POST GRADUATE DEGREE PROGRAMME (CBCS) IN

MATHEMATICS

SEMESTER IV

SELF LEARNING MATERIAL

PAPER : DSE 4.3
(Pure Stream)

Advanced Complex Analysis I

Directorate of Open and Distance Learning
University of Kalyani
Kalyani, Nadia
West Bengal, India

Content Writers

DSE 4.3 Advanced Complex Analysis I	Ms. Audrija Choudhury Assistant Professor Department of Mathematics Directorate of Open and Distance Learning University of Kalyani

April, 2024

Directorate of Open and Distance Learning, University of Kalyani

Published by the Directorate of Open and Distance Learning

University of Kalyani, 741235, West Bengal

Director’s Message

Satisfying the varied needs of distance learners, overcoming the obstacle of Distance and reaching the unreached students are the three fold functions catered by Open and Distance Learning (ODL) systems. The onus lies on writers, editors, production professionals and other personnel involved in the process to overcome the challenges inherent to curriculum design and production of relevant Self Learning Materials (SLMs). At the University of Kalyani a dedicated team under the able guidance of the Hon’ble Vice-Chancellorhas invested its best efforts, professionally and in keeping with the demands of Post Graduate CBCS Programmes in Distance Mode to devise a self-sufficient curriculum for each course offered by the Directorate of Open and Distance Learning (DODL), University of Kalyani.

Development of printed SLMs for students admitted to the DODL within a limited time to cater to the academic requirements of the Course as per standards set by Distance Education Bureau of the University Grants Commission, New Delhi, India under Open and Distance Mode UGC Regulations, 2020 had been our endeavor. We are happy to have achieved our goal.

Utmost care and precision have been ensured in the development of the SLMs, making them useful to the learners, besides avoiding errors as far as practicable. Further suggestions from the stakeholders in this would be welcome.

During the production-process of the SLMs, the team continuously received positive stimulations and feedback from Professor **(Dr.) Amalendu Bhunia, Hon’ble Vice-Chancellor, University of Kalyani**, who kindly accorded directions, encouragements and suggestions, offered constructive criticism to develop it with in proper requirements. We gracefully, acknowledge his inspiration and guidance.

Sincere gratitude is due to the respective chairpersons as well as each and every member of PGBOS (DODL), University of Kalyani. Heartfelt thanks are also due to the Course Writers-faculty members at the DODL, subject-experts serving at University Post Graduate departments and also to the authors and academicians whose academic contributions have enriched the SLMs. We humbly acknowledge their valuable academic contributions. I would especially like to convey gratitude to all other University dignitaries and personnel involved either at the conceptual or operational level of the DODL of University of Kalyani.

Their persistent and coordinated efforts have resulted in the compilation of comprehensive, learner-friendly, flexible texts that meet the curriculum requirements of the Post Graduate Programme through Distance Mode.

All the Self Learning Materials are self written and collected from e-book, journals and websites.

Director

Directorate of Open and Distance Learning

University of Kalyani

Post Graduate Board of Studies

Department of Mathematics
Directorate of Open and Distance Learning
University of Kalyani

Sl No.	Name & Designation	Role
1	Dr. Samares Pal, Professor & Head Department of Mathematics, University of Kalyani	Chairperson
2	Dr. Pulak Sahoo, Professor Department of Mathematics, University of Kalyani	Member
3	Dr. Sahidul Islam, Associate Professor Department of Mathematics, University of Kalyani	Member
4	Dr. Sushanta Kumar Mohanta, Professor Department of Mathematics, West Bengal State University	External Nominated Member
5	Ms. Audrija Choudhury, Assistant Professor Department of Mathematics Directorate of Open and Distance Learning University of Kalyani	Member
6	Director Directorate of Open and Distance Learning University of Kalyani	Convener

Discipline Specific Elective Paper

DSE 4.3
Marks : 100 (SEE : 80; IA : 20)

Advanced Complex Analysis I
(Pure Stream)

Syllabus

- **Unit 1:** The functions- $M(r)$ and $A(r)$. Hadamard theorem on the growth of $\log M(r)$
- **Unit 2:** Schwarz inequality, Borel-Caratheodory inequality, Open mapping theorem.
- **Unit 3:** Dirichlet series, abscissa of convergence and abscissa of absolute convergence, their representations in terms of the coefficients of the Dirichlet series.
- **Unit 4:** The Riemann Zeta function, the product development and the zeros of the zeta functions.
- **Unit 5:** Entire functions, growth of an entire function, order and type and their representations in terms of the Taylor coefficients.
- **Unit 6:** Distribution of zeros of entire functions
- **Unit 7:** The exponent of convergence of zeros.
- **Unit 8:** Weierstrass factorization theorem.
- **Unit 9:** Canonical product, Borel's first theorem. Borel's second theorem (statement only),
- **Unit 10:** Hadamard's factorization theorem, Schottky's theorem (no proof), Picard's first theorem.
- **Unit 11:** Multiple-valued functions
- **Unit 12:** Riemann surface for the functions $\sqrt{z}$, $\log z$.
- **Unit 13:** Analytic continuation, uniqueness
- **Unit 14:** Continuation by the method of power series
- **Unit 15:** Continuation by the method of natural boundary. Existence of singularity on the circle of convergence.
- **Unit 16:** Functions element, germ and complete analytic functions. Monodromy theorem.
- **Unit 17:** Conformal transformations, Riemann's theorems for circle
- **Unit 18:** Schwarz principle of symmetry, Schwarz-Christoffel formula (statement only) with applications.

- **Unit 19:** Univalent functions, general theorems
- **Unit 20:** Sequence of univalent functions, sufficient conditions for univalence.

Contents

Unit 1

Course Structure

- The functions- $M(r)$ and $A(r)$.
- Hadamard theorem on the growth of $\log M(r)$

1.1 Introduction

We have read about the Maximum and Minimum modulus theorems for a non-constant analytic functions on bounded set G. We mainly introduce two new terms, viz., $M(r)$ and $A(r)$, and derive their properties. The main motive is to study the growth of an analytic function f. f, being a complex function, is not comparable with the real functions when we come to measure their orders of growth. So, to be able to measure the order of growth of such functions, we need to define real function and find the desired results with respect to them.

Objectives

After reading this unit, you will be able to

- define $M(r)$ and $A(r)$ of an analytic function
- deduce the properties of them with the help of the maximum modulus theorem
- learn preliminary definitions of convex functions
- deduce the Hadamard's three-circles theorem

1.2 The functions $M(r)$ and $A(r)$

Let us first state the maximum modulus theorem and then we will define the new terms $M(r)$ and $A(r)$ and derive its properties as a consequence of the maximum modulus theorem.

Theorem 1.2.1. If a function f is analytic in a bounded region G, and continuous on $\overline{G}$ and $M = \max\{|f(z)| : z \in \partial G\}$, where ∂G is the boundary of G, then $|f(z)| < M$ in G, unless f is a constant function.

Example 1.2.1. Consider the function $f(z) = z^2$ defined on the closed disc $D = \{z : |z - 1 - i| \leq 1\}$. Let us show that the maximum value of $|f(z)|$ is attained at $z = (1 + 1/\sqrt{2})(1 + i)$. To do this, set

$$z = 1 + i + \mathrm{e}^{i\theta} = (1 + \cos\theta) + i(1 + \sin\theta),\ \theta \in [0, 2\pi).$$

Then $|f(z)| = 3 + 2(\cos\theta + \sin\theta)$. It follows that the maximum value of $|f(z)|$ is attained at $\theta = \pi/4$ and the maximum value is $3 + 2\sqrt{2}$. The maximum value is attained at $z = 1 + i + \mathrm{e}^{i\frac{\pi}{4}}$.

Corollary 1.2.1. Suppose that f is analytic in a bounded region G and continuous on $\overline{G}$. Then, each of $\mathrm{Re}f(z)$, $-\mathrm{Re}f(z)$, $\mathrm{Im}f(z)$ and $-\mathrm{Im}f(z)$ attains its maximum at some point on the boundary ∂G of G.

Proof. Let $u(x, y) = \mathrm{Re}\ f(z)$ and $g(z) = \mathrm{e}^{f(z)}$. By the Maximum Modulus theorem, $|g(z)| = \mathrm{e}^{u(x,y)}$ cannot assume the maximum value in G. Since e^u is maximized when u is maximized, obtain that $u(x, y)$ cannot attain its maximum value in G. Similarly, the other cases can be proved. □

The minimum modulus theorem comes as a direct corollary of the above theorem which is

Theorem 1.2.2. Let f be a non-constant analytic function in a bounded region G and continuous on $\overline{G}$. If $f(z) \neq 0$ inside ∂G, then $|f(z)|$ must attain its minimum value on ∂G.

Example 1.2.2. Suppose that f and g are analytic on the closed unit disc $|z| \leq 1$ such that

1. $|f(z)| \leq M$ for all $|z| \leq 1$;
2. $f(z) = z^n g(z)$ for all $|z| \leq 1/3$ and for some $n \in \mathbb{N}$.

We wish to use the Maximum modulus theorem to find the maximum value of $|f(z)|$ on $|z| \leq 1/3$. To do this, we proceed as follows. On $|z| = 1$, we have

$$m \geq |f(z)| = |z^n g(z)| = |g(z)|$$

and so, $|g(z)| \leq M$ for $|z| \leq 1$. Now, for $|z| \leq 1/3$, we have,

$$|f(z)| = |z^n g(z)| = |z^n||g(z)| = 3^{-n}|g(z)| \leq 3^{-n}M.$$

Thus, $|f(z)| \leq 3^{-n}M$ for all $|z| \leq 1/3$.

The hypothesis that G is bounded can't be dropped however as we see in the following example.

Example 1.2.3. Define $f(z) = \mathrm{e}^{-iz}$ on $G = \{z :\ \mathrm{Im}\ z > 0\}$. Then $|f(z)| = 1$ on the boundary $\partial G = \{z :\ \mathrm{Im}\ z = 0\}$, that is, the real axis. But, for $z = x + iy \in G$, we have,

$$|f(x + iy)| = \mathrm{e}^y \to \infty \text{ as } y \to +\infty;$$

that is, f itself is not bounded. And the Maximum modulus theorem fails.

We will now define the terms $M(r)$ and $A(r)$ related to an analytic function f as follows

Definition 1.2.1. Let f be a non-constant analytic function defined in $|z| \leq R$. Then, for $0 \leq r < R$, we define

1. $M(r) = \max\{|f(z)| :\ |z| = r\}$; and
2. $A(r) = \max\{\mathrm{Re}\ f(z) :\ |z| = r\}$.

Theorem 1.2.3. Let f be a non-constant analytic function defined in $|z| \leq R$. Then, $0 \leq r < R$

1. $M(r)$ is a strictly increasing function of r;
2. $A(r)$ is a strictly increasing function of r.

Proof. 1. Let $0 \leq r_1 < r_2 < R$. Since f is analytic in $|z| \leq r_2$, the maximum value of $|f(z)|$ for $|z| \leq r_2$ is attained on $|z| = r_2$. Let z_2 be a point on $|z| = r_2$ such that $|f(z_2)| = M(r_2)$. Similarly, the maximum value of $|f(z)|$ for $|z| \leq r_1$ is attained on $|z| = r_1$. Let z_1 be a point on $|z| = r_1$ such that $|f(z_1)| = M(r_1)$. Since z_1 is an interior point of the closed region $|z| \leq r_2$, hence by maximum modulus theorem, $|f(z_1)| < M(r_2)$, that is, $M(r_1) < M(r_1)$. This completes the proof of the first part of the theorem.

2. Let $0 \leq r_1 < r_2 < R$. Since f is analytic in $|z| \leq r_2$, the maximum value of Re $f(z)$ for $|z| \leq r_2$ is attained on $|z| = r_2$. Let z_2 be a point on $|z| = r_2$ such that Re $f(z_2) = A(r_2)$. Similarly, the maximum value of Re $f(z)$ for $|z| \leq r_1$ is attained on $|z| = r_1$. Let z_1 be a point on $|z| = r_1$ such that Re $f(z_1) = A(r_1)$. Since z_1 is an interior point of the closed region $|z| \leq r_2$, hence by corollary 1.2.1, Re $f(z_1) < A(r_2)$, that is, $A(r_1) < A(r_2)$.

□

1.3 Hadamard's theorem on the growth of $\log M(r)$

Definition 1.3.1. Let $[a, b]$ be an interval in $\mathbb{R}$. A function $f : [a, b] \to \mathbb{R}$ is said to be **convex** if for any two points $x_1, x_2 \in [a, b]$,

$$f(tx_2 + (1-t)x_1) \leq f(x_2) + (1-t)f(x_1) \quad \text{for} \quad 0 \leq t \leq 1.$$

Geometrically, $f(x)$ is said to be convex downwards, or simply convex in $[a, b]$ if the curve $y = f(x)$ between any two points x_1 and x_2 in $[a, b]$ always lies below the chord joining the points $(x_1, f(x_1))$ and $(x_2, f(x_2))$.

1.3.1 Analytical condition for convexity

Consider the figure 1.1.

Let $y = f(x)$ be the curve and $(x_1, f(x_1))$, $(x_2, f(x_2))$ be two points as shown in fig. 1.1. Let $x = tx_1 + (1-t)x_2$ be any point between x_1 and x_2 and $0 \leq t \leq 1$. Then, $x_1 \leq x \leq x_2$ and the equation of the chord joining $(x_1, f(x_1))$ and $(x_2, f(x_2))$ is

$$\begin{aligned} y - f(x_1) &= \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \\ \text{or,} \qquad y &= f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1). \end{aligned}$$

Let the coordinates of any point on the chord joining $(x_1, f(x_1))$ and $(x_2, f(x_2))$ be (x, y). According to the definition of convexity, $f(x)$ will be convex if and only if $f(x) \leq y$, that is,

$$f(x) \leq f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1)$$

that is, if and only if,

$$f(x) \leq \frac{x_2 - x}{x_2 - x_1} f(x_1) + \frac{x - x_1}{x_2 - x_1} f(x_2) \tag{1.3.1}$$

holds for all $x_1 \leq x \leq x_2$.

Let us check through some quick results on convex functions that we will need in the sequel.

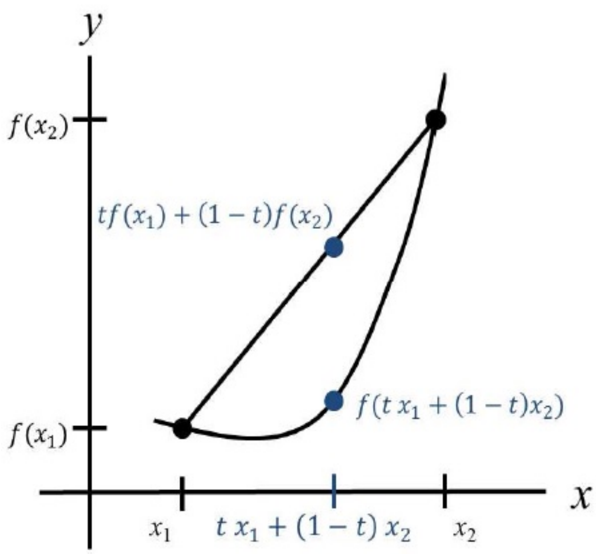


Figure 1.1

Result 1.3.1. 1. A differentiable function f on $[a, b]$ is convex if and only if f' is increasing.

2. A sufficient condition for f to be convex is that, $f''(x) \geq 0$.

3. If $f : (a, b) \to \mathbb{R}$ is convex, then it is continuous there.

We now state the Hadamard's three circles theorem.

Theorem 1.3.1. (Hadamard's three-circles theorem) Let f be analytic on the closed annulus $0 < r_1 \leq |z| \leq r_3$ (see fig. 1.2). If $r_1 < r_2 < r_3$, then

$$M(r_2)^{\log\left(\frac{r_3}{r_1}\right)} \leq M(r_1)^{\log\left(\frac{r_3}{r_2}\right)} M(r_3)^{\log\left(\frac{r_2}{r_1}\right)}.$$

Proof. Let $\phi(z) = z^{\lambda} f(z)$, where λ is a real constant to be chosen later. If λ is not an integer, $\phi(z)$ is multi-valued in $r_1 \leq |z| \leq r_3$. So we cut the annulus along the negative part of the real axis obtaining a simply connected region G in which the principal branch of ϕ is analytic.

The maximum modulus of this branch of ϕ in G, that is, the cut-annulus is obtained on the boundary of G. Since λ is real, all the branches of ϕ have the same modulus. By considering another branch of ϕ which is analytic in another cut-annulus obtained by using a different cut, it is clear that the principal branch of ϕ must attain its maximum modulus on at least one of the bounding circles of the annulus. Thus, $|\phi(z)| \leq \max\{r_1^{\lambda} M(r_1), r_3^{\lambda} M(r_3)\}$. Hence, on $|z| = r_2$, we have,

$$r_2^{\lambda} M(r_2) \leq \max\{r_1^{\lambda} M(r_1), r_3^{\lambda} M(r_3)\}. \tag{1.3.2}$$

We now choose λ such that

$$r_1^{\lambda} M(r_1) = r_3^{\lambda} M(r_3);$$

that is,

$$\lambda = -\frac{\log\left(\frac{M(r_3)}{M(r_1)}\right)}{\log\left(\frac{r_3}{r_1}\right)}.$$

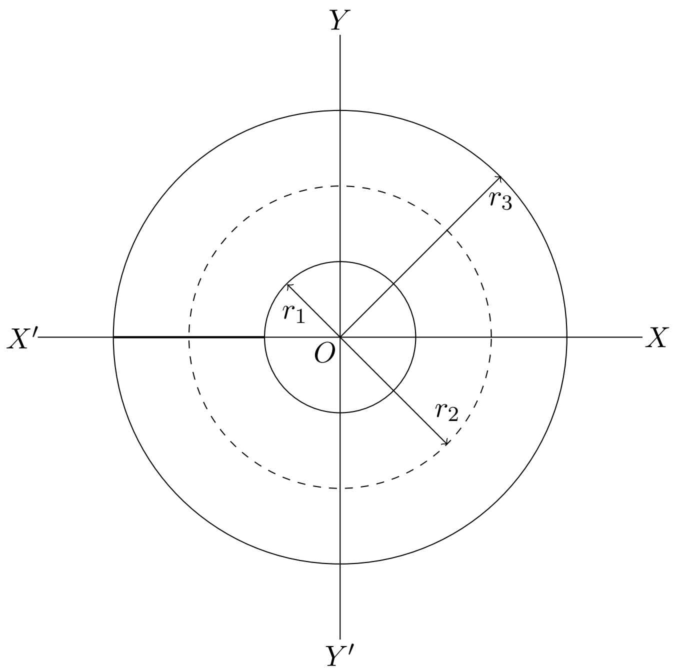


Figure 1.2: Hadamard's Three-Circles Theorem

With this λ, we get from (1.3.2), and the fact that $a^{\log b} = b^{\log a}$ holds for all positive real numbers a and b, we get

$$
\begin{aligned}
& M(r_2) \le \left(\frac{r_1}{r_2}\right)^{\lambda} M(r_1) \\
\Rightarrow \quad & M(r_2)^{\log\left(\frac{r_3}{r_1}\right)} \le \left(\frac{r_1}{r_2}\right)^{\log\left(\frac{M(r_1)}{M(r_3)}\right)} \cdot M(r_1)^{\log\left(\frac{r_3}{r_1}\right)} \\
\Rightarrow \quad & M(r_2)^{\log\left(\frac{r_3}{r_1}\right)} \le \left(\frac{M(r_1)}{M(r_3)}\right)^{\log\left(\frac{r_1}{r_2}\right)} \cdot M(r_1)^{\log\left(\frac{r_3}{r_1}\right)} \\
\Rightarrow \quad & M(r_2)^{\log\left(\frac{r_3}{r_1}\right)} \le M(r_1)^{\log\left(\frac{r_3}{r_2}\right)} \cdot M(r_3)^{\log\left(\frac{r_2}{r_1}\right)}.
\end{aligned}
$$

□

Note 1.3.1. The equality in the above theorem can be achieved when $\phi(z)$ is constant, that is, $f(z)$ is of the form $f(z) = kz^{\lambda}$, for some real λ, k being a constant.

Theorem 1.3.2. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be analytic in $|z| \le r$. Then

$$|a_n| r^n \le \max\{4A(r), 0\} - 2\text{Re}\, f(0), \ \forall n > 0.$$

Proof. Let $z = r\,\mathrm{e}^{i\theta}$, $f(z) = \sum_{n=0}^{\infty} a_n z^n = u(r,\theta) + iv(r,\theta)$ and $a_n = \alpha_n + i\beta_n$. Thus,

$$\begin{aligned} u(r,\theta) + iv(r,\theta) &= \sum_{n=0}^{\infty} (\alpha_n + i\beta_n) r^n\,\mathrm{e}^{in\theta} \\ &= \sum_{n=0}^{\infty} (\alpha_n + i\beta_n) r^n (\cos n\theta + i \sin n\theta) \\ &= \sum_{n=0}^{\infty} r^n \{(\alpha_n \cos n\theta - \beta_n \sin n\theta) + i(\alpha_n \sin n\theta + \beta_n \cos n\theta)\}. \end{aligned}$$

Equating real parts, we get,

$$u(r,\theta) = \sum_{n=0}^{\infty} r^n (\alpha_n \cos n\theta - \beta_n \sin n\theta). \tag{1.3.3}$$

The series (1.3.3) converges uniformly with respect to θ. Hence we may multiply it by $\sin n\theta$ or $\cos n\theta$ and integrate it term by term. Now,

$$u(r,\theta) = \alpha_0 + (\alpha_1 \cos\theta - \beta_1 \sin\theta) r + \cdots + (\alpha_n \cos n\theta - \beta_n \sin n\theta) r^n + \cdots$$

Hence,

$$\int_0^{2\pi} u(r,\theta) d\theta = \alpha_0 2\pi \Rightarrow \alpha_0 = \frac{1}{2\pi} \int_0^{2\pi} u(r,\theta) d\theta.$$

Also,

$$u(r,\theta)\cos n\theta = \alpha_0 \cos n\theta + (\alpha_1 \cos\theta \cos n\theta - \beta_1 \sin\theta \cos n\theta) r + \cdots + (\alpha_n \cos^2 n\theta - \beta_n \sin n\theta \cos n\theta) r^n + \cdots$$

So,

$$\int_0^{2\pi} u(r,\theta) \cos n\theta d\theta = \alpha_n r^n \int_0^{2\pi} \cos^2 n\theta d\theta = \pi \alpha_n r^n.$$

Hence, for $n > 0$,

$$\alpha_n r^n = \frac{1}{\pi} \int_0^{2\pi} u(r,\theta) \cos n\theta d\theta.$$

Similarly, multiplying (1.3.3) by $\sin n\theta$ and integrating term by term from 0 to 2π, we get,

$$-\beta_n r^n = \frac{1}{\pi} \int_0^{2\pi} u(r,\theta) \sin n\theta d\theta, \quad \text{for} \quad n > 0.$$

Hence,

$$\begin{aligned} a_n r^n = (\alpha_n + i\beta_n) r^n &= \frac{1}{\pi} \int_0^{2\pi} u(r,\theta) \cos n\theta d\theta - \frac{1}{\pi} \int_0^{2\pi} u(r,\theta) \sin n\theta d\theta \\ &= \frac{1}{\pi} \int_0^{2\pi} u(r,\theta)\,\mathrm{e}^{-in\theta}\,d\theta, \quad n > 0. \end{aligned}$$

Thus,

$$|a_n| r^n \leq \frac{1}{\pi} \int_0^{2\pi} |u(r,\theta)| d\theta, \quad n > 0.$$

Hence,

$$|a_n| r^n + 2\alpha_0 = \frac{1}{\pi} \int_0^{2\pi} \{|u(r,\theta)| + u(r,\theta)\} d\theta. \tag{1.3.4}$$

Now, $|u(r,\theta)| + u(r,\theta) = 0$ when $u(r,\theta) < 0$. Hence if $A(r) < 0$, the right hand side of (1.3.4) is 0. Again, if $A(r) \geq 0$, the right hand side of (1.3.4) does not exceed

$$\frac{1}{\pi}\int_0^{2\pi} 2A(r)d\theta = 4A(r).$$

Thus,

$$\begin{aligned} & |a_n|r^n + 2\alpha_0 \leq \max\{4A(r), 0\} \\ \Rightarrow\quad & |a_n|r^n + 2\text{Re } f(0) \leq \max\{4A(r), 0\} \\ \Rightarrow\quad & |a_n|r^n \leq \max\{4A(r), 0\} - 2\text{Re } f(0). \end{aligned}$$

□

Few Probable Questions

---

1. State the Maximum modulus theorem. Show that $M(r)$ is an increasing function of r.
2. Show that for any non-constant analytic function f defined on a bounded region G and continuous on $\overline{G}$, Re $f(z)$ attains its maximum at some point on the boundary ∂G of G. Show that $A(r)$ is an increasing function of r.
3. State and prove the Hadamard's three-circles theorem.

Unit 2

Course Structure

- Schwarz Lemma, Open mapping theorem.
- Borel-Caratheodory inequality

2.1 Introduction

Let us denote $\mathbb{D} = \{z \in \mathbb{C} : \ |z| < 1\}$. In this unit, we start with a simple but one of the classical theorems in complex analysis, named Schwarz' Lemma, which states that if f is analytic and satisfies $|f(z)| < 1$ in $\mathbb{D}$ and $f(0) = 0$, then $|f(z)| \leq |z|$ for each $z \in \mathbb{D}$ with equality sign if and only if f has the form $f(z) = \mathrm{e}^{i\alpha} z$, for some $\alpha \in \mathbb{R}$. Furthermore, $|f'(0)| \leq 1$ with the equality if and only if f has the form as stated previously. This result has important role in the proof of Riemann mapping theorem.

This unit also deals with deducing the Open mapping theorem and the Borel-Caratheodory theorem.

Objectives

After reading this unit, you will be able to

- deduce Schwarz lemma and its various variants; also applying it in various problems
- deduce the Borel Caratheodory inequality from Schwarz lemma and discuss some of its consequences
- derive the Open mapping theorem

2.2 Schwarz Lemma

Let us now begin with a sharp version of the classical Schwarz lemma.

Theorem 2.2.1. (Schwarz Lemma) Let $f : \mathbb{D} \to \overline{\mathbb{D}}$ be analytic having a zero of order n at the origin. Then

1. $|f(z)| \leq |z|^n$ for all $z \in \mathbb{D}$,
2. $|f^{(n)}(0)| \leq n!$

and the equality holds either in 1 for some point $0 \neq z_0 \in \mathbb{D}$ or in 2 occurs if and only if $f(z) = \epsilon z^n$ with $|\epsilon| = 1$.

Proof. Let $f : \mathbb{D} \to \overline{\mathbb{D}}$ be analytic on $\mathbb{D}$ and has nth order zero at the origin. Then we have,

$$f(0) = 0 = f'(0) = \cdots = f^{(n-1)}(0) \quad \text{and} \quad f^{(n)}(0) \neq 0.$$

So we can write

$$f(z) = \sum_{k=n}^{\infty} a_k z^k = z^n g(z), \quad \text{for} \quad z \in \mathbb{D},$$

where

$$a_k = \frac{f^{(k)}(0)}{k!} \quad \text{and} \quad g(z) = \sum_{k=n}^{\infty} a_k z^{k-n}.$$

The function $g(z) = f(z)/z^n$ has a removable singularity at the origin so that if

$$\begin{aligned} g(z) &= z^{-n} f(z) \quad \text{for } z \in \mathbb{D} \setminus \{0\} \\ &= a_n \qquad \text{for } z = 0. \end{aligned}$$

then g is analytic in $\mathbb{D} \setminus \{0\}$ and continuous in $\mathbb{D}$. Since we will have by Cauchy's theorem for a disc, $\int_C g(z)dz = 0$ for all closed contours C inside $\mathbb{D}$ so by Morera's theorem, g is analytic on $\mathbb{D}$.

We claim that $|g(z)| \leq 1$ for all $z \in \mathbb{D}$. Now, for $0 < r < 1$, g is analytic on the bounded domain $\mathbb{D}_r = \{z : \ |z| < r\}$ and g is continuous on the closure of $\mathbb{D}_r$. Thus, the maximum modulus theorem is applicable here. As $|f(z)| \leq 1$ for all $z \in \mathbb{D}$, it follows that for $|z| = r$,

$$|g(z)| = \frac{|f(z)|}{|z|^n} \leq \frac{1}{r^n}.$$

By Maximum modulus theorem, $|g(z)| \leq r^{-n}$, for all z with $|z| \leq r$. Since r is arbitrary, by letting $r \to 1$, we find that $|g(z)| \leq 1$, that is,

$$|g(z)| \leq 1 \quad \text{for all} \quad z \in \mathbb{D} \tag{2.2.1}$$

and this implies that

$$|f(z)| \leq |z|^n \quad \text{for all} \quad z \in \mathbb{D}.$$

Equality in 1 holds for some point z_0 in $\mathbb{D} \setminus \{0\}$ implies that $|g(z_0)| = 1$. It follows that g achieves its maximum modulus at an interior point z_0. Consequently, by the Maximum modulus theorem, g must reduce to a constant, say ϵ. Then, $f(z) = \epsilon z^n$, where $|\epsilon| = 1$.

Also, note that $|g(z)| \leq 1$ throughout the disc $\mathbb{D}$. Since $|a_n| = |g(0)|$, we get by equation (2.2.1), $|g(0)| \leq 1$ so, we get $|f^{(n)}(0)|/n! \leq 1$ and hence, 2 follows.

Again, if $|f^{(n)}(0)| = n!$ then $|g(0)| = 1$ showing that g achieves its maximum modulus 1 at some interior point $'0'$. So, g is a constant function of absolute value 1 and as before, it means that $f(z) = \epsilon z^n$ with $|\epsilon| = 1$. □

Remark 2.2.1. Note that the case $n = 1$ of the previous theorem is the original Schwarz lemma state in the beginning of this unit.

For example, if f is an analytic function over $\mathbb{D}$ with $|f(z)| \leq 1$ and $f(0) = 0$, then what kind of function is f if $f(1/2) = 1/2$? It must be none other than the identity function since the equality in the preceding theorem holds with $n = 1$ and $z = 1/2 \in \mathbb{D}$.

Corollary 2.2.1. If f is analytic and satisfies $|f(z)| \leq M$ in $B(a; R)$ and $f(a) = 0$, then

1. $|f(z)| \leq M|z-a|/R$ for every $z \in B(a;R)$,
2. $|f'(a)| \leq M/R$

with the equality sign if and only if f has the form $f(z) = M\epsilon(z-a)/R$, for some constant ϵ with $|\epsilon| \leq 1$.

Proof. Use Schwarz lemma with $g(z) = f(Rz+a)/M$, $|z| < 1$. ☐

Another generalisation of the above theorem is as follows

Corollary 2.2.2. If f is analytic and satisfies $|f(z)| \leq M$ in $B(a;R)$ and a is a zero of f of order n. Then

1. $|f(z)| \leq M|z-a|^n/R^n$ for every $z \in B(a;R)$,
2. $|f^{(n)}(a)| \leq M/R^n$

with the equality sign if and only if f has the form $f(z) = M\epsilon(z-a)^n/R^n$, for some constant ϵ with $|\epsilon| \leq 1$.

Proof. Prove the corollary independently without using the Schwarz lemma. ☐

Does the Schwarz lemma hold for real-valued functions? Consider the function

$$u(x) = \frac{2x}{x^2+1}.$$

Then u is infinitely differentiable on $\mathbb{R}$. In particular, $u'(x)$ is continuous on $[-1,1]$, $u(0) = 0$ and $|u(x)| \leq 1$. But $|u(x)| > |x|$ for $0 < |x| < 1$.

Example 2.2.1. Let $\omega = e^{2\pi i/n}$ be the nth root of unity, where $n \in \mathbb{N}$ is fixed. Suppose that $f : \mathbb{D} \to \mathbb{D}$ is analytic such that $f(0) = 0$. We wish to apply Schwarz lemma to show that

$$|f(z) + f(\omega z) + f(\omega^2 z) + \cdots + f(\omega^{n-1}z)| \leq n|z|^n \tag{2.2.2}$$

and the equality for some point $0 \neq z_0 \in \mathbb{D}$ occurs if and only if $f(z) = \epsilon z^n$ with $|\epsilon| = 1$. To do this, we define $F : \mathbb{D} \to \mathbb{D}$ by

$$F(z) = \frac{1}{n}\sum_{k=0}^{n-1} f(\omega^k z).$$

Clearly F is analytic on $\mathbb{D}$, $F(0) = 0$ and, for $1 \leq m \leq n-1$,

$$F^{(m)}(z) = \frac{1}{n}\sum_{k=0}^{n-1}(\omega^k)^m f^{(m)}(\omega^k z)$$

so that (as $\omega^n = 1$)

$$F^{(m)}(0) = \frac{1}{n}\sum_{k=0}^{n-1}(\omega^k)^m f^{(m)}(0) = \frac{f^{(m)}(0)}{n}\left(\frac{1-(\omega^m)^n}{1-\omega^m}\right) = 0.$$

By Schwarz lemma, it follows that $|F(z)| \leq |z|^n$ for all $z \in \mathbb{D}$ which is the same as (2.2.2). The equality in this inequality for some point $z_0 \neq 0$ occurs if and only if $F(z) = \epsilon z^n$ with $|\epsilon| = 1$, or equivalently,

$$\sum_{k=0}^{n-1}[f(\omega^k z) - \epsilon z^n] = 0. \tag{2.2.3}$$

We claim that the above equation implies that $f(z) = \epsilon z^n$. If we let $f(z) = \sum_{m=1}^{\infty} a_m z^m$, then (2.2.3) becomes

$$\sum_{m=1}^{\infty} a_m \left(\sum_{k=0}^{n-1} \omega^{km} \right) z^m = n\epsilon z^n.$$

In view of the identity

$$\begin{aligned}\sum_{k=0}^{n-1} \omega^{km} &= n \text{ if } m \text{ is a multiple of } n \\ &= 0 \text{ otherwise}\end{aligned}$$

the last equation implies that $a_n = \epsilon$ and $a_{2n} = a_{3n} = \cdots = 0$. On the other hand, as $|f(z)| \leq 1$ on $\mathbb{D}$ and $|a_n| = 1$, we have

$$\lim_{r\to 1^-} \frac{1}{2\pi} \int_0^{2\pi} |f(r\,e^{i\theta})|^2 d\theta = \sum_{m=1}^{\infty} |a_m|^2 \leq 1$$

which shows that all the Taylor's coefficients of f (except a_n) must vanish and so, $f(z) = e^{i\theta}\, z^n$.

2.2.1 Borel-Caratheodory theorem

Borel-Caratheodory theorem is an important theorem that establishes the relationship between $M(r)$ and $A(r)$. We deduce this with the help of Schwarz theorem.

Theorem 2.2.2. (Borel Caratheodory theorem) Let f be analytic on $D : |z| \leq R$ and $M(r)$ and $A(r)$ are as we defined in the previous unit. Then for $0 < r < R$,

$$M(r) \leq \frac{2r}{R-r} A(R) + \frac{R+r}{R-r} |f(0)|.$$

Proof. We consider the following cases.

Case I. When $f(z) = \text{constant} = a + ib$, where a and b are real constants. Then $M(r) = \sqrt{a^2+b^2}$, and $|f(0)| = \sqrt{a^2+b^2}$, and $A(R) = a$. Now,

$$\begin{aligned}\frac{2r}{R-r} A(R) + \frac{R+r}{R-r} |f(0)| - M(r) &= \frac{2r}{R-r} a + \left(\frac{R+r}{R-r} - 1 \right) \sqrt{a^2+b^2} \\ &= \frac{2r}{R-r} (a + \sqrt{a^2+b^2}) \geq 0.\end{aligned}$$

Hence,

$$M(r) \leq \frac{2r}{R-r} A(R) + \frac{R+r}{R-r} |f(0)|.$$

Case II. When $f(z) \neq \text{constant}$ and $f(0) = 0$. Then $A(R) > A(0) = 0$, since $A(r)$ is an increasing function of r. Let,

$$g(z) = \frac{f(z)}{2A(R) - f(z)}. \tag{2.2.4}$$

$2A(R) - f(z) \neq 0$ for all $z \in D$, since the real part of $2A(R) - f(z)$ does not vanish in D and $g(0) = 0$. Let $f(z) = u + iv$. Then

$$|g(z)|^2 = \frac{u^2+v^2}{(2A(R)-u)^2+v^2} \leq 1, \quad z \in D,$$

since $2A(R) - u \geq u$. Hence by Schwarz lemma,

$$|g(z)| \leq \frac{r}{R} \quad \text{for} \quad |z| = r < R.$$

From (2.2.4), we get,

$$f(z) = 2A(R)g(z) - f(z)g(z)$$
$$\text{or} \quad , \quad f(z)(1 + g(z)) = 2A(R)g(z).$$

Hence,

$$|f(z)| = \left|\frac{2A(R)g(z)}{1 + g(z)}\right| \leq \frac{2A(R)|g(z)|}{1 - |g(z)|} \leq \frac{2A(R)\frac{r}{R}}{1 - \frac{r}{R}} = \frac{2r}{R - r}A(R),$$

for $|z| = r < R$. Thus, $M(r) \leq \frac{2r}{R-r}A(R)$ and

$$M(r) \leq \frac{2r}{R - r}A(R) + \frac{R + r}{R - r}|f(0)|.$$

Case III. When $f(z) \neq$ constant and $f(0) \neq 0$. Let $h(z) = f(z) - f(0)$. Then $h(0) = 0$. So, by case II, we have

$$\max\{|h(z)| : \ |z| = r\} \leq \frac{2r}{R - r} \max\{\text{Re } h(z) : \ |z| = R\}. \tag{2.2.5}$$

Now,

$$\begin{aligned} \max\{|h(z)| : \ |z| = r\} &= \max\{|f(z) - f(0)| : \ |z| = r\} \\ &\geq \max\{|f(z)| : \ |z| = r\} - |f(0)| \\ &= M(r) - |f(0)|, \end{aligned}$$

and

$$\begin{aligned} \max\{\text{Re } h(z) : \ |z| = R\} &= \max\{\text{Re } (f(z) - f(0)) : \ |z| = R\} \\ &= \max\{(\text{Re } f(z) - \text{Re } f(0)) : \ |z| = R\} \\ &\leq \max\{\text{Re } f(z) : \ |z| = R\} + |f(0)| \\ &= A(R) + |f(0)|. \end{aligned}$$

Hence from (2.2.5), we get,

$$M(r) - |f(0)| \leq \frac{2r}{R - r}(A(R) + |f(0)|)$$
$$\text{or,} \quad M(r) \leq \frac{2r}{R - r}A(R) + \frac{R + r}{R - r}|f(0)|.$$

□

Corollary 2.2.3. If $A(R) \geq 0$, then

$$M(r) \leq \frac{2r}{R - r}A(R) + \frac{R + r}{R - r}|f(0)| \leq \frac{R + r}{R - r}(A(R) + |f(0)|)$$

$\left[\text{since } \frac{2r}{R-r} < \frac{r+r}{R-r} \text{ as } R + r > 2r\right]$.

Corollary 2.2.4. If $A(R) \geq 0$, then

$$\max\{|f^{(n)}(z)| : |z| = r\} \leq \frac{2^{n+2} \cdot n! R}{(R-r)^{n+1}}(A(R) + |f(0)|).$$

Proof. By Cauchy's Integral formula for derivatives, we have,

$$f^{(n)}(z) = \frac{n!}{2\pi i} \int_\gamma \frac{f(t)dt}{(t-z)^{n+1}} \tag{2.2.6}$$

where, $\gamma : |t - z| = \delta = (R - r)/2$.

On γ,

$$|t| = |t - z + z| \leq |t - z| + |z| = \frac{1}{2}(R - r) + r = \frac{1}{2}(R + r) < R,$$

which ensures that γ lies within $|z| = R$. By Borel Caratheodory theorem, we have,

$$\begin{aligned} \max|f(t)| &\leq \frac{R + \frac{1}{2}(R+r)}{R - \frac{1}{2}(R+r)}(A(R) + |f(0)|) \\ &= \frac{3R + r}{R - r}(A(R) + |f(0)|) \\ &< \frac{4R}{R - r}(A(R) + |f(0)|). \end{aligned}$$

Hence from (2.2.6),

$$\begin{aligned} |f^{(n)}(z)| &= \frac{n!}{2\pi}\left|\int_\gamma \frac{f(t)dt}{(t-z)^{n+1}}\right| \\ &\leq \frac{n!}{2\pi\delta^{n+1}} \frac{4R}{R-r}(A(R) + |f(0)|) \cdot 2\pi\delta \\ &= \frac{n!}{\delta^n} \cdot \frac{4R}{R-r}(A(R) + |f(0)|) \\ &= \frac{2^{n+2} \cdot n! R}{(R-r)^{n+1}}(A(R) + |f(0)|). \end{aligned}$$

Hence,

$$\max\{|f^{(n)}(z)| : |z| = r\} \leq \frac{2^{n+2} \cdot n! R}{(R-r)^{n+1}}(A(R) + |f(0)|).$$

□

Exercise 2.2.1. If f is an analytic function defined on $\mathbb{D}$ such that $|f(z)| < 1$ for all z in $\mathbb{D}$ and f fixes two distinct points of $\mathbb{D}$, then show that f is the identity function.

2.3 Open Mapping Theorem

In this section, we are interested in functions that send open sets to open sets. A function f defined on an open set U is said to be an **open mapping** if for every open subset V of U, the image $f(V)$ is open. Consider the mappings $f, g, h : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = x^2, \quad g(x) = \sin x, \quad h(x) = \frac{e^x + e^{-x}}{2},$$

respectively. Clearly,

$$f(\mathbb{R}) = [0, \infty), \quad g(\mathbb{R}) = (0, 1], \quad h(\mathbb{R}) = [1, \infty)$$

showing that each of f, g and h are not open mappings. Each of the above functions are infinitely differentiable non-constant real valued functions defined on the real line. Thus, the above examples show that the following theorem does not hold for real line $\mathbb{R}$. Let us now state the open mapping theorem.

Theorem 2.3.1. Let G be a region and suppose that f is a non-constant analytic function on G. Then for any open set U in G, $f(U)$ is open.

Proof. Let $U \subset G$ be open. To show that $f(U)$ is open, we show that for each $a \in U$, $\exists\ \delta > 0$ such that the open ball $B(f(a); \delta) \subset f(U)$. Let $\phi(z) = f(z) - f(a)$. Then a is a zero of ϕ. Since the zeros of a non-constant analytic functions are isolated points, so there exists an open ball $B(a; r)$ with $\overline{B}(a; r) \subset U$ such that $\phi(z) \neq 0$ in $0 < |z - a| < r$. In particular, $\phi(\alpha) \neq 0$ for $\alpha \in \partial B(a; \rho)$ where $\rho < r$.

Let

$$2\delta = \min\{|\phi(\alpha)| : \ \alpha \in \partial B(a; \rho)\}.$$

Then $\delta > 0$. Now, for any $w \in B(f(a); \delta)$, we have,

$$\begin{aligned}
|f(\alpha) - w| &\geq |f(\alpha) - f(a)| - |f(a) - w| \\
&= |\phi(\alpha)| - |f(a) - w| \\
&> 2\delta - \delta \\
&= \delta \\
&> |f(a) - w|,
\end{aligned}$$

for all $\alpha \in \partial B(a; \rho)$. This implies that

$$\min\{|f(\alpha) - w| : \ \alpha \in \partial B(a; \rho)\} > |f(a) - w|. \tag{2.3.1}$$

Let $F(z) = f(z) - w$. Then F has a zero in $B(a; \rho)$. For, if $F(z) \neq 0$ in $B(a; \rho)$, there exists a nbd $N(a)$ of a containing $\overline{B}(a; \rho)$ lying in G such that $F(z) \neq 0$ in $N(a)$. Then $1/F(z)$ will be analytic in $N(a)$ and

$$\left|\frac{1}{F(a)}\right| < \max\left\{\left|\frac{1}{F(\alpha)}\right| : \ \alpha \in \partial B(a; \rho)\right\} = \frac{1}{\min\{|F(\alpha)| : \ \alpha \in \partial B(a; \rho)\}}$$

that is,

$$\min\{|f(\alpha) - w| : \ \alpha \in \partial B(a; \rho)\} < |f(a) - w|,$$

which contradicts (2.3.1). Hence $\exists\ z_0 \in B(a; \rho)$ such that $f(z_0) = w$. Since w is an arbitrary point of $B(f(a); \delta)$, it follows that $B(f(a); \delta) \subset f(U)$, and hence the theorem. □

Few Probable Questions

---

1. State and prove Schwarz lemma.
2. State and prove Borel-Caratheodory theorem.
3. State and prove the Open Mapping theorem.

Unit 3

Course Structure

- Dirichlet series, abscissa of convergence and abscissa of absolute convergence,
- Their representations in terms of the coefficients of the Dirichlet series.

3.1 Introduction

In mathematics, a Dirichlet series is any series of the form

$$f(s) = \sum_{n=1}^{\infty} \frac{a_n}{n^s}$$

where s is a complex number, and a_n is a complex sequence. It is a special case of general Dirichlet series.

