Nils Phillip Talgø

Picard's little theorem

Bachelor's thesis in Mathematical sciences
Supervisor: Karl-Mikael Perfekt
Co-supervisor: Athanasios Kouroupis
June 2024

Bachelor's thesis

NTNU
Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Nils Phillip Talgø

Picard's little theorem

Bachelor’s thesis in Mathematical sciences
Supervisor: Karl-Mikael Perfekt
Co-supervisor: Athanasios Kouroupis
June 2024

Norwegian University of Science and Technology
Faculty of Information Technology and Electrical Engineering
Department of Mathematical Sciences

NTNU
Norwegian University of
Science and Technology

Abstract

This thesis gives a proof of Picard's little theorem by using a modular function, analytic continuation and Liouville's theorem. After some necessary preparations, we start the construction of the modular function with a Riemann mapping from a modified vertical strip in the upper half-plane to the unit disk. Then by using some möbius transformations and Schwartz reflection principle, we get a mapping from a set in the upper half-plane to the twice punctured complex plane. The inverse of this mapping is almost everything we need to use Liouville's theorem, except that it won't be continuous because of the reflection of the borders. This is solved by repeating the domain of the constructed mapping by defining a modular function on a möbius group. Then by using analytic continuation on the composition of an inverse of the modular function with an entire function with two lacunary points, we get an entire function by the monodromy theorem, which maps the complex plane to a region contained in the upper half-plane. After using one more möbius transformation that maps the upper half-plane to the unit disk, we can then finally use Liouville's theorem to conclude that the entire function must be constant, hence proving Picard's little theorem.

Sammendrag

Denne oppgaven gir et bevis av Picard's lille teorem ved å bruke en modulær funksjon, analytisk fortsettelse og Liouville's teorem. Etter noen nødvendige forberedelser, starter vi konstruksjonen av den modulære funksjonen med en Riemann avbildning fra en modifisert vertikal stripe i det øvre halv-plan til enhetsdisken. Ved å bruke noen möbius transformasjoner og Schwartz refleksjonsprinsipp, får vi så en avbildning fra en mengde i det øvre halv-plan til det dobbelt punkterte komplekse plan. Inversen av dette er nesten alt vi trenger for å kunne bruke Liouville's teorem, bortsett fra at den ikke vil være kontinuerlig på grunn av refleksjonen av randen. Dette problemet er løst ved å repetere domenet av den konstruerte avbildningen ved å definere en modulær funksjon på en möbius gruppe. Så bruker vi analytisk fortsettelse på komposisjonen av en invers av den modulære funksjonen med en hel funksjon med to lakunære punkter for å få en hel funksjon ved monodromi teoremet som sender det komplekse planet til en region i det øvre halv-plan. Med én möbius transformasjon til som sender det øvre halv-plan til enhetsdisken, kan vi deretter endelig bruke Liouville's teorem for å konkludere at den hele funksjonen må være konstant, som dermed beviser Picard's lille teorem.

Contents

1 Introduction

Picard's theorem has two versions: Picard's great theorem and Picard's little theorem. In this thesis, I will prove Picard's little theorem. It states that all non-constant entire functions attain all values in the complex plane, with at most a single exception. This implies that all equations of the form $f(z) = z_0$, where f is an entire, non-constant function, has at least one solution in $\mathbb{C}$ except for possibly one value of z_0. The theorem extends to functions meromorphic in the entire plane, but then it allows for two exceptions. Picard's little theorem also follows from Picard's great theorem, which is a lot stronger. It states that if a function is holomorphic in a punctured neighborhood of an essential singularity, then it attains all values in the complex plane *infinitely often*, again with at most a single exception. This implies that if f is an entire non-polynomial function, then $f(z) = z_0$ has infinitely many solutions in $\mathbb{C}$, except for possibly one value of z_0. This theorem extends to functions meromorphic in the entire plane except on a set of isolated essential singularities, but then allows for two exceptions, just like the little version. For most of this thesis, I will be following the book *Real and complex analysis* by Walter Rudin [1].

This thesis will show a clear and concise way to prove a composite and complex theorem in abstract mathematics, which could be useful in education in universities. That, in turn, helps with achieving the UN's fourth Sustainable Development Goal: Quality education.

Terminology and notation

- **Lacunary points** are points not in the range of a function. [2] (p. 97-98)
- **Entire functions** are functions that are holomorphic in the entire complex plane. [3] (p. 630)
- **Holomorphic functions** on a region are functions that have a complex derivative everywhere within that region. [4] (p. xv)
- A function is **biholomorphic** if both it and its inverse are holomorphic. [4] (p. 206)
- A **meromorphic function** is holomorphic everywhere within a region, except on a set of isolated singularities that are poles. [5] (p. 138)
- **Analytic functions** are functions that can locally be represented by a convergent Taylor series. [3] (p. 172) A famous result in complex analysis is that all holomorphic functions are analytic. [5] (p. 82-83)
- A **Laurent series** is a Taylor series representation of an analytic function, but where you allow for all integer powers of z. [3] (p. 708-709)
- **The principal part** of a Laurent series are the terms with a negative power of z. [3] (p. 708-709)

- ∞ is used to denote complex infinity at the north pole of the Riemann sphere.
- $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$ denotes the Riemann sphere.
- Π^+ denotes the open upper half-plane. That is, all z such that $\text{Im}(z) > 0$.
- I will use $D(\alpha, r)$ and $C(\alpha, r)$ to denote the open disk and circle centered at α with radius r, respectively.
- ∂U is the boundary of a set U.

1.1 Definitions

Definition 1.1. *For a function $f : X \to Y$, we call X the domain of f, where f is defined for all x in X. We call Y the codomain of f, where for all x in X, $f(x)$ is in Y. The set of all points in Y that gets mapped to from a set $S \subseteq X$ by f, denoted $f(S)$, is called the image of S, and is defined as $f(S) = \{y \in Y : \exists s \in S, f(s) = y\}$. The image of the domain is called the range of f and is always contained in the codomain.*

Definition 1.2. *Let $f : X \to Y$ be a function.*

- *If for all y in $f(X)$, there is at most one x in X so that $f(x) = y$, then f is injective.*
- *If for all y in Y, there is at least one x in X, so that $f(x) = y$, then f is surjective.*
- *If f is both injective and surjective, it is bijective.*

An equivalent definition of surjectiveness, is that the range equals the codomain.

Definition 1.3. *A set is connected if it cannot be written as the union of two open and disjoint sets.*

If a set is both connected and open, there must exist a continuous curve from any point to any other.

Definition 1.4. *A region in the complex plane is a non-empty, open and connected set. [1] (p. 197)*

The next term I will need to define is simply connectedness. There are many possible formulations of this, but the one that gives the best intuitive understanding, is that a region is simply connected if for any closed curve in the region, you can continuously transform it to a point. That is to say, the region has no holes. Here is another equivalent definition:

Definition 1.5. *A region in the complex plane is simply connected if the complement on the Riemann sphere is connected.*

Lastly, a very useful transformation I will need, is möbius transformations. I will come back to its properties and motivations later. For now, I will simply state its definition:

Definition 1.6. *Let $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$, then a möbius transformation is a map of the form*

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

1.2 Open mapping theorem

This theorem is needed for some technical arguments later on:

Theorem 1.1. *If U is a region, and f is a function holomorphic on U, then $f(U)$ is either a region or a point. [1] (p. 214)*

1.3 Cauchy's integral theorem

Theorem 1.2. *Let U be a simply connected region, let f be holomorphic on U and let γ be a closed and simple curve in U. Then*

$$\oint_\gamma f(z)\, dz = 0$$

1.4 Morera's theorem

Theorem 1.3. *Let U be a simply connected region. If*

$$\oint_\gamma f(z)\, dz = 0$$

for all closed, continuous, and simple paths γ in U, then f is holomorphic in U.

This is the converse of Cauchy's integral theorem.

1.5 Residue theorem

Let γ be a closed curve with a positive orientation (counter-clockwise). Let f be holomorphic on and within γ except on a set of isolated singularities within γ. Let $a_1, a_2, \cdots, a_n$ be those singularities of f within γ. Let $W(\gamma, z)$ denote the winding number of γ at z, that is, the number of times γ goes around the point z, and let $\operatorname{Res}(f, z)$ denote the residue of f at z. The residue theorem then states:

$$\oint_\gamma f(z)\, dz = 2\pi i \sum_{k=1}^{n} W(\gamma, a_k) \operatorname{Res}(f, a_k)$$

[4] (p. 76-77)

1.6 The ML-inequality

This is also known as the Estimation lemma, and says that a complex contour integral is bounded by the maximum of the absolute value of the integrand on the contour, multiplied by the length of the contour.

$$\left| \int_\gamma f(z)\, dz \right| \leq \max_{z \in \gamma} |f(z)| * \text{Length}(\gamma)$$

Here we call M the maximum, and L the length, hence the name "ML-inequality".

1.7 Cauchy's integral formula

This formula is very useful in much of complex analysis and gives a way to find the $n'th$ derivative of a function using a closed contour integral. This integral can then be computed using the residue theorem, or be bounded by the ML-inequality to prove a certain derivative is zero. Let $r > 0$ and $\gamma = \{z : |z - z_0| = r\}$. Cauchy's integral formula is then as follows: [4] (p. 47-48)

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \oint_\gamma \frac{f(z)}{(z - z_0)^{n+1}}\, dz$$

Proof. Let γ be a circle centered at w with radius r, and let f be a function holomorphic on and within γ. Then we have that

$$\begin{aligned} \oint_\gamma \frac{f(z)}{z-w}\, dz &= \oint_\gamma \frac{f(z) - f(w) + f(w)}{z - w}\, dz & (1) \\ &= \oint_\gamma \frac{f(z) - f(w)}{z - w}\, dz + f(w) \oint_\gamma \frac{1}{z-w}\, dz & (2) \\ &= \oint_\gamma \frac{f(z) - f(w)}{z - w}\, dz + 2\pi i f(w) & (3) \end{aligned}$$

(2) follows from the linearity of integrals, and to get (3), one can compute the integral of $1/(z - w)$ directly by substituting $z = re^{it} + w$:

$$\oint_\gamma \frac{1}{z-w}\, dz = \int_0^{2\pi} \frac{ire^{it}dt}{(re^{it} + w) - w} = \int_0^{2\pi} idt = 2\pi i$$

