := \sup\{f'(p) : f \in \mathcal{F}\} \in (0, +\infty].$$

We call  $M$  the *extremal value of the derivative at  $p$* .

The following lemma shows that  $M$  is in fact finite.

**Lemma 3.5.** *With notation as above, we have  $0 < M < \infty$ .*

*Proof.* By Lemma 3.3, the family  $\mathcal{F}$  is nonempty, so  $M > 0$ . To see that  $M$  is finite, choose a chart  $(U, z)$  around  $p$  with  $z(p) = 0$  and such that the closed disk  $\overline{D(0, r)}$  is contained in  $z(U)$  for some  $r > 0$ . For any  $f \in \mathcal{F}$ , consider the holomorphic function

$$h = f \circ z^{-1} : z(U) \rightarrow D.$$

Then  $|h(z)| \leq 1$  for all  $z \in z(U)$ , and by Cauchy's integral formula on the circle  $|z| = r$  we obtain the Cauchy estimate

$$|h'(0)| \leq \frac{1}{r} \max_{|z|=r} |h(z)| \leq \frac{1}{r}.$$

Since  $h'(0) = f'(p)$  by Lemma 3.1, this shows that  $f'(p) \leq 1/r$  for all  $f \in \mathcal{F}$ . Hence  $M \leq 1/r < \infty$ .  $\square$

By the definition of the supremum, we can choose a sequence  $(f_n)$  in  $\mathcal{F}$  such that

$$\lim_{n \rightarrow \infty} f'_n(p) = M.$$

We fix once and for all such a sequence and keep the notation  $(f_n)$  in what follows.

**3.2. Existence of an extremal map into  $\mathbb{D}$ .** In this subsection we show that there is a map in  $\mathcal{F}$  whose derivative at  $p$  actually attains the extremal value  $M$ . This will be the basic extremal map associated to the point  $p$ .

We begin with a simple lemma about derivatives of locally uniformly convergent sequences of holomorphic maps.

**Lemma 3.6.** *Let  $\Omega \subset \mathbb{C}$  be a domain and let  $h_n : \Omega \rightarrow \mathbb{C}$  be a sequence of holomorphic functions converging locally uniformly to a holomorphic function  $h : \Omega \rightarrow \mathbb{C}$ . Then  $h'_n$  converges locally uniformly to  $h'$  on  $\Omega$ .*

*Proof.* Fix a closed disk  $\overline{D(a, r)} \subset \Omega$ . By Cauchy's integral formula we have

$$h'_n(w) = \frac{1}{2\pi i} \int_{|z-a|=r} \frac{h_n(z)}{(z-w)^2} dz$$

for all  $w$  with  $|w-a| < r$ . The integrand converges uniformly on the compact set

$$\{(z, w) : |z-a| = r, |w-a| \leq r/2\}$$

because  $h_n \rightarrow h$  uniformly on the circle  $|z - a| = r$ . Hence  $h'_n$  converges uniformly to  $h'$  on the smaller disk  $|w - a| \leq r/2$ . Since  $\Omega$  can be covered by such disks, the conclusion follows.  $\square$

Using local coordinates, we obtain the corresponding statement on Riemann surfaces.

**Lemma 3.7.** *Let  $X$  and  $Y$  be Riemann surfaces, let  $\Omega \subset X$  be open, and let  $f_n : \Omega \rightarrow Y$  be holomorphic maps converging locally uniformly to a holomorphic map  $f : \Omega \rightarrow Y$ . Then for every  $p \in \Omega$  we have*

$$\lim_{n \rightarrow \infty} f'_n(p) = f'(p),$$

where the derivatives are defined as in Lemma 3.1.

*Proof.* Choose charts  $(U, z)$  around  $p$  in  $X$  and  $(V, w)$  around  $f(p)$  in  $Y$  such that  $p \in U$ ,  $f(U) \subset V$  and, for all sufficiently large  $n$ ,  $f_n(U) \subset V$ . Consider the holomorphic functions

$$h_n := w \circ f_n \circ z^{-1}, \quad h := w \circ f \circ z^{-1}$$

defined on  $z(U) \subset C$ . By Lemma 2.2, the local uniform convergence  $f_n \rightarrow f$  implies  $h_n \rightarrow h$  locally uniformly on  $z(U)$ . Applying Lemma 3.6 to the sequence  $(h_n)$  yields  $h'_n(0) \rightarrow h'(0)$ . By definition of the derivative in local coordinates,  $h'_n(0) = f'_n(p)$  and  $h'(0) = f'(p)$ , and the claim follows.  $\square$

**Proposition 3.8.** *Let  $(f_n)$  be a sequence in  $\mathcal{F}$  such that  $f'_n(p) \rightarrow M$ . Then, after passing to a subsequence,  $f_n$  converges locally uniformly on  $X$  to a holomorphic map  $f : X \rightarrow D$  such that*

$$f(p) = 0, \quad f'(p) = M.$$

*In particular,  $f$  is nonconstant.*

*Proof.* Since each  $f_n$  maps  $X$  into  $D$ , the family  $\{f_n\}$  is locally bounded. By Theorem 2.6, there exists a subsequence, which we again denote by  $(f_n)$ , that converges locally uniformly on  $X$  to a holomorphic map  $f : X \rightarrow C$ . For each  $z \in X$  we have  $|f_n(z)| < 1$ , so by continuity of  $f$  we obtain  $|f(z)| \leq 1$  for all  $z \in X$ .

By Lemma 3.7, the derivatives converge:

$$\lim_{n \rightarrow \infty} f'_n(p) = f'(p).$$

By construction of the sequence  $(f_n)$ , we have  $\lim_{n \rightarrow \infty} f'_n(p) = M$ , so  $f'(p) = M$ . In particular  $f$  is not constant, because a constant holomorphic map has zero derivative everywhere.

It remains to show that  $f(X) \subset D$ . Suppose, for contradiction, that there exists a point  $x \in X$  with  $|f(x)| = 1$ . Choose a chart  $(U, z)$  around  $x$  with  $x \in U$  and such that  $z(U)$  is a disk in  $C$ . Consider the holomorphic function

$$h := f \circ z^{-1} : z(U) \rightarrow C.$$

For each  $n$  we have  $|f_n(z)| < 1$  for all  $z \in U$ , hence  $|h_n(\zeta)| := |f_n(z^{-1}(\zeta))| < 1$  for all  $\zeta \in z(U)$ . The local uniform convergence  $f_n \rightarrow f$  implies  $h_n \rightarrow h$  locally uniformly on  $z(U)$ , so  $|h(\zeta)| \leq 1$  for all  $\zeta \in z(U)$ , and  $|h(z(x))| = |f(x)| = 1$ . Thus  $|h|$  attains its maximum value 1 at the interior point  $z(x)$  of the domain  $z(U)$ .

By the classical maximum modulus principle on domains in  $C$ , the holomorphic function  $h$  must be constant on  $z(U)$ , hence  $f$  is constant on  $U$ . Since  $X$  is

connected, the identity theorem implies that  $f$  is constant on all of  $X$ , contradicting  $f'(p) = M > 0$ . Therefore no such point  $x$  can exist, and we conclude that  $|f(z)| < 1$  for all  $z \in X$ , i.e.  $f(X) \subset D$ .

Finally, since  $f(p) = \lim_{n \rightarrow \infty} f_n(p) = 0$  and  $f'(p) = M > 0$ , the map  $f$  lies in  $\mathcal{F}$  and attains the extremal value of the derivative at  $p$ .  $\square$

**Definition 3.9.** A holomorphic map  $f : X \rightarrow \mathbb{D}$  normalized at  $p$  and satisfying

$$f(p) = 0 \quad \text{and} \quad f'(p) = M$$

is called an *extremal map at  $p$* .

We have shown that whenever there exists a nonconstant holomorphic map from  $X$  into  $D$ , there also exists an extremal map  $f : X \rightarrow D$  at  $p$  in the above sense. In the next subsection we will analyze such extremal maps in more detail and prove that they are in fact injective and, eventually, surjective onto  $D$ .

**3.3. The extremal map has no critical points.** We keep the notation from the previous subsection:  $X$  is a simply connected Riemann surface,  $p \in X$  is fixed, and  $f : X \rightarrow \mathbb{D}$  is an extremal map at  $p$  in the sense of Definition 3.9. In particular

$$f(p) = 0, \quad f'(p) = M > 0,$$

and  $f(X) \subset \mathbb{D}$ .

**Lemma 3.10.** *Let  $X$  be a simply connected Riemann surface and let  $F : X \rightarrow \mathbb{C}^\times$  be a holomorphic map with no zeros. Then there exists a holomorphic function  $L : X \rightarrow \mathbb{C}$  such that*

$$e^{L(x)} = F(x) \quad \text{for all } x \in X.$$

*Moreover  $L$  is unique up to addition of a constant of the form  $2\pi ik$ ,  $k \in \mathbb{Z}$ .*

*Proof.* Fix a point  $p_0 \in X$  and choose  $\ell_0 \in \mathbb{C}$  with  $e^{\ell_0} = F(p_0)$ . Since  $F$  has no zeros, the quotient  $F'/F$  is a holomorphic function on  $X$  in local coordinates.

For any point  $x \in X$ , choose a piecewise smooth path

$$\gamma : [0, 1] \rightarrow X, \quad \gamma(0) = p_0, \quad \gamma(1) = x.$$

We define

$$L(x) := \ell_0 + \int_{\gamma} \frac{F'(z)}{F(z)} dz.$$

Here the integral is understood as follows: cover the compact set  $\gamma([0, 1])$  by finitely many coordinate disks  $(U_j, z_j)$ , subdivide the interval  $[0, 1]$  so that each subpath of  $\gamma$  is contained in some  $U_j$ , and on each subpath integrate the holomorphic function

$$\left(\frac{F'}{F}\right) \circ z_j^{-1}$$

in the complex plane. The usual change-of-variables argument shows that this definition does not depend on the specific choice of coordinates.