Dirichlet' s series were, as their name implies, first introduced into analysis by Dirichlet, primarily with a view to applications in the theory of numbers. A number of important theorems concerning them were proved by Dedekind, and incorporated by him in his later editions of Dirichlet's Vorlesungen uber Zahlentheorle. Dirichlet and Dedekind, however, considered only real values of the variable s. The first theorems involving complex values of s are due to Jensen, who determined the nature of the region of convergence of the general series; and the first attempt to construct a systematic theory of the function $f(s)$ was made by Cahent in a memoir which, although much of the analysis which it contains is open to serious criticism, has served and possibly just for that reason as the starting point of most of the later researches in the subject. We will however, not go into a very vigorous treatment of the subject. We will mainly concern ourselves with the preliminaries of Dirichlet series and gain some idea about their convergence.

Objectives

After reading this unit, you will be able to

- define general and ordinary Dirichlet's series and its examples
- deduce various conditions for convergence of Dirichlet's series
- define certain terms related to the convergence of Dirichlet's series and deduce certain properties

3.2 Dirichlet Series

We formally define the Dirichlet series as follows.

Definition 3.2.1. The series of the form

$$f(s) = \sum_{n=1}^{\infty} a_n \, \mathrm{e}^{-\lambda_n s} \tag{3.2.1}$$

where, $\{\lambda_n\}$ is an increasing sequence of real numbers whose limit is infinity, and $s = \sigma + it$ is a complex variable, whose real and imaginary parts are σ and t respectively. Such a series is called a Dirichlet's series of type λ_n. If $\lambda_n = n$, then (3.2.1) is a power series in e^{-n}. If $\lambda_n = \log n$, then (3.2.1) becomes

$$f(s) = \sum_{n=1}^{\infty} a_n n^{-s} \tag{3.2.2}$$

is called an ordinary Dirichlet's series. In this unit, we will mainly deal with the ordinary Dirichlet's series.

It is clear that all but a finite number of the numbers λ_n must be positive. It is often convenient to suppose that they are all positive, or at any rate $\lambda_1 \geq 0$. Sometimes, an additional assumption is needed, such as the Bohr condition, namely $\lambda_{n+1} - \lambda_n \geq c/n$ for some $c > 0$.

We will look into certain examples of Dirichlet's series now.

Example 3.2.1. A very important example is the Riemann zeta function which is

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

For $t = 0$, that is, $s = \sigma \in \mathbb{R}$, it is proved from elementary calculus, that $\zeta(\sigma)$ diverges for $\sigma = 1$ and is absolutely convergent for $\sigma > 1$. This is called the "p-test", where $p = \sigma$. We will learn more about this in our next unit.

Example 3.2.2. Another familiar example is the alternating zeta series

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^s},$$

again when $s = \sigma \in \mathbb{R}$ known as the Euler-Dedekind function. It is proved in elementary calculus that this series converges for $\sigma > 0$, where the convergence is conditional for $0 < \sigma \leq 1$ and absolute for $1 < \sigma$.

In this section we shall prove that very similar results hold, with appropriate hypotheses on the coefficients a_n, for $s \in \mathbb{C}$, that is, dropping the condition $t = 0$.

3.3 Convergence of Dirichlet's series

Recall that for a power series $\sum_{n=1}^{\infty} a_n z^n$, there exists a value $R \in [0, \infty]$, called the radius of convergence, such that

1. if $|z| < R$, then the power series converges;
2. if $|z| < R$, then the power series diverges;

3. for any $r < R$, the series converges uniformly and absolutely in $\{|z| \leq R\}$ and the sum is bounded on this set;

4. on the circle $\{|z| = R\}$, the behavior is more delicate.

As we will see, a Dirichlet's series has an **abscissa of convergence** $\sigma_0(f)$ such that the series converges for all $s \in \mathbb{C}$ with $\mathrm{Re}s > \sigma_0(f)$ and diverges for all $s \in \mathbb{C}$ with $\mathrm{Re}s < \sigma_0(f)$. For instance, the abscissa of convergence for the Riemann zeta function $\zeta(s)$ is 1.

Before we go into further details, we prove the following lemmas.

Lemma 3.3.1. Let $\alpha, \beta, \sigma \in \mathbb{R}$, $0 < \sigma$, $0 < \alpha < \beta$. Then

$$|\, \mathrm{e}^{-\alpha s} - \mathrm{e}^{-\beta s} \,| \leq \frac{|s|}{\sigma}(\mathrm{e}^{-\alpha\sigma} - \mathrm{e}^{-\beta\sigma}).$$

Proof. We have,

$$\mathrm{e}^{-\alpha s} - \mathrm{e}^{-\beta s} = s \int_\alpha^\beta \mathrm{e}^{-us}\, du,$$

hence,

$$|\, \mathrm{e}^{-\alpha s} - \mathrm{e}^{-\beta s} \,| \leq |s| \int_\alpha^\beta |\, \mathrm{e}^{-us} \,| du = |s| \int_\alpha^\beta \mathrm{e}^{-u\sigma}\, du = \frac{|s|}{\sigma}(\mathrm{e}^{-\alpha\sigma} - \mathrm{e}^{-\beta\sigma}).$$

□

Note 3.3.1. Setting $\alpha = \log(m),\ \beta = \log(n)$ in the above lemma, $0 < m < n,\ \sigma > 0$, then

$$|m^{-s} - n^{-s}| \leq \frac{|s|}{\sigma}(m^{-\sigma} - n^{-\sigma}).$$

Lemma 3.3.2. (Abel's Summation by parts formula) Let $A_n = \sum_{k=1}^n a_k$, then

$$\sum_{k=1}^n a_k b_k = A_n b_{n+1} - \sum_{k=1}^n A_k (b_{k+1} - b_k).$$

Proof. Since $a_k = A_k - A_{k-1}$, we have

$$\begin{aligned} \sum_{k=1}^n a_k b_k &= \sum_{k=1}^n [A_k - A_{k-1}] b_k \\ &= \sum_{k=1}^n A_k b_k - \sum_{k=1}^n A_k b_{k+1} + A_n b_{n+1}. \end{aligned}$$

Hence the result. □

Corollary 3.3.1. The sum $\sum_{k=1}^\infty a_k b_k$ converges if both $\sum_{k=1}^\infty A_k(b_{k+1} - b_k)$ and $\{A_n b_{n+1}\}$ are convergent.

We remark that Abel's summation formula can be thought of as a discrete version of the familiar integration by parts formula from calculus, this should be clear by writing them side by side as

$$\sum_{k=1}^n a_k b_k = A_n b_{n+1} - \sum_{k=1}^n A_k (b_{k+1} - b_k), \quad \int u dv = vu - \int v du.$$

Theorem 3.3.1. Consider $\sum_{n=1}^\infty a_n n^{-s}$, $a_n \in \mathbb{C}$. Let $A_n = a_1 + a_2 + \cdots$. If $\{|A_n|\}$ is bounded, then the series converges for $\sigma > 0$.

Proof. We have, $|A_n| \leq C$, for some $C > 0$ and for all n. We shall use corollary 3.3.1, with $a_n = A_n$ and $b_n = n^{-s}$. Then

$$|A_n b_{n+1}| = |A_n| \cdot |b_{n+1}| \leq C \cdot (n+1)^{-\sigma} \to 0, \text{ as } n \to \infty.$$

Hence the second condition of Corollary (3.3.1) is satisfied, that is, $\{A_n b_{n+1}\}$ converges (in this case, to 0).

For the first condition, we apply the Cauchy convergence criterion to $\sum_{k=1}^{\infty} A_k((k+1)^{-s} - k^{-s})$. Given $\epsilon > 0$ and using note 3.3.1, if $\{S_n\}$ is the partial sum associated with the series, then we have,

$$\begin{aligned} |S_n - S_m| &= \left| \sum_{k=m+1}^{n} A_k((k+1)^{-s} - k^{-s}) \right| \\ &\leq C \sum_{k=m+1}^{n} |(k+1)^{-s} - k^{-s})| \\ &\leq \frac{C|s|}{\sigma} \sum_{k=m+1}^{n} \left(\frac{1}{k^{\sigma}} - \frac{1}{(k+1)^{\sigma}} \right) \\ &= \frac{C|s|}{\sigma} \left(\frac{1}{(m+1)^{\sigma}} - \frac{1}{(n+1)^{\sigma}} \right) \\ &\leq \frac{C|s|}{\sigma(m+1)^{\sigma}} < \epsilon \end{aligned}$$

for sufficiently large m. Hence, the result follows. □

We now state an elementary theorem for the convergence of Dirichlet's series.

Theorem 3.3.2. If the series is convergent for $s = \sigma + it$, then it is convergent for any value of s whose real part is greater than σ.

This theorem is included in the more general and less elementary theorem which follows. The above theorem can be obtained as a corollary of the theorem that follows.

Theorem 3.3.3. If the series $\sum_{n=1}^{\infty} a_n n^{-s}$ converges at some $s_0 \in \mathbb{C}$, then, for every $\delta > 0$, it converges uniformly in the sector

$$\left\{ s : \ -\frac{\pi}{2} + \delta < \arg(s - s_0) < \frac{\pi}{2} - \delta \right\}.$$

Proof. Without any loss of generality, we may assume that $s_0 = 0$, that is, the series $\sum_{n=1}^{\infty} a_n$ converges. Let $r_n = \sum_{k=n+1}^{\infty} a_k$, and fix $\epsilon > 0$. Then there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$, $|r_n| < \epsilon$. Using summation by parts, for s in the sector and $M, N > n_0$, we get,

$$\begin{aligned} \sum_{n=M}^{N} a_n n^{-s} &= \sum_{n=M}^{N} (r_{n-1} - r_n) n^{-s} \\ &= \sum_{n=M}^{N-1} r_n \left[\frac{1}{(n+1)^s} - \frac{1}{n^s} \right] + \frac{r_{M-1}}{M^s} - \frac{r_N}{N^s} \end{aligned} \tag{3.3.1}$$

The absolute values of the last two terms are bounded by ϵ, numerators are bounded by ϵ, while the denominators have absolute value at least 1. To estimate the summation part of (3.3.1), note that

$$\frac{1}{(n+1)^s} - \frac{1}{n^s} = \int_n^{n+1} \frac{-s}{x^{s+1}} dx,$$

so that

$$\left|\frac{1}{(n+1)^s} - \frac{1}{n^s}\right| \le |s| \int_n^{n+1} \frac{dx}{|x^{s+1}|} = \frac{|s|}{\sigma}\left[\frac{1}{n^\sigma} - \frac{1}{(n+1)^\sigma}\right]. \tag{3.3.2}$$

Thus the absolute value of the summation part of (3.3.1) satisfies for $M, N > n_0$,

$$\begin{aligned}\left|\sum_{n=M}^{N-1} r_n \left[\frac{1}{(n+1)^s} - \frac{1}{n^s}\right]\right| &\le \sum_{n=M}^{N-1} |r_n| \frac{|s|}{\sigma}\left[\frac{1}{n^\sigma} - \frac{1}{(n+1)^\sigma}\right] \\ &\le \epsilon \frac{|s|}{\sigma} \sum_{n=M}^{N-1} \left[\frac{1}{n^\sigma} - \frac{1}{(n+1)^\sigma}\right] \\ &\le \epsilon \frac{|s|}{\sigma}\left[\frac{1}{M^\sigma} - \frac{1}{N^\sigma}\right] \\ &\le c(\delta)\epsilon,\end{aligned}$$

since

$$\frac{|s|}{\sigma} = \left|\frac{1}{\cos(\arg s)}\right| \le \frac{1}{\cos\left(\frac{\pi}{2} - \delta\right)} =: c(\delta).$$

This proves that the series is uniformly Cauchy, and hence uniformly convergent. □

There are now three possibilities as regards the convergence of the series. It may converge for all, or no, or some values of s. In the last case it follows from theorem 3.3.2, by a classical argument, that we can find a number σ_0 such that the series is convergent for $\sigma > \sigma_0$ and divergent or oscillatory for $\sigma < \sigma_0$.

Theorem 3.3.4. The series may be convergent for all values of s, or for none, or for some only. In the last case there is a number σ_0 such that the series is convergent for $\sigma > \sigma_0$ and divergent or oscillatory for $\sigma < \sigma_0$.

Proof. If the series converges at some $s_0 \in \mathbb{C}$, the theorem follows from the inclusion

$$\{s : \text{ Re } s = \sigma > \sigma_0\} \subset \bigcup_{\delta > 0} \{s : \ |s - s_0| < \pi/2 - \delta\}.$$

□

In other words the region of convergence is a half-plane. We call σ_0 as the **abscissa of convergence**, and the line $\sigma = \sigma_0$ as the **line of convergence**. It is convenient to write $\sigma_0 = -\infty$ or $\sigma_0 = \infty$ when the series is convergent for all or no values of s. On the line of convergence the question of the convergence of the series remains open, and requires considerations of a much more delicate character.

We formally define the abscissa of convergence as follows:

Definition 3.3.1. The abscissa of convergence of the Dirichlet series $\sum_{n=1}^{\infty} a_n n^{-s}$ is the extended real number $\sigma_0 \in [-\infty, \infty]$ with the following properties

1. If Re $s > \sigma_0$, then the series converges;

2. If Re $s < \sigma_0$, then the series diverges; If Re $s = \sigma_0$, nothing can be said about the convergence of the series.

Note 3.3.2. To determine the abscissa of convergence, it is enough to look at convergence of the series for $s \in \mathbb{R}$.

Example 3.3.1. 1. The series $\sum_{n=1}^{\infty} a^n n^{-s}$, where $|a| < 1$, is convergent for all s. If $|a| > 1$, then the series converges for no value of s. And for $a = 1$, it is not convergent at any point of the line of convergence, diverging to $+\infty$ for $s = 1$ and oscillating finitely for other values of s.

2. The series $\sum_{n=2}^{\infty}(\log n)^{-2} n^{-s}$ has the same line of convergence as the last series, but is convergent (indeed absolutely convergent) at all points of the line.

3. The series $\sum_{n=2}^{\infty} a_n n^{-s}$ where $a_n = (-1)^n + (\log n)^{-2}$, has the same line of convergence, and is convergent (though not absolutely) at all points of it.

We also have an abscissa of absolute convergence of a Dirichlet's series $\sum_{n=1}^{\infty} a_n n^{-s}$.

Definition 3.3.2. Given a Dirichlet's series $\sum_{n=1}^{\infty} a_n n^{-s}$, the abscissa of absolute convergence is defined as

$$\begin{aligned}\sigma_a &= \inf\left\{\rho : \sum_{n=1}^{\infty} a_n n^{-s} \text{ converges absolutely for some s with } Re\ s = \rho\right\} \\ &= \inf\left\{\rho : \sum_{n=1}^{\infty} a_n n^{-s} \text{ converges absolutely for all s with } Re\ s \geq \rho\right\}.\end{aligned}$$

The following theorem gives the relationship between σ_0 and σ_a for a Dirichlet's series.

Theorem 3.3.5. For any Dirichlet series, we have

$$\sigma_0 \leq \sigma_a \leq \sigma_0 + 1.$$

Proof. The first inequality is obvious. For the second, assume, that $\sigma_0 = 0$. We need to show that for $\sigma > 1$, $\sum_{n=1}^{\infty} |a_n n^{-s}|$ converges. Take $\epsilon > 0$ such that $\sigma - \epsilon > 1$. Then,

$$\sum_{n=1}^{\infty} |a_n n^{-s}| = \sum_{n=1}^{\infty} \frac{|a_n|}{n^{\sigma}} = \sum_{n=1}^{\infty} \frac{|a_n|}{n^{\epsilon}} \cdot \frac{1}{n^{\sigma-\epsilon}} \leq C \sum_{n=1}^{\infty} \frac{1}{n^{\sigma-\epsilon}} < \infty,$$

where, $C := \sup_n |a_n/n^{\epsilon}|$ is finite, since $\sigma_0 = 0$. □

Remark 3.3.1. If $a_n > 0$ for all $n \in \mathbb{N}$, then $\sigma_0 = \sigma_a$. This follows immediately by considering $s \in \mathbb{R}$.

Recall that for the radius of convergence of a power series, we have the following formula

$$1/R = \limsup_{n \to \infty} |a_n|^{\frac{1}{n}}.$$

The following is an analogous formula for the abscissa of convergence of a Dirichlet's series.

Theorem 3.3.6. Let $\sum_{n=1}^{\infty} a_n n^{-s}$ be a Dirichlet's series, and let σ_0 be its abscissa of convergence. Let $s_n = a_1 + a_2 + \cdots + a_n$ and $r_n = a_{n+1} + a_{n+1} + \cdots$

1. If $\sum a_n$ diverges, then
$$0 \leq \sigma_0 = \limsup_{n \to \infty} \frac{\log |s_n|}{\log n}.$$

2. If $\sum a_n$ converges, then
$$0 \geq \sigma_0 = \limsup_{n \to \infty} \frac{\log |r_n|}{\log n}.$$

Proof. 1. We assume that the series $\sum_{n=1}^{\infty} a_n n^{-s}$ diverges and define

$$\alpha = \limsup_{n\to\infty} \frac{\log|s_n|}{\log n}.$$

We will first show that $\alpha \leq \sigma_0$. Assume that $\sum_{n=1}^{\infty} a_n n^{-\sigma}$ converges. Thus, $\sigma > 0$ and we need to show that $\sigma \geq \alpha$. Let $b_n = a_n n^{-s}$ and $B_n = \sum_{k=1}^{n} b_k$ (so that $B_0 = 0$). By assumption, the sequence $\{B_n\}$ is bounded, say by M, and we can use the summation by parts as follows:

$$\begin{aligned} s_N &= \sum_{n=1}^{N} a_n \\ &= \sum_{n=1}^{N} b_n n^{\sigma} \\ &= \sum_{n=1}^{N-1} B_n[n^{\sigma} - (n+1)^{\sigma}] + B_N N^{\sigma} \end{aligned}$$

so that

$$|s_N| \leq M \sum_{n=1}^{N-1} [(n+1)^{\sigma} - n^{\sigma}] + MN^{\sigma} \leq 2MN^{\sigma}.$$

Applying the natural logarithm to both sides yields

$$\log|s_N| \leq \sigma \log N + \log 2M,$$

so,

$$\frac{\log|s_n|}{\log N} \leq \sigma + \frac{\log 2M}{\log N},$$

and this tends to σ as $N \to \infty$, giving the desired upper bound for α.

We need to show the other inequality $\sigma_0 \leq \alpha$. Suppose $\sigma > \alpha$. We need to show that $\sum_{n=1}^{\infty} a_n n^{-\sigma}$ converges. Chosen an $\epsilon > 0$ such that $\alpha + \epsilon < \sigma$. By definition, there exists $n_0 \in \mathbb{N}$ such that for all $n \geq n_0$,

$$\frac{\log|s_n|}{\log n} \leq \alpha + \epsilon.$$

This implies that

$$\log|s_n| \leq (\alpha+\epsilon)\log n = \log\left(n^{\alpha+\epsilon}\right).$$

Using summation by parts, we can compute

$$\begin{aligned} \sum_{n=M+1}^{N} \frac{a_n}{n^{\sigma}} &= \sum_{n=M}^{N} s_n[n^{-\sigma} - (n+1)^{\sigma}] + S_N(N+1)^{\sigma} - s_M M^{-\sigma} \\ &\leq \sum_{n=M}^{N} n^{\alpha+\epsilon}[\sigma n^{-\sigma-1}] + N^{\alpha+\epsilon}N^{-\sigma} + M^{\alpha+\epsilon}M^{-\sigma} \\ &\lesssim (M-1)^{\alpha+\epsilon-\sigma}, \end{aligned}$$

and the last quantity tends to zero as M tends to ∞.

We estimated $\sum_{n=M}^{N} n^{\alpha+\epsilon-\sigma-1}$ by the integral

$$\int_{M-1}^{N-1} x^{\alpha+\epsilon-\sigma-1} dx \lesssim (M-1)^{\alpha+\epsilon-\sigma},$$

and the symbol $\lesssim$ means less than or equal to a constant times the right hand-side (where the constant depends on $\alpha+\epsilon-\sigma$, but, critically, not on M).

2. Similar to the first part.

□

From the formulae above we can simply deduce formulae for the abscissa of absolute convergence, although these can be derived easily on their own.

Corollary 3.3.2. For a Dirichlet's series $\sum_{n=1}^{\infty} a_n n^{-s}$, we have

1. if $\sum |a_n|$ diverges, then

$$\sigma_a = \limsup_{n\to\infty} \frac{\log(|a_1| + |a_2| + \cdots + |a_n|)}{\log n} \geq 0,$$

2. if $\sum |a_n|$ converges, then

$$\sigma_a = \limsup_{n\to\infty} \frac{\log(|a_{n+1}| + |a_{n+2}| + \cdots)}{\log n} \leq 0.$$

Example 3.3.2. The series

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{p_n^s}$$

(where, p_n are primes) has $\sigma_0 = 0$ and $\sigma_a = 1$.

The series of coefficients diverges and so we use the first of the pair of formulae for each abscissae

$$\sigma_0 = \limsup_{n\to\infty} \frac{\log 1}{\log n} = 0,$$

and, using the prime number theorem,

$$\sigma_a = \limsup_{n\to\infty} \frac{\log(\pi(n))}{\log n} = \limsup_{n\to\infty} \frac{\log n - \log(\log n)}{\log n} = 1,$$

where, $\pi(x)$ denotes the number of primes less than or equal to x.

Theorem 3.3.7. Suppose that the series $\sum_{n=1}^{\infty} a_n n^{-s}$ converges absolutely to some $f(s)$ in some half-plane $\mathbb{H}_c = \{s : \text{ Re } s > c\}$ and $f(s) \equiv 0$ in the half-plane $\mathbb{H}_c$. Then $a_n = 0$ for all $n \in \mathbb{N}$.

Proof. We may assume that $c < 0$, so, in particular, $\sum |a_n| < \infty$. Suppose that all a_n's are not zero, and let n_0 be the be the smallest natural number such that $a_{n_0} \neq 0$.

We claim that $\lim\limits_{\sigma\to\infty} f(\sigma) n_0^{\sigma} = a_{n_0}$. To prove the claim that

$$\begin{aligned} 0 &\leq n_0^{\sigma} \left| \sum_{n>n_0} a_n n^{-\sigma} \right| \\ &\leq \sum_{n>n_0} |a_n| \left(\frac{n_0}{n}\right)^{\sigma} \\ &\leq \left(\frac{n_0}{n_0+1}\right)^{\sigma} \sum_{n>n_0} |a_n|, \end{aligned}$$

and the last term tends to 0 as $\sigma \to \infty$, since $\sum |a_n|$ converges. As

$$f(\sigma)n_0^\sigma = a_{n_0} + n_0^\sigma \sum_{n>n_0} a_n n^{-\sigma},$$

the claim is proved.

The proof is also finished, because the limit in the claim is obviously 0, a contradiction. □

Few Probable Questions

---

1. Define the abscissa of convergence of a Dirichlet's series $\sum_{n=1}^{\infty} a_n n^{-s}$. Show that if the series diverges, then
$$0 \leq \sigma_0 = \limsup_{n\to\infty} \frac{\log |s_n|}{\log n}.$$
2. Define the abscissa of absolute convergence of a Dirichlet's series $\sum_{n=1}^{\infty} a_n n^{-s}$. Show that $\sigma_0 \leq \sigma_a \leq \sigma_0 + 1$.
3. Show that if the series $\sum_{n=1}^{\infty} a_n n^{-s}$ converges for some $s_0 \in \mathbb{C}$, then, for every $\delta > 0$, it converges uniformly in the sector
$$\left\{ s : \ -\frac{\pi}{2} + \delta < \arg(s - s_0) < \frac{\pi}{2} - \delta \right\}.$$

Unit 4

Course Structure

- The Riemann Zeta function
- The product development and the zeros of the zeta functions.

4.1 Introduction

As we have introduced in the previous unit, the Riemann zeta function $\zeta(s)$ is a function of the complex variable s, defined as

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

which plays a pivotal role in the analytic number theory and has applications in physics, probability theory, and applied statistics.

As a function of a real variable, Leonhard Euler first introduced and studied it in the first half of the eighteenth century without using complex analysis, which was not available at the time. Bernhard Riemann's 1859 article "On the Number of Primes Less Than a Given Magnitude" extended the Euler definition to a complex variable, proved its meromorphic continuation and functional equation, and established a relation between its zeros and the distribution of prime numbers.

The values of the Riemann zeta function at even positive integers were computed by Euler. The first of them, $\zeta(2)$, provides a solution to the Basel problem. In 1979 Roger Apéry proved the irrationality of $\zeta(3)$. The values at negative integer points, also found by Euler, are rational numbers and play an important role in the theory of modular forms. Many generalizations of the Riemann zeta function, such as Dirichlet series, Dirichlet L-functions and L-functions, are known. We will however not indulge into such rigorous treatments of the zeta function. We will only restrict ourselves to some preliminary ideas, starting with the definition, convergence, etc.

Objectives

After reading this unit, you will be able to

- define the Riemann zeta function and know about its origins in a preliminary level

- deduce the product development of the zeta function
- deduce the functional equation and its other forms

4.2 Riemann Zeta Function

The Riemann zeta function, as we have already seen, is the function

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

where, $s = \sigma + it$ is a complex number. First, we will discuss the convergence of the function. See that

$$\begin{aligned} |\zeta(s)| &= \left|\sum_{n=1}^{\infty} \frac{1}{n^s}\right| \\ &\leq \sum_{n=1}^{\infty} \frac{1}{|n^s|} \\ &= \sum_{n=1}^{\infty} \frac{1}{|n^{\sigma+it}|} \\ &= \sum_{n=1}^{\infty} \frac{1}{|n^{\sigma}|\cdot|n^{it}|} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^{\sigma}|\,\mathrm{e}^{it\log(n)}\,|} \\ &= \sum_{n=1}^{\infty} \frac{1}{n^{\sigma}}. \end{aligned}$$

which converges for all $\sigma > 1$. Hence, $\sigma_0 = 1$ is the abscissa of convergence of the series $\sum n^{-s}$. Also, the series is not convergent on the line of convergence. Also, since for $\sigma > 1$,

$$\left|\frac{1}{n^s}\right| \leq \frac{1}{n^{\sigma}},$$

and since the series $\sum_{n=1}^{\infty} n^{-\sigma}$ converges in the said region, so by Weierstrass M-test, the series $\sum n^{-s}$ converges uniformly and absolutely in the hlf-plane $\sigma > 1$, and thus, defines an analytic function in the plane $\mathbb{H}_1 = \{s \in \mathbb{C} : \text{ Re } s > 1\}$.

4.3 The Product Development

The number-theoretic properties of $\zeta(s)$ are inherent in the following connection between the zeta function and the ascending sequence of primes $p_1, p_2, \dots, p_n, \dots$

Theorem 4.3.1. For $\sigma > 1$, we have

$$\frac{1}{\zeta(s)} = (1 - p_1^{-s})(1 - p_2^{-s}) \cdots (1 - p_n^{-s}) \cdots = \Pi_{n=1}^{\infty}(1 - p_n^{-s}),$$

where, $p_1, p_2, \dots, p_n, \dots$, are prime numbers and the term on the right hand side of the above equation is the infinite product of the numbers $(1 - p_n^{-s})$.

Proof. Under the assumption that $\sigma > 1$, we see that

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}. \tag{4.3.1}$$

Multiplying equation (4.3.1) by 2^{-s}, we get

$$\zeta(s) \cdot \frac{1}{2^s} = \sum_{n=1}^{\infty} \frac{1}{n^s} \cdot \frac{1}{2^s} = \sum_{n=1}^{\infty} \frac{1}{(2n)^s}. \tag{4.3.2}$$

Subtracting equation (4.3.2) from (4.3.1), we get,

$$\zeta(s)\left(1 - \frac{1}{2^s}\right) = \sum_{n=1}^{\infty} \frac{1}{n^s} - \sum_{n=1}^{\infty} \frac{1}{(2n)^s} = \sum_{n=1;\ n\neq 2k}^{\infty} \frac{1}{n^s}. \tag{4.3.3}$$

The last term in the above equation is the sum of all terms n^{-s}, excluding the terms n which are multiples of 2.

Again, multiplying the equation (4.3.3) by 3^{-s}, we get

$$\zeta(s)\left(1 - \frac{1}{2^s}\right) \cdot \frac{1}{3^s} = \sum_{n=1;\ n\neq 2k}^{\infty} \frac{1}{n^s} \cdot \frac{1}{3^s} = \sum_{n=1;\ n\neq 2k}^{\infty} \frac{1}{(3n)^s}. \tag{4.3.4}$$

Subtracting equation (4.3.4) from (4.3.3), we get,

$$\zeta(s)\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right) = \sum_{n=1;\ n\neq 2k;\ n\neq 3k}^{\infty} \frac{1}{n^s}.$$

The last term in the above equation is the sum of all terms n^{-s}, excluding the terms n which are multiples of 2 and 3.

Continuing in this way, we get,

$$\zeta(s)\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right)\cdots\left(1 - \frac{1}{p_n^s}\right) = \sum_{\cdots\ n\neq p_n k}^{\infty} \frac{1}{n^s},$$

where, the term on the right hand side of the above equation is the sum of all those terms n^{-s}, which are not the multiples of the primes $2, 3, 5, \ldots, p_n$ arranged in ascending order. Thus, taking limit as $n \to \infty$, we get

$$\lim_{n\to\infty} \zeta(s)\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right)\cdots\left(1 - \frac{1}{p_n^s}\right) = \sum_{n\neq p_n k} \frac{1}{n^s},$$

where, the sum is taken over all such n^{-s}, such that n is not a multiple of any prime p_n and such number can be none other than 1. So, the above equation becomes,

$$\zeta(s)\left(1 - \frac{1}{2^s}\right)\left(1 - \frac{1}{3^s}\right)\cdots\left(1 - \frac{1}{p_n^s}\right)\cdots = 1,$$

which finally gives,

$$\frac{1}{\zeta(s)} = (1 - p_1^{-s})(1 - p_2^{-s})\cdots(1 - p_n^{-s})\cdots,$$

where, $p_1, p_2, p_3, \ldots, p_n, \ldots$ is the complete list of prime numbers arranged in ascending order. $\square$

The above representation of the zeta function is called the Euler product representation of the zeta function. Also, notice that the product development explained above, includes the introduction of an infinite product. Infinite products, like the infinite series, are convergent when the sequence of partial products converge as we will see in subsequent units. The infinite product in this case is convergent uniformly in the region $\sigma > 1$.

We have taken for granted that there are infinitely many primes. Actually, the reasoning can be used to prove this fact. For if p_n were the largest prime, then we would have got

$$\zeta(s)\left(1-\frac{1}{2^s}\right)\left(1-\frac{1}{3^s}\right)\cdots\left(1-\frac{1}{p_n^s}\right)=1$$

and it would follow that $\zeta(\sigma)$ has a finite limit when $\sigma \to 1$. This contradicts the divergence of the series $\sum_{n=1}^{\infty} n^{-1}$.

4.4 Functional Equations

4.4.1 Relationship with the Gamma Function

We are familiar with the gamma function which is written as

$$\Gamma(s)=\int_0^{\infty} t^{s-1}\,\mathrm{e}^{-t}\,dt,$$

and the integral converges for all values of Re $s > 0$. We make the substitution

$$t = nu \Rightarrow dt = ndu,$$

where n is positive integer. So the above integral changes to

$$\Gamma(s)\int_0^{\infty}(nu)^{s-1}\,\mathrm{e}^{-nu}\,ndu=\int_0^{\infty} n^s u^{s-1}\,\mathrm{e}^{-nu}\,du,$$

which gives

$$\Gamma(s)\frac{1}{n^s}\int_0^{\infty} u^{s-1}\,\mathrm{e}^{-nu}\,du.$$

Summing over n from 1 to ∞, we get,

$$\Gamma(s)\sum_{n=1}^{\infty}\frac{1}{n^s}=\sum_{n=1}^{\infty}\int_0^{\infty} u^{s-1}\,\mathrm{e}^{-nu}\,du$$

Since the integral on the right hand side is absolutely converging, so the sum and integral can be exchanged. Thus, the above equation changes to

$$\begin{aligned}
\Gamma(s)\zeta(s) &= \int_0^{\infty} u^{s-1}\sum_{n=1}^{\infty}\mathrm{e}^{-nu}\,du\\
&= \int_0^{\infty} u^{s-1}\left(\frac{1}{1-\mathrm{e}^{-u}}-1\right)du\\
&= \int_0^{\infty} u^{s-1}\frac{\mathrm{e}^{-u}}{1-\mathrm{e}^{-u}}du\\
&= \int_0^{\infty}\frac{u^{s-1}}{\mathrm{e}^{u}-1}du.
\end{aligned}$$

for Re $s > 1$. Thus, the celebrated relationship between the gamma and zeta function is given by

$$\Gamma(s)\zeta(s) = \int_0^\infty \frac{u^{s-1}}{\mathrm{e}^u - 1} du$$

for Re $s > 1$.

4.4.2 Theta Function

We need to learn certain basics of the Jacobi theta function that we will need in the sequel. The theta function is given by

$$\vartheta(x) = \sum_{n \in \mathbb{Z}} \mathrm{e}^{-\pi n^2 x}.$$

Let f be any complex function that is analytic in the strip $\{z \in \mathbb{C} : |\text{Im } z| < a\}$, and $|f(x+iy)| \leq A/(1+x^2)$ for some constant $A > 0$ and all $x \in \mathbb{R}$ such that $|y| < a$ for $a > 0$. $\mathrm{e}^{-\pi n^2 z}$ satisfies the properties stated thus.

By Poisson summation formula, we have for such f as described above,

$$\sum_{n \in \mathbb{Z}} f(n) = \sum_{n \in \mathbb{Z}} \hat{f}(n),$$

where, $\hat{f}$ is the Fourier transform of f. The Fourier transform of the function $\mathrm{e}^{\pi x^2}$ is the function itself, that is,

$$\int_{-\infty}^{\infty} \mathrm{e}^{-\pi x^2} \mathrm{e}^{-2\pi i x \xi} \, dx = \mathrm{e}^{-\pi \xi^2}.$$

For fixed values of $t > 0$ and $a \in \mathbb{R}$, the change of variables $x \mapsto t^{1/2}(x + a)$ in the above integral show that the Fourier transform of the function

$$f(x) = \mathrm{e}^{-\pi t (x+a)^2},$$

for fixed values of $t > 0$ and $a \in \mathbb{R}$, we get

$$\hat{f}(\xi) = t^{-1/2} \mathrm{e}^{-\pi \xi^2 / t} \mathrm{e}^{2\pi i a \xi}.$$

Applying the Poisson summation formula to the above pair we get,

$$\sum_{n \in \mathbb{Z}} \mathrm{e}^{-\pi t (n+a)^2} = \sum_{n \in \mathbb{Z}} t^{-1/2} \mathrm{e}^{-\pi n^2 / t} \mathrm{e}^{2\pi i n a}.$$

This identity has noteworthy consequences. For instance, the special case $a = 0$ is the transformation law for the theta function we defined above. Thus, we get,

$$\vartheta(t) = t^{-1/2} \vartheta(1/t), \tag{4.4.1}$$

for $t > 0$.

4.4.3 Functional equations

Now, we will derive the Riemann functional equation. The equation is

$$\pi^{-\frac{s}{2}} \Gamma\left(\frac{s}{2}\right) \zeta(s) = \pi^{-\frac{1-s}{2}} \Gamma\left(\frac{1-s}{2}\right) \zeta(1-s).$$

We have

$$\Gamma(s) = \int_0^\infty t^{s-1}\,\mathrm{e}^{-t}\,dt.$$

Thus,

$$\Gamma\left(\frac{s}{2}\right) = \int_0^\infty t^{\frac{s}{2}-1}\,\mathrm{e}^{-t}\,dt. \tag{4.4.2}$$

We use the substitution

$$t = \pi n^2 x \Rightarrow dt = \pi n^2 dx$$

Thus, equation (4.4.2) becomes

$$\Gamma\left(\frac{s}{2}\right) = \int_0^\infty (\pi n^2 x)^{\frac{s}{2}-1}\,\mathrm{e}^{-\pi n^2 x}\,\pi n^2 dx = \int_0^\infty \pi^{\frac{s}{2}} n^s x^{\frac{s}{2}-1}\,\mathrm{e}^{-\pi n^2 x}\,dx. \tag{4.4.3}$$

Multiplying (4.4.3) by $\pi^{-s/2}\cdot n^{-s}$, we get

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\cdot\frac{1}{n^s} = \int_0^\infty x^{\frac{s}{2}-1}\,\mathrm{e}^{-\pi n^2 x}\,dx.$$

Summing the above equation over, from 1 to ∞, we get,

$$\sum_{n=1}^\infty \pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\cdot\frac{1}{n^s} = \sum_{n=1}^\infty \int_0^\infty x^{\frac{s}{2}-1}\,\mathrm{e}^{-\pi n^2 x}\,dx$$

which gives,

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \int_0^\infty x^{\frac{s}{2}-1}\sum_{n=1}^\infty \mathrm{e}^{-\pi n^2 x}\,dx. \tag{4.4.4}$$

We have,

$$\vartheta(x) = \sum_{n\in\mathbb{Z}} \mathrm{e}^{-\pi n^2 x} = 1 + 2\sum_{n=1}^\infty \mathrm{e}^{-\pi n^2 x} = 1 + 2\psi(x).$$

Replacing this in the right hand side of equation (4.4.4) we get,

$$\begin{aligned}\int_0^\infty x^{\frac{s}{2}-1}\sum_{n=1}^\infty \mathrm{e}^{-\pi n^2 x}\,dx &= \int_0^\infty x^{\frac{s}{2}-1}\psi(x)dx\\ &= \int_0^1 x^{\frac{s}{2}-1}\psi(x)dx + \int_1^\infty x^{\frac{s}{2}-1}\psi(x)dx. \end{aligned} \tag{4.4.5}$$

We have, by (4.4.1),

$$2\psi(x) + 1 = \frac{1}{\sqrt{x}}\left(2\psi\left(\frac{1}{x}\right) + 1\right)$$

which gives,

$$\psi(x) = \frac{1}{\sqrt{x}}\psi\left(\frac{1}{x}\right) + \frac{1}{2\sqrt{x}} - \frac{1}{2}.$$

Thus,

$$\begin{aligned}\int_0^1 x^{\frac{s}{2}-1}\psi(x)dx &= \int_0^1 x^{\frac{s}{2}-1}\left(\frac{1}{\sqrt{x}}\psi\left(\frac{1}{x}\right) + \frac{1}{2\sqrt{x}} - \frac{1}{2}\right)dx\\ &= \int_0^1 \left[x^{\frac{s}{2}-\frac{3}{2}}\psi\left(\frac{1}{x}\right) + \frac{1}{2}\left(x^{\frac{s}{2}-\frac{3}{2}} - x^{\frac{s}{2}-1}\right)\right]dx\\ &= \int_0^1 x^{\frac{s}{2}-\frac{3}{2}}\psi\left(\frac{1}{x}\right)dx + \frac{1}{s(s-1)}. \end{aligned} \tag{4.4.6}$$

Using the substitution in the integral on the right hand of the above equation

$$x = \frac{1}{u} \Rightarrow dx = -\frac{1}{u^2}du$$

and the limits also change accordingly and the equation (4.4.6) becomes, with a change in the dummy variable u to x,

$$\int_0^1 x^{\frac{s}{2}-1}\psi(x)dx = \int_1^\infty x^{-\frac{s}{2}-\frac{1}{2}}\psi(x)dx + \frac{1}{s(s-1)}.$$

Finally, (4.4.5) becomes

$$\begin{aligned}\int_0^\infty x^{\frac{s}{2}-1}\psi(x)dx &= \int_1^\infty x^{\frac{s}{2}-1}\psi(x)dx + \int_1^\infty x^{-\frac{s}{2}-\frac{1}{2}}\psi(x)dx + \frac{1}{s(s-1)} \\ &= \int_1^\infty \left(x^{\frac{s}{2}-1} + x^{-\frac{s}{2}-\frac{1}{2}}\right)\psi(x)dx + \frac{1}{s(s-1)}.\end{aligned}$$

Thus, (4.4.4) and (4.4.5) together give

$$\begin{aligned}\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) &= \int_1^\infty \left(x^{\frac{s}{2}-1} + x^{-\frac{s}{2}-\frac{1}{2}}\right)\psi(x)dx + \frac{1}{s(s-1)} \\ &= \int_1^\infty \left(x^{\frac{s}{2}} + x^{\frac{1-s}{2}}\right)\frac{\psi(x)}{x}dx + \frac{1}{s(s-1)}. \qquad (4.4.7)\end{aligned}$$

Putting s by $1-s$ in the above equation, we get,

$$\begin{aligned}\pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s) &= \int_1^\infty \left(x^{\frac{1-s}{2}} + x^{\frac{1-(1-s)}{2}}\right)\frac{\psi(x)}{x}dx + \frac{1}{(1-s)(1-s-1)} \\ &= \int_1^\infty \left(x^{\frac{1-s}{2}} + x^{\frac{s)}{2}}\right)\frac{\psi(x)}{x}dx + \frac{1}{(1-s)s}. \qquad (4.4.8)\end{aligned}$$

Notice that the right hand sides of the equations (4.4.7) and (4.4.8) are the same. So, we get

$$\pi^{-\frac{s}{2}}\Gamma\left(\frac{s}{2}\right)\zeta(s) = \pi^{-\frac{1-s}{2}}\Gamma\left(\frac{1-s}{2}\right)\zeta(1-s).$$

Which is our required functional equation.

Few Probable Questions

1. With proper justification, find the abscissa of convergence of the zeta function.
2. Establish the Euler product representation of the zeta function.
3. Establish the relation between zeta function and the gamma function.
4. Deduce the Riemann functional equation.
5. Write the Riemann functional equation. Hence deduce that

$$\zeta(s) = 2^s\pi^{s-1}\sin\frac{\pi s}{2}\Gamma(1-s)\zeta(1-s).$$

Unit 5

Course Structure

- Entire functions, growth of an entire function
- Order and type and their representations in terms of the Taylor coefficients.

5.1 Introduction

In complex analysis, an entire function, also called an integral function, is a complex-valued function that is holomorphic at all finite points over the whole complex plane. Typical examples of entire functions are polynomials and the exponential function, and any finite sums, products and compositions of these, such as the trigonometric functions sine and cosine and their hyperbolic counterparts sinh and cosh, as well as derivatives and integrals of entire functions such as the error function. If an entire function $f(z)$ has a root at w, then $f(z)/(z-w)$, taking the limit value at w, is an entire function. On the other hand, neither the natural logarithm nor the square root is an entire function, nor can they be continued analytically to an entire function. Also, a transcendental entire function is an entire function that is not a polynomial. For example, the exponential function, sine and cosine functions are most common transcendental entire functions. This unit is dedicated to the study of the growth of entire functions.

Objectives

After reading this unit, you will be able to

- get more idea about the various entire functions
- learn the behaviour of the maximum modulus function $M(r)$ for entire functions
- define the order and type of entire functions
- define the order and type of entire functions with the help of the Taylor coefficients

5.2 Entire Functions

An entire function is a single valued function having Taylor series expansion

$$f(z) = a_0 + a_1 z + \cdots + a_n z^n + \cdots$$

which converges for all finite z. Moreover, according to the Cauchy Hadamard formula,

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} = 0.$$

There are three possible ways in which an entire function $f(z)$ can behave at infinity:

1. $f(z)$ can have a regular point at infinity, then, according to Liouville's theorem, f is a constant function;
2. f can have a pole of order $k \geq 1$ at infinity, and then f reduces to a polynomial;
3. f can have an essential singularity at infinity, and then f is said to be an entire transcendental function.

Note that, the behaviour of $f(z)$ at infinity is determined by the action of $f(1/z)$ at 0. We will be mainly concerned with the transcendental entire functions from now on. If f is such a function, then we will clearly have, since $M(r)$ is a strictly increasing function of r,

$$\lim_{r\to\infty} M(r) = \infty.$$

We have the following theorem for transcendental entire functions f.

Theorem 5.2.1. If $f(z)$ is a transcendental entire function, with maximum modulus function $M(r)$, then

$$\liminf_{r\to\infty} \frac{\log M(r)}{\log r} = \infty.$$

Proof. Let $f(z) = \sum_{k=0}^{\infty} a_k z^k$ be a transcendental entire function with maximum modulus function $M(r)$. If possible, let

$$\liminf_{r\to\infty} \frac{\log M(r)}{\log r} = \mu < \infty.$$

Then for $\epsilon > 0$, we can find an increasing sequence $\{r_n\}$, tending to ∞, such that

$$\frac{\log M(r_n)}{\log r_n} < \mu + \epsilon$$

for every r_n, that is,

$$\log M(r_n) < (\mu + \epsilon) \log r_n \Rightarrow M(r_n) < r_n^{\mu+\epsilon}$$

for every r_n. Hence, by Cauchy's inequality,

$$|a_k| \leq \frac{M(r_n)}{r_n^k} < r_n^{\mu+\epsilon-k}$$

for $k = 0, 1, 2, \ldots$. Since r_n can be chosen arbitrarily large, it follows that $a_k = 0$ for all $k > \mu + \epsilon > \mu$. Hence, f is a polynomial of degree not greater than $[\mu]$, that is, the largest integer less than or equal to μ. This contradicts our assumption that f is transcendental. Hence the result. □

We have another result analogous to the above for non-transcendental entire functions as follows.

Theorem 5.2.2. For an entire function f, if there exists a positive integer k such that

$$\lim_{r\to\infty} \frac{M(r)}{r^k} < \infty,$$

then f is a polynomial of degree k atmost.

