The Gamma Function

James Bonnar

Ψ

TREASURE TROVE OF
MATHEMATICS

February 2017

$$
\int_0^\infty \frac{1+x^2/(b+1)^2}{1+x^2/(a)^2} \times \frac{1+x^2/(b+2)^2}{1+x^2/(a+1)^2} \times \cdots \,\mathrm{d}x
$$

$$
= \frac{\sqrt{\pi}}{2} \times \frac{\Gamma(a+\frac{1}{2})\Gamma(b+1)\Gamma(b-a+\frac{1}{2})}{\Gamma(a)\Gamma(b+1/2)\Gamma(b-a+1)}
$$

for $0 < a < b + 1/2$.

Srinivasa Ramanujan, 16 January 1913

Preface

The Gamma function is an extension of the factorial function, with its argument shifted down by 1, to real and complex numbers. The Gamma function is defined by an improper integral that converges for all real numbers except the non-positive integers, and converges for all complex numbers with nonzero imaginary part. The factorial is extended by analytic continuation to all complex numbers except the non-positive integers (where the integral function has simple poles), yielding the meromorphic function we know as the Gamma function.

The Gamma function has very many extremely important applications in probability theory, combinatorics, statistical and quantum mechanics, solid-state physics, plasma physics, nuclear physics, and in the decades-long quest to unify quantum mechanics with the theory of relativity – the development of the theory of quantum gravity – the objective of string theory.

The problem of extending the factorial to non-integer arguments was apparently first considered by Daniel Bernoulli and Christian Goldbach in the 1720s, and was solved at the end of the same decade by Leonard Euler. Euler gave two different definitions: the first was an infinite product, of which he informed Goldbach in a letter dated October 13, 1729. He wrote to Goldbach again on January 8, 1730, to announce his discovery of the integral representation. Euler further discovered some of the Gamma function's important functional properties, notably the reflection formula.

Carl Friedrich Gauss rewrote Euler's product and then used his formula to discover new properties of the Gamma function. Although Euler was a pioneer in the theory of complex variables, he does not appear to have considered the factorial of a complex number, as Gauss first did. Gauss also proved the multiplication theorem of the Gamma function and investigated the connection between the Gamma function and elliptic integrals.

Karl Weierstrass further established the role of the Gamma function in complex analysis, starting from yet another product representation. Weierstrass originally wrote his product as one for $1/\Gamma$, in which case it is taken over the function's zeros rather than its poles. Inspired by this result, he proved what is known as the Weierstrass factorization theorem – that any entire function can be written as a product over its zeros in the complex plane; a generalization of the fundamental theorem of algebra.

The name of the Gamma function and its symbol Γ were introduced by Adrien-Marie Legendre around 1811; Legendre also rewrote Euler's integral definition in its modern form. The alternative "Pi function" notation $\Pi(z) = z!$ due to Gauss is sometimes encountered in older literature, but Legendre's notation is dominant in modern works. It is justified to ask why we distinguish between the "ordinary factorial" and the Gamma function by using distinct symbols, and particularly why the Gamma function should be normalized to $\Gamma(n+1) = n!$ instead of simply using "$\Gamma(n) = n!$". Legendre's motivation for the normalization does not appear to be known, and has been criticized as cumbersome by some (the 20th-century mathematician Cornelius Lanczos, for example, called it "void of any rationality" and would instead use $z!$). Legendre's normalization does simplify a few formulas, but complicates most others.

A large number of definitions have been given for the Gamma function. Although they describe the same function, it is not entirely straightforward to prove their equivalence. Instead of having to find a specialized proof for each formula, it would be highly desirable to have a general method of identifying the Gamma function given any particular form.

One way to prove equivalence would be to find a differential equation that characterizes the Gamma function. Most special functions in applied mathematics arise as solutions to differential equations, whose solutions are unique. However, the Gamma function does not appear to satisfy any simple differential equation. Otto Hölder proved in 1887 that the Gamma function at least does not satisfy any algebraic differential equation by showing that a solution to such an equation could not satisfy the Gamma function's recurrence formula. This result is known as Hölder's theorem.

A definite and generally applicable characterization of the Gamma function was not given until 1922. Harald Bohr and Johannes Mollerup then proved what is known as the Bohr-Mollerup theorem: that the Gamma function is the unique solution to the factorial recurrence relation that is positive and logarithmically convex for positive z and whose value at 1 is 1 (a function is logarithmically convex if its logarithm is convex).

The Bohr-Mollerup theorem is useful because it is relatively easy to prove logarithmic convexity for any of the different formulas used to define the Gamma function. Taking things further, instead of defining the Gamma function by any particular formula, we can choose the conditions of the Bohr-Mollerup theorem as the definition, and then pick any formula we like that satisfies the conditions as a starting point for studying the Gamma function. This approach was used by the Bourbaki group.

G.P. Michon describes the Gamma function as "Arguably, the most common special function, or the least 'special' of them. The other transcendental functions . . . are called 'special' because you could conceivably avoid some of them by staying away from many specialized mathematical topics. On the other hand, the Gamma function is most difficult to avoid."

The Gamma function finds application in diverse areas such as quantum physics, statistical mechanics and fluid dynamics. The Gamma distribution, which is formulated in terms of the Gamma function, is used in statistics to model a wide range of processes; for example, the time between occurrences of time-series events. The primary reason for the Gamma function's usefulness is the prevalence of expressions of the type $f(t)\exp(-g(t))$ which describe processes that decay exponentially in time or space.† Integrals of such expressions can often be solved in terms of the Gamma function when no elementary solution exists. For example, if f is a power function and g is a linear function, a simple change of variables yields

† See chapter 2 for further discussion of this topic.

$$\int_0^\infty t^b e^{-at}\,\mathrm{d}t = \frac{\Gamma(b+1)}{a^{b+1}}.$$

The fact that the integration is performed along the entire positive real line might signify that the Gamma function describes the cumulation of a time-dependent process that continues indefinitely, or the value might be the total of a distribution in an infinite space. It is of course frequently useful to take limits of integration other than 0 and ∞ to describe the cumulation of a finite process, in which case the ordinary Gamma function is no longer a solution; the solution is then called an incomplete Gamma function. (The ordinary Gamma function, obtained by integrating across the entire positive real line, is sometimes called the complete Gamma function for contrast.)

The Gamma function's ability to generalize factorial products immediately leads to applications in many areas of mathematics; in combinatorics, and by extension in areas such as probability theory and the calculation of power series. Many expressions involving products of successive integers can be written as some combination of factorials, the most important example perhaps being that of the binomial coefficient.

By taking limits, certain rational products with infinitely many factors can be evaluated in terms of the Gamma function as well. Due to the Weierstrass factorization theorem, analytic functions can be written as infinite products, and these can sometimes be represented as finite products or quotients of the Gamma function. For one example, the reflection formula essentially represents the sine function as the product of two Gamma functions. Starting from this formula, the exponential function as well as all the trigonometric and and hyperbolic functions can be expressed in terms of the Gamma function.

The hypergeometric function and special cases thereof, can be represented by means of complex contour integrals of products and quotients of the Gamma function, called Mellin-Barnes integrals.

The Gamma function can also be used to calculate the "volume" and "area" of n-dimensional hyperspheres.

An elegant and deep application of the Gamma function is in the study of the Riemann zeta function. A fundamental property of the Riemann zeta function is its functional equation. Among other things, it provides an explicit form for the analytic continuation of the zeta function to a meromorphic function in the complex plane and leads to an immediate proof that the zeta function has infinitely many so-called "trivial" zeros on the real line. Borwein et. al call this formula "one of the most beautiful findings in mathematics".

The Gamma function has caught the interest of some of the most prominent mathematicians of all time. In the words of Philip J. Davis, "each generation has found something of interest to say about the Gamma function. Perhaps the next generation will also." Its history reflects many of the major developments within mathematics since the 18th century.

TABLE OF CONTENTS

1

Analytic Continuation of the Factorials

Analytic continuation is a technique used to extend the domain of a given analytic function. Analytic continuation often succeeds in defining further values of a function. Consider the sum of the first n natural numbers:

$$S : \mathbb{N} \to \mathbb{N} \qquad S(n) := \sum_{k=1}^{n} k = 1 + 2 + \cdots + (n-1) + n.$$

We may simply construct a new formula, not involving summation, in the following way:

$$\begin{array}{c} 1 + 2 + \cdots + (n-1) + n \\ n + (n-1) + \cdots + 2 + 1 \\ \hline \underbrace{(n+1) \ \ + \cdots + \ \ (n+1)}_{n \text{ terms}} \end{array}$$

Gauss famously discovered this independently in his early childhood. In doing this, we have performed the sum twice, so

$$S(n) = \frac{1}{2}n(n+1). \tag{1.1}$$

It is not meaningful to speak of the first 3/2 natural numbers, for instance, but the above equation does interpolate between the integer values. We may apply the function to non-integer values and get, for example,

$$S(\frac{3}{2}) = \frac{1}{2} \cdot \frac{3}{2} \cdot \frac{5}{2} = \frac{15}{8}.$$

The factorial function $n!$ is also only defined over the positive integers:

$$n! := n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 3 \cdot 2 \cdot 1 \tag{1.2}$$

with the convention that $0! = 1$. A formula such as Equation 1.1 which would allow computation of $n!$ without having to perform the multiplication was sought by the mathematics community for many years.

Euler first discovered a formula for the factorials that allows us to compute a factorial without performing the multiplication – the so-called elementary definition of the Gamma function, is Euler's integral, which is as follows:

Definition 1.1. (Gamma function). For $0 < x < \infty$, $x \in \mathbb{R}$,

$$\Gamma(x) := \int_0^\infty t^{x-1} e^{-t}\,\mathrm{d}t. \tag{1.3}$$

The defining improper integral itself, in the real case, converges if and only if $0 < x < \infty$. This integral function is extended by analytic continuation to all real numbers except the non-positive integers, $x \notin \{\{0\} \cup \mathbb{Z}^-\}$, by the use of a recurrence formula to yield the Gamma function. The question may be asked as to how the Gamma function is computed on the negative nonintegers given that the defining integral does not converge on them. Indeed, Euler's integral does not converge for $x \leq 0$, but the function it defines over the positive reals has a unique analytic continuation to the negative reals. One way to find that analytic continuation is to use Euler's integral for positive arguments and extend the domain to negative numbers through the application of the identity,

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

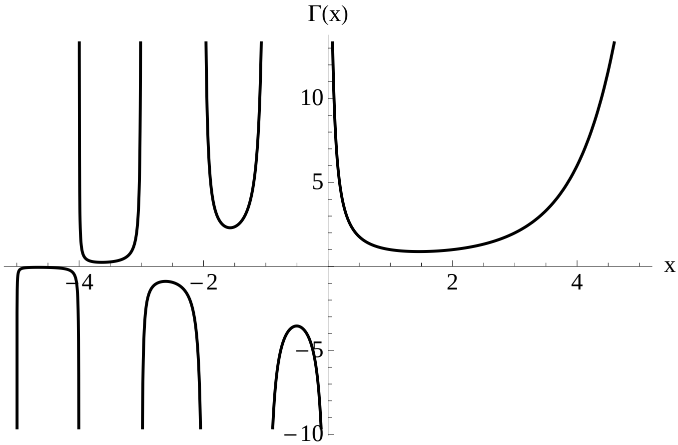


Figure 1.1: The Gamma function plotted over a portion of the real line.

choosing n sufficiently large such that $z + n$ is positive. The product in the denominator is zero when z equals any of the integers $0, -1, -2, \ldots$. Thus, the gamma function must be undefined at those points.

A restriction on the domain of the Gamma function does not equate to a restriction on the range of integration over the variable of integration (the argument t of the integrand). But rather, the argument of the Gamma function parametrizes the integrand, say $g_x(t) := t^{x-1}e^{-t}$, given in the Gamma function's definition – restrictions on the domain of the Gamma function restrict the allowed values of the parameter x defining $g_x(t)$.

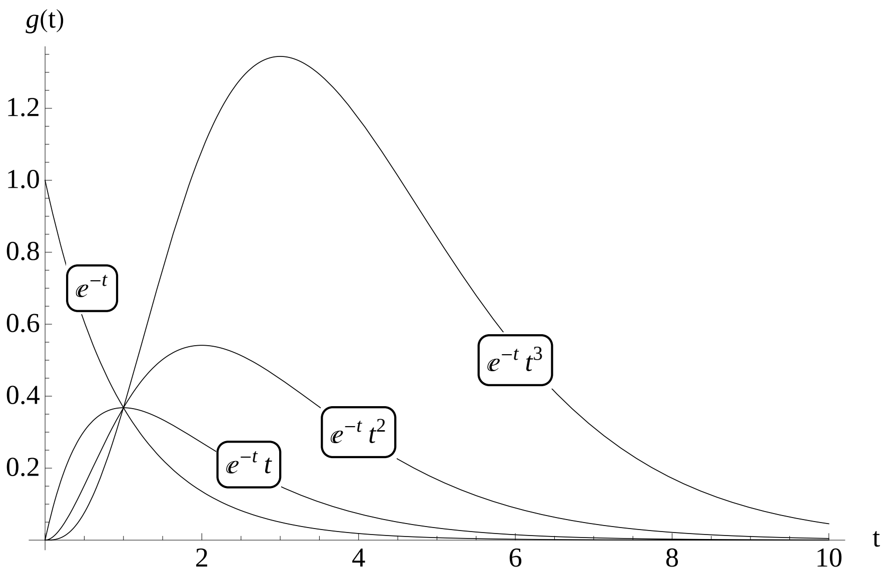


Figure 1.2: The argument of the Gamma function parametrizes the curve representing the integrand $g_x(t)$ in the Gamma function definition. It is a single curve, such as one of the curves shown here, which is integrated over the range of integration (from $t = 0$ to ∞) in the evaluation of the Gamma function. Here are shown plots of the integrand $g_x(t) = e^{-t}t^{x-1}$ for parametrizations $x = 1, 2, 3$ and 4. If one were to smoothly change the value of the parameter x from 1 to 4 then the shape of the curve would smoothly change from that of the $x = 1$ curve to that of the $x = 4$ curve.

We shall prove a result known as the *functional equation* (a recurrence relation) later in this chapter. It is defined as

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

Inverting the functional equation we have

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

When $z = 0$ it diverges because

$$\Gamma(1) = \int_0^\infty e^{-t}\,\mathrm{d}t = \lim_{m\to\infty} \left[-e^{-t}\right]_0^m = \lim_{m\to\infty} (-e^{-m}+1) = 1$$

is finite.

Formally, we can discover that the Gamma function has simple poles in all negative integers by simply iterating the recursion:

$$\begin{aligned}\Gamma(-n) &= \frac{1}{-n}\Gamma(-n+1) \\ &= \cdots \\ &= \frac{(-1)^n}{n!}\Gamma(0) \qquad (n \in \mathbb{N}).\end{aligned}$$

The power function factor that appears within the integrand of the Gamma function definition may have a complex exponent. What does it mean to raise a real number to a complex power? If t is a positive real number, and z is any complex number, then t^{z-1} is defined as $e^{(z-1)\log(t)}$, where $x = \log(t)$ is the unique real solution to the equation $e^x = t$. By Euler's formula, $e^{i\theta} = \cos\theta + i\sin\theta$, so we have for example:

$$\begin{aligned}2^{3+4i} &= 2^3 \cdot 2^{4i} = 8e^{4i\log 2} \\ &= 8(\cos(\log 2) + i\cdot\sin(\log 2))^4 \\ &\approx -7.4615 + 2.88549i\end{aligned}$$

Though we can straightforwardly compute the complex power of a real number in the analytical sense, a more intuitive interpretation of taking the complex power of a real number is elusive. Nonetheless, by virtue of the recurrence relation, $\Gamma(z)$ can be interpreted as a generalization of the factorial function to complex numbers. Since $\Gamma(z) = \Gamma(z+1)/z$ and $\Gamma(z+1)$ is defined for $\mathrm{Re}(z) > -1$ we can extend analytically the domain of $\Gamma(z)$ to the strip $-1 < \mathrm{Re}(z) < 0$. Repeating the same argument, we can extend $\Gamma(z)$ to all values of the complex plane, except for $z = 0, -1, -2, \ldots$.

Since the argument of the Gamma function may be complex, we may restate the definition of the Gamma function in the following way:

Definition 1.2. (Gamma function). For $|z| < \infty$, $z \in \mathbb{C}$,

$$\Gamma(z) := \int_0^\infty t^{z-1} e^{-t}\,\mathrm{d}t. \tag{1.4}$$

The improper integral itself converges iff $\mathrm{Im}(z) \neq 0$, or $z \in \mathbb{R}$ and $0 < z < \infty$. The integral function is extended by analytic continuation to all complex numbers except the non-positive integers, $z \notin \{\{0\} \cup \mathbb{Z}^-\}$, through the use of a recurrence formula yielding the Gamma function.

Let's take a closer look at definition 1.2. Notice that we are integrating over the range $[0, \infty)$ – here we have an

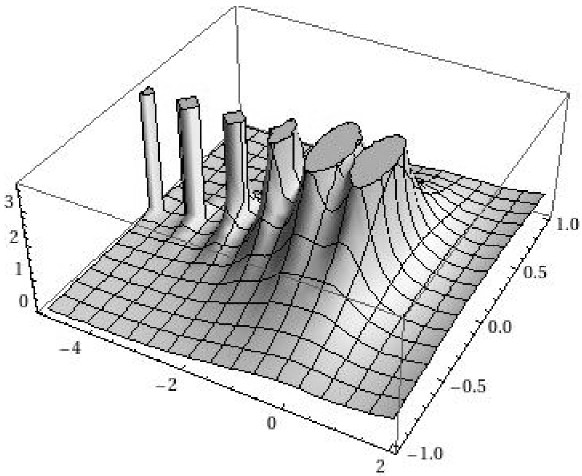


Figure 1.3: A surface plot of the absolute value of the Gamma function over a portion of the complex plane that illustrates the pattern of its domain of convergence, defined as $z \in \mathbb{C}$ such that $z \in \mathbb{R}$ and $z \notin \{\{0\} \cup \mathbb{Z}^-\}$, or $\mathrm{Im}(z) \neq 0$.

example of an *improper integral*, which are defined by a limiting process:

$$\int_m^\infty f(t)\,\mathrm{d}t := \lim_{n\to\infty} \int_m^n f(t)\,\mathrm{d}t.$$

If $\mathrm{Im}(z) = 0$, then when $z \leq 1$, we have that t^{z-1} is undefined at $t = 0$, so we must, in that case, define our improper integral in terms of two limits:

$$\int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t := \lim_{n\to\infty} \lim_{m\to 0} \int_m^n t^{z-1}e^{-t}\,\mathrm{d}t.$$

Next, let's verify that the integral in definition 1.2 converges. To this effect, we shall need the following lemma:

Lemma 1.1. ***Comparison Test for Improper Integrals.*** *If* $f(x) \geq g(x) \geq 0$ *on the interval* $[a, \infty)$ *then,*

- *If* $\int_a^\infty f(x)\,\mathrm{d}x$ *converges, then so does* $\int_a^\infty g(x)\,\mathrm{d}x$*.*
- *If* $\int_a^\infty g(x)\,\mathrm{d}x$ *diverges, then so does* $\int_a^\infty f(x)\,\mathrm{d}x$*.*

We shall take this as intuitively obvious and not pursue a formal proof here.

Theorem 1.1. $\int_0^\infty t^{z-1}e^{-t}\,\mathrm{d}t$ *converges for all* $\mathrm{Re}(z) \in (0, \infty)$*.*

Proof. Let $x = \mathrm{Re}(z)$ and split the integral into a sum of two terms:

$$\int_0^\infty t^{x-1}e^{-t}\,\mathrm{d}t = \int_0^1 t^{x-1}e^{-t}\,\mathrm{d}t + \int_1^\infty t^{x-1}e^{-t}\,\mathrm{d}t.$$

The strategy is to show that each of the two integrals on the right-hand side converge – in both cases we construct $f(t)$ such that $0 \leq t^{x-1}e^{-t} \leq f(t)$ and show that the integral of $f(t)$ over $(0, \infty)$ is finite, and hence so is the integral of $t^{x-1}e^{-t}$.

In the case of the first integral, since $e^{-t} \leq 1$ for $t \geq 0$, we have that $0 \leq t^{x-1}e^{-t} \leq t^{x-1}$ for $t \geq 0$. Then

$$\begin{aligned} 0 \leq \int_0^1 t^{x-1}e^{-t}\,\mathrm{d}t &\leq \int_0^1 t^{x-1}\,\mathrm{d}t \\ = \lim_{\alpha\to 0^+} \left[\frac{t^x}{x}\right]_{t=\alpha}^1 &= \lim_{\alpha\to 0^+}\left(\frac{1}{x} - \frac{\alpha^x}{x}\right). \end{aligned}$$

If $x > 0$, then $\alpha^x \to 0$ as $\alpha \to 0^+$, so that $\int_0^1 t^{x-1}\,\mathrm{d}t$ converges to $\frac{1}{x}$. Hence $\int_0^1 t^{x-1}e^{-t}\,\mathrm{d}t$ converges.