Subtracting $2\pi i f(w)$ from both sides of (3) and taking the absolute value, we then get:

$$\begin{aligned} \left| \oint_\gamma \frac{f(z)}{z-w}\, dz - 2\pi i f(w) \right| &= \left| \oint_\gamma \frac{f(z) - f(w)}{z - w}\, dz \right| & (4) \\ &\leq \max_{z \in \gamma} \left| \frac{f(z) - f(w)}{z - w} \right| * 2\pi r & (5) \\ &= \max_{z \in \gamma} |f(z) - f(w)| * 2\pi \to 0 \text{ as } r \to 0 & (6) \end{aligned}$$

We get (5) by the ML-inequality, and since $z \in \gamma$, we know that $|z - w| = r$, which gives us the first part of (6) from (5). The limit in (6) is due to the continuity of f, since f is holomorphic. Because the integral remains unchanged for all $r > 0$, we get that (4) is zero, which gives us cauchy's formula for $n = 0$:

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

Finally, to get the general formula for all $n > 0$, we can differentiate with respect to w:

$$\begin{aligned}
f'(w) &= \frac{1}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - w)^2} \, dz \\
f''(w) &= \frac{2}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - w)^3} \, dz \\
f'''(w) &= \frac{2 * 3}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - w)^4} \, dz \\
&\vdots \\
f^{(n)}(w) &= \frac{n!}{2\pi i} \oint_{\gamma} \frac{f(z)}{(z - w)^{n+1}} \, dz
\end{aligned}$$

For each step, the exponent of $(z - w)$ increases, and we get new incremental factors, which gives us the factorial. You also get a negative sign from the exponent, but then another negative from the chain rule, so the sign never changes. □

1.8 Isolated singularities

Let f be holomorphic in a punctured neighborhood of z_0. If f is not defined at z_0 or not holomorphic there, then z_0 is an isolated singularity of f. Isolated singularities are classified into three types. Let the following be the Laurent series of f centered at z_0:

$$f(z) = \sum_{n=-\infty}^{\infty} a_n (z - z_0)^n$$

Definition 1.7. *If $a_n = 0$ for all negative n, then z_0 is a removable singularity.*

Definition 1.8. *If $a_n = 0$ for all $n < -m < 0$, and $a_{-m} \neq 0$, then z_0 is a pole of order m.*

Definition 1.9. *If $\forall N < 0, \exists n \leq N$, such that $a_n \neq 0$, then z_0 is an essential singularity.*

[5] (p. 102-103)
In other words, if the principle part of the Laurent series of f, centered at z_0,

has an infinite number of terms, then z_0 is an essential singularity of f. This can be thought of as a pole of order ∞. There is also a corresponding classification of zeros:

Definition 1.10. *If $a_n = 0$ for all $n < m$ where $m > 0$ and $a_m \neq 0$, then z_0 is a zero of order m.*

A pole or zero of order 1 is called a simple pole or a simple zero, respectively. As for removable singularities, as the name suggests, can be removed by defining the function as its limit value there, and thereby becoming holomorphic at z_0. This is always possible, because if you have that z_0 is a removable singularity, you can simply define $f(z_0) := a_0$ from its Laurent series centered at z_0. This also means that f is bounded near removable singularities. The converse however, is slightly less trivial:

Theorem 1.4 (Riemann removable singularity theorem)**.** *Let U be a region. If $z_0 \in U \subset \mathbb{C}$, and f is a function holomorphic and bounded on $U \backslash \{z_0\}$, then z_0 is a removable singularity of f. [5] (p. 105)*

Proof. Let f be a bounded and holomorphic function on $U \backslash \{z_0\}$, where U is a region. Then $|f(z)| \leq M, \forall z \in U \backslash \{z_0\}$ for some $M \geq 0$. Let f have a Laurent series as described above. Then for all $n > 0$, for sufficiently small $r > 0$,

$$a_{-n} = \frac{1}{2\pi i} \oint_\gamma (z - z_0)^{n-1} f(z)\, dz$$

where γ is a circle centered at z_0 with radius r. By the ML-inequality, we have that

$$|a_{-n}| \leq \frac{1}{2\pi} * r^{n-1} M * 2\pi r = r^n M \to 0 \text{ as } r \to 0$$

Thus $a_{-n} = 0$ for all $n > 0$. □

A similar, but more powerful and useful result, which will prove to be essential in the final proof of Picard's little theorem, is Liouville's theorem. If f is entire and bounded globally, then f' is zero everywhere.

1.9 Liouville's theorem

Theorem 1.5. *If a function is entire and bounded, then it is constant.*

Proof. Let f be an entire and bounded function, so that $|f(z)| \leq M \geq 0, \forall z \in \mathbb{C}$. By Cauchy's integral formula, we have that

$$f'(z_0) = \frac{1}{2\pi i} \oint_\gamma \frac{f(z)}{(z - z_0)^2}\, dz$$

where γ is a circle centered at $z_0 \in \mathbb{C}$ with radius $R > 0$. We can then use the ML-inequality and get a bound for the absolute value:

$$|f'(z_0)| \leq \frac{1}{2\pi} * \frac{M}{R^2} * 2\pi R = \frac{M}{R} \to 0 \text{ as } R \to \infty$$

Since f is entire, this bound holds for all R, and thus $f'(z_0) = 0$, and since this result holds for all z_0, we get that $f'(z) \equiv 0$, and thus f must be constant. □

1.10 Extending Picard's little theorem

Picard's little theorem is a very strong and rich theorem that says that all entire functions are almost surjective in the complex plane. It can only miss one point. This one exception is like the *Achilles heel* of holomorphic functions, so close to being surjective. If it wasn't for this missing point, we could always guarantee a solution for equations only involving entire functions. There is unfortunately nothing we can do to salvage this. However, we can extend the family of functions Picard's little theorem works with, to meromorphic functions. This allows us to almost always solve equations that also include poles. For example, using the Riemann sphere, we can solve equations like $1/z = 0$ by simply using the well-defined complex infinity ∞. A regrettable consequence with extending to meromorphic functions however, is that we then need to allow for *two* lacunary points. To show this, I will first need a result concerning constant compositions:

Theorem 1.6. *If f and g are holomorphic on a region U, and $f \circ g$ is constant on U, then at least one of f and g must be constant on U.*

I will prove this in the next chapter.

Theorem 1.7. *Meromorphic functions on the complex plane are either constant, or have at most two lacunary points.*

Proof. Let f be a meromorphic function, and let φ be a möbius transformation:

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

Suppose that f omits three points w_1, w_2, w_3. We can then find coefficients for φ so that:

$$\varphi(0) = w_1$$
$$\varphi(1) = w_2$$
$$\varphi(\infty) = w_3$$

Here is one possibility:

$$a = w_3$$
$$b = w_1 * d$$
$$c = 1$$
$$d = \frac{w_2 - w_3}{w_1 - w_2}$$

Let $g = \varphi^{-1} \circ f$. Since f omits w_3, g is entire, and since f also omits w_1 and w_2, g has two lacunary points. By Picard's little theorem, g must therefore be constant, and then by Theorem 1.6, f must be constant. □

So entire functions have at most one lacunary value, and meromorphic functions have at most two, but when we have essential singularities, we still have at most only one, as seen with Picard's great theorem. Functions that are holomorphic everywhere, except at a set of isolated singularities, are therefore always almost surjective with at most two lacunary points. A well known example of an entire function with a lacunary point, is the exponential function $\exp(z) = e^z$. It attains all values in the complex plane except 0. Extending the domain to $\hat{\mathbb{C}}$, $\exp(z)$ also has a lacunary point at ∞, which is also an essential singularity. This means that $\exp(1/z)$ has an essential singularity at 0. In this case, Picard's great theorem tells us that around 0, it attains all values in the complex plane infinitely often. I will here prove a weaker version of this, called the Casorati-Weierstrass theorem.

1.11 Casorati-Weierstrass

Theorem 1.8. *Let U be a region. If f is holomorphic in $U\backslash\{z_0\}$, where z_0 is an essential singularity of f, then the image of any punctured neighborhood of z_0 within U is dense in the complex plane. [4] (p. 86-87)*

In other words, you can get arbitrarily close to any point in $\mathbb{C}$ from around an essential singularity.

Proof. Let f be holomorphic on $U\backslash\{z_0\}$ where z_0 is an essential singularity of f. Assume that there exists a value w that f cannot get close to. More precisely, that for an $\varepsilon > 0$, $\exists w \in f(U\backslash\{z_0\})$, such that $|f(z) - w| > \varepsilon, \forall z \in U\backslash\{z_0\}$.
Let

$$g(z) = \frac{1}{f(z) - w}$$

We see that g must be holomorphic on $U\backslash\{z_0\}$ and that it is bounded by $1/\varepsilon$. By Theorem 1.4, z_0 must therefore be a removable singularity of g. This means that the limit of g at z_0 must exist.
Let

$$\lambda = \lim_{z \to z_0} g(z)$$

If $\lambda = 0$, f has a pole at z_0, if $\lambda \neq 0$, f has a removable singularity at z_0. Either case contradicts that z_0 was an essential singularity of f. Hence the assumption must be false, and such a w cannot exist. □

Another nice way to make use of an essential singularity, is proving Picard's little theorem using Picard's great theorem. The proof makes use of the essential singularity at ∞ of non-polynomial entire functions.

1.12 Picard's little theorem follows from Picard's great theorem

Proof. Let f be an entire function. This means that f must have a Taylor series at 0.

$$f(z) = \sum_{n=0}^{\infty} a_n z^n$$

If this series is finite, f is a polynomial, but if the series is infinite, the principal part of the Laurent series for $f\left(\frac{1}{z}\right)$ has an infinite number of terms:

$$f\left(\frac{1}{z}\right) = \sum_{n=0}^{\infty} \frac{a_n}{z^n}$$

Then by definition, $f\left(\frac{1}{z}\right)$ has an essential singularity at 0, which in turn means that $f(z)$ has an essential singularity at ∞.

- If f is a polynomial, then Picard's little theorem follows from the fundamental theorem of algebra.
- If f is not a polynomial, by Picard's great theorem, the image of any punctured neighborhood of ∞ on $\hat{\mathbb{C}}$ by f, attains all complex values infinitely often, with at most one exception. This proves Picard's little theorem.

□

This also proves that if you have an equation of the form $f(z) = z_0$, where f is an entire, non-polynomial function, it has infinitely many solutions, except possibly for one value of z_0, as stated in the introduction.

2 Analytic continuation

In the proof of Picard's little theorem, in order to be able to use Liouville's theorem, I will need to analytically continue functions to the entire plane. I will need Schwarz reflection principle, the identity theorem, and the monodromy theorem, the last of which, relies on analytic continuation along curves.

