Chapter 7

Analytic Continuation

It often happens in solving complex variables problems, particularly where the solution is obtained as the sum of a series, that we are led to an analytic function f, defined, as far as we know, just on a certain open set U. Yet it may turn out that f can be extended to be an analytic function on a much larger open set. For example, the power series

$$\sum_{n=0}^{\infty} z^n$$

converges on the open unit disc D to a function analytic on D. In fact, this function is

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

on D, and it can obviously be extended to a function, given by the same algebraic expression, which is analytic on the much larger set $\mathbb{C} \setminus \{1\}$.

According to the Inverse Function Theorem, A function f which is analytic on an open set U and which has a non-zero derivative at a point $z_0 \in U$ has an analytic inverse function defined in some neighborhood W of $f(z_0)$. Although the theorem assures us of an inverse function just in some possibly very small neighborhood of $f(z_0)$, we would naturally like to know on just how large a set can the inverse function be defined.

There are many problems of this nature, problems where a solution in the form of an analytic function is known or is known to exist locally, in a neighborhood of a point, and the question arises as to just how much we can enlarge the domain on which this function is defined, is analytic, and is still a solution to the original problem. Questions of this type are questions of *analytic continuation.*

7.1 The Schwarz Reflection Principle

One way to extend an analytic function to a larger domain, is to take advantage of some kind of symmetry operation. The Schwarz Reflection Principle does

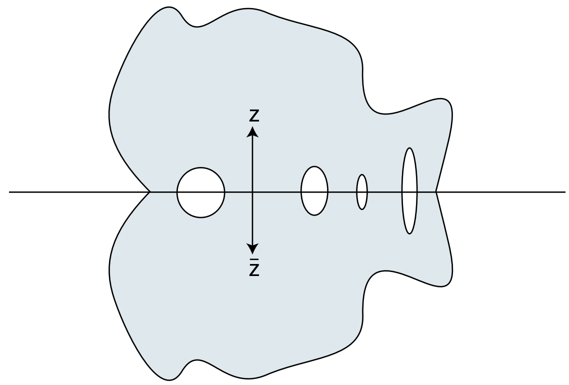

Figure 7.1: Continuation by reflection.

exactly this by making use of the symmetry operation $z \to \overline{z}$, which sends a point to its reflection through the x axis.

Schwarz Reflection

Figure 7.1 illustrates the situation described in the next theorem.

Theorem 7.1.1. (Schwarz Reflection Principle) *Let U be an open set which is symmetric about the x-axis $\mathbb{R}$. Let $A = U \cap \mathbb{R}$ and let V be the part of U in the open upper half plane. If f is a function which is continuous on $V \cup A$, analytic on V, and real valued on A, then f can be continued to a function analytic on all of U by setting its continuation equal to $\overline{f(\overline{z})}$ on $\overline{V}$.*

Proof. The function g, defined by

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

is analytic on the set

$$\overline{V} = \{\overline{z} : z \in V\}.$$

(Exercise 2.2.15)

Note that $\overline{V}$ is the part of U which lies in the open lower half plane, since U is symmetric about the x-axis. Note also that the function g is continuous on $\overline{V} \cup A$, since f is continuous on $V \cup A$.

The fact that f is real on A implies that f and g agree on A and, hence, that they define a single continuous function on $U = V \cup A \cup \overline{V}$. This function is analytic on U, except possibly at points of A.

Now A is an open subset of the real line and, hence, is a union of open line segments. Let L be one of these line segments. Then the set $W = V \cup L \cup \overline{V}$ is an open subset of U and f is continuous on W and analytic on W except possibly at points of L. However, it is a simple consequence of Morera's Theorem that a function which is continuous on an open set W and analytic on W except possibly at the points of a line segment in W is actually analytic on all of W (Exercise 3.2.14). Since f is analytic on each of the open sets $W = V \cup L \cup \overline{V}$ for L one of the line segments making up A, it is anaytic on the union of these sets, and U is this union. This completes the proof. □

Continuation Across a Simple Analytic Curve

Recall from our discussion of the Riemann Sphere S^2 in Section 6.3, that lines in $\mathbb{C}$ can be thought of as circles in S^2 that pass through the point at infinity. In fact, the set of all lines and circles in $\mathbb{C}$ is just the set of all circles in S^2.

Definition 7.1.2. An analytic curve is the image under a conformal equivalence $h : W \to V$, between domains in S^2, of a set $I \subset W$ which is a circle or an open arc on a circle in S^2.

Of course, by using a linear fractional transformation to move W, we can always assume that the circle in the above definition is $\mathbb{R} \cup \{\infty\}$ and I is either an open interval on the line or is $\mathbb{R} \cup \{\infty\}$ itself. In the later case, the curve will be a simple closed curve.

In the proof of the Schwarz Reflection Principle, we used the fact that a function which is continuous on an open set U and analytic on U except possibly at the points of a line segment in U is actually analytic on all of U (Exercise 3.2.14). We can strengthen this as follows:

Theorem 7.1.3. *If f is a function which is continuous on an open set U and analytic on U except possibly at the points of a simple analytic curve C in U, then f is actually analytic on all of U.*

Proof. By removing a point p from C, if necessary, we obtain a simple analytic curve C_1 which is not a closed curve. Then there is a domain V containing C_1, a domain W which contains an open interval $L \subset \mathbb{R}$, and a conformal equivalence $h : W \to V$ which maps L onto C_1. We may assume that $V \subset U$. Let $h^{-1} : V \to W$ denote the inverse function of h.

The composition $f \circ h$ is a continuous function on W which is analytic on $W \setminus L$. It is, therefore, analytic on all of W. Since, on V,

$$f = f \circ h \circ h^{-1},$$

we conclude that f is analytic on all of V. Since it is analytic on each of two open sets, V and $U \setminus C$, whose union is $U \setminus \{p\}$, f is actually analytic on all of $U \setminus \{p\}$. Since it is continuous at p, it is analytic on all of U, by Theorem 3.4.8. □

Example 7.1.4. Prove that a function f which is continuous on an open set U and analytic on $U \setminus C$, where C is an open arc of a circle, is actually analytic on all of U.

Solution: An open arc of a circle is a simple analytic curve. Thus, by the above theorem, f is analytic on all of U.

Reflection Through an Analytic Curve

The symmetry consisting of reflection through the x-axis played the key role in the Schwarz Reflection Principle. It turns out there is a similar reflection symmetry for each simple analytic curve and it leads to a similar reflection principle. To describe this requires the notion of a *conjugate analytic* function.

A function f on an open set U is said to be *conjugate analytic* if $\overline{f}$ is analytic on U, where

$$\overline{f}(z) = \overline{f(z)}.$$

Definition 7.1.5. Suppose U is a domain in $\mathbb{C}$ and C is a simple analytic curve in U such that $U \setminus C$ has two connected components V and W. If ρ is a conjugate analytic function from U to U which fixes each point of C and interchanges V and W, and if

$$\rho \circ \rho(z) = z \quad \text{for every} \quad z \in U, \tag{7.1.1}$$

then ρ is called a reflection through C defined on U.

Note that (7.1.1) implies that ρ is its own inverse function on U. If $\kappa(z) = \overline{z}$, then $\kappa \circ \rho$ and $\rho \circ \kappa$ are both analytic functions and are inverse functions of one another. This implies that $\kappa \circ \rho$ is a one to one conformal map. Thus, each reflection defined on U is the conjugate of a conformal equivalence from U to $\overline{U}$.