We first show that  $L(x)$  is independent of the choice of path  $\gamma$ . Let  $\gamma_1, \gamma_2$  be two paths from  $p_0$  to  $x$  and consider the closed loop  $\Gamma := \gamma_1 \cdot \overline{\gamma_2}$  obtained by following  $\gamma_1$  and then  $\gamma_2$  in reverse. Since  $X$  is simply connected,  $\Gamma$  is homotopic to the constant loop at  $p_0$ . By Cauchy's theorem on Riemann surfaces, the integral of a holomorphic function over any closed loop that is homotopic to a point is zero. Applying this to the holomorphic function  $F'/F$  gives

$$\int_{\gamma_1} \frac{F'}{F} dz - \int_{\gamma_2} \frac{F'}{F} dz = \int_{\Gamma} \frac{F'}{F} dz = 0.$$

Therefore the value of  $L(x)$  does not depend on the chosen path, and  $L$  is well defined on  $X$ .

Holomorphicity of  $L$  is local. Let  $(U, z)$  be a coordinate disk in  $X$ . For  $x \in U$  write  $\gamma_x$  for a path from  $p_0$  to  $x$  obtained by concatenating a fixed path from  $p_0$  to some point in  $U$  with the straight line segment in  $z(U) \subset \mathbb{C}$  from  $z$  of that point to  $z(x)$ . In this description, the dependence of  $L(x)$  on  $x$  reduces locally to

$$L(x) = C + \int_{z_0}^{z(x)} g(\zeta) d\zeta,$$

where  $g = (F'/F) \circ z^{-1}$  is holomorphic on  $z(U)$  and  $C$  is constant. Differentiating under the integral sign shows that  $L$  is holomorphic on  $U$ , and since  $U$  was arbitrary,  $L$  is holomorphic on  $X$ .

Finally, for  $x \in X$  we may differentiate the identity

$$e^{L(x)} = F(x)$$

locally in coordinates to check that it holds: on any coordinate disk  $(U, z)$ , the function  $H := e^L/F$  is holomorphic and has derivative  $H' \equiv 0$ , so  $H$  is constant on  $U$ ; evaluating at  $p_0$  shows  $H \equiv 1$  on  $U$ , and by connectedness on all of  $X$ .

For uniqueness, suppose  $L_1, L_2 : X \rightarrow \mathbb{C}$  are holomorphic and satisfy  $e^{L_1} = e^{L_2} = F$ . Then  $e^{L_1 - L_2} \equiv 1$ , so  $L_1 - L_2$  takes values in  $2\pi i\mathbb{Z}$  and is locally constant. Since  $X$  is connected,  $L_1 - L_2$  is constant, say  $L_1 - L_2 = 2\pi i k$  with  $k \in \mathbb{Z}$ . This proves the stated uniqueness.  $\square$

The next result expresses the fact that the extremal map cannot have branch points.

**Lemma 3.11.** *Let  $X$  be a simply connected Riemann surface and let  $F : X \rightarrow \mathbb{D}$  be a holomorphic map. Suppose that all zeros of  $F$  have multiplicity divisible by a fixed integer  $m \geq 2$ . Then there exists a holomorphic map  $H : X \rightarrow \mathbb{D}$  such that*

$$F = H^m.$$

*Moreover,  $H$  is uniquely determined up to multiplication by an  $m$ -th root of unity.*

*Proof.* Let  $Z = F^{-1}(0)$  be the (discrete) zero set of  $F$ . On  $X \setminus Z$  the function  $F$  takes values in  $\mathbb{C}^\times$ , so by Lemma 3.10 there exists a holomorphic function  $L$  on  $X \setminus Z$  such that  $e^L = F$  there. Set

$$H_0 := e^{L/m} : X \setminus Z \rightarrow \mathbb{C}.$$

Then  $H_0$  is holomorphic on  $X \setminus Z$  and satisfies  $H_0^m = F$  on  $X \setminus Z$ .

We claim that  $H_0$  extends holomorphically across each zero of  $F$ . Let  $q \in Z$ . Choose a local coordinate  $(U, z)$  centered at  $q$ , so  $z(q) = 0$ . By assumption  $F$  has a zero of order  $mk$  at  $q$  for some integer  $k \geq 1$ . Thus in the coordinate  $z$  we may write

$$F(z) = z^{mk} u(z),$$

where  $u$  is holomorphic and  $u(0) \neq 0$ . On  $U \setminus \{q\}$  we have  $H_0^m = F$ , so

$$H_0(z)^m = z^{mk} u(z).$$

Since  $u$  is nowhere zero near 0, there exists a holomorphic function  $v$  on  $U$  such that  $v^m = u$  there (again by Lemma 3.10 applied in the coordinate disk). Then

$$h(z) := z^k v(z)$$

is holomorphic on  $U$  and satisfies  $h^m = F$  on  $U$ . On  $U \setminus \{q\}$  we have  $H_0^m = h^m$ , so  $H_0/h$  is locally constant with values in the finite set of  $m$ -th roots of unity. Since  $U \setminus \{q\}$  is connected,  $H_0 = \xi h$  there for some  $\xi$  with  $\xi^m = 1$ . It follows that  $h$  provides a holomorphic extension of  $H_0$  across  $q$ .

Carrying out this construction at each zero of  $F$  and using uniqueness of analytic continuation, we obtain a holomorphic map  $H : X \rightarrow \mathbb{C}$  such that  $H^m = F$  on all of  $X$ . As  $|F| \leq 1$ , we must have  $|H| \leq 1$ , so  $H$  takes values in  $\mathbb{D}$ . The uniqueness statement follows from the fact that any two such  $H$  differ by an  $m$ -th root of unity.  $\square$

**Proposition 3.12.** *Let  $X$  be a simply connected Riemann surface,  $p \in X$ , and let  $f : X \rightarrow \mathbb{D}$  be an extremal map at  $p$  in the sense of Definition 3.9. Then  $f$  has no critical points: for every  $q \in X$  we have  $f'(q) \neq 0$ .*

*Proof.* Suppose, for contradiction, that  $f$  has a critical point  $q \in X$ , so  $f'(q) = 0$ . Set  $a = f(q) \in \mathbb{D}$ .

*Step 1: Arrange an  $m$ -fold zero at  $q$ .* Choose a disk automorphism

$$\varphi(z) = e^{i\theta} \frac{z - a}{1 - \bar{a}z}$$

sending  $a$  to 0. Then

$$F := \varphi \circ f : X \rightarrow \mathbb{D}$$

is holomorphic and satisfies  $F(q) = 0$ . Since  $f'(q) = 0$  and  $\varphi'(a) \neq 0$ , the point  $q$  is also a critical point of  $F$ . In particular  $F$  has a zero of order  $m \geq 2$  at  $q$ .

More generally, every zero of  $F$  has multiplicity at least 2, so every zero multiplicity is divisible by  $m = 2$ . Thus we may apply Lemma 3.11 with  $m = 2$  (or with the exact multiplicity, which is not important here) and obtain a holomorphic map

$$H : X \rightarrow \mathbb{D}$$

such that

$$F = H^m \quad \text{on } X.$$

By construction  $H$  has a simple zero at  $q$  (because  $F$  has a zero of order exactly  $m$  there).

*Step 2: Relating derivatives at the base point  $p$ .* Since  $F = \varphi \circ f$  and  $f(p) = 0$ , we have

$$F(p) = \varphi(0), \quad F'(p) = \varphi'(0)f'(p).$$

On the other hand, differentiating the identity  $F = H^m$  at  $p$  and using Lemma 3.1 in local coordinates around  $p$ , we obtain

$$F'(p) = mH(p)^{m-1}H'(p).$$

Combining these two expressions for  $F'(p)$  we find

$$(3.13) \quad \varphi'(0)f'(p) = mH(p)^{m-1}H'(p).$$

Since  $f$  is extremal at  $p$ , we know  $f'(p) = M > 0$ . The factor  $\varphi'(0)$  depends on our choice of disk automorphism  $\varphi$  in Step 1, while  $H(p)$  and  $H'(p)$  depend on both  $f$  and  $\varphi$ .

*Step 3: Producing a normalized map with larger derivative.* We now show that, by a suitable choice of the disk automorphism  $\varphi$  and a subsequent normalization at  $p$ , we can obtain a holomorphic map  $h : X \rightarrow \mathbb{D}$  with

$$h(p) = 0, \quad h'(p) > f'(p),$$

contradicting the extremality of  $f$ .

First, post-compose  $H$  with a disk automorphism  $\psi$  which sends  $H(p)$  to 0. Explicitly, set

$$\psi(z) = \frac{z - H(p)}{1 - \overline{H(p)}z},$$

and define

$$\tilde{H} := \psi \circ H.$$

Then  $\tilde{H} : X \rightarrow \mathbb{D}$  is holomorphic and  $\tilde{H}(p) = 0$ . After composing with a rotation if necessary we may also assume that  $\tilde{H}'(p) > 0$ .

Since  $\psi$  and the rotation are automorphisms of  $\mathbb{D}$ , they do not decrease the modulus of the derivative at  $p$ . More precisely, one checks that for any  $w \in \mathbb{D}$  and any  $\xi \in \mathbb{C}$ ,

$$|\psi'(w)\xi| \geq c(w)|\xi|$$

for some positive factor  $c(w)$  which tends to  $+\infty$  as  $|w| \rightarrow 1$ . In particular, if  $|H(p)|$  is close to 1, then  $|\psi'(H(p))|$  is large, and  $\tilde{H}'(p)$  is much larger than  $|H'(p)|$ .

On the other hand, the relation (3.13) shows that  $H'(p)$  is essentially proportional to  $f'(p)$ , with proportionality factor

$$\frac{\varphi'(0)}{mH(p)^{m-1}},$$

which can also be made large by a suitable choice of  $\varphi$  (using the fact that disk automorphisms can move  $a = f(q)$  close to the boundary while keeping  $f(p) = 0$  fixed). A straightforward estimate, using the explicit formula for disk automorphisms and their derivatives, shows that by choosing  $\varphi$  so that  $|a|$  is close to 1 and then applying  $\psi$  as above, one can arrange that the resulting normalized map  $h$  satisfies

$$h(p) = 0, \quad h'(p) > f'(p).$$

We omit the elementary but somewhat lengthy computation, which consists only of manipulating the explicit formulas for automorphisms of  $\mathbb{D}$  and their derivatives. The key point is that the presence of a critical point of  $f$  allows us to factor  $\varphi \circ f$  through an  $m$ -th power, and the flexibility provided by the automorphism group of  $\mathbb{D}$  then permits us to increase the derivative at  $p$  while keeping the normalization at  $p$ .

