COMPLEX ANALYSIS NOTES

TANG-KAI LEE

ABSTRACT. These are my notes for the honors complex variables course taught in fall 2025.

CONTENTS

1. **Introduction**

Complex analysis is the study of functions of a complex variable, a subject that reveals a deep and beautiful interplay between algebra, analysis, and geometry. At first glance, one might think of complex numbers as merely an extension of the real number system, but the introduction of a complex variable fundamentally changes the nature of calculus. Functions that are differentiable in the complex sense—holomorphic functions—exhibit remarkable properties that set them apart from their real-variable counterparts.

Unlike real differentiable functions, holomorphic functions are automatically infinitely differentiable and analytic, meaning they can be expressed as convergent power series wherever they are defined. This rigidity leads to powerful results such as Cauchy's integral formula, which establishes profound connections between differentiation, integration, and series expansions. These properties make complex analysis not just a refinement of real analysis but an entirely new and far-reaching mathematical framework.

Beyond its theoretical elegance, complex analysis has deep applications across mathematics and physics. It plays a crucial role in number theory, algebraic geometry, dynamical systems, and representation theory, and it provides essential tools for solving problems in mathematical physics, engineering, and even fluid dynamics. Many results that seem difficult or inaccessible in real analysis become transparent through the lens of complex analysis.

In this course, we will develop the fundamental tools of complex analysis and explore their consequences. We will see why complex analysis is not only one of the most elegant subjects in mathematics but also one of the most powerful. To illustrate this, consider one of the most important functions in math, the **Riemann zeta function** $\zeta(s)$. It is defined to be

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s}$$

for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$. At first, this sum looks like just a function of real numbers when s is real. But it turns out that by analytically continuing it, we get a function defined on almost the entire complex plane. This reveals deep insights into prime numbers. One way this connection emerges is through the relation

$$\zeta(s) = \prod_{p \text{ is a prime}} \frac{1}{1 - p^{-s}}$$

for $s \in \mathbb{C}$ with $\mathrm{Re}(s) > 1$. This is how prime numbers are encoded inside $\zeta(s)$. Using the tools of complex analysis, we can study the analytic behavior of this function and eventually arrive at profound results—such as the Prime Number Theorem, which tells us that the number of primes less than x, denoted $\pi(x)$, grows approximately like $x/\log x$.

Another beautiful application of complex analysis is its usage to classify "surfaces," that is, two-dimensional manifolds. Different from the general case, it is possible to equip a "complex structure" to any surfaces, so tools related to single-variable complex analysis can help in the classification program. We will return to both of these ideas (prime number theorem and classification of surfaces) after developing the necessary background.

In the notes, the notation $^{\text{ex}}$ means that the conclusion follows from a relatively straightforward argument, and will be left as a practice exercise. Feel free to contact me/come to my office hour if you have any question about them (or about any other things).

1.1. **Preliminaries.**

1.1.1. *Complex numbers.* We recall the notations of the common sets of numbers:

- $\mathbb{N}$: the set of positive integers.
- $\mathbb{Z}$: the set of integers.
- $\mathbb{Q}$: the set of rational numbers.
- $\mathbb{R}$: the set of real numbers.
- $\mathbb{C}$: the set of complex numbers.

Each of the constructions of a new number system from the previous one involves a new extension of some important properties.

- $\mathbb{N}$: the set $\{1, 2, 3, \cdots\}$ that is closed under addition and multiplication.
- $\mathbb{Z}$: the smallest additive "group" that contains $\mathbb{N}$.
- $\mathbb{Q}$: the smallest "field" that contains $\mathbb{Z}$.
- $\mathbb{R}$: the smallest complete metric space that contains $\mathbb{Q}$.
- $\mathbb{C}$: the smallest algebraically closed field that contains $\mathbb{R}$.

Another aspect is about solving polynomial equations. Each extension allows one to solve more polynomial equations in the extended system.

- $\mathbb{N}$ contains solutions to $x - 1 = 0$.
- $\mathbb{Z}$ contains solutions to $x + 1 = 0$.
- $\mathbb{Q}$ contains solutions to $2x - 1 = 0$.
- $\mathbb{R}$ contains solutions to $2x^2 - 1 = 0$.
- $\mathbb{C}$ contains solutions to $2x^2 + 1 = 0$.

We won't talk about the detailed constructions of any of them. Here, we briefly mention a few equivalent ways to construct $\mathbb{C}$ from $\mathbb{R}$.

(1) Field extension: $\mathbb{C} := \mathbb{R}(i) := \{a + bi : a, b \in \mathbb{R}\}$ where i satisfies $i^2 = -1$.

(2) Quotient field: $\mathbb{C} := \mathbb{R}[x]/\left\langle x^2 + 1\right\rangle$, where we let i be the image of the image x under the quotient map $\mathbb{R}[x] \to \mathbb{R}[x]/\left\langle x^2 + 1\right\rangle$.

(3) Algebraic closure: $\mathbb{C} :=$ the algebraic closure of $\mathbb{R}$.

Further discussions about these constructions will be too algebraic. For our purpose, we recall some basic operations that we will need in the course.

(1) Given a complex number $z = a + bi$ with $a, b \in \mathbb{R}$, $\mathrm{Re}z := a$ is the real part of z, and $\mathrm{Im}z := b$ is the imaginary part of z.

(2) The basic operations are based on the commutative, associative, and distributive laws, with the rule $i^2 = -1$. For example,

$$(a + bi) \cdot (c + di) = (ac - bd) + (ad + bc)i.$$

(3) The complex conjugation is a map $\overline{(\cdot)} \colon \mathbb{C} \to \mathbb{C}$ defined by $a + bi \mapsto a - bi$ for $a, b \in \mathbb{R}$. It is then direct to check

$$\mathrm{Re}z = \frac{z + \overline{z}}{2} \text{ and } \mathrm{Im}z = \frac{z - \overline{z}}{2i}$$

for any $z \in \mathbb{C}$.

(4) The length or the absolute value of a complex number z is defined by $|z| := \sqrt{z\overline{z}}$. To be explicit, if $z = a + bi$ where $a, b \in \mathbb{R}$, then $|z| = \sqrt{a^2 + b^2}$. It is its distance to the origin on the complex plane.

(5) With the distance given by $d(z, w) := |z - w|$, the complex plane $\mathbb{C}$ is a "metric space." In particular, it satisfies the triangle inequality:

$$|z + w| \leq |z| + |w|$$

for any $z, w \in \mathbb{C}$.

We recall that a metric space is a set X with a distance function $d \colon X \times X \to \mathbb{R}_{\geq 0}$[1] satisfying the following three properties for any $x, y, z \in X$.

(1) $d(x, y) = 0$ if and only if $x = y$.

(2) $d(x, y) = d(y, z)$.

(3) $d(x, z) \leq d(x, y) + d(y, z)$.

Based on the metric space structure, many analytic tools can be applied and will be crucial in the course.

1.1.2. *Review on analysis and topology.* Most of the content in this section applies to an arbitrary metric space, but we focus on the case of $\mathbb{C}$, where the distance is given by the absolute value of the difference. Much of this material should be familiar, and we encourage students to recall the concepts by working through a few **examples**, though few will be provided here since this section serves mainly as a recap.

Recall that the metric space structure gives a natural topological structure on $\mathbb{C}$. To talk about the topology in detail, we first mention that we use $B_r(z_0)$ to be the open ball centered at z_0 with radius r. That is,

$$B_r(z_0) := \{z \in \mathbb{C} : |z - z_0| < r\}.$$

We let $\overline{B}_r(z_0)$ be the corresponding closed ball, including those points on the boundary of $B_r(z_0)$. A subset U in $\mathbb{C}$ is then called **open** if for any point $z_0 \in U$, there exists $r > 0$ such that $B_r(z_0) \subseteq U$.

[1] We occasionally use subscriptions to denote different subset of the number sets. For example, $\mathbb{R}_{\geq 0} := \{r \in \mathbb{R} : r \geq 0\}$, and $\mathbb{Z}_{<0} := \{n \in \mathbb{Z} : n < 0\}$.

In general, such a point is called an **interior point** of U (even when U is not an open set, which means that not every point is an interior point). A subset K is called **closed** if its complement $\mathbb{C} \setminus K$ is an open set in $\mathbb{C}$. We leave all the other topological notions to the first time when we use them, including compactness, connectivity, and more.

Since we want to use calculus to study complex numbers and functions of them, we first recall the notions of taking limits. We will use limits of both sequences and functions.

Given a sequence a_n ($n \in \mathbb{N}$) of complex numbers, we say it converges to a complex number L, denoted by

$$\lim_{n\to\infty} a_n = L,$$

if given any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $|a_n - L| < \varepsilon$ for $n \geq N$.

Given a function $f\colon U \to \mathbb{C}$ defined on an open set U and a point $z_0 \in U$, we say f converges to a complex number L as z tend to z_0, denoted by

$$\lim_{z\to z_0} f(z) = L,$$

if given any ε, there exists $\delta > 0$ such that $|f(z) - L| < \varepsilon$ for $z \in B_\delta(z_0) \setminus \{z_0\}$. Note that the limit of f as $z \to z_0$ does not depend on $f(z_0)$. If the function $f\colon U \to \mathbb{C}$ satisfies

$$\lim_{z\to z_0} f(z) = f(z_0),$$

then we say f is **continuous** at the point z_0. We say f is a continuous function on U if f is continuous at every point $z \in U$.

Based on these two definitions, it is natural to think about the convergence of a given sequence of functions. Thus, we assume f_n is a sequence of complex-valued functions defined on U.

(1) We say f_n converges **pointwisely** to a function $f\colon U \to \mathbb{C}$ if given any $z \in U$ and $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that $|f_n(z) - f(z)| < \varepsilon$ for $n \geq N$.

(2) We say f_n converges **uniformly** to a function $f\colon U \to \mathbb{C}$ if given any $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for any $z \in U$, $|f_n(z) - f(z)| < \varepsilon$ for $n \geq N$.

The notion of uniform convergence arose after a mistake had been found in Cauchy's original proof of the statement that a pointwise limit of a sequence of continuous functions is still continuous. The way how Cauchy corrected his proof is the introduction of the notion of uniform convergence (if written in modern languages). We record this and another consequence of uniform convergence here, and will use them without quoting in the rest of the lecture.

Proposition 1.1. Let $f_n\colon U \to \mathbb{C}$ be a sequence of functions. Suppose f_n uniformly converges to a function $f\colon U \to \mathbb{C}$.

(1) If each f_n is continuous, then f is also continuous.

(2) If each f_n is (Riemann) integrable[2], then f is also (Riemann) integrable, and

$$\lim_{n\to\infty} \int_U f_n \, dz = \int_U f \, dz. \tag{1.2}$$

[2]Here, we mean that for each f_n, $\mathrm{Re} f_n$ and $\mathrm{Im} f_n$ are integrable on U, as a subset of $\mathbb{R}^2$. The integral in (1.2) is also interpreted in this sense.

A more useful localized notion of uniform convergence will be used later many times. We will mention it after it first appears in Section 2.1.

1.2. **Holomorphic functions: basics and examples.**

1.2.1. *Cauchy–Riemann equations.* We want to do calculus on $\mathbb{C}$, so we care about functions that are differentiable in the complex variable z.

Definition 1.3. Let U be an open set[3] in $\mathbb{C}$ and let $f\colon U \to \mathbb{C}$ be a function. We say f is **holomorphic** at a point $z_0 \in U$ if the limit

$$\lim_{z\to z_0} \frac{f(z)-f(z_0)}{z-z_0} = \lim_{h\to 0} \frac{f(z_0+h)-f(z_0)}{h} \tag{1.4}$$

exists. When it does, the limit is denoted by $f'(z_0)$.

We remark that the limit is taken by viewing $\mathbb{C}$ as a metric space using the distance function given by the absolute value $|z| := \sqrt{z\overline{z}}$. Recall that using this, we say the limit (1.4) exists and is $L \in \mathbb{C}$ if given any $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\left| \frac{f(z)-f(z_0)}{z-z_0} - L \right| \leq \varepsilon$$

for $z \in B_\delta(z_0) \setminus \{z_0\}$. This limit is the **complex derivative** of f at z_0, and as mentioned, denoted by $f'(z_0)$. We say f is holomorphic on U if f is holomorphic at any point in U. In this case, we get a function $f'\colon U \to \mathbb{C}$. Most of the properties we learn for real derivatives are true, and we mention a few of them.

(1) If f is holomorphic on U, then f is, in particular, continuous.

(2) If f and g are holomorphic, then $f \pm g$ and $f \cdot g$ are holomorphic, with
$$(f \pm g)' = f' \pm g' \text{ and}$$
$$(f \cdot g)' = f' \cdot g + f \cdot g'.$$

(3) If g is holomorphic and $g \neq 0$ at a point, then at the point, $1/g$ is holomorphic with
$$\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}.$$

(4) If f and g are holomorphic, then at a point z where $f \circ g$ is well-defined, $f \circ g$ is holomorphic and
$$(f \circ g)'(z) = f'\left(g(z)\right) \cdot g'(z).$$

However, holomorphic functions also behave completely differently from real differentiable functions. They are much more “rigid.” To illustrate this, suppose a holomorphic function f is defined in an open set U. For $z \in U$, we write $z = x + iy$ and

$$f(z) = f(x,y) = u(x,y) + iv(x,y) \tag{1.5}$$

where u and v are real-valued functions. We will see that f being holomorphic is much stronger then u and v being differentiable. We state this observation as a lemma.

[3] In the rest of the notes, U will be an open set in $\mathbb{C}$ unless otherwise stated.

Lemma 1.6. Suppose $f\colon U \to \mathbb{C}$ is holomorphic and is written in the form (1.5). Then

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \text{ and } \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}. \tag{1.7}$$

These equations are called the **Cauchy–Riemann equations**.

Proof. The lemma follows from the existence of the limit (1.4). In fact, at a point $z_0 = x_0 + iy_0 \in U$, we can approximate z_0 from the real direction and get

$$\begin{aligned} f'(z_0) = \lim_{r\to 0} \frac{f(z_0 + r) - f(z_0)}{r} &= \lim_{r\to 0} \left(\frac{u(x_0 + r, y_0) - u(x_0, y_0)}{r} + i \frac{v(x_0 + r, y_0) - v(x_0, y_0)}{r} \right) \\ &= \frac{\partial u}{\partial x}(x_0, y_0) + i \frac{\partial v}{\partial x}(x_0, y_0). \end{aligned}$$

We can also approximate z_0 from the imaginary direction and get

$$\begin{aligned} f'(z_0) = \lim_{r\to 0} \frac{f(z_0 + ri) - f(z_0)}{ri} &= \lim_{r\to 0} \left(\frac{u(x_0, y_0 + r) - u(x_0, y_0)}{ri} + i \frac{v(x_0, y_0 + r) - v(x_0, y_0)}{ri} \right) \\ &= -i \frac{\partial u}{\partial y}(x_0, y_0) + \frac{\partial v}{\partial y}(x_0, y_0) \end{aligned}$$

where we use $1/i = -i$. Comparing these two equations, we get the conclusion. □

Based on Lemma 1.6, we know that holomorphic functions form a much more restricted class of differentiable functions. Let's see some consequences of the Cauchy–Riemann equation.

(1) If u and v are C^2 and satisfy (1.7), then in particular, we have

$$\frac{\partial^2 u}{\partial x^2} = \frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 v}{\partial y \partial x} = -\frac{\partial^2 u}{\partial y^2},$$

so u is a **harmonic function**, in the sense that

$$\Delta u := \frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0.$$

One can similarly see that v is also harmonic. Because of this property, v is sometimes called a **harmonic conjugate** of u. Harmonic functions play an important role in the study of partial differential equations, and later we will see that many good properties of them also apply to holomorphic functions. Here, we just mention the converse about the existence of harmonic conjugates as an evidence for the connection between holomorphic and harmonic functions. We remark that, however, harmonic conjugates do not always exist on a general domain (cf. Corollary 4.5).

Lemma 1.8. Let u be a harmonic function on the unit disc B_1. Then there exists a unique harmonic conjugate of u up to a constant. That is, there exists a harmonic function v such that u and v satisfy (1.7), and if $\widetilde{v}$ is another harmonic function with the same property, then $v = \widetilde{v} + c$ where c is a constant.

Proof. If such a v exists, then for any $(x_0, y_0) \in B_1$, the fundamental theorem of calculus and (1.7) imply

$$\begin{aligned} v(x_0, y_0) &= v(0) + \int_0^{x_0} \frac{\partial v}{\partial x}(x,0)dx + \int_0^{y_0} \frac{\partial v}{\partial y}(x_0, y)dy \\ &= v(0) - \int_0^{x_0} \frac{\partial u}{\partial y}(x,0)dx + \int_0^{y_0} \frac{\partial u}{\partial x}(x_0, y)dy. \end{aligned}$$

The same equation holds for $\widetilde{v}$, so taking their difference, we get

$$v(x_0, y_0) - \widetilde{v}(x_0, y_0) = v(0) - \widetilde{v}(0),$$

so the uniqueness follows.

For existence, inspired by the equation above, we define

$$v(x_0, y_0) := -\int_0^{x_0} \frac{\partial u}{\partial y}(x,0)dx + \int_0^{y_0} \frac{\partial u}{\partial x}(x_0, y)dy. \tag{1.9}$$

for any $(x_0, y_0) \in B_1$. By the fundamental theorem of calculus, differentiating (1.9) in y gives

$$\frac{\partial v}{\partial y}(x_0, y_0) = \frac{\partial u}{\partial x}(x_0, y_0),$$

noting that the first term in (1.9) does not depend on y_0. For the other equation in (1.7), differentiating (1.9) in x (noting that u being C^2 allows us to do so) leads to

$$\begin{aligned} \frac{\partial v}{\partial x}(x_0, y_0) &= -\frac{\partial u}{\partial y}(x_0, 0) + \int_0^{y_0} \frac{\partial^2 u}{\partial x^2}(x_0, y_0)dy \\ &= -\frac{\partial u}{\partial y}(x_0, 0) - \int_0^{y_0} \frac{\partial^2 u}{\partial y^2}(x_0, y_0)dy \\ &= -\frac{\partial u}{\partial y}(x_0, y_0) \end{aligned}$$

where we use the harmonicity of u and the fundamental theorem of calculus again (and again). Thus, (1.7) follows. The equations and the C^2-regularity of u imply that v is also C^2, and hence harmonic. $\square$

We go back to the discussion of the Cauchy–Riemann equation.

(2) If u and v are C^1 and satisfy (1.7), then $f := u + iv$ is holomorphic.[4] This is the converse of Lemma 1.6, and this can be seen by checking an equivalent condition of the existence of the limit (1.4). That is, f is holomorphic at z_0 if and only if there exists $L \in \mathbb{C}$ such that

$$f(z_0 + h) - f(z_0) = Lh + o(|h|)$$

as $h \to 0$. This implies that L is the first-order approximation of f at z_0 and means that $f'(z_0) = L$. Here, we recall the little o-notation, and we say that two functions $g_1(z)$ and $g_2(z)$ satisfies $g_1 = o(g_2)$ as $z \to a$ for some $a \in [-\infty, \infty]$ if $|g_1(z)|/|g_2(z)| \to 0$ as $z \to a$. By writing out the first-order approximations of u and v, one can see that their derivatives together determine the derivative of of f.[ex]

[4] Note that the a priori regularity here is important. Otherwise, there are counterexamples. See Assignment 1.

(3) Cauchy–Riemann equations also allow us to introduce the first-order operator

$$\partial_z := \frac{1}{2}\left(\frac{\partial}{\partial x} - i\frac{\partial}{\partial y}\right) \text{ and}$$
$$\partial_{\overline{z}} := \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right).$$

With these notations, one can calculate that

$$\partial_{\overline{z}} f = \frac{1}{2}\left(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y}\right)(u+iv) = \frac{1}{2}\left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y}\right) + \frac{1}{2}i\left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y}\right).$$

Thus, a holomorphic function satisfies (1.7) if and only if $\partial_{\overline{z}} f = 0$. Intuitively, this means that if we write

$$f(z,\overline{z}) = f(x,y) = f\left(\frac{z+\overline{z}}{2}, \frac{z-\overline{z}}{2i}\right),$$

then it is "independent" of $\overline{z}$.

Example 1.10. We end this section by mentioning a few examples.

(1) A **polynomial** $f(z) = a_0 + a_1 z + \cdots + a_n z^n$ is a holomorphic function. One can check that, as in the real case, if $f(z) = z^n$, then $f'(z) = nz^{n-1}$. Polynomials are **entire** functions, functions that are holomorphic on the whole $\mathbb{C}$.

(2) The conjugate function $f(z) = \overline{z}$ is not holomorphic. This can be seen by either the definition or by the Cauchy–Riemann equation. In the languages introduced in (3), this is because $\partial_{\overline{z}} f = 1 \neq 0$. A problem in Assignment 1 will ask you to check whether a function in z and $\overline{z}$ is holomorphic based on the definition (of complex differentiability).

(3) A **rational function** $f(z) = P(z)/Q(z)$ is holomorphic at which Q does not vanish, where P and Q are polynomials. The zeros of Q are called the **poles** of f. The study of the behavior of f near its poles is crucial. Common examples are when P and Q are polynomials of degree one, and we will see some discussion in Assignment 1. Some properties of polynomial and rational functions are discussed in Section 1.2.2.

(4) The exponential function

$$e^z := \sum_{n\geq 0} \frac{1}{n!} z^n = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \cdots$$

is a holomorphic function. It is direct to verify it is a well-defined function for all $z \in \mathbb{C}$. By the binomial formula, as in the real case, one can verify that

$$e^z \cdot e^w = e^{z+w}$$

for any $z, w \in \mathbb{C}$. We will see that, in general, a **power series**

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

is holomorphic at which the series converges. This will be in Section 1.2.3. This will also be a main topic in Section 2.1, as we will see all holomorphic functions are locally power series.

(5) We mention a few examples of harmonic conjugates.

$$\begin{aligned} u(x,y) = x,\ v(x,y) = y+1; \\ u(x,y) = y,\ v(x,y) = -x; \\ u(x,y) = x^2 - y^2,\ v(x,y) = 2xy; \\ u(x,y) = e^x \cos y,\ v(x,y) = e^x \sin y. \end{aligned}$$

One can try to find out holomorphic functions they represent.

1.2.2. *Polynomial and rational functions.* We will have a brief discussion about polynomial and rational functions, as they are the first non-trivial class of holomorphic functions we can work on. They are also, respectively, finite cases for power series and Laurent series, which are important expansions for holomorphic functions and meromorphic functions, and some notions we introduce here will be generalized later as we proceed.

If $P(z)$ is a polynomial of degree n, then the fundamental theorem of algebra[5] implies that P has exactly n roots, counted with multiplicities. If a zero z_1 appears exactly n_1 times, that is, we can write $P(z) = P_1(z) \cdot (z - z_1)^{n_1}$ for another polynomial P_1 with $P_1(z_1) \neq 0$, then we say z_1 is a zero of P of **order** n_1. Here, we mention an interesting results about zeros of P and its derivatives.

Lemma 1.11 (Gauss–Lucas theorem)**.** Let P be a non-constant polynomial. Then any zero of P' lies in the **convex hall** of the set of all zeros of P.

The convex hall of a subset $X \subseteq \mathbb{C}$ is the smallest convex set that contains X. In practice, one can take the intersection of all half-planes that contain X to find the convex hall of X.

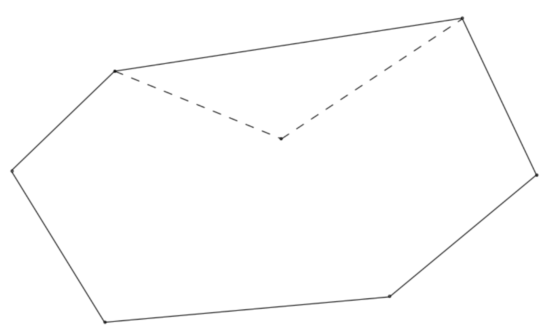

FIGURE 1. When there is a finite set, its convex hall is a convex polygon that contains all the segments between any two points of the set.

Proof. It is sufficient to show that if all the zeros of P lie in a half-plane, then all the zeros of P' also do. Thus, we assume all the zeros $z_1, \cdots, z_n$ of P lie in the half-plane defined by $\mathrm{Im} z \geq 0$. Take a point z_0 that is not in the half-plane defined by $\mathrm{Im} z \geq 0$, that is,

$$\mathrm{Im} z_0 < 0. \tag{1.12}$$

[5]The theorem can be proven based on calculus, so we take it for granted here. Later, we will see that it can be proven based on the tools we develop for holomorphic functions, cf. Section 2.1.3.

In particular, we know $\operatorname{Im}(z_0 - z_k) < 0$ for all $k = 1, \cdots, n$ and $P(z_0) \neq 0$, so

$$\frac{P'(z_0)}{P(z_0)} = \frac{1}{z_0 - z_1} + \cdots + \frac{1}{z_0 - z_n}$$

has positive imaginary part. Thus, $P'(z_0) \neq 0$. □

Next, we talk about rational functions. For a rational function $R(z) = P(z)/Q(z)$, where we assume P and Q do not have common zeros, we know that $R(z)$ is holomorphic away from its poles. One can also show that at a pole p_1 of R, we have $\lim_{z \to p_1} R(z) = \infty$, which we will rigorously define next. It is thus sometimes convenient to introduce the space with a "formal infinity." That is, we consider

$$\hat{\mathbb{C}} := \mathbb{C} \cup \{\infty\}.$$

We only formally add it to the complex plane so that it matches all the calculations we have here.[6] For ∞-valued point, we define $R(z_0) = \infty$ if

$$\lim_{z \to z_0} \frac{1}{R(z)} = 0.$$

For ∞-input, we define

$$R(\infty) := \lim_{w \to 0} R\left(\frac{1}{w}\right).$$

When $z_0 \in \mathbb{C}$, $R(z_0) = \infty$ if and only if $Q(z_0) = 0$. If z_0 is a zero of Q of order n_0, then we say z_0 is a pole of R of **order** n_0.

We can write down $R(\infty)$ very explicitly. Suppose $\deg P = n$ and $\deg Q = m$, and

$$R(z) = \frac{P(z)}{Q(z)} = \frac{a_n z^n + \cdots + a_1 z + a_0}{b_m z^m + \cdots + b_1 z + b_0}.$$

Then we know that

$$R\left(\frac{1}{w}\right) = \frac{a_n w^{-n} + \cdots + a_1 w^{-1} + a_0}{b_m w^{-m} + \cdots + b_1 w^{-1} + b_0} = w^{m-n} \cdot \frac{a_n + \cdots + a_1 w^{n-1} + a_0 w^n}{b_m + \cdots + b_1 w^{m-1} + b_0 w^m},$$

so we can conclude that

$$R(\infty) = \lim_{w \to 0} R\left(\frac{1}{w}\right) = \begin{cases} 0 & \text{if } m > n \text{ (so } R \text{ has a zero of order } m-n) \\ \frac{a_n}{b_m} & \text{if } m = n \\ \infty & \text{if } m < n \text{ (so } R \text{ has a pole of order } n-m) \end{cases}.$$

The reason why this information is useful is as follows. We let $d := \max\{m, n\}$, called the **degree** of R. Then for any $y_0 \in \hat{\mathbb{C}}$, we have

$$\#\left\{z \in \hat{\mathbb{C}} : R(z) = y_0\right\} = d.$$

For example, when $y_0 = \infty$, we know that $R(z) = \infty$ if either $z \in \mathbb{C}$ and $Q(z) = 0$, and we get m solutions, or $z = \infty$, which happens only when $m < n$, and we get $n - m$ solutions. Thus, there are $\max\{m, n\} = d$ solutions. The case for a general y_0 is similar.

[6] It does have some geometric meaning, and as a Riemann surface, $\hat{\mathbb{C}}$ is an important model space, cf. Section 5.5.

1.2.3. *Power series.* We've encountered some special power series in calculus. For example, we know

$$\frac{1}{1-r} = \sum_{n\geq 0} r^n \tag{1.13}$$

when $|r| < 1$. This is also an example of writing a rational function in the form of a power series.

Remark 1.14. A direct calculation (the same as the real case) shows that (1.13) also works for complex r with $|r| < 1$. That is,

$$\frac{1}{1-z} = \sum_{n\geq 0} z^n. \tag{1.15}$$

Note, however, that the left hand side of (1.15) makes sense for any $z \neq 1$, for which it even defined a holomorphic function. Therefore, one can view $1/(1-z)$ as a "natural" extension of the power series to a much larger domain. This process is called an **analytic continuation**, and we will encounter this a few times later.

We will first see that, in general, when a power series converges, it defines a holomorphic function and its derivative can be derived easily. We recall the notation of an open ball

$$B_r(z_0) := \{z \in \mathbb{C} : |z - z_0| < r\},$$

and we sometimes write $B_r := B_r(0)$ when the center is the origin. We let $\overline{B}_r(z_0)$ be the corresponding closed ball, which is the closure of $B_r(z_0)$.

Theorem 1.16 (Abel's theorem)**.** Given a sequence $a_n \in \mathbb{C}$, consider the power series $\sum_{n=0}^{\infty} a_n z^n$. The limit

$$R := \left(\limsup_{n\to\infty} |a_n|^{1/n}\right)^{-1} \in [0, \infty] \tag{1.17}$$

is called the **radius of convergence** of the power series and satisfies the following properties.

(1) The power series converges absolutely when $z \in B_R$ and diverges when $z \in \mathbb{C} \setminus \overline{B}_R$.

(2) For $\rho < R$, the convergence is uniform on $\overline{B}_\rho$.

(3) The function

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

is a holomorphic function in B_R, and its derivative is given by

$$f'(z) = \sum_{n=0}^{\infty} n a_n z^{n-1}.$$

The formula (1.17) is Hadamard's formula. It is direct to see why the radius should be of the form, since we hope for, at least, $|a_n z^n| \leq 1$ for large enough n, compared with the geometric series

(1.13). It turns out that it is also a sufficient condition for the convergence. We also recall that given a sequence b_n, its limit superior is defined by

$$\limsup_{n\to\infty} b_n := \lim_{n\to\infty}\left(\sup_{j\geq n} b_j\right).$$

Proof of Theorem 1.16. We prove (1) and (2) first. If $z \in B_R$, then we can take $\rho \in (|z|, R)$. This means that $1/\rho > 1/R$, so there exists $N \in \mathbb{N}$ such that for $n \geq N$, $|a_n|^{1/n} < 1/\rho$, which implies[ex]

$$|a_n z^n| < \left(\frac{|z|}{\rho}\right)^n.$$

Thus, the convergence follows from the root test. For the uniform convergence, by taking $\overline{\rho} \in (\rho, R)$, we can similarly show that for $z \in \overline{B}_\rho$,

$$|a_n z^n| < \left(\frac{\rho}{\overline{\rho}}\right)^n$$

for large enough n, and the uniform convergence follows from Weierstrass' M-test.[ex]

On the other hand, if $|z| > R$, we can take $\rho \in (R, |z|)$ and since $1/\rho < 1/R$, we can find a subsequence a_{n_j} of a_n such that $|a_{n_j}|^{1/n_j} > 1/\rho$ for all j, which means

$$|a_{n_j} z^{n_j}| > \left(\frac{|z|}{\rho}\right)^{n_j}.$$

The divergence then also follows from the root test.

Finally, we prove (3). First, from the limit $\lim_{n\to\infty} n^{1/n} = 1$, it is clear that the power series $\sum_{n=0}^{\infty} n a_n z^{n-1}$ has the same radius of convergence. Then at a point $z_0 \in B_R$, for a nearby point $z \in B_R \setminus \{z_0\}$, we analyze the difference

$$\frac{f(z)-f(z_0)}{z-z_0} - \sum_{n=0}^{\infty} n a_n z_0^{n-1} = \sum_{n=0}^{\infty}\left(a_n \frac{z^n - z_0^n}{z-z_0} - n a_n z_0^{n-1}\right).$$

Using the relation $\frac{z^n-z_0^n}{z-z_0} = \sum_{j=0}^{n-1} z^j z_0^{n-1-j}$ twice, if we take $\rho \in (|z_0|, R)$, then for $z \in B_\rho$, we can estimate

$$\begin{aligned}
\left|a_n \frac{z^n - z_0^n}{z-z_0} - n a_n z_0^{n-1}\right| &= |a_n|\left|\sum_{j=0}^{n-1}\left(z^j z_0^{n-1-j} - z_0^{n-1}\right)\right| \\
&= |a_n|\left|\sum_{j=0}^{n-1} z_0^{n-1-j}\left(z^j - z_0^j\right)\right| \\
&= |a_n|\left|\sum_{j=1}^{n-1} z_0^{n-1-j}\cdot(z-z_0)\sum_{k=0}^{j-1} z^k z_0^{j-1-k}\right| \leq |a_n|\cdot|z-z_0|\cdot\frac{1}{2}n(n-1)\rho^{n-2}.
\end{aligned}$$

We only need to do this for $n \geq 2$, since the term vanishes when $n = 0$ and 1. Thus, we can estimate

$$\left|\frac{f(z)-f(z_0)}{z-z_0} - \sum_{n=0}^{\infty} n a_n z_0^{n-1}\right| \leq \frac{1}{2}|z-z_0|\sum_{n=2}^{\infty} n(n-1)|a_n|\rho^{n-2}.$$

The sum converges since it comes from the power series expansion of $f''(z)$. Thus, we see that the difference tends to zero as $z \to z_0$, and the conclusion follows. □

We give two remarks here. First, we talk about power series centered at the origin. One can do the same thing for power series centered at any other point, like

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

and all the arguments are the same. Second, note that we do not talk about the case when $|z|$ is exactly the radius of convergence. That is because many things can happen in this case. One can look at the series

$$\sum_{n=0}^{\infty} z^n, \sum_{n=0}^{\infty} \frac{1}{n} z^n, \text{ and } \sum_{n=0}^{\infty} \frac{1}{n^2} z^n$$

when $|z| = 1$ and see whether they converge or not.

As a consequence of Theorem 1.16, we know that given a power series

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

with R being its radius of convergence, f is a holomorphic function on B_R, and

$$f'(z) = \sum_{n=0}^{\infty} n a_n z^{n-1}$$

is also a holomorphic function on B_R. Inductively, we can show that $f^{(n)}(z)$ is holomorphic on B_R for all $n \in \mathbb{N}$. In particular, f is a smooth function, though being analytic is a much more rigid condition than smoothness. We note that we can also integrate and get a "primitive" of f. That is,

$$F(z) = \sum_{n=0}^{\infty} \frac{a_n}{n+1} z^{n+1}$$

is holomorphic on B_R and satisfies $F' = f$.

A function that can locally be defined by a power series is called an **analytic function**. From above, we know that a complex analytic function is a holomorphic function. We will later see that any holomorphic function is also an analytic function, so basically all of them are locally represented by power series. That will then show a fundamental difference between holomorphic functions and real differentiable functions. Many interesting properties of holomorphic functions can, in fact, be derived now using power series, but we delay this until actually showing that holomorphic functions are analytic, and at that time we will have more tools to get these consequences directly. See Section 2.1.3.

At this point, using power series, we can extend many real-valued functions we know to holomorphic function.

Example 1.18. One can check that the following functions all have radii of convergence being ∞ and hence they are all holomorphic on the whole $\mathbb{C}$.

(1) $\sin z = \frac{e^{iz}-e^{-iz}}{2i} = z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 - \cdots$.
(2) $\cos z = \frac{e^{iz}+e^{-iz}}{2} = 1 - \frac{1}{2!}z^2 + \frac{1}{4!}z^4 - \cdots$.
(3) One can check directly, as in the real case, that $(e^z)' = e^z$, $(\sin z)' = \cos z$, and $(\cos z)' = -\sin z$.

2. Cahchy–Goursat theory: integral representation of holomorphic functions

We will show that a holomorphic function can be locally represented by an integral given by the Cauchy integral formula. As a consequence, we will see that a holomorphic function is analytic. This will then tell us a lot of good properties of holomorphic functions.

2.1. Local results and consequences.

2.1.1. *Line integrals and indices.* We first talk about how to integrate a holomorphic function along a reasonable curve. In this course, we mostly talk about **piecewise smooth curves**, meaning that they are continuous curves and are smooth away from finitely many times.

Definition 2.1. We say $\gamma\colon [a,b] \to \mathbb{C}$ is a **path** if it is continuous and piecewise smooth, in the sense that there exist finitely many times $t_n \in [a,b]$ for $n = 0, \cdots, \ell$ such that $a = t_0 < t_1 < \cdots < t_\ell$ and $\gamma|_{[t_{n-1},t_n]}$ is a smooth curve. If the curve further satisfies $\gamma(a) = \gamma(b)$, then we say γ is a **closed path**.

For simplicity, we start from a smooth curve $\gamma\colon [a,b] \to U \subseteq \mathbb{C}$. We will sometimes abuse the notation and denote the image of γ, that is, $\gamma([a,b])$, also by γ. The image is definitely independent of the parametrization of γ, so sometimes we have to be careful. If there is a continuous function $f\colon U \to \mathbb{C}$, then the **line integral** of f along γ is defined to be

$$\int_\gamma f dz := \int_a^b f(\gamma(t)) \cdot \gamma'(t) dt.$$

Note that this definition does on depend on the parametrization of γ. That is, given any strictly increasing smooth map $\varphi\colon [c,d] \to [a,b]$ between two intervals, which leads to a reparametrization $\widetilde{\gamma} := \gamma \circ \varphi$ of γ, then one can check that

$$\int_{\gamma\circ\varphi} f dz = \int_\gamma f dz$$

based on the chain rule.[ex] Thus, this justifies the usage of γ to denote its image since the line integral only depends on the curve and its orientation. We remark that most of the curves in this course will be compact.

When γ is a piecewise smooth curve (or a path), this means that we can find a parametrization $\gamma\colon [a,b] \to U$ and finitely many times $t_n \in [a,b]$ for $n = 0, \cdots, \ell$ such that $a = t_0 < t_1 < \cdots < t_\ell$ and $\gamma|_{[t_{n-1},t_n]}$ is a smooth curve for all $n = 1, \cdots, \ell$. We can then define

$$\int_\gamma f dz := \sum_{i=1}^{\ell} \int_{\gamma|_{[t_{n-1},t_n]}} f dz,$$

which is again well-defined in the sense that it is invariant under reparametrization.

We remark that the line integral of a function along a curve can be defined for much more general functions and curves, as Riemann integrals make sense for functions with less regularity. In most situations, we deal with paths. We discuss a general notion of the length of a curve in Assignment 2, and one can similarly discuss a general notion of line integrals. We end the introduction by mentioning a common property that we will use, the triangle inequality

$$\left|\int_\gamma f dz\right| \leq \int_\gamma |f|\,|dz|. \tag{2.2}$$

Here, the line integral of a real-valued scalar function with respect to arc length is defined by the same way. That is, if $h\colon U \to \mathbb{R}$ is continuous and $\gamma\colon [a,b] \to U$ is smooth, then

$$\int_\gamma h|dz| := \int_a^b h\left(\gamma(t)\right) \cdot \left|\gamma'(t)\right| dt,$$

which can be similarly generalized to the path case. A proof of (2.2) will be discussed in Assignment 2.

Example 2.3. We see some examples of line integrals.

(1) We do a specific example. Suppose γ is the part of the unit circle $\partial B_1(0)$ from 1 to i counterclockwise and let $f(z) = \mathrm{Re} z$. We can parametrize γ by $\gamma(z) = e^{it}$ for $t \in [0, \pi/2]$. Then

$$\int_\gamma f dz = \int_0^{\pi/2} \left(\mathrm{Re}\, e^{it}\right) \cdot ie^{it} dt = \int_0^{\pi/2} \cos t \cdot (-\sin t + i\cos t)\, dt = -\frac{1}{2} + \frac{\pi}{4} i.$$

(2) Given a path $\gamma\colon [a,b] \to U \subseteq \mathbb{C}$, it comes with an orientation. We can look at the same curve with the other orientation, denoted by

$$\begin{aligned} -\gamma\colon [-b,-a] &\to U \\ t &\mapsto \gamma(-t). \end{aligned}$$

The images of γ and $-\gamma$ are the same, and one can derive that

$$\int_{-\gamma} f dz = -\int_\gamma f dz$$

for any continuous function f on U.

(3) Given two paths $\gamma_1\colon [a,b] \to U$ and $\gamma_2\colon [b,c] \to U$, one can consider their concatenation

$$\begin{aligned} \gamma_1 + \gamma_2\colon [a,c] &\to U \\ t &\mapsto \begin{cases} \gamma_1(t) & \text{if } t \in [a,b] \\ \gamma_2(t) & \text{if } t \in [b,c] \end{cases} \end{aligned}$$

if $\gamma_1(b) = \gamma_2(b)$. The image of $\gamma_1 + \gamma_2$ is the union of those of γ_1 and γ_2, and one can show that

$$\int_{\gamma_1+\gamma_2} f dz = \int_{\gamma_1} f dz + \int_{\gamma_2} f dz$$

for any continuous function f on U.

The notations in (2) and (3) look natural, and we will see that they are compatible with a more general operation acting on the space of "cycles," introduced in Section 2.2.

The next example is separated, since it tells us that line integrals always provide ways to construct more holomorphic functions.

Proposition 2.4. Let γ be a path in an open set U and $\varphi\colon U \to \mathbb{C}$ be a continuous function. Then the function

$$f(w) := \int_\gamma \frac{\varphi(z)}{z-w} dz$$

is analytic (and hence holomorphic) on $U \setminus \gamma$.

Proof. Fix an open ball $B_r(a) \Subset U \setminus \gamma$. Then for $z \in \gamma$ and $w \in B_r(a)$, we can write

$$\frac{\varphi(z)}{z-w} = \frac{\varphi(z)}{z-a} \cdot \frac{z-a}{z-w} = \frac{\varphi(z)}{z-a} \cdot \frac{1}{1-\frac{w-a}{z-a}}.$$

Note that the conditions $z \in \gamma$ and $w \in B_r(a)$ then imply

$$\left| \frac{w-a}{z-a} \right| \leq \frac{|w-a|}{r} < 1.$$

Thus, we get a power series expansion

$$\frac{\varphi(z)}{z-w} = \sum_{n=0}^{\infty} \frac{\varphi(z)}{(z-a)^{n+1}} (w-a)^n.$$

As we know in Abel's theorem, the convergence is uniform in any compact set in $U \setminus \gamma$, so we can take the integral and switch it with the sum, deriving

$$\int_\gamma \frac{\varphi(z)}{z-w} dz = \sum_{n=0}^{\infty} \left(\int_\gamma \frac{\varphi(z)}{(z-a)^{n+1}} dz \right) (w-a)^n.$$

This proves that f is locally a power series, and hence holomorphic. $\square$

Note that the assumption on φ is very weak, only requiring its continuity. Thus, the proposition does produce many holomorphic functions. We remark that the proof extends to the case when φ is a complex-valued function on a more general space when the integration makes sense. That will need more advanced measure theory.

Since this local notion of uniform convergence will be used a lot in the rest of the course, we formally give it a definition here.

Definition 2.5. Let f_n's and f be functions on an open set U. We say f_n converges **locally uniformly** to f if given any compact set $K \subseteq U$, f_n converges uniformly to f on K.

We then look at a special example for Proposition 2.4. It is called the **index** function, and will be important later.

Corollary 2.6. Let γ be a closed path in an open set U. Then the function

$$\operatorname{Ind}_\gamma(w) := \frac{1}{2\pi i}\int_\gamma \frac{dz}{z-w}$$

is integer-valued for $w \in U \setminus \gamma$. In particular, it is constant on each component of $U \setminus \gamma$.

This index function "counts" the number of times that γ winds around w. It is thus sometimes called the winding number. A particular case is when γ is the boundary of a small ball, and we have

$$\frac{1}{2\pi i}\int_{\partial B_\varepsilon(a)} \frac{dz}{z-w} = \begin{cases} 1 & \text{if } w \in B_r(a) \\ 0 & \text{if } w \in \mathbb{C} \setminus \overline{B}_r(a) \end{cases}$$

where $\partial B_\varepsilon(a)$ is oriented counterclockwise.