In the case of the second integral, first note that $t^q e^{-t/2} \to 0$ as $t \to \infty$ for any $q \in \mathbb{R}$. Hence for any $x \in \mathbb{R}$, there exists ϵ such that $0 \leq t^{x-1}e^{-t/2} \leq 1$ for $t = \epsilon$. So we further split the second integral at ϵ:

$$\int_1^\infty t^{x-1}e^{-t}\,\mathrm{d}t = \int_1^\epsilon t^{x-1}e^{-t}\,\mathrm{d}t + \int_\epsilon^\infty t^{x-1}e^{-t}\,\mathrm{d}t.$$

The first term is finite, being a finite integral. For the second term, when $t \geq \epsilon$ we have

$$t^{x-1}e^{-t} = (t^{x-1}e^{-t/2})e^{-t/2} \leq e^{-t/2},$$

thus

$$\int_\epsilon^\infty t^{x-1}e^{-t}\,\mathrm{d}t \leq \int_\epsilon^\infty e^{-t/2}\,\mathrm{d}t = \lim_{\alpha\to\infty}\left[-2e^{-t/2}\right]_{t=\epsilon}^{\alpha}$$

$$= \lim_{\alpha\to\infty}(2e^{-\epsilon/2} - 2e^{-\alpha/2}) = 2e^{-\epsilon/2}.$$

This shows the second term is convergent, thus $\int_1^\infty t^{x-1}e^{-t}\,\mathrm{d}t$ converges. All said, $\int_0^\infty t^{x-1}e^{-t}\,\mathrm{d}t$ converges for all x such that $0 < x < \infty$ by lemma 1.1.

$\square$

Next, it is proper to ask what the definition of $\Gamma(x)$ *means*. It has already been said that $\Gamma(x)$ extends the factorial function to $0 < x < \infty$, but we have not shown this – in order to do so, we need the following result:

Proposition 1.2. *For all $x > 0$, $\Gamma(x+1) = x\Gamma(x)$.*

Proof. From definition 1.1,

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

Integration by parts[†], with $f(t) = t^x$ and $g(t) = -e^{-t}$, yields

$$\int_0^\infty t^x e^{-t}\,\mathrm{d}t - \int_0^\infty x t^{x-1} e^{-t}\,\mathrm{d}t = \left[-t^x e^{-t}\right]_0^\infty.$$

Rearranging,

$$\begin{aligned}\Gamma(x+1) &= \lim_{m\to\infty} \left[-t^x e^{-t}\right]_0^m + \int_0^\infty x t^{x-1} e^{-t}\,\mathrm{d}t \\ &= \lim_{m\to\infty} (-e^{-m} m^x) + x\int_0^\infty t^{x-1} e^{-t}\,\mathrm{d}t.\end{aligned}$$

† The product rule says that

$$(f \cdot g)' = f \cdot g' + f' \cdot g,$$

from which it follows that

$$\int_a^b f \cdot g'\,\mathrm{d}x = -\int_a^b f' \cdot g\ \mathrm{d}x + f \cdot g\ \big|_a^b.$$

Now, as $m \to \infty$, $m^x e^{-m} \to 0$, so we have

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

$\square$

The recurrence relation, also known as the *functional equation*,

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

is valid for all complex arguments z for which the integral in definition 1.2 converges.

We may extend the functional equation to negative values through inversion,

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

so for example $\Gamma(-1/2) = -2\Gamma(1/2)$. Reiteration of this identity allows us to define the Gamma function on the whole real axis except on the nonpositive integers $(0, -1, -2, \dots)$.

That said, we may now show the following:

Theorem 1.3. *For all $n \in \mathbb{N}$, $\Gamma(n+1) = n!$.*

Proof. From proposition 1.2, we have

$$\begin{aligned}\Gamma(n+1) &= n \cdot \Gamma(n) \\ &= n \cdot (n-1) \cdot \Gamma(n-1) \\ &\vdots \\ &= n \cdot (n-1) \cdot \dots \cdot 2 \cdot 1 \cdot \Gamma(1) \\ &= n! \cdot \Gamma(1).\end{aligned}$$

It remains to show that $\Gamma(1) = 1 = 0!$. By definition 1.1, we have

$$\Gamma(1) = \int_0^\infty e^{-t}\,\mathrm{d}t = \lim_{m\to\infty}\left[-e^{-t}\right]_0^m = \lim_{m\to\infty}(-e^{-m}+1) = 1.$$

□

Euler wrote a letter on January 8, 1730 to Christian Goldbach in which he proposed the following definition for the Gamma function:

Definition 1.3. Let $x > 0$,

$$\Gamma(x) := \int_0^1 (-\log(t))^{x-1}\,\mathrm{d}t \tag{1.5}$$

We can show this historical definition equivalent to its more usual form, definition 1.1, through a simple change of variable $u = -\log(t)$.

Proof. Let $u = -\log(t)$. Then we have

$$t = e^{-u} \quad \text{and} \quad \mathrm{d}u = -\frac{\mathrm{d}t}{t}.$$

Our corresponding limits of integration are

$$t = 0 \to u = \infty \quad \text{and} \quad t = 1 \to u = 0.$$

We thus have

$$-\int_{\infty}^{0} u^{x-1} e^{-u}\,\mathrm{d}u = \int_{0}^{\infty} u^{x-1} e^{-u}\,\mathrm{d}u$$

where a simple relabeling yields definition 1.1. $\square$

Other forms of the Gamma function are obtained through a simple change of variables, as follows:

$$\Gamma(z) = 2\int_0^{\infty} y^{2z-1} e^{-y^2}\,\mathrm{d}y \qquad \text{by letting } t = y^2$$

$$\Gamma(z) = \int_0^1 \left(\ln \frac{1}{y}\right)^{z-1} \mathrm{d}y \qquad \text{by letting } e^{-t} = y.$$

Definition 1.4. Derivatives of the Gamma Function. The derivatives of the Gamma function can be calculated by straightforward differentiation under the integral sign:

$$\Gamma'(x) = \int_0^{\infty} t^{x-1} e^{-t} \log(t)\,\mathrm{d}t$$

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

where the identity $\frac{\mathrm{d}}{\mathrm{d}x}a^x = a^x \log(a)$ is used[†].

So how do we actually compute the Gamma function? It was previously mentioned that we can compute the Gamma function for negative argument using the recursion formula:

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

where we choose n sufficiently large such that $z+n$ is positive. Another formula for computing the Gamma function for negative argument involves Euler's reflection formula (discussed later):

$$\Gamma(-z) = \frac{-\pi}{z\Gamma(z)\sin \pi z}.$$

However both of the above two formulas rely upon an ability to compute the Gamma function for positive argument.

The integral definition is not very useful in terms of efficiency; to produce an accurate result, an extremely high

† This derivative identity follows from a simple application of the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}a^x = \frac{\mathrm{d}}{\mathrm{d}x}e^{x \log a} = e^{\log a^x} \cdot \log a = a^x \cdot \log a.$$

number of terms would have to be added during some numerical integration procedure. There are several other ways, possessing various degrees of efficiency, to numerically compute the Gamma function. The simplest to understand is as follows: A series expansion (derived later) for $\log(\Gamma(1+x))$ exists such that for $|x| < 1$ we have

$$\log(\Gamma(1+x)) = -\gamma x + \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k} x^k. \tag{1.6}$$

It follows easily from the equation above and the functional equation that

$$\frac{1}{\Gamma(z)} = z \exp\left(\gamma z - \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k) z^k}{k}\right). \tag{1.7}$$

Other formulas for computing the Gamma function include the *Lanczos approximation*, which is used to compute the Gamma function to arbitrary precision, and *Stirling's formula*, which is an asymptotic formula used to approximate the value of the Gamma function given very large arguments when absolute precision is less important, and is discussed in a later chapter.

2

Integrals Involving a Decaying Exponential

We begin with the Gamma function, which is defined by Euler's integral:

$$\Gamma(z) \equiv \int_0^\infty \mathrm{d}x\, x^{z-1} e^{-x}, \quad \text{for } \operatorname{Re}(z) > 0.$$

It can be immediately recognized that if we set $z = n+1$ then we have

$$\Gamma(n+1) = \int_0^\infty t^n e^{-t}\, \mathrm{d}t, \quad \text{for } \operatorname{Re}(n) > -1.$$

We arrive at a useful relation if we make the substitution $x = \nu t$ such that $\mathrm{d}x = \nu\, \mathrm{d}t$ in the evaluation of the following integral:

$$\int_0^\infty \mathrm{d}t\, t^n e^{-\nu t} = \int_0^\infty \frac{\mathrm{d}x}{\nu}\left(\frac{x}{\nu}\right)^n e^{-x}$$
$$= \frac{\int_0^\infty x^n e^{-x}\,\mathrm{d}x}{\nu^{n+1}} = \frac{\Gamma(n+1)}{\nu^{n+1}}.$$

Integrals over temporally or spatially decaying processes (such as collisional damping at rate $\nu \sim 1/\tau$) often result in integrals of the form

$$\int_0^\infty \mathrm{d}t\, t^n e^{-t/\tau} = \tau^{n+1}\int_0^\infty \mathrm{d}x\, x^n e^{-x} \quad \text{where } x \equiv t/\tau.$$

Two values of the argument z of fundamental interest for Gamma functions are $z = 1$ and $z = 1/2$. For $z = 1$ the Gamma function becomes simply the integral of a decaying exponential:

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

For $z = 1/2$, by using the substitution $x = u^2$ the Gamma function becomes the integral of a Gaussian distribution over an infinite domain:†

$$\Gamma(\frac{1}{2}) = 2\int_0^\infty \mathrm{d}u\, u(u^2)^{-1/2}e^{-u^2} = 2\int_0^\infty \mathrm{d}u\, e^{-u^2} = \sqrt{\pi}.$$

† Refer to Chapter 5 for a proof of this result.

When the argument of the Gamma function is a positive integer (given by the map $z \to n > 0$), the Gamma function simplifies to a factorial function:

$$\Gamma(n+1) = n\Gamma(n) = \cdots = n(n-1)(n-2)\cdots 1 \equiv n!.$$

Using this factorial form for the Gamma function, one thus finds that

$$\int_0^\infty \mathrm{d}t\, t^n e^{-t/\tau} = \tau^{n+1} n!, \quad \text{for } n = 0, 1, 2, \ldots$$

using the usual convention that $0! \equiv 1$. The first few of these integrals are

$$\int_0^\infty \frac{\mathrm{d}t}{\tau} \left\{ \begin{array}{c} 1 \\ t/\tau \\ t^2/\tau^2 \end{array} \right\} e^{-t/\tau} = \int_0^\infty \mathrm{d}x \left\{ \begin{array}{c} 1 \\ x \\ x^2 \end{array} \right\} e^{-x} = \left\{ \begin{array}{c} 1 \\ 1 \\ 2 \end{array} \right\}.$$

When the argument of the Gamma function is a positive half-integer (given by the map $z \to n + 1/2 > 0$), the Gamma function simplifies to a double factorial:†

$$\begin{aligned} \Gamma\left(n + \frac{1}{2}\right) &= \left(n - \frac{1}{2}\right) \Gamma\left(n - \frac{1}{2}\right) \\ &= \left(n - \frac{1}{2}\right)\left(n - \frac{3}{2}\right) \Gamma\left(n - \frac{3}{2}\right) \\ &= [(2n-1)(2n-3)\cdots 1]\Gamma\left(\frac{1}{2}\right)/2^n \\ &\equiv (2n-1)!!\sqrt{\pi}/2^n. \end{aligned}$$

† The product of all odd integers up to some odd positive integer n is often called the double factorial of n. It is denoted by $n!!$.

3

Hölder's Theorem

A large number of definitions have been given for the Gamma function. Although they describe the same function, it is not entirely straightforward to prove their equivalence. Instead of having to find a specialized proof for each formula, it would be highly desirable to have a general method of identifying the Gamma function given any particular form.

One way to prove equivalence would be to find a differential equation that characterizes the Gamma function. Most special functions in applied mathematics arise as solutions to differential equations, whose solutions are unique. However, the Gamma function does not appear to satisfy any simple differential equation. Otto Hölder proved in 1887 that the Gamma function at least does not satisfy any algebraic differential equation by showing that a solution to such an equation could not satisfy the Gamma function's recurrence formula. This result is known as Hölder's theorem.

A definite and generally applicable characterization of the Gamma function was not given until 1922. Harald Bohr and Johannes Mollerup then proved what is known as the Bohr-Mollerup theorem: that the Gamma function is the unique solution to the factorial recurrence relation that is positive and logarithmically convex for positive z and whose value at 1 is 1 (a function is logarithmically convex if its logarithm is convex).

The Bohr-Mollerup theorem (discussed in the next chapter) is useful because it is relatively easy to prove logarithmic convexity for any of the different formulas used to define the Gamma function. Taking things further, instead of defining the Gamma function by any particular formula, we can choose the conditions of the Bohr-Mollerup theorem as the definition, and then pick any formula we like that satisfies the conditions as a starting point for studying the Gamma function.

The constant $\pi = 3.14159\ldots$ represents the ratio of the circumference of a circle to its diameter. The number π, like many other fundamental mathematical constants such as $e = 2.71828\ldots$, is a *transcendental number*. Both π and e possess an infinite number of digits which appear to have no orderly pattern to their arrangement. Transcendental numbers cannot be expressed as the root of any algebraic equation that contains only rational numbers. In general, *transcendental* means nonalgebraic.

To be clear, consider the numbers $\sqrt{2}$ and π. Irrational means 'cannot be expressed as a fraction'. $\sqrt{2}$ is an irrational number but is not transcendental, whereas π is both transcendental and (is therefore) irrational. Any algebraic number raised to the power of an irrational number is necessarily a transcendental number.

A *transcendental function* does not satisfy any polynomial equation whose coefficients are themselves polynomials, in contrast to an algebraic function, which does satisfy such an equation. A transcendental function $f(x)$ is not expressible as a finite combination of the algebraic operations (employing only rationals) of addition, subtraction, multiplication, division, raising to a power, and extracting a root applied to x. Examples include the functions $\log x$, $\sin x$, $\cos x$, e^x and any functions containing them. Such functions are expressible in algebraic terms only as infinite series – in this sense a transcendental function is a function that "transcends" algebra.

The following functions are transcendental:

$$\begin{aligned} f_1(x) &= x^\pi \\ f_2(x) &= c^x, \quad c \neq 0,1 \\ f_3(x) &= x^x \\ f_4(x) &= x^{\frac{1}{x}} \\ f_5(x) &= \log_c x, \quad c \neq 0,1 \\ f_6(x) &= \sin x \end{aligned}$$

Formally, an analytic function $f(z)$ of one real or complex variable z is transcendental if it is algebraically independent of that variable. This can be extended to functions of several variables.

Hölder's theorem states that the Gamma function does not satisfy any algebraic differential equation whose coefficients are rational functions. The result was first proved by Otto Hölder in 1887. The theorem also generalizes to the q-gamma function.

Theorem 3.1. ***(Hölder's Theorem).*** *There exists no non-constant polynomial $P(x; y_0, y_1, \ldots, y_n)$ such that*

$$P(x; \Gamma(x), \Gamma'(x), \ldots, \Gamma^{(n)}(x)) \equiv 0$$

where $y_0, y_1, \ldots, y_n$ are functions of x, $\Gamma(x)$ is the Gamma function, and P is a polynomial in $y_0, y_1, \ldots, y_n$ with coefficients drawn from the ring of polynomials in x. That is to say,

$$P(x; y_0, y_1, \ldots, y_n) = \sum_{(a_0,a_1,\ldots,a_n)} A_{(a_0,a_1,\ldots,a_n)}(x) \cdot (y_0)^{a_0} \cdot \ldots \cdot (y_n)^{a_n}$$

where the $(a_0, a_1, \ldots, a_n)$ *index all possible terms of the polynomial and the* $A_{(a_0,a_1,\ldots,a_n)}(x)$ *are polynomials in* x *acting as coefficients of polynomial* P*. The* $A_{(a_0,a_1,\ldots,a_n)}(x)$ *may be constants or zero.*

Example 3.1. If $P(x; y_0, y_1, y_2) = x^2y_2+xy_1+(x^2-\alpha^2)y_0$ then $A_{(0,0,1)}(x) = x^2$, $A_{(0,1,0)}(x) = x$ and $A_{(1,0,0)}(x) = (x^2-\alpha^2)$ where α is a constant. All the other coefficients in the summation are zero. Then

$$P(z; f, f', f'') = x^2f'' + xf' + (x^2 - \alpha^2)f = 0$$

is an algebraic differential equation which, in this example, has solutions $f = J_\alpha(x)$ and $f = Y_\alpha(x)$, the Bessel functions of either the first or second kind. So

$$P(x; J_\alpha(x), J'_\alpha(x), J''_\alpha(x)) \equiv 0$$

and therefore both $J_\alpha(x)$ and $Y_\alpha(x)$ are differentially algebraic (but are algebraically transcendental).

The vast majority of the familiar special functions of mathematical physics are differentially algebraic. All algebraic combinations of differentially algebraic functions are also differentially algebraic. Also, all compositions of differentially algebraic functions are differentially algebraic. Hölder's theorem simply states that the gamma function, $\Gamma(x)$, is not differentially algebraic and is, therefore, a transcendentally transcendental (or equivalently, hypertranscendental) function.

A hypertranscendental function is a function which is not the solution of an algebraic differential equation with coefficients in $\mathbb{Z}$ (the integers) and with algebraic initial conditions. The term was introduced by D. D. Morduhai-Boltovskoi in 1949. Hypertranscendental functions usually arise as the solutions to functional equations, the Gamma function being one example.

A functional equation is any equation that specifies a function in implicit form. Often, the equation relates the value of a function at some point with its values at other points. For instance, properties of functions can be determined by considering the types of functional equations they satisfy. The term functional equation usually refers to equations that cannot be simply reduced to algebraic equations.

A few simple lemmas will greatly assist us in proving Hölder's theorem.

In what follows, the minimal polynomial of an algebraic number ζ is the unique irreducible monic polynomial of smallest degree $m(x)$ with rational coefficients such that $m(\zeta) = 0$ and whose leading coefficient is 1. The minimal polynomial of an algebraic number α divides any other polynomial with rational coefficients $p(x)$ such that $p(\alpha) = 0$. It follows that it has minimal degree among all polynomials f with this property.

The following table lists some algebraic numbers ζ alongside their minimal polynomials $m(x)$, computed using the Mathematica function `MinimalPolynomial`:

ζ	$m(x)$
2	$x-2$
$2^{1/2}$	x^2-2
$2^{1/3}$	x^3-2
$2^{1/3}+5$	$x^3-15x^2+75x-127$
$2^{1/2}+3^{1/2}$	x^4-10x^2+1

Lemma 3.1. *The minimal polynomial of an algebraic number ζ is unique.*

Proof. If we had two such polynomials, they must both have the same degree and the same leading coefficient 1, and so their difference is a polynomial of smaller degree which still gives 0 when applied to ζ. But this would contradict the minimality of m. $\square$

Lemma 3.2. *If p is some polynomial such that $p(\zeta)=0$, then m divides p.*

Proof. By definition, $\deg(p) \geq \deg(m)$. We may write $p = qm + r$ for some polynomials q, r, such that $\deg(r) < \deg(m)$. Then since $m(\zeta) = 0$ and $p(\zeta) = 0$, we have that $r(\zeta) = p(\zeta) - q(\zeta)m(\zeta) = 0$, which contradicts the minimality of m, unless $r(x) \equiv 0$. Therefore $r(x) \equiv 0$ and m divides p. $\square$

Proof. **(Hölder's Theorem).** Assume the existence of P as described in the statement of the theorem, that is

$$P(x; \Gamma(x), \Gamma'(x), \ldots, \Gamma^{(n)}(x)) \equiv 0$$

with

$$P(x; y_0, y_1, \ldots, y_n) = \sum_{(a_0, a_1, \ldots, a_n)} A_{(a_0, a_1, \ldots, a_n)}(x) \cdot (y_0)^{a_0} \cdot \ldots \cdot (y_n)^{a_n}.$$

Also assume that P is of lowest possible degree. This means that all the coefficients $A_{(a_0, a_1, \ldots, a_n)}$ have no common factor of the form $(x - \gamma)$ and so P is not divisible by any factor of $(x - \gamma)$. It also means that P is not the product of any two polynomials of lower degree. Consider the relations

$$\begin{aligned} &P(x+1; \Gamma(x+1), \Gamma^{(1)}(x+1), \ldots, \Gamma^{(n)}(x+1)) = \\ &\quad = P\left(x+1; x\Gamma(x), [x\Gamma(x)]^{(1)}, [x\Gamma(x)]^{(2)}, \ldots, [x\Gamma(x)]^{(n)}\right) \\ &\quad = P\left(x+1; x\Gamma(x), x\Gamma^{(1)}(x) + \Gamma(x), \ldots, x\Gamma^{n}(x) + n\Gamma^{(n-1)}(x)\right) \end{aligned}$$

so we can define a second polynomial Q, defined by the transformation

$$Q(x; y_0, y_1, \ldots, y_n) = P\left(x+1; xy_0, xy_1 + y_0, \ldots, xy_n + ny_{(n-1)}\right)$$

and $Q\left(x; \Gamma(x), \Gamma'(x), \ldots, \Gamma^{(n)}(x)\right) = 0$ is also an algebraic differential equation for $\Gamma(x)$. Q and P both have the same degree and P must divide Q otherwise there would be a remainder and that would mean P was not of minimal degree. Since $(xy_n + ny_{n-1})^{h_n} = x^{h_n} y_n^{h_n} + \cdots$, the highest degree term of Q is

$$x^{h_0+h_1+\cdots+h_n} A_{(h_0,h_1,\ldots,h_n)}(x+1) \cdot (y_0)^{h_0} \cdot (y_1)^{h_1} \cdot \ldots \cdot (y_n)^{h_n}$$

where $(h_0, h_1, \ldots, h_n)$ are the exponents of the highest degree term of P.

Let $R(x)$ be the ratio between P and Q:

$$\begin{aligned} Q(x; y_0, y_1, \ldots, y_n) &= P\left(x+1; xy_0, xy_1 + y_0, \ldots, xy_n + ny_{(n-1)}\right) \\ &= R(x) P(x; y_0, y_1, \ldots, y_n). \end{aligned}$$

Suppose $R(x)$ has a zero, say γ. Substitute γ into

$$P(\gamma+1; \gamma y_0, \gamma y_1 + y_0, \ldots, \gamma y_n + ny_{(n-1)}) = 0 \cdot P(x; y_0, y_1, \ldots, y_n) = 0.$$

This last equality indicates that $(z-(\gamma+1))$ is a factor of P, contradicting the assumption that P was of minimal degree.

