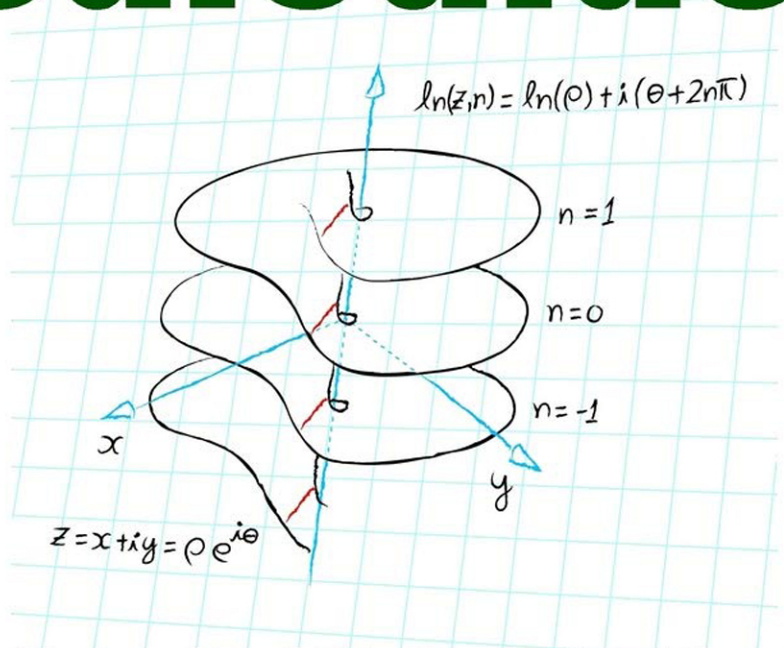
Complex Calculus



Mathematical Methods for Physics and Engineering - Volume 1

Jorge L. deLyra

Preface

The teaching of mathematics to physics students is a longstanding problem. The main reason for this is that there is an enormous amount of mathematics that must be learned by physicists. Unlike what may happen in the case of the professional work of a mathematician, the physicist cannot be limited to studying a few specific areas of mathematics and, instead of that, he must become able and fluent in a large number of different and varied areas, since almost all mathematics has important and, in fact, many times essential, uses in physics. Mathematics is for the physicist, first and foremost, the language in which science is formulated, thought, discussed, developed, recorded and transmitted.

On the other hand, since mathematics is just a part of the learning necessities of a physicist, and not his central area of dedication, usually there is very little time available for the absorption of all this universe of mathematics. This often leads to an excessively superficial approach to the mathematics in its teaching to physics students, which also entails significant losses, since mathematics is used in physics in an essential way as a language. This conflict between extension and depth exists even within the rather restrictive scope of what is traditionally described as the "mathematical methods of physics", which usually already includes a rather limited selection of topics, centered on those traditionally considered of greater relevance. This text intends to present an approach that tries to track what could be called the "middle way" in this conflict.

This text is the result of several years of experience of the author teaching the Mathematical Physics courses at the Physics Institute of the University of São Paulo. These are disciplines that are offered to all the undergraduate students of the Institute, starting usually at the beginning of the third year of the undergraduate curriculum. Occasionally, good students may be able to take it already in the second term of the second year. The presentation of the content assumes, as preliminary material, the basic physics introductory courses, as well as those of calculus and linear algebra, that are usually taken during the first two years. The typical level of preliminary preparation of the students was one of the factors taken into consideration for the elaboration of the text, and the level of abstraction adopted is compatible with a characteristically undergraduate course. One of the important objectives of the text is to serve the mathematical needs of the more advanced physics courses of the undergraduate physics curriculum.

The content included, as well as the approach adopted, are the result of an effort at updating and reorganizing these courses, which are traditionally part of the central core of all undergraduate physics curricula, aiming at a more modern presentation, and also at an adaptation of the content to the possibilities and limitations of a pair of one-semester courses. The text is organized in several volumes. Experience shows that it is possible to go through the contents of the first three in one semester: "Complex Calculus", "Fourier Transforms" and "Differential Equations". The volume on differential equations is limited to the use of Cartesian coordinates. The subsequent volumes complete a second semester, dedicated to the study of differential equations in cylindrical and spherical coordinates, as well as to the introduction of the corresponding special functions, in two volumes: "Bessel Functions" and "Spherical Harmonics". The text is organized in the form of relatively short chapters, each appropriate for exposition in one lecture.

The text is first and foremost a text on physics, which focuses attention on the mathematical tools needed for the treatment of more complex problems in physics. By and large, a considerable influence of the books by Ruel V. Churchill on complex analysis, Fourier series and boundary value problems, which are referred to in the bibliography, is reflected in the text. The mathematics is treated seriously and as completely as possible within the time limitations, but the motivation is predominantly the application of this mathematics to physics. Therefore, whenever possible the mathematics is approached in the context of physics applications, with the widest possible use of examples and motivations with a geometrical character, as well as of analogies with topics of physics. For example, the study of analytic functions is developed with the aid of an analogy with electrostatics and its scalar and vector fields, by means of the use of tools of real calculus in two dimensions.

Many theorems are proven along the text, with varied levels of rigor and generality; however, not all results presented are proven. The criteria for the choice of what to prove or not include the effort and time required, the mathematical complexity inherent to the topic, the level of difficulty of the mathematics involved, and the usefulness of the proofs in promoting a better understanding of the structures by the students. The emphasis in the presentation of the mathematics is the clear exhibition and understanding of the structures and procedures involved, with an algorithmic approach aiming at their utilization in the applications, as well as its use as a language in physics, and not the aspects of rigor and generality. The theorems of real analysis and linear algebra which are assumed known, and which are used without proof, are explicited in the text. Some of them are also discussed in an appendix.

Each chapter includes a list of proposed problems, which have varied levels of difficulty, including exercises and practice problems, problems that complete and extend the material presented in the text, and some longer and more difficult problems, which are presented as challenges to the students. These longer and more difficult problems are marked as "challenge problems". Some of these challenge problems involve proofs, but these proofs are not considered a central part of the material, in the sense of what should be required in tests and exams. The intent of these problems is to give even to the best students something non-trivial to face and against which to test their abilities. The results of some of the problems are used in later parts of the text, and such problems are marked as "reference problems". A high priority should be given to such problems in recitation classes. There are complete solutions available, detailed and commented, to all the problems proposed, which are presented in separate volumes, organized according to the sequence of corresponding chapters.

Although the text itself does not include an explicitly computational approach, the presentation is such that it allows for a relatively direct connection with the algorithmic and computational aspects. For example, the topic of Fourier transforms and series is developed from a discrete realization on a finite lattice, which is directly applicable to calculations of a computational nature. Starting from this structure, one then takes the continuum limit in order to derive the usual structures of the Fourier series and transforms, with various types of boundary conditions. Due to the fact that it treats the mathematics with a good level of precision, as well as with an appropriate level of rigor, keeping at the same time a practical and geometrical character, this text can be used both by physics students and by students of other areas, such as applied mathematics and engineering.

This Volume

This volume is dedicated to the complex calculus. This is a more practical and less abstract version of complex analysis and of the study of analytic functions. This does not mean that there are no proofs in the text, since all the fundamental theorems are proved with a good level of rigor. The text starts from the very beginning, with the definition of complex numbers, and proceeds up to the study of integrals on the complex plane and on Riemann surfaces. The facts and theorems established here will be used routinely in all the subsequent volumes of this series of books. The development is based on an analogy with vector fields and with electrostatics, emphasizing interpretations and proofs that have a geometrical character. The approach is algorithmic and emphasizes the representation of functions by series, with detailed discussion of the convergence issues.

Acknowledgements

The author would like to thank the contributions of the several teaching assistants that acted in the courses on Mathematical Physics under his supervision, as well as to the many students that took the courses, throughout the years of development of this text, and that in a way played the role of guinea pigs during that development. Without such able and willing test subjects, it would not have been possible to develop the text to the considerable extent which was achieved.

Many students and teaching assistants contributed by finding errors in the text and in the equations, both in the text itself and in the solutions to the problems, but the main contribution of the students is always to read and use the text with a critical mind-set, as well as to ask intelligent and challenging questions, of which there were many. This certainly contributed to add both quality and scope to the text.

Contents

1 Number, the Language of Science 1

2 The Simplicity of Complex Numbers 15

3 Elementary Functions, but Not Quite 33

4 Even Less Elementary Functions 49

5 Geometrical Aspects of the Functions 65

6 Border Effects in Capacitors 81

7 Complex Calculus I: Differentiation 99

8 Complex Calculus II: Integration 115

9 Complex Derivatives and Integrals 137

10 Complex Inequalities and Series 153

11 Series, Limits and Convergence 169

12 Representation of Functions by Series 187

13 Convergence Criteria and Proofs 205

14 Laurent Series and Residues 227

15 Calculation of Integrals by Residues 245

16 Residues on Riemann Surfaces 267

Appendices

A Continued Fractions 285

B Series for the Square Roots of Integers 291

C The Stirling Approximation 299

D Some Facts from Real Analysis 305

E Complete Singularity Classification 311

F The Complex Gamma Function 319

Chapter 1

Number, the Language of Science

Mathematics is the language in which natural science is formulated [1]. In particular, the use of numbers is essential for the representation of the quantitative aspects of science. In physics we are used to the utilization of the field of real numbers as a basic tool for the description of nature. The numbers are used not only in the quantitative description of experimental results but also in the mathematical structure of the physical theories, the function of which is to describe the relations that exist among those experimental results, thus leading to the understanding of nature.

The precise mathematical definition of a *field* $\mathbb{K}$ consists of a set of numbers and of two arithmetic operations defined on those numbers, the addition or sum and the multiplication or product. The existence of these two operations is essential for the majority of the physical applications. The set is a field if the two operations can be defined for all the pairs of elements of the set and have the properties that follow.

Operation of Addition:

Closure:

$a, b \in \mathbb{K} \Longrightarrow a + b \in \mathbb{K}.$

Existence of the identity:

$\exists\, 0 \in \mathbb{K} \;/\; a + 0 = a,\ \forall a \in \mathbb{K}.$

Existence of the inverse:

$\exists (-a) \in \mathbb{K} \;/\; a + (-a) = 0,\ \forall a \in \mathbb{K}.$

Commutativity:

$a, b \in \mathbb{K} \Longrightarrow a + b = b + a.$

Associativity:

$a, b, c \in \mathbb{K} \Longrightarrow (a + b) + c = a + (b + c).$

Operation of Multiplication:

Closure:

$a, b \in \mathbb{K} \Longrightarrow ab \in \mathbb{K}.$

Existence of the identity:

$\exists\, 1 \in \mathbb{K} \;/\; 1a = a,\ \forall a \in \mathbb{K}.$

Existence of the inverse:

$\exists\, a^{-1} \in \mathbb{K} \;/\; aa^{-1} = 1,\ \forall a \in \mathbb{K}$ except $a = 0.$

Commutativity:

$a, b \in \mathbb{K} \Longrightarrow ab = ba.$

Associativity:

$a, b, c \in \mathbb{K} \Longrightarrow (ab)c = a(bc).$

Distributivity of the product with respect to the sum:

$a, b, c \in \mathbb{K} \Longrightarrow a(b + c) = ab + ac.$

All the real arithmetics that we use every day in physics can be derived from this set of basic properties, which holds for the field $\mathbb{R}$ of real numbers. They are essential to allow us to interpret notions of quantification or measurements in terms of the real numbers, or possibly of some subset of the real numbers.

We can consider using the operations and properties listed above to make an explicit *construction* of the set of all real numbers starting from a simpler set that we can assume to be intuitively understood *a priori*. A very simple set, for example, is the infinite but discrete set $\mathbb{N}$ of the natural numbers, that is, the positive integers, which is a subset of the field of real numbers. The set $\mathbb{N}$ is not, however, a field, and it will be necessary to enlarge it in order to obtain a set that would satisfy all the properties of a field.

Taking this simple and intuitive set as the initial set of our construction, we impose the existence of an identity element of the addition, and that forces us to add the element 0 (zero) to the set. Besides, if we impose the existence of the inverse of the sum for these numbers, then we are led to add to the set all the negative integers. In this way we obtain the set $\mathbb{I}$ of all the integer numbers, in which all the properties related only to the addition hold. However, we still no not have a field here, because not all the properties related to the multiplication hold within $\mathbb{I}$.

We can now continue the construction by imposing the existence of the inverse of the product of these numbers, and that leads us to introduce the integer fractions and, eventually, to the set of all the numbers that can be expressed as p/q where p and q are integers, that is, as the product of an integer by the inverse of some other integer. This takes us to the set $\mathbb{Q}$ of the rational numbers, that satisfies in fact *all* the properties of a field, including the closure properties. Therefore we already have here a field, the field of rational numbers.

This field also has the property of being *ordered*, that is, given two different numbers of the set, it is always possible to decide which one of them is greater than the other, a property which also holds for the real numbers. We say that the pairs of numbers have a property of trichotomy: given $a, b \in \mathbb{Q}$, one and only one of three things holds: $a > b$, $a = b$ or $a < b$. Since in practice we are always limited to a finite amount of precision in physical measurements, the field of the rational numbers is in fact sufficient to describe all the possible results of experiments and the relations among them. However, this is often not the *simplest* way to proceed.

Since we already obtained a field, for a moment it may seem that at this point we have completed our construction. However, we can in fact continue the construction so as to obtain fields larger than $\mathbb{Q}$ that contain it as a subset. If we try to represent geometrically the rational numbers, putting them over a straight line, we can see that they do not fill the line completely, because there are points on the line, which can be defined by means of purely geometrical arguments, that cannot be expressed as ratios of two integers. In fact, it can be shown that the set $\mathbb{Q}$ is denumerable, that is, that it has the same number of elements as the set $\mathbb{N}$. On the other hand, the same cannot be said about the real numbers, which has infinitely more elements than either $\mathbb{Q}$ or $\mathbb{N}$.

One can show that the set $\mathbb{Q}$ fills the real line in a way that, generally speaking, is sufficient for physics applications, but is still somewhat limited from the mathematical point of view: once we put all the rational numbers on the line, there is no empty interval with non-zero length left, that is, there are no finite holes left. There is, however, a very large infinite set of infinitesimal holes, of length zero. We say that the set $\mathbb{Q}$ is *dense* in the real line, but that it does not complete it, that is, it is not a *complete* set.

It is relatively simple to prove this property of denseness by contradiction (reductio ad absurdum). In order to do this, let us imagine that there is an interval of the line inside which there is no rational number. Let us say that this interval is $(r_\ominus, r_\oplus)$, were we take the open interval, that is, only the interior, without the two ends, since we do not know whether or not the real numbers $r_\ominus$ or $r_\oplus$ are rational. The length of this interval is the real number $\Delta = r_\oplus - r_\ominus > 0$. We may assume that this interval cannot be enlarged, since otherwise it would suffice to enlarge it to the maximum possible length before starting the argument. This enlargement cannot ever make the interval infinite, due to the existence of the integers, which are rational. In fact, due to this we know beforehand that $\Delta < 1$.

Certainly there is a rational number $q_\oplus$ larger than $r_\oplus$, for example the first integer greater than $r_\oplus$. The same is true for a rational number $q_\ominus$ which is less than $r_\ominus$. Surely it is possible to choose $q_\oplus$ and $q_\ominus$ such that they are at a shorter distance than, for example, $\Delta/10$ from the respective ends of the interval, otherwise it would be possible to increase the interval. It now suffices to consider the arithmetic average of $q_\oplus$ and $q_\ominus$,

$$q = \frac{q_\oplus + q_\ominus}{2},$$

which is also a rational number, since $q_\oplus$, $q_\ominus$ and 2 are rational and the rational numbers form a field. It is not difficult to see, from the geometry of the problem, that q is contained in the original interval, which is absurd because this interval by hypothesis should not contain any rational number. Thus, we show that there can be no interval of non-zero length Δ that contains no rational number.

The precise definition of a dense subset of the real line is that, given any point on it and a distance δ, it is always possible to find an element of $\mathbb{Q}$ at a distance less than δ from that point. From the above argument, it follows that $\mathbb{Q}$ is a dense subset of the line. However, not every point that one can define on the line belongs in fact to $\mathbb{Q}$. A traditional example of this is the number $\sqrt{2}$, the length of the diagonal of the unit square, which is therefore defined by a geometric construction. We can show that $\sqrt{2}$ is not rational using the fact that the decomposition of an integer number into prime factors is unique. Let us first discuss the latter result. If an integer n is decomposed into a product of prime factors p_i

$$n = p_1^{l_1} \dots p_j^{l_j},$$

where all p_i are different from each other, and can also be decomposed into a different product of prime factors q_i

$$n = q_1^{m_1} \dots q_k^{m_k},$$

where all q_i are different from each other, then we have

$$p_1^{l_1} \dots p_j^{l_j} = q_1^{m_1} \dots q_k^{m_k}.$$

Simplifying any common factors from the two sides of this equation, and assuming that the two decompositions are not identical, there remains an equality between products of primes

$$p_1^{l_1} \dots p_{j'}^{l_{j'}} = q_1^{m_1} \dots q_{k'}^{m_{k'}},$$

where every p_i is different from each one of the q_i. Isolating one of prime numbers we get from this

$$p_1^{l_1} = \frac{q_1^{m_1} \dots q_{k'}^{m_{k'}}}{p_2^{l_2} \dots p_{j'}^{l_{j'}}},$$

which is impossible, because the right side of the equation is not an integer, since the product of the prime numbers in the numerator is not divisible by the set of all prime number in denominator. Thus we see that it is impossible to have an equality such as

$$p_1^{l_1} \dots p_j^{l_j} = q_1^{m_1} \dots q_k^{m_k},$$

where every p_i is different from each one of the q_i. We may thus conclude that the decomposition of an integer into prime factors is unique. Returning to the question of the number $r = \sqrt{2}$, if we assume that it is rational, then there are integers p and q such that

$$\begin{aligned} r &= \frac{p}{q} \\ &= \frac{p_1^{l_1} \dots p_j^{l_j}}{q_1^{m_1} \dots q_k^{m_k}}, \end{aligned}$$

where we decomposed p and q in its prime factors and canceled all common factors. Since p and q are two different numbers, we are left with two sets of prime numbers p_i and q_i, all different, and since r is not an integer, there is at least one element q_i in the denominator. If we now take the square of this equation, where $r^2 = 2$, we obtain

$$2 = \frac{p_1^{2l_1} \dots p_j^{2l_j}}{q_1^{2m_1} \dots q_k^{2m_k}},$$

which is not possible because the right-hand side is not an integer. Thinking about this in another way, this equation implies that

$$2q_1^{2m_1} \dots q_k^{2m_k} = p_1^{2l_1} \dots p_j^{2l_j},$$

that is, we have here two different decompositions of the same number in prime factors, which we have shown not to be possible. It follows that $\sqrt{2}$ cannot be written as p/q for p and q integers, so that $\sqrt{2}$ is not a rational number.

Considering that the roots of integers are typically not rational, at this point we can consider a generalization of the rational numbers, adding to them the roots of integers or, more generally, all the real roots of the polynomials of all orders with rational coefficients. The resulting set is called the set of algebraic numbers, but actually it is still an incomplete generalization, not including, for example, the number π. As in the case of rational numbers, it can be shown that this set has only a countable infinity of elements. In any case, here we see that this generalization of the rational numbers is associated to the concept of function, in this case to the polynomials, which involve algorithms for manipulating these numbers, based on the field properties.

The greatest possible generalization of the rational numbers, in order to cover all infinitesimal holes left by the rational numbers on the line, involves introducing in our discussion the concept of *infinite processes*, which is one of the most difficult and important aspects of abstract mathematics. This is where the concept of limit, which is one of the foundations of the integral and differential calculus, appears. It can be shown that the greatest possible generalization is to add to the set $\mathbb{Q}$ the limits of all possible infinite convergent sequences of rational numbers, which leads us to the set $\mathbb{R}$ of the real numbers.

The proof that this is a complete generalization is difficult and is beyond our scope here. The proof that the resulting set is a field can be worked out through the generalization of the arithmetic operations to infinite sequences of rational numbers. From the geometric point of view the field $\mathbb{R}$ of the real numbers can be identified in an unequivocal way with an infinite oriented line, which we then call the real line, where the numbers are interpreted as the distances of each point of the line from the origin, which is associated to the number zero.

The completeness of $\mathbb{R}$ in relation to the real line means no more than the existence of a complete one-to-one relation between the real numbers and the points of the line, that is, there is no point on the line that does not correspond to a real number, and there is no real number that does not correspond to a point on the line. Another way, this one more algorithmic, to express this relation between arithmetic and geometric elements is to say that there is no way to produce a real number, using the two arithmetic operations that are defined on them, even with an infinite number of steps, which is not representable as a point on the line, and that there is no way to produce a point on the line by purely geometrical means that cannot be represented by a sequence, possibly infinite, of arithmetic operations with real numbers.

We can illustrate both the rational numbers in $\mathbb{Q}$ and the real numbers in $\mathbb{R}$ in terms of their decimal representation. Note however that there is nothing mathematically fundamental about the representation on the decimal base, and that the same thing could be done with any other representation, such as the binary base, for example. Typically, rational numbers can be represented by a finite number of decimal digits, such as

$$\frac{3}{2} = 1.5,$$

or by an infinite number of digits which display, however, a periodic pattern starting from a given digit, such as

$$\frac{8}{3} = 2.666666\ldots\ .$$

On the other hand, real numbers that are not rational, which are called irrational numbers, are represented by an infinite number of digits that display no periodic pattern with any cycle length, such as, for example

$$\sqrt{2} = 1.414213562373095\ldots\ ,$$

or

$$\pi = 3.141592653589793\ldots\ .$$

The use of the continuum of the real numbers is often a very desirable simplifying resource. We will often make use of infinite processes in our mathematical work. The limits leading to the concepts of the derivative and of the integral are assumed known and familiar from previous work in basic real calculus. We will also often use sequences and series, and discuss their limits. As a simple and elegant example of a sequence of rational numbers that converges to an irrational number, illustrating the final step of the construction of the field of the real numbers, we can mention the sequence of finite continued fractions converging to $\sqrt{2}$, which is discussed in detail in [2] (available in Appendix A). As discussed there, from the simple identity

$$1 + \sqrt{2} = 2 + \frac{1}{1 + \sqrt{2}},$$

which one can easily check directly, it can be shown that the sequence of manifestly rational numbers

$$\begin{array}{l}
1, \\
1 + \dfrac{1}{2}, \\
1 + \dfrac{1}{2 + \dfrac{1}{2}}\,, \\
1 + \dfrac{1}{2 + \dfrac{1}{2 + \dfrac{1}{2}}}\,, \\
\ldots
\end{array}$$

converges quickly to $\sqrt{2}$. In Table 1.1 one can see the initial part of this infinite sequence of rational numbers.

It is also possible to construct power series involving only rational numbers, which converge to an irrational number such as $\sqrt{2}$, but this will only be discussed in detail later on in this book. This approximation scheme is discussed in detail in [3] (available in Appendix B). In both cases, computer programs [4] are also available for the partial construction of the sequences and series, thus allowing one to make an interesting

$\left(\frac{1}{1}\right)$	$=$	1.00000000	$\left(\frac{1393}{985}\right)$	$\approx$	1.41421320
$\left(\frac{3}{2}\right)$	$=$	1.50000000	$\left(\frac{3363}{2378}\right)$	$\approx$	1.41421362
$\left(\frac{7}{5}\right)$	$=$	1.40000000	$\left(\frac{8119}{5741}\right)$	$\approx$	1.41421355
$\left(\frac{17}{12}\right)$	$\approx$	1.41666667	$\left(\frac{19601}{13860}\right)$	$\approx$	1.41421356
$\left(\frac{41}{29}\right)$	$\approx$	1.41379310	$\left(\frac{47321}{33461}\right)$	$\approx$	1.41421356
$\left(\frac{99}{70}\right)$	$\approx$	1.41428571	$\left(\frac{114243}{80782}\right)$	$\approx$	1.41421356
$\left(\frac{239}{169}\right)$	$\approx$	1.41420118	$\left(\frac{275807}{195025}\right)$	$\approx$	1.41421356
$\left(\frac{577}{408}\right)$	$\approx$	1.41421569	$\vdots$		$\vdots$

Table 1.1: A sequence of rational numbers that tends to $\sqrt{2}$.

practical analysis through numerical methods, of their convergence to the limit.

Given the importance to physics of the mathematical structure of an ordered field, and having built the field $\mathbb{Q}$ of the rational numbers and the field $\mathbb{R}$ of the real numbers, of which $\mathbb{Q}$ is a subset, both ordered and the second complete on the real line, we can now ask ourselves if there is a complete and ordered field whose structure is different from the structure of $\mathbb{R}$. The answer, perhaps surprising, given to us by advanced algebra, is that no, there is not, and that this is a unique structure.

However, it is a remarkable fact that if we lift a single requirement, namely that the field be ordered, then we find that there is a more general structure, which has all the properties of a field, in addition to completeness and some additional structures. This is the field $\mathbb{C}$ of the *complex numbers*, which is in fact a larger field than the real numbers, and which contains it. According to theorems in advanced algebra, $\mathbb{C}$ is in fact the largest (commutative) field possible. Just as the use of $\mathbb{R}$

can simplify things in physics, although it is not conceptually essential, the use of $\mathbb{C}$ can simplify things even more radically. It is therefore not surprising that its use is widespread in physics.

Structures larger than $\mathbb{C}$ can be obtained only if we lift some of the requirements that define a field of numbers. For example, it is possible to obtain a set of elements named *quaternions*, formed by ordered quadruples of real numbers, provided that we give up the commutative property of the multiplication operation. We sometimes call this structure a non-commutative field, but strictly speaking it is no longer a field. It can also define an even larger set of elements named octonions, provided that we lift the requirement of the associativity of multiplication. Although the set of quaternions has something to do with quantum mechanics, since it can be represented through the Pauli matrices, which represent the concept of *spin* in that theory, the usefulness of these larger structures is much more limited than the usefulness of the complex numbers, which is truly immense.

The complex numbers are defined as ordered pairs of real numbers. Let $z = (x, y)$ be such an ordered pair with $x, y \in \mathbb{R}$, then $z \in \mathbb{C}$, and the operations of addition and multiplication of complex numbers are defined in terms of the corresponding operations of real numbers, as follows:

$$\begin{aligned} &\textbf{sum:} & z_1 + z_2 &= (x_1 + x_2, y_1 + y_2); \\ &\textbf{product:} & z_1 z_2 &= (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1). \end{aligned}$$

It is easy to verify that the identity element (or neutral element) of the sum is $(0, 0)$, and that the identity element (or neutral element) of the product is $(1, 0)$. The first element of the ordered pair is called its *real part* and the second element is called its *imaginary part*. Just as we have a real unit $(1, 0)$, we can also define an *imaginary unit*, which is the element $(0, 1)$. Since this element is particularly important, we denote it by a dedicated symbol, namely $\boldsymbol{\imath} = (0, 1)$. From the definition of the multiplication operation in $\mathbb{C}$ it follows that $\boldsymbol{\imath}^2 = (-1, 0)$.

Since the identity element of the multiplication is $(1, 0)$, it is natural to identify it with the symbol 1, without introducing any possibility of confusion. Thus we can also say that $\boldsymbol{\imath}^2 = -1$. The real or complex nature of numbers and symbols used will always be clear from the context of the arguments developed. Note that $\boldsymbol{\imath}$ is *not* defined as $\sqrt{-1}$, but as the ordered pair $(0, 1)$. In terms of symbols 1 and $\boldsymbol{\imath}$, any complex number $z = (x, y)$ can be written as a linear combination, $z = x + \boldsymbol{\imath} y$, which may be verified using only the definitions of the complex arithmetic operations.

If we consider the subset of the complex numbers which are of the form $(x, 0)$, we can easily see that we recover the set of real numbers x, with its usual properties of addition and multiplication, so that the field $\mathbb{R}$ is a subset of $\mathbb{C}$, which preserves the arithmetic operations of $\mathbb{C}$. It is also not difficult to see directly that the above-defined arithmetic operations for complex numbers satisfy all the properties necessary to define a field of numbers, by reducing the complex properties to the corresponding real properties, which are already known to hold.

Just as $\mathbb{R}$ is complete on the real line, $\mathbb{C}$ is complete in a two-dimensional plane (x, y), which we call the *complex plane*. The notion of completeness of the set, as well as the notion of limit, extend naturally to the complex framework, and are essentially the same that we know in the real context. Moreover, all the algebra of $\mathbb{C}$ is operationally identical to the algebra (arithmetic) of the real numbers, so that we can handle complex numbers in exactly the same way that we handle real numbers, without having to bother to indicate each time that the numbers involved are not purely real.

Although the field of complex numbers is not ordered, one can associate to them a notion of magnitude, which is very useful both in physics and in mathematical analysis, through the concept of the *absolute value* or *modulus* $|z|$ of a complex number $z = (x, y)$, which is the positive real number given by

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

Finally, we have a very useful additional operation in $\mathbb{C}$, which we call *complex conjugation*. Given the complex number $z = (x, y) = x + \imath y$, we define its complex conjugate z^*, often also denoted by $\bar{z}$, as

$$\begin{aligned} z^* &= (x, -y) \\ &= x - \imath y. \end{aligned}$$

It is easy to show that the absolute value of z can be written in terms of z and of its complex conjugate, as $|z| = \sqrt{z^* z}$. Note that the absolute value of a complex number has the same familiar form of a two-dimensional modulus or magnitude of a vector with two Cartesian components. Of course, just as we may consider variations of a real number x along a real axis, we may also consider variations of a complex number

z, which occur along the plane described by the Cartesian coordinates x and y. In each case, the magnitude of the variation is given by the absolute value of the difference of two numbers, either $|x_2 - x_1|$ in the real case, or $|z_2 - z_1|$ in the complex case. In the real case, the variation is always linear as it occurs along an axis, that is, on a straight line, while in the complex case it may occur along any continuous curve in the coordinate plane (x, y).

For those who are beginning to suspect that there is an unusual similarity between the structure of complex numbers and the structure of a two-dimensional vector space, since both can be represented on a plane (x, y), we end by observing that the similarity does exist, and that it will be used extensively in this book. Indeed, complex numbers can be identified with two-dimensional vectors over which a multiplication operation with inverse is defined. This geometric identification with the two-dimensional plane is very useful to facilitate the learning of the complex numbers, of the complex functions and of the complex calculus that we will develop, involving all these elements.

Problem Set

1. Consider the set $\mathbb{I}$ of the integers. Show that it does *not* satisfy some of the 11 properties of a field of numbers, among those that do not relate only to the operation of addition.

2. Consider the set $\mathbb{Q}$ of the rational numbers. Show that it satisfies all the 11 properties of a field of numbers.

3. Show that $\mathbb{Q}$ is a dense subset of $\mathbb{R}$, considering the integers, the arithmetic averages of pairs of these numbers, the arithmetic averages of pairs of the resulting numbers, and so on ad infinitum.

4. Consider a regular polygon with N sides, in which the distance from the center to a vertex is 1.

   (a) Calculate the perimeter of the polygon as a function of N.

   (b) Show that, in the limit $N \to \infty$, the perimeter approaches the value 2π.

5. Consider the complex numbers defined as ordered pairs of real numbers, and the definitions of the operations of addition and multiplication on them. We call this structure $\mathbb{C}$.

(a) Show that $(0, 0)$ is the identity element of the addition and that $(1, 0)$ is the identity element of the multiplication.

(b) Verify that the complex numbers and the two operations satisfy all the 11 properties of a field of numbers.

(c) Verify that $\mathbb{C}$ reduces to the field $\mathbb{R}$ of the real numbers for the subset of the complex numbers for which the second element of the ordered pair is zero.

(d) Consider the complex number $(0, 1)$, which is a very special case. Show that $(0, 1)^2 = (-1, 0)$.

6. Consider the complex field $\mathbb{C}$. Find a representation of this field by 2×2 real matrices, through the steps below.

   (a) Find a 2×2 real matrix whose square is the negative of the identity 2×2 real matrix.

   (b) Identifying the complex number $(1, 0)$ with the 2×2 identity matrix and the complex number $(0, 1)$ with the matrix found above, write an arbitrary complex number $z = (x, y)$ in terms of these matrices.

   (c) Write the arbitrary complex number $z = (x, y)$ as a real 2×2 matrix, and verify explicitly that this new representation preserves the form of the arithmetic operations on complex numbers.

7. **(Challenge Problem)** Show that the field of rational numbers $\mathbb{Q}$ is countable, building explicitly a one-to-one relation between $\mathbb{Q}$ and $\mathbb{N}$.

8. **(Challenge Problem)** Show that the field of real numbers $\mathbb{R}$ is *not* countable. In order to do this, by reductio ad absurdum, show that, if we assume that a sequence enumerating the real numbers is presented, then it is always possible to construct a real number that is none of the elements in that sequence.

   Hint: use the decimal representation of the real numbers.

Chapter 2

The Simplicity of Complex Numbers

Let us summarize the main points about complex numbers that we have so far. We have the definition of complex numbers as ordered pairs of real numbers, $z = (x, y)$, with the two arithmetic operations that follow:

$$\textbf{sum:} \qquad z_1 + z_2 = (x_1 + x_2, y_1 + y_2);$$

$$\textbf{product:} \qquad z_1 z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1).$$

The set $\mathbb{C}$ is a field, the largest there is. The identity elements are $(0, 0)$ (of the addition) and $(1, 0)$ (of the multiplication), and there is also the imaginary unit $\boldsymbol{\imath}$, which is defined as $\boldsymbol{\imath} = (0, 1)$, so that any complex number z can be written as

$$\begin{aligned} z &= (x, y) \\ &= x + \boldsymbol{\imath} y. \end{aligned}$$

There are some additional operations that can be defined. One is the operation of complex conjugation,

$$\begin{aligned} z &= (x, y) \\ &= x + \boldsymbol{\imath} y \quad \Rightarrow \\ z^* &= (x, -y) \\ &= x - \boldsymbol{\imath} y, \end{aligned}$$

which consists of changing the sign of the imaginary part y, or of changing $\boldsymbol{\imath}$ to $-\boldsymbol{\imath}$. There is also the absolute value of a complex number,

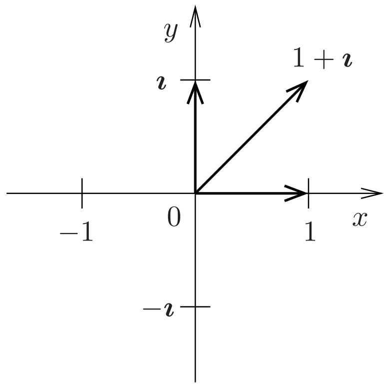


Figure 2.1: The complex plane, showing the numbers $(0,0)$, $(1,0)$, $(0,1)$ and $(1,1)$.

$$
\begin{aligned}
|z| &= \sqrt{x^2+y^2} \\
&= \sqrt{z^* z},
\end{aligned}
$$

which is always a non-negative real number, and which is zero only for $z=0$. Despite the fact that $\mathbb{C}$ is not an ordered field, one can assign to complex numbers a concept of magnitude, with the use of the absolute value. Thus, we can compare magnitudes within $\mathbb{C}$.

All the complex arithmetic and algebra are identical to those of the real numbers, so that we can handle complex numbers just as we handle real numbers. Note that, just as in the case of real numbers, division is only problematic if we have $z=(0,0)$ (or $z=0$, for short) in the denominator. Consider dividing any given complex number z by any another given complex number (a,b),

$$
\begin{aligned}
\frac{z}{(a,b)} &= \frac{z}{a+\imath b} \\
&= \frac{z(a-\imath b)}{(a+\imath b)(a-\imath b)} \\
&= \frac{z(a-\imath b)}{a^2+b^2},
\end{aligned}
$$

where we did the *rationalization* of the expression, eliminating any factors of $\imath$ in the denominator, so that, in fact, we only have a zero in the denominator for $a=0$ and $b=0$, that is, if the number in the denominator is the complex number zero.

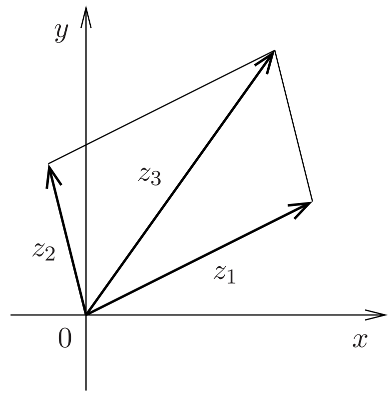


Figure 2.2: The sum of two generic complex numbers, interpreted as vectors in the complex plane.

It is possible to make a geometric representation of complex numbers as vectors on a plane, using the real and imaginary parts as Cartesian coordinates of a *complex plane*, as shown in Figure 2.1.

As is the case for vectors we can represent complex numbers in both Cartesian coordinates and polar coordinates. While the coordinates in the Cartesian representation are the real and imaginary parts, for the polar representation we can define the variables $\rho = |z|$ and θ, which is the angle that the vector z makes with the x axis, where we put the real parts of the vectors. Thus, in addition to $z = x + \boldsymbol{\imath} y$ we can write

$$z = \rho[\cos(\theta) + \boldsymbol{\imath}\sin(\theta)].$$

The arithmetic operations with complex numbers correspond to operations with these vectors. The sum of two complex numbers becomes a vector sum, as illustrated in Figure 2.2 for $z_3 = z_1 + z_2$.

The multiplication of two complex numbers involves a rotation of the vectors, that corresponds to the sum of the angles of each one of them. If the two vectors have unit absolute value, then this is purely a rotation. We can illustrate this by simply making the product of the numbers $\sqrt{2}/2 + \boldsymbol{\imath}\sqrt{2}/2$, that has unit absolute value and an angle of $\pi/4$, and $\boldsymbol{\imath}$, which also has unit absolute value and an angle of $\pi/2$, as shown in Figure 2.3, to get $z_3 = z_1 z_2$, with $z_1 = (1 + \boldsymbol{\imath})/\sqrt{2}$ and $z_2 = \boldsymbol{\imath}$.

The result of the multiplication, by the purely arithmetic definition, is $z_3 = (\boldsymbol{\imath} - 1)/\sqrt{2}$, so that $z_3 = -\sqrt{2}/2 + \boldsymbol{\imath}\sqrt{2}/2$, that has unit absolute value and angle $3\pi/4$, which is the sum of the other two angles. Another

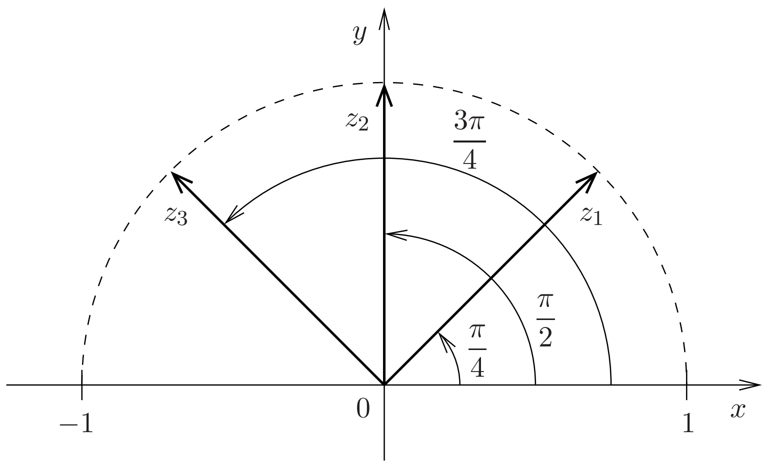


Figure 2.3: The product of two complex numbers of unit absolute value interpreted as vectors in the complex plane.

simple example illustrating this is $\boldsymbol{\imath}\,\boldsymbol{\imath} = \boldsymbol{\imath}^2 = -1$, because the result makes the angle $\pi = \pi/2 + \pi/2$ with the real axis. If the vectors have absolute values $\rho \neq 1$, then the operation also involves the product of these absolute values. As we shall see, the polar representation is particularly convenient to represent products.

Examining the case of the sum, illustrated in Figure 2.2, we can derive an extremely important relation involving the absolute values of the vectors in a sum, called the *triangle inequality*. As shown in Figure 2.2, the sum $z_3 = z_1 + z_2$ of two vectors z_1 and z_2 forms a triangle with the two vectors being added. It follows from the geometry that the absolute value of z_3 cannot be greater than the sum of the absolute values of z_1 and z_2, that is, we have the relation

$$\begin{aligned} z_3 &= z_1 + z_2 \quad \Rightarrow \\ |z_3| &\leq |z_1| + |z_2|, \end{aligned}$$

for all z_1, z_2 and z_3, the equality holding only in the case in which z_1 and z_2 are co-linear and point in the same direction. From this, it is not difficult to show a similar relation for the sum of any number of vectors, which will be left as an exercise. These triangle inequalities will be of capital importance in the development of the complex calculus.

Note that, since $\mathbb{C}$ is a field, the inverse of the multiplication is well defined for all $z \neq 0$, so that it is indeed possible to divide by a vector. This does not conflict with the general situation in vector spaces in which there is no division operation, because this is a very special particular case. The real vector space of two dimensions that we have here is the only one in which one can define a multiplication operation with a well-defined inverse.

Using the inverse of the sum, we can easily define a variation δz of a complex number z, which is no more than the difference of two complex numbers close to each other, z and $z + \delta z$. Similarly, if we have a complex function $w(z)$, that to each given complex number z associates another complex number w, this variation of z induces a corresponding variation δw of the complex number w. Each one of these variations is simply some other complex number, so that it is possible to divide δw by δz, and this allows us to define, through a limiting process, *complex derivatives*, thus opening the door to the development of a differential calculus, as we will do later.

The concept of limit for complex numbers is basically the same as in the real case, when dealing with functions of more than one real variable. It is in fact essentially identical to the concept of limit that applies to real functions of two variables. The difference is that the magnitudes of the numbers and of the distances between points, rather than being one-dimensional as in the real line, are those that apply to the plane $\mathbb{R}^2$. The absolute values of differences of real numbers appearing in the definitions of limits are simply exchanged by corresponding complex absolute values, which are the absolute values of vector differences. Another important difference is that the limits are taken along curves in $\mathbb{R}^2$. For example, the limit which leads to the derivative, as discussed above, is $\delta z \to 0$, which means that the complex number δz should approach the origin of the complex plane. However, this limit can be taken in many different ways, along different curves that pass through the origin, and not only along a single straight line.

Back to real analysis, we will now discuss a relation between certain functions, the *Euler formula*, which is associated with the DeMoivre theorem about integer powers of a complex number of unit absolute value, which states that

$$[\cos(\theta) + \imath \sin(\theta)]^n = \cos(n\theta) + \imath \sin(n\theta).$$

As one can see here, this is actually a generalization of the multiplication examples of complex numbers that we have examined above.

Due to the existence and uniqueness theorem of solutions of linear ordinary differential equations, associated with suitable auxiliary conditions, these equations can be understood as a way to define functions. For example, we can define the exponential function by means of the differential equation and auxiliary condition given by

$$\begin{aligned} \frac{df(x)}{dx} &= f(x), \\ f(0) &= 1 \quad \Rightarrow \\ f(x) &= e^x \\ &= \exp(x). \end{aligned} \tag{2.1}$$

This is a first-order equation, but we can also do the same with second-order equations, provided that two auxiliary conditions be given,

$$\begin{aligned} \frac{d^2 f(x)}{dx^2} &= -f(x), \quad f(0) = 1 \text{ and } f'(0) = 0 \quad \Rightarrow \\ & \qquad f(x) = \cos(x), \\ \frac{d^2 f(x)}{dx^2} &= -f(x), \quad f(0) = 0 \text{ and } f'(0) = 1 \quad \Rightarrow \\ & \qquad f(x) = \sin(x), \end{aligned}$$

where f' is an abbreviation for the derivative of f. Since the complex arithmetic is the same as the real one, and since the concept of limit is also the same in both cases, we can use Equation (2.1) in order to define the exponential for imaginary arguments. Consider $f(\omega) = \exp(\imath\omega)$ with ω real; in the equation that defines the exponential we will replace x by $\imath\omega$ and we will assume that $f = u + \imath v = u(\omega) + \imath v(\omega)$, since f must be some complex number. With this, we have $dx = \imath d\omega$, $df = du + \imath dv$, and thus the differential equation in Equation (2.1) can be written as $df = f dx$, so that we have

$$\begin{aligned} du + \imath dv &= (u + \imath v)\imath d\omega \\ &= \imath u\, d\omega - v\, d\omega. \end{aligned}$$

Equating separately the real and imaginary parts we obtain

$$\begin{aligned} du &= -v\, d\omega \quad \Rightarrow \quad \frac{du}{d\omega} = -v, \\ dv &= u\, d\omega \quad \Rightarrow \quad \frac{dv}{d\omega} = u, \end{aligned} \tag{2.2}$$

where the first line corresponds to the real part and the second to the imaginary part of the equation. Note that all the derivatives in question here are real, those with which we are familiar in the real calculus. If we combine the above two equations, we get separate differential equations for u and for v,

$$\frac{d^2u}{d\omega^2} = -u \quad \text{and} \quad \frac{d^2v}{d\omega^2} = -v.$$

Considering now the auxiliary conditions for these equations, we start from $f(0) = u(0) + \imath v(0) = 1$, implying that $u(0) = 1$ and $v(0) = 0$, and therefore, from the relations in Equation (2.2), that $u'(0) = -v(0) = 0$ and $v'(0) = u(0) = 1$, finally leading to the separate equations for u and v, with their respective auxiliary conditions,

$$\begin{aligned} \frac{d^2u}{d\omega^2} &= -u, \quad u(0) = 1 \quad \text{and} \quad u'(0) = 0, \\ \frac{d^2v}{d\omega^2} &= -v, \quad v(0) = 0 \quad \text{and} \quad v'(0) = 1. \end{aligned}$$

These are the equations and auxiliary conditions that define the real functions $\cos(\omega)$ and $\sin(\omega)$ respectively, so that we have $u(\omega) = \cos(\omega)$ and $v(\omega) = \sin(\omega)$, respectively. It follows that we have a (very useful) relation

$$e^{\imath\omega} = \cos(\omega) + \imath \sin(\omega),$$

which is the Euler formula. From this, assuming the usual properties of the exponential function, which can be understood as consequences of the complex arithmetic, we get the definition of the exponential function for a complex argument $z = x + \imath y$, or complex exponential function,

$$\begin{aligned} e^z &= e^{x+\imath y} \\ &= e^x[\cos(y) + \imath \sin(y)], \end{aligned}$$

which is the general form of the exponential function for complex arguments in terms of corresponding real functions. This is our first important example of a complex function. Note that it is obtained by the replacement $x \to z = x + \imath y$ on the real function $\exp(x)$.

The polar representation of a complex number z can now be written in terms of its absolute value ρ and the angle θ as

$$\begin{aligned} z &= \rho[\cos(\theta) + \imath \sin(\theta)] \\ &= \rho\, e^{\imath\theta}. \end{aligned}$$

Note how the interpretation of the multiplication of complex numbers in terms of the sum of angles is now clearer, due to the properties of the exponential. The division is now clearly related to the difference of the angles, and the inversion to the change of the sign of the angle. Finally, once more due to the properties of the exponential, it is now clear that the absolute value of the product is the product of the absolute values, that is, if we have that $z_3 = z_1 z_2$, then it follows that $|z_3| = |z_1||z_2|$.

We are now in a position to give a complete definition of a complex function: it is a mapping from $\mathbb{C}$ to $\mathbb{C}$. Given the complex variable $z = x + \imath y$, we can define a complex function $w(z)$ through the definition of its real and imaginary parts, with

$$w(z) = u(x, y) + \imath v(x, y),$$

where u and v are two real functions of two real variables, the variables x and y.

Although this is not clear in our first example, the exponential function may also be defined by the complex arithmetic, by means of a power series which we will see later. The definition by induction from real examples through the $x \to z$ replacement, as we did here, consists of a very particular type of choice of functions u and v. We can now show some simple examples of this using just the arithmetic. We start with the powers,

$$\begin{aligned} w_0(z) &= 1, \\ w_1(z) &= z, \\ w_n(z) &= z^n, \end{aligned}$$

with n a positive integer. We may then define polynomials and series, in terms of sets of constant complex numbers A_i,

$$\begin{aligned} w_p(z) &= \sum_{i=0}^{n} A_i z^i, \\ w_s(z) &= \sum_{i=0}^{\infty} A_i z^i. \end{aligned}$$

Knowing the real series which converge to certain real functions, we can use them to extend these functions to the complex domain, by means of the $x \to z$ replacement. We may therefore write series for functions such as $\exp(z)$, $\cos(z)$, $\sin(z)$ etc. In other words, from known real

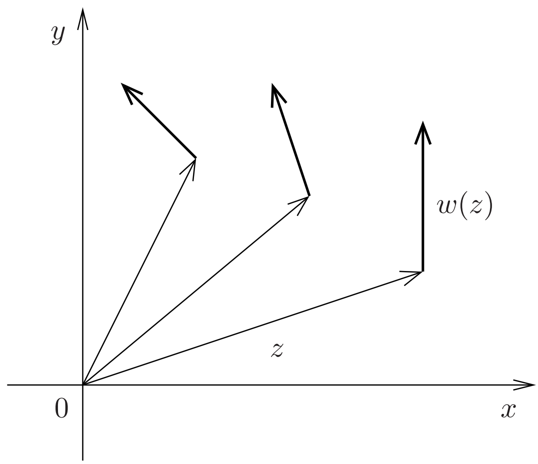


Figure 2.4: A vector field as an example of graphic representation of a complex function.

functions we can induce corresponding complex functions. These are, in fact, generalizations of the corresponding real functions since $\mathbb{C}$ is reduced back to $\mathbb{R}$ when the imaginary part of z vanishes.

From the geometric point of view, a complex function can be understood as a *vector field* in the complex plane, as illustrated in the diagram of Figure 2.4. If $z = (x, y)$ represents a position in the complex plane, then $w(z) = (u, v)$ is a two-dimensional vector at that position. This geometric representation will be very useful in the development of the differential and integral calculus in $\mathbb{C}$.

It is worth mentioning here that there is another way, quite different from this one, to interpret the complex functions geometrically: the functions can be understood as mappings from the plane to the plane, that can be illustrated by their action on given curves in the complex plane. For the most important functions, which will be named *analytic*, these mappings are very special and under some conditions are called *conformal mappings*. This will be discussed in more detail later on.

We will now verify a very simple property that is truly extraordinary, and that holds for all the above examples. Consider for example the case of the identity function $w(z) = z$, with $w = u + iv$ and $z = x + iy$, so that $u = x$ and $v = y$, implying that we have the relations

$$\begin{aligned}
\frac{\partial u}{\partial x} &= 1 = \frac{\partial v}{\partial y}, \\
\frac{\partial u}{\partial y} &= 0 = -\frac{\partial v}{\partial x}.
\end{aligned}$$

This may seem somewhat useless in this trivial case, but the interesting thing is that it is possible to generalize the relations above to other cases. Let us go to the next case, $w(z) = z^2$, in which we have $w(z) = u + \imath v = (x^2 - y^2) + 2\imath xy$, which gives $u = x^2 - y^2$ and $v = 2xy$. It is easy to see that the same conditions on the derivatives hold,

$$\begin{aligned}
\frac{\partial u}{\partial x} &= 2x = \frac{\partial v}{\partial y}, \\
\frac{\partial u}{\partial y} &= -2y = -\frac{\partial v}{\partial x}.
\end{aligned}$$

The two *Cauchy-Riemann conditions*,

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y}, \\
\frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x},
\end{aligned}$$

apply in general to complex functions induced from real functions as those shown above. We can generalize this to any power z^n, $n > 2$, using the method of finite induction. We start by assuming the result for the case $w_{n-1}(z)$,

$$\begin{aligned}
w_{n-1}(z) &= z^{n-1} \\
&= u_{n-1} + \imath v_{n-1}, \\
\frac{\partial u_{n-1}}{\partial x} &= \frac{\partial v_{n-1}}{\partial y}, \\
\frac{\partial u_{n-1}}{\partial y} &= -\frac{\partial v_{n-1}}{\partial x}.
\end{aligned}$$

Then we find out what the situation is for the following function,

$$\begin{aligned}
w_n(z) &= u_n + \imath v_n \\
&= z^n \\
&= w_{n-1}(z)\, z \\
&= (u_{n-1} + \imath v_{n-1})(x + \imath y).
\end{aligned}$$

After a short algebraic manipulation, we obtain for u_n and v_n

$$\begin{aligned} u_n &= u_{n-1}\, x - v_{n-1}\, y, \\ v_n &= u_{n-1}\, y + v_{n-1}\, x. \end{aligned}$$

Taking the partial derivatives of u_n with respect to x and of v_n with respect to y we get

$$\begin{aligned} \frac{\partial u_n}{\partial x} &= \frac{\partial u_{n-1}}{\partial x}\, x + u_{n-1} - \frac{\partial v_{n-1}}{\partial x}\, y, \\ \frac{\partial v_n}{\partial y} &= \frac{\partial u_{n-1}}{\partial y}\, y + u_{n-1} + \frac{\partial v_{n-1}}{\partial y}\, x, \end{aligned}$$

and using the Cauchy-Riemann relations for the case $n-1$ in the first equation above, we find that its terms become identical to those of the second equation, so that we obtain the first Cauchy-Riemann condition

$$\frac{\partial u_n}{\partial x} = \frac{\partial v_n}{\partial y},$$

which completes the proof by induction. The other case is similar, it suffices to consider the two other partial derivatives u_n and v_n, instead of the ones which were considered here, and the proof will be left as an exercise. It results in the second Cauchy-Riemann condition,

$$\frac{\partial u_n}{\partial y} = -\frac{\partial v_n}{\partial x}.$$

We can also show that the Cauchy-Riemann conditions hold to the function $w(z) = 1/z$, in almost all points of the complex plane. Using the complex arithmetic with $z = x + \boldsymbol{\imath} y$, we see that

$$\begin{aligned} w &= \frac{1}{z} \\ &= \frac{1}{x + \boldsymbol{\imath} y} \\ &= \frac{x - \boldsymbol{\imath} y}{x^2 + y^2} \\ &= \frac{x}{x^2 + y^2} - \boldsymbol{\imath}\, \frac{y}{x^2 + y^2}, \end{aligned}$$

where we rationalized the expression, eliminating any factors of $\boldsymbol{\imath}$ in the denominator, so that, since $w(z) = u + \boldsymbol{\imath} v$, it results that we have

$$\begin{aligned} u(x, y) &= \frac{x}{x^2 + y^2}, \\ v(x, y) &= -\frac{y}{x^2 + y^2}. \end{aligned}$$

Taking the appropriate partial derivatives of these two functions we obtain

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{1}{x^2+y^2} - \frac{2x^2}{(x^2+y^2)^2} \\ &= \frac{-x^2+y^2}{(x^2+y^2)^2}\end{aligned}$$

and

$$\begin{aligned}\frac{\partial v}{\partial y} &= \frac{-1}{x^2+y^2} + \frac{2y^2}{(x^2+y^2)^2} \\ &= \frac{-x^2+y^2}{(x^2+y^2)^2},\end{aligned}$$

from which follows the first relation,

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}.$$

Similarly, we also get the second relation,

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

whose derivation is left as an exercise. Thus we see that the Cauchy-Riemann relations are satisfied for the function $w = 1/z$. It should be noted however that in this case the relations do not hold or make any sense at the point $z = 0$, due to the factors $x^2 + y^2$ appearing in the denominators of all derivatives. Therefore, in this case the Cauchy-Riemann conditions are valid only away from the point $z = 0$, where the function has a singularity, and is not well defined.

Just as we did in the case of positive powers, we can extend this result to other negative powers of z by means of finite induction, except at the point $z = 0$, a fact whose proof we leave as an exercise. Moreover, it is trivial to verify that the relations also hold for the constant function $w = A + \imath B$, with A and B real constants, in which case all partial derivatives vanish. Thus, we see that the Cauchy-Riemann conditions hold for all powers, positive and negative, including the power zero, with the exception of point the $z = 0$ in the case of strictly negative powers.

Since the Cauchy-Riemann conditions hold for any power, and since partial differentiation is a linear operation, it follows that they also hold for polynomials and thus for power series, provided that these are convergent. We can even include sums and series containing negative powers. There is therefore a huge number of complex functions that satisfy these conditions, but it is important to note that they are *not all*. One can easily make an arbitrary choice of $u(x, y)$ and of $v(x, y)$, so as not to satisfy these conditions.

We will now reverse this whole argument and consider the set of all complex functions for which the two Cauchy-Riemann conditions are satisfied, regardless of the particular way in which they may have been defined. This brings us to the famous set of the so-called *analytic functions*, whose full definition is: a complex function $w(z) = u(x, y) + \imath v(x, y)$ is analytic if it is continuous and differentiable, that is, if $u(x, y)$ and $v(x, y)$ are continuous and differentiable real functions, and if $u(x, y)$ and $v(x, y)$ satisfy the Cauchy-Riemann conditions. The function is said to be analytic in the region of the complex plane within which all these conditions hold.

These functions are very special and very important. Although they have very special properties, they lend themselves to numerous applications in physics, engineering, and in mathematics itself [5]. We will be using these functions, explicitly or implicitly, throughout this series of books. Note that, since the analyticity criterion is based on a linear operation, the taking of partial derivatives of functions of two variables, we immediately have that the sum of two analytic functions is also analytic within the common domain of analyticity of the two functions being added together. The proof is immediately reduced to the fact that the derivative of the sum is the sum of the derivatives of each term of the sum.

Likewise, we can see that the product of two analytic functions is also analytic, in the common domain of analyticity of the two functions being multiplied together. In order to verify this we assume that $f(z)$ and $g(z)$ are two analytic functions, for which we have

$$\begin{aligned} f &= u_f + \imath v_f, \quad \text{with} \quad \frac{\partial u_f}{\partial x} = \frac{\partial v_f}{\partial y} \quad \text{and} \quad \frac{\partial u_f}{\partial y} = -\frac{\partial v_f}{\partial x}, \\ g &= u_g + \imath v_g, \quad \text{with} \quad \frac{\partial u_g}{\partial x} = \frac{\partial v_g}{\partial y} \quad \text{and} \quad \frac{\partial u_g}{\partial y} = -\frac{\partial v_g}{\partial x}, \end{aligned}$$

and consider the function $w = fg = u + \imath v$. We have for u and v in terms of u_f, v_f, u_g and v_g,

$$u = u_f u_g - v_f v_g,$$

$$v = u_f v_g + v_f u_g.$$

Taking the appropriate partial derivatives of u and v, and using the Leibniz rule for the derivative of the product of two real functions, we obtain

$$\begin{aligned}\frac{\partial u}{\partial x} &= \frac{\partial u_f}{\partial x} u_g + u_f \frac{\partial u_g}{\partial x} - \frac{\partial v_f}{\partial x} v_g - v_f \frac{\partial v_g}{\partial x}, \\ \frac{\partial v}{\partial y} &= \frac{\partial u_f}{\partial y} v_g + u_f \frac{\partial v_g}{\partial y} + \frac{\partial v_f}{\partial y} u_g + v_f \frac{\partial u_g}{\partial y}.\end{aligned}$$

Using now the Cauchy-Riemann conditions for u_f, v_f, u_g and v_g, the first condition for u and v follows. The second condition may be obtained similarly, and is left as an exercise.

Finally, it can be shown that the composition of two analytic functions is also analytic within the chained domain of analyticity of the two functions. Just as in the case of the product the proof reduces to the Leibniz formula for the derivative of a product and to the Cauchy-Riemann relations for the two original functions, in the case of the composition the proof reduces to the formula of the chain rule for the derivative of a composite function, and to the Cauchy-Riemann relations for the two original functions. The proof of this case is left entirely as an exercise.

Problem Set

1. Starting from the Euler formula, prove the DeMoivre theorem.

2. Show that the constant function $w(z) = 1$ is analytic, that is, that it satisfies the Cauchy-Riemann conditions. Starting from this fact, prove by finite induction that the function $w(z) = z^n$, with n a positive integer, is analytic on the entire complex plane.

3. Complete the proof given in the text that the function $w(z) = 1/z$ is analytic. From this fact prove, by finite induction, that the function $w(z) = 1/z^n$, with n a strictly positive integer, is analytic on the whole complex plane except at the origin $z = 0$.

4. Consider the movement of a solid disk which rotates around its axis with a constant angular speed ω. Consider a point of this disk at a distance ρ from the center, whose initial position forms an angle θ_0 with the x axis of a system of Cartesian coordinates (x, y) on the plane of the disk.

(a) Given that (x, y) are the coordinates of a point of the disk in this coordinate system, represent the position of the point as a complex number $z = (x, y)$, giving z as a function of the time t and of the initial position θ_0 of the point.

(b) Give the complex velocity $\dot{z}$ of the point of the disk as a function of the time t and of the initial position θ_0 of the point. Also, write $\dot{z}$ in terms of z.

(c) What is the effect on the two-dimensional vector $z = (x, y)$ of the operation of multiplying it by $\imath = (0, 1)$?

(d) Give the complex acceleration $\ddot{z}$ of the point of the disk as a function of the time t and of the initial position θ_0 of the point. Also, write $\ddot{z}$ in terms of z.

(e) Rewrite the results for z, $\dot{z}$ and $\ddot{z}$ using the exponential form of the complex numbers.

5. Consider the complex number $z = \exp(\imath\theta)$, where θ is a real number.

   (a) Determine the real and imaginary parts of z and of its complex conjugate z^*.

   (b) Calculate the absolute value of z, that is, $|z| = \sqrt{z^* z}$.

   (c) Determine the complex number z^2, the square of z, using the exponential form, and write its real and imaginary parts.

   (d) Calculate z^2 starting from z, without using the exponential form, that is, first writing z in terms of its real and imaginary parts, and then taking the square.

   (e) Comparing the results obtained in the two previous items, derive the trigonometric identities for the sine and the cosine of the double arc.

6. Show that the sum-function $w(z) = f(z) + g(z)$, of two analytic functions $f(z)$ and $g(z)$, is also analytic in the common domain of analyticity of $f(z)$ and $g(z)$.

7. Show that the product-function $w(z) = f(z)g(z)$, of two analytic functions $f(z)$ and $g(z)$, is also analytic in the common domain of analyticity of $f(z)$ and $g(z)$.

8. **(Reference Problem)** Show that the composite function $w(z) = f[g(z)]$ obtained by the composition of two analytic functions $f(z)$ and $g(z)$ is also analytic, within the domain that is given by the intersection of the image of $g(z)$ with the analyticity domain of $f(z)$.

   Hint: consider the composition of two complex functions involving three complex planes, so that one of the functions maps the first plane onto the second and the other function maps the second plane onto the third.

9. Starting from the fact that, for any complex numbers z_1, z_2 and z_3, it is true that

$$\begin{aligned} z_3 &= z_1 + z_2 \quad \Rightarrow \\ |z_3| &\leq |z_1| + |z_2|, \end{aligned}$$

   show that if we have a sum of n complex numbers z_i,

$$\begin{aligned} z_s &= \sum_{i=1}^{n} z_i \\ &= z_1 + z_2 + z_3 + \ldots + z_{n-1} + z_n, \end{aligned}$$

   it follows that

$$|z_s| \leq \sum_{i=1}^{n} |z_i|.$$

10. **(Challenge Problem)** Consider the complex number given by $z = \exp(\imath 2\pi/N)$, where N is a strictly positive integer.

    (a) Determine the real and imaginary parts of the powers z^k of z, for $k \in \{1, \ldots, N\}$, and write z^k in terms of their real and imaginary parts, leaving the result in terms of k and N.

    (b) Determine the sum of all the N numbers z^k, that is, calculate

$$S = \sum_{k=1}^{N} z^k.$$

Hint: draw vectors z^k in the complex plane, for a definite and not too large value of N, for example, for $N = 6$, $N = 7$ or $N = 8$.

Chapter 3

Elementary Functions, but Not Quite

Here we will examine some of the properties of some elementary analytic functions, in order to establish some familiarity with them. We can do this, for example, by plotting the real and imaginary parts of each function restricted to the real axis, or other restrictions of the same functions in the complex plane. We can also plot vector fields on the complex plane as a way to represent the functions, as mentioned in the previous chapter (Chapter 2). However, this is a less common way to proceed. In addition, we can examine the relations between various functions.

In general, we get some intuition about complex functions when they are written in terms of known real functions. Given a complex function, we can ask several relevant questions about it, such as those that follow.

- Whether or not it is analytic.
- How to define it algorithmically in terms of real functions.
- What is the domain within $\mathbb{C}$ where it is defined and analytic.
- What is its image within $\mathbb{C}$.
- Which differential equations are satisfied by it.
- How to write the function in terms of the polar representation.
- What are its algebraic properties.

- How it relates to other functions.

With the exception of the analytic issues involving derivatives and differential equations, these are the kinds of questions we will try to answer here in some cases. All functions which we will examine here are generalizations of real functions. Usually, these generalizations result in analytic functions when we exchange the real argument x by the complex argument $z = x + \imath y$, assuming that the algorithms used in the definition can be extended beyond the real domain to the complex domain through the generalization of the field $\mathbb{R}$ to the field $\mathbb{C}$. The resulting function is thus a function of the two variables (x, y), but its analyticity is related to the fact that this dependence is introduced solely through the particular linear combination $(x + \imath y)$. For example,

$$\begin{aligned} w(x, y) &= \exp(z) \\ &= \mathrm{e}^x[\cos(y) + \imath \sin(y)] \end{aligned}$$

is an analytic function, but

$$\begin{aligned} w(x, y) &= \exp(\imath x) \\ &= \cos(x) + \imath \sin(x) \end{aligned}$$

is not.

Since we have shown before that the sum, the product and the composition of analytic functions are also analytic functions, the simplest way to generate new analytic functions is by algebraic means, using these operations to combine analytic functions which are already known, such as the powers z^n for any integer n. Thus, from positive and negative powers we can generate rational functions such as

$$w(z) = \frac{1}{1+z}.$$

We see immediately that this function is defined on the whole complex plane except for the point $z = -1$, where it has a singularity due to a division by zero. In order to better understand the function, we can write it explicitly in terms of its real and imaginary parts,

$$\begin{aligned} w(z) &= \frac{1}{1 + x + \imath y} \\ &= \frac{(1+x) - \imath y}{(1+x)^2 + y^2} \\ &= \frac{(1+x)}{(1+x)^2 + y^2} + \imath \frac{-y}{(1+x)^2 + y^2}. \end{aligned}$$

The real part can be made arbitrarily small by making x approach -1 with $y \neq 0$, and arbitrarily large by making x approach -1 with $y = 0$. The imaginary part can be made arbitrarily small by making y approach 0 with $x \neq -1$, and arbitrarily large by making y approach 0 with $x = -1$. The value $(0, 0)$, however, is not accessible to the function, that is, it is never assumed by the function for finite values of x and y. This fundamentally different behavior for various limits that tend to the same point, indicating that the limit of the function to that point is not well defined, is characteristic of many singular points, as is the case here for the point $(-1, 0)$.

Restricting z to the real or imaginary axes, we can plot restrictions of the function. On the real axis, that is for $y = 0$, the function reverts back to the real function which was originally generalized,

$$w(x, 0) = \frac{1}{1 + x},$$

which is a quite familiar real function, with zero imaginary part. On the imaginary axis we have

$$\begin{aligned} w(0, y) &= \frac{1 - \boldsymbol{\imath} y}{1 + y^2} \\ &= \frac{1}{1 + y^2} + \boldsymbol{\imath}\, \frac{-y}{1 + y^2}, \end{aligned}$$

which is also written in terms of familiar real functions. Note that in this way we may consider that several related real functions are encoded into a single complex function, in its restrictions. We may, for example, examine the function on the axis $x = y$, getting

$$\begin{aligned} w(x, x) &= \frac{(1 + x)^2 - \boldsymbol{\imath} x}{(1 + x)^2 + x^2} \\ &= \frac{(1 + x)^2}{2x^2 + 2x + 1} + \boldsymbol{\imath}\, \frac{-x}{2x^2 + 2x + 1}. \end{aligned}$$

An interesting exercise is to examine more closely the behavior of a singular function such as $w(z) = 1/z$, close to the singularity. In this case this can be done simply and transparently using the polar representation $z = \rho \exp(\boldsymbol{\imath}\theta)$,

$$\begin{aligned} w(z) &= \frac{1}{z} \\ &= \frac{1}{\rho}\, \mathrm{e}^{-\boldsymbol{\imath}\theta}. \end{aligned}$$

The essence of this singularity is not just that the function diverges to infinity in limits to that point, but that the outcome depends on how one takes the limit.

For limits along the positive real semi-axis we have that $\theta = 0$, $\exp(-\imath\theta) = 1$ and the function is real, with zero imaginary part, and diverges to ∞ as $1/\rho$. For limits along the negative real semi-axis, we have that $\theta = \pi$, $\exp(-\imath\theta) = -1$ and the function is real and diverges to $-\infty$ as $(-1)/\rho$. On the other hand, for limits along the imaginary axis, with $\theta = \pi/2$ or $\theta = 3\pi/2$, we have that $\exp(-\imath\theta) = -\imath$ or $\exp(-\imath\theta) = \imath$, so that the function is purely imaginary, with zero real part, and diverges to $\mp\imath\infty$. For comparison, note that in the case of the function $w(z) = z = \rho\exp(\imath\theta)$ the result of the limit is always zero, whatever the value of θ.

The next case of interest, where things happen somewhat differently, is the function $w(z) = \sqrt{z}$. It is not difficult to define this function in the polar representation, because we simply have

$$\begin{aligned} w(z) &= \sqrt{z} \\ &= \sqrt{\rho}\,\mathrm{e}^{\imath\theta/2}, \end{aligned}$$

where $\sqrt{\rho}$ is the positive real square root of the non-negative real number ρ, in the usual sense of the real square root function. In this case, the function seems to be well defined in $z = 0$, because the limit of the function when $z \to 0$ is zero for any value of θ. However, this function is *not* analytic at $z = 0$. This is due to the fact that, as we shall see, it does not have a well-defined complex derivative at this point, that is, it is not differentiable. One way to check this, that we can do right now, is in terms of the Cauchy-Riemann conditions for u and v, because we can write

$$\begin{aligned} w(z) &= \sqrt{\rho}[\cos(\theta/2) + \imath\sin(\theta/2)] \\ &= u + \imath v, \end{aligned}$$

where we have the relations between (x, y) and (ρ, θ),

$$\begin{aligned} \rho &= \sqrt{x^2 + y^2}, \\ x &= \rho\cos(\theta), \\ y &= \rho\sin(\theta), \end{aligned}$$

in addition to the half-angle trigonometric relations,

$$\begin{aligned}
\sin(\theta/2) &= \pm\sqrt{\frac{1-\cos(\theta)}{2}} \\
&= \pm\sqrt{\frac{\rho - x}{2\rho}}, \\
\cos(\theta/2) &= \pm\sqrt{\frac{1+\cos(\theta)}{2}} \\
&= \pm\sqrt{\frac{\rho + x}{2\rho}},
\end{aligned}$$

so that we have for the functions u and v in terms of x and y,

$$\begin{aligned}
u(x,y) &= \pm\sqrt{\frac{\rho + x}{2}}, \\
v(x,y) &= \pm\sqrt{\frac{\rho - x}{2}},
\end{aligned}$$

where ρ is a function of x and y and the signs depend on the quadrant where θ is. We have therefore for the partial derivatives that appear in the first Cauchy-Riemann condition,

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \pm\frac{\sqrt{\rho + x}}{\sqrt{8}\,\rho}, \\
\frac{\partial v}{\partial y} &= \pm\frac{\sqrt{\rho + x}}{\sqrt{8}\,\rho},
\end{aligned}$$

so that the first Cauchy-Riemann relation is satisfied, whatever sign is chosen, so long as these expressions exist. Note that the derivatives diverge for $\rho \to 0$, and therefore they are not defined at the origin. The same holds for the other partial derivatives, so that the Cauchy-Riemann relations are not satisfied at that point, although it can be seen that they are satisfied arbitrarily close to the point and, indeed, throughout the rest of the complex plane. Here we see one example of a complex function that is not analytic at a given point, although it is not divergent at that singular point.