Proof. We let $\gamma\colon [a,b] \to U$ be a parametrization, and we have

$$\operatorname{Ind}_\gamma(w) = \frac{1}{2\pi i}\int_a^b \frac{\gamma'(t)}{\gamma(t)-w}dt$$

for $w \in U \setminus \gamma$. We then look at the function

$$h(s) := \int_a^s \frac{\gamma'(t)}{\gamma(t)-w}dt.$$

Suppose γ is smooth on $[a,b] \setminus S$ where S is a finite set. Then we know that for $s \in [a,b] \setminus S$,

$$h'(s) = \frac{\gamma'(s)}{\gamma(s)-w}.$$

This implies

$$0 = \gamma' - (\gamma - w)h' = e^h\left((\gamma-w)e^{-h}\right)'$$

on $[a,b] \setminus S$. Thus, $(\gamma - w)e^{-h}$ is constant on $[a,b] \setminus S$ and hence on $[a,b]$. We then, using $h(a) = 0$, get

$$(\gamma(b) - w)\, e^{-h(b)} = (\gamma(a) - w)\, e^{-h(a)} = \gamma(a) - w.$$

Then by $\gamma(a) = \gamma(b)$, we derive $e^{h(b)} = 1$, so $\frac{1}{2\pi i}h(b) \in \mathbb{Z}$. The second conclusion then follows from this and Proposition 2.4 since analytic functions are continuous. $\square$

We end this section by mentioning a common situation in which it is clear how the index function behaves. We say γ is a **simple closed path** if it is a closed path that does not have self-intersection. That is, it is of the form $\gamma\colon [a,b] \to \mathbb{C}$ with $\gamma(a) = \gamma(b)$ and $\gamma|_{[a,b)}$ being an injective map. For such a curve, it divides the plane $\mathbb{C}$ into exactly two parts. This is a special case of the Jordan curve theorem, which in fact holds in a much more general situation for continuous curves. Some related discussion will be in Assignment 2.

Theorem 2.7 (Jordan curve theorem, simple version)**.** Let γ be a simple closed path. Then $\mathbb{C} \setminus \gamma$ has exactly two components. One of them is a bounded open set on which $\operatorname{Ind}_\gamma = \pm 1$ (depending on the orientation of γ); the other is an unbounded component on which $\operatorname{Ind}_\gamma = 0$.

From now on, when the curve γ we are integrating along is a Jordan curve (that is, a simple closed curve), we will always implicitly assume that it is oriented so that on the bounded component it encloses, the index function Ind_γ is $+1$. For example, when we write

$$\int_{\partial B_1(0)} f dz,$$

unless otherwise stated, we implicitly assume the circle $\partial B_1(0)$ is oriented counterclockwise.

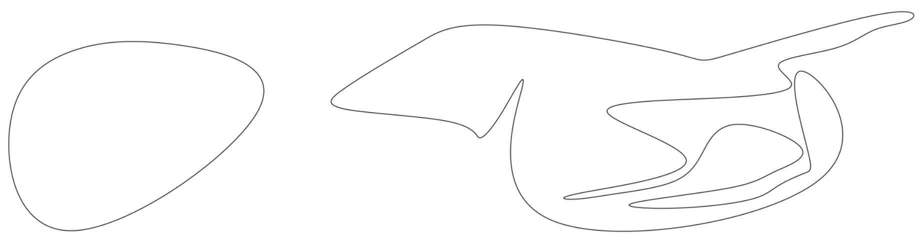

FIGURE 2. Usually it is intuitive to believe the content of the Jordan curve theorem, but even for a smooth curve, sometimes it is not completely trivial, as there are many more complicated curves than the right one that one can find on Google.

2.1.2. *Cauchy's integral formula.* The important property we are going to investigate is the dependence of the line integral on the curve. This reminds us of Green's theorem in calculus. Similar arguments tell us that if we can "integrate" the function, then the line integral should only depend on the endpoints of the curve. We know that we can integrate a convergence power series, and this is in general true if a continuous function f has a **primitive**, in the sense that $f(z) = F'(z)$ for some differentiable function F. We remark that, unlike the one-dimensional case, this is usually not a trivial condition.

Proposition 2.8. Let $f\colon U \to \mathbb{C}$ be a continuous function and γ be a closed path in U. If f has a primitive on U, then $\int_\gamma f dz = 0$.

The proof will tell us that the line integral only depends on the endpoints of the curves, and hence vanishes when the endpoints are the same.

Proof. We do the case when γ is a smooth curve and let $\gamma\colon [a,b] \to U$ be a parametrization. The piecewise smooth case is similar.

Let F be a primitive of f on U. Then by definition, we have

$$\begin{aligned}\int_\gamma f dz = \int_a^b f\left(\gamma(t)\right)\gamma'(t)dt = &\int_a^b F'\left(\gamma(t)\right)\gamma'(t)dt \\ = &\int_a^b \left(F\circ\gamma\right)'(t)dt = F\left(\gamma(b)\right) - F\left(\gamma(a)\right) = 0,\end{aligned}$$

where we use the chain rule and the fundamental theorem of calculus. □

This observation then directly applies to those holomorphic functions defined by power series. Our goal in this section is to show that the property is in fact correct for all holomorphic functions. To achieve, we first deal with the following special but important case when the curve is triangular.

Theorem 2.9 (Goursat's theorem)**.** Let $f\colon U \to \mathbb{C}$ be a holomorphic function. If R is a (solid) closed triangle in U, then

$$\int_{\partial R} f dz = 0,$$

where ∂R is the boundary of R and we view it as a piecewise smooth curve.

We remark that the orientation of a curve is usually important for a line integral, but in this case it doesn't matter since the integral is always zero. In the proof, all triangles are closed and solid (so not just the boundary).

Proof. We write $R = R_1^0$, and we divide it into the four triangles R_1^1, R_2^1, R_3^1, and R_4^1 of the same shapes and similar to R^0. We can inductively do this once we have R_k^i $(k = 1, \cdots, 4^i)$. For each $i \in \mathbb{N}$, we let D^i be the diameter of R_1^i (and hence R_k^i for all k). Note that $D^i = D^1/2^i \to 0$ as $i \to \infty$. We orient all R_k^i's counterclockwise.

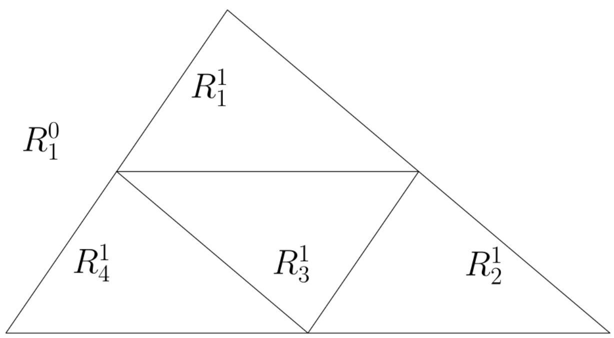

FIGURE 3. One can divide a triangle into four similar smaller triangles.

The orientation implies the equality

$$\int_{\partial R_1^0} f dz = \sum_{k=1}^{4} \int_{\partial R_k^1} f dz.$$

Thus, we know that at least one of the k's satisfies

$$\left| \int_{\partial R_1^0} f dz \right| \leq 4 \left| \int_{\partial R_k^1} f dz \right|.$$

After reordering, we may assume this $k = 1$. We can inductively, for each i, choose R_1^{i+1} such that

$$\left| \int_{\partial R_1^i} f dz \right| \leq 4 \left| \int_{\partial R_1^{i+1}} f dz \right|$$

for all $i \in \mathbb{N}$. In particular, we have

$$\left| \int_{\partial R} f dz \right| \leq 4^i \left| \int_{\partial R_1^i} f dz \right|. \tag{2.10}$$

Next, we look at the sequence $R_1^1 \supseteq R_1^2 \supseteq \cdots \supseteq R_1^i \supseteq \cdots$ of decreasing triangles. From the compactness of each R_1^i, the completeness of $\mathbb{R}$, and the fact that $D^i \to 0$ as $i \to \infty$, we know that the intersection

$$\bigcap_{i \in \mathbb{N}} R_1^i$$

consists of exactly one point. After translation, we may assume this point is the origin, that is,

$$\bigcap_{i \in \mathbb{N}} R_1^i = \{0\}\,. \tag{2.11}$$

We now look at the first order approximation of f at the origin. That is, we can write

$$f(z) = f(0) + f'(0)z + h(z)$$

where $h(z) = o(z)$ as $z \to 0$. Recall that this means $h(z)/z \to 0$ as $z \to 0$. Using the fact that $f(0) + f'(0)z$ has a primitive (since it is linear), Proposition 2.8 then implies

$$\int_{\partial R_1^i} f dz = \int_{\partial R_1^i} \big(f(0) + f'(z)z\big)\, dz + \int_{\partial R_1^i} h dz = \int_{\partial R_1^i} h dz. \tag{2.12}$$

We start the analysis now. Given $\varepsilon > 0$, using $h(z) = o(z)$ as $z \to 0$, we can take $\rho > 0$ such that

$$|h(z)| \le \varepsilon |z| \tag{2.13}$$

for all $z \in B_\rho$. Using (2.11), the fact that R_1^i's are all similar, and the fact that $D^i \to 0$, we can find $N \in \mathbb{N}$ such that $R_1^i \subseteq B_\rho$ for $i \ge N$. Combining this with (2.12) and (2.13), we obtain that for $i \ge N$,

$$\left| \int_{\partial R_1^i} f dz \right| \le \int_{\partial R_1^i} |h|\, |dz| \le 4D^i \cdot \varepsilon D^i$$

where we use the triangle inequality (2.2) and a rough bound of the perimeter of ∂R_1^i by $4D^i$. Putting this together with (2.10) and using $D^i = D^1/2^i$, we get

$$\left| \int_{\partial R} f dz \right| \le 4^i \left| \int_{\partial R_1^i} f dz \right| \le 4^i \cdot \left(4D^i \cdot \varepsilon D^i\right) = \varepsilon \cdot 4 \left(D^1\right)^2 .$$

This is true for any $\varepsilon > 0$, so we are done. $\square$

We remark that if f is assumed to be C^1, then one can derive Theorem 2.9 based on Green's theorem in calculus and the Cauchy–Riemann equations. Here, we do not make this assumption, although later we will see that a holomorphic function is always C^1.

Before exploring more consequences of this important theorem, we first notice that we can make it slightly more general. That is, we do not need the function to be holomorphic on the whole triangle.

Theorem 2.14 (Goursat's theorem, missing a point)**.** Let $f\colon U \to \mathbb{C}$ be a continuous function and $p \in U$. Suppose f is holomorphic in $U \setminus \{p\}$. If R is a (solid) closed triangle in U, then

$$\int_{\partial R} f dz = 0,$$

where ∂R is the boundary of R and we view it as a piecewise smooth curve.

In the proof, for two points A and B on the complex plane, we let $\overline{AB}$ be the segment connecting A and B. Given three points, we let ΔABC be the triangle with A, B, and C being its vertices, and its boundary is oriented in the direction from A to B, and so on.

Proof. If p is not in R, then the theorem already follows from Theorem 2.9.

Now, suppose p is one of the vertices of R, and let the other two vertices be B and C so that ΔpBC is oriented counterclockwise. Given $\varepsilon > 0$, we then choose $X \in \overline{pB}$ and $Y \in \overline{pC}$ such that the lengths of them are less than ε. We can then estimate

$$\int_{\partial R} f dz = \int_{\partial \Delta pXY} f dz + \int_{\partial \Delta XBY} f dz + \int_{\partial \Delta BCY} f dz = \int_{\partial \Delta pXY} f dz,$$

where we apply Theorem 2.9 to the triangles ΔBCY and ΔpXY. Since f is bounded on R, say bounded by $M < \infty$, we can then estimate

$$\left| \int_{\partial R} f dz \right| = \left| \int_{\partial \Delta pXY} f dz \right| \leq M \cdot 4\varepsilon.$$

This is true for all ε, so the conclusion follows.

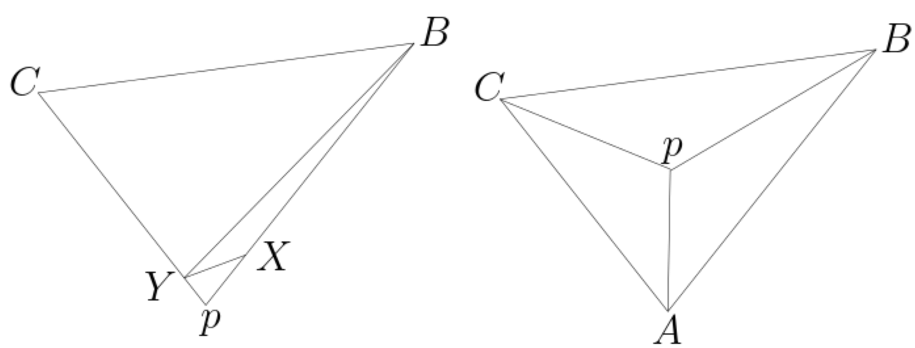

FIGURE 4. Possible choices of X and Y in the second case and the picture in the third case.

Finally, we deal with the case when $p \in R \setminus \{A, B, C\}$ where we write $R = \Delta ABC$. Then we can apply the previous case to $\Delta ABp, \Delta BCp$, and ΔCAp to get

$$\int_{\partial R} f dz = \int_{\partial \Delta ABp} f dz + \int_{\partial \Delta BCp} f dz + \int_{\partial \Delta CAp} f dz = 0 + 0 + 0 = 0.$$

This finishes the proof of the theorem. $\square$

One can imagine that Goursat's theorem basically implies the line integral of a holomorphic function along any reasonable curve should vanish, since every bounded domain can be approximated by triangles as long as its boundary is not too bad. We can make this guess rigorous by using the simple observation, Proposition 2.8. That is, we can now use Goursat's theorem to get a primitive of a holomorphic function. We do it in a convex set.

Corollary 2.15. Let $f\colon U \to \mathbb{C}$ be a continuous function and $p \in U$. Suppose f is holomorphic in $U \setminus \{p\}$. If U is a convex set, then

$$\int_\gamma f dz = 0$$

for any closed path γ in U.

Proof. We will show that f admits a primitive and apply Proposition 2.8. Given two points A and B, we let $\overline{AB}$ be the segment oriented from A to B.

Fix a point O in U. Given any $w \in U$, we define

$$F(w) := \int_{\overline{Ow}} f dz.$$

We will show that $F' = f$. To see this, fix $z_0 \in U$. Given any $\varepsilon > 0$, by the continuity of f, there exists $\delta > 0$ such that $B_\delta(z_0) \subseteq U$ and $|f(z) - f(z_0)| < \varepsilon$ for all $z \in B_\delta(z_0)$. For such a $z \in B_\delta(z_0)$, we then have

$$\begin{aligned} \left|\frac{F(z) - F(z_0)}{z - z_0} - f(z_0)\right| = \left|\frac{1}{z - z_0}\int_{\overline{zz_0}} f dz - f(z_0)\right| &= \left|\frac{1}{z - z_0}\int_{\overline{zz_0}} (f(z) - f(z_0))\, dz\right| \\ &\leq \frac{1}{|z - z_0|} \cdot \varepsilon \cdot |\overline{zz_0}| = \varepsilon, \end{aligned}$$

where we apply Theorem 2.14 to the triangle formed by O, z, and z_0 and use the triangle inequality (2.2). This then shows that

$$F'(z_0) = \lim_{z \to z_0} \frac{F(z) - F(z_0)}{z - z_0} = f(z_0),$$

and the result follows from Proposition 2.8. $\square$

We remark that the proof also works for a star-shaped set in $\mathbb{C}$. This does not make a big difference, as we are going to apply it to a local open ball to get the Cauchy integral formula.

Theorem 2.16 (Cauchy's integral formula for convex sets)**.** Let γ be a closed path in a convex open set U. If f is a holomorphic function on U, then for $w \in U \setminus \gamma$,

$$f(w) \cdot \mathrm{Ind}_\gamma(w) = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z - w} dz.$$

In particular, if $\overline{B}_r(a) \subseteq U$, we have

$$f(w) = \frac{1}{2\pi i}\int_{\partial B_r(a)} \frac{f(z)}{z - w} dz$$

for any $w \in B_r(a)$.

Proof. Fix $w \in U$. We consider the following function

$$g(z) := \begin{cases} \frac{f(z)-f(w)}{z-w} & \text{if } z \neq w \\ f'(w) & \text{for } z = w \end{cases}.$$

By definition, g is a continuous function and is holomorphic on $U \setminus \{w\}$. Thus, Corollary 2.15 implies $\int_\gamma g dz = 0$. Since $w \notin \gamma$, this means

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

where we use the definition of $\operatorname{Ind}_\gamma$ in Corollary 2.6. □

2.1.3. *Consequences of Cauchy's formula.* We will now see that even only stated for convex sets, Corollary 2.15 and Theorem 2.16 already have many applications, including the fact that holomorphic functions are analytic. The general case of the Cauchy integral formula will be dealt with in Theorem 2.38.

First, as mentioned many times, we can now show that holomorphic functions are analytic.

Theorem 2.17. Let f be a holomorphic function on an open set U. Then f is analytic. In fact, given any $\overline{B}_r(z_0) \subseteq U$, $f(z)$ is equal to a convergent power series

$$f(z) = \frac{1}{2\pi i}\sum_{n=0}^{\infty}\left(\int_{\partial B_r(z_0)} \frac{f(w)}{(w-z_0)^{n+1}}dw\right)(z-z_0)^n.$$

Proof. This is a direct consequence of Theorem 2.16 and Proposition 2.4. □

Before seeing more consequences, we first note that by using the expansion techniques in Proposition 2.4, we can get a formula of remainder terms in the Taylor expansion if we expand $1/(z-w)$ up to finitely many terms.[ex] We record it as a corollary of this theorem and the proof of Proposition 2.4.

Corollary 2.18. Let f be a holomorphic function on an open set U. Then given any $\overline{B}_r(z_0) \subseteq U$ and $k \in \mathbb{N}$, we can write

$$\begin{aligned} f(z) &= \sum_{n=0}^{k-1}\frac{f^{(n)}(z_0)}{n!}(z-z_0)^n + \frac{1}{2\pi i}\left(\int_{\partial B_r(z_0)} \frac{f(w)}{(w-z_0)^k(w-z)}dw\right)(z-z_0)^k \\ &= \frac{1}{2\pi i}\sum_{n=0}^{k-1}\left(\int_{\partial B_r(z_0)} \frac{f(w)}{(w-z_0)^{n+1}}dw\right)(z-z_0)^n + \frac{1}{2\pi i}\left(\int_{\partial B_r(z_0)} \frac{f(w)}{(w-z_0)^k(w-z)}dw\right)(z-z_0)^k. \end{aligned}$$

Now we know that holomorphic functions and analytic functions are the same. We derive more interesting properties of them. First, we can get a kind of converse to the Cauchy–Goursat theorem.

Theorem 2.19 (Morera's theorem)**.** Let $f\colon U \to \mathbb{C}$ be a continuous function. If for any closed triangle $R \subseteq U$, it satisfies $\int_{\partial R} f dz = 0$, then f is holomorphic.

Proof. For any $\overline{B}_r(z_0) \subseteq U$, following the proof in Corollary 2.15, we can find a converse F of f by taking $F(z) := \int_{\overline{z_0 z}} f dz$. Since $F' = f$, F is holomorphic, and hence analytic. Thus, f is holomorphic on $B_r(z_0)$. This is true for any $\overline{B}_r(z_0) \subseteq U$, so the conclusion follows. □

We remark that both Theorem 2.17 and Theorem 2.19 are improvements of "regularity" of a function that is a priori only differentiable or even continuous. This is commonly seen for second-order elliptic operators in PDE theories. However, here, note that holomorphic functions satisfy a first-order equation, either the Cauchy–Riemann equation (1.7) or the $\overline{\partial}$-equation $\partial_{\overline{z}} f = 0$.

We also remark that Morera's theorem does not require convexity for U. Morera's theorem can be used to prove that a locally uniformly convergent sequence of holomorphic functions has a nice limit. This result is due to Weierstrass.

Corollary 2.20 (Weierstrass)**.** Let f_n be a sequence of holomorphic functions on an open set U. If f_n converges to a function f locally uniformly, then the limit f is also holomorphic.

Proof. Let R be a closed triangle in U. Then since each f_n is holomorphic, by the convergence on the compact set R, we have

$$\int_{\partial R} f dz = \lim_{n\to\infty} \int_{\partial R} f_n dz = \lim_{n\to\infty} 0 = 0.$$

This is true for any closed triangle in U, so Theorem 2.19 implies that f is holomorphic. □

This corollary can be used to produce holomorphic functions. In fact, we can say more about the convergence in the corollary. To see this, we first see that the power series expansion gives us good estimates on high derivatives of a holomorphic function.

Proposition 2.21. Let f be a holomorphic function on an open set U. Then for any $\overline{B}_r(z_0) \subseteq U$, we have

$$|f^{(n)}(z_0)| \le \frac{n!}{r^n} \cdot \sup_{\overline{B}_r(z_0)} |f|.$$

Proof. By the expansion in Theorem 2.17, we have

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_{\partial B_r(z_0)} \frac{f(w)}{(w-z_0)^{n+1}} dw.$$

If we parametrize $\partial B_r(z_0)$ by $\gamma(t) := z_0 + re^{it}$ for $t \in [0, 2\pi]$, we have

$$f^{(n)}(z_0) = \frac{n!}{2\pi i} \int_0^{2\pi} \frac{f(z_0 + re^{it})}{(re^{it})^{n+1}} \cdot ire^{it} dt = \frac{n!}{2\pi} \int_0^{2\pi} \frac{f(z_0 + re^{it})}{(re^{it})^n} dt. \tag{2.22}$$

The estimate then follows directly. □

Using Proposition 2.21, we can then strengthen Corollary 2.20 to the following. As Corollary 2.20, this is also sometimes attributed to Weierstrass.

Corollary 2.23 (Weierstrass)**.** Let f_n be a sequence of holomorphic functions on an open set U. If f_n converges to a function f locally uniformly, then the limit f is holomorphic and for all $k \in \mathbb{N}$, $f_n^{(k)}$ converges to $f^{(k)}$ locally uniformly.

Proof. Let K be a compact subset in U. We can then take $r > 0$ such that $\overline{B}_r(z) \subseteq U$ for all $z \in K$, and hence $\overline{B}_r(K) := \cup_{z\in K}\overline{B}_r(z)$ is still a compact set.$^{\text{ex}}$ Then for any $z_0 \in K$, and any n, Proposition 2.21 implies

$$|f'(z_0) - f_n'(z_0)| \leq \frac{1}{r} \cdot \sup_{\overline{B}_r(K)} |f - f_n|. \tag{2.24}$$

Now we can conclude the proof. Given any ε, by the uniform convergence of f_n on $\overline{B}_r(K)$, there exists $N \in \mathbb{N}$ such that $|f - f_n| < r\varepsilon$ on $\overline{B}_r(K)$ for $n \geq N$. This and (2.24) then imply $|f' - f_n| \leq \varepsilon$ on K for $n \geq N$. This shows that f_n' converges uniformly on K. The argument for $f_n^{(k)}$ is similar. $\square$

Note that both Corollaries 2.20 and 2.23 are far from being true in the real case. The uniform convergence of the sequence itself definitely says nothing about their derivatives.

The estimate in Proposition 2.21 is very powerful. First, it gives a growth restriction for non-trivial entire holomorphic function. The proof only uses the $n = 0$ case, and will be left as an exercise in Assignment. One can also try to look at a version that depends on the high-order derivative estimates.

Corollary 2.25 (Liouville's theorem)**.** Let f be a holomorphic function on $\mathbb{C}$. If f is bounded, then f is constant.

In view of the estimate in Proposition 2.21, one can in fact prove Liouville's theorem for entire holomorphic functions with "sublinear" growth. On the other hand, based on Liouville's theorem, one can prove the fundamental theorem of algebra in a relatively simple way. We will also see this in Assignment 2, and here just list it as another corollary.

Corollary 2.26 (Fundamental theorem of algebra)**.** Every non-constant polynomial with complex coefficients has at least a complex root.

Another useful consequence of Proposition 2.21 is the **maximum principle**. This kind of result says that the maximum of a function must be achieved on its boundary. It has a stronger version, usually called the **strong maximum principle**, saying that if the maximum is achieved in the interior, then the function must be constant. We will show this stronger version for holomorphic function.

Theorem 2.27 (Strong maximum principle)**.** Let f be a holomorphic function on a **connected** open set U. If $\sup_U |f|$ is achieved at a point in U, then f is a constant function.

From its proof, we will see that one can also replace $|f|$ with $\operatorname{Re} f$ or $\operatorname{Im} f$. Before proving the theorem, we first single out the useful formula (2.22) especially when $n = 0$.

Corollary 2.28. Let f be a holomorphic function on an open set U. Then for any $\overline{B}_r(z_0) \subseteq U$, we have

$$f(z_0) = \frac{1}{2\pi}\int_0^{2\pi} f(z_0 + re^{it})dt.$$

This formula is sometimes also called Cauchy's integral formula, as a special case of its more generalized version. Note that it is also a mean-value-type equality, a property that is also true for harmonic functions.

Proof of Theorem 2.27. Suppose the supremum of $|f|$ is achieved at a point $w \in U$, say $|f(w)| = M$. We will show that

$$|f|^{-1}(M) := \{z \in U : |f(z)| = M\} = U.$$

First, if $z_0 \in |f|^{-1}(M)$, then by taking $r > 0$ such that $\overline{B}_r(z_0) \subseteq U$, Corollary 2.28 implies that for all $\varepsilon \in (0, r)$,

$$M = |f(z_0)| = \left|\frac{1}{2\pi}\int_0^{2\pi} f(z_0 + \varepsilon e^{it})dt\right| \leq \frac{1}{2\pi}\int_0^{2\pi} \left|f(z_0 + \varepsilon e^{it})\right| dt.$$

In particular, this implies

$$\frac{1}{2\pi}\int_0^{2\pi} \left(\left|f\left(z_0 + \varepsilon e^{it}\right)\right| - M\right) dt \geq 0.$$

However, since $|f(z_0 + \varepsilon e^{it})| \leq M$ for all $t \in [0, 2\pi]$, the integrand is a non-positive continuous function, so it must vanish. This is true for all $\varepsilon \in (0, r)$ and $t \in [0, 2\pi]$, so we have $B_r(z_0) \subseteq |f|^{-1}(M)$. This shows that $|f|^{-1}(M)$ is an open set in U.

On the other hand, $|f|^{-1}(M)$ is automatically a closed set in U since $|f|$ is continuous. Since U is connected, we have either $|f|^{-1}(M) = \emptyset$ or $|f|^{-1}(M) = U$. We know $w \in |f|^{-1}(M)$, so $|f|^{-1}(M)$ can only be U, and the proof is complete. □

A particular consequence of the strong maximum principle is the weak maximum principle. It says that if f is a holomorphic function on U and K is a compact set in U, then

$$\max_K |f| = \max_{\partial K} |f|.$$

This follows almost directly from Theorem 2.27.$^{\text{ex}}$

There are a lot more applications of the maximum principle. For example, one can use it to prove the following Schwarz lemma (Assignment 2), which implies a classification result of the automorphism group of the unit disc D.

Corollary 2.29 (Schwarz lemma)**.** Let D be the unit disc and $f\colon D \to D$ be a holomorphic function. If $f(0) = 0$, then

(1) for all $z \in D$, $|f(z)| \leq |z|$;

(2) $|f'(0)| \leq 1$;

(3) if either $|f(z)| = |z|$ for some $z \in D \setminus \{0\}$ or $|f'(0)| = 1$, then $f = e^{it}$ is a rotation about 0.

As a consequence, one can derive that the **automorphism group** of D,

$$\mathrm{Aut}(D) := \{f\colon D \to D \text{ is biholomorphic}\},$$

consists of elements of the form

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

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

We mention another immediate application of Theorem 2.17 and the Cauchy integral formula (Corollary 2.28). Recall that from Lemma 1.8, we see a connection between harmonic functions and holomorphic functions. Based on this, we can obtain the same regularity property for harmonic functions from these Cauchy's theorems, and these are also left in Assignment 2.

Corollary 2.30. Let $u\colon U \to \mathbb{R}$ be a harmonic function on an open set U. Then u is (real) analytic. Also, u satisfies the mean-value equality, in the sense that given any $\overline{B}_r(z_0) \subseteq U$, we have

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} u\left(z_0 + re^{it}\right) dt.$$

This is another example of elliptic regularity. A priori, a harmonic function is a C^2 function that satisfies $\Delta u = 0$. This corollary tells us that they are in fact automatically analytic. More properties and applications of harmonic functions will be discussed in Section 5.3.

Finally, we mention an application of the Cauchy–Goursat theorem that is about Cauchy's original motivation. It is about using the integral formula to calculate some definite integrals. Currently, we have Corollary 2.15 that only works for convex sets, but we will see that it is enough to derive some non-trivial integrals.

Example 2.31. We will use Corollary 2.15 to calculate

$$e^{-\pi\xi^2} = \int_{-\infty}^{\infty} e^{-\pi x^2} e^{-2\pi i x\xi} dx \tag{2.32}$$

for any $\xi \in \mathbb{R}$. This means that the Fourier transform of $e^{-\pi x^2}$ is still itself. To see this, we consider a rectangular curve γ with vertices $L, L+i\xi, -L+i\xi$, and $-L$.

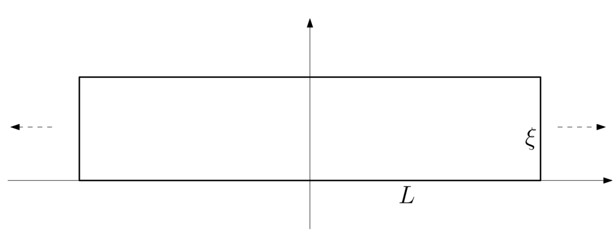

FIGURE 5. A rectangular curve whose top and bottom sides are longer and longer.

By Corollary 2.15, we know that the line integral of $f(z) = e^{-\pi z^2}$ along γ vanishes, and we can calculate the line integrals of the four sides directly. The bottom part of γ can be parametrized by

$\gamma_{\text{bot}}(t) = t$, $t \in [-L, L]$, so

$$\int_{\gamma_{\text{bot}}} f dz = \int_{-L}^{L} e^{-\pi t^2} dt \to 1 \tag{2.33}$$

as $L \to \infty$. For the right side, we can parametrize it by $\gamma_{\text{right}}(t) = L + it$ for $t \in [0, \xi]$, and we derive

$$\int_{\gamma_{\text{right}}} f dz = \int_0^{\xi} e^{-\pi(L+it)^2} i\, dt = i \int_0^{\xi} e^{-\pi\left(L^2 - t^2\right) + 2\pi L t i} dt.$$

Its magnitude can then be estimated by

$$\left| \int_{\gamma_{\text{right}}} f dz \right| = \int_0^{\xi} e^{-\pi L^2 + \pi t^2} dt \le e^{-\pi L^2} \cdot C_\xi \to 0 \tag{2.34}$$

as $L \to \infty$, where $C_\xi = \int_0^\xi e^{\pi t^2} dt$ is a constant depending only on ξ. One can similarly show that

$$\int_{\gamma_{\text{left}}} f dz \to 0 \text{ as } L \to \infty. \tag{2.35}$$

Finally, for the top part, we parametrize it by $\gamma_{\text{top}}(t) = -t + i\xi$ for $t \in [-L, L]$ and get

$$\begin{aligned} \int_{\gamma_{\text{top}}} f dz = \int_{-L}^{L} e^{-\pi(-t+i\xi)^2} \cdot (-1) dt &= -e^{\pi\xi^2} \int_{-L}^{L} e^{-\pi t^2 + 2\pi i \xi t} dt \\ &= -e^{\pi \xi^2} \int_{-L}^{L} e^{-\pi x^2 - 2\pi i \xi x} dx, \end{aligned} \tag{2.36}$$

where in the last step, we change the variables $x := -t$. Combining (2.33), (2.34), (2.35), and (2.36) with Corollary 2.15, we get

$$0 = \lim_{L \to \infty} \int_\gamma f dz = 1 - e^{\pi \xi^2} \int_{-\infty}^{\infty} e^{-\pi x^2 - 2\pi i \xi x} dx,$$

which implies (2.32).

2.2. **Global Cauchy theory.** We will develop some integral theories and prove a general form of the Cauchy integral formula in this section. The formula will be used many times in the rest of the course, and the tools developed will lead to further nice properties of holomorphic functions, as we will see at the end of this section.

To allow a general class of curves to integrate over, given an open set $U \subseteq \mathbb{C}$, we consider the **free abelian group**, $\mathcal{P}(U)$, generated by all oriented paths in U. That is, any element $\gamma \in \mathcal{P}(U)$ can be formally written as

$$\gamma = c_1 \cdot \gamma_1 + \cdots + c_N \cdot \gamma_N \tag{2.37}$$

where $c_n \in \mathbb{Z}$ and each γ_n is a path in U. An element in $\mathcal{P}(U)$ will be called a chain in U. When each path γ_n in (2.37) is a closed path, then γ will be called a **cycle** in U.

The reason that we introduce these is that it allows us to deal with line integral over a less restricted class of curves. For example, we can now talk about the integral over the union of two or more curves directly, without a limiting process. This will be helpful when we apply the Cauchy theorem in different situations. **However**, the proof of the general Cauchy theorem (Theorem 2.38) still works for a single closed path, the object we studied before, so if one is not comfortable with the notion of a general cycle, there is no problem to skip it for a while.

Given a continuous function f on U, for a chain γ of the form in (2.37), we define

$$\int_\gamma f dz := \sum_{n=1}^N c_n \int_{\gamma_n} f dz.$$

When γ is a cycle, we define

$$\mathrm{Ind}_\gamma(w) := \sum_{n=1}^N c_n \cdot \mathrm{Ind}_{\gamma_n}(w)$$

for $w \in U \setminus \gamma$. Here, we still use the same notation γ to denote the union of all the images of γ_n's. One can then interpret Ind_γ as the "total winding number" that all γ_n's wind around w. When $\gamma = 1 \cdot \gamma_1$ is just a usual (closed) path, the line integral and the index are just the ones we used to work with.

Theorem 2.38 (Cauchy's theorem)**.** Let f be a holomorphic function on U. If γ is a cycle in U such that

$$\mathrm{Ind}_\gamma(z) = 0 \text{ for all } z \in \mathbb{C} \setminus U, \tag{2.39}$$

then we have

(1) $\int_\gamma f dz = 0$, and

(2) $f(w) \cdot \mathrm{Ind}_\gamma(w) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z-w} dz$ for $w \in U \setminus \gamma$.

Note that condition (2.39) is always true when U is a convex open set. Thus, this generalizes the special case we proved in Theorem 2.16. It is interesting to note that in the proof, we will in fact use a few consequences of the first Cauchy integral formula in Section 2.1.

To prove the theorem, we introduce a continuity lemma. It will also be used later when we prove the inverse function theorem (Theorem 2.47).

Lemma 2.40. Let f be a holomorphic function on an open set U. Then the function

$$F(z,w) := \begin{cases} \frac{f(z)-f(w)}{z-w} & \text{if } z \neq w \\ f'(z) & \text{if } z = w \end{cases} \tag{2.41}$$

is a continuous function on $U \times U$.

Proof of Lemma 2.40. We need to verify the continuity at a point $(z_0, z_0) \in U \times U$. Note that for $z \neq w$ in a ball $B_r(z_0) \subseteq U$, we can write $\sigma(t) := tz + (1-t)w$ to be the segment from w to z, and hence

$$F(z,w) = \frac{f(z)-f(w)}{z-w} = \frac{1}{z-w}\int_0^1 f'(\sigma(t))\,\sigma'(t)dt = \int_0^1 f'(\sigma(t))\,dt.$$

This allows us to finish the proof. Given $\varepsilon > 0$, we can find $r > 0$ such that $B_r(z_0) \subseteq U$ and $|f'(z) - f'(z_0)| \leq \varepsilon$ for $z \in B_r(z_0)$. Then given any $z, w \in B_r(z_0)$, choosing the path σ as above, which lies in $B_r(z_0)$, we have

$$|F(z,w) - F(z_0,z_0)| = \left| \int_0^1 f'(\sigma(t))\, dt - f'(z_0) \right| \leq \int_0^1 \left| f'(\sigma(t)) - f'(z_0) \right| dt \leq \varepsilon.$$

This proves the continuity of F. □

A crucial idea of proving Theorem 2.38 is about analytic continuation. We want to prove a function vanishes on U, and to do this, we extend it to the whole complex plane and use its structure to show that it is bounded. Thus, in the course of the proof, many consequences of the Cauchy theorem in the convex case will play a role.

Proof of Theorem 2.38. We will prove (2) first. Note that (2) is equivalent to

$$0 = \frac{1}{2\pi i} \int_\gamma \frac{f(z)}{z-w} dz - f(w) \cdot \frac{1}{2\pi i} \int_\gamma \frac{dz}{z-w} = \frac{1}{2\pi i} \int_\gamma \frac{f(z) - f(w)}{z-w} dz.$$

In view of this, we define F as in (2.41), and our goal is then to prove

$$h(w) := \int_\gamma F(z,w) dz = 0$$

on $U \setminus \gamma$. Note that now this function still makes sense for $w \in \gamma$ based on Lemma 2.40, and in fact, we will prove that it can be extended to the whole $\mathbb{C}$ as a globally zero function.

First, we show that h is continuous. Since γ (as its image) is a compact set, given any $w_n \to w_0$ in U, $F(\cdot, w_n) \to F(\cdot, w_0)$ uniformly on γ. Thus, $h(w_n) \to h(w_0)$, and hence h is continuous.

Next, we show that h is holomorphic. Given any closed triangle $R \subseteq U$, we can write

$$\int_{\partial R} h(w) dw = \int_{\partial R} \left(\int_\gamma F(z,w) dz \right) dw = \int_\gamma \left(\int_{\partial R} F(z,w) dw \right) dz = 0,$$

where the last equality follows from Goursat's theorem (Theorem 2.14), since for any fixed z, $F(z, \cdot)$ is holomorphic on $U \setminus \{z\}$ and continuous on U. Thus, Morera's theorem (Theorem 2.19) implies that h is holomorphic.

Finally, we extend h to an entire function. We consider the set

$$V := \mathrm{Ind}_\gamma^{-1}(0) = \{ z \in \mathbb{C} \setminus \gamma : \mathrm{Ind}_\gamma(z) = 0 \}.$$

Recall that Ind_γ is an integer-valued continuous function (Corollary 2.6), so V is an open set. By the assumption (2.39), we have $\mathbb{C} \setminus U \subseteq V$, so $U \cup V = \mathbb{C}$. For $w \in V$, we consider

$$\widetilde{h}(w) := \int_\gamma \frac{f(z)}{z-w} dz,$$

which differs from $h(w)$ by Ind_γ. This implies that for $w \in U \cap V$, we have

$$h(w) = \int_\gamma \frac{f(z)}{z-w} dz - 2\pi i \cdot \mathrm{Ind}_\gamma(w) = \widetilde{h}(w).$$

Moreover, $\widetilde{h}$ is a holomorphic function on V by Proposition 2.4. Thus, we can define a function

$$h_0(w) := \begin{cases} h(w) & \text{if } w \in U \\ \widetilde{h}(w) & \text{if } w \in V \end{cases}$$

which is well-defined and holomorphic on $U \cup V = \mathbb{C}$. To see that it vanishes, note that since γ is compact, we can find $R < \infty$ such that $\gamma \subseteq B_R$. In particular, $\mathbb{C} \setminus B_R \subseteq V$, so we can estimate

$$|h_0(w)| \leq \sup_\gamma |f| \cdot \operatorname{Length}(\gamma) \cdot \frac{1}{|w| - R} \tag{2.42}$$

for $w \in \mathbb{C} \setminus B_R$. In particular, h_0 is globally bounded. By Liouville's theorem (Corollary 2.25), h_0 is a constant, and hence must be zero by letting $|w| \to \infty$ in (2.42). This finishes the proof of (2).

To prove (1), take any $z_0 \in U \setminus \gamma$ and consider the holomorphic function $f(z) \cdot (z - z_0)$. Applying (2) with $w = z_0$, we obtain

$$0 = f(z_0)(z_0 - z_0) \cdot \operatorname{Ind}_\gamma(z_0) = \frac{1}{2\pi i} \int_\gamma \frac{f(z)(z - z_0)}{z - z_0} dz = \frac{1}{2\pi i} \int_\gamma f dz.$$

Hence, (1) follows. □

A consequence of Theorem 2.38 is that when we smoothly perturb a curve, the Cauchy integral does not change. To state it rigorously, we introduce a way to formulate such a perturbation. Let $\gamma_0, \gamma_1 \colon [0,1] \to \mathbb{C}$ be two oriented paths in an open set U. We say that these two paths are **homotopic** to each other in U if there exists a continuous map $H \colon [0,1] \times [0,1] \to U$ such that

$$H(0, \cdot) = \gamma_0 \text{ and } H(1, \cdot) = \gamma_1. \tag{2.43}$$

That is, there is a "continuous process" that deforms γ_0 to γ_1 in U.

The notion of homotopy plays an important role in the study of topology and geometry. Here, we mention one particular consequence. That is, if γ_0 and γ_1 are two closed paths that are homotopic to each other in an open set U, then$^{\text{ex}}$

$$\operatorname{Ind}_{\gamma_0}(w) = \operatorname{Ind}_{\gamma_1}(w) \text{ for all } w \in \mathbb{C} \setminus U. \tag{2.44}$$

The reason why we separate this conclusion is that it implies

$$\int_{\gamma_0} f dz = \int_{\gamma_1} f dz \tag{2.45}$$

for any holomorphic function f on U. This is a direct consequence of Theorem 2.38 (1) if we let $\gamma := \gamma_1 - \gamma_0 \in \mathcal{P}(U)$.

Forget about two homotopic paths. In general, if two cycles γ_0 and γ_1 satisfy (2.44), then (2.45) is true for any holomorphic function f on U. Homotopic closed paths are just a special and important case of this general principle, which is usually regarded as part of the Cauchy theorem.

Example 2.46. We mention an example why this generalized Cauchy theorem is useful. It allows us to change the integral path freely. For example, consider γ_1 and γ_2 as in Figure 6.

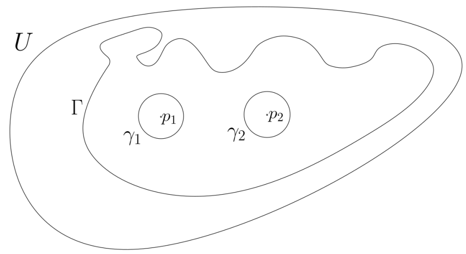

FIGURE 6. When a curve is complicated, the Cauchy theorem can be used to possibly simplify the line integral.

The open set U does not contain the two points p_1 and p_2, so a holomorphic function f on U may have a non-trivial line integral along the curve Γ in Figure 6. However, Γ and the two circles γ_1 and γ_2 satisfy (cf. Assignment 2)

$$\operatorname{Ind}_\Gamma(w) = \sum_{n=1}^{2} \operatorname{Ind}_{\gamma_n}(w)$$

for all $w \in \mathbb{C} \setminus U$. Hence, the Cauchy theorem implies

$$\int_\Gamma f dz = \sum_{n=1}^{2} \int_{\gamma_n} f dz$$

for any holomorphic function f on U. This may allow us to simplify the calculations. We will see more examples like this in the rest of this course. In particular, this idea of looking at small circles around points where a function is not defined is a key idea in the residue theorem in Section 3.3.

As mentioned, Lemma 2.40 can lead to other local information of holomorphic functions. First, it implies the inverse function theorem.

Theorem 2.47 (Inverse function theorem)**.** Let f be a holomorphic function on an open set U and $z_0 \in U$. If $f'(z_0) \neq 0$, then there exists $r > 0$ such that

(1) $f|_{B_r(z_0)}$ is injective,

(2) $f(B_r(z_0))$ is open in $\mathbb{C}$, and

(3) the inverse $g\colon f(B_r(z_0)) \to B_r(z_0)$ is holomorphic.

Note that the statement (3) is valid based on (1) and (2). In this case, $f|_{B_r(z_0)}\colon B_r(z_0) \to f(B_r(z_0))$ is called a biholomorphic function.

Proof. For (1), we apply Lemma 2.40, which implies that there exists $r > 0$ such that $B_r(z_0) \subseteq U$ and

$$\left| \frac{f(z) - f(w)}{z - w} \right| \geq \frac{1}{2}|f'(z_0)|$$

for $z \neq w$ in $B_r(z_0)$ based on the continuity of the function F defined in (2.41) at (z_0, z_0). It means

$$|z-w| \leq \frac{2}{|f'(z_0)|}|f(z)-f(w)| \tag{2.48}$$

for all $z, w \in B_r(z_0)$ (including the case when $z = w$). This implies the injectivity of f on $B_r(z_0)$.

For (2), given $w_0 \in B_r(z_0)$, we want to show that $f(w_0)$ is an interior point of $f(B_r(z_0))$. By (2.48), we can find $\rho > 0$ such that

$$\left|f(w_0+\rho e^{it})-f(w_0)\right| \geq \frac{|f'(z_0)|}{2}\rho =: L.$$

We claim that this implies

$$B_{L/2}\left(f(w_0)\right) \subseteq f\left(B_r(z_0)\right), \tag{2.49}$$

which implies the openness of $f\left(B_r(z_0)\right)$. To see (2.49), note that for $y \in B_{L/2}\left(f(w_0)\right)$, we have

$$\left|y-f(w_0+\rho e^{it})\right| \geq \left|f(w_0+\rho e^{it})-f(w_0)\right| - |f(w_0)-y| > \frac{L}{2}$$

for all $t \in [0, 2\pi]$, and hence

$$\left|\frac{1}{y-f(w_0+\rho e^{it})}\right| < \frac{2}{L} < \left|\frac{1}{y-f(w_0)}\right|. \tag{2.50}$$

This implies y is mapped by f. Otherwise, if y were not mapped by f, then $1/(y-f)$ would be a holomorphic function on $\overline{B}_\rho(w_0)$, and then (2.50) could not happen based on the maximum principle (Theorem 2.27) since f is non-constant.