Proof. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be an entire function with maximum modulus function $M(r)$. Let

$$\lim_{r\to\infty} \frac{M(r)}{r^k} = \mu < \infty.$$

Then,

$$M(r) < (\mu + \epsilon) r^k,$$

for any positive ϵ and all $r \geq r_0$, for some r_0. By Cauchy's inequality,

$$|a_n| \leq \frac{M(r)}{r^n} < (\mu + \epsilon) r^{k-n}$$

for all $r \geq r_0$. Since we can choose r sufficiently large, $|a_n| \to 0$ as $r \to \infty$ for $n > k$. Hence, $a_n = 0$ for all $n > k$ and thus f is a polynomial of degree at most k. □

Note that the above theorem would also be true if

$$\liminf_{r\to\infty} \frac{M(r)}{r^k} < \infty.$$

5.3 Order of an entire function

The preceding discussions assert that a transcendental entire function $f(z)$ grows faster than any fixed positive power of r. This suggests that using the exponential function (that is, the simplest "rapidly growing function") to measure the growth of $f(z)$. We now formally define the order of an entire function.

Definition 5.3.1. An entire function f is said to be of **finite** order if there exists a positive number k such that the inequality

$$\log M(r) < r^k,$$

or

$$M(r) < \mathrm{e}^{r^k}$$

holds for sufficiently large r. Then

$$\rho = \inf \ \{ \ k : \ M(r) < \mathrm{e}^{r^k} \text{ holds for sufficiently large } r\}$$

is called the **order** of f. If $\rho = \infty$, that is, for any number k, there exists arbitrarily large values of r such that $\log M(r) > r^k$, then f is said to be of infinite order.

For example, e^z if of finite order (in fact, of order 1), while $\mathrm{e}^{\mathrm{e}^z}$ is of infinite order. From the definition, it is clear that the order of an entire function is always non-negative.

Theorem 5.3.1. The order ρ of an entire function f is given by the formula

$$\rho = \limsup_{r\to\infty} \frac{\log\log M(r)}{\log r}.$$

Proof. Let ρ be the order of f. Then, from the definition of ρ, we have, for any $\epsilon > 0$, there exists a number $r_0(\epsilon) > 0$ such that

$$\log M(r) < r^{\rho+\epsilon}$$

holds for all $r > r_0$. On the other hand, there exists an increasing sequence $\{r_n\}$ tending to infinity, such that

$$\log M(r) > r^{\rho-\epsilon}$$

for $r = r_n$. In other words,

$$\frac{\log\log M(r)}{\log r} < \rho + \epsilon, \quad \forall r > r_0 \tag{5.3.1}$$

and

$$\frac{\log\log M(r)}{\log r} > \rho - \epsilon, \tag{5.3.2}$$

for a sequence of values of r tending to infinity. Equations (5.3.1) and (5.3.2) precisely means

$$\rho = \limsup_{r\to\infty} \frac{\log\log M(r)}{\log r}.$$

□

5.4 Type of an entire function of finite non-zero order

Next, we subdivide the class of entire functions of finite non-zero order ρ.

Definition 5.4.1. Let f be an entire function with finite non-zero order ρ. By the **type** τ of f, we mean the greatest lower bound of positive numbers k such that the inequality

$$\log M(r) < kr^{\rho} \tag{5.4.1}$$

holds for all sufficiently large r, that is,

$$\tau = \inf\{k : \text{(5.4.1) holds for all } r > r_0(k)\}.$$

However, suppose that $\tau = \infty$, that is, suppose that given any positive number k, there exists arbitrarily large values of r such that

$$\log M(r) > kr^{\rho}.$$

Then f is said to be of **infinite** (or, maximum) type. And when, $\tau = 0$, then f is said to be of minimum type. From the definition of type, it is clear that τ is always non-negative. When $0 < \tau < \infty$, then f is said to be of normal type.

Example 5.4.1.

1. If $f(z) = e^z$, then $\rho = 1$ and $\tau = 1$.
2. If $f(z) = e^{pz^n}$, where $p > 0$ and n is a positive integer, then $\rho = n$ and $\tau = p$.

Theorem 5.4.1. The type τ of an entire function f with finite non-zero order ρ is given by the formula

$$\tau = \limsup_{r\to\infty} \frac{\log M(r)}{r^{\rho}}.$$

Proof. Let τ be the type of an entire function f of order $\rho(\neq 0)$. Then, from the definition of τ, we have, for any $\epsilon > 0$, there exists a number $r_0(\epsilon) > 0$ such that

$$\log M(r) < (\tau + \epsilon)r^\rho, \quad \forall\, r > r_0.$$

On the other hand, there exists an increasing sequence $\{r_n\}$ tending to infinity, such that

$$\log M(r) > (\tau - \epsilon)r^\rho, \quad \forall\, r = r_n.$$

In other words,

$$\frac{\log M(r)}{r^\rho} < \tau + \epsilon, \quad \forall\, r > r_0 \tag{5.4.2}$$

and

$$\frac{\log M(r)}{r^\rho} > \tau - \epsilon, \tag{5.4.3}$$

for a sequence if values of r tending to infinity. Equations (5.4.2) and (5.4.3) together means

$$\tau = \limsup_{r\to\infty} \frac{\log M(r)}{r^\rho}.$$

□

Example 5.4.2. We show that the order of any polynomial is zero. Let $f(z) = a_0 + a_1 z + \cdots + a_n z^n$ be a polynomial. Then

$$\begin{aligned} |f(z)| &= |a_0 + a_1 z + \cdots + a_n z^n| \\ &\leq |a_0| + |a_1||z| + \cdots + |a_n||z|^n. \end{aligned}$$

Thus,

$$M(r) \leq |a_0| + |a_1|r + \cdots + |a_n|r^n \leq r^n(|a_0| + |a_1| + \cdots + |a_n|) = Br^n.$$

(taking $r \geq 1$. This choice is justified since ultimately $r \to \infty$.) where,

$$B_n = |a_0| + |a_1| + \cdots + |a_n|.$$

Hence,

$$\log M(r) \leq \log B + n \log r \leq logr + n \log r$$

taking r sufficiently large. Thus,

$$\log M(r) \leq (n+1)\log r$$

for large r. Hence,

$$\rho = \limsup_{r\to\infty} \frac{\log\log M(r)}{\log r} \leq \limsup_{r\to\infty} \frac{\log(n+1) + \log\log r}{\log r} = 0,$$

that is, $\rho \leq 0$. Also, we know that by definition, $\rho \geq 0$. Hence, $\rho = 0$.

From the above example, it is clear that the order of any constant function is zero. But, it does not mean that any zero order entire function is always a polynomial.

Example 5.4.3. The order of a transcendental entire function may also be zero. For example, if

$$f(z) = \sum_{n=0}^{\infty} \frac{z^n}{n^{n^{1+\delta}}}, \quad \delta > 0,$$

is a transcendental entire function having order $\rho = 0$ as we will soon see.

5.5 Order for sum and multiplications of entire functions

This section is dedicated to study the growth of the sum and multiplication of entire functions with reference to the orders of the parent functions.

Theorem 5.5.1. Let ρ_1 and ρ_2 be the orders of the entire functions f_1 and f_2 respectively. Then

1. order of $f_1 f_2 \leq \max\{\rho_1, \rho_2\}$
2. order of $f_1 \pm f_2 \leq \max\{\rho_1, \rho_2\}$

Proof. 1. Let $\phi(z) = f_1(z) f_2(z)$ and let ρ be the order of ϕ and let $\rho_1 \geq \rho_2$. Also, let

$$M(r, \phi) = \max_{|z|=r} |\phi(z)|, \quad M(r, f_1) = \max_{|z|=r} |f_1(z)|, \quad M(r, f_2) = \max_{|z|=r} |f_2(z)|.$$

Since ρ_1 and ρ_2 are orders of f_1 and f_2 respectively, we have for any given $\epsilon > 0$,

$$\log M(r, f_1) < r^{\rho_1 + \epsilon}$$
$$\log M(r, f_2) < r^{\rho_2 + \epsilon}$$

for all sufficiently large r. Now,

$$|\phi(z)| = |f_1(z)||f_2(z)| \leq M(r, f_1) \cdot M(r, f_2), \quad \forall\, z \text{ in } |z| \leq r.$$

Hence,

$$\begin{aligned} M(r, \phi) &\leq M(r, f_1) \cdot M(r, f_2) \\ \text{or,} \quad \log M(r, \phi) &\leq \log M(r, f_1) + \log M(r, f_2) \\ &\leq r^{\rho_1 + \epsilon} + r^{\rho_2 + \epsilon} \\ &\leq r^{\rho_1 + \epsilon} + r^{\rho_1 + \epsilon} \\ &\leq 2r^{\rho_1 + \epsilon} < r^{\epsilon} \cdot r^{\rho_1 + \epsilon} = r^{\rho_1 + 2\epsilon} \end{aligned}$$

for large r. Thus,

$$M(r, \phi) < r^{\rho_1 + 2\epsilon} \tag{5.5.1}$$

for sufficiently large r. Since $\epsilon > 0$ is arbitrary, so from (5.5.1), it follows that

$$\rho \leq \rho_1 = \max\{\rho_1, \rho_2\}.$$

Similarly, the result follows when $\rho_2 > \rho_1$.

2. Let $\psi(z) = f_1(z) \pm f_2(z)$ be of order ρ and let $\rho_1 \geq \rho_2$. Also, let $M(r, \psi) = \max_{|z|=r} |\psi(z)|$. Since ρ_1 and ρ_2 are orders of f_1 and f_2 respectively, we have for any given $\epsilon > 0$,

$$M(r, f_1) < \mathrm{e}^{r^{\rho_1 + \epsilon}}, \quad \text{and} \quad M(r, f_2) < \mathrm{e}^{r^{\rho_2 + \epsilon}},$$

for sufficiently large r. Now,

$$\begin{aligned} |\psi(z)| \leq |f_1(z)| + |f_2(z)| \Rightarrow M(r, \psi) &\leq M(r, f_1) + M(r, f_2) \\ &< \mathrm{e}^{r^{\rho_1 + \epsilon}} + \mathrm{e}^{r^{\rho_2 + \epsilon}} < 2\,\mathrm{e}^{r^{\rho_1 + \epsilon}} \end{aligned}$$

for sufficiently large r [since exponential function and r^n are both increasing]. Thus,

$$M(r, \psi) < \mathrm{e}^{r^{\rho_1 + 2\epsilon}} \Rightarrow \log M(r) < r^{\rho_1 + 2\epsilon}$$

for all large r. Since $\epsilon > 0$ is arbitrary, it follows that the order of ψ can't exceed ρ_1, that is, $\rho \leq \rho_1 = \max\{\rho_1, \rho_2\}$. Similarly, if $\rho_2 > \rho_1$, the result can be proved.

$\square$

Corollary 5.5.1. If ρ be the order of $f_1 \pm f_2$ and $\rho_1 \neq \rho_2$, then $\rho = \max\{\rho_1, \rho_2\}$.

Proof. Let $\rho_1 > \rho_2$. There exists a sequence $r_n \to \infty$ such that

$$M(r_n, f_1) > \mathrm{e}^{r_n^{\rho_1-\epsilon}}.$$

Hence,

$$M(r,\psi) \geq \mathrm{e}^{r_n^{\rho_1-\epsilon}} - \mathrm{e}^{r_n^{\rho_2+\epsilon}} = \exp(r_n^{\rho_1-\epsilon})\{1-\exp(r_n^{\rho_2+\epsilon}-r_n^{\rho_1-\epsilon})\} > \frac{1}{2}\exp(r_n^{\rho_1-\epsilon})$$

if ϵ is chosen so small that $\rho_2+\epsilon < \rho_1-\epsilon$ for sufficiently large n. Hence, $\rho \geq \rho_1$. But already we have, $\rho \leq \rho_1$. Thus, $\rho = \rho_1 = \max\{\rho_1,\rho_2\}$. □

Corollary 5.5.2. If ρ be the order of $f_1 f_2$ and $\rho_1 \neq \rho_2$, then $\rho = \max\{\rho_1, \rho_2\}$.

Remark 5.5.1. The result of the above two corollaries are not true if $\rho_1 = \rho_2$. For example, let $f_1(z) = \mathrm{e}^z$ and $f_2(z) = -\mathrm{e}^z$. Then the orders of f_1 and f_2 are both 1. But the order of f_1+f_2 is 0. Similarly, if $f_1(z) = \mathrm{e}^z$ and $f_2(z) = \mathrm{e}^{-z}$ then the orders of f_1 and f_2 are both 1 and the order of $f_1 f_2$ is 0.

Example 5.5.1. Let $P(z)$ be a polynomial of degree n, then the order of $\mathrm{e}^{P(z)}$ is n and the type of $\mathrm{e}^{P(z)}$ is the modulus of the coefficient of the highest degree term in $P(z)$.

Let $P(z) = a_0 + a_1 z + \cdots + a_n z^n$, $a_n \neq 0$ and $f(z) = \exp(a_0 + a_1 z + \cdots + a_n z^n)$. Then,

$$\begin{aligned} M(r,f) = \max_{|z|=r}|f(z)| &= \max_{|z|=r}|\exp(a_0+a_1z+\cdots+a_nz^n)| \\ &= \max_{|z|=r}|\,\mathrm{e}^{a_0}\cdot\mathrm{e}^{a_1z}\cdots\mathrm{e}^{a_nz^n}\,| \\ &= \max_{|z|=r}\{|\,\mathrm{e}^{a_0}\,|\cdot|\,\mathrm{e}^{a_1z}\,|\cdots|\,\mathrm{e}^{a_nz^n}\,|\}. \end{aligned} \tag{5.5.2}$$

Let us find $\max_{|z|=r}|\,\mathrm{e}^{a_m z^m}\,|$. Let $a_m = t\,\mathrm{e}^{i\phi}$ and $z = r\,\mathrm{e}^{i\theta}$. Then

$$a_m z^m = t\,\mathrm{e}^{i\phi}\cdot r^m\,\mathrm{e}^{im\theta} = tr^m\,\mathrm{e}^{i(m\theta+\phi)}.$$

Hence,

$$\begin{aligned} \max_{|z|=r}|\,\mathrm{e}^{a_m z^m}\,| &= \max_{\theta}|\exp\{tr^m\,\mathrm{e}^{i(m\theta+\phi)}\}| \\ &= \max_{\theta}|\exp(tr^m\{\cos(m\theta+\phi)+i\sin(m\theta+\phi)\})| \\ &= \max_{\theta}|\exp(tr^m\cos(m\theta+\phi))| = \exp(tr^m) = \exp\{|a_m|r^m\}. \end{aligned}$$

Hence, from (5.5.2),

$$M(r,f) = \mathrm{e}^{|a_0|}\cdot\mathrm{e}^{|a_1|r}\cdots\mathrm{e}^{|a_n|r^n} = \mathrm{e}^{|a_0|+|a_1|r+\cdots+|a_n|r^n}.$$

Hence,

$$\log M(r,f) = |a_0| + |a_1|r + \cdots + |a_n|r^n.$$

Hence,

$$\begin{aligned}
\rho = \limsup_{r\to\infty} \frac{\log\log M(r,f)}{\log r} &= \limsup_{r\to\infty} \frac{\log(|a_0| + |a_1|r + \cdots + |a_n|r^n)}{\log r} \\
&= \limsup_{r\to\infty} \frac{\log r^n \left(\frac{|a_0|}{r^n} + \frac{|a_1|}{r^{n-1}} + \cdots + |a_n|\right)}{\log r} \\
&= \limsup_{r\to\infty} \frac{n\log r + \log \left(\frac{|a_0|}{r^n} + \frac{|a_1|}{r^{n-1}} + \cdots + |a_n|\right)}{\log r} \\
&= \limsup_{r\to\infty} \left(n + \frac{\log \left(\frac{|a_0|}{r^n} + \frac{|a_1|}{r^{n-1}} + \cdots + |a_n|\right)}{\log r}\right) = n.
\end{aligned}$$

And,

$$\begin{aligned}
\tau = \limsup_{r\to\infty} \frac{\log M(r)}{r^\rho} &= \limsup_{r\to\infty} \frac{|a_0| + |a_1|r + \cdots |a_n|r^n}{r^n} \\
&= \limsup_{r\to\infty} \left\{ \frac{|a_0|}{r^n} + \frac{|a_1|}{r^{n-1}} + \cdots + |a_n| \right\} = |a_n|.
\end{aligned}$$

5.6 Order and coefficients in terms of Taylor's Coefficients

Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be an entire function with order ρ and type τ. We now state the formulae for order and type of f in terms of the Taylor's coefficients.

Theorem 5.6.1. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be an entire function of finite order ρ. Then

$$\rho = \limsup_{n\to\infty} \frac{\log n}{\log\left(\frac{1}{|a_n|^{1/n}}\right)}.$$

Theorem 5.6.2. Let $f(z) = \sum_{n=0}^{\infty} a_n z^n$ be an entire function of finite order ρ. Then

$$\tau = \frac{1}{\mathrm{e}\,\rho} \limsup_{n\to\infty} n|a_n|^{\rho/n}.$$

Example 5.6.1. Let

$$f(z) = \sum_{n=0}^{\infty} \frac{z^n}{(n!)^\alpha}, \quad \alpha > 0.$$

Let ρ and τ be the order and type respectively for f. Here,

$$a_n = \frac{1}{(n!)^\alpha} \Rightarrow \log\left(\frac{1}{|a_n|}\right) = \alpha \log n!$$

Hence,

$$\begin{aligned}
\rho = \limsup_{r\to\infty} \frac{n\log n}{\alpha \log n!} &= \frac{1}{\alpha}\limsup_{r\to\infty} \frac{(n+1)\log(n+1) - n\log n}{\log(n+1)! - \log n!} \\
&= \frac{1}{\alpha}\limsup_{r\to\infty} \frac{(n+1)\log n(1+1/n) - n\log n}{\log(n+1)} \\
&= \frac{1}{\alpha}\limsup_{r\to\infty} \frac{(n+1)\log n + (n+1)\log(1+1/n) - n\log n}{\log(n+1)} \\
&= \frac{1}{\alpha}\limsup_{r\to\infty} \frac{\log n + \log(1+1/n)^n + \log(1+1/n)}{\log n + \log(1+1/n)} \\
&= \frac{1}{\alpha}\limsup_{r\to\infty} \frac{1 + \frac{\log(1+1/n)^n}{\log n} + \frac{\log(1+1/n)}{\log n}}{1 + \frac{\log(1+1/n)}{\log n}} = \frac{1}{\alpha}.
\end{aligned}$$

Hence $\rho = 1/\alpha$. Thus,

$$\begin{aligned}
\tau = \frac{\alpha}{\mathrm{e}}\limsup_{r\to\infty} n\left(\frac{1}{(n!)^\alpha}\right)^{\frac{1}{n\alpha}} &= \frac{\alpha}{\mathrm{e}}\limsup_{r\to\infty} n\left(\frac{1}{n!}\right)^{\alpha\cdot\frac{1}{n\alpha}} \\
&= \frac{\alpha}{\mathrm{e}}\limsup_{r\to\infty} n\left(\frac{1}{n!}\right)^{\frac{1}{n}} \\
&= \frac{\alpha}{\mathrm{e}}\limsup_{r\to\infty} \left(\frac{n^n}{n!}\right)^{\frac{1}{n}} = \frac{\alpha}{\mathrm{e}}\cdot\alpha = \alpha.
\end{aligned}$$

Hence, $\tau = \alpha$.

Exercise 5.6.1. 1. Find the order and type of the following functions

(a) e^z

(b) $\mathrm{e}^{z^4} \cdot z^4$

(c) $\sin z$

2. Find the order and type of the following functions

$$\sum_{n=0}^{\infty}\left(\frac{z}{n}\right)^n, \quad \sum_{n=0}^{\infty}\left(\frac{\log n}{n}\right)^{n/a} z^n, \quad a > 0.$$

Few Probable Questions

1. Show that for a transcendental entire function f with maximum modulus function $M(r)$,

$$\liminf_{r\to\infty} \frac{\log M(r)}{\log r} = \infty.$$

2. If for an entire function f with maximum modulus $M(r)$, the relation

$$\lim_{r\to\infty} \frac{M(r)}{r^k} < \infty,$$

holds, then show that f is a polynomial of degree at most k.

3. Show that the order ρ of an entire function with maximum modulus function $M(r)$ is given by

$$\rho = \limsup_{r\to\infty} \frac{\log\log M(r)}{\log r}.$$

Hence find the order of $\cos z$.

4. Show that any polynomial is of order zero. Is the converse true? Justify.

5. Define the type of an entire function f having finite non-zero order ρ. Also, show that

$$\tau = \limsup_{r\to\infty} \frac{\log M(r)}{r^\rho}.$$

6. Show that for two entire functions f_1 and f_2 of orders ρ_1 and ρ_2 respectively,

$$\text{order of } f_1 f_2 \leq \max\{\rho_1, \rho_2\}.$$

7. Show that for two entire functions f_1 and f_2 of orders ρ_1 and ρ_2 respectively,

$$\text{order of } f_1 \pm f_2 \leq \max\{\rho_1, \rho_2\}.$$

Unit 6

Course Structure

- Distribution of zeros of entire functions.

6.1 Introduction

The zeroes of entire functions play an important role in determining their growth rates. We will start off with the Jensen's theorems for analytic functions. In complex analysis, Jensen's formula, introduced by Johan Jensen (1899), relates the average magnitude of an analytic function on a circle with the number of its zeros inside the circle. It forms an important statement in the study of entire functions as we will soon come to see.

Objectives

After reading this unit, you will be able to

- study Jensen's theorems and related results

6.2 Distribution of zeros of analytic functions

Theorem 6.2.1. (Jensen's theorem) Let f be analytic on $|z| \leq R$, $f(0) \neq 0$ and $f(z) \neq 0$ on $|z| = R$. If $a_1, a_2, \ldots, a_n$ are the zeros of f in $|z| < R$, multiple zeros being repeated, and $|a_i| = r_i$, then

$$\log \frac{R^n}{r_1 r_2 \cdots r_n} = \frac{1}{2\pi} \int_0^{2\pi} \log |f(R\,\mathrm{e}^{i\theta})| d\theta - \log |f(0)|.$$

Proof. Let

$$\begin{aligned} \phi(z) &= f(z) \cdot \frac{R^2 - \overline{a_1} z}{R(z - a_1)} \cdot \frac{R^2 - \overline{a_2} z}{R(z - a_2)} \cdots \frac{R^2 - \overline{a_n} z}{R(z - a_n)} \\ &= f(z) \prod_{k=1}^{n} \frac{R^2 - \overline{a_k} z}{R(z - a_k)}. \end{aligned} \tag{6.2.1}$$

The zeroes of the denominator of ϕ are also the zeros of f of the same order. Hence the zeros of f cancel the poles a_n in the product and so ϕ is analytic on $|z| \leq R$. Also, $\phi(z) \neq 0$ on $|z| \leq R$. Since

$$R^2 - \overline{a_k}z = 0 \Rightarrow z = \frac{R^2}{\overline{a_k}} \Rightarrow |z| = \frac{R^2}{|\overline{a_k}|} = \frac{R^2}{|a_k|} > R$$

since $|a_k| < R$ for all $k = 1, 2, \ldots, n$. Thus, any zero of $\phi(z)$ lies outside the circle $|z| = R$. So, ϕ has neither zeros nor poles in $|z| \leq R$. Thus, the function $\log \phi(z)$ is analytic in $|z| \leq R$. Thus, by Cauchy's Integral theorem, we have

$$\log \phi(0) = \frac{1}{2\pi i} \int_{|z|=R} \frac{1}{z} \log \left(f(z) \prod_{k=1}^{n} \frac{R^2 - \overline{a_k}z}{R(z - a_k)} \right) dz. \tag{6.2.2}$$

On $|z| = R$, we have, $z = R\,\mathrm{e}^{i\theta}$, $\theta \in [0, 2\pi]$, which implies that $dz = R\,\mathrm{e}^{i\theta}\, i d\theta$. Also,

$$|\phi(z)| = |f(z)| \left| \frac{R^2 - \overline{a_1}z}{R(z - a_1)} \right| \cdot \left| \frac{R^2 - \overline{a_2}z}{R(z - a_2)} \right| \cdots \left| \frac{R^2 - \overline{a_n}z}{R(z - a_n)} \right|.$$

On $|z| = R$, we have

$$\left| \frac{R^2 - \overline{a_k}z}{R(z - a_k)} \right| = \left| \frac{z\overline{z} - \overline{a_k}z}{R(z - a_k)} \right| = \frac{|z|}{R} \left| \frac{\overline{z} - \overline{a_k}}{z - a_k} \right| = \left| \frac{\overline{z - a_k}}{z - a_k} \right| = 1$$

and thus,

$$|\phi(z)| = |f(z)|, \quad \text{on } \ |z| = R.$$

The equation (6.2.2) changes to

$$\begin{aligned} \log \phi(0) &= \frac{1}{2\pi i} \int_0^{2\pi} \frac{1}{R\,\mathrm{e}^{i\theta}} \log \left(f(R\,\mathrm{e}^{i\theta}) \prod_{k=1}^{n} \frac{R^2 - \overline{a_k}z}{R(z - a_k)} \right) R\,\mathrm{e}^{i\theta}\, i d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} \left[\log f(R\,\mathrm{e}^{i\theta}) + \sum_{k=1}^{n} \log \left(\frac{R^2 - \overline{a_k}z}{R(z - a_k)} \right) \right] d\theta. \end{aligned} \tag{6.2.3}$$

Taking real parts of equation (6.2.3), we get, by using the conditions deduced in the previous discussions,

$$\log |\phi(0)| = \frac{1}{2\pi} \int_0^{2\pi} \log |f(R\,\mathrm{e}^{i\theta})| d\theta. \tag{6.2.4}$$

Since equation (6.2.1) gives

$$|\phi(0)| = |f(0)| \prod_{k=1}^{n} \frac{R}{|a_k|} = |f(0)| \prod_{k=1}^{n} \frac{R}{r_k} = |f(0)| \frac{R^n}{r_1 r_2 \cdots r_n}$$

Thus,

$$\log |\phi(0)| = \log |f(0)| + \log \frac{R^n}{r_1 r_2 \cdots r_n}$$

and thus, equation (6.2.4) gives

$$\log \frac{R^n}{r_1 r_2 \cdots r_n} = \frac{1}{2\pi} \int_0^{2\pi} \log |f(R\,\mathrm{e}^{i\theta})| d\theta - \log |f(0)|.$$

$\square$

Note 6.2.1. Jensen's theorem can also be written as

$$\log|f(0)| + \sum_{i=1}^{n} \log\frac{R}{|a_k|} = \frac{1}{2\pi}\int_0^{2\pi} \log|f(R\,\mathrm{e}^{i\theta})|d\theta.$$

Theorem 6.2.2. **(Jensen's Inequality)** Let f be analytic on $|z| \le R$, $f(0) \ne 0$ and $f(z) \ne R$ on $|z| = R$. If $a_1, a_2, \ldots, a_n$ are the zeros of f in $|z| < R$, multiple zeros being repeated, and $|a_i| = r_i$, then

$$\frac{R^n|f(0)|}{r_1 r_2 \cdots r_n} \le M(R).$$

Proof. Let

$$\begin{aligned}\phi(z) &= f(z)\cdot\frac{R^2-\overline{a_1}z}{R(z-a_1)}\cdot\frac{R^2-\overline{a_2}z}{R(z-a_2)}\cdots\frac{R^2-\overline{a_n}z}{R(z-a_n)}\\ &= f(z)\prod_{k=1}^{n}\frac{R^2-\overline{a_k}z}{R(z-a_k)}.\end{aligned}$$

The zeroes of the denominator of ϕ are also the zeros of f of the same order. Hence the zeros of f cancel the poles a_n in the product and so ϕ is analytic on $|z| \le R$. Also,

$$|\phi(z)| = |f(z)|\left|\frac{R^2-\overline{a_1}z}{R(z-a_1)}\right|\cdot\left|\frac{R^2-\overline{a_2}z}{R(z-a_2)}\right|\cdots\left|\frac{R^2-\overline{a_n}z}{R(z-a_n)}\right|.$$

On $|z| = R$, we have

$$\left|\frac{R^2-\overline{a_k}z}{R(z-a_k)}\right| = \left|\frac{z\overline{z}-\overline{a_k}z}{R(z-a_k)}\right| = \frac{|z|}{R}\left|\frac{\overline{z}-\overline{a_k}}{z-a_k}\right| = \left|\frac{\overline{z-a_k}}{z-a_k}\right| = 1$$

and thus,

$$|\phi(z)| = |f(z)|, \quad \text{on } |z| = R.$$

By Maximum modulus theorem, $|\phi(z)| \le M(R)$ for $|z| \le R$. In particular, $|\phi(0)| \le M(R)$, that is,

$$|f(0)|\left|\frac{R}{-a_1}\right|\cdots\left|\frac{R}{-a_n}\right| \le M(R) \Rightarrow \frac{R^n|f(0)|}{r_1 r_2\cdots r_n} \le M(R).$$

□

Definition 6.2.1. Let f be analytic on $|z| \le R$, with zeros at the points $a_1, a_2, \ldots,$ arranged in the order of non-decreasing modulus, multiple zeros being repeated. We define the function $n(r)$ as the number of zeros of f in $|z| \le r$, $r \le R$. Evidently, $n(r)$ is a non-negative, non-decreasing function of r which is constant in any interval which does not contain the modulus of a zero of f. Observe that if $f(0) \ne 0$, then $n(r) = 0$ for $r < |a_1|$. Also, $n(r) = n$ for $|a_n| \le r < |a_{n+1}|$.

We will rewrite Jensen's inequality in terms of $n(r)$ as follows.

Theorem 6.2.3. Let f be analytic on $|z| \le R$, $f(0) \ne 0$. Let its zeros, arranged in order of non-decreasing modulus be $a_1, a_2, \ldots$, multiple zeros being repeated according to their multiplicities. If $|a_n| \le r < |a_{n+1}|$, then

$$\int_0^x \frac{n(x)}{x}dx \le \log M(r) - \log|f(0)|.$$

Proof. Let $|a_i| = r_i$, $i = 1, 2, \ldots,$ and r be a positive number such that $r_N \le r < r_{N+1}$, $(r \le R)$. Let $x_1, x_2, \ldots, x_m$ be the distinct numbers of the set $E = \{r_1, r_2, \ldots, r_N\}$ so that $x_1 = r_1, \ldots, x_m = r_N$. Suppose x_i is repeated p_i times in E. Then $p_1 + \cdots + p_m = N$. Also, $s_i = p_1 + \cdots + p_i$, $i = 1, 2, \ldots, m$. We consider two cases.

Case I: Let $r_N < r$. Then,

$$\begin{aligned}
\int_0^x \frac{n(x)}{x}dx &= \lim_{\epsilon\to 0}\left\{\int_{x_1}^{x_2-\epsilon}\frac{n(x)}{x}dx + \int_{x_2}^{x_3-\epsilon}\frac{n(x)}{x}dx + \cdots + \int_{x_{m-1}}^{x_m-\epsilon}\frac{n(x)}{x}dx\right\} + \int_{x_m}^{r}\frac{n(x)}{x}dx \\
&\qquad\qquad \left[\text{since } \int_0^{x_1-\epsilon}\frac{n(x)}{x}dx = 0 \text{ as } n(x) = 0 \text{ when } 0 \le x < x_1\right] \\
&= \lim_{\epsilon\to 0}\left\{\int_{x_1}^{x_2-\epsilon}\frac{s_1}{x}dx + \int_{x_2}^{x_3-\epsilon}\frac{s_2}{x}dx + \cdots + \int_{x_{m-1}}^{x_m-\epsilon}\frac{s_{m-1}}{x}dx\right\} + \int_{x_m}^{r}\frac{N}{x}dx \\
&= \lim_{\epsilon\to 0}\left\{[s_1\log x]_{x_1}^{x_2-\epsilon} + [s_2\log x]_{x_2}^{x_3-\epsilon} + \cdots + [s_{m-1}\log x]_{x_{m-1}}^{x_m-\epsilon}\right\} + [N\log x]_{x_N}^{r} \\
&= \lim_{\epsilon\to 0}[s_1\{\log(x_2-\epsilon) - \log x_1\} + s_2\{\log(x_3-\epsilon) - \log x_2\} + \cdots \\
&\qquad + s_{m-1}\{\log(x_m-\epsilon) - \log x_{m-1}\}] + N(\log r - \log r_N) \\
&= s_1(\log x_2 - \log x_1) + s_2(\log x_3 - \log x_2) + \cdots + s_{m-1}(\log x_m - \log x_{m-1}) \\
&\qquad + N(\log r - \log r_N) \\
&= p_1\log x_2 - p_1\log x_1 + (p_1+p_2)\log x_3 - (p_1+p_2)\log x_2 + \cdots \\
&\quad + (p_1+\cdots+p_{m-1})\log x_m - (p_1+\cdots+p_{m-1})\log x_{m-1} \\
&\qquad + N\log r - (p_1+\cdots+p_m)\log x_m \\
&= N\log r - (p_1\log x_1 + p_2\log x_2 + \cdots + p_m\log x_m) \\
&= \log r^N - \log x_1^{p_1}x_2^{p_2}\cdots x_m^{p_m} \\
&= \log\frac{r^N}{x_1^{p_1}x_2^{p_2}\cdots x_m^{p_m}} = \log\frac{r^N}{r_1r_2\cdots r_N}.
\end{aligned}$$

Case II: Let $r_N = r$. Then as before,

$$\begin{aligned}
\int_0^r \frac{n(x)}{x}dx &= \lim_{\epsilon\to 0}\left\{\int_{x_1}^{x_2-\epsilon}\frac{s_1}{x}dx + \int_{x_2}^{x_3-\epsilon}\frac{s_2}{x}dx + \cdots + \int_{x_{m-1}}^{x_m-\epsilon}\frac{s_{m-1}}{x}dx\right\} \\
&= \sum_{i=1}^{m-1} s_i(\log x_{i+1} - \log x_i) + s_m(\log r - \log r_N) \qquad [\text{ since } r = r_N] \\
&= \log\frac{r^N}{r_1\cdots r_N} \qquad (\text{ Proceeding as in Case I})
\end{aligned}$$

Thus, in any case,

$$\int_0^r \frac{n(x)}{x}dx = \log\frac{r^N}{r_1r_2\cdots r_N}$$

But Jensen's inequality gives us

$$\frac{r^N}{r_1r_2\cdots r_N} \le \frac{M(r)}{|f(0)|}$$

Hence,

$$\int_0^r \frac{n(x)}{x}dx = \log\frac{r^N}{r_1\cdots r^N} \le \log M(r) - \log|f(0)|$$

□

Note 6.2.2. Jensen's inequality is also true for entire functions where $R \to \infty$.

6.3 Distribution of zeros of entire functions

Theorem 6.3.1. Let f be an entire function with finite order ρ and $f(0) \neq 0$. Then $n(r) = O\left(r^{\rho+\epsilon}\right)$ for any $\epsilon > 0$ and for sufficiently large values of r.

Proof. By Jensen's inequality,

$$\int_0^r \frac{n(x)}{x} dx \leq \log M(r) - \log |f(0)| \tag{6.3.1}$$

Replacing r by $2r$ in (6.3.1) we get,

$$\int_0^{2r} \frac{n(x)}{x} dx \leq \log M(2r) - \log |f(0)| \tag{6.3.2}$$

Since ρ is the order of f we have for each $\epsilon > 0$,

$$\log M(2r) < (2r)^{(\rho+\epsilon)} = 2^{(\rho+\epsilon)} \cdot r^{(\rho+\epsilon)} = kr^{(\rho+\epsilon)}$$

for all large values of r, where $k = 2^{(\rho+\epsilon)}$ = a constant. Hence from (6.3.2)

$$\int_0^{2r} \frac{n(x)}{x} dx < Ar^{(\rho+\epsilon)}, \forall \text{ large } r, A = \text{ constant independent of } r.$$

Now, since $n(x)$ is non-negative and non-decreasing function of x we have,

$$\int_r^{2r} \frac{n(x)}{x} dx \leq \int_0^{2r} \frac{n(x)}{x} dx < Ar^{(\rho+\epsilon)}$$

and also

$$\int_r^{2r} \frac{n(x)}{x} dx \geq \int_r^{2r} \frac{n(x)}{x} dx = n(r) \int_r^{2r} \frac{dx}{x} = n(r) \log 2$$

Hence,

$$n(r) \log 2 \leq \int_r^{2r} \frac{n(x)}{x} dx < Ar^{(\rho+\epsilon)};$$

that is,

$$n(r) < \frac{A}{\log 2} r^{(\rho+\epsilon)}, \text{ for all large } r.$$

Hence, $n(r) = O\left(r^{(\rho+\epsilon)}\right)$ □

Few Probable Questions

1. State and prove Jensen's theorem.

2. State and prove Jensen's inequality.

3. For a function f analytic in $|z| \leq R$, $f(0) \neq 0$, if $a_1, a_2, \ldots$ are its zeros, arranged in the order of non-decreasing modulus, multiple zeros repeated according to their multiplicities, show that, for $|a_n| \leq r < |a_{n+1}|$,
$$\int_0^x \frac{n(x)}{x} dx \leq \log M(r) - \log|f(0)|.$$

4. Show that for an entire function f of finite order ρ and $f(0) \neq 0$, $n(r) = O(r^{\rho+\epsilon})$ for any $\epsilon > 0$ and for sufficiently large values of r.

5. If f is an entire function of order $\rho < \infty$ and has an infinity of zeros with $f(0) \neq 0$, then show that for $\epsilon > 0$, there exists R_0 such that for $R \geq R_0$
$$n\left(\frac{R}{3}\right) \leq \frac{1}{\log 2} \cdot \log\left\{\frac{\exp R^{\rho+\epsilon}}{|f(0)|}\right\},$$
where $n(R)$ denotes the number of zeros of f in $|z| \leq R$.

Unit 7

Course Structure

- The exponent of convergence of zeros.

7.1 Introduction

Let f be an entire function with the zeros $z_1, z_2, \ldots$, arranged in order of non-decreasing modulus. We associate with this sequence of zeros a number ρ_1 defined by the equation

$$\rho_1 = \lim_{n\to\infty} \frac{\log n}{\log r_n}$$

$\left[\text{or}\ , \rho_1 = \lim\limits_{r\to\infty} \dfrac{\log n\,(r)}{\log r}\right]$, where $|z_n| = r_n$. This number ρ_1 is called the convergence exponent or exponent of convergence of the zeros of function f.

Objectives

After reading this unit, you will be able to

- define convergence exponent of the zeros of an entire function and deduce various related results

7.2 Convergence exponent of zeros of entire functions

Theorem 7.2.1. Let f be an entire function with zeros $z_1, z_2, \ldots$, arranged in order of non-decreasing modulus and $|z_n| = r_n$. If the convergence exponent ρ_1 of the zeros of f be finite, then the series $\sum\limits_{n=1}^{\infty} \dfrac{1}{{r_n}^{\alpha}}$ converges when $\alpha > \rho_1$ and diverges when $\alpha < \rho_1$. If ρ_1 is infinite, the above series diverges for all positive values of α.

Proof. Let ρ_1 be finite and $\alpha > \rho_1$. Then $\rho_1 < \frac{1}{2}(\rho_1 + \alpha)$. Hence, from the definition of ρ_1 we have, $\frac{\log n}{\log r_n} < \frac{1}{2}(\rho_1 + \alpha)$ for all large n. Hence,

$$\log n < \frac{1}{2}(\rho_1 + \alpha)\log r_n = \log {r_n}^{\frac{1}{2}(\rho_1+\alpha)}$$

that is,

$$n < {r_n}^{\frac{1}{2}(\rho_1+\alpha)}, \text{ or, } n^{\frac{2}{\rho_1+\alpha}} < r_n,$$

$$r_n^\alpha > n^{\frac{2\alpha}{\rho_1+\alpha}} = n^{1+\frac{\alpha-\rho_1}{\alpha+\rho_1}} = n^{1+p}, \text{ where } p = \frac{\alpha - \rho_1}{\alpha + \rho_1} > 0$$

Hence, $\frac{1}{{r_n}^\alpha} < \frac{1}{n^{1+p}}$ for large n. Hence, $\sum_{n=1}^{\infty} \frac{1}{{r_n}^\alpha}$ converges.

Next, let $\alpha < \rho_1$. Then, $\frac{\log n}{\log r_n} > \alpha$ for a sequence of values of n, tending to ∞, that is, $\log n > \alpha \log r_n$ $= \log {r_n}^\alpha$. Hence, $n > {r_n}^\alpha$, or $\frac{1}{{r_n}^\alpha} > \frac{1}{n}$ for a sequence of values of n tending to infinity. Let N be such a value of n for which the above inequality holds, that is, $\frac{1}{{r_N}^\alpha} > \frac{1}{N}$ and let m be the least integer greater than $\frac{N}{2}$. Then, since r_n is non-decreasing with n, we have

$$\sum_{N-m}^{N} = \frac{1}{{r_{N-m}}^\alpha} + \frac{1}{{r_{N-m+1}}^\alpha} + \cdots + \frac{1}{{r_N}^\alpha} \geq \frac{1}{{r_N}^\alpha} + \cdots + \frac{1}{{r_N}^\alpha} = \frac{m+1}{{r_N}^\alpha} > \frac{m}{N} > \frac{1}{2}$$

Since these are values of N as large as we please, by Cauchy's principle of convergence, the series $\sum_{n=1}^{\infty} \frac{1}{{r_n}^\alpha}$ diverges.

If ρ_1 is infinite, then for any value of α, $\frac{\log n}{\log r_n} > \alpha$ for a sequence of values of n tending to infinity, that is, $\log n > \log {r_n}^\alpha$, that is, $n > {r_n}^\alpha$ for a sequence of values of n tending to infinity, from which we may similarly conclude that the series $\sum_{n=1}^{\infty} \frac{1}{{r_n}^\alpha}$ diverges for any positive α. □

Note 7.2.1. We may also define convergence exponent ρ_1 as the $g.l.b$ of the positive numbers α for which the series $\sum_{n=1}^{\infty} \frac{1}{{r_n}^\alpha}$ is convergent. For an entire function with no zeros we define $\rho_1 = 0$ and if the series $\sum_{n=1}^{\infty} \frac{1}{{r_n}^\alpha}$ diverges for all positive α, then $\rho_1 = \infty$.

Note 7.2.2. If ρ_1 is finite, the series $\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\rho_1}}$ may be either convergent or divergent. For example, if $r_n = n$ we have, $\rho_1 = \limsup_{n\to\infty} \frac{\log n}{\log r_n} = \limsup_{n\to\infty} \frac{\log n}{\log n} = 1$ and $\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\rho_1}} = \sum_{n=1}^{\infty} \frac{1}{n}$ diverges.

Again, if $r_n = n(\log n)^2$ we have,

$$\rho_1 = \limsup_{n\to\infty} \frac{\log n}{\log r_n} = \limsup_{n\to\infty} \frac{\log n}{\log n + 2\log\log n} = \limsup_{n\to\infty} \frac{1}{1 + 2\frac{\log\log n}{\log n}} = 1$$

and $\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\rho_1}} = \sum_{n=1}^{\infty} \frac{1}{n\left(\log n\right)^2}$ converges.

Theorem 7.2.2. If f is an entire function with finite order ρ and $r_1, r_2, \ldots$ are the moduli of the zeros of f, then $\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\alpha}}$ converges if $\alpha > \rho$.

Proof. Let β be a number such that $\rho < \beta < \alpha$. Since $n\left(r\right) = O\left(r^{\rho+\epsilon}\right)$ for any $\epsilon > 0$. We have, $n\left(r\right) < Ar^{\beta}$ for all large r, A being a constant.

Putting $r = r_n$, n being large, this inequality gives $n < A{r_n}^{\beta}$, that is, ${r_n}^{\beta} > \frac{n}{A}$, or, $r_n > \frac{n^{\frac{1}{\beta}}}{A^{\frac{1}{\beta}}}$ or, ${r_n}^{\alpha} > \frac{n^{\frac{\alpha}{\beta}}}{A^{\frac{\alpha}{\beta}}} = Bn^{\frac{\alpha}{\beta}}$, $B =$ constant. Hence, $\frac{1}{{r_n}^{\alpha}} < \frac{B_1}{n^{\frac{\alpha}{\beta}}}$ for large n, $B_1 =$ constant. Since $\frac{\alpha}{\beta} > 1$, it follows that $\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\alpha}}$ converges. □

Corollary 7.2.1. Since convergence exponent ρ_1 of the zeros of f is the lower bound of the positive numbers α for which $\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\alpha}}$ is convergent, it follows from the above theorem that $\rho_1 \leq \rho$.

Remark 7.2.1. We can prove that the result $\rho_1 \leq \rho$ without using the last theorem.

Proof. We have

$$\begin{aligned}
\rho_1 &= \limsup_{n\to\infty} \frac{\log n}{\log r_n} \\
&= \limsup_{r\to\infty} \frac{\log n\left(r\right)}{\log r} \leq \limsup_{r\to\infty} \frac{\log\left(Ar^{\rho+\epsilon}\right)}{\log r} \\
&= \limsup_{r\to\infty} \frac{\log A + \left(\rho+\epsilon\right)\log r}{\log r}, \quad A = \text{ constant} \\
&= \limsup_{r\to\infty} \left\{\rho + \epsilon + \frac{\log A}{\log r}\right\} \\
&= \rho + \epsilon, \text{ for any } \epsilon > 0
\end{aligned}$$

Hence, $\rho_1 \leq \rho$. □

Note 7.2.3. Convergence exponent may be 0 or ∞. For example, if $r_n = e^n$, then $\rho_1 = \limsup_{n\to\infty} \frac{\log n}{n} = 0$. Also,if $r_n = \log n$, then

$$\rho_1 = \limsup_{n\to\infty} \frac{\log n}{\log\log n} = \infty$$

We may have, $\rho_1 < \rho$. For example, if $f\left(z\right) = e^z$, then $\rho = 1$, $\rho_1 = 0$, since there are no zeros of f. For $\sin z$ or $\cos z$, $\rho = \rho_1 = 1$.

