UMEÅ UNIVERSITY

The Great Picard Theorem

Dennis Wahlström

Bachelor's Thesis, 15 Credits
Bachelor of Mathematics, 180 Credits
Spring 2018
Department of Mathematics and Mathematical Statistics

Abstract

In this essay, we present a proof of the great Picard theorem by showing that a holomorphic function with an essential singularity attains infinitely many complex values in the vicinity of the singularity.

Sammanfattning

Vi kommer att presentera ett bevis på Picards stora sats genom att visa att en holomorf funktion med en väsentlig singularitet antar oändligt många komplexa värden i ett område av singulariteten.

Contents

Acknowledgements

Firstly, I would like to thank my supervisor, Per Åhag with my deepest gratitude for all his supports and advises during my difficult time, that was one the most challenging period of my life. Without Per, it would have been more difficult to be able to spell the word *success*. Secondly, many thanks to Aron Persson, Jesper Karlsson and Robin Törnkvist for took times to read and comment this essay. Lastly, I would like to thank my examiner, Olow Sande for his advises on this essay.

1. Introduction

Complex analysis is one of the mathematical branches, and without any doubt it is an important mathematical field which has contributed to many study applications related to our daily life experience, such as in physics and engineering.

The complex analysis is the analysis of functions of so called complex numbers, which include the square root of negative numbers. These numbers were first encountered in the first century of Anno Domini by Hero of Alexandria, at that time they were considered nonsensical due to the lack of awareness [7]. Even so, it took the works of many mathematicians before the theory of complex variables was commenced in the eighteenth century by Leonhard Euler [17]. The theory of complex variables was heired by the three pioneers Augustin-Louis Cauchy, Bernhard Riemann and Karl Weierstrass with distinct thought on the subject [17]. Cauchy's view on the theory of complex variables was circular functions and the exponential functions, meanwhile Riemann focused on the geometry and Weierstrass had presented a vision on the convergent power series [17]. These visions have their own significative meaning in the today's mathematics.

We will focus ourself on one of the succeeders in the theory of complex variables, Charles Émile Picard. In the nineteenth century, Picard presented his thesis in an application of linear complexes where he mainly dealt with geometrical aspect and as the result, it became an important property to the Riccati's equation [5]. Picard fully became a professor at the age of thirty and during his time as professor, he taught countless students e.g. in rational mechanics, differential and integral calculus, higher algebra and analysis [5]. Moreover, his personalities and teaching skills were acknowledged by many including by one of his PhD student Jacques Hadamard, who had chosen to described his advisor as: "A notable feature about Picard's scientific personality was the perfection of his teaching, one of the most, if not the most, marvelous I have experienced. One could say of it, thinking of what Mozart is reported to have said of one of his own works, that 'there was not one word too much, nor one word lacking'." [5].

A few years after Picard obtained his professorship, he realised that the Casorati-Weierstrass theorem for a holomophic function with essential singularities could be strengthened, which today is known as the great Picard theorem [5, 17].

Theorem 6.5 (The Great Picard Theorem). *If f has an essential singularity at z_0, then with at most one exception, f attains every complex value infinitely many times in every neighborhood of z_0.*

The original proof of the great Picard theorem can be found in [15, 16]. It has many generalisations in the area of analytic mappings to an extended plane for essential functions as well as for a punctured disk [11, 14, 18]. Besides these, it also has generalitions to e.g. Montel's theorem, Schottky's theorem and Bloch's theorem [3, 19].

The prerequisites of this essay are a first course in complex analysis, a common property in calculus e.g. the real-differentiability and a course in linear algebra. Those who are familiar with algebraic topology and algebraic structure or advanced complex analysis will have a greater advantage in pursuing the great Picard theorem in Section 6, but we encourage to all the readers to explore the other sections filled with interesting proofs.

The layout of this essay is as follows: first a review of complex analysis, and in Section 3 we shall define the automorphism group and construct the extended complex plane. Section 4 contains a construction of a modular function, and in Section 5 we will tie most of our results together preparing for Section 6, which is the great Picard theorem. The main theorems in this thesis are presented in Figure 1.1.

This thesis is based on [20] with the focus on simplifying and clarifying for the readers who has not studied the advanced calculus course as it was one of prerequisites for the book. Moreover, a few definitions were alternatively taken from [9].

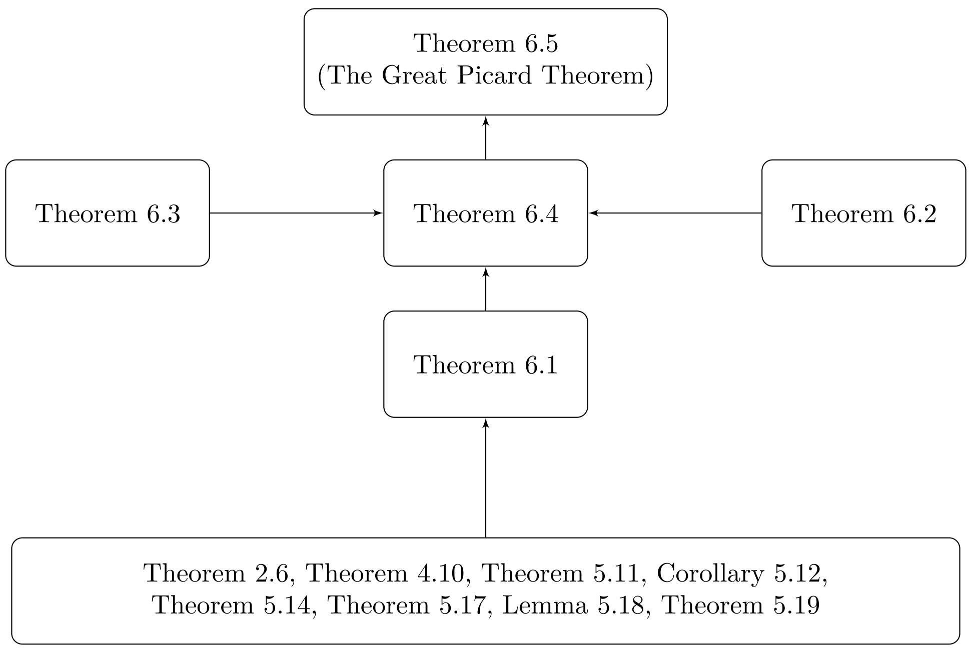

Figure 1.1 – Architecture of the essay presented in a flowchart.

2. Complex Analysis

The goal in this section is to reconstruct the prerequisites from complex analysis in the usage to construct conformal maps and modular functions. We will start off by introducing the concept of *complex differentiable*.

Proposition 2.1. *Suppose that V is an open subset of $\mathbb{C}$, that $f : V \to \mathbb{C}$, and $z \in V$. Let u and v be the real and imaginary of f. Then $f = u + iv$ is complex-differentiable at z, if, and only if, it is real-differentiable at z and the real and imaginary parts satisfy the Cauchy-Riemann equations*

$$u_x(z) = v_y(z), \qquad u_y(z) = -v_x(z)$$

where x and y are the real part in $\mathbb{R}$, and u_x, u_y v_x and v_y are the partial derivative.

Throughout the essay we will denote V as an open set of the set of all complex numbers $\mathbb{C}$. The Proposition 2.1 gives us authority to define the concept of holomorphic function, which is a complex function that is differentiable everywhere in the domain.

Definition 2.2. Suppose that $V \subset \mathbb{C}$ is open and $f : V \to \mathbb{C}$. We say that f is *holomorphic* in V if f is complex differentiable at every point of V. The class of all functions holomorphic in V is denoted $H(V)$.

There are various type of singularities, *isolated singularity*, *removable singularity* and *essential singularity*. By any means, the singularity is a point in which the mathematical object is not defined, e.g. differentiability fails. These singularities have their definition as following.

Definition 2.3. Let f be a function in $\mathbb{C}$ and $z \in \mathbb{C}$, suppose that f has an isolated singularity at z. The singularity is *removable* if f can be defined at z such that f becomes holomorphic in a neighborhood of z. The singularity is a *pole* if $\lim_{w \to z} f(w) = \infty$. Finally, an isolated singularity which is neither a pole nor a removable singularity is an *essential singularity*.

Let's introduce an open and a closed disk with z as its center where $z \in \mathbb{C}$ and a finite radius $r > 0$:

$$D(z,r) = \{w \in \mathbb{C} : |z - w| < r\} \text{ and } \overline{D}(z,r) = \{w \in \mathbb{C} : |z - w| \leq r\}. \tag{2.1}$$

Further through the essay, we will encounter a bounded holomorphic function on the boundary of its domain which can be divided into two different cases, a constant and a nonconstant holomorphic function. However, we will state these in form of two famous theorems that goes by the name of *open mapping theorem with bounds* and *Maximum modulus theorem*.

Lets start with the open mapping theorem with bounds.

Theorem 2.4. *Suppose that $f \in H(V)$ and its second-order complex derivative $|f''(z)| \leq B$ for all $z \in V$, where B is a bound of V. Suppose that $z_0 \in V$ and first-order complex derivative $f'(z_0) \neq 0$. If $r > 0$ is small enough that $\overline{D}(z_0, r) \subset V$ and also $r < \frac{2|f'(z_0)|}{B}$ then*

$$D(f(z_0), \delta) \subset f(D(z_0, r))$$

for

$$\delta = r|f'(z_0)| - \frac{Br^2}{2}.$$

Proof. See e.g. in ([17], p. 256-257) or Theorem 5.7 in [20]. □

Before we proceed with the proof of the Maximum modulus theorem let's define power series. Furthermore, we shall observe that the power series contains an important property to the analytical functions as well to the Maximum modulus theorem. Hence, the Maximum modulus theorem will be used frequently through essay, where it shows a great deal to handle the boundness of the holomorphic functions, which is what we needed in our proof of the great Picard theorem.

Definition 2.5. A series on the form $\sum_{j=0}^{\infty} a_j(z - z_0)^j$ is called a *power series.* The constants a_j are the coefficients of the power series.

Theorem 2.6. *Suppose that $f \in H(V)$ and V is connected. Then $|f|$ cannot achieve a (local) maximum in V unless f is constant: If f is nonconstant, then for every $a \in V$ and $\delta > 0$ there exists $z \in V$ with $|f(z)| > |f(a)|$ and $|z - a| < \delta$.*

Proof. Let $f \in H(V)$ and f has a zero order of N at $z \in V$, then there is a function $h \in H(V)$ with $h(z) \neq 0$ which can be presented together with the power series of f as:

$$f(z) = f(a) + c(z - a)^N + (z - a)^{N+1}h(z).$$

From the previous step and select $z - a$ appropriately we are able to get:

$$|f(a) + c(z - a)^N| = |f(a)| + |c(z - a)^N| > |f(a)|.$$

This shows that, if the modulus $(z - a)$ is small enough, then the growth $(z - a)^N$ is negligible.

In particular, let $\alpha, \beta, \theta \in \mathbb{R}$ be selected in such way that $f(a) = |f(a)|e^{i\alpha}$, $c = |c|e^{i\beta}$ and $\beta + N\theta = \alpha$ for a $N \neq 0$. If $z = a + re^{i\theta}$ with $r > 0$, then

$$f(a) + c(z - a)^N = \big(|f(a)| + |c|r^N\big)e^{i(\beta+N\theta)}$$

and

$$|f(a) + c(z - a)^N| = |f(a)| + |c|r^N.$$

We know that h is continuous in V and bounded near a. Thus, there is a bound M such that $|h(z)| \leq M$ for $|z - a| \leq \rho$, where $\rho > 0$ and $M < \infty$. The triangle inequality shows that, if $0 < r < \rho$, then

$$\begin{aligned} |f(z)| &\geq |f(a) + c(z - a)^N| - |(z - a)^{N+1}h(z)| \\ &\geq |f(a)| + |c|r^N - Mr^{N+1} = |f(a)| + (|c| - Mr)r^N. \end{aligned}$$

Herein, we are free to select $\delta \in (0, \rho)$ so that $M\rho < |c|/2$, hence $0 < r < \rho$,

$$|f(z)| \geq |f(a)| + \frac{|c|r^N}{2} > |f(a)|.$$

□

Besides the open mapping theorem and the maximum modulus theorem, we will also be needing the Schwarz lemma, but first lets define the unit disk $\mathbb{D}$ as $\mathbb{D} = D(0, 1)$, where $D(z, r)$ is defined as (2.1). The following theorem is the Schwarz lemma.

Theorem 2.7. *Suppose that $f : \mathbb{D} \to \mathbb{D}$ is holomorphic and $f(0) = 0$. Then $|f(z)| \leq |z|$ for all $z \in \mathbb{D}$ and $|f'(0)| \leq 1$. Furthermore, if $|f'(0)| = 1$ or $|f(z)| = |z|$ for some nonzero $z \in \mathbb{D}$, then f is a rotation: $f(z) = \beta z$ for some constant β with $|\beta| = 1$.*

Proof. Since $f(0) = 0$, we will define a function $g \in H(\mathbb{D})$ in such manner:

$$g(z) = \begin{cases} \frac{f(z)}{z}, & \text{if } z \neq 0, \\ f'(0), & \text{if } z = 0. \end{cases}$$

To begin with, we will recall the result from Theorem 2.6 which states that if D is a bounded open set in the complex plane with $f \in H(D)$, then there is a bound $M \in D$ preventing f to achieve a local maximum in D unless f is constant. It is also indicating that $|g| \leq 1$ in $\mathbb{D}$. However, we want to show that $|g| \leq 1$ on the boundary of the unit disk, $\partial\mathbb{D}$, problematically, g is undefined on $\partial\mathbb{D}$.

To handle the issue above: suppose that $r \in (0, 1)$, the function $|g| \leq 1/r$ is restricted on $\partial D(0, r)$ and Theorem 2.6 shows that $|g| \leq 1$ in $\mathbb{D}$. If $|f'(0)| = 1$ or $|f(z)| = |z|$ for a nonzero $z \in \mathbb{D}$, then $|g(z)| = 1$ for some $z \in \mathbb{D}$ since g is constant. Thus, the proof is then concluded. □

Historical note: Carathéodory recognised the strength of the Maximum modulus theorem, where he was first to apply it to prove the Schwarz lemma in such manner [1].

We will continue with the definition of compact sets.

Definition 2.8. A *compact set* in $\mathbb{C}$ is a set that is both closed and bounded.

Now with the definition of a compact set we can state the Montel's theorem for a uniformly bounded family. A family of functions is said to be *uniformly bounded* if there is a constant in the domain restricting all the functions. The Montel's theorem is stated as the following.

Theorem 2.9. *If $\mathcal{F} \subset H(D)$ is a uniformly bounded family and (f_n) is a sequence in $\mathcal{F}$, then there exists a subsequence f_{n_j} such that $f_{n_j} \to f \in H(D)$ uniformly on compact subsets of D.*

Proof. See e.g. in ([2], p. 182-183). □

One of the results from a first course in complex analysis is the *Riemann mapping theorem*, and it will be strengthened in the connection with the extended complex plane.

Theorem 2.10. *Suppose that $D \subset \mathbb{C}$ is open, connected, nonempty, and has the property that any nonvanishing function holomorphic in D has a holomorphic square root. If $D \neq \mathbb{C}$, then D is conformally equivalent to $\mathbb{D}$.*

Proof. See e.g. Theorem 6.4.2 in [4]. □

Through this essay, we shall encounter functions with poles and singularities, and realise that such particular functions are suitable to be described as a harmonic function. Accordingly, the property of the harmonic function is essential to the proof of the great Picard theorem. The definition of harmonic function is stated as following.

Definition 2.11. A real-valued function $\phi(x, y)$ is said to be *harmonic* in a domain D if all its second-order partial derivatives are continuous in D and if, at each point of D, ϕ satisfies Laplace's equation.

Hence, the domain $D \in \mathbb{R}^2$ with a fact that $\mathbb{R}^2 \in \mathbb{C}$, this contributes a connection between holomorphic function and harmonic function stated as in Lemma 2.12. Explicitly, the harmonic function satisfies the Cauchy-Riemann equations, and hence it is considered as a holomorphic function.

