

Internet Supplement for
Basic Complex Analysis
Third Edition

Jerrold E. Marsden and Michael J. Hoffman

November 3, 1998

Preface

This document consists of a series of supplements to the third edition of our book *Basic Complex Analysis*. Some of the topics give additional technical details of results that were stated but not proved in the textbook while others treat additional topics and applications that may be of interest to some readers.

Pasadena, CA
Fall, 1998

Jerry Marsden and Mike Hoffman

Contents

Preface	ii
1 Analytic Functions	1
2 Cauchy's Theorem	7
3 Series Representation of Analytic Functions	15
4 Calculus of Residues	21
5 Conformal Mappings	27
6 Further Development of the Theory	29
7 Asymptotic Methods	39
8 Laplace Transform and Applications	49

Chapter 1

Analytic Functions

Uniqueness of the Complex Numbers

In the text we constructed the field \mathbb{C} of complex numbers, which contains the reals and in which every quadratic equation has a solution. It is only natural to ask whether there are any other such fields. We shall address this question somewhat informally. A precise explanation would be tantamount to a short course in abstract algebra. However, the student should nevertheless be able to grasp the important points.

The answer is that the complex numbers are the *smallest* field containing \mathbb{R} in which all quadratic equations are solvable and any two such fields are “the same”. In that sense it is unique. The reason for this is quite simple. Let F be a field containing \mathbb{R} and in which quadratic equations are solvable. Let j be any solution in F to the equation $z^2 + 1 = 0$. Consider, in F , all numbers of the form $a + jb$ for real numbers a and b . This set is, algebraically, “the same” as \mathbb{C} , because of the simple fact that since $j^2 = -1$, j plays the role of and may be identified with i . We must also check that $a + jb = c + jd$ implies that $a = c$ and $b = d$ to be sure that equality in this set coincides with that in \mathbb{C} . Indeed, $(a - c) + j(b - d) = 0$, so we must prove that $e + if = 0$ implies that $e = 0$ and $f = 0$ (where $e = a - c$ and $f = b - d$). If $f = 0$, then clearly $e = 0$ as well. But if $f \neq 0$, then $j = -e/f$, which is real. However, no *real* number satisfies $j^2 = -1$ because the square of any real number is nonnegative. Therefore, f must be zero. This proves our claim. We can rephrase our result by saying that \mathbb{C} is the *smallest field extension of \mathbb{R} in which all quadratic equations are solvable*.¹

¹Another question arises at this point. We made \mathbb{R}^2 into a field. For what other n can \mathbb{R}^n be made into a field? Let us demand at the outset that the algebraic operations agree with those on \mathbb{R} , assuming that \mathbb{R} is the x axis. The answer is, only in the case $n = 2$. A fieldlike structure, called the *quaternions*, can be obtained for $n = 4$, except that the rule $zw = wz$ fails. Such a structure is called a *noncommutative field*. The proof of these facts can be found in an advanced abstract algebra text.

Solution of a Cubic Equation

Complex numbers are often introduced by appealing to the quadratic formula,

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a},$$

which supplies roots for the quadratic polynomial

$$ax^2 + bx + c = 0.$$

The text points out how leads to solutions such as $-1 \pm \sqrt{-3}$ to the equation $x^2 + x + 1$, which has no real number solutions. This is done by creating “imaginary” square roots for negative numbers, and then all such quadratic polynomials will have roots. This is certainly very elegant and may be aesthetically pleasing, but it may be less clear that anything really important has actually occurred. After all, the only situations in which one needs this extension to complex numbers are those for which there are no real solutions anyway. Even geometrically the graph of $y = x^2 + x + 1 = (x + (1/2))^2 + (3/4)$ is a parabola that never crosses the x -axis.

The situation for cubic, or third degree, equations may be more striking. Every cubic polynomial

$$y = x^3 + Ax^2 + Bx + C$$

with real coefficients must have at least one real root and perhaps as many as three. Its graph must go up for large positive x and down for large negative x . By continuity it must cross the axis at least once somewhere in between. The fact that complex numbers are deeply involved in an effective way of finding these solutions may make it clearer that there is something meaningful and important going on with them.

The solution to the cubic equation

$$x^3 + Ax^2 + Bx + C = 0 \tag{1}$$

was discovered by Scipione del Ferro and Niccolo Tartaglia in the 1500s and published by Giordano Cardano in 1545. To see how to get at the solution, consider how the coefficients of the equation are related to its roots.

$$(x - \alpha)(x - \beta)(x - \gamma) = x^3 - (\alpha + \beta + \gamma)x^2 + (\alpha\beta + \alpha\gamma + \beta\gamma)x - \alpha\beta\gamma.$$

That is,

$$A = -(\alpha + \beta + \gamma), \quad B = \alpha\beta + \alpha\gamma + \beta\gamma, \quad \text{and} \quad C = -\alpha\beta\gamma.$$

If we make the change of variables $t = x + A/3$, then the corresponding roots are

$$t = \alpha + \frac{A}{3}, \beta + \frac{A}{3}, \quad \text{and} \quad \gamma + \frac{A}{3}.$$

Since these add up to 0, there should be no t^2 term in the transformed equation. Indeed, substitution of $x = t - A/3$ into (1) produces

$$t^3 + \left(B - \frac{A^2}{3}\right)t + \left(C - \frac{AB}{3} + \frac{2A^3}{27}\right) = 0.$$

Thus we really need to be able to solve only the more special cubic equations of the form

$$t^3 + pt + q = 0. \quad (2)$$

For this we need some cleverness.

Letting $t = u + v$, equation (2) becomes

$$(u^3 + v^3) + (3uv + p)(u + v) + q = 0.$$

We will have a solution if we can pick u and v with

$$u^3 + v^3 = -q \quad \text{and} \quad 3uv = -p. \quad (3)$$

To solve (3), we begin by eliminating v . Since $v = -p/3u$, we have

$$u^3 - \frac{p^3}{27u^3} = -q$$

or

$$(u^3)^2 + (u^3)q - \frac{p^3}{27} = 0.$$

Solving this quadratic for u^3 gives

$$u^3 = -\frac{q}{2} \pm \frac{1}{2}\sqrt{q^2 + \frac{4p^3}{27}} = -\frac{q}{2} \pm \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}.$$

Thus

$$v^3 = -q - u^3 = -\frac{q}{2} \mp \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}.$$

We can take

$$u = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} \quad \text{and} \quad v = \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

This gives a solution

$$t = \sqrt[3]{-\frac{q}{2} + \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}} + \sqrt[3]{-\frac{q}{2} - \sqrt{\frac{q^2}{4} + \frac{p^3}{27}}}.$$

This looks innocent enough, but there are a few interesting things lurking under all those radicals. First, a cubic with real coefficients must have at least one real

root. What happens if the quantity under the inner square root signs is negative? Second, every cubic has three solutions (if we count multiplicity and allow complex roots), and all three could be real. Where are the other two roots hiding? In the algebra section of *Standard Mathematical Tables* published by the Chemical Rubber Company, the roots are given as

$$u + v \quad ; \quad -\frac{u+v}{2} + \frac{u-v}{2}\sqrt{-3} \quad ; \quad -\frac{u+v}{2} - \frac{u-v}{2}\sqrt{-3}.$$

Where do these formulas come from?

Every real number has a real cube root, but, like every nonzero complex number, it actually has three complex cube roots. They are distributed at 120° intervals around a circle whose radius is the real cube root of the absolute value of the number. Thus there are three possibilities for each of the cube roots, u, u', u'' and v, v', v'' . These must be combined in appropriate pairs guided by the relation $3uv = -p$ to get the three roots.

It is instructive to consider an example with three real roots. The equation

$$0 = (t-3)(t+1)(t+2) = t^3 - 7t - 6$$

has roots $-1, -2,$ and 3 . Here we have $p = -7$ and $q = -6$. Our equations become $u^3 + v^3 = 6$ and $3uv = 7$. We find

$$\begin{aligned} u^3 &= 3 + \sqrt{9 - \frac{343}{27}} = 3 + \sqrt{-\frac{100}{27}} = 3 + \frac{10\sqrt{3}}{9}i \approx 3 + 1.9245i \\ v^3 &= 3 - \sqrt{9 - \frac{343}{27}} = 3 - \sqrt{-\frac{100}{27}} = 3 - \frac{10\sqrt{3}}{9}i \approx 3 - 1.9245i. \end{aligned}$$

We have

$$|u^3|^2 = |v^3|^2 = 9 + \frac{11}{27} = \frac{343}{27} \approx 12.7037,$$

so

$$|u| = |v| \approx 1.5275.$$

Also

$$\begin{aligned} \arg(u^3) &= \tan^{-1}\left(\frac{10\sqrt{3}}{27}\right) \approx 0.57\text{rad} \approx 32.68^\circ \\ \arg(v^3) &= -\arg(u^3) \\ \arg(u) &= \frac{1}{3}\arg(u^3) \approx 10.89^\circ \pm 120^\circ \\ \arg(v) &= -\frac{1}{3}\arg(u^3) \approx -10.89^\circ \pm 120^\circ. \end{aligned}$$

We get one pair from

$$\begin{aligned} \operatorname{Re}(u) &= |u| \cos(\arg(u)) \approx 1.5 & \operatorname{Im}(u) &= |u| \sin(\arg(u)) \approx 0.28867 \\ \operatorname{Re}(v) &= |v| \cos(\arg(v)) \approx 1.5 & \operatorname{Im}(v) &= |v| \sin(\arg(v)) \approx -0.28867. \end{aligned}$$

This pair gives the root $u + v = 3$. These and the other two pairs are plotted in Figure 1.1. Notice that none of these are real. Nevertheless, when combined in the proper pairs, they produce the three real roots, -1 , -1 , and 3 , of our equation. This ability of complex numbers to produce the real roots for a polynomial equation with real coefficients was more convincing to many that there really was something important going on here than were the purely formal complex solutions to quadratics which did not have any real roots anyway.

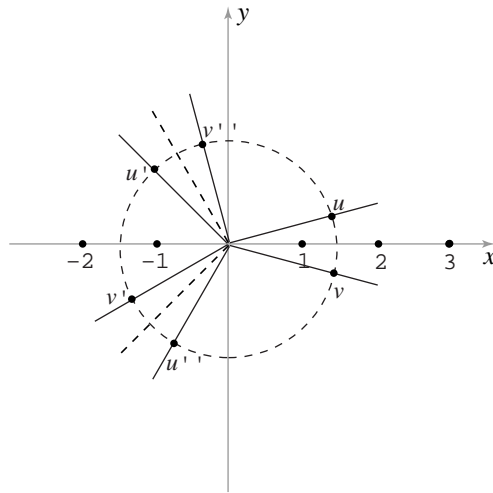


Figure 1.1: Solution of a cubic polynomial.

Chapter 2

Cauchy's Theorem

Riemann Sums

The theory of complex contour integrals can be based directly on a definition in terms of approximation by Riemann sums, as in calculus. If γ is a curve from a to b in the complex plane and f is a function defined along γ , we can choose intermediate points $a = z_0, z_1, z_2, \dots, z_{n-1}, z_n = b$ on γ and form the sum

$$\sum_{k=1}^n f(z_k)(z_k - z_{k-1})$$

(see Figure 2.1). As in calculus, if these sums approach a limit as the maximum of the oriented arc length from z_{k-1} to z_k tends toward 0, we take that limit to be the value of the integral $\int_{\gamma} f(z)dz$.

The properties of the integral given in Proposition 2.1.3 follow from this approach much as the corresponding properties in real-variable calculus. To see that this leads to the same result as Definition 2.1.1 when γ is a C^1 curve, suppose that $z(t) = u(t) + iv(t)$ is a continuously differentiable parametrization of γ with $z(t_k) = z_k$. The mean value theorem guarantees numbers t'_k and t''_k between t_{k-1} and t_k such that $z_k - z_{k-1} = [u'(t'_k) + iv'(t''_k)](t_k - t_{k-1})$. Thus, the Riemann sums $\sum f(z_k)(z_k - z_{k-1})$ correspond to Riemann sums for $\int f(\gamma(t))\gamma'(t)dt$ after sorting out real and imaginary parts.

This approach to the integral allows the use of more general curves and is sometimes useful in writing approximations to the integral. For example, Proposition 2.1.6 may be established by using the triangle inequality: for any approximating Riemann sum, we have

$$\begin{aligned} \left| \sum f(z_k)(z_k - z_{k-1}) \right| &\leq \sum |f(z_k)||z_k - z_{k-1}| \\ &\leq M \sum |z_k - z_{k-1}| \\ &\leq Ml(\gamma). \end{aligned}$$

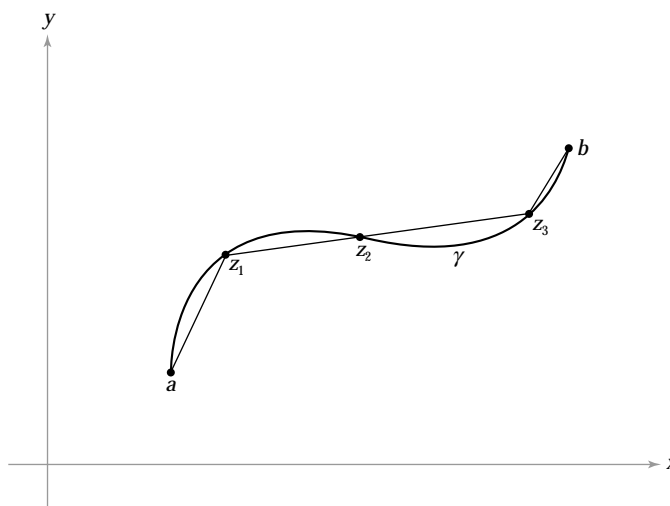


Figure 2.1: A polygonal approximation of γ .

The last step uses the fact that $|z_k - z_{k-1}|$ is the length of the line segment from z_{k-1} to z_k , which is no greater than the distance between them along γ . Since the estimate holds for each approximating sum, it must hold for the integral, which is their limit.

A More General Definition of the Integral

The topics here are separated from the body of the section because they are essential to neither an understanding of Cauchy's Theorem nor of the material in subsequent chapters. We supply the material promised in the text's discussion of the Deformation Theorem. The Smooth Deformation Theorem is used to show how the integral of an analytic function may be defined along a curve which is continuous but not necessarily piecewise C^1 . This definition and the Smooth Deformation Theorem itself are used to finish the proof of the Deformation Theorem. Following this, we explore, without proof, the relationship of Cauchy's Theorem to a geometric result known as the Jordan Curve Theorem, which discusses what we mean by the inside and outside of a simple continuous closed curve.

Integrals along Continuous Curves In the proof of the Deformation Theorem, that is, the homotopy version of Cauchy's Theorem, we made the provisional assumption that the deformation was smooth in the sense that each intermediate curve $\gamma_s(t) = H(s, t)$ and each cross curve $\lambda_t(s) = H(s, t)$, thought of as curves traced out by the point $H(s, t)$ as either s or t , respectively, is held constant, are piecewise C^1 . It was stated that the condition of being C^1 is not actually neces-

sary. We really need only to assume that $H(s, t)$ is a continuous function of s and t (which implies that each $\gamma_s(t)$ is a continuous curve).

For the time being we will refer to the theorem with the C^1 assumption as the “Smooth Deformation Theorem.” The main reason for the assumption was that our whole definition of contour integrals was based on piecewise C^1 curves—after all, the derivative of the curve appears explicitly in the definition! In general we do not know what the integral of a function along a curve which is continuous but not piecewise C^1 really is. In fact, such a general theory is not within our grasp. However, the situation is saved by the fact that we are interested not in general functions but in analytic functions. This extra assumption about the function to be integrated makes up for the weaker information about the curve along which it is to be integrated.

The approach taken here to overcome this difficulty may not be the most direct route to the Deformation Theorem, but it has the advantage of showing how we can make sense of the integral of an analytic function along a continuous curve. It also has the interesting feature of using the Smooth Deformation Theorem in the process of showing that the smoothness assumption is not really needed.¹

Suppose f is an analytic function on an open set G and that $\gamma : [0, 1] \rightarrow G$ is a continuous (but not necessarily piecewise C^1) curve from z_0 to z_1 in G . We want to find a reasonable way to define $\int_\gamma f$. The outline of the program is this:

- (i) We know what $\int_\lambda f$ means if λ is a piecewise C^1 curve in G from z_0 to z_1 .
- (ii) We show that there is at least one such λ that is “close to” γ by using the Path-covering Lemma 1.4.24.
- (iii) We show that if λ_0 and λ_1 are two such curves that are “close to” γ , then they are “close to” each other, and we use the Smooth Deformation Theorem to show that $\int_{\lambda_0} f = \int_{\lambda_1} f$.
- (iv) Because of (iii), $\int_\lambda f$ is the same for all the piecewise C^1 curves λ that are “close to” γ with the same endpoints, and we can take that common value as a reasonable definition for $\int_\gamma f$.

To carry out this program, we must first define “close to”. To do this, we define a type of distance between two parametrized curves with the same parameter interval by moving out along both curves, recording at each parameter value t the distance between the corresponding points in the curves and then taking the largest of these distances. This is illustrated in Figure 2.2.

Definition of the “distance” between curves If $\lambda : [0, 1] \rightarrow \mathbb{C}$ and $\gamma : [0, 1] \rightarrow \mathbb{C}$ are parametrized curves in \mathbb{C} , let

$$\text{dist}(\lambda, \gamma) = \max\{|\lambda(t) - \gamma(t)| \text{ such that } 0 \leq t \leq 1\}.$$

¹Many of the ideas here are presented more completely and a bit differently in the paper by R. Redheffer, “The homotopy theorems of function theory,” *American Mathematical Monthly*, **76** (1969), 778–787, and are used there to do several other interesting things.

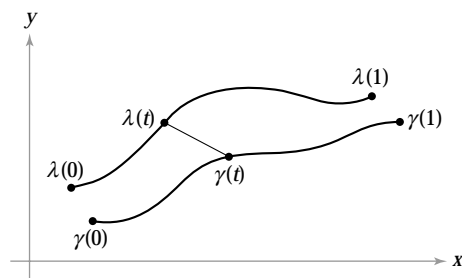


Figure 2.2: A “distance” between parametrized curves.