However, the function $\sqrt{z}$ has another special property. Since the polar representation of z is periodic in θ, it is clear that we can add 2π to θ without changing the number, that is

$$\begin{aligned}
z &= \rho\,\mathrm{e}^{\imath\theta} \\
&= \rho\,\mathrm{e}^{\imath\theta + 2\imath\pi} \\
&= \rho\,\mathrm{e}^{\imath\theta + 2n\imath\pi},
\end{aligned}$$

for any integer n. Thus, if we add π to the angle of the image of the function $\sqrt{z}$, its square value will not change,

$$\begin{aligned}
\sqrt{z} &= \rho_s\, e^{\imath\theta_s} \quad \Rightarrow \\
\sqrt{z}^2 &= \rho_s^2\, e^{\imath 2\theta_s} \\
&= z, \\
\sqrt{z}' &= \rho_s\, e^{\imath(\theta_s+\pi)} \quad \Rightarrow \\
\sqrt{z}'^2 &= \rho_s^2\, e^{\imath 2\theta_s}\, e^{2\imath\pi} \\
&= z,
\end{aligned}$$

where $\rho_s = \sqrt{\rho}$ and $\theta_s = \theta/2$, because $\exp(2\imath\pi) = 1$. Thus, we always have two complex numbers whose squares are the same number z,

$$\begin{aligned}
\sqrt{z} &= \rho_s\, e^{\imath\theta_s}, \\
\sqrt{z} &= \rho_s\, e^{\imath\theta_s}\, e^{\imath\pi} \\
&= -\rho_s\, e^{\imath\theta_s},
\end{aligned}$$

because $\exp(\imath\pi) = -1$. That is, regarded as the inverse function of $w(z) = z^2$, the function $w^{-1}(z) = \sqrt{z}$ is a function that assigns two different values for each value of z. In fact, this is outside the scope of the usual definition of a function, but nonetheless we will call $\sqrt{z}$ a *function with multiple values* or alternatively a *multivalued function*.

This duality of values is an extension to the complex plane of the corresponding duality of the real function $\sqrt{x}$, which is well known. However, as is often the case in the complex plane, on it we have additional structures. If we examine the values of $\sqrt{z}$ when we go around $z = 0$ in the complex plane, for example along a circle of radius ρ, we see that the function has a discontinuity at some point of the circle. This contradicts, of course, the fact that we have already determined that the only singularity of the function is at $z = 0$, and that the function is in fact analytic on the rest of the plane. If we use the interval $[-\pi, \pi]$ for the values of θ, by starting the process at $\theta = 0$ and approaching the point $\theta = \pi$ by two different paths, one by positive values of θ and another by negative values, as shown in the diagram of Figure 3.1, we find that the behavior of the function $w(z) = \sqrt{z}$ is the following,

$$\begin{aligned}
w(z) &= \sqrt{\rho}\, e^{\imath\theta/2} \quad \Rightarrow \\
w(\rho,\theta) &= \sqrt{\rho}\cos(\theta/2) + \imath\sqrt{\rho}\sin(\theta/2) \quad \to \quad \imath\sqrt{\rho}, \\
w(\rho,-\theta) &= \sqrt{\rho}\cos(\theta/2) - \imath\sqrt{\rho}\sin(\theta/2) \quad \to \quad -\imath\sqrt{\rho},
\end{aligned}$$

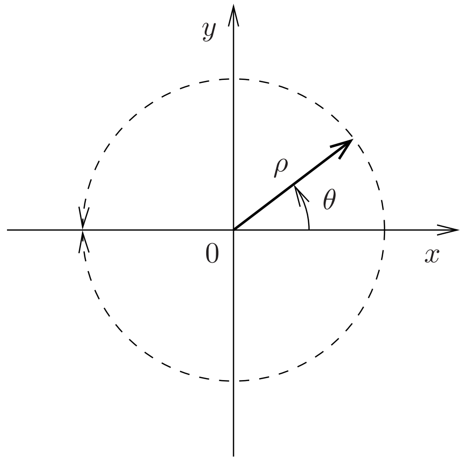


Figure 3.1: The plane and the paths of the two limits.

as $\theta \to \pi$. We see that the two values differ by a change of sign, except at the point $z = 0$, and therefore that the function defined in this way is not continuous and hence cannot be analytic along the negative real semi-axis. The most curious fact is that the location of this line of apparent discontinuities depends on the interval that we use for θ, despite the fact that the choice of this interval is clearly irrelevant for the polar description of z! For example, if we use the interval $[0, 2\pi]$ for θ, and repeat the argument, then the line of apparent discontinuities will be on the positive real semi-axis. Whatever the interval used to θ, there will always be a line of apparent discontinuities connecting the point $z = 0$ of the singularity of the function to infinity in some direction.

This apparent discontinuity and the apparent non-analyticity which follows from it are a curious phenomenon because, if we are interested in dealing with the function in a particular region of the complex plane that does not include the origin $z = 0$, then we can always avoid these singularities through an appropriate change in the interval for θ! The only thing we cannot do is to make a complete turn around $z = 0$.

The most interesting thing of all this, however, is that one can in fact redefine this function as a double-valued complex function, in order to completely eliminate these apparent singularities. The idea is due to Riemann and it leads to the concept of *Riemann surfaces*, composed in this case of two leaves, or *Riemann leaves*, which are connected to one another via a *branch cut*. In this case the singular point $z = 0$ is called a *branch point*.

For the function $w(z) = \sqrt{z}$ the solution is simple: in order to return to the initial value and define a continuous function, just make *two* turns around the origin, instead of just one. So, we use the interval $[0, 4\pi]$ for values of θ, thus covering the complex plane not once but twice. Since we have that

$$\begin{aligned} w(\rho, \theta = 0) &= \sqrt{\rho}\,\mathrm{e}^{\imath\theta/2} &= \sqrt{\rho}\,\mathrm{e}^{\imath 0} &= \sqrt{\rho}, \\ w(\rho, \theta = 4\pi) &= \sqrt{\rho}\,\mathrm{e}^{\imath\theta/2} &= \sqrt{\rho}\,\mathrm{e}^{\imath 2\pi} &= \sqrt{\rho}, \end{aligned}$$

we see that, after the second turn, the function returns to the value it had at the beginning, being therefore continuous and, as we will show later, even differentiable. We can easily show that the function is analytic at all points of the Riemann surface, with the exception of the origin, using the Cauchy-Riemann conditions written in polar coordinates. It is not difficult to change variables in the Cauchy-Riemann conditions from (x, y) to (ρ, θ). The derivation is left as an exercise, and the result is

$$\begin{aligned} \frac{\partial u(\rho,\theta)}{\partial \rho} &= \frac{1}{\rho}\frac{\partial v(\rho,\theta)}{\partial \theta}, \\ \frac{\partial v(\rho,\theta)}{\partial \rho} &= -\frac{1}{\rho}\frac{\partial u(\rho,\theta)}{\partial \theta}. \end{aligned}$$

Using this and writing

$$\begin{aligned} w(z) &= \sqrt{z} \\ &= \sqrt{\rho}\,[\cos(\theta/2) + \imath \sin(\theta/2)] \\ &= u + \imath v, \end{aligned}$$

it is not difficult to show the validity of the two Cauchy-Riemann conditions away form the point $\rho = 0$.

We have therefore a double-valued analytic function in such a way that, when we go once around the zero, the value of the function changes from one value to the other, and such that after the second turn it returns to its initial value. If we define the function on a double cover of the complex plane, namely on two Riemann leaves, each one with one of the two possible values associated to it, then the two leaves may be appropriately connected to one another as a single Riemann surface, as shown in Figure 3.2. On this surface the function is continuous, differentiable and analytic, at all points except at the point $z = 0$, which is the branch point, a particular type of singularity.

A simple curve of arbitrary shape, drawn by connecting the origin to infinity, the branch cut, marks the place where we agree to move from

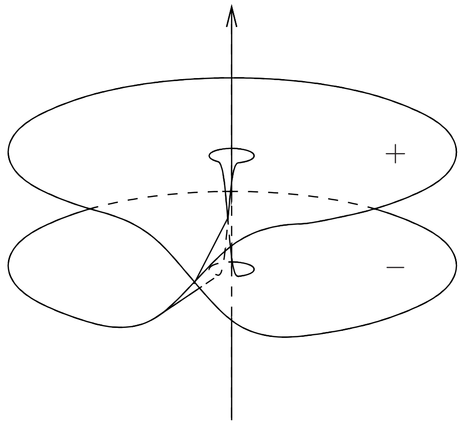


Figure 3.2: The two Riemann leaves and their connection with each other.

one leaf to the other. The points of this line are not singularities, except for the point $z = 0$, while its position is arbitrary and it can be deformed in any way we wish, provided that it does not auto-intersect and that it connects the origin to infinity. If in the course of some calculation we are forced to cross this line, meaning that we cannot deform the line to avoid the crossing, then we have to worry about the passage from one leaf of the function to the other.

The introduction of Riemann surfaces simplifies and unifies the definition of functions with multiple values, taking all possible values into consideration and giving the structure a geometric substrate which is useful and intuitive. Note that what we have here is a change in the *domain* of definition of the function, which is generalized from the simple complex plane to the Riemann surface, so that the function becomes dependent not solely on the single continuous variable z, but instead on the pair (z, n), where n is a discrete integer variable.

The Riemann surface becomes the geometric manifold in which the function is defined, that is, its domain. The image of the function is contained in the simple complex plane. We can imagine that we mark on each point of the Riemann surface of the function its value at that point, as a way of representing the function. In our case here, if we change from one leaf to the other at any position on the complex plane, the function simply changes sign. Riemann surfaces for functions with multiple values often appear when we consider the inverse functions of simple analytic functions, such as $w(z) = z^2$ in this case.

Note that there is no real equivalent of these Riemann surfaces, because in order to define them it is absolutely necessary to leave the real axis and include imaginary values in the derivations. Therefore, for the real function $\sqrt{x}$ we must make the usual arbitrary choice and define the function as positive, and then only on the positive real semi-axis.

Turning now to the exponential function, we already have its definition in real terms,

$$\begin{aligned}\exp(z) &= \mathrm{e}^{x}\,\mathrm{e}^{\imath y} \\ &= \mathrm{e}^{x}[\cos(y) + \imath\sin(y)].\end{aligned}$$

Note that this formula shows that both the real function $\exp(x)$ and the real functions $\sin(x)$ and $\cos(x)$ are different aspects of a single complex function, which somehow unifies them. On the real axis, the above function is reduced to $\exp(x)$, but on the imaginary axis it is reduced to a complex linear combination of $\sin(y)$ and $\cos(y)$.

It is not difficult to see that the function can be calculated for any values of x and y, meaning that its domain is the whole complex plane. Furthermore, since $\exp(x)$ is not zero for any value of x, and the combination $[\cos(y) + \imath\sin(y)]$ does not vanish for any value of y, we see that the value zero is not reached by the function, that is, that its image is the whole complex plane except for the point $z = 0$. Note that $\exp(z)$ is *periodic* in the imaginary direction, with period $2\pi\imath$. The periodicity of the complex function which constitutes the factor dependent on y, $[\cos(y) + \imath\sin(y)]$, may be represented by making infinitely many loops around the unit circle of the complex plane, as illustrated in Figure 3.3, which is a type of complex periodicity, that never passes through the point $(0, 0)$.

We can just as easily generalize the functions $\sin(x)$ and $\cos(x)$ to the complex plane. Starting from the relations we already have, derived for the real version of the functions,

$$\begin{aligned}\exp(\imath x) &= \cos(x) + \imath\sin(x), \\ \exp(-\imath x) &= \cos(x) - \imath\sin(x),\end{aligned}$$

we can derive, adding and subtracting the two equations, that

$$\begin{aligned}\cos(x) &= \frac{\mathrm{e}^{\imath x} + \mathrm{e}^{-\imath x}}{2}, \\ \sin(x) &= \frac{\mathrm{e}^{\imath x} - \mathrm{e}^{-\imath x}}{2\imath}.\end{aligned}$$

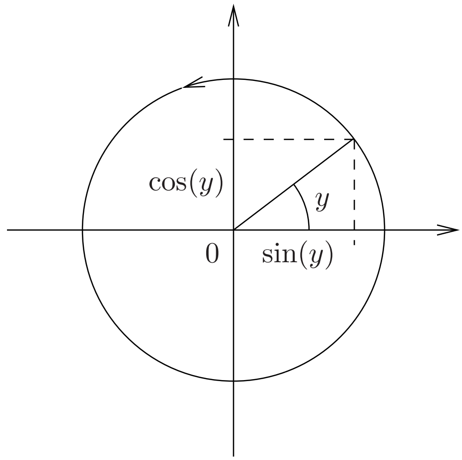


Figure 3.3: The unit circle with $\cos(y)$ and $\sin(y)$ on the axes.

The generalization to the complex plane can be made with the simple exchange of x for z,

$$\begin{aligned}\cos(z) &= \frac{e^{\imath z} + e^{-\imath z}}{2}, \\ \sin(z) &= \frac{e^{\imath z} - e^{-\imath z}}{2\imath}.\end{aligned}$$

Making the substitution $z = x + \imath y$ we can now write $\cos(z)$ explicitly in terms of real functions,

$$\begin{aligned}\cos(z) &= \frac{1}{2}\left[e^{(\imath x - y)} + e^{(-\imath x + y)}\right] \\ &= \frac{1}{2}\left\{e^{-y}[\cos(x) + \imath\sin(x)] + e^{y}[\cos(x) - \imath\sin(x)]\right\} \\ &= \frac{1}{2}\left[\cos(x)(e^{y} + e^{-y}) - \imath\sin(x)(e^{y} - e^{-y})\right],\end{aligned}$$

so that we have

$$\cos(z) = \cos(x)\cosh(y) - \imath\sin(x)\sinh(y),$$

a result written in terms of the familiar real hyperbolic functions

$$\begin{aligned}\sinh(x) &= \frac{e^{x} - e^{-x}}{2}, \\ \cosh(x) &= \frac{e^{x} + e^{-x}}{2}.\end{aligned}$$

We can also make a similar calculation for the other function, obtaining

$$\sin(z) = \sin(x)\cosh(y) + \imath\cos(x)\sinh(y).$$

A very important point is that, unlike their real versions, $\sin(z)$ and $\cos(z)$ are *not* limited functions. These are other complex functions that unify the real functions $\exp(x)$, $\sin(x)$ and $\cos(x)$. Trigonometric functions are periodic along the real axis x, but not along the imaginary axis y or other vertical lines, where their real and imaginary parts increase or decrease without limit. Making $y = 0$ in the above formulas we see that every complex function is reduced, on the real axis, to the corresponding real function. Making $x = 0$, we get on the imaginary axis relations between the real functions $\sin(x)$, $\sinh(x)$, $\cos(x)$ and $\cosh(x)$,

$$\begin{aligned} \cos(\imath y) &= \cosh(y), \\ \sin(\imath y) &= \imath\sinh(y). \end{aligned}$$

The other trigonometric functions can be defined in terms of $\sin(z)$ and $\cos(z)$ in the usual way,

$$\begin{aligned} \tan(z) &= \frac{\sin(z)}{\cos(z)}, \\ \cot(z) &= \frac{\cos(z)}{\sin(z)}, \\ \sec(z) &= \frac{1}{\cos(z)}, \\ \csc(z) &= \frac{1}{\sin(z)}. \end{aligned}$$

Since the function $\cos(z)$ appears in the denominator in these formulas, it is interesting to examine here the question of the zeros of this function. Just like the location of the singularities, the location of the zeros of an analytic function is also an important criterion for characterizing these functions. We know that the real function $\cos(x)$ has zeros, and their locations. The question is whether there are additional zeros along the complex plane. If we impose that $\cos(z) = 0$, we obtain

$$\cos(x)\cosh(y) - \imath\sin(x)\sinh(y) = 0,$$

which implies that we must have the two conditions

$$\begin{aligned}\cos(x)\cosh(y) &= 0,\\ \sin(x)\sinh(y) &= 0.\end{aligned}$$

Since the function $\cosh(y)$ is never zero, it follows that we must have $\cos(x) = 0$. But at the points where $\cos(x)$ is zero the function $\sin(x)$ never vanishes, so that it is necessary that $\sinh(y) = 0$, implying that $y = 0$. Thus we see that all the zeros of the complex function are on the real axis $y = 0$, and are the known zeros of $\cos(x)$, which are therefore the *only* zeros of the function $\cos(z)$.

We can also define the complex hyperbolic functions, by direct analogy with the corresponding real definitions,

$$\begin{aligned}\sinh(z) &= \frac{e^{z} - e^{-z}}{2},\\ \cosh(z) &= \frac{e^{z} + e^{-z}}{2},\end{aligned}$$

as well as the other usual hyperbolic functions, $\tanh(z)$, $\coth(z)$, etc. Writing $\cosh(z)$ explicitly in terms of real functions, we have

$$\begin{aligned}\cosh(z) &= \frac{1}{2}\left[e^{(x+\imath y)} + e^{(-x-\imath y)}\right]\\ &= \frac{1}{2}\left\{e^{x}[\cos(y) + \imath\sin(y)] + e^{-x}[\cos(y) - \imath\sin(y)]\right\},\end{aligned}$$

so that we get

$$\cosh(z) = \cosh(x)\cos(y) + \imath\sinh(x)\sin(y),$$

and similarly for the other function,

$$\sinh(z) = \sinh(x)\cos(y) + \imath\cosh(x)\sin(y).$$

Many properties can be easily derived from these definitions. For example, restricting the relations to the imaginary axis we obtain

$$\begin{aligned}\cosh(\imath y) &= \cos(y),\\ \sinh(\imath y) &= \imath\sin(y).\end{aligned}$$

All these analytic functions are complex combinations of the real functions $\exp(x)$, $\sin(x)$ and $\cos(x)$, and all are periodic in one direction of the complex plane and unbounded in the other direction. Because of this they are called *simply periodic functions*.

There is also another class of functions that are much less elementary, which are periodic in both directions of the complex plane. The functions of this type are called *doubly-periodic functions*. These are the *elliptic functions*, that are related to the calculation of the arc length of an ellipse, in the same way as the trigonometric functions are related to the arc length of a circle. These functions appear from time to time in physics, for example relating to the exact solution for the motion of a simple pendulum, without the approximation of small amplitudes. These doubly periodic functions, as well as being less familiar, have a considerably more complicated and difficult structure than the structure of the trigonometric functions, and there are whole treatises dedicated to them.

Problem Set

1. By means of a transformation of variables in the complex plane, from the Cartesian coordinates (x, y) to the polar coordinates (ρ, θ), write the two Cauchy-Riemann conditions in terms of the variables ρ and θ, that is, in terms of derivatives with respect to these variables, where

$$\begin{aligned} x &= \rho\cos(\theta), \\ y &= \rho\sin(\theta). \end{aligned}$$

2. Show that the complex function $w(z) = \sqrt{z}$ is analytic, that is, that it satisfies the Cauchy-Riemann conditions throughout the complex plane, except for the origin $z = 0$.

3. Given the complex number $z = x + \imath y$, consider the complex function $w(z) = u(x, y) + \imath v(x, y)$ in each case below. In each case, determine the real part $u(x, y)$ and the imaginary part $v(x, y)$ of $w(z)$, and show that the function $w(z)$ satisfies the two Cauchy-Riemann conditions.

   (a) $w(z) = \exp(z)$.

   (b) $w(z) = \cos(z)$.

   (c) $w(z) = \sin(z)$.

   (d) $w(z) = \cosh(z)$.

(e) $w(z) = \sinh(z)$.

4. Show that the complex function of the variable $z = x + \imath y$ given by $w(z) = x^2 + \imath y^2$ is *not* analytic. Show the same for the complex function $w(z) = z^* z$.

5. In each case below, determine the domain of analyticity of the function, that is, determine the points of the complex plane where the function is *not* analytic.

   (a) $w(z) = \tan(z)$.
   (b) $w(z) = \cot(z)$.
   (c) $w(z) = \sec(z)$.
   (d) $w(z) = \csc(z)$.

6. **(Challenge Problem)** Determine the branch points, the branch cuts and the Riemann leaves of the functions that follow. Cut and paste the leaves to produce Riemann surfaces on which the functions are completely well defined and analytic, except for the singularities at the branch points.

   (a) $w(z) = \sqrt{z^2 - 1}$.
   (b) $w(z) = 1/\sqrt{z^2 - 1}$.

   Hint: consider the behavior of the functions along circles around the points where the square root vanishes.

Chapter 4

Even Less Elementary Functions

We will now continue our exploration of analytic functions, examining some of the properties of some less elementary functions, which means only that they are less familiar because of their less frequent use in introductory courses on physics and mathematics. It is interesting to emphasize here that *all* the functions we will be studying throughout the rest of this series of books are examples of analytic functions, progressively less and less familiar or "basic". One of the things that determines this classification in more or less elementary functions is the degree of algorithmic difficulty involved in the practical determination of the values of the functions at a given arbitrary point of their domain within $\mathbb{C}$.

A very important function that we will examine here is the inverse function of the function $\exp(z)$, that is, the logarithm $\ln(z)$. It is relatively easy to do this by writing z in the polar representation, $z = \rho \exp(\imath\theta)$, because then we can see at once, given the known properties of the logarithm, that the definition should be

$$\ln(z) = \ln(\rho) + \imath\theta,$$

that is, the real part is the usual real logarithm of the positive real quantity ρ, and the imaginary part is θ, the polar angle of z. We see immediately that this function reduces to $\ln(x)$ for $\theta = 0$, that is, on the positive real semi-axis. As we shall see below, it is also very easy to see that this function is in fact analytic, except for the case of the point $z = 0$.

Two notable things happen here: first, as in the case of the real function, the point $z = 0$, for which $\rho = 0$, is a singular point of the function, because it diverges to $-\infty$ when we go to this point along the positive real semi-axis. Moreover, if we go around the origin along circles centered at the origin of the complex plane, we have that $\ln(\rho)$ is constant, and therefore the function simply varies linearly with θ. This means that when we go around the full circle, the function does not return to the initial value, and instead jumps by $2\pi\boldsymbol{\imath}$.

The situation is similar to that of the function $\sqrt{z}$ that we examined before, with the apparent line of singularities dependent on the interval of values that we choose for θ. The traditional choice is $\theta \in [-\pi, \pi]$, so that the line is located on the negative real semi-axis. Instead of a change of sign as before, the discontinuity of the function is now an imaginary constant, which is independent of ρ and θ. In this case, if we take a second turn around the origin, the function does not return to the initial value, but simply jumps again by $2\pi\boldsymbol{\imath}$. Nevertheless, as we shall see, it is still possible to redefine the function so that it is continuous, differentiable and analytic on the whole complex plane, except for the origin.

Just as before, we can show that this function is naturally a multi-valued function. Consider the value for $\ln(z)$ given by

$$\ln(z) = \ln(\rho) + \boldsymbol{\imath}(\theta + 2n\pi),$$

where n is an arbitrary integer. Taking the exponential of either side of this equality we obtain

$$\begin{aligned}
\exp[\ln(z)] &= \mathrm{e}^{\ln(\rho)}\,\mathrm{e}^{\boldsymbol{\imath}(\theta+2n\pi)} \\
&= \rho\,\mathrm{e}^{\boldsymbol{\imath}\theta}\,\mathrm{e}^{\boldsymbol{\imath}2n\pi} \\
&= \rho\,\mathrm{e}^{\boldsymbol{\imath}\theta} \\
&= z,
\end{aligned}$$

since $\exp(\boldsymbol{\imath}2n\pi) = 1$ for all n, so that any one of this infinite collection of different values, each one spaced $2\pi\boldsymbol{\imath}$ from the previous one, is a possible logarithm of z. It follows that once again we can redefine the function so that it is continuous and, in fact, analytic on the whole complex plane except at $z = 0$, using a Riemann surface. In this case, however, it will be necessary to employ an infinite number of Riemann leaves, numbered by n from $-\infty$ to ∞, each one connected to the next through a branch

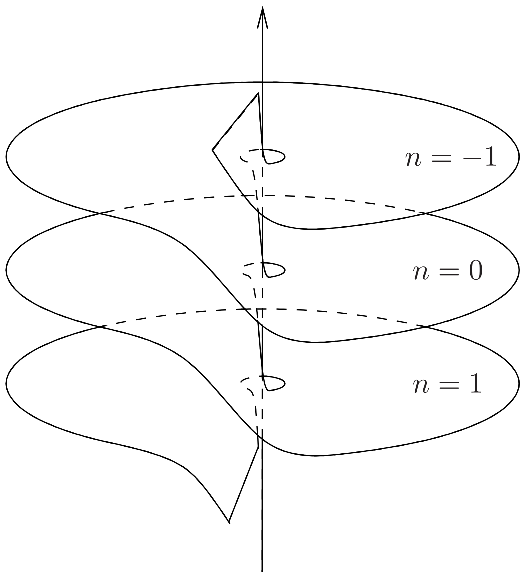


Figure 4.1: The Riemann leaves and the connection of each one to the next.

cut connecting the origin to infinity, as illustrated in the diagram of Figure 4.1. The central leaf is the one given by $n = 0$.

Note that we always have $\exp[\ln(z)] = z$, independently of n, but that $\ln[\exp(z)] = z + 2n\pi\boldsymbol{\imath}$, for a value of n which remains indeterminate. Hence, it is necessary to choose a Riemann leaf, in this case the central leaf $n = 0$, in order to have a well-defined inversion in the opposite order, $\ln[\exp(z)] = z$. Just as in the case of $\sqrt{z}$, we can see that $\ln(z)$ is analytic away from the origin, using the Cauchy-Riemann conditions in their polar form. In the general case we have

$$\begin{aligned} \ln(z) &= \ln(\rho) + \boldsymbol{\imath}(\theta + 2n\pi) \\ &= u + \boldsymbol{\imath}v, \end{aligned}$$

so that we have

$$\begin{aligned} u &= \ln(\rho), \\ v &= \theta + 2n\pi, \end{aligned}$$

and therefore the relevant partial derivatives are

$$\begin{aligned}
\frac{\partial u}{\partial \rho} &= \frac{1}{\rho} \quad \text{and} \quad \frac{1}{\rho}\frac{\partial v}{\partial \theta} = \frac{1}{\rho}, \\
\frac{\partial v}{\partial \rho} &= 0 \quad \text{and} \quad \frac{1}{\rho}\frac{\partial u}{\partial \theta} = 0,
\end{aligned}$$

so that the two conditions are satisfied provided that $\rho \neq 0$, that is, away from the point $z = 0$. Note that, since the dependence on n does not appear in these relations, they are valid for all leaves, that is, over the entire Riemann surface of the function.

Having the logarithm function, we can now define powers with complex exponents. As in the case of real exponents, we must use the logarithm function in order to define powers with exponents other than just integers or rationals. In the real case, we can define integer powers in terms of products,

$$x^n = x \times x \times \ldots \text{ (n factors) } \ldots \times x \times x.$$

We can also define powers with rational exponents p/q, using order-p powers and order-q roots,

$$x^{p/q} = \sqrt[q]{x \times x \times \ldots \text{ (p factors) } \ldots \times x \times x}.$$

However, in order to define real exponents, not necessarily rational, we must use real logarithms and the exponential function,

$$x^r = \exp[r \ln(x)].$$

Since the functions $\exp(x)$ and $\ln(x)$ are algorithmically defined for real values (by power series, as we shall see later on), provided that x is positive, this is an algorithmic definition of x^r. The same reasoning applies to the case of complex numbers, because we can define, using only the field properties,

$$z^n = z \times z \times \ldots \text{ (n factors) } \ldots \times z \times z.$$

Analogously to the real case, using the concept of complex roots we can also define

$$z^{p/q} = \sqrt[q]{z \times z \times \ldots \text{ (p factors) } \ldots \times z \times z}.$$

It is not difficult to verify, using the polar form of z, that the q-th root of z, with q integer as in the case above, has in fact q different complex values. If $z = \rho \exp(\imath\theta)$, then these q possible values of the q-th root are distributed along the circle of radius $\rho^{1/q}$ in the complex plane, equally spaced along the circle, that is, with angles $2\pi/q$ between each one and the next. Of course, in this case we can regularize the function by introducing a Riemann surface, just like we did in the case of the square root, except that we will now have a surface with q leaves.

Continuing with our development of powers of complex numbers, in this complex case we can generalize directly from the case of rational exponents to the case of complex exponents c, which includes the case of real exponents r, using

$$z^c = \exp[c \ln(z)].$$

Note that here we have no concern that the real part x of z be positive, as we had in the real case. Instead we have the preoccupation of defining the Riemann leaf at which the logarithm is to be taken, always excepting the point $z = 0$. Since $z = \rho \exp(\imath\theta)$, we have

$$z^c = \exp\{c\,[\ln(\rho) + \imath(\theta + 2n\pi)]\},$$

where n determines the leaf of the logarithm, and thus we have the n-dependent factor

$$\exp(c\,2n\pi\imath).$$

If c is a real integer then this exponential always equals 1 and the power has a single value, but otherwise there is more than one possible value for the power. If c is a real rational number p/q, then this exponential reduces to the q-th complex root of 1, which in general generates q different values for the power z^c. In all other cases there will be an infinite number of possible values for z^c, corresponding to the infinite number of leaves of the Riemann surface of the function $\ln(z)$.

Note that if $c = \chi$ is real then the absolute value of z^c does not depend on n, which only affects the phase angle of the function,

$$\begin{aligned} z^c &= \rho^\chi \exp[\imath\chi(\theta + 2n\pi)] \quad \Rightarrow \\ |z^c| &= \rho^\chi, \end{aligned}$$

but if c has an imaginary part, $c = \chi + \imath\lambda$, then both the absolute value and the phase become dependent on n, because in this case we will have for the absolute value

$$|z^c| = \rho^\chi \exp[-\lambda(\theta + 2n\pi)].$$

Finally, besides the inverse functions of the powers and of the exponential, we can also define the inverse of trigonometric and hyperbolic functions. For example, if $w = \sin^{-1}(z)$ then

$$\begin{aligned} z &= \sin(w) \\ &= \frac{e^{\imath w} - e^{-\imath w}}{2\imath}. \end{aligned}$$

We can solve this in terms of the variable $\exp(\imath w)$, since this equation is quadratic on this variable,

$$(e^{\imath w})^2 - 2\imath z\,(e^{\imath w}) - 1 = 0,$$

so that the Baskara formula, which is valid for complex numbers since it depends only on the field properties, gives us

$$e^{\imath w} = \imath z \pm \sqrt{1 - z^2},$$

where the fact that the square root function is double-valued is expressed explicitly. Taking the logarithm we have now

$$\begin{aligned} w(z) &= \sin^{-1}(z) \\ &= -\imath \ln\left(\imath z \pm \sqrt{1 - z^2}\right), \end{aligned}$$

showing that the inverse function has infinitely many different values, related to the Riemann leaves of the complex logarithm function. Choosing a leaf of the logarithm and a leaf of the square root, the resulting function is analytic on that leaf of the Riemann surface because it is a composition of analytic functions. Similarly, one can obtain for the other inverse functions

$$\begin{aligned} \cos^{-1}(z) &= -\imath \ln\left(z \pm \sqrt{z^2 - 1}\right), \\ \tan^{-1}(z) &= \frac{\imath}{2} \ln\left(\frac{\imath + z}{\imath - z}\right), \\ \sinh^{-1}(z) &= \ln\left(z \pm \sqrt{z^2 + 1}\right), \\ \cosh^{-1}(z) &= \ln\left(z \pm \sqrt{z^2 - 1}\right), \\ \tanh^{-1}(z) &= \frac{1}{2} \ln\left(\frac{1 + z}{1 - z}\right). \end{aligned}$$

Just as the real functions $\exp(x)$, $\sin(x)$ and $\cos(x)$ are all related to each other in the complex context, so the corresponding inverse functions are also all related, so that we can write all of them in terms of the logarithm function.

We have been examining the so-called elementary functions, starting with examples that, in fact, are quite simple, and proceeding to others that are no longer so elementary. Let us end by giving an example of an analytic function that is definitely not in this so-called elementary class. It is the *gamma function*, or $\Gamma(x)$, a function that can be defined in the real context by a parametric integral,

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

and which will often be very useful, as we shall see later. Since this definition involves a real exponent of t, it implicitly involves the definition of the real logarithm. Writing this explicitly, we have the alternative expressions

$$\begin{aligned}
\Gamma(x) &= \int_0^\infty \mathrm{e}^{-t}\, \mathrm{e}^{(x-1)\ln(t)}\, dt \\
&= \int_0^\infty \frac{dt}{t}\, \mathrm{e}^{-t}\, \mathrm{e}^{x\ln(t)} \\
&= \int_{t=0}^{t=\infty} d[\ln(t)]\, \mathrm{e}^{-t}\, \mathrm{e}^{x\ln(t)}.
\end{aligned}$$

Since the exponential $\exp(-t)$ goes to zero for $t \to \infty$ faster than any power t^{x-1} goes to infinity, the parametric integral always converges in the asymptotic region, that is for $t \to \infty$. However, close to $t = 0$ we have that $\ln(t)$ is negative and goes to $-\infty$ for $t \to 0$, so that the exponential $\exp[x\ln(t)]$ in the second and third forms of the equation above is limited, and the integral convergent, only if x is strictly positive. The particular case $x = 0$ leads to

$$\Gamma(0) = \int_0^\infty \frac{dt}{t}\, \mathrm{e}^{-t},$$

that corresponds to an integral of $1/t$ to the origin $t \to 0$, which is also divergent. Thus, the necessary and sufficient condition for the existence of the parametric integral on t is that $x > 0$.

The function can be calculated without too much difficulty for some particular values of x, such as in the cases $\Gamma(1) = 1$ and $\Gamma(2) = 1$. One can also show, with a little more work, that $\Gamma(1/2) = \sqrt{\pi}$. Through manipulations of the integral it can be shown that the function has the fundamental property that

$$\Gamma(x+1) = x\Gamma(x).$$

In order to prove this it suffices to write the definition of $\Gamma(x+1)$ and integrate by parts, recalling that we must have x strictly positive in this definition, in which case the integrated term vanishes,

$$\begin{aligned}
\Gamma(x+1) &= \int_0^\infty \mathrm{e}^{-t} t^x \, dt, \\
&= \int_0^\infty \mathrm{e}^{-t} \, \mathrm{e}^{x\ln(t)} \, dt \\
&= -\,\mathrm{e}^{-t} \, \mathrm{e}^{x\ln(t)} \Big[_0^\infty + \int_0^\infty \mathrm{e}^{-t} \left(\frac{x}{t}\right) \mathrm{e}^{x\ln(t)} \, dt \\
&= -\,\mathrm{e}^{-t} t^x \Big[_0^\infty + x \int_0^\infty \mathrm{e}^{-t} \left(\frac{1}{t}\right) t^x \, dt \\
&= x \int_0^\infty \mathrm{e}^{-t} t^{x-1} \, dt \\
&= x\Gamma(x).
\end{aligned}$$

With this, and the particular values mentioned above, one can verify by induction that $\Gamma(n+1) = n!$, where $n!$ is the factorial of n, that is, the product of all integers from 1 to n,

$$n! = n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1.$$

This is a very useful quantity appearing, for example, in the formula for the coefficients of the Taylor series, as we shall see later. One can also define the double factorial, which will also be useful later. In this case, the definition must be considered separately for the cases in which n is odd or even. For even n we have that the double factorial is given by

$$n!! = n \times (n-2) \times (n-4) \times \ldots \times 6 \times 4 \times 2,$$

that is, it is the product of all even integers from 2 to n. In the case of odd n the double factorial is given by

$$n!! = n \times (n-2) \times (n-4) \times \ldots \times 5 \times 3 \times 1,$$

that is, it is the product of all odd integers from 1 to n. Clearly, we have that $n!!(n-1)!! = n!$, for all n. It is not difficult to show that, for even n, we have that

$$n!! = 2^{(n/2)}(n/2)!,$$

so that this double factorial can be written in terms of common factorials, whereas for odd n we have

$$\begin{aligned} n!! &= \frac{n!}{(n-1)!!} \\ &= \frac{n!}{2^{[(n-1)/2]}[(n-1)/2]!}, \end{aligned}$$

where we used the previous result, since in this case $n-1$ is even. Thus we see that these double factorials can also be written in terms of the function $\Gamma(x)$. This function was discovered by Euler precisely as an answer to the problem of interpolating continuously between the values of the integer factorials. Interestingly, unlike what happens with the elementary functions, it is not possible to define this function as the unique solution of some ordinary differential equation. Thus, it is very different from other functions studied so far. However, there is no difficulty in extending the definition of the function to the complex domain. In order to do this it suffices to write that

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

where z is a complex variable, and the condition $x > 0$ translates to $\Re(z) > 0$. Note that the parametric integral, which is on t, remains real, although there is no difficulty in interpreting it as a complex integral that is taken over the positive real semi-axis, as we shall see later on.

We can now show that the function is analytic, except for isolated singularities. In order to do this we need to write the function in terms of the (still real) logarithm of t,

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

where, using $z = x + \imath y$, we can explicitly write the real and imaginary parts of the function, we get

$$\begin{aligned}
\Gamma(z) &= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{(x+\imath y)\ln(t)} \\
&= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{x\ln(t)}\{\cos[y\ln(t)] + \imath\sin[y\ln(t)]\} \\
&= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{x\ln(t)}\cos[y\ln(t)] + \\
&\quad +\imath\int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{x\ln(t)}\sin[y\ln(t)] \\
&= \int_0^\infty \mathrm{e}^{-t}t^{x-1}\cos[y\ln(t)]\,dt + \imath\int_0^\infty \mathrm{e}^{-t}t^{x-1}\sin[y\ln(t)]\,dt.
\end{aligned}$$

Note that because these real trigonometric functions are limited and periodic, they do not affect the convergence of the integrals, so that the existence conditions of the integrals here are the same as in the real case. In short, the necessary and sufficient condition for the existence of the integrals is that $\Re(z) > 0$. With these formulas established, we can identify the functions $u(x, y)$ and $v(x, y)$. The most convenient form for our immediate use here is

$$\begin{aligned}
u(x, y) &= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{x\ln(t)}\cos[y\ln(t)], \\
v(x, y) &= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{x\ln(t)}\sin[y\ln(t)],
\end{aligned}$$

so that we can calculate the partial derivatives involved in the first Cauchy-Riemann condition,

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\ln(t)\,\mathrm{e}^{x\ln(t)}\cos[y\ln(t)], \\
\frac{\partial v}{\partial y} &= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{x\ln(t)}\ln(t)\cos[y\ln(t)].
\end{aligned}$$

Thus we see that the first condition is satisfied whenever it is possible to calculate these integrals, whose convergence conditions are the same as those of the integral that defines $\Gamma(x)$. For the other partial derivatives we have

$$\begin{aligned}
\frac{\partial u}{\partial y} &= -\int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\,\mathrm{e}^{x\ln(t)}\ln(t)\sin[y\ln(t)], \\
\frac{\partial v}{\partial x} &= \int_0^\infty \frac{dt}{t}\,\mathrm{e}^{-t}\ln(t)\,\mathrm{e}^{x\ln(t)}\sin[y\ln(t)],
\end{aligned}$$

so that the second condition is also satisfied. It follows that the function is analytic on all the half-plane $x > 0$.

The fundamental property of the function is generalized immediately to complex values of the argument, since the integration by parts that was involved in its proof was for the parametric integral over t, in which x was only a parameter. The derivation can be repeated in the complex case without any change, leading to

$$\Gamma(z+1) = z\Gamma(z).$$

Although the original definition in terms of the parametric integral only applies for $\Re(z) > 0$, this property allows us to extend the definition of the function to the whole complex plane, except for a set of isolated singularities. For example, considering the region of the complex plane defined by $-1 < x \leq 0$, for any value of y, and considering that the function is well defined in the region defined by $0 < x \leq 1$, we can *define* the function in this first negative strip imposing this property for the negative values of x, that is, we can define within this region that

$$\Gamma(z) = \frac{\Gamma(z+1)}{z},$$

for

$$-1 < \Re(z) \leq 0,$$

where $\Gamma(z+1)$ is defined by the parametric integral as before. Since $\Gamma(z+1)$ is perfectly well defined for z within the strip $-1 < x \leq 0$, because in this case $z+1$ is within the strip $0 < x \leq 1$, and furthermore, since $\Gamma(z)$ is analytic in the strip $0 < x \leq 1$, its composition with the analytic function $(z+1)$ produces a function that is also analytic. Finally, the ratio of the two analytic functions is also analytic, except for the points where the function in the denominator vanishes. It follows that the function $\Gamma(z)$ defined above is analytic throughout the region $-1 < x \leq 0$, except for the point $z = 0$, at which it has a simple pole.

This is the *analytic continuation* of the function to this strip with negative x. This analytic continuation process is a unique process in the sense that it defines the function in the new region in a unique way. Note that if we have $0 < x$, then we can use the parametric integral to write $\Gamma(z+1) = z\Gamma(z)$ on the right-hand side of the definition, so that it is reduced to the above definition in terms of the parametric integral. In other words, this new definition can be extended to all the half-plane

$-1 < x$, without changing the definition of the function where it was already defined by the previous criterion. We can therefore write

$$\begin{aligned}
\Gamma(z) &= \frac{\Gamma(z+1)}{z}, \\
\Gamma(z+1) &= \int_0^\infty \mathrm{e}^{-t} t^z \, dt \quad \Rightarrow \\
\Gamma(z) &= \frac{1}{z} \int_0^\infty \mathrm{e}^{-t} t^z \, dt,
\end{aligned}$$

as a definition of the function throughout the half-plane $-1 < \Re(z)$. We may then continue to extend the function in this fashion, for example to the strip $-2 < x \leq -1$, by the corresponding definition, based on the same fundamental property,

$$\Gamma(z) = \frac{\Gamma(z+2)}{z(z+1)},$$

for

$$-2 < \Re(z) \leq -1.$$

Once again we see that the function is analytic within the new region, except for a new simple pole, this one located at $z = -1$. Therefore, we can generalize the definition once again, obtaining

$$\begin{aligned}
\Gamma(z) &= \frac{\Gamma(z+2)}{z(z+1)}, \\
\Gamma(z+2) &= \int_0^\infty \mathrm{e}^{-t} t^{z+1} \, dt \quad \Rightarrow \\
\Gamma(z) &= \frac{1}{z} \frac{1}{(z+1)} \int_0^\infty \mathrm{e}^{-t} t^{z+1} \, dt,
\end{aligned}$$

as a definition of the function throughout the half-plane $-2 < \Re(z)$. It is clear that we can continue this indefinitely, thereby generalizing the function $\Gamma(z)$ to the whole complex plane except only for an infinite set of isolated singularities, which are simple poles located at zero and at the negative integers. This fact could be induced from the fact that the factor t^{x-1} that appears in the parametric integral acquires the behavior of a pole of order n at $t = 0$ for $x = 1 - n$ with n a positive integer.

Note that this definition has to be broken down separately for each negative unit interval located between two negative real integers. In other words, we can only give a single definition for the half-plane given by values of x above a certain negative integer. It is possible to write a more general definition of $\Gamma(z)$ for the whole complex plane, through the use of complex integrals, also knows as contour integrals, on the complex t plane, but this is somewhat beyond the scope of this book. A derivation of the general formula can be found in [6] (available in Appendix F).

The function $\Gamma(z)$ will be very useful in future studies, to permit the generalization of factorials to numbers that are not integers. In fact, this function is useful on many occasions throughout the study of physics and mathematics. In particular, in various situations, especially those involving analyzes of a statistical nature, and also in our future discussions about the convergence of series of special functions, it is necessary to consider the values of this function for large values of the argument, usually taken on the real axis. In this case the function can be approximated in a very useful way by an asymptotic formula, which is a good approximation for large values of the function. A simple derivation of this formula, called the Stirling formula, or Stirling approximation, can be found in [7] (available in Appendix C).

Problem Set

1. Calculate, giving their real and imaginary parts, the numbers that follow.

   (a) $z = 1^{\boldsymbol{\imath}}$.

   (b) $z = \mathrm{e}^{\boldsymbol{\imath}}$.

   (c) $z = \sin(\boldsymbol{\imath})$.

   (d) $z = \cos(\boldsymbol{\imath})$.

2. Consider the complex function $w(z) = z^{1/n} = \sqrt[n]{z}$ where $n > 1$ is an integer.

   (a) Show that $w(z)$ is a multivalued function, with n different values for each z.

   (b) Show that these n values are evenly distributed along a circle in the complex plane.

   (c) Determine the radius of this circle and the angles corresponding to each one of the n possible values.

(d) Build a Riemann surface with n leaves to represent the domain of the function. Determine the singular point.

(e) Show that the function is analytic throughout this Riemann surface.

3. For each one of the inverse functions listed below, prove that they can be written in terms of the logarithm as shown in each case.

(a) $\sin^{-1}(z) = -\imath \ln\left(\imath z \pm \sqrt{1-z^2}\right)$.

(b) $\cos^{-1}(z) = -\imath \ln\left(z \pm \sqrt{z^2-1}\right)$.

(c) $\tan^{-1}(z) = \frac{\imath}{2} \ln\left(\frac{\imath + z}{\imath - z}\right)$.

(d) $\sinh^{-1}(z) = \ln\left(z \pm \sqrt{z^2+1}\right)$.

(e) $\cosh^{-1}(z) = \ln\left(z \pm \sqrt{z^2-1}\right)$.

(f) $\tanh^{-1}(z) = \frac{1}{2} \ln\left(\frac{1+z}{1-z}\right)$.

4. Starting from the definition of the function $\Gamma(x)$ in terms of a parametric integral, show that it can be written in each one of the three forms below.

(a) $\Gamma(x) = \displaystyle\int_0^\infty dt\; \mathrm{e}^{-t}\, \mathrm{e}^{(x-1)\ln(t)}$.

(b) $\Gamma(x) = \displaystyle\int_0^\infty \frac{dt}{t}\, \mathrm{e}^{-t}\, \mathrm{e}^{x\ln(t)}$.

(c) $\Gamma(x) = \displaystyle\int_{t=0}^{t=\infty} d[\ln(t)]\, \mathrm{e}^{-t}\, \mathrm{e}^{x\ln(t)}$.

5. Calculate explicitly the following real values of the function $\Gamma(x)$.

(a) $\Gamma(1)$.

(b) $\Gamma(2)$.

(c) $\Gamma(1/2)$.

Answer: $\sqrt{\pi}$.

6. Show that $\Gamma(z+1) = z\Gamma(z)$ for any complex z except for the origin $z = 0$ and the negative integers.

7. Consider the function

$$f(t) = \exp\left[-\frac{(t-t_0)^2}{2\tau^2}\right],$$

where t_0 is a real constant and $\tau \neq 0$ is a strictly positive real constant, and consider also the definition of the average value of another function $g(t)$ in the statistical distribution defined by $f(t)$, which is given by

$$\langle g \rangle = \frac{\displaystyle\int_{-\infty}^{\infty} g(t) f(t)\, dt}{\displaystyle\int_{-\infty}^{\infty} f(t)\, dt}.$$

   (a) Calculate the average value of t, that is, $\langle t \rangle$.

   Answer: t_0.

   (b) Calculate the dispersion of t, that is the quantity

$$\sigma_t = \sqrt{\langle (t - \langle t \rangle)^2 \rangle}.$$

   Answer: τ.

8. **(Challenge Problem)** Show that, for any real, strictly positive number α,

$$\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{\alpha}\right) dt = \sqrt{\alpha\pi}.$$

Hint: calculate the square of the integral, and transform to a polar coordinate system on the plane.

Chapter 5

Geometrical Aspects of the Functions

Being in control of the concepts of complex numbers, complex functions and analytic functions, let us now examine some properties of the analytic functions that have a geometric character, and which follow from the Cauchy-Riemann conditions that they satisfy. First, let us examine some properties of the two real functions of two variables $u(x,y)$ and $v(x,y)$, which constitute an analytic function $w = u + \imath v$. The Cauchy-Riemann conditions tell us that

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y}, \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x}. \end{aligned}$$

Differentiating the first of these two relations with respect x and then using the second relations to write the result only in terms of u, we obtain

$$\begin{aligned} \frac{\partial^2 u}{\partial x^2} &= \frac{\partial}{\partial x}\frac{\partial v}{\partial y} \\ &= \frac{\partial}{\partial y}\frac{\partial v}{\partial x} \\ &= -\frac{\partial^2 u}{\partial y^2}, \end{aligned}$$

so that we have the result that u satisfies a familiar differential equation,

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} = 0,$$

the Laplace equation in two dimensions. The same can be done for the other function, resulting in

$$\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} = 0,$$

or, using the usual shorthand for this differential operator,

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

where the underscore is a placeholder for the object the operator is to act on, we see that the two functions satisfy the Laplace equation, and are therefore, by definition, *harmonic functions*,

$$\begin{aligned} \nabla^2 u &= 0, \\ \nabla^2 v &= 0. \end{aligned}$$

This establishes a surprising relation with electrostatics in two dimensions, in the absence of free charges. It is enough to write an *arbitrary* analytic function in order to get without further effort two solutions of the electrostatic equation, that is, two possible electric potentials in two dimensions. Since we have several methods to generate analytic functions from real functions, or from other analytic functions, we have here truly a factory of solutions of problems in two-dimensional electrostatics.

What we have here is an unusual situation because we can easily generate a large set of solutions that then stand looking for their problems! Often, with a bit of art, we may be able to guess a function $w(z)$ whose real part u satisfies the appropriate boundary conditions on certain curves in the two-dimensional space, that correspond to appropriate surfaces in three-dimensional space, and therefore to find in this way the solution of a two-dimensional electrostatic problem.

For example, in this case a very simple example, consider the function $\phi(z) = c - i\alpha z$, where $z = x + \imath y$, $c = a + \imath b$ is an arbitrary complex constant, and α is an arbitrary real constant. Since this is a linear combination of complex powers, which are analytic functions, we have here an analytic function. The real and imaginary parts of this function are given by

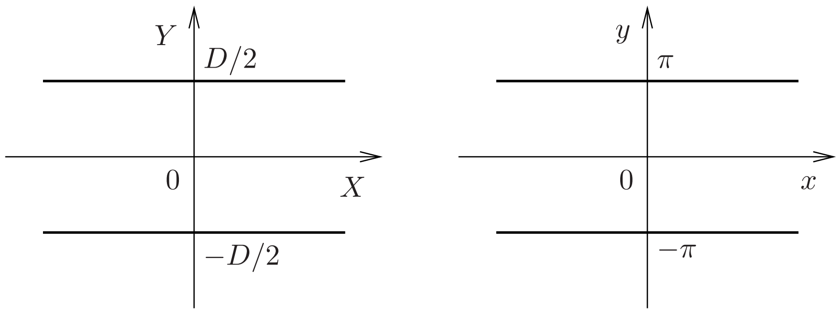


Figure 5.1: The (X, Y) plane with the surfaces $Y = \pm D/2$ and the corresponding (x, y) plane with the surfaces $y = \pm\pi$.

$$\begin{aligned} \phi(z) &= u(x,y) + \imath v(x,y) \quad \Rightarrow \\ u(x,y) &= a + \alpha y, \\ v(x,y) &= b - \alpha x, \end{aligned}$$

which are linear functions of y and x, and can therefore be interpreted as electric potentials in a region of space where the electric field is uniform. Using u as an example, it could be the electric potential between two flat and infinite metal plates defined for each value of y by a certain constant for all values of x, that is, the potential within an infinite plane capacitor. In this case we are ignoring the direction z of a three-dimensional space, which can be done in cases like this, in which there is a complete translation symmetry in the direction of the third Cartesian coordinate, so that no quantity is dependent on that coordinate, effectively reducing the three-dimensional physical problem to a two-dimensional problem.

In order to complete this mapping of our mathematical structure onto a physical problem, we must recall that here x and y are dimensionless quantities, and relate them to the corresponding coordinates with dimension of length of the physical problem, which we call X and Y. For reasons that will become clear later, we put the two plates at positions corresponding to $y = \pm\pi$, so that we have the disposition of physical elements and corresponding dimensionless variables shown in Figure 5.1. It follows that the relations between these variables are as follows,

$$\begin{aligned} \frac{x}{\pi} &= \frac{X}{D/2} \\ &= \frac{2X}{D}, \end{aligned}$$

$$\begin{aligned}
\frac{y}{\pi} &= \frac{Y}{D/2} \\
&= \frac{2Y}{D},
\end{aligned}$$

so that we can write our complex potential as

$$\begin{aligned}
\phi(z) &= c - i\alpha z \\
&= (a + \alpha y) + \imath(b - \alpha x) \\
&= \left(a + \frac{2\pi\alpha Y}{D}\right) + \imath\left(b - \frac{2\pi\alpha X}{D}\right).
\end{aligned}$$

Naturally, the potential itself must also have appropriate dimensions in a physical problem, which means that in practice we must work with a potential $\Phi = B\phi$, where B is some constant with dimensions of Volts,

$$\Phi(z) = B\left(a + \frac{2\pi\alpha Y}{D}\right) + \imath B\left(b - \frac{2\pi\alpha X}{D}\right).$$

Placing definite boundary conditions, namely voltages defined on the two plates of the capacitor, we determine the constants a and α. Assuming that the potential is $B/2$ on the top plate, and $-B/2$ on the bottom plate, so that the voltage on the capacitor is B, it is easy to see that we should have $a = 0$ and $\alpha = 1/(2\pi)$ so that the real part of Φ can represent the electric potential. The constant b remains undetermined, and we just choose it to be zero, thus obtaining for the complex function Φ, whose real part is the potential,

$$\begin{aligned}
\Phi(z) &= B\phi(z) \\
&= B\left(\frac{1}{2\pi}\,y - \imath\,\frac{1}{2\pi}\,x\right) \\
&= B\left(\frac{1}{D}\,Y - \imath\,\frac{1}{D}\,X\right).
\end{aligned}$$

Another interesting fact that we can show about the behavior of u and v is related to the gradients of these functions of two variables or, more precisely, the *field lines* or *integral curves* of the two-dimensional vector fields formed by these gradients. The gradients are the vector fields given by

$$\begin{aligned}
\vec{\nabla}u &= \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y}\right), \\
\vec{\nabla}v &= \left(\frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}\right).
\end{aligned}$$

Using the Cauchy-Riemann conditions we can write the gradient of v in terms of u, getting

$$\vec{\nabla}v = \left(-\frac{\partial u}{\partial y}, \frac{\partial u}{\partial x}\right).$$

The exchange of the two components and the inversion of the sign betray the fact that what we have here is a rotation by $\pi/2$, that is, the two gradients are orthogonal each other, as can be seen directly,

$$\begin{aligned}\vec{\nabla}u \cdot \vec{\nabla}v &= -\frac{\partial u}{\partial x}\frac{\partial u}{\partial y} + \frac{\partial u}{\partial y}\frac{\partial u}{\partial x} \\ &= 0.\end{aligned}$$

It follows that the two gradients are perpendicular to one another, not only at some particular point, but *at all points in the complex plane* $\mathbb{C}$ where the function $w(z)$ is analytic. Since the integral curves of a vector field are tangent to the field vector at all its points, it follows that the integral curves of u and of v intersect at right angles at every point of the complex plane.

In this way, if u is interpreted as an electric potential, so that $-\vec{\nabla}u$ is proportional to an electric field, then the integral curves of v represent equipotential curves, which correspond to the equipotential surfaces in three-dimensional space. This follows from the fact that $\vec{\nabla}v$ points in the direction in which u does not vary at all, which can be seen from the fact that the variation of u due to a displacement $\vec{d\ell}$ of absolute value $d\ell$ in the direction of the versor of $\vec{\nabla}v$ is given by

$$\begin{aligned}du &= \vec{\nabla}u \cdot \vec{d\ell} \\ &= \vec{\nabla}u \cdot \frac{\vec{\nabla}v}{|\vec{\nabla}v|}\, d\ell \\ &= 0,\end{aligned}$$

that vanishes due to the vanishing of the dot-product of the two gradients. Having determined this fact, we can now put conducting metal surfaces along the equipotentials defined by the integral curves of the gradient of v, in order to construct the boundary conditions that, in addition to the Laplace equation, completely define an electrostatic problem. An example is given by our complex potential function Φ, because starting from

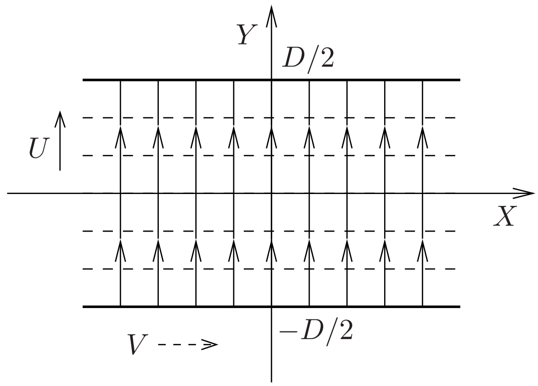


Figure 5.2: The (X, Y) plane with the integral curves of U and V.

$$
\begin{aligned}
\Phi(z) &= \frac{B}{D} Y - \imath \frac{B}{D} X \\
&= U + \imath V
\end{aligned}
$$

we can calculate the gradients of U and V, using partial derivatives with respect to X and Y,

$$
\begin{aligned}
\vec{\nabla} U &= (0, B/D), \\
\vec{\nabla} V &= (-B/D, 0),
\end{aligned}
$$

that have the dimensions of electric field, thus leading to the two sets of integral curves shown in Figure 5.2. Thus we meet here, for the very first time, with what is in fact a *boundary value problem*, involving a partial differential equation, as well as its solution. This subject will be studied in detail in later parts of this series of books.

In addition to this orthogonality relation involving the two gradients, we can also show that they always have the same absolute value, at all points where the function is analytic. For this purpose it suffices to verify that, while we have

$$
\begin{aligned}
\vec{\nabla} u &= \left(\frac{\partial u}{\partial x}, \frac{\partial u}{\partial y} \right) \quad \Rightarrow \\
|\vec{\nabla} u|^2 &= \left(\frac{\partial u}{\partial x} \right)^2 + \left(\frac{\partial u}{\partial y} \right)^2,
\end{aligned}
$$

we can use the Cauchy-Riemann conditions in order to write

$$\begin{aligned}\vec{\nabla} v &= \left(\frac{\partial v}{\partial x}, \frac{\partial v}{\partial y}\right) \\ &= \left(-\frac{\partial u}{\partial y}, \frac{\partial u}{\partial x}\right) \quad \Rightarrow \\ |\vec{\nabla} v|^2 &= \left(\frac{\partial u}{\partial y}\right)^2 + \left(\frac{\partial u}{\partial x}\right)^2,\end{aligned}$$

so that the absolute value of $\vec{\nabla} v$ ends up being the same as the absolute value of $\vec{\nabla} u$. This is due, of course, to the fact that the gradient of v is obtained from the gradient of u by a rotation, that does not change its absolute value. In particular, if one of the two gradients is zero then the other is also, an occurrence that happens only in this case. Note that if the two gradients are zero at a certain point, then both the real part and the imaginary part of the complex function do not vary to first order at that point. We shall see later on that this is related to the existence of a zero of the complex derivative of the function at the point in question.

Another very important aspect that has a geometrical character is the interpretation of analytic functions as mappings from the complex plane to the complex plane, that is, if we have a complex function $w(z) = u + \imath v$, where $z = x + \imath y$, then we can interpret the function as a mapping of points on a plane (x, y) to points on a plane (u, v),

$$(x, y) \xrightarrow{w} (u, v).$$

This is a point-to-point mapping between the two planes, which is not always invertible, that is, there may be points in which the inverse mapping does not exist, because two or more different points (x, y) are mapped to the same point (u, v). In general, it is more interesting to examine this type of mapping in terms of their action on curves given in the domain-plane (x, y). In general, since we are talking here of continuous functions, these curves are mapped to corresponding curves in the image-plane (u, v). If the function $w(z)$ is analytic, then the mapping defined by it has special properties, that are reflected on this relation between the domain and image curves.

Let us illustrate this type of mapping using as an example the function

$$w(z) = z + \frac{1}{z},$$

where $z = x + \imath y = \rho \exp(\imath\theta)$. Since it is the sum of two analytic functions, this is also an analytic function on the whole complex plane

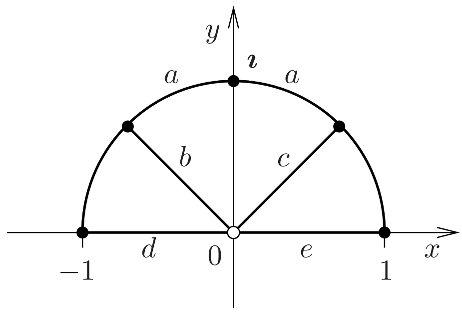


Figure 5.3: The unit upper semicircle in the (x, y) plane.

except for the point $z = 0$. If we had only the first term, we would have the identity map, $u = x$ and $v = y$, which maps each point of (x, y) to the identical point (u, v). The examination of the behavior of the second term alone will be left as an exercise. We will determine what this transformation does to a certain curve, the unit-radius semicircle with positive imaginary part shown in Figure 5.3.

Writing the function in terms of the polar representation we have

$$\begin{aligned} w(z) &= \rho\,\mathrm{e}^{\imath\theta} + \frac{1}{\rho}\,\mathrm{e}^{-\imath\theta} \\ &= \left(\rho + \frac{1}{\rho}\right)\cos(\theta) + \imath\left(\rho - \frac{1}{\rho}\right)\sin(\theta), \end{aligned}$$

and using this it is not difficult to see that the semicircle in the (x, y) plane determined by $\rho = 1$ and $\theta \in [0, \pi]$ is mapped onto the real segment given by $v = 0$ and $u \in [-2, 2]$ on the (u, v) plane. Similarly, it is not difficult to see that the radii of the upper half-disc defined in the (x, y) plane by θ constant within the interval $[0, \pi]$ and $\rho \in [0, 1]$ are mapped onto curves that begin at the segment just described and extend to infinity on the lower half-plane of the (u, v) plane. In particular, the segment $y = 0$, $x \in [0, 1]$ is mapped onto the semi-axis $v = 0$, $u \in [2, \infty]$, and the segment $y = 0$, $x \in [-1, 0]$ is mapped onto the semi-axis $v = 0$, $u \in [-\infty, 2]$, as shown in Figure 5.4.

With this it is clear that the interior of the upper half-disk in the (x, y) plane is mapped onto the entire lower half-plane in the (u, v) plane, that is, it is an infinite deformation of the upper half-disk, which is deformed and stretched to fill the entire lower half-plane.

One can also show that the curves in the (u, v) plane to which the radii are mapped depart at straight angles from the segment $u \in (-2, 2)$

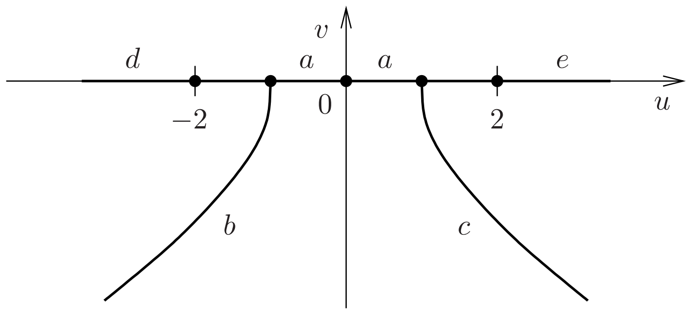


Figure 5.4: The (u, v) plane with the segment, the straight lines and the mapped curves.

onto which the semicircle is mapped. Moreover, it is interesting to note that the semicircle and the radii form those same right angles in the (x, y) plane. Thus these angles are being preserved by the transformation that maps one plane onto the other. One can verify that this happens at almost all the points, but there are two exceptions, since at the points $(-2, 0)$ and $(2, 0)$ of the (u, v) plane this angle changes from $\pi/2$ to π, so that suddenly we have a change of behavior.

Transformations that preserve the angles between curves in this way are called *conformal mappings*. In what follows, we will show that, under certain conditions, the mapping defined by an analytic function is always a conformal mapping. As we shall see, the relevant condition for the angles to be preserved at a given point is that the mapping be analytically invertible at the point in question, that is, that the *inverse* mapping exist and be non-singular at that point.

Consider then an analytic function $w(z)$, which maps points $z = (x, y)$ in the (x, y) plane to points $w = (u, v)$ in the (u, v) plane. Throughout this proof we will always be considering z and w as vectors in two dimensions, and will freely use vector notation in order to clarify the ideas. Given a curve defined by an equation $f(x, y) = 0$ in (x, y), this transformation produces a corresponding curve $g(u, v) = 0$ in (u, v). We will show here that, given two oriented curves C_1 and C_2 which intersect at an angle θ in (x, y), the two corresponding curves C_1' and C_2' in (u, v) intersect at the same angle $\theta' = \theta$, as illustrated in the diagram of Figure 5.5.

Let us consider, for this purpose, two variations of z, $\vec{dz}_1$ and $\vec{dz}_2$, that are tangent to the two curves, which intersect each other at a point (x, y), taken in the positive direction of each curve. The angle θ between

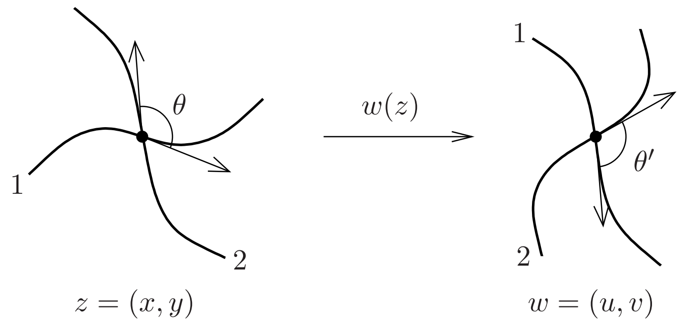


Figure 5.5: Two curves C_1 and C_2 mapped between the (x, y) plane and the (u, v) plane by the analytic function $w(z)$.

the two curves is given by the scalar product of versors

$$\cos(\theta) = \frac{\vec{dz}_1 \cdot \vec{dz}_2}{\left|\vec{dz}_1\right|\left|\vec{dz}_2\right|},$$

where $\vec{dz}_1 = (dx_1, dy_1)$, $\vec{dz}_2 = (dx_2, dy_2)$, and where we have

$$\begin{aligned} \left|\vec{dz}_1\right| &= \sqrt{(dx_1)^2 + (dy_1)^2}, \\ \left|\vec{dz}_2\right| &= \sqrt{(dx_2)^2 + (dy_2)^2}. \end{aligned}$$

These two variations $\vec{dz}_1$ and $\vec{dz}_2$ cause, by the mapping $w(z)$, two corresponding variations $\vec{dw}_1$ and $\vec{dw}_2$ on the (u, v) plane. The angle between them is given by

$$\cos(\theta') = \frac{\vec{dw}_1 \cdot \vec{dw}_2}{\left|\vec{dw}_1\right|\left|\vec{dw}_2\right|},$$

where $\vec{dw}_1 = (du_1, dv_1)$, $\vec{dw}_2 = (du_2, dv_2)$, and so on. Using the expression of the differentials of $u(x, y)$ and $v(x, y)$ in terms of the variations of x and y, we can write $\vec{dw}_{1,2}$ as functions of the corresponding $(dx, dy)_{1,2}$. Omitting for the moment the indices that identify the curves, we have

$$\begin{aligned} \vec{dw} &= (du, dv) \\ &= \left(\frac{\partial u}{\partial x}\, dx + \frac{\partial u}{\partial y}\, dy, \frac{\partial v}{\partial x}\, dx + \frac{\partial v}{\partial y}\, dy \right). \end{aligned}$$

Using the Cauchy-Riemann conditions we can write this exclusively in terms of u,

$$\vec{dw} = \left(\frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy, -\,\frac{\partial u}{\partial y}\,dx + \frac{\partial u}{\partial x}\,dy \right).$$

Note that this holds true for each of the two curves, and that the partial derivatives are the same in each case, that is, if we write this explicitly for $\vec{dw}_1$ and $\vec{dw}_2$, then the identification indices of each curve apply only to the variations (dx, dy). Calculating now the square modulus of $\vec{dw}$, we can see that the two mixed products that come from the square of each component cancel off due to the Cauchy-Riemann conditions, and we have therefore

$$\begin{aligned} \left|\vec{dw}\right|^2 &= \left(\frac{\partial u}{\partial x}\right)^2 (dx)^2 + \left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial u}{\partial y}\right) dxdy + \left(\frac{\partial u}{\partial y}\right)^2 (dy)^2 \\ &\quad + \left(\frac{\partial u}{\partial y}\right)^2 (dx)^2 - \left(\frac{\partial u}{\partial y}\right)\left(\frac{\partial u}{\partial x}\right) dxdy + \left(\frac{\partial u}{\partial x}\right)^2 (dy)^2 \\ &= \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right]\left[(dx)^2 + (dy)^2\right]. \end{aligned}$$

The quantities that remain are the square of the absolute value of the gradient of u and the square of the absolute value of $\vec{dz}$,

$$\left|\vec{dw}\right|^2 = |\vec{\nabla} u|^2 \left|\vec{dz}\right|^2.$$

There is nothing special about the presence of $|\vec{\nabla} u|$ here, because as we have seen $|\vec{\nabla} v|$ has the same value, and we could also write this equation in terms of it. We note again that when we write this separately for each curve, the gradient is the same, and the indices should be applied only to the variations $\vec{dw}$ and $\vec{dz}$.

We can also calculate the dot product of the variations $\vec{dw}_1$ and $\vec{dw}_2$. Of course, this time we need to keep the indices of the curves explicitly at every step. Again we verify that due to the Cauchy-Riemann conditions the mixed products that appear cancel off, so that it results that

$$\begin{aligned} \vec{dw}_1 \cdot \vec{dw}_2 &= \left(\frac{\partial u}{\partial x}\right)^2 dx_1 dx_2 + \left(\frac{\partial u}{\partial y}\right)^2 dy_1 dy_2 + \\ &\quad + \left(\frac{\partial u}{\partial x}\right)\left(\frac{\partial u}{\partial y}\right)(dx_1 dy_2 + dx_2 dy_1) + \\ &\quad + \left(\frac{\partial u}{\partial y}\right)^2 dx_1 dx_2 + \left(\frac{\partial u}{\partial x}\right)^2 dy_1 dy_2 + \end{aligned}$$

$$
\begin{aligned}
& -\left(\frac{\partial u}{\partial y}\right)\left(\frac{\partial u}{\partial x}\right)(dx_1 dy_2 + dx_2 dy_1) \\
= \; & \left[\left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial u}{\partial y}\right)^2\right](dx_1 dx_2 + dy_1 dy_2).
\end{aligned}
$$

This time the remaining quantities are the square of the absolute value of the gradient of u, once again, and the dot product of the variations $\vec{dz}_1$ and $\vec{dz}_2$,

$$\vec{dw}_1 \cdot \vec{dw}_2 = |\vec{\nabla} u|^2 \left(\vec{dz}_1 \cdot \vec{dz}_2\right).$$

Using the results above for the absolute values and the dot product of the variations $\vec{dw}_1$ and $\vec{dw}_2$, we can write for the expression of the angle in the (u, v) plane

$$
\begin{aligned}
\cos(\theta') &= \frac{\vec{dw}_1 \cdot \vec{dw}_2}{|\vec{dw}_1||\vec{dw}_2|} \\
&= \frac{|\vec{\nabla} u|^2 \left(\vec{dz}_1 \cdot \vec{dz}_2\right)}{|\vec{\nabla} u||\vec{dz}_1||\vec{\nabla} u||\vec{dz}_2|}.
\end{aligned}
$$

Before making the cancellations that are suggested by this formula, let us draw attention to the fact that we can only write it if $|\vec{\nabla} u| \neq 0$, which also implies that $|\vec{\nabla} v| \neq 0$. Otherwise, the quantities in the numerator and in the denominator are both zero and the expression that defines $\cos(\theta')$ remains undetermined. If we have $\vec{\nabla} u = 0 = \vec{\nabla} v$ at a certain point, then it follows that $w = (u, v)$ does not vary at that point when x and y vary in an infinitesimal way around the corresponding point in the domain-plane. Thus, all values of $z = (x, y)$ infinitesimally near that point are mapped to essentially the same value w, indicating that the mapping is not analytically invertible at that point. Thus we see that this theorem only holds at points where the transformation is invertible, with an analytic inverse (which corresponds to the associated analytic function having non-zero complex derivative at the point). In this case we can make the cancellations, and we then have

$$
\begin{aligned}
\cos(\theta') &= \frac{\vec{dw}_1 \cdot \vec{dw}_2}{|\vec{dw}_1||\vec{dw}_2|} \\
&= \frac{\vec{dz}_1 \cdot \vec{dz}_2}{|\vec{dz}_1||\vec{dz}_2|},
\end{aligned}
$$

where, in the last version of the formula, we recognize the expression of $\cos(\theta)$, which gives the angle in the (x, y) plane, so it results that

$$\cos(\theta') = \cos(\theta),$$

that is, the angle between the curves is left invariant by the transformation defined by $w(z)$. In truth, in order to conclude this we must also show that $\sin(\theta') = \sin(\theta)$, which we will leave as an exercise. Thus, we see that the mapping defined by an analytic function is a conformal transformation at all points where both the mapping and its inverse transformation are well defined and are analytic, without singularities.

In a way, this conservation of angles is also related to the possibility of using these mappings to transform the solutions of electrostatic problems into one another. As we have seen, the integral curves of the gradients of the real and imaginary parts of analytic functions can be interpreted as the field lines and equipotential surfaces of electrostatics, which are always perpendicular to each other. It is therefore necessary that these mappings preserve angles, so that this property of the electrostatic solution in the domain-plane can be preserved in the corresponding solution in the image-plane. This is also guaranteed by the analyticity of the functions involved, because as we shall see the potential Φ' on the (u, v) plane is associated with the potential Φ on the (x, y) plane and with the function $w(z)$ by the composition of functions. The point here is that the composition of two analytic functions generates a function that is also analytic, as we have shown before.