Lemma 2.12. *Suppose that $D \subset \mathbb{C}$ is open and $u : D \to \mathbb{C}$ is twice continuously differentiable, then these hold:*

(i) *A holomorphic function is harmonic;*

(ii) *If u is harmonic, then the function $f = \delta u/\delta z$ is holomorphic;*

(iii) *If D is simply connected and u is a real-valued harmonic function in D, then there exists $f \in H(D)$ with $u = \mathrm{Re}(f)$.*

Proof. See e.g. Lemma 10.0.1 in [20]. □

3. Conformal Mappings and Automorphism Groups

The goal with this section is to introduce *conformal mappings*, *automorphism groups* and the *extended complex plane*. We will start with the conformal map.

A function $f \in H(V)$ is said to be a *conformal map* when its derivative f' has no zero in V. In other words, the function f preserves angles locally as presented in the following proposition:

Proposition 3.1. *Suppose that $f \in H(V), z_0 \in V$, and $f'(z_0) \neq 0$. Also, suppose that $\gamma_0, \gamma_1 : [0,1] \to V$ are curves with nonvanishing (right) derivatives at* 0*, with $\gamma_j(0) = z_0$. Let $\tilde{\gamma}_j(t) = f(\gamma_j(t))$. Then*

$$\arg\left(\frac{\tilde{\gamma}_1'(0)}{\tilde{\gamma}_0'(0)}\right) = \arg\left(\frac{\gamma_1'(0)}{\gamma_0'(0)}\right).$$

Proof. Let

$$\gamma_j^{'}(0) = r_j e^{i\theta_j}, \quad \text{for } j = 0, 1$$

with $r_j > 0$ and $\theta_j \in \mathbb{R}$ and let

$$f^{'}(z_0) = re^{i\theta},$$

then

$$\tilde{\gamma}_j^{'}(0) = f^{'}(z_0)\gamma_j^{'}(0) = rr_j e^{i(\theta_j + \theta)},$$

gives

$$\arg\left(\frac{\tilde{\gamma}_1^{'}(0)}{\tilde{\gamma}_0^{'}(0)}\right) = (\theta_1 + \theta) - (\theta_0 + \theta) = \theta_1 - \theta_0 = \arg\left(\frac{\gamma_1^{'}(0)}{\gamma_0^{'}(0)}\right).$$

The angles are preserved. □

Remark 3.2. If $f \in H(V)$ and it is a one-to-one function in V with $W = f(V)$, then the function f is said to be a *conformal equivalence* and W is *conformally equivalent* to V.

The conformal map is a type of transformation which is important in complex analysis, and in our concern it is one of the main pieces of introduction to an other type of transformation that allows us to extend the complex plane, that is a *linear-fractional transformation*. Before then, we need a few more components.

Definition 3.3. A function f is meromorphic in V, if there exists a relatively closed set $S \subset V$ such that every point of S is isolated with $f \in H(V \setminus S)$, and f has a pole or a removable singularity at every point of S.

If a function is meromorphic in a domain V that is connected, then the function is a holomorphic mapping of V into the *extended plane*, $\mathbb{C}_\infty$ (also knows as the *Riemann sphere*):

$$\mathbb{C}_\infty = \mathbb{C} \cup \{\infty\}.$$

To illustrate this we define:

$$S_1 = \{x \in \mathbb{R}^3 : x_1^2 + x_2^2 + x_3^2 = 1\}.$$

This is the unit sphere in $\mathbb{R}^3$.

The point here is to construct the holomorphic mapping by conformally projecting the unit sphere onto a plane with the *stereographic projection*. We define $P : \mathbb{C}_\infty \to S_1$

and $P(z)$ for $z \in \mathbb{C}$ by letting $P(\infty) = (0,0,1)$ with a straight line intersecting $(0,0,1)$ and $(z,0) \in \mathbb{R}^3$. The illustration of this can be observed in Figure 3.1. The stereographic projection is:

$$P(z) = \left(\frac{2x}{|z|^2+1}, \frac{2y}{|z|^2+1}, \frac{|z|^2-1}{|z|^2+1} \right), \quad \text{for } z = x+iy \in \mathbb{C}.$$

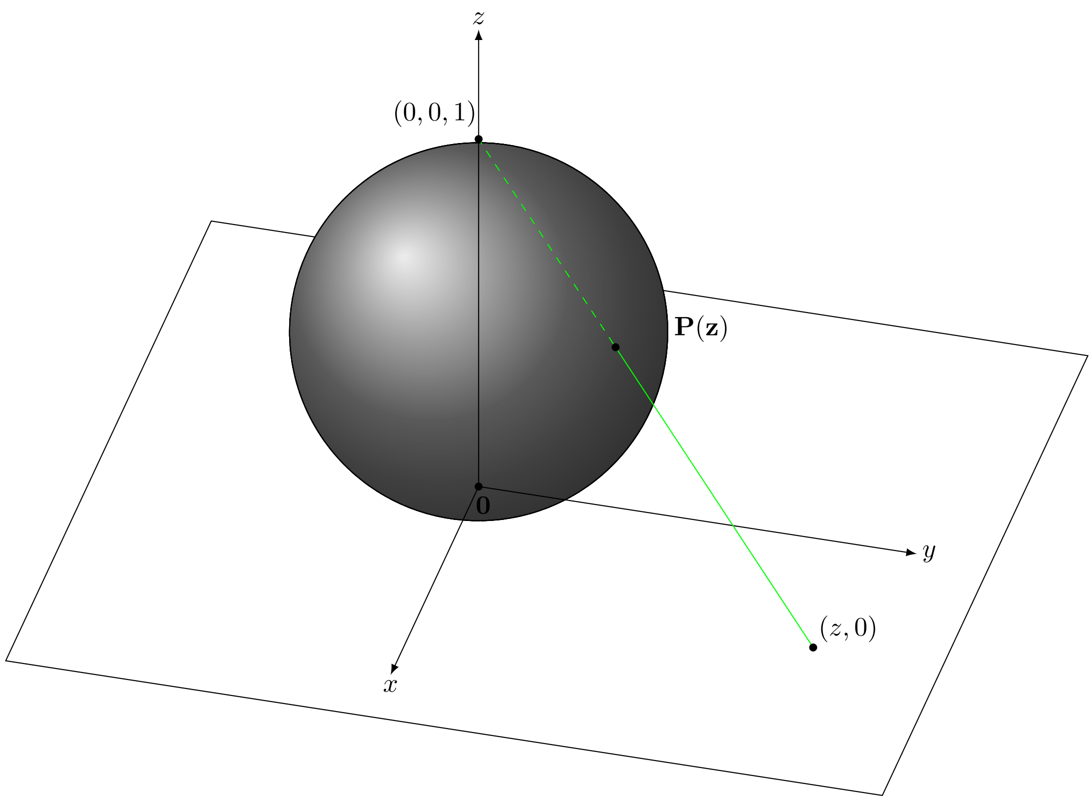

Figure 3.1 – A sphere on a plane.

The stereographic projection can be used to define a metric on $\mathbb{C}_\infty$:

$$d_\infty(z,w) = \|P(z) - P(w)\|, \quad \text{for } z, w \in \mathbb{C}_\infty. \tag{3.1}$$

where d_∞ is the default metric on $\mathbb{C}_\infty$ and $\|\cdot\|$ is the normed vector space. Analogously, $(\mathbb{C}_\infty, d_\infty)$ is a compact metric space due to the fact that S_1 is the unit sphere, a compact subset of $\mathbb{R}^3$.

Herein, $\mathbb{C}_\infty$ is a *complex manifold* and we shall not break it down in details, except our interest is to show that functions from the subsets of $\mathbb{C}_\infty$ to $\mathbb{C}_\infty$ are holomorphic (if one is interested in details in the complex manifold one can turn e.g. to [12]). This suggests that there is a bijection of $\mathbb{C}_\infty$ to itself. Particularly, if we define $1/0 = \infty$ and

$1/\infty = 0$, then this shows that there is a function $z \to 1/z$ such that $\mathbb{C}_\infty$ is a bijection onto itself. Specifically, suppose that V is an open subset of $\mathbb{C}_\infty$ and $f : V \to \mathbb{C}_\infty$ is a function. The function f is holomorphic, if the following four functions are continuous $f(z), f(1/z), 1/f(z)$ and $1/f(1/z)$. If this is true, then the function is holomorphic in the open subset and finite on $\mathbb{C}$. We will call f a *holomorphic mapping*.

Theorem 3.4. *If $f : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is holomorphic, then there exist polynomials P and Q such that*

$$f(z) = \frac{P(z)}{Q(z)}, \tag{3.2}$$

for all $z \in \mathbb{C}_\infty$.

Proof. If $f(z) = \infty$ for all $z \in \mathbb{C}_\infty$, then (3.2) is true with $P = 1$ and $Q = 0$. We know that $\mathbb{C}_\infty$ is a complex manifold and f is meromorphic, by assuming that f has a pole or removable singularity at ∞, we need to show that f is a rational function.

If f has an isolated singularity at ∞, then there exists $R > 0$ such that f is holomorphic in $\{z \in \mathbb{C} : |z| > R\}$. This implies that f has finitely number of poles in $\mathbb{C}$, since zeroes of f are isolated at ∞, then f has also finitely many zeroes in $\mathbb{C}$. Suppose that the zeroes of f in $\mathbb{C}$ are $\{z_1, ..., z_n\}$ and the poles are $\{p_1, ..., p_m\}$. Let

$$g(z) = f(z)\frac{\prod_{j=1}^{m}(z - p_j)}{\prod_{k=1}^{n}(z - z_j)},$$

this shows that g is an entire function with a limit at ∞ and has no zeroes in $\mathbb{C}$. Additionally,

$$\lim_{z\to\infty} \frac{f(z)}{cz^N} = 1,$$

for some $N \in \mathbb{Z}$ and nonzero $c \in \mathbb{C}$. Thus,

$$\lim_{z\to\infty} \frac{g(z)}{cz^{N+m-n}} = 1,$$

which requires g to has a pole or removable singularity at ∞, as well as for the function f. If g has a pole at ∞, then it would imply that $1/g$ is an entire function that approaches to 0 at ∞, and can also be interpreted as $1/g$ is identically 0. Therefore, g must has a finite limit at ∞ with a condition of $N + m - n \leq 0$, which implies that g is a bounded entire function and constant. □

3.1. **Automorphism Groups.** We let $\text{Aut}(V)$ denote a group of invertible holomorphic mappings from V to itself (V is an open subset of the complex plane or $V \subseteq \mathbb{C}_\infty$).

Definition 3.5. A binary operation $*$ on a set S is a function mapping $S \times S$ into S. For each $(a, b) \in S \times S$, we will denote the element $*(a, b)$ of S by $(a * b)$.

In group theory, a *group* is a set with elements and an operator that satisfies the group axioms. Let X be the set with a binary operator $* : X \times X \to X$. The group axioms are:

$$G_0 \text{ :(Closure), } a * b \in X \quad \text{for all } a, b \in X;$$
$$G_1 \text{ :(Associative), } (a * b) * c = a * (b * c);$$
$$G_2 \text{ :(Identity), } \quad e \in X : a * e = e * a = a \quad \text{for all } a \in X;$$
$$G_3 \text{ :(Inverse), } a^{-1} \in X : a * a^{-1} = a^{-1} * a = e \text{ for all } a \in X.$$

We are going to define several definitions related to the group theory and not less to the automorphism group for future conventions.

Definition 3.6. Given two sets X and Y. A set isomorphism between X and Y is a bijection $f : X \to Y$, for some function f.

Definition 3.7. If a subset H of a group G is closed under the binary operation of G and if H with the induced operation from G is itself a group, then H is a *subgroup* of G.

Definition 3.8. In group theory, an order of a group is its cardinality. The order of an element a is the smallest positive integer m such that $a^m = e$, then a is said to have a *finite order*. If no such m exists, then a is said to have *infinite order*.

As a result from defining the conformal map and introducing the stereographic projection, we are now able to obtain other type of transformation.

Definition 3.9. A *linear-fractional transformation* (or *Möbius transformation*) is a mapping in the form of:

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

where $a, b, c, d \in \mathbb{C}$ satisfy

$$ad - bc \neq 0.$$

The condition $ad - bc \neq 0$ implies that c and d cannot both vanish, otherwise, the function becomes rational and reduces to a meromorphic function (a function that is holomorphic on a plane except on a set of isolated points) in the complex plane or a holomorphic mapping from $\mathbb{C}_\infty$ to itself.

Theorem 3.10. $\text{Aut}(\mathbb{C}_\infty)$ *is the set of all linear-fractional transformations.*

Proof. Let $\phi \in \text{Aut}(\mathbb{C}_\infty)$ and let ψ be the inverse of ϕ, then

$$\psi(z) = \frac{dz - b}{-cz + a}$$

then,

$$\phi(\psi(z)) = \psi(\phi(z)) = z$$

for all $z \in \mathbb{C}_\infty$ and $\phi \in \mathrm{Aut}(\mathbb{C}_\infty)$. Since $\phi : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is holomorphic, thus, Theorem 3.4 implies that ϕ is rational,

$$\phi(z) = c\frac{\prod_{k=1}^{n}(z - z_k)}{\prod_{j=1}^{m}(z - p_j)}$$

where $z_1, ..., z_n$ are the zeroes of $\phi \in \mathbb{C}$ and $p_1, ..., p_m$ are the poles in $\mathbb{C}$.

The fact that $\phi : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is one-to-one shows that ϕ has at most one zero in $\mathbb{C}$ and similarly to ψ, so that $n \leq 1$ and $m \leq 1$. If $n = m = 0$, then the function $\phi(z) = c$ is constant, it would also imply that $\phi \notin \mathrm{Aut}(\mathbb{C}_\infty)$. Hence, it leaves us with three possibilities,

$$\phi(z) = c\frac{z - z_1}{z - p_1}, \quad \text{for } (z_1 \neq p_1),$$

$$\phi(z) = c\frac{1}{z - p_1}$$

and

$$\phi(z) = c(z - z_1).$$

Since all the cases are linear-fractional transformation in $\mathbb{C}_\infty$. Hence, $\mathrm{Aut}(\mathbb{C}_\infty)$ is the set of all linear-fractional transformations. □

The Schwarz lemma can be applied to the holomorphic mapping which shows that there is a rotation in $\mathrm{Aut}(\mathbb{D})$.

Theorem 3.11. *Suppose that* $\phi \in \mathrm{Aut}(\mathbb{D})$ *and* $\phi(0) = 0$. *Then* ϕ *is a rotation:*

$$\phi(z) = \beta z$$

for some $\beta \in \mathbb{C}$ *with* $|\beta| = 1$.

Proof. Theorem 2.7 shows that $|\phi'(0)| \leq 1$, and if we let ψ be the inverse of ϕ, then ψ would also satisfy $|\psi'(0)| \leq 1$. Furthermore, if we apply a chain rule, then it would give $\psi'(0) = 1/\phi'(0)$ with $|\phi'(0)| = 1$. Hence, this is a rotation according to Theorem 2.7. □

Our next subject is the *Cayley transform* and will be using for construction of a map and extend a plane. First, we need to define a few topological properties.

Definition 3.12. A collection $\mathcal{T}$ of subsets of X is a topology on X if:

(*i*) $\varnothing, X \in \mathcal{T}$;

(*ii*) if $G_\alpha \in \mathcal{T}$ for $\alpha \in \mathcal{A}$, then $\bigcup_{\alpha \in \mathcal{A}} G_\alpha \in \mathcal{T}$;

(*iii*) if $G_i \in \mathcal{T}$ for $i = 1, 2, ..., n$, then $\bigcap_{i=1}^{n} G_i \in \mathcal{T}$.

The pair $(X, \mathcal{T})$ is called *a topological space* or if $\mathcal{T}$ is not in the context, then X is the topological space.