Now suppose G is an open set in \mathbb{C} and $\gamma : [0, 1] \rightarrow G$ is a continuous curve from z_0 to z_1 in G . By the Distance Lemma 1.4.21, there is a positive distance ρ between the compact image of γ and the closed complement of G , that is, $|\gamma(t) - w| \geq \rho$ for $w \in \mathbb{C} \setminus G$, so $|\gamma(t) - z| < \rho$ implies that z is in G . The Path-covering Lemma 1.4.24 provides a covering of the curve γ by a finite number of disks centered at points $\gamma(t_k)$ along the curve in such a way that each disk is contained in G and each contains the centers of the succeeding and preceding disks. The radius of these disks may be taken to be ρ for purposes of this proof.

We construct a piecewise C^1 curve λ in G by putting $\lambda(t_k) = \gamma(t_k)$ for $k = 0, 1, 2, \dots, n$ and then connecting these points by straight-line segments. More precisely, for $t_{k-1} \leq t \leq t_k$, we put

$$\lambda(t) = \frac{(t - t_{k-1})\lambda(t_k) + (t_k - t)\lambda(t_{k-1})}{t_k - t_{k-1}}.$$

Since the numbers $(t - t_{k-1})/(t_k - t_{k-1})$ and $(t_k - t)/(t_k - t_{k-1})$ are positive and add up to 1, the point $\lambda(t)$ traces out the straight line segment from $\gamma(t_{k-1})$ to $\gamma(t_k)$ as t goes from t_{k-1} to t_k , as in Figure 2.3.

The function $\lambda(t)$ is linear and therefore is a differentiable function of t between t_{k-1} and t_k , so λ is a piecewise C^1 path from z_0 to z_1 . Furthermore, for each t , the points $\lambda(t)$ and $\gamma(t)$ both lie in the disk $D(\gamma(t_{k-1}); \rho)$, so the curve λ lies in the set G and $\text{dist}(\lambda, \gamma) \leq 2\rho$. In fact, since $\lambda(t)$ is on the line between the centers and $\gamma(t)$ is in both disks $D(\gamma(t_{k-1}); \rho)$ and $D(\gamma(t_k); \rho)$, we have $\text{dist}(\lambda, \gamma) \leq \rho$. Since all three sides of the triangle shown have length less than ρ , the distance from $\lambda(t)$ to $\gamma(t)$ is also less than ρ . (See Figure 2.4.)

This gives us the existence of at least one piecewise C^1 path that is “close to” γ . Step (iii) of the program outlined earlier is to show that the integrals along all such paths are the same. Suppose λ_0 and λ_1 are piecewise C^1 paths from z_0 to z_1 such that $\text{dist}(\lambda_0, \gamma) < \rho$ and $\text{dist}(\lambda_1, \gamma) < \rho$. Then both λ_0 and λ_1 lie in G . The Smooth Deformation Theorem can be used to show that $\int_{\lambda_0} f = \int_{\lambda_1} f$. The required homotopy between the two curves can be accomplished by following the straight line from $\lambda_0(t)$ to $\lambda_1(t)$. (See Figure 2.5.) For s and t between 0 and 1,

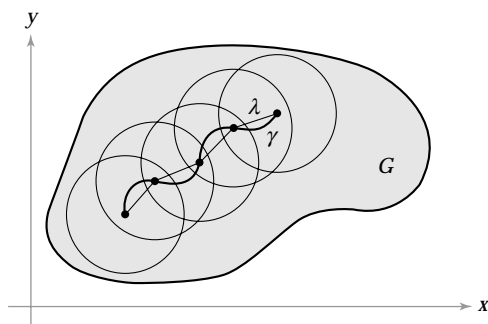


Figure 2.3: A piecewise smooth, in fact linear, approximation to a continuous curve.

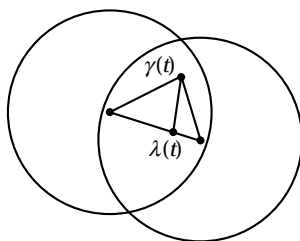


Figure 2.4: $\text{dist}(\lambda, \gamma) < \rho$.

define

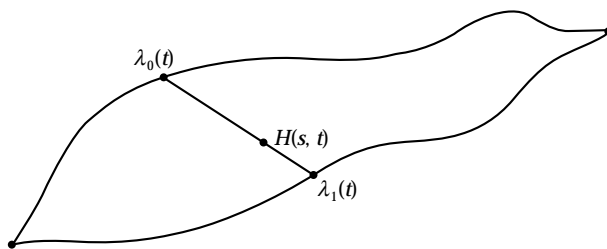
$$H(s, t) = s\lambda_1(t) + (1 - s)\lambda_0(t).$$

The function $H(s, t)$ is a piecewise C^1 function of s and of t . Trouble can occur only when $t = t_k, k = 0, 1, 2, \dots, n$, so we need only check that the image always lies in G . But

$$\begin{aligned} |H(s, t) - \gamma(t)| &= |s\lambda_1(t) + (1 - s)\lambda_0(t) - \gamma(t)| \\ &= |s[\lambda_1(t) - \gamma(t)] + (1 - s)[\lambda_0(t) - \gamma(t)]| \\ &\leq s|\lambda_1(t) - \gamma(t)| + (1 - s)|\lambda_0(t) - \gamma(t)| \\ &\leq s\rho + (1 - s)\rho = \rho. \end{aligned}$$

Thus $H(s, t) \in D(\gamma(t); \rho) \subset G$, so the Smooth Deformation Theorem applies to λ_0 and λ_1 and shows that $\int_{\lambda_0} f = \int_{\lambda_1} f$. This completes step (iii) of the program and shows that it makes sense to define the integral of an analytic function along a continuous curve as follows.

Definition of the integral along continuous curves Suppose that f is analytic on an open set G and that $\gamma : [0, 1] \rightarrow G$ is a continuous curve in G . If

Figure 2.5: Smooth homotopy from λ_0 to λ_1 .

the distance from γ to the complement of G is ρ , let $\int_\gamma f = \int_\lambda f$, where λ is any piecewise C^1 curve in G that has the same endpoints as γ and that is “close to” γ in the sense that $\text{dist}(\lambda, \gamma) < \rho$.

The Deformation Theorem With a bit of care, essentially the same idea used in the proof of step (iii) can be used to obtain the deformation theorem (for both fixed endpoints and closed curves) from the Smooth Deformation Theorem. If H is a continuous homotopy from γ_0 to γ_1 , then for s^* close to s , $\gamma_{s^*}(t)$ is close to $\gamma_s(t)$, so γ_{s^*} is “close to” γ_s . If we choose piecewise C^1 curves λ and μ sufficiently “close to” γ_s and γ_{s^*} , respectively, then λ will be “close to” μ , and following along the short straight-line segment between $\lambda(t)$ and $\mu(t)$ will provide a smooth deformation from λ to μ . (See Figure 2.6.)

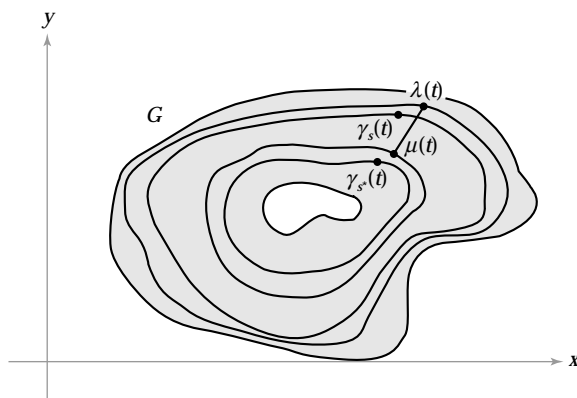


Figure 2.6: The deformation theorems can be obtained from the Smooth Deformation Theorem.

The Smooth Deformation Theorem says that $\int_\lambda f = \int_\mu f$, so the integral along γ_s is the same as that along γ_{s^*} . Thus if we shift s from 0 to 1 in steps sufficiently

small that this argument applies at every step, the integral will never change and the integral along γ_0 will be the same as that along γ_1 . That this actually can be done in a finite number of sufficiently small steps follows because H is a continuous function from the compact square $[0, 1] \times [0, 1]$, so its image is a compact subset of G and lies at a positive distance from the closed complement of G .

The Jordan Curve Theorem

An understanding of the Jordan Curve Theorem is not essential to an understanding of Cauchy's Theorem or of the material in subsequent chapters. However, the Jordan Curve Theorem is closely related to the hypotheses in Cauchy's Theorem, and therefore it is briefly considered here. In many practical examples the result of the Jordan Curve Theorem is geometrically obvious and can usually be proven directly. The general case of the theorem is quite difficult and is not proven here.

Jordan Curve Theorem *Let $\gamma : [a, b] \rightarrow \mathbb{C}$ be a simple closed continuous curve in \mathbb{C} . Then $\mathbb{C} \setminus \gamma([a, b])$ can be written uniquely as the disjoint union of two regions I and O such that I is bounded (that is, lies in some large disk). The region I is called the **inside** of γ and O is called the **outside**. Region I is simply connected and γ is contractible to any point in $I \cup \gamma([a, b])$. The boundary of each of the two regions is $\gamma([a, b])$.*

The proof of this theorem uses more advanced mathematics and is beyond the scope of this book.²

Thus the Jordan Curve Theorem, combined with Cauchy's Theorem, yields the following: *If f is analytic on a region A and γ is a simple closed curve in A and the inside of γ lies in A , then $\int_{\gamma} f = 0$.* This is one classical way of stating Cauchy's Theorem. Although convenient in practice, it is theoretically awkward for two reasons: (1) It depends on the Jordan Curve Theorem for defining the concept of "inside"; (2) γ is restricted to being a simple curve. The versions of Cauchy's Theorem stated in §2.3 do not depend on the difficult Jordan Curve Theorem, are more general, and are just as easy to apply. On the other hand, the Jordan Curve Theorem reassures us that regions we intuitively expect to be simply connected indeed are. (There is another way to describe the inside of a simple closed curve using the index, or winding number, of a curve; this method is discussed in the next section.)

By applying the Jordan Curve Theorem, one can prove that a region is simply connected iff, for every simple closed curve γ in A , the inside of γ also lies in A . This conclusion should seem reasonable. We can also apply the theorem to prove that the inside of a simple closed curve is simply connected.

²See, for example, G. T. Whyburn, *Topological Analysis* (Princeton, N.J.: Princeton University Press, 1964).

The general philosophy of this text is that we should use our geometric intuition to justify that a given region is simply connected or that two curves are homotopic—but with the realization that such knowledge is based on intuition and that to attempt to make it precise could be tedious. On the other hand, a precise argument should be used whenever possible and practical (see, for instance, the argument in the text that a convex region is simply connected).

Chapter 3

Series Representation of Analytic Functions

The following two results illustrate how the Cauchy integral formula can sometimes be used to obtain uniformity where it might not be expected. We begin with some useful terminology.

Definition 3.1 A family of functions \mathcal{S} defined on a set G is said to be **uniformly bounded on closed disks in G** if for each closed disk $B \subset G$ there is a number $M(B)$ such that $|f(z)| \leq M(B)$ for all z in B and for all f in \mathcal{S} .

The word “uniformly” refers to the fact that the constant $M(B)$ does not depend on the particular function used from the family, but may depend on the family \mathcal{S} itself and on the disk B chosen.

Theorem 3.2 If f_1, f_2, f_3, \dots is a sequence of functions analytic on a region G that is uniformly bounded on closed disks in G , then the sequence of derivatives f'_1, f'_2, f'_3, \dots is also uniformly bounded on closed disks in G .

Proof Suppose $B = \{z \text{ such that } |z - z_0| \leq r\}$ is a closed disk in G . Since B is closed and G is open, Worked Example 1.4.27 shows that there is a number ρ with $B \subset D(z_0; \rho) \subset G$. Let $R = (r + \rho)/2$ and $D = \{z \text{ such that } |z - z_0| \leq R\}$. By hypothesis, there is a number $N(D)$ such that $|f_n(z)| \leq N(D)$ for all n and all z in D . If Γ is the boundary circle of D , the Cauchy integral formula for derivatives gives, for any z in B ,

$$|f'_n(z)| = \left| \frac{1}{2\pi i} \int_{\Gamma} \frac{f_n(\zeta)}{(\zeta - z)^2} d\zeta \right| \leq \frac{1}{2\pi} \left[\frac{N(D)}{(R - r)^2} \right] 2\pi R.$$

Thus, if we put $M(B) = N(D)R/(R - r)^2$, we will have $|f'_n(z)| \leq M(B)$ for all n and for all z in B , as desired. ■

Definition 3.3 A family of functions \mathcal{S} defined on a set B is said to be **uniformly equicontinuous on B** if for each $\epsilon > 0$ there is a number $\delta > 0$ such that $|f(\zeta) - f(\xi)| < \epsilon$ for all f in \mathcal{S} whenever ζ and ξ are in B and $|\zeta - \xi| < \delta$. That is, for each ϵ , the same δ can be made to work for all functions in the family \mathcal{S} and everywhere in the set B .

The following shows how one can use Cauchy's Theorem to verify that a given family is uniformly equicontinuous. Why one would want to do so will become clear in the supplementary material for Chapter 6.

Theorem 3.4 Prove that if f_1, f_2, f_3, \dots is a sequence of functions analytic on a region G that is uniformly bounded on closed disks in G , then this family of functions is uniformly equicontinuous on every closed disk in G .

Proof Let B be a closed disk in G . By the last example, there is a number $M(B)$ such that $|f'_n(z)| \leq M(B)$ for every n and for all z in B . Let γ be the straight line from ζ to ξ in B . Since that straight line is contained in B , we have

$$|f_n(\zeta) - f_n(\xi)| = \left| \int_{\gamma} f'_n(z) dz \right| \leq \int_{\gamma} |f'_n(z)| |dz| \leq M(B) |\zeta - \xi|.$$

Thus, given $\epsilon > 0$, we can satisfy the definition of uniform equicontinuity on B by setting $\delta = \epsilon/M(B)$. ■

Some applications of these results are given in the supplementary material for Chapter 6.

Power Series via Hadamard's Formula

In this section, the basic facts about convergence of a power series are proved by directly involving a formula due to Hadamard for the radius of convergence. A sketch is given showing how these ideas can be applied to operator theory, in particular to the spectral radius of a matrix or continuous linear operator. This can be used to link the material here with other analysis courses the student may be taking.

The basic facts about power series can be developed in a way that also supplies a formula for the radius of convergence with the help of a few facts from intermediate analysis. The notion needed is that of the *limit superior* of a sequence of real numbers. If c_1, c_2, c_3, \dots is a sequence in \mathbb{R} , then the largest cluster point (possibly infinite) of the sequence is called the **limit superior**, and the smallest is called the **limit inferior**. More precisely,

$$\begin{aligned} \limsup_{k \rightarrow \infty} c_k &= \lim_{k \rightarrow \infty} [\sup\{c_{k+1}, c_{k+2}, c_{k+3}, \dots\}] \\ \liminf_{k \rightarrow \infty} c_k &= \lim_{k \rightarrow \infty} [\inf\{c_{k+1}, c_{k+2}, c_{k+3}, \dots\}]. \end{aligned}$$

These could be infinite. The facts we need are

- (i) If $B < \limsup_{k \rightarrow \infty} c_k$, then $c_k > B$ for infinitely many values of k .
- (ii) If $B > \limsup_{k \rightarrow \infty} c_k$, then there is an index N such that $c_k < B$ whenever $k \geq N$.

These facts are proved in, for example, J. Marsden and M. Hoffman *Elementary Classical Analysis*, Second Edition (New York: W. H. Freeman and Company, 1993).

To apply this concept, we use the following:

- (iii) If a sequence $f_0, f_1, f_2, f_3, \dots$ of functions with values in a complete space such as \mathbb{R} or \mathbb{C} is uniformly Cauchy on a domain A in the sense that for each ϵ there is an index $N(\epsilon)$ such that

$$|f_{n+p}(z) - f_n(z)| < \epsilon$$

for every z in A whenever $n \geq N(\epsilon)$ and $p > 0$, then there is a function f to which the sequence converges uniformly on A .

- (iv) If a series whose terms are functions on a domain A with values in a complete space such as \mathbb{R} or \mathbb{C} is such that the series of absolute values converges uniformly on A , then the series itself converges uniformly on A .

Fact (iii) is used to obtain (iv) by taking the partial sums of the series as the f_n . They are then used to obtain the *Weierstrass M test*.

Proposition 3.5 *Suppose $g_0, g_1, g_2, g_3, \dots$ are functions on a domain A with values in a complete space such as \mathbb{R} or \mathbb{C} . If there is a sequence of positive constants M_k such that $\sum_{k=0}^{\infty} M_k$ converges and $|g_k(z)| \leq M_k$ for every z in A and for every k , then the series $\sum_{k=0}^{\infty} g_k(z)$ converges uniformly and absolutely on A .*

Proof Let $\epsilon > 0$. The partial sums of the series $\sum M_k$ form a uniformly Cauchy sequence in \mathbb{R} , so there is an index $N(\epsilon)$ such that $\sum_{k=n+1}^{n+p} M_k < \epsilon$ whenever $n \geq N$ and $p > 0$. For such n and for z in A , we have both

$$\left| \sum_{k=0}^{n+p} g_k(z) - \sum_{k=0}^n g_k(z) \right| = \left| \sum_{k=n+1}^{n+p} g_k(z) \right| \leq \sum_{k=n+1}^{n+p} |g_k(z)| \leq \sum_{k=n+1}^{n+p} M_k < \epsilon$$

and

$$\sum_{k=0}^{n+p} |g_k(z)| - \sum_{k=0}^n |g_k(z)| = \sum_{k=n+1}^{n+p} |g_k(z)| \leq \sum_{k=n+1}^{n+p} M_k < \epsilon.$$

The series of absolute values and the series itself are uniformly Cauchy and hence uniformly convergent on the domain A . ■

We are now ready to obtain the fundamental theorem about power series.

Theorem 3.6 Suppose $\sum_{k=0}^{\infty} a_k(z - z_0)^k$ is a power series in a complex variable z with complex coefficients a_k . Let $L = \limsup_{k \rightarrow \infty} |a_k|^{1/k}$ and define R by

$$R = \begin{cases} 0 & \text{if } L = +\infty \\ 1/L & \text{if } 0 < L < +\infty. \\ +\infty & \text{if } L = 0 \end{cases}$$

Then the series converges absolutely if $|z - z_0| < R$ and diverges if $|z - z_0| > R$. Furthermore,

- (i) If $R = 0$, the series converges only for $z = z_0$.
- (ii) If $R = +\infty$, the convergence is absolute and uniform on each closed disk $D_r = \{z \in \mathbb{C} \mid |z - z_0| \leq r\}$ with $0 < r < \infty$.
- (iii) If $0 < R < +\infty$, the convergence is absolute and uniform on each closed disk $D_r = \{z \in \mathbb{C} \mid |z - z_0| \leq r\}$ with $0 < r < R$.