Theorem 7.2.3. Let f be an entire function of finite order. If convergence exponent ρ_1 of the zeros of f is greater than zero,then f has infinite number of zeros.

Proof. If possible, let f has finite number of zeros. Let $r_1, r_2, \ldots, r_N$ be the moduli of the zeros of f arranged in non-decreasing order. The series $\sum_{n=1}^{N} \frac{1}{{r_n}^{\alpha}}$, being a series of finite number of terms, converges for every positive value of α. It follows $\rho_1 = 0$ which contradicts our assumption. Hence f contains infinite number of zeros. $\square$

Note 7.2.4. For an entire function with finite number of zeros, $\rho_1 = 0$.

Example 7.2.1. Find the convergence exponent of the zeros of $\cos z$.

Solution. The zeros of $\cos z$ are $\frac{\pi}{2}, \frac{-\pi}{2}, \frac{3\pi}{2}, \frac{-3\pi}{2}, \ldots$. Now,

$$\begin{aligned}\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\alpha}} &= \left(\frac{2}{\pi}\right)^{\alpha} + \left(\frac{2}{\pi}\right)^{\alpha} + \left(\frac{2}{\pi}\right)^{\alpha} \cdot \frac{1}{3^{\alpha}} + \left(\frac{2}{\pi}\right)^{\alpha} \cdot \frac{1}{3^{\alpha}} + \cdots \\ &= 2\left(\frac{2}{\pi}\right)^{\alpha}\left(1 + \frac{1}{3^{\alpha}} + \frac{1}{5^{\alpha}} + \cdots\right)\end{aligned}$$

The series $\frac{1}{1^{\alpha}} + \frac{1}{3^{\alpha}} + \frac{1}{5^{\alpha}} + \cdots$ converges when $\alpha > 1$ and diverges when $\alpha < 1$. Hence, the lower bound of positive numbers α for which the series $\sum_{n=1}^{\infty} \frac{1}{{r_n}^{\alpha}}$ converges us 1. Hence, $\rho_1 = 1$.

Aliter: The zeros of $\cos z$ are $(2n+1)\frac{\pi}{2}$, $n = 0, \pm 1, \pm 2, \ldots$. Let $z_1 = \frac{\pi}{2}$, $z_1' = -\frac{\pi}{2}$, $z_2 = \frac{3\pi}{2}$, $z_2' = -\frac{3\pi}{2}, \ldots, z_n = (2n-1)\frac{\pi}{2}$, $z_n' = -(2n-1)\frac{\pi}{2}, \ldots$. Hence, $r_1 = |z_1| = |z_1'| = \frac{\pi}{2}$, $r_2 = |z_2| = |z_2'| = \frac{3\pi}{2}, \ldots, r_n = |z_n| = |z_n'| = (2n-1)\frac{\pi}{2}, \ldots$

Hence,

$$\begin{aligned}\rho_1 &= \limsup_{n\to\infty} \frac{\log n}{\log r_n} \\ &= \limsup_{n\to\infty} \frac{\log n}{\log(2n-1)\pi/2} \\ &= \limsup_{n\to\infty} \frac{\log n}{\log(2n-1) + \log \pi/2} \\ &= \limsup_{n\to\infty} \frac{\log n}{\log n(2-1/n) + \log \pi/2} \\ &= \limsup_{n\to\infty} \frac{1}{1 + \frac{\log(2-1/n)}{\log n} + \frac{\log \pi/2}{\log n}} = 1.\end{aligned}$$

$\blacksquare$

Theorem 7.2.4. If f is an entire function having no zeros, then f is of the form $f(z) = \mathrm{e}^{g(z)}$, where $g(z)$ is an entire function.

Proof. Since $f(z) \neq 0$ for all $z \in \mathbb{C}$, then the function

$$h(z) = \frac{f'(z)}{f(z)} \tag{7.2.1}$$

is also an entire function. Integrating (7.2.1) along any path joining the two points z_0 and z, we get

$$\int_{z_0}^{z} h(z)dz = \int_{z_0}^{z} \frac{f'(z)}{f(z)}dz = \log f(z) - \log f(z_0),$$

where principal branch of logarithm is taken. Hence,

$$\log f(z) = \log f(z_0) + \int_{z_0}^{z} h(z)dz. \tag{7.2.2}$$

The right hand side of equation (7.2.2) is an entire function, say $g(z)$. Hence,

$$\log f(z) = g(z) \Rightarrow f(z) = \mathrm{e}^{g(z)},$$

where $g(z)$ is an entire function. □

Exercise 7.2.1. Find the convergence exponent of the zeros of $\sin z$.

Few Probable Questions

1. Define convergence exponent ρ_1 of the zeros of an entire function. Show that if r_i be the moduli of the zeros of an entire function f, arranged in order of non-decreasing modulus, then the series $\sum_{n=1}^{\infty} \frac{1}{r_n^{\alpha}}$ converges for $\alpha > \rho_1$ and diverges for $\alpha < \rho_1$.
2. Show that an entire function f having no zeros, is of the form $f(z) = \mathrm{e}^{g(z)}$, where $g(z)$ is an entire function.

Unit 8

Course Structure

- Infinite products and infinite product of functions
- Weierstrass factorization theorem.

8.1 Introduction

In this unit, our main objective is to deduce Weierstrass' factorization theorem, which asserts that every entire function can be represented as a (possibly infinite) product involving its zeroes. The theorem may be viewed as an extension of the fundamental theorem of algebra, which asserts that every polynomial may be factored into linear factors, one for each root.

The theorem, which is named for Karl Weierstrass, is closely related to a second result that every sequence tending to infinity has an associated entire function with zeroes at precisely the points of that sequence.

A generalization of the theorem extends it to meromorphic functions and allows one to consider a given meromorphic function as a product of three factors: terms depending on the function's zeros and poles, and an associated non-zero analytic function.

We will start off with the infinite products of complex numbers and study the conditions required for their convergence and thereafter establish the results for factorisation of entire functions in the upcoming units.

Objectives

After reading this unit, you will be able to

- define the conditions of convergence of infinite products and also the infinite product of functions
- define the Weierstrass' primary factors and related results
- learn about the factorisations of entire functions and deduce the Weierstrass' factorisation theorem
- deduce related results for the structures of entire functions

8.2 Infinite Products

An expression of the form

$$\prod_{n=1}^{\infty} u_n = u_1 u_2 \cdots u_n \cdots \tag{8.2.1}$$

where $\{u_n\}$ is a sequence of non-zero complex numbers is called an infinite product. Let $P_n = \prod_{n=1}^{\infty} u_n = u_1 u_2 \cdots u_n$. Then $\{P_n\}$ is called the sequence of partial products of (8.2.1). The infinite product (8.2.1) is said to be convergent if the sequence $\{P_n\}$ converges to a non-zero limit u as $n \to \infty$. If $\lim_{n\to\infty} P_n = u \neq 0$, then u is called the value the value of the infinite product (8.2.1) and we write $\prod_{n=1}^{\infty} u_n = u$. If an infinite product does not converge to a non-zero limit, it is said to be divergent.

However, sometimes we need to modify the definition as follows:

An infinite product $\prod_{n=1}^{\infty} u_n$ is said to be convergent if atmost a finite number of factors u_n are zero and the sequence of partial products formed by the non-vanishing factors tends to a non-zero finite limit.

Analogous to the necessary condition for the convergence of series, we have the following theorem for infinite products.

Theorem 8.2.1. (Necessary Condition of convergence) If an infinite product $\prod_{n=1}^{\infty} u_n$ is convergent, then, $\lim_{n\to\infty} u_n = 1$.

Proof. Let $\prod_{n=1}^{\infty} u_n = u$, then $P_n = u_1 u_2 \cdots u_n \to u \neq 0$ as $n \to \infty$, and $P_{n-1} = u_1 u_2 \cdots u_{n-1} \to u$. Hence

$$\lim_{n\to\infty} u_n = \lim_{n\to\infty} \frac{P_n}{P_{n-1}} = \frac{\lim_{n\to\infty} P_n}{\lim_{n\to\infty} P_{n-1}} = \frac{u}{u} = 1.$$

□

Remark 8.2.1. The condition is however, not sufficient. For example, if we take the product $\prod_{n=1}^{\infty} \frac{n}{n+1}$, then we have,

$$P_n = \frac{1}{2} \cdot \frac{2}{3} \cdots \frac{n-1}{n} \cdot \frac{n}{n+1} = \frac{1}{n+1} \to 0 \quad \text{as } n \to \infty$$

and thus the product is divergent. But, $\lim_{n\to\infty} u_n = \frac{n}{n+1} = 1$. So, the condition in the preceding theorem is not sufficient.

In view of the necessary condition for convergence, we write the general term of the product (8.2.1) in the form $u_n = (1 + a_n),\ (a_n \neq 1)$, so that a necessary condition for convergence of the product $\prod_{n=1}^{\infty}(1 + a_n)$ is $\lim_{n\to\infty} a_n = 0$.

Definition 8.2.1. (Absolute Convergence) An infinite product $\prod_{n=1}^{\infty}(1+a_n)$ is said to be absolutely convergent if the product $\prod_{n=1}^{\infty}(1+|a_n|)$ is convergent.

In case of absolute convergence, the factors of the product can be rearranged arbitrarily without affecting the convergence of the product or changing the value of the product.

Theorem 8.2.2. The infinite product $\prod_{n=1}^{\infty}(1+a_n)$, $(a_n \neq 1)$, converges if and only if the series $\sum_{n=1}^{\infty}\log(1+a_n)$ converges, where each logarithm has its principal value. Also,

$$\prod_{n=1}^{\infty}(1+a_n) = \exp\left(\sum_{n=1}^{\infty}\log(1+a_n)\right).$$

Proof. Let $P_n = \prod_{k=1}^{n}(1+a_k)$, $S_n = \sum_{k=1}^{n}\log(1+a_k)$ and $\lim_{n\to\infty} S_n = S$. Since $\mathrm{e}^{z+w} = \mathrm{e}^z \cdot \mathrm{e}^w$ for all $z, w \in \mathbb{C}$, we have,

$$\mathrm{e}^{S_n} = \mathrm{e}^{\log(1+a_1)} \cdot \mathrm{e}^{\log(1+a_2)} \cdots \mathrm{e}^{\log(1+a_n)} = (1+a_1)(1+a_2)\cdots(1+a_n) = P_n.$$

Since e^z is a continuous function on $\mathbb{C}$, $S_n \to S$ implies $P_n = \mathrm{e}^{S_n} \to \mathrm{e}^S \neq 0$. Hence, if the series $\sum_{n=1}^{\infty}\log(1+a_n)$ converges to S, then the product $\prod_{n=1}^{\infty}(1+a_n)$ converges to e^S.

Conversely, suppose that $P_n \to P(\neq 0)$ as $n \to \infty$. Without any loss of generality, we may assume that $P \notin (-\infty, 0]$. For, if $P \in (-\infty, 0]$, then we may consider in place of $\{1+a_n\}$ a new sequence $\{1+b_n\}$ with $1+b_1 = -(1+a_1)$ and $1+b_n = 1+a_n$ for $n \geq 2$. Then

$$\prod_{n=1}^{\infty}(1+b_n) = -\prod_{n=1}^{\infty}(1+a_n) = -P \notin (-\infty, 0].$$

As $P_n \to P \neq (-\infty, 0]$, we have, $P_n \in \mathbb{C} \setminus (-\infty, 0]$ for large n and since $\log z$ is continuous at P, $\log P_n \to \log P$ as $n \to \infty$. Since $\mathrm{e}^{S_n} = P_n$, we can write

$$S_n = \log P_n + 2\pi k_n i. \tag{8.2.2}$$

Then,

$$\log(1+a_{n+1}) = S_{n+1} - S_n = \log P_{n+1} - \log P_n + 2\pi i(k_{n+1} - k_n), \tag{8.2.3}$$

where k_{n+1} and k_n are integers. Equating the imaginary parts on both sides of (8.2.3), we get

$$\arg(1+a_{n+1}) = \arg P_{n+1} - \arg P_n + 2\pi(k_{n+1} - k_n). \tag{8.2.4}$$

Since the product $\prod_{n=1}^{\infty}(1+a_n)$ converges, we have $(1+a_n) \to 1$ and since $\arg(1+a_n)$ is continuous at 1, we have, $\arg(1+a_n) \to \arg 1 = 0$ as $n \to \infty$. Taking limit as $n \to \infty$ in (8.2.4), it follows that $(k_{n+1} - k_n) \to 0$ as $n \to \infty$. Since k_{n+1} and k_n are integers, there exists an integer N such that $k_{n+1} = k_n = \text{constant} = m$, for $n \geq N$, m being an integer. By (8.2.2), we conclude that $S_n \to \log P_n + 2\pi m i$ as $n \to \infty$ for some integer m. This completes the proof. $\square$

Theorem 8.2.3. If $a_n \geq 0$ for all $n \in \mathbb{N}$, then the product $\prod_{n=1}^{\infty}(1+a_n)$ converges if and only if the series $\sum_{n=1}^{\infty} a_n$ converges.

Proof. Let $P_n = \prod_{k=1}^{n}(1+a_k)$, and $S_n = \sum_{k=1}^{n} a_k$. Since $a_n \geq 0$ for all n, $\{P_n\}$ and $\{S_n\}$ are both increasing sequences. Since $1+x \leq \mathrm{e}^x$ for $x \geq 0$, we have, $a_1 < 1+a_1 \leq \mathrm{e}^{a_1}$. Hence,

$$(a_1+a_2+\cdots+a_n) < (1+a_1)(1+a_2)\cdots(1+a_n) \leq \mathrm{e}^{a_1+a_2+\cdots+a_n},$$

that is, $S_n < P_n \leq \mathrm{e}^{S_n}$. Thus, by the monotonic bounded principle, the series and the product are both convergent or both divergent according as both are bounded or unbounded. Let S_n be bounded. Then $S_n \leq M$ for some $M > 0$ and for all n. Now, $P_n \leq \mathrm{e}^{S_n}$ and so $P_n \leq \mathrm{e}^M$ for all n, that is P_n is bounded. Thus, both the series and the product are either both bounded or both unbounded. Hence the product $\prod_{n=1}^{\infty}(1+a_n)$ converges if and only if the series $\sum_{n=1}^{\infty} a_n$ converges. $\square$

Corollary 8.2.1. The product $\prod_{n=1}^{\infty}(1+a_n)$ is absolutely convergent if and only if the series $\sum_{n=1}^{\infty} a_n$ is absolutely convergent.

Proof. Since $|a_n| \geq 0$ for all $n \in \mathbb{N}$, so by the previous theorem, the product $\prod_{n=1}^{\infty}(1+|a_n|)$ is convergent if and only if the series $\sum_{n=1}^{\infty}|a_n|$ is convergent. Hence the result. $\square$

Theorem 8.2.4. The three series $\sum_{n=1}^{\infty}|a_n|$, $\sum_{n=1}^{\infty}|\log(1+a_n)|$ and $\sum_{n=1}^{\infty}\log(1+|a_n|)$ either converge or diverge together.

Proof. We know that $\lim_{z\to 0}\dfrac{\log(1+z)}{z} = 1$. Hence for ϵ with $0 < \epsilon < 1$, we get, $\left|\dfrac{\log(1+z)}{z} - 1\right| < \epsilon$ whenever $z \to 0$. The triangle inequality shows that

$$(1-\epsilon) < \left|\frac{\log(1+z)}{z}\right| < 1+\epsilon, \quad \text{whenever } z \to 0$$

that is,

$$(1-\epsilon)|z| < |\log(1+z)| < (1+\epsilon)|z|, \quad \text{whenever } z \to 0. \tag{8.2.5}$$

Similarly, $\lim_{t\to 0}\dfrac{\log(1+t)}{t} = 1$ for $0 < t \leq 1$. Then,

$$(1-\epsilon)t < \log(1+t) < (1+\epsilon)t, \quad \text{whenever } t \to 0. \tag{8.2.6}$$

We now assume that $a_n \to 0$ as $n \to \infty$. Consequently, from (8.2.5) and (8.2.6), we have

$$(1-\epsilon)|a_n| < |\log(1+a_n)| < (1+\epsilon)|a_n|, \quad \text{as } n \to \infty \tag{8.2.7}$$

$$\text{and} \quad (1-\epsilon)|a_n| < \log(1+|a_n|) < (1+\epsilon)|a_n|, \quad \text{as } n \to \infty. \tag{8.2.8}$$

From (8.2.7) and (8.2.8), by comparison test, for any sequence $\{a_n\}$ convergent to 0, the three series $\sum_{n=1}^{\infty} |a_n|$, $\sum_{n=1}^{\infty} |\log(1+a_n)|$ and $\sum_{n=1}^{\infty} \log(1+|a_n|)$ converge or diverge together. □

8.3 Infinite product of functions

We now consider the infinite products whose factors are functions defined on a set.

Given a sequence $\{f_n\}$ of functions defined on some set $E \subseteq \mathbb{C}$, the infinite product $\prod_{n=1}^{\infty}(1+f_n(z))$ is said to be convergent on E if for each $a \in E$,

$$\lim_{n\to\infty} P_n(a) = \lim_{n\to\infty} \prod_{k=1}^{n}(1+f_k(a)) \text{ exists and is non-zero.}$$

And, the infinite product $\prod_{n=1}^{\infty}(1+f_n(z))$ is said to be uniformly convergent to a function $f(z)$ in E if the sequence $\{P_n(z)\}$ of partial products, defined by $P_n(z) = \prod_{k=1}^{n}(1+f_k(z))$ is uniformly convergent to the function $f(z)$ in E, with $f(z) \neq 0$ in E.

Theorem 8.3.1. Let every term of the sequence of functions $\{f_n(z)\}$ be analytic in a region G and suppose the infinite series $\sum_{n=1}^{\infty} \log(1+f_n(z))$ is uniformly convergent on every compact subset of G (in particular, none of the terms $f_n(z)$ can take the value -1 at any point of G). Then the infinite product $\prod_{n=1}^{\infty}(1+f_n(z))$ converges uniformly on every compact subset of G.

Theorem 8.3.2. (M test) Let every term of the sequence of functions $\{f_n(z)\}$ be analytic in a region G and suppose none of the terms $f_n(z)$ takes the value -1 at any point of G. Moreover, suppose that there is a convergent series $\sum_{n=1}^{\infty} M_n$, whose terms are non-negative constants, such that $|f_n(z)| \leq M_n$ for all $z \in G$, and for all $n \geq N$, N being a positive integer. Then the infinite product $\prod_{n=1}^{\infty}(1+f_n(z))$ converges uniformly and absolutely to a non-vanishing analytic function $f(z)$ on every compact subset of G.

Exercise 8.3.1.

1. Prove that every absolutely convergent product is convergent.
2. Prove the following
   (a) $\prod_{n=2}^{\infty}\left(1-\frac{1}{n^2}\right) = \frac{1}{2}$
   (b) $\prod_{n=1}^{\infty}\left(1+\frac{(-1)^{n-1}}{n}\right) = 1$

8.4 Factorization of Entire functions

In this section, we try to construct an entire function $f(z)$ with the given zeros. Let f be an entire function with inly finite number of zeros $z_1, z_2, \dots, z_n$ (multiple zeros being repeated according to their multiplicities). Then, the function $\dfrac{f(z)}{(z-z_1)(z-z_2)\cdots(z-z_n)}$ is an entire function with no zeros. Hence, by the last theorem in the preceding unit, $\dfrac{f(z)}{(z-z_1)(z-z_2)\cdots(z-z_n)} = \mathrm{e}^{g(z)}$, where $g(z)$ is an entire function. Then, $f(z) = (z-z_1)(z-z_2)\cdots(z-z_n)\,\mathrm{e}^{g(z)}$, $g(z)$ is an entire function.

If, however, an entire function $f \neq 0$ has an infinite number of zeros, then set of zeros cannot have a limit point in any finite region of $\mathbb{C}$ since such a limit point would be a singularity of f. The only limit point is therefore, the point at infinity. We now consider the construction of an entire function with prescribed infinite number of zeros.

Definition 8.4.1. (Weierstrass' Primary factors): The functions

$$E(z,0) = 1-z, \;\; E(z,p) = (1-z)\exp\left(z + \frac{z^2}{2} + \cdots \frac{z^p}{p}\right), \;\; p = 1, 2, \dots$$

are called Weierstrass' primary factors. Each primary factor is an entire function with only one zero, a simple zero at $z = 1$.

Lemma 8.4.1. If $|z| \leq \dfrac{1}{2}$, $|\log E(z,p)| \leq 2|z|^{p+1}$.

Proof. First let $|z| < 1$. Then,

$$\begin{aligned} \log E(z,p) &= \log(1-z) + z + \frac{z^2}{2} + \cdots \frac{z^p}{p} \\ &= \left(-z - \frac{z^2}{2} - \cdots \frac{z^p}{p} - \frac{z^{p+1}}{p+1} - \cdots\right) + \left(z + \frac{z^2}{2} + \cdots \frac{z^p}{p}\right) \\ &= -\frac{z^{p+1}}{p+1} - \frac{z^{p+2}}{p+2} - \cdots \end{aligned} \tag{8.4.1}$$

Now if $|z| \leq \frac{1}{2}$,

$$\begin{aligned} |\log E(z,p)| &\leq |z|^{p+1} + |z|^{p+2} + \cdots \\ &= |z|^{p+1}\left(1 + |z| + |z|^2 + \cdots\right) \\ &\leq |z|^{p+1}\left(1 + \frac{1}{2} + \frac{1}{2^2} + \cdots\right) \\ &= 2|z|^{p+1} \end{aligned}$$

□

Theorem 8.4.1. Weierstrass' Factorization theorem: Let $\{a_n\}$ be an arbitrary sequence of complex numbers whose only limit point is ∞, that is, $a_n \to \infty$ as $n \to \infty$. Then it is possible to construct an entire function $f(z)$ with zeros precisely at these points.

Proof. We may suppose that the origin is not a zero of the entire function $f(z)$ to be constructed so that $a_n \neq 0\ \forall n$. Because if origin is a zero of $f(z)$ of order n, we need only multiply the constructed function by

z^m. We also arrange the zeros in order of non-decreasing modulus(if several distinct points a_n have the same modulus, we take them in any order) so that $|a_1| \leq |a_2| \leq \cdots$.

Let $|a_n| = r_n$. Since $r_n \to \infty$, we can always find a sequence of positive integers $\{p_n\}$ such that $\sum_{n=1}^{\infty} \left(\frac{r}{r_n}\right)^{p_n}$ converges for all $r > 0$. In fact, if $p_n = n$, for any given value of r, the inequality

$$\left(\frac{r}{r_n}\right)^n < \frac{1}{2^n}$$

holds for all sufficiently large values of n and hence the series is convergent.

Next, we take an arbitrary positive number R and choose the integer N such that $r_N \leq 2R < r_{N+1}$. Then, for $n > N$ and $|z| \leq R$, we have

$$\left|\frac{z}{a_n}\right| \leq \frac{R}{r_n} \leq \frac{R}{r_{N+1}} < \frac{1}{2}.$$

From the previous lemma, we get $\left|\log E\left(\frac{z}{a_n}, p_n - 1\right)\right| \leq 2\left(\frac{R}{r_n}\right)^{p_n}$. By Weierstrass' M-test, the series $\log E\left(\frac{z}{a_n}, p_n - 1\right)$ converges absolutely and uniformly in $|z| \leq R$. This implies that the infinite product $\prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, p_n - 1\right)$ converges absolutely and uniformly in the disk $|z| \leq R$, however large R may be. Hence, the above product represents an entire function, say $G(z)$. Thus $G(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, p_n - 1\right)$ with the same value of R. We choose another integer K such that $r_K \leq R < r_{K+1}$. Then each of the functions of the sequence $\prod_{n=1}^{m} E\left(\frac{z}{a_n}, p_n - 1\right)$, $m = K+1, K+2, \ldots$ vanishes at the points $a_1, a_2, \ldots a_k$ and nowhere else in $|z| \leq R$. Hence by **Hurwitz's theorem**(Let each function of the sequence $\{f_n\}$ be analytic in the closed region D and bounded by a closed contour γ. The sequence $\{f_n\}$ converges uniformly to f in D. If $f(z) \neq 0$ on γ, then f and the functions f_n, for all large values of n have the same number of zeros within γ. Also, a zero of f is either a zero of f_n for all sufficiently large values of n, or, else is a limit point of the set of zeros of the functions of the sequence), the only zeros of G in $|z| \leq R$ are $a_1, a_2, \ldots, a_K$. Since R is arbitrary this implies that the only zeros of G are the points of the sequence $\{a_n\}$. Thus, G is our required entire function. Now, if the origin is a zero function of order m of the required entire function $f(z)$, then $f(z)$ is of the form:

$$f(z) = z^m G(z)$$

□

Note 8.4.1. Since there are many possible sequences $\{p_n\}$ in the construction of the function $G(z)$ and ultimately of $f(z)$, the function $f(z)$ is not uniquely determined. Again, for any entire function $g(z)$, $\mathrm{e}^{g(z)}$ is also an entire function without any zeros. Hence, the general form of the required entire function $f(z)$ is of the form:

$$f(z) = z^m \,\mathrm{e}^{g(z)}\, G(z) = z^m \,\mathrm{e}^{g(z)} \prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, p_n - 1\right).$$

Theorem 8.4.2. If f is an entire function and $f(0) \neq 0$, then $f(z) = f(0)G(z)\,\mathrm{e}^{g(z)}$, where $G(z)$ is a product of primary factors and $g(z)$ is an entire function.

Proof. We form $G(z)$ as in the previous theorem from the zeros of f. Let $\phi(z) = \dfrac{f'(z)}{f(z)} - \dfrac{G'(z)}{G(z)}$. Then ϕ is an entire function, since the poles of one term are cancelled by those of the other. Let

$$g(z) = \int_0^z \phi(t)dt.$$

Then $g(z)$ is also an entire function. Now,

$$\begin{aligned}
g(z) &= \int_0^z \left(\frac{f'(t)}{f(t)} - \frac{G'(t)}{G(t)}\right) dt \\
&= \int_0^z \frac{d}{dt}\left(\log f(t) - \log G(t)\right) dt \\
&= \left[\log f(t) - \log G(t)\right]_0^z \\
&= \log f(z) - \log f(0) - \log G(z) + \log G(0) \\
&= \log f(z) - \log f(0) - \log G(z) \quad [\text{since } \ \log G(0) = 1] \\
&= \log \frac{f(z)}{f(0)G(z)}.
\end{aligned}$$

Hence,

$$\mathrm{e}^{g(z)} = \frac{f(z)}{f(0)G(z)} \Rightarrow f(z) = f(0)G(z)\,\mathrm{e}^{g(z)}.$$

□

Theorem 8.4.3. If the real part of an entire function f satisfies the inequality Re $f < r^{k+\epsilon}$ for any $\epsilon > 0$ and for a sequence of values of r tending to infinity, then f is a polynomial of degree not exceeding k.

Proof. By Taylor's theorem, $f(z) = \displaystyle\sum_{n=0}^{\infty} a_n z^n$, where $a_n = \dfrac{1}{2\pi i}\displaystyle\int_\gamma \frac{f(z)}{z^{n+1}}dz$, $\gamma : |z| = r$, r being a positive number. On γ, $z = r\,\mathrm{e}^{i\theta}$, $0 \le \theta \le 2\pi$. Now, when $n > 0$,

$$\begin{aligned}
\int_\gamma \frac{\overline{f(z)}}{z^{n+1}}dz &= \int_\gamma \sum_{m=0}^{\infty} \overline{a}_m \overline{z}^m \frac{dz}{z^{n+1}} \\
&= \sum_{m=0}^{\infty} \int_\gamma \overline{a}_m \overline{z}^m \frac{dz}{z^{n+1}} \\
&= \sum_{m=0}^{\infty} \int_0^{2\pi} \frac{\overline{a}_m r^m\,\mathrm{e}^{-im\theta} \cdot ir\,\mathrm{e}^{i\theta}}{r^{n+1}\,\mathrm{e}^{i(n+1)\theta}} d\theta \\
&= \sum_{m=0}^{\infty} \int_0^{2\pi} \overline{a}_m r^{m-n}\,\mathrm{e}^{-i(m+n)\theta}\, i d\theta = 0,
\end{aligned}$$

the term by term integration is valid since the series $\displaystyle\sum_{m=0}^{\infty} \overline{a}_m \overline{z}^m$ converges uniformly. Hence, for $n > 0$,

$$\begin{aligned}
a_n &= \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z^{n+1}}dz + \frac{1}{2\pi i}\int_\gamma \frac{\overline{f(z)}}{z^{n+1}}dz \\
&= \frac{1}{2\pi i}\int_\gamma \left(f(z) + \overline{f(z)}\right)\frac{dz}{z^{n+1}} \\
&= \frac{1}{\pi i}\int_\gamma \text{Re } f(z)\frac{dz}{z^{n+1}} = \frac{1}{\pi}\int_0^{2\pi} \frac{\text{Re } f(r\,\mathrm{e}^{i\theta})}{r^n\,\mathrm{e}^{in\theta}} d\theta.
\end{aligned}$$

Hence,

$$a_n r^n = \frac{1}{\pi}\int_0^{2\pi} \frac{\text{Re } f(r\,\mathrm{e}^{i\theta})}{\mathrm{e}^{in\theta}} d\theta. \Rightarrow |a_n| r^n \le \frac{1}{\pi}\int_0^{2\pi} |\text{Re } f(r\,\mathrm{e}^{i\theta})| d\theta, \text{ for } n > 0.$$

Also,

$$a_0 = \frac{1}{2\pi i}\int_\gamma \frac{f(z)}{z} dz = \frac{1}{2\pi}\int_0^{2\pi} f(r\,\mathrm{e}^{i\theta}) d\theta$$

and so,

$$\text{Re } a_0 = \frac{1}{2\pi}\int_0^{2\pi} \text{Re } f(r\,\mathrm{e}^{i\theta}) d\theta.$$

Hence,

$$\begin{aligned} 2\text{Re } a_0 + |a_n| r^n \le \frac{1}{\pi}\int_0^{2\pi} \left(|\text{Re } f| + \text{Re } f\right) d\theta &= 2\text{Re } f; \text{ if Re } f > 0 \\ &= 0; \text{ if Re } f \le 0. \end{aligned}$$

Since Re $f < r^{k+\epsilon}$ for any $\epsilon > 0$ and for a sequence of values $\{r_m\}$ of r tending to infinity, we have,

$$2\text{Re } a_0 + |a_n| r_m^n < 4 r_m^{k+\epsilon} \Rightarrow |a_n| < 4 r_m^{k+\epsilon-n} - 2\text{Re } a_0 r_m^{-n}$$

for any $\epsilon > 0$. Taking limit as $r_m \to \infty$, we have, $a_n = 0$ when $n > k$ and so, f is a polynomial of degree not exceeding k. □

Theorem 8.4.4. The function $\mathrm{e}^{f(z)}$ is an entire function of finite order with no zeros if and only if $f(z)$ is a polynomial.

Proof. We already know that $\mathrm{e}^{f(z)}$ is an entire function with no zeros if and only if f is an entire function. Moreover, if f is a polynomial of degree k, then $\mathrm{e}^{f(z)}$ is of finite order k.

Conversely, we assume that $\mathrm{e}^{f(z)}$ is an entire function with finite order ρ and without any zeros. Then, f is an entire function. Also, $\left|\mathrm{e}^{f(z)}\right| < \mathrm{e}^{r^{\rho+\epsilon}}$, for all $r > r_0$, $r = |z|$ and $\epsilon > 0$ is arbitrary, that is,

$$\mathrm{e}^{\text{Re } f} < \mathrm{e}^{r^{\rho+\epsilon}} \Rightarrow \text{Re } f < r^{\rho+\epsilon}, \ \forall r > r_0 \text{ and } \epsilon > 0.$$

Hence, by the previous theorem, f is a polynomial of degree not exceeding ρ. □

Few Probable Questions

1. When is an infinite product $\prod_{n=1}^{\infty} u_n$, $u_n \neq 0$ for all n, said to be convergent. Deduce a necessary condition for the convergence of the product.

2. Show that if an infinite product $\prod_{n=1}^{\infty} u_n$, $u_n \neq 0$ for all n, is convergent, then $\lim_{n\to\infty} u_n = 1$. Is the condition sufficient? Justify your answer.

3. Show that the infinite product $\prod_{n=1}^{\infty}(1+a_n)$, $(a_n \neq -1)$ converges if and only if the series $\sum_{n=1}^{\infty}\log(1+a_n)$ converges and

$$\prod_{n=1}^{\infty}(1+a_n) = \exp\left(\sum_{n=1}^{\infty}\log(1+a_n)\right).$$

4. Show that the infinite product $\prod_{n=1}^{\infty}(1+a_n)$, $(a_n \geq 0)$ converges if and only if the series $\sum_{n=1}^{\infty} a_n$ converges.

5. State and prove Weierstrass' Factorization theorem.

6. If for an entire function f satisfies the inequality Re $f < r^{k+\epsilon}$ for any $\epsilon > 0$ and a sequence of values of r tending to infinity, then show that f is a polynomial of degree not exceeding k.

Unit 9

Course Structure

- Canonical product, Borel's first theorem. Borel's second theorem (statement only).

9.1 Introduction

This unit deals with the factorization of entire functions of finite order with the help of the newly defined canonical product. Hence we have deduced several results related to the relationship between the order of an entire function f and the convergence exponent of its zeros culminating in the Picard's little theorem that we will come across in the next unit.

Objectives

After reading this unit, you will be able to

- define the canonical product and genus of entire functions
- deduce the Borel's first theorem

9.2 Canonical Product

Let f be an entire function with infinite number of zeros a_n, $n = 1, 2, \ldots$ where $a_n \neq 0$ for all n and $|a_n| = r_n$. If there exists a least non-negative integer p such that the series $\sum\limits_{n=1}^{\infty} \dfrac{1}{r_n^{p+1}}$ is convergent. We form the infinite product $G(z) = \prod\limits_{n=1}^{\infty} E\left(\dfrac{z}{a_n}, p\right)$. By Weierstrass' factorization theorem, $G(z)$ represents an entire function having zeros precisely at the points a_n. We call $G(z)$ as the canonical product formed with the sequence $\{a_n\}$ of zeros of f and the integer p is called its **genus**.

If $z = 0$ is a zero of f of order m, then the canonical product is $z^m G(z)$.

Observe that, if the convergence exponent $\rho_1 \neq$ an integer, then $p = [\rho_1]$ and if $\rho_1 =$ an integer, then

1. $p = \rho_1$, if $\displaystyle\sum_{n=1}^{\infty} \frac{1}{r_n^{\rho_1}}$ is divergent;

2. $p = \rho_1 - 1$, if $\displaystyle\sum_{n=1}^{\infty} \frac{1}{r_n^{\rho_1}}$ is convergent.

In any case, $\rho_1 - 1 \le p \le \rho_1 \le \rho$, where $\rho =$ order of $f(z)$.

Example 9.2.1. 1. Let $a_n = n$. Then $\displaystyle\sum_{n=1}^{\infty} \frac{1}{r_n^2} = \sum_{n=1}^{\infty} \frac{1}{n^2}$ is convergent, while $\displaystyle\sum_{n=1}^{\infty} \frac{1}{r_n} = \sum_{n=1}^{\infty} \frac{1}{n}$ is divergent. Hence, the genus $p = 1$.

2. Let $a_n = \mathrm{e}^n$. Then the genus is $p = 0$.

3. Let $a_1 = \dfrac{1}{2}\log 2$, $a_n = \log n$, $n \ge 2$. Then there exists no finite p such that the series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{r_n^{p+1}}$ is convergent.

Theorem 9.2.1. (Borel's theorem) The order of a canonical product is equal to the convergence exponent of its zeros.

Proof. Let

$$G(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, p\right)$$

be a canonical product with zeros at the points $a_1, a_2, \ldots$ and genus p. Let ρ_1 and ρ be the convergence exponent and order respectively of $G(z)$. Since $\rho_1 \le \rho$ for any entire function, we only need to prove that $\rho \le \rho_1$ for $G(z)$. Let $|a_n| = r_n$ and $K(> 1)$ be a constant. Then for $|z| = r$,

$$\log|G(z)| = \sum_{r_n \le Kr} \log\left|E\left(\frac{z}{a_n}, p\right)\right| + \sum_{r_n > Kr} \log\left|E\left(\frac{z}{a_n}, p\right)\right| = \Sigma_1 + \Sigma_2 \text{ (say)}.$$

We first estimate Σ_2. In Σ_2, $\dfrac{r}{r_n} < \dfrac{1}{K} < 1$. Hence,

$$\begin{aligned}
\log E\left(\frac{z}{a_n}, p\right) &= -\frac{1}{p+1}\left(\frac{z}{a_n}\right)^{p+1} - \frac{1}{p+2}\left(\frac{z}{a_n}\right)^{p+2} - \cdots \\
\Rightarrow \left|\log E\left(\frac{z}{a_n}, p\right)\right| &< \frac{1}{p+1}\left\{\left(\frac{r}{r_n}\right)^{p+1} + \left(\frac{r}{r_n}\right)^{p+2} + \cdots\right\} \\
&= \frac{1}{p+1}\left(\frac{\left(\frac{r}{r_n}\right)^{p+1}}{1 - \frac{r}{r_n}}\right) < A\left(\frac{r}{r_n}\right)^{p+1},
\end{aligned}$$

A being a constant. Also, we know that $\log|f| = \mathrm{Re}(\log f) \le |\log f|$ [since $\log f(z) = \log|f(z)| + i\arg f(z)$], for any function f.

Hence,

$$\Sigma_2 = \sum_{r_n > Kr} \log\left|E\left(\frac{z}{a_n}, p\right)\right| = O\left(\sum_{r_n > Kr}\left(\frac{r}{r_n}\right)^{p+1}\right) = O\left(r^{p+1}\sum_{r_n > Kr}\frac{1}{{r_n}^{p+1}}\right) = O\left(r^{p+1}\right)$$

$\left[\text{since} \displaystyle\sum_{r_n > Kr} \frac{1}{{r_n}^{p+1}} \text{ is convergent by the definition of } p \text{ and converges to } B(\text{say})\right]$. If $p+1 = \rho_1$, then $\sum_2 = O\left(r^{\rho_1}\right)$. Otherwise, $p+1 > \rho_1 + \epsilon$, $\epsilon > 0$ being small enough. Then

$$r^{p+1} \sum_{r_n > Kr} r_n^{-p-1} = r^{p+1} \sum_{r_n > Kr} \left(r_n^{\rho_1+\epsilon-p-1} r_n^{-\rho_1-\epsilon}\right) < r^{p+1} (Kr)^{\rho_1-\epsilon-1} \sum_{r_n > Kr} r_n^{-\rho_1-\epsilon} = O\left(r^{\rho_1+\epsilon}\right)$$

$\left[\text{since } \sum_{r_n > Kr} r_n^{-\rho_1-\epsilon} \text{ is convergent}\right]$. Thus, in any case,

$$\Sigma_2 = O\left(r^{\rho_1+\epsilon}\right). \tag{9.2.1}$$

Next, we estimate $\sum_1$. In $\sum_1$, $\dfrac{r}{r_n} \geq \dfrac{1}{K}$. Now,

$$\begin{aligned} \log\left|E\left(\frac{z}{a_n}, p\right)\right| &= \log\left|\left(1 - \frac{z}{a_n}\right) \exp\left(\frac{z}{a_n} + \dots \frac{1}{p}\left(\frac{z}{a_n}\right)^p\right)\right| \\ &\leq \log\left(1 + \frac{r}{r_n}\right) + \frac{r}{r_n} + \cdots \frac{1}{p}\left(\frac{r}{r_n}\right)^p. \end{aligned}$$

Also, $\log\left(1 + \dfrac{r}{r_n}\right) < \dfrac{r}{r_n}$ [since $1+|x| < \mathrm{e}^{|x|}$, hence $\log(1+|x|) < |x|$]. Hence, $\log\left|E\left(\dfrac{z}{a_n}, p\right)\right| < A\left(\dfrac{r}{r_n}\right)^p$ where A depends only on K. Hence $\sum_1 = O\left(\displaystyle\sum_{r_n \leq Kr} \left(\frac{r}{r_n}\right)^p\right) = O\left(r^p \displaystyle\sum_{r_n \leq Kr} {r_n}^{-p}\right)$
$= O\left(r^p \displaystyle\sum_{r_n \leq Kr} r_n^{\rho_1+\epsilon-p} \cdot r_n^{-\rho_1-\epsilon}\right) = O\left(r^p (Kr)^{\rho_1+\epsilon-p} \cdot \displaystyle\sum_{r_n \leq Kr} r_n^{-\rho_1-\epsilon}\right) = O\left(r^{\rho_1+\epsilon}\right)$. Using this and equation (9.2.1), we get, $\log|G(z)| = O(r^{\rho_1+\epsilon})$. This implies that $\rho \leq \rho_1$. Combining, we have $\rho = \rho_1$. □

Example 9.2.2. We find the canonical product of $\sin z$.

Let $f(z) = \sin z$. Then f is an entire function with infinite number of zeros at $z = n\pi$, n being an integer. First we consider the zeros of f exceeding the simple zero at $z = 0$. Let $a_n = n\pi$, $n = \pm 1, \pm 2, \dots$ and $|a_n| = r_n = |n\pi|$. Now,

$$\sum_{n=1}^{\infty} \frac{1}{r_n} = \sum_{n=1}^{\infty} \frac{1}{|n\pi|} = \frac{1}{\pi} \sum_{n=1}^{\infty} \frac{1}{n}$$

is divergent, but

$$\sum_{n=1}^{\infty} \frac{1}{r_n^2} = \frac{1}{\pi^2} \sum_{n=1}^{\infty} \frac{1}{n^2}$$

is convergent. Hence the genus of the required canonical product is 1. Hence the canonical product is given by

$$\begin{aligned} G(z) &= \prod_{n=-\infty}^{-1} \left(1 - \frac{z}{a_n}\right) \mathrm{e}^{\frac{z}{a_n}} \cdot \prod_{n=1}^{\infty} \left(1 - \frac{z}{a_n}\right) \mathrm{e}^{\frac{z}{a_n}} \\ &= \prod_{n=1}^{\infty} \left(1 - \frac{z}{n\pi}\right) \mathrm{e}^{\frac{z}{n\pi}} \left(1 + \frac{z}{n\pi}\right) \mathrm{e}^{-\frac{z}{n\pi}} \\ &= \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2\pi^2}\right). \end{aligned}$$

Now, since zero is a simple zero of f, the canonical product of f will be

$$G(z) = z\prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2\pi^2}\right).$$

Exercise 9.2.1. Find the canonical product of $\cos z$.

Few Probable Questions

1. Define canonical product of the zeros of an entire function. Find the canonical product of the function $\sin z$.
2. State and prove Borel's theorem.
3. Show that for an entire function f of finite order ρ, which is not an integer, and convergence exponent ρ_1, $\rho = \rho_1$.
4. Show that
$$\sin \pi z = \pi z\prod_{n=1}^{\infty}\left(1 - \frac{z^2}{n^2}\right).$$

Unit 10

Course Structure

- Hadamard's factorization theorem,
- Schottky's theorem (no proof), Picard's first theorem.

10.1 Introduction

Picard's little theorem, named after Charles E. Picard, is an important theorem that puts light on the range set of a non-constant entire function. In 1879 Picard proved that an entire function takes on every value with at most one exception, (Picard's "Little Theorem"), and that in any neighborhood of an isolated essential singularity, an analytic function takes on every value except at most one, (Picard's "Big Theorem"). We will discuss few examples in this light.

Objectives

After reading this unit, you will be able to

- deduce Hadamard's factorization theorem
- deduce related results and the celebrated Picard's little theorem

10.2 Hadamard's Factorization theorem and results

Theorem 10.2.1. (Hadamard's Factorization theorem) If f is an entire function of order ρ with infinite number of zeros and $f(0) \neq 0$, then $f(z) = \mathrm{e}^{Q(z)}\, G(z)$, where $G(z)$ is the canonical product formed with the zeros of f and $Q(z)$ is a polynomial of degree not greater than ρ.