One of the relations between topological spaces is the homeomorphic function.

Definition 3.13. A function $f : X \to Y$ is said to be *homeomorphic* between the two topological spaces (X, τ_x) and (Y, τ_y), if they follow these properties:

(i) f is a bijection;

(ii) f is continuous;

(iii) the inverse function f^{-1} is continuous.

Now we are ready to introduce the Cayley transform. The Cayley transform is a conformal mapping from the unit disk onto the upper half-plane, $\Pi^+ = \{z \in \mathbb{C} : \text{Im}(z) > 0\}$. The principal behind the Cayley transform is based on the linear-fractional transformation, principally on the topological mapping. Our interest here is to provide a way to map different topological spaces with the focus on the upper half-plane, but first we will consider this question: "When do we have a complex number in Π^+ in a signature of the Cayley transform?" (The question might be obvious to some, since it is necessary we will carefully explain its meaning.) Consider this: For all $z \in \Pi^+$ that $|z - i| < |z + i|$, since z needs to be closer to $+i$ than $-i$, thus,

$$z \in \Pi^+ \leftrightarrow \left| \frac{z - i}{z + i} \right| < 1.$$

Let ψ_C be a conformal map from the upper half-plane onto the unit disk.

$$\psi_C(z) = \frac{z - i}{z + i}. \tag{3.3}$$

Since there is a conformal mapping from the upper half-plane onto the unit disk, then there exists a function such that $\psi_C^{-1} = \phi_C$, where

$$\phi_C(z) = i\frac{1 + z}{1 - z}. \tag{3.4}$$

The function ϕ_C is the Cayley transform from a unit disk onto Π^+, it follows that ψ_C is the inverse of the Cayley transform.

The next step is to use the Cayley transform to classify different elements of $\text{Aut}(\mathbb{D})$, by showing that $\phi_C : \mathbb{D} \to \Pi^+$ extends to a homeomorphism of $\overline{\mathbb{D}}$ and $\overline{\Pi^+_\infty}$. Define $\overline{\Pi^+_\infty}$ as the closure of the upper half-plane in the extended plane:

$$\overline{\Pi^+_\infty} = \{z \in \mathbb{C} : \text{Im}(z) \geq 0\} \cup \{\infty\}$$

and $\partial_\infty \Pi^+$ to be the extended boundary of $\overline{\Pi^+_\infty}$:

$$\partial_\infty \Pi^+ = \mathbb{R}_\infty = \mathbb{R} \cup \{\infty\}.$$

Note that in some occasion, $\text{Aut}(\mathbb{D})$ and $\text{Aut}(\Pi^+)$ can be induced isomorphically, if there is a function $\phi' \in \text{Aut}(\mathbb{D})$ such that the element of $\text{Aut}(\Pi^+)$ is

$$\psi = \phi_C \circ \phi' \circ \phi_C^{-1}.$$

For non-trivial elements of $\text{Aut}(\mathbb{D})$ we will classify them as *parabolic*, *hyperbolic* or *elliptic* based on the number of their fixed points and locations. More specifically, the non-trivial elements $\phi \in \text{Aut}(\mathbb{D})$ with exactly one fixed point in $\mathbb{D}$ and none in $\partial\mathbb{D}$, ϕ is said to be elliptic. On the other hand, if ϕ has no fixed point in $\mathbb{D}$ but exactly one fixed point in

$\partial\mathbb{D}$, then ϕ is said to be parabolic. In a similar way as in the parabolic case but with two fixed point in $\partial\mathbb{D}$, ϕ is said to be hyperbolic.

Definition 3.14. Let G be a group and $a, b \in G$. A *conjugate* of b to a is when $b = c^{-1}ac$ for some $c \in G$.

In the next few theorems, we will construct the element classification of $\text{Aut}(\mathbb{D})$. Also, note that for now we will overlook the identity case of non-trivial elements of $\phi \in \text{Aut}(\mathbb{D})$, since, this will be elaborated in Section 5, for now we will refer to ([20], p. 269).

Theorem 3.15. *Suppose that $\phi \in \text{Aut}(\mathbb{D})$ is not the identity.*

(i) *If ϕ is elliptic, then ϕ is conjugate to a rotation: There exists a complex number $\alpha \neq 1$ with $|\alpha| = 1$, such that ϕ is conjugate to ψ, where*

$$\psi(z) = \alpha z, \quad \textit{for } z \in \mathbb{D}.$$

(ii) *If ϕ is parabolic, then ϕ is conjugate to a translation in Π^+ : There exists $B \in \mathbb{R}$ with $B \neq 0$ such that the element of $\text{Aut}(\Pi^+)$ corresponding to ϕ is conjugate to ψ, where*

$$\psi(z) = z + B, \quad \textit{for } z \in \Pi^+.$$

(iii) *If ϕ is hyperbolic, then ϕ is conjugate to a dilation in Π^+ : There exists $\delta > 0$ with $\delta \neq 1$ such that the element of $\text{Aut}(\Pi^+)$ corresponding to ϕ is conjugate to ψ, where*

$$\psi(z) = \delta z, \quad \textit{for } z \in \Pi^+.$$

Proof. For part (i), suppose that ϕ is elliptic with $\phi(0) = 0$, Theorem 3.11 shows that ϕ is a rotation.

For part (ii) and (iii), suppose that ϕ is either parabolic or hyperbolic in $\text{Aut}(\Pi^+)$ and has no fixed point in Π^+, and either one or two fixed points on $\partial_\infty\Pi^+$. Then, we can make a change of variables so that $\phi(\infty) = \infty$. Since ϕ can be expressed as a linear-fractional transformation:

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

for some $a, b, c, d \in \mathbb{R}$ with $ad - bc \neq 0$. It may happen that $c = 0$, then

$$\phi(z) = Az + B,$$

for some $A, B \in \mathbb{R}$ with $(A, B) \neq (1, 0)$. If $A = 1$, then ϕ is a translation since ∞ is the only fixed point of ϕ. Suppose now that $A \neq 1$: let $c = B/(1 - A)$, set $\psi(z) = z + c$, and define $\tilde{\phi} = \psi^{-1} \circ \phi \circ \psi$. This shows that $\tilde{\phi}(z) = Az$, which implies that 0 is the only fixed point of ψ, similarly for ∞. In conclusion, ϕ is a translation.

□

We will now use the previous theorem and show that there is a bounded nonconstant function for the nontrivial elements of $\text{Aut}(\mathbb{D})$.

Theorem 3.16. *Suppose that* $\phi \in \text{Aut}(\mathbb{D})$*.*

(i) *If* ϕ *is either elliptic of finite order, parabolic, or hyperbolic, then there exists a bounded nonconstant* $f \in H(\mathbb{D})$ *such that*

$$f \circ \phi = f,$$

and in fact such that $f(z) = f(w)$ *if, and only if,* $w = \phi^n(z)$ *for some integer* n*.*

(ii) *If* u *is elliptic of infinite order and* $f \in H(\mathbb{D})$ *with* $f \circ \phi = f$*, then* f *is constant.*

Proof. We are dividing the proof into three different conditions where we prove each of condition separately, elliptic, parabolic and hyperbolic, starting with the elliptic. The meaning of finite order and infinite order will be referred back to Definition 3.8. Theorem 3.15 shows that ϕ is elliptic and rotational with $\phi(z) = \alpha z$ and $|\alpha| = 1$. Since ϕ has a finite order of N with $\alpha^N = 1$ and $\alpha^k \neq 1$ for $k = 1, 2, ..., N-1$, it follows that $f(z) = z^N$ works for $f \circ \phi = f$. On the other hand, if ϕ has infinite order, then the powers of α are dense in the unit circle so that $f \in H(\mathbb{D})$ and $f \circ \phi = f$, then f must be a constant.

Suppose now that ϕ is parabolic; based on Theorem 3.15 we may use that $\phi \in \text{Aut}(\Pi^+)$ and the parabolic function can be expressed as:

$$\phi(z) = z + B$$

where $B \neq 0$. Furthermore, let

$$f(z) = e^{2\pi iz/|B|}.$$

This shows that f is bounded in Π^+, since $z = x + iy \in \Pi^+$,

$$|f(x + iy)| = e^{\text{Re}(2\pi i(x+iy)/|B|)} = e^{-2\pi y/|B|} < 1.$$

If $f(z) = f(w)$, then $2\pi iw/|B| = 2\pi iz/|B| + 2\pi ik$ for some $k \in \mathbb{Z}$. Thus, $w = z + nB$ with $n = \pm k$, which is the same as $w = \phi^n(z)$.

Suppose that ϕ is hyperbolic; the function is (similarly to earlier, $\phi \in \text{Aut}(\Pi^+)$):

$$\phi(z) = \delta z$$

with a positive $\delta \neq 1$. We need to define a branch of logarithm in Π^+ as

$$\log(re^{it}) = \ln(r) + it, \quad \text{for } r > 0,\ 0 < t < \pi,$$

and set,

$$f(z) = z^{ic} = e^{ic\log(z)}.$$

The idea here is to find $c \neq 0$ such that $f(w) = f(z)$ is satisfied.

Note that

$$|f(re^{it})| = e^{\text{Re}(ic(\ln(r)+it))} = e^{-(ct)}, \quad \text{for } r > 0,\ 0 < t < \pi,$$

this shows that f is bounded in Π^+ for some $c \in \mathbb{R}$. Hence, z and w are both in Π^+. We have $f(z) = f(w)$ if, and only if, $ic\log(z) - ic\log(w) = 2\pi ik$, for some $k \in \mathbb{Z}$ with $c\ln(r_1/r_2) = 2\pi k$. This results that $z = r_1e^{it}$ and $w = r_2e^{it}$, or

$$r_1 = e^{2\pi k/c}r_2,$$

and

$$z = e^{2\pi k/c} w.$$

In conclusion, there exists $c = 2\pi/\ln(\delta)$ that fulfills the relation of $f(w) = f(z)$ if, and only if, $z = \delta^k w$ or $z = \phi^k(w)$ for $k \in \mathbb{Z}$. □

3.2. **Schwarz Reflection Principle.** Schwarz reflection principle is a way to extend a domain of complex functions on a real axis. Suppose that D is an open set of the complex plane, and D is symmetric on the real axis such that $\bar{z} \in D$ whenever $z \in D$.

$$D^+ = \{z \in D : \operatorname{Im}(z) > 0\}, \tag{3.5}$$

$$D^- = \{z \in D : \operatorname{Im}(z) < 0\}, \tag{3.6}$$

$$D^0 = \{z \in D : \operatorname{Im}(z) = 0\}. \tag{3.7}$$

There are various flavors of the Schwarz reflection principle and for our beneficial toward the great Picard theorem, we will only consider the Schwarz reflection principle for holomorphic functions. However, we need the result from the Schwarz reflection principle for harmonic functions, therefore, we will refer e.g. to [10]. Below is the proof of the Schwarz reflection principle for holomorphic function.

Theorem 3.17. *Suppose that D is an open subset of the plane which is symmetric about the real axis. Suppose that $f \in H(D^+)$ satisfies*

$$\lim_{n\to\infty} \operatorname{Im}(f(z_n)) = 0$$

for every sequence (z_n) in D^+ which converges to a point of D^0. Then, f extends to a function $F \in H(D)$, which satisfies

$$F(\bar{z}) = \overline{F(z)}, \quad \text{for } z \in D.$$

Before proceeding with the proof, we need to be aware of that we are assuming that $\operatorname{Im}(f)$ extends continuously to $D^+ \cup D^0$. The problem with this is that we cannot define $F(z) = f(z)$ for $z \in D^0$ since f is undefined on D^0.

Proof. The Schwarz reflection principle for harmonic functions states that there is a function $f = u + iv \in D$ such that v can be extended to a function V in D, and satisfies

$$V(\bar{z}) = -V(z), \quad \text{for } z \in D.$$

Introducing a disk $D_t \subset D$ with t as its center such that $t \in D^0$. Since V is harmonic in D_t, and D_t is simply connected, then there exists a function $f_t \in H(D_t)$ such that $\operatorname{Im}(f_t) = V$ in D_t. Moreover, we know that $\operatorname{Im}(f) = V$ in D_t^+, hence we can shift its position by applying a constant to f_t so that $f = f_t$ in D_t^+.

It follows from (3.7) that $V = 0$ on D^0, which indicates that the function $f_t^k(t)$ takes real values on D_t^0, and therefore, coefficients produced from its derivatives must also be real.

$$f_t(\bar{z}) = \overline{f_t(z)}, \quad \text{for } z \in D_t.$$

We may define $F : D \to \mathbb{C}$ by,

$$F(z) = \begin{cases} f(z), & \text{if } z \in D^+, \\ f_t(z), & \text{if } z \in D_t, \\ \overline{f(\bar{z})}, & \text{if } z \in D^-. \end{cases}$$

We know that that $f_t = f$ in $D^+ \cap D_t = D_t^+$, and it follows that $f_t(z) = \overline{f(\bar{z})}$ for $z \in D_t \cap D^- = D_t^-$. Then, $f_t = f_s$ in $D_t \cap D_s$. The reason we wrote the definition of F with overlapping cases is that the domain of the function defined in each case is an open set. Thus, in order to show that F is holomorphic we need to show that $g \in H(D^-)$, where $g(z) = \overline{f(\bar{z})}$ (since we already know that $f \in H(D^+)$ and $f_t \in H(D_t)$).

We will use power series to show that g is holomorphic in D^-. Suppose that $a \in D^-$ then $\bar{a} \in D^+$, since $f \in H(D^+)$ there exists $r > 0$ and a sequence of complex numbers, (c_n), such that

$$f(z) = \sum_{n=1}^{\infty} c_n (z - \bar{a})^n, \quad \text{where } |z - \bar{a}| < r.$$

It follows that

$$f(\bar{z}) = \sum_{n=1}^{\infty} c_n (\bar{z} - \bar{a})^n, \quad \text{where } |\bar{z} - \bar{a}| < r,$$

so,

$$g(z) = \overline{f(\bar{z})} = \overline{\sum_{n=1}^{\infty} c_n (\bar{z} - \bar{a})^n} = \sum_{n=0}^{\infty} \bar{c}_n (z - a)^n, \quad \text{where } |z - a| < r.$$

Finally, the last step shows that $g \in H(D(a, r))$ for any $a \in D^-$, therefore $g \in H(D^-)$. □

3.3. **The Extended Plane and Carathéodory's Theorem.** Carathéodory's theorem is a way to conformally map an open and simply connected subset of the complex plane onto the unit disk, it is sometimes known as an extension of the Riemann mapping theorem. We shall also provide theorems needed in our construction of a *modular function* that will be introduced in the next section. Herein, we will state a weaker version of the Carathéodory's theorem which will suffice for us to show that all the boundary points are simple, since we have an explicit region bounded by a simply closed curve. (The original Carathéodory's theorem has a more general statement which can be seen e.g. in [1].)

Lets start with a few results on bounded holomorphic functions.