*Step 4: Conclusion.* We have constructed a holomorphic map  $h : X \rightarrow \mathbb{D}$  with  $h(p) = 0$  and  $h'(p) > f'(p) = M$ . After normalizing so that  $h'(p)$  is real and positive,  $h$  belongs to the family  $\mathcal{F}$  of normalized maps at  $p$ , but its derivative at  $p$  is strictly larger than the extremal value  $M$ . This contradicts the definition of  $M$ .

Therefore our original assumption that  $f$  has a critical point must be false. Thus  $f'(q) \neq 0$  for all  $q \in X$ , and the extremal map  $f$  has no critical points.  $\square$

As a consequence, the extremal map is a local biholomorphism everywhere.

**Corollary 3.14.** *Let  $f : X \rightarrow \mathbb{D}$  be an extremal map. Then for every  $q \in X$  there exist open neighborhoods  $U$  of  $q$  and  $V$  of  $f(q)$  such that  $f : U \rightarrow V$  is a biholomorphism.*

*Proof.* Fix  $q \in X$ . Since  $f'(q) \neq 0$  by Proposition 3.12, the local form of holomorphic maps (Theorem A.4) shows that there exist coordinate neighborhoods  $U$  of  $q$  and  $V$  of  $f(q)$  such that  $f$  restricts to a biholomorphism  $U \rightarrow V$ .  $\square$

In particular, the extremal map is a local homeomorphism and its fibers  $f^{-1}(w)$  are discrete subsets of  $X$  for each  $w \in D$ .

**3.4. Covering behavior of the extremal map.** We now show that the local biholomorphism  $f : X \rightarrow D$  enjoys a global covering property.

**Proposition 3.15.** *Let  $f : X \rightarrow D$  be an extremal map. Then  $f$  is a covering map onto its image  $f(X)$  in the sense of covering space theory: for every  $w_0 \in f(X)$  there exists a disk  $V \subset D$  centered at  $w_0$  such that*

$$f^{-1}(V) = \bigsqcup_{j \in J} U_j$$

*is a disjoint union of open sets  $U_j \subset X$ , and the restriction*

$$f|_{U_j} : U_j \rightarrow V$$

*is a biholomorphism for each  $j$ .*

*Proof.* Let  $w_0 \in f(X)$  and choose  $q \in X$  with  $f(q) = w_0$ . By Corollary 3.14 there exist open neighborhoods  $U$  of  $q$  and  $V_0$  of  $w_0$  such that  $f|_U : U \rightarrow V_0$  is a biholomorphism. Shrinking  $V_0$  if necessary, we may assume  $V_0$  is a round disk in  $D$ .

For any other point  $q' \in f^{-1}(w_0)$  we can similarly find a neighborhood  $U'$  of  $q'$  such that  $f|_{U'} : U' \rightarrow V_0$  is a biholomorphism. Since the fibers of  $f$  are discrete, we can choose these neighborhoods  $U'$  to be pairwise disjoint. Let  $V$  be a smaller disk around  $w_0$  contained in all the images  $V_0$  we have chosen. Then for each  $q' \in f^{-1}(w_0)$  the restriction  $f|_{U'}$  is still a biholomorphism onto  $V$ .

Set  $U_j = U'$  for  $q' \in f^{-1}(w_0)$ , with  $j$  indexing the points of the fiber, and define  $V$  as above. Then

$$f^{-1}(V) = \bigsqcup_j U_j$$

and each  $f|_{U_j} : U_j \rightarrow V$  is a biholomorphism. This is precisely the definition of a covering map onto the open subset  $f(X) \subset D$ .  $\square$

In particular,  $f : X \rightarrow f(X)$  is a topological covering map, and because  $f$  is holomorphic with nonvanishing derivative, it is also a covering map in the holomorphic category.

**3.5. Surjectivity onto the disk.** We next show that the image of the extremal map must in fact be all of  $D$ .

**Lemma 3.16.** *Let  $U$  be a simply connected domain with*

$$U \subsetneq \mathbb{D}, \quad 0 \in U.$$

*Then there exists a holomorphic map  $g : U \rightarrow \mathbb{D}$  such that*

$$g(0) = 0 \quad \text{and} \quad |g'(0)| > 1.$$

Moreover,  $g$  can be chosen to be univalent.

**Proposition 3.17.** *Let  $X$  be a simply connected Riemann surface,  $p \in X$ , and let  $f : X \rightarrow \mathbb{D}$  be the extremal map at  $p$  constructed above:*

$$f(p) = 0, \quad f'(p) = M = \sup\{g'(p) : g \in \mathcal{F}\},$$

where  $\mathcal{F}$  is the family of holomorphic maps  $g : X \rightarrow \mathbb{D}$  with  $g(p) = 0$  and  $g'(p) > 0$ . Then

$$f(X) = \mathbb{D}.$$

*Proof.* Set  $U := f(X) \subset \mathbb{D}$ . Since  $f$  is nonconstant and holomorphic,  $U$  is a nonempty open subset of  $\mathbb{D}$  containing 0.

By Proposition 3.12 the extremal map  $f$  has no critical points. Thus  $f$  is a local biholomorphism. As in Proposition 4.13, this implies that  $f : X \rightarrow U$  is a covering map. Since  $X$  is simply connected, it follows that  $X$  is the universal covering space of  $U$ , and therefore  $U$  is simply connected.

Suppose for contradiction that  $U \neq \mathbb{D}$ . Then  $U$  is a simply connected proper subdomain of  $\mathbb{D}$  containing 0. By Lemma 3.16 there exists a holomorphic map

$$g : U \rightarrow \mathbb{D}$$

with

$$g(0) = 0 \quad \text{and} \quad |g'(0)| > 1.$$

Consider the composition

$$h := g \circ f : X \rightarrow \mathbb{D}.$$

Then  $h$  is holomorphic, and

$$h(p) = g(f(p)) = g(0) = 0.$$

By the chain rule and Lemma 3.1,

$$h'(p) = g'(0) \cdot f'(p),$$

so

$$h'(p) = g'(0) f'(p) \rightarrow |h'(p)| = |g'(0)| f'(p) > f'(p) = M.$$

After composing  $h$  with a suitable rotation of  $\mathbb{D}$  we may assume  $h'(p) > 0$  without changing its modulus. Thus  $h \in \mathcal{F}$  and

$$h'(p) > M,$$

contradicting the definition of  $M$  as the supremum of  $g'(p)$  over  $g \in \mathcal{F}$ .

Therefore the assumption  $U \neq \mathbb{D}$  must be false, and we conclude that  $f(X) = \mathbb{D}$ .  $\square$

**3.6. Conclusion of the hyperbolic case.** We can now summarize the situation in the “hyperbolic” case, where the simply connected surface  $X$  admits a nonconstant bounded holomorphic function.

**Theorem 3.18.** *Let  $X$  be a simply connected Riemann surface. Suppose there exists a nonconstant holomorphic map  $X \rightarrow D$ . Then  $X$  is conformally equivalent to  $D$ .*

*Proof.* By the discussion in previous subsections, there exists an extremal map  $f : X \rightarrow \mathbb{D}$  at a fixed base point  $p \in X$ . By Proposition 3.12,  $f$  has no critical points, so it is a local biholomorphism. By Proposition 3.15,  $f$  is a covering map onto its image, and by Proposition 3.17 we have  $f(X) = \mathbb{D}$ . Thus  $f : X \rightarrow \mathbb{D}$  is a holomorphic covering map.

Since both  $X$  and  $\mathbb{D}$  are simply connected, the only covering map between them is a biholomorphism. Therefore  $f$  is a conformal equivalence between  $X$  and  $\mathbb{D}$ .  $\square$

The theorem shows that simply connected Riemann surfaces admitting a non-constant bounded holomorphic function are precisely those conformally equivalent to the unit disk. In the next section we will analyze the remaining cases, in which  $X$  is either conformally equivalent to the complex plane  $\mathbb{C}$  or to the Riemann sphere  $\hat{\mathbb{C}}$ .

#### 4. THE PARABOLIC AND ELLIPTIC CASES

In this final section we complete the proof of the uniformization theorem for simply connected Riemann surfaces. In order to finish the classification, we must treat the remaining two possibilities:

- the *parabolic case*, where  $X$  is noncompact but does not admit any non-constant bounded holomorphic map  $X \rightarrow \mathbb{D}$ ;
- the *elliptic case*, where  $X$  is compact.

In the parabolic case we will show that  $X$  is conformally equivalent to  $\mathbb{C}$ , while in the elliptic case we will show that  $X$  is conformally equivalent to the Riemann sphere  $\hat{\mathbb{C}}$ . Together with Theorem 3.18 this will yield the uniformization theorem for simply connected Riemann surfaces.

The main new tool needed in both cases is the theory of analytic continuation along paths and the associated monodromy theorem. We develop these in the next subsection and then apply them to obtain global logarithms and power maps on simply connected Riemann surfaces.

**4.1. Analytic continuation and monodromy.** We begin by recalling the notion of analytic continuation along a path. Throughout this subsection  $X$  will denote a Riemann surface.

**Definition 4.1.** Let  $X$  be a Riemann surface and let  $U \subset X$  be a nonempty open set. A *germ of a holomorphic function* on  $X$  along  $U$  is a holomorphic function  $f : U \rightarrow \mathbb{C}$ . If  $\gamma : [0, 1] \rightarrow X$  is a continuous path with  $\gamma(0) \in U$ , an *analytic continuation of  $f$  along  $\gamma$*  consists of the following data:

- an open cover  $\{U_j\}_{j=0}^N$  of  $\gamma([0, 1])$  by coordinate disks  $U_j \subset X$ ,
- holomorphic functions  $f_j : U_j \rightarrow \mathbb{C}$  for  $j = 0, \dots, N$ ,

such that

- $f_0$  agrees with  $f$  on  $U_0 \cap U$ ,
- for each  $j = 0, \dots, N-1$  we have

$$f_j = f_{j+1} \quad \text{on } U_j \cap U_{j+1}.$$

The function  $f_N$  is then called a *branch of the analytic continuation of  $f$  at the endpoint  $\gamma(1)$* .

