

Contents

1	Review: Complex numbers and functions of a complex variable	6
2	Convergent power series	7
2.1	Series	7
2.2	Power series	8
3	Continuity and differentiability	8
3.1	Differentiability of power series	9
3.2	Some basic functions	11
3.3	Operations with power series	13
3.4	The Cauchy–Riemann equations	13
3.5	Analyticity at infinity	14
4	Integrals	14
5	Cauchy’s formula	17
5.1	Homotopic curves	17
5.2	Independence of the integral on the path	18
5.3	Cauchy’s Formula	18
6	Taylor series of analytic functions	20
7	The Fundamental Theorem of Algebra	22
8	More properties of analytic functions	22
9	Harmonic functions	24
9.1	Potential and Hamiltonian flows	25
10	The maximum modulus principle	26
10.1	Application: soap films	27
10.2	Cauchy principal value integrals (PV)	28
10.3	Bi-analytic bijections (automorphisms) of $\mathbb{D}$	29
11	Poisson’s formula	29
11.1	The Dirichlet problem for the Laplacian in $\mathbb{D}$	31
12	Isolated singularities, Laurent series	31
13	Laurent series and Fourier series	33
14	Calculating the Taylor series of simple functions	34
15	Residues and integrals	34
16	Calculating the Taylor series of simple functions	36
17	Residues and integrals	36

18 Integrals of trigonometric functions	38
19 Counting zeros and poles	38
19.1 Constancy of the number of roots of $f(s) - \zeta$	39
19.2 Some further properties of the logarithm	40
19.3 The argument principle	40
19.4 Hurwitz's theorem	41
19.5 Rouché's Theorem	41
20 The inverse function theorem	42
21 Analytic continuation	44
22 The Schwarz reflection principle	46
23 Multi-valued functions	48
23.1 Generalization: log of a function	49
23.2 General powers of z	50
24 Conformal Mapping	50
24.1 Uniqueness	51
24.2 Preservation of angles and small shapes: heuristics	52
24.3 Preservation of angles	52
24.4 Rescaling of arc length	53
24.5 Transformation of areas	53
24.6 Puiseux series	54
25 Evaluation of definite integrals	56
25.1 Fourier-type integrals	57
25.2 Further examples	58
26 Integrals of branched functions	60
26.0.1 Conformally Equivalent Domains	61
26.1 Automorphisms	62
26.2 Automorphisms of the plane	62
26.3 The Riemann sphere	63
26.4 Linear fractional transformations (LFTs): a first view	63
26.5 Automorphisms of the Riemann sphere	64
27 Linear fractional transformations (Möbius transformations)	65
27.1 Finding specific LFTs	66
27.1.1 Cross ratio	66
27.2 Mappings of regions	66
27.3 The Schwarz lemma	67
27.4 Automorphisms of the unit disk	68
27.5 Automorphisms of the unit disk	68
28 Some special biholomorphic transformations	69

29	The Riemann Mapping Theorem	76
29.1	Equicontinuity and Normal Families	79
29.1.1	The Ascoli–Arzelà Theorem	80
29.2	The Riemann Mapping Theorem	81
29.3	Behavior at the boundary of biholomorphisms: a general but weaker result	84
30	Boundary behavior.	84
30.1	A reflection principle for harmonic functions	85
30.2	Analytic arcs	86
30.3	An extension of the Schwarz reflection principle	86
30.3.1	Discussion: Tilings with Parallelograms	88
30.4	Behavior at the Boundary: A Stronger Result	89
30.4.1	A Negative Result	89
31	Conformal Mappings of Polygons and the Schwarz–Christoffel Formula	91
31.1	Heuristics	92
31.2	Conformal map of the UHP onto the interior of a triangle of angles $\pi\alpha, \pi\beta, \pi\gamma$. . .	97
31.3	Schwarz triangle functions and hypergeometric functions	98
32	Optional material: curvilinear triangles	99
33	Asymptotics of Integrals and of Taylor Coefficients of Analytic Functions	103
34	Asymptotic Power Series	105
35	The Phragmén–Lindelöf Theorem	107
35.1	Laplace and Inverse Laplace Transforms	108
35.2	A Laplace Inversion Formula	109
35.3	Abstract Stokes Phenomena	111
35.4	Watson’s Lemma	112
35.5	Applications of Watson’s Lemma	113
36	Riemann surfaces, an introduction	117
36.1	Universal covers.	120
36.1.1	Basic examples	121
36.2	Properly discontinuous groups	122
36.3	The uniformization theorem: simply connected case	125
36.4	General form of the uniformization theorem	125
36.5	Appendix: Deck Transformations of Riemann surfaces are holomorphic	126
37	The modular group	126
37.1	Lattices and tessellations	127
37.2	Example: mapping rectangles	127
37.3	Continuation to $\mathbb{C}$ and double periodicity	129
37.4	Jacobi elliptic functions sn , cn , dn	129
37.5	Bases of lattices	130
37.6	The fundamental region of $PSL(2, \mathbb{Z})$	130
37.7	The generators of the modular group	133
37.8	The hyperbolic plane	134

38 Elliptic functions	135
38.1 The period module	135
38.2 General properties of elliptic functions	135
38.3 The Weierstrass elliptic functions	137
38.3.1 The Weierstrass zeta function	138
38.3.2 The differential equation of $\wp$	138
38.4 The modular function λ	139
38.5 The action of the $\text{PSL}(2, \mathbb{Z})$ on λ	140
38.6 The conformal mapping of λ	141
38.7 An example: M , the universal cover of $\hat{\mathbb{C}} \setminus \{0, 1, \infty\}$	143
39 Padé Approximants	144
39.1 Motivation	145
39.2 Definition of the Padé approximant	145
39.3 First examples	145
39.4 Padé Approximants and C-Fractions	146
39.5 Analytic continuation	147
39.6 Error and convergence	147
39.7 Stahl's Theory and the Role of the Conformal Map	148
39.7.1 List of mfs [?]	153
39.7.2 Transformation theory	154
39.8 Continuation Through the Natural Boundary	162
39.8.1 General Solution of Mock Modular Equations for $ q < 1$	162
40 Applications of Mock Theta functions	163
40.1 Schwarz reflection of domains about a circle	163
41 Entire and Meromorphic functions	164
41.1 A historical context	164
41.2 Partial fraction decompositions	165
41.3 The Mittag-Leffler theorem	166
41.4 Further examples	168
42 Infinite products	168
42.1 Uniform convergence of products	170
42.2 Example: the sin function	170
42.3 Canonical products	171
42.4 Counting zeros of analytic functions. Jensen's formula	172
42.5 Entire functions of finite order	173
42.6 Estimating analytic functions by their real part	175
43 Hadamard's theorem	176
43.1 Canonical products	177
44 The minimum modulus theorem; end of proof of Theorem 280	177
44.1 Some applications	180
45 The little Picard theorem	181

46 Riemann-Hilbert problems: an introduction	182
46.1 A simple Riemann-Hilbert problem	183
47 Cauchy type integrals	183
47.1 Asymptotic behavior of $\Phi(z)$ for large z	184
47.2 Regularity and singularities	184
47.3 Examples	186
47.3.1 A very simple example	186
47.3.2 Another simple example	186
47.3.3 A solution	187
47.3.4 Calculating principal value integrals	187
47.3.5 Uniqueness issues	187
47.3.6 Degree of a function at infinity	188
48 More general scalar R-H problems	188
48.1 Scalar homogeneous R-H problems	188
48.1.1 Scalar homogeneous R-H problems	188
48.1.2 Inhomogeneous R-H problems	189
48.2 Applications	189
48.2.1 Inhomogeneous singular integral equations	189
49 Asymptotic series	190
49.1 More general asymptotic series	192
49.2 Asymptotic power series	193
49.3 Integration and differentiation of asymptotic power series.	195
49.4 Lebesgue's Dominated Convergence Theorem	197
49.5 Fubini's theorem	197
49.6 Appendix to Chapter 8	198
49.7 Some facts about the topology of $\mathbb{C}$	198
49.7.1 Connected sets	199
50 Soap films and the minimal surface equation	200

It is hard to overstate the importance of complex analysis in virtually every area of mathematics. Its power stems from several sources; one of them is that it represents the “ultimate” extension of the real numbers: it satisfies all the axioms of $\mathbb{R}$ except for the ordering property. Furthermore, the Gelfand–Mazur theorem shows that the only normed division algebras over $\mathbb{R}$ are $\mathbb{R}$, $\mathbb{C}$, and the quaternions (the latter being non-commutative).

Functions that extend from $\mathbb{R}$ to $\mathbb{C}$ —the analytic functions—preserve all their local properties, by the so-called *principle of permanence of relations* (to be discussed later). Hence, to understand such functions, it is essential to consider their full complex domain of definition.

A simple example is the function

$$\frac{1}{1+x^2}.$$

It admits a convergent power series at $x = 0$ of radius 1. There is nothing remarkable about the points $x = \pm 1$ on the real axis, but as soon as we extend the function to $\mathbb{C}$, we see that $\pm i$ are singular points, being the roots of the denominator.

More generally, the fundamental theorem of algebra holds in $\mathbb{C}$ but not in $\mathbb{R}$.

As another illustration, the exponential decay of Fourier transforms of real-analytic functions cannot be shown by purely real methods; one “needs” complex-analytic arguments.

1 Review: Complex numbers and functions of a complex variable

- Complex numbers $\mathbb{C}$ form a field. Addition and multiplication of complex numbers obey the same rules as their real counterparts.
- There is no “good” order relation on $\mathbb{C}$. Aside from this, we operate with complex numbers in essentially the same way as with real numbers.
- A function f of a complex variable is a function defined on some subset of $\mathbb{C}$ with values in $\mathbb{C}$. Equivalently, we can view it as a pair of real-valued functions of two real variables. Writing $z = x + iy$ with $x, y \in \mathbb{R}$ and $i^2 = -1$, we have

$$f(z) = f(x + iy) = u(x, y) + iv(x, y).$$

- Note that $i^2 = (-i)^2 = -1$. There is no intrinsic distinction between i and $-i$. This entails a fundamental symmetry of the theory: invariance under complex conjugation.¹
- Using only the algebraic properties of complex numbers we can already define several elementary functions: z , $1/z$, and more generally, for $m \in \mathbb{Z}$, the power z^m , as well as any polynomial

$$\sum_{m=0}^K c_m (z - z_0)^m.$$

- To define and study more general functions we must introduce continuity, derivatives, and limits. Writing $z = x + iy$, the modulus

$$|z| = \sqrt{x^2 + y^2}$$

¹More precisely, the map $z = x + iy \mapsto \bar{z} = x - iy$ is an involution and a field automorphism of $\mathbb{C}$.

provides a natural notion of length (or smallness). It induces a norm under which $\mathbb{C}$ becomes a **complete metric space**. Convergence then reduces to convergence of real numbers:

$$z_n \rightarrow z \iff |z - z_n| \rightarrow 0 \text{ as } n \rightarrow \infty. \quad (1.2)$$

- The *topology* of $\mathbb{C}$ coincides with that of $\mathbb{R}^2$, under the identification $z = x + iy \leftrightarrow (x, y) \in \mathbb{R}^2$. Some basic topological facts are reviewed in Appendix §49.6.
- In what follows, a *domain* in $\mathbb{C}$ means an open connected subset of $\mathbb{C}$. Examples include disks of radius $r > 0$ centered at $z_0 \in \mathbb{C}$:

$$\mathbb{D}(z_0, r) := \{z \in \mathbb{C} : |z - z_0| < r\}. \quad (1.3)$$

The special cases $r = 0$ (the empty set, $\emptyset$) and $r = \infty$ (the whole of $\mathbb{C}$) are also open sets. The unit disk

$$\mathbb{D} := \mathbb{D}(0, 1)$$

will play a special role as a canonical choice of a disk.

Exercise 1. Show that $z_n \rightarrow z$ if and only if $\operatorname{Re}(z_n) \rightarrow \operatorname{Re}(z)$ and $\operatorname{Im}(z_n) \rightarrow \operatorname{Im}(z)$. Using the completeness of $\mathbb{R}$, show that $\mathbb{C}$ is a complete normed space.

Definition 2

For functions, limits reduce to the real case: we write

$$\lim_{z \rightarrow z_0} f(z) = a \iff |f(z) - a| \rightarrow 0 \text{ as } z \rightarrow z_0.$$

2 Convergent power series

2.1 Series

A complex series is an infinite sum

$$\sum_{k=0}^{\infty} a_k, \quad (2.2)$$

where $a_k \in \mathbb{C}$ for $k \geq 0$. It is said to *converge* if the sequence of *partial sums*

$$S_N := \sum_{k=0}^N a_k \quad (2.3)$$

converges as $N \rightarrow \infty$.

The series is said to *converge absolutely* if the real series

$$\sum_{k=0}^{\infty} |a_k| \quad (2.4)$$

converges.

Exercise 3. (a) Show that a necessary condition for convergence is $a_k \rightarrow 0$ as $k \rightarrow \infty$.

(b) Show that absolute convergence implies convergence.

(c) Verify that the standard convergence tests from real analysis (ratio test, n -th root test, etc. which do not involve ordering) apply equally to complex series.

2.2 Power series

A *power series* centered at $z_0 \in \mathbb{C}$ is a series of the form

$$\sum_{k=0}^{\infty} c_k (z - z_0)^k, \quad (2.5)$$

where $c_k, z, z_0 \in \mathbb{C}$.

Theorem 4: Abel

If for some $z_1 \neq z_0$ the series

$$\sum_{k=0}^{\infty} c_k (z_1 - z_0)^k \quad (2.6)$$

converges, then

$$\sum_{k=0}^{\infty} c_k (z - z_0)^k \quad (2.7)$$

converges absolutely and uniformly in any disk $\mathbb{D}(z_0, r)$ with $r < |z_1 - z_0|$.

Exercise 5. Prove Theorem 4 by reducing it to the corresponding fact about real series and using the completeness of $\mathbb{C}$.

Abel's theorem shows that the domain of convergence of a power series is always a disk; convergence may also extend to parts of the boundary. The largest r such that the series (2.7) converges for all $z \in \mathbb{D}(z_0, r)$ is called the *radius of convergence*.

The disk of convergence can be degenerate: in one extreme case, it reduces to the point $z = z_0$ (radius of convergence 0), while in the other extreme it covers all of $\mathbb{C}$ ("infinite radius of convergence").

3 Continuity and differentiability

Definition 6

A complex function f is said to be *continuous* at z_0 if

$$f(z) \rightarrow f(z_0) \quad \text{as } z \rightarrow z_0.$$

Exercise 7. Show that polynomials are continuous in $\mathbb{C}$.

We can now define differentiability.

Definition 8

A function f is *differentiable* at z_0 if there exists a number $f'(z_0)$ such that

$$\frac{f(z) - f(z_0)}{z - z_0} \rightarrow f'(z_0) \quad \text{as } z \rightarrow z_0.$$

Exercise 9. Show that differentiation in $\mathbb{C}$ satisfies the familiar rules from real analysis. In particular, the sum rule, product rule, and chain rule all hold. (Prove this by mimicking the real-variable arguments.)

Also prove that differentiability implies continuity.

Differentiability in $\mathbb{C}$ is much more restrictive than in $\mathbb{R}$. For the same reason, complex differentiable functions enjoy far stronger regularity properties than merely real-differentiable functions.

Exercise 10. Show that the function $f(z) = |z|$ is **nowhere differentiable** in $\mathbb{C}$.

Definition 11

Let $\mathcal{D}$ be a domain in $\mathbb{C}$. A function f is said to be *analytic* in $\mathcal{D}$ if it is differentiable at every point $z \in \mathcal{D}$.

Functions analytic in a domain are also called *holomorphic*. The two terms are synonymous in this course.

If A is a set that contains a domain $\mathcal{D}$ together with parts of its boundary, we say that f is *analytic in A* if there exists a domain $\mathcal{D}' \supset A$ such that f is analytic in $\mathcal{D}'$.

We will see that if f is analytic in some domain $\mathcal{D}$, then f' is also analytic there. In particular, f has continuous derivatives of all orders in $\mathcal{D}$. This stems from the crucial fact that in complex analysis the derivative has an integral representation. We will also later establish that analyticity in a domain $\mathcal{D}$ is equivalent to the convergence of the Taylor series of f at every $z_0 \in \mathcal{D}$.

Clearly, in $\mathbb{R}$ differentiability does not imply C^∞ (infinite differentiability).

3.1 Differentiability of power series

Theorem 12

Let

$$S(z) = \sum_{k=0}^{\infty} c_k (z - z_0)^k \quad (3.2)$$

be a power series convergent in the open disk $\mathbb{D}(z_0, r)$ with $r > 0$. Then S has derivatives of all orders in $\mathbb{D}(z_0, r)$. In particular,

$$S'(z) = \sum_{k=0}^{\infty} k c_k (z - z_0)^{k-1}, \quad (3.3)$$

$$S''(z) = \sum_{k=0}^{\infty} k(k-1) c_k (z - z_0)^{k-2}, \quad (3.4)$$

$\vdots$

$$S^{(p)}(z) = \sum_{k=0}^{\infty} k(k-1) \cdots (k-p+1) c_k (z - z_0)^{k-p}, \quad (3.5)$$

$\vdots$

and each of these series converges in $\mathbb{D}(z_0, r)$ to the corresponding derivative of S .

Remark 13. The derivative of the sum of a convergent power series S is obtained by differentiating each term individually. In other words, S can be differentiated term by term.

Proof. It suffices to prove the case $p = 1$, since higher derivatives then follow by induction. By translation, assume without loss of generality that $z_0 = 0$.

Fix $|z| < \rho < r$, and choose h small enough so that $|z| + |h| < \rho$. Consider the partial sums

$$S_N(z) = \sum_{k=0}^N c_k z^k, \quad \frac{S_N(z+h) - S_N(z)}{h}.$$

For **real, positive** z, h , the binomial formula and Taylor's theorem with remainder give

$$\begin{aligned} (z+h)^n - z^n - nz^{n-1}h &= \frac{n(n-1)}{2}z^{n-2}h^2 + \sum_{k=3}^n \binom{n}{k} z^{n-k}h^k \\ &\leq \frac{n(n-1)}{2}(z+a)^{n-2}h^2 \leq \frac{n(n-1)}{2}\rho^{n-2}h^2, \end{aligned} \quad (3.6)$$

for some $a \in (0, h)$. The **equality** remains valid for complex z, h . Thus,

$$\begin{aligned} \left| \frac{(z+h)^n - z^n}{h} - nz^{n-1} \right| &= \left| \frac{n(n-1)}{2}z^{n-2}h + \sum_{k=3}^n \binom{n}{k} z^{n-k}h^k \right| \\ &\leq \frac{n(n-1)}{2}|z|^{n-2}|h| + \sum_{k=3}^n \binom{n}{k}|z|^{n-k}|h|^k \\ &\leq \frac{n(n-1)}{2}\rho^{n-2}|h|. \end{aligned}$$

Therefore,

$$\left| \frac{S_N(z+h) - S_N(z)}{h} - S'_N(z) \right| \leq |h| \sum_{k=0}^N \frac{k(k-1)}{2} |c_k| \rho^k.$$

The series

$$C = \sum_{k=0}^{\infty} \frac{k(k-1)}{2} |c_k| \rho^k$$

converges because $\rho < r$. Indeed, since the original power series $\sum c_k z^k$ has radius of convergence r , the series $\sum |c_k| \rho^k$ converges absolutely for every $\rho < r$. Multiplying the coefficients by the polynomial factor $k(k-1)/2$ does not affect convergence (apply, for instance, the n th root test).

Hence the right-hand side is bounded by $C|h|$, which tends to 0 as $h \rightarrow 0$.

Thus $\lim_{h \rightarrow 0} \frac{S(z+h) - S(z)}{h}$ exists and equals $S'(z) = \sum_{k=0}^{\infty} k c_k z^{k-1}$, as claimed. ■

Exercise 14. Show that

$$\frac{(z+h)^n - z^n}{hn} = \int_0^1 (z+th)^{n-1} dt.$$

Use the dominated convergence theorem, the triangle inequality, and the fact that $\lim_{n \rightarrow \infty} n^{1/n} = 1$ to derive a shorter proof of Theorem 12.

Corollary 15. Assume S is a power series convergent in the disk $\mathbb{D}(z_0, r)$ with $r > 0$. Then S is analytic (holomorphic) in $\mathbb{D}(z_0, r)$.

Corollary 16. For the series (3.2), we have

$$S^{(k)}(z_0) = k! c_k.$$

Hence (3.2) is precisely the Taylor series of S centered at z_0 .

Corollary 17. Suppose $S(z)$ converges in a disk $\mathbb{D}(z_0, r)$, and there exists a sequence $\{z_n\}_{n \in \mathbb{N}}$ with an accumulation point at z_0 such that $S(z_n) = 0$ for all n . Then $S(z) \equiv 0$.

Remark 18. Corollary 17 is a crucial ingredient in the important principle of permanence of relations.

Proof. Without loss of generality, assume $z_0 = 0$. Write

$$S(z) = c_0 + zT(z),$$

where $T(z)$ is a power series convergent in $\mathbb{D}(0, r)$.

By assumption,

$$0 = S(z_n) = c_0 + z_n T(z_n), \quad n \in \mathbb{N}.$$

Letting $n \rightarrow \infty$ shows that $c_0 = 0$.

Now divide by z to obtain $S(z)/z = T(z) = c_1 + zT_1(z)$, where $T_1(z)$ is again a power series convergent in $\mathbb{D}(0, r)$. Since $T(z_n) = 0$ for infinitely many $z_n \rightarrow 0$, the same reasoning shows $c_1 = 0$.

Repeating the argument inductively, we conclude that all coefficients c_k vanish. Hence $S \equiv 0$. ■

3.2 Some basic functions

◦ *The exponential.* We define

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}. \quad (3.7)$$

This series converges for every $z \in \mathbb{C}$, hence e^z is analytic in $\mathbb{C}$ by Theorem 12. Differentiating term by term (see (3.5)) yields

$$(e^z)' = e^z. \quad (3.8)$$

In particular,

$$(e^z e^{-z})' = 0, \quad (3.9)$$

so $e^z e^{-z}$ is constant throughout $\mathbb{C}$. Evaluating at $z = 0$ shows $e^0 = 1$, hence

$$e^z e^{-z} = 1 \quad \Leftrightarrow \quad e^{-z} = \frac{1}{e^z}. \quad (3.10)$$

Similarly, for any $a \in \mathbb{C}$ we compute

$$\begin{aligned} (e^{z+a} e^{-z})' &= (e^{z+a})' e^{-z} + e^{z+a} (e^{-z})' = 0 \\ \implies e^{z+a} e^{-z} &= e^a e^0 = e^a \implies e^{z+a} = e^z e^a, \end{aligned} \quad (3.11)$$

which is the fundamental exponential identity.

For real $\varphi \in \mathbb{R}$, one immediately checks Euler's formula:

$$e^{i\varphi} = \cos \varphi + i \sin \varphi. \quad (3.12)$$

Remark 19. Euler's formula shows that the mapping $\varphi \mapsto e^{i\varphi}$ parametrizes the unit circle in the complex plane. As φ varies over $\mathbb{R}$, the point $e^{i\varphi}$ traces the circle $|z| = 1$ counterclockwise, starting at 1 when $\varphi = 0$. Thus the exponential provides a natural bridge between trigonometry and complex geometry.

Exercise 20. (a) In the above arguments we tacitly used an additional fact. Identify it and fill in the missing details using results already established.

(b) Define $\cos$ and $\sin$ in $\mathbb{C}$ by their power series and prove that (3.12) extends to all $z \in \mathbb{C}$:

$$e^{iz} = \cos z + i \sin z.$$

Deduce the important exponential representations

$$\sin z = \frac{e^{iz} - e^{-iz}}{2i}, \quad \cos z = \frac{e^{iz} + e^{-iz}}{2}.$$

Use these to show that $\sin^2 z + \cos^2 z = 1$ for all $z \in \mathbb{C}$.

Exercise 21 (A first form of permanence of analytic relations). (a) Let f be analytic in a domain $\mathcal{D}$, and let F be analytic in a domain $\mathcal{D}'$ with $\mathcal{D}' \supset f(\mathcal{D})$. Suppose $\mathcal{D}$ contains a real interval and that

$$F(f(x)) = 0 \quad \text{for all real } x \text{ in that interval.}$$

Show that $F(f(z)) = 0$ for all $z \in \mathcal{D}$.

(b) Use part (a) to give an alternative proof of the fundamental exponential identities, relying on Corollary 17.

(c) Use part (a), together with the power series extensions of $\sin$ and $\cos$, to prove that

$$\sin^2 z + \cos^2 z = 1 \quad \text{for all } z \in \mathbb{C}.$$

Exercise 22. Show that

$$e^s = 1 \Leftrightarrow s = 2N\pi i, \quad N \in \mathbb{Z} \quad (3.13)$$

◦ *The logarithm.* In the complex domain, the logarithm is a more delicate function. For the moment, we consider the simpler case of defining $\log(1+z)$ for $|z| < 1$. This is done via the convergent Taylor series

$$\log(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \cdots. \quad (3.14)$$

By Theorem 12, term-by-term differentiation yields

$$\frac{d}{dz} \log(1+z) = 1 - z + z^2 - z^3 + \cdots = \frac{1}{1+z}, \quad |z| < 1. \quad (3.15)$$

Exercise 23. Show that if $|s|$ is small then

$$\log(e^s) = s, \quad e^{\log(1+s)} = 1 + s.$$

We will return later to the more general problem of defining $\log z$ for arbitrary $z \neq 0$. Its natural maximal domain of definition is not a subset of $\mathbb{C}$ but rather a *Riemann surface*, which we will study in detail later. The logarithm is one of the fundamental examples of a *branched function*, and in fact many other multivalued functions in complex analysis inherit their branching behavior from it.

3.3 Operations with power series

If S and T are power series convergent in a neighborhood of $z_0 = 0$, then the following combinations are also convergent in some (possibly smaller) neighborhood of 0:

$$S + T, \quad S \cdot T, \quad \frac{S}{T} \text{ (provided } T(0) \neq 0), \quad S \circ T \text{ (provided } T(0) = 0).$$

If S and T are power series convergent in a neighborhood of $z_0 = 0$, then the combinations

$$S + T, \quad S \cdot T, \quad \frac{S}{T} \text{ (if } T(0) \neq 0), \quad S \circ T \text{ (if } T(0) = 0)$$

are also convergent in some neighborhood of 0. Formulas for the coefficients of these new series are obtained by observing that power series are uniform limits of polynomials and that the operations agree with those for polynomials.

For example,

$$ST = s_0t_0 + (s_1t_0 + s_0t_1)z + (s_2t_0 + s_1t_1 + s_0t_2)z^2 + \cdots \quad (3.16)$$

Exercise 24. (a) If S and T are two power series with radius of convergence r , show that ST has radius of convergence at least r .

(b) Write three terms of the series S/T if $T(0) \neq 0$.

(c)* Under the assumptions above, show that S/T has nonzero radius of convergence.

(d)* Replace the assumption in (b) by $T(0) = 0$. Show that in this case $S \circ T$ has nonzero radius of convergence.

Exercise 25. Formulate these statements for points other than $z_0 = 0$.

3.4 The Cauchy–Riemann equations

Analytic functions admit several equivalent characterizations. So far, we have defined them by requiring differentiability at every point of a domain (Definition 8). An alternative, equally fundamental characterization is given by the Cauchy–Riemann equations:

Theorem 26: Cauchy–Riemann equations

(1) If $f = u + iv$ is analytic in a domain $\mathcal{D} \subset \mathbb{C}$, then the partial derivatives exist and satisfy

$$\begin{aligned} u_x &= v_y, \\ u_y &= -v_x, \end{aligned} \quad (3.17)$$

throughout $\mathcal{D}$.

(2) Conversely, if u, v are differentiable functions in $\mathcal{D}$ that satisfy (3.17), then $f = u + iv$ is analytic in $\mathcal{D}$.

Proof. (1) Let $f(z) = u(x, y) + iv(x, y)$ and suppose $f'(z_0) = a + ib$. By translation we may assume $z_0 = 0$. Then

$$\begin{aligned} f(z) - f(0) &= u(x, y) - u(0, 0) + i(v(x, y) - v(0, 0)) \\ &= [f'(0) + \varepsilon(z)]z = (a + ib)(x + iy) + \varepsilon(z)z, \end{aligned} \quad (3.18)$$

where $\varepsilon(z) \rightarrow 0$ as $z \rightarrow 0$. This implies the existence of u_x, u_y, v_x, v_y at $(0,0)$ and the validity of the Cauchy–Riemann equations there.

(2) Conversely, if u, v are differentiable at $(0,0)$, then

$$\begin{aligned} f(z) - f(0) &= u(x, y) - u(0, 0) + i(v(x, y) - v(0, 0)) \\ &= u_x(0, 0)x + u_y(0, 0)y + i(v_x(0, 0)x + v_y(0, 0)y) + \varepsilon(x, y)x + i\eta(x, y)y, \end{aligned} \quad (3.19)$$

where $\varepsilon, \eta \rightarrow 0$ as $z \rightarrow 0$. Comparing (3.18) and (3.19) shows that f is differentiable iff (3.17) holds. ■

Exercise 27. Show that (3.18) and (3.19) are compatible if and only if (3.17) holds. (Hint: equate the real and imaginary parts and note that x and y are independent variables.)

3.5 Analyticity at infinity

As $|z| \rightarrow \infty$, we have $1/z \rightarrow 0$. By definition, a function f is said to be *analytic at infinity* if the function $f(1/z)$ is analytic at $z = 0$.

4 Integrals

Integration plays an important role in complex analysis. As we shall see, the derivative of a function can be written as an integral, and many of the nice properties of analytic functions originate in this fact.

If $f(t) = u(t) + iv(t)$ is a complex-valued function of one real variable t , then

$$\int_a^b f(t) dt = \int_a^b u(t) dt + i \int_a^b v(t) dt. \quad (4.2)$$

Thus questions of complex integration reduce to familiar real integration.

Note 28

In the following, unless otherwise specified, all curves are assumed to be piecewise differentiable.

Let $\gamma(t) = x(t) + iy(t)$, $t \in [a, b]$, be a piecewise C^1 parametrized curve. We define

$$\int_{\gamma} f(z) dz := \int_a^b f(\gamma(t)) \gamma'(t) dt,$$

in the same way as in (4.2), by breaking γ , γ' , and f into their real and imaginary parts. (Do this calculation explicitly.)

The value of the integral depends on the *orientation* of γ , specified by stating that t increases from a to b . It is natural to say that $\gamma(a)$ is the *starting point* and $\gamma(b)$ the *endpoint*.

The same geometric curve with the opposite orientation is denoted $-\gamma$. A convenient formula is

$$(-\gamma)(t) = x(a + b - t) + iy(a + b - t), \quad t \in [a, b].$$

It follows that

$$\int_{-\gamma} f(z) dz = \int_b^a f(\gamma(t)) \gamma'(t) dt = - \int_{\gamma} f(z) dz. \quad (4.3)$$

Positively oriented γ (counterclockwise)

Reversed orientation $-\gamma$ (clockwise)

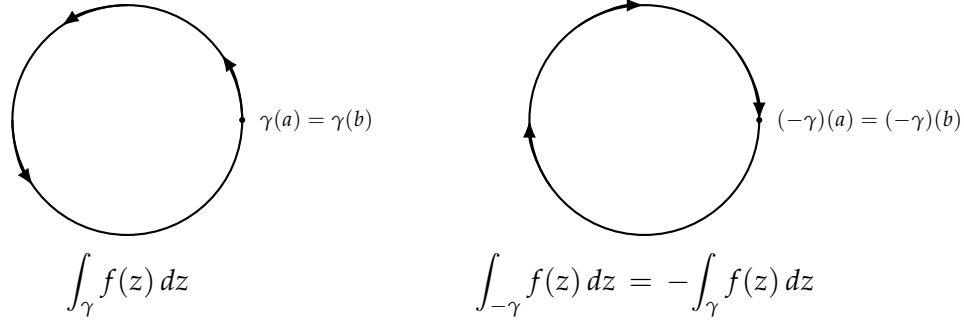


Figure 1: Orientation of a closed curve: γ counterclockwise (left) and $-\gamma$ clockwise (right).

If $\gamma(a) = \gamma(b)$ the curve is called *closed*. Unless otherwise specified, positive orientation (counterclockwise) is assumed, in which case we write

$$\oint_{\gamma} := \int_{\gamma}.$$

Exercise 29. Show that

$$\int_{\gamma} f(z) dz = \int_{\gamma} (u dx - v dy) + i \int_{\gamma} (u dy + v dx). \quad (4.4)$$

Exercise 30. Show that the integral along a given geometric curve depends on its parametrization only through the orientation (i.e. reparametrizations preserving orientation leave the integral unchanged, while reversing orientation multiplies it by -1).

A curve is called *simple* if it has no self-intersections. For example, the circle is simple, but the figure “8” is not.

A domain $\mathcal{D}$ is called *simply connected* if any simple closed curve γ contained in $\mathcal{D}$ can be deformed continuously to a point through curves completely contained in $\mathcal{D}$. Formally, this means:

There exists a continuous function of two variables

$$F(t, s) = x(t, s) + iy(t, s), \quad (t, s) \in [a, b] \times [0, 1],$$

with values in $\mathcal{D}$ such that $F(t, 0) = \gamma(t)$ and $F(t, 1) = p$ for some point $p \in \mathcal{D}$.

Intuitively, a simply connected domain has no holes. For example, a disk is simply connected, but a punctured disk (such as $\mathbb{D} \setminus \{0\}$) or an *annulus*

$$\{z \in \mathbb{C} : r < |z| < R\}$$

are not simply connected.

Theorem 31: Cauchy

Assume $\mathcal{D}$ is a simply connected domain and that f is continuously differentiable in $\mathcal{D}$. If γ is a piecewise differentiable simple closed curve contained in $\mathcal{D}$, then

$$\oint_{\gamma} f(z) dz = 0. \quad (4.5)$$

Proof. Start from the decomposition (4.4) and apply Green's theorem:

$$\int_{\gamma} (u dx - v dy) = - \iint_{\text{Int}(\gamma)} \left(\frac{\partial v}{\partial x} + \frac{\partial u}{\partial y} \right) dx dy. \quad (4.6)$$

This vanishes by the Cauchy–Riemann equations (3.17). The second integral in (4.4) is handled in the same way. ■

Cauchy originally assumed continuity of f' was required. Later, Montel and others weakened this hypothesis, requiring only boundedness, L^1 , or other mild conditions. Finally, Goursat (1884) observed that mere differentiability suffices.

Theorem 32: Cauchy–Goursat

Assume $\mathcal{D}$ is a simply connected domain and that f is differentiable in $\mathcal{D}$. If γ is a piecewise differentiable simple closed curve contained in $\mathcal{D}$, then

$$\oint_{\gamma} f(z) dz = 0. \quad (4.7)$$

Proof. By continuity of f and piecewise differentiability of γ , the integral $\int_{\gamma} f(s) ds$ can be approximated with arbitrary accuracy by integrals along polygonal curves P , i.e. $\int_P f(s) ds$. Decomposing the polygon into triangles, it suffices to show that

$$\int_T f(s) ds = 0$$

for any triangle $T \subset \mathcal{D}$.

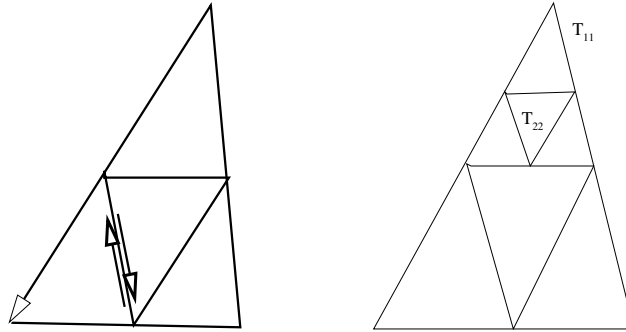


Figure 2: Orientation of integration after the first subdivision (left); further partitions (right).

Split T into four smaller triangles and write $\int_T = \sum_{j=1}^4 \int_{T_j}$, choosing orientations so that integrals along inner sides cancel. Let T_{11} denote the subtriangle such that

$$\left| \int_{T_{11}} f(s) ds \right| = \max_j \left\{ \left| \int_{T_j} f(s) ds \right| \right\}.$$

Then

$$\left| \int_T f(s) ds \right| \leq 4 \left| \int_{T_{11}} f(s) ds \right|.$$

Repeating inductively, we obtain nested closed triangles T_{nn} with

$$\left| \int_T f(s) ds \right| \leq 4^n \left| \int_{T_{nn}} f(s) ds \right|, \quad n \in \mathbb{N}.$$

Since $\{T_{nn}\}$ is a nested sequence of compact sets, their intersection reduces to a point $z_0 \in T$.

Now for any fixed n and $z \in \partial T_{nn}$, by differentiability, we have

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \varepsilon(z; z_0)(z - z_0),$$

with $\varepsilon(z)$ as $z \rightarrow z_0$. Since $\oint_{T_{nn}} f(z_0) dz = \oint_{T_{nn}} f'(z_0)(z - z_0) dz = 0$ (by Cauchy's theorem), it follows that

$$\int_{T_{nn}} f(s) ds = \int_{T_{nn}} \varepsilon(s; z_0)(s - z_0) ds.$$

Estimating the diameter of T_{nn} by $C2^{-n}$ for some constant C (why can we choose C independent of n ?), we conclude

$$\left| \int_T f(s) ds \right| \leq C^2 4^n 2^{-n} 2^{-n} \sup_{T_{nn}} |\varepsilon(s, t)| \rightarrow 0 \quad (n \rightarrow \infty).$$

■

Exercise 33. Why is the supremum above tending to zero?

It is sometimes useful to integrate analytic functions along the boundary of their domain of analyticity. For instance, if f is continuous up to the boundary, Cauchy's theorem still holds:

Exercise 34. Assume that f is analytic in $\mathcal{D}$, continuous on $\overline{\mathcal{D}}$, and that $\partial\mathcal{D}$ is a rectifiable curve of winding number one. Show that (4.5) holds if γ is a simple closed curve in $\mathcal{D}$.

5 Cauchy's formula

5.1 Homotopic curves

Let $\mathcal{D}$ be a domain in $\mathbb{C}$. Two curves in $\mathcal{D}$ are said to be *homotopic in $\mathcal{D}$* if one can be continuously deformed into the other through curves lying entirely in $\mathcal{D}$ (cf. the discussion preceding Theorem 31).

For example, if $\mathcal{D}$ is simply connected then any simple closed curve is homotopic to a point (see Fig. 3). As another example, if $\mathcal{D}$ is the annulus

$$\{z \in \mathbb{C} : 1 < |z| < 2\},$$

then all circles $|z| = r$ with $1 < r < 2$ are homotopic to each other, but not to a point. On the other hand, any simple closed curve that does not wind around 0 is homotopic to a point.

It is also useful to consider curves γ_1, γ_2 in $\mathcal{D}$, given by functions $\gamma_1(t), \gamma_2(t)$ for $t \in [a, b]$, which have the same endpoints, $\gamma_1(a) = \gamma_2(a)$ and $\gamma_1(b) = \gamma_2(b)$. Two such curves are called *homotopic with fixed endpoints* if they can be continuously deformed into each other through a transformation that preserves the endpoints and remains entirely within $\mathcal{D}$ (see Fig. 4).

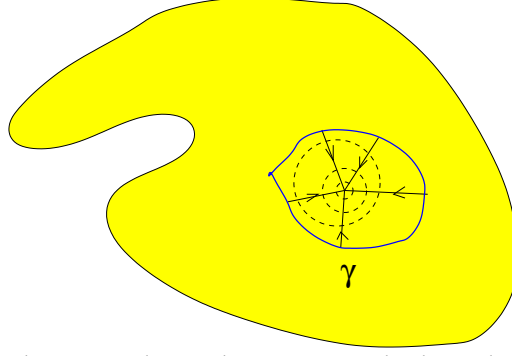


Figure 3: All dotted curves inside γ are homotopic to each other and to the central point.

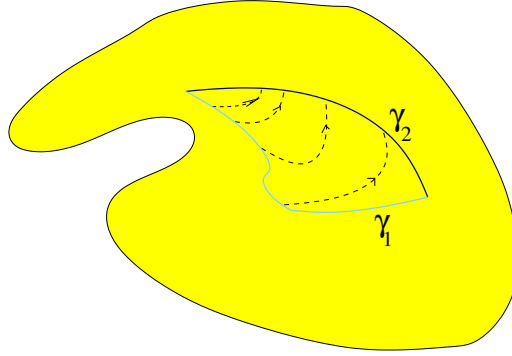


Figure 4: Two curves γ_1 and γ_2 homotopic with fixed endpoints.

5.2 Independence of the integral on the path

Line integrals are additive with respect to concatenation of curves. Let γ_1 be an oriented curve, and let γ_2 start at the final point of γ_1 . If $\gamma_1(t)$ is defined for $t \in [a, b]$ and $\gamma_2(t)$ for $t \in [b, c]$ with $\gamma_1(b) = \gamma_2(b)$, then the concatenated curve $\gamma_1 + \gamma_2$ runs from $t = a$ to $t = c$. By definition,

$$\int_{\gamma_1 + \gamma_2} f(z) dz = \int_{\gamma_1} f(z) dz + \int_{\gamma_2} f(z) dz. \quad (5.2)$$

(i) Suppose f is analytic in a simply connected domain $\mathcal{D}$, and let γ_1, γ_2 be two oriented curves in $\mathcal{D}$ with the same endpoints. Then $\gamma = \gamma_1 - \gamma_2$ is a closed curve. If there exists a subdomain $\mathcal{D}' \subset \mathcal{D}$ containing γ such that γ is homotopic to a point in $\mathcal{D}'$, then by (4.5), (5.2), and (4.3) we obtain

$$0 = \int_{\gamma_1 - \gamma_2} f(z) dz = \int_{\gamma_1} f(z) dz - \int_{\gamma_2} f(z) dz.$$

Therefore, the integral of an analytic function on a simply connected domain depends only on the endpoints of the path, provided the paths are homotopic:

$$\int_{\gamma_1} f(z) dz = \int_{\gamma_2} f(z) dz. \quad (5.3)$$

5.3 Cauchy's Formula

Let $\mathcal{D}$ be a domain in $\mathbb{C}$ and $z_0 \in \mathcal{D}$. The functions $(z - z_0)^{-n}$, $n = 1, 2, \dots$ are analytic in $\mathcal{D} \setminus \{z_0\}$. Thus, if γ_1 and γ_2 are two closed curves in $\mathcal{D}$ not passing through z_0 , and homotopic

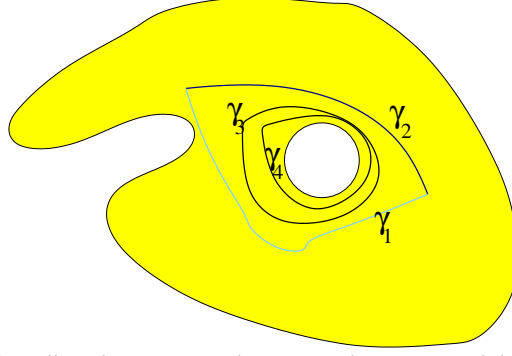


Figure 5: In the yellow domain, γ_1 and γ_2 are not homotopic, while γ_3 and γ_4 are.

to each-other in $\mathcal{D} \setminus \{z_0\}$ then

$$\int_{\gamma_1} (z - z_0)^{-n} dz = \int_{\gamma_2} (z - z_0)^{-n} dz \quad (5.4)$$

Clearly, these integrals are zero if γ_i does not contain z_0 inside. To calculate the integrals on a simple closed curve encircling z_0 it suffices, by (5.4), to do the calculation when the curve is a circle, which can be done explicitly. Indeed, a circle centered at z_0 with radius ρ is parametrized by $z = z_0 + \rho e^{it}$, $t \in [0, 2\pi]$ (where we used Euler's formula), and we get

$$\oint \frac{dz}{(z - z_0)^n} = \frac{i}{\rho^{n-1}} \int_0^{2\pi} e^{-i(n-1)t} dt = \begin{cases} 2\pi i & \text{if } n = 1 \\ 0 & \text{otherwise} \end{cases} \quad (5.5)$$

Definition 35

If γ is a curve and $z_0 \notin \gamma$, then the index of z_0 w.r.t. γ is defined as

$$\text{Ind}_\gamma(z_0) = \frac{1}{2\pi i} \oint_\gamma \frac{d\zeta}{\zeta - z_0}$$

and it represents the “number of times γ winds around z_0 ”.

Theorem 36: Cauchy's formula

If f is analytic in a simply connected domain $\mathcal{D}$ and γ is a piecewise differentiable simple closed curve in $\mathcal{D}$ enclosing z , then

$$f(z) = \frac{1}{2\pi i} \oint_\gamma \frac{f(s)}{s - z} ds. \quad (5.6)$$

Proof. Write

$$f(s) - f(z) = f'(z)(s - z) + \varepsilon(s)(s - z),$$

where $\varepsilon(s) \rightarrow 0$ as $s \rightarrow z$. Note that ε is continuous in $\mathcal{D} \setminus \{z\}$. Choose $\rho > 0$ small enough that $\mathbb{D} := \mathbb{D}_\rho(z) \subset \mathcal{D}$, and fix $\delta > 0$ with $|\varepsilon(s)| < \delta/(2\pi)$ for $s \in \partial\mathbb{D}$. Then

$$\begin{aligned} \frac{1}{2\pi i} \oint_\gamma \frac{f(s)}{s - z} ds &= \frac{1}{2\pi i} \oint_{\partial\mathbb{D}} \frac{f(s)}{s - z} ds \\ &= \frac{f(z)}{2\pi i} \oint_{\partial\mathbb{D}} \frac{ds}{s - z} + \frac{f'(z)}{2\pi i} \oint_{\partial\mathbb{D}} ds + \frac{1}{2\pi i} \oint_{\partial\mathbb{D}} \varepsilon(s) ds. \end{aligned} \quad (5.7)$$

By (5.5), the first integral equals $f(z)$, the second vanishes, and the last has absolute value $\leq \delta$. Letting $\delta \rightarrow 0$ gives the result. ■

6 Taylor series of analytic functions

Theorem 37: Taylor series; Cauchy's formula for higher derivatives

If f is analytic in a domain $\mathcal{D}$ and $z_0 \in \mathcal{D}$, then there exists $\rho > 0$ such that for all $z \in \mathbb{D}(z_0; \rho)$

$$f(z) = \sum_{k=0}^{\infty} c_k (z - z_0)^k, \quad (6.2)$$

where

$$c_k = \frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(z_0)} \frac{f(s)}{(s - z_0)^{k+1}} ds = \frac{f^{(k)}(z_0)}{k!}. \quad (6.3)$$

Proof. Assume f is analytic in $\mathcal{D}$ and fix $z_0 \in \mathcal{D}$. Choose $\rho > 0$ small enough so that $\overline{\mathbb{D}_\rho(z_0)} \subset \mathcal{D}$. By Theorem 36, for $z \in \mathbb{D}_\rho(z_0)$

$$f(z) = \frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(z_0)} \frac{f(s)}{s - z} ds. \quad (6.4)$$

Expand the kernel:

$$\frac{1}{s - z} = \frac{1}{s - z_0 - (z - z_0)} = \frac{1}{s - z_0} \sum_{k=0}^{\infty} \left(\frac{z - z_0}{s - z_0} \right)^k, \quad (6.5)$$

which converges absolutely for $|z - z_0| < |s - z_0| = \rho$. Since the series is bounded by

$$\frac{1}{|s - z_0|} \sum_{k=0}^{\infty} \left| \frac{z - z_0}{s - z_0} \right|^k,$$

the dominated convergence theorem allows us to interchange sum and integral, yielding

$$f(z) = \sum_{k=0}^{\infty} c_k (z - z_0)^k, \quad c_k = \frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(z_0)} \frac{f(s)}{(s - z_0)^{k+1}} ds. \quad (6.6)$$

■

Remark 38. *This theorem shows the central equivalence of complex analysis: analytic functions are exactly those that admit a convergent power series expansion in a neighborhood of every point of their domain. In contrast with real analysis, where differentiability need not imply such an expansion, holomorphic functions are “rigid”—their local behavior is completely determined by their power series.*

Exercise 39 (Cauchy remainder). *Show that*

$$f(z) = \sum_{k=0}^n c_k (z - z_0)^k + E(z, z_0, n), \quad (6.7)$$

where

$$E(z, z_0, n) = \frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(z_0)} \frac{f(s)}{s - z} \frac{(z - z_0)^{n+1}}{(s - z_0)^{n+1}} ds. \quad (6.8)$$

Prove that, if $\overline{\mathbb{D}_\rho(z_0)} \subset \mathcal{D}$, then

$$|E(z, z_0, n)| \leq \max_{\mathbb{D}_\rho(z_0)} |f| \frac{|z - z_0|^{n+1}}{\rho - |z - z_0|}. \quad (6.9)$$

If $\mathbb{D}_\rho(z_0)$ is the maximal disk centered at z_0 contained in $\mathcal{D}$, then the optimal estimate is

$$|E(z, z_0, n)| \leq \inf_{\rho' \leq \rho} \sup_{\mathbb{D}_{\rho'}(z_0)} |f| \frac{|z - z_0|^{n+1}}{\rho' - |z - z_0|}. \quad (6.10)$$

Theorem 40

If f is analytic in a domain $\mathcal{D}$ and $z_0 \in \mathcal{D}$, then f has derivatives of all orders at z_0 . In particular, $f', f'', \dots$ are analytic in $\mathcal{D}$.

Remark 41. The disk of convergence of the Taylor series of an analytic function cannot be zero. In fact, its radius of convergence equals the distance from z_0 to the nearest singularity of f (see Fig. 6).

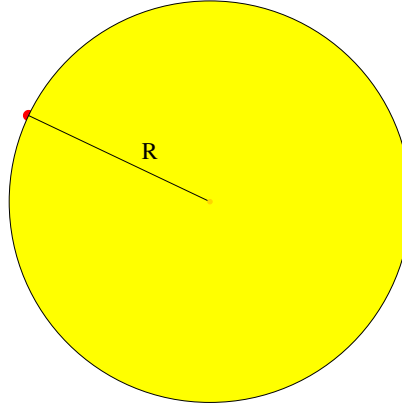


Figure 6: Disk of convergence of a Taylor series. The yellow region is a domain of analyticity and the red dot a singularity.

Indeed, in any smaller disk Theorem 37 applies. If the radius of convergence were larger than the distance to the nearest singularity, f would extend analytically beyond its given domain (why?), contradicting maximality.

Example. Consider $f(z) = \frac{1}{1+z^2}$. Its Taylor series at $z = 0$ is

$$\frac{1}{1+z^2} = \sum_{k=0}^{\infty} (-1)^k z^{2k}, \quad |z| < 1.$$

The singularities at $z = \pm i$ lie exactly on the boundary of the disk of convergence.

Theorem 42: Liouville

An **entire** function (analytic in all of $\mathbb{C}$) that is bounded on $\mathbb{C}$ must be constant.

Proof. Let $M = \sup_{\mathbb{C}} |f|$. By Theorem 37, for any $\rho > 0$,

$$f'(z) = \frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(0)} \frac{f(s)}{(s-z)^2} ds. \quad (6.11)$$

Therefore

$$|f'(z)| \leq \frac{1}{2\pi} M \frac{1}{\rho^2} (2\pi\rho) = \frac{M}{\rho}. \quad (6.12)$$

Letting $\rho \rightarrow \infty$ yields $f'(z) \equiv 0$, so f is constant. ■

Exercise 43. Show that an entire function that is not a polynomial must grow faster than any power of $|z|$ along some path as $z \rightarrow \infty$.

7 The Fundamental Theorem of Algebra

Exercise 44. Let P be a polynomial of degree $n \geq 1$. Show that there exists $R > 0$ such that $1/P$ is analytic in the domain $\{z : |z| > R\}$ and that for some nonzero c , $z^n/P(z) \rightarrow c$ as $|z| \rightarrow \infty$.

One classical application of Liouville's Theorem is the Fundamental Theorem of Algebra:

Theorem 45: Fundamental Theorem of Algebra

A polynomial $P_n(z)$ of degree $n \geq 1$ has exactly n roots in $\mathbb{C}$, counting multiplicity.

Proof. It suffices to show the existence of one root when $n \geq 1$, since the general statement then follows inductively by factoring.

By Exercise 44, if P_n had no zeros, then $1/P_n$ would be entire and vanish at infinity. But then $1/P_n$ would be bounded (why?) and entire, hence constant by Liouville's theorem. This is impossible, so P_n must have at least one zero. ■

8 More properties of analytic functions

Assume f is analytic in $\mathbb{D}_\rho(z_0)$ and that all derivatives of f vanish at z_0 . Then $f \equiv 0$ in $\mathbb{D}_\rho(z_0)$ (check!). In fact, more is true:

Proposition 8.1. If f is analytic in a domain $\mathcal{D}$ and all derivatives at some $z_0 \in \mathcal{D}$ vanish, then f is identically zero in $\mathcal{D}$.

Proof. For any $z \in \mathcal{D}$ there is a polygonal path P from z_0 to z , composed of segments $[z_{j-1}, z_j]$, $j = 1, \dots, n$, with disks $\mathbb{D}_{r_j}(z_j) \subset \mathcal{D}$ covering each step (see Proposition 49.3).

By simple geometry, we may choose $\rho > 0$ such that every point along P is contained in a disk of radius ρ lying in $\mathcal{D}$. Cover P by a finite chain of such overlapping disks, starting with one centered at z_0 .

Since f vanishes identically in the first disk, all its derivatives vanish at the center of the second disk, so f vanishes there too. Repeating this argument along the chain of disks, we conclude $f(z) = 0$ for all $z \in \mathcal{D}$. ■

Theorem 46: Morera

Let f be continuous in a simply connected domain $\mathcal{D}$ and assume that

$$\oint_{\gamma} f(z) dz = 0$$

for every simple piecewise C^1 closed curve $\gamma \subset \mathcal{D}$. Then f is analytic in $\mathcal{D}$. (The same conclusion holds when the curves γ are boundaries of triangles; see Exercise 47 below.)

Proof. Fix $z_0 \in \mathcal{D}$ and define

$$F(z) = \int_{z_0}^z f(s) ds,$$

where the path of integration is any piecewise C^1 path in $\mathcal{D}$. By the assumption $\oint_{\gamma} f = 0$, this definition is path-independent.

Let δ be small and choose a path γ' from z_0 to $z + \delta$. Then

$$\int_{\gamma} f(s) ds - \int_{\gamma'} f(s) ds = \int_z^{z+\delta} f(s) ds,$$

where the last integral can be taken along a straight line. Writing $f(s) = f(z) + \varepsilon(s)$ with $\varepsilon(s) \rightarrow 0$ as $s \rightarrow z$, we obtain

$$\int_z^{z+\delta} f(s) ds = f(z)\delta + o(\delta).$$

Thus F is differentiable with $F'(z) = f(z)$. Since F is analytic, so is f . The restriction to triangles follows from polygonal connectivity and is left as an exercise. ■

Exercise 47 (Restriction to triangles). (a) Prove Morera's theorem assuming $\oint_{\gamma} f = 0$ only for triangular paths, using polygonal-path connectedness. (b) Give an alternative proof based on approximation of curves by polygonal lines.

We now have three equivalent viewpoints on analytic functions in a domain $\mathcal{D}$:

- as functions differentiable with respect in $\mathcal{D}$,
- as sums of convergent power series centered at any point in $\mathcal{D}$,
- as continuous functions with vanishing integrals along closed loops.

Each characterization is fundamental and useful.

Theorem 48: Weierstrass

Let f_n be analytic in a domain Ω and suppose $f_n \rightarrow f$ uniformly on compact subsets of Ω . Then f is analytic in Ω . Moreover, $f'_n \rightarrow f'$ uniformly on compact subsets of Ω .

Note 49

By iterating, it follows that $f_n^{(k)} \rightarrow f^{(k)}$ uniformly on compact sets for each $k \in \mathbb{N}$.

Proof. Let T be a triangle contained in a compact $K \subset \Omega$. Since f_n is analytic,

$$\int_T f_n(z) dz = 0.$$

By uniform convergence, f is continuous on K and dominated convergence applies, giving

$$\int_T f(z) dz = 0.$$

By Morera's theorem, f is analytic. The convergence of the derivatives follows directly from Cauchy's integral formula and dominated convergence. ■

9 Harmonic functions

A real-valued C^2 function $u(x, y)$ satisfying Laplace's equation

$$u_{xx} + u_{yy} = 0 \tag{9.2}$$

in a domain U is called *harmonic* in U .

Theorem 50

A function u is harmonic in a simply connected domain $\mathcal{D} \subset \mathbb{C}$ if and only if it is the real (or imaginary) part of some analytic function f in $\mathcal{D}$. The function f is unique up to an additive purely imaginary constant.

In other words: for each harmonic u there exists a harmonic conjugate v (unique up to a constant) such that $u + iv$ is analytic.

Proof. If $u = \operatorname{Re}(f)$ for some analytic f , then $f = u + iv$ satisfies the Cauchy–Riemann equations

$$u_x = v_y, \quad u_y = -v_x.$$

Differentiating once more, we obtain

$$u_{xx} = v_{yx}, \quad u_{yy} = -v_{xy}.$$

Since $v_{xy} = v_{yx}$, it follows that $u_{xx} + u_{yy} = 0$, i.e. u is harmonic.

Conversely, suppose u is harmonic in $\mathcal{D}$. Define the vector field

$$\mathbf{E} = (-u_y, u_x).$$

By the harmonicity condition $u_{xx} + u_{yy} = 0$, one checks that

$$\frac{\partial}{\partial x}(u_x) = \frac{\partial}{\partial y}(-u_y),$$

which shows that $\mathbf{E}$ is curl-free. Since $\mathcal{D}$ is simply connected, $\mathbf{E}$ is conservative, so there exists a potential v (unique up to an additive constant) such that

$$v_x = -u_y, \quad v_y = u_x.$$

These are exactly the Cauchy–Riemann equations for $u + iv$, so $f = u + iv$ is analytic in $\mathcal{D}$. ■

Exercise 51. Assume f is analytic in a domain $\mathcal{D}$ and write $f = u + iv$ with $u = \operatorname{Re} f$ and $v = \operatorname{Im} f$ on $\mathcal{D}$. Show that

$$\begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} \operatorname{Re} f'(z) & -\operatorname{Im} f'(z) \\ \operatorname{Im} f'(z) & \operatorname{Re} f'(z) \end{pmatrix}. \quad (9.3)$$

Hint: See the section on the C-R equations.

Lemma 52. Let $f = u + iv$ be analytic near z_0 with $f'(z_0) \neq 0$. Then the constant level curves $u(x, y) = u(x_0, y_0)$ and $v(x, y) = v(x_0, y_0)$ exist near z_0 , are smooth, and intersect orthogonally.

Proof. Left as an exercise, since the result follows immediately from (9.3). ■

9.1 Potential and Hamiltonian flows

Consider an autonomous system of ODEs in $\mathcal{D}$:

$$\dot{x} = E_1(x, p), \quad \dot{p} = E_2(x, p).$$

The system is *Hamiltonian* if there exists $H \in C^1(\mathcal{D})$ such that

$$E_1 = \frac{\partial H}{\partial p}, \quad E_2 = -\frac{\partial H}{\partial x}.$$

It is a *potential system* if there exists $V \in C^1(\mathcal{D})$ with

$$E_1 = \frac{\partial V}{\partial x}, \quad E_2 = \frac{\partial V}{\partial p}.$$

A system is simultaneously potential and Hamiltonian precisely when there exist H and V such that

$$\frac{\partial H}{\partial p} = \frac{\partial V}{\partial x}, \quad \frac{\partial H}{\partial x} = -\frac{\partial V}{\partial p}.$$

Thus H and V must be harmonic conjugates, i.e. the real and imaginary parts of an analytic function.

In Hamiltonian systems, H is conserved:

$$\frac{d}{dt}H(x(t), p(t)) = H_x \dot{x} + H_p \dot{p} = 0,$$

so H remains constant along trajectories. In gradient systems, $(\dot{x}, \dot{p})$ points in the direction of steepest ascent of V .

Note 53

The steepest descent/ascent lines of a harmonic function u satisfy the ODE system

$$\dot{x} = v_x, \quad \dot{y} = v_y,$$

where v is the harmonic conjugate. The solution curves are smooth wherever $v_x^2 + v_y^2 \neq 0$. This viewpoint will be important in the method of steepest descent.

10 The maximum modulus principle

One of the most striking features of analytic functions is that their absolute value cannot attain an interior maximum unless the function is constant:

Theorem 54: Maximum Modulus Principle

If f is analytic and nonconstant in a domain $\mathcal{D}$, then $|f|$ cannot achieve a maximum at any point of $\mathcal{D}$.

In other words, any maximum of $|f|$, if achieved, must occur on the boundary of $\mathcal{D}$.

Proof. Assume, for contradiction, that $z_0 \in \mathcal{D}$ is a point where $|f|$ attains its maximum. Since f is nonconstant, it cannot be identically zero. Let $M = \max_{\mathcal{D}} |f| = |f(z_0)| > 0$. By scaling, we may assume without loss of generality that $M = 1$ and $z_0 = 0$.

If $f \equiv 1$ identically, this contradicts the assumption that f is nonconstant. Otherwise, in the Taylor expansion of f about 0,

$$f(z) = 1 + c_k z^k + c_{k+1} z^{k+1} + \cdots,$$

let $k \geq 1$ be the smallest index with $c_k = |c_k|e^{i\theta} \neq 0$. Factor

$$f(z) = 1 + c_k z^k (1 + d_k z E_1(z)), \quad (10.2)$$

where E_1 is analytic near 0 and bounded.

Choose $z_1 = |z_1|e^{-i\theta}$ small enough so that $|z_1 d_k E_1(z_1)| < 1$. Then $c_k z_1^k = |c_k||z_1|^k \in \mathbb{R}^+$. By the triangle inequality,

$$|f(z_1)| = |1 + c_k z_1^k (1 + d_k z_1 E_1(z_1))| > 1,$$

contradicting the maximality of $|f|$ at 0.

Therefore f must be constant. ■

Exercise 55. Show that if $|f|$ has a minimum in $\mathcal{D}$, then this minimum must be 0.

Exercise 56. * Find the maximum and minimum values of $|\sin z|$ inside the closed unit disk.

Steepest ascent and descent lines

Let f be analytic near z_0 , and let $k \geq 1$ be the smallest index such that $c_k \neq 0$ in the Taylor expansion

$$f(z) = f(z_0) + c_k (z - z_0)^k + \cdots.$$

Definition 57

Let $f(z) = \sum_{j=0}^{\infty} c_j (z - z_0)^j$ with $k \geq 1$ the smallest index such that $c_k \neq 0$. The direction of $d \in \mathbb{C}$ is called a *steepest ascent direction* for $|f|$ at z_0 if $c_k d^k \in \mathbb{R}^+$, and a *steepest descent direction* if $c_k d^k \in \mathbb{R}^-$.

A piecewise differentiable oriented curve γ is called a *steepest ascent (descent) line* for f if, at every point of differentiability $z \in \gamma$, its tangent vector is a steepest ascent (descent) direction for $|f|$ at z .

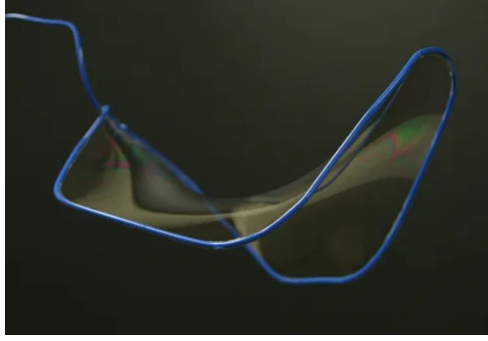


Figure 7: A soap film (image from the google images, <https://math.hmc.edu/~jacobson/demolab/soap-film-2/>)

Note 58

By Note 53, the lines of steepest ascent and descent are smooth and unique through any z_0 such that

$$f(z) = f(z_0) + f'(z_0)(z - z_0) + \cdots, \quad f'(z_0) \neq 0.$$

A **saddle point** is a point where $f'(z_0) = 0$. Then f is locally of the form $z^k + \cdots$ with $k \geq 2$. One checks that there are exactly k distinct directions of steepest ascent at such a point. Consequently, lines of steepest descent can be continued, in k possible ways, through a saddle point. These notions are central in the *steepest descent method* (also called the *saddle point method*).

Theorem 59

Assume u is harmonic and nonconstant in a domain $\mathcal{D}$. Then u has no maximum and no minimum in $\mathcal{D}$.

Proof. Let $u = \operatorname{Re}(f)$ and define $g = e^f$, which is analytic in $\mathcal{D}$. Writing $e^f = e^u e^{iv}$, we see that $|e^f| = e^u$. Thus u has a maximum if and only if $|g|$ has a maximum. But by the maximum modulus principle, $|g|$ cannot attain a maximum strictly inside $\mathcal{D}$. For a minimum, note that $\min(u) = -\max(-u)$, which reduces to the previous case. ■

10.1 Application: soap films

A soap film picked up by a thin closed wire assumes the surface of *least area* compatible with the constraint that it is bounded by the wire (since the potential energy is proportional to the surface area).

If the wire is close to planar (say lying near the (x, y) -plane), the local height $u(x, y)$ of the film satisfies Laplace's equation. (This can be shown by elementary differential geometry.)²

If the wire is exactly planar, then $u = 0$ on the boundary, and by Theorem 59, the minimal surface is flat. This is unsurprising.

More generally, we will be able to solve Laplace's equation with arbitrary boundary conditions, yielding insight into minimal surfaces. Conversely, the intuition we have about soap

²If the film can be written locally as a graph $z = u(x, y)$ over a planar domain $\Omega \subset \mathbb{R}^2$, then u solves the *minimal surface equation*

$$\operatorname{div} \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = 0 \quad \text{in } \Omega, \quad u|_{\partial\Omega} = u_{\partial\Omega}, \quad (10.3)$$

where the boundary data $u_{\partial\Omega}$ encodes the wire (after projection to the plane).

films translates directly to solutions of Laplace's equation: for example, a soap film has no local extremum, since any local flattening would reduce the surface area.

10.2 Cauchy principal value integrals (PV)

Suppose f is analytic in a domain containing the simple closed, piecewise differentiable curve C . By Cauchy's theorem we have

$$\frac{1}{2\pi i} \oint_C \frac{f(s)}{s-z} ds = \begin{cases} f(z), & z \text{ inside } C, \\ 0, & z \text{ outside } C. \end{cases} \quad (10.4)$$

What if z lies on C ? Then the integral requires interpretation. One definition is to excise from C a symmetric segment of length ε centered at z , and then let $\varepsilon \rightarrow 0$. This yields the *Cauchy principal value integral*, denoted $\text{PV} \oint$. Another equivalent definition is to take the half-sum of the integrals along curves circumventing z from the inside and from the outside.

Exercise 60. Show that if C is a smooth closed curve and f is analytic in a neighborhood of C , then

$$\frac{1}{2\pi i} \text{PV} \oint_C \frac{f(s)}{s-z} ds = \frac{1}{2} f(z), \quad (10.5)$$

and that the two definitions above coincide. In particular, if $f(s)/(s-z)$ extends analytically across z , then the PV integral agrees with the usual one.

Definition 61: Cauchy PV on the line

If for every small $\varepsilon > 0$ one has $f \in L^1((a, L-\varepsilon)) \cap L^1((L+\varepsilon, b))$, then the principal value is defined by

$$\text{PV} \int_a^b f(s) ds := \lim_{\varepsilon \downarrow 0} \left(\int_a^{L-\varepsilon} f(s) ds + \int_{L+\varepsilon}^b f(s) ds \right), \quad (10.6)$$

provided the limit exists.

Example 62. The *Hilbert transform*, an important operator in analysis and applications, is defined by

$$H(u)(t) = \frac{1}{\pi} \text{PV} \int_{-\infty}^{+\infty} \frac{u(\tau)}{t-\tau} d\tau. \quad (10.7)$$

It is bounded on $L^p(\mathbb{R})$ for $1 < p < \infty$, where the limit exists pointwise almost everywhere, and maps L^1 to weak L^1 . With some Hölder regularity, the limit exists everywhere. The Hilbert transform is a central tool in solving Riemann–Hilbert problems. to which we will return later.

Exercise 63. Show that if $f \in L^1(a, b)$ then

$$\text{PV} \int_a^b f(s) ds = \int_a^b f(s) ds. \quad (10.8)$$

10.3 Bi-analytic bijections (automorphisms) of $\mathbb{D}$

Exercise 64. Let $a \in \mathbb{C}$ with $|a| < 1$ and $\theta \in \mathbb{R}$. Show that

$$A(z) := e^{i\theta} \frac{a+z}{1+\bar{a}z} \quad (10.9)$$

is an automorphism of $\mathbb{D}$.

Remark 65. We will prove later (using Schwarz's Lemma) that every automorphism of $\mathbb{D}$ is of the form (10.9) for some $a \in \mathbb{C}$ with $|a| < 1$ and $\theta \in \mathbb{R}$.

11 Poisson's formula

A function u is said to be harmonic in an open set $\Omega \subset \mathbb{R}^n$ if it is of class C^2 in Ω and satisfies Laplace's equation

$$\Delta u = 0 \quad \text{in } \Omega.$$

The **mean value property** for harmonic functions states that if $B(x, r)$ is a ball with center x and radius r completely contained in Ω , then the value of u at the center of the ball is given by the average value of u :

$$u(x) = \frac{1}{n\omega_n r^{n-1}} \int_{\partial B(x,r)} u d\sigma = \frac{1}{\omega_n r^n} \int_{B(x,r)} u dV,$$

where ω_n is the volume of the unit ball in $\mathbb{R}^n$ and $d\sigma$ denotes the $(n-1)$ -dimensional surface measure.

We shall first prove this property in the case $n = 2$. Clearly, we can change variables so that $x = 0$ and $r = 1$. In complex notation, the statement becomes:

Proposition 11.2 (Mean value property for harmonic functions). Assume u is harmonic in the open unit disk and continuous on the closed unit disk. Then

$$u(0) = \frac{1}{2\pi} \int_0^{2\pi} u(e^{it}) dt. \quad (11.2)$$

Proof. If v is a harmonic conjugate of u , then $f := u + iv$ is analytic in the open unit disk. By Cauchy's integral formula, for any $\rho < 1$ we have

$$\begin{aligned} u(0) + iv(0) &= f(0) = \frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(0)} \frac{f(s)}{s} ds = \frac{1}{2\pi} \int_0^{2\pi} f(\rho e^{it}) dt \\ &= \frac{1}{2\pi} \int_0^{2\pi} u(\rho e^{it}) dt + i \frac{1}{2\pi} \int_0^{2\pi} v(\rho e^{it}) dt. \end{aligned} \quad (11.3)$$

Taking the real part and letting $\rho \rightarrow 1^-$ yields (11.2). ■

Exercise 66. (i) Let u be as in Proposition 11.2 and let A be an automorphism of the disk, as in Exercise 64. Show that

$$U(z) = u(A(z))$$

is harmonic in the open unit disk and continuous on the closed unit disk. (Hint: Write the Laplacian as $\Delta = 4\partial_z\partial_{\bar{z}}$.)

(ii) Show that, if $z_0 = re^{i\theta}$, then

$$u(z_0) = \frac{1}{2\pi} \int_0^{2\pi} u\left(e^{i\theta} \frac{r + e^{is}}{1 + re^{is}}\right) ds = \frac{1}{2\pi} \int_{-\pi}^{\pi} u\left(e^{i\theta} \frac{r + e^{is}}{1 + re^{is}}\right) ds. \quad (11.4)$$

(Hint for (ii): Use (i) with $A(w) = e^{i\theta} \frac{r + w}{1 + rw}$, so that $A(0) = z_0$, and apply Proposition 11.2 to $U = u \circ A$.)

Proposition 11.3 (Poisson's formula). Let $f(e^{it})$ be continuous on $[-\pi, \pi)$. Define

$$u(re^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} \frac{1 - r^2}{1 - 2r \cos(t - \theta) + r^2} f(e^{it}) dt. \quad (11.5)$$

Then u is harmonic in $\mathbb{D}$, extends continuously to $\overline{\mathbb{D}}$, and satisfies $u(e^{i\theta}) = f(e^{i\theta})$.

Moreover, the **Poisson kernel**

$$P_r(\alpha) := \frac{1 - r^2}{1 - 2r \cos \alpha + r^2}$$

satisfies $P_r(\alpha) \geq 0$ and

$$\frac{1}{2\pi} \int_0^{2\pi} P_r(\alpha) d\alpha = 1.$$

Proof. Harmonicity follows either from a direct computation $\Delta u = 0$.