Lemma 3.18. *Suppose that $f \in H(\mathbb{D})$ and $|f(z)| \le 1$ for all $z \in \mathbb{D}$. Suppose that N is a positive integer, $a \in \mathbb{R}$, and let $S = \{re^{it} : 0 < r < 1, a < t < a + 2\pi/(2N)\}$. Suppose that $\gamma : [0, 1] \to \mathbb{D}$ is a continuous map, $\gamma(0) = r_0 e^{ai}$ and $\gamma(1) = r_1 e^{(a+2\pi/(2N))i}$ for some $r_j > 0$, and $\gamma(t) \in S$ for $0 < t < 1$. If $|f(\gamma(t))| < \epsilon < 1$ for all $t \in [0, 1]$, then $|f(0)| < \epsilon^{1/(2N)}$.*

Proof. For simplicity, assume that $a = 0$. Let $g \in H(\mathbb{D})$ be a function and define

$$g(z) = f(z)\overline{f(\bar{z})}$$

and set

$$h(z) = \prod_{j=0}^{N-1} g\Big(e^{2\pi ij/N} z\Big).$$

This shows that $|g| < \epsilon$ at every point of the set $\gamma^* \cup \tilde{\gamma}^*$, so that $|h| < \epsilon$ at every point of the set

$$\Gamma = \bigcup_{j=0}^{N-1} e^{-2\pi ij/N}, \quad \text{for } \gamma^* \bigcup \tilde{\gamma}^*.$$

There exists $0 \in \Omega$ and $\partial\Omega \subset \Gamma$ in a bounded open set $\Omega \subset \mathbb{D}$. Theorem 2.6 shows that there exists $|h(0)| \leq \epsilon$, thus $|f(0)| < \epsilon^{1/(2N)}$. □

Lemma 3.19. *Suppose that $f \in H(\mathbb{D})$ is bounded. Suppose that $0 < \theta < \pi$ and let S be the region*

$$S = \{re^{it} : 1/2 < r < 1, 0 < t < \theta\}.$$

Suppose that the map $\gamma_n : [0,1] \to \mathbb{D}$ is continuous, $\gamma_n((0,1)) \subset S, \gamma_n(0) > 0$ and $\gamma_n(1) = r_n e^{i\theta} (r_n > 0)$ for $n = 1, 2, \ldots$. If $L \in \mathbb{C}$ and $f \circ \gamma_n \to L$ uniformly on $[0,1]$ as $n \to \infty$, then $f = L$ everywhere in $\mathbb{D}$.

Proof. Subtracting a constant and dividing by another constant we may assume that $L = 0$ and $|f| < 1$ in $\mathbb{D}$. If we apply Lemma 3.18 and define $g, h \in H(\mathbb{D})$ with an integer N such that $2\pi/(2N) < \theta$, then there exists $|f| < \epsilon$ at every point of the set γ_n^*. If we replace the curve γ_n^* by γ as in Lemma 3.18, then we receive $|h| \leq \epsilon$ at every point of $D(0, 1/2)$ with the fact that $\epsilon > 0$ and h must disappear identically on $D(0, 1/2)$. Thus, f is identically zero. (If f is not zero, then it would have only finitely many zeroes in $D(0, 1/2)$, since h has finitely many zeroes.) □

The next subject is the Lindelöf's theorem and it states that a holomorphic function on the bounded region grows "moderately" in the unbounded direction. Hereby, we will use Lemma 3.18 and Lemma 3.19 in the proof, where we have shown that there is a restriction to the functions with limits, $L = 0$, and $|f| < 1$ for $f \in H(\mathbb{D})$. Below is the Lindelöf's theorem.

Theorem 3.20. *Suppose that $\gamma : [0,1] \to \overline{\mathbb{D}}$ is continuous, $\gamma(t) \in \mathbb{D}$ for $0 \leq t < 1$ and $\gamma(1) = 1$. Suppose that $f \in H(\mathbb{D})$ is bounded. If $f(\gamma(t)) \to L$ as $t \to 1$, then $\lim_{r \to 1^-} f(r) = L$.*

Proof. We are going to show that there is a function $|f(r)| < \epsilon^{1/6}$ such that $R < r < 1$ for an arbitrary $\epsilon > 0$, then $\lim_{r \to 1^-} f(r) = 0$. We may also assume that $L = 0$ with the same argument we used in Lemma 3.19.

Let $\psi_r \in \mathrm{Aut}(\mathbb{D})$,

$$\psi_r(z) = \frac{z + r}{1 + rz},$$

where $-1 < r < 1$.

Furthermore, let $C_0 = [-i, i] \subset \overline{\mathbb{D}}$ be a curve and $C_r = \psi_r(C_0)$, and there exists $t_r \in (0, 1)$ such that $\gamma(t_r) \in C_r$ and $\gamma(t) \notin C_r$ for $t_r < t < 1$. Also let $\gamma_r = \gamma|_{[t_r,1)}$.

Define $\Gamma_r = \psi_r^{-1} \circ \gamma_r$ and $p_r = \psi_r^{-1}(\gamma_r(t_r)) = \Gamma_r(t_r)$. This shows that $p_r \in [-i, i]$, since $\gamma_r(t_r) \in C_r$ and $C_r = \psi_r(C_0)$ where $C_0 = [-i, i]$. If $p_r = 0$, then $f(r) = f(\gamma_r(t_r))$, so that $|f(r)| < \epsilon < \epsilon^{1/6}$. Suppose $p_r \neq 0$, then $p_r = iy$ for some $y \neq 0$, by the assumption of $y > 0$.

If $F = f \circ \psi_r$, then $|F| < \epsilon$ on Γ_r^*. The curve Γ_r lies in the right half-plane with end-points at 1 and iy (from what we have defined earlier). Thus, this could lead to a problem since 1 lies outside $\mathbb{D}$, and in order to handle this we apply Lemma 3.18 with

$$S = \{\rho e^{it} : 0 < \rho < 1, \pi/6 < t < \pi/2\}.$$

Hence, $|F(0)| < \epsilon^{1/6}$. But we already know that $f(r) = F(0)$, which conclusively gives $|f(r)| < \epsilon^{1/6}$. □

We will tie the previous lemmas into a proposition that will mostly cover all the context of the Carathéodory's theorem, but first we need a definition on the simple boundary point.

Definition 3.21. Let D be an open subset of the complex plane and $\zeta \in \partial D$. We will say that ζ is a *simple boundary point*, if (z_n) is any sequence of points in D such that $z_n \to \zeta$, then there exists a continuous map $\gamma : [0, 1) \to D$ and a sequence (t_n) in $[0, 1)$ such that $t_n \leq t_{n+1}$, $t_n \to 1$, $\gamma(t_n) = z_n$, and $\lim_{t\to 1} \gamma(t) = \zeta$.

Proposition 3.22. *Suppose D is a bounded simply connected open set in the plane, and let $\phi : D \to \mathbb{D}$ be a conformal equivalence.*

- (i) *If ζ is a simple boundary point of D, then there exists $L \in \partial\mathbb{D}$ such that $\phi(z_n) \to L$ whenever (z_n) is a sequence in D tending to ζ. (In other words, ϕ can be extended continuously to the set $D \cup \{\zeta\}$, and the extended function satisfies $\phi(\zeta) \in \partial\mathbb{D}$).*
- (ii) *Suppose that ζ_1 and ζ_2 are two simple boundary points; choose corresponding complex numbers L_1 and L_2 as in (i). If $\zeta_1 \neq \zeta_2$, then $L_1 \neq L_2$.*

Proof. First note that if (z_n) is any sequence in D tending to a point of the boundary, then $|\phi(z_n)| \to 1$: if $r < 1$ with K as a compact subset of D and $K = \phi^{-1}(\overline{D}(0, r))$, then there exists N such that $z_n \notin K$ for all $n > N$. Hence, $|\phi(z_n)| > r$ for every $n > N$.

Part (i), let ζ be a simple boundary point of D with a sequence (z_n) tends to ζ. Since there is a bounded and closed function in the disk, then there exists $L \in \overline{\mathbb{D}}$ and a subsequence (z_{n_j}) such that $\phi(z_{n_j}) \to L$. To prove this, we need to show that every sequence (z_n) of $\phi(z_n)$ tends to L.

A simple way to prove this is by proving its contradiction, which would be stated as: the function $(\phi(z_n))$ has two subsequences that are pursuing to various limits L_1 and L_2, and that $L_1 \neq L_2$, see Figure 3.2. By adjusting the notation we can assume that $\phi(z_{2n}) \to L_1$ and $\phi(z_{2n+1}) \to L_2$, where $L_j \in \partial\mathbb{D}$. Moreover, we know that ζ is a simple bounded point and there exists a continuous curve $\gamma : [0, 1) \to D$ and a sequence (t_n) in $[0, 1)$ such that $t_n \leq t_{n+1}$ where $t_n \to 1$, implies that $\gamma(t_n) = z_n$ and $\lim_{t\to 1}\gamma(t) = \zeta$.

If we now let $\Gamma = \phi \circ \gamma$, then Γ should be lying in $[0, 1)$ such that $\Gamma : [0, 1) \to \mathbb{D}$, thus, $|\Gamma(t)| \to 1$ as $t \to 1$. However, the curve Γ is diverging from a single point of $\partial\mathbb{D}$, since

$\Gamma(t_{2n}) \to L_1$ and $\Gamma(t_{2n+1}) \to L_2$. For this to occur, a reasonable presumption is that the curve Γ must cross some sector S infinitely many times as it tends to $\partial\mathbb{D}$, it would either be S_1 or S_2 and this can be seen in Figure 3.2. When n is large enough, there exist numbers a_n and b_n lying in between $t_{2n} < a_n < b_n < t_{2n+1}$, with curves $\gamma(a_n)$ and $\gamma(b_n)$ locating on edges of the S (not at the same edge) and $\gamma((a_n, b_n))$ lies in the interior of S. Let $\gamma_n = \gamma|_{[a_n,b_n]}$.

If we let $f = \phi^{-1}$, then $f \in H(\mathbb{D})$ and f is bounded. Then, we have $f(\Gamma(t)) \to \zeta$ as $t \to 1$, so $f(\Gamma(t)) = \gamma(t)$, since $f \circ \gamma_n \to \zeta$ uniformly. Thus, f is constant according to Lemma 3.19, which is a contradiction to the fact that f is a bijection from $\mathbb{D}$ to D.

Part (ii), suppose that ζ_1 and ζ_2 are simple boundary points of D and $\zeta_1 \neq \zeta_2$. Part (i) shows that, ϕ can be extended to a function continuous on $D \cup \{\zeta_1, \zeta_2\}$, and that $|L_j| = 1$, if $L_j = \phi(\zeta_j)$, where $j = 1, 2$.

We need to show $L_1 \neq L_2$; start by employing the continuous curves $\gamma_j : [0, 1) \to D$ in such way that $\gamma_j(t) \to \zeta_j$ as $t \to 1$. Also, let $\Gamma_j = \phi \circ \gamma_j$ and as before, there exists an inverse function ϕ^{-1} such that $f = \phi^{-1}$. Hence, f is bounded in $H(\mathbb{D})$ and the curve Γ_j lies in $\mathbb{D}$ tending to L_j, so $f(\Gamma_j(t)) \to \zeta_j$ as $t \to 1$. Lastly, Theorem 3.20 shows that $f(rL_j) \to \zeta_j$ as r approaches to 1, thus the points $\zeta_1 \neq \zeta_2$, and therefore, $L_1 \neq L_2$. □

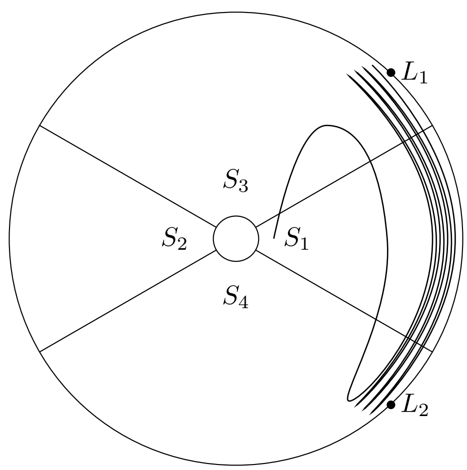

Figure 3.2 – Circle with four sections and two boundary points

Lemma 3.23. *Suppose that D is a bounded open set in the complex plane, $\phi : D \to \mathbb{C}$ is continuous, and for every $\zeta \in \partial D$ there exists a function ϕ_ζ continuous on $D \cup \{\zeta\}$ which agrees with ϕ on D. Then ϕ extends to a function continuous on $\overline{D}$.*

Proof. We start by extending ϕ to a function $\psi : \overline{D} \to \mathbb{C}$ such that

$$\psi(\zeta) = \phi_\zeta(\zeta), \quad \text{for } \zeta \in \partial D.$$

Suppose that $\zeta \in \partial D$ and $|z - \zeta| < \delta$ for all $z \in D$, we want to show that ψ is continuous at ζ. If ϕ_ζ is continuous on $D \cup \{\zeta\}$, and there exists $\delta > 0$ such that $|\psi(z) - \psi(\zeta)| < \epsilon$

with $\epsilon > 0$, then $|z - \zeta| < \delta$. Let $w \in \partial D$, then $\psi(z) = \lim_{w \to z} \psi(w)$, and this follows that $|\psi(z) - \psi(\zeta)| \leq \epsilon$. $\square$

Finally, we are now able to state our anticipated theorem, the Carathéodory's theorem. To demonstrate this, we are going to use that a function can be continuously extended on $\overline{\mathbb{D}}$ and that there is a sequence tending to a constant on the plane, which has been shown in Lemma 3.23 and Proposition 3.22.

Theorem 3.24. *Suppose D is a bounded simply connected domain in the complex plane and every boundary point of D is simple. Let $\phi : D \to \mathbb{D}$ be a conformal equivalence. Then ϕ extends to a homeomorphism of $\overline{D}$ and $\overline{\mathbb{D}}$.*

Proof. We have seen that ϕ can be continuously extended to $D \cup \{\zeta\}$ for every $\zeta \in \partial D$. Also that ϕ extends continuously to the closure of D from Lemma 3.23 and Proposition 3.22 (*i*).

For the extended function, $\phi(\overline{D})$ must be a compact subset of $\overline{\mathbb{D}}$ containing $\mathbb{D}$ such that $\phi(\overline{D}) = \overline{\mathbb{D}}$. Proposition 3.22 (*ii*) shows that ϕ is a bijection of $\overline{D}$, so that there exists an inverse function $\phi^{-1} : \overline{\mathbb{D}} \to \overline{D}$. The only thing that remains is to show that the inverse of ϕ is continuous on $\overline{\mathbb{D}}$. Since $\overline{D}$ is compact and there is a continuous bijection $\phi : \overline{D} \to \overline{\mathbb{D}}$, then there exists an open set $\mathcal{O} \subset \overline{D}$ such that $(\phi^{-1})^{-1}(\mathcal{O}) = \phi(\mathcal{O}) \subset \overline{\mathbb{D}}$. Thus, $\phi^{-1} : \overline{\mathbb{D}} \to \overline{D}$ is continuous. $\square$

4. A Modular Function

The goal with this section is to construct a subgroup $\Gamma(2)$ of the modular group Γ, which is a group consisting of all the linear-fractional transformations. Furthermore, we will demonstrate that the subgroup $\Gamma(2)$ is freely generated by elements τ and σ, and see that this is the implication from Lemma 4.4.

A modular function is a meromorphic function in an upper half-plane Π^+, that is invariant under certain actions of the modular group Γ. Let us recall the linear-fractional transformations, ϕ from Definition 3.9,

$$\phi(z) = \frac{az+b}{cz+d} \tag{4.1}$$

where a, b, c and d are integers and $ad - bc \neq 0$.

The modular group, Γ is a subgroup of all the transformations of $\mathrm{Aut}(\Pi^+)$ correspond to the group $\mathrm{SL}_2(\mathbb{Z})$ (special linear group, $\mathrm{SL}_2(\mathbb{Z})$ is the group of 2×2 matrices belonging to a so called *Lie group*. We will leave details of the Lie groups e.g. to [6]). The subgroup, $\Gamma(2)$ is consisting of all ϕ with requirements of a and d being odd, b and c being even. If so, this constructs

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}). \tag{4.2}$$

Hence, the condition required for $\Gamma(2)$,

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \equiv \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mod 2.$$

This shows that, $\Gamma(2)$ is a subgroup of Γ, because every matrix of $\Gamma(2)$ is the identity matrix with mod 2. Now, let σ and τ be defined by

$$\sigma(z) = \frac{z}{-2z+1} \tag{4.3}$$

and

$$\tau(z) = z + 2. \tag{4.4}$$

Note that (4.3) and (4.4) can be written in matrix form by applying the relations of (4.1) and (4.2):

$$\tau^n = \begin{bmatrix} 1 & 2n \\ 0 & 1 \end{bmatrix}, \ \sigma^n = \begin{bmatrix} 1 & 0 \\ -2n & 1 \end{bmatrix}$$

and

$$\tau^n \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a+2nc & b+2nd \\ c & d \end{bmatrix},$$

similarly,

$$\sigma^n \begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} a & b \\ c-2na & d-2nb \end{bmatrix}.$$

Lemma 4.1. *The group* $\Gamma(2)$ *is generated by* σ *and* τ*.*

Proof. It is clear that σ and τ are elements from $\Gamma(2)$. Let G be the subgroup of $\Gamma(2)$ that generated by σ and τ. Meaning, we need to show that $\Gamma(2) = G$.