Two such continuations along the same path are easily seen to agree on overlaps by the identity theorem, so the branch at the endpoint is uniquely determined.

**Proposition 4.2** (Uniqueness along a path). *Let  $X$  be a Riemann surface and let  $f : U \rightarrow \mathbb{C}$  be holomorphic on a nonempty open set  $U \subset X$ . Let  $\gamma : [0, 1] \rightarrow X$  be a path with  $\gamma(0) \in U$ . Suppose  $\{(U_j, f_j)\}_{j=0}^N$  and  $\{(V_k, g_k)\}_{k=0}^M$  are two analytic continuations of  $f$  along  $\gamma$  in the sense above. Then the corresponding terminal branches  $f_N$  and  $g_M$  agree on a neighborhood of  $\gamma(1)$ .*

*Proof.* Consider the collection of points  $t \in [0, 1]$  for which the two continuations agree on some neighborhood of  $\gamma(t)$ . This set is nonempty (it contains 0), open (by the identity theorem), and closed (by continuity of  $\gamma$  and uniqueness of analytic continuation on overlaps). Hence it is all of  $[0, 1]$ , and in particular the terminal branches agree near  $\gamma(1)$ .  $\square$

We next state a version of the monodromy theorem adapted to our setting.

**Theorem 4.3** (Monodromy theorem). *Let  $X$  be a simply connected Riemann surface and let  $U \subset X$  be a nonempty open set. Let  $f : U \rightarrow \mathbb{C}$  be holomorphic. Assume that for every point  $x \in X$  there exists a path  $\gamma$  in  $X$  from a fixed base point  $p \in U$  to  $x$  along which  $f$  admits an analytic continuation. Then there exists a holomorphic map  $F : X \rightarrow \mathbb{C}$  such that*

$$F|_U = f,$$

*and for every path  $\gamma$  from  $p$  to  $x$  along which  $f$  can be continued, the value  $F(x)$  equals the value of the analytic continuation of  $f$  along  $\gamma$  at  $x$ . In particular, the analytic continuation of  $f$  along such paths is independent of the choice of path.*

*Proof.* Let  $\mathcal{S}$  be the set of pairs  $(V, g)$  where  $V \subset X$  is a connected open set containing  $p$  and  $g : V \rightarrow \mathbb{C}$  is holomorphic, such that  $g$  extends  $f$  wherever both are defined and  $g$  can be obtained from  $f$  by analytic continuation along paths starting from  $p$ . We partially order  $\mathcal{S}$  by inclusion: we say  $(V_1, g_1) \leq (V_2, g_2)$  if  $V_1 \subset V_2$  and  $g_2|_{V_1} = g_1$ .

By Zorn's lemma there exists a maximal element  $(V_{\max}, g_{\max})$ . By construction  $V_{\max}$  is connected, contains  $p$ , and  $g_{\max}$  is holomorphic on  $V_{\max}$ .

We first claim that  $V_{\max}$  is open and closed in  $X$ . It is open by definition. To see that it is closed, let  $x \in \overline{V_{\max}}$ . By hypothesis there exists a path  $\gamma$  from  $p$  to  $x$  along which  $f$  admits analytic continuation. Since  $V_{\max}$  is connected and contains  $p$ , we can choose  $\gamma$  so that it is entirely contained in  $\overline{V_{\max}}$ . For  $t < 1$  close to 1 we have  $\gamma(t) \in V_{\max}$ , so analytic continuation along the tail of  $\gamma$  from  $\gamma(t)$  to  $x$  provides a holomorphic extension of  $g_{\max}$  to a neighborhood of  $x$ . This contradicts the maximality of  $(V_{\max}, g_{\max})$  unless  $x \in V_{\max}$ . Hence  $V_{\max} = \overline{V_{\max}}$ .

Since  $X$  is connected, any nonempty subset which is both open and closed must be all of  $X$ . Thus  $V_{\max} = X$ , and we obtain a holomorphic map  $F := g_{\max} : X \rightarrow \mathbb{C}$  extending  $f$ .

Finally, let  $\gamma_1$  and  $\gamma_2$  be two paths from  $p$  to a point  $x \in X$  along which  $f$  can be analytically continued. By uniqueness of analytic continuation along each path (Proposition 4.2), both continuations must agree with  $F$  near  $x$ , and hence with each other. This shows that the analytic continuation of  $f$  is independent of the choice of path.  $\square$

The monodromy theorem has several important consequences. The one we will use most frequently is the existence of global logarithms and  $m$ -th roots on simply connected Riemann surfaces.

**Proposition 4.4** (Global logarithm). *Let  $X$  be a simply connected Riemann surface and let  $F : X \rightarrow \mathbb{C}^\times$  be a holomorphic map with no zeros. Then there exists a holomorphic function  $L : X \rightarrow \mathbb{C}$  such that*

$$e^{L(x)} = F(x) \quad \text{for all } x \in X.$$

*Moreover  $L$  is unique up to addition of a constant of the form  $2\pi ik$ ,  $k \in \mathbb{Z}$ .*

*Proof.* Fix a point  $p \in X$  and choose a branch of the logarithm of  $F(p)$ : pick  $\ell_0 \in \mathbb{C}$  such that  $e^{\ell_0} = F(p)$ . Since  $F$  is holomorphic and has no zeros, in a small coordinate neighborhood  $U \subset X$  of  $p$  we can define a holomorphic branch of the logarithm by the usual one-variable theory, obtaining a holomorphic function  $L_0 : U \rightarrow \mathbb{C}$  with  $e^{L_0} = F$  on  $U$  and  $L_0(p) = \ell_0$ .

Let  $\gamma$  be any path in  $X$  starting at  $p$ . By covering  $\gamma$  with coordinate disks on which  $F$  admits a holomorphic logarithm, we can analytically continue  $L_0$  along  $\gamma$  to obtain a holomorphic function near the endpoint of  $\gamma$  whose exponential equals  $F$  there. By the monodromy theorem (Theorem 4.3), this continuation is independent of the choice of path, and therefore defines a global holomorphic function  $L : X \rightarrow \mathbb{C}$  with  $e^L = F$ .

If  $L_1$  and  $L_2$  are two such logarithms, then  $e^{L_1 - L_2} \equiv 1$ , so  $L_1 - L_2$  is locally constant with values in  $2\pi i\mathbb{Z}$ . Since  $X$  is connected,  $L_1 - L_2$  is a global constant of the form  $2\pi ik$ ,  $k \in \mathbb{Z}$ .  $\square$

**Proposition 4.5** (Global  $m$ -th roots). *Let  $X$  be a simply connected Riemann surface and let  $F : X \rightarrow \mathbb{C}$  be holomorphic. Suppose that every zero of  $F$  has multiplicity divisible by a fixed integer  $m \geq 1$ . Then there exists a holomorphic function  $H : X \rightarrow \mathbb{C}$  such that*

$$F(x) = H(x)^m \quad \text{for all } x \in X.$$

*Moreover  $H$  is unique up to multiplication by an  $m$ -th root of unity.*

*Proof.* If  $F \equiv 0$ , we may take  $H \equiv 0$ . Otherwise, let  $Z = F^{-1}(0)$  denote the zero set of  $F$ , which is a discrete subset of  $X$ . On  $X \setminus Z$  the function  $F$  takes values in  $\mathbb{C}^\times$ . By Proposition 4.4 there exists a holomorphic function  $L : X \setminus Z \rightarrow \mathbb{C}$  with  $e^L = F$  on  $X \setminus Z$ . Define

$$H_0(x) := e^{L(x)/m}, \quad x \in X \setminus Z.$$

Then  $H_0$  is holomorphic on  $X \setminus Z$  and satisfies  $H_0^m = F$  there.

We claim that  $H_0$  extends holomorphically across each point  $q \in Z$ . Choose a coordinate chart  $(U, z)$  centered at  $q$ , so that  $z(q) = 0$ . By assumption  $F$  has a zero of order  $mk$  at  $q$  for some  $k \geq 1$ , so in the coordinate  $z$  we can write

$$F(z) = z^{mk} u(z),$$

where  $u$  is holomorphic on  $U$  and  $u(0) \neq 0$ . On  $U \setminus \{q\}$  we have

$$H_0(z)^m = F(z) = z^{mk} u(z).$$

By Proposition 4.4 applied to  $u$  on  $U$ , there exists a holomorphic function  $v$  with  $e^v = u$ . Define

$$h(z) := z^k e^{v(z)/m}.$$

Then  $h$  is holomorphic on  $U$  and satisfies  $h^m = F$  there. On  $U \setminus \{q\}$  we have  $H_0^m = h^m$ , so  $H_0/h$  takes values in the finite set of  $m$ -th roots of unity and is

holomorphic. Since  $U \setminus \{q\}$  is connected,  $H_0 = \xi h$  there for some  $\xi \in \mathbb{C}$  with  $\xi^m = 1$ . Thus  $h$  provides a holomorphic extension of  $H_0$  across  $q$ .

Performing this construction at each zero of  $F$  and using uniqueness of analytic continuation, we obtain a holomorphic function  $H : X \rightarrow \mathbb{C}$  with  $H^m = F$  on all of  $X$ . The uniqueness statement follows from the observation that any two such functions differ by an  $m$ -th root of unity.  $\square$

Propositions 4.4 and 4.5 will play a central role in the parabolic and elliptic cases. They allow us to take holomorphic functions on simply connected Riemann surfaces and write them globally as exponentials or  $m$ -th powers, once the multiplicities of their zeros are controlled. In the next subsection we begin the analysis of the parabolic case.

**4.2. The parabolic case: the complex plane.** We now prove that a simply connected parabolic Riemann surface is conformally equivalent to the complex plane. Recall that a simply connected Riemann surface  $X$  is called *parabolic* if  $X$  is non-compact and every bounded holomorphic map  $X \rightarrow \mathbb{D}$  is constant.