Proof. By Weierstrass' factorization theorem, we have,

$$f(z) = \mathrm{e}^{Q(z)}\, G(z), \tag{10.2.1}$$

where $Q(z)$ is an entire function and $G(z)$ is the canonical product with genus p formed with the zeros $a_1, a_2, \ldots$ of f. Since ρ is finite, we need to show that $Q(z)$ is a polynomial of degree less than or equal to ρ. Let $m = [\rho]$. Then genus $p \leq m$. Taking logarithms on both sides of (10.2.1), we get

$$\begin{aligned}\log f(z) &= Q(z) + \log G(z)\\ &= Q(z) + \sum_{n=1}^{\infty} \log\left(1 - \frac{z}{a_n}\right) + \sum_{n=1}^{\infty}\left\{\frac{z}{a_n} + \frac{1}{2}\left(\frac{z}{a_n}\right)^2 + \cdots + \frac{1}{p}\left(\frac{z}{a_n}\right)^p\right\}\end{aligned} \tag{10.2.2}$$

$$\left[\text{since } G(z) = \prod_{n=1}^{\infty} E\left(\frac{z}{a_n}, p\right) = \prod_{n=1}^{\infty}\left(1 - \frac{z}{a_n}\right)\exp\left(\frac{z}{a_n} + \cdots + \frac{1}{p}\left(\frac{z}{a_n}\right)^p\right)\right].$$

Differentiating (10.2.2) $m+1$ times, we get

$$\frac{d^m}{dz^m}\left(\frac{f'(z)}{f(z)}\right) = Q^{(m+1)}(z) - m! \sum_{n=1}^{\infty} \frac{1}{(a_n - z)}^{m+1} \tag{10.2.3}$$

$$\left[\text{since } p \leq m,\ \frac{d^{m+1}}{dz^{m+1}} \sum_{n=1}^{\infty}\left\{\frac{z}{a_n} + \frac{1}{2}\left(\frac{z}{a_n}\right)^2 + \cdots + \frac{1}{p}\left(\frac{z}{a_n}\right)^p\right\} = 0\right]$$

$$\left[\text{and } \frac{d^{m+1}}{dz^{m+1}} \log\left(1 - \frac{z}{a_n}\right) = \frac{d^{m+1}}{dz^{m+1}} \log(a_n - z) = -m! \frac{1}{(a_n - z)}^{m+1}\right].$$

Now, $Q(z)$ will be a polynomial of degree n at most if we can show that $Q^{(m+1)}(z) = 0$.

Let

$$g_R(z) = \frac{f(z)}{f(0)} \prod_{|a_n| \leq R}\left(1 - \frac{z}{a_n}\right)^{-1}.$$

Then since $f(z)$ is entire and $f(0) \neq 0$ and $\displaystyle\prod_{|a_n| \leq R}\left(1 - \frac{z}{a_n}\right)^{-1}$ $g_R(z)$ cancels with the factors in $f(z)$, so $g_R(z)$ is an entire function and $g_R(z) \neq 0$ in $|z| \leq R$. For $|z| = 2R$ and $|a_n| \leq R$, we have, $\left|1 - \frac{z}{a_n}\right| \geq 1$. Hence,

$$|g_R(z)| \leq \frac{|f(z)|}{|f(0)|} < A \exp\left((2R)^{\rho+\epsilon}\right), \quad \text{for } |z| = 2R, \ A = \text{constant}.$$

By maximum modulus theorem,

$$|g_R(z)| < A \exp\left((2R)^{\rho+\epsilon}\right), \quad \text{for } |z| < 2R. \tag{10.2.4}$$

Let $h_R(z) = \log g_R(z)$, the logarithm being determined such that $h_R(0) = 0$. Then $h_R(z)$ is analytic in $|z| \leq R$. Now, from (10.2.4), we have

$$\text{Re } h_R(z) = \log|g_R(z)| < \log A + 2^{\rho+\epsilon} R^{\rho+\epsilon}.$$

Hence,

$$\text{Re } h_R(z) < K R^{\rho+\epsilon}, \quad K = \text{constant}. \tag{10.2.5}$$

Hence, from the second corollary of Borel Caratheodory theorem, we have,

$$|h_R^{(m+1)}(z)| \leq \frac{2^{m+3}(m+1)! \cdot R}{(R-r)^{m+2}} K R^{\rho+\epsilon}, \quad \text{for } |z| = r < R.$$

Hence, for $|z| = r = \dfrac{R}{2}$,
$$h_R^{(m+1)}(z) = O\left(R^{\rho+\epsilon-m-1}\right). \tag{10.2.6}$$
But,
$$h_R(z) = \log g_R(z) = \log f(z) - \log f(0) - \sum_{|a_n|\le R} \log\left(1 - \frac{z}{a_n}\right).$$
Hence,
$$h_R^{(m+1)}(z) = \frac{d^m}{dz^m}\left(\frac{f'(z)}{f(z)}\right) + m! \sum_{|a_n|\le R} \frac{1}{(a_n - z)^{m+1}}.$$
Hence, from (10.2.2), we have
$$\begin{aligned} Q^{(m+1)}(z) &= h_R^{(m+1)}(z) + m! \sum_{|a_n|\le R} \frac{1}{(a_n - z)^{m+1}} \\ &= O\left(R^{\rho+\epsilon-m-1}\right) + O\left(\sum_{|a_n|>R} \frac{1}{|a_n|^{m+1}}\right), \quad \text{for } |z| = \frac{R}{2} \text{ [by (10.2.6)]} \end{aligned} \tag{10.2.7}$$
and so also for $|z| < \dfrac{R}{2}$ by maximum modulus theorem. The first term on the right hand side of (10.2.7) tends to zero as $R \to \infty$ if $\epsilon > 0$ is small enough, since $m + 1 > \rho$. Also, the second term tends to 0 since $\displaystyle\sum_{n=1}^{\infty} \frac{1}{|a_n|^{m+1}}$ is convergent. In fact, $\displaystyle\sum_{|a_n|>R} \frac{1}{|a_n|^{m+1}}$ becomes the remainder term for large R. Hence, $Q^{(m+1)}(z) = 0$, since $Q^{(m+1)}$ is independent of R. Thus, $Q(z)$ is a polynomial of degree not greater than ρ. □

Note 10.2.1. If $z = 0$ is a zero of f of order m, then $f(z) = z^m\, e^{Q(z)}\, G(z)$.

Theorem 10.2.2. If f is an entire function of order ρ and ρ_1 is the convergence exponent of its zeros, then $\rho_1 = \rho$ if ρ is not an integer.

Proof. Since the zeros of f coincide with the zeros of its canonical product $G(z)$, we can take ρ_1 to be the convergence exponent of the zeros of $G(z)$. By Hadamard's factorization theorem, we have, $f(z) = e^{Q(z)}\, G(z)$, where $Q(z)$ is a polynomial of degree not exceeding ρ. In any case, $\rho_1 \le \rho$. Suppose if possible $\rho_1 < \rho$. Also, if degree of $Q(z) = q$, then $e^{Q(z)}$ is of order $q \le \rho$. In this case, $q < \rho$, since q is an integer and ρ is not an integer. Thus, f is the product of two entire functions each of order less than ρ. Hence, order of f is less than ρ which contradicts the given hypothesis. Hence, $\rho_1 = \rho$. □

Theorem 10.2.3. Let f be an entire function with order ρ and g be an entire function with order $\le \rho$. If the zeros of g are all zeros of f, then $H(z) = \dfrac{f(z)}{g(z)}$ is an entire function of order ρ at most.

Proof. Since the zeros of g are all zeros of f, $H(z)$ is an entire function. Let $G_1(z)$ and $G_2(z)$ be the canonical products formed with the zeros of f and g respectively. By Hadamard's factorization theorem, we have
$$f(z) = G_1(z)\, e^{Q_1(z)} \quad \text{and} \quad g(z) = G_2(z)\, e^{Q_2(z)},$$
where $Q_1(z)$ and $Q_2(z)$ are polynomials with degrees less than or equal to ρ. Then,
$$H(z) = G(z)\, e^{Q_1(z)-Q_2(z)}, \quad \text{where } G(z) = \frac{G_1(z)}{G_2(z)}$$

is the canonical product formed with the zeros of $G_1(z)$ that are not zeros of $G_2(z)$. Since the convergence exponent of a sequence is not increased by removing some of the terms, the convergence exponent and hence the order of $G(z)$ does not exceed ρ. Also, $Q_1(z) - Q_2(z)$ is a polynomial of degree $\leq \rho$. Hence, order of $e^{Q_1(z)-Q_2(z)}$ is $\leq \rho$. Thus, H is the product of two entire functions, each of order $\leq \rho$. Hence, order of $H(z)$ is ρ at most. □

Theorem 10.2.4. (Picard's little theorem) An entire function of finite order takes any complex number except at most one number.

Proof. Let f be an entire function of finite order. If possible, let f do not take two values a and b. Then $f(z)-a \neq 0$ and $f(z)-b \neq 0$ for all $z \in \mathbb{C}$. Thus, there exists an entire function g such that $f(z) - a = e^{g(z)}$. Since f is of finite order, the function $f(z) - a$ is also of finite order. By Hadamard's factorization theorem, $g(z)$ must be a polynomial. Now,

$$f(z) - b = f(z) - a + (a - b) = e^{g(z)} + (a - b) \neq 0 \ \ \forall z \in \mathbb{C}.$$

Hence, $e^{g(z)} \neq b - a$ for all $z \in \mathbb{C}$. This is a contradiction, since $g(z)$ being a polynomial, by fundamental theorem of algebra, $g(\mathbb{C}) = \mathbb{C}$. Hence the theorem. □

Example 10.2.1.

1. The most common example is the non-constant entire function e^z which omits only the value 0.
2. Any non-constant polynomial f takes all the values of the finite complex plane. This is due to the fact that for any complex number a, the function $f(z) - a$ is also a polynomial in $\mathbb{C}$ having zero in $\mathbb{C}$ by the fundamental theorem of algebra.

Theorem 10.2.5. Let f be an entire function of finite order ρ which is not an integer. Then f has infinitely many zeros.

Proof. Let f be an entire function of finite order ρ which is not an integer. If possible, suppose that the zeros of f are $\{a_1, a_2, \dots, a_n\}$, finite in number, counted according to multiplicities. Then $f(z)$ can be expressed as $f(z) = (z - a_1)(z - a_2) \cdots (z - a_n)\, e^{g(z)}$, where $g(z)$ is an entire function. By Hadamard's factorization theorem, $g(z)$ is a polynomial whose degree is less than or equal to ρ. Clearly, $f(z)$ and $e^{g(z)}$ are of same order. But the order of $e^{g(z)}$ is exactly the degree of $g(z)$, which is an integer. This implies that ρ is an integer. This is a contradiction and hence the result. □

Example 10.2.2. If $\alpha > 1$, we show that the entire function $\displaystyle\prod_{n=1}^{\infty}\left(1 - \frac{z}{n^\alpha}\right)$ is of finite order $\dfrac{1}{\alpha}$. Firstly, the infinite product $\displaystyle\prod_{n=1}^{\infty}\left(1 - \frac{z}{n^\alpha}\right)$ is of the form $\displaystyle\prod_{n=1}^{\infty} E\left(\frac{z}{n^\alpha}, 0\right)$. Here, $p = 0$ is the least non-negative integer for which,

$$\sum_{n=1}^{\infty} \frac{1}{r_n^{p+1}} = \sum_{n=1}^{\infty} \frac{1}{n^\alpha}$$

is convergent since $\alpha > 1$. Hence, the given infinite product is a canonical product. So, the order of $\displaystyle\prod_{n=1}^{\infty}\left(1 - \frac{z}{n^\alpha}\right)$ is the same as the convergence exponent of its zeros. Here, zeros are $a_n = n^\alpha$. Hence, $r_n = |a_n| = n^\alpha$. Hence the convergence exponent

$$\rho_1 = \limsup_{n\to\infty} \frac{\log n}{\log r_n} = \limsup_{n\to\infty} \frac{\log n}{\alpha \log n} = \frac{1}{\alpha}.$$

Exercise 10.2.1. 1. Prove the following.

(a) $\sin \pi z = \pi z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{n^2}\right)$

(b) $\cos z = \prod_{n=1}^{\infty} \left(1 - \frac{4z^2}{\pi^2(2n-1)^2}\right)$

2. Prove Fundamental theorem of algebra using Picard's little theorem.

3. With proper justifications, write which of the following functions is constant.

(a) An entire function that omit two values on the complex plane;
(b) An entire function that omit values on the negative real axis;
(c) An entire function whose range set is a dense set;
(d) An entire function whose range set is not a dense set;
(e) An entire function whose range set is a straight line;
(f) An entire function whose range set is $B(0; R)$, $R > 0$ is a real number;
(g) An entire function for which $M(r_1) = M(r_2)$ for $r_1 < r_2$;
(h) An entire function whose range set is a closed set;
(i) An entire function that takes the value $a \in \mathbb{C}$ at all points in the set E that has a limit point in $\mathbb{C}$.

Few Probable Questions

1. State and prove Hadamard's Factorization theorem.
2. State the Hadamard's factorization theorem. If f and g be entire functions of order ρ and ρ' respectively, such that $\rho' \leq \rho$, and also if the zeros of g are all zeros of f, the show that the function $H(z) = \frac{f(z)}{g(z)}$ is an entire function of order at most ρ.
3. State and prove the Picard's little theorem.
4. Show that an entire function f of finite order ρ, which is not an integer, has infinitely many zeros.
5. State Picard's Little theorem. Hence prove the Fundamental theorem of Algebra.

Unit 11

Course Structure

- Multiple-valued functions

11.1 Introduction

Multifunctions are hard to avoid. Many complex functions, like the complex exponential, are not globally one-to-one. We may view such a function as having an inverse, so long as we allow the inverse to be a multifunction. Constructing, at least locally, a well-behaved functional inverse will involve extracting a suitable value from this multifunction at each point of the domain. But in order to treat more complicated examples, such as $\sqrt{z}$, $z^{2/3}$, etc., we begin a deeper analysis of many-valuedness. which will enable us to handle logarithms and powers of rational functions.

Objectives

After reading this unit, you will be able to

- define multifunctions and visualize certain examples related to them
- discuss in detail the argument function and its rle in the many-valuedness of complex functions
- define branches and branch cuts of multifunctions

11.2 Multiple-Valued Functions

Thus far, we have considered a complex function f to be a rule that assigns to each point z, a single complex number $f(z)$. This familiar conception of a function is unduly restricted. Using examples, we now discuss how we may broaden the definition of a function to allow $f(z)$ to have many different values for a single value of z. In such a case, f is called a many-valued function or a multifunction.

For example, if we consider the function $\sqrt[3]{z}$, then it has three different values (if z is non-zero) for a single value of z and hence is a three-valued multifunction.

Let us see how. The function $z \mapsto \sqrt[3]{z}$, that is, if $w = \sqrt[3]{z}$, and a is a solution of this equation, let us find the other two solutions too. If $z = r\,\mathrm{e}^{i\theta}$ orbits round an origin-centred circles, $z^3 = r^3\,\mathrm{e}^{3i\theta}$ orbits three times

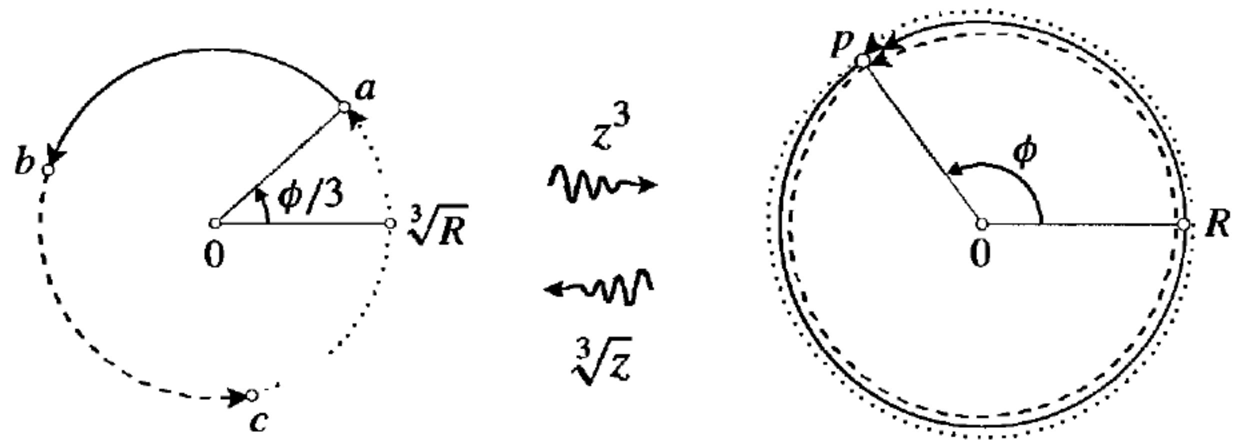


Figure 11.1

faster executing a complete revolution each time z executes one third of a revolution. Put differently, reversing the direction of the mapping divides the angular speed by three. This is an essential ingredient, which we will now study in detail.

Writing $z = r\,\mathrm{e}^{i\theta}$, we have $\sqrt[3]{z} = \sqrt[3]{r}\,\mathrm{e}^{i(\theta/3)}$. Here, $\sqrt[3]{r}$ uniquely defined as the real cube root of the length of z; the sole source of the three-fold ambiguity in the formula is the fact that there are infinitely many different choices for the angle θ of a given point z.

Think of z as a moving point that is initially at $z = p$. If we arbitrarily choose θ to be the angle ϕ as shown in fig. 11.1, then $\sqrt[3]{p} = a$. As z gradually moves away from p, θ gradually changes from its initial value ϕ, and $\sqrt[3]{z} = \sqrt[3]{r}\,\mathrm{e}^{i(\theta/3)}$ gradually moves away from its initial position a, but in a completely determined way-its distance from the origin is the cubic root of the distance of z, and its speed of movement is one-third that of z.

11.3 Argument as a function

The angle θ in the expression $z = |z|\,\mathrm{e}^{i\theta}$ is not uniquely determined. Indeed, this is the fundamental cause of many-valuedness in complex function theory. For $z \neq 0$, we define the argument of z to be

$$[\arg z] = \{\theta \in \mathbb{R} :\ z = |z|\,\mathrm{e}^{i\theta}\}.$$

The bracket notation $[\arg z]$ is designed to emphasize that the argument of z is a set of numbers, not a single number. In fact, $[\arg z]$ is an infinite set, consisting of all numbers from $\theta + 2k\pi$ for $k \in \mathbb{Z}$, where θ is any fixed real number such that $\mathrm{e}^{i\theta} = z/|z|$.

The restriction $-\pi < \theta \leq \pi$, or alternatively, $0 \leq \theta < 2\pi$, uniquely determines θ in the equation $0 \neq z = |z|\,\mathrm{e}^{i\theta}$.

Now, consider what happens to a principal value determination of argument Arg $z = \theta$, where $z = |z|\,\mathrm{e}^{i\theta}$, $0 \leq \theta < 2\pi$, where z performs a complete anticlockwise circuit round the unit circle starting from $z = 1$, with $\theta \in [0, 2\pi)$. Within the chosen range $[0, 2\pi)$, θ has value 0 at the start and increases steadily towards 2π as z moves round the circle until it arrives back at 1, when θ must jump back to 0. Thus, Arg z has a jump discontinuity. On the other hand, if we insist on choosing θ so that it varies continuously with z, then its final value has to be 2π, a different choice from $[\arg 1]$ from that we made at the start. We can give a more formal treatment of the issues just discussed.

We show that there is no way to impose a restriction which selects $\theta(z) \in [\arg z]$ for all $z \in \mathbb{C} \setminus \{0\}$, so $\theta : z \mapsto \theta(z)$ is continuous as a function of z. We assume for a contradiction that such a continuous function

θ does not exist and consider

$$k(t) = \frac{1}{2\pi}\left(\theta(\mathrm{e}^{it}) + \theta(\mathrm{e}^{-it})\right), \ \ t \in \mathbb{R}.$$

Then k is continuous and

$$k(t) = \frac{1}{2\pi}\left((t + 2m_t\pi) + (-t + 2n_t\pi)\right), \ \ m_t, n_t \in \mathbb{Z},$$

so k takes only integer values. Also $k(0)$ is even and $k(\pi)$ is odd, so k is non-constant. This contradicts the intermediate value theorem from real analysis.

This result has implications for other multifunctions. For example, it tells us that there cannot be a continuous logarithm in $\mathbb{C} \setminus \{0\}$: if there were one, then its imaginary part - an argument function - would be continuous too.

11.4 Branch Points

Take a multifunction $f(z)$, so that $w(z)$ is a non-empty subset of $\mathbb{C}$ for each z in the domain of definition of f. Assume that the many-valuedness arises because, for one or more points a, the definition of $f(z)$ explicitly or implicitly involves the angle θ, where $z - a = |z - a|\,\mathrm{e}^{i\theta}$. Such points are called branch points. Any branch point is excluded from the domain of definition of $f(z)$. More formally, a is a branch point for $f(z)$, if for all sufficiently small $r > 0$, it is not possible to choose a particular value of $w = f(z)$ such that f is continuous in the open ball $B(a; r)$. The motivation comes from the previous section, no continuous argument function can be drawn from $[\arg(z - a)]$ for z on a circle with centre at a.

We illustrate this with the example of $\sqrt[3]{z}$ (see fig. 11.2). Usually, we draw mappings from left to right, but here we have reversed this convention to to facilitate comparison with 11.1. As z travels along the closed

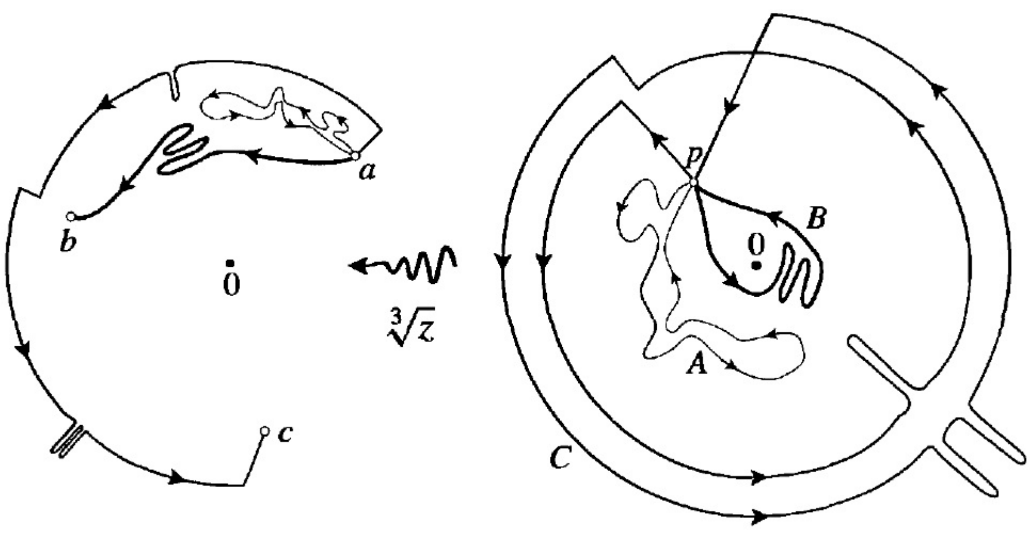


Figure 11.2

loop A (finally returning to p), $\sqrt[3]{z}$ travels along the illustrated closed loop and returns to its original value a. However, if z instead travels along the closed loop B, which goes round the origin once, then $\sqrt[3]{z}$ does not return to its original value, but instead it ends up at a different cube root of p, namely b. Note that the detailed shape of B is irrelevant, all that matters is that it encircles the origin once. Similarly, if z travels along C, encircling the origin twice, then $\sqrt[3]{z}$ ends up at c, the third and final cube root of p. Clearly, if z ere to travel along the loop (not shown) that encircled the origin three times, then $\sqrt[3]{z}$ would return to the original value a.

The premise for this picture of $z \mapsto \sqrt[3]{z}$ was the arbitrary choice of $\sqrt[3]{p} = a$, rather than b or c. If we instead chose $\sqrt[3]{p} = b$, then the orbits on the left of fig. 11.2 would simply be rotated by $2\pi/3$. Similarly, if we chose $\sqrt[3]{p} = c$, then the orbits would be rotated by $4\pi/3$.

The point $z = 0$ is the branch point of $\sqrt[3]{z}$. More generally, let $f(z)$ be a multifunction and let $a = f(p)$ be one of its values at some point $z = p$. Arbitrarily choosing the initial position of $f(z)$ to be a, we may follow the movement of $f(z)$ as z travels along a closed loop beginning and ending at p. When z returns to p, $f(z)$ will either return to a or it will not. A branch point $z = q$ of f is a point such that $f(z)$ fails to return to a as z travels along any loop that encircles q once.

Returning to the specific example $f(z) = \sqrt[3]{z}$, we have seen that if z executes three revolutions round the branch point at $z = 0$ then $f(z)$ returns to its original value. If $f(z)$ were an ordinary, single-valued function then it would return to its original value after only one revolution. Thus, relative to an ordinary function, two $extra$ revolutions are needed to restore the original value of $f(z)$. We summarize this by saying that 0 is a branch point of $\sqrt[3]{z}$ of $order\ two$.

Definition 11.4.1. If q is a branch point of some multifunction $f(z)$, and $f(z)$ first returns to its original value after N revolutions round q, then q is called an $algebraic\ branch\ point$ of order $(N-1)$; an algebraic branch point of order 1 is called a $simple\ branch\ point$. We should stress that it is perfectly possible that $f(z)$ never returns to its original value, no matter how many times z travels round q. In this case q is the $logarithmic$ $branch\ point$-the name will be explained in the next section.

By extending the above discussion of $\sqrt[3]{z}$, check for yourself that if n is an integer, then $z^{1/n}$ is an n-valued multifunction whose only (finite) branch point is at $z = 0$, the order of this branch point being $(n-1)$. More generally, the same is true for any fractional power $z^{m/n}$, where m/n is a fraction reduced to lowest terms.

11.4.1 Multibranches

Suppose we are given a multifunction $f(z)$. Our goal is to select a value $w = f(z)$ from various possible values, for each z in as large a domain as possible, so that f is holomorphic. In particular. f has to be continuous. We now introduce multibranches. These provide a stepping stone on the way to our goal.

There is a sense in which we can make continuous selections from multifunctions in a natural way. The key idea is the following. Rather than considering z as our variable we introduce, for each branch point a, new variables (r, θ), where $z = a + r\,e^{i\theta}$. Let us illustrate this with the example of $\sqrt[3]{z}$.

By arbitrarily picking one of the three values of $\sqrt[3]{p}$ at $z = p$, and then allowing z to move, we see that we obtain a unique value of $\sqrt[3]{Z}$ associated with any particular path from p to Z. However, we are still dealing with multifunction: by going round the branch point 0, we can end up at any one of the three possible values of $\sqrt[3]{Z}$.

On the other hand, the value of $\sqrt[3]{Z}$ does not depend on the detailed shape of the path: *if we continuously deform the path without crossing the branch point, then we obtain the same value of* $\sqrt[3]{Z}$. This shows us how we may obtain a single-valued function. If we restrict z to a simply connected set S that contains p, but does not contain the branch point, then every path in S from p to Z will yield the same value of $\sqrt[3]{Z}$, which we will call $f_1(Z)$. Since the path is irrelevant, f_1 is an ordinary, single-valued function of position in S, which is a branch of the $\sqrt[3]{z}$.

Fig. 11.3 illustrates such a set S, together with its image under the branch f_1 of $\sqrt[3]{z}$. Here, we have reverted to our normal practice of depicting the mapping going from left to right. If we instead choose $\sqrt[3]{p} = b$, then we obtain a second branch f_2 of $\sqrt[3]{z}$, while $\sqrt[3]{p} = c$ yields a third and final branch f_3.

We can simplify it as considering three different single-valued functions

$$f_1(r, \theta) = \sqrt[3]{r}\,e^{i\theta/3}, \quad f_2(r, \theta) = \sqrt[3]{r}\,e^{i(\theta+2\pi)/3}, \quad f_3(r, \theta) = \sqrt[3]{r}\,e^{i(\theta+4\pi)/3}$$

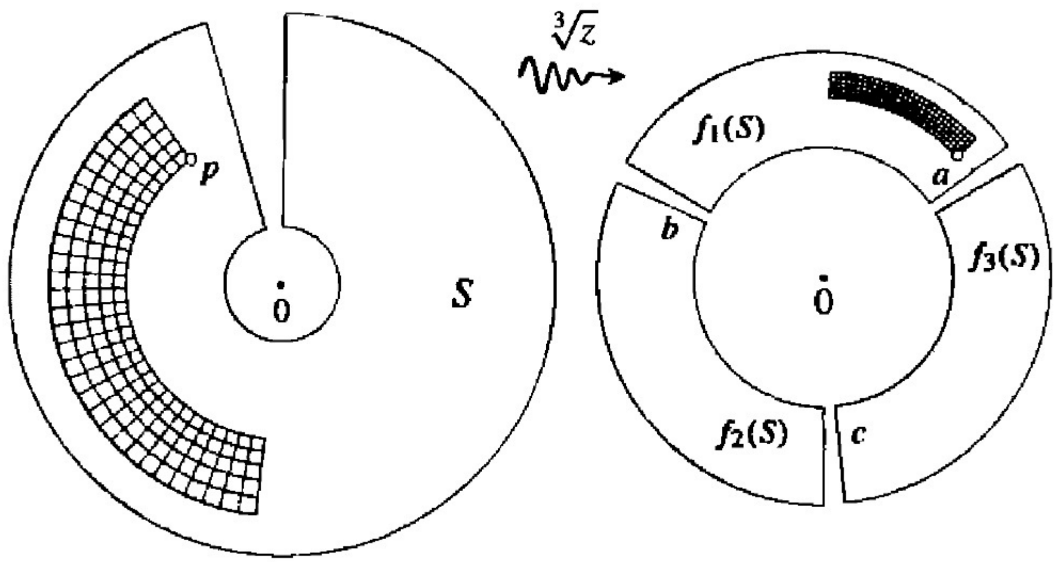


Figure 11.3

such that for $z = r\,e^{i\theta}$,

$$\begin{aligned} \sqrt[3]{z} &= f_1(r,\theta) \quad \text{if } \theta \in [0, 2\pi) \\ &= f_2(r,\theta) \quad \text{if } \theta \in [2\pi, 4\pi) \\ &= f_3(r,\theta) \quad \text{if } \theta \in [4\pi, 6\pi). \end{aligned}$$

And, as z circles the origin and θ changes from 4π through 6π, $\sqrt[3]{z}$ jumps back to the previous branch $f_1(r,\theta)$.

11.4.2 Branch Cuts

How can we prevent the interchange of f_3 to f_1? f_3 changes to f_1 due to the change of argument θ. So, if we restrict the movement of θ, then we can remain in a particular branch of $\sqrt[3]{z}$, which is single-valued and holomorphic. We draw an arbitrary curve C from the branch point 0 to infinity. We are now restricting the domain of z which includes all points of S excluding those on C. This prevents the closed path in S from encircling the branch point; or simplifying, we can say that we restrict the increase of θ by a value of 2π since changing the value of θ by 2π changes the branch of $\sqrt[3]{z}$.

This C is called a *branch cut*. As we have just seen, this shows up in the fact that the resulting branches are discontinuous on C, despite the fact that the three values of $\sqrt[3]{z}$ move continuously as z move continuously. As z crosses C travelling counterclockwise, then we must switch from one branch to the next in order to maintain continuous motion of $\sqrt[3]{z}$. If z executes three counterclockwise revolutions round the branch point, then the branches permute cyclically as shown in fig. 11.4.

A common choice for C is the negative real axis. If we do not allow z to cross the cut, then we restrict the angle θ to lie in the range $-\pi < \theta \leq \pi$. This is called the *principal value of argument*, written as Arg z, as we have encountered in the beginning of this unit. With this choice of θ, the single-valued function $\sqrt[3]{r}\,e^{i\theta/3}$ is called the *principal branch* of the cube root. Let us denote this as $[\sqrt[3]{z}]$. Note that the principal branch agrees with the real cube root function on the positive real axis, but not the negative real axis and note that the other two branches can be expressed in terms of the principal branch as $e^{i(2\pi/3)}\,[\sqrt[3]{z}]$ and $e^{i(4\pi/3)}\,[\sqrt[3]{z}]$.

$$\left\{\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array}\right\} \to \left\{\begin{array}{c} f_2 \\ f_3 \\ f_1 \end{array}\right\} \to \left\{\begin{array}{c} f_3 \\ f_1 \\ f_2 \end{array}\right\} \to \left\{\begin{array}{c} f_1 \\ f_2 \\ f_3 \end{array}\right\}$$

Figure 11.4

Few Probable Questions

1. Which of the following are multifunctions?

   A. $\cos z$ B. $\sqrt{z}$ C. e^z D. None

2. The branch point of the function $\log(z)$ is at

   A. 0 B. 1 C. ∞ D. -1

3. The function $\sqrt[3]{z^2}$ has an algebraic branch point of order ...... at $z = 0$.

   A. $3/2$ B. $2/3$ C. 2 D. 3

4. For the function $\sqrt{z-1}$, the point $z = 1$ is a/an ...... branch point.

   A. transcendental B. logarithmic C. algebraic D. not a branch point

5. Define branch of a multifunction. Show that the branches of the square root function are discontinuous at each point of the negative real axis.

Unit 12

Course Structure

- Riemann surface for the functions $\sqrt{z}$, $\log z$.

12.1 Introduction

Riemann surfaces were first studied by Bernhard Riemann. Riemann surfaces are the easiest way to geometrically understand the multifunctions. The main interest in Riemann surfaces is that holomorphic functions may be defined between them. We discuss the Riemann surfaces of two major functions: the complex square root function and the logarithm function.

Objectives

After reading this unit, you will be able to

- discuss the Riemann surfaces of the square root function and logarithm function

12.2 Riemann Surfaces

12.2.1 Square Root function

Just as we have discussed the case of $\sqrt[3]{z}$, we can simply show that $w = u+iv = f(z) = \sqrt{z}$ is a multifunction having a branch at 0. We define two branches

$$f_1(z) = \sqrt{r}\,\mathrm{e}^{i\theta/2}, \qquad f_2(z) = \sqrt{r}\,\mathrm{e}^{i(\theta+2\pi)/2} = -\,\mathrm{e}^{i\theta/2} = -f_1(z);$$

so f_1 and f_2 can be thought of as ”plus” and ”minus” square root functions. The negative real axis is called a branch cut for the functions f_1 and f_2. Each point on the branch cut is a point of discontinuity for both functions f_1 and f_2.

Example 12.2.1. We show that f_1 is discontinuous along the negative real axis. Let $z_0 = r\,\mathrm{e}^{i\pi}$ denote a negative real number. We compute the limit as z approaches z_0 through the upper half plane and the limit as

z approaches z_0 through the lower half plane. In polar coordinates, these are given by

$$\lim_{(r,\theta)\to(r_0,\pi)} f_1(r\,e^{i\theta}) = \lim_{(r,\theta)\to(r_0,\pi)} \sqrt{r_0}\left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right) = i\sqrt{r_0}, \quad \text{and}$$

$$\lim_{(r,\theta)\to(r_0,-\pi)} f_1(r\,e^{i\theta}) = \lim_{(r,\theta)\to(r_0,-\pi)} \sqrt{r_0}\left(\cos\frac{\theta}{2} + i\sin\frac{\theta}{2}\right) = -i\sqrt{r_0}.$$

The two limits are distinct, so f_1 is discontinuous at z_0. Since z_0 is arbitrary, so f_1 is discontinuous on the whole negative real axis.

We will now draw the Riemann surface for f. $f(z)$ has two values for any $z \neq 0$. Each functions f_1 and f_2 are single-valued on the domain formed by cutting the z plane along the negative real axis. Let D_1 (see fig. 12.1) and D_2 (see fig. 12.2) be the domains of f_1 and f_2 respectively. The range set for f_1 is H_1 consisting of the right-half plane and the positive v-axis; and the range set for f_2 is H_2 consisting of the left-half plane and the negative v-axis. The sets H_1 and H_2 are "glued together" along the positive v-axis and the negative v-axis to form the w plane with the origin deleted.

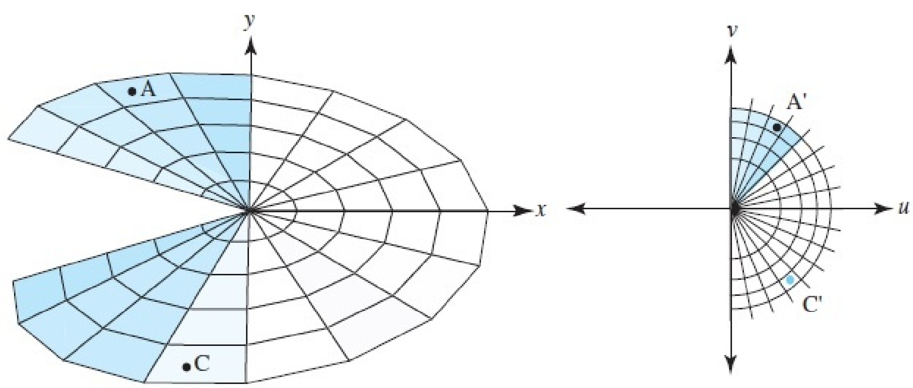


Figure 12.1: A portion of D_1 and its image under $w = \sqrt{z}$

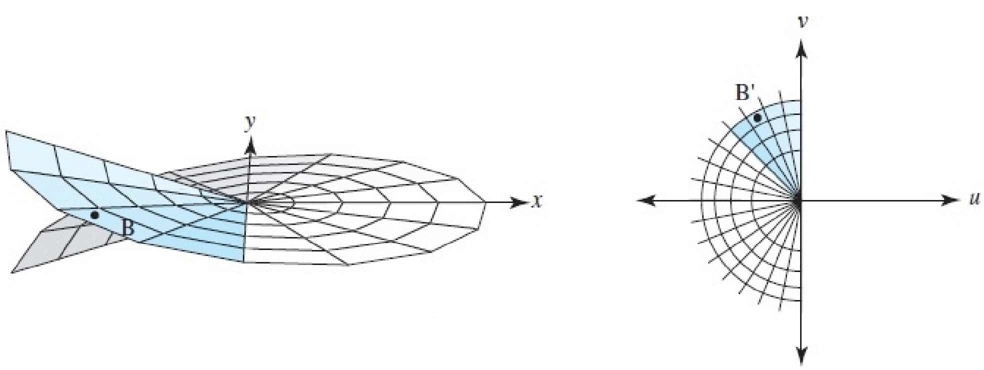


Figure 12.2: A portion of D_2 and its image under $w = \sqrt{z}$

We stack D_1 directly above D_2. The edge of D_1 in the upper half-plane is joined to the edge of D_2 in the lower half-plane, and the edge of D_1 in the lower half-plane is joined to the edge of D_2 in the upper half-plane. When these domains are glued together in this manner, they form R, which is a Riemann surface domain for

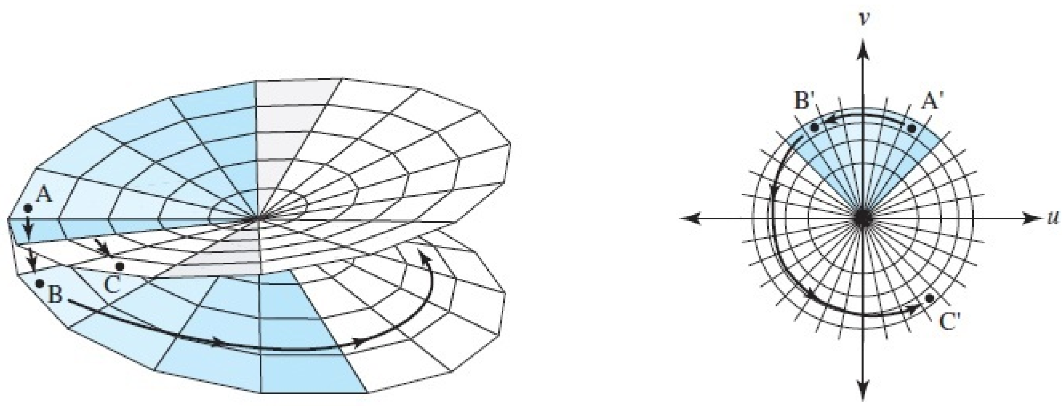


Figure 12.3: A portion of R and its image under $w = \sqrt{z}$

the mapping $w = f(z) = \sqrt{z}$. The portions of $D_1,\ D_2$ and R that lie in $\{z :\ |z| < 1\}$ are shown in fig. 12.3.

The beauty of this structure is that it makes this "full square root function" continuous for all $z \neq 0$. Normally, the principal square root function would be discontinuous along the negative real axis, as points near -1 but above that axis would get mapped to points close to i, and points near -1 but below the axis would get mapped to points close to $-i$. As fig. 12.3 indicates, however, between the point A and the point B, the domain switches from the edge of D_1 in the upper half-plane to the edge of D_2 in the lower half plane. The corresponding mapped points A' and B' are exactly where they should be. The surface works in such a way that going directly between the edges of D_1 in the upper and lower half planes is impossible (likewise for D_2). Going counterclockwise, the only way to get from the point A to the point C, for example, is to follow the path indicated by the arrows in fig. 12.3.

We now move on to the logarithmic function.

Exercise 12.2.1. Show that f_2 is discontinuous at every point on the negative real axis.

12.2.2 Logarithm Function

The complex logarithm function $\log(z)$ may be introduced as the "inverse" of e^z. More precisely, we define $\log z$ to be any complex number z that satisfies

$$\mathrm{e}^{\log z} = z.$$

It follows that

$$\log z = \ln|z| + i\arg(z).$$

Since $\arg(z)$ takes infinitely many values, differing from each other by multiples of 2π, we see that $\log(z)$ is a multifunction taking infinitely many values, differing from each other by multiples of $2\pi i$. For example,

$$\log(2+2i) = \ln 2\sqrt{2} + i\frac{\pi}{4} + 2n\pi i,$$

where n is an arbitrary integer. The reason we get infinitely many values is clear if we see the behaviour of the exponential function e^z. Each time z travels straight upwards by $2\pi i$, e^z executes a complete revolution and returns to its original value. Clearly, $\log(z)$ has a branch point at 0. However, this branch point is quite unlike

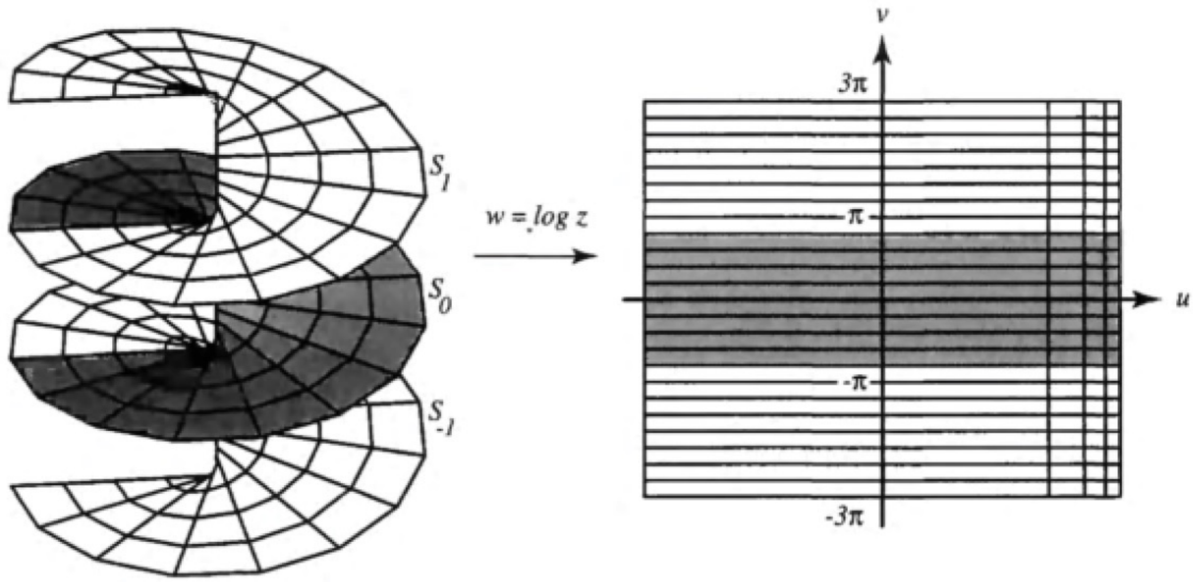


Figure 12.4

that of $\sqrt[n]{z}$, for no matter how many times we loop around the origin, $\log(z)$ never returns to its original value, rather it continues moving upwards forever. You can now understand previously introduced term ”logarithmic branch point”.

Here is another difference between the branch point of $\sqrt[n]{z}$ and $\log(z)$. As z approaches the origin, say along a ray, $|\sqrt[n]{z}|$ tends to zero, but $|\log(z)|$ tends to infinity, and in this sense, origin is a singularity as well as a branch point.

To define single-valuedness of $\log(z)$, we make a branch cut from 0 out to infinity. The most common choice for this cut is the negative real axis. In this cut plane, we may restrict $\arg(z)$ to its principal value Arg z. This yields the principal branch of logarithm, written as Log z, defined as

$$\text{Log } z = \ln|z| + \text{Arg } z.$$

We are in a position to draw the Riemann surface for $\log(z)$. Let define the multibranches of logarithm function as follows

$$F_k(z) = \log|z| + i(\theta + 2k\pi), \quad k \in \mathbb{Z}.$$

Each F_k is a continuous function of z. Furthermore, for any fixed $c \in \mathbb{R}$, and for $0 \neq z = r\,\mathrm{e}^{i\theta}$,

$$\log(z) = \{F_k(z) : \quad k \in \mathbb{Z}, \quad \theta \in [c, c + 2\pi)\},$$

with no values repeated, and if similarly, θ is restricted to any interval $(c, c + 2\pi]$.