Theorem 7.1.6. *If C is any simple analytic curve in $\mathbb{C}$, then there is a reflection ρ through C defined on some domain V containing C. This reflection is unique in the sense that another reflection through C, defined on a neighborhood V_1 of C, must be equal to ρ on the connected component of $V \cap V_1$ which contains C.*

Proof. Since C is a simple analytic curve, there is a conformal equivalence

$$h : W \to V,$$

where W is a domain in S^2 which meets $\mathbb{R} \cup \{\infty\}$ in a set L which is either $\mathbb{R} \cup \{\infty\}$ or a line segment in $\mathbb{R}$, V contains C, and C is the image of L under h.

We may as well assume that $\overline{W} = W$, since, otherwise, we can replace W by the connected component of $W \cap \overline{W}$ containing L without affecting its intersection, L, with $\mathbb{R} \cup \{\infty\}$, or the fact that its image under h is a domain containing C.

With the assumption that $\overline{W} = W$, conjugation defines a reflection κ through L on W ($\kappa(z) = \overline{z}$). We then get a reflection ρ through C on V by setting

$$\rho = h \circ \kappa \circ h^{-1}$$

on V, where $h^{-1} : V \to W$ is the inverse function for h. The mapping ρ is conjugate analytic because it is the composition of an analytic function h with a conjugate analytic function $\kappa \circ h^{-1} = \overline{h^{-1}}$ (Exercise 7.1.6). It also satisfies

$$\begin{aligned}\rho \circ \rho(z) &= h \circ \kappa \circ h^{-1} \circ h \circ \kappa \circ h^{-1}(z) \\ &= h \circ \kappa \circ \kappa \circ h^{-1}(z) = z.\end{aligned}$$

on V, and

$$\rho(z) = z$$

on C, because $\kappa(z) = z$ on L and h^{-1} maps C to L. Thus, ρ is a reflection through C defined on V.

If σ is any other reflection through C, defined on V_1, then $\sigma \circ \rho(z)$ is an analytic function of z on $V \cap V_1$, since the composition of two conjugate analytic functions is analytic (Exercise 7.1.7). However, it is also equal to z on C, because this is true of both $\sigma(z)$ and $\rho(z)$. But then it must be equal to z on all of the connected component of $V \cap V_1$ containing C. This implies $\rho = \sigma^{-1} = \sigma$ on this component. Thus, ρ is unique in the sense described in the theorem. □

It is now a simple matter to prove that a version of the Schwarz Reflection Principle holds for reflection through a simple analytic curve.

Theorem 7.1.7. *If U is a domain, $C \subset U$ is a simple analytic curve with $U \setminus C$ having two components V and W, and ρ is a reflection through C defined on U, then any function f which is continuous on $V \cup C$, analytic on V, and real valued on C can be continued to an analytic function defined on all of U by setting its continuation equal to $\overline{f(\rho(z))}$ on W.*

Proof. The proof is the same as the proof of the Schwarz Reflection Principle. We set

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

on $W \cup C$ and note that g is continuous on $W \cup C$ and analytic on W because it is the composition of two conjugate analytic functions, $\overline{f}$ and ρ (Exercise 7.1.7). The functions g and f agree on C because ρ fixes points of C and f and g are real valued on C. Hence, they define a single function on U which is continuous on U and analytic on $U \setminus C$. By Theorem 7.1.3, this function is analytic on all of U. Since it agrees with f on $V \cup C$, it is an analytic continuation of f to U. □

Example 7.1.8. What is the reflection through an arc on the unit circle and what does the previous theorem say about continuing an analytic function across such an arc?

Solution: The map

$$\rho(z) = \frac{1}{\overline{z}} \quad \text{or} \quad \rho(re^{i\theta}) = \frac{1}{r}e^{i\theta}$$

is defined and conjugate analytic on $\mathbb{C} \setminus \{0\}$. It satisfies $\rho \circ \rho(z) = z$. It also fixes each point on the unit circle since, for z on the unit circle, $|z|^2 = z\overline{z} = 1$, which implies $z = 1/\overline{z}$. It follows that, for any domain U which meets the unit circle in an arc C and is taken to itself by ρ, the unique reflection through C, defined on U, is the map ρ.

For such a U and C, the previous theorem implies that any function analytic on the part V of U lying on one side of C, continuous on $V \cup C$, and real valued on C can be analytically continued to an analytic function on all of U.

Exercise Set 7.1

1. The power series $\sum_{n=0}^{\infty}(n+1)z^n$ defines an analytic function g on the open unit disc. What is the largest open set in the plane to which g can be analytically continued.

2. Suppose that the hypotheses for the Schwarz Reflection Principle (Theorem 7.1.1) are satisfied for f, U, V and A, except that the function f is purely imaginary on A rather than purely real. Is it still true that f can be analytically continued to all of U? If so, prove it, and describe the formula for the continuation of f to $\overline{V}$.

3. If L is the line through the origin with equation $ax + by = 0$, find a formula for the map which is reflection through L, and verify that this map is conjugate analytic.

4. The function $\sqrt{z^2 - 1}$ is defined and analytic on the set V consisting of the open right half plane with the interval $(0, 1]$ removed, provided we use the principle branch of the log function to define the square root (this is because $z^2 - 1$ lies in the complement of the non-negative real axis if $z \in V$). This function can be extended to be continuous on the closed right half plane with the interval $[0, 1]$ removed. Show how to use the Schwarz Reflection Principle to extend this function to an analytic function on $\mathbb{C} \setminus [-1, 1]$.

5. Formulate and prove a version of the Schwarz Reflection Principle (Theorem 7.1.1) for meromorphic functions.

6. Prove that the composition $f \circ g$ of an analytic function f with a conjugate analytic function g is conjugate analytic.

7. Prove that the composition $f \circ g$ of two conjugate analytic functions is analytic.

8. Find a formula for the reflection through the circle of radius r centered at z_0.

9. Prove that the curve in the plane with equation $y = x^2$ is a simple analytic curve.

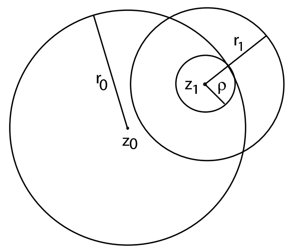

Figure 7.2: Continuation to Another Disc.

10. Prove the following form of the Schwarz Reflection Principle: Suppose C_1 and C_2 are circles and the corresponding reflections through these circles are ρ_1 and ρ_2. If U is a domain containing C_1 which is symmetric under ρ_1, and V is the part of U which lies on one side of C_1 (inside or outside), then any analytic function f on V which is continuous on $V \cup C_1$ and maps C_1 to $\mathbb{C}_2$ may be analytically continued to a function defined on all of U by setting it equal to $\rho_2 \circ f \circ \rho_1$ on $\rho_1(V)$ (the part of U which lies on the other side of C_1).

11. The power series $\displaystyle\sum_{n=0}^{\infty} z^{n!}$ defines an analytic function f on the open unit disc. Prove that f cannot be analytically continued to any larger open set. Hint: consider the values of f along rays of the form $z = re^{2\pi ip/q}$, where p and q are integers.

7.2 Continuation Along a Curve

Suppose f is a function which is analytic on a disc $D_{r_0}(z_0)$. One way to try to extend f to a larger domain is to consider its power series expansion about a point $z_1 \in D_{r_0}(z_0)$ which is different from z_0. We know this power series expansion converges to f in the disc $D_\rho(z_1)$, where $\rho = r_0 - |z_1 - z_0|$, but it might converge in a larger disc $D_{r_1}(z_1)$, where $r_1 > \rho$ (see Figure 7.2). If it does, then it converges to an analytic functiion g, defined on $D_{r_1}(z_1)$, which is equal to f on $D_\rho(z_1)$.