For (3), fix $w_0 \in B_r(z_0)$ and let $y_0 := f(w_0)$. Since y_0 is an interior point of the image of f, given any y close to y_0, we can write

$$\frac{g(y)-g(y_0)}{y-y_0} = \frac{w-w_0}{f(w)-f(w_0)}$$

for a unique w close to w_0, with $w \to w_0$ if and only if $y \to y_0$. Hence, taking the limit, we have $g'(y_0) = 1/f'(w_0)$, and hence g is holomorphic at y_0. □

Example 2.51. We mention a direct example. That is, $f(z) = e^z$. Since $f'(z) = e^z \neq 0$ for any $z \in \mathbb{C}$, Theorem 2.47 implies that there is always a local inverse function of e^z. That is a logarithm function. An issue about it is that it is impossible to find a global inverse of e^z. We will discuss this more in Section 4.1.1.

Another consequence of the inverse function theorem is the open mapping theorem (Corollary 3.6). Since it pertains to the local behavior of zeros, we postpone its discussion until the next section. We summarize some key properties of holomorphic functions that we have encountered or will soon encounter later.

- Analyticity: Theorem 2.17.
- Cauchy's theorem: Theorem 2.38.
- Maximum principle: Theorem 2.27.
- Mean-value equality: Corollary 2.28.

- Morera's theorem: Theorem 2.19.
- Liouville's theorem: Corollary 2.25.
- Inverse function theorem: Theorem 2.47.
- Strong unique continuation: Corollary 3.4.
- Open mapping theorem: Corollary 3.6.

We end this section by mentioning another example of calculating definite integrals. With a more general version of the Cauchy theorem (Theorem 2.38), we can now deal with more complicated curve.

Example 2.52. We will use Theorem 2.38 to calculate

$$\int_0^\infty \frac{1-\cos x}{x^2}dx = \frac{\pi}{2}. \tag{2.53}$$

The idea is to take the line integral of $f(z) = (1-e^{iz})/z^2$ along the following curve γ.

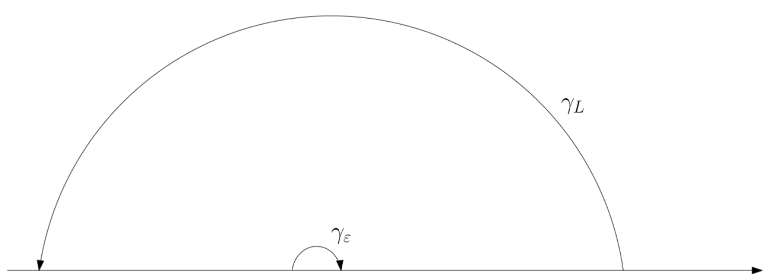

FIGURE 7. Semicircles tending to infinity and zero.

By Theorem 2.38, $\int_\gamma f dz = 0$ for any ε and L.[7] Thus, we will calculate the line integrals for γ_L and γ_ε, We can parametrize γ_L by $\gamma_L(t) = Le^{it}$, $t \in [0,\pi]$, so

$$\begin{aligned} \left|\int_{\gamma_L} f dz\right| = \left|\int_0^\pi \frac{1-e^{iLe^{it}}}{L^2e^{2it}} iLe^{it}dt\right| = \left|\int_0^\pi \frac{1-e^{iL\cos t - L\sin t}}{Le^{it}} idt\right| &\le \frac{1}{L}\int_0^\pi \left(1+e^{-L\sin t}\right)dt \\ &\le \frac{1}{L}\cdot 2\pi \to 0 \end{aligned} \tag{2.54}$$

as $L\to\infty$, using $\sin t \ge 0$ for $t\in[0,\pi]$. For γ_ε, we can parametrize it by $\gamma_\varepsilon(t) = \varepsilon e^{-it}$, $t \in [-\pi, 0]$. Moreover, for $z = \varepsilon e^{-it}$, we can estimate,

$$f(z) = \frac{1-e^{iz}}{z^2} = \frac{1-\left(1+iz-\frac{z^2}{2}-\frac{z^3}{6}i+\cdots\right)}{z^2} = \frac{-iz+O(z^2)}{z^2} = -\frac{i}{z}+O(1)$$

[7]Note that without the generalization in Theorem 2.38, we cannot deal with the curve γ directly by Corollary 2.15.

as $\varepsilon \to 0$.[8] Thus, we can estimate

$$\int_{\gamma_\varepsilon} f dz = \int_{-\pi}^{0} -\frac{i}{\varepsilon e^{-it}} \cdot \left(-i\varepsilon e^{-it}\right) dt + \int_{-\pi}^{0} O(1) \cdot \left(-i\varepsilon e^{-it}\right) dt \tag{2.55}$$
$$= \int_{-\pi}^{0} i^2 dt - \varepsilon \cdot O(1) \int_{-\pi}^{0} ie^{-it} dt \to -\pi + 0$$

as $\varepsilon \to 0$. Combining (2.54) and (2.55) with Theorem 2.38, we get

$$\int_{-\infty}^{\infty} \frac{1-e^{iz}}{z^2} dz := \lim_{L\to\infty} \lim_{\varepsilon \to 0^+} \left(\int_{-L}^{-\varepsilon} f dz + \int_{\varepsilon}^{L} f dz \right)$$
$$= \lim_{L\to\infty} \lim_{\varepsilon \to 0^+} \left(-\int_{\gamma_L} f dz - \int_{\gamma_\varepsilon} f dz \right) = \pi.$$

Taking the real parts of the both sides, we conclude

$$\int_{-\infty}^{\infty} \frac{1-\cos x}{x^2} dx = \pi,$$

which is equivalent to (2.53).

3. **Laurent series: local behavior near zeros and poles**

We will use the tools developed in the previous sections to further study local behavior of holomorphic functions. This includes their behavior near a zero or near a singularity, a point where a priori the function is not defined.

3.1. **Behavior near isolated zeros.** We start from the study of zeros of a holomorphic function. Given a function $f \colon U \to \mathbb{C}$, its zero set will be denoted by $Z(f)$. That is,

$$Z(f) := f^{-1}(0) = \{z \in U : f(z) = 0\}.$$

The first result is about the structure of $Z(f)$ when f is holomorphic.

Proposition 3.1. Let f be a holomorphic function on a connected open set U and let $Z(f)$ be its zero set. Then either $Z(f) = U$ or $Z(f)$ only contains discrete points.

Proof. This is a consequence of being an analytic function. For any $z_0 \in U$, since f is analytic, we can find $r > 0$ such that $\overline{B}_r(z_0) \subseteq U$ and for $z \in B_r(z_0)$, we have

$$f(z) = \sum_{n=0}^{\infty} a_n (z-z_0)^n.$$

There are then two situations. Either all a_n's are zeros, or we can find $n_0 \in \mathbb{Z}_{\geq 0}$ that is the smallest n such that $a_n \neq 0$. For the first case, we get $f|_{B_r(z_0)} = 0$. For the second case, we can then define a holomorphic function

$$h(z) := \begin{cases} \sum\limits_{n=0}^{\infty} a_n (z-z_0)^{n-n_0} & \text{if } z \in B_r(z_0) \\ \frac{f(z)}{(z-z_0)^{n_0}} & \text{if } z \in U \setminus \{z_0\} \end{cases}.$$

[8]By definition, one can just view $O(1)$ as a bounded constant as $\varepsilon \to 0$.

Note that the two definitions agree when $z \in B_r(z_0) \setminus \{z_0\}$. Hence, we can write

$$f(z) = h(z) \cdot (z - z_0)^{n_0},$$

where h satisfies $h(z_0) \neq 0$ (since $a_{n_0} \neq 0$), and hence we can shrink r so that $h|_{B_r(z_0)} \neq 0$. Thus, if $n_0 \geq 1$, then f has a unique zero z_0 in $B_r(z_0)$; if $n_0 = 0$, then f is non-vanishing on $B_r(z_0)$. In the first case, we say f has a zero of **order** n_0 at z_0.

Now, we can conclude the proof. We let U_1 be the set of the points in U of the first case and U_2 be the set of the points in U of the second case. That is,

$$\begin{aligned} U_1 &:= \{z \in U : f \text{ vanishes on } B_r(z) \text{ for some } r > 0\}, \text{ and} \\ U_2 &:= \{z \in U : f \text{ is non-vanishing on } B_r(z) \setminus \{z\} \text{ for some } r > 0\}. \end{aligned}$$

Then by the above discussion, we know $U = U_1 \cup U_2$ and $U_1 \cap U_2 = \emptyset$. Moreover, they are both open sets by their definitions, so the connectivity implies either $U = U_1$ or $U = U_2$. In the first case, we have $f = 0$ on U. In the second case, we know that all zeros are isolated. ☐

We remark that in the case when f has a discrete set of zeros, we can deduce that there are at most countably many zeros.[ex]

Example 3.2. We look at the function $\sin z$ at $z = 0$. We can write

$$\begin{aligned} \sin z = z - \frac{1}{3!}z^3 + \frac{1}{5!}z^5 &= z\left(1 - \frac{1}{3!}z^2 + \frac{1}{5!}z^4 - \cdots\right) \\ &=: z \cdot h(z), \end{aligned}$$

where h is a holomorphic function with $h(0) \neq 0$. Thus, $\sin z$ has a zero at 0 of order 1. In this case, 0 is called a simple zero of $\sin z$. One can also see this by noticing $\sin'(0) = \cos(0) \neq 0$.

Remark 3.3. Another perspective to view Proposition 3.1 is that the zero set of a holomorphic function is locally either zero-dimensional or (complex) one-dimensional. It is not easy to see why this is a distinguished property, but this amazing property is also roughly true for several complex variables. That allows us to look at "analytic subsets," which may include singularities so are more general than complex manifolds that we are going to look at in Section 5.5. For now, a related thing is that one can use zero sets of holomorphic functions to define a topology (by viewing them as closed sets). This is called the Zariski topology.

Proposition 3.1 poses strong restriction on the structure of zero sets of holomorphic functions. Its proof then leads to important consequences, including a **strong unique continuation** result.

Corollary 3.4. Let f be a holomorphic function on a connected open set U. Then f vanishes identically on U if
(1) either there exists $z_0 \in U$ such that $f^{(n)}(z_0) = 0$ for all $n \geq 0$,
(2) or f vanishes on a subset V which admits a limit point in U.

Note that the second case includes the situation when V is an open subset of U. The reason why this result is called unique continuation is as follows. Let $\widetilde{f}$ be a holomorphic function on a connected open set U and let S be an arbitrary subset in U. Then clearly $f := \widetilde{f}|_S$ is a holomorphic function on S. In this situation, we can view $\widetilde{f}$ as an **analytic continuation** of f from S to a

larger set U, a process we have seen many times. What Corollary 3.4 is telling us is as follows.
(1) If we are finding an extension F of f such that all of its derivatives are prescribed at a fixed point, then $\widetilde{f}$ is the unique possible extension.
(2) If we are finding an extension F of f from a subset V admitting a limit point, then $\widetilde{f}$ is still the unique possible extension.
In either case, if there is another extension F, then the difference $F - \widetilde{f}$ satisfies the condition in Corollary 3.4 so it must vanish identically, implying $F = \widetilde{f}$ on U. Thus, Corollary 3.4 is a very strong unique continuation property of holomorphic functions.

We can further understand the behavior of a non-constant holomorphic function near a zero. It will tell us that near a zero of finite order, the map has a "multi-sheeted" structure.

Theorem 3.5. Let f be a non-constant holomorphic function on a connected open set U. Suppose $z_0 \in U$, $y_0 = f(z_0)$, and as a zero of $f - y_0$, the order of z_0 is n_0. Then we can find $r > 0$ and a holomorphic function φ on $B_r(z_0) \subseteq U$ such that

(1) $f(z) = y_0 + \varphi(z)^{n_0}$ for $z \in B_r(z_0)$, and

(2) φ' is non-vanishing in $B_r(z_0)$ and is biholomorphic onto a disc.

Note that by Proposition 3.1, the order of z_0 as a zero of $f - f(z_0)$ is finite since f is non-constant. The theorem is a consequence of the inverse function theorem 2.47. As a corollary, we have the following open mapping theorem.[ex]

Corollary 3.6 (Open mapping theorem)**.** Let f be a non-constant holomorphic function on a connected open set U. Then $f(U)$ is still an open set.

The construction in the proof of Theorem 3.5 is related to defining **logarithm** for complex numbers. We will talk about this in a more general context later, but for a given complex number, this is not a big deal. Recall that the map $t \in [0, 2\pi] \mapsto e^{it}$ is a surjective map onto the unit circle on the complex plane. Thus, given any non-zero complex number $z \in \mathbb{C} \setminus \{0\}$, we can always find $t \in [0, 2\pi]$ such that

$$z = |z| \cdot \frac{z}{|z|} = |z| \cdot e^{it},$$

so it makes sense to define

$$\log z := \log |z| + it,$$

where $\log |z|$ is the usual logarithm for positive real numbers. The problem of globally defining such a function is finding a continuous choice of $t \in [0, 2\pi]$, but this is not an issue if we only look at a point.

Proof of Theorem 3.5. By Proposition 3.1, z_0 is an isolated zero of $f - y_0$. Thus, we can find $r > 0$ such that $B_r(z_0) \subseteq U$ and write

$$f(z) = y_0 + h(z) \cdot (z - z_0)^{n_0}$$

where n_0 is the order of z_0 as a zero of $f - y_0$ and h is a non-vanishing holomorphic function on $B_r(z_0)$. Note that it remains to take the "n_0-th root" of h. This can be done by first taking the logarithm of h.

To do this, first note that h'/h is holomorphic (since h does not vanish) and hence analytic on $B_r(z_0)$. Thus, we know that h'/h admits a primitive, that is, a holomorphic function F such that $F' = h'/h$. This implies

$$\left(he^{-F}\right)' = h'e^{-F} - h \cdot e^{-F} \cdot \frac{h'}{h} = 0.$$

Thus, we can shift F so that $h = e^F$. In fact, since we have

$$he^{-F} = h(z_0)e^{-F(z_0)}$$

on $B_r(z_0)$ by its constancy, we could consider $F_1 := F - F(z_0) + \log h(z_0)$, which satisfies

$$h = e^{F+\log h(z_0)-F(z_0)} = e^{F_1}.$$

Thus, (1) follows by letting $\varphi := e^{F_1/n_0} \cdot (z - z_0)$.

For (2), note that

$$\varphi'(z_0) = e^{F_1(z_0)/n_0} \neq 0,$$

so the conclusion follows from Theorem 2.47, after possibly shrinking r, and the fact that the power map $z \mapsto z^{n_0}$ is an open map. □

3.2. **Behavior near isolated singularities.**

3.2.1. *Laurent series.* Next, we talk about some possible singularities of a holomorphic function. For simplicity, we introduce a notation

$$\hat{B}_r(z_0) := B_r(z_0) \setminus \{z_0\}$$

for $r > 0$ and $z_0 \in \mathbb{C}$. This is not a standard notation, but we will use it from time to time.

Definition 3.7. If f is a holomorphic function on $\hat{B}_r(z_0)$ for some $r > 0$ and $z_0 \in \mathbb{C}$, then we say f has an **isolated singularity** at z_0.

This is the simplest type of singularities, since it is isolated. It already has some complicated behavior. However, we can still see some very straightforward examples.

Example 3.8. We consider the 1-ball B_1.

(1) If f is a holomorphic function on B_1, by considering $g := f|_{\hat{B}_1(0)}$, we get a holomorphic function g with an isolated singularity at the origin. We can see that this is a "fake" singularity since we can in fact extend g to the whole ball. Thus, such a singularity is not too bad.

(2) If f is a holomorphic function on B_1 with $f(0) = 0$ and $f(z) \neq 0$ for $z \in \hat{B}_1(0)$, then $g := 1/f$ defines a holomorphic function on $\hat{B}_1(0)$, and hence has an isolated singularity. We know that a non-trivial zero of a holomorphic function is isolated, so there are many such examples.

We will start from a general discussion of isolated singularities. The main tool will be Proposition 2.4 and the Cauchy theorem 2.38. After that, we will classify singularities into three categories and talk about them separately. We introduce a notation

$$A_{R,r}(z_0) := B_R(z_0) \setminus \overline{B}_r(z_0)$$

for $z_0 \in \mathbb{C}$ and $R > r \geq 0$. This is an open annulus centered at z_0. When $z_0 = 0$, we may just write $A_{R,r} := A_{R,r}(0)$. We will look at holomorphic functions on an annulus, since a holomorphic function with an isolated singularity can be viewed as a special case of these because $A_{R,0} = \hat{B}_R(0)$.

Theorem 3.9. Let f be a holomorphic function on an annulus $A_{R,r}(z_0)$. Then we can define

$$a_n := \frac{1}{2\pi i} \int_{\partial B_s(z_0)} \frac{f(z)}{(z - z_0)^{n+1}} dz$$

for any $s \in (r, R)$ and $n \in \mathbb{Z}$, so that

$$f(w) = \sum_{n=-\infty}^{\infty} a_n (w - z_0)^n \tag{3.10}$$

for all $w \in A_{R,r}$. The convergence is locally uniform in $A_{R,r}$ and the expression is unique. (In particular, the coefficients a_n's are independent of s.)

Proof. For simplicity, we assume $z_0 = 0$ and work on the annulus $A_{R,r}$. Then for $w \in A_{R,r}$, we can choose r_1 and r_2 such that $r < r_1 < |w| < r_2 < R$ and

$$f(w) = \frac{1}{2\pi i} \int_{\partial B_{r_2}} \frac{f(z)}{z - w} dz - \frac{1}{2\pi i} \int_{\partial B_{r_1}} \frac{f(z)}{z - w} dz. \tag{3.11}$$

This follows from the Cauchy theorem (Theorem 2.38) when we take $\gamma = \partial B_{r_2} - \partial B_{r_1}$, both oriented counterclockwise.

Now, we analyze the sum (3.11). We know that they are both analytic function in w by Proposition 2.4, and we can follow the proof idea of Proposition 2.4 (with $a = 0$) to get an expansion. That is, for $z \in \partial B_{r_2}$, we have

$$\frac{1}{z - w} = \frac{1}{z} \cdot \frac{1}{1 - \frac{w}{z}} = \frac{1}{z} \sum_{n=0}^{\infty} \left(\frac{w}{z}\right)^n = \sum_{n=0}^{\infty} \frac{w^n}{z^{n+1}},$$

while for $z \in \partial B_{r_1}$, we have

$$-\frac{1}{z - w} = \frac{1}{w} \cdot \frac{1}{1 - \frac{z}{w}} = \frac{1}{w} \sum_{m=0}^{\infty} \left(\frac{z}{w}\right)^m = \sum_{m=0}^{\infty} \frac{z^m}{w^{m+1}}.$$

As in Proposition 2.4, these both converge locally uniformly. Thus, combining these, we derive

$$f(w) = \frac{1}{2\pi i} \sum_{n=0}^{\infty} \left(\int_{\partial B_{r_2}} \frac{f(z)}{z^{n+1}} dz \right) w^n + \frac{1}{2\pi i} \sum_{m=0}^{\infty} \left(\int_{\partial B_{r_1}} f(z) \cdot z^m dz \right) \frac{1}{w^{m+1}}.$$

By letting $m = -n - 1$, we get a summation in n, and the result follows since for any $n \in \mathbb{Z}$,

$$\int_{\partial B_{r_2}} \frac{f(z)}{z^{n+1}} dz = \int_{\partial B_{r_1}} \frac{f(z)}{z^{n+1}} dz = \int_{\partial B_s} \frac{f(z)}{z^{n+1}} dz$$

for any $s \in (r, R)$ because $f(z)/z^{n+1}$ is holomorphic in $A_{R,r}$.

To get the uniqueness of the expression, suppose

$$f(w) = \sum_{n=-\infty}^{\infty} c_n (w - z_0)^n$$

is another expansion that converges absolutely and locally uniformly. Then for any $s \in (r, R)$,

$$a_j = \frac{1}{2\pi i} \int_{\partial B_s(z_0)} \frac{f(z)}{(z - z_0)^{j+1}} dz = \sum_{n=-\infty}^{\infty} c_n \cdot \frac{1}{2\pi i} \int_{\partial B_s(z_0)} \frac{(z - z_0)^n}{(z - z_0)^{j+1}} dz.$$

Note that by parametrizing $\partial B_s(z_0)$ by $z_0 + se^{it}$ for $t \in [0, 2\pi]$, we can calculate that for each n,

$$\begin{aligned} \frac{1}{2\pi i} \int_{\partial B_s(z_0)} \frac{(z - z_0)^n}{(z - z_0)^{j+1}} dz &= \frac{1}{2\pi i} \int_0^{2\pi} \left(se^{it}\right)^{n-j-1} \cdot ise^{it} dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left(se^{it}\right)^{n-j} dt = \delta_{nj} := \begin{cases} 1 & \text{if } n = j \\ 0 & \text{if } n \neq j \end{cases}. \end{aligned}$$

Putting this back to the summation, we get $a_j = \sum_n c_n \cdot \delta_{nj} = c_j$. This proves the uniqueness and the theorem follows. □

The expansion (3.10) is called the **Laurent series** of f centered at z_0. It can be thought of as a generalization of power series expansion, which is exactly when $a_n = 0$ for all $n < 0$. Note that this happens when f is a holomorphic function on the whole ball $B_R(z_0)$, and the Cauchy theorem implies $a_n = 0$ for $n < 0$. In general, if f has an isolated singularity at z_0, then based on the general behavior of the function near z_0, we can have the following classification.

Definition 3.12. Suppose f is a holomorphic function on $A_{R,0}(z_0) = \hat{B}_R(z_0)$, and let

$$f(z) = \sum_{n \in \mathbb{Z}} a_n (z - z_0)^n$$

be the Laurent series of f at z_0.

(1) We say z_0 is a **removable singularity** of f if $a_n = 0$ for $n < 0$.
(2) We say z_0 is a **pole** of f if there exists $n_0 < 0$ such that $a_{n_0} \neq 0$ and $a_n = 0$ for $n < n_0$.
(3) We say z_0 is an **essential singularity** of f if there are infinitely many $n < 0$ with $a_n \neq 0$.

Example 3.13. We mention a few examples of singularities before starting analyzing them.

(1) A rational function of the form $R(z) = P(z)/Q(z)$ with P and Q having no common factors has poles at the zeros of Q.
(2) Let $f(z) = e^{1/z}$ for $z \in \mathbb{C} \setminus \{0\}$. Then
$$e^{1/z} = \sum_{n=0}^{\infty} \frac{1}{n!} \left(\frac{1}{z}\right)^n = \sum_{n=0}^{\infty} \frac{1}{n!} \frac{1}{z^n}$$
for $z \neq 0$. Thus, f has an essential singularity at 0.

Singularities are not necessarily isolated (unlike zeros). We mention two examples.

(3) The function $\frac{1}{\sin(1/z)}$ has poles at $z = 1/(2\pi n)$ for all $n \in \mathbb{N}$. Thus, 0 is an accumulated singularity of the function.

(4) A function defined as in Proposition 2.4 has a curve of singularities. That is, if γ is a path in U and $\varphi\colon U \to \mathbb{C}$ is a continuous function, then the function

$$f(w) := \int_\gamma \frac{\varphi(z)}{z-w} dz$$

is holomorphic on $U \setminus \gamma$ but has the whole path γ as its singular set.

We can also look at possible singular behavior at ∞. When a function $f(z)$ is holomorphic outside a ball, we can consider $g(w) := f(1/w)$. We say that f has a removable singularity/pole/essential singularity at ∞ if g does at 0.

(5) Given a non-constant polynomial $f(z) = \sum_{n=0}^{N} a_n z^n$, if we let $g(w) = f(1/w)$, we get

$$g(w) = \sum_{n=0}^{N} \frac{a_n}{w^n},$$

so g has a pole at 0 and hence f has a pole at ∞. We will see that this property in fact characterizes a polynomial in Assignment 4.

(6) A non-constant entire function cannot have a removable singularity at ∞; otherwise, it will be bounded, which cannot be the case by the Liouville theorem.

(7) By (2), $f(z) = e^z$ has an essential singularity at ∞.

3.2.2. *Classification of singularities.* We will now see many different properties of different types of singularities.

First, for a removable singularity of a holomorphic function f, from its definition, we know that it literally means that f can extend to a holomorphic function on the whole disc. This is also a common definition of a removable singularity. We know that such a singularity can trivially happen from Example 3.8. In that example, the function is naturally bounded near the singularity. We will see that in general, this property characterizes removable singularities.

Proposition 3.14. Let $z_0 \in U$ and f be a holomorphic function on $U \setminus \{z_0\}$. Then the following are equivalent.

(1) z_0 is a removable singularity of f.

(2) f extends to a holomorphic function on U.

(3) There exists $r > 0$ such that f is bounded on $\hat{B}_r(z_0)$.

(4) $\lim_{z \to z_0} (z - z_0) f(z) = 0$.

The implication from (3) to (1) is often referred as the simplest case of the Riemann extension theorem.

Proof. The first two conditions are equivalent, as mentioned, based on the definition. It is straightforward to see (1)⇒(3)⇒(4). Thus, we will prove (4)⇒(1).

One approach is to use the Laurent series directly. We instead proceed without relying on that. We again assume $z_0 = 0$ for simplicity. We consider the function

$$g(z) := \begin{cases} z^2 f(z) & \text{if } z \in \hat{B}_r(0) \\ 0 & \text{if } z = 0 \end{cases}.$$

By (4), g is a continuous function. Moreover, we can see that

$$\lim_{z\to 0} \frac{g(z) - g(0)}{z - 0} = \lim_{z\to 0} \frac{z^2 f(z)}{z} = \lim_{z\to 0} z f(z) = 0$$

by (4) again. Thus, g is a holomorphic function (since it is evidently holomorphic on $\hat{B}_r(0)$). In particular, we can write

$$g(z) = \sum_{n=2}^{\infty} a_n z^n$$

in $B_r(z_0)$. We can then define $f(0) := a_2$, which makes $f(z) = g(z)/z^2$ a holomorphic function. ☐

We note that condition (4) is a priori weaker than (3). In fact, (4) includes the possibility that $f \sim (z - z_0)^{-\alpha}$ for some $\alpha \in (0, 1)$. It turns out that this cannot happen based on the nature of Laurent series.

Next, we study poles. If a holomorphic function f on $\hat{B}_r(z_0)$ has a pole at z_0, with

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

and $a_{-N} \neq 0$, then we say z_0 is a pole of order N. When $N = 1$, z_0 is called a **simple pole**.

Proposition 3.15. Let $z_0 \in U$ and f be a holomorphic function on $U \setminus \{z_0\}$. Then the following are equivalent.

(1) z_0 is a pole of f.

(2) There exist $N \in \mathbb{N}$ and $a_{-1}, \cdots, a_{-N}$ where $a_{-N} \neq 0$ such that

$$f(z) - \sum_{n=-N}^{-1} a_n (z - z_0)^n$$

has a removable singularity at z_0.

(3) There exists $N \in \mathbb{N}$ such that $(z - z_0)^N f(z)$ has a removable singularity and extends to a non-zero value at z_0.

(4) There exists $r > 0$ such that $1/f$ is holomorphic in $\hat{B}_r(z_0)$ and has a removable singularity and extends to zero at z_0.

(5) $\lim_{z\to z_0} f(z) = \infty$. This means $\lim_{z\to z_0} |f(z)| = \infty$.

When (2) and (3) happen, the number N is exactly the order of the pole z_0. The part of functions $\sum_{n=-N}^{-1} a_n(z-z_0)^n$ is called the **principal part** of f at z_0. As said in (2), it is a rational function $P(z)$ which is a polynomial in $1/(z-z_0)$ such that $f-P$ is holomorphic near z_0. It captures the singular behavior of the function near the pole.

Example 3.16. Let $f(z) = \frac{1}{z(z-2)}$. Then f has a simple pole at 0. Moreover, we can write

$$\frac{1}{z(z-2)} = \frac{1}{2}\left(\frac{1}{z-2} - \frac{1}{z}\right).$$

Since $1/(z-2)$ is holomorphic near 0, we know that the principal part of f at 0 is $-\frac{1}{2z}$.

Proof of Proposition 3.15. The equivalence among (1), (2), and (3) is clearer from the Laurent series expansion, so we focus on (4) and (5). We will more or less see (3) from the proof. We again assume $z_0 = 0$ for simplicity.

For (1)⇒(5), note that if 0 is a pole of order N, say

$$f(z) = \sum_{n \geq -N} a_n z^n$$

on $\hat{B}_r(0)$ with $a_{-N} \neq 0$, then we can write $f(z) = h(z)/z^N$ if we let

$$h(z) := \sum_{n \geq 0} a_{n-N} z^n,$$

which defines a holomorphic function on B_r (which proves (3)). Since $h(0) \neq 0$, we have

$$\lim_{z \to 0} |f(z)| = \lim_{z \to 0} \frac{|h(z)|}{|z|^N} = \infty.$$

For (5)⇒(4), note that (5) implies that we can find $r > 0$ such that $|f| > 10$ on $\hat{B}_r(0)$. In particular, $f \neq 0$ on $\hat{B}_r(0)$. Thus, we can define $g := 1/f$, a holomorphic function on $\hat{B}_r(0)$. Since $|g| \leq 1/10$ on $\hat{B}_r(0)$, by Proposition 3.14 (3), g extends to a holomorphic function on B_r, with

$$g(0) = \lim_{z \to 0} \frac{1}{f(z)} = 0.$$

Hence, we get (4).

For (4)⇒(1), finally, we can write

$$g(z) = \frac{1}{f(z)} = \sum_{n \geq n_0} a_n z^n$$

in B_r with $a_{n_0} \neq 0$ for some $n_0 \geq 1$. Thus, we get a holomorphic function

$$h(z) := \frac{g(z)}{z^{n_0}} = \sum_{n \geq 0} a_{n+n_0} z^n$$

with $h(0) \neq 0$. That is, h is a non-vanishing function on B_r after shrinking r. In particular, $1/h$ is still holomorphic. We then get

$$f(z) = \frac{1}{z^{n_0}} \cdot \frac{1}{h(z)},$$

and we are done. □

Finally, we look at essential singularities.

Proposition 3.17 (Casorati–Weierstrass theorem)**.** Let $z_0 \in U$ and f be a holomorphic function on $U \setminus \{z_0\}$. Then the following are equivalent.

(1) z_0 is an essential singularity of f.

(2) For $r > 0$ such that $B_r(z_0) \subseteq U$, $\overline{f\left(\hat{B}_r(z_0)\right)} = \mathbb{C}$.

Example 3.18. Let $f(z) = e^{1/z}$. Then the image of f is $\mathbb{C} \setminus \{0\}$, which is dense in $\mathbb{C}$.

The Casorati–Weierstrass says that if z_0 is an essential singularity of f, then given any $w \in \mathbb{C}$, there exists a sequence $z_n \to z_0$ such that $f(z_n) \to w$ as $n \to \infty$. This is how the name, "essential" singularity, comes from. Note that, combined with Proposition 3.14 (2) and Proposition 3.15 (5), Proposition 3.17 also allows us to summarize the three different kinds of behavior based on the asymptotics near a singularity.

Corollary 3.19. Let $z_0 \in U$ and f be a holomorphic function on $U \setminus \{z_0\}$.

(1) z_0 is a removable singularity if and only if $\lim_{z \to z_0} f(z)$ exists and is finite.

(2) z_0 is a pole if and only if $\lim_{z \to z_0} f(z) = \infty$.

(3) z_0 is an essential singularity if and only if $\lim_{z \to z_0} f(z)$ does not exist (and can be any number sequentially).

Proof of Proposition 3.17. It is clear that (2) implies (1) based on Proposition 3.14 (2) and Proposition 3.15 (5). Thus, we will prove (1) implies (2).

We prove it by contradiction. Suppose not, that is, there exists $r > 0$ and $w \in \mathbb{C}$ such that $B_r(z_0) \subseteq U$ but $w \notin \overline{f\left(\hat{B}_r(z_0)\right)}$. Then we can find $\delta > 0$ such that

$$B_\delta(w) \subseteq \mathbb{C} \setminus \overline{f\left(\hat{B}_r(z_0)\right)}. \tag{3.20}$$

This then suggests considering the function

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

which is still holomorphic in $\hat{B}_r(z_0)$. Note that the property (3.20) implies that

$$|h(z)| = \frac{1}{|f(z) - w|} \leq \frac{1}{\delta}.$$

Hence, Proposition 3.14 (3) implies that h extends to a holomorphic function on the whole $B_r(z_0)$.

We can now conclude the proof by looking into two case. If $h(z_0) = 0$, then $f(z)$ has a pole at z_0, a contradiction. If $h(z_0) \neq 0$, then $f(z)$ has a removable singularity at z_0, another contradiction. Hence, the conclusion follows. □

For an essential singularity, we can't really say much, but if time permits, we will see a generalization of the Casorati–Weierstrass theorem, Picard's theorem, which gives more precise information about the image of a holomorphic function near an essential singularity. (See Theorem 4.10 and Theorem 5.24.) Among different singularities, the most interesting ones are poles. Especially, when there is a simple pole of a holomorphic function, it is important to study its Laurent coefficients. We will talk about some of them in the next section.

3.3. **Meromorphic functions and residue theory.** Let U be an open set in $\mathbb{C}$. We say f is a **meromorphic function** on U if there is a discrete subset E of points in U such that

(1) f is holomorphic in $U \setminus E$, and
(2) f has a pole at each $p \in E$.

The set of all meromorphic functions on U will be denoted by $\mathcal{M}(U)$.

We've seen some examples of meromorphic functions. They may have singularities, but by allowing them, the set of all meromorphic functions has a nicer structure. In fact, if U is connected, then $\mathcal{M}(U)$ is a **field**, in the sense that given $f, g \in \mathcal{M}(U)$, we have

- $f \pm g$ and $f \cdot g$ are both in $\mathcal{M}(U)$.
- $f/g \in \mathcal{M}(U)$ if $g \in \mathcal{M}(U) \setminus \{0\}$.

The identity element of this field is the constant function 1. Moreover, one can show that the order of poles at a given point is a **valuation**, in terms of the language of commutative algebra.[ex] We will see more structures or consequences about meromorphic functions in this section.

A useful consequence about the theory of meromorphic functions comes from calculating their resides. Suppose a meromorphic function f on U has a pole at $z_0 \in U$, say

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

near z_0. The coefficient a_{-1} is called the **residue** of f at the pole z_0. We will write

$$\mathrm{Res}_f(z_0) := a_{-1}.$$

We can first see why this term will stand out from the Laurent series by looking at the integral $\int_{\partial B_r(z_0)} f dz$. First, we take $r > 0$ such that $B_r(z_0) \subseteq U$ and f is holomorphic in $\hat{B}_r(z_0)$. Then by the Cauchy theorem, the integral of the holomorphic part vanishes, and by parametrizing $\partial B_r(z_0)$

with $\gamma(t) = z_0 + re^{it}$ $(t \in [0, 2\pi])$, we can calculate

$$\begin{aligned}\frac{1}{2\pi i}\int_{\partial B_r(z_0)} f dz &= \frac{1}{2\pi i}\sum_{n=-N}^{-1}\int_{\partial B_r(z_0)} a_n (z-z_0)^n dz \\ &= \frac{1}{2\pi i}\sum_{n=-N}^{-1}\int_0^{2\pi} a_n \left(re^{it}\right)^n \cdot ire^{it} dt \\ &= \frac{1}{2\pi}\sum_{n=-N}^{-1}\int_0^{2\pi} a_n \left(re^{it}\right)^{n+1} dt = a_{-1};\end{aligned}$$

that is,

$$\frac{1}{2\pi i}\int_{\partial B_r(z_0)} f dz = \mathrm{Res}_f(z_0). \tag{3.21}$$

Thus, the line integral only detects the term $a_{-1}(z-z_0)^{-1}$. As in the holomorphic case, this fact is generally true when we take the index of a cycle into account. This is recorded in the following residue theorem, whose proof will make use of the model case (3.21).

Theorem 3.22. Let f be a meromorphic function on U and let E be the set of the poles of f. If γ is a cycle in $U \setminus E$ such that

$$\mathrm{Ind}_\gamma(z) = 0 \text{ for all } z \in \mathbb{C} \setminus U, \tag{3.23}$$

then we have

$$\frac{1}{2\pi i}\int_\gamma f dz = \sum_{p \in E} \mathrm{Res}_f(p) \cdot \mathrm{Ind}_\gamma(p). \tag{3.24}$$

We remark that the sum in (3.24) is essentially finite, which we will discuss in the proof. Another remark for the condition (3.23) is that a similar condition appears in the Cauchy theorem, see (2.39) in Theorem 2.38. Since every holomorphic function is particularly meromorphic, Theorem 3.22 can be viewed as a generalization of Theorem 2.38 (1). In fact, we will use Theorem 2.38 in the proof. The idea is that we can apply the Cauchy theorem to the difference of the cycle γ and the sum of some circles around the poles whose indices (of γ) are non-zero. See Figure 6.

Proof. We consider

$$E_\gamma := \{p \in E : \mathrm{Ind}_\gamma(p) \neq 0\},$$

the set of the poles that we care about. Note that if we let U_γ be the unbounded connected component of $\mathbb{C} \setminus \gamma$, we have $E_\gamma \subseteq \mathbb{C} \setminus U_\gamma$, a compact set. Since E is a discrete set, in particular, this implies that E_γ is finite, say $E_\gamma = \{p_n : n = 1, \cdots, N\}$.

For each p_n, since $p_n \notin \gamma$, we can find $r_n > 0$ such that $\overline{B}_{r_n}(p_n) \subseteq U \setminus \gamma$. On the open set

$$U' := U \setminus \left(E \cup \bigcup_{n=1}^{N} \overline{B}_{r_n/2}(p_n)\right),$$

f is a holomorphic function. We want to apply the Cauchy theorem 2.38 on this open set, so we look at the cycle

$$\Gamma := \gamma - \sum_{n=1}^{N} \mathrm{Ind}_\gamma(p_n) \cdot \partial B_{r_n}(p_n),$$

where we recall that each $\partial B_{r_n}(p_n)$ is oriented counterclockwise. For any $z \in \mathbb{C} \setminus U'$, there are three situations.

(1) $z \in \mathbb{C} \setminus U$,

(2) $z \in E \setminus E_\gamma$, or

(3) $z \in \overline{B}_{r_j/2}(p_j)$ for some j.

If $z \in \mathbb{C} \setminus U$, then obviously $\mathrm{Ind}_\Gamma(z) = 0$. If $z \in E$, we have

$$\mathrm{Ind}_\Gamma(z) = \mathrm{Ind}_\gamma(z) - \sum_{n=1}^{N} \mathrm{Ind}_\gamma(p_n) \cdot \mathrm{Ind}_{\partial B_{r_n}(p_n)}(z) = 0 - 0 = 0,$$

where the second zero comes from the fact that z and p_n's are in different components of $U \setminus \gamma$, so $z \notin \overline{B}_{r_n/2}(p_n)$ for all n. If $z \in \overline{B}_{r_j/2}(p_j)$ for some j, then

$$\mathrm{Ind}_\Gamma(z) = \mathrm{Ind}_\gamma(z) - \sum_{n=1}^{N} \mathrm{Ind}_\gamma(p_n) \cdot \mathrm{Ind}_{\partial B_{r_n}(p_n)}(z) = \mathrm{Ind}_\gamma(p_j) - \mathrm{Ind}_\gamma(p_j) \cdot 1 = 0.$$

In conclusion, $\mathrm{Ind}_\Gamma(z) = 0$ for all $\mathbb{C} \setminus U'$.

Now we can use the Cauchy theorem to finish the proof. Since f is holomorphic on U' and $\mathrm{Ind}_\Gamma(z) = 0$ for all $\mathbb{C} \setminus U'$, Theorem 2.38 implies $\int_\Gamma f dz = 0$. Based on the definition of Γ, this means

$$\frac{1}{2\pi i} \int_\gamma f dz = \frac{1}{2\pi i} \sum_{n=1}^{N} \mathrm{Ind}_\gamma(p_n) \int_{\partial B_{r_n}(p_n)} f dz.$$

For each n, the model case (3.21) implies $\frac{1}{2\pi i} \int_{\partial B_{r_n}(p_n)} f dz = \mathrm{Res}_f(p_n)$. Combining these then leads to (3.24). □

The main applications of the residue theorem (Theorem 3.22) usually include two parts. The first one is its usage of calculating some definite integrals; the second is about counting zeros and poles. We leave the latter to the next section, and see some examples of integrals in the rest of this section.

We start from a simple observation about calculating residues. When f has a simple pole at p, then we know

$$f(z) = \frac{a_{-1}}{z - p} + h_1(z) = \frac{1}{h_2(z)}$$

in some $\hat{B}_r(p)$, where h_1 and h_2 are holomorphic in $B_r(p)$ and we know that h_2 has a simple zero at p with $h_2'(p) \neq 0$. Thus, we get two ways to calculate a_{-1}. That is,

$$\mathrm{Res}_f(p) = \lim_{z \to p}(z-p)f(z) = \frac{1}{h_2'(p)}.$$

One can derive similar formulas when the order of a pole is higher, and we record it as a lemma.[ex]

Lemma 3.25. Suppose f has a pole of order N at z_0, say

$$f(z) = \frac{h(z)}{(z-z_0)^N}$$

in $\hat{B}_r(z_0)$ for some $r > 0$ where h is a holomorphic function on $B_r(z_0)$. Then

$$\mathrm{Res}_f(z_0) = \frac{1}{(N-1)!}h^{(N-1)}(z_0).$$

Now we look at some examples. We first look at integrals of the type

$$\int_{-\infty}^{\infty} f(x)dx. \tag{3.26}$$

Complex analysis can help if the function f can extend to a meromorphic function with suitable decay at infinity.

Example 3.27. We look at an example, which asks the value of

$$\int_{-\infty}^{\infty} \frac{dx}{1+x^4}.$$

The idea is to approximate the real line by the boundary of a large half disc. Here we use the half disc in the upper half plane. Note that the function $f(z) := 1/(1+z^4)$ has four poles, two of which are in the upper half plane. They are $p_1 = e^{i\pi/4}$ and $p_2 = e^{3i\pi/4}$. They are both simple poles, so we can calculate

$$\mathrm{Res}_f(p_1) = \lim_{z \to p_1} \frac{1}{(1+z^4)'} = \frac{1}{4p_1^3} = \frac{1}{4}e^{-3i\pi/4}$$

and similarly

$$\mathrm{Res}_f(p_2) = \lim_{z \to p_2} \frac{1}{(1+z^4)'} = \frac{1}{4p_2^3} = \frac{1}{4}e^{-i\pi/4}.$$

Hence, if we let the upper circular arc be S_R, the residue theorem (Theorem 3.22) implies

$$\int_{-R}^{R} f(x)dx + \int_{S_R} f dz = 2\pi i \left(\frac{1}{4}e^{-3i\pi/4} + \frac{1}{4}e^{-i\pi/4}\right). \tag{3.28}$$

The key is to deal with the additional term $\int_{S_R} f dz$ when $R \to \infty$. In this case, note that

$$\left|\int_{S_R} f dz\right| \leq 2\pi R \cdot \frac{1}{R^4 - 1},$$

which tends to 0 as $R \to \infty$. Thus, (3.28) implies

$$\int_{-\infty}^{\infty} f(x)dx = 2\pi i \left(\frac{1}{4}e^{-3i\pi/4} + \frac{1}{4}e^{-i\pi/4}\right) = \frac{\pi}{\sqrt{2}}.$$

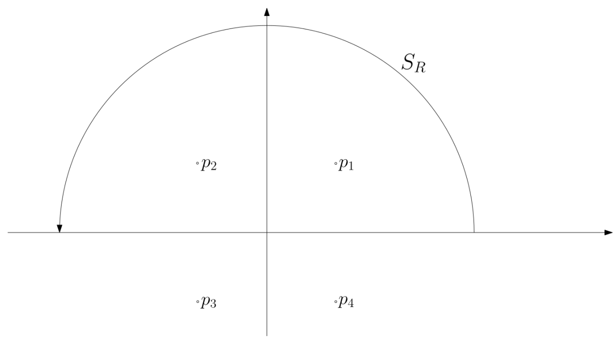

FIGURE 8. A semicircle enclosing two poles.