Define $\chi : \Gamma(2) \to \mathbb{Z}^+$ by

$$\chi\left(\begin{bmatrix} a & b \\ c & d \end{bmatrix}\right) = |a| + |c|.$$

Suppose that $\phi \in \Gamma(2)$. Hence, we need to show that $\phi \in G$. Define

$$G \circ \phi = \{\psi \circ \phi : \psi \in G\}.$$

Since, χ can only attain positive integers, then there must exists

$$\phi_0 = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in G \circ \phi,$$

with $\chi(\phi_0)$ minimal.

Suppose that $c \neq 0$. If $|a| > |c|$, then this would lead to a contradiction which is that there exists $n \in \mathbb{Z}$ such that $-|c| \leq a + 2nc < |c|$. This implies that $-|c| < a + 2nc < |c|$ since, a is odd and c is even, hence,

$$\chi(\tau^n \phi_0) = |a + 2nc| + |c| < |c| + |c| < |a| + |c| = \chi(\phi_0),$$

which contradicts our choice of ϕ_0. If we now let, $|a| \leq |c|$, then $0 < |a| < |c|$ since a is odd, and with a similar to the previous argument that there exists $n \in \mathbb{Z}$ such that $\chi(\sigma^n \phi_0) < \chi(\phi)$, once again our choice of ϕ_0 is unsatisfied.

So this only works if $c = 0$. Since, $ad - bc = 1$ implies that $a = d = \pm 1$, so that $\phi_0 = \tau^n$ for some $n \in \mathbb{Z}$. In particular, $\phi_0 \in G$. Hence, $\phi \in G$. $\square$

From Lemma 4.1, one can see that σ and τ are conjugate in Γ. We define $j \in \Gamma$ by

$$j(z) = \frac{-1}{z}. \tag{4.5}$$

So that

$$\sigma = j^{-1} \circ \tau \circ j.$$

Now, we define new σ and τ that only cover the upper half-plane:

$$T = \{z \in \Pi^+ : -1 \leq \operatorname{Re}(z) < 1\} \tag{4.6}$$

and

$$S = \left\{z \in \Pi^+ : \left|z + \frac{1}{2}\right| > \frac{1}{2},\ \ \left|z - \frac{1}{2}\right| \geq \frac{1}{2}\right\}. \tag{4.7}$$

The illustrations of (4.6) and (4.7) are shown in Figure 4.1 and Figure 4.2.

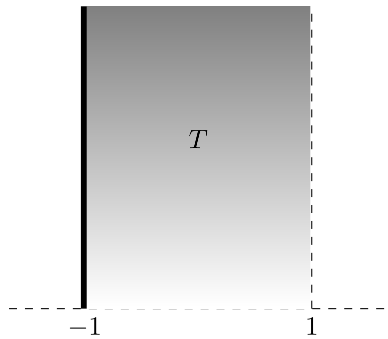

Figure 4.1

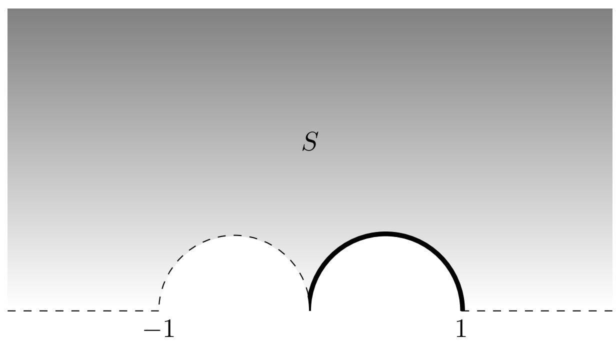

Figure 4.2

Lemma 4.2. *The following statements are true:*

(*i*) *the sets* $\tau^n(T), (n \in \mathbb{Z})$ *are disjoint, and their union is* Π^+*;*

(*ii*) *the sets* $\sigma^n(S), (n \in \mathbb{Z})$ *are disjoint, and their union is* Π^+*;*

(*iii*) *if* $n \in \mathbb{Z}$ *and* $n \neq 0$*, then* $\tau^n(T) \subset S$*;*

(*iv*) *if* $n \in \mathbb{Z}$ *and* $n \neq 0$*, then* $\sigma^n(S) \subset T$*.*

Proof. The statements (*i*) and (*ii*) can be drawn from Figure 4.1 and Figure 4.2, which indicate that T is a fundamental domain of the subgroup $\Gamma(2)$ generated by τ, and similarly for S generated by σ. Part (*iii*) and (*iv*) follow from part (*i*) and (*ii*), since $\sigma = j^{-1} \circ \tau \circ j$ and $S = j(T)$, where $j(z) = -1/z$ takes lines and circles to lines and circles with the preservation of angles. □

Now, we are going to introduce another concept in a connection with the previous results for determining the fundamental domain of the modular functions. This concept is often known as the Ping-Pong lemma (also as Klein's criterion), which is a group action on a freely generated group (more of this can be seen e.g. in [8]).

Let G be a group generated by $a, b \in G$. Then, a *nontrivial reduced word* of a and b is expressed as:

$$g_1^{n_1} g_2^{n_2} ... g_k^{n_k} \tag{4.8}$$

where $k \geq 1$, each n_j is a nonzero integer, and each g_j is either a or b. For $k = 0$, we would have the identity element of G.

Remark 4.3. A group G is *freely generated* by a and b, meaning that G is generated by a and b, and no nontrivial reduced word in a and b is the identity element of G.

Lemma 4.4 is known as the Ping-Pong lemma:

Lemma 4.4. *Suppose that G is a group of bijections of a set X (with composition as the group operation), and G is generated by a, $b \in G$. Suppose that $A, B \subset X$ satisfy the following:*

- (i) *the sets $a^n(A), (n \in \mathbb{Z})$ are disjoint, and their union is X;*
- (ii) *the sets $b^n(B), (n \in \mathbb{Z})$ are disjoint, and their union is X;*
- (iii) *if $n \in \mathbb{Z}$ and $n \neq 0$, then $a^n(A) \subset B$;*
- (iv) *if $n \in \mathbb{Z}$ and $n \neq 0$, then $b^n(B) \subset A$.*

Let $E = A \cap B$, and assume that $E \neq \varnothing$. If w is a nontrivial reduced word in a and b, then

$$E \cap w(E) = \varnothing.$$

Proof. Assume that, w starts and ends with a power of a as in (4.8), and as the following $w = a^{n_1} b^{n_2} ... a^{n_k}$. If we apply (iii) and (iv) repeatably, then this shows that

$$w(A) = a^{n_1} b^{n_2} ... b^{n_{k-1}} a^{n_k}(A) \subset a^{n_1} b^{n_2} ... b^{n_{k-1}}(B) \subset ... \subset a^{n_1}(A).$$

Part (i) shows that $A \cap w(A) = \varnothing$, which implies that $(A \cap B) \cap w(A \cap B) = \varnothing$. Similarly, we can use this to prove part (ii), if we let w ends with a power of b so that $w = a^{n_1} b^{n_2} ... b^{n_k}$. This gives: $w(B) \subset a^{n_1}(A)$ and $A \cap w(B) = \varnothing$, thus $(A \cap B) \cap w(A \cap B) = \varnothing$. The results show that the sets $a^n(A)$ and $b^n(B)$ are disjoint. □

We are one step closer to be able to construct a fundamental domain of the modular group, $\Gamma(2)$ with the aids from Lemma 4.2 and Lemma 4.4. The fundamental domain we want to construct contains the following properties:

Remark 4.5. We want to construct a set $Q \subset \Pi^+$ such that:

- (i) the sets $\phi(Q)$ for $(\phi \in \Gamma(2))$ are disjoint;
- (ii) $\Pi^+ = \bigcup_{\phi \in \Gamma(2)} \phi(Q)$.

Again, the modular function we want is a meromorphic function in Π^+. Part (i) follows from Lemma 4.2 and Lemma 4.4. On the other hand, this might be unclear for part (ii) because the set E does not have to be a fundamental domain of G. We know that $\Gamma(2)$ is a group and by showing that $z \in \Pi^+$ there exists $\phi \in \Gamma(2)$ with $\phi(z) \in Q$, which can be used to prove part (ii). (A proof of these is the last in this section)

One might wonder what was the connection between the Ping-Pong lemma and our modular group of the upper half-plane. The definitions below might be able to clarify that inquiry.

$$\text{ping} : \Pi^+ \to T \quad \text{and} \quad \text{pong} : \Pi^+ \to S.$$

The unique element of T is $\text{ping}(z) = \tau^n(z)$ for some $n \in \mathbb{Z}$, similarly for $\text{pong}(z)$ but for element S, $\text{pong}(z) = \sigma^n(z)$ for some $n \in \mathbb{Z}$. The reason to define them in such way is that we want to find $\phi \in \Gamma(2)$ such that $\phi(z) \in Q$ for given $z \in \Pi^+$. Moreover, if there is such ϕ, then there is a unique sequence that must converge at a point of Q, hence it is more reasonable to define sets $z_1 = \text{ping}(z)$ and $z_2 = \text{pong}(z_1)$.

The purpose here is to make ping and pong move closer to i, since i is the upper half-plane. We will also see that ping and pong are constants and invariants under j defined as in (4.5), otherwise, they will approach to the extended upper half-plane, $\partial_\infty \Pi^+$. Thus, we need to define a certain function:

Define $\rho : \Pi^+ \to [0, \infty)$ by

$$\rho(z) = \left| \frac{z - i}{z + i} \right|. \tag{4.9}$$

If we have $\Phi : \Pi^+ \to \mathbb{D}$ as the Cayley transform, then this would give $\rho(z) = |\Phi(z)|$. As we have stated in Section 3, that such function has limits $0 \le \rho(z) < 1$ for all $z \in \Pi^+$. Furthermore, the set of all z with $\rho(z) \le r$ is a compact subset of Π^+ for any $r \in (0, 1)$.

The function ρ as in (4.9) satisfies the following properties:

Lemma 4.6. *For ρ as in (4.9), j as in (4.5) and for every $z \in \Pi^+$, we have*

(*i*) $\rho(j(z)) = \rho(z)$;

(*ii*) $\rho(\text{ping}(z)) \le \rho(z)$;

(*iii*) $\rho(\text{pong}(z)) \le \rho(z)$.

Proof. Part (*i*) is straightforward, since $j(z) = -\frac{1}{z}$ and with some algebra it can be shown that (*i*) is satisfied:

$$\rho(j(z)) = \left| \frac{\frac{-1}{z} - i}{\frac{-1}{z} + i} \right| = \left| \frac{(1 + zi)}{(1 - zi)} \right| = \left| \frac{i}{i} \frac{(1 + zi)}{(1 - zi)} \right| = \left| \frac{z - i}{z + i} \right| = \rho(z)$$

For part (*ii*), we are using the fact that $z = x + iy$ and assuming $z \notin T$, where T is defined as (4.6), thus,

$$(\rho(z))^2 = 1 - \frac{4y}{x^2 + (y + 1)^2}.$$

If $z \in T$, then we would have $\text{ping}(z) = z$ since our $z \notin T$, and $\text{ping}(z) = x' + iy$ for some $x' \in \Pi^+$. This leads to $|x| \ge 1$ and $|x'| \le 1$, which concludes (*ii*). Similarly for part (*iii*), which follows from part (*i*) and (*ii*), together with the fact that $\sigma = j^{-1} \circ \tau \circ j$. □

One of the main concerns here is to ensure that sequences of the holomorphic function in a disk D are bounded, which is significant to Remark 4.5 when we apply the Ping-Pong lemma to construct the fundamental domain. Now, it is a suitable moment to look at the compactness of our holomorphic function in Π^+ and $\partial_\infty \Pi^+$. With the reason of that sequences ping and pong will eventually tend to the extended upper half-plane, $\partial_\infty \Pi^+$.

Lemma 4.7. *Suppose that $f_n : \mathbb{D} \to \mathbb{D}$ is holomorphic. If $z \in \mathbb{D}$ and $|f_n(z)| \to 1$, then $|f_n(w)| \to 1$ for every $w \in \mathbb{D}$.*

Proof. We will prove a contradiction of the statement. It follows from Theorem 2.9, that there exists a subsequence (f_{n_j}) of (f_n) such that f converges uniformly on compact sets to $f \in H(\mathbb{D})$. Suppose that there is $w \in \mathbb{D}$ with the subsequence (f_{n_j}) such that $f_{n_j}(w) \to \alpha$ with $|\alpha| < 1$. It implies that $|f| \leq 1$ in $\mathbb{D}$. Since $|f(z)| = 1$, and f is a constant according to Theorem 2.6, which contradicts the fact that $|f(w)| < 1$. □

From Lemma 4.7 follows by two corollaries.

Corollary 4.8. *Suppose that $f_n : \Pi^+ \to \Pi^+$ is holomorphic, $z \in \Pi^+$ and $f_n(z) \to \partial_\infty \Pi^+$. Then $f_n(w) \to \partial_\infty \Pi^+$ for all $w \in \Pi^+$.*

Proof. If $(z_n) \subset \Pi^+$, then we will say that $z_n \to \partial_\infty \Pi^+$. If for every compact set $K \subset \Pi^+$, then there exists N so that $z_n \in \Pi^+ \setminus K$ for all $n > N$. Moreover, if we let $\Phi : \Pi^+ \to \mathbb{D}$ be the Cayley transform that leads to $z_n \to \partial_\infty \Pi^+$ with the only requirement, which is $|\Phi(z_n)| \to 1$. The proof is concluded according to Lemma 4.7. □

A short notice, that if we have ρ as in (4.9) and $z \in \Pi^+$ with a, b, c and d as elements in a group of $\mathrm{SL}_2(\mathbb{Z})$, then

$$\left(\rho\left(\frac{ai+b}{ci+d}\right)\right)^2 = \frac{a^2+b^2+c^2+d^2-2}{a^2+b^2+c^2+d^2+2}. \tag{4.10}$$

Corollary 4.9. *If (ϕ_n) is a sequence of distinct elements of Γ (that is, $\phi_n \neq \phi_m$ for $n \neq m$), then $\phi_n(z) \to \partial_\infty \Pi^+$ for all $z \in \Pi^+$.*

Proof. As we have shown in (4.10) that $\rho(\phi_n(i)) \to 1$, since for a given $R < \infty$ there are only finitely many elements $(a, b, c, d) \in \mathbb{Z}^4$ with $a^2 + b^2 + c^2 + d^2 < R$. This implies that $|\Phi(\phi_n(i))| \to 1$, so that $\phi_n(i) \to \partial_\infty \Pi^+$. Hence, $\phi_n(z) \to \partial_\infty \Pi^+$ for all $z \in \Pi^+$ according to Corollary 4.8. □

For future convention we will use this opportunity to tie up a few theorems and lemmas in this section into one single theorem, for constructing the normal subgroup, $\Gamma(2)$. First, we need to show that the group $\Gamma(2)$ is in Γ.

Theorem 4.10. *The group $\Gamma(2)$ is freely generated by τ and σ. Further, $\Gamma(2)$ is a normal subgroup of Γ, every nontrivial element of $\Gamma(2)$ is either parabolic or hyperbolic, and if $Q = S \cap T$, then*

(*i*) *the sets of $\phi(Q)$ for ($\phi \in \Gamma(2)$) are disjoint;*

(*ii*) $\bigcup_{\phi \in \Gamma(2)} \phi(Q) = \Pi^+$.