Proof If $z = z_0$, the only non-zero term is a_0 , and the series certainly converges.

Consider divergence first. If $|z - z_0| > R$, we can select a nonzero number ρ with $|z - z_0| > \rho > R$ such that

$$\frac{1}{|z - z_0|} < \frac{1}{\rho} < \limsup_{k \rightarrow \infty} |a_k|^{1/k}.$$

Thus

$$\frac{1}{|z - z_0|} < \frac{1}{\rho} < |a_k|^{1/k}$$

for infinitely many values of k . Taking k th powers and multiplying by $|z - z_0|^k$ gives

$$1 < |a_k(z - z_0)^k|$$

for those values of k . The terms cannot converge to 0 and the series must diverge. This settles the divergence and case (i).

For the general convergence claim and (ii) and (iii), it suffices to show uniform absolute convergence on each disk D_r with $0 < r < R$ since any z with $|z - z_0| < R$ is contained in such a disk. If $|z - z_0| \leq r < R$, we can select a finite nonzero number ρ with $|z - z_0| \leq r < \rho < R$. Then

$$\frac{1}{r} > \frac{1}{\rho} > \limsup_{k \rightarrow \infty} |a_k|^{1/k}.$$

There is an index N such that

$$\frac{1}{r} > \frac{1}{\rho} > |a_k|^{1/k}$$

whenever $k \geq N$. Multiplying by r and taking k th powers gives

$$1 > \frac{r}{\rho} > |a_k|^{1/k} r \geq |a_k|^{1/k} |z - z_0|$$

$$1 > \left(\frac{r}{\rho}\right)^k \geq |a_k| |z - z_0|^k = |a_k(z - z_0)^k|.$$

Since $r/\rho < 1$, the series $\sum_{k=N}^{\infty} (r/\rho)^k$ converges. Thus, $\sum_{k=N}^{\infty} a_k(z - z_0)^k$ converges uniformly and absolutely on D_r by the Weierstrass M test. Adding in the finitely many terms for $k = 0, 1, 2, 3, \dots$ does not change this conclusion. ■

The formula

$$R = \frac{1}{\limsup_{k \rightarrow \infty} |a_k|^{1/k}}$$

for the radius of convergence of the power series $\sum_{k=0}^{\infty} a_k(z - z_0)^k$ is usually called **Hadamard's Formula** after the French mathematician Jacques Hadamard who lived from 1865 to 1963. For any particular series, direct appeal to the ratio test or the root test may well be a more practical way of actually finding the radius of convergence, but this formula is often valuable, particularly as a theoretical tool.

Application in Operator Theory: The Spectral Radius The preceding argument does not really demand that the coefficients a_k be numbers. What is required is something corresponding to the absolute value. This can be the norm or length of a vector much as the absolute value of a complex number is the same as its length as a vector in the plane. We also want the terms of the series to be in a complete space, so the statements about the Cauchy property implying convergence apply. A setting in which this is particularly useful is that of the space of square matrices or of continuous linear operators on a normed vector space. In this setting one wants to know as much as possible about invertibility of the matrices or operators. If T is a matrix or operator, I is the identity matrix or operator, and μ is a number, one studies the invertibility of $\mu I - T$. The set of numbers for which this is not invertible is called the **spectrum** of T , and for square matrices it is the same as the set of eigenvalues. It is tempting to try to use a geometric series. Will this work?

$$(\mu I - T)^{-1} = \frac{1}{\mu} \left(I - \frac{1}{\mu} T \right)^{-1} = \frac{1}{\mu} \left(I + \frac{1}{\mu} T + \frac{1}{\mu^2} T^2 + \frac{1}{\mu^3} T^3 + \dots \right).$$

To study this, one can define the operator norm of a matrix or operator T as the maximum amount it can stretch a unit vector.

$$\|T\| = \sup\{\|Tu\| \mid u \text{ is a unit vector}\}$$

This turns out to have the usual properties of a norm or length for a vector, namely,

1. $\|T\| \geq 0$
2. $\|T\| = 0$ if and only if $T = 0$

3. $\|\alpha T\| = |\alpha| \|T\|$ for every number α

4. $\|S + T\| \leq \|S\| + \|T\|$

in addition to a particularly useful property relating it to products;

5. $\|ST\| \leq \|S\| \cdot \|T\|$

It is not too hard to show that if the series converges with respect to this norm, then it gives the desired inverse. Furthermore, properties 3. and 5. show that

$$\left\| \frac{1}{\mu^k} T^k \right\| \leq \frac{1}{|\mu|^k} \|T^k\| \leq \frac{1}{|\mu|^k} \|T\|^k = \left(\frac{1}{|\mu|} \|T\| \right)^k.$$

Thus, if $|\mu| > \|T\|$, the series of norms converges by comparison to a geometric series of numbers. Once we know that the space of matrices or of operators is complete, we can use the fact that absolute convergence implies convergence in any complete space to conclude that our series converges. However, if we use Hadamard's Formula or the techniques that lead to it, we get a much sharper result. The series converges if

$$|\mu| > \limsup \left(\|T^k\|^{1/k} \right).$$

As a consequence we have the following.

Proposition 3.7 *If μ is an eigenvalue of a square matrix T (or is in the spectrum of a continuous linear operator T), then*

$$|\mu| \leq \limsup \left(\|T^k\|^{1/k} \right).$$

Further use of ideas from complex analysis such as the Laurent series expansion and Liouville's theorem show that this estimate is precise. The number

$$\rho(T) = \limsup \left(\|T^k\|^{1/k} \right)$$

is called the *spectral radius* of T and is equal to the largest absolute value of points in the spectrum.

Chapter 4

Calculus of Residues

Technical Lemma In the text, the following technical lemma was of interest in the evaluation of definite integrals along the whole real line. We provide a proof here.

Lemma 4.1 *If*

$$\lim_{A \rightarrow \infty, B \rightarrow \infty} \int_{-A}^B f(x) dx$$

exists, then

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

exists and is equal to this limit.

Proof To say that the limit exists and is L is to say that for each $\epsilon > 0$ there is an $R(\epsilon)$ such that

$$\left| L - \int_{-A}^B f(x) dx \right| < \epsilon$$

whenever $A \geq R(\epsilon)$ and $B \geq R(\epsilon)$. The assertion is that this implies the independent existence of the two limits $\lim_{A \rightarrow \infty} \int_{-A}^0 f(x) dx$ and $\lim_{B \rightarrow \infty} \int_0^B f(x) dx$. We show how to do the second of these. The first is similar. Notice that if α and β are both larger than $R(\epsilon)$, then

$$\begin{aligned} \left| \int_{\alpha}^{\beta} f(x) dx \right| &= \left| \left(\int_{-R(\epsilon)}^{\beta} f(x) dx - L \right) - \left(\int_{-R(\epsilon)}^{\alpha} f(x) dx - L \right) \right| \\ &\leq \left| L - \int_{-R(\epsilon)}^{\beta} f(x) dx \right| + \left| L - \int_{-R(\epsilon)}^{\alpha} f(x) dx \right| < 2\epsilon. \end{aligned}$$

We use this observation twice. First suppose b_1, b_2, b_3, \dots is any sequence tending to $+\infty$. The observation shows that the integrals $\int_0^{b_k} f(x) dx$ form a Cauchy sequence and must converge to some limit. Our desired conclusion follows if the value of that limit is independent of the particular sequence used. Suppose

$$\begin{aligned} b_1, b_2, b_3, \dots \rightarrow +\infty & \quad \text{and} \quad \int_0^{b_k} f(x) dx \rightarrow L_b \\ \beta_1, \beta_2, \beta_3, \dots \rightarrow +\infty & \quad \text{and} \quad \int_0^{\beta_k} f(x) dx \rightarrow L_\beta. \end{aligned}$$

If k is large enough such that b_k and β_k are both larger than $R(\epsilon)$ and each of the integrals is within ϵ of its respective limit, then

$$\begin{aligned} |L_b - L_\beta| &\leq \left| L_b - \int_0^{b_k} f(x) dx \right| + \left| \int_0^{\beta_k} f(x) dx - L_\beta \right| \\ &\quad + \left| \int_0^{b_k} f(x) dx - \int_0^{\beta_k} f(x) dx \right| \\ &\leq \epsilon + \epsilon + \left| \int_{\beta_k}^{b_k} f(x) dx \right| < 4\epsilon. \end{aligned}$$

Since this is true for every positive ϵ , we must have $L_b = L_\beta$ as required. ■

Fresnel Integrals Next we treat some special integrals that can be evaluated using the methods of contour integrals. These types of integrals are useful in optics.

Example 4.2 (Fresnel Integrals) Show that

$$\int_{-\infty}^{\infty} \cos(x^2) dx \quad \text{and} \quad \int_{-\infty}^{\infty} \sin(x^2) dx$$

both exist and equal $\sqrt{\pi/2}$.

Solution First we show the integrals exist. Observe that $\sin(x^2)$ has zeros at $x_n = \sqrt{\pi n}$ for integers n . Since

$$\sqrt{n+1} - \sqrt{n} = 1/(\sqrt{n+1} + \sqrt{n}),$$

the distance between these zeros shrinks to zero as n increases, so the quantities

$$a_n = \left| \int_{x_{n-1}}^{x_n} \sin(x^2) dx \right|$$

decrease monotonically to 0. Thus, $\sum_0^\infty (-1)^n a_n$ converges by the alternating series test to some number A . If R is any real number, then $x_{N-1} \leq R < x_N$ for a unique N , and $\int_0^R \sin(x^2) dx$ is between the partial sums

$$\sum_0^{N-1} (-1)^n a_n \quad \text{and} \quad \sum_0^N (-1)^n a_n.$$

Thus, $\lim_{R \rightarrow \infty} \int_0^R \sin(x^2) dx$ exists and is equal to A . Similarly, $\lim_{R \rightarrow \infty} \int_0^R \cos(x^2) dx$ exists.

Consider the integral of $f(z) = e^{iz^2} / \sin(\sqrt{\pi}z)$ around the contour

$$\gamma = \text{I} + \text{II} + \text{III} + \text{IV}$$

shown in Figure 4.1.

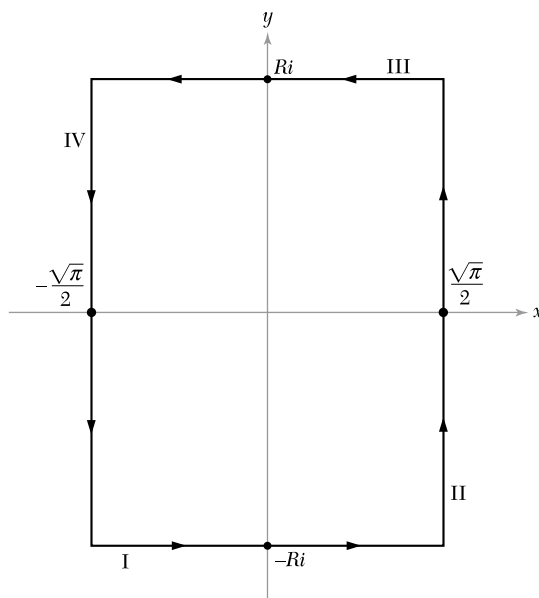


Figure 4.1: Contour used for evaluating the Fresnel integrals.

The function f has a simple pole at 0 inside γ with residue $1/\sqrt{\pi}$, so

$$\int_{\gamma} f = 2\sqrt{\pi}i.$$

Along I, $z = x - Ri$, so

$$|e^{iz^2}| = |e^{i(x^2 - 2Rix - R^2)}| = e^{2Rx}$$

and

$$|\sin \sqrt{\pi}z| = \frac{1}{2} |e^{i\sqrt{\pi}x - R\sqrt{\pi}} - e^{-i\sqrt{\pi}x + R\sqrt{\pi}}| \geq \frac{1}{2} (e^{R\sqrt{\pi}} - 1).$$

Thus, along I we have

$$\left| \int_I f \right| \leq \frac{2}{e^{R\sqrt{\pi}} - 1} \int_{-\sqrt{\pi}/2}^{\sqrt{\pi}/2} e^{2Rx} dx = \frac{1}{R} \frac{e^{R\sqrt{\pi}} - e^{-R\sqrt{\pi}}}{e^{R\sqrt{\pi}} - 1},$$

which goes to 0 as $R \rightarrow \infty$. Similarly

$$\int_{III} f \rightarrow 0 \quad \text{as } R \rightarrow \infty.$$

The contribution from the vertical sides is

$$\begin{aligned} \int_{II} f + \int_{IV} f &= \int_{-R}^R \frac{e^{i(\sqrt{\pi}/2 + iy)^2}}{\sin(\pi/2 + \sqrt{\pi}yi)} idy + \int_R^{-R} \frac{e^{i(-\sqrt{\pi}/2 + iy)^2}}{\sin(-\pi/2 + \sqrt{\pi}yi)} idy \\ &= \int_{-R}^R \frac{e^{i(\pi/4 - y^2)}(e^{-\sqrt{\pi}y} + e^{\sqrt{\pi}y})}{\cos(i\sqrt{\pi}y)} idy = 2i \int_{-R}^R e^{i(\pi/4 - y^2)} dy \\ &= 2e^{3\pi i/4} \int_{-R}^R e^{-iy^2} dy \\ &= \sqrt{2}(-1 + i) \left[\int_{-R}^R \cos(x^2) dx - i \int_{-R}^R \sin(x^2) dx \right]. \end{aligned}$$

Letting $R \rightarrow \infty$, we obtain

$$2\sqrt{\pi}i = \sqrt{2}(-1 + i) \left[\int_{-\infty}^{\infty} \cos(x^2) dx - i \int_{-\infty}^{\infty} \sin(x^2) dx \right]$$

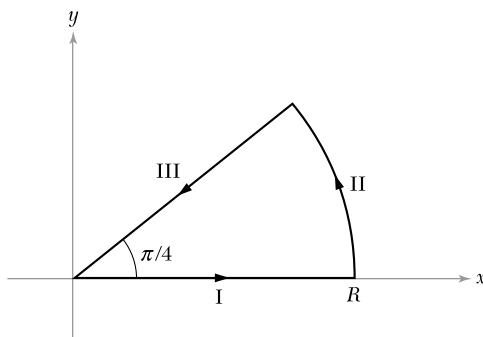
and

$$\sqrt{2\pi}i = \left[- \int_{-\infty}^{\infty} \cos(x^2) dx + \int_{-\infty}^{\infty} \sin(x^2) dx \right] + i \left[\int_{-\infty}^{\infty} \cos(x^2) dx + \int_{-\infty}^{\infty} \sin(x^2) dx \right].$$

The real part of this equation shows that our integrals are equal, while the imaginary part shows that their common value is $\sqrt{2\pi}/2 = \sqrt{\pi}/2$.

Example 4.3 Show that

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi}.$$

Figure 4.2: Contour for $\int_{-\infty}^{\infty} e^{-x^2} dx$.

Solution¹ Let $f(z) = e^{-z^2}$ and consider the integral of f along the contour $\gamma = \text{I} + \text{II} + \text{III}$ shown in Figure 4.2.

Notice that

$$\int_{\text{I}} f = \int_0^R e^{-x^2} dx$$

and

$$\int_{\text{III}} f = \int_R^0 e^{-ir^2} e^{\pi i/4} dr = e^{5\pi i/4} \int_0^R (\cos r^2 - i \sin r^2) dr.$$

Along II, $z = Re^{i\theta}$, so

$$|f(z)| = |e^{-R^2(\cos 2\theta + i \sin 2\theta)}| = e^{-R^2 \cos 2\theta}.$$

But for $0 \leq \theta \leq \pi/4$, we have $\cos 2\theta \geq 1 - 4\theta/\pi$ (see Figure 4.3).

Therefore,

$$|f(z)| \leq e^{-R^2} e^{4R^2\theta/\pi},$$

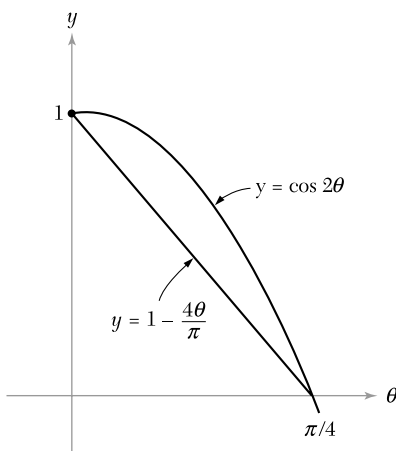
and thus

$$\left| \int_{\text{II}} f \right| \leq \int_0^{\pi/4} e^{-R^2} e^{4R^2\theta/\pi} |iRe^{i\theta}| d\theta = Re^{-R^2} \frac{\pi}{4R^2} (e^{R^2} - 1) = \frac{\pi}{4R} (1 - e^{-R^2}).$$

This goes to 0 as $R \rightarrow \infty$. Since f is entire,

$$0 = \int_{\text{I}} f + \int_{\text{II}} f + \int_{\text{III}} f.$$

¹The method that follows is usually attributed to R. Courant.

Figure 4.3: Proof that $\cos 2\theta \geq 1 - 4\theta/\pi$.

Letting $R \rightarrow \infty$, we obtain

$$0 = \int_0^\infty e^{-x^2} dx - \frac{1+i}{\sqrt{2}} \left[\int_0^\infty \cos(x^2) dx - i \int_0^\infty \sin(x^2) dx \right],$$

since we already know from the last example that both of these integrals exist. Both integrands are even, and by the last example, both integrals equal $\sqrt{\pi}/2\sqrt{2}$. We are left with

$$\int_0^\infty e^{-x^2} dx = \frac{1+i}{\sqrt{2}} (1-i) \frac{\sqrt{\pi}}{2\sqrt{2}} = \frac{\sqrt{\pi}}{2}.$$

Again, the integrand is even, so

$$\int_{-\infty}^\infty e^{-x^2} dx = 2 \int_0^\infty e^{-x^2} dx = \sqrt{\pi}.$$

Chapter 5

Conformal Mappings

The main supplementary material for this chapter is the proof of the Riemann Mapping Theorem. However, the proof we use requires some tools from Chapter 6, so it is deferred to the next chapter.