Observe that the path we choose in Example 3.27 works when the function f satisfies a decaying estimate[ex]

$$f(z) = o\left(\frac{1}{z}\right) \text{ as } z \to \infty. \tag{3.29}$$

Example 3.30 (Fourier transform)**.** We look at the integral

$$\int_{-\infty}^{\infty} f(x)e^{ix}dx, \tag{3.31}$$

which is still of the form (3.26), but it is related to the Fourier transform of the function f. We calculate an example in Example 2.31, and now we can look at more general functions. When f can be extended to a meromorphic function with finitely many poles away from the real line, we can calculate (3.31) in terms of the residues of $f(z)e^{iaz}$ when

$$f(z) = O\left(\frac{1}{z}\right) \tag{3.32}$$

as $z \to \infty$. Note that this is a weaker condition compared with (3.29). To this end, we choose a different path. Consider the square of side length L, with vertices L, $L + 2Li$, $-L + 2Li$, and $-L$ in $\mathbb{C}$. We estimate the top side and the right side. This is similar to the path in Figure 5, but the difference is that all the sides are growing in the case here.

For the top side from $L + 2Li$ to $-L + 2Li$, we can parametrize it by $\gamma_{\text{top}}(t) = -tL + 2Li$ for $t \in [-1, 1]$. Then for L large enough,

$$\left|\int_{\gamma_{\text{top}}} f(z)e^{iz}dz\right| = \left|\int_{-1}^{1} f\left(-tL + 2Li\right) e^{i(-tL+2Li)} \cdot (-L)\, dt\right| \le 2 \cdot \frac{C}{2L} \cdot e^{-2L} \cdot L, \tag{3.33}$$

where we use (3.32) to get some $C < \infty$ such that $|f(z)| \le C/|z|$ when $|z|$ is large.

For the right side from L to $L + 2Li$, we can parametrize it by $\gamma_{\text{right}}(t) = L + tLi$ for $t \in [0, 2]$. Then for L large enough,

$$\left|\int_{\gamma_{\text{right}}} f(z)e^{iz}dz\right| = \left|\int_{0}^{2} f\left(L + tLi\right) e^{i(L+tLi)} \cdot (Li)\, dt\right| \le \frac{C}{L}\int_0^2 e^{-tL}Ldt = \frac{C}{L}\left(1 - e^{2L}\right). \tag{3.34}$$

Note that both (3.33) and (3.34) tend to 0 as $L \to \infty$. One can estimate the left side similarly like (3.34). Thus, the residue theorem (Theorem 3.22) implies

$$\int_{-\infty}^{\infty} f(x)e^{ix}dx = 2\pi i \cdot \sum_{\mathrm{Im}(p)>0} \mathrm{Res}_h(p)$$

where $h(z) := f(z)e^{iaz}$.

3.4. **Argument principle and other applications.** We now talk about other applications of the residue theorem. The first one is called the argument principle. It counts the number of zeros and poles inside a closed path.

In the following, we talk about the simplest case when the curve γ is a simple closed path. That is, it is a closed embedded curve which is piecewise smooth. Thus, by Theorem 2.7, we know that $\mathbb{C} \setminus \gamma$ has two components, one of which is bounded, called Ω_γ. We assume the curve γ is positively oriented so that for any $z \in \Omega_\gamma$, we have $\mathrm{Ind}_\gamma(z) = 1$, and for $z \in \mathbb{C} \setminus (\gamma \cup \Omega_\gamma)$, $\mathrm{Ind}_\gamma(z) = 0$. In this setting, we can count the number of the zeros and the poles of a meromorphic function using a line integral.

Theorem 3.35 (Argument principle)**.** Let γ be a simple closed path such that both γ and the bounded component Ω_γ it bounds are contained in an open set U. Let f be a meromorphic function on U such that f does not have a pole or a zero on γ. If we write

$$\begin{aligned} N_0(\gamma, f) &:= \text{the number of zeros of } f \text{ in } \Omega, \text{ and} \\ N_\infty(\gamma, f) &:= \text{the number of poles of } f \text{ in } \Omega, \end{aligned}$$

both counted with multiplicities, then we have

$$\frac{1}{2\pi i}\int_\gamma \frac{f'(z)}{f(z)}dz = N_0(\gamma, f) - N_\infty(\gamma, f).$$

For example, $f(z) = (z-i)^2/(z-4)^3$ has a zero of order 2 at i, and hence i will contribute two to the number $N_0(\gamma, f)$ if i is enclosed by γ. Similarly, 4 will contribute three to the number $N_\infty(\gamma, f)$ if 4 is enclosed by γ.

Proof. First, we note that f only has finitely many zeros in $\overline{\Omega_\gamma} = \Omega_\gamma \cup \gamma$; otherwise it would vanish identically on the component of U that contains γ by Theorem 3.1, and hence would also vanish on γ. This in particular implies that f'/f is also a meromorphic function on U.

Note also that a pole of f'/f is either a zero of f or a pole of f. If $z_0 \in \Omega_\gamma$ is a zero of f of order n_0, then we can write

$$f(z) = h(z) \cdot (z - z_0)^{n_0}$$

where h is holomorphic near z_0 with $h(z_0) \neq 0$. Thus, near z_0,

$$\frac{f'(z)}{f(z)} = \frac{h'(z)\cdot(z-z_0)^{n_0} + h(z)\cdot n_0(z-z_0)^{n_0-1}}{h(z)\cdot(z-z_0)^{n_0}} = \frac{h'(z)}{h(z)} + \frac{n_0}{z-z_0}.$$

Since h'/h is holomorphic near z_0, we then derive

$$\mathrm{Res}_{f'/f}(z_0) = n_0. \tag{3.36}$$

On the other hand, if $p \in \Omega_\gamma$ is a pole of f of order N, then we can write

$$f(z) = \frac{g(z)}{(z-z_0)^N}$$

where g is holomorphic near z_0 with $g(z_0) \neq 0$. Thus, near z_0,

$$\frac{f'(z)}{f(z)} = \frac{\frac{g'(z)}{(z-z_0)^N} - \frac{g(z)\cdot N(z-z_0)^{N-1}}{(z-z_0)^{2N}}}{g(z)/(z-z_0)^N} = \frac{g'(z)}{g(z)} - \frac{N}{z-z_0}.$$

Since g'/g is holomorphic near p, we then derive

$$\mathrm{Res}_{f'/f}(p) = -N. \tag{3.37}$$

Combining (3.36) and (3.37) with the residue theorem (Theorem 3.22), the conclusion follows. □

Note that the argument principle gives topological/analytical restriction on the distributions of zeros and poles of a meromorphic function. We mention a few consequences of the argument principles in this aspect. First, it tells us a simple fact about the limit of a sequence of non-vanishing holomorphic functions.

Corollary 3.38. Let f_n be a sequence of non-vanishing holomorphic functions on a connected open set U. If f_n converges to a function f locally uniformly on U, then f is a holomorphic function that is either non-vanishing or identically zero.

Remember, from Corollary 2.20, we know that such a limit f must be holomorphic, and from Corollary 2.23, we also have convergence of their derivatives. Now, we know that such a limit cannot vanish unless it is trivial if the sequence doesn't. A variant of it says that if all f_n's omit a value $z_0 \in \mathbb{C}$, then the locally uniform limit f also omits z_0. None of these holds in the case of real differentiable functions, as we stress many times.

Proof. Suppose f is not identically zero. We will show that f is then non-vanishing on U.

If f had a zero at $z_0 \in U$, since f is not identically zero, Proposition 3.1 would imply that 0 is an isolated zero of f. Suppose the order of z_0 is n_0, and hence we can take $r > 0$ such that $\overline{B}_r(z_0) \subseteq U$ and Then Theorem 3.35 and the convergence of f_n's imply

$$n_0 = \frac{1}{2\pi i}\int_{\partial B_r(z_0)} \frac{f'(z)}{f(z)}dz = \lim_{n\to\infty} \frac{1}{2\pi i}\int_{\partial B_r(z_0)} \frac{f_n'(z)}{f_n(z)}dz = 0$$

where the last equality is because each f_n'/f_n is holomorphic since f_n is non-vanishing. This is a contradiction since $n_0 > 0$. □

Corollary 3.38 is sometimes referred as Hurwitz's theorem, which has another part about limits of injective holomorphic functions. It can be proven similarly or be derived from its first version (Corollary 3.38), so we state it as follows and leave it to Assignment 5.

Corollary 3.39 (Hurwitz)**.** Let f_n be a sequence of injective holomorphic functions on a connected open set U. If f_n converges to a function f locally uniformly on U, then f is a holomorphic function that is either injective or constant.

The argument principle can also be used to compare the roots of two holomorphic functions that are close enough to each other. An example is the following theorem by Rouché.

Corollary 3.40 (Rouché's theorem)**.** Let γ be a simple closed path such that both γ and the bounded component Ω it bounds are contained in an open set U. Let f and g be holomorphic functions on U. If $|f| > |g|$ on γ, that is,

$$|f(z)| > |g(z)| \text{ for all } z \in \gamma, \tag{3.41}$$

then f and $f+g$ have the same number of zeros in Ω (again, counted with multiplicities).

Proof. The proof is based on a continuity argument. We consider a family

$$f_t(z) := f(z) + t \cdot g(z)$$

of holomorphic functions on U for $t \in [0,1]$. The condition (3.41) implies that f_t does not have a zero on γ for all $t \in [0,1]$. Then by Theorem 3.35, we know that the number of zeros of f_t in Ω is

$$n_t := \frac{1}{2\pi i} \int_\gamma \frac{f_t'(z)}{f_t(z)} dz.$$

Thus, it remains to show $n_0 = n_1$.

Since $f_t(z)$ can be viewed as a continuous function on $[0,1] \times U \ni (t,z)$, n_t is a continuous function in t. Because for each $t \in [0,1]$, n_t is an integer, the continuity forces it to be a constant. In particular, $n_0 = n_1$. $\square$

Rouché's theorem can be viewed as a stability result for numbers of zeros. The result has many applications. For example, in Assignment 5, we will see that we can give another proof of the fundamental theorem of algebra based on Rouché's theorem. Second, it can really help us to count the number of roots in a bounded region. We discuss this in the following example.

Example 3.42. Let $f(z) = z^6 - 4z^2 + 3$. We know that it has six roots in $\mathbb{C}$, and we can use Rouché's theorem to get a rough range of them. In fact, we want to take $R > 0$ such that

$$\left|z^6\right| > \left|4z^2 + 3\right|$$

on ∂B_R so that we can apply Rouché's theorem to this circle. For this, the inequality

$$\left|z^6\right| > 4\left|z^2\right| + 3$$

suffices. For example, we can take $R = 2$. Thus, Rouché's theorem implies that f and z^6 have the same number of zeros in B_2, so we know that all the roots of f lie in B_2. This is the simplest situation, and one can even deal with equations involving other functions, like exponential or trigonometric functions.

The last application that we are going to see is the open mapping theorem, which has been mentioned in Corollary 3.6 as a consequence of the inverse function theorem. Using Rouché's theorem, we can see another simple proof of the open mapping theorem.

Another proof of Corollary 3.6. Suppose $z_0 \in U$ and $y_0 = f(z_0)$. Since f is not a constant map, Proposition 3.1 imply that z_0 is an isolated zero of $f - y_0$. Thus, by taking a small enough $r > 0$, we can make $\overline{B}_r(z_0) \subseteq U$ and z_0 is the unique zero of $f - y_0$ in $\overline{B}_r(z_0)$. In particular, $f - y_0$ is non-vanishing on $\partial B_r(z_0)$ and hence

$$R := \inf_{\partial B_r(z_0)} |f - y_0| > 0.$$

Therefore, for any complex number w with $|w| < R$, Rouché's theorem (Corollary 3.40) implies that $f - y_0 + w$ also has at least a zero in $B_r(z_0)$. This means that $f(U)$ contains $B_R(y_0)$, and the conclusion follows. $\square$

A more quantitative version of the open mapping theorem, Bloch's theorem, can be found in Assignment 5. It can also be viewed as a consequence of Rouché's theorem, and we state it here.

Theorem 3.43 (Bloch's theorem)**.** Let f be a holomorphic function on U which contains $B_R(z_0)$. Then there exists a universal constant $C > 0$ such that $f(B_R(z_0))$ contains a ball of radius $CR|f'(z_0)|$.

4. **Examples and techniques of constructing meromorphic functions**

We will see more examples of holomorphic functions and meromorphic functions with some of their applications in this section. It includes some explicit examples and some ways to construct holomorphic or meromorphic functions with some prescribed information.

4.1. **Logarithm and entire functions.**

We will talk about an important extension of a real-valued function, the logarithm, $\log x$. Using it, we will briefly talk about how to look at the growth of a holomorphic function.

4.1.1. *Logarithm function on a simply connected domain.* When $x \in \mathbb{R}_{>0}$, we know that there is a well-defined analytic function $\log x$ such that

$$e^{\log x} = \log e^x = x. \tag{4.1}$$

It turns out that it is not straightforward to extend this to a well-defined continuous function $\log z$ for $z \in \mathbb{C} \setminus \{0\}$. An obvious reason is that for any $k \in \mathbb{Z}$, we have

$$e^{z+2k\pi i} = e^z \cdot \left(e^{2\pi i}\right)^k = e^z.$$

Thus, to make it well-defined, one has to fix a choice of the argument. For example, we can define $\log 1 = 0$ as in the real case. However, if we would like to extend this to the whole $\mathbb{C} \setminus \{0\}$, even just continuously, we will have to define

$$\log e^{it} = it$$

for $t \in [0, 2\pi)$. The problem then comes when t tends to 2π. On the one hand, continuity forces $\log e^{2\pi i} = 2\pi$. On the other hand, if the function is well-defined, we have

$$\log e^{2\pi i} = \log 1 = 0.$$

Thus, it is impossible to have a well-defined continuous log function on $\mathbb{C} \setminus \{0\}$.

We can make this more concrete if we require $\log z$ to be a holomorphic function. In fact, it has to be holomorphic if it exists as the inverse of the exponential function, based on the inverse function theorem 2.47. We can use this property to say that we cannot define log on a domain where there is a closed path that winds the origin, that is, a closed path γ such that $\mathrm{Ind}_\gamma(0) = 1$.

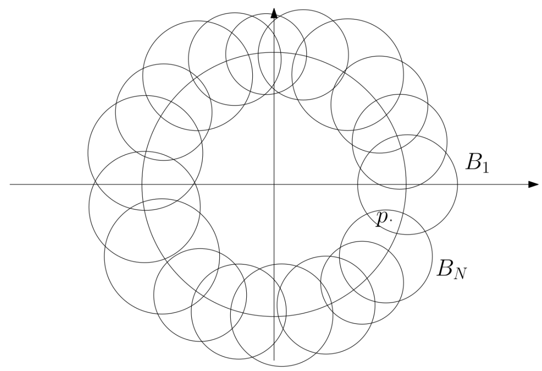

FIGURE 9. We know that one can always define a logarithm function locally, so one can extend it along a path by looking at more and more balls. The argument above tells us that if we start the extension from the ball B_1 and end in the ball B_N, then we will get two different values at p which is in their intersection.

In fact, if $\log\colon U \to \mathbb{C}$ is a holomorphic function satisfying (4.1), then differentiating (4.1) (using the complex variable z) leads to

$$1 = e^{\log z} \cdot \frac{d}{dz} \log z,$$

implying that

$$\frac{d}{dz} \log z = \left(e^{\log z}\right)^{-1} = \frac{1}{z}$$

on U. Then, if there is a closed path γ in U with $\mathrm{Ind}_\gamma(0) = 1$, then by the residue theorem (Theorem 3.22), as a meromorphic function on $\mathbb{C}$, $1/z$ satisfies

$$\int_\gamma \frac{1}{z} dz = 2\pi i,$$

while the fundamental theorem of calculus implies

$$\int_\gamma \frac{1}{z} dz = \int_\gamma \left(\frac{d}{dz} \log z\right) dz = 0$$

since γ is closed. We then get a contradiction. Formally, we would say that it is not possible to analytically continuate a logarithm function along the closed path γ. We will talk about this in more detail in Section 5.5.3.

We now formalize a related topological property, and will show that it characterizes regions on which we can define $\log$.

Definition 4.2. A subset S of $\mathbb{C}$ is called **simply connected** if it is connected and any closed curve in S is homotopic to a constant map in S.

Recall the definition of homotopy in (2.43). We will especially talked about simply connected open subsets. Thus, an open set U is simply connected if any closed curve in U can be continuously contracted to a point, and the whole contraction process happens in U.

Example 4.3. We mention some examples of domains in $\mathbb{C}$.

(1) A convex set is simply connected.

(2) More generally (compared with (1)), a star-shaped domain is simply connected. In fact, if U is a star-shaped open set in $\mathbb{C}$, say every point p can be connected to the origin by a segment in U, which is $\overline{Op}$, then every closed curve is homotopic to the constant map $\gamma(t) = p$.

(3) A typical example which is not simply connected is a punctured disc. For example, a round circle centered at the origin is not homotopic to a constant map in $\mathbb{C} \setminus 0$.

We recall that two homotopic closed paths satisfy (2.45) as a consequence of the Cauchy theorem. It turns out that it is also true for any two homotopic paths with the same endpoints. This formulation is helpful because we have the following characterization of simply connected domains.[ex]

Proposition 4.4. A subset S in $\mathbb{C}$ is simply connected if and only if it is connected and any two curves with the same endpoints are homotopic to each other. As a consequence, combined with (2.45), it implies that for a holomorphic function f on S and two paths γ_1 and γ_2 with the same endpoints, we have

$$\int_{\gamma_1} f dz = \int_{\gamma_2} f dz.$$

As a corollary, one can generalize Lemma 1.8 to a simply connected domain.[ex] The proof idea is to find a primitive directly, as performed in Lemma 1.8.

Corollary 4.5. Let U be a simply connected open set in $\mathbb{C}$ and let u be a harmonic function on U. Then there exists a unique harmonic conjugate of u up to a constant.

On a simply connected domain away from the origin, we can then have a well-defined logarithm. In general, we can take the logarithm of a function if its domain is simply connected and the function is non-vanishing.

Theorem 4.6. Let U be a simply connected open set in $\mathbb{C}$ and f be a non-vanishing holomorphic function on U. Then there exists a holomorphic function $\log f$ such that $e^{\log f} = f$.

Such a function $\log f$ is called a **branch** of the logarithm of f. One can also see that such a branch is unique up to an integer multiple of $2\pi i$. When $f(z) = z$ and $U = \mathbb{C} \setminus (-\infty, 0]$, we can then find a branch $g(z) = \log z$ such that $g(1) = 0$. Its value is

$$\log\left(re^{it}\right) = \log r + it$$

if $r \in \mathbb{R}_{\geq 0}$ and $t \in (-\pi, \pi)$. This is called the **principal branch** of the logarithm. One can also find out its power series expansion[ex]

$$\log z = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(z-1)^n \tag{4.7}$$

for $z \in B_1(1)$.

Proof of Theorem 4.6. We do the case when $f(z) = z$, and leave the general case as an exercise.[ex] In this special case, the assumption implies that U does not contain the origin.

Fix a point $p \in U$. Then given any $z \in U$, we find a path $\gamma\colon [0,1] \to U$ with $\gamma(0) = p$ and $\gamma(1) = z$, and define

$$g(z) := \int_\gamma \frac{dz}{z} + C_p$$

for some $C_p \in \mathbb{C}$ to be defined. Note that this integral is valid since $1/z$ is holomorphic on U. Moreover, g is a well-defined function, since Proposition 4.4 implies that the expression does not depend on the path γ as long as it connects p to z. Thus, $g(z)$ is a continuous function whose derivative is $1/z$ (cf. the proof of Corollary 2.15), and hence g is a holomorphic function.

It remains to show that $e^{g(z)} = z$, which depends on our choice of C_p. To see this, using $g'(z) = 1/z$, we have

$$\left(ze^{-g(z)}\right)' = e^{-g(z)} + ze^{-g(z)} \cdot \left(-\frac{1}{z}\right) = e^{-g(z)}\,(1-1) = 0,$$

which implies that $ze^{-g(z)}$ is constant by the connectivity of U. Thus, we have

$$ze^{-g(z)} = pe^{-g(p)} = p \cdot e^{-C_p}. \tag{4.8}$$

We can then choose $C_p \in \mathbb{C}$ such that $e^{C_p} = p$, and then (4.8) implies $e^{g(z)} = z$. Note that the choice of C_p is unique up to an integer multiple of $2\pi i$. □

Besides its many applications, the existence of a branch of logarithm on a simply connected domain allows us to define inverses of many other common functions. We record them here and omit the details.[ex]

Corollary 4.9. Let U be a simply connected open set in $\mathbb{C}$.

(1) If $f\colon U \to \mathbb{C} \setminus \{0\}$ is a holomorphic function, then there exists a holomorphic function $f^{1/n}$ such that $\left(f^{1/n}\right)^n = f$.

(2) If $g\colon U \to \mathbb{C} \setminus \{\pm 1\}$ is a holomorphic function, then there exists a holomorphic function $\cos^{-1} g$ such that $\cos\left(\cos^{-1} g\right) = g$.

As an application, we can prove the little Picard theorem. It is a generalization of the fact that a non-constant entire function has a dense image (Assignment 4).

Theorem 4.10 (Little Picard)**.** If $f\colon \mathbb{C} \to \mathbb{C}$ is a non-constant entire function, then f omits at most one point in $\mathbb{C}$.

First, note that this theorem is sharp in view of the example given by the exponential function, missing exactly the origin in its image. Also, note that Theorem 4.10 is related to the Casorati–Weierstrass theorem (Proposition 3.17). In fact, as we see in Assignment 4 (cf. Example 3.13 (5) and (6)), if $f(z)$ is a non-constant entire function, then f has a pole or an essential singularity at ∞. Hence, either f is a polynomial (so its image is $\mathbb{C}$), or Proposition 3.17 implies that $f(\mathbb{C})$ is dense in $\mathbb{C}$. Theorem 4.10 then strengthens this result by providing more precise information of the image of such a function.

Proof of Theorem 4.10. We will assume the function f misses at least two points, and prove that it must be constant. After applying a linear-type transformation (of the form $az+b$), we may assume that f misses at least 0 and 1.

Since $f(\mathbb{C})$ does not include 0 and $\mathbb{C}$ is simply connected, by Theorem 4.6, we can define a holomorphic function $g := \log f$ on $\mathbb{C}$. Since $f(\mathbb{C})$ does not include 1, $g(\mathbb{C})$ then does not include any integer multiple of $2\pi i$. In particular, the image of $g/2\pi i$ does not include ± 1, so Corollary 4.9 implies that we can find a holomorphic function

$$h = \cos^{-1} \frac{1}{2\pi i} g = \cos^{-1} \left(\frac{1}{2\pi i} \log f \right)$$

on $\mathbb{C}$. Then notice that if $h(z) = \pm i \cosh^{-1} m + 2\pi n$ for some $m, n \in \mathbb{Z}$, we would have

$$f(z) = e^{2\pi i \cos h(z)} = e^{2\pi i \cos\left(\pm i \cosh^{-1} m\right)} = e^{2\pi i \cdot m} = 1,$$

where we used the formula $\cos(iz) = \cosh z$. Hence, the image of h does not contain

$$\left\{ \pm i \cosh^{-1} m + 2\pi n : m, n \in \mathbb{Z} \right\}.$$

Recalling the formula[ex]

$$\cosh^{-1} m = \log\left(m + \sqrt{m^2 - 1}\right),$$

we conclude that $h(\mathbb{C})$ does not contain an open ball of radius larger than a finite number C_h.

Now, we could apply Bloch's theorem (Theorem 3.43) to conclude that there exists $C > 0$ such that

$$CR \cdot \left|h'(z)\right| \leq C_h$$

for any $z \in \mathbb{C}$ and $R > 0$, since $h\left(B_R(z)\right) \subseteq h(\mathbb{C})$ does not contain an open ball of radius at most C_h. The relation means

$$\left|h'(z)\right| \leq \frac{C_h}{C} \cdot \frac{1}{R}$$

for any $R > 0$, so h' vanishes identically, and hence h is a constant. This finishes the proof. □

The little Picard theorem tells us about how dense the image of an entire function has to be. There is a stronger result, the great Picard theorem, stating more precise information of the behavior of a holomorphic function near an essential singularity. See Theorem 5.24.

4.1.2. *Growth of zeros.* We study the density property of the image of an entire function in the previous section. Next, we look at the growth of an entire function. It turns out that the behavior of a holomorphic function on the boundary of a disc is related to zeros in the interior of the disc. This is encoded in the following formula by Jensen.

Theorem 4.11 (Jensen's formula)**.** Let f be a holomorphic function on an open set U and suppose $\overline{B}_R \subseteq U$ for some $R > 0$. If $f(0) \neq 0$ and $f|_{\partial B_R}$ is non-vanishing, then

$$\log|f(0)| = \sum_{z \in f^{-1}(0) \cap B_R} \log \frac{|z|}{R} + \frac{1}{2\pi} \int_0^{2\pi} \log\left|f\left(Re^{it}\right)\right| dt,$$

where the zeros are counted with multiplicities.

In fact, the formula also extends to the case of meromorphic functions. Here, since we are going to look at entire functions later, we focus on holomorphic functions. Also note that if f is non-vanishing on the disc, the formula follows directly from the mean-value property of $\log|f|$, which is the real part of a holomorphic function $\log f$. This is one way to prove the formula, which we will see in the following.

Proof. We do the case when $R = 1$, and the general case follows from scaling (and one can also get a version when the center is not the origin).

The trick we are going to use is to multiply f by a good function so that the number of its zeros decreases. That is, for $z_0 \in B_1$, consider the function

$$h_{z_0}(z) := \frac{z_0 - z}{1 - \overline{z_0} z}.$$

In Assignment 1, the following properties are proven.

(1) $h_{z_0}(0) = z_0$,

(2) h_{z_0} has a unique zero z_0 in B_1, and

(3) $|h_{z_0}| = 1$ on ∂B_1.

As a result, we can look at the function

$$F := f \cdot \prod_{z \in f^{-1}(0) \cap B_1} h_z^{-1},$$

a non-vanishing function on $\overline{B}_1$. We can then apply the mean-value theorem to the real part of $\log F$ to get

$$\log|F(0)| = \frac{1}{2\pi} \int_0^{2\pi} \log\left|F\left(e^{it}\right)\right| dt. \tag{4.12}$$

Note that for each $z \in f^{-1}(0) \cap B_1$, the properties (1), (2), and (3) imply

$$\log|h_z(0)| = \log|z| = \log|z| + \frac{1}{2\pi} \int_0^{2\pi} \log\left|h_z\left(e^{it}\right)\right| dt. \tag{4.13}$$

Combing (4.12) with (4.13) for each $z \in f^{-1}(0) \cap B_1$, the conclusion follows. $\square$

There are many applications of Jensen's formula. For example, one can use it to prove the Liouville theorem and the fundamental theorem of algebra again. A typical usage of Jensen's formula when applied to entire functions is to relate the growth of an entire function with the density of its zeros. We state one type of this result here.

Corollary 4.14. Let f be a non-constant entire function such that for some $C, \alpha < \infty$,

$$|f(z)| \leq Ce^{C|z|^\alpha} \tag{4.15}$$

for all $z \in \mathbb{C}$. Then there exists $C_Z = C_Z(f) < \infty$ such that f contains at most $C_Z R^\alpha$ zeros in B_R.

For a function satisfying (4.15), we say f is **of order at most** α. For such a holomorphic function, if we let

$$N_0(B_R, f) := \text{the number of zeros of } f \text{ in } B_R,$$

the theorem says that the number $N_0(B_R, f)$ of zeros of f does not grow faster than R^α.

Proof. First, we assume $f(0) \neq 0$. Then applying Theorem 4.11 with radius $2R$ and putting the integral term on one side, we have

$$\log|f(0)| + \sum_{z \in f^{-1}(0) \cap B_{2R}} \log \frac{2R}{|z|} = \frac{1}{2\pi} \int_0^{2\pi} \log\left|f\left(2Re^{it}\right)\right| dt.$$

On the one hand, we can use the growth assumption (4.15) to estimate

$$\frac{1}{2\pi} \int_0^{2\pi} \log\left|f\left(2Re^{it}\right)\right| dt \leq \frac{1}{2\pi} \cdot 2\pi \cdot \log Ce^{C\cdot(2R)^\alpha} = \log C + 2^\alpha C \cdot R^\alpha.$$

On the other hand, for the summation term, we can just look at those zeros in B_R, and get

$$\sum_{z \in f^{-1}(0) \cap B_{2R}} \log \frac{2R}{|z|} \geq \sum_{z \in f^{-1}(0) \cap B_{R}} \log \frac{2R}{|z|} \geq \log 2 \cdot N_0(B_R, f).$$

Combining these, we get

$$\log|f(0)| + \log 2 \cdot N_0(B_R, f) \leq \log C + 2^\alpha C \cdot R^\alpha,$$

which implies

$$N_0(B_R, f) \leq \frac{1}{\log 2}\left(2^\alpha C R^\alpha - \log|f(0)|\right),$$

the desired growth.

If $f(0) = 0$, since f is non-constant, 0 must be an isolated zero of f. Thus, for small enough $r > 0$, we have $f(r) \neq 0$. Applying the argument above to the point r, we get the same form of estimates for large enough R.[ex] □

4.2. **Approximating meromorphic functions.** We understand how the number of the zeros of an entire function may grow. We next can ask for refined information, such as patterns or distribution of those zeros. For example, given a set of points, can we find an entire function whose zero set is exactly the given set? From Proposition 3.1, we know that there are some restrictions. That is, if f is an entire function and its zero set $Z(f)$ is not the whole complex plane $\mathbb{C}$, then $Z(f)$ only contains discrete points. Thus, we have a necessary condition for a set to be the zero set of an entire function. We will see that the condition is also sufficient, due to Weierstrass, at the end of the seciotn.

We will start by Runge's approximation theorem. Using it, we will be able to derive two results about prescribing vanishing or singular behavior of a holomorphic functions by Mittag-Leffler and Weierstrass. For the Weierstrass theorem, we will in fact use a more direct method to prove it.

4.2.1. *Runge's approximation theorem.* The question Runge asked is the possibility of approximating a holomorphic function by rational functions or polynomial functions. This reminds us of the Weierstrass approximation theorem, which says that every continuous function defined on a closed interval can be uniformly approximated as closely as desired by a polynomial function.

For holomorphic functions, recall that we know they are analytic (Theorem 2.17), so it is true that they can locally be approximated by polynomial functions. In general, we would like to understand what condition is necessary or sufficient to guarantee such an approximation on a compact set. One can see that using polynomials is not always enough by looking at $1/z$ on the compact set ∂B_1, along which every polynomial has trivial line integral. Restrictions come from non-trivial topology of the set K, as in the case of ∂B_1, its complement is not connected. If we allow rational functions, that it works for all compact sets.

Theorem 4.16 (Runge)**.** Let K be a compact set in $\mathbb{C}$ and let f be a holomorphic function defined on a neighborhood of K.

(1) f can be approximated uniformly on K by rational functions whose poles are away from K.

(2) f can be approximated uniformly on K by polynomial functions if $\mathbb{C} \setminus K$ is connected.

We will first prove the first part, that is, Theorem 4.16(1). After that, we will discuss how to approximate a rational function by another with prescribed poles.

Proof of Theorem 4.16(1)*.* Suppose f is a holomorphic function on U which is a bounded open neighborhood of K. We will first divide the boundary of a (slightly smaller) neighborhood of K.

Let $d := d_{\mathrm{Euc}}(K, \partial U)/100$. Consider a grid formed by closed solid squares with sides parallel to the axes and of length d. Let $\mathcal{R}$ be the set $\{R_1, \cdots, R_N\}$ that collects all squares that intersect K, and let $\gamma_1, \cdots, \gamma_M$ be those sides of squares in $\mathcal{R}$ that do not belong to two adjacent squares. By the construction, each $\gamma_m \subseteq U \setminus K$.

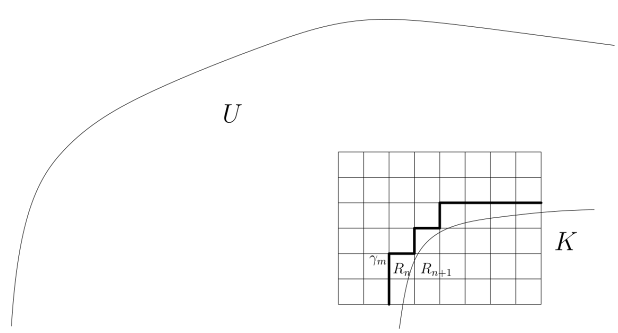

FIGURE 10. A possible choice of $\mathcal{R}$.

We let

$$\partial \mathcal{R} := \bigcup_{n=1}^{N} \partial R_n$$

be the union of the boundaries of R_n's. For eah $z \in K \setminus \partial\mathcal{R}$, we can find $R_{n_z} \in \mathcal{R}$ such that $z \in \mathrm{int} R_{n_z}$, and hence

$$\frac{1}{2\pi i}\int_{\partial R_n} \frac{f(w)}{w-z}dw = \begin{cases} f(z) & \text{if } n = n_z \\ 0 & \text{otherwise} \end{cases}.$$

Hence, summing over $n = 1, \cdots, N$ so that the contribution from adjacent sides cancel, we get

$$\sum_{m=1}^{M} \frac{1}{2\pi i}\int_{\gamma_m} \frac{f(w)}{w-z}dw = f(z) \tag{4.17}$$

for $z \in K \setminus \partial\mathcal{R}$. For $z \in K \cap \partial\mathcal{R}$, since $\gamma_m \subseteq U \setminus K$ for each m, the left hand side of (4.17) still makes sense and defines a holomorphic function (Proposition 2.4). The continuity then implies that (4.17) is true for all $z \in K$.

Now we use the expression (4.17) to get an approximation. For each γ_m, since it is a compact segment that is disjoint from K, we can find a closed ball $\overline{B}_{r_m}(z_m)$ that contains γ_m and is still disjoint from K. Then the function

$$f_m(z) := \frac{1}{2\pi i}\int_{\gamma_m} \frac{f(w)}{w-z}dw$$

is holomorphic on $\mathbb{C} \setminus \overline{B}_{r_m}(z_m)$, and Theorem 3.9 implies that we have a Laurent series expansion

$$f_m(z) = \sum_{n=-\infty}^{\infty} a_n^m (z - z_m)^n$$

where the convergence is locally uniform in $\mathbb{C} \setminus \overline{B}_{r_m}(z_m)$. Therefore, the sequence

$$\sum_{m=1}^{M} \sum_{n=-j}^{j} a_j^m (z - z_m)^j \tag{4.18}$$

is a sequence of rational functions that converges uniformly to f on K (since the sum of $f_1, \cdots, f_M$ is f on K). Note that all the poles z_n's are in $\mathbb{C} \setminus K$. $\square$

Based on the construction (4.18) in the proof, to achieve the second part of the theorem, one would like to move the poles of rational functions to infinity. We will do this in a general way in the following result.

Proposition 4.19. Let K be a compact set in $\mathbb{C}$ and let U be a connected open set in $\hat{\mathbb{C}}$ such that $U \cap K = \emptyset$. Given $z_0 \in U$, a rational function with poles in U can be approximated uniformly on K by rational functions with exactly a pole at z_0.

Proof. We let $\mathcal{R}(z_0)$ be the set of rational functions with exactly a pole at z_0, that is,

$$\mathcal{R}(z_0) := \left\{ \sum_{j=0}^{N} \frac{p_j(z)}{(z-z_0)^j} : \text{each } p_j \text{ is a polynomial} \right\}$$

when $z_0 \in \mathbb{C}$, and $\mathcal{R}(z_0)$ is the set of polynomials when $z_0 = \infty$. Consider the subset

$$V := \left\{ w \in U : \frac{1}{z-w} \text{ can be approximated uniformly on } K \text{ by rational functions in } \mathcal{R}(z_0) \right\}.$$

Since a rational functional can be expressed by partial fractions, it suffices to show that $V = U$. By its definition, V contains z_0 automatically, and we will show that it is both closed and open.

The closedness is more direct. If $z_n \in V$ is a sequence and $z_n \to z_\infty \in U$ as $n \to \infty$, then $\frac{1}{z-z_n}$ converges uniformly to $\frac{1}{z-z_\infty}$ on K as $n \to \infty$. Thus, given $\varepsilon > 0$, we can find n large such that

$$\left| \frac{1}{z - z_\infty} - \frac{1}{z - z_n} \right| < \frac{\varepsilon}{2}$$

for $z \in K$, and we can find $R \in \mathcal{R}(z_0)$ with

$$\left| R(z) - \frac{1}{z - z_n} \right| < \frac{\varepsilon}{2}$$

for $z \in K$. Combining this, we see that $1/(z - z_\infty)$ can be approximated uniformly on K by functions in $\mathcal{R}(z_0)$, and hence $z_\infty \in V$. This proves that V is closed in U.

The difficulty mainly lies in proving the openness of V. Suppose $w_0 \in V$. We want to approximate $1/(z-w)$ by functions in $\mathcal{R}(w_0)$ for w close enough to w_0. We first deal with the case when $w_0 = \infty$, which means $\mathcal{R}(w_0)$ consists of polynomials. First, we take $R < \infty$ large such that $K \subseteq B_R$. For w close to w_0, in the sense that $|w|$ is large, say $|w| \geq 2R$, we get

$$\frac{1}{z-w} = -\frac{1}{w}\frac{1}{1 - z/w} = -\frac{1}{w}\sum_{n=0}^{\infty} \frac{z^n}{w^n}$$

with $|z/w|^n \leq (1/2)^n$. Thus, Weierstrass' M-test implies that the series converges uniformly on K, and hence we get $w \in \mathcal{R}(\infty)$.

The case when $w_0 \in \mathbb{C}$ is similar, using the technique in the proof of Proposition 2.4.[9] That is, we write

$$\frac{1}{z-w} = \frac{1}{z-w_0} \cdot \frac{z-w_0}{z-w} = \frac{1}{z-w_0} \cdot \frac{1}{1 - \frac{w-w_0}{z-w_0}}.$$

For such a $w_0 \in V \subseteq U$, take a positive $\varepsilon < d_{\mathrm{Euc}}(w_0, K)$. Then for $w \in B_{\varepsilon/2}(w_0)$, we have

$$\frac{1}{z-w} = \frac{1}{z-w_0} \sum_{n=0}^{\infty} \left(\frac{w-w_0}{z-w_0} \right)^n$$

with $\left| \frac{w-w_0}{z-w_0} \right|^n < (1/2)^n$. Weierstrass' M-test again implies that the series converges uniformly on K, and hence we get $w \in \mathcal{R}(w_0)$. This finishes the proof of the openness of V, and hence the proposition follows. $\square$

As a corollary, we can now prove the second part of Theorem 4.16.

[9] In fact, it is exactly the same proof after a change of variables $z \mapsto 1/(z_0 - w_0)$.

Proof of Thoemre 4.16(2). Given f, from Theorem 4.16(1), f can be approximated by rational functions with poles away from K. That is, given $\varepsilon > 0$, f can be $\varepsilon/2$-approximated on K by $R_1 + \cdots + R_N$, where R_n has exactly one pole at $z_n \in \mathbb{C} \setminus K$. By Proposition 4.19, each R_n can be $\varepsilon/(2N)$-approximated on K by a function in $\mathcal{R}(\infty)$, the set of polynomials. Combining these, we complete the proof. □

In fact, the same argument leads to a slightly stronger result. We record it here and leave the (repeated) arguments to exercises.

Corollary 4.20. Let K be a compact set in $\mathbb{C}$ and let f be a holomorphic function defined on a neighborhood of K. If S is a subset of $\hat{\mathbb{C}} \setminus K$ such that each component of $\hat{\mathbb{C}} \setminus K$ intersects S, then f can be approximated uniformly on K by rational functions whose poles are in S.

All of these results will be referred to as Runge's approximation theorems. Runge's theorems are very useful when one tries to construct holomorphic and meromorphic functions. We will see one example in the next section, and before that, we first show that it provides an analytic way to characterize simple connectivity (cf. Assignment 5).

Corollary 4.21. Let U be a bounded connected open set in $\mathbb{C}$. If $\mathbb{C} \setminus U$ is connected, then U is "holomorphically simply connected," in the sense that given any holomorphic function f on U and a closed path $\gamma \subseteq U$, $\int_\gamma f dz = 0$.

We know from Proposition 4.4 that a simply connected set is holomorphically simply connected. The converse of this can be obtained easily after we learn the Riemann mapping theorem later.

Proof of Corollary 4.21. Let f be a holomorphic function on U and γ be a closed path in U. Consider the compact set

$$K := \{z \in U : d_{\text{Euc}}(z, \mathbb{C} \setminus U) \geq \varepsilon\}$$

for small $\varepsilon > 0$. We want to show that $\mathbb{C} \setminus K$ is connected so that we can apply Runge's theorem. This is a purely topological fact, and we can do it directly.

To this end, suppose for a contradiction that we can write $\mathbb{C} \setminus K = V_1 \cup V_2$ where V_1 and V_2 are non-empty disjoint open sets. With this, consider

$$C_i := V_i \setminus U$$

for $i = 1, 2$. We claim that C_1 and C_2 are non-empty disjoint closed sets. If this is done, then we get a contradiction since $\mathbb{C} \setminus U = C_1 \cup C_2$ while $\mathbb{C} \setminus U$ is connected.

For the claim, it is clear that C_1 and C_2 are disjoint. To show that they are closed, suppose $z_n \in C_1$ is a sequence that converges to a point z_0. Then since U is open, we know $z_0 \in \mathbb{C} \setminus U$, so $z_0 \notin K$ by the definition of K. Hence $z_0 \in V_1 \cup V_2$, and the disjoint property implies $z_0 \in V_1$. This proves that C_1 is closed, and similarly so is C_2.

Finally, we show that C_1 is non-empty (and similarly so is C_2). If $C_1 = \emptyset$, then we would get $V_1 \subseteq U$. For any $w \in V_1$, since $w \notin K$, there exists $z \in \mathbb{C} \setminus U$ such that $|w - z| < \varepsilon$ and $\overline{wz} \subseteq \mathbb{C} \setminus K$.

The idea of the rest of the proof is that because $z \in V_2$ (since $V_1 \subseteq U$), a limiting point on $\overline{zw}$ would make a contradiction. In practice, one can consider

$$\bar{t} := \sup \{t \in [0, 1] : (1 - t)z + tw \in V_2\}$$

and see that it cannot belong to either V_1 or V_2. This finishes the proof of the claim.

Now, by Runge's theorem (Theorem 4.16(2)), we can approximate f uniformly on K by polynomials. Since γ is compact, by choosing ε small, we can assume $\gamma \subseteq K$. This then finishes the proof by the Cauchy theorem. □

4.2.2. *Mittag-Leffler theorem.* Recall that given a meromorphic function f and one of its poles z_0, its principal part is a rational function P, which is a polynomial in $1/(z - z_0)$ such that $f - P$ has a removable singularity at z_0 (and hence can be viewed as holomorphic near z_0), see Proposition 3.15. Using Runge's theorem 4.16, we will see that we can prescribe poles and even their principal parts for a meromorphic function.

Theorem 4.22 (Mittag-Leffler)**.** Let $Z = \{z_n\}$ be a countable set of distinct points in $U \subsetneq \mathbb{C}$. Suppose Z does not have accumulation points in U. Given a sequence of rational functions P_n's which are polynomials in $1/(z - z_n)$, there exists a meromorphic function $f \in \mathcal{M}(U)$ whose poles are exactly z_n's with each $f - P_n$ holomorphic near z_n for all n.

The idea is as follows. A naive way to get such a meromorphic function is to just take the sum of all the principal parts, which works when there are only finitely many z_n's. When there are infinitely many poles, there is no reason to believe that the sum of P_n's converges. We will then use Runge's theorem to approximate P_n by rational functions, and hence their differences can form a convergent series.

Proof. For $j \in \mathbb{N}$, consider the compact set

$$K_j := \left\{z \in U : z \in \overline{B}_j \text{ and } d_{\text{Euc}}(z, \partial U) \geq 1/j\right\}. \tag{4.23}$$

When $\partial U = \emptyset$, e.g., when $U = \mathbb{C}$, $d_{\text{Euc}}(z, \partial U)$ is always considered to be ∞. Note that each component of $\hat{\mathbb{C}} \setminus K_j$ intersects $\hat{\mathbb{C}} \setminus U$.

For each $j \in \mathbb{N}$, let I_j be the index set of those n's such that $z_n \in K_{j+1} \setminus K_j$. This implies that I_j is a finite set, and we consider

$$f_j(z) := \sum_{n \in I_j} P_n(z),$$

which is a holomorphic function on a neighborhood of K_j. By Runge's theorem (Corollary 4.20), there exists a rational function g_j with poles in $\hat{\mathbb{C}} \setminus U$ such that

$$\sup_{K_j} |f_j - g_j| \leq \frac{1}{2^j}.$$

Since $K_j \subseteq K_{j+1}$ for all j, by Weierstrass' M-test, the function

$$\sum_{j=m}^{\infty} (f_j - g_j)$$

converges uniformly and hence is holomorphic on K_m. On the other hand,

$$\sum_{j=1}^{n}(f_j - g_j)$$

has poles at z_n for $n \in I_1 \cup \cdots \cup I_j$ with the prescribed principal parts. Thus, the function

$$\sum_{j=1}^{\infty}(f_j - g_j)$$

satisfies all the conditions we list. $\square$

Example 4.24. We mention an example of meromorphic function that has simple poles at positive integers with residues all equal to 1. The corresponding principal part is $\frac{1}{z-n}$ for $z_n = n \in \mathbb{N}$. Clearly, the sum (consisting of f_j in the proof)

$$\sum_{n=1}^{\infty} \frac{1}{z-n}$$

does not converge. To find out suitable g_j's, we note that the zero-order term in the power series expansion of $1/(z-n)$ at 0 is $-1/n$, so we consider

$$\sum_{n=0}^{\infty} \left(\frac{1}{z-n} + \frac{1}{n} \right).$$

Note that for $z \in \overline{B}_R$ and $n > 2R$, we have $|z-n| > n/2$, so

$$\left| \frac{1}{z-n} + \frac{1}{n} \right| = z \left| \frac{1}{(z-n)n} \right| \leq z \left| \frac{1}{n/2 \cdot n} \right|.$$

Hence, the series converges locally uniformly, and hence defines a meromorphic function with the prescribed data.