Of course this same technique of solution, which we will examine in the next chapter, can be used to help solve other physical problems that can be reduced to the Laplace equation in two dimensions, which includes for example heat conduction problems and problems involving two-dimensional fluid flow.

Problem Set

1. Let us call stationary those points where a function $f(x, y)$ of two variables (x, y) has its two first partial derivatives with respect to these variables equal to zero. Assuming that at least one of the two second partial derivatives is non-zero, a stationary point of this type can be a point of local maximum of the function, a point of local minimum of the function, an inflection point or a saddle point. At a point of local minimum the two second partial

derivatives are strictly positive, at a point of local maximum the two second partial derivatives are strictly negative, at an inflection point one of the second derivatives is zero while the other is strictly positive or strictly negative, and at a saddle point one of the second derivatives is strictly positive while the other is strictly negative. Prove that the functions $u(x, y)$ and $v(x, y)$ which constitute an analytic function cannot have any points of local minimum, local maximum, or inflection, but only saddle points.

2. Consider an arbitrary analytic function $w(z) = u(x, y) + \imath v(x, y)$ of the complex variable $z = x + \imath y$.

   (a) Write the gradient vectors in the (x, y) plane of the two harmonic functions $u(x, y)$ and $v(x, y)$.

   (b) Show that these two gradient vectors are orthogonal to each other, at any point of the (x, y) plane where $w(z)$ is analytic.

   (c) Show that these two gradient vectors have the same absolute value at every point of the (x, y) plane where $w(z)$ is analytic.

3. Consider the analytic function $w(z) = 1/z$ as a mapping from the complex plane $z = (x, y)$ onto the complex plane $w = (u, v)$.

   (a) Show that the unit circle is mapped onto itself, but not as the identity map.

   (b) Find the fixed points of the mapping, that is, points that are mapped to themselves.

   Answer: ± 1.

   (c) Show that the interior of the unit circle is mapped onto its exterior, and vice-versa.

4. Consider the complex function

$$w(z) = z + \frac{1}{z},$$

where $z = x + \imath y$ and $w = u + \imath v$, seen as a transformation that maps the complex (x, y) plane onto the complex (u, v) plane.

   (a) Show that this transformation maps the unit semicircle with positive imaginary part of the (x, y) plane onto the real segment $(-2, 2)$ of the (u, v) plane.

(b) Show that this transformation maps the radii of the unit disk that fall strictly within this upper unit semicircle onto curves starting in the real segment $(-2, 2)$, extending to infinity on the (u, v) half-plane with negative imaginary part.

(c) Show that the curves mentioned in the previous item intersect the segment $(-2, 2)$ perpendicularly, on the (u, v) plane.

(d) Show that the interior of the semicircle is mapped onto the whole lower half-plane in the (u, v) plane.

5. Consider the two oriented curves C_1 and C_2 in the complex (x, y) plane, which intersect at a certain point, and the transformation defined by the analytic function $w(z)$, which maps these two curves onto two other curves C_1' and C_2' in the complex (u, v) plane, as discussed in the text. Assume that the gradients of $u(x, y)$ and $v(x, y)$ are not zero at the intersection point. Consider two infinitesimal variations of z at the intersection point in the (x, y) plane, denoted in vector language by $\vec{dz}_1$ and $\vec{dz}_2$, each one tangent to and pointing in the positive direction of the corresponding curve.

   (a) Show that the sine of the angle θ between the two curves in the (x, y) plane is given by

   $$\sin(\theta) = \frac{dx_1 dy_2 - dx_2 dy_1}{|\vec{dz}_1||\vec{dz}_2|},$$

   where $\vec{dz}_1 = dx_1 + \imath dy_1$ and $\vec{dz}_2 = dx_2 + \imath dy_2$. Show in the same way that the sine of the angle θ' between the corresponding curves C_1' and C_2' in the complex (u, v) plane, at the intersection point, is given by

   $$\sin(\theta') = \frac{du_1 dv_2 - du_2 dv_1}{|\vec{dw}_1||\vec{dw}_2|},$$

   where $\vec{dw}_1 = du_1 + \imath dv_1$ and $\vec{dw}_2 = du_2 + \imath dv_2$.

   Hint: consider using vector (cross) products.

   (b) Prove, using the analyticity properties of the transformation function $w(z)$, that $\sin(\theta') = \sin(\theta)$. Together with the result that was shown in the text, $\cos(\theta') = \cos(\theta)$, this suffices to ensure that $\theta' = \theta$, thus completing the proof started in the text.

Chapter 6

Border Effects in Capacitors

The interpretation of analytic functions as conformal mappings, that was studied in the previous chapter (Chapter 5), can be used as a technique for solving difficult problems that often cannot be solved otherwise except by numerical methods. This may be applied to any problems that can be reduced to the resolution of the two-dimensional Laplace equation with certain boundary conditions. This includes, for example, electrostatic problems, heat conduction problems and certain problems involving fluid flow around solids. We will now illustrate the usefulness of these transformations with a physical example in electrostatics. It can be shown that the complex function given by

$$w(z) = z + \mathrm{e}^{z}$$

maps the lines $y = \pm\pi$ of the plane $z = (x, y)$ onto the semi-axes $v = \pm\pi$, $u \leq -1$ of the plane $w = (u, v)$. The mapping is twofold, that is, both the positive half and the negative half of each one of the two lines in the (x, y) plane are mapped separately onto the corresponding semi-axes of the (u, v) plane. Furthermore, one can show that the strip $-\pi \leq y \leq \pi$ of the (x, y) plane is mapped onto the entire (u, v) plane. One can also show that this conformal transformation is invertible in the strip of the (x, y) plane between $y = -\pi$ and $y = \pi$, wherein the only points the conformal transformation is not analytically invertible are $(0, -\pi)$ and $(0, \pi)$. In order to do all this we must recall that $z = x + \imath y$ and that $w = u + \imath v$, so that we have

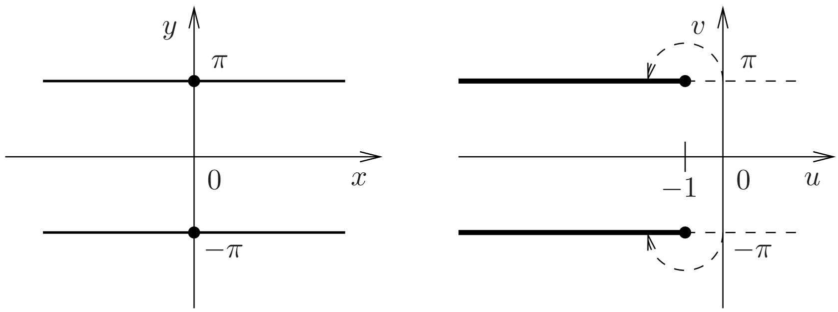


Figure 6.1: The (x, y) and (u, v) planes with straight lines and mapped semi-axes.

$$\begin{aligned} w(z) &= z + \mathrm{e}^z \\ &= x + \imath y + \mathrm{e}^x[\cos(y) + \imath \sin(y)] \\ &= [x + \mathrm{e}^x \cos(y)] + \imath[y + \mathrm{e}^x \sin(y)], \end{aligned}$$

which implies that in terms of x, y, $u(x, y)$ and $v(x, y)$ the transformation equations are

$$\begin{aligned} u(x, y) &= x + \mathrm{e}^x \cos(y), \\ v(x, y) &= y + \mathrm{e}^x \sin(y). \end{aligned}$$

The details of these derivations will be left as exercises. This is a remarkable deformation, as can be seen. It is as if two open hinges had been placed at the points $(0, -\pi)$ and $(0, \pi)$ of the (x, y) plane, these points moving to the points $(-1, -\pi)$ and $(-1, \pi)$ of the (u, v) plane while the hinges are closed, so that the two positive parts of the lines rotate and sweep the rest of the plane, the top semi-axis in the counterclockwise direction and the lower one in the clockwise direction. Thus, the strip of the (x, y) plane is stretched and extended to the whole (u, v) plane, as illustrated in the diagram of Figure 6.1.

Note that this deformation maps the interior of the strip between two metal plates, that is, the interior of an infinite plane capacitor, where the solution of the electrostatic problem is simple and well known, onto the problem of the capacitor which is finite on one side in the horizontal direction, at which it is open to the rest of the space. The latter problem is a very difficult electrostatic problem, one which we do not know how to solve analytically in explicit form with the most common mathematical tools. The issue at hand is the determination of the electric field at the edge of a capacitor, which turns out to be represented by the deformation implemented by this transformation. Note that in the (u, v) plane the distance between the edges of the plates and the axis $u = 0$ is relatively small, because the ratio between it and the distance between the two plates is $1/(2\pi)$ or about one-sixth.

As we saw in the previous chapter (Chapter 5), we have a complex function $\Phi(Z)$ whose real part $\Lambda(X, Y)$ represents the electric potential within the infinite strip in (X, Y), where $\Phi(Z)$, $\Lambda(X, Y)$, X and Y have the appropriate physical dimensions. It is related to a corresponding dimensionless analytic function $\phi(z)$, with real part $\lambda(x, y)$, which has the same mathematical properties, but where all quantities are dimensionless. For simplicity of presentation, in this chapter we will use only the dimensionless quantities throughout most of the arguments. The relations with the corresponding quantities that have the correct physical dimensions will be listed in a dictionary of dimensionfull quantities at the end of the chapter. The dimensionless complex electric potential is given by

$$\phi(z) = \frac{1}{2\pi} y - \imath \frac{1}{2\pi} x,$$

for $-\infty < x < \infty$ and $-\pi \leq y \leq \pi$, whose real part $\lambda(x, y) = y/(2\pi)$ is a dimensionless version of the electrostatic potential of the physical problem. Note that the electric field is zero and therefore the potential is constant for $y < -\pi$ and for $\pi < y$, so that the two straight lines $y = \pm\pi$ represent the boundaries of the region of analyticity of $\phi(z)$, in the context of this application.

This function is a solution of the electrostatic problem in (x, y) because, since $\phi(z)$ is an analytic function, its real part $\lambda(x, y) = y/(2\pi)$ satisfies the Laplace equation on the variables (x, y), and also satisfies the boundary conditions, because it has the correct values $\lambda(x, \pm\pi) = \pm 1/2$ at the two planes $y = \pm\pi$ forming the capacitor plates. This real part $\lambda(x, y)$ is therefore the electric potential in the domain that corresponds to the interior of the capacitor in the (x, y) plane. If we compose this function with the function $w = z + \exp(z)$ which defines the conformal transformation the result is a function $\phi'(w)$ on the (u, v) plane, which is also analytic, because it is a composition of analytic functions. It follows that the real part $\lambda'(u, v)$ of $\phi'(w)$ satisfies the Laplace equation on (u, v), and also satisfies the boundary conditions, since the boundaries are curves that have been mapped from (x, y) to (u, v). Thus we see that the conformal transformation maps the known solution of the problem in (x, y) onto the solution we want to determine in (u, v).

It is important to keep in mind that, in order to represent the physical problem on the image-plane (u, v), one must introduce in it dimensionfull coordinates, just as we did in (x, y). It suffices to take in this case new dimensionfull variables U and V defined by relations similar to those that define X and Y, in order to preserve the distance D between the capacitor plates. The definition of the composite function $\phi'(w)$ from $\phi(z)$ and $w(z)$ is given as follows: given a point $w = (u, v)$ on the image-plane, first determine the corresponding value $z = (x, y)$ on the domain-plane, inverting the function $w(z)$ and thereby determining x and y. Having done this, we verify what is the value of $\phi(x, y)$ on the domain-plane, and associate this value to $\phi'(u, v)$. In other words, we define that $\phi'(u, v) = \phi(x, y)$, where (u, v) is the image of (x, y) by the conformal transformation. This clearly preserves the values of the potential on the surfaces used for the boundary conditions. Thus we see that in fact the function $\phi'(w)$ is the composition of $\phi(z)$ with the *inverse* transformation function $z(w)$ of the conformal transformation given by $w(z)$, and thus we can write for the definition of $\phi'(w)$

$$\phi'(w) = \phi(z(w)) .$$

The practical problem in this procedure is of an algebraic order, because in this case the function $w(z)$ cannot be inverted by purely algebraic means. In order to see this it suffices to write the transformation explicitly in terms of real functions and real variables, and therefore recall that in terms of the variables x, y, u and v the transformation equations are

$$\begin{aligned} u &= x + \mathrm{e}^x \cos(y), \\ v &= y + \mathrm{e}^x \sin(y). \end{aligned}$$

Since we have here a couple of transcendental equations, we cannot invert them algebraically in order to obtain explicit forms for $x(u, v)$ and $y(u, v)$. Note that there are two regimes for the mapping between x and u on the plates of the capacitors, which correspond to the fact that there is a double mapping between the plates in the (x, y) plane and the half-plates in the (u, v) plane. For example on the top plate, that is, for $y = \pi$, since $\cos(\pi) = -1$ and $\sin(\pi) = 0$ we have that $v = y$, and the relation

$$u = x - e^x,$$

so that the semi-axis $x \in (-\infty, 0]$ is mapped onto the semi-axis $u \in (-\infty, -1]$, while the semi-axis $x \in [0, \infty)$ is mapped onto that same semi-axis $u \in (-\infty, -1]$. The first of these two mappings describes in fact the *inner surface* of the upper plate in the (u, v) plane, while the second describes the *outer surface* of that same plate. Note that the first mapping tends to the identity $u = x$ for large negative values of x, unlike what happens with the second. These are two different mappings, thus showing once again the lack of analyticity over the plate.

Since we have these transcendental equations describing the mappings, the solution to the potential $\phi'(w)$ can be given only implicitly, by means of

$$\begin{aligned} \phi'(u, v) &= \phi(x, y) \\ &= \frac{1}{2\pi}\, y(u, v) - \boldsymbol{\imath}\, \frac{1}{2\pi}\, x(u, v), \end{aligned}$$

where the values of x and y to be used in the right-hand side must be calculated in terms of the given values of u and v by means of the inversion of the conformal transformation. We have therefore for the real part $\lambda'(u, v)$ of $\phi'(w)$, which is the electric potential on the (u, v) plane,

$$\lambda'(u, v) = \frac{1}{2\pi}\, y(u, v),$$

where

$$\begin{aligned} u &= x + e^x \cos(y), \\ v &= y + e^x \sin(y). \end{aligned}$$

Although this limitation of algebraic manipulation does not allow us to explicitly show the solution $\lambda'(u, v)$, we must recognize that the mathematical technique is powerful because it transformed an analytical problem, that of solving analytically or numerically a partial differential equation, into an algebraic problem, a much simpler one, that of solving the pair of algebraic equations above, something that can be done numerically without too much difficulty.

However, although we have no explicit solution, we can still proceed with the calculation of other physical quantities of interest. A quantity of great interest in capacitors is the amount of electric charge stored on the plates. In order to examine this aspect of the problem, let us first calculate the surface charge density on the plates. This is related to the component of the electric field that is normal to the plates, which in turn relates to the normal derivative of the potential, that is, to the derivative with respect to the outer normal to the surface of the metal plates. In the MKSA system of units, we have for the surface charge density, with the usual physical dimensions,

$$\begin{aligned} \Sigma &= \epsilon_0 E_n \\ &= -\epsilon_0 \partial_n \Lambda, \end{aligned}$$

where ϵ_0 is the dielectric permeability of the vacuum and ∂_n is the normal derivative. On the (x, y) plane we can execute this immediately, and we have in this case the known result, in its dimensionless version,

$$\begin{aligned} \sigma &= \epsilon_0 \frac{\partial \lambda}{\partial y} \\ &= \frac{\epsilon_0}{2\pi}, \end{aligned}$$

which is the charge density on the *inner* surface of the upper plate, which is constant over the entire plate. On the outer surface of the plate the charge density is zero. Naturally, the result on the lower plate is the same, except for the inverted sign. On the (U, V) plane we have, on the inner surface of the upper plate, the surface charge density, in its dimensionless version,

$$\begin{aligned} \sigma'(u) &= \epsilon_0 \frac{\partial \lambda'}{\partial v} \\ &= \frac{\epsilon_0}{2\pi} \frac{\partial y}{\partial v}. \end{aligned}$$

Note that this expression applies to the *inner* surface of the top plate, but that in this case we can have normal components of the electric field and surface charge densities both at the internal surface and at the external surface. In the case of the outer surface the direction of the normal to the surface changes sign, so that there would be a corresponding change of sign in this expression. We can calculate the derivative shown here, using for this the equations of transformation between (x, y) and (u, v). This is not as hard as it may seem, because we are interested in calculating the derivative only in the vicinity of the plates, that is for

$v = \pi$, in which case the equations are simplified. Since $\cos(\pi) = -1$ and $\sin(\pi) = 0$, in this case we have the simpler relations between the dimensionless variables (x, y) and (u, v),

$$\begin{aligned} u &= x - \mathrm{e}^x, \\ v &= y, \end{aligned}$$

that is, the fact that u is constant (since we are interested in a partial derivative with respect to v) implies that x is constant, so that it is sufficient to calculate the derivative of y with respect to v with both x and u constant. Another way to see this is to use the fact that both the field lines of $\vec{\nabla}\lambda$ are normal to the plates in the (x, y) plane and the field lines of $\vec{\nabla}\lambda'$ are normal to the plates in the (u, v) plane. Since we are taking normal derivatives in the vicinity of the plates, these derivatives are in reality tangent to the field lines along y or v, which may be regarded as equipotentials along x or u. It follows that these derivatives can be taken indifferently in the direction of y or in the direction of v, and that both x and u are constant on the variations involved in these derivatives. The easiest quantity to calculate is

$$\frac{\partial v}{\partial y} = 1 + \mathrm{e}^x \cos(y),$$

which, inverted and written at the point $y = \pi$, gives us

$$\frac{\partial y}{\partial v} = \frac{1}{1 - \mathrm{e}^x}.$$

It follows therefore that we have for the surface charge density on the inner surface of the upper plate of the (u, v) plane, in its dimensionless version,

$$\sigma'(u) = \frac{\epsilon_0}{2\pi} \frac{1}{1 - \mathrm{e}^x}.$$

We now see that this density is not a constant, but a function of u, which is given through x. This is still written in implicitly form because given a value of u, with $v = \pi$, so that we are somewhere on the top plate, it is necessary to invert the transformation to find the corresponding value of x. The situation is, however, simpler than the one we found before for the potential because we have $y = v = \pi$, so that there is only one transcendental algebraic equation to solve,

$$u(x) = x - \mathrm{e}^x.$$

It is clear that in principle we would need to calculate the charge density both on the inner surface and on the outer surface, but here we will do the calculation only on the inner surface, leaving the examination of the case of the outer surface, which is more difficult, to another opportunity. In this equation we use x in the interval $(-\infty, 0]$ and u in the interval $(-\infty, -1]$, thus characterizing the calculation on the inner surface. It is not difficult to show that the only point of extreme of the function $u(x)$ is at $x = 0$, and therefore the maximum value $u = -1$ of u corresponds to the maximum value $x = 0$ of x. Thus, examining the behavior of the factor

$$\frac{1}{1 - \mathrm{e}^x}$$

contained in the solution for $\sigma'(u)$, which represents a multiplicative correction with respect to the constant approximate result that ignores the edge effects, we see that $\sigma'(u)$ tends very quickly to become constant as we enter into the capacitor, making x decrease from zero. This represents the approach to the typical regime of a infinite plane capacitor, and it is exponentially fast. The value of $\sigma'(u)$ in this limit is equal to the original value of σ on the (x, y) plane.

On the other hand, $\sigma'(u)$ increases quickly as we approach the edge, making x approach zero, and indeed diverges at the position of the edge. This is a well-known physical phenomenon, the migration of charges to the edges and tips, and their higher concentration there. The divergence of $\sigma'(u)$ on the edge indicates that there is a large concentration of charges there, and can even give the impression that along the edge a linear distribution of charge with a singular character arises. However, this is *not* the case here. The initial impression that this seems to be a non-integrable divergence is just that, an impression, since although

$$\int \frac{dx}{1 - \mathrm{e}^x}$$

is in fact non-integrable near $x = 0$, the integral that is of interest to us for the calculation of the total charge is in fact with respect to u,

$$\int \frac{du}{1 - \mathrm{e}^x},$$

where x is an implicit function of u. We can verify this fact by writing this explicitly in terms of u, which may be done in an approximate way in the vicinity of the edge where $y = v = \pi$, $x \approx 0$ with $x < 0$, and $u \approx -1$ with $u < -1$. However, in reality we can do better than this, for it is indeed possible to make an accurate and explicit calculation of the total charge on the internal surface, as discussed below. Note that we see in this way that the usual approximation for calculating the capacitance of a finite plane capacitor, in which we ignore the edge effects, ends up *minimizing* the value of the capacitance, which is actually somewhat *larger* than the result of that approximate calculation.

We may now proceed to consider the calculation of the total charge on the internal surface of the upper plate, which is expressed as an integral of $\sigma'(u)$ on the surface of the plate. Of course, in order to do this we must make the plate big but finite, because otherwise the result will inevitably diverge. Therefore, let us assume that we have rectangular plates with a dimension l_z in a direction of z, perpendicular to the complex planes, and with a dimension l_u in the direction of u, such that $2\pi \ll l_u \ll l_z$. We can avoid the divergence of the integral in the direction of u by calculating only the variation between the two calculations, the exact one and the traditional approximation, which assumes that the charge density is constant along the inner surface of the plate. In this case it is not necessary to introduce l_u, but only l_z, and the integral to be calculated is that of the difference $\sigma'(u) - \sigma$, which can be written as

$$\begin{aligned} \sigma'(u) - \sigma &= \frac{\epsilon_0}{2\pi}\left(\frac{1}{1 - \mathrm{e}^x} - 1\right) \\ &= \frac{\epsilon_0}{2\pi}\left(\frac{\mathrm{e}^x}{1 - \mathrm{e}^x}\right). \end{aligned}$$

The total charge on the inner surface of the upper plate in its dimensionless version is given by

$$q' = l_z \int_{-\infty}^{-1} du\, \sigma'(u).$$

We thus have for the additional charge on the inner surface of the upper plate, as compared to the infinite-plane approximation,

$$\begin{aligned} \Delta q' &= l_z \int_{-\infty}^{-1} du\, \left[\sigma'(u) - \sigma\right] \\ &= l_z\, \frac{\epsilon_0}{2\pi} \int_{-\infty}^{-1} du\, \left(\frac{\mathrm{e}^x}{1 - \mathrm{e}^x}\right). \end{aligned}$$

We now verify that it is very simple to change integration variables from u to x in this integral, since due to the transformation among the variables for the case $y = v = \pi$ we have

$$\begin{aligned} u &= x - \mathrm{e}^x \quad \Rightarrow \\ du &= dx - \mathrm{e}^x dx \\ &= (1 - \mathrm{e}^x) dx, \end{aligned}$$

that is,

$$dx = \frac{du}{1 - \mathrm{e}^x},$$

so that we can rewrite our integral in a trivial way as

$$\Delta q' = l_z \frac{\epsilon_0}{2\pi} \int_{-\infty}^{0} dx \ \mathrm{e}^x,$$

where we have changed the upper extreme of the integration interval appropriately. Again we see here the exponential decay involved in the exact solution, since the integrand goes to zero exponentially fast when we move away from the edge, showing the fact that almost all the additional charge $\Delta q'$ is close to the edge.

It is interesting to observe that this same transformation, if applied to the integral that gives the total charge q' on the inner surface of the upper plate on the (u, v) plane, transforms the integrand into a constant, but integrated over the semi-axis $(-\infty, 0]$ instead of $(-\infty, -1]$, the constant being the same that appears in the case of the integral that gives the total charge q in the case of the (x, y) plane. This gives us an interesting physical picture of what happens to the charge during this transformation, going from the infinite capacitor to the semi-finite capacitor: the charge that would be distributed evenly in the interval $[-1, 0]$ ends up migrating to the region just to the left of $u = -1$, concentrating there and thus increasing the charge density in this edge region.

We can now calculate the integral, in order to obtain a result in explicit form,

$$\begin{aligned} \Delta q' &= l_z \frac{\epsilon_0}{2\pi} \left(\mathrm{e}^0 - \mathrm{e}^{-\infty} \right) \\ &= l_z \frac{\epsilon_0}{2\pi}. \end{aligned}$$

Note that this result depends on l_z, but does not dependent on any dimension related to a length l_u, which could involve in some way an approximation relating to how deeply we should enter the capacitor in the negative u direction, in order to reach the constant charge density limit. What we see here is that the variation of the charge density is strictly an edge effect, measured in terms of the distance 2π between the plates. The corresponding variation of the capacitance C is obtained by dividing the corresponding variation of the dimensionfull charge $\Delta Q'$ by the voltage B between the two capacitor plates,

$$\begin{aligned}\Delta C &= \frac{\Delta Q'}{B} \\ &= \frac{BD}{2\pi}\frac{\Delta q'}{B} \\ &= \frac{D}{2\pi}\, l_z\, \frac{\epsilon_0}{2\pi} \\ &= L_Z\, \frac{\epsilon_0}{2\pi}.\end{aligned}$$

From this point on we will use the dimensionfull variables to express the correction to the capacitance. It is important to keep in mind that the correction calculated here is only the part that relates to the charges on the internal surfaces of the plates. In order to get a better understanding of the physical meaning of this result, we normalize it by dividing it by the result of the usual approximate calculation for C, in which we disregard the edge effects, which is $C_0 = \epsilon_0 L_Z L_U/D$, thus obtaining the relative variation of C,

$$\frac{\Delta C}{C_0} = \frac{D}{2\pi L_U}.$$

Here the quantity $D/(2\pi)$ is the horizontal distance between the edges of the plates and the V axis, that is, the width of the region from which came the charge concentrated at the edge of the plate. This quantity also gives the length scale of the exponential decay of the correction to the charge density as one enters the interior of the capacitor. We can make explicit the correction to the approximate value C_0 writing this expression in the form

$$C = C_0\left(1 + \frac{D}{2\pi L_U}\right).$$

In order to write a result with a more practical applicability, we can consider a capacitor in the shape of an infinite strip with width L_U, in which case one needs to assume the situation $D \ll L_U \ll L_Z$, and recall that in this case we have two edges, with a charge build-up of this type in each one. So long as $D/(2\pi)$ is sufficiently smaller than L_U, the central region of the plates will be very approximately in the constant surface density regime, so that we can simply add the edge corrections for each of the two edges, and the following result will be valid to good approximation

$$\frac{\Delta C}{C_0} = \frac{D}{\pi L_U},$$

where one should observe that we multiplied the correction by a factor of 2, in such a way as to take into account the two edges of the strip-shaped plate, that is, we can explicitly write the correction to the approximate value C_0 as

$$C = C_0 \left(1 + \frac{D}{\pi L_U} \right).$$

Since the convergence to the constant surface density regime is exponentially fast when one enters the interior of the capacitor from the edge, on the scale of the dimensionfull quantity $D/(2\pi)$, it follows that L_U does not actually need to be much larger than D for this result to be valid with good accuracy. Anyway, for typical values of D and L_U the error introduced by this approximation is very small.

Furthermore, we can use the exact result for the semi-infinite plate in order to build a practical result which is even better than this, taking into consideration all the edges of the plates. This result is perfectly applicable to three-dimensional situations. Returning to the formula for the additional capacitance ΔC, we see that the quantity

$$\frac{\Delta C}{L_Z} = \frac{\epsilon_0}{2\pi}$$

is the additional capacitance per unit length of the edge of the plate, since L_Z is the length of that edge. So long as both L_U and L_Z are sufficiently larger than D, we can calculate the total additional capacitance using this formula for the whole perimeter of the plates. If P is this perimeter, and ignoring the format of the plates except for the conditions $D \ll L_U$ and $D \ll L_Z$, then we have that

$$\Delta C = \frac{\epsilon_0}{2\pi} P.$$

In this case we have for the corrected capacitance

$$C = C_0 \left(1 + \frac{PD}{2\pi A} \right),$$

where A is the area of the plates, whose shape is left undefined. The only things that are not being taken into account explicitly in this result are eventual *tip* corrections, such as those at the four corners of square or rectangular plates, which presumably are even smaller than the edge corrections that we calculated here. Note that our result applies equally to plates of other formats, possibly without any sharp tips, such as to circular plates. The result will be even more precise when the edges of the plate curve as little as possible, as in the case of the circle. In the case of rectangular plates such as the ones we are considering here, the above result reduces to

$$C = C_0 \left[1 + \frac{(L_U + L_Z)D}{\pi L_U L_Z} \right],$$

which, in the case that $L_Z \gg L_U$, reduces to our previous result.

Thus we see that the mathematical techniques examined here can be of great value in dealing with problems that would be very hard to deal with otherwise. Even when one cannot exhibit exact results in explicit form, these analytic methods allow us to understand physically relevant aspects of the problem, and answer qualitative questions about it. In some cases it is possible to present approximate high-quality results for some relevant quantities. Of course, in order to permit a direct comparison of our results here with experiments, first it would be necessary to complete our analysis, taking into account the possibility that there is a charge build-up also on the external surfaces of the plates, near the edges.

Dimensionary

This is a dictionary of dimensionfull quantities. The dimensionless quantities are represented by symbols using lower-case letters, and the corresponding quantities with physical dimensions by symbols using upper-case letters. The dimensionfull quantities are always to the left in the equations.

- Cartesian coordinates and the complex variable in the complex Z plane:

$$\begin{aligned}\frac{2X}{D} &= \frac{x}{\pi},\\ \frac{2Y}{D} &= \frac{y}{\pi},\\ \frac{2Z}{D} &= \frac{z}{\pi},\end{aligned}$$

where $Z = X + \imath Y$ and D is the spacing between the two plates of the capacitor.

- Cartesian coordinates and the complex variable in the complex W plane:

$$\begin{aligned}\frac{2U}{D} &= \frac{u}{\pi},\\ \frac{2V}{D} &= \frac{v}{\pi},\\ \frac{2W}{D} &= \frac{w}{\pi},\end{aligned}$$

where $W = U + \imath V$ and D is the spacing between the two plates of the capacitor.

- Complex electric potential in the complex Z plane:

$$\Phi(Z) = B\phi(z),$$

where B is the potential difference applied to the capacitor. The electric potential itself is the real part, which is given by

$$\Lambda(X, Y) = B\lambda(x, y).$$

- Complex electric potential in the complex W plane:

$$\Phi'(W) = B\phi'(w),$$

where B is the potential difference applied to the capacitor. The electric potential itself is the real part, which is given by

$$\Lambda'(U, V) = B\lambda'(u, v).$$

- Surface charge density:

$$\begin{aligned}\frac{D}{2B}\Sigma &= \pi\sigma,\\ \frac{D}{2B}\Sigma'(U) &= \pi\sigma'(u).\end{aligned}$$

- Particular lengths:

$$\begin{aligned}\frac{2L_Z}{D} &= \frac{l_z}{\pi},\\ \frac{2L_U}{D} &= \frac{l_u}{\pi}.\end{aligned}$$

- Electric charge:

$$\frac{2Q'}{BD} = \frac{q'}{\pi}.$$

Problem Set

1. Consider the complex variables $z = x + \imath y$ and $w = u + \imath v$, and the analytic function

$$w(z) = z + \mathrm{e}^z,$$

interpreted as a mapping or transformation from the complex plane $z = (x, y)$ onto the complex plane $w = (u, v)$.

(a) Show that this transformation maps the straight line $y = 0$ of the (x, y) plane onto the straight line $v = 0$ of the (u, v) plane.

(b) Show that this transformation maps the two semi-axes of the (x, y) plane given by $y = \pm\pi$ with $x \geq 0$ onto the two semi-axes of the (u, v) plane given by $v = \pm\pi$ with $u \leq -1$, respectively, and that it maps the two semi-axes of the (x, y) plane given by $y = \pm\pi$ with $x \leq 0$ onto these same two semi-axes of the (u, v) plane.

(c) Show that this transformation maps the strip given by the interval $-\pi \leq y \leq \pi$ of the (x, y) plane onto the entire (u, v) plane. In order to do this consider the mappings of families of curves.

(d) Calculating $|\vec{\nabla} u|$ and/or $|\vec{\nabla} v|$, show that this transformation is analytically invertible within the strip of the (x, y) plane between $y = -\pi$ and $y = \pi$, including the boundary of this region, except for the two points $(0, \pi)$ and $(0, -\pi)$ located on this boundary.

2. Show that the integral

$$\int_{-u_0}^{-1} \frac{du}{1 - \mathrm{e}^{x(u)}},$$

where

$$u(x) = x - \mathrm{e}^{x},$$

is finite for any finite value of the constant u_0 in the interval $[1, \infty)$.

3. Consider the complex function $w(z) = z^2$ where $z = x + \imath y$ and $w = u + \imath v$.

(a) Find the equipotential curves of u and v, that is, those in which these quantities are constant.

(b) Calculate the electric-field vectors $\vec{E}_u = -\vec{\nabla} u$ and $\vec{E}_v = -\vec{\nabla} v$.

(c) Show that $\vec{E}_u \cdot \vec{E}_v = 0$ at all points, where the dot represents the dot-product of vectors.

(d) Show that $\left|\vec{E}_u\right| = \left|\vec{E}_v\right|$ at all points, where the absolute values shown are vector magnitudes in the usual sense, that is,

$$\left|\vec{E}_u\right| = \sqrt{\vec{E}_u \cdot \vec{E}_u}.$$

and similarly for $\left|\vec{E}_v\right|$.

(e) Show that the equipotential curves of u and v are the integral curves (field lines) of $\vec{E}_v$ and $\vec{E}_u$ respectively.

4. Consider the complex function $w(z) = \sqrt{z}$ where $z = x + \imath y$ and $w = u + \imath v$.

   (a) Write the real part $u(x, y)$ using polar coordinates and show that it is the potential of the electrostatic problem of a semi-infinite grounded metal plate whose cross section takes up the negative real semi-axis. Note: the conducting plate being grounded means that it is at zero electrical potential.

   (b) For angles $\theta \neq \pm\pi$ determine the behavior of the potential for $\rho \to \infty$.

   (c) Calculate the electric-field vector $\vec{E} = -\vec{\nabla} u$. Write the result in terms of the variables ρ and θ.

   Answer:

$$\vec{E} = -\frac{\cos(\theta/2)}{2\sqrt{\rho}}\,\hat{x} - \frac{\sin(\theta/2)}{2\sqrt{\rho}}\,\hat{y}.$$

   (d) Calculate the limits of $\vec{E}$ for ρ constant and $\theta \to \pm\pi$.

   (e) After calculating these limits, determine the limits of the results obtained when we make $\rho \to \infty$.

   (f) Calculate the surface charge density σ on the plate, using the fact that

$$E_n = \frac{\sigma}{2\varepsilon_0},$$

   where E_n is the component of the electric field that is normal to the plate, pointing away from it. Determine at which point $\sigma(\rho)$ has a singular behavior.

   (g) Since we are, in fact, looking at a two-dimensional section of a three-dimensional problem, consider that the negative real semi-axis is a slice of unit width of the infinite half-plane. Calculate the total electric charge within this slice between the origin $\rho = 0$ and a particular value ρ_0 of ρ.

   Answer: $-2\varepsilon_0\sqrt{\rho_0}$.

5. **(Challenge Problem)** Consider a capacitor formed by two semi-infinite plates. Each plate has an edge with the form of an infinite straight line and the two plates are placed forming a wedge with

an angle $\theta_0 \ll \pi/4$, with the two edges placed together but electrically isolated from one another. Determine the capacitance per unit area for a unit-width slice of the plates located at distances between $r_1 > 0$ and $r_2 > r_1$ from the edges. Do this by means of a conformal transformation in the complex plane, that maps an infinite capacitor on this wedge capacitor, except for the point corresponding to the edge of the plates.

Answer:

$$\frac{\varepsilon_0}{\theta_0} \frac{\ln(r_2) - \ln(r_1)}{r_2 - r_1}.$$

The solution to this problem was found by Prof. Henrique Fleming. It will be addressed again later in this series of books, through the use of other techniques.

Chapter 7

Complex Calculus I: Differentiation

From now on we will work to establish the differential and integral calculus that can be developed for the analytic functions. In this chapter we will begin this program by the definition of the complex derivative, which can be understood as an extension of the real derivative. First of all, all limits involved for the complex variables are reduced to the corresponding limits for the real and imaginary parts, so that the concept of limit itself is the same, the one we know from the integral and differential calculus of real functions.

Since we can calculate variations, or differences Δz and Δw of complex variables, as well as divide by complex numbers, in principle there is no difficulty in writing

$$\frac{dw}{dz} = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z},$$

provided that this limit exists, thus defining the *complex derivative* of the analytic function given by $w(z) = u(x, y) + \imath v(x, y)$. Note that this is a new concept which, as we will see, is associated with the concept of partial derivatives of the real functions of two variables $u(x, y)$ and $v(x, y)$, but that is not identical to it. The question is centered on the existence of the limit. This is ensured in part by the fact that

$$\frac{dw(z)}{dz} = \frac{du(x, y) + \imath dv(x, y)}{dx + \imath dy}$$

and by the fact that both $u(x, y)$ and $v(x, y)$ are continuous and differentiable real functions, for the variations of which we can write

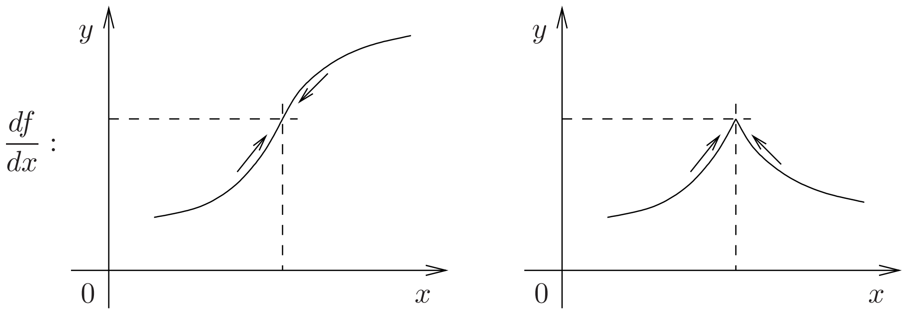


Figure 7.1: Real derivatives of a function $f(x)$ in two cases, a smooth one and one with a kink.

$$\begin{aligned} du &= \frac{\partial u}{\partial x}\,dx + \frac{\partial u}{\partial y}\,dy, \\ dv &= \frac{\partial v}{\partial x}\,dx + \frac{\partial v}{\partial y}\,dy. \end{aligned}$$

There is however the requirement that the limit should not depend on *how* one makes the variations and then takes the limit of vanishing variations. The analog to this for the definition of the derivative of a real function $f(x)$ of a single real variable can be illustrated by the fact that the limits from the left and from the right should exist and be equal in order for the derivative to exist, as shown in the diagram of Figure 7.1.

If both limits exist, but are not equal, then the derivative on the point at issue does not exist, and the function is said to be *non-differentiable* at that point. In our case here, the structure is somewhat more complex since dz is a small *vector* in the complex plane. Instead of just two possibilities of approach to the point where one wants to compute the derivative (from one side or from the other side of an axis), we have now a continuum of possibilities, parametrized for example by the angle α that Δz makes with the real axis x, thus describing all possible directions in which one can vary z, as shown in the diagram of Figure 7.2.

It is necessary to impose that the result of the limit

$$\frac{dw}{dz} = \lim_{\Delta z \to 0} \frac{\Delta w}{\Delta z}$$

be independent of the angle α, that is, of the way in which we take the limit. So let us write dw/dz and check whether or not the condition that $w(z)$ is analytic brings some useful result. We have

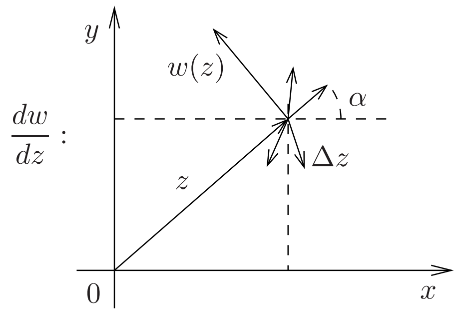


Figure 7.2: The derivative in the complex plane, the variation Δz, the vector field $w(z)$ and the angle α. The relevant question here is whether or not the limit of the ratio of variations depends on α.

$$
\begin{aligned}
\frac{dw}{dz} &= \frac{du + \imath dv}{dx + \imath dy} \\
&= \frac{(du + \imath dv)(dx - \imath dy)}{dx^2 + dy^2} \\
&= \frac{(du\,dx + dv\,dy) + \imath(dv\,dx - du\,dy)}{dx^2 + dy^2}.
\end{aligned}
$$

Using the form of the differentials du and dv, we have for the real part of the numerator in the expression above,

$$du\,dx + dv\,dy = \frac{\partial u}{\partial x}\,dx^2 + \frac{\partial u}{\partial y}\,dx\,dy + \frac{\partial v}{\partial x}\,dx\,dy + \frac{\partial v}{\partial y}\,dy^2.$$

Using now the Cauchy-Riemann conditions, we see that the terms containing the products $dx\,dy$ cancel off, so that this is reduced to

$$du\,dx + dv\,dy = \frac{\partial u}{\partial x}\,(dx^2 + dy^2),$$

in which only a partial derivative of u appears, implying that we have for the real part of the derivative of $w(z)$

$$
\begin{aligned}
\Re\left(\frac{dw}{dz}\right) &= \frac{\partial u}{\partial x}\,\frac{dx^2 + dy^2}{dx^2 + dy^2} \\
&= \frac{\partial u}{\partial x}.
\end{aligned}
$$

For the imaginary part of the numerator we have, similarly,

$$dv\,dx - du\,dy = \frac{\partial v}{\partial x}\,dx^2 + \frac{\partial v}{\partial y}\,dx\,dy - \frac{\partial u}{\partial x}\,dx\,dy - \frac{\partial u}{\partial y}\,dy^2,$$

that by the Cauchy-Riemann conditions is reduced to

$$dv\,dx - du\,dy = \frac{\partial v}{\partial x}\,(dx^2 + dy^2),$$

so that we have for the imaginary part of the derivative of $w(z)$

$$\begin{aligned} \Im\left(\frac{dw}{dz}\right) &= \frac{\partial v}{\partial x}\,\frac{dx^2 + dy^2}{dx^2 + dy^2} \\ &= \frac{\partial v}{\partial x}. \end{aligned}$$

That is, we have an extremely simple result for the complex derivative,

$$\frac{dw(z)}{dz} = \frac{\partial u(x,y)}{\partial x} + \boldsymbol{\imath}\,\frac{\partial v(x,y)}{\partial x},$$

in which only derivatives with respect to x appear, because those with respect to y are automatically determined by the Cauchy-Riemann conditions. Due to the validity of these conditions, this result could also be written in terms only of the two other partial derivatives of u and v, those with respect to y, or as some other combination, as for example in terms of partial derivatives of u,

$$\frac{dw(z)}{dz} = \frac{\partial u(x,y)}{\partial x} - \boldsymbol{\imath}\,\frac{\partial u(x,y)}{\partial y}.$$

Note that, as we saw in previous chapters (Chapters 5 and 6), where we examined conformal mapping defined by analytic functions, the gradients of u and v always have the same absolute value, at all points where the function w is analytic. On the other hand, if the gradient of u is zero, and only in this case, then the two partial derivatives that appear above vanish, which implies that the complex derivative of the function also vanishes. In fact, in this case the gradient v is also zero and therefore *all* four partial derivatives vanish. As we saw before, the angle-preserving property of conformal mappings is determined by the condition that the two gradients be non-vanishing. We now see that this new condition is identical to the analytic invertibility condition of the complex function $w(z)$ at a certain point, namely the condition that its derivative be non-zero at that point, in close analogy with the case of the real functions,

$$\begin{aligned}
|\vec{\nabla} u| &= 0 \iff \\
|\vec{\nabla} v| &= 0 \iff \\
\frac{dw(z)}{dz} &= 0.
\end{aligned}$$

Examining the first result that we have obtained above for the complex derivative one can already conclude that it cannot depend on the direction of the variation of z, since we can write the complex derivative only in terms of the partial derivatives with respect to x. But we can easily examine this question in another way. Using again the Cauchy-Riemann conditions we can show that these results for the derivative of an analytic function have an invariant geometric meaning, if we interpret the quantities involved in terms of two-dimensional vectors. Since both the real part and the imaginary part of the derivative may be written in terms of two different partial derivatives of u and v, we can construct certain linear combinations of these partial derivatives,

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \frac{1}{2}\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right), \\
\frac{\partial v}{\partial x} &= \frac{1}{2}\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right),
\end{aligned}$$

thereby obtaining vector expressions for these quantities

$$\begin{aligned}
\frac{\partial u}{\partial x} &= \frac{1}{2}\vec{\nabla} \cdot (u, v), \\
\frac{\partial v}{\partial x} &= \frac{1}{2}\vec{\nabla} \times (u, v).
\end{aligned}$$

Here we have the *divergence* and the *curl* of the vector field $\vec{w} = (u, v)$, where in the second relation we are exerting a certain license with respect to the usual notation, which is permissible because in this case the two-dimensional curl vector has a single component, which is always normal to the complex plane. We can therefore write the real and imaginary parts of the complex derivative of $w(z)$ in terms of these vector operators,

$$\begin{aligned}
\Re\left(\frac{dw}{dz}\right) &= \frac{1}{2}\vec{\nabla} \cdot \vec{w}, \\
\Im\left(\frac{dw}{dz}\right) &= \frac{1}{2}\vec{\nabla} \times \vec{w}.
\end{aligned}$$

These are results with an invariant geometric character because the divergence is always a scalar and, in two dimensions, the curl also behaves as a scalar or, more precisely, as a pseudo-scalar, that changes sign when we reverse the direction of a coordinate axis (which we will never do). We have therefore for the derivative of an analytic function,

$$\frac{dw(z)}{dz} = \frac{1}{2}\left(\vec{\nabla}\cdot\vec{w} + \boldsymbol{\imath}\vec{\nabla}\times\vec{w}\right).$$

It is interesting to observe that, since this result has a geometrically invariant meaning, it can depend on position, but cannot in fact depend on how we make the variation to define the derivative, that is, it cannot depend on the angle α, because given a differentiable vector field $\vec{w}$ both its divergence and its curl are uniquely determined, regardless of α. We can however make this fact more explicit, writing for $dz = dx + \boldsymbol{\imath}dy$

$$dz = d\ell[\cos(\alpha) + \boldsymbol{\imath}\sin(\alpha)],$$

where $d\ell$ is the absolute value of dz, and α is the angle between the vector dz and the axis x, that is,

$$\begin{aligned} dx &= d\ell\cos(\alpha), \\ dy &= d\ell\sin(\alpha), \\ d\ell^2 &= dx^2 + dy^2. \end{aligned}$$

Retracing now the calculation of the derivative of the analytic function $w(z)$ in terms of $d\ell$ and of α, we have

$$\begin{aligned} \frac{dw}{dz} &= \frac{du + \boldsymbol{\imath}dv}{d\ell[\cos(\alpha) + \boldsymbol{\imath}\sin(\alpha)]} \\ &= \frac{(du + \boldsymbol{\imath}dv)[\cos(\alpha) - \boldsymbol{\imath}\sin(\alpha)]}{d\ell} \\ &= \frac{[du\cos(\alpha) + dv\sin(\alpha)] + \boldsymbol{\imath}[dv\cos(\alpha) - du\sin(\alpha)]}{d\ell}. \end{aligned}$$

Using now the form of the differentials of $u(x, y)$ and $v(x, y)$ in terms of $d\ell$ and α,

$$\begin{aligned} du &= \left[\frac{\partial u}{\partial x}\cos(\alpha) + \frac{\partial u}{\partial y}\sin(\alpha)\right] d\ell, \\ dv &= \left[\frac{\partial v}{\partial x}\cos(\alpha) + \frac{\partial v}{\partial y}\sin(\alpha)\right] d\ell, \end{aligned}$$

we get, for the real part of the derivative of w,

$$\begin{aligned}\frac{du\cos(\alpha)+dv\sin(\alpha)}{d\ell} &= \frac{\partial u}{\partial x}\cos^2(\alpha)+\frac{\partial u}{\partial y}\sin(\alpha)\cos(\alpha)+\\ &\quad +\frac{\partial v}{\partial x}\sin(\alpha)\cos(\alpha)+\frac{\partial v}{\partial y}\sin^2(\alpha).\end{aligned}$$

Again, the Cauchy-Riemann conditions imply that the terms containing the product $\sin(\alpha)\cos(\alpha)$ cancel off, just as before, and we get

$$\begin{aligned}\frac{du\cos(\alpha)+dv\sin(\alpha)}{d\ell} &= \frac{\partial u}{\partial x}\left[\cos^2(\alpha)+\sin^2(\alpha)\right]\\ &= \frac{\partial u}{\partial x},\end{aligned}$$

so that we now see, explicitly, that the dependence on α vanishes. The corresponding calculation for the imaginary part is similar, and we obtain once again

$$\begin{aligned}\Re\left(\frac{dw}{dz}\right) &= \frac{\partial u}{\partial x},\\ \Im\left(\frac{dw}{dz}\right) &= \frac{\partial v}{\partial x},\end{aligned}$$

that is, we are once more led to the invariant result

$$\frac{dw(z)}{dz}=\frac{1}{2}\left(\vec{\nabla}\cdot\vec{w}+\boldsymbol{\imath}\vec{\nabla}\times\vec{w}\right).$$

Thus we see that the analyticity of the function $w(z)$ implies the existence of its complex derivative. On the other hand, if $w(z)$ is not analytic, then the Cauchy-Riemann conditions are not valid, the above results become in fact dependent on the angle α, and therefore the complex derivative does not exist. For an analytic function $w(z)$, the complex derivative dw/dz has a unique meaning, regardless of the direction in which one takes the variation dz. When we change the direction of Δz, automatically Δw changes so as to maintain the ratio dw/dz constant, that is, Δw changes to maintain constant the limit

$$\lim_{\Delta z\to 0}\frac{\Delta w}{\Delta z},$$

as illustrated in Figure 7.3.

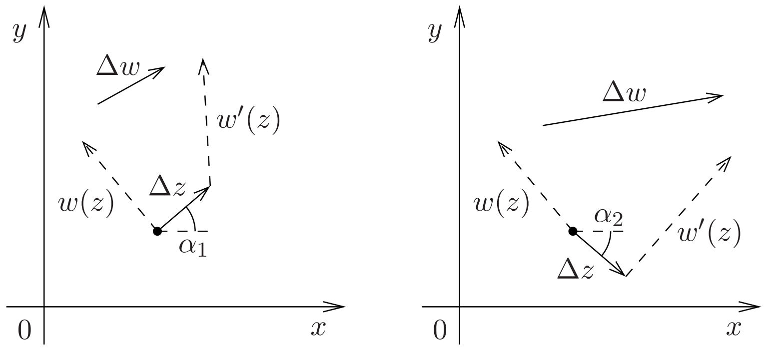


Figure 7.3: Two complex planes with $w(z)$, each one with dz taken in a different direction, illustrating the fact that the limit of the ratio of the two variations must be the same, although both Δw and Δz change. Note that this is merely illustrative and that it is *not* drawn to scale.

Since the algebraic operations are the same in $\mathbb{R}$ and in $\mathbb{C}$, as well as the concept of limit, it is also apparent that the rules of operation and manipulation of complex derivatives are the same as those that we use in the case of real functions, but in many cases it is not difficult to show this directly, for example by calculating the derivatives of the integer powers of z. To begin with, of course we have

$$\begin{aligned} w_0(z) &= z^0 \\ &= 1 \quad \Rightarrow \\ \frac{dw_0}{dz} &= 0, \end{aligned}$$

and

$$\begin{aligned} w_1(z) &= z^1 \\ &= z \quad \Rightarrow \\ \frac{dw_1}{dz} &= \frac{dz}{dz} \\ &= 1\,z^0 \\ &= 1. \end{aligned}$$

We can now proceed by finite induction to prove the result for the general case $w_n(z) = z^n$, assuming it to be true in the case $n - 1$ and showing that this implies its validity in the case n. We assume therefore that

$$\begin{aligned} w_{n-1}(z) &= z^{n-1} \quad \Rightarrow \\ \frac{dw_{n-1}}{dz} &= (n-1)z^{n-2}, \end{aligned}$$

and we calculate the derivative of $w_n = z^n$, using the Leibniz rule for the derivative of a product, since $w_n = z\, w_{n-1}$. Thus we have

$$\begin{aligned} \frac{dw_n}{dz} &= \frac{d(z\, w_{n-1})}{dz} \\ &= w_{n-1} + z\, \frac{dw_{n-1}}{dz} \\ &= w_{n-1} + z(n-1)w_{n-2} \\ &= z^{n-1} + z(n-1)z^{n-2} \\ &= z^{n-1} + (n-1)z^{n-1} \\ &= n\, z^{n-1}, \end{aligned}$$

which proves the results for $w_n(z)$ and therefore for the general case,

$$\frac{dz^n}{dz} = nz^{n-1}.$$

Having this result, we can immediately generalize it to polynomials and to power series, that is, essentially to all analytic functions. Moreover, it is not difficult to repeat this inductive argument for negative powers, away from the point $z = 0$, starting from the formula for the function $w(z) = 1/z$.

However, there is another way to obtain the derivatives, which is often faster and more powerful. The idea is that, having established that the limit that defines the derivative does not depend on the direction in which the variation dz is taken, we are free to choose a particular direction, which can simplify things by reducing the problem quickly to a real derivation problem which is already well known. For example, for the exponential function, which can be written as

$$\begin{aligned} w(z) &= \mathrm{e}^{z} \\ &= \mathrm{e}^{x}[\cos(y) + \imath \sin(y)], \end{aligned}$$

we can put dz in the direction of the axis x, making $dz = dx$, so that only the exponential in x in the expression of $w(z)$ varies when we vary the position in the plane (x, y). In this way, we have that

$$\frac{dw(z)}{dz} = \frac{d(\,\mathrm{e}^{x})}{dx}\,[\cos(y) + \imath \sin(y)].$$

The derivative of the real exponential function is well known, so that we have

$$\begin{aligned}\frac{dw(z)}{dz} &= \mathrm{e}^{x}[\cos(y) + \imath\sin(y)] \\ &= \mathrm{e}^{z},\end{aligned}$$

thus obtaining the corresponding result for the complex function,

$$\frac{dw(z)}{dz} = \mathrm{e}^{z}.$$

Note that this shows that we can also define ordinary differential equations for the case of complex variables, such as the one suggested by the above result,

$$\frac{dw(z)}{dz} = w(z),$$

for which the analytic function $w(z) = \exp(z)$ is the only solution that satisfies the auxiliary condition $w(0) = 1$. As it turns out, the differential equation satisfied by the complex function is an immediate generalization of the real differential equation satisfied by the real function that it generalizes.

This technique of calculation of complex derivatives can be used without too much trouble for all elementary functions. In the case of the function $w(z) = 1/z$, just write it in terms of the polar representation,

$$\begin{aligned}w(z) &= \frac{1}{z} \\ &= \frac{1}{\rho}\,\mathrm{e}^{-\imath\theta},\end{aligned}$$

where the point $z = 0$ is of course excepted. Given a point away from the origin, it now suffices to make the variation of z in the radial direction from this point, by selecting

$$dz = d\rho\,\mathrm{e}^{\imath\theta},$$

with θ kept constant. In this case we have for the derivative of $w(z)$

$$\begin{aligned}\frac{dw(z)}{dz} &= \frac{d(\rho^{-1})}{d\rho\,\mathrm{e}^{\imath\theta}}\,\mathrm{e}^{-\imath\theta} \\ &= \frac{d(\rho^{-1})}{d\rho}\,\mathrm{e}^{-2\imath\theta}.\end{aligned}$$

Again the derivative of the real function is well known, and therefore we have

$$\frac{dw(z)}{dz} = \frac{-1}{\rho^2}\,\mathrm{e}^{-2\imath\theta},$$

so that the result in this case is

$$\frac{dw(z)}{dz} = \frac{-1}{z^2}.$$

A calculation such as this one can be used with equal ease for all integers powers, positive or negative. In all of them we simply write the function in the polar representation and choose dz in the radial direction. This can be done also for fractional powers, as in the case of square root, $w(z) = \sqrt{z}$, which can be written as

$$\begin{aligned} w(z) &= \sqrt{z} \\ &= \sqrt{\rho}\,\mathrm{e}^{\imath\theta/2}, \end{aligned}$$

so that we have for the derivative

$$\begin{aligned} \frac{dw(z)}{dz} &= \frac{d(\sqrt{\rho})}{d\rho\,\mathrm{e}^{\imath\theta}}\,\mathrm{e}^{\imath\theta/2} \\ &= \frac{d(\sqrt{\rho})}{d\rho}\,\mathrm{e}^{-\imath\theta/2} \\ &= \frac{1}{2\sqrt{\rho}}\,\mathrm{e}^{-\imath\theta/2}, \end{aligned}$$

and the final result is, as expected,

$$\frac{dw(z)}{dz} = \frac{1}{2\sqrt{z}}.$$

This can be generalized to any rational powers, and to more complex cases, as for example to the case of the logarithm function, in which the polar representation also solves the problem.

It is important to give here a counterexample, a case in which the derivative of the complex function does not exist, due to the fact that the function is not analytic. A very simple example of this is the function

$$\begin{aligned} w(z) &= |z| \\ &= \sqrt{x^2 + y^2} + 0\imath \\ &= \rho + 0\imath. \end{aligned}$$

In this case we have $u = \sqrt{x^2 + y^2} = \rho$ and $v = 0$, so that the function does not satisfy the Cauchy-Riemann conditions and therefore is not analytic. However, the partial derivatives of $u(x, y)$ and $v(x, y)$ do exist outside the origin $z = 0$. If we write the derivative at an arbitrary point (x, y), using a variation written explicitly in a direction α,

$$dz = d\ell[\cos(\alpha) + \imath \sin(\alpha)],$$

where $d\ell = |dz|$, that is

$$\begin{aligned} dx &= d\ell \cos(\alpha), \\ dy &= d\ell \sin(\alpha), \\ d\ell^2 &= dx^2 + dy^2, \end{aligned}$$

we have for the derivative

$$\begin{aligned} \frac{dw(z)}{dz} &= \frac{\dfrac{\partial u}{\partial x}\, dx + \dfrac{\partial u}{\partial y}\, dy}{dx + \imath dy} \\ &= \frac{\left(\dfrac{x}{\rho}\, dx + \dfrac{y}{\rho}\, dy\right)(dx - \imath dy)}{dx^2 + dy^2} \\ &= \frac{\left[\dfrac{x}{\rho}\, \cos(\alpha) + \dfrac{y}{\rho}\, \sin(\alpha)\right][\cos(\alpha) - \imath \sin(\alpha)]}{d\ell^2}\, d\ell^2 \\ &= [\cos(\theta)\cos(\alpha) + \sin(\theta)\sin(\alpha)][\cos(\alpha) - \imath \sin(\alpha)], \end{aligned}$$

where we used the facts that $x = \rho \cos(\theta)$ and $y = \rho \sin(\theta)$. We see that the formula for the cosine of the arc $(\theta - \alpha)$ appears, so that we have

$$\frac{dw(z)}{dz} = \cos(\theta - \alpha)\cos(\alpha) - \imath \cos(\theta - \alpha)\sin(\alpha).$$

This clearly depends on α, and therefore it is not possible to define the complex derivative of $w(z) = |z|$. It is interesting to observe that the derivative is zero whenever $\cos(\theta - \alpha) = 0$, that is, for the choice of the direction of variation given by

$$\alpha = \theta + \pi/2.$$

This shows that the function has a conical structure, since it is constant whenever we make variations of the position that are perpendicular to the radius. It follows that the function is constant on circles centered at the origin.

We will finish by showing that the complex derivative of an analytic function is itself analytic. Let $w(z)$ be an analytic function and $w'(z)$ its derivative. We can write $w'(z)$ in terms of its real and imaginary parts,

$$w'(z) = U(x, y) + \imath V(x, y).$$

Assuming that the functions U and V are differentiable, the question is whether they satisfy the Cauchy-Riemann conditions. As we saw before, when we discussed the complex derivative, we have for U and V the relations

$$\begin{aligned} U(x, y) &= \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \\ V(x, y) &= \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}, \end{aligned}$$

so that we can calculate the following partial derivatives,

$$\begin{aligned} \frac{\partial U}{\partial x} &= \frac{\partial^2 v}{\partial x \partial y}, \\ \frac{\partial V}{\partial y} &= \frac{\partial^2 v}{\partial y \partial x}, \end{aligned}$$

which are the same, which shows that the first Cauchy-Riemann condition for U and V is satisfied. Analogously we calculate

$$\begin{aligned} \frac{\partial U}{\partial y} &= \frac{\partial^2 u}{\partial y \partial x}, \\ -\frac{\partial V}{\partial x} &= \frac{\partial^2 u}{\partial x \partial y}, \end{aligned}$$

which shows that the second Cauchy-Riemann condition for U and V is also satisfied. Thus we see that the complex derivative of an analytic function is also an analytic function, provided that the real part u and the imaginary part v of the original function have regular derivatives up to the second order. Later on we will see an even stronger theorem on the analyticity of the derivatives of an analytic function.

As we study analytic functions more deeply, it is possible to acquire an increasingly strong feeling that the analyticity condition is extremely restrictive, to the point in which one is led to ask how such a restricted set of functions can have such a general usefulness in physics and mathematics. While it is true that the analyticity condition is very restrictive indeed, we must recall that the restriction takes place in the space of all possible complex functions, that is, pairs of functions of two variables. The restriction to analytic functions seems much less severe if we consider the numerous real restrictions that a single analytic function can have.

Moreover, we shall see that it is perfectly possible to deal with functions that have singular points, which in reality play a key role in the integration theory of complex functions. And to finalize, add to this the fact that, as we shall see later, through the use of infinite series of analytic functions, almost arbitrary functions can be represented. Thus, despite the analyticity condition in the complex plane being very restrictive indeed, this does no harm to its usefulness in physics and mathematics. Rather, it is quite the opposite, this more rigid structure provides us with powerful tools to deal with the real functions we need in physics.

Problem Set

1. Prove that the Leibniz formula for the derivative of a product applies to the product of two analytic functions. In other words, given two analytic functions $w_1(z)$ and $w_2(z)$, each one of which has, therefore, a well-defined complex derivative, show that the derivative of the product-function $w(z) = w_1(z)w_2(z)$ is given by

$$\frac{dw}{dz}(z) = \frac{dw_1}{dz}(z)\, w_2(z) + w_1(z)\, \frac{dw_2}{dz}(z).$$

2. Starting from the complex derivative

$$\frac{d}{dz} z^{-1} = -\frac{1}{z^2},$$

prove by finite induction the more general formula, for a positive integer n,

$$\frac{d}{dz} z^{-n} = -\frac{n}{z^{n+1}}.$$

3. Starting from the known real derivatives, derive the formulas for the derivatives of the analytic functions that follow.

(a) $w(z) = \cosh(z)$.

(b) $w(z) = \sinh(z)$.

(c) $w(z) = \cos(z)$.

(d) $w(z) = \sin(z)$.

4. Derive the second-order complex differential equations and the auxiliary conditions that are satisfied by the following analytic functions.

   (a) $w(z) = \cosh(z)$.

   (b) $w(z) = \sinh(z)$.

   (c) $w(z) = \cos(z)$.

   (d) $w(z) = \sin(z)$.

5. Show that the chain rule applies to the composition of analytic functions. In other words, show that, if $f(z)$ and $g(z)$ are two analytic functions, and $w(z) = f(g(z))$ is the composite function, then

$$\frac{dw(z)}{dz} = \left[\frac{df(z)}{dz}\right](g)\,\frac{dg(z)}{dz}.$$

6. Consider the following analytic function as a conformal transformation, that is, a transformation that preserves the angles between oriented curves that intersect each other, mapping the (x, y) plane onto the (u, v) plane,

$$\begin{aligned} w(z) &= u(x, y) + \imath v(x, y) \\ &= z + \frac{1}{z}. \end{aligned}$$

   (a) Calculate the electric field vectors associated with u and v, $\vec{E}_u = -\vec{\nabla} u$ and $\vec{E}_v = -\vec{\nabla} v$.

   (b) Show that $\left|\vec{E}_u\right| = 0$ at the points $(-1, 0)$ and $(1, 0)$ of the complex plane (x, y), and that therefore we also have $\left|\vec{E}_v\right| = 0$ at these points.

(c) Show that a radius inside the semicircle of the (x, y) plane with θ very small and $\rho \in [\rho_0, 1]$, where $\rho_0 > 0$ is a small number, is mapped on a curve which is very close to the positive real semi-axis of the (u, v) plane, a curve which is, however, still perpendicular to the real segment $(-2, 2)$ of the (u, v) plane, crossing this segment at a point near $(2, 0)$.

(d) Show that in the $\theta \to 0$ limit the curve tends to the positive real semi-axis $[2, \infty)$, which is obviously *not* perpendicular to the segment $(-2, 2)$ of the (u, v) plane.

(e) Show that at the points $(-1, 0)$ and $(1, 0)$, the complex derivative of the function $w(z)$ vanishes, that is, that

$$\begin{aligned} \frac{dw}{dz}(-1, 0) &= 0, \\ \frac{dw}{dz}(1, 0) &= 0. \end{aligned}$$

Just as an informative note, let us recall that at the points where we have $\left|\vec{E}_u\right| = 0$ the proof that the conformal transformation preserves angles does *not* apply, so that the preservation of angles does not hold for curves that cross at these points. Since the derivative of the function is zero at these points, it is not analytically invertible at them, which means that the inverse function of $w(z)$ has singularities at these points. It can be said that these points are singular points of the conformal transformation, in a sense that is *not* the same as the notion of singularity of the complex function, which is analytic at these points. It is the *inverse* of this analytic function that has singularities at these points.

7. Starting from the known real derivatives, derive the formula for the complex derivative of the analytic function $w(z) = \ln(z)$. Show explicitly that this derivative gives the same result on all the infinitely many leaves of the Riemann surface of the logarithm function, thus mapping them all onto a function that has a Riemann surface with a single leaf, which is simply the complex plane except for one point.

Chapter 8

Complex Calculus II: Integration

Having established the fundamentals of the differential calculus of analytic functions, we will now discuss the integral calculus. Our first step is to decide what we mean by an integral of an analytic function. The immediate generalization of the real integral of a function $f(x)$

$$\int f(x)\,dx$$

to a complex function $w(z)$ is simply something like

$$\int w(z)\,dz.$$

There is in fact no conceptual difficulty in defining this, for all arithmetic operations involved are those of a field such as $\mathbb{C}$. The concept of limit that is involved here is the same as in the real case, and the definition of the integration process is that of Riemann, in which we have an infinite sum of infinitesimal quantities on some domain D,

$$\int_D w(z)\,dz = \lim_{\Delta z \to 0} \sum_D w(z)\Delta z.$$

However, since the Δz's and dz's are now vectors in a two-dimensional space, it is necessary to discuss the meaning and nature of the integration domain. At first glance one might think that this is something equivalent to a real integral in two dimensions, but this is not the case. If we

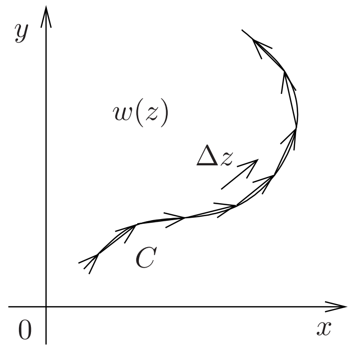


Figure 8.1: The complex plane showing the complex integral as a line integral over a curve C.

concatenate all the dx's of a real integral, we get the real line or a part of it, which is the integration domain. What happens if we do the same with the dz's of a complex integral? Since each dz is a small vector in the plane, we will produce with this concatenation a *curve* in the complex plane, of course. In other words, we see that what we have here is a kind of *line integral*, defining the complex quantity

$$\int_C w(z)\,dz = \lim_{\Delta z \to 0} \sum_C w(z)\Delta z$$

on a curve C, as illustrated in the diagram of Figure 8.1.

With this established, we see that we can reduce this type of complex integration to the real case, just as we could do the same with the complex derivative. In the case of the derivative, it is sufficient to limit the domain to the real line and make variations in the direction of this line, which does not alter the results if the function is analytic, as we have shown before. In our case here, it suffices to choose as the integration curve the real line or some part of it.

Since the integral $\int w(z)\,dz$ is a sum of products of complex numbers, the integral is itself a complex number with a real part and an imaginary part. Therefore, instead of a double real integral what we have here is something like a pair of line integrals on the plane. Let us explicitly write these real and imaginary parts, in terms of the real and imaginary parts of $w(z) = u + \imath v$ and $dz = dx + \imath dy$,

$$\int_C w(z)\,dz \;\; = \;\; \int_C (u + \imath v)(dx + \imath dy)$$

$$\begin{aligned} &= \int_C (u\,dx - v\,dy) + \imath \int_C (v\,dx + u\,dy) \\ &= \int_C (u, -v) \cdot \vec{d\ell} + \imath \int_C (v, u) \cdot \vec{d\ell}, \end{aligned}$$

where we used the vector notation $\vec{d\ell} = dz = (dx, dy)$. Thus we see that in addition to the vector field $\vec{w} = (u, v)$ that represents $w(z)$, two other vector fields related to it arise in our arguments, the field

$$\vec{w}_R = (u, -v),$$

which appears in the real part of the integral, and the field

$$\vec{w}_I = (v, u),$$

which appears in the imaginary part of the integral. Integrals such as this one are known as *complex contour integrals*, in this case on a *closed contour*.

We ask now whether the indefinite integral of a complex function, taken between two points of the plane, one fixed and one variable, may itself be understood as a complex function of the variable point, expressed as follows,

$$F(z) = \int_{z_0}^{z} w\left(z'\right)\, dz'.$$

Integrals such as this one are known as *open-contour integrals*. At the moment we cannot be sure, due to the possibility that the result may depend on the curve that is chosen to go from one point to the other. We can explore this issue by examining a few examples. For example, consider the integral of an analytic function such as

$$\int_A^B z^n\, dz,$$

for a non-negative integer n, between the points $A = (0, 0)$ and $B = (1, 1)$, by the two different paths shown in Figure 8.2.

Making the integration in the two ways indicated one can easily verify that in this case the result is the same, independently of the path, namely $(1 + \imath)^{n+1}/(n + 1)$. Moreover, considering the integral of a function that is not analytic, such as

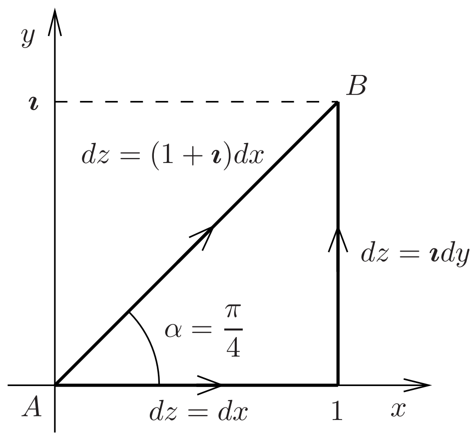


Figure 8.2: The complex plane with two paths for the integral, a straight one between the points A and B and the other with two segments parallel to the coordinate axes.

$$\int_A^B z^*\,dz$$

between points $A = (-1, 0)$ and $B = (1, 0)$, by the two paths consisting of an arc of the unit circle and a straight line, as illustrated in the diagram of Figure 8.3, we find that in this case the two results are different, one being zero and the other not. The first example illustrates a very important theorem, the Cauchy theorem, while the second illustrates its antithesis.

We will now discuss the Cauchy theorem, which is one of the foundations of the complex calculus. It can be obtained from Green's theorem for pairs of real functions of two variables. Green's theorem tells us that, given two functions P and Q, continuous and differentiable (and with continuous derivatives) in a given two-dimensional region S of the plane (x, y), whose boundary is a closed curve C, the following relation holds for them,

$$\oint_{C=\partial S}(P\,dx + Q\,dy) = \int_S \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) dxdy,$$

where we used the notation $C = \partial S$ to indicate that the curve C is the boundary of the surface S. It is very useful to reinterpret this theorem in

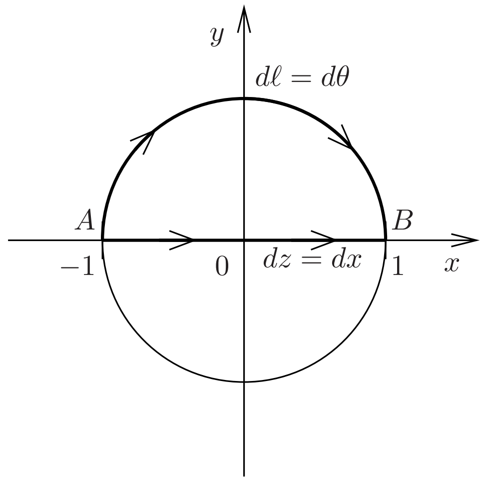


Figure 8.3: The complex plane with the two paths for the integral from A to B. Note that in the case of the arc of circle, while it is true that $d\ell = d\theta$, we have that $dz = \imath z d\theta$.

vector language. We can interpret the pair (P, Q) as the two components of a real vector $\vec{V} = (P, Q)$ in two dimensions, and write this equation in the form

$$\oint_{C=\partial S} \vec{V} \cdot \vec{d\ell} = \int_S \vec{\nabla} \times \vec{V} \cdot \vec{da}.$$

We now have, on the left-hand side, a line integral on C, and on the right-hand side a surface integral on S, whose integrand is the component of the curl of V, which in two dimensions is always in the direction of the versor normal to S. Thus we see that this is a particular case of the Stokes theorem in three dimensions, or the curl theorem, which is used in electromagnetism. We will make an intuitive geometric discussion of these theorems at the end of this chapter.