Now consider the two leading terms, which must satisfy the equality

$$\begin{aligned} &R(x) A_{(h_0,\ldots,h_n)}(x) \cdot (y_0)^{h_0} \cdot \ldots \cdot (y_n)^{h_n} \\ &\quad = x^{h_0+\cdots+h_n} A_{(h_0,\ldots,h_n)}(x+1) \cdot (y_0)^{h_0} \cdot \ldots \cdot (y_n)^{h_n} \end{aligned}$$

$$R(x) A_{(h_0,\ldots,h_n)}(x) = x^{h_0+\cdots+h_n} A_{(h_0,\ldots,h_n)}(x+1).$$

This equality cannot be satisfied for arbitrary x if $R \equiv c$, where c is constant. Therefore, no such P exists and $\Gamma(x)$ is not differentially algebraic. $\square$

4

Bohr-Mullerup Theorem

Theorem 1.3 shows that the Gamma function extends the factorial function from the set of natural numbers $\mathbb{N}$ to real non-null positive numbers. But the Gamma function is not the only way to do so. Consider, for instance, functions of the form $\cos(2m\pi x)\Gamma(x)$, where m is any non-null integer. We already know $\Gamma(x+1) = x\Gamma(x)$, so it follows

$$\begin{aligned}
\cos(2m\pi(x+1))\Gamma(x+1) &= x\cos(2m\pi(x+1))\Gamma(x) \\
&= x\cos(2m\pi x + 2m\pi)\Gamma(x) \\
&= x\cos(2m\pi x)\Gamma(x)
\end{aligned}$$

thereby satisfying the functional equation. Hadamard proposed another alternative:

$$F(x) := \frac{1}{\Gamma(1-x)} \frac{d}{dx} \log\left(\frac{\Gamma(\frac{1-x}{2})}{\Gamma(1-\frac{x}{2})}\right).$$

which gives $F(n+1) = n!$ for all $n \in \mathbb{N}$. In principle, there are an infinite number of possibilities, since we may draw *any* curve through the points $(1, 1), (2, 2), (3, 6), (4, 24), \ldots,$ and assume it represents *some* function that returns factorials at integer values. The purpose of this section is to explain why we consider the Gamma function as *the* function which extends the factorial function to all real $x > 0$. We seek a condition which implies that any such function cannot be anything else other than $\Gamma(x)$.

The functional equation does not only apply to the natural numbers; the relation $\Gamma(x+1) = x\Gamma(x)$ is valid for any $x > 0$. This is a stronger requirement than $\Gamma(n+1) = n!$ alone is, since it implies that the values of Γ in any range $[x, x+1]$ determine the value of Γ on the entire real line. This is a fair amount of restriction, but not enough to meet our goal.

To distinguish the Gamma function amongst all the possible continuations of the factorial function, the notion of *convexity* is introduced. A standard definition of a convex function is a necessarily continuous function whose value at the midpoint of every interval in its domain does not exceed the arithmetic mean of its values at the ends of the interval. Convex functions can be defined formally as follows:

Definition 4.1. A function $f : (a,b) \to \mathbb{R}$ is called convex if and only if for any $x, y \in (a,b)$, we have

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda) f(y)$$

for all $\lambda \in (0,1)$.

We have parametrized the interval (x,y) as $\{\lambda x + (1-\lambda)y \mid 0 < \lambda < 1\}$. The above definition means that as you move from x to y, the line joining $(x, f(x))$ to $(y, f(y))$ always lies above the graph of f.

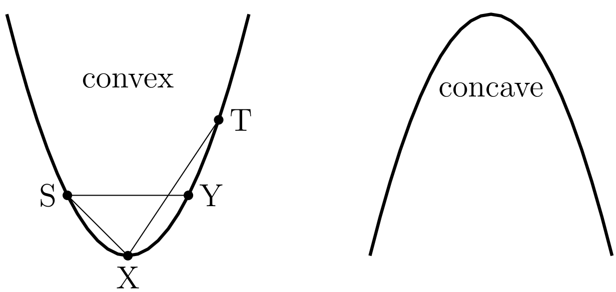


Figure 4.1: Convex versus concave functions.

Any convex function on an open interval must be continuous, which we will now show.

Lemma 4.1. *Any convex function $f : (a,b) \to \mathbb{R}$ is continuous.*

Proof. We shall follow the proof given by Rudin. The proof uses geometric language. See Figure 4.1 for the corresponding diagram.

Suppose $a < s < x < y < t < b$. The points should not be considered fixed, just that the preceding inequalities are satisfied. Write S for the point $(s, f(s))$, and similarly for x, y and t. Convexity implies X must lie below $\overline{SY}$, and Y is above the line $\overline{SX}$. Also, Y is below the line $\overline{XT}$. As $y \to x^+$, the point Y is sandwiched between these two lines, and hence $f(y) \to f(x)$. Left-hand limits are handled similarly, and continuity of f follows. $\square$

As defined, convexity requires openess of the domain; otherwise, for instance, we may have a convex function on $[a, b]$ which is not continuous, through the presence of a point discontinuity at a or b.

Given two points x and y in the domain of a real function f, we may form the *difference quotient* $\frac{f(y)-f(x)}{y-x}$, which is equivalent to the slope of the line segment from $(x, f(x))$ to $(y, f(y))$. For convex functions, the difference quotient always increases as we increase x and y:

Proposition 4.1. *If $f : (a, b) \to \mathbb{R}$ is convex and if $a < s < t < u < b$, then*

$$\frac{f(t) - f(s)}{t - s} \leq \frac{f(u) - f(s)}{u - s} \leq \frac{f(u) - f(t)}{u - t}.$$

Proof. Let $a < s < t < u < b$. Then

$$f(\lambda s + (1-\lambda)u) \leq \lambda f(s) + (1-\lambda) f(u).$$

Let $\lambda = \frac{u-t}{u-s}$, (When $t = u$, $\lambda = 0$ and when $t = s$, $\lambda = 1$.)

Then

$$\begin{aligned}
\lambda s + (1-\lambda)u &= \frac{u-t}{u-s} s + \left(1 - \frac{u-t}{u-s}\right) u \\
&= \frac{(u-t)s + (u-s)u - (u-t)u}{u-s} \\
&= \frac{us - ts + u^2 - us - u^2 + ut}{u-s} \\
&= \frac{ut - ts}{u-s} = t
\end{aligned}$$

It thus follows

$$f(t) \leq f(u) + \frac{u-t}{u-s} \cdot (f(s) - f(u)). \tag{4.1}$$

Use of the identity $\frac{u-t}{u-s} = 1 + \frac{s-t}{u-s}$ in equation 4.1 gives

$$f(t) \leq f(s) + \frac{s-t}{u-s}(f(s) - f(u)). \tag{4.2}$$

Rearrangement of equation 4.1 gives,

$$\frac{f(u) - f(s)}{u-s} \leq \frac{f(u) - f(t)}{u-t},$$

whereas rearrangement of equation 4.2 gives,

$$\frac{f(t)-f(s)}{t-s} \leq \frac{f(u)-f(s)}{u-s}.$$

Hence

$$\frac{f(t)-f(s)}{t-s} \leq \frac{f(u)-f(s)}{u-s} \leq \frac{f(u)-f(t)}{u-t}$$

as was meant to be shown. $\square$

If we take the logarithm of both sides of the convex function $f(x) = x^2$, we get $\log f = 2\log x$. Now, $\log x$ is not a convex function (for each pair of points, the line segment joining them lies below the graph, not above it); so $\log x$ is in fact concave, not convex. Thus, $\log f$ is not convex either. But if we consider $h(x) = e^x$, which is convex, we have that $\log h = x$, which is also convex. We use the terminology *log-convex* to describe h, which is a stronger property than convexity is. To show this, we need the following preliminary lemma:

Lemma 4.2. *Any increasing convex function of a convex function is convex.*

Proof. Let $f : (a,b) \to (h,k)$ be convex and let $g : (h,k) \to \mathbb{R}$ be convex and increasing, i.e., $x \leq y \Rightarrow g(x) \leq g(y)$. By the convexity of f, for any $a < x < y < b$ we have

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y)$$

for any $\lambda \in [0,1]$. Furthermore by the convexity of g we get,

$$\begin{aligned} g(f(\lambda x + (1-\lambda)y)) &\leq g(\lambda f(x) + (1-\lambda)f(y)) \\ &\leq \lambda g(f(x)) + (1-\lambda)g(f(y)). \end{aligned}$$

Hence $g \circ f$ is convex. $\square$

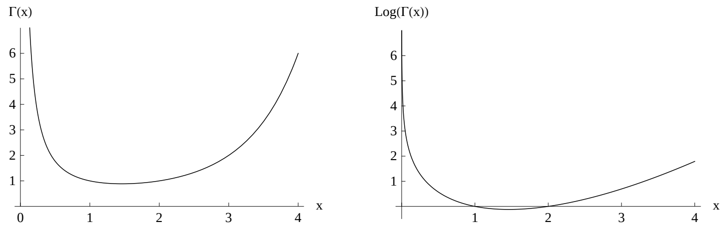


Figure 4.2: The graphs of $\Gamma(x)$ and $\log\Gamma(x)$ plotted over $(0,4)$.

Proposition 4.2. *Given a function $f : (a,b) \to \mathbb{R}$, if $\log f$ is convex, then so is f itself.*

Proof. If $\log f$ is convex, then by lemma 4.2, since e^x is increasing and convex, we have that $e^{\log f} = f$ is also convex. $\square$

We now return to the Gamma function. It is obvious from figure 4.2 that $\Gamma(x)$ with $x \in (0,\infty)$ is convex. In fact, $\Gamma(x)$ increases steeply enough as $x \to \infty$ that $\Gamma : (0,\infty) \to \mathbb{R}$ is log-convex, which we shall momentarily prove. To prove that $\log \Gamma$ is convex (which implies that Γ is convex) we need an inequality known as *Hölder's inequality*, which we prove in kind. To this end, consider the following:

Lemma 4.3. *If* $a > 0$, $b < 1$, $s > 0$, $t > 0$ *and* $s + t = 1$, *then*

$$(1-a)^s(1-b)^t \leq 1 - a^s b^t. \tag{4.3}$$

Proof. Consider $f(a) = (1-a)^s(1-b)^t + a^s b^t$ where b is fixed. Then

$$f'(a) = -s(1-a)^{s-1}(1-b)^t + sa^{s-1}b^t$$

from which it follows

$$f'(a) = 0 \;\Rightarrow\; \left(\frac{1}{a} - 1\right)^{1-s} = \left(\frac{1}{b} - 1\right)^t \;\Rightarrow\; a = b.$$

Now

$$f''(a) = s(s-1)(1-a)^{s-2}(1-b)^t + s(s-1)a^{s-2}b^t$$

from which we see

$$f''(a)\big|_{a=b} = s(s-1)\left[(1-a)^{-1} + a^{-1}\right] < 0.$$

This means f attains its maximum at $a = b$, and we have

$$f(a)\big|_{a=b} = 1 - a + a = 1.$$

Hence $f(a) \leq 1$ which gives inequality 4.3. □

Lemma 4.4. ***(Hölder's inequality).*** *Let p and q be positive real numbers such that $\frac{1}{p}+\frac{1}{q}=1$. Then for any integrable functions $f, g : [a, b] \to \mathbb{R}$, we have*

$$\left|\int_a^b f(x)g(x)\,\mathrm{d}x\right| \leq \left(\int_a^b |f|^p\,\mathrm{d}x\right)^{1/p}\left(\int_a^b |g|^q\,\mathrm{d}x\right)^{1/q}. \tag{4.4}$$

Proof. $\frac{1}{p}+\frac{1}{q}=1 \Rightarrow (1-p)(1-q)=1$, thus inequality 4.3 is satisfied. Now, using inequality 4.3 as well as the Cauchy-Schwarz inequality, we have in the limit $\epsilon \to 0^+$,

$$\left|\int_a^b f \cdot g \;\mathrm{d}x\right| \leq \left|\int_a^{b-\epsilon} f \cdot g \;\mathrm{d}x\right| + \big|f(b)g(b)\big|\epsilon$$
$$\leq \left(\int_a^{b-\epsilon} |f|^p\,\mathrm{d}x\right)^{1/p}\left(\int_a^{b-\epsilon} |g|^q\,\mathrm{d}x\right)^{1/q} + |f(b)|\cdot|g(b)|\cdot\epsilon$$

$$
\begin{aligned}
&= \left(\int_a^b |f|^p \,\mathrm{d}x - |f(b)|^p \cdot \epsilon \right)^{1/p} \left(\int_a^b |g|^q \,\mathrm{d}x - |g(b)|^q \cdot \epsilon \right)^{1/q} \\
&\quad + |f(b)| \cdot |g(b)| \cdot \epsilon \\
&= \left(\textstyle\int_a^b |f|^p \,\mathrm{d}x \right)^{1/p} \left(\textstyle\int_a^b |g|^q \,\mathrm{d}x \right)^{1/q} \left(1 - \tfrac{|f|^p \cdot \epsilon}{\int_a^b |f|^p \,\mathrm{d}x} \right)^{1/p} \left(1 - \tfrac{|g|^q \cdot \epsilon}{\int_a^b |g|^q \,\mathrm{d}x} \right)^{1/q} \\
&\quad + |f(b)| \cdot |g(b)| \cdot \epsilon \\
&\leq \left(\textstyle\int_a^b |f|^p \,\mathrm{d}x \right)^{1/p} \left(\textstyle\int_a^b |g|^q \,\mathrm{d}x \right)^{1/q} \left(1 - \tfrac{|f||g|\epsilon}{(\int_a^b |f|^p \,\mathrm{d}x)^{1/p} (\int_a^b |g|^q \,\mathrm{d}x)^{1/q}} \right) \\
&\quad + |f(b)| \cdot |g(b)| \cdot \epsilon \\
&= \left(\int_a^b |f|^p \,\mathrm{d}x \right)^{1/p} \left(\int_a^b |g|^q \,\mathrm{d}x \right)^{1/q}
\end{aligned}
$$

and the proof is complete. $\square$

Now consider that

Theorem 4.3. $\Gamma : (0, \infty) \to \mathbb{R}$ *is log-convex.*

Proof. Let $1 < p < \infty$ and $\frac{1}{p} + \frac{1}{q} = 1$. As in Rudin, consider

$$\begin{aligned}
\Gamma\left(\frac{x}{p}+\frac{y}{q}\right) &= \int_0^\infty t^{(x/p+y/q-1)}e^{-t}\,\mathrm{d}t \\
&= \int_0^\infty t^{x/p}t^{y/q}t^{-1/p}t^{-1/q}e^{-t/p}e^{-t/q}\,\mathrm{d}t \\
&= \int_0^\infty \left(t^{x-1}e^{-t}\right)^{1/p}\left(t^{y-1}e^{-t}\right)^{1/q}\mathrm{d}t \\
&\leq \left(\int_0^\infty t^{x-1}e^{-t}\,\mathrm{d}t\right)^{1/p}\left(\int_0^\infty t^{y-1}e^{-t}\,\mathrm{d}t\right)^{1/q} \\
&= \Gamma(x)^{1/p}\ \Gamma(y)^{1/q}.
\end{aligned}$$

Let $\lambda = \frac{1}{p}$, and hence $1-\lambda = \frac{1}{q}$. Then $\lambda \in (0,1)$, and

$$\begin{aligned}
\Gamma(\lambda x + (1-\lambda)y) &\leq \Gamma(x)^\lambda\,\Gamma(y)^{1-\lambda} \\
\log\Gamma(\lambda x + (1-\lambda)y) &\leq \log\left[\Gamma(x)^\lambda\,\Gamma(y)^{1-\lambda}\right] \\
&= \lambda\log\Gamma(x) + (1-\lambda)\log\Gamma(y)
\end{aligned}$$

for any $x, y \in (0,\infty)$. Hence $\log\Gamma$ is convex. □

A famous theorem of Bohr and Mullerup says that proposition 1.2, theorem 1.3 and theorem 4.3, taken together, distinguish Γ as *the* function which extends the factorials to all real $x > 0$.

Theorem 4.4. ***(Bohr-Mullerup).*** *If $f : (0,\infty) \to (0,\infty)$ satisfies*

1. $f(1) = 1$,

2. $f(x+1) = xf(x)$, *and*

3. $\log f$ *is convex,*

then $f(x) = \Gamma(x)$ *for all* $x \in (0, \infty)$*.*

Proof. Rudin's proof is very elegant. Since we have already shown Γ to satisfy conditions 1 through 3, it suffices to prove that $f(x)$ is uniquely determined by these conditions. Furthermore, condition 2 asserts it is enough to prove this only for $x \in (0, 1)$.

Set $\varphi = \log f$. Condition 1 says that $\varphi(1) = 0$ and condition 2 says

$$\varphi(x+1) = \varphi(x) + \log x. \tag{4.5}$$

Condition 3 means that φ is convex.

Let $0 < x < 1$, and let $n \in \mathbb{N}$. Consider the difference quotients of φ; let $s = n$, $t = n+1$ and $u = n+1+x$ to get

$$\varphi(n+1) - \varphi(n) \leq \frac{\varphi(n+1+x) - \varphi(n+1)}{x}.$$

The difference quotients of φ with $s = n+1$, $t = n+1+x$ and $u = n+2$ give

$$\frac{\varphi(n+1+x) - \varphi(n+1)}{x} \leq \varphi(n+2) - \varphi(n+1).$$

Note that by equation 4.5, $\varphi(n+1) - \varphi(n) = \log n$, and in addition by condition 2, $\varphi(n+2) - \varphi(n+1) = \log(n+1)$.

Combining these gives

$$\log n \leq \frac{\varphi(n+1+x) - \varphi(n+1)}{x} \leq \log(n+1).$$

Repeatedly applying equation 4.5 gives

$$\begin{aligned}
&\varphi(x+n+1) \\
&\quad = \varphi(x+n) + \log(x+n) \\
&\quad = \varphi(x+n-1) + \log(x+n) + \log(x+n-1) \\
&\quad = \varphi(x+n-1) + \log[(x+n)(x+n-1)] \\
&\quad = \varphi(x+n-2) + \log[(x+n)(x+n-1)(x+n-2)] \\
&\qquad\quad \vdots \\
&\quad = \varphi(x) + \log[(x+n)(x+n-1)\cdots(x+1)x].
\end{aligned}$$

Also by equation 4.5, we have $\varphi(n+1) = \log(n!)$. So

$$\begin{aligned}
&\frac{\varphi(n+1+x) - \varphi(n+1)}{x} \\
&\quad = \frac{1}{x}[\varphi(x) + \log[(x+n)\cdots(x+1)x] - \log(n!)]
\end{aligned}$$

giving us

$$\log n \leq \tfrac{1}{x}[\varphi(x) + \log[(x+n)\cdots(x+1)x] - \log(n!)] \leq \log(n+1).$$

Multiplying through by x yields

$$\log n^x \leq \varphi(x) + \log[(x+n)\cdots(x+1)x] - \log(n!) \leq \log(n+1)^x.$$

Subtracting $\log n^x$ from each term gives

$$0 \le \varphi(x) + \log[(x+n)\cdots(x+1)x] - \log(n!) - \log n^x \le \log(n+1)^x - \log n^x.$$

Simplifying,

$$0 \le \varphi(x) - \log\left[\frac{n!n^x}{x(x+1)\cdots(x+n)}\right] \le x\log\left(1+\frac{1}{n}\right).$$

Now let $n \to \infty$, so that $\log\left(1+\frac{1}{n}\right) \to 0$, and hence

$$\varphi(x) = \lim_{n\to\infty} \log\left[\frac{n!n^x}{x(x+1)\cdots(x+n)}\right].$$

In any case φ is uniquely determined and the proof is complete. □

The last equation in the above proof brings us to an alternative definition for the Gamma function. In another (earlier) letter written October 13, 1729 also to his friend Goldbach, Euler gave the following equivalent definition for the Gamma function:

Definition 4.2. Let $0 < x < \infty$ and define

$$\Gamma(x) = \lim_{n\to\infty}\left[\frac{n!n^x}{x(x+1)\cdots(x+n)}\right] = \lim_{n\to\infty}\left[\frac{n!n^x}{x(1+x/1)\cdots(1+x/n)}\right].$$

This approach, using an infinite product, was also chosen by Gauss, in 1811, in his study of the Gamma function. Using this formulation is often more convenient in establishing new properties of the Gamma function.

Proof. Since we are dealing with a definition, this proof is only to show consistency with the previous definitions of Γ already given. The proof of theorem 4.4 shows that

$$\varphi(x) = \lim_{n\to\infty} \log \left[\frac{n! n^x}{x(x+1)\cdots(x+n)} \right]$$

for $0 < x < 1$. Since the log function is continuous, we can exchange lim and log this way:

$$\varphi(x) = \log \lim_{n\to\infty} \left[\frac{n! n^x}{x(x+1)\cdots(x+n)} \right].$$

Exponentiating both sides gives

$$\Gamma(x) = \lim_{n\to\infty} \left[\frac{n! n^x}{x(x+1)\cdots(x+n)} \right] \tag{4.6}$$

for $0 < x < 1$. Equation 4.6 also holds for $x = 1$, in which case we have

$$\Gamma(1) = \lim_{n\to\infty} \left[\frac{n! n}{1 \cdot 2 \cdot \cdots \cdot (n+1)} \right] = 1.$$

Thus, equation 4.6 holds for $0 < x \leq 1$. Using proposition 1.2, we see that

$$\begin{aligned}
\Gamma(x+1) &= x \lim_{n\to\infty} \left[\frac{n!n^x}{x(x+1)\cdots(x+n)}\right] \\
&= \lim_{n\to\infty} \frac{x+n+1}{n} \left[\frac{n!n^{x+1}}{(x+1)\cdots(x+n)(x+n+1)}\right] \\
&= \lim_{n\to\infty} \left[1 + \tfrac{1+x}{n}\right] \cdot \lim_{n\to\infty} \left[\tfrac{n!n^{x+1}}{(x+1)\cdots(x+n)(x+n+1)}\right] \\
&= \lim_{n\to\infty} \left[\frac{n!n^{x+1}}{(x+1)\cdots(x+n)(x+n+1)}\right].
\end{aligned}$$

From this we see that equation 4.6 holds for $1 < x \leq 2$ as well; repeatedly applying this procedure shows that it applies for all $x > 0$, as required. $\square$

5

The Beta Function

The integral in the definition of the Gamma function (definition 1.1) is known as *Euler's second integral.* Now, *Euler's first integral* (1730) is another integral related to the Gamma function, which he also proposed:

Definition 5.1. Beta Function. For $\mathrm{Re}(x), \mathrm{Re}(y) > 0$, define

$$\begin{aligned} B(x,y) &:= \int_0^1 t^{x-1}(1-t)^{y-1}\,\mathrm{d}t \\ &= 2\int_0^{\pi/2} \sin(t)^{2x-1}\cos(t)^{2y-1}\,\mathrm{d}t \\ &= \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = B(y,x). \end{aligned}$$

This integral is commonly known as the Beta function. The definition above involves three equivalent identities – an integral over trigonometric functions, an integral over polynomials and a ratio of Gamma functions. The Beta function is symmetric.

To interrelate these various representations of the Beta function, consider the product of two factorials written in terms of the Gamma function:

$$m!\,n! = \int_0^\infty e^{-u} u^m \,\mathrm{d}u \int_0^\infty e^{-v} v^n \,\mathrm{d}v.$$

Let $u = x^2$ and $v = y^2$ so that

$$\begin{aligned} m!\,n! &= 4\int_0^\infty e^{-x^2} x^{2m+1} \,\mathrm{d}x \int_0^\infty e^{-y^2} y^{2n+1} \,\mathrm{d}y \\ &= \int_{-\infty}^\infty \int_{-\infty}^\infty e^{-(x^2+y^2)} |x|^{2m+1} |y|^{2n+1} \,\mathrm{d}x \,\mathrm{d}y. \end{aligned}$$

Switch to polar coordinates with $x = r\cos\theta$ and $y = r\sin\theta$,

$$\begin{aligned} m!\,n! &= \int_0^{2\pi} \int_0^\infty e^{-r^2} |r\cos\theta|^{2m+1} |r\sin\theta|^{2n+1} r \,\mathrm{d}r \,\mathrm{d}\theta \\ &= \int_0^\infty e^{-r^2} r^{2m+2n+3} \,\mathrm{d}r \int_0^{2\pi} |\cos^{2m+1}\theta \sin^{2n+1}\theta| \,\mathrm{d}\theta \\ &= 4\int_0^\infty e^{-r^2} r^{2(m+n+1)+1} \,\mathrm{d}r \int_0^{\pi/2} \cos^{2m+1}\theta \sin^{2n+1}\theta \,\mathrm{d}\theta. \end{aligned}$$

Now make the substitutions $t = r^2$ and $dt = 2r\,\mathrm{d}r$ in the first integral,

$$\begin{aligned} m!\,n! &= 2\int_0^\infty e^{-t} t^{m+n+1}\,\mathrm{d}t \int_0^{\pi/2} \cos^{2m+1}\theta \sin^{2n+1}\theta\,\mathrm{d}\theta \\ &= 2\cdot\Gamma(m+n+2)\int_0^{\pi/2} \cos^{2m+1}\theta \sin^{2n+1}\theta\,\mathrm{d}\theta \\ &= 2(m+n+1)!\int_0^{\pi/2} \cos^{2m+1}\theta \sin^{2n+1}\theta\,\mathrm{d}\theta. \end{aligned}$$

In terms of the Beta function we then have

$$\begin{aligned} B(m+1,n+1) &\equiv 2\int_0^{\pi/2} \cos^{2m+1}\theta \sin^{2n+1}\theta\,\mathrm{d}\theta \\ &= \frac{m!\,n!}{(m+n+1)!}. \end{aligned}$$

Adjusting the arguments then gives

$$\begin{aligned} B(m,n) &\equiv \frac{\Gamma(m)\Gamma(n)}{\Gamma(m+n)} \\ &= \frac{(m-1)!\,(n-1)!}{(m+n-1)!}. \end{aligned}$$

From the trigonometric form

$$B(x+1,y+1) = 2\int_0^{\pi/2} \sin(t)^{2x+1}\cos(t)^{2y+1}\,\mathrm{d}t$$

we recover the integral over polynomials by making the change of variables $u = \cos^2 t$ and $du = 2\cos t \sin t\, dt$.

Doing this,

$$B(x+1, y+1) = \int_0^1 (1-u)^x (u)^y \,\mathrm{d}u$$

so readjustment of the arguments and relabeling gives us

$$B(x, y) = \int_0^1 (1-t)^{x-1} (t)^{y-1} \,\mathrm{d}t.$$

Directly from the definition follows the *beta function functional equation*

$$B(x+1, y) = \frac{x}{x+y} B(x, y).$$

Proof.

$$B(x+1, y) = \tfrac{\Gamma(x+1)\Gamma(y)}{\Gamma(x+y+1)} = \tfrac{x\Gamma(x)\Gamma(y)}{(x+y)\Gamma(x+y)} = \tfrac{x}{x+y} B(x, y).$$

□

By substituting $x = y = \frac{1}{2}$ into the definition, we obtain the following:

Corollary. $\Gamma(\frac{1}{2}) = \sqrt{\pi}$.

Proof. Considering

$$\frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} = 2\int_0^{\pi/2} (\sin\theta)^{2x-1}(\cos\theta)^{2y-1}\,\mathrm{d}\theta$$

the special case $x = y = \frac{1}{2}$ gives

$$\left[\Gamma\left(\frac{1}{2}\right)\right]^2 = 2\int_0^{\pi/2} \mathrm{d}\theta = \pi.$$

Since Γ is positive for all $x \in (0,\infty)$, we have $\Gamma(\frac{1}{2}) = \sqrt{\pi}$ as required. □

Using this result, we can easily perform an integral very fundamental in probability theory:

Proposition 5.1.

$$\int_{-\infty}^{\infty} e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

Proof. We have $\Gamma(\frac{1}{2}) = \sqrt{\pi}$. By definition,

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} t^{-1/2}e^{-t}\,\mathrm{d}t.$$

Substituting $t = x^2$, $dt = 2x\;dx$ yields

$$\Gamma\left(\frac{1}{2}\right) = \int_0^{\infty} x^{-1}e^{-x^2}2x\,\mathrm{d}x = \sqrt{\pi}.$$

Therefore
$$\int_0^\infty e^{-x^2}\,\mathrm{d}x = \frac{\sqrt{\pi}}{2}.$$
Since e^{-x^2} is symmetric about $x = 0$, we have
$$\int_{-\infty}^\infty e^{-x^2}\,\mathrm{d}x = 2\int_0^\infty e^{-x^2}\,\mathrm{d}x = \sqrt{\pi}.$$

□

6

Wallis's Integrals

The following integrals (Wallis's integrals)

$$W_n = \int_0^{\pi/2} \sin^n \theta \, \mathrm{d}\theta = \int_0^{\pi/2} \cos^n \theta \, \mathrm{d}\theta$$

may be computed by means of the Beta and Gamma functions. By the definition of the Beta function, we have

$$W_n = \frac{1}{2} B\left(\frac{n+1}{2}, \frac{1}{2}\right)$$

which gives rise to the two cases $n = 2p + 1$ and $n = 2p$. For the odd values of the argument n:

$$W_{2p+1} = \frac{1}{2} B\left(p+1, \frac{1}{2}\right) \frac{\Gamma(p+1)\Gamma(1/2)}{2\Gamma(p+3/2)} = \frac{p!\,\Gamma(1/2)}{(2p+1)\Gamma(p+1/2)}.$$

Using the formula

$$\Gamma(n+1/2) = \frac{1\cdot 3\cdot 5\cdots(2n-1)}{2^n}\sqrt{\pi}$$

produces the result

$$W_{2p+1} = \frac{2^p p!}{1\cdot 3\cdot 5\cdots(2p+1)} = \frac{4^p p!^2}{(2p+1)!}.$$

For the even values of the argument n:

$$W_{2p} = \frac{1}{2}B\left(p+\frac{1}{2},\frac{1}{2}\right)\frac{\Gamma(p+1/2)\Gamma(1/2)}{2\Gamma(p+1)}$$

and

$$W_{2p} = \frac{1\cdot 3\cdot 5\cdot(2p-1)}{2^{p+1}p!}\pi = \frac{(2p)!}{4^p p!^2}\frac{\pi}{2}.$$

We obviously have

$$W_{n+2} = \tfrac{1}{2}B\left(\tfrac{n+2+1}{2},\tfrac{1}{2}\right) = \tfrac{(n+1)/2}{n/2+1}W_n = \left(\tfrac{n+1}{n+2}\right)W_n$$

according to the Beta function functional equation.

Note that

$$W_\alpha = \frac{1}{2}B\left(\frac{\alpha+1}{2},\frac{1}{2}\right)$$

works for any real $\alpha > -1$ and we can therefore deduce using the definition of the Beta function (respectively with $\alpha = -1/2$ and $\alpha = 1/2$) that

$$\int_0^{\pi/2} \frac{\mathrm{d}\theta}{\sqrt{\sin\theta}} = \int_0^1 \frac{2\,\mathrm{d}t}{\sqrt{1-t^4}} = \frac{\Gamma^2(1/4)}{2\sqrt{2\pi}}$$
$$\int_0^{\pi/2} \sqrt{\sin\theta}\,\mathrm{d}\theta = \int_0^1 \frac{2t^2\,\mathrm{d}t}{\sqrt{1-t^4}} = \frac{(2\pi)^{3/2}}{\Gamma^2(1/4)}.$$

The product of those two integrals allows us to derive a relation due to Euler:

$$\int_0^1 \frac{\mathrm{d}t}{\sqrt{1-t^4}} \int_0^1 \frac{t^2\,\mathrm{d}t}{\sqrt{1-t^4}} = \frac{\pi}{4}.$$

7

Wallis's Product

Let's establish an infinite product for $\pi/2$ known as "Wallis's product."

Theorem 7.1.

$$\frac{\pi}{2} = \frac{2}{1}\frac{2}{3}\frac{4}{3}\frac{4}{5}\frac{6}{5}\frac{6}{7}\cdots\frac{2k}{2k-1}\frac{2k}{2k+1}\cdots$$

By this is meant that if P_n is the product of the first n factors on the right-hand side,

$$\lim_{n\to\infty} P_n = \frac{\pi}{2}.$$

Proof. Directly from the definition of the Beta function,

$$\begin{aligned} B\left(n+\frac{1}{2},\frac{1}{2}\right): \quad & \int_0^{\pi/2} \sin^{2n} x \,\mathrm{d}x = \frac{\sqrt{\pi}\,\Gamma(n+1/2)}{2(n!)} \\ B\left(n+1,\frac{1}{2}\right): \quad & \int_0^{\pi/2} \sin^{2n+1} x \,\mathrm{d}x = \frac{\sqrt{\pi}\, n!}{2\,\Gamma(n+3/2)} \\ & \text{where } n = 0, 1, \ldots . \end{aligned} \tag{7.1}$$

Hence, the quotient of these two integrals is

$$\begin{aligned} \frac{\int_0^{\pi/2} \sin^{2n} x \,\mathrm{d}x}{\int_0^{\pi/2} \sin^{2n+1} x \,\mathrm{d}x} &= \frac{\Gamma(n+\frac{1}{2})}{n!} \frac{\Gamma(n+\frac{3}{2})}{n!} \qquad (7.2) \\ &= \frac{2n+1}{2n} \frac{2n-1}{2n} \frac{2n-1}{2n-2} \cdots \frac{3}{4} \frac{3}{2} \frac{1}{2} \frac{\pi}{2} \\ &= \frac{1}{P_{2n}} \frac{\pi}{2}. \end{aligned}$$

We shall now show that the left-hand side of equation 7.2 approaches 1 as $n \to \infty$. By equation 7.1 formed for n and for $n-1$ we have

$$\int_0^{\pi/2} \sin^{2n+1} x \,\mathrm{d}x = \frac{2n}{2n+1} \int_0^{\pi/2} \sin^{2n-1} x \,\mathrm{d}x. \tag{7.3}$$

Since $0 \leq \sin x \leq 1$ in the interval $(0, \pi/2)$, we have

$$0 < \int_0^{\pi/2} \sin^{2n+1} x \,\mathrm{d}x < \int_0^{\pi/2} \sin^{2n} x \,\mathrm{d}x < \int_0^{\pi/2} \sin^{2n-1} x \,\mathrm{d}x.$$

Dividing this inequality by the first of its integrals and allowing n to become infinite, we have by equation 7.3 that the left-hand side of equation 7.2 approaches 1.

Hence,

$$\lim_{n\to\infty} P_{2n} = \frac{\pi}{2}.$$

Also

$$\lim_{n\to\infty} P_{2n+1} = \lim_{n\to\infty} \frac{2n+2}{2n+1} P_{2n} = \frac{\pi}{2}$$

and the proof is complete. □

Corollary.

$$\lim_{n\to\infty} \frac{(n!)^2 2^{2n}}{(2n)!\sqrt{n}} = \sqrt{\pi}.$$

Proof. To prove this, multiply and divide the right-hand side of the equation

$$P_{2n} = \frac{2}{1}\frac{2}{3}\cdots\frac{2n}{2n-1}\frac{2n}{2n+1}$$

by $2{\cdot}2\cdots 2n{\cdot}2n$, thus introducing factorials in the denominator. If then factors 2 are segregated in the numerator, the result becomes apparent. □

8

Product & Reflection Formulas

Using the fact that $(1-t/n)^n$ converges to e^{-t} as $n \to \infty$, one may write

$$\begin{aligned}\Gamma(z) &= \lim_{n\to\infty} \int_0^n t^{z-1}\left(1 - \frac{t}{n}\right)^n \mathrm{d}t \\ &= \lim_{n\to\infty} \frac{1}{n^n}\int_0^n t^{z-1}(n-t)^n \, \mathrm{d}t, \qquad \operatorname{Re} z > 0.\end{aligned}$$

Integrating by parts, we have

$$\begin{aligned}\Gamma(z) &= \lim_{n\to\infty} \frac{1}{n^n}\cdot\frac{n}{z}\int_0^n t^z(n-t)^{n-1}\,\mathrm{d}t\\ &= \lim_{n\to\infty}\frac{1}{n^n}\frac{n(n-1)\cdots 1}{z(z+1)\cdots(z+n-1)}\int_0^n t^{z+n-1}\,\mathrm{d}t\\ &= \lim_{n\to\infty}\frac{n^z}{z}\left(\frac{1}{z+1}\right)\left(\frac{2}{z+2}\right)\cdots\left(\frac{n}{z+n}\right).\end{aligned}$$

Thus,

$$\begin{aligned}\frac{1}{\Gamma(z)} &= \lim_{n\to\infty} zn^{-z}(1+z)\left(1+\frac{z}{2}\right)\cdots\left(1+\frac{z}{n}\right)\\ &= \lim_{n\to\infty} zn^{-z}\prod_{k=1}^{n}\left(1+\frac{z}{k}\right).\end{aligned}$$

To evaluate the above limit, we insert convergence factors $e^{-z/k}$ to get

$$\begin{aligned}\frac{1}{\Gamma(z)} &= \lim_{n\to\infty} zn^{-z}e^{z(1+1/2+\cdots+1/n)}\prod_{k=1}^{n}\left(1+\frac{z}{k}\right)e^{-z/k}\\ &= \lim_{n\to\infty} e^{z(1+1/2+\cdots+1/n-\log n)}\left[z\prod_{k=1}^{n}\left(1+\frac{z}{k}\right)e^{-z/k}\right].\end{aligned}$$

We shall shortly prove that $1+\frac{1}{2}+\cdots+\frac{1}{n}-\log n$ approaches a positive limit γ, known as the Euler constant, so that

$$\frac{1}{\Gamma(z)} = ze^{\gamma z}\prod_{k=1}^{\infty}\left(1+\frac{z}{k}\right)e^{-z/k}. \tag{8.1}$$

which is **Weierstrass' product form of the Gamma function**.

Euler's constant, also known as the Euler-Mascheroni constant, has the numerical value

$$\gamma \sim 0.577\,215\,664\,901\,532\,860\,606\,512\ldots.$$

Using the Weierstrass identity to define an extension of the Gamma function to the left half-plane, we get

$$\frac{1}{\Gamma(z)\Gamma(-z)} = -z^2 \prod_{k=1}^{\infty}\left(1 - \frac{z^2}{k^2}\right) = -z\frac{\sin \pi z}{\pi}.$$

where we used the identity

$$\frac{\sin \pi z}{\pi z} = \prod_{k=1}^{\infty}\left(1 - \frac{z^2}{k^2}\right)$$

which is proved in Appendix A.

Thus

$$\Gamma(z)\Gamma(-z) = \frac{-\pi}{z \sin \pi z} \tag{8.2}$$

and given that $\Gamma(1-z) = -z\Gamma(-z)$,

$$\Gamma(z)\Gamma(1-z) = B(z, 1-z) = \frac{\pi}{\sin \pi z} \tag{8.3}$$

which is **Euler's reflection formula**.

Two immediate results which follow from equation 8.3 are

1. Γ is zero-free,
2. $\Gamma(1/2) = \sqrt{\pi}$. Applying $\Gamma(z+1) = z\Gamma(z)$, we also find $\Gamma(3/2) = \frac{1}{2}\sqrt{\pi}$, $\Gamma(5/2) = 3\sqrt{\pi}/4$, and so on.

Lemma 8.1. ***(Euler's constant*** γ***).*** *If* $s_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} - \log n$, *then* $\lim_{n\to\infty} s_n$ *exists.*

Proof. $t_n = 1 + \frac{1}{2} + \cdots + 1/(n-1) - \log n$ increases with n since, in the geometric sense, t_n represents the area of $n-1$ regions between an upper Reimann sum and the exact value of $\int_1^n (1/x)\,\mathrm{d}x$. We may write

$$t_n = \sum_{k=1}^{n-1} \left[\frac{1}{k} - \log\left(\frac{k+1}{k} \right) \right]$$

and

$$\lim_{n\to\infty} t_n = \sum_{k=1}^{\infty} \left[\frac{1}{k} - \log\left(1 + \frac{1}{k}. \right) \right]$$

Now since

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$$

the series above converges to a positive constant since

$$0 < \frac{1}{k} - \log\left(1 + \frac{1}{k}\right) = \frac{1}{2k^2} - \frac{1}{3k^3} + \frac{1}{4k^4} - \cdots \leq \frac{1}{2k^2}.$$

This proves the lemma, because $\lim_{n\to\infty} s_n = \lim_{n\to\infty} t_n$. $\square$

Using equation 8.3, the gamma function $\Gamma(r)$ of a rational number r can be reduced. For instance,

$$\Gamma\left(\frac{2}{3}\right) = \frac{2\pi}{\sqrt{3}\,\Gamma(\frac{1}{3})}.$$

9

Half-Integer Values

From the functional equation $\Gamma(n+1) = n\Gamma(n)$ the value of the Gamma function at half-integer values is determined by a single one of them; one has

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

from which it follows, through repeated application of the functional equation, that for $n \in \mathbb{N}$,

$$\Gamma\left(\frac{1}{2} + n\right) = \sqrt{\pi}\prod_{k=1}^{n} \frac{2k-1}{2} = \frac{(2n-1)!!}{2^n}\sqrt{\pi}$$

and

$$\Gamma\left(\frac{1}{2} - n\right) = \sqrt{\pi}\prod_{k=1}^{n} -\frac{2}{2k-1} = \frac{(-1)^n 2^n}{(2n-1)!!}\sqrt{\pi}.$$

Similarly,

$$\Gamma\left(\frac{n}{2}\right) = \frac{(n-2)!!}{2^{(n-1)/2}}\sqrt{\pi}.$$

10

Digamma and Polygamma Functions

An entire theory revolves around Digamma and Polygamma functions, which we shall not pursue in great depth here. We briefly define them, then first show their involvement in the derivation of an expression for the derivative of the Gamma function which follows from Weierstrass' product. Series expansions shall prove useful.

The Digamma function is given by the logarithmic derivative of the Gamma function:

$$\Psi(z) \equiv \frac{d}{dz} \log \Gamma(z) = \frac{\Gamma'(z)}{\Gamma(z)}.$$

The nth derivative of $\Psi(z)$ is called the polygamma function, denoted $\psi_n(z)$. The notation

$$\psi_0(z) \equiv \Psi(z)$$

is often used for the digamma function itself.