Remark 4.25. We remark that the idea of subtracting zero-order parts from a power series is a commonly used technique in analysis. For example, it can be used to construct doubly periodic meromorphic functions with double poles. We will not talk about it in detail but just mention that given two linearly independent complex numbers ω_1 and ω_2, the Weierstrass $\wp$-function associated to them is

$$\wp(z) := \frac{1}{z^2} + \sum_{(m,n)\neq(0,0)} \left(\frac{1}{(z - m\omega_1 - n\omega_2)^2} - \frac{1}{(m\omega_1 + n\omega_2)^2} \right).$$

Weierstrass $\wp$-functions transform the study of doubly periodic meromorphic functions into a study of elliptic curves via a geometric relationship, providing elegant geometric proofs for the addition formula of elliptic functions and creating a link between these two fields of mathematics.

4.2.3. *Weierstrass factorization theorem.* In this section, we will see another consequence of Runge's theorem. It is parallel to the Mittag-Leffler theorem, but know we want to prescribes zeros and poles with given orders.

Different from prescribing principal parts, which involves taking infinite sums, prescribing zeros will require taking infinite products. Thus, we start with a brief discussion on this.

Let a_n ($n \in \mathbb{N}$) be a sequence of complex numbers. We say the infinite product

$$\prod_{n=1}^{\infty} a_n$$

converges non-degenerately if the limit

$$\lim_{N \to \infty} \prod_{n=n_0}^{N} a_n$$

exists and is a non-zero complex number for some $n_0 \in \mathbb{N}$. As expected, this non-degenerate convergence is related to the infinite sum of the logarithm series. Note that if $\prod_{n=1}^{\infty} a_n$ converges non-degenerately, the sequence a_n necessarily converges to 1 as $n \to \infty$. Thus, for large n, we have $a_n \in \mathbb{C} \setminus (-\infty, 0]$, on which we define the principal branch of the logarithm, which we will keep using the notation $\log$.

By the continuity of the exponential function, we get the following criterion of the convergence of infinite product.

Lemma 4.26. Let a_n ($n \in \mathbb{N}$) be a complex sequence such that $\{n : a_n = 0\}$ is finite. If there exists $n_0 \in \mathbb{N}$ such that $\sum_{n=n_0}^{\infty} \log a_n$ converges, then $\prod_{n=1}^{\infty} a_n$ converges non-degenerately.

In fact, the two conditions are equivalent. We will mainly use this direction, so we do not discuss the other direction here. When the sum converges uniformly on a set, the continuity of the exponential function implies the product also converges uniformly. This will be used when proving Weierstrass' theorem.

Example 4.27. We mention some examples here.

(1) The infinite product

$$\prod_{n=1}^{\infty} \left(1 + \frac{(-1)^{k+1}}{k}\right) = (1+1)\left(1 - \frac{1}{2}\right)\left(1 + \frac{1}{3}\right) \cdots$$

converges to 1. This follows from

$$\prod_{n=1}^{m} \left(1 + \frac{(-1)^{k+1}}{k}\right) = \begin{cases} 1 & \text{if } m \text{ is even} \\ 1 + \frac{1}{m} & \text{if } m \text{ is odd} \end{cases}.$$

(2) The product $\prod_{n=1}^{\infty} \left(1 + \frac{i}{n}\right)$ does not converge, but $\prod_{n=1}^{\infty} \left|1 + \frac{i}{n}\right|$ does converge. Thus, the convergence of the two products are not related as in the sum case.

Now, we state Weierstrass' theorem. A proof of the theorem is parallel to that of Mittag-Leffler's theorem (Theorem 4.22).

Theorem 4.28 (Weierstrass)**.** Let $Z = \{z_n\}$ be a countable set of distinct points in $U \subseteq \mathbb{C}$. Suppose Z does not have accumulation points in U. Given a sequence of non-zero integers N_n's, there exists a meromorphic function $f \in \mathcal{M}(U)$ whose zeros and poles are exactly z_n's with orders N_n's.

In the statement, we use the convention that if z_n is a pole, then its order is a **negative** integer. This is different from our definition in Section 3.2.2, but we use this convention here since the theorem can be stated in this way so that one can determine whether z_n is a zero or a pole by the parity of N_n.

We are not going to talk about the proof using the Runge theorem. A drawback of using Runge's theorem is that it does not produce an explicit function directly. Instead, we look at the following special case, for which we can write down a holomorphic function explicitly.

Theorem 4.29 (Weierstrass factorization theorem)**.** Let $Z = \{z_n\}$ be a countable set in $\mathbb{C}$, and we allow multiplicities in Z. Suppose Z is either finite or diverging to infinity, in the sense that

$$\lim_{n\to\infty} |z_n| = \infty. \tag{4.30}$$

Then there is an entire function f such that $Z(f) = Z$.

Different from Theorem 4.28, we only talk about zeros, and we do not use a separate index to record their multiplicities but just list them out.

Before giving a proof, we first observe that it is reasonable to consider the product

$$\prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right).$$

This doesn't make sense when $z_n = 0$ for some n, which can be fixed by simply taking all the zero points out and get

$$z^m \cdot \prod_{n=m+1}^{\infty} \left(1 - \frac{z}{z_n}\right) \tag{4.31}$$

if we assume $z_n = 0$ if and only if $n \leq m$. Conceptually, at least pointwise at a point z, what we need is the convergence of the sum $\sum\limits_{n=m+1}^{\infty} \left|\frac{z}{z_n}\right|$, which, since it should be true for all z, is basically equivalent to the convergence of

$$\sum_{n=m+1}^{\infty} \frac{1}{|z_n|}.$$

This is definitely not true in general, so we need to adjust more. Let's complete it in the proof.

Proof of Theorem 4.29. To make each term in (4.31) decay fast enough, we consider

$$z^m \cdot \prod_{n=m+1}^{\infty} \left(1 - \frac{z}{z_n}\right) e^{P_n(z)} \tag{4.32}$$

where each P_n is a polynomial and we hope that they will make

$$\sum_{n=m+1}^{\infty} \log \left(1 - \frac{z}{z_n}\right) e^{P_n(z)} = \sum_{n=m+1}^{\infty} \left(\log \left(1 - \frac{z}{z_n}\right) + P_n(z)\right)$$

converge so that we can apply Lemma 4.26. To this end, first, we observe that for a fixed radius R, given any $z \in B_R$, we have $|z/z_n| < 1/2$ for n large enough by the assumption $z_n \to \infty$. Thus, from the expansion (4.7), we have

$$\log \left(1 - \frac{z}{z_n}\right) = -\frac{z}{z_n} - \frac{1}{2}\left(\frac{z}{z_n}\right)^2 - \frac{1}{3}\left(\frac{z}{z_n}\right)^3 - \cdots .$$

Therefore, we consider

$$P_n(z) = \sum_{k=1}^{m_n} \frac{1}{k}\left(\frac{z}{z_n}\right)^k$$

for some m_n to be chosen. For $z \in B_R$ and n large as above, we can estimate

$$\begin{aligned} \left|\log \left(1 - \frac{z}{z_n}\right) + P_n(z)\right| = \left|\sum_{k=m_n+1}^{\infty} \frac{1}{k}\left(\frac{z}{z_n}\right)^k\right| &\le \frac{1}{m_n+1}\left(\frac{|z|}{|z_n|}\right)^{m_n+1} \cdot \left(1 - \frac{|z|}{|z_n|}\right)^{-1} \\ &\le \frac{1}{m_n+1} \cdot \left(\frac{1}{2}\right)^{m_n+1} \cdot 2. \end{aligned} \tag{4.33}$$

Thus, if we choose, for example, $m_n = n$ for each P_n, then the series

$$\sum_{n=m+1}^{\infty} \left(\log \left(1 - \frac{z}{z_n}\right) + P_n(z)\right)$$

converges locally uniformly. This tells us that (4.32) defines an entire function with those prescribed zeros. □

To make notations simpler, we let

$$E_N(z) := (1-z)e^{\sum_{k=1}^{N} \frac{1}{k} z^k}.$$

Thus, from (4.32), we obtain an entire function of the form

$$z^m \cdot \prod_{n=m+1}^{\infty} E_n\left(\frac{z}{z_n}\right).$$

In fact, since given an entire function $g(z)$, its exponential $e^g(z)$ is always an non-vanishing entire function, we obtain a lot of functions of the form

$$e^{g(z)} \cdot z^m \cdot \prod_{n=m+1}^{\infty} E_n\left(\frac{z}{z_n}\right) \tag{4.34}$$

satisfying Theorem 4.29. It is good that it has exactly the zero set that we want. A problem is that we do not have a controlled growth of this function. In the expression (4.34), the growth is essentially controlled by the growth of g and the n in E_n. For the part of E_n, from the estimate in (4.33), one can first see that if the growth of the zeros is controlled, then we can first get a uniform E_d for all n. We formulate it as a corollary of the Weierstrass factorization theorem here.[ex]

Corollary 4.35. Let $Z = \{z_n\}$ be a countable set in $\mathbb{C}$. Suppose Z satisfies that $z_n = 0$ if and only if $n \leq m$ and

$$\sum_{i=m+1}^{\infty} \frac{1}{|z_n|^{\alpha+1}} < \infty. \tag{4.36}$$

If we let $d := \lfloor \alpha \rfloor$, then for any polynomial g of degree at most d, the function

$$f(z) := e^{g(z)} \cdot z^m \cdot \prod_{n=m+1}^{\infty} E_d\left(\frac{z}{z_n}\right)$$

satisfies that $Z(f) = Z$ and is of order at most α. As before, we allow multiplicities in both the sets Z and $Z(f)$.

The converse of this corollary is a characterization for entire function of finite order. It is the Hadamard factorization theorem.

Theorem 4.37 (Hadamard factorization theorem)**.** Let $\alpha \geq 0$ and f an entire function of order at most α. Suppose $Z = \{z_n\}$ is the zero set of f with $z_n = 0$ if and only if $n \leq m$. If we let $d := \lfloor \alpha \rfloor$, then there exists a polynomial g of degree at most d such that

$$f(z) := e^{g(z)} \cdot z^m \cdot \prod_{n=m+1}^{\infty} E_d\left(\frac{z}{z_n}\right).$$

4.3. **Gamma function and Zeta function.**

4.3.1. *Gamma function.* We look at one example for the Hadamard factorization theorem (Theorem 4.37). Take $f(z) = \sin \pi z$. We know that

$$Z(f) = \mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\}$$

and all the zeros have order one. Moreover, we have

$$|\sin \pi z| = \frac{1}{2}\left|e^{i\pi z} - e^{-i\pi z}\right| \leq e^{\pi |z|},$$

and hence the order of f is at most 1. (Recall how this is defined in (4.15).) Thus, Theorem 4.37 implies that there is a polynomial $g(z)$ of degree at most one such that

$$\sin \pi z = e^{g(z)} \cdot z \cdot \prod_{n \neq 0} E_1\left(\frac{z}{n}\right).$$

Since we can switch the order of the terms in the product, we can first calculate

$$E_1\left(\frac{z}{n}\right) \cdot E_1\left(-\frac{z}{n}\right) = e^{\frac{z}{n}}\left(1 - \frac{z}{n}\right) \cdot e^{-\frac{z}{n}}\left(1 + \frac{z}{n}\right) = 1 - \frac{z^2}{n^2}.$$

That is,

$$\sin \pi z = e^{g(z)} \cdot z \cdot \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right). \tag{4.38}$$

To find out g, we look at

$$\log \frac{\sin \pi z}{z} = g(z) + \sum_{n=1}^{\infty} \log\left(1 - \frac{z^2}{n^2}\right),$$

where the logarithm on the left hand side is a suitable branch. We can look at the Taylor series of both of the functions near 0. Recall again that we have (4.7). Thus, we know

$$\sum_{n=1}^{\infty} \log\left(1 - \frac{z^2}{n^2}\right) = \sum_{n=1}^{\infty}\left(-\frac{z^2}{n^2} + \cdots\right) = -\left(\sum_{n=1}^{\infty} \frac{1}{n^2}\right) z^2 + \cdots, \tag{4.39}$$

where we do not need an explicit expression of the second-order term.[10] We keep track of the terms up to second order since we know that the degree of g is at most one. Similarly, we have

$$\log \frac{\sin \pi z}{\pi z} = \log\left(1 - \frac{\pi^2 z^2}{3!} + \cdots\right) = k \cdot 2\pi i - \frac{\pi^2 z^2}{6} + \cdots, \tag{4.40}$$

where the multiple of $2\pi i$ comes from the possibility of different branches. Thus, comparing (4.39) and (4.40) and noting that $\deg g \leq 1$, we know that $g(z) = \log \pi + k \cdot 2\pi i$ for some k, and hence we know that (4.38) can be now explicitly written as

$$\sin \pi z = \pi z \cdot \prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2}\right), \tag{4.41}$$

the famous Euler sine product formula. Note that comparing the coefficients of z^2 leads to another proof of $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$.

Another consequence of (4.41) is that any entire function f of order at most one and with $Z(f) = \mathbb{Z}$ is of the form

$$f(z) = e^{az+b} \cdot \sin \pi z$$

for some $a, b \in \mathbb{C}$. One then can ask what happens if we would like to find f with $Z(f) = \mathbb{N}$.

To follow the tradition, we assume $Z(f)$ to be the set of non-positive integers as simple zeros. That is,

$$Z(f) = \mathbb{Z}_{\leq 0} = \{0, -1, -2, -3, \cdots\}.$$

If we require, as above, that the order of f is at most one, Theorem 4.37 then implies

$$f(z) = e^{az+b} \cdot z \cdot \prod_{n=1}^{\infty} E_1\left(-\frac{z}{n}\right) = e^{az+b} \cdot z \cdot \prod_{n=1}^{\infty} e^{-z/n}\left(1 + \frac{z}{n}\right)$$

for some $a, b \in \mathbb{C}$. We first choose $b = 0$ for simplicity, and we want to find a so that the resulting f is an "interesting function."

[10] We could have used the formula $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$ to get the coefficient.

Since the zero set is almost translation-invariant (at least in one direction), we first look at how $f(z)$ and $f(z+1)$ are related. Recall that we've chosen $b=0$, and then note that

$$\begin{aligned} f(z+1) &= e^{az+a}(z+1)\cdot\prod_{n=1}^{\infty} e^{-\frac{z+1}{n}}\left(1+\frac{z+1}{n}\right) \\ &= e^{az+a}(z+1)\cdot\prod_{n=2}^{\infty} e^{-\frac{z+1}{n-1}}\left(1+\frac{z+1}{n-1}\right) \\ &= e^{az+a}(1+z)\cdot\prod_{n=2}^{\infty} e^{-\frac{1}{n-1}}\left(1+\frac{1}{n-1}\right)e^{-\frac{z}{n-1}}\left(1+\frac{z}{n}\right) \end{aligned}$$

after a change of variables $n\mapsto n-1$. We compare this with

$$f(z)=e^{az}z\cdot\prod_{n=1}^{\infty} e^{-\frac{z}{n}}\left(1+\frac{z}{n}\right)=e^{az}z\cdot e^{-z}(1+z)\cdot\prod_{n=2}^{\infty} e^{-\frac{z}{n}}\left(1+\frac{z}{n}\right),$$

where we single out the first term in the product. Notice that most of the terms can be compared directly, and for the exponential terms, we have

$$\prod_{n=2}^{\infty} e^{-\frac{z}{n-1}}\Bigg/\prod_{n=2}^{\infty} e^{-\frac{z}{n}}=e^{-z\sum_{n=2}^{\infty}\left(\frac{1}{n-1}-\frac{1}{n}\right)}=e^{-z}.$$

Thus, we can write

$$\begin{aligned} z\cdot f(z+1) &= f(z)\cdot e^{a}\cdot\prod_{n=2}^{\infty} e^{-\frac{1}{n-1}}\left(1+\frac{1}{n-1}\right) \\ &= f(z)\cdot e^{a}\cdot\prod_{n=1}^{\infty} e^{-\frac{1}{n}}\left(1+\frac{1}{n}\right). \end{aligned}$$

Thus, we choose $a=\gamma\in\mathbb{R}$ such that

$$e^{\gamma}\cdot\prod_{n=1}^{\infty} e^{-\frac{1}{n}}\left(1+\frac{1}{n}\right)=1,$$

and such a number γ is called the Euler—Mascheroni constant. A familiar form is

$$\gamma=\sum_{n=1}^{\infty}\left(\frac{1}{n}-\log\left(1+\frac{1}{n}\right)\right)=\lim_{N\to\infty}\left(1+\frac{1}{2}+\cdots+\frac{1}{N}-\log(N+1)\right)=0.577\cdots.$$

With this choice of $a=\gamma$, we know that f satisfies $zf(z+1)=f(z)$. This f then has (simple) zero set exactly consisting of non-positive integers. The **Gamma function** is then defined by its reciprocal; that is,

$$\Gamma(z):=f(z)^{-1}=e^{-\gamma z}\cdot z^{-1}\cdot\prod_{n=1}^{\infty} e^{\frac{z}{n}}\left(1+\frac{z}{n}\right)^{-1}.$$

We summarize some properties of Γ.

(1) Γ has simple poles at $z\in\mathbb{Z}_{\le 0}$ and is holomorphic and non-vanishing elsewhere.

(2) The relation of f implies

$$\Gamma(z+1) = z\Gamma(z) \tag{4.42}$$

for $z \notin \mathbb{Z}_{\leq 0}$.

(3) The recurrence relation (4.42) implies

$$\Gamma(1) = \lim_{z\to 0} z\Gamma(z) = \lim_{z\to 0} e^{-\gamma z} \cdot \prod_{n=1}^{\infty} e^{\frac{z}{n}} \left(1+\frac{z}{n}\right)^{-1} = e^0 \cdot \prod_{n=1}^{\infty} e^0 \cdot 1 = 1.$$

Hence, applying the relation (4.42) again leads to

$$\Gamma(n) = (n-1)!$$

for $n \in \mathbb{N}$. That is, $\Gamma(z)$ can be regarded as an extension of the factorial function to the complex plane.

4.3.2. *Zeta function and prime number theorem.* We now talk about the function that we've mentioned in the introduction. That is, the **Riemann zeta function**

$$\zeta(s) := \sum_{n=1}^{\infty} \frac{1}{n^s},$$

which converges locally uniformly in the half plane[11]

$$H_{>1} := \{s \in \mathbb{C} : \mathrm{Re}(s) > 1\}.$$

In particular, ζ is a holomorphic function on $H_{>1}$. This function is one of the most important functions in math. We know its value when $s = 2$ is

$$\zeta(2) = \frac{\pi^2}{6},$$

which has already appeared before.

In this section, we will relate the function ζ to a theorem about the distribution of prime numbers. The theorem says that the counting function

$$\pi(x) := \text{the number of primes} \leq x$$

grows like $x/\log x$. That is, the "density" of primes grows like $1/\log x$.

Theorem 4.43 (Prime number theorem)**.** As $x \to \infty$, $\frac{\pi(x)}{x/\log x} \to 1$.

We will mention a heuristic why such a result holds. See the discussion below Corollary 4.46. First, we can notice a relation between ζ and the set P of prime numbers, encoded in the following Euler's product identity.

Lemma 4.44 (Euler product identity)**.** For $s \in H_{>1}$, we have

$$\zeta(s) = \prod_{p \in P} \frac{1}{1-p^{-s}}. \tag{4.45}$$

In particular, we know that $\zeta\colon H_{>1} \to \mathbb{C}$ is a non-vanishing holomorphic function.

[11]We use the principal branch of the logarithm to define $n^s := e^{s\log n}$ for $s \in H_{>1}$.

A priori, we have to fix an order in P before knowing the convergence of the product. We can simply write it in the increasing order; that is, $P = \{p_j : j \in \mathbb{N}\}$ with $p_j < p_{j+1}$.

Proof of Lemma 4.44. The proof is basically based on the fundamental theorem of arithmetic. Recall that it says that every positive integer greater than one can be written as a product of prime numbers, and the expression is unique up to the order of the factors. That is, for any $n \in \mathbb{N}$, we can write

$$n = \prod_{p\in P} p^{a_{n,p}}$$

where $a_{n,p} = 0$ for all but finitely many $p \in P$. Thus, for $x \geq 1$ and $N \in \mathbb{N}$, we can write

$$\prod_{p\leq x}\sum_{a=0}^{N}\frac{1}{p^{as}} = \sum_{n}\frac{1}{n^s},$$

where the sum is running over all n with $a_{n,p} \leq N$ for $p \in x$ and $a_{n,p} = 0$ for $p > x$. When s is real with $s > 1$, letting N and x tend to infinity, we see that the right hand side is monotonically converging to $\zeta(s)$ and hence

$$\zeta(s) = \prod_{p\in P}\sum_{a=0}^{\infty}\frac{1}{p^{as}} = \prod_{p\in P}\frac{1}{1-p^{-s}}.$$

For a complex s with $\mathrm{Re}(s) > 1$, one can use

$$|\zeta(s)| \leq \sum_{n=1}^{\infty}\frac{1}{n^{\mathrm{Re}(s)}}$$

and the dominated convergence to conclude the same formula. □

By looking at the formula (4.45) as $s \to 1$, one can get the following consequence which is a well-known fact first proven by Euclid.[ex] This is a sign that the study of ζ can help us understand the distribution of prime numbers, and this is a reason why we intentionally avoid using $|P| = \infty$ in the proof of Lemma 4.44.

Corollary 4.46. $|P| = \infty$. That is, there are infinitely many prime numbers. In particular, we have $\pi(x) \to \infty$ as $x \to \infty$.

Let's now briefly mention why we should believe the asymptotics in Theorem 4.43, as now we know that $\pi(x) \to \infty$. This is not a formal argument, and the symbol $\approx$ means nothing in this paragraph. A very rough idea is that by the idea of removing composite numbers, we have

$$\pi(x) \approx x \cdot \prod_{p\leq x}\left(1 - \frac{1}{p}\right).$$

The reciprocal of this is, similar to above, then

$$\frac{x}{\pi(x)} \approx \prod_{p\leq x}\sum_{a=0}^{\infty}\frac{1}{p^a} = \sum_{n}\frac{1}{n}$$

where the sum is running over n with $a_{n,p} \leq x$. When x is large, in the sum, one can imagine that those terms with $n > x$ will contribute (relatively) very little. If so, then we get

$$\frac{x}{\pi(x)} \approx \sum_{n=1}^{x} \frac{1}{n} \approx \log x.$$

To make this rigorous, one crucial step is to prove that ζ has no zeros on the line $1 + i\mathbb{R} = H_{=1}$. This is also related to the infamous conjecture of Riemann (see Riemann hypothesis).

5. Introduction to Riemann surfaces

In the second part of this semester, we will talk about topics related to different aspects of Riemann surfaces. It will mainly be about harmonic functions, conformal mapping, and the uniformization problem. All of these topics require techniques of taking limits of a sequence of holomorphic or meromorphic functions, and hence we will start from this.

5.1. Normal families in function spaces.

5.1.1. *Normal families of continuous functions.* We will talk about compactness properties of a family of holomorphic functions on an open set U. This question already appears in the study of real-valued functions (for example, the Arzelà–Ascoli theorem), and, as expected, will have many interesting applications. We will see again that because of many nice properties of holomorphic functions, compactness may hold under much weaker conditions.

We first recall some properties for families of continuous functions. They can be discussed in a very general situation, but we just focus on functions on an open set U in $\mathbb{C}$. For simplicity, we will let $\mathcal{C}(U)$ be the set of (real-valued) continuous functions on U.

Definition 5.1. Let $\mathcal{F}$ be a family of continuous functions on U; that is, $\mathcal{F} \subseteq \mathcal{C}(U)$.

(1) We say $\mathcal{F}$ is **equicontinuous** on a set $S \subseteq U$ if given any $\varepsilon > 0$, there exists $\delta > 0$ such that for any $f \in \mathcal{F}$ and $z, w \in S$ with $|z - w| < \delta$, we have $|f(z) - f(w)| < \delta$.

(2) We say $\mathcal{F}$ is **normal** if any sequence $f_n \in \mathcal{F}$ has a subsequence which converges locally uniformly on U.[12]

This may remind us of the sequential compactness properties for metric spaces. Remember it says that in a metric space X, a subset K is compact if and only if it is sequentially compact.$^{\text{ex}}$ Thus, a natural question is whether we can find a "reasonable metric" on the space $\mathcal{C}(U)$. If so, then it may give us good criteria to check property (2).

There is a natural norm that one can put on the space $\mathcal{C}(K)$ when K is a compact subset of U. That is, the **sup norm**, defined by

$$\|f\|_{\mathcal{C}(K)} := \sup_K |f| \tag{5.2}$$

for $f \in \mathcal{C}(K)$. This is a well-defined complete norm when K is compact, and we know that convergence with respect to $\|\cdot\|_{\mathcal{C}(K)}$ means uniform convergence on K.$^{\text{ex}}$ However, we would like to work with $\mathcal{C}(U)$ where U is an open set, so not necessarily (and mostly not) compact. An idea we

[12] Note that the limit does not have to lie in $\mathcal{F}$.

can use is to try to write U as a union of compact sets K_j's. Then try to connect the convergence property we want to a combination of metrics on these K_j's. We divide this idea into three steps.

(a) **Exhaustion.** We suppose there exist compact sets K_j's such that

$$U = \bigcup_{j=1}^{\infty} K_j$$

with $K_j \subseteq \mathrm{int} K_{j+1}$ for all $j \in \mathbb{N}$. This then tells us that knowing how to control a family $\mathcal{F}$ in each K_j is enough for us to control it in an arbitrary compact set $K \subseteq U$. This is because given a compact set $K \subseteq U$, the assumptions above then allow us to find some $j \in \mathbb{N}$ such that $K \subseteq K_j$.$^{\mathrm{ex}}$ Such an exhaustion always exists, and one possible construction have been used before; see (4.23) in the proof of the Mittag-Leffler theorem.

(b) **Normalized norms.** For $f, g \in \mathcal{C}(U)$, we can naturally view them as elements in $\mathcal{C}(K_j)$ for any j by restricting them to $K_j \subseteq U$. Thus, as in (5.2), we can consider

$$d_j(f, g) := \sup_{K_j} |f - g|,$$

well-defined on $\mathcal{C}(U)$ (though not a metric yet). Since we are going to combine all information from K_j's, we first normalize a single d_j by considering

$$\delta_j(f, g) := \frac{d_j(f, g)}{1 + d_j(f, g)}.$$

A few simple consequences of this normalization are listed below.

(1) $\delta_j(f, g) < 1$ and $\delta_j(f, g) \leq d_j(f, g)$.

(2) $d_j(f, g) \leq 1$ if and only if $\frac{1}{2} d_j(f, g) \leq \delta_j(f, g) \leq \frac{1}{2}$.

(c) **Locally uniform convergence property.** For $f, g \in \mathcal{C}(U)$, we now combine all $\delta_j(f, g)$'s by taking

$$\rho_U(f, g) = \rho(f, g) := \sum_{j=1}^{\infty} \frac{1}{2^j} \delta_j(f, g),$$

which is always finite by property (1) above. One can show that ρ defines a metric on $\mathcal{C}(U)$.$^{\mathrm{ex}}$ Our desired convergence property is summarized in the following lemma.

Lemma 5.3. Let U be an open set in $\mathbb{C}$ and f_n's and f be elements in $\mathcal{C}(U)$. Then the following are equivalent.

(1) In the metric space $(\mathcal{C}(U), \rho_U)$, $f_n \to f$ as $n \to \infty$.

(2) On any compact subset of U, f_n converges to f uniformly.

Proof. First, we prove (1)$\Rightarrow$(2). Given a compact set $K \subseteq U$, we can take j large enough such that $K \subseteq K_j$. Since $\rho(f_n, f) \to 0$, in particular, we have $\delta_j(f_n, f) \to 0$. This implies $d_j(f_n, f) \to 0$. Thus, f_n converges to f uniformly on K_j, and hence, on K.

Next, we prove (2)⇒(1). By the assumption, we know that f_n converges to f on every K_j. To show that $\rho(f_n, f) \to 0$, we are going to bound it in two parts as follows. For some $J \in \mathbb{N}$ to be chosen, we write

$$\begin{aligned}\rho(f_n, f) &= \sum_{j=1}^{\infty} \frac{1}{2^j} \delta_j(f_n, f) \\ &\leq \sum_{j=1}^{J} \frac{1}{2^j} \delta_j(f_n, f) + \sum_{j=J+1}^{\infty} \frac{1}{2^j}\end{aligned}$$

using $\delta_j < 1$ in property (1) of δ_j's. To this end, given any $\varepsilon > 0$, we find J so large that $2^{-J} < \varepsilon/2$. After fixing this J, using the uniform convergence on K_J, we can find N such that for any $n \geq N$, $\delta_J(f_n, f) < \varepsilon/(2J)$. In particular, $\delta_j(f_n, f) < \varepsilon/(2J)$ for all $j \leq J$ and $n \geq N$. Putting these into the decomposition above, we have

$$\begin{aligned}\rho(f_n, f) &\leq \sum_{j=1}^{J} \frac{1}{2^j} \delta_j(f_n, f) + \sum_{j=J+1}^{\infty} \frac{1}{2^j} \\ &\leq \sum_{j=1}^{J} \frac{\varepsilon}{2J} + \frac{\varepsilon}{2} \sum_{j'=1}^{\infty} \frac{1}{2^{j'}} \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\end{aligned}$$

for $n \geq N$. This proves $\rho(f_n, f) \to 0$. □

This characterization also tells us that the metric ρ_U is a complete metric. This follows directly from Lemma 5.3 and the the completeness of $\left(\mathcal{C}(K), \|\cdot\|_{\mathcal{C}(K)}\right)$ for a compact set K.

Corollary 5.4. Given an open set U in $\mathbb{C}$, the metric space $(\mathcal{C}(U), \rho_U)$ is complete.

Now we can rephrase the definition in (2). That is, a family $\mathcal{F} \subseteq \mathcal{C}(U)$ is normal if and only if $\overline{\mathcal{F}}$ is compact in $\mathcal{C}(U)$. What we are looking for is "good description" that guarantees normality. We mention two equivalent conditions here. The first one is through the notion of total boundedness.

Definition 5.5. A subset S in a metric space (X, d) is called **totally bounded** if given any $\varepsilon > 0$, S can be covered by finitely many open balls of radius ε. That is, there exist $x_1, \cdots, x_N \in X$ for some $N \in \mathbb{N}$ such that

$$S \subseteq \bigcup_{n=1}^{N} B_\varepsilon(x_n).$$

Recall that a metric space is compact if and only if it is complete and totally bounded. As in Lemma 5.3, we may expect that we can translate this equivalence to the local setting, that is, a description on a compact set.

Proposition 5.6. Given an open set U in $\mathbb{C}$ and a family $\mathcal{F} \subseteq \mathcal{C}(U)$, the following are equivalent.

(1) $\mathcal{F}$ is normal.

(2) $\mathcal{F}$ is totally bounded in $(\mathcal{C}(U), \rho_U)$.

(3) For any compact $K \subseteq U$ and $\varepsilon > 0$, there exist finitely many $f_1, \cdots, f_N \in \mathcal{F}$ such that for any $f \in \mathcal{F}$, there exists f_ℓ with

$$\sup_K |f - f_\ell| < \varepsilon.$$

Note that (3) may appear to be a straightforward reformulation of (2). As in Lemma 5.3, however, the connection is more subtle due to the openness of U. We leave the proof, which uses the compactness of $\overline{\mathcal{F}}$ ((1)⇔(2)) and the same idea as in Lemma 5.3 ((2)⇔(3)), to Assignment 6.

One consequence of Proposition 5.6 is the famous Arzelà–Ascoli theorem.

Theorem 5.7 (Arzelà–Ascoli)**.** Let U be an open set in $\mathbb{C}$ and $\mathcal{F} \subseteq \mathcal{C}(U)$. Then $\mathcal{F}$ is a normal family if and only if the following conditions are true.

(1) $\mathcal{F}$ is equicontinuous on any compact set $K \subseteq U$.

(2) For any $z \in U$, the set $\{f(z) : f \in \mathcal{F}\}$ is bounded.

Proof. First we show that the normality of $\mathcal{F}$ implies (1) and (2). For (1), given a compact set $K \subseteq U$ and a positive number ε, we can take $f_1, \cdots, f_N$ such that for any $f \in \mathcal{F}$, there exists f_ℓ such that

$$\sup_K |f - f_\ell| < \frac{\varepsilon}{3} \tag{5.8}$$

by Proposition 5.6 (3). Now, since each of f_n's is uniformly continuous on K, we can find $\delta > 0$ such that for any $\ell = 1, \cdots, N$ and $z, w \in K$ with $|z - w| < \delta$, we have

$$|f_\ell(z) - f_\ell(w)| < \frac{\varepsilon}{3}. \tag{5.9}$$

Combining (5.8) and (5.9), we get (1), since for any $f \in \mathcal{F}$, we first find f_ℓ with (5.8), and hence

$$|f(z) - f(w)| \leq |f(z) - f_\ell(z)| + |f_\ell(z) - f_\ell(w)| + |f_\ell(w) - f(w)| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

so (1) follows.

For (2), given $z \in U$, one can apply Proposition 5.6 (3) again to a small compact ball around z and a positive number ε, to get finitely many $f_1, \cdots, f_\ell$. Then for any $f \in \mathcal{F}$, we have

$$|f(z)| \leq \max\{|f_n(z)| + \varepsilon : n = 1, \cdots, \ell\}.$$

This shows the boundedness of the set.

In the rest of the proof, we show that (1) and (2) imply $\mathcal{F}$ is normal. Let Q be the set of rational points in U, and hence Q is a countable dense set. We number these rational points by

$$Q = \{q_k : k \in \mathbb{N}\}.$$

Now, suppose we are given a sequence f_n in $\mathcal{F}$. We look at q_1. By condition (2), we know that $f_n(q_1)$ is a bounded sequence in $\mathbb{C}$, so the Bolzano–Weierstrass theorem implies that it admits a convergent subsequence. This gives a subsequence f_{n1} of f_n. Next, we look at this new subsequence at q_2, that is, the sequence $f_{n1}(q_2)$. This is again a bounded sequence, so we can find a further subsequence f_{n2} of f_{n1} such that $f_{n2}(q_2)$ converges. Iteratively, for each k, we can find a sequence

f_{nk}, a subsequence of f_n such that f_{nk} is a subsequence of $f_{n(k-1)}$ and $f_{nk}(q_k)$ is a convergent sequence.

We take the "diagonal subsequence" f_{nn}, and we claim that this is a subsequence we desire. Let $K \subseteq U$ be a compact set. Given $\varepsilon > 0$, by condition (1), we can find $\delta > 0$ such that for any $f \in \mathcal{F}$ and $z, w \in K$ with $|z - w| < \delta$,

$$|f(z) - f(w)| < \frac{\varepsilon}{3}. \tag{5.10}$$

Since K is compact, we can cover it with finitely many balls $B_\delta(x_1), \cdots, B_\delta(x_N)$ with radius δ. For eacu x_j, we can choose a rational point $q_{k_j} \in B_\delta(x_j)$. We know that the sequences $f_{nn}\left(q_{k_j}\right)$ are convergent, and hence Cauchy, for each j. Thus, we can find $n_0 \in \mathbb{N}$ such that

$$\left|f_{nn}(q_{k_j}) - f_{mm}(q_{k_j})\right| \leq \frac{\varepsilon}{3} \tag{5.11}$$

for any $j = 1, \cdots, N$ and $n, m \geq n_0$. Combining (5.10) and (5.11), we can conclude the proof. In fact, using the uniform number n_0 above (only depending on K and ε), given any $z \in K$, we can find some $j = 1, \cdots, N$ such that $z \in B_\delta(x_j)$, and hence

$$\begin{aligned}|f_{nn}(z) - f_{mm}(z)| &\leq |f_{nn}(z) - f_{nn}(x_j)| + |f_{nn}(x_j) - f_{mm}(x_j)| + |f_{mm}(x_j) - f_{mm}(z)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,\end{aligned}$$

for $n, m \geq n_0$. This shows that f_{nn} is a Cauchy sequence in $\mathcal{C}(K)$ and hence finishes the proof. □

5.1.2. *Normal families of holomorphic functions.* Now, as promised, we will look at families of holomorphic functions. To this extend, we let $\mathcal{H}(U)$ be the set of holomorphic functions on U.

Because holomorphic functions behave much more rigidly than continuous functions, one should expect a (pre)compactness result with a weaker condition than the Arzelà–Ascoli theorem. The condition we will look at is the following.

Definition 5.12. Let $\mathcal{F} \subseteq \mathcal{H}(U)$. We say $\mathcal{F}$ is **locally bounded** if given any compact set $K \subseteq U$, there exists $C_K < \infty$ such that $\sup_K |f| \leq C_K$ for any $f \in \mathcal{F}$.

Note that if a family $\mathcal{F}$ of functions satisfies (1) and (2) in Theorem 5.7, then $\mathcal{F}$ is locally bounded.[ex] This is a much weaker condition then (1) and (2), especially for families of continuous functions. However, we will see that it is sufficient to guarantee normality for families in $\mathcal{H}(U)$.

Theorem 5.13 (Montel)**.** Let U be an open set in $\mathbb{C}$ and $\mathcal{F} \subseteq \mathcal{H}(U)$. Then $\mathcal{F}$ is normal if and only if $\mathcal{F}$ is locally bounded.

We remark that once we have a normal family $\mathcal{F} \subseteq \mathcal{H}(U)$, then given a sequence in $\mathcal{F}$, after taking a subsequence that converges locally uniformly on U, Weierstrass's theorem (Corollary 2.23) implies that the derivatives of this subsequence also converges locally uniformly on U.

Proof. The only if part is based on Theorem 5.7 and the exercise above that (2) in Theorem 5.7 imply local boundedness. In the rest, we assume $\mathcal{F} \subseteq \mathcal{H}(U)$ is locally bounded.

By Theorem 5.7, it suffices to show that $\mathcal{F}$ is equicontinuous on any compact $K \subseteq U$, which is equivalent to show that $\mathcal{F}$ is equicontinuous on any small open ball.$^{\text{ex}}$ At a point $z_0 \in U$, take $\delta > 0$ so small that $\overline{B}_{3\delta}(z_0) \subseteq U$. Since $\mathcal{F}$ is locally bounded, we can find $M < \infty$ such that $\sup_{\overline{B}_{3\delta}(z_0)} |f| \leq M$ for all $f \in \mathcal{F}$. Now, for any $f \in \mathcal{F}$ and $z_1, z_2 \in B_\delta(z_0)$, we can then estimate

$$\begin{aligned} |f(z_1) - f(z_2)| = \left| \int_{\partial B_{2\delta}(z_0)} \left(\frac{f(z)}{z - z_1} - \frac{f(z)}{z - z_2} \right) dz \right| &\leq M \int_{\partial B_{2\delta}(z_0)} \frac{|z_1 - z_2|}{|z - z_1| \cdot |z - z_2|} dz \\ &\leq M \cdot \frac{|z_1 - z_2|}{\delta^2} \cdot 2\pi \cdot 2\delta. \end{aligned}$$

This implies the equicontinuity, and the proof is done. □

5.1.3. *Normal families of meromorphic functions.* Montel's theorem has many applications, as we will see later. The first application we will see is its application after being extended to the case of meromorphic functions.

Recall that given an open set $U \subseteq \mathbb{C}$, we let $\mathcal{M}(U)$ be the set of meromorphic functions. It collects functions that are holomorphic away from a discrete set of points, at which the functions have poles. We now introduce a new notion for looking at them, serving as a precursor to our introduction to Riemann surfaces. We consider a new space

$$\hat{\mathbb{C}} := \mathbb{C} \dot{\cup} \{\infty\}$$

by "adding an infinity point" to the complex plane. We have seen this construction in Section 1.2.2, and now we can give it more structures. This space is called the **extended complex plane** or the **Riemann sphere**, and by viewing it as the one-point compactification of $\mathbb{C}$, we get an induced topological structure on it (cf. Exercise 6 in Assignment 2).

A natural way to look at the space is to use the stereographic projection from the unit two-sphere S^2. That is, we look at the map $P\colon S^2 \to \hat{\mathbb{C}}$ defined by

$$P(x, y, w) := \frac{x}{1 - w} + \frac{y}{1 - w} i$$

where we view S^2 as a unit sphere centered at the origin in $\mathbb{R}^3$. Note that when we look at the north pole (i.e., when $w = 1$), we naturally get $P(0, 0, 1) = \infty \in \hat{\mathbb{C}}$. One can show that this map is a homeomorphism and hence $\hat{\mathbb{C}}$ is homeomorphic to the two-sphere S^2.$^{\text{ex}}$

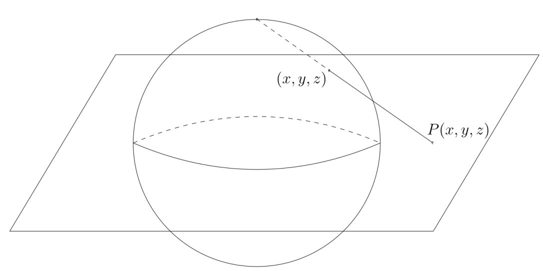

FIGURE 11. A homeomorphism between S^2 and $\hat{\mathbb{C}}$ maps $(0, 0, 1)$ to ∞.

We want to view a meromorphic function $f \in \mathcal{M}(U)$ as a "usual map" into $\hat{\mathbb{C}}$. To this end, we pull make the metric structure on S^2 to $\hat{\mathbb{C}}$. Note that given $z \in \hat{\mathbb{C}}$, one can compute$^{\text{ex}}$

$$P^{-1}(z) = \left(\frac{2z}{1+|z|^2}, \frac{|z|^2-1}{|z|^2+1}\right) \in \mathbb{C} \times \mathbb{R} \simeq \mathbb{R}^3. \tag{5.14}$$

Therefore, if we define $d_{\hat{\mathbb{C}}}(z_1, z_2)$ by $d_{\text{Euc}}\left(P^{-1}(z_1), P^{-1}(z_2)\right)$, we can finally get$^{\text{ex}}$

$$d_{\hat{\mathbb{C}}}(z_1, z_2) = \frac{2|z_1 - z_2|}{\sqrt{1+|z_1|^2} \cdot \sqrt{1+|z_2|^2}}$$

and we have a well-defined metric structure that is compatible with the topology given by the one-point compactification on $\hat{\mathbb{C}}$. The metric $d_{\hat{\mathbb{C}}}$ is called the **spherical metric** on $\hat{\mathbb{C}}$.

Before studying families of meromorphic functions, as in the holomorphic case, we first look at the space $\mathcal{C}(U; \hat{\mathbb{C}})$ of continuous functions from an open set U to $\hat{\mathbb{C}}$. The discussion before Theorem 5.7, as we mentioned, holds for functions valued in a metric space, and therefore, we can again endow a metric $\hat{\rho}_U$ on the space $\mathcal{C}(U; \hat{\mathbb{C}})$ by the same process, and have the same conclusion of Theorem 5.7 and all the results before it. Moreover, notice that condition (2) is now vacuous since $\left(\hat{\mathbb{C}}, d_{\hat{\mathbb{C}}}\right)$ is a bounded metric space. With this, we summarize them in the following corollary.

Corollary 5.15. Let U be an open set in $\mathbb{C}$ and $\mathcal{F} \subseteq \mathcal{C}(U; \hat{\mathbb{C}})$. Then $\mathcal{F}$ is a normal family (in the sense of Definition 5.1(2)) if and only if $\mathcal{F}$ is equicontinuous on any compact set $K \subseteq U$.

The problem now is that the space $\mathcal{C}(U; \hat{\mathbb{C}})$ contains too many functions. Given a function $f \in \mathcal{C}(U; \hat{\mathbb{C}})$, $f^{-1}(\infty)$ may be very large. Thus, to get structural results for families in $\mathcal{M}(U)$, we have to understand possible behavior near infinity along a convergent sequence first. That is done in the following lemma.

Lemma 5.16. Let U be a connected open set in $\mathbb{C}$ and f_n be a sequence in $\mathcal{M}(U)$. If f_n converges to a function f locally uniformly, then f is either meromorphic on U or identically ∞. For the first case, if $f_n \in \mathcal{H}(U)$, then f is holomorphic.