2.1 Schwarz reflection principle

Theorem 2.1. *Let L be a segment on the real axis, Ω^+ a region in Π^+, and every $t \in L$ be the center of an open disc D_t, such that $D_t \cap \Pi^+ \subset \Omega^+$ for all t. Let Ω^- be the reflection of Ω^+:*

$$\Omega^- = \{z : \bar{z} \in \Omega^+\}$$

If f is a function holomorphic in Ω^+, and Im$(f(z)) \to 0$, as $z \to t \in L$, then there exists a function F that is holomorphic on $\Omega^+ \cup L \cup \Omega^-$, with $f(z) = F(z)$ when $z \in \Omega^+$, and where F satisfies the following: $F(\bar{z}) = \overline{F(z)}$. [1] (p. 237-238)

2.2 Identity theorem

Definition 2.1. *A limit point z_0 of a set S, which may or may not be in S, is a point where all neighborhoods in S around z_0 also include points from the set.*

All regions must therefore contain uncountably infinitely many points which must all be limit points. This definition becomes useful when considering points in a discrete set, for example the zero-set of a function, that is, the set of all zeros.

Definition 2.2. *If $z_0 \in S$, and z_0 is not a limit point of S, then z_0 is an isolated point of S.*

Theorem 2.2. *If $f \not\equiv 0$ is a function holomorphic on a region U, then all zeros of f in U are isolated.*

Proof. A function with an infinite order zero would have a taylor series identically equal to zero. Thus the only holomorphic function with an infinite order zero, is the zero function. [5] (p. 88) Let $f \not\equiv 0$ be a holomorphic function on a region U, with a zero at z_0 of finite order m. Then $f(z) = (z - z_0)^m g(z)$ where g is a holomorphic function on U that is nonzero at z_0. Since g is continuous, it is also non-zero in a neighborhood of z_0. Thus z_0 has a positive distance to all other zeros $f^{-1}(0)\backslash\{z_0\}$. In other words, z_0 is an isolated point of the zero-set of f in U. □

Theorem 2.3. *Let f and g be functions holomorphic on a region $U \subseteq \mathbb{C}$, and let $\{z_n\} \subset U$ be a sequence with a limit point. If $f(z_n) = g(z_n)$ for all n, then $f = g$ in all of U. [5] (p. 89)*

Proof. Let f and g be holomorphic functions on a region U, and suppose they agree on a subset of U that has a limit point. Then $(f - g)^{-1}(0)$ has a limit point. In other words, $f - g$ has a non-isolated zero, and then by Theorem 2.2, $f - g$ must therefore be the zero function. Thus $f = g$ on U. □

This theorem allows us to do the proof of Theorem 1.6 which stated that if $f \circ g$ is constant, then at least one of f and g is constant:

Proof. Let f and g be holomorphic functions on a region $U \subseteq \mathbb{C}$, and suppose their composition is constant, that $f(g(z)) = z_0, \forall z \in U$. If g is constant, this holds, and we are done. If g is not constant, then there exists two distinct values α and β, such that $g(\alpha) \neq g(\beta)$. Since U is a region and $g(U)$ is not a single point, by the Open mapping theorem, $g(U)$ must also be a region. Therefore there must exist a continuous path in $g(U)$ from $g(\alpha)$ to $g(\beta)$. $g(U)$ must therefore contain a limit point, and since $f(z) = z_0$ for all $z \in g(U)$, by the Identity theorem, $f \equiv z_0$ in all of U. □

2.3 Monodromy theorem

To be able to state the monodromy theorem, I will first need to define analytic continuation along curves.

Definition 2.3. *A function element is an ordered pair* (f, D)*, where* D *is an open disc, and* f *is holomorphic on* D*. [1] (p. 323)*

Definition 2.4. *A direct continuation between two function elements, denoted* $(f_0, D_0) \sim (f_1, D_1)$*, is the relation where* D_0 *and* D_1 *are not disjoint and where* $f_0(z) = f_1(z)$ *on* $D_0 \cap D_1$*. [1] (p. 323)*

Definition 2.5. *A chain* $C = \{D_0, D_1, \ldots, D_n\}$ *is a finite sequence of open disks, where* $D_i \cap D_{i+1} \neq \emptyset$ *for all* $i = 0, 1, \ldots, n-1$*.*

Since all continuous curves in an open set can be covered by a chain, we can now define analytic continuation along a curve:

Definition 2.6. *Let* $z_0, z_1 \in \mathbb{C}$ *be the centers of two open disks* D_0*,* D_n*, let* γ *be a curve from* z_0 *to* z_1*, and let* C *be a chain that covers* γ *with* n *open disks.* $C = \{D_0, \ldots, D_n\}$*. If the function element* (f_0, D_0) *is given, and there exists function elements* (f_i, D_i) *such that* $(f_i, D_i) \sim (f_{i+1}, D_{i+1})$ *for* $i = 0, 1, \ldots, n-1$*, then* (f_n, D_n) *is the analytic continuation of* (f_0, D_0) *along the curve* γ*, and along the chain* C*.*

Theorem 2.4. *If* (f, D) *is a function element, and* γ *is a curve starting at the center of* D*, and ending in a point* $z_1 \in \mathbb{C}$*, then there is at most one analytic continuation along* γ*. [1] (p. 324-325)*

I can now state the monodromy theorem:

Theorem 2.5. *Let* U *be a simply connected region and* (f, D) *a function element with* $D \subset U$*. If* (f, D) *can be analytically continued along all curves in* U *that start at the center of* D*, then there exists a function* g *holomorphic in* U *and where* $g(z) = f(z)$ *within* D*. [1] (p. 326-327)*

The proof of this is outside the scope of the thesis and will not be proven here. A proof can be found in [1].

3 Conformal maps

Definition 3.1. *A mapping is conformal if it preserves angles locally.*

This is equivalent to being holomorphic and having a non-zero derivative everywhere within its domain, if the domain is a region. [1] (p. 278)

3.1 Möbius transformations

Recall, a möbius transformation is a ratio of two linear functions where $a, b, c, d \in \mathbb{C}$ and $ad - bc \neq 0$:

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

This map is also called a *linear fractional transformation.* [1] (p. 279) We see that this mapping has a simple pole at $-d/c$ and a simple zero at $-b/a$. Taking a look at its derivative;

$$\varphi'(z) = \frac{a(cz+d)-(az+b)c}{(cz+d)^2} = \frac{ad-bc}{(cz+d)^2}$$

we see the motivation for the requirement on the constants, which makes φ have a non-zero derivative everywhere, which then means it is conformal everywhere.

3.1.1 Möbius transformations are bijective

From the definition of φ, we can find the inverse directly:

$$\begin{aligned} \frac{az+b}{cz+d} &= w \\ az+b &= w(cz+d) \\ az - wcz &= wd - b \\ z &= \frac{dw-b}{-cw+a} \end{aligned}$$

We see that φ^{-1} is of the same form as φ, where a and d have swapped places, and b and c have swapped signs. Since the inverse is defined everywhere except at a/c, φ is almost surjective with a/c as the only lacunary point. Extending the domain and codomain to $\hat{\mathbb{C}}$ allows us to reach it: $\varphi(\infty) = a/c$. Hence φ is surjective. What remains to be shown, is that it is injective:

$$\begin{aligned} \frac{az_1+b}{cz_1+d} &= \frac{az_2+b}{cz_2+d} \\ (az_1+b)(cz_2+d) &= (az_2+b)(cz_1+d) \\ acz_1z_2 + adz_1 + bcz_2 + bd &= acz_2z_1 + adz_2 + bcz_1 + bd \\ adz_1 + bcz_2 &= adz_2 + bcz_1 \\ adz_1 - bcz_1 &= adz_2 - bcz_2 \\ (ad-bc)z_1 &= (ad-bc)z_2 \\ z_1 &= z_2 \end{aligned}$$

Therefore φ is bijective on $\hat{\mathbb{C}}$, and indeed also biholomorphic.

3.1.2 Fixed points

Another topic of interest, is the fixed points of möbius transformations. Assuming $z \neq -d/c$, we can set $\varphi(z) = z$ and solve for z:

$$\begin{aligned} \frac{az+b}{cz+d} &= z \\ az+b &= z(cz+d) \\ cz^2+(d-a)z-b &= 0 \end{aligned}$$

We can see the number of solutions depend only on the constants a, b, c, d. We can categorize it like this:

1. $c \neq 0$: One or two fixed points: $z = \frac{1}{2c}(a - d \pm \sqrt{(a-d)^2 + 4bc})$
2. $c = 0$:
   (a) $a \neq d$: One fixed point: $z = \frac{b}{d-a}$
   (b) $a = d$:
      i. $b \neq 0$: No fixed points
      ii. $b = 0$: This is the identity map $\varphi(z) \equiv z$.

In conclusion, except for the trivial identity map, möbius transformations can have at most 2 fixed points.

3.1.3 Möbius transformations preserve lines and circles

To show this, I will first show that any möbius transformation can be written as the composition of the following transformations:

- $f_1(z, \beta) = z + \beta$ (Translation)
- $f_2(z) = 1/z$ (Inversion)
- $f_3(z, \alpha) = \alpha z$ (Scaling and rotation)

$$\begin{aligned} \varphi(z) &= \frac{az+b}{cz+d} \\ &= \frac{a}{c}\frac{z+b/a}{z+d/c} \\ &= \frac{a}{c}\frac{z+d/c+b/a-d/c}{z+d/c} \\ &= \frac{a}{c}\left(1+\frac{b/a-d/c}{z+d/c}\right) \\ &= \frac{a}{c}\frac{1}{ac}\frac{bc-ad}{z+d/c}+\frac{a}{c} \\ &= \frac{bc-ad}{c^2}\frac{1}{z+d/c}+\frac{a}{c} \end{aligned}$$

We see that $\varphi(z) = f_1(f_3(f_2(f_1(z, \frac{d}{c})), \frac{bc-ad}{c^2}), \frac{a}{c})$. The only exception to this formula, is if $c = 0$. In this case, it is easy to see that φ has no inversion, and is only a composition of translation with scaling and rotation. What remains to be shown is that the individual transformations preserve lines and circles. In $\mathbb{R}^2$, all points lying on lines or circles satisfy the following quadratic equation:

$$A(x^2 + y^2) + Bx + Cy + D = 0 \tag{1}$$

An equivalent equation for $\mathbb{C}$ can be deduced by substituting $x = \frac{1}{2}(z + \bar{z})$ and $y = \frac{1}{2i}(z - \bar{z})$ and using that $x^2 + y^2 = z\bar{z}$ when $z = x + yi$:

$$\begin{aligned}
Az\bar{z} + B\left(\frac{z+\bar{z}}{2}\right) + C\left(\frac{z-\bar{z}}{2i}\right) + D &= 0 &(2)\\
Az\bar{z} + \frac{B}{2}z + \frac{B}{2}\bar{z} + \frac{C}{2i}z - \frac{C}{2i}\bar{z} + D &= 0 &(3)\\
Az\bar{z} + \left(\frac{Bi+C}{2i}\right)z + \left(\frac{Bi-C}{2i}\right)\bar{z} + D &= 0 &(4)\\
Az\bar{z} + B'z + C'\bar{z} + D &= 0 &(5)
\end{aligned}$$

In (2), I have made the subtitutions from (1). In (3), I have expanded the paranthesis from (2). In (4), I have grouped the terms with z and $\bar{z}$, which we can see are simply some different constants, which we can relabel, like in (5). Also, notice that B' is always the complex conjugate of C'. I will now use the exact same strategy for the transformations:
Translation:

$$\begin{aligned}
A(z+\beta)\overline{(z+\beta)} + B(z+\beta) + C\overline{(z+\beta)} + D &= 0\\
\iff A(z\bar{z} + z\bar{\beta} + \beta\bar{z} + \beta\bar{\beta}) + Bz + B\beta + C\bar{z} + C\bar{\beta} + D &= 0\\
\iff Az\bar{z} + (A\bar{\beta} + B)z + (A\beta + C)\bar{z} + (A\beta\bar{\beta} + B\beta + C\bar{\beta} + D) &= 0
\end{aligned}$$

Scaling and rotation:

$$\begin{aligned}
A(\alpha z)\overline{(\alpha z)} + B(\alpha z) + C\overline{(\alpha z)} + D &= 0\\
\iff (A\alpha\bar{\alpha})z\bar{z} + (B\alpha)z + (C\bar{\alpha})\bar{z} + D &= 0
\end{aligned}$$

For inversion, we know that 0 gets mapped to ∞, and if $z \neq 0$, we can multiply both sides of the equation with $z\bar{z}$:

$$\begin{aligned}
A\frac{1}{z\bar{z}} + B\frac{1}{z} + C\frac{1}{\bar{z}} + D &= 0\\
\iff A + B\bar{z} + Cz + Dz\bar{z} &= 0\\
\iff Dz\bar{z} + Cz + B\bar{z} + A &= 0
\end{aligned}$$

This proves that all the constituent transformations preserve lines and circles, and hence so do all möbius transformations.

Using the equations above, I have also been able to find two corresponding quadratic equations after a möbius transformation by composing the above equations in the order described at the beginning of this section. One for when $c \neq 0$, and one for when $c = 0$. Using those equations, I found a shortcut for finding the center and radius of the image of a curve described by (1). If $c \neq 0$, this will always be a circle, as long as the curve does not go through the pole. Let A, B, C, D be the coefficients of (1), and a, b, c, d be the coefficients of a möbius transformation. Let $z_0 \in \mathbb{C}$ be the center and $r > 0$ be the radius of the image.

Center:

$$z_0 = \frac{A\frac{\bar{d}}{\bar{c}} - \frac{1}{2}(B - Ci)}{\frac{c^2}{bc-ad}(A\left|\frac{d}{c}\right|^2 - B\mathrm{Re}(\frac{d}{c}) - C\mathrm{Im}(\frac{d}{c}) + D)} + \frac{a}{c}$$

Radius:

$$r^2 = |z_0|^2 - \left|\frac{a}{c}\right|^2 + \frac{\mathrm{Re}(\frac{ac}{bc-ad})(B - 2A\mathrm{Re}(\frac{d}{c})) - \mathrm{Im}(\frac{ac}{bc-ad})(C - 2A\mathrm{Im}(\frac{d}{c})) - A}{\left|\frac{c^2}{bc-ad}\right|^2 (A\left|\frac{d}{c}\right|^2 - B\mathrm{Re}(\frac{d}{c}) - C\mathrm{Im}(\frac{d}{c}) + D)}$$

Notice how the center and radius both go to infinity when the point $(\mathrm{Re}(-d/c), \mathrm{Im}(-d/c))$ is in (1). As mentioned, this is because the image is a line when the pole $-d/c$ is on the curve. On the other hand, when $c = 0$, we get ∞/∞. In this case, the formulas you get simplify somewhat:

Center:

$$z_0 = \frac{b}{d} - \frac{\frac{1}{2}(B + Ci)}{A\frac{d}{a}}$$

Radius:

$$r^2 = |z_0|^2 + \frac{B\mathrm{Re}(\frac{b}{a}) + C\mathrm{Im}(\frac{b}{a}) - A\left|\frac{b}{a}\right|^2 - D}{A\left|\frac{d}{a}\right|^2}$$

Notice, similar to before, if $A = 0$, both the radius and center to go infinity. This is because if $c = 0$, φ has no inversion, and the image is therefore a circle whenever the input curve is too. The only singularities here besides $A = 0$, is when $a = 0$ or $d = 0$. But since we have assumed $c = 0$, we know that $d \neq 0$, and if $a = 0$, φ would be a constant.

3.1.4 Useful examples

A very useful example of a möbius transformation is the case when $a = 0, b = r > 0, c = 1$ and $d = -z_0$:

$$\varphi_1(z) = \frac{r}{z - z_0}$$

This is a biholomorphic mapping from the complement of an open disk on the Riemann sphere with radius r and center z_0 to the closed unit disk. To see this, notice that when $z \in \hat{\mathbb{C}} \backslash D(z_0, r)$, then $|z - z_0| \geq r$, and thus $|\varphi_1(z)| \leq 1$.

Another useful example is the mapping from the upper half-plane Π^+, to the unit disk:

$$\varphi_2(z) = \frac{z-i}{z+i}$$

[1] (p. 281)
If $z \in \Pi^+$, then $|z-i| < |z+i|$, and thus $|\varphi_2(z)| < 1$. These examples are so useful, because if you had an entire function that omitted a disk or a half-plane, you could compose one of these transformations with it to get an entire and bounded function, and then by Liouville's theorem conclude that the composition must be constant, which means the original function must be constant.

3.1.5 Möbius transformations can map any triple to any triple

Möbius transformations have 4 free variables, but only 3 degrees of freedom. Let $\varphi(z) = (az+b)/(cz+d)$ be a möbius transformation. Notice that

$$\varphi(z) = \frac{\frac{a}{c}z + \frac{b}{c}}{z + \frac{d}{c}}$$

It therefore makes sense that φ can map any three distinct points to any three distinct points. I will now deduce an explicit formula. Suppose that φ maps the triple (z_1, z_2, z_3) to (w_1, w_2, w_3), where $z_1 \neq z_2 \neq z_3 \neq z_1$ and $w_1 \neq w_2 \neq w_3 \neq w_1$. We then get 3 equations and we need to solve for 4 variables, but where one of them is just a scalar for the other three. From the algebra that follows, it is easiest to let c be the free variable.

$$\begin{aligned}
\varphi(z_1) = w_1 \quad &\implies \quad az_1 + b = w_1(cz_1 + d) \\
&\implies \quad b = w_1(cz_1 + d) - az_1
\end{aligned}$$

$$\begin{aligned}
\varphi(z_2) = w_2 \quad &\implies \quad az_2 + b = w_2(cz_2 + d) \\
&\implies \quad az_2 + (w_1(cz_1 + d) - az_1) = w_2(cz_2 + d) \\
&\implies \quad a(z_2 - z_1) + w_1(cz_1 + d) = w_2(cz_2 + d) \\
&\implies \quad a = \frac{w_2(cz_2 + d) - w_1(cz_1 + d)}{z_2 - z_1}
\end{aligned}$$

For the last equation, $\varphi(z_3) = w_3$, we can substitute the formulas for b and then for a, and then expand the terms with d, to finally solve for d.

$$
\begin{aligned}
&\varphi(z_3) = w_3 \\
\implies & az_3 + b = w_3(cz_3 + d) \\
\implies & az_3 + (w_1(cz_1 + d) - az_1) = w_3(cz_3 + d) \\
\implies & a(z_3 - z_1) + dw_1 - dw_3 = cw_3z_3 - cw_1z_1 \\
\implies & \frac{w_2(cz_2 + d) - w_1(cz_1 + d)}{z_2 - z_1}(z_3 - z_1) + d(w_1 - w_3) = cw_3z_3 - cw_1z_1 \\
\implies & (dw_2 - dw_1 + cw_2z_2 - cw_1z_1)(z_3 - z_1) + d(w_1 - w_3)(z_2 - z_1) \\
& = (cw_3z_3 - cw_1z_1)(z_2 - z_1) \\
\implies & d(w_2 - w_1)(z_3 - z_1) + (cw_2z_2 - cw_1z_1)(z_3 - z_1) \\
& + d(w_1 - w_3)(z_2 - z_1) = (cw_3z_3 - cw_1z_1)(z_2 - z_1) \\
\implies & d = \frac{(cw_3z_3 - cw_1z_1)(z_2 - z_1) - (cw_2z_2 - cw_1z_1)(z_3 - z_1)}{(w_2 - w_1)(z_3 - z_1) + (w_1 - w_3)(z_2 - z_1)}
\end{aligned}
$$

Notice that d has a common factor of c. Hence so does a and b. These formulas work as long as the points are distinct, as described in the beginning.

3.2 Riemann mapping theorem

Biholomorphic mappings like the ones shown in the previous section that map to the unit disk, always exists as long as the set is non-empty, simply connected, open, and is not the whole complex plane. [1] (p.283). This is called the Riemann mapping theorem. The reason for why we can't allow the whole plane, is that the mapping would then be constant, by Liouville's theorem.

Theorem 3.1 (Riemann mapping theorem)**.** *For every non-empty, simply connected, open and proper subset $U \subset \mathbb{C}$, there exists a biholomorphic mapping from U to the open unit disk. [5] (p. 142)*

Definition 3.2 (Simple boundary point)**.** *Let $U \subset \mathbb{C}$ be a region, and $\beta \in \partial U$, then β is a simple boundary point if for every sequence $\{\alpha_n\} \subset U$, where $\alpha_n \to \beta$ as $n \to \infty$, there is a continuous curve $\gamma : [0, 1] \to \mathbb{C}$, where $\gamma(t) \in U$ when $t \in [0, 1)$, and a sequence $\{t_n\} \subset [0, 1)$ where $t_n \to 1$ and $t_n < t_{n+1}$, such that $\gamma(t_n) = \alpha_n$ for all $n \in \mathbb{N}$. [1] (p.289)*

In other words, a boundary point is simple if all sequences converging to it, can be connected by a continuous curve. This also extends to a boundary point at infinity. If you have a sequence in a region that converges to infinity, and a continuous curve within the region connecting them that also converges to infinity, then the boundary point there is simple. An example of a region with non-simple boundary points is an open disk with radius 1, centered at zero, with the segment $[0, 1)$ removed. For points on this segment, you could have a convergent sequence that jumped above and below the segment, so that any continuous curve connecting them, would have to go around 0 infinitely often. Such a curve would not be continuous in the end points, and those boundary points are therefore not simple. Another theorem I will need later on for the proof of Picard is the following:

Theorem 3.2. *If U is a bounded and simply connected region, and if every boundary point of U is simple, then every conformal mapping of U onto the open unit disk D, extends to a bijective and continuous map from $\overline{U}$ onto $\overline{D}$. [1] (p. 290)*

Theorem 3.3. *Let U be an unbounded and simply connected region, where every boundary point of U is simple, including at ∞, and let φ be a möbius transformation. If $\varphi(U)$ is bounded and $\varphi(\infty)$ is a simple boundary point of $\varphi(U)$, then every bijective and conformal mapping of U onto the open unit disk D, extends to a bijective and continuous map from $\overline{U}$ to $\overline{D}$.*

Proof. Let U be an unbounded, simply connected and open region with simple boundary points, also at ∞. Let D be the open unit disk, let φ be a möbius transformation where $\varphi(U)$ is bounded, and let $\varphi(\infty) \in \partial\varphi(U)$ be a simple boundary point. Since every finite boundary point of U is simple, the same is true for $\varphi(U)$. Let f be a bijective and conformal mapping from U to D. Since f is bijective, so is f^{-1}, and $f^{-1}(D) = U$. Then $\varphi \circ f^{-1}$ maps D to $\varphi(U)$, and so $(\varphi \circ f^{-1})^{-1}$ maps $\varphi(U)$ to D. Since $\partial\varphi(U)$ is simple and $\varphi(U)$ is bounded, by theorem 3.2, this extends to a bijective and continuous map g from $\overline{\varphi(U)}$ to $\overline{D}$. Then g^{-1} maps $\overline{D}$ to $\overline{\varphi(U)}$, and $\varphi^{-1} \circ g^{-1}$ maps $\overline{D}$ to $\overline{U}$. Finally, $(\varphi^{-1} \circ g^{-1})^{-1}$ maps $\overline{U}$ to $\overline{D}$. □

4 Construction of a modular function

The final piece I will need for the proof of Picard's little theorem, is a specific modular function λ which will help transform the range of an entire function with two lacunary points to the upper half-plane. Then, by using analytic continuation on the composition, I will end up with an entire function that maps the entire plane to just a region in the upper half-plane. Then by using a möbius transformation, as mentioned in chapter 3.1.4, we get an entire function from the plane to the unit disk, which implies the function is constant, by Liouville's theorem. Almost everything of this chapter is following the book of Rudin [1] (p. 328-332), but with more explanations, intermediary steps and some graphics to help visualize the regions described.

Definition 4.1. *A modular function f is a function that is invariant under möbius transformations from a group G. That is, for all $\varphi \in G$,*

$$f \circ \varphi = f$$

To define λ, we will first take a look at a special möbius group with integer coefficients, and determinant 1.

4.1 The möbius group

This is the group G of all möbius transformations

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

where a, b, c, d are integers, and $ad - bc = 1$. As shown in the previous chapter, all such φ have an inverse:

$$\varphi^{-1}(z) = \frac{dz-b}{-cz+a}$$

We see that φ^{-1} also has integer coefficients, and that $da - (-b)(-c) = 1 \iff ad - bc = 1$. Thus $\forall \varphi \in G, \exists! \varphi^{-1} \in G$. I will now show that G is closed under compositions. Let $\varphi_1, \varphi_2 \in G$.

$$\begin{aligned} \varphi_1 \circ \varphi_2 &= \frac{a_1 \frac{a_2z+b_2}{c_2z+d_2} + b_1}{c_1 \frac{a_2z+b_2}{c_2z+d_2} + d_1} \frac{c_2z+d_2}{c_2z+d_2} \\ &= \frac{a_1(a_2z+b_2) + b_1(c_2z+d_2)}{c_1(a_2z+b_2) + d_1(c_2z+d_2)} \\ &= \frac{(a_1a_2+b_1c_2)z + a_1b_2 + b_1d_2}{(c_1a_2+d_1c_2)z + c_1b_2 + d_1d_2} \end{aligned}$$

The coefficients consists of products and sums of integers, and are therefore also integers. Lastly, we must confirm the determinant is still 1.

$$\begin{aligned} ad - bc &= (a_1a_2+b_1c_2)(c_1b_2+d_1d_2) - (a_1b_2+b_1d_2)(c_1a_2+d_1c_2) \\ &= a_1a_2c_1b_2 + a_1a_2d_1d_2 + b_1c_2c_1b_2 + b_1c_2d_1d_2 \\ &- (a_1b_2c_1a_2 + a_1b_2d_1c_2 + b_1d_2c_1a_2 + b_1d_2d_1c_2) \\ &= a_2d_2(a_1d_1) - a_2d_2(b_1c_1) + b_2c_2(b_1c_1) - b_2c_2(a_1d_1) \\ &= a_2d_2(a_1d_1 - b_1c_1) + b_2c_2(b_1c_1 - a_1d_1) \\ &= a_2d_2 + b_2c_2(-1) \\ &= 1 \end{aligned}$$

Hence $\varphi_1 \in G \wedge \varphi_2 \in G \implies \varphi_1 \circ \varphi_2 \in G$. Thus, G is a group with composition as the group operation and with the identity map as the identity element.

4.2 A subgroup Γ of G

Let Γ be the subgroup of G generated by the following transformations:

$$\tau(z) = z + 2$$
$$\sigma(z) = \frac{z}{2z+1}$$

As proved in the previous section, since $\tau, \sigma \in G$, Γ is a subgroup with composition as the group operation. It is under this group that λ will be defined.

Following [1] (p. 329), I will need an explicit formula for the imaginary part of φ in the upcoming proofs. Though since the coefficients are real, we can find this without much effort:

$$\begin{aligned}\varphi(z) &= \frac{az+b}{cz+d}\frac{c\bar{z}+d}{c\bar{z}+d} \\ &= \frac{acz\bar{z}+adz+bc\bar{z}+bd}{|cz+d|^2}\end{aligned}$$

Identifying the imaginary part:

$$\begin{aligned}\mathrm{Im}(\varphi(z)) &= \mathrm{Im}\left(\frac{adz+bc\bar{z}}{|cz+d|^2}\right) \\ &= \frac{\mathrm{Im}(z)(ad-bc)}{|cz+d|^2}\end{aligned}$$

We now also see the motivation for the requirement that $ad-bc=1$:

$$\mathrm{Im}(\varphi(z)) = \frac{\mathrm{Im}(z)}{|cz+d|^2} \tag{1}$$

Let Q be the set of all $z = x+yi$ satisfying the following inequalites:

$$y>0, \quad -1 \leq x < 1, \quad |2z+1| \geq 1, \quad |2z-1| > 1$$

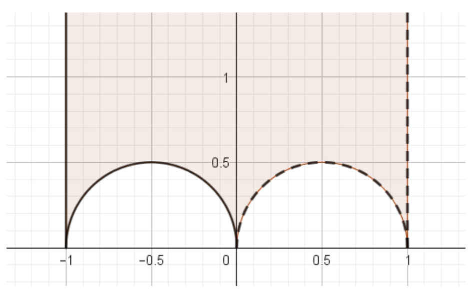


Figure 1: Q

Theorem 4.1. *Let Γ and Q be as above.*

(a) If $\varphi_1, \varphi_2 \in \Gamma$ and $\varphi_1 \neq \varphi_2$, then $\varphi_1(Q) \cap \varphi_2(Q) = \emptyset$

(b) $\bigcup_{\varphi\in\Gamma} \varphi(Q) = \Pi^+$

(c) If $\varphi \in \Gamma$ and $\varphi(z) = (az+b)/(cz+d)$, then $\varphi \in G$ where a, d are odd integers, and b, c are even integers.

[1] (p. 329)

(a) and (b) means that Q is a *fundamental domain* of Γ, while (c) is used to prove (a).

Proof. Let Γ_1 be the group described in (c), and let (a') be the statement (a) for Γ_1. From the previous section, in the proof that G is closed under composition, we see that the coefficients a, d remain odd and b, c remain even after compositions, so Γ_1 is indeed a group. Notice that proving (c) is the same as proving that $\Gamma = \Gamma_1$. For τ, we see that $a = d = 1, b = 2, c = 0$. For σ, $a = d = 1, b = 0, c = 2$. Therefore τ and σ are in Γ_1. This must mean that $\Gamma \subseteq \Gamma_1$, as Γ_1 contains the generators of Γ.

I will now show that (a') $\wedge$ (b) $\implies$ (c). Suppose that this isn't true. That we have (a') and (b), and yet $\Gamma \subset \Gamma_1$. This means that there is a $\psi \in \Gamma_1$ where $\psi \notin \Gamma$. By (b), the images $\varphi(Q)$ where $\varphi \in \Gamma$ cover Π^+. By (1), we see that since $Q \subseteq \Pi^+$, $\psi(Q) \subseteq \Pi^+$. Thus $\psi(Q)$ intersects $\varphi'(Q)$ for some $\varphi' \in \Gamma$. But this contradicts (a'), and thus ψ cannot exist, and $\Gamma = \Gamma_1$. Also, notice that if you have (a') and (c), you automatically also get (a). Therefore, proving (a') and (b) is enough to prove the whole theorem.

To prove (a'), there is one more simplification we can do. Let $\varphi_1, \varphi_2 \in \Gamma_1$, where $\varphi_1 \neq \varphi_2$, and let $\varphi = \varphi_1^{-1} \circ \varphi_2$. (a') then states that $\varphi_1(Q) \cap \varphi_2(Q) = \emptyset$. Suppose that $z \in \varphi_1(Q) \cap \varphi_2(Q)$. Then $\varphi_1^{-1}(z) \in \varphi_1^{-1}(\varphi_1(Q) \cap \varphi_2(Q))$. And since the image of an intersection is a subset of the intersection of the images, we get that $\varphi_1^{-1}(z) \in \varphi_1^{-1}(\varphi_1(Q)) \cap \varphi_1^{-1}(\varphi_2(Q)) = Q \cap \varphi(Q)$. Therefore, proving (a') is equivalent to proving that for all $\varphi \in \Gamma_1$ except the identity map, the following must hold:

$$Q \cap \varphi(Q) = \emptyset \tag{2}$$

To justify the last step, notice that for any two sets A and B and function f, $A \cap B \subseteq A \implies f(A \cap B) \subseteq f(A)$, and similarily, $f(A \cap B) \subseteq f(B)$. Thus $f(A \cap B) \subseteq f(A) \cap f(B)$.

I will now prove (a') by considering three cases. Let $\varphi \in \Gamma_1$ and $\varphi \not\equiv z$.

- If $c = 0$, then $ad - bc = 1 \implies ad = 1 \implies a = d = \pm 1$, since the coefficients are integers. Then $\varphi(z) = z + b = z + 2n, n \in \mathbb{Z}$. Since Q only includes its left border, and since the width of Q is 2, shifting Q by any even integer does not intersect Q. Hence (2) holds.

- If $c = 2d$, then $ad - bc = 1 \implies d(a - 2b) = 1 \implies d = \pm 1 \implies c = \pm 2$. In this case,
$$\varphi(z) = \frac{az + b}{2z + 1} = \frac{(2n + 1)z + (2m)}{2z + 1}$$
and letting $n = 2m$ then gives us
$$\varphi(z) = \frac{4mz + z + 2m}{2z + 1} = \sigma(z) + 2m, m \in \mathbb{Z}$$

Since σ has a pole at $-1/2$, and since $-1/2 \notin \partial Q$, we know that $\partial\sigma(Q)$ only consists of semicircles. Using the explicit formulas found in chapter 3.1.3, or just by finding the image of three points on each of the boundaries of Q to find the corresponding circle, we get that the left vertical border of Q gets mapped to the upper half of $C(\frac{3}{4}, \frac{1}{4})$, the left semicircle gets mapped to the right semicircle which is the upper half of $C(\frac{1}{2}, \frac{1}{2})$, the right semicircle gets mapped to the upper half of $C(\frac{1}{6}, \frac{1}{6})$, and the right vertical border gets mapped to the upper half of $C(\frac{5}{12}, \frac{1}{12})$. Travelling along the border of Q counter-clockwise, the interior is on the left, and travelling along the semicircles mentioned, in the same order, we see that the left side gives the area between the semicircles.

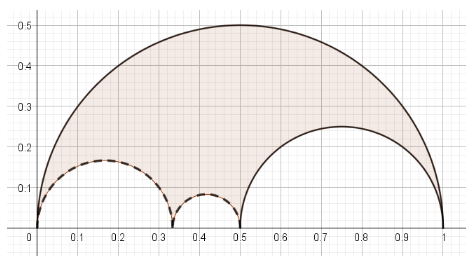
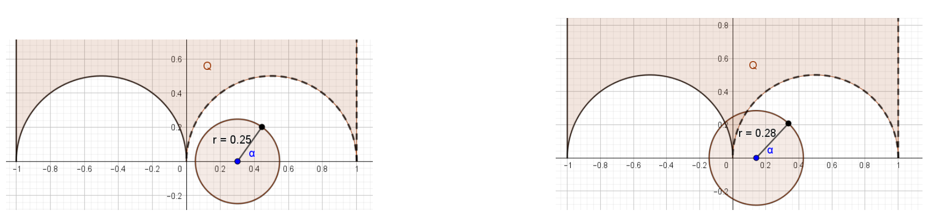


Figure 2: $\sigma(Q)$

We see that $\sigma(Q) \subseteq \overline{D(\frac{1}{2}, \frac{1}{2})}$, and since $\overline{D(\frac{1}{2}, \frac{1}{2})} \cap Q = \emptyset$, $\varphi(Q) \cap Q = \emptyset$ by the same argument as above. Hence (2) holds.

- Finally, we can let $c \neq 0$ and $c \neq 2d$. To show (2), I will make a contradiction using (1). To do this, I will need to show that $|cz + d| > 1$ for all $z \in Q$. Notice that $|cz + d| > 1 \iff z \notin \overline{D(-\frac{d}{c}, \frac{1}{|c|})}$. From the description of Q, one can see that if $\alpha \in \mathbb{R}$ and $\alpha \neq -1/2$ and $r > 0$, then $\overline{D(\alpha, r)} \cap Q \neq \emptyset$ only if at least one of the points $-1, 0, 1$ is in $D(\alpha, r)$.



(a) Not intersecting

(b) Intersecting

Figure 3: The closure of a disk on $\mathbb{R}$ intersects Q iff -1, 0 or 1 is in the disk

Let $\alpha = -d/c$ and $r = \frac{1}{|c|}$, then $\overline{D(-\frac{d}{c}, \frac{1}{|c|})} \cap Q \neq \emptyset \iff w \in D(-\frac{d}{c}, \frac{1}{|c|}) \iff |cw + d| < 1$ where w is one of -1, 0 and 1. But this is impossible, because since c is an even integer, and d is an odd integer,

$|cw+d| \geq 1$. Hence $\overline{D(-\frac{d}{c}, \frac{1}{|c|})} \cap Q = \emptyset$ and $|cz+d| > 1$ for all $z \in Q$. From (1), we therefore get that for all $z \in Q$:

$$\operatorname{Im}(\varphi(z)) < \operatorname{Im}(z) \tag{3}$$

Notice that if $c = 2d$, then $-d/c = -1/2$, and $w \in D(-\frac{1}{2}, r)$ only if the open $D(-\frac{1}{2}, r) \cap Q \neq \emptyset$. Thus the inequality would not be strict. As for $\varphi^{-1}(z) = (dz-b)/(-cz+a)$, from (1), we get that $\operatorname{Im}(\varphi^{-1}(z)) = \operatorname{Im}(z)/|-cz+a|^2$. Since $c \neq 0$, we know that $-c \neq 0$, but we can't guarantee $-c \neq 2a$. The inequality is therefore not strict for the inverse in general. If there was a point $z \in Q$ where also $\varphi(z) \in Q$, we could use the inequality with the inverse on the point $\varphi(z)$:

$$\operatorname{Im}(z) = \operatorname{Im}(\varphi^{-1}(\varphi(z))) \leq \operatorname{Im}(\varphi(z)) \tag{4}$$

But now we get that $\operatorname{Im}(\varphi(z))$ is both greater and strictly smaller than $\operatorname{Im}(z)$. A contradiction. Thus such a z cannot exist and (2) holds.

This completes the proof of (a').
What remains to be shown, is (b). Let $\Sigma = \bigcup_{\varphi \in \Gamma} \varphi(Q)$. We want to show that $\Sigma = \Pi^+$. By (1), we know that $\Sigma \subseteq \Pi^+$. We also see that $\tau^n(Q) \subseteq \Sigma, \forall n \in \mathbb{Z}$ where $\tau^n(z) = z + 2n$.

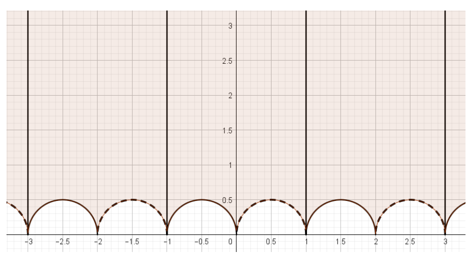


Figure 4: $\bigcup \tau^n(Q)$

Since σ maps $|2z+1| = 1$ to $|2z-1| = 1$, we see that Σ also contains the right semicircle boundary of the shifted Q regions. Therefore, Σ contains all $z \in \Pi^+$ satisfying all inequalities

$$|2z - (2m+1)| \geq 1, m \in \mathbb{Z} \tag{5}$$

Fix $w \in \Pi^+$. Since c, d are integers, there are only finitely many pairs of c, d such that $|cw+d|$ is below a given bound. We can therefore choose φ_0 so that $|cw+d|$ is minimized. Hence by (1), we have:

$$\operatorname{Im}(\varphi(w)) \leq \operatorname{Im}(\varphi_0(w)) \quad \forall \varphi \in \Gamma \tag{6}$$

Let $z = \varphi_0(w)$. (6) is then equivalent to $\operatorname{Im}(\varphi(\varphi_0^{-1}(z))) \leq \operatorname{Im}(z)$ by substitution. For any $\varphi' \in \Gamma$, $\varphi' = (\varphi' \circ \varphi_0) \circ \varphi_0^{-1}$ and since Γ is a group, we know that $\varphi' \circ \varphi_0 \in \Gamma$. We therefore get

$$\operatorname{Im}(\varphi(z)) \leq \operatorname{Im}(z) \quad \forall \varphi \in \Gamma \tag{7}$$

I will now show that z satisfies (5). Consider the following compositions ($n \in \mathbb{Z}$):

$$(\sigma \circ \tau^{-n})(z) = \frac{(z-2n)}{2(z-2n)+1} = \frac{z-2n}{2z-4n+1} \tag{8}$$

$$(\sigma^{-1} \circ \tau^{-n})(z) = \frac{(z-2n)}{-2(z-2n)+1} = \frac{z-2n}{-2z+4n+1} \tag{9}$$

By (1) we have

$$\operatorname{Im}(\sigma \circ \tau^{-n})(z) = \frac{\operatorname{Im}(z)}{|2z-4n+1|^2} \tag{10}$$

$$\operatorname{Im}(\sigma^{-1} \circ \tau^{-n})(z) = \frac{\operatorname{Im}(z)}{|2z-4n-1|^2} \tag{11}$$

and since they are both in Γ, we can use (7) to get

$$\operatorname{Im}(\sigma \circ \tau^{-n})(z) \leq \operatorname{Im}(z) \implies |2z-4n+1| \geq 1 \tag{12}$$

$$\operatorname{Im}(\sigma^{-1} \circ \tau^{-n})(z) \leq \operatorname{Im}(z) \implies |2z-4n-1| \geq 1 \tag{13}$$

We see that z therefore satisfies (5), and thus $z \in \Sigma$. There must therefore be a point $z_0 \in Q$ and $\varphi \in \Gamma$ where $\varphi(z_0) = z = \varphi_0(w) \implies w = \varphi_0^{-1}(\varphi(z_0))$. Since Γ is a group, we know that $\varphi_0^{-1} \circ \varphi \in \Gamma$. Thus $w \in \Sigma$. This completes the proof of (b), and hence the theorem is proved. □

4.3 The modular function λ

This theorem and proof is from Rudin [1] (p. 330-331). Let Γ and Q be as in the previous section.

Theorem 4.2. *There is a function λ holomorphic on Π^+, where*

(a) $\lambda \circ \varphi = \lambda, \forall \varphi \in \Gamma$

(b) λ is injective on Q

(c) $\lambda(\Pi^+) = \mathbb{C} \backslash \{0, 1\} =: \Omega$

(d) λ has $\mathbb{R}$ as its natural boundary

Proof. Notice that by (a), $\lambda(Q) = \lambda(\varphi(Q)), \forall \varphi \in \Gamma$. And by Theorem 4.1 (b), $\bigcup_{\varphi \in \Gamma} \varphi(Q) = \Pi^+$. Hence $\lambda(Q) = \lambda(\Pi^+)$.

Let Q_0 be the open right half of Q, i.e that for $z = x + yi$,

$$y > 0, \quad 0 < x < 1, \quad |2z-1| > 1$$

By theorem 3.1 (Riemann mapping theorem), since Q_0 is non-empty, open and simply connected, there exists a biholomorphic function $f : Q_0 \to D$ where D is the unit disk centered at 0. Recall from the previous section that $\sigma(Q)$ is

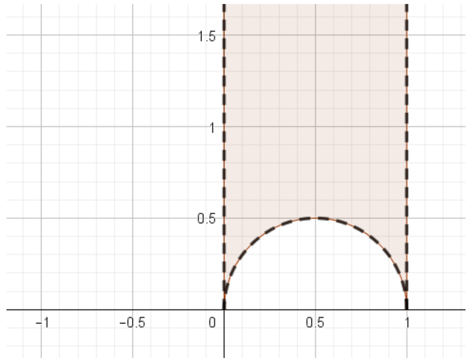


Figure 5: Q_0

bounded. Since $Q_0 \subseteq Q$, $\sigma(Q_0)$ is also bounded. We see that every finite boundary point of Q_0 is simple, but also that every sequence in Q_0 that converges to ∞ on $\hat{\mathbb{C}}$, can be connected by a continuous curve within Q_0 that also converges to ∞. Thus the boundary of Q_0 is simple. Hence, by Theorem 3.3, f extends to a bijective and continous function from $\overline{Q_0}$ to $\overline{D}$. Since $0, 1, \infty$ are on the boundary of Q_0, $f(0), f(1), f(\infty)$ are on the boundary of D, and since f is injective, they are distinct. Travelling along ∂Q_0 counter-clockwise, the interior is on the left, hence travelling along ∂D counter-clockwise, the interior must also be on the left. Therefore $f(0), f(1), f(\infty)$ must be mapped to, counter-clockwise. Consider $\psi(z) = (z-i)/(z+i)$. Since $|z-i| < |z+i|, \forall z \in \Pi^+$, this maps Π^+ to D. Also, $\psi(0) = -1, \psi(1) = -i, \psi(\infty) = 1$. Let φ be the möbius transformation that maps $f(0), f(1), f(\infty)$ to $-1, -i, 1$ respectively. Since φ maps three points on the unit circle to three points on the unit circle, the unit circle is preserved. And since $-1, -i, 1$ are also ordered counter-clockwise, the interior of the disk is preserved.

Let $h = \psi^{-1} \circ \varphi \circ f$. Then for example $h(1) = \psi^{-1}(\varphi(f(1))) = \psi^{-1}(-i) = 1$. We see that 1 is a fixed point of h. The same is true of 0 and ∞. We also see that $h(Q_0) = \Pi^+$, which means that $h(\partial Q_0) = \mathbb{R}$. Thus $h : \overline{Q_0} \to \Pi^+ \cup \mathbb{R}$ with $0, 1$ as fixed points and $h(\infty) = \infty$.

Let $i\mathbb{R}^+$ denote the positive imaginary axis. Since h is real on ∂Q_0, $h(0) = 0$, $h(1) = 1$, and because h is continuous and injective on ∂Q_0, we see that $h(i\mathbb{R}^+) = \mathbb{R}^-$, the semicircle centered at $1/2$ gets mapped to $(0, 1)$ and $h(1 + i\mathbb{R}^+) = (1, \infty)$. Therefore by the Schwartz reflection principle,

$$h(-x + yi) = \overline{h(x + yi)} \tag{1}$$

extends h to be holomorphic on interior(Q) and injective and continuous in Q. Since h is real on ∂Q_0, using (1), we see that

$$h(-1 + yi) = h(1 + yi) = h(\tau(-1 + yi)), \quad (y \geq 0) \tag{2}$$

$$h(-\frac{1}{2} + \frac{1}{2}e^{i\theta}) = h(\frac{1}{2} + \frac{1}{2}e^{i(\pi-\theta)}) = h(\sigma(-\frac{1}{2} + \frac{1}{2}e^{i\theta})), \quad (\theta \in [0, \pi]) \tag{3}$$

From these observations, since the left boundaries was in Q, we see that $h(Q)$ contain $(1, \infty)$ and $(0, 1)$, and because $i\mathbb{R}^+ \subset Q$, $h(Q)$ also contain $\mathbb{R}^- = (-\infty, 0)$. Thus $h(Q) = \mathbb{C}\backslash\{0, 1\} = \Omega$. I will now define λ:

$$\lambda(z) = h(\varphi^{-1}(z)) \quad z \in \Pi^+, \varphi \in \Gamma \tag{4}$$

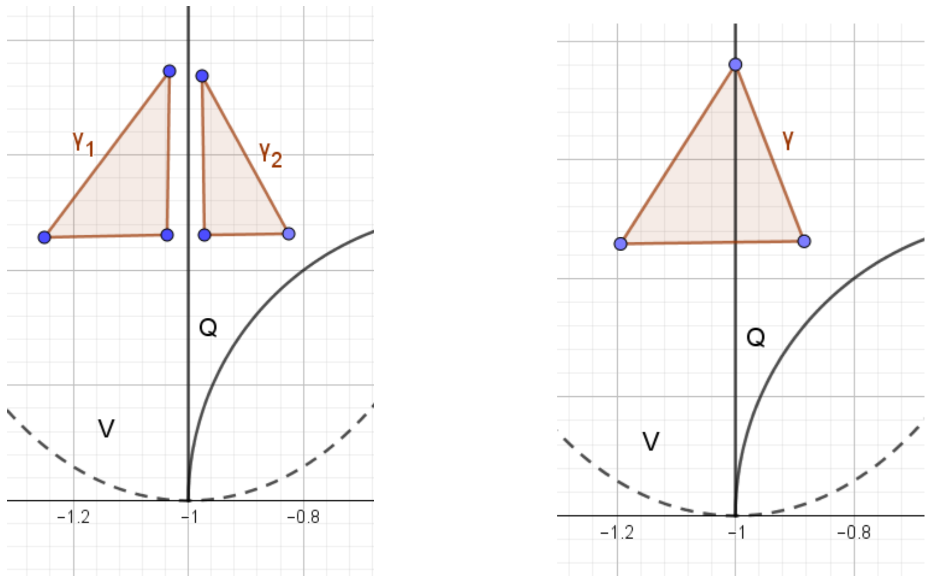
By theorem 4.1, $\forall z \in \Pi^+, \exists! \varphi \in \Gamma, z \in \varphi(Q)$. Thus (4) defines λ uniquely in the upper half-plane. If $z \in Q$, then the corresponding φ^{-1} is the identity map. Thus $\lambda(z) = h(z)$ in Q, and since h is injective in Q, so is λ. Hence (b) holds. And since then $\lambda(Q) = h(Q) = \Omega$, (c) holds if (a) holds. I will now prove (a). Let $z \in \Pi^+$ and let $\varphi \in \Gamma$ be the unique möbius transformation where $z \in \varphi(Q)$. Then $\lambda(z) = h(\varphi^{-1}(z))$. Let ψ be any möbius transformation in Γ, and let $\varphi' \in \Gamma$ be the unique möbius transformation where $\psi(z) \in \varphi'(Q)$. Then $\lambda(\psi(z)) = h(\varphi'^{-1}(\psi(z)))$. But from theorem 4.1, we know that for a fixed $w \in \Pi^+$, $\exists! \Phi \in \Gamma$ where $\Phi^{-1}(w) \in Q$. Therefore $\varphi^{-1} = \varphi'^{-1} \circ \psi$ and thus $\lambda(z) = \lambda(\psi(z))$, and (a) is proved.

Since h is holomorphic on interior(Q), λ is holomorphic on all interior$(\varphi(Q))$ for $\varphi \in \Gamma$, by (a). For $z \in \tau^{-1}(Q)$, $\lambda(z) = h(\tau(z))$, and on $-1 + i\mathbb{R}^+ \subset Q$, $\lambda(z) = h(z)$. Hence by (2), λ is continuous on $Q \cup \tau^{-1}(Q)$. Similarily on $\sigma^{-1}(Q)$, $\lambda(z) = h(\sigma(z))$, and on the left semicircle border of Q, $\lambda(z) = h(z)$. Hence by (3), λ is continuous on $Q \cup \sigma^{-1}(Q)$. Together, this tells us that λ is continuous on $Q \cup \tau^{-1}(Q) \cup \sigma^{-1}(Q)$. Therefore, λ is also continous on an open set V which contains Q.

Let γ_1 be a triangular path outside Q, but within V, where one side is close to and parallel to the line $-1 + i\mathbb{R}^+$. Let γ_2 be a triangular path in Q with one side close to and parallel to the same line. Both γ_1 and γ_2 are oriented counter-clockwise. Since V is simply connected, and λ is holomorphic in the interior of Q and outside Q, by cauchy's integral theorem,

$$\oint_{\gamma_1} \lambda(z)\, dz = 0 = \oint_{\gamma_2} \lambda(z)\, dz$$

But since λ is continuous on V, this still holds even if we move γ_1 and γ_2 together, so that the sides parallell to the line coincide on the line. On these overlapping segments, γ_1 goes upwards, and γ_2 goes downwards, thus their contribution to the sum of the contour integrals is zero. Let γ be a triangle constructed in this way when excluding the overlapping sides. Then the closed contour integral of λ over γ is zero for all triangles γ on the line. Hence, by Morera's theorem, λ is holomorphic also on the line. By the use of a möbius transformation ψ, you can move the left semicircle of Q to the line. This shows that $\lambda(\psi(z))$ is holomorphic on the left semicircle, which means that λ is also holomorphic on the left semicircle. Thus λ is holomorphic in all of V, and since $\varphi(V)$ covers Π^+ by theorem 4.1 (b), where $\varphi \in \Gamma$, and since $\lambda \circ \varphi = \lambda$, we can conclude that λ is holomorphic in all of Π^+.



(a) γ_1 and γ_2 (b) γ

Figure 6: The contour integral remains zero because λ is continuous

What remains to be shown, is that λ has the real axis as its natural boundary. This means that it cannot be analytically continued there. Notice that $\lambda(b/d) = h(\varphi^{-1}(b/d)) = h(0) = 0$. And since $\lambda \circ \varphi = \lambda$, every fraction b/d is a zero. But this is a dense subset of $\mathbb{R}$, thus if λ could be analytically continued to $\mathbb{R}$, its zeros would have a limit point, and therefore λ would be identically equal to zero by the identity theorem, but this is not possible, as λ is not constant. $\square$

The fact that $\mathbb{R}$ is a natural boundary of λ, is why we can't use the proof of Picard for entire functions that omit only one point. To do that, one would have to extend the domain of λ to a region containing 0 or 1, which would then make it constant.

One might think that h would be enough to prove Picard, since it maps something in the upper half-plane to Ω, which means that h^{-1} would map Ω to something in Π^+. You could then use a möbius transformation to get a mapping from Ω to a region contained in the unit disk. Then composing with an entire function that maps to Ω to get a constant by Liouville's theorem. However, because of the reflection of the boundaries of Q, h^{-1} is actually not continuous, and Liouville's theorem would therefore not apply. Recall, h maps Q_0 to Π^+, the open left half of Q gets mapped to Π^- (the lower half-plane), both semicircle boundaries of Q to $(0, 1)$, and both vertical boundaries of Q to $(1, \infty)$. This means that if $D \subset \Omega$ is a disk that intersects $(1, \infty)$, then h^{-1} would map the open upper half of D to a region in Q_0 adjacent to the right vertical boundary of Q, and $D \cap (1, \infty)$ and the lower half of D would map to a set in the left half of Q adjacent to and on the left vertical border of Q. Notice that neither part of $h^{-1}(D)$ include 0 or 1, since 0 and 1 are fixed points of h, and since $D \subset \Omega$.

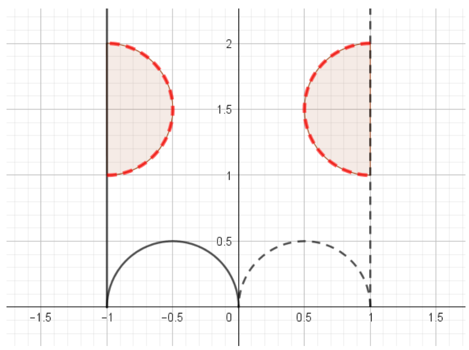


Figure 7: $h^{-1}(D)$ where D is a disk in Ω that intersects $(1, \infty)$

5 Proof of Picard's little theorem

Theorem. *If f is an entire function that omits two distinct values in $\mathbb{C}$, then f is constant. [1] (p. 332)*

Let f be an entire function that omits two values. If f omits α and β, with $\alpha \neq \beta$, we could consider $\frac{f-\alpha}{\beta-\alpha}$ which omits 0 and 1. We can therefore assume without loss of generality that f omits 0 and 1.

Let $Q, \Gamma, h, \lambda, \Omega$ be as in chapter 4. ($\Omega = \mathbb{C} \backslash \{0, 1\}$) For each disk $D_1 \subset \Omega$, and for each $\varphi \in \Gamma$, there is a region $V_1 \subset \Pi^+$ where λ is injective on V_1 and $\lambda(V_1) = D_1$. If D_1 does not intersect $(0, 1)$ or $(1, \infty)$, then V_1 is a branch of $\lambda^{-1}(D_1)$ and is contained in $\varphi(Q)$. If D_1 does intersect $(0, 1)$ or $(1, \infty)$, then $h^{-1}(D_1)$ is disjoint, and since $\lambda^{-1} = \varphi \circ h^{-1}, \forall \varphi \in \Gamma$, each branch of $\lambda^{-1}(D_1) = \varphi(h^{-1}(D_1))$ is disjoint. But since the transformations $\varphi \in \Gamma$ map the two pieces from Q to countably infinitely many sets in Π^+ that are adjacent, the images of the pieces can be glued together at the boundaries of the sets $\varphi(Q)$ to create regions. V_1 is in this case one of these glued pieces. Since 0 and 1 are not in D_1, D_1 can intersect at most one of the intervals $(-\infty, 0), (0, 1), (1, \infty)$. Hence by the mapping properties of h^{-1}, V_1 intersects exactly one of the images of a vertical border of Q, a semicircle border of Q and the positive imaginary axis in Q by φ. Thus V_1 intersects exactly two of the sets $\varphi(Q)$. Since λ is modular, it makes no difference if we consider V_1 to be the glued pieces from the left half or the right half of $\varphi(Q)$. In fact, we see that $\lambda(\varphi(h^{-1}(D_1))) = \lambda(V_1) = D_1$ for all $\varphi \in \Gamma$, even though $\varphi(h^{-1}(D_1))$ is disjoint, and V_1 is not.

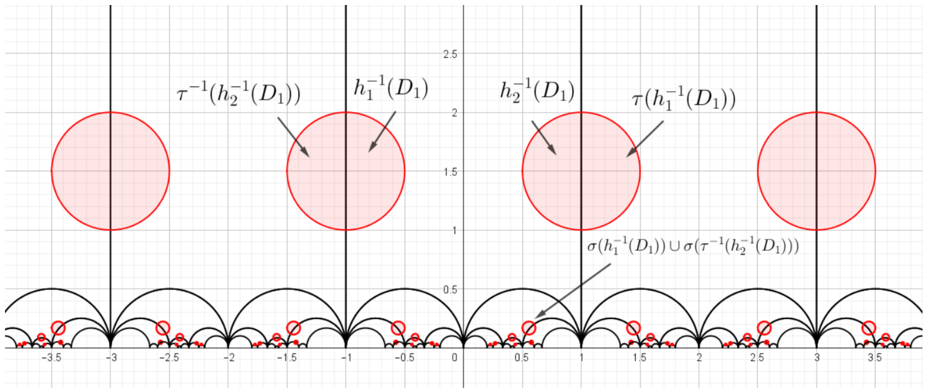


Figure 8: Possible choices of V_1 (in red) from the glued pieces of the branches of $\lambda^{-1}(D_1)$ where $D_1 \subset \Omega$ is a disk that intersects $(1, \infty)$. These regions are open, but is shown with a full border here, to make it easier to see. This figure is also only 2 layers deep. There are countably infinitely many choices of V_1 within each vertical strip. All figures in this thesis have been constructed in Geogebra version 5.

For each choice of V_1, there is an inverse ψ_1 holomorphic on D_1 so that $\psi_1(\lambda(z)) = z$ for all $z \in V_1$. Let D_2 be another disk in Ω, where $D_2 \cap D_1 \neq \emptyset$. Then there is a corresponding inverse ψ_2 and region V_2 where $V_2 \cap V_1 \neq \emptyset$. The function elements (ψ_1, D_1) and (ψ_2, D_2) are then direct analytic continuations of each other. Since $\lambda : \Pi^+ \to \Omega$, we know that the codomain of ψ_i is Π^+, and in particular that $\psi_i(D_i) \subset \Pi^+$.

Since $f(\mathbb{C}) = \Omega$, there is a disk $A_0 \subseteq \mathbb{C}$ centered at 0, so that $f(A_0) \subseteq D_0 \subset \Omega$, where D_0 is a disk. Choose $V_0 \subset \Pi^+$ and ψ_0 as above, so that ψ_0 is holomorphic on D_0, λ is injective on V_0, $\lambda(V_0) = D_0$ and $\psi_0(\lambda(z)) = z$ for all $z \in V_0$. Let $g = \psi_0 \circ f$ in A_0, and let γ be any continuous curve in $\mathbb{C}$ starting at 0. We can then choose chains of disks $A_0, A_1, \ldots, A_n$ in $\mathbb{C}$ that covers γ, and $D_0, D_1, \ldots, D_n$ in Ω so that each $f(A_i)$ is contained in the disk D_i, where $A_i \cap A_{i+1} \neq \emptyset$ and $D_i \cap D_{i+1} \neq \emptyset$ for all $i = 0, \ldots, n-1$. We can then also choose regions V_i so that $V_i \cap V_{i+1} \neq \emptyset$ and the corresponding ψ_i so that $(\psi_i, D_i) \sim (\psi_{n+1}, D_{n+1})$ for all $i = 0, \ldots, n-1$. This gives the analytic continuation (ψ_n, D_n) of (ψ_0, D_0) along the chain $D_0, \ldots, D_n$, and the analytic continuation $(\psi_n \circ f, A_n)$ of (g, A_0) along the chain $A_0, \ldots, A_n$. Since $\psi_i(\Omega) \subset \Pi^+$, we know that $\psi_n(f(\mathbb{C})) \subset \Pi^+$.

Since the function element (g, A_0) can be analytically continued along any curve γ in the complex plane, and since the plane is simply connected, by the monodromy theorem, g can be extended to an entire function G which then also must map to the upper half-plane. This is how we circumvent the problem mentioned in chapter 4.3. By using analytic continuation on an inverse of λ, we get a function with the same mapping properties as $h^{-1} \circ f$, mapping $\mathbb{C}$ to something contained in Π^+, but this time being continuous and entire. $G(\mathbb{C})$ will be a region similar to one of the $\varphi(Q)$ sets. As alluded to in chapter 3.1.4, we now get that $F := (G - i)/(G + i)$ maps the complex plane to a region contained in the open unit disk, which is therefore also bounded. And since G is entire, so is F, and then by Liouville's theorem, F is constant in all of $\mathbb{C}$. Since $(z - i)/(z + i)$ is not constant, by theorem 1.6, G is constant in all of $\mathbb{C}$, which must mean that g is constant on A_0. Since ψ_0 is injective on $D_0 \supseteq f(A_0)$, and since ψ_0 is non-constant, f must be constant on A_0. And then finally, since A_0 is a non-empty open disk, by the Identity theorem, f is constant in all of $\mathbb{C}$.

QED

References

[1] W. Rudin, *Real and complex analysis.* McGraw-Hill Book Co., 1987.

[2] D. N. Clark, *Dictionary of Analysis, Calculus, and Differential Equations.* CRC Press, 1999.

[3] E. J. N. Erwin Kreyszig, Herbert Kreyzig, *Advanced engineering mathematics.* John Wiley & Sons Pte Ltd., 10 th edition, International Student Version, 2012.

[4] R. S. Elias M. Stein, *Complex Analysis (Princeton Lectures in Analysis, Volume II).* Princeton university press, 2003.

[5] D. Sarason, *Complex function theory.* American Mathematical Soc., 2007.

NTNU

Norwegian University of
Science and Technology