The domain of definition of each F_k are none other than the whole complex plane with different arguments. Let S_k be the domain of definition of the branch F_k of $\log(z)$. (These copies of the complex plane are each cut along the negative real axis) These cut planes are then stacked directly on top of each other and joined as follows. For each integer k, the edge of S_k in the upper half plane is joined to the edge of S_{k+1} in the lower half plane. The resulting Riemann surface of $\log(z)$ looks like a spiral staircase that extends upwards on $S_1, S_2, \ldots$ and downwards on $S_{-1}, S_{-2}, \ldots$ as shown in fig. 12.4. If we start on S_0 and make a counterclockwise circuit around the origin, we end up on S_1, and the next circuit brings us to S_2, etc, so each time we cross the negative real axis, we end up on a new branch of $\log(z)$.

Unit 13

Course Structure

- Analytic continuation, uniqueness of analytic continuation

13.1 Introduction

Analytic continuation is an important idea since it provides a method for making the domain of definition of an analytic function as large as possible. Usually, analytic functions are defined by means of some mathematical expressions such as polynomials, infinite series, integrals, etc. The domain of definition of such an analytic function is often restricted by the manner of defining the function. For instance the power series representation of such analytic functions does not provide any direct information as to whether we could have a function analytic in a domain larger than disc of convergence which coincides with the given function. We have previously seen that an analytic function is determined by its behaviour at a sequence of points having limit point. This was precisely the content of the identity theorem which is also referred to as the principle of analytic continuation. For example, as a consequence, there is precisely a unique entire function on $\mathbb{C}$ which agrees with $\sin x$ on the real axis, namely $\sin z$.

Objectives

After reading this unit, you will be able to

- define analytic continuation of an analytic function and consider examples of such process
- show that the analytic continuation of an analytic function is always unique
- define chain and function elements and discuss the condition for analytic continuation from a domain into another
- discuss the power series method and its examples

13.2 Analytic Continuation

As we have seen in the introduction that if G is a region and f is an analytic function on G. Also, let g be an analytic function defined on an open set $D \subseteq G$, such that $f(z) = g(z)$ in D. Then, by uniqueness theorem (identity theorem), we have $f \equiv g$ on G. It is a simple process to extend the domain of g over to G. A natural question is the following: Is it always possible to have such an extension? Clearly that is not the case. For example,

$$f(z) = \frac{1}{z}, \ \ z \in \mathbb{C} \setminus \{0\}$$

does not have an extension to $\mathbb{C}$. Similarly, if we take two domains

$$D_1 = \mathbb{C} \setminus \{z : \text{ Re } z \leq 0, \text{ Im } z = 0\}, \quad \text{and} \quad D_2 = \mathbb{C},$$

then for $f(z) = \text{Log }(z)$, which is the principal branch of logarithm function that is analytic on the domain D_1, but can't be extended on to D_2.

However, if the extension is possible, there are ways to carry out the process of continuation so that the given analytic function becomes analytic on a larger domain. To make this point more precise, let us start by examining the analytic continuation of the function

$$f(z) = \sum_{n \geq 0} z^n. \tag{13.2.1}$$

The series on the right hand side of (13.2.1), as is well known, is convergent for $|z| < 1$ and diverges for $|z| \geq 1$. On the other hand, we know that the series given by the formula (13.2.1) represents an analytic function for $|z| < 1$ and the sum of the series (13.2.1) for $|z| < 1$ is $1/(1-z)$. However, the function F defined by the formula

$$F(z) = \frac{1}{1-z}$$

is analytic for $z \in \mathbb{C}_\infty \setminus \{1\} = D$, since

$$F\left(\frac{1}{z}\right) = \frac{1}{1-z^{-1}} = \frac{z}{z-1}$$

is analytic at $z = \infty$. Now, $f(z) = F(z)$ for all $z \in \mathbb{D} \cap D$, and we call F an analytic continuation of f from $\mathbb{D}$ into D, that is, the function f, given at first for $|z| < 1$, has been extended to the extended complex plane but for the point 1, at which the function has a simple pole. Thus, it seems that F, which is analytic globally, is represented by a power series only locally.

We now, formally define analytic continuation of a function f as follows.

Definition 13.2.1. Suppose that f and F are two functions such that

1. f is analytic on some domain $D \subset \mathbb{C}$;
2. F is analytic in a domain D_1 such that $D_1 \cap D \neq \emptyset$ and $D \subset D_1$, such that $f(z) = F(z)$ for $z \in D \cap D_1$.

Then we call F an analytic continuation or holomorphic extension of f from domain D into D_1. In other words, f is said to be analytically continuable into D_1.

The definition can also be given as follows.

Definition 13.2.2. A function $f(z)$, together with a domain D in which it is analytic, is said to be a **function element** and is denoted by (f, D). Two function elements (f_1, D_1) and (f_2, D_2) are called **direct analytic continuations** of each other if and only if

$$D_1 \cap D_2 \neq \emptyset \quad \text{and} \quad f_1 = f_2 \ \text{ on } \ D_1 \cap D_2.$$

Remark 13.2.1. Whenever there exists a direct analytic continuation of (f_1, D_1) into a domain D_2, it must be uniquely determined, for any two direct analytic continuations would have to agree on $D_1 \cap D_2$, and by identity theorem, would consequently have to agree throughout D_2. That is, given an analytic function f_1 on D_1, there is at most one way to extend f_1 from D_1 into D_2 so that the extended function is analytic in D_2.

The property of being a direct analytic continuation is not transitive. That is, even if (f_1, D_1) and (f_2, D_2) are direct analytic continuations of each other, and (f_2, D_2) and (f_3, D_3) are direct analytic continuations of each other, we cannot conclude that (f_1, D_1) and (f_3, D_3) are direct analytic continuations of each other. A simple example of this occurs whenever D_1 and D_3 have no points in common. However, there is a relationship between $f_1(z)$ and $f_3(z)$ that is worth explaining.

Definition 13.2.3. Suppose that $\{(f_1, D_1), (f_2, D_2), \ldots, (f_n, D_n)\}$ is a finite set of function elements with the property that (f_k, D_k) and (f_{k+1}, D_{k+1}) are direct analytic continuations of each other for $k = 1, 2, \ldots, n-1$. Then the set of function elements are said to be analytic continuations of one another. Such a set of function elements is then called a **chain**.

Example 13.2.1. Consider the figure 13.1. Let

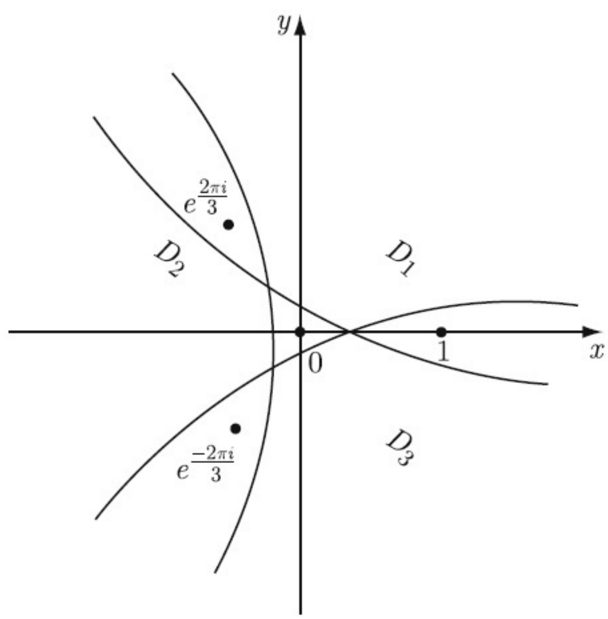


Figure 13.1

$$
\begin{aligned}
f_1(z) &= \text{Log}\, z, \quad z \in D_1 \\
f_2(z) &= \text{Log}\, z, \quad z \in D_2 \\
f_3(z) &= \text{Log}\, z + 2\pi i, \quad z \in D_3.
\end{aligned}
$$

Then $\{(f_1, D_1), (f_2, D_2), (f_3, D_3)\}$ is a chain with $n = 3$. Note that $0 = f_1(1) \neq f_3(1) = 2\pi i$.

Note that (f_i, D_i) and (f_j, D_j) are analytic continuations of each other if and only if they can be connected by finitely many direct analytic continuations.

13.3 Analytic Continuation along a curve

Definition 13.3.1. If $\gamma : [0, 1] \to \mathbb{C}$ is a curve and if there exists a chain $\{(f_i, D_i)\}_i$ of function elements such that

$$\gamma([0, 1]) \subset \cup_{i=1}^{n} D_i, \;\; z_0 = \gamma(0) \in D_1, \;\; z_n = \gamma(1) \in D_n,$$

then we say that the function element (f_n, D_n) is an analytic continuation of (f_1, D_1) along the curve γ. That is a function element (f, D) can be analytically continued along a curve if there is a chain containing (f, D) such that each point on the curve is contained in the domain of some function element of the chain.

As another example, the domains of a chain are also shown in Figure 13.2.

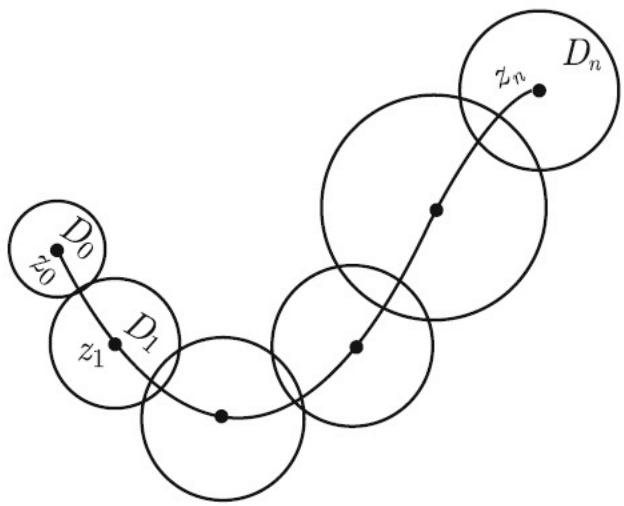


Figure 13.2

The definition can also be given as

Definition 13.3.2. Let $\gamma : [a, b] \to \mathbb{C}$ be a curve such that

$$\gamma(t) = z(t) = x(t) + iy(t), \;\; a \le t \le b.$$

Let us consider a partition $a = t_0 < t_1 < \ldots < t_n = b$ of $[a, b]$. If there is a chain $\{(f_1, D_1), (f_2, D_2), \ldots, (f_n, D_n)\}$ of function elements such that (f_{k+1}, D_{k+1}) is a direct analytic continuation of (f_k, D_k) for $k = 1, 2, \ldots, n-1$ and $z(t) \in D_k$ for $t_{k-1} \le t \le t_k$, $k = 1, 2, \ldots, n$, then (f_n, D_n) is said to be an analytic continuation of (f_1, D_1) along the curve γ.

Thus we shall obtain a well defined analytic function in a nbd. of the end point of the path, which is called the analytic continuation of (f_1, D_1) along the path γ. Here, D_k may be taken as discs containing $z(t_{k-1})$. Further, we say that the sequence $\{D_1, D_2, \ldots, D_n\}$ connected by the curve γ along the partition if the image $z([t_{k-1}, t_k])$ is contained in D_k.

Theorem 13.3.1. (Uniqueness of Analytic Continuation along a Curve) Analytic continuation of a given function element along a given curve is unique. In other words, if (f_n, D_n) and (g_m, D_m) are two analytic continuations of (f_1, D_1) along the curve γ defined by

$$\gamma(t) = z(t) = x(t) + iy(t), \;\; a \le t \le b.$$

Then $f_n = g_m$ on $D_n \cap E_m$.

Proof. Suppose there are two analytic continuations of (f_1, D_1) along the curve γ, namely,

$$(f_1, D_1), (f_2, D_2), \dots, (f_n, D_n)$$
$$(g_1, E_1), (g_2, E_2), \dots, (g_m, E_m)$$

where $f_1 = g_1$ and $E_1 = D_1$. Then there exist partitions

$$a = t_0 < t_1 < \dots < t_n = b$$
$$a = s_0 < s_1 < \dots < s_m = b$$

such that $z(t) \in D_i$ for $t_{i-1} \le t \le t_i$, for $i = 1, 2, \dots, n$ and $z(t) \in E_j$ for $s_{j-1} \le t \le s_j$ for $j = 1, 2, \dots, m$. We claim that if $1 \le i \le n$, $1 \le j \le m$ and

$$[t_{i-1}, t_i] \cap [s_{j-1}, s_j] \ne \emptyset$$

then (f_i, D_i) and (g_j, E_j) are direct analytic continuations of each other. This is certainly true when $i = j = 1$, since $f_1 = g_1$ and $E_1 = D_1$. If it is not true for all i and j, then we may pick from all (i, j), for which the statement is false and such that $i + j$ is minimal. Suppose that $t_{i-1} \ge s_{j-1}$, where $i \ge 2$. Since $[t_{i-1}, t_i] \cap [s_{j-1}, s_j] \ne \emptyset$ and $s_{j-1} \le t_{i-1}$, we must have $t_{i-1} \le s_j$. Thus, $s_{j-1} \le t_{i-1} \le s_j$. It follows that $z(t_{i-1}) \in D_{i-1} \cap E_i \cap E_j$. In particular, this intersection is non-empty. None of (f_i, D_i) is a direct analytic continuation of (f_{i-1}, D_{i-1}). Moreover, (f_{i-1}, D_{i-1}) is a direct analytic continuation of (g_j, E_j) since $i+j$ is minimal, where we observe that $t_{i-1} \in [t_{i-2}, t_{i-1}] \cap [s_{j-1}, s_j]$ so that the hypothesis of the claim is satisfied. Since $D_{i-1} \cap D_i \cap E_j \ne \emptyset$, (f_i, D_i) must be direct analytic continuation of (g_j, E_j) which is a contradiction. Hence our claim holds for all i and j. In particular, it holds for $i = n$ and $j = m$, which proves the theorem. $\square$

Given a chain $\{(f_1, D_1), (f_2, D_2), \dots, (f_n, D_n)\}$, can a function $f(z)$ be defined such that $f(z)$ is analytic in the domain $D_1 \cup D_2 \cup \dots \cup D_n$? Certainly this can be done when $n = 2$. The function

$$\begin{aligned} f(z) &= f_1(z) \text{ if } z \in D_1 \\ &= f_2(z) \text{ if } z \in D_2 \end{aligned}$$

is analytic in $D_1 \cup D_2$. If $D_1 \cap D_2 \cap \dots \cap D_n \ne \emptyset$, we can show by induction that f, defined by $f(z) = f_i(z)$ for $z \in D_i$, $i = 1, 2, \dots, n$ is analytic.

However, the proof for the general case fails. Consider the four domains illustrated in Figure 13.3.

For a fixed branch of $\log z$, set $f_1(z) = \log z$ in D_1. The function element (f_1, D_1) determines a unique direct analytic continuation (f_2, D_2) which determines (f_3, D_3) which determines (f_4, D_4). We thus have the chain $\{(f_1, D_1), (f_2, D_2), (f_3, D_3), (f_4, D_4)\}$. However, in the domain $D_1 \cap D_4$, it is not true that $f_1(z) = f_4(z)$. We actually have $f_4(z) = f_1(z) + 2\pi i$ for all points in $D_1 \cap D_4$. The difference in the two functions lies in the fact that the argument of the multiple-valued logarithmic function has increased by 2π after making a complete revolution around the origin. Note also that we can continue (f_1, D_1) into the domain D_3 by different chains and come up with different functions. For the chains $\{(f_1, D_1), (f_2, D_2), (f_3, D_3)\}$ and $\{(f_1, D_1), (g_1, D_4), (g_2, D_3)\}$, we have the values of f_3 and g_2 differing by $2\pi i$. We shall continue this discussion which ultimately culminates into Monodromy theorem, in the upcoming units.

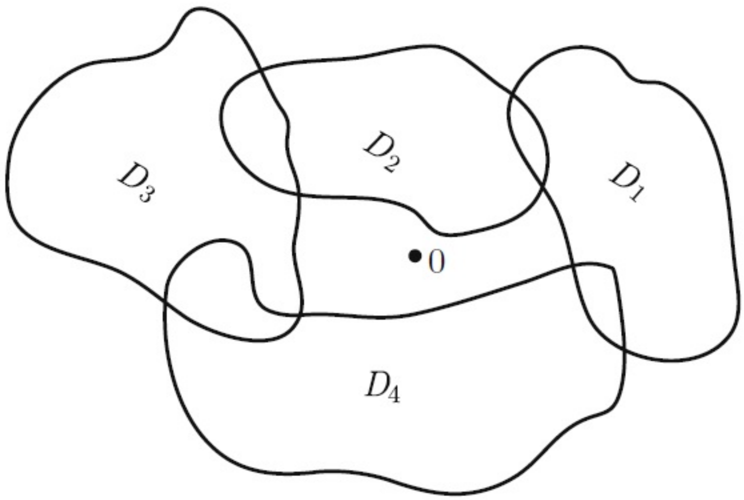


Figure 13.3

Few Probable Questions

1. Define analytic continuation of an analytic function f from a domain D_1 into another domain D_2. Show that such a continuation is unique. Is the analytic continuation of an analytic function always possible? Justify your answer.

2. Define analytic continuation along a curve. Show that it is unique.

3. Let $\{(f_1, D_1), (f_2, D_2), \ldots, (f_n, D_n)\}$ be a chain. With proper justifications, show that a function defined by $f(z) = f_i(z)$, for $z \in D_i$ is analytic in $D_1 \cup D_2 \cup \cdots \cup D_n$ if $D_1 \cap D_2 \cap \cdots \cap D_n \neq \emptyset$. Is the result true for any general case? Justify.

Unit 14

Course Structure

- Continuation by the method of power series

14.1 Introduction

We have seen that for a given analytic function f_1 on D_1, if there exists an analytic continuation f_2 on D_2, then it is unique. When does a power series represent a function which is analytic beyond the disc of convergence of the original series? One way to provide an affirmative answer is by the power series method. Let us start our discussion on this method and see how one can use the power series to go beyond the boundary of the disc of convergence.

Objectives

After reading this unit, you will be able to

- discuss the power series method and its examples

14.2 Power Series Method

A fundamental fact about a function f, analytic in a domain D, is that, for each $a \in D$, there exists a sequence $\{a_n\}_{n\geq 0}$ and a number $r_a \in (0, \infty]$ such that

$$f(z) = \sum_{n=0}^{\infty} a_n (z-a)^n \quad \text{for all } \ z \in B(a; r_a).$$

To extend f, we choose a point b other than a in the disc of convergence $B(a;r_a)$. Then $|b-a|<r_a$ and

$$\begin{aligned}\sum_{n=0}^{\infty} a_n(z-a)^n &= \sum_{n=0}^{\infty} a_n[z-b+b-a]^n \\ &= \sum_{n=0}^{\infty} a_n\left(\sum_{k=0}^{n}\binom{n}{k}(b-a)^{n-k}(z-b)^k\right) \\ &= \sum_{k=0}^{\infty}\left(\sum_{n=k}^{\infty} a_n\binom{n}{k}(b-a)^{n-k}\right)(z-b)^k \\ &= \sum_{k=0}^{\infty} A_k(z-b)^k.\end{aligned}$$

The interchange of summation is justified since

$$\sum_{n=0}^{\infty}|a_n|\sum_{k=0}^{n}\binom{n}{k}|b-a|^{n-k}|z-b|^k = \sum_{n=0}^{\infty}|a_n|\left(|z-b|+|b-a|\right)^n<\infty$$

whenever $|z-b|+|b-a|<r_a$. Therefore, the series about b converges at least for $|z-b|<r_a-|b-a|$. However, this may happen that the disc of convergence $B(b;r_b)$ for this new series extends outside $B(a;r_a)$, that is, it may be possible that $r_b>r_a-|b-a|$. In this case, the function can be analytically continued to the union of these two discs. This process may be continued.

Example 14.2.1. Let

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

Then, for $z\in\mathbb{D}$, with $a=0$, $r_a=1$, we have

$$\frac{1}{1-z}=\sum_{n\geq 0} z^n,\quad z\in\mathbb{D}.$$

Take $b=i$. In order to get the expression for $z\in B(i;r_b)$, we write

$$\begin{aligned}\frac{1}{1-z} &= \frac{1}{1-i-(z-i)} \\ &= \sum_{n=0}^{\infty}\frac{1}{(1-i)^{n+1}}(z-i)^n,\ \ |z-i|<|1-i|=\sqrt{2}, \\ &= \sum_{n=0}^{\infty} A_n(z-i)^n,\quad A_n=\frac{(1+i)^{n+1}}{2^{n+1}},\ |z-i|<r_b=\sqrt{2}.\end{aligned}$$

Thus, $\sum_{n=0}^{\infty} A_n(z-i)^n$ is an analytic continuation of $\sum_{n=0}^{\infty} z^n$ in $\mathbb{D}$ to the disc $B(i;\sqrt{2})$. Similarly, one can see that $\sum_{n=0}^{\infty}\frac{1}{2^{n+1}}(z+1)^n$ is an analytic continuation of $\sum_{n=0}^{\infty} z^n$ from $\mathbb{D}$ to the disc $B(-1;2)$.

Example 14.2.2. We show that the function

$$f(z)=\frac{1}{a}+\frac{z}{a^2}+\frac{z^2}{a^3}+\cdots$$

can be continued analytically. This series converges within the circle $C_0 : |z| = |a|$ and has the sum

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

The only singularity of $f(z)$ on C_0 is at $z = a$. Hence the analytic continuation of $f(z)$ beyond C_0 is possible. For this purpose we take a point $z = b$ not lying on the line segment joining $z = 0$ and $z = a$. We draw a circle C_1 with centre b and radius $|a - b|$ that is, C_1 is $|z - b| = |a - b|$. This new circle C_1 clearly extends beyond C_0 as shown in the figure 14.1

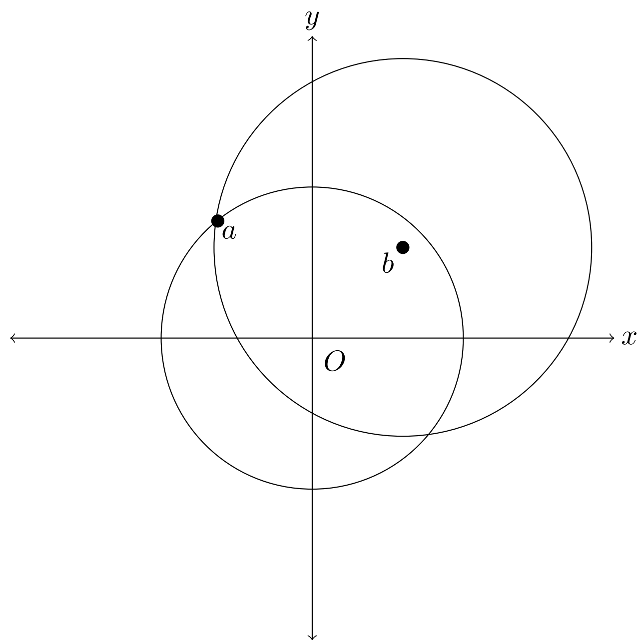


Figure 14.1

Now, we reconstruct the power series given in powers of $(z - b)$ in the form

$$\sum_{n=0}^{\infty} \frac{(z-b)^n}{(a-b)^{n+1}}, \quad \text{where } f^{(n)}(b) = \frac{n!}{(a-b)^{n+1}}. \tag{14.2.1}$$

This power series has the circle of convergence C_1 and has cum function $\dfrac{1}{a-z}$. Thus, the power series (14.2.1) and the given power series represent the same function in the region common to C_0 and C_1. Hence (14.2.1) represents an analytic continuation of the given series.

Few Probable Questions

---

1. Check whether the function

$$f(z) = \frac{1}{a} + \frac{z}{a^2} + \frac{z^2}{a^3} + \dots$$

can be continued analytically.

2. Show that the functions defined by

$$a + az + a^2z^2 + \dots$$
$$\frac{1}{1-z} - \frac{(1-a)z}{(1-z)^2} + \frac{(1-a)^2z^2}{(1-z)^3}$$

are analytic continuations of each other.

Unit 15

Course Structure

- Continuation by the method of natural boundary,
- Existence of singularity on the circle of convergence

15.1 Introduction

This unit deals with the continuation by natural boundary. Suppose that a power series has radius of convergence R and defines an analytic function f inside that disc. Consider points on the circle of convergence. A point for which there is a neighbourhood on which f has an analytic extension is regular, otherwise singular. Convergence is limited to within by the presence of at least one singularity on the boundary of . If the singularities on are so densely packed on the circle, that analytic continuation cannot be carried out on a path that crosses, then it is said to form a natural boundary. In particular, the circle is a natural boundary if all its points are singular. More generally, we may apply the definition to any open connected domain on which f is analytic, and classify the points of the boundary of the domain as regular or singular: the domain boundary is then a natural boundary if all points are singular. We will study about this in details.

Objectives

After reading this unit, you will be able to

- define natural boundary of an analytic function f on a domain D
- deduce certain results on the existence of singularities on the circle of convergence

15.2 Continuation by method of natural boundary

We start by the definition of natural boundary.

Definition 15.2.1. (Natural Boundary) Let f be analytic on a domain D. If f cannot be continued analytically across the boundary ∂D, then ∂D is called $natural\ boundary$ of f. A point $z_0 \in \partial D$ is said to be a regular point of $f(z)$ if f can be continued analytically to a region D_1 with $z_0 \in D_1$. Otherwise, $f(z)$ is said to have a $singular\ point$ at z_0.

Example 15.2.1. Consider the power series

$$f(z) = \sum_{k\geq 0} z^{2^k} \tag{15.2.1}$$

A direct consequence of the Root test is that the radius of convergence of the above series is 1 and so, $f(z)$ defined as above is analytic for $|z| < 1$. If $|z| \geq 1$, then $\lim_{n\to\infty} |z^{2n}| \neq 0$ is therefore, the series diverges for $|z| \geq 1$.

Let $\zeta = \mathrm{e}^{2\pi i m/2^n}$, $m = 0, 1, 2, \ldots, 2^n - 1$, $(n \in \mathbb{N})$ be the 2^nth root of unity. If $z = r\,\mathrm{e}^{2\pi i m/2^n} \in \mathbb{D}$, then

$$f(z) = \sum_{k=0}^{n-1} z^{2^k} + \sum_{k=n}^{\infty} z^{2^k}$$

and so for $r \to 1^-$, we have

$$|f(\zeta r)| \geq \sum_{k=n}^{\infty} r^{2^k} - \left| \sum_{k=0}^{n-1} z^{2^k} \right| \geq \sum_{k=n}^{\infty} r^{2^k} - n,$$

and hence, for every 2^nth root of ζ, we have,

$$\lim_{r\to 1^-} |f(\zeta r)| = \infty.$$

Therefore, if D is a domain containing points of $\mathbb{D}$ and of its complement, then D contains the points $\zeta = \mathrm{e}^{2\pi i m/2^n}$ and so any function F in D which coincides with f in $\mathbb{D} \cap D$ cannot be continued analytically through $\zeta^{2n} = 1$ for each $n \in \mathbb{N}$. In other words, any root of the equation

$$z^2 = 1, \;\; z^4 = 1, \;\; \ldots, z^{2n} = 1 \;\; (n \in \mathbb{N})$$

is a singular point of f and hence any arc, however small it may be, of $\partial\mathbb{D}$ contains an infinite number of singularities. Thus, f on $\mathbb{D}$ cannot be continued analytically across the boundary $\partial\mathbb{D}$ of $\mathbb{D}$. This observation shows that the unit circle $|z| = 1$ is a natural boundary for the power series defined by (15.2.1).

Example 15.2.2. Similarly, if

$$f(z) = \sum_{k\geq 0} z^{k!} \tag{15.2.2}$$

then f is analytic in $\mathbb{D}$. Upon taking $\zeta = \mathrm{e}^{2\pi i m/n}$, $m = 0, 1, 2, \ldots, n-1$, $z = r\zeta$,(where m/n is the irreducible fraction), and choosing r close to 1 from below along a radius of the unit circle it can be seen that $\lim_{r\to 1^-} |f(\zeta r)| = \infty$. Hence, f is singular at every n-th root of unity for any $n \in \mathbb{N}$. Since, every point on $|z| = 1$ is a singular point, f cannot be continued analytically through the n-th root of unity for any natural number n. In other words, there can be no continuation anywhere across $|z| = 1$ and hence, $|z| = 1$ is a natural boundary for the power series defined by (15.2.2).

15.3 Existence of singularities on the circle of convergence

Theorem 15.3.1. If $f(z) = \sum_{n\geq 0} a_n z^n$ has a radius of convergence $R > 0$, then f must have at least one singularity on $|z| = R$.

Proof. Suppose, on contrary that f has no singularity on $|z| = R$. Then f must be analytic at all points of $|z| = R$. This implies f is analytic on $|z| \leq R$. It follows, from the definition of analyticity at a point, that for each $\zeta \in \partial\mathbb{D}_R$ there exists for some $R_\zeta > 0$ and a function f_ζ which is analytic in $B(\zeta; R_\zeta)$ and

$$f = f_\zeta \quad \text{on } \mathrm{B}(\zeta; R_\zeta) \cap \mathbb{D}_R$$

In this way, if ζ_k and $\zeta_l \in \partial\mathbb{D}_R(k \neq l)$ with $G = B(\zeta_k; R_{\zeta_k}) \cap B(\zeta_l; R_{\zeta_l}) \neq \phi$, then we have two functions f_{ζ_k} and f_{ζ_l} which are respectively analytic in $B(\zeta_k; R_{\zeta_k})$ and $B(\zeta_l; R_{\zeta_l})$ such that

$$f = f_{\zeta_k} = f_{\zeta_l} \quad G \cap \mathbb{D}_R$$

Since G is connected and $G \cap \mathbb{D}_R$ is an open subset of G, by the uniqueness theorem, $f_{\zeta_k} = f_{\zeta_l}$ on G. Since

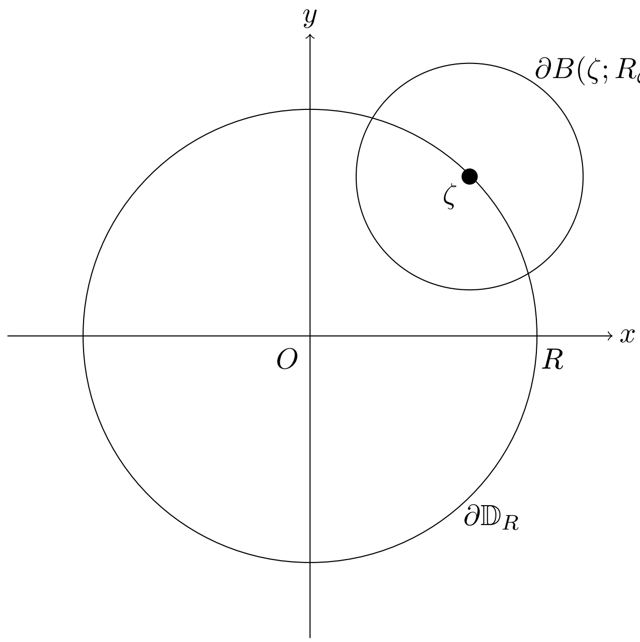


Figure 15.1: Illustration for singularity on circle $|z| = R$.

$|\zeta| = R$ is compact, by the Heine-Borel theorem, we may select a finite number of $B(\zeta_1; R_{\zeta_1}), B(\zeta_2; R_{\zeta_2}), \ldots,$ $B(\zeta_n; R_{\zeta_n})$ from the collection $\{B(\zeta; R_\zeta) : \zeta \in \partial\mathbb{D}_R\}$ such that it covers the circle $\partial\mathbb{D}_R$. Let

$$\Omega = \cup_{k=1}^n B(\zeta_k; R_{\zeta_k}) \text{ and } \delta = \text{ dist}(\partial\mathbb{D}_R\, \Omega)$$

Then, as $R_{\zeta_k} > 0$ for each k, we have $\delta > 0$. Moreover,

$$\{z : R - \delta < |z| < R + \delta\} \subset \Omega \quad \text{ and } \quad \mathbb{D}_{R+\delta} \subset D = \mathbb{D}_R \cup \Omega$$

Then g is defined by

$$\begin{aligned} g(z) &= f(z) \qquad \text{for } |z| < R \\ &= f_{\zeta_k}(z) \qquad \text{for } |z - \zeta_k| < R_{\zeta_k}, \quad k = 1, 2, \ldots, n \end{aligned}$$

as well defined, single-valued and analytic on D and has same power series representation as f for $|z| < R$. Thus there exists an analytic function, say ϕ, in $\mathbb{D}_{R+\delta}$, which coincides with f on $\mathbb{D}_R$. But, then by Taylor's theorem we have the power series representation

$$\phi(z) = \sum_{n \geq 0} b_n z^n \quad \text{for} \quad z \in \mathbb{D}_{R+\delta}$$

Since $f = g$ on $\mathbb{D}_R$, by the uniqueness theorem, we have $a_n = b_n$ for each n. This shows that the radius of convergence of f is $R + \delta$, which is a contradiction. $\square$

Theorem 15.3.2. If $a_n \geq 0$ and $f(z) = \sum_{n\geq 0} a_n z^n$ has radius of convergence 1, then $(f, \mathbb{D})$ has no direct analytic continuation to a function element (F, D) with $1 \in D$.

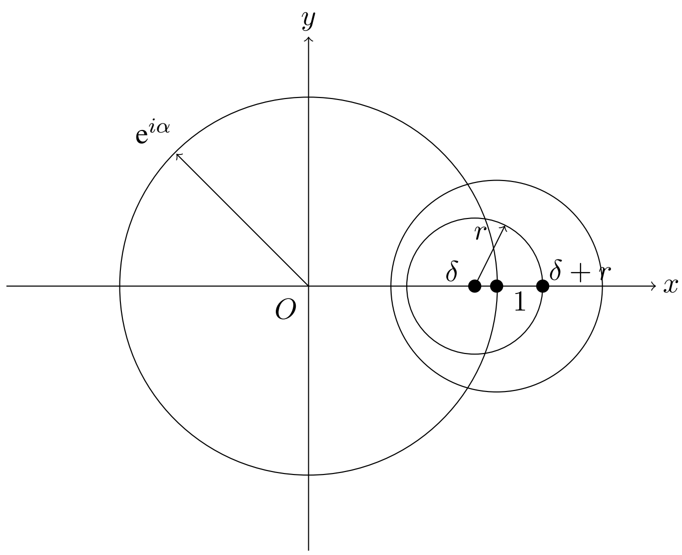


Figure 15.2: Existence of singularity on the circle of convergence

Proof. For each $z = r\,e^{i\theta} \in \mathbb{D}(0 < r < 1; \theta \in [0, 2\pi])$, we have

$$f^{(k)}(z) = \sum_{n\geq k} n(n-1)\cdots(n-(k-1))a_n z^{n-k} \tag{15.3.1}$$

so that since $a_n \geq 0$

$$|f^{(k)}(r\,e^{i\theta})| \leq \sum_{n\geq k} n(n-1)\cdots(n-(k-1))a_n z^{n-k} = f^{(k)}(r) \tag{15.3.2}$$

We have to show that 1 is a singular point of f. Suppose, on the contrary that 1 is a regular point of f. Then, f can be analytically continued in a neighbourhood of $z = 1$ and so there is a δ with $0 < \delta < 1$ (see fig. (15.2)) for which the Taylor's series expansion of f about δ, namely the series

$$\sum_{k\geq 0} \frac{f^{(k)}(\delta}{k!}(z-\delta)^k, \tag{15.3.3}$$

would be convergent for $|z - \delta| < r$ with $\delta + r > 1$. Then by (15.3.2), we find that

$$\frac{|f^{(k)}(\delta\,e^{i\theta})|}{k!} \leq \frac{f^{(k)}(\delta)}{k!}.$$

From this, the root test and the comparison test with (15.3.3), it follows that the radius of convergence of the Taylor series about $\delta\,e^{i\theta}$ is at least r. This observation implies that the Taylor series

$$\sum_{k\geq 0} \frac{f^{(k)}(\delta\,e^{i\theta})}{k!}\left(z - \delta\,e^{i\theta}\right)^k$$

would be convergent in the disc $|z-\delta\,\mathrm{e}^{i\theta}\,|<r$ for each θ, with $\delta+r>1$. In other words, the Taylor series

$$\sum_{k\geq 0}\frac{f^{(k)}(z_0)}{k!}\,(z-z_0)^k$$

about each z_0 with $|z_0|=\delta$ would have radius of convergence $\geq r>1-\delta$. Since this contradicts the previous theorem, and hence 1 must be a singular point of f. This completes the proof. □

Notice that the last series is actually a rearrangement of $\sum_{n\geq 0} a_n z^n$. Indeed, by (15.3.1),

$$\begin{aligned}\sum_{k\geq 0}\left(\sum_{n\geq k}\binom{n}{k}a_n z_0^{n-k}\right)(z-z_0)^k &= \sum_{n\geq 0}\sum_{k=0}^{n}\binom{n}{k}a_n z_0^{n-k}(z-z_0)^k\\ &= \sum_{n\geq 0}a_n(z-z_0+z_0)^n\\ &= \sum_{n\geq 0}a_n z^n.\end{aligned}$$

Corollary 15.3.1. If $a_n\geq 0$ and $f(z)=\sum_{n\geq 0}a_n z^n$ has the radius of convergence $R>0$, then $z=R$ is a singularity of $f(z)$.

Finally, we state the following result.

Theorem 15.3.3. If $f(z)=\sum_{k\geq 0}a_k z^{n_k}$ and $\liminf_{k\to\infty}\dfrac{n_{k+1}}{n_k}>1$. Then the circle of convergence of the power series is the natural boundary for f.

Example 15.3.1. Consider the function

$$f(z)=\sum_{k=0}^{\infty}\frac{z^{3^k}}{3^k}.$$

Clearly, $n_k=3^k$. Thus,

$$\liminf_{k\to\infty}\frac{n_{k+1}}{n_k}=\liminf_{k\to\infty}\frac{3^{k+1}}{3^k}=\liminf_{k\to\infty}3\frac{3^k}{3^k}=3>1.$$

Thus, by the previous theorem, the circle of convergence of the given series is the natural boundary for f. We find the radius of convergence by Cauchy-Hadamard's theorem. Let R be the radius of convergence of the power series. Writing the given power series as $\sum_{n=0}^{\infty}a_n z^n$, we get the terms a_n of the power series as

$$f(z)=0\cdot z^0+1\cdot z^1+0\cdot z^2+\frac{1}{3}z^3+\cdots+\frac{1}{9}z^9+\cdots$$

Thus, the set $|a_n|^{\frac{1}{n}}$ is given below

$$\left\{|1|^1,\ |0|^{\frac{1}{2}},\ \left|\frac{1}{3}\right|^{\frac{1}{3}},\ 0,\ 0,\ \ldots,\left|\frac{1}{9}\right|^{\frac{1}{9}},\ldots\right\}.$$

Thus, Cauchy Hadamard's theorem yields

$$\begin{aligned} R = \frac{1}{\limsup\limits_{n\to\infty} |a_n|^{\frac{1}{n}}} &= \frac{1}{\limsup\limits_{k\to\infty} |a_{3^k}|^{\frac{1}{3^k}}} \\ &= \frac{1}{\limsup\limits_{k\to\infty} \left|\frac{1}{3^k}\right|^{\frac{1}{3^k}}} \\ &= \limsup_{k\to\infty} \left(3^k\right)^{\frac{1}{3^k}} = 1. \end{aligned}$$

Hence, $R = 1$ is the radius of convergence of the given power series and $|z| = 1$ is the natural boundary of f.

Few Probable Questions

---

1. Show that the function $f(z) = \sum\limits_{n\geq 0} a_n z^n$ having radius of convergence $R > 0$ must have at least one singularity on $|z| = R$.
2. Find the natural boundary of the function $f(z) = \sum\limits_{n\geq 0} a_n z^n$ has radius of convergence 1, the show that $(f, \mathbb{D})$ has no direct analytic continuation to a function element (F, D), with $1 \in D$.

$$f(z) = \sum_{k=0}^{\infty} \frac{z^{2^k}}{2^{k^2}}.$$

Unit 16

Course Structure

- Monodromy theorem
- germs

16.1 Introduction

This unit is a continuation of the previous unit an deals in the Monodromy theorem. In complex analysis, the Monodromy theorem is an important result about analytic continuation of a complex-analytic function to a larger set. The idea is that one can extend a complex-analytic function (from here on called simply analytic function) along curves starting in the original domain of the function and ending in the larger set. A potential problem of this analytic continuation along a curve strategy is there are usually many curves which end up at the same point in the larger set. The Monodromy theorem gives sufficient conditions for analytic continuation to give the same value at a given point regardless of the curve used to get there, so that the resulting extended analytic function is well-defined and single-valued.

Objectives

After reading this unit, you will be able to

- define homotopy of two curves
- define germ of an analytic function f at a point a
- deduce Monodromy theorem

16.2 Monodromy Theorem

We first give some definitions.

Definition 16.2.1. Let $\gamma_0, \gamma_1 : [0,1] \to G$ be two closed rectifiable curves in a region G then γ_0 is homotopic to γ_1 in G if there is a continuous function

$$F : [0,1] \times [0,1] \to G$$

such that

$$\begin{array}{rcll} F(s,0) & = & \gamma_0(s) & \\ F(s,1) & = & \gamma_1(s) & (0 \le s \le 1) \\ F(0,t) & = & F(1,t) & (0 \le t \le 1) \end{array}$$

Definition 16.2.2. Let $\gamma_0, \gamma_1 : [0,1] \to G$ be two closed rectifiable curves in G such that $\gamma_0(0) = \gamma_1(0) = a$ and $\gamma_0(1) = \gamma_1(1) = b$. Then γ_0 and γ_1 are fixed-end-point homotopic (FEP homotopic) if there is a continuous map $F : [0,1] \times [0,1] \to G$ such that

$$\begin{array}{l} F(s,0) = \gamma_0(s), \;\; F(s,1) = \gamma_1(s) \\ F(0,t) = a, \;\; F(1,t) = b, \quad \text{for } \; 0 \le s,t \le 1. \end{array}$$

We note that the relation of FEP homotopic is an equivalence relation on the curves from one given point to another.

Definition 16.2.3. An open set G is called simply connected if G is connected and every closed curve in G is homotopic to zero.

This is equivalent to the definition of simply connected region which we had learnt previously which states that a set is simply connected if every closed rectifiable curve can be continuously deformed to a single point without passing through any point outside the set. Now, we define the germ of a function f.

Definition 16.2.4. Let (f, G) be a function element. Then the germ of f at a is the collection of all function elements (g, D) such that $a \in D$ and $f(z) = g(z)$ for all z in a neighbourhood of a. The germ of f at a is denoted by $[f]_a$.

Notice that $[f]_a$ is a collection of function elements.

Definition 16.2.5. Let $\gamma : [0,1] \to \mathbb{C}$ be a path and suppose that for each $t \in [0,1]$ there is a function element (f_t, D_t) such that

1. $\gamma(t) \in D_t$;
2. for each $t \in [0,1]$, there is a $\delta > 0$ such that $|s - t| < \delta$ implies that $\gamma(s) \in D_t$ and

$$[f_s]_{\gamma(s)} = [f_t]_{\gamma(s)}.$$

Then (f_1, D_1) is called analytic continuation of (f_0, D_0) along the path γ.

Remark 16.2.1. Since γ is a continuous function and $\gamma(t)$ is in the open set D_t, so there is a $\delta > 0$ such that $\gamma(s) \in D_t$ for $|s - t| < \delta$.

So, part 2 of the previous definition implies

$$f_s(z) = f_t(z) \;\; \text{for all} \;\; z \in D_s \cap D_t,$$

whenever $|s - t| < \delta$.

Theorem 16.2.1. Let $\gamma : [0,1] \to \mathbb{C}$ be a path from a to b and let $\{(f_t, D_t) : 0 \leq 1\}$ and $\{(g_t, B_t) : 0 \leq 1\}$ be analytic continuation along γ such that $[f_0]_a = [g_0]_a$. Then $[f_1]_b = [g_1]_b$.

Proof. Consider the set

$$T = \{t \in [0,1] : [f_t]_{\gamma(t)} = [g_t]_{\gamma(t)}\}$$

Since $[f_0]_a = [g_0]_a$, so $0 \in T$. Thus $T \neq \phi$.

We shall show that T is both open and closed. To show T is open, let t be a fixed point of T such that $t \neq 0$. By definition of analytic continuation, there is $\delta > 0$ such that for $|s - t| < \delta$.

$$\begin{aligned}
&\gamma(s) \in D_t \cap B_t \quad \text{and} \\
&[f_s]_{\gamma(s)} = [f_t]_{\gamma(s)} \\
&[g_s]_{\gamma(s)} = [g_t]_{\gamma(s)}
\end{aligned}$$

But $t \in T$ implies

$$f_t(z) = g_t(z) \qquad \forall z \in D_t \cap B_t$$

Hence, $[f_t]_{\gamma(s)} = [g_t]_{\gamma(s)}$ for all $\gamma(s) \in D_t \cap B_t$. So, $[f_s]_{\gamma(s)} = [g_s]_{\gamma(s)}$ whenever $|s - t| < \delta$. That is, $s \in T$ whenever $|s - t| < \delta$ or $(t - \delta, t + \delta) \subset T$.

If $t = 0$ then the above argument shows that $[a, a + \delta) \subset T$ for some $\delta > 0$. Hence T is open.