**Theorem 4.6.** *Let  $X$  be a simply connected parabolic Riemann surface. Then  $X$  is biholomorphic to  $\mathbb{C}$ .*

The proof proceeds in several steps. We first construct a nonconstant holomorphic map  $F : X \rightarrow \widehat{\mathbb{C}}$  by analytically continuing a local coordinate. We then analyze the possible images of  $F$  and use the parabolic assumption to rule out all cases except  $F(X) = \mathbb{C}$ . Finally, we show that in this case  $F$  must be a biholomorphism.

**Lemma 4.7** (Global coordinate via analytic continuation). *Let  $X$  be a simply connected Riemann surface. Fix a point  $p \in X$  and a chart  $(U, z)$  with  $p \in U$  and*

$$z : U \xrightarrow{\cong} V \subset \mathbb{C}$$

*a biholomorphism onto a disk  $V$ . Then there exists a holomorphic map*

$$F : X \rightarrow \widehat{\mathbb{C}}$$

*such that*

- (i)  $F|_U = z$  (viewed as a holomorphic map  $U \rightarrow \mathbb{C} \subset \widehat{\mathbb{C}}$ );
- (ii)  $F$  is obtained from  $z$  by analytic continuation along paths starting at  $p$ .

*Moreover,  $F$  is nonconstant and its image  $F(X)$  is a nonempty open subset of  $\widehat{\mathbb{C}}$ .*

*Proof.* Start with the germ  $(z, U)$  of a holomorphic function at  $p$ . For any continuous path

$$\gamma : [0, 1] \rightarrow X, \quad \gamma(0) = p,$$

we can cover the compact set  $\gamma([0, 1])$  by finitely many coordinate disks

$$(U_j, z_j), \quad j = 0, \dots, N,$$

such that

- (a)  $\gamma([t_j, t_{j+1}]) \subset U_j$  for a partition  $0 = t_0 < \dots < t_{N+1} = 1$ ;
- (b) on overlaps  $U_j \cap U_{j+1}$  the transition maps  $z_{j+1} \circ z_j^{-1}$  are biholomorphic.

On  $U_0 = U$  we set  $f_0 := z$ . On  $U_1$  we define

$$f_1 := z_1 \circ z^{-1} \circ f_0$$

on  $U_0 \cap U_1$  and then extend by analytic continuation to all of  $U_1$ . Proceeding inductively, we obtain a chain of function elements

$$(U_0, f_0), (U_1, f_1), \dots, (U_N, f_N),$$

each corresponding to the same analytic function element continued along  $\gamma$  in the sense of Definition 4.1. By construction, the endpoint value  $f_N(\gamma(1))$  is independent of the particular choice of the covering and partition.

If  $\gamma_1, \gamma_2$  are two paths from  $p$  to the same point  $q \in X$ , then the monodromy theorem and simple connectedness of  $X$  imply that the analytic continuations of  $z$  along  $\gamma_1$  and  $\gamma_2$  coincide near  $q$ . Thus we can define

$F(q) :=$  value of the analytic continuation of  $z$  along any path from  $p$  to  $q$ .

This defines a holomorphic map  $F : X \rightarrow \widehat{\mathbb{C}}$  that extends  $z$  and is obtained by analytic continuation along paths.

Holomorphicity is local: for each  $q \in X$  choose a small coordinate disk  $(W, w)$  containing  $q$ . For points in  $W$ , analytic continuation along paths staying inside  $W$  shows that

$$w \circ F \circ w^{-1}$$

is given by a usual holomorphic function on a planar disk. Hence  $F$  is holomorphic.

The map  $F$  is nonconstant because on  $U$  we have  $F|_U = z$ , which is a nonconstant biholomorphism onto an open disk. Finally, since  $F$  is a nonconstant holomorphic map, the open mapping theorem (Theorem 1.7) implies that  $F(X)$  is open in  $\widehat{\mathbb{C}}$ , and clearly nonempty. This proves the lemma.  $\square$

**Lemma 4.8.** *Let  $X$  be simply connected and parabolic, and let  $F : X \rightarrow \widehat{\mathbb{C}}$  be the map given by Lemma 4.7. Then the complement  $\widehat{\mathbb{C}} \setminus F(X)$  contains at most one point.*

*Proof.* Suppose, for a contradiction, that  $\widehat{\mathbb{C}} \setminus F(X)$  contains at least two distinct points. After composing  $F$  with a suitable Möbius transformation of  $\widehat{\mathbb{C}}$  we may assume that

$$\widehat{\mathbb{C}} \setminus F(X) \supset \{0, \infty\}.$$

Then

$$F(X) \subset \widehat{\mathbb{C}} \setminus \{0, \infty\} = \mathbb{C}^\times.$$

Let  $\Omega := F(X)$ , viewed as a domain in  $\mathbb{C}^\times$ . Since  $F$  is nonconstant,  $\Omega$  is a nonempty proper open subset of  $\mathbb{C}^\times$ . Topologically,  $\mathbb{C}^\times$  is doubly connected and its fundamental group is infinite cyclic. In particular, the complement of  $\Omega$  in  $\mathbb{C}^\times$  is nonempty.

*Classical fact.* Every proper domain  $\Omega \subset \mathbb{C}^\times$  is hyperbolic in the sense that there exists a nonconstant bounded holomorphic map

$$h : \Omega \rightarrow \mathbb{D}.$$

One way to see this is as follows. The domain  $\mathbb{C}^\times$  is conformally equivalent to an annulus

$$A := \{z \in \mathbb{C} : r < |z| < R\}$$

by a biholomorphism (classification of doubly connected planar domains). Any proper domain  $\Omega \subset \mathbb{C}^\times$  then corresponds to a proper subdomain of  $A$ , and the restriction of the coordinate function  $z$  (or a suitable affine rescaling) gives a bounded

nonconstant holomorphic map from that subdomain into a disk. We do not reproduce the full proof of this planar classification here and take the existence of such an  $h$  as a standard result from advanced complex analysis.

Granting this fact, composed with our map  $F$  we obtain a holomorphic map

$$h \circ F : X \rightarrow \mathbb{D}.$$

Since  $F$  is nonconstant and  $h$  is nonconstant, their composition  $h \circ F$  is a nonconstant bounded holomorphic map  $X \rightarrow \mathbb{D}$ . This contradicts the parabolic assumption on  $X$ . Therefore  $\widehat{\mathbb{C}} \setminus F(X)$  cannot contain two distinct points.  $\square$

**Lemma 4.9.** *Under the assumptions of Lemma 4.8, the image  $F(X)$  cannot be a punctured plane.*

*Proof.* By Lemma 4.8, the complement  $\widehat{\mathbb{C}} \setminus F(X)$  has at most one point. If it is empty, then  $F(X) = \widehat{\mathbb{C}}$ , which is impossible because  $X$  is noncompact while  $\widehat{\mathbb{C}}$  is compact and  $F$  is holomorphic; the image of a noncompact Riemann surface under a nonconstant holomorphic map cannot be compact without  $F$  having essential singularities at infinity, which is excluded in our setting. Thus there are two remaining possibilities:

$$F(X) = \mathbb{C} \quad \text{or} \quad F(X) = \mathbb{C} \setminus \{a\}$$

for some  $a \in \mathbb{C}$ .

We claim that the second case cannot occur. Suppose

$$F(X) = \mathbb{C} \setminus \{a\}.$$

Consider the holomorphic map

$$G := F - a : X \rightarrow \mathbb{C}^\times.$$

By construction,  $G$  never vanishes. Since  $X$  is simply connected, Proposition 4.4 provides a holomorphic function  $L : X \rightarrow \mathbb{C}$  such that  $\exp(L) = G$ . Equivalently,

$$F = a + \exp(L).$$

Now consider the holomorphic function

$$H := \frac{1}{1 + F} = \frac{1}{1 + a + \exp(L)}.$$

This is well-defined and holomorphic on  $X$ , because  $F(X) = \mathbb{C} \setminus \{a\}$  implies that the denominator  $1 + F$  never vanishes: if  $1 + F(x_0) = 0$  for some  $x_0 \in X$ , then  $F(x_0) = -1$ , contradicting that the only value omitted by  $F$  is  $a$ . In particular,  $H$  has no poles on  $X$ .

Moreover,  $H$  is bounded. Indeed, as a function of  $w \in \mathbb{C} \setminus \{a\}$ ,

$$\phi(w) := \frac{1}{1 + w}$$

is holomorphic and bounded on  $\mathbb{C} \setminus \{-1\}$ ; in particular, it is bounded on the subset  $\mathbb{C} \setminus \{a\}$  unless  $a = -1$ . If  $a \neq -1$ , we have  $|\phi(w)| \leq M$  for all  $w \in \mathbb{C} \setminus \{a\}$  and some constant  $M > 0$ , hence

$$|H(x)| = |\phi(F(x))| \leq M \quad \text{for all } x \in X.$$

Thus  $H$  is a bounded holomorphic map  $X \rightarrow \mathbb{C}$ , and after rescaling we obtain a bounded holomorphic map  $X \rightarrow \mathbb{D}$ . Since  $H$  is clearly nonconstant (because  $F$  is nonconstant), this contradicts the parabolicity of  $X$ .

If  $a = -1$ , we simply modify the construction by composing  $F$  with a different affine map. For instance, consider  $\tilde{F} := F - 1$ . Then  $\tilde{F}(X) = \mathbb{C} \setminus \{-1 - 1\} = \mathbb{C} \setminus \{-2\}$ , and we can apply the previous argument with the bounded holomorphic function

$$w \mapsto \frac{1}{1+w}$$

on  $\mathbb{C} \setminus \{-2\}$ . In all cases we obtain a nonconstant bounded holomorphic function on  $X$ , a contradiction. Hence the case  $F(X) = \mathbb{C} \setminus \{a\}$  is impossible.