Proof. To prove (*i*), we will use Lemma 4.2 and Lemma 4.4, these show that $\Gamma(2)$ is freely generated by σ and τ. Thus, (*i*) holds. In part (*ii*), we need to show for $z \in \Pi^+$ there exists $\phi \in \Gamma(2)$ such that $\phi(z) \in Q$. Assume that $z \in \Pi^+ \setminus Q$.

Since, $\Pi^+ = S \cup T$ we must either have $z \in S \setminus T$ or $z \in T \setminus S$. Also, by the fact that these two cases are equivalent by j as in (4.5), hereby, we will use the case of $z \in S \setminus T$. Let $z_1 = \mathrm{ping}(z), z_2 = \mathrm{pong}(z_1), z_3 = \mathrm{ping}(z_1)$, etc. In particular, $z_1 \in T, z_2 \in S, z_3 \in T$, etc., and there exists $\phi_n \in \Gamma(2)$ with $z_n = \phi_n(z)$. If there exists n with $z_n \in Q$, then we are finished with the proof. Suppose not, then we would have $z_{n+1} \neq z_n$ for all $n \geq 2$.

Now, the definitions of ping and pong show that for every m and n with $m > n$, there exists w, a nontrivial word in σ and τ such that $\phi_m = w \circ \phi_n$. This implies that $\phi_m \neq \phi_n$ for $m > n$. Corollary 4.9 states that there is a sequence $z_n \to \partial_\infty \Pi^+$, which is the contradiction to Lemma 4.6 since, $\rho(z_n) \leq \rho(z) < 1$ for all n, so that the sequences (z_n) are all contained in some fixed compact subset of Π^+.

Suppose that $\psi \in \Gamma(2), z \in \Pi^+$ and $\psi(z) = z$. Choose $\phi \in \Gamma(2)$ so that $z \in \phi(Q)$. We also have $z = \psi(z) \in \psi \circ \phi(Q)$ and (i) implies that $\phi = \psi \circ \phi$. Hence, ψ is the identity.

Furthermore, suppose that

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \Gamma(2),$$

then

$$A = I + 2M$$

for some matrix M with integer entries. Suppose that $B \in \mathrm{SL}_2(\mathbb{Z})$, then

$$BAB^{-1} = I + 2BMB^{-1},$$

since B^{-1} also has integer entries it follows that $BAB^{-1} \in \Gamma(2)$. Hence, $\Gamma(2)$ is a normal subgroup of Γ. Also that Q is a fundamental domain of $\Gamma(2)$. $\square$

5. Covering Maps

We have introduced the conformal map in Section 3, which is a function that preserves all angles locally. Now, we are going to give more topological properties and introduce another kind of projection, namely *covering projection* or *covering maps*. At the end of this section, finally, the readers should be able to see the connection between the fundamental domain that we just proved in the previous section and the covering maps. Lets start by defining covering maps.

Definition 5.1. Let X and Y be topological spaces such that there is a map $p : X \to Y$ that is continuous, then p is said to be a *covering map*, if each point of z in Y has a connected open neighborhood N such that

$$p^{-1}(N) = \bigcup_{\alpha \in A} \tilde{N}_\alpha,$$

where each $\tilde{N}$ is an open subset of X and $\tilde{N}$ is pairwise disjoint. Each of the point $\alpha \in A$ is mapped homeomorphically from $\tilde{N}$ onto N by p. For any open neighborhood N satisfies the condition above, N is then called *elementary neighborhood*.

Besides the definition of the covering maps we also need several definitions from topology.

Definition 5.2. In a topological context, a *homotopy* is a continuous deformation of one continuous function to another continuous function.

Definition 5.3. A topological space is *simply connected*, if any two paths between two points are homotopic to each other, i.e. there exists a homotopy between them.

Definition 5.4. Let X, Y and $\tilde{X}$ be topological spaces, and let $g : X \to \tilde{X}$ be a covering map. A *lift* of a function $f : X \to Y$ is a function $h : \tilde{X} \to Y$ such that $h \circ g = f$.

Now that we have defined the covering maps, we need to return to Section 4 and tie up one last resource for our modular function. So far, we have shown that the normal subgroup of Γ is generated by τ and σ precisely as in Theorem 4.10. Moreover, we have constructed a fundamental domain of $\Gamma(2)$ such that $Q \in \Pi^+$. The next theorem is one of the main pieces for the great Picard theorem.

Theorem 5.5. *There exists a holomorphic covering map* λ *for* Π^+ *onto* $\mathbb{C} \setminus \{0, 1\}$, *with the following properties:*

- (i) $\lambda \circ \phi = \lambda$ *for all* $\phi \in \Gamma(2)$*;*
- (ii) λ *is one-to-one on* Q*;*
- (iii) *if* (z_j) *is a sequence in* Π^+ *with* $\operatorname{Im}(z_j) \to \infty$, *then* $\lambda(z_j) \to \infty$.

Proof. In this, we will start off with the conformal maps case and then afterward, we will show that λ is a holomorphic covering map. For part (i) and (ii), we will start our proof by defining Q^+ as an interior of the right half of Q where Q is a fundamental domain of $\Gamma(2)$,

$$Q^+ = \left\{ z = x + iy : y > 0, 0 < x < 1, \left| z - \frac{1}{2} \right| > \frac{1}{2} \right\}.$$

At this point, we have proved various theorems, and a few of them will be recalled with the significative interest to this proof. Theorem 2.10 shows that there is a conformal map h from Q^+ onto Π^+, and together with Theorem 3.24, it shows that h extends to a homeomorphism of the extended closure of Q^+, that is, $\overline{Q^+} \cup \{\infty\}$, onto $\overline{\Pi^+_\infty}$. Moreover, if there is a bounded simply connected open set and every point of it is simple, for a positively oriented triple of distinct points e.g. $(0, 1, \infty)$, there is a conformal equivalence of h such that $h(0) = 0, h(1) = 1$ and $h(\infty) = \infty$ (proof of this can be seen e.g. in ([20], p. 310)).

Let us also recall Theorem 3.17, which indicates that there is a reflection on a real axis such that:

$$h(x + iy) = \overline{h(-x + iy)}$$

for $h(iy) \in \mathbb{R}$ and $y > 0$. Notice that, $x + iy$ signifies the upper right of the circle and $-x + iy$ the upper left of the circle, which are symmetric according to Theorem 3.17. Since there is a reflection $h^{-1}(\{1\}) = \{1, -1\}, h^{-1}(\{0\}) = \{0\}, h^{-1}(\{\infty\}) = \{\infty\}$ and $(\overline{Q} \cup \{\infty\}) \setminus \Pi^+ = \{-1, 0, 1, \infty\}$, which in fact gives $h(\overline{Q} \cap \Pi^+) = \mathbb{C} \setminus \{0, 1\}$, so

$$h(Q) = \mathbb{C} \setminus \{0, 1\}.$$

The conformal map, h is one-to-one on the closure of Q^+, additionally, there exists a reflection so that h is one-to-one on Q, and containing sequences of z_n in Q such that if $\text{Im}(z_n) \to \infty$, then $h(z_n) \to \infty$. Define a function $\tilde{h} : Q^+ \to \mathbb{C}$ by

$$\tilde{h}(z) = \frac{1}{h(j(z))},$$

where $j(z)$ is a function in the modular group Γ defined as in (4.5). Since, j defines the conformal map from Q^+ onto the interior of the left half of Q, thus, $h \circ j$ defines a conformal map from Q^+ onto the lower half-plane, then this indicates that $\tilde{h}$ is a conformal map from Q^+ onto Π^+.

This shows that $\tilde{h}$ extends continuously to the extended closure, $\overline{Q^+} \cup \{\infty\}$, and one can verify that by: $\tilde{h}(1) = 1 = h(1), \tilde{h}(0) = 0 = h(0)$ and $\tilde{h}(\infty) = \infty = h(\infty)$ (which is similar as when we argued for h). Then, $\tilde{h} = h$ in Q^+, and by the continuity and uniqueness of analytic continuation (can be seen e.g. in ([20], p. 298-299)) it follows that

$$h(j(z)) = \frac{1}{h(z)}, \quad \text{for } z \in Q.$$

We now define $\lambda : \Pi^+ \to \mathbb{C}$ by

$$\lambda(z) = h(\phi^{-1}(z)) \;\; \text{where} \;\; \phi \in \Gamma(2) \text{ and } z \in \phi(Q).$$

Combining Theorem 4.10 with the properties of h shown above yields that $\lambda(\Pi^+) = \lambda(Q) = \mathbb{C} \setminus \{0, 1\}$, $\lambda \circ \phi = \phi$ for all $\phi \in \Gamma(2)$ and that λ is one-to-one. This concludes the proof of part (i) and (ii) for the conformal map case. (So far we have only proved for the conformal map case and we need to show that λ is a covering map)

Part (iii), if $\text{Im}(z) > 1$, then there exists $k \in \mathbb{Z}$ such that $\tau^k(z) \in Q$. Let (z_n) be sequences in Q such that $\tilde{z}_n = \tau^k z_n$ and $\tilde{z}_n \in Q$ where $\text{Im}(z_n) > 1$. Thus, this leads to $\lambda(z_n) = \lambda(\tilde{z}_n) \to \infty$ since $\text{Im}(\tilde{z}_n) \to \infty$.

Primarily, we need to show that λ is a holomorphic covering map. To begin with, let us show that λ is holomorphic in $(Q \cup \tau^{-1}(Q))^{\circ}$. Suppose that $-1-x+iy \in \tau^{-1}(Q)$ and $-1-x > -3$. Since $\tau(-1-x+iy) = 1-x+iy \in Q$, we see that

$$\lambda(-1-x+iy) = \lambda(1-x+iy) = h(1-x+iy) = \overline{h(-1+x+iy)}.$$

Once again, Theorem 3.17 implies that λ is holomorphic in $(Q \cup \tau^{-1}(Q))^{\circ}$ at any point in z with $\text{Im}(z) = -1$. Similarly, λ is also holomorphic in $(Q \cup \sigma^{-1}(Q))^{\circ}$, since $\Gamma(2)$ is freely generated by τ and σ according to Lemma 4.1.

Notice that, $\Gamma(2)$ is a normal subgroup of Γ, and there is a function $\lambda \circ j$ that is invariant under $\Gamma(2)$. If $\phi \in \Gamma(2)$, then there exists $\psi \in \Gamma(2)$ such that $j \circ \phi = \psi \circ j$,

$$\lambda \circ j \circ \phi = \lambda \circ \psi \circ j = \lambda \circ j.$$

But we already know that $\lambda \circ j$ is invariant in Q, since both sides are invariant under $\Gamma(2)$. It follows that $\lambda \circ j = 1/\lambda$ in all of Π^{+}. Thus,

$$j(Q \cup \tau^{-1}(Q)) = j(Q) \cup j(\tau^{-1}(Q)) = j(Q) \cup \sigma^{-1}(j(Q)) = Q \cup \sigma^{-1}(Q)$$

shows that $\lambda \circ j$ and λ are holomorphic in $(Q \cup \sigma^{-1}(Q))^{\circ}$, since $\lambda \circ j = 1/\lambda$. In addition, λ is holomorphic in $\Omega = (Q \cup \tau^{-1}(Q))^{\circ} \cup (Q \cup \sigma^{-1}(Q))^{\circ}$ and with the fact that λ is invariant under $\Gamma(2)$ shows the holomorphism of λ in $\bigcup_{\phi \in \Gamma(2)} \phi(\Omega)$.

Let us define

$$Q \setminus T^{\circ} = \{-1+iy : y > 0\},$$

this gives that $Q \setminus T^{\circ} \subset (Q \cup \tau^{-1}(Q))^{\circ}$. Again, if we apply $j(z) = -1/z$, then $Q \setminus S^{\circ} \subset (Q \cup \sigma^{-1}Q)^{\circ}$ and $Q^{\circ} = T^{\circ} \cap S^{\circ}$, this shows that

$$Q \setminus Q^{\circ} = (Q \setminus T^{\circ}) \cup (Q \setminus S^{\circ}) \subset (Q \cup \tau^{-1}(Q))^{\circ} \cup (Q \cup \sigma^{-1}(Q))^{\circ} = \Omega.$$

Since we have $Q^{\circ} \subset \Omega$, it follows that $Q \subset \Omega$, which would give $\Pi^{+} = \bigcup_{\phi \in \Gamma(2)} \phi(Q) \subset \bigcup_{\phi \in \Gamma(2)} \phi(\Omega)$, so that λ is holomorpic in Π^{+}.

Now, we define an open set N such that $N = \lambda(Q^{\circ})$ with λ holomorphic in $\bigcup_{\phi \in \Gamma(2)} \phi(\Omega)$, then

$$\lambda^{-1}(N) = \bigcup_{\phi \in \Gamma(2)} \phi(Q^{\circ}). \tag{5.1}$$

This indicates that λ^{-1} is the union collection of disjoint open sets $\mathcal{O}_{\phi}$, such that λ restricted to any $\mathcal{O}_{\phi}$ and homeomorphic onto N. Conclusively, N has an elementary neighborhood of each of its points. We have handled most points of $\mathbb{C} \setminus \{0,1\}$, and the next step here is to show that every points of $\lambda(Q) = \mathbb{C} \setminus \{0,1\}$ has an elementary neighborhood.

Suppose there is a $g \in \Gamma$ with $Q^{'} = g(Q)$ and $\phi_0 \in \Gamma(2)$ is not the identity, then there exists $\psi \in \Gamma(2)$ such that $\phi_0 \circ g = g \circ \psi$. It holds that

$$\phi_0(Q') = \phi_0(g(Q)) = g(\psi(Q)).$$

Since g is a bijection and $Q \cap \psi(Q) = \varnothing$, it follows that $Q^{'} \cap \phi_0(Q^{'}) = \varnothing$. Hence, $\phi(Q^{'})$ and $\phi \in \Gamma(2)$ are pairwise disjoint and together with (5.1), it shows that $N^{'} = \lambda((Q^{'})^{\circ})$, thus, N' is elementary for λ.

Let define functions in Γ, such that $g_1 \in \Gamma$ by $g_1(z) = z-1$ and $Q_1 = g(Q)$. Also define: $g_2 = j \circ g_1$ and $Q_2 = g_2(Q)$. These have an elementary neighborhood in $\lambda(Q_2^{\circ})$, and similarly $\lambda(Q_1^\circ)$. As stated earlier: $Q = Q^\circ \cup (Q \setminus T^\circ) \cup (Q \setminus S^\circ)$, and thus we now know that $Q \setminus T^\circ \subset Q_1^\circ$ and $Q \setminus S^\circ \subset Q_2^\circ$. Finally, $Q \subset Q^\circ \cup Q_1^\circ \cup Q_2^\circ$ and every points of $\lambda(Q) = \mathbb{C} \setminus \{0, 1\}$ has an elementary neighborhood, hence λ is a covering map. □

Let us now carry on with the covering maps, and at the end of this section we will show that there is a covering map generated by $\phi \in \Gamma(2)$ with sequences tending to $L \in \{0, 1, \infty\}$.