Note that the locally uniform convergence is equivalent to the convergence in the metric $\hat{\rho}_U$. Also, note that $\mathcal{M}(U) \subseteq \mathcal{C}(U; \hat{\mathbb{C}})$ by the blow-up characterization of poles. The lemma tells us that the situation is not too bad, since the only bad case when one has too many singularities is exactly when the limit is identically infinity.

Proof. First, we know that f is a continuous function in the spherical metric $d_{\hat{\mathbb{C}}}$. Take $z_0 \in U$. We deal with two cases.

If $f(z_0) \neq \infty$, then in a neighborhood of z_0, $f_n \neq \infty$ for n large. Hence, Weierstrass's theorem (Corollary 2.23) implies that f is holomorphic in such a neighborhood of z_0.

If $f(z_0) = \infty$, then in a neighborhood of z_0, $f_n \neq 0$ for n large. The sequence of analytic functions $1/f_n$ then converges to $1/f$ uniformly in a slightly smaller neighborhood (when using the spherical metric, and hence when using the usual Euclidean metric), and hence Weierstrass's theorem implies $1/f$ is holomorphic. If z_0 is and isolated zero of $1/f$, then f is meromorphic at z_0. Otherwise, we know $1/f \equiv 0$ and hence $f \equiv \infty$.[13]

For the last conclusion, if $f_n \in \mathcal{H}(U)$ but $f(z_0) = \infty$, then in the neighborhood chosen in the preceding paragraph, f_n cannot be non-vanishing by Hurwitz's theorem (Corollary 3.38). This implies f_n is not holomorphic, a contradiction. □

In view of Lemma 5.16, it may be reasonable to allow the function ∞ when considering normal families of meromorphic functions. Thus, we relax the condition in Definition 5.1(2) and propose the following definition for meromorphic functions.

Definition 5.17. Let $\mathcal{F} \subseteq \mathcal{M}(U) \subseteq \mathcal{C}(U; \hat{\mathbb{C}})$ be a family of meromorphic functions on a connected open set U. We say $\mathcal{F}$ is **normal** if any sequence $f_n \in \mathcal{F}$ has a subsequence which, on any compact subset of U, either converges uniformly to a meromorphic function or converges uniformly to ∞.

Lemma 5.16 then tells us that this definition agrees with Definition 5.1(2) if we replace $(\mathbb{C}, d_{\text{Euc}})$ with $(\hat{\mathbb{C}}, d_{\hat{\mathbb{C}}})$. Moreover, combined with Corollary 5.15, we know that in the new definition of normality for meromorphic families, we can still use equicontinuity to characterize it.

Corollary 5.18. Let U be a connected open set in $\mathbb{C}$ and $\mathcal{F} \subseteq \mathcal{C}(U; \hat{\mathbb{C}})$ be a family of meromorphic functions on U. Then $\mathcal{F}$ is a normal family (in the sense of Definition 5.17) if and only if $\mathcal{F}$ is equicontinuous on any compact set $K \subseteq U$.

In the rest of the section, we will see other ways to characterize normality for meromorphic families. In the following theorem, given a function $f \in U$, we define a new function

$$\hat{f}(z) := \frac{2|f'(z)|}{1 + |f(z)|^2}.$$

We use the notation $\hat{f}$ since, as we will see soon, this is related to arc length on the Riemann sphere $\hat{\mathbb{C}}$.

Theorem 5.19 (Marty)**.** Let $\mathcal{F} \subseteq \mathcal{C}(U; \hat{\mathbb{C}})$ be a family of meromorphic functions on a connected open set U. Then $\mathcal{F}$ is normal if and only if the family

$$\hat{\mathcal{F}} := \left\{ \hat{f} : f \in \mathcal{F} \right\}$$

is locally bounded (in the sense of Definition 5.12).

Proof. First, we proof the "if" part. The idea is to try to use the arc length on the Riemann sphere to bound the spherical distance. To this end, we look at a smooth curve $\gamma \colon [a, b] \to \hat{\mathbb{C}}$. Recall that by (5.14), its trajectory before the stereographic is

$$\widetilde{\gamma}(t) := \left(\frac{2\gamma(t)}{1 + |\gamma(t)|^2}, 1 - \frac{2}{1 + |\gamma(t)|^2} \right).$$

[13]Note that based on Hurwitz's theorem (Corollary 3.38), the second case necessarily happens when f_n's are holomorphic.

To calculate its arc length, we need to know the magnitude of $\widetilde{\gamma}'(t)$, which can be given by a straightforward calculation$^{\mathrm{ex}}$

$$|\widetilde{\gamma}'(t)| = \frac{2|\gamma'(t)|}{1+|\gamma(t)|^2}.$$

Now, to show the equicontinuity in a ball $\overline{B} \subseteq U$, given $z, w \in \overline{B}$, we take the straight line $\ell\colon [0, |z-w|] \to U$ connecting z and w in $\overline{B}$.[14] Then for any $f \in \mathcal{F}$, the curve we will consider in $\hat{\mathbb{C}}$ is $f(\ell(t))$. Since the chord is not longer than the arc, we have

$$d_{\hat{\mathbb{C}}}(f(z), f(w)) \leq \mathrm{Length}(f \circ \ell) = \int_0^{|z-w|} \frac{2|f'(\ell(t))|}{1+|f(\ell(t))|^2} dt$$

using $|\ell'(t)| = 1$. By the assumption, we can find a bound $M < \infty$ such that for all $f \in \mathcal{F}$,

$$\sup_{\overline{B}} |\hat{f}| \leq M.$$

Thus, we can further estimate

$$d_{\hat{\mathbb{C}}}(f(z), f(w)) \leq M \cdot |z-w|.$$

This proves the equicontinuity.

For the "only if" part, we again fix a ball $\overline{B} \subseteq U$. For any small ε, say $\varepsilon < 10^{-10}$, the equicontinuity gives us $\delta > 0$ such that

$$d_{\hat{\mathbb{C}}}(f(z), f(w)) < \varepsilon$$

for any $f \in \mathcal{F}$ and $z, w \in \overline{B}$ with $|z-w| \leq 2\delta$. Let z_0 be the center of the ball $\overline{B}$ and assume $\overline{B}_{2\delta}(z_0) \subseteq \overline{B} \subseteq U$ after shrinking δ if needed. Then we look at $\hat{f}$ in $B_\delta(z_0)$.

Suppose $|f(z_0)| \leq 1$. Then we have $|f(z)| \leq 1 + \varepsilon < 2$ for $z \in \overline{B}_{2\delta}(z_0)$. For $z \in B_\delta(z_0)$, we can estimate

$$|f'(z)| = \left| \frac{1}{2\pi i} \int_{\partial B_{2\delta}(z_0)} \frac{f(w)}{(w-z)^2} dw \right| \leq 2\delta \cdot \frac{2}{\delta^2} = \frac{4}{\delta}.$$

Thus,

$$\hat{f}(z) = \frac{2|f'(z)|}{1+|f(z)|^2} \leq \frac{8/\delta}{1} = \frac{8}{\delta}.$$

Suppose $|f(z_0)| \geq 1$ on the other hand, including the case $f(z_0) = \infty$. Then $|f(z)| \geq 1 - \varepsilon > 1/2$ for $z \in \overline{B}_{2\delta}(z_0)$. Notice that if we consider $g := 1/f$ in $\overline{B}_{2\delta}(z_0)$, we get

$$\hat{g} = \frac{2|f'/f^2|}{1+1/|f|^2} = \frac{2|f'|}{|f|^2+1} = \hat{f}.$$

Moreover, we have $|g| < 2$ in $\overline{B}_{2\delta}(z_0)$, so the previous case implies $\hat{g}(z) \leq 8/\delta$ for $z \in B_\delta(z_0)$.

Combining the two cases, we get $\hat{f}(z) \leq 8/\delta$ for $f \in \mathcal{F}$ and $z \in B_\delta(p)$. This proves the local boundedness. □

[14] We can take ℓ in such a way that it has unit speed, e.g., $\ell(t) = \frac{t}{|z-w|} w + \left(1 - \frac{t}{|z-w|}\right) z$.

Marty's theorem will later be used to get another result of Montel. To achieve that, we want to understand what happens when we are given a "non-normal" sequence. That is, a sequence f_n without a convergent subsequence. Note that by Marty's theorem, this means that the corresponding sequence $\hat{f}_n$ is not locally bounded, so we can always find a "blow-up sequence" in a compact set. By normalizing the sequence using this blow-up phenomenon, we can produce a rescaled limit, and it captures the "singular structure" of the original sequence. This will be implemented in the following result, and we remark that this technique has been extremely successful in the study of singularities in many areas of geometry and analysis.

Proposition 5.20 (Zalcman's lemma)**.** Let U be a connected open set in $\mathbb{C}$ and let f_n be a sequence of meromorphic functions on U. If the family $\{f_n : n \in \mathbb{N}\}$ is not normal, then there exist a sequence z_n in U converging to $z_0 \in U$ and a sequence of positive numbers $M_n \to \infty$ such that the functions

$$g_n(w) := f_n\left(z_n + \frac{w}{M_n}\right)$$

converge to a non-constant meromorphic function on $\mathbb{C}$ with $\hat{g} \leq 1 = \hat{g}(0)$.

Proof. Since the sequence is not normal, by Marty's theorem (Theorem 5.19), there is a compact set $K \subseteq U$ and a sequence $w_n \in K$ such that

$$\lim_{n\to\infty} \hat{f}_n(w_n) = \infty.$$

After rescaling and translating U, we may assume $\overline{B}_1(0) \subseteq U$ and $w_n \in \overline{B}_1(0)$ with $w_n \to 0$. Instead of using w_n and $\hat{f}_n(w_n)$ to rescale, we take another sequence of points z_n such that[15]

$$\max_{z\in\overline{B}_1}(1-|z|)\hat{f}_n(z) = (1-|z_n|)\hat{f}_n(z_n).$$

Note that $z_n \notin \partial B_1$ when n is large as $\hat{f}_n$ diverges. Since $\hat{f}_n(w_n) \to \infty$ and $w_n \to 0$, we know the maximum sequence above diverges. In particular, we also have

$$\lim_{n\to\infty} \hat{f}_n(z_n) = \infty,$$

and we let $M_n := \hat{f}_n(z_n)$ be the blow-up scale.

Now, we use z_n and M_n to define g_n as in the theorem. Then we have that for each n,

$$\hat{g}_n(0) = \frac{\hat{f}_n(z_n)}{M_n} = 1, \tag{5.21}$$

and for $|w| < (1-|z_n|)M_n$, we can estimate

$$\begin{aligned} \hat{g}_n(w) = \frac{\hat{f}_n\left(z_n + \frac{w}{M_n}\right)}{M_n} &\leq \frac{1}{M_n} \cdot \frac{1-|z_n|}{1-\left|z_n + \frac{w}{M_n}\right|} M_n \\ &\leq \frac{1-|z_n|}{1-|z_n|-\left|\frac{w}{M_n}\right|} = \frac{1}{1-\frac{|w|}{(1-|z_n|)M_n}} \end{aligned} \tag{5.22}$$

[15]Directly using w_n to get $g_n := f_n(w_n + w/M_n)$ does not allow one to control $\hat{g}_n$

where the first inequality uses that $(1-|z_n|)M_n$ is the maximum of $(1-|z|)\hat{f}_n(z)$ in $\overline{B}_1$. Hence, for a given $R<\infty$, by taking N so large that $\frac{1}{2}(1-|z_N|)M_N > R$, we have $\hat{g}_n$ is uniformly bounded on B_R for $n \geq N$. Hence, Marty's theorem implies the conclusion. The bounds on $\hat{g}$ follow from taking the limits of (5.21) and (5.22). □

We can now state Montel's theorem for meromorphic families.

Theorem 5.23 (Montel)**.** Let $\mathcal{F}$ be a family of meromorphic functions on a connected open set $U \subseteq \mathbb{C}$. If $\mathcal{F}$ omits three values in $\hat{\mathbb{C}}$, then $\mathcal{F}$ is normal.

Proof. By applying a rational linear transformation (see Assignment 2), we may assume $\mathcal{F}$ omits the three points $0, 1$, and ∞ in $\hat{\mathbb{C}}$. That is, $\mathcal{F}$ is a family of non-vanishing holomorphic functions that omit the value 1.

We prove the theorem by contradiction. Suppose $\mathcal{F}$ is not normal. Then by Marty's theorem, we can find a small ball $\overline{B} \subseteq U$, a sequence $f_n \in \mathcal{F}$, and points $w_n \in \overline{B}$ with $w_n \to 0 \in \overline{B}$ such that

$$\lim_{n\to\infty} \hat{f}_n(w_n) = \infty.$$

We now consider some new families of functions constructed from f_n. Recall that by Corollary 4.9, we can take the roots of these functions on $\overline{B}_1$. Thus, for each $k \in \mathbb{N}\cup\{0\}$, we consider the sequence

$$\mathcal{F}_k := \left\{ f_n^{1/2^k} : n \in \mathbb{N} \right\}.$$

None of these families is normal by our choice of f_n. Hence, by Zalcman's lemma (Proposition 5.20), for each $k \in \mathbb{N}$, we can get a rescaled limit g_k, a non-constant meromorphic function on $\mathbb{C}$ with

$$\hat{g}_k \leq 1 = \hat{g}_k(0)$$

Marty's theorem (Theorem 5.19) then implies that $\{g_k : k \in \mathbb{N}\}$ is a normal family, so a subsequence of it locally uniformly converges to a limit g with $\hat{g}(0) = 1$. In particular, g is a non-constant meromorphic function on $\mathbb{C}$.

Now we look at what values these functions omit. Since each element in $\mathcal{F}_k$ omits all 2^k-th roots of unity, so does their rescaled limit g_k by Hurwitz's theorem (Corollary 3.38). Similar, we know that g omits all 2^k-th roots of unity for all k. By the open mapping theorem (Corollary 3.6 and Assignment 6), we know that $\hat{\mathbb{C}} \setminus g(\mathbb{C})$ is closed. Since the set of all 2^k-th roots of unity for all k is dense in the unit circle, we can conclude that $g(\mathbb{C})$ is disjoint from the unit circle.

If $|g| < 1$, then the Liouville theorem (Corollary 2.25) implies that h is a constant. This contradicts $\hat{g}(0) = 1$.

If $|g| > 1$, by looking at $h := 1/g$, we know that $\hat{h}(0) = \hat{g}(0) = 1$, and applying Liouville to h leads to a contradiction again. □

As we've mentioned, Montel's theorems (Theorem 5.13 and Theorem 5.23) have many applications. Now, we can first use Theorem 5.23 to improve the little Picard theorem we prove in Theorem 4.10.

Theorem 5.24 (Great Picard)**.** If $f\colon B_1 \setminus \{0\} \to \mathbb{C}$ is a non-constant holomorphic function with an essential singularity at 0, then f omits at most one point in $\mathbb{C}$.

Another way to phrase the theorem is that if a meromorphic function omits three values on $\hat{B}_1(0)$, then it extends meromorphically to 0. We know that the little Picard theorem (Theorem 4.10) is related to Proposition 3.17. Theorem 5.24 is a stronger result than both of them.

Proof. Assume f omits at least two values in $\mathbb{C}$. We will show that f must be a constant function on $\hat{B}_1 = B_1 \setminus \{0\}$.

Applying a rational linear transformation, we may assume f omits $0, 1$, and ∞ (in $\hat{\mathbb{C}}$). Thus, f is a holomorphic function. We consider a family by rescaling f; that is, consider $f_\varepsilon(z) := f(\varepsilon z)$ and

$$\mathcal{F} := \{f_\varepsilon(z) : \varepsilon \in (0,1)\}.$$

Each f_ε is defined on a punctured ball larger than $\hat{B}_1$, but we only look at its restriction on $\hat{B}_1$. By Montel's theorem (Theorem 5.23), $\mathcal{F}$ is normal, so there exists a sequence $\varepsilon_n \to 0$ such that f_{ε_n} locally uniformly converges to a limit g. By Lemma 5.16, g is either holomorphic or identically infinity on $\hat{B}_1$.

If g is holomorphic on $\hat{B}_1$, then we can take

$$M := \sup_{\partial B_{1/2}} |g| < \infty.$$

This implies that for large n,

$$\sup_{\partial B_{1/2}} |f_{\varepsilon_n}| \leq 2M,$$

which means that for $z \in \partial B_{1/2}$,

$$|f(\varepsilon_n z)| \leq 2M.$$

Therefore, we can bound

$$\sup_{\partial B_{1/2} \cup \partial B_{\varepsilon_n/2}} |f| \leq 2M.$$

Since $\partial B_{1/2} \cup \partial B_{\varepsilon_n/2} = \partial\left(B_{1/2} \setminus B_{\varepsilon_n/2}\right)$, the maximum principle (Theorem 2.27) implies

$$\sup_{B_{1/2} \setminus B_{\varepsilon_n/2}} |f| \leq 2M.$$

Since this true for all n large and $\varepsilon_n \to 0$, we get

$$\sup_{\hat{B}_{1/2}} |f| \leq 2M.$$

By the Riemann extension theorem (Proposition 3.14), 0 is a removable singularity of f, a contradiction.

If $g \equiv \infty$ on $\hat{B}_1$, we can run the arguments above identically for $1/f$, $1/f_{\varepsilon_n}$, and $1/g$ to show that 0 is a pole of f. This contradicts the assumption again, and the proof is complete. $\square$

5.2. **Introduction to complex dynamics.** We will apply Montel's theorem to study some simple situations in the study of complex dynamics. Complex dynamics studies the behavior of iteration of holomorphic or meromorphic maps. The question is whether these iterations have a "good pattern." To be precise, given a map $f: U \to U$, we will look at

$$f^n := \overbrace{f \circ f \circ \cdots \circ f}^{n \text{ times}}$$

and let $n \to \infty$ to see the behavior. Let's see an example.

Example 5.25. We consider $f: \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ by letting $f(z) := z^2$. For $z_0 \in \hat{\mathbb{C}} \setminus \partial B_1$, the behavior of f^n is well-controlled. In fact, it is clear that if $z_0 \in B_1$, then $f^n(z) \to 0$ for z near z_0; if $z_0 \in \hat{\mathbb{C}}$, then $f^n(z) \to \infty$ for z near z_0. However, if $|z_0| = 1$, then the behavior of f^n near z_0 is not that simple.

We will try to see the different situations mentioned in Example 5.25 for a large class of meromorphic functions. The original ideas of Fatou and Julia are to decompose $\hat{\mathbb{C}}$ to two invariant sets on which f^n has different behavior as $n \to \infty$, as in Example 5.25. Given a function $f: \hat{\mathbb{C}} \to \hat{\mathbb{C}}$, we let the **Fatou set** of f be

$$\mathcal{F}_f := \left\{ z_0 \in \hat{\mathbb{C}} : z_0 \text{ has a neighborhood near which } f^n\text{'s form a normal family} \right\},$$

and we let the **Julia set** of f be its complement, that is,

$$\mathcal{J}_f := \hat{\mathbb{C}} \setminus \mathcal{F}_f.$$

The Julia set is then the set on which the behavior of f^n may become chaotic. For example, when $f(z) = z^2$, we see that $\mathcal{J}_f$ is the unit circle. The first basic property is that both $\mathcal{F}_f$ and $\mathcal{J}_f$ are invariant.

Proposition 5.26. Given a meromorphic function $f \in \mathcal{M}(\mathbb{C})$, $f(\mathcal{F}_f) \subseteq \mathcal{F}_f$ and $f(\mathcal{J}_f) \subseteq \mathcal{J}_f$.

Proof. By their definitions $\mathcal{J}_f = \hat{\mathbb{C}} \backslash \mathcal{F}_f$, it suffices to prove that $f(\mathcal{F}_f) \subseteq \mathcal{F}_f$ and that $f^{-1}(\mathcal{F}_f) \subseteq \mathcal{F}_f$. That is, $z_0 \in \mathcal{F}_f$ if and only if $f(z_0) \in \mathcal{F}_f$. We assume f is not a constant map (otherwise the conclusion is obvious).

If $f(z_0) \in \mathcal{F}_f$, there is a neighborhood V of $f(z_0)$ so that $\{f^n\}$ is normal on V. Then $f^{-1}(V)$ is a neighborhood around z_0 and $\{f^n\}$ is normal on $f^{-1}(V)$, so $z_0 \in \mathcal{F}_f$.

If $z_0 \in \mathcal{F}_f$, there is a neighborhood U of z_0 so that $\{f^n\}$ is normal on U. Then the open mapping theorem (Corollary 3.6) implies $f(U)$ is a neighborhood of $f(z_0)$ and $\{f^n\}$ is normal on $f(U)$. ☐

In the rest of this section, we will focus on the case when f is a polynomial or rational function on $\hat{\mathbb{C}}$. We will see that even for polynomial functions, the corresponding dynamics can already be complicated.

We notice that in Example 5.25, the Julia set of z^2 is the boundary of the set of those points converging to ∞. Let's look at it for the simplest case and assume f is a monic polynomial of degree $f \geq 2$. For such an f, we can find $R < \infty$ large enough that$^{\text{ex}}$

$$|f(z)| > 2|z| \text{ for } |z| \geq R.$$

We let $U_R := \hat{\mathbb{C}} \setminus \overline{B}_R$. Then clearly $U_R \subseteq \mathcal{F}_f$. Moreover, we define the **basin of attraction of** ∞ to be

$$\mathcal{A}_f(\infty) := \left\{ z \in \hat{\mathbb{C}} : \lim_{n\to\infty} f^n(z) = \infty \right\}.$$

We know that $z \in \mathcal{A}_f(\infty)$ if and only if $f^n(z) \in U_R$ for some n_0 (and then for all $n \geq n_0$). Thus,

$$\mathcal{A}_\infty(f) = \bigcup_{n=1}^{\infty} f^{-n}(U_R),$$

where we write $f^{-n}(U_R) := (f^n)^{-1}(U_R)$. In particular, $\mathcal{A}_\infty(f)$ is an open set contained in $\mathcal{F}_f$.

Theorem 5.27. Suppose f is a monic polynomial of degree $d \geq 2$. Then $\mathcal{A}_f(\infty)$ is an open connected subset of $\hat{\mathbb{C}}$ with $\infty \in \mathcal{A}_\infty(f)$. Moreover, $\mathcal{J}_f = \partial \mathcal{A}_\infty(f)$, which is a non-empty compact subset of $\mathbb{C}$.

This theorem provides a way to visualize $\mathcal{J}_f$. We look at another example.

Example 5.28. We consider $f_2 \colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ by letting $f(z) := z^2 - 2$. A trick is to apply a change of variables given by

$$\varphi(z) := z + \frac{1}{z},$$

which is a bijective holomorphic map from $\hat{\mathbb{C}} \setminus \overline{B}_1$ to $\hat{\mathbb{C}} \setminus [-2, 2]$.[ex] It gives

$$f_2 \circ \varphi(z) = \left(z + \frac{1}{z}\right)^2 - 2 = z^2 + \frac{1}{z^2} = \varphi(z^2).$$

This implies that the iteration of f_2 and $f_0(z) := z^2$ are basically the same, by the commutative diagram

$$\begin{tikzcd}
\hat{\mathbb{C}} \setminus \overline{B}_1 \arrow[r, "f_0"] \arrow[d, "\varphi"] & \hat{\mathbb{C}} \setminus \overline{B}_1 \arrow[d, "\varphi"] \\
\hat{\mathbb{C}} \setminus [-2,2] \arrow[r, "f_2"] & \hat{\mathbb{C}} \setminus [-2,2]
\end{tikzcd}.$$

In other words, f_1 is conjugate to $f_2 = \varphi \circ f_0 \circ \varphi^{-1}$. In particular, based on Example 5.25, we know

$$\hat{\mathbb{C}} \setminus [-2, 2] \subseteq \mathcal{A}_{f_2}(\infty) \subseteq \mathcal{F}_{f_2}.$$

On the other hand, for $x \in [-2, 2]$, $f_2(x) = x^2 - 2 \in [-2, 2]$, so it won't escape to ∞. Therefore, we have

$$\mathcal{A}_{f_2}(\infty) = \hat{\mathbb{C}} \setminus [-2, 2],$$

so by Theorem 5.27, $\mathcal{J}_{f_2} = \partial \mathcal{A}_{f_2}(\infty) = [-2.2]$.

We remark that Examples 5.25 and 5.28 are exceptions, in the sense that for the map $f_c(z) := z^2 + c$, its Julia set $\mathcal{J}_{f_c}$ is contained in a smooth curve if and only if $c = 0$ or -2. For example, $\mathcal{J}_{f_{-0.6}}$ is a Jordan curve that is nowhere differentiable.

Proof of Theorem 5.27. It is clear that $f\left(\mathcal{A}_f(\infty)\right) \subseteq \mathcal{A}_f(\infty)$ and $f^{-1}\left(\mathcal{A}_f(\infty)\right) \subseteq \mathcal{A}_f(\infty)$. Hence, $f\left(\partial\mathcal{A}_f(\infty)\right)$ does not intersect $\mathcal{A}_f(\infty)$ and $f\left(\partial\mathcal{A}_f(\infty)\right) \subseteq \partial\mathcal{A}_f(\infty)$. Thus, for any $n \in \mathbb{N}$,

$$f^n\left(\partial\mathcal{A}_f(\infty)\right) \subseteq \partial\mathcal{A}_f(\infty). \tag{5.29}$$

Noting that $\infty \in \mathcal{A}_f(\infty)$, we know that $\partial\mathcal{A}_f(\infty)$ is a bounded set and hence that f^n is uniformly bounded on $\partial\mathcal{A}_f(\infty)$. Thus, if U is a connected component of $\hat{\mathbb{C}} \setminus \partial\mathcal{A}_f(\infty)$ that does not contain ∞, then the maximum principle implies

$$\max_U |f^n| \le \max_{\partial U} |f^n| \le \max_{\partial\mathcal{A}_f(\infty)} |f^n| \le \max_{z\in\partial\mathcal{A}_f(\infty)} |z| < \infty \tag{5.30}$$

for all n. Therefore, U cannot intersect $\mathcal{A}_f(\infty)$ and hence such a component must be in $\hat{\mathbb{C}} \setminus \overline{\mathcal{A}_f(\infty)}$. Consequently, $\mathcal{A}_f(\infty)$ has only one component (the one containing ∞) and hence it is connected.

Nest, on each bounded open set U above, (5.30) also implies that f^n's form a normal family by Montel's theorem (Theorem 5.13). Thus, $U \subseteq \mathcal{F}_f$ for all such U, and hence $\mathcal{J}_f \subseteq \partial\mathcal{A}_f(\infty)$. On the other hand, given $z_0 \in \partial\mathcal{A}_f(\infty)$, by (5.29), $\{f^n(z_0)\}$ is a bounded set. However, for any open neighborhood V near z_0, $V \cap \mathcal{A}_f(\infty) \neq \emptyset$ so $f^n(w) \to \infty$ for any $w \in V \cap \mathcal{A}_f(\infty)$. Thus, $z_0 \notin \mathcal{F}_f$, and hence $z_0 \in \mathcal{J}_f$. This proves $\mathcal{J}_f = \partial\mathcal{A}_f(\infty)$. □

Theorem 5.27, as we mentioned, provide a way to find Julia sets. The next result tells us more refined information about the structure of Julia sets.

Theorem 5.31. Suppose f is a monic polynomial of degree $d \ge 2$. Given a point $z_0 \in \mathcal{J}_f$ and an open neighborhood U of z_0, there exists $N \in \mathbb{N}$ such that

$$\mathcal{J}_f \subseteq \bigcup_{n=0}^{N} f^n\left(U\right). \tag{5.32}$$

In particular, the union of the inverse iterates

$$\bigcup_{k=1}^{\infty} f^{-k}\left(z_0\right) \tag{5.33}$$

is dense in $\mathcal{J}_f$.

For the second statement of Theorem 5.31, it is about the "fractal" property of Julia sets. Roughly speaking, it means that shapes in a Julia set reappear infinitely many times at different scales near any point in the set. Both Theorems 5.27 and 5.31 provide different ways to produce computer images of Julia sets (cf. [Gam01, Section XII.3]).

Proof. For an open neighborhood U of $z_0 \in \mathcal{J}_f$, we know that f^n's do not form a normal family on U. Since each f^n omits ∞ on U, Montel's theorem (Theorem 5.23) implies that f^n's can omit at most one more common value in $\mathbb{C}$. We split it into two cases.

If f^n's do not omit any other common value, that is,

$$\bigcup_{n=1}^{\infty} f^n(U) = \mathbb{C},$$

then the union contains $\mathcal{J}_f$, and (5.32) follows from the compactness of $\mathcal{J}_f$.

If f^n's omit $w_0 \in \mathbb{C}$, that is,

$$\bigcup_{n=1}^{\infty} f^n(U) = \mathbb{C} \setminus \{w_0\}\,, \tag{5.34}$$

then $f^{-1}(w_0) = \{w_0\}$. (Otherwise, if there were $w \neq w_0$ such that $f(w) = w_0$, then (5.34) would imply $w = f^n(z)$ for some $z \in U$ and n and hence $w_0 = f^{n+1}(z)$, contradicting (5.34).) This implies that $f(z) - w_0$ is a constant multiple of $(z - w_0)^d$, so

$$f(z) = w_0 + (z - w_0)^d.$$

This then implies that $w_0 \in \mathcal{F}_f$. Thus, we again derive that the union (5.34) contains $\mathcal{J}_f$, and (5.32) follows from the compactness of $\mathcal{J}_f$.

For the second assertion, we note that the first assertion implies that given an open set U of $\mathcal{J}_f$, there exists $N \in \mathbb{N}$ such taht (5.32) is true. Thus, given $z_0 \in \mathcal{J}_f$, we have $z_0 \in f^n(U)$ for some $n \in \{0, \cdots, N\}$. This means that

$$f^{-n}(z_0) \cap U \neq \emptyset.$$

We can find such a number n for any open set U in $\mathcal{J}_f$, which means that the set (5.33) is dense in $\mathcal{J}_f$. □

We have used complex analysis to get some structure theorems for $\mathcal{F}_f$ and $\mathcal{J}_f$. Now, we will see that how they are related to some important objects in a dynamical system. Given a map $f\colon U \to U \subseteq \hat{\mathbb{C}}$, a central object in the study of dynamical systems is the behavior of "orbits." Given $z \in U$, the **orbit** of z is the set

$$\mathcal{O}_f(z) := \{f^n(z) : z \in \mathbb{N} \cup \{0\}\}\,.$$

Clearly, all the orbits form a partition of U. Based on this definition, we mention a few terminologies.

Definition 5.35. Given $f\colon U \to U \subseteq \hat{\mathbb{C}}$, a point $z_0 \in U$ is called **periodic** if $f^n(z_0) = z_0$ for some $n \in \mathbb{N}$. The smallest number n such that this holds is called the **period** of z_0. The multiplier of z_0 is defined to be

$$\lambda := (f^n)'(z) = \prod_{j=0}^{n} f'(z_j)$$

where $z_j := f^j(z_0)$'s are the elements of $\mathcal{O}_f(z_0)$. When z_0 is periodic, $\mathcal{O}_f(z_0)$ is called a periodic orbit. Moreover, given such a periodic orbit $\mathcal{O}_f(z_0)$,

(1) we say $\mathcal{O}_f(z_0)$ is attracting if $|\lambda| \in (0, 1)$;
(2) we say $\mathcal{O}_f(z_0)$ is superattracting if $\lambda = 0$;
(3) we say $\mathcal{O}_f(z_0)$ is repelling if $|\lambda| > 1$;
(4) we say $\mathcal{O}_f(z_0)$ is indifferent (or neutral) if $|\lambda| = 1$.

In Case (4), if λ is a root of unity but $f^k \neq \mathrm{id}_U$ for all k, then the points in $\mathcal{O}_f(z_0)$ are called parabolic periodic points. In Case (1), the basin of attraction of $\mathcal{O}_f(z_0)$, denoted by $\mathcal{A}_f(\mathcal{O}_f(z_0))$, collects all z such that $f^{jn}(z) \to z_\infty$ as $j \to \infty$ for some $z_\infty \in \mathcal{O}_f(z_0)$.

We do not have to remember these names, though some of them are pretty self-explanatory. The definition of $\mathcal{A}_f(\mathcal{O}_f(z_0))$ may be viewed as a generalization of $\mathcal{A}_f(\infty)$ we have encountered. It is straightforward to see whether each of them belongs to $\mathcal{F}_f$ or $\mathcal{J}_f$.[ex]

Lemma 5.36. Let $f\colon U \to U \subseteq \hat{\mathbb{C}}$ and suppose $z_0 \in U$ is periodic.

(1) If $\mathcal{O}_f(z_0)$ is attracting or superattracting, then $\mathcal{O}_f(z_0) \subseteq \mathcal{A}_f(\mathcal{O}_f(z_0)) \subseteq \mathcal{F}_f$.
(2) If $\mathcal{O}_f(z_0)$ is repelling, then $\mathcal{O}_f(z_0) \subseteq \mathcal{J}_f$.
(3) If z_0 is a parabolic periodic point, then $\mathcal{O}_f(z_0) \subseteq \mathcal{J}_f$.

A generalization of the concept of orbits also involves preimages under f. Given $f\colon U \to U$ and $z_0 \in U$, the **grand orbit** of z_0 is the set

$$\begin{aligned}\mathcal{G}_f(z_0) :&= \{z \in U : \mathcal{O}_f(z) \cap \mathcal{O}_f(z_0) \neq \emptyset\} \\ &= \{z \in U : f^m(z) = f^n(z_0) \text{ for some } m, n \in \mathbb{N} \cup \{0\}\}.\end{aligned}$$

When $|\mathcal{G}_f(z_0)| < \infty$, we say z_0 is an **exceptional point** of f. The set of exceptional points is denoted by $\mathcal{E}_f$. As its name suggests, a point in $\mathcal{E}_f$ can rarely be approached by f, say, if one starts iterating f from a generic point in U. We can show that this does not happen for many points when f is a rational function.[ex]

Proposition 5.37. If $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a rational function of degree $d \geq 2$, then $\mathcal{E}_f$ contains at most two points. Moreover, each element in $\mathcal{E}_f$ is superattracting and hence lies in $\mathcal{F}_f$.

The importance of the exceptional sets of rational functions can be seen in the following transitivity property.

Theorem 5.38. If $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ is a rational function of degree $d \geq 2$, then given $z_0 \in \mathcal{J}_f$ and an open neighborhood U of z_0 that is disjoint from $\mathcal{E}_f$, we have

$$\bigcup_{n=0}^{\infty} f^n(U) = \hat{\mathbb{C}} \setminus \mathcal{E}_f.$$

One can compare Theorem 5.38 with Theorem 5.31. Theorem 5.38 says that $\mathcal{J}_f$ is "transitive," in the sense that f "expands" any small neighborhood of points in $\mathcal{J}_f$ to the whole Riemann sphere except at most two points (Proposition 5.37).

Proof of Theorem 5.38. Assume f is not a constant map. Then the set $V := \bigcup_{n=0}^{\infty} f^n(U)$ is open. On V, f^n's do not form a normal family, so by Montel's theorem (Theorem 5.23), we know that f^n's omit at most two points. Since $f(V) \subseteq V$ by its definition, this implies $\hat{\mathbb{C}} \setminus V$ contains at most two points.

On the other hand, $f(V) \subseteq V$ means $\hat{\mathbb{C}} \setminus V \subseteq \hat{\mathbb{C}} \setminus f(V)$, so given $z \in \hat{\mathbb{C}} \setminus V$, we have $z \in \hat{\mathbb{C}} \setminus f(V)$ and hence $f^{-1}(z) \in \hat{\mathbb{C}} \setminus V$. Inductively, we get $f^{-n}(z) \in \hat{\mathbb{C}} \setminus V$ for all $n \geq 0$. That is,

$$f\left(\hat{\mathbb{C}} \setminus V\right) \subseteq \hat{\mathbb{C}} \setminus V.$$

Thus, in particular, $\mathcal{G}_f(z) \subseteq \hat{\mathbb{C}} \setminus V$, which is a finite set (having at most two points). This implies $z \in \mathcal{E}_f$.

Conversely, if $z \in \mathcal{E}_f$, then since U does not touch $\mathcal{E}_f$, we know $\mathcal{G}_f(z) \cap U = \emptyset$ so z does not lie in V. Hence, $z \in \hat{\mathbb{C}} \setminus V$ and the proof is done. □

This tells us more properties about Julia sets of rational functions. We again list them here and refer to Milnor's book [Mil06] for reading.[ex]

Corollary 5.39. Let $f \colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a rational function of degree $d \geq 2$.

(1) $\mathcal{J}_f$ has an interior point if and only if $\mathcal{J}_f = \mathbb{C}$.

(2) $\mathcal{J}_f$ does not contain an isolated point.

(3) Given $z_0 \in \hat{\mathbb{C}} \setminus \mathcal{E}_f$, the iterated preimage $\bigcup_{k=1}^{\infty} f^{-k}(z_0)$ is dense in $\mathcal{J}_f$, which we see in Theorem 5.31 for $z_0 \in \mathcal{J}_f$.

(4) For all but finitely many $z_0 \in \mathcal{J}_f$, the set $\{z \in \mathcal{J}_f : (\mathcal{J}_f, z) \text{ is locally biholomorphic to } (\mathcal{J}_f, z_0)\}$ is dense in $\mathcal{J}_f$.

Property (4) can be viewed as "self-similarity." What one can actually show is that if $z_0 \in \mathcal{J}_f$ with $f'(z_0) \neq 0$, then $(\mathcal{J}_f, f(z_0))$ is locally biholomorphic to $(\mathcal{J}_f, z_0)$.

We can only briefly look at some aspects of complex dynamics. This is a fruitful subject which is later related to ergodic theory, hyperbolic geometry, and Teichmüller theory. For interested readers, a good starting point is Milnor's book [Mil06], on which some content in this section was based. (e.g., One can find Lemma 5.36 and Proposition 5.37 in Milnor's book.)

5.3. **Conformal maps and harmonic functions.** We will now turn to the topics about conformal maps and harmonic functions.

5.3.1. *Conformal maps and harmonic functions.* A map $f \colon U \to \mathbb{C}$ is called **conformal** if it "preserves the angles." To be precise, let $f \colon U \to \mathbb{C}$ be a C^1 map on $U \subseteq \mathbb{C}$. As before, we can write

$$f(z) = f(x, y) = u(x, y) + iv(x, y)$$

by viewing $z = x + iy \in U$. Then the tangents in the x and y directions will be sent to

$$\begin{aligned}(1,0) &\mapsto \left(\frac{\partial u}{\partial x}, \frac{\partial v}{\partial x}\right), \text{ and}\\ (0,1) &\mapsto \left(\frac{\partial u}{\partial y}, \frac{\partial v}{\partial y}\right).\end{aligned}$$

Thus, if we let

$$J := \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix}$$

be the Jacobian matrix of f, then the conformality condition means

$$\frac{\langle v, w\rangle}{|v|\cdot|w|} = \frac{\langle Jv, Jw\rangle}{|Jv|\cdot|Jw|} \tag{5.40}$$

given any $v, w \in \mathbb{R}^2 \setminus \{0\}$. By this, we can see that holomorphic functions are included when their derivatives do not vanish.

Proposition 5.41. If $f\colon U \to \mathbb{C}$ is holomorphic at $z_0 \in U$ with $f'(z_0) \neq 0$, then f is conformal at z_0.

Proof. By the Cauchy–Riemann equation (1.7), the Jacobian matrix of f at z_0 is

$$J = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \end{pmatrix} = \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ -\frac{\partial u}{\partial y} & \frac{\partial u}{\partial x} \end{pmatrix},$$

which is a rotation matrix when $f'(z_0) \neq 0$. Thus, (5.40) holds. $\square$

One can similarly show that the converse statement is also true. That is, a conformal map is in fact a holomorphic function with a non-vanishing derivative.$^{\text{ex}}$

When Riemann originally studied conformal maps, he was trying to find conformal maps using harmonic functions. For example, if one is able to solve the system

$$\begin{cases} \Delta u = 0 & \text{in } U \\ u = f & \text{on } \partial U \end{cases} \tag{5.42}$$

for a domain U with a suitably given f on ∂U, then it is possible to relate it back to finding conformal maps into a disc. We will go back to this in Section 5.4.

Besides the usage related to conformal map, the boundary value problem (5.42) is also important in many other geometric questions, and even physics. The problem of solving the boundary value problem (5.42) is called the **Dirichlet problem**. Our goal in the following two sections is to see whether we can solve it for a general class of domains and boundary data. On the other hand, we will also see that (5.42) is not always solvable, depending on the given U and f. See Remark 5.60 and Assignment 8.

Before studying harmonic functions in a systematic way, we can already use the most fundamental property of them to solve the Dirichlet problem for a disc. Recall that harmonic functions satisfy the mean-value equality based on Corollary 2.30. One can also prove this fact by starting from the definition ($\Delta u = 0$) directly.$^{\text{ex}}$ In fact, we can get a slightly stronger version than Corollary 2.30 by approximating the boundary values. We list it as a lemma here.

Lemma 5.43. Let $u\colon \overline{B}_1 \to \mathbb{R}$ be a continuous function that is harmonic in B_1. Then

$$u(0) = \frac{1}{2\pi}\int_0^{2\pi} u\left(e^{it}\right) dt.$$

The Dirichlet problem for B_1 is then related to generalizing the mean-value equality above to an arbitrary point in B_1. To achieve, we recall again that in Assignment 1, we study some good properties of the function

$$h_{z_0}(z) := \frac{z_0 - z}{1 - \overline{z_0} z},$$

which, especially sends 0 to z_0. Thus, given a continuous $u \colon \overline{B}_1 \to \mathbb{R}$ that is harmonic inside B_1, for any $z_0 \in B_1$, using h_{z_0}, we can consider

$$\widetilde{u}(w) := u\left(h_{z_0}(w)\right) = u\left(\frac{z_0 - w}{1 - \overline{z}_0 w}\right),$$

which is still harmonic in B_1.$^{\text{ex}}$ We can then calculate

$$u(z_0) = \widetilde{u}(0) = \frac{1}{2\pi}\int_0^{2\pi} \widetilde{u}\left(e^{it}\right) dt = \frac{1}{2\pi}\int_0^{2\pi} u\left(\frac{z_0 - e^{it}}{1 - \overline{z}_0 e^{it}}\right) dt. \tag{5.44}$$

Note that $\frac{z_0 - e^{it}}{1 - \overline{z}_0 e^{it}}$ is still on the unit circle by the property of h_{z_0}, and by writting $\frac{z_0 - e^{it}}{1 - \overline{z}_0 e^{it}} =: e^{is}$, we can get

$$e^{it} = h_{z_0}\left(e^{is}\right) = \frac{z_0 - e^{is}}{1 - \overline{z}_0 e^{is}}.$$

Thus, differentiating leads to

$$e^{it} dt = \left(\frac{-e^{is}}{1 - \overline{z}_0 e^{is}} + \frac{\left(z_0 - e^{is}\right) \cdot \overline{z}_0 e^{is}}{\left(1 - \overline{z}_0 e^{is}\right)^2}\right) ds.$$

Dividing both sides by $e^{it} = \frac{z_0 - e^{is}}{1 - \overline{z}_0 e^{is}}$, we get

$$dt = \left(\frac{e^{is}}{e^{is} - z_0} + \frac{\overline{z}_0 e^{is}}{1 - \overline{z}_0 e^{is}}\right) ds = \frac{1 - |z_0|^2}{|e^{is} - z_0|^2} ds.$$

Putting this back to (5.44), we derive

$$u(z_0) = \frac{1}{2\pi}\int_0^{2\pi} u\left(e^{is}\right) \frac{1 - |z_0|^2}{|e^{is} - z_0|^2} ds. \tag{5.45}$$

That is, the value of u in B_1 is completely determined by its value on ∂B_1. The equality (5.45) is called the **Poisson formula**.

On the other hand, formula (5.45) actually can be used to produce harmonic functions with a prescribed continuous boundary datum. Thus, it provides a solution to the Dirichlet problem for B_1.

Proposition 5.46. Let $v \colon \partial B_1 \to \mathbb{R}$ be a continuous function on the unit circle. Then if we define

$$u(z) := \begin{cases} \frac{1}{2\pi}\int_0^{2\pi} v\left(e^{is}\right) \frac{1 - |z|^2}{|e^{is} - z|^2} ds & \text{if } z \in B_1 \\ v(z) & \text{if } z \in \partial B_1 \end{cases},$$

then u solves (5.42) for $U = B_1$ and $f = v$.

Proof. It is relatively direct to see the harmonicity in the interior. In fact, we can write[ex]

$$\frac{1-|z|^2}{|e^{is}-z|^2} = \mathrm{Re}\left(\frac{e^{is}+z}{e^{is}-z}\right),$$

so, on B_1, u is the real part of the holomorphic function

$$\frac{1}{2\pi i}\int_{\partial B_1} v(w)\cdot\frac{w+z}{w-z}\cdot\frac{1}{w}dw. \tag{5.47}$$

This part in fact does not require the continuity of v.