Chapter 6

Further Development of the Theory

Normal Families and the Riemann Mapping Theorem

The main objective of this supplement is to outline a proof of the Riemann Mapping Theorem. The material is separated from the main text since it is somewhat more advanced than the rest of the chapter and is not needed for understanding or using the theorem in succeeding chapters. However, it does illustrate several powerful tools and techniques of complex analysis.

Throughout this section, G represents a connected, simply connected open set properly contained in the complex plane \mathbb{C} , and D , the open unit disk $D = D(0; 1) = \{z \text{ such that } |z| < 1\}$. Given $z_0 \in \mathbb{C}$, the *Riemann Mapping Theorem* asserts:

There is a function f that is analytic on G and maps G one-to-one onto D with $f(z_0) = 0$. Furthermore, if it is required that $f'(z_0) > 0$, then there is exactly one such function.

The uniqueness has already been established in Chapter 5; that is, there can be no more than one such function. We still need to show there is at least one. The idea of the proof is to look at all the analytic functions that map G one-to-one into D taking z_0 to 0 with positive derivative at z_0 , find one among them that maximizes $f'(z_0)$, and show that this function must take G onto D .

Montel's Theorem on Normal Families The proof of the existence of a function that maximizes $f'(z_0)$ rests on the material of §3.1 concerning uniform convergence on closed disks. We learned there that if a sequence of analytic functions on a region converges uniformly on closed disks contained in the region, then the

limit function must be analytic. The existence of such sequences is addressed by the theorem of Montel on normal families.

Definition 6.1 (Definition of Normal Family) *If A is an open subset of \mathbb{C} , a set \mathcal{S} of functions analytic on A is called a **normal family** if every sequence of functions in \mathcal{S} has a subsequence that converges uniformly on closed disks in A .*

By the Analytic Convergence Theorem, the limit of such a subsequence must be analytic on A .

Theorem 6.2 (Montel's Theorem) *If A is an open subset of \mathbb{C} and \mathcal{S} is a set of functions analytic on A that is uniformly bounded on closed disks in A , then every sequence of functions in \mathcal{S} has a subsequence that converges uniformly on closed disks in A . That is, \mathcal{S} is a normal family.*

Proof The plan of attack is as follows:¹

- (i) Select a countable set of points $C = \{z_1, z_2, z_3, \dots\}$ that are scattered densely throughout A in the sense that $A \subset \text{cl}(C)$.
- (ii) Show that there is a subsequence of the original sequence of functions that converges at all of these points.
- (iii) Show that convergence on this dense set of points is enough to force the subsequence to converge at all points of A .
- (iv) Check that this convergence is uniform on every closed disk in A .

The first step may be accomplished by taking those points whose real and imaginary parts are both rational numbers. There are only countably many of these, so they may be arranged in a sequence, and they are scattered densely in A in the sense that some of them are arbitrarily close to anything in A .

Let f_1, f_2, f_3, \dots be a sequence of functions in \mathcal{S} . The assumption of uniform boundedness on closed disks is that for each closed disk $B \subset A$, there is a number $M(B)$ such that $|f_n(z)| < M(B)$ for all n and for all z in B . In particular, the numbers $f_1(z_1), f_2(z_1), f_3(z_1), \dots$ are all smaller than $M(\{z_1\})$. Thus there must be a subsequence of them that converges to a point w_1 with $|w_1| \leq M(\{z_1\})$. Relabel this subsequence as

$$f_{1,1}(z_1), f_{1,2}(z_1), f_{1,3}(z_1), \dots \rightarrow w_1.$$

Evaluating these functions at z_2 gives another sequence of numbers,

$$f_{1,1}(z_2), f_{1,2}(z_2), f_{1,3}(z_2), \dots,$$

¹The student who has seen the Arzela-Ascoli theorem (see, for example, J. Marsden and M. Hoffmann, *Elementary Classical Analysis*, Second Edition (New York: W. H. Freeman and Company, 1993)) can give a quick proof of Montel's theorem by using the assumed uniform boundedness and Worked Example 3.1.19 of this book to prove equicontinuity.

which are bounded by $M(\{z_2\})$. Some subsequence of these must converge to a point w_2 . Relabel this subsubsequence as

$$f_{2,1}(z_2), f_{2,2}(z_2), f_{2,3}(z_2), \dots \rightarrow w_2.$$

It is important to notice that the functions $f_{2,1}, f_{2,2}, f_{2,3}, \dots$ are selected from among $f_{1,1}, f_{1,2}, f_{1,3}, \dots$. Continuing in this way, selecting subsequences of subsequences, produces an array,

$$\begin{array}{cccccc} f_{1,1}(z_1), & f_{1,2}(z_1), & f_{1,3}(z_1), & \dots & \rightarrow & w_1 \\ f_{2,1}(z_2), & f_{2,2}(z_2), & f_{2,3}(z_2), & \dots & \rightarrow & w_2 \\ f_{3,1}(z_3), & f_{3,2}(z_3), & f_{3,3}(z_3), & \dots & \rightarrow & w_3 \\ f_{4,1}(z_4), & f_{4,2}(z_4), & f_{4,3}(z_4), & \dots & \rightarrow & w_4 \\ \vdots & \vdots & \vdots & \vdots & & \vdots \end{array}$$

in which the k th horizontal row converges to some complex number w_k and the functions used in each row are selected from among those in the row above. The proof uses a procedure, called the **diagonal construction**, which is sometimes useful in other contexts. Let $g_n = f_{n,n}$. Then g_1, g_2, g_3, \dots is a subsequence of the original sequence of functions, and $\lim_{l \rightarrow \infty} g_l(z_k) = w_k$ for each k . This is because $g_n = f_{n,n}$ is a subsequence of $f_{k,1}, f_{k,2}, f_{k,3}, \dots$ as soon as $n > k$. Thus the subsequence g_n converges at a set of points that are scattered densely throughout A . Steps (iii) and (iv) of the program are to show that the fact that the g_n 's are uniformly bounded on closed disks in A is enough to force them to converge everywhere in G and in fact to do so uniformly on closed disks in A . We accomplish this by showing that the sequence satisfies the Cauchy condition uniformly on closed disks.

Let B be a closed disk contained in A , and let $\epsilon > 0$. By the supplementary results for Chapter 3 (see Theorem 3.4 in this Supplement), the functions g_n are uniformly equicontinuous on B ; that is, there is a number $\delta > 0$ such that $|g_l(\zeta) - g_l(\xi)| < \epsilon/3$ for all l whenever ζ and ξ are in B and $|\zeta - \xi| < \delta$. By using only finitely many of the points z_k we can guarantee that everything in B is within a distance δ of at least one of them. That is, there is an integer $K(B)$ such that for each $z \in B$ there is at least one $k \in \{1, 2, 3, \dots, K(B)\}$ with $|z - z_k| < \delta$ and hence $|g_l(z) - g_l(z_k)| < \epsilon/3$ for all l . One way to do this would be to take a square grid of points with rational coordinates and separation less than δ (see Figure 6.1). Since $\lim_{l \rightarrow \infty} g_l(z_k) = w_k$ for each k , each of these sequences satisfies the Cauchy condition, and as there are only finitely many of them, there is an integer $N(B)$ such that $|g_n(z_k) - g_m(z_k)| < \epsilon/3$ whenever $n \geq N(B), m \geq N(B)$, and $1 \leq k \leq K(B)$.

Putting all this together, suppose $n \geq N(B)$ and $m \geq N(B)$. If $z \in A$, then z is within δ of z_k for some $k \leq K(B)$, so

$$\begin{aligned} |g_n(z) - g_m(z)| &\leq |g_n(z) - g_n(z_k)| + |g_n(z_k) - g_m(z_k)| + |g_m(z_k) - g_m(z)| \\ &\leq \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon. \end{aligned}$$

The sequence g_n thus uniformly satisfies the Cauchy condition on B , so converges uniformly on B to some limit function, as desired. ■

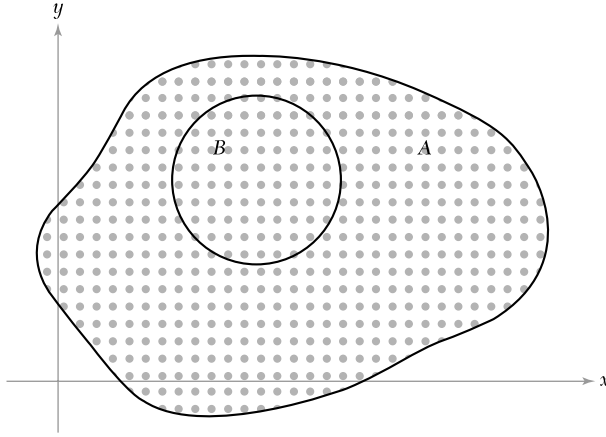


Figure 6.1: Finitely many of the z_k 's give one within δ of anything in B .

Proof of the Riemann Mapping Theorem We are now in a position to prove the Riemann Mapping Theorem. Let G be a connected, simply connected, open set properly contained in the complex plane \mathbb{C} . Let $z_0 \in G$, and let $D = D(0; 1)$ be the open unit disk. We must show that there is a function f analytic on G that maps G one-to-one onto D with $f(z_0) = 0$ and $f'(z_0) > 0$. To do this, let

$$\mathcal{S} = \{f : G \rightarrow D \mid f \text{ is analytic and one-to-one on } G, f(z_0) = 0, \text{ and } f'(z_0) > 0\}.$$

The main steps of the proof are:

- (i) Show that \mathcal{S} is not empty.
- (ii) Show that the numbers $\{f'(z_0) \mid f \in \mathcal{S}\}$ are bounded above, so have a finite least upper bound M .
- (iii) Use Montel's theorem to extract from a sequence of functions in \mathcal{S} whose derivatives at z_0 converge to M a subsequence that converges uniformly on closed disks in G . The limit function f is analytic in G and $f'(z_0) = M$.
- (iv) Show that $f \in \mathcal{S}$.
- (v) Show that f must map G onto D .

To show that \mathcal{S} is not empty, it is enough to show that we can map G analytically into the unit disk. Once that is done, we need only compose with a linear fractional transformation of the disk onto itself, which takes z_0 to 0, and then multiply by a constant $e^{i\theta}$, chosen so that the derivative of the resulting map at z_0 is positive. If G is bounded, for example, if $|z - z_0| < R$ for all z in G , the map $z \mapsto (z - z_0)/R$ does the job. If G is not bounded, it at least omits a point a . The translation

$z \mapsto z - a$ takes G to a simply connected region G_1 not containing 0. By Theorem 2.2.6, there is a branch of logarithm defined on G_1 , which we will call F . Then the map g defined by $z \mapsto e^{(1/2)F(z)}$ is a branch of the square root function; by the open mapping or inverse mapping theorem, one sees that $G_2 = g(G_1)$ contains some disk $D(b; r)$. By properties of the square root function, $D(-b; r)$ fails to meet G_2 . The map $f(z) = r/[b + z]$ then maps G_2 into the unit disk. See Figure 6.2.

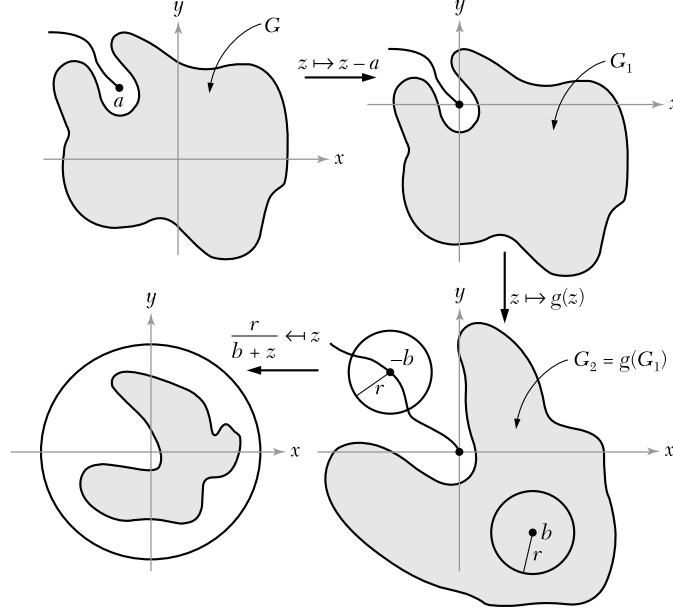


Figure 6.2: Mapping G into the unit disk.

Having shown that \mathcal{S} is not empty, we must establish step (ii). The family \mathcal{S} is uniformly bounded by 1 on G , so by the supplementary Theorem 3.2, the derivatives are uniformly bounded on closed disks in G . In particular, there is a finite number $M(\{z_0\})$ such that $f'(z_0) \leq M(\{z_0\})$ for all f in \mathcal{S} . Let M be the least upper bound of these derivatives. There must be a sequence f_1, f_2, f_3, \dots of functions in \mathcal{S} with the property that $\lim_{n \rightarrow \infty} f'_n(z_0) = M$. Since the family \mathcal{S} is uniformly bounded, it is normal by Montel's Theorem and there must be a subsequence that converges uniformly on closed disks in G . We may as well throw away the functions we don't need and assume that we have a sequence that converges uniformly on closed disks in G . By the Analytic Convergence Theorem 3.1.8 they converge to a limit function f , which is analytic on G and $f'(z_0) = M$.

We next want to know that f is a member of \mathcal{S} . Each of the functions f_n maps G into the open unit disk, so f certainly maps G into the closed unit disk. Since f is not constant, the Maximum Modulus Principle says that $|f(z)|$ cannot have a maximum anywhere in G , so the image never touches the boundary of the disk and

f maps G into D . Certainly $f(z_0) = \lim_{n \rightarrow \infty} f_n(z_0) = 0$. Finally, the corollary of Hurwitz' Theorem 6.2.8 shows that f must be one-to-one since it is a nonconstant limit of one-to-one functions that converge uniformly on closed disks. Thus, $f \in \mathcal{S}$.

The final step, (iv), is to show that f must actually map G onto D . This follows from the following assertion.

Claim *If A is a connected and simply connected open set properly contained in D and $0 \in A$, then there is a function F analytic on A that maps A one-to-one into D with $F(0) = 0$ and $F'(0) > 1$.*

To see how (iv) follows from this assertion, suppose that f does not map G onto D . Then $A = f(G)$ satisfies the conditions of the claim. (That A is open follows from the Open Mapping Theorem 6.3.3. Consider $g(z) = F(f(z))$. Then $g \in \mathcal{S}$, but $g'(z_0) = F'(f(z_0))f'(z_0) = F'(0)M > M$, contradicting the maximality of M .)

Thus it remains to check the claim. The construction is a bit like that used in step (i) and is seen by following the diagrams in Figure 6.3.

The region A is shaded by diagonal lines in the first diagram. It misses a point a indicated by an open circle in the diagram. The successive images of a and 0 are indicated by open dots and solid dots, respectively, in each of the following diagrams. Map F_1 is a linear fractional transformation of the disk to itself taking a to 0 and 0 somewhere. The purpose of map F_2 is to guarantee a situation in which the image of A misses a neighborhood of a point on the boundary circle. This is done just as in step (i) by using a branch of logarithm on the simply connected region $F_1(A)$ that misses 0 . Map F_3 is another linear fractional transformation that returns the image of 0 to 0 . At this stage the image of A misses a small circle γ that intersects the unit circle C at right angles at two points. An appropriate linear fractional transformation F_4 taking these points to 0 and ∞ will take the circles to lines through 0 and ∞ and the region between them to a quarter plane. Squaring F_5 opens this up to a half plane. Finally another linear fractional transformation takes the half plane to the unit disk with the black dot going to 0 and the correct rotation making the derivative of the whole thing at 0 positive. The function F is

$$F(z) = F_6(F_5(F_4(F_3(F_2(F_1(z))))) = w.$$

The inverse function $g(w) = F^{-1}(w) = z$ satisfies the conditions of the Schwarz Lemma. Since it is not a rotation, we have strict inequality $|g'(0)| < 1$ by the Schwarz Lemma, but $F'(0) = 1/g'(0)$. Therefore, $F'(0) > 1$, as required. All the pieces have been assembled, so the proof of the Riemann Mapping Theorem is now complete. ■

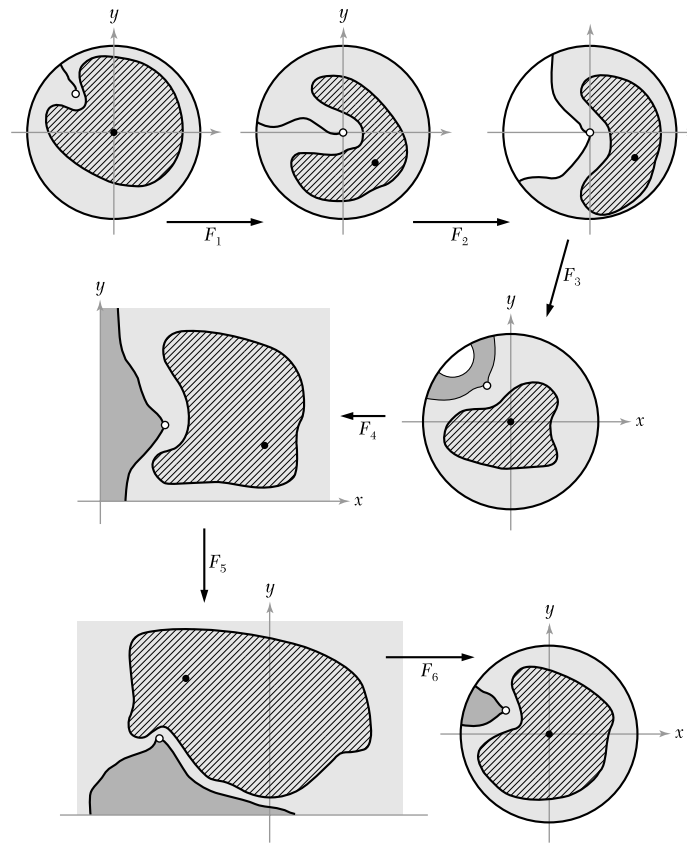


Figure 6.3: Construction for the claim in the proof of the Riemann Mapping Theorem.