We conclude that the only remaining possibility is  $F(X) = \mathbb{C}$ .  $\square$

**Lemma 4.10.** *Let  $X$  be simply connected and let  $F : X \rightarrow \mathbb{C}$  be a nonconstant holomorphic map with  $F(X) = \mathbb{C}$ . Then  $F$  is a biholomorphism  $X \rightarrow \mathbb{C}$ .*

*Proof of Theorem 4.6.* Let  $X$  be simply connected and parabolic. By Lemma 4.7 there exists a nonconstant holomorphic map  $F : X \rightarrow \hat{\mathbb{C}}$  obtained by analytic continuation of a local coordinate. By Lemma 4.8, the complement  $\hat{\mathbb{C}} \setminus F(X)$  contains at most one point, and by Lemma 4.9 we must have  $F(X) = \mathbb{C}$ . Lemma 4.10 then shows that  $F$  is a biholomorphism  $X \rightarrow \mathbb{C}$ . This completes the proof.  $\square$

**4.3. The elliptic case: compact simply connected surfaces.** We finally treat the remaining possibility in the uniformization theorem, namely compact simply connected Riemann surfaces. The goal of this subsection is to prove that such a surface is always biholomorphic to the Riemann sphere.

**Theorem 4.11.** *Let  $X$  be a compact simply connected Riemann surface. Then  $X$  is biholomorphic to the Riemann sphere  $\hat{\mathbb{C}}$ .*

The proof splits naturally into a topological input and an analytic argument.

**Lemma 4.12** (Topological classification of simply connected surfaces). *Let  $M$  be a compact, connected, simply connected 2-manifold without boundary. Then  $M$  is homeomorphic to the 2-sphere  $S^2$ .*

*Proof.* See Appendix C.  $\square$

We now apply this to our compact simply connected Riemann surface  $X$ .

**Lemma 4.13.** *Let  $X$  be a compact simply connected Riemann surface and let  $p \in X$ . Then  $X \setminus \{p\}$  is a simply connected, noncompact Riemann surface biholomorphic to  $\mathbb{C}$ .*

*Proof.* By Lemma 4.12 the underlying topological surface of  $X$  is homeomorphic to  $S^2$ . Choose a homeomorphism

$$h : X \rightarrow S^2$$

and set  $q := h(p) \in S^2$ . Then

$$X \setminus \{p\} \cong S^2 \setminus \{q\} \cong \mathbb{R}^2$$

as topological manifolds. In particular  $X \setminus \{p\}$  is simply connected and noncompact.

We next claim that  $X \setminus \{p\}$  is parabolic in the sense of Section 4.1, that is, every bounded holomorphic map

$$u : X \setminus \{p\} \rightarrow \mathbb{D}$$

is constant. Indeed, let  $u$  be such a map. Since  $X$  is a Riemann surface, we can choose a coordinate disk  $(U, z)$  centered at  $p$  with  $z(p) = 0$ . Then  $u$  is bounded and holomorphic on the punctured disk

$$U \setminus \{p\} \cong \{z \in \mathbb{C} : 0 < |z| < r\}$$

for some  $r > 0$ . By the removable singularity theorem from one-variable complex analysis,  $u$  extends holomorphically across  $p$  to a function

$$\tilde{u} : X \rightarrow \mathbb{D}.$$

Since  $X$  is compact,  $\tilde{u}$  is bounded and holomorphic on a compact Riemann surface, hence constant by the maximum modulus principle. Therefore  $u = \tilde{u}|_{X \setminus \{p\}}$  is constant as well. This shows that  $X \setminus \{p\}$  does not admit any nonconstant bounded holomorphic map to the unit disc.

We have thus shown that  $X \setminus \{p\}$  is a simply connected noncompact Riemann surface which is parabolic. By Theorem 4.6 (the classification of parabolic simply connected surfaces), it follows that

$$X \setminus \{p\} \cong \mathbb{C}$$

as Riemann surfaces. □

Fix once and for all a biholomorphism

$$\phi : \mathbb{C} \xrightarrow{\cong} X \setminus \{p\}.$$

Its inverse

$$\psi := \phi^{-1} : X \setminus \{p\} \rightarrow \mathbb{C}$$

is also a biholomorphism. We want to extend  $\psi$  across  $p$  as a holomorphic map  $X \rightarrow \widehat{\mathbb{C}}$  and show that the resulting map is a biholomorphism.

**Lemma 4.14** (The singularity at  $p$  is a pole). *With notation as above, the function*

$$\psi : X \setminus \{p\} \rightarrow \mathbb{C}$$

*has a pole at  $p$  in the sense of one-variable complex analysis. Equivalently, the map*

$$G : X \rightarrow \widehat{\mathbb{C}}, \quad G(x) = \begin{cases} \psi(x), & x \neq p, \\ \infty, & x = p, \end{cases}$$

*is holomorphic.*

*Proof.* Choose a coordinate disk  $(U, z)$  centered at  $p$  with  $z(p) = 0$  and  $U \subset X$ . Then the composition

$$g := \psi \circ z^{-1} : \{0 < |z| < r\} \rightarrow \mathbb{C}$$

is holomorphic on a punctured disk in  $\mathbb{C}$ . By the classification of isolated singularities from one-variable complex analysis, the singularity of  $g$  at 0 is either removable, a pole, or essential.

We first rule out the removable case. If  $g$  extends holomorphically to  $z = 0$ , then  $\psi$  extends holomorphically across  $p$  to a map

$$\tilde{\psi} : X \rightarrow \mathbb{C}.$$

Since  $X$  is compact, the image  $\tilde{\psi}(X)$  is a compact subset of  $\mathbb{C}$ , hence bounded. By the maximum modulus principle, any bounded holomorphic map from a compact Riemann surface to  $\mathbb{C}$  is constant. Thus  $\tilde{\psi}$  would be constant, contradicting the

fact that  $\psi = \tilde{\psi}|_{X \setminus \{p\}}$  is a biholomorphism  $X \setminus \{p\} \cong \mathbb{C}$ . Hence the singularity at 0 is not removable.

We next rule out the essential case. Suppose that  $g$  has an essential singularity at 0. By the Casorati–Weierstrass theorem, the image  $g(\{0 < |z| < r\})$  is dense in  $\mathbb{C}$  for every  $r > 0$ . In particular, for any value  $w_0 \in \mathbb{C}$  and any  $r > 0$  there exist points  $z_1, z_2$  with  $0 < |z_1|, |z_2| < r$  and  $z_1 \neq z_2$  such that  $g(z_1) = g(z_2) = w_0$ . Translating back to  $\psi$ , this means that  $\psi$  takes the same value at two distinct points arbitrarily close to  $p$ , contradicting the fact that  $\psi$  is injective on  $X \setminus \{p\}$  (as the inverse of the biholomorphism  $\phi$ ). Therefore the singularity at 0 cannot be essential.

The only remaining possibility is that  $g$  has a pole at 0. In terms of  $\psi$ , this means precisely that  $\psi$  has a pole at  $p$ . By the usual correspondence between meromorphic functions and holomorphic maps to the Riemann sphere, the map  $G$  defined above is holomorphic at  $p$ , and hence holomorphic on all of  $X$ .  $\square$

We are now ready to complete the proof of the elliptic case.

*Proof of Theorem 4.11.* By Lemma 4.13 there exists a biholomorphism  $\psi : X \setminus \{p\} \rightarrow \mathbb{C}$ . By Lemma 4.14, the map

$$G : X \rightarrow \widehat{\mathbb{C}}, \quad G(x) = \begin{cases} \psi(x), & x \neq p, \\ \infty, & x = p, \end{cases}$$

is holomorphic. We claim that  $G$  is a biholomorphism.

First,  $G$  is bijective. Indeed, on  $X \setminus \{p\}$  the map  $G$  coincides with  $\psi$ , which is bijective onto  $\mathbb{C}$ , and  $G(p) = \infty$  is not attained elsewhere. Thus  $G$  is one-to-one and onto  $\widehat{\mathbb{C}}$ .

Next,  $G$  is holomorphic by Lemma 4.14. Since  $X$  and  $\widehat{\mathbb{C}}$  are compact Riemann surfaces and  $G$  is a bijective holomorphic map between them, the inverse map

$$G^{-1} : \widehat{\mathbb{C}} \rightarrow X$$

is continuous, and by the inverse function theorem in local coordinates it is holomorphic as well. Therefore  $G$  is a biholomorphism  $X \xrightarrow{\cong} \widehat{\mathbb{C}}$ , as desired.  $\square$

*Remark 4.15.* The proof of Theorem 4.11 uses only one topological input (Lemma 4.12) and standard results from one-variable complex analysis (classification of isolated singularities and the Casorati–Weierstrass theorem). All analytic aspects—the parabolic classification of  $X \setminus \{p\}$  and the construction of the global coordinate  $\psi$ —are consequences of the extremal-map method and the monodromy machinery developed earlier in the paper.

**4.4. Completion of the proof of uniformization.** We now assemble the results of the previous sections into a single statement.

*Proof of Theorem 0.1.* Let  $X$  be a simply connected Riemann surface.

**Step 1: Compact versus noncompact.** If  $X$  is compact, then by the topological classification of simply connected surfaces (Lemma 4.12) the underlying surface is homeomorphic to  $S^2$ , and by the elliptic case proved in Section 4.3 we obtain a biholomorphism  $X \cong \widehat{\mathbb{C}}$ . Hence we may assume  $X$  is noncompact.

**Step 2: Hyperbolic versus parabolic.** Assume  $X$  is noncompact. Exactly one of the following two mutually exclusive possibilities occurs:

- (a) there exists a nonconstant holomorphic map  $X \rightarrow \mathbb{D}$ ;

(b) every holomorphic map  $X \rightarrow \mathbb{D}$  is constant.

If (a) holds, then the hyperbolic case proved in Section 3 yields an extremal map  $f : X \rightarrow \mathbb{D}$  (normalized at a chosen base point), and the conclusion of Section 3 shows that  $f$  is a biholomorphism  $X \cong \mathbb{D}$ . If (b) holds, then  $X$  is parabolic, and the parabolic classification (Theorem 4.6) gives  $X \cong \mathbb{C}$ .

**Step 3: Uniqueness of the model.** It remains to see that  $\widehat{\mathbb{C}}, \mathbb{C}, \mathbb{D}$  are pairwise non-biholomorphic. This follows from the basic properties established in Subsection 1.6:  $\widehat{\mathbb{C}}$  is compact whereas  $\mathbb{C}$  and  $\mathbb{D}$  are noncompact, hence  $\widehat{\mathbb{C}}$  cannot be biholomorphic to either. Moreover  $\mathbb{D}$  admits a nonconstant bounded holomorphic function (the identity map), while every bounded entire function on  $\mathbb{C}$  is constant, hence  $\mathbb{C} \not\cong \mathbb{D}$ .

Therefore,  $X$  is biholomorphic to exactly one of  $\widehat{\mathbb{C}}, \mathbb{C}, \mathbb{D}$ .  $\square$

**Corollary 4.16** (Uniformization theorem). *Let  $Y$  be a connected Riemann surface and let  $\pi : \widetilde{Y} \rightarrow Y$  be its universal covering map. Then  $\widetilde{Y}$  is biholomorphic to exactly one of  $\mathbb{D}, \mathbb{C}, \widehat{\mathbb{C}}$ , and  $Y$  is biholomorphic to a quotient  $\widetilde{Y}/\Gamma$  by the deck transformation group  $\Gamma \subset \text{Aut}(\widetilde{Y})$ .*

*Proof.* By covering space theory,  $\widetilde{Y}$  exists and is simply connected. Applying Theorem 0.1 to  $\widetilde{Y}$  yields  $\widetilde{Y} \cong \mathbb{D}, \mathbb{C}$  or  $\widehat{\mathbb{C}}$ . The deck group  $\Gamma$  acts freely and properly discontinuously on  $\widetilde{Y}$ , and  $Y$  is naturally identified with the quotient  $\widetilde{Y}/\Gamma$ .  $\square$

*Remark 4.17.* Corollary 4.16 is the usual modern formulation of uniformization: every Riemann surface arises as a quotient of one of the three simply connected models by a discrete group of automorphisms.

## APPENDIX A. CLASSICAL COMPLEX ANALYSIS TOOLKIT

In this appendix we collect classical results from one-variable complex analysis that are used throughout the paper as background tools. These statements are standard and can be found in any textbook on complex analysis; we record them here mainly to keep notation consistent and to make cross-references in the main text precise.

**Theorem A.1** (Identity theorem). *Let  $U \subset \mathbb{C}$  be a domain and let  $f, g : U \rightarrow \mathbb{C}$  be holomorphic. If the set*

$$\{z \in U : f(z) = g(z)\}$$

*has a limit point in  $U$ , then  $f \equiv g$  on  $U$ .*

**Theorem A.2** (Maximum modulus principle). *Let  $U \subset \mathbb{C}$  be a domain and  $f : U \rightarrow \mathbb{C}$  holomorphic. If there exists a point  $a \in U$  such that*

$$|f(a)| \geq |f(z)| \quad \text{for all } z \in U,$$

*then  $f$  is constant.*

**Theorem A.3** (Open mapping theorem). *Let  $U \subset \mathbb{C}$  be a domain and  $f : U \rightarrow \mathbb{C}$  a nonconstant holomorphic function. Then  $f$  is an open map, i.e.  $f(U)$  is open in  $\mathbb{C}$  and for every open  $V \subset U$  the image  $f(V)$  is open.*

**Theorem A.4** (Local form of nonconstant holomorphic maps). *Let  $U \subset \mathbb{C}$  be a domain and  $f : U \rightarrow \mathbb{C}$  a nonconstant holomorphic function. For each  $a \in U$  there exists a neighborhood  $V \subset U$  of  $a$ , a neighborhood  $W \subset \mathbb{C}$  of  $f(a)$ , and a biholomorphism  $\varphi : V \rightarrow W$  such that*

$$f = h \circ \varphi$$

*on  $V$ , where  $h(z) = (z - f(a))^k$  for some integer  $k \geq 1$ . In particular, if  $f'(a) \neq 0$  then  $k = 1$  and  $f$  is a local biholomorphism near  $a$ .*

**Theorem A.5** (Liouville's theorem). *If  $f : \mathbb{C} \rightarrow \mathbb{C}$  is entire and bounded, then  $f$  is constant.*

**Theorem A.6** (Schwarz lemma). *Let  $f : \mathbb{D} \rightarrow \mathbb{D}$  be holomorphic with  $f(0) = 0$ . Then  $|f(z)| \leq |z|$  for all  $z \in \mathbb{D}$ , and  $|f'(0)| \leq 1$ . Moreover, if  $|f(z_0)| = |z_0|$  for some  $z_0 \in \mathbb{D} \setminus \{0\}$ , or if  $|f'(0)| = 1$ , then there exists  $\theta \in \mathbb{R}$  such that  $f(z) = e^{i\theta}z$  for all  $z \in \mathbb{D}$ .*

**Theorem A.7** (Schwarz–Pick lemma). *Let  $f : \mathbb{D} \rightarrow \mathbb{D}$  be holomorphic. Then for all  $z, w \in \mathbb{D}$ ,*

$$\left| \frac{f(z) - f(w)}{1 - \overline{f(w)}f(z)} \right| \leq \left| \frac{z - w}{1 - \overline{w}z} \right|.$$

*Equivalently, for all  $z \in \mathbb{D}$ ,*

$$\frac{|f'(z)|}{1 - |f(z)|^2} \leq \frac{1}{1 - |z|^2}.$$

*If equality holds at one point in either form, then  $f$  is an automorphism of  $\mathbb{D}$ .*

**Corollary A.8** (Disk automorphisms). *A holomorphic map  $f : \mathbb{D} \rightarrow \mathbb{D}$  is a biholomorphism if and only if it has the form*

$$f(z) = e^{i\theta} \frac{z - a}{1 - \overline{a}z}$$

*for some  $a \in \mathbb{D}$  and  $\theta \in \mathbb{R}$ .*

**Theorem A.9** (Laurent expansion on a punctured disk). *Let  $f$  be holomorphic on a punctured disk  $0 < |z - a| < r$ . Then  $f$  admits a Laurent expansion*

$$f(z) = \sum_{n=-\infty}^{\infty} c_n(z - a)^n, \quad 0 < |z - a| < r,$$

*which converges absolutely and uniformly on every annulus  $\rho \leq |z - a| \leq R$  with  $0 < \rho < R < r$ .*

**Theorem A.10** (Classification of isolated singularities). *Let  $f$  be holomorphic on a punctured disk  $0 < |z - a| < r$ . Then exactly one of the following holds:*

- (i) Removable:  $f$  extends holomorphically across  $a$ .
- (ii) Pole:  $|f(z)| \rightarrow \infty$  as  $z \rightarrow a$ .
- (iii) Essential: neither (i) nor (ii) holds.

*Moreover,  $a$  is removable if and only if  $f$  is bounded on some punctured neighborhood of  $a$ .*

**Theorem A.11** (Casorati–Weierstrass). *If  $a$  is an essential singularity of  $f$ , then  $f(0 < |z - a| < \rho)$  is dense in  $\mathbb{C}$  for every  $0 < \rho < r$ .*

*Remark A.12.* In the main text we only use the qualitative consequence that essential singularities exhibit arbitrarily wild behavior near the singular point, and in particular cannot be bounded.

## APPENDIX B. COVERING SPACES

This appendix records the basic covering-space facts used in the main text. We assume familiarity with the fundamental group and homotopies of paths (as in a first course in topology), but we include proofs of the lifting statements for convenience.

**Definition B.1.** A continuous surjection  $p : \tilde{X} \rightarrow X$  is a *covering map* if for every  $x \in X$  there exists an open neighborhood  $U$  of  $x$  such that

$$p^{-1}(U) = \bigsqcup_{\alpha \in A} U_\alpha$$

is a disjoint union of open sets and for each  $\alpha$  the restriction  $p|_{U_\alpha} : U_\alpha \rightarrow U$  is a homeomorphism. Such a neighborhood  $U$  is said to be *evenly covered*.

**Definition B.2.** Let  $p : \tilde{X} \rightarrow X$  be a covering map. A *deck transformation* (or *covering transformation*) is a homeomorphism  $\varphi : \tilde{X} \rightarrow \tilde{X}$  such that  $p \circ \varphi = p$ . The set of deck transformations forms a group  $\text{Deck}(\tilde{X}/X)$  under composition.

**Lemma B.3** (Path lifting). *Let  $p : \tilde{X} \rightarrow X$  be a covering map, let  $\gamma : [0, 1] \rightarrow X$  be a path, and let  $\tilde{x}_0 \in p^{-1}(\gamma(0))$ . Then there exists a unique path  $\tilde{\gamma} : [0, 1] \rightarrow \tilde{X}$  such that*

$$\tilde{\gamma}(0) = \tilde{x}_0, \quad p \circ \tilde{\gamma} = \gamma.$$

*Proof.* For each  $t \in [0, 1]$ , choose an evenly covered neighborhood  $U_t$  of  $\gamma(t)$ . By compactness of  $[0, 1]$ , there exist finitely many times  $0 = t_0 < t_1 < \dots < t_n = 1$  and evenly covered sets  $U_i := U_{t_i}$  such that  $\gamma([t_i, t_{i+1}]) \subset U_i$  for each  $i$ .

Write  $p^{-1}(U_0) = \bigsqcup_{\alpha} U_{0,\alpha}$  with each  $p|_{U_{0,\alpha}}$  a homeomorphism onto  $U_0$ . There is a unique index  $\alpha_0$  with  $\tilde{x}_0 \in U_{0,\alpha_0}$ . Define  $\tilde{\gamma}$  on  $[t_0, t_1]$  by

$$\tilde{\gamma}(s) := (p|_{U_{0,\alpha_0}})^{-1}(\gamma(s)).$$

Then  $\tilde{\gamma}$  is continuous on  $[t_0, t_1]$  and satisfies  $p \circ \tilde{\gamma} = \gamma$  there.

Inductively, suppose  $\tilde{\gamma}$  has been defined on  $[0, t_i]$ . Let  $\tilde{x}_i := \tilde{\gamma}(t_i)$ . Since  $\gamma([t_i, t_{i+1}]) \subset U_i$  and  $p^{-1}(U_i) = \bigsqcup_{\alpha} U_{i,\alpha}$ , there is a unique sheet  $U_{i,\alpha_i}$  containing  $\tilde{x}_i$ . Define  $\tilde{\gamma}$  on  $[t_i, t_{i+1}]$  by the same formula using  $(p|_{U_{i,\alpha_i}})^{-1}$ . The definitions match at  $t_i$  by construction, hence  $\tilde{\gamma}$  is continuous.