For the boundary continuity, we do need the continuity of v. For simplicity, we write

$$P_{v_0}(z) := \frac{1}{2\pi}\int_0^{2\pi} v_0\left(e^{is}\right)\frac{1-|z|^2}{|e^{is}-z|^2}ds$$

be the Poisson integral constructed from v_0, which makes sense for any integrable function v_0 and $z \in B_1$. Thus, viewing P as a map from the space of integrable functions on ∂B_1 to the space of harmonic functions on B_1, we can derive some basic properties of P, including the linearity and the preservation of non-negativity. Based on this, to prove the boundary continuity, that is,

$$\lim_{z\to z_0} P_v(z) = v(z_0)$$

for any $z_0 \in \partial B_1$, we may assume $v(z_0) = 0$ by possibly adding a constant to v.

Now we start to estimate and fix an arbitrarily small number $\varepsilon > 0$. By the continuity of v, we can take $\delta > 0$ such that $|v| < \varepsilon/2$ on $\partial B_1 \cap B_\delta(z_0)$. For the part in the arc $C_1 := \partial B_1 \cap B_\delta(z_0)$, we will make use of the smallness of v, and on the rest $C_2 := \partial B_1 \setminus B_\delta(z_0)$, we will estimate directly. We consider

$$v_i := v \cdot \mathrm{id}_{C_i}$$

where id_{C_i} is the characteristic function of C_i. The pointwise bound $|v_1| < \varepsilon/2$ then implies

$$|P_{v_1}| < \frac{\varepsilon}{2} \text{ on } B_1. \tag{5.48}$$

On the other hand, since v_2 vanishes on C_1, if we write $z_0 = e^{is_0}$ and

$$C_1 = \left\{e^{is} : s \in (s_0 - d, s_0 + d)\right\}$$

for some $d > 0$, we have, when z is close to z_0,

$$\begin{aligned}
|P_{v_2}(z)| = \left|\frac{1}{2\pi}\int_0^{2\pi} v_2\left(e^{is}\right)\frac{1-|z|^2}{|e^{is}-z|^2}ds\right| &= \left|\frac{1}{2\pi}\int_{s_0+d}^{s_0-d+2\pi} v_2\left(e^{is}\right)\frac{1-|z|^2}{|e^{is}-z|^2}ds\right| \\
&\le \frac{2\pi-2d}{2\pi}\cdot\sup_{\partial B_1}|v|\cdot\frac{\left|1-|z|^2\right|}{2C_d^2}
\end{aligned}$$

where $C_d^2 = 2 - 2\cos(2d) = |e^{id} - 1|$ is the minimum of $|e^{is} - z_0|$ on C_2. Hence, this implies that

$$|P_{v_2}(z)| < \frac{\varepsilon}{2} \tag{5.49}$$

when $z \to z_0$. Combining (5.48) and (5.49), we get the continuity of u at z_0. □

The question we want to answer by the end of next section is to what extend we can solve (5.42). We will use Perron's method of approximation by subharmonic functions to achieve in some general situations. We will first study properties of harmonic functions in this section, and then introduce subharmonic functions and Perron's method in Section 5.3.3.

First, we can use the mean-value property (the equality in Lemma 5.43) to derive the maximum principle. Note that this is also what we've done for holomorphic functions. What we can do more here is, we can instead start from the mean-value equality. For a continuous function $u\colon U \to \mathbb{R}$ on an open set $U \subseteq \mathbb{C}$, we say that u satisfies the mean-value property if for any $\overline{B}_r(z_0) \subseteq U$, it satisfies

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u\left(z_0 + re^{it}\right) dt.$$

We know that harmonic functions satisfy the mean-value property by Lemma 5.43. As in Theorem 2.27, it implies the strong maximum principle, and similarly the strong minimum principle.

Lemma 5.50. Suppose U is a connected open set in $\mathbb{C}$ and $u\colon U \to \mathbb{R}$ satisfies the mean-value property. If $\sup_U u$ or $\inf_U u$ is achieved at a point in U, then u is a constant function.

The proof is the same as that of Theorem 2.27 and we do not repeat it here. As a result, we can see that the mean-value property in fact characterizes harmonicity. Note that the mean-value property does not assume C^2 regularity.

Corollary 5.51. Suppose U is an open set in $\mathbb{C}$ and $u\colon U \to \mathbb{R}$ satisfies the mean-value property. Then u is harmonic.

Proof. We fix a ball $\overline{B}_r(z_0) \subseteq U$. Since u is continuous, we can use it to produce a harmonic function by the Dirichlet problem for B_1. That is, Proposition 5.46 implies that there is a continuous function $\widetilde{u}\colon \overline{B}_r(z_0) \to \mathbb{R}$ which is harmonic in $B_r(z_0)$ and equals v on $\partial B_r(z_0)$. Thus, $u - v$ is identically zero on $\partial B_r(z_0)$ and the maximum principle (Lemma 5.50) implies $u - v$ is identically zero in $B_r(z_0)$. This implies that $u = v$, and hence u is harmonic near z_0, and this argument applies to any z_0. $\square$

Thus, functions satisfying the mean-value property are automatically harmonic. Note that the regularity then suddenly becomes much stronger (from C^1 to analytic).

Finally, we prove an important estimate for harmonic functions. It is called the Harnack inequality. It allows us to compare harmonic functions at different points.

Theorem 5.52 (Harnack inequality)**.** Let $u\colon U \to \mathbb{R}$ be a positive harmonic function. Given a compact set $K \subseteq U$, there exists $C_K < \infty$ such that

$$\sup_K u \le C_K \cdot \inf_K u.$$

The crucial point is that the constant C_K does not depend on the function itself. Thus, the estimate can be applied to a large class of harmonic functions, and thus can be useful when we want to get a uniform bound from a family or a sequence of harmonic functions.

Proof. We first deal with the case when K is a ball. Recall that we have the Poisson formula (5.45), so for any $z \in B_1$, we have

$$u(z) = \frac{1}{2\pi} \int_0^{2\pi} u\left(e^{it}\right) \frac{1 - |z|^2}{|e^{it} - z|^2} dt.$$

To estimate the integrand, for $r \in (0,1)$ and $z \in B_r$, the triangle inequality implies

$$1 - r \leq |e^{it} - z| \leq 1 + r.$$

Thus,

$$\frac{1-r}{1+r} \leq \frac{1-r^2}{(1+r)^2} \leq \frac{1-|z|^2}{|e^{it}-z|^2} \leq \frac{1-r^2}{(1-r)^2} \leq \frac{1+r}{1-r}.$$

This and the usual mean-value property then imply

$$\frac{1-r}{1+r} u(0) \leq u(z) \leq \frac{1+r}{1-r} u(0).$$

This is true for all $z \in B_r$, and hence we get

$$\sup_{B_r} u \leq \frac{1+r}{1-r} u(0) = \left(\frac{1+r}{1-r}\right)^2 \cdot \frac{1-r}{1+r} u(0) \leq \left(\frac{1+r}{1-r}\right)^2 \inf_{B_r} u \tag{5.53}$$

for any $r \in (0,1)$.

We sketch the idea for the case of general compact sets. When we are given a compact $K \subseteq U$, we can assume K is connected by looking at each of its connected components. Then we can find z_M and z_m such that

$$u(z_M) = \sup_K u \text{ and } u(z_m) = \inf_K u.$$

Take a curve γ joining z_M and z_m, and cover the curve by finitely many balls. This curve can be bounded by at most N balls with radius $r < d_{\mathrm{Euc}}(K, \partial U)$, where N only depends on K and U. The inequality (5.53) can be applied to each of the balls, and combining them leads to the desired estimate on K. $\square$

Harnack's estimate, as expected, has many applications. We end the section by mentioning a direct application about monotone sequences of harmonic functions.

Corollary 5.54. Let U be a connected open set in $\mathbb{C}$ and suppose u_n is a sequence of harmonic functions on U. Suppose u_n is an increasing sequence, in the sense that for $u_1 \leq u_2 \leq \cdots$. Then u_n either locally uniformly converges to a harmonic function on U, or locally uniformly converges to ∞.

Proof. We may assume $u_n > 0$; otherwise, we can replace u_n with $u_n - u_1 + 1$. Fix $z_0 \in U$ and take $\delta > 0$ so small that $\overline{B}_\delta(z_0) \subseteq U$. We look at the limit $L := \lim_{n\to\infty} u_n(z_0)$, which exists by the monotonicity.

If $L < \infty$, then for n large, we have $u_n(z_0) \leq 2L$. We can apply Harnack's estimate and get a constant $C < \infty$ (independent of n) such that for all n large as above,

$$\sup_{\overline{B}_\delta(z_0)} u_n \leq C \cdot \inf_{\overline{B}_\delta(z_0)} u_n \leq 2CL.$$

This finite bound and the monotonicity then imply u_n converges to a limit u. The monotone convergence theorem implies that u still satisfies the mean-value property, so u is harmonic in $B_\delta(z_0)$.

If $L = \infty$, then given any $M < \infty$, we have $u_n \geq M$ for n large. We can apply Harnack's estimate and get a constant $C' > 0$ (independent of n) such that for all n large as above,

$$\inf_{\overline{B}_\delta(z_0)} u_n \geq C' \sup_{\overline{B}_\delta(z_0)} u_n \geq C'M.$$

This implies $u_n \to \infty$ uniformly on $\overline{B}_\delta(z_0)$.

By the two cases above, we know that the two sets

$$\left\{z \in U : \lim_{n\to\infty} u_n(z) < \infty\right\} \text{ and } \left\{z \in U : \lim_{n\to\infty} u_n(z) = \infty\right\}$$

are both open. The connectivity of U then forces one of them is empty, and hence exactly one of them holds for all $z \in U$. $\square$

We remark that by assuming the functions in Corollary 5.54 are positive, one can drop the monotonicity assumption by using the Harnack estimate directly.

5.3.2. *Subharmonic functions.* The goal of this section is to study the Dirichlet problem (5.42) for a general domain.

The method we will study was first employed by Perron, and it is based on the approximation idea using "subharmonic functions," which are subsolutions to the Laplace equation. In the one-dimensional case, harmonic functions are just straight lines, and then subsolutions to the Laplace equation are exactly convex functions (functions with $u'' \geq 0$). Perron's method in the (simple) one-dimensional case can be illustrated in Figure 12, which motivates the definition of subharmonic functions.

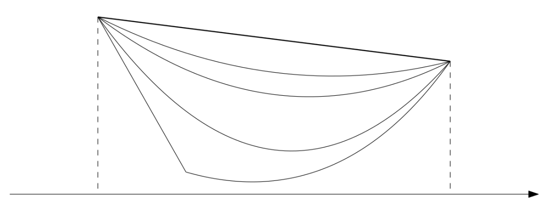

FIGURE 12. When the boundary points are fixed, the straight line, the graph of a harmonic function, maximizes the value at each point among all convex (subharmonic) graphs.

Definition 5.55. Let U be an open set and $v\colon U \to \mathbb{R}$ be a continuous function. We say v is **subharmonic** if for any $\overline{B}_r(z_0) \subseteq U$, it satisfies the mean-value inequality

$$v(z_0) \leq \frac{1}{2\pi} \int_0^{2\pi} v\left(z_0 + re^{it}\right) dt.$$

Clearly, different from the mean-value equality, subharmonic functions are much less rigid, and we do not expect the regularity can be upgraded as in the harmonic case. We will still see in Assignment 8 that subharmonic functions can be related to the Laplace equation, as mentioned above, when they are assumed to be C^2. In general, we can use the maximum principle to characterize subharmonicity.

Theorem 5.56. Let v be a continuous function on a connected open set $U \subseteq \mathbb{C}$. Then v is subharmonic if and only if for any connected subdomain $U' \subseteq U$ and any harmonic function u on U', $v-u$ satisfies the strong maximum principle on U'; that is, if $v-u$ achieves its maximum at an interior point $p \in U'$, then it must be constant on U'.

Proof. The argument for the "only if" direction is the same as the proof of Theorem 2.27, and we go through it again as we did not repeat it for Lemma 5.50. Suppose the maximum M of $v-u$ is achieved in U'. We then consider

$$(v-u)^{-1}(M) = \left\{z \in U' : v(z) - u(z) = M\right\}.$$

Since v satisfies the mean-value inequality and u satisfies the mean-value equality, for any $z_0 \in (v-u)^{-1}(M)$, we can take $r > 0$ such that $\overline{B}_r(z_0) \subseteq U'$ and hence for any $\varepsilon \in (0,r)$,

$$M = v(z_0) - u(z_0) \leq \frac{1}{2\pi} \int_0^{2\pi} (v-u)\left(z_0 + \varepsilon e^{it}\right).$$

The maximality of M then implies $v-u$ is constantly M in the ball. This implies $(v-u)^{-1}(M)$ is open, and hence $(v-u)^{-1}(M) = U'$.

For the "if" part, we will use the Dirichlet problem for balls. Fix a ball $\overline{B}_r(z_0) \subseteq U$. Since v is continuous, by Proposition 5.46, there is a continuous function $u \colon \overline{B}_r(z_0) \to \mathbb{R}$ which is harmonic in $B_r(z_0)$. The assumption implies that the maximum of $v-u$ is achieved on the boundary $\partial B_r(z_0)$. However, we know that $v-u=0$ on the boundary, so we have $v-u \leq 0$ on the whole ball $B_r(z_0)$. Using the mean-value equality for u, we get

$$v(z_0) \leq u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u\left(z_0 + re^{it}\right) dt = \frac{1}{2\pi} \int_0^{2\pi} v\left(z_0 + re^{it}\right) dt,$$

and the proof is done. □

Example 5.57. Although subharmonic functions do not have much rigidity, this means that it may be easier to construct them. We mention two main sources here.

(1) Any harmonic function is obviously a subharmonic function.

(2) If $f(z)$ is holomorphic and non-vanishing, then $|f(z)|$ is subharmonic and $\log|f(z)|$ is harmonic. Thus, we have a large class of examples from holomorphic functions. See Assignment 8.

Next, we list some basic constructions for subharmonic functions.

Proposition 5.58. Let v and v_1 be subharmonic functions on an open set $U \subseteq \mathbb{C}$.

(1) kv ($k \geq 0$) is subharmonic.

(2) $\max\{v, v_1\}$ is subharmonic.

(3) Given $\overline{B}_r(z_0) \subseteq U$, replacing the values in $B_r(z_0)$ with the Poisson integral leads to another subharmonic function. That is, the function

$$v_2(z) := \begin{cases} v(z) & \text{if } z \in U \setminus B_r(z_0) \\ \frac{1}{2\pi}\int_0^{2\pi} v\left(z_0 + re^{it}\right) \frac{r^2 - |z-z_0|^2}{|re^{it} - (z-z_0)|^2} dt & \text{if } z \in B_r(z_0) \end{cases}$$

is subharmonic on U. This is called the **harmonic lifting** of v in $B_r(z_0)$.

Proof. It is not hard to see from the definition of the mean-value inequality that kv is subharmonic for $k \geq 0$.

For $s := \max\{v, v_1\}$, we look at a connected subdomain $U' \subseteq U$ and a harmonic function u on U. Suppose $s - u$ achieves its maximum at $z_0 \in U'$, say $s(z_0) = v(z_0)$. Then for any $z \in U'$,

$$v(z) - u(z) \leq s(z) - u(z) \leq s(z_0) - u(z_0) = v(z_0) - u(z_0).$$

Thus, $v - u$ also achieves its maximum at $z_0 \in U'$, so $v - u$ is constant on U'. This implies that the inequalities above are all equalities, and hence $s - u = v - u$ is also constant.

For v_2, first we know that $v \leq v_2$. We again look at a connected subdomain $U' \subseteq U$ and a harmonic function u on U'. Suppose $v_2 - u$ achieves its maximum at $z_0 \in U'$. We consider the two situations.

If $z_0 \in U \setminus B_r(z_0)$, then $v_2(z_0) = v(z_0)$, and hence for any $z \in U'$,

$$v(z) - u(z) \leq v_2(z) - u(z) \leq v_2(z_0) - u(z_0) = v(z_0) - u(z_0).$$

Thus, $v - u$ also achieves its maximum at $z_0 \in U'$, so $v - u$ is constant on U', and the inequalities above are equalities.

If $z_0 \in B_r(z_0)$, the harmonicity of v_2 implies $v_2 - u$ is constant in $U' \cap \overline{B}_r(z_0)$. We can then run the same argument in the preceding paragraph to a point on $U' \cap \partial B_r(z_0)$, on which $v_2 = v$. $\square$

5.3.3. *Perron's method to solve the Dirichlet problem.* Now, we start to set up for Perron's method of solving the Dirichlet problem. Let U be a bounded and connected open set in $\mathbb{C}$ and f be a function on ∂U. We first do not assume f is continuous, but we assume it is bounded, say $|f| \leq M < \infty$.

As we mentioned, we want to use subharmonic function to approximate the solution to the Dirichlet problem (if exists). Thus, we consider

$$\mathcal{S}(f) := \left\{ v \in \mathcal{C}\left(\overline{U}\right) : v \text{ is subharmonic in } U \text{ with } v|_{\partial U} \leq f \right\}.$$

Note that by the bound $|f| \leq M$, any constant functions less than $-M$ belong to $\mathcal{S}(f)$.

Perron's strategy consists of the following two steps.

(1) Since subharmonic functions are bounded by harmonic functions (with the same boundary data), it is reasonable to approximate the solution from below. Thus, we define a function $\mathcal{P}f$ on U by letting

$$(\mathcal{P}f)(z) := \sup_{v \in \mathcal{S}(f)} v(z) \tag{5.59}$$

for each $z \in U$. Is $\mathcal{P}f : U \to \mathbb{R} \cup \{\infty\}$ harmonic?

(2) Does $\mathcal{P}f$ satisfy the boundary condition? That is, is it true that

$$\lim_{z\to p} (\mathcal{P}f)(z) = f(p)$$

for $p \in \partial U$?

Remark 5.60. . We can't hope to get both of (1) and (2) for all U's. For example, if one lets $U = B_1 \setminus \{0\}$ and sets $f = 1$ on ∂B_1 and $f(0) = 0$, then one cannot solve the Dirichlet problem for such data. We will work on this example in Assignment 8.

As mentioned in Remark 5.60, it is not always possible to have the nice boundary behavior as in (2). However, we can first see that (1) is in general true.

Theorem 5.61. The function $\mathcal{P}f$ constructed in (5.59) is harmonic in U. That is, (1) is correct.

The function $\mathcal{P}f$ is called the **Perron solution** on U with boundary f.[16] Note that when the Dirichlet problem (5.42) is solvable, the solution must equal $\mathcal{P}f$ by the maximum principle.

Proof. First, for each $v \in \mathcal{S}(f)$, the maximum principle (applied to $v-0$) implies $v \leq \sup_{\partial U} f \leq M$. As a result, we first know that $\mathcal{P}f \colon U \to \mathbb{R}$ is a bounded well-defined function. We write $u := \mathcal{P}f$ and will prove that it is harmonic.

Given any $B := B_r(z_0)$ with $\overline{B} \subseteq U$, by the definition of u, we can find a sequence of subharmonic functions $v_n \in \mathcal{S}(f)$ such that

$$\lim_{n\to\infty} v_n(z_0) = u(z_0). \tag{5.62}$$

We will now produce a harmonic function from these v_n's, and show that it agrees with u.

We may assume v_n is an increasing sequence (i.e., $v_1 \leq v_2 \leq \cdots$) by replacing v_n with $\max\{v_1, \cdots, v_n\}$ inductively. The new sequence is still in $\mathcal{S}(f)$ by Proposition 5.58 and still produces the same limit (5.62) at z_0. We then consider $\widetilde{v}_n$, the harmonic lifting of v_n in B. Since $v_n \leq \widetilde{v}_n$ and $\widetilde{v}_n$ is still in $\mathcal{S}(f)$, we have the (monotone) limit (5.62) is still preserved; that is,

$$\lim_{n\to\infty} \widetilde{v}_n(z_0) = u(z_0). \tag{5.63}$$

Moreover, for any $z \in B$, the maximum principle implies

$$\widetilde{v}_n(z) - \widetilde{v}_{n+1}(z) \leq \max_{\partial B} (\widetilde{v}_n - \widetilde{v}_{n+1}) = \max_{\partial B} (v_n - v_{n+1}) \leq 0 \tag{5.64}$$

by the monotonicity of v_n's. Thus, $\widetilde{v}_n|_B$ is an increasing sequence of harmonic functions that are uniformly bounded by M. Harnack's compactness theorem (Corollary 5.54) then tells us that $\widetilde{v}_n|_B$ locally uniformly converges to a harmonic function $\widetilde{v}$ on B with $\widetilde{v}(z_0) = u(z_0)$ by (5.63).

We are going to prove $\widetilde{v} = u$ on the whole ball B. Given another point $z_1 \in B$, we can do a similar process as above, with v_n's added in the construction. This allows us to compare the corresponding limit with $\widetilde{u}$. To be more specific, we can take $w_n \in \mathcal{S}(f)$ such that

$$\lim_{n\to\infty} w_n(z_1) = u(z_1) \tag{5.65}$$

[16] This is a bit misleading since we do not know any actual boundary information about $\mathcal{P}f$.

and by replacing w_n with $\max\{v_j, w_j : j \le n\}$ inductively, we get an increasing sequence w_n with the same limit (5.65) and $w_n \ge v_n$. Then again, consider $\widetilde{w}_n$, the harmonic lifting of w_n in B. The same reason implies

$$\lim_{n\to\infty} \widetilde{w}_n(z_1) = u(z_1) \tag{5.66}$$

and $\widetilde{w}_n$ is increasing. Harnack's compactness theorem then implies that $\widetilde{w}_n|_B$ locally uniformly converges to a harmonic functions $\widetilde{w}$ on B with $\widetilde{w}(z_1) = u(z_1)$ by (5.66).

Now, we compare $\widetilde{w}$ and $\widetilde{v}$. Since $w_n \ge v_n$, the same argument in (5.64) implies $\widetilde{w}_n \ge \widetilde{v}_n$ in B. In particular, $\widetilde{v}_n(z_0) \le \widetilde{w}_n(z_0)$. Moreover, we have $\widetilde{w}_n(z_0) \le u(z_0) = \widetilde{v}(z_0)$ by (5.63). Combining these two, we get $\widetilde{w}(z_0) = \widetilde{v}(z_0)$. On the other hand, $\widetilde{w}_n \ge \widetilde{v}_n$ implies $\widetilde{w} \ge \widetilde{v}$ in B. Thus, the maximum principle implies $\widetilde{w} = \widetilde{v}$. This implies

$$\widetilde{v}(z_1) = \widetilde{w}(z_1) = u(z_1)$$

by (5.66). This is true for any $z_0 \in B$, so we get $u = \widetilde{v}$ on the whole ball B. In particular, u is harmonic near z_0, and the proof is complete. □

Note that we do not use much explicit information of f in the construction of $\mathcal{P}f$ except the bound $|f| \le M$. Thus, Theorem 5.61 is true for very arbitrary f, and we may need more to achieve (2). To understand what really happens on the boundary, we first try to find some necessary conditions for (2) to be true. To this end, we say $p_0 \in \partial U$ is a **regular boundary point** if given any bounded function $f \colon \partial U \to \mathbb{R}$ that is continuous near p_0, we have

$$\lim_{z\to p_0} (\mathcal{P}f)(z) = f(p_0).$$

That is, (2) is true at p_0 when f is continuous. We will see that this boundary regularity is related to some "barrier property" near the boundary point. We start with the following observation.

Suppose $p_0 \in \partial U$ is a regular boundary point. We look at the continuous function $f(p) := |p - p_0|$ for $p \in \partial U$ and get its Perron solution $u := \mathcal{P}f$. The boundary regularity implies

$$\lim_{z\to p_0} u(z) = f(p_0) = 0.$$

Note that by Example 5.57(2), the function $|z - p_0|$ is subharmonic and hence in $\mathcal{S}(f)$. This implies $u(z) \ge |z - p_0|$ for all $z \in U$. Hence, the function u satisfies the following properties.

(1) u is harmonic on U.

(2) $u \ge |(\cdot) - p_0|$ on U.

(3) $\lim_{z\to p_0} u(z) = 0$.

This means that u has a limit at p_0 that is strictly less than its limit sup as z tends to other boundary points. A relaxation of this necessary condition is recorded in the following definition.

Definition 5.67. A **harmonic barrier** at $p_0 \in \partial U$ is a continuous function w on $\overline{U}$ such that

(1) w is harmonic on U,

(2) $w > 0$ on $\partial U \setminus \{p_0\}$, and

(3) $w(p_0) = 0$.

We remark thata we will only need local barrier functions for the purpose of solving the Dirichlet problem. The definition above can also be localized.

Example 5.68. In a few important situations, such a harmonic barrier function exists.

(1) When $U = B_1$, one can check that $w(z) := 1 - \mathrm{Re}\,(\overline{p_0}z)$ is a harmonic barrier for $p_0 \in \partial B_1$.
(2) Suppose that at $p_0 \in \partial U$, there is a line segment $I \subseteq \mathbb{C} \setminus U$ with $p_0 \in \partial I$. We claim that there is a subharmonic barrier at p_0. To see this, after a rotation and a translation, we may assum $p_0 = 0$ and $I = [-1, 0]$ is contained in the real axis. Then one can check that $w(z) := \mathrm{Re}\sqrt{\frac{z}{z+1}}$ is a harmonic barrier.
(3) Based on (2), if ∂U is a union of finitely many paths (see Definition 2.1), then every point on ∂U admits a harmonic barrier.

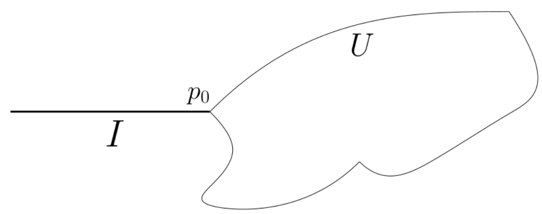

FIGURE 13. When the boundary of an open set is a path, in particular, every boundary point admits a harmonic barrier.

The existence of subharmonic barriers, as we know, is a necessary consequence of boundary regularity. It turns out that it characterizes boundary regularity and hence gives us a general condition to answer (2) and solve the Dirichlet problem.

Theorem 5.69. *Let U be a bounded, connected, open set and let $p_0 \in \partial U$. If there is a harmonic barrier at p_0, then p_0 is a regular boundary point on ∂U.*

The idea is that $\mathcal{P}f$ can be almost bounded by f near its continuous point. This boundedness can be achieved with the help of a barrier function.

Proof. Let f be a bounded function on ∂U that is continuous near p_0 and let $u := \mathcal{P}f$. Let $w\colon U \to \mathbb{R}$ be a harmonic barrier at p_0. Given any $\varepsilon > 0$, we will prove

$$\limsup_{z\to p_0} u(z) \le f(p_0) + \varepsilon \tag{5.70}$$

and

$$\liminf_{z\to p_0} u(z) \ge f(p_0) - \varepsilon. \tag{5.71}$$

Clearly, they imply $u(z) \to f(p_0)$ as $z \to p_0$, which is what we want.

To see (5.70), we first take $\delta > 0$ such that

$$|f(p) - f(p_0)| < \varepsilon \text{ for } p \in \partial U \cap B_\delta(p_0).$$

We take

$$w_0 := \min_{\partial U \setminus B_\delta(p_0)} w,$$

which is positive (by (2)). Then we consider a continuous function

$$w_1(z) := f(p_0) + \frac{M - f(p_0)}{w_0} w(z) + \varepsilon,$$

which is still harmonic in U. Note that for $p \in \partial U \cap B_\delta(p_0)$,

$$w_1(p) \geq f(p_0) + \varepsilon \geq f(p),$$

while for $p \in \partial U \setminus B_\delta(p_0)$,

$$w_1(p) \geq f(p_0) + (M - f(p_0)) + \varepsilon \geq f(p).$$

Hence, given $v \in \mathcal{S}(f)$, the maximum principle (Theorem 5.56) applied to $v - w_1$ implies $v \leq w_1$ also in U. Taking the supremum over $v \in \mathcal{S}(f)$, we get $u(z) \leq w_1(z)$. Thus,

$$\limsup_{z \to p_0} u(z) \leq \limsup_{z \to p_0} w_1(z) = w_1(p_0) = f(p_0) + \varepsilon.$$

This proves (5.70).

For (5.71), the argument is similar by considering

$$w_2(z) := f(p_0) - \frac{M + f(p_0)}{w_0} w(z) - \varepsilon$$

and proving $w_2 \leq f$ on ∂U.[ex] Hence, $w_2 \in \mathcal{S}(f)$, so $u(z) \geq w_2$, implying

$$\liminf_{z \to p_0} u(z) \geq \limsup_{z \to p_0} w_2(z) = w_2(p_0) = f(p_0) - \varepsilon.$$

This proves (5.71) and the theorem. □

Combining Theorem 5.61, Example 5.68(3), and Theorem 5.69, we can derive that the Dirichlet problem is solvable for a continuous function on ∂U, which is a union of finitely many paths. We record this as a corollary.

Corollary 5.72. When U is a bounded connected open set with ∂U being a union of finitely many paths, given a continuous function f on ∂U, there exists a unique solution to the Dirichlet problem (5.42).

Later, we will see how this is related to finding conformal maps between complicated domains. See Theorem 5.75 in Section 5.4.

5.4. **Riemann mapping theorem.** We defined a local notion of conformality in Section 5.3. Now, we look at a global version of it that gives us a way to classify spaces.

Definition 5.73. A **conformal mapping** $f\colon U \to V$ is a bijective C^1 map that is conformal at any $z \in U$. In this case, we say U is **conformally equivalent** to V.

Thus, the set of conformal mappings is a more restrictive class, instead of those maps that are conformal at any point. One can find a map that is conformal at any point but fails to be injective globally. Note that based on the inverse function theorem (Theorem 2.47), the inverse of a conformal mapping is also holomorphic, so a conformal mapping is actually the same as a biholomorphic map.

Example 5.74. We mention examples of conformal mappings that can be written down explicitly.

(1) Any automorphism $f \in \mathrm{Aut}(D)$ is a conformal equivalence between D and itself.

(2) Let $H := \{z \in \mathbb{C} : \mathrm{Im} z > 0\}$ be the upper half plane. Then the map

$$\varphi \colon H \to D$$
$$z \mapsto \frac{z-i}{z+i}$$

is a conformal equivalence between H and D. Its inverse is $w \mapsto \frac{1+w}{1-w} i$.

(3) The "cylinder" $\mathbb{C}/\mathbb{Z}$, which is defined by

$$\{a + bi : a \in \mathbb{R} \text{ and } b \in \mathbb{R}/\mathbb{Z}\},$$

is conformally equivalent to the punctured plane $\mathbb{C}^* := \mathbb{C} \setminus \{0\}$ by the map $e^{2\pi i z}$.

(4) The exponential map e^z defines a conformal equivalence between the (horizontal) strip $\{z \in \mathbb{C} : \mathrm{Im} z \in (0, \pi/2)\}$ and the sector $\{z \in \mathbb{C} : \mathrm{Re} z, \mathrm{Im} z > 0\}$.

We also provide a case when a conformal equivalence does not exist.

(5) The unit disc D is not conformally equivalent to $\mathbb{C}$. Otherwise, there would exist a non-constant holomorphic $f \colon \mathbb{C} \to D$. This cannot happen because of the Liouville theorem (Corollary 2.25).

We will see that these examples are not special, based on the Riemann mapping theorem we will prove in this section. At the beginning of Section 5.3, we mention that this was one of Riemann's original motivations to study the Dirichlet problem. When he did this, he was not able to get a general result due to the lack of some topological facts. Let's now try to implement his ideas at least in a very nice and still fairly general situation.

Theorem 5.75 (Riemann mapping theorem, simple version)**.** Let U be a bounded, connected, open set in $\mathbb{C}$ such that ∂U is a union of finitely many paths. If U is simply connected, then U is conformally equivalent to the (open) unit disc D.

Proof. After a translation, we may assume $0 \in U$. Then the function $f(z) = \log |z|$ is a continuous function on ∂U. Thus, by Corollary 5.72, we can find a harmonic function $u \colon U \to \mathbb{R}$ with $u|_{\partial U} = f$. Next, since U is simply connected, by Corollary 4.5, we can find a harmonic conjugate $v \colon U \to \mathbb{R}$ of u on U. We then consider the function

$$\varphi(z) := z \cdot e^{-(u(z)+iv(z))},$$

a holomorphic function on U. Given $z_0 \in \partial U$, we have

$$\lim_{z \to z_0} |\varphi(z)| = |z_0| \cdot e^{-u(z_0)} = |z_0| \cdot e^{-\log |z_0|} = 1. \tag{5.76}$$

Thus, the maximum principle implies $|\varphi(z)| < 1$ for $z \in U$. That is, φ is a holomorphic map into the unit disc.

We now verify that φ is a conformal mapping. First, we will see φ is bijective from U to D. Notice that $\varphi(z) = 0$ if and only if $z = 0$. Moreover, given $y_0 \in D$, using (5.76), we can take $\varepsilon > 0$ so small that

$$|\varphi| > |y_0| \text{ on } \partial\varphi^{-1}\left(B_{1-\varepsilon}\right).$$

Thus, Rouché's theorem (Corollary 3.40) implies φ and $\varphi - y_0$ have the same numbers of zeros in $\varphi^{-1}\left(B_{1-\varepsilon}\right)$ for any ε small enough, and hence in U. Therefore, there exists a unique $z_0 \in U$ such that $\varphi(z_0) = y_0$.

Next, note that the injectivity implies that φ' cannot vanish. This follows from the local description of holomorphic functions, see Theorem 3.5. In conclusion, φ is a conformal mapping. $\square$

We see that Theorem 5.75 deals with a large class of domains, but the method can only apply for bounded domains with nice enough boundaries. In fact, for a simply connected domain, one can show that the boundary can't be too bad, but this has not been known in Riemann's time.

The first proof of the most general form of the Riemann mapping theorem was done by Koebe, or Osgood, who proved another theorem from which the Riemann mapping theorem follows. We now state the theorem.

Theorem 5.77 (Riemann mapping theorem)**.** Let $U \subsetneq \mathbb{C}$ be a simply connected open set. Then U is conformally equivalent to the (open) unit disc D. Furthermore, given $z_0 \in U$, if we require a conformal mapping $h\colon U \to D$ to satisfy $h(z_0) = 0$ and $h'(z_0) > 0$, then such an h is unique.

We remark that all the assumptions are necessary. If U is conformally equivalent to the unit disc, then the conformal equivalence is, in particular, a homeomorphism, so U must be simply connected. Moreover, we also see that the whole complex plane $\mathbb{C}$ is not conformally equivalent to the unit disc in Example 5.74(5). Thus, Theorem 5.77 actually characterizes those domains that are conformally equivalent to D.

To motivate the proof, we reinterpret the Schwarz lemma we prove for the unit disc in Corollary 2.29. It can be viewed as a consequence of Schwarz's lemma.

Lemma 5.78. Let U be a connected open set in $\mathbb{C}$ and $z_0 \in U$. Let

$$\mathcal{F} := \{f \in \mathcal{H}(U) : |f| < 1 \text{ and } f(z_0) = 0\}.$$

If $\mathcal{F}$ contains a conformal mapping φ_0 from U to the unit disc D, then for any $f \in \mathcal{F}$,

$$|f(z)| \leq |\varphi_0(z)| \text{ for all } z \in U \text{ and } |f'(z_0)| \leq |\varphi_0'(z_0)|.$$

Moreover, if either $|f(z)| = |\varphi_0(z)|$ for some $z \in U$ or $|f'(z_0)| = |\varphi_0'(z_0)|$, then $f = e^{it}\varphi_0$ for some $t \in \mathbb{R}$.

Proof. Given $f \in \mathcal{F}$, consider the function $\widetilde{f} := f \circ \varphi_0^{-1}\colon D \to D$. Then $\widetilde{f}$ is holomorphic and satisfies $\widetilde{f}(0) = f(z_0) = 0$. Thus, Schwarz's lemma implies $|\widetilde{f}(z)| \leq |z|$ and $|\widetilde{f}'(0)| \leq 1$, with characterizations of the equality cases. After plugging in $\widetilde{f} = f \circ \varphi_0^{-1}$, this is exactly what we want. $\square$

Based on Lemma 5.78, we now fix $z_0 \in U$ and look at the class $\mathcal{F}$ defined above. If the Riemann mapping theorem is correct, then a conformal mapping $\varphi_0 \colon U \to D$ will maximize the value $|f'(z_0)|$ among all f's in $\mathcal{F}$. Thus, we will try to find such an element in $\mathcal{F}$. Practically, we will restrict $\mathcal{F}$ to an even smaller class by assuming each f is injective, and that is the class that we will work on.

Proof of Theorem 5.77. Fix $z_0 \in U$. The uniqueness follows from the classification of automorphisms of D. See Corollary 2.29.

For existence, as we discussed above, we consider the family

$$\mathcal{F}_0 := \{f \in \mathcal{H}(U) : |f| < 1, f(z_0) = 0, \text{ and } f \text{ is injective}\}.$$

First, we show that $\mathcal{F}_0$ is a non-trivial family (so that later we can do an approximation argument). Since $U \subsetneq \mathbb{C}$, after a translation, we may assume $0 \notin U$. To get a map into D, we want to consider a function of the form

$$\mathrm{Inv}_{r,z_1}(z) := \frac{r}{2(z - z_1)}$$

for some $r > 0$ and $z_1 \in \mathbb{C}$, since if U is disjoint from the disc $B_r(z_1)$, then Inv_r is an injection into the unit disc D. To achieve this, we have to "squeeze" U first. (Otherwise, U may not avoid any open ball.) To this end, we note that since U is simply connected, by Corollary 4.9, we can find a holomorphic branch $s(z) = \sqrt{z}$ of the square root function on U. s is injective since s^2 is injective. Moreover, if $s(U)$ contains a ball $B_r(z_1)$ then $s(U)$ will not intersect $B_r(-z_1)$; otherwise s^2 would not be injective. We can find such $r > 0$ and $z_1 \in \mathbb{C}$ by the open mapping theorem. Thus, we are done, and

$$\mathrm{Inv}_{r,z_1} \circ s \in \mathcal{F}_0.$$

Next, by using the map $\mathrm{Inv}_{r,z_1} \circ s$, we may assume U is an open subset of D and $z_0 = 0$, and hence we are going to maximize $|f'(0)|$ for $f \in \mathcal{F}_0$. We let

$$M := \sup_{f \in \mathcal{F}_0} |f'(0)|.$$

Since $\mathrm{id}_U \in \mathcal{F}_0$, we know $M \geq 1$. Moreover, the derivative estimate, Proposition 2.21, implies that

$$|f'(0)| \leq \frac{1}{1/2} \cdot \sup_{\overline{B}_{1/2}} |f| \leq 2$$

for all $f \in \mathcal{F}_0$, so $M \leq 2$. Thus, we can now take a sequence $f_n \in \mathcal{F}_0$ such that

$$\lim_{n \to \infty} |f_n'(0)| = M.$$

Since $\mathcal{F}_0$ is a normal family by Montel's theorem (Theorem 5.13), after passing to a subsequence, we may assume f_n locally uniformly converges to a holomorphic function f on U, which satisfies $|f| \leq 1$ and $f(0) = 0$. We are now going to show that f is the desired conformal mapping by proving that f is both injective and surjective.

To see that f is injective, we note that since f_n' also converges locally uniformly (Weierstrass' theorem, Corollary 2.23), $|f'(0)| = M \geq 1$. In particular, f is not a constant map, so Hurwitz's theorem (Theorem 3.39) implies that f must be injective.

To see that f is surjective, we will use a contradiction argument and use the maximality of M. Suppose $a \in D \backslash f(U)$. A very rough idea is that the square root function will increase the derivative, to get which one needs to use the information of a missing point. We take an automorphism

$A_a \in \mathrm{Aut}(D)$ that sends a to 0. Then $A_a \circ f$ is non-vanishing on a simply connected open set, so by Corollary 4.9 again, we can take a holomorphic branch $h\colon U \to D$ with

$$h^2 = A_a \circ f,$$

which means

$$h = s \circ A_a \circ f$$

where $s(z) = \sqrt{z}$ is a branch of the square root function on $A_a \circ f(U)$ (so it only makes sense after composing with $A_a \circ f$). To compare their derivatives, we now take another automorphism $A_h \in \mathrm{Aut}(D)$ that sends $h(0)$ to 0, and consider

$$\widetilde{f} := A_h \circ h = A_h \circ s \circ A_a \circ f.$$

Then $\widetilde{f}$ is also an injective map from D to D with $\widetilde{f}(0) = 0$, and if we let $t(z) := z^2$ be the squaring function, we get

$$f = A_a^{-1} \circ t \circ A_h^{-1} \circ \widetilde{f}.$$

Now, $A_a^{-1} \circ t \circ A_h^{-1}$ is a map from D to D that sends 0 to 0 and it is not a rotation (because of t), so Schwarz's lemma implies

$$\left|\left(A_a^{-1} \circ t \circ A_h^{-1}\right)'(0)\right| < 1.$$

This implies

$$M = |f'(0)| = \left|\left(A_a^{-1} \circ t \circ A_h^{-1}\right)'\left(\widetilde{f}(0)\right)\right| \cdot \left|\widetilde{f}'(0)\right| < \left|\widetilde{f}'(0)\right|.$$

Since $\widetilde{f} \in \mathcal{F}_0$, this is a contradiction. Thus, f is surjective, and the proof is done. □

Note that, as in the existence part of the Dirichlet problem, in Theorem 5.77, no assertion about the boundary behavior of the constructed conformal map is mentioned. In some good situations, one can extend the map to the boundary; see, for example, [Ahl78, Chapter 6.2]. In general, Carathéodory's theorem asserts that this extension is always possible provided the boundary of U is a Jordan curve; see this page.

On the other hand, from Assignment 9, we know that there does not exist a simple version of the Riemann mapping theorem for multiply connected regions. It is, however, still possible to find a list of "standard models." One can see [Ahl78, Chapter 6.5].

5.5. **Riemann surfaces and uniformization.** We are now going to talk about what a Riemann surface really is. Roughly speaking, such an object is a space that locally looks like an open set in the complex plane. This local structure is based on the notion of local charts, on which there is a coordinate system from the Euclidean space.

Definition 5.79. Let M be a Hausdorff[17] topological space.

[17]Recall that this means that any two different points can be separated by disjoint open sets.

(1) A **holomorphic atlas** on M is an open cover U_α ($\alpha \in A$) of M and a family of homeomorphisms $\varphi_\alpha \colon U_\alpha \to V_\alpha$ where each V_α is an open set in $\mathbb{C}$ such that for any $\alpha, \beta \in A$, the **transition map**

$$\varphi_\beta \circ \varphi_\alpha^{-1} \colon \varphi_\alpha \left(U_\alpha \cap U_\beta\right) \to \varphi_\beta \left(U_\alpha \cap U_\beta\right)$$

is a holomorphic map between subsets in $\mathbb{C}$. With U_α's and φ_α's, (M, φ_α) is called a **Riemann surface**. Each U_α is called a **holomorphic chart**.

(2) Suppose M and N are two Riemann surfaces, with holomorphic charts $\varphi_\alpha \colon U_\alpha \to V_\alpha$ ($\alpha \in A$) and $\varphi'_{\alpha'} \colon U'_{\alpha'} \to V'_{\alpha'}$ ($\alpha' \in A'$). A continuous map $f \colon M \to N$ is called a **holomorphic map** between M and N if given any $\alpha \in A$ and $\alpha' \in A'$, the map

$$\varphi'_{\alpha'} \circ f \circ \varphi_\alpha^{-1} \colon \varphi_\alpha \left(f^{-1}\left(U'_{\alpha'}\right) \cap U_\alpha\right) \to V'_{\alpha'}$$

is a holomorphic map between subsets in $\mathbb{C}$.

(3) A holomorphic map $f \colon M \to N$ between Riemann surfaces is called a **(complex) diffeomorphism** if f^{-1} exists, is continuous, and is holomorphic.

(4) If M admits two holomorphic atlases (M, φ_α) and $(M, \varphi'_{\alpha'})$, we say that these two atlases are equivalent if the identity map $(M, \varphi_\alpha) \to (M, \varphi'_{\alpha'})$ is a complex diffeomorphism An equivalent class of holomorphic atlases is called a **complex structure** on M.

Example 5.80. We have encountered many Riemann surfaces. Let's find some holomorphic structures on them.

(1) The complex plane $\mathbb{C}$ is a Riemann surface, and itself can be a holomorphic atlas with a single chart $\mathbb{C}$. Similarly, any open subset of $\mathbb{C}$ is a Riemann surface.