Since $f - g$ is defined on $D_{r_0}(z_0) \cap D_{r_1}(z_1)$ and is zero on $D_\rho(z_1)$, it follows from the Identity Theorem (Theorem 3.4.4) that $f - g = 0$ on all of $D_{r_0}(z_0) \cap D_{r_1}(z_1)$; that is, $f = g$ on this intersection. Thus, we succeed in extending f to an analytic function on the larger domain $D_{r_0}(z_0) \cup D_{r_1}(z_1)$ by setting it equal to g on $D_{r_1}(z_1)$.

Analytic continuation of an analytic function along a curve is an extension of this idea to the situation where a curve is covered by an overlapping sequence of discs, and an analytic function, given on the first disc, can be extended successively to each disc in the sequence (Figure 7.3). We will make this more precise

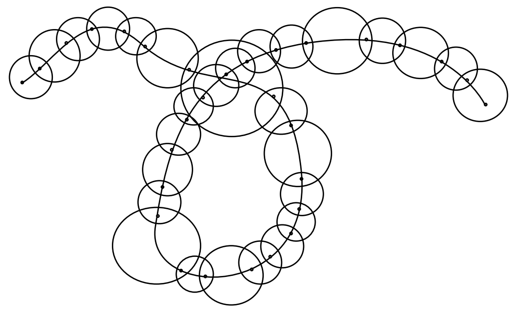

Figure 7.3: Continuation Along a Curve.

in the next definition, but first we need to introduce some useful terminology.

An *analytic function element*, (f, D), is an analytic function f defined on an open disc D. We will say that it is an analytic function element *at* w if $w \in D$. Two analytic function elements (f_1, D_1) and (f_2, D_2) at w are said to be *equivalent* at w if $f = g$ on $D_1 \cap D_2$.

Definition 7.2.1. Let $\gamma : [0, 1] \to \mathbb{C}$ be a curve and (f_0, D_0) an analytic function element at $z_0 = \gamma(0)$. Suppose there exist

(a) a partition $0 = t_0 < t_1 < \cdots < t_{n+1} = 1$ of $[0, 1]$; and

(b) a sequence of function elements

$$(f_1, D_1), (f_2, D_2), \cdots, (f_n, D_n),$$

with $\gamma([t_j, t_{j+1}]) \subset D_j$ for $j = 0, \cdots, n$ and $f_j = f_{j+1}$ on $D_j \cap D_{j+1}$ for $j = 0, \cdots, n - 1$.

Then we will say that (f_n, D_n) is an analytic continuation of (f_0, D_0) along γ.

Although the function element (f_n, D_n) of the above definition, seems to depend on the choice of partition $0 = t_0 < t_1 < \cdots < t_{n+1} = 1$ of $[0, 1]$ and sequence $(f_1, D_1), (f_2, D_2), \cdots, (f_n, D_n)$ of function elements, it turns out that, up to equivalence, it is actually independent of these choices.

Theorem 7.2.2. *Given a curve beginning at z_0 and ending at w and an analytic function element (f_0, D_0) at z_0, any two analytic continuations of (f_0, D_0) along γ are equivalent as analytic function elements at w.*

Proof. We first observe that, given an analytic continuation of (f_0, D_0) along γ, involving a particular partition $0 = t_0 < t_1 < \cdots < t_{n+1} = 1$ of $[0, 1]$ and corresponding function elements $(f, D) = (f_1, D_1), (f_2, D_2), \cdots, (f_n D_n)$, we can replace the original partition with any given refinement $0 = s_0, s_1, \cdots, s_{m+1} = 1$ and modify the sequence of function elements so that this new set of data still determines the same analytic continuation of (f_0, D_0). In fact, if $[t_j, t_{j+1}]$ is partitioned into k subintervals in the new partition, we simply assign (f_j, D_j) to each of these new subintervals (relabeling them in accordance with the labeling of the subintervals in the new partition). Obviously, this does not change the function element assigned to the last subinterval.

Given two analytic continuations of (f_0, D_0) along γ, by passing to a common refinement of the corresponding partitions, we may assume that the two continuations are defined using the same partition

$$0 = s_0 < s_1 < \cdots < s_{m+1} = 1.$$

Suppose one of them is defined from the sequence of function elements

$$(f_1, D_1), \cdots, (f_m, D_m),$$

while the other is defined from

$$(\tilde{f}_1, \tilde{D}_1), \cdots, (\tilde{f}_m, \tilde{D}_m)$$

(we set $(\tilde{f}_0, \tilde{D}_0) = (f_0, D_0)$). If we set $V_j = D_j \cap \tilde{D}_j$ for $j = 1, \cdots, m$, then, for each j, V_j is a connected (in fact, convex) open set containing $\gamma(t_j)$ and $\gamma(t_{j+1})$. To prove the theorem we need to show that $f_m = \tilde{f}_m$ on V_m. In fact, we will prove by induction on j that $f_j = \tilde{f}_j$ on V_j for each j.

We have $f_0 = \tilde{f}_0$ on all of V_0 by definition. Suppose $j < m$ and $f_j = \tilde{f}_j$ on V_j. We have $f_{j+1} = f_j$ on $D_j \cap D_{j+1}$ and $\tilde{f}_{j+1} = \tilde{f}_j$ on $\tilde{D}_j \cap \tilde{D}_{j+1}$. Since $\gamma(t_j)$ is in both of these open sets, it follows that $f_{j+1} = \tilde{f}_{j+1}$ in a neighborhood of $\gamma(t_j)$ and, hence, on all of V_{j+1}, since V_{j+1} is connected. This completes the induction step and finishes the proof.

□

The next theorem is almost obvious, but there is some work to be done to show the required chain of overlapping discs can be chosen as in Definition 7.2.1. We leave the details to the exercises (Exercise 7.2.1).

Theorem 7.2.3. *Suppose (f_0, D_0) is a function element at z_0 and γ is a curve joining z_0 to w. If there is an open set U, containing D_0 and $\gamma(I)$, and an analytic function f on U such that $f = f_0$ on D_0, then (f_0, D_0) can be analytically continued along γ.*

The converse of the above theorem is not true. It may seem that it should be true, since the union of the overlapping discs in Definition 7.2.1 is an open set U containing $\gamma(I)$, and it appears that the functions f_j fit together to define a single function, on this union, that agrees with f_0 on D_0 . However, this is not the case, due to the fact that the curve γ may cross itself, as in Figure 7.3 and, the crossing point may be contained in two of the discs D_j and D_k. The corresponding functions f_j and f_k may not be equivalent as function elements at this point. In fact, the curve might be a closed curve, so that $z_0 = \gamma(0)$ and $w = \gamma(1)$ are the same point. In this case the analytic continuation of (f_0, D_0) along γ leads to another function element (f_n, D_n) at z_0. This function element need not be equivalent to (f_0, D_0).

Example 7.2.4. Give an example of the phenomenon refered to in the previous paragraph.

Solution Let log denote the principle branch of the log function and consider the analytic function element $(\log, D_1(1))$. If γ is the unit circle traversed once in the counterclockwise direction beginning at $\gamma(0) = \gamma(1) = 1$, then $(\log, D_1(1))$ can be analytically continued along γ (Exercise 7.2.2) but the resulting function element on the final disc will differ from that on the first disc by $2\pi i$.

There is no domain containing the unit circle to which $(\log, D_1(1))$ can be analytically continued. If there were such a domain U and a continuation f of $(\log, D_1(1))$ to U, then f would have to agree with the principle branch of the log function on all of $U \setminus [-\infty, 0]$, but this function has a $2\pi i$ jump discontinuity across the half line $[-\infty, 0]$ and certainly cannot be analytically continued across it.

The previous example is closely related to the following question: If γ_1 and γ_2 are two curves joining z_0 to w, (f_0, D_0) is an analytic function element at z_0, and (f_0, D_0) can be analytically continued along γ_1 and along γ_2, then are the resulting function elements at w equivalent? Not necessarily, as the following example shows:

Example 7.2.5. Can analytic continuations of a function element at z_0 along different paths from z_0 to w lead to non-equivalent function elements at w?

Solution: Let $z_0 = 1$, $w = -1$ and let the initial function element be $(\log, D_1(1))$ where log is the principle branch of the log function. Clearly this can be continued along the curve γ_1 consisting of the upper half of the unit circle traversed from 1 to -1 and along the curve γ_2 consisting of the lower half of the unit circle traversed from 1 to -1 . However, the resulting function elements at $-1 = \gamma_1(1) = \gamma_2(1)$ differ by $2\pi i$.

The Monodromy Theorem

The next theorem gives conditions under which we can be sure that the analytic continuations of a function along two curves from z_0 to w are necessarily equivalent function elements, in contrast to the preceding example.

Theorem 7.2.6. (Monodromy Theorem) *Let U be a connected open set in $\mathbb{C}$, z_0 and w points of U, and (f_0, D_0) an analytic function element at z_0, with $D_0 \subset U$. Suppose*

(a) (f_0, D_0) can be analytically continued along every curve in U; and

(b) γ_0 and γ_1 are homotopic curves from z_0 to w.

Then the continuations of f along γ_0 and γ_1 are equivalent function elements at w.

Proof. We refer to terminology and notation developed in Section 4.6.

A homotopy from γ_0 to γ_1 determines a continuous one parameter family of curves $\{\gamma_s\}$ from z_0 to w. The continuity of this family means that, given $\epsilon > 0$, there is a $\delta > 0$ such that

$$||\gamma_s - \gamma_r|| < \epsilon \quad \text{whenever} \quad |s - r| < \delta.$$

The analytic function element (f_0, D_0) has an analytic continuation along each of the curves γ_s, by hypothesis. Denote the terminal function element for the continuation along γ_s by ϕ_s. We claim that, for each $r \in [0, 1]$, there is a $\delta > 0$ such that ϕ_s is equivalent to ϕ_r whenever $|s - r| < \delta$.

Let $0 = t_0 < t_1 < \cdots < t_{n+1} = 1$ be a partition and $(f_1, D_1), \cdots, (f_n, D_n)$ a sequence of function elements defining $\phi_r = (f_n, D_n)$ as an analytic continuation of (f_0, D_0) along γ_r. Then

$$\gamma_r([t_j, t_{j+1}]) \subset D_j \quad \text{for} \quad j = 0 \cdots n.$$

For each $j = 0, \cdots, n$, let ϵ_j be the distance from the compact set $\gamma_r([t_j, t_{j+1}])$ to the boundary of the disc D_j. If $||\gamma_s - \gamma_r|| < \epsilon_j$, then it will also be true that $\gamma_s([t_j, t_{j+1}]) \subset D_j$. Thus, if $\epsilon = \min\{\epsilon_0, \cdots, \epsilon_n\}$, and we choose $\delta > 0$ such that

$$||\gamma_s - \gamma_r|| < \epsilon \quad \text{whenever} \quad |s - r| < \delta,$$

then, for each s with $|s - r| < \delta$, the partition $0 = t_0 < t_1 < \cdots < t_{n+1} = 1$ and sequence of function elements $(f_1, D_1), \cdots, (f_n, D_n)$ also defines (f_n, D_n) as an analytic continuation of (f_0, D_0) along γ_s. Since, by the previous theorem, any other continuation of (f_0, D_0) along γ_s is equivalent to this one, we conclude that ϕ_r is equivalent to ϕ_s. This proves that ϕ_s is equivalent to ϕ_r whenever $|r - s| < \delta$.

The remainder of the proof is a standard connectedness argument. We define a function h on $[0, 1]$ by setting $h(s) = 0$ if ϕ_s is equivalent to ϕ_0, and $h(s) = 1$ otherwise. Then, by the result of the previous paragraph, h is a continuous function on $[0, 1]$. If $h(1) = 1$, then, since $h(0) = 0$, the Intermediate Value Theorem implies that h must take on every value between 0 and 1. Since this is not the case, we conclude that $h(1) = 0$. This means that ϕ_1 is equivalent to ϕ_0. □

Example 7.2.7. Consider the function element ϕ consisting of the principal branch of the log function and the disc $D_1(1)$. Set

$$A = \{: |z| > 0, \ -\pi/2 < \arg(z) < 3\pi/2\}.$$

Show that the analytic continuations along any two curves from 1 to -1 in A have equivalent terminal function elements at -1.

Solution: The branch of the log function defined by restricting $\arg(z)$ to lie in the inteval $(-\pi/2, 3\pi/2)$ agrees with the principle branch of the log function on $D_1(1)$. By Theorem 7.2.3, ϕ can be analytically continued along any curve in A. The set A is simply connected and so any two curves in A from 1 to -1 are homotopic. In view of the Monodromy Theorem, continuations of ϕ along two such curves will have equivalent terminal function elements at -1.

The Monodromy Theorem allows us, in the next theorem, to conclude that if a function element can be analytically continued along every curve in an open set U and if the open set is simply connected, then there is an analytic continuation of f to all of U.

Theorem 7.2.8. *Suppose U is a simply connected open set and (f_0, D_0) is an analytic function element at $z_0 \in U$ (with $D_0 \subset U$). If (f_0, D_0) can be analytically continued along every curve in U, then there is a function f which is analytic on U and equal to f_0 on D_0.*

Proof. If $z \in U$, then since any two curves from z_0 to z in U are homotopic in U, the Monodromy Theorem implies that any two terminal elements of analytic continuations of (f_0, D_0) along curves from z_0 to z in U will be equivalent and, hence, will determine the same analytic function in some neighborhood of z. Hence all such analytic continuations determine the same function value at z. We define $f(z)$ to be this value.

Clearly $f_0(z) = f(z)$ on D_0. It remains to prove that f is analytic on U. Let w be a point of U, let γ be a curve in U joining z_0 to w, and let (D_n, f_n) be the terminal function element of some continuation of f along γ. If $z \in D_n$, then (D_n, f_n) is also the terminal element of a continuation of (f_0, D_0) along a curve γ_1 from z_0 to z in U – we simply extend γ to a curve γ_1, ending at z, by joining γ with the line segment from w to z. It follows that $f(z) = f_n(z)$ on all of D_n, not just when $z = w$. Since f_n is analytic on D_n, f is also analytic on D_n. Since w was an arbitrary point of U, we conclude that f is analytic everywhere on U. □

Exercise Set 7.2

1. Prove Theorem 7.2.3.

2. Prove that if log is the principle branch of the log function and D_0 is any disc centered at 1 and not containing 0, then $(\log, D_0)$ can be analytically continued along any curve γ in $\mathbb{C} \setminus \{0\}$.

3. Prove that any analytic continuation of the element $(\log, D_0)$ of the previous exercise, along a curve γ in $\mathbb{C} \setminus \{0\}$, yields a function element (f, D) at $\gamma(1)$ which is some branch of the log function restricted to D. That is, prove that it satisfies $e^{f(z)} = z$ on D.

In the next four exercises U is a domain and (g_0, D_0) is a function element in U which can be analytically continued along a curve γ in U to a function element (g, D) in U.

4. If f is an entire function, prove that $(f \circ g_0, D_0)$ can be analytically continued along γ to $(f \circ g, D)$.

5. If f is an entire function, h is a function analytic on U and $f \circ g_0 = h$ on D_0, then $f \circ g = h$ on D.

6. If a linear differential equation on U, with analytic coeficients, has g_0 for a solution on D_0, then it has g for a solution on D.

7. If u is a harmonic function on U and $u = \mathrm{Re}\,(g_0)$ on D_0, then $u = \mathrm{Re}\,(g)$ on D.

   In the next 3 exercises, use Theorem 7.2.8 to give a proof, different from the one in Theorem 4.6.15, that the indicated property holds for a simply connected domain U. In each case, show there is a solution in a disc D_0 centered at a point $z_0 \in U$ and then show that this solution can be analytically continued along any curve in U.

8. Suppose g is a non-vanishing analytic function on U. Prove that g has an analytic square root on U.

9. Prove that every analytic function on U has an analytic antiderivative.

10. Prove that every real valued harmonic function on U is the real part of an analytic function on U.

7.3 Analytic Covering Maps

If U, V and W are open subsets of $\mathbb{C}$ and

$$h : V \to W \quad \text{and} \quad f : U \to W$$

are analytic maps, then we say that f *lifts* through h if there is an analytic map

$$g : U \to V \quad \text{with} \quad f = h \circ g$$

(See Figure 7.4). If h is a conformal equivalence from V to W, then f trivially lifts through h. In this case, h has an analytic inverse function $h^{-1} : W \to V$ and we can simply set

$$g = h^{-1} \circ f : U \to V.$$

There is another class of maps h for which a kind of lifting result holds. This is the class of *analytic covering maps*.

Definition 7.3.1. An analytic map $h : V \to W$ is called an *analytic covering map* if, for each $w_0 \in W$, there is a neighborhood A of w_0, contained in W, such that $h^{-1}(A)$ is the disjoint union of a collection $\{B_j\}$ of open subsets of V with the property that, for each j, h is a conformal equivalence of B_j onto A.

Example 7.3.2. Show that $z \to e^z$ is an analytic covering map from $\mathbb{C}$ to $\mathbb{C} \setminus \{0\}$.

Solution: The function exp, defined by $\exp(z) = e^z$, is certainly an analytic map of $\mathbb{C}$ onto $\mathbb{C} \setminus \{0\}$. We claim that each disc D in $\mathbb{C} \setminus \{0\}$ has inverse image $\exp^{-1}(D)$ consisting of a disjoint union of open sets on each of which exp is a conformal equivalence onto D.

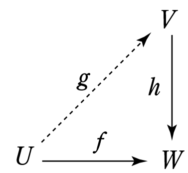

Figure 7.4: Lifting a Map f Through Another Map h.

Given a disc $D \subset \mathbb{C} \setminus \{0\}$, there is a ray from 0 to ∞ disjoint from D (e. g. take the ray from 0 to ∞ which passes through the center of D and rotate it by the angle π). This means that there is an angle θ_0 such that every z for which $e^z \in D$ satisfies

$$\theta_0 + 2n\pi < \mathrm{Im}\,(z) < \theta_0 + 2(n+1)\pi$$

for some integer n. Thus, $\exp^{-1}(D)$ is a disjoint union of open sets

$$B_n = \exp^{-1}(D) \cap \{z : \theta_0 + 2n\pi < \mathrm{Im}\,(z) < \theta_0 + 2(n+1)\pi\}.$$

On each B_n, exp is a conformal equivalence from B_n to D, with inverse function equal to the branch of the log function for which arg takes values between $\theta_0 + 2n\pi$ and $\theta_0 + 2(n+1)\pi$ (see Figure 7.5). Thus, by definition, exp is an analytic covering map.

An analytic covering map $h : V \to W$ is, in particular, a conformal map, since, for each point z_0 of V it is a conformal equivalence of a neighborhood of z_0 onto a neighborhood of $h(z_0)$. However, not every conformal map is a covering map.

Example 7.3.3. Give an example of a conformal map h of a region V onto a region W such that h is not an analytic covering map.

Solution: We use the exponential map again. However, this time we restrict its domain to be the set

$$V = \{z : -\pi < \mathrm{Im}\,(z) < 2\pi\}.$$

The image of this map is still $\mathbb{C} \setminus \{0\}$, and it is clearly a conformal map. However, any disc $D \subset \mathbb{C} \setminus \{0\}$, centered at 1, has an inverse image which is a disjoint union of two open sets – one which lies in the strip $-\pi < \mathrm{Im}\,(z) < \pi$ and one which lies in the strip $\pi < \mathrm{Im}\,(z) < 2\pi$. The map exp is a conformal equivalence of the first of these onto D, but only maps the second one onto the lower half of D (see Figure 7.6). Since this problem persists no matter how small a disc we choose centered at 1, the map exp is not an analytic covering map from V to $\mathbb{C} \setminus \{0\}$.

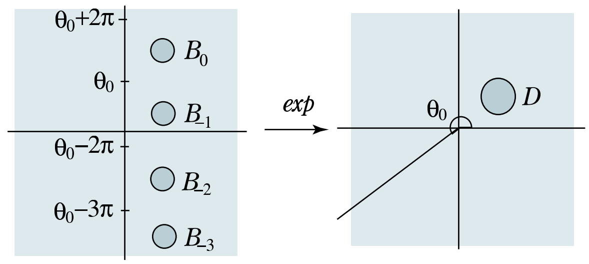

Figure 7.5: The Exponential as a Covering Map.

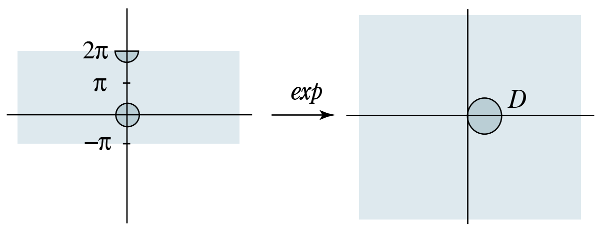

Figure 7.6: A Conformal Map Which is Not a Covering Map.

Lifting Through an Analytic Covering Map

Theorem 7.3.4. *Let $h : V \to W$ be an analytic covering map and U a simply connected domain in $\mathbb{C}$. Then each analytic function $f : U \to W$ can be lifted through h. That is, there is an analytic map $g : U \to V$ with $f = h \circ g$.*

Proof. We fix a $z_0 \in U$ and let $w_0 = f(z_0)$. For this element w_0, we then choose a neighborhood A_0 of w_0 such that $h^{-1}(A_0)$ is a disjoint union of open sets on each of which h is a conformal equivalence onto A_0. We choose one of these, call it B_0, and denote by h_0^{-1} the inverse of $h : B_0 \to A_0$. We then choose an open disc D_0, centered at z_0 and contained in $f^{-1}(A_0)$. On this disc, we define a function g_0 by

$$g_0 = h_0^{-1} \circ f.$$

This function serves to lift f through h, but only on the set D_0. The main work of the proof is to show that the analytic function element (g_0, D_0) can be analytically continued along any curve in U beginning at z_0. If we can do this, then the Monodromy Theorem implies that (g_0, D_0) can be analytically continued to an analytic function defined on all of U (since U is simply connected). This function g serves to lift f through h on all of U, since it satisfies the identity

$$f = h \circ g$$

on D_0 and, hence, on all of U, by Theorem 3.4.2.

So it remains to prove that (g_0, D_0) can be analytically continued along any curve in U beginning at z_0. Let $\gamma : [0,1] \to U$ be such a curve. Let S be the subset of $I = [0,1]$ consisting of those points s such that (g_0, D_0) can be analytically continued along γ as far as s. Clearly, if $s \in S$ and $0 \le s_1 \le s$, then $s_1 \in S$ as well. Also, if $s \in S$, then there is a chain of open discs along γ restricted to $[0,s]$ that serves to analytically continue (g_0, D_0) along this curve to $\gamma(s)$. If $s < 1$, then there is an $s_1 > s$ such that s_1 is also in the last disc in this chain of open discs. This implies that (g_0, D_0) can also be analytically continued along γ restricted to $[0, s_1]$ and, hence, that s_1 is also in S. We conclude that S is either all of $[0,1]$, in which case the proof is complete, or it is a half open interval of the form $[0,r)$ with $r = \sup S \le 1$.

If S is a half open interval $[0,r)$, we let $z_r = \gamma(r)$ and choose a neighborhood A of $f(z_r)$ in V such that $h^{-1}(A)$ is a disjoint union of open sets B_j, each of which is conformally equivalent to A under h. We choose an open disc $D \subset U$ such that $f(D) \subset A$. This is possible because f is continuous. We also choose an $s \in [0,r)$ such that $w = f(\gamma(s))$ belongs to D. This is possible because D is open and $f \circ \gamma$ is continuous.

Because $s \in S$, the function element (g_0, D_0) may be analytically continued along γ to a function element (g_n, D_n) with $\gamma(s) \in D_n$. Since $f = h \circ g_0$ on D_0, the same thing will be true of the function element (g_n, D_n). That is,

$$f = h \circ g_n \quad \text{on} \quad D_n,$$

In particular, $w = f(\gamma(s)) = h(g_n(\gamma(s)))$. This means that $g_n(\gamma(s))$ belongs to $h^{-1}(A)$ and, hence, to exactly one of the sets B_j, call it B_k. We define a new function element (g_{n+1}, D_{n+1}) by choosing $D_{n+1} = D$, setting g_{n+1} equal to the composition of $f : D \to A$ with an inverse function for $h : B_k \to A$. This new function element certainly satisfies

$$f = h \circ g_{n+1} \quad \text{on} \quad D_{n+1},$$

but does it agree with g_n on $D_n \cap D_{n+1}$? It does because, not only is the point $g_n(\gamma(s))$ in B_k, but all of $g_n(D_n \cap D_{n+1})$ is contained in B_k, otherwise, the connected open set $D_n \cap D_{n+1}$ would be separated by $g_n^{-1}(B_k)$ and the union of the sets $g_n^{-1}(B_j)$ for $j \ne k$. But this means that the two inverse functions for h used in the definitions of g_n and g_{n+1} agree on $f(D_n \cap D_{n+1})$ which implies that g_n and g_{n+1} agree on $D_n \cap D_{n+1}$. But now, since $\gamma(r)$ is an interior point of D_{n+1}, this implies that (g_0, D_0) can be analytically continued to $\gamma(r)$ and, hence, that $r \in S$. This contradicts the assumption that S has the form $[0,r)$. The only other possiblility is that $S = [0,1]$ and this means that (g_0, D_0) can be analytically continued along γ. □

Exercise Set 7.3

1. Is the function $h(z) = z^2$ an analytic covering map of $\mathbb{C}$ onto $\mathbb{C}$? Is it an analytic covering map of $\mathbb{C} \setminus \{0\}$ onto $\mathbb{C} \setminus \{0\}$? Justify your answers.

2. Give a proof, using Theorem 7.3.4, that if U is a simply connected open set and f is a non-vanishing analytic function on U, then there is an analytic logarithm of f – that is, an analytic function g on U such that $f = e^g$.

3. Prove that if $h : V \to W$ is an analytic covering map from a connected open set U to a simply connected open set V, then h is a conformal equivalence.

4. Prove that if $h : V \to W$ is an analytic covering map and U is any simply connected open subset of W, then $h^{-1}(U)$ is a disjoint union of open sets on each of which h is a conformal equivalence onto U.

5. Is there an analytic covering map from the unit disc to $\mathbb{C} \setminus \{0\}$? Justify your answer.

6. Prove that if $h : V \to W$ is an analytic cover and $\gamma : [0, 1] \to W$ is a curve, then γ can be lifted through h in the sense that there is an curve $\lambda : [0, 1] \to V$ such that $\gamma = h \circ \lambda$.

7. Let $p(z)$ be a polynomial of degree n in z and let w be a point of $\mathbb{C}$. Prove that the polynomial $p(z) - w$ fails to have n distinct roots (that is, it has a repeated root) if and only if $w = p(z_0)$ for some point z_0 at which $p'(z_0) = 0$.

8. Let $p(z)$ be a polynomial of degree n in z and let S be the set of all points $w \in C$ at which the polynomial $p(z) - w$ fails to have n distinct roots. If $W = \mathbb{C} \setminus S$ and $V = p^{-1}(W)$, prove that $p : V \to W$ is an analytic covering map. Hint: use the Inverse Mapping Theorem and the preceding exercise.

9. Prove that the polynomial $p(z) = z^3 - 3z$ is an analytic covering map from $\mathbb{C} \setminus \{-1, 1\}$ to $\mathbb{C} \setminus \{-2, 2\}$.

10. Use the previous exercise to prove that if g is an analytic function on a simply connected domain U and if g does not take on the values -2 and 2, then there is an analytic function f on U which satisfies $f^3(z) - 3f(z) = g(z)$ on U.

7.4 The Picard Theorems

The Little Picard Theorem states that a non-constant entire function takes on every value in $\mathbb{C}$ except possibly one. This is a vast generalization of Liouville's Theorem and an impressive application of the analytic continuation techniques we have been developing in this chapter. The Big Picard Theorem states that an analytic function with an essential sigularity at z_0 takes on every complex value but one infinitely often in every neighborhood of z_0.

The strategy for proving the Picard Theorems is to construct an analytic covering map h from the unit disc to the plane with two points removed. The

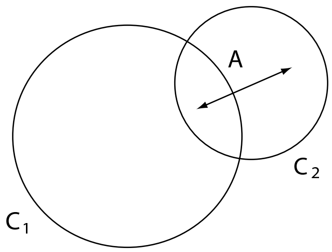

Figure 7.7: Reflection Through an Arc of a Perpendicular Circle

Little Picard Theorem follows easily from this since any analytic function from the plane to the plane with two points removed will lift through this covering map to an analytic function from the plane to the unit disc. By Liouville's Theorem the only such functions are constants. The Big Picard Theorem uses the existence of h along with Montel's Theorem. The construction of h relies heavily on analytic continuation by reflection through a circle (Example 7.1.8 – the Schwarz Reflection Principle for circles). We begin by proving a few simple facts about reflection through a circle.

Reflection Through a Circle

Suppose C_1 and C_2 are two circles in the plane that intersect one another in two points. If the tangents to the two circles are perpendicular at each of these two points, then we will say that the circles meet at right angles (Figure 7.7). Actually, it turns out that if the tangents are perpendicular at one of the points of intersection, then they are also perpendicular at the other point of intersection (Exercise 7.4.1).

Theorem 7.4.1. *If circles C_1 and C_2 meet at right angles, and A is the arc of C_1 that lies inside C_2, then reflection through A maps C_2 onto itself.*

Proof. There is a linear fractional transformation h that maps C_1 onto $\mathbb{R} \cup \{\infty\}$ and maps the inside of C_1 onto the upper half plane. We can choose which point on C_1 is sent to ∞ and we choose it so that it is not on $\overline{A}$; in particular, it is not one of the points where C_1 and C_2 intersect.

The fact that linear fractional transformations are conformal maps (angle preserving) and the fact that C_1 and C_2 meet at right angles means that the image of C_2 under h meets the real line at two points and it has vertical tangents at these points. The image of a circle under a linear fractional transformation is either a line or a circle, and so $h(C_2)$ must be a circle which meets the real axis at right angles. This implies that the real axis passes through a diameter of $h(C_2)$. Reflection through the x–axis, therefore, maps the inside of the circle $h(C_2)$ onto itself.

Now reflection through the arc A may be described as h composed with reflection through the horizontal diameter of $h(C_2)$ followed by h^{-1}. Why?

Well, this certainly describes a reflection through A defined on the inside of C_2 and, by the uniqueness part of Theorem 7.1.6, this is the only reflection through A.

Since reflection through the x-axis takes the inside of $h(C_2)$ onto itself, reflection through A takes the inside of C_2 onto itself. □

If two circles meet at right angles, then we will say that the circular arc consisting of that part of the first circle which is inside the second circle meets the second circle at right angles.

Theorem 7.4.2. *With C_1, C_2 as above, the reflection through C_1 takes any other circle C, which meets C_2 at right angles, to another circle which meets C_2 at right angles or to a line through the center of C_2.*

Proof. A reflection through a circle is a linear fractional transformation followed by conjugation (Exercise 7.1.8). It therefore takes a circle C to either a circle or a line.

The reflection through C_1 takes the inside of C_2 onto itself. It follows that the image of C under this reflection still meets C_2 in two points.

A conformal map preserves angles, while conjugation reverses the orientation of any angle – that is, it takes an angle to its negative. It follows that a reflection, which is a conformal map followed by conjugation, is also an angle reversing map. This implies that if two curves have perpendicular tangents at a point of intersection, then the same will be true of their images under a reflection. Therefore, reflection through C_1 takes C to a circle or line which also meets C_2 at right angles. The only way the image of C could be a line is if it passes through the center of C_2. Thus, the reflection of C through C_1 is another circle which meets C_2 at right angles or is a line through the center of C_2. □

An Analytic Covering Map

The three points -1, $1/2 + i\sqrt{3}/2$, $1/2 - i\sqrt{3}/2$ are equidistant points on the unit circle C. According to Exercise 7.42, we can join each pair of these points with an arc of a circle that meets C at right angles. These three arcs then bound an open curvilinear triange V_0 (see Figure 7.8).

The open set V_0 is clearly simply connected. By the Riemann Mapping Theorem, there is a conformal equivalence $h : V_0 \to H$, where H is the upper half plane. By Theorem 6.5.2, the map h extends to a continous map of the closure of V_0 to the closure of H in S^2, which takes the boundary of V_0 to the boundary of H in S^2. By composing with a linear fractional transformation, if necessary, we may choose this map in such a way that it takes the points $-1, 1/2 - i\sqrt{3}/2, 1/2 + i\sqrt{3}/2$ to the points $\infty, 0, 1$.

The function h is real valued on each of the three "edges" of V_0. By the circle version of the Schwarz Reflection Principle, h can be continued by reflection across each of these edges. This results in h being defined and analytic in the set V_1 described in Figure 7.9. Note that each of the three new cells on which h is defined in this way is contained in the unit disc D (Theorem 7.4.1) and is

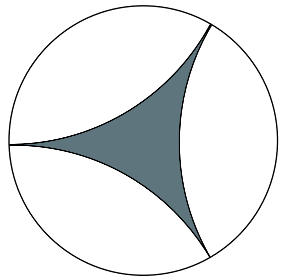

Figure 7.8: The Curvilinear Triangle V_0

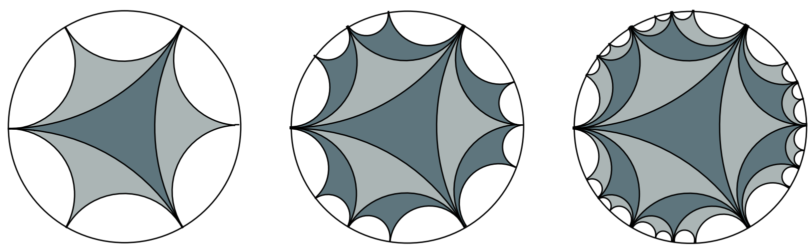

Figure 7.9: The Sets V_1, V_2 and V_3 of Theorem 7.4.3

bounded by three circles that meet the unit circle C in right angles (Theorem 7.4.2).

Note also that, while h maps V_0 to the upper half plane, it maps each of the new cells of V_1 to the lower half plane, since each is a reflection of V_0 through one of its edges.

We can now analytically continue h to a still larger domain V_2 by reflecting across each of the circular arcs that bound V_1. This results in h being defined in the set V_2 represented in Figure 7.9. Note that h now maps each of the new cells in V_2 into the upper half plane. We can clearly continue this process by induction to create an increasing sequence $\{V_n\}$ of open sets to which h may be analytically continued. The union of these open sets is D (Exercise 7.4.5), and so the result is an analytic function h which maps the open unit disc D onto a subset of the plane which contains the upper and lower open half planes and the intervals $(-\infty, 0)$, $(0, 1)$, and $(1, \infty)$ on the real line. The points ∞, 0, and 1 are not in the image because every triangular cell that occurs in the above construction has all of its vertices on the unit circle. Thus, h is an analytic map of D onto $\mathbb{C} \setminus \{0, 1\}$.

Theorem 7.4.3. *The analytic map $h : D \to \mathbb{C} \setminus \{0, 1\}$, described above, is an analytic covering map.*

Proof. Think of the disc D as being partitioned into light cells and dark cells, separated by arcs of circles, as in Figure 7.9. The dark cells are mapped into the upper half plane by h and the light cells are mapped into the lower half plane. Each arc separating two of these cells is mapped to one of the intervals $I_1 = (-\infty, 0)$, $I_2 = (0, 1)$, $I_3 = (1, \infty)$.

For $j = 1, 2, 3$, let P_j be the union of the open upper and lower half planes and the open interval I_j. Then each P_j is an open set consisting of S^2 with a closed arc removed.

Let Δ be any open disc in $\mathbb{C} \setminus \{0, 1\}$. Then Δ lies entirely inside P_j for at least one j. We fix a j for which this is true. The inverse image of P_j under h consists of the union of all the light and dark open cells together with some of the arcs separating them – those arcs which h maps to I_j. The collection of these arcs is paiwise disjoint and each one of them separates exactly one light cell from one dark cell. The union of the arc and the two cells it separates is an open set on which h is a conformal equivalence onto P_j. No two distinct open sets of this form can overlap since, if they did, the overlap would have to be a cell with boundary containing two arcs mapping to I_j under h. Let $\{U_k\}_{k=1}^{\infty}$ be the collection of open sets of this form. This is a pairwise disjoint collection of open sets with union equal to $h^{-1}(P_j)$. Also,

$$h^{-1}(\Delta) \subset \bigcup_k U_k,$$

and so $h^{-1}(\Delta)$ is the disjoint union of the open sets $B_k = h^{-1}(\Delta) \cap U_k$. The map h is a conformal equivalence of B_k onto Δ for each k. It follows that $h : D \to \mathbb{C} \setminus \{0, 1\}$ is an analytic covering map. □

The Little Picard Theorem

Theorem 7.4.4. (Little Picard Theorem) *If f is an entire function and there are two distinct points points in the plane that are not in the image of f, then f is constant.*

Proof. If f does not take on the values z_0 and z_1, then the function

$$\frac{f(z) - z_0}{z_1 - z_0}$$

does not take on the values 0 and 1. Thus, we may as well assume that f itself has this property. Then f is an analytic function from the simply connected set $\mathbb{C}$ to $\mathbb{C} \setminus \{0, 1\}$. Since the map $h : D \to \mathbb{C} \setminus \{0, 1\}$ of the previous theorem is an analytic covering map, Theorem 7.3.4 implies that f may be lifted through h to a analytic function $g : \mathbb{C} \to D$ such that

$$f = h \circ g.$$

However, this means that g is a bounded entire function and, hence, is a constant, by Liouville's Theorem. This, of course, forces f to also be a constant. □

The Big Picard Theorem

The proof of the big Picard Theorem depends on the following theorem. Recall the definition of a normal family (Definition 6.4.1).

Theorem 7.4.5. *Let U be a connected subset of the plane. Then the set of analytic functions on U with values in the set $\mathbb{C} \setminus \{0, 1\}$ is a normal family.*

Proof. We will prove the theorem in the case where U is an open disc Δ. The proof that the theorem for general U follows from this special case will be left to the exercises.

Let Δ be any open disc and let z_0 be its center. Let $\mathcal{F}$ denote the set of analytic function on U with values in $\mathbb{C} \setminus \{0, 1\}$. Let $\{f_n\}$ be a sequence in $\mathcal{F}$. We will show that this sequence converges uniformly to ∞ on compact subsets of Δ or it has a subsequence which converges uniformly on compact subsets of Δ to a function analytic on Δ.

Suppose $\{f_n\}$ does not converge uniformly to ∞ on compact subsets of Δ. Then there is a disc Δ_1 with the same center and with $\overline{\Delta_1} \subset \Delta$ and an $R > 0$ such that $f_n(\Delta_1)$ has elements with modulus less than R for infinitely many n. This means that we can choose a subsequence of $\{f_n\}$ which itself has no subsequence converging uniformly to ∞ on Δ_1. Without loss of generality, we may replace $\{f_n\}$ by this subsequence. That is, we may assume that $\{f_n\}$ has no subsequence converging uniformly to ∞ on Δ_1.

Let $h : D \to \mathbb{C} \setminus \{0, 1\}$ be the analytic covering map of Theorem 7.4.3. Then, since Δ is simply connected, Theorem 7.3.4 implies each f_n can be lifted through h to an analytic function $g_n : \Delta \to D$ with $f_n = h \circ g_n$. Furthermore, since the set V_1 in Figure 7.9 maps onto $\mathbb{C} \setminus \{0, 1\}$ under h, we can choose the g_n in such a way that $g_n(z_0) \in V_1$.

The sequence $\{g_n\}$ is a uniformly bounded sequence of analytic functions on Δ and, hence, by Montel's Theorem (Theorem 6.4.2), it has a subsequence which converges uniformly on compact subsets of Δ to an analytic function g.

The function g has its image in $\overline{D}$. However, unless g is constant, this image must be an open set (by the Open Mapping Theorem – Theorem 4.5.8) and, hence, must be contained in D. In this case $f = h \circ g$ is an analytic function on Δ with values in $\mathbb{C} \setminus \{0, 1\}$ and is the uniform limit on compact subsets of Δ of a subsequence of $\{f_n\}$.

On the other hand, if g is constant, then this constant must lie in the closure of V_1, since each $g_n(z_0)$ lies in V_1. This means the constant is either in D or is one of the six points of $\overline{V}_1$ which lie on the unit circle. This means that $f = h \circ g$ is has constant value 0, 1, or ∞.

Thus, each sequence $\{f_n\}$ has a subsequence which converges uniformly on compact subsets of Δ to ∞ or to an analytic function. Since there is no subsequence converging to ∞, it has a subsequence converging uniformly on compact subsets to an analytic function. This completes the proof of the theorem in the case where U is a disc. The general case follows from Exercise 7.4.11.

□

Theorem 7.4.6. (Big Picard Theorem) *Let f be a function which is analytic in a neighborhood U of z_0 except at z_0 itself, where it has an essential singularity. Then f takes on every value but one infinitely often in every neighborhood of z_0.*

Proof. We may as well assume that $z_0 = 0$. If f does not have the property stated in the conclusion of the theorem, then there is a disc $D_r(0)$ on which f fails to take on at least two complex values. We may assume these values are 0 and 1, since otherwise we may compose f with a linear fractional transformation which takes the two values missed by f to 0 and 1.

Thus, we may assume f is an analytic function from $D_r(0) \setminus \{0\}$ to $\mathbb{C} \setminus \{0, 1\}$ with an essential singularity at 0. We define a sequence of functions with the same properties by setting

$$f_n(z) = f(z/n) \quad \text{for} \quad n = 1, 2, \cdots .$$

By the previous theorem, this sequence converges uniformly on compact subsets of $D_r(0)$ to ∞ or it has a subsequence converging uniformly on compact subsets of $D_r(0)$ to an analytic function. In the first case, $1/f$ is bounded and, hence, has a removable singularity at 0. This means it extends to an analytic function with a zero of some finite order at 0. This is impossible, since it implies that f has a pole at 0 rather than an essential singularity. The second case implies that f itself is bounded and, hence, has a removable singularity at 0. This also violates the hypothesis that f has an essential singularity at 0. Thus, our assumption that f misses two values in some disc centered at 0 has led to a contradiction. This completes the proof. □

Exercise Set 7.4

1. Prove that if two circles intersect in two points and their tangents are perpendicular at one point of intersection, then the tangents are also perpendicular at the other point of intersection.

2. Prove that, given two distinct points on a circle C_1, there is a unique circle C_2 which meets C_1 at right angles at these two points.

3. Prove that if C_r and C_R are circles of radius r and R, respectively, which are tangent at a point, then the reflection of C_R through C_r is a circle of radius
$$\frac{rR}{2R + r}$$
if neither circle is inside the other and is
$$\frac{rR}{2R - r}$$
if C_r is inside C_R.

4. The set V_n in the construction preceding Theorem 7.4.3 has boundary consisting of a set of circular arcs of different sizes with endpoints on the boundary of the unit disc. Prove by induction that the largest of these circular arcs has radius $(2n+1)^{-1}R_0$, where R_0 is the radius of each of the three circular arcs comprising the boundary of V_0. Hint: use the calculation of the preceding exercise.

5. Prove that the union of the open sets V_n, described in the paragraph preceding Theorem 7.4.3, is the unit disc D. Hint: use the result of the preceding exercise.

6. Prove that if f and $1/f$ are both entire functions, then the image of f is exactly $\mathbb{C} \setminus \{0\}$.

7. Prove that if f is a non-constant entire function and $b^2 \neq 4ac$, then the function
$$g(z) = af^2(z) + bf(z) + c$$
must have a zero.

8. Is there a non-constant analytic function from the plane with one point removed to the plane with two points removed? Justify your answer.

9. If U is a simply connected open set, show that the set of all analytic functions on U with values in $\mathbb{C} \setminus \{0, 1\}$ is a normal family (see Section 6.4).

10. Suppose $r < 1$, $R < 1$ and h is a conformal equivalence from the annulus $A_R = \{z : R < |z| < 1\}$ to the annulus $A_r = \{z : r < |z| < 1\}$. Suppose also that h extends to a continuous map from $\overline{A}_R$ to $\overline{A}_r$ which takes the unit circle to itself and the disc of radius R to the disc of radius r. Prove that h can be continued to a conformal equivalence of the unit disc onto the unit disc that takes 0 to 0. Conclude from this that $R = r$. Hint: do the analytic continuation through a series of steps involving reflection through a circle, as in Exercise 7.1.10; the first of these steps involves reflecting the annulus A_R through the circle of radius R.

11. Prove that if $\mathcal{F}$ is a family of analytic functions on a connected open set U and if for each disc Δ in U the family of restrictions of elements of $\mathcal{F}$ to Δ is a normal family, then $\mathcal{F}$ itself is a normal family.

12. Prove that an entire function which is not a polynomial takes on every complex value but one infinitely often.