Since we have here the two real functions of two variables u and v, with exactly these properties, we can use this theorem for the vector $\vec{w}$ which represents a complex function and write, using the same license that we used before in relation to the curl vector notation,

$$\oint_{C=\partial S} \vec{w} \cdot \vec{d\ell} = \int_S \vec{\nabla} \times \vec{w}\, da.$$

The curve C is now a contour in the complex plane, that is, a curve that is the domain of a complex integral, which should be oriented by

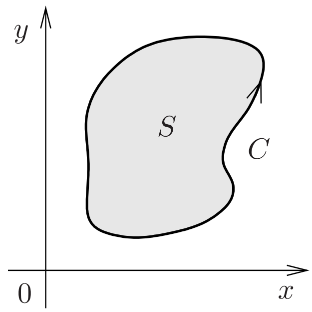


Figure 8.4: The complex plane with the region S and the corresponding closed integration contour $C = \partial S$ properly oriented.

the right-hand rule with respect to the normal to the plane (x, y), which is the normal direction pointing out of the paper, in the third spatial direction, as in the illustration shown in Figure 8.4.

This formula is true, but it is not very interesting, because in fact it does *not* appear in our development of the integral of a complex function. The most interesting thing is, therefore, to consider the two vector fields that appear in the complex integral, as we saw earlier,

$$\begin{aligned}\oint_C w\,dz &= \oint_C (u\,dx - v\,dy) + \imath \oint_C (v\,dx + u\,dy)\\ &= \oint_C (u, -v)\cdot d\vec{\ell} + \imath \oint_C (v, u)\cdot d\vec{\ell}.\end{aligned}$$

If we now use Green's theorem for the fields $\vec{w}_R = (u, -v)$ and $\vec{w}_I = (v, u)$, we get a very interesting and even surprising result. For the real part, for example, we have

$$\begin{aligned}\oint_{C=\partial S} \vec{w}_R \cdot d\vec{\ell} &= \int_S da\,\vec{\nabla}\times\vec{w}_R\\ &= \int_S da\left(-\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right).\end{aligned}$$

Note now that, if the function $w(z)$ is analytic, then the Cauchy-Riemann conditions tell us that

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

at every point where the function is analytic. If the analyticity region contains the whole region S, including its boundary, then it follows that the right-hand side vanishes identically, so that we have

$$\oint_C \vec{w}_R \cdot d\vec{\ell} = 0,$$

that is, for any closed curve C within which and on which the function is analytic. It is important to emphasize that, for this result to be valid, the function $w(z)$ must be analytic *throughout the interior* of the region S, and not only at its boundary. The same result holds for the case of $\vec{w}_I$,

$$\begin{aligned} \oint_{C=\partial S} \vec{w}_I \cdot d\vec{\ell} &= \int_S da\, \vec{\nabla} \times \vec{w}_I \\ &= \int_S da \left(\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \right). \end{aligned}$$

This time it is the other Cauchy-Riemann relation,

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

which implies that we have

$$\oint_C \vec{w}_I \cdot d\vec{\ell} = 0.$$

It follows, therefore, that the complex integral of an analytic function vanishes for *any* closed curve C,

$$\oint_C w(z) dz = 0,$$

if the function is analytic, that is, has no singularities within the curve and over it. This is the important *Cauchy theorem*, also called the Cauchy-Goursat theorem, to extend the honor to this other mathematician who proved the theorem in a somewhat stronger form, with one less hypothesis, that of the continuity of the partial derivatives.

Let us take a moment to ponder this extraordinary fact: $w(z)$ is *any* analytic function, and S *any* region on which the function is analytic, and *all* these integrals are zero. It may even seem that the theory ends here, for lack of anything that is not zero! However, we must draw attention to the fact that we consider only *closed* curves in this theorem. The integrals are not always zero if the curves are open. This theorem is extremely important and powerful, but it does *not* trivialize the theory.

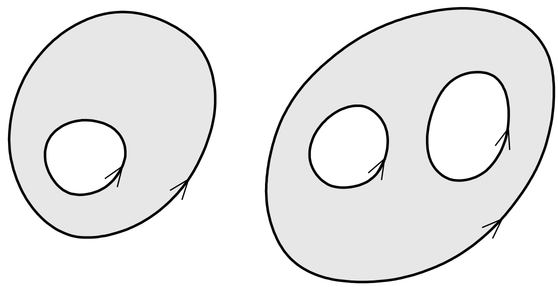


Figure 8.5: The complex plane with regions S with holes, and the corresponding oriented boundaries $C = \partial S$.

Note that the region S need not be simply connected, that is, it could be a region with one or more holes in it, such as illustrated in Figure 8.5. In this case the boundary of the region is a set of curves, instead of a single curve. As a matter of fact S does not even need to be connected, as we may have several separate regions, and what holds for each one also holds for the set, as illustrated in Figure 8.6.

Given the Cauchy-Goursat theorem, it is easy to see that the complex integral of an analytic function between two points does not depend on the path that is chosen for the integration, so long as it is possible to deform each path onto the others without leaving the region where the function is analytic. From this we can conclude that it is indeed possible to define the indefinite integral or primitive of a complex function, as suggested above, if and only if the function under consideration is analytic,

$$F(z) = \int_{z_0}^{z} w\big(z'\big)\, dz'.$$

What the Cauchy-Goursat theorem and the independence of the path tell us is that there is a complex *potential* Φ, that is actually a pair of real potentials Φ_R and Φ_I, one for the real part of the integral and one for the imaginary part, the definition of which is associated with this property of the process of integration of an analytic function. In other words, given an analytic function $w(z)$, represented by the vector field $\vec{w}$, and given the associated vector fields $\vec{w}_R$ and $\vec{w}_I$, there are functions Φ_R and Φ_I such that

$$\begin{aligned} \vec{w}_R &= \vec{\nabla}\Phi_R, \\ \vec{w}_I &= \vec{\nabla}\Phi_I. \end{aligned}$$

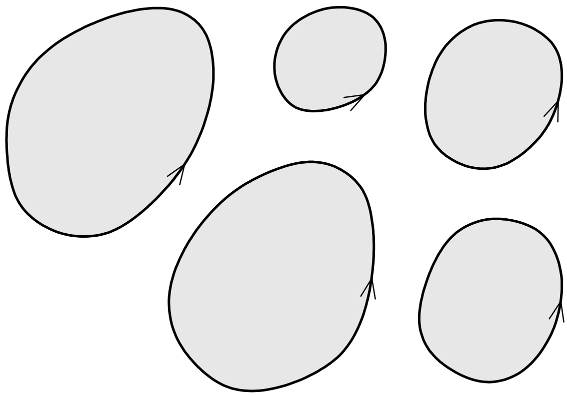


Figure 8.6: The complex plane with several separate regions S and the oriented boundaries $C = \partial S$ of each one of them.

This means that the associated fields $\vec{w}_R$ and $\vec{w}_I$ are the gradients of these functions, that is, they are proportional to the electrostatic fields associated with each one of the two potentials. Thus we see that Φ_R and Φ_I resemble electrostatic potentials, and therefore seem to have similar behavior to that of the functions u and v of an analytic function. In order to elaborate this argument, we begin by noting that the indefinite integral can be written as a variation of these potentials,

$$\begin{aligned} F(z) &= \int_{z_0}^{z} w(z')\, dz' \\ &= (\Phi_R - \Phi_{R,0}, \Phi_I - \Phi_{I,0}) \\ &= U + \imath V, \end{aligned}$$

where $(\Phi_{R,0}, \Phi_{I,0})$ are the values that the potentials have at the starting point z_0. Here we see that the functions $U = (\Phi_R - \Phi_{R,0})$ and $V = (\Phi_I - \Phi_{I,0})$ are the real and imaginary parts of the complex function $F(z)$. Since we know that the vector fields $\vec{w}_R = \vec{\nabla} U$ and $\vec{w}_I = \vec{\nabla} V$ exist, it is clear that these functions are differentiable and therefore continuous. Since we know that $\vec{w}_R = (u, -v)$ and $\vec{w}_I = (v, u)$, we can write the partial derivatives of U and V explicitly in terms of u and v,

$$\begin{aligned} \frac{\partial U}{\partial x} &= u, \\ \frac{\partial U}{\partial y} &= -v, \end{aligned}$$

$$\begin{aligned}\frac{\partial V}{\partial x} &= v,\\ \frac{\partial V}{\partial y} &= u.\end{aligned}$$

From an examination of the first and last equations above, we conclude that

$$\frac{\partial U}{\partial x} = \frac{\partial V}{\partial y},$$

while the other two equations give us

$$\frac{\partial U}{\partial y} = -\frac{\partial V}{\partial x}.$$

With this, the Cauchy-Riemann conditions are established for the functions U and V forming the complex function $F(z)$, which is now manifestly an analytic function. With this it is also established that the functions U and V are not only electrostatic potentials, but potentials in a region without electric charges, which satisfy the homogeneous Laplace equation,

$$\begin{aligned}\nabla^2 U &= 0,\\ \nabla^2 V &= 0.\end{aligned}$$

It is important to note here, for later use, that in this proof that $F(z)$ is analytic, the fact that $w(z)$ is analytic was *not* actually used, but only the fact that all the integrals of $w(z)$ on closed curves contained within a given region are zero.

Thus we see that the primitive of an analytic function is always another analytic function without the need for any additional hypothesis. Since this argument can be iterated, it follows that an analytic function is infinitely integrable, that is, we can integrate it indefinitely, do the same with the primitive that results, and so on infinitely, thus generating an infinite sequence of analytic functions. This sequence reminds us of another sequence, the one formed by the successive derivation of analytic functions, although in that case the proof we presented that the derivative of an analytic function is also analytic depended on the additional hypothesis that the functions $u(x, y)$ and $v(x, y)$ have second-order derivatives. Later on we will see that this hypothesis can be raised.

Anyway, this leads us to think that these two sequences of functions, the one generated by iterated derivation and the one generated by iterated integration, are related, being in reality the same, except for constants added to the functions, provided that they have an element in common. Recalling that, in the proof that the primitive $F(z)$ of an analytic function $w(z)$ is also analytic, we have for $w = u + \boldsymbol{\imath} v$ and $F = U + \boldsymbol{\imath} V$, that

$$\begin{aligned}
\frac{\partial U}{\partial x} &= u, \\
\frac{\partial U}{\partial y} &= -v, \\
\frac{\partial V}{\partial x} &= v, \\
\frac{\partial V}{\partial y} &= u,
\end{aligned}$$

and recalling that the derivative of an analytic function $w(z)$ can be written as

$$\frac{dw(z)}{dz} = \frac{\partial u}{\partial x} + \boldsymbol{\imath}\,\frac{\partial v}{\partial x},$$

it follows that the derivative of the primitive $F(z)$ can be written as

$$\begin{aligned}
\frac{dF(z)}{dz} &= \frac{\partial U}{\partial x} + \boldsymbol{\imath}\,\frac{\partial V}{\partial x}, \\
&= u + \boldsymbol{\imath} v \\
&= w(z).
\end{aligned}$$

That is, just as in the real case, complex integration and complex differentiation are inverse operations to each other, so long as we are in a region of the complex plane where $w(z)$ is analytic, and provided that we do not use two separate curves whose union goes around a singular point of $w(z)$ in order to define the primitive $F(z)$. We can say therefore that the fundamental theorem of the real calculus also holds for our analytic functions,

$$\begin{aligned}
F(z) &= \int^{z} w(z')\,dz' \quad \Longleftrightarrow \\
\frac{dF(z)}{dz} &= w(z).
\end{aligned}$$

Note that the Cauchy-Goursat theorem allows us to define the derivative of the primitive of an analytic function in a simple and geometric form, since we have that, given the definition of the primitive,

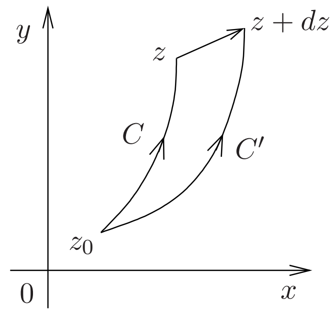


Figure 8.7: The complex plane showing the complex indefinite integral and its variation.

$$F(z) = \int_{z_0}^{z} w(z)\, dz,$$

taking the derivative of $F(z)$ means varying z, which corresponds to a variation of the upper extreme of the integral. This variation involves adding a new segment to the integral, taking us from the endpoint of C to a new point, as illustrated in Figure 8.7.

Since $w(z)$ is analytic, because of the Cauchy-Goursat theorem we can go to the new point by any path from the endpoint of C, without changing the corresponding variation of F. On the other hand, since $F(z)$ is analytic, it should be possible to take this variation in any direction without changing the result for the ratio dF/dz. In specific circumstances it may be advantageous to choose a specific direction for the variation, in order to simplify the calculations.

Since Green's theorem, that is, the two-dimensional version of the Stokes or curl theorem, plays such a central role in our theory here, it is important that we examine it in order to understand the origin of the result. As we shall see, it is an analytical theorem fundamentally based on geometrical ideas. We can construct a draft of the proof of the theorem, considering a generic vector $\vec{w} = (u, v)$, an infinitesimal rectangle (dx, dy) and the two functions u and v as associated with *links*, namely connections between consecutive points or *sites* of a rectangular *lattice* of points. We adopt the convention that u and v are indexed by the endpoints of the links in the negative direction of the coordinate axes, as illustrated in Figure 8.8.

An element of the line integral of $\vec{w} \cdot \vec{d\ell}$ around the rectangle, with

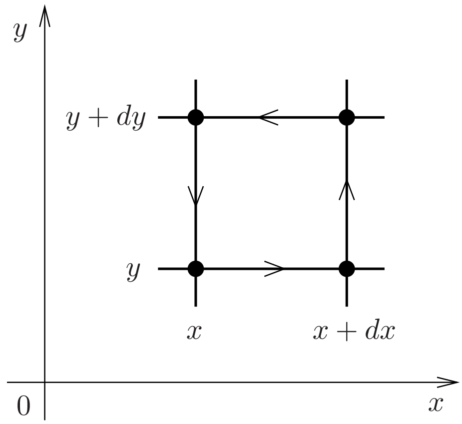


Figure 8.8: The axes (x, y), the oriented rectangle and its elements.

the positive orientation shown, can be written as

$$\begin{aligned}\delta I_C &= u(x,y)\,dx + v(x+dx,y)\,dy - u(x,y+dy)\,dx - v(x,y)\,dy \\ &= dy\,[v(x+dx,y) - v(x,y)] - dx\,[u(x,y+dy) - u(x,y)].\end{aligned}$$

The two differences in the last form of this equation can be represented in terms of the appropriate partial derivatives of u and of v, and of the variations dx and dy of the coordinates,

$$\begin{aligned}\delta I_C &= dydx\,\frac{\partial v(x,y)}{\partial x} - dxdy\,\frac{\partial u(x,y)}{\partial y} \\ &= da\left[\frac{\partial v(x,y)}{\partial x} - \frac{\partial u(x,y)}{\partial y}\right] \\ &= \delta I_S,\end{aligned}$$

where $da = dxdy$ is the area element associated with the surface of the infinitesimal rectangle. This establishes the theorem on the infinitesimal rectangle defined by dx and dy. Now we can build a larger region S by concatenating numerous infinitesimal rectangles like this. Both on the left-hand side and on the right-hand side of the equation, the complete integral over S is the sum of integrals over each one of the infinitesimal rectangles.

It is easy to see that the contributions to the line integral in the left-hand side cancel off in pairs on the links which are internal to the set of rectangles, so that from the point of view of the line integral, what is left is only the integral over the external curve, as illustrated in

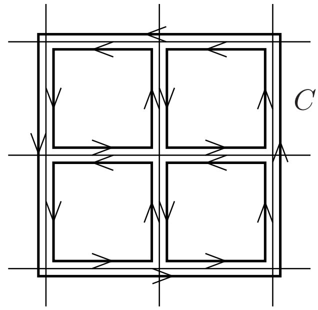


Figure 8.9: Various infinitesimal rectangles oriented and linked together, forming a larger surface and also forming the oriented curve which constitutes its boundary.

Figure 8.9. On the other side of the equation, the surface integrals are added together to result in the integral over the whole area of S, and we have

$$\oint_{C=\partial S} \vec{w} \cdot \vec{d\ell} = \int_S \vec{\nabla} \times \vec{w}\, da,$$

that holds for any region formed by the union of rectangles. In order to build a more complete proof it is also necessary to use infinitesimal triangles, in order to be able to approximate with perfection any given curve C. This is necessary in principle because we cannot approximate the length of the curve C, which is the domain of the line integral, using only the sides of the rectangles. One can check this fact trying to represent the length of the diagonal of a square by a "staircase" with steps parallel to the sides. Of course, for the integral do be correct, it is necessary that the concatenation of the elements on which it is calculated faithfully reproduce the integration domain in the integration limit, with its proper measure.

In order to approximate a curve with the correct measure we must use a polygon whose vertices are all on the curve, which introduces the need for the use of triangles to represent the area within the curve. It is not difficult to show that it is always possible to represent the interior of a simple closed curve by a set of rectangles and triangles, all smaller than a given size, such that the external polygon formed by all of these rectangles and triangles constitutes a closed polygon with all its vertices on the curve.

It is not necessary to consider the contribution of these additional

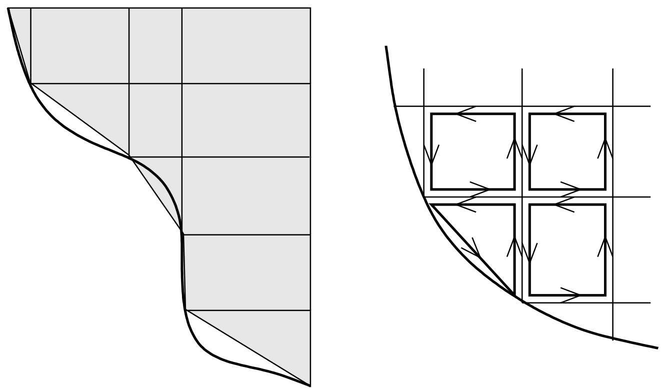


Figure 8.10: An area S composed of inner rectangles and triangles along the boundary, showing in detail the cancellations along the boundary of one of the triangles.

triangles to the area in the case of the right-hand side of the equation, since the total area thereof tends to zero in the limit, as it fits within a perimeter of S with infinitesimal width and finite length, whose area goes to zero in the integration limit. Since the integral is two-dimensional, the total contribution of this sub-domain of measure zero is zero. However, one must consider the contributions of the perimeters of the triangles in the case of the left-hand side of the equation, as illustrated in Figure 8.10.

When we add these triangular contributions, all terms associated with the vertical and horizontal sides of the triangles cancel off in pairs with the corresponding terms associated with the sides of the inner rectangles next to which the triangles are placed. Thus, there remain only the terms that are associated with the slanted sides of the triangles, which run all the way around the curve C. Thus we see that what the inclusion of the triangles does is to replace every two rectangle sides by a segment whose ends are on the curve C.

With this, one can see that the theorem holds for any surfaces S whose boundary curves C can be approximated by sequences of infinitesimal segments, with the ends of the segments on the curves. Theorems like this can also be proved in higher dimensions. An example of this is the Gauss or divergence theorem, which is very useful in electromagnetism, as is also the curl theorem in three dimensions. The simplest example, in a single dimension, reduces to the familiar fundamental theorem of the calculus. The most general case is known as a the general Stokes theorem. Note that the proof of the theorem we discussed here could be extended even to curved surfaces immersed in three-dimensional space, without any important conceptual change in the argument presented.

Problem Set

1. For each function $w(z)$ below, make an analysis of the possible singularities in order to determine whether or not one can use the Cauchy-Goursat theorem in order to determine that, on the unit circle C,

$$\oint_C w(z)\,dz = 0.$$

   (a) $w(z) = \dfrac{z^2}{z-3}$.

   (b) $w(z) = z\,\mathrm{e}^z$.

   (c) $w(z) = \dfrac{1}{z^2 + 2z + 2}$.

   (d) $w(z) = \mathrm{sech}(z)$.

   (e) $w(z) = \tan(z)$.

   (f) $w(z) = \ln(z+2)$.

2. Consider an arbitrary closed simple curve C on the (x, y) plane, with the condition that it do not pass through the point $z = 0$. Show that

$$\oint_C \frac{dz}{z^2} = 0,$$

   in the following cases.

   (a) If the interior of the curve does not contain the origin $z = 0$.

   (b) If the interior of the curve contains the origin $z = 0$.

   Note: a closed simple curve is a closed curve that does not intersect itself.

3. Consider an arbitrary closed simple curve C on the (x, y) plane, with the condition that it do not pass through the point $z = 0$. Show that

$$\oint_C \frac{dz}{z^3} = 0,$$

in the following cases.

   (a) If the interior of the curve does not contain the origin $z = 0$.
   (b) If the interior of the curve contains the origin $z = 0$.

   Note: a closed simple curve is a closed curve that does not intersect itself.

4. Consider an arbitrary closed simple curve C on the (x, y) plane, with the condition that it do not pass through the point $z = 1$. Show that

$$\oint_C \frac{dz}{(z-1)^2} = 0,$$

in the following cases.

   (a) If the interior of the curve does not contain the point $z = 1$.
   (b) If the interior of the curve contains the point $z = 1$.

   Note: a closed simple curve is a closed curve that does not intersect itself.

5. Consider the analytic function $w(z) = \sqrt{z}$.

   (a) Calculate the integral

$$\oint_C w(z)\, dz,$$

   on the unit circle C.

   (b) Calculate the integral of this function over a curve that goes twice around the unit circle.

(c) Examine the situation for the integrals over other circular contours centered at the origin, in order to determine whether it is possible to say something about the values of these integrals based on the results obtained here.

Hint: use polar coordinates.

6. Consider the analytic function $w(z) = 1/\sqrt{z}$.

(a) Calculate the integral

$$\oint_C w(z)\, dz,$$

on the unit circle C.

(b) Calculate the integral of this function over a curve that goes twice around the unit circle.

(c) Examine the situation for the integrals over other circular contours centered at the origin, in order to determine whether it is possible to say something about the values of these integrals based on the results obtained here.

Hint: use polar coordinates.

7. Consider the analytic function $w(z) = \ln(z)$.

(a) Calculate the integrals

$$\oint_C w(z)\, dz,$$

on the unit circle C.

(b) Calculate the integral of this function over a curve that goes twice around the unit circle.

(c) Examine the situation for the integrals over other circular contours centered at the origin, in order to determine whether it is possible to say something about the values of these integrals based on the results obtained here.

Hint: use polar coordinates.

8. Consider the proof of Green's theorem in two dimensions that was presented in the text. Consider therefore an infinitesimal rectangle of dimensions dx and dy, a vector field (u, v) with Cartesian components $u(x, y)$ and $v(x, y)$, the line integral of $(u\,dx + v\,dy)$ over the perimeter of the rectangle, and the surface integral

$$\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy$$

over the area of the rectangle. In the case of the line integral, consider that each value of u and v is associated with a *link*, like those which form the perimeter of the rectangle, and represent these values as functions of the midpoint of each link. For example, in the first part of the integral we have $u(x + dx/2, y)$, and so on.

   (a) Show the central fact of the theorem, that is, that in this infinitesimal rectangle, with the perimeter positively oriented,

$$\sum_{\text{links}}^{(4)} (u\,dx + v\,dy) = \left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) dx\,dy.$$

   (b) Determine to which point one should associate each partial derivative in the most natural and symmetric way possible, and show that the expression

$$\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right)$$

   is naturally associated with the center of the rectangle, which is also called, in some circumstances, a *plaquette*.

9. Consider the integral of an analytic function given by

$$I = \int_A^B z^n\,dz,$$

for n a non-negative integer, between the points $A = (0, 0)$ and $B = (1, 1)$.

   (a) Calculate the integral along the path formed by a straight segment which links the point A directly to the point B.

(b) Calculate the integral along the path formed by a straight segment from A to $(1,0)$ and another straight segment from $(1,0)$ to B.

(c) Calculate the integral along the path formed by a straight segment from A to $(0,1)$ and another straight segment from $(0,1)$ to B.

10. Consider the integral of a complex function, which is not analytic, given by

$$I = \int_A^B z^* \, dz,$$

between the points $A = (-1,0)$ and $B = (1,0)$.

(a) Calculate the integral along the path formed by a straight segment which links the point A directly to the point B.

(b) Calculate the integral along the arc of the unit circle on the upper half-plane connecting the points A and B.

(c) Calculate the integral along the arc of the unit circle on the lower half-plane connecting the points A and B.

11. Consider the unit-side square and its diagonal going from the point $(0,0)$ to the point $(1,1)$ in a Cartesian coordinate system (x,y). Consider approximating this diagonal by a "staircase" that consists of horizontal and vertical segments, all the steps being equal, going from the point $(0,0)$ to the point $(1,1)$. Consider the limit in which the size of the steps goes to zero.

(a) What is the limit of the distance between any point on the "staircase" and the diagonal of the square, when the size of the steps goes to zero? The "staircase" can be considered as a good representation of the diagonal in this limit, in terms of the points of the plane that make up each object?

(b) What is the limit for the total length of the "staircase", when the size of the steps goes to zero? The "staircase" can be considered as a good representation of the diagonal in this limit, if we now think in terms of the *measure* associated with each object?

12. **(Challenge Problem)** Consider a closed simple curve C in the plane, the interior of which is the surface S. For simplicity, assume that the curve is differentiable. A certain maximum size is given, represented by the strictly positive real number ϵ, which can be chosen as small as needed.

    (a) Construct a set of rectangles and triangles, all with vertical and horizontal dimensions smaller than ϵ, which is almost entirely contained within the interior of the curve, in the sense that its boundary is a polygon with all its vertices located on the curve.

    (b) Given a continuous and differentiable vector field $\vec{w} = (u, v)$, construct a discrete representation of the surface integral

    $$\int_S \vec{\nabla} \times \vec{w}\, da,$$

    using this set of rectangles and triangles. Show that in the limit $\epsilon \to 0$ this discrete representation approaches the integral.

    (c) Given the same vector field $\vec{w}$, construct a discrete representation of the line integral

    $$\oint_C \vec{w} \cdot d\vec{\ell},$$

    using the polygon which is the external boundary of the set of rectangles and triangles. Show that in the limit $\epsilon \to 0$ this discrete representation approaches the integral.

    (d) Use these constructions to elaborate a more complete proof of Green's theorem, that is, in our vector language, of the result

    $$\oint_C \vec{w} \cdot d\vec{\ell} = \int_S \vec{\nabla} \times \vec{w}\, da.$$

    Consider in detail the cancellations of integration elements, in order to transform the surface integral into a line integral.

Chapter 9

Complex Derivatives and Integrals

Having studied before, and separately, the differential calculus and the integral calculus of analytic functions, let us now examine in greater detail the important relations between these two structures. Let us begin by reviewing the basic results we have so far. In general, the complex differential calculus was shown to be operationally identical to the differential calculus of real functions. In the case of the integral calculus, some powerful results give it a whole new personality. In the previous chapter (Chapter 8) we saw the Cauchy-Goursat theorem, which tells us that if a function $w(z)$ is analytic along the interior of a closed curve C, and also on the curve, then

$$\oint_C w(z)\, dz = 0,$$

which holds for all closed curves C which do not contain singularities of the function $w(z)$. In addition to that we have seen that, given this theorem, we can define the indefinite integral of an analytic function, that is, its primitive,

$$F(z) = \int_{z_0}^{z} w(z')\ dz',$$

on an open curve C going from z_0 to z, so long as w is analytic in the region where C is located. What the theorem tells us is that there is a pair of potentials Φ_R and Φ_I, associated with the real and imaginary parts of $F(z)$, whose gradients are $\vec{w}_R = \vec{\nabla}\Phi_R$ and $\vec{w}_I = \vec{\nabla}\Phi_I$, such that

$$w(z)\,dz = \left(\vec{w}_R \cdot \vec{d\ell}, \vec{w}_I \cdot \vec{d\ell}\right),$$

so that the indefinite integral can be written as

$$F(z) = (\Phi_R - \Phi_{R,0}) + \imath(\Phi_I - \Phi_{I,0}),$$

where $(\Phi_{R,0}, \Phi_{I,0})$ are the values that the potentials have at the starting point z_0, that is, we have $F(z) = U + \imath V$, with the values of $U(x, y)$ and $V(x, y)$ given above. Studying these two functions we have also shown that $F(z)$ is analytic when and where $w(z)$ is.

Turning back now to the Cauchy-Goursat theorem, note that if the function $w(z)$ has a singularity in the interior region to the curve C, then it is no longer necessary for the complex line integral of the function $w(z)$ along C to be zero, although this may happen. As an example of an integral that vanishes due to the analyticity of the integrand, we can calculate

$$I = \oint_C z^n dz,$$

with n a positive integer, for example on circles C centered at the origin. In order to explicitly implement the integration it suffices to write $z^n = \rho^n \exp(n\imath\theta)$, and to verify that an integration element dz tangent to the circle in the positive direction, and therefore in the direction of increasing θ, with constant ρ, can be written as $dz = \imath\rho\exp(\imath\theta)\,d\theta$, so that we have

$$\begin{aligned} I &= \int_0^{2\pi} \rho^n\, \mathrm{e}^{n\imath\theta} \imath\rho\, \mathrm{e}^{\imath\theta} d\theta \\ &= \imath\rho^{n+1} \int_0^{2\pi} \mathrm{e}^{(n+1)\imath\theta} d\theta \\ &= \imath\rho^{n+1} \frac{1}{(n+1)\imath} \left[\mathrm{e}^{(n+1)2\pi\imath} - \mathrm{e}^0\right], \end{aligned}$$

that vanishes due to the periodicity of the exponential. Note that due to the factor $(n+1)$ in the denominator, this result does not hold in the case $n = -1$. Indeed, as a particularly important example of an integral that is *not* zero, due to a singularity, we may calculate

$$\oint_C \frac{dz}{z},$$

on circles C centered at the origin. Using again the same ideas that were used in the previous integral, we have

$$\begin{aligned} I &= \int_0^{2\pi} \frac{1}{\rho}\,\mathrm{e}^{-\imath\theta}\imath\rho\,\mathrm{e}^{\imath\theta}d\theta \\ &= \imath\int_0^{2\pi} d\theta \\ &= 2\pi\imath. \end{aligned}$$

We see that in this case the result is not zero. Examining the arguments above, it is not hard to see that the result would be zero for all values of n, positive or negative, except for the value -1, which makes it a very special case. The cases in which $n < -1$ are examples of integrals that vanish *despite* the non-analyticity of the integrand. A geometrical intuition of the reason for this might be achieved considering explicitly the vector fields $\vec{w}_R$ and $\vec{w}_I$ of the real and imaginary parts of the integral for functions $w(z) = z^n$, for various positive and negative values of n.

Doing this it is not difficult to see that in the case $n = -1$, and only in this case, the imaginary part $\vec{w}_I$ of the integral is tangent to the circles centered at the origin, thus giving a non-zero value for the line integral on these circles. However, the real part $\vec{w}_R$ is perpendicular to the circles and therefore results in zero for the line integral, which is why the non-zero result is purely imaginary. For all other values of n these vectors revolve around the tangent to the circle as we move along the circle, producing cancellations and resulting in zero line integrals for both the real part and the imaginary part.

Armed with these two results, we can then use the Cauchy-Goursat theorem to generalize any of them to any closed curve around the origin. We can also consider integrals in regions that are not simply connected, such as a ring between two disjoint curves that surround the origin, or even in regions that are not connected at all. In addition to this, through a simple change of variables it is possible to generalize the second result to functions that have a singularity at some other point z_0, rather than at the origin, such as $z' = z + z_0$ and $dz' = dz$, leading to

$$\oint_C \frac{dz}{z - z_0},$$

for example on closed curves C containing the point z_0. Integrals of functions like this, exhibiting what we call *simple poles*, play a central role in the development of the theory, as we shall see. We can systematize these results for a generic power n, not necessarily positive, writing for the integral of $w = z^n$ around the unit circle

$$\oint_C z^n \, dz = 2\pi \imath \delta_{n,-1},$$

where $\delta_{i,j}$ is the discrete delta function, that is, the so-called Kronecker delta, a symbol which equals 1 if the two indices are equal, and 0 if they are different. This result can be interpreted intuitively and geometrically, by plotting the vector fields $\vec{w}_R$ and $\vec{w}_I$ associated with z^n along the unit circle and thus verifying the existence of a kind of resonance between the rotation along C around the origin and the rotation of the vectors associated with z^n around themselves.

Using the ease provided to us by the Cauchy-Goursat theorem in order to relate different curves, considering the change of variables $z' = z + z_0$, $dz' = dz$ that allows us to generalize the result above for the integral,

$$\oint_C \frac{dz}{z} = \oint_{C'} \frac{dz'}{z' - z_0},$$

and since the singular points of the integrands are within the respective curves, we conclude that

$$\oint_C (z - z_0)^n \, dz = 2\pi \imath \delta_{n,-1},$$

on any closed contour C containing the point z_0, the value of the integral being zero for any other closed curve.

We will now show an interesting and fundamental fact about these closed-contour integrals of functions with a singularity of the simple-pole type, which will give us a new insight into the analytic functions and into the complex differential and integral calculus. As we have seen, we have that, on a closed curve C around z_0,

$$\oint_C \frac{dz}{z - z_0} = 2\pi \imath.$$

We can now ask what happens if we try to calculate, for a function $f(z)$ that is analytic inside this closed curve, as well as on it, the integral

$$I = \oint_C \frac{f(z)}{z - z_0} \, dz.$$

We already know the behavior of the vectors of $(z - z_0)^{-1}$ around z_0, they grow without limit, in modulus, as we approach z_0, and the corresponding vectors $\vec{w}_I$ are oriented in the tangent direction to the circles centered at z_0, for which reason they give rise to finite and non-zero integrals over curves around that point. In other words, close to and around that point the function $(z - z_0)^{-1}$ is large in magnitude and its variations are large but purely angular. On the other hand, the function $f(z)$ is analytic by hypothesis, and therefore regular at z_0, so that it approaches a constant when we take the limit $z \to z_0$, given by the value $f(z_0)$. Since we can deform the contour to a circular contour around z_0, with a radius as small as we wish, on which $f(z)$ is practically constant, it follows that the simple pole $(z - z_0)^{-1}$ dominates the behavior of the integrand at that point, so that one is tempted to conclude that

$$\begin{aligned} I &= f(z_0) \oint_C \frac{dz}{z - z_0} \\ &= 2\pi \boldsymbol{\imath} f(z_0). \end{aligned}$$

Interestingly, this equation is indeed true, but the analysis needs to be a little more careful and precise, as the same kind of argument could lead us to the conclusion that

$$\begin{aligned} \oint_C \frac{f(z)}{(z - z_0)^2}\, dz &= f(z_0) \oint \frac{dz}{(z - z_0)^2} \\ &= 0, \end{aligned}$$

which is *not* true. In order to improve and refine our argument, let us consider the sum

$$I = \oint_C \frac{f(z)}{z - z_0}\, dz - \oint_C \frac{f(z_0)}{z - z_0}\, dz + \oint_C \frac{f(z_0)}{z - z_0}\, dz,$$

which of course is the same as its first term. Joining the first two terms and considering that $f(z_0)$ is a constant we can write

$$I = \oint_C \frac{f(z) - f(z_0)}{z - z_0}\, dz + f(z_0) \oint_C \frac{dz}{z - z_0}.$$

The second term gives indeed the expected result, while the integrand in the first term, if we take the limit $z \to z_0$, by definition gives us the derivative of $f(z)$ at z_0. We can actually take this limit without difficulty, simply by deforming the contour C to an infinitesimal circle around z_0, thereby making every point of this circle tend to z_0, so that we can write for the integral,

$$I = \oint_C \frac{df}{dz}(z_0)\, dz + 2\pi\imath f(z_0),$$

since the result of the limit that defines the derivative does not depend on the direction in which one makes the variation of z. Since $f(z)$ is analytic at z_0 by hypothesis, we know that the derivative exists and is finite. It is in fact a constant in the integrand of the integral in the first term. This is therefore the integral of an analytic function (the constant function) on a closed contour, which would be sufficient to show that the integral vanishes, even without the fact, also found here, that the contour in question has zero length on the $z \to z_0$ limit. It should be noted here that it is *not* necessary to use the fact that the derivative of $f(z)$ is also an analytic function, but only the fact that it exists and is finite, which stems from the fact that $f(z)$ is analytic. We can therefore conclude that the first term is in fact zero and it finally results that $I = 2\pi\imath f(z_0)$, that is, it results that

$$f(z_0) = \frac{1}{2\pi\imath} \oint_C \frac{f(z)}{z - z_0}\, dz,$$

for any curve C containing the point z_0. This is the famous and important *Cauchy integral formula.* Let us take a moment to interpret the meaning of this formula, as its derivation is so easy and quick that this meaning is not immediately apparent. What we see above is the fact that the value $f(z_0)$ of an analytic function at a point z_0 can be written as an integral, over a closed curve around the point, of another function,

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

built very simply from the original function. Note that the function $g(z)$ is also analytic, at every point where $f(z)$ is analytic, except for the point z_0. In particular, since z_0 is within the curve and not on it, $g(z)$ is regular over the entire length of the curve C, which has finite length, which in turn shows that there cannot be any problem as to the existence of the integral.

Thus we see that the value of the analytic function $f(z)$ at the point z_0 is completely determined by the values of the function on the curve. Furthermore, the point z_0 could be *any* point internal to the curve, so that it follows an extraordinary fact, which we put in evidence here.

The values of the function $f(z)$ in the whole interior of a region of the complex plane in which it is analytic are determined by the values of $f(z)$ only on a closed curve which is the boundary of that region.

Note that this is reminiscent of the similar situation relating to the solution of a partial differential equation, given the values on the boundary. Indeed, $u(x, y)$ and $v(x, y)$ are solutions of the Laplace equation with the values given on the boundary, as we have seen before. We see here, once again, how these analytic functions are special, because we find that there are strong constraints between their values at various different points.

We will now explore the possible generalizations of this formula. A simple way to do this is to consider the possibility of differentiating this formula with respect to z_0, taking the derivative inside the integral, acting directly on the only part of the integrand which depends on z_0, which leads us to write tentatively that

$$f'(z_0) = \frac{1}{2\pi\imath} \oint_C \frac{f(z)}{(z - z_0)^2}\, dz,$$

an expression that now determines the derivative $f'(z)$ of the function $f(z)$ in terms of an integral involving $f(z)$ on the same contour C. Again, let us look at this with a little more care to disperse the initial impression that the right-hand side of this equation vanishes, as indicated by the superficial analysis that we discussed before. We compute the derivative of $f(z)$ by its definition, involving the difference of the values at two points, which we call z_1 and z_0, using therefore the Cauchy integral formula to write that

$$\begin{aligned}
\frac{f(z_1) - f(z_0)}{z_1 - z_0} &= \frac{1}{2\pi\imath(z_1 - z_0)} \left[\oint_C \frac{f(z)}{z - z_1}\, dz - \oint_C \frac{f(z)}{z - z_0}\, dz \right] \\
&= \frac{1}{2\pi\imath(z_1 - z_0)} \oint_C \left[\frac{1}{z - z_1} - \frac{1}{z - z_0} \right] f(z)\, dz \\
&= \frac{1}{2\pi\imath(z_1 - z_0)} \oint_C \frac{(z - z_0) - (z - z_1)}{(z - z_1)(z - z_0)}\, f(z)\, dz \\
&= \frac{1}{2\pi\imath(z_1 - z_0)} \oint_C \frac{(z_1 - z_0)}{(z - z_1)(z - z_0)}\, f(z)\, dz \\
&= \frac{1}{2\pi\imath} \oint_C \frac{1}{(z - z_1)(z - z_0)}\, f(z)\, dz.
\end{aligned}$$

Since the derivative of $f(z)$ at z_0 is the limit of the left-hand side of this expression when $z_1 \to z_0$, and since in this same limit we have the square of $(z - z_0)$ in the denominator on the right-hand side, it follows in fact that

$$f'(z_0) = \frac{1}{2\pi\imath} \oint_C \frac{f(z)}{(z - z_0)^2} \, dz.$$

Thus we see that the derivative of f can also be written in terms of a contour integral. Having shown that this is not zero, we can ask ourselves what was wrong with the more superficial analysis considered initially. The point is that in the case of the integral of $w(z) = 1/(z - z_0)$, the sum involving the vectors $\vec{w}_I$ around the circle was not zero, so that the multiplication by the regular function $f(z)$, which is almost constant near $z = z_0$, cannot significantly change the result. On the other hand, in the case of the integral of $w(z) = 1/(z - z_0)^2$ the sum involving the vectors $\vec{w}_I$ around the circle is zero, consisting of a precise cancellation of large quantities when we approach z_0. In this case the multiplication by $f(z)$ affects the result, because although $f(z)$ does not vary too much over the sum, the existing small variations are sufficient to disturb this precise cancellation between large quantities, resulting in the non-zero result shown above.

Considering the two results so far, we are tempted to try to generalize them, taking again a derivative of the above expression with respect to z_0, and thereby obtaining the second derivative

$$f''(z_0) = \frac{2}{2\pi\imath} \oint_C \frac{f(z)}{(z - z_0)^3} \, dz.$$

Again, we can analyze this in more detail, using as before two points z_1 and z_0, and using the formula for the first derivative in order to write

$$\frac{f'(z_1) - f'(z_0)}{z_1 - z_0} = \frac{1}{2\pi\imath(z_1 - z_0)} \oint_C f(z) \left[\frac{1}{(z - z_1)^2} - \frac{1}{(z - z_0)^2} \right] dz,$$

where the expression in square brackets can be worked as follows,

$$\begin{aligned}
\frac{1}{(z - z_1)^2} - \frac{1}{(z - z_0)^2} &= \frac{(z^2 + z_0^2 - 2zz_0) - (z^2 + z_1^2 - 2zz_1)}{(z - z_1)^2 (z - z_0)^2} \\
&= \frac{-(z_1^2 - z_0^2) + 2z(z_1 - z_0)}{(z - z_1)^2 (z - z_0)^2} \\
&= (z_1 - z_0) \frac{2z - (z_1 + z_0)}{(z - z_1)^2 (z - z_0)^2}
\end{aligned}$$

$$= \quad (z_1 - z_0)\frac{(z - z_1) + (z - z_0)}{(z - z_1)^2(z - z_0)^2},$$

so that we have

$$\frac{f'(z_1) - f'(z_0)}{z_1 - z_0} = \frac{1}{2\pi\imath}\oint_C f(z)\frac{(z - z_1) + (z - z_0)}{(z - z_1)^2(z - z_0)^2}\,dz.$$

On the $z_1 \to z_0$ limit the expression on the left approaches the second derivative, whereas the numerator on the right reduces to $2(z - z_0)$, canceling with one of the factors in the denominator, so that we obtain the expected result,

$$f''(z_0) = \frac{2}{2\pi\imath}\oint_C \frac{f(z)}{(z - z_0)^3}\,dz.$$

Thus we see that the second derivative of the analytic function $f(z)$ can also be written as an integral over the curve C. The generalization of the result for the case of the n-th derivative is now immediate, since it is clear that it suffices to take various derivatives with respect to z_0, and thus generalize the result to

$$f^{n\prime}(z_0) = \frac{n!}{2\pi\imath}\oint_C \frac{f(z)}{(z - z_0)^{n+1}}\,dz.$$

This result can be established more precisely by finite induction, assuming that it holds for the case $(n - 1)$,

$$f^{(n-1)\prime}(z_0) = \frac{(n - 1)!}{2\pi\imath}\oint_C \frac{f(z)}{(z - z_0)^n}\,dz,$$

and then proving that this implies the validity of the case n, which can be written in terms of this formula as

$$\frac{f^{(n-1)\prime}(z_1) - f^{(n-1)\prime}(z_0)}{z_1 - z_0}$$
$$= \quad \frac{(n - 1)!}{2\pi\imath(z_1 - z_0)}\oint_C f(z)\left[\frac{1}{(z - z_1)^n} - \frac{1}{(z - z_0)^n}\right]dz.$$

However, the derivation is rather long and will be left as an exercise, with the following recommendations regarding the handling of the integrand above: do not separate the expression $(z - z_0)$ during the calculations; write $\delta z = z_1 - z_0$, recalling that δz will be eventually taken to zero;

and write the difference $z - z_1$ as $z - z_1 = (z - z_0) - \delta z$. Note that it will be necessary to use the formula of the binomial expansion,

$$(a+b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\, a^k b^{n-k},$$

that, being of a purely arithmetic nature, holds for complex numbers as well as for real numbers.

Let us now return to the question of showing that the derivative $f'(z)$ of an analytic function $f(z)$ is also an analytic function, which we discussed earlier. As we saw earlier, if $f = u + \imath v$, then the derivative has the real and imaginary components

$$f'(z) = \frac{\partial u}{\partial x} + \imath \frac{\partial v}{\partial x},$$

that is, we have $f' = u' + \imath v'$ with

$$\begin{aligned} u' &= \frac{\partial u}{\partial x} = \frac{\partial v}{\partial y}, \\ v' &= \frac{\partial v}{\partial x} = -\frac{\partial u}{\partial y}, \end{aligned}$$

given that u and v satisfy the Cauchy-Riemann conditions. As we saw earlier, so long as the partial derivatives of u' and v' exist, they also satisfy the Cauchy-Riemann conditions. One of the things that the relation shown above for the case of the first derivative, which is given by

$$f'(z_0) = \frac{1}{2\pi\imath} \oint_C \frac{f(z)}{(z-z_0)^2}\, dz,$$

provides us with is the certainty that the derivative exists, that is, that the partial derivatives u' and v' of u and v exist, since we have shown that they can be written explicitly in terms of integrals of a function over a finite domain where this function is fully regular. It follows that if f is analytic, then u and v are necessarily differentiable, as we know. However, one can go beyond that. Previously, in order to prove that the derivative of $f(z)$ is also an analytic function, we had to *assume* that the second partial derivatives of u and v exist. However, we now have the relation

$$f''(z_0) = \frac{2}{2\pi\imath} \oint_C \frac{f(z)}{(z-z_0)^3}\, dz,$$

which gives us the certainty that the second derivative exists, that is, the partial derivatives of u' and v', which are the components u'' and v'' of $f''(z)$, also exist, for once again we have an explicit expression for them in terms of integrals that clearly exist. Thus, it is not necessary to assume separately that these second derivatives exist because we have an explicit construction of them in terms of clearly convergent integrals.

It follows therefore that the derivative of an analytic function is also an analytic function without the need for any additional hypothesis. It is enough that the original function be analytic at a certain point for its derivative to exist and to be analytic at that point. Since this argument can be iterated indefinitely, we have that an analytic function always has derivatives of all orders, and that these are all analytic. We say that analytic functions are of class C^∞.

Armed with this fact, the identification of sequences of analytic functions obtained by iterated differentiation and by iterated integration, which was discussed in the previous chapter (Chapter 8) in connection to the fundamental theorem of the calculus, is now fully established. Both in the case of integration and in the case of differentiation, the operation performed on an analytic function produces another analytic function without the need for any additional hypotheses.

Due to this fact there is one more additional thing that we can observe in the case of the integration to define the primitive, and of the proof that this is analytic, which were discussed in the previous chapter (Chapter 8). All that was used in order to enable the unambiguous definition of the primitive $F(z)$ was that the indefinite integral of $f(z) = u + \imath v$ be independent of the path. The fact that $f(z)$ is analytic was *not* used. Of course, if $f(z)$ is analytic it follows that the integral does not depend on the path, but we can repeat the proof without any change, only with the weaker hypothesis that $f(z)$ is a complex function such that

$$\oint_C f(z)\, dz = 0$$

for all closed curves C within the region under consideration. With this, we have that the indefinite integral of $f(z)$ does not depend on the path and so we can define the primitive $F(z)$ unequivocally. In addition to this, for the proof that $F(z)$ is analytic the Cauchy-Riemann relations for $f(z) = u + \imath v$ were not used, and not even the existence of the partial derivatives of $u(x, y)$ and $v(x, y)$ was assumed. All that was used was the existence of $u(x, y)$ and $v(x, y)$, and the expression of $F(z)$

$$F(z) = U + \imath V,$$

in terms of $U(x, y)$ and $V(x, y)$, which relate to the vector fields $\vec{w}_R$ and $\vec{w}_I$ by

$$\begin{aligned} \vec{w}_R &= \vec{\nabla} U(x, y), \\ \vec{w}_I &= \vec{\nabla} V(x, y), \end{aligned}$$

vector fields that in turn are given by

$$\begin{aligned} \vec{w}_R &= (u, -v), \\ \vec{w}_I &= (v, u). \end{aligned}$$

Putting all these elements together we have the relations between (U, V) and (u, v), repeating here the equations derived in the previous chapter (Chapter 8),

$$\begin{aligned} \frac{\partial U}{\partial x} &= u, \\ \frac{\partial U}{\partial y} &= -v, \\ \frac{\partial V}{\partial x} &= v, \\ \frac{\partial V}{\partial y} &= u. \end{aligned}$$

It follows that the primitive $F(z)$ of a complex function $f(z)$, whose integral has the path independence property, is always analytic, regardless of the hypothesis that $f(z)$ be analytic. However, we can now use what we learned in the previous chapter (Chapter 8) on the fundamental theorem of calculus to come to a new conclusion. If $F(z)$ is analytic, then its derivative is also analytic, as we saw earlier in this chapter. However, this derivative is precisely $f(z)$, as we saw in the previous chapter (Chapter 8). It follows that in this way we have a new theorem, known as the *Morera theorem*, which is the reverse of the Cauchy theorem: if we have a complex function $f(z)$ such that

$$\oint_C f(z)\, dz = 0,$$

for all curves C in a particular region of the complex plane, then it follows that $f(z)$ is analytic in that region. Thus we see that two seemingly very different things are in fact the same thing for a complex function $f(z)$: to be an analytic function, that is, to be differentiable and satisfy the Cauchy-Riemann conditions, and to be a "conservative" function, that is, to have zero integrals on all closed curves.

In short, besides having previously verified the validity of the fundamental theorem of calculus for the analytic functions, similarly to the real case, we verify here that the relations between integrals and derivatives are much richer in the case of the complex calculus. As we shall see, these relations will be central and instrumental in the study ahead, involving the representation of analytic functions by power series, which is a key instrument for the applications in physics, besides being mathematically important on their own.

Problem Set

1. Calculate the integral

$$\oint_C \frac{z^n}{(z-z_0)^2}\,dz,$$

for a non-negative integer n, on an arbitrary closed curve C that makes a single turn around the point z_0. Determine for which combinations of values of n and z_0 the result is zero.

2. Calculate the integral

$$\oint_C \frac{\sqrt{z}}{(z-z_0)^n}\,dz,$$

for a non-negative integer n, on an arbitrary closed curve C that makes a single turn around the point z_0 and does not contain the origin. Calculate the $z_0 \to 0$ limit of the result obtained. How can we interpret the result in this limit?

Hint: use the Cauchy integral formulas.

3. Calculate the integral

$$\oint_C \frac{\sin(z)}{(z - z_0)^2}\, dz,$$

on an arbitrary closed curve C that makes a single turn around the point z_0. Determine for which values of z_0 the result is zero.

Hint: use the Cauchy integral formulas.

4. Calculate the integral

$$\oint_C \frac{\sin(z)}{(z - z_0)^3}\, dz,$$

on an arbitrary closed curve C that makes a single turn around the point z_0. Determine for which values for z_0 the result is zero.

Answer: $-\imath\pi \sin(z_0)$.

Hint: use the Cauchy integral formulas.

5. Prove by finite induction the Newton binomial formula,

$$(a + b)^n = \sum_{k=0}^{n} \frac{n!}{k!(n-k)!}\, a^k b^{n-k},$$

where a and b are any complex numbers. Make sure that the proof holds for complex numbers.

6. Prove by finite induction the Cauchy integral formula for the n-th derivative of an analytic function,

$$f^{n\prime}(z_0) = \frac{n!}{2\pi\imath} \oint_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz,$$

that is, assume the formula for the case $n - 1$ and show that this implies that it holds for the case n.

Hints: do not separate the expression $(z - z_0)$ during the calculations; write $\delta z = z_1 - z_0$, recalling that δz will be eventually taken to zero; and write the difference $z - z_1$ as $z - z_1 = (z - z_0) - \delta z$.

7. Consider an analytic function $f(z)$. Suppose that we differentiate it a certain number n of times, and then integrate it the same number of times in the sense of indefinite integration or primitivization, thus obtaining an analytic function $g(z)$. Write the most general possible form of the difference $g(z) - f(z)$.

Chapter 10

Complex Inequalities and Series

Our next goal is to work our way to the series of complex powers and the representation of analytic functions through them. However, before addressing the issue of the series, we have to go back and talk a little more extensively about inequalities involving absolute values of complex numbers, that we have discussed in a previous chapter (Chapter 2), as these will be needed for the proofs of convergence. First, if z_1 and z_2 are two complex numbers, let us recall how to connect the absolute value of their product with the absolute values of each number. In this case it is simple to see that

$$|z_1 z_2| = |z_1|\,|z_2|.$$

In order to show this, it suffices to write that $z_1 = \rho_1 \exp(\imath\theta_1)$ and that $z_2 = \rho_2 \exp(\imath\theta_2)$, so that we have

$$\begin{aligned} |z_1 z_2| &= \left|\rho_1 \rho_2 \, e^{\imath(\theta_1+\theta_2)}\right| \\ &= \rho_1 \rho_2 \\ &= |z_1|\,|z_2|. \end{aligned}$$

In the case of the sum of two or more complex numbers things get a bit more complicated. As we have seen in a previous chapter (Chapter 2), in this case we can think of z_1, z_2 and $z_1 + z_2$ as vectors, as shown in the diagram of Figure 10.1.

It is immediately clear that $|z_1 + z_2|$ is smaller than $|z_1| + |z_2|$, unless z_1 and z_2 are aligned, the only case in which the equality holds. That is,

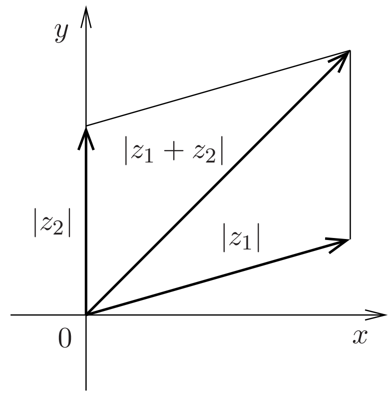


Figure 10.1: The complex plane with the sum of two vectors, illustrating the relation between their absolute values.

it only holds for $z_1 = Cz_2$ with some real positive constant C. Therefore we have the basic triangle inequality

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

As we also discussed in an earlier chapter (Chapter 2), it is simple to generalize this relation to the sum of more than two numbers, because we have immediately, for example, that

$$\begin{aligned} |z_1 + z_2 + z_3| &\leq |z_1| + |z_2 + z_3| \\ &\leq |z_1| + |z_2| + |z_3|, \end{aligned}$$

so that for the sum of N numbers, for any value of N, it holds that

$$\left| \sum_{i=1}^{N} z_i \right| \leq \sum_{i=1}^{N} |z_i|.$$

Since an integral is fundamentally an infinite sum, it is not difficult to verify that a similar relation holds for complex integrals on an integration contour C,

$$\left| \oint_C f(z)\, dz \right| \leq \oint_C |f(z)\, dz|\,,$$

so long as the two integrals involved exist. The details of this case will be left as exercises. All inequalities of this type are called triangle inequalities. This last inequality allows us to make the following observation about the Cauchy integral formula which we derived in the previous chapter (Chapter 9),

$$f(z_0) = \frac{1}{2\pi\imath} \oint_C \frac{f(z)}{z - z_0} \, dz,$$

where we are considering, for simplicity, a circular contour C of radius r_0 around z_0, that is, we are considering a polar coordinate system (r, θ) centered at z_0. If we take absolute values we have

$$|f(z_0)| = \frac{1}{2\pi} \left| \oint_C \frac{f(z)}{z - z_0} \, dz \right|,$$

where we used the rule for the product. Since the integral is a sum, using the triangle inequalities we also have the relation of inequality

$$\begin{aligned} |f(z_0)| &\leq \frac{1}{2\pi} \oint_C \left| \frac{f(z)}{z - z_0} \, dz \right| \\ &= \frac{1}{2\pi} \oint_C \frac{|f(z)|}{|z - z_0|} \, |dz|, \end{aligned}$$

where we used again the product rule in this last passage. We write now all the quantities in terms of r and θ. Recalling that we are on a circle of radius r_0 centered at z_0, we have that $z - z_0 = r_0 \exp(\imath\theta)$, and therefore that $|z - z_0| = r_0$, and that $|dz| = r_0 \, d\theta$, where θ ranges from 0 to 2π along the contour. It follows that for our relation of inequality we have

$$\begin{aligned} |f(z_0)| &\leq \frac{1}{2\pi} \oint_C |f(z)| \, d\theta \\ &= \frac{1}{2\pi} \int_0^{2\pi} |f(z)| \, d\theta \\ &= \overline{|f|}_C, \end{aligned}$$

which is the average value of $|f|$ on the contour C. With this result we can already see that if $f(z)$ does not have constant absolute value on the contour C, then surely there is some point on C where the absolute value of $f(z)$ is larger than $|f(z_0)|$. Let us now verify, with more precision, when the equality is realized. For this to happen it is necessary that in the sum involved in the integral

$$\oint_C \frac{f(z)}{z - z_0} \, dz$$

all terms be aligned with each other. Writing all complex numbers in terms of the polar coordinate system (r, θ) centered at z_0, and recalling that we are still on a circle of radius r_0, we have $z - z_0 = r_0 \exp(\imath\theta)$, for the variation of z we have $dz = \imath r_0 \exp(\imath\theta)\,d\theta$ and for the function $f(z)$ we have $f = |f|\exp(\imath\alpha)$, so that we can write for this integral

$$\begin{aligned}\oint_C \frac{f(z)}{z - z_0}\,dz &= \oint_C \frac{|f|\,\mathrm{e}^{\imath\alpha}}{r_0\,\mathrm{e}^{\imath\theta}}\,\imath r_0\,\mathrm{e}^{\imath\theta}d\theta \\ &= \imath\oint_C |f|\,\mathrm{e}^{\imath\alpha}d\theta,\end{aligned}$$

where the direction of the complex numbers being added together in this last integral is given by the exponential with imaginary argument $\imath\alpha$. Thus, in order to have the equality when we take the absolute value on both sides and use the triangle inequality, we need $\exp(\imath\alpha)$ to always have the same direction in the complex plane along the curve C, being therefore constant over it, thus being removable from inside the integral, and subsequently made equal to 1 by the extraction of the absolute value. What is of interest to us here is that for $\exp(\imath\alpha)$ to always have the same direction, we need α to be constant along the curve, so that we can write the integral as

$$\oint_C \frac{f(z)}{z - z_0}\,dz = \imath\,\mathrm{e}^{\imath\alpha}\oint_C |f|\,d\theta.$$

If we now recall that, by the Cauchy-Goursat theorem, the integral on the left-hand side of this equation has the same value for circles with any radius $r_0 > 0$ around z_0, which is also evidenced by the fact that $2\pi\imath f(z_0)$ is equal to this integral, and obviously does not depend on r_0, we conclude that α must be constant throughout the interior of the circle, and thus in a neighborhood of z_0. This is true because it follows from the above argument that the complex quantity

$$\mathrm{e}^{\imath\alpha}\oint_C |f|\,d\theta = \rho_0\,\mathrm{e}^{\imath\alpha},$$

where ρ_0 is some real number, is independent of r_0. Since this complex number is determined jointly by its absolute value ρ_0 and by its direction $\exp(\imath\alpha)$, it follows that neither ρ_0 nor α may depend on r_0. The fact that α is a real constant implies that we can write for the function $f(z)$, in the whole interior of the circle of radius r_0,

$$f(z) = |f|[\cos(\alpha) + \imath\sin(\alpha)],$$

which means that it has a constant direction, that is, the real and imaginary parts of $f(z)$ are proportional, with a real multiplicative constant. Assuming that $\cos(\alpha) \neq 0$, this means that $v = Au$ for a real constant $A = \tan(\alpha)$, and thus we have

$$f(z) = u(x, y) + \imath A u(x, y).$$

In the case that $\cos(\alpha) = 0$ we have that $\sin(\alpha) \neq 0$, and we can write instead of this

$$f(z) = A v(x, y) + \imath v(x, y),$$

ultimately leading to the same results that we will obtain here for the case $\cos(\alpha) \neq 0$, by means of the argument that follows. Considering therefore the case $\cos(\alpha) \neq 0$, the Cauchy-Riemann conditions give us then

$$\begin{aligned} \frac{\partial u}{\partial x} &= \frac{\partial v}{\partial y} \\ &= A \frac{\partial u}{\partial y}, \\ \frac{\partial u}{\partial y} &= -\frac{\partial v}{\partial x} \\ &= -A \frac{\partial u}{\partial x}. \end{aligned}$$

Combining these two equations we can write for u that

$$\begin{aligned} \frac{\partial u}{\partial x} &= A \frac{\partial u}{\partial y} \\ &= -A^2 \frac{\partial u}{\partial x}. \end{aligned}$$

We conclude from this that, unless we have

$$\frac{\partial u}{\partial x} = 0,$$

the resulting equation implies that $A^2 = -1$, which is impossible because A is real. It follows that the partial derivative of u must vanish, which also implies the vanishing of the other partial derivative of u,

$$\begin{aligned} \frac{\partial u}{\partial y} &= -A \frac{\partial u}{\partial x} \\ &= 0. \end{aligned}$$

Since the Cauchy-Riemann conditions imply that the partial derivatives of v are also proportional to those of u, we have

$$\begin{aligned} \frac{\partial u}{\partial x} &= 0, \\ \frac{\partial u}{\partial y} &= 0, \\ \frac{\partial v}{\partial x} &= 0, \\ \frac{\partial v}{\partial y} &= 0. \end{aligned}$$

This means that $u(x,y)$ and $v(x,y)$ are real constants and therefore that $f(z)$ is some complex constant. It follows therefore that this is the *only* case in which the equality is realized in the triangle relations, for a function $f(z)$ that is analytic. We therefore have that, unless $f(z)$ is a constant,

$$\begin{aligned} |f(z_0)| &< \frac{1}{2\pi} \oint_C |f(z)| \, d\theta \\ &= \overline{|f|}_C. \end{aligned}$$

It follows that, unless $f(z)$ is a constant, $|f(z)|$ cannot assume its maximum value strictly within the contour, but only on one or more boundary points on the contour C itself. This result is called the *maximum modulus theorem*. It is interesting to observe that the two harmonic functions $u(x,y)$ and $v(x,y)$ which constitute $f(z)$ also have a property similar to this one, namely they only assume their maximum and minimum values at the boundary of the region, never inside. This is a property of all solutions of the Laplace equation, which is satisfied by these functions within the region where $f(z)$ is analytic.

We observe here that the maximum modulus theorem can now be generalized to any contour, that is, to the boundary of a region that need not be necessarily circular. It is sufficient to note that, given any region where the non-constant function $f(z)$ is analytic, and given any internal point z_0 in this region, there is always an open disk, that is, the interior of a disk without its boundary, with a radius different from zero and centered at z_0, that is contained in the region. It follows that there is some point z_1 on the boundary of this disc, and therefore within the region, where $|f(z)|$ is strictly larger than $|f(z_0)|$, that is, where we have that

$$|f(z_1)| > |f(z_0)|,$$

and thus we conclude that $|f(z)|$ does not assume its maximum at z_0. Since this argument holds for *all* internal points, it follows that $|f(z)|$ does not assume its maximum at any internal point, but just on the boundary, whatever the shape of the region.

Of course, if the point z_0 is on the boundary, then there is no disk centered on it and contained in the region, so that the argument we just used is not valid in this case. The interior of the disk centered at z_0 is called an *open neighborhood* of z_0. It is interesting to observe that, given any starting point z_0, we can build in this way a sequence of points and neighborhoods, points at which $|f(z)|$ is monotonically increasing, and that inevitably leads us to the boundary of the area.

The maximum modulus theorem has another very interesting consequence. As a preliminary point for this discussion let us note that, just as from the Cauchy integral formula we can derive the inequality

$$\begin{aligned} |f(z_0)| &\leq \frac{1}{2\pi} \oint_C |f(z)|\, d\theta \\ &= \overline{|f|}_C, \end{aligned}$$

we can apply the same reasoning and derivations to the related formulas that give the derivatives of $f(z)$ in terms of contour integrals,

$$f^{n\prime}(z_0) = \frac{n!}{2\pi\imath} \oint_C \frac{f(z)}{(z-z_0)^{n+1}}\, dz,$$

thus obtaining the inequalities, for all $n \geq 0$,

$$\begin{aligned} \left|f^{n\prime}(z_0)\right| &\leq \frac{n!}{2\pi} \oint_C \frac{|f(z)|}{r_0^n}\, d\theta \\ &= \frac{n!}{r_0^n}\, \overline{|f|}_C. \end{aligned}$$

We will leave the details of this derivation as an exercise. In order to discuss the next theorem of interest, let us make the hypothesis that a function $f(z)$ is analytic throughout the plane and that it has upper and lower bounds, that is, there are real numbers f_M and f_m such that

$$f_m < |f(z)| < f_M,$$

for all values of z. Of course, since this is an absolute value, we always have the lower bound given by $f_m = 0$, so that the only relevant hypothesis is that there is an upper bound. Since $f(z)$ is continuous, limited and is defined on the entire plane, this implies that $|f(z)|$ has a maximum at some point z_0 and assumes its maximum at that point, or tends to this maximum when $z \to \infty$ in some direction of the complex plane. Using the fact that $|f(z)|$ cannot assume its maximum in any region within a contour containing z_0, unless $f(z)$ is a constant, it follows that $f(z)$ must be constant in all these regions, or else tend to its maximum when $z \to \infty$ in some direction of the complex plane. In order to decide on the latter case, we write the inequality involving the derivative of $f(z)$,

$$\left|f'(z_0)\right| \leq \frac{1}{r_0}\,\overline{|f|}_C.$$

This holds for contours C of unlimited size, since $f(z)$ is analytic throughout the plane. Since $|f(z)|$ is limited by f_M throughout the whole complex plane, its average over the contour is also limited, so that we have

$$\left|f'(z_0)\right| \leq \frac{1}{r_0}\,f_M.$$

Making $r_0 \to \infty$, we concluded that $|f'(z_0)| = 0$, which implies that the derivative of $f(z_0)$ is zero. Since this applies to all points in the plane, it follows that the derivative is identically zero and therefore that $f(z)$ is constant over the entire complex plane. In short, we have the result that, if $f(z)$ is analytic on the entire plane and is limited, then $f(z)$ is constant. This is the *Liouville theorem.*

The Liouville theorem can be used to prove in a very simple way the fundamental theorem of algebra, serving as an example of the power of this formalism and its applicability in other areas of mathematics. The fundamental theorem of algebra states that, given the polynomial of degree n

$$P_n(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_{n-1} z^{n-1} + a_n z^n,$$

with complex coefficients $a_0, \ldots, a_n$, where $a_n \neq 0$ and $n > 0$, it holds that the polynomial has at least one complex root, that is, a value of z for which $P_n(z) = 0$. We will prove the theorem by reductio ad absurdum. If we assume that $P_n(z) \neq 0$ for all z, then the rational function

$$f(z) = \frac{1}{P_n(z)}$$

is analytic on the entire complex plane, and is also limited, because we have that

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

and since the polynomial is not limited, we have that $|f| \to 0$ when $|z| \to \infty$. It follows, by the Liouville theorem, that $f(z)$, and therefore $P_n(z)$, are constant, which is clearly absurd, since by hypothesis $n > 0$. It follows therefore, very simply, that P has at least one root, some number z_1 along the complex plane such that $P(z_1) = 0$. This establishes the result with great ease and simplicity. This theorem is, however, extremely difficult to prove by purely algebraic means.

Note that one can iterate the argument. Having determined the existence of a complex root z_1 of $P_n(z)$, one can then factor the polynomial in the form

$$\begin{aligned} P_n(z) &= (z - z_1)P_{n-1}(z), \\ P_{n-1}(z) &= b_0 + b_1 z + b_2 z^2 + \ldots + b_{n-2} z^{n-2} + b_{n-1} z^{n-1}, \end{aligned}$$

and repeat the argument for $P_{n-1}(z)$, thus showing that there is a second root of $P_n(z)$, and so on, until only a constant term remains in the polynomial P_0, which in fact has no roots. Thus, we can see that a polynomial of order n always has n complex roots, which is how this theorem is often remembered.

As a last preliminary topic to allow us to discuss the complex power series, we must examine the issue of the sum of a complex geometric progression. Let us assume that, for some complex α, we want to calculate the finite sum S given by

$$S = 1 + \alpha + \alpha^2 + \alpha^3 + \ldots + \alpha^n.$$

Since the sum is finite, there are no convergence issues here, so that the problem is purely algebraic. The case $\alpha = 1$ needs to be examined in separate, but is trivial, because in this case all the terms are equal to 1, and the sum is equal to $n + 1$. In the other case, in which $\alpha \neq 1$, we can calculate the sum algebraically, multiplying and dividing the above expression by $\alpha - 1$, which gives us

$$\begin{aligned}
S &= \frac{(\alpha-1)(1+\alpha+\alpha^2+\ldots+\alpha^n)}{\alpha-1} \\
&= \frac{1}{\alpha-1}\times \\
&\quad \times \Big(\quad\quad \alpha+\alpha^2+\ldots+\alpha^{n-1}+\alpha^n+\alpha^{n+1} \\
&\quad\quad -1-\alpha-\alpha^2-\ldots-\alpha^{n-1}-\alpha^n \quad\quad \Big) \\
&= \frac{\alpha^{n+1}-1}{\alpha-1} \\
&= \frac{a_n q - a_0}{q-1},
\end{aligned}$$

where, in order to use the most common notation, we define $a_0 = 1$, $a_n = \alpha^n$ and $q = \alpha$. We see that this is indeed the usual formula. It is very important to note that since the derivation is purely arithmetic, using only the properties of the field, it is clear that the result holds equally well for real numbers and for complex numbers. It will be useful, in the development that follows, to use the sum written with the terms from 1 to $n-1$, for which we have

$$1+\alpha+\alpha^2+\ldots+\alpha^{n-1} = \frac{1-\alpha^n}{1-\alpha}.$$

We are now ready to face the question of the series of complex powers, including the determination and proof of the conditions that ensure their convergence. Let us start with the problem of, given an analytic function $f(z)$, developing a power series that converges to it. The construction of the series begins with a purely algebraic argument. For this, consider the Cauchy integral formula for $f(z)$,

$$f(z) = \frac{1}{2\pi\imath}\oint_C \frac{f(z')}{z'-z}\,dz',$$

where the integral is defined on a contour C within the region of analyticity of $f(z)$, z is a generic point strictly within the contour, and z' is a point located on the integration contour. Consider also a point z_0 which is a given fixed point strictly within the contour. The contour and the points involved are shown in the diagram of Figure 10.2.