Dynamics of Complex Analytic Mappings

The pictures shown in Figure 6.4 are representations of the dynamics of complex analytic mappings. The purpose of this section is to provide a brief introduction to this subject—mainly to inspire the reader to find out more by consulting a reference on the subject.²

The subject we will be looking at has to do with the way points in the complex plane behave under iteration of an analytic function. It has its origins in classical

²Such as R. L. Devany, *An Introduction to Chaotic Dynamical Systems* (Reading, Mass.: Addison-Wesley, 1985); P. Blanchard, Complex dynamics on the Riemann sphere, *Bulletin of the American Mathematical Society*, **11** (1984), 85–141; or B. Mandelbrot, *The Fractal Geometry of Nature* (New York: W. H. Freeman and Company, 1982).

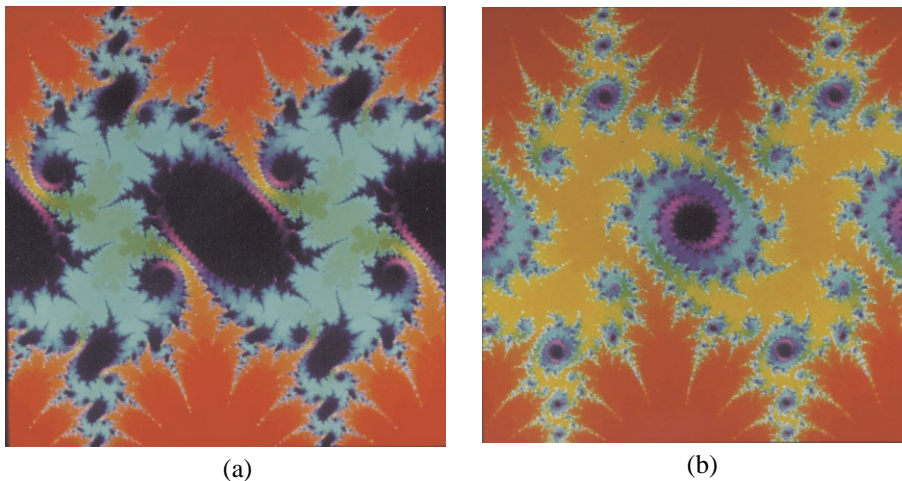


Figure 6.4: The different shadings represent the rate of approach of points to infinity under iteration of the mapping; the black region (assuming that you are viewing the figure in color) consists of “stable” points that remain bounded under iteration. In part (a) the mapping is $(1+0.1i)\sin z$, while in (b) it is $(1+0.2i)\sin z$. (*Courtesy of R. Devany of Boston University, with the assistance of C. Mayberry, C. Small, and S. Smith*)

and beautiful work of G. Julia³ and P. Fatou.⁴ In this study normal families play an important role. In fact, Montel himself was interested in these questions.⁵

Let us fix an entire function $f : \mathbb{C} \rightarrow \mathbb{C}$. We need a little terminology to get going. Given a point $z \in \mathbb{C}$, the **orbit** of z is the sequence of points

$$z, f(z), f(f(z)), f(f(f(z))), \dots,$$

which we also write as $z, f(z), f^2(z), f^3(z), \dots$. We think of the point z as moving successively under the mapping f to new locations. A **fixed point** is a point z such that $f(z) = z$, that is, a point z that does not move when we apply f . A **periodic point** is a point z such that $f^n(z) = z$ for some integer n (called the period), where f^n means f composed with itself n times.

A fixed point z is called an **attracting fixed point** if $|f'(z)| < 1$. The reason for this terminology is that the orbits of nearby points converge to z ; this is so because near z , f behaves like a mapping that rotates by an amount $\arg f'(z)$ and magnifies by an amount $|f'(z)|$, so every time f is applied, points will be pulled toward z by a factor $|f'(z)|$, so as this is repeated, the point tends to z . Likewise,

³Memoiré sur l'itérations des fonctions rationnelles, *J. Math.*, **8** (1918), 47–245.

⁴Sur l'itérations des fonctions transcendentes entière, *Acta Math.*, **47** (1926), 337–370.

⁵See his *Leçons sur les familles normales de fonctions analytiques et leurs applications* (1927; reprinted New York: Chelsea, 1974), Chapter VIII.

a point z is called a **repelling fixed point** if $|f'(z)| > 1$; points near repelling points will be pushed away under iteration of the function f . Similarly, a periodic point z with period n is called an **attracting periodic point** if $|(f^n)'(z)| < 1$; such points have the property that the orbits of points close to z tend to the orbit of z . Likewise, a **repelling periodic point** has the property that $|(f^n)'(z)| > 1$; orbits of points near such points will be shoved away from the orbit of z .

The **Julia set** $J(f)$ of f is defined to be the closure of the set of repelling periodic points of f . This set can have remarkable and beautiful complexity usually called a **fractal**; in fact, in the picture in Figure 6.4 the nonblack region is the Julia set. This statement rests on a theorem, which we shall not prove, stating that the Julia set is the closure of the points that go to infinity under iteration of f . It is this characterization that is useful for computational purposes. Figure 6.5 shows two more Julia sets for quadratic maps.

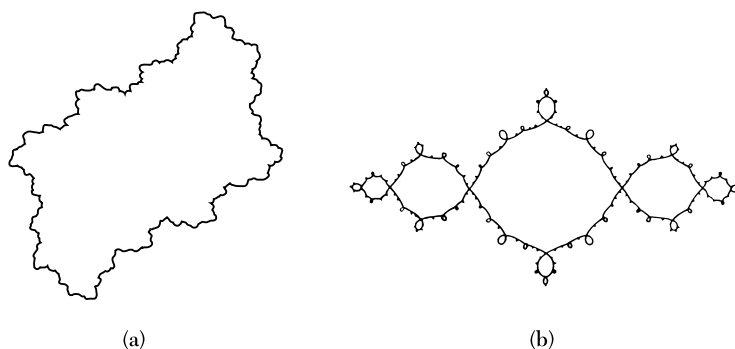


Figure 6.5: (a) Julia set of $f(z) = z^2 + \frac{1}{2}i$, which is a simple closed curve but is nowhere differentiable. (b) Julia set of $f(z) = z^2 - 1$, which contains infinitely many closed curves.

As far as complex analysis is concerned, one of the most important results is the following:

The Julia set of f is the set of points at which the family of functions f^n is not normal.

This result can be used as an alternative definition for the Julia set, and in fact this was the original definition of Fatou and Julia. We will prove only the following statement here to give a flavor of how the arguments go:

If f is a repelling fixed point of f (and therefore is in the Julia set), then the family of iterates f^n fails to be normal at z .

Let us assume that this family is normal at z and derive a contradiction. Normal at z means normal in a neighborhood of z , in the same way we used the terminology “analytic at z .” By Definition 6.1, the family f^n has a subsequence that converges

uniformly on a neighborhood of z . Since $f(z) = z$ and $|f'(z)| > 1$, it follows from the chain rule that

$$|f^{n'}(z)| = |f'(z)|^n \rightarrow \infty,$$

that is, that the sequence of derivatives of f^n evaluated at z must tend to infinity as $n \rightarrow \infty$. However, the sequence of derivatives must converge to the derivative of the limit function by the Analytic Convergence Theorem 3.1.8, which is finite, giving us the required contradiction.

This discussion represents only the tip of a large collection of very interesting and beautiful results. We hope the reader will be inspired to look up some of the references on the subject we have given as well as further references found in those sources and will explore the subject further. We hasten to point out that the iteration of complex mappings is just one part of a larger and growing field called *chaotic dynamics*. For the more general aspects, the reader can consult Devany's book cited in footnote 2 or the book *Nonlinear Oscillations, Dynamical Systems, and Bifurcation of Vector Fields*, by J. Guckenheimer and P. Holmes (New York: Springer-Verlag, 1983).

Chapter 7

Asymptotic Methods

Proof of the Steepest Descent Theorem

The goal of this section is to provide the proof of the Steepest Descent Theorem that we omitted in the textbook. We first recall the statement of the theorem.

Theorem 7.1 (Steepest Descent Theorem) *Let $\gamma :] - \infty, \infty[\rightarrow \mathbb{C}$ be a C^1 curve. (γ may also be defined only on a finite interval.) Let $\zeta_0 = \gamma(t_0)$ be a point on γ and let $h(\zeta)$ be a function continuous along γ and analytic at ζ_0 . Make the following hypotheses: For $|z| \geq \mathbb{R}$ and $\arg z$ fixed,*

(i) *The integral*

$$f(z) = \int_{\gamma} e^{zh(\zeta)} d\zeta$$

converges absolutely.

(ii) $h'(\zeta_0) = 0; h''(\zeta_0) \neq 0$.

(iii) $\text{Im}[zh(\zeta)]$ is constant for ζ on γ in some neighborhood of ζ_0 .

(iv) $\text{Re}[zh(\zeta)]$ has a strict maximum at ζ_0 along the entire curve γ .

Then

$$f(z) \sim \frac{e^{zh(\zeta_0)} \sqrt{2\pi}}{\sqrt{z} \sqrt{-h''(\zeta_0)}}$$

as $z \rightarrow \infty, \arg z$ fixed. The sign of the square root is chosen such that

$$\sqrt{z} \sqrt{-h''(\zeta_0)} \cdot \gamma'(t_0) > 0.$$

Proof We begin by breaking up the curve γ into three portions, γ_1 , C , and γ_2 , as illustrated in Figure 7.1. Choose C such that it lies in a neighborhood of ζ_0 small enough that $h(\zeta)$ is analytic and that condition (iii) holds. Clearly,

$$f(z) = I_1(z) + I_2(z) + J(z),$$

where we use the notations

$$J(z) = \int_C e^{zh(\zeta)} d\zeta \quad \text{and} \quad I_k(z) = \int_{\gamma_k} e^{zh(\zeta)} d\zeta, \quad k = 1, 2.$$

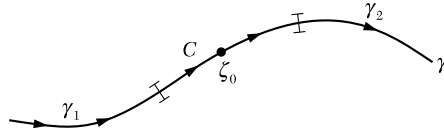


Figure 7.1: Method of steepest descent.

We will show that for large z the part of the integral that really matters is $J(z)$ so that an asymptotic approximation for $J(z)$ will also give one for $f(z)$. Worked Example 7.2.12 says that to do this, it is enough to show that

$$\frac{I_k(z)}{J(z)} = O\left(\frac{1}{z^n}\right)$$

for all positive n . To prove this, note that

$$|I_k(z)| = \left| \int_{\gamma_k} e^{zh(\zeta)} d\zeta \right| \leq \int_{\gamma_k} e^{\operatorname{Re} zh(\zeta)} |d\zeta|.$$

However,

$$J(z) = \int_C e^{zh(\zeta)} d\zeta = \int_C e^{\operatorname{Re} zh(\zeta)} e^{i \operatorname{Im} zh(\zeta)} d\zeta.$$

Since $\operatorname{Im}[zh(\zeta)]$ is constant on C , we get

$$|J(z)| = \left| \int_C e^{\operatorname{Re} zh(\zeta)} d\zeta \right|.$$

If C is short enough that $\arg(\gamma')$ changes by less than $\pi/4$ along C , then we obtain

$$|J(z)| > (1/\sqrt{2}) \int_C e^{\operatorname{Re} zh(\zeta)} |d\zeta|.$$

Thus,

$$\left| \frac{I_k(z)}{J(z)} \right| \leq \frac{\int_{\gamma_k} e^{\operatorname{Re} zh(\zeta)} |d\zeta|}{\int_C e^{\operatorname{Re} zh(\zeta)} |d\zeta|} \sqrt{2}.$$

Let \tilde{C} be a strictly smaller subinterval of C , centered at ζ_0 . Then

$$\left| \frac{I_k(z)}{J(z)} \right| \leq \frac{\int_{\gamma_k} e^{\operatorname{Re} zh(\zeta)} |d\zeta|}{\int_{\tilde{C}} e^{\operatorname{Re} zh(\zeta)} |d\zeta|} \sqrt{2}.$$

Fix z_0 and let α be the minimum of $\operatorname{Re}[zh(\zeta)]$ on \tilde{C} . There is an $\epsilon > 0$ such that $\operatorname{Re}[z_0 h(\zeta)] \leq \alpha - \epsilon$ for all $\zeta \in \gamma_k$. Thus, using the fact that z lies on the same ray as z_0 ,

$$\begin{aligned} \frac{\int_{\gamma_k} e^{\operatorname{Re} zh(\zeta)} |d\zeta|}{\int_{\tilde{C}} e^{\operatorname{Re} zh(\zeta)} |d\zeta|} &= \frac{\int_{\gamma_k} e^{\operatorname{Re} z_0 h(\zeta)} e^{\operatorname{Re}(z-z_0)h(\zeta)} |d\zeta|}{\int_{\tilde{C}} e^{\operatorname{Re} z_0 h(\zeta)} e^{\operatorname{Re}(z-z_0)h(\zeta)} |d\zeta|} \\ &\leq \frac{\left(\int_{\gamma_k} e^{\operatorname{Re} z_0 h(\zeta)} |d\zeta| \right) e^{|z-z_0|(\alpha-\epsilon)/|z_0|}}{\left(\int_{\tilde{C}} e^{\operatorname{Re} z_0 h(\zeta)} |d\zeta| \right) e^{|z-z_0|\alpha/|z_0|}}. \end{aligned}$$

This expression is a constant factor, say M , times $e^{-|z-z_0|\epsilon/|z_0|}$. The latter is certainly $O(1/z)$ (and in fact is $O(1/z^n)$ for all $n \geq 1$), so we have proved that $I_k(z)/J(z) = O(1/z^n)$ for all $n \geq 1$. This localizes the problem to a neighborhood around ζ_0 where the bulk of the contribution to the integral is made. Also, we can shrink the length of C without affecting the conclusion that $f(z) \sim J(z)$ as $z \rightarrow \infty$.

Next, we write

$$h(\zeta) = h(\zeta_0) - w(\zeta)^2,$$

where $w(\zeta)$ is analytic and invertible (abusing notation, we denote the inverse by $\zeta(w)$), where $w(\zeta_0) = 0$, and where

$$[w'(\zeta_0)]^2 = \frac{-h''(\zeta_0)}{2}$$

(see Worked Example 6.3.7). Since

$$\operatorname{Im}(zh(\zeta)) = \operatorname{Im}[zh(\zeta_0)]$$

on C and

$$\operatorname{Re}[zh(\zeta)] < \operatorname{Re}[zh(\zeta_0)],$$

we see that $z[w(\zeta)]^2$ is real and greater than zero on C ; also, by our choice of branch for the square root, $\sqrt{z}w(\zeta)$ is real and, as a function of the curve parameter t , has positive derivative at t_0 . Thus, by shrinking C if necessary, we can assume that $\sqrt{z}w(\zeta)$ is increasing along C .

Note that

$$J(z) = \int_C e^{zh(\zeta)} d\zeta = \int_C e^{zh(\zeta_0)} \cdot e^{-zw(\zeta)^2} d\zeta = e^{zh(\zeta_0)} \int_C e^{-z[w(\zeta)]^2} d\zeta.$$

We change variables by setting $\sqrt{z}w(\zeta) = y$, and we get

$$J(z) = e^{zh(\zeta_0)} \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} \frac{d\zeta}{dw} \frac{dy}{\sqrt{z}} = \frac{e^{zh(\zeta_0)}}{\sqrt{z}} \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} \frac{d\zeta}{dw} dy,$$

since y is real on C ; we choose positive numbers ϵ_1 and ϵ_2 such that $[-\sqrt{|z|\epsilon_1}, \sqrt{|z|\epsilon_2}]$ is in the range of y corresponding to ζ on C . Next we write

$$\zeta = \zeta_0 + a_1 w + a_2 w^2 + \dots,$$

so

$$\frac{d\zeta}{dw} = a_1 + 2a_2 w + 3a_3 w^2 + \dots,$$

where $w = y/\sqrt{z}$. Thus,

$$\begin{aligned} \frac{J(z)\sqrt{z}}{e^{zh(\zeta_0)}} &= \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} \left[\sum_{k=1}^{\infty} k a_k \left(\frac{y}{\sqrt{z}} \right)^{k-1} \right] dy \\ &= \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} \left[\sum_{k=0}^N (k+1) a_{k+1} \left(\frac{y}{\sqrt{z}} \right)^k \right] dy \\ &\quad + \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} O \left(\left(\frac{y}{\sqrt{z}} \right)^{N+1} \right) dy \\ &= \sum_{k=0}^N \frac{(k+1) a_{k+1}}{(\sqrt{z})^k} \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} y^k dy \\ &\quad + \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} O \left(\left(\frac{y}{\sqrt{z}} \right)^{N+1} \right) dy. \end{aligned}$$

By Exercise 7,

$$\int_{-\infty}^{\infty} e^{-y^2} y^k dy = \frac{(2m)! \sqrt{\pi}}{m! 2^{2m}}$$

if $k = 2m$ is even and it is zero if $k = 2m + 1$ is odd, so we are led to the series

$$S \equiv \sum \frac{(k+1) a_{k+1}}{(\sqrt{z})^k} \int_{-\infty}^{\infty} e^{-y^2} y^k dy = \sum_{m=0}^{\infty} \frac{(2m)! \sqrt{\pi}}{m! 2^{2m}} \frac{(2m+1) a_{2m+1}}{z^m}.$$

This gives

$$\begin{aligned} \frac{J(z)}{e^{zh(\zeta_0)}/\sqrt{z}} - S_M &= - \sum_{k=0}^{2M} \frac{(k+1) a_{k+1}}{(\sqrt{z})^k} \left(\int_{-\infty}^{\sqrt{|z|\epsilon_1}} e^{-y^2} y^k dy + \int_{\sqrt{|z|\epsilon_2}}^{\infty} e^{-y^2} y^k dy \right) \\ &\quad + \int_{-\sqrt{|z|\epsilon_1}}^{\sqrt{|z|\epsilon_2}} e^{-y^2} O \left(\left(\frac{y}{\sqrt{z}} \right)^{2M+1} \right) dy. \end{aligned}$$

The first two integrals are $o(1/(\sqrt{z})^{2M})$ by Proposition 7.2.3(v), since $e^{-y^2} y^k = o(1/y^{2M+1})$. In the third, there is a constant B_M such that the integrand is bounded by

$$\frac{B_M e^{-y^2} |y|^{2M+1}}{|\sqrt{z}|^{2M+1}} = \left(\frac{B_M}{|z|^M} \right) \left(\frac{1}{\sqrt{|z|}} \right) e^{-y^2} |y|^{2M+1}.$$

Since

$$\int_{-\infty}^{\infty} e^{-y^2} |y|^{2M+1} dy < \infty,$$

this term is also $o(1/|z|^M)$. Thus,

$$J(z) \sim \frac{e^{zh(\zeta_0)} S}{\sqrt{z}},$$

and by Worked Example 7.2.12, the same is true of $f(z)$. Thus,

$$f(z) \sim e^{zh(\zeta_0)} \sqrt{\frac{\pi}{z}} \left(a_1 + \frac{1 \cdot 3 a_3}{z} + \frac{1 \cdot 3 \cdot 5 a_5}{z^2} + \dots \right).$$

To complete the proof, note that

$$a_1 = \frac{d\zeta}{dw}(0) = \frac{1}{\frac{dw}{d\zeta}(\zeta_0)} = \frac{\sqrt{2}}{\sqrt{-h''(\zeta_0)}},$$

so

$$f(z) \sim \frac{e^{zh(\zeta_0)} \sqrt{2\pi}}{\sqrt{z} \sqrt{-h''(\zeta_0)}},$$

as desired. ■

Bounded Variation and the Stationary Phase Formula

We saw in the method of stationary phase that we needed to impose a condition on the amplitude that limits the amount of high-frequency oscillation. This type of condition is often needed in theory involving integrals; the notion of bounded variation provides the appropriate tools. We will then use it to prove the Stationary Phase Formula.