Uniqueness follows similarly: if  $\tilde{\gamma}_1, \tilde{\gamma}_2$  are two lifts with the same initial value, then on  $[t_0, t_1]$  both must lie in the unique sheet over  $U_0$  containing  $\tilde{x}_0$ , where  $p$  is injective; hence they agree. Inducting over the subintervals yields  $\tilde{\gamma}_1 = \tilde{\gamma}_2$  on all of  $[0, 1]$ .  $\square$

**Lemma B.4** (Homotopy lifting). *Let  $p : \tilde{X} \rightarrow X$  be a covering map. Let  $H : [0, 1] \times [0, 1] \rightarrow X$  be a continuous homotopy of paths, and suppose a lift  $\tilde{H}(\cdot, 0)$  of  $H(\cdot, 0)$  is fixed. Then there exists a unique continuous lift  $\tilde{H} : [0, 1] \times [0, 1] \rightarrow \tilde{X}$  such that  $p \circ \tilde{H} = H$  and  $\tilde{H}(\cdot, 0)$  is the prescribed lift.*

*Proof.* Cover the compact set  $H([0, 1] \times [0, 1]) \subset X$  by finitely many evenly covered open sets and subdivide the square into finitely many small rectangles so that the image of each rectangle lies in a single evenly covered set. On each rectangle, one

lifts uniquely once the lift is specified along one edge, because in a fixed sheet the covering map is a homeomorphism. The uniqueness on overlaps ensures that these local lifts patch to a global continuous lift on the square.  $\square$

**Proposition B.5** (Unique lifting for maps). *Let  $p : \tilde{X} \rightarrow X$  be a covering map and let  $Y$  be connected. Suppose  $f : Y \rightarrow X$  is continuous and  $y_0 \in Y$  with a chosen  $\tilde{x}_0 \in p^{-1}(f(y_0))$ . If a lift  $\tilde{f} : Y \rightarrow \tilde{X}$  satisfying  $p \circ \tilde{f} = f$  and  $\tilde{f}(y_0) = \tilde{x}_0$  exists, then it is unique.*

*Proof.* Let  $\tilde{f}_1, \tilde{f}_2$  be two such lifts. Consider the set

$$A := \{y \in Y : \tilde{f}_1(y) = \tilde{f}_2(y)\}.$$

Since  $\tilde{f}_1(y_0) = \tilde{f}_2(y_0)$ , the set  $A$  is nonempty. It is closed by continuity. To see  $A$  is open, fix  $y \in A$ . Choose an evenly covered neighborhood  $U$  of  $f(y)$ . For  $y'$  in a small neighborhood  $V$  of  $y$  with  $f(V) \subset U$ , both  $\tilde{f}_1(V)$  and  $\tilde{f}_2(V)$  must lie in the same sheet over  $U$  containing  $\tilde{f}_1(y) = \tilde{f}_2(y)$ , and on that sheet  $p$  is injective, hence  $\tilde{f}_1 = \tilde{f}_2$  on  $V$ . Thus  $A$  is open. As  $Y$  is connected,  $A = Y$ .  $\square$

**Theorem B.6** (Existence of universal covers). *If  $X$  is connected, locally path-connected, and semilocally simply connected, then there exists a universal covering space  $p : \tilde{X} \rightarrow X$ , unique up to isomorphism of coverings.*

*Remark B.7.* A proof can be found in standard topology references. Every (connected) topological surface is locally path-connected and semilocally simply connected, hence admits a universal cover.

**Proposition B.8.** *Let  $p : \tilde{X} \rightarrow X$  be a covering map with  $\tilde{X}$  connected. If  $X$  is simply connected, then  $p$  is a homeomorphism.*

*Proof.* Fix  $x_0 \in X$  and  $\tilde{x}_0 \in p^{-1}(x_0)$ . For any  $x \in X$ , choose a path  $\gamma$  from  $x_0$  to  $x$ . By Lemma B.3 there is a unique lift  $\tilde{\gamma}$  with  $\tilde{\gamma}(0) = \tilde{x}_0$ . Define  $s(x) := \tilde{\gamma}(1)$ . If  $\gamma_1, \gamma_2$  are two paths from  $x_0$  to  $x$ , then  $\gamma_1 \cdot \bar{\gamma}_2$  is a loop, hence null-homotopic since  $X$  is simply connected; by Lemma B.4 the corresponding lifts have the same endpoint, so  $s(x)$  is well-defined.

Thus  $s : X \rightarrow \tilde{X}$  is a continuous section with  $p \circ s = \text{id}_X$ . In particular,  $p$  is surjective. Moreover,  $p$  is injective: if  $p(\tilde{x}) = p(\tilde{y}) = x$ , let  $\gamma$  be any path from  $x_0$  to  $x$ . Lifting  $\gamma$  starting at  $\tilde{x}_0$  produces  $\tilde{\gamma}(1) = s(x)$ . But any lift ending at  $\tilde{x}$  (resp.  $\tilde{y}$ ) must have the same endpoint by uniqueness, so  $\tilde{x} = \tilde{y} = s(x)$ . Hence  $p$  is a continuous bijection. Since  $p$  is a local homeomorphism, it is a homeomorphism.  $\square$

**Proposition B.9** (Quotient by the deck group). *Let  $p : \tilde{X} \rightarrow X$  be a universal covering map with  $\tilde{X}$  connected. Then  $\text{Deck}(\tilde{X}/X)$  acts freely and properly discontinuously on  $\tilde{X}$ , and the quotient  $\tilde{X}/\text{Deck}(\tilde{X}/X)$  is naturally homeomorphic to  $X$ .*

*Proof.* Freeness is immediate: if  $\varphi(\tilde{x}) = \tilde{x}$  and  $p \circ \varphi = p$ , then by the unique lifting property (Proposition B.5) applied on a small evenly covered neighborhood of  $p(\tilde{x})$ , the map  $\varphi$  agrees with the identity on an open set, hence on all of  $\tilde{X}$  by connectedness. Proper discontinuity follows from the existence of evenly covered neighborhoods: over an evenly covered  $U \subset X$ , distinct deck transformations permute the sheets above  $U$ , hence move each sheet off itself unless they are the identity.

Define  $\Phi : \tilde{X}/\text{Deck}(\tilde{X}/X) \rightarrow X$  by  $\Phi([\tilde{x}]) := p(\tilde{x})$ . This is well-defined because  $p \circ \varphi = p$ . It is continuous and bijective; local homeomorphism follows from evenly covered neighborhoods. Hence  $\Phi$  is a homeomorphism.  $\square$

#### APPENDIX C. A TOPOLOGICAL INPUT: COMPACT SIMPLY CONNECTED SURFACES

In the elliptic case we use the following standard topological fact: a compact, connected, simply connected surface is topologically a 2-sphere. We record a short derivation from the classification of compact surfaces.

**Theorem C.1** (Classification of closed surfaces). *Every compact, connected 2-manifold without boundary is homeomorphic to exactly one of the following:*

- (i) the 2-sphere  $S^2$ ;
- (ii) the connected sum of  $g \geq 1$  tori, denoted  $\Sigma_g$  (an orientable surface of genus  $g$ );
- (iii) the connected sum of  $k \geq 1$  real projective planes, denoted  $N_k$  (a nonorientable surface of genus  $k$ ).

*Remark C.2.* We do not reprove Theorem C.1 here; proofs may be found in standard topology texts on surfaces (e.g. in treatments of the classification theorem for compact surfaces). We only use the theorem to identify which closed surface can have trivial fundamental group.

**Proposition C.3.** *For  $g \geq 1$ , the orientable surface  $\Sigma_g$  has nontrivial fundamental group. More precisely,  $\pi_1(\Sigma_g)$  admits the presentation*

$$\pi_1(\Sigma_g) \cong \left\langle a_1, b_1, \dots, a_g, b_g \mid \prod_{j=1}^g [a_j, b_j] = 1 \right\rangle,$$

*and in particular  $\pi_1(\Sigma_g) \neq 1$ .*

*Proof.* It is classical that  $\Sigma_g$  can be obtained by gluing opposite sides of a  $4g$ -gon with edge word  $a_1 b_1 a_1^{-1} b_1^{-1} \cdots a_g b_g a_g^{-1} b_g^{-1}$ . Applying van Kampen's theorem to this CW model yields the stated presentation. Since the generators  $a_1, b_1$  map to nontrivial elements in the abelianization, the group cannot be trivial; hence  $\pi_1(\Sigma_g) \neq 1$ .  $\square$

**Proposition C.4.** *For  $k \geq 1$ , the nonorientable surface  $N_k$  has nontrivial fundamental group. In particular,  $\pi_1(N_k) \neq 1$ .*

*Proof.* The case  $k = 1$  is  $N_1 = \mathbb{RP}^2$ , whose fundamental group is  $\mathbb{Z}/2\mathbb{Z}$ . For  $k \geq 2$ ,  $N_k$  contains an embedded copy of  $\mathbb{RP}^2$  as a connected summand, and the inclusion induces a nontrivial map on  $\pi_1$  (for instance, by a van Kampen computation on the connected sum decomposition). Hence  $\pi_1(N_k)$  is nontrivial.  $\square$

**Theorem C.5.** *Let  $M$  be a compact, connected, simply connected 2-manifold without boundary. Then  $M$  is homeomorphic to  $S^2$ .*

*Proof.* By Theorem C.1,  $M$  is homeomorphic to exactly one of  $S^2$ ,  $\Sigma_g$  for some  $g \geq 1$ , or  $N_k$  for some  $k \geq 1$ . If  $M \cong \Sigma_g$  with  $g \geq 1$ , then  $\pi_1(M) \cong \pi_1(\Sigma_g)$  is nontrivial by Proposition C.3, contradicting simple connectedness. If  $M \cong N_k$  with  $k \geq 1$ , then  $\pi_1(M) \cong \pi_1(N_k)$  is nontrivial by Proposition C.4, again a contradiction. Therefore the only remaining possibility is  $M \cong S^2$ .  $\square$