To show that T is closed let t be a limit point of T. Again by definition of analytic continuation there is a $\delta > 0$ such that $|s - t| < \delta$, $\gamma(s) \in D_t \cap B_t$ and

$$\begin{aligned}
[f_s]_{\gamma(s)} &= [f_t]_{\gamma(s)} \\
[g_s]_{\gamma(s)} &= [g_t]_{\gamma(s)}
\end{aligned} \tag{16.2.1}$$

Since t is a limit point of T, there is a point s in T such that $|s - t| < \delta$. Let $G = D_t \cap B_t \cap D_s \cap B_s$. Then $\gamma(s) \in G$. So, G is non-empty open set. Thus by definition of T, $f_s(z) = g_s(z)$ for all $z \in G$. But, (16.2.1) implies

$$\begin{aligned}
&f_s(z) = f_t(z) \qquad \text{and} \qquad g_s(z) = g_t(z) \qquad \text{for all } z \in G \\
&f_t(z) = g_t(z) \qquad \forall z \in G.
\end{aligned}$$

Since, G has a limit point in $D_t \cap B_t$, this gives $[f_t]_{\gamma(t)} = [g_t]_{\gamma(t)}$. Thus, $t \in T$ and so T is closed.

Now,T is non-empty subset of $[0,1]$ such that T is both open and closed. So, connectedness of $[0,1]$ implies $T = [0,1]$. Thus $1 \in T$ and hence $[f_1]_{\gamma(1)} = [g_1]_{\gamma(1)}$, that is, $[f_1]_b = [g_1]_b$ as $\gamma(1) = b$ □

Definition 16.2.6. If $\gamma : [0,1] \to \mathbb{C}$ is a path from a to b and $\{(f_t, D_t) : 0 \leq t \leq 1\}$ is an analytic continuation along γ then the germ $[f_1]_b$ is the analytic continuation of $[f_0]_a$ along γ.

Remark 16.2.2. Suppose a and b are two complex numbers and let γ and σ be two paths from a to b. Suppose, $\{(f_t, D_t)\}$ and $\{(g_t, D_t)\}$ are analytic continuations along γ and σ respectively such that $[f_0]_a = [g_0]_a$. Now, the question is, does it follow that $[f_1]_b = [g_1]_b$? If γ and σ are the same path then the above result gives an affirmative answer. However, if γ and σ are distinct then the answer can be no.

Lemma 16.2.1. Let $\gamma : [0,1] \to \mathbb{C}$ be a path and let $\{(f_t, D_t) : 0 \leq t \leq 1\}$ be an analytic continuation along γ. For $0 \leq t \leq 1$, let $R(t)$ be the radius of convergence of the power series expansion of f about $z = \gamma(t)$. Then either $R(t) \equiv \infty$ or $R : [0,1] \to (0, \infty)$ is continuous.

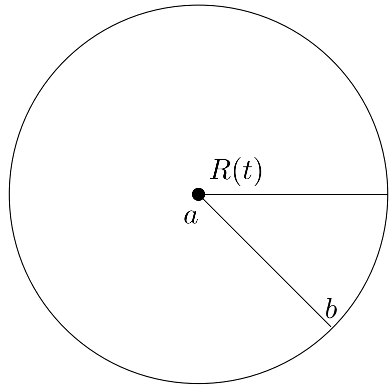


Figure 16.1

Proof. Suppose $R(t) = \infty$ for some value of t. Then, f_t can be extended to an entire function. It follows that $f_s(z) = f_t(z)$ for all $z \in D_s$ so that $R(s) = \infty$ for all $s \in [0, 1]$. That is $R(s) \equiv \infty$. Now, suppose that $R(t) < \infty$ for all t. Let t be a fixed number in $[0, 1]$ and let $a = \gamma(t)$. Let

$$f_t(z) = \sum_{n=0}^{\infty} a_n (z-a)^n$$

be the power series expansion of f_t about a. Now, let $\delta_1 > 0$ be such that $|s-t| < \delta_1$ implies that $\gamma(s) \in D_t \cap B(a; R(t))$ and $[f_s]_{\gamma(s)} = [f_t]_{\gamma(s)}$. Fix s with $|s-t| < \delta_1$ and let $b = \gamma(s)$. Now, f_t can be extended to an analytic function on $B(a; R(t))$. Since, f_s agrees with f_t on a neighbourhood of f_s can be extended so that it is also analytic on $B(a; R(t)) \cup D_s$. If f_s has power series expansion

$$f_s(z) = \sum_{n=0}^{\infty} b_n (z-b)^n \quad \text{about} \quad z = b$$

Then the radius of convergence $R(s)$ must be at least as big as the distance from b to the circle $|z-a| = R(t)$; that is,

$$\begin{aligned} R(s) &\geq d(b, \{z : |z-a| = R(t)\}) \\ &\geq R(t) - |a-b| \end{aligned}$$

This implies $R(t) - R(s) \leq |a-b|$ that is $R(t) - R(s) \leq |\gamma(t) - \gamma(s)|$. Similarly, we can show $R(s) - R(t) \leq |\gamma(t) - \gamma(s)|$. Hence,

$$|R(s) - R(t)| \leq |\gamma(t) - \gamma(s)| \quad \text{for} \quad |s-t| < \delta_1.$$

Since, $\gamma : [0, 1] \to \mathbb{C}$ is continuous so given $\epsilon > 0$, $\exists \delta_2 > 0$ so that $|\gamma(t) - \gamma(s)| < \epsilon$ for $|s-t| < \delta_2$. Let $\delta = \min\{\delta_1, \delta_2\}$. Then $\delta > 0$ and $|R(s) - R(t)| < \epsilon$ for $|s-t| < \delta$. Hence R is continuous at t.

□

Lemma 16.2.2. Let $\gamma : [0, 1] \to \mathbb{C}$ be a path from a to b and let $\{(f_t, D_t) : 0 \leq t \leq 1\}$ be an analytic continuation along γ. There is a number $\epsilon > 0$ such that if $\sigma : [0, 1] \to \mathbb{C}$ is any path from a to b with $|\gamma(t) - \sigma(t)| < \epsilon$ for all t and if $\{(g_t, B_t) : 0 \leq t \leq 1\}$ is any continuation along γ with $[g_0]_a = [f_0]_a$; the $[g_1]_b = [f_1]_b$.

Proof. For $0 \leq t \leq 1$, let $R(t)$ be the radius of convergence of the power series expansion of f_t about $z = \gamma(t)$. If $R(t) \equiv \infty$ then any value of ϵ will be sufficient. So, suppose $R(t) < \infty$ for all t. Since R is a continuous function and $R(t) > 0$ for all t, R has a positive minimum value. Let $0 < \epsilon < \frac{1}{2}\min\{R(t) : 0 \leq t \leq 1\}$. Suppose $\sigma : [0,1] \to \mathbb{C}$ is any path from a to b with $|\gamma(t) - \sigma(s)| < \epsilon$ for all t and $\{(g_t, B_t) : 0 \leq t \leq 1\}$ is any continuation along σ with $[g_0]_a = [f_0]_a$. Suppose D_t is a disk of radius $R(t)$ about $\gamma(t)$. Since $|\sigma(t) - \gamma(t)| < \epsilon < R(t)$, $\sigma(t) \in B_t \cap D_t$ for all t.

Define the set $T = \{t \in [0,1] : f_t(z) = g_t(z)\ \forall\ z \in B_t \cap D_t\}$. Then $0 \in T$, since $[g_0]_a = [f_0]_a$. So, $T \neq \phi$. We will show that $1 \in T$. For this, it is sufficient to show that T is both open and closed subset of $[0,1]$.

To show T is open, let t be any fixed point of T. Choose $\delta > 0$

$$\begin{aligned} |\gamma(s) - \gamma(t)| < \epsilon, \quad [f_s]_{\gamma(s)} = [f_t]_{\gamma(s)} \\ |\sigma(s) - \sigma(t)| < \epsilon, \quad [g_s]_{\sigma(s)} = [g_t]_{\sigma(s)} \end{aligned} \tag{16.2.2}$$

and $\sigma(s) \in B_t$ whenever $|s - t| < \delta$.

We now show that $B_s \cap B_t \cap D_s \cap D_t \neq \phi$ for $|s-t| < \delta$. For this we will show $\sigma(s) \in B_s \cap B_t \cap D_s \cap D_t$ for $|s - t| < \delta$. If $|s - t| < \delta$, then

$$|\sigma(s) - \gamma(s)| < \epsilon < R(s)$$

so that $\sigma(s) \in D_s$. Also

$$|\sigma(s) - \gamma(t)| = |\sigma(s) - \gamma(s) + \gamma(s) - \gamma(t)| \leq |\sigma(s) - \gamma(s)| + |\gamma(s) - \gamma(t)| < 2\epsilon < R(t)$$

So, $\sigma(s) \in D_t$. Since we already have $\sigma(s) \in B_s \cap B_t$, so (16.2.2) we have $\sigma(s) \in B_s \cap B_t \cap D_s \cap D_t = G$. Since, $t \in T$, it follows that $f_t(z) = g_t(z)$ for all $z \in G$. Also, (16.2.2) implies $f_s(z) = f_t(z)$ and $g_s(z) = g_t(z)$ for all $z \in G$. Thus, $f_s(z) = g_s(z)$ for all $z \in G$. But since G has a limit point in $B_s \cap D_s$, we must have $s \in T$. That is, $(t - \delta, t + \delta) \subset T$. Hence, T is open.

Similarly, we can show T is closed. T is non-empty and closed subset of $[0,1]$. As, $[0,1]$ is connected, we have $[0,1] = T$. Thus $1 \in T$ and the result follows. □

Definition 16.2.7. Let (f, D) be a function element and let G be a region which contains D. Then (f, D) admits unrestricted analytic continuation in G if for any path γ in G with initial point in D there is an analytic continuation of (f, D) along γ.

Theorem 16.2.2. (Monodromy Theorem) Let (f, D) be a function element and let G be a region containing D such that (f, D) admits unrestricted continuation in G. Let $a \in D$, $b \in G$ and let γ_0 and γ_1 be paths in G from a to b; let $\{(f_t, D_t) :\ 0 \leq t \leq 1\}$ and $\{(g_t, D_t) :\ 0 \leq t \leq 1\}$ be analytic continuations of (f, D) along γ_0 and γ_1 respectively. If γ_0 and γ_1 are FEP homotopic in G, then

$$[f_1]_b = [g_1]_b.$$

Proof. Since γ_0 and γ_1 are fixed end point homotopic in G, there is a continuous function $F : [0,1] \times [0,1] \to G$ such that

$$\begin{aligned} F(t,0) = \gamma_0(t), \quad F(t,1) = \gamma_1(t) \\ F(0,u) = a, \qquad F(1,u) = b. \end{aligned}$$

For all t and u in $[0,1]$. Let u be a fixed point of $[0,1]$. Consider the path γ_u, defined by

$$\gamma_u(t) = F(t,u) \ \text{ for } \ t \in [0,1].$$

Then,

$$\gamma_u(0) = F(0,u) = a, \ \ \gamma_u(1) = F(1,u) = b.$$

Thus, γ_u is a path from a to b. By hypothesis, there is an analytic continuation

$$\{(h_{t,u}, D_{t,u}) : \ 0 \leq t \leq 1\}$$

of (f, D) along γ_u. Now, $\{(h_{t,u}, D_{t,u}) : \ 0 \leq t \leq 1\}$ and $\{(f_t, D_t) : \ 0 \leq t \leq 1\}$ are analytic continuations along γ_0 so by theorem 16.2.1, we have

$$[f_1]_b = [h_{1,0}]_b.$$

Similarly,

$$[g_1]_b = [h_{1,1}]_b.$$

To prove the theorem, it is sufficient to show

$$[h_{1,0}]_b = [h_{1,1}]_b.$$

Consider the set

$$U = \{u \in [0,1] : \ [h_{1,u}]_b = [h_{1,0}]_b\}.$$

We will show that $1 \in U$. Now, $0 \in U$. So, $U \neq \emptyset$. We claim that U is both open and closed subset of $[0,1]$.

Let $u \in [0,1]$ be arbitrary. We assert that there is $\delta > 0$ such that if $|u - v| < \delta$, then

$$[h_{1,u}]_b = [h_{1,v}]_b. \tag{16.2.3}$$

By lemma 16.2.2, there is an $\epsilon > 0$ such that if σ is any path from a to b with $|\gamma_u(t) - \sigma(t)| < \epsilon$ for all t and if $\{(k_t, E_t)\}$ is any continuation of (f, D) along σ, then

$$[h_{1,u}]_b = [k_1]_b. \tag{16.2.4}$$

Now, F is uniformly continuous function so there is $\delta > 0$ such that

$$\begin{aligned} & |F(t,u) - F(t,v)| < \epsilon \ \text{ whenever } \ |u - v| < \delta \\ \Rightarrow \quad & |\gamma_u(t) - \gamma_v(t)| < \epsilon \quad \text{ whenever } \ |u - v| < \delta. \end{aligned}$$

So, for $|u - v| < \delta$, γ_v is a path from a to b with

$$\gamma_u(t) - \gamma_v(t)| < \epsilon$$

for all t and $\{(h_{t,v}, D_{t,v})\}$ is a continuation of (f, D) along γ_v, so by (16.2.4),

$$[h_{1,u}]_b = [h_{1,v}]_b.$$

Suppose $u \in U$ such that $[h_{1,u}]_b = [h_{1,0}]_b$. Then as proved above, there is a $\delta > 0$ such that $|u - v| < \delta$ which implies that

$$\begin{aligned} & [h_{1,u}]_b = [h_{1,v}]_b \\ \text{i.e.} \quad & v \in (u - \delta, u + \delta) \Rightarrow [h_{1,v}]_b = [h_{1,0}]_b \\ \text{i.e.} \quad & v \in (u - \delta, u + \delta) \Rightarrow v \in U \\ \text{i.e.} \quad & (u - \delta, u + \delta) \subset U. \end{aligned}$$

Hence U is open.

To show that U is closed, we show that $\overline{U} = U$. Let $u \in U$ and δ be the positive number satisfying (16.2.3). Then there is a $v \in U$ such that $|u - v| < \delta$. So, by (16.2.3), $[h_{1,u}]_b = [h_{1,v}]_b$. Since $v \in U$, so $[h_{1,v}]_b = [h_{1,0}]_b$. Thus, $[h_{1,u}]_b = [h_{1,0}]_b$ so that $u \in U$. Thus, U is closed as $\overline{U} = U$.

Now, U is a non-empty open and closed subset of $[0,1]$ and since $[0,1]$ is connected, so, $U = [0,1]$. So, $1 \in U$ and the result is proved. $\square$

The following corollary is the main consequence of the Monodromy theorem.

Corollary 16.2.1. Let (f, D) be a function element which admits unrestricted continuation in the simply connected region G. Then there is an analytic function $F : G \to \mathbb{C}$ such that $F(z) = f(z)$ for all $z \in D$.

Proof. Let a be a fixed point in D and z is any point in G. If γ is a path in G from a to z and $\{(f_t, D_t) : 0 \leq t \leq 1\}$ is an analytic continuation of (f, D) along γ, then let $F(z, \gamma) = f_1(z)$ since G is simply connected.

$F(z, \gamma) = F(z, \sigma)$ for any two paths γ and σ in G from a to z. Thus, $F(z) = F(z, \gamma)$ is a well defined function from G to $\mathbb{C}$. To show that F is analytic, let $z \in G$. Let γ be a path in G from a to z and $\{(f_t, D_t)\}$ be the analytic continuation of (f, D) along γ. Then $F(\omega) = f_1(\omega)$ for all ω in a neighbourhood of z. Hence F must be analytic. $\square$

Few Probable Questions

---

1. Let (f, G) be a function element. Define germ of f at a. If $\gamma : [0, 1] \to \mathbb{C}$ be a path from a to b and $\{(f_t, D_t) : 0 \leq t \leq 1\}$ and $\{(g_t, B_t) : 0 \leq t \leq 1\}$ be analytic continuation along γ such that $[f_0]_a = [g_0]_a$, then show that $[f_1]_b = [g_1]_b$.

2. State and prove Monodromy theorem.

3. If $\gamma : [0, 1] \to \mathbb{C}$ be a path and $\{(f_t, D_t) : 0 \leq t \leq 1\}$ is an analytic continuation along γ. For $0 \leq t \leq 1$, if $R(t)$ is the radius of convergence of the power series expansion of f_t about $z = \gamma(t)$, then show that either $R(t) \equiv \infty$ or $R : [0, 1] \to (0, \infty)$ is continuous.

Unit 17

Course Structure

- Conformal transformations
- Riemann's theorems for circle.

17.1 Introduction

In mathematics, a conformal map is a function that locally preserves angles, but not necessarily lengths. We shall see that the derivative relates the angle between two curves to the angle between their images. In addition, the derivative will be seen to measure the "distortion" of image curves. They are also worth studying because of their usefulness in solving certain physical problems, for example, problems about two-dimensional fluid flow, the idea being to transform a given problem into an equivalent one which is easier to solve. So we wish to consider the problem of mapping a given region G onto a geometrically simpler region G'. For example the open unit disc or the open upper half-plane.

Objectives

After reading this unit, you will be able to

- define conformal and isogonal maps and see certain examples
- deduce further conditions satisfied by conformal maps
- define conformally equivalent regions and see certain examples of them
- define Mӧbius transformation and related terms and deduce few results related to symmetry

17.2 Conformal Transformations

Any straight line in the plane that passes through the origin may be parameterized by $\sigma(s) = s\,\mathrm{e}^{i\alpha}$, where s traverses the set of real numbers and α is the angle measured in radians between the positive real axis and the line. More generally, a straight line passing through the point z_0 and making an angle α with the real axis can be expressed as $\sigma(s) = z_0 + s\,\mathrm{e}^{i\alpha}$, s is real.

Suppose now that a function f is analytic on a smooth (parameterized) curve whose derivative is given by $f'(z(t))z'(t)$ (by chain rule). A smooth curve is characterized by having a tangent at each point. So, we interpret $z'(t)$ as a vector in the direction of the tangent vector at the point $z(t)$. Our purpose is to compare the inclination of the tangent to the curve at a point with the inclination of the tangent to the image curve at the image of the point.

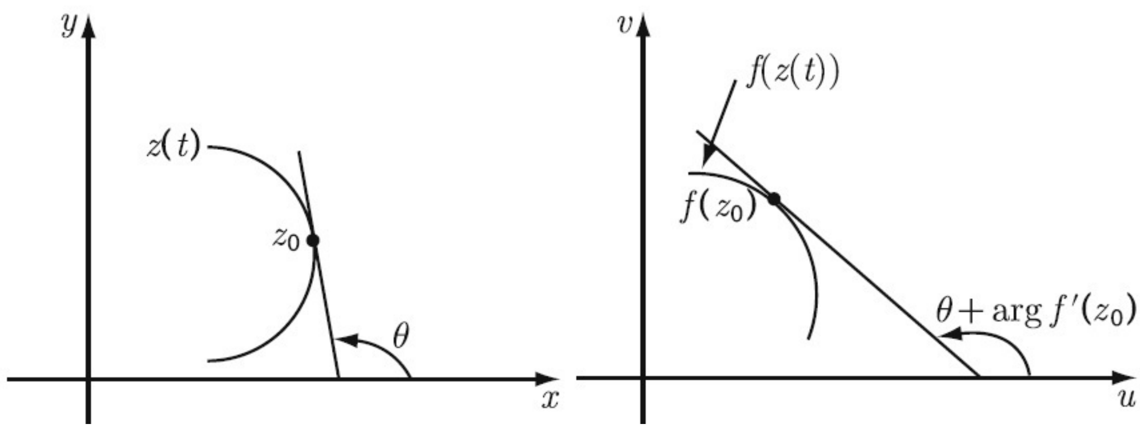
Let $z_0 = z(t_0)$ be a point on the curve $z = z(t)$. Then the vector $z'(t_0)$ is tangent to the curve at the point z_0 and $\arg z'(t_0)$ is the angle this directed tangent makes with the positive x-axis. Suppose that $w = w(t) = f(z(t))$ with $w_0 = f(z_0)$. For any point z on the curve other than z_0, we have the identity

$$w - w_0 = \frac{f(z) - f(z_0)}{z - z_0}(z - z_0).$$

Thus,

$$\arg(w - w_0) = \arg \frac{f(z) - f(z_0)}{z - z_0} + \arg(z - z_0) \qquad (\mod 2\pi), \tag{17.2.1}$$

where it is assumed that $f(z) \neq f(z_0)$ so that (17.2.1) has meaning. Note that $\arg(z - z_0)$ is the angle in the z plane between the x axis and the straight line passing through the points z and z_0, while $\arg(w - w_0)$ is the angle in the w plane between the u axis and the straight line passing through the points w and w_0. Hence, as



z approaches z_0 along the curve $z(t)$, $\arg(z - z_0)$ approaches a value θ, which is the angle that the tangent to the curve $z(t)$ at z_0 makes with the x-axis. Similarly, $\arg(w - w_0)$ approaches a value ϕ, the angle that the tangent to the curve $f(z(t))$ at w_0 makes with the u axis.

Suppose that $f'(z_0) \neq 0$ so that $\arg f'(z_0)$ has a meaning. Then, taking limits in (17.2.1), we find ($\mod 2\pi$) that

$$\phi = \arg f'(z_0) + \theta, \quad \text{or} \quad \arg w'(t_0) = \arg f'(z_0) + \arg z'(t_0). \tag{17.2.2}$$

That is, the difference between the tangent to a curve at a point and the tangent to the image curve at the image of the point depends only on the derivative of the function at the point.

Theorem 17.2.1. Suppose $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$. Let $C_1 : z_1(t)$ and $C_2 : z_2(t)$ be smooth curves in the z plane that intersect at $z_0 = z_1(t_0) = z_2(t_0)$ with $C_1' : w_1(t)$ and $C_2' : w_2(t)$ the images of C_1 and C_2, respectively. Then the angle between C_1 and C_2, measured from C_1 to C_2, is equal to the angle between C_1' and C_2' measured from C_1' to C_2'.

Proof. Let the tangents to C_1 and C_2 make angles θ_1 and θ_2 respectively with the x-axis. Then the angle between C_1 and C_2 at z_0 is $\theta_2 - \theta_1$ (see fig. 17.1). According to (17.2.2), the angle between C_1' and C_2', which

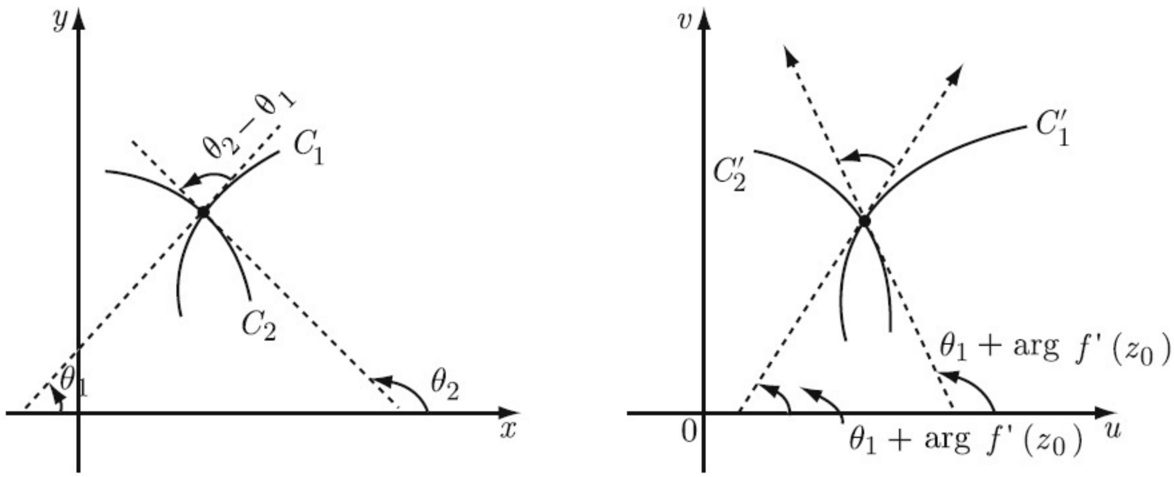


Figure 17.1

is the angle between the tangent vectors $f'(z_0)z_1'(t_0)$ and $f'(z_0)z_2'(t_0)$, of the image curves is

$$\theta_2 + \arg f'(z_0) - (\theta_1 + \arg f'(z_0)) = \theta_2 - \theta_1,$$

and the theorem is proved. □

A function that preserves both angle size and orientation is said to be **conformal**. Theorem 17.2 says that an analytic function is conformal at all points where the derivative is non-zero. For example, the function $f(z) = e^z$ maps vertical and horizontal lines into circles and orthogonal radial rays, respectively.

A function that preserves angle size but not orientation is said to be **isogonal**. An example of such a function is $f(z) = \overline{z}$. To illustrate, $\overline{z}$ maps maps the positive real axis and the positive imaginary axis onto the positive real axis and the negative real axis respectively (see fig. 17.2). Although the two curves intersect at right angles in each plane, a ”counterclockwise” angle is mapped onto a ”clockwise” angle.

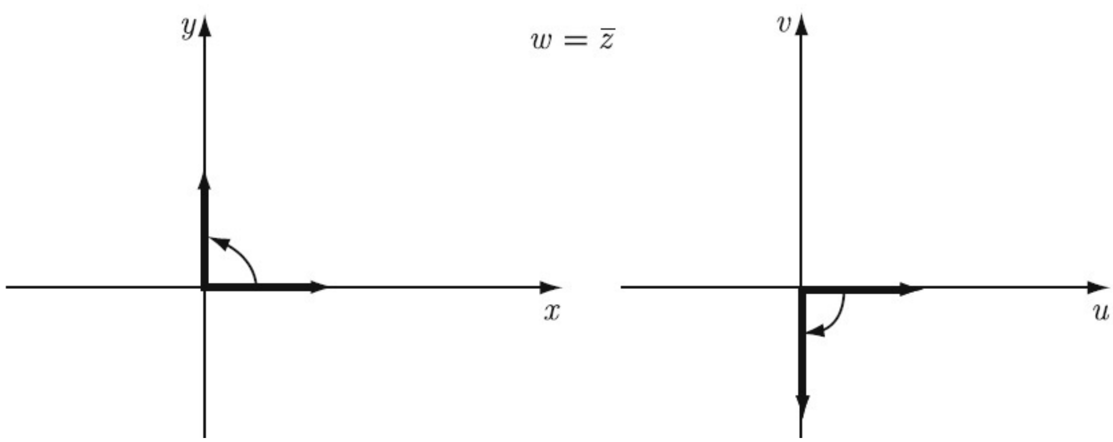


Figure 17.2

The non-zero derivatives of f has certain implications which we shall see now.

Theorem 17.2.2. If $f(z)$ is analytic at z_0 with $f'(z_0) \neq 0$, then $f(z)$ is one-to-one in some neighbourhood of z_0.

Proof. Since $f'(z_0) \neq 0$ and $f'(z)$ is continuous at z_0, there exists a $\delta > 0$ such that

$$|f'(z) - f'(z_0)| < \frac{|f'(z_0)|}{2} \quad \text{for all} \quad |z| < \delta$$

Let z_1 and z_2 be two distinct points in $|z| < \delta$, and γ be a line segment connecting z_1 and z_2. Set $\phi(z) = f(z) - f'(z_0)z$ so that $|\phi'(z)| < |f'(z_0)|/2$ for all $|z| < \delta$. Now we have,

$$|\phi(z_2) - \phi(z_1)| = \left| \int_\gamma \phi'(z) dz \right| < (|f'(z_0)|/2)|z_2 - z_1|$$

or equivalently,

$$|f(z_2) - f(z_1) - f'(z_0)(z_2 - z_1)| < (|f'(z_0)|/2)|z_2 - z_1|.$$

Thus, by the triangle inequality, we obtain

$$|f(z_2) - f(z_1)| > (|f'(z_0)|/2)|z_2 - z_1| > 0.$$

That is, $f(z)$ is one-to-one in $|z| < \delta$. □

The vanishing of a derivative does not preclude the possibility of real function being one-to-one. Although the derivative of $f(x) = x^3$ is zero at the origin, the function is still one-to-one on the real line. That this cannot occur for complex functions is seen by

Theorem 17.2.3. If $f(z)$ is analytic and one-to-one in a domain D, then $f'(z) \neq 0$ in D, so that f is conformal in D.

Proof. If $f'(z) = 0$ at some point z_0 in D, then

$$f(z) - f(z_0) = \frac{f''(z_0)}{2!}(z - z_0)^2 + \cdots$$

has a zero of order $k (k \geq 2)$ at z_0. Since zeros of an analytic function are isolated, there exists an $r > 0$ so small that both $f(z) - f(z_0)$ and $f'(z)$ have no zeros in the punctured disk $0 < |z - z_0| \leq r$. Let $g(z) := f(z) - f(z_0)$, $C = \{z : |z - z_0| = r\}$ and $m = \min_{z \in C} |g(z)|$.

Then, g has a zero of order $k (k \geq 2)$ and $m > 0$. Let $b \in \mathbb{C}$ be such that $0 < |b - f(z_0)| < m$. Then, as $m \leq |g(z)|$ on C,

$$|f(z_0) - b| < |g(z)| \quad \text{on} \quad C$$

It follows from Rouche's theorem that $g(z)$ and $g(z) + (f(z_0) - b) = f(z) - b$ have same number of zeros inside C. Thus, $f(z) - b$ has at least two zeros inside C. Observe that none of these zeros can be at z_0. Since $f'(z) \neq 0$ in the punctured disk $0 < |z - z_0| \leq r$, these zeros must be simple and so, distinct. Thus, $f(z) = b$ at two or more points inside C. This contradicts the fact that f is one-to-one on D. □

We sum up our results for differentiable functions. In the real case,the nonvanishing of a derivative on an interval is a sufficient but not a necessary condition for the function to be one-to-one on the interval; whereas in the complex case, the nonvanishing of a derivative on a domain is a necessary but not a sufficient condition for the function to be one-to-one on the domain.

An analytic function $f : D \to \mathbb{C}$ is called *locally bianalytic* at $z_0 \in D$ if there exists a neighbourhood N of z_0 such that the restriction of f from N onto $f(N)$ is bianalytic. Clearly, a locally bianalytic map on D need not be bianalytic on D, as the example $f(z) = z^n (n > 2)$ on $\mathbb{C} - \{0\}$ illustrates.

Combining 17.2.2 and 17.2.3 leads to the following criterion for local bianalytic maps.

Theorem 17.2.4. Let $f(z)$ be analytic in a domain D and $z_0 \in D$. Then f is bianalytic at z_0 iff $f'(z_0) \neq 0$.

A sufficient condition for an analytic function to be one-to-one in a simply connected domain is that it be one-to-one on its boundary. More formally, we have

Theorem 17.2.5. Let $f(z)$ be analytic in a simply connected domain D and on its boundary, the simple closed contour C. If $f(z)$ is one-to-one on C, then $f(z)$ is one-to-one in D.

Proof. (See fig. 17.3) Choose a point $z_0 \in D$ such that $w_0 = f(z_0) \neq f(z)$ for z on C. According to the argument principle, the number of zeros of $f(z) - f(z_0)$ in D is given by $(1/2\pi i)\int_C \{f(z) - f(z_0)\}dz$. By hypothesis, the image of C must be a simple closed contour, which we shall denote by C'. Thus the net change in the argument of $w - w_0 = f(z) - f(z_0)$ as $w = f(z)$ traverses the contour C' is either $+2\pi$ or -2π, according to whether the contour is traversed counterclockwise or clockwise. Since $f(z)$ assumes the value w_0 at least once in D, we must have

$$\frac{1}{2\pi i}\int_C \{f(z) - f(z_0)\}dz = \frac{1}{2\pi i}\int_C \{w - w_0\} = 1.$$

That is, $f(z)$ assumes the value $f(z_0)$ exactly once in D.

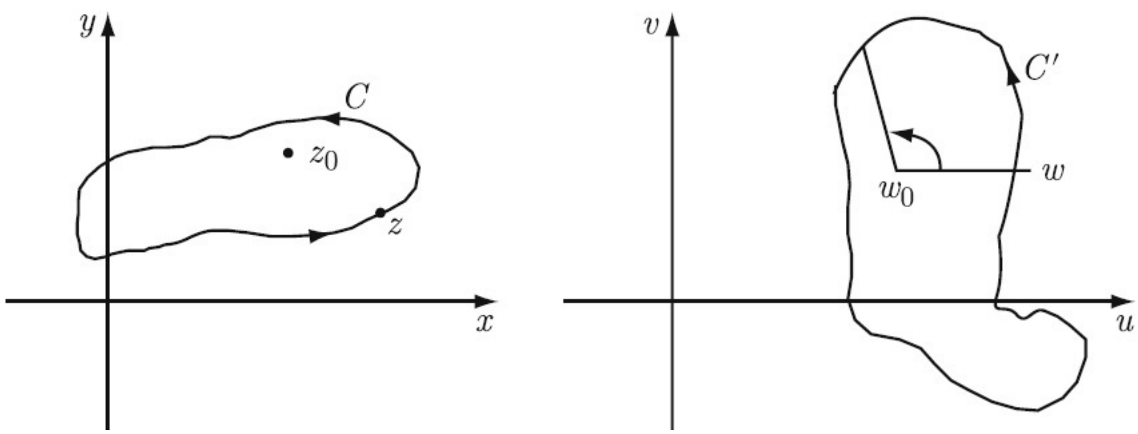


Figure 17.3

This proves the theorem for all points z_0 in D at which $f(z) \neq f(z_0)$ when z is on C. If $f(z) = f(z_0)$ at some point on C, then the expression $\int_C \{f(z) - f(z_0)\}dz$ is not defined. We leave for the reader the completion of the proof in this special case. □

17.3 Conformal Equivalences and Examples

An analytic map $f : G \to G'$ which is bijective is called a bianalytic map as we have already come across. Given such a map f, we say that G and G' are conformally equivalent or simply biholomorphic. An important fact is that the inverse of f is analytic in that case automatically. We have also seen that if an analytic map $f : G \to G'$ is injective, then $f'(z) \neq 0$ for all $z \in G$, that is, f is conformal. We begin our study of conformal mappings by looking at a number of specific examples. The first gives the conformal equivalence between the unit disc and the upper half-plane, which plays an important role in many problems.

Example 17.3.1. Let $\mathbb{H} = \{z \in \mathbb{C} : \ \text{Im}\, z > 0\}$ be the upper half plane. A remarkable fact, which at first seems surprising, is that the unbounded set $\mathbb{H}$ is conformally equivalent to the unit disc. Moreover, an explicit formula giving this equivalence exists. Indeed, let

$$F(z) = \frac{i - z}{i + z} \quad \text{and} \quad G(w) = i\frac{1 - w}{1 + w}.$$

Then it is a regular exercise to check that map $F : \mathbb{H} \to \mathbb{D}$ is conformal with inverse $G : \mathbb{D} \to \mathbb{H}$. An interesting aspect of these functions is their behaviour on the boundaries of our open sets. Observe that F is analytic everywhere on $\mathbb{C}$ except at $z = -i$, and in particular it is continuous everywhere on the boundary of $\mathbb{H}$, namely, the real line. If we take $z = x$ real, then the distance from x to i is the same as the distance from x to $-i$, therefore $|F(x)| = 1$. Thus, F maps $\mathbb{R}$ onto the boundary of $\mathbb{D}$. We get more information by writing

$$F(z) = \frac{i - x}{i + x} = \frac{1 - x^2}{1 + x^2} + i\frac{2x}{1 + x^2},$$

and parametrizing the real line by $x = \tan t$ with $t \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$. Since

$$\sin 2a = \frac{2\tan a}{1 + \tan^2 a} \quad \text{and} \quad \cos 2a = \frac{1 - \tan^2 a}{1 + \tan^2 a},$$

we have, $F(x) = \cos 2t + i \sin 2t = \mathrm{e}^{i2t}$. Hence the image of the real line is the arc consisting of the circle omitting the point -1. Moreover, as x travels from $-\infty$ to ∞, $F(x)$ travels along the arc starting from -1 and first going through that part of the circle that lies in the lower half-plane. The point -1 on the circle corresponds to the "point at infinity" of the upper half-plane.

Example 17.3.2. Mappings of the form

$$z \mapsto \frac{az + b}{cz + d},$$

where a, b, c and d are complex numbers, and where the denominator is assumed not to be a multiple of the numerator, are usually referred to as **fractional linear transformations**.

Example 17.3.3. The map

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

takes the upper half-disc $\{z = x + iy : \ |z| < 1 \ \text{ and } \ y > 0\}$ conformally to the first quadrant $\{w = u + iv : \ u > 0, \ v > 0\}$ (see fig. 17.4).

Indeed, if $z = x + iy$, we have so f maps the half-disc in the upper half-plane into the first quadrant. The inverse map, given by

$$g(w) = \frac{w - 1}{w + 1},$$

is clearly analytic in the first quadrant. Moreover, $|w + 1| > |w - 1|$ for all w in the first quadrant because the distance from w to -1 is greater than the distance from w to 1; thus g maps into the unit disc. Finally, an easy calculation shows that the imaginary part of $g(w)$ is positive whenever w is in the first quadrant. So g transforms the first quadrant into the desired half-disc and we conclude that f is conformal because g is the inverse of f.

To examine the action of f on the boundary, note that if $z = \mathrm{e}^{i\theta}$ belongs to the upper half-circle, then

$$f(z) = \frac{1 + \mathrm{e}^{i\theta}}{1 - \mathrm{e}^{i\theta}} = \frac{\mathrm{e}^{-i\theta/2} + \mathrm{e}^{i\theta/2}}{\mathrm{e}^{-i\theta/2} - \mathrm{e}^{i\theta/2}} = \frac{i}{\tan(\theta/2)}.$$

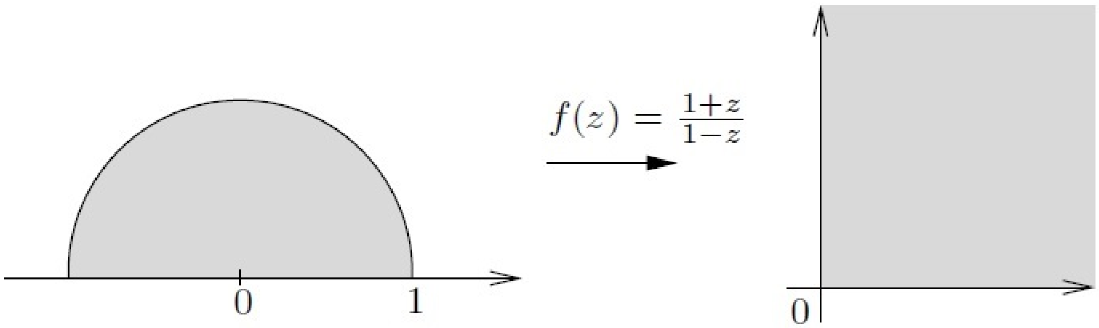


Figure 17.4

As θ travels from 0 to π we see that $f(\mathrm{e}^{i\theta})$ travels along the imaginary axis from infinity to 0. Moreover, if $z = x$ is real, then

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

is also real; and one sees from this, that f is actually a bijection from $(-1, 1)$ to the positive real axis, with $f(x)$ increasing from 0 to infinity as x travels from -1 to 1. Note also that $f(0) = 1$.

Exercise 17.3.1. 1. Show that for $h \in \mathbb{C}$, the translation map $f(z) = z + h$ is a conformal map from $\mathbb{C}$ to itself.

2. Show that the map $f(z) = \mathrm{e}^{iz}$ takes the half-strip $\left\{z = x + iy : \ -\frac{\pi}{2} < x < \frac{\pi}{2}, \ y > 0\right\}$ conformally to the half-disc $\{w = u + iv : \ |w| < 1, \ u > 0\}$.

17.4 Mӧbius Transformations

We have already seen that the functions of the form

$$f(z) = \frac{az+b}{cz+d} \tag{17.4.1}$$

is a linear fractional transformation. If $ad - bc \neq 0$, then $f(z)$ is called a Mӧbius Transformation. If f is a Mӧbius Transformation, then

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

is the inverse map of f. Also, if f and g are two linear fractional transformations, then their composition $f \circ g$ is also so. Hence, the set of all Mӧbius Transformations form a group under group composition.

Theorem 17.4.1. If f is a Mӧbius Transformation, then f is the composition of translations, dilations and inversion.

The fixed points of a Mӧbius Transformation (17.4.1) are the points where $f(z) = z$, that is,

$$cz^2 + (d-a)z - b = 0.$$

Hence a Mӧbius Transformation has at most two fixed points unless it is the identity transformation.

Now, let f be a Möbius Transformation and let a, b, c be distinct points in $\mathbb{C}_\infty$ such that $f(a) = \alpha,\ f(b) = \beta,\ f(c) = \gamma$. Suppose that g is another Möbius Transformation with the same property. Then $g^{-1} \circ f$ has a, b and c as fixed points and hence it is the identity transformation and thus, $f \equiv g$. Thus a Möbius Transformation is uniquely determined by its action on three points in $\mathbb{C}_\infty$.

Let z_2, z_3 and z_4 be points on $\mathbb{C}_\infty$. Define $f : \mathbb{C}_\infty \to \mathbb{C}_\infty$ by

$$\begin{aligned} f(z) &= \frac{z - z_3}{z - z_4} \cdot \frac{z_2 - z_4}{z_2 - z_3} && \text{if } z_2, z_3, z_4 \in \mathbb{C}_\infty \\ &= \frac{z - z_3}{z - z_4} && \text{if } z_2 = \infty \\ &= \frac{z_2 - z_4}{z - z_4} && \text{if } z_3 = \infty \\ &= \frac{z - z_3}{z_2 - z_3} && \text{if } z_4 = \infty. \end{aligned}$$

In any case, $f(z_2) = 1,\ f(z_3) = 0,\ f(z_4) = \infty$ and f is the only transformation having this property.

Definition 17.4.1. If $z_1 \in \mathbb{C}_\infty$, then the **cross ratio** of z_1, z_2, z_3 and z_4 is the image of z_1 under the unique Mö transformation which takes z_2 to 1, z_3 to 0 and z_4 to ∞. The cross ratio of z_1, z_2, z_3 and z_4 is denoted by (z_1, z_2, z_3, z_4).

For example, $(z_2, z_2, z_3, z_4) = 1$ and $(z, 1, 0, \infty) = z$. Also, if M is a Möbius map and w_2, w_3, w_4 are the points such that $Mw_2 = 1,\ Mw_3 = 0$ and $Mw_4 = \infty$, then $Mz = (z, w_2, w_3, w_4)$.

Theorem 17.4.2. If z_2, z_3 and z_4 are distinct points and T is any Möbius transformation, then

$$(z_1, z_2, z_3, z_4) = (T(z_1), T(z_2), T(z_3), T(z_4))$$

for any point z_1.

Proof. Let $S(z) = (z, z_2, z_3, z_4)$. Then S is a Möbius map. If $M = ST^{-1}$, then $M(T(z_2)) = 1,\ M(T(z_3)) = 0,\ M(T(z_4)) = \infty$. Hence, $ST^{-1}(z) = (z, T(z_2), T(z_3), T(z_4))$ for all $z \in \mathbb{C}_\infty$. In particular, if $z = T(z_1)$, the desired result follows. □

Theorem 17.4.3. If z_2, z_3, z_4 are distinct points in $\mathbb{C}_\infty$ and w_2, w_3, w_4 are also distinct points of $\mathbb{C}_\infty$, then there is one and only one Möbius transformation S such that $S(z_2) = w_2, S(z_3) = w_3, S(z_4) = w_4$.

Proof. Let $T(z) = (z, z_2, z_3, z_4)$, $M(z) = (z, w_2, w_3, w_4)$ and put $S = M^{-1}T$. Clearly, S has the desired property. If R is another Möbius transformation with $Rz_j = w_j$ for $j = 2, 3, 4$ then $R^{-1} \cdot S$ has three fixed points (z_2, z_3 and z_4). Hence, $R^{-1} \cdot S = I$ or $S = R$. □

It is well known that three points in the plane determine a circle. The next result explains when four points lie on a circle.

Theorem 17.4.4. Let z_1, z_2, z_3, z_4 be four distinct points in $\mathbb{C}_\infty$. Then (z_1, z_2, z_3, z_4) is a real number iff all four points lie on a circle.

Proof. Let $S : \mathbb{C}_\infty \to \mathbb{C}_\infty$ be defined by $S(z) = (z, z_2, z_3, z_4)$; then $S^{-1}(\mathbb{R})$=the set of z such that (z, z_2, z_3, z_4) is real. Hence, we will be finished if we can show that the image of $\mathbb{R}_\infty$ under a Möbius map us a circle.

Let

$$S(z) = \frac{az + b}{cz + d} \tag{17.4.2}$$

If $z = w \in \mathbb{R}$ and $w = S^{-1}(x)$ then $x = S(w)$ implies that $S(w) = \overline{S(w)}$. That is,

$$\frac{aw+b}{cw+d} = \frac{\overline{a}\overline{w}+\overline{b}}{\overline{c}\overline{w}+\overline{d}}$$

Cross multiplying this gives

$$(a\overline{c} - \overline{a}c)|w|^2 + (a\overline{d} - \overline{b}c)w + (b\overline{c} - d\overline{a})\overline{w} + (b\overline{d} - \overline{b}d) = 0 \tag{17.4.3}$$

If $a\overline{c}$ is real then $a\overline{c} - \overline{a}c = 0$; putting $\alpha = 2(a\overline{d} - \overline{b}c)$, $\beta = i(b\overline{d} - \overline{b}d)$ and multiplying (17.4.3) by i gives

$$0 = \text{Im}\,(\alpha w) - \beta = \text{Im}\,(\alpha w - \beta) \tag{17.4.4}$$

since β is real. That is, w lies on the line determined by (17.4.4) for fixed α and β. If $a\overline{c}$ is not real then (17.4.3) becomes

$$|w|^2 + \overline{\gamma}w + \gamma\overline{w} - \delta = 0 \tag{17.4.5}$$

for some constants γ in $\mathbb{C}$, δ in $\mathbb{R}$. Hence,

$$|w + \gamma| = \lambda \tag{17.4.6}$$

where

$$\lambda = \sqrt{|\gamma|^2 + \delta} = \left|\frac{ad - bc}{\overline{a}c - a\overline{c}}\right| > 0.$$

Since γ and λ are independent of x and since (17.4.6) is the equation of a circle, the proof is done. □

Theorem 17.4.5. A Mӧbius transformation takes circles into circles.