What we would like do is to develop a series expansion for $f(z)$ around the reference point z_0. We can write the following purely algebraic relation among these points, involving the denominator that appears in the Cauchy integral formula,

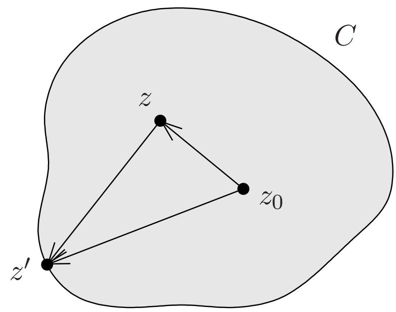


Figure 10.2: The complex plane with the contour, the three points and the three difference vectors.

$$z' - z = (z' - z_0) - (z - z_0),$$

that relates the three difference vectors, which connect in pairs the three points. Using this we can write the fraction that appears in the Cauchy integral formula as follows,

$$\begin{aligned} \frac{1}{z' - z} &= \frac{1}{(z' - z_0) - (z - z_0)} \\ &= \frac{1}{z' - z_0} \frac{1}{1 - \left(\dfrac{z - z_0}{z' - z_0} \right)}. \end{aligned}$$

Note now that the first fraction on the right-hand side of the equation involves only the reference point and the points on the contour, and that the dependence on z, the point where we will calculate $f(z)$, is now only in the fraction in parenthesis on the right-hand side. In order to handle the second fraction on the right-hand side of this equation, we use the expression of the sum of the geometric progression, written as

$$1 + \alpha + \alpha^2 + \ldots + \alpha^{n-1} = \frac{1}{1 - \alpha} - \frac{\alpha^n}{1 - \alpha},$$

that is, we have that

$$\frac{1}{1 - \alpha} = 1 + \alpha + \alpha^2 + \ldots + \alpha^{n-1} + \frac{\alpha^n}{1 - \alpha},$$

where we will use for α the expression

$$\alpha = \frac{z - z_0}{z' - z_0}.$$

Therefore we have for the second fraction that we are examining,

$$\begin{aligned}
&\frac{1}{1 - \left(\dfrac{z - z_0}{z' - z_0}\right)} \\
&\quad = \quad 1 + \left(\frac{z - z_0}{z' - z_0}\right) + \left(\frac{z - z_0}{z' - z_0}\right)^2 + \ldots \\
&\qquad\qquad \ldots + \left(\frac{z - z_0}{z' - z_0}\right)^{n-1} + \frac{1}{1 - \left(\dfrac{z - z_0}{z' - z_0}\right)} \left(\frac{z - z_0}{z' - z_0}\right)^n,
\end{aligned}$$

and therefore, including the first fraction, we have the expression

$$\begin{aligned}
\frac{1}{z' - z} &= \frac{1}{z' - z_0} \frac{1}{1 - \left(\dfrac{z - z_0}{z' - z_0}\right)} \\
&= \frac{1}{(z' - z_0)} + \frac{(z - z_0)}{(z' - z_0)^2} + \frac{(z - z_0)^2}{(z' - z_0)^3} + \ldots \\
&\qquad \ldots + \frac{(z - z_0)^{n-1}}{(z' - z_0)^n} + \frac{1}{(z' - z)} \left(\frac{z - z_0}{z' - z_0}\right)^n.
\end{aligned}$$

We will use this in the expression of the Cauchy integral formula, thereby obtaining

$$\begin{aligned}
f(z) &= \frac{1}{2\pi \imath} \oint_C dz' f(z') \times \\
&\quad \times \left[\frac{1}{(z' - z_0)} + \frac{(z - z_0)}{(z' - z_0)^2} + \frac{(z - z_0)^2}{(z' - z_0)^3} + \ldots \right. \\
&\qquad \left. \ldots + \frac{(z - z_0)^{n-1}}{(z' - z_0)^n} + \frac{1}{(z' - z)} \left(\frac{z - z_0}{z' - z_0}\right)^n \right].
\end{aligned}$$

Distributing the integral to each term of the sum we obtain the expansion

$$\begin{aligned}
f(z) &= \frac{1}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)} \, dz' + \frac{(z - z_0)}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)^2} \, dz' + \\
&\quad + \frac{(z - z_0)^2}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)^3} \, dz' +
\end{aligned}$$

$$+ \ldots + \frac{(z - z_0)^{n-2}}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)^{n-1}} \, dz' +$$
$$+ \frac{(z - z_0)^{n-1}}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)^{n}} \, dz' + R_n,$$

where the *remainder* R_n is given by

$$R_n = \frac{(z - z_0)^n}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z)(z' - z_0)^n} \, dz'.$$

We now recognize, in each term of the sum, the Cauchy integral formulas that give the successive derivatives of $f(z)$ at the point z_0 in terms of integrals on the contour C around z_0,

$$f^{n\prime}(z_0) = \frac{n!}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)^{n+1}} \, dz',$$

that is, we have for the integrals appearing in the sum the expressions

$$\frac{1}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)^{n+1}} \, dz' = \frac{f^{n\prime}(z_0)}{n!},$$

so that we can write for our expansion

$$\begin{aligned} f(z) \quad = \quad & f(z_0) + f'(z_0)(z - z_0) + \frac{f''(z_0)}{2!} (z - z_0)^2 + \\ & + \frac{f^{3\prime}(z_0)}{3!} (z - z_0)^3 + \ldots + \\ & + \frac{f^{(n-2)\prime}(z_0)}{(n-2)!} (z - z_0)^{n-2} + \\ & + \frac{f^{(n-1)\prime}(z_0)}{(n-1)!} (z - z_0)^{n-1} + R_n. \end{aligned}$$

We recognize now the $(n - 1)$ first terms of this sum as the initial part of the Taylor series of the function $f(z)$, in this case written for complex functions, but that can immediately be restricted to the real axis to reproduce the known result of that case.

Note that, since we started our argument with a closed and exact expression for $f(z)$, we can immediately interpret the quantity R_n as the remainder, that is, the difference between the exact value and that obtained by adding the first so many terms of the series. In other words, R_n is the error that remains when we add the first $(n - 1)$ terms of the series, with respect to the exact value of $f(z)$, a fact that can be expressed as

$$R_n = f(z) - \sum_{i=0}^{n-1} \frac{f^{i\prime}(z_0)}{i!} (z - z_0)^i,$$

a quantity for which we in fact have an explicit expression, as seen above. We are thus led to discuss the question of convergence of the series, that is, of whether or not it is true that if we add an increasing number of terms of the series, making n increase without limit, the result will approach arbitrarily well the exact result. For this to be true, it is necessary that the value of R_n tend to zero when n tends to infinity. We will discuss this in terms of the absolute value $|R_n|$, that is, the question of convergence can be reduced to the discussion of the limit

$$\lim_{n\to\infty} |R_n|.$$

If we can show that this limit is zero, we will have proved the convergence of the Taylor series of an analytic function. This is what we will do in the next chapter, in which we will also discuss the issue of how to express the idea of convergence in a mathematically more precise way, so that it can be applied also to some more general cases than this one. As we shall see, along this process we will obtain explicitly the conditions under which the series is convergent, and also some idea about the speed of convergence, which is important in practice for the approximate calculation of the values of various functions of interest.

Problem Set

1. Starting from the discrete triangle inequality for an arbitrary integer N,

$$\left| \sum_{n=1}^{N} z_n \right| \leq \sum_{n=1}^{N} |z_n|,$$

show that the corresponding inequality for complex integrals over finite integration contours C holds,

$$\left| \int_C f(z)\, dz \right| \leq \int_C |f(z)|\, |dz|\,,$$

so long as the two integrals involved exist.

Hint: use the Riemann definition of the integrals.

2. Starting from the Cauchy integral formulas that give the derivatives of an analytic function $f(z)$ in terms of contour integrals,

$$f^{n\prime}(z_0) = \frac{n!}{2\pi\imath} \oint_C \frac{f(z)}{(z - z_0)^{n+1}}\, dz,$$

where the integration contour C contains z_0, show that the inequalities

$$\left|f^{n\prime}(z_0)\right| \leq \frac{n!}{r_0^n}\, \overline{|f|}_C$$

hold, where the indicated average is taken on a circular contour C of radius r_0 centered at z_0.

3. Consider the Taylor series of the functions $\exp(x)$, $\sin(x)$ and $\cos(x)$, for real x, around $x = 0$.

   (a) Write each one of the series as an infinite sum of a general term.

   (b) Write the series of $\exp(\imath x)$, with x real, separating its real and imaginary parts.

   (c) Identify the real and imaginary parts of the series, and write the resulting relation between the functions involved.

   Answer: $\exp(\imath x) = \cos(x) + \imath \sin(x)$.

4. Consider the function $f(z) = \cos(z)$.

   (a) Expand the function around the point $z = \pi/2$.

   (b) Identify the coefficients of the terms of the resulting series and write the trigonometric identity that follows from them.

   Answer: $\cos(z) = \sin(\pi/2 - z)$.

5. Consider the function $f(z) = \sinh(z)$.

   (a) Expand the function around the point $z = \imath\pi$.

(b) Identify the coefficients of the terms of the resulting series and write the identity between functions that follows from them.

Answer: $\sinh(z) = \sinh(\imath\pi - z)$.

6. **(Challenge Problem)** For each of the following functions, use its Taylor series around $z = 0$ in order to explicitly write the real and imaginary parts of the function in explicit form, in terms of other known functions.

(a) $\sin(z)$.

Answer: $\sin(x)\cosh(y) + \imath\cos(x)\sinh(y)$.

(b) $\cos(z)$.

Answer: $\cos(x)\cosh(y) - \imath\sin(x)\sinh(y)$.

(c) $\sinh(z)$.

Answer: $\sinh(x)\cos(y) + \imath\cosh(x)\sin(y)$.

(d) $\cosh(z)$.

Answer: $\cosh(x)\cos(y) + \imath\sinh(x)\sin(y)$.

Hints: it will be necessary to use the Newton binomial formula, and to figure out how to make changes of summation variables in infinite double sums with two indices.

Chapter 11

Series, Limits and Convergence

In the previous chapter (Chapter 10), starting from an explicit form for the value of an analytic function $f(z)$ at a point z, the Cauchy integral formula, we derived for this function, using the Cauchy integral formulas for the derivatives of $f(z)$, a finite expansion with a polynomial character around a reference point z_0,

$$\begin{aligned} f(z) &= f(z_0) + f'(z_0)(z - z_0) + \frac{f''(z_0)}{2!}(z - z_0)^2 + \\ &\quad + \frac{f^{3\prime}(z_0)}{3!}(z - z_0)^3 + \ldots + \\ &\quad + \frac{f^{(n-2)\prime}(z_0)}{(n-2)!}(z - z_0)^{n-2} + \\ &\quad + \frac{f^{(n-1)\prime}(z_0)}{(n-1)!}(z - z_0)^{n-1} + R_n \\ &= \sum_{i=0}^{n-1} \frac{f^{i\prime}(z_0)}{i!}(z - z_0)^i + R_n, \end{aligned}$$

where n is a positive integer and the remainder R_n is given explicitly by

$$R_n = \frac{(z - z_0)^n}{2\pi\imath} \oint_C \frac{f(z')}{(z' - z)(z' - z_0)^n}\, dz'.$$

We observed that the quantity R_n is indeed the remainder, that is, the error that is left when one adds the first n terms of the series, a fact that can be expressed by the relation

$$R_n = f(z) - \sum_{i=0}^{n-1} \frac{f^{i\prime}(z_0)}{i!} (z - z_0)^i.$$

We will now discuss the question of the *convergence* of the series, that is, of whether or not it is true that, if we add an increasing number of terms of the series, making n increase without limit, the result will approach arbitrarily well the exact result. For this to be true, it is necessary and sufficient that the value of R_n tend to zero when n tends to infinity, as shown by the formula above. Since in our case here we have a closed expression for the remainder, we can easily discuss the question of convergence, by calculating directly this limit,

$$\lim_{n\to\infty} R_n.$$

Since the condition that the limit of a complex number be zero is equivalent to the condition that the limit of its absolute value be zero, we will discuss the convergence in terms of absolute value $|R_n|$, that is, the convergence issue can be reduced to the discussion of the limit

$$\lim_{n\to\infty} |R_n|.$$

We will therefore examine the behavior of $|R_n|$ for large values of n. We first write an expression for the absolute value of the remainder using the triangle inequalities,

$$\begin{aligned} |R_n| &= \frac{|z - z_0|^n}{2\pi} \left| \oint_C \frac{f(z')}{(z' - z)(z' - z_0)^n} \, dz' \right| \\ &\leq \frac{|z - z_0|^n}{2\pi} \oint_C \frac{|f(z')|}{|z' - z|\,|z' - z_0|^n} \, |dz'| \, . \end{aligned}$$

Recalling that the integral has the same value for any contour that includes the points z and z_0, we will now specialize the contour to a circle of radius r_0 centered at z_0, as shown in the diagram of Figure 11.1.

Considering the distances between the various points involved, we see that $|z' - z_0| = r_0$ and we define r such that $|z - z_0| = r$. For $|z' - z|$ we can write, as we did before, that

$$\begin{aligned} (z' - z) &= (z' - z_0) - (z - z_0) \quad \Rightarrow \\ (z' - z_0) &= (z' - z) + (z - z_0) \quad \Rightarrow \\ |z' - z_0| &\leq |z' - z| + |z - z_0|. \end{aligned}$$

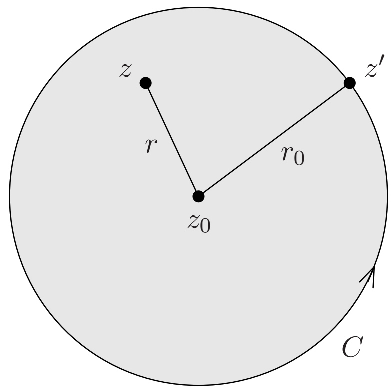


Figure 11.1: The complex plane, showing the circular contour, the three points, the radius r_0 and the vector-difference of z_0 and z, of absolute value r.

This implies that the distance between z' and z can be related to the radii r and r_0, as it is not difficult to see in Figure 11.1, so that we have

$$\begin{aligned} r_0 &\leq |z' - z| + r \quad \Rightarrow \\ |z' - z| &\geq (r_0 - r), \end{aligned}$$

where $(r_0 - r)$ is a positive real quantity. Turning to the expression of $|R_n|$, we can now write it in terms of the variables r_0 and θ on the circular contour, thus obtaining

$$\begin{aligned} |R_n| &\leq \frac{r^n}{2\pi} \oint_C \frac{|\,f(z')\,|}{|z' - z|\, r_0^n} \left|dz'\right| \\ &\leq \frac{r^n}{2\pi r_0^n} \oint_C \frac{|\,f(z')\,|}{|z' - z|}\, r_0\, d\theta, \end{aligned}$$

where we wrote the integration element on the circle of radius r_0 as $|dz'| = r_0\, d\theta$. Since we have that $|z' - z| \geq (r_0 - r)$, we can replace $|z' - z|$ by a smaller quantity in the denominator, without violating the inequality above, thus obtaining

$$\begin{aligned} |R_n| &\leq \frac{1}{2\pi} \left(\frac{r}{r_0}\right)^n \oint_C \frac{r_0}{r_0 - r} \left|\,f\left(z'\right)\,\right| d\theta \\ &= \frac{1}{2\pi} \left(\frac{r}{r_0}\right)^n \frac{r_0}{r_0 - r} \oint_C \left|\,f\left(z'\right)\,\right| d\theta, \end{aligned}$$

where we took out of the integral all quantities that no longer involve the integration variable z'. Since $f(z)$ is analytic and therefore regular throughout the disk within the circular contour and on the contour, we can now limit the remaining integral from above, exchanging $|f(z')|$ by its maximum value f_M on the contour, still without violating the inequality. Thus we can write

$$\begin{aligned} |R_n| &\leq \frac{1}{2\pi}\left(\frac{r}{r_0}\right)^n \frac{r_0}{r_0 - r} f_M \oint_C d\theta \\ &= \left(\frac{r}{r_0}\right)^n \frac{r_0}{r_0 - r} f_M, \end{aligned}$$

where we have used the fact that the value of the remaining integral is 2π. From this result for the remainder R_n we can conclude that, since $r < r_0$, that is, since we are using the series to calculate $f(z)$ on a point which is *strictly internal* to the contour C, the remainder R_n tends to zero as $n \to \infty$, that is,

$$\lim_{n\to\infty} |R_n| = 0,$$

if $r < r_0$, due to the fact that in this case

$$\lim_{n\to\infty} \left(\frac{r}{r_0}\right)^n = 0,$$

while the other factors are fixed numbers in this limit. Note that, the closer the ratio r/r_0 is to 1, the slower the latter limit will approach zero, apart from the fact that the fixed factor $r_0/(r_0 - r)$ will be considerably larger in this case. Thus we see that the closer the ratio r/r_0 is to 1, the slower the initial error will approach zero, which means that the series will converge more slowly. However, so long as $r/r_0 < 1$ the series is in fact convergent.

In this discussion we can get the impression that another way to improve the convergence behavior of the series would be to increase the value of r_0, instead of decreasing the value of r. This may give the false impression that for functions that are analytic throughout the complex plane, in which case we can make r_0 tend to infinity, the error is zero even if we add only a small part of the terms of the series, keeping n finite. However, we can easily verify that this is not the case, by simply observing that the expression of the series does not depend at all on the value of r_0, because there appear in it only the derivatives of the function at z_0 and powers of $z - z_0$.

We can understand this by noting that, although the factors which involve r_0 decrease when we increase r_0, the factor f_M tends to increase with r_0, as we saw earlier, because an analytic function always has the maximum value f_M of its absolute value on the boundary of the disk, so that this value can only increase with r_0. In this way we verify that the behaviors of the factors involving r_0 and f_M are opposite to one another, so that the speed of convergence of the series turns out not to depend at all of r_0, but only on the distance between z and z_0. The range for this distance will be given, in fact, by the maximum convergence radius, which we will discuss momentarily.

In any case, provided that $r/r_0 < 1$, the series converges, that is, as we add more and more terms of the series, the sum will approach the exact result for $f(z)$ with increasing accuracy. Yet another way to express this fact is to say that the sequence of partial sums S_N of the series converges to $f(z)$,

$$\begin{aligned} f(z) &= \lim_{N\to\infty} S_N \\ &= \lim_{N\to\infty} \sum_{n=0}^{N} \frac{f^{n\prime}(z_0)}{n!} (z-z_0)^n. \end{aligned}$$

This limit of the partial sums is a complete criterion for the convergence of the series to the function, which can be used even if we do not know the value of the remainder R_N. The series that we have here is the *Taylor series* of an analytic function, which naturally reduces to the usual real Taylor series for z, z_0 and $f(z)$ real, and that we write symbolically as

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

In the case in which we have $z_0 = 0$, the Taylor series reduces to a particular case named the *Maclaurin series*,

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{n\prime}(0)}{n!} z^n.$$

Thus we see that a Taylor series developed around a point z_0 converges to the function from which it is developed within a circle centered at z_0, provided that the function be analytic in the whole interior of this circle. We can now define without too much difficulty the *radius of convergence* of the series around the point z_0. For this, it is enough to consider the largest circle centered at that point within which the function is still analytic, so that within it the Cauchy integral formula holds. If we increase the radius of the circle until it reaches a singular point of the function in some direction, we will have reached the maximum limit of validity of the proof of convergence of the series. Thus we see that the radius of convergence of the series is the distance from the reference point z_0 to the nearest singularity of the function $f(z)$, which gives us a very easy way to determine a region where the series converges, based on the properties of the analytic function from which we obtain the series. This region is therefore an open disk, the *maximum convergence disk* of the series.

Note that, since we majored several times the expression that gives the remainder of the series, this argument can assure us that the series converges strictly within the disk, but cannot guarantee that it diverges on the circle, or even outside of it. In order to complete the analysis, additional theorems will be required. We will discuss this later on, however we can already mention here that, in general, the situation is that the series converges strictly within the maximum convergence disk, diverges strictly outside of it, and may or may not diverge on the circle which constitutes its boundary. If it does not diverge on the circle, it usually converges very slowly. In addition to this, the further away the point z where we want to calculate the function is from the center z_0 and the closer it is to the boundary of the maximum convergence disk, the slower the convergence of the series.

Note also that it is very easy to decide when and where the Taylor series of a real function converges. It suffices to extend the real function to the complex domain and find the singular point of the resulting analytic function that is closer to the point of the real axis around which the function has been expanded. The intersection of the resulting disk with the real axis gives us immediately the convergence interval of the real series. In particular, we see immediately that the convergence interval is centered at the reference point of the expansion. In general it would be much more difficult to determine the convergence interval without resorting to the extension to the complex plane.

Our convergence analysis was greatly simplified by the fact that we had an explicit expression for the remainder R_n. In other circumstances we will find that the situation may not be that simple. As preliminary work in order to face more complex convergence issues in the future, we will now proceed to a more general and mathematically more precise discussion of the concept of convergence. First of all, we need to recall some convergence concepts of series, or of infinite sums, from real analysis.

The concepts of simple convergence and absolute convergence are concepts that apply point-by-point, and are the same for series of numbers and for series of functions. Of course, a series of functions can always be understood as an infinite set of series of numbers, one for each point on the domain of the functions. Thus, the concept of convergence in its simplest form applies not only to series, but also to simple sequences of real numbers. Consider an infinite but countable set of given real numbers $r_n \in \mathbb{R}$, where $n \in \{0, 1, 2, 3, \ldots, \infty\}$. We say that this set, taken on the given order, is a sequence. The sequence is convergent to a real number r_∞ if r_n approaches r_∞ when n grows without limit. We then write that

$$\lim_{n\to\infty} r_n = r_\infty.$$

In order to put it somewhat more precisely, what this means is that there is a real number, the limit r_∞, such that when n increases without limit, the *distance* between r_n and r_∞ decreases and eventually goes to zero. With complete precision, we can say that this is true if and only if the following condition is satisfied: given an arbitrary distance $\epsilon > 0$, there is always a value of n, which generally depends on ϵ, and that we will denote as $n(\epsilon)$, such that

$$n > n(\epsilon) \Longrightarrow |r_\infty - r_n| < \epsilon.$$

This is the complete and precise convergence criterion. This same concept applies without modification to series, which are sums of sequences of numbers. This is true because we can interpret the convergence of a series in terms of the convergence of a sequence like the one we dealt with above, the sequence of its partial sums. Consider then, once again, the infinite but countable set of real numbers $r_n \in \mathbb{R}$, where $n \in \{0, 1, 2, 3, \ldots, \infty\}$, of which we take the first $N + 1$ in order to build the sum

$$S_N = \sum_{n=0}^{N} r_n.$$

This is a finite sum, which we call a *partial sum* of the series, and therefore it results in a finite real number for each value of N. The set of all values of the partial sums S_N, for $N \in \{0, 1, 2, 3, \ldots, \infty\}$, forms therefore a sequence of real numbers. We say that the series converges to a real number S_∞ if this sequence of sums converges to that number, that is, if it is true that

$$S_\infty = \lim_{N \to \infty} S_N,$$

where the limit exists and is finite, according to the convergence criterion described above. Note that in this case the distance $|S_\infty - S_N|$ between S_∞ and S_N is the error we have when we approximate the limit S_∞ by the partial sum S_N. We call the difference $(S_\infty - S_N)$ the *remainder* of the partial sum S_N, and we denote it by R_N. The convergence criterion is then that, given ϵ, there exists $N(\epsilon)$ such that

$$N > N(\epsilon) \Longrightarrow |R_N| < \epsilon.$$

It is often possible to show that this condition is satisfied, even if an explicit form for the remainder R_N of the series is not available. If we have an explicit form for R_N, as was the case for the Taylor series, we can think of analyzing this condition directly, as we did in that case. One can also write R_N as an infinite sum, that is, as the sum of all terms of the series S_∞ beyond a certain N, and therefore another way of stating the same thing is to say that we have for the remainder R_N associated to the sum S_N that

$$\begin{aligned} \lim_{N \to \infty} R_N &= \lim_{N \to \infty} \sum_{n=N+1}^{\infty} r_n \\ &= 0, \end{aligned}$$

that is, $R_\infty = 0$. We can represent the fact of the convergence of the series symbolically by saying that there is a real number S_∞ such that it is the sum of the infinite series,

$$S_\infty = \sum_{n=0}^{\infty} r_n,$$

which is nothing more than an abbreviation for the limiting process described above.

A second relevant concept is that of absolute convergence, which only applies to series, not to simple sequences. We say that a series

S_∞ converges absolutely if a second series which is associated with it, namely

$$\bar{S}_\infty = \sum_{n=0}^{\infty} |r_n|,$$

is convergent. The central idea here is that in this second series all terms are positive, and therefore that, for that series, the finiteness of the limit of the sum cannot be a consequence of the cancellation of terms with opposite signs. Thus it is clear that this is a more restrictive convergence criterion than the previous one. It can be shown without too much difficulty that absolute convergence implies simple convergence, that is, that the convergence of $\bar{S}_\infty$ implies the convergence of S_∞. It suffices to observe that the absolute value of the remainder R_N associated with each partial sum S_N satisfies

$$\begin{aligned} |R_N| &= \left| \sum_{n=N+1}^{\infty} r_n \right| \\ &\leq \sum_{n=N+1}^{\infty} |r_n| \\ &= \bar{R}_N, \end{aligned}$$

where $\bar{R}_N$ is the remainder associated with the corresponding partial sum of absolute values $\bar{S}_N$, and where we used the triangle inequalities. From this fact it follows that R_N is limited between $\bar{R}_N$ and $-\bar{R}_N$, so that, if we have that $\bar{R}_N \to 0$ when $N \to \infty$, then we also have that $R_N \to 0$ in the same limit. Note that if we have for the convergence condition of $\bar{S}_\infty$ that there exists $N(\epsilon)$ such that $N > N(\epsilon)$ implies that $\bar{R}_N < \epsilon$, it follows immediately that $|R_N| \leq \bar{R}_N < \epsilon$, so that, given the same value of ϵ for the two cases, the same solution $N(\epsilon)$ that works for $\bar{S}_\infty$ can also be used to satisfy the convergence criterion for S_∞.

We see in this way that absolute convergence of a series implies its simple convergence, that is, all absolutely convergent series are also simply convergent. However, the opposite is not true in general, that is, there are series that are convergent but that are not absolutely convergent. Examples of this are, of course, real series whose terms exchange sign, such as

$$S = \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1}.$$

In order to verify that this series converges as it is, we separate it in terms with even n and terms with odd n, bringing the terms together two by two, for the values $2k$ and $2k+1$ of n, so as to rewrite the series as

$$\begin{aligned} S &= \sum_{k=0}^{\infty}\left(\frac{1}{2k+1}-\frac{1}{2k+2}\right) \\ &= \sum_{k=0}^{\infty}\frac{1}{(2k+1)(2k+2)} \\ &< \sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}. \end{aligned}$$

In this form, in which the terms are now all positive, it is not difficult to show that the terms of the sum can be limited from above by a function that decays as $1/k^2$ when k goes to infinity, and that the sum can be limited from above by the asymptotic integral of this function. Since this function has a finite asymptotic integral, it can be seen that the partial sums of the series are monotonically increasing and limited from above, which implies the convergence in accordance with a basic theorem of real analysis. However, an argument of the same type shows that the associated series $\bar{S}$, in which we take the absolute value of each term,

$$\bar{S} = \sum_{n=0}^{\infty}\frac{1}{n+1},$$

is not convergent, but instead diverges to infinity when we take the infinite sum limit, as can be shown by limiting the terms of the sum from below by a function that decays as $1/n$ when n goes to infinity, and by the corresponding lower bound of the limit of the sum by the asymptotic integral of this function. This time the function has a diverging asymptotic integral, showing that the sum cannot have a finite limit. The comparison of series whose terms are all positive with other series that are already known to be convergent or divergent, or with asymptotic integrals that are known to converge or diverge, is a very common way to determine the convergence or divergence of these series.

Observe that the convergence of S depends on the order in which the terms are added. If we first add all the terms with even n this sub-sum alone diverges to ∞. On the other hand, if we first add all terms with odd n, this other sub-sum diverges to $-\infty$. In fact, adding the terms in several different orders, we can get as the limit any real number we want. Therefore we can say that the criterion of absolute convergence is a way to ensure that the sum of the series is not dependent on the order of the sum of the terms. If a series is convergent but not absolutely convergent, then the convergence depends on the order of the sum and, unless expressly stated otherwise, the order is taken as the natural order of the series, the one in which the index of the sum of the terms increases monotonically.

Unlike the concepts of simple convergence and absolute convergence, which are applied point-by-point, the following concept, which is that of uniform convergence, relates to an interval of values of a variable, and applies only to series of functions, not to simple series of numbers. Another way to interpret this is to note that a series of functions is actually a collection of series of numbers, one for each value of the variable on the domain of the functions, and that the criterion of uniform convergence relates to this collection of series. Thus, we must consider series of functions such as power series, that is, we need to generalize our infinite sum for something like

$$S_\infty(x) = \sum_{n=0}^{\infty} f_n(x),$$

for x within some real domain, typically a closed interval of the real line. In this case we have to consider the convergence of the sum for each value of x within this domain. Let us assume that this series is convergent in the entire domain, in which case we say that the series is convergent point-by-point on this domain. We say that the series is uniformly convergent if it converges with the same "speed" at all points of the domain. What we mean by this is that there is no point on the domain at which we have to add a much larger number of terms of the series in order to reach a certain level of accuracy of the estimate obtained. That is, if we want the result of a partial sum of N terms to be at a distance smaller than ϵ of the exact value, we impose on the quantity

$$R_N = S_\infty(x) - S_N(x)$$

the condition that

$$|R_N| < \epsilon,$$

for a given real and positive ϵ, and the series is convergent if it is possible to find a value $N(\epsilon)$ of N above which this condition is satisfied. The series is said to be uniformly convergent if it is true that, for a given arbitrary value of ϵ, one can find a finite value $N(\epsilon)$ of N above which the above inequality is satisfied for *all* values of x within the domain. The point is that the value of $N(\epsilon)$ should not depend on x. It is clear that if a series is uniformly convergent, then it is convergent, but the opposite is not always true, as we shall see later in this series of books when we study Fourier series.

The central point of all this for our purposes here is that, for us to be able to differentiate term-by-term, with respect to x, a series of functions such as $S_\infty(x)$, and thus obtain from it another series that is also convergent, it is always necessary that the original series be uniformly convergent. In the case of power series, this condition is necessary and sufficient. In other cases, such as the Fourier series that we will study in a further volume of this series of books, it is necessary but not sufficient in the case of the differentiation. On the other hand, uniform convergence in closed domains is always a sufficient condition for us to be able to integrate the series term-by-term, and thereby obtain other convergent series. Moreover, uniform convergence is also a necessary and sufficient condition for $S_\infty(x)$ to be a continuous function of x within the domain, if $f_n(x)$ are continuous functions. These facts about the continuity of the limit, the term-by-term differentiability and term-by-term integrability of the series are theorems of real analysis, which will be used frequently, but that we will not discuss here in greater detail.

All these convergence concepts can be generalized without difficulty to the complex context. The simple convergence of a series of complex numbers

$$S_\infty = \sum_{n=0}^{\infty} z_n,$$

where $z_n = x_n + \imath y_n$ and $S_\infty = X_\infty + \imath Y_\infty$, is associated with the simple convergence of the two real series

$$\begin{aligned} X_\infty &= \sum_{n=0}^{\infty} x_n, \\ Y_\infty &= \sum_{n=0}^{\infty} y_n. \end{aligned}$$

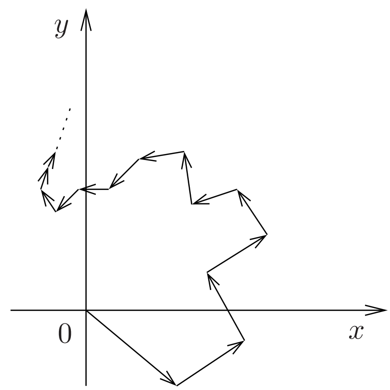


Figure 11.2: A string of concatenated vectors in the complex plane, representing their sum starting at the origin.

The absolute convergence of the series S_∞ now means that the series of absolute values of complex numbers

$$\bar{S}_\infty = \sum_{n=0}^{\infty} |z_n|$$

converges. Note however that this is always a real series, even if S_∞ is a complex series. The uniform convergence, for example of a complex powers series,

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

for z within any given domain, has the same meaning as before, where z now belongs to a two-dimensional domain in the complex plane, typically a closed disk instead of a closed real interval.

It is interesting and useful to note that there is in the complex plane a geometric interpretation for the convergence of a complex series. Since each term of the sum is a complex number and can therefore be represented as a vector in the complex plane, a partial sum S_N can be represented as N vectors of this type linked together, forming a chain of segments that propagates along the plane, as illustrated in the diagram of Figure 11.2.

The complete series is represented by a string of this type that consists of an infinite but countable and ordered set of chained vectors, according to the natural order defined by the index of the sum. Each

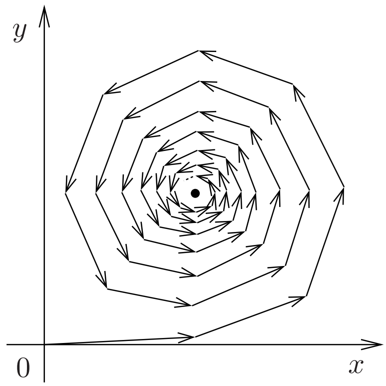


Figure 11.3: A chain of infinite length, possibly converging in spiral to a point of the complex plane, illustrating the corresponding oscillations of the coordinates in both axes.

partial sum is represented as a point in the plane, which is one of the successive vertices of this polygonal chain. The convergence of the series is thus represented by the convergence of the sequence of vertices of the chain to some fixed point of the complex plane. The chain may have either a finite or an infinite total length, but may converge even if the total length is infinite, as suggested by the diagram in Figure 11.3.

Since the absolute value of each vector is the length of the corresponding segment of the chain, the sum of the absolute values of the vectors in a partial sum S_N gives us the total length of the corresponding chain. However, this sum of absolute values is exactly the partial sum $\bar{S}_N$ of the absolute values of the terms of the corresponding series, which appears in the absolute-convergence criterion.

We therefore have a very simple interpretation of the concept of absolute convergence: a series being absolutely convergent is equivalent to its chain in the complex plane having finite total length. In this case it is clear that the series converges because the chain can only spread around the plane for a finite length, and must have its endpoint on some fixed point of the complex plane. On the other hand, if the series is not absolutely convergent, then its chain has infinite total length, and may diverge by extending indefinitely to infinity, or by oscillating and circling indefinitely along the complex plane as shown in Figure 11.3, in the case in which the spiral does not converge to a fixed point, but instead circles indefinitely around a limiting closed curve.

Problem Set

1. Consider an arbitrary series whose terms are all positive real numbers, such as

$$S_\infty = \sum_{k=0}^{\infty} |a_k|.$$

   Show that, if this series is limited from above, that is, if there is a positive real number S_M such that $S_\infty \leq S_M$, then the series necessarily converges to some positive real number less than or equal to S_M.

   Hint: the statement $S_\infty \leq S_M$ actually means that all the partial sums S_N of this series are limited in this way, for all N.

2. Show that the real numerical series of positive terms

$$S_\infty = \sum_{k=0}^{\infty} \frac{1}{(2k+1)(2k+2)}$$

   converges to a finite value, limiting it from above by an asymptotic integral of the real function $1/k^2$, such an integral being finite. Recall that we can ignore some initial terms of the sum without affecting the convergence of the series, if this is necessary for the argument.

3. Show that the real numerical series of positive terms

$$S_\infty = \sum_{n=0}^{\infty} \frac{1}{n+1}$$

   diverges to positive infinity, limiting it from below by as asymptotic integral of the real function $1/n$, such an integral being infinite. Recall that we can ignore some initial terms of the sum without affecting the convergence of the series, if this is necessary for the argument.

4. Consider the function $f(z) = 1/(1+z)$.

(a) Expand the function $f(z)$ in a power series around $z = 0$.

(b) Identify the convergence disk of the power series.

(c) Change to the variable $w = 1 + z$ and write a power series for the function $g(w) = 1/w$.

(d) Identify the convergence disk in terms of w.

(e) Show that this power series is a Taylor series of $g(w)$ and identify the point z_0 around which it is developed.

5. Consider the function $w(z) = 2/\sqrt{4 - z}$, where $z = x + \imath y$ is a complex number, while x and y are real numbers.

   (a) Write the Taylor expansion of $w(z)$ around $z = 0$, that is, the Maclaurin expansion of the function. Write explicitly the general term of the series.

   (b) Determine and describe the interval of convergence of the series on the real x axis, that is, for $y = 0$.

   (c) Use this series in order to define a method for calculating the irrational number $\sqrt{2}$ via a limit involving only rational numbers. Show that each term of the series is a rational number.

   (d) Using this series, identify an infinite *sequence* of rational numbers that converges to $\sqrt{2}$.

6. Consider the function $w(z) = 3/\sqrt{9 - z}$, where $z = x + \imath y$ is a complex number, while x and y are real numbers.

   (a) Write the Taylor expansion of $w(z)$ around $z = 0$, that is, the Maclaurin expansion of the function. Write explicitly the general term of the series.

   (b) Determine and describe the interval of convergence of the series on the real x axis, that is, for $y = 0$.

   (c) Use this series in order to define a method for calculating the irrational number $\sqrt{3}$ via a limit involving only rational numbers. Show that each term of the series is a rational number.

   (d) Using this series, identify an infinite *sequence* of rational numbers that converges to $\sqrt{3}$.

7. Consider the function $w(z) = 3/\sqrt{4 - z}$, where $z = x + \imath y$ is a complex number, while x and y are real numbers.

   (a) Write the Taylor expansion of $w(z)$ around $z = 0$, that is, the Maclaurin expansion of the function. Write explicitly the general term of the series.

   (b) Determine and describe the interval of convergence of the series on the real x axis, that is, for $y = 0$.

   (c) Use this series in order to define a method for calculating the irrational number $\sqrt{3}$ via a limit involving only rational numbers. Show that each term of the series is a rational number.

   (d) Using this series, identify an infinite *sequence* of rational numbers that converges to $\sqrt{3}$.

8. Consider the function $w(z) = 5/\sqrt{2^2 + z}$, where $z = x + \imath y$ is a complex number, while x and y are real numbers.

   (a) Calculate the Taylor expansion $w(z)$ around $z = 0$, that is, the Maclaurin expansion of the function. Write explicitly the general term of the series.

   (b) Determine and describe the interval of convergence of the series on the real x axis, that is, for $y = 0$.

   (c) Use this series in order to define a method for calculating the irrational number $\sqrt{5}$ via a limit involving only rational numbers. Show that each term of the series is a rational number.

   (d) Using this series, identify an infinite *sequence* of rational numbers that converges to $\sqrt{5}$.

Chapter 12

Representation of Functions by Series

As we have seen before, given an analytic function in a region of the complex plane around a point of reference z_0, a power series is uniquely determined, the Taylor series of the function around that point. Moreover, given the Taylor series of this analytic function $f(z)$ around the point z_0,

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

the series converges to the function, within its radius of convergence, which is determined by the singularity of $f(z)$ which is closer to z_0. Thus, it is reasonable to think of these series as faithful representations of their functions. This is very important, because what the series give us is an *algorithmic* representation of the functions, based on sequences of elementary arithmetic operations. It is only through algorithmic representations like this that we can gain control over these functions in the practice of applications.

These sequences of operations are generally infinite, which means that in practice they cannot be fully implemented. However, if the series converges, this means that the error that is left when we add a finite number of terms goes to zero, when the number of added terms goes to infinity. Therefore, if we add a sufficient number of terms of the series, possibly large but finite, we will have an approximation with a level of accuracy which is sufficient for its utilization in physics, because in practice we always have a finite level of accuracy, both in the physical measurements and in our ability to describe the results, as well as in our ability to theoretically predict these results.

A mathematically more complete analysis of the relationship between power series and analytic functions involves a more complete and thorough discussion of the convergence of a power series in the complex context. Such a mathematically more precise discussion will be postponed to the next chapter. For now, we will verify that it is possible to manipulate these series, just as we would do with the corresponding functions, so that we can in fact use them in practice to represent the functions. When it becomes necessary to use the convergence properties of the series during this development, we will draw attention to the fact and will state the necessary theorems without proof, leaving the discussion of some of these proofs to the next chapter.

We will consider here several series, all related to the same reference point z_0, and all restricted to the domain of convergence common to all of them. Since each one has a certain disk of convergence, we will always be within that with the smallest radius of convergence among the various series being discussed, so that all of them are convergent within one and the same disk. For example, it is easy to see that the sum of two Taylor series, of two particular functions, is the series of the sum of the two functions, because if we have

$$\begin{aligned} f_1(z) &= \sum_{n=0}^{\infty} \frac{f_1^{n\prime}(z_0)}{n!}\,(z-z_0)^n, \\ f_2(z) &= \sum_{n=0}^{\infty} \frac{f_2^{n\prime}(z_0)}{n!}\,(z-z_0)^n, \end{aligned}$$

it suffices to add the two equations and collect terms in order to have

$$f_1(z) + f_2(z) = \sum_{n=0}^{\infty} \frac{f_1^{n\prime}(z_0) + f_2^{n\prime}(z_0)}{n!}\,(z-z_0)^n.$$

Since the multiple differentiation is a linear operation,

$$\frac{d^n}{dz^n}(f_1+f_2)(z) = f_1^{n\prime}(z) + f_2^{n\prime}(z),$$

the result follows,

$$(f_1+f_2)(z) = \sum_{n=0}^{\infty} \frac{(f_1+f_2)^{n\prime}(z_0)}{n!}\,(z-z_0)^n.$$

We can therefore add the series of two functions in order to obtain as the result the series of the sum of the two functions. Still due to the linearity of the differentiation operation we can also multiply each of the two series by arbitrary complex numbers before adding them. Thus, we can use the series to represent arbitrary linear combinations of their respective functions and, in particular, to represent the difference of two functions instead of their sum. Thus we see that the representation of the function by the series is faithful with respect to one of the two operations of a field, the addition operation, as well as with respect to its inverse operation.

Recalling that the derivative of an analytic function is also an analytic function, it is also easy to verify that we can differentiate a Taylor series term-by-term, thereby obtaining the series of the derivative of the function. It is important to observe that this term-by-term differentiation is not possible in general, because in principle it can result in a series that does not converge, even if the original series is convergent. However, we will verify below that for the Taylor series of analytic functions, this always works and in fact generates the correct Taylor series of the derivative of the original function, which we already know to be convergent, since the derivative of the analytic function is another analytic function in the same domain. Let us then take the term-by-term derivative,

$$\begin{aligned} f(z) &= \sum_{n=0}^{\infty} \frac{f^{n\prime}(z_0)}{n!}(z-z_0)^n \quad \Rightarrow \\ \frac{df}{dz}(z) &= \sum_{n=0}^{\infty} \frac{f^{n\prime}(z_0)}{n!}\, n(z-z_0)^{n-1}, \end{aligned}$$

where, due to the factor of n, the term $n = 0$ vanishes. Therefore, we can reduce the variation interval of n and then change the variable of the sum from n to $m = n - 1$, in order to obtain

$$\begin{aligned} \frac{df}{dz}(z) &= \sum_{n=1}^{\infty} \frac{f^{n\prime}(z_0)}{(n-1)!}(z-z_0)^{n-1} \\ &= \sum_{m=0}^{\infty} \frac{f^{(m+1)\prime}(z_0)}{m!}(z-z_0)^m. \end{aligned}$$

Since $f^{(m+1)\prime}(z)$ is the m-th derivative of df/dz, and changing again the name of the summation variable from m to n, we obtain

$$\frac{df}{dz}(z) = \sum_{n=0}^{\infty} \frac{(df/dz)^{n\prime}(z_0)}{n!} (z - z_0)^n,$$

that is, we have in fact the Taylor series of df/dz, which, as previously proved, converges to this analytic function. It is important to note that it was not necessary to assume that this new series converges based on the convergence of the original series. We simply recognize the resulting series, whose convergence has been proved before. Note also that, since the derivative of an analytic function always exists and is itself analytic, throughout the analyticity domain of that function, it follows that the resulting series has the same radius of convergence of the original series.

Similarly, and recalling that we have already shown that the primitive of an analytic function is also analytic, we can try to make the term-by-term indefinite integration of a series in order to obtain the series of the primitive. The same observations about the preservation of the convergence of the series in such an operation, already mentioned in the case of the term-by-term derivation, also apply in this case. Doing then the term-by-term integration we have

$$\begin{aligned} f(z) &= \sum_{n=0}^{\infty} \frac{f^{n\prime}(z_0)}{n!} (z - z_0)^n \quad \Rightarrow \\ F(z) &= \int^z f(z')\, dz' \\ &= C_0 + \sum_{n=0}^{\infty} \frac{f^{n\prime}(z_0)}{n!(n+1)} (z - z_0)^{n+1}, \end{aligned}$$

where C_0 is an arbitrary complex constant, and $n!(n + 1) = (n + 1)!$. Making the transformation $m = n + 1$ of the summation variable, we obtain

$$F(z) = C_0 + \sum_{m=1}^{\infty} \frac{f^{(m-1)\prime}(z_0)}{m!} (z - z_0)^m.$$

As we have shown before, $F'(z) = f(z)$, so that we have that $f^{(m-1)\prime}(z) = F^{m\prime}(z)$ is the m-th derivative of the function $F(z)$ and, since it is immediate to verify that $C_0 = F(z_0)$, we get for this series, changing once again the name of the new summation variable from m back to n,

$$F(z) = \sum_{n=0}^{\infty} \frac{F^{n\prime}(z_0)}{n!} (z - z_0)^n.$$

Here we recognize once again, in its usual form, the Taylor series of the analytic function $F(z)$, which converges to this function, since it is analytic, as we have shown before. The convergence radius, once again, is the same as that of the original series, for the same reasons explained in the case of the term-by-term differentiation.

The fact that, as we have shown here, we can integrate and differentiate the Taylor series term-by-term, thereby obtaining the correct Taylor series of the primitive and of the derivative of the original function, all of which are already known to be convergent to their respective functions, shows that the convergence of these series is in fact stronger than it looks, because it is in fact *uniform convergence*, rather than simple convergence. As we mentioned in the previous chapter (Chapter 11), uniform convergence is always a necessary condition for us to be able to differentiate a series term-by-term and obtain as a result another equally convergent series. Since for the case of Taylor series this always happens, as we have just proven, it follows that these series are always uniformly convergent on closed disks contained within the maximum convergence disk.

Finally, we will discuss the question of the product and of the division of two series, which is a bit more complicated to analyze. For the case of the product of two series, we start from

$$\begin{aligned} f_1(z) &= \sum_{n=0}^{\infty} \frac{f_1^{n\prime}(z_0)}{n!} (z - z_0)^n, \\ f_2(z) &= \sum_{n=0}^{\infty} \frac{f_2^{n\prime}(z_0)}{n!} (z - z_0)^n, \end{aligned}$$

so that we have for the product of the two functions

$$f_1(z) f_2(z) = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \frac{f_1^{k\prime}(z_0) f_2^{m\prime}(z_0)}{k!\, m!} (z - z_0)^{k+m}.$$

In order to write this as an ordered series of powers, we have to collect all the terms for which the power $k + m = n$ is a given constant. Let us write only a few terms, in order to exemplify the results of such a collecting of terms,

$$\begin{aligned} f_1(z)f_2(z) \quad = \quad & f_1(z_0)f_2(z_0) \\ & + \frac{f_1(z_0)f_2'(z_0) + f_1'(z_0)f_2(z_0)}{1!}(z - z_0)^1 \\ & + \frac{f_1(z_0)f_2''(z_0) + 2f_1'(z_0)f_2'(z_0) + f_1''(z_0)f_2(z_0)}{2!}(z - z_0)^2 \\ & + \ldots \; . \end{aligned}$$

We can thus verify that the numerators of the coefficients appearing are those that result from the Leibniz rule for the derivative of a product, applied multiple times,

$$\begin{aligned} (f_1f_2)(z) &= f_1(z)f_2(z), \\ (f_1f_2)'(z) &= f_1(z)f_2'(z) + f_1'(z)f_2(z), \\ (f_1f_2)''(z) &= f_1(z)f_2''(z) + 2f_1'(z)f_2'(z) + f_1''(z)f_2(z) \\ \ldots &= \ldots \; , \end{aligned}$$

and so on. Thus we see that the coefficients that appear are in fact the derivatives of the product of the two functions, and we can *induce* the general result,

$$(f_1f_2)(z) = \sum_{n=0}^{\infty} \frac{(f_1f_2)^{n\prime}(z_0)}{n!}(z - z_0)^n.$$

This argument can be generalized from the first few terms of the series, which we discussed here, to the general case of the n-th term, through a process of induction, including a complete proof by finite induction, which will be left as an exercise.

The case of the division is an extension of the case of the multiplication, but still a little more complicated to analyze. In this case we must also worry about the possibility of having zeros in the denominator. If we consider the division as the construction of the function $g(z) = f_2(z)/f_1(z)$, eventual zeros of $f_1(z)$ within the convergence domain of the series will produce singularities in $g(z)$, so that we cannot expect that the radius of convergence of the series will be maintained in this case. Assuming that $f_1(z)$ does not have any zeros within the region of interest around z_0, we can define the function $g(z)$,

$$g(z) = \frac{f_2(z)}{f_1(z)},$$

which is such that

$$f_2(z) = f_1(z)g(z),$$

that is, we can write for $g(z)$ a tentative Taylor series

$$g(z) = \sum_{n=0}^{\infty} \frac{g^{n\prime}(z_0)}{n!}\,(z - z_0)^n,$$

so that we have in this case the set of relations between the coefficients of the series given by

$$\sum_{n=0}^{\infty} \frac{f_2^{n\prime}(z_0)}{n!}\,(z - z_0)^n = \sum_{k=0}^{\infty}\sum_{m=0}^{\infty} \frac{f_1^{k\prime}(z_0)g^{m\prime}(z_0)}{k!m!}\,(z - z_0)^{k+m}.$$

We will now assume that, for this equation to be true, it is necessary that the coefficients of each power, in each of the two sides of the equation, be equal. While it is clear that this is indeed a possible solution of the problem, it is *not* clear that it is the *only* one. The uniqueness is a consequence of the fact that the analytic functions we are discussing here can be understood as a vector space, within which the set of all the integer powers is a basis, thus consisting of linearly independent elements. These facts and concepts will be discussed later on. For now, it is necessary to collect terms from the right-hand side, in such a way as to gather all the terms in which $n = k + m$ is constant, just as we did in the case of the product, and then to solve recursively for the unknown coefficients $g^{m\prime}(z_0)$. We will do only the first two cases as examples. For $n = 0$ we simply have that

$$\begin{aligned} f_2(z_0) &= f_1(z_0)g(z_0) \quad \Rightarrow \\ g(z_0) &= \frac{f_2(z_0)}{f_1(z_0)}, \end{aligned}$$

so long as $f_1(z_0) \neq 0$, which is guaranteed by our hypothesis that $f_1(z)$ does not have any zeros within the region of interest around the point z_0. Having the result of this case, we can do the case $n = 1$,

$$\begin{aligned} f_2'(z_0) &= f_1'(z_0)g(z_0) + f_1(z_0)g'(z_0) \\ &= f_1'(z_0)\frac{f_2(z_0)}{f_1(z_0)} + f_1(z_0)g'(z_0) \quad \Rightarrow \\ \frac{f_2'(z_0)}{f_1(z_0)} &= f_1'(z_0)\frac{f_2(z_0)}{[f_1(z_0)]^2} + g'(z_0) \quad \Rightarrow \\ g'(z_0) &= \frac{f_1(z_0)f_2'(z_0) - f_1'(z_0)f_2(z_0)}{[f_1(z_0)]^2}, \end{aligned}$$

that is the correct expression for the derivative of the ratio of the two functions at the point z_0. As in the case of the multiplication, we can continue to do this kind of analysis indefinitely, for increasing values of n, with the expected results, and with increasing algebraic work. For the division we induce in this way the expected result,

$$\left(\frac{f_2}{f_1}\right)(z) = \sum_{n=0}^{\infty} \frac{(f_2/f_1)^{n\prime}(z_0)}{n!}(z - z_0)^n.$$

The full proof of this result is considerably more complex and difficult than that of the case of the multiplication. Since division is the product of one function by the inverse of the other, we can reduce this problem to the problem of the product, which we discussed above, and to the problem of finding the power series of the multiplicative inverse of a function. We will leave the discussion of this case as an exercise, in the form of a challenge problem for which we do not really have a truly complete solution available.

In summary, we have shown so far that each analytic function is associated with a power series, its Taylor series around a given point of reference, that represents it in a faithful way, within a certain convergence disk. To put it more precisely, each analytic function is associated to an infinite *set* of power series, each one associated with a point z_0 at which the function is analytic, and it is the set of all these power series that, in fact, represents the function completely. We can imagine this set of power series, each one with its convergence disk, as a way to propagate the definition of the function from one point to another in the complex plane.

This process of propagation of the definition of an analytic function that promotes the extension of the definition of the function, in an analytic way, from one region to another, is called a process of *analytic continuation*. Given the Taylor series of a certain function around a point z_0, this series determines the function and all its infinite derivatives at a point z_1 that is within its convergence disk. It follows that, by using these values, we can build around z_1 a second Taylor series of the same function. In general, the convergence disk of this second series is not contained within the convergence disk of the first one, and therefore constitutes an extension of the definition of the function to a region that goes beyond the region of validity of the definition by the first series. Continuing this process indefinitely, we end up covering an entire connected region in which the function is analytic.

With what we have seen here it is already clear that, in the common region of convergence of two series within this set of series, their limits coincide, because both series are Taylor series of the same analytic function, and therefore both, if convergent at all, converge to that same function. This idea can be generalized to the case of two series whose limits coincide within some open region, even if the reference point of each one of them is not within the convergence disk of the other. There is a theorem, called the *theorem of analytic continuation*, which ensures that, whatever the way in which one extends the definition of an analytic function, the result is always unique, and always represents the same analytic function. We have seen a different example of this sort of thing when we discussed the function $\Gamma(z)$ in a previous chapter (Chapter 4).

It is interesting to observe that the uniqueness of the analytic continuation of a function can also be understood in terms of the theorem of uniqueness of the solution of a partial differential equation, given certain boundary conditions. This is true because, as we have seen before, both the real part $u(x, y)$ and the imaginary part $v(x, y)$ of an analytic function satisfy the Laplace equation. Given a region where the function is defined, for example by means of a power series, it follows that we know the value of this function and of its derivatives on the boundary of the region. Thus, we can integrate the differential equation starting from this boundary, using the boundary conditions, and thus obtain, in a unique way, the values of $u(x, y)$ and $v(x, y)$ in areas contiguous to the initially given area. In this way we make contact here with a subject which we will examine in detail later on in this series of books, that of the solution of these partial differential equations.

Having established the relation of each analytic function with its Taylor series, let us now address the inverse problem, that is, the problem of determining, given a certain power series, whether there is an analytic function that corresponds to it, and in what circumstances this happens. A mathematically complete discussion of this problem involves the use of a number of theorems about the convergence of power series. These theorems will be used here without proof, and a more detailed discussion of some of them will be made in the next chapter. We will limit ourselves here to define quickly and intuitively the mathematical concepts involved, and to use the theorems to establish the facts that are relevant for our purposes. In the next chapter there will be a more detailed discussion of the mathematical facts involved, and some of the theorems will be proven, which will require the use of a little more abstract and formal methods of proof, involving the famous ϵ's and δ's.

We have used before some basic theorems of real analysis, which we will often use again in the future, in particular for the proofs contained in the next chapter. For reference, we will list here the real analysis theorems that we will be using without proof. Simple proofs of some of these mathematical facts are available in [8] (available in Appendix D). We will assume a certain familiarity of the reader with the contents of these theorems, all of which have an intuitively transparent content, in a fairly immediately way.

1. A monotonically increasing (or decreasing) real sequence which is bounded from above (or below) is necessarily convergent.

   Because of this, a series of positive terms that is bounded from above is necessarily convergent.

2. If a series is convergent, then the limit of its terms when the summation index n tends to infinity is necessarily zero. In other words, the sequence of terms of the series converges to zero.

   Associated with this, if a sequence is convergent, then it is bounded from above and from below.

3. The uniform convergence of a series in closed domains is a sufficient condition for one to be able to integrate the series term-by-term, and a necessary condition for one to be able to differentiate the series term-by-term.

   In the particular case of power series, the condition becomes necessary and sufficient for both differentiation and integration.

4. The uniform convergence of series of continuous functions is a necessary and sufficient condition for the series to converge to a continuous function.

At this point we need to simply quote the fundamental mathematical facts that are involved and which we need in order to be able to analyze the problem posed above. The mathematical concept involved is that of uniform convergence. As we discussed in the previous chapter (Chapter 11), what this means intuitively is that the series converge basically at the same speed at all points of the convergence disk, that is, that there is no point where it converges so much slower than at other points as to force us to add many more terms in order to reach a certain accuracy of the result.

The important point for us here is that uniform convergence is a necessary and sufficient condition to ensure that one can integrate and differentiate a convergent power series without spoiling its convergence properties, that is, thereby producing two other series that are also convergent. More generally, that is, not specifically for power series, uniform convergence is a necessary but not sufficient condition for one to be able to differentiate term-by-term convergent series, and is also a sufficient but not necessary condition for one to be able to integrate term-by-term convergent series, thereby producing other convergent series.

The relevant property of the series of complex powers, without reference to their being a Taylor series of some function, is that under very weak hypotheses, it can be shown that they are always uniformly convergent within a certain disk. In order to show this, first one shows that the series are always absolutely convergent inside this disk. As we discussed in the previous chapter (Chapter 11), the absolute convergence is the convergence of the series formed by taking the absolute values, that is, real or complex absolute values as applicable, of each term of the original series. Absolute convergence always implies the simple convergence of the original series, regardless of the type of series that we discuss, but in the case of power series it also implies the uniform convergence of the original series. Making explicit the facts, given an arbitrary complex power series,

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

written in terms of a reference point z_0, with the only condition that the given coefficients a_n be such that the series converges at a point $z_1 \neq z_0$, that is, so long as the complex numerical series

$$S_\infty(z_1) = \sum_{n=0}^{\infty} a_n (z_1 - z_0)^n$$

exists and is finite, it can be shown that the series converges uniformly within the disk centered at z_0 with radius $|z_1 - z_0|$, that is, for the entire region where z is such that $|z - z_0| < |z_1 - z_0|$. In fact, since uniform convergence relates not to points but to regions, and is only truly useful when it holds in closed domains (including their boundaries), the more precise statement is that the series is uniformly convergent in any closed disk strictly contained within the convergence disk, such that

$|z - z_0| < |z_1 - z_0|$. Note that this is relevant because the series may not even be convergent at all at some point on the boundary of the maximum convergence disk.

Thus, given an arbitrary power series that satisfies only a very weak hypothesis, that it be convergent at two different points, since the uniform convergence is a sufficient condition to allow its term-by-term integration, we know that we can integrate the series term-by-term and that the resulting series is also convergent. The two points at issue are z_0 itself, where we have trivially that $S_\infty = a_0$, and the point z_1 which was given in the hypothesis. With this established, it is very easy to finish the analysis of our problem. Consider therefore the complex function defined, at the points where it converges, by the power series

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

that converges within the open disk with non-zero radius $|z_1 - z_0|$. We will now integrate this complex function in a closed contour C, arbitrary except for the fact that it must be contained strictly within this open disk. On the right-hand term, we will integrate the series term-by-term, since it is uniformly convergent, in order to obtain

$$\oint_C S_\infty(z)\, dz = \sum_{n=0}^{\infty} a_n \oint_C (z - z_0)^n dz.$$

Because the functions $(z-z_0)^n$ are analytic on the whole complex plane, for all $n \geq 0$, by the Cauchy-Goursat theorem the right-hand side of the equation vanishes, and therefore we have that

$$\oint_C S_\infty(z)\, dz = 0,$$

for all contours C within the disk $|z - z_0| < |z_1 - z_0|$. Given this fact, then by the Morera theorem, which is the reverse of the Cauchy-Goursat theorem, it follows that the function $S_\infty(z)$ is analytic inside the disk. Thus we see that the given series in fact represents, strictly within the disk defined by z_0 and z_1, an analytic function.

We can now also show that the given series is in fact the Taylor series of that analytic function around the point z_0, that is, that the representation of an analytic function by a power series around a given point z_0 is in fact *unique*. Since it was shown that the function is analytic, it follows that it has derivatives of all orders, and therefore we can differentiate the function repeated times, differentiating the series term-by-term repeatedly, in order to obtain the collection of series

$$\begin{aligned}
S_\infty(z) &= \sum_{n=0}^{\infty} a_n(z-z_0)^n, \\
S'_\infty(z) &= \sum_{n=1}^{\infty} na_n(z-z_0)^{n-1}, \\
S''_\infty(z) &= \sum_{n=2}^{\infty} n(n-1)a_n(z-z_0)^{n-2}, \\
\ldots &\quad \ldots, \\
S^{k\prime}_\infty(z) &= \sum_{n=k}^{\infty} n(n-1)\ldots(n-k+2)(n-k+1)a_n(z-z_0)^{n-k}, \\
\ldots &\quad \ldots.
\end{aligned}$$

Applying these relations at the point $z = z_0$, only the first term of each one of the sums survives, thus showing once again that all of these series are in fact convergent at this point, so that we get

$$\begin{aligned}
S_\infty(z_0) &= a_0, \\
S'_\infty(z_0) &= a_1, \\
S''_\infty(z_0) &= 2a_2, \\
\ldots &\quad \ldots, \\
S^{k\prime}_\infty(z_0) &= k!a_k \\
\ldots &\quad \ldots,
\end{aligned}$$

and we can then solve for the coefficients a_n, obtaining therefore for the n-th coefficient

$$a_n = \frac{S^{n\prime}_\infty(z_0)}{n!},$$

that is, we have in fact the coefficients of the Taylor series of the function around z_0, thus showing the uniqueness of this series for this function, with this point of reference, since we see here that, if a power series converges to an analytic function, then it is the Taylor series of that analytic function, around the point of reference.

Thus we see that there is a one-to-one relation between the complex functions that are analytic around the point z_0 and the convergent power series around that point. Given a function, a corresponding series is determined, and given a series that is convergent at least at one point besides z_0, a corresponding analytic function is determined. Thus, we see that in fact the power series are faithful and complete representations of the analytic functions within their convergence disks. Note that the uniqueness of the power series that represents an analytic function justifies, a posteriori, the procedure of matching coefficients of each power, which we used before in this chapter in the discussion of the case of the division of two Taylor series.

Armed with these new facts, we can now think of expanding our knowledge of the convergence of complex powers series. First of all, given a series defined by z_0 and convergent at z_1, and considering that it is the Taylor series of a function within the disk $|z - z_0| < |z_1 - z_0|$, we can now extend this disk to the maximum disk of convergence of the Taylor series. In order to do this, we simply increase the radius of the disk until its boundary reaches a singular point of the function in some direction in the complex plane. We have therefore the guarantee of convergence of the series to the function, for all the points that are strictly within this maximum convergence disk.

That done, we can now ensure that the series is *divergent* at any point strictly *outside* of the maximum convergence disk. The proof of this is simple, because if the series did converge at a point z_2 external to the disk, then we have a power series that converges in z_0 and at a second point, in this case z_2. We know, therefore, that in accordance with the above theorems, the series converges in the entire disk $|z-z_0| < |z_2-z_0|$, which is greater than the maximum convergence disk, and contains it. But this implies that the singular point, which is on the boundary of the maximum convergence disk, is contained within this larger disk, and that the series, which is the Taylor series of the function, converges to the function at this singular point, which is absurd, since a power series which is convergent at a given point always converges to a function which is analytic at that point, and therefore *not* singular. It follows that we can ensure that the series is *not* convergent at any point strictly external to the maximum convergence disk.

We will finish with some remarks on the interpretation of the set of integer power functions as a *basis* of the space of analytic functions. As we have seen, given a reference point z_0 on the complex plane, and the set of all functions that are analytic at that point, the infinite but discrete set of power functions $w_n(z) = (z - z_0)^n$, with $n \in \{0, 1, 2, 3, \dots\}$, can be used to generate series which converge to each one of these functions, on some neighborhood of that point. As we have seen, there is a one-to-one relation between analytic functions and convergent power series centered at that point, every analytic function corresponds to a single series and every series to a unique analytic function. In the spirit of the use of bases within a vector space, we can think that each one of these series is the expansion of the corresponding analytic function on the basis formed by the particular collection of analytic functions $w_n(z)$.

Since the sum of analytic functions is analytic and the product of an analytic function with a complex constant is also analytic, it follows that the set of analytic functions can indeed be understood as a vector space, with the identically zero function playing the role of the null vector. Since the number of basis elements is infinite, this is a vector space of infinite dimension. However, just as is the case of a usual finite-dimensional vector space, given an analytic function there exists one and only one linear combination of elements of the basis, that is, a power series, that represents that function. Furthermore, given a linear combination of any elements of the basis, that is, a power series, which has the property to converge at least at one point besides z_0, there is one and only one analytic function which is represented by it. As we shall see, this interpretation of this type of structure in terms of a vector space is very powerful and useful, and of very widespread use in physics.

Problem Set

1. Consider the Taylor series of the function $w(z) = \exp(z)$ around $z_0 = 0$.

   (a) Differentiate the series term by term and show that it remains invariant.

   (b) Integrate the series term by term and show that it remains invariant except for a complex constant.

2. Consider the Taylor series of the function $w(z) = \sin(z)$ around $z_0 = 0$.

   (a) Differentiate the series term by term and show that the resulting series is that of the function $\cos(z)$.

(b) Integrate the series term by term and show that, except for a complex constant, the resulting series is that of the function $-\cos(z)$.

3. Consider the function $f(z) = 1/(1-z)$.

   (a) Write the Maclaurin series of $f(z)$ as the sum of a general term.
   (b) Determine the convergence disk of this power series.
   (c) Differentiate the power series in order to obtain a representation of the function $1/(1-z)^2$.
   (d) Differentiate the power series once again in order to obtain a representation of the function $1/(1-z)^3$.

4. Consider the function $f(z) = 1/(1+z)$.

   (a) Write the Maclaurin series of $f(z)$ as the sum of a general term.
   (b) Determine the convergence disk of this power series.
   (c) Integrate the power series from 0 to z in order to obtain a representation of the function $\ln(1+z)$.
   (d) Compare the result with the Taylor series of $\ln(1+z)$ around the point $z = 0$.

5. Consider the following differential equation for a real function $f(x)$, with $k \neq 0$,

$$\frac{\partial^2}{\partial x^2} f(x) - k^2 f(x) = 0.$$

Assume that $f(x)$ can be faithfully represented by a power series,

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

substitute this power series in the equation, differentiating it term by term, and manipulate the summation indices in order to write the equation as a certain power series equaled to zero. Determine which should be the relations between the coefficients a_n of the series above so that it is indeed a solution of the equation. Pay particular attention to the first two coefficients. Determine in this way two different functions that are solutions of the equation and identify these functions.

6. **(Challenge Problem)** Consider the question of the calculation of the sucessive derivatives of the product of two analytic functions, $f_1(z)$ and $f_2(z)$, that is, derivatives of $f(z) = f_1(z)f_2(z)$, where $f_1(z)$ and $f_2(z)$ are known only in terms of their expressions in series around a reference point z_0.

   (a) Using the Leibniz rule, write the first, second and third derivatives of $f(z)$ in terms of the derivatives of $f_1(z)$ and $f_2(z)$.

   (b) From this kind of experimentation, induce a general formula for the n-th derivative of $f(z)$, in terms of a sum involving the appropriate combinatorial factors.

   (c) Prove by finite induction the formula that was induced, that is, assume that the formula is true for the case $n-1$ and show that, as a consequence of this, the case n also holds.

   (d) Consider now the product $f(z)$ of the two functions, represented by the product of the Taylor series of $f_1(z)$ and of $f_2(z)$. Collect the terms with a definite power n of $z - z_0$ and write a general formula for the coefficient of the power n.

   (e) Use the results to prove that the Taylor series of the function $f(z)$ that is the product of $f_1(z)$ and $f_2(z)$ is obtained as the product of the two series.

7. **(Challenge Problem)** Assuming that $f_2(z)$ does not have any zeros in the common region of analyticity of $f_1(z)$ and $f_2(z)$, initially show that if $f(z)$ is the *ratio* of the two analytic functions $f_1(z)$ and $f_2(z)$,

$$f(z) = \frac{f_1(z)}{f_2(z)},$$

   then the Taylor series $S(z)$ of $f(z)$ is such that $S_2(z)S(z) = S_1(z)$, where $S_1(z)$ is the Taylor series of $f_1(z)$ and $S_2(z)$ is the Taylor series of $f_2(z)$, all with respect to the same reference point z_0. Prove this fact *without* using the explicit expression of the coefficients of $S(z)$ in terms of the coefficients $S_1(z)$ and $S_2(z)$.

Hint: use the result for the product of two analytic functions, which was proved before.

Then consider the question of the calculation of the derivatives of $f(z)$ where $f_1(z)$ and $f_2(z)$ are known only in terms of their expressions in series. The idea is to try to repeat the inductive-deductive scheme that was used in the corresponding deduction for the product of two analytic functions, and to use the result to explicitly show how the Taylor series of $f(z)$ is in fact obtained from the Taylor series of $f_1(z)$ and $f_2(z)$.

Hint: consider simplifying the problem by first trying to solve the case where $f_1(z) = 1$, that is, first try to solve the problem of finding the Taylor series of the multiplicative inverse of a function,

$$f(z) = \frac{1}{f_2(z)},$$

which naturally means that

$$f(z)f_2(z) = 1.$$

After that one may consider combining this solution with that of the problem of the product, which was solved before.

Warning: this problem has a very difficult combinatorial part and, for the time being, a truly complete solution is not available.

Chapter 13

Convergence Criteria and Proofs

We will now describe and prove the sequence of theorems which ensures that a series of complex powers is always uniformly convergent in closed disks contained within its maximum disk of simple convergence. In addition to this, we will also elaborate on the ideas related to the fact that this is a necessary and sufficient condition for us to be able to integrate and differentiate this type of series term-by-term. We will describe more accurately the relevant mathematical facts, and present simple proofs of the three main theorems, with no intention of giving precise and complete formal proofs of all the theorems that are involved and that were used in the previous chapter (Chapter 12).

As was seen in previous chapters (Chapters 11 and 12), there are three types of convergence criteria that can be applied to series of functions, simple or point-by-point convergence, absolute convergence and uniform convergence, which is a criterion which applies to closed domains. If we have a series of functions such as

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

that is convergent within a certain domain of values of z, uniform convergence is always a necessary criterion for us to be able to differentiate the series term-by-term, thus obtaining another series that is also convergent in this same domain. What matters for our purposes here is that for power series this condition is also sufficient. Moreover, uniform convergence is always a sufficient criterion for us to be able to integrate the series term-by-term, while for power series this condition is also necessary. Thus we have available for our purposes here the following result. Given a series of functions that are powers of z,

$$\begin{aligned} S_\infty(z) &= \sum_{n=0}^{\infty} f_n(z) \\ &= \sum_{n=0}^{\infty} a_n z^n, \end{aligned}$$

which is uniformly convergent for z on a given closed domain, then the two series associated to it by term-by-term differentiation and term-by-term indefinite integration,

$$S'_\infty(z) = \sum_{n=0}^{\infty} \frac{df_n}{dz}(z),$$

and

$$\Sigma_\infty(z) = \sum_{n=0}^{\infty} \int^z f_n\left(z'\right)\, dz',$$

will also be convergent on that same domain, and will represent well-defined functions, which are respectively the derivative and the primitive of the function given by the original series.

The need for uniform convergence to ensure this is related to an interference between two limiting processes. Both the differentiation operation and the indefinite integration operation are abbreviations for limiting processes. On the other hand, the sum of an infinite series is also a limiting process. When we pass the differentiation or integration operation into the infinite sum, we are effectively reversing the order in which the two limits are taken, one being the limit of the sum of the series, and the other the limit involved in the differentiation or integration operation. The exchange of the order of two limits is a subtle and delicate thing, and can often change the final result. The uniform convergence of the series ensures that one can interchange the order of the two limits without changing the results.

With all this stated, we can now describe the sequence of results that allow us us to state that a series of complex powers that is convergent within a convergence disk, is also uniformly convergent on any closed disk contained within that convergence disk. To begin with, given an arbitrary power series

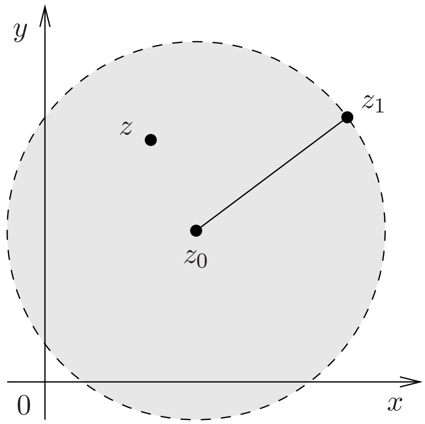


Figure 13.1: The disk defined by z_0 and z_1, and an arbitrary point z strictly inside this disk.

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

around the reference point z_0, we assume the single condition that the given coefficients a_n be such that the series converges at least at one point $z_1 \neq z_0$, that is, we assume that in addition to having the trivial convergence at z_0, with $S_\infty(z_0) = a_0$, we also have that the infinite numerical sum

$$S_\infty(z_1) = \sum_{n=0}^{\infty} a_n (z_1 - z_0)^n$$

exists and has a finite value. The first step of the sequence is to show that, starting from this, we have that the series converges absolutely for each and every point z such that $|z - z_0| < |z_1 - z_0|$, that is, that the series of absolute values

$$\bar{S}_\infty(z) = \sum_{n=0}^{\infty} |a_n (z - z_0)^n|$$

converges strictly inside the disk of radius $|z_1 - z_0|$ centered at z_0, as illustrated in the diagram of Figure 13.1. Let us formalize this statement as the following theorem.