To conclude the proof of the first part, it remains to show that u extends continuously to $\partial\mathbb{D}$ and that

$$\lim_{r \rightarrow 1^-} u(re^{i\theta}) = f(e^{i\theta}).$$

Fix θ and, by rotation, assume $\theta = 0$. Then

$$u(r) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{1 - r^2}{1 - 2r \cos t + r^2} f(e^{it}) dt.$$

Split the integral at $|t| = \delta$. For $|t| > \delta$,

$$0 \leq \frac{1 - r^2}{1 - 2r \cos t + r^2} \leq \frac{(1 + r)(1 - r)}{2r(1 - \cos \delta)} \rightarrow 0 \quad (r \rightarrow 1^-),$$

so by dominated convergence this part tends to 0.

For $|t| \leq \delta$, write $\cos t = 1 - \frac{t^2}{2} \varphi(t)$ with $\varphi(t) \rightarrow 1$. Then

$$\frac{1 - r^2}{1 - 2r \cos t + r^2} = \frac{(1 + r)(1 - r)}{(1 - r)^2 + rt^2 \varphi(t)}.$$

Setting $t = (1 - r)\tau$ gives

$$u(r) = \frac{1}{2\pi} \int_{-\frac{\delta}{1-r}}^{\frac{\delta}{1-r}} \frac{1 + r}{1 + r\tau^2 \varphi((1 - r)\tau)} f(e^{i(1-r)\tau}) d\tau,$$

and dominated convergence (check the details here) yields

$$\lim_{r \rightarrow 1^-} u(r) = \frac{2}{2\pi} f(1) \int_{-\infty}^{\infty} \frac{d\tau}{1 + \tau^2} = f(1).$$

Thus u extends continuously to $\partial\mathbb{D}$ with boundary values f . The positivity and normalization properties of the Poisson kernel are straightforward to verify. ■

11.1 The Dirichlet problem for the Laplacian in $\mathbb{D}$

Formula (11.5) provides the solution of Laplace's equation in two dimensions with Dirichlet boundary conditions, namely with u prescribed on the boundary, when the domain is $\mathbb{D}$. A simple change of variables adapts this formula to any disk. More generally, the formula can be extended to apply to a general simply connected domain bounded by a simple, closed, piecewise differentiable curve. This follows from the *Riemann mapping theorem*.

Example 67 (The Faraday cage). “A region surrounded by a conductor does not feel the electrical influence of static outside charges.” We illustrate this phenomenon in two dimensions.

At equilibrium, the electric potential along a conductor is constant. Indeed, if two points of the conductor were at different potentials, there would exist a potential difference, hence an electric current $I = V/R$ (with R the resistivity). This would contradict equilibrium. This is, of course, a standard physics argument, though not a mathematically rigorous one. A full mathematical treatment would require setting up the precise PDE model and proving existence and uniqueness results, which lies beyond the scope of our course.

Thus we consider Laplace's equation (24.2) with boundary condition $V = C$ on $\partial\mathcal{D}$. Since the constant function $V = C$ is a solution, uniqueness of solutions implies that it is *the* solution. Hence

$$\mathbf{E} = -\nabla V = 0,$$

which is the desired conclusion.

12 Isolated singularities, Laurent series

Definition 68

A function f has an *isolated singularity* at z_0 if f is analytic in a *punctured disk* $\mathbb{D}_\rho(z_0) \setminus \{z_0\}$ for some $\rho > 0$.

More generally, we will analyze functions analytic in annuli

$$\mathbb{D}_\rho(z_0) \setminus \mathbb{D}_{\rho'}(z_0), \quad 0 < \rho' < \rho.$$

For example, the functions $e^{1/z}$ and $1/\sin z$ have isolated singularities at $z = 0$, whereas the singularity of $\ln z$ at 0 is not isolated (we will see that $\ln$ is not well defined in $\mathbb{D} \setminus \{0\}$).

Definition 69

An isolated singularity z_0 of f is classified as follows:

1. a **removable singularity** if f extends to a function $\tilde{f}$ analytic in the whole disk. By a slight abuse of notation, we typically do not distinguish $\tilde{f}$ from f itself;
2. a **pole of order** $M \geq 1$ if $(z - z_0)^M f$ has a removable singularity at z_0 , but $(z - z_0)^{M-1} f$ does not;
3. an **essential singularity** if it is neither removable nor a pole.

A function all of whose singularities are poles is called **meromorphic**.

In the following, we denote the punctured disk $\mathbb{D}_\rho(z_0) \setminus \{z_0\}$ by $\mathbb{D}_\rho^*(z_0)$.

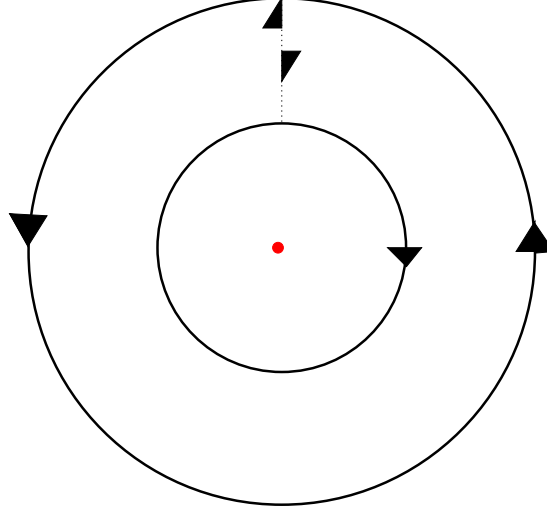


Figure 8: Circles of integration and cut in the proof of the Laurent series expansion.

Proposition 12.4 (Laurent series). *If f is analytic in $\mathbb{D}_\rho(z_0) \setminus \mathbb{D}_{\rho'}(z_0)$, where $0 < \rho' < \rho$, then f has a convergent representation*

$$f(z) = \sum_{k \in \mathbb{Z}} a_k (z - z_0)^k, \quad z \in \mathbb{D}_\rho(z_0) \setminus \mathbb{D}_{\rho'}(z_0), \quad (12.2)$$

where

$$a_k = \frac{1}{2\pi i} \oint \frac{f(s)}{(s - z_0)^{k+1}} ds, \quad k \in \mathbb{Z},$$

with the integral taken along any circle centered at z_0 with radius between ρ' and ρ .

Proof. Without loss of generality, take $z_0 = 0$. Consider the annulus between two circles C_o and C_i in $A = \mathbb{D}_\rho \setminus \mathbb{D}_{\rho'}$, as in Fig. 8 (C_o is the outer boundary and C_i the inner one). Introduce a cut L in the annulus as shown, so that $A \setminus L$ is simply connected. By Cauchy's integral formula, for $z \in A \setminus L$ we have

$$f(z) = \frac{1}{2\pi i} \oint_{C_o} \frac{f(s)}{s - z} ds - \frac{1}{2\pi i} \oint_{C_i} \frac{f(s)}{s - z} ds. \quad (12.3)$$

Here we used the fact that the boundary of the cut annulus is the closed path composed of C_o (counterclockwise), followed by the cut segment ℓ , then C_i (clockwise), and finally $-\ell$. The rest of the argument parallels the proof of Taylor's theorem and is left as an exercise. ■

Exercise 70. Complete the proof of formula (12.2) by expanding the integrands in (12.3) into power series in z/s and s/z , respectively, and estimating the remainders as we did in deriving formula (6.4).

Note. Check that (12.3) provides a decomposition of f into two parts: a function f_1 , analytic in the disk enclosed by C_o , and a function f_2 , analytic in $1/z$ in the exterior of the disk bounded by C_i . In [3] this decomposition is used to give an elegant proof of (12.2). For example, $e^{1/z}$ has an essential singularity at $z = 0$. Application of (12.2) yields

$$e^{1/z} = \sum_{k=0}^{\infty} \frac{1}{k!} z^{-k}. \quad (12.4)$$

(Of course, this formula could also be obtained directly from the power series of the exponential by replacing z with z^{-1} .)

Note. The part of the Laurent series consisting of the terms with negative powers is called the *principal part* of the series.

Note. Laurent series are of great theoretical value. However, they are not practical for *computing* the values of functions near an essential singularity, since the convergence becomes slower as the singularity is approached.

Theorem 71: Casorati–Weierstrass; also known as the “little Picard theorem”

Assume f is analytic in $\mathbb{D}_\rho^*(z_0)$ and has an essential singularity at z_0 . Then, for any $\rho_1 < \rho$, the set $f(\mathbb{D}_{\rho_1}^*(z_0))$ is dense in $\mathbb{C}$.

Proof. Without loss of generality, assume $z_0 = 0$. Suppose, for contradiction, that there exists $\zeta \in \mathbb{C}$ and $\rho_1 > 0$ such that

$$|f(z) - \zeta| > c > 0 \quad \text{for all } z \in \mathbb{D}_{\rho_1}^*(0).$$

Then

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

is bounded in $\mathbb{D}_{\rho_1}^*(0)$, and hence 0 is a removable singularity of h (why?). But then

$$f(z) = \zeta + \frac{1}{h(z)}$$

would have either a removable singularity or a pole at 0 (why?), contradicting the assumption that 0 is an essential singularity. ■

13 Laurent series and Fourier series

Let f be 2π -periodic and assume:

Assumption. The Fourier coefficients $\{c_k\}_{k \in \mathbb{Z}}$ of f satisfy $\{c_k\} \in \ell^1(\mathbb{Z})$ ³.

Then

$$g(z) := \sum_{k=0}^{\infty} c_k z^k$$

converges absolutely and uniformly in $\overline{\mathbb{D}}$ (why?), hence g is analytic in $\mathbb{D}$ and continuous up to $\partial\mathbb{D}$ (why?). For all $k \leq -1$ we then have

$$\oint_{\partial\mathbb{D}} \frac{g(s)}{s^{k+1}} ds = 0.$$

Similarly, define

$$H(\zeta) := \sum_{k=1}^{\infty} c_{-k} \zeta^k.$$

³ $\ell^1(\mathbb{Z})$ denotes the space of sequences $\{c_k\}_{k \in \mathbb{Z}}$ which are absolutely summable, i.e. $\sum_{k \in \mathbb{Z}} |c_k| < \infty$.

Then H is analytic in $\mathbb{D}$ and continuous up to $\overline{\mathbb{D}}$. Thus

$$h(z) := H\left(\frac{1}{z}\right)$$

is analytic in $\text{ext}(\overline{\mathbb{D}})$ and continuous down to $\partial\mathbb{D}$. For all $k \geq 0$ one checks that

$$\oint_{\partial\mathbb{D}} \frac{h(s)}{s^{k+1}} ds = 0.$$

Hence, for $s \in \mathbb{R}$,

$$f(s) = g(e^{is}) + h(e^{is}).$$

We have therefore proved the following result.

Proposition 13.5. *If f is 2π -periodic and its Fourier coefficients satisfy the assumption above, then f decomposes as the sum of two functions: one analytic in $\mathbb{D}$ and continuous in $\overline{\mathbb{D}}$, and another analytic in $\text{ext}(\overline{\mathbb{D}})$ and continuous in $\text{ext}(\mathbb{D})$.*

14 Calculating the Taylor series of simple functions

One easy way to calculate Taylor series is to use operations with series; see §3.3.

Example. (1) The Taylor series of the function $z^{-1} \sin z$ is

$$\frac{\sin z}{z} = 1 - \frac{z^2}{6} + \frac{z^4}{120} - \dots \quad (14.2)$$

(2) The Taylor series of the function $z / \sin z$ is

$$\begin{aligned} \frac{z}{\sin z} &= \frac{1}{1 - \left(\frac{z^2}{6} - \frac{z^4}{120} + \dots\right)} = 1 + \left(\frac{z^2}{6} - \frac{z^4}{120} + \dots\right) \\ &\quad + \left(\frac{z^2}{6} - \frac{z^4}{120} + \dots\right)^2 + \dots \\ &= 1 + \frac{z^2}{6} - \frac{z^4}{120} + \frac{z^4}{36} + \dots = 1 + \frac{z^2}{6} + \frac{7z^4}{360} + \dots \end{aligned} \quad (14.3)$$

The first function is entire, while the second one is not. *Question:* What is the radius of convergence of the second series?

Exercise 72. Find the integral of $1/\cos z$ on a circle of radius $1/2$ centered at $z_0 = \pi/2$.

15 Residues and integrals

Definition 73

A function f is *meromorphic* in a neighborhood of a point z_0 if it is analytic in some punctured domain $\mathcal{D} \setminus \{z_0\}$ and there exists $k \geq 0$ such that $(z - z_0)^k f(z)$ has a removable singularity at z_0 .

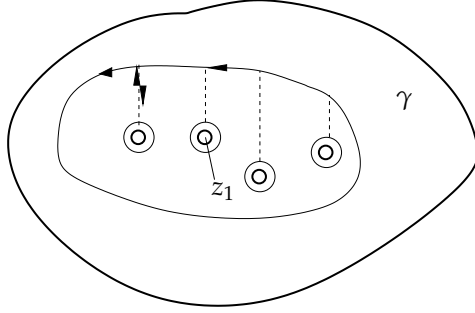


Figure 9: Multiply connected domains

Note 74

Meromorphic functions have *one-sided* convergent Laurent expansions in $\mathcal{D} \setminus \{z_0\}$, in the sense that only finitely many Laurent coefficients of negative order are nonzero.

Definition 75: Residues

Let f be meromorphic in $\mathcal{D} \setminus \{z_0\}$. The coefficient c_{-1} in its Laurent expansion at z_0 is called the **residue** of f at z_0 , denoted $\text{Res}(f, z_0)$. If $\mathbb{D}_\rho(z_0) \subset \mathcal{D}$, then

$$\frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(z_0)} f(s) ds = c_{-1},$$

hence the terminology “residue.”

Proposition 15.6. Let $\mathcal{D}$ be a simply connected domain. Suppose f is analytic in $\mathcal{D} \setminus \bigcup_{k=1}^n \mathbb{D}_k$, where $\mathbb{D}_k$ are disjoint disks centered at $z_k \in \mathcal{D}$. Let γ be a simple closed, piecewise differentiable curve which encircles each $\mathbb{D}_k$ once (see Definition 89 and Figure 15). Then

$$\oint_{\gamma} f(s) ds = 2\pi i \sum_{k=1}^n \text{Res}(f, z_k). \quad (15.2)$$

Exercise 76. Prove Proposition 17.7 by deforming and cutting the curve of integration appropriately.

Example. Calculate

$$\oint \frac{dz}{\sin^3 z}$$

on a circle of radius 1/2 around the origin.

Solution. In $\mathbb{D}(0, 1/2)$ we have

$$\begin{aligned} \frac{1}{\sin^3 z} &= \frac{1}{\left(z - \frac{z^3}{6} + \frac{z^5}{120} - \dots\right)^3} = \frac{1}{z^3} \frac{1}{\left(1 - \frac{z^2}{6} + \frac{z^4}{120} - \dots\right)^3} \\ &= \frac{1}{z^3} \left(1 + \frac{z^2}{2} + \frac{17z^4}{120} + \dots\right). \end{aligned} \quad (15.3)$$

Thus the residue of $\sin^{-3}(z)$ at $z = 0$ is 1/2, and the integral equals πi .

Exercise 77. Show that if f has a pole of order m at $z = z_i$, then

$$\text{Res}(f, z_i) = \frac{1}{(m-1)!} \left[\frac{d^{m-1}}{dz^{m-1}} \left((z - z_i)^m f(z) \right) \right]_{z=z_i} \quad (15.4)$$

by applying Laurent's formula near $z = z_i$.

Note 78: Computational tool

Formula (17.4) is the standard method for computing residues at poles of finite order. It reduces the problem to taking a finite number of derivatives of an analytic function. In practice, this formula, together with the method used in 15.3, are often the main “go-to” tools for residue computations; for low order poles the latter method is usually faster.

16 Calculating the Taylor series of simple functions

One easy way to calculate Taylor series is to use operations with series; see §3.3.

Example.

1. The Taylor series of the function $z^{-1} \sin z$ is

$$\frac{\sin z}{z} = 1 - \frac{z^2}{6} + \frac{z^4}{120} - \dots \quad (16.2)$$

2. The Taylor series of the function $z / \sin z$ is

$$\begin{aligned} \frac{z}{\sin z} &= \frac{1}{1 - (z^2/6 - z^4/120 + \dots)} \\ &= 1 + (z^2/6 - z^4/120 + \dots) + (z^2/6 - z^4/120 + \dots)^2 + \dots \\ &= 1 + \frac{z^2}{6} - \frac{z^4}{120} + \frac{z^4}{36} + \dots = 1 + \frac{z^2}{6} + \frac{7z^4}{360} + \dots \end{aligned} \quad (16.3)$$

The first function is entire; the second one is not. *Question:* What is the radius of convergence of the second series?⁴

Exercise 79. Find the integral of $1/\cos z$ on a circle of radius $1/2$ centered at $z_0 = \pi/2$.

17 Residues and integrals

Definition 80

A function f is meromorphic in a neighborhood of a point z_0 if it is analytic in some punctured domain $\mathcal{D} \setminus \{z_0\}$ and there is a $k \geq 0$ so that $(z - z_0)^k f$ has a removable singularity at z_0 .

⁴Hint: the radius of convergence is determined by the distance from 0 to the nearest pole of $z / \sin z$.

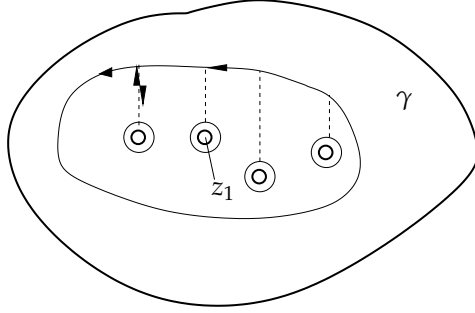


Figure 10: Multiply connected domains

Note 81

Meromorphic functions have *one-sided* convergent Laurent expansions in $\mathcal{D} \setminus \{z_0\}$, in the sense that only a finite number of Laurent coefficients of negative order are nonzero (why?).

Definition 82: Residues

Let f be meromorphic in $\mathcal{D} \setminus \{z_0\}$. The coefficient c_{-1} in its Laurent expansion at z_0 is called the **residue** of f at z_0 , denoted $\text{Res}(f, z_0)$. If $\mathbb{D}_\rho(z_0) \subset \mathcal{D}$, then

$$\frac{1}{2\pi i} \oint_{\partial \mathbb{D}_\rho(z_0)} f(s) ds = c_{-1},$$

hence the name “residue.”

Proposition 17.7. Let $\mathcal{D}$ be a simply connected domain. Suppose f is analytic in $\mathcal{D} \setminus \bigcup_{k=1}^n \mathbb{D}_k$, where $\mathbb{D}_k$ are disks centered at $z_k \in \mathcal{D}$. Let γ be a simple closed, piecewise differentiable curve which encircles each $\mathbb{D}_k$ once (see Definition 89 and Figure 15). Then

$$\oint_{\gamma} f(s) ds = 2\pi i \sum_{k=1}^n \text{Res}(f, z_k). \quad (17.2)$$

Exercise 83. Prove Proposition 17.7 by deforming and cutting the curve of integration appropriately.

Example. Calculate

$$\oint \frac{dz}{\sin^3 z}$$

on a circle of radius $1/2$ around the origin.

Solution. In $\mathbb{D}(0, 1/2)$ $\sin$ vanishes only at the origin. We have

$$\begin{aligned} \frac{1}{\sin^3 z} &= \frac{1}{\left(z - \frac{z^3}{6} + \frac{z^5}{120} - \dots\right)^3} = \frac{1}{z^3} \frac{1}{\left(1 - \frac{z^2}{6} + \frac{z^4}{120} - \dots\right)^3} \\ &= \frac{1}{z^3} \left(1 + \frac{z^2}{2} + \frac{17z^4}{120} + \dots\right). \end{aligned} \quad (17.3)$$

Thus the residue of $\sin^{-3}(z)$ at $z = 0$ is $1/2$, and the integral equals πi .

Exercise 84. Show that if f has a pole of order m at $z = z_i$, then

$$\text{Res}(f, z_i) = \frac{1}{(m-1)!} \left[\frac{d^{m-1}}{dz^{m-1}} \left((z - z_i)^m f(z) \right) \right]_{z=z_i} \quad (17.4)$$

by applying Laurent’s formula near $z = z_i$.

18 Integrals of trigonometric functions

Contour integration is very useful in calculating or estimating Fourier coefficients of periodic functions. Consider the integral

$$I = \int_0^{2\pi} \frac{\cos(nt)}{2 + \cos t} dt = \operatorname{Re}(J), \text{ where } J := \int_0^{2\pi} \frac{e^{int}}{2 + \cos t} dt. \quad (18.2)$$

Let $z = e^{it}$. Then

$$J = -i \oint_C \frac{z^{n-1}}{2 + (z + 1/z)/2} dz = -2i \oint_C \frac{z^n}{z^2 + 4z + 1} dz, \quad (18.3)$$

where C is the unit circle. The roots of $z^2 + 4z + 1$ are $-2 \pm \sqrt{3}$, and only one, $z_0 = -2 + \sqrt{3}$, lies inside the unit disk. Thus (check)

$$J = -2i \cdot 2\pi i \frac{z_0^n}{2z_0 + 4} \Rightarrow I = \frac{4\pi z_0^n}{2z_0 + 4} = \frac{2\pi z_0^n}{\sqrt{3}}. \quad (18.4)$$

19 Counting zeros and poles

Notations and definitions. (1) Assume f is analytic in a disk $\mathbb{D}_\rho(z_0)$ and $f(z_0) = 0$. Then, in $\mathbb{D}_\rho(z_0)$,

$$f(z) = \sum_{k=1}^{\infty} c_k (z - z_0)^k. \quad (19.2)$$

If $f \not\equiv 0$, there exists some k_0 such that $c_{k_0} \neq 0$ (see Proposition 8.1). The smallest such k_0 is called the **order** (or **multiplicity**) of the zero z_0 . For a meromorphic function g (see Definition 69), the order of a pole at z_0 is the multiplicity of the zero of $1/g$ at z_0 .

Exercise 85 (The zeros of an analytic function are isolated). Assume $f \not\equiv 0$ is analytic near z_0 and $f(z_0) = 0$. Use the Taylor series expansion to show that there is some disk around z_0 in which

$$f(z) = 0 \Rightarrow z = z_0.$$

Now assume f is meromorphic in $\mathcal{D}$, and let γ be a simple, piecewise C^1 closed curve contained in $\mathcal{D}$ together with its interior Γ . Since $\gamma \cup \Gamma$ is a closed subset of $\mathcal{D}$, by definition, the region of analyticity of f strictly contains Γ . For the purpose of the next theorem, the assumptions can be relaxed, allowing γ to be the boundary of the analyticity domain of f , provided suitable continuity conditions are imposed on f and f' .

Theorem 86: Counting zeros and poles

Let f be as above, and assume f has no zeros on γ . Let N be the total number of zeros of f in Γ , counting multiplicities, and let P be the total number of poles of f in Γ , each counted with its order. Then

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f'(s)}{f(s)} ds = N - P. \quad (19.3)$$

Proof. We leave some details to the reader. The function f'/f is meromorphic. At a zero of f of order n_i , f'/f has a simple pole with residue n_i ; at a pole of f of order p_i , f'/f has a simple pole with residue $-p_i$. The formula then follows immediately from the residue theorem (Proposition 17.7). ■

19.1 Constancy of the number of roots of $f(s) - \zeta$

Let γ and Γ be as before. Assume that f is analytic in $\mathcal{D} \supset \Gamma$.

Let V be the set of values of f attained in Γ but not on γ , that is,

$$V = \{ \zeta \in \mathbb{C} : \zeta \in f(\Gamma) \setminus f(\gamma) \}.$$

Theorem 87

Assume f is not constant and for each $\zeta \in V$, let $N(\zeta)$ be the number of roots of $f(z) - \zeta = 0$, counting multiplicity. Then $N(\zeta)$ is independent of ζ .

Hence, every value that f takes in Γ is attained the same finite number of times. This integer is called the **order** of f in Γ .

Proof. ■

For $\zeta \in V$ define

$$g(\zeta) = \frac{1}{2\pi i} \oint_{\gamma} \frac{(f(s) - \zeta)'}{f(s) - \zeta} ds = \frac{1}{2\pi i} \oint_{\gamma} \frac{f'(s)}{f(s) - \zeta} ds.$$

By Theorem 86, $g(\zeta)$ equals the number of zeros of $f(s) - \zeta$ in Γ (since there are no poles).

Moreover, g is continuous in $\zeta \in V$. But an integer-valued function is continuous on a connected set if and only if it is constant.

Remark 88. The integer $N(\zeta)$, independent of $\zeta \in V$, is what we called the *order* of f in Γ . In other contexts, especially in topology and the theory of Riemann surfaces, the same invariant is referred to as the *degree* of the holomorphic map f . It measures how many times f covers its image, counting multiplicity, and anticipates the notion of branched coverings.

Definition 90

The **winding number** of a point z_0 with respect to a closed, piecewise C^1 curve γ not passing through z_0 is defined by

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{dz}{z - z_0}. \quad (19.4)$$

Exercise 90. Show that the winding number defined above is always an integer. (Hint: see §19.1.)

Remark 91. Geometrically, the winding number measures how many times the curve γ winds around the point z_0 , counted with orientation. A positively oriented simple closed curve enclosing z_0 has winding number 1, while if z_0 lies outside γ the winding number is 0. More complicated curves may yield larger positive or negative integers depending on how many times they encircle z_0 and in which direction.

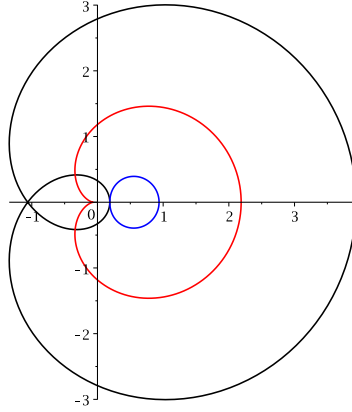


Figure 11: Images under $(z-1)(z-1/2)$ of circles centered at 0 with radii $1/4$ (innermost), $3/4$ (middle), and $5/4$ (outer).

19.2 Some further properties of the logarithm

If γ is a curve without self-intersections joining 1 to z and not passing through 0, then

$$\int_{\gamma} \frac{ds}{s} =: \log_{\gamma} z$$

defines an analytic function in a neighborhood of γ . We may homotopically deform γ so that it begins with the interval $[1, 1 + \varepsilon]$ for some small $\varepsilon > 0$. Since for positive real x we have $e^{\log x} = x$, analyticity in a neighborhood of γ allows us to apply the permanence of relations. It follows that $e^{\log_{\gamma} s} = s$ for s near γ , and in particular $e^{\log_{\gamma} z} = z$. If we take another curve without self-intersections, $\tilde{\gamma}$ joining 1 to z , the same argument yields $e^{\log_{\tilde{\gamma}} z} = z$. We thus have

$$\begin{aligned} e^{\log_{\gamma} z} = e^{\log_{\tilde{\gamma}} z} &\Rightarrow e^{\log_{\gamma} z - \log_{\tilde{\gamma}} z} = 1 \\ &\Rightarrow \log_{\gamma} z - \log_{\tilde{\gamma}} z = 2N\pi i \text{ for some } N \in \mathbb{Z} \end{aligned} \quad (19.5)$$

Thus, logarithms defined along different curves may differ only by an integer multiple of $2\pi i$. This multiple determines a particular *branch* of the log. Writing $z = \rho e^{i\varphi}$ and $\log z = u + iv$, we obtain $e^u e^{iv} = \rho e^{i\varphi}$, whence

$$\log z = \log |\rho| + i\psi,$$

where $\log |\rho|$ is the usual real log of a real positive number, and the angle $\psi = \varphi + 2N\pi$ for some $N \in \mathbb{Z}$ is the *argument* of z denoted by $\arg(z)$.

19.3 The argument principle

Consider again the setting of Theorem 86. Take $g = \ln f$ for any particular branch of the log. We have $g' = f'/f$ and then (19.3) shows that the change in $g = \ln f$ as we traverse positively γ is $N - P$. Since $\ln f = \ln |f| + 2\pi i \arg f$, and f is single-valued, $f(z)$ and $|f(z)|$ do not change as we traverse the curve. In other words,

$$\frac{1}{2\pi i} \oint_{\gamma} \frac{f'(s)}{f(s)} ds = \frac{1}{2\pi i} \oint_{f \circ \gamma} \frac{d\zeta}{\zeta} = N - P = \frac{1}{2\pi i} \Delta \log f = \frac{1}{2\pi i} \Delta \arg f. \quad (19.6)$$

Thus $2\pi i(N - P)$ is the **variation of the argument of f along γ** . Equivalently, it is the winding number of 0 with respect to $f \circ \gamma$ (see Definition 89), that is, the number of times — positive or negative — that $f \circ \gamma$ winds around 0.

Finally, this says that if we take the image of a parametrization of γ under f , then, $N - P$ counts the number of times the image $f(\gamma)$ turns around zero.

Exercise 92. Let $f(z) = \exp(1/z) - 1$. Clearly 0 is an essential singularity and Proposition 86 does not apply. Determine, as a function of $\varepsilon > 0$, the number of times the curve

$$\{f(\varepsilon e^{it}) : t \in [0, 2\pi)\}$$

winds around the origin.

19.4 Hurwitz's theorem

This theorem shows that a uniform limit of analytic functions with no zeros in a domain is either identically zero, or else an analytic function with no zeros in the domain:

Theorem 93: Hurwitz

Let $\{f_n\}_{n \in \mathbb{N}}$ be analytic and nonvanishing in a domain $\Omega \subset \mathbb{C}$, and assume that $\{f_n\}_{n \in \mathbb{N}}$ converges uniformly on compact sets to a function $f \not\equiv 0$. Then f has no zeros in Ω .

Proof. By standard results, the uniform convergence of $\{f_n\}$ on compacta implies that f is analytic. Consider the neighborhood of a point $z_0 \in \Omega$; by translation we may take $z_0 = 0$. Since $f \not\equiv 0$, there exists $\delta > 0$ such that $f(z) \neq 0$ for all $z \in \partial \mathbb{D}_\delta(0)$ (by Corollary 17).

Moreover, by Weierstrass's theorem (Theorem 48), the sequence $\{f'_n\}$ converges uniformly to f' . Hence $\{f'_n/f_n\}_{n \in \mathbb{N}}$ converges uniformly to f'/f on $\partial \mathbb{D}_{\delta/2}(0)$. Applying Theorem 86 to each f_n and passing to the limit shows that f has no zeros in $\mathbb{D}_{\delta/2}(0)$. Since z_0 was arbitrary, the conclusion holds throughout Ω . ■

19.5 Rouché's Theorem

Theorem 94: Rouché

Assume f and h are analytic in a domain $\mathcal{D}$ containing the piecewise differentiable simple closed curve γ and its interior Γ . Suppose that on γ we have $|h| < |f|$. Then the number of zeros of f and of $f + h$ in Γ is the same (counting multiplicities).

Note 95

Often h is viewed as a perturbation of f . The theorem is quantitative: it does not merely assert that a “small enough” perturbation preserves the number of zeros, but gives the explicit sufficient condition $|h| < |f|$ on γ . Whenever this inequality holds, f and $f + h$ have the same number of zeros inside Γ .

Proof. Since $|h| < |f|$ on γ , there is a neighborhood $\mathcal{N}$ of γ in which f has no zeros. In $\mathcal{N}$ we can write

$$f + h = f \cdot (1 + h/f) =: fQ, \tag{19.7}$$

where $Q = 1 + g$ with $g = h/f$. On $\mathcal{N}$ we have $|g| < 1$, so the series

$$q(z) = \log(1 + g(z)) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} g(z)^k$$

converges uniformly, and hence q is analytic in $\mathcal{N}$ by Theorem 48.

Differentiating gives

$$q' = \frac{g'}{1+g} = \frac{Q'}{Q}.$$

Thus

$$\frac{(fQ)'}{fQ} = \frac{f'}{f} + \frac{Q'}{Q}.$$

But since Q is analytic and nonvanishing in $\mathcal{N}$, we have

$$\oint_{\gamma} \frac{Q'}{Q} dz = 0.$$

Therefore

$$\oint_{\gamma} \frac{(f+h)'}{f+h} dz = \oint_{\gamma} \frac{f'}{f} dz,$$

and the result follows from Theorem 86. ■

Definition 96

A mapping $F : \mathcal{D} \rightarrow \mathbb{C}$ is called **open** if for every open set $\mathcal{O} \subset \mathcal{D}$, the image $F(\mathcal{O})$ is open in $\mathbb{C}$.

Theorem 97: Open Mapping Theorem

Let f be analytic and non-constant in a domain $\mathcal{D}$. Then f is an open mapping.

Proof. This is left as an exercise. (Hint: factor out the first nonzero term in the Taylor expansion of f at z_0 , and use the continuity of $\arg$ near 1.)

Namely, if $w_0 \in f(\mathcal{D})$, then there exists $z_0 \in \mathcal{D}$ such that $f(z_0) = w_0$. Consider the Taylor expansion

$$f(z) = w_0 + \sum_{k=1}^{\infty} f_k(z - z_0)^k.$$

Since f is non-constant, there is a smallest $k \geq 1$ with $f_k \neq 0$. Then we may factor

$$f(z) - w_0 = f_k(z - z_0)^k(1 + g(z)),$$

where g is analytic and $g(z) \rightarrow 0$ as $z \rightarrow z_0$. A key point is that $1 + g(z)$ remains close to 1 near z_0 , so $\log(1 + g(z))$ is analytic there and hence $\arg(1 + g(z)) = \text{Im}(\log(1 + g(z)))$ varies continuously (as does $\arg(z - z_0)^k$). From this, it follows that along every ray from w_0 there is a nearby point in $f(\mathcal{D})$ (check!). Finally, continuity of f implies that $f(\mathcal{D})$ contains a neighborhood of w_0 , proving that f is open. ■

20 The inverse function theorem

Theorem 98

Assume f is analytic at z_0 and $f'(z_0) = a \neq 0$. Then there exists a disk $\mathbb{D}_\varepsilon(z_0)$ such that f is injective on $\mathbb{D}_\varepsilon(z_0)$ and maps it biholomorphically onto the open set $f(\mathbb{D}_\varepsilon(z_0))$; in particular, the inverse is analytic.

In this situation we also say that f is a *conformal map* between $\mathbb{D}_\varepsilon(z_0)$ and $f(\mathbb{D}_\varepsilon(z_0))$. (We will return to this notion in more detail later.)

Proof. Without loss of generality, assume $z_0 = 0$ and $f(0) = 0$. Then

$$f(z) = az + z^2g(z),$$

where g is analytic near 0. For z, z_1 near 0,

$$\begin{aligned} f(z) - f(z_1) &= a(z - z_1) + (z^2g(z) - z_1^2g(z_1)) \\ &= a(z - z_1) + (z - z_1)(z + z_1)g(z) + z_1^2(g(z) - g(z_1)) \\ &= (z - z_1) \left[a + (z + z_1)g(z) + z_1^2 \frac{g(z) - g(z_1)}{z - z_1} \right]. \end{aligned} \quad (20.2)$$

By analyticity of g , for z, z_1 in a small disk $\mathbb{D}_\rho(0)$,

$$\frac{g(z) - g(z_1)}{z - z_1} = \int_0^1 g'(z_1 + t(z - z_1)) dt,$$

so the bracket in (20.2) has a removable singularity at z_1 . In particular, it extends continuously and tends to a as $z, z_1 \rightarrow 0$. Hence we can choose $\rho > 0$ so small that

$$\left| (z + z_1)g(z) + z_1^2 \frac{g(z) - g(z_1)}{z - z_1} \right| < |a| \quad \text{for all } z, z_1 \in \mathbb{D}_\rho(0).$$

Then the bracket in (20.2) never vanishes on $\mathbb{D}_\rho(0) \times \mathbb{D}_\rho(0)$, so $f(z) = f(z_1)$ implies $z = z_1$. Therefore f is injective on $\mathbb{D}_\rho(0)$.

Now fix $w_0 = f(z_0)$ with $z_0 \in \mathbb{D}_\rho(0)$. From (20.2) we have, as $z \rightarrow z_0$,

$$\frac{f(z) - f(z_0)}{z - z_0} \rightarrow a \neq 0.$$

Thus

$$\lim_{z \rightarrow z_0} \frac{z - z_0}{f(z) - f(z_0)} = \frac{1}{a},$$

which shows that f^{-1} is differentiable at w_0 with derivative $1/f'(z_0)$. Since z_0 was arbitrary, f^{-1} is analytic on $f(\mathbb{D}_\rho(0))$. ■

Exercise 99 (Important generalization). (a) Assume f is analytic in a neighborhood of z_0 (say $z_0 = 0$) and

$$f(0) = f'(0) = \cdots = f^{(k-1)}(0) = 0, \quad f^{(k)}(0) = a \neq 0.$$

Show that there exists $\varepsilon > 0$ such that, for every $t \in \mathbb{D}_\varepsilon(0)$, the equation $f(z) = t$ has *exactly* k roots near 0 (counted with multiplicity).

Hint: factor $f(z) = a(z^k + \text{higher order terms})$ and argue as in the proof of Theorem 98, or apply Rouché's theorem.

(b) Deduce from (a) an alternative proof of the open mapping theorem.

21 Analytic continuation

Assume f is analytic in $\mathcal{D}$ and f_1 is analytic in $\mathcal{D}_1 \supsetneq \mathcal{D}$ with $f = f_1$ in $\mathcal{D}$. Then f_1 is called an *analytic extension* of f . We also say that f_1 is obtained from f by *analytic continuation*.

Weierstrass's point of view was to regard analytic functions as properly defined *chains of Taylor series*, up to a natural equivalence (see below), each Taylor series being the analytic continuation of the adjacent ones.

If f is analytic at z_0 , then there exists $\varepsilon > 0$ such that f is given by its Taylor series in $\mathbb{D}_\varepsilon(z_0)$. Let ε_0 be the maximal such radius. If $z_1 \in \mathbb{D}_\varepsilon(z_0)$, then f is also analytic at z_1 , hence represented by a Taylor series centered at z_1 . The radius of convergence of this series is at least $d(z_1, \partial\mathbb{D}_\varepsilon(z_0))$, but possibly larger. (Example: the function $1/(1+z)$ with disks centered at $z = 0$ and $z = 1/2$.)

In the latter case, we obtain a function f_1 , piecewise given by the two Taylor series, which is analytic in $\mathbb{D}_\varepsilon(z_0) \cup \mathbb{D}_{\varepsilon_1}(z_1)$. Continuing in this way, we construct chains $z_0, z_1, \dots$ with corresponding disks $\mathbb{D}_{\varepsilon_i}(z_i)$. There is then a natural equivalence relation: f_α and f_β are equivalent if they are analytic continuations of each other.

Definition 100: Global analytic functions (Weierstrass)

An analytic function in the sense of Weierstrass is an equivalence class under this relation.

Uniqueness. If there is an analytic continuation of f to $\mathbb{D}_\varepsilon(z_0) \cup \mathbb{D}_{\varepsilon_1}(z_1)$, then it is unique (use Proposition 8.1).

Ahlfors makes this more precise as follows.

Definition 101: Function elements and global analytic functions

An *analytic function* defined on an open set $U \subset \mathbb{C}$ is called a *function element* and denoted (f, U) . Two function elements (f_1, U_1) and (f_2, U_2) are said to be *analytic continuations* of one another if $U_1 \cap U_2 \neq \emptyset$ and $f_1 = f_2$ on $U_1 \cap U_2$.

A *chain of analytic continuations* is a finite sequence of function elements

$$(f_1, U_1), (f_2, U_2), \dots, (f_n, U_n)$$

such that each consecutive pair are analytic continuations of one another.

A *global analytic function* is a family $\mathcal{F}$ of function elements such that, for any $(f, U), (g, V) \in \mathcal{F}$, there exists a chain of analytic continuations in $\mathcal{F}$ beginning at (f, U) and ending at (g, V) .

A *complete global analytic function* is a global analytic function $\mathcal{F}$ that contains every analytic continuation of each of its elements.

But the *values of the continuations* may differ. For instance, starting from (f, U_0) with $f = \log$ defined near 1, we can build two chains of analytic continuations, one passing along the upper semicircle from 1 to -1 , the other along the lower semicircle. Both chains end in function elements defined on the same domain U near -1 , yet the resulting functions differ by $2\pi i$. Thus the same open set U may occur in two different function elements with distinct values, reflecting the multivalued nature of functions like $\log$.

This phenomenon leads naturally to the notion of *Riemann surfaces*, which provide the correct setting for treating multivalued analytic functions (such as $\log z$, $z^{1/n}$, or $\arcsin z$) as genuinely single-valued. We will return to this perspective later.

Cuts. One way to restore single-valuedness is to define $\log$ in $\mathbb{C} \setminus (-\infty, 0]$ where $(-\infty, 0]$ is called a **cut** and its function is to prevent curves to circle around 0.

We note that the cut $(-\infty, 0]$ is **highly arbitrary**. It can be replaced by any ray $\{te^{i\varphi} : \varphi \in [0, 2\pi)\}$ or any curve that has the same functionality, namely to ensure there are no paths of continuation that encircle 0. This is worth emphasizing this, because many books tend to give the wrong impression that the cut $(-\infty, 0]$, giving the "principal branch of the log", is somehow set in stone. Notably too, the **value of the log depends on the cut**.

Exercise 102. What is the value of $\log e^{3\pi i/4}$ if the cut is $(-\infty, 0]$ and what is it if the cut is $\{it : t \geq 0\}$?

Remark. The choice of cut is a way to make a multivalued function (like $\log z$) single-valued on a simply connected domain. Different choices of cuts give different branches of the function. From the perspective of Riemann surfaces, these branches are just different "sheets" of a single global object. We will return to this viewpoint later when discussing Riemann surfaces and branch points.

Natural boundaries. For general analytic functions, it might happen that one encounters lines, or more generally closed regions, through which no analytic continuation exists.

Such a boundary is called a *natural boundary*. It represents a set of "terminus" points for a global analytic function in the sense of Weierstrass. Yet, *some* natural boundaries can be traversed in a precise mathematical sense. This idea was pioneered by Borel (see Borel, *Leçons sur les fonctions monogènes*), who developed methods of analytic continuation through natural boundaries. We will return to this question later. Separately, Borel also introduced methods of resumming divergent (zero radius of convergence) series, which later evolved into the modern theory of Borel summation and resurgent analysis.

The standard example of such a function is

$$f(z) = \sum_{k=0}^{\infty} z^{2^k}.$$

We have $f(z) \rightarrow \infty$ as $z \rightarrow 1$, and also as $z \rightarrow -1, i$, and $-i$, and more generally as $z \rightarrow e^{2\pi i N/2^M}$ with $(N, M) \in \mathbb{N}^2$. The points $N/2^M$ with $(N, M) \in \mathbb{N}^2$ are the *binary rationals*, i.e. the numbers whose base-2 expansion has only finitely many digits. They form a dense subset of $\mathbb{R}$, hence the points $e^{2\pi i N/2^M}$ are dense on the unit circle. This shows that **no point** of the unit circle is a point of analyticity (why?).

Exercise 103 (Another example of natural boundary). Consider the lacunary series

$$f(z) = \sum_{k=0}^{\infty} z^{p_k}.$$

Show that ∂D is a natural boundary for f .

Hint: For each n , the $n!$ -th roots of unity are singular points of f . Observe that these roots are dense on ∂D , and argue as in the case of $\sum z^{2^k}$ that no boundary point can be a point of analyticity.

This is a special case of a lacunary series: if the sequence of natural numbers p_k has the property that $\liminf_{n \rightarrow \infty} p_{n+1}/p_n = 1 + \delta$, then the series $\sum_{k=0}^{\infty} c_k z^{p_k}$ is called **lacunary**. There is a whole literature about them. The following theorem is due to Hadamard, with a proof in [10]:

Theorem 104: Hadamard–Mandelbrojt

If the exponents p_n satisfy $\liminf_{n \rightarrow \infty} p_{n+1}/p_n > 1$, and $\sum_{n=0}^{\infty} c_n z^{p_n}$ has radius of convergence 1, then $\partial\mathbb{D}$ is a natural boundary for f .

Hadamard proved this under the stronger hypothesis $p_{n+1}/p_n > 1 + 1/n$, and Mandelbrojt later weakened it to the general form above.

Remark 105. Mandelbrojt’s condition is essentially sharp.

If $\liminf_{n \rightarrow \infty} \frac{p_{n+1}}{p_n} > 1$, then for *any* coefficients c_n , the unit circle is a natural boundary of f . On the other hand, if $\liminf_{n \rightarrow \infty} \frac{p_{n+1}}{p_n} = 1$, then there exist lacunary series of radius of convergence 1 which admit analytic continuation across parts of $\partial\mathbb{D}$. Thus Mandelbrojt’s criterion gives a necessary and sufficient condition for natural boundaries in this class. (see [19]).

Remark 106 (Genericity of natural boundaries). The examples above (such as $\sum z^{2^n}$ or $\sum z^{n!}$) show that highly sparse (“lacunary”) series typically have the unit circle as a natural boundary. Even more striking is the fact that natural boundaries are in fact *generic*: if one considers power series

$$f(z) = \sum_{k=0}^{\infty} c_k z^k, \quad c_k \in \{0, 1\},$$

and interprets the sequence $\{c_k\}_k$ as the binary expansion of a number in $[0, 1]$, then for a set of full measure (with respect to Lebesgue measure on $[0, 1]$), the function f has the unit circle as a natural boundary.

This phenomenon was first observed in work of Steinhaus (1929), and later developed in the probabilistic theory of random power series (see Kahane, *Some Random Series of Functions*, 2nd ed., 1985). In this sense, having the unit circle as a natural boundary is the “typical” behavior for analytic functions defined by such series.

As we shall see later, natural boundaries occur in the *uniformization of nontrivial Riemann surfaces*. For example, let $\mathcal{R}$ be the set of curves in $\mathbb{C} \setminus \{0, 1\}$ modulo homotopies. A uniformization ψ is a bi-analytic bijection between $\mathcal{R}$ and $\mathbb{D}$. Then $\partial\mathbb{D}$ is a natural boundary for ψ^{-1} .

Exercise 107 (Example of natural boundary due to Poincaré). Consider the rational numbers $r = p/q$ (with p, q relatively prime) and associate to r $N_{pq} = 7^{|p|} 5^{|q|}$ (other primes would work too) (check that this is injective as a function from $\mathbb{Q}^+$ to $\mathbb{N}$). Similarly, find an isomorphism between $\mathbb{Q}^-$ and $\mathbb{Z}^-$. Define

$$f(z) = \sum_{N_{pq}}^{\infty} \frac{2^{-N_{pq}}}{z - p/q}. \quad (21.2)$$

Show that the series converges for $z \in \mathbb{C} \setminus \mathbb{R}$ and that $\mathbb{R}$ is a natural boundary for f (the poles accumulate densely on $\mathbb{R}$).

How can this example be modified to obtain an analytic function f in any domain bounded by a simple closed curve γ , such that γ is a natural boundary of f ?

22 The Schwarz reflection principle

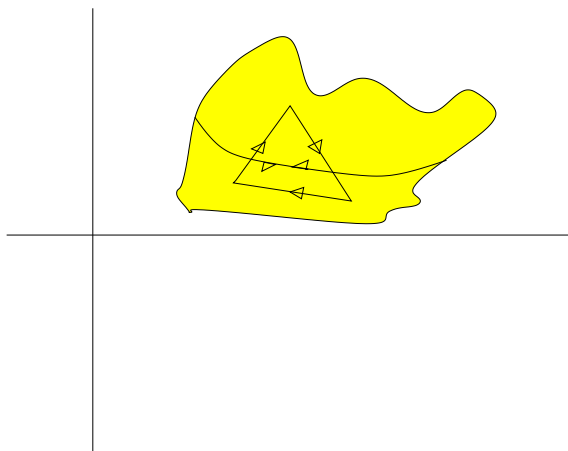


Figure 12: Analytic continuation through a line along which a function is continuous.

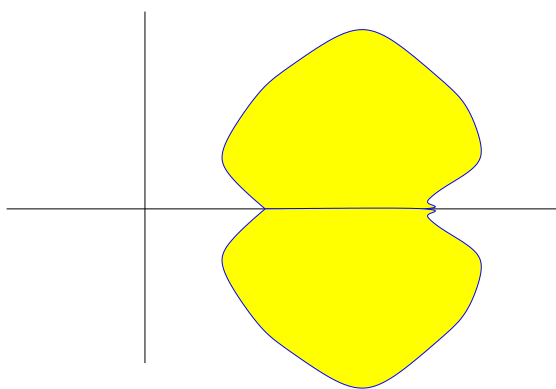


Figure 13: Schwarz's reflection principle.

Theorem 108

Assume f is analytic in two domains $\mathcal{D}_1$ and $\mathcal{D}_2$ which share a common boundary portion given by a piecewise differentiable curve γ . Suppose further that f is continuous on $\mathcal{D}_1 \cup \mathcal{D}_2 \cup \gamma$. Then f is analytic in

$$\mathcal{D} = \mathcal{D}_1 \cup \mathcal{D}_2.$$

Proof. By Morera's theorem, it suffices to show that the integral of f around any triangular contour in $\mathcal{D}$ vanishes. If the triangle lies entirely inside $\mathcal{D}_1$ or $\mathcal{D}_2$, this follows immediately from analyticity there. If the triangle intersects γ , split it along γ into two smaller triangles (see Fig. 14). Each piece lies in $\mathcal{D}_1$ or $\mathcal{D}_2$, so their integrals vanish. The contributions along γ cancel due to continuity of f , giving zero for the whole integral. ■

If all is known about a function f is analyticity in $\mathcal{D}_1$, and the fact that a piece of the boundary is an *analytic curve* (meaning, there is a parametrization by a function $\gamma : [0,1] \rightarrow \mathbb{C}$ which extends analytically in a complex neighborhood of $[0,1]$) up to which f is continuous, and a further condition on the values of f on γ holds, then it is possible to extend f analytically past γ . We start with the simplest such case.

Theorem 109: The Schwarz reflection principle

Assume f is analytic in a domain $\mathcal{D}$ in the upper half plane (UHP, also denoted in these

notes by $\mathbb{H}^u$) whose boundary contains an interval $I \subset \mathbb{R}$ and assume f is continuous on $\mathcal{D} \cup I$ and real valued on I . Then f has analytic continuation across I , in a domain $\mathcal{D} \cup \mathcal{D}^* \cup I$ where $\mathcal{D}^* = \{\bar{z} : z \in \mathcal{D}\}$.

Note 110

see §30.3 for a generalization of this result.

Proof. Define

$$F(z) = \begin{cases} f(z), & z \in \mathcal{D} \cup I, \\ \overline{f(\bar{z})}, & z \in \mathcal{D}^* \cup I. \end{cases}$$

Then F is continuous on $\mathcal{D} \cup I \cup \mathcal{D}^*$ (check). It is analytic in $\mathcal{D}^*$ by a Taylor expansion argument. For Morera's theorem: integrals over closed curves lying entirely in $\mathcal{D}$ or in $\mathcal{D}^*$ vanish. If a curve crosses I , it can be split into two parts symmetric across I , which cancel because f is real on I (see again Fig. 14). Hence F is analytic in $\mathcal{D} \cup \mathcal{D}^*$, as desired. ■

Example 111. The square root $\sqrt{z} = \rho^{1/2}e^{i\varphi/2}$, $z = \rho e^{i\varphi}$ with $\varphi \in [0, \pi)$, is analytic in the upper half-plane, continuous down to $[0, \infty)$, and real there. By Schwarz reflection, it continues analytically to the lower half-plane. For $z = \rho e^{-i\theta} \in \mathbb{H}_\ell$, $\theta \in (0, \pi)$, we have $\bar{z} = \rho e^{i\theta} \in \mathbb{H}^u$, so

$$\sqrt{z} = \overline{\sqrt{\bar{z}}} = \rho^{1/2}e^{-i\theta/2}.$$

23 Multi-valued functions

As we discussed, as a result of analytic continuation in the complex plane we may get a *global analytic function* which is not necessarily a function **defined on** $\mathbb{C}$ since the definition is path-dependent; the function is thus defined on a space of paths or curves, modulo homotopies.

As long as the domain of continuation is simply connected, we still get a function in the usual sense:

Theorem 112: The monodromy theorem

Assume $\mathcal{D}$ is a simply connected domain, $\mathbb{D}_\varepsilon(z_0) \subset \mathcal{D}$ that f is analytic in $\mathbb{D}_\varepsilon(z_0)$. Assume further that for any $z_1 \in \mathcal{D}$ there is a curve γ joining z_0 to z_1 such that f has analytic continuation along γ ^a. Then, f has analytic continuation in $\mathcal{D}$.

^aMore precisely, since γ is compact, γ is contained in a finite union of open disks in $\mathcal{D}$ such that $U_1 = \mathbb{D}_\varepsilon(z_0), \dots, U_n = \mathbb{D}_\delta(z_1) \dots$ such that $(f, U_1), (f_2, U_2), \dots, (f_n, U_n)$ form a chain of analytic continuations.

Proof. Take $z_1 \in \mathcal{D}$ and let $\mathcal{D}' \subset \mathcal{D}$ be a simply connected subdomain containing $\{z_0, z_1\}$ whose boundary decomposes as $\gamma_1 \cup \gamma_2$, where γ_1 and γ_2 are smooth curves with common endpoints.

Choose a curve γ_3 sufficiently close to γ_1 so that: (a) γ_3 can be covered by disks $U'_1 = U_1 = \mathbb{D}_\varepsilon(z_0), \dots, U'_k = U_n = \mathbb{D}_\delta(z_1)$; (b) the domain $\mathcal{D}_3 = U_1 \cup U_2 \cup \dots \cup U_n$ is connected and simply connected; and (c) $(f, U_1), (g_2, U_2), \dots, (g_k, U_k)$ forms a chain of analytic continuations.

Thus f is analytic in each U_j , and by construction the restrictions agree on overlaps. By uniqueness of analytic continuation (Proposition 8.1), the local definitions glue together to a single analytic function on $\mathcal{D}_3$.

Repeating this deformation step, one can deform γ_1 continuously into γ_2 without losing analyticity. Hence f admits analytic continuation in $\mathcal{D}'$. Since $\mathcal{D}'$ was arbitrary, f extends analytically to the whole of $\mathcal{D}$. ■

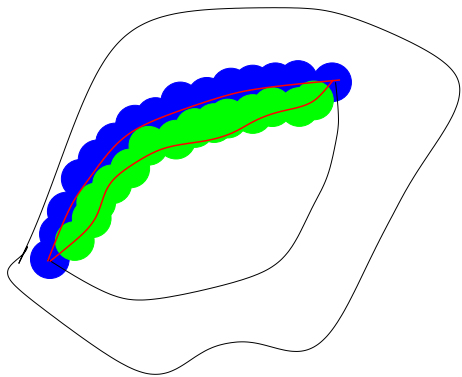


Figure 14: The monodromy theorem.

In a nutshell: if f admits analytic continuation along every path in a simply connected domain $\mathcal{D}$, then f extends to an analytic function on $\mathcal{D}$. Equivalently, for analyticity to fail in $\mathcal{D}$, there must exist singular points of f within $\mathcal{D}$ that obstruct continuation.

Exercise 113 (Alternative proof of monodromy). Let $\overline{\mathcal{D}_1} \subsetneq \mathcal{D}$ be a connected, simply connected compact domain with smooth boundary that contains z_0 . Cover $\overline{\mathcal{D}_1}$ by finitely many disks $\{U_j\}_{j=1}^N$, with $U_1 = \mathbb{D}_\rho(z_0)$, arranged so that consecutive disks intersect.

Since analytic continuation exists along any polygonal arc contained in $\overline{\mathcal{D}_1}$ that joins adjacent disks, define F on $\overline{\mathcal{D}_1}$ by continuing f from U_1 to any $p \in \overline{\mathcal{D}_1}$ along such a polygonal chain of overlapping disks.

Show that F is *well-defined* (path-independent): if two polygonal chains join z_0 to p , together they form a closed loop in $\overline{\mathcal{D}_1}$; since $\overline{\mathcal{D}_1}$ is simply connected, the loop is null-homotopic. By uniqueness of analytic continuation on overlaps, the two continuations agree. Conclude that F is single-valued and analytic on $\mathcal{D}_1$.

Because $\overline{\mathcal{D}_1} \subsetneq \mathcal{D}$ was arbitrary, exhaust $\mathcal{D}$ by an increasing sequence of such compact, simply connected domains $\overline{\mathcal{D}_k}$ with $\bigcup_{k \in \mathbb{N}} \mathcal{D}_k = \mathcal{D}$. Construct a sequence $(\mathcal{D}_k, f_k)_{k \in \mathbb{N}}$ where f_k is the analytic continuation on $\mathcal{D}_k$ and $f_k|_{\mathcal{D}_{k-1}} = f_{k-1}$. Then $f := \bigcup_{k \in \mathbb{N}} f_k$ defines the desired analytic continuation on all of $\mathcal{D}$.

23.1 Generalization: log of a function

If g is a function defined in a region in $\mathbb{C}$, we can define $\ln g$ by

$$\ln g(z) = \int_a^z \frac{g'(s)}{g(s)} ds. \quad (23.2)$$

Now, depending on the properties of g , the homotopy classes will in general be more complicated than those of $\log z$.

If, for instance, g is a meromorphic function in $\mathbb{C}$, then the set S consisting of all the zeros and poles of g are points where the integral is undefined. The logarithm of g is defined on homotopy classes of curves over $\mathbb{C} \setminus S$.

It is sometimes convenient to define a branch of $\ln g$ by cutting the plane along rays originating at the points in S . (Check that such cuts exist for any meromorphic g —hint: join each singularity to infinity with non-intersecting rays.) See also Fig. 15.

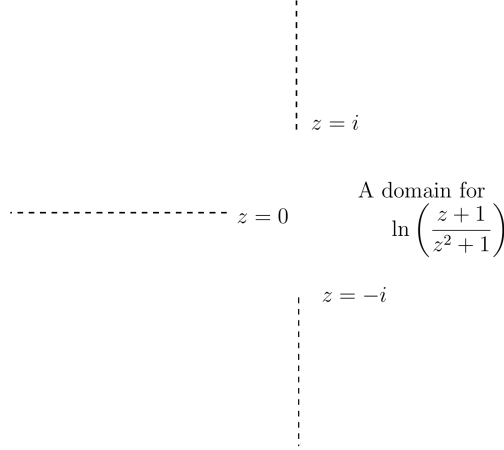


Figure 15: Cuts defining the log of a rational function.

23.2 General powers of z

Once we have defined the log, it is natural to define

$$z^\alpha = e^{\alpha \ln z} \quad (23.3)$$

Since $\ln z$ is defined on the homotopy classes of curves over $\mathbb{C} \setminus \{0\}$, so is z^α . For a general $\alpha \in \mathbb{C}$, the multivaluedness of z^α is inherited from the multivaluedness of the log: $e^{\alpha(\ln z + 2N\pi i)} = e^{\alpha \ln z} e^{2N\pi i \alpha}$. Note however that if $\alpha \in \mathbb{Z}$ then the value does not depend on the homotopy class and the definition (23.3) defines a function in $\mathbb{C} \setminus \{0\}$, with a pole at zero if $\alpha < 0$ and a removable singularity if $\alpha \geq 0$. Check that this definition coincides with the usual power on $\mathbb{R}^+$.

If $\alpha = p/q$ with p, q relatively prime, then $z^{p/q}$ admits exactly q distinct values corresponding to the q -th roots of unity.

Note. Beware of possible pitfalls.

$$e^{\ln z_1 + \ln z_2} = e^{\ln z_1} e^{\ln z_2} = z_1 z_2 \quad (23.4)$$

However, this does not mean $\ln z_1 + \ln z_2 = \ln z_1 z_2$, but just that

$$\ln z_1 + \ln z_2 = \ln z_1 z_2 + 2N\pi i, N \in \mathbb{Z}. \quad (23.5)$$

For the same reason, $z^{\alpha_1} z^{\alpha_2}$ is not necessarily $z^{\alpha_1 + \alpha_2}$. Note the fallacious calculation⁵

$$1 = \sqrt{1} = \sqrt{(-1)(-1)} = \sqrt{-1}\sqrt{-1} = i \cdot i = -1 \quad (!) \quad (23.6)$$

24 Conformal Mapping

Laplace's equation in two dimensions,

$$\Delta f = f_{xx} + f_{yy} = 0, \quad (24.2)$$

arises in many physical problems.

⁵This calculation was done as presented by an early version of a computer algebra program.

In electrostatics, for instance, the equation $\Delta V = 0$ describes the potential V in a charge-free region, following from two of Maxwell's equations:

$$\nabla \cdot \mathbf{E} = \varepsilon_0^{-1} \rho = 0, \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} = 0,$$

which imply $\mathbf{E} = -\nabla V$ and $\Delta V = 0$. Boundary conditions are imposed by the distribution of charges, and the physical problem becomes: solve (24.2) with V prescribed on $\partial\mathcal{D}$ (the *Dirichlet problem*).

In fluid dynamics, consider a two-dimensional incompressible velocity field $\mathbf{V} = (v, u)$. Incompressibility gives

$$\operatorname{div}(v, u) = v_x + u_y = 0, \quad (24.3)$$

while irrotationality implies

$$\nabla \times \mathbf{V} = 0 \Rightarrow u_x - v_y = 0. \quad (24.4)$$

Equations (24.3)–(24.4) show that u and v are harmonic conjugates. Thus, in any simply connected domain, there exists a *velocity potential* φ with $\mathbf{V} = \nabla\varphi$ and $\Delta\varphi = 0$. Its harmonic conjugate ψ is called the *stream function*, and the curves $\psi = \text{const.}$ coincide with the streamlines of the flow. Physically, the natural boundary condition is impermeability of the boundary $\partial\mathcal{D}$:

$$(v, u) \cdot (n_1, n_2) = 0,$$

the so-called *no-penetration condition*. Laplace's equation with prescribed normal derivative on $\partial\mathcal{D}$ is known as the *Neumann problem*.

We already know the explicit solution of the Dirichlet problem for the special case $\mathcal{D}$ a disk: Poisson's formula (11.5). More generally, existence and uniqueness of solutions hold for any bounded domain $\mathcal{D}$ with sufficiently smooth boundary and continuous boundary data.

24.1 Uniqueness

Uniqueness follows from the maximum principle: if u_1, u_2 are two solutions of (24.2) with the same boundary conditions, then $u = u_1 - u_2$ satisfies (24.2) with $u = 0$ on $\partial\mathcal{D}$. But a harmonic function attains both its maximum and its minimum on the boundary; hence $u \equiv 0$. Similarly, in the Neumann problem, u is determined uniquely up to an arbitrary additive constant.

Often, a PDE possesses a symmetry group, which can be exploited both for solving the equation and for understanding its qualitative behavior. For Laplace's equation in $\mathbb{C}$, the symmetry group is especially large: it is invariant under conformal maps (analytic homeomorphisms). More precisely, if $f : \mathcal{D}_1 \rightarrow \mathcal{D}_2$ is a biholomorphism between simply connected domains, then solutions in $\mathcal{D}_2$ pull back to solutions in $\mathcal{D}_1$. In this sense, Laplace's equation is said to be *conformally invariant*.

Proposition 24.8. *If u solves (24.2) ($\Delta u = 0$) in a simply connected domain $\mathcal{D}_2$, and $f = f_1 + if_2 : \mathcal{D}_1 \rightarrow \mathcal{D}_2$ is analytic, then*

$$(x, y) \mapsto u(f_1(x, y), f_2(x, y))$$

is a solution of (24.2) in $\mathcal{D}_1$.

Proof. Since u is harmonic in $\mathcal{D}_2$, it has a harmonic conjugate v (unique up to an additive constant). Thus $g = u + iv$ is analytic in $\mathcal{D}_2$. Then $g \circ f$ is analytic in $\mathcal{D}_1$, and hence its real and imaginary parts, $u \circ f$ and $v \circ f$, satisfy the Cauchy–Riemann equations in $\mathcal{D}_1$. It follows that $u \circ f$ is harmonic in $\mathcal{D}_1$. ■

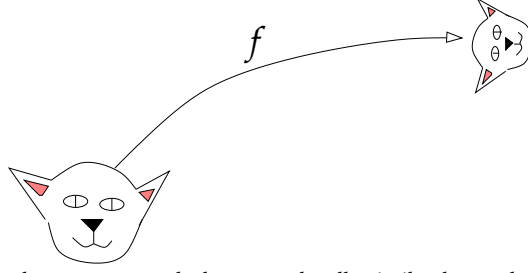


Figure 16: Conformal kitty. Angles are preserved, shapes are locally similar, but scales and orientations are changed.

We will be especially interested in analytic *homeomorphisms*, which preserve the topological structure of regions. Two regions related by such a map are said to be *conformally equivalent*.

The celebrated **Riemann mapping theorem** (to be proved later) states that any simply connected domain other than $\mathbb{C}$ itself is conformally equivalent to the unit disk. In this correspondence, the boundary of the region is mapped onto the unit circle. Thus

Remark 114. *The orbit of the unit disk under all conformal transformations contains every simply connected region (except $\mathbb{C}$ itself).*

Because of this invariance, the conformal group is powerful enough that, in principle, Laplace's equation in any simply connected domain (with sufficiently smooth boundary, say $C^{1,\alpha}$) can be reduced to the disk case. This is one of the key motivations for studying conformal mappings carefully.

24.2 Preservation of angles and small shapes: heuristics

Let f be analytic at z_0 with $f'(z_0) = a \neq 0$. (Without loss of generality we may assume $z_0 = 0$ and $f(0) = 0$.) Consider a tiny neighborhood $\mathcal{N}$ of 0. If $z_\beta \in \mathcal{N}$ then

$$f(z_\beta) \approx az_\beta. \quad (24.5)$$

Thus all points near 0 are multiplied by the *same* complex number a . Multiplication by a rescales vectors by $|a|$ and rotates them by $\arg(a)$. If we think of the collection $\{z_\beta\}$ as forming a small figure, then $\{f(z_\beta)\}$ is essentially the same figure, uniformly rotated and rescaled. In particular, *angles are preserved*, and the local shape is unchanged up to similarity. This is precisely what it means for f to be *conformal*.

Since a tiny square of side ε is mapped to a square of side $|a|\varepsilon$, areas are multiplied by the factor $|a|^2$.

We make this rigorous in what follows.

24.3 Preservation of angles

Recall that additivity of arguments holds up to an integer multiple of 2π . In the calculations below, this ambiguity will be immaterial.

Assume f is analytic in a disk D , that $z_0 \in D$, and that $f'(z_0) \neq 0$. Let $\gamma(t)$ and $\Gamma(t)$ be two smooth curves crossing at z_0 , and without loss of generality assume $\gamma(0) = \Gamma(0) = z_0$. The angle between the curves is defined as the difference of the arguments of their tangent vectors:

$$\theta = \arg \gamma'(0) - \arg \Gamma'(0),$$

provided these derivatives are nonzero.

Now consider the images of these curves under f . The corresponding tangent vectors at $f(z_0)$ are

$$(f \circ \gamma)'(0) = f'(z_0)\gamma'(0), \quad (f \circ \Gamma)'(0) = f'(z_0)\Gamma'(0).$$

Hence the angle between the images is

$$\begin{aligned} \arg[(f \circ \gamma)'(0)] - \arg[(f \circ \Gamma)'(0)] &= \arg(f'(z_0)\gamma'(0)) - \arg(f'(z_0)\Gamma'(0)) \\ &= \arg f'(z_0) + \arg \gamma'(0) - (\arg f'(z_0) + \arg \Gamma'(0)) \\ &= \arg \gamma'(0) - \arg \Gamma'(0) = \theta. \end{aligned} \quad (24.6)$$

Thus the image of two curves intersecting at an angle θ is a pair of curves intersecting at the same angle θ . In other words, analytic maps with $f'(z_0) \neq 0$ preserve angles (and orientation) at z_0 . Equivalently, a sufficiently small domain is transformed into a similar one — rescaled and rotated, but not distorted.

Exercise 115. More precisely, consider a sequence of similar triangles whose sides go to zero and with a common point z_0 , with $f'(z_0) \neq 0$. Consider their images through f . For each triangle, rescale both the triangle by the size L of one of its sides, and rescale their images through f by $L|f'(z_0)|$. Show that in the limit of vanishing sides, the limit of the rescaled triangles and the limit of the rescaled images of these triangles are congruent. Hence, applying a similar procedure to polygons, we will obtain congruent polygons. The same extends to shapes that are limits of polygons.

24.4 Rescaling of arc length

The arc length along a curve $\gamma(t)$ is given by

$$L(\gamma) = \int_{\gamma} d|z| := \int_a^b |\gamma'(t)| dt. \quad (24.7)$$

If f is analytic, then

$$L(f(\gamma)) = \int_a^b |(f \circ \gamma)'(t)| dt = \int_a^b |f'(\gamma(t))| |\gamma'(t)| dt = \int_{\gamma} |f'(z)| d|z|. \quad (24.8)$$

Thus the arc length is locally stretched by the factor $|f'(z)|$, in agreement with Exercise 115.

24.5 Transformation of areas

The area of a domain A is its Lebesgue measure,

$$\iint_A dx dy. \quad (24.9)$$

Let $f = u + iv : B \rightarrow A$ be biholomorphic. Under the change of variables $(x, y) \mapsto (u(x, y), v(x, y))$, the area of A can be computed as

$$\iint_B |J(x, y)| dx dy, \quad (24.10)$$

where J is the Jacobian determinant

$$J(x, y) = \det \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix}.$$

Using the Cauchy–Riemann equations (see (9.3)) one finds that $J(x, y) = |f'(z)|^2$.

Geometrically, this means that areas are locally stretched by the factor $|f'(z)|^2$, just as lengths are stretched by $|f'(z)|$ (cf. Exercise 115).

Note 116

It is interesting to observe that if a smooth transformation (u, v) of the plane preserves angles, then $u + iv$ must be analytic (i.e., conformal). Likewise, if it rescales every infinitesimal figure by the same factor, then the map is either analytic, or it coincides with the complex conjugate of an analytic function.^a This is not difficult to prove; see [3, §II.3], p. 74.

Thus analytic functions can be characterized geometrically as those maps which are *locally Euclidean*: they transform small shapes into similar shapes, preserving angles (orientation-preserving in the analytic case, orientation-reversing in the conjugate case).

^aIn this case the transformation is *orientation-reversing*. Geometrically, this means that locally the map looks like $z \mapsto \overline{f(z)}$ for some analytic f .

Note 117

Observe that we did not require f to be globally one-to-one. The simple fact that f is analytic with nonzero derivative makes it conformal.

Remark 118 (On bijectivity). In $\mathbb{R}$, it suffices that $f' \neq 0$ everywhere for f to be conformal. This is **not** the case in $\mathbb{C}$. For instance, $\exp'(z) = \exp(z) \neq 0$ for all $z \in \mathbb{C}$, yet the exponential function is not injective: $\exp$ wraps the plane around the punctured plane infinitely many times.

24.6 Puiseux series

Proposition 24.9 (Ramified expansions: Puiseux series). Assume that $f : \mathcal{D}_1 \rightarrow \mathcal{D}_2$ is analytic, and that for some $z_0 \in \mathcal{D}_1$ we have

$$f^{(j)}(z_0) = 0 \quad \text{for } j = 1, \dots, m-1, \quad f^{(m)}(z_0) = a \neq 0.$$

(As usual, we arrange $z_0 = 0$ and $f(z_0) = 0$.) Let $\{\omega_k\}_{1 \leq k \leq m}$ be the m th roots of unity. Then there exists a biholomorphic function φ in a neighborhood of 0 with $\varphi(0) = 0$ such that, for a chosen branch of $t^{1/m}$, all the solutions of the equation $f(z) = t$ near 0 are given by

$$z_k = \varphi(\omega_k t^{1/m}), \quad 1 \leq k \leq m.$$

Proof. In a neighborhood of 0 we may write

$$f(z) = z^m(a + zE(z)) = z^m g(z),$$

where g is analytic and $g(0) = a \neq 0$. Thus $g(z) \neq 0$ in some disk $\mathbb{D}_{\varepsilon_1}(0)$. There exists an analytic branch $h = g^{1/m}$ in $\mathbb{D}_{\varepsilon_1}(0)$. Choosing also a branch of $t^{1/m}$, our equation becomes

$$(zh(z))^m = t,$$

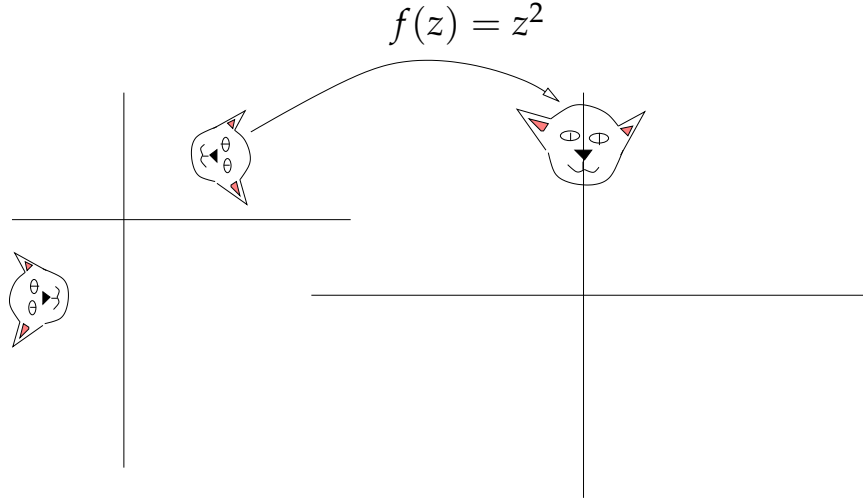


Figure 17: Ramified expansions (to be made rigorous in the sequel).

equivalent to m equations

$$zh(z) = \omega_k t^{1/m}, \quad k = 1, \dots, m.$$

The function $zh(z)$ is analytic at 0, and

$$(zh)'(0) = h(0) \neq 0.$$

By the inverse function theorem, each equation $zh = \omega_k t^{1/m}$ has exactly one solution near 0, namely

$$z_k(t) = \varphi(\omega_k t^{1/m}),$$

where φ is a biholomorphism in a neighborhood of 0. ■

Note 119

Expanding φ as a Taylor series,

$$\varphi(s) = \sum_{j \geq 0} a_j s^j,$$

we obtain

$$z_k(t) = \varphi(\omega_k t^{1/m}) = \sum_{j \geq 0} a_j (\omega_k t^{1/m})^j = \sum_{j \geq 0} a_j \omega_k^j t^{j/m}.$$

This is a *ramified expansion*, i.e. a convergent series in fractional powers of t . Such expansions are called *Puiseux series*.

This gives another proof of the open mapping theorem:

Theorem 120

Let $\mathcal{D}$ be a domain and $f : \mathcal{D} \rightarrow \mathbb{C}$ a nonconstant analytic function. Then f is open; that is, the image of every open set in $\mathcal{D}$ is open.

Proof. This follows immediately from the fact that φ is a biholomorphism. ■

25 Evaluation of definite integrals

Contour integrals can be evaluated using the residue theorem. Many improper integrals with endpoints at infinity can also be computed this way: they can be regarded as contour integrals on the Riemann sphere, where ∞ is often a singular point of the integrand. Similarly, many integrals with singular endpoints can be evaluated in closed form. We begin with a simple but important case.

Proposition 25.10. *Let R be a rational function, continuous on $\mathbb{R}$, such that there exists $C > 0$ with*

$$|R(z)| \leq C|z|^{-2} \quad \text{for all } z \in \mathbb{C}.$$

(In particular, this holds if the degree of the numerator is at least two less than that of the denominator.) Then

$$\int_{-\infty}^{\infty} R(x) dx = 2\pi i \sum_{z_i \in \text{UHP}} \text{Res}(R; z = z_i) = 2\pi i \sum_{z_i \in \text{LHP}} \text{Res}(R; z = z_i), \quad (25.2)$$

where the sums run over the poles of R in the upper (resp. lower) half-plane.

In particular, if all poles of R lie strictly in the upper (or strictly in the lower) half-plane, then the integral vanishes.

Proof. Consider the rectangular contour shown in Fig. 41, consisting of the interval $[-A, A]$ on the real line together with the top horizontal segment C_1 and two vertical segments. By the residue theorem,

$$\oint_{[-A, A] \cup C_1} R(z) dz = 2\pi i \sum_{z_i \in \text{UHP}} \text{Res}(R; z = z_i).$$

Hence

$$\begin{aligned} \int_{-\infty}^{\infty} R(x) dx &= \lim_{A \rightarrow \infty} \int_{-A}^A R(x) dx = \lim_{A \rightarrow \infty} \left[\oint_{[-A, A] \cup C_1} R(z) dz - \int_{C_1} R(z) dz \right] \\ &= 2\pi i \sum_{z_i \in \text{UHP}} \text{Res}(R; z = z_i) - \lim_{A \rightarrow \infty} \int_{C_1} R(z) dz. \end{aligned} \quad (25.3)$$

But by the assumption $|R(z)| = O(|z|^{-2})$ as $|z| \rightarrow \infty$,

$$\left| \int_{C_1} R(z) dz \right| \leq \left(\sup_{z \in C_1} |R(z)| \right) \cdot \text{length}(C_1) = O(A^{-2}) \cdot O(A) = O(A^{-1}) \rightarrow 0,$$

so the integral over C_1 vanishes in the limit and we obtain the claim. ■

Note. It is useful to interpret the procedure above as “pushing” (or deforming) the original contour of integration $\mathbb{R}$ upwards toward $+i\infty$, collecting residues along the way. Since a rational function has only finitely many poles, once the contour has been pushed above the highest pole (say to $\text{Im } z = h$), it can be moved further to $+i\infty$ without encountering new singularities. On this final horizontal segment, the integral tends to zero as $h \rightarrow +\infty$ because of the decay assumption on R . By contour independence in singularity-free regions, the contribution of this part is therefore exactly zero.

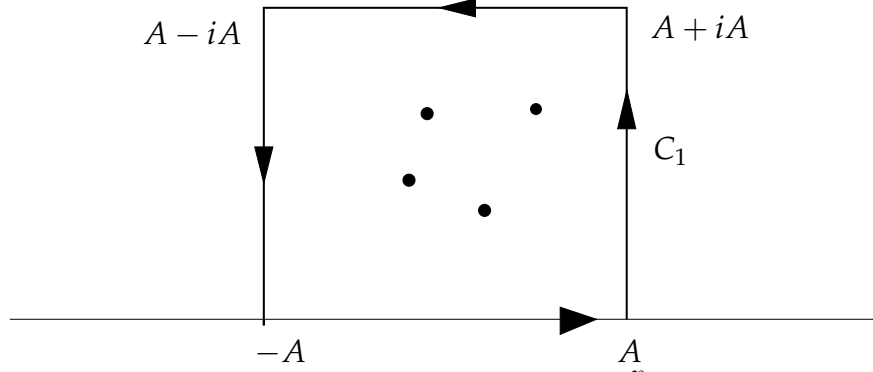


Figure 18: Rectangular contour used to evaluate $\int_{-\infty}^{\infty} R(x) dx$.

Example 121. Evaluate

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

Solution. The poles of $R(x) = \frac{1}{1+x^4}$ are the fourth roots of -1 , namely

$$e^{i\pi/4}, e^{3i\pi/4}, e^{5i\pi/4}, e^{7i\pi/4}.$$

In the upper half-plane we have $z_1 = e^{i\pi/4}$ and $z_2 = e^{3i\pi/4}$. The residue at each is

$$\text{Res}\left(\frac{1}{1+z^4}; z = z_i\right) = \frac{1}{(1+z^4)'|_{z=z_i}} = \frac{1}{4z_i^3}.$$

Therefore,

$$I = 2\pi i \left(\frac{1}{4z_1^3} + \frac{1}{4z_2^3} \right).$$

A short computation gives

$$\frac{1}{z_1^3} + \frac{1}{z_2^3} = \frac{1}{e^{3i\pi/4}} + \frac{1}{e^{9i\pi/4}} = e^{-3i\pi/4} + e^{-i\pi/4} = \sqrt{2}i.$$

Hence

$$I = 2\pi i \cdot \frac{1}{4} \cdot (\sqrt{2}i) = \frac{\pi}{\sqrt{2}}.$$

25.1 Fourier-type integrals

In calculating Fourier transforms, we encounter integrals of the form

$$\int_{-\infty}^{\infty} e^{iax} Q(x) dx, \quad (25.4)$$

where Q is a rational function, analytic on $\mathbb{R}$, and decaying at least like $1/x$ at infinity so that the (possibly improper) integral is well defined. We take $a > 0$ for definiteness, as in the inverse Fourier transform. As in the previous section, one way to evaluate the integral is to close the contour in the upper half-plane and push it towards $+i\infty$, where the exponential decays. To justify that this additional arc contributes nothing in the limit, Jordan's lemma is the standard tool.

Lemma 122 (Jordan). Assume $a > 0$, that Q is analytic in

$$\mathcal{D} = \{z : \operatorname{Im} z \geq 0, |z| > c\},$$

and let γ be the semicircle of radius $\rho > c$ in the UHP centered at the origin. Assume further that for large $|z|$, we have $|Q(z)| \leq M$. Then

$$\left| \int_{\gamma} e^{iaz} Q(z) dz \right| \leq \frac{M\pi}{a}. \quad (25.5)$$

In particular, if $Q(z) \rightarrow 0$ as $z \rightarrow \infty$, then

$$\int_{\gamma} e^{iaz} Q(z) dz \rightarrow 0 \quad \text{as } \rho \rightarrow \infty. \quad (25.6)$$

Proof. Let ρ_0 be such that $|Q(z)| \leq M$ for $|z| > \rho_0$. For $\rho > \rho_0$ and γ as above,

$$\begin{aligned} \left| \int_{\gamma} e^{iaz} Q(z) dz \right| &= \left| \int_0^{\pi} e^{ia\rho e^{i\varphi}} Q(\rho e^{i\varphi}) \rho i e^{i\varphi} d\varphi \right| \\ &\leq M \int_0^{\pi} \rho e^{-\rho a \sin \varphi} d\varphi = 2M \int_0^{\pi/2} \rho e^{-\rho a \sin \varphi} d\varphi, \end{aligned} \quad (25.7)$$

where the last equality uses the symmetry $\sin \varphi = \sin(\pi - \varphi)$.

Next, note that $\theta^{-1} \sin \theta$ is decreasing on $[0, \pi/2]$, so

$$\sin \theta \geq \frac{2}{\pi} \theta, \quad 0 \leq \theta \leq \frac{\pi}{2}.$$

Hence

$$\left| \int_{\gamma} e^{iaz} Q(z) dz \right| \leq 2M \int_0^{\pi/2} \rho e^{-\frac{2\rho a}{\pi} \varphi} d\varphi \leq \frac{M\pi}{a}, \quad (25.8)$$

which proves (25.5). The convergence to zero in the case $Q(z) \rightarrow 0$ as $|z| \rightarrow \infty$ is then immediate. ■

Proposition 25.11. Assume $a > 0$ and Q is a rational function, continuous on $\mathbb{R}$, and vanishing as $|z| \rightarrow \infty$. Then

$$\int_{-\infty}^{\infty} Q(x) e^{iax} dx = 2\pi i \sum_{z_i \in \mathbb{H}^u} \operatorname{Res}(Q(z) e^{iaz}; z = z_i), \quad (25.9)$$

where the sum runs over the poles of Q in the open upper half-plane $\mathbb{H}^u$.

Proof. This follows immediately by combining Jordan's lemma (Lemma 122) with the residue computation in Proposition 25.10. Details are left as an exercise. ■

25.2 Further examples

Example 123. Let $\tau > 0$ and evaluate

$$I(\tau) = \int_0^{\infty} \frac{\cos(\tau x)}{x^2 + 1} dx. \quad (25.10)$$

Solution. The integrand is even, so

$$2I(\tau) = \int_{-\infty}^{\infty} \frac{\cos(\tau x)}{x^2 + 1} dx = \operatorname{Re} \int_{-\infty}^{\infty} \frac{e^{i\tau x}}{x^2 + 1} dx. \quad (25.11)$$

This is of the form in Proposition 25.11, with poles in $\mathbb{H}^u$ at $z_1 = e^{i\pi/4}$ and $z_2 = e^{3i\pi/4}$. A short residue calculation gives

$$I(\tau) = \frac{\pi}{2} e^{-\tau}.$$

Note that $I(\tau)$ is the cosine Fourier transform of a function analytic in a neighborhood of $\mathbb{R}$, and it decays exponentially as $\tau \rightarrow \infty$.

Exercise 124 ([17], p. 116). Assume $\operatorname{Re} z > 0$. Show that

$$I(z) = \int_0^{\infty} \frac{e^{-t} - e^{-tz}}{t} dt = \log z. \quad (25.12)$$

Hint. The integrand is continuous at $t = 0$, so the integral is well defined. Moreover $I(z)$ depends analytically on z , and by dominated convergence,

$$I'(z) = \int_0^{\infty} e^{-tz} dt = \frac{1}{z}. \quad (25.13)$$

Thus $I(z) = \log z + C$ for some constant C . Show that $C = 0$.

Example 125 (An important definite integral). Show that

$$\int_0^{\infty} \frac{\sin t}{t} dt = \frac{\pi}{2}. \quad (25.14)$$

Solution. A naive attempt to write

$$\int_{-\infty}^{\infty} \frac{\sin t}{t} dt \stackrel{?}{=} \operatorname{Im} \int_{-\infty}^{\infty} \frac{e^{it}}{t} dt \quad (25.15)$$

fails, since the right-hand side is ill defined. Instead, proceed as follows:

1. Using the “box” contour argument (see figure below), with $A = (2N + 1)\pi/2$, $N \in \mathbb{Z}$, show that

$$\int_{-\infty}^{\infty} \frac{\sin t}{t} dt = \int_{-\infty+i}^{\infty+i} \frac{\sin t}{t} dt = \int_{-\infty}^{\infty} \frac{\sin(t+i)}{t+i} dt.$$

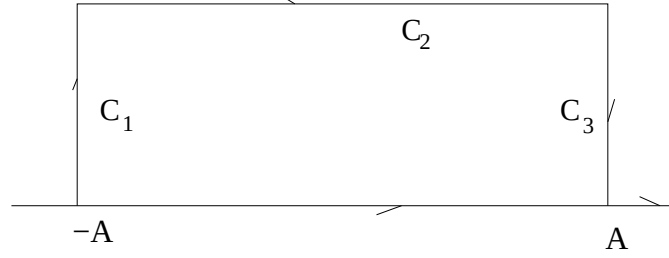
2. Now compute

$$\int_{-\infty}^{\infty} \frac{\sin(t+i)}{t+i} dt = \int_{-\infty}^{\infty} \frac{e^{i(t+i)} - e^{-i(t+i)}}{2i(t+i)} dt.$$

The first term vanishes by Proposition 25.11. The second term equals

$$\overline{\int_{-\infty}^{\infty} \frac{-e^{i(t-i)}}{-2i(t-i)} dt},$$

to which Proposition 25.11 applies again. Carrying out the residue computation yields the desired result.



Remark 126. The same method shows more generally that

$$\int_0^\infty \frac{\sin(ax)}{x} dx = \begin{cases} \frac{\pi}{2}, & a > 0, \\ -\frac{\pi}{2}, & a < 0. \end{cases}$$

This fact will reappear naturally in Fourier analysis, where it corresponds to the Fourier transform of the sign function.

26 Integrals of branched functions

We now show that, for $\alpha \in (0, 1)$,

$$\int_0^\infty \frac{t^{-\alpha}}{t+1} dt = \frac{\pi}{\sin(\pi\alpha)}. \quad (26.2)$$

Note first that the improper integral is well defined.

We place a cut along $\mathbb{R}^+$, and adopt the convention that $t^{-\alpha} = |t|^{-\alpha}$ on the upper side of the cut. If we turn clockwise around the origin, then on the lower side of the cut

$$t^{-\alpha} = |t|^{-\alpha} e^{-2\pi i \alpha}.$$

Hence the integral in (26.2) equals

$$\frac{1}{1 - e^{-2\pi i \alpha}} \int_\gamma \frac{t^{-\alpha}}{t+1} dt = \frac{1}{1 - e^{-2\pi i \alpha}} \lim_{R \rightarrow \infty} \int_{\partial D_R} \frac{t^{-\alpha}}{t+1} dt, \quad (26.3)$$

where γ is the boundary of $\mathbb{C} \setminus [0, \infty)$ traversed from 0 to ∞ on the upper side of the cut and from ∞ back to 0 on the lower side, and ∂D_R is the boundary of the disk D_R of radius R centered at 0 with a cut along $[0, R]$. This is a simply connected domain containing the pole of the integrand at -1 , traversed counterclockwise. See Fig. 19.

The latter integral can be calculated by residues: the residue of the integrand at $t = -1$ is

$$\text{Res}_{t=-1} \frac{t^{-\alpha}}{t+1} = e^{-\pi i \alpha},$$

which yields the result.

More generally, we have the following result.

Proposition 26.12. Assume $\text{Re } a \in (0, 1)$ and let Q be a rational function, continuous on $\mathbb{R}^+$, such that $x^a Q(x) \rightarrow 0$ both as $x \rightarrow 0$ and as $x \rightarrow \infty$. Then

$$\int_0^\infty x^{a-1} Q(x) dx = -\frac{\pi e^{-\pi i a}}{\sin(\pi a)} \sum_{z_i} \text{Res}(z^{a-1} Q(z); z_i), \quad (26.4)$$

where the sum runs over all poles z_i of Q .

Exercise 127. Prove Proposition 26.12.

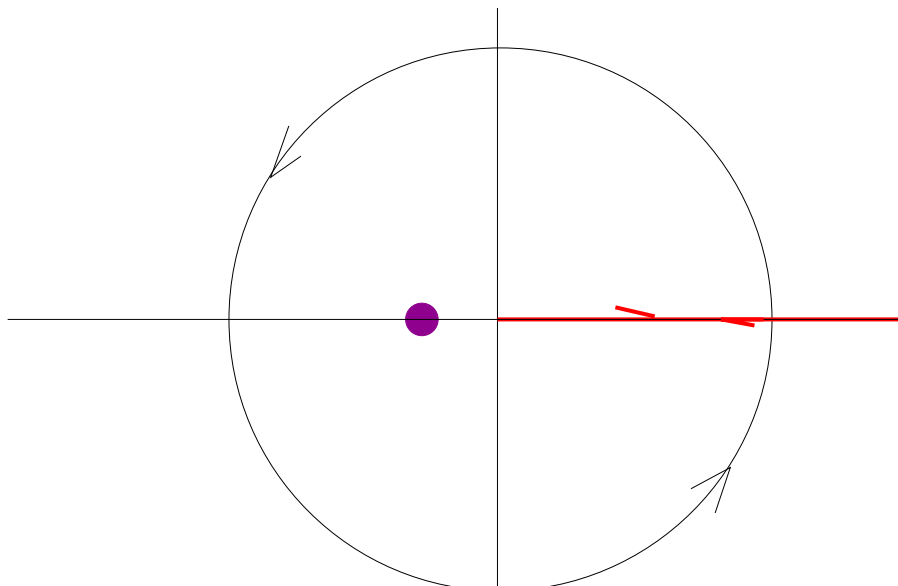


Figure 19: Contour for the evaluation of (26.2).

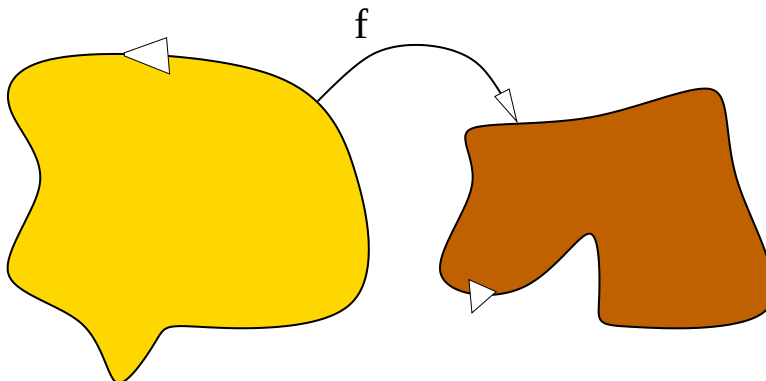
26.0.1 Conformally Equivalent Domains

Definition 128

Two domains $\mathcal{D}_1$ and $\mathcal{D}_2$ are **conformally equivalent** if there exists a biholomorphic bijection $f: \mathcal{D}_1 \rightarrow \mathcal{D}_2$. In this case, we say that f **maps conformally** $\mathcal{D}_1$ onto $\mathcal{D}_2$.

The boundary behavior of analytic maps is a crucial tool in determining their conformal mapping properties.

Recall that a curve is traversed in a positive (counterclockwise) direction if its parameterization is such that the interior lies to the left of the curve as the parameter increases.



Proposition 26.13. Assume that $f: \mathcal{D} \rightarrow \mathbb{C}$ is analytic and that γ is a simple, piecewise differentiable closed curve contained in $\mathcal{D}$ along with its interior.

If f is one-to-one from γ onto $f(\gamma)$, then f maps $\text{Int}(\gamma)$ one-to-one and conformally onto $f(\text{Int}(\gamma))$, and it preserves the orientation of the curve.

Proof. Since f is continuous and one-to-one on the simple closed curve γ , its image $f(\gamma)$ is also a simple closed curve. Let $w_0 \in \text{Int}(f(\gamma))$. By Theorem 86 (the argument principle), the number

of roots of $f(z) - w_0$ in $\text{Int}(\gamma)$ equals

$$I := \frac{1}{2\pi i} \int_{\gamma} \frac{f'(z)}{f(z) - w_0} dz, \quad (26.5)$$

which is some nonnegative integer n .

By the change of variables $w = f(z)$, and since f is injective on γ , this becomes

$$I = \frac{1}{2\pi i} \int_{f(\gamma)} \frac{dw}{w - w_0}.$$

This integral equals ± 1 , depending on the orientation of $f(\gamma)$. The minus sign is impossible since $I = n \geq 0$. Hence $n = 1$, and the orientation of $f(\gamma)$ is positive.

Therefore, f maps $\text{Int}(\gamma)$ injectively and conformally onto $\text{Int}(f(\gamma))$. ■

Note 129

(1) We can relax the condition that f is one-to-one on γ and replace it with the requirement that $\Gamma = f(\gamma)$ is a piecewise differentiable simple curve, and the winding number of $f(\gamma)$ is ± 1 (in other words, $f(\gamma)$ is a simple curve, but the parametrization $f(\gamma(t))$ can retrace itself). Indeed, since f is analytic but not one-to-one, the set $\{g \in \gamma \mid f(g) = 0\} = \{g_i = \gamma(t_i)\}_{i=1, \dots, n}$ is finite and we break the integral into $\sum_i \int_{\gamma(t_i)}^{\gamma(t_{i+1})}$; the parts of the curve that are retraced over themselves cancel out and we are left with

$$I = \frac{1}{2\pi i} \int_{\Gamma} \frac{dw}{w - w_0}.$$

(2) The differentiable simple curve could close at infinity on the Riemann sphere, such as $\mathbb{R}$ traversed positively, since by a conformal Möbius transformation it can be mapped into a differentiable closed curve.

26.1 Automorphisms

Automorphisms are biholomorphic maps from a domain $\mathcal{D}$ onto itself. They play an important role in complex analysis, since they describe the symmetries of the domain and often form groups with rich geometric and analytic structure. The group of automorphisms of $\mathcal{D}$ is denoted by $\text{Aut}(\mathcal{D})$.

They are also central in understanding the celebrated uniformization theorem, which we anticipate here and will return to in more detail later.

The very important and deep *uniformization theorem* states that, up to such isomorphisms, there are only 3 distinct *simply connected* Riemann surfaces: $\hat{\mathbb{C}}$, $\mathbb{C}$ and $\mathbb{D}$. Riemann surfaces that are not “very simple” are conformally equivalent to $\mathbb{D}$.

26.2 Automorphisms of the plane

Theorem 130

The group of biholomorphic self-maps of $\mathbb{C}$ is exactly the affine group $z \mapsto az + b : a \in \mathbb{C} \setminus 0, b \in \mathbb{C}$.

Proof. The maps $z \mapsto az + b$ with $a \neq 0$ are entire and bijective with holomorphic inverse, so they lie in $\text{Aut}(\mathbb{C})$. For the converse, let $f \in \text{Aut}(\mathbb{C})$. After a translation we may assume $f(0) = 0$.

Because f is a homeomorphism, its inverse is continuous at 0. Hence there exists $c > 0$ such that

$$|w| < c \implies |f^{-1}(w)| < 1.$$

Equivalently,

$$|z| \geq 1 \implies |f(z)| \geq c. \quad (26.6)$$

Consider $h(z) = f(1/z)$ on $0 < |z| < \infty$. If 0 were an essential singularity of h , then by the Casorati–Weierstrass theorem the image of every punctured neighborhood of 0 would be dense in $\mathbb{C}$, contradicting injectivity. Thus ∞ is either a removable singularity or a pole of f .

If ∞ were a removable singularity of f , then f would be bounded and hence, by Liouville's theorem, be constant—contradicting injectivity. Thus ∞ is a pole of f , so f is a polynomial (cf. Exercise 43). By the Fundamental Theorem of Algebra, an injective polynomial must have degree 1, and therefore $f(z) = az + b$ with $a \neq 0$. ■

26.3 The Riemann sphere

The Riemann sphere is a model of the extended complex plane (also called the closed complex plane): the complex plane plus one point at infinity. This extended plane represents the extended complex numbers, that is, the complex numbers together with the value ∞ . Analytic functions on $\hat{\mathbb{C}}$ are defined using bi-holomorphic charts.

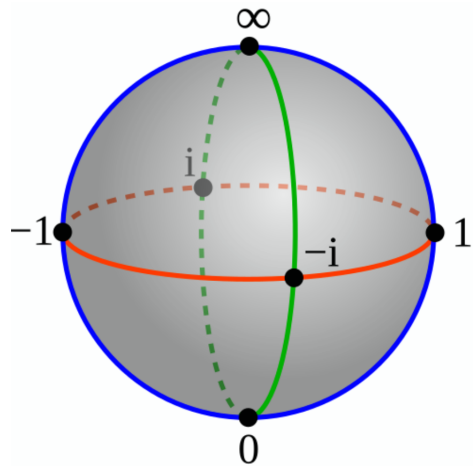


Figure 20: The Riemann sphere (adapted from Wikimedia Commons).

26.4 Linear fractional transformations (LFTs): a first view

A linear fractional transformation, or Möbius transformation, is a map of the form

$$L(z) = \frac{az + b}{cz + d}, \quad ad - bc \neq 0, \quad (26.7)$$

which is meromorphic on $\hat{\mathbb{C}}$ with a single pole at $z = -d/c$.

Exercise 131. Associate to an LFT its coefficient matrix

$$\hat{M}(L) := \begin{pmatrix} a & b \\ c & d \end{pmatrix}. \quad (26.8)$$

If L_1 and L_2 are LFTs, show that

$$\hat{M}(L_1 \circ L_2) = \hat{M}(L_1)\hat{M}(L_2), \quad (26.9)$$

Hence LFTs with composition are (projectively) isomorphic to $\text{GL}(2, \mathbb{C})$ with multiplication.⁶

26.5 Automorphisms of the Riemann sphere

Theorem 132

A map $F : \hat{\mathbb{C}} \rightarrow \hat{\mathbb{C}}$ is analytic if and only if F is a rational function.

Proof. If F is analytic on $\hat{\mathbb{C}}$, then it has no essential singularities anywhere on the sphere, though it may have poles (since ∞ is a regular point of $\hat{\mathbb{C}}$). In particular, the restriction of F to $\mathbb{C} = \hat{\mathbb{C}} \setminus \{\infty\}$ is meromorphic.

Because $\hat{\mathbb{C}}$ is compact, the set of poles must be finite—otherwise the poles would have an accumulation point (necessarily at ∞), contradicting analyticity there (why?). Let P be a polynomial whose zeros are exactly the poles of F (counted with multiplicity). Then PF is entire on $\mathbb{C}$.

Consider the function h defined on a punctured neighborhood of 0 in $\mathbb{C}$ by

$$h(z) = P(1/z) F(1/z).$$

Since F is analytic at ∞ , h has at most a pole at 0. Equivalently, the Laurent expansion of h about 0 has only finitely many negative powers of z ; otherwise PF would have an essential singularity at ∞ . Thus

$$PF(z) = \sum_{j=-m}^n c_j z^j, \quad m, n \geq 0.$$

Analyticity of PF at $z = 0$ forces $m = 0$, so PF is a polynomial Q . Hence $F = Q/P$ is a rational function.

Conversely, any rational function $F = Q/P$ is meromorphic on $\mathbb{C}$ and has at most a pole at ∞ , so it defines an analytic map on $\hat{\mathbb{C}}$. ■

Theorem 133

The automorphism group of $\hat{\mathbb{C}}$ is precisely the Möbius group:

$$\text{Aut}(\hat{\mathbb{C}}) = \left\{ z \mapsto \frac{az + b}{cz + d} : ad - bc \neq 0 \right\}.$$

⁶Multiplying all entries of a 2×2 matrix by the same nonzero scalar does not change the associated linear fractional transformation. Thus what really matters is the matrix up to overall scaling. This is what is meant by “projectively isomorphic” here.

Proof. Every Möbius map is a rational function with at most one simple pole, and its inverse is again a Möbius map, so it is clearly an automorphism of $\hat{\mathbb{C}}$.

Conversely, let f be an automorphism of $\hat{\mathbb{C}}$. By Theorem 132, $f = P/Q$ with P, Q polynomials having no common factors. Since f is bijective, it can have at most one pole, and that pole must be simple; hence $Q(z) = cz + d$ is linear.

If $\deg P \geq 2$, then, by the Fundamental Theorem of Algebra (Theorem 45), the equation $f(z) = 0$ has $\deg P$ solutions (counting multiplicities). Hence $\deg P = 1$, so $P(z) = az + b$ is linear.

Finally, $ad - bc \neq 0$, otherwise f would be constant. Thus f is a Möbius map. ■

27 Linear fractional transformations (Möbius transformations)

Linear fractional transformations were defined in (26.7).

If $c = 0$ we have a linear function. If $c \neq 0$ we write

$$\frac{az + b}{cz + d} = \frac{a}{c} - \frac{ad - bc}{c^2(z + d/c)}; \quad z \neq -d/c \quad (27.1)$$

It is clear from (27.1) that $S(z_1) = S(z_2)$ iff $z_1 = z_2$ and in particular $S'(z) \neq 0$: If a LFT L is defined in the domain $\mathcal{D}$ (that is, $-d/c \notin \mathcal{D}$), then $\mathcal{D}$ and $L(\mathcal{D})$ are conformally equivalent. It also follows from the decomposition (27.1) that

Proposition 27.1. The Möbius group is generated by $z \mapsto z + a$, $z \mapsto az$ and $z \mapsto 1/z$. This implies that $\text{PSL}(2, \mathbb{C})$ is generated by the matrices of the form

$$\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}, \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} \text{ and } \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad (27.2)$$

(we took a negative sign in the last matrix to have determinant one.)

Hence, the automorphisms of $\hat{\mathbb{C}}$ correspond to $\text{PSL}(2, \mathbb{C})$ while those of $\mathbb{C}$ correspond to the subgroup generated by the matrices $\begin{pmatrix} a & 0 \\ 0 & 1 \end{pmatrix}$ and $\begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$.

Exercise 134. Show that the inversion $z \mapsto 1/z$ maps a line or a circle into a line or a circle. Hint: Show first that the equation

$$\alpha z \bar{z} + \beta z + \bar{\beta} \bar{z} + \gamma = 0$$

where $\alpha, \gamma \in \mathbb{R}$ and $\beta \in \mathbb{C}$ and $|\beta|^2 > \alpha\gamma$ is the most general equation of a line or a circle. Then apply the inversion to the equation.

As a result we have an important property of LFTs

Proposition 27.2. A LFT maps a line or a circle into a line or a circle.

Proof. By Exercise 134 inversion has the property above, and clearly, so do Euclidean transformations. □

27.1 Finding specific LFTs

A line is determined by two points and a circle is determined by three. We now show that for any two circles or lines there exists a LFT mapping one into the other, and we describe how to find it explicitly.

Let z_1, z_2, z_3 be three distinct points in $\mathbb{C}$. Then the transformation

$$S(z) = \frac{z_1 - z_3}{z_1 - z_2} \frac{z - z_2}{z - z_3} \quad (27.3)$$

maps z_1, z_2, z_3 into $1, 0, \infty$ in this order; this is straightforward to check. If one of z_1, z_2, z_3 is ∞ , we pass to the limit. For example, the transformation taking (∞, z_1, z_2) to $(1, 0, \infty)$ is

$$L(z) = \frac{z - z_1}{z - z_2}. \quad (27.4)$$

To find a transformation that maps z_1, z_2, z_3 into $\tilde{z}_1, \tilde{z}_2, \tilde{z}_3$ in this order, we may write it as $\tilde{L} = \tilde{L}^{-1} \circ L$, where L maps (z_1, z_2, z_3) to $(1, 0, \infty)$ and $\tilde{L}$ maps $(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3)$ to $(1, 0, \infty)$. By Exercise 135, this transformation is unique.

Exercise 135. Show that a LFT that takes $(1, 0, \infty)$ into itself is the identity.

27.1.1 Cross ratio

If $z_i, i = 1, \dots, 4$ are four distinct points and $w_i = L(z_i)$, then (check!)

$$\frac{w_1 - w_2}{w_1 - w_3} \cdot \frac{w_3 - w_4}{w_2 - w_4} = \frac{z_1 - z_2}{z_1 - z_3} \cdot \frac{z_3 - z_4}{z_2 - z_4}.$$

The quantity

$$[z_1, z_2; z_3, z_4] = \frac{z_1 - z_2}{z_1 - z_3} \cdot \frac{z_3 - z_4}{z_2 - z_4}$$

is called the *cross ratio* of the four points. It is invariant under Möbius transformations. This provides a handy way to determine the image of a fourth point once a transformation has been fixed using three points.

27.2 Mappings of regions

We know that LFTs are *conformal and one-to-one* and transform circles and lines onto circles and lines. Proposition 26.13 shows that LFTs map disks and half-planes to disks and half-planes. To see the latter, note that on the Riemann sphere LFTs map circles to circles, hence disks to disks as well. A rigorous proof requires verifying that the interior/exterior of a circle are mapped to interiors/exterior respectively — we leave this as a straightforward exercise.

Exercise 136. (i) Show that the *Cayley transform*

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

maps the upper half-plane (UHP) conformally onto the unit disk $\mathbb{D}$. Its inverse

$$C^{-1}(w) = i \frac{1 + w}{1 - w}$$

maps $\mathbb{D}$ conformally back onto the UHP.

- (ii) Find an LFT that maps the disk $(x-1)^2 + (y-2)^2 = 4$ onto the unit disk and sends the center $(1, 2)$ to $i/2$.
- (iii)* Find the most general LFT that maps the unit disk onto itself.
Hint: The general solution has the form

$$z \mapsto e^{i\theta} \frac{z-a}{1-\bar{a}z}, \quad |a| < 1, \theta \in \mathbb{R}.$$

27.3 The Schwarz lemma

As usual, we let $\mathbb{D}$ denote the open unit disk.

Theorem 137: Schwarz lemma

Let $f : \mathbb{D} \rightarrow \mathbb{D}$ be analytic with $f(0) = 0$. Then:

- (i) For all $z \in \mathbb{D}$,
- $$|f(z)| \leq |z|. \quad (27.5)$$
- (ii) If equality in (27.5) holds for some nonzero $z_0 \in \mathbb{D}$, then there exists $\varphi \in \mathbb{R}$ such that $f(z) = e^{i\varphi}z$.
- (iii) We have $|f'(0)| \leq 1$, with equality if and only if $f(z) = e^{i\varphi}z$ for some $\varphi \in \mathbb{R}$.

Proof. (i) Since $f(0) = 0$, the quotient $f(z)/z$ extends analytically to all of $\mathbb{D}$ (by defining its value at 0 to be $f'(0)$). Because $f(\mathbb{D}) \subset \mathbb{D}$, the maximum modulus principle gives

$$\left| \frac{f(z)}{z} \right| \leq \limsup_{r \uparrow 1} \max_{|z|=r} \left| \frac{f(z)}{z} \right| \leq 1,$$

which implies (27.5).⁷

(ii) Suppose equality holds at some nonzero z_0 . Then $|f(z_0)/z_0| = 1$, so the analytic function $f(z)/z$ attains its maximum modulus inside $\mathbb{D}$. By the maximum modulus principle, $f(z)/z$ must be constant, say C . Thus $f(z) = Cz$, and since $|C| = 1$, we may write $C = e^{i\varphi}$ with $\varphi \in \mathbb{R}$.

(iii) The bound $|f'(0)| \leq 1$ follows immediately by taking the limit $z \rightarrow 0$ in (27.5). If equality holds, write $f(z) = e^{i\varphi}z + a_m z^{m+1} + \dots$ with $m \geq 1$. Then

$$\frac{f(z)}{z} = e^{i\varphi}(1 + a_m z^m + \dots).$$

Choosing $z = \varepsilon e^{-i \arg(a_m)/m}$ with ε small, we obtain $|f(z)| > |z|$, contradicting (i). Hence $f(z) \equiv e^{i\varphi}z$. $\square$

Corollary 138. If h is an automorphism of the unit disk and $h(0) = 0$, then

$$h(z) = e^{i\varphi}z \quad \text{for some } \varphi \in \mathbb{R}.$$

Proof. By Theorem 137, we have $|h(z)| \leq |z|$ for all $z \in \mathbb{D}$. Since h is an automorphism, its inverse h^{-1} is also an automorphism of $\mathbb{D}$ and satisfies $h^{-1}(0) = 0$. Applying Theorem 137 to h^{-1} gives $|h^{-1}(w)| \leq |w|$ for all $w \in \mathbb{D}$. In particular, with $w = h(z)$ we obtain

$$|z| = |h^{-1}(h(z))| \leq |h(z)|.$$

⁷You can show that $\limsup$ can be replaced by $\lim$ based on the monotonicity implied by the maximum principle.

Combining this with the first inequality, we deduce $|h(z)| = |z|$ for all $z \in \mathbb{D}$. By Theorem 137(ii), it follows that $h(z) = e^{i\varphi}z$ for some $\varphi \in \mathbb{R}$. $\square$

The automorphisms of $\hat{\mathbb{C}}, \mathbb{C}, \mathbb{D}$ are therefore subgroups of $\text{PSL}(2, \mathbb{C})$.

Of course, knowing the generators of a group can be very useful, for instance when we need to prove a property of the whole group.

27.4 Automorphisms of the unit disk

We have seen that the subgroup of LFTs

$$L(z) = e^{i\varphi} \frac{z - \alpha}{1 - \bar{\alpha}z}, \quad \varphi \in \mathbb{R}, |\alpha| < 1, \quad (27.6)$$

maps the unit disk one-to-one onto itself.

The converse is also true:

Theorem 139

Every automorphism f of $\mathbb{D}$ is of the form (27.7), with $\alpha = f^{-1}(0)$.

Proof. Let $\alpha = f^{-1}(0)$. Consider the map $S(z) = \frac{z - \alpha}{1 - \bar{\alpha}z}$, which sends α to 0. Then $h := S \circ f^{-1}$ is an automorphism of $\mathbb{D}$ satisfying $h(0) = 0$. By Corollary 138, we must have $h(z) = e^{i\varphi}z$ for some $\varphi \in \mathbb{R}$. Thus f is of the form (27.7). $\square$

Exercise 140. Show that the automorphisms of the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} : \text{Im } z > 0\}$ are precisely the LFTs

$$z \mapsto \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{R}, ad - bc > 0.$$

- (i) Verify directly that such maps preserve $\mathbb{H}$ and are bijective.
- (ii) Show that if $\varphi \in \text{Aut}(\mathbb{H})$ satisfies $\varphi(0) = 0$ and $\varphi(1) = 1$, then $\varphi(z) = z$.
- (iii) Conclude that this subgroup of LFTs is isomorphic to $\text{PSL}(2, \mathbb{R})$.

27.5 Automorphisms of the unit disk

We have seen that the subgroup of LFTs

$$S(z) = e^{i\varphi} \frac{z - \alpha}{1 - \bar{\alpha}z}, \quad \varphi \in \mathbb{R}, |\alpha| < 1, \quad (27.7)$$

maps the unit disk one-to-one onto itself.

The converse is also true:

Theorem 141

Every automorphism f of $\mathbb{D}$ is of the form (27.7), with $\alpha = f^{-1}(0)$.

Proof. Let $\alpha = f^{-1}(0)$. Consider the map $S(z) = \frac{z - \alpha}{1 - \bar{\alpha}z}$, which sends α to 0. Then $h := S \circ f^{-1}$ is an automorphism of $\mathbb{D}$ satisfying $h(0) = 0$. By Corollary 138, we must have $h(z) = e^{i\varphi}z$ for some $\varphi \in \mathbb{R}$. Thus f is of the form (27.7). $\square$

Exercise 142. Show that the automorphisms of the upper half-plane $\mathbb{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ are precisely the LFTs

$$z \mapsto \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{R}, \quad ad - bc > 0.$$

- (i) Verify directly that such maps preserve $\mathbb{H}$ and are bijective.
- (ii) Show that if $\varphi \in \operatorname{Aut}(\mathbb{H})$ satisfies $\varphi(0) = 0$ and $\varphi(1) = 1$, then $\varphi(z) = z$.
- (iii) Conclude that this subgroup of LFTs is isomorphic to $PSL(2, \mathbb{R})$.

Remark 143. The automorphism group of $\mathbb{D}$ acts *transitively* on $\mathbb{D}$: for any $z_1, z_2 \in \mathbb{D}$ there exists an automorphism sending z_1 to z_2 . In fact, $\operatorname{Aut}(\mathbb{D})$ acts even more strongly on the boundary: it is *triply transitive* on $\partial\mathbb{D}$, meaning that any three distinct boundary points can be mapped to any other three distinct boundary points. This illustrates the very rich symmetry of the unit disk.

Remark 144. We see that the automorphism groups of all three “canonical” simply connected Riemann surfaces (cf. Theorem 227) are realized as subgroups of $PSL(2, \mathbb{C})$:

- For the Riemann sphere $\hat{\mathbb{C}}$, $\operatorname{Aut}(\hat{\mathbb{C}}) = PSL(2, \mathbb{C})$.
- For the complex plane $\mathbb{C}$, the automorphisms are the affine maps $z \mapsto az + b$ with $a \in \mathbb{C}^\times$, $b \in \mathbb{C}$, corresponding to the subgroup of upper-triangular matrices in $PSL(2, \mathbb{C})$.
- For the unit disk $\mathbb{D}$, the automorphisms are the Blaschke factors

$$z \mapsto e^{i\varphi} \frac{z - \alpha}{1 - \bar{\alpha}z}, \quad |\alpha| < 1, \quad \varphi \in \mathbb{R},$$

forming a subgroup of $PSL(2, \mathbb{C})$.

Thus, in each canonical case the symmetry group of the surface embeds naturally into $PSL(2, \mathbb{C})$.

28 Some special biholomorphic transformations

Many elementary conformal maps are compositions of Möbius maps, ramifications ($z \mapsto z^\alpha$), exponential and trigonometric maps, and other simple transformations such as the Joukowski map. We illustrate below a number of useful transformations; see [1], [15] for further examples. A good number of interesting domains can be mapped to the unit disk using combinations of these transformations. Note that by Proposition 26.13 it suffices to examine carefully how the boundaries are mapped in order to understand the action of a map on a domain.

1. **Möbius transformations.** We have already discussed these in detail. A particularly useful special case is the Cayley transform

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

which maps the upper half-plane conformally onto $\mathbb{D}$.

2. **Power maps.** The map $z \mapsto z^{1/\alpha}$ transforms sectors into half-planes. See Fig. 21.
3. **Maps of the cut plane.**



Figure 21: The map $z \mapsto z^{1/\alpha}$ maps the open sector of opening $\pi\alpha$ conformally onto the open upper half-plane. The boundary point 0 maps to 0, and the upper boundary of the sector is rotated to $-\mathbb{R}$.

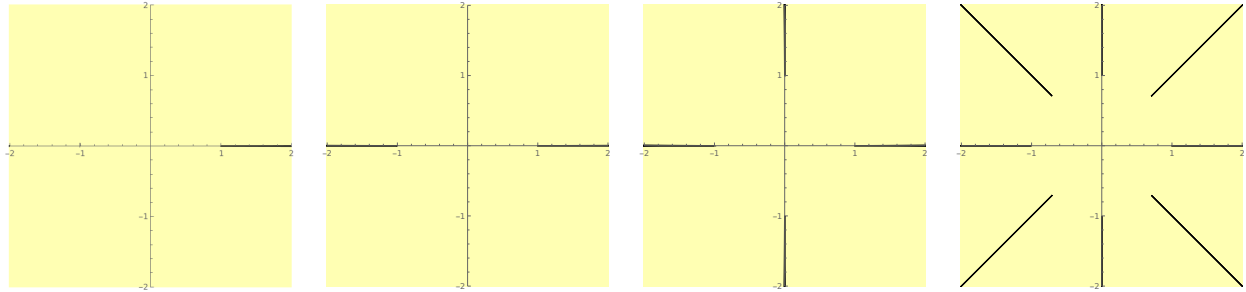


Figure 22: $\mathbb{C}$ with one, two, four, and eight symmetric cuts. More generally, the directions of the cuts can be taken along the roots of unity $\omega_1, \dots, \omega_k, k \in \mathbb{N}$.

Exercise 145. (a) Show that $\varphi_1(z) = \frac{4z}{(1+z)^2}$ maps conformally the unit disk onto $\mathbb{C} \setminus [1, \infty)$. Verify, using Proposition 26.13, that the two sides of $(1, \infty)$ are treated as distinct boundary arcs, as discussed in §26.

(b) Show that for any $n \in \mathbb{N}$ the map

$$\varphi_n(z) = [\varphi_1(z^n)]^{1/n}$$

maps $\mathbb{D}$ conformally onto $\mathbb{C}$ with symmetric cuts along $\omega_k, k = 1, \dots, n$.

4. The Joukowski transformation. This map sends the region in the upper half-plane lying outside the unit circle conformally onto the upper half-plane (see Fig. ??). By rescalings and translations, one can adjust the radius or center along $\mathbb{R}^+$. It is defined by

$$J(z) = z + \frac{1}{z}. \quad (28.8)$$

Exercise 146. Check, by examining boundary behavior and applying Proposition 26.13, that J indeed maps the indicated region conformally onto the UHP. What is the image under J of the reflection across $\mathbb{R}$ of the yellow region?

5. Exponential map. The exponential function maps conformally the strip

$$S = \{z \in \mathbb{C} : 0 < \operatorname{Im} z < \pi\}$$

onto $\mathbb{C} \setminus (-\infty, 0]$.

Exercise 147. Use some of the maps discussed above to construct a conformal map from $\mathbb{D}$ onto the strip S .

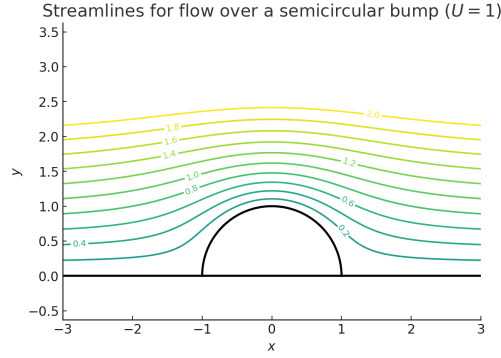


Figure 23: Velocity field lines in Joukowski's domain

5. The cosine maps the region $\{z \mid \operatorname{Im} z > 0, 0 < x < \pi\}$ conformally to the UHP.

Exercise 148. Note that the cosine and the exponential are linked via a Joukowski map. Can you explain the mapping of $\cos$ based on 5. and J ?

Note 149

Keep in mind that the conformal maps above **are not tied to the domains presented**, except loosely through the fact that these domains are often chosen to be some maximal region of conformality. Often subdomains of a maximal domain are interesting in applications.

Application: Ideal flow above a semicircular bump (normalized case $U = 1$)

We consider irrotational, incompressible, inviscid flow in the upper half-plane above the semicircle $\{|z| = 1, \operatorname{Im} z \geq 0\}$. The boundary condition is free-slip (no penetration). As $x \rightarrow \pm\infty$, the flow approaches the uniform stream $(v, u) = (1, 0)$; see Fig. ??.

Complex potential. The Joukowski map $w = J(z) = z + \frac{1}{z}$ flattens the boundary, sending the upper half-plane above the semicircle to the upper half w -plane, with the boundary mapping to the real line. In the w -plane, the uniform stream with unit speed has potential

$$F(w) = w.$$

Therefore the complex potential in the z -plane is

$$W(z) = F(J(z)) = z + \frac{1}{z}.$$

Velocity field. The complex velocity is

$$\frac{dW}{dz} = 1 - \frac{1}{z^2}.$$

Thus, writing $z = x + iy$,

$$v(x, y) = 1 - \frac{x^2 - y^2}{(x^2 + y^2)^2}, \quad u(x, y) = \frac{2xy}{(x^2 + y^2)^2}, \quad |z| > 1, \operatorname{Im} z > 0.$$

Potential and stream function. The potential and stream function are

$$\varphi(x, y) = \operatorname{Re} W(z) = x + \frac{x}{x^2 + y^2}, \quad \psi(x, y) = \operatorname{Im} W(z) = y - \frac{y}{x^2 + y^2}.$$

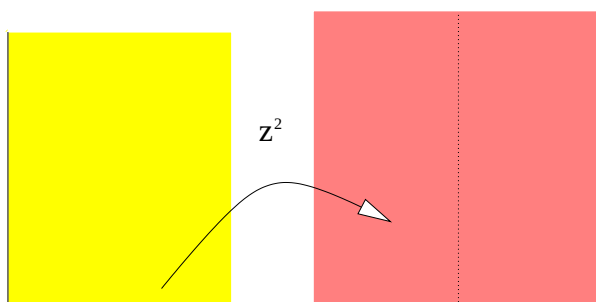
Boundary conditions. On the boundary $|z| = 1$, one checks $\psi = 0$, so the semicircle and real axis are streamlines (no penetration). As $x \rightarrow \pm\infty$, $(v, u) \rightarrow (1, 0)$, the uniform flow. Stagnation points occur at $z = \pm 1$, where $W'(z) = 0$.

Other common conformal maps are depicted in the following figures. The Cayley transform

$$\frac{z - i}{z + i}$$

maps conformally the upper half-plane (UHP) onto $\mathbb{D}$. It plays an important role in the theory of unbounded operators. Any other function f that maps conformally the UHP to $\mathbb{D}$ must be the Cayley transform composed on the left with an automorphism of the disk.

Exercise 150. Write the general form of such a map.



1

$$\mathbb{C} \setminus \overline{\mathbb{R}^-}$$

Figure 24: A conformal mapping example.

Exercise 151. 1. Check the transformations illustrated in the figures above.

2. Draw a similar picture for the mapping $\sin z$ from the upper half-strip bordered by the half-lines $x = \pm\pi/2$, $y > 0$.
3. Find a conformal map of the quarter disk $|z| < 1$, $\arg(z) \in (0, \pi/2)$ onto the upper half-plane.

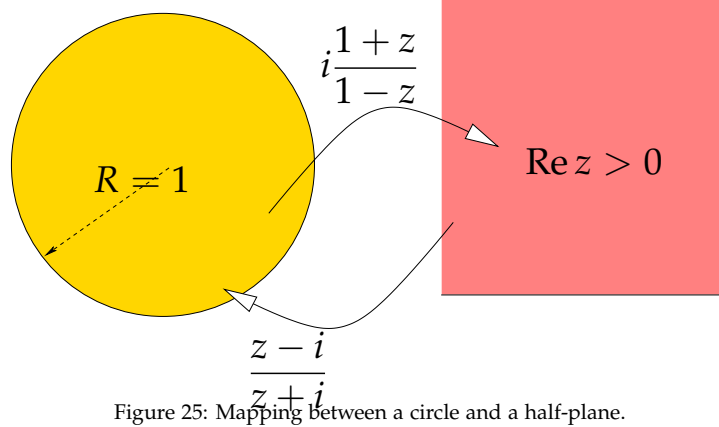


Figure 25: Mapping between a circle and a half-plane.

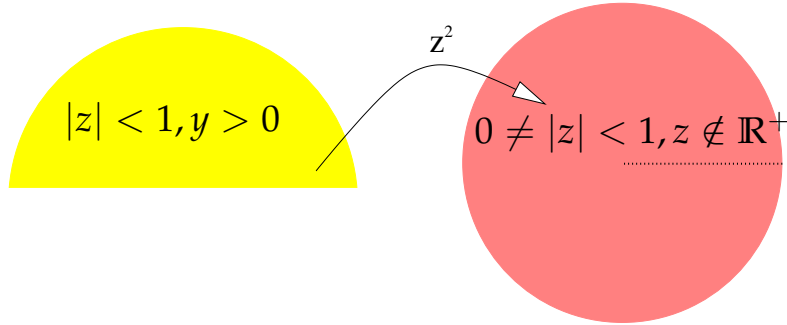


Figure 26: Another conformal mapping illustration.

4. Find a conformal map of the half-disk $|z| < 1, \arg(z) \in (0, \pi)$ onto the half-strip $x < 0, y \in (0, \pi)$.
5. Find a conformal map of the right half-plane (RHP) with a cut along $[0, 1]$ in the RHP.

It is useful to remark that we can find linear fractional transformations that map conformally a region between two circles into a half-plane (or disk) using very simple transformations. As an illustration, let us map the “moon crescent” M in Fig. 27 onto a half-plane.

The equations of the bounding circles are

$$x^2 + y^2 = 16, \quad (x - 3)^2 + y^2 = 9.$$

Solving for their intersection points $a_{\pm}$ we obtain

$$a_{\pm} = \frac{8}{3} \pm \frac{4i\sqrt{5}}{3}.$$

The angle between the two circles is the same as the angle between the tangent vectors at the intersection, which can be computed at a_+ as

$$\arg\left(\frac{a_+ - 3}{a_+}\right) = \arctan\left(\frac{\sqrt{5}}{2}\right) =: \varphi_0.$$

The key idea is that if a Möbius map sends one of the intersection points to infinity, the circular arcs are mapped to straight rays, which simplifies the geometry of the domain. Specifically, if we choose an LFT sending

$$a_- \mapsto 0, \quad 0 \mapsto 1, \quad a_+ \mapsto \infty,$$

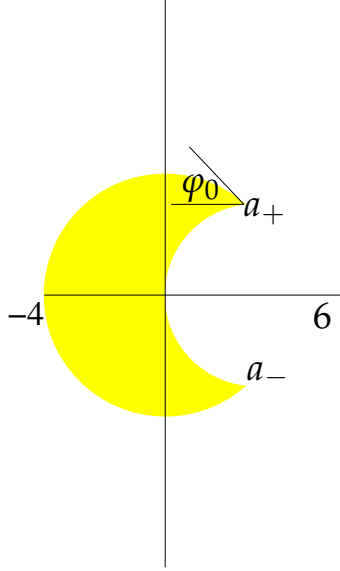


Figure 27: Mapping a crescent-shaped domain to a half-plane.

then both arcs become rays (since they meet at ∞). One of them is mapped to the positive real axis $\mathbb{R}^+$, while the other is mapped to a ray making angle φ_0 with $\mathbb{R}^+$. Such a transformation is for instance

$$S(z) = \frac{a_+}{a_-} \frac{z - a_-}{z - a_+}. \quad (28.9)$$

Finally, to straighten the resulting sector of opening φ_0 into the upper half-plane, we apply the power map

$$z \mapsto z^{\pi/\varphi_0}.$$

Exercise 152. 1. Verify the coordinates $a_{\pm}$ of the intersection points.

2. Check that the angle between the circles at a_+ equals $\varphi_0 = \arctan(\sqrt{5}/2)$.
3. Confirm that under the Möbius transformation (28.9) the small arc becomes $\mathbb{R}^+$ and the larger one becomes a ray at angle φ_0 .
4. Show that the ramified map $z \mapsto z^{\pi/\varphi_0}$ sends the resulting sector conformally onto the upper half-plane.

Example 2. Solve $\Delta u = 0$ in the quarter disk

$$\Omega = \{ |z| < 1, \arg(z) \in (0, \frac{\pi}{2}) \}$$

with boundary conditions: $u = 1$ on the circular arc, and $u = 0$ on the two radii.

Solution. The strategy is to find a conformal map of Ω onto a horizontal strip

$$S = \{ w = x + iy : 0 < y < 1 \}$$

such that the circular arc maps to the line $y = 1$, while the radii map to the line $y = 0$. In the strip S , the harmonic solution of the boundary problem is immediate:

$$u(w) = \operatorname{Im} w.$$

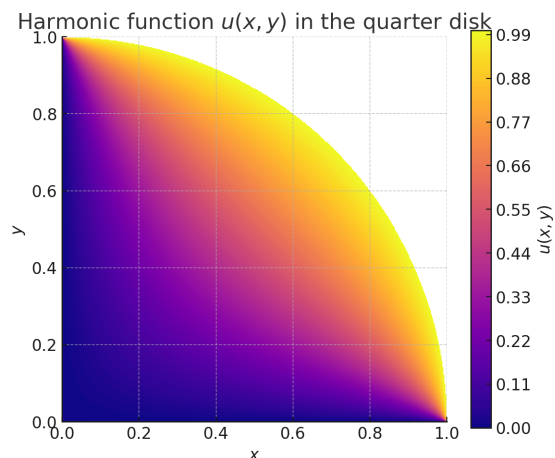


Figure 28: Color density plot of the heat equation solution for a quarter disk

We then pull back u through the conformal transformations.

Step 1. The map $z \mapsto z^2$ sends the quarter disk onto the upper half-disk $\{|z| < 1, \operatorname{Im} z > 0\}$. On the boundary we still have $u = 1$ on the arc, $u = 0$ on the interval $[-1, 1]$.

Step 2. Map the upper half-disk to the first quadrant by a Möbius transformation sending $z = 1$ to ∞ :

$$z \mapsto i \frac{1+z}{1-z}.$$

This sends the interval $[-1, 1]$ to $\mathbb{R}^+$ and the semicircle to $i\mathbb{R}^+$, so the quadrant is obtained with boundary data $u = 0$ on $\mathbb{R}^+$ and $u = 1$ on $i\mathbb{R}^+$.

Step 3. The squaring map $z \mapsto z^2$ opens the quadrant into the upper half-plane $\{\operatorname{Im} z > 0\}$, with boundary values $u = 0$ on $\mathbb{R}^+$ and $u = 1$ on $\mathbb{R}^-$.

Step 4. Finally, the logarithm maps the upper half-plane to a horizontal strip. Specifically,

$$w = \frac{2}{\pi} \ln \left(\frac{1+z^2}{1-z^2} \right)$$

maps Ω conformally onto the strip $\{0 < \operatorname{Im} w < 1\}$, with the desired correspondence of boundary parts.

Thus the solution is

$$u(z) = \operatorname{Im} \left(\frac{2}{\pi} \ln \frac{1+z^2}{1-z^2} \right),$$

or, explicitly,

$$u(x, y) = 2 - \frac{2}{\pi} \arctan \left(\frac{4xy}{-1 + (x^2 + y^2)^2} \right).$$

Exercise* The temperature distribution also satisfies Laplace's equation. (1) Map M onto a strip as in Example 2. What is the distribution of temperature in the domain M if the temperature on the larger arc is 1 and 0 on the smaller one? What shape do the lines of constant temperature have?

(2) What is the distribution of temperature in the domain and with the boundary conditions described in example 2? Draw an approximate picture of the lines of

29 The Riemann Mapping Theorem

Using elementary transformations we can conformally map a limited family of domains onto $\mathbb{D}$; with the maps we used in the previous section, the boundary is always very simple. The Schwarz–Christoffel formulas in §30, generally nonelementary, provide conformal maps between any polygons and the upper half plane (which, as we saw, can be mapped onto the unit disk). For polygons with more than four sides, some numerical calculations are needed. Conformally mapping more general domains, even in a computer-assisted fashion, is a difficult problem.

In principle, however, **any simply connected domain other than $\mathbb{C}$** can be mapped onto $\mathbb{D}$:

Theorem 153: Riemann Mapping Theorem

Given any simply connected domain $\mathcal{D}$ other than $\mathbb{C}$, there exists a biholomorphic map f between $\mathcal{D}$ and $\mathbb{D}$.

Normalizing f by $f(z_0) = 0$ and $f'(z_0) \in \mathbb{R}^+$ for some $z_0 \in \mathcal{D}$, such a map is unique.

Riemann stated this theorem in 1851 for domains with piecewise regular boundary, and gave a proof based on solving a Dirichlet problem for the Laplacian. The method, *Dirichlet's principle*, contained a gap: the existence of a minimizer for the associated variational problem was not established. This gap was pointed out by Weierstrass in 1870.

In full generality, the theorem finally received a rigorous proof in 1900 by Osgood. Further important proofs were given by Koebe (1907) and Carathéodory (1912).

The next two pages, reproduced from [source, see header], describe Riemann's original proof in detail.

holomorphic function F on the set with $F' = f$. We shall say then that U is *holomorphically simply connected*.

It is easy to show that topological simple connectivity as defined implies the holomorphic antiderivative property just described. The theorem itself in the following form will show among other things the converse, namely, that the holomorphic simple connectivity implies topological simple connectivity.

Theorem 1.1 (Riemann) *Suppose that U is a connected open subset of $\mathbb{C}$ with $U \neq \mathbb{C}$. If U is holomorphically simply connected, then U is biholomorphic to the unit disc, i.e., there is a one-to-one holomorphic function from U onto the unit disc $D = \{z \in \mathbb{C} : |z| < 1\}$.*

In the proof of this result, it will be useful to be able to assume that U is bounded. For this, we recall the familiar fact that such a U as in the theorem is always biholomorphic to a bounded open set. The proof of this in summary form goes like this. Since $U \neq \mathbb{C}$, we can replace U by a translate to suppose that $0 \notin U$. The function $1/z$ is then holomorphic on U and hence has an antiderivative $L(z)$, say. Changing L by an additive constant will arrange that $\exp(L(z)) = z$ (this is the usual process for finding complex logarithms). Then $\exp(L(z)/2)$ is one-to-one on U . Choose an open disc in the image of $\exp(L(z)/2)$. The negative of this disc is disjoint from the image of $\exp(L(z)/2)$. So the image of $\exp(L(z)/2)$, which is biholomorphic to U , is itself biholomorphic to a bounded open subset of $\mathbb{C}$, via a linear fractional transformation.

Note that it is not clear by definition that the holomorphic simple connectivity is preserved by a biholomorphic mapping since the meaning of taking the derivative is different when the coordinates change; but by the complex chain rule this is a matter of a holomorphic factor which can be assimilated into the original function. Checking the details of this is left to the reader as an exercise.

So now we can assume without loss of generality that the open set U is bounded. And by translation we can now assume $0 \in U$. We shall look for a biholomorphic mapping from U to the unit disc D which takes 0 to 0. Of course, if there is a biholomorphic map from U to the unit disc at all, there is one that takes 0 to 0 since a linear fractional transformation taking the unit disc D to itself will take any given point to the origin, and in particular the image of 0 to begin with can be moved to the origin.

Now if $H:U \rightarrow D$ is biholomorphic and has $H(0) = 0$, then $H(z)/z$ has a removable singularity at 0. Hence, H can be written as $zh(z)$, where h is holomorphic on U and $h(0) \neq 0$. That $h(0) \neq 0$ follows because H is supposed to be one-to-one and hence must have derivative vanishing nowhere. Of course, $h(z)$ is also nonzero for every other $z \in U$ because 0 is the only point of U with $H(z) = 0$.

Now there is an antiderivative L of h'/h on U . The product $h(z) \exp(-L(z))$ is constant since it has derivative identically equal to 0 and U is connected. Changing L by an additive constant, we can assume $h(z) \exp(-L(z)) = 1$ for all $z \in U$. (This familiar argument will occur several times here).

The essential point of Riemann's method was to consider the harmonic function $\operatorname{Re} L(z)$. This is of course equal to $\ln |h(z)|$. Since the "boundary values" of $|H(z)|$ have to be 1, it must be that the boundary value of $|h|$ at a boundary point z_0 of U has to be

$1/|z_0|$. In particular, the harmonic function $\ln |h|$ has to have boundary value at z_0 equal to $-\ln |z_0|$.

At this point, Riemann appealed to what he referred to as the Dirichlet Principle. The Euler–Lagrange equation for the variational problem of minimizing the so-called Dirichlet (energy) integral for a real-valued function $f(x, y)$, namely minimizing this integral

$$\int_U \left[\left(\frac{\partial f}{\partial x} \right)^2 + \left(\frac{\partial f}{\partial y} \right)^2 \right] dx dy$$

under the condition that $f = g$ on the boundary ∂U of U , is easily computed to satisfy $\Delta f = 0$ (§18 of [10, 11]).

So Riemann proposed that the harmonic function with the boundary values $-\ln |z_0|$ at each boundary point z_0 could be found by minimization of the Dirichlet integral. And Riemann was well aware of how to construct h and hence H from knowing $\ln |h(z)|$. Riemann actually expressed this all in terms of $\ln |H|$ and the idea of Green's function, a function with boundary value 0 and a specified singularity at (in our case) the point 0, namely the function had to be of the form $\ln |z| + u(z)$ with u harmonic near the point 0. This is equivalent for open sets in $\mathbb{C}$ to our discussion, though the Green's function notion is useful when one tries to extend the Riemann Mapping Theorem to the uniformization problem where there is no *a priori* global z -coordinate.

The main difficulty is that there is no particular reason to suppose that there is in fact any minimum for the Dirichlet integral in this situation. There is also a less serious difficulty of explaining why the resulting function is one-to-one and onto—intuitively this is just a matter of winding numbers if one can approximate U from inside by domains with smooth or piecewise-smooth closed curve boundaries. One supposes that Riemann may have taken this part for obvious, though it is actually quite subtle if one does not appeal to any pre-existing topological intuitions. We shall give a precise argument later on. Riemann apparently considers only domains, the boundary of which is smooth in some sense. Osgood made the major forward step treating simply connected open sets in general, thus proving what we call today the Riemann Mapping Theorem. The Osgood proof is acknowledged directly by Carathéodory [2] where the ideas involved in the usual proof of today, via normal families, are presented. See the footnote (**) of page 108 of [2]. But, for some reason, Osgood's proof fell from favor or even recognition for the history of the Theorem in [9]; there is a reference to Osgood's paper but no comment on it, no acknowledgment that this is in fact the reasonably complete first proof of the general result.

2 The application of the Perron method

Even simply connected bounded open sets in $\mathbb{C}$ can have complicated boundaries. The boundary of the Koch snowflake, for example, has Hausdorff dimension greater than 1 [15]. And Osgood [6] already gave an example with boundary having positive (2-dimensional) measure. Thus, it is appropriate to introduce carefully what is to be meant by finding functions with specified boundary values. For this purpose, let U be a bounded open set in $\mathbb{C}$ and ∂U be its boundary, that is the complement of U within the closure $\text{cl}(U)$ of U in $\mathbb{C}$, or equivalently the intersection of $\text{cl}(U)$ with $\mathbb{C} - U$. Suppose that $b: \partial U \rightarrow \mathbb{R}$ is a continuous function. Then we say that a harmonic function $h: U \rightarrow \mathbb{R}$ is

29.1 Equicontinuity and Normal Families

We consider functions $f : M \rightarrow M'$, where M and M' are metric spaces with metrics d and d' , respectively. Recall that f is said to be *uniformly continuous* if the constants involved in the continuity condition are independent of the point. More precisely,

$$\forall \varepsilon > 0 \exists \delta > 0 \left(\forall (z, z_0) \in M^2, d(z, z_0) < \delta \Rightarrow d'(f(z), f(z_0)) < \varepsilon \right). \quad (29.10)$$

We can assume that the metric d' is a **bounded function**, for we can always replace it by

$$d'' = \frac{d'}{1 + d'},$$

which is a metric topologically equivalent to d' , with $d'' \leq 1$. (This standard trick allows us to avoid unbounded distances.)

Definition 154: Equicontinuity

A family $\mathcal{F}$ of continuous functions $f : M \rightarrow M'$, where M and M' are metric spaces with metrics d and d' , is said to be *equicontinuous* if the continuity modulus can be chosen independently of the function. Equivalently, for every $x \in M$ and every $\varepsilon > 0$, there exists $\delta(x, \varepsilon) > 0$ such that

$$d(x, y) < \delta(x, \varepsilon) \implies d'(f(x), f(y)) < \varepsilon \quad \text{for all } f \in \mathcal{F}. \quad (29.11)$$

Definition 155: Normal Families

Let M, M' be complete metric spaces. A collection of continuous functions $\mathcal{F}$ from M to M' is a **normal family** if it is pre-compact in the topology of uniform convergence on compact sets.^a

^aA subset of a metric space is called *pre-compact* (or *relatively compact*) if its closure is compact. Equivalently, every sequence in the set has a subsequence converging to a point in the ambient space.

Exhaustion by Compact Sets

If the metric space M is σ -compact, then by definition it can be covered by a countable nested family of compact sets. For instance, in $\mathbb{C}$ or $\mathbb{R}^n$ we may take

$$M = \bigcup_{n \in \mathbb{N}} K_n, \quad K_n = \{x \in M : |x| \leq n\}. \quad (29.12)$$

Metrizability of Uniform Convergence on Compact Sets

On each K_n in (29.12) we define the distance between two functions f, g analogously to the L^∞ norm:

$$\delta_n(f, g) = \sup_{x \in K_n} d''(f(x), g(x)). \quad (29.13)$$

We then define a distance between the functions f and g by

$$\rho(f, g) = \sum_{n=1}^{\infty} 2^{-n} \delta_n(f, g). \quad (29.14)$$

Since $d'' \leq 1$, the sum converges. This construction metrizes the topology of uniform convergence on compact sets. The details are left as an exercise, below.

Exercise 156. Let M be a σ -compact metric space. Check that ρ in (29.14) is a metric on $C(M)$. Check that convergence with respect to ρ is equivalent to uniform convergence on compact sets. Finally, check that $C(M)$ is complete with respect to ρ if M' is a complete metric space.

Proposition 29.3. A family $\mathcal{F}$ is **normal** if and only if its closure $\overline{\mathcal{F}}$ with respect to ρ is compact in $C(M)$.

29.1.1 The Ascoli–Arzelà Theorem

Theorem 157: Ascoli–Arzelà

A family $\mathcal{F}$ of continuous functions in the region $\Omega \subset \mathbb{C}$ with values in a metric space M' is normal in Ω if and only if the following conditions are satisfied:

- (i) $\mathcal{F}$ is equicontinuous at every $x \in \Omega$.
- (ii) $\mathcal{F}$ is equibounded: for every $z \in \Omega$ there exists a compact set $K_1 = K_1(z) \subset M'$ such that $f(z) \in K_1$ for all $f \in \mathcal{F}$.

This is a standard theorem; we leave the proof for the Appendix, see §157.

Proposition 29.4. Let $M' \subset \mathbb{C}$ and $\mathcal{F}$ be a normal family from Ω to M' . Let $K \subset \Omega$ be compact. Then $\mathcal{F}$ is uniformly bounded on K :

$$\sup_{z \in K, f \in \mathcal{F}} |f(z)| = m < \infty. \quad (29.15)$$

Proof. Since $\mathcal{F}$ is a normal family, by Theorem 157 and equicontinuity, it follows that for any point a we can find $\delta(a) > 0$ such that

$$|a - b| < \delta(a) \Rightarrow |f(a) - f(b)| < 1, \quad \forall f \in \mathcal{F}. \quad (29.16)$$

Cover K by finitely many disks $\mathbb{D}_{\delta(a_j)}(a_j)$, and let δ_0 be the minimum radius in this cover. For each j , set $m_j = \sup\{|f(a_j)| : f \in \mathcal{F}\}$ and let $m = 1 + \max_j m_j$. Then, for any $x \in K$, there exists a_j with $|x - a_j| < \delta_0$, and for all $f \in \mathcal{F}$ we have

$$|f(x)| \leq |f(x) - f(a_j)| + |f(a_j)| < 1 + m_j \leq m, \quad (29.17)$$

as desired. ■

Theorem 158: (Montel's theorem)

Let $\mathcal{D} \subset \mathbb{C}$ be a domain and let $\mathcal{F}$ be a family of analytic functions on $\mathcal{D}$ such that for every compact set $K \subset \mathcal{D}$,

$$\sup\{|f(z)| : z \in K, f \in \mathcal{F}\} = m(K) < \infty.$$

Then $\mathcal{F}$ is a normal family on $\mathcal{D}$.

Proof. The family is clearly equibounded. We show that the derivatives are also equibounded, from which equicontinuity follows.

Let $K \subset \mathcal{D}$ be compact. Take $r < \text{dist}(K, \partial\mathcal{D})$, and cover K by finitely many disks $\mathbb{D}_r(a_j)$. Let $O = \bigcup_{j=1}^m \mathbb{D}_r(a_j)$ and set $r' = \text{dist}(\partial O, K)$. Then $\bar{O} \subset \mathcal{D}$ and $r' > 0$.

By Cauchy's integral formula, for $z \in K$ we have

$$|f'(z)| = \left| \frac{1}{2\pi i} \oint_{\partial O} \frac{f(s)}{(s-z)^2} ds \right| \leq \frac{1}{2\pi} \cdot \frac{2\pi \cdot m(K)}{(r')^2} \cdot \text{length}(\partial O) \leq \frac{2m(K)}{r'}, \quad (29.18)$$

so $|f'(z)|$ is uniformly bounded on K . Equicontinuity follows, and the result is proved. ■

29.2 The Riemann Mapping Theorem

Definition 159

An injective analytic function is called *univalent* (or *schlicht*).

Theorem 160: The Riemann Mapping Theorem

Let $\Omega \subset \mathbb{C}$ be a nonempty simply connected domain other than $\mathbb{C}$ itself, and let $z_0 \in \Omega$. Then there exists a unique biholomorphism $\varphi : \Omega \rightarrow \mathbb{D}$ satisfying the normalization conditions

$$\varphi(z_0) = 0, \quad \varphi'(z_0) \in \mathbb{R}^+.$$

Note 161

The exclusion of $\mathbb{C}$ is necessary because any bounded entire function is constant (by Liouville's Theorem), and hence no biholomorphism from $\mathbb{C}$ to $\mathbb{D}$ can exist.

Note 162

As a consequence, any two nonempty simply connected domains $\mathcal{D}_1$ and $\mathcal{D}_2$ different from $\mathbb{C}$ are conformally equivalent. Indeed, if $\varphi_i : \mathcal{D}_i \rightarrow \mathbb{D}$ are conformal maps for $i = 1, 2$, then the composition $\varphi_2^{-1} \circ \varphi_1$ is a conformal map from $\mathcal{D}_1$ onto $\mathcal{D}_2$. (See also Exercise 142.)

Proof of the Riemann Mapping Theorem.

Uniqueness. This part is easier. If φ_1 and φ_2 are two functions with the stated properties, then $S := \varphi_1(\varphi_2)^{-1}$ is an automorphism of $\mathbb{D}$ and $S(0) = 0$. By Schwarz's lemma $S(z) = ze^{i\theta}$ for some θ . $S'(0) > 0$ implies $\theta = 0$.

Existence.

Note 163

Since conformal equivalence is transitive, we may replace the given domain $\mathcal{D}$ by any domain conformally equivalent to it. We shall therefore begin by simplifying $\mathcal{D}$ through a sequence of elementary transformations.

Furthermore, because connectedness and simple connectedness are topological invariants,^a and openness is preserved by the open mapping theorem, if Ω satisfies the hypotheses of the Riemann Mapping Theorem and is conformally equivalent to Ω' , then Ω' does as well.

We shall show that Ω is conformally equivalent to a bounded domain, which, by an

affine transformation, can be mapped onto the interior of the unit disk with z_0 sent to 0. Hence, without loss of generality, we may assume that Ω itself is a simply connected domain contained in $\mathbb{D}$ and that $z_0 = 0$.

^aIf γ is a path joining z_1 to z_2 and f is conformal, then $f(\gamma)$ joins $f(z_1)$ to $f(z_2)$. Since f is biholomorphic (hence bicontinuous), any path γ in $f(\Omega)$ can be lifted to a path $\gamma' = f^{-1} \circ \gamma$ in Ω . As Ω is simply connected, there exists a continuous map $g : [0, 1]^2 \rightarrow \Omega$ that shrinks γ' to a point, and since f is continuous, $f \circ g$ shrinks γ to a point.

Lemma 164. Assume $\Omega \not\equiv 0$ satisfies the conditions of Theorem 160. Then:

- (i) There exists a biholomorphic branch of the logarithm between Ω and $L = \log(\Omega)$ (with $\exp$ as its inverse, of course). The image $L(\Omega)$ is open by the open mapping theorem.
- (ii) $L \cap (L + 2\pi i) = \emptyset$.
- (iii) L omits an open disk.

Proof. (i) Fix a basepoint $a \in \Omega$, and define

$$\text{Log}(z) = \int_1^a \frac{ds}{s} + \int_a^z \frac{ds}{s}.$$

This defines an analytic branch of the logarithm in Ω , since any closed curve in Ω can be homotopically shrunk to a point without crossing 0 (which lies outside Ω). If $\log z_1 = \log z_2$, then exponentiating yields $z_1 = z_2$, establishing injectivity. The inverse mapping theorem completes the proof of (i).

(ii) Suppose that w and $w - 2\pi i$ both lie in L . Then $w = \log z_1$ and $w + 2\pi i = \log z_2$ for some $z_1, z_2 \in \Omega$. Exponentiating both expressions gives $z_1 = z_2$, a contradiction.

(iii) Let $w_0 \in L + 2\pi i$. Since $L + 2\pi i$ is open, there exists $\varepsilon > 0$ such that $\mathbb{D}_\varepsilon(w_0) \subset L + 2\pi i$, so $\mathbb{D}_\varepsilon(w_0)$ is an omitted disk in L . ■

Note 165

If w_0 is the center of an omitted disk, then the function $f(w) = \frac{1}{w-w_0}$ is conformal and bounded in L , and therefore $f(L)$ is bounded. Based on Note 163, we may assume, without loss of generality, that $\Omega \subset \mathbb{D}$ and that $z_0 = 0$.

Preview of the rest of the proof.

Note 166

1. Let φ be any biholomorphism between Ω and $S \subset \mathbb{D}$. We want to understand how $\varphi'(0)$ relates to the “size” of S , in the sense of set inclusion. Suppose for the moment that the Riemann Mapping Theorem is already known. Let $\varphi_1 : \Omega \rightarrow S_1$ and $\varphi_2 : \Omega \rightarrow S_2$ with $S_1 \subsetneq S_2$. Consider the following chains of biholomorphisms:

$$\mathbb{D} \xrightarrow{\psi} \Omega \xrightarrow{\varphi_2} S_2 \xrightarrow{\chi} \mathbb{D}, \quad (29.19)$$

normalized so that $\chi \circ \varphi_2 \circ \psi$ vanishes at 0 with positive derivative there. By Schwarz’s lemma it must be the identity. Similarly,

$$\mathbb{D} \xrightarrow{\psi} \Omega \xrightarrow{\varphi_1} S_1 \xrightarrow{\chi} \tilde{S}_1 \subsetneq \mathbb{D}. \quad (29.20)$$

The strict inclusion follows from $S_1 \subsetneq S_2$. By the chain rule and Schwarz's lemma, we conclude that $\varphi'_1(0) \leq \varphi'_2(0)$.

Hence, our strategy will be to maximize $\varphi'(0)$.

2. Next, we consider the supremum M of the values $\varphi'(0)$ over all such maps. Since $\varphi_n(\mathbb{D}) \subset \mathbb{D}$, Montel's theorem applies and there exists a subsequence $\varphi_n \rightarrow \varphi$ with $\varphi'(0) = M$. We then show that $\varphi(\Omega) = \mathbb{D}$: if not, one can construct a new map φ_1 with $\varphi'_1(0) > M$ by composing with a suitable branch of the square root, see below, contradicting maximality.

3. Since, for $z \in \mathbb{D}$, we have $0 < |z| < 1$, it follows that $|z| < \sqrt{|z|} < 1$; that is, the map $z \mapsto \sqrt{z}$ is *expanding*.

If $\Omega \neq \mathbb{D}$, one can select a point outside $\varphi(\Omega)$, choose a branch of the square root, and enlarge $\varphi(\Omega)$ by applying this branch, thereby increasing $\varphi'(0)$.

Note 167

After applying the inversion $z \mapsto 1/z$, the quantity $\varphi'(0)$ acquires an important interpretation: it measures the *logarithmic capacity* (or transfinite diameter) of the domain mapped by φ onto the unit disk. Logarithmic capacity is a set function that is *monotone increasing* with respect to set inclusion.

Definition 168

Let

$$\mathcal{F} = \left\{ \varphi \mid \varphi : \Omega \longrightarrow \varphi(\Omega) \subset \mathbb{D}, \varphi(0) = 0, \varphi'(0) > 0, \varphi \text{ biholomorphic} \right\}.$$

Proposition 29.5. (i) $M := \sup\{f'(0) : f \in \mathcal{F}\}$ is attained by some $F \in \mathcal{F}$.

(ii) If $F(\Omega) \neq \mathbb{D}$, then there exists $F_1 \in \mathcal{F}$ with $F'_1(0) > M$, a contradiction which completes the proof of the theorem.

Proof. (i) For any $f \in \mathcal{F}$ and $z \in \Omega$, we have $|f(z)| < 1$. By Montel's Theorem 29.4, $\mathcal{F}$ is a normal family. Thus, if $f_n \in \mathcal{F}$ with $f'_n(0) = m_n \rightarrow M$, then a subsequence converges uniformly on compacta to some F . By Weierstrass's theorem 48, F is analytic and $F'(0) = M$. Hurwitz's theorem 93 ensures F is injective: indeed, F cannot be constant since $F'(0) = M \neq 0$, and if $F(z_1) = F(z_2)$ for some $z_1 \neq z_2$, then $F(z) - F(z_1)$ would vanish in $\Omega^* = \Omega \setminus \{z_1\}$, contradicting Hurwitz's theorem.

(ii) Suppose $\mathbb{D} \setminus F(\Omega) \neq \emptyset$, and let $a \in \mathbb{D} \setminus F(\Omega)$. First apply the automorphism

$$f_1(z) = \frac{F(z) - a}{1 - \bar{a}F(z)}, \quad (29.21)$$

which maps Ω biholomorphically into $\mathbb{D} \setminus \{0\}$. On $f_1(\Omega)$ we construct, as before, an analytic branch of $\log$ ⁸, and from it an analytic branch of the square root. Defining

$$f_2(z) = \sqrt{f_1(z)}, \quad z \in \Omega, \quad (29.22)$$

⁸This time around the log may be unbounded since 0 may be on the boundary of $f_1(\Omega)$; this will be inconsequential since we are only using it to define a consistent branch of the square root.

we see that f_2 is biholomorphic and that $f_2(\Omega) \subset \mathbb{D}$. Moreover, $f_2(0) = b = \sqrt{-a} \neq 0$.

Now apply another automorphism of the disk to move b to 0 and normalize the derivative:

$$F_1(z) = \frac{f_2(z) - b}{1 - \bar{b}f_2(z)} \cdot \frac{|f_2'(0)|}{f_2'(0)}, \quad z \in \Omega. \quad (29.23)$$

A direct calculation gives

$$F_1'(0) = \frac{1 + |a|}{2\sqrt{|a|}} F'(0) > F'(0) = M, \quad (29.24)$$

a contradiction. ■

Note 169

This proof is nonconstructive: it does not provide a concrete sequence realizing the extremal map F , nor a direct procedure for obtaining F . The square-root trick ensures that any failure of surjectivity would yield a larger derivative at 0, but in practice the convergent sequence with $f_n'(0) \rightarrow M$ cannot be prescribed explicitly.

29.3 Behavior at the boundary of biholomorphisms: a general but weaker result

We now establish a weaker, but easier and useful fact [3]: if $\mathcal{D}$ is a simply connected domain and φ maps it conformally onto $\mathbb{D}$, then $\varphi(z)$ approaches $\partial\mathbb{D}$ whenever z approaches $\partial\mathcal{D}$, in the sense defined below (which does not necessarily imply that $\varphi(z)$ converges to a boundary point).

Definition 170

Let $\mathcal{D}$ be a domain.

- A sequence $\{z_n\} \subset \mathcal{D}$ satisfies $z_n \rightarrow \partial\mathcal{D}$ as $n \rightarrow \infty$ if for every compact set $K \subset \mathcal{D}$ there exists n_0 such that $z_n \notin K$ for all $n > n_0$.

- Similarly, for an arc $\gamma : [0, 1] \rightarrow \mathcal{D}$ we say that $\gamma(t) \rightarrow \partial\mathcal{D}$ as $t \rightarrow 1$ if for every compact set $K \subset \mathcal{D}$ there exists $t_0 \in (0, 1)$ such that $\gamma(t) \notin K$ for all $t > t_0$.

30 Boundary behavior.

Carathéodory's theorem (see [7, pp. 110–118]), whose proof would take us some time, states that if $\mathcal{D}$ and $\mathcal{D}'$ are Jordan regions (that is, domains bounded by Jordan curves), then the conformal map between $\mathcal{D}$ and $\mathcal{D}'$ extends to a homeomorphism between $\overline{\mathcal{D}}$ and $\overline{\mathcal{D}'}$.

Theorem 171

If $\varphi : \mathcal{D} \rightarrow \mathcal{D}'$ is a domain biholomorphism and $z_n \rightarrow \partial\mathcal{D}$, as $n \rightarrow \infty$ then $\varphi(z_n) \rightarrow \partial\mathcal{D}'$ as $n \rightarrow \infty$.

Proof. We prove the statement for sequences; the one for arcs is almost identical. Since φ is biholomorphic, any compact covering of $\mathcal{D}$ generates a compact covering of $\mathcal{D}'$ and vice-versa. Let $z_n \rightarrow \partial\mathcal{D}$ and let $K' \subset \mathcal{D}'$ be any compact set and $K = \varphi^{-1}(K')$. By definition,

$$z_n \rightarrow \partial\mathcal{D} \Rightarrow \exists n_0 \text{ s.t. } \forall n > n_0 \quad z_n \notin K$$

Since φ is one to one, for any $n > n_0$, $\varphi(z_n) \notin K'$ either.

Corollary 172. If $\varphi : \mathcal{D} \rightarrow \mathbb{D}$ is a biholomorphism, then $|\varphi(z)| \rightarrow 1$ as $z \rightarrow \partial\mathcal{D}$.

30.1 A reflection principle for harmonic functions

Let $\mathcal{D}^+$ be a domain in the upper half-plane (UHP) such that $\partial\mathcal{D}^+ \supset I := [a, b] \subset \mathbb{R}$. We denote by $\mathcal{D}_-$ the reflection of $\mathcal{D}^+$ across I .

Theorem 173

Assume v is harmonic in $\mathcal{D}^+$, continuous on $\mathcal{D}^+ \cup I$, and satisfies $v = 0$ on I . Then v extends to a harmonic function on

$$\mathcal{D} = \mathcal{D}^+ \cup I \cup \mathcal{D}_-.$$

Note 174

We cannot simply invoke the Schwarz reflection principle to prove this **stronger** theorem: nothing is known *a priori* about the continuity on $\mathcal{D}^+ \cup I$ of u , the harmonic conjugate of v . In fact, the Schwarz reflection principle itself follows as a consequence of Theorem 173; see Exercise 175.

Proof. (See [13].) As in the Schwarz reflection principle, we define the extension of v by setting $V(z) = v(z)$ for $z \in \mathcal{D}^+$, and

$$V(\bar{z}) := -v(z), \quad \forall \bar{z} \in \mathcal{D}_-; \quad (\text{i.e., } V(x, -y) = -V(x, y)), \quad (30.1)$$

and $V(z) = 0$ on I .

The property of being harmonic is a local one, so it suffices to verify that V is harmonic in every disk contained in $\mathcal{D} = \mathcal{D}^+ \cup I \cup \mathcal{D}_-$ for some family of disks covering $\mathcal{D}$. For disks in $\mathcal{D}^+$ this holds by assumption, while for disks contained in $\mathcal{D}_-$ it follows directly from (30.1).

Now consider a disk $\mathbb{D}_\varepsilon$ whose diameter lies on I . Laplace's equation $\Delta v = 0$ with continuous boundary condition $V(z)$, $z \in \partial\mathbb{D}_\varepsilon$, has a unique solution v_1 . Because $V(\bar{z}) = -V(z)$ on $\partial\mathbb{D}_\varepsilon$, we have $v_1(z) = -v_1(\bar{z})$ ⁹; in particular, $v_1(\mathbb{D}_\varepsilon \cap I) = \{0\}$. Both V and v_1 satisfy Laplace's equation in the interior of the upper half-disk $\mathbb{D}_\varepsilon \cap \mathcal{D}^+$, with identical boundary data $V(z)$ on $\partial\mathbb{D}_\varepsilon \cap \mathcal{D}^+$ and zero on $\mathbb{D}_\varepsilon \cap I$; hence, by uniqueness, $v_1 = V$ in the closed upper half-disk. By the same argument, $v_1 = V$ in the closed lower half-disk, and therefore $V = v_1$ is harmonic in the entire disk $\mathbb{D}_\varepsilon$. ■

Exercise 175. Show that if $f = u + iv$ is analytic in a domain $\mathcal{D}^+$ as in the theorem and v is zero on I , then u is continuous at I . In particular, for the Schwarz reflection principle for analytic functions f in a domain $\mathcal{D}$ in the UHP whose closure contains a real segment I , we only need to show that $\text{Im } f$ is continuous and vanishes on I .

Exercise 176. Assume that $\mathcal{D}$ is a simply connected domain in the UHP such that $\bar{\mathcal{D}} \supset I \subset \mathbb{R}$. Let φ be the conformal map from $\mathbb{D}$ to $\mathcal{D}$. Consider a half disk strictly contained in $\mathcal{D}$ with diameter on I as in the figure. The arc is the image of an analytic curve in $\mathbb{D}$ bounding a domain $\mathcal{D}'$ in $\mathbb{D}$. Show that $\text{Im } \varphi(z) \rightarrow 0$ as $z \rightarrow \partial\mathbb{D} \cap \bar{\mathcal{D}}'$ and apply Theorem 173 and Exercise 175 to show that φ has analytic continuation to the reflection of $\mathcal{D}$ through I . In particular φ^{-1} extends continuously to I , and φ maps part of $\partial\mathbb{D}$ to I .

⁹Let $\tilde{y} = -y$. Then $-V(x, -y) = -V(x, \tilde{y})$ and

$$-\frac{\partial^2}{\partial y^2} V(x, -y) = -\frac{\partial^2}{\partial \tilde{y}^2} V(x, \tilde{y}) = \frac{\partial^2}{\partial x^2} V(x, \tilde{y}),$$

so the Laplace equation $\Delta V = 0$ remains valid under reflection.

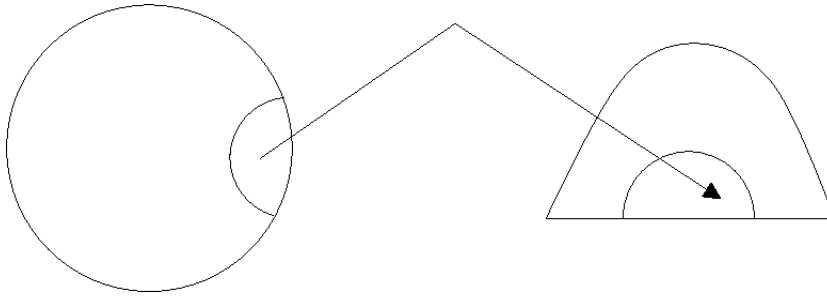


Figure 29: Boundary behavior across line segments.

Note 177

The following ingenious construction can be found in Ahlfors [3, p. 27]. If f is analytic, then so is its *reflection*,

$$f^*(z) := \overline{f(\bar{z})}.$$

We observe that

$$2u(x, y) = f(x + iy) + \overline{f(x - iy)} = f(z) + f^*(z), \quad (30.2)$$

for all $z = x + iy$ in the region where the right-hand side is defined.

Assuming further that the expression $u(z/2, z/2i)$ is meaningful^a, we find

$$\begin{aligned} 2u\left(\frac{z}{2}, \frac{z}{2i}\right) &= f(z) + f^*(0) \\ &= f(z) + C \quad \Rightarrow \quad f(z) = 2u\left(\frac{z}{2}, \frac{z}{2i}\right) - C, \end{aligned} \quad (30.3)$$

where $C = f^*(0)$ is a constant.

Consequently,

$$v(z) = \operatorname{Im} \left(2u\left(\frac{z}{2}, \frac{z}{2i}\right) - C \right).$$

Thus, at least in simple cases (for instance, for rational functions), one can recover f and v directly from u , without performing any integration. Can you extend this construction to arbitrary analytic functions?

^aSuch is the case, for instance, if u is a polynomial.

30.2 Analytic arcs

Definition 178

Let $I = [0, 1]$. A *proper analytic arc* is the image of I under a function γ which extends analytically and injectively to a neighborhood of I in $\mathbb{C}$.

Equivalently, a proper analytic arc is one whose parametrization

$$\gamma = \Gamma(I) \quad \text{where } \Gamma : I \longrightarrow \mathbb{C}$$

extends to a biholomorphism between a neighborhood of I and its image.

30.3 An extension of the Schwarz reflection principle

Informally, the principle states that if f is analytic in a domain $\mathcal{D}$ which contains an analytic

arc γ , if f is continuous up to γ , and if the image of γ is itself an analytic arc, then f extends analytically beyond γ .

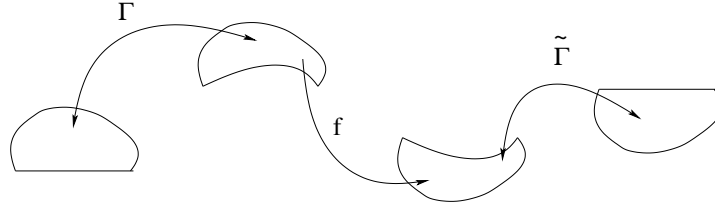


Figure 30: Analytic continuation across an analytic arc.

Definition 179: One-sided boundary

Let γ be a proper analytic arc bounding a domain $\mathcal{D}$. We say that γ is a *one-sided boundary* of $\mathcal{D}$ if for every $z_0 \in \Gamma(I)$, there exists a disk $\mathbb{D}_\varepsilon(z_0)$ such that

$$\Gamma^{-1}(\mathbb{D}_\varepsilon(z_0) \cap \mathcal{D})$$

lies entirely in either the upper half-plane or the lower half-plane (independently of z_0).^a

In this case, we also say that $\mathcal{D}$ lies *on one side* of γ . As usual, by analytic continuation of a function f across a curve, we mean the existence of an analytic function $\hat{f}$ in a neighborhood of the curve which coincides with f wherever both are defined. Note that such analytic continuation is necessarily unique.

^aRecall that Γ extends to a biholomorphism in a neighborhood of I .

The following result may be viewed as the “conformally mapped” Schwarz reflection principle.

Theorem 180: Analytic reflection across arcs

Let $\mathcal{D} \subset \mathbb{C}$ be a simply connected domain, and assume that $\mathcal{D}$ lies on one side of a proper analytic arc γ . Let f be analytic in $\mathcal{D}$ and continuous on $\mathcal{D} \cup \gamma$. Assume that $f(\gamma) \subset \tilde{\gamma}$, where $\tilde{\gamma}$ is a proper analytic arc, and that $f(\mathcal{D})$ lies on one side of $\tilde{\gamma}$. Then f extends analytically across γ .

Proof. By assumption, γ and $\tilde{\gamma}$ are the images of closed intervals I and $\tilde{I}$ under biholomorphisms Γ and $\tilde{\Gamma}$, respectively. The function f is analytic in $\mathcal{D}$ and continuous up to γ , while Γ and $\tilde{\Gamma}$ are biholomorphic in neighborhoods of I and $\tilde{I}$.

Without loss of generality, we may assume that I is approached from above. Define

$$\psi := \tilde{\Gamma}^{-1} \circ f \circ \Gamma.$$

Then ψ is analytic in a domain in the upper half-plane, continuous on I , and real-valued on I ; hence, the classical Schwarz reflection principle applies to ψ . The desired analytic continuation of f is therefore given by

$$\hat{f} := \tilde{\Gamma} \circ \psi^* \circ \Gamma^{-1},$$

where ψ^* denotes the analytic continuation of ψ reflected across I . ■

Exercise 181. Let $\mathcal{D}$ be a simply connected nonempty domain with nonempty complement and let φ be the conformal map of $\mathbb{D}$ to $\mathcal{D}$ assumed continuous up to the boundary. Show that φ has analytic continuation through an arc of $\partial\mathbb{D}$ iff the image of the arc is an analytic arc in $\partial\mathcal{D}$.

Theorem 182

Let v be harmonic in a domain $\mathcal{D}$, and let $\gamma \subset \partial\mathcal{D}$ be a proper analytic arc such that $\mathcal{D}$ lies on one side of γ . Assume that v is continuous up to γ and vanishes on γ . Then v extends to a harmonic function in an open set containing $\mathcal{D} \cup \gamma$.

Proof. Exercise: use Theorem 173 together with the strategy of Theorem 180. ■

Theorem 183

Let $\mathcal{D} \subsetneq \mathbb{C}$ be a simply connected domain, and let φ be a conformal map between $\mathcal{D}$ and $\mathbb{D}$. If $\partial\mathcal{D}$ contains a proper analytic arc γ and $\mathcal{D}$ lies on one side of γ , then

- (i) φ extends analytically across γ ;
- (ii) if for $n \in \mathbb{N}$, $z_n \in \mathcal{D}$, and $\lim_n z_n = z_0 \in \gamma$, then $\lim_n \varphi(z_n) = \varphi(z_0) \in \partial\mathbb{D}$;
- (iii) moreover, this extension is one-to-one on γ , hence also in a neighborhood of γ .

Note 184

By applying a linear fractional transformation, one obtains an analogous statement if $\mathbb{D}$ is replaced by a half-plane.

In particular, if the boundary of $\mathcal{D}$ is piecewise analytic, then φ extends analytically to a domain $\mathcal{D}' \supset \mathcal{D}$. (It is also biholomorphic in some domain $\mathcal{D}'' \supset \mathcal{D}$.)

Proof. (i) follows from Theorem 180.

(ii) follows immediately from (i).

(iii) The analytic extension in (i) implies that φ is continuous in a neighborhood of γ . If φ were not injective on an open subarc of γ , that portion would be mapped onto a closed curve, contradicting the continuity of φ near the endpoints of that arc. ■

30.3.1 Discussion: Tilings with Parallelograms

An elegant application of the Schwarz reflection principle in conformal mapping arises from the study of tilings by rectangles. Consider a conformal map

$$\varphi : P \longrightarrow \text{UHP},$$

where P is one of the rectangles in Figure 31. With a suitable normalization, φ coincides with the elliptic sine function sn , one of the Jacobi elliptic functions (to which we shall return later). By Theorem 180, φ admits analytic continuation across any side of the rectangle to its adjacent one.

By successive reflections across the sides of the rectangles, any point in the plane can be reached from a fixed rectangle. Although a given point can generally be obtained by several distinct sequences of reflections, the parity of the number of reflections (even or odd) is independent of the chosen path.

This observation implies that the function obtained by successive reflections provides a single-valued analytic continuation of φ to the entire complex plane $\mathbb{C}$, except at the corners of the

rectangles. The extended function is therefore *meromorphic*, since the point at infinity must correspond to a boundary point of P . Moreover, the analytically continued φ maps conformally each rectangle of the tiling onto the upper half-plane when the number of reflections is even, and onto the lower half-plane when it is odd. Because two successive reflections in parallel (horizontal or vertical) sides return each point to itself, the extended function is *doubly periodic*. Near any corner of the tiling where the mapping φ is bounded, it locally doubles the angle; hence, locally,

$$\varphi(z) = a + (z - z_0)^2(1 + o(1)).$$

At all corners except one, the continuation of φ is single-valued and bounded, and therefore analytic there. At the exceptional corner corresponding to iK' , however, φ has a simple pole; this is consistent with the fact that the image of iK' is the point at infinity on the boundary of the upper half-plane.

We will later show, more precisely, that for $k \in [0, 1]$, the function $\operatorname{sn}(z, k)$ gives a conformal map of the closed rectangle $[-K, K] \times [0, K']$ onto the upper half-plane, with

$$0, \pm K, \pm K + iK', iK' \longmapsto 0, \pm 1, \pm k^{-2}, \infty,$$

respectively, and that iK' is indeed a simple pole.

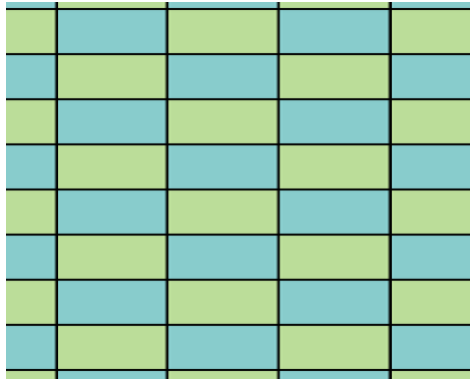


Figure 31: Tiling of the plane by reflections of a rectangle.

Source: https://upload.wikimedia.org/wikipedia/commons/3/33/Stacked_bond.png

30.4 Behavior at the Boundary: A Stronger Result

30.4.1 A Negative Result

The following classical example (see [15]) shows that, in full generality, continuity of a conformal mapping up to the boundary cannot be expected. Consider the case of an *inaccessible point*.

Let $\mathcal{D}$ be an open horizontal rectangle with a vertex at the origin, from which vertical line segments of length $1/2$ have been removed (see Fig. 32). It is easy to verify that 0 is indeed inaccessible. The image of 0 on $\partial\mathbb{D}$ (which, without loss of generality, we may take to be $z = 1$) must then be a point of discontinuity for the conformal map φ onto the unit disk. Indeed, $\varphi^{-1}(1) = 0$, yet every neighborhood of 1 contains infinitely many points whose preimage under φ lies at height $1/2$.

Interestingly, this is, in a precise sense, the only way in which boundary limits can fail to exist. The following result, due to Lindelöf and Koebe, makes this statement rigorous. We omit the proof for reasons of space.

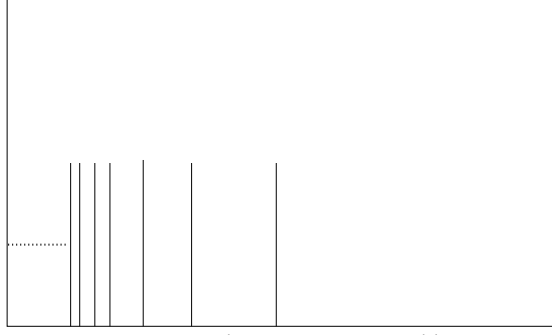


Figure 32: The origin is inaccessible.

Definition 185

A point $\zeta \in \partial\mathbb{D}$ is called *accessible* if there exists a sequence $\{z_n\}_{n \in \mathbb{N}} \subset \mathcal{D}$ with $\lim_{n \rightarrow \infty} z_n = \zeta$, and a continuous curve $\gamma : [0, 1] \rightarrow \overline{\mathcal{D}}$ such that $\gamma(0) \in \mathcal{D}$, $\gamma(1) = \zeta$, and γ passes through all z_n , that is, there exist $0 < t_1 < t_2 < \cdots < t_n < 1$ with $\lim_{n \rightarrow \infty} t_n = 1$ and $\gamma(t_n) = z_n$.

Theorem 186: Boundary Behavior

Let $\mathcal{D}$ be a bounded domain.

- (i) If z_0 is an accessible point of $\partial\mathbb{D}$, then the biholomorphic map $\varphi : \mathcal{D} \rightarrow \mathbb{D}$ has a finite limit (denoted $\varphi(z_0)$) as $z \rightarrow z_0$, $z \in \mathcal{D}$, and $\varphi(z_0) \in \partial\mathbb{D}$.
- (ii) If z_1 and z_2 are distinct accessible points of $\partial\mathbb{D}$, then $\varphi(z_1) \neq \varphi(z_2)$.

Illustration: nonanalytic boundary behavior for a conformal image of the unit disk

Exercise 187. Show that:

- The functional equation

$$G(z^2) = \lambda^{-1} \frac{G(z)^2}{1 - G(z)}, \quad z \in \mathbb{D}, \quad G'(0) = 1, \quad (30.4)$$

with $\lambda = \frac{i}{2}$, defines a unique function G analytic on the unit disk $\mathbb{D}$.

- The points in $G(\partial\mathbb{D})$ are exactly those $w \in \mathbb{C}$ for which the forward orbit of the one-step recurrence

$$x_{n+1} = \lambda x_n(1 - x_n), \quad x_0 = w,$$

converges to 0.

- The map $G : \mathbb{D} \rightarrow G(\mathbb{D})$ is conformal.

Figure 33 depicts the domain $G(\mathbb{D})$. Its boundary $J = \partial G(\mathbb{D})$ is the *Julia set* for iteration of the quadratic map $x \mapsto \lambda x(1 - x)$.

Theorem 188

The unit circle $\partial\mathbb{D}$ is a *natural boundary* for G .

Sketch of proof. Assuming 1 is a point of continuity for G on $\partial\mathbb{D}$, the functional equation yields

$$G(1) = \frac{\lambda}{1 + \lambda}. \quad (30.5)$$

We show that, at every binary rational angle (i.e., points of the form $e^{2\pi i n/2^k}$, $n \in \mathbb{Z}$), either the nontangential limit of G' is 0 or it does not exist; in either case G cannot extend analytically across any such point.

Differentiating (30.4) and passing to the boundary along radii, one sees that if $G'(1) = 0$, then by repeated use of the functional equation (doubling angles) G' must vanish at *all* binary rational angles.

If instead $G'(1)$ exists and is nonzero, normalize by dividing the differentiated identity by $G'(1)$ at $z = 1$ to obtain, in contradiction with (30.5),

$$G(1) = 1 \pm (1 + 2\lambda)^{-1/2}.$$

Iterating the same boundary differentiation at the dyadic preimages of 1 shows that in this case the boundary limit of G' fails to exist at *all* binary rational angles.

Since the set of binary rational angles is dense in $\partial\mathbb{D}$, in either scenario G admits no analytic continuation across any boundary point. Thus $\partial\mathbb{D}$ is a natural boundary for G . See [9] for details.

■

Exercise 189. Fill in the missing steps in the argument above: justify the boundary differentiation of (30.4), the propagation to dyadic angles via $z \mapsto z^2$, and the dichotomy for G' at binary rational angles. Conclude that $\partial\mathbb{D}$ is indeed a natural boundary of G .

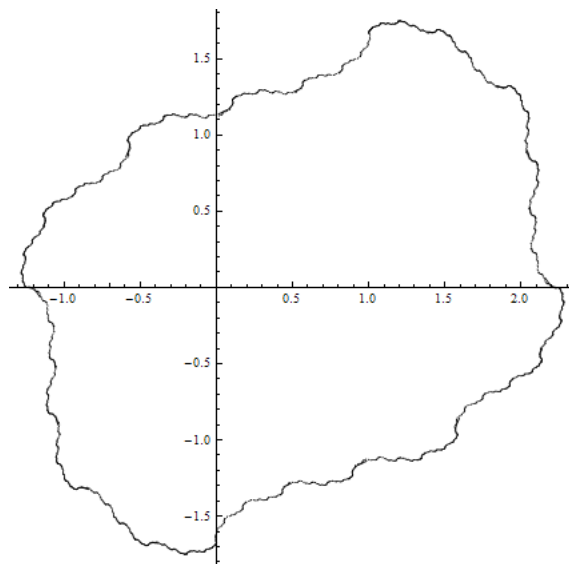


Figure 33: The conformal image of the unit disk under the map G .

31 Conformal Mappings of Polygons and the Schwarz–Christoffel Formula

For polygonal regions, a conformal map from the unit disk—or equivalently from the upper half-plane $\mathbb{H}^u$ —can be expressed in terms of quadratures. Although the resulting transforma-

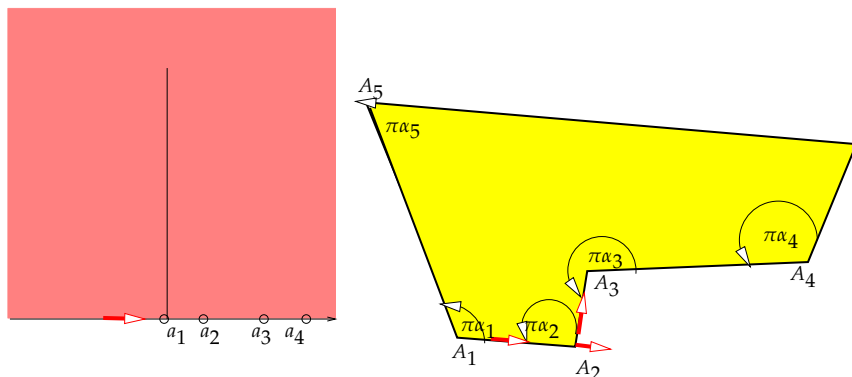


Figure 34: Schwarz–Christoffel transformation: a closed polygonal domain.

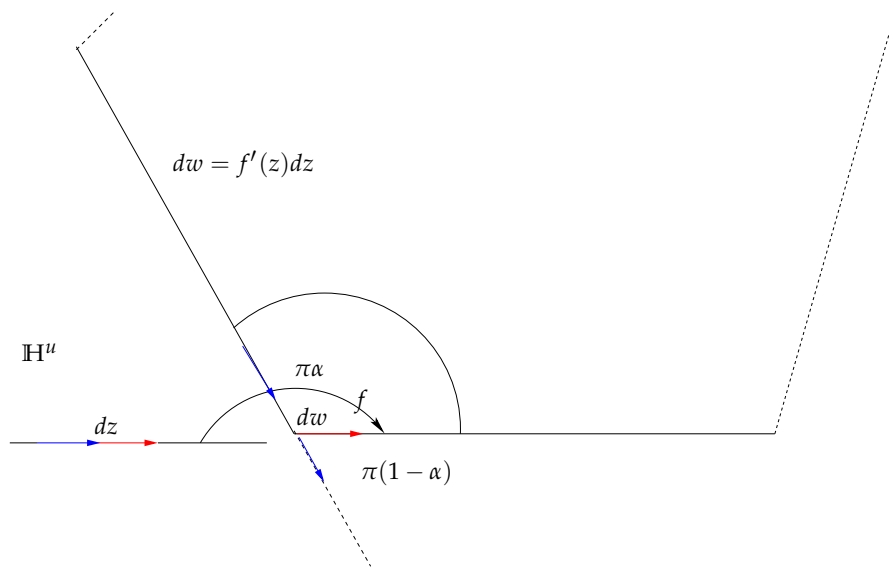


Figure 35: Schwarz–Christoffel transformation near one vertex: two adjacent sides of the polygon.

tion is generally nonelementary, its integral representation provides sufficient analytic control to describe its behavior with precision.

31.1 Heuristics

If ψ is biholomorphic at z_0 , then the angle between the tangent to a curve γ passing through z_0 and the tangent to its image curve under ψ at $\psi(z_0)$ is given by $\arg \psi'(z_0)$. In differential form, this is expressed as

$$dw = \psi'(z) dz.$$

We seek a conformal mapping that sends $\mathbb{H}^u$ onto the interior of a polygon. Choosing the positive orientation of $\partial\mathbb{H}^u$ —that is, traversing the boundary so that the mapped domain lies to the left—corresponds to moving along the real axis from $\mathbb{R}^-$ to $\mathbb{R}^+$. For convenience, denote by $\pi\alpha_i$ the interior angles of the polygon.

Without loss of generality, place one vertex of the polygon at the origin and rotate the polygon so that one of its sides lies along $\mathbb{R}^+$. The red arrow in Figure 35 indicates the positive orientation of the polygon. Suppose that we wish to map $0 \mapsto 0$ and the blue segment of $\partial\mathbb{H}^u$ onto the corresponding blue edge of the polygon, as shown in Figure 35.

Examining the local behavior near $z = 0$, we observe that to the left of $z = 0$ the differential dw is rotated by $-\pi(1 - \alpha)$ relative to dz , while to the right of $z = 0$ (in the direction of the red arrow) the directions of dz and dw coincide. A function whose derivative possesses exactly this property along the boundary is

$$\psi(z) = z^\alpha,$$

defined using the principal branch of the logarithm, positive for $z > 0$.

There are two ways to see this. First, consider the mapping properties of $z^{-\alpha}$. The argument of $z^{-\alpha}$ remains constant along the real axis, except for a jump at the discontinuity of $\arg z$ at $z = 0$. Take some $\varepsilon > 0$. Moving from $-\varepsilon$ to ε corresponds, in the z -plane, to a rotation by $-\pi$, which under the transformation $z^{-\alpha}$ becomes a rotation by $\alpha\pi$, ending at $\varepsilon^{-\alpha} > 0$. This precisely corresponds to traversing the vertex 0 of the polygon.

Alternatively, we can write

$$\log \psi' = (-\alpha - 1) \log z + C,$$

so that

$$(\log \psi')' = \frac{\alpha - 1}{z} \in \mathbb{R} \quad (\text{since } z \in \mathbb{R}), \quad \Rightarrow \quad d(\arg \psi') = 0 \text{ for } z \neq 0. \quad (31.1)$$

Hence $(\operatorname{Im} \psi')' = 0$.¹⁰

Proposition 31.1. *Let ψ map conformally a one-sided neighborhood $\mathcal{N} \subset \mathbb{H}^u$ of a segment $I = [-a, b]$, with $a, b > 0$, to a one-sided neighborhood on the interior side of a polygon, such that $\psi(0) = 0$ and ψ extends continuously to the boundary. Then*

$$\psi'(z) = z^{\alpha-1} h(z),$$

where h is holomorphic in $\mathcal{N} \cup I \cup \overline{\mathcal{N}}$. In particular,

$$\psi'(z) = z^{\alpha-1} \sum_{k=0}^{\infty} h_k z^k, \quad h_0 = h(0) \neq 0,$$

and the series converges in a neighborhood of 0.

Proof. By the reflection theorem across analytic arcs (Theorem 180), ψ admits analytic continuation through both $[-a, 0)$ and $(0, b]$. Define

$$H(z) = \int_i^z s^{-(\alpha-1)} \psi'(s) ds.$$

Then

$$dH = z^{-(\alpha-1)} \psi'(z) dz.$$

The prefactor $z^{-(\alpha-1)}$ exactly compensates for the rotation produced by ψ , so that H maps a one-sided neighborhood of 0 in $\mathbb{H}^u$ to a one-sided neighborhood of 0 in $\mathbb{H}^u$. Hence, by the Schwarz reflection principle, there exists $\varepsilon > 0$ such that H has analytic continuation through the interval $[-\varepsilon, \varepsilon]$. ■

¹⁰ A subtle point: the two harmonic functions given by the real and imaginary parts of an analytic function are generally not complex-differentiable in z . However, along the real axis we are differentiating with respect to x alone, so the derivative exists in the ordinary real sense.

Definition 190

A function that is (locally) of the form

$$(z - z_0)^a H(z),$$

where H is holomorphic, will be called **ramified-analytic**.

More generally, a transformation of the form

$$\psi'(z) = \prod_{i=1}^{n-1} (z - a_i)^{\alpha_i - 1}, \quad a_i \in \mathbb{R}, \quad \alpha_i \in (0, 2], \quad (31.2)$$

satisfies, for $z \in \mathbb{R}$,¹⁰

$$(\ln \psi')' = \sum_{i=1}^{n-1} \frac{\alpha_i - 1}{z - a_i} \in \mathbb{R}, \quad \Rightarrow \quad d(\arg \psi') = 0 \quad \text{for } z \neq a_i. \quad (31.3)$$

Hence $\arg \psi'$ remains constant along each interval between consecutive prevertices a_i and changes only when z crosses one of them. Specifically, when z passes through a_i , $\arg \psi'$ changes by $\pi(1 - \alpha_i)$. The corresponding image points

$$A_i := \psi(a_i)$$

on the polygon are therefore the vertices at which $\arg \psi'$ changes, and the change in argument is $+\pi(1 - \alpha_i)$ (the sign convention must be verified carefully).

If the polygon P is closed, then the point at infinity corresponds to one of its vertices. Indeed, infinity must map to a point on the boundary of P , and this point cannot lie strictly between two vertices, since by reflection across analytic arcs, ψ^{-1} admits analytic continuation through the sides of the polygon. Because ψ is conformal and hence injective on the boundary, this vertex cannot coincide with any of $\{A_j : j = 1, \dots, n-1\}$. We denote it by A_n .

Recalling that the sum of the exterior angles of a closed polygon is 2π , we have

$$B = \sum_{i=1}^{n-1} \beta_i = 1 + \alpha_n, \quad \text{where } \beta_i := 1 - \alpha_i.$$

Each factor $(1 - a_i/z)$ is analytic at infinity. Thus, writing

$$\psi'(z) = z^B \prod_{i=1}^{n-1} (1 - a_i/z)^{\beta_i} = \zeta^{-B} g(\zeta), \quad g(\zeta) = \prod_{i=1}^{n-1} (1 - a_i \zeta)^{\beta_i}, \quad \zeta = \frac{1}{z}, \quad (31.4)$$

we see that g is analytic at 0 and satisfies $g(\zeta) = 1 + \zeta h(\zeta)$, with h analytic near 0. Hence $\ln g$ is well defined for sufficiently small ζ .

In the limit $R \rightarrow \infty$, when z traverses the semicircular arc $\partial \mathbb{D}_R$ from $+R$ to $-R$ in the positive (counterclockwise) direction, the total change in $\arg \psi'$ arises solely from the factor z^B and equals πB . This is precisely the change in argument produced by a factor $(z - A)^{1 - \alpha_n}$. Consequently, the mapping “closes the polygon” with the final vertex

$$A_n = \lim_{z \rightarrow \infty} \psi(z).$$

In the case of a closed polygon, ψ is bounded on $\mathbb{R}$; for an open polygonal chain, the argument is entirely analogous.

Theorem 191: Schwarz–Christoffel

Let P be a polygonal domain (closed polygon or open polygonal chain) with ordered vertices $A_1, \dots, A_n$ and interior angles $\pi\alpha_1, \dots, \pi\alpha_n$, where $0 < \alpha_k < 2$. Then:

(i) Upper half-plane. There exist distinct prevertices $a_1, \dots, a_n \in \mathbb{R} \cup \{\infty\}$ (in increasing order along $\mathbb{R}$), a nonzero constant $C \in \mathbb{C}$, and a constant $C' \in \mathbb{C}$ such that

$$\psi(z) = C \int_0^z \prod_{k=1}^n (s - a_k)^{\alpha_k - 1} ds + C', \quad (31.5)$$

where, if $a_k = \infty$, the corresponding factor $(s - a_k)^{\alpha_k - 1}$ is omitted. With appropriate branch choices (holomorphic in $\mathbb{H}^u$ and continuous up to $\mathbb{R}$ away from the a_k), ψ maps $\mathbb{H}^u$ onto P and satisfies $\psi(a_k) = A_k$.

(ii) Unit disk. Let $b_1, \dots, b_n \in \partial\mathbb{D}$ be distinct boundary prevertices in counterclockwise order. Then there exist $C \neq 0$ and $C' \in \mathbb{C}$ such that

$$\psi(z) = C \int_0^z \prod_{k=1}^n (s - b_k)^{\alpha_k - 1} ds + C'. \quad (31.6)$$

(Equivalently, one may write $\psi'(z) = \tilde{C} \prod_{k=1}^n (1 - z/b_k)^{\alpha_k - 1}$; the two forms differ by an overall constant factor absorbed into C .)

(iii) Conversely, every conformal map from $\mathbb{H}^u$ (resp. $\mathbb{D}$) onto a polygonal domain is of the form (31.5) (resp. (31.6)) for a suitable choice of prevertices and constants, unique up to postcomposition by an affine map $w \mapsto \lambda w + \mu$ ($\lambda \neq 0$) and precomposition by an automorphism of $\mathbb{H}^u$ or $\mathbb{D}$.

(iv) A prevertex at infinity. One of the points a_k in (i) may be chosen to be ∞ . In that case,

$$\psi(z) = C \int_0^z \prod_{k=1}^{n-1} (s - a'_k)^{\alpha_k - 1} ds + C', \quad (31.7)$$

where $a'_1, \dots, a'_{n-1} \in \mathbb{R}$ are the remaining finite prevertices.

Note 192

For an *open* polygonal chain, one of the α_k encodes the “vertex at infinity”; this is handled by taking one prevertex $a_k = \infty$ and using (31.7).

Remark 193. (i) It is important to note the freedom we have in such transformations. Suppose we wish to map the upper half-plane $\mathbb{H}^u$ onto a triangle Δ . All triangles with the same angles are similar, and a mapping between two similar triangles reduces to scaling, rotation, and translation. Thus it suffices to understand one triangle with given interior angles α_i . By a rescaling and shift of $\mathbb{H}^u$ we arrange $a_1 = 0$ and $a_2 = 1$. This results in a rescaling and shift of the original triangle, leading to a similar one Δ_1 . We can choose C and C' so as to map onto Δ .^{??} Hence we can map any triangle onto $\mathbb{H}^u$, prescribing at will the positions of the images of its prevertices. See §31.2.

(ii) For $n > 3$, we can still place three points at will, but the positions of the remaining prevertices cannot be chosen arbitrarily. We have $n - 1$ *real* constants $(a_1, \dots, a_{n-1})$ and two

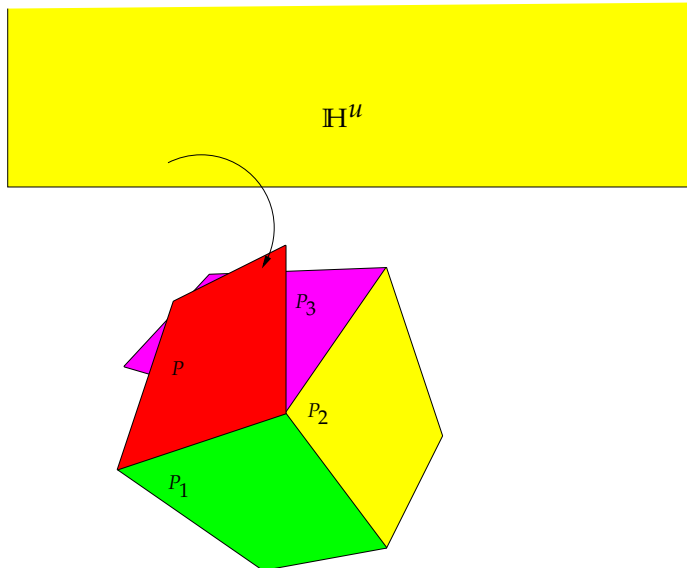


Figure 36: Successive reflections.

complex ones (C and C'), a total of $n + 3$ real degrees of freedom, whereas an arbitrary closed n -gon has $2n$ real degrees of freedom (the positions of its vertices). For $n > 3$ one has $2n > n + 3$, so the analytic expression provides fewer free parameters than the polygon has degrees of freedom. The remaining quantities a_k with $k \geq 4$, whose values must be chosen so that the image domain has the prescribed shape, are traditionally called *accessory parameters*.

They are determined by the shape of the polygon together with the values of a_1, a_2, a_3 . Except in very symmetric cases (such as regular polygons), these parameters cannot be expressed in closed form.

Given this disparity between the number of available parameters and the degrees of freedom of the problem, it is not a priori clear that the Schwarz–Christoffel transformation should always exist—but it does.

Proof. We analyze the case of closed polygons; the argument for open polygons is entirely analogous. Once the formula is established for the upper half-plane, the corresponding one for $\mathbb{D}$ follows immediately via the Cayley transform. This proof follows closely [15]; a slightly different argument appears in [3]. We use the notations indicated in Figure 36.

Let ψ be the conformal map from $\mathbb{H}^u$ onto P . We assume, without loss of generality, that no α_i is zero, and denote by $a_1, \dots, a_{n-1} \in \mathbb{R}$ the prevertices that map to the vertices $A_1, \dots, A_{n-1}$ of P . The figure shows successive reflections of the polygon P across its sides. Since ψ maps an interval I of $\mathbb{R}^+$ to the common side $\ell = \partial P \cap \partial P_1$, it admits analytic continuation across I . The image of the upper half-plane is P , while the image of the lower half-plane is P_1 . This continuation may be reflected again, producing a new polygon P_2 . The polygon P_2 is obtained from P by two reflections and is therefore congruent to P ; hence $P_2 = aP + c$ for some constants $a, c \in \mathbb{C}$ with $|a| = 1$. The domains P and P_1 are related by a reflection (a flip).

By Theorem 180, the function ψ is analytic at every point of $\mathbb{R}$ except at the prevertices $a_1, \dots, a_n$, where it is singular. Indeed, ψ fails to be conformal at a_i , because it maps two consecutive intervals of $\mathbb{R}$ into two straight segments meeting at an interior angle $\pi\alpha_i$. Consider now the auxiliary function

$$h(z) = [\psi(z) - \psi(a_i)]^{1/\alpha_i}.$$

Then h maps an interval to a straight line, and by reflection across analytic arcs, it is analytic at a_i . Consequently,

$$\psi(z) = \psi(a_i) + [h(z)]^{\alpha_i}. \quad (31.8)$$

Since $h(a_i) = 0$, each a_i is a branch point of ψ . We have

$$h(z) = (z - a_i)H(z) \text{ where } H(a_i) \neq 0, \text{ hence } h'(a_i) \neq 0 \quad (31.9)$$

(for the polygon interior angle to be α_i).

Consider now

$$g(z) = \frac{\psi''(z)}{\psi'(z)} = \frac{(\alpha_i - 1)h'(z)}{h(z)} + \frac{h''(z)}{h'(z)} = \frac{\alpha_i - 1}{z - a_i} + \frac{(\alpha_i - 1)H'(z)}{H(z)} + \frac{h''(z)}{h'(z)}. \quad (31.10)$$

From (31.8) and 31.9 we see that g has a simple pole at a_i with residue $\alpha_i - 1$. Since the a_i are the only singularities of g , it is single-valued and meromorphic in $\mathbb{C}$. Define

$$G(z) = g(z) + \sum_{k=1}^{n-1} \frac{1 - \alpha_k}{z - a_k}. \quad (31.11)$$

The sum on the right cancels the poles of g , so that G is entire.

From the behavior at the last vertex, which corresponds to $z = \infty$, we obtain as in (31.11), $\psi''/\psi' = O(1/z)$ as $z \rightarrow \infty$, and the same holds for the sum in (31.11); thus $G(z) \rightarrow 0$ as $z \rightarrow \infty$. By Liouville's theorem, G must vanish identically, hence

$$\frac{\psi''(z)}{\psi'(z)} = \sum_{k=1}^{n-1} \frac{\alpha_k - 1}{z - a_k}. \quad (31.12)$$

Integrating (31.12) gives

$$\psi(z) = c_1 + c_2 \int_0^z \prod_{k=1}^{n-1} (t - a_k)^{\alpha_k - 1} dt,$$

which establishes part (iv) of Theorem 191. As discussed earlier, composition with the Cayley transform yields (ii). To make one of the a_k equal to ∞ , we use a linear fractional substitution $z = a_k - 1/z_1$; the rest of the proof of the theorem is similar. ■

Note 194

Multiplying (31.12) by $\psi'(z)$ yields a second-order linear differential equation for ψ . This equation is a *Fuchsian equation*: all its singularities on $\hat{\mathbb{C}}$ (the extended complex plane) are *regular singular points* [4].

31.2 Conformal map of the UHP onto the interior of a triangle of angles $\pi\alpha, \pi\beta, \pi\gamma$

We choose for convenience $a_1 = 0, a_2 = 1, a_3 = \infty$ resp. Theorem 191 (iv) gives

$$\psi(z) = A \int_0^z z^{\alpha-1} (1-z)^{\beta-1} dz + B \quad (31.13)$$

We choose $A = 1$. From the choice above, we know that the length of the first side is

$$\int_0^1 |\psi'(s)| ds = \int_0^1 s^{\alpha-1} (1-s)^{\beta-1} ds = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)}$$

The lengths of the other sides can be most easily determined from the relation

$$\frac{a}{\sin \pi \alpha} = \frac{b}{\sin \pi \beta} = \frac{c}{\sin \pi \gamma}$$

We note that $\psi(z) = B_z(\alpha, \beta)$, the incomplete Beta function, which, for $\alpha, \beta \notin -\mathbb{N}$ satisfies

$$B_z(\alpha, \beta) = \Gamma(\alpha) z^\alpha {}_2F_1(\alpha, 1 - \beta; \alpha + 1; z)$$

where ${}_2F_1$ is the hypergeometric function. This is seen from the differential equation giving ψ . **Differential equation.** We can derive a differential equation with polynomial coefficients which gives ψ as follows. Starting from

$$\psi'(z) = z^{\alpha-1}(1-z)^{\beta-1},$$

we write $\psi = z^\alpha \psi_1$ and, after dividing by $z^{\alpha-1}$, obtain

$$z\psi_1' + \alpha\psi_1 - (1-z)^{\beta-1} = 0 \Rightarrow \frac{z}{(1-z)^{\beta-1}}\psi_1' + \frac{\alpha}{(1-z)^{\beta-1}}\psi_1 - 1 = 0. \quad (31.14)$$

Differentiating once more to eliminate the constant 1 gives

$$z(1-z)\psi_1'' + [\alpha + 1 - (2 - \beta - \alpha)z]\psi_1' + \alpha(\beta - 1)\psi_1 = 0. \quad (31.15)$$

The differential equation satisfied by the hypergeometric function ${}_2F_1(a, b; c; z)$ is the *Riemann equation*

$$z(1-z)h'' + [c - (a + b + 1)z]h' - abh = 0. \quad (31.16)$$

From the way we passed from (31.14) to (31.15) (or by direct verification), one solution of (31.15) is $z^{-\alpha}$. Comparing with (31.16), the second independent solution is

$${}_2F_1(\alpha, 1 - \beta; \alpha + 1; z) \Rightarrow \psi(z) = z^\alpha {}_2F_1(\alpha, 1 - \beta; \alpha + 1; z). \quad (31.17)$$

Thus, in this case, the Schwarz–Christoffel transformation is expressed as the ratio of two independent solutions of the special hypergeometric equation (31.15).

31.3 Schwarz triangle functions and hypergeometric functions

We can attempt the reflection–continuation procedure of §???. Consider successive reflections of the triangle about its sides that share a common vertex. For the resulting function to be single valued, the sequence of reflections must return to the original triangle without overlap or gap. This happens precisely when the rotations about that vertex have finite order, that is, when

$$\alpha_i = \frac{1}{n_i}, \quad n_i \in \mathbb{N}.$$

The constraints are therefore

$$\alpha + \beta + \gamma = 1, \quad 1/\alpha, 1/\beta, 1/\gamma \in \mathbb{N}. \quad (31.18)$$

Check that the only solutions are

$$\begin{aligned} \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right) & \text{ equilateral triangle,} \\ \left(\frac{1}{2}, \frac{1}{3}, \frac{1}{6}\right) & \text{ half of an equilateral triangle,} \\ \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{4}\right) & \text{ isosceles right triangle.} \end{aligned}$$

In each case, the reflected images tile the entire plane, and the corresponding mapping functions are restrictions of meromorphic functions. These special cases are known as the *Schwarz triangle functions*. The associated group of reflections is a special case of an infinite *Coxeter group*.

Each triangle function corresponds to an elliptic function; we will return to this topic later.

32 Optional material: curvilinear triangles

In general, the map from $\mathbb{H}^u$ into a *curvilinear triangle*, one whose sides are arccircles is given by the ratio of two independent solutions of (31.16), where the angles α, β, γ are related to a, b, c by (cf. [15])

$$a = \frac{1}{2}(1 + \beta - \alpha - \gamma), \quad b = \frac{1}{2}(1 - \alpha - \beta - \gamma), \quad c = 1 - \alpha \quad (32.19)$$

Note 195

To see qualitatively why that is, before we work out the mathematical details, since the sought-for function f transforms segments into arccircles, Möbius transforms of f map real segments into real segments. Möbius transforms are conformal wherever defined, so they preserve angles, and thus f should be ramified analytic at three points, say $\{0, 1, \infty\}$.

Secondly, note that any function which is real-analytic (with real values) on some interval I and such that $f' \neq 0$ on I conformally maps a neighborhood of I into a neighborhood of $f(I)$ (where, of course, $f(I)$ is an interval in $\mathbb{R}$).

Finally, this approach would apply to any curvilinear polygon, and the final ODE that we construct will still be second order linear, but the solutions are generally quite complicated.

To find an f as in Note 195, a natural candidate would be a ratio of two solutions of a real-valued linear second order ODE with analytic coefficients, whose solutions have only ramified singularities, at in $\{0, 1, \infty\}$. Such an ODE will have two linearly independent solutions, f_1 and f_2 ¹¹. Define $S = \{a_1 f_1 + a_2 f_2 \mid A_{1,2} \in \mathbb{C}\}$ (this of course is the space of all solutions). Then,

1. $f_{1,2}$ are analytic in $\mathbb{C}^* \setminus \{0, 1, \infty\}$ while at the points $0, 1, \infty$ they are *ramified-analytic*;
2. on each interval $(-\infty, 0)$, $(0, 1)$ and $(1, \infty)$ there are two positive linearly independent functions in S ;
3. S is invariant under Schwarz reflections¹². This is because any analytic continuation of a solution is a solution, by permanence of relations.
4. $(f_1/f_2)' \neq 0 \forall x \in \mathbb{R}$.¹³

Then some ratio F_1/F_2 of two linearly independent $F_{1,2} \in S$ maps the upper half plane into a curvilinear triangle, that is, one whose sides are line segments or arccircles (a line segment is a limiting case of an arccircle, so we will call arccircle too). Indeed, take F_1 and F_2 in S linearly independent, choose one of the intervals in (2) above, and let $\tilde{F}_1$ and $\tilde{F}_2$ the functions which are real valued on the interval I . We choose $\tilde{F}_1$ and $\tilde{F}_2$ such that $\tilde{F}_1/\tilde{F}_2$ is bounded at zero. By assumption, $F_{1,2} = A_{1,2}\tilde{F}_1 + B_{1,2}\tilde{F}_2$ and

$$\frac{F_1}{F_2} = \frac{1 + a\tilde{F}_1/\tilde{F}_2}{b + c\tilde{F}_1/\tilde{F}_2} \quad (32.20)$$

Now, $\tilde{F}_1/\tilde{F}_2$ is real-valued and one-to-one, and thus the right side of (32.20) is a Möbius transformation of a line segment: an arccircle. The image of $\mathbb{R}$ through F_1/F_2 consists of three arccircles,

¹¹Meaning that $A, B \in \mathbb{C}$ and $Af_1 + Bf_2 = 0 \Rightarrow A = B = 0$.

¹²Equivalently, the space generated by $f_{1,2}$, $S := \{C_1 f_1 + C_2 f_2 : C_1, C_2 \in \mathbb{C}\}$ is closed under continuations at the branch points (e.g., $f_1(ze^{2\pi i}) = C_1 f_1(z) + C_2 f_2(z)$).

¹³This in fact is implied by linear independence. Check this by first showing that two functions are linearly independent iff their Wronskian $W(f_1, f_2) := f_1' f_2 - f_2' f_1 \neq 0$.

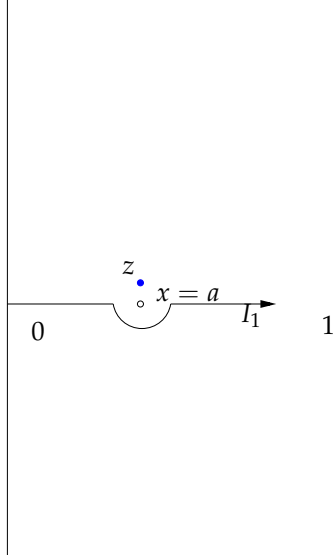


Figure 37: Hypergeometric contours

a general curvilinear triangle, provided that the singularities are such that F_1/F_2 are continuous on $\mathbb{R}$.

We now show that (31.16) is such an equation, for a, b, c satisfying (32.19). The following integral representation due to Euler can be checked to solve (31.16):

$${}_2F_1(a, b; c; z) = K \int_0^1 H(z, x) dx, \quad \text{where } H(z, x) = x^{b-1}(1-x)^{c-b-1}(1-zx)^{-a} \text{ assuming } c > b > 0 \quad (32.21)$$

The standard choice of K , immaterial here, is $K^{-1} = B(b, c-b)$ with B the Beta function¹⁴. For the integral to exist, we need

$$b-1 > -1, c-b-1 > -1, -a > -1 \quad (32.22)$$

which are satisfied if $\alpha + \beta + \gamma < 1$, $\alpha, \beta, \gamma \in \mathbb{R}^+$, which is the case for a hyperbolic triangle (with concave sides) as can be verified from (32.19). Under this same condition, we have

$$(b-1) + c-b-1-a < -1 \Rightarrow H(x, z) \sim s^{-p}, \quad p > 1 \text{ as } x \rightarrow \infty \quad (32.23)$$

Singularities of ${}_2F_1$ For general a, b, c the behavior of h in (31.16) at the singular points of the ODE, $0, 1, \infty$ ¹⁵ follows general results about regular singular points of ODEs (Frobenius theory, cf. [4]); we will not assume this here however. We will find the behavior of the solutions in three different way, to illustrate various approaches.

Directly from the integral representation, the behavior can be calculated in the following way.

Clearly, (32.21) and Corollary ?? imply that ${}_2F_1$ is analytic except possibly on $\mathbb{R}^+$. If $z \rightarrow a \in \mathbb{R}^+ \setminus \{1\}$ then we can use the analyticity of the integrand at a to first homotopically deform the contour as shown in Fig. 37. The new integral (of course, equal to the original one) is manifestly

¹⁴The formula is valid under the more general condition $\text{Re}(c) > \text{Re}(b) > 0$, but here we only need real a, b, c .

¹⁵Writing the equation in the form $h'' + Q(x)h + R(x) = 0$, $\{0, 1\}$ are singular points of P, Q . Ditto after the change of variable $z = 1/t$, $y(z) = t^{-c}Y(t)$ at $t = 0$.

analytic in z near a . Thus the only possible singularity is at 1 (since the integrand is manifestly analytic for small z , cf. Corollary ??).

At $z = 0$ we get

$$\frac{{}_2F_1(a, b; c; z)}{K} = \int_0^1 x^{b-1} (1-x)^{c-b-1} dx > 0 \quad (32.24)$$

To find the type of behavior at $z = 1$ it is convenient to take $z = 1 - \varepsilon$ and change the variable to $x = 1 - s$ in the integral:

$$K \int_0^1 h(z, (1-s)) ds = K \int_0^1 (1-s)^{b-1} s^{c-b-1} (\varepsilon + s - \varepsilon s)^{-a} ds \quad (32.25)$$

By (32.23) We can then push the contour up toward $+i\infty$,

$$\int_0^1 h(z, (1-s)) ds = \left(\int_0^{+i\infty} - \int_1^{+i\infty} \right) (1-s)^{b-1} s^{c-b-1} (\varepsilon + s - \varepsilon s)^{-a} ds \quad (32.26)$$

The second integral is analytic in ε by the same Corollary ??.

In the first one, we change variable to $s = \varepsilon u$, to get

$$\begin{aligned} -\varepsilon^{c-b-1-a+1} \int_0^{+i\infty} (1-\varepsilon u)^{b-1} u^{c-b-1} (1+u-\varepsilon u)^{-a} ds \\ = -\varepsilon^{c-b-a} \int_0^{+\infty} (1-\varepsilon u)^{b-1} u^{c-b-1} (1+u-\varepsilon u)^{-a} ds \end{aligned} \quad (32.27)$$

where the contour change is justified by (32.23).

Exercise 196. What is the phase of the last integral in (32.27)?

The integral in (32.27) is analytic in ε and thus

$${}_2F_1(a, b; c; z) = A_1 + A_2(1-z)^\gamma \text{ where } A_{1,2} \text{ are analytic at 1} \quad (32.28)$$

and $A_2(1) = \int_1^{+\infty} u^{c-b-1} (1+u)^{-a} ds > 0$. A second solution of (31.16) is obtained, cf. [15], by noticing that the substitution $h(z) = g(1-z)$ leads to the equation (31.16) with c replaced by $C = a + b + 1 - c$. The integral representation (32.21) still holds with c replaced by $C = 1 - \gamma$, and converges under the same conditions on α, β, γ as (32.21). The behavior of the integrand at infinity is still of the form in (32.23). Thus the same analysis applies, to show that there is a second solution which is analytic at $z = 1$ and has a singularity at zero, with $c - b - a$ replaced by $1 - c = \alpha$. This also suggests making the substitution $h_1 = z^{1-c} h = z^\alpha h$ directly into (31.16); we get an equation of the form (31.16), with A, B, C replaced by

$$A = a + 1 - c, B = b + 1 - c, C = c - 2 \quad (32.29)$$

which has an integral representation of the form (32.21) valid under the same conditions on α, β, γ . Thus, near $z = 0$ we have two linearly independent solutions, A_1 analytic and A_2 of the form $z^\alpha A(z)$ with A analytic and $A(0) \neq 0$. These are clearly linearly independent since the second solution is not analytic at zero. In the same way, there are two solutions, one analytic and positive at $z = 1$ and another one of the form $z^\gamma A(z)$, A analytic and positive for $z < 1$.

From the ODE. Here we are assuming knowledge of basic properties of linear ODEs: a linear combination (with constant coefficients) of solutions is a solution (this can be checked directly) and the fact that the space of solutions of a second order linear equation is two-dimensional

(i.e., there are exactly two free constants, or, in other words, there are two linearly independent solutions which form a basis in the space of all solutions).

With F_{01} given by (32.21), we notice as before that F is analytic near zero and $F(0) > 0$. We look for a second solution in the form $F_{02} = F_{01}g$. The equation for g is

$$\begin{aligned} \frac{g''}{g'} &= -2\frac{F'}{F} + q(x); \quad \left(q(x) := \frac{(a+b+c)x-c}{x(1-x)} \right) = -2\frac{F'}{F} - \frac{c}{x} + A(x) \\ &\Rightarrow g' = -\frac{c}{x}A_1(x) \Rightarrow F_{01}(x) = x^{-c+1}A_2(x) = x^\alpha A_2(x); \quad (32.30) \end{aligned}$$

with A, A_1, A_2 analytic (check the conclusions above). Since F is real-valued for real $z \in (0, 1)$, g is also real-valued, and it is an independent solution (it has a manifestly different behavior at zero). At $x = 1$ we make the substitution $z = 1 - y$ and we get

$$y(1-y)h'' + [C - (a+b+1)y]h' - abh = 0, \quad C = a+b+c-1 = 1-\gamma \quad (32.31)$$

In the same way as above we get two independent real valued solutions for y real, $F_{10}(y)$ analytic and a second one F_{11} of the form $y^\gamma A_1$ with A_1 analytic. They are in general different from the solutions $F_{00}(y)$ and $F_{00}(y)$.

Finally the substitution $h(z) = z^\alpha H(1/z)$ $z = 1/Z$ in (31.16) results in an equation of the same type as (31.16). The ratio of two solutions behaves like $z^{-\beta}$ for large z . Now, any transformation of the type F_1/F_2 where F_1 and F_2 are real maps the interval $(0, 1)$ into an interval. A different choice f_1/f_2 is a Möbius transformation of F_1/F_2 , and thus it maps the interval $(0, 1)$ into an arccircle. Keeping one combination F_1/F_2 chosen so that $F_1/F_2 \rightarrow 0$ as $z \rightarrow \infty$ will then map $\mathbb{R}$ into a curvilinear triangle, of angles α, β, γ .

Note 197

In the simplest nondegenerate case, Frobenius theory allows to determine the behavior of solutions of linear meromorphic ODEs, Lf , with regular singularities, say at z_0 by a very simple method: take $L(z - z_0)^r$ and keep only the lowest power of $(z - z_0)$. This gives a quadratic equation for r . If the solutions $r_{1,2}$ do not differ by an integer, then $Lf = 0$ has two linearly independent solutions in the form $(z - z_0)_{1,2}^r A_{1,2}(z)$, A analytic at z_0 .

Note 198

The equation satisfied by f_1/f_2 , a ratio of solutions of the hypergeometric equation is

$$\{w, z\} = \frac{1-\alpha^2}{2z^2} + \frac{1-\gamma^2}{2(z-1)^2} + \frac{\alpha^2 + \gamma^2 - \beta^2 - 1}{2z(z-1)} = 0 \quad (32.32)$$

where $\{w, z\}$ is the important *Schwarzian derivative* [15]:

$$\{w, z\} = \left(\frac{w''}{w'} \right)' - \frac{1}{2} \left(\frac{w''}{w'} \right)^2 \quad \left(w' = \frac{dw}{dz} \right) \quad (32.33)$$

The Schwarzian derivative is *invariant under any Möbius transformations* of w as expected from our discussion and can be checked by straightforward calculation.

33 Asymptotics of Integrals and of Taylor Coefficients of Analytic Functions

Many special functions—solutions of ordinary or partial differential equations—admit integral representations in which one or more variables appear as parameters. For instance, the Airy function can be expressed as

$$\text{Ai}(x) = \frac{1}{\pi} \int_0^\infty \cos\left(\frac{1}{3}t^3 + xt\right) dt,$$

or, equivalently, as a contour integral in the complex plane:

$$\text{Ai}(z) = \frac{1}{2\pi i} \int_{\infty e^{-\pi i/3}}^{\infty e^{\pi i/3}} \exp\left(\frac{1}{3}t^3 - zt\right) dt.$$

Another example is the sine integral,

$$\text{Si}(z) = \int_0^z \frac{\sin t}{t} dt = \int_0^1 \frac{\sin(zt)}{t} dt.$$

Likewise, the Maclaurin coefficients of an analytic function f are given by Cauchy's integral formula:

$$c_k = \frac{1}{2\pi i} \oint f(s) s^{-k-1} ds. \quad (33.34)$$

Asymptotic analysis of such integrals plays a central role throughout mathematics and physics, providing a unified approach to estimating the behavior of special functions and the growth of their Taylor coefficients.

As an illustration, consider the case of orthogonal polynomials defined through generating functions.

Proposition 33.2 ([12]). The Legendre polynomials admit the generating function

$$\frac{1}{\sqrt{1 - 2x\zeta + \zeta^2}} = \sum_{n=0}^{\infty} P_n(x) \zeta^n, \quad (33.35)$$

where the branch of the square root is chosen so that the expression is positive for small positive x and ζ . The polynomials $\{P_n\}_{n \geq 0}$ form a complete orthogonal system in $L^2([-1, 1])$, satisfying

$$\int_{-1}^1 P_m(x) P_n(x) dx = \begin{cases} 0, & n \neq m, \\ \frac{2}{2n+1}, & n = m. \end{cases} \quad (33.36)$$

Applying Cauchy's integral formula to (33.35), we obtain, for $x \in (-1, 1)$,

$$P_n(x) = \frac{1}{2\pi i} \oint_C \frac{ds}{s^{n+1} \sqrt{s^2 - 2xs + 1}}, \quad (33.37)$$

where C is a positively oriented, simple closed contour lying inside the unit circle.

The quadratic factor in the denominator can be written as

$$(s_+ - s)(s_- - s), \quad s_{\pm} = x \pm \sqrt{x^2 - 1}.$$

We have $|s_{\pm}| = 1$ and $s_+ = \overline{s_-}$. Writing $s_+ = e^{i\zeta}$ and choosing the branch so that $\zeta \in (0, \pi)$, we deform the contour C into a pair of Hankel contours surrounding the branch cuts emanating from $s_{\pm}$ in the directions $\arg s_{\pm}$.

It suffices to consider the contour issuing from s_+ . By analyticity and dominated convergence (and since the integral around a circle of radius $\varepsilon > 0$ centered at s_+ vanishes as $\varepsilon \rightarrow 0$, because the integrand is locally in L^1 near s_+), the Hankel contour can be collapsed onto the two sides of the branch cut. The right-hand side of the cut is traversed from infinity to s_+ , and the left-hand side from s_+ to infinity. These two contributions are equal in magnitude, since they are traversed in opposite directions but the square root changes sign across the cut.

Making the substitution $s = s_{\pm}t$ and proceeding similarly for both branches yields¹⁶

$$P_n(x) = \frac{1}{\pi} \sum_{\pm} s_{\pm}^{-(n+1)} \int_1^{\infty} \frac{t^{-n-1} dt}{\sqrt{(t-1)(t-s_{\mp}/s_{\pm})}}. \quad (33.38)$$

With the substitution $x = \cos \zeta$ (so that $s_{\pm} = e^{\pm i\zeta}$ and $s_{\mp}/s_{\pm} = e^{\mp 2i\zeta}$), we obtain

$$P_n(\cos \zeta) = \frac{2}{\pi} \operatorname{Re} \left[e^{i(n+1)\zeta} \int_1^{\infty} \frac{t^{-n-1} dt}{\sqrt{t-1} \sqrt{t-e^{-2i\zeta}}} \right]. \quad (33.39)$$

Finally, with the change of variable $t = e^p$, we obtain the Laplace-type representation

$$P_n(\cos \zeta) = \frac{2}{\pi} \operatorname{Re} \left[e^{i(n+1)\zeta} \int_0^{\infty} \frac{e^{-np} dp}{\sqrt{e^p-1} \sqrt{e^p-e^{-2i\zeta}}} \right]. \quad (33.40)$$

The Airy integral can likewise be transformed into a Laplace-type form. Setting $t = \sqrt{z}u$, we obtain

$$\operatorname{Ai}(z) = \frac{\sqrt{z}}{2\pi i} \int_C \exp\left(\frac{1}{3}z^{3/2}u(u^2-3)\right) du,$$

for a contour C depending on the argument of z . Letting $z^{3/2} = -\mu$, introducing the change of variable

$$u(u^2-3) = -3p,$$

and writing $u = f(p)$, we obtain

$$\operatorname{Ai}(z) = \frac{\mu^{1/6}}{2\pi i} \int_C e^{-\mu p} f'(p) dp.$$

The contour C and the function $f(p)$ will be analyzed in detail when we discuss the saddle point method; at that stage, the integral above will take the canonical Laplace form.

A similar transformation can be carried out for the sine integral $\operatorname{Si}(z)$. Other examples further illustrate this correspondence between special functions and Laplace-type integrals.

Example 1 (Error function). For large $y > 0$, the complementary error function can be written as

$$\frac{\sqrt{\pi}}{2} \operatorname{erfc}(y) := \int_y^{\infty} e^{-s^2} ds = \frac{1}{2} \int_{y^2}^{\infty} t^{-1/2} e^{-t} dt = \frac{\sqrt{x} e^{-x}}{2} \int_0^{\infty} \frac{e^{-xp}}{\sqrt{p+1}} dp, \quad x = y^2, \quad (33.41)$$

where we made the change of variables $t = y^2(p+1)$. Up to the prefactor $\frac{\sqrt{x}e^{-x}}{2}$, this representation shows that the asymptotic behavior of $\int_y^{\infty} e^{-s^2} ds$ as $y \rightarrow \infty$ is governed by the Laplace-type integral on the right-hand side, which controls the power-law corrections in $1/s$ for large s .

¹⁶As mentioned, the branch of $\sqrt{1-t}$ is the natural one for $t < 1$. Hence, for $t > 1$, we have $\sqrt{1-t} = -i\sqrt{t-1}$.

Example 2 (Gamma function). The Gamma function can be rewritten in a form that leads naturally to a Laplace-type representation:

$$\Gamma(n+1) = \int_0^\infty e^{-t} t^n dt \stackrel{t=ns}{=} n^{n+1} \int_0^\infty e^{n(-s+\ln s)} ds. \quad (33.42)$$

The function $-s + \ln s$ is monotonic on each of the intervals $[0, 1]$ and $[1, \infty)$. Splitting the integral as $\int_0^\infty = \int_0^1 + \int_1^\infty$, we perform in each part the substitution

$$p = t - \ln t, \quad dp = \left(-1 + \frac{1}{t}\right) dt.$$

Let f_1 and f_2 denote the inverse functions of $t - \ln t$ on $[0, 1]$ and $[1, \infty)$, respectively, and define

$$G(p) = \frac{1}{f_2(p)} - \frac{1}{f_1(p)}.$$

Then

$$\Gamma(n+1) = n^{n+1} e^{-n} \int_0^\infty e^{-np} G(p) dp. \quad (33.43)$$

Note 199

The function G admits an explicit expression in terms of the Lambert W function—the multivalued inverse of $x \mapsto xe^x$. Exponentiating the relation $f(p) - \ln f(p) = p$ gives

$$-f(p)e^{-f(p)} = -e^{-p-1}, \quad (33.44)$$

hence

$$f(p) = -W(-e^{-p-1}).$$

The two branches of f correspond to the two real branches of the Lambert function: for $p < 1$, $f_1(p) = -W(0, -e^{-p-1})$, and for $p > 1$, $f_2(p) = -W(-1, -e^{-p-1})$. Differentiating and substituting into the definition of G gives

$$G(p) = \frac{1}{1 + W(-1, -e^{-p-1})} - \frac{1}{1 + W(0, -e^{-p-1})}. \quad (33.45)$$

This Laplace-type representation of $\Gamma(n+1)$ forms the basis of Stirling's asymptotic expansion, which follows from the application of Watson's lemma to (33.43).

In the examples above—and indeed quite generally—the dependence on parameters is analytic (including the dependence on k in (33.34)), except for isolated special values. Infinity, however, is often a point of nonanalyticity: in such cases, the integral representations yield *divergent* asymptotic series, whose domains of convergence are empty.

34 Asymptotic Power Series

Definition 200

- An *asymptotic power series in powers of z at $z = 0$* is a formal expression

$$\sum_{k=0}^{\infty} c_k z^k, \quad (34.46)$$

where the coefficients c_k are arbitrary complex numbers. No assumption is made regarding convergence; the series is understood purely *formally*.

- An *asymptotic series at a point a* is an asymptotic power series at zero in the shifted variable $z = t - a$.
- Formal power series at infinity are defined analogously as

$$\sum_{k=0}^{\infty} \frac{c_k}{t^k}. \quad (34.47)$$

- Let $\{\beta_k\}_{k \geq 0}$ be a sequence satisfying

$$\operatorname{Re} \beta_0 < \operatorname{Re} \beta_1 < \operatorname{Re} \beta_2 < \cdots.$$

Asymptotic power series at $z = 0$ in noninteger (or more generally, complex) powers of the variable are formal sums of the type

$$\sum_{k=0}^{\infty} c_k z^{\beta_k}, \quad c_k, \beta_k \in \mathbb{C}. \quad (34.48)$$

- More generally, a sequence of functions $\{f_k\}_{k \geq 0}$ forms an *asymptotic scale* at $z = a$ if, for every k ,

$$\lim_{z \rightarrow a} \frac{f_{k+1}(z)}{f_k(z)} = 0.$$

An *asymptotic expansion at a in that scale* is then written formally as

$$f(z) \sim \sum_{k=0}^{\infty} f_k(z), \quad z \rightarrow a.$$

In this limit, z may be constrained to be in a sector $\mathcal{S}$ or along some curve γ .

Without loss of generality, one can usually assume that the point of expansion is 0 (or ∞ , depending on convenience for the problem at hand).

A simple example of a divergent asymptotic series (its radius of convergence is zero, as follows from the ratio test) is

$$\sum_{k=0}^{\infty} \frac{k!}{t^{k+1}}. \quad (34.49)$$

Definition 201: Operations on formal power series

Algebraic operations on formal power series—addition, multiplication, differentiation, integration, composition, and division (when the leading coefficient is nonzero)—are defined

coefficientwise, by imitating the relations that hold for convergent series. For instance,

$$\left(\sum_{k=0}^{\infty} c_k z^k \right) + \left(\sum_{k=0}^{\infty} d_k z^k \right) := \sum_{k=0}^{\infty} (c_k + d_k) z^k. \quad (34.50)$$

Definition 202: Big-O and little-o notation

We write $f(z) = O(z^\beta)$ as $z \rightarrow 0$ if $z^{-\beta} f(z)$ remains bounded near $z = 0$. If boundedness holds only within a sector

$$\mathcal{S} = \{z : \arg z \in (a, b)\}$$

or along a curve γ , we write $f(z) = O(z^\beta)$ as $z \rightarrow 0$ in $\mathcal{S}$ (or along γ).

Similarly, $f(z) = o(z^\beta)$ as $z \rightarrow 0$ means that $z^{-\beta} f(z) \rightarrow 0$ as $z \rightarrow 0$, possibly within a specified sector or along a curve. These definitions extend verbatim to other limits, such as $t \rightarrow \infty$.

Functions Asymptotic to a Power Series.

Definition 203

Let $\{\beta_k\}_{k \geq 0}$ be a sequence satisfying

$$\operatorname{Re} \beta_0 < \operatorname{Re} \beta_1 < \operatorname{Re} \beta_2 < \cdots.$$

We say that a function f admits the *asymptotic expansion*

$$f(z) \sim \sum_{k=0}^{\infty} c_k z^{\beta_k}, \quad c_k \in \mathbb{C},$$

as $z \rightarrow 0$ in a sector $\mathcal{S}$ or along a curve γ , if for every $n \in \mathbb{N}$,

$$f(z) - \sum_{k=0}^n c_k z^{\beta_k} = o(z^{\beta_n}), \quad z \rightarrow 0 \text{ in } \mathcal{S} \text{ or along } \gamma.$$

Example 204. If $f : \mathbb{R} \rightarrow \mathbb{R}$ is a C^∞ function, then at any point $a \in \mathbb{R}$ we have the asymptotic expansion

$$f(x) \sim \sum_{k=0}^{\infty} c_k (x - a)^k, \quad \text{as } x \rightarrow a \text{ along the real line,}$$

where $c_k = f^{(k)}(a)/k!$.

35 The Phragmén–Lindelöf Theorem

Theorem 205: Phragmén–Lindelöf, sectorial formulation

Let U be the open sector between two rays from the origin forming an angle π/β , where $\beta > \frac{1}{2}$. Assume that f is analytic in U , continuous on $\bar{U}$, and satisfies, for some constants $C_1, C_2, M > 0$ and $\alpha \in (0, \beta)$,

$$|f(z)| \leq C_1 e^{C_2 |z|^\alpha} \quad \text{for all } z \in U, \quad |f(z)| \leq M \quad \text{for all } z \in \partial U. \quad (35.1)$$

Then

$$|f(z)| \leq M, \quad \text{for all } z \in U. \quad (35.2)$$

Proof. By a rotation we may assume

$$U = \{z : 2|\arg z| < \pi/\beta\}.$$

Placing a branch cut in the complement of U , we can define an analytic branch of $\log z$ in U , and consequently of z^β . Setting

$$g(w) = f(w^{1/\beta}),$$

we may, without loss of generality, assume $\beta = 1$; then $\alpha \in (0, 1)$ and

$$U = \{z : |\arg z| < \frac{\pi}{2}\}.$$

Let $\alpha' \in (\alpha, 1)$ and $\varepsilon > 0$, and consider the analytic function

$$h(z) = e^{-\varepsilon z^{\alpha'}} f(z). \quad (35.3)$$

Since $\operatorname{Re} z^{\alpha'} = \operatorname{Re} |z| e^{i \arg(z) \alpha'} = |z| \cos \alpha' \arg(z) > 0$ for $z \in U$, we have $|e^{-\varepsilon z^{\alpha'}}| < 1$ there, and

$$|e^{-\varepsilon z^{\alpha'} + C_2 z^\alpha}| \longrightarrow 0 \quad \text{as } |z| \rightarrow \infty, \text{ for } z \in U.$$

Applying the maximum modulus principle to h on the truncated sector

$$U_R = \{z \in U : |z| < R\},$$

we obtain $|h(z)| \leq M$ on $\overline{U_R}$. Letting $R \rightarrow \infty$ gives $|h(z)| \leq M$ throughout $\overline{U}$.

Finally, letting $\varepsilon \rightarrow 0$ in (35.3) yields (35.2). ■

35.1 Laplace and Inverse Laplace Transforms

We will study Laplace and inverse Laplace transforms in more detail later. For now, let $F \in L^1(\mathbb{R}^+)$. By the dominated convergence theorem, the Laplace transform

$$(\mathcal{L}F)(x) := \int_0^\infty e^{-px} F(p) \, dp \quad (35.4)$$

is well defined and **continuous** in x on the closed right half-plane $\overline{\mathbb{H}^+}$. Using Fubini's theorem and Morera's theorem, one verifies that $(\mathcal{L}F)(x)$ is **analytic** in the open right half-plane $\mathbb{H}^+$. (More generally, if $F e^{-|\alpha|p} \in L^1$, then $\mathcal{L}F$ is defined for $\operatorname{Re} x > |\alpha|$.) The function F is uniquely determined by its Laplace transform, as shown below.

Lemma 206 (Uniqueness). *Assume $F \in L^1(\mathbb{R}^+)$ and that $\mathcal{L}F = 0$ on a set of x having an accumulation point in $\mathbb{H}^+$. Then $F = 0$ almost everywhere on $[0, \infty)$.*

Proof (OC). Since $\mathcal{L}F$ is analytic, it vanishes identically throughout the open right half-plane. By continuity on $\overline{\mathbb{H}^+}$, we may evaluate (35.4) at $x = -ik$, with $k \in \mathbb{R}$. Let Θ denote the Heaviside function. Then, for all $k \in \mathbb{R}$,

$$(\mathcal{L}F)(-ik) = 0 = (\mathcal{F}[F\Theta])(k),$$

where $\mathcal{F}$ denotes the Fourier transform. Since $F\Theta \in L^1$, the Fourier inversion theorem [10] implies that $F = 0$ almost everywhere on $\mathbb{R}^+$. ■

The following result (and its corollary) is of central importance in resurgence theory, showing that Laplace transforms *cannot exhibit asymptotic behavior involving exponentially small corrections*. We now make this statement precise.

Proposition 35.1 (Lower bound on decay rates of Laplace transforms, OC). *Assume $F \in L^1(\mathbb{R}^+)$ and that for some $\varepsilon > 0$,*

$$\mathcal{L}F(x) = O(e^{-\varepsilon x}) \quad \text{as } x \rightarrow +\infty. \quad (35.5)$$

Then $F = 0$ almost everywhere on $[0, \varepsilon]$.

Proof. We decompose the Laplace integral as

$$\int_0^\infty e^{-px} F(p) \, dp = \int_0^\varepsilon e^{-px} F(p) \, dp + \int_\varepsilon^\infty e^{-px} F(p) \, dp. \quad (35.6)$$

For the second term, we have

$$\left| \int_\varepsilon^\infty e^{-px} F(p) \, dp \right| \leq e^{-\varepsilon x} \int_\varepsilon^\infty |F(p)| \, dp \leq e^{-\varepsilon x} \|F\|_1 = O(e^{-\varepsilon x}). \quad (35.7)$$

Therefore,

$$g(x) := \int_0^\varepsilon e^{-px} F(p) \, dp = O(e^{-\varepsilon x}) \quad \text{as } x \rightarrow +\infty. \quad (35.8)$$

By Fubini's and Morera's theorems, the function g is entire. Define $h(x) = e^{\varepsilon x} g(x)$. Then h is entire and uniformly bounded on the real axis: indeed, by assumption there exists x_0 such that $|h(x)| \leq C$ for $x > x_0$, and by continuity $\max_{[0, x_0]} |h| = C' < \infty$, hence $|h(x)| \leq C + C'$ for all real x . Moreover, h is bounded on $i\mathbb{R}$ (by $\|F\|_1$). By the Phragmén–Lindelöf theorem (applied first in the first quadrant and then in the fourth, with $\beta = 2$, $\alpha = 1$), h is bounded in the closed right half-plane.

For $x = -s < 0$ we have

$$\left| e^{-se} \int_0^\varepsilon e^{sp} F(p) \, dp \right| = \left| \int_0^\varepsilon e^{s(p-e)} F(p) \, dp \right| \leq \int_0^\varepsilon |F(p)| \, dp \leq \|F\|_1. \quad (35.9)$$

Applying the Phragmén–Lindelöf theorem again (in the second and third quadrants) shows that h is bounded in the closed left half-plane as well; hence h is bounded on all of $\mathbb{C}$, and is therefore constant. By the Riemann–Lebesgue lemma, $h(x) \rightarrow 0$ for $x = is$ as $s \rightarrow +\infty$, so $h \equiv 0$. Consequently,

$$\int_0^\varepsilon F(p) e^{-isp} \, dp = \widehat{\mathcal{F}}(\chi_{[0, \varepsilon]} F)(s) = 0, \quad (35.10)$$

for all $s \in \mathbb{R}$, which implies the claim. ■

Corollary 207 (OC). *Assume $F \in L^1(\mathbb{R}^+)$ and $\mathcal{L}F = O(e^{-Ax})$ as $x \rightarrow +\infty$ for all $A > 0$. Then $F = 0$.*

Proof. This follows immediately from Proposition 35.1. ■

In this sense, no two distinct functions in $L^1(\mathbb{R}^+)$ that are real analytic on $(0, \infty)$ can have Laplace transforms differing by terms that are exponentially small. This observation will play an important role later on.

35.2 A Laplace Inversion Formula

Theorem 208

Assume that $c \geq 0$ and that f is analytic in the closed half-plane

$$H_c := \{z \in \mathbb{C} : \operatorname{Re} z \geq c\}.$$

Assume further that for all $c' \geq c$ and $t \in \mathbb{R}$ we have

$$|f(c' + it)| \leq g(t), \quad \text{where } g \in L^1(\mathbb{R}).$$

Define

$$F(p) = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{px} f(x) dx =: (\mathcal{L}^{-1}f)(p). \quad (35.11)$$

Then, for any $x \in \{z : \operatorname{Re} z > c\}$,

$$\mathcal{L}F(x) = \int_0^\infty e^{-px} F(p) dp = f(x). \quad (35.12)$$

Proof (OC). Let $x' = x'_1 + iy'_1$ with $x'_1 > c$. Then

$$\int_0^\infty dp \int_{c-i\infty}^{c+i\infty} |e^{p(s-x')} f(s)| |ds| \leq \int_0^\infty e^{p(c-x'_1)} dp \|g\|_1 = \frac{\|g\|_1}{x'_1 - c}. \quad (35.13)$$

Hence, by Fubini's theorem, the order of integration may be interchanged:

$$\begin{aligned} U(x') &= \int_0^\infty e^{-px'} \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} e^{ps} f(s) ds dp \\ &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} f(s) \left(\int_0^\infty e^{-p(x'-s)} dp \right) ds \\ &= \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{f(s)}{x' - s} ds. \end{aligned} \quad (35.14)$$

Since $g \in L^1(\mathbb{R})$, there exist sequences $\tau_n, \tau'_n \rightarrow +\infty$ such that $|g(\tau_n)|, |g(-\tau'_n)| \rightarrow 0$. Let x_2 be such that $c < x'_1 < x_2$, and consider the rectangular contour

$$B_n := \{z : \operatorname{Re} z \in [c, x_2], \operatorname{Im} z \in [-\tau'_n, \tau_n]\},$$

traversed clockwise. By the residue theorem,

$$\frac{1}{2\pi i} \int_{B_n} \frac{f(s)}{x' - s} ds = f(x'). \quad (35.15)$$

On the other hand,

$$\lim_{n \rightarrow \infty} \frac{1}{2\pi i} \int_{B_n} \frac{f(s)}{x' - s} ds = \frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{f(s)}{x' - s} ds - \frac{1}{2\pi i} \int_{x_2-i\infty}^{x_2+i\infty} \frac{f(s)}{x' - s} ds. \quad (35.16)$$

Moreover, by dominated convergence,

$$\int_{x'-i\infty}^{x'+i\infty} \frac{f(s)}{x' - s} ds \rightarrow 0 \quad \text{as } x' \rightarrow \infty. \quad (35.17)$$

Combining (35.15)–(35.17), we obtain

$$\frac{1}{2\pi i} \int_{c-i\infty}^{c+i\infty} \frac{f(s)}{x' - s} ds = f(x'),$$

which establishes (35.12). ■

35.3 Abstract Stokes Phenomena

This theorem shows that if an analytic function decays rapidly along some direction, then it must grow “correspondingly” along a complementary direction. The statement below is reminiscent of a theorem of Carlson (see [16]).

Theorem 209: (OC)

Assume $f \not\equiv 0$ is analytic in the closed right half-plane $\overline{\mathbb{H}^+}$ and that, for every $a > 0$,

$$f(t) = O(e^{-at}) \quad (t \in \mathbb{R}^+, t \rightarrow \infty).$$

Then, for every $b > 0$, the function

$$z \mapsto e^{-bz} f(z) \tag{35.18}$$

is unbounded in $\overline{\mathbb{H}^+}$.

Proof. Suppose, to the contrary, that there exist $b > 0$ and $M > 0$ such that

$$|e^{-bz} f(z)| \leq M \quad (z \in \overline{\mathbb{H}^+}).$$

Consider

$$\psi(z) = \frac{e^{-bz} f(z)}{(z+1)^2}. \tag{35.19}$$

Then ψ is analytic in $\overline{\mathbb{H}^+}$ and, for each vertical line $\operatorname{Re} z = c' \geq 0$,

$$|\psi(c' + it)| \leq \frac{M}{|(c' + 1) + it|^2},$$

so that $\sup_{c' \geq 0} |\psi(c' + it)| \leq g(t)$, where $g(t) = (1 + t^2)^{-1} \in L^1(\mathbb{R})$. Therefore ψ satisfies the hypotheses of Theorem 208, and its inverse Laplace transform $F = \mathcal{L}^{-1}\psi$ is well defined, with

$$\mathcal{L}F(x) = \psi(x) \quad (\operatorname{Re} x > 0).$$

Along the positive real axis,

$$\psi(x) = \frac{e^{-bx} f(x)}{(x+1)^2} = O(e^{-(b+a)x}) \quad \text{for every } a > 0,$$

hence $\psi(x) = O(e^{-Ax})$ as $x \rightarrow +\infty$ for every $A > 0$. Applying Corollary 207 to F (with $\mathcal{L}F = \psi$) yields $F \equiv 0$, and therefore $\psi \equiv 0$. By analyticity, $e^{-bz} f(z) \equiv 0$ in $\overline{\mathbb{H}^+}$, which forces $f \equiv 0$, a contradiction. ■

Let $\alpha > 2$.

Corollary 210 (OC). Assume $f \not\equiv 0$ is analytic in the closed sector

$$S = \{z : 2|\arg z| \leq \pi/\alpha\}, \quad \alpha > \frac{1}{2},$$

and satisfies

$$|f(t)| \leq Ce^{-t^\beta} \quad (t \in \mathbb{R}^+)$$

for some $\beta > \alpha$. Then, for every $\beta' \in (0, \beta)$, there exists a sequence $z_n \in S$ such that

$$|f(z_n) e^{-z_n^{\beta'}}| \longrightarrow \infty \quad \text{as } n \rightarrow \infty. \tag{35.20}$$

Proof. Map the sector S conformally to a right half-plane via

$$w = z^\alpha,$$

with the branch chosen so that S maps onto $\overline{\mathbb{H}^+}$. Apply Theorem 209 to the function

$$g(w) = f(w^{1/\alpha})$$

and to $e^{-bw^{\beta/\alpha}}g(w)$. The unboundedness of the latter along a subsequence in $\overline{\mathbb{H}^+}$ translates, under the inverse map, into the asymptotic behavior (35.20). ■

Exercise 211. *Give the details of the reduction used in the corollary:

- (i) Verify that $w = z^\alpha$ maps S onto $\overline{\mathbb{H}^+}$ and preserves analyticity up to the boundary.
- (ii) Translate the decay $|f(t)| \leq Ce^{-t^\beta}$ on $\mathbb{R}^+$ into a decay statement for $g(w)$ on $\mathbb{R}^+$.
- (iii) Conclude (35.20) by applying Theorem 209 to a suitable choice of $b > 0$.

35.4 Watson's Lemma

In many instances, integral representations of functions are amenable to analysis via Laplace transforms:

$$(\mathcal{L}F)(x) := \int_0^\infty e^{-xp} F(p) dp. \quad (35.21)$$

Watson's Lemma

This fundamental result states that the asymptotic series at infinity of $(\mathcal{L}F)(x)$ is obtained by formal term-by-term integration of the asymptotic series of $F(p)$ for small p , provided F admits such a series.

Lemma 212 (Watson). Let $\{\alpha_k\}_{k \geq 0} \subset \mathbb{C}$ satisfy

$$0 < \operatorname{Re} \alpha_0 < \operatorname{Re} \alpha_1 < \cdots,$$

and let $F \in L^\infty(\mathbb{R}^+)$. Assume that, as $p \rightarrow 0^+$,

$$F(p) = \sum_{k=0}^m c_k p^{\alpha_k-1} + o(p^{\alpha_m-1}). \quad (35.22)$$

Then, as $x \rightarrow \infty$ along any ray ρ in the open right half-plane $H = \{x : |\arg x| < \frac{\pi}{2}\}$,

$$f(x) = \mathcal{L}F(x) = \int_0^\infty e^{-xp} F(p) dp = \sum_{k=0}^m c_k \Gamma(\alpha_k) x^{-\alpha_k} + o(x^{-\alpha_m}).$$

Note 213: A generalization

1. Taking $F\chi_{[0,a]}$ instead of F in Lemma 212, we see that for any $a \in (0, \infty]$, as $x \rightarrow \infty$ along any ray ρ in the open right half-plane H ,

$$\int_0^a e^{-xp} F(p) dp = \sum_{k=0}^m c_k \Gamma(\alpha_k) x^{-\alpha_k} + o(x^{-\alpha_m}). \quad (35.23)$$

2. Since the question is the large x behavior of the Laplace transform, we can relax the assumption $F \in L^\infty(\mathbb{R}^+)$ to $e^{-ap}f \in L^\infty$ for some $a > 0$. L^∞ can be replaced by L^1 , by

integration by parts:

$$\int_0^\infty e^{-xp} F(p) dp = x \int_0^\infty e^{-xp} \left(\int_0^p F(s) ds \right) dp \quad (35.24)$$

after verifying that the relation (35.22) can be integrated term-by-term.

Proof of Watson's Lemma. Since for $\operatorname{Re} \alpha > 0$ we have $\mathcal{L}[p^{\alpha-1}] = \Gamma(\alpha) x^{-\alpha}$, the result follows by induction on m from the one-term lemma below. ■

Lemma 214. Let $F \in L^\infty(\mathbb{R}^+)$ and assume that for some $\beta \in \mathbb{C}$ with $\operatorname{Re} \beta > -1$ and some $c \in \mathbb{C}$,

$$\lim_{p \rightarrow 0^+} p^{-\beta} F(p) = c. \quad (35.25)$$

Then, for $x = \rho e^{i\varphi}$,

$$\lim_{\rho \rightarrow \infty} x^{\beta+1} \int_0^\infty F(p) e^{-px} dp = c \Gamma(\beta+1).$$

Proof (OC). After the change of variables $s = p\rho$, we obtain, by dominated convergence and using (35.25),

$$x^{\beta+1} \int_0^\infty F(p) e^{-px} dp = e^{i\varphi(\beta+1)} \int_0^\infty \frac{F(s/\rho)}{(s^\beta/\rho^\beta)} s^\beta e^{-se^{i\varphi}} ds \longrightarrow c e^{i\varphi(\beta+1)} \int_0^\infty s^\beta e^{-se^{i\varphi}} ds, \quad (\rho \rightarrow \infty). \quad (35.26)$$

Finally, by a change of contour,

$$e^{i\varphi(\beta+1)} \int_0^\infty s^\beta e^{-se^{i\varphi}} ds = \int_0^{\infty e^{-i\varphi}} u^\beta e^{-u} du = \int_0^\infty u^\beta e^{-u} du = \Gamma(\beta+1),$$

which proves the claim. ■

35.5 Applications of Watson's Lemma

The Gamma Function.

Remark 215. The function $G(p)$ is an analytic function in $\sqrt{p}$ and thus $G'(p)$ has a convergent Puiseux series

$$\sum_{k=-1}^\infty c_k p^{k/2} = \sqrt{2} p^{-1/2} + \frac{\sqrt{2}}{6} p^{1/2} + \frac{\sqrt{2}}{216} p^{3/2} - \frac{139\sqrt{2}}{97200} p^{5/2} + \dots$$

Thus, by Watson's Lemma, for large n we have

$$n! \sim \sqrt{2\pi n} n^n e^{-n} \left(1 + \frac{1}{12n} + \frac{1}{288n^2} - \frac{139}{51840n^3} + \dots \right) \quad (35.27)$$

Proof. It is easy to calculate this expansion directly from the definition, $s - \ln s = p$, without using the Lambert function.

We write $s = 1 + S$ and $t = 1 + p$ and the equation $s - \ln(s) = t$ becomes $S - \ln(1 + S) = p$. Note that $S - \ln(1 + S) = S^2 U(S)/2$ where $U(0) = 1$ and $U(S)$ is analytic for small S ; with the natural branch of the square root, $\sqrt{U(S)} = H(S)$ is also analytic. We rewrite $S - \ln(1 + S) = p$ as $SH(S) = \pm \sqrt{2}\sigma$ where $\sigma^2 = p$. Since $(SH(S))'|_{S=0} = 1$ the implicit function theorem ensures the existence of two functions $S_{1,2}(\sigma)$ (corresponding to the two choices of sign) which are analytic in σ . The concrete expansion may be gotten by implicit differentiation in $SH(S) = \pm \sqrt{2}\sigma$, for instance. ■

Legendre Polynomials. We start with the representation (33.40) and apply Watson's Lemma. We have the following convergent Puiseux expansion near $p = 0$:

$$\frac{1}{\sqrt{e^p - 1} \sqrt{e^p - e^{-2i\xi}}} = \frac{1}{\sqrt{1 - e^{-2i\xi}} \sqrt{p}} + \left(-\frac{1}{2(1 - e^{-2i\xi})^{3/2}} - \frac{\frac{1}{4}}{\sqrt{1 - e^{-2i\xi}}} \right) \sqrt{p} + \dots \quad (35.28)$$

(Indeed, this follows from $\sqrt{e^p - 1} = \sqrt{p} \sqrt{p^{-1}(e^p - 1)} = \sqrt{p} H(p)$ where $H(p) = \sqrt{1 + \frac{p}{2} + \frac{p^2}{6} + \dots}$ is analytic at zero.) Since the Laplace transform of $p^{-1/2}$ is

$$\mathcal{L}(p^{-1/2})(n) = \int_0^\infty e^{-np} p^{-1/2} dp = \sqrt{\pi} n^{-1/2},$$

the leading contribution is

$$P_n(\cos \xi) \sim \frac{2\sqrt{\pi}}{\pi} \operatorname{Re} e^{i(n+1)\xi} \frac{1}{\sqrt{1 - e^{-2i\xi}}} n^{-1/2}.$$

Using the elementary identity

$$\frac{1}{\sqrt{1 - e^{-2i\xi}}} = \frac{e^{i\xi/2}}{\sqrt{2} \sqrt{\sin \xi} \sqrt{i}} \quad (35.29)$$

we obtain after simplification

$$P_n(\cos \xi) = \sqrt{\frac{2}{\pi n \sin \xi}} \cos\left((n + \frac{1}{2})\xi - \frac{\pi}{4}\right) + O(n^{-3/2}), \quad n \rightarrow \infty. \quad (35.30)$$

Watson's lemma applies equally for complex ξ , provided that $e^{\pm 2i\xi} \notin \mathbb{R}^+$; hence the asymptotic expression (35.30) remains valid under the same condition for complexified ξ .

The Error Function. We start from (33.41). For small p , the following convergent expansion holds:

$$\frac{1}{\sqrt{1+p}} = 1 - \frac{1}{2}p + \frac{3}{8}p^2 - \frac{5}{16}p^3 + \frac{35}{128}p^4 - \frac{63}{256}p^5 + \dots \quad (35.31)$$

Applying Watson's lemma to the Laplace-type integral representation in (33.41), we obtain the classical asymptotic expansion, valid in the open right half-plane:

$$\operatorname{erfc}(y) \sim \frac{e^{-y^2}}{y\sqrt{\pi}} \left(1 - \frac{1}{2y^2} + \frac{3}{4y^4} - \frac{15}{8y^6} + \frac{105}{16y^8} - \frac{945}{32y^{10}} + \dots \right), \quad y \rightarrow \infty. \quad (35.32)$$

Asymptotics of Taylor Coefficients. There exists a duality between the behavior of the Maclaurin coefficients of a function and the nature and location of its singularities in the complex domain, assuming for simplicity that these singularities form a discrete set without accumulation points. We consider the following representative case.

Assumptions. Let f be analytic in

$$D = \mathbb{D}_{1+\varepsilon}(0) \setminus \bigcup_{i=1}^n C_i,$$

where the C_i are radial cuts joining $z_i = e^{i\varphi_i} \in \partial\mathbb{D}$ to $z'_i \in \partial\mathbb{D}_{1+\varepsilon}(0)$, with $1/2 > \varepsilon > 0$. Assume further that the singularities are power-like: at each z_i , we have

$$f(z) = A_{i0}(z) + \sum_{j=1}^{n_i} (z_i - z)^{\beta_{ij}} A_{ij}(z), \quad A_{ij} \text{ analytic at } z_i = |z_i|e^{i\varphi_i}, \quad \beta_{ij} \notin \mathbb{N} \cup \{0\}. \quad (35.33)$$

The condition $\beta_{ij} \notin \mathbb{N} \cup \{0\}$ ensures that $(z_i - z)^{\beta_{ij}} A_{ij}(z)$ is genuinely nonanalytic at z_i —otherwise such a term could be absorbed into $A_{i0}(z)$.

The branch is chosen as follows: setting $s = e^{i\varphi}t$, for small t we have

$$(e^{i\varphi} - z)^\beta = e^{i\varphi\beta}(1 - t)^\beta = e^{i\varphi\beta}(1 - \beta t + o(t)). \quad (35.34)$$

Let $\mathbb{D}_{i\varepsilon}$ be disks of radius ε centered at z_i and assume f is continuous in $D_1 = \overline{\mathbb{D} \setminus \bigcup_i \mathbb{D}_{i\varepsilon}}$.

Write

$$f(z) = \sum_{m=0}^{\infty} c_m z^m, \quad z \in \mathbb{D}. \quad (35.35)$$

Theorem 216

Under these assumptions,

$$c_m \sim \sum_{i=1}^n \sum_{j=1}^{n_i} e^{-im\varphi_i} m^{-\beta_{ij}-1} \sum_{k=0}^{\infty} \frac{C_{ijk}}{m^k}, \quad m \rightarrow \infty, \quad (35.36)$$

where the coefficients C_{ijk} can be determined in closed form from the φ_i , the β_{ij} , and the Taylor coefficients of A_{ij} at z_i .

More generally, the same arguments yield the following result.

Note 217

(a) Let $R > 0$ and assume f be analytic in

$$D = \mathbb{D}_R(0) \setminus \bigcup_{i=1}^n C_i,$$

where $r_i < R$; $i = 1, \dots, n$ and the C_i are radial cuts joining $z_i = r_i e^{i\varphi_i} \in \partial\mathbb{D}$ to $z'_i \in \partial\mathbb{D}_R(0)$. Assume further that the singularities are power-like: at each z_i , we have

$$f(z) = A_{i0}(z) + \sum_{j=1}^{n_i} (z_i - z)^{\beta_{ij}} A_{ij}(z), \quad A_{ij} \text{ analytic at } z_i = |z_i|e^{i\varphi_i}, \quad \beta_{ij} \notin \mathbb{N} \cup \{0\}. \quad (35.37)$$

and let $\mathbb{D}_{i\varepsilon}$ be disks of radius ε centered at z_i and assume f is continuous in $D_1 = \overline{\mathbb{D} \setminus \bigcup_i \mathbb{D}_{i\varepsilon}}$. Let $r = \min_i r_i$. Under these assumptions,

$$c_m \sim \sum_{j \leq n_i; i \leq n; r_i = r} r_i^{-m} e^{-im\varphi_i} m^{-\beta_{ij}-1} \sum_{k=0}^{\infty} \frac{C_{ijk}}{m^k}, \quad m \rightarrow \infty, \quad (35.38)$$

The sum is over $r_i = r$ since the terms with $r_i > r$ are relatively exponentially small.

(b) The disk $\mathbb{D}_{1+\varepsilon}$ can be replaced by any $\mathbb{D}_R$, substituting the circles $\mathbb{D}_{i\varepsilon}$ with elliptical domains tangent to the outer circle and containing the cuts strictly inside. This allows for pushing the contours in the representation (35.38) below all the way to infinity, if the large z behavior of $f(z)$ allows for it.

We first prove two auxiliary lemmas that reduce the problem to a simpler one. The first lemma expresses a superposition principle: up to corrections that are exponentially small relative to the coefficients, the contributions of the terms $(z_i - z)^{\beta_{ij}} A_{ij}(z)$ are additive.

Lemma 218. *We have*

$$c_m = \frac{1}{2\pi i} \sum_{i=1}^n \sum_{j=1}^{n_i} \oint_{\partial \mathbb{D}_{i\varepsilon}} s^{-m-1} (z_i - s)^{\beta_{ij}} A_{ij}(s) ds + O(e^{-m \ln(1+\varepsilon)}). \quad (35.39)$$

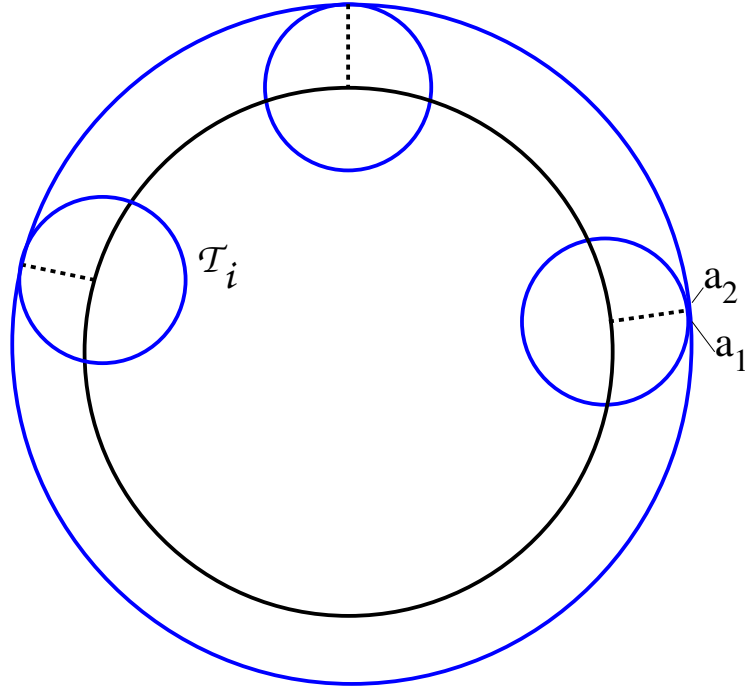


Figure 38: Deformation of the integration contour.

Proof of Lemma 218. We apply Cauchy's integral formula for the coefficients c_m on a small circle centered at the origin, deforming the contour homotopically to the boundary of D_1 . The representation of f near z_i is additive in $A_{i0}(z)$ and the $\{(z_i - z)^{\beta_{ij}} A_{ij}(z)\}_{j \leq n_i}$. The total contribution of the integral along ∂D_1 is bounded by $2\pi(1+\varepsilon)^{-m} \max_{\partial D_1} |f(z)|$, which yields the result. ■

Lemma 219. *Up to corrections that are exponentially small relative to the coefficients, one may assume that $\operatorname{Re} \beta_{ij} > -1$ for all i, j .*

Proof. Each term in the double sum in (35.39) is of the form

$$J = \oint_{\mathbb{D}_\varepsilon} s^{-m-1} (\alpha - s)^\beta A(s) ds. \quad (35.40)$$

Repeated integration by parts increases the real part of β . After one integration by parts, we have

$$(\beta + 1)J = [s^{-m-1}(\alpha - s)^{\beta+1}A(s)]_{a_1}^{a_2} + \oint_{\mathbb{D}_\varepsilon} s^{-m-1}(\alpha - s)^{\beta+1}\tilde{A}(s) ds \\ = \oint_{\mathbb{D}_\varepsilon} s^{-m-1}(\alpha - s)^{\beta+1}\tilde{A}(s) ds + O\left(e^{-(m+1)\ln(1+\varepsilon)}\right), \quad (35.41)$$

where a_1 and a_2 are the right and left sides of the cut on $\mathbb{D}_{1+\varepsilon}$ and $\tilde{A}(s) = (m+1)A(s) - sA'(s)$ is analytic at α . Thus, each integration by parts increases $\operatorname{Re} \beta$ by one. ■

Proof of Theorem 216. Using Lemmas 218 and 219, it suffices to analyze the large- m behavior of J in (35.40) under the assumption $\operatorname{Re} \beta > -1$. By homotopy and dominated convergence, we can shrink the contour $\partial\mathbb{D}_\varepsilon$ to the two sides of the cut, traversed in opposite ways.

Let $s = te^{i\varphi}$. As t moves from 0 to a point on the left side of the cut, $\arg(1-t)$ changes by $-\pi$, giving $(1-t)^\beta = e^{-i\pi\beta}|1-t|^\beta$; on the opposite side we have $(1-t)^\beta = e^{i\pi\beta}|1-t|^\beta$. Hence,

$$J = -e^{-im\varphi} \frac{e^{i\pi\beta} - e^{-i\pi\beta}}{2\pi i} \int_1^{1+\varepsilon} t^{-m-1}(t-1)^\beta A(te^{i\varphi}) dt \stackrel{t=e^p}{=} -\frac{\sin(\pi\beta) e^{-im\varphi}}{\pi} \int_0^b e^{-mp}(e^p-1)^\beta A(e^{p+i\varphi}) dp. \quad (35.42)$$

Since $A(s)$ is analytic at $s = \alpha$, we can expand

$$A(e^{p+i\varphi}) = \sum_{k=0}^{\infty} a_k p^k,$$

and therefore

$$(e^p - 1)^\beta A(e^{p+i\varphi}) = a_0 p^\beta + \frac{1}{2}(a_0\beta + 2a_1)p^{\beta+1} \\ + \frac{1}{24}[a_0\beta(3\beta+1) + 12(a_1\beta + 2a_2)]p^{\beta+2} + o(p^{\beta+2}). \quad (35.43)$$

Applying Watson's lemma then yields, as $m \rightarrow \infty$,

$$J \sim -\frac{\sin(\pi\beta) e^{-im\varphi} \Gamma(1+\beta)}{\pi m^{1+\beta}} \left(a_0 + \frac{(1+\beta)(a_0\beta + 2a_1)}{2m} \right. \\ \left. + \frac{(1+\beta)(2+\beta)[a_0\beta(3\beta+1) + 12(a_1\beta + 2a_2)]}{24m^2} + o(m^{-2}) \right). \quad (35.44)$$

■

36 Riemann surfaces, an introduction

Here I follow losely the approach in [5].

By definition, an n -dimensional manifold is a Hausdorff topological space Ω such that every point $x \in \Omega$ belongs to an open set in Ω which is homeomorphic to an open subset of $\mathbb{R}^n$.

In particular, when $n = 2$, one can endow a two-dimensional manifold with additional structure that reflects the complex plane. This leads to the notion of a *complex chart* and, more generally, of a *Riemann surface*.

Definition 220

Let Ω be a two-dimensional manifold. A *complex chart* on Ω is a homeomorphism

$$\varphi : U \rightarrow V$$

of an open subset $U \subset \Omega$ onto an open subset $V \subset \mathbb{C}$. Two complex charts $\varphi_i : U_i \rightarrow V_i$ ($i = 1, 2$) are said to be *holomorphically compatible* if the transition map

$$\varphi_2 \circ \varphi_1^{-1} : \varphi_1(U_1 \cap U_2) \rightarrow \varphi_2(U_1 \cap U_2)$$

is biholomorphic.

A *complex atlas* on Ω is a family $\mathcal{F} = \{\varphi_i : U_i \rightarrow V_i\}_{i \in I}$ such that $\{U_i\}_{i \in I}$ covers Ω and the charts are pairwise holomorphically compatible.

Two complex atlases $\mathcal{F}$ and $\mathcal{F}'$ are said to be *analytically equivalent* if every chart in $\mathcal{F}$ is holomorphically compatible with every chart in $\mathcal{F}'$.

Remark 221.

- The restriction of a complex chart φ to a smaller open set $U' \subset U$ is holomorphically compatible with φ .
- Since biholomorphisms preserve analyticity, analytic equivalence of atlases is indeed an equivalence relation.

Definition 222

A *complex structure* Σ on a two-dimensional manifold Ω is an equivalence class of analytically equivalent atlases.

This complex structure Σ is determined by any representative atlas $\mathcal{F}$, and there exists a unique *maximal atlas* $\mathcal{F}^*$ consisting of all complex charts on Ω that are holomorphically compatible with every chart in $\mathcal{F}$.

Definition 223

A *Riemann surface* Ω is a pair (Ω, Σ) where Ω is a connected two-dimensional manifold and Σ is a complex structure on Ω .

Bianalytic transformations are the natural isomorphisms of domains and by Remark 221, more generally, of Riemann surfaces in complex analysis.

By convention, the complex structure Σ is often omitted from the notation when it is clear from the context. By definition, a Riemann surface is locally equivalent to an open set of $\mathbb{C}$. Since each point of Ω comes equipped with a whole family of indistinguishable charts, the only notions and structures that pass from $\mathbb{C}$ to Ω are those which are *independent of the choice of chart*, i.e. those invariant under biholomorphic maps.

Typically, we cannot embed Riemann surfaces in $\mathbb{R}^3$ and they are not easy to depict. In common cases, one useful way to handle Riemann surfaces is to think of them as spaces of curves with a fixed point p (see Note 224 below) modulo homotopies. The variable endpoints of these curves are the points on the Riemann surface. Thus, the log that we have defined by integrating ds/s over various paths avoiding zero is naturally generating such a set of equivalence classes.

An important object is **the fundamental group** $\pi_1(\Omega, p)$ generated by the homotopy classes of *closed curves* joining p to p under composition. Here, the product is concatenation of curves, the inverse of γ , γ^{-1} is $t \mapsto \gamma(1 - t)$ and $1 = \gamma\gamma^{-1}$.

Loops. Given a topological space Ω , a *loop based at p* is, by definition, a continuous curve that starts and ends at p :

$$\gamma: [0, 1] \rightarrow \Omega; \quad \gamma(0) = \gamma(1) = p$$

Homotopy of loops. A homotopy is a continuous interpolation between two loops. More precisely, a homotopy between two loops

$$\gamma, \gamma': [0, 1] \rightarrow \Omega$$

(based at the same point p) is a continuous map

$$h: [0, 1] \times [0, 1] \rightarrow \Omega,$$

such that

$$h(0, t) = h(1, t) = p \quad \text{for all } t \in [0, 1],$$

When such a homotopy h exists, γ and γ' are said to be *homotopic*. The relation “ γ is homotopic to γ' ” is an equivalence relation. The set of equivalence classes can be considered:

$$\pi_1(\Omega, p) := \{\text{all loops } \gamma \text{ based at } p\} / \text{homotopy}.$$

with the group structure below, is called the *fundamental group* of the topological space Ω at the base point p .

Note 224

It turns out that, up to isomorphism, the choice of p choice makes no difference as long as the space Ω is path-connected: more precisely, one obtains an isomorphism by pre- and post-concatenating with a path between the two basepoints.

The group structure on $\pi_1(\Omega, p)$. The set $\pi_1(\Omega, p)$ becomes a group, the *fundamental group* of Ω at p using the concatenation of loops as their product. Given

$$\gamma_0, \gamma_1: [0, 1] \rightarrow \Omega,$$

their product is defined to be the loop

$$\gamma_0 \cdot \gamma_1: [0, 1] \rightarrow \Omega, \quad (\gamma_0 \cdot \gamma_1)(t) = \begin{cases} \gamma_0(2t), & 0 \leq t \leq \frac{1}{2}, \\ \gamma_1(2t - 1), & \frac{1}{2} \leq t \leq 1. \end{cases}$$

We then define the product of *the equivalence classes* $[\gamma_0]$ and $[\gamma_1]$ by

$$[\gamma_0] \cdot [\gamma_1] := [\gamma_0 \cdot \gamma_1].$$

You can check that the product defined above does not depend on the choice of representatives and therefore is a well-defined operation on the set $\pi_1(\Omega, p)$ and is an associative operation on equivalence classes.

The neutral element is the equivalence (homotopy) class of the constant loop, $\gamma([0, 1]) = \{p\}$. The inverse of a loop is the same loop but traversed in the opposite direction, namely

$$\gamma^{-1}(t) := \gamma(1 - t).$$

and $[\gamma]^{-1} := [\gamma^{-1}]$. The structure above turns $\pi_1(\Omega, p)$ into a group.

Covering maps. A covering map (also called a covering or projection) is a surjective open map $f : Y \rightarrow \Omega$ that is locally a homeomorphism. In a covering map, the preimages $f^{-1}(z)$ are discrete sets in Y whose cardinality (finite or infinite) do not depend on z .

Exercise 225. (a) Check that the fundamental group of $\mathbb{C} \setminus \{0\}$ is the free group generated by one element, an abelian group isomorphic to $\mathbb{Z}$. Hence, $\pi_1(\mathbb{C} \setminus \{0\}) \cong \mathbb{Z}$.

(b) For each subgroup $n\mathbb{Z}$, we can form the quotient space

$$X_n = \mathbb{C} / (2\pi i n \mathbb{Z}),$$

and the map

$$p_n : X_n \longrightarrow \mathbb{C} \setminus \{0\}, \quad [z] \longmapsto e^z,$$

is an n -fold covering.

Intuitively, we have “wrapped” the complex plane n times before identifying points that differ by $2\pi i n$.

36.1 Universal covers.

The universal cover of a connected topological space Ω is a simply connected space Y with a map $f : Y \rightarrow \Omega$ that is a covering map. If Ω is simply connected, i.e., has a trivial fundamental group, then it is its own universal cover. For instance, the sphere $S^2 = \hat{\mathbb{C}}$ is its own universal cover. The universal cover is always unique (up to an isomorphism) and, under very mild assumptions, always exists. In fact, the universal cover of a topological space Ω exists iff the space Ω is connected, locally pathwise-connected, and semilocally simply connected ¹⁷.

Deck transformations. Let $\pi : Y \rightarrow \Omega$ be a covering map of topological spaces. A *deck transformation* (or *covering transformation*) is a homeomorphism

$$\varphi : Y \longrightarrow Y$$

such that

$$\pi \circ \varphi = \pi.$$

You can check that any continuous map between Riemann surfaces that is a local homeomorphism and is compatible with a holomorphic structure is automatically holomorphic (see §36.5).

The set of all deck transformations forms a group under composition, called the *deck transformation group* or *group of covering transformations*, usually denoted

$$\text{Deck}(\pi) \quad \text{or} \quad \text{Aut}(Y/\Omega).$$

Heuristics. A deck transformation is a symmetry of the covering space that

- moves points around *within the same fiber*,
- keeps the projection to the base space unchanged, and
- is continuous from Y to Y .

¹⁷A space Ω is locally simply connected if a small enough neighborhood of any $x \in \Omega$ is simply connected. Semilocally simply connected is a weaker notion than locally simply connected. See [6].

1. **Square root surface over the punctured plane. Example.** Consider the map

$$\pi : \mathbb{C} \setminus \{0\} \longrightarrow \mathbb{C} \setminus \{0\}, \quad \pi(z) = z^2.$$

This is a 2-to-1 covering map.

A deck transformation must satisfy $\pi \circ \varphi = \pi$, and the only nontrivial such map is

$$\varphi(z) = -z.$$

Thus the deck transformation group is

$$\text{Deck}(\pi) = \{ \text{id}, z \mapsto -z \} \cong \mathbb{Z}/2\mathbb{Z}.$$

2. **Torus.** The deck transformations of $\mathbb{C} \xrightarrow{\pi} \mathbb{C}/\Lambda$, are the translations $z \mapsto z + \lambda$, $\lambda \in \Lambda$, so

$$\text{Deck}(\pi) \cong \Lambda.$$

36.1.1 Basic examples

$\mathbb{C} \setminus \{0\}$ as a Riemann surface. **The exponential map.** The map

$$\pi : \mathbb{C} \rightarrow \mathbb{C} \setminus \{0\}, \quad \pi(z) = e^z.$$

is a covering map of $\Omega = \mathbb{C} \setminus \{0\}$, with universal cover $Y = \mathbb{C}$. The deck transformations are

$$\text{Deck}(\pi) = \{ z \mapsto z + 2\pi i k : k \in \mathbb{Z} \} \cong \mathbb{Z},$$

Hence, as found in Exercise 225, $\pi_1(\mathbb{C} \setminus \{0\}) \cong \mathbb{Z}$.

Note 226

Not every bijection between “sheets” of a covering is a deck transformation. Deck transformations must be global homeomorphisms compatible with the covering.

of this countable fiber. Note however the **continuity implications** for the exponential covering map: so the induced permutations σ_w must vary continuously with w . Choose a branch of the log. The fibers of the exponential map are of the form $\exp^{-1}(w) = \{\log w + 2\pi i k : k \in \mathbb{Z}\}$. Continuity is imposed on the transformation of the whole covering space, and not on fibers. In particular, a map that, for example, “swaps two horizontal strips” of $\mathbb{C}$ is not continuous, as you can easily check.

The cylinder as a Riemann surface. Let

$$\Lambda = 2\pi i \mathbb{Z},$$

the rank-one lattice generated by $2\pi i$. The quotient

$$C = \mathbb{C}/\Lambda$$

is a Riemann surface obtained from the complex plane by identifying each point with its vertical translates $z \mapsto z + 2\pi in$, $n \in \mathbb{Z}$. Topologically, C is a cylinder:

$$C \cong S^1 \times \mathbb{R}.$$

The canonical projection

$$\pi : \mathbb{C} \longrightarrow C, \quad \pi(z) = z \bmod 2\pi i,$$

is the *universal covering map* of the cylinder. It is holomorphic and locally biholomorphic. Its deck transformation group is exactly the lattice:

$$\text{Deck}(\pi) = \{ z \mapsto z + 2\pi in \mid n \in \mathbb{Z} \} \cong \mathbb{Z}.$$

Relation with $\mathbb{C} \setminus \{0\}$. The composition $\widehat{\exp} = \exp \circ \pi$ is well-defined, bijective and bianalytic between $C = \mathbb{C}/\Lambda$ and $\mathbb{C} \setminus \{0\}$ so the cylinder is conformally equivalent to the punctured plane.

The Universal Cover of the Torus. Let

$$\Lambda = \{ m\omega_1 + n\omega_2 \mid m, n \in \mathbb{Z} \},$$

with $\omega_1, \omega_2 \in \mathbb{C}$ linearly independent over $\mathbb{R}$. The (complex) torus is the quotient

$$T = \mathbb{C}/\Lambda,$$

and the canonical projection

$$\pi : \mathbb{C} \longrightarrow T, \quad \pi(z) = z \bmod \Lambda,$$

is the *universal covering map* of the torus. It is holomorphic and locally biholomorphic.

36.2 Properly discontinuous groups

Intuitive idea. A group G of homeomorphisms (or diffeomorphisms, or automorphisms) of a topological space Ω acts *properly discontinuously* if point there is a neighborhood U such that only finitely many group elements move it back into U . Then, the quotient Ω/G is manifold, if Ω is.

Formal definition.

A group G acting on a topological space Ω is said to act properly discontinuously if any of the following equivalent conditions hold:

(Neighborhood condition). For every $x \in \Omega$, there exists an open neighborhood U of x such that

$$g(U) \cap U = \emptyset \quad \text{for all but finitely many } g \in G.$$

(In most geometric cases, only the identity element fixes U .)

(Compact-intersection condition). For every compact subset $K \subset \Omega$, the set

$$\{g \in G : g(K) \cap K \neq \emptyset\}$$

is finite.

Key theorem (regular coverings). For a universal covering $Y \rightarrow \Omega$:

- the deck transformation group acts freely on Y ,
- properly discontinuously, and
- $\Omega \cong Y / \text{Deck}(\pi)$.
- Moreover,

$$\text{Deck}(\pi) \cong \pi_1(\Omega)$$

(up to isomorphism), though not canonically without choosing a basepoint and a lift.

Examples of Riemann surfaces. Clearly, $\mathbb{C}$ itself is a Riemann surface. So are domains $\mathcal{D} \subset \mathbb{C}$.

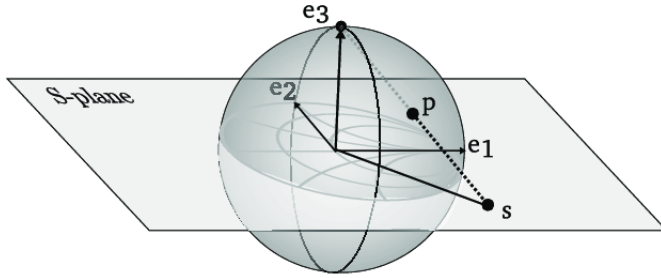


Figure 39: The Riemann sphere (taken from Google images)

An important surface in complex analysis is the Riemann sphere $S^2 = \hat{\mathbb{C}}$. The complex plane contains the equator of $\hat{\mathbb{C}}$, represented by $\partial\mathbb{D} \subset \mathbb{C}$. For any point z in the plane, one draws a line segment from the north pole of the sphere through the sphere, cutting the sphere at some point $p(z)$. Clearly, to any point z in the plane, there corresponds a unique $p(z)$ on $\hat{\mathbb{C}}$. $0 \in \mathbb{C}$ corresponds to the south pole in $\hat{\mathbb{C}}$. There is no point in $\mathbb{C}$ corresponding to the north pole on $\hat{\mathbb{C}}$; this is “the point at infinity”. Points in the southern hemisphere map to points in $\mathbb{D}$ while those in the northern hemisphere are in $\overline{\mathbb{D}}^c$.

The Riemann Surface of the Log. We can realize the Riemann surface $\mathcal{R}$ of the log as the space of equivalence classes of curves in $\mathbb{C} \setminus \{0\}$ originating at 1 modulo homotopies in $\mathbb{C} \setminus \{0\}$. The final result is a connected, simply connected, surface that can be viewed as a spiraling parking garage with infinitely many levels extending both upward and downward. You can check that $\log(z) = \int_{\gamma_{1,z}} s^{-1} ds$ where $\gamma_{1,z}$ is a smooth curve in $\mathbb{C} \setminus \{0\}$ joining 1 and z only depends on $[\gamma_{1,z}]$, and thus log is well defined on $\mathcal{R}$.

The Riemann surface of the square root $\mathcal{R}_2$. As the log, the square root, definable in terms of the log is analytic on the Riemann surface of the log. We note however, from its log-based definition, that closed curves starting at some $z_0 \neq 0$, say $z_0 = 1$ that encircle 0 twice yield the same value of the square root, here 1. What is different from the case of the log is that the fundamental group has a relation: $a^2 = 1$. We have $\pi_1(\mathcal{R}_2) = \mathbb{Z}/(2\mathbb{Z}) = \mathbb{Z}_2$.

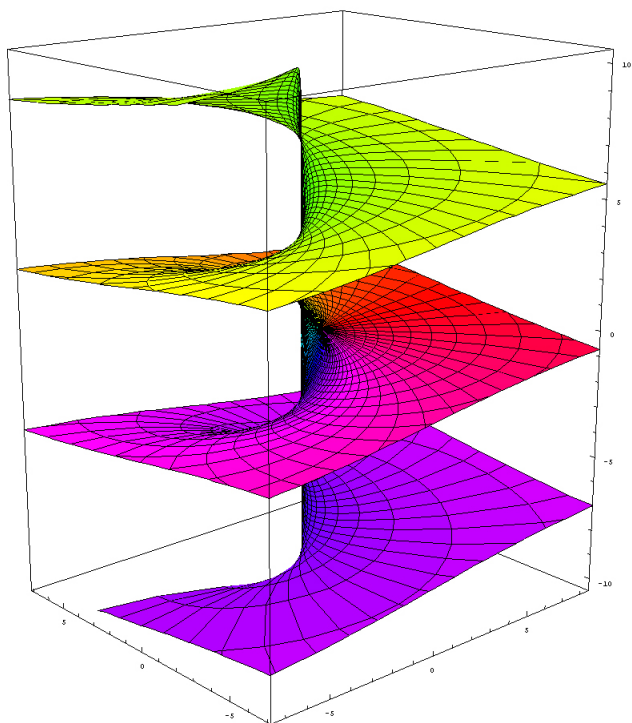


Figure 40: The Riemann surface of the log (from https://en.wikipedia.org/wiki/Complex_logarithm)

The Riemann surface of the square root is different from that of the log: the second floor of the “infinite parking lot” is glued to the first. It is not possible to embed this surface in $\mathbb{R}^3$, but we can draw a more or less suggestive picture, below.

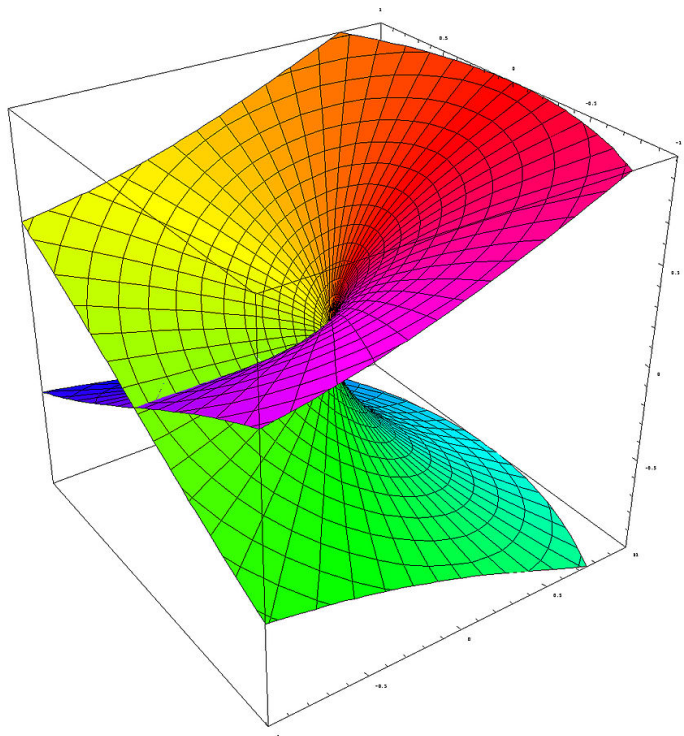


Figure 41: An approximate depiction of the Riemann surface of the square root (from <https://commons.wikimedia.org/wiki/>)

36.3 The uniformization theorem: simply connected case

For details and proofs of the results in this and the next subsection see, e.g., [11].

Theorem 227

Every simply connected Riemann surface is conformally equivalent to exactly one of the following three domains:

1. the Riemann sphere $\hat{\mathbb{C}}$,
2. the complex plane $\mathbb{C}$,
3. the unit disk $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$.

Note 228

Only special, relatively simple, Riemann surfaces are conformally equivalent to the Riemann sphere or with the plane. It is case 3. that typically occurs.

36.4 General form of the uniformization theorem

Recall that every Riemann surface Ω admits a universal covering surface $\tilde{\Omega}$ which is simply connected. More generally,

Theorem 229

A Riemann surface Ω is conformally equivalent to a quotient

$$\Omega \cong \tilde{\Omega}/\Gamma,$$

where Γ is a properly discontinuous group of automorphisms of $\tilde{\Omega}$.

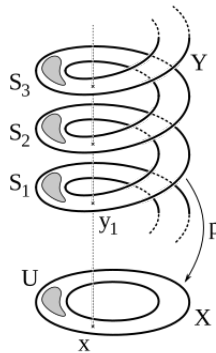


Figure 42: The covering space of an annulus (adapted from Wikimedia Commons).

Example. Let us uniformize Ω_0 , the Riemann surface of the log. That means finding a bianalytic bijection from Ω_0 to one of $\hat{\mathbb{C}}, \mathbb{C}, \mathbb{D}$.

The log is well defined by $\log z = \int_1^z ds/s$ on curves from 1 to z modulo homotopies in $\hat{\mathbb{C}} \setminus \{0\}$, and is manifestly analytic in z . The fact that it is injective follows from the calculation

$\log z_1 = \log z_2 \Rightarrow \exp \log z_1 = \exp \log z_2 \Rightarrow z_1 = z_2$. It is also onto since any $z \neq 0$ can be written as $\log x + i\varphi + 2N\pi i$ for some $x \in \mathbb{R}^+$, $\varphi \in [0, 2\pi)$ and $N \in \mathbb{Z}$.

The Riemann surface of the log is uniformized by the log onto $\mathbb{C}$.

In the opposite direction, let us look again at the action of the exponential on $\mathbb{C}$. As we have seen, $\exp$ is one-to-one bi-analytic from any vertical strip

$$S_n\{z = u + iv : u \in \mathbb{R}, v \in [2n\pi i, 2(n+1)\pi i)\} \rightarrow \mathbb{C} \setminus \{0\}$$

(think again of the polar representation of complex numbers). If we take $\exp$ to be defined on $\mathbb{C}$, it is not invertible because it is many-to-one. It is not univalent. But in each strip it is one-to-one and we can take as its inverse $\log |z| + i\varphi + 2n\pi i$. By taking $\cup_n S_n$ we cover all the possible values of $\log$. We see that, if we define $\exp$ from $\mathbb{C}$ to the Riemann surface of the log, then it is one-to-one, onto and bi-analytic.

36.5 Apendix: Deck Transformations of Riemann surfaces are holomorphic

Let $\pi : Y \rightarrow \Omega$ be a holomorphic covering map between Riemann surfaces, and let $\varphi : Y \rightarrow Y$ be a deck transformation, so $\pi \circ \varphi = \pi$.

Fix $z \in Y$. Choose charts

$$(U, \psi), \quad (U', \psi'), \quad (V, \chi)$$

around z , $\varphi(z)$, and $\pi(z)$ respectively, such that $\pi(U) = V$, $\pi(U') = V$, and in these charts the restrictions $\pi|_U$ and $\pi|_{U'}$ become biholomorphisms. Then on U we have

$$\varphi = (\pi|_{U'})^{-1} \circ \pi|_U.$$

Passing to charts, the local expression of φ is

$$\psi' \circ \varphi \circ \psi^{-1} = (\psi' \circ (\pi|_{U'})^{-1} \circ \chi^{-1}) \circ (\chi \circ \pi|_U \circ \psi^{-1}),$$

a composition of holomorphic maps. Hence φ is holomorphic in charts. Applying the same argument to φ^{-1} shows that φ is a biholomorphism of Y .

37 The modular group

In preparation for elliptic function theory, define the **modular group** as the group of linear fractional transformations

$$z \mapsto \frac{az + b}{cz + d}, \quad a, b, c, d \in \mathbb{Z}, \quad ad - bc = \pm 1,$$

which is isomorphic to $\Gamma = PSL(2, \mathbb{Z})$, the group of matrices

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad a, b, c, d \in \mathbb{Z}, \quad ad - bc = \pm 1,$$

with matrix multiplication and with the identification $M \sim -M$.

Note 230

The fact that $ad - bc = \pm 1$ implies that the fractions $a/b, a/c, c/d, b/d$ are all irreducible. Check this.

37.1 Lattices and tessellations

Definition 231: Lattices

Given complex numbers α, β with $\alpha/\beta \notin \mathbb{R}$ (periods) the associated *lattice* is

$$\Lambda = \{ m\alpha + n\beta : m, n \in \mathbb{Z} \}. \quad (37.45)$$

Such a Λ is a free $\mathbb{Z}$ -module¹⁸ of rank¹⁹ 2 and tessellates the plane by parallelograms

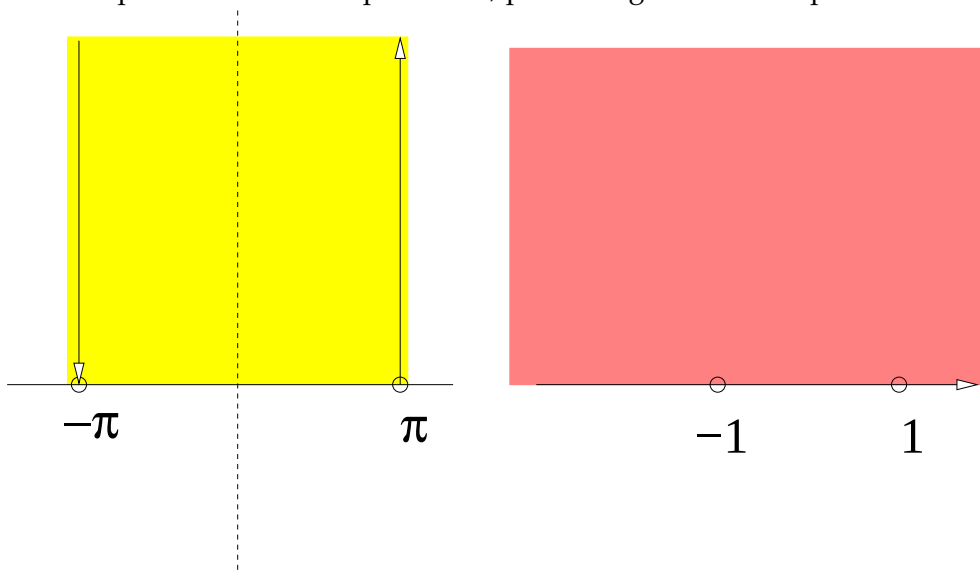
Definition 232

A **tessellation** (tiling) of a region is a family of disjoint open sets whose closures cover the region.

37.2 Example: mapping rectangles

A degenerate rectangle: The sine function revisited. To warm up, consider the map from a vertical strip to the upper half-plane $\mathbb{H}^u$.

Problem. Map the indicated strip into $\mathbb{H}^u$, preserving the marked points and orientation.



Solution. As the boundary angles at $-\pi$ and π are $\pi/2$, the Schwarz–Christoffel formula $\psi(z) = C \int_0^z \prod_{k=1}^n (s - a_k)^{\alpha_k - 1} ds + C'$, reduces to

$$\psi = C \arcsin z + C',$$

hence $f = \psi^{-1}$ has the form $\sin(cw + c')$. Matching the boundary points leads to $c' = 0$ and $c = \frac{1}{2}$, so

$$f(w) = \sin\left(\frac{w}{2}\right),$$

which maps each vertical strip (a degenerate rectangle) conformally onto $\mathbb{H}^u$.

¹⁸In abstract algebra, a module is a generalization of a vector space where the scalars come from a ring instead of a field.

¹⁹The rank is the maximum number of elements in a subset of the module that are "linearly independent" over the ring.

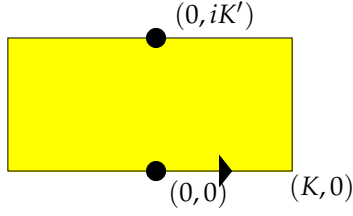


Figure 43: The fundamental rectangle.

R_{-11}	R_{01} iK'	R_{11}
R_{-10}	R_{00}	R_{10}
R_{-1-1}	$-K$ 0 K R_{0-1}	R_{1-1}

Figure 44: Tiling by fundamental rectangles.