Definition 7.2 Suppose $f : [a, b] \rightarrow \mathbb{R}$.

- (i) If P is a partition of $[a, b]$ given by $a = t_0 < t_1 < \dots < t_n = b$, then the **variation of f on $[a, b]$ relative to P** is defined to be

$$V_P f = \sum_{k=1}^n |f(t_k) - f(t_{k-1})|.$$

- (ii) The **total variation of f on $[a, b]$** is $V_{[a,b]}f = \sup\{V_P f\}$, where the least upper bound is taken over all possible partitions. (It might be $+\infty$.)
- (iii) If $V_{[a,b]}f < \infty$ we say that f is of **bounded variation** and write

$$f \in BV([a, b]).$$

Some important examples of such functions are included in the following.

Proposition 7.3

- (i) If f is monotone and bounded on $[a, b]$, then $f \in BV([a, b])$ and $V_{[a,b]}f = |f(b) - f(a)|$.
- (ii) If f is differentiable on a bounded interval $[a, b]$ and $|f'(x)| < M$ for all $x \in [a, b]$, then $f \in BV([a, b])$ and $V_{[a,b]}f \leq |b - a|M$.
- (iii) If f has a continuous derivative on the bounded interval $[a, b]$ —that is, if $f \in C^1([a, b])$ —then $f \in BV([a, b])$.

Proof The first result holds since the succeeding differences from point to point along any partition are all of the same sign and values at intermediate points cancel out. The second is shown by applying the mean value theorem to each subinterval of any partition, and the third follows from it since if f' is continuous on the compact interval $[a, b]$, then it is bounded. ■

It is possible for a continuous function not to have bounded variation. On $[-1, 1]$ set $f(0) = 0$ and $f(x) = x \cos(1/x)$ for $x \neq 0$. (See Figure 7.2.) Then we have

$$|f(1/n\pi) - f(1/(n+1)\pi)| = (2n+1)/n(n+1)\pi > 1/n\pi.$$

Since the harmonic series diverges, partitions may be created using these points that give arbitrarily large variation.

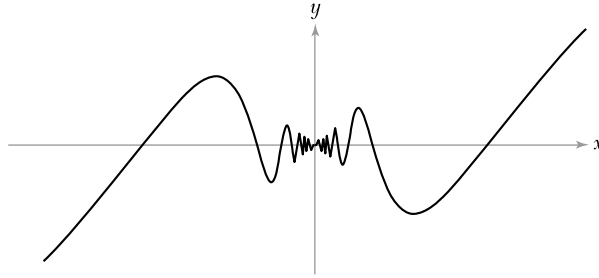


Figure 7.2: The continuous function $x \cos(1/x)$ has unbounded variation.

Some of the important properties of functions of bounded variation are outlined in the following proposition.

Proposition 7.4 Suppose $f \in BV([a, b])$.

- (i) If $[c, d] \subset [a, b]$, then $f \in BV([c, d])$ and $V_{[c,d]}f \leq V_{[a,b]}f$.
- (ii) $V_{[a,c]}f + V_{[c,b]}f = V_{[a,b]}f$ if $a < c < b$.
- (iii) $(Vf)(x) = V_{[a,x]}f$ is a bounded increasing function on $[a, b]$ with $(Vf)(a) = 0$ and $(Vf)(b) = V_{[a,b]}f$.
- (iv) If $a \leq x \leq y \leq b$, then $(Vf)(y) - (Vf)(x) = V_{[x,y]}f$.
- (v) f is the difference of two bounded increasing functions: $f = f_1 - f_2$ with $f_1 = (Vf + f)/2$ and $f_2 = (Vf - f)/2$.

Proof The first assertion follows since any partition of $[c, d]$ can be extended by the intervals $[a, c]$ and $[d, b]$ to obtain a partition of $[a, b]$ offering a larger candidate for $V_{[a,b]}f$. For the second, adjoin partitions of $[a, c]$ and $[c, b]$ to get a partition of $[a, b]$ and show $V_{[a,c]}f + V_{[c,b]}f \leq V_{[a,b]}f$. For the opposite inequality let $a = t_0 < t_1 < \dots < t_n = b$ be any partition of $[a, b]$ with

$$\sum_{k=1}^n |f(t_k) - f(t_{k-1})| > V_{[a,b]}f - \epsilon.$$

Pick N with $t_N \leq c \leq t_{N+1}$. Then

$$\begin{aligned} V_{[a,b]}f &< \sum_{k=0}^N |f(t_k) - f(t_{k-1})| + \sum_{k=N}^{N+1} |f(t_k) - f(t_{k-1})| + \epsilon \\ &< \sum_{k=0}^N |f(t_k) - f(t_{k-1})| + |f(c) - f(t_N)| + |f(t_{N+1}) - f(c)| \\ &\quad + \sum_{k=N+2}^n |f(t_k) - f(t_{k-1})| + \epsilon \\ &\leq V_{[a,c]}f + V_{[c,b]}f + \epsilon. \end{aligned}$$

Since this holds for any $\epsilon \geq 0$, we have the desired inequality. The third assertion is clear and the fourth follows from it and the second. For the last assertion, use (iv) to show that the functions indicated are increasing. ■

The last property will be the one directly utilized in the proof of Theorem 7.2.10. The tool by which we will use it is the second mean value theorem for integrals.

Theorem 7.5 (Second Mean Value Theorem for Integrals) If f is bounded and increasing on $[a, b]$ and g is integrable, then there is a point c in $[a, b]$ such that

$$\int_a^b f(t)g(t)dt = f(a) \int_a^c g(t)dt + f(b) \int_c^b g(t)dt.$$

Proof Let

$$F(x) = f(a) \int_a^x g(t)dt + f(b) \int_x^b g(t)dt.$$

Then

$$F(a) = f(b) \int_a^b g(t)dt \quad \text{and} \quad F(b) = f(a) \int_a^b g(t)dt$$

and

$$f(a) \int_a^b g(t)dt \leq \int_a^b f(t)g(t)dt \leq f(b) \int_a^b g(t)dt.$$

Since F is continuous on $[a, b]$, the conclusion follows from the intermediate value theorem. ■

We will also need the following estimates.

Lemma 7.6 *If f is differentiable on $[a, b]$ and $|f'(x)| \leq M$ for all x in $[a, b]$, then $|f(y) - f(x)|$, $|(Vf)(y) - (Vf)(x)|$, $|f_1(y) - f_1(x)|$, and $|f_2(y) - f_2(x)|$ are each bounded above by $M|y - x|$.*

Proof The first conclusion follows from the mean value theorem and the second from Proposition 7.3(ii) and Proposition 7.4(iv). The last two follow from the first two and the formulas for f_1 and f_2 given in Proposition 7.4(v). ■

The first step in the intuitive derivation of Theorem 7.2.10 was that contributions to the integral from parts of the interval away from t_0 tended to cancel out and could be neglected by comparison with the contribution from a short interval near t_0 . The only critical point of h was at t_0 and h' and h'' were continuous, so away from t_0 the derivative stays away from 0 and we can apply the following lemma.

Lemma 7.7 *Suppose h has a continuous second derivative on $[a, b]$, that $h'(x)$ is never 0 in $[a, b]$, and that g has a continuous derivative on $[a, b]$. Then*

$$\int_a^b e^{izh(t)} g(t)dt = O(1/z).$$

Proof The function $\psi(x) = g(x)/h'(x)$ has a continuous derivative on $[a, b]$ and thus has bounded variation and may be written as a difference of two increasing functions, $\psi = \psi_1 - \psi_2$. Then

$$\begin{aligned} \int_a^b e^{izh(t)} g(t)dt &= \frac{1}{z} \int_a^b \psi_1(t) \cos(zh(t))zh'(t)dt + \frac{i}{z} \int_a^b \psi_1(t) \sin(zh(t))zh'(t)dt \\ &\quad - \frac{1}{z} \int_a^b \psi_2(t) \cos(zh(t))zh'(t)dt - \frac{i}{z} \int_a^b \psi_2(t) \sin(zh(t))zh'(t)dt. \end{aligned}$$

Each of these integrals may be estimated using the second mean value theorem for integrals. There is a point x between a and b with

$$\begin{aligned} \left| \int_a^b \psi_1(t) \cos(zh(t)) zh'(t) dt \right| &= \left| \psi_1(a) \int_a^x \cos(zh(t)) zh'(t) dt \right. \\ &\quad \left. + \psi_1(b) \int_x^b \cos(zh(t)) zh'(t) dt \right| \\ &= |\psi_1(a)[\sin(zh(x)) - \sin(zh(a))] \\ &\quad + \psi_1(b)[\sin(zh(b)) - \sin(zh(x))]| \\ &\leq 2|\psi_1(a)| + 2|\psi_1(b)|. \end{aligned}$$

The others are treated similarly to obtain

$$\left| \int_a^b e^{izh(t)} g(t) dt \right| \leq \frac{4}{z} [|\psi_1(a)| + |\psi_1(b)| + |\psi_2(a)| + |\psi_2(b)|],$$

as needed. ■

We now complete the proof of Theorem 7.2.10. Since t_0 is the only critical point of h in $[a, b]$, we know that for any $\delta > 0$, $h'(t)$ is never 0 on $[a, t_0 - \delta]$ or $[t_0 + \delta, b]$, so by the last lemma, the integrals of $e^{izh(t)} g(t)$ over each are $O(1/z)$, so are $o(1/\sqrt{z})$. Thus, to establish Theorem 7.2.10 it is enough to show that

$$\lim_{z \rightarrow \infty} \sqrt{z} e^{-ih(t_0)} \int_J e^{izh(t)} g(t) dt = \frac{\sqrt{2\pi}}{\sqrt{\pm h''(t_0)}} e^{\pm \pi i/4} g(t_0),$$

where $J = [t_0 - \delta, t_0 + \delta]$. We may fix δ as small as we please as long as its choice does not depend on z . In the course of the proof we shall find conditions for that choice.

We know h is analytic in a neighborhood of t_0 . By Worked Example 6.3.7 there is an analytic function $w(t)$ such that $h(t) = h(t_0) \pm [w(t)]^2$ for t near t_0 and w is locally one-to-one. We may choose w to be real and strictly increasing on J if δ is selected small enough. This is our first criterion for δ . We choose the plus sign if $h''(t_0) > 0$ and the minus sign if $h''(t_0) < 0$. Since $w(t_0) = 0$ and w is continuous, $w(t_0 + \delta) = c$ and $w(t_0 - \delta) = d$, where $c < 0 < d$. The change of variables $x = w(t)$ gives

$$\int_J e^{izh(t)} g(t) dt = e^{izh(t_0)} \int_c^d e^{\pm izx^2} \psi(x) dx,$$

where $\psi(x) = g(w^{-1}(x))/(w^{-1})'(x)$. The function ψ has a continuous derivative on $[c, d]$. The point $x = 0$ corresponds to $t = t_0$, and

$$h''(t_0) = \pm 2w(t_0)w''(t_0) \pm 2[w'(t_0)]^2.$$

Thus, $(w^{-1})'(0) = 1/w'(t_0) = \sqrt{\pm h''(t_0)}/2$. Since ψ' is continuous, ψ has bounded variation and can be written as a difference $\psi_1 - \psi_2$ of two increasing functions. Let $\epsilon > 0$. Since c and d got to 0 as $\delta \rightarrow 0$, we can use Lemma 7.6 to select δ small enough so that the quantities $|\psi_1(c) - \psi_1(0)|$, $|\psi_1(d) - \psi_1(0)|$, $|\psi_2(c) - \psi_2(0)|$, and $|\psi_2(d) - \psi_2(0)|$ are all smaller than ϵ . Thus,

$$\begin{aligned} \sqrt{z}e^{-izh(t_0)} \int_J e^{izh(t)} g(t) dt &= \sqrt{z} \int_c^d e^{\pm izx^2} \psi(x) dx \\ &= \int_c^d \cos(zx^2) \psi_1(x) \sqrt{z} dx \pm i \int_c^d \sin(zx^2) \psi_1(x) \sqrt{z} dx \\ &\quad - \int_c^d \cos(zx^2) \psi_2(x) \sqrt{z} dx \mp i \int_c^d \sin(zx^2) \psi_2(x) \sqrt{z} dx. \end{aligned}$$

As in the proof of Lemma 7.7, each integral may be handled by the second mean value theorem for integrals and the first is typical. There is a point y between c and d such that

$$\begin{aligned} \int_c^d \cos(zx^2) \psi_1(x) \sqrt{z} dx &= \psi_1(c) \int_c^y \cos(zx^2) \sqrt{z} dx + \psi_1(d) \int_y^d \cos(zx^2) \sqrt{z} dx \\ &= \psi_1(c) \int_{c\sqrt{z}}^{y\sqrt{z}} \cos(u^2) du + \psi_1(d) \int_{y\sqrt{z}}^{d\sqrt{z}} \cos(u^2) du. \end{aligned}$$

Using the Fresnel integrals from the above supplementary material for Chapter 4, these integrals converge as z goes to $+\infty$. Since $c < 0 < d$, the limit is $\psi_1(d)\sqrt{\pi/2}$ if $y < 0$, is $\psi_1(c)\sqrt{\pi/2}$ if $y > 0$, and is $\{[\psi_1(c) + \psi_1(d)]/2\}\sqrt{\pi/2}$ if $y = 0$. But each of these is within $\epsilon\sqrt{\pi/2}$ of $\psi_1(0)\sqrt{\pi/2}$. Similar arguments for the other three integrals show that the whole sum converges to a limit that is

$$\psi_1(0)\sqrt{\frac{\pi}{2}} \pm i\psi_1(0)\sqrt{\frac{\pi}{2}} - \psi_2(0)\sqrt{\frac{\pi}{2}} \mp i\psi_2(0)\sqrt{\frac{\pi}{2}},$$

with an error of no more than $\epsilon\sqrt{\pi/2}$ in each term. Thus, we do get a limit that is no more than $2\epsilon\sqrt{2\pi}$ away from the point

$$[\psi_1(0) - \psi_2(0)](1 \pm i)\sqrt{\pi/2} = \psi(0)\sqrt{\pi}e^{\pm\pi i/4} = \sqrt{2\pi}g(t_0)e^{\pm\pi i/4}/\sqrt{\pm h''(t_0)},$$

just as desired. This completes the proof of Theorem 7.2.10. \blacksquare

Chapter 8

Laplace Transform and Applications

Fourier Transform and Wave Equation

The Fourier transform, which was introduced in §4.3, provides an alternative to the Laplace transform for solving differential equations. We illustrate this use and the role of complex variables by focusing on the wave equation. Our discussion will be somewhat informal and we shall forgo the rigorous formulation of theorems.

Wave Equation The wave equation is the equation of motion that describes the development of a wave disturbance propagating in a medium. It describes, for example, the vertical displacement of a vibrating string (see Figure 8.1), the propagation of an electromagnetic wave through space and of a sound wave in a concert hall, and some types of water wave motion.

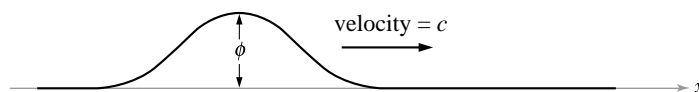


Figure 8.1: ϕ is the wave amplitude.

First consider the homogeneous problem, the simplest case of which is a wave traveling down a string of constant density ρ and under constant tension T . The vertical displacement $\phi(x, t)$ at position x and time t satisfies the *wave equation*

$$\frac{1}{c^2} \cdot \frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial x^2},$$

where $c = \sqrt{T/\rho}$ is the velocity of propagation, a constant. We accept this fact from elementary physics. (The derivation assumes that the amplitude is small.)

Note that if we were to have $c = \sqrt{-1}$ in the wave equation, we would recover the Laplace equation (see §2.5 and §5.3). Indeed, just as that equation admitted solutions of the form $f(x \pm \sqrt{-1}y)$, the solutions to the wave equation take the form $f(x \pm ct)$. The fact that the wave equation is of second order in the t variable suggests that a solution is uniquely given when two pieces of initial data at $t = 0$ are specified. These data consist of $\phi(x, 0)$ and $d\phi/dt$ at $(x, 0)$; the wave equation then gives the development of $\phi(x, t)$ for subsequent t .

To solve the wave equation, we perform a transform on the x variable to obtain a simpler equation involving the transform variable k . However, here x runs from $-\infty$ to $+\infty$, so instead of using the Laplace transform we use the Fourier transform. Let $f : \mathbb{R} \rightarrow \mathbb{C}$; the **Fourier transform** \hat{f} of f is defined by

$$\hat{f}(k) = \int_{-\infty}^{+\infty} e^{-ikx} f(x) dx.$$

There is an inversion formula that is analogous to the Laplace inversion formula:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ikx} \hat{f}(k) dk.$$

The Fourier transform of the function $\phi(x, t)$ is defined by

$$\hat{\phi}(k, t) = \int_{-\infty}^{+\infty} e^{-ikx} \phi(x, t) dx.$$

Here we perform the integral with respect to the x variable, regarding t as a fixed parameter. The Fourier inversion formula now reads

$$\phi(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ikx} \hat{\phi}(k, t) dk.$$

We are now ready to solve the wave equation. Taking the Fourier transform and differentiating under the integral, we obtain

$$\frac{1}{c^2} \cdot \frac{\partial^2 \hat{\phi}}{\partial t^2}(k, t) + k^2 \hat{\phi}(k, t) = 0.$$

In other words, our transformation technique has replaced the partial differential equation for $\phi(x, t)$ with an ordinary differential equation for $\hat{\phi}(k, t)$, which is easily solved. The solution is

$$\hat{\phi}(k, t) = A(k)e^{ikct} + B(k)e^{-ikct},$$

where $A(k), B(k)$ are two constants of integration that may depend on the parameter k . Applying the inversion formula, we get

$$\phi(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} [A(k)e^{ik(x+ct)} + B(k)e^{ik(x-ct)}] dk.$$

This is our solution to the wave equation. The functions $A(k), B(k)$ are determined by the initial data $\phi(x, 0)$ and $\partial\phi(x, 0)/\partial t$.