Theorem 13.1: *Given a power series $S_\infty(z)$ written with respect to the reference point z_0, which is convergent at a point $z_1 \neq z_0$, it follows that the series is absolutely convergent within an open disk centered at z_0 with radius $|z_1 - z_0|$.*

Proof 13.1.1:

In order to prove this, we begin by noting that, since the series converges at z_1, each of the terms in the sum $S_\infty(z_1)$ is finite, which of course is also a consequence of the fact that both a_n and $(z_1 - z_0)$ are finite complex numbers, for all values of n. Moreover, a necessary condition for any series to be convergent is that its terms tend to zero when n goes to infinity. In the case of our power series it is therefore necessary that we have

$$\lim_{n\to\infty} a_n(z_1 - z_0)^n = 0.$$

This is true for real series, and one can see that it can be extended immediately to complex series, by just considering separately the real and imaginary parts of these series. It follows that, if the series converges at z_1, then this must be true at that point. Since the fact that a complex number is zero is equivalent to the fact that its absolute value is zero, we can write this necessary condition as

$$\lim_{n\to\infty} |a_n||z_1 - z_0|^n = 0.$$

This means, in particular, that the sequence of terms of the series converges, and therefore that the complete set of all the real numbers $|a_n||z_1 - z_0|^n$ has an upper limit ρ_M, such that, for all n,

$$|a_n||z_1 - z_0|^n \leq \rho_M,$$

which, since $|z_1 - z_0| \neq 0$, implies that the coefficients satisfy

$$|a_n| \leq \frac{\rho_M}{|z_1 - z_0|^n}.$$

If we now consider a partial sum $\bar{S}_N(z)$ of the series $\bar{S}_\infty(z)$, we can state that

$$\begin{aligned}
\bar{S}_N(z) &= \sum_{n=0}^{N} |a_n||z-z_0|^n \\
&\leq \sum_{n=0}^{N} \rho_M \frac{|z-z_0|^n}{|z_1-z_0|^n} \\
&= \rho_M \sum_{n=0}^{N} \left(\frac{|z-z_0|}{|z_1-z_0|} \right)^n ,
\end{aligned}$$

where ρ_M is a finite real constant. We now see that the sum on the right-hand side of this equation is the sum of a geometric progression, so that we can write that

$$\begin{aligned}
\bar{S}_N(z) &\leq \rho_M \frac{1-\left(\dfrac{|z-z_0|}{|z_1-z_0|} \right)^{N+1}}{1-\left(\dfrac{|z-z_0|}{|z_1-z_0|} \right)} \\
&= \rho_M \frac{1-\alpha^{N+1}}{1-\alpha} \\
&< \frac{\rho_M}{1-\alpha},
\end{aligned}$$

where we call the real quantity in parenthesis α, which by hypothesis is less that one. Here we see that all the partial sums $\bar{S}_N(z)$ of the series are bounded by a positive real number which is independent of N, which suffices to imply that the series converges. In other words, if α is strictly less than one, which is true if z is strictly internal to the disk, we can now take the $N \to \infty$ limit, obtaining a finite result on the right-hand side, because in this limit the power $N+1$ of the quantity α will go to zero exponentially fast. We thus obtain

$$\begin{aligned}
\bar{S}_\infty(z) &= \lim_{N\to\infty} \bar{S}_N(z) \\
&\leq \rho_M \lim_{N\to\infty} \frac{1-\alpha^{N+1}}{1-\alpha} \\
&= \frac{\rho_M}{1-\alpha}.
\end{aligned}$$

We see in this way that the series $\bar{S}_\infty(z)$, which is a sum of positive terms, and whose partial sums are therefore monotonically increasing, has an upper limit, thus being finite and convergent to some real number between zero and the upper limit shown above. The convergence condition is just that $\alpha < 1$, that is, that $|z - z_0|$ is strictly less than $|z_1 - z_0|$, or that z is contained strictly within the disk of radius $|z_1 - z_0|$ centered at z_0. Since $\bar{S}_\infty(z)$ converges at any point z strictly within this disk, it follows that $S_\infty(z)$ is absolutely convergent in this same set of points. This completes the proof of Theorem 13.1.

Thus we see that, with a very weak condition on $S_\infty(z)$, that is, that it converge at a single point z_1 besides z_0, we can verify that it is in fact absolutely convergent in the whole interior of the disk defined by z_0 and z_1. Also, since absolute convergence implies simple convergence of the series, we have that $S_\infty(z)$ converges within this same disk.

In other words, given *any* power series, we can state that, either it converges inside a disk, or it converges only at z_0, that is, at a single point, which can be interpreted as a zero-radius disk. It is clear that the power series have very special properties.

With the convergence and absolute convergence established, strictly within the disk defined by z_0 and z_1, one can then show that the series converges uniformly in any closed disk contained within the original open disk given by $|z - z_0| < |z_1 - z_0|$. Therefore, let us now prove the following theorem.

Theorem 13.2: *Given a power series $S_\infty(z)$ written with respect to the reference point z_0, which is convergent at a point $z_1 \neq z_0$, it follows that the series is uniformly convergent within any closed disk centered at z_0 and contained within the open disk centered at that point and with radius $|z_1 - z_0|$.*

Proof 13.2.1:

In order to prove this, we begin by noting that, since the series $\bar{S}_\infty(z)$ converges strictly within the disk, the remainders $\bar{R}_N(z)$ associated with the partial sums $\bar{S}_N(z)$ have the property that, given an arbitrary strictly positive real number ϵ, there is an integer $N(\epsilon)$ such that $N > N(\epsilon)$ implies that the remainder $\bar{R}_N(z)$ is less than ϵ,

$$\begin{aligned} \bar{R}_N(z) &= \sum_{n=N+1}^{\infty} |a_n||z - z_0|^n \\ &< \epsilon, \end{aligned}$$

which is true for any point z strictly internal to the disk. Let us now consider a closed disk centered at z_0, with a radius r_m that is strictly less than $|z_1 - z_0|$, as illustrated in Figure 13.2.

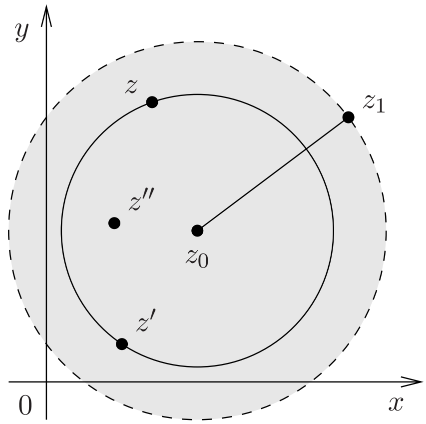


Figure 13.2: The closed disk contained within the original open disk, with the points z and z' on its boundary, and a point z'' strictly within this closed disk.

The disk being closed means that it includes its own boundary, so that we can consider a point z on this boundary as part of the disk. Since $r_m < |z_1 - z_0|$, this point is internal to the original open disk, so that the condition above applies to it, that is, given an ϵ there is an $N(\epsilon)$ such that $N > N(\epsilon)$ implies that

$$\begin{aligned}
\bar{R}_N(z) &= \sum_{n=N+1}^{\infty} |a_n||z - z_0|^n \\
&= \sum_{n=N+1}^{\infty} |a_n| r_m^n \\
&< \epsilon,
\end{aligned}$$

for this particular point. However, we now see that, if we consider any other point z' over the boundary of the closed disk, we have both that $|z' - z_0| = r_m$ and that $|z - z_0| = r_m$, that is, we have that $|z' - z_0| = |z - z_0|$, so that the above condition does not change if we exchange z' for z, and hence we have that

$$\begin{aligned}
\bar{R}_N(z') &= \sum_{n=N+1}^{\infty} |a_n||z' - z_0|^n \\
&= \sum_{n=N+1}^{\infty} |a_n| r_m^n
\end{aligned}$$

$$< \quad \epsilon.$$

This implies that the convergence condition for $\bar{S}_\infty(z)$ not only holds at all points of the boundary of the closed disk, but also holds with exactly the *same* value of $N(\epsilon)$, which in fact does not depend on the position of the point along the boundary of the closed disk. Furthermore, if we now consider what happens in the case of a point z'' which is strictly *within* the closed disk, for which $|z'' - z_0| < r_m$, that is, $|z'' - z_0| < |z - z_0|$, we see that the sum that appears in the inequality above can only decrease, thus maintaining the inequality satisfied if we exchange z'' for z, so that we have

$$\begin{aligned}
\bar{R}_N\left(z''\right) &= \sum_{n=N+1}^{\infty} |a_n| \left|z'' - z_0\right|^n \\
&< \sum_{n=N+1}^{\infty} |a_n||z - z_0|^n \\
&= \sum_{n=N+1}^{\infty} |a_n| r_m^n \\
&< \epsilon,
\end{aligned}$$

still for the same value of $N(\epsilon)$. Thus we see that, given a value of ϵ, the solution $N(\epsilon)$ that satisfies the convergence condition for a point on the boundary of the closed disk also satisfies that condition for any internal point on that disk. In summary, given a value of ϵ, there is a solution $N(\epsilon)$ which satisfies the absolute convergence condition for all points on the closed disk, and is thus independent of z on that domain. Since absolute convergence implies simple convergence, with the *same* value of $N(\epsilon)$, it follows therefore that the series $S_\infty(z)$ is uniformly convergent on the whole closed disk. This holds for any closed disk that is contained strictly within the original open convergence disk. This completes the proof of Theorem 13.2.

It is therefore established that, if a power series is convergent at all, then it is absolutely and uniformly convergent within a certain closed disk, which can be, in fact, any closed disk centered at the point z_0 and contained within the original open disk centered there and with radius $|z_1 - z_0|$. This establishes in turn that $S_\infty(z)$ is a continuous function on that closed disk, as well as that we can integrate the series term-by-term, thereby producing another series which is also convergent within that same closed disk. This also justifies our use of this result in the previous chapter (Chapter 12), in which we showed that it implies, in fact, that $S_\infty(z)$ is an analytic function on the closed disk, and is therefore infinitely differentiable as well as infinitely integrable there.

Note that the facts that $S_\infty(z)$ is an analytic function within the open disk and that the power series is its Taylor series, as was shown in the previous chapter (Chapter 12), also ensure that these two other series, obtained by term-by-term integration and differentiation, are not only convergent, but also absolutely convergent within the open disk, and uniformly convergent on any closed disk contained within it. This is true because, as we have shown, if we integrate or differentiate the Taylor series of an analytic function term-by-term, we always get automatically the Taylor series of either the primitive or the derivative of that function, which in either case is another analytic function on exactly the same domain as the original function, the resulting series being therefore convergent to the corresponding functions on that domain.

It is important to note, however, that the fact that the series $S_\infty(z)$ converges at z_1 does *not* guarantee that the function given by that series is analytic *at* that point, because z_1 may be on the border of the maximum convergence disk of the series, and the function may be singular there. Therefore, while it is true that the series obtained by term-by-term differentiation will be convergent within the open disk, it is *not* necessarily true that it will be convergent at all points where the original series converges. In fact, the differentiated series may be divergent at z_1, if it so happens that this point is on the border of the maximum convergence disk of the original series, and that the function has a singularity there. On the other hand, the integrated series will always be convergent at z_1. And yes, the Taylor series of a function may be convergent at a point where the function is singular, so long as the limit of the function to that point exists. However, further discussion of these facts here would take us too far afield, and is beyond the scope of this text.

Note that we were careful to distinguish here between open and closed disks. By definition, a closed disk is one that includes its boundary, while the open disk is only the interior, without the boundary. This is important because of course it cannot be said that the power series of a complex function converges, let alone that it converges uniformly, in its maximum disk of convergence, seen as a closed disk. This is due to the fact that the boundary of that disk necessarily includes a point of singularity of the function somewhere, at which point the series may not converge at all. Let us recall that, for us to be able to integrate and differentiate the series term-by-term, it is necessary that it be uniformly convergent on a closed domain, including the boundary points, which therefore excludes the maximum disk. It may be that the series converges at some point at the boundary of the maximum disk, but typically it will do it very slowly in comparison to the speed of convergence at the internal points, so that in general one cannot expect uniform convergence.

The arguments involving uniform convergence always involve closed disks, and thus the most that can be stated is that uniform convergence holds for any closed disk that is contained in the open disk that is the maximum convergence disk of the series. In order to be contained in the maximum open disk, these closed disks must have radius at least slightly smaller than the radius of the maximum convergence disk. However, none of this disturbs the use that we want to make here of the result that power series are uniformly convergent. Our goal is to differentiate or integrate the series term-by-term, the differentiation being at strictly internal points of the maximum convergence disk, and the integration on contour strictly internal to that same disk. Therefore, it is enough to know that the series is uniformly convergent on all closed disks strictly internal to the maximum convergence disk. It is clear that the results obtained can be as close as one wants to the boundary, they just cannot be on the boundary. Thus in the end the results obtained are valid for all internal points of the maximum convergence disk, which is as much as we could wish for in any case.

We will end this chapter with a discussion of the convergence criteria based solely on the examination of the terms of the series. In the discussion we have had so far, on the issue of determining the convergence domain, we are implicitly assuming that we know in advance the function whose representation through power series is being discussed. If indeed we know the functions and know where their singularities are, then we can determine whether and when the series converge. However, sometimes we have available only the representation of the function by a series, and no other information about it, as will be the case in later parts of this series of books. In this case, it is important to know how to determine the convergence of the series, or lack thereof, only through an examination of its terms. In order to be able to accomplish this, here we will examine one of these convergence criteria, which is the most popular, simple and useful for our purposes.

Most of the convergence criteria of this type are intended to identify whether or not a series is absolutely convergent, and therefore they apply to the series of absolute values of the terms of the original series. They are generally sufficient but not necessary conditions for the convergence. The series does not necessarily need to be a power series, or even a series of functions, it can be any type of series, including a purely numerical series. In the case of power series, it is clear that the only variable part of the structure is constituted by the set of constant coefficients of the terms, whose behavior will then be decisive for the convergence of the series. In the general case, we just consider the terms of the series, be them dependent on a variable z or not. Therefore, consider the complex series given by

$$S_\infty = \sum_{n=0}^{\infty} t_n,$$

where the term t_n is given by $a_n(z - z_0)^n$ in the case of a power series, but where we are considering the more general case. The convergence criterion that we will discuss, which is called the *criterion of the ratio*, or *ratio test* for short, is as follows: if there exist a real number q such that $0 < q < 1$ and a certain value n_0 of n such that for all $n \geq n_0$ we have that

$$\frac{|t_{n+1}|}{|t_n|} < q,$$

it follows that the series is absolutely convergent. Note that this can be, in fact, a statement about the limit of the ratio $|t_{n+1}|/|t_n|$ when $n \to \infty$, if this limit exists, because the inequality above immediately implies that

$$\begin{aligned} \lim_{n\to\infty} \frac{|t_{n+1}|}{|t_n|} &\leq q \\ &< 1, \end{aligned}$$

since we have here a limited sequence of positive numbers $|t_{n+1}|/|t_n|$, so that, if it is convergent at all, then it must converge to a positive number less than one. We cannot, however, be sure that this sequence converges, because it need not be monotonic. It should be noted however that the ratio test is valid and applicable regardless of whether or not this limit exists. On the other hand, if the limit exists and is less than one, then the ratio test is satisfied and hence the series converges. We may call this new criterion involving a limit, that is, the ratio test in the event that the limit exists, the *criterion of the limit of the ratio*, or *ratio-limit test* for short. It is usually simpler and easier to apply than the ratio test itself. Let us establish, then, the ratio test, as the following theorem.

Theorem 13.3: *Given a complex series S_∞ with terms t_n, if there exist a real number q such that $0 < q < 1$ and a certain value n_0 of n such that for all $n \geq n_0$ we have that*

$$\frac{|t_{n+1}|}{|t_n|} < q,$$

then it follows that the series S_∞ is absolutely convergent.

Proof 13.3.1:

The proof of this criterion is worked out by the comparison of the given series with a well-known series, the geometric series of ratio q. As we have seen before in this chapter, the geometric series, that is, the infinite-sum limit of a complex geometric progression, has a central role to play in the discussion of complex power series. Beginning by the term t_{n_0} we can iterate the above relation to obtain a relation between t_n and t_{n_0}. Writing the first two cases, we have

$$\begin{aligned} |t_{n_0+1}| &< q|t_{n_0}|, \\ |t_{n_0+2}| &< q|t_{n_0+1}| \quad \Rightarrow \\ |t_{n_0+2}| &< q^2|t_{n_0}|, \end{aligned}$$

and so on, until we have obtained the relation, for an arbitrary value of $n = n_0 + k$,

$$\begin{aligned} |t_{n_0+k}| &< q^k|t_{n_0}| \quad \Rightarrow \\ |t_n| &< q^{n-n_0}|t_{n_0}|. \end{aligned}$$

We can now add over n on the two sides of this inequality, going from $n = n_0+1$ to any value $N > n_0+1$, thus obtaining the corresponding relation between certain partial sums, that of the series $\bar{S}_\infty$ associated with the original series S_∞ from which we started, and that of a geometric series,

$$\begin{aligned}
\sum_{n=n_0+1}^{N} |t_n| &< \sum_{n=n_0+1}^{N} q^{n-n_0}|t_{n_0}| \quad \Rightarrow \\
\sum_{n=n_0+1}^{N} |t_n| &< \frac{|t_{n_0}|}{q^{n_0}} \sum_{n=n_0+1}^{N} q^n.
\end{aligned}$$

Note that the sum on the left-hand side is associated to the remainder $\bar{R}_{n_0}$, which in turn is associated with the partial sum $\bar{S}_{n_0}$ of the series $\bar{S}_\infty$. The sum on the right-hand side is now the sum of a real geometric progression, and has a well-known value, so that we can write that

$$\begin{aligned}
\sum_{n=n_0+1}^{N} |t_n| &< \frac{|t_{n_0}|}{q^{n_0}} \frac{q^{n_0+1} - q^{N+1}}{1-q} \\
&= |t_{n_0}|\, q\, \frac{1 - q^{N-n_0}}{1-q} \\
&< |t_{n_0}| \frac{q}{1-q},
\end{aligned}$$

where by hypothesis $q < 1$. Here we see that all the partial sums $\bar{S}_N$ are bounded by a positive real number independent of N, which suffices to imply that the series converges. Since for $q < 1$ the geometric series converges, and since the sum on the left-hand side is a sum of positive terms, that is limited by the result on the right-hand side, it follows that the series on the left-hand side converges, and taking the limit $N \to \infty$ we have the relation,

$$\sum_{n=n_0+1}^{\infty} |t_n| < |t_{n_0}| \frac{q}{1-q}, \tag{13.1}$$

where we recall that the left-hand side is the remainder $\bar{R}_{n_0}$ that is associated with the partial sum $\bar{S}_{n_0}$ of the series $\bar{S}_\infty$. Adding the initial terms of the sum giving $\bar{S}_\infty$, of which there is a finite number and whose sum is therefore equal to some finite complex number, we have for the series $\bar{S}_\infty$,

$$\begin{aligned}
\bar{S}_\infty &= \sum_{n=0}^{\infty} |t_n| \\
&< \sum_{n=0}^{n_0} |t_n| + |t_{n_0}| \frac{q}{1-q}.
\end{aligned}$$

We have expressed here the fact that $\bar{S}_\infty$ is a sum of real positive terms that is limited from above, which implies that this series converges. We have therefore that the series S_∞ is absolutely convergent, from which it follows that it is also simply convergent. This establishes the criterion of the ratio of the absolute values of successive terms as sufficient to ensure the convergence of the series.

Proof 13.3.2:

It is interesting to observe that there is an alternative way to verify the convergence of $\bar{S}_\infty$, using the same elements that went into the argument above. The inequality that we previously obtained for the terms of the series,

$$|t_n| < q^{n-n_0}|t_{n_0}|,$$

implies that, since $q < 1$, the sequence of terms of the series tends to zero for $n \to \infty$, because we have that

$$\begin{aligned} \lim_{n\to\infty} |t_n| &= \frac{|t_{n_0}|}{q^{n_0}} \lim_{n\to\infty} q^n \\ &= 0 \quad \Rightarrow \\ \lim_{n\to\infty} t_n &= 0, \end{aligned}$$

which, as we already know, is a necessary but not sufficient condition for the convergence of the series. On the other hand, on the left-hand side of Equation (13.1) above we have the remainder $\bar{R}_{n_0}$ associated to the partial sum $\bar{S}_{n_0}$ of the series $\bar{S}_\infty$, and on the right-hand side we have an upper limit for this remainder, that is, we have that

$$\begin{aligned} \bar{R}_{n_0} &= \sum_{n=n_0+1}^{\infty} |t_n| \\ &< |t_{n_0}| \frac{q}{1-q}. \end{aligned}$$

Of course we can make $n_0 \to \infty$ in this expression, and since we have just shown that the sequence $|t_n|$ has zero limit, it follows that in the $n_0 \to \infty$ limit the upper limit of the positive remainder $\bar{R}_{n_0}$ vanishes, which in turn implies that

$$\lim_{n_0\to\infty} \bar{R}_{n_0} = 0,$$

which is another way, perhaps a little more direct, to verify the convergence of the series $\bar{S}_\infty$. This completes the proof of Theorem 13.3.

This comparison of a given series of positive terms with another series, this one known to be convergent or divergent, is one of the most common ways to determine the convergence or divergence of the given series. Another common way to proceed is to compare the series with an asymptotic integral that is known to be either convergent or divergent. The geometric series is particularly useful for this purpose, since its convergence and its limiting value are easily determinable. Whatever the way in which one is able to limit the partial sums of the given real series of positive terms from above, this bound is always sufficient to ensure their convergence. The ratio test is only the most common and simplest way to do this. As a simple example of use of the test, let us consider the Maclaurin series of the exponential,

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

By simply calculating the ratio of the absolute values of a given term and the previous one, we obtain

$$\left| \frac{z^{n+1}/(n+1)!}{z^n/n!} \right| = \frac{1}{n+1} |z|.$$

If we use the ratio-limit test, it is very simple to determine the convergence of the series. Given any finite value of z, the limit of this ratio when n tends to infinity is zero, which is smaller than one. Thus, the convergence of the series is guaranteed, at every point on the complex plane, as expected, since we know that the function is analytic on the whole complex plane, and that is has no singularities. As one can see, the radius of convergence of this series is infinite, and the maximum convergence disk is the whole complex plane. Here we have, therefore, a very simple example of the usefulness of the ratio-limit test to determine the structure of singularities of a function.

Just as an illustration, we will discuss how to determine the convergence of the series using only the ratio test, regardless of the limit of the ratio. Considering once more the ratio of the absolute values of a term and of the previous term, we now impose the condition that

$$\frac{1}{n+1} |z| < q.$$

The question now is whether, given a value of $\rho = |z|$, there exists a real number $q < 1$ and a certain integer n_0 such that this inequality is valid for all $n > n_0$. In order to determine this, we write this inequality as

$$\begin{aligned} \rho &< (n+1)q \quad \Rightarrow \\ n &> \frac{\rho}{q} - 1. \end{aligned}$$

It is now enough to choose n_0 as the smallest integer greater than the right-hand side of the latter inequality, so that it is satisfied when $n > n_0$. Since this can always be done, whatever the value of q, it is enough to choose any value of q in $(0, 1)$ and adopt the corresponding n_0 for the ratio criterion to be satisfied. Since this can be done regardless of the value of ρ, it follows that the series is convergent on the entire complex plane. As one can see, the application of this test is a bit more elaborate than that of the ratio-limit test, and this is the reason why, whenever the limit exists, it is more convenient to apply that simpler and more straightforward criterion.

Let us discuss for a moment the question of the necessity of the ratio-limit test. In the most general case this test, just as the ratio test itself, is sufficient but not necessary. However, in the case of power series it is possible to identify the circumstances under which the ratio-limit test is actually necessary, if we limit ourselves to the cases in which the limit of the ratio exists, of course. What has been shown above is that the ratio-limit test is sufficient, that is, if the limit of the ratio of the absolute values of the two terms is less than one, then the series $\bar{S}_\infty$ is convergent. Moreover, it is clear that if the limit of that ratio is strictly greater than one, then the terms increase in magnitude with n and thus do not go to zero for $n \to \infty$, which prevents the convergence of this real series, since the limit of the sequence of terms of the series being zero is a necessary condition for its convergence. Therefore it follows that the series S_∞ is not absolutely convergent. Moreover, since for power series simple convergence implies absolute convergence, as was shown above, it follows that for this type of series we also have that S_∞ is in fact divergent.

In short, for the case of power series for which the limits involved exist, the limit of the ratio of the absolute values of the two terms being strictly greater than one implies the divergence of the series. It remains to be discussed only the possibility that the limit of that ratio be exactly equal to one. So far we have concluded here that in the general case, in so far as the possible values of the limit are concerned, the ratio-limit test is sufficient but not necessary, because there may be convergent series in which this limit is exactly equal to one. Let us see, however, what happens in the specific case of a power series when the limit of this ratio is equal to one. Let us consider that a certain power series developed around the reference point z_0 is convergent at a point $z_1 \neq z_0$,

$$S_\infty(z_1) = \sum_{n=0}^{\infty} a_n (z_1 - z_0)^n.$$

For this series the ratio-limit test involves the behavior of the quantity

$$\frac{|a_{n+1}|}{|a_n|} |z_1 - z_0|$$

for large values of n. If the series is convergent at the point z_1, and assuming that the limit of the ratio exists, then the limit of this ratio as $n \to \infty$ cannot be larger than one, since this would imply the divergence of the series. Thus, this limit must be equal to one or less than one. If it is less than one, then the ratio-limit test is satisfied and the series converges at that point. If the limit is equal to one, then it follows that any infinitesimal increase in the quantity $|z_1 - z_0|$ makes the limit larger than one, and thus makes the series divergent. It follows that a point where the series converges and the limit is equal to one cannot be strictly inside the maximum disk of convergence of the series. It is necessary that it be located on the boundary, because otherwise there would be an open neighborhood around it within which the function is still analytic and the series is still convergent, which would allow us to increase $|z_1 - z_0|$, while still remaining within this open neighborhood, and thus still having a convergent series.

In conclusion, any point at which the limit of the ratio of the absolute values of the consecutive terms of a power series exists and is equal to one is necessarily a point on the boundary of the maximum convergence disk, and any point at which the limit of the ratio of the terms exists and is strictly less than one is necessarily a point strictly within this disk, and at this point the series converges. Therefore, if we exclude the points on the boundary of the maximum convergence disk, the ratio-limit test becomes a necessary and sufficient criterion for the convergence of the series. Since the discussion of uniform convergence must be made on closed domains and therefore must always exclude the points of the boundary of the maximum convergence disk, in the context of a discussion involving uniform convergence the ratio-limit test is a necessary and sufficient condition for the convergence of power series. Looking for values of z for which the limit is exactly equal to one we can try to locate the boundary of the maximum convergence disk of the series, along which the function has at least one singularity.

Mathematical Epilogue: Uniform Convergence

We are now in a good position to apply the ratio-limit criterion to the case of power series in a way that will allow us to better clarify the relation between term-by-term differentiability, term-by-term integrability and uniform convergence. Note, however, that we will limit this discussion to the cases in which the limit of the ratio of the absolute values of two consecutive terms of the series exists. Hence, let us consider once again a series which is convergent within the open disk defined by z_0 and z_1, that is, such that

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

is a finite number when z is within this open disk, and which is also such that the limit of the ratio exists. Let us also assume that the point z is on a closed disk contained in the maximum convergence disk, that is, that z is a point strictly internal to this maximum convergence disk. If this series is convergent at this point, then it follows that the ratio-limit condition is valid, since we have shown that in this case it is a necessary and sufficient condition. Therefore, it follows that we have

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|} |z - z_0| < 1.$$

We also know that this series is in fact uniformly convergent within the closed disk, since it is a power series. The series we get from this one, differentiating it term-by-term, is

$$\begin{aligned} S'_\infty(z) &= \sum_{n=0}^{\infty} \frac{df_n}{dz}(z) \\ &= \sum_{n=1}^{\infty} n a_n (z - z_0)^{n-1}. \end{aligned}$$

If we calculate the ratio of the absolute values of a term and the previous one for this series, we get

$$\frac{(n+1)|a_{n+1}|}{n|a_n|}\,|z-z_0| = \frac{n+1}{n}\,\frac{|a_{n+1}|}{|a_n|}\,|z-z_0|.$$

The first fraction of this expression has as its limit when $n \to \infty$ the value one, although it is approached by values greater than one, and the rest of the expression is exactly the same as we had before, so that we obtain, calculating the limit,

$$\begin{aligned}
&\lim_{n\to\infty}\left(\frac{n+1}{n}\,\frac{|a_{n+1}|}{|a_n|}\,|z-z_0|\right) \\
&= \lim_{n\to\infty}\left(\frac{n+1}{n}\right) \times \lim_{n\to\infty}\left(\frac{|a_{n+1}|}{|a_n|}\,|z-z_0|\right) \\
&= \lim_{n\to\infty}\frac{|a_{n+1}|}{|a_n|}\,|z-z_0|.
\end{aligned}$$

We therefore conclude that, if the limit associated with the original series is less than one, then this one is as well, and therefore the series obtained by differentiating the first one term-by-term is also convergent at that same point.

Let us consider where we are with this argument. We already know that the uniform convergence condition on some closed domain is a necessary condition for the series obtained by term-by-term derivation to be convergent. We already know that any convergent power series is also uniformly convergent on closed sets contained strictly within the maximum convergence disk. Thus, if this uniform convergence condition in some closed domain were not sufficient, for the case of power series, then there would be some uniformly convergent power series on such a closed domain, such that the series obtained from it by term-by-term differentiation would not be convergent.

However, we have just proved, using the ratio-limit test, that the series obtained from any power series by term-by-term differentiation is always convergent, as well as uniformly convergent, within the same closed domain in which the original series is uniformly convergent. Thus, it is proven that there is no case that violates the sufficiency of the uniform convergence condition, and therefore it is a necessary and sufficient condition for term-by-term differentiability in the case of power series.

The same argument can be applied to the case of the term-by-term integration of the series $S_\infty(z)$, but this time we do not assume that this series is convergent at z. The process of term-by-term integration applied to it produces the series

$$\begin{aligned}\Sigma_\infty(z) &= \sum_{n=0}^{\infty} \int^{z} f_n(z')\, dz' \\ &= \sum_{n=0}^{\infty} \frac{a_n}{n+1}\,(z-z_0)^{n+1},\end{aligned}$$

where we ignored the integration constant, which is irrelevant. If we calculate the ratio of the absolute values of a term with the previous one for this series, we obtain

$$\frac{(n+1)|a_{n+1}|}{(n+2)|a_n|}\,|z-z_0| = \frac{n+1}{n+2}\,\frac{|a_{n+1}|}{|a_n|}\,|z-z_0|.$$

As one can see, in this case the first fraction on the right-hand side also has as its limit the number one, this time approaching the limit from below. In this case, since we know that in general uniform convergence is a sufficient condition to ensure the term-by-term integrability, and since what we want to verify is that this condition is also necessary, what matters is that the convergence of the integrated series implies that the original series is also convergent. We can verify this by using the above ratio, because if its limit is less than one,

$$\lim_{n\to\infty} \frac{n+1}{n+2}\,\frac{|a_{n+1}|}{|a_n|}\,|z-z_0| < 1,$$

then it follows that the ratio of the absolute values of the terms of the series $S_\infty(z)$ also has this property,

$$\lim_{n\to\infty} \frac{|a_{n+1}|}{|a_n|}\,|z-z_0| < 1,$$

that is, the series $S_\infty(z)$ is also convergent. Also, we must not forget that, when it comes to power series, simple convergence is equivalent to uniform convergence.

Let us now consider where we are in this case. We already know that the condition of uniform convergence on some closed domain is a sufficient condition for the convergence of the series obtained by term-by-term integration. Thus, if this uniform convergence condition on some closed domain was not necessary in the case of power series, then there would exist some divergent power series at points on such a closed domain, which would be such that the series obtained from it by term-by-term integration would still be convergent on all the domain.

However, we have just proven, using the ratio-limit test, that the convergence of the series obtained from any power series by term-by-term integration implies that the original series is convergent and also uniformly convergent on the same closed domain on which the integrated series is convergent. Thus, it is proven that there is no case that violates the necessity of the uniform convergence condition, and therefore that it is a necessary and sufficient condition for the term-by-term indefinite integrability in the case of power series.

Thus we have established, independently of the previous arguments based on the analyticity of the functions, that the uniform convergence of a power series is the necessary and sufficient condition for term-by-term differentiability and integrability. Since they are power series, it follows that the two new series obtained are also uniformly convergent on the closed disk where the original series is uniformly convergent. Also, since simple convergence and uniform convergence are equivalent for power series, it follows that the simple convergence is also a necessary and sufficient condition for this same purpose.

Problem Set

1. Use the ratio-limit test in order to show that the Maclaurin series of each of the following functions is convergent, and determine in each case the radius of the convergence disk.

   (a) $w(z) = \cos(z)$.
   (b) $w(z) = \sin(z)$.
   (c) $w(z) = \cosh(z)$.
   (d) $w(z) = \sinh(z)$.
   (e) $w(z) = \ln(1 + z)$.

2. Consider the series of positive real numbers

$$S_\infty = \sum_{n=0}^{\infty} |a_n|,$$

   where a_n are complex numbers.

   (a) Write in symbolic mathematical language, involving a positive real number ϵ, the convergence condition of the series.

(b) Write in symbolic mathematical language the condition that the limit of $|a_n|$ be zero when n goes to infinity.

(c) Show that if the series converges, then we do have this value for the limit, that is,

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

Hint: one can try using the method of reductio ad absurdum, but this time a constructive proof is simpler.

3. Consider the series of positive real numbers

$$S_\infty = \sum_{n=0}^{\infty} |a_n|,$$

where a_n are complex numbers.

(a) Write in symbolic mathematical language, involving a positive real number ϵ, the convergence condition of the series.

(b) Show that if the series converges, then there is a real number A such that $|a_n| < A$, for all n.

Hint: use the method of reductio ad absurdum.

4. Consider the Maclaurin series of the complex function

$$w(z) = \sqrt{z + \epsilon^2},$$

where ϵ is a strictly positive real number.

(a) Write the series as an infinite sum of a general term, involving factorials, double factorials and powers.

(b) Use the ratio-limit test in order to establish the convergence of the series, and calculate the radius of the convergence disk.

(c) Consider the limit in which $\epsilon \to 0$. What is the convergence domain of the series in this case?

Chapter 14

Laurent Series and Residues

As we have seen, the complex power series provide us with an important algorithmic representation of analytic functions that have no singularities in a particular region. Furthermore, we found that the construction of analytic functions with a specific singularity, such as

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

which is singular at z_0, is very useful, and encodes important information about the analytic function $f(z)$ used in the construction, obtainable by closed-contour integrals around the singular point. Thus, it would be important to have an algorithmic representation also for functions with singularities of this type, which we call *isolated singularities*, and that in cases such as that of this example are called *poles*. The Taylor series gives us an algorithmic representation,

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

within the convergence disk of the series, but only for functions that are analytic, without any such singularities. If we use this series expansion for the function $f(z)$ which we used above in the definition of $g(z)$, we can induce that we would have for $g(z)$ something like

$$g(z) \quad = \quad \frac{1}{(z - z_0)} f(z)$$

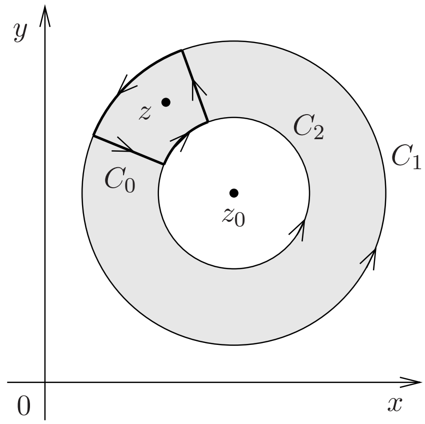


Figure 14.1: The complex plane, showing two circles C_1 and C_2, the annular region, the contour C_0, and the points z and z_0.

$$
\begin{aligned}
&= \sum_{n=0}^{\infty} \frac{f^{n\prime}(z_0)}{n!}(z-z_0)^{n-1} \\
&= \frac{f(z_0)}{z-z_0} + \sum_{n=1}^{\infty} \frac{f^{n\prime}(z_0)}{n!}(z-z_0)^{n-1} \\
&= \frac{f(z_0)}{z-z_0} + \sum_{m=0}^{\infty} \frac{f^{(m+1)\prime}(z_0)}{(m+1)!}(z-z_0)^{m},
\end{aligned}
$$

where we made a simple transformation of the summation index, and which would represent the function within the convergence disk of the original series, but away from the point z_0. Thus, the expectation is that, in order to represent a function with a pole, a power series would have to start with a sufficiently large *negative* power. Note that, by integrating both sides of this equation on a closed contour around z_0, and using the Cauchy-Goursat theorem, we obtain back the Cauchy integral formula for $f(z)$.

We will now extend the concept of the Taylor series, in such a way that it can be applied to functions with one or more isolated singularities within the disk. Consider a function $f(z)$ which is analytic on two circles around z_0, as well as in the annular region between the two circles as illustrated in Figure 14.1.

We will allow $f(z)$ to have isolated singularities strictly within the internal disk whose boundary is C_2, the nature of which need not be specified in detail for now, other than that they do not involve any

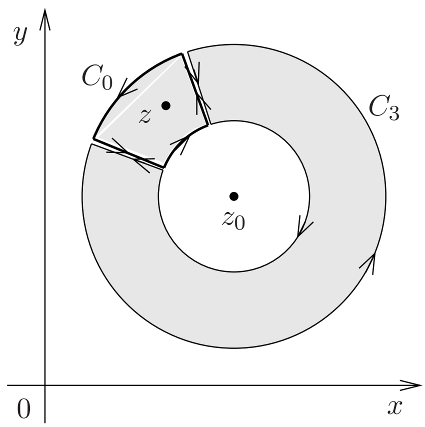


Figure 14.2: The complex plane, showing the annular region, the contours C_0 and C_3, and the points z and z_0.

branch cuts that cross the annular region between the circles C_2 and C_1. The point z_0 is the center of the two circles that define the annular region, and may be but need not be the location of one of the singularities. In this case we can use the Cauchy integral formula in order to write

$$f(z) = \frac{1}{2\pi\imath} \oint_{C_0} \frac{f(z')}{z' - z}\, dz',$$

where z is an internal point to the contour C_0, which is contained in the annular region, as shown in Figure 14.1. Using the Cauchy-Goursat theorem, and manipulating the integration contours, we can decompose the circular contours C_1 and C_2 as is shown in Figure 14.2, in such a way that the integral over C_0 can be deformed into an integral over C_1 and C_2, by attaching to it the inner region of C_3.

In order to see this it is enough to consider that, since the function $f(z')/(z' - z)$ is analytic throughout the interior of C_3, given that the point z is outside the interior of C_3, and therefore the denominator never goes to zero within or over C_3, we have that

$$\oint_{C_3} \frac{f(z')}{z' - z}\, dz' = 0.$$

It follows that we can add the integral over C_3 to the integral over C_0, thus obtaining the relation

$$f(z) = \frac{1}{2\pi\imath} \left[\oint_{C_0} \frac{f(z')}{z' - z}\, dz' + \oint_{C_3} \frac{f(z')}{z' - z}\, dz' \right].$$

Examining the combinations of the various segments of the contours and the cancellations among the corresponding parts of the integrals over them, we see that this sum of integrals can be written as a sum of integrals over C_1 and C_2,

$$\oint_{C_0} \frac{f(z')}{z'-z}\,dz' + \oint_{C_3} \frac{f(z')}{z'-z}\,dz' = \oint_{C_1} \frac{f(z')}{z'-z}\,dz' - \oint_{C_2} \frac{f(z')}{z'-z}\,dz',$$

where the sign of the term on C_2 was reversed due to the reversal of the orientation of that contour. It follows that we can write that, for a point z within the annular region,

$$f(z) = \frac{1}{2\pi\imath} \oint_{C_1} \frac{f(z')}{z'-z}\,dz' - \frac{1}{2\pi\imath} \oint_{C_2} \frac{f(z')}{z'-z}\,dz'.$$

This formula is a generalization of the Cauchy integral formula, which gives $f(z)$ in terms of a closed-contour integral around the point z, from which we previously derived the Taylor series. Indeed, the Cauchy integral formula is precisely the first of these two terms. Since the point z is within the annular region and thus outside of C_2, if $f(z)$ is analytic within C_2 then the second integral is zero, and therefore we go back to having the Cauchy integral formula in its original form.

Before we continue with the development, it is interesting to note that so far nothing we have done here depends on the fact that we chose two circles to build the annular region. It would all hold equally well for any connected region that has a single internal hole, regardless of the precise format of the region or of the hole. As we shall see, and analogously to the case of the Taylor series, it is only for the proof of convergence, that is, that the remainders go to zero in the summation limit of the series, that it is necessary that the annular region be made up of two circles.

Just as we used the Cauchy integral formula in order to derive from it the Taylor series, we can consider the possibility of deriving from this generalized formula a series to represent algorithmically the function $f(z)$, which will be valid even in cases in which $f(z)$ has singularities within the two disks whose boundaries are C_1 and C_2. For the first integral we can proceed exactly as we did in the case of the Taylor series, writing the denominator that appears in the integrand in the form

$$\begin{aligned}\frac{1}{z'-z} \;&=\; \frac{1}{(z'-z_0)} + \frac{(z-z_0)}{(z'-z_0)^2} + \frac{(z-z_0)^2}{(z'-z_0)^3} + \ldots \\ &\qquad \ldots + \frac{(z-z_0)^{n-1}}{(z'-z_0)^n} + \frac{(z-z_0)^n}{(z'-z_0)^n(z'-z)},\end{aligned}$$

where the last term is what will eventually give origin to a part of the remainder of this new series. Therefore, we can manipulate the first integral, just as we did before, in order to obtain

$$\begin{aligned} f(z) & \\ = \quad & a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \ldots + a_{n-1}(z - z_0)^{n-1} + R_n + \\ & - \frac{1}{2\pi\imath} \oint_{C_2} \frac{f(z')}{z' - z}\, dz', \end{aligned}$$

where the remainder R_n is the same that was derived before,

$$R_n = \frac{(z - z_0)^n}{2\pi\imath} \oint_{C_1} \frac{f(z')}{(z' - z_0)^n (z' - z)}\, dz',$$

and where the coefficients a_n, $n \in \{0, 1, 2, 3, \ldots, \infty\}$ are given, as before, by the integrals

$$a_n = \frac{1}{2\pi\imath} \oint_{C_1} \frac{f(z')}{(z' - z_0)^{n+1}}\, dz'.$$

These are the formulas that previously gave us the n-th derivative of $f(z)$ at the point z_0, but note that we can no longer write the coefficients a_n in this way, since by hypothesis $f(z)$ may not be analytic throughout the interior of the contour C_1, and the analyticity throughout the interior of the contour is essential for the formulas that associate these integrals with the derivatives to hold. Note however that, if it happens to be true that $(z - z_0)f(z)$ is analytic at z_0, as in our original example, that is, if it is true that

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

where $h(z)$ is analytic throughout the interior, then we can write that

$$a_n = \frac{1}{2\pi\imath} \oint_{C_1} \frac{h(z')}{(z' - z_0)^{n+2}}\, dz',$$

and then we can use the Cauchy integral formulas for the derivatives of $h(z)$, thus obtaining

$$a_n = \frac{h^{(n+1)\prime}(z_0)}{(n + 1)!}.$$

This is precisely the result that we obtain manipulating directly the series of the analytic function and the explicit singularity, just as we did earlier in this chapter. That is, in this particular case the coefficient a_n of the series of $f(z)$ is, in fact, equal to the coefficient a_{n+1} of the Taylor series of $h(z)$.

For the second integral, we can make manipulations similar to those made for the Taylor series, but we must first realize that the role of the points z and z' in relation to the contour changes in this case. For the outer contour C_1, z is within the contour and z' on it, so that we have the inequality $|z - z_0| \leq |z' - z_0|$. For the internal contour C_2, on the other hand, the point z is outside the contour, which reverses this inequality, that becomes $|z' - z_0| \leq |z - z_0|$. Thus we should expect that the differences $(z - z_0)$ and $(z' - z_0)$ will exchange their roles, so that in this case we manipulate the denominator appearing in the integrand, including the sign, as follows,

$$\begin{aligned}
\frac{-1}{z' - z} &= \frac{1}{(z - z_0) - (z' - z_0)} \\
&= \frac{1}{(z - z_0)} \frac{1}{1 - \dfrac{(z' - z_0)}{(z - z_0)}} \\
&= \frac{1}{(z - z_0)} + \frac{(z' - z_0)}{(z - z_0)^2} + \frac{(z' - z_0)^2}{(z - z_0)^3} + \ldots \\
&\qquad \ldots + \frac{(z' - z_0)^{n-1}}{(z - z_0)^n} + \frac{(z' - z_0)^n}{(z - z_0)^n \, (z - z')},
\end{aligned}$$

where the last term will eventually give origin to the other part of the remainder of this new series, and one may observe that in fact the variables z and z' have interchanged positions in relation to the development used for the Taylor series. Using this expansion in the second integral, we now obtain for the full expansion of $f(z)$ within the annular region around the point z_0,

$$\begin{aligned}
& f(z) \\
&= a_0 + a_1(z - z_0) + a_2(z - z_0)^2 + \ldots + a_{n-1}(z - z_0)^{n-1} + R_n + \\
&\quad + \frac{b_1}{(z - z_0)} + \frac{b_2}{(z - z_0)^2} + \ldots + \frac{b_n}{(z - z_0)^n} + Q_n,
\end{aligned}$$

where the remainder R_n is the same as before, the remainder Q_n is given by

$$Q_n = \frac{1}{2\pi\imath(z - z_0)^n} \oint_{C_2} \frac{(z' - z_0)^n f(z')}{(z - z')} dz',$$

and where the coefficients b_n, $n \in \{1, 2, 3, \ldots, \infty\}$, are given by

$$b_n = \frac{1}{2\pi\imath} \oint_{C_2} (z' - z_0)^{n-1} f(z') \, dz'.$$

The proof that $R_n \to 0$ when $n \to \infty$ proceeds as before, and from this point it is necessary that the annular region have as its outer boundary a circle centered at z_0. This time we obtain, in terms of the radius r_1 of the outer circle,

$$|R_n| \leq \left(\frac{r}{r_1}\right)^n \frac{r_1}{r_1 - r} f_{M_1},$$

where f_{M_1} is the maximum of the absolute value of $f(z)$ on the external contour, which implies that $R_n \to 0$ when $n \to \infty$, because $r/r_1 < 1$, since z is within the annular region and hence within C_1. In the case of Q_n, the proof is worked out in a similar way and requires that the inner boundary of the annular region also be a circle centered at z_0. In terms of the radius r_2 of the inner circle, we obtain

$$|Q_n| \leq \left(\frac{r_2}{r}\right)^n \frac{r_2}{r - r_2} f_{M_2},$$

where f_{M_2} is the maximum of the absolute value of $f(z)$ on the internal contour. Thus we also have that $Q_n \to 0$ when $n \to \infty$, because $r_2/r < 1$, since z is within the annular region and thus outside of C_2. In this way, we are allowed to write that

$$\begin{aligned} f(z) &= \sum_{n=0}^{\infty} a_n (z - z_0)^n + \sum_{n=1}^{\infty} \frac{b_n}{(z - z_0)^n}, \\ a_n &= \frac{1}{2\pi\imath} \oint_{C_1} \frac{f(z')}{(z' - z_0)^{n+1}} dz', \\ b_n &= \frac{1}{2\pi\imath} \oint_{C_2} (z' - z_0)^{n-1} f(z') \, dz'. \end{aligned}$$

This generalized series is the *Laurent series* of the function $f(z)$ around the point z_0. It gives us a representation of the function $f(z)$, as faithful as that of the Taylor series, that holds in the annular region within which the function is analytic. This representation now involves both positive and negative powers. It is interesting to note that strictly speaking the convergence proofs do not require that the two circles that define the annular region be necessarily concentric, as we assumed here. All we need is two circles, one contained within the other, such that $f(z)$ is analytic within the (possibly irregular) annular region between the two. If the two circles are not concentric, then a different value of z_0 must be used in each one of the two parts of the series, each one of them making reference to the corresponding center. The role of the inner circle is to effectively remove from the inner region the points where the function has singularities. In usual practice this expansion is in fact useful only if we have two concentric circles, a situation that we will be assuming from now on.

Note that if $f(z)$ is analytic within the entire external disk with boundary C_1, then all the integrals that appear in the expressions of the coefficients b_n are zero because of the Cauchy-Goursat theorem, while the integrals that appear in the expressions of the coefficients a_n become the derivatives of $f(z)$ at the point z_0, so that in this case the generalized series that we have obtained is reduced back to the Taylor series. Note also that the coefficient b_1 has a somewhat special character, being given in terms of $f(z)$ by an integral involving only this function without any explicit additional powers of $(z' - z_0)$,

$$\begin{aligned} b_1 &= \frac{1}{2\pi\imath} \oint_{C_2} f(z')\, dz' \quad \Rightarrow \\ \oint_{C_2} f(z)\, dz &= 2\pi\imath b_1, \end{aligned}$$

which allows us to calculate the integral of $f(z)$ by simply finding the coefficient b_1, which may be easy to do if there is only one singularity of the function within the contour C_2. Note also that, once again due to the Cauchy-Goursat theorem, the coefficients a_n and b_n can be written not only in terms of integrals over C_1 and C_2, as shown above, but also in terms of integrals over any contour C, not necessarily circular, which makes exactly one turn around the inner circle, within the annular region between C_1 and C_2, as shown in the diagram of Figure 14.3.

Thus, we see that we can write the Laurent series around z_0 in a unified and more succinct form,

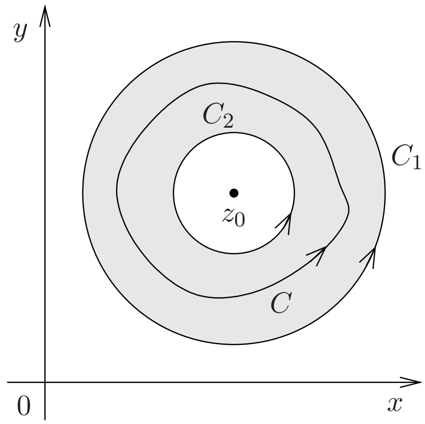


Figure 14.3: The complex plane, showing the two circles C_1 and C_2, the annular region and the contour C contained within it.

$$
\begin{aligned}
f(z) &= \sum_{n=-\infty}^{\infty} A_n (z - z_0)^n, \\
A_n &= \frac{1}{2\pi \imath} \oint_C \frac{f(z')}{(z' - z_0)^{n+1}} \, dz',
\end{aligned}
$$

where, for $n \geq 0$, we have that $A_n = a_n$ and, for $n < 0$, we have that $A_n = b_{-n}$. Just as we did in the case of the convergence disk of the Taylor series, we can think about extending the annular region of convergence or *convergence ring* of the Laurent series, both outward and inward. In each case, we can vary the radii, increasing r_1 and decreasing r_2, until the first singularity of $f(z)$ is reached, in each one of the two directions. However, unlike what happened in the case of the Taylor series, in which the coefficients were written in terms of the values of the function and of its derivatives at a single point, the point z_0 where presumably these values are known, in our case here the coefficients are written as contour integrals, so that in the general case there is no interest in keeping the reference point z_0 fixed during this process of deformation. Therefore we can freely vary the radii of the inner and outer circles, keeping them concentric, possibly also varying the position of their common center, so long as the inner circle still includes all internal singularities, and so long as the outer circle does not contain any singularities external to the initial outer circle.

Regarding the internal singular points, there is a special case which is particularly important. If the function $f(z)$ is analytic throughout the interior of C_1 except for a single point, which therefore constitutes a single *isolated singularity*, then the radius r_2 of the inner circle can be made as small as desired, so long as the center z_0 of this circle approaches indefinitely the singularity. In order to simplify things a bit, we can simply choose the singular point as the common center of the two circles, placing z_0 at that point, and therefore the contour C can be arbitrarily close to the singularity, which is now located at z_0. Thus we see that all the coefficients of the series can be written as integrals over a contour C arbitrarily close to the singularity, so that it is clear that they depend only on the behavior of the function near the singularity. When $f(z)$ has only one such isolated singularity at z_0, we say that the coefficient $b_1 = A_{-1}$ is the *residue* of $f(z)$ at that singular point.

In order to emphasize the importance of this residue, that is, of this particular coefficient of the series, it suffices to note that the integral of the function $f(z)$ on a contour that makes a single turn around the singularity point z_0, and which contains no other singularities, is given in a very simple way by $2\pi\imath b_1$, that is, only by the term b_1. If we had written the integral

$$\oint_C f(z)dz$$

using the representation of $f(z)$ by its Laurent series, and then integrated that Laurent series term-by-term, the integrals of all the other terms would vanish, so that only the term with the coefficient b_1 would remain. Once again, this can be understood geometrically, if we think of the complex integral as a pair of line integrals. As we have seen before, there is a kind of resonance between the turns of the vector $\vec{w}_I$ about itself and the turn of the integration process along the integration contour around the point z_0, which causes the integral of the term with $1/(z - z_0)$ to be the only one different from zero, because only in this case the vector $\vec{w}_I$ rotates in synchronism with the tangent to the integration contour. Thus we see that the value of the closed contour integrals of the function $f(z)$ are closely related to their residues at the singular points.

The remaining question is how to calculate the coefficients of the Laurent series, and in particular the coefficient b_1 which gives the residue in the case of a single singularity. In the case of the Taylor series this problem is rendered trivial by the Cauchy integral formulas for the multiple derivative of the function, by means of which the coefficients can be written in terms of the derivatives of the function at the point z_0. In the case of the Laurent series a general method, as simple as that one, is not available. However, in many cases the coefficients may also be determined in a simple way. Note that in many cases, including most of the cases of greater interest to us, many of the coefficients b_n may be zero, and in fact it is very common in the applications that only a finite number of them is not zero. For example, the singular function

$$f(z) = \frac{1}{(z-1)^2}$$

is already in the form of a Laurent series, with $z_0 = 1$, $A_{-2} = 1$ and $A_n = 0$ for $n \neq -2$. In particular, we have that $A_{-1} = b_1 = 0$, so that we get, once again, a result that we already know. Since the residue is zero we have that, for a closed contour C around $z = 1$,

$$\oint_C \frac{1}{(z-1)^2}\, dz = 0.$$

In many cases the coefficients a_n, b_n or A_n can be found by means other than the direct calculation of their expressions in terms of integrals. One can often obtain the coefficients combining two Taylor or Maclaurin series, or through simple changes of variables. For example, from the Maclaurin series of the exponential we can derive immediately that

$$\frac{e^z}{z^2} = \frac{1}{z^2} + \frac{1}{z} + \frac{1}{2} + \frac{z}{3!} + \frac{z^2}{4!} + \ldots\,,$$

giving us very easily the Laurent series of this function. In particular, we have that $b_1 = 1$, so that for a closed contour C around $z = 0$ we have that

$$\oint_C \frac{e^z}{z^2}\, dz = 2\pi\imath.$$

A less trivial example is obtained from the Maclaurin series of the exponential and a change of variables, by composing the exponential with the function $1/z$, in order to write that

$$\begin{aligned} e^{1/z} &= \sum_{n=0}^{\infty} \frac{1}{n!\, z^n} \\ &= 1 + \frac{1}{z} + \sum_{n=2}^{\infty} \frac{1}{n!\, z^n}. \end{aligned}$$

In particular, we have that $b_1 = 1$, so that we obtain immediately that for a contour C around $z = 0$,

$$\oint_C \mathrm{e}^{1/z}\, dz = 2\pi\imath.$$

Since it is true that the Laurent series of a function, just as was the case for the Taylor series, is unique, any of these legal manipulations produce as the result the Laurent series of the function around the point z_0. In fact, all the most important properties of the positive-power series, such as the Taylor series, can be generalized to the Laurent series, within their domains of convergence.

It is interesting to make here some comments of a more general character, before continuing with our development. Because they always exist and are faithful representations of the analytic functions, we can use the Taylor and Laurent series to prove many important facts about these functions. For example, an important theorem that relates to them is the theorem of analytic continuation (or analytic extension), about which we have talked before. According to this theorem, if two power series, one around z_0 and another around z_1, have convergence disks that intersect at more than one point, and the two series coincide (have the same limits at each point) within the common domain of convergence, then they are representations (unique, each relative to its point of reference) of the same analytic function, each one within its convergence domain. Thus we see that by using power series with convergence disks or rings which intersect, it is possible to extend the definition of an analytic function, in a unique and consistent way, from any region of the complex plane to any other region connected with the first, in which it exists.

Another example is that one can use this fact, or the Taylor series directly, in order to show that the zeros of an analytic function that is not identically zero are isolated, that is, there is an open disk around each one of them within which there is no other zero. This has a certain importance for us because the zeros of an analytic function define the singularities of a rational function in which it appears in the denominator and, as we saw, because isolated singularities are particularly important for the calculation of integrals. However, unlike the case of the zeros, in general the singularities of analytic functions are not necessarily isolated.

We will end this chapter using the Taylor series to show that the zeros of analytic functions are always isolated, and examining some immediate consequences of this fact. Let us consider, therefore, a function $f(z)$ that is analytic at a point z_0. It follows that there is a disk centered at z_0

within which the function is faithfully represented by a Taylor series with respect to that point,

$$\begin{aligned} f(z) &= \sum_{n=0}^{\infty} a_n (z - z_0)^n \\ &= \sum_{n=0}^{\infty} \frac{f^{n\prime}(z_0)}{n!} (z - z_0)^n. \end{aligned}$$

Let us imagine that $f(z)$ has a zero at z_0. Of course, around any point at which $f(z)$ has a zero and is analytic, we can build a Taylor series. In this case, $f(z_0) = 0$ implies that $a_0 = 0$. Now, it may also happen that the function has zero derivative at this point, so that $a_1 = 0$. However, except for the case in which the function is identically zero within the convergence disk of the series, it cannot happen that *all* the derivatives are zero at z_0. This is true because the Taylor series of a function around a given point is unique, so that the Taylor series of the identically zero function is the only one that has all zero coefficients.

Thus, if the function has a zero at z_0 but is not identically zero, it follows that there must be some coefficient a_m that is not zero. When all coefficients up to a_m are zero, but $a_m \neq 0$, we say that the function has a zero of order m at z_0. Thus, the Taylor series of this function around z_0 can be written as

$$\begin{aligned} f(z) &= \sum_{n=m}^{\infty} a_n (z - z_0)^n \\ &= (z - z_0)^m \sum_{k=0}^{\infty} a_{k+m} (z - z_0)^k \\ &= (z - z_0)^m g(z), \end{aligned}$$

where $g(z)$ is an analytic function such that $g(z_0) \neq 0$, since it is represented by a convergent power series, and does *not* have a zero at z_0, since $a_m \neq 0$. Let us imagine that there is a point $z_1 \neq z_0$ within the convergence disk of the series of $f(z)$ in which $f(z_1) = 0$. Since we have $(z_1 - z_0) \neq 0$ it follows that $(z_1 - z_0)^m \neq 0$, so that if $f(z_1) = 0$, then it is necessary that $g(z_1) = 0$. If this is the only other zero of $f(z)$ within the disk centered at z_0 with radius $|z_1 - z_0|$, then it follows that there is a disk around z_0 where there is no other zero of $f(z)$, so that the zero of $f(z)$ at z_0 is an isolated zero. In fact, if there is a zero of $g(z)$ that is the one that is closer to z_0 among all the zeros of $g(z)$, and it is at a point $z_1 \neq z_0$, then it follows that the zero of $f(z)$ at z_0 is an isolated zero.

Thus, the only way for the zero of $f(z)$ at z_0 to *not* be isolated is that there exist zeros of $g(z)$ which are arbitrarily close to z_0, at which point $g(z)$ does *not* have a zero, since we have $g(z_0) = a_m \neq 0$. However, it is easy to see that this would violate the continuity of $g(z)$ at z_0. Since $g(z)$ is analytic at z_0, it follows, in particular, that $g(z)$ is also continuous at that point, meaning that the following limit exists and has the value shown,

$$\begin{aligned} \lim_{z\to z_0} g(z) &= g(z_0) \\ &= a_m \\ &\neq 0. \end{aligned}$$

Formally, what this means is that, given a positive real number ϵ, it is always possible to find a positive real number δ such that, for any z within the disk of radius δ centered at z_0, it is true that

$$|g(z) - g(z_0)| < \epsilon.$$

However, if there are points z_1 arbitrarily close to z_0 at which $g(z_1) = 0$, then the limit does not exist. In order to see this, it is enough to formally write the limiting condition by choosing a real and positive value of ϵ such that $\epsilon < |a_m|$. It becomes clear, then, that there is no disk of radius δ centered at z_0, for any positive real value of δ, such that $|g(z) - g(z_0)|$ is less than ϵ for z within the disk, because no matter how small δ is, there is a point z_1 within the disk at which $g(z_1) = 0$ and therefore at which

$$\begin{aligned} |g(z_1) - g(z_0)| &= |g(z_0)| \\ &= |a_m| \\ &> \epsilon. \end{aligned}$$

Thus we see that if the zero of $f(z)$ at z_0 is not isolated, then the function $g(z)$ cannot be continuous at z_0. However, the function $g(z)$ is clearly continuous, and in fact analytic, at z_0. It follows that the zero of $f(z)$ at z_0 is necessarily isolated. Since everything that was done here holds for any zero of any analytic function that is not identically zero, it follows that an analytic function always has all its zeros isolated unless it is the identically zero function. Note that this constitutes, in fact, a trivial version of the analytic continuation theorem, which allows us to extend a function that is zero on a localized continuous region to the whole complex plane.

Since the identically zero function is a particular case of the constant function, and both are analytic, in fact, on the whole complex plane, the combination of the theorem proved above with the trivial version of the analytic continuation theorem involved in that proof, leads to some interesting consequences. First, as we saw above, there is the fact that an analytic function having a non-isolated zero is always identically zero over the *entire* complex plane. It follows from this that an analytic function that is not identically constant cannot be constant over a curve or a two dimensional region of the complex plane. If a function $f(z)$ is a constant A on a segment of curve, then the analytic function $g(z) = f(z)-A$ has zeros on this segment, and it follows that it is the identically zero function on the whole complex plane. It follows then that $f(z) = A$ on the whole complex plane.

Finally, we can see that if two analytic functions $f_1(z)$ and $f_2(z)$ coincide on a segment of curve, then they are identical. This is true because in this case the analytic function $g(z) = f_1(z) - f_2(z)$ is zero over the segment and it follows that it is the identically zero function over the whole complex plane, that is $f_1(z) = f_2(z)$. Note that if $f_1(z)$ and $f_2(z)$ are the limiting functions of two power series with a common part of their maximum convergence disks, and coincide within this common part, it follows that the two power series converge to the same analytic function in different domains. In fact, this constitutes a simple proof of the analytic continuation theorem.

In conclusion, since an analytic function always has their zeros isolated, a function that has singularities due to the fact that it has an analytic function in the denominator, has the singularities generated in this way also isolated. Although it is possible for an analytic function to have singularities that are not isolated, in most cases of interest they are isolated, and therefore we will be able to handle each of them separately, and to consider their residues for the calculation of closed-contour complex integrals.

Problem Set

1. Consider the function $f(z) = \cos(z)/(z-z_0)$, for an arbitrary value of z_0.

(a) Write the Taylor series of $\cos(z)$ around z_0.

(b) Write the Laurent series of $f(z)$ around z_0.

(c) Determine the radii of the convergence ring of this series.

(d) Determine the value of the residue of $f(z)$ at the point z_0.

Answer: $\cos(z_0)$.

2. Consider the function $f(z) = \ln(z)/(z-z_0)$, for a z_0 whose real part is strictly positive. Consider just the $n = 0$ leaf of the Riemann surface of the logarithm.

(a) Write the Taylor series of $\ln(z)$ around z_0.

(b) Write the Laurent series of $f(z)$ around z_0.

(c) Determine the radii of the convergence ring of this series.

(d) Determine the value of the residue of $f(z)$ at the point z_0.

Answer: $\ln(z_0)$.

3. Consider the function $f(z) = 1/P_n(z)$, where $P_n(z)$ is a polynomial of degree n,

$$P_n(z) = a_0 + a_1 z + a_2 z^2 + \ldots + a_{n-1} z^{n-1} + a_n z^n,$$

with complex coefficients $a_0, \ldots, a_n$, where $a_n \neq 0$ and $n > 0$, and where the coefficients are such that the polynomial has n distinct complex roots $z_1, \ldots, z_n$.

(a) Write the first three terms of the Laurent series of $f(z)$ around z_n.

(b) Determine the radii of the convergence ring of this series.

(c) Determine the value of the residue of $f(z)$ at the point z_n.

Answer:

$$\frac{1}{a_n} \prod_{k=1}^{n-1} \frac{1}{z_n - z_k}.$$

4. Consider the function $f(z) = \sin\left(z^2\right)/z^2$, for $z \neq 0$.

(a) Write the Taylor series of $\sin\left(z^2\right)$ around $z = 0$.

(b) Write a series based on the series obtained above, which represents the function $f(z)$.

(c) Determine the convergence ring or disk of this series, noting that the series of $\sin(z)$ converges on the entire complex plane.

(d) Calculate the limit of $f(z)$ when $z \to 0$.

(e) How should we define the value of $f(z)$ at $z = 0$, in such a way that it becomes analytic at that point?

Chapter 15

Calculation of Integrals by Residues

We will now discuss the issue of the determination of the nature of the singularities and of the corresponding residues of an analytic function, as well as the important *residue theorem*, which gives us a powerful way to calculate certain types of integrals, including certain real integrals. In the general case of the Laurent series, developed in the previous chapter (Chapter 14), the coefficient b_1 was related to an integral of the function on a contour that could contain multiple singularities of the function. However, one could only use this fact to actually calculate the integral with ease in the case in which there was a single isolated singularity within the contour. We will now extend this method for the calculation of integrals to cases in which there are several singularities of the function in the region of interest.

Let us suppose that a function $f(z)$ is analytic in a given region, except at a *finite* number of singular points. If the number of singular points is finite, then these singularities are necessarily isolated from each other. We will also assume that there are no branch cuts involved. Therefore, around each of these singularities there is a convergence ring of a Laurent series that represents the function, as shown in the diagram of Figure 15.1.

The integrals of $f(z)$ around each of these singular points give the *residues*, that is, the coefficients b_1 of each one of the corresponding Laurent series, one in each annular region,

$$2\pi\imath b_{1,j} = \oint_{C_j} f(z)\, dz,$$

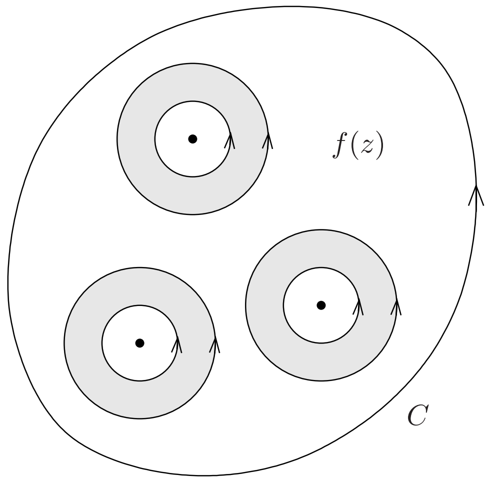


Figure 15.1: The complex plane, showing the isolated singularities of $f(z)$, each with its annular region of convergence, and the integration contour C.

where the index j runs through the set of singular points. It follows that all analytic functions have a residue at each one of its isolated singularities, and such residues may often be determined by very simple methods, without the need of an explicit calculation of the integrals. Let us now consider an integral of $f(z)$ on a contour C which goes around the whole region where the singularities are located. Using the Cauchy-Goursat theorem and manipulating the integration contours we can decompose this integral into a sum of integrals over small contours, with one of these small contours around each one of the singularities, as shown in the diagram of Figure 15.2.

If we consider the contour C' shown in Figure 15.2, we will see that it has no internal singularities, so that by the Cauchy-Goursat theorem the integral over C' is zero. On the other hand, this integral can be written as a sum of integrals involving the original integral over the contour C and the integrals over the small contours C_j,

$$\begin{aligned} \oint_{C'} f(z)\,dz &= \oint_{C} f(z)\,dz - \sum_j \oint_{C_j} f(z)\,dz \\ &= 0. \end{aligned}$$

Therefore, we can write the integral over the contour C in terms of the integrals around each singularity, and we have that

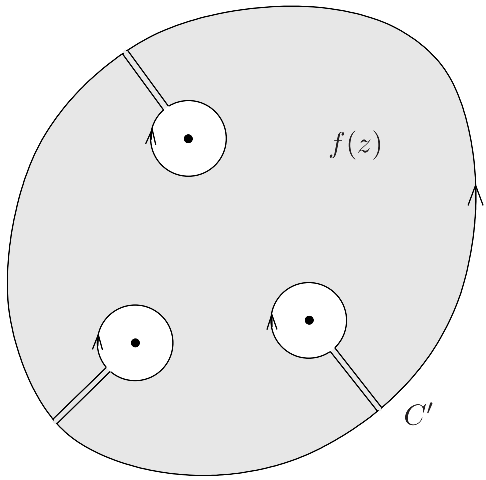


Figure 15.2: The complex plane, showing the isolated singularities of $f(z)$, and the decomposition of the overall integral into integrals on small contours.

$$\begin{aligned}\oint_C f(z)\,dz &= \sum_j \oint_{C_j} f(z)\,dz \\ &= 2\pi\imath \sum_j b_{1,j},\end{aligned}$$

that is, we can write the integral on the contour C in terms of a sum of residues, including a residue for each singularity that lies within C. This is the *residue theorem.* So long as we can determine the residues, it provides us with a tool to calculate integrals. Note that it is not really necessary to have only a finite number of singularities within the contour, provided that all existing singularities be isolated. The restriction to a finite number of singularities was used here only as a way to simplify the argument. If the contour is extended indefinitely in some direction, then it is clearly possible that there may exist within it an infinite number of isolated singularities. In this case, the sum in the result above becomes a numerical series, that is, an infinite discrete sum of complex numbers. In this case, the issue of the convergence of an asymptotic integral is reduced to the convergence problem of this numerical series.

We saw at the beginning of this study of analytic functions that when we calculate integrals around poles of order n given by z^{-n}, we have zero as a result except in the case where $n = 1$. We see now, in a

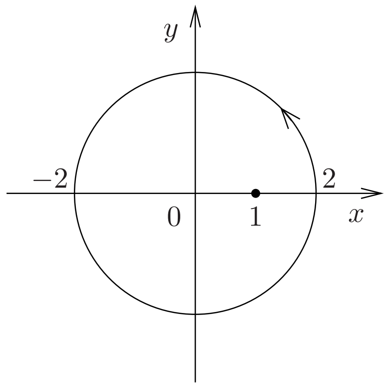


Figure 15.3: The complex plane, showing a simple example of a contour.

very general way, that the contributions to a closed contour integral come exclusively from the singular points z_j at which the Laurent expansion of the function that appear in the integrand has a component $1/(z-z_j)$, and only from this component of each of the expansions.

The residue theorem effectively reduces a definite integration problem to a problem of expansion in series, or rather of the determination of a single term of a series expansion. Since the objective here is to manipulate a power series, this often reduces the integration problem to a problem of differentiation, or even to a simple algebraic problem involving series which are already well known. Sometimes finding the relevant residues is very simple indeed. For example, consider the integral

$$\oint_C \frac{\mathrm{e}^{-z}}{(z-1)^2}\,dz$$

on the contour drawn in Figure 15.3, the circle of radius 2 centered at the origin. We can use the Taylor series of $\exp(-z)$ around $z_0 = 1$ in order to write

$$\begin{aligned}\frac{\mathrm{e}^{-z}}{(z-1)^2} &= \frac{\mathrm{e}^{-1}}{(z-1)^2}\sum_{n=0}^{\infty}(-1)^n\frac{(z-1)^n}{n!}\\ &= \frac{\mathrm{e}^{-1}}{(z-1)^2} - \frac{\mathrm{e}^{-1}}{(z-1)} + \mathrm{e}^{-1}\sum_{n=2}^{\infty}(-1)^n\frac{(z-1)^{n-2}}{n!},\end{aligned}$$

which is the Laurent series of this function, concluding therefore that the residue at the sole singularity at $z=1$, is $-1/e$, so that we have

$$\oint_C \frac{\mathrm{e}^{-z}}{(z-1)^2}\, dz = -\frac{2\pi\imath}{e}.$$

This is a case in which the function has what is called a *pole*, which in this case is of *order* 2, at $z = 1$. Note however that the residue is not associated with the term of the series with power -2, which is the term with the largest negative power, but instead with the term with power -1 exactly. Note that in this simple case the value of the integral can also be obtained from the Cauchy integral formula for the first derivative. A case which is a little more complicated is the following integral on the same contour,

$$\oint_C z\, \mathrm{e}^{1/z^2}\, dz.$$

We write, starting from the Maclaurin series of $\exp(z)$, substituting variables and multiplying by z, the expansion

$$z\, \mathrm{e}^{1/z^2} = z + \frac{1}{z} + \frac{1}{2!\, z^3} + \frac{1}{3!\, z^5} + \frac{1}{4!\, z^7} + \dots\ ,$$

which again is the Laurent series of the function. From this we see that $b_1 = 1$, and therefore we have for the integral

$$\oint_C z\, \mathrm{e}^{1/z^2}\, dz = 2\pi\imath.$$

This is a case in which we say that the function has at the point $z = 0$ an *essential* singularity because there is no limit to the powers of $1/z$ appearing in the expansion. Note that the existence and calculation of the residue are not limited to cases in which the function has only one or more poles, whatever their orders, but also applies to essential singularities.

As shown here, the technical difficulty of the calculation of integrals through the residue theorem is reduced to the calculation or determination, in one way or another, of the residues of the function at the singularities involved. In many cases this can be done simply by algebraic manipulations or operations with the Taylor or Maclaurin series. In order to generalize a little more this technique, one first needs to better systematize the classification of the types of isolated singularities, a classification that we have mentioned before several times. A more complete and universal classification of all possible isolated singularities of analytic functions can be found in [9] (available in Appendix E). For the time being, let us recall that if the Laurent series of a function around a point where it is singular has a finite number of non-zero terms with negative powers, then we say that the singularity is a *pole* and the largest negative power $-m$ existing in the series is the negative of the *order* m of the pole. If $m = 1$, we say that it is a *simple pole*. On the other hand, if the number of terms with negative powers is infinite, then we say that the singularity is not a pole, but an *essential singularity*.

Let us take a rational function as a simple example involving poles. It is easy to verify, by manipulating algebraically the numerator, that

$$\frac{z^2 - 2z + 3}{z - 2} = \frac{3}{z - 2} + 2 + (z - 2),$$

so that we see that this function has a simple pole at $z = 2$, and that the residue at that point is $b_1 = 3$. Note the need to write all terms consistently in terms of $(z - 2)$, in order to generate the series correctly. We thus obtain the integration formula,

$$\oint_C \frac{z^2 - 2z + 3}{z - 2}\, dz = 6\pi \boldsymbol{\imath},$$

for closed contours C around $z = 2$. For another example, consider that the function

$$\frac{\sinh(z)}{z^4} = \frac{1}{z^3} + \frac{1}{3!\, z} + \frac{z}{5!} + \frac{z^3}{7!} + \ldots$$

has a pole of order 3 at $z = 0$, and the residue there is $1/6$. We obtain this time the integration formula

$$\oint_C \frac{\sinh(z)}{z^4}\, dz = \frac{\pi \boldsymbol{\imath}}{3},$$

for closed contours C around $z = 0$. Finally, the function

$$\cosh\left(\frac{1}{z}\right) = 1 + \sum_{n=1}^{\infty} \frac{1}{(2n)!\, z^{2n}}$$

has an essential singularity at $z = 0$, and the residue there is zero, so that this time we have

$$\oint_C \cosh\left(\frac{1}{z}\right)\, dz = 0,$$

for closed contours C around $z = 0$. It is relatively easy to devise many other examples of non-trivial integrals that can be easily calculated using this technique. In some cases there is the need for a bit of algebraic work with the terms of the series, such as in a case involving the function $\exp(1 + z + z^2)/z^3$, for example.

If a function $f(z)$ has a pole of finite order m at z_0, it is always possible to define, from it, a new function $\phi(z)$ that is analytic throughout the analyticity domain of $f(z)$, and that furthermore is also analytic at z_0. The function $\phi(z)$ is defined as follows,

$$\begin{aligned} \phi(z) &= (z - z_0)^m f(z), && \text{for} \quad z \neq z_0, \\ \phi(z_0) &= \lim_{z \to z_0} (z - z_0)^m f(z) && \\ &= b_m, && \text{for} \quad z = z_0, \end{aligned}$$

so that the value of $\phi(z)$ at the point z_0 is defined by a continuity criterion. Note that it is not possible to do this in the case of an essential singularity, because the limit would not exist in that case. Since we have for the Laurent series of $f(z)$ that

$$f(z) = \frac{b_m}{(z - z_0)^m} + \frac{b_{m-1}}{(z - z_0)^{m-1}} + \ldots + \frac{b_1}{(z - z_0)} + \sum_{n=0}^{\infty} a_n (z - z_0)^n,$$

we find that $\phi(z_0) = b_m$, and in addition to this we also have the complete Laurent series (in fact, the Taylor series) of $\phi(z)$,

$$\phi(z) = b_m + b_{m-1}(z - z_0) + \ldots + b_1(z - z_0)^{m-1} + \sum_{n=0}^{\infty} a_n (z - z_0)^{n+m}.$$

Since the series of $f(z)$ is convergent, it follows that the infinite sum appearing in these expressions is convergent. Thus it is clear that this series for $\phi(z)$ is a convergent power series of positive powers that converges to b_m at $z = z_0$, so that $\phi(z)$ is analytic, since it was defined at z_0 so as to be continuous there. We say that the function $\phi(z)$, defined only as the product $(z - z_0)^m f(z)$, has a *removable* singularity, a kind of simple "hole" in its domain, a singularity that can be removed by defining the function by continuity at the singularity point, thus effectively filling up the "hole".

We can now write the residue b_1 of $f(z)$ in terms of the function $\phi(z)$. Since the convergent series of positive powers that converges to $\phi(z)$ must be its Taylor series, we have that

$$b_1(z-z_0)^{m-1} = \frac{\phi^{(m-1)\prime}(z_0)}{(m-1)!}(z-z_0)^{m-1},$$

so that we have for the residue

$$b_1 = \frac{\phi^{(m-1)\prime}(z_0)}{(m-1)!}.$$

This can often be a simple way to calculate the residue for $m > 1$. If one has a simple pole with $m = 1$, then we simply have that

$$\begin{aligned} b_1 &= \phi(z_0) \\ &= \lim_{z\to z_0}(z-z_0)f(z), \end{aligned}$$

which is the definition of $\phi(z_0)$ by continuity. This is a very useful way to calculate the residue of a simple pole, that comes up very often. If the singularity of $f(z)$ is an essential one, then there is no way to calculate the residue other than to obtain the Laurent series of the function. However, if the singularity is a pole, then the expression of the residue in terms of the derivative of $\phi(z)$ is often very helpful.

The technique of integration by residues can be used to calculate certain definite real integrals, which would otherwise be much more difficult to calculate. An example of this are integrals involving trigonometric functions, so long as they extend over the period of these functions. In general, they are integrals of the type

$$\int_0^{2\pi} F[\cos(\theta), \sin(\theta)]\, d\theta,$$

for some rational function F. Integrals of this type can be interpreted as complex integrals on the unit circle on a complex plane described by the variable $z = \rho \exp(\imath\theta)$, with $\rho = 1$ and with θ ranging from 0 to 2π. In order to write the integral in this way, we use the transformation of variables

$$\begin{aligned} z &= \mathrm{e}^{\imath\theta}, \\ dz &= \imath z\, d\theta, \\ \cos(\theta) &= \frac{1}{2}\left(z + \frac{1}{z}\right), \\ \sin(\theta) &= \frac{1}{2\imath}\left(z - \frac{1}{z}\right), \end{aligned}$$

which often result in the integral of a rational function of z on the unit circle. Note that only on the unit circle we can write the functions $\cos(\theta)$ and $\sin(\theta)$ in terms of z in this simple way. If one can locate the singularities of the resulting function and determine their residues, then one can determine the value of the integral. As a simple example of this type of calculation, let us consider the real definite integral

$$I = \int_0^{2\pi} d\theta \, \frac{1}{5 + 4\cos(\theta)}.$$

Making the changes indicated above, we obtain

$$\begin{aligned} I &= \int_0^{2\pi} \imath z \, d\theta \, \frac{1}{\imath z} \frac{1}{5 + 2(z + 1/z)} \\ &= \int_0^{2\pi} dz \, \frac{1}{\imath} \frac{1}{5z + 2(z^2 + 1)} \\ &= -\imath \oint_C dz \, \frac{1}{5z + 2z^2 + 2}, \end{aligned}$$

where the contour C is the unit circle. It is necessary now to factor the polynomial in the denominator. Using the Baskara formula it is not difficult to see that the two roots are -2 and $-1/2$, so that we have

$$\begin{aligned} I &= -\imath \oint dz \, \frac{1}{2z^2 + 5z + 2} \\ &= -\imath \oint dz \, \frac{1}{2(z + 2)(z + 1/2)}. \end{aligned}$$

Only the pole given by $z = -1/2$ is located within the integration contour, so that only its residue b_1 will contribute to the integral. This residue can be calculated very simply by the limit

$$\begin{aligned} b_1 &= \lim_{z \to -1/2} \frac{(z + 1/2)}{2(z + 2)(z + 1/2)} \\ &= \lim_{z \to -1/2} \frac{1}{2(z + 2)} \\ &= \frac{1}{-1 + 4} \\ &= \frac{1}{3}. \end{aligned}$$

It follows that the value of the integral is given, without difficulty, by

$$\begin{aligned} I &= -\imath 2\pi \imath b_1 \\ &= \frac{2\pi}{3}. \end{aligned}$$

Another type of real integral where the method of integration by residues can be used to great advantage are asymptotic integral on the whole real line. This is one of the most common and useful applications in physics problems. As a typical example of the calculation of an integral of this type by the residue theorem, let us consider the asymptotic integral given by

$$I = 2\int_0^\infty \frac{1}{x^2+1}\,dx.$$

Since the integrand is even, we can write for the integral

$$I = \int_{-\infty}^\infty \frac{1}{x^2+1}\,dx.$$

We can now extend the function in the integrand to complex values, so that we have the analytic function

$$\begin{aligned} f(z) &= \frac{1}{z^2+1} \\ &= \frac{1}{(z+\imath)(z-\imath)}, \end{aligned}$$

which has two simple poles, one at $z = \imath$ and the other at $z = -\imath$. Next we try to build a closed contour integral in the complex plane, which in some limit reproduces the real integral we want to calculate. Let us consider, for that purpose, the integration contour C illustrated in the diagram of Figure 15.4, consisting of the real interval $[-R, R]$ and the semicircle C_R of radius R, ranging from $\theta = 0$ to $\theta = \pi$ on the upper half-plane.

In the $R \to \infty$ limit, the integral on the segment $[-R, R]$ of the real line reproduces the integral over the whole real axis, that is, the integration domain of the original real integral,

$$I = \lim_{R\to\infty}\int_{-R}^R \frac{1}{z^2+1}\,dz.$$

The remaining question is what happens in this limit to the integral over the semicircle C_R, that we have to add to the other in order to close the contour,

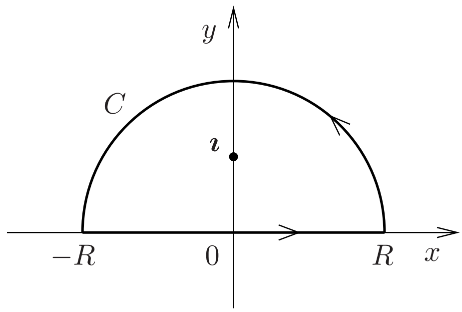


Figure 15.4: The complex plane, showing the contour C formed by the segment and the semicircle, including the relevant pole.

$$I_{C_R} = \int_{C_R} \frac{dz}{z^2+1}.$$

Since over the semicircle we have that $|dz| = R\,d\theta$ and $|z| = R$, we can write for this integral, using the triangle inequalities,

$$\begin{aligned} |I_{C_R}| &= \left| \int_{C_R} \frac{dz}{z^2+1} \right| \\ &\leq \int_{C_R} \frac{|dz|}{|z^2+1|} \\ &= \int_0^{\pi} \frac{R\,d\theta}{|z^2+1|}. \end{aligned}$$

We now have to analyze the behavior of the denominator. Starting from $(z^2+1)-1 = z^2$, we have that $|(z^2+1)-1| = R^2$. Distributing the absolute value on the sum in the left-hand side of this equation, and using the triangle inequality, we can write this inequality as

$$\begin{aligned} R^2 &= |(z^2+1)-1| \\ &\leq |z^2+1| + |-1| \\ &= |z^2+1| + 1. \end{aligned}$$

From this it follows that

$$|z^2+1| \geq R^2 - 1.$$

We can therefore limit our integral from above, exchanging the factor in the denominator for a smaller quantity, $R^2 - 1$. With this we obtain for the integral I_{C_R}

$$\begin{aligned} |I_{C_R}| &\leq \int_0^\pi \frac{R\,d\theta}{|z^2+1|} \\ &\leq \int_0^\pi \frac{R\,d\theta}{R^2-1} \\ &= \frac{R}{R^2-1} \int_0^\pi d\theta \\ &= \pi \frac{R}{R^2-1}. \end{aligned}$$

Since the last form of the upper bound of the integral goes to zero when $R \to \infty$, it follows that the integral over the semicircle goes to zero in this limit, so that the integral on the complex contour becomes in fact equal to the original real integral, which can therefore be written as a complex integral over a closed contour,

$$\begin{aligned} I &= \lim_{R\to\infty} \int_{-R}^{R} \frac{1}{z^2+1}\,dz + \lim_{R\to\infty} \int_{C_R} \frac{1}{z^2+1}\,dz \\ &= \lim_{R\to\infty} \oint_C \frac{1}{z^2+1}\,dz, \end{aligned}$$

where this closed-contour integral does not depend, in fact, on R, so long as it is large enough to contain the singularity of the function at $z = \boldsymbol{\imath}$. The complex contour contains a single simple pole, the pole of the integrand at $z = \boldsymbol{\imath}$, and the residue at this pole is given by

$$\lim_{z\to z_0} (z - \boldsymbol{\imath}) f(z) = \frac{1}{2\boldsymbol{\imath}}.$$

The integral is therefore $2\pi\boldsymbol{\imath}$ times this residue, that is, we have $I = \pi$.

We will finish by commenting on a special situation which is sometimes of interest in the applications. We may wonder what happens if a closed integration contour passes right on top of a singularity. Of course this case is not included in the hypotheses of our theorems about complex integrals on closed contours, so that in principle we cannot say anything about it. In fact, the integral may just not be well defined, and that is all we can say. Depending on the direction in which we cross the singularity, or on the way in which we take limits along the curve, from the two sides of the singularity toward it, the integral may take many different values. A real case analogous to this one would be, for example, the real integral of $1/x$ between -1 and 1,

$$\int_{-1}^{1} dx\,\frac{1}{x}.$$

Since the integral diverges to $+\infty$ on the right-hand side of the point $x = 0$ and to $-\infty$ on the left-hand side, depending on how we take the limits on both sides of this point we can adjust things so that the integral has any real value we want. There is, however, a particularly simple special case, because if we take the limits symmetrically, then the result of the integral is identically zero along the limit, and thus the limit of the integral is zero,

$$\begin{aligned}
\lim_{\epsilon\to 0}\left(\int_{-1}^{-\epsilon} dx\,\frac{1}{x} + \int_{\epsilon}^{1} dx\,\frac{1}{x}\right) &= \lim_{\epsilon\to 0}\left(-\ln|x|\,\Big[_{-1}^{-\epsilon} + \ln|x|\,\Big[_{\epsilon}^{1}\right)\\
&= \lim_{\epsilon\to 0}\left[-\ln(1) - \ln(\epsilon) + \ln(1) + \ln(\epsilon)\right]\\
&= \lim_{\epsilon\to 0} 0\\
&= 0.
\end{aligned}$$

This special way of defining the integral gives it a special value, which we call the *Cauchy principal value*. In many cases, singular integrals can be made regular or *regularized* in this way, through this additional condition on how the limit should be taken, so that they can still be used and manipulated in a meaningful and useful way. Note that we could do something similar to integrate this same function on the whole real line, although the integral diverges both at zero and at $\pm\infty$, because we have that

$$\begin{aligned}
&\lim_{\epsilon\to 0}\lim_{L\to\infty}\left(\int_{-L}^{-\epsilon} dx\,\frac{1}{x} + \int_{\epsilon}^{L} dx\,\frac{1}{x}\right)\\
&= \lim_{\epsilon\to 0}\lim_{L\to\infty}\left(-\ln|x|\,\Big[_{-L}^{-\epsilon} + \ln|x|\,\Big[_{\epsilon}^{L}\right)\\
&= \lim_{\epsilon\to 0}\lim_{L\to\infty}\left[-\ln(L) - \ln(\epsilon) + \ln(L) + \ln(\epsilon)\right]\\
&= \lim_{\epsilon\to 0}\lim_{L\to\infty} 0\\
&= 0,
\end{aligned}$$

irrespective of the order or manner in which one takes the two limits, relative to one another. We can, for example, take simultaneous limits by choosing the relation $L = 1/\epsilon$ between these two parameters.

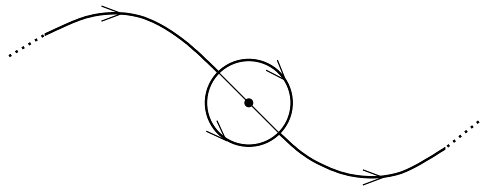


Figure 15.5: The complex plane, showing the integration contour passing over a singularity, and two small deformations which can be used to avoid the singular point.

We will discuss here a version of this idea, known as the Sokhotskii-Plemelj theorem, that applies to integrals over closed contours in the complex plane. Let us, then, imagine that we have an integral over a closed contour that runs right on top of an isolated singularity. We can give a complete definition of the integral by slightly deforming the contour, such that it passes on either side of the singularity. Since the singularity is isolated, we can do this without going through any other singularity. Note that the deformation can be arbitrarily small, and in fact even infinitesimal, as illustrated in the diagram of Figure 15.5.

On the other hand, we can do this in two different ways, since we can go around the singularity by each one of its two sides, as illustrated in Figure 15.5. Additional deformations that keep to one side of the singularity change nothing, but those that cross from one side of the singularity to the other indeed change the integral, because they include in or exclude from the integral, depending on the side on which the integration contour is closed, the contribution due to the residue of the function at that singularity.

Thus we can give a special value to the singular integral, in a very natural and intuitive way, by doing the following: first we deform the contour a bit to one side, and calculate the value of the integral, which becomes well defined; afterwards we deform the contour slightly to the other side, and calculate the integral anew; finally, considering that integration is a linear operation, we calculate the arithmetic average of the two values obtained for the integral. Since by the Sokhotskii-Plemelj theorem [10] the Cauchy principal value of the integral over the contour is equal to this arithmetic average of these two integrals, we may look at this as the Cauchy principal value for this singular complex integral. Since in one of the two cases the singularity does not contribute its residue to the integral, while in the other it contributes $2\pi\imath b_1$, where

b_1 is the residue, on average we will have as a contribution to the integral, from the singularity that is crossed by the contour, the value $\pi \imath b_1$, according to this criterion.

It is as if we had "cut the singularity in half". It should be noted, however, that this separation of the residue in two has a topological character, and not a geometrical one. The shape of the curve that passes over the singular point does not matter, what matters is only that one of the two curves that deviates from the singular point passes by it on one side, and the other on the other side. The original curve can even make an angle, with the vertex at the singular point, and therefore not be differentiable at that point, without this changing anything in this situation. It can be shown that this criterion, applied to the complex function $w(z) = 1/z$, is equivalent to the idea of taking symmetric limits, which was discussed above for the real function $f(x) = 1/x$. This is, of course, a consequence of the Sokhotskii-Plemelj theorem [10]. We will leave a closer examination of this paradigmatic case as an exercise.

It is important to keep in mind that this is not an inevitable value for the integral, but rather a choice, albeit a very special and important choice. As we will see later, this choice allows us to give an accurate and consistent meaning to integrals that, otherwise, would be meaningless, and enables us to deal properly and meaningfully with them.

Problem Set

1. Calculate by residues the following integral,

$$\int_{-\infty}^{\infty} \frac{1}{x^2 + 2x + 2}\, dx.$$

   Consider the following steps.

   (a) Factor completely the polynomial in the denominator.

   (b) Determine how to close the contour and what are the relevant singularities.

   (c) Show that the integral over the additional part of the contour, used to close it, vanishes.

   (d) Calculate the relevant residues and use the residue theorem in order to find the value of the integral.

   Answer: π.

2. Consider the following integral, to be calculated by residues,

$$\int_0^\infty \frac{x^2}{x^4 + 5x^2 + 4}\,dx.$$

Consider the following steps.

(a) Discover how to extend the integral to the interval $(-\infty, \infty)$.

(b) Factor completely the polynomial in the denominator.

(c) Determine how to close the contour in the complex plane and what are the relevant singularities.

(d) Show that the integral over the additional part of the contour, used to close it, vanishes.

(e) Calculate the relevant residues and use the residue theorem in order to find the value of the integral.

Answer: $\pi/6$.

3. Consider the following integral, to be calculated by residues,

$$\int_{-\infty}^\infty \frac{x^2}{x^4 + 3x^2 + 2}\,dx.$$

Consider the following steps.

(a) Factor completely the polynomial in the denominator.

(b) Determine how to close the contour in the complex plane and what are the relevant singularities.

(c) Show that the integral over the additional part of the contour, used to close it, vanishes.

(d) Calculate the relevant residues and use the residue theorem in order to find the value of the integral.

Answer: $\pi\left(\sqrt{2} - 1\right)$.

4. Calculate by residues the following real asymptotic integral,

$$\int_{-\infty}^\infty \frac{\sin(x)}{x^2 + 4x + 5}\,dx.$$

Consider the following steps.

(a) Write $\sin(x)$ in terms of the complex exponentials $\exp(\pm\imath x)$. Note that we will have two integrals to calculate.

(b) Factor completely the polynomial in the denominator.

(c) Determine how to close the contour in each case, and what are the relevant singularities.

(d) Show that the integral over the additional parts of the contours, used to close them, vanish.

(e) Calculate the relevant residues and use the residue theorem in order to find the value of the original integral.

Answer: $-\pi \sin(2)/e$.

5. Calculate by residues the following integral, where $a > 0$,

$$\int_0^\infty \frac{\cos(ax)}{x^2+1}\, dx.$$

Consider the following steps.

(a) Discover how to extend the integral to the interval $(-\infty, \infty)$.

(b) Consider another integral similar to this extended integral, but with $\cos(ax)$ replaced by $\sin(ax)$. What is its value?

(c) Combine the two integrals in order to write a third one, whose calculation is simpler.

(d) Determine how to close the contour and what are the relevant singularities. Consider the relevance of the sign of a.

(e) Show that the integral over the additional part of the contour, used to close it, vanishes.

(f) Calculate the relevant residues and use the residue theorem in order to find the value of the integral.

Answer: $\pi \exp(-|a|)/2$.

6. Calculate by residues the following real asymptotic integral,

$$\int_{-\infty}^\infty \frac{\cos(x)}{x^2+4x+5}\, dx.$$

Consider the following steps.

(a) Write $\cos(x)$ in terms of the complex exponentials $\exp(\pm\imath x)$. Note that we will have two integrals to calculate.

(b) Factor completely the polynomial in the denominator.

(c) Determine how to close the contour in each case, and what are the relevant singularities.

(d) Show that the integrals over the additional parts of the contours, used to close them, vanish.

(e) Calculate the relevant residues and use the residue theorem in order to find the value of the original integral.

Answer: $\pi\cos(2)/e$.

7. Consider the following real asymptotic integral, to be calculated by residues,

$$\int_{-\infty}^{\infty} \frac{\cos(x+1)}{x^2+4x+5}\,dx.$$

Consider the following steps.

(a) Determine and describe one or more closed contours in the complex plane that can be used to compute this integral.

(b) Show that the integrals over the additional parts of the contours, used to close the original contour over the real line, vanish.

(c) Calculate the residues of the relevant singularities and use the residue theorem in order to find the value of the integral.

Answer: $\pi\cos(1)/e$.

8. Calculate by residues the following integral,

$$\int_{0}^{2\pi} \frac{4}{5+4\sin(\theta)}\,d\theta.$$

Consider the following steps.

(a) Write $\sin(\theta)$ in terms of the complex exponentials $\exp(\pm\imath\theta)$.

(b) Change variables in the integral, from θ to $z = \exp(\imath\theta)$. For this purpose write $d\theta$ in terms of dz. Also determine what is the integration contour in the complex z plane.

(c) Write the transformed integral as the integral of a rational function, that is, the ratio of two polynomials on z.

(d) Factor completely the polynomial in the denominator.

(e) Determine which are the relevant singularities.

(f) Calculate the relevant residues and use the residue theorem in order to find the value of the integral.

Answer: $8\pi/3$.

9. Calculate by residues the following integral,

$$\int_{-\pi}^{\pi} \frac{1}{1 + \sin^2(\theta)} \, d\theta.$$

Consider the following steps.

(a) Write $\sin^2(\theta)$ in terms of the complex exponentials $\exp(\pm\imath\theta)$.

(b) Change variables in the integral, from θ to $z = \exp(\imath\theta)$. For this purpose write $d\theta$ in terms of dz. Also determine what is the integration contour in the complex z plane.

(c) Write the transformed integral as the integral of a rational function, that is, the ratio of two polynomials on z.

(d) Factor completely the polynomial in the denominator.

(e) Determine which are the relevant singularities.

(f) Calculate the relevant residues and use the residue theorem in order to find the value of the integral.

Answer: $\pi\sqrt{2}$.

10. Consider the complex function $w(z) = 1/z$, and the unit circle in the complex z plane.

(a) Calculate the closed-contour integral of $w(z)$ on the complete unit circle.

(b) Calculate the open integral of $w(z)$ on the upper semicircle of the unit circle, that is, the integral over the circle between the points $(-1, 0)$ and $(1, 0)$.

(c) Calculate the Cauchy principal value of the closed-contour integral of $w(z)$ on the upper semicircle of the unit circle, closed by the real line segment between the points $(1,0)$ and $(-1,0)$.

(d) Show that the complex version of the criterion for the Cauchy principal value is equivalent to the real version of the criterion, that is, to the use of the symmetric limit defined in the text for the calculation of the real integral between -1 and 1.

11. Consider the complex function $w(z) = 1/z$, and the real axis in the complex z plane. Calculate the integral of the function on the real axis according to the criterion of the Cauchy principal value, that is, according to the definition

$$\begin{aligned} I &= \int_{-\infty}^{\infty} dz\, w(z) \\ &= \lim_{R\to\infty} \int_{-R}^{R} dz\, w(z), \end{aligned}$$

with respect to the asymptotic part, and the Cauchy criterion discussed in the text for the integration through the singularity at the origin. Do this in the two ways listed below and interpret the meaning of the Cauchy principal value of all the integrals that appear.

(a) Do the calculation directly by elementary means, that is, do the real integral using the fundamental theorem of the calculus, for finite R, and then take the limit $R \to \infty$.

(b) Close the contour in the complex plane by means of a circular arc, with finite radius R, consider the integral over this arc of circle, use the residue theorem in order to calculate the integral, and then take the limit $R \to \infty$.

Answer: $I = 0$.

12. Calculate by residues the following integral,

$$\int_{-\infty}^{\infty} \frac{1}{\cosh(x/2)}\, dx.$$

Consider the following steps.

(a) Write $\cosh(x/2)$ in terms of the real exponentials $\exp(\pm x)$.

(b) Extend the variable x to the complex plane of $z = x + \imath y$, and verify to what the integral reduces over two infinite straight lines: the real axis $y = 0$ and the line $y = 2\pi$.

(c) Consider how to close a contour which comprises these two straight lines, in order to enable the calculation of the integral.

(d) Locate the singularities of the function being integrated and determine which ones are relevant.

(e) Show that the integrals over the two vertical segments which are included in order to close the contour vanish when these additional segments are taken to $x \to \pm\infty$.

(f) Assume that the relevant pole is of the first order and calculate the residue by means of the limit that applies to simple poles. Use the residue theorem in order to find the value of the integral.

Answer: 2π.

(g) **(Challenge Item)** Find the Laurent series of the function being integrated, around the relevant pole, including a sufficient number of terms to determine the order of the pole and the value of the residue.

Chapter 16

Residues on Riemann Surfaces

In this chapter we will extend a little more the integration techniques that use the residue theorem. As we saw in the previous chapter (Chapter 15), we can calculate real asymptotic integrals using the residue theorem. A simple example of this is the integral

$$I = \int_0^\infty dx\, \frac{1}{x^2+1}.$$

In this case, due to the fact that the function in the integrand is even on x, we can easily transform this integral into an integral over the whole real line,

$$I = \frac{1}{2}\int_{-\infty}^\infty dx\, \frac{1}{x^2+1}.$$

Then, we extend the integral to the complex plane, noting that the function in the integrand has two simple poles, which are illustrated in Figure 16.1,

$$I = \frac{1}{2}\int_{-\infty}^\infty dz\, \frac{1}{(z-\boldsymbol{\imath})(z+\boldsymbol{\imath})},$$

where the integration contour is still the real line. After that, we close the contour by an arc at infinity and show that the integral over this additional contour is zero. In this way, we then have a closed contour and we can thus calculate the integral by the residue theorem. This is, in essence, the sequence of steps for this integration technique.

Let us now see what happens when we cannot, for some reason, extend the integral to the whole real axis. By doing this we will develop techniques for definite asymptotic integrals on a real semi-axis. A simple example of this is the integral

$$I = \int_0^\infty dx \, \frac{\sqrt{x}}{x^2 + 1}.$$

Not only the function in the integrand does not permit a simple extension to the negative real semi-axis, but also the existence of the two leaves of a Riemann surface in the domain of this function appears to complicate, at first glance, the analysis of the problem. However, what really happens is that the use of the Riemann surface in fact *solves* the problem, as we shall see. First of all, we must verify whether or not this integral converges at the two integration extremes. This is sufficient to establish the existence of the integral, since there are no singularities on the positive real semi-axis. Close to $x = 0$ we have the behavior $\sqrt{x}$, which is finite and integrable. In the asymptotic limit we have the behavior $x^{-3/2}$, which has zero limit and is asymptotically integrable. Thus, the integral certainly exists and is finite. We will then consider the complex integral, inspired by this real integral,

$$\begin{aligned} \bar{I} &= \oint_C dz \, \frac{\sqrt{z}}{z^2 + 1} \\ &= \oint_C dz \, \frac{\sqrt{z}}{(z - \imath)(z + \imath)}, \end{aligned}$$

where the closed integration contour C is that shown in Figure 16.1. Besides the two simple poles shown above, there is, of course, a branch point at $z = 0$, as shown in Figure 16.1. In this figure the two horizontal semi-axes are drawn slightly above and below the real positive semi-axis only for clarity of illustration. We must in fact consider the limit in which these two semi-axes approach indefinitely the positive real semi-axis. Note that, due to the existence of a Riemann surface with two leaves in the domain of the integrand, unlike the contour C the small circular contour of radius ε around the origin, and the large circular contour of radius R, are *not* closed contours over this Riemann surface. However, they can still be expressed as integrals over the parameter θ, a variable which varies in the interval from 0 to 2π. The branch point is at $z = 0$, and we are putting the corresponding branch cut over the positive real semi-axis.

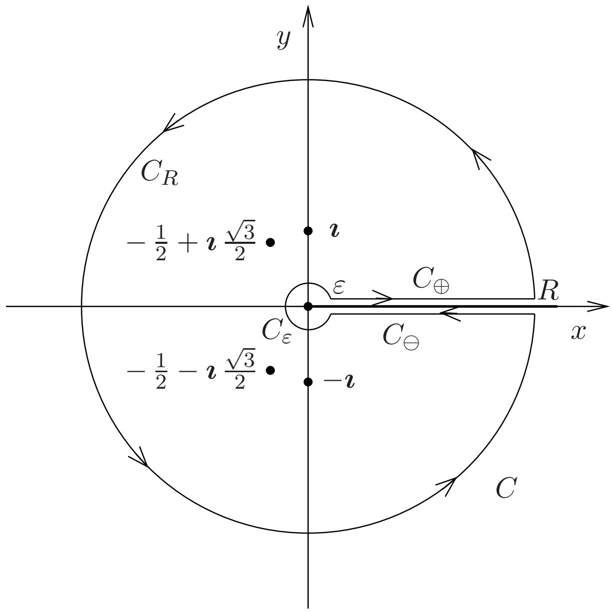


Figure 16.1: The integration contour C on the complex plane, showing the branch point at the origin and the poles involved in the two examples discussed in the text.

Let us now examine the behavior of the integral over each part of the contour C, in the limit in which $\varepsilon \to 0$ and $R \to \infty$. On the small circular contour C_ε around the origin we have $z = \varepsilon \exp(\imath\theta)$ and thus

$$
\begin{aligned}
\lim_{\varepsilon\to 0} \bar{I}_{C_\varepsilon} &= \lim_{\varepsilon\to 0} \int_{C_\varepsilon} dz \, \frac{\sqrt{z}}{z^2+1} \\
&= \lim_{\varepsilon\to 0} \int_0^{2\pi} d\theta \, \imath\varepsilon \, e^{\imath\theta} \, \frac{\sqrt{\varepsilon \, e^{\imath\theta}}}{\varepsilon^2 \, e^{2\imath\theta}+1} \\
&\approx \lim_{\varepsilon\to 0} \int_0^{2\pi} d\theta \, \imath\varepsilon \, e^{\imath\theta} \sqrt{\varepsilon \, e^{\imath\theta}} \\
&= \imath \left(\lim_{\varepsilon\to 0} \varepsilon^{3/2} \right) \left(\int_0^{2\pi} d\theta \, e^{3\imath\theta/2} \right).
\end{aligned}
$$

Since the remaining integral is clearly limited and the limit shown is zero, it follows that this part of the integral is zero in the limit. A similar argument holds for the integral over the large circle C_R, where $z = R\exp(\imath\theta)$, on which we have

$$\begin{aligned}
\lim_{R\to\infty} \bar{I}_{C_R} &= \lim_{R\to\infty} \int_{C_R} dz\, \frac{\sqrt{z}}{z^2+1} \\
&= \lim_{R\to\infty} \int_0^{2\pi} d\theta\, \imath R\, \mathrm{e}^{\imath\theta}\, \frac{\sqrt{R\,\mathrm{e}^{\imath\theta}}}{R^2\,\mathrm{e}^{2\imath\theta}+1} \\
&\approx \lim_{R\to\infty} \int_0^{2\pi} d\theta\, \imath R\, \mathrm{e}^{\imath\theta}\, \frac{\sqrt{R\,\mathrm{e}^{\imath\theta}}}{R^2\,\mathrm{e}^{2\imath\theta}} \\
&= \imath \left(\lim_{R\to\infty} \frac{1}{\sqrt{R}} \right) \left(\int_0^{2\pi} d\theta\, \mathrm{e}^{-\imath\theta/2} \right).
\end{aligned}$$

Once again, since the remaining integral is clearly limited and the limit shown is zero, it follows that this part of the integral is also zero in the limit. Note that we do not insist here on producing a rigorous argument to show that these integrals vanish. There is no need for a lot of care in this case, but if there was any doubt about this, it would not be difficult to take absolute values and use the triangle inequalities, which we will leave as an exercise.

Therefore there remain only the two horizontal parts of the contour C for the completion of this analysis. On the contour $C_\oplus$, which is at $\theta = 0$, and where $z = x$, we have

$$\begin{aligned}
\lim_{R\to\infty} \lim_{\varepsilon\to 0} \bar{I}_{C_\oplus} &= \lim_{R\to\infty} \lim_{\varepsilon\to 0} \int_{C_\oplus} dz\, \frac{\sqrt{z}}{z^2+1} \\
&= \lim_{R\to\infty} \lim_{\varepsilon\to 0} \int_{\varepsilon}^{R} dx\, \frac{\sqrt{x}}{x^2+1} \\
&= \int_0^{\infty} dx\, \frac{\sqrt{x}}{x^2+1},
\end{aligned}$$

given the orientation of the contour, and where we chose, as usual, the positive value of the real square root. Thus we see that in the limit we obtain the original integral I on the contour $C_\oplus$. On the contour $C_\ominus$, which is at $\theta = 2\pi$, and where $z = x$, we have

$$\begin{aligned}
\lim_{R\to\infty} \lim_{\varepsilon\to 0} \bar{I}_{C_\ominus} &= \lim_{R\to\infty} \lim_{\varepsilon\to 0} \int_{C_\ominus} dz\, \frac{\sqrt{z}}{z^2+1} \\
&= \lim_{R\to\infty} \lim_{\varepsilon\to 0} \int_{R}^{\varepsilon} dx\, \frac{-\sqrt{x}}{x^2+1} \\
&= \int_0^{\infty} dx\, \frac{\sqrt{x}}{x^2+1},
\end{aligned}$$

where we took into account the inverted orientation of the contour as well as the fact that, having rotated around the origin to the second leaf of the Riemann surface, we must now choose the negative sign for the square root. Thus, on this contour we obtain once again the original integral I. Bringing together all the parts of the complex integral over the closed contour C, we have as a result the relation

$$\bar{I} = 2I.$$

It follows that the calculation of $\bar{I}$ is sufficient to determine I. We can now calculate the complex integral $\bar{I}$ using the residue theorem, according to which we have

$$\bar{I} = 2\pi\imath(b_{1\oplus} + b_{1\ominus}),$$

since the two poles of the integrand, at $\imath$ and at $-\imath$, are within the closed contour C. It remains only to calculate the residues. The two poles are simple, and we can therefore use the method of the limit, which gives us

$$\begin{aligned}
b_{1\oplus} &= \lim_{z\to\imath}(z-\imath)\,\frac{\sqrt{z}}{(z-\imath)(z+\imath)} \\
&= \lim_{z\to\imath}\frac{\sqrt{z}}{z+\imath} \\
&= \frac{e^{\imath\pi/4}}{2\imath} \\
&= \frac{1}{2\imath}\left(\frac{\sqrt{2}}{2}+\imath\frac{\sqrt{2}}{2}\right) \\
&= \frac{1}{2\imath}\,\frac{\sqrt{2}}{2}\,(1+\imath), \\
b_{1\ominus} &= \lim_{z\to-\imath}(z+\imath)\,\frac{\sqrt{z}}{(z-\imath)(z+\imath)} \\
&= \lim_{z\to-\imath}\frac{\sqrt{z}}{z-\imath} \\
&= \frac{e^{3\imath\pi/4}}{-2\imath} \\
&= \frac{1}{2\imath}\left(\frac{\sqrt{2}}{2}-\imath\frac{\sqrt{2}}{2}\right) \\
&= \frac{1}{2\imath}\,\frac{\sqrt{2}}{2}\,(1-\imath).
\end{aligned}$$

It follows that we have for the complex integral $\bar{I}$

$$\begin{aligned}\bar{I} &= 2\pi\imath\frac{1}{2\imath}\frac{\sqrt{2}}{2}\left[(1+\imath)+(1-\imath)\right]\\ &= \pi\sqrt{2},\end{aligned}$$

and therefore we get for the original real integral I

$$I = \pi\,\frac{\sqrt{2}}{2}.$$

As a partial verification of the correctness of our procedures, we see that we obtain in this way a real and positive value, as was necessary in the case of the real integral in question. Moreover, it is not difficult to make a rough verification by estimating numerically the value of the real integral.

As a second example of an integral over the positive real semi-axis that we cannot extend easily to the whole real axis, we can consider

$$I = \int_0^\infty dx\,\frac{1}{1+x+x^2}.$$

In this case the straightforward extension cannot be made because the integrand has no definite parity. However, it is very interesting that by using the properties of the complex function $\ln(z)$, and of the leaves of the Riemann surface of this function, we can calculate this integral. Let us consider, perhaps in a somewhat surprising way, a kind of extension of this integral to the complex plane in which we *include* a factor of $\ln(z)$ in the integrand,

$$\bar{I} = \oint_C dz\,\frac{\ln(z)}{1+z+z^2},$$

where the contour is the same as before, shown in Figure 16.1. The apparent additional complication due to the introduction of the logarithm is in fact what simplifies and solves the problem. First of all, we determine the singularity structure of the function in the integrand. Of course there is a branch point at $z=0$, due to the presence of the logarithm, and again we put the branch cut on the positive real semi-axis. In order to determine the remaining singularities we must factor the polynomial in the denominator. The Baskara formula gives us for the roots

$$z_\pm = \frac{-1\pm\sqrt{1-4}}{2} \quad\Rightarrow$$

$$
\begin{aligned}
z_{\oplus} &= -\frac{1}{2} + \imath\frac{\sqrt{3}}{2}, \\
z_{\ominus} &= -\frac{1}{2} - \imath\frac{\sqrt{3}}{2}.
\end{aligned}
$$

Note that these roots are on the unit circle and correspond to the angles $\theta = 2\pi/3$ and $\theta = 4\pi/3$. Our complex integral can now be written as

$$
\bar{I} = \oint_C dz \, \frac{\ln(z)}{(z - z_{\oplus})(z - z_{\ominus})},
$$

so that besides the branch point there are two simple poles on the unit circle, which are also illustrated in Figure 16.1. Once again we will analyze the behavior of the integral on each segment of this contour, and once again we will do this without insisting on rigor of proof. On the contour C_ε we have $z = \varepsilon \exp(\imath\theta)$ and thus we get

$$
\begin{aligned}
\lim_{\varepsilon\to 0} \bar{I}_{C_\varepsilon} &= \lim_{\varepsilon\to 0} \int_{C_\varepsilon} dz \, \frac{\ln(z)}{1 + z + z^2} \\
&= \lim_{\varepsilon\to 0} \int_0^{2\pi} \imath z \, d\theta \, \frac{\ln(z)}{1 + z + z^2} \\
&\approx \lim_{\varepsilon\to 0} \int_0^{2\pi} \imath d\theta \, z \ln(z) \\
&= \imath \lim_{\varepsilon\to 0} \int_0^{2\pi} d\theta \, \varepsilon \, \mathrm{e}^{\imath\theta} \left[\ln(\varepsilon) + \imath\theta\right] \\
&= \imath \left[\lim_{\varepsilon\to 0} \varepsilon \ln(\varepsilon)\right] \left(\int_0^{2\pi} d\theta \ \mathrm{e}^{\imath\theta}\right) - \left(\lim_{\varepsilon\to 0} \varepsilon\right) \left(\int_0^{2\pi} d\theta \, \theta \ \mathrm{e}^{\imath\theta}\right).
\end{aligned}
$$

Since the remaining integrals are limited (the first of which is, in fact, zero), and the limits are zero, we see that the integral over this segment is zero in the limit. A similar argument can be used for the contour C_R, where $z = R \exp(\imath\theta)$,

$$
\begin{aligned}
\lim_{R\to\infty} \bar{I}_{C_R} &= \lim_{R\to\infty} \int_{C_R} dz \, \frac{\ln(z)}{1 + z + z^2} \\
&= \lim_{R\to\infty} \int_0^{2\pi} \imath z \, d\theta \, \frac{\ln(z)}{1 + z + z^2} \\
&\approx \lim_{R\to\infty} \int_0^{2\pi} \imath d\theta \, \frac{\ln(z)}{z} \\
&= \imath \lim_{R\to\infty} \int_0^{2\pi} d\theta \ \mathrm{e}^{-\imath\theta} \left[\frac{\ln(R)}{R} + \frac{\imath\theta}{R}\right]
\end{aligned}
$$

$$
\begin{aligned}
= \; & \imath \left[\lim_{R\to\infty} \frac{\ln(R)}{R} \right] \left(\int_0^{2\pi} d\theta \; \mathrm{e}^{-\imath\theta} \right) \\
& - \left(\lim_{R\to\infty} \frac{1}{R} \right) \left(\int_0^{2\pi} d\theta \, \theta \; \mathrm{e}^{-\imath\theta} \right).
\end{aligned}
$$

Again the remaining integrals are limited (the first of which is zero), and the limits are zero, so that we see that the integral over this segment is also zero in the limit. We will now consider the integral over the upper straight contour $C_\oplus$, where $z = x$. Given the orientation of the contour, and assuming that in this case we are at the central leaf of the Riemann surface of the logarithm, we simply have

$$
\bar{I}_{C_\oplus} = \int_\varepsilon^R dx \, \frac{\ln(x)}{1 + x + x^2}.
$$

In the lower straight contour $C_\ominus$, considering the orientation and the fact that we vary θ from 0 to 2π to get to it, and therefore that we are in this way on the next leaf of the Riemann surface of the logarithm, we have

$$
\begin{aligned}
\bar{I}_{C_\ominus} &= -\int_\varepsilon^R dx \, \frac{\ln(x) + 2\pi\imath}{1 + x + x^2} \\
&= -\int_\varepsilon^R dx \, \frac{\ln(x)}{1 + x + x^2} - 2\pi\imath \int_\varepsilon^R dx \, \frac{1}{1 + x + x^2}.
\end{aligned}
$$

Adding the two integrals over the straight contours we find that the terms containing $\ln(x)$ cancel each other, so that we are left with

$$
\bar{I}_{C_\oplus} + \bar{I}_{C_\ominus} = -2\pi\imath \int_\varepsilon^R dx \, \frac{1}{1 + x + x^2}.
$$

In the limit in which $R \to \infty$ and $\varepsilon \to 0$ the remaining integral converges to the real integral with which we originally started. This fact and the cancellation of the part containing the logarithm are the central facts that allow us to calculate the original real integral. Adding all the parts of the complex integral $\bar{I}$ and taking the limit, in which case only the two straight segments give non-zero contributions, we have therefore the relation

$$
\lim_{R\to\infty} \lim_{\varepsilon\to 0} \bar{I} = -2\pi\imath I.
$$

Since the integral $\bar{I}$ does not change with the limit, provided that R is sufficiently large and ε sufficiently small so that the two poles are within the contour C, we can write

$$I = -\frac{\bar{I}}{2\pi\imath},$$

with the understanding that the integral $\bar{I}$ is to be calculated by residues, including the contribution of the two poles. Therefore, let us carry out this calculation. The function in the integrand can be written as

$$f(z) = \frac{\ln(z)}{(z - z_\oplus)(z - z_\ominus)}.$$

Since the two poles are simple, we can calculate the residues by the method of the limit,

$$\begin{aligned}
b_{1\oplus} &= \lim_{z \to z_\oplus} (z - z_\oplus) \frac{\ln(z)}{(z - z_\oplus)(z - z_\ominus)} \\
&= \lim_{z \to z_\oplus} \frac{\ln(z)}{(z - z_\ominus)} \\
&= \frac{\ln(z_\oplus)}{z_\oplus - z_\ominus}, \\
b_{1\ominus} &= \lim_{z \to z_\ominus} (z - z_\ominus) \frac{\ln(z)}{(z - z_\oplus)(z - z_\ominus)} \\
&= \lim_{z \to z_\ominus} \frac{\ln(z)}{(z - z_\oplus)} \\
&= \frac{\ln(z_\ominus)}{z_\ominus - z_\oplus}.
\end{aligned}$$

We have therefore for the complex integral $\bar{I}$

$$\begin{aligned}
\bar{I} &= 2\pi\imath(b_{1\oplus} + b_{1\ominus}) \\
&= 2\pi\imath \left[\frac{\ln(z_\oplus)}{z_\oplus - z_\ominus} - \frac{\ln(z_\ominus)}{z_\oplus - z_\ominus} \right] \\
&= \frac{2\pi\imath}{z_\oplus - z_\ominus} \left[\ln(z_\oplus) - \ln(z_\ominus) \right].
\end{aligned}$$

Using the previously calculated values for $z_\oplus$ and $z_\ominus$, we have for the difference of the two values

$$z_\oplus - z_\ominus = \imath\sqrt{3}.$$

In order to calculate the logarithms we use the polar form of $z_\oplus$ and $z_\ominus$, which results in

$$\begin{aligned}
\ln(z_\oplus) - \ln(z_\ominus) &= \ln\left(\mathrm{e}^{2\pi\imath/3}\right) - \ln\left(\mathrm{e}^{4\pi\imath/3}\right) \\
&= \frac{2\pi\imath}{3} - \frac{4\pi\imath}{3} \\
&= -\frac{2\pi\imath}{3}.
\end{aligned}$$

We therefore have for $\bar{I}$

$$\begin{aligned}
\bar{I} &= -\frac{2\pi\imath}{\imath\sqrt{3}}\frac{2\pi\imath}{3} \\
&= -\frac{(2\pi\imath)^2}{\imath 3\sqrt{3}},
\end{aligned}$$

and it thus follows that we have for I

$$\begin{aligned}
I &= \frac{1}{2\pi\imath}\frac{(2\pi\imath)^2}{\imath 3\sqrt{3}} \\
&= \frac{2\pi\imath}{\imath 3\sqrt{3}} \\
&= \frac{2\pi}{3\sqrt{3}}.
\end{aligned}$$

As one can see, the final result is real and positive, as is to be expected. It is also not difficult to make an approximate numerical verification of the result.

It is not difficult to find a more general set of functions for which this second instance of this method of calculation works. The essential conditions are that the asymptotic integral converge, that the integral vanish on the arc at infinity, that is, on C_R in the limit $R \to \infty$, that the integrand be analytic on the whole complex plane except for isolated singularities, that it be a single-valued function, and that it have no singularities on the positive real semi-axis, including the zero. Therefore, if we have a polynomial $P_n(x)$ of order $n \geq 2$, which has no roots on the positive real semi-axis, including the zero, then the integral

$$I = \int_0^\infty dx\, \frac{1}{P_n(x)}$$

converges and can be calculated using this method. This function has at most n poles, and the integral can be written in terms of the residues at these poles. The number of poles may be smaller than n if some of them are multiple poles. Also, by multiplying this function by any function which is analytic on the whole complex plane and that does not diverge at infinity faster than $P_n(x)$, we immediately have other cases, for example

$$I = \int_0^\infty dx \, \frac{Q_m(x)}{P_n(x)},$$

where $Q_m(x)$ is a polynomial of order $m \leq n - 2$. More generally, the asymptotic integral on the positive real semi-axis of any single-valued function that is analytic throughout the complex plane except for isolated singularities, whose integral over the arc at infinity vanishes, and that has no singularities on the positive real semi-axis and at zero, can be calculated in this way.

Going back to our first example, it is also not difficult to identify more general cases where that instance of the technique works. These are cases where it is not necessary to include the logarithm function, because it is already indirectly present in the integrand. A simple generalization of the first example is that given by the integral

$$I = \int_0^\infty dx \, \frac{x^\alpha}{x^2 + 1},$$

with $0 < \alpha < 1$. We will make this case the last example that we will deal with in detail here. The complex integral, always on the same contour C shown in Figure 16.1, is given by

$$\begin{aligned} \bar{I} &= \oint_C dz \, \frac{z^\alpha}{z^2 + 1} \\ &= \oint_C dz \, \frac{\mathrm{e}^{\alpha \ln(z)}}{z^2 + 1} \\ &= \oint_C dz \, \frac{\mathrm{e}^{\alpha \ln(z)}}{(z - \imath)(z + \imath)}, \end{aligned}$$

where we used the definition of a real power in terms of the exponential and of the logarithm, and we explicitly exhibit the poles of the integrand. Let us now consider the complex integral on each one of the segments of the contour C. For the segment C_ε we have

$$\lim_{\varepsilon \to 0} \bar{I}_{C_\varepsilon} = \lim_{\varepsilon \to 0} \int_{C_\varepsilon} dz \, \frac{z^\alpha}{z^2 + 1}$$

$$
\begin{aligned}
&\approx \lim_{\varepsilon\to 0}\int_{C_\varepsilon} \boldsymbol{\imath} z\, d\theta\, z^\alpha \\
&= \boldsymbol{\imath} \lim_{\varepsilon\to 0}\int_0^{2\pi} d\theta\, \varepsilon^{1+\alpha}\ \mathrm{e}^{(1+\alpha)\boldsymbol{\imath}\theta} \\
&= \boldsymbol{\imath}\left(\lim_{\varepsilon\to 0}\varepsilon^{1+\alpha}\right)\left[\int_0^{2\pi} d\theta\ \mathrm{e}^{(1+\alpha)\boldsymbol{\imath}\theta}\right].
\end{aligned}
$$

Since the integral is limited and the limit shown is zero, we have that this part of the integral vanishes in the limit. Similarly, for the segment C_R we have

$$
\begin{aligned}
\lim_{R\to\infty}\bar{I}_{C_R} &= \lim_{R\to\infty}\int_{C_\varepsilon} dz\,\frac{z^\alpha}{z^2+1} \\
&\approx \lim_{R\to\infty}\int_{C_\varepsilon} \boldsymbol{\imath} z\, d\theta\,\frac{z^\alpha}{z^2} \\
&= \boldsymbol{\imath}\lim_{R\to\infty}\int_0^{2\pi} d\theta\,\frac{1}{\varepsilon^{1-\alpha}}\ \mathrm{e}^{-(1-\alpha)\boldsymbol{\imath}\theta} \\
&= \boldsymbol{\imath}\left(\lim_{R\to\infty}\frac{1}{\varepsilon^{1-\alpha}}\right)\left[\int_0^{2\pi} d\theta\ \mathrm{e}^{-(1-\alpha)\boldsymbol{\imath}\theta}\right].
\end{aligned}
$$

Again the integral is limited and the limit shown is zero so that this part of the integral also vanishes in the limit. For the segment $C_\oplus$ we simply have

$$
\begin{aligned}
\bar{I}_{C_\oplus} &= \int_{C_\oplus} dz\,\frac{z^\alpha}{z^2+1} \\
&= \int_\varepsilon^R dx\,\frac{x^\alpha}{x^2+1},
\end{aligned}
$$

that converges to our original real integral in the limit in which $R\to\infty$ and $\varepsilon\to 0$. For the segment $C_\ominus$, on the other hand, we have

$$
\begin{aligned}
\bar{I}_{C_\ominus} &= \int_{C_\ominus} dz\,\frac{\mathrm{e}^{\alpha[\ln(z)+2\pi\boldsymbol{\imath}]}}{z^2+1} \\
&= \int_{C_\ominus} dz\,\frac{z^\alpha\,\mathrm{e}^{2\pi\boldsymbol{\imath}\alpha}}{z^2+1} \\
&= \int_R^\varepsilon dx\ \mathrm{e}^{2\pi\boldsymbol{\imath}\alpha}\ \frac{x^\alpha}{x^2+1} \\
&= -\,\mathrm{e}^{2\pi\boldsymbol{\imath}\alpha}\int_\varepsilon^R dx\,\frac{x^\alpha}{x^2+1},
\end{aligned}
$$

where we take into account the orientation of the contour and the fact that we moved to the next leaf of the Riemann surface of the logarithm. We have therefore for the sum of the two integrals over the straight contours

$$\begin{aligned}\bar{I}_{C_\oplus} + \bar{I}_{C_\ominus} &= \int_\varepsilon^R dx\,\frac{x^\alpha}{x^2+1} - \mathrm{e}^{2\pi\imath\alpha}\int_\varepsilon^R dx\,\frac{x^\alpha}{x^2+1} \\ &= \left(1-\mathrm{e}^{2\pi\imath\alpha}\right)\int_\varepsilon^R dx\,\frac{x^\alpha}{x^2+1},\end{aligned}$$

which is therefore proportional to our original real integral in the limit in which $R \to \infty$ and $\varepsilon \to 0$. Taking the limits and taking into consideration the fact that the other parts of the integral go to zero, we obtain

$$\begin{aligned}\lim_{R\to\infty}\lim_{\varepsilon\to 0}\bar{I} &= \left(1-\mathrm{e}^{2\pi\imath\alpha}\right)\int_0^\infty dx\,\frac{x^\alpha}{x^2+1} \\ &= \left(1-\mathrm{e}^{2\pi\imath\alpha}\right) I.\end{aligned}$$

We can now calculate the complex integral by residues, without further reference to R or ϵ, except for the fact that they must be such that the two poles are within the closed contour C. Since we have two simple poles, they are given by the limits

$$\begin{aligned}b_{1\oplus} &= \lim_{z\to\imath}(z-\imath)\,\frac{\mathrm{e}^{\alpha\ln(z)}}{(z-\imath)(z+\imath)} \\ &= \lim_{z\to\imath}\frac{\mathrm{e}^{\alpha\ln(z)}}{z+\imath} \\ &= \frac{\mathrm{e}^{\alpha\ln[\exp(\imath\pi/2)]}}{2\imath} \\ &= \frac{\mathrm{e}^{\imath\pi\alpha/2}}{2\imath}, \\ b_{1\ominus} &= \lim_{z\to-\imath}(z+\imath)\,\frac{\mathrm{e}^{\alpha\ln(z)}}{(z-\imath)(z+\imath)} \\ &= \lim_{z\to-\imath}\frac{\mathrm{e}^{\alpha\ln(z)}}{z-\imath} \\ &= \frac{\mathrm{e}^{\alpha\ln[\exp(3\imath\pi/2)]}}{-2\imath} \\ &= -\frac{\mathrm{e}^{3\imath\pi\alpha/2}}{2\imath}.\end{aligned}$$

We have therefore for the complex integral

$$\begin{aligned}
\lim_{R\to\infty}\lim_{\varepsilon\to 0}\bar{I} &= (2\pi\imath)\left(\frac{\mathrm{e}^{\imath\pi\alpha/2}}{2\imath}-\frac{\mathrm{e}^{3\imath\pi\alpha/2}}{2\imath}\right)\\
&= \pi\ \mathrm{e}^{2\imath\pi\alpha/2}\left(\mathrm{e}^{-\imath\pi\alpha/2}-\ \mathrm{e}^{\imath\pi\alpha/2}\right)\\
&= -2\pi\imath\ \mathrm{e}^{\imath\pi\alpha}\ \sin\!\left(\frac{\pi\alpha}{2}\right),
\end{aligned}$$

and therefore we get for our original real integral

$$\begin{aligned}
I &= \frac{1}{1-\ \mathrm{e}^{2\pi\imath\alpha}}\lim_{R\to\infty}\lim_{\varepsilon\to 0}\bar{I}\\
&= -2\pi\imath\sin\!\left(\frac{\pi\alpha}{2}\right)\frac{\mathrm{e}^{\imath\pi\alpha}}{1-\ \mathrm{e}^{2\pi\imath\alpha}}\\
&= 2\pi\imath\sin\!\left(\frac{\pi\alpha}{2}\right)\frac{1}{\mathrm{e}^{\pi\imath\alpha}-\ \mathrm{e}^{-\imath\pi\alpha}}\\
&= \pi\,\frac{\sin\!\left(\frac{\pi\alpha}{2}\right)}{\sin(\pi\alpha)}\\
&= \pi\,\frac{\sin\!\left(\frac{\pi\alpha}{2}\right)}{2\sin\!\left(\frac{\pi\alpha}{2}\right)\cos\!\left(\frac{\pi\alpha}{2}\right)}\\
&= \frac{\pi}{2\cos\!\left(\frac{\pi\alpha}{2}\right)}.
\end{aligned}$$

This is a regular result, without singularities within the open interval $(0,1)$ postulated for α. It is not difficult to see that for $\alpha = 1/2$ this result reduces to the result of our first example in this chapter.

Once again it is not difficult to generalize this instance of the integration technique by residues to other collections of integrals. For example, using the same polynomial $P_n(x)$ with $n \geq 2$ discussed above, we can construct the class of integrals

$$I = \int_0^\infty dx\,\frac{x^\alpha}{P_n(x)}.$$

We can also construct examples of cases in which there are double poles, such as that of the integral

$$I = \int_0^\infty dx\,\frac{x^\alpha}{(x+1)^2},$$

with $0 < \alpha < 1$, involving a pole of second order, which we will leave as an exercise. Yet another example, which interpolates between this and the last of the cases that we have examined in detail here, is given by

$$I = \int_0^\infty dx\, \frac{x^\alpha}{x^2 + \beta x + 1},$$

with $0 < \alpha < 1$ and $0 \leq \beta \leq 2$. For $\beta = 0$ we have the example we worked out before, and for $\beta = 2$ we have the example with a double pole mentioned just above, which was left as an exercise. When we change β continuously from one value to another, the two simple poles move along the unit circle, approaching each other and then coalescing into a single double pole at $z = -1$. We will leave this case as an exercise as well.

Problem Set

1. Consider the first complex integral that was discussed in the text,

$$\bar{I} = \oint_C dz\, \frac{\sqrt{z}}{z^2 + 1},$$

on the closed contour C defined in the text.

   (a) Show with all rigor that the part of the integral over the segment C_ε of the contour is zero in the limit $\varepsilon \to 0$.

   (b) Show with all rigor that the part of the integral over the segment C_R of the contour is zero in the limit $R \to \infty$.

Hint: take absolute values and use the triangle inequalities.

2. Consider the second complex integral that was discussed in the text,

$$\bar{I} = \oint_C dz\, \frac{\ln(z)}{1 + z + z^2},$$

on the closed contour C defined in the text.

   (a) Show with all rigor that the part of the integral over the segment C_ε of the contour is zero in the limit $\varepsilon \to 0$.

(b) Show with all rigor that the part of the integral over the segment C_R of the contour is zero in the limit $R \to \infty$.

Hint: take absolute values and use the triangle inequalities.

3. Consider the third complex integral that was discussed in the text,

$$\bar{I} = \oint_C dz \, \frac{z^\alpha}{z^2 + 1},$$

with $0 < \alpha < 1$, on the closed contour C defined in the text.

(a) Show with all rigor that the part of the integral over the segment C_ε of the contour is zero in the limit $\varepsilon \to 0$.

(b) Show with all rigor that the part of the integral over the segment C_R of the contour is zero in the limit $R \to \infty$.

Hint: take absolute values and use the triangle inequalities.

4. Calculate by residues the real asymptotic integral given by

$$I = \int_0^\infty dx \, \frac{\alpha + \beta x}{1 + x + x^2 + x^3},$$

for arbitrary real α and β. Justify explicitly every step of the procedure.

Answer: $\pi(\alpha + \beta)/4$.

5. Calculate by residues the real asymptotic integral given by

$$I = \int_0^\infty dx \, \frac{x^\alpha}{(x+1)^2},$$

for a real constant α in the open interval $(0, 1)$. Justify explicitly every step of the procedure.

Answer: $\pi\alpha/[\sin(\pi\alpha)]$.

6. Calculate by residues the real asymptotic integral given by

$$I = \int_0^\infty dx \, \frac{x^\alpha}{1 + \beta x + x^2},$$

for real constants $0 < \alpha < 1$ and $0 \le \beta \le 2$. Justify explicitly every step of the procedure. Discuss in particular the cases $\beta = 0$ and $\beta = 2$.

Answer:

$$\begin{aligned} I &= \pi \, \frac{\sin[\alpha(\pi - \theta_\oplus)]}{\sin(\pi\alpha)\sin(\theta_\oplus)}, \\ \sin(\theta_\oplus) &= \sqrt{1 - (\beta/2)^2}, \\ \cos(\theta_\oplus) &= -\beta/2. \end{aligned}$$

7. Calculate by residues the real asymptotic integral given by

$$I = \int_0^\infty dx \, \frac{x^\alpha}{1 + 2x^2 + x^4},$$

for a real constant α in the open interval $(0, 3)$. Justify explicitly every step of the procedure.

Answer: $-\pi(\alpha - 1)/[4\cos(\pi\alpha/2)]$.

8. **(Challenge Problem)** Calculate by residues the real asymptotic integral given by

$$I = \int_0^\infty dx \, \frac{\ln(x)}{x^2 + 1}.$$

Justify explicitly every step of the procedure.

Answer: 0.

Hint: consider a contour contained in the upper half-plane.

Appendix A

Continued Fractions

We will see here how to use a sequence of finite continued fractions in order to represent square roots of integers. The arithmetic operations required for this process involve only integers and rational numbers. In this way we will construct representations of some irrational numbers (in fact, some algebraic numbers, which are the roots of polynomials with integer coefficients), as the limits of sequences of rational numbers. These representations are examples of the construction of the real numbers starting from the rational numbers.

Given an integer p and therefore another integer $n = p^2 + 1$, we can build a continued fraction involving $\sqrt{n}$ in the following way. It is easy to show directly that

$$p + \sqrt{n} = 2p + \frac{1}{p + \sqrt{n}}.$$

We can iterate this formula by replacing the expression in the denominator by the complete formula, obtaining for example, in the first iteration,

$$p + \sqrt{n} = 2p + \cfrac{1}{2p + \cfrac{1}{p + \sqrt{n}}} .$$

We can actually proceed with the iteration indefinitely, producing something like

$$p + \sqrt{n} = 2p + \cfrac{1}{2p + \cfrac{1}{2p + \cfrac{1}{2p + \cfrac{1}{\cdots}}}} .$$

Note that in the infinite iteration limit, the right-hand side of this equation does not depend on n, and thus the equation becomes a way to determine $\sqrt{n}$ in terms of p. Of course, the operations indicated on the right-hand side of the equation are infinite in number and are in the reverse order, in the sense that it is not immediately clear how the result could be approximated by a sequence of finite operations.

We can overcome this obstacle considering a sequence of finite continued fractions, each with a given number of iterations, that converges to the infinite fraction defined above. Consider then the sequence of relations

$$
\begin{aligned}
p+\sqrt{n} &= 2p+\frac{1}{p+\sqrt{n}}, \\
p+\sqrt{n} &= 2p+\cfrac{1}{2p+\cfrac{1}{p+\sqrt{n}}}, \\
p+\sqrt{n} &= 2p+\cfrac{1}{2p+\cfrac{1}{2p+\cfrac{1}{p+\sqrt{n}}}}, \\
\cdots & \quad \cdots .
\end{aligned}
$$

Each of these equations is satisfied exactly, as can be verified algebraically. We can now use as an approximation for $\sqrt{n}$, on the right-hand side of each equation, $\sqrt{n} \approx p$, which produces a sequence of approximations for $\sqrt{n}$,

$$
\begin{aligned}
p+\sqrt{n}_{(0)} &= 2p, \\
p+\sqrt{n}_{(1)} &= 2p+\frac{1}{2p}, \\
p+\sqrt{n}_{(2)} &= 2p+\cfrac{1}{2p+\cfrac{1}{2p}}, \\
p+\sqrt{n}_{(3)} &= 2p+\cfrac{1}{2p+\cfrac{1}{2p+\cfrac{1}{2p}}}, \\
\cdots & \quad \cdots ,
\end{aligned}
$$

where we added at the beginning of the sequence the approximation of order zero. This sequence converges quickly to $p+\sqrt{n}$, that is, the sequence

$$
\begin{aligned}
&p,\\
&p+\frac{1}{2p},\\
&p+\cfrac{1}{2p+\cfrac{1}{2p}},\\
&p+\cfrac{1}{2p+\cfrac{1}{2p+\cfrac{1}{2p}}},\\
&\ldots
\end{aligned}
$$

of manifestly rational numbers converges quickly to $\sqrt{n}$. This allows us to calculate the square roots of some integers, those which can be written as p^2+1, such as $\sqrt{2}$, $\sqrt{3}$, $\sqrt{10}$, $\sqrt{17}$, etc. These are the square roots of the integers succeeding the perfect squares.

This process can be generalized to all integers as follows. Let us consider any integer n. It can be written in the form $n = p^2 + q$, where p^2 is the largest perfect square less than or equal to n and q is limited to the interval $[0, 2p)$, such that q interpolates between p^2 and $p^2 + 2p = (p+1)^2 - 1$, which is the number that precedes the next perfect square. We can now define a generalized continued fraction, first showing that

$$p+\sqrt{n} = 2p + \frac{q}{p+\sqrt{n}},$$

and iterating to obtain

$$p+\sqrt{n} = 2p + \cfrac{q}{2p+\cfrac{q}{p+\sqrt{n}}},$$

and so on. The infinite continued fraction is now given by

$$p+\sqrt{n} = 2p + \cfrac{q}{2p+\cfrac{q}{2p+\cfrac{q}{2p+\cfrac{q}{\ldots}}}},$$

which determines $\sqrt{n}$ in terms of p and q. The sequence of finite continued fractions is given now by

$$
\begin{aligned}
p+\sqrt{n} &= 2p+\frac{q}{p+\sqrt{n}}, \\
p+\sqrt{n} &= 2p+\cfrac{q}{2p+\cfrac{q}{p+\sqrt{n}}}, \\
p+\sqrt{n} &= 2p+\cfrac{q}{2p+\cfrac{q}{2p+\cfrac{q}{p+\sqrt{n}}}}, \\
\cdots & \quad \cdots,
\end{aligned}
$$

and using the approximation $\sqrt{n} \approx p$ on the right-hand side of each equation we have the following approximations for $\sqrt{n}$,

$$
\begin{aligned}
p+\sqrt{n}_{(0)} &= 2p, \\
p+\sqrt{n}_{(1)} &= 2p+\frac{q}{2p}, \\
p+\sqrt{n}_{(2)} &= 2p+\cfrac{q}{2p+\cfrac{q}{2p}}, \\
p+\sqrt{n}_{(3)} &= 2p+\cfrac{q}{2p+\cfrac{q}{2p+\cfrac{q}{2p}}}, \\
\cdots & \quad \cdots.
\end{aligned}
$$

Again, this sequence converges quickly to $p+\sqrt{n}$, that is, the sequence

$$
\begin{aligned}
& p, \\
& p+\frac{q}{2p}, \\
& p+\cfrac{q}{2p+\cfrac{q}{2p}}, \\
& p+\cfrac{q}{2p+\cfrac{q}{2p+\cfrac{q}{2p}}}, \\
& \cdots
\end{aligned}
$$

of manifestly rational numbers converges quickly to $\sqrt{n}$.

For $q=0$ we return to the perfect squares, that is, $\sqrt{n}=p$, and for $q=1$ we return to the previously examined particular case. When $q=0$ the sequence immediately converges in the first step. In the general case the sequence always converges very quickly. The smaller q becomes and the larger p becomes, that is, the smaller the ratio $q/(2p)$ becomes, the faster it converges. Since $q \leq 2p$, this ratio is always less than or equal to 1.

The `continued_fractions` program [4] implements this algorithm. Given a value of n, it determines the corresponding values of p and q (the program refers to p as `i` and to q as `j`) using the following algorithm involving integer truncation:

```
p=int(sqrt(real(n))), q=n-p**2.
```

Of all the possible values of n the case in which the convergence is the slowest is $n = 3$, wherein the program converges with precision 10^{-10}, in 18 steps. For large values of n the convergence is very fast, typically in no more than 6 steps. Besides giving results in real format, the program also shows explicitly the sequence of ratios of two integers obtained through the iteration process. In Table A.1 one can see an example of the data generated by the program for $n = 2$.

```
n = 2
i = 1
j = 1

 0        1.0                1 / 1
 1        1.5                3 / 2
 2        1.4                7 / 5
 3        1.41666667        17 / 12
 4        1.4137931         41 / 29
 5        1.41428571        99 / 70
 6        1.41420118       239 / 169
 7        1.41421569       577 / 408
 8        1.4142132       1393 / 985
 9        1.41421362      3363 / 2378
10        1.41421355      8119 / 5741
11        1.41421356     19601 / 13860
12        1.41421356     47321 / 33461
13        1.41421356    114243 / 80782
14        1.41421356    275807 / 195025

sqrt(2) = 1.41421356
```

Table A.1: Data generated by the `continued_fractions` program, showing a sequence of rational numbers which tends to $\sqrt{2}$.

Appendix B

Series for the Square Roots of Integers

We will examine here the question of how to build a Taylor series with rational coefficients and variables in order to approximate the square root of a given integer.

General Case

Consider an arbitrary integer n. It can be written in the form $n = p^2 + q$, where p^2 is the largest perfect square less than or equal to n and q is limited to the interval $[0, 2p)$, such that q interpolates between p^2 and $p^2 + 2p = (p+1)^2 - 1$, which is the number that precedes the next perfect square. Let us consider the function

$$f(x) = \frac{n}{\sqrt{p^2 + x}},$$

which, for $x = q$, is equal to $\sqrt{n}$. We can easily calculate the successive derivatives of this function, thus obtaining

$$\begin{aligned}
f^{0\prime}(x) &= \frac{n}{\sqrt{p^2 + x}}, \\
f^{1\prime}(x) &= \left(-\frac{1}{2}\right) \frac{n}{\sqrt{p^2 + x}^{\,3}}, \\
f^{2\prime}(x) &= \left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right) \frac{n}{\sqrt{p^2 + x}^{\,5}},
\end{aligned}$$

$$\begin{array}{rcl} f^{3\prime}(x) & = & \left(-\frac{1}{2}\right)\left(-\frac{3}{2}\right)\left(-\frac{5}{2}\right)\frac{n}{\sqrt{p^2+x}^{7}}, \\ \cdots & & \cdots, \end{array}$$

so that the Taylor series of the function around $x = 0$ can be written as

$$f(x) = n\sum_{i=0}^{\infty}(-1)^i \, \frac{(2i-1)!!}{2^i p^{2i+1}} \, \frac{x^i}{i!}.$$

Since we have that $(2i+1)!! = (2i+1)(2i-1)!!$, and if in addition to this we multiply and divide by $(2i)!!$, then we can write this as

$$f(x) = n\sum_{i=0}^{\infty}(-1)^i \, \frac{(2i+1)!!(2i)!!}{(2i+1)2^i p^{2i+1}(2i)!!} \, \frac{x^i}{i!}.$$

Considering now that $(2i+1)!!(2i)!! = (2i+1)!$, and that $(2i)!! = 2^i i!$, we obtain

$$f(x) = n\sum_{i=0}^{\infty}(-1)^i \, \frac{(2i+1)!}{(2i+1)2^{2i} p^{2i+1}} \, \frac{x^i}{(i!)^2},$$

and making some simplifications we finally have, for the case $x = q$,

$$\sqrt{n} = \frac{n}{p}\sum_{i=0}^{\infty}(-1)^i \, \frac{(2i)!}{(i!)^2} \, \frac{q^i}{(2p)^{2i}}.$$

Note that we have here a series of manifestly rational numbers representing the irrational number $\sqrt{n}$. Since the sum of rational numbers is also rational, the partial sums in this series form a sequence of rational numbers which is such that, if the series converges, then this sequence converges to an irrational number.

Let us examine the radius of convergence of this series. Since the only singularity of the function is at $x = -p^2$ and we are expanding around $x = 0$, it follows that the radius of the convergence disk is p^2, and the disk's center is at zero. It follows that for $q < p^2$ the series must converge fast, for $q = p^2$ it must converge very slowly, if it converges at all, and for $q > p^2$ it must not converge at all. We can confirm this behavior calculating the ratio of the absolute values of two successive terms t_i and t_{i+1} of the series, as a function of i, thus obtaining

$$\frac{|t_{i+1}|}{|t_i|} = \frac{q}{p^2} \, \frac{4i^2+6i+2}{4i^2+8i+4},$$

so that for $i \to \infty$ we have

$$\lim_{i\to\infty} \frac{|t_{i+1}|}{|t_i|} = \frac{q}{p^2}.$$

In order to examine the convergence for all possible values of n it suffices to look at those values for which q is maximum for a given value of p, that is, for those integers that precede a perfect square, given by $p^2 - 1$. If there is no convergence in this case, then one verifies also the integer preceding q, for the same p, that is, $p^2 - 2$. Examining case by case the values of n, p and q, we see that there will be convergence problems in only three cases,

$$\begin{aligned} n = 2 \quad &\Rightarrow \quad p = 1, \quad q = 1, \quad \frac{q}{p^2} = 1, \\ n = 3 \quad &\Rightarrow \quad p = 1, \quad q = 2, \quad \frac{q}{p^2} = 2, \\ n = 8 \quad &\Rightarrow \quad p = 2, \quad q = 4, \quad \frac{q}{p^2} = 1. \end{aligned}$$

In all other cases there will be rapid convergence. The only case that really diverges is $n = 3$. In the case $n = 8$, since the square root is $2\sqrt{2}$, the problem is reduced again to the case $n = 2$. We therefore need to develop fast-converging alternative series for the cases $n = 2$ and $n = 3$, which is what we will do next.

Case $n = 2$

In the case $n = 2$ we can construct a series that converges faster following the same steps used in the general case. It suffices to consider the function

$$f(x) = \frac{2}{\sqrt{4 - x}},$$

that, for $x = 2$, is equal to $\sqrt{2}$. We can easily calculate the successive derivatives of this function, thus obtaining

$$\begin{aligned} f^{0\prime}(x) &= \frac{2}{\sqrt{4 - x}}, \\ f^{1\prime}(x) &= \left(\frac{1}{2}\right) \frac{2}{\sqrt{4 - x}^3}, \end{aligned}$$

$$\begin{aligned}
f^{2\prime}(x) &= \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\frac{2}{\sqrt{4-x}^5}, \\
f^{3\prime}(x) &= \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\left(\frac{5}{2}\right)\frac{2}{\sqrt{4-x}^7}, \\
\cdots & \quad \cdots,
\end{aligned}$$

so that the Taylor series of the function around $x = 0$ can be written as

$$f(x) = 2\sum_{i=0}^{\infty} \frac{(2i-1)!!}{2^{3i+1}}\frac{x^i}{i!}.$$

Using the same manipulations as before we can write this as

$$f(x) = \sum_{i=0}^{\infty} \frac{(2i)!}{2^{4i}}\frac{x^i}{(i!)^2},$$

so that for the case $x = 2$ we have

$$\sqrt{2} = \sum_{i=0}^{\infty} \frac{(2i)!}{(i!)^2}\frac{1}{2^{3i}}.$$

Since the singularity of the function is at $x = 4$ and the expansion is around $x = 0$, the convergence disk radius is 4. Since we are using the series for $x = 2$ we are well within the convergence disk, so that the series will converge quickly.

Case $n = 3$

In the case $n = 3$ we again proceed similarly. In this case it suffices to consider the function

$$f(x) = \frac{3}{\sqrt{9-x}},$$

that, for $x = 6$, is equal to $\sqrt{3}$. We can easily calculate the successive derivatives of this function, thus obtaining

$$\begin{aligned}
f^{0\prime}(x) &= \frac{3}{\sqrt{9-x}}, \\
f^{1\prime}(x) &= \left(\frac{1}{2}\right)\frac{3}{\sqrt{9-x}^3},
\end{aligned}$$

$$\begin{aligned}
f^{2\prime}(x) &= \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\frac{3}{\sqrt{9-x}^5}, \\
f^{3\prime}(x) &= \left(\frac{1}{2}\right)\left(\frac{3}{2}\right)\left(\frac{5}{2}\right)\frac{3}{\sqrt{9-x}^7}, \\
\cdots \quad & \quad \cdots,
\end{aligned}$$

so that the Taylor series of the function around $x = 0$ can be written as

$$f(x) = 3\sum_{i=0}^{\infty} \frac{(2i-1)!!}{2^i 3^{2i+1}} \frac{x^i}{i!}.$$

Using the same manipulations as before we can write this as

$$f(x) = \sum_{i=0}^{\infty} \frac{(2i)!}{2^{2i} 3^{2i}} \frac{x^i}{(i!)^2},$$

so that for the case $x = 6$ we have

$$\sqrt{3} = \sum_{i=0}^{\infty} \frac{(2i)!}{(i!)^2} \frac{1}{2^i 3^i}.$$

Since the singularity of the function is at $x = 9$ and the expansion is around $x = 0$, the convergence disk radius is 9. Since we are using the series for $x = 6$ we are well within the convergence disk so that the series will not only converge, but will do it very quickly.

Programs

There are three programs available, called `series_sqrtn`, `series_sqrt2` and `series_sqrt3`, which implement the series described herein. The first implements the general case, and the other two implement the two particular cases needed to complete the values of n. Given a value of n, the program for the general case determines the corresponding values of p and q using the following algorithm involving integer truncation:

```
p=int(sqrt(real(n))), q=n-p**2.
```

Note that in the general case the terms of the series alternate signs, and thus in this case it is advantageous, in order to accelerate the convergence of the series, to add the terms in pairs, as is done by the program.

Note that the factorials that appears in the terms diverge very rapidly with i, so that the direct calculation of the factors in each term can cause "overflows" even for relatively small values of i. Since these divergences appear in both numerator and denominator, and end up canceling off, it is advantageous to calculate the terms logarithmically. The cases of slower convergence are those for $n = 7$, which converges with precision 10^{-10} with about 70 terms of the series, and $n = 3$, which converges with about 50 terms. All other cases converge with about 30 terms or less. For large values of n the convergence typically occurs with less than 10 terms. In Table B.1 one can see an example of the data generated by the program for $n = 2$.

0	1.0	1.0
1	0.2	1.25
2	0.0697674419	1.34375
3	0.0282485876	1.3828125
4	0.0122078828	1.39990234
5	0.00546353308	1.40759277
6	0.00249786438	1.41111755
7	0.00115837935	1.41275406
8	0.000542695641	1.41352117
9	0.000256207283	1.41388342
10	0.000121683651	1.41405549
11	5.80729151E-05	1.41413761
12	2.78258309E-05	1.41417696
13	1.33776243E-05	1.41419588
14	6.44988442E-06	1.414205
15	3.11743442E-06	1.41420941
16	1.51000502E-06	1.41421154
17	7.32796015E-07	1.41421258
18	3.56220158E-07	1.41421308
19	1.73422942E-07	1.41421333
20	8.45436769E-08	1.41421345
21	4.12653644E-08	1.41421351
22	2.01637572E-08	1.41421354
23	9.86270717E-09	1.41421355
24	4.82861699E-09	1.41421356
25	2.36602242E-09	1.41421356
26	1.16026096E-09	1.41421356
27	5.69387279E-10	1.41421356
28	2.79609916E-10	1.41421356
29	1.37394552E-10	1.41421356
30	6.7552286E-11	1.41421356
sqrt(2) =		1.41421356

Table B.1: Data generated by the `series_sqrt2` program, showing a sequence of rational numbers which tends to $\sqrt{2}$.

Appendix C

The Stirling Approximation

Here we will examine the asymptotic behavior of the function $\Gamma(x)$, for real $x \gg 1$, using the approximation process called asymptotic approximation or saddle point approximation. This will generate an approximation of the function for the case of large arguments that will be very useful in the study of differential equations in curvilinear coordinates and of the special functions that appear on that study. As was seen in the text, the function can be written for real x as

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

for $x > 0$. Since the function in the integrand decreases exponentially for large values of x, and since for $x > 1$ it goes to zero both for $t \to 0$ and for $t \to \infty$, it follows that the main contribution to the integral should come from the vicinity of the point of maximum of this function on this semi-axis of positive t. Since the exponential function is monotonic and positive, we first locate the maximum of the exponent,

$$\begin{aligned}
\frac{d}{dt}[(x-1)\ln(t) - t] &= \frac{x-1}{t} - 1 \\
&= 0 \quad \Rightarrow \\
t &= x - 1.
\end{aligned}$$

Thus we see that for $x \gg 1$ the only point of maximum of the function is located at $t \gg 1$. We will now expand the exponent in a power series on t around the point $t = x - 1$, up to second order. Note that we have here a purely real Taylor series. The second derivative of the exponent is given by

$$\begin{aligned}\frac{d^2}{dt^2}[(x-1)\ln(t)-t] &= \frac{d}{dt}\left[\frac{x-1}{t}-1\right]\\ &= -\frac{x-1}{t^2},\end{aligned}$$

and therefore we have for the expansion around $t = x - 1$,

$$\begin{aligned}[(x-1)\ln(t)-t]\Big|_{t=x-1} &= (x-1)\ln(x-1)-(x-1),\\ \frac{d}{dt}[(x-1)\ln(t)-t]\Big|_{t=x-1} &= 0,\\ \frac{d^2}{dt^2}[(x-1)\ln(t)-t]\Big|_{t=x-1} &= -\frac{1}{x-1} \quad\Rightarrow\\ [(x-1)\ln(t)-t] &= (x-1)\ln(x-1)-(x-1)\\ &\quad -\frac{1}{2}\frac{[t-(x-1)]^2}{x-1}+\ldots\,.\end{aligned}$$

As we shall see, it can be shown that the behavior of the exponent near $t = x - 1$ becomes Gaussian in the limit $x \to \infty$, so that in this limit the approximation that we are developing here becomes exact. We can now write for $\Gamma(x)$

$$\Gamma(x) \approx \mathrm{e}^{(x-1)\ln(x-1)-(x-1)} \int_0^\infty \exp\left\{-\frac{[t-(x-1)]^2}{2(x-1)}\right\} dt.$$

Note that we have a Gaussian integral to do. In the limit of large values of x the integrand goes to zero in the whole $t < 0$ semi-axis, so that we can complete the integration domain, which becomes $(-\infty, \infty)$, without appreciably changing the value of the integral. With the further change of variables $t' = t - (x - 1)$, we can write

$$\Gamma(x) \approx \mathrm{e}^{(x-1)\ln(x-1)-(x-1)} \int_{-\infty}^\infty \exp\left\{-\frac{[t']^2}{2(x-1)}\right\} dt'.$$

Making now the change of variables $t'' = t'\sqrt{x-1}$, which is such that $dt'' = dt'\sqrt{x-1}$, we have

$$\Gamma(x) \approx \mathrm{e}^{(x-1)\ln(x-1)-(x-1)}\sqrt{x-1} \int_{-\infty}^\infty \exp\left(-\frac{[t'']^2}{2}\right) dt''.$$

The Gaussian integral that remains is given by

$$\int_{-\infty}^{\infty} \exp\left(-\frac{t^2}{2}\right) dt = \sqrt{2\pi},$$

so that we get for the function $\Gamma(x)$ the approximation

$$\Gamma(x) \approx \sqrt{2\pi}\sqrt{x-1}(x-1)^{(x-1)}\,\mathrm{e}^{-(x-1)}.$$

Using now the property that $\Gamma(x) = (x-1)\Gamma(x-1)$ one can write that

$$\Gamma(x-1) \approx \sqrt{\frac{2\pi}{x-1}}\,(x-1)^{(x-1)}\,\mathrm{e}^{-(x-1)}.$$

Replacing now $x-1$ with x we finally obtain the relation

$$\Gamma(x) \approx \sqrt{2\pi}\,x^{(x-1/2)}\,\mathrm{e}^{-x},$$

which is the *Stirling approximation* for the function $\Gamma(x)$, for large arguments, that is, for the factorials of large numbers. In general, this approximation is presented in terms of logarithms, as

$$\ln[\Gamma(x)] \approx x\ln(x) - x - \frac{1}{2}\ln(x) + \frac{1}{2}\ln(2\pi).$$

The two dominant terms of this expression, for large values of x, are the first two, so that in general we write this approximation in its simplest form as $\ln[\Gamma(x)] \sim x\ln(x) - x$. Note that the Gaussian integral that gives the first multiplicative correction with respect to this simple approximation,

$$\int_0^{\infty} \exp\left\{-\frac{[t-(x-1)]^2}{2(x-1)}\right\} dt,$$

has as the point of maximum of the integrand $t_M = (x-1)$, while the width of the Gaussian is given by $\sigma = \sqrt{x-1}$. Thus we see that, in relation to the position of the peak of the Gaussian, the width goes to zero for large values of x, and the relevant part of the integral is increasingly concentrated around the maximum, indicating that the approximation becomes increasingly better. The third derivative of the exponent of the original expression is given by

$$\frac{d^3}{dt^3}[(x-1)\ln(t) - t] = 2\,\frac{x-1}{t^3},$$

which, when applied at the point of maximum $t = (x-1)$ results in the coefficient

$$\frac{d^3}{dt^3}[(x-1)\ln(t) - t]\,\Big\lfloor_{t=x-1} = \frac{2}{(x-1)^2},$$

which goes to zero faster than the coefficient of the second-order term. The same is true for the subsequent coefficients, which shows that the Gaussian approximation becomes increasingly more precise for x tending to infinity.

It is interesting to observe how this Gaussian integrand behaves when we extend the variable t to complex values, maintaining the parameter x at real values. Starting at the point of maximum $t = (x-1)$ and going in the imaginary direction of complex t, we will have that the variable $t-(x-1)$ is purely imaginary, that is, we can write $t-(x-1) = \imath\tau$ for a variable real τ. Since $\imath^2 = -1$, this reverses the sign of the exponent, and we have in terms of τ

$$\int \exp\left\{\frac{\tau^2}{2(x-1)}\right\}\, dt,$$

which again is a real integrand, but this time has a point of *minimum* at $t = (x-1)$, instead of a point of maximum. This is a general property of analytic functions, they never have local maxima or minima, but only points at which the derivative is zero, and which are points of maximum in some direction and of minimum in another direction. For this reason these possible stationary points of analytic functions are called *saddle points*. Note that if the integral was taken on a straight line passing through the point $t = (x-1)$ and extending infinitely in the imaginary direction of the plane of complex t, it would indeed be divergent, and therefore would not be well defined.

If we have some given integral of an analytic function and it is possible to change the path of integration, without changing the value of the integral, in such a way that it passes through a stationary point of the function, exactly in the direction in which it decreases as fast as possible from that point, then it is possible to concentrate the dominant contributions to the integral in the vicinity of this point, rather than having them scattered in a complicated way along the integration contour. As it is discussed in the text, due to the Cauchy-Goursat theorem it is indeed possible to make this kind of deformation of an integration contour in the complex plane without changing the value of the integral of an analytic function. This gives rise to a general technique of approximation for this type of integral, which is called the saddle point method, or "steepest descent" method.

Appendix D

Some Facts from Real Analysis

We will prove here some simple theorems of real analysis, which are used in the text. Generally speaking, these theorems have contents that are intuitively reasonable and clear. What we will do here is to formulate the theorems precisely and also to give precise proofs.

Theorem D.1: *A monotonically increasing sequence of real numbers that is bounded from above is necessarily convergent.*

The content of this theorem is not difficult to induce intuitively. Since the sequence is monotonically increasing, it cannot go back once a certain point is reached, which prevents it from oscillating indefinitely. Moreover, if there is an upper limit, it is also not possible for it to diverge to infinity. Thus, the two usual modes of divergence are discarded. A rigorous proof ensures that there is no other possibility of divergence of which we have not thought.

Proof D.1.1:

Consider a sequence of real numbers a_n, that is indexed by integers $n \in \{0, 1, 2, 3, \ldots, \infty\}$, and that is monotonically increasing and bounded from above by a number B_0,

$$\begin{aligned} a_n &\leq a_{n+1}, \\ a_n &\leq B_0, \end{aligned}$$

for all n. Consider the closed interval $[A_0, B_0]$ where $A_0 = a_0$, which has its length given by

$$L_0 = B_0 - A_0.$$

It follows immediately that the whole sequence is contained in this interval, that is, a_n is contained within the interval for all $n \geq 0$. Consider now the midpoint of the interval, which is given by

$$C_0 = \frac{A_0 + B_0}{2}.$$

About the occupation of the two sub-intervals $[A_0, C_0]$ and $(C_0, B_0]$ by the elements of the sequence, we can now state what follows. Either there is some element of the sequence greater than the midpoint, or there is none, and there is no third possibility. If there is no element a_{n_0} such that $a_{n_0} > C_0$, then all the elements of the sequence are within the left closed interval, that is, C_0 is a new upper limit. In this case we define a new closed interval $[A_1, B_1]$ making $A_1 = A_0$ and $B_1 = C_0$, which has its length given by

$$\begin{aligned} L_1 &= B_1 - A_1 \\ &= \frac{1}{2} L_0. \end{aligned}$$

We have that the sequence is completely contained within this lower interval, that is, a_n is contained within the interval for all $n \geq 0$.

On the other hand, if there is any element a_{n_0} of the sequence such that $a_{n_0} > C_0$ then, since the sequence is monotonically increasing, all subsequent elements are in the right sub-interval, that is $a_n > C_0$ for all $n \geq n_0$. In this case, there is only a finite number of sequence elements in the left sub-interval, those ranging from a_0 to a_{n_0-1}. In this case we define the new closed interval $[A_1, B_1]$ making $A_1 = C_0$ and $B_1 = B_0$, which has the same length L_1 shown above.

In any case, the new interval $[A_1, B_1]$ is such that its length is half the length of the previous interval, that is, $L_1 = L_0/2$. There is always an infinite number of sequence elements within the new interval, and only a finite number outside. Therefore, there is always a certain number n_1 such that if $n > n_1$, then the element a_n is within the new interval. In the first of the two cases above we have $n_1 = 0$, and in the second case we have $n_1 = n_0$.

Repeating indefinitely this procedure we construct a sequence of intervals $[A_m, B_m]$, with $m \in \{0, 1, 2, 3, \ldots, \infty\}$, each nested within the previous interval, with lengths given by

$$L_m = \frac{1}{2^m} L_0,$$

such that for each one of these intervals it is true that there is only a finite number of sequence elements outside. In other words, for each of these intervals there is a number n_m such that if $n > n_m$, then a_n is within the interval $[A_m, B_m]$.

Note that in the limit $m \to \infty$ the length L_m goes to zero. It follows that there is only one real number a_∞ that is contained in all these intervals. This is an accumulation point of the sequence, and therefore its limit. This set of intervals is an example of the nested intervals of Bolzano, and the completeness axiom of the real line ensures that this limiting point exists in $\mathbb{R}$. Note that we have, from our construction of the sequence of nested intervals, the limits

$$\begin{aligned} \lim_{m\to\infty} A_m &= a_\infty, \\ \lim_{m\to\infty} B_m &= a_\infty, \\ \lim_{m\to\infty} C_m &= a_\infty. \end{aligned}$$

Now we can explicitly construct the criterion for the convergence of the sequence. Given $\varepsilon > 0$, there is a value of m such that $L_m < \varepsilon$, because

$$\begin{aligned} L_m &= \frac{1}{2^m} L_0 \\ &< \varepsilon \quad \Rightarrow \\ 2^m &> \frac{L_0}{\varepsilon} \quad \Rightarrow \\ m \ln(2) &> \ln\left(\frac{L_0}{\varepsilon}\right) \quad \Rightarrow \\ m &> \frac{\ln(L_0/\varepsilon)}{\ln(2)}. \end{aligned}$$

If we choose as the value of m the first integer larger than the right-hand side of this last equation, the condition $L_m < \varepsilon$ will be satisfied. We will call this value $m(\varepsilon)$. Consider now the interval $[A_m, B_m]$. As we saw earlier, there is a n_m such that if $n > n_m$ then a_n is within this interval. In the case of the interval given by $m(\varepsilon)$, we will refer to this value of n_m as $n(\varepsilon)$.

Now, given ε, we consider the interval given by $m(\varepsilon)$, which has length smaller than ε, and which contains the point a_∞, which is true for all $m(\varepsilon)$. Taking now the value $n(\varepsilon)$, we know that if $n > n(\varepsilon)$ then a_n is also within the same interval, and thus at a shorter distance than L_m from any other point within the interval, including a_∞. Since we have $L_m < \varepsilon$, we conclude in particular that, for the element a_n,

$$|a_n - a_\infty| < \varepsilon.$$

Thus we have shown that, for any $\varepsilon > 0$, there is a value $n(\varepsilon)$ such that if $n > n(\varepsilon)$ then the convergence condition is satisfied. This establishes the convergence of the sequence to the point a_∞, and therefore completes the proof of Theorem D.1.

It is clear that the same theorem holds for monotonically decreasing sequences that are limited from below, because it suffices to change the sign of both the elements of the sequence and the limit, in order to reduce the problem to the one we just solved. In addition to this, since a series of positive terms has monotonically increasing partial sums, and since the convergence of the series is the convergence of this sequence of partial sums, the following corollary immediately follows.

Corollary D.1.1: *A series of positive terms that is limited from above is necessarily convergent.*

Theorem D.2: *If a series converges, then the sequence of terms of the series is necessarily convergent to zero.*

Again there is a quite clear intuitive approach here, because if the terms did not tend to zero, then every step of the infinite sum would always add a quantity with a certain minimum absolute value to the value of the series, which therefore could never stay near a fixed limit. The sum would have to diverge to infinity or else would have to oscillate between two numbers distant from one another by at least that minimum amount. A rigorous proof shows that this intuition is in fact correct, and that there is no other possibility of which we have not thought.

Proof D.2.1:

Consider a series with partial sums given by

$$S_N = \sum_{n=0}^{N} t_n,$$

which by hypothesis converges to a number S_∞. This means that the sequence of partial sums converges to S_∞, that is, the following condition is valid: given $\varepsilon > 0$, there is an $N(\varepsilon)$ such that if $N > N(\varepsilon)$, then

$$|S_\infty - S_N| < \varepsilon.$$

Since this holds for all N larger than $N(\varepsilon)$, then it holds in particular for $N + 1$, so that we also have that

$$|S_\infty - S_{N+1}| < \varepsilon.$$

Consider now that we can write for the term t_{N+1} of the series that

$$S_{N+1} - S_N = t_{N+1},$$

from which it follows that

$$(S_\infty - S_N) - (S_\infty - S_{N+1}) = t_{N+1}.$$

Taking absolute values and using the triangle inequality, we obtain

$$\begin{array}{rcl} |t_{N+1}| & \leq & |S_\infty - S_N| + |S_\infty - S_{N+1}| \\ & < & 2\varepsilon \quad \Rightarrow \\ |t_{N+1}| & < & 2\varepsilon. \end{array}$$

Since all this argument holds for any strictly positive value of ε, we can repeat it using $\bar{\varepsilon} = \varepsilon/2$. The argument now takes the following form. Given $\varepsilon > 0$, we define $\bar{\varepsilon} = \varepsilon/2$, which also satisfies the condition of strict positivity. Now, given this $\bar{\varepsilon} > 0$, as we have seen, there is an $N(\bar{\varepsilon})$ such that if $N > N(\bar{\varepsilon})$ then

$$\begin{array}{rcl} |S_\infty - S_N| & < & \bar{\varepsilon}, \\ |S_\infty - S_{N+1}| & < & \bar{\varepsilon}, \end{array}$$

As we also noted above, this in turn implies that

$$\begin{array}{rcl} |t_{N+1}| & < & 2\bar{\varepsilon} \quad \Rightarrow \\ |t_{N+1} - 0| & < & \varepsilon. \end{array}$$

In order to write this relation in its usual form, we can now define a new number $N'(\varepsilon) = N(\bar{\varepsilon}) + 1$, such that, if $N > N'(\varepsilon)$, then it follows that

$$|t_N - 0| < \varepsilon.$$

This is the condition that the sequence t_n converges, showing also that the limit is zero. This completes the proof of Theorem 13.2.

Since a convergent numerical sequence consists of a collection of finite numbers with a finite limit, the following corollary follows.

Corollary D.2.1: *If a series converges, then the sequence t_n of terms of the series is limited both from below and from above.*

We can establish this as follows. Since the sequence is convergent, we can take some arbitrary value of ε, for example $\varepsilon = 1$, and we know that there is an $N(\varepsilon)$ such that for $N > N(\varepsilon)$ we have that $|t_N - t_\infty| < \varepsilon$, where t_∞ is the finite limit of the sequence. This implies that t_n is within an interval of length 2ε around t_∞, because if we take squares we have

$$(t_N - t_\infty)^2 < \varepsilon^2.$$

The equality of the two sides implies that

$$\begin{aligned} t_N - t_\infty &= \pm\varepsilon \quad \Rightarrow \\ t_N &= t_\infty \pm \varepsilon, \end{aligned}$$

so that the inequality can be written as

$$t_\infty - \varepsilon < t_N < t_\infty + \varepsilon.$$

With this we establish that for $N > N(\varepsilon)$ all elements t_n are limited within this interval. Since only a finite number of elements is outside this interval, these are necessarily also limited, as one can see by simply considering as the limits the smallest and the largest among them. It follows that the whole sequence is limited between some maximum and minimum limits.

Appendix E

Complete Singularity Classification

In most texts on the subject one examines a few types of singularity of complex analytic functions, most notably poles and branch points. While poles can be isolated singularities, branch points cannot be completely isolated, in a certain sense. Isolated poles have two important properties, the first being that there is a neighborhood of the singular point that contains no other singularities, and the other being that it is possible to draw in the domain of the function a closed integration contour that makes a single turn around them. While branch points can be isolated in the sense of the first property, the second property never holds for them. This is so because a contour that makes a single turn around a branch point will always cross the branch cut and therefore change to another leaf of a Riemann surface, not being therefore truly a closed contour within that Riemann surface, which in this case is the actual domain of the function under consideration.

It is an interesting fact that it is possible to set up a complete classification of all possible singularities of analytic functions that is independent of the two properties mentioned above. All that is needed in order to do this is the concept of the limit of an analytic function $w(z)$ to a given point on its domain, or on its closure, and the usual concepts of differentiation and of indefinite integration of that analytic function, leading to its derivative and primitive, for which we will use here the notations that follow,

$$w'(z) \quad = \quad w^{1\prime}(z)$$

$$
\begin{aligned}
&= \frac{dw(z)}{dz}, \\
w^{-1\prime}(z) &= \int_{z_0}^{z} dz'\, w(z')\,,
\end{aligned}
$$

where z_0 is some arbitrary reference point within the domain of analyticity of the function, and where the integration contour from z_0 to z is contained within that domain, with the single possible exception of the upper end-point $z = z_1$, where z_1 is a singular point. In order to proceed we must first recall that any analytic function is always both differentiable and integrable within its whole domain of analyticity, thus producing other analytic functions on exactly the same domain. Therefore, the derivative and the primitive always exist, and are always analytic function on the same domain of analyticity of the original function. As a consequence of this, if an analytic function has a singularity at a certain point, then both its derivative and its primitive also have singularities at that same point. This invariance of the singularity structure under the analytic operations will constitute a foundation of the singularity classification. Another instance of this classification of singularities, somewhat different from the one we will discuss here, can be found in [11].

Let then $w(z)$ be an arbitrary analytic function and let z_1 be a point where it has a singularity. The first step in the construction of the classification will depend only of the concept of the limit of $w(z)$ when we take the limit $z \to z_1$. As this first step, we therefore adopt the following definition.

Definition E.1: *Classification of Singularities: Soft and Hard*

The singularity of $w(z)$ at z_1 is a *soft singularity* if the limit of $w(z)$ to that point exists and is finite. Otherwise, it is a *hard singularity.*

This is already a complete classification of all possible singularities because, given the function $w(z)$ and the point z_1, either the limit of $w(z)$ when $z \to z_1$ exists, or it does not. There is no third alternative, and therefore every singularity is either soft or hard. The intuition behind this definitions is given by the fact that, if the singularity at z_1 is soft, then the function $w(z)$ is continuous and bounded there, but if the singularity is hard, then the function is discontinuous and either oscillates indefinitely or diverges to infinity there.

We will now prove a couple of important properties of this general singularity classification, one for soft singularities and one for hard singularities. We begin by discussing a property of hard singularities, which is related to the operation of differentiation.

Property E.1.1: *If $w(z)$ has a hard singularity at z_1, then the derivative $w'(z)$ also has a hard singularity at that point.*

This can be establishing very simply by the observation that, if $w(z)$ has a hard singularity at z_1, then it is not well defined there, implying that it is not continuous there, and therefore that it is also not differentiable there. This clearly implies that the derivative $w'(z)$ of $w(z)$, which we already know to also have a singularity at z_1, is not well defined there as well. This in turn implies that the singularity of $w'(z)$ at z_1 must be a hard one, which thus establishes this property.

Let us discuss now a property of soft singularities, which is related to the operation of integration.

Property E.1.2: *If $w(z)$ has a soft singularity at z_1, then the primitive $w^{-1\prime}(z)$ also has a soft singularity at that point.*

This can be easily established by reductio ad absurdum, using the previous property and the fact that differentiation and integration are inverse operation to one another. Suppose that we have a function $w(z)$ with a soft singularity at z_1, and that the primitive $w^{-1\prime}(z)$ of $w(z)$ has a hard singularity at that point. Then by the previous property it follows that the derivative of this primitive also has a hard singularity at z_1. However, this is absurd, because when we differentiate $w^{-1\prime}(z)$ we get back the function $w(z)$ itself, which has a soft singularity there. Therefore, we conclude that $w^{-1\prime}(z)$ must have a soft singularity at z_1, and thus we have established this property.

Note that these two properties imply that, if we have a hard singularity at z_1, then successive differentiations of $w(z)$ produce an infinite sequence of hard singularities at that point, and that if we have a soft singularity at z_1, then successive integrations of $w(z)$ produce an infinite sequence of soft singularities at that point. On the other hand, the differentiation of a soft singularity may he either soft or hard, and the integration of a hard singularity may he either hard or soft. In either case, once the character of the singularity changes under one of these two analytic operations, that change is never reversed by further application of the same operation. Therefore, if there is a point along the sequence of analytic operation where the character of the singularity changes, then this point is unique. In other words, there is a single transition between two functions on any such chain of analytic operations where the character of the singularity changes, unless it never changes along the whole chain.

Next, as our second step in the construction of the complete singularity classification, we will refine this general classification of singularities by the introduction of *gradations* of both the soft singularities and the hard singularities. This can be done very simply with the use of the operations of differentiation and of integration of analytic functions. The first step in order to establish these gradations is the introduction of the concept of a *borderline hard* singularity, to which we will assign the degree of hardness zero. We therefore adopt the following definition.

Definition E.2: *Classification of Singularities: Borderline Hard Singularities*

Given the function $w(z)$ and the point z_1 where it has a *hard* singularity, if a *single* integration of $w(z)$ produces a primitive $w^{-1\prime}(z)$ that has at z_1 a *soft* singularity, then the function $w(z)$ has at z_1 a *borderline hard* singularity, that is, a hard singularity with degree of hardness zero.

When we integrate $w(z)$, the primitive $w^{-1\prime}(z)$ certainly has at z_1 a singularity. This singularity may be either soft or hard. It the singularity of the primitive is soft, then the integration has mapped the original hard singularity onto a soft one. Therefore, this characterizes the single point of transition at which the character of the singularity changes under the application of the analytic operations.

In order to complete the second step in the construction of our complete classification of singularities, involving the gradations of soft and hard singularities, and having defined the borderline hard singularities as having degree of hardness zero, we will now use the operations of differentiation and integration in order to assign to each singularity either a *degree of softness* or a *degree of hardness*. Given the function $w(z)$ and the singular point z_1, we adopt the following two definitions.

Definition E.3: *Classification of Singularities: Gradation of Soft Singularities*

Let us assume that the singularity of $w(z)$ at z_1 is soft. If an arbitrarily large number of successive differentiations of $w(z)$ always results in a singularity at z_1 which is still soft, then we say that the singularity of $w(z)$ at z_1 is an *infinitely soft* singularity. Otherwise, if n is the number of differentiations that have to be applied to $w(z)$ in order for the singularity at z_1 to become a borderline hard one, then we define n as the *degree of softness* of the original singularity of $w(z)$ at z_1. Therefore, a degree of softness is an integer $n \in \{1, 2, 3, \ldots, \infty\}$.

Definition E.4: *Classification of Singularities: Gradation of Hard Singularities*

Let us assume that singularity of $w(z)$ at z_1 is hard. If an arbitrarily large number of successive integrations of $w(z)$ always results in a singularity at z_1 which is still hard, then we say that the singularity of $w(z)$ at z_1 is an *infinitely hard* singularity. Otherwise, if n is the number of integrations that have to be applied to $w(z)$ in order for the singularity at z_1 to become a borderline hard one, then we define n as the *degree of hardness* of the original singularity of $w(z)$ at z_1. Therefore, a degree of hardness is an integer $n \in \{0, 1, 2, 3, \ldots, \infty\}$.

In order to see that this establishes a complete classification of all possible singularities of analytic functions, let us examine all the possible outcomes when we apply differentiations and integrations to analytic functions. We already saw that, if we apply an integration to a soft singularity, then the result is always another soft singularity. We also saw that, if we apply a differentiation to a hard singularity, then the result is always another hard singularity. The two remaining possibilities are the application of an integration to a hard singularity, and the application of a differentiation to a soft singularity. In these two cases the resulting singularity may be either soft or hard, and the remaining possibilities were dealt with in Definitions E.3 and E.4. Since this applies to all the singularities of any given analytic function, it applies to all the possible singularities of all analytic functions. This constitutes, therefore, a complete and universal classification of singularities.

Let us end this discussion by establishing a few more important properties of soft and hard singularities, regarding their integrability, that is, the integrability of the function $w(z)$ on integration contours connecting to the singular point z_1. To begin with, we now establish the following important property of soft singularities.

Property E.1.3: *A soft singularity of the function $w(z)$ at the point z_1 is necessarily an integrable one.*

This is so because the integration of $w(z)$ produces an analytic function $w^{-1\prime}(z)$ which has at z_1 a soft singularity, and therefore is well defined at that point. Since the value of $w^{-1\prime}(z)$ at z_1 is given by an integral of $w(z)$ along an integration contour connecting to that point, that integral must therefore exist and result in a finite complex number. In addition to this, since $w^{-1\prime}(z)$ has at z_1 a well-defined and unique complex value, the integral defining it, starting from a reference point z_0, cannot depend on the integration contour from z_0 to z_1. We may thus conclude that all soft singularities are integrable ones, which establishes this property.

Next we establish the following important property of borderline hard singularities.

Property E.1.4: *A borderline hard singularity of the function $w(z)$ at the point z_1 in necessarily an integrable one.*

This is so because the integration of $w(z)$ produces an analytic function $w^{-1\prime}(z)$ which has at z_1 a soft singularity, given that the singularity of $w(z)$ is a borderline hard one, and therefore $w^{-1\prime}(z)$ is well defined at that point. Since the value of $w^{-1\prime}(z)$ at z_1 is given by an integral of $w(z)$ along an integration contour connecting to that point, that integral must therefore exist and result in a finite complex number. In addition to this, since $w^{-1\prime}(z)$ has at z_1 a well-defined and unique complex value, the integral defining it, starting from a reference point z_0, cannot depend on the integration contour from z_0 to z_1. We may thus conclude that all borderline hard singularities are integrable ones, which establishes this property.

As one can see this last argument is similar to that used for the previous property. In order to complete the picture, we establish now the final property regarding integrability.

Property E.1.5: *The borderline hard singularities are the only hard singularities that are integrable.*

If a hard singularity of $w(z)$ at z_1 has a degree of hardness of one or larger, then by integration it is mapped to another hard singularity, the hard singularity of $w^{-1\prime}(z)$ at z_1. If the hard singularity of $w(z)$ were integrable, then $w^{-1\prime}(z)$ would be well defined at z_1, and therefore its singularity would be soft rather than hard. Since we know that the singularity of $w^{-1\prime}(z)$ is hard, it follows that the singularity of $w(z)$

cannot be integrable. In other words, all hard singularities with strictly positive degrees of hardness are necessarily non-integrable singularities, which establishes this property.

In some cases examples of this classification are well known. For instance, a simple example of an infinitely hard singularity is any essential singularity. Examples of infinitely soft singularities are harder to come by, and they are related to functions which are infinitely differentiable but not analytic at a certain point. A simple example of a hard singularity with degree of hardness $n \geq 1$ is a pole of order n. Examples of soft singularities are the square root, and products of strictly positive powers with the logarithm. A simple example of a borderline hard singularity is a purely logarithmic singularity.

If a singularity at a given singular point z_1 is either infinitely soft or infinitely hard, then the corresponding chain of singularities obtained by multiple differentiation or multiple integration contains either only soft singularities or only hard singularities. If the singularity is neither infinitely soft nor infinitely hard, then at some point along that chain the character of the singularity changes, and from that point on the soft or hard character remains constant at the new value throughout the rest of the chain in that direction.

The transition between a borderline hard singularity and a soft singularity is therefore the single point of transition of the soft or hard character of the singularities along the corresponding chain. Starting from a borderline hard singularity, n integrations produce a soft singularity with degree of softness n, and n differentiations produce a hard singularity with degree or hardness n. Note that a strictly positive degree of softness given by n can be identified with a negative degree of hardness given by $-n$, and vice-versa.

The degrees of softness and hardness establish the analytic character of the function at the singular point. For example, a degree of softness $n = 1$ implies that the function is continuous but not differentiable at z_1, while a degree of softness $n > 1$ implies that the function is continuous and differentiable at z_1. A degree of hardness $n = 0$ implies that the function is discontinuous but integrable at z_1, while a degree of hardness $n > 0$ implies that the function is discontinuous and non-integrable at z_1, meaning that its integral fails to converge when one approaches z_1 along at least some directions.

It is interesting to observe that it seems that all soft singularities can be characterized as branch points associated to Riemann surfaces, but in fact we do not yet have a complete proof of this. Of course, hard singularities are not limited to poles of some order, and can be branch points associated to Riemann surfaces as well. Examples of this would be products of strictly negative powers with the logarithm.

Appendix F

The Complex Gamma Function

The gamma function $\Gamma(z)$ of the complex variable $z = x + \imath y$ is a very important example of an analytic function, with some very interesting properties. As we saw in the text, $\Gamma(z)$ can be defined in terms of a real parametric integral over the positive real semi-axis,

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

an integral that converges so long as $x > 0$, as was verified in the text. Defined in this way $\Gamma(z)$ is in fact an analytic function of z, as was shown in the text. Extensions of this definition are possible through the use of the identity $\Gamma(z+1) = z\,\Gamma(z)$ satisfied by this function, possibly iterated several times. For example, we have the alternative definition

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

which is valid for $x > -1$. In a similar way, we may write the definition

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

which is valid for $x > -2$. If we iterate this n times we get the definition

$$\Gamma(z) = \frac{1}{z(z+1)\ldots(z+n-2)(z+n-1)} \int_0^\infty dt\, t^{z+n-1}\, \mathrm{e}^{-t},$$

which is valid for $x > -n$. In this way we can extend the definition of $\Gamma(z)$ to the whole complex plane, where it is manifestly analytic except for the simple poles at the origin and at the negative integers, which are shown explicitly in these formulas. However, this process does not generate a single general formula for the definition of $\Gamma(z)$ on the whole complex plane. As we will see here, it is indeed possible to derive such a formula, involving a complex contour integral around the positive real semi-axis, which constitutes a complex integral representation of the function $\Gamma(z)$. Here we will do this, using the techniques of integration around a branch cut which were discussed in the text.

In order to do this we will extend the parametric integral on t given in Equation (F.1) from the positive real semi-axis to the complex plane, using a complex variable $s = t + \imath\sigma$. This will allow us to extend the definition of $\Gamma(z)$ to the whole complex plane, by a process of analytic continuation. In order to do this, first of all we must examine the singularity structure of the function that appears in the integrand of Equation (F.1), when it is extended to the complex plane,

$$f(s) = s^{z-1}\,\mathrm{e}^{-s}.$$

Unless the number $z - 1$ is real and an integer, the power of s has a branch point at the origin $s = 0$, since it is, by the very definition of a complex power, given by

$$s^{z-1} = \mathrm{e}^{(z-1)\ln(s)},$$

where $\ln(s)$ has a branch point at the origin, and where we will assume that the branch cut extends to infinity along the positive real semi-axis, as illustrated in Figure F.1.

Let us now consider the complex contour integral of the complex function $f(s)$, over the contour C_R described in Figure F.1, consisting, in sequence, of the segments $C_{R\oplus}$, $C_\oplus$, C_ε, $C_\ominus$ and $C_{R\ominus}$ shown there, with the orientation indicated,

$$\begin{aligned}\int_{C_R} ds\, f(s) &= \int_{C_R} ds\, s^{z-1}\,\mathrm{e}^{-s} \\ &= \int_{C_R} ds\, \mathrm{e}^{(z-1)\ln(s)}\,\mathrm{e}^{-s}.\end{aligned}$$

The contour C_R includes the two straight segments located at a distance ε above and below the positive real semi-axis, the semicircle of radius ε around the origin, and the two vertical segments of length ε at the position $t = R$. The idea here is to consider the integral in the $R \to \infty$

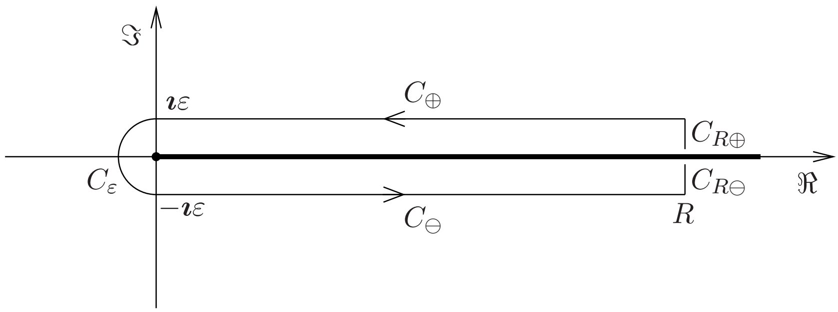


Figure F.1: The complex integration contour C_R used for the definition by analytic continuation of the function $\Gamma(z)$, when one takes the $R \to \infty$ limit, showing the single point of singularity of the integrand $f(s)$ at $s = 0$.

limit of this contour. It is important to note that C_R is *not* a closed contour, since it is defined on the Riemann surface of the logarithm, so that the two ends of the segments $C_{R\oplus}$ and $C_{R\ominus}$ located near the real line do not really meet. Not only this is not a closed contour, but it cannot be closed unless z is real and an integer. However, since the integrand is analytic along the whole integration contour, and goes to zero exponentially fast when $t \to \infty$, the integral over C_R certainly exists, for any value of R and for any value of z, and therefore defines a certain function of z on the whole complex plane.

The integral of $f(s)$ over the contour C_R can be separated into integrals on the five segments $C_{R\oplus}$, $C_\oplus$, C_ε, $C_\ominus$ and $C_{R\ominus}$, and hence we will have to discuss and evaluate in turn each one of these five parts of the complete integral over C_R. The first thing that we will show here is that, when we take the $R \to \infty$ limit, the integrals of $f(s)$ on the segments $C_{R\oplus}$ and $C_{R\ominus}$ tend to zero, regardless of the complex value of z and of the real value of ε. Therefore, this allows us to take the $R \to \infty$ limit and thus consider the resulting infinite contour C_∞, for arbitrary values of z and ε.

In the case of the segment $C_{R\oplus}$ our objective is therefore to show that, for any values of z and ε, the integral goes to zero in the $R \to \infty$ limit. On this segment we must therefore evaluate the integral

$$\int_{C_{R\oplus}} ds\, f(s) = \int_{C_{R\oplus}} ds\ \mathrm{e}^{(z-1)\ln(s)}\ \mathrm{e}^{-s}.$$

Over this segment we have that $s = R + \imath\sigma$ with $\sigma \in [0, \varepsilon]$, leading to

$ds = \imath d\sigma$. Before we work on the integral itself, let us first analyze that happens with the exponent of the first exponential in the integrand of this integral, in the $R \to \infty$ limit. We may write s in polar form as

$$s = \sqrt{R^2 + \sigma^2}\, \mathrm{e}^{\imath\theta},$$

for $\theta \in [-\pi, \pi]$, and where, over the contour $C_{R\oplus}$, for large values of R, the angle $\theta(R, \sigma)$ is small, with absolute value of the order of ε/R, being given in fact by

$$\begin{aligned} \sin(\theta) &= \frac{\sigma}{\sqrt{R^2 + \sigma^2}}, \\ \cos(\theta) &= \frac{R}{\sqrt{R^2 + \sigma^2}}. \end{aligned}$$

We therefore have for the exponent of the first exponential in the integrand at hand

$$\begin{aligned} (z-1)\ln(s) &= (z-1)\ln\!\left(\sqrt{R^2+\sigma^2}\, \mathrm{e}^{\imath\theta}\right) \\ &= [(x-1) + \imath y]\left[\ln\!\left(\sqrt{R^2+\sigma^2}\right) + \imath\theta\right] \\ &= \left[(x-1)\ln\!\left(\sqrt{R^2+\sigma^2}\right) - y\theta\right] \\ &\quad +\imath\left[(x-1)\theta + y\ln\!\left(\sqrt{R^2+\sigma^2}\right)\right], \end{aligned}$$

where the real and imaginary parts are now clearly identified. We have therefore for our integral on this segment

$$\begin{aligned} &\int_{C_{R\oplus}} ds\, f(s) \\ &\quad = \imath \int_0^{\varepsilon} d\sigma\, \mathrm{e}^{\left[(x-1)\ln\left(\sqrt{R^2+\sigma^2}\right) - y\theta\right] + \imath\left[(x-1)\theta + y\ln\left(\sqrt{R^2+\sigma^2}\right)\right]}\, \mathrm{e}^{-R}\, \mathrm{e}^{-\imath\sigma} \\ &\quad = \imath\, \mathrm{e}^{-R} \int_0^{\varepsilon} d\sigma\, \mathrm{e}^{\left[(x-1)\ln\left(\sqrt{R^2+\sigma^2}\right) - y\theta\right]}\, \mathrm{e}^{\imath\left[(x-1)\theta + y\ln\left(\sqrt{R^2+\sigma^2}\right)\right]}\, \mathrm{e}^{-\imath\sigma}. \end{aligned}$$

Taking absolute values and using the triangle inequalities we get

$$\begin{aligned} \left|\int_{C_{R\oplus}} ds\, f(s)\right| &\leq \mathrm{e}^{-R} \int_0^{\varepsilon} d\sigma\, \mathrm{e}^{(x-1)\ln\left(\sqrt{R^2+\sigma^2}\right) - y\theta} \\ &= \int_0^{\varepsilon} d\sigma\, \mathrm{e}^{-y\theta}\, \mathrm{e}^{-R + (x-1)\ln\left(\sqrt{R^2+\sigma^2}\right)}. \end{aligned} \tag{F.2}$$

Given any finite values of x and y, and considering that θ is limited in the $R \to \infty$ limit, the behavior of this integral in that limit is determined by the behavior of the exponent of this last exponential factor in the integrand. Note that we do *not* propose to take here the $\varepsilon \to 0$ limit, but *only* the $R \to \infty$ limit. We may write the relevant exponent above as

$$\begin{aligned} & -R + (x-1)\ln\left(\sqrt{R^2+\sigma^2}\right) \\ = & \; -R\left[1-(x-1)\frac{1}{R}\ln\left(R\sqrt{1+\sigma^2/R^2}\right)\right]. \end{aligned}$$

Taking the $R \to \infty$ limit we get, given that σ is limited by ε,

$$\begin{aligned} & \lim_{R\to\infty}\left[-R + (x-1)\ln\left(\sqrt{R^2+\sigma^2}\right)\right] \\ = & \lim_{R\to\infty}\left\{-R\left[1-(x-1)\frac{1}{R}\ln\left(R\sqrt{1+\sigma^2/R^2}\right)\right]\right\} \\ = & \lim_{R\to\infty}\left\{-R\left[1-(x-1)\frac{1}{R}\ln(R)\right]\right\} \\ = & \lim_{R\to\infty}(-R), \end{aligned}$$

which diverges to $-\infty$, regardless of the value of x. As a consequence of this, we see that the $R \to \infty$ limit of the integral on the right-hand side of Equation (F.2) is zero, and therefore we have that

$$\begin{aligned} \lim_{R\to\infty}\left|\int_{C_{R\oplus}} ds\, f(s)\right| & \leq \lim_{R\to\infty}\left\{\int_0^\varepsilon d\sigma\ \mathrm{e}^{-y\theta}\ \mathrm{e}^{-R+(x-1)\ln\left(\sqrt{R^2+\sigma^2}\right)}\right\} \\ & = 0 \quad \Rightarrow \\ \lim_{R\to\infty}\int_{C_{R\oplus}} ds\, f(s) & = 0. \end{aligned}$$

Note that this result holds for all x and for all y, and therefore for all z, and that it is also independent of the value of ε, so that we do *not* have to take the $\varepsilon \to 0$ limit in order for it to hold.

The case of the segment $C_{R\ominus}$ is similar to that of the segment $C_{R\oplus}$, and our objective remains the same, namely to show that, for any values of z and ε, the integral goes to zero in the $R \to \infty$ limit. On this segment we must therefore evaluate the integral

$$\int_{C_{R\ominus}} ds\, f(s) = \int_{C_{R\ominus}} ds\ \mathrm{e}^{(z-1)\ln(s)}\ \mathrm{e}^{-s}.$$

Over this segment we again have that $s = R + \imath\sigma$, but now with $\sigma \in [-\varepsilon, 0]$, leading once more to $ds = \imath d\sigma$. Except for the integration interval on σ, that plays no role in the argument presented above, the situation is exactly the same as before, so that we may conclude at once that

$$\begin{aligned}
\lim_{R\to\infty}\left|\int_{C_{R\ominus}} ds\, f(s)\right| &\leq \lim_{R\to\infty}\left\{\int_{-\varepsilon}^{0} d\sigma\ \mathrm{e}^{-y\theta}\ \mathrm{e}^{-R+(x-1)\ln\left(\sqrt{R^2+\sigma^2}\right)}\right\} \\
&= 0 \quad\Rightarrow \\
\lim_{R\to\infty}\int_{C_{R\ominus}} ds\, f(s) &= 0.
\end{aligned}$$

The conclusion that we can draw so far is that we may directly consider the integral of $f(s)$ over the contour C_∞, which turns out to be independent of the value of ε, and that can now be separated into just three parts, where the segments $C_\oplus$ and $C_\ominus$ now extend to infinity, thus becoming semi-axes,

$$\int_{C_\infty} ds\, f(s) = \int_{C_\oplus} ds\, f(s) + \int_{C_\varepsilon} ds\, f(s) + \int_{C_\ominus} ds\, f(s).$$

We will now show that, if $x > 0$, then in the $\varepsilon \to 0$ limit this integral reduces to the real integral that appears in the definition of $\Gamma(z)$ given in Equation (F.1). In order to do this, let us start by showing that, under the condition that $x > 0$, the integral over the segment C_ε vanishes in this limit.

Therefore, in the case of the segment C_ε our objective is to show that, so long as $x > 0$, the integral goes to zero in the $\varepsilon \to 0$ limit. Note that this integral does not depend at all on the value of R, regardless of whether or not we take the $R \to \infty$ limit. We start by writing this integral as

$$\int_{C_\varepsilon} ds\, f(s) = \int_{C_\varepsilon} ds\, \frac{1}{s}\ \mathrm{e}^{z\ln(s)}\ \mathrm{e}^{-s}.$$

We now observe that the complex function given by

$$g(s) = \mathrm{e}^{-s}$$

is analytic on any neighborhood of the origin, and hence is limited there. We have therefore

$$\int_{C_\varepsilon} ds\, f(s) = \int_{C_\varepsilon} ds\, \frac{1}{s}\, \mathrm{e}^{z\ln(s)}\, g(s),$$

where $g(s)$ is limited on the whole integration domain C_ε, for any value of ε. Over this segment we have that $s = \varepsilon \exp(\imath\theta)$ with $\theta \in [\pi/2, 3\pi/2]$, thus leading to $ds = \imath\varepsilon \exp(\imath\theta) d\theta$. In addition to this we have that $\ln(s) = \ln(\varepsilon) + \imath\theta$, and therefore we may write that

$$\begin{aligned}
\int_{C_\varepsilon} ds\, f(s) &= \int_{\pi/2}^{3\pi/2} d\theta\, \imath\varepsilon\, \mathrm{e}^{\imath\theta}\, \frac{1}{\varepsilon\, \mathrm{e}^{\imath\theta}}\, \mathrm{e}^{z[\ln(\varepsilon)+\imath\theta]}\, g(s) \\
&= \imath \int_{\pi/2}^{3\pi/2} d\theta\, \mathrm{e}^{z\ln(\varepsilon)}\, \mathrm{e}^{\imath z\theta}\, g(s) \\
&= \imath\, \mathrm{e}^{z\ln(\varepsilon)} \int_{\pi/2}^{3\pi/2} d\theta\, \mathrm{e}^{\imath z\theta}\, g(s) \\
&= \imath\, \mathrm{e}^{x\ln(\varepsilon)}\, \mathrm{e}^{\imath y\ln(\varepsilon)} \int_{\pi/2}^{3\pi/2} d\theta\, \mathrm{e}^{\imath x\theta}\, \mathrm{e}^{-y\theta}\, g(s),
\end{aligned}$$

where we used the fact that $z = x + \imath y$. Taking absolute values and using the triangle inequalities we get

$$\left| \int_{C_\varepsilon} ds\, f(s) \right| \leq \mathrm{e}^{x\ln(\varepsilon)} \int_{\pi/2}^{3\pi/2} d\theta\, \mathrm{e}^{-y\theta}\, |g(s)|.$$

The remaining integral on θ is the real integral of a limited function on a finite domain, and is therefore a finite real number, for any values of y and ε. If we call this real integral $I(y,\varepsilon)$, which therefore tends to a finite value in the $\varepsilon \to 0$ limit, regardless of the value of y, and take the $\varepsilon \to 0$ limit of the equation above, we get

$$\lim_{\varepsilon\to 0} \left| \int_{C_\varepsilon} ds\, f(s) \right| \leq \lim_{\varepsilon\to 0} \left[\mathrm{e}^{x\ln(\varepsilon)} I(y,\varepsilon) \right],$$

where we see that, if $x > 0$, then the exponential goes to zero in the $\varepsilon \to 0$ limit, given that $\ln(\varepsilon) \to -\infty$, while $I(y,\varepsilon)$ approaches a finite limit. Therefore, since the left-hand side of the equation above is positive and bounded from above by zero, we have that

$$\begin{aligned}
\lim_{\varepsilon\to 0} \left| \int_{C_\varepsilon} ds\, f(s) \right| &= 0 \quad \Rightarrow \\
\lim_{\varepsilon\to 0} \int_{C_\varepsilon} ds\, f(s) &= 0,
\end{aligned}$$

thus showing that, if $x > 0$, then in the $\varepsilon \to 0$ limit the integral over the segment C_ε of the contour C_∞ does not contribute to the complete integral over C_∞. The conclusion of all that we have done so far is that, if we take the $R \to \infty$ limit, and consider the corresponding integration contour C_∞, still for $\varepsilon > 0$, we may write that

$$\int_{C_\infty} ds\, f(s) = \int_{C_\oplus} ds\, f(s) + \int_{C_\varepsilon} ds\, f(s) + \int_{C_\ominus} ds\, f(s),$$

which is valid for all z. In order to relate this integral with the definition of the function $\Gamma(z)$, we consider the $\varepsilon \to 0$ limit under the condition that $x > 0$, in which case the integral over the segment C_ε vanishes, as we have just shown. Therefore, all that is left to do here is to work out what happens to the other two integrals in the $\varepsilon \to 0$ limit.

Over the segment $C_\oplus$ we have that $s = t + \imath\varepsilon$, leading to $ds = dt$ and therefore to

$$\begin{aligned}\int_{C_\oplus} ds\, f(s) &= \int_{C_\oplus} ds\ \mathrm{e}^{(z-1)\ln(s)}\ \mathrm{e}^{-s}\\ &= \int_\infty^0 dt\ \mathrm{e}^{(z-1)\ln(t+\imath\varepsilon)}\ \mathrm{e}^{-t}\ \mathrm{e}^{-\imath\varepsilon}\\ &= -\,\mathrm{e}^{-\imath\varepsilon}\int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t+\imath\varepsilon)}\ \mathrm{e}^{-t}.\end{aligned}$$

The situation over the segment $C_\ominus$ is similar to the one just discussed, with $s = t - \imath\varepsilon$ and $ds = dt$, but in this case we must observe that we have changed to the next leaf of the Riemann surface of the logarithm, and also that the orientation of the circuit is reversed, so that we get

$$\begin{aligned}\int_{C_\ominus} ds\, f(s) &= \int_{C_\ominus} ds\ \mathrm{e}^{(z-1)\ln(s)}\ \mathrm{e}^{-s}\\ &= \int_0^\infty dt\ \mathrm{e}^{(z-1)[\ln(t-\imath\varepsilon)+2\imath\pi]}\ \mathrm{e}^{-t}\ \mathrm{e}^{\imath\varepsilon}\\ &= \mathrm{e}^{2\imath\pi(z-1)}\ \mathrm{e}^{\imath\varepsilon}\int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t-\imath\varepsilon)}\ \mathrm{e}^{-t}\\ &= \mathrm{e}^{2\imath\pi z}\ \mathrm{e}^{\imath\varepsilon}\int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t-\imath\varepsilon)}\ \mathrm{e}^{-t},\end{aligned}$$

since $\exp(-2\imath\pi) = 1$. We have therefore for the sum of these two integrals

$$\begin{aligned}
&\int_{C_\oplus} ds\, f(s) + \int_{C_\ominus} ds\, f(s) \\
&\quad = \mathrm{e}^{2\imath\pi z}\ \mathrm{e}^{\imath\varepsilon} \int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t-\imath\varepsilon)}\ \mathrm{e}^{-t} \\
&\qquad - \mathrm{e}^{-\imath\varepsilon} \int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t+\imath\varepsilon)}\ \mathrm{e}^{-t}.
\end{aligned}$$

If we now take the $\varepsilon \to 0$ limit with $x > 0$ this sum of integrals becomes the integral over the contour C_∞, given that the integral over C_ε vanishes in the limit, and therefore we get

$$\begin{aligned}
&\lim_{\varepsilon\to 0} \int_{C_\infty} ds\, f(s) \\
&\quad = \lim_{\varepsilon\to 0} \left[\int_{C_\oplus} ds\, f(s) + \int_{C_\ominus} ds\, f(s) \right] \\
&\quad = \mathrm{e}^{2\imath\pi z} \int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t)}\ \mathrm{e}^{-t} - \int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t)}\ \mathrm{e}^{-t} \\
&\quad = \left(\mathrm{e}^{2\imath\pi z} - 1 \right) \int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t)}\ \mathrm{e}^{-t} \\
&\quad = 2\imath\ \mathrm{e}^{\imath\pi z}\ \frac{\mathrm{e}^{\imath\pi z} - \mathrm{e}^{-\imath\pi z}}{2\imath} \int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t)}\ \mathrm{e}^{-t} \\
&\quad = 2\imath\ \mathrm{e}^{\imath\pi z}\ \sin(\pi z) \int_0^\infty dt\ \mathrm{e}^{(z-1)\ln(t)}\ \mathrm{e}^{-t} \\
&\quad = 2\imath\,(-1)^z \sin(\pi z) \int_0^\infty dt\, t^{z-1}\, \mathrm{e}^{-t},
\end{aligned}$$

since $\exp(\imath\pi) = -1$. Here we see that the integral used to define $\Gamma(z)$ in Equation (F.1) has turned up, so that we may write that

$$\begin{aligned}
\Gamma(z) &= \int_0^\infty dt\, t^{z-1}\, \mathrm{e}^{-t} \\
&= \frac{1}{2\imath\,(-1)^z \sin(\pi z)} \lim_{\varepsilon\to 0} \int_{C_\infty} ds\, s^{z-1}\ \mathrm{e}^{-s} \\
&= \frac{\imath}{2\sin(\pi z)} \lim_{\varepsilon\to 0} \int_{C_\infty} ds\, (-s)^{z-1}\ \mathrm{e}^{-s}.
\end{aligned}$$

We see therefore that the formula on the right-hand side of this equation reproduces the known definition of $\Gamma(z)$ in the $\varepsilon \to 0$ limit. That definition is valid only for $x > 0$, but the contour integral we started with is subject to no such limitation. In addition to this we must recall that,

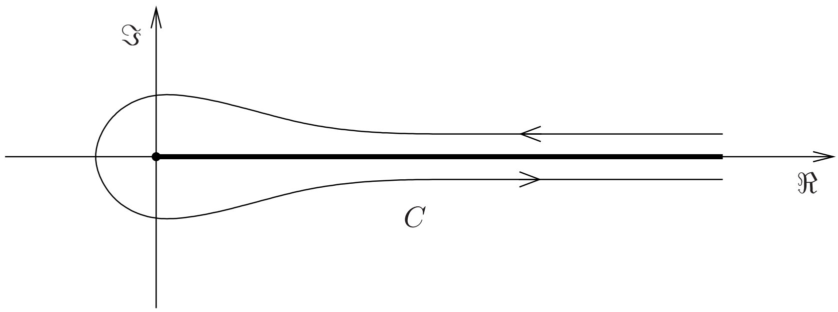


Figure F.2: The deformed complex integration contour C used for the definition by analytic continuation of the function $\Gamma(z)$, showing the single point of singularity of the integrand $f(s)$ at $s = 0$.

due to the Cauchy-Goursat theorem, we may deform the contour C_∞ at will, without changing the integral, so long as the deformation does not cross the singularity at the origin, and so long as the contour extends to infinity along the positive real semi-axis, both above and below it. Note that there is no need for the contour C resulting from the deformation to be asymptotic to the positive real semi-axis when $t \to \infty$. For example, it might be asymptotic to any pair of horizontal lines $\sigma = \pm A$, one above and another below the real axis.

Therefore, given any contour C which satisfies these conditions, such as the example shown in Figure F.2, the contour integral given above can be used to extend the definition of $\Gamma(z)$ to the whole complex plane, with the possible exception of the points at which z is an integer, as indicated by the zeros of the function $\sin(\pi z)$ in the denominator. We have therefore a new integral representation for $\Gamma(z)$,

$$\Gamma(z) = \frac{\imath}{2\sin(\pi z)} \int_C ds\, (-s)^{z-1}\, \mathrm{e}^{-s},$$

for z not an integer. This integral representation constitutes a single formula that gives $\Gamma(z)$ for all values of z except for the origin and the negative integers. Proof of the analyticity of $\Gamma(z)$ according to this more general definition can be obtained along the same lines used in the text to prove the analyticity according to the definition in Equation (F.1). This extension of the definition is another instance of the analytic continuation theorem. Note that the divergences of this formula for z equal to zero or to a negative integer are expected from what we already know about this function, since it does have simple poles at these points. However, the explicit divergences at the strictly *positive* integers are just an artifact of this representation. Note that for z an integer n with $n \geq 1$, the integration contour can in fact be closed, and hence we may write this integral representation as

$$\Gamma(n) = \frac{\imath}{2\sin(\pi n)} \oint_C ds\,(-s)^{n-1}\ \mathrm{e}^{-s},$$

where we see that the factor in front of the integral diverges when n approaches an integer, but the integral itself is zero, since in this case we have a non-negative power in the integrand, and therefore the integrand is analytic on and within C, which causes the integral to be zero by the Cauchy-Goursat theorem. We see, therefore, that what we have here is an indeterminate result of the type $0/0$. For all points other than these, this complex contour integral is a faithful representation of the function $\Gamma(z)$ over the whole complex plane.

Bibliography

[1] T. Dantzig, *Number: The Language of Science*. Plume, January 30 2007. Recommended reading; paperback, 416 pages; ISBN-10: 0452288118; ISBN-13: 978-0452288119.

[2] J. L. deLyra, "Continued fractions." This little paper can be obtained in the PS or PDF formats at the URLs

`http://latt.if.usp.br/fismat/micro-papers/continued-fractions_en.ps`
`http://latt.if.usp.br/fismat/micro-papers/continued-fractions_en.pdf`

on the web.

[3] J. L. deLyra, "Series for the square roots of integers." This little paper can be obtained in the PS or PDF formats at the URLs

`http://latt.if.usp.br/fismat/micro-papers/square-root-series_en.ps`
`http://latt.if.usp.br/fismat/micro-papers/square-root-series_en.pdf`

on the web.

[4] Fortran programs that implement the approximation schemes described in [2] and [3] can be found at the URL

`http://latt.if.usp.br/fismat/micro-papers/micropapers-src.tar.gz`

on the web.

[5] R. V. Churchill, *Complex Variables and Applications*. McGraw-Hill, second ed., 1960.

[6] J. L. deLyra, "The complex gamma function." This little paper can be obtained in the PS or PDF formats at the URLs

http://latt.if.usp.br/fismat/micro-papers/complex-gamma-func_en.ps
http://latt.if.usp.br/fismat/micro-papers/complex-gamma-func_en.pdf

on the web.

[7] J. L. deLyra, "The stirling approximation." This little paper can be obtained in the PS or PDF formats at the URLs

http://latt.if.usp.br/fismat/micro-papers/stirling-approx_en.ps
http://latt.if.usp.br/fismat/micro-papers/stirling-approx_en.pdf

on the web.

[8] J. L. deLyra, "Some facts from real analysis." This little paper can be obtained in the PS or PDF formats at the URLs

http://latt.if.usp.br/fismat/micro-papers/real-analysis_en.ps
http://latt.if.usp.br/fismat/micro-papers/real-analysis_en.pdf

on the web.

[9] J. L. deLyra, "Complete singularity classification." This little paper can be obtained in the PS or PDF formats at the URLs

http://latt.if.usp.br/fismat/micro-papers/singularity-class_en.ps
http://latt.if.usp.br/fismat/micro-papers/singularity-class_en.pdf

on the web.

[10] Y. W. Sokhotskii, *On definite integrals and functions used in series expansions*. St. Petersburg, 1873.

[11] J. L. deLyra, "Complex analysis of real functions I – complex-analytic structure and integrable real functions," *Transnational Journal of Mathematical Analysis and Applications*, vol. 6, no. 1, pp. 15–61, 2018. (ArXiv: 1708.06182, 2017).

Index

absolute convergence, 177, 181, 182, 207
  complex plane, 182
  complex scope, 181
  real definition, 177
  Taylor series, 207
algorithmic representation, 187
analytic continuation, 59, 194
  gamma function, 59
  Taylor series, 194
analytic extension, *see* analytic continuation
analytic function, 23, 27
  conformal mapping, 23
  definition, 27

basis of functions, 200
  Taylor, 200

Cauchy integral formula, 142
  for the derivatives, 145
  for the function, 142
Cauchy principal value, 257
Cauchy-Riemann conditions, 24
complete basis, 201
  Taylor, 201
complete set, 4
complex conjugation, 11
complex derivative, 19, 71, 99, 103, 111, 146
  analyticity, 111, 146
  definition, 99
  example of use, 71
  geometrical meaning, 103
complex geometrical progression, 161
complex integral, 116, 124
  analyticity, 124
complex numbers, 9, 10, 12, 17
  definition, 10
  imaginary part, 10
  imaginary unit, 10
  product, geometric, 17
  real part, 10
  sum, geometric, 17
  vectors, 12
complex plane, 17
complex potential, 122
complex rationalization, 16, 25
conformal mappings, 23, 73, 81
  example of use, 81
  proof, 73
construction of the reals, 2
convergence, 170, 175, 180, 182
  complex plane, 182
  complex scope, 180
  real definition, 175
  Taylor series, 170
convergence radius, 173
corrected capacitance formula, 92

dense subset, 4
  definition, 4
doubly-periodic functions, 46

elliptic functions, 46
equation, 66
 Laplace, $d = 2$, 66
Euler formula, 19, 21
 definition, 21
expansion, *see* series

field lines, *see* integral curves
field of numbers, 1–3, 9
 complex, 9
 ordered, 3
 rational, 3
 real, 2

gamma function, 55

harmonic functions, 66

integral curves, 68

line integral, 116

multivalued function, 38

octonions, 10
open disk, 158
open interval, 4
open neighborhood, 159
operator, 68, 103
 curl, 103
 divergence, 103
 gradient, 68

partial sum, 175
Pauli matrices, 10
pole, *see* singularity

quaternions of Hamilton, 10

regularization, 257
 of integrals, 257

residue, 236, 245
Riemann leaves, 39
Riemann surfaces, 39
 integration over, 268

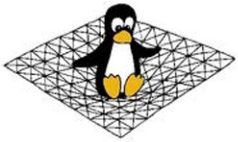
series, 162, 170, 173, 233
 Laurent, 233
 Maclaurin, 173
 Taylor, 162, 170, 173
  analytic, 173
  construction, 162
  convergence, 170
simply-periodic functions, 45
singularity, 39, 139, 227, 249–251
 branch cut, 39
 branch point, 39
 essential, definition, 250
 essential, example, 249
 isolated, 227
 order of a pole, 249
 removable, 251
 simple pole, 139, 250
spin, 10

test, 215
 ratio, 215
theorem, 19, 118, 119, 121, 126, 148, 158, 160, 245, 247
 algebra, fundamental, 160
 Cauchy-Goursat, 121
 curl, *see* Stokes theorem
 DeMoivre, 19
 Green, 118, 126
  partial proof, 126
 Liouville, 160
 maximum modulus, 158
 Morera, 148
 residue, 245, 247
  proof, 247
 Stokes, 119

triangle inequality, 18

uniform convergence, 179, 181, 191, 210, 222
complex scope, 181
integration and differentiation, 191, 222
Taylor series, 222
real definition, 179
Taylor series, 210

vector field, 23



The author is a research physicist at the Department of Mathematical Physics of the University of São Paulo. He holds a PhD degree from Yale University, and works on theoretical physics, mostly on quantum field theory, including a computational component. He has also done some work on mathematics, and is also something of a free-software, open-hardware computer enthusiast and expert.

There is a longstanding conflict between extension and depth in the teaching of mathematics to physics students. This text intends to present an approach that tries to track what could be called the ``middle way'' in this conflict. It is the result of several years of experience of the author teaching the mathematical physics courses at the Physics Institute of the University of São Paulo. The text is organized in the form of relatively short chapters, each appropriate for exposition in one lecture. Each chapter includes a list of proposed problems, which have varied levels of difficulty, including practice problems, problems that complete and extend the material presented in the text, and some longer and more difficult problems, which are presented as challenges to the students. There are complete solutions available, detailed and commented, to all the problems proposed, which are presented in separate volumes.

This volume is dedicated to the complex calculus. This is a more practical and less abstract version of complex analysis and of the study of analytic functions. This does not mean that there are no proofs in the text, since all the fundamental theorems are proved with a good level of rigor. The text starts from the very beginning, with the definition of complex numbers, and proceeds up to the study of integrals on the complex plane and on Riemann surfaces. The facts and theorems established here will be used routinely in all the subsequent volumes of this series of books. The development is based on an analogy with vector fields and with electrostatics, emphasizing interpretations and proofs that have a geometrical character. The approach is algorithmic and emphasizes the representation of functions by series, with detailed discussion of the convergence issues.

ISBN 9781793012050
90000
9 781793 012050