Rectangle with sides $2K, iK'$; the elliptic function sn . To obtain a nondegenerate rectangle symmetric about both axes, choose the Schwarz–Christoffel exponents symmetrically, giving

$$\psi(z, k) = \int_0^z \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}}, \quad k \in (0, 1). \quad (37.46)$$

This is the classical incomplete elliptic integral of the first kind.

We take the branch of the square root positive on $(0, 1)$ and continue analytically in the upper half-plane. Since the integrand is even, ψ is odd.

Define the complete integrals

$$K = \int_0^1 \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}} \in \mathbb{R}^+, \quad (37.47)$$

$$iK' = \int_1^{1/k} \frac{ds}{\sqrt{(1-s^2)(1-k^2s^2)}} \in i\mathbb{R}^+. \quad (37.48)$$

Note 233

The analytic continuation from 0^+i to $1^- + 0^+i$ changes the argument of the square root by π ; this accounts for the imaginary increment iK' .

By symmetry, the image $\psi(\mathbb{H}^u)$ is therefore the rectangle with vertices

$$(-K, 0), (K, 0), (K, iK'), (-K, iK').$$

Going from 0^+i to $1^+ + 0^+i$, the argument of the square root changes by 2π , hence the square root becomes negative:

$$\sqrt{(1-s^2)(1-k^2s^2)} = -ks^2 \sqrt{(1-s^{-2})(1-(ks)^{-2})}, \quad s > 1, \quad (37.49)$$

where the square root on the right-hand side is positive for large s and analytic at infinity. For large z , we write $\int_0^z = \int_0^\infty - \int_z^\infty$, expand in powers of s^{-2} and integrate term-by-term. This gives the classical expansion

$$\psi(z, k) = iK' + \frac{1}{kz} + \frac{k^{-1} + k^{-3}}{6z^3} + \cdots. \quad (37.50)$$

37.3 Continuation to $\mathbb{C}$ and double periodicity

Successive Schwarz reflections in the sides of the fundamental rectangle extend ψ to a doubly periodic map; after an even number of reflections the value returns to its original one (see §30.3.1).

The inverse of $k^{-1}\psi$ is the *elliptic sine function* $\text{sn}(z, k)$, a doubly periodic meromorphic function with fundamental periods

$$4K, \quad 2iK'.$$

(One Schwarz reflection across a side of the rectangle flips the rectangle to its mirror image; hence a full period requires two successive reflections. This explains why the periods are *twice* the side lengths of the fundamental rectangle.) Thus

$$\text{sn}(z, k) = \text{sn}(z + 4K, k) = \text{sn}(z + 2iK', k), \quad \text{sn}(iK') = \infty.$$

(The last equality follows from the fact that $\psi(\infty) = iK'$, so iK' corresponds to the point at infinity under the inverse map.) In terms of theta functions (defined later),

$$\text{sn}(u, k) = \frac{\theta_3(\tau) \theta_1(\zeta|\tau)}{\theta_2(\tau) \theta_4(\zeta|\tau)}, \quad k = \lambda(\tau) = \frac{\theta_2^4(\tau)}{\theta_3^4(\tau)}, \quad \zeta = \frac{u}{\theta_3^2(\tau)}.$$

Exercise 234. * Placing one prevertex at ∞ and the others at 0, 1, and $\rho > 0$, the SC formula for a rectangle becomes

$$F(z) = \int_0^z \frac{ds}{\sqrt{s(s-1)(s-\rho)}}.$$

Find changes of variables relating (37.46) to this expression.

37.4 Jacobi elliptic functions sn , cn , dn

We define the **Jacobi elliptic sine** $\text{sn}(u, k)$ as the inverse of $k^{-1}\psi$:

$$u = k^{-1}\psi(z, k) \iff z = \text{sn}(u, k).$$

The elliptic functions cn and dn are defined by

$$\begin{aligned} \text{cn}(u, k) &= \sqrt{1 - \text{sn}^2(u, k)}, \\ \text{dn}(u, k) &= \sqrt{1 - k^2 \text{sn}^2(u, k)}, \end{aligned}$$

with branches chosen so that

$$\text{cn}(0, k) = 1, \quad \text{dn}(0, k) = 1.$$

They satisfy the algebraic identities

$$\text{sn}^2(u, k) + \text{cn}^2(u, k) = 1, \quad k^2 \text{sn}^2(u, k) + \text{dn}^2(u, k) = 1.$$

These are the complete elliptic integrals of the first kind. The corresponding Jacobi functions are *doubly periodic* with periods $4K$ and $2iK'$ (see (37.47)).

37.5 Bases of lattices

Lattices –see (37.45)–are discrete sets. Hence there exists an element $\omega_1 \in \Lambda$ of minimal absolute value. After this choice, pick an $\omega_2 \in \Lambda$ of minimal absolute value that is not in $\mathbb{Z}\omega_1$. We must have $\omega_2/\omega_1 \notin \mathbb{R}$, for otherwise if $n < \omega_2/\omega_1 < n+1$ then $n\omega_1 - \omega_2 \notin \omega_1\mathbb{Z}$ while $|n\omega_1 - \omega_2| < |\omega_1|$, a contradiction.

Proposition 37.2. *$PSL(2, \mathbb{Z})$ acts by symmetries on any lattice Λ ; that is,*

$$PSL(2, \mathbb{Z}) \Lambda = \Lambda.$$

Proof (simplified from [3]). Let

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL(2, \mathbb{Z}), \quad a, b, c, d \in \mathbb{Z}, \quad ad - bc = \pm 1,$$

and let $\Lambda = \Lambda_{\omega_1, \omega_2}$ with $\omega_2/\omega_1 \notin \mathbb{R}$. Define

$$\begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}. \quad (37.51)$$

The pair (ω'_1, ω'_2) plays the same role as (ω_1, ω_2) , since M^{-1} also has integer entries because $ad - bc = \pm 1$ and

$$\begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix} = M^{-1} \begin{pmatrix} \omega'_1 \\ \omega'_2 \end{pmatrix}, \quad (37.52)$$

Equation (37.52) in particular implies $\omega'_2/\omega'_1 \notin \mathbb{R}$: otherwise, ω_1, ω_2 would be real multiples of one another, contradicting the assumption that they form a basis.

Thus every point of the lattice generated by (ω_1, ω_2) is an integer combination of (ω'_1, ω'_2) , and vice versa, yielding the inclusions

$$\Lambda_{\omega_1, \omega_2} \subset \Lambda_{\omega'_1, \omega'_2} \quad \text{and} \quad \Lambda_{\omega'_1, \omega'_2} \subset \Lambda_{\omega_1, \omega_2}.$$

Hence $M\Lambda = \Lambda$. ■

Corollary 235.

- Any two bases of a lattice Λ are related by a modular transformation (37.51).
- $PSL(2, \mathbb{Z})$ are automorphisms of lattices, and hence of the 2d torus $\mathbb{T}$ (in fact, these are all).

37.6 The fundamental region of $PSL(2, \mathbb{Z})$

The group $PSL(2, \mathbb{Z})$ has a *fundamental region* D , described by the conditions 1.–4. in the theorem below. Every lattice has a unique, canonical representative $\tau = \omega_2/\omega_1 \in D$, and exactly 2, 4, or 6 choices of bases realizing this τ .

Theorem 236

There exists a basis (ω_1, ω_2) of a lattice such that $\tau = \omega_2/\omega_1$ satisfies

1. $\text{Im } \tau > 0$,
2. $-\frac{1}{2} < \text{Re } \tau \leq \frac{1}{2}$,

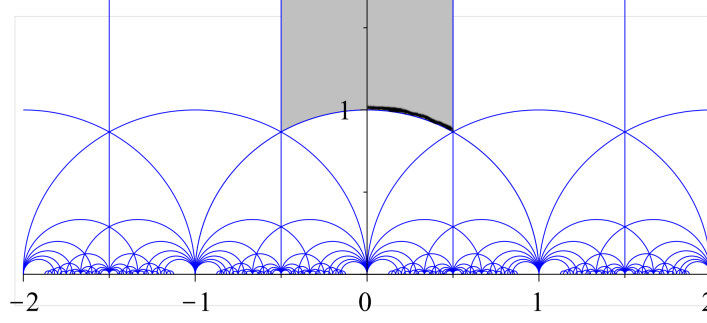


Figure 45: The fundamental region D (the grey region together with the boundary from i to $e^{\pi i/3}$). Figure from [18].

$$3. |\tau| \geq 1,$$

$$4. \operatorname{Re} \tau \geq 0 \text{ whenever } |\tau| = 1.$$

Moreover, τ is uniquely determined by these conditions; depending on the symmetries of τ , the corresponding basis is determined up to 2, 4, or 6 possibilities. The region D is shown in Fig. 45.

Proof (simplified from [3]). Choose ω_1, ω_2 as in §37.5: ω_1 has minimal absolute value among all nonzero lattice points, and ω_2 has minimal absolute value among those not proportional to ω_1 .

Step 1: Basic inequalities for τ . From these minimality properties we obtain

$$|\omega_1| \leq |\omega_2|, \quad |\omega_2| \leq |\omega_1 + \omega_2|, \quad |\omega_2| \leq |\omega_1 - \omega_2|.$$

Writing $\tau = \omega_2/\omega_1 = \alpha + i\beta$, these inequalities give

$$|\tau| \geq 1, \quad |\operatorname{Re} \tau| \leq \frac{1}{2}.$$

Step 2: Ensuring $\operatorname{Im} \tau > 0$. If $\operatorname{Im} \tau < 0$, replace (ω_1, ω_2) by $(-\omega_1, \omega_2)$. This changes the sign of $\operatorname{Im} \tau$ while preserving the preceding inequalities.

Step 3: Ensuring $-\frac{1}{2} < \operatorname{Re} \tau \leq \frac{1}{2}$. If $\operatorname{Re} \tau = -\frac{1}{2}$, replace (ω_1, ω_2) by $(\omega_1, \omega_1 + \omega_2)$. This sends τ to $\tau + 1$ and places its real part in the desired interval.

Step 4: Ensuring 4 on the unit semicircle. If $|\tau| = 1$ and $\operatorname{Re} \tau < 0$, then τ lies on the left semicircle. Apply the modular transformation $\tau \mapsto -1/\tau$, corresponding to

$$(\omega_1, \omega_2) \mapsto (-\omega_2, \omega_1),$$

which maps the left semicircle to the right one while preserving the previous conditions. This gives $\operatorname{Re} \tau \geq 0$.

Thus we have constructed a basis for which $\tau \in D$.

Step 5: Uniqueness. Suppose τ and τ' correspond to two bases and both lie in D . By Cor. 235,

$$\tau' = \frac{a\tau + b}{c\tau + d}, \quad a, b, c, d \in \mathbb{Z}, \quad ad - bc = \pm 1. \quad (37.53)$$

We also have

$$\operatorname{Im} \tau' = \frac{\operatorname{sign}(ad - bc) \operatorname{Im} \tau}{|c\tau + d|^2} = \frac{\operatorname{Im} \tau}{|c\tau + d|^2}, \quad (37.54)$$

so $ad - bc = 1$. Since τ and τ' play symmetric roles, we may assume

$$\operatorname{Im} \tau' \geq \operatorname{Im} \tau, \quad (37.55)$$

which implies

$$|c\tau + d| \leq 1. \quad (37.56)$$

Step 6. Write $\tau = \alpha + i\beta$ with $\alpha \in (-\frac{1}{2}, \frac{1}{2}]$ and $\alpha^2 + \beta^2 \geq 1$. Then (37.56) becomes

$$(\alpha c + d)^2 + \beta^2 c^2 \leq 1.$$

• If $\alpha c + d = 0$, then either $d = -c/2$ with $\alpha = 1/2$, contradicting co-primeness, or $d = 0$ and hence $c = 0$, contradicting $ad - bc = 1$. Thus this case is impossible. Next, if $|c| \geq 2$, then $\beta < 1/2$, implying $\tau' \notin D$. Hence,

$$c \in \{-1, 0, 1\}.$$

Case $c = 0$. Here $ad = 1$, so $a = d = \pm 1$. From (37.53), $\tau' = \tau \pm b$ and hence $\operatorname{Im} \tau = \operatorname{Im} \tau'$. Condition (2) gives $|b| < 1$, therefore $b = 0$, so $\tau = \tau'$.

Case $|c| = 1$. Then $|c\tau + d| \leq 1$ becomes $|d \pm \tau| \leq 1$. Using (2) again, we find $d \in \{-1, 0, 1\}$.

If $d = 0$, condition 3 gives $|\tau| = 1$ and $bc = -1$, hence

$$\tau' = \pm a - \frac{1}{\tau} = \pm a - \bar{\tau}.$$

Because $|\operatorname{Re}(\tau + \tau')| < 1$, we must have $a = 0$, and so $\tau' = -1/\tau$ hence $\tau = i$. Thus $\tau' = \tau$.

If $d = \pm 1$, then (37.56) forces $|\tau - 1| \leq 1$ (the alternative $|\tau + 1| \leq 1$ is impossible, since the closest point to -1 in $\bar{D}$ is $(-\frac{1}{2}, \frac{\sqrt{3}}{2}) \notin D$). Thus $|\tau - 1| \leq 1$ and $|\tau + 1| > 1$, so the denominator in (37.54) must be $|\tau - 1|^2$. The closest point of $\bar{D}$ to 1 is $e^{i\pi/3} = (\frac{1}{2}, \frac{\sqrt{3}}{2}) \in D$, where $|\tau - 1| = 1$ and $\operatorname{Im} \tau$ is minimal. Thus $\operatorname{Im} \tau' = \operatorname{Im} \tau$, and by uniqueness of this boundary point, $\tau' = \tau$.

In all cases, $\tau' = \tau$, completing the proof.

Classification of bases. To determine how many bases correspond to a given $\tau \in D$, we must find the fixed points of the action of $PSL(2, \mathbb{Z})$ on D . We always have at least two bases for the same τ , since

$$(\omega_1, \omega_2) \neq (-\omega_1, -\omega_2) = (-I)(\omega_1, \omega_2), \quad \text{while } -I \sim I \text{ in } PSL(2, \mathbb{Z}).$$

Now, suppose that a fractional linear transformation

$$\tau \mapsto \frac{a\tau + b}{c\tau + d}$$

fixes a point $\tau \in D$. Because the transformed point still lies in D , the entire analysis in the uniqueness proof above applies. The only compatible possibilities are:

(1) $d = 0, c = \pm 1$. This gives, as above, the fixed point $\tau = i$, which is fixed by the involution

$$\tau \mapsto -\frac{1}{\tau}.$$

(2) $|\tau - 1| = 1$. This forces, again as above, the point

$$\tau = e^{\pi i/3},$$

which is fixed simultaneously by the two order-3 elements

$$\tau \mapsto -\frac{\tau+1}{\tau}, \quad \tau \mapsto -\frac{1}{\tau+1}.$$

Thus the only points of D with nontrivial stabilizer are i and $e^{\pi i/3}$. A generic lattice has exactly two corresponding bases, the square lattice with $\tau = i$ has four, and the equilateral triangular lattice with $\tau = e^{\pi i/3}$ has six.

■

Exercise 237. Find all $z \in \text{UHP}$ for which there exists $M \in \text{PSL}(2, \mathbb{Z})$ such that $Mz = z$.

We denote the interior of a set A by A° .

Corollary 238. If $M \in \text{PSL}(2, \mathbb{Z})$, then $(MD)^\circ \cap D^\circ \neq \emptyset$ if and only if $M = 1$. Equivalently, for $A = MD$ and $A' = M'D$, we have

$$A^\circ \cap (A')^\circ \neq \emptyset \iff M = M'.$$

Proof. If $\tau \in D^\circ \cap (MD)^\circ$ and $M \neq 1$, then $\tau = M\tau'$ for some $\tau' \in D^\circ$. Theorem 236 gives uniqueness of the representative in D , so this would force $\tau = M\tau$ with $M \neq 1$, a contradiction.

■

Exercise 239. Does $\text{PSL}(2, \mathbb{Z})$ act properly discontinuously on the upper half-plane?

37.7 The generators of the modular group

Proposition 37.3. The modular group is generated by $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ and $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, or in terms of LFTs, S is given by $z \mapsto -1/z$ and T by $\tau = z \mapsto z + 1$.

Proof. Let $F(z) = \frac{az+b}{cz+d}$. We distinguish three cases.

1. $a = 0$. Then $bc = -1$ giving $b = -c = \pm 1$. In this case, $F(z) = bc/(c^2z + dc) = -1/(z + cd) = -1/(z \pm 1) = ST^{\pm d}(z)$.
2. $|a| = 1$. By projective equivalence, we may assume $a = 1$, and then $d - bc = 1$. We bring this transformation back to case 1 by S, T :

$$T^c S \left(\frac{z+b}{cz+d} \right) = T^c \frac{-cz-d}{z+b} = \frac{-cz-d}{z+b} + c = -\frac{1}{z+b} = ST^b$$

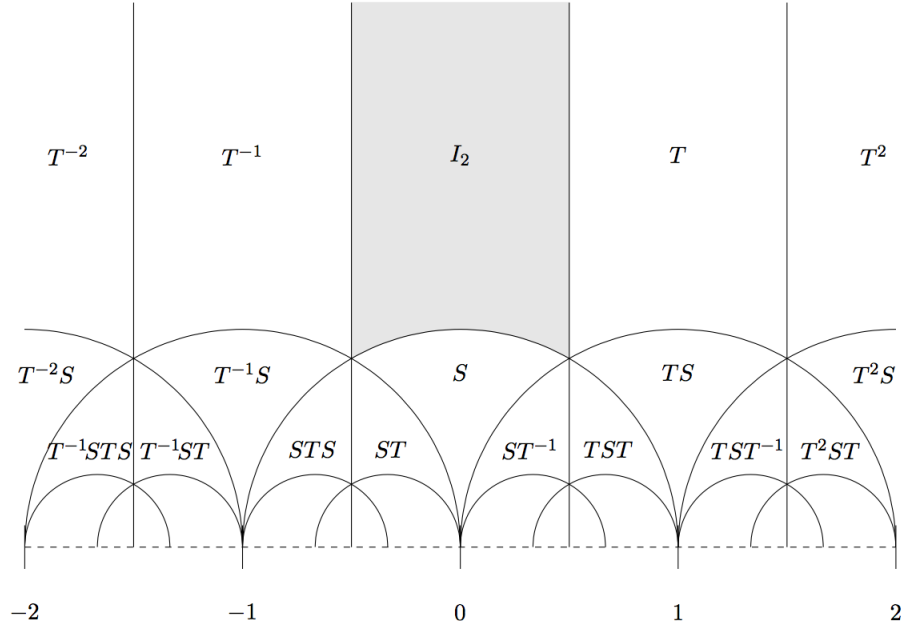


Figure 46: Marked action of S, T

3. $|a| > 1$. By Note 230 we have $|a| \neq |c|$. Applying T shows that we may interchange (a, b) with $(-c, -d)$, and hence we may assume without loss of generality that $|a| > |c|$ and $c \neq 0$ (otherwise SF falls under case 1). Now

$$T^{\pm 1}F(z) = \frac{(a \pm c)z + (b - d)}{cz + d}. \quad (37.57)$$

Since $|a| > |c|$, one of $|a + c|$ or $|a - c|$ is strictly smaller than $|a|$. Thus successive applications of $T^{\pm 1}$ strictly decrease $|a|$, and after finitely many steps we reach case 2.

■

Exercise 240. * 1. Show that the modular group induces a tessellation of the UHP, namely $\{(MD)^\circ : M \in PSL(2, \mathbb{Z})\}$. One way is to apply Corollary 238 and note that for any $z \in D$ and $x \in \mathbb{R}^+$, xz is in the closure of the tessellation by taking $r = p/q \in \mathbb{Q}^+$ and taking LFTs of the form $(npz + b)/(cz + nq)$, $n \in \mathbb{N}, n \rightarrow \infty$. This shows that the strip $\{|\operatorname{Re} z| \leq 1/2, \operatorname{Im} z \geq 0\}$ is in the closure of the tessellation.

2. Find the tessellations of the unit disk induced by conjugation of $PSL(2, \mathbb{Z})$ with the Cayley transform. Find the Poincaré metric on $\mathbb{D}$ using the Cayley transform.

37.8 The hyperbolic plane

On the upper half-plane, define the metric

$$ds^2 = \frac{dx^2 + dy^2}{y^2} = \frac{dz \overline{dz}}{y^2}. \quad (37.58)$$

The corresponding distance between two points $z_1, z_2 \in \mathbb{H}$ is

$$\rho(z_1, z_2) = 2 \tanh^{-1} \left| \frac{z_1 - z_2}{z_1 - \bar{z}_2} \right| = \log \left(\frac{|z_1 - \bar{z}_2| + |z_1 - z_2|}{|z_1 - \bar{z}_2| - |z_1 - z_2|} \right). \quad (37.59)$$

The hyperbolic plane equipped with this metric has constant negative curvature.

Exercise 241. Verify that this metric is invariant under linear fractional transformations, and hence under the group $SL(2, \mathbb{R})$. Show that the geodesics for this metric are precisely the vertical lines and the Euclidean semicircles centered on the real axis (note that such semicircles meet $\mathbb{R}$ orthogonally).

38 Elliptic functions

This material is based on [3] and [15], with some simplifications and additions.

Elliptic functions are doubly periodic meromorphic functions, and they play an important role in analysis, algebra, and number theory. We have already encountered an example of an elliptic function, sn , in §37.3.

38.1 The period module

We have already established (see §37.5) that for a function f to have two distinct periods ω_1 and ω_2 , we must have $\omega_1/\omega_2 \notin \mathbb{R}$. The lattice $\Lambda = \Lambda_{\omega_1, \omega_2}$ is called the *period module* of f .

From now on we restrict attention to *elliptic functions* which are *meromorphic* doubly periodic functions.

38.2 General properties of elliptic functions

In general, we will not assume that Λ contains all the periods of f . The equivalence relation $z_1 = z_2 \pmod{\Lambda}$ means, naturally, that $z_1 - z_2 \in \Lambda$. Since the values of f depend only on congruence classes modulo the lattice, we may regard f as a function defined on these classes.

One convenient way to do this is to restrict f to a fundamental parallelogram

$$P_a = \text{convex hull}\{a, a + \omega_1, a + \omega_2, a + \omega_1 + \omega_2\},$$

for some $a \in \mathbb{C}$. To represent all values of f , part of the boundary of P_a must also be included (for instance, two adjacent sides). The choice of a is immaterial, but it is convenient to take a so that f has no poles on ∂P_a .

When counting zeros or poles of an elliptic function, we always count those lying **inside a fixed fundamental parallelogram**.

Theorem 242

An elliptic function without poles is constant.

Proof. Such a function is entire and, by double periodicity, bounded. Hence it is constant by Liouville's theorem. ■

By definition, a meromorphic function has isolated singularities; thus an elliptic function has only finitely many poles in a fundamental parallelogram.

Theorem 243

The sum of the residues of an elliptic function is zero.

Proof. Choose a fundamental parallelogram P_a so that no poles lie on ∂P_a . Then the sum of the residues is

$$\frac{1}{2\pi i} \int_{\partial P_a} f(s) ds. \quad (38.60)$$

The integral is zero, since the contributions from opposite sides cancel each other. ■

Theorem 244

If f is a non-constant elliptic function, then the numbers of zeros and poles coincide.

Proof. The function $g = f'/f$ is also meromorphic and doubly periodic, with poles of residue -1 for each pole of f and $+1$ for each zero. Now Theorem 244 follows from Theorem 243. ■

Note 245

The three theorems above are due to Liouville.

Corollary 246. Any value of f is assumed equally many times.

Proof. If $c \in \mathbb{C}$, then $f - c$ is an elliptic function. Its poles are exactly those of f , and hence the number of its zeros equals the number of poles of f . ■

Theorem 247

Let $a_1, \dots, a_n$ be the zeros and $b_1, \dots, b_n$ the poles of an elliptic function f in a fundamental parallelogram. Then

$$\sum_{i=1}^n a_i \equiv \sum_{i=1}^n b_i \pmod{\Lambda}.$$

Proof. Take a parallelogram P_a as in Theorem 243. Then

$$\sum_{i=1}^n a_i - \sum_{i=1}^n b_i = \frac{1}{2\pi i} \int_{\partial P_a} \frac{s f'(s)}{f(s)} ds. \quad (38.61)$$

Note that the values of s (and therefore of the integrand) on opposite sides of the parallelogram differ by the corresponding period ω_i of f . Thus,

$$\frac{1}{2\pi i} \left(\int_a^{a+\omega_1} + \int_{a+\omega_1+\omega_2}^{a+\omega_2} \right) \frac{s f'(s)}{f(s)} ds = -\frac{\omega_2}{2\pi i} \int_a^{a+\omega_1} \frac{f'(s)}{f(s)} ds = -\frac{\omega_2}{2\pi i} N_1, \quad (38.62)$$

where $N_1 \in \mathbb{Z}$ is the winding number of 0 with respect to the closed curve $f([a, a + \omega_1])$.

Repeating the same argument with the other pair of sides yields

$$\sum_{i=1}^n a_i - \sum_{i=1}^n b_i = n_1 \omega_1 + n_2 \omega_2, \quad n_1, n_2 \in \mathbb{Z}. \quad (38.63)$$

This is precisely equality modulo Λ . ■

38.3 The Weierstrass elliptic functions

The starting point of the Weierstrass theory is the study of elliptic functions with exactly one double pole per fundamental parallelogram, which we agree to place at the origin. Since multiplying such a function by a constant does not change the theory, we may normalize the singular part to be simply z^{-2} .

The function must be even: indeed, $f(z) - f(-z)$ is elliptic and has no poles, hence is constant; and since $f(\omega_1/2) = f(-\omega_1/2)$, this constant is zero. Since adding a constant to f is also irrelevant, we choose the constant term in the Laurent expansion to be zero.

With all these normalizations, we arrive at the Weierstrass elliptic function

$$\wp(z) = z^{-2} + \sum_{k \geq 1} a_k z^{2k}. \quad (38.64)$$

All of this, however, presupposes the existence of such a function, which we now prove.

Theorem 248

The function

$$\wp(z) = \frac{1}{z^2} + \sum_{0 \neq \omega \in \Lambda_{\omega_1, \omega_2}} \left(\frac{1}{(z - \omega)^2} - \frac{1}{\omega^2} \right) \quad (38.65)$$

is doubly periodic and has a Laurent expansion of the form (38.64) at the origin.

Proof. Until the proof is complete, write $f(z)$ for the sum on the right-hand side of (38.65). First check convergence on compact sets away from Λ . We have

$$\left| \frac{1}{(z - \omega)^2} - \frac{1}{\omega^2} \right| = \frac{2|z|}{|\omega|^3} \frac{|1 - \frac{1}{2}z\omega^{-1}|}{|1 - z\omega^{-1}|^2} \leq \frac{4|z|}{|\omega|^3} \quad (|\omega| \rightarrow \infty). \quad (38.66)$$

By periodicity of Λ , the number of lattice points in the annulus $\{m < |s| \leq m+1\}$ satisfies

$$\#\{s \in \Lambda : m < |s| \leq m+1\} \leq C'(m+1)$$

for some constant $C' > 0$. Hence the right-hand side of (38.66) is bounded by a constant multiple of $\sum_{m=1}^{\infty} m^{-2}$.

Next we check periodicity. Differentiating (38.65) term-by-term (using dominated convergence) yields

$$f'(z) = -2 \sum_{0 \neq \omega \in \Lambda} \frac{1}{(z - \omega)^3}. \quad (38.67)$$

Since f' is doubly periodic, we have

$$f(z + \omega_1) - f(z) = \int_z^{z+\omega_1} f'(s) ds = c_1, \quad f(z + \omega_2) - f(z) = \int_z^{z+\omega_2} f'(s) ds = c_2$$

for some constants c_1, c_2 .

By (38.65), the function f is even. Evaluating at $z_1 = -\omega_1/2$ and $z_2 = -\omega_2/2$ shows that $c_1 = c_2 = 0$. It is then straightforward to check that the Laurent expansion of f at 0 has the form (38.64). ■

38.3.1 The Weierstrass zeta function

Since $\wp$ has zero residues, it possesses a meromorphic antiderivative. We choose the odd antiderivative $-\zeta(z)$, where

$$\zeta(z) = \frac{1}{z} + \sum_{0 \neq \omega \in \Lambda} \left(\frac{1}{z - \omega} + \frac{z}{\omega^2} + \frac{1}{\omega} \right) = \frac{z^2}{\omega^2(\omega - z)}. \quad (38.68)$$

from which convergence is clear. Furthermore, The Maclaurin expansion in z is $|\omega| \rightarrow \infty$ is

$$\frac{1}{z - \omega} + \frac{1}{\omega} + \frac{z}{\omega^2} = -\omega^{-1} \sum_{k \geq 2} (z/\omega)^k, \quad (38.69)$$

Since, by parity, $\sum_{0 \neq \omega \in \Lambda} \omega^{-2k-1} = 0$, summing over Λ gives

$$\zeta(z) = \frac{1}{z} - \sum_{k \geq 2} G_k z^{2k-1}, \quad (38.70)$$

where

$$G_k = \sum_{0 \neq \omega \in \Lambda} \frac{1}{\omega^{2k}}. \quad (38.71)$$

Since $\wp$ is doubly periodic, we have

$$\zeta(z + \omega_i) - \zeta(z) = \eta_i, \quad i = 1, 2,$$

where $\eta_i = \int_z^{z+\omega_i} \wp(s) ds$ are constants.

Because ζ has one simple pole of residue 1 in each fundamental parallelogram,

$$\int_{\partial P_a} \zeta(s) ds = 2\pi i. \quad (38.72)$$

Adding the contributions of opposite sides, using $\zeta(z + \omega_i) - \zeta(z) = \eta_i$, and comparing with (38.72), yields **Legendre's relation**

$$\eta_1 \omega_2 - \eta_2 \omega_1 = 2\pi i. \quad (38.73)$$

38.3.2 The differential equation of $\wp$

Using $-\zeta' = \wp$ and (38.70) (or directly from the definition of $\wp$), we obtain

$$\wp(z) = \frac{1}{z^2} + \sum_{k=2}^{\infty} (2k-1) G_k z^{2k-2}. \quad (38.74)$$

We have

$$\begin{aligned} (\wp')^2 &= \frac{4}{z^6} - \frac{24G_2}{z^2} - 80G_3 + O(z^2) \\ 4\wp^3 &= \frac{4}{z^6} + \frac{36}{z^2} + 60G_3 + O(z^2) \\ G_2\wp &= \frac{G_2}{z^2} + O(z^2) \end{aligned}$$

we have

$$(\wp')^2 - 4\wp^3 + 60G_2\wp + 140G_3 = O(z^2)$$

Here the left side, hence the right side, is doubly periodic with no poles hence a constant. Since the right side vanishes at zero, the left side is zero. Writing $g_2 = 60G_2$ and $g_3 = 140G_3$, we obtain the separable ODE

$$\wp'(z)^2 = 4\wp(z)^3 - g_2\wp(z) - g_3 = 4(\wp(z) - e_1)(\wp(z) - e_2)(\wp(z) - e_3), \quad (38.75)$$

where e_1, e_2, e_3 are the roots of $4z^3 - g_2z - g_3$. Thus

$$z - z_0 = \int_{\wp(z_0)}^{\wp(z)} \frac{ds}{\sqrt{4s^3 - g_2s - g_3}}, \quad (38.76)$$

along any path avoiding the zeros of the denominator.

Proposition 38.4. *The roots e_1, e_2, e_3 are distinct.*

Proof. If all three roots coincided, then $\wp(s) = 4(s + C)^{-2} + 1$, whose inverse is not periodic. If exactly two coincide, the result is an elementary simply periodic function. ■

38.4 The modular function λ

Equation (38.75) implies that the roots of $\wp - e_1, \wp - e_2, \wp - e_3$ of $\wp$ are zeros of $\wp'$. Since $\wp$ is periodic and even, for any period ω we have $\wp(\omega - z) = \wp(z)$, implying $\wp'(\omega - z) = -\wp'(z)$ and therefore $\wp'(\omega/2) = 0$. Taking $\omega = \omega_1, \omega_2, \omega_1 + \omega_2$ gives the three distinct zeros of $\wp'$ in a fundamental parallelogram:

$$e_1 = \wp(\omega_1/2), \quad e_2 = \wp(\omega_2/2), \quad e_3 = \wp((\omega_1 + \omega_2)/2). \quad (38.77)$$

From (38.65) and (38.77), each e_i is homogeneous of degree -2 : replacing ω_1, ω_2 by $\lambda\omega_1, \lambda\omega_2$ multiplies e_i by λ^{-2} . Hence the function

$$\lambda(\tau) = \frac{e_3 - e_2}{e_1 - e_2} \quad (38.78)$$

depends **only on** $\tau = \omega_2/\omega_1$.

Analyticity of λ . From (38.65) we have

$$\omega_1^2 e_1 = 4 + \sum_{\omega \neq 0} \left(\frac{1}{(\frac{1}{2} - n_1 - n_2\tau)^2} - \frac{1}{(n_1 + n_2\tau)^2} \right), \quad (38.79)$$

which is analytic for $\text{Im } \tau > 0$ (and also for $\text{Im } \tau < 0$). The same holds for e_2, e_3 . Since $e_1 \neq e_2$ and $e_2 \neq e_3$, the function λ is analytic in the upper half-plane and never assumes the values 0 or 1.

38.5 The action of the $PSL(2, \mathbb{Z})$ on λ

Recall that $PSL(2, \mathbb{Z})$ is an automorphism of Λ , and since $\wp$ is periodic with any $\omega \in \Lambda$, if we take

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in PSL(2, \mathbb{Z})$$

and change basis to

$$\begin{pmatrix} \omega'_2 \\ \omega'_1 \end{pmatrix} = M \begin{pmatrix} \omega_2 \\ \omega_1 \end{pmatrix} \quad (38.80)$$

the roots of $\wp$ cannot change; they can only be permuted. Recall also that $PSL(2, \mathbb{Z})$ is generated by $T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$, $S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ corresponding to $\tau \mapsto \tau + 1$ and $\tau \mapsto -1/\tau$ resp.

The transformation $\tau \mapsto -1/\tau$ implies $\omega'_1 = \omega_2$, $\omega'_2 = -\omega_1 \equiv \omega_1$ which means e_1 and e_2 get permuted. This means that the new λ is

$$\lambda(-1/\tau) = \frac{e_3 - e_1}{e_2 - e_1} = \frac{e_3 - e_2 + e_2 - e_1}{e_2 - e_1} = 1 - \lambda(\tau)$$

Similarly, $\tau \mapsto \tau + 1$ amounts to changing ω_2/ω_1 to $(\omega_2 + \omega_1)/\omega_1$ or equivalently to $\frac{(\omega_1 + \omega_2)/2}{\omega_1/2}$, fixing e_1 and permuting e_3 and e_2 . Hence,

$$\lambda(\tau + 1) = -\frac{e_3 - e_2}{e_1 - e_3} = \frac{\frac{e_3 - e_2}{e_1 - e_2}}{\frac{e_3 - e_1}{e_1 - e_2}} = \frac{\frac{e_3 - e_2}{e_1 - e_2}}{\frac{e_3 - e_2 + e_2 - e_1}{e_1 - e_2}} = \frac{\lambda}{\lambda - 1}$$

Both transformations are involutions. It follows that $\lambda(\tau + 2) = \lambda(\tau)$. This means that λ is invariant under the transformations with the property that $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \equiv \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \pmod{2}$ which is the subgroup of matrices such that $a \equiv d \equiv 1 \pmod{2}$ and $b \equiv c \equiv 0 \pmod{2}$, a subgroup of $SL(2, \mathbb{Z})$ denoted by $SL(2, \mathbb{Z})/\Gamma(2)$. A function invariant under a subgroup is called a *modular function* for that subgroup.

In turn, this implies that λ is an analytic function of $e^{i\pi\tau}$.

Exercise 249. Using the transformations of λ under $\tau \rightarrow \tau + 1$ and $\tau \mapsto -1/\tau$, check that

$$\lambda\left(\frac{a\tau + b}{c\tau + d}\right) = \begin{cases} \lambda(\tau), & (a, b, c, d) \equiv (1, 0, 0, 1) \pmod{2} \\ 1 - \lambda(\tau), & (a, b, c, d) \equiv (0, 1, 1, 0) \pmod{2} \\ \frac{1}{\lambda(\tau)}, & (a, b, c, d) \equiv (1, 0, 1, 1) \pmod{2} \\ \frac{1}{1 - \lambda(\tau)}, & (a, b, c, d) \equiv (0, 1, 1, 1) \pmod{2} \\ \frac{\lambda(\tau) - 1}{\lambda(\tau)}, & (a, b, c, d) \equiv (1, 1, 1, 0) \pmod{2} \\ \frac{\lambda(\tau)}{\lambda(\tau) - 1}, & (a, b, c, d) \equiv (1, 1, 0, 1) \pmod{2} \end{cases};$$

$$\tau \in \mathbb{H} \text{ and } \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2, \mathbb{Z}) \quad (38.81)$$

We note that the transformations above generate the so-called **anharmonic group**:

$$\left\{ \lambda, 1 - \lambda, \frac{1}{\lambda}, \frac{1}{1 - \lambda}, \frac{\lambda - 1}{\lambda}, \frac{\lambda}{\lambda - 1} \right\}$$

Recall that by the definition of e_1, e_2, e_3 and the double periodicity of $\wp$ any transformation as in (38.81) can only permute the e_i . The table above can also be obtained based on this observation.

38.6 The conformal mapping of λ

We may, without loss of generality, normalize $\omega_1 = 1$ and $\omega_2 = \tau$. In this case,

$$\begin{aligned} e_3 - e_2 &= \sum_{m,n \in \mathbb{Z}} \left[\frac{1}{(m - \frac{1}{2} + (n + \frac{1}{2})\tau)^2} - \frac{1}{(m + (n - \frac{1}{2})\tau)^2} \right] \\ e_1 - e_2 &= \sum_{m,n \in \mathbb{Z}} \left[\frac{1}{(m - \frac{1}{2} + n\tau)^2} - \frac{1}{(m + (n - \frac{1}{2})\tau)^2} \right] \end{aligned} \quad (38.82)$$

As discussed, these functions are analytic in the UHP. Note also that the sums are *even*, and therefore both are real on $i\mathbb{R}^+$.

Proposition 38.5. *We have*

$$\frac{\pi^2}{\sin^2(\pi z)} = \sum_{m \in \mathbb{Z}} \frac{1}{(z - m)^2}.$$

Proof. The functions

$$\frac{\pi^2}{\sin^2(\pi z)} \quad \text{and} \quad \sum_{m \in \mathbb{Z}} \frac{1}{(z - m)^2} \quad (38.83)$$

have the same poles and residues and are both 1-periodic. Hence

$$E(z) = \frac{\pi^2}{\sin^2(\pi z)} - \sum_{m \in \mathbb{Z}} \frac{1}{(z - m)^2} \quad (38.84)$$

is entire.

For real t ,

$$\frac{\pi^2}{\sin^2(\pi it)} = -\frac{\pi^2}{\sinh^2(\pi t)} \rightarrow 0 \quad (t \rightarrow \pm\infty). \quad (38.85)$$

For real t we also have $|m - it|^2 = m^2 + t^2 \geq m^2 + 1$ for large $|t|$, so dominated convergence applies and

$$\sum_{m \in \mathbb{Z}} \frac{1}{(it - m)^2} \rightarrow 0 \quad \text{as } t \rightarrow \pm\infty.$$

The function $E(z)$ is bounded on

$$\{z : \operatorname{Re} z \in [0, 1], \operatorname{Im} z \in [-M, M]\},$$

and by periodicity it is bounded on the entire strip $\{z : \operatorname{Im} z \in [-M, M]\}$. Since $E(z) \rightarrow 0$ as $\operatorname{Im} z \rightarrow \infty$, it is bounded on all of $\mathbb{C}$, hence entire and bounded. Because $E(z) \rightarrow 0$ as $\operatorname{Im} z \rightarrow \infty$, the identity theorem implies $E \equiv 0$. ■

Using the proposition, we obtain

$$\begin{aligned} e_3 - e_2 &= \pi^2 \sum_{n \in \mathbb{Z}} \left(\frac{1}{\cos^2(\pi(n + \frac{1}{2})\tau)} - \frac{1}{\sin^2(\pi(n - \frac{1}{2})\tau)} \right) \\ e_1 - e_2 &= \pi^2 \sum_{n \in \mathbb{Z}} \left(\frac{1}{\cos^2(\pi n\tau)} - \frac{1}{\sin^2(\pi(n - \frac{1}{2})\tau)} \right) \end{aligned} \quad (38.86)$$

and these converge uniformly on $\{\tau : \operatorname{Im} \tau \geq \delta > 0\}$.

Theorem 250

The modular function λ is a conformal map from the domain Ω bounded by $i[0, \infty)$, the circle of radius $\frac{1}{2}$ centered at $\frac{1}{2}$, and the line $1 + i[0, \infty)$ onto the upper half plane.

Proof. We first follow the image of $\partial\Omega$ and then apply the argument principle.

Dominated convergence shows that as $\text{Im } \tau \rightarrow \infty$,

$$e_3 - e_2 \rightarrow 0, \quad e_1 - e_2 \rightarrow \pi^2$$

(the latter because of the $n = 0$ term). Thus $\lambda(\tau) \rightarrow 0$ uniformly in $\text{Re } \tau$. More precisely, from (38.86),

$$e_3 - e_2 = 2\pi^2 \left[\frac{4e^{\pi i \tau}}{(1 + e^{\pi i \tau})^2} + \frac{4e^{\pi i \tau}}{(1 - e^{\pi i \tau})^2} \right] (1 + o(1)) \quad (\text{Im } \tau \rightarrow +\infty), \quad (38.87)$$

hence

$$\lambda(\tau) e^{-\pi i \tau} \rightarrow 16 \quad (\text{Im } \tau \rightarrow +\infty). \quad (38.88)$$

Since $\lambda(-1/\tau) = 1 - \lambda(\tau)$, we obtain $\lambda \rightarrow 1$ as $\tau \rightarrow 0$ along $i\mathbb{R}^+$. By (38.86), λ is real on $i\mathbb{R}^+$. Thus the segment $i[0, \infty]$, traversed from $i\infty$ down to 0, is mapped onto $[0, 1]$, traversed from 0 to 1, possibly retracing itself finitely many times. Indeed, (38.88) shows that λ is monotone near $i\infty$, and $\lambda(-1/\tau) = 1 - \lambda(\tau)$ shows the same near 0. Hence λ' can vanish at only finitely many points on $i\mathbb{R}^+$.

Next, the automorphism $\tau \mapsto \frac{\tau}{\tau+1} \in \text{Aut}(\mathbb{H})$ maps segments and circular arcs to segments and circular arcs. It sends $0 \mapsto 0$, $i \mapsto \frac{1}{2} + \frac{i}{2}$, and $i\infty \mapsto 1$, hence it maps $i[0, \infty)$ to the semicircle in Fig. 47.

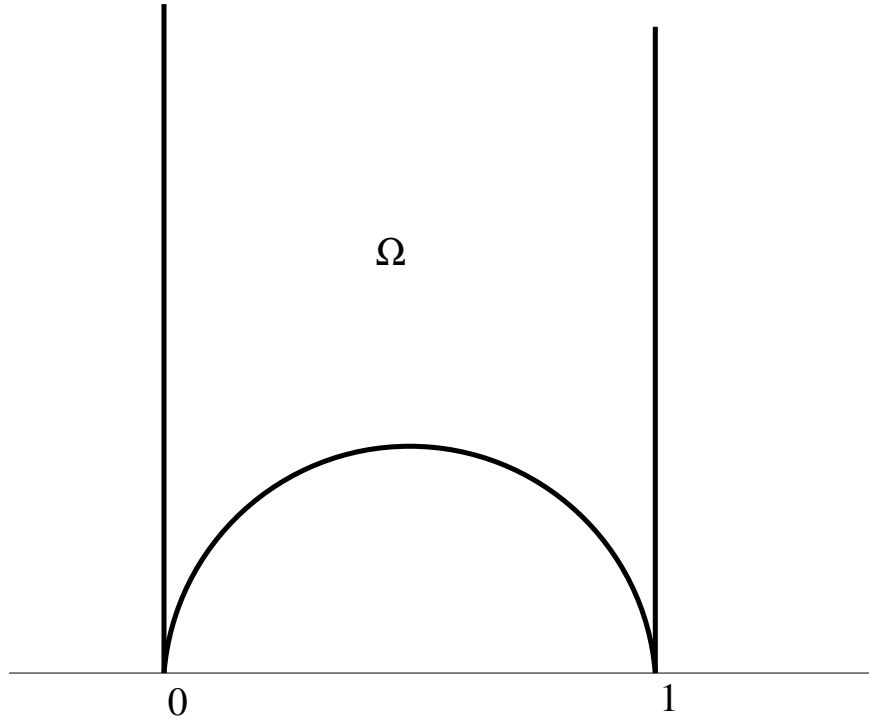


Figure 47: The domain Ω .

We have the transformation rule

$$\lambda\left(\frac{\tau}{\tau+1}\right) = \frac{1}{\lambda(\tau)}.$$

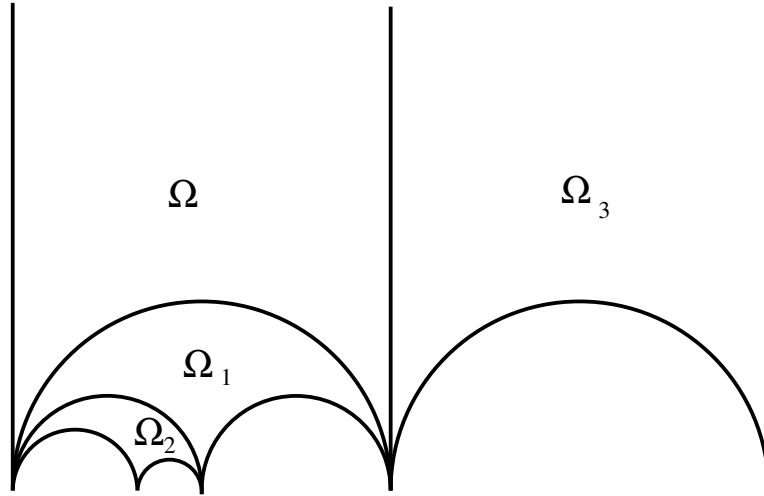


Figure 48: Successive reflections

Since λ is real on $i\mathbb{R}^+$, it is also real on this semicircle. Thus 0 is mapped to 1, and as $\tau \rightarrow i\infty$, $\frac{\tau}{\tau+1} \rightarrow 1$, so 1 is mapped to $1/\lambda(i\infty) = +\infty$.

Finally,

$$\lambda(1+it) = \frac{\lambda(it)}{\lambda(it)-1} = 1 - \frac{1}{\lambda(\frac{i}{t})}. \quad (38.89)$$

Since λ is real on $i\mathbb{R}^+$, it is real on $1+i\mathbb{R}^+$ as well. Thus $1+i0^+$ is mapped to $-\infty$ and $1+i\infty$ is mapped to 0.

Hence $\lambda(\partial\Omega) = \mathbb{R}$, and tracking the boundary shows that it traverses $\mathbb{R}$ exactly once, in the positive direction (with possible retracings, see Note 129). Thus $\lambda(\Omega)$ takes every value in the UHP exactly once and is analytic there; hence λ is a biholomorphism from Ω onto the upper half plane.

Finally, λ' cannot vanish in the UHP: if locally

$$\lambda(s) = \lambda_0 + a(s-s_0)^m,$$

then for $m \geq 2$ the image would enter the lower half plane, where the winding number of $\lambda(\mathbb{R})$ is zero. Thus $\lambda(\partial\Omega)$ is traversed monotonically. ■

38.7 An example: M , the universal cover of $\hat{\mathbb{C}} \setminus \{0, 1, \infty\}$

We use the modular function λ to illustrate the uniformization of this Riemann surface on the upper half plane; composing with a Cayley transform then yields uniformization on the unit disk $\mathbb{D}$. The fundamental group $\pi_1(M)$ is the free group on two generators, a_0, a_1 , corresponding to loops rotating once around 0 and 1, respectively. We may represent M by equivalence classes of curves in $\mathbb{C} \setminus \{0, 1\}$, winding around 0 and 1 according to the word associated with the point (e.g. $a_0 a_1 a_0^{-1} a_1, \dots$).

Exercise 251. * Show that the Schwarz reflection of a line in a circle is again a line or a circle. Show that reflecting Ω successively across the boundary arcs indicated in Fig. 49 covers the UHP.

Each domain obtained by Schwarz reflection in Fig. 49 is mapped by λ alternately and conformally onto the UHP and LHP. Because $1+i\mathbb{R}^+$ is mapped to $(-\infty, 0)$, a reflection of Ω across

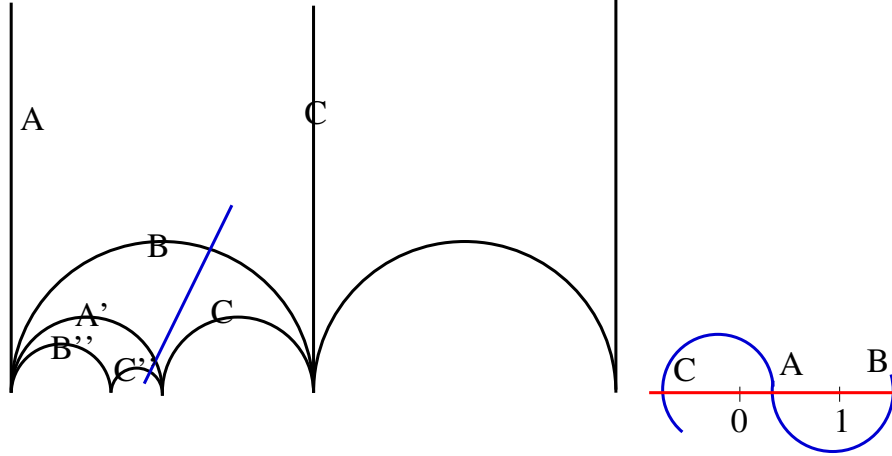


Figure 49: A winding curve in $\hat{\mathbb{C}} \setminus \{0, 1, \infty\}$ becomes an open curve in the UHP.

its right boundary corresponds, under λ , to reflecting the UHP across the interval $(-\infty, 0)$. Likewise, reflecting across the left boundary corresponds to reflecting the UHP across $(0, 1)$, and reflecting across the circular boundary arc corresponds to reflecting across $(1, \infty)$.

All these arcs are analytic, so $\psi = \lambda^{-1}$ admits analytic continuation across each of them. Iterating these reflections, we see that ψ has analytic continuation along any curve in the universal cover of $\hat{\mathbb{C}} \setminus \{0, 1, \infty\}$, and that its singularities occur only at $\{0, 1, \infty\}$. Hence:

Proposition 38.6. *ψ has analytic continuation on M .*

Moreover, any word $abc \dots$ corresponds uniquely to a sequence of reflections across the intervals $(-\infty, 0)$, $(0, 1)$, and $(1, \infty)$. Two distinct words $abcd \dots$ and $abcD \dots$ correspond to reflections across different arcs (here d and D indicate reflections across different intervals), and therefore produce different analytic continuations. Thus two distinct points in M are mapped by ψ to distinct points in the UHP. Using the result of Exercise 251, we obtain:

Proposition 38.7. *ψ is a conformal map from M onto the UHP.*

Exercise 252. Clearly λ is conformal from the UHP onto M . Show that $\mathbb{R}$ is a natural boundary for λ . Use Exercises 249 and 240 to give an alternative proof of this fact.

Padé approximants

39 Padé Approximants

Padé approximants provide a powerful refinement of Taylor expansions. Whereas a Taylor polynomial captures only local behavior of an analytic function, a Padé approximant is a *rational function* that often continues the function analytically far beyond the radius of convergence of its series. Rational functions naturally represent meromorphic behavior, and Padé approximants detect poles automatically.

39.1 Motivation

Let f be analytic at 0 with Taylor expansion

$$f(z) = \sum_{k=0}^{\infty} c_k z^k.$$

The degree- n Taylor polynomial interpolates f and its derivatives at 0:

$$f(z) = T_n(z) + O(z^{n+1}).$$

However, T_n cannot capture singularities: if f has a pole somewhere, T_n remains entire. A Padé approximant uses a rational function and therefore can represent poles directly. Since they are single-valued, rational functions cannot reproduce branch points, but diagonal and near-diagonal Padé approximants mimic branch singularities by placing poles increasingly densely along the boundary of the maximal domain where the function is single-valued, as we shall see.

39.2 Definition of the Padé approximant

Fix integers $m, n \geq 0$. We seek a rational function

$$R_{m,n}(z) = \frac{P_m(z)}{Q_n(z)}$$

where P_m, Q_n are polynomials with

$$\deg P_m \leq m, \quad \deg Q_n \leq n, \quad Q_n(0) = 1,$$

such that it represents the best approximant in a tiny neighborhood of the origin. This means in particular that the Maclaurin series of $R_{m,n}$ agrees with that of f to the highest possible order: $P_m/Q_n - f = O(z^{n+m+1})$ or equivalently

$$f(z)Q_n(z) - P_m(z) = O(z^{m+n+1}). \quad (39.90)$$

$R_{m,n}$ is called the *Padé approximant of type (m, n)* of f at 0.

Proposition 39.8. *If f is a rational function P_m/Q_n , then, for $m' \geq m, n' \geq n$ we have $\text{Padé}[m', n'](f) = f$*

Proof. It is clear that $g = P_m/Q_n$ satisfies the assumptions of being a Padé approximant of f since $\deg(P_m) \leq m', \deg(Q_n) \leq n'$ and $g - f = 0 = O(z^{m'+n'+1})$. ■

Condition (39.90) amounts to solving $m + n + 1$ linear equations for the $m + n$ unknown coefficients of P_m and Q_n (after normalizing $Q_n(0) = 1$).

39.3 First examples

Geometric series. For $f(z) = \frac{1}{1-z}$, the Padé approximant of type $(0, 1)$ is exactly f :

$$R_{0,1}(z) = \frac{1}{1-z}.$$

Already here we see that a Padé approximant can reproduce a singularity that a Taylor polynomial of any degree cannot.

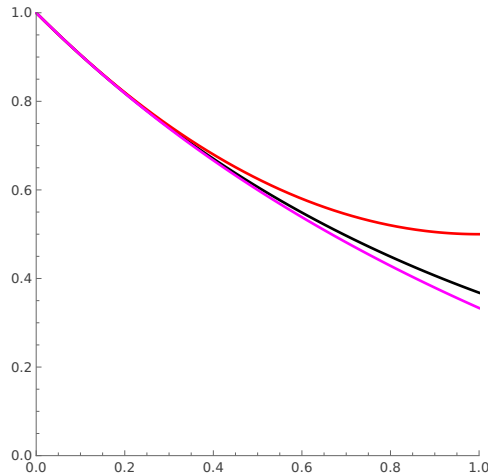


Figure 50: Calculating exp from Padé (1, 1) (magenta), Maclaurin series of order 2 (red) versus the exponential itself (black)

Exponential function. For $f(z) = e^{-z}$, the Padé approximant of type (1, 1) at zero is

$$R_{1,1}(z) = \frac{1 - z/2}{1 + z/2}.$$

It has a pole at $z = -2$, even though e^z has none. The pole does not correspond to an actual singularity of e^z . On $[0, 1]$ the Padé approximant gives a much better approximation than the Taylor polynomial of the same total degree, $1 - z + \frac{z^2}{2}$.

$\tan(z)$ The Maclaurin series of $\tan(z)$ is

$$z + \frac{z^3}{3} + \frac{2z^5}{15} + \frac{17z^7}{315} + \cdots$$

The Padé (3, 4) approximant of $\tan(z)$ at zero is

$$\frac{z - \frac{2z^3}{21}}{\frac{z^4}{105} - \frac{3z^2}{7} + 1}$$

39.4 Padé Approximants and C-Fractions

Padé approximants have a deep connection with continued fractions. If $R_{n,n}$ denotes the diagonal Padé approximants of a function f analytic at 0, then (under mild nondegeneracy assumptions) these approximants arise as the successive *convergents* of a continued fraction expansion of f , known as a *C-fraction* (or *regular continued fraction* in the analytic sense).

C-fractions. A C-fraction for f is an expansion of the form

$$f(z) = \frac{a_0}{1 + \frac{a_1 z}{1 + \frac{a_2 z}{1 + \cdots}}}.$$

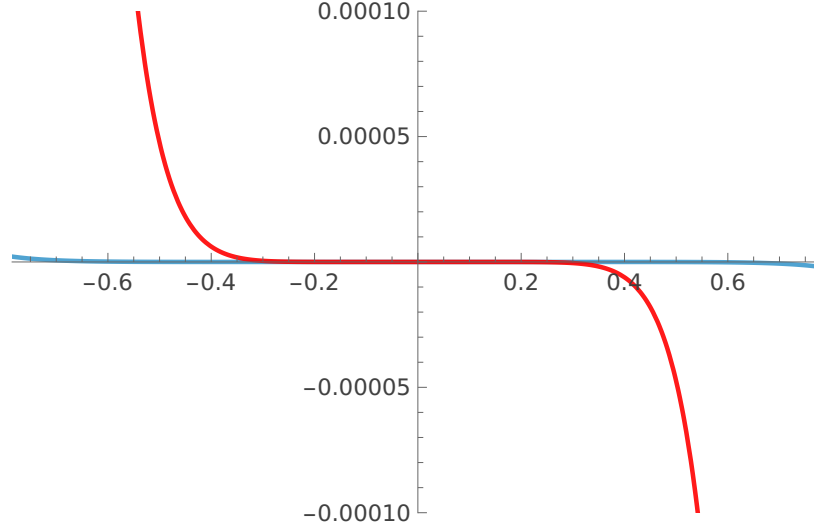


Figure 51: The errors in calculating $\tan(z)$ from Padé (3,4) (blue) and the order 7 Maclaurin series (red)

Truncating this continued fraction after nn stages gives a rational function $F_n(z)$, called the n th *convergent* of the C-fraction and it's a diagonal or near-diagonal Padé (n, n) or (m, n) with $m/n \approx 1$ for large n . This C-fraction can be gotten “by hand” as follows. Let $f(z)$ be analytic at 0, $f(z) = c_0 + c_1z + c_2z^2 + \dots$. Assuming the c_i are nonzero, we write

$$f(z) = c_0 + c_1z(1 + \frac{c_2}{c_1}z + \dots) = c_0 + \frac{c_1z}{\frac{1}{1 + \frac{c_2}{c_1}z + \dots}} = c_0 + \frac{c_1z}{1 - (c_2/c_1z)/\dots}$$

If some $c_l = 0$ we just skip it and factor out the first power with nonzero coefficient. The algorithm is

$$f_0(z) = f(z), \quad f_{k+1}(z) = \frac{1}{f_k(z) - a_k}, \quad a_k = f_k(0).$$

39.5 Analytic continuation

A Taylor series has a disk of convergence limited by the nearest singularity of f . In contrast, $R_{m,n}$ is meromorphic on the entire Riemann sphere. Thus Padé approximants often extend analytic information beyond the original radius of convergence. In many applications, the poles of $R_{m,n}$ approximate the actual singularities of f , providing *analytic continuation* and *singularity detection*.

39.6 Error and convergence

A deep theorem of Stahl shows that if f admits analytic continuation along any path in $\mathbb{C} \setminus S$, where S is a compact set of *zero logarithmic capacity*²⁰ (the “Stahl compact set”), then the diagonal Padé approximants converge to f in capacity on $\mathbb{C} \setminus S$, and this domain is maximal (up to sets of capacity zero).

Outside this class, essentially any behavior is possible. In particular, Lubinsky constructed examples of analytic functions whose analytic continuation has a singular set of *positive* capacity

²⁰Geometric intuition: (a) any smooth or rectifiable curve has positive capacity, hence a set of zero logarithmic capacity is “smaller than a curve”: (b) any set with Hausdorff dimension > 0 has positive capacity, but a countable set, a finite union of points, or even a Cantor-type dust with Hausdorff dimension zero may have zero capacity.

²¹, for which *no subsequence* of the diagonal Padé approximants converges anywhere in the sense of Stahl. Thus Stahl's hypotheses are sharp: if the singular set has zero capacity, diagonal Padé approximants converge in capacity; if it has positive capacity, convergence may fail completely.

39.7 Stahl's Theory and the Role of the Conformal Map

Padé approximants provide rational functions that often extend the analytic behavior of a function far beyond the radius of convergence of its Taylor series. Stahl's fundamental theorem describes the asymptotic behavior of the diagonal Padé approximants $R_{n,n}$ for functions that admit analytic continuation in a sufficiently regular manner.

Setup. Let f be analytic at ∞ and single-valued in $\mathbb{C} \setminus S$, where $S \subset \mathbb{C}$ is a compact set of *zero logarithmic capacity*. There is a unique compact set S_* of *minimal logarithmic capacity* for which f admits single-valued analytic continuation in $\mathbb{C} \setminus S_*$. The set S_* is called the *Stahl compact set*. Its complement $D = \mathbb{C} \setminus S_*$ is a maximal domain of single-valued analytic continuation of f .

Connection with conformal mapping (simply connected case). If the Stahl domain D is simply connected, Stahl's theorem says that, in capacity,

$$\frac{1}{n} \log |f(z) - R_{n,n}(z)| + 2 \log |\Phi(z)| \rightarrow 0 \quad \text{on } D.$$

Where Φ maps D onto the exterior of the closed unit disk. Informally, this means that for “quasi every” $z \in D$ one has

$$|f(z) - R_{n,n}(z)| \approx |\Phi(z)|^{-2n},$$

with the exceptional sets where this fails having capacity tending to zero as $n \rightarrow \infty$.

Informally, this means that for “quasi every” $z \in D$ one has

$$|f(z) - R_{n,n}(z)| \approx |\Phi(z)|^{-2n},$$

with the exceptional sets where this fails having capacity tending to zero as $n \rightarrow \infty$.

Thus Stahl's theorem identifies the unique “best possible” domain on which diagonal Padé approximants can converge, describes where their poles must accumulate, and gives an explicit asymptotic rate of convergence.

The rate of Padé convergence is determined entirely by the conformal map of D onto the exterior disk. Equivalently, the diagonal Padé approximants behave like the truncated Laurent series of the analytic continuation of f in the variable $w = \Phi(z)$.

In particular, points z for which $|\Phi(z)|$ is close to 1 (i.e. points near the boundary S_*) are those where Padé approximants converge the slowest. This parallels classical results for polynomial approximation on compact sets and shows that diagonal Padé approximants detect the shape of the maximal analytic continuation domain.

²¹Let E be a simply connected compact set in the complex plane. By the Riemann mapping theorem, there exists a unique analytic function

$$f(z) = \alpha z + \alpha_0 + \frac{\alpha_1}{z} + \frac{\alpha_2}{z^2} + \cdots, \quad \alpha > 0,$$

which maps the exterior of the unit disk conformally onto the exterior of E and satisfies $f(\infty) = \infty$. The constant α is called the *conformal radius* of E , and α_0 is called the *conformal center* of E .

The function $f(z)$ encodes important geometric information about E . In particular, the constant α is equal to the *logarithmic capacity* of E . For any compact set, its capacity equals that of its outermost boundary.

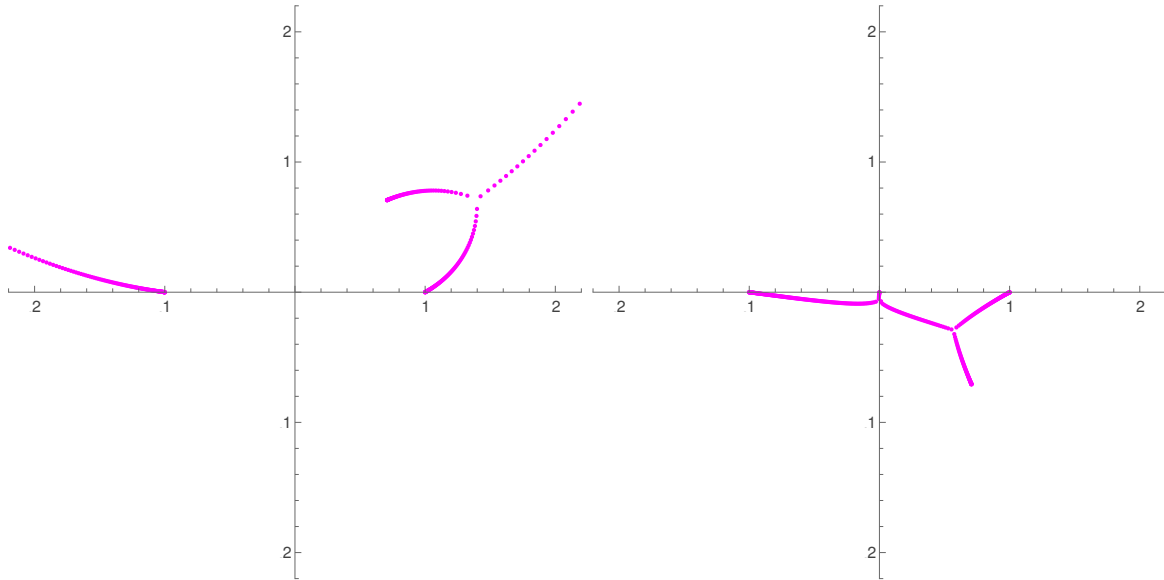


Figure 52: Padé poles for the function $\sqrt{1-z} + \sqrt{i-z} + \sqrt{z+1}$ seen in the z plane (left) and in $1/z$ (right)—from infinity. The picture from infinity is the minimal capacitor. On the right, there are poles accumulating at zero. Indeed, infinity is a singular point for the function.

Poles of Padé. Moreover, the poles of the convergents (hence the approximants) distribute (according to the equilibrium measure) close to the Stahl compact set.

Thus C -fractions, orthogonal polynomials, Padé approximants, and the potential-theoretic geometry of analytic continuation all describe the same underlying structure.

In practice, most (but not always, not even “with probability one”!) often Padé approximants converge pointwise uniformly outside the capacitor. What happens sometimes is that Padé produce spurious poles. These are poles that have no immediate connection with special properties of the associated function, and where evidently the approximation blows up. These poles come in pairs with zeros, and their effect is tiny outside a small neighborhood of their location.

The Hyperbolic Plane and Complex Analysis

- The hyperbolic plane, a model of non-Euclidean geometry with constant negative curvature, plays a pivotal role in complex analysis.
- The uniformization theorem states that every simply connected Riemann surface is conformally equivalent to one of three canonical surfaces: the Riemann sphere, the complex plane, or the unit disk (or equivalently, the upper half-plane). Among these, the unit disk (or upper half-plane) represents the generic case for nontrivial Riemann surfaces.
- The automorphism group of the upper half-plane consists of Möbius transformations. When the upper half-plane is endowed with the metric preserved by this group, it becomes the hyperbolic plane.
- The Möbius transformations preserving the upper half-plane are closely related to the group $SL(2, \mathbb{R})$. Specifically, the group of Möbius transformations is isomorphic to the projective group $PSL(2, \mathbb{R}) = SL(2, \mathbb{R}) / \{\pm I\}$, where $\pm I$ denotes the equivalence of matrices M and $-M$ in $SL(2, \mathbb{R})$, as they induce the same Möbius transformation.

The subgroup $SL(2, \mathbb{Z})$. The group $SL(2, \mathbb{Z})$, consisting of 2×2 integer matrices with determinant 1, is a discrete subgroup with finite covolume (the quotient group has finite Haar measure). Its central role in number theory, geometry, and dynamical systems is due to its connections with:

- **Modular Forms:** The group $SL(2, \mathbb{Z})$ acts on the upper half-plane $\mathbb{H}$ via Möbius transformations. Modular forms are holomorphic functions that exhibit specific invariance properties under this action, playing a central role in number theory and complex analysis.
- **Tessellations and Geometry:** The action of $SL(2, \mathbb{Z})$ tessellates the upper half-plane $\mathbb{H}$, dividing it into fundamental domains. This connects the group's algebraic structure to hyperbolic geometry, providing a geometric interpretation of its action.
- **Arithmetic Groups:** As a prototype of arithmetic groups, $SL(2, \mathbb{Z})$ links diverse areas such as algebraic number theory, Galois representations, and the study of elliptic curves.
- **Dynamical Systems:** The group's action on hyperbolic surfaces illustrates chaotic dynamics and ergodic theory, serving as a fundamental example in the study of dynamical systems.

Generators and Symmetry: The group is generated by two fundamental matrices:

$$S = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Under Möbius transformations, S corresponds to the inversion $\tau \mapsto -1/\tau$, and T represents the translation $\tau \mapsto \tau + 1$.

There is an interesting link to the heat equation

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2}.$$

The equation is invariant under translations, $x \rightarrow x + a$. The general periodic solution of period 2π is

$$u(x, t) = \sum_{n=-\infty}^{\infty} a_n(t) e^{inx},$$

Substituting this into the heat equation, we get

$$a'_n = -n^2 a_n$$

Solving this equation yielding:

$$a_n(t) = C_n e^{-n^2 t}$$

Taking $C_n = 1$ gives the solution with $u(x, 0) = 2\pi\delta(x)$ (check this by calculating the Fourier coefficients of the delta function). With $x = 2\pi z$ and $t = -\pi i\tau$, this is $u(x, t)$ precisely the theta function:

$$\theta(z; \tau) = \sum_{n=-\infty}^{\infty} q^{n^2} e^{2\pi i n z},$$

where $q = e^{\pi i \tau}$, τ is a complex parameter in the upper half-plane, and $z = x$ is a complex variable, and $\tau = t$. Recall now the Poisson summation formula,

$$\sum_{n \in \mathbb{Z}} s(n) = \sum_{n \in \mathbb{Z}} S(k)$$

where

$$S(y) := \mathcal{F}s(y) = \int_{\mathbb{R}} s(x) e^{-2\pi i xy} dx$$

is the Fourier transform of s . On the other hand,

$$\mathcal{F}e^{-n^2 t + 2\pi i n z} = \sqrt{\frac{\pi}{t}} e^{-\frac{(k-z)^2 \pi^2}{t}}$$

For the theta function, this manifests as the modular transformation:

$$\theta\left(\frac{z}{\tau}, -\frac{1}{\tau}\right) = (-i\tau)^{1/2} e^{\pi i \frac{z^2}{\tau}} \theta(z, \tau).$$

and reflects another fundamental symmetry of the heat equation ²², its invariance under the transformation

$$u(x, t) \mapsto t^{-1/2} u\left(-\frac{1}{t}, \frac{x}{t}\right)$$

The usual theta functions $\theta_i(q)$ (where $i = 1, 2, 3, 4$) are specific cases of the general theta function $\theta(z; \tau)$, corresponding to different choices of the argument z and modifications of the series. Here is how they relate:

General Theta Function.

$$\theta(z; \tau) = \sum_{n=-\infty}^{\infty} q^{n^2} e^{2\pi i n z}, \quad q = e^{\pi i \tau}.$$

Usual Theta Functions 1. $\theta_1(z, \tau)$ (Odd Theta Function):

$$\theta_1(z; \tau) = 2 \sum_{n=0}^{\infty} (-1)^n q^{(n+\frac{1}{2})^2} \sin((2n+1)\pi z).$$

This is derived from the general definition by introducing an alternating sign and modifying the summation index.

2. $\theta_2(z, \tau)$ (Semi-Odd Theta Function):

$$\theta_2(z; \tau) = 2 \sum_{n=0}^{\infty} q^{(n+\frac{1}{2})^2} \cos((2n+1)\pi z).$$

This is obtained by using a cosine term instead of a sine term, without the alternating sign.

22

Theorem 1 ([?]). The point symmetry pseudogroup G of the (1+1)-dimensional linear heat equation is constituted by the point transformations of the form

$$\begin{aligned} \tilde{t} &= \frac{\alpha t + \beta}{\gamma t + \delta}, \quad \tilde{x} = \frac{x + \lambda_1 t + \lambda_0}{\gamma t + \delta}, \\ \tilde{u} &= \sigma \sqrt{|\gamma t + \delta|} \exp\left(\frac{\gamma(x + \lambda_1 t + \lambda_0)^2}{4(\gamma t + \delta)} - \frac{\lambda_1}{2} x - \frac{\lambda_1^2}{4} t\right) (u + h(t, x)), \end{aligned} \tag{39.91}$$

where $\alpha, \beta, \gamma, \delta, \lambda_1, \lambda_0$ and σ are arbitrary constants with $\alpha\delta - \beta\gamma = 1$ and $\sigma \neq 0$, and h is an arbitrary solution of the heat equation.

3. $\theta_3(z, \tau)$ (Even Theta Function):

$$\theta_3(z; \tau) = \sum_{n=-\infty}^{\infty} q^{n^2} \cos(2n\pi z).$$

This corresponds to the general theta function for $z \rightarrow 2z$.

4. $\theta_4(z, \tau)$ (Even Alternating Theta Function):

$$\theta_4(z; \tau) = \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2} \cos(2n\pi z).$$

This is similar to $\theta_3(z; \tau)$ but introduces an alternating sign.

Usual Theta Functions at $z = 0$ When $z = 0$, the usual theta functions reduce to simpler q -series:

$$\theta_3(0; \tau) = \sum_{n=-\infty}^{\infty} q^{n^2}, \quad \theta_4(0; \tau) = \sum_{n=-\infty}^{\infty} (-1)^n q^{n^2}, \quad \theta_2(0; \tau) = 2 \sum_{n=0}^{\infty} q^{(n+\frac{1}{2})^2}.$$

The modular transformations of the usual theta functions at $z = 0$ are given as follows:

$$\begin{aligned} \theta_3(0; -1/\tau) &= \sqrt{-i\tau} \theta_3(0; \tau). \\ \theta_4(0; -1/\tau) &= \sqrt{-i\tau} \theta_2(0; \tau). \\ \theta_2(0; -1/\tau) &= \sqrt{-i\tau} \theta_4(0; \tau). \end{aligned}$$

Under the shift $\tau \mapsto \tau + 1$, the transformations are:

$$\begin{aligned} \theta_3(0; \tau + 1) &= \theta_4(0; \tau). \\ \theta_4(0; \tau + 1) &= \theta_3(0; \tau). \\ \theta_2(0; \tau + 1) &= e^{\pi i/4} \theta_2(0; \tau). \end{aligned}$$

The modular transformations of the theta functions in Watson-Ramanujan notation²³, $q = e^{-\alpha}$, $q_1 = e^{-\beta} = e^{-\pi^2/\alpha}$, are as follows:

$$\begin{aligned} \theta_3(q) &= \sqrt{\frac{\pi}{\alpha}} \theta_3(q_1) = \sqrt{\frac{\pi}{\alpha}} \left(1 + 2q_1 + 2q_1^4 + 2q_1^9 + \cdots \right) \\ \theta_4(q) &= \sqrt{\frac{\pi}{\alpha}} \theta_2(q_1) = \sqrt{\frac{\pi}{\alpha}} \left(2q_1^{1/4} + 2q_1^{9/4} + 2q_1^{25/4} + \cdots \right), \end{aligned}$$

We see that these theta functions have convergent transseries as $q \rightarrow 1$ (and similarly at other rational points on the unit circle).

In this notation, the relation between q and q_1 corresponds to the modular transformation $\tau \mapsto -1/\tau$, where $\tau = i\alpha/\pi$ and $\tau_1 = i\beta/\pi$.

Ramanujan considered more general q series in *Eulerian form* such as the third-order mock theta function (mf) $f(q)$, which he defined as: The third-order mock theta function $f(q)$ is defined as:

$$f(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(q; q^2)_{n+1}}, \quad (39.92)$$

²³For historical reasons, notations in the literature have ended up quite inconsistent. In the standard $SL(2, \mathbb{Z})$ theory, we have $q = e^{2\pi i \tau}$.

where

$$(q; q^2)_n = (1 - q)(1 - q^3)(1 - q^5) \cdots (1 - q^{2n-1})$$

is the q -Pochhammer symbol. It can be shown that the unit circle S^1 is a natural boundary for f . As q approaches a rational angle on S^1 , e.g., $(q \rightarrow -1)$, i.e. $\alpha \rightarrow \pi i$ we have

$$f(q) + \sqrt{\frac{\pi}{\alpha}} \exp\left(\frac{\pi^2}{24\alpha} - \frac{\alpha}{t}\right) = 4 + o(1)$$

The leading growth expressed in exponential form terminates: the *exponential part* of the transseries converges. The same holds for other rational points on the unit circle. This mirrors properties seen in classical modular forms and theta functions.

39.7.1 List of mfs [?]

In the following we use the q -Pochhammer symbol:

$$(a; q)_n = \begin{cases} 1 & \text{if } n = 0, \\ (1 - a)(1 - aq) \cdots (1 - aq^{n-1}) & \text{if } n > 0, \end{cases}$$

$$(a; q)_\infty = \prod_{k=0}^{\infty} (1 - aq^k),$$

$$(q)_n = (q; q)_n, \quad (q)_\infty = (q; q)_\infty.$$

Order 3

$$f(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q)_{2n}}, \quad \varphi(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q^2; q^2)_n},$$

$$\psi(q) = \sum_{n=1}^{\infty} \frac{q^{n^2}}{(q; q^2)_n}, \quad \chi(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q)_n (-q^3; q^3)_n}.$$

Order 5

$$f_0(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q)_n}, \quad f_1(q) = \sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(-q; q)_n},$$

$$F_0(q) = \sum_{n=0}^{\infty} \frac{q^{2n^2}}{(q; q^2)_n}, \quad F_1(q) = \sum_{n=0}^{\infty} \frac{q^{2n(n+1)}}{(q; q^2)_{n+1}},$$

$$\varphi_0(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q^2)_n}, \quad \varphi_1(q) = \sum_{n=0}^{\infty} \frac{q^{(n+1)^2}}{(-q; q^2)_n},$$

$$\psi_0(q) = \sum_{n=0}^{\infty} q^{\frac{1}{2}(n+1)(n+2)} (-q; q)_n, \quad \psi_1(q) = \sum_{n=0}^{\infty} q^{\frac{1}{2}n(n+1)} (-q; q)_n,$$

$$\chi_0(q) = \sum_{n=0}^{\infty} \frac{q^n}{(q^{n+1}; q)_n}, \quad \chi_1(q) = \sum_{n=0}^{\infty} \frac{q^n}{(q^{n+1}; q)_{n+1}}.$$

Order 7

$$F_0(q) = \sum_{n=0}^{\infty} \frac{q^{n^2}}{(q^{n+1}; q)_n}, \quad F_1(q) = \sum_{n=0}^{\infty} \frac{q^{(n+1)^2}}{(q^{n+1}; q)_{n+1}}, \quad F_2(q) = \sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(q^{n+1}; q)_{n+1}}.$$

Additional mfs (discovered later)

$$\omega(q) = \sum_{n=0}^{\infty} \frac{q^{2n(n+1)}}{(q; q^2)_{2n+1}}, \quad v(q) = \sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(-q; q^2)_{n+1}}, \quad \rho(q) = \sum_{n=0}^{\infty} \frac{q^{2n(n+1)}(q; q^2)_{n+1}}{(q^3; q^6)_{n+1}}.$$

$$\xi(q) = 1 + 2 \sum_{n=1}^{\infty} \frac{q^{6n(n-1)}}{(q; q^6)_n (q^5; q^6)_n}, \quad \sigma(q) = \sum_{n=1}^{\infty} \frac{q^{3n(n-1)}}{(-q; q^3)_n (-q^2; q^3)_n}.$$

39.7.2 Transformation theory

Watson [?] discovered *the transformation theory of mfs, which he completed for mf3*. Let

$$W(\alpha) = \int_0^{\infty} e^{-\frac{3}{2}\alpha x^2} \frac{\cosh \frac{5}{2}\alpha x + \cosh \frac{1}{2}\alpha x}{\cosh 3x} dx \quad (39.93)$$

$$W_1(\alpha) = \int_0^{\infty} e^{-\frac{3}{2}\alpha x^2} \frac{\sinh \alpha x}{\sinh \frac{3}{2}\alpha x} dx \quad (39.94)$$

$$W_2(\alpha) = \int_0^{\infty} e^{-\frac{3}{2}\alpha x^2} \frac{\cosh \alpha x}{\cosh 3\alpha x} dx \quad (39.95)$$

$$W_3(\alpha) = \int_0^{\infty} e^{-3\alpha x^2} \frac{\sinh \alpha x}{\sinh 3\alpha x} dx \quad (39.96)$$

They satisfy the inversion formulas

$$W(\beta) = \sqrt{\frac{\alpha^3}{\pi^3}} W(\alpha), \quad W_1(\beta) = \sqrt{\frac{2\alpha^3}{\pi^3}} W_2(\alpha),$$

$$W_2(\beta) = \sqrt{\frac{\alpha^3}{2\pi^3}} W_1(\alpha), \quad W_3(\beta) = \sqrt{\frac{\alpha^3}{\pi^3}} W_3(\alpha).$$

We then have

$$\begin{aligned} q^{-\frac{1}{24}} f(q) &= \sqrt{\frac{8\pi}{\alpha}} q_1^{\frac{4}{3}} \omega(q_1^2) + \sqrt{\frac{24\alpha}{\pi}} W_1(\alpha) \\ q^{-\frac{1}{24}} f(-q) &= -\sqrt{\frac{\pi}{\alpha}} q_1^{-\frac{1}{24}} f(-q_1) + \sqrt{\frac{24\alpha}{\pi}} W(\alpha) \\ q^{\frac{2}{3}} \omega(q) &= \sqrt{\frac{\pi}{4\alpha}} q_1^{-\frac{1}{12}} \omega(q_1^2) - \sqrt{\frac{3\alpha}{\pi}} W_2\left(\frac{\alpha}{2}\right) \\ q^{\frac{2}{3}} \omega(-q) &= -\sqrt{\frac{\pi}{\alpha}} q_1^{\frac{2}{3}} \omega(-q_1) + \sqrt{\frac{12\alpha}{\pi}} W_3(\alpha) \end{aligned} \quad (39.97)$$

Remark 253. 1. The transformations $\tau \rightarrow \tau + 1$ and $\tau \rightarrow -\frac{1}{\tau}$ correspond to $q \rightarrow -q$ and $q \rightarrow q_1$, respectively.

2. Caveat: typically, in the literature, the phase is **not** included when changing q to $-q$; it is left to the reader.
3. These two transformations are sufficient to determine the modular transformations under the full $\text{SL}(2, \mathbb{Z})$.
4. Unlike theta functions, the modular transformations of mock theta functions involve additional terms, the W -functions, which are called non-holomorphic parts or *shadows*.
5. The shadows are, in fact, Borel-summed divergent series in disguise. For instance, after the changes of variables $\alpha x = q$ and $3q^2 = p$, we obtain

$$W_3 = \frac{\sqrt{3}}{6\alpha} \int_0^\infty \frac{e^{-\frac{p}{\alpha}} \sinh\left(\frac{\sqrt{3}}{3}\sqrt{p}\right)}{\sqrt{p} \sinh\left(\sqrt{3}\sqrt{p}\right)} dp, \quad (39.98)$$

with analogous forms for the other W -integrals. The exponential part of the transseries in α converges.

6. The shadows of all listed mock modular forms are Laplace transforms of functions of the form $p^{-1/2}R(p)$ where $R(p)$ is a rational function in $e^{i\sqrt{r}p}$, with r being rational numbers.
7. The unit circle $|q| = 1$ is a natural boundary for both f and q (and, in fact, for any mock modular form of any order); it corresponds to $\alpha \in i\mathbb{R}$.
8. *However*, all shadows are *analytic* across the imaginary axis $i\mathbb{R}$.
9. The Stokes lines for the shadows are $\mathbb{R}^-$. The integrals become singular when the denominator vanishes, which corresponds to $\sqrt{p} \in i\mathbb{R}$.
10. All singularities of the shadows are simple poles, spaced quadratically.
11. **Fact 1.** Rotating α by $\pi - 0$ and the contour by $-\pi + 0$, the transseries of the rotated integral can be expressed as follows: the Borel-summed part is a Cauchy principal value (PV) integral, due to 10, and equals the *real* part of the integral. The imaginary part consists of a convergent transseries, which is a sum of pure exponentials. **This holds for all mock modular forms, as shown in the examples below.**
12. **Fact 2.** The Borel-summed part matches the structure on the left-hand side of Equations (39.97), with the functions of q replaced by *unary series* in $1/q$, and the right-hand side of Equations (39.97) similarly modified for q_1 . A unary series is an integer power series with coefficients ± 1 or 0.

For example, denoting $\check{q} = 1/q$, we have, for $\alpha \in \mathbb{R}^-$:

$$\check{q}^{\frac{2}{3}}\omega(-\check{q}) = -\sqrt{\frac{\pi}{\alpha}}\check{q}_1^{\frac{2}{3}}\omega(-\check{q}_1) + \sqrt{\frac{12\alpha}{\pi}}W_3(\alpha), \quad (39.99)$$

where

$$\omega(-\check{q}) = -\sum_{k \in \mathbb{Z}} (-\check{q})^{m(3m+2)+1}.$$

The transformation laws for the mf_3s (2.1), (2.4), (2.5) are [W1, 78–79], [GM2, 204]

$$\begin{aligned}
q^{-\frac{1}{24}}f(q) &= \sqrt{\frac{8\pi}{\alpha}}q_1^{\frac{4}{3}}\omega(q_1^2) + \sqrt{\frac{24\alpha}{\pi}}W_1(\alpha), \\
q^{-\frac{1}{24}}f(-q) &= -\sqrt{\frac{\pi}{\alpha}}q_1^{-\frac{1}{24}}f(-q_1) + \sqrt{\frac{24\alpha}{\pi}}W(\alpha), \\
q^{\frac{2}{3}}\omega(q) &= \sqrt{\frac{\pi}{4\alpha}}q_1^{-\frac{1}{12}}f(q_1^2) - \sqrt{\frac{3\alpha}{\pi}}W_2\left(\frac{\alpha}{2}\right), \\
q^{\frac{2}{3}}\omega(-q) &= -\sqrt{\frac{\pi}{\alpha}}q_1^{\frac{2}{3}}\omega(-q_1) + \sqrt{\frac{12\alpha}{\pi}}W_3(\alpha), \\
q^{-\frac{1}{24}}\phi(q) &= \sqrt{\frac{4\pi}{\alpha}}q_1^{-\frac{1}{24}}\psi(q_1) + \sqrt{\frac{6\alpha}{\pi}}W(\alpha), \\
q^{-\frac{1}{24}}\phi(-q) &= \sqrt{\frac{2\pi}{\alpha}}q_1^{\frac{1}{3}}v(-q_1) + \sqrt{\frac{6\alpha}{\pi}}W_1(\alpha), \\
q^{-\frac{1}{24}}\psi(q) &= \sqrt{\frac{\pi}{4\alpha}}q_1^{-\frac{1}{24}}\phi(q_1) - \sqrt{\frac{3\alpha}{2\pi}}W(\alpha), \\
q^{-\frac{1}{24}}\psi(-q) &= \sqrt{\frac{\pi}{2\alpha}}q_1^{\frac{1}{3}}v(q_1) - \sqrt{\frac{3\alpha}{2\pi}}W_1(\alpha), \\
q^{\frac{1}{3}}v(q) &= \sqrt{\frac{2\pi}{\alpha}}q_1^{-\frac{1}{24}}\psi(-q_1) + \sqrt{\frac{6\alpha}{\pi}}W_2(\alpha), \\
q^{\frac{1}{3}}v(-q) &= \sqrt{\frac{\pi}{2\alpha}}q_1^{-\frac{1}{24}}\phi(-q_1) - \sqrt{\frac{6\alpha}{\pi}}W_2(\alpha), \\
q^{-\frac{1}{24}}\chi(q) &= \sqrt{\frac{\pi}{2\alpha}}\xi\left(q_1^{\frac{2}{3}}\right) + \sqrt{\frac{3\alpha}{2\pi}}W_1(\alpha), \\
q^{-\frac{1}{24}}\chi(-q) &= \sqrt{\frac{\pi}{\alpha}}q_1^{\frac{7}{24}}\sigma\left(-q_1^{\frac{1}{3}}\right) + \sqrt{\frac{3\alpha}{2\pi}}W(\alpha),
\end{aligned}$$

Figure 53: From [?]

$$\begin{aligned}
q^{\frac{2}{3}}\rho(q) &= -\sqrt{\frac{\pi}{\alpha}}q_1^{\frac{7}{12}}\sigma\left(q_1^{\frac{2}{3}}\right) + \sqrt{\frac{3\alpha}{4\pi}}W_2\left(\frac{\alpha}{2}\right), \\
q^{\frac{2}{3}}\rho(-q) &= \sqrt{\frac{\pi}{4\alpha}}\xi\left(-q_1^{\frac{1}{3}}\right) - \sqrt{\frac{3\alpha}{\pi}}W_3(\alpha), \\
\xi(q) &= \sqrt{\frac{4\pi}{3\alpha}}q_1^{-\frac{1}{36}}\chi\left(q_1^{\frac{2}{3}}\right) - \sqrt{\frac{9\alpha}{\pi}}W_2\left(\frac{3\alpha}{2}\right), \\
\xi(-q) &= \sqrt{\frac{4\pi}{3\alpha}}q_1^{\frac{2}{9}}\rho\left(-q_1^{\frac{1}{3}}\right) + \sqrt{\frac{36\alpha}{\pi}}W_3(3\alpha), \\
q^{\frac{7}{8}}\sigma(q) &= -\sqrt{\frac{2\pi}{3\alpha}}q_1^{\frac{4}{9}}\rho\left(q_1^{\frac{2}{3}}\right) + \sqrt{\frac{9\alpha}{2\pi}}W_1(3\alpha), \\
q^{\frac{7}{8}}\sigma(-q) &= \sqrt{\frac{\pi}{3\alpha}}q_1^{-\frac{1}{72}}\chi\left(-q_1^{\frac{1}{3}}\right) - \sqrt{\frac{9\alpha}{2\pi}}W(3\alpha),
\end{aligned}$$

where the Mordell integrals W , W_1 , W_2 , W_3 are defined in [W1] by

$$\begin{aligned}
W(\alpha) &= \int_0^\infty e^{-\frac{3}{2}\alpha x^2} \frac{\cosh \frac{5}{2}\alpha x + \cosh \frac{1}{2}\alpha x}{\cosh 3\alpha x} dx, \\
W_1(\alpha) &= \int_0^\infty e^{-\frac{3}{2}\alpha x^2} \frac{\sinh \alpha x}{\sinh \frac{3}{2}\alpha x} dx, \\
W_2(\alpha) &= \int_0^\infty e^{-\frac{3}{2}\alpha x^2} \frac{\cosh \alpha x}{\cosh 3\alpha x} dx, \\
W_3(\alpha) &= \frac{1}{4}W_2\left(\frac{\alpha}{8}\right) - W_1(2\alpha) = \int_0^\infty e^{-3\alpha x^2} \frac{\sinh \alpha x}{\sinh 3\alpha x} dx.
\end{aligned}$$

They satisfy the inversion rules

$$\begin{aligned}
W(\beta) &= \sqrt{\frac{\alpha^3}{\pi^3}}W(\alpha), \quad W_1(\beta) = \sqrt{\frac{2\alpha^3}{\pi^3}}W_2(\alpha), \\
W_2(\beta) &= \sqrt{\frac{\alpha^3}{2\pi^3}}W_1(\alpha), \quad W_3(\beta) = \sqrt{\frac{\alpha^3}{\pi^3}}W_3(\alpha).
\end{aligned}$$

Figure 54: From [?]

The last eight of these transformation laws involve mf_3s in the wide sense, which may account for their omission in [W1]. The transformation laws for the mf_5s (2.2) are [GM2, 207–209]

$$\begin{aligned}
q^{-\frac{1}{60}} f_0(q) &= \sqrt{\frac{2\pi(5-\sqrt{5})}{5\alpha}} q_1^{-\frac{1}{60}} (F_0(q_1^2) - 1) + \sqrt{\frac{2\pi(5+\sqrt{5})}{5\alpha}} q_1^{\frac{71}{60}} F_1(q_1^2) \\
&\quad + \sqrt{\frac{60\alpha}{\pi}} L\left(\frac{1}{5}, 10\alpha\right), \\
q^{\frac{11}{60}} f_1(q) &= \sqrt{\frac{2\pi(5+\sqrt{5})}{5\alpha}} q_1^{-\frac{1}{60}} (F_0(q_1^2) - 1) - \sqrt{\frac{2\pi(5-\sqrt{5})}{5\alpha}} q_1^{\frac{71}{60}} F_1(q_1^2) \\
&\quad + \sqrt{\frac{60\alpha}{\pi}} L\left(\frac{2}{5}, 10\alpha\right), \\
q^{-\frac{1}{60}} f_0(-q) &= -\sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{60}} f_0(-q_1) + \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{\frac{11}{60}} f_1(-q_1) \\
&\quad + \sqrt{\frac{60\alpha}{\pi}} L\left(\frac{1}{5}, 10\alpha\right), \\
q^{\frac{11}{60}} f_1(-q) &= \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{60}} f_0(-q_1) + \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{\frac{11}{60}} f_1(-q_1) \\
&\quad - \sqrt{\frac{60\alpha}{\pi}} L\left(\frac{2}{5}, 10\alpha\right), \\
q^{-\frac{1}{120}} (F_0(q) - 1) &= \sqrt{\frac{\pi(5-\sqrt{5})}{20\alpha}} q_1^{-\frac{1}{30}} f_0(q_1^2) + \sqrt{\frac{\pi(5+\sqrt{5})}{20\alpha}} q_1^{\frac{11}{30}} f_1(q_1^2) \\
&\quad - \sqrt{\frac{15\alpha}{2\pi}} L\left(\frac{1}{5}, 5\alpha\right), \\
q^{\frac{71}{120}} F_1(q) &= \sqrt{\frac{\pi(5+\sqrt{5})}{20\alpha}} q_1^{-\frac{1}{30}} f_0(q_1^2) - \sqrt{\frac{\pi(5-\sqrt{5})}{20\alpha}} q_1^{\frac{11}{30}} f_1(q_1^2) \\
&\quad - \sqrt{\frac{15\alpha}{2\pi}} L\left(\frac{2}{5}, 5\alpha\right),
\end{aligned}$$

Figure 55: From [?]

$$\begin{aligned}
q^{-\frac{1}{120}}(F_0(-q)-1) &= \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{120}}(F_0(-q_1)-1) + \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{\frac{71}{120}} F_1(-q_1) \\
&\quad + \sqrt{\frac{15\alpha}{2\pi}} L_1\left(\frac{4}{5}, 5\alpha\right), \\
q^{\frac{71}{120}} F_1(-q) &= \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{120}}(F_0(-q_1)-1) - \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{\frac{71}{120}} F_1(-q_1) \\
&\quad + \sqrt{\frac{15\alpha}{2\pi}} L_1\left(\frac{2}{5}, 5\alpha\right), \\
q^{-\frac{1}{120}} \phi_0(q) &= \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{120}} \phi_0(q_1) + \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{-\frac{49}{120}} \phi_1(q_1) \\
&\quad - \sqrt{\frac{15\alpha}{2\pi}} L_1\left(\frac{4}{5}, 5\alpha\right), \\
q^{-\frac{49}{120}} \phi_1(q) &= \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{120}} \phi_0(q_1) - \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{-\frac{49}{120}} \phi_1(q_1) \\
&\quad - \sqrt{\frac{15\alpha}{2\pi}} L_1\left(\frac{2}{5}, 5\alpha\right), \\
q^{-\frac{1}{120}} \phi_0(-q) &= \sqrt{\frac{\pi(5-\sqrt{5})}{5\alpha}} q_1^{-\frac{1}{30}} \psi_0(q_1^2) + \sqrt{\frac{\pi(5+\sqrt{5})}{5\alpha}} q_1^{\frac{11}{30}} \psi_1(q_1^2) \\
&\quad + \sqrt{\frac{15\alpha}{2\pi}} L\left(\frac{1}{5}, 5\alpha\right), \\
q^{-\frac{49}{120}} \phi_1(-q) &= -\sqrt{\frac{\pi(5+\sqrt{5})}{5\alpha}} q_1^{-\frac{1}{30}} \psi_0(q_1^2) + \sqrt{\frac{\pi(5-\sqrt{5})}{5\alpha}} q_1^{\frac{11}{30}} \psi_1(q_1^2) \\
&\quad - \sqrt{\frac{15\alpha}{2\pi}} L\left(\frac{2}{5}, 5\alpha\right), \\
q^{-\frac{1}{60}} \psi_0(q) &= \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{60}} \phi_0(-q_1^2) - \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{-\frac{49}{60}} \phi_1(-q_1^2) \\
&\quad - \sqrt{\frac{15\alpha}{\pi}} L\left(\frac{1}{5}, 10\alpha\right), \\
q^{\frac{11}{60}} \psi_1(q) &= \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}} q_1^{-\frac{1}{60}} \phi_0(-q_1^2) + \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}} q_1^{-\frac{49}{60}} \phi_1(-q_1^2) \\
&\quad - \sqrt{\frac{15\alpha}{\pi}} L\left(\frac{2}{5}, 10\alpha\right),
\end{aligned}$$

Figure 56: From [?]

$$\begin{aligned}
q^{-\frac{1}{60}}\psi_0(-q) &= -\sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}}q_1^{-\frac{1}{60}}\psi_0(-q_1) + \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}}q_1^{\frac{11}{60}}\psi_1(-q_1) \\
&\quad - \sqrt{\frac{15\alpha}{\pi}}L\left(\frac{1}{5}, 10\alpha\right), \\
q^{\frac{11}{60}}\psi_1(-q) &= \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}}q_1^{-\frac{1}{60}}\psi_0(-q_1) + \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}}q_1^{\frac{11}{60}}\psi_1(-q_1) \\
&\quad + \sqrt{\frac{15\alpha}{\pi}}L\left(\frac{2}{5}, 10\alpha\right), \\
q^{-\frac{1}{120}}(\chi_0(q)-2) &= -\sqrt{\frac{\pi(5-\sqrt{5})}{5\alpha}}q_1^{-\frac{1}{30}}(\chi_0(q_1^4)-2) - \sqrt{\frac{\pi(5+\sqrt{5})}{5\alpha}}q_1^{\frac{71}{30}}\chi_1(q_1^4) \\
&\quad - \sqrt{\frac{135\alpha}{2\pi}}L\left(\frac{1}{5}, 5\alpha\right), \\
q^{\frac{71}{120}}\chi_1(q) &= -\sqrt{\frac{\pi(5+\sqrt{5})}{5\alpha}}q_1^{-\frac{1}{30}}(\chi_0(q_1^4)-2) + \sqrt{\frac{\pi(5-\sqrt{5})}{5\alpha}}q_1^{\frac{71}{30}}\chi_1(q_1^4) \\
&\quad - \sqrt{\frac{135\alpha}{2\pi}}L\left(\frac{2}{5}, 5\alpha\right), \\
q^{-\frac{1}{120}}(\chi_0(-q)-2) &= \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}}q_1^{-\frac{1}{120}}(\chi_0(-q_1)-2) + \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}}q_1^{\frac{71}{120}}\chi_1(-q_1) \\
&\quad + \sqrt{\frac{135\alpha}{2\pi}}L_1\left(\frac{4}{5}, 5\alpha\right), \\
q^{\frac{71}{120}}\chi_1(-q) &= \sqrt{\frac{\pi(5-\sqrt{5})}{10\alpha}}q_1^{-\frac{1}{120}}(\chi_0(-q_1)-2) - \sqrt{\frac{\pi(5+\sqrt{5})}{10\alpha}}q_1^{\frac{71}{120}}\chi_1(-q_1) \\
&\quad + \sqrt{\frac{135\alpha}{2\pi}}L_1\left(\frac{2}{5}, 5\alpha\right),
\end{aligned}$$

where

$$L(r, \alpha) = \int_0^\infty e^{-\frac{3}{2}\alpha x^2} \frac{\cosh(3r-2)\alpha x + \cosh(3r-1)\alpha x}{\cosh \frac{3}{2}\alpha x} dx$$

and

$$\begin{aligned}
L_1(r, \alpha) &= \int_0^\infty e^{-\frac{3}{2}\alpha x^2} \left\{ \cosh\left(3r - \frac{7}{2}\right)\alpha x + \cosh\left(3r - \frac{5}{2}\right)\alpha x \right. \\
&\quad \left. + \cosh\left(3r - \frac{1}{2}\right)\alpha x - \cosh\left(3r + \frac{1}{2}\right)\alpha x \right\} / \cosh 3\alpha x dx.
\end{aligned}$$

Figure 57: From [?]

The transformation laws for the mf₇s (2.3) are [GM2, 218–219]

$$\begin{aligned}
q^{-\frac{1}{168}}(\mathcal{F}_0(q) - 2) &= \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{\pi}{7}\right) q_1^{-\frac{1}{42}} \mathcal{F}_0(q_1^4) + \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{2\pi}{7}\right) q_1^{-\frac{25}{42}} \mathcal{F}_1(q_1^4) \\
&\quad + \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{3\pi}{7}\right) q_1^{\frac{47}{42}} \mathcal{F}_2(q_1^4) - \sqrt{\frac{42\alpha}{\pi}} L\left(\frac{1}{7}, 7\alpha\right), \\
q^{-\frac{25}{168}} \mathcal{F}_1(q) &= \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{2\pi}{7}\right) q_1^{-\frac{1}{42}} \mathcal{F}_0(q_1^4) - \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{3\pi}{7}\right) q_1^{-\frac{25}{42}} \mathcal{F}_1(q_1^4) \\
&\quad + \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{\pi}{7}\right) q_1^{\frac{47}{42}} \mathcal{F}_2(q_1^4) - \sqrt{\frac{42\alpha}{\pi}} L\left(\frac{2}{7}, 7\alpha\right), \\
q^{\frac{47}{168}} \mathcal{F}_2(q) &= \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{3\pi}{7}\right) q_1^{-\frac{1}{42}} \mathcal{F}_0(q_1^4) + \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{\pi}{7}\right) q_1^{-\frac{25}{42}} \mathcal{F}_1(q_1^4) \\
&\quad - \sqrt{\frac{8\pi}{7\alpha}} \sin\left(\frac{2\pi}{7}\right) q_1^{\frac{47}{42}} \mathcal{F}_2(q_1^4) - \sqrt{\frac{42\alpha}{\pi}} L\left(\frac{3}{7}, 7\alpha\right), \\
q^{-\frac{1}{168}}(\mathcal{F}_0(-q) - 2) &= -\sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{3\pi}{7}\right) q_1^{-\frac{1}{168}} \mathcal{F}_0(-q_1) + \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{\pi}{7}\right) q_1^{-\frac{25}{168}} \mathcal{F}_1(-q_1) \\
&\quad + \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{2\pi}{7}\right) q_1^{\frac{47}{168}} \mathcal{F}_2(-q_1) + \sqrt{\frac{42\alpha}{\pi}} L_1\left(\frac{6}{7}, 7\alpha\right), \\
q^{-\frac{25}{168}} \mathcal{F}_1(-q) &= \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{\pi}{7}\right) q_1^{-\frac{1}{168}} \mathcal{F}_0(-q_1) - \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{2\pi}{7}\right) q_1^{-\frac{25}{168}} \mathcal{F}_1(-q_1) \\
&\quad + \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{3\pi}{7}\right) q_1^{\frac{47}{168}} \mathcal{F}_2(-q_1) - \sqrt{\frac{42\alpha}{\pi}} L_1\left(\frac{2}{7}, 7\alpha\right), \\
q^{\frac{47}{168}} \mathcal{F}_2(-q) &= \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{2\pi}{7}\right) q_1^{-\frac{1}{168}} \mathcal{F}_0(-q_1) + \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{3\pi}{7}\right) q_1^{-\frac{25}{168}} \mathcal{F}_1(-q_1) \\
&\quad + \sqrt{\frac{4\pi}{7\alpha}} \sin\left(\frac{\pi}{7}\right) q_1^{\frac{47}{168}} \mathcal{F}_2(-q_1) - \sqrt{\frac{42\alpha}{\pi}} L_1\left(\frac{4}{7}, 7\alpha\right).
\end{aligned}$$

Figure 58: From [?]

Similarly,

$$\check{q}^{\frac{2}{3}}\omega(\check{q}) - \sqrt{\frac{\pi}{4\alpha}}\check{q}_1^{-\frac{1}{12}}f(\check{q}_1^2) = -\sqrt{\frac{3\alpha}{\pi}}W_2\left(\frac{\alpha}{2}\right), \quad (39.100)$$

where

$$f(\check{q}) = \sum_{k \in \mathbb{Z}} (-\check{q})^{m(3m-1)/2}.$$

Remark 254. *Modular transformations can be viewed as the hyperbolic plane analog of difference equations. In the Euclidean plane, shifts preserve the Euclidean metric, while in the hyperbolic plane, the hyperbolic metric is invariant under the action of the Möbius group. This invariance is specifically governed by the generators S and T of the modular group.*

Remark 255. *We have developed broadly applicable methods for solving systems of modular equations.*

39.8 Continuation Through the Natural Boundary

To achieve unique continuation across natural boundaries, we rely on a **one-to-one, property-preserving correspondence** between the "continuant" and the "continuee." The key properties preserved under this correspondence are the modular transformations (see Remark 254). This approach is analogous to the philosophy underlying Borel and generalized Borel summation.

39.8.1 General Solution of Mock Modular Equations for $|q| < 1$

Proposition 39.9. The general solution of the homogeneous part of the equation for $\omega(-q)$,

$$q^{\frac{2}{3}}\omega(-q) = -\sqrt{\frac{\pi}{\alpha}}q^{\frac{2}{3}}\omega(-q_1)$$

is $w(q)$, satisfying

$$\frac{(q^{2/3}w(q))^6}{\vartheta_3^6(q)} = F(\lambda_q(q)),$$

where F is an entire function, ϑ_3 denotes the theta function, and λ_q is the inverse elliptic nome function.

By linearity, this provides the general solution of (39.99). Here, we have utilized only **one** of the modular transformations of ω , while the other transformation, given by (39.100), involves f (see [?]).

For $|q| > 1$, there is a unique solution to (39.99) and (39.100), expressed in terms of the real and imaginary parts of W_3 and W_2 .

For $|q| < 1$, the following theorem applies:

Theorem 256

There exists a unique function ω satisfying (39.99) for $|q| < 1$, such that $q^{-2/3}\omega(q)$ is analytic in the unit disk $\mathbb{D}$ and $\omega(q)q_1^{1/12}$ remains bounded as $q \rightarrow 1$. Consequently, for $|q| < 1$, there is a unique solution to (39.99) and (39.100), given by the pair (ω, f) , which corresponds to the Ramanujan-Watson functions.

The proof relies on the results established above and the observation that $w(q)$ in Proposition 39.9, if nonzero, exhibits growth that is too rapid.

It follows that the pair (f, ω) has a **unique continuation** across the natural boundary. Specifically:

- For $|q| > 1$, the solution is given by the unary series discussed earlier.
- For $|q| < 1$, the solution corresponds to the Watson-Ramanujan pair.

40 Applications of Mock Theta functions

Mock Theta functions appear naturally in various applications in mathematics and physics.

Mathematical Applications

Mock Theta	Definition	Field	Applications
$\varphi(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q)_n}$	Number Theory	Partition identities, integer partitions, . . .
$\chi_0(q)$	$\sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(-q; q)_n}$	q-Series and Special Functions	Asymptotics of q -series, modular-like behavior
$\psi(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(q; q^2)_{n+1}}$	Combinatorics	Restricted partitions, lattice paths
$f(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q)_n}$	Modular Forms and Mock Modular Forms	Connection to harmonic Maass forms
$g_2(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2+n}}{(1+q)(1+q^2)\cdots(1+q^n)}$	Moonshine Theory, Algebra	Connections to sporadic groups and representation theory

Physics applications

Mock Theta	Definition	Area	Applications
$\varphi(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q)_n}$	Black Hole Physics, String Theory	BPS state counting, entropy of black holes
$\chi_0(q)$	$\sum_{n=0}^{\infty} \frac{q^{n(n+1)}}{(-q; q)_n}$	Statistical Mechanics, Lattice Models	Partition functions in RSOS models
$\psi(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(q; q^2)_{n+1}}$	Topological Quantum Field Theory	Quantum invariants of 3-manifolds
$f(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2}}{(-q; q)_n}$	Logarithmic Conformal Field Theory	Virasoro characters, LCFT partition functions
$g_2(q)$	$\sum_{n=0}^{\infty} \frac{q^{n^2+n}}{(1+q)(1+q^2)\cdots(1+q^n)}$	Moonshine Theory, Quantum Gravity	Partition functions, holographic dualities

40.1 Schwarz reflection of domains about a circle

We have seen in Theorem 109 that we can analytically continue functions from the upper half plane into a symmetric domain in the lower half plane provided they are continuous up to the

real line and real-valued there. Consider a domain $\mathcal{D}$ bounded by a curve and an arccircle C , and let F be analytic in $\mathcal{D}$ and continuous up to C . At any point $z \in C$ we can decompose $f(z) = f_1 + f_2$ where the direction of f_1 is tangent to the circle and f_2 is perpendicular to it. Assume that $f_2 = 0$. Let M be a Möbius transformation that takes C to a real interval J and $\mathcal{D}$ to a domain $\mathcal{D}'$ in the UHP. Then $f \circ M^{-1}$ is analytic in $\mathcal{D}'$, continuous down to J and real valued on J . Then there is a function analytic F in the domain $\mathcal{D}' \cup \mathcal{D}'' \cup J$ where $\mathcal{D}''$ is the reflection of $\mathcal{D}$ across J s.t. $F|_{\mathcal{D}} = f \circ M^{-1}$ and $F \circ M^{-1}$ provides analytic continuation of f to $\mathcal{D} \cup R\mathcal{D} \cup C$ where $R\mathcal{D}$ is the reflection of $\mathcal{D}$ across the arccircle C .

Exercise 257. Find the formula of the reflection of a point across an circle, defined as in the construction above.

Exercise 258 (Ahlfors). Show that the Klein j -invariant

$$j(\tau) = 256 \frac{(1-x)^3}{x^2}$$

where $x = \lambda(1-\lambda)$ is invariant under the modular group.

41 Entire and Meromorphic functions

Analytic and meromorphic functions share with polynomials and rational functions a number of very useful properties, such as decomposition by partial fractions and root-factorization. These notions have to be carefully analyzed though, since questions of convergence arise.

41.1 A historical context

Finding the *exact* value of the sum

$$S := \sum_{n=1}^{\infty} \frac{1}{n^2} \quad (41.1)$$

known as the “Basel problem” had been open for almost a century, in spite of efforts by great mathematicians of the time (the Bernoulli brothers, Leibniz, Goldbach, Stirling, Moivre and others) before Euler solved it in 1735 when he was 28; because the problem stumped so many brilliant minds, this attracted a lot of attention. Euler’s solution, well before there was any systematic theory of complex functions proceeds (roughly) as follows. Take the function

$$f(x) = \frac{\sin x}{x} \quad (41.2)$$

This has a Taylor series which converges for all $x \in \mathbb{C}$. If f were a polynomial with roots at $z = a_i$, then we would be able to write

$$f(x) = A \prod (x - a_i) \quad (41.3)$$

Assuming the same were true for an “infinite order polynomial” and noting that the roots of $\sin(x)/x$ are at $n\pi$, $n \in \mathbb{Z} \setminus \{0\}$ we get

$$\frac{\sin x}{x} = \prod_{n \geq 1} \left(1 - \frac{x}{n\pi}\right) \left(1 + \frac{x}{n\pi}\right) = \prod_{n \geq 1} \left(1 - \frac{x^2}{(n\pi)^2}\right) \quad (41.4)$$

(here the free constant must be 1, comparing the values at 0) Expanding out and collecting the coefficient of x^2 we get $\frac{\sin x}{x} = 1 - cx^2 + O(x^4)$ where

$$c = \frac{1}{\pi^2} + \frac{1}{4\pi^2} + \cdots = \frac{1}{\pi^2} S \quad (41.5)$$

On the other hand from the Maclaurin expansion of $f(x)$ we have $c = \frac{1}{3!}$. Thus

$$S = \frac{\pi^2}{6} \quad (41.6)$$

In fact, Euler went further and calculated $\sum \frac{1}{n^{2k}}$ for $k \geq 1$, in for all even k . This was mostly fine by the standards of the day, though it led to some criticisms that prompted more rigorous proofs later by Euler. Weierstrass was apparently inspired by this solution when he developed the theory of decomposition of entire functions as products. For a rigorous proof of (41.5) see §42.2

41.2 Partial fraction decompositions

First let $R = P_0/Q = P_1 + P/Q$ be a rational function. where P_i and Q are polynomials and $\deg(P) < \deg(Q)$. We aim at a partial fraction decomposition of R ; if $\deg(Q) = 0$ there is nothing further to do. Otherwise let $z_1, \dots, z_n$, $n \geq 1$, be the zeros of Q , where we don't count the multiplicities, and let m_j be the multiplicities of these roots. Let's look at the singular part of the Laurent expansion of P/Q at z_j :

$$\frac{P}{Q} = \sum_{k=1}^{m_j} \frac{c_{jk}}{(z - z_j)^k} + \text{analytic at } z_j \quad (41.7)$$

We claim that

$$\frac{P}{Q} = \sum_{j=1}^n \sum_{k=1}^{m_j} \frac{c_{jk}}{(z - z_j)^k} \quad (41.8)$$

Indeed,

$$E(z) := \frac{P}{Q} - \sum_{j=1}^n \sum_{k=1}^{m_j} \frac{c_{jk}}{(z - z_j)^k} \quad (41.9)$$

is an entire function. By assumption, $P/Q \rightarrow 0$ as $z \rightarrow \infty$ and the rhs of (41.8) also, clearly, goes to zero as $z \rightarrow \infty$. Thus $E(z) \rightarrow 0$ as $z \rightarrow \infty$ and therefore $E \equiv 0$.

Second example. We examine in detail an example that has some of the features of the general case. The function

$$\frac{\pi^2}{\sin^2 \pi z} \quad (41.10)$$

has poles for $z_i = k$, $k \in \mathbb{Z}$, and the singular part of the Laurent series at $z = k$ is, as it can be quickly checked

$$\frac{1}{(z - k)^2} \quad (41.11)$$

We first note that

$$g(z) = \sum_{k \in \mathbb{Z}} \frac{1}{(z - k)^2}; \quad z \in \mathbb{C} \setminus \mathbb{Z}$$

converges uniformly and absolutely on compact sets in $\mathbb{C} \setminus \mathbb{Z}$ to an analytic function: take z_0 s.t. $\text{dist}(z, \mathbb{Z}) > \varepsilon$ and the disk $\overline{\mathbb{D}_{\varepsilon/2}(z_0)}$. By translation symmetry we may assume that $\lfloor x_0 \rfloor = \alpha \in [0, 1)$. We can write

$$g(z) = \sum_{|k| \leq 1} \frac{1}{(z-k)^2} + \sum_{|k| \geq 2} \frac{1}{(z-k)^2}$$

The second sum is bounded in absolute value by $\sum_{|k| \geq 1} \frac{1}{k^2}$, while the finite sum is a manifestly analytic function. Next we note that

$$\frac{\pi^2}{\sin^2 \pi z} - g(z) \quad (41.12)$$

extended by continuity at the integers is entire. This can be immediately checked by noting that the only singularities of $h(z) = \frac{\pi^2}{\sin^2 \pi z}$ are $z = k \in \mathbb{Z}$, and calculating the Laurent series of h at $z = k$.

Proposition 41.1. *We have*

$$\frac{\pi^2}{\sin^2 \pi z} = \sum_{k \in \mathbb{Z}} \frac{1}{(z-k)^2}; \quad z \notin \mathbb{Z} \quad (41.13)$$

Proof. In the notations introduced above the proposition, we saw that, after extending it by continuity at the removable singularities, the function $h - g$ is entire. This function is clearly periodic of period 1, and it is clearly bounded in any compact subset of the strip $\text{Re } z \in [0, 1]$. By dominated convergence $g(z) \rightarrow 0$ as $\text{Im } z \rightarrow \infty$. Now, with $z = x + iy$ we have

$$|\sin(x + iy)^2| = \frac{1}{2} (\cosh(2y) - \cos(2x)) \geq \frac{1}{2} (\cosh(2y) - 1)$$

It follows that $h - z$ is entire and vanishes at infinity, thus it is identically zero.

■

Exercise 259. *By integrating (41.13) show that*

$$\cot(x) = \frac{1}{x} + \sum_{n=1}^{\infty} \frac{2x}{n^2 \pi^2 - x^2}$$

By one more integration, show the product formula Euler used. (See also Lemma 268 below.)

41.3 The Mittag-Leffler theorem

As such, the procedure in Proposition 41.1 does not extend in general. For instance, for the function Γ'/Γ the poles are located at the negative integers and have residue 1, and the partial fraction decomposition would not converge without some modifications. Furthermore, even when it does converge, a function can be determined by a sum over its poles only up to an entire function. For a general theorem, we need to address these issues.

The theorem below shows that for any sequence of one-sided Laurent series centered at the points $\{b_n\}_{n \in \mathbb{N}} \in \mathbb{C}$ with no accumulation point, there is a meromorphic function with exactly that singular behavior and analytic elsewhere. Conversely, any meromorphic function can be decomposed as a sum of its negative powers-part of its Laurent series at the poles and an entire function. More precisely,

Theorem 260: Mittag-Leffler

(i) Let $\{b_n\}_{n \in \mathbb{N}}$ be a sequence of complex numbers with no accumulation point (cf. Note 261) in $\mathbb{C}$ and let $\{P_n\}_{n \in \mathbb{N}}$ be a sequence of polynomials without constant term. Then there are meromorphic functions f in $\mathbb{C}$ whose only poles are at $z = b_n$ and having singular part $P_n((z - b_n)^{-1})$ at $z = b_n$.

(ii) Conversely let f be meromorphic in $\mathbb{C}$ whose poles are $\{b_n\}_{n \in \mathbb{N}}$ and singular part of the Laurent expansion $P_n((z - b_n)^{-1})$ at b_n , $n \in \mathbb{N}$. Then there exists a sequence of polynomials $\{p_n\}_{n \in \mathbb{N}}$ and an entire function g such that

$$f = \sum_{n \in \mathbb{N}} \left[P_n \left(\frac{1}{z - b_n} \right) - p_n(z) \right] + g(z) := S(z) + g(z) \quad (41.14)$$

where the series converges uniformly on compact sets in $\mathbb{C} \setminus \{b_n\}_{n \in \mathbb{N}}$.

Note 261

The condition that $\{b_n\}_{n \in \mathbb{N}}$ has no accumulation point in $\mathbb{C}$ automatically implies that $|b_n| \rightarrow \infty$ as $n \rightarrow \infty$.

Note 262: Idea of the construction

The possible divergence of the infinite series $\sum_{n \in \mathbb{N}} P_n((z - b_n)^{-1})$ is due to the behavior for large n of the terms of the series. We fix z and take n large enough so that $|z/b_n| < 1$. For large n , $\frac{1}{z - b_n}$ roughly behaves like $-1/b_n$. If we subtract out $-1/b_n$, the behavior is roughly $-zb_n^{-2}$. Continuing in this way, and keeping in mind that $|b_n| \rightarrow \infty$ as $n \rightarrow \infty$, we obtain for each n a bound $|z|^{m_n-1} |b_n|^{-m_n}$ ultimately achieving convergence.

Proof. We start by proving (i). We can assume without loss of generality that $b_n \neq 0$ for if say $b_1 = 0$ then we can prove the theorem for $\tilde{f} = f - P_1(1/z)$. Note first that the series at infinity in $1/b_n$ of $P_n(z - b_n)^{-1}$ and its series in z at zero coincide.

Let p_n be a Maclaurin polynomial of P_n such that

$$\left| P_n((z - b_n)^{-1}) - p_n(z) \right| \leq 2^{-n}; \quad \forall z \text{ s.t. } |z| < |b_n/4| \quad (41.15)$$

Now we look at the series

$$f_1 = \sum_{n=1}^{\infty} \left[P_n((z - b_n)^{-1}) - p_n(z) \right] \quad (41.16)$$

and fix an R and analyze the series for $z \in \mathbb{D}_R$. We split the sum (41.15) into two parts:

$$f_1 = \sum_{n: |b_n| \leq 4R} \left[P_n((z - b_n)^{-1}) - p_n(z) \right] + \sum_{n: |b_n| > 4R} \left[P_n((z - b_n)^{-1}) - p_n(z) \right] \quad (41.17)$$

The first sum is finite, while for the second one we have $|z| < |b_n/4|$ and the estimate (41.15) applies. Thus (41.16) is convergent away from the poles uniformly on compact sets.

(ii) Clearly, $f - f_1$ is entire. ■

41.4 Further examples

We write the result in Exercise 259 in the form

$$\pi \cot \pi z = \frac{1}{z} + \sum_{n \in \mathbb{N} \setminus \{0\}} \frac{2z}{z^2 - n^2} \quad (41.18)$$

We can now use this identity to calculate easily some familiar sums. Note that the lhs of (41.18) has the Laurent expansion at $z = 0$

$$\pi \cot \pi z = \frac{1}{z} - \frac{\pi^2 z}{3} - \frac{\pi^4 z^3}{45} - \frac{2\pi^6 z^5}{945} - \dots \quad (41.19)$$

Since the series on the rhs of (41.18) converges uniformly near $z = 0$, by Weierstrass's theorem it converges together with all derivatives. On the other hand we have

$$\frac{2z}{z^2 - n^2} = -2 \left(\frac{z}{n^2} + \frac{z^3}{n^4} + \frac{z^5}{n^6} + \dots \right) \quad (41.20)$$

and we get immediately,

$$\sum_{n \geq 1} \frac{1}{n^2} = \frac{\pi^2}{6}, \quad \sum_{n \geq 1} \frac{1}{n^4} = \frac{\pi^4}{90}, \quad \sum_{n \geq 1} \frac{1}{n^6} = \frac{\pi^6}{945} \dots \quad (41.21)$$

Exercise 263. * The definition of the Bernoulli numbers B_k is

$$\frac{1}{e^z - 1} = \frac{1}{z} - \frac{1}{2} + \sum_{k=1}^{\infty} (-1)^{k-1} \frac{B_k}{(2k)!} z^{2k-1} \quad (41.22)$$

Show that

$$\sum_{n=1}^{\infty} \frac{1}{n^{2k}} = 2^{2k-1} \frac{B_k}{(2k)!} \pi^{2k} \quad (41.23)$$

42 Infinite products

An infinite product is the limit

$$\prod_{n=1}^{\infty} p_n := \lim_{k \rightarrow \infty} \prod_{n=1}^k p_n = \lim_{k \rightarrow \infty} \Pi_k \quad (42.1)$$

We wish to express entire functions as infinite products. Then the following conventions are natural:

1. Only finitely many p_j are nonzero.

Hence there exists an n_0 such that $p_k \neq 0$ for any $k \geq n_0$ and we require:

2.

$$\prod_{k=n_0}^n p_k \rightarrow \ell \neq 0 \quad \text{as } n \rightarrow \infty$$

In issues of convergence we ignore the zero terms, so in the following we assume they simply do not exist. Write now $p_n = P_n / P_{n-1}$, with $P_0 = 1$ we see that the limit of Π_k is the same as the limit of the P_k , and since $P_k \rightarrow \ell \neq 0$ we have $P_{k+1} / P_k \rightarrow 1$ as $k \rightarrow \infty$. Thus $p_n \rightarrow 1$ is a necessary condition of convergence of the infinite product. We should then better write the products as

$$\prod_{n=1}^{\infty} (1 + a_n) \quad (42.2)$$

and then a necessary condition of convergence is $a_n \rightarrow 0$.

Theorem 264

The infinite product (42.2) converges iff

$$\sum_{n=1}^{\infty} \ln(1 + a_n) \quad (42.3)$$

converges. We omit, as before, the terms with $a_n = -1$. The log is defined with a cut along $(-\infty, 0]$, extended by continuity on the upper side of the cut.

Proof. If the sum (42.3) converges, then $\prod_n$ converges, since the exponential of a finite sum is a finite product.

In the opposite direction, a word of caution. We know that in the complex domain, $\ln ab$ is not always $\ln a + \ln b$. The limit of the sum will not, in general, be the log of the infinite product. So the reasoning is not that obvious. We write $(1 + a_n) = \rho_n e^{i\varphi_n}$ and note that $\prod \rho_n$ is convergent, hence the real part of (3.14) converges. Since $(1 + a_{n+1})/(1 + a_n) \rightarrow 1$ we can inductively choose a multiple of $2\pi i$ in $\arg(1 + a_n)$ so that

$$\varphi_{n+1} - \varphi_n \rightarrow 0 \text{ as } n \rightarrow \infty \quad (42.4)$$

With this choice, we see that $\sum_k \varphi_k$ must converge. Otherwise, there are two possibilities. If $S_n = \sum_{k=1}^n \varphi_k$ is unbounded, then, by (42.4), $\text{dist}(\{S_n\}_{n \in \mathbb{N}}, 2\mathbb{Z} + 1) = \text{dist}(\{S_n\}_{n \in \mathbb{N}}, 2\mathbb{Z}) = 0$ entailing infinite oscillation and divergence of the product. If instead $\{S_n\}_{n \in \mathbb{N}}$ is confined to a compact set and has $a \neq b$ as accumulation points, then the conclusion is the same. Hence, the imaginary part of (3.14) converges as well. ■

Absolute convergence is easier to control in terms of series. An infinite product is absolutely convergent, by definition, iff

$$\sum_{n=1}^{\infty} |\ln(1 + a_n)| \quad (42.5)$$

is convergent.

Theorem 265

The sum (42.5) is absolutely convergent iff $\sum a_k$ is absolutely convergent.

Proof. Assume $\sum a_k$ converges absolutely. Then in particular $a_n \rightarrow 0$. Also, if $\sum_{n=1}^{\infty} \ln(1 + a_n)$ converges absolutely then $\ln(1 + a_n) \rightarrow 0$ and $a_n \rightarrow 0$. But (eliminating all the irrelevant zero terms which are zero) we have, as $n \rightarrow \infty$ $\lim_{n \rightarrow \infty} |a_n|^{-1} \ln(1 + |a_n|) = 1$, and the rest follows from the limit ratio theorem. ■

Note 266

Conditional (not absolute) convergence of $\sum a_n$ and of $\prod(1 + a_n)$ are unrelated notions. (Consider, e.g., the product $\prod(1 - (-1)^n n^{-1/2})$. Is the associated series $\sum(-1)^n n^{-1/2}$ convergent? Is the product convergent?)

42.1 Uniform convergence of products

Exercise 267. Assume that $p_n(z)$ are analytic in the domain $\mathcal{D}$. Show that $f(z) = \prod_{n \geq 1} p_n(z)$ converges absolutely and uniformly on every compact set in the domain $\mathcal{D}$ iff

$$f'(z) = \sum_{k=1}^{\infty} \prod_{n=1}^{\infty} \frac{p'_k}{p_k} p_n \quad (42.6)$$

converges absolutely and uniformly on every compact set in the domain $\mathcal{D}$ (if some p_k are zero, the terms are understood in the sense of removable singularities).

42.2 Example: the sin function

Let us prove directly Euler's product formula for the sin function. A good product formula candidate is indeed

$$C\pi \sin \pi z \stackrel{?}{=} z \prod_{n>0} \left(1 - \frac{z^2}{n^2}\right) \quad (42.7)$$

The constant C can only be $1/\pi^2$ if we look at the behavior near $z = 0$. Thus,

$$\frac{\sin \pi z}{\pi} \stackrel{?}{=} z \prod_{n>0} \left(1 - \frac{z^2}{n^2}\right) \quad (42.8)$$

This equality of course needs to be proved, but this is not difficult.

First we note that the product on the rhs of (42.8) is absolutely and uniformly convergent on any compact z set; this can be easily checked. It thus defines an entire function $g(z)$. Motivated by the way we obtained this possible identity, let us look at the expression $f'/f - g'/g$ where $\pi f(z) = \sin \pi z$. We get, using Exercise 267,

$$f'/f - g'/g = \pi \cot \pi z - \frac{1}{z} + \sum_{n \in \mathbb{N} \setminus \{0\}} \frac{2z}{z^2 - n^2} = 0 \quad (42.9)$$

This means that

$$\frac{f'g - fg'}{fg} = 0 \quad (42.10)$$

in $\mathbb{C} \setminus \mathbb{Z}$ or, equivalently,

$$\frac{f'g - fg'}{g^2} = 0 = \left(\frac{f}{g}\right)' \quad (42.11)$$

or $f/g = \text{const}$; we already calculated the constant based on the behavior at zero, it is one. Thus indeed,

$$\frac{\sin \pi z}{\pi z} = \prod_{n>0} \left(1 - \frac{z^2}{n^2}\right) \quad (42.12)$$

proving (41.5). More generally, we have the following:

Lemma 268. Assume that $\{f_n\}$ are analytic and nonzero in a domain $\mathcal{D}$ containing zero, that $f_n(0) = 1$ and that $\sum_{k=1}^n f'_k/f_k$ converges uniformly on compact sets in $\mathcal{D}$. Then $\{\prod_{k=1}^n f_k\}_{n \in \mathbb{N}}$ converges uniformly on compact sets in $\mathcal{D}$, to a function analytic in $\mathcal{D}$.

Proof. Under these assumptions, $h_n(z) = \int_0^z \sum_{k=1}^n f'_k(s)/f_k(s) ds$ also converges uniformly on compact sets in $\mathcal{D}$. Defining $g(z) = e^{-h_n(z)} \prod_{k=1}^n f_k$ and using Exercise 267 we have $g'_n(z) = 0$ and $g(0) = 1$. The rest is immediate. ■

42.3 Canonical products

The simplest possible case is that in which we have a function with no zeros.

Theorem 269

Assume f is entire and $f \neq 0$ in $\mathbb{C}$ Then f is of the form

$$f = e^g \quad (42.13)$$

where g is also entire.

Proof. Since f'/f is entire and $\mathbb{C}$ is simply connected, $h(z) = \int_0^z f'(s)/f(s)ds$ is well defined and also entire. Now we note that $(fe^{-h})' = 0$ in $\mathbb{C}$ and thus $f = \exp(h + C)$ proving the result. Another proof is by using the monodromy theorem and the fact that $\log f$ has no singularities in $\mathbb{C}$. ■

Assume now that f has finitely many zeros, a zero of order $m \geq 0$ at the origin, and the nonzero ones, possibly repeated are $a_1, \dots, a_n$.

Then

$$f = z^m \prod_{k=1}^n \left(1 - \frac{z}{a_n}\right) e^{g(z)}$$

where g is entire.

This is clear, since if we divide f by the prefactor of e^g we get an entire function with no zeros.

We cannot expect, in general, such a simple formula to hold if there are infinitely many zeros. Again we have to take care of convergence problems. This is done in a manner similar to that used in the Mittag-Leffler construction.

Theorem 270: Weierstrass

(i) If $\{a_n\}_{n \in \mathbb{N}}$ is a sequence with no accumulation points, and $\{m_n\}_{n \in \mathbb{N}} \subset \mathbb{N}$ is a set of multiplicities then there exists an entire function with zeros at a_n of multiplicities m_n and no other zeros.

(ii) Assume f is an entire function with zeros at a_n of multiplicities m_n . Then there exist polynomials q_n , and an entire function $g(z)$ such that

$$f(z) = e^{g(z)} z^m \prod_{n \in \mathbb{N}} \left[\left(1 - \frac{z}{a_n}\right)^{m_n} e^{q_n} \right] \quad (42.14)$$

Proof. This is a consequence of Mittag-Leffler. For the proof, we may assume that $f(0) \neq 0$, for otherwise, if f has a zero of order m at 0 we prove the theorem first for $\tilde{f}(z) = f(z)z^{-m}$ understood as usual after removing the singularity at zero.

Take a meromorphic function h with simple poles at a_n of residues m_n as provided by Mittag-Leffler:

$$h(z) = \sum_{n \in \mathbb{N}} \left(\frac{m_n}{z - a_n} - p_n \right) \quad (42.15)$$

We claim that $e^{\int_0^z h(s)ds}$ understood by analytic continuation from a neighborhood of zero has the desired properties. Indeed, in a neighborhood of 0 we write

$$\frac{m_n}{z - a_n} = \frac{f'_n(z)}{f_n(z)} \text{ where } f_n(z) = \left(1 - \frac{z}{a_n}\right) \quad (42.16)$$

we choose any branch of $\log(1 - z/a_n)$ (a multiple of $2\pi i$ will not matter); the natural choice is that given by the power series. We choose a disk of radius r $\mathbb{D}_r$ and choose N_r so that for $n \geq N_r$ $a_n \notin \mathbb{D}_r$. In $\mathbb{D}_r$ we have

$$\int_0^z h(s)ds = \sum_{n \leq N(r)} m_n \log(1 - z/a_n) - \int_0^z p_n(s)ds + \int_{n \geq N(r)} \left(\frac{m_n}{z - a_n} - p_n \right) \quad (42.17)$$

Once more, the integral is understood as being well defined near zero, and analytically continued to $\mathbb{D}_r$. In $\mathbb{D}_r$ we have

$$e^{\int_0^z h(s)ds} = \prod_{n \leq N(r)} (1 - z/a_n)^{m_n} e^{\int_{n \geq N(r)} \left(\frac{m_n}{s - a_n} - p_n(s) \right) ds} \quad (42.18)$$

where the exponential on the right side of (42.18) is analytic. For (ii) we note that if f is entire with the prescribed zeros and multiplicities and H is as in (i), then f/H is entire with no zeros and the result follows from Theorem 269. ■

Note 271

Historically, Weierstrass' factorization theorem came first. Gösta Mittag-Leffler was inspired by Weierstrass and found his decomposition theorem.

Corollary 272. *Any meromorphic function is a ratio of entire functions.*

Proof. Let F be meromorphic with poles at b_n of order m_n . Let G be any entire function with zeros at b_n of order m_n . Then FG has only removable singularities. ■

42.4 Counting zeros of analytic functions. Jensen's formula

The rate of growth of an analytic function is closely related to the density of zeros. A quite effective counting theorem is due to Jensen.

Theorem 273: Jensen

Assume $f \not\equiv 0$ is analytic in the closed disk $\overline{\mathbb{D}_r}$ and $f(z) = cz^m g(z)$ with $m \geq 0$ and $g(0) = 1$. Let a_i be the nonzero roots of f in $\mathbb{D}_r$, repeated according to their multiplicity. Then

$$\ln |c| = -m \ln r - \sum_{i=1}^n \ln \left(\frac{r}{|a_i|} \right) + \frac{1}{2\pi} \int_0^{2\pi} \ln |f(re^{i\theta})| d\theta \quad (42.19)$$

Proof. The proof will follow quite easily from the case when $f \neq 0$ in $\mathbb{D}_r$. In this case we can define a consistent branch of $\ln f$ in $\mathbb{D}_r$ (see also the second proof of Theorem 269), and $\operatorname{Re} \ln f = \ln |f|$ is harmonic in $\mathbb{D}_r$. For $r' < r$, by the mean value theorem for harmonic functions, we have

$$\ln |f(0)| = \frac{1}{2\pi} \int_0^{2\pi} \ln |f(r'e^{i\theta})| d\theta \quad (42.20)$$

Since f is analytic in the closed disk and $\ln |x|$ is in $L^1(\mathbb{R})$, it is easy to see by dominated convergence (check) that (42.20) holds in the limit $r = r'$ too, even if there are zeros on the circle of radius r :

$$\ln |f(0)| = \frac{1}{2\pi} \int_0^{2\pi} \ln |f(re^{i\theta})| d\theta \quad (42.21)$$

Assume now f has zeros, with the convention in the statement of the theorem. We then build a function which has no zeros inside $\mathbb{D}_r$ and has the same absolute value for $|z| = r$. Such a function is

$$h(z) = \frac{r^m}{z^m} f(z) \prod_{i=1}^n \frac{r^2 - \bar{a}_i z}{r(a_i - z)} \quad (42.22)$$

(understood as usual with the removable singularities removed). Then

$$\ln |h(0)| = \frac{1}{2\pi} \int_0^{2\pi} \ln |h(re^{i\theta})| d\theta = \frac{1}{2\pi} \int_0^{2\pi} \ln |f(re^{i\theta})| d\theta \quad (42.23)$$

We have $h(0) = r^m c \prod_{i=1}^n \frac{r}{a_i}$ and the formula follows from the case with no roots. ■

Corollary 274. Assume f is analytic in the closed disk of radius R and $f(0) \neq 0$. Let $v(r)$ denote the number of zeros of f in the disk of radius $r \leq R$. Then

$$\int_0^R \frac{v(x)}{x} dx \leq \ln \max_{|z|=R} |f(z)| - \ln |f(0)| \quad (42.24)$$

(Note that for some $\varepsilon > 0$, $v(x) = 0$ for $x < \varepsilon$.) Clearly, $v(x)$ is an increasing (discontinuous) function of x .

Proof. Note that

$$\ln(R/|a_i|) = \int_{|a_i|}^R \frac{dx}{x} = \int_0^R \chi([|a_i|, x]) \frac{dx}{x}$$

Thus

$$\sum_{i=1}^n \ln \left(\frac{R}{|a_i|} \right) = \int_0^R \sum_{i=1}^n \chi([|a_i|, x]) \frac{dx}{x} = \int_0^R \frac{v(x)}{x} dx$$

The rest follows immediately from (42.19). ■

42.5 Entire functions of finite order

Let f be an entire function. We denote

$$\|f\|_R = \sup_{|z| \leq R} |f(z)| = \sup_{|z|=R} |f(z)|$$

A function is of **order** $\leq \rho$ if for all R large enough we have

$$\forall \varepsilon > 0 \exists c_\varepsilon \text{ s.t. } \|f\|_R \leq \exp(c_\varepsilon R^{\rho+\varepsilon}) \quad (42.25)$$

or equivalently

$$\forall \varepsilon > 0 \ln \|f\|_R = O(R^{\rho+\varepsilon}) \quad (42.26)$$

Note 275

We can always check the condition for $R \in \mathbb{N}$ large enough since $(N+1)^\rho = O(N^\rho)$.

The function f has **strict order** $\leq \rho$ if there is some $c > 0$ such that for all R large enough we have

$$\|f\|_R \leq \exp(cR^\rho) \quad (42.27)$$

A function has **order equal** ρ if

$$\rho = \inf\{\rho' : \forall \varepsilon > 0 \exists c_\varepsilon \text{ s.t. } \|f\|_R \leq \exp(c_\varepsilon R^{\rho'+\varepsilon})\}$$

A function has **strict order equal** ρ if

$$\rho = \inf\{\rho' : \forall R > 0, \|f\|_R \leq \exp(cR^{\rho'})\}$$

Exercise 276. * Check that $\cosh z^{1/2}$ has strict order $1/2$ and $\cos z^{1/4} + \cosh z^{1/4}$ has strict order $1/4$. Variations of this construction lets you find entire functions of any rational order. But can you find an entire function of exact order π ?

Proposition 42.1. Assume $f(z)$ is entire, and for large $|z|$ there are positive constants C, c and ρ such that $|f(z)| \leq Ce^{c|z|^\rho}$. Then, for $R > 0$ there is a $c_2 > 0$ such that

$$\nu(R) \leq c_2 R^\rho \quad (42.28)$$

Indeed, the zeros at zero do not change the shape of the inequality, and we can thus assume $f(0) \neq 0$. Then,

$$c|R|^\rho \geq \int_0^R \frac{\nu(x)}{x} dx \geq \int_{R/2}^R \frac{\nu(x)}{x} dx \geq \frac{\nu(R/2)}{R} \frac{R}{2} \quad (42.29)$$

and the rest is immediate. The constants in the inequality can be optimized by choosing R/τ , $\tau > 1$ instead of $R/2$ and finding the best τ .

Theorem 277

Let f be entire of strict order $\leq \rho$ and let $\{z_n\}$ be its nonzero zeros, repeated according to their multiplicity and ordered increasingly by their absolute value. Then for any $\varepsilon > 0$, $\{|z_n|^{-1}\}_{n \in \mathbb{N}} \in \ell^{\rho+\varepsilon}(\mathbb{N})$, i.e., the series

$$\sum_{n=1}^{\infty} \frac{1}{|z_n|^{\rho+\varepsilon}} \quad (42.30)$$

is convergent.

Proof. We can obviously discard the roots with $|z_i| \leq 1$ which are in finite number. Without loss of generality we assume there are none. We have, with $N \in \mathbb{N}$ and estimating the sum by annuli,

$$\sum_{|z_n| \leq N} \frac{1}{|z|^{\rho+\varepsilon}} \leq \sum_{k=1}^N \frac{\nu(k+1) - \nu(k)}{k^{\rho+\varepsilon}} \quad (42.31)$$

we can now use the method of Abel summation by parts. We write

$$\frac{\nu(k+1) - \nu(k)}{k^{\rho+\varepsilon}} = \nu(k+1) \left(\frac{1}{k^{\rho+\varepsilon}} - \frac{1}{(k+1)^{\rho+\varepsilon}} \right) + \left(\frac{\nu(k+1)}{(k+1)^{\rho+\varepsilon}} - \frac{\nu(k)}{k^{\rho+\varepsilon}} \right) \quad (42.32)$$

and note that by summation, the terms in the last parenthesis cancel out to

$$\frac{\nu(N+1)}{(N+1)^{\rho+\varepsilon}} - \nu(1)$$

Note that by the mean value theorem we have for some $\gamma = \gamma(k)$

$$\frac{\nu(k+1)}{k^{\rho+\varepsilon}} - \frac{\nu(k+1)}{(k+1)^{\rho+\varepsilon}} = \frac{(\rho+\varepsilon)\nu(k+1)}{(k+\gamma)^{\rho+\varepsilon+1}} \leq \frac{Ck^\rho(\rho+\varepsilon)}{k^{\rho+\varepsilon+1}} \quad (42.33)$$

and the sum converges. ■

42.6 Estimating analytic functions by their real part

Theorem 278: Borel-Carathéodory

Let $R > 0$ and assume $f = u + iv$ is analytic $\overline{\mathbb{D}_R}$. Let $A_R = \max_{|z|=R} u(z)$. Then for $r < R$ we have

$$\max_{|z| \leq r} |f(z)| \leq \frac{2rA_R}{R-r} + \frac{R+r}{R-r} |f(0)| \quad (42.34)$$

Note that if f is entire and nonconstant, then $\|f\|_R \rightarrow \infty$ as $R \rightarrow \infty$. Then, since $|u| \leq |f|$, the theorem above shows that

$$\lim_{R \rightarrow \infty} \frac{\max_{|z| \leq R} u(z)}{\max_{|z| \leq R} |f(z)|} \geq \frac{1}{2} \quad (42.35)$$

Proof. Assume first that $f(0) = 0$. Then $u(0) = 0$ and by the mean value theorem $A_R \geq 0$. If $A_R = 0$ then by the same argument $u \equiv 0$ on $\partial\mathbb{D}_R$ and by Poisson's formula $u \equiv 0$ in $\mathbb{D}_R$. Then $v \equiv \text{const} = 0$ since $f(0) = 0$, thus $f \equiv 0$ and the formula holds trivially.

We now assume $A_R > 0$. Since the maximum of a harmonic function is reached on the boundary, we have $2A_R - u \geq u$ in $\mathbb{D}_R$ and the inequality is strict in the interior. Also note that if at some point $u < 0$, then again $2A_R - u = 2A_R + |u| \geq |u|$. In $\mathbb{D}_R$ we have

$$|2A_R - f| = \sqrt{(2A_R - u)^2 + v^2} > \sqrt{u^2 + v^2} = |f| \quad (42.36)$$

Hence the function

$$g(z) = \frac{1}{2A_R - f(z)} \frac{f(z)}{z} \quad (42.37)$$

is holomorphic in $\mathbb{D}_R$ and we have, by passing to the limit $|z| \rightarrow R$,

$$|g(z)| = \left| \frac{1}{2A_R - f(z)} \frac{f(z)}{z} \right| \leq \frac{1}{R} \quad (42.38)$$

hence

$$\left| \frac{f(z)}{z} \right| \leq \frac{1}{R} |2A_R - f(z)| \leq \frac{1}{R} (2A_R + |f(z)|) \quad (42.39)$$

Solving for $|f(z)|$ we get

$$|f(z)| \leq \frac{2|z|A_R}{R - |z|} \quad (42.40)$$

as claimed. The general case is obtained by applying this inequality to $f(z) - f(0)$ (exercise). ■

Corollary 279. Assume $\rho \geq 0$, $f = u + iv$ is entire and as $|z| \rightarrow \infty$ we have

$$|u(z)| \leq C|z|^\rho \quad (42.41)$$

Then f is a polynomial of degree at most ρ .

Proof. From (42.35), for any $\varepsilon > 0$ we have, for large R $|f(z)| \leq (2C + \varepsilon)|z|^\rho$ and the rest is a simple exercise. ■

43 Hadamard's theorem

Let $\rho > 0$ and let k_ρ be the smallest integer strictly greater than ρ , $k_\rho = \lfloor \rho \rfloor + 1$. We consider again the truncates of the series of $-\ln(1 - z)$, namely, with $k = k_\rho$,

$$P_k(z) = z + \frac{z^2}{2} + \cdots + \frac{z^{k-1}}{k-1} \quad (43.1)$$

Theorem 280: Hadamard

Let f be entire of order ρ , let z_n be its nonzero zeros and let $k = k_\rho$. Then, with $m \geq 0$ the order of the zero of f at zero, there is a polynomial h of degree $\leq \rho$ such that

$$f(z) = e^{h(z)} z^m \prod_{n=1}^{\infty} \left[\left(1 - \frac{z}{z_n} \right) e^{P_k(z/z_n)} \right] = e^{h(z)} E(z) \quad (43.2)$$

The proof of this important theorem requires a number of intermediate results, notably the *minimum modulus theorem* proved in the following section, a very useful result in its own right.

Lemma 281. Let ε be such that $\lambda := \rho + \varepsilon < k_\rho := k$. There is a $c > 0$ such that

$$|(1 - \zeta) \exp P_k(\zeta)| \leq \exp(c|\zeta|^\lambda) \quad (43.3)$$

Proof. For $|\zeta| \leq 1/2$ we have

$$\ln(1 - \zeta) + P_k(\zeta) = \sum_{n=k}^{\infty} \frac{\zeta^n}{n} = \zeta^k C_k; \quad |C_k| \leq \sum_{n=k}^{\infty} 2^{-n} \leq 2 \quad (43.4)$$

$$\Rightarrow (1 - \zeta) e^{P_k(\zeta)} \leq e^{2|\zeta|^k} \leq e^{2|\zeta|^\lambda} \quad (43.5)$$

For $|\zeta| \in [1/2, 1]$ we have

$$\begin{aligned} |(1 - \zeta) \exp P_k(\zeta)| &\leq \frac{1}{2} \exp \left[|\zeta|^k \left(\frac{1}{|\zeta|^{k-1}} + \cdots + \frac{1}{|\zeta|(k-1)} \right) \right] \\ &\leq \frac{1}{2} \exp(2^k |\zeta|^k) \leq \frac{1}{2} \exp(2^k |\zeta|^\lambda) \end{aligned} \quad (43.6)$$

For $|\zeta| > 1$ we have

$$\begin{aligned} |(1 - \zeta) \exp P_k(\zeta)| &\leq |1 - \zeta| \exp \left[|\zeta|^{k-1} \left(\frac{1}{k-1} + \cdots + \frac{1}{|\zeta|^{k-2}} \right) \right] \\ &\leq \exp \left(k|\zeta|^{k-1} + \ln |1 + |\zeta|| \right) \leq \exp \left(k|\zeta|^\lambda + \ln |1 + |\zeta|| \right) \leq \exp \left(C_2 |\zeta|^\lambda \right) \end{aligned} \quad (43.7)$$

for some C_2 independent of ζ , $|\zeta| > 1$. This is because $t^{-\lambda} \ln(1 + t)$ is continuous on $[1, \infty)$ and goes to zero at infinity (fill in the details). ■

43.1 Canonical products

Take any sequence $\{z_n\}_n$ where the terms are ordered by absolute value, with the property that for some $\rho > 0$ and any $\varepsilon > 0$ we have

$$\sum_{n=1}^{\infty} \frac{1}{|z_n|^{\rho+\varepsilon}} < \infty \quad (43.8)$$

Definition 282

The canonical product determined by the sequence $\{z_n\}$, denoted by $E^{(k)}(z, \{z_n\})$ or simply $E(z)$ is defined by

$$E(z) = \prod_{n=1}^{\infty} \left(1 - \frac{z}{z_n}\right) \exp[P_k(z/z_n)] \quad (43.9)$$

Theorem 283

$E(z)$ is an entire function of order $\leq \rho$.

Proof. Take again any ε be such that $\lambda := \rho + \varepsilon < k_\rho$. Then, by Lemma 281 we have

$$|E(z)| \leq \prod_{n=1}^{\infty} \exp(c|z/z_n|^\lambda) = \exp\left(c|z|^\lambda \sum_{k=1}^{\infty} |z_n|^{-\lambda}\right) \leq \exp\left(c_1|z|^\lambda\right) \quad (43.10)$$

proving, in the process, uniform convergence of the product. ■

44 The minimum modulus theorem; end of proof of Theorem 280

This important theorem tells us, roughly, that if a function does not grow too fast it cannot decrease too quickly either, aside from zeros. More precisely we have

Theorem 284: Minimum modulus theorem

Let f be an entire function of order $\leq \rho$. As before, let $\{z_n\}$ be its zeros with $|z_i| > 1$, repeated according to their multiplicity and let $\varepsilon > 0$. At every root, remove a disk $\mathbb{D}_{r_n}(z_n)$ with $r_n = |z_n|^{-\rho-\varepsilon}$ $U = \mathbb{C} \setminus \cup_n \mathbb{D}_{r_n}(z_n)$. Then, there is a constant c such that for all $z \in U$ we have

$$|f(z)| \geq \exp(-c|z|^{\rho+\varepsilon}) \quad \text{or} \quad \frac{1}{|f(z)|} = O(\exp(|z|^{\rho+\varepsilon})) \quad (44.1)$$

Proof. We start with the case when the entire function is a canonical product. We take $|z| = r$ and write

$$E(z) = \prod_{|z_n| < 2r} E_k(z, z_n) \prod_{|z_n| \geq 2r} E_k(z, z_n) \quad (44.2)$$

and estimate the two terms separately. We note that in the second product, all ratios $\zeta =: \zeta_n = z/z_n$ have the property $|\zeta| \leq 1/2$. Taking one term of the product, we have to estimate *below*

$$E(\zeta) = (1 - \zeta)e^{P(\zeta)} \quad (44.3)$$

Since $|\zeta| \leq 1/2$, $\ln(1 - \zeta)$ exists; we take the principal branch and write

$$\begin{aligned} |(1 - \zeta)e^{P(\zeta)}| &= |e^{\ln(1 - \zeta) + P(\zeta)}| = \left| \exp \left(- \sum_{n=k}^{\infty} \frac{\zeta^n}{n} \right) \right| \\ &= \exp \left(- \operatorname{Re} \sum_{n=k}^{\infty} \frac{\zeta^n}{n} \right) \geq \exp \left(- \left| \sum_{n=k}^{\infty} \frac{\zeta^n}{n} \right| \right) \geq \exp \left(- \sum_{n=k}^{\infty} \frac{|\zeta|^n}{n} \right) \\ &\geq \exp \left(- |\zeta|^k \sum_{n=0}^{\infty} \frac{(1/2)^n}{n+k} \right) \geq e^{-2|\zeta|^k} \geq e^{-2|\zeta|^{\rho+\varepsilon}} \end{aligned} \quad (44.4)$$

Thus for the second product in (44.2) we have

$$\prod_{|z_n| \geq 2r} E_k(z, z_n) \geq \exp \left(-2|z|^{\rho+\varepsilon} \sum_{|z_n| \geq 2r} \frac{1}{|z_n|^{\rho+\varepsilon}} \right) \geq e^{-c|z|^{\rho+\varepsilon}} \quad (44.5)$$

since the infinite sum converges by Theorem 277.

We split the remaining region into $z_k \in [1, r]$ and $z_k \in (r, 2r)$. Here the factors $(1 - z/z_k)$ have to be bounded from below, and it is here that we use the conditions on the removed disks. We have, on $[1, r]$,

$$\begin{aligned} \prod_{|z_n| \leq r} |1 - z/z_n| &= \frac{|z - z_n|}{|z_n|} \geq \prod_{|z_n| \leq r} |z_n|^{-\rho-\varepsilon-1} \geq \prod_{|z_n| \leq r} r^{-\rho-\varepsilon-1} = \\ &= \left(r^{-\rho-\varepsilon-1} \right)^{v(r)} \gtrsim e^{-r^{\rho+\varepsilon} \ln r (\rho+\varepsilon+1)} \geq e^{-c_1 r^{\rho+\varepsilon'}} \end{aligned} \quad (44.6)$$

for some C_1 and $\varepsilon' > \varepsilon$, since $(r^{\varepsilon-\varepsilon'} \ln r \rightarrow 0$ as $r \rightarrow \infty$.

On $(r, 2r)$ we have

$$|1 - z/z_n| = \frac{|z - z_n|}{|z_n|} \geq |z_n|^{-\rho-\varepsilon-1} \geq (2r)^{-\rho-\varepsilon-1} \quad (44.7)$$

and thus

$$\prod_{|z_n| < 2r} |1 - z/z_n| \geq [(2r)^{-\rho-\varepsilon-1}]^{v(2r)} = e^{-v(2r)(\rho+\varepsilon+1) \ln(2r)} \geq e^{-c_6 r^{\rho+\varepsilon'}} \quad (44.8)$$

for any $\varepsilon' > \varepsilon$ if r is large enough.

We now examine the convergence improving factors, for $|z_n| < 2r$.

$$\left| \sum_{|z_n| < 2r} P_k(z/z_n) \right| \leq \left| \sum_{r < |z_n| < 2r} P_k(z/z_n) \right| + \left| \sum_{|z_n| \leq r} P_k(z/z_n) \right| \quad (44.9)$$

For the first term on the right we note that when $|z/z_n| = |\zeta_n| =: |\zeta| < 1$ and we have

$$\left| \sum_{n=1}^{k-1} \frac{\zeta^n}{n} \right| \leq \sum_{n=1}^{k-1} \frac{1}{n} =: c_1 \quad (44.10)$$

and thus

$$\sum_{r < |z_n| < 2r} |P_k(z/z_n)| \leq v(2r)c_1 \leq c_2 r^{\rho+\varepsilon} \quad (44.11)$$

For the second term on the right of (44.9) we note that $|z/z_n| \geq 1$ and thus, with $\zeta = z/z_n$ we have

$$\left| \sum_{n=1}^{k-1} \frac{\zeta^n}{n} \right| \leq |\zeta|^{k-1} \sum_{n=1}^{k-1} \frac{1}{n} =: c_1 |\zeta|^{k-1} = c_1 r^{k-1} |z_n|^{-k+1} \quad (44.12)$$

We use Abel summation by parts (we are careful that r is not necessarily an integer)

$$\begin{aligned} \sum_{|z_n| \leq r} |z_n|^{-k+1} &\leq \sum_{m \leq r} \frac{\nu(m+1) - \nu(m)}{m^{k-1}} \\ &= \sum_{m \leq r} \nu(m+1) \left(\frac{1}{m^{k-1}} - \frac{1}{(m+1)^{k-1}} \right) + \frac{\nu(r+1)}{r^{k-1}} - \nu(1) \\ &\leq \sum_{m \leq r} \nu(m+1) \left(\frac{1}{m^{k-1}} - \frac{1}{(m+1)^{k-1}} \right) + \frac{\nu(r+1)}{r^{k-1}} \\ &\leq \sum_{m \leq r} \frac{kCm^{\rho+\varepsilon}}{m^k} + c_3 r^{\rho+\varepsilon-k+1} \\ &\leq C_1 \sum_{m \leq r} m^{\rho+\varepsilon-k} + c_3 r^{\rho+\varepsilon-k+1} \leq C_3 r^{\rho+\varepsilon-k+1} \end{aligned} \quad (44.13)$$

where we majorized the sum by an integral in the usual way. Multiplying by $c_1 r^{k-1}$ we get that the second term on the right of (44.9) is bounded by

$$C_4 r^{\rho+\varepsilon} \quad (44.14)$$

We now finish the proof of Theorem 280.

Proof. We take $\varepsilon > 0$ and $s = \rho + \varepsilon$. We order the roots nondecreasingly by $|z_n|$. For each root z_n we consider the annulus $A_n = \{z : |z| \in [|z_n| - 2|z_n|^{-s}, |z_n| + 2|z_n|^{-s}]\}$ (for the purpose of this argument, we can as well assume that all roots are on $\mathbb{R}^+$, since the angular position is irrelevant). Consider $J^c := \mathbb{R}^+ \setminus J$ where J is the union of all intersections of the A_n with $\mathbb{R}^+$. Since the Lebesgue measure of J does not exceed $4 \sum_n |z_n|^{-s} < \infty$, there exist arbitrarily large numbers in the complement J^c . Take r be any number in J^c and consider the circle $\partial \mathbb{D}_r$. Consider the function $g = f/E$. g is clearly an entire function with no zeros. Then, by Theorem 269, $g = e^h$ with h entire. Since $\operatorname{Re} h \leq (C_f + C_E) r^{\rho+\varepsilon}$ for some $C_f + C_E$ independent of r in $\mathbb{D}_r$ for arbitrarily large r (check), we have by Corollary 279 that h is a polynomial of degree at most $\rho + \varepsilon$. Since ε is arbitrary, h is a polynomial of degree at most ρ . ■

To finish the proof of the minimum modulus principle, we use Hadamard's theorem and the fact that e^{-h} satisfies the required bounds. (Exercise: fill in the details.) ■

Example 285. Let us show that $f(z) = e^z - z$ has infinitely many roots in $\mathbb{C}$. Indeed, first note that $f(z)$ has order 1 since $|z| \leq e^{|z|}$ for all z . Suppose f had finitely many zeros. Then

$$e^z - z = P(z) e^{h(z)} \quad (44.15)$$

where $P(z)$ is a polynomial and $h(z)$ is a polynomial of degree one, and without loss of generality we can take $h(z) = cz$, $c = \alpha + i\beta$. As $z = t \rightarrow +\infty$ we have

$$P(t) e^{(c-1)t} = 1 - t e^{-t} \rightarrow 1 \quad (44.16)$$

In particular $|P(t)|e^{(\alpha-1)t} \rightarrow 1$ which is only possible if $\alpha = 1$. But then $|P(t)| \rightarrow 1$ which is only possible if $P(t) = \text{const} = e^{i\varphi}$. We are then left with

$$e^{i(\varphi+t\beta)} = \cos(t\beta + \varphi) + i \sin(t\beta + \varphi) \rightarrow 1 \text{ as } t \rightarrow +\infty \quad (44.17)$$

which clearly implies $\beta = 0$. Then $e^{i\varphi} = 1$. We are left with the identity

$$e^z - z = e^z \quad \forall z \quad (44.18)$$

which is obviously false.

Exercise 286. * Let $P \not\equiv 0$ be a polynomial. Show that the equation $e^z = P(z)$ has infinitely many roots in $\mathbb{C}$.

Exercise 287. ** (i) Rederive formula (42.12) using Hadamard's theorem.
(ii) Write down a product formula of the function

$$f(z) = \sin z + 3 \sin(3z) + 5 \sin(5z) + 7 \sin(7z)$$

The final formula should be explicit except for arcsins of roots of a cubic polynomial.

44.1 Some applications

Corollary 288 (Borel). Assume that ρ is not an integer and f has strict order ρ . Then f takes every value in $\mathbb{C}$ infinitely many times.

Proof. It suffices to show that such a function has infinitely many zeros, since f and $f - z_0$ have the same strict order. Assume to get a contradiction f had finitely many zeros. Then $g(z) = f(z) \prod_{i=1}^n (z - z_i)^{-1}$ would be entire, with no zeros, and as it is easy to check, of strict order ρ . Then $g = e^h$ with h a polynomial whose degree can only be an integer. ■

Definition 289

Let $\exp_n$ be the exponential composed with itself n times.

Corollary 290 (A weak form of Picard's theorem). A nonconstant entire function which is bounded by $\exp_n(C|z|)$ for some n and large z takes every value with at most one exception.

Proof. We prove this by induction on n . We first show that a nonconstant entire function of finite order takes every value with at most one exception. Assume a is an exceptional (lacunary) value. Then $f(z) - a$ is entire with no zeros, thus of the form e^h with h a polynomial, $f = e^h - a$. If the degree of h is zero, then f is a constant. Otherwise, we must show that $e^h - a$ takes all values with at most one exception ($-a$ of course), or, which is the same, e^h takes all values with at most one exception. The equation $e^h = b$, $b \neq 0$ is solved if $h - \ln b$ has roots, which is true by the fundamental theorem of algebra.

Assume now the property holds for $n \leq k-1$ and we wish to prove it for $n = k$. Let f be an entire function bounded by $\exp^{(n)}(C|z|)$ which avoids the value a . Then $f - a$ is entire with no zeros, $f - a = e^h$ with h entire. It is easy to show that h is bounded by $\exp^{(n-1)}(C|z|)$. Thus it avoids at most one value, by the induction hypothesis. The equation $e^h - a = b$, for $b \neq -a$ always has a solution. Indeed, if $\ln(b + a)$ is not an avoided value of h this is obvious. On the other hand, if $\ln(b + a)$ is avoided by h , then again by the induction hypothesis $\ln(b + a) + 2\pi i$ is not avoided. ■

Exercise 291. ** Show that the equation

$$\cos(z) = z^4 + 5z^2 + 13 \quad (44.19)$$

has infinitely many roots in $\mathbb{C}$.

Exercise 292. ** (Bonus) Show that the error function

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt \quad (44.20)$$

takes every complex value infinitely many times. (Hint: using L'Hospital show that $\operatorname{erf}(is)/(e^{s^2}/s) \rightarrow \text{const.}$ as $s \rightarrow +\infty$.)

We will return to the error function later and use asymptotic methods to locate, for large x , these special points.

45 The little Picard theorem

Theorem 293: Little Picard theorem

If $f : \mathbb{C} \rightarrow \mathbb{C}$ is entire and there are two points $a \neq b$ which are not in the range of f , then f is a constant.

Points that are not in the range of f are called *lacunary values*.

Proof. Assume that f is entire and has two lacunary values. Without loss of generality, we can assume that these are 0 and 1. Let $h = C \circ \psi \circ f$ where C is the Cayley transform and ψ is as in Proposition 38.7. Since the only singularities of ψ are at 0 and 1, ψ takes M to the UHP and C is analytic in the UHP, h is entire with values in $\mathbb{D}$, meaning h , hence f are constants. ■

Theorem 294: Montel's fundamental normality test

Let $\mathcal{D}$ be a domain in $\mathbb{C}$, let α and β be two distinct complex numbers, and let $\mathcal{F}$ be the set of holomorphic functions in $\mathcal{D}$ whose range omits the two values α and β . Then $\mathcal{F}$ is a normal family in $\hat{\mathbb{C}}$.

A stronger version, using the uniformization theorem is ([8])

Theorem 295

Let M be a Riemann surface and let $\mathcal{F}$ be the set of holomorphic functions on M with values in $\hat{\mathbb{C}}$ whose range omits the three values α , β and γ . Then $\mathcal{F}$ is a normal family in $\hat{\mathbb{C}}$.

Proof of Theorem 294. We first arrange the following: $\alpha = 0$, $\beta = 1$, and, since normality is a local property, that $\mathcal{D} = \mathbb{D}$.

With $\psi = \lambda^{-1}$, let $\eta = C \circ \psi$ where C is the Cayley transform; the function η is a conformal map from M , the universal cover of $\hat{\mathbb{C}} \setminus \{0, 1, \infty\}$ to $\mathbb{D}$.

Let $f \in \mathcal{F}$ and $a \in \mathbb{D}$ and choose any branch of $\eta(f(a))$. We note, as in the uniformization proof, that for that branch, $\tilde{f} = \eta \circ f$ defined by analytic continuation from a to $\mathbb{D}$ is analytic in $\mathbb{D}$.

To construct a convergent subsequence of a given sequence $\{f_n \in \mathcal{F}\}_{n \in \mathbb{N}}$ we first note that for any $a \in \mathbb{D}$, there is a subsequence of $\{f_n(a)\}$ which converges in $\hat{\mathbb{C}}$, which, by passing to this subsequence, we can assume is $\{f_n(a)\}_{n \in \mathbb{N}}$ itself. If for all a and any subsequence the limit is infinity, then f_n converges to the constant $\infty \in \mathbb{C}$ proving the result. Otherwise, there is a point, say $a = 0$ s.t. a subsequence of $\{f_n(0)\}_{n \in \mathbb{N}}$ converges to $\ell \in \mathbb{C}$. We assume first that $\ell \notin \{0, 1\}$.

Note now that $\tilde{f}_n, n \in \mathbb{N}$ have range in $\overline{\mathbb{D}}$, hence they are a normal family and by Theorem 158, they have a convergent subsequence, that by passing again to a subsequence, we can assume is $\{\tilde{f}_n\}_{n \in \mathbb{N}}$; let the limit be g . If g is a constant, then $g = \eta(\ell) \in \mathbb{D}$. Otherwise, since $\text{ran}(g) \subset \overline{\mathbb{D}}$, by the open mapping theorem, $\text{ran}(g) \subset \mathbb{D}$. Now we note that $\eta^{-1} \circ g$ is well-defined, and that $\lim_{n \rightarrow \infty} f_n = \lim_{n \rightarrow \infty} \eta^{-1} \tilde{f}_n = \eta^{-1} \circ g$, which finishes the proof if $\ell \notin \{0, 1\}$.

The remaining case is $\ell = 1$ since, if $\ell = 0$, then we can take the family $1 - f, f \in \mathcal{F}$. Since $f_n \neq 0$, there is a *holomorphic* branch of the square root such that $h_n = \sqrt{f_n} \rightarrow -1$. A holomorphic branch of the square root cannot take both values 1 and -1 , so h_n also avoid 1 and -1 . For h_n the proof above applies to extract a convergent subsequence, which by squaring gives the desired subsequence of f_n . ■

Note 296

The fundamental normality test plays an important role in complex dynamics, see [8].

We are now going to prove the celebrated Great Picard's theorem:

Theorem 297

If an analytic function f has an essential singularity at a point z_0 , then on any punctured neighborhood of z_0 , f takes on all possible complex values, with at most a single exception, infinitely often.

Proof. Without loss of generality, we take $z_0 = 0$, assume that f is analytic in $\mathbb{D}^* = \mathbb{D} \setminus \{0\}$ and assume, to get a contradiction, that f omits the values 0 and 1. Let $\{z_n\}_{n \in \mathbb{N}}$ be any sequence converging to 0 and $f_n(z) = f(z_n z)$. By Theorem 294 the family $\{f_n\}_{n \in \mathbb{N}}$ is normal. Hence we can extract a subsequence $\{f_{n_k}\}_{k \in \mathbb{N}}$ convergent on compact sets in $\mathbb{D}_r \setminus \{0\}$ to, say g which is analytic in $\mathbb{D}_r \setminus \{0\}$ with values in $\hat{\mathbb{C}}$. Assume that g is not the constant $\infty \in \hat{\mathbb{C}}$. Note that, for s in the annulus between the circle of radius $12|z_{n_k}|$ and the circle of radius $\frac{1}{4}|z_n|$, we have $f(s) - g(s') \rightarrow 0$ as $n \rightarrow \infty$ and s' is in the annulus between the circle of radius $\frac{1}{2}$ and the circle of radius $\frac{1}{4}$. Hence f is bounded along any sequence converging to zero, contradiction. If $g = \infty$ then $1/f$ has a removable singularity at 0, thus f has a pole at zero, again a contradiction. ■

46 Riemann-Hilbert problems: an introduction

An impressive number of problems coming from integrable world can be reduced to so-called Riemann Hilbert problems, and for many of them the only known way to get a closed form solution is via the associated Riemann-Hilbert problem.

Problems which can be solved with R-H techniques include

(1) integrable models such as the transcendental Painlevé equations e.g. $y'' = 6y^2 + x$ (P_I) and many others;

(2) integrable PDEs s.a. nonlinear initial value problem for the KdV (Korteweg–deVries) equation

$$u_t + u_{xxx} + uu_x = 0; \quad u(x, 0) = u_0(x), \quad u \rightarrow 0 \text{ as } |x| \rightarrow \infty \quad (x \in \mathbb{R}, t \in \mathbb{R}^+) \quad (46.1)$$

(3) inverse scattering problems: find, from the scattering data the potential $q(x)$ in the time-independent Schrödinger equation

$$-\psi_{xx} + (k^2 + q(x))\psi = 0 \quad (46.2)$$

(4) questions in orthogonal polynomials, random matrices, and combinatorial probability;

(5) integral equations of the type

$$f(t) + \int_0^\infty \alpha(t - t')f(t')dt' = \beta(t) \quad (46.3)$$

(under suitable integrability conditions);

(6) finding the inverse Radon transform, a transform which is measured in tomography.

46.1 A simple Riemann-Hilbert problem

Perhaps the simplest R-H problem is: given a simple smooth contour $\mathcal{C}$ and $f(t)$ a suitably regular function on $\mathcal{C}$, find analytic functions Φ^+, Φ^- , defined to the left and right of $\mathcal{C}$ such that the limits of $\Phi^\pm$ on $\mathcal{C}$ exist and satisfy

$$\Phi^+(z) - \Phi^-(z) = f(z) \quad (46.4)$$

Note 298

Φ is clearly determined up to an entire function. To hope for a unique solution one needs to impose more conditions on Φ , such as behavior at infinity and other special points.

47 Cauchy type integrals

We recall that a function is Hölder continuous of order β on a smooth curve $\mathcal{C}$ if

$$\exists \beta > 0 \text{ and } C > 0 \text{ s.t. } \forall x, y \in \mathcal{C}, |f(x) - f(y)| \leq C|x - y|^\beta \quad (47.1)$$

The condition implies continuity if $\beta > 0$ and it is nontrivial if $\beta \leq 1$ (if $\beta > 1$ then $df/ds = 0$).

Let $\mathcal{C}$ for now be a compact smooth curve and φ be Hölder continuous on $\mathcal{C}$. Then the function

$$\Phi(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\varphi(s)}{s - z} ds \quad (47.2)$$

is manifestly analytic for $z \notin \mathcal{C}$ (you can check this by Morera's theorem using Fubini, and in a good number of other ways).

Revisit §10.2 and Definition 179.

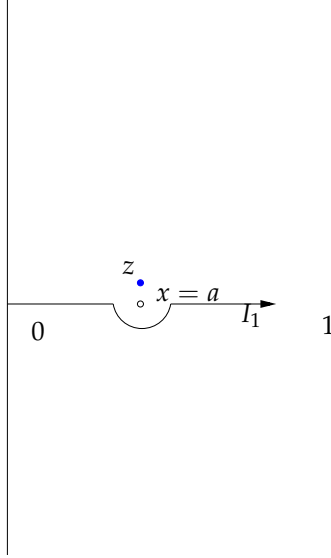


Figure 59: Analyticity of Φ when φ is analytic in a neighborhood of $\mathcal{C}$.

47.1 Asymptotic behavior of $\Phi(z)$ for large z

Exercise 299. Assume $\mathcal{C}$ is compact. Show that $\Phi(z)$ is analytic at infinity in z and that

$$\Phi(z) = - \left(\frac{1}{2\pi i} \int_{\mathcal{C}} \varphi(s) ds \right) \frac{1}{z} (1 + g(1/z)) \quad \text{as } z \rightarrow \infty \quad (47.3)$$

as where $g(1/z) \rightarrow 0$ as $z \rightarrow \infty$.

47.2 Regularity and singularities

Let us first take a simple example, in which φ is analytic in a neighborhood $\mathcal{N}$ of $(0, 1)$ (note that this does not exclude multi-valued functions singular at zero and one s.a. $\ln z(1 - z)$).

Then Φ has analytic continuation from $\mathbb{H}^u$ and $\mathbb{H}_l$ through $\mathcal{C} \setminus \{c_i\}$ in $\mathcal{N}$, where c_i are endpoints of $\mathcal{C}$ (if any). (These analytic continuations are in general different.) Indeed, to perform analytic continuation in $z \in \mathbb{H}^u \cup (\mathcal{N} \setminus \{c_i\})$, we can first deform the contour as shown in de-form as shown in Fig. 59. To calculate the values of Φ on $\mathcal{C}$, we can use Exercise 60 to see that $A(a) = \pm \frac{1}{2} \varphi(a) + \frac{1}{2\pi i} PV \int_0^1 \varphi(s)(s - t)^{-1} ds$.

The same, interestingly, holds more generally:

Theorem 300: Plemelj's formulas

Assume φ is Hölder continuous of exponent β on the simple smooth curve $\mathcal{C}$. let t be an interior point of $\mathcal{C}$. Then,

(i)

$$PV \int \frac{\varphi(s)}{s - t} ds \quad (47.4)$$

(cf. Definition 61) exists.

(ii) Let $\{z_n\}_{n \in \mathbb{N}}$ be a sequence approaching t from the left (right). If $\mathcal{C}$ is not bounded, assume also that $\varphi \in L^1(\mathcal{C})$. Then, with the $\pm$ sign being $+$ for left limit and $-$ for right

limit,

$$\lim_{n \rightarrow \infty} \Phi(z_n) = \Phi^\pm(t) \quad (47.5)$$

where

$$\Phi^\pm(t) = \pm \frac{1}{2} \varphi(t) + \frac{1}{2\pi i} PV \int \frac{\varphi(s)}{s-t} ds \quad (47.6)$$

and

Note 301

(i) The properties in the theorem are local; it is enough to prove them for compact pieces of $\mathcal{C}$.

(ii) A similar statement can be obviously made when t is approached along a curve, since all limits along subsequences coincide.

We leave it as an exercise to extend the proof from the case when $\mathcal{C}$ is a piece of $\mathbb{R}$, say $[a, b]$ with $t = 0 \in (a, b)$, approached from above, to a more general smooth bounded curve (open or not): parametrize the curve.

Proof. We can assume without loss of generality that $0 \in [a, b]$, $t = 0$ and $|a| > |b|$.

(i) We write $\varphi(s) = (\varphi(s) - \varphi(0)) + \varphi(0)$ and show that both PV integrals exist. By symmetry,

$$PV \int_{-b}^b s^{-1} \varphi(0) ds = 0 \quad (47.7)$$

implying

$$PV \int_a^b \frac{\varphi(0)}{s} ds = \int_a^{-b} \frac{\varphi(0)}{s} ds \quad (47.8)$$

We then note that $|\varphi(s) - \varphi(0)|/|s| = (|\varphi(s) - \varphi(0)|s|^{-\beta})|s|^{\beta-1} \in L^1([a, b])$, hence

$$PV \int_a^b \frac{\varphi(s) - \varphi(0)}{s} ds = \int_a^b \frac{\varphi(s) - \varphi(0)}{s} ds \quad (47.9)$$

(ii) Take $\varepsilon > 0$ small enough, n large enough, and $0 < c \in (a, b)$, change variable to $u = s - x_n$ and break the interval of integration into $[a - x_n, -c]$, $[-c, c]$ and $[c, b - x_n]$.

$$\lim_{n \rightarrow \infty} \left(\int_{a-x_n}^{-c} + \int_c^{b-x_n} \right) \frac{\varphi(u+x_n) du}{u - iy_n} = \left(\int_a^{-c} + \int_c^b \right) \frac{\varphi(u) du}{u} \quad (47.10)$$

We decompose the integral on $[-c, c]$ into

$$\int_{-c}^c \frac{\varphi(u+x_n) - \varphi(x_n)}{|u|^\beta} \frac{|u|^\beta du}{u - iy_n} + \varphi(x_n) \int_c \frac{du}{u} \quad (47.11)$$

where in the last integral we homotopically deformed the contour into a half-circle of radius c in the LHP centered at zero. For all n , the integrand in the first integral is bounded in absolute value, up to a constant, by $|u|^{\beta-1} \in L^1([-c, c])$ while the second integral evaluates to πi . The result now follows by using dominated convergence and the symmetry argument in (i). ■

The following result follows immediately.

Theorem 302: Existence

Under the conditions of Theorem 300, the function in (47.2) solves the Riemann-Hilbert problem in §46.1.

Note 303

The Hölder condition can be replaced with the weaker Dini condition, with the same proof.

Note 304

(a) The function defined by the Cauchy type integral (47.2) is called **sectionally analytic**. With the convention about the sides of the curve mentioned before, functions that are boundary values of Cauchy type integrals are sometimes denoted $\oplus$ and $\ominus$ functions. respectively.

(b) Branch jumps of analytic functions are an essential ingredient in Sato's theory of hyperfunctions.

47.3 Examples**47.3.1 A very simple example**

Find a function Φ analytic in $\mathbb{C} \setminus \partial\mathbb{D}$ such that along $\partial\mathbb{D}$ we have

$$\Phi^+(t) - \Phi^-(t) = 1 \quad (47.12)$$

Note that for this problem the set of analyticity of Φ is a union of two disjoint domains. Since Φ is discontinuous across $\partial\mathbb{D}$, we are dealing with two separate analytic functions. Plemelj's formula reads

$$\Phi(z) = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{1}{s-z} ds \quad (47.13)$$

Clearly, if $z \in \mathbb{D}$ (which is to the left of $\partial\mathbb{D}$ oriented positively), we have

$$\Phi_{\text{in}}(z) = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{1}{s-z} ds = 1 \quad (47.14)$$

Likewise, if z is outside $\mathbb{D}$ we have

$$\Phi_{\text{out}}(z) = \frac{1}{2\pi i} \int_{\partial\mathbb{D}} \frac{1}{s-z} ds = 0 \quad (47.15)$$

Both Φ_{in} and Φ_{out} are analytic, but not analytic continuations of each-other, so in this case our sectionally analytic function is really a pair of distinct analytic functions. We leave the question of uniqueness to the next subsection when the contour is open and which leads to a more interesting discussion.

47.3.2 Another simple example

Find a function Φ analytic in $\mathbb{C} \setminus [-1, 1]$ such the upper and lower limits across $[-1, 1]$ satisfy

$$\Phi^+(z) - \Phi^-(z) = 1 \quad (47.16)$$



Figure 60: Contour used for applying Morera's theorem.

47.3.3 A solution

According to Plemelj's formulas a function satisfying (47.16) is given by

$$\Phi(z) = \frac{1}{2\pi i} \int_{-1}^1 \frac{1}{s-z} ds \quad (47.17)$$

This formula shows Φ is analytic on the universal cover of $\hat{\mathbb{C}} \setminus \{-1, 1\}$ (Φ is analytic at infinity). We get explicitly

$$\Phi(z) = -\frac{1}{2\pi i} (\ln(z-1) - \ln(z+1)) \quad (47.18)$$

where the log is determined by the condition $\Phi(z) < 0$ for $z > 1$. It is useful to check directly on (47.18) that $\Phi^+ - \Phi^- = 1$ on $(-1, 1)$

47.3.4 Calculating principal value integrals

Plemelj's formulas help us calculate principal values integrals as well, sometimes in a simpler way. Let $\mathcal{C}$ be a simple smooth closed curve and assume that $f(z)$ is analytic in $\text{Int}(\mathcal{C})$ and Hölder continuous in the closure of $\text{Int}(\mathcal{C})$. Then,

$$\frac{1}{2\pi i} PV \int_{\mathcal{C}} \frac{f(s)}{s-t} ds = \frac{1}{2} f(t) \quad (47.19)$$

a "limiting case" of a Cauchy formula.

47.3.5 Uniqueness issues

What obvious freedoms do we have in solving the problem in §47.3.2? As mentioned, we can add to Φ any entire function. More generally, we can add any function analytic in $\mathbb{C} \setminus \{-1, 1\}$ (note that this means single-valued). The following theorem shows that these are all the freedoms.

Theorem 305: Uniqueness

Consider the problem in §46.1 with the following further conditions:

- (1) $\Phi(z) \rightarrow 0$ as $|z| \rightarrow \infty$;
- (2) $(1 \mp z)\Phi(z) \rightarrow 0$ as $z \rightarrow \pm 1$

Then the solution is unique, namely (47.2).

Proof. The Φ in (47.2) satisfies this condition, as it is easy to verify. Assume Φ_1 is another solution with the same properties. Then $f = \Phi - \Phi_1$ is analytic in $\mathbb{C} \setminus [-1, 1]$ and continuous on $(-1, 1)$, entailing continuity in $\mathbb{C} \setminus \{-1, 1\}$. Show that the contour integral of $\Phi - \Phi_1$ on a circle of radius

ε around 1 and -1 is zero, hence it vanishes on any closed contour in $\mathbb{C}$ (see Fig. 60) implying, by Morera, that $\Phi - \Phi_1$ is entire. Since $\lim_{z \rightarrow \infty} (\Phi(z) - \Phi_1(z)) = 0$, $\Phi - \Phi_1 = 0$. ■

Exercise 306. Check using Plemelj's formulas that any simple smooth curve $\mathcal{C}$ in $\mathbb{C}$ is the natural boundary of a large class analytic of functions, and furthermore the limit of such functions on $\mathcal{C}$ can be as smooth as we want, short of analytic. Use the Riemann mapping theorem to show that any simple Jordan curve is a natural boundary of a large class of analytic functions.

We see from

47.3.6 Degree of a function at infinity

By definition Φ has degree k at infinity if for some $C \neq 0$ we have

$$\Phi(z) = Cz^k + O(z^{k-1}) \quad \text{as } z \rightarrow \infty \quad (47.20)$$

The function Φ has finite degree at infinity if $\Phi = o(z^m)$ for some m .

48 More general scalar R-H problems

48.1 Scalar homogeneous R-H problems

48.1.1 Scalar homogeneous R-H problems

This is a problem of the type

$$\Phi^+ = g\Phi^- \quad \text{on } \mathcal{C} \quad (48.1)$$

where $\mathcal{C}$ is a smooth simple closed contour, g nonzero on $\mathcal{C}$ and satisfying a Hölder condition on $\mathcal{C}$. We are looking for solutions of degree m at infinity.

Solution to (48.1) First we note that if Φ is a solution and H is entire, then by homogeneity ΦH is also a solution.

Assume that the *index of g w.r.t. $\mathcal{C}$* is k . Without loss of generality we assume $0 \in \text{Int}(\mathcal{C})$. We can rewrite the problem as

$$\Phi^+(t) = (t^{-k}g(t))(t^k\Phi^-) \quad \text{on } \mathcal{C} \quad (48.2)$$

Formally for now, taking the log of both sides, we get

$$\ln \Phi^+(t) = \ln(t^{-k}g(t)) + \ln(t^k\Phi^-) \quad \text{on } \mathcal{C} \quad (48.3)$$

or finally, with obvious notation,

$$\Gamma^+(t) = f(t) + \Gamma^-(t) \quad \text{on } \mathcal{C} \quad (48.4)$$

The reason we formed the combination $t^{-k}g(t)$, where we choose k to be the index of g w.r.t $\mathcal{C}$, is to ensure Hölder continuity of f . Otherwise, since $\arg g$ changes by $2k\pi$ upon traversing $\mathcal{C}$, f would have a jump discontinuity somewhere on $\mathcal{C}$.

A solution to (48.4) is given by Plemelj's formulas:

$$\Gamma(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{f(s)}{s-z} ds \quad (48.5)$$

The exterior of $\mathcal{C}$ corresponds to the negative sign and we have $\Gamma^- = O(z^{-1})$ for large z , and thus $\exp(\Gamma^-) = z^k\Phi^- = 1 + o(1)$ for large z , hence $\Phi^- = z^{-k} + o(z^{-k})$. We finally get a solution of

degree m at infinity by multiplying by P_{m+k} , a polynomial of degree $m+k$. The solution of (48.1) is then

$$\Phi = \begin{cases} P_{m+k}(z) \exp(\Gamma(z)), & z \text{ inside } \mathcal{C} \\ z^{-k} P_{m+k}(z) \exp(\Gamma(z)), & z \text{ outside } \mathcal{C} \end{cases} \quad (48.6)$$

We will not, for reasons of space, discuss uniqueness issues here.

48.1.2 Inhomogeneous R-H problems

These are equations of the form

$$\Phi^+ = g\Phi^- + f \quad (48.7)$$

Where f and g are Hölder continuous. These can be brought to Plemelj's formulas in the following way. We first solve the homogeneous problem

$$\Psi^+ = g\Psi^- \quad (48.8)$$

which we dealt with in §48.1.1, and look for a solution of (48.7) in the form $\Phi = X\Psi$. We get

$$X^+\Psi^+ = gX^-\Psi^- + f \Rightarrow X^-g\Psi^+ = X^-g\Psi^- + f \Rightarrow \Psi^+ - \Psi^- = \frac{f}{gX^-} \quad (48.9)$$

which is of the form we already solved.

48.2 Applications

48.2.1 Inhomogeneous singular integral equations

These are equations of the form

$$a(t)\varphi(t) + b(t)PV \int_{\mathcal{C}} \frac{\varphi(s)}{s-z} ds = c(t) \quad (48.10)$$

with a, b, c Hölder continuous and the further condition $i\pi a(t) \pm b(t) \neq 0$. We attempt to write, guided by Plemelj's formulas

$$\varphi(t) = \Phi^+(t) - \Phi^-(t) \quad (48.11)$$

and

$$\Phi(z) = \frac{1}{2\pi i} \int_{\mathcal{C}} \frac{\varphi(s)}{s-z} ds \quad (48.12)$$

and then

$$PV \int_{\mathcal{C}} \frac{\varphi(s)}{s-t} ds = i\pi [\Phi^+(t) + \Phi^-(t)] \quad (48.13)$$

where, of course φ is still unknown. The equation becomes

$$a(t) [\Phi^+(t) - \Phi^-(t)] + b(t)i\pi [\Phi^+(t) + \Phi^-(t)] = c(t) \quad (48.14)$$

or

$$\Phi^+(t)(a(t) + b(t)i\pi) + \Phi^-(t)(b(t)i\pi - a(t)) = c(t) \quad (48.15)$$

or, finally,

$$\Phi^+(t) = \frac{a(t) - b(t)i\pi}{a(t) + b(t)i\pi} \Phi^-(t) + \frac{c(t)}{a(t) + b(t)i\pi} \quad (48.16)$$

which is of the form (48.7) which we addressed already. Care must be taken that the chosen solution Φ is such that $\Phi^+ + \Phi^-$ has the behavior (47.3) at infinity. Then the substitution is a posteriori justified. We did not discuss whether there are other solutions of the integral equation. A complete discussion of this and related equations can be found in [14].

We choose one of the applications in [1], the solution of the Dirichlet problem for the Laplacian in the upper half plane, with condition f on the boundary, $\mathbb{R}$. The problem can be reformulated as a Riemann-Hilbert problem of the form (48.16), see [1], but in this case, the solution is obtained easily from Plemelj's formulas.

For this purpose we look for an analytic function in $\mathbb{H}^u$ generated by u . As we know, this is

$$\Phi^+ = u + iv \quad (48.17)$$

where v is the harmonic conjugate of u , unique up to a constant. Now note that the function

$$\Phi(z) = \frac{1}{\pi i} \int_{\mathbb{R}} \frac{f(s)}{s - z} ds \quad (48.18)$$

is analytic in the upper half plane and, if f is Hölder continuous, then by Plemelj's formulas we have

$$\lim_{z \downarrow t \in \mathbb{R}} \Phi(z) = f(t) + \frac{1}{\pi i} PV \int_{\mathbb{R}} \frac{f(s)}{s - t} ds \quad (48.19)$$

In particular $u = \operatorname{Re} \Phi$ is harmonic and has the limit $f(x, y)$ as $(x, y) \rightarrow (t, 0)$. Now, simply writing $z = x + iy$ and taking the real part we get the solution in the Poisson kernel form,

$$u(x, y) = \frac{y}{\pi} \int_{-\infty}^{\infty} \frac{f(\tau) d\tau}{(\tau - x)^2 + y^2} \quad (48.20)$$

The condition that f is Hölder can be relaxed to mere continuity as follows. To better adapt to the limit $y \rightarrow 0$ we change variable to $t = x + \beta y$ and obtain

$$u(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{f(x + \beta y)}{\beta^2 + 1} d\beta \quad (48.21)$$

which, by dominated convergence, has $f(x)$ as the limit when $y \rightarrow 0$.

49 Asymptotic series

We have seen in the Schwarz-Christoffel section that the behavior of analytic functions near a point of nonanalyticity can be given by a series in noninteger powers of the distance to the singularity. The behavior can be more complicated, containing exponentially small corrections, logarithmic terms and so on. The series themselves may have zero radius of convergence. It is not the purpose of this part of the course to classify these behaviors, but it can be done for a fairly large class of functions. Here we look how simple behaviors can be determined for relatively simple functions.

Example 307. Consider the following integral related to the so-called error function

$$F(z) = e^{z^{-2}} \int_0^z s^{-2} e^{-s^{-2}} ds$$

It is clear that the integral converges at the origin, if the origin is approached through real values (see also the change of variable below).

Definition of $F(z)$. We define the integral to $z \in \mathbb{C}$ as being taken on a curve γ with $\gamma'(0) > 0$, and define $F(0) = 0$.

Check that this is a consistent definition and the resulting function is analytic except at $z = 0$ (this is essentially the contents of Exercise 309 below).

What about the behavior at $z = 0$? It depends on the direction in which 0 is approached! Let's look more carefully. Replace z by $1/x$, make a corresponding change of variable in the integral and you are led to

$$E(x) = e^{x^2} \int_x^\infty e^{-s^2} ds =: \frac{\sqrt{\pi}}{2} e^{x^2} \operatorname{erfc}(x) \quad (49.1)$$

Let us take x (and thus z) real and integrate by parts m times

$$\begin{aligned} E(x) &= \frac{1}{2x} - \frac{e^{x^2}}{2} \int_x^\infty \frac{e^{-s^2}}{s^2} ds = \frac{1}{2x} - \frac{1}{4x^3} + \frac{3e^{x^2}}{4} \int_x^\infty \frac{e^{-s^2}}{s^4} ds = \dots \\ &= \sum_{k=0}^{m-1} \frac{(-1)^k}{2\sqrt{\pi}} \frac{\Gamma(k + \frac{1}{2})}{x^{2k+1}} + \frac{(-1)^m e^{x^2} \Gamma(m + \frac{1}{2})}{\sqrt{\pi}} \int_x^\infty \frac{e^{-s^2}}{s^{2m}} ds \end{aligned} \quad (49.2)$$

On the other hand, we have, by L'Hospital

$$\left(\int_x^\infty \frac{e^{-s^2}}{s^{2m}} ds \right) \left(\frac{e^{-x^2}}{x^{2m+1}} \right)^{-1} \rightarrow \frac{1}{2} \text{ as } x \rightarrow \infty \quad (49.3)$$

and the last term in (49.2) is $O(x^{-2m-1})$ as well. On the other hand it is also clear that the series in (49.2) is alternating and thus

$$\sum_{k=0}^{m-1} \frac{(-1)^k}{2\sqrt{\pi}} \frac{\Gamma(k + \frac{1}{2})}{x^{2k+1}} \leq E(x) \leq \sum_{k=0}^m \frac{(-1)^k}{2\sqrt{\pi}} \frac{\Gamma(k + \frac{1}{2})}{x^{2k+1}} \quad (49.4)$$

if m is even.

Remark 308. Using (49.3) and Exercise 317 below we conclude that $F(z)$ has a Taylor series at zero,

$$\tilde{F}(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{2\sqrt{\pi}} \Gamma(k + \frac{1}{2}) z^{2m+1} \quad (49.5)$$

that $F(z)$ is C^∞ on $\mathbb{R}$ and analytic away from zero.

Exercise 309. ** Show that $z = 0$ is an isolated singularity of $F(z)$. Using Remark 308, show that F is unbounded as 0 is approached along some directions in the complex plane.

Notes (1) It is *not* the Laurent series of f at 0 that we calculated! Laurent series converge. Think carefully about this distinction and why the positive index coefficients do not coincide.

(2) The rate of convergence of the Laurent series is slower as 0 is approached, quickly becoming numerically useless. By contrast, the precision gotten from (49.4) near zero is such that for $z = 0.1$ the error in calculating f is of order 10^{-45} ! However, of course (49.4) is divergent and it cannot be used to calculate *exactly* for any nontrivial value of z .

(3) We have illustrated here a simple method of evaluating the behavior of integrals, the method of integration by parts.

49.1 More general asymptotic series

Classical asymptotic analysis typically deals with the qualitative and quantitative description of the behavior of a function close to a point, usually a singular point of the function. This description is provided in the form of an *asymptotic expansion*. The expansion certainly depends on the point studied and, as we have noted, often on the direction along which the point is approached (in the case of several variables, it also depends on the relation between the variables as the point is approached). If the direction matters, it is often convenient to change variables to place the special point at infinity.

Asymptotic expansions are formal series²⁴ of simpler functions f_k ,

$$\tilde{f} = \sum_{k=0}^{\infty} f_k(t) \quad (49.6)$$

in which each successive term is much smaller than its predecessors (one variable is assumed for clarity). For instance if the limiting point is t_0 approached from above along the real line this requirement is written

$$f_{k+1}(t) = o(f_k(t)) \quad \text{or} \quad f_{k+1}(t) \ll f_k(t) \quad \text{as } t \downarrow t_0 \quad (49.7)$$

denoting

$$\lim_{t \rightarrow t_0^+} f_{k+1}(t)/f_k(t) = 0 \quad (49.8)$$

We will often use the variable x when the limiting point is $+\infty$ and z when the limiting point is zero. Simple examples are the Taylor series, e.g.

$$\sin z + e^{-\frac{1}{z}} \sim z - \frac{z^3}{6} + \dots \quad (z \rightarrow 0^+)$$

and the expansion in the Stirling formula

$$\ln \Gamma(x) \sim x \ln x - x - \frac{1}{2} \ln x + \frac{1}{2} \ln(2\pi) + \sum_{n=1}^{\infty} \frac{B_{2n}}{2n(2n-1)x^{2n-1}}, \quad x \rightarrow +\infty$$

where B_k are the Bernoulli numbers.

(The asymptotic expansions in the examples above are the formal sums following the “ $\sim$ ” sign, the meaning of which will be explained shortly.)

Examples of expansions that are *not* asymptotic expansions are

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} \quad (x \rightarrow +\infty)$$

which converges to $\exp(x)$, but it is not an asymptotic series for large x since it fails (49.7); another example is the series

$$\sum_{k=0}^{\infty} \frac{x^{-k}}{k!} + e^{-x} \quad (x \rightarrow +\infty) \quad (49.9)$$

²⁴That is, there are no convergence requirements. More precisely, they are defined as sequences $\{f_k\}_{k \in \mathbb{N} \cup \{0\}}$, the operations being defined in the same way as if they represented convergent series; see also §49.2.

(because of the exponential terms, this is not an ordered *simple series* satisfying (49.7)). Note however expansion (49.9), *does* satisfies all requirements in the *left* half plane, if we write e^{-x} in the first position.

We also note that in this particular case the first series is convergent, and if we replace (49.9) by

$$e^{1/x} + e^{-x} \quad (49.10)$$

then (49.10) is a valid asymptotic expansion, of a very simple kind, with two nonzero terms. Since convergence is relative to a topology, this elementary remark will play a crucial role when we will speak of Borel summation.

Functions asymptotic to a series, in the sense of Poincaré. The relation $f \sim \tilde{f}$ between an actual function and a formal expansion is defined as a sequence of limits:

Definition 310

A function f is asymptotic to the formal series $\tilde{f}$ as $t \rightarrow t_0^+$ if

$$f(t) - \sum_{k=0}^N \tilde{f}_k(t) =: f(t) - \tilde{f}^{[N]}(t) = o(\tilde{f}_N(t)) \quad (\forall N \in \mathbb{N}) \quad (49.11)$$

We note that condition (49.11) can then be also written as

$$f(t) - \sum_{k=0}^N \tilde{f}_k(t) = O(\tilde{f}_{N+1}(t)) \quad (\forall N \in \mathbb{N}) \quad (49.12)$$

where $g(t) = O(h(t))$ means $\limsup_{t \rightarrow t_0^+} |g(t)/h(t)| < \infty$. Indeed, this follows from (49.11) and the fact that $f(t) - \sum_{k=0}^{N+1} \tilde{f}_k(t) = o(\tilde{f}_{N+1}(t))$.

49.2 Asymptotic power series

In many instances the functions f_k are exponentials, powers and logarithms. This is not simply a matter of choice or an accident, but reflects some important fact about the relation between asymptotic expansions and functions which will be clarified later.

A special role is played by power series which are series of the form

$$\tilde{S} = \sum_{k=0}^{\infty} c_k z^k, \quad z \rightarrow 0^+ \quad (49.13)$$

With the transformation $z = t - t_0$ (or $z = x^{-1}$) the series can be centered at t_0 (or $+\infty$, respectively).

Remark. If a c_k is zero then Definition 310 fails trivially in which case (49.13) is not an asymptotic series. This motivates the following definition.

Definition 311: Asymptotic power series

A function possesses an asymptotic *power* series if

$$f(z) - \sum_{k=0}^N c_k z^k = O(z^{N+1}) \quad (\forall N \in \mathbb{N}) \quad (49.14)$$

We use the boldface notation $\sim$ for the stronger asymptoticity condition in (49.11) when confusion is possible.

Example Check that the Taylor series of an analytic function at zero is its asymptotic series there.

In the sense of (49.14), the asymptotic power series at zero of e^{-1/x^2} is the zero series. It is however surely not the case that e^{-1/x^2} behaves like zero as $x \rightarrow 0$ on $\mathbb{R}$. Rather, in this case, the asymptotic *behavior* of e^{-1/x^2} is e^{-1/x^2} itself (only exponentials and powers involved).

Asymptotic power series form an algebra; addition of asymptotic power series is defined in the usual way:

$$A \sum_{k=0}^{\infty} c_k z^k + B \sum_{k=0}^{\infty} c'_k z^k = \sum_{k=0}^{\infty} (Ac_k + Bc'_k) z^k$$

while multiplication is defined as in the convergent case

$$\left(\sum_{k=0}^{\infty} c_k z^k \right) \left(\sum_{k=0}^{\infty} c'_k z^k \right) = \sum_{k=0}^{\infty} \left(\sum_{j=0}^k c_j c'_{k-j} \right) z^k$$

Remark 312. If the series $\tilde{f}$ is convergent and f is its sum (note the ambiguity of the “sum” notation) $f = \sum_{k=0}^{\infty} c_k z^k$ then $f \sim \tilde{f}$.

The proof of this remark follows directly from the definition of convergence.

Lemma 313. (Uniqueness of the asymptotic series to a function) If $f(z) \sim \tilde{f} = \sum_{k=0}^{\infty} \tilde{f}_k z^k$ as $z \rightarrow 0$ then the $\tilde{f}_k$ are unique.

Proof. Assume that we also have $f(z) \sim \tilde{F} = \sum_{k=0}^{\infty} \tilde{F}_k z^k$. We then have (cf. (49.11))

$$\tilde{F}^{[N]}(z) - \tilde{f}^{[N]}(z) = o(z^N)$$

which is impossible unless $g_N(z) = \tilde{F}^{[N]}(z) - \tilde{f}^{[N]}(z) = 0$, since g_N is a polynomial of degree N in z . ■

Corollary 314. The asymptotic series at the origin of an analytic function is its Taylor series at zero. More generally, if F has a Taylor series at 0 then that series is its asymptotic series as well.

The proof of the following lemma is immediate:

Lemma 315. (Algebraic properties of asymptoticity to a power series) If $f \sim \tilde{f} = \sum_{k=0}^{\infty} c_k z^k$ and $g \sim \tilde{g} = \sum_{k=0}^{\infty} d_k z^k$ then

- (i) $Af + Bg \sim A\tilde{f} + B\tilde{g}$
- (ii) $fg \sim \tilde{f}\tilde{g}$

Sometimes it is convenient to check a formally weaker condition of asymptoticity:

Lemma 316. Let $\tilde{f} = \sum_{n=0}^{\infty} a_n z^n$. If f is such that there exists a sequence $p_n \rightarrow \infty$ such that

$$(\forall n \exists p_n) \text{ s.t. } f(z) - \tilde{f}^{[p_n]}(z) = o(z^n) \text{ as } z \rightarrow 0$$

then $f \sim \tilde{f}$.

Proof. We let m be arbitrary and choose $n > m$ such that $p_n > m$. We have

$$f(z) - \tilde{f}^{[m]} = (f(z) - \tilde{f}^{[p_n]}) + (\tilde{f}^{[p_n]} - \tilde{f}^{[m]}) = o(z^m) \quad (z \rightarrow 0)$$

by assumption and since $\tilde{f}^{[p_n]} - \tilde{f}^{[m]}$ is a polynomial for which the smallest power is z^{m+1} (we are dealing with truncates of the same series). ■

49.3 Integration and differentiation of asymptotic power series.

While asymptotic power series can be safely integrated term by term as the next proposition shows, differentiation is more delicate. In suitable spaces of functions and expansions, we will see the asymmetry largely disappears if we are dealing with analytic functions in suitable regions.

Anyway, for the moment note that the function $e^{-1/z} \sin(e^{1/z^2})$ is asymptotic to the zero power series as $z \rightarrow 0^+$ although the derivative is unbounded and thus not asymptotic to the zero series.

Proposition 49.1. Assume f is integrable near $z = 0$ and that

$$f(z) \sim \tilde{f}(z) = \sum_{k=0}^{\infty} \tilde{f}_k z^k$$

Then

$$\int_0^z f(s) ds \sim \int \tilde{f} := \sum_{k=0}^{\infty} \frac{\tilde{f}_k}{k+1} z^{k+1}$$

Proof. This follows from the fact that $\int_0^z o(s^n) ds = o(z^{n+1})$ as can be seen by immediate estimates. ■

Asymptotic power series of analytic function, if they are valid in wide enough regions can be differentiated.

Asymptotics in a strip. Assume $f(x)$ is analytic in the strip $S_a = \{x : |x| > R, |\operatorname{Im}(x)| < a\}$. Let $\alpha < a$ and and $S_\alpha = \{x : |x| > R, |\operatorname{Im}(x)| < \alpha\}$ and assume that

$$f(x) \sim \tilde{f}(x) = \sum_{k=0}^{\infty} c_k x^{-k} \quad (|x| \rightarrow \infty, x \in S_\alpha) \quad (49.15)$$

It is assumed that that the limits implied in (49.15) hold uniformly in the given strip.

Proposition 49.2. If (49.15) holds, then, for $\alpha' < \alpha$ we have

$$f'(x) \sim \tilde{f}'(x) := \sum_{k=0}^{\infty} -\frac{k c_k}{x^{k+1}} \quad (|x| \rightarrow \infty, x \in S_{\alpha'})$$

Proof. We have $f(x) = \tilde{f}^{[N]}(x) + g_N(x)$ where clearly g is analytic in S_a and $|g_N(x)| \leq \operatorname{Const} \cdot |x|^{-N-1}$ in S_α . But then, for $x \in S_{\alpha'}$ and $\delta = \frac{1}{2}(\alpha - \alpha')$ we get

$$\begin{aligned} |g'_N(x)| &= \frac{1}{2\pi} \left| \oint_{|x-s|=\delta} \frac{g_N(s) ds}{(s-x)^2} \right| \leq \frac{1}{\delta} \frac{\operatorname{Const.}}{(|x| - |\delta|)^{N+1}} \\ &= O(x^{-N-1}) \quad (|x| \rightarrow \infty, x \in S_{\alpha'}) \end{aligned}$$

By Lemma 316, the proof follows. ■

Exercise 317. ** Show that if $f(x)$ is continuous on $[0, 1]$ and differentiable on $(0, 1)$ and $f'(x) \rightarrow L$ as $x \downarrow 0$, then f is differentiable to the right at zero and this derivative equals L . Use this fact, Proposition 49.2 and induction to show that the Taylor series at the origin of $F(z)$ is indeed given by (49.5).

49.4 Lebesgue's Dominated Convergence Theorem

Let $\{f_n\}_{n \in \mathbb{N}}$ be a sequence of complex-valued measurable functions on a measure space (S, Σ, μ) . (Frequently, in our applications the measure is of the form $d\mu = h(x) dx$ where h is some “nice” weight function.)

Suppose that the sequence converges pointwise to a function f , i.e.

$$\lim_{n \rightarrow \infty} f_n(x) = f(x), \quad \forall x \in S.$$

Assume moreover that the sequence $\{f_n\}$ is dominated by some integrable function $g \in L^1(S, \mu)$, i.e.

$$|f_n(x)| \leq |g|(x), \quad \forall x \in S, \forall n,$$

with $\int_S |g| d\mu < \infty$.

Then $f_n, f \in L^1(S, \mu)$ and

$$\lim_{n \rightarrow \infty} \int_S f_n d\mu = \int_S \lim_{n \rightarrow \infty} f_n d\mu = \int_S f d\mu.$$

In fact, more is true:

$$\lim_{n \rightarrow \infty} \int_S |f_n - f| d\mu = 0.$$

49.5 Fubini's theorem

Theorem 318: Fubini

(i) If a function is Lebesgue integrable on a rectangle $X \times Y$ then one the double integral equals the iterated integrals:

$$\iint_{X \times Y} f(x, y) d(x, y) = \int_X \left(\int_Y f(x, y) dy \right) dx = \int_Y \left(\int_X f(x, y) dx \right) dy.$$

(ii) Assume X and Y are σ -finite measure spaces and $X \times Y$ is endowed with the product measure (which is unique as X and Y are σ -). Assume f is a measurable function and

$$\int_{X \times Y} |f(x, y)| d(x, y) < \infty,$$

then

$$\int_X \left(\int_Y f(x, y) dy \right) dx = \int_Y \left(\int_X f(x, y) dx \right) dy = \int_{X \times Y} f(x, y) d(x, y).$$

(iii) Under the assumptions in (ii) on X, Y, σ and f ,

$$\int_X \left(\int_Y |f(x, y)| dy \right) dx = \int_Y \left(\int_X |f(x, y)| dx \right) dy = \int_{X \times Y} |f(x, y)| d(x, y)$$

49.6 Appendix to Chapter 8

49.7 Some facts about the topology of $\mathbb{C}$

From a topological point of view, $\mathbb{C}$ is isomorphic to $\mathbb{R}^2$: we identify the point $z = x + iy$ with $(x, y) = (\operatorname{Re} z, \operatorname{Im} z)$ and the length $|z|$ with $|(x, y)| = \sqrt{x^2 + y^2}$. This induces the metric

$$d(z_1, z_2) := |z_1 - z_2|.$$

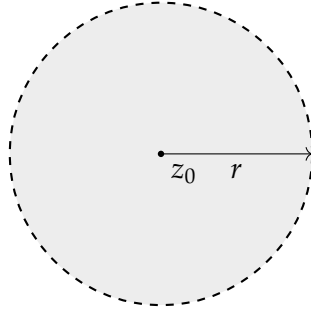


Figure 61: Open disk $\mathbb{D}(z_0, r) = \{z : |z - z_0| < r\}$: boundary dashed to indicate exclusion.

A region $O \subset \mathbb{C}$ is *open* if with each $z_0 \in O$ it contains some nonempty disk centered at z_0 . Examples include the open disk $\mathbb{D}(z_0, r)$, the punctured disk $\mathbb{D}_p(z_0, r)$, the upper half-plane $\mathbb{H}^u$, $\mathbb{C}$ itself, and $\emptyset$. By contrast, the *closed disk*

$$\overline{\mathbb{D}}(z_0, r) := \{z \in \mathbb{C} : |z - z_0| \leq r\}$$

is not open.

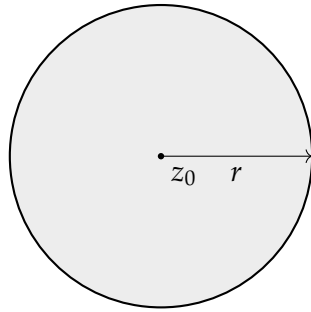


Figure 62: Closed disk $\overline{\mathbb{D}}(z_0, r)$: solid boundary indicates inclusion.

Finite intersections of open sets are open, and arbitrary unions (finite or infinite) of open sets are open.

Definition 319

A **topology** on a set X is a family $\mathcal{O}$ of subsets, called open sets, such that: (1) $X, \emptyset \in \mathcal{O}$; (2) $O_1, O_2 \in \mathcal{O} \implies O_1 \cap O_2 \in \mathcal{O}$; (3) Arbitrary unions of sets in $\mathcal{O}$ are also in $\mathcal{O}$.

Complements of open sets are called *closed* sets. Note that X and $\emptyset$ are both open and closed. For $\mathbb{C}$, the topology can be taken as the collection of all unions of open disks.

The *boundary* of $S \subset \mathbb{C}$, denoted ∂S , consists of points which can be approached both from within S and from outside S . For example,

$$\partial \mathbb{D}(z_0, r) = \{z \in \mathbb{C} : |z - z_0| = r\}.$$

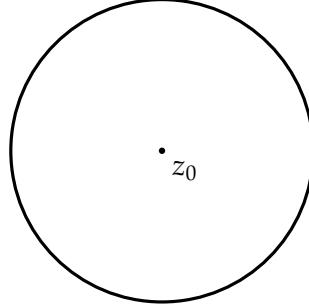


Figure 63: Boundary of the disk: the circle $|z - z_0| = r$.

Since

$$\max(|x|, |y|) \leq \sqrt{x^2 + y^2} \leq |x| + |y|,$$

a sequence $\{z_n\}$ is Cauchy in $\mathbb{C}$ iff both $\{x_n\}$ and $\{y_n\}$ are Cauchy in $\mathbb{R}$. Similarly, $z_n \rightarrow 0$ in $\mathbb{C}$ iff $x_n \rightarrow 0$ and $y_n \rightarrow 0$ in $\mathbb{R}$. In $\mathbb{R}^n$, Cauchy sequences are exactly the convergent ones, hence $\mathbb{C}$ is complete.

Consequently, if a series $\sum_{n=0}^{\infty} c_n z_0^n$ converges, then $c_n z_0^n \rightarrow 0$ as $n \rightarrow \infty$. This implies that the real-valued series $\sum_{n=0}^{\infty} |c_n| |z_1|^n$ converges for every $|z_1| < |z_0|$, i.e. $\sum_{n=0}^{\infty} c_n z_1^n$ converges uniformly and absolutely for $|z| < |z_0|$. This is Abel's theorem in the complex setting: the domain of convergence is a disk rather than an interval.

A region is *closed* iff it contains all of its boundary points; it is *open* iff it contains none of them.

49.7.1 Connected sets

An open set $O \subset \mathbb{C}$ is *connected* if it is not the union of two disjoint nonempty open sets. More generally, $X_1 \subset X$ is connected if it is not the disjoint union of two nonempty sets open in the induced topology. Equivalently, there is no proper subset of X_1 that is simultaneously open and closed in the induced topology. For example, any disk or punctured disk in $\mathbb{C}$ is connected.

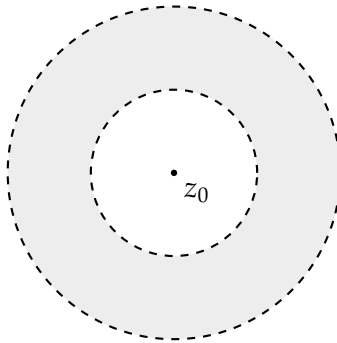


Figure 64: Annulus $\{z : r < |z - z_0| < R\}$: region between two excluded boundary circles.

Exercise 320. Is the annulus $\{z \in \mathbb{C} : r \leq |z| < R\}$ open? closed? connected? What is its boundary?

Proposition 49.3. Domains $\mathcal{D} \subset \mathbb{C}$ are path connected. Moreover, the connecting path can be chosen polygonal.

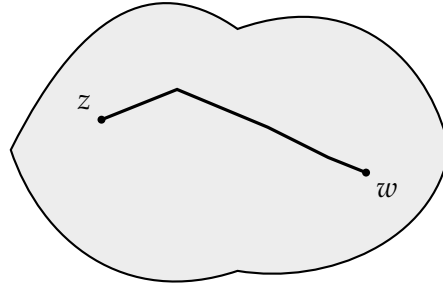


Figure 65: Any two points in a domain can be joined by a polygonal path (Proposition 49.3).

50 Soap films and the minimal surface equation

Let $\Gamma \subset \mathbb{R}^3$ be a simple smooth closed curve (the wire). An equilibrium soap film spanning Γ minimizes the area among admissible surfaces with boundary Γ (Plateau's problem). The Euler–Lagrange equation of the area functional says that the *mean curvature* vanishes:

$$H \equiv 0 \quad \text{on the film.}$$

Equivalently, the *mean curvature vector* $\mathbf{H} = H \mathbf{n}$ is zero.

Graph formulation. If the film can be written locally as a graph $z = u(x, y)$ over a planar domain $\Omega \subset \mathbb{R}^2$, then u solves the *minimal surface equation*

$$\operatorname{div} \left(\frac{\nabla u}{\sqrt{1 + |\nabla u|^2}} \right) = 0 \quad \text{in } \Omega, \quad u|_{\partial\Omega} = u_{\partial\Omega}, \quad (50.2)$$

where the boundary data $u_{\partial\Omega}$ encodes the wire (after projection to the plane).

Parametric (intrinsic) formulation. If $X(u, v) : \Sigma \rightarrow \mathbb{R}^3$ is a smooth immersion with $X(\partial\Sigma) = \Gamma$, the area functional is

$$\mathcal{A}[X] = \iint_{\Sigma} |X_u \times X_v| \, du \, dv.$$

Critical points satisfy $H \equiv 0$. In *isothermal (conformal) parameters* one has the equivalent harmonic map form

$$\Delta X = 0 \quad \text{on } \Sigma, \quad X|_{\partial\Sigma} = \Gamma,$$

where Δ is the flat Laplacian in the (u, v) -coordinates (the conformality constraints ensure $H = 0$).

Remark (bubbles vs. films). A *soap film* has negligible pressure difference across it, hence $H = 0$. A *soap bubble* encloses a volume and has constant mean curvature:

$$H = \frac{\Delta p}{2\sigma} \quad (\text{Laplace–Young law}),$$

with surface tension σ and pressure jump Δp .

Claim. For the classically normalized Legendre polynomials P_ℓ with $P_\ell(1) = 1$,

$$|P_\ell(x)| < 1 \quad \text{for all } |x| < 1, \quad \text{and} \quad \max_{x \in [-1, 1]} |P_\ell(x)| = |P_\ell(\pm 1)| = 1.$$

Proof (via Schläfli's integral representation). Write $x = \cos \theta$ with $\theta \in [0, \pi]$. A standard integral representation (see, e.g., Szegő) is

$$P_\ell(\cos \theta) = \frac{1}{\pi} \int_0^\pi (\cos \theta + i \sin \theta \cos t)^\ell dt.$$

Let $a = \cos \theta$ and $b = \sin \theta$. For $t \in [0, \pi]$,

$$r(t) := |a + ib \cos t| = \sqrt{a^2 + b^2 \cos^2 t} = \sqrt{\cos^2 \theta + \sin^2 \theta \cos^2 t} = \sqrt{1 - \sin^2 \theta \sin^2 t}.$$

If $0 < \theta < \pi$ then $0 < b = \sin \theta < 1$, and hence $r(t) < 1$ for all $t \in (0, \pi)$ (indeed, $r(t) = 1$ only at the endpoints $t = 0, \pi$). Thus,

$$|P_\ell(\cos \theta)| \leq \frac{1}{\pi} \int_0^\pi r(t)^\ell dt < \frac{1}{\pi} \int_0^\pi 1 dt = 1,$$

because the integrand is strictly less than 1 on a set of positive measure. By continuity and the normalization,

$$P_\ell(1) = 1, \quad P_\ell(-1) = (-1)^\ell,$$

so the maximum of $|P_\ell|$ on $[-1, 1]$ is attained at the endpoints and equals 1. ■

Remarks. 1) The bound is sharp and reflects the normalization $P_\ell(1) = 1$. 2) For the L^2 -normalized polynomials $\tilde{P}_\ell(x) = \sqrt{\frac{2\ell+1}{2}} P_\ell(x)$ one has $\|\tilde{P}_\ell\|_{L^\infty([-1,1])} \asymp \ell^{1/2}$, consistent with the endpoint behavior.

References

- [1] Mark J. Ablowitz and Athanassios S. Fokas, *Complex variables*, 2nd edition, Cambridge University Press, 2003.
- [2] M Abramowitz, I A Stegun, *Handbook of mathematical functions with formulas, graphs, and mathematical tables* New York : Wiley-Interscience (1984).
- [3] L. Ahlfors, *Complex Analysis*, McGraw-Hill (1979).
- [4] E.A. Coddington and N. Levinson, *Theory of Ordinary Differential Equations*, McGraw-Hill, New York, (1955). (2009).
- [5] O. Forster, *Lectures on Riemann Surfaces*, Springer New York (1981).
- [6] Hatcher, Allen, *Algebraic Topology*, Cambridge Univ. Press (2002).
- [7] Krantz, Steven G., *Geometric function theory: explorations in complex analysis*, Birkhäuser, (2006)
- [8] J. Milnor, *Complex Dynamics*, Annals of Mathematics Studies, Third Edition, (2006).
- [9] O. Costin and M. Huang, *Geometric construction and analytic representation of Julia sets of polynomial maps*, Nonlinearity 24, 1311–1327 (2011)

- [10] Rudin, Real and Complex Analysis
- [11] W. Schlag, A Course in Complex Analysis and Riemann Surfaces, Graduate Studies in Mathematics vol. 154, American Mathematical Society, 2014.
- [12] https://en.wikipedia.org/wiki/Legendre_polynomials
- [13] S. Lang, Complex Analysis, fourth edition, (1999).
- [14] N.I. Muskhelishvili, Singular integral equations, Dover (2008).
- [15] Z. Nehari, Conformal Mapping, Dover (1975).
- [16] E. C. Titchmarsh, The theory of functions, Oxford University Press, 1964.
- [17]
- [18] Alexander Hulpke - Own work based on: ModularGroup-FundamentalDomain-01.png,
<https://commons.wikimedia.org/w/index.php?curid=59963451>
- [19] G. Pólya & G. Szegő, Problems and Theorems in Analysis II –see Problem IV.4.8 and its discussion.