Note that the first integral in this solution depends only on the variable $x + ct$, whereas the second depends only on $x - ct$; that is, $\phi(x, t)$ has the form

$$\phi(x, t) = f(x + ct) + g(x - ct),$$

where the functions, f, g are again determined by ϕ and $\partial\phi/\partial t$ at $t = 0$. We can verify by substitution that this formula for ϕ does indeed give a solution of the wave equation.

Some special solutions deserve separate attention. These are monochromatic (single-frequency) waves and are of the form

$$\phi(x, t) = e^{i(x/c-t)\omega},$$

where ω is the frequency. This ϕ represents a wave of frequency ω traveling to the right down the string. Generally, $f(x + ct)$ is a wave moving to the left whose shape is that of the graph of f , with velocity c . Similarly, $g(x - ct)$ is a wave moving to the right.

Next we shall deal with the inhomogeneous problem, which occurs when an external force is applied to the wave. For example, suppose that the string illustrated in Figure 8.1 were given a constant charge density q and then placed in an external electric field $E(x, t)$ pointing in the y direction. This would result in the application to the string of a force $F(x, t)$ proportional to $qE(x, t)$. We must then solve the following equation of motion for the displacement $\phi(x, t)$, which is called the *inhomogeneous wave equation*:

$$\frac{1}{c^2} \cdot \frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial x^2} + F(x, t).$$

For simplicity, we take $F(x, t)$ to be periodic with frequency ω :

$$F = f(x, \omega)e^{i\omega t}.$$

This allows us to consider the simpler problem

$$\frac{1}{c^2} \cdot \frac{\partial^2 \phi}{\partial t^2} = \frac{\partial^2 \phi}{\partial x^2} + f(x, \omega)e^{i\omega t}.$$

We write the solution as $\phi(x, t, \omega)$. Once we have solved this simpler problem, we can deal with the problem of a general force $F(x, t)$. First we “Fourier-analyze” it; that is, we write F as follows:

$$F(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} f(x, \omega)e^{-i\omega t} d\omega,$$

where

$$f(x, \omega) = \int_{-\infty}^{+\infty} F(x, t)e^{it\omega} dt.$$

We then superimpose the solutions of the simpler problem to obtain

$$\phi(x, t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \phi(x, t, \omega) d\omega.$$

We solve the simpler inhomogeneous wave equation by taking the Fourier transform with respect to the variable x to get

$$\frac{1}{c^2} \cdot \frac{\partial^2 \hat{\phi}}{\partial t^2} + k^2 \hat{\phi} = \hat{f}(k, \omega) e^{i\omega t},$$

where

$$\hat{f}(k, \omega) = \int_{-\infty}^{+\infty} e^{ikx} f(x, \omega) dx.$$

This is a simple inhomogeneous second-order differential equation. Adding the particular solution

$$\frac{\hat{f}(k, \omega) e^{i\omega t}}{k^2 - (\omega/c)^2}$$

to the homogeneous solutions obtained earlier, we obtain the general solution

$$\hat{\phi}(t, \omega) = A(k) e^{ikct} + B(k) e^{-ikct} + \frac{\hat{f}(k, \omega)}{k^2 - (\omega/c)^2} e^{i\omega t}.$$

The solution is thus

$$\phi(x, t, \omega) = h(x + ct) + g(x - ct) + (G * f) e^{i\omega t}.$$

The terms in this equation are explained as follows. The first two terms are solutions to the homogeneous wave equation, and again they are to be chosen so that the initial data at $t = 0$ are satisfied. The last term, a particular solution to the inhomogeneous equation, is given by taking the inverse Fourier transform of the last term in the expression for $\hat{\phi}$:

$$\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ikx} \left[\frac{\hat{f}(k, \omega)}{k^2 - (\omega/c)^2} \right] dk.$$

As with the Laplace transform, this term is the convolution of G and f where $\hat{G} = 1/[k^2 - (\omega/c)^2]$. This function \hat{G} plays a central role in the theory of partial differential equations. Its transform,

$$G(x, \omega) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left[\frac{e^{ikx}}{k^2 - (\omega/c)^2} \right] dk,$$

is called the **Green's function**¹, and we can use contour integration to evaluate it in closed form as follows.

The integrand of $G(x, \omega)$ has simple poles at $k = \pm(\omega/c)$. In its present form, the integral is not convergent. To specify its value, we use the Cauchy principal value. Several possible values may be obtained depending on how we interpret our integrals. To select the value we want, we shall evaluate the integral by closing the contour of integration in the upper half of the complex k plane for $x > 0$ and in the lower half of the plane for $x < 0$. This is necessary if the integral over the semicircle is to approach zero as the radius approaches infinity. By Cauchy's Theorem, we pick up the residues of the enclosed poles. We still must specify how we are to go around the singularities at $k = \pm\omega/c$. Different choices will lead to different but still mathematically acceptable values of G . Our final choice is determined by the asymptotic behavior we want G to have as $x \rightarrow \infty$. The homogeneous solutions to the wave equation in which we are interested behave like $\exp(\pm ikx)$ as a function of x , and we will require the same behavior of G . This can be specified by the “ $i\epsilon$ prescription”:

$$G(x, \omega) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} \frac{1}{2\pi} \int_{-\infty}^{+\infty} \left[\frac{e^{ikx}}{k^2 - (\omega/c - i\epsilon)^2} \right] dk,$$

in which we still close the contour (as shown in Figure 8.2) according to the sign of x .

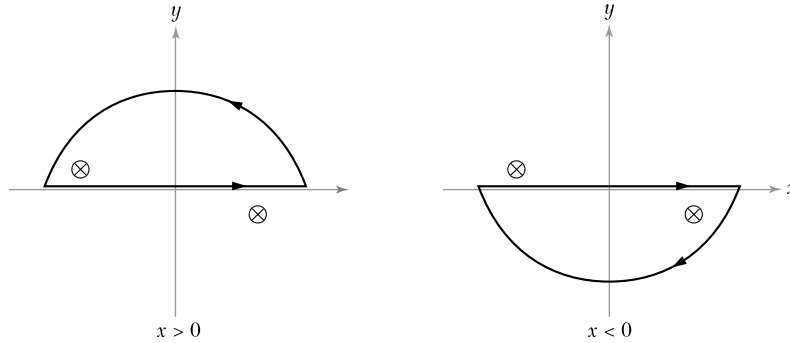


Figure 8.2: Contours for $G(x, \omega)$.

We can now evaluate $G(x, \omega)$. From the preceding equation and the Residue Theorem we obtain, for $x > 0$,

$$G(x, \omega) = \lim_{\epsilon \rightarrow 0, \epsilon > 0} \frac{1}{2\pi} \left[2\pi i \cdot \frac{e^{ix(i\epsilon + \omega/c)}}{-2(\omega/c - i\epsilon)} \right] = -\frac{ic}{2\omega} e^{i\omega x/c}.$$

¹Gramatically, the use of the term “the Green's function” is incorrect, just as it would be to say “the Cauchy's Theorem”, but it is, unfortunately, how it is commonly expressed.

Making a similar computation for $x < 0$ and using the contour at the right in Figure 8.2, we obtain

$$G(x, \omega) = \begin{cases} \frac{-ic}{2\omega} e^{i\omega x/c} & x > 0 \\ \frac{-ic}{2\omega} e^{-i\omega x/c} & x < 0 \end{cases}.$$

Equivalently,

$$G(x, \omega) = \frac{c}{2i\omega} e^{i\omega|x|/c}.$$

In textbooks on differential equations, G is often obtained as the solution to

$$\frac{d^2 G}{dx^2} + \omega^2 G = \delta(x - y),$$

where δ is the “Dirac δ function.” The solution is found by the general formula

$$G(x, y, \omega) = \begin{cases} -u(x)v(y)/w & x > y \\ -u(y)v(x)/w & x < y \end{cases}.$$

Here u, v are solutions of the corresponding homogeneous equations and w is their Wronskian, $w = uv' - vu'$. In this case, we have $u = e^{i\omega x}$ and $v = e^{-i\omega x}$. We recover the formula for $G(x, \omega)$ by setting $y = 0$.

Scattering Problem When the medium through which the wave propagates is not homogeneous, we encounter the scattering problem. For example, suppose that the vibrating string of Figure 8.1 now consist of three pieces smoothly joined together, with one piece, of length a , having a density of ρ_2 (region II in Figure 8.3) and the other two pieces each having a density of ρ_1 (region I, III in Figure 8.3).

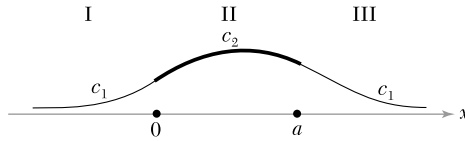


Figure 8.3: One-dimensional scattering.

Let c_1 and c_2 denote the corresponding velocities of propagation. Assume that $\rho_2 > \rho_1$. Imagine that an incident wave $e^{i(x/c_1 - t)\omega}$ from the left travels down the string. As the wave moves onto the denser material at $x = 0$, part of it will be reflected backward, while part will be transmitted onward. At $x = a$, some of the wave will again be reflected backward while the rest travels forward (see Figure 8.4).

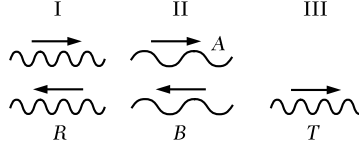


Figure 8.4: Reflection in one-dimensional scattering.

We must solve the wave equation in each region:

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c_1^2} \cdot \frac{\partial^2 \phi}{\partial t^2} \quad \text{for regions I, III}$$

$$\frac{\partial^2 \phi}{\partial x^2} = \frac{1}{c_2^2} \cdot \frac{\partial^2 \phi}{\partial t^2} \quad \text{for region II.}$$

It is not unreasonable to expect solutions of the following forms:

$$\begin{aligned} \phi_{\text{I}}(x, t) &= e^{i(x/c_1 - t)\omega} + R e^{i(-x/c_1 - t)\omega} \\ \phi_{\text{II}}(x, t) &= A e^{i(x/c_2 - t)\omega} + B e^{i(-x/c_2 - t)\omega} \\ \phi_{\text{III}}(x, t) &= T e^{i(x/c_1 - t)\omega}. \end{aligned}$$

We shall require that, at $x = 0$ and $x = a$, the solutions continuously join each other and have no sharp bend. Mathematically, this means that we impose the boundary conditions that ϕ and $\partial\phi/\partial x$ be continuous at $x = 0, a$. In other words,

$$\begin{aligned} \phi_{\text{I}}(0, t) &= \phi_{\text{II}}(0, t) & \phi_{\text{II}}(a, t) &= \phi_{\text{III}}(a, t) \\ \frac{\partial \phi_{\text{I}}}{\partial x}(0, t) &= \frac{\partial \phi_{\text{II}}}{\partial x}(0, t) & \frac{\partial \phi_{\text{II}}}{\partial x}(a, t) &= \frac{\partial \phi_{\text{III}}}{\partial x}(a, t). \end{aligned}$$

These four equations allow us to solve for the coefficients R, A, B, T . We are particularly interested in T . After performing some algebraic manipulations we find that

$$T = \frac{4c_1 c_2 e^{it[(1/c_2) - (1/c_1)]\omega a}}{(c_1 + c_2)^2 - (c_1 - c_2)^2 e^{2ia\omega/c_2}}$$

(see Exercise 5). The quantity T is called the **scattering amplitude**, and the square of its absolute value represents the intensity of the wave transmitted into region III.

We now allow ω in this equation to become a complex variable, and we see that $T(\omega)$ has the following property:

- (i) T is meromorphic in ω and has poles in the lower half plane at $\omega = (c_2/a)(n\pi - i\rho)$, where ρ is determined by $e^{2\rho} = (c_1 + c_2)^2 / (c_1 - c_2)^2$.
- (ii) T has absolute value 1 at those values of ω for which $e^{2i\omega/c_2} = 1$.
- (iii) As $\omega \rightarrow +i\infty$, $T \rightarrow 0$.
- (iv) As $\omega \rightarrow -i\infty$, $T \rightarrow \infty$.

Dispersion Relations When a function of a complex variable $f(z)$ shares the same four properties as $T(\omega)$, the Cauchy theorem can be used to obtain an interesting and useful representation for $f(z)$, as shown in the following.

Theorem 8.1 (Hilbert Transform Theorem) *If $f(z)$ is analytic for $\text{Im}(z) \geq 0$ and $f(z) \rightarrow 0$ uniformly as $z \rightarrow \infty$ in the half plane $0 < \arg z < \pi$, then $f(z)$ satisfies the following integral relationships.*

(i) *If $z_0 = x_0 + iy_0$ with $y_0 > 0$, then*

$$f(z_0) = \frac{y_0}{\pi} \int_{-\infty}^{+\infty} \frac{f(x, 0)}{(x_0 - x)^2 + y_0^2} dx$$

and

$$f(z_0) = \frac{1}{\pi i} \int_{-\infty}^{+\infty} \frac{f(x, 0)(x - x_0)}{(x - x_0)^2 + y_0^2} dx.$$

(ii) *If $z_0 = x_0$ is real, then*

$$f(z_0) = \frac{1}{\pi i} \int_{-\infty}^{+\infty} \frac{f(x, 0)}{x - x_0} dx.$$

Proof Because of the assumptions, we can apply Cauchy's Theorem, using a large semicircle in the upper half plane, to give

$$f(z_0) = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{f(x, 0)}{x - z_0} dx$$

(see §4.3). We also have

$$0 = \frac{1}{2\pi i} \int_{-\infty}^{+\infty} \frac{f(x, 0)}{x - \bar{z}_0} dx,$$

where \bar{z}_0 lies in the lower half plane. If we subtract the last two equations we obtain the first equation for $f(z_0)$; if we add them we get the second. The third follows from formula 6 of Table 4.2.1. ■

As a corollary, by taking real and imaginary parts of each side of the first equation for $f(z_0)$, we get

$$u(x_0, y_0) = \frac{y_0}{\pi} \int_{-\infty}^{+\infty} \frac{u(x, 0)}{(x - x_0)^2 + y_0^2} dx,$$

where $f = u + iv$; a similar equation holds for $v(x, y)$. From the second we have

$$\begin{aligned} u(x_0, y_0) &= \frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{(x - x_0)v(x, 0)}{(x - x_0)^2 + y_0^2} dx \\ v(x_0, y_0) &= -\frac{1}{\pi} \int_{-\infty}^{+\infty} \frac{(x - x_0)u(x, 0)}{(x - x_0)^2 + y_0^2} dx. \end{aligned}$$

Finally, from the third, we obtain

$$\begin{aligned} u(x_0, 0) &= \frac{1}{\pi} \text{P.V.} \int_{-\infty}^{+\infty} \frac{v(x, 0)}{x - x_0} dx \\ v(x_0, 0) &= \frac{-1}{\pi} \text{P.V.} \int_{-\infty}^{+\infty} \frac{u(x, 0)}{x - x_0} dx. \end{aligned}$$

The first formula for $u(x_0, y_0)$ gives us the values of a harmonic function in the upper half plane, in terms of its boundary values on the real axis, and thus provides a solution to the Laplace equation in the upper half plane.

Note that if $f(z)$ satisfies the symmetry property $\overline{f(-x)} = f(x)$, then we can write

$$\text{Re } f(x_0) = \frac{2}{\pi} \text{P.V.} \int_0^{\infty} \frac{\text{Im } f(x)}{x^2 - x_0^2} x dx,$$

whereas if $f(z)$ satisfies $\overline{f(-x)} = -f(x)$, then

$$\text{Re } f(x_0) = \frac{2x_0}{\pi} \text{P.V.} \int_0^{\infty} \frac{\text{Im } f(x)}{x^2 - x_0^2} dx.$$

These equations for u and v can be regarded as integral versions of the Cauchy-Riemann equations; they simply tell us, for example, the values that the real part of an analytic function must take when the imaginary part is specified. When functions u, v satisfy these equations, we say that u and v are ‘‘Hilbert transforms’’ of each other. Historically, the Hilbert transforms were the forerunners of a series of such relations called dispersion relations. They were first observed to hold for the complex dielectric constant as a function of incident frequency by H. A. Kramers and R. de L. Kronig in 1924. Since approximately 1950, they have been systematically studied and applied to the scattering amplitude $T(\omega)$ and to quite general classes of scattering problems for which this amplitude is defined. The extension of these relations to three-dimensional scattering problems will be considered later in this supplement.

The relations derived assumed that $f(z)$ is analytic only for $\text{Im } z \geq 0$. However, there is a second class of dispersion relations for functions that are analytic in the z plane except for a branch line along the real axis.

Proposition 8.2 *If $f(z)$ is analytic in the z plane with a branch line from $z = a$ to ∞ , and if $|f(z)| = O(1/z)$, then*

$$f(z) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{2\pi i} \int_a^{\infty} \frac{1}{x - z} [f(x + i\epsilon) - f(x - i\epsilon)] dx.$$

(The notation $O(1/z)$ is explained in §7.2 and $\lim_{\epsilon \rightarrow 0^+}$ means the limit is taken through $\epsilon > 0$.)

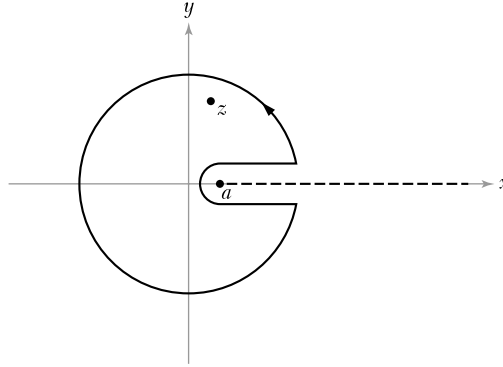


Figure 8.5: Contour used in the proof of the preceding proposition.

Proof Take the contour of Figure 8.5 and apply the Cauchy Theorem to the function $f(\zeta)/(\zeta - z)$. ■

If, in addition to the hypotheses of the preceding proposition, $f(z)$ also satisfies the relation $\overline{f(\bar{z})} = f(z)$, that is, if $u(x, \epsilon) + iv(x, \epsilon) = u(x, -\epsilon) - iv(x, -\epsilon)$, so the real part of f is continuous across the real axis whereas the imaginary part is discontinuous, then we obtain

$$f(z) = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \int_a^\infty \frac{1}{x - z} \operatorname{Im} f(x + i\epsilon) dx.$$

When z actually moves onto the real axis, we can take real parts of each side of the preceding equation to obtain

$$\operatorname{Re} f(x_0) = \lim_{\epsilon \rightarrow 0^+} \frac{\text{P.V.}}{\pi} \int_a^{+\infty} \frac{1}{x - x_0} \operatorname{Im} f(x + i\epsilon) dx.$$

Wave Equation in Three Dimensions The ideas that have been developed thus far in this section for wave motion in one dimension can easily be extended to higher-dimensional problems. In two dimensions the vibrating string is replaced by a vibrating membrane. In three dimensions we can think of sound waves propagating in air. The pressure $\phi(\mathbf{r}, t)$ then satisfies the equation of motion,

$$\frac{1}{c^2} \cdot \frac{\partial^2 \phi}{\partial t^2}(\mathbf{r}, t) = \nabla^2 \phi(\mathbf{r}, t) + F(\mathbf{r}, t),$$

where F represents some external source of waves, $\mathbf{r} = (x, y, z)$, and

$$\nabla^2 = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2}$$

is the Laplace operator. When the equation of motion is expressed in terms of rectangular coordinates, its homogeneous and inhomogeneous solutions are obtained

in much the same way they were previously. The really new and exciting features not present in one dimension arise in the scattering problem, and these are most interesting and tractable when the scattering medium has spherical symmetry.

Consider an incident plane wave $e^{i(x/c - \omega)t}$ traveling from the left down the x axis that impinges on a ball located at the origin (Figure 8.6). Part of the wave may penetrate the ball, part of the wave is scattered by the surface of the ball and then travels radially outward, and the remainder of the wave simply bypasses the ball. To solve for ϕ , we proceed as previously. We first obtain the solutions for ϕ in region I and region II separately and then require that ϕ and the radial derivative $\partial\phi/\partial r$ be continuous at the surface of the ball. This procedure eventually specifies the total wave in region I that results from the “impurity” of the medium in region II.

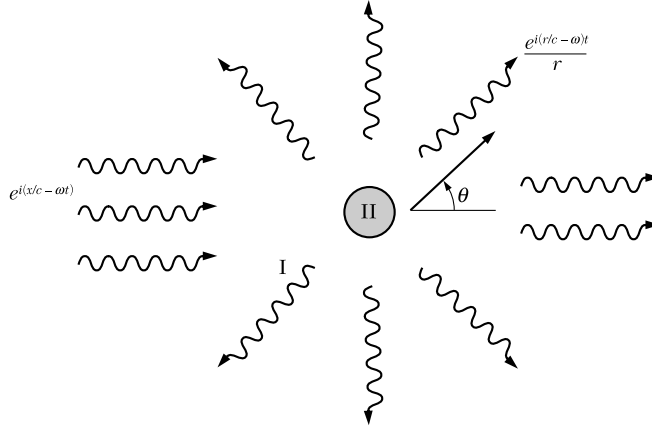


Figure 8.6: Scattering of a wave by a spherical obstacle.

These calculations are not detailed here, because such a task would take us too far afield into the subject of partial differential equations. However, the form of the final result is not too difficult to anticipate. The wave in region I will be a sum of incident wave and the outgoing radial wave, and this will take the asymptotic form

$$\phi_{\text{I}}(\mathbf{r}, t) \sim \left[e^{ix/c} + \frac{e^{ir/c}}{r} f(\omega, \theta) \right] e^{-i\omega t}$$

as $|\mathbf{r}| = r \rightarrow \infty$. Here $f(\omega, \theta)$ is the amount of scattering wave that is traveling outward at an angle θ to the axis of symmetry (see Figure 8.6); it is the three-dimensional analogue of the scattering amplitude $T(\omega)$ of the one-dimensional problem discussed earlier in the section.

Note that f is now a function of *two* complex variables, ω and θ . Physically observable scattering occurs, of course, only when ω and θ are real with $0 \leq \theta \leq \pi$. But by studying the properties of f for complex values of ω and θ , as we will do in

the following paragraphs, we do gain a deeper understanding of the characteristics of f .

It will be convenient to change variables as follows:

$$s = \frac{4\omega^2}{c^2} \quad t = -2\frac{\omega^2}{c^2}(1 - \cos\theta).$$

Define the function A of the two complex variables s, t by $A(s, t) = f(\omega, \theta)$. For a large class of scattering problems, it can be shown that A has the following properties:

- (i) $A(s, t)$ is analytic in the two complex variables s and t with branch lines from $s = a$ to ∞ and $t = b$ to ∞ .
- (ii) $A(s, t) = O(1/s)$ as $s \rightarrow \infty$ for each t .
- (iii) $A(\bar{s}, t) = \overline{A(s, t)}$ for t real.

Applying the preceding proposition,

$$A(s_0, t_0) = \begin{cases} \lim_{\epsilon \rightarrow 0^+} \frac{1}{2\pi i} \int_a^{\infty} \frac{1}{s - s_0} [A(s + i\epsilon, t_0) - A(s - i\epsilon, t_0)] ds & s_0, t_0 \text{ in } \mathbb{C} \\ \lim_{\epsilon \rightarrow 0^+} \frac{1}{\pi} \int_b^{\infty} \frac{1}{s - s_0} \text{Im } A(s + i\epsilon, t_0) ds & t_0 \text{ real} \end{cases}.$$

In the first equation for $A(s_0, t_0)$, we are integrating $A(\zeta, t)$ for ζ slightly above and below the real axis. Now consider the integrand. By property (i) of $A(s, t)$ we can write a dispersion relation in the t variable as follows:

$$A(s + i\epsilon, t_0) = \lim_{\delta \rightarrow 0^+} \frac{1}{2\pi i} \int_b^{\infty} \frac{1}{t - t_0} [A(s + i\epsilon, t + i\delta) - A(s + i\epsilon, t - i\delta)] dt.$$

Using a similar representation for the second term in the first equation for $A(s_0, t_0)$, we finally obtain the **double dispersion relation**

$$A(s_0, t_0) = \frac{1}{\pi^2} \int_a^{\infty} \frac{1}{s - s_0} \left[\int_b^{\infty} \frac{1}{t - t_0} \rho(s, t) dt \right] ds,$$

where

$$\rho(s, t) = \lim_{\epsilon, \delta \rightarrow 0^+} \left(\frac{1}{2i} \right)^2 [A(s + i\epsilon, t + i\delta) - A(s + i\epsilon, t - i\delta) - A(s - i\epsilon, t + i\delta) + A(s - i\epsilon, t - i\delta)].$$

The equation for $A(s_0, t_0)$ is a representation first obtained by S. Mandelstam in 1958.

Exercises for Supplement to Chapter 8

1. Apply the boundary conditions

$$\begin{aligned}\phi_{\text{I}}(0, t) &= \phi_{\text{II}}(0, t) & \phi_{\text{II}}(a, t) &= \phi_{\text{III}}(a, t) \\ \frac{\partial \phi_{\text{I}}}{\partial x}(0, t) &= \frac{\partial \phi_{\text{II}}}{\partial x}(0, t) & \frac{\partial \phi_{\text{II}}}{\partial x}(a, t) &= \frac{\partial \phi_{\text{III}}}{\partial x}(a, t)\end{aligned}$$

to obtain expressions for $T(\omega)$, $R(\omega)$, $A(\omega)$, $B(\omega)$. Verify that when ω is real,

$$|T(\omega)|^2 + |R(\omega)|^2 = 1.$$

(This relation expresses “conservation of intensity”: In a medium without dissipative loss, the unit intensity of the incident wave of Figure 8.4 is equal to the sum of the intensity of the wave R reflected back into region I and intensity of the wave T transmitted into region III.)

2. Since

$$f(z_0) = \frac{y_0}{\pi} \int_{-\infty}^{+\infty} \frac{f(x, 0)}{(x_0 - x)^2 + y_0^2} dx$$

solves the Dirichlet problem for the upper half plane, apply it to the problem illustrated at the right in Figure 5.3.5 of the text and obtain the same solution as in Example 5.3.1. Thus, we have two methods for solving Laplace’s equation: conformal mapping and integral relations of the type described by the preceding displayed equation.

3. As an application of the Hilbert Transform, let

$$f(z) = \frac{1}{\sqrt{z}} = r^{-1/2} e^{-i\theta/2} \quad \text{for } 0 < \theta < 2\pi$$

be defined on the z plane with a branch line along the positive real axis. What equality results?

4. As yet another way to solve the Laplace equation, take Fourier transforms of

$$\frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y) = 0 \tag{8.1}$$

with respect to the variable x (use the method described in the Wave Equation subsection); show that this method gives the ordinary differential equation

$$\frac{\partial^2 \hat{u}}{\partial y^2}(k, y) - k^2 \hat{u}(k, y) = 0. \tag{8.2}$$

By solving, summing over all solutions, and keeping only those that decay exponentially in the limit $y \rightarrow +\infty$, obtain the formula

$$u(x, y) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(k) e^{ikx - k/y} dk. \tag{8.3}$$

With relative ease, $A(k)$ can now be evaluated in terms of the boundary data at $y = 0$. Since at $y = 0$, equation (8.2) reduces to

$$u(x, 0) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} A(k) e^{ikx} dk,$$

$A(k)$ must in fact be the Fourier transform of $u(x, 0)$. Using this result, substituting in equation (8.2), and interchanging orders of integration, obtain

$$u(x, y) = \int_{-\infty}^{+\infty} u(z, 0) \left(\frac{1}{2\pi} \int_{-\infty}^{+\infty} e^{ik(x-z)-|k|/y} dk \right) dz.$$

Perform the integral on k to get

$$u(x, y) = \int_{-\infty}^{+\infty} u(z, 0) \cdot \frac{1}{\pi} \cdot \frac{1}{(x-z)^2 + y^2} dz. \quad (8.4)$$

Referring to equation (8.4), show that

$$\frac{y}{(x-z)^2 + y^2} = \frac{-\partial}{\partial w} G(x, y|z, w) \Big|_{w=0},$$

where

$$G(x, y|z, w) = \log |\mathbf{r} - \mathbf{r}'|$$

with $\mathbf{r} = x + iy$, $\mathbf{r}' = z + iw$. (The function G is the Green's function and equals the potential at \mathbf{r} caused by a unit charge at the point \mathbf{r}' in the plane.)

5. Many problems in applied mathematics include an infinite series

$$F(z) = \sum_{n=0}^{\infty} f_n(z)$$

in which the function $F(z)$ has singularities in addition to all those of each $f_n(z)$. These singularities are introduced by the failure of the series to converge. The simplest example of this type of series is

$$\frac{1}{1-z} = 1 + z + z^2 + \dots$$

The individual terms on the right of the equation are entire functions; the function their sum defines has a simple pole at $z = 1$. As a second example, consider

$$\frac{z}{z-1} = 1 + \frac{1}{z} + \frac{1}{z^2} + \dots$$

The terms on the right of this equation have singularities at $z = 0$; their sum is singular at $z = 1$.

Consider

$$F(z) = \sum_{n=0}^{\infty} [g \log(z - \alpha)]^n,$$

where α, g are constants.

- (a) What are the analytic properties of the individual terms on the right side of the last equation?
- (b) By summing the series in closed form, verify that the sum in this equation has a pole at $z = \alpha + e^{1/g}$.

Comments on Selected Exercises

1. Substituting the expressions for $\phi_I, \phi_{II}, \phi_{III}$ into the boundary conditions that follow, we have

$$1 + R = A + B \quad (8.5)$$

$$1 - R = \frac{c_1}{c_2}(A - B) \quad (8.6)$$

$$Ae^{ia\omega/c_2} + Be^{-ia\omega/c_2} = Te^{ia\omega/c_1} \quad (8.7)$$

$$Ae^{ia\omega/c_2} - Be^{-ia\omega/c_2} = \frac{c_2}{c_1}Te^{ia\omega/c_1}. \quad (8.8)$$

Solve Equations (8.7) and (8.8) to get

$$A = T \left(\frac{c_1 + c_2}{2c_1} \right) e^{ia\omega(1/c_1 - 1/c_2)} \quad (8.9)$$

$$B = T \left(\frac{c_1 - c_2}{2c_1} \right) e^{ia\omega(1/c_1 + 1/c_2)}. \quad (8.10)$$

Adding equations (8.5) and (8.6) gives

$$2 = A \left(\frac{c_1 + c_2}{2} \right) + B \left(\frac{c_2 - c_1}{c_2} \right). \quad (8.11)$$

Now substitute (8.9), and (8.10) into (8.11) to obtain the expression for T in the text. If we subtract (8.5) and (8.6), we have

$$2R = A \left(\frac{c_2 - c_1}{c_2} \right) + B \left(\frac{c_2 + c_1}{c_2} \right). \quad (8.12)$$

Substituting (8.9) and (8.10) into (8.12) gives R in terms of T ,

$$R = \left(\frac{T}{2ic_1c_2} \right) (c_2^2 - c_1^2) e^{ia\omega/c_1} \sin \left(\frac{a\omega}{c_2} \right).$$

We can now obtain the desired result. We have

$$|T|^2 + |R|^2 = |T|^2 \left[1 + \frac{(c_1^2 - c_2^2)^2}{4c_1^2c_2^2} \sin^2 \left(\frac{a\omega}{c_2} \right) \right]. \quad (8.13)$$

The denominator of $|T|^2$ is

$$\begin{aligned} & (c_1 + c_2)^2 + (c_1 - c_2)^4 - (c_1^2 - c_2^2)^2 (e^{2ia\omega/c_2} + e^{-2ia\omega/c_2}) \\ &= (c_1 + c_2)^4 + (c_1 - c_2)^4 - (c_1^2 - c_2^2)^2 \cdot 2 \cdot \left(1 - 2 \sin^2 \left(\frac{a\omega}{c_2} \right) \right) \\ &= [(c_1 + c_2)^2 - (c_1 - c_2)^2]^2 + 4(c_1^2 - c_2^2)^2 \sin^2 \left(\frac{a\omega}{c_2} \right) \\ &= 4 \left(4c_1^2c_2^2 + (c_1^2 - c_2^2)^2 \sin^2 \left(\frac{a\omega}{c_2} \right) \right). \end{aligned}$$

Thus,

$$|T|^2 = \frac{4c_1^2c_2^2}{4c_1^2c_2^2 + (c_1^2 - c_2^2)^2 \sin^2 \left(\frac{a\omega}{c_2} \right)}.$$

This, with Equation (8.13), gives the desired result.

3. By Proposition 8.3.5,

$$f(z_0) = \lim_{\epsilon \rightarrow 0} \frac{1}{2\pi i} \int_0^\infty \frac{dx}{x - z_0} [f(x + i\epsilon) - f(x - i\epsilon)],$$

where

$$f(z) = \frac{1}{\sqrt{z}} = \frac{e^{-i\theta/2}}{\sqrt{r}}, \quad 0 < \theta < 2\pi.$$

Now

$$f(x + i\epsilon) \rightarrow \frac{1}{\sqrt{x}} \quad \text{as } \epsilon \rightarrow 0+$$

and

$$f(x - i\epsilon) \rightarrow -\frac{1}{\sqrt{x}} \quad \text{as } \epsilon \rightarrow 0+,$$

so the identity which results is

$$\frac{1}{\sqrt{z_0}} = f(x_0 + iy_0) = \frac{1}{\pi i} \int_0^\infty \frac{dx}{x - z_0} \frac{1}{\sqrt{x}} = -\frac{i}{\pi} \int_0^\infty \frac{dx}{x - z_0} \frac{1}{\sqrt{x}}.$$

This can be verified directly by changing variables using $w^2 = x$.

$$-\frac{i}{\pi} \int_0^\infty \frac{dx}{x - z_0} \frac{1}{\sqrt{x}} = -\frac{2i}{\pi} \int_0^\infty \frac{dw}{w^2 - z_0} = -\frac{i}{\pi} \int_{-\infty}^\infty \frac{dw}{w^2 - z_0}.$$

This has poles at $\pm\sqrt{z_0}$. If z_0 is not positive real, one is in the upper half plane and one in the lower with residues $\pm 1/(2\sqrt{z_0})$. The last expression becomes

$$-\frac{i}{\pi} \int_{-\infty}^\infty \frac{dw}{w^2 - z_0} = \left(-\frac{i}{\pi}\right) (2\pi i) \left(\frac{1}{2\sqrt{z_0}}\right) = \frac{1}{\sqrt{z_0}}$$

as before.

4. We have

$$G(x, y|z, w) = \log[(x - z)^2 + (y - w)^2]^{1/2}.$$

Then

$$\frac{\partial G}{\partial w} = \frac{w - y}{(x - z)^2 + (y - w)^2}$$

and hence $-\frac{\partial G}{\partial w}\big|_{w=0}$ is the desired expression. Thus,

$$u(x_0, y_0) = \frac{y_0}{\pi} \int_{-\infty}^{+\infty} \frac{u(x, 0)}{(x - x_0)^2 + y_0^2} dx,$$

becomes

$$u(x, y) = - \int_{-\infty}^{+\infty} u(z, 0) \frac{1}{\pi} \frac{\partial G}{\partial w}(x, y|z, 0).$$

This expression for the solution to Laplace's equation is a special case of the Green's function solution obtained in courses on partial differential equations.

5. Each term of the series has a logarithmic singularity at $z = \alpha$. Provided $|g \log(z - \alpha)| < 1$, we can use the geometric series to exactly sum the series as

$$F(z) = \frac{1}{1 - g \log(z - \alpha)}.$$

The denominator is 0 at $z = \alpha + e^{1/g}$, but its derivative, $-g/(z - \alpha)$, is not 0 there. Thus, there is a pole of order 1 at $z = \alpha + e^{1/g}$.