Theorem 5.6. *Let X and Y be topological spaces such that $p : X \to Y$ is a covering map, D is a simply connected, path-connected and locally path-connected topological space, and $f : D \to Y$ is continuous. Suppose that $a \in D$. Fix $x_0 \in X$ with $p(x_0) = f(a)$. There exists a unique continuous function $\tilde{f} : D \to X$ such that $\tilde{f}(a) = x_0$ and $p \circ \tilde{f} = f$.*

Proof. Suppose that $b \in D$, and let $\gamma : [0, 1] \to D$ be a continuous function with $\gamma(0) = a$ and $\gamma(1) = b$. Also, let $\tilde{\gamma}$ be a lifting of $f \circ \gamma$ with $\tilde{\gamma}(0) = x_0$ and define $\tilde{f}(b) = \tilde{\gamma}(1)$. Since D is connected, then there is a unique path-homotopy of γ_0 and γ_1 such that f is well-defined. □

This is a light version of the proof, since it is straightforward by our assumptions of topological simply connected and path-homotopic, indicating that there is a lift between $X \to Y$. A more rigorous proof of Theorem 5.6 can be seen e.g. in [13]. Statement following from this is:

Corollary 5.7. *Let X and Y be domains such that $p : X \to Y$ is a covering map and X is simply connected. If $f : X \to X$ is continuous, $p \circ f = p$, and f has a fixed point in X, then $f(x) = x$ for all $x \in X$.*

Proof. Let $a \in X$ with $f(a) = a$ and define $g : X \to Y$ by $g = p$. Theorem 5.6 shows that there is a unique continuous function $\tilde{g} : X \to X$ such that $\tilde{g}(a) = a$ and $p \circ \tilde{g} = g$. These functions satisfy $\tilde{g} = f$ and $\tilde{g}(x) = x$, thus, they must be the same function. □

Our next concern is the *covering group*. Here, we want to apply group actions onto covering map and will be starting off by showing the meaning of the covering group, follows by a few usages and a clear connection between automorphism group and covering map.

Remark 5.8. We will call $\mathrm{Aut}(X, p)$ a covering group of p. The elements of $\mathrm{Aut}(X, p)$ are sometimes called *deck transformations* or *covering transformations*.

Theorem 5.9. *Suppose that $p : X \to Y$ is a covering map and X is simply connected. There exists a group $\mathrm{Aut}(X, p)$ of homeomorphisms of X such that for $x_0, x_1 \in X$ we have $p(x_0) = p(x_1)$ if, and only if, $x_1 = g(x_0)$ for some $g \in \mathrm{Aut}(X, p)$.*

Proof. Suppose that $x_0, x_1 \in X$ and $p(x_0) = p(x_1)$. Now, Theorem 5.6 shows that there exists a continuous $g : X \to X$ such that $g(x_0) = x_1$ and $p \circ g = p$. In the same way, this can be said that there exists a continuous $g' : X \to X$ such that $g'(x_1) = x_0$ and $p \circ g' = p$, which results $g \circ g'(x_1) = x_1$.

From Corollary 5.7, we saw that $g \circ g'$ is the identity, and similarly, $g' \circ g$ is the identity. Hence, g is a homeomorphism of X onto X.

Now, let $\mathrm{Aut}(X, p)$ be the set of all homeomorphisms g. Since, g is homeomorphic and by its definition there exists an inverse between the topological spaces. Thus, there exists an inverse in $\mathrm{Aut}(X, p)$. Lastly, Theorem 5.6 implies that $\mathrm{Aut}(X, p)$ is closed. □

We will rephrase Corollary 5.7 that is more suitable to our usage.

Corollary 5.10. *Suppose that $p : X \to Y$ is a covering map and X is simply connected. If $g \in \mathrm{Aut}(X, p)$ has a fixed point in X then g is the identity.*

In the next few theorems we will construct an automorphism covering map that is a covering map with group actions defined as in Subsection 3.1. Purposely, to unite covering map with the group actions of automorphism.

Theorem 5.11. *Let X and Y be topological spaces, let X be a simply connected topological space and $p_0, p_1 : X \to Y$ be covering maps. If $\mathrm{Aut}(X, p_0) = \mathrm{Aut}(X, p_1)$, then there exists a homeomorphism χ of Y such that*

$$p_1 = \chi \circ p_0.$$

Proof. Define $\chi : Y \to Y$ by $\chi(y) = p_1(x)$, where $x \in X$ satisfies $p_0(x) = y$. This is well-defined if $p_0(x) = p_0(x') = y$, then $x' = g(x)$ for some $g \in \mathrm{Aut}(X, p_0) = \mathrm{Aut}(X, p_1)$ as in Theorem 5.6. This implies that p_0, p_1 and their inverses are continuous. The continuity of p_0 and p_1 implies that χ is at least locally continuous on Y, then there must exist $\chi : Y \to Y$ such that $p_1 = \chi \circ p_0$. Similarly, $\chi' : Y \to Y$ such that $p_0 = \chi' \circ p_1$. This shows that $\chi' \circ \chi$ is the identity, thus χ is homeomorphism. □

Theorem 5.11 implies the following corollary.

Corollary 5.12. *Let $\mathbb{D}' = \mathbb{D}\backslash\{0\}$ and suppose that $f \in H(\Pi^+)$. There exists $F \in H(\mathbb{D}')$ with $f(z) = F(e^{2\pi i z})$ for all $z \in \Pi^+$ if, and only if, $f(z + 1) = f(z)$ for all $z \in \Pi^+$.*

Proof. This is the following result of Theorem 5.11, with the fact that Π^+ and $\mathbb{D}'$ are connected open subsets of the plane, such that $p : \Pi^+ \to \mathbb{D}'$ is a holomorphic covering map. Suppose that $f \in H(\Pi^+)$, there exists $F \in H(\mathbb{D}')$ such that

$$f = F \circ p,$$

if, and only if,

$$f \circ \phi = f,$$

for all $\phi \in \mathrm{Aut}(\Pi^+, \mathrm{p})$.

Define $f(z) = F(e^{2\pi i z})$ for all $z \in \Pi^+$. This shows that $F \in H(\mathbb{D}')$ if, and only if, $f(z + 1) = f(z)$ for all $z \in \Pi^+$. □

Theorem 5.13. *Suppose that Ω and V are connected open subsets of the complex plane, and $p : \Omega \to V$ is a holomorphic covering map. Suppose that D is a simply connected open subset of the complex plane, and $f : D \to V$ is holomorphic. Suppose that $a \in D$. Fix $z \in \Omega$ with $p(z) = f(a)$. Then, there exists a unique holomorphic function $\tilde{f} : D \to \Omega$ such that $\tilde{f}(a) = z$ and $p \circ \tilde{f} = f$.*

Proof. Theorem 5.6 shows that there exists a unique continuous function $\tilde{f} : D \to \Omega$ such that $\tilde{f}(a) = z$ and $p \circ \tilde{f} = f$. We also know that $\tilde{f}$ is holomorphic from the following: for $w \in D$ and $W \subset V$, there exists a holomorphic function $\phi : W \to \Omega$ such that $p \circ \phi(\alpha) = \alpha$ for all $\alpha \in W$ and $\tilde{f}(w) \in \phi(W)$, then

$$\tilde{f} = \phi \circ f$$

in $f^{-1}(W)$ and it is an open set containing w, thus $\tilde{f}$ is holomorphic in a vicinity of w. □

In Section 3, we have introduced different classifications of the non-trivial element of the automorphism group as elliptic, parabolic and hyperbolic. Our goal here is to show that these are covering maps and provide an approach to handle poles and singularities for the holomorphic functions in the considerate domain.

Theorem 5.14. *Suppose that* $\phi \in \mathrm{Aut}(\Pi^+)$*, and let* G *be the subgroup of* $\mathrm{Aut}(\Pi^+)$ *generated by* ϕ*. If* ϕ *is not elliptic, then there exists a covering map* h *mapping* Π^+ *onto a bounded open set in the complex plane, such that* $\mathrm{Aut}(\Pi^+, h) = G$*. In more detail:*

(*i*) *If* ϕ *is the identity, then we can take* h *to be a covering map onto* $\mathbb{D}$ *(in fact* h *is a homeomorphism in this case).*

(*ii*) *If* ϕ *is elliptic of infinite order, then there is no non-constant* $h \in H(\Pi^+)$ *with* $h \circ \phi = h$*. Suppose that* ϕ *is elliptic of finite order, and let* $a \in \Pi^+$ *be the fixed point of* ϕ*. There exists a holomorphic map* $h : \Pi^+ \to \mathbb{D}$ *such that* $h|_{\Pi^+ \setminus \{a\}}$ *is a covering map onto* $\mathbb{D}' = \mathbb{D} \setminus \{0\}$*, with* $\mathrm{Aut}(\Pi^+, h|_{\Pi^+ \setminus \{a\}}) = G$*.*

(*iii*) *If* ϕ *is parabolic, then we can take* h *to be a covering map onto* $\mathbb{D}' = \mathbb{D} \setminus \{0\}$*.*

(*iv*) *If* ϕ *is hyperbolic, then we can take* h *to be a covering map onto a proper annulus* $\{z \in \mathbb{C} : r < |z| < R\}$ *for some* r, R *with* $0 < r < R < \infty$*.*

Proof. Suppose that $\phi(z) = z$ and define $h \in H(\Pi^+)$ by

$$h(z) = \frac{z - i}{z + i}.$$

We know from Section 3, that $h(z)$ is the Cayley transform and can be extended homeomorphically onto $\mathbb{D}$. Theorem 5.5 implies that a homeomorphism is certainly a covering map. Part (*ii*), (*iii*) and (*iv*) follow from Theorem 3.16 and with the previous argument of homeomorphisms concluded the fact, that the maps constructed here are covering maps. □

From Theorem 4.10, we have shown that there is a covering map from Π^+ onto $\mathbb{C} \setminus \{0, 1\}$ and that $\mathrm{Aut}(\Pi^+, \lambda) = \Gamma(2)$. We will use these to show that $\mathrm{Aut}(\mathbb{C} \setminus \{0, 1\})$ is isomorphic to the group of all permutations of the set $\{0, 1, \infty\}$.

Theorem 5.15. *If* $\phi \in \mathrm{Aut}(\mathbb{D}')$ *and* $\mathbb{D}' = \mathbb{D} \setminus \{0\}$*, then* ϕ *is a rotation: There exists* $\alpha \in \mathbb{C}$ *with* $|\alpha| = 1$ *such that*

$$\phi(z) = \alpha z, \quad \text{for} \quad z \in \mathbb{D}.$$

Proof. Theorem 3.16 shows that ϕ is a rotation and a bounded nonconstant in $\mathbb{D}$ with a removable singularity at the origin. Hence, ϕ is holomorphic in $\mathbb{D}$ and one-to-one of $\mathbb{D}'$

such that $\phi(0) = 0$, it follows from the continuity of ϕ and Theorem 2.6 that $|\phi(0)| < 1$. The fact that ϕ is one-to-one in $\mathbb{D}^{'}$ shows that $\phi(0) = 0$. If $0 < |\phi(0)| < 1$, then there exists a value $a \in \mathbb{D}'$ such that $\phi(0) = \phi(a)$. Since the function ϕ is nonconstant and according to Theorem 2.4 there must exist values near the points 0 and a, therefore the function $|\phi(0)|$ has two values, which is violating the fact of being one-on-one function onto $\mathbb{D}'$. Hence, ϕ is a one-to-one from $\mathbb{D}$ to $\mathbb{D}$ and $\phi \in \mathrm{Aut}(\mathbb{D})$. Lastly, Theorem 3.11 shows that ϕ is a rotation. □

Theorem 5.16. *If $\phi \in \mathrm{Aut}(\mathbb{C} \setminus \{0, 1\})$, then ϕ is a linear-fractional transformation mapping the set $\{0, 1, \infty\}$ onto itself.*

Proof. Suppose there is a function $\phi \in \mathrm{Aut}(\mathbb{C} \setminus \{0, 1\})$ and ϕ extends to a holomorphic map from $\mathbb{C}_\infty$ to $\mathbb{C}_\infty$. Thus, the function ϕ is one-to-one and with that fact it must contain essential singularities at $0, 1$ and ∞. We need to show that these singularities are distinct, e.g. there exist $\phi(0) = 0$, $\phi(1) = 1$ and $\phi(\infty) = \infty$.

Our argument above implies that, $\phi(0) \in \{0, 1, \infty\}$, and suppose that it is not the case, then there exists $a \in \mathbb{C} \setminus \{0, 1\}$ such that $\phi(0) = \phi(a)$. Theorem 5.15 with Theorem 2.4 show that ϕ is not one-to-one in $\mathbb{C} \setminus \{0, 1\}$, since the values near 0 and 1 can be obtained twice. Similarly, for $\phi(1) \in \{0, 1, \infty\}$ and $\phi(\infty) \in \{0, 1, \infty\}$, which indicates that these three values, $\phi(0), \phi(1)$ and $\phi(\infty)$ must be distinct. Thus, $\phi : \mathbb{C}_\infty \to \mathbb{C}_\infty$ is one-to-one and Theorem 3.10 shows that the set $\mathrm{Aut}(\mathbb{C}_\infty)$ is the set of all linear-fractional transformation. □

We will now look at a certain behavior of the modular function.

Theorem 5.17. *Suppose $g \in \Gamma$ and $g(w_j)$ is a sequence of points in Π^+ such that*

$$\mathrm{Im}(g(w_j)) \to \infty.$$

Then, for the covering map λ from Π^+, it holds that

$$\lambda(w_j) \to L,$$

for some $L \in \{0, 1, \infty\}$.

Proof. Theorem 5.5 shows that there is a sequence (w_j) such that $\lambda(g(w_j)) \to \infty$ as $\mathrm{Im}(w_j) \to \infty$. Theorem 4.10 and Theorem 5.13 show that there is a function χ in $\mathrm{Aut}(\mathbb{C} \setminus \{0, 1\})$ such that $\lambda \circ g = \chi \circ \lambda$ and χ^{-1} is a continuous map from $\mathbb{C}_\infty$ to $\mathbb{C}_\infty$. It follows from Theorem 5.16 that

$$\lambda(w_j) = \chi^{-1}(\lambda(g(w_j))) \to \chi^{-1}(\infty) \in \{0, 1, \infty\}.$$

□

Lemma 5.18.

(*i*) *Suppose that $\phi \in \Gamma$ is parabolic and let α be the unique fixed point of ϕ in $\overline{\Pi^+_\infty}$. Then $\alpha \in \mathbb{Q} \bigcup \{\infty\}$.*

(*ii*) *If $\alpha \in \mathbb{Q}$, then there exists $g \in \Gamma$ with $g(\infty) = \alpha$.*

Proof. To prove (i) we let ϕ be represented by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathrm{SL}_2(\mathbb{Z}).$$

For the case of $c = 0$, it would result in $\alpha = \infty$, but if we assume that $c \neq 0$, then the function $\phi = z$ is quadratic with roots as

$$z = \frac{p \pm \sqrt{r}}{q}.$$

This shows that ϕ has two solutions/fixed points if $r > 0$, on the other hands, if $r < 0$, then ϕ would have a fixed point in Π^+, therefore, $r = 0$, thus, $\alpha \in \mathbb{Q}$.

For part (ii), note first that if $\alpha = 0$, then $g(\infty) = \alpha$ if $g(z) = -1/z$. Suppose $\alpha \neq 0$, and let $\alpha = p/q$ where $p, q \in \mathbb{Z}$. Now, there is $n, m \in \mathbb{Z}$ such that $np + mq = 1$ contributing $g(z) = (pz - m)/(qz + n)$. Thus, $g \in \Gamma$ and $g(\infty) = \alpha$. □

At last, before we continue to the final section we will show an important case to the proof of the great Picard theorem that there is a convergence in the punctured disk, when we have a parabolic function with singularities at $\{0, 1, \infty\}$.

Theorem 5.19. *Suppose that $\phi \in \Gamma(2)$ is parabolic; let $h : \Pi^+ \to \mathbb{D}' = \mathbb{D}\backslash\{0\}$ be a covering map such that $\mathrm{Aut}(\Pi^+, h)$ is generated by ϕ, as in Theorem 5.14. If sequence $(z_n) \subset \Pi^+$ and $h(z_n) \to 0$, then there exists $L \in \{0, 1, \infty\}$ such that the covering map λ from Π^+, so that $\lambda(z_n) \to L$.*

Proof. Firstly, by letting $\alpha \in \partial_\infty \Pi^+$ be the unique fixed point of ϕ in $\overline{\Pi^+_\infty}$, Lemma 5.18 shows that $\alpha \in \mathbb{Q} \cup \{\infty\}$. Then, there exists $g \in \Gamma$ with $g(\infty) = \alpha$.