Proof. Let Γ be any circle in $\mathbb{C}_\infty$ and let S be any Mӧbius transformation. Let z_2, z_3, z_4 be three distinct points on Γ and put $w_j = S(z_j)$ for $j = 2, 3, 4$. Then w_2, w_3, w_4 determine a circle Γ'. We claim that $S(\Gamma) = \Gamma'$. In fact, for any z in $\mathbb{C}_\infty$,

$$(z, z_2, z_3, z_4) = (S(z), w_2, w_3, w_4) \tag{17.4.7}$$

by theorem 17.4.2. By the preceding theorem, if z is on Γ, then both sides of (17.4.7) are real. But this says that $S(z) \in \Gamma'$. □

Now, let Γ and Γ' be two circles in $\mathbb{C}_\infty$ and let $z_2, z_3, z_4 \in \Gamma$; $w_2, w_3, w_4 \in \Gamma'$. Put $R(z) = (z, z_2, z_3, z_4)$, $S(z) = (z, w_2, w_3, w_4)$. Then $T = S^{-1} \circ R$ maps Γ onto Γ'. In fact, $T(z_j) = w_j$ for $j = 2, 3, 4$ and, as in the above proof, it follows that $T(\Gamma) = \Gamma'$.

Theorem 17.4.6. For any given circles Γ and Γ' in $\mathbb{C}_\infty$, there is a Mӧbius transformation T such that $T(\Gamma) = \Gamma'$. Furthermore we can specify that T takes any three points in Γ onto any three points on Γ'. If we specify $T(z_j)$ for $j = 2, 3, 4$ (distinct z_j in Γ) then T is unique.

Now that we know that a Mӧbius map takes circles to circles, the next question is: What happens to the inside and the outside of these circles? To answer this we introduce some new concepts.

Definition 17.4.2. Let Γ be a circle through points z_2, z_3, z_4. The points z, z^* in $\mathbb{C}_\infty$ are said to be symmetric with respect to Γ is

$$(z^*, z_2, z_3, z_4) = \overline{(z, z_2, z_3, z_4)}. \tag{17.4.8}$$

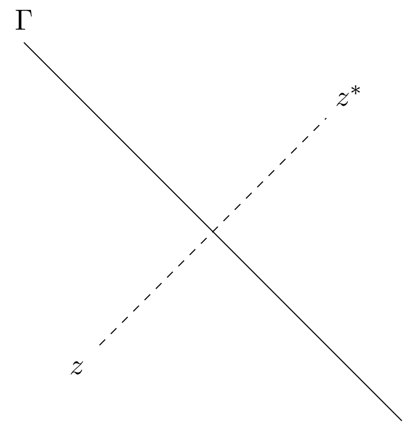


Figure 17.5

As it stands, this definition not only depends on the circle but also on the points z_2, z_3, z_4.

Also by theorem 17.4.4, z is symmetric to itself with respect to Γ if and only if $z \in \Gamma$. Let us investigate what it means for z and z^* to be symmetric. If Γ is a straight line then our linguistic prejudices lead us to believe that z and z^* are symmetric with respect to Γ if the line through z and z^* are the same distance from Γ but on the opposite sides of it (see fig. 17.5.

If Γ is a straight line then, choosing $z_4 = \infty$, (17.4.8) becomes

$$\frac{z^* - z_3}{z_2 - z_3} = \frac{\overline{z} - \overline{z}_3}{\overline{z}_2 - \overline{z}_3}.$$

This gives $|z^* - z_3| = |z - z_3$. Since z_3 was not specified, we have that z and z^* are equidistant from each point on Γ. Also,

$$\mathrm{Im}\frac{z^* - z_3}{z_2 - z_3} = \mathrm{Im}\frac{\overline{z} - \overline{z}_3}{\overline{z}_2 - \overline{z}_3} = -\mathrm{Im}\frac{z - z_3}{z_2 - z_3}.$$

Hence, we have (unless $z \in \Gamma$) that z and z^* in different half planes determined by Γ. It now follows that $[z, z^*]$ is perpendicular to Γ.

Now, suppose that $\Gamma = \{z : \ |z - a| = R\}$ $(0 < R < \infty)$. Let z_2, z_3, z_4 be points in Γ. Using (17.4.8) and theorem 17.4.2 for a number of Mӧbius transformations gives

$$\begin{aligned}
(z^*, z_2, z_3, z_4) &= \overline{(z, z_2, z_3, z_4)} \\
&= \overline{(z - a, z_2 - a, z_3 - a, z_4 - a)} \\
&= \left(\overline{z} - \overline{a}, \frac{R^2}{z_2 - a}, \frac{R^2}{z_3 - a}, \frac{R^2}{z_4 - a}\right) \\
&= \left(\frac{R^2}{\overline{z} - \overline{a}}, z_2 - a, z_3 - a, z_4 - a\right) \\
&= \left(\frac{R^2}{\overline{z} - \overline{a}} + a, z_2, z_3, z_4\right).
\end{aligned}$$

Hence $z^* = a + R^2(\overline{z} - \overline{a})^{-1}$ or $(z^* - z)(\overline{z} - \overline{a}) = R^2$. From this it follows that

$$\frac{z^* - a}{z - a} = \frac{R^2}{|z - a|^2} > 0,$$

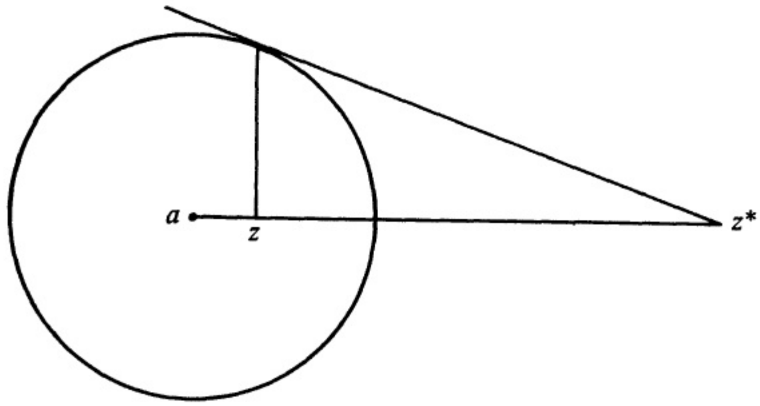


Figure 17.6

so that z^* lies on the ray $\{a+t(z-a) : 0 < t < \infty\}$ from a through z. Using the fact that $|z-a||z^*-a| = R^2$, we obtain z^* from z (if z lies inside Γ) as in the figure 17.6. That is, let L be the ray from a through z. Construct a line P perpendicular to L at z and at the point where P intersects Γ; construct the tangent to Γ. The point of intersection of this tangent with L is the point z^*. Thus, the points a and ∞ are symmetric with respect to Γ.

Theorem 17.4.7. (Symmetry Principle) If a Mӧbius transformation T takes a circle Γ_1 onto the circle Γ_2, then any pair of points symmetric with respect to Γ_1 are mapped by T onto a pair of points symmetric with respect to Γ_2.

Few Probable Questions

---

1. Define conformal maps. Show that a map f, analytic at z_0 with $f'(z_0) \neq 0$ is one-to-one in a neighbourhood of z_0.

2. Show that a one-to-one analytic function in a domain is conformal there.

3. If a function f is analytic in a simply connected domain D and on its boundary C (which is a simple closed contour), then f one-to-one on C implies it is so in D.

4. Define conformally equivalent regions. Show that the upper half disc $\{z : |z| < 1, \text{ Im } z > 0\}$ is conformally equivalent to the first quadrant $\{w = u + iv : u > 0, v > 0\}$.

5. Define cross ratio of z_1, z_2, z_3, z_4. For $z_1, z_2, z_3, z_4 \in \mathbb{C}_\infty$, show that the cross ratio is a real number if and only if all the four points lie on a circle.

6. Show that a Mӧbius transformation takes circles into circles. When are two points said to be symmetric with respect to a circle Γ?

Unit 18

Course Structure

- Schwarz principle of symmetry
- Schwarz-Christoffel formula (statement only)
- Applications of Schwarz-Christoffel formula.

18.1 Introduction

In mathematics, the Schwarz reflection principle, or the Schwarz principle of symmetry, is a way to extend the domain of definition of a complex analytic function, i.e., it is a form of analytic continuation. It states that if an analytic function is defined on the upper half-plane, and has well-defined (non-singular) real values on the real axis, then it can be extended to the conjugate function on the lower half-plane as we shall see. This unit is also dedicated to a preliminary study of the Schwarz-Christoffel mapping which is mainly a conformal transformation of the upper half-plane onto the interior of a simple polygon. Schwarz–Christoffel mappings are used in potential theory and some of its applications, including minimal surfaces and fluid dynamics. They are named after Elwin Bruno Christoffel and Hermann Amandus Schwarz.

Objectives

After reading this unit, you will be able to

- define symmetric open set and deduce the symmetry principle
- deduce the Schwarz principle of symmetry
- have some preliminary idea about the Schwarz Christoffel mappings

18.2 Schwarz Principle of Symmetry

In real analysis, there are various situations where one wishes to extend a function from a given set to a larger one. Several techniques exist that provide extensions for continuous functions, and more generally for functions with varying degrees of smoothness. Of course, the difficulty of the technique increases as we impose more conditions on the extension. The situation is very different for holomorphic functions. Not only are these functions indefinitely differentiable in their domain of definition, but they also have additional characteristically rigid properties, which make them difficult to mould. For example, there exist holomorphic functions in a disc which are continuous on the closure of the disc, but which cannot be continued (analytically) into any region larger than the disc.

Let Ω be an open subset of $\mathbb{C}$ that is symmetric with respect to the real line, that is

$$z \in \Omega \text{ if and only if } \overline{z} \in \Omega.$$

Let Ω^+ denote a part of Ω that lies in the upper half-plane and Ω^- that part that lies in the lower half-plane (see fig.18.1 for illustration).

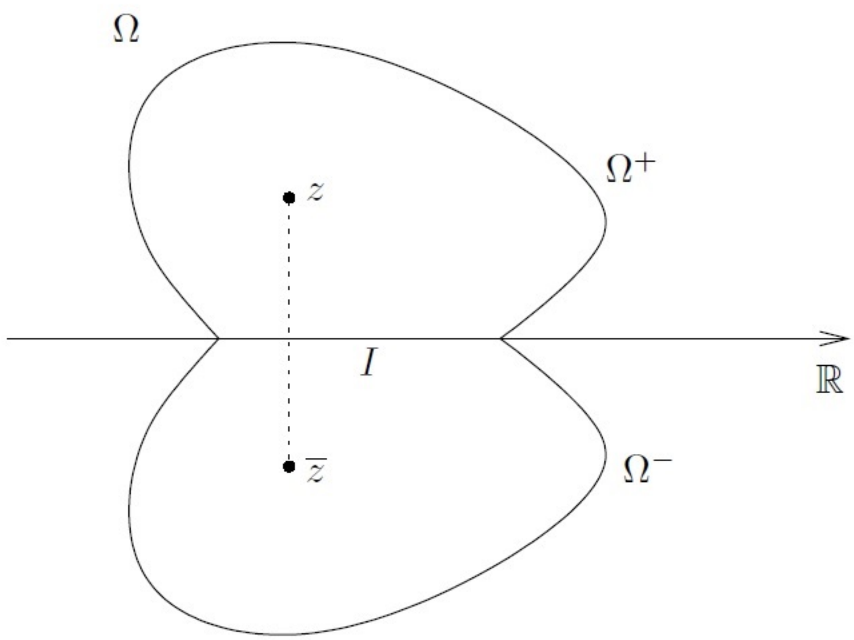


Figure 18.1

Also, let $I = \Omega \cap \mathbb{R}$ so that I denotes the interior of that part of the boundary of Ω^+ and Ω^- that lies on the real axis. Then we have

$$\Omega^+ \cup I \cup \Omega^- = \Omega$$

and the only interesting case of the next theorem occurs, of course, when I is non-empty.

Theorem 18.2.1. (Symmetry principle) If f^+ and f^- are analytic in Ω^+ and Ω^- respectively, that extend continuously to I and

$$f^+(x) = f^-(x), \;\; \forall\, x \in I,$$

then the function f defined on Ω by

$$\begin{aligned} f(z) &= f^+(z) \quad \text{if } z \in \Omega^+ \\ &= f^+(z) = f^-(z) \quad \text{if } z \in I \\ &= f^-(z) \quad \text{if } z \in \Omega^- \end{aligned}$$

is analytic on all of Ω.

Proof. One notes first that f is continuous throughout Ω. The only difficulty is to prove that f is analytic at points of I. Suppose D is a disc centred at a point on I and entirely contained in Ω. We prove that f is analytic in D by Morera's theorem. Suppose T is a triangle in D. If T does not intersect I, then

$$\int_T f(z)dz = 0$$

since f is analytic in the upper and lower half-discs. Suppose now that one side or vertex of T is contained in I, and the rest of T is in, say, the upper half-disc. If T_ϵ is the triangle obtained from T by slightly raising the edge or vertex which lies on I, we have $\int_{T_\epsilon} f = 0$ since T_ϵ is entirely contained in the upper half-disc an (illustration of the case when an edge lies on I is given in Figure 18.2). When we let $\epsilon \to 0$, and by continuity, we conclude that

$$\int_T f(z)dz = 0$$

If the interior of T intersects I, we can reduce the situation to the previous one by writing T as the union of

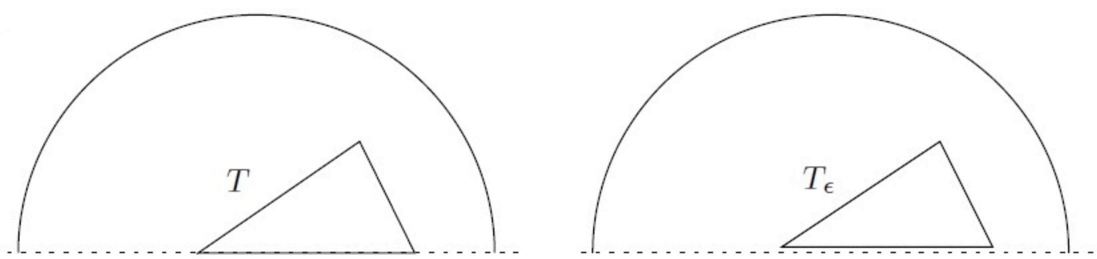


Figure 18.2: Raising a vertex

triangles each of which has an edge or vertex on I as shown in Figure 18.3. By Morera's theorem we conclude that f is analytic in D, as was to be shown. □

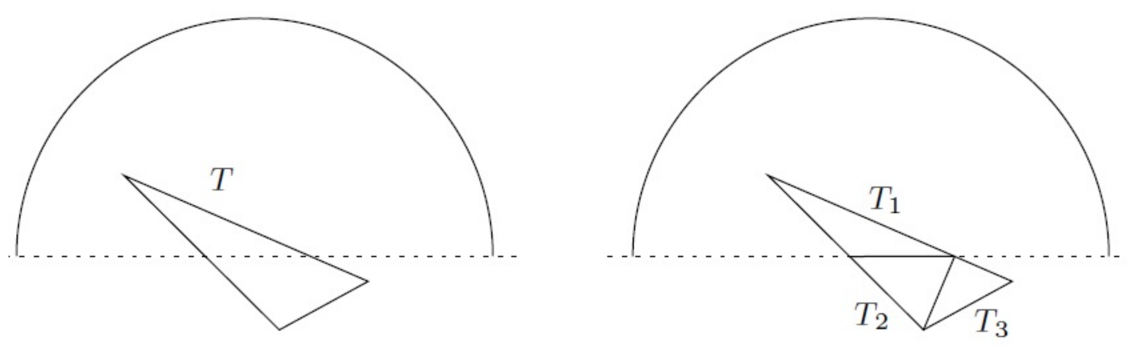


Figure 18.3: Splitting a triangle

We can now state the extension principle, where we use the above notation.

Theorem 18.2.2. (Schwarz reflection principle) Suppose that f is a analytic function in Ω^+ that extends continuously to I and such that f is real-valued on I. Then there exists a function F analytic in all of Ω such that $F = f$ on Ω^+.

Proof. The idea is simply to define $F(z)$ for $z \in \Omega^-$ by

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

To prove that F is analytic in Ω^- we note that if $z, z_0 \in \Omega^-$, then $\overline{z}, \overline{z}_0 \in \Omega^+$ and hence, the power series expansion of f near $\overline{z}_0$ gives

$$f(\overline{z}) = \sum a_n(\overline{z} - \overline{z}_0)^n.$$

As a consequence we see that

$$F(z) = \sum \overline{a}_n(z - z_0)^n$$

and F is analytic in Ω^-. Since f is real valued on I we have, $\overline{f(x)} = f(x)$, whenever $x \in I$ and hence F extends continuously up to I. The proof is complete once we invoke the symmetry principle. $\square$

18.3 Schwarz Christoffel formula

We represent the unit vector which is tangent to a smooth are C at a point z_0 by the complex number t, and we let the number τ denote the unit vector tangent to the image Γ of C at the corresponding point w_0 under a transformation $w = f(z)$. We assume that f is analytic at z_0 and that $f'(z_0) \neq 0$. We know,

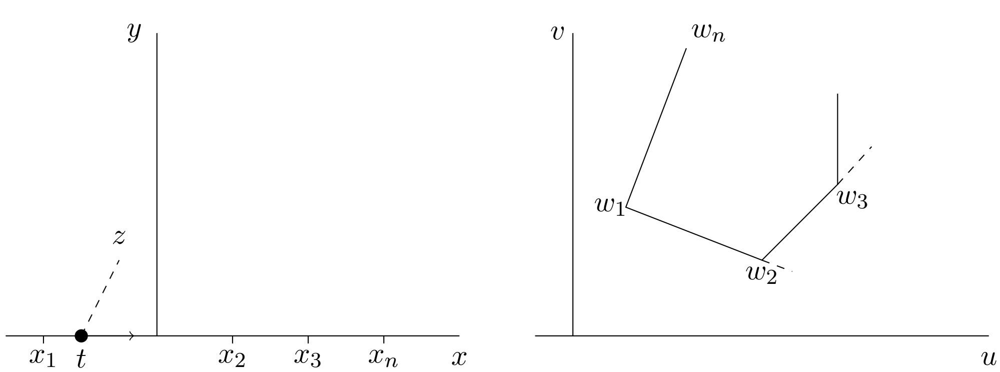
$$\arg \tau = \arg f'(z_0) + \arg t \tag{18.3.1}$$

In particular, if C is a segment of the x-axis with positive sense to the right then $t = 1$ and $\arg t = 0$ at each point $z_0 = x$ on C. In that case, equation (18.3.1) becomes

$$\arg \tau = \arg f'(x) \tag{18.3.2}$$

If $f'(z)$ has a constant argument along that segment, it follows that $\arg \tau$ is constant. Hence, the image Γ of C is also a segment of a straight line.

Let us now construct a transformation $w = f(z)$ that maps the whole x-axis onto a polygon of n sides, where $x_1, x_2, \ldots, x_{n-1}$ and ∞ are the points on that axis whose images are to be the vertices of the polygon and where $x_1 < x_2 < \cdots < x_{n-1}$. The vertices are the n points $w_j = f(x_j)(j = 1, 2, \ldots n-1)$ and $w_n = f(\infty)$. The function f should be such that $\arg f'(z)$ jumps from one constant value to another at the points $z = x_j$ as the point z traces out the x-axis. If the function f is chosen such that



$$f'(z) = A(z - x_1)^{k_1}(z - x_2)^{k_2} \cdots (z - x_{n-1})^{k_{n-1}}. \tag{18.3.3}$$

where A is a complex constant and each k_j is a real constant, then the argument of $f'(z)$ changes in the prescribed manner as z describes the real axis. This is seen writing the argument of the derivative (18.3.3) as

$$\arg f'(z) = \arg A - k_1 \arg(z - x_1) - k_2 \arg(z - x_2) - \cdots - k_{n-1} \arg(z - x_{n-1}) \tag{18.3.4}$$

When $z = x$ and $x < x_1$,

$$\arg(z - x_1) = \arg(z - x_2) = \cdots = \arg(z - x_{n-1}) = \pi$$

When $x_1 < x < x_2$, the argument $\arg(z - x_1)$ is 0 and each of the other arguments is π. According to equation (18.3.4), then $\arg f'(z)$ increases abruptly by the angle $k_1\pi$ as z moves to the right through the point $z = x_1$. It again jumps in value, by the amount $k_2\pi$, as z passes through the point x_2, etc.

In view of (18.3.2), the unit vector τ is constant in direction as z moves from x_{j-1} to x_j; the point w thus moves in that fixed direction along a straight line. The direction of τ changes abruptly by the angle $k_j\pi$ at the image point w_j of x_j. Those angles $k_j\pi$ are the exterior angles of the polygon described by the point w.

The exterior angles can be limited to angles between $-\pi$ to π, in which case $-1 < k_j < 1$. We assume that the sides of the polygon never cross one another and that the polygon is given is given a positive or counterclockwise orientation. The sum of the exterior angles of a *closed* polygon is, then 2π and the exterior angle at the vertex w_n which is the image of the point $z = \infty$, can be written

$$k_n\pi = 2\pi - (k_1 + k_2 + \cdots + k_{n-1})\pi$$

Thus the numbers k_j must necessarily satisfy the conditions

$$k_1 + k_2 + \cdots + k_{n-1} + k_n = 2, \qquad -1 < k_j < 1 \quad (j = 1, 2, \ldots, n) \tag{18.3.5}$$

Note that $k_n = 0$ if $k_1 + k_2 + \cdots + k_{n-1} = 2$. This means that the direction of τ does not change at the point w_n. So, w_n is not a vertex, and the polygon has $n - 1$ sides.

Few Probable Questions

---

1. State and prove the symmetry principle.
2. State and prove the Schwarz principle of symmetry.

Unit 19

Course Structure

- Univalent functions, general theorems

19.1 Introduction

In mathematics, in the branch of complex analysis, an analytic function on an open subset of the complex plane is called univalent if it is injective. The theory of univalent functions is an old subject, born around the turn of the century, yet it remains an active field of research. This unit introduces the class S of univalent functions and some of its subclasses defined by geometric conditions. A number of basic questions are answered by elementary methods. Most of the results concerning the class S are direct consequences of the area theorem, which may be regarded as the cornerstone of the entire subject.

Objectives

After reading this unit, you will be able to

- define univalent functions
- define normal families of analytic functions and related terms
- learn about preliminary types of univalent functions

19.2 Normal Families

A family $\mathscr{F}$ of functions analytic in a domain D is called a **normal family** if every sequence of functions $f_n \in \mathscr{F}$ has a subsequence which converges uniformly on each compact subset of D.

A family $\mathscr{F}$ is compact if whenever $f_n \in \mathscr{F}$ and $f_n(z) \to f(z)$ uniformly on compact subsets of D, it follows that $f \in \mathscr{F}$. The defining property of a normal family is analogous to the Bolzano-Weierstrass property of a bounded set of points in Euclidean space. Compact families are analogous to closed sets.

A family $\mathscr{F}$ of functions analytic in D is said to be **locally bounded** is the functions are **uniformly bounded** on each closed disc $B \subset D$; that is, if $|f(z)| \leq M$ for all $z \in B$ and for every $f \in \mathscr{F}$, where the bound M depends only on B. In view of the Heine Borel theorem, it then follows that the functions are uniformly bounded on each compact subset of D. If $\mathscr{F}$ is a locally bounded family of analytic functions, then by the Cauchy integral formula, the family of derivatives $\{f' : f \in \mathscr{F}\}$ is also locally bounded.

We have the following theorem concerning locally bounded family of analytic functions.

Theorem 19.2.1. A necessary and sufficient condition for a family of analytic functions to be locally bounded is that, it is normal.

19.3 Univalent Functions

Definition 19.3.1. A single valued function f is said to be univalent (or schlicht) in a domain $D \subset \mathbb{C}$ if it never takes the same value twice; that is, if $f(z_1) \neq f(z_2)$ for all points z_1 and z_2 in D with $z_1 \neq z_2$. The function f is said to be locally univalent at a point $z_0 \in D$ if it is univalent in some neighbourhood of z_0.

For analytic functions f, the condition $f'(z_0) \neq 0$ is equivalent to local univalence at z_0 as we have seen previously. An analytic univalent function is a conformal mapping because of its angle-preserving property.

We shall be concerned primarily with the class S of functions f analytic and univalent in the unit disc $\mathbb{D}$, satisfied by the conditions $f(0) = 0$ and $f'(0) = 1$. Thus each $f \in S$ has a Taylor series expansion of the form

$$f(z) = z + a_2 z^2 + a_3 z^3 + \cdots, \quad |z| < 1.$$

The Riemann mapping theorem states that for any simply connected domain D, which is a proper subset of the complex plane and any point $\zeta \in D$, there is a unique function f which maps D conformally onto the unit disc and has properties $f(\zeta) = 0$ and $f'(\zeta) > 0$. That is, it says that any simply connected domain D, which is a proper subset of $\mathbb{C}$, is conformally equivalent to the unit disc $\mathbb{D}$.

In view of the Riemann mapping theorem, most of the geometric theorems concerning functions of class S are readily translated to statements about univalent functions in arbitrary simply connected domains with more than one boundary point. The leading example of a function of class S is the *Koebe* function

$$k(z) = z(1-z)^{-2} = z + 2z^2 + 3z^3 + \cdots$$

The Koebe function maps the disc $\mathbb{D}$ onto the entire plane minus the part of the negative real axis from $-1/4$ to infinity. This is best seen by writing

$$k(z) = \frac{1}{4}\left(\frac{1+z}{1-z}\right)^2 - \frac{1}{4}$$

and observing that the function

$$w = \frac{1+z}{1-z}$$

maps $\mathbb{D}$ conformally onto the right half-plane $\text{Re}\{w\} > 0$.

Other simple examples of functions in S are

1. $f(z) = z$, the identity mapping;
2. $f(z) = z(1-z)^{-1}$, which maps $\mathbb{D}$ conformally onto the half plane $\text{Re}\{w\} > -1/2$;
3. $f(z) = z(1-z^2)^{-1}$, which maps $\mathbb{D}$ onto the entire plane minus the two half lines $\frac{1}{2} \leq x < \infty$ and $-\infty < x \leq -\frac{1}{2}$;
4. $f(z) = z - \frac{1}{2}z^2 = \frac{1}{2}[1-(1-z)^2]$, which maps $\mathbb{D}$ onto the interior of a cardioid.

The sum of two functions in S need not be univalent. For example, the sum of $z(1-z)^{-1}$ and $z(1+iz)^{-1}$ has a derivative which vanishes at $\frac{1}{2}(1+i)$ (verify!). However, the class S is preserved under a number of elementary transformations.

1. **Conjugation:** If $f \in S$ and
$$g(z) = \overline{f(\overline{z})} = z + \overline{a}_2 z^2 + \overline{a}_3 z^3 + \cdots ,$$
then $g \in S$.

2. **Rotation:** If $f \in S$ and
$$g(z) = \mathrm{e}^{-i\theta} f(\mathrm{e}^{-i\theta} z),$$
then $g \in S$.

3. **Dilation:** If $f \in S$ and
$$g(z) = \frac{1}{r} f(rz), \quad \text{where } 0 < r < 1,$$
then $g \in S$.

4. **Range Transformation:** If $f \in S$ and ψ is a function analytic and univalent on the range of f, with $\psi(0) = 0$ and $\psi'(0) = 1$, then $g = \psi \circ f \in S$.

5. **Omitted-value transformation:** If $f \in S$ and $f(z) \neq \omega$, then
$$g = \frac{\omega f}{\omega - f} \in S.$$

6. **Square-root transformation:** If $f \in S$ and $g(z) = \sqrt{f(z^2)}$, the $g \in S$.

The square root transformation requires a word of explanation. Since $f(z) = 0$ only at the origin, a single valued branch of the square root may be chosen by writing

$$\begin{aligned} g(z) &= \sqrt{f(z^2)} = z\{1 + a_2 z^2 + a_3 z^4 + \cdots\}^{\frac{1}{2}} \\ &= z + c_3 z^3 + c_5 z^5 + \cdots , \quad |z| < 1. \end{aligned}$$

Note that g is an odd analytic function, so that $g(-z) = -g(z)$. If $g(z_1) = g(z_2)$, then $f(z_1^2) = f(z_2^2)$ and $z_1^2 = z_2^2$, which gives $z_1 = \pm z_2$. But, if $z_1 = -z_2$, then $g(z_1) = g(z_2) = -g(z_1)$. Thus $g(z_1) = 0$, and $z_1 = 0$. This shows that $z_1 = z_2$ in either case, proving that g is univalent.

Closely related to S, is the class Σ of functions

$$g(z) = z + b_0 + b_1 z^{-1} + b_2 z^{-2} + \cdots$$

is analytic and univalent in the domain $\mathbb{E} = \{z : |z| > 1\}$, exterior to the domain $\mathbb{D}$, except for a simple pole at infinity with residue 1. Each function $g \in \Sigma$ maps $\mathbb{E}$ onto the complement of a compact connected set E. It is useful to consider the subclass Σ' of functions $g \in \Sigma$ for which $0 \in E$; that is, for which $g(z) \neq 0$ in $\mathbb{E}$. Any function $g \in \Sigma$ will belong to Σ' after suitable adjustment of the constant term b_0. Such an adjustment will only translate the range of g and will not destroy the univalence.

For each $f \in S$, the function

$$g(z) = \left\{ f\left(\frac{1}{z}\right) \right\}^{-1} = z - a_2 + (a_2^2 - a_3) z^{-1} + \cdots$$

belongs to Σ'. This transformation is called an inversion. It actually establishes a one-to-one correspondence between S and Σ'. The class Σ' is preserved under the square-root transformation

$$G(z) = \sqrt{g(z^2)} = z\{1 + b_0 z^{-2} + b_1 z^{-4} + \cdots\}^{\frac{1}{2}}.$$

It is important to observe that this operation cannot be applied to every function $g \in \Sigma$, but is permissible only if $g \in \Sigma'$, because the square root will introduce a branch point wherever $g(z^2) = 0$.

Sometimes, it is convenient to consider the subclass Σ_0 consisting of all $g \in \Sigma$ with $b_0 = 0$. Obviously this can be achieved by suitable translation, but it may not be possible to translate a given function $g \in \Sigma$ simultaneously to both Σ_0 and Σ'.

It is also useful to distinguish the subclass $\tilde{\Sigma}$ of all functions $g \in \Sigma$ whose omitted set E has two dimensional Lebesgue measure zero. The functions $g \in \tilde{\Sigma}$ will be called full mappings.

Few Probable Questions

---

1. Define normal family.
2. Define locally bounded family of analytic functions in a domain D.
3. Define univalent function on a domain D. Show that for an analytic function f on D, $f'(z_0) \neq 0$ at $z_0 \in D$ is equivalent to the local univalence of f at z_0.

Unit 20

Course Structure

- Area theorem
- Growth and Distortion theorems

20.1 Introduction

The univalence of a function

$$g(z) = z + b_0 + \sum_{n=1}^{\infty} b_n z^{-1}, \qquad |z| > 1$$

places strong restriction on the size of the Laurent coefficients b_n, $n = 1, 2, \ldots$. This is expressed by the area theorem, which is fundamental to the theory of univalent functions. The reason for the name will be apparent from the proof. Gronwall discovered the theorem in 1914.

Objectives

After reading this unit, you will be able to

- deduce the area theorem and related results
- deduce the growth and distortion theorems

20.2 Area Theorem

Theorem 20.2.1. Area Theorem: If $g \in \sum$, then

$$\sum_{n=1}^{\infty} n|b_n|^2 \leq 1$$

with equality if and only if $g \in \tilde{\sum}$.

Proof. Let E be the set omitted by g. For $r > 1$, let C_r be the image under g of the circle $|z| = r$. Since g is univalent, C_r is a simple closed curve which encloses a domain $E_r \supset E$. By Green's theorem, the area of E_r is

$$\begin{aligned} A_r &= \frac{1}{2i}\int_{C_r} \overline{w}dw = \frac{1}{2i}\int_{|z|=r} \overline{g(z)}g'(z)dz \\ &= \frac{1}{2}\int_0^{2\pi} \{r\,\mathrm{e}^{-i\theta} + \sum_{n=0}^{\infty} \overline{b_n} r^{-n}\,\mathrm{e}^{in\theta}\} \times \{1 - \sum_{v=1}^{\infty} v b_v r^{-v-1}\,\mathrm{e}^{-i(v+1)\theta}\} r\,\mathrm{e}^{i\theta}\,d\theta \\ &= \pi\{r^2 - \sum_{n=1}^{\infty} n|b_n|^2 r^{-2n}\}, \qquad r > 1 \end{aligned} \tag{20.2.1}$$

Let r decrease to 1, we obtain

$$m(E) = \pi\{r^2 - \sum_{n=1}^{\infty} n|b_n|^2\}$$

where $m(E)$ is the outer measure of E. Since, $m(E) \geq 0$, this proves the theorem. □

An immediate corollary is the inequality $|b_n| \leq n^{-1/2}$, $n = 1, 2, \ldots$. This inequality is not sharp if $n \geq 2$, since the function

$$g(z) = z + n^{-1/2} z^{-n}$$

is not univalent. Indeed, its derivative

$$g'(z) = 1 - n^{1/2} z^{-n-1}$$

vanishes at certain points in $\mathbb{E}$ if $n \geq 2$. However, the inequality $|b_1| \leq 1$ is sharp and has important consequences.

Corollary 20.2.1. If $g \in \sum$, then $|b_1| \leq 1$, with equality if and only if g has the form

$$g(z) = z + b_0 + \frac{b_1}{z}, \qquad |b_1| = 1$$

This is a conformal mapping of $\mathbb{E}$ onto the complement of a line segment of length 4.

From this result it is a short step to a theorem of Bieberbach estimating the second coefficient a_2 of a function of class S. This theorem was given in 1916 and was the main basis for the famous $Bieberbach$ $conjecture$.

Theorem 20.2.2. (Bieberbach's Theorem). If $f \in S$, then $|a_2| \leq 2$, with equality if and only if f is a rotation of the Koebe function.

Proof. A square-root transformation and an inversion applied to $f \in S$ will produce a function

$$g(z) = \{f(1/z)\}^{-1/2} = z - (a_2/2)z^{-1} + \cdots$$

of class $\sum$. Thus $|a_2| \leq 2$, by the corollary to the area theorem. Equality occurs if and only if g has the form

$$g(z) = z - \mathrm{e}^{i\theta}/z$$

A simple calculation shows that this is equivalent to

$$f(\zeta) = \zeta(1 - \mathrm{e}^{i\theta}\zeta)^{-2} = \mathrm{e}^{-i\theta} k(\mathrm{e}^{i\theta}\zeta),$$

a rotation of the Koebe function. □

As a first application of Bieberbach's theorem, we shall now prove a famous covering theorem due to Koebe. Each function $f \in S$ is an open mapping with $f(0) = 0$, so its range contains some disk centered at the origin. As early as 1907, Koebe discovered that the ranges of all functions in S contain a common disk $|w| < \rho$, where ρ is an absolute constant. The Koebe function shows that $\rho \leq \frac{1}{4}$, and Bieberbach later established Koebe's conjecture that ρ may be taken to be $\frac{1}{4}$.

Theorem 20.2.3. (Koebe One-Quarter Theorem): The range of every function of class S contains the disk $\{w : |w| < \frac{1}{4}\}$.

Proof. If a function $f \in S$ omits the value $\omega \in \mathbb{C}$, then

$$g(z) = \frac{\omega f(z)}{\omega - f(z)} = z + \left(a_2 + \frac{1}{\omega}\right) z^2 + \cdots$$

is analytic and univalent in $\mathbb{D}$. This is the omitted-value transformation, which is the composition of f with a linear fractional mapping. Since, $g \in S$, Bieberbach's theorem gives

$$\left|a_2 + \frac{1}{\omega}\right| \leq 2$$

Combined with the inequality $|a_2| \leq 2$ this shows that $|1/\omega| \leq 4$, or $|\omega| \geq \frac{1}{4}$. Thus every omitted value must lie outside the disk $|w| < \frac{1}{4}$. □

This proof actually shows that the Koebe function and its rotations are the only functions in S which omit a value of modulus $\frac{1}{4}$. Thus the range of every other function in S covers a disk of larger radius.

It should be observed that $univalence$ is the key to Koebe's theorem. For example, the analytic functions

$$f_n(z) = \frac{1}{n}(\mathrm{e}^{nz} - 1), \qquad n = 1, 2, \ldots,$$

have the properties $f_n(0) = 0$ and $f_n'(0) = 1$, yet f_n omits the value $-1/n$, which may be chosen arbitrarily close to the origin.

20.3 Growth and Distortion Theorems

Bieberbach's inequality $|a_2| \leq 2$ has further implications in the geometric theory of conformal mapping. One important consequence is the $Koebe\ distortion\ theorem$, which provides sharp upper and lower bounds for $|f'(z)|$ as f ranges over the class S, The term "distortion" arises from the geometric interpretation $|f'(z)|$ as the infinitesimal magnification factor of arclength under the mapping f, or from that of the Jacobian $|f'(z)|^2$ as the infinitesimal magnification factor of area. The following theorem gives a basic estimate which leads to the distortion theorem and related results.

Theorem 20.3.1. For each $f \in S$,

$$\left|\frac{zf''(z)}{f'(z)} - \frac{2r^2}{1-r^2}\right| \leq \frac{4r}{1-r^2}, \qquad |z| = r < 1 \tag{20.3.1}$$

Proof. Given $f \in S$, fix $\zeta \in \mathbb{D}$ and perform a disk automorphism to construct

$$F(z) = \frac{f\left(\frac{z+\zeta}{1+\bar{\zeta}z}\right) - f(\zeta}{(1-|\zeta|^2)f'(\zeta)} = z + A_2(\zeta)z^2 + \cdots. \tag{20.3.2}$$

Then $F \in S$ and a calculation gives

$$A_2(\zeta) = \frac{1}{2}\{(1-|\zeta|^2)\frac{f''(\zeta)}{f'(\zeta)} - 2\zeta\}.$$

But by Bieberbach's theorem, $|A_2(\zeta)| \leq 2$. Simplifying this inequality and replacing ζ by z, we obtain the inequality (20.3.1). A suitable rotation of the Koebe function shows that the estimate is sharp for each $z \in \mathbb{D}$. □

Theorem 20.3.2. (Distortion Theorem). For each $f \in S$

$$\frac{1-r}{(1+r)^3} \leq |f'(z)| \leq \frac{1+r}{(1-r)^3}, \qquad |z| = r < 1 \tag{20.3.3}$$

For each $z \in \mathbb{D}$, $z \neq 0$, equality occurs if and only if f is a suitable rotation of the Koebe function.

Proof. Since, an inequality $|\alpha| \leq c$ implies $-c \leq Re\{\alpha\} \leq c$, it follows from (20.3.1) that

$$\frac{2r^2-4r}{1-r^2} \leq Re\{\frac{zf''(z)}{f'(z)}\} \leq \frac{2r^2+4r}{1-r^2}$$

Because $f'(z) \neq 0$ and $f'(0) = 1$, we can choose a single-valued branch of $\log f'(z)$ which vanishes at the origin. Now, observe that

$$Re\{\frac{zf''(z)}{f'(z)}\} = r\frac{\partial}{\partial r}Re\{\log f'(z)\}, \qquad z = \mathrm{e}^{i\theta}.$$

Hence,

$$\frac{2r^2-4r}{1-r^2} \leq \frac{\partial}{\partial r}|f'(r\,\mathrm{e}^{r\theta})| \leq \frac{2r^2+4r}{1-r^2} \tag{20.3.4}$$

Holding θ fixed, integrate with respect to r from 0 to R. A calculation then produces the inequality

$$\log\frac{1-R}{(1+R)^3} \leq \log|f'(R\,\mathrm{e}^{i\theta})| \leq \log\frac{1+R}{(1-R)^3},$$

ad the distortion theorem follows by exponentiation.

A suitable rotation of the Koebe function, whose derivative is

$$k'(z) = \frac{1+z}{(1-z)^3}, \tag{20.3.5}$$

shows that both estimates of $|f'(z)|$ are best possible. Furthermore, whenever equality occurs for $z = R\,\mathrm{e}^{i\theta}$ in either the upper or the lower estimate of (20.3.3), the equality must hold in the corresponding part of (20.3.4) for all r, $0 \leq r \leq R$. In particular,

$$Re\{\mathrm{e}^{i\theta}\frac{f''(0)}{f'(0)}\} = \pm 4,$$

which implies that $|a_2| = 2$. Hence by Bieberbach's theorem, f must be a rotation of the Koebe function. □

The distortion theorem will now be applied to obtain the sharp upper and lower bounds for $|f(z)|$. This result is as follows.

Theorem 20.3.3. (Growth Theorem). For each $f \in S$,

$$\frac{r}{(1+r)^2} \leq |f(z)| \leq \frac{r}{(1-r)^2}, \qquad |z| = r < 1. \tag{20.3.6}$$

For each $z \in \mathbb{D}$, $z \neq 0$, equality occurs if and only if f is a suitable rotation of the Koebe function.

Proof. Let $f \in S$ and fix $z = \mathrm{e}^{i\theta}$ with $0 < r < 1$. Observe that

$$f(z) = \int_0^r f'(\rho\,\mathrm{e}^{i\theta})\,\mathrm{e}^{i\theta}\,d\rho,$$

since $f(0) = 0$. Thus by the distortion theorem,

$$|f(z)| \leq \int_0^r |f'(\rho\,\mathrm{e}^{i\theta})| d\rho \leq \int_0^r \frac{1+\rho}{(1-\rho)^3} d\rho = \frac{r}{(1-r)^2}.$$

The lower estimate is more subtle. It holds trivially if $|f(z)| \geq \frac{1}{4}$, since $r(1+r)^{-2} < \frac{1}{4}$ for $0 < r < 1$. If $|f(z)| < \frac{1}{4}$, the Koebe one-quarter theorem implies that the radial segment from 0 to $f(z)$ lies entirely in the range of f. Let C be the preimage of this segment. Then C is a simple arc from 0 to z, and

$$f(z) = \int_C f'(\zeta) d\zeta$$

But $f'(\zeta)d\zeta$ has constant signum along C, by construction, so the distortion theorem gives

$$|f(z)| = \int_C |f'(\zeta)||d\zeta| \geq \int_0^r \frac{1-\rho}{(1+\rho)^3} d\rho = \frac{r}{(1+r)^2}.$$

Equality in either part of (20.3.6) implies equality in the corresponding part of (20.3.3), which implies that f is a rotation of the Koebe function.

All of this information was obtained by passing to the real part in the basic inequality (20.3.1). Taking the imaginary part instead, one finds

$$-\frac{4r}{1-r^2} \leq Im\{\frac{zf''(z)}{f'(z)}\} \leq \frac{4r}{1-r^2}$$

$$-\frac{4r}{1-r^2} \leq \frac{\partial}{\partial r} \arg f'(r\,\mathrm{e}^{i\theta}) \leq \frac{4r}{1-r^2}$$

Radial integration now produces the inequality

$$|\arg f'(z)| \leq 2\log\frac{1+r}{1-r}, \qquad f \in S \tag{20.3.7}$$

Here it is understood that $\arg f'(z)$ is the branch which vanishes at the origin. The quantity $\arg f'(z)$ can be interpreted geometrically as the local rotation factor under the conformal mapping f. For this reason the inequality (20.3.7) may be called a *rotation theorem*. Unfortunately, however, it is not sharp at any point $z \neq 0$ in the disk. The true rotation theorem

$$\begin{aligned} |\arg f'(z)| &\leq 4\sin^{-1} r, \quad r \leq 1/\sqrt{2} \\ &\leq \pi + \log\frac{r^2}{1-r^2}, \quad r \geq 1/\sqrt{2}, \end{aligned}$$

lies much deeper. The splitting of the sharp bound at $r = 1/\sqrt{2}$ is one of the most remarkable phenomena in the univalent function theory. □

One further inequality, a combined growth and distortion theorem, is sometimes useful.

Few Probable Questions

1. State and prove the Area theorem.
2. State and prove Bieberbach's theorem.
3. State and Koebe One-Quarter theorem.
4. State and prove Distortion theorem.
5. State and prove Growth theorem.

References

1. E. T. Copson : An Introduction to the Theory of Functions of a Complex Variable.
2. E. C. Titthmarsh : The Theory of Functions.
3. A. I. Markushevich : Theory of Functions of a Complex Variable (Vol. I, II & III).
4. L. V. Ahlfors : Complex Analysis.
5. J. B. Conway : Functions of One Complex Variable.
6. A. I. Markushevich : The Theory of Analytic Functions, A Brief Course.
7. G. Valiron : Integral Functions.
8. C. Caratheodory : Theory of Functions of a Complex Variable.
9. R. P. Boas : Entire Functions.