(2) The Riemann sphere $\hat{\mathbb{C}}$ is a Riemann surface. We take an atlas that consists of two charts $U_1 = \mathbb{C}$ and $U_2 = \hat{\mathbb{C}} \setminus \{0\}$, with

$$\varphi_1 \colon U_1 \to \mathbb{C}, z \mapsto z, \text{ and}$$
$$\varphi_2 \colon U_2 \to \mathbb{C}, z \mapsto \frac{1}{z}.$$

One can then check that the transition map is

$$\varphi_2 \circ \varphi_1^{-1} \colon \mathbb{C} \setminus \{0\} \to \mathbb{C} \setminus \{0\}, z \mapsto \frac{1}{z},$$

which is a holomorphic map. Thus, they form a holomorphic atlas and $\hat{\mathbb{C}}$ is a Riemann surface. Without specification, when we mention $\hat{\mathbb{C}}$ as a Riemann surface, we mean using the standard complex structure induced by this holomorphic atlas.

(3) A complex cylinder $\mathbb{C}/\mathbb{Z}$ is a Riemann surface (see Example 5.74(3)). It is topologically equivalent to $S^1 \times \mathbb{R}$.

(4) Another common example are complex tori. Given two linearly independent complex numbers ω_1 and ω_2, they generate a lattice

$$\Lambda := \mathbb{Z}\omega_1 + \mathbb{Z}\omega_2.$$

Then $\mathbb{C}/\Lambda$ is a complex torus. Note that all these complex tori are homeomorphic to each other, but these homeomorphisms may not be complex diffeomorphisms.

Now after Riemann surfaces are rigorously introduced, we can talk about holomorphic maps into $\hat{\mathbb{C}}$. Suppose U is an open set in $\mathbb{C}$ and $f\colon U \to \hat{\mathbb{C}}$ is a holomorphic map between the two Riemann surfaces U and $\hat{\mathbb{C}}$ with their standard complex structures. Then by definition, this means that f is continuous and when composed with the local charts in Example 5.80(2), it is holomorphic. That is, both

$$\varphi_1 \circ f\colon \{z \in U : f(z) \neq \infty\} \to \mathbb{C}, z \mapsto f(z), \text{ and}$$
$$\varphi_2 \circ f\colon \{z \in U : f(z) \neq 0\} \to \mathbb{C}, z \mapsto \frac{1}{f(z)}$$

are holomorphic functions. By Proposition 3.15, this means that f is either a meromorphic function on U or identically ∞. One can similarly treat the case when U is an open set in $\hat{\mathbb{C}}$.$^{\text{ex}}$

5.5.1. *Model spaces and maps between them.* We now mention one of the most important results in complex analysis of one variable. It is the **uniformization theorem**, which classifies all possible simply connected Riemann surfaces up to complex diffeomorphisms.

Theorem 5.81 (Uniformization theorem)**.** Let M be a simply connected Riemann surface. Then M is complex diffeomorphic to $\mathbb{C}$, $\hat{\mathbb{C}}$, or the unit disc D.

As open sets in $\mathbb{C}$ are also naturally Riemann surfaces, one can view this as a much stronger generalization of the Riemann mapping theorem (Theorem 5.77). We will first see how we can use it to roughly get a whole classification of all Riemann surfaces. This will also be based on topological theories about covering spaces, and to implement these, we have to understand possible automorphisms of these model spaces.

Given a Riemann surface M, we define its automorphism group to be

$$\mathrm{Aut}(M) := \{f\colon M \to M \text{ is a complex diffeomorphism}\}.$$

Note that $\mathrm{Aut}(M)$ has a natural group structure by composition. We will combine our knowledge for the automorphism groups of the three model spaces with the following fact from topology about the universal cover of a topological space.

Theorem 5.82. Let M be a connected Riemann surface. Then there exists a simply connected Riemann surface $\widetilde{M}$ and a covering map $\pi\colon \widetilde{M} \to M$.

The non-trivial part is about the existence of $\widetilde{M}$ as a topological space and the covering map π. Once such a covering map is found, it can be used to pull back the complex structure from M to $\widetilde{M}$.$^{\text{ex}}$ As a corollary of Theorems 5.81 and 5.82, we obtain a classification result.

Corollary 5.83. Let M be a connected Riemann surface. Then there exist a model space $\widetilde{M} \in \left\{\mathbb{C}, \hat{\mathbb{C}}, D\right\}$ and a discrete subgroup Γ of $\mathrm{Aut}(\widetilde{M})$ such that M is complex diffeomorphic to $\widetilde{M}/\Gamma$.

Here, the discrete group Γ has to act **freely** on M so that the quotient is still a manifold. That is, any non-identity element has to have no fixed points when acting on M.[18] We will explain more in Corollary 5.86.

[18]The group Γ is the fundamental group of M.

To say more in this direction, we quickly find out the automorphism groups of these model spaces. For automorphisms of D, recall that from Schwarz's lemma (Corollary 2.29), we know that

$$\mathrm{Aut}(D) = \left\{ f(z) = e^{it} \cdot \frac{a - z}{1 - \bar{a}z} : t \in \mathbb{R} \text{ and } a \in D \right\}.$$

Next, we look at automorphisms of $\hat{\mathbb{C}}$. For the cases of both $\mathbb{C}$ and $\hat{\mathbb{C}}$, the following lemma will be useful.

Lemma 5.84. Let $f\colon \hat{\mathbb{C}} \to \hat{\mathbb{C}}$ be a holomorphic map. Then f is either a rational function or identically ∞.

Proof. Based on the discussion above, $f|_{\mathbb{C}}$ is a meromorphic function or identically ∞. Suppose f is not identically ∞. Then $f^{-1}(\infty)$ is a discrete set in $\hat{\mathbb{C}}$, so it is a finite set (by the compactness of $\hat{\mathbb{C}}$.

Let $p_1, \cdots, p_k$ be the poles of f and suppose $P_1, \cdots, P_k$ are the principal parts of f at p_j's (cf. Proposition 3.15). Then we know that the function

$$h(z) := f(z) - \sum_{j=1}^{k} P_j(z)$$

is a holomorphic map between $\hat{\mathbb{C}}$ and is a holomorphic function on $\mathbb{C}$. If $h(\infty) = f(\infty) \neq \infty$, then $h|_{\mathbb{C}}$ is a bounded holomorphic function and hence a constant by Liouville's theorem (Corollary 2.25), so f is a rational function. If $h(\infty) = f(\infty) = \infty$, then h has a pole at ∞, and hence h is a polynomial so f is still a rational function. $\square$

We can use this lemma to find out automorphisms of $\hat{\mathbb{C}}$ and $\mathbb{C}$.

Corollary 5.85. An automorphism of $\hat{\mathbb{C}}$ is a Möbius transformation, and an automorphism of $\mathbb{C}$ is a non-constant linear function. That is,

$$\begin{aligned} \mathrm{Aut}(\hat{\mathbb{C}}) &= \left\{ f(z) = \frac{az+b}{cz+d} : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C}) \right\} \text{ and} \\ \mathrm{Aut}(\mathbb{C}) &= \{ f(z) = az + b : a \neq 0 \}. \end{aligned}$$

Proof. Suppose $f \in \mathrm{Aut}(\hat{\mathbb{C}})$. Then by Lemma 5.84, $f = \frac{P(z)}{Q(z)}$ where P and Q are polynomials. Assume P and Q do not have common roots. Since f is injective, P and Q must be linear, so we can write

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

for some $a, b, c, d \in \mathbb{C}$. As f is non-constant, we have $ad \neq bc$. Thus, after rescaling, we may assume $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{C})$.

For $\mathrm{Aut}(\mathbb{C})$, by Assignment 4, we already know that an element must be of the form $az + b$ due to its injectivity. We include an argument here for completeness. Suppose $f \in \mathrm{Aut}(\mathbb{C})$. By the open mapping theorem (Corollary 3.6), $f(B_1)$ is a non-trivial open set in $\mathbb{C}$. In particular, since

f is bijective, we have $f(\mathbb{C} \setminus \overline{B}_1)$ does not contain the open set $f(B_1)$. Hence, f does not have an essential singularity at ∞ by the Casorati–Weierstrass theorem (Proposition 3.17), noting that $\mathbb{C} \setminus B_1$ is an open (punctured) neighborhood of ∞. Thus, f has either a removable singularity or a pole at ∞. For the former case, f is bounded and hence constant, which cannot happen. Thus, f has a pole at ∞, so f is a polynomial. Since f is bijective, f must be a non-constant linear function. □

We can now use this information about automorphism groups to get a finer classification result.

Corollary 5.86. A connected Riemann surface is complex diffeomorphic to one of the following spaces.

(1) (Elliptic type) The Riemann sphere $\hat{\mathbb{C}}$.

(2) (Parabolic type) The complex plane $\mathbb{C}$, the punctured plane $\mathbb{C}^*$, and a complex torus $\mathbb{C}/\Lambda$.

(3) (Hyperbolic type) A hyperbolic surface D/G for some $G \subseteq \operatorname{Aut}(D)$.

We basically do not get anything more specific in the third case. However, we do get cleaner lists for the other two types.

Proof. Suppose $M = \hat{\mathbb{C}}/\Gamma$ for some discrete $\Gamma \leq \operatorname{Aut}(\hat{\mathbb{C}})$ that acts freely on $\hat{\mathbb{C}}$. However, in $\operatorname{Aut}(\hat{\mathbb{C}})$, all the non-identity elements have at least a fixed point (cf. Problem 10(3) in Assignment 1). Thus, the only possibility of Γ is the trivial group, so $M = \hat{\mathbb{C}}$.

Suppose $M = \mathbb{C}/\Gamma$ for some discrete $\Gamma \leq \operatorname{Aut}(\mathbb{C})$ that acts freely on $\mathbb{C}$. Since a non-identity element cannot have a fixed point, an automorphism $az + b$ must satisfy $a = 1$, and hence it is of the form $f_c(z) = z + c$. Since Γ is discrete, we can find f_{c_0} so that

$$|c_1| = \min_{f_c \in \Gamma \setminus \{0\}} |c|.$$

After a conjugation, we may assume $c_1 = 1$ (cf. Problem 10(5) in Assignment 1). We now deal with three situations.

If Γ is generated by f_1, then $\Gamma = \mathbb{Z}$ and hence $M = \mathbb{C}/\mathbb{Z}$, which is equivalent to $\mathbb{C}^*$ as observed in Example 5.74(3).

If there exists f_c that is not generated by f_1 (so that $c \in \mathbb{C} \setminus \mathbb{Z}$ with $|c| \geq 1$), then we can take such a $c = c_2$ with the smallest length (again using the discrete property) and conclude that $\Gamma = \mathbb{Z} \cdot f_1 + \mathbb{Z} \cdot f_{c_2}$. In fact, if there were another element not generated by f_1 and f_{c_2}, then Γ won't be discrete.$^{\text{ex}}$. Thus, $M = \mathbb{C}/(\mathbb{Z} \cdot f_1 + \mathbb{Z} \cdot f_{c_2})$ is a complex torus (cf. Example 5.80(4)). □

5.5.2. *Green's function on a Riemann surface.* We will talk about half of the proof of Theorem 5.81. The proof is based on **Green's function**. This is related to finding harmonic functions on a Riemann surface. Crucial ingredients include most of the tools we develop in Section 5.3. Many of them are local, and hence we expect that they also hold in a general Riemann surface. We first talk about what harmonicity means on M, on which sometimes we will use z to represent a holomorphic chart (when it is not too confusing, otherwise, we will still use φ).

Definition 5.87. Let M be a Riemann surface, $u \colon M \to \mathbb{R}$ be a continuous function, and $p \in M$.

(1) We say u is harmonic near p if there is an open neighborhood U of p and a holomorphic chart $\varphi\colon U \to V \subseteq \mathbb{C}$ such that $u \circ \varphi^{-1}$ is a harmonic function on V.

(2) We say $u\colon U \to [-\infty, \infty)$ is subharmonic on an open set $U \subseteq M$ if for any connected compact set $K \subseteq U$ with $\partial K \neq \emptyset$, given a harmonic function v on $\mathrm{int}K$ with $u|_{\partial K} \leq v|_{\partial K}$, it follows that $u \leq v$ on K.

Importantly, the definition of harmonicity does not depend on the choice of holomorphic charts, as this is equivalent to having a local harmonic conjugate by Lemma 1.8. Also note that here we use the maximum principle to define subharmonic functions, which makes sense as we have Theorem 5.56 in the planar case.

We need a way to produce harmonic functions. In fact, Perron's method introduced in Section 5.3.3 is general enough for a large class of families of functions. To say this, we recall the harmonic lifting introduced in Proposition 5.58. Let v be a subharmonic function on M. We say an open set $U \subseteq M$ is a **coordinate ball** if there exists a holomorphic chart $\varphi\colon U \to B$ where B is an open ball in $\mathbb{C}$. Then the **harmonic lifting** of v in U is defined as in the planar case. To be explicit, $\widetilde{v}|_{M\setminus U} = v|_{M\setminus U}$, and for $p \in U$,

$$\widetilde{v}(p) = \frac{1}{2\pi}\int_0^{2\pi} v \circ \varphi^{-1}\left(re^{it}\right) \cdot \frac{r^2 - |\varphi(p)|^2}{|re^{it} - \varphi(p)|^2} dt$$

if $B = B_r(0) \subseteq \mathbb{C}$ and $v|_{\partial U} > -\infty$. Thus, we can always find a harmonic lifting of a (non-trivial) subharmonic function in a coordinate ball.

We say a family $\mathcal{F}$ of subharmonic functions on M is a **Perron family** if it satisfies the following closure properties.

(1) If $v_1, v_2 \in \mathcal{F}$, then $\max\{v_1, v_2\} \in \mathcal{F}$.

(2) If $v \in \mathcal{F}$ and U is a coordinate ball, then the harmonic lifting of v in U is also in $\mathcal{F}$.

Then the proof of Theorem 5.61 essentially tells us the following result.

Theorem 5.88. Let $\mathcal{F}$ be a Perron family on M and consider

$$u(p) := \sup\{v(p) : v \in \mathcal{F}\}.$$

Then either u is identically ∞ or u is a harmonic function on M.

This theorem requires Harnack's estimate and compactness. The proofs of them in the planar case are local, so the results are still true in the general case on M. We reiterate the results here.

Proposition 5.89. Let $u\colon U \to \mathbb{C}$ be a positive harmonic function on an open set in a Riemann surface M. Given a compact set $K \subseteq U$, there exists $C_K < \infty$ such that

$$\sup_K u \leq C_K \cdot \inf_K u.$$

As a result, if U is connected and u_n is an increasing sequence of harmonic functions on U, then u_n either locally uniformly converges to a harmonic function on U, or locally uniformly converges to ∞.

Now, we can introduce Green's function. Such a function is a harmonic function with a singularity modeled on that of $\log|z|$ at 0. Given a Riemann surface M and $q \in M$, we take a coordinate ball U near q, and we can talk about the function $\log|z(p)|$ for $p \in z^{-1}(B \setminus \{0\})$ if $z\colon U \to B = B_1(0)$ is a holomorphic chart of M near q. For simplicity, we may write $\log|z| = \log|z(p)|$ and view it as a function on $U \setminus \{q\}$. To get Green's function, consider

$$\mathcal{F}_q := \left\{ v \in \mathcal{C}\left(M \setminus \{q\}\right) : \substack{v=0 \text{ away from a compact set in } M, \\ v \text{ is subharmonic on } M\setminus\{q\}, \text{ and} \\ v+\log|z| \text{ is subharmonic on } U.} \right\}.$$

The collection $\mathcal{F}_q$ forms a Perron family.[ex] This family is non-empty since $0 \in \mathcal{F}_q$, so we may use it to produce a harmonic function possibly with a singularity.

Definition 5.90. If there exists $q \in M$ such that the function

$$g(p,q) := \sup\{v(p) : v \in \mathcal{F}_q\}$$

is a harmonic function on $M \setminus \{q\}$ (that is, not identically infinity), then we say Green's function exists on M. Otherwise, we say M does not have Green's function.

When $g(\cdot,q)$ exists, we will say it is Green's function with a pole at q, or Green's function based at q. Right now the definition seems to depend on a point, but later we will see that when M is connected, it doesn't matter which point q we choose. We start with some basic properties of Green's function.

Theorem 5.91. Suppose M has Green's function $g(\cdot,q)$ based at $q \in M$ near which there is a coordinate ball U. Then

(1) $g(\cdot,q)$ is a positive harmonic function on $M \setminus \{q\}$;

(2) $g(\cdot,q) + \log|z|$ is harmonic on U;

(3) given a positive harmonic function h on $M \setminus \{q\}$ such that $h + \log|z|$ is harmonic on U, we have

$$h(p) \geq g(p,q) \tag{5.92}$$

for $p \in M \setminus \{q\}$.

One can interpret the harmonicity of

$$g(\cdot,q) + \log|z| = g(\cdot,q) - \log\frac{1}{|z|}$$

by saying that $g(\cdot,q)$ has a logarithmic singularity at q. A particular consequence of Theorem 5.91 is that $g(\cdot,q)$ is the minimal positive harmonic function with the singularity modeled on $\log|z|$. Especially, if Green's function with a pole at q exists, then it is unique and is characterized by this minimality property.

Example 5.93. We look at the Riemann surface $M = D = B_1(0) \subseteq \mathbb{C}$. Given $z_0 \in D$, Green's function based at z_0 exists and is

$$g(z,z_0) = \log\left|\frac{1-\overline{z_0}z}{z-z_0}\right|.$$

We look at the case when $z_0 = 0$, which leads to

$$g(z, 0) = \log \frac{1}{|z|}.$$

To show this, suppose h is a positive harmonic function on $D \setminus \{0\}$ such that $h(z) + \log |z|$ is harmonic. The maximum principle then implies

$$h(z) + \log |z| \geq \min_{z \in \partial D} (h(z) + \log |z|) = \min_{\partial D} h \geq 0$$

for any $z \in D$, and in particular, for $z \in D \setminus \{0\}$. This implies

$$h(z) \geq -\log |z| = \log \frac{1}{|z|},$$

which means that $\log \frac{1}{|z|}$ satisfies the minimality (5.92). Thus, $g(z, 0) = \log \frac{1}{|z|}$.

Proof of Theorem 5.91. By the construction, we know that $g(\cdot, q)$ is a non-negative harmonic function on $M \setminus \{q\}$. To show that it is positive, suppose $U = z^{-1}(B_1)$ is a coordinate ball given by $z \colon U_1 \to B_2$ with $U \subseteq U_1 \subseteq M$ and $z(q) = 0$. Then we consider a function

$$v(p) := \begin{cases} -\log |z(p)| & \text{if } p \in U \\ 0 & \text{otherwise} \end{cases},$$

which lies in $\mathcal{F}_q$. Thus,

$$g(p, q) \geq -\log |z(p)| > 0 \text{ for } p \in U \setminus \{q\} \tag{5.94}$$

since $|z(p)| < 1$. By the strong maximum principle, $g(\cdot, q) > 0$ on $M \setminus \{q\}$. Otherwise, it would attain its minimum at an interior point.

Next, we show that $g(\cdot, q) + \log |z|$ is harmonic on U. We let

$$M := \sup_{\partial U} g(\cdot, q) = \sup_{|z(p)|=1} g(p, q).$$

This implies that for any $v \in \mathcal{F}_q$, $v(p) \leq M$ for $p \in \partial U$, and the subharmonicity of $v + \log |z|$ implies that for any $p \in U$,

$$v(p) + \log |z(p)| \leq \sup_{\partial U} (v + \log |z|) \leq M + 0,$$

and taking its supremum over $v \in \mathcal{F}_q$ leads to

$$g(p, q) + \log |z(p)| \leq M$$

for any $p \in U$. Thus, this and (5.94) imply that $g(\cdot, q) + \log |z|$ is a bounded harmonic function on U, and hence the singularity at q is removable. (The local version is done in Assignment 8, which also applies for a harmonic function on a Riemann surface.$^{\text{ex}}$)

Finally, we prove the minimality property. Let h be a positive harmonic function on $M \setminus \{q\}$ such that $h + \log |z|$ is harmonic on U. Take any $v \in \mathcal{F}_q$. We know that $v - h$ is subharmonic on $M \setminus \{q\}$. Moreover, since

$$v - h = (v + \log |z|) - (h + \log |z|),$$

it is also subharmonic on U. Thus, $v-h$ is subharmonic on M. If we let K be the (compact) support of v, we get $v-h<0$ on $M \setminus K$. Thus, the maximum principle implies $v-h \leq 0$ also on K. Hence, we get that $v \leq h$ for all $v \in \mathcal{F}_q$, so $g(\cdot,q) \leq h$. □

As a consequence, we know that when $g(\cdot,q)$ exists, it satisfies

$$\inf_{M\setminus\{q\}} g(\cdot,q)=0. \tag{5.95}$$

Otherwise, if it had positive infimum m, then the function $h(p):=g(p,q)-m$ would be a positive harmonic function that violates (5.92).

We will talk about two more properties of Green's functions. First, we talk about changing basepoints of Green's functions.

Theorem 5.96. Let M be a connected Riemann surface and suppose $g(\cdot,q)$ exists for some $q \in M$. Then $g(\cdot,p)$ exists for all $p \in M$.

This will be proven by the help of an auxiliary harmonic function. Again, assume $q_0 \in U \subseteq U_1 \subseteq M$ where $z\colon U_1 \to B_2 \subseteq \mathbb{C}$ is a holomorphic chart with $z(q_0)=0$ and $U=z^{-1}(B_1)$. We consider a family

$$\mathcal{F}:=\left\{\text{subharmonic function } v \text{ on } M\setminus\overline{U} : v \leq 1 \text{ and } v \text{ has compact support}\right\}.$$

This defines a Perron family,[ex] and the bound in the construction implies $v \leq 1$, and hence the function

$$w(p):=\sup\left\{v(p): v\in\mathcal{F}\right\}$$

is harmonic on $M\setminus\overline{U}$ with $0\leq w\leq 1$. Since there is a harmonic barrier at each point of ∂U, one can show that

$$\lim_{p\to p_0} w(p)=1 \tag{5.97}$$

for any $p_0 \in \partial U$ essentially by Theorem 5.69. The strong maximum principle then implies that either w is identically 1, or

$$0<w<1 \text{ on } M\setminus\overline{U}. \tag{5.98}$$

This second situation in fact characterizes the existence of Green's function.

Proposition 5.99. Let q_0, U, and w be constructed as above. If $g(\cdot,q_0)$ exists, then (5.98) holds. Conversely, if (5.98) holds, then $g(\cdot,q)$ exists for all $q\in U$.

Proof. First we assume the existence of $g(\cdot,q_0)$. Consider

$$m:=\min_{p\in\partial U} g(p,q_0),$$

which is positive by Theorem 5.91(1). Then for any $v\in\mathcal{F}$, the maximum principle implies

$$v(p)\leq 1=\frac{1}{m}\min_{\partial U} g(\cdot,q_0)\leq\frac{1}{m}g(p,q_0) \text{ for } p\in M\setminus\overline{U}.$$

Thus, taking its supremum over $v \in \mathcal{F}$ leads to

$$w \leq \frac{1}{m} g(\cdot, q_0) \text{ on } M \setminus \overline{U}.$$

With this, if $\inf_{M \setminus \overline{U}} w > 0$, then it would imply

$$\inf_{M \setminus \{q\}} g(\cdot, q_0) \geq \min \left\{ m \cdot \inf_{M \setminus \overline{U}} w, \inf_{U \setminus \{q_0\}} g(\cdot, q_0) \right\} > 0,$$

which contradicts (5.95). Hence, $\inf_{M \setminus \overline{U}} w = 0$, and hence w cannot be a constant.

Next, we assume $0 < w < 1$. For any $q \in U$, we want to show that there is a uniform upper bound for functions in $\mathcal{F}_q$. We take $U_2 := z^{-1}(B_{3/2})$ that contains U. For any $v \in \mathcal{F}_q$, we let

$$M := \sup_{\partial U_2} v.$$

It remains to show that M can be bounded independently of v (which implies that the supremum of u is finite on ∂U_2). We take the bounds

$$\begin{aligned} C_1 &:= \sup_{\partial U_2} \left| \log |z - z(q_0)| \right|, \text{and} \\ C_2 &:= \sup_{\partial U_2} w < 1. \end{aligned}$$

Since $v + \log |z - z(q_0)|$ is subharmonic, by the maximum principle, for any $p \in U_2$, we have

$$v(p) + \log |z(p) - z(q_0)| \leq \sup_{\partial U_2} \left(v + \log |z - z(q_0)| \right) \leq M + C_1.$$

In particular, for $p \in \partial U$, we can estimate

$$\begin{aligned} v(p) &\leq M + 2C \\ &= (M + 2C) \cdot w(p) \end{aligned}$$

where we use $w|_{\partial U} = 1$ from (5.97). The maximum principle then implies

$$v(p) \leq (M + 2C_1) w(p)$$

for $p \in M \setminus U$, which contains ∂U_2. Taking the supremum over $p \in \partial U_2$, we then derive

$$M \leq (M + 2C_1) \cdot C_2$$

Since $C_2 < 1$, we get

$$M \leq \frac{2C_1 C_2}{1 - C_2}.$$

This proves the boundedness of v, and the existence of $g(\cdot, q)$ follows. □

Using this, we can then prove Theorem 5.96. It follows from the continuity argument again.

Proof of Theorem 5.96. By Proposition 5.99, the set of q such that $g(\cdot, q)$ exists is open and the set of q such that $g(\cdot, q)$ does not exist is also open. Thus, the theorem follows from the connectivity of M. □

Thus, from now on, when we talk about the existence of Green's function, we will usually say that g or $g(p,q)$ exists or simply that Green's function exists.

Finally, we talk about the symmetry property of Green's function. We state the result, and will give a rough idea about how to prove it.

Proposition 5.100. Let M be a connected Riemann surface on which Green's function g exists. Then for $p \neq q$ in M, $g(p,q) = g(q,p)$.

The main idea of the proof is to utilize Green's second identity, which says that for any C^2 functions u and v on an open set U, one has

$$\int_U (v\Delta u - u\Delta v)\, dA = \int_{\partial U} \left(v\frac{\partial u}{\partial \mathbf{n}} - u\frac{\partial v}{\partial \mathbf{n}} \right) ds,$$

where $\frac{\partial}{\partial \mathbf{n}}$ is the directional derivative in the normal direction. The formula basically follows from integration by parts, but one has to be careful when trying to apply it to Green's function g for two reasons: (1) g has singularities, and (2) g does not have compact support. However, if we take it for granted that it applies to g and also g is the "fundamental solution" to the Laplacian equation, meaning that

$$\int_M f \cdot \Delta g(\cdot,q) dA = -2\pi f(q)$$

for a smooth function f with suitable behavior at infinity, then we can (formally) derive

$$-2\pi g(p,q) = \int_M g(\cdot,q)\Delta g(\cdot,p) dA = \int_M \Delta g(\cdot,q) g(\cdot,p) dA = -2\pi g(q,p)$$

for $p \neq q$. Based on the positivity of g, Proposition 5.100 then follows.

If one is interested in a rigorous discussion about Proposition 5.100, one can see [Gam01, Section XVI.4].

We end this section by mentioning that, in general, Green's function may not exist. There is a quick criterion to check this.

Lemma 5.101. Let M be a connected Riemann surface. If there is a non-constant bounded holomorphic function on M, then Green's function exists on M.

This gives us another reason why we can find Green's function on a disc in Example 5.93.

Proof. Fix a point $q \in M$. Take a non-constant bounded holomorphic function f, and after a translation and scaling, we may assume $f(q) = 0$ and $|f| < 1$ on M. Then given any $v \in \mathcal{F}_q$, say with support K, $v + \log|f|$ is subharmonic on M with

$$v + \log|f| < 0 + \log 1 = 0 \text{ on } M \setminus K.$$

The maximum principle then implies $v + \log|f| \leq 0$ on M, so

$$v(p) \leq -\log|f(p)|$$

for any $p \in M \setminus \{q\}$. Finding a point p such that $f(p) \neq 0$, we get a uniform upper bound of $v(p)$ for all $v \in \mathcal{F}_q$, and the existence of g follows. $\square$

In fact, the existence of Green's function is equivalent to the existence of a non-constant subharmonic function bounded from above.[ex19] From this, for example, one can deduce that Green's function does not exists on $\mathbb{C}$ or $\hat{\mathbb{C}}$.

However, if we allow two poles, then such a **bipolar Green's function** does exist in general. An example on $\hat{\mathbb{C}}$ is the function $G(z) = -\log|z|$, which has poles at both 0 and ∞. In general, given a connected Riemann surface M and two distinct points q_1 and q_2 in M, we can find such a bipolar Green's function. We state the existence result here.

Theorem 5.102. Let M be a connected Riemann surface and U_1 and U_2 be two disjoint coordinate balls around q_1 and q_2 in M. Then there exists a bipolar Green's function $G(\cdots, q_1, q_2)$ with the following properties, where we let z_1 and z_2 be holomorphic charts on U_1 and U_2.

(1) $G(\cdot, q_1, q_2)$ is harmonic on $M \setminus \{q_1, q_2\}$.

(2) $G(\cdot, q_1, q_2) + \log|z_1|$ is harmonic on U_1.

(3) $G(\cdot, q_1, q_2) - \log|z_2|$ is harmonic on U_2.

(4) $G(\cdot, q_1, q_2)$ is bounded on $M \setminus (U_1 \cup U_2)$.

Bipolar Green's function can be use to prove the uniformization of elliptic and parabolic Riemann surfaces. We won't talk about it, and interested readers can see [Gam01, Section XVI.5] for discussion on bipolar Green's function.

5.5.3. *Monodromy theorem.* Besides Green's function, another tool we will need in the uniformization theorem is analytic continuation, as we have seen many times in this course. Here, we will use the topological condition to say when such an extension is well-defined.

We first make rigorous what we mean by it here. Let $\gamma\colon [a,b] \to U$ be a path in an open set U in a Riemann surface M, and suppose a holomorphic function f is defined near $\gamma(a)$. We say f is **analytically continuable** along γ if for any $t \in [a,b]$, there exist $r_t > 0$ and a coordinate ball $z_t\colon U_t \to B_{r_t} \subseteq \mathbb{C}$ near $\gamma(t)$ with $z_t(\gamma(t)) = 0$ and a convergent power series P_t on B_{r_t} such that

(1) $P_a \circ z_a(\gamma(a)) = f(\gamma(a))$, and

(2) for $p \in U_{t_1} \cap U_{t_2}$ for some $t_1, t_2 \in [a,b]$, they satisfy

$$P_{t_1} \circ z_{t_1}(p) = P_{t_2} \circ z_{t_2}(p). \tag{5.103}$$

In this case, we can extend the domain of f from a neighborhood of $\gamma(a)$ to a neighborhood of the whole curve γ. By the unique continuation principle (Corollary 3.4), we know that such an extension is unique if it exists.

Analytic continuation is, however, not always possible. In fact, the discussion in Section 4.1.1 essentially tells us that starting from $f(z) = \log|z| + it$ for $z = e^{it}$ where $t \in (-1/10, 1/10)$, this local function f is NOT analytically continuable along $\gamma(s) = e^{is}$ for $s \in [0, 2\pi]$; see Figure 9. We want to use a topological condition to rule out this bad situation, and it is called the **monodromy theorem**.

[19] One can see Exercises 5 and 6 in [Gam01, Section XVI.3].

Theorem 5.104. Let M be a Riemann surface and $p, q \in M$, and let f is a holomorphic function defined on a neighborhood of p. Let $\gamma_0, \gamma_1 \colon [0,1] \to M$ be two paths with $\gamma_0(0) = \gamma_1(0) = p$ and $\gamma_0(1) = \gamma_1(1) = q$. Suppose the following conditions are satisfied.

(1) γ_0 and γ_1 are homotopic to each other through a homotopy $\gamma_s := H(s, \cdot)$ where $H \colon [0,1] \times [0,1] \to M$ is a continuous map with $H(s,0) = p$ and $H(s,1) = q$ for all $s \in [0,1]$.

(2) For any $s \in [0,1]$, f is analytically continuable along γ_s.

Then the analytic continuations of f along γ_0 and γ_1 coincide at q.

In the monodromy theorem, by saying that the two continuations coincide at q, we mean that there exists an open neighborhood of q on which the two functions are the same. In this statement, we only care about the "local information" of holomorphic functions, and two functions are viewed as the same when they coincide locally. This inspires the notion of the **germ** of a function at a point, which is useful in many different settings.

We will prove the theorem after some preliminary examinations of the continuation process. We start with the continuity of the convergence radius in a coordinate ball.

Lemma 5.105. Let M be a Riemann surface and let $z \colon U \to B_1$ be a coordinate ball near $p_0 \in M$. Let f be a holomorphic function on U, and for any $p \in U$, we let $R(p)$ be the radius of convergence of the power series expansion of $f \circ z^{-1}$ centered at $z(p)$. Then $R \colon U \to \mathbb{R} \cup \{\infty\}$ is a continuous function.

One can think about the Euclidean case. For example, for $f(z) = \frac{1}{1-z}$, we have $R(z) = |z-1|$.

Proof. We work on points in B_1, on which $g := f \circ B_1$ is a holomorphic function. For $z_1, z_2 \in B_1$, we know that $B_{R(z_2)}(z_2)$ cannot contain $\overline{B}_{R(z_1)}(z_1)$ by the definition of R. Thus, this implies

$$R(z_2) \le R(z_1) + |z_1 - z_2|.$$

By symmetry, we also have $R(z_1) \le R(z_2) + |z_1 - z_2|$, so we can conclude

$$|R(z_1) - R(z_2)| \le |z_1 - z_2|.$$

Thus, R is Lipschitz, and hence, in particular, R is continuous. ☐

Note that the radius function R does depend on the choice of a coordinate function z. However, when a holomorphic function can be continuated, the compatibility condition (5.103) guarantees that there is a well-defined radius function R on the neighborhood $U_{\gamma,g} := \bigcup_{t \in [a,b]} U_t$ of γ. Based on this, we can get a preliminary version of the monodromy theorem.

Corollary 5.106 (Baby monodromy)**.** Let $\gamma \colon [a,b] \to M$ be a path in a Riemann surface M and let f be a holomorphic function locally defined near $\gamma(a)$. Suppose f is analytically continuable along γ, and let δ be the infimum of the convergence radius function R, defined on a neighborhood $U_{\gamma,f}$, on γ. That is,

$$\delta := \inf_{\gamma} R.$$

If $\sigma\colon [a,b] \to M$ is another path with $\sigma(a) = \gamma(a)$ and $\sigma(b) = \gamma(b)$ such that

$$\sup_{[a,b]} |\sigma - \gamma| < \delta, \tag{5.107}$$

then f can also be continuated along σ. Moreover, the analytic continuations of f along γ and σ coincide at $\gamma(b) = \sigma(b)$.

Remark 5.108. Note that (5.107) has to be interpreted properly since we do not define a metric space structure on M. Here, (5.107) specifically means that using the coordinate charts z_t's, for any $t \in [a,b]$, $\sigma(t) \in U_t$ and

$$|z_t(\sigma(t)) - z_t(\gamma(t))| < \delta.$$

Also note that $\delta > 0$ by Lemma 5.105.

Proof of Corollary 5.106. For each $t \in [a,b]$, since $\sigma(t) \in U_t$, we can use the same function P_t on B_{r_t} directly. To be specific, we can take a neighborhood $V_t \subseteq U_t$ of $\sigma(t)$ and define

$$w_t(p) := z_t(p) - z_t\left(\sigma(t)\right)$$

so that $w_t(\sigma(t)) = 0$. Then we define

$$Q_t(z) := P_t\left(z + z_t\left(\sigma(t)\right)\right),$$

which is well-defined by (5.107). Thus, if $p \in V_{t_1} \cap V_{t_2}$ for some $t_1, t_2 \in [a,b]$, we get

$$\begin{aligned} Q_{t_1} \circ w_{t_1}(p) &= P_{t_1} \circ z_{t_1}(p) \\ &= P_{t_2} \circ z_{t_2}(p) = Q_{t_1} \circ w_{t_2}(p), \end{aligned}$$

so we can continuate f along σ. Moreover, we have

$$Q_b \circ w_b(\sigma(b)) = P_b \circ z_b(\gamma(b)),$$

so the two continuations agree near $\gamma(b)$. $\square$

We can now repeat this process to prove the monodromy theorem.

Proof of Theorem 5.104. For each $s \in [0,1]$, we let $z_{s,t}\colon U_{s,t} \to B_{s,t}$ be the coordinate balls used in the continuation of f along γ_s, and let $P_{s,t}$ be the analytic function on $B_{s,t}$. By choosing $z_{s,0}$ to be the same coordinate ball for all $s \in [0,1]$, we know that $P_{s,0}$'s are the same analytic functions on $B_{s,0}$, and hence we get a well-defined convergence radius function R on

$$U := \bigcup_{s\in[0,1]} \bigcup_{t\in[0,1]} U_{s,t},$$

which is a neighborhood of the compact set

$$K := \bigcup_{s\in[0,1]} \bigcup_{t\in[0,1]} \gamma_s\left(t\right).$$

By Lemma 5.105, we can take $\delta := \inf_K R > 0$. Then by the compactness, we can choose finitely many $0 < s_1 < \cdots < s_N < s_{N+1} = 1$ such that

$$\sup_{t\in[0,1]} \left|\gamma_{s_n}(t) - \gamma_{s_{n+1}(t)}\right| < \delta$$

for all $n = 1, \cdots, N$ in the sense of Remark 5.108 The theorem now follows from the unique continuation theorem and Corollary 5.106 since if we let $f_{s,1}$ be the continuation of f along γ_s near $\gamma(1)$, we get

$$f_{0,1}(p) = f_{s_1,1}(p) = \cdots = f_{s_{N+1},1}(p) = f_{1,1}(p)$$

for p near $\gamma(1)$. $\square$

The monodromy theorem is particularly useful when M is simply connected. We will use it to deal with one of the cases of the uniformization theorem (Theorem 5.81).

5.5.4. *Uniformization of hyperbolic Riemann surfaces.* When Green's function exists on a Riemann surface M, we say M is **hyperbolic**. In this section, we will show that a simply connected hyperbolic Riemann surface is complex diffeomorphic to D. We know that the existence of Green's function is related to the existence of a bounded non-trivial holomorphic function by Lemma 5.101. This then characterizes hyperbolic Riemann surfaces.

Theorem 5.109 (Uniformization of hyperbolic Riemann surfaces)**.** Let M be a simply connected Riemann surface. If M is hyperbolic, then M is complex diffeomorphic to the unit disc D.

Proof. Fix a point $q_0 \in M$ and let $g := g(\cdot, q_0)$ be Green's function based at q_0. We will use g to define a holomorphic function on M.

By Theorem 5.91(2), we can find a coordinate ball $z\colon U \to B_1$ near q_0 such that $z(q_0) = 0$ and $g + \log|z|$ is a harmonic function on U. Thus, by Lemma 1.8 (applied on B_1), we can find a harmonic conjugate h of $g + \log|z|$ on U. That is, $f := (g + \log|z|) + ih$ is a holomorphic function on U. We then take

$$\varphi := z \cdot e^{-f},$$

which is a holomorphic function on U that satisfies

(1) $|\varphi| = |z| \cdot e^{-g-\log|z|} = e^{-g}$, and

(2) $\varphi(q_0) = \lim_{q\to q_0} |z(q)| \cdot e^{-f(q_0)} = 0$.

These properties and the finiteness of g away from q_0 imply that q_0 is the unique (simple) zero of φ on U.

Next, we can extend φ to the whole surface M. In fact, given any $p \in M$, we can find a path γ connecting q_0 and p and finitely many coordinate balls $z_n\colon U_n \to B_n$ for $n = 1, \cdots, N$ such that U_n's cover γ. To be more specific, we may assume $U_1 = U$ chosen above and we can choose the rest of U_n's such that

(1) $q_0 \in U_1 \setminus \bigcup_{n=2}^{N} U_n$,

(2) $p \in U_N$, and

(3) $U_n \cap U_{n+1} \neq \emptyset$ for all $n = 1, \cdots, N-1$.

Then by finding a harmonic conjugate of g on each U_n for $n > 1$, we obtain a continuation of φ_γ along γ with $\varphi_\gamma = \varphi$ on U. To be more specific, if the extension to U_n has been done, then we know $\varphi = e^{-g-ih_n}$ for some harmonic h_n. (This includes the case when $n = 1$.) Then on U_{n+1}, we can find a harmonic conjugate h_{n+1} of g such that $h_{n+1} = h_n$ on $U_n \cap U_{n+1}$. This is because of the uniqueness of harmonic conjugates up to a constant. We then extend φ to U_{n+1} by defining it to be $e^{-g-ih_{n+1}}$.

We can now apply the monodromy theorem to finish the extension. Given two paths γ_0 and γ_1, both of which connect q_0 to p, we can find a homotopy γ_s such that each γ_s also connects q_0 to p. The argument in the preceding paragraph implies that φ (on U) is analytically continuable along each γ_s. Thus, Theorem 5.104 implies that $\varphi_{\gamma_0} = \varphi_{\gamma_1}$ on the intersection of their domains. Thus, we get a well-defined continuation φ on M such that $|\varphi| = e^{-g}$ on M. The finiteness of g away from q_0 again implies that q_0 is the unique zero of φ on M.

We've successfully obtained a holomorphic map φ on M. We notice that the positivity of g implies that $|\varphi| = e^{-g} < 1$, so we actually obtain a holomorphic map (or, in fact, a holomorphic function)

$$\varphi\colon M \to D.$$

It then remains to show that φ is an injective map. If this is done, then we know that φ is a complex diffeomorphism. Moreover, $\varphi(M)$ is also simply connected so it is conformally equivalent to D by the Riemann mapping theorem (Theorem 5.77).

To show that φ is injective, given $p_0 \in M \setminus \{q_0\}$, we consider[20]

$$\psi(p) := \frac{\varphi(p) - \varphi(p_0)}{1 - \overline{\varphi(p_0)}\varphi(p)},$$

which again defines a holomorphic function from M to D. Clearly $\psi(p_0) = 0$, and the injectivity of φ will follow if we prove that p_0 is the unique zero of ψ.

We start by relating ψ to Green's function. Let n be the order of the zero p_0 of ψ. Given $u \in \mathcal{F}_{p_0}$, by definition, we know that

$$u + \frac{1}{n}\log|\psi|$$

is subharmonic on M. If we let K be the (compact) support of u, we know that $u + \frac{1}{n}\log|\psi| < 0$ on $M \setminus K$, so the maximum principle implies that

$$u + \frac{1}{n}\log|\psi| < 0$$

on M. Taking the supremum over $u \in \mathcal{F}_{p_0}$, we conclude that

$$g(\cdot, p_0) + \frac{1}{n}\log|\psi| \le 0$$

[20]Recall that this comes from the automorphism group of D, as our goal is to relate M to D, on which we know how to move the zeros of a function by composing with those automorphisms.

on M. Thus, plugging in q_0 and use the symmetry of Green's function (Proposition 5.100), we get

$$\begin{aligned} 0 &\geq g(q_0, p_0) + \frac{1}{n} \log |\psi(q_0)| \\ &= g(p_0, q_0) + \frac{1}{n} \cdot \log \left| \frac{-\varphi(p_0)}{1} \right| \\ &= g(p_0) + \frac{1}{n} \log e^{-g(p_0)} = \left(1 - \frac{1}{n}\right) g(p_0). \end{aligned}$$

Since $p_0 \neq q_0$, we know $g(p_0) > 0$, so $n = 1$. Then the (in)equality also implies that the harmonic function $g(\cdot, p_0) + \log |\psi|$ attains a local maximum at an interior point p_0, so the maximum principle implies that

$$g(\cdot, p_0) = -\log |\psi|.$$

This implies that $|\psi| = e^{-g(\cdot, p_0)}$ has a unique zero at p_0.

This is true for any $p_0 \in M \setminus \{q_0\}$, so the injectivity of φ follows. As mentioned, this finishes the proof based on the Riemann mapping theorem. □

Acknowledgment

I thank Chen-Yu Chi for sharing his notes and experience on teaching complex analysis. I thank Jacob Bernstein, Gábor Székelyhidi, Terry Tao, and Chung-Jun Tsai, whose publicly available teaching materials on complex analysis have been helpful. I thank the students in GU4065, 2025, for their thoughtful questions and feedback, which have helped improve these notes.

In preparation of these notes, I have freely used several books and notes from people listed above and those in the reference below. Especially, I do not claim originality on most of the content.

References

[Ahl78] L. V. Ahlfors, *Complex analysis*, 3rd ed., International Series in Pure and Applied Mathematics, McGraw-Hill Book Co., New York, 1978. An introduction to the theory of analytic functions of one complex variable. MR510197

[Gam01] T. W. Gamelin, *Complex analysis*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 2001. MR1830078

[Mil06] John Milnor, *Dynamics in one complex variable*, 3rd ed., Annals of Mathematics Studies, vol. 160, Princeton University Press, Princeton, NJ, 2006. MR2193309

[Rud87] W. Rudin, *Real and complex analysis*, 3rd ed., McGraw-Hill Book Co., New York, 1987. MR924157

[SS03] E. M. Stein and Rami Shakarchi, *Complex analysis*, Princeton Lectures in Analysis, vol. 2, Princeton University Press, Princeton, NJ, 2003. MR1976398