Let $h \circ g : \Pi^+ \to \mathbb{D}'$ be a covering map, and let $\psi = g^{-1} \circ \phi \circ g \in \mathrm{Aut}(\Pi^+, h \circ g)$. As previously, it shows that ∞ is a fixed point of ψ in $\overline{\Pi^+_\infty}$. Since ϕ is parabolic and from Theorem 3.15 there is an element $a \in \mathbb{R}$ with $a \neq 0$ such that ϕ is conjugated to ψ, with a translation: $\psi(z) = z + a$. Define a covering map $H : \Pi^+ \to \mathbb{D}'$ and let

$$H(z) = e^{2\pi i z/|a|},$$

and let $\mathrm{Aut}(\Pi^+, H)$ be generated by ψ. Herein, Theorem 5.13 shows that there exists a holomorphic function $\chi \in \mathrm{Aut}(\mathbb{D}')$, such that $h \circ g = \chi \circ H$. Theorem 5.15 implies that the function χ is a rotation; This is true if there is β such that $|\beta| = 1$ in a way that

$$h \circ g(z) = \beta e^{2\pi i z/|a|}.$$

However, $h \circ g(g^{-1}(z_n)) \to 0$ and causing $\mathrm{Im}(g^{-1}(z_n)) \to \infty$. Thus, Theorem 5.17 shows that $\lambda(z_n) \to L$ for some $L \in \{0, 1, \infty\}$. □

6. The Great Picard Theorem

In Section 3, we constructed an extended plane and demonstrated that holomorphic functions could be extended homeomorphically from a closed disk $\overline{D}$ onto the closed unit disk $\overline{\mathbb{D}}$. In Section 4, we defined the modular function Γ, and showed that its normal subgroup $\Gamma(2)$ was generated by elements σ and τ. Besides, we had shown that the nontrivial subgroup element of $\Gamma(2)$ was either parabolic or hyperbolic. In Section 5, we introduced the holomorphic covering maps and showed that there existed a holomorphic covering map, h, and that $\text{Aut}(\Pi^+, h)$ is generated by the linear-fractional transformation, ϕ. Moreover, we combined the group actions of automorphism with the covering maps, and we demonstrated that there existed a sequence in Π^+, such that $\lambda(w_j) \to L \in \{0, 1, \infty\}$ where λ is the holomorphic covering map of Π^+ and $w_j \subset \Pi^+$.

At this point, we have almost all the tools needed to prove the great Picard theorem. We will tie most of the results that we have previously mentioned into Theorem 6.1.

Theorem 6.1. *Suppose that $F : \Pi^+ \to \mathbb{C} \setminus \{0, 1\}$ is holomorphic and satisfies $F(z+1) = F(z)$ for all $z \in \Pi^+$. Then there exists $L \in \mathbb{C}_\infty$ such that $F(z_j) \to L$ for every sequence (z_j) in Π^+ with $\text{Im}(z_j) \to \infty$.*

Proof. The upper half-plane in the extended plane, Π^+ is connected with $\lambda : \Pi^+ \to \mathbb{C} \setminus \{0, 1\}$ as a covering map, then by Theorem 5.11 there exists a unique holomorphic function $\tilde{F} : \Pi^+ \to \Pi^+$ such that

$$F = \lambda \circ \tilde{F},$$

and

$$\lambda(\tilde{F}(i + 1)) = F(i + 1) = F(i) = \lambda(\tilde{F}(i)).$$

This, together with Theorem 4.10 imply that there exist $\phi \in \Gamma(2) = \text{Aut}(\Pi^+, \lambda)$ such that,

$$\tilde{F}(i + 1) = \phi(\tilde{F}(i)),$$

and define $\tau_1(z) = z + 1$, thus

$$\lambda \circ (\tilde{F} \circ \tau_1) = F \circ \tau_1 = F = \lambda \circ \tilde{F} = \lambda \circ (\phi \circ \tilde{F}),$$

so that

$$\tilde{F} \circ \tau_1 = \phi \circ \tilde{F}.$$

This satisfies the uniqueness of the holomorphic function $\tilde{F}$.

Since $\phi \in \Gamma(2)$, then Theorem 4.10 shows that ϕ cannot be elliptic, hence, it must either be parabolic or hyperbolic. Therefore, Theorem 5.14 implies that there exists a covering map h from Π^+ onto a bounded open set in the plane such that $\text{Aut}(\Pi^+, h) = G$, where G is the subgroup of $\text{Aut}(\Pi^+)$ generated by ϕ.

Now let $\tilde{H} = h \circ \tilde{F}$. This shows that $\tilde{H}$ is bounded since h is bounded, and that

$$\tilde{H} \circ \tau_1 = h \circ \tilde{F} \circ \tau_1 = h \circ \phi\tilde{F} = h \circ \tilde{F} = \tilde{H}.$$

Recall from Corollary 5.12 for a disk $\mathbb{D}' = \mathbb{D} \setminus \{0\}$ and $f \in \tilde{H}(\Pi^+)$, there exists a function $\tilde{H}^* \in \tilde{H}(\mathbb{D}')$ with $\tilde{H}(z) = \tilde{H}^*(e^{2\pi i z})$. Since $\tilde{H}$ is bounded, then it must have a removable singularity at origin and Theorem 5.17 implies that there is $L = \tilde{H}^*(0)$. If there is sequence $z_j \in \Pi^+$ with $\text{Im}(z_j) \to \infty$ and $w_j = \tilde{F}(z_j)$. Hence, we need to show that $(\lambda(w_j))$ has a limit in $\mathbb{C}_\infty$,

$$h(w_j) = h(\tilde{F}(z_j)) = \tilde{H}(z_j) = \tilde{H}^*(e^{2\pi i z_j}) \to L.$$

Let $L \in h(\Pi^+)$, since h is covering map onto $h(\Pi^+)$, it follows from Lemma 5.18 that there exists a sequence $(w_j^{'})$ in Π^+ such that $h(w_j^{'}) = h(w_j)$ and $w_j^{'} \to w \in \Pi^+$. We know that $\text{Aut}(\Pi^+, h)$ is a subgroup of $\text{Aut}(\Pi^+)$ generated by ϕ, and $\phi \in \text{Aut}(\Pi^+, \lambda)$. Hence,

$$\lambda(w_j) = \lambda(w_j^{'}) \to \lambda(w).$$

This indicates that there exists a limit for $\lambda(w_j)$ as $\text{Im}(z_j) \to \infty$. In other words, there is a function $f(0) = \lambda(w)$ and $f(0) \in \mathbb{C} \setminus \{0, 1\}$. Particularly, the sequence $(\lambda(w_j))$ converged in $\mathbb{C}_\infty$, which is what we needed to prove.

Finally, all that is left is to show that $L \in h(\Pi^+)$ and it can be done by using Theorem 2.6 and Theorem 5.19. Furthermore, we have mentioned before that ϕ can either be parabolic or hyperbolic, and here we will prove the two cases separately.

First, if ϕ is the identity, then

$$\tilde{H}^*(\mathbb{D}') = \tilde{H}(\Pi^+) \subset h(\Pi^+) = \mathbb{D},$$

since $L = \tilde{H}^*(0)$. Theorem 2.6 shows that $|L| < 1$ and $L \in h(\Pi^+)$. Similarly, if ϕ is hyperbolic, then h is a covering map onto an annulus defined by $r < |z| < R$ for some $0 < r < R < \infty$. If we apply Theorem 2.6 to $\tilde{H}^*$ and $1/\tilde{H}^*$, then this turn out that $r < |L| < R$ and so $L \in h(\Pi^+)$.

Now that we have shown for the hyperbolic case: suppose that ϕ is parabolic and $h(\Pi^+)$ is the punctured disk $\mathbb{D}' = \mathbb{D} \setminus \{0\}$ and similar as in the previous paragraph, Theorem 2.6 shows that $|L| < 1$, hence $0 < |L| < 1$.

The remaining of the parabolic case is when we have a parabolic ϕ and $L = 0$. However, we have already seen this in Theorem 5.19, where we show that the sequence $(\lambda(w_j))$ converges in $\mathbb{C}_\infty$. □

At this point, one might has realised that one of the the holomorphic functions satisfying $F(z + 1) = f(z)$ for all $z \in \Pi^+$ is a periodic function. The next few theorems will be dedicated to such function.

Theorem 6.2. *If $u : \mathbb{C} \to \mathbb{R}$ is a bounded harmonic function, then u is constant.*

Proof. Since $\mathbb{C}$ is simply connected, it follows from Lemma 2.12 that there is a real-valued harmonic function v, such that $u + iv$ satisfies the Cauchy-Riemann equation. Thus, the harmonic function is holomorphic. If we let $h = e^{u+iv}$, then h is an entire function and bounded, such that $|h| = e^u$, which implies that h is constant. Thus, u is constant. □

We are now going to elaborate the boundness of harmonic function based on Theorem 6.2, by showing that the function can be harmonically extended and still bounded in the domain.

Theorem 6.3. *If $f : \mathbb{D}' \to \mathbb{R}$ is harmonic and bounded, then $u : \mathbb{C} \to \mathbb{R}$ extends to a function harmonic in $\mathbb{D}$.*

Proof. We will start by defining $U : \Pi^+ \to \mathbb{R}$ by $U(z) = u(e^{2\pi iz})$, then U is harmonic. Since Π^+ is simply connected, there exists a function $F \in H(\Pi^+)$ such that $U = \mathrm{Re}(F)$. Hence, $U(z+1) = U(z)$ but this is not necessary to have $F(z+1) = F(z)$. Since $\mathrm{Re}(F(z+1) - F(z)) = U(z+1) - U(z) = 0$, which indicates that there must exists an imaginary constant such that $F(z+1) - F(z) = ic$ for $c \in \mathbb{R}$ and for all $z \in \Pi^+$. By choosing a real number $\alpha \neq 0$ so that $c\alpha$ is a multiple of 2π, and let

$$E(z) = e^{\alpha F(z)},$$

then the following holds: $E(z+1) = E(z)$, and that there is $f \in H(\mathbb{D}')$ so that

$$E(z) = f(e^{2\pi iz}).$$

Since u is bounded and $U = \mathrm{Re}(F)$, thus, F is bounded. This indicates that f has a removable singularity at origin. Hence, we have previously proved that $f(0) \neq 0$, which shows that u has a limit at 0, since $u = \log(|f|)/\alpha$. Thus, u extended to $\mathbb{D}$. From our previous steps $f(e^{2\pi iz}) = e^{\alpha F(z)}$ then,

$$\log(|f(e^{2\pi iz})|) = \alpha\,\mathrm{Re}(F(z)) = \alpha U(z) = \alpha u(e^{2\pi iz}).$$

□

Combining results from Theorem 6.2 and Theorem 6.3 the statement below is clear.

Theorem 6.4. *If $f \in H(\mathbb{D}')$ and $f(\mathbb{D}') \subset \mathbb{C} \setminus \{0,1\}$, then f has a pole or a removable singularity at* 0*.*

Proof. The statement follows from Theorem 6.1 by replacing the upper half-plane Π^+ with $\mathbb{D}'$ and together with the result of Theorem 6.3, which shows that f has a pole or a removable singularity at 0 and that f is bounded.

□

Finally, from Theorem 6.4 one can see that it is conceivable to show that a function with an essential singularity is boundless in the vicinity of a point. Herein, we are able to prove the great Picard theorem and it is stated as the following.

Theorem 6.5 (The Great Picard Theorem)**.** *If f has an essential singularity at z_0, then with at most one exception, f attains every complex value infinitely many times in every neighborhood of z_0.*

Proof. Based on Theorem 6.4 and a few changed of variables: if f has an essential singularity at z_0, there exists values α and β that can be attained infinitely many times by f in a vicinity of z_0. By considering, that f lies in $D(z_0, r)$ for some sufficiently small $r > 0$ so that f would never attain values α and β. If we let $f_1 = (f - \alpha)/(\beta - \alpha)$, this shows that f_1 never attains values 0 and 1 in $D(0, r)$. If now we let $f_2(z) = f_1(z_0 + rz)$, then $f_2 \in H(\mathbb{D}')$ and $f_2(\mathbb{D}') \subset \mathbb{C} \setminus \{0,1\}$. According to Theorem 6.4, this shows that there exists a pole or a removable singularity at the origin in f_2. Hence, it is the same to f. This finalises the proof of the great Picard Theorem. □

One of the applications from the great Picard theorem is presented in Corrollary 6.6.

Corollary 6.6. *If f is an entire function and f is not a polynomial, then with at most one exception, f attains every complex value infinitely many times.*

Proof. This is an immediate result from Theorem 6.5. If f is an entire function and not a polynomial, by applying the power series on f, the function $f(1/z)$ becomes the Laurent series. Theorem 6.3 has shown that there is a periodic function in the upper half-plane that can be replaced by the automorphisms of the punctured disk $\mathbb{D}' = \mathbb{D} \setminus \{0\}$. Thus, the function f is entire. Theorem 6.5 shows that f can attains complex values infinitely many times. □

7. References

[1] C. Carathéodory, *Sur quelques Applications de Théorème de Landau-Picard*, Mathematische Annalen **144** (1907), no. 3, 1203–1206.

[2] ———, *Theory of Functions of a Complex Variable*, 2nd ed., translated by F. Steinhardt, Vol. 1, AMS Chelsea publishing, 1958.

[3] J. B. Conway, *Functions of One Complex Variable*, 2nd ed. (F. W. Gehring, P. R. Halmos, and C. C. Moore, eds.), Vol. 11, Springer-Verlag, 1978.

[4] R. E. Greene and S. G. Krantz, *Function Theory of One Complex Variable*, Vol. 40, American Mathematical Society, 2006.

[5] J. Hadamard, *Emile Picard, 1856-1941*, The Royal Society **4** (1942), no. 11.

[6] B. C. Hall, *Lie Groups, Lie Algebras and Representations: an Elementary Introduction*, 2nd ed., Vol. 222, Springer, 2015.

[7] I. Hargittai, *Fivefold Symmetry*, 2nd ed., World Scientific, 1992.

[8] P. D. L. Harpe, *Topics in Geometric Group Theory*, University of Chicago Press, 2000.

[9] A. Hatcher, *Algebraic Topology*, Cambridge University Press, 2002.

[10] M. Hazewinkel, *Riemann-Schwarz Principle, Encyclopedia of Mathematics*, Springer Science and Business Media B.V. and Kluwer Academic Publishers, 1994.

[11] H. Heinz, *Über Analytische Abbildungen Riemannscher Flächen in Sich*, Comment. Math. Helv. **27** (1953), 1–81.

[12] K. Kodaira, *Complex Manifolds and Deformation of Complex Structures*, Classics in Mathematics, Springer-Verlag Berlin Heidelberg, 2005.

[13] W. S. Massey, *Algebraic Topology: an Introduction*, Vol. 11, Springer-Verlag, 1978.

[14] M. Ohtsuka, *On the Behavior of an Analytic Function About an Isolated Boundary Point*, Nagoya Math. J. **4** (1952), 103–108.

[15] C. E. Picard, *Sur les Fonctions Analytiques Uniformes dans le Voisinage d'un Point Singulier Essentiel*, C.R. Acad. Sci. Paris **89** (1879), 852–854.

[16] ———, *Mémoire sur les Fonctions Entières*, Ann. Ecole Norm. Sup. **9** (1880), 145–166.

[17] R. Remmert, *Theory of Complex Functions*, 3rd ed., translated by R. B. Burckel, Vol. 122, Springer, 1991.

[18] H.L. Royden, *The Picard Theorem for Riemann Surfaces*, Proc. Amer. Math. Soc. **90** (1984), 571–574.

[19] F. Schottky, *Über den Picardschen Satz und die Borelschen Ungleichungen*, Sitzungsberichte der Preussischen Akademie der Wissenschaften zu Berlin (1904), 1244–1263.

[20] D. C. Ullrich, *Complex Made Simple*, Vol. 97, American Mathematical Society, 2008.