Begin with the Weierstrass form

$$\Gamma(z) = \left[z e^{\gamma z} \prod_{k=1}^{\infty} \left(1 + \frac{z}{k}\right) e^{-z/k} \right]^{-1}.$$

Take the logarithm of both sides,

$$-\log[\Gamma(z)] = \log z + \gamma z + \sum_{k=1}^{\infty} \left[\log\left(1 + \frac{z}{k}\right) - \frac{z}{k} \right].$$

Now differentiate,

$$\begin{aligned} -\frac{\Gamma'(z)}{\Gamma(z)} &= \frac{1}{z} + \gamma + \sum_{k=1}^{\infty} \left(\frac{1/k}{1 + z/k} - \frac{1}{k} \right) \\ &= \frac{1}{z} + \gamma + \sum_{k=1}^{\infty} \left(\frac{1}{k+z} - \frac{1}{k} \right) \end{aligned}$$

$$\begin{aligned} \Gamma'(z) = -\Gamma(z) &\left[\frac{1}{z} + \gamma + \sum_{k=1}^{\infty} \left(\frac{1}{k+z} - \frac{1}{k} \right) \right] \\ &\equiv \Gamma(z)\Psi(z) = \Gamma(z)\psi_0(z). \end{aligned}$$

It immediately follows that

$$\begin{aligned}\Gamma'(1) &= -\Gamma(1)\left\{1+\gamma+\left[\left(\tfrac{1}{2}-1\right)+\left(\tfrac{1}{3}-\tfrac{1}{2}\right)+\cdots+\left(\tfrac{1}{k+1}-\tfrac{1}{k}\right)+\cdots\right]\right\}\\ &= -(1+\gamma-1)\\ &= -\gamma\end{aligned}$$

and

$$\begin{aligned}\Gamma'(n) &= -\Gamma(n)\left\{\tfrac{1}{n}+\gamma\left[\left(\tfrac{1}{1+k}-1\right)+\left(\tfrac{1}{2+k}-\tfrac{1}{2}\right)+\left(\tfrac{1}{3+k}-\tfrac{1}{3}\right)+\cdots\right]\right\}\\ &= -\Gamma(n)\left(\frac{1}{n}+\gamma-\sum_{k=1}^{n}\frac{1}{k}\right).\end{aligned}$$

Given the expression for $\Gamma'(z)$, the minimum value x_m of $\Gamma(x)$ for real positive $x = x_m$ occurs when

$$\Gamma'(x_m) = \Gamma(x_m)\psi_0(x_m) = 0.$$

In other words, when

$$\psi_0(x_m) = 0.$$

Numerically, this is solved to give $x_m = 1.46163\ldots.$

We have that

$$\Psi(z) = \frac{\Gamma'(z)}{\Gamma(z)} = -\gamma - \frac{1}{z} + \sum_{k=1}^{\infty}\left(\frac{1}{k} - \frac{1}{z+k}\right)$$

$$= -\gamma + \sum_{k=1}^{\infty}\left(\frac{1}{k} - \frac{1}{z+k-1}\right) \quad z \neq 0, -1, -2, \ldots \tag{10.1}$$

$$= -\gamma + \sum_{k=1}^{\infty}\left(\frac{z-1}{k(z+k-1)}\right) \quad z \neq 0, -1, -2, \ldots$$

If we differentiate relation 10.1 many times, we find

$$\Psi'(z) = \frac{\Gamma(z)\Gamma''(z) - \Gamma'^2(z)}{\Gamma^2(z)} = \sum_{k=1}^{\infty}\frac{1}{(k+z-1)^2} \tag{10.2}$$

$$\Psi''(z) = -\sum_{k=1}^{\infty}\frac{2}{(k+z-1)^3}$$

$$\Psi^{(n)}(z) = \sum_{k=1}^{\infty}\frac{(-1)^{n+1}n!}{(k+z-1)^{(n+1)}} \tag{10.3}$$

where the $\Psi_n = \Psi^{(n)}$ functions are the polygamma functions,

$$\Psi_n(z) = \frac{d^{n+1}}{dz^{n+1}}(\log(\Gamma(z)))$$

$$\Psi_0(z) = \Psi(z).$$

The series expansion relation 10.1 suggests

$$\Psi(z+1) - \Psi(z) = \sum_{k=1}^{\infty} \left(\frac{1}{z+k-1} - \frac{1}{z+k} \right)$$

which gives the recurrence formula

$$\begin{aligned} \Psi(z+1) &= \Psi(z) + \frac{1}{z} \\ \Psi(z+n) &= \Psi(z) + \frac{1}{z} + \frac{1}{z+1} + \cdots + \frac{1}{z+n-1} \quad n \geq 1 \end{aligned}$$

and by differentiating the first of these relations,

$$\Psi_n(z+1) = \Psi_n(z) + \frac{(-1)^n n!}{z^{n+1}}. \tag{10.4}$$

Now, the Riemann zeta function, $\zeta(s)$, is a function of a complex variable s that analytically continues the sum of the infinite series

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

The Riemann zeta function is defined as the analytic continuation of the function defined for $\mathrm{Re}(s) > 1$ by the sum of the given series.

The values of the zeta function obtained from integral arguments are called *zeta constants*. The following are some common values of the Riemann zeta function:

$$\zeta(0) = -\frac{1}{2}$$
$$\zeta(1) = 1 + \frac{1}{2} + \frac{1}{3} + \cdots = \infty$$
$$\zeta(2) = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \cdots = \frac{\pi^2}{6}$$
$$\zeta(3) = 1 + \frac{1}{2^3} + \frac{1}{3^3} + \cdots \approx 1.202\ldots$$
$$\zeta(4) = 1 + \frac{1}{2^4} + \frac{1}{3^4} + \cdots = \frac{\pi^4}{90} \approx 1.0823\ldots$$

Let's consider the Basel problem in greater detail. The Basel problem is a famous problem in number theory, first posed by Pietro Mengoli in 1644, and solved by Euler in 1735. Seeing that the problem had withstood the attacks of the leading mathematicians of the day, Euler's accomplishment brought him immediate fame when he was twenty-eight. Over time, Euler generalized the problem considerably, and his ideas were taken up later by Riemann when he defined his zeta function and proved its fundamental properties. The problem is named after Basel, hometown of Euler as well as of the Bernoulli family, who unsuccessfully attacked the problem.

The Basel problem asks for the precise summation of the reciprocals of the squares of the natural numbers, i.e., a closed form solution to (with proof) of the infinite series:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \lim_{n \to +\infty} \left(\frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} \right).$$

Euler's arguments were based on manipulations that were not justified at the time, and it was not until 1741 that he was able to produce a truly rigorous proof.

Euler's "derivation" of the value $\frac{\pi^2}{6}$ is clever and original. He extended observations about finite polynomials and assumed that these properties hold true for infinite series. By simply obtaining the correct value, he was able to verify it numerically against partial sums of the series. The agreement he observed gave him sufficient confidence to announce his result to the mathematical community.

To follow Euler's argument, consider the Taylor series expansion of the sine function

$$\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots .$$

Dividing through by x, we have

$$\frac{\sin(x)}{x} = 1 - \frac{x^2}{3!} + \frac{x^4}{5!} - \frac{x^6}{7!} + \cdots .$$

Now, the zeros of $\sin(x)/x$ occur at $x = n \cdot \pi$ where $n = \pm 1, \pm 2, \pm 3, \ldots$. Assume we can express this infinite series as a product of linear factors given by its zeros, just as is commonly done for finite polynomials:

$$\begin{aligned}\frac{\sin(x)}{x} &= \left(1-\tfrac{x}{\pi}\right)\left(1+\tfrac{x}{\pi}\right)\left(1-\tfrac{x}{2\pi}\right)\left(1+\tfrac{x}{2\pi}\right)\left(1-\tfrac{x}{3\pi}\right)\left(1+\tfrac{x}{3\pi}\right)\cdots \\ &= \left(1-\frac{x^2}{\pi^2}\right)\left(1-\frac{x^2}{4\pi^2}\right)\left(1-\frac{x^2}{9\pi^2}\right)\cdots.\end{aligned}$$

If we formally multiply out this product and collect all the x^2 terms, we see that the x^2 coefficient of $\sin(x)/x$ is

$$-\left(\frac{1}{\pi^2}+\frac{1}{4\pi^2}+\frac{1}{9\pi^2}+\cdots\right) = -\frac{1}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}.$$

But from the original infinite series expansion of $\sin(x)/x$, the coefficient of x^2 is $-1/(3!) = -1/6$. These two coefficients must be equal. Thus,

$$-\frac{1}{6} = -\frac{1}{\pi^2}\sum_{n=1}^{\infty}\frac{1}{n^2}.$$

Multiplying through both sides of this equation by $-\pi^2$ gives the sum of the reciprocals of the positive square integers,

$$\sum_{n=1}^{\infty}\frac{1}{n^2} = \frac{\pi^2}{6}.$$

Now, returning to the Polygamma function, from the relations 10.1 and 10.3 we have

$$\begin{aligned} \Psi(1) &= -\gamma \\ \Psi_1(1) &= \zeta(2) = \pi^2/6 \\ \Psi_2(1) &= -2\,\zeta(3) \\ \Psi_n(1) &= (-1)^{n+1} n!\,\zeta(n+1) \end{aligned} \tag{10.5}$$

Using the recurrence relation 10.4 allows us to compute these values for any positive integer,

$$\begin{aligned} \Psi(n) = \frac{\Gamma'(n)}{\Gamma(n)} &= -\gamma + \sum_{k=1}^{n-1} \frac{1}{k} \\ &= -\gamma + H_{n-1}. \end{aligned} \tag{10.6}$$

The following series expansions are obvious consequences of relations 10.5 and of the series

$$\frac{1}{1+x} - 1 = -\sum_{k=2}^{\infty} (-1)^k x^{k-1}.$$

Theorem 10.1. ***(Digamma series).***

$$\Psi(1+x) = -\gamma + \sum_{k=2}^{\infty} (-1)^k \zeta(k) x^{k-1} \quad |x| < 1, \tag{10.7}$$

$$\Psi(1+x) = -\frac{1}{1+x} - (\gamma - 1) + \sum_{k=2}^{\infty} (-1)^k (\zeta(k) - 1) x^{k-1} \quad |x| < 1. \tag{10.8}$$

11

Series Expansions

Finding series expansions for the Gamma function is now a direct consequence of the series expansions for the Digamma function:

Theorem 11.1. *For* $|x| < 1$,

$$\log(\Gamma(1+x)) = -\gamma x + \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k} x^k \tag{11.1}$$

$$\log(\Gamma(1+x)) = -\log(1+x) - (\gamma - 1)x + \sum_{k=2}^{\infty} \frac{(-1)^k (\zeta(k)-1)}{k} x^k \tag{11.2}$$

Proof. Use term by term integration of the Taylor series 10.7 and 10.8. $\square$

It follows easily from equation 11.1 and the functional equation that

$$\frac{1}{\Gamma(z)} = z \exp\left(\gamma z - \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k) z^k}{k}\right). \tag{11.3}$$

Series formulas involving $\zeta(k)$ can be derived from equation 11.1. For example, setting $x = 1$ gives

$$\log(\Gamma(2)) = -\gamma + \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k},$$

thus we arrive at a result due to Euler,

$$\gamma = \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k}.$$

Setting $x = 1/2$ into equation 11.1 yields

$$\log\left(\Gamma\left(\frac{3}{2}\right)\right) = \log(\sqrt{\pi}/2) = -\frac{\gamma}{2} + \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k)}{k} \frac{1}{2^k},$$

so we have

$$\gamma = \log\left(\frac{4}{\pi}\right) + 2\sum_{k=2}^{\infty} (-1)^k \frac{\zeta(k)}{2^k k}.$$

12

Euler-Mascheroni Integrals

Using the integral representation of $\Gamma'(x)$ gives the interesting integral formula for Euler's constant

$$\Gamma'(1) = \int_0^\infty e^{-t} \log(t)\,\mathrm{d}t = -\gamma$$

and from

$$\Psi'(z) = \frac{\Gamma(z)\Gamma''(z) - \Gamma'^2(z)}{\Gamma^2(z)} = \sum_{k=1}^\infty \frac{1}{(k+z-1)^2}$$

comes the relation

$$\Psi'(1)\Gamma^2(1) + \Gamma'^2(1) = \Gamma(1)\Gamma''(1)$$

hence

$$\Gamma''(1) = \int_0^\infty e^{-t} \log^2(t)\,\mathrm{d}t = \gamma^2 + \frac{\pi^2}{6}.$$

We may continue this, computing the Euler-Mascheroni integrals –

$$
\begin{aligned}
\Gamma^{(3)}(1) &= -\gamma^3 - \frac{1}{2}\pi^2\gamma - 2\zeta(3) \\
\Gamma^{(4)}(1) &= \gamma^4 + \pi^2\gamma^2 + 8\zeta(3)\gamma + \frac{3}{20}\pi^4 \\
\Gamma^{(5)}(1) &= -\gamma^5 - \frac{5}{3}\pi^2\gamma^3 - 20\zeta(3)\gamma^2 - \frac{3}{4}\pi^4\gamma - 24\zeta(5) - \frac{10}{3}\zeta(3)\pi^2 \\
&\vdots
\end{aligned}
$$

13

Duplication & Multiplication Formulas

Theorem 13.1. ***(Gauss Multiplication Formula).***

$$\Gamma(z)\Gamma\left(z+\frac{1}{n}\right)\cdots\Gamma\left(z+\frac{n-1}{n}\right) = (2\pi)^{(n-1)/2} n^{1/2-nz}\Gamma(nz). \tag{13.1}$$

A brilliant proof of the multiplication formula was produced by Liouville in 1855. We here present a modernized version of that proof.

Proof. The product of Gamma functions on the left-hand side of equation 13.1 can be written

$$\int_0^\infty e^{-t_1} t_1^{z-1}\,\mathrm{d}t_1 \int_0^\infty e^{-t_2} t_2^{z+(1/n)-1}\,\mathrm{d}t_2 \cdots \int_0^\infty e^{-t_n} t_n^{z+((n-1)/n)-1}\,\mathrm{d}t_n.$$

Rearranging,

$$\int_0^\infty \int_0^\infty \cdots \int_0^\infty e^{-(t_1+t_2+\cdots+t_n)} t_1^{z-1} t_2^{z+(1/n)-1} \cdots t_n^{z+((n-1)/n)-1} \, \mathrm{d}t_1 \cdots \mathrm{d}t_n.$$

Next, introduce a change of variables:

$$t_1 = \frac{q^n}{t_2 \cdots t_n}, t_2 = t_2, \ldots, t_n = t_n.$$

The Jacobian is thus

$$\frac{nq^{n-1}}{t_2 t_3 \cdots t_n}$$

so the integral can be written

$$\int_0^\infty \cdots \int_0^\infty \exp\left[-\left(t_2 + t_3 + \cdots + t_n + \frac{q^n}{t_2 t_3 \cdots t_n} \right) \right] \times \left(\frac{q^n}{t_2 \cdots t_n} \right)^{z-1} t_2^{z+(1/n)-1} \cdots t_n^{z+((n-1)/n)-1} \frac{nq^{n-1}}{t_2 t_3 \cdots t_n} \, \mathrm{d}q \, \mathrm{d}t_2 \cdots \mathrm{d}t_n.$$

Set $s = t_2 + t_3 + \cdots + t_n + q^n/(t_2 t_3 \cdots t_n)$, so we have

$$n \int_0^\infty \int_0^\infty \cdots \int_0^\infty e^{-s} q^{nz-1} t_2^{(1/n)-1} t_3^{(2/n)-1} \cdots t_n^{((n-1)/n)-1} \, \mathrm{d}q \, \mathrm{d}t_2 \cdots \mathrm{d}t_n. \tag{13.2}$$

Now evaluate

$$\mathcal{I} = \int_0^\infty \cdots \int_0^\infty e^{-s} \prod_{k=1}^{n-1} t_{k+1}^{(k/n)-1} \, \mathrm{d}t_2 \, \mathrm{d}t_3 \cdots \mathrm{d}t_n.$$

Obviously,

$$\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}q} = -nq^{n-1}\int_0^\infty \cdots \int_0^\infty e^{-s}\prod_{k=1}^{n-1} t_{k+1}^{(k/n)-1}\frac{\mathrm{d}t_2\cdots\,\mathrm{d}t_n}{t_2\cdots t_n}.$$

Now introduce a second change of variables,

$$t_2 = q^n/(t_1 t_3\cdots t_n), t_3 = t_3, \ldots, t_n = t_n$$

and

$$s_1 = t_3 + t_4 + \cdots + t_n + t_1 + q^n/(t_3\cdots t_n t_1).$$

Now the Jacobian is

$$J = \frac{-q^n}{t_1^2 t_3\cdots t_{n-1}}$$

and we have

$$\begin{aligned}\frac{\mathrm{d}\mathcal{I}}{\mathrm{d}q} &= nq^{n-1}\int_0^\infty \cdots \int_0^\infty e^{-s_1}|J|\left(\tfrac{q^n}{t_1 t_3\cdots t_n}\right)^{(1/n)-1}\cdot\prod_{k=2}^{n-1} t_{k+1}^{(k/n)-1}\tfrac{\mathrm{d}t_1\,\mathrm{d}t_3\cdot\,\mathrm{d}t_n}{q^n/t_1}\\ &= -n\int_0^\infty \cdots \int_0^\infty e^{-s_1}\prod_{k=2}^{n-1} t_{k+1}^{(k/n)-1} t_1^{((n-1)/n)-1}\,\mathrm{d}t_3\cdots\,\mathrm{d}t_n\,\mathrm{d}t_1\\ &= -n\mathcal{I}.\end{aligned}$$

Thus,

$$\mathcal{I} = Ce^{-nq}.$$

To find C, we set $q = 0$ in the $\mathcal{I}$ integral and in the equation above, then set them equal to each other to get

$$\Gamma\left(\frac{1}{n}\right)\Gamma\left(\frac{2}{n}\right)\cdots\Gamma\left(\frac{n-1}{n}\right) = C.$$

Next we factor the integer n in the following way:

$$\begin{aligned} n &= \prod_{a=1}^{n-1}(1-\exp(2a\pi i/n)) \\ &= 2^{n-1}\sin\frac{\pi}{n}\sin\frac{2\pi}{n}\cdots\sin\frac{(n-1)\pi}{n}. \end{aligned} \tag{13.3}$$

Here we used the identities

$$\begin{aligned} 1-\exp\left(\frac{a}{n}2\pi i\right) &= 1-\left[\cos\frac{2a\pi}{n}+i\sin\frac{2a\pi}{n}\right] \\ &= 1-\cos^2\tfrac{a\pi}{n}+\sin^2\tfrac{a\pi}{n}-2i\sin\tfrac{a\pi}{n}\cos\tfrac{a\pi}{n} \\ &= 2\sin\frac{a\pi}{n}\left[\sin\frac{a\pi}{n}-i\cos\frac{a\pi}{n}\right] \end{aligned}$$

and we also have

$$\begin{aligned} \prod_{a=1}^{n-1}\left[\sin\frac{a\pi}{n}-i\cos\frac{a\pi}{n}\right] &= (-i)^{n-1}\exp\left(\frac{i\pi}{n}\sum_{a=1}^{n-1}a\right) \\ &= (-1)^{n-1}(i)^{n-1}(-1)^{\frac{1}{n}\sum_{a=1}^{n-1}a} \\ &= (-1)^{n-1}(i)^{n-1}(-1)^{\frac{n-1}{2}} \\ &= i^{4(n-1)} = 1 \end{aligned}$$

so that equation 13.3 follows.

Now by the reflection formula

$$r(b) = \Gamma\left(\frac{b}{n}\right)\left(1-\frac{b}{n}\right) = \frac{\pi}{\sin\frac{b\pi}{n}}$$

so by letting b in $r(b)$ run from 1 to $n-1$, we essentially perform the product given by C twice, so that $C = (2\pi)^{(n-1)/2} n^{-1/2}$ and $\mathcal{I} = (2\pi)^{(n-1)/2} n^{-1/2} e^{-nq}$.

Substitution in equation 13.2 gives

$$\begin{aligned}
&\Gamma(z)\Gamma(z+1/n)\cdots\Gamma(z+(n-1)/n) \\
&= n^{1/2}(2\pi)^{(n-1)/2}\int_0^\infty e^{-nq} q^{nz-1}\,\mathrm{d}q \\
&= n^{1/2-nz}(2\pi)^{\frac{n-1}{2}}\Gamma(nz)
\end{aligned}$$

which completes the proof. □

Legendre's duplication formula is found simply by setting $n = 2$ in the multiplication formula.

14

The Gamma and Zeta Function Relationship

We have already shown in a previous section that

$$\frac{1}{\Gamma(z)} = z \exp\left(\gamma z - \sum_{k=2}^{\infty} \frac{(-1)^k \zeta(k) z^k}{k}\right).$$

Now we will derive two more relationships between the Gamma function and the zeta function.

Theorem 14.1. ***(Gamma and zeta functions relation).***

$$\zeta(z)\Gamma(z) = \int_0^{\infty} \frac{t^{z-1}}{e^t - 1}\, \mathrm{d}t \qquad \textit{for } \operatorname{Re}(z) > 1. \tag{14.1}$$

Proof. The integral definition of the Gamma function

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

together with the change of variables $t = ku$ (with k a positive integer) gives

$$\Gamma(z) = \int_0^\infty (ku)^{z-1} e^{-ku} k \, \mathrm{d}u = k^z \int_0^\infty u^{z-1} e^{-ku} \, \mathrm{d}u.$$

Rewrite this in the form

$$\frac{1}{k^z} = \frac{1}{\Gamma(z)} \int_0^\infty u^{z-1} e^{-ku} \, \mathrm{d}u,$$

hence by summation

$$\begin{aligned} \sum_{k=1}^\infty \frac{1}{k^z} &= \frac{1}{\Gamma(z)} \int_0^\infty u^{z-1} \sum_{k=1}^\infty (e^{-ku}) \, \mathrm{d}u \\ &= \frac{1}{\Gamma(z)} \int_0^\infty u^{z-1} \left(\frac{1}{1 - e^{-u}} - 1 \right) \mathrm{d}u. \end{aligned}$$

We thus obtain

$$\zeta(z)\Gamma(z) = \int_0^\infty \frac{t^{z-1}}{e^t - 1} \, \mathrm{d}t.$$

$\square$

For $z = 2$, this becomes

$$\frac{\pi^2}{6} = \int_0^\infty \frac{t}{e^t - 1}\,\mathrm{d}t.$$

There is another important functional equation between the two functions, the *Riemann zeta function functional equation*:

Theorem 14.2. ***(Functional Equation).*** *The function $\zeta(s)$ is regular for all values of s except $s = 1$, where there is a simple pole with residue 1. It satisfies the functional equation*

$$\zeta(s) = 2^s \pi^{s-1} \sin\left(\frac{s\pi}{2}\right)\Gamma(1-s)\zeta(1-s). \tag{14.2}$$

Proof. This can be proved in many different ways. Here we give a proof depending on the following summation formula.

Let $\phi(x)$ be any function with a continuous derivative in the interval $[a, b]$. Then, if $[x]$ denotes the greatest integer not exceeding x,

$$\sum_{a<n\leq b} \phi(n) = \int_a^b \phi(x)\,\mathrm{d}x + \int_a^b (x - [x] - 1/2)\phi'(x)\,\mathrm{d}x$$
$$+ (a - [a] - 1/2)\phi(a) - (b - [b] - 1/2)\phi(b). \tag{14.3}$$

Since the formula is additive with respect to the interval $(a, b]$, we may suppose $n \leq a < b \leq n+1$ so that we have

$$\int_a^b (x-n-1/2)\phi'(x)\,\mathrm{d}x = (b-n-1/2)\phi(b) - (a-n-1/2)\phi(a) - \int_a^b \phi(x)\,\mathrm{d}x.$$

The right hand side of equation 14.3 thus reduces to $([b]-n)\phi(b)$. This vanishes unless $b = n+1$, in which case it is $\phi(n+1)$.

Let $\phi(n) = n^{-s}$, where $s \neq 1$, and let a and b be positive integers. Then

$$\sum_{n=a+1}^{b} \frac{1}{n^s} = \frac{b^{1-s} - a^{1-s}}{1-s} - s\int_a^b \frac{x - [x] - \frac{1}{2}}{x^{s+1}}\,\mathrm{d}x + \frac{1}{2}(b^{-s} - a^{-s}). \tag{14.4}$$

Take the half-plane $\sigma > 1$, $a = 1$, and make $b \to \infty$. Adding 1 to each side, we obtain

$$\zeta(s) = s\int_1^\infty \frac{[x] - x + \frac{1}{2}}{x^{s+1}}\,\mathrm{d}x + \frac{1}{s-1} + \frac{1}{2}. \tag{14.5}$$

Since $[x] - x + 1/2$ is bounded, this integral is convergent for $\sigma > 0$, and uniformly convergent in any finite region to the right of $\sigma = 0$. It therefore defines an analytic function of s, regular for $\sigma > 0$. The right-hand side therefore provides the analytic continuation of $\zeta(s)$ up to $\sigma = 0$, and there is clearly a simple pole at $s = 1$ with residue 1.

For $0 < \sigma < 1$ we have

$$\int_0^1 \frac{[x]-x}{x^{s+1}}\,\mathrm{d}x = -\int_0^1 x^{-s}\,\mathrm{d}x = \frac{1}{s-1}$$

$$\frac{s}{2}\int_1^\infty \frac{\mathrm{d}x}{x^{s+1}} = \frac{1}{2}$$

and equation 14.5 may be written

$$\zeta(s) = s\int_0^\infty \frac{[x]-x}{x^{s+1}}\,\mathrm{d}x \qquad (0 < \sigma < 1). \tag{14.6}$$

Actually equation 14.5 gives the analytic continuation of $\zeta(s)$ for $\sigma > 1$, for if we have

$$f(x) = [x] - x + \frac{1}{2}, \qquad f_1(x) = \int_1^x f(y)\,\mathrm{d}y,$$

then $f_1(x)$ is also bounded, since, as is easily seen,

$$\int_k^{k+1} f(y)\,\mathrm{d}y = 0$$

for any integer k. Hence

$$\int_{x_1}^{x_2} \frac{f(x)}{x^{s+1}}\,\mathrm{d}x = \left[\frac{f_1(x)}{x^{s+1}}\right]_{x_1}^{x_2} + (s+1)\int_{x_1}^{x_2} \frac{f_1(x)}{x^{s+2}}\,\mathrm{d}x$$

which tends to 0 as $x_1 \to \infty$, $x_2 \to \infty$, if $\sigma > -1$. Hence the integral in equation 14.5 is convergent for $\sigma > -1$.

Also it is easily verified that

$$s\int_0^1 \frac{[x]-x+\frac{1}{2}}{x^{s+1}}\,\mathrm{d}x = \frac{1}{s-1}+\frac{1}{2} \qquad (\sigma<0).$$

Hence

$$\zeta(s) = s\int_0^\infty \frac{[x]-x+\frac{1}{2}}{x^{s+1}}\,\mathrm{d}x \qquad (-1<\sigma<0). \quad (14.7)$$

Now we have the Fourier series

$$[x]-x+\frac{1}{2} = \sum_{n=1}^{\infty}\frac{\sin 2n\pi x}{n\pi} \qquad (14.8)$$

where x is not an integer. Substituting in equation 14.7, and integrating term by term, we obtain

$$\begin{aligned}\zeta(s) &= \frac{s}{\pi}\sum_{n=1}^{\infty}\frac{1}{n}\int_0^\infty \frac{\sin 2n\pi x}{x^{s+1}}\,\mathrm{d}x \\ &= \frac{s}{\pi}\sum_{n=1}^{\infty}\frac{(2n\pi)^s}{n}\int_0^\infty \frac{\sin y}{y^{s+1}}\,\mathrm{d}y \\ &= \frac{s}{\pi}(2\pi)^s\{-\Gamma(-s)\}\sin\left(\frac{s\pi}{2}\right)\zeta(1-s)\end{aligned}$$

i.e., equation 14.2. This is valid primarily for $-1<\sigma<0$. Here, however, the right-hand side is analytic for all values of s such that $\sigma<0$. It therefore provides the analytic continuation of $\zeta(s)$ over the remainder of the plane, and there are no singularities other than the pole already encountered at $s = 1$.

We have still to justify the term-by-term integration. Since the series 14.8 is boundedly convergent, term-by-term integration over any finite range is permissible. It is therefore sufficient to prove that

$$\lim_{\lambda \to \infty} \sum_{n=1}^{\infty} \frac{1}{n} \int_{\lambda}^{\infty} \frac{\sin 2n\pi x}{x^{s+1}}\, \mathrm{d}x = 0 \qquad (-1 < \sigma < 0).$$

Now

$$\int_{\lambda}^{\infty} \frac{\sin 2n\pi x}{x^{s+1}}\, \mathrm{d}x = \left[-\tfrac{\cos 2n\pi x}{2n\pi x^{s+1}}\right]_{\lambda}^{\infty} - \tfrac{s+1}{2n\pi} \int_{\lambda}^{\infty} \tfrac{\cos 2n\pi x}{x^{s+2}}\, \mathrm{d}x$$

$$= \mathcal{O}\left(\tfrac{1}{n\lambda^{\sigma+1}}\right) + \mathcal{O}\left(\tfrac{1}{n} \int_{\lambda}^{\infty} \tfrac{\mathrm{d}x}{x^{\sigma+2}}\right) = \mathcal{O}\left(\tfrac{1}{n\lambda^{\sigma+1}}\right)$$

and the desired result clearly follows. □

Corollary. ***(Functional Equation).*** *Let*

$$\Lambda(s) = \pi^{-s/2}\, \Gamma\left(\frac{s}{2}\right) \zeta(s)$$

be an analytic function except at poles 0 and 1, then

$$\Lambda(s) = \Lambda(1-s).$$

Proof. Changing s into $1 - s$, the functional equation is

$$\zeta(1-s) = 2^{1-s} \pi^{-s} \cos\left(\frac{s\pi}{2}\right) \Gamma(s) \zeta(s). \tag{14.9}$$

It may also be written

$$\zeta(s) = \chi(s)\zeta(1-s) \tag{14.10}$$

where

$$\chi(s) = 2^s \pi^{s-1} \sin\left(\frac{s\pi}{2}\right)\Gamma(1-s) = \pi^{s-1/2}\frac{\Gamma(1/2 - s/2)}{\Gamma(s/2)}$$

and

$$\chi(s)\chi(1-s) = 1.$$

The corollary is at once verified from 14.9 and 14.10. $\square$

15

Stirling's Formula

In this section we study how the Gamma function behaves when the argument x becomes large. If we restrict the argument x to integral values n, the following result, due to James Stirling and Abraham de Moivre is quite famous and extremely important:

Theorem 15.1. ***(Stirling's formula).*** *If the integer n tends to infinity, we have the asymptotic formula*

$$\Gamma(n+1) = n! \sim \sqrt{2\pi n}n^n e^{-n}. \tag{15.1}$$

For greater clarity in the proof of equation 15.1, we introduce several simple lemmas.

Lemma 15.1.

$$\int_n^{n+1} \frac{\mathrm{d}x}{x} = \log\left(1 + \frac{1}{n}\right) > \frac{2}{2n+1} \qquad n = 1, 2, \ldots.$$

Proof. Since the curve $y = 1/x$ is convex, the area under the curve from $x = n$ to $x = n + 1$ is greater than the area of the trapezoid bounded by these two ordinates, the x-axis, and the lower tangent to the curve at the point $(n+1/2, 2/(2n+1))$, i.e., the lower Riemann sum. Note that the area of a trapezoid is equal the product of the length of the base by the length of the median, which, in this case, is

$$\frac{1}{n+\frac{1}{2}} = \frac{2}{2n+1}.$$

□

Lemma 15.2.

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{n!}{(n/e)^n\sqrt{n}} \quad \textit{exists.}$$

Proof. We have

$$\frac{a_n}{a_{n+1}} = \frac{\left(1+\frac{1}{n}\right)^{n+1/2}}{e} > 1$$

since by lemma 15.1,

$$\left(n+\frac{1}{2}\right)\log\left(1+\frac{1}{n}\right) > 1.$$

Since $a_n > 0$ for all n, the proof is complete. □

Lemma 15.3.

$$\lim_{n\to\infty} a_n > 0.$$

Proof. Use an argument similar to that of lemma 15.2, but this time use an upper Riemann sum and compare areas with the function $y = \log x$. We break the interval up by bounding a series of trapezoids by two rectangles, one at each end of the interval. The altitudes of the two rectangles at the endpoints are 2 and $\log n$. Thus, the area of the trapezoids and the two rectangles is

$$\begin{aligned} 1 + \log 2 + \log 3 + \cdots + \log(n-1) + \frac{1}{2}\log n \\ = 1 + \log n! - \log\sqrt{n} \end{aligned}$$

The area under the curve is

$$\int_1^n \log x \,\mathrm{d}x = n\log n - n + 1 = \log(n/e)^n + 1.$$

Hence,

$$\log\left(\frac{n}{e}\right)^n < \log\frac{n!}{\sqrt{n}}$$

$$\frac{(n/e)^n\sqrt{n}}{n!} < 1 \qquad n = 1, 2, \ldots.$$

Consequently,

$$a_n > 1, \quad \lim_{n\to\infty} a_n \geq 1.$$

$\square$

We have proved more than stated. It is only the nonvanishing of the limit which is needed.

Proof. Let's restate Stirling's formula this way:

$$\lim_{n\to\infty} \frac{(n/e)^n\sqrt{2\pi n}}{n!} = 1.$$

To prove the result, we need only to show that

$$r = \lim_{n\to\infty} a_n = \sqrt{2\pi}.$$

We use the fact that

$$\lim_{n\to\infty} \frac{(n!)^2 2^{2n}}{(2n)!\sqrt{n}} = \sqrt{\pi}$$

which was derived in an earlier section, to evaluate this limit r. The function of n appearing here can be rewritten in terms of a_n as follows:

$$\frac{(n!)^2 2^{2n}}{(2n)!\sqrt{n}} = \frac{a_n^2}{a_{2n}} \frac{1}{\sqrt{2}}.$$

As n becomes infinite, this quotient approaches $\sqrt{\pi}$ on the one hand and $r^2/(r\sqrt{2})$ on the other. Hence, $r = \sqrt{2\pi}$, and the proof is complete. □

Corollary.

$$\lim_{n\to\infty} \frac{(2n)!e^{2n}}{(2n)^{2n}} = +\infty.$$

Proof.

$$\lim_{n\to\infty}\frac{(2n)!e^{2n}}{(2n)^{2n}}=\lim_{n\to\infty}\left(\frac{(2n)!}{(2n)^{2n}e^{-2n}\sqrt{4\pi n}}\right)\sqrt{4\pi n}=+\infty$$

□

Corollary.

$$\frac{(n+p)!}{n!}\sim n^p \qquad n\to\infty;\ p=1,2,\ldots.$$

Proof. By Stirling's formula,

$$\lim_{n\to\infty}\frac{(n+p)!}{n!n^p}=\lim_{n\to\infty}\frac{(n+p)!}{(n+p)^{n+p}e^{-n-p}\sqrt{2\pi(n+p)}}\cdot \frac{n^n e^{-n}\sqrt{2\pi n}}{n!}\,\frac{(1+(p/n))^{n+p+1/2}}{e^p}=1.$$

Each of the three quotients on the right approaches 1. □

Corollary.

$$\lim_{n\to\infty}\frac{1}{n}\sqrt[n]{n!}=\frac{1}{e}.$$

Proof. By Stirling's formula,

$$\lim_{n\to\infty}\frac{\sqrt[n]{n!}}{n}=\lim_{n\to\infty}\left(\frac{n!}{(n/e)^n\sqrt{2\pi n}}\right)^{1/n}\frac{(2\pi n)^{1/(2n)}}{e}=\frac{1}{e}.$$

□

16

Residues of the Gamma Function

If a complex function is analytic on a region R, it is infinitely differentiable in R.[†] A complex function may fail to be analytic at one or more points through the presence of singularities, or through the presence of branch cuts. A single-valued function that is analytic in all but a discrete subset of its domain, and at those singularities goes to infinity like a polynomial (i.e., these exceptional points are poles), is called a meromorphic function. In this context, the word "pole" is used to denote a singularity of a complex function. f has a pole of order n at z_0 if n is the smallest positive integer for which $(z-z_0)^n f(z)$ is analytic at z_0.

† An intermediate understanding of complex analysis is necessary to understand the material in this and the next chapter.

Definition 16.1. A function f has a pole at z_0 if it can be represented by a Laurent series centered about z_0 with only finitely many terms of negative exponent, i.e.,

$$f(z) = \sum_{k=-n}^{\infty} a_k (z - z_0)^k$$

in some neighborhood of z_0, with $a_{-n} \neq 0$, for some $n \in \mathbb{N}$. The number n is called the order of the pole. A simple pole is of pole of order 1.

Definition 16.2. The constant a_{-1} in the Laurent series

$$f(z) = \sum_{k=-\infty}^{\infty} a_k (z - z_0)^k$$

of $f(z)$ about a point z_0 is called the residue of $f(z)$. The residue of a function f at a point z_0 is often denoted $\mathrm{Res}(f; z_0)$.

If f is analytic at z_0, its residue is zero, but the converse is not always true.

The residues of a meromorphic function at its poles characterize a great deal of the structure of a function – residues appear in the residue theorem of contour integration, which, briefly, is a very powerful theorem used to evaluate contour integrals of analytic functions.

The residues of a function $f(z)$ may be found without explicitly expanding into a Laurent series. If $f(z)$ has a pole of order n at z_0, then $a_k = 0$ for $k < -n$ and $a_{-n} \neq 0$.

Thus,

$$f(z) = \sum_{k=-n}^{\infty} a_k (z - z_0)^k = \sum_{k=0}^{\infty} a_{k-n} (z - z_0)^{k-n}$$

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

Differentiating,

$$
\begin{aligned}
\frac{d}{dz}\left[(z-z_0)^n f(z)\right] &= \sum_{k=0}^{\infty} k a_{k-n}(z-z_0)^{k-1} \\
&= \sum_{k=1}^{\infty} k a_{k-n}(z-z_0)^{k-1} \\
&= \sum_{k=0}^{\infty} (k+1) a_{k-n+1}(z-z_0)^{k}
\end{aligned}
$$

$$
\begin{aligned}
\frac{d^2}{dz^2}\left[(z-z_0)^n f(z)\right] &= \sum_{k=0}^{\infty} k(k+1) a_{k-n+1}(z-z_0)^{k-1} \\
&= \sum_{k=1}^{\infty} k(k+1) a_{k-n+1}(z-z_0)^{k-1} \\
&= \sum_{k=0}^{\infty} (k+1)(k+2) a_{k-n+2}(z-z_0)^{k}
\end{aligned}
$$

Iterating,

$$
\frac{d^{n-1}}{dz^{n-1}}\left[(z-z_0)^n f(z)\right] = \sum_{k=0}^{\infty} (k+1)(k+2)\cdots(k+n-1)\; a_{k-1}(z-z_0)^k
$$

$$
= (n-1)!\; a_{-1} + \sum_{k=1}^{\infty} (k+1)(k+2)\cdots(k+n-1)\; a_{k-1}(z-z_0)^k
$$

so we have

$$\lim_{z\to z_0} \frac{d^{n-1}}{dz^{n-1}}\big[(z-z_0)^n f(z)\big] = \lim_{z\to z_0} (n-1)!\, a_{-1} + 0$$
$$= (n-1)!\, a_{-1}$$

and the residue is

$$a_{-1} = \frac{1}{(n-1)!} \frac{d^{n-1}}{dz^{n-1}}\big[(z-z_0)^n f(z)\big]_{z=z_0}$$

Proposition 16.1. *In the case that the limit* $lim_{z\to z_0}(z-z_0)f(z)$ *exists and has a non-zero value* r*, the point* $z = z_0$ *is a pole of order 1 for the function* f *and*

$$\mathrm{Res}(f; z_0) = r$$

This result follows directly from the preceding discussion.

Now, the Gamma function may be expressed

$$\Gamma(z) = \frac{\Gamma(z+n)}{z(z+1)(z+2)\cdots(z+n-1)}. \tag{16.1}$$

According to the standard definition

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

$\Gamma(z)$ is defined only in the right half-plane $\mathrm{Re}(z) > 0$, whereas equation 16.1 is defined and meromorphic in the half-plane $\mathrm{Re}(z) > -n$ where it has poles of order 1 at the points $0, -1, -2, \ldots, -(n-1)$. Equation 16.1 is the analytic continuation of $\Gamma(z)$ to the half-plane $\mathrm{Re}(z) > -n$. Since the positive integer n can be chosen arbitrarily, Euler's Gamma function has been analytically continued to the whole complex plane.

For determining the residues of the Gamma function at the poles, rewrite equation 16.1 as

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

from which we see

$$(z+n)\Gamma(z) = \frac{\Gamma(z+n+1)}{z(z+1)\cdots(z+n-1)}.$$

At the point $z = -n$,

$$\Gamma(z+n+1) = \Gamma(1) = 0! = 1$$

and

$$z(z+1)\cdots(z+n-1) = (-1)^n n!$$

which, taken together, imply that

$$\mathrm{Res}(\Gamma; -n) = \frac{(-1)^n}{n!}.$$

17

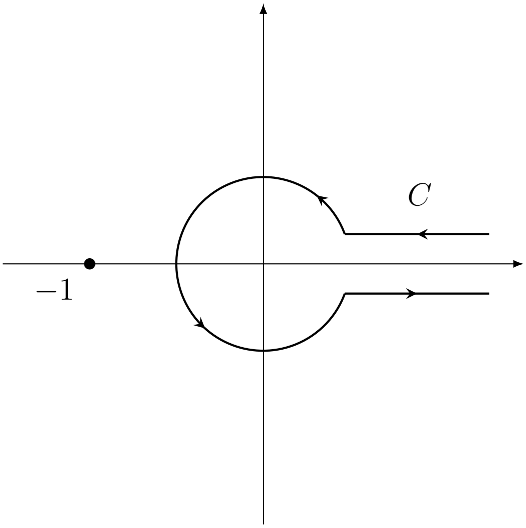
Hankel's Contour Integral Representation

In this chapter we find integral representations of the Gamma function and its reciprocal. Numerical evaluation of Hankel's integral is the basis of some of the best methods for computing the Gamma function.

Consider the integral

$$\mathcal{I} = \int_C (-t)^{z-1} e^{-t} \, \mathrm{d}t,$$

where $\mathrm{Re}(z) > 0$ and where the contour C starts slightly above the real axis at $+\infty$, then runs down to $t = 0$ where it winds around counterclockwise in a small circle, and then returns to $+\infty$ just below the real axis.



The t-plane is cut from 0 to ∞. We define the phase such that $\arg(-t) \equiv 0$ on the negative real t-axis. On the contour C we have $-\pi < \arg(t) < \pi$. Therefore, just above the positive real t-axis we have $\arg(-t) = -\pi$, whereas just below we have $\arg(-t) = +\pi$, the angle measured counterclockwise.

Given these definitions, it then follows that

$$(-t)^{z-1} = e^{-i\pi(z-1)} t^{z-1}$$

just above the positive real axis, and

$$(-t)^{z-1} = e^{+i\pi(z-1)} t^{z-1}$$

just below the positive real axis. And on the small circle enclosing $t = 0$ we have $-t = \delta e^{i\theta}$. So then our integral becomes

$$\begin{aligned}\mathcal{I} = \int_{\infty}^{\delta} e^{-i\pi(z-1)} t^{z-1} e^{-t}\,\mathrm{d}t + \int_{\delta}^{\infty} e^{+i\pi(z-1)} t^{z-1} e^{-t}\,\mathrm{d}t \\ + \int_{-\pi}^{\pi} \left(\delta e^{i\theta}\right)^{z-1} e^{\delta(\cos\theta + i\sin\theta)} \delta e^{i\theta} i\,\mathrm{d}\theta \\ \to -2i\sin(\pi z) \int_0^{\infty} t^{z-1} e^{-t}\,\mathrm{d}t \quad \text{for} \quad \delta \to 0.\end{aligned}$$

The expression above contains the usual representation for the Gamma function, so $\mathcal{I} = -2i\sin(\pi z)\Gamma(z)$, or

$$\Gamma(z) = \frac{-1}{2i\sin\pi z} \int_C (-t)^{z-1} e^{-t}\,\mathrm{d}t. \tag{17.1}$$

This is Hankel's integral representation for the Gamma function, valid for all $z \neq 0, \pm1, \pm2, \ldots$. It has several equivalent forms. Considering the fact that $e^{\pm i\pi} = -1$, we find

$$\begin{aligned}&\frac{-1}{2i\sin\pi z} \int_C (-t)^{z-1} e^{-t}\,\mathrm{d}t \\ &= \frac{-1\cdot(-1)^{z-1}}{2i\sin\pi z} \int_C t^{z-1} e^{-t}\,\mathrm{d}t \\ &= \frac{e^{-i\pi z}}{e^{i\pi z} - e^{-i\pi z}} \int_C t^{z-1} e^{-t}\,\mathrm{d}t \\ &= \frac{1}{e^{2\pi i z} - 1} \int_C t^{z-1} e^{-t}\,\mathrm{d}t,\end{aligned}$$

so another form of the contour integral representation is given by

$$\Gamma(z) = \frac{1}{e^{2\pi i z} - 1} \int_C t^{z-1} e^{-t} \, \mathrm{d}t.$$

The trivial substitution $s = -t$ produces yet another equivalent form,

$$\Gamma(z) = \frac{1}{2i \sin \pi z} \int_C s^{z-1} e^{s} \, \mathrm{d}s.$$

An application of Euler's reflection formula,

$$\Gamma(z)\Gamma(1-z) = \frac{\pi}{\sin \pi z},$$

to Equation 17.1 leads immediately to the contour integral representation for the reciprocal of the Gamma function,

$$\begin{aligned}
\frac{1}{\Gamma(z)} &= \frac{\sin \pi z}{\pi} \Gamma(1-z) \\
&= \frac{\sin \pi z}{\pi} \cdot \frac{-1}{2i \sin(\pi(1-z))} \int_C (-t)^{-z} e^{-t} \, \mathrm{d}t \\
&= \frac{-1}{2\pi i} \int_C (-t)^{-z} e^{-t} \, \mathrm{d}t.
\end{aligned}$$

So we have for the contour integral representation of the reciprocal gamma function,

$$\frac{1}{\Gamma(z)} = \frac{i}{2\pi} \int_C (-t)^{-z} e^{-t} \, \mathrm{d}t.$$

An equivalent representation is found by making the substitution $s = -t$,

$$\frac{1}{\Gamma(z)} = \frac{1}{2\pi i} \int_C (s)^{-z} e^s \, \mathrm{d}s.$$

Sometimes we write $\int_\infty^{(0+)}$ for $\int_C$ meaning a path starting at infinity on the real axis, encircling zero in the positive (counterclockwise) sense and then returning to infinity along the real axis, respecting the cut along the positive real axis.

A

The Weierstrass Factor Theorem

The most important property of the entire rationals is expressed in the fundamental theorem of algebra: *Every non-constant entire rational function has zeros.* If we have an arbitrary, non-constant entire rational function,

$$g_0(z) = a_0 + a_1 z + \cdots + a_m z^m, \qquad (m \geq 1, a_m \neq 0)$$

then it follows from the fundamental theorem of algebra that $g_0(z)$ can be written

$$g_0(z) = a_m (z - z_1)^{\alpha_1} (z - z_2)^{\alpha_2} \cdots (z - z_k)^{\alpha_k}$$

where $z_1, z_2, \ldots, z_k$ denote all the distinct zeros of $g_0(z)$, and $\alpha_1, \alpha_2, \ldots, \alpha_k$ denote their respective orders.

For every entire rational function there is a so-called *factor representation.* From this representation, we infer that every other entire rational function $g(z)$ which has the same zeros to the same respective orders can differ from $g_0(z)$ only in the factor a_m. It is thus always possible to construct an entire rational function (with prescribed zeros) that can be represented as a product which displays these zeros.

Suppose the entire function to be constructed is to have no zeros at all. Then the constant 1, or the function e^z, or e^{z^2}, or more generally, $e^{h(z)}$ is a solution of the problem, if $h(z)$ is a completely arbitrary entire function.

Theorem A.1. *If $h(z)$ denotes an arbitrary entire function, then $e^{h(z)}$ is the most general entire function with no zeros.*

Proof. We have only to show that if $H(z) = a_0 + a_1 z + a_2 z^2 + \cdots$ is a given entire function with no zeros, then we can determine another entire function $h(z) = b_0 + b_1 z + b_2 z^2 + \cdots$ such that $e^{h(z)} = H(z)$. Now, since $H(z) \neq 0$, we have in particular $a_0 = H(0) \neq 0$. Hence, b_0 can be chosen such that $e^{b_0} = a_0$; for, e^z takes on every value except zero. Likewise, $\frac{1}{H(z)}$ is everywhere single-valued and regular, and is therefore an entire function.
The same is true of $H'(z)$, so that

$$\frac{H'(z)}{H(z)} = c_0 + c_1 z + c_2 z^2 + \cdots$$

is also an entire function, and this series is everywhere convergent.

The same is true of the series

$$\begin{aligned} &b_0 + c_0 z + \frac{c_1}{2} z^2 + \cdots + \frac{c_{n-1}}{n} z^n + \cdots \\ &= b_0 + b_1 z + \cdots + b_n z^n + \cdots \end{aligned}$$

which represents an entire function, $h(z)$. If we set $e^{h(z)} = H_1(z)$, then

$$\frac{H_1'(z)}{H_1(z)} = c_0 + c_1 z + c_2 z^2 + \cdots = \frac{H'(z)}{H(z)}$$

and hence $H_1 \cdot H' - H \cdot H_1' = 0$. Therefore

$$\frac{H \cdot H_1' - H_1 \cdot H'}{H^2} = \frac{d}{dz}\left(\frac{H_1(z)}{H(z)}\right) = 0$$

and the quotient of the two functions $H_1(z)$ and $H(z)$ is constant. For $z = 0$ the value of this constant is 1. Thus,

$$H(z) = H_1(z) = e^{h(z)}.$$

□

Having thus described the case for which no zeros are prescribed, its easy to see the extent to which a general entire function is determined by its zeros. If $G_0(z)$ and

$G(z)$ are two entire functions which coincide in the positions and orders of their zeros, then their quotient is also an entire function, but one with no zeros. $G(z)$ and $G_0(z)$ thus differ by at most a multiplicative entire function with no zeros. Conversely, the presence of such a factor of $G_0(z)$ does not alter the positions or orders of its zeros.

Theorem A.2. *Let $G_0(z)$ be a particular entire function. Then, if $h(z)$ denotes an arbitrary entire function,*

$$G(z) = e^{h(z)} \cdot G_0(z)$$

is the most general entire function whose zeros coincide with those of $G_0(z)$ in position and order.

The question of the possibility and method of constructing a particular entire function with arbitrarily prescribed zeros remains. An entire function has no singularity in the finite part of the plane; therefore it can have only a finite number of zeros in every finite region. The prescribed points must not have a finite limit point as a consequence. If we make this single restriction, an entire function can always be constructed. It can be set up in the form of a product which has the positions and orders of its zeros.

Theorem A.3. ***(Weierstrass factor theorem).*** *Let any finite or infinite set of points having no limit point* *be prescribed, and associate with each of its points a definite positive integer as order. Then there exists an entire function which has zeros to the prescribed orders at precisely the prescribed points, and is otherwise different from zero. It can be represented as a product, from which one can read off again the positions and orders of the zeros. If $G_0(z)$ is one such function,*

$$G(z) = e^{h(z)} \cdot G_0(z)$$

is the most general function satisfying the conditions of the problem, if $h(z)$ denotes an arbitrary entire function.

The entire function satisfying the conditions of the Weierstrass factor theorem will be set up in the form of a product; in general, in the form of an infinite product. As with infinite series, we shall assume the reader is familiar with the simplest facts in the theory of infinite products. We thus present, without proofs, the most important definitions and theorems for our purposes.

Definition A.1. The infinite product

$$u_1 \cdot u_2 \cdots u_v \cdots = \prod_{v=1}^{\infty} u_v \tag{A.1}$$

in which the factors are arbitrary complex numbers, is said to be convergent iff from a certain index on, say for all $v > m$, no factor vanishes, and

$$\lim_{n\to\infty} (u_{m+1} \cdot u_{m+2} \cdots u_n)$$

exists and has a finite value distinct from zero. If we call this limit U_m, then the number

$$U = u_1 \cdot u_2 \cdots u_m \cdot U_m$$

which is obviously independent of m, is regarded as the value of the infinite product A.1.

Theorem A.4. *A convergent product has the value zero iff one of its factors vanishes.*

Theorem A.5. *The infinite product A.1 is convergent iff having chosen an arbitrary $\epsilon > 0$, an index n_0 can be determined such that*

$$|u_{n+1} \cdot u_{n+2} \cdots u_{n+r} - 1| < \epsilon$$

for all $n > n_0$ and all $r \geq 1$.

Since on the basis of this theorem (let $r = 1$ and $n+1 = v$) it is necessary that $\lim_{v\to\infty} u_v = 1$, one usually sets the factors of the product equal to $1 + c_v$, so that instead of dealing with (A.1) one is concerned with products of the form

$$\prod_{v=1}^{\infty}(1 + c_v). \tag{A.2}$$

For these, $c_v \to 0$ is a necessary, but insufficient, condition for convergence.

Definition A.2. The product A.2 is said to be absolutely convergent if

$$\prod_{v=1}^{\infty}(1+|c_v|) \qquad \text{converges.}$$

Theorem A.6. *Absolute convergence is a sufficient condition for ordinary convergence; in other words, the convergence of* $\prod(1+|c_v|)$ *implies convergence of* $\prod(1+c_v)$.

On the basis of this theorem it is sufficient for our purposes to have convergence criteria for absolutely convergent products. The next two theorems settle the question of convergence for such products.

Theorem A.7. *The product* $\prod(1+\gamma_v)$, *with* $\gamma_v \geq 0$, *is convergent iff the series* $\sum \gamma_v$ *converges.*

Theorem A.8. *For* $\prod(1+c_v)$ *to converge absolutely, it is necessary and sufficient that* $\sum c_v$ *converge absolutely.*

The following theorem is similar to one on absolutely convergent series:

Theorem A.9. *If the order in which the factors of an absolutely convergent product occur is changed in a completely arbitrary manner, the product remains convergent and has the same value.*

In addition to products with constant factors, we need products whose factors are functions of a complex variable z. We write these products in the form

$$\prod_{v=1}^{\infty}(1 + f_v(z)). \tag{A.3}$$

We designate as the region of convergence of such a product the set $\mathcal{M}$ of all those points z which (a) belong to the domain of the definition of every $f_v(z)$, and for which (b) the product A.3 is convergent. Accordingly, the product assumes a certain value for every z of $\mathcal{M}$; thus, the product represents in $\mathcal{M}$ a certain (single-valued) function. For our purposes, it is important to possess useful conditions under which such a product, in its region of convergence, represents an analytic function. The following theorem suffices:

Theorem A.10. *Let $f_1(z), f_2(z), \ldots, f_v(z), \ldots$ be an infinite sequence of functions, and suppose a region $\mathcal{G}$ exists in which all these functions are regular. Let $\sum_{v=1}^{\infty} |f_v(z)|$ be uniformly convergent in every closed subregion $\mathcal{G}'$ of $\mathcal{G}$. Then the product A.3 is convergent in the entire region $\mathcal{G}$, and represents a regular function $f(z)$ in $\mathcal{G}$. Moreover, by theorem A.1, this function has a zero at those, and only those, points of $\mathcal{G}$ at which at least one of the factors is equal to zero. The order of such a zero is equal to the sum of the orders to which these factors vanish there.*

Proof. Let $\mathcal{G}'$ be an arbitrary closed subregion of $\mathcal{G}$. For every $m \geq 0$,

$$\sum_{v=m+1}^{\infty} |f_v(z)|, \quad \text{along with } \sum_{v=1}^{\infty} |f_v(z)|$$

converges uniformly in $\mathcal{G}'$. By theorem A.8, the product

$$\prod_{v=m+1}^{\infty} (1 + f_v(z)) \tag{A.4}$$

is absolutely convergent in $\mathcal{G}'$, and represents a certain function there. Let us call this function $F_m(z)$. Choose m such that

$$|f_{n+1}(z)| + |f_{n+2}(z)| + \cdots + |f_{n+r}(z)| < \frac{1}{2} \tag{A.5}$$

for all $n \geq m$, all $r \geq 1$, and all z in $\mathcal{G}'$ – then $F_m(z)$ is regular and nonzero in $\mathcal{G}'$. Indeed, if, for $n > m$, we set

$$\prod_{v=m+1}^{n} (1 + f_v(z)) = P_n \quad \text{and} \quad P_m = 0$$

then we have

$$F_m(z) = \lim_{n\to\infty} P_n$$

$$= \lim_{n\to\infty} [(P_{m+1} - P_m) + (P_{m+2} - P_{m+1}) + \cdots + (P_n - P_{n-1})]$$

or

$$F_m(z) = \sum_{v=m+1}^{\infty} (P_v - P_{v-1}) \tag{A.6}$$

and $F_m(z)$ is thus represented by an infinite series. Now, for $n > m$,

$$\begin{aligned}|P_n| &\leq (1 + |f_{m+1}(z)|) \cdots (1 + |f_n(z)|) \\ &\leq e^{|f_{m+1}(z)| + \cdots + |f_n(z)|} < e^{\frac{1}{2}} < 2\end{aligned}$$

so the inequality

$$|P_v - P_{v-1}| = |P_{v-1}| \cdot |f_v(z)| < 2|f_v(z)|$$

is valid for the terms (from the second onward) of the series just obtained. Consequently, the new series A.6, along with $\sum |f_v(z)|$, is uniformly convergent in $\mathcal{G}'$, and the function $F_m(z)$ defined by that series is a regular function in $\mathcal{G}'$. It is also distinct from zero there. For, by (A.5), we have in $\mathcal{G}'$, for $n \geq m$,

$$|f_{n+1}(z)| < \frac{1}{2}$$

and hence, for $v \geq m + 1$,

$$|1 + f_v(z)| \geq 1 - |f_v(z)| > \frac{1}{2}$$

so that no factor of F_m can be equal to zero. Since

$$f(z) = (1 + f_1(z)) \cdots (1 + f_m(z)) \cdot F_m(z)$$

we have that $f(z)$, together with $F_m(z)$, is regular at every point z of $\mathcal{G}'$, and can vanish at such a point only if one of the factors appearing before $F_m(z)$ vanishes. The order of such a zero is then equal to the sum of the orders to which these factors vanish there.

Now let z be an arbitrary point of $\mathcal{G}$. Since z is an interior point of $\mathcal{G}$, it is always possible to choose $\mathcal{G}'$ such that z also belongs to $\mathcal{G}'$. Hence, the above considerations hold for all of $\mathcal{G}$, and the proof is complete. □

We can also make an assertion concerning the derivative of $f(z)$. Since the ordinary derivative of a product of many factors is difficult to calculate, we introduce the logarithmic derivative. We have the following theorem concerning this derivative:

Theorem A.11. *Given theorem A.10, we have that*

$$\frac{f'(z)}{f(z)} = \sum_{v=1}^{\infty} \frac{f_v'(z)}{1 + f_v(z)} \tag{A.7}$$

for every point z of $\mathcal{G}$ at which $f(z) \neq 0$; i.e., the series on the right is convergent for every such z and is equivalent to the logarithmic derivative of $f(z)$.

Proof. If z is a particular point of the type mentioned above, and if the subregion $\mathcal{G}'$ is chosen so as to contain z, then

$$\frac{f'(z)}{f(z)} = \frac{f_1'(z)}{1 + f_1(z)} + \cdots + \frac{f_m'(z)}{1 + f_m(z)} + \frac{F_m'(z)}{F_m(z)}. \tag{A.8}$$

Since the series A.6 converges uniformly in $\mathcal{G}'$,

$$F'_m(z) = \sum_{v=m+1}^{\infty} (P'_v - P'_{v-1}) = \lim_{n\to\infty} P'_n$$

according to theorem A.6. Here P'_n denotes the derivative of P_n. Since $F_m(z)$ and all P_n for $n > m$ are not zero,

$$\begin{aligned}\frac{F'_m(z)}{F_m(z)} = \lim_{n\to\infty} \frac{P'_n}{P_n} &= \lim_{n\to\infty} \left(\frac{f'_{m+1}(z)}{1 + f_{m+1}(z)} + \cdots + \frac{f'_n(z)}{1 + f_n(z)} \right) \\ &= \sum_{v=m+1}^{\infty} \frac{f'_v(z)}{1 + f_v(z)}\end{aligned}$$

which, with (A.8), proves the assertion. □

Theorem A.12. *The series (A.7) converges absolutely and uniformly in every closed subregion $\mathcal{G}''$ of $\mathcal{G}$ containing no zero of $f(z)$, and hence may be repeatedly differentiated there any number of times term by term.*

Proof. Since none of the factors $(1 + f_v(z))$ can vanish in $\mathcal{G}''$, the absolute value of each remains greater than a positive bound, γ_v say. Since this is certainly greater than $1/2$ for all $v > m$, a positive number γ exists, such that $\gamma_v \geq \gamma > 0$ for all v. Then, for all v and all z in $\mathcal{G}''$,

$$\left| \frac{f'_v(z)}{1 + f_v(z)} \right| < \frac{1}{\gamma} \cdot |f'_v(z)|.$$

From the proof of theorem A.6 it follows that $\sum |f'_v(z)|$ converges uniformly in $\mathcal{G}''$. By the last inequality, this is also true then of the series A.7. □

Having become familiar with infinite products, it is straightforward to prove the Weierstrass factor theorem.

Proof. If only a finite number of points $z_1, z_2, \ldots, z_k$ having the respective orders $\alpha_1, \alpha_2, \ldots, \alpha_k$ are prescribed, then the product

$$(z - z_1)^{\alpha_1}(z - z_2)^{\alpha_2} \cdots (z - z_k)^{\alpha_k} \tag{A.9}$$

is already a solution of the problem, so that this case is settled immediately. If, however, an infinite number of points are prescribed as zeros, because the corresponding product would be meaningless generally. This would still be the case if, with regard to the infinite products we are dealing with here, we were to replace (A.9) by the product

$$\left(1 - \frac{z}{z_1}\right)^{\alpha_1}\left(1 - \frac{z}{z_2}\right)^{\alpha_2} \cdots \left(1 - \frac{z}{z_k}\right)^{\alpha_k} \tag{A.10}$$

which serves the same purpose. We therefore modify our approach – and therein lies the utility of Weierstrass's method.

The set of prescribed points is enumerable, since every finite region can contain only a finite number of them. They can therefore be arranged in a sequence. The way in which the points are numbered is not important. However, if the origin, with the order α_0, is contained among them, we shall call this point z_0 and, leaving it aside for the present, arrange the remaining points in an arbitrary, but then fixed, sequence: $z_1, z_2, \ldots, z_v, \ldots$. Let the corresponding orders be $\alpha_1, \alpha_2, \ldots, \alpha_v, \ldots$. The z_v are all different from zero; and since they have no finite limit point,

$$z_v \to \infty, \quad |z_v| \to +\infty.$$

Consequently, it is possible to assign a sequence of positive integers $k_1, k_2, \ldots, k_v, \ldots$ such that

$$\sum_{v=1}^{\infty} \alpha_v \left(\frac{z}{z_v} \right)^{k_v} \tag{A.11}$$

is absolutely convergent for every z. In fact, it suffices, e.g., to take $k_v = v + \alpha_v$. For, no matter what fixed value z may have, since $z_v \to \infty$, we have for all sufficiently large v

$$\left| \frac{z}{z_v} \right| < \frac{1}{2}$$

and hence

$$\left| \alpha_v \left(\frac{z}{z_v} \right)^{v+\alpha_v} \right| < \alpha_v \left(\frac{1}{2} \right)^{v+\alpha_v} < \left(\frac{1}{2} \right)^{v}$$

and the absolute convergence of the series is thus assured.

Let the numbers k_v be chosen subject to this condition, but otherwise arbitrarily, and keep them fixed. Then we shall prove that the product

$$G_0(z) = z^{\alpha_0} \cdot \prod_{v=1}^{\infty} \left[\left(1 - \tfrac{z}{z_v}\right) \cdot \exp\left\{ \tfrac{z}{z_v} + \tfrac{1}{2}\left(\tfrac{z}{z_v}\right)^2 + \cdots + \tfrac{1}{k_v - 1}\left(\tfrac{z}{z_v}\right)^{k_v - 1} \right\} \right]^{\alpha_v}$$

represents an entire function with the required properties. (Here the factor z^{α_0} appearing before the product symbol is to be suppressed in case the origin is not one of the prescribed zeros. Likewise, if one of the numbers k_v is equal to unity, the corresponding exponential factor does not appear.)

The proof of this assertion is now very simple. To be able to apply our theorems on products, we set the factors of our infinite product equal to $1 + f_v(z)$. According to theorem A.10, we must then merely prove that

$$\sum_{v=1}^{\infty} |f_v(z)| = \sum_{v=1}^{\infty} \left| \left[\left(1 - \frac{z}{z_v}\right) \cdot \right.\right.$$
$$\left.\left. \cdot \exp\left\{ \frac{z}{z_v} + \cdots + \frac{1}{k_v - 1}\left(\frac{z}{z_v}\right)^{k_v - 1} \right\} \right]^{\alpha_v} - 1 \right| \qquad \text{(A.12)}$$

converges uniformly in every bounded region. For then the entire plane can be taken as the region $\mathcal{G}$ of theorem A.10, according to which the infinite product, and consequently also $G_0(z)$, is an entire function. On account of the form of the factors of $G_0(z)$, the second part of theorem A.10 at once yields that $G_0(z)$ also possesses the required properties. The uniform convergence of the series (A.12) in the circle about the origin with radius R ($R > 0$ arbitrary, but fixed) is established as follows:

Since the series (A.11) also converges for $z = R$, and since $z_v \to \infty$, m can be chosen so large that

$$\alpha_v \left| \frac{R}{z_v} \right|^{k_v} < \frac{1}{2} \quad \text{and} \quad \frac{R}{|z_v|} < \frac{1}{2} \tag{A.13}$$

for all $v > m$. Let us for the moment replace z/z_v by u, k_v by k, and α_v by α. Then, for $v > m$, the vth term of the series (A.12) has the form

$$\left| \left[(1-u) \exp \left\{ u + \frac{u^2}{2} + \cdots + \frac{u^{k-1}}{k-1} \right\} \right]^{\alpha} - 1 \right|$$

with $|u| < \frac{1}{2}$ and $\alpha |u|^k < \frac{1}{2}$. Now for $|u| < 1$ we can set

$$1 - u = \exp \left\{ -u - \frac{u^2}{2} - \frac{u^3}{3} \cdots \right\}$$

so that this vth term is further equal to

$$\left| \exp \left\{ \alpha \left(-\frac{u^k}{k} - \frac{u^{k+1}}{k+1} - \cdots \right) \right\} - 1 \right|$$

and hence

$$\leq \exp\left\{\alpha\left(\frac{|u|^k}{k}+\frac{|u|^{k+1}}{k+1}+\cdots\right)\right\}-1$$

$$\leq \exp\left\{\alpha|u|^k(1+|u|+|u|^2+\cdots)\right\}-1 < e^{2\alpha|u|^k}-1$$

because $|u| < \frac{1}{2}$. Further, since $e^x - 1 \leq xe^x$ for $x \geq 0$, the vth term is less than or equal to

$$2\alpha|u|^k e^{2\alpha|u|^k} < 6\alpha|u|^k$$

the exponent of e being smaller than one, according to (A.13). Hence, for all sufficiently large v and all $|z| \leq R$ we have

$$|f_v(z)| < 6\alpha_v\left|\frac{z}{z_v}\right|^{k_v} \leq 6\alpha_v\left|\frac{R}{z_v}\right|^{k_v}.$$

But these are positive numbers whose sum converges (because of the manner in which the k_v were chosen). Therefore, by Weierstrass's M-test, $|f_v(z)|$ is uniformly convergent in the circle with radius R about the origin as center, so the proof of the Weierstrass factor theorem is complete. □

The product is simplest if the prescribed zeros and orders are such, that the series $\sum \alpha_v/z_v$, and consequently, for every z, the series $\sum \alpha_v(z/z_v)$, converges absolutely for our sequence $z_1, z_2, \ldots$. For then it is possible to take all $k_v = 1$, and the desired function is obtained simply in the form

$$G_0(z) = z^{\alpha_0} \cdot \prod_{v=1}^{\infty} \left(1 - \frac{z}{z_v}\right)^{\alpha_v}$$

If, for example, the points $0, 1, 4, 9, \ldots, v^2, \ldots$ are to be zeros of order unity, then we have that

$$G(z) = e^{h(z)} \cdot z \cdot \prod_{v=1}^{\infty} \left(1 - \frac{z}{v^2}\right)$$

with $h(z)$ an arbitrary entire function, is the most general solution of the problem. If the points $1, 8, \ldots, v^3, \ldots$ are to be zeros of respective orders $1, 2, \ldots, v, \ldots$, then

$$G(z) = e^{h(z)} \cdot \prod_{v=1}^{\infty} \left(1 - \frac{z}{v^3}\right)^{v}$$

is the most general solution.

Now we present an application of the factor theorem which is of importance to an earlier topic in the text, namely, in the derivation of Euler's reflection formula where we used an identity known as the product representation of the sine function.

Consider constructing an entire function which has zeros, of order unity, at precisely all the real lattice points (i.e., at $0, \pm 1, \pm 2, \ldots$). We number these points so that $z_0 = 0, z_1 = +1, z_2 = -1, \ldots, z_{2v-1} = v, z_{2v} = -v, \ldots,$ with $(v = 1, 2, \ldots)$. The series

$$\sum_{v=1}^{\infty} \left(\frac{z}{z_v} \right)^2 = z^2 \cdot \sum_{v=1}^{\infty} \frac{1}{z_v^2}$$

is absolutely convergent for every z, and we can therefore take all $k_v = 2$.

Then we have

$$\begin{aligned} G(z) &= e^{h(z)} \cdot z \cdot \prod_{v=1}^{\infty} \left[\left(1 - \frac{z}{z_v} \right) e^{z/z_v} \right] \\ &= e^{h(z)} \cdot z \cdot \prod_{v=1}^{\infty} \left[\left(1 - \frac{z}{v} \right) e^{z/v} \right] \left[\left(1 + \frac{z}{v} \right) e^{-z/v} \right] \\ &= e^{h(z)} \cdot z \cdot \prod_{v=1}^{\infty} \left(1 - \frac{z^2}{v^2} \right) \end{aligned}$$

as the most general solution of the problem.

Since the function $\sin \pi z$ is evidently also a solution of the problem, it must be contained in the expression just found. That is, there exists a certain entire function, which we shall call $h_0(z)$, such that

$$\sin \pi z = e^{h_0(z)} \cdot z \cdot \prod_{v=1}^{\infty} \left(1 - \frac{z^2}{v^2}\right). \tag{A.14}$$

If we can succeed in obtaining this function $h_0(z)$, we shall have the factor representation of $\sin \pi z$.

The function $h_0(z)$ certainly cannot be ascertained from a knowledge of the zeros alone. On the contrary, for its determination we must make use of further properties of the function $\sin \pi z$; e.g., its power series expansion, its periodicity properties, the conformal map effected by it, its behavior at infinity, and so on. Now to determine $h_0(z)$.

First, to show that $h_0''(z)$ is a constant. According theorem A.11, it follows from (A.14) that

$$\pi \cot \pi z = h_0'(z) + \frac{1}{z} + \sum_{v=1}^{\infty} \left(\frac{1}{z-v} + \frac{1}{z+v}\right). \tag{A.15}$$

According to theorem A.12, this expression may be differentiated repeatedly term by term. Thus,

$$-\frac{\pi^2}{\sin^2 \pi z} = h_0''(z) - \frac{1}{z^2} - \sum_{v=1}^{\infty} \left(\frac{1}{(z-v)^2} + \frac{1}{(z+v)^2}\right)$$

or, more succinctly,

$$h_0''(z) = \sum_{v=-\infty}^{+\infty} \frac{1}{(z-v)^2} - \frac{\pi^2}{\sin^2 \pi z}.$$

This relation holds in every closed region which contains no real lattice points. If we replace z by $z+1$ in the rightmost member, it is not altered because $sin^2 \pi z$ has the period $+1$, and

$$\sum_{v=-\infty}^{+\infty} \frac{1}{(z+1-v)^2} = \sum_{v=-\infty}^{\infty} \frac{1}{(z-(v-1))^2} = \sum_{\mu=-\infty}^{+\infty} \frac{1}{(z-\mu)^2}.$$

Hence, $h_0''(z)$ is an entire function with the period $+1$. In order to show that $h_0''(z)$ is a constant, it is sufficient to show that $|h_0''(z)|$ cannot become arbitrarily large. On account of the periodicity of $h_0''(z)$ which we just established, it is sufficient, for this purpose, to show that a constant K exists such that $|h_0''(z)| < K$ for all $z = x+iy$ for which $0 \leq x \leq 1$ and $|y| \geq 1$.

Now for these z,

$$\left| \sum_{v=-\infty}^{+\infty} \frac{1}{(z-v)^2} \right| \leq \sum_{v=-\infty}^{+\infty} \frac{1}{(x-v)^2 + y^2} \leq 2 \sum_{n=0}^{\infty} \frac{1}{n^2 + y^2}$$

and, since $\sin \pi z = (1/2i)(e^{i\pi z} - e^{-i\pi z})$,

$$\left| \frac{\pi^2}{\sin^2 \pi z} \right| = \frac{4\pi^2}{e^{2\pi y} + e^{-2\pi y} - 2\cos 2\pi x} < \frac{4\pi^2}{e^{2\pi|y|} - 2}$$

for those z. Consequently,

$$\left| h_0''(z) \right| < 2 \sum_{n=0}^{\infty} \frac{1}{n^2 + y^2} + \frac{4\pi^2}{e^{2\pi|y|} - 2}$$

there, and this expression certainly remains less than a fixed bound for all $|y| \geq 1$. Hence,

$$h_0''(z) = \text{ constant } = c''.$$

According to the inequality just obtained, $|h_0''(z)|$ is arbitrarily small if $|y|$ is sufficiently large; hence c'' must be equal to zero. Therefore

$$h_0''(z) = 0, \qquad h_0'(z) = \text{ constant } = c'$$

and hence by A.15

$$\pi \cot \pi z = c' + \frac{1}{z} + \sum_{v=1}^{\infty} \frac{2z}{z^2 - v^2}.$$

Now if we substitute $-z$ for z in this equality, we see that $c' = -c'$, and hence $c' = 0$. Then $h_0(z)$ and $e^{h_0(z)}$ are also constant. Therefore

$$\sin \pi z = c \cdot z \cdot \prod_{v=1}^{\infty} \left(1 - \frac{z^2}{v^2} \right).$$

If we divide through by z and allow z to approach zero, we obtain $\pi = c$. We thus have,

$$\sin \pi z = \pi z \cdot \prod_{v=1}^{\infty} \left(1 - \frac{z^2}{v^2} \right)$$

valid for all z, which is the product representation of the sine function we set out to find.

B

The Mittag-Leffler Theorem

For every fractional rational function there is a so-called *decomposition into partial fractions*, in which the poles and corresponding principal parts are apparent. Thus, let $f_0(z)$ be the given rational function, and let $z_1, z_2, \ldots, z_k$ be its poles with the corresponding principal parts

$$h_v(z) = \frac{a_{-1}^{(v)}}{z - z_v} + \frac{a_{-2}^{(v)}}{(z - z_v)^2} + \cdots + \frac{a_{-\alpha_v}^{(v)}}{(z - z_v)^{\alpha_v}} \tag{B.1}$$

where $v = 1, 2, \ldots, k$. Then we can set

$$f_0(z) = g_0(z) + h_1(z) + h_2(z) + \cdots + h_k(z) \tag{B.2}$$

where $g_0(z)$ is a suitable entire rational function. From this decomposition into partial fractions, we infer that every other rational function $f(z)$ having the same poles with the same respective principal parts can differ from

$f_0(z)$ in the term $g_0(z)$ alone. Furthermore, one can arbitrarily assign these poles and their principal parts. In other words, it is always possible to construct a rational function whose poles and their principal parts are prescribed. This function can be represented as a partial-fractions decomposition which displays these poles and their principal parts. The most general function of this kind is obtained from a particular one by adding to it an arbitrary entire rational function.

These fundamentals concerning rational functions can be carried over to the more general class of meromorphic functions.

Definition B.1. A single-valued function shall – without regard to its behavior at infinity – be called meromorphic, if it has no singularities other than at poles in the entire plane.

Theorem B.1. *A meromorphic function has in every finite region at most a finite number of poles.*

For otherwise there would exist a finite limit point of poles, and this point would be singular, but certainly not a pole.

According to this, the rational functions are special cases of meromorphic functions, and the entire functions must also be regarded as such.

The function $1/\sin z$ is meromorphic because in the finite part of the plane it has a singularity, namely a pole of order unity, only wherever $\sin z$ has a zero. We see, likewise, that $\cot z = \cos z/\sin z$ and $\tan z$ are meromorphic functions. More generally, if $G(z)$ denotes any entire function, its reciprocal, $1/G(z)$, is a meromorphic function. For, $1/G(z)$ has poles (but otherwise no singularities) at those, and only those, points at which $G(z)$ has zeros; and the orders of both are the same. If $G_1(z)$ is an entire function which has no zeros in common with $G(z)$, we see that $G_1(z)/G(z)$ is a meromorphic function whose poles coincide in position and order (although, in general, not in their principal parts!) with those of $1/G(z)$.

We inquire whether, and how, one can construct a meromorphic function if its poles and the corresponding principal parts are prescribed, and to what extent a meromorphic function is determined by these conditions.

If $M_0(z)$ and $M(z)$ are two meromorphic functions which coincide in their poles and the corresponding principal parts, then their difference, $M(z)-M_0(z)$, is evidently an entire function. Consequently, they differ by at most an additive entire function (a meromorphic function with no poles). Conversely, since the addition of such a function to $M_0(z)$ does not alter its poles or the corresponding principal parts, we are able to say:

Theorem B.2. *Let $M_0(z)$ be a particular meromorphic* *function. Then, if $G(z)$ denotes an arbitrary entire function,*

$$M(z) = M_0(z) + G(z)$$

is the most general meromorphic function which coincides with $M_0(z)$ in its poles and the corresponding principal parts.

There remains only the possibility and method of constructing a particular meromorphic function with arbitrarily prescribed poles.

According to theorem B.1, the set of assigned poles cannot have a finite limit point. If this is excluded, however, then the problem posed can be solved without any further restriction.

Theorem B.3. ***(Mittag-Leffler partial-fractions theorem).*** *Let any finite or infinite set of points having no finite limit point be prescribed, and associate with each of its points a principal part, i.e., a rational function of the special form (B.1). Then there exists a meromorphic function which has poles with the prescribed principal parts at precisely the prescribed points, and is otherwise regular. It can be represented in the form of a partial-fractions decomposition from which one can read off again the poles along with their principal parts. Further, by theorem B.2, if $M_0(z)$ is one such function,*

$$M(z) = M_0(z) + G(z)$$

is the most general function satisfying the conditions of the problem, if $G(z)$ denotes an arbitrary entire function.

If we let $M(z)$ be an arbitrarily given meromorphic function, the set of its poles has no finite limit point. Hence, according to Mittag-Leffler's theorem, another meromorphic function, $M_0(z)$, having the same poles and principal parts as $M(z)$, can be constructed in the form of a partial-fractions decomposition displaying these. Then by theorem B.2,

$$M(z) = M_0(z) + G_0(z)$$

where $G_0(z)$ denotes a suitable entire function. We have thus actually obtained a decomposition of the given meromorphic function $M(z)$ into partial fractions, from which its poles and the corresponding principal parts can be read off.

Proof. If the function to be constructed is to have no poles at all, then every entire function is a solution of the problem. If it is to have the finitely many poles $z_1, z_2, \ldots, z_k$ with the respective principal parts $h_1(z)$, $h_2(z)$, ..., $h_k(z)$, then

$$M_0(z) = h_1(z) + h_2(z) + \cdots + h_k(z)$$

is a solution. If, however, an infinite number of poles are prescribed, we cannot attain our goal simply because the analogous series, being infinite, would generally diverge. Nevertheless, we can produce the convergence by means of a suitable modification of the terms of the series.

If the origin is a prescribed pole, we denote it by z_0 and leave it aside for now. Let $h_0(z), h_1(z), \ldots, h_v(z), \ldots$ be the principal parts corresponding to the points z_0, z_1, ..., z_v, ...; $h_v(z)$ is understood to be an expression of the type appearing in (B.1). Each of these functions $h_v(z)$, $v = 1, 2, 3, \ldots$, is regular in a neighborhood of the origin. Its power-series expansion

$$h_v(z) = a_0^{(v)} + a_1^{(v)} z + a_2^{(v)} z^2 + \cdots \qquad (v = 1, 2, \ldots)$$

for this neighborhood converges for all $|z| < |z_v|$; hence, it is uniformly convergent for all $|z| \leq \frac{1}{2}|z_v|$. Consequently (for every $v = 1, 2, 3, \ldots$) an integer n_v can be determined such that the remainder of the power series after the n_vth term remains, in absolute value, less than any preassigned positive number, e.g., $1/2^v$. Denote the sum of the first n_v terms of the series by $g_v(z)$. Thus, $g_v(z)$ is an entire rational function of degree n_v:

$$g_v(z) = a_0^{(v)} + a_1^{(v)} z + \cdots + a_{n_v}^{(v)} z^{n_v} \qquad (v = 1, 2, 3, \ldots)$$

and for all $|z| \leq \frac{1}{2}|z_v|$ we have

$$|h_v(z) - g_v(z)| < \frac{1}{2^v}.$$

Then

$$M_0(z) = h_0(z) + \sum_{v=1}^{\infty} [h_v(z) - g_v(z)]$$

is a meromorphic function satisfying the conditions of the theorem. (If the origin is not assigned as a pole, the term $h_0(z)$ must be omitted.

To prove this, we must merely show that the right-hand side defines an analytic function having in every finite domain, e.g., a circle with radius R about the origin as center, exactly the prescribed singularities and no others.

Now, $|z_v| \to +\infty$. Therefore it is possible to choose m so large, that $|z_v| > 2R$, and hence $R < \frac{1}{2}|z_v|$, for all $v > m$. Then, for all $|z| \le R$ and all $v > m$,

$$|z| < \frac{1}{2}|z_v| \quad \text{and consequently} \quad |h_v(z) - g_v(z)| < \frac{1}{2^v}.$$

Hence, for all $|z| \le R$, the series

$$\sum_{v=m+1}^{\infty} [h_v(z) - g_v(z)]$$

is absolutely and uniformly convergent. Since its terms are regular for $|z| \le R$ (because the poles of the $h_v(z)$ with $v > m$ lie outside the circle $|z| = R$), it defines there a regular function which we shall denote by $F_m(z)$. Then evidently

$$M_0(z) = h_0(z) + \sum_{v=1}^{m}[h_v(z) - g_v(z)] + F_m(z)$$

is also an analytic function which is regular in the circle with radius R about the origin as center, with the exception of those points z_v in this circle which are poles with principal parts $h_v(z)$. The same is valid for every finite region, because R was completely arbitrary and hence, $M_0(z)$ is a meromorphic function with the necessary properties. $\square$

From the proof it follows that it is sufficient to take the degree n_v of the polynomial $g_v(z)$ (the sum of the first n_v terms of the power series for $h_v(z)$) so large that having chosen an arbitrary $R > 0$, the terms $|h_v(z) - g_v(z)|$ for all $|z| \leq R$ finally (i.e., for all sufficiently large v) remain less than the terms of a convergent series of positive terms.

The convergence terms $g_v(z)$ are not always necessary. Then, of course, the function to be constructed is especially simple. If, e.g., the points $0, 1, 4, \ldots, v^2, \ldots$ are to be poles of order unity with respective principal parts $1/(z - v^2)$, then

$$M_0(z) = \frac{1}{z} + \sum_{v=1}^{\infty} \frac{1}{z - v^2} = \sum_{v=0}^{\infty} \frac{1}{z - v^2}$$

is a solution. For, let $R > 0$ be chosen arbitrarily, and $m > \sqrt{2R}$. Then the series from $v = m + 1$ on is evidently uniformly convergent in $|z| \leq R$, which proves the assertion.

Consider the case of $\cot \pi z$. The real lattice points are to be poles of order unity with the residue $+1$, and hence, with the principal parts

$$h_v(z) = \frac{1}{z - z_v}, \qquad (z_0 = 0, z_{2v-1} = v, z_{2v} = -v).$$

For $v = 1, 2, 3, \dots,$

$$h_v(z) = -\frac{1}{z_v} - \frac{z}{z_v^2} - \frac{z^2}{z_v^3} - \cdots$$

and it suffices to take all $n_v = 0$, and hence,

$$g_v(z) = -\frac{1}{z_v}$$

because then for all sufficiently large v (namely, for all $v > 4R$) and all $|z| \leq R$,

$$|h_v(z) - g_v(z)| \leq \frac{R}{|z_v|(|z_v| - R)} < \frac{2R}{|z_v|^2}$$

so that the $|h_v(z) - g_v(z)|$ finally remain less than the terms of an obviously convergent series of positive terms. Consequently, according to the concluding remark of the preceding paragraph, if $G(z)$ is an arbitrary entire function,

$$
\begin{aligned}
M(z) &= G(z) + \frac{1}{z} + \sum_{v=1}^{\infty} \left[\frac{1}{z - z_v} + \frac{1}{z_v} \right] \\
&= G(z) + \frac{1}{z} + \sum_{v=1}^{\infty} \left(\left[\frac{1}{z - v} + \frac{1}{v} \right] + \left[\frac{1}{z + v} - \frac{1}{v} \right] \right) \\
&= G(z) + \frac{1}{z} + \sum_{v=1}^{\infty} \left[\frac{1}{z - v} + \frac{1}{z + v} \right]
\end{aligned}
$$

is the most general function of the kind required. The function $\cot \pi z$ also has poles of order unity at the points $0, \pm 1, \pm 2, \dots$. If n is one of them, the residue at this point is

$$
\lim_{z \to n} \frac{(z - n) \cos \pi z}{\sin \pi z} = \lim_{z \to n} \frac{(z - n)[(-1)^n + \cdots]}{(-1)^n \pi (z - n) + \cdots} = \frac{1}{\pi}
$$

which can be read off from the series expansion for a neighborhood of the point $z = n$. Hence, the function $\pi \cot \pi z$ is contained among the functions $M(z)$ which we constructed.

The still undetermined entire function $G(z)$, which there was called $h_0'(z)$, cannot be ascertained solely from the nature and position of the poles. We should, as before, have to make use of special properties of the function. However, in determining the product $\sin \pi z$, we have already discovered that we have to set $h_0'(z)$, that is, $G(z)$, equal to zero. Therefore

$$\pi \cot \pi z = \frac{1}{z} + \sum_{v=1}^{\infty} \left[\frac{1}{z-v} + \frac{1}{z+v} \right]$$

which is the partial-fractions decomposition of the cotangent function.