CHAPTER 9

THE CALCULUS AND MATHEMATICAL ANALYSIS

THE ORIGINS AND BASIC NOTIONS OF THE CALCULUS

IN THE SIXTEENTH CENTURY the central concern of *physics* was the investigation of *motion*. The motion of a body in a given trajectory, a straight line, for instance, is determined by the way in which the *distance* changes with *time*. Thus *Galileo Galilei* (1564–1642) made the important discovery that the distance covered by a falling body increases in direct proportion to the square of the time. In general, descriptions of motion express the distance s in terms of the time t: Galileo's law may then be expressed in the form

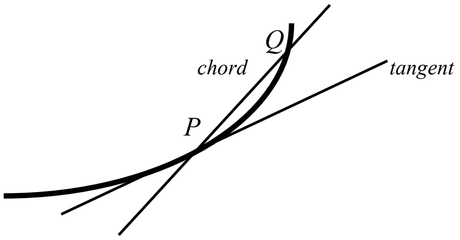
$$s = \tfrac{1}{2}gt^2,$$

where g is the acceleration due to gravity (9.81 m/s^2). Here both s and t are examples of *variable quantities* or *variables*, since they take (continuously) varying values. The values of s and t are related in such a way that to each value of t there corresponds a unique value of s: this state of affairs is expressed—as we have observed—by asserting that the *dependent variable s* is a *function* of the *independent variable t*.

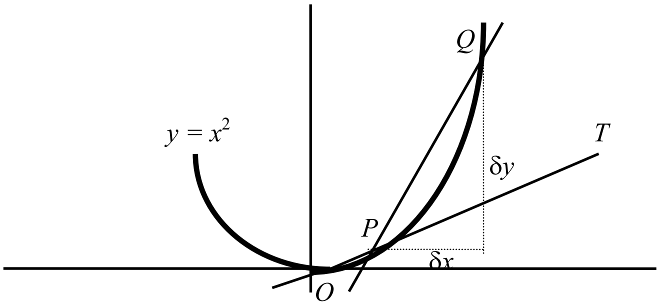
The mathematicians of the seventeenth century found that the problems arising in the analysis of motion were closely related to certain geometric problems which had long resisted solution, and that these problems could be broadly divided into two classes. In the first class was to be found, for example, the problem of determining the velocity at any instant of a given nonuniform motion, or, more generally, finding the rate of change of a continuously varying magnitude, and the associated geometric problem of constructing the tangent to a given curve at a point on it. Efforts to solve problems of this kind led to what is now known as the *differential calculus*. The second class of problems included that of calculating the area of a given curved figure (the problem of *quadrature*) and the associated kinematical problem of finding the distance traversed in a nonuniform motion, or, more generally, determining the total effect of the action of a continuously changing magnitude. These problems issued in what is now known as the *integral calculus*. In this way two fundamental problems came to be identified—that of *tangents* and that of *quadrature.* The most important discovery in the mathematics arising from the analysis of motion was made independently by

Newton and *Georg Friedrich Wilhelm Leibniz* (1646–1716) , who showed that in a certain sense these two problems are *inverse* to one another. Thus the differential and integral calculus came to be fused into a single mathematical discipline, known ever since as the *calculus.*

Let us first consider the problem of constructing the tangent to a curve at a given point. If we are given a curve and *two* points *P* and *Q* on it, it is a simple matter to draw the line or *chord* passing through them. As *Q* approaches *P,* the chord *PQ*, Suitably produced, evidently provides a better and better approximation to the tangent at *P.* However, the chord *PQ* will differ (in general) from the tangent at *P* unless *P* and *Q* coincide. But when *P* and *Q* are coincident, they do not determine a unique line, since there are infinitely many straight lines passing through a given point. In the modern version of the differential calculus this problem is circumvented by obtaining the tangent to the curve at *P* as the line that *limits* the position of the chord *PQ* as *Q* approaches *P*—limiting in the sense that the angle between the tangent at *P* and the chord *PQ* can be made as small as we please by taking *Q* sufficiently close to *P*.



As an example, let us determine the slope of the tangent at the point $P = (x_0, y_0)$ on the parabola $y = x^2$. Let *Q* be the point $(x_0+\delta x, y_0+\delta y)$, where $\delta x, \delta y$ are small increments in x_0 , y_0 respectively. The slope of the chord *PQ* is given by the equation



$$\delta y/\delta x = \frac{(x_0+\delta x)^2 - x_0^2}{\delta x} = 2x_0 + \delta x.$$

When *Q* is close to *P*, the increment δx is small, and by making it small enough we can bring *Q* as close to *P* as we please. Let *PT* be the line through *P* whose slope is $2x_0$. Then this line *PT* is the tangent at *P,* because the slope of *PT* is $2x_0$, that of *PQ* is $2x_0 + \delta x$, and we can take *Q* so close to *P* that the difference between these two slopes, namely δx, will be as small as we please. The slope of the tangent at (x_0, y_0) is, accordingly, that of *PT,* namely $2x_0$.

The process by which we passed from the slope $\delta y/\delta x$ of the chord *PQ* to the slope of the tangent at *P* is known as the *method of limits*. Thus, in our example above, where $\delta y/\delta x$ is equal to $2x_0 + \delta x$, the slope of the tangent, namely $2x_0$, is said to be the *limit* of the slope $2x_0 + \delta x$ as δx *tends to zero*. In this definition, there is present a quantity δx which we suppose assumes smaller and smaller values tending towards zero, that is, δx is a *variable tending to zero*. In a figurative sense, δx may be thought of as an *infinitesimal quantity*.

The concept of limit, one of the most important and far-reaching in mathematics, is extended to arbitrary functions as follows. Given a function $y = f(x)$, we say that *A* is the *limit of* $f(x)$ *as x tends to a* if the difference between $f(x)$ and *A* remains as small as we please so long as *x* remains sufficiently near to *a*, *while remaining distinct from a*. The notation for a limit is *lim*, and the statement that the function $y = f(x)$ has limit *A* as *x* tends to *a* is then written

$$\lim_{x\to a} f(x) = A \tag{1}$$

In the nineteenth century, assertion (1) was provided with a watertight definition: *the function* $y = f(x)$ *is said to have* limit *A* as *x* tends to *a if corresponding to any* $\varepsilon > 0$ *there is* $\delta > 0$ *for which*[1] $|f(x) - A| < \varepsilon$ *whenever* $0 < |x - a| < \delta$. Further, we say that $y = f(x)$ *has a limit* as $x \to a$, or that the limit *exists,* if there is a number *A* for which (1) holds.

Intuitively, the *gradient* of a function at a point is the slope of the tangent—assuming it to be well-defined—at that point to the curve determined by the function. This is given the following precise definition in terms of the limit concept. The function $y = f(x)$ is said to *have a gradient* or to be *differentiable* at $x = x_0$ if the limit as δx tends to 0 of the function

$$\frac{f(x_0 + \delta x) - f(x_0)}{\delta x}$$

exists. This limit, which is called the *gradient* of $y = f(x)$ at $x = x_0$., is often written in the less accurate but more suggestive form

$$lim_{\delta x \to 0}\ \delta y/\delta x. \tag{2}$$

---

[1] If *r* is a real number, we write $|r|$ for the *absolute value* of *r,* that is, $|r| = r$ if *r* is positive, and $|r| = -r$ if *r* is negative.

If f is differentiable at each point of its domain, then its gradient is a function of the argument x, being the function which associates, with each value of x, the slope of the tangent to the curve determined by $f(x)$. This function is called the *derivative* or *differential coefficient* of $f(x)$ with respect to x, and is written variously as $\mathrm{D}f(x)$, $\mathrm{D}y$, $f'(x)$ or $\mathrm{d}y/\mathrm{d}x$, the latter notation being suggested by analogy with the expression in (2) above. The process of obtaining the differential coefficient is called *differentiation.*

In passing we note that a function differentiable at a point is also *continuous* at that point. Here $y = f(x)$ is said to be continuous at $x = a$ if

$$\lim_{x \to a} f(x) = f(a).$$

This means simply that the curve determined by f does not have a "jump" or "gap" at $x = a$.

It is now straightforward to establish the usual laws for differentiation, for example:

$$\mathrm{D}(ax^n) = nax^{n-1},$$

and, if $u(x)$ and $v(x)$ are both functions of x, then

$$\mathrm{D}(uv) = u\mathrm{D}v + v\mathrm{D}u.$$

If we calculate the value of $\delta y/\delta x$ defined above for various functions $y = f(x)$, we find that in every case it consists of two terms, namely (i) the derivative, and (ii) a term of the form Ax_0. Accordingly, we have, for any value of x, an equation of the form

$$\delta y/\delta x = f'(x) + A\ x.$$

If we multiply this equation by x, we obtain

$$\delta y = f'(x)\delta x + A(\delta x)^2. \tag{3}$$

Now if δx is small, $(\delta x)^2$ is considerably smaller, and it follows that $f'(x)\delta x$ is a good approximation to δy. In recent years an approach to the calculus has been developed on the assumption of the presence of quantities δx which, while not actually being equal to zero, are nevertheless so small that their squares $(\delta x)^2$ are *literally* equal to zero. Quantities of this sort are called *square zero* or *nilsquare infinitesimals*: for any such quantity δx equation (3) becomes

$$f(x + \delta x) - f(x) = \ y = f'(x)\delta x.$$

In other words the change in the value of $f(x)$ attendant upon a nilsquare infinitesimal change δx in x is not merely approximately, but *exactly equal to* $f'(x)\delta x$. This enables the derivative $f'(x)$ to be *defined* as that number H which satisfies the equation

$$f(x + \delta x) - f(x) = H\delta x$$

for all nilsquare infinitesimal δx. This approach, which does not involve limits, considerably simplifies the development of the differential calculus: a further account of these ideas is given in Appendix 3.

The derivative of a function $y = f(x)$ of x is itself a function of x and so may itself have a derivative. For example, if $y = x^4$, then $\mathrm{D}y = 4x^3$, so that $\mathrm{D}(\mathrm{D}y) = 12x^2$. The function $\mathrm{D}(\mathrm{D}y)$, in this case, $12x^2$, is called the *second derivative* of y and is written D^2y, $\mathrm{d}^2y/\mathrm{d}x^2$, or $f''(x)$. Similarly, $\mathrm{D}(12x^2)$, that is, $24x$, is called the *third derivative* of $y = x^4$. This process may be iterated indefinitely to yield, for each natural number n, the n^{th} derivative D^ny or $f^{(n)}(y)$ of $y = f(x)$.

Use of the derivative enables various physical concepts to be given precise mathematical formulations. For example, suppose that at time t (seconds) from a given instant, a body of mass m (kilos), moving in a straight line, is at distance x (metres) from a fixed point O on the line. If the velocity of the body is v (metres per second) and the acceleration a (metres per second per second), then

$$v = \mathrm{d}x/\mathrm{d}t \qquad a = \mathrm{d}v/\mathrm{d}t = \mathrm{d}^2x/\mathrm{d}t^2.$$

That is, velocity is the time derivative of distance, and acceleration is the time derivative of velocity, or the second time derivative of distance.

Newton called v the *fluxion* of x and a the *second fluxion* of x; x and v he also called the *fluents* (i.e., "flowing quantities"). His notation for a fluxion was a dot placed over the fluent; thus

$$v = \dot{x} \qquad a = \dot{v} = \ddot{x}$$

By his *second law of motion*, the *force* F in the direction of motion is the time rate of change of the *momentum* mv. Therefore

$$F = \mathrm{d}(mv)/\mathrm{d}t = m\mathrm{d}v/\mathrm{d}t = ma.$$

This equation is the mathematical presentation of Newton's second law.

An important application of the differential calculus is in determining *maximum and minimum values.* The method goes back in essence to Fermat. Suppose we are given a function $y = f(x)$; then the value of its derivative $f'(x_0)$ at x_0 is the slope of the (tangent to the) curve represented by $y = f(x)$ at x_0. If $f'(x_0) = 0$, then the curve is said to have a *stationary point* at x_0. It is obvious that a point at which f assumes a maximum or minimum value must be a stationary point: this fact makes the determination of these values a simple matter, as the following example illustrates.

Suppose that a closed cylindrical can is to be made to contain a given volume V, and we want to determine the shape of the can so that the amount of material used in its manufacture is a minimum. We proceed as follows. Writing r for the radius of the can's base and h for its height, its total surface area is then

$$S = 2\pi rh + 2\pi r^2.$$

Thus the problem is to minimize $2\pi rh + 2\pi r^2$, subject to the constraint $V = \ r^2h$. Here we may simplify the algebra by saying that the problem is to minimize $rh + r^2$ subject to $r^2h = c$. If we write

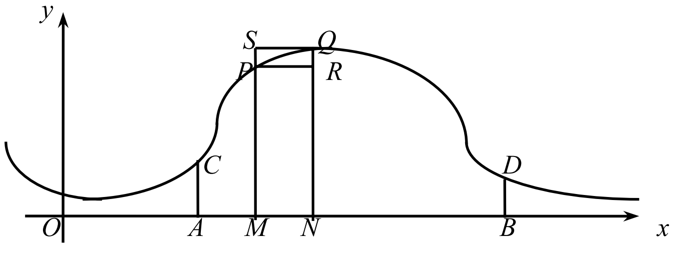
$$f(r) = r^2 + rh = r^2 + c/r,$$

then

$$f'(r) = 2r - c/r^2,$$

so that $f'(r) = 0$ if $2r^3 = c$. Thus a stationary point for $f(r)$ is obtained when $2r^3 = c$, i.e. when $2r^3 = r^2h$, so that $h = 2r$. This value is a minimum because it is plain that $f(r)$ increases indefinitely both when r tends to zero and when r increases indefinitely. We conclude that the surface area, and hence also the amount of material, is a minimum when *the can's height and base diameter coincide.*

We turn now to the problem of *quadrature*, that is, of determining the area enclosed in, or bounded by, a curve. Thus (to follow Newton) let CPD be the graph of the function $y = f(x)$, $OA = a$, $AC = f(a)$, $OM = x$, $MP = f(x)$. Let z denote the area $AMPC$; this area may be thought of as being generated by a line parallel to the y-axis which sets out from the position AC and moves to the right. Thus z is a function of x which is zero when $x = a$. We wish to calculate z, and to do this we find dz/dx. (For simplicity we assume that $f(x)$ is always positive.) When x is increased by the amount δx (= MN), the area z increases by the amount δz (= area $MNQP$). Draw PR, SQ parallel to MN. Then $MNQP$ is equal to the rectangle $MNRP$, together with the area PRQ, which is less than the area of the rectangle $PRQS$; therefore



$$\delta z = MP.\ \delta x + \text{a quantity} < RQ.\ \delta x,$$

so that

$$\delta z/\delta x = MP + \text{a quantity} < RQ.$$

As δx tends to zero, so does RQ, and accordingly

$$dz/dx = MP = f(x). \tag{4}$$

Note incidentally that the use of nilsquare infinitesimals makes this demonstration even simpler, for when δx is nilsquare, the area of PRQ, being proportional to $(\delta x)^2$, is *literally* zero. Therefore $z = MP.\, x$: since this holds for arbitrary x, equation (4) again follows (see Appendix 3).

The area z under the curve $y = f(x)$ is therefore *that function of x whose derivative is* $f(x)$ (and which vanishes when $x = a$). We write

$$z = \int f(x)dx$$

reading the right-hand side of this equation as the (indefinite) *integral of f(x)*; the operation of forming the integral is called *integration.* By definition the operation of differentiation is *inverse* to that of integration, i.e.,

$$d/dx\,[z = \int f(x)dx] = f(x).$$

The theorem that we have just demonstrated, namely, that the derivative of the function representing the area under a curve given by $y = f(x)$ is just $f(x)$ itself, is called the *Fundamental Theorem of the Calculus*: it forges the link between the two types of problem—that of tangents and that of quadrature—with which we began our discussion.

Returning to our figure on the previous page, there is a special notation for the integral that is equal to the area $ABDC$, namely,

$$\int_a^b f(x)dx\,.$$

This is called a *definite integral*; it is computed by finding the indefinite integral of $f(x)$, replacing x by b, then by a, and subtracting the second result from the first.

For example, the area under the curve $y = 4x^3 + 3x^2 + 2x$ between $x = 0$ and $x = 2$ is calculated as follows. Since

$$d/dx[x^4 + x^3 + x^2] = 4x^3 + 3x^2 + 2x,$$

it follows from the definition of the integral that

$$\int (4x^3 + 3x^2 + 2x)\mathrm{d}x = x^4 + x^3 + x^2,$$

so the area in question is

$$\int_0^2 (4x^3 + 3x^2 + 2x)\mathrm{d}x = (4x^3 + 3x^2 + 2x)\mathrm{d}x = (x^4 + x^3 + x^2)_{x=2} - (x^4 + x^3 + x^2)_{x=0}$$

$$= 2^4 + 2^3 + 2^2 - 0 = 28.$$

The interpretation of the area z under a curve will depend on the nature of the quantities represented by x and y. Here are some examples.

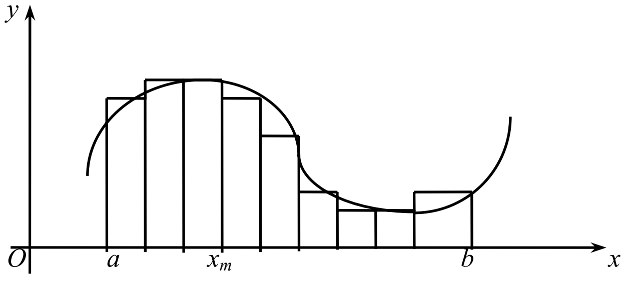
(i) If *OM* or x represents *time* and *MP velocity,* then the area z represents the *gain in distance* during the time represented by *AM.*

(ii) If *OM* represents time and *MP acceleration,* then z represents the *velocity* gained in the time represented by *AM.*

(iii) If *OM* represents the distance a *force* moves its point of application and *MP* the force, then z represents the *work done* by the force in moving its point of application through the distance represented by *AM.*

(iv) If *OM* represents time and *MP* the *temperature* in a room, then z represents the total *energy* expended in heating the room during the time represented[2] by *AM.*

There is an alternative way of arriving at the definite integral, namely as *the limit of a sum*, a method which, in essence at least, goes back to the ancient Greek method of exhaustion but which in its modern form is due to Riemann. Thus, consider once again the curve defined by the function $y = f(x)$ between $x = a$ and $x = b$. We will see first how the area under the curve, which we shall now denote by A, can be approximated by a *sum of rectangles.* We divide the interval from a to b into N small subintervals (x_m, x_{m+1}), of equal width $(b - a)/N$, with $x_0 = a$ and $x_n = b$, and define $\delta x = x_m - x_{m-1} = (b - a)/N$. We choose δx as the common length of the bases of



---

[2] Thus the energy expended in heating a room over a stretch of time depends only on the *average* temperature during that time: when one is out it is cheaper to turn the thermostat down and allow the room to cool than to maintain it at a constant temperature, even though the process of reheating it on one's return may require a disconcertingly steady expenditure of heat energy.

our "approximating rectangles". As their heights we choose the value of $y = f(x)$ at the right-hand endpoint of the subinterval. Thus the sum of the areas of these rectangles will be

$$S_N = f(x_1)\delta x + f(x_2)\delta x + \ldots + f(x_N)\delta x,$$

which we abbreviate as [3]

$$S_N = \sum_{j=1}^{N} f(x_j)\delta x\,.$$

Now we form a sequence of such approximations in which N increases indefinitely, so that the number of terms in each sum increases while, owing to the presence of the factor $\delta x = (b - a)/N$, each single term $f(x_j)\delta x$ tends to zero. By taking N sufficiently large, we can make S_N differ from the area A by an amount as small as we please. This we express briefly by saying that A is the *limit* of the S_N as N tends to infinity. We may then define the definite integral $\int_a^b f(x)\mathrm{d}x = A$ as the *limit of the sums* $S_N = \sum_{j=1}^{N} f(x_j)\delta x$ as *N tends to infinity*.

Leibniz symbolized this passage to the limit from the approximating sum S_N to A by replacing the summation sign $\sum$ —which in his time was written "S"—by the symbol $\int$ (a stylized "S"), and the difference symbol δ by the symbol d. This notation is standard today.

The *natural logarithm* $\log(x)$ is an important example of a function defined by means of integration. It is defined for positive x by

$$\log(x) = \int_1^x \frac{\mathrm{d}t}{t}$$

The number e—like π, one of the fundamental constants of mathematics—is defined to be that number whose logarithm is 1, that is,

$$1 = \int_1^e \frac{\mathrm{d}t}{t}$$

e has the value 2.7182818…. The function e^x, also written $\exp(x)$, is called the *exponential function*; it can be shown that e^x is the function *inverse* to $\log(x)$, that is, $e^{\log(x)} = \log(e^x) = x$. The exponential function has the key property of being equal to its own derivative: $\mathrm{D}(e^x) = e^x$.

---

[3] Here the symbol $\sum_{j=1}^{N}$ signifies the sum of all the expressions obtained by allowing j to assume successively the values 1, 2, 3,.., N.

The definition of the integral as the limit of a sum of plane areas can be extended to three (and higher) dimensions so as to enable the differential calculus to be employed in determining areas of curved surfaces and volumes of solids. This definition also provides the basis for the general theories of integration which have been developed in the present century.

MATHEMATICAL ANALYSIS

Within both the differential and integral calculus the limit concept plays a central role: the derivative of a function is the limit of a quotient, and the definite integral is the limit of a sum. A characteristic feature of the limit concept is that it involves the idea of an *infinite process*: in forming a limit as x tends to zero, say, we imagine x becoming smaller and smaller endlessly without ever actually vanishing. The use of infinite processes is the hallmark of the extensive area of mathematics known as *mathematical analysis*, or *analysis* for short. Analysis may be considered, along with geometry and algebra (here taken to include number theory), as one of the three main divisions of mathematics. Analysis is, perhaps, the largest of these divisions, and includes many topics besides the calculus; only a few of these will be touched on here. We defer until Chapter 11 our discussion of *set theory*, the branch of mathematics concerned with the infinite in its purest form, and which may be considered to have emerged from analysis.

Infinite Series

Let a_0, a_1, a_2, … be a sequence of real numbers. We say that the sequence—which we shall denote by $\langle a_n \rangle$—*converges to limit a as n tends to infinity*—written $a_n \to a$ as[4] $n \to \infty$—if the differences between a_n and a become arbitrarily small as n becomes arbitrarily large. To be precise, $\langle a_n \rangle$ converges to limit a if, corresponding to any $\varepsilon > 0$ there is an integer N such that $|a_n - a| < \varepsilon$ for all $n \geq N$. In the opposite event, the sequence is said to *diverge.*

Given a sequence $\langle a_n \rangle$, the sequence $\langle S_n \rangle$ defined by

$$S_n = \sum_{j=1}^{n} a_j = a_1 + a_2 + \ldots + a_n$$

is called the sequence of partial sums of the *infinite series*

$$\sum_{n=1}^{\infty} a_n ,$$

---

[4] The sign "∞" to denote an infinite number was introduced by Wallis in his work *De Sectionibus Conicis* of 1655. It has been conjectured that Wallis, who was a classical scholar, adopted this sign from the late Roman symbol "∞" (possibly a form of "M") for 1000.

sometimes written simply

$$\sum a_n .$$

If $S_n \to s$ as $n \to \infty$, the series is said to *converge* to the *sum s,* and we write

$$s = \sum_{n=1}^{\infty} a_n = a_1 + a_2 + \ldots + a_n + \ldots .$$

In the opposite event the series is said to *diverge.*

The most familiar type of convergent series is the *geometric series*

$$\sum_{n=1}^{\infty} x^n$$

with fixed x satisfying $|x| < 1$. In this case the partial sums satisfy

$$S_n = \sum_{j=1}^{n} x^j = 1 + x + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}.$$

If $|x| < 1$, then $x^{n+1} \to 0$ as $n \to \infty$, so that

$$S_n \to 1/(1 - x) \quad \text{as } n \to \infty.$$

Accordingly $\sum_{n=1}^{\infty} x^n$ converges with sum $1/(1 - x)$ provided $|x| < 1$.

The ancient Greeks were conversant with the idea that an infinite series could converge, even if they lacked a precise definition of the concept. In particular they had grasped the fact that

$$\tfrac{1}{2} + \tfrac{1}{4} + \tfrac{1}{8} + \ldots + \tfrac{1}{2^n} + \ldots = 1.$$

In Chapter 6 of his *Physics* Aristotle observes, in effect, that such a series should have a sum.

Infinite series make an occasional appearance in medieval mathematics. *Nicolas Oresme* (c. 1323–1382), for example, observed that the so-called *harmonic series*

$$\tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{4} + \ldots \tfrac{1}{n} + \ldots$$

is *divergent.* For it can be replaced by the series of lesser terms

$$(\tfrac{1}{4} + \tfrac{1}{4}) + (\tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8} + \tfrac{1}{8}) + \ldots,$$

which yields as many terms of magnitude ½ as one pleases.

In 1593 Viète derived a formula for the sum of an infinite geometric series, and in 1647 *Gregory of St. Vincent* (1584–1687), in his *Opus Geometricum*, is the first to state explicitly that a convergent infinite series represents a magnitude, which he calls the limit of the series.

The principal use of infinite series has been to represent *functions*, and thereby to provide a means for calculating their values. In the mid-seventeenth century, for instance, Newton obtained a series for the natural logarithm

$$\log(1+x) = x - \tfrac{1}{2}x^2 + \tfrac{1}{3}x^3 - \ldots,$$

which converges for $|x| \leq 1$. This is a typical example of what is known as a *power series.* Newton also obtained the power series

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots, \tag{1}$$

so that in particular

$$e = 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \ldots,$$

as well as power series for the trigonometric functions[5]

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots \tag{2}$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots . \tag{3}$$

An important general method, still very much in use today, for representing a function as a power series was developed by *Brook Taylor* (1685–1731) in his *Methodus Incrementum Directa et Inversa* of 1715. Building on the work of Newton and James Gregory, Taylor obtains for the function $y = f(x)$ the equation

$$f(x+h) = f(x) + hf'(x) + \frac{h^2}{2!}f''(x) + \frac{h^3}{3!}f'''(x) + \ldots$$

Putting $x = 0$ yields the equation

---

[5] Here $\sin x$ and $\cos x$ are given as functions of a variable angle x measured in *radians*, where a radian is the angle subtended at the centre of a circle by an arc equal to the radius. Thus π radians = 180°.

$$f(h) = f(0) + hf'(0) + \frac{h^2}{2!}f''(0) + \frac{h^3}{3!}f'''(0) + \ldots$$

This is known as *Maclaurin's theorem,* after *Colin Maclaurin* (1698–1746). In obtaining their series neither Taylor nor Maclaurin employ rigorous argument, nor do they consider the question of convergence. Maclaurin derives his series by means of the method of *undetermined coefficients.* He begins by assuming that $f(x)$ can be expanded as a series

$$f(x) = A + Bx + Cx^2 + Dx^3 + \ldots .$$

Then by successive term-by-term differentiation he obtains

$$f'(x) = B + 2Cx + 3Dx^2 + \ldots .$$

$$f''(x) = 2C + 6Dx + \ldots .$$

Putting $x = 0$ in each equation then determines $A, B, C, D, \ldots$.

Euler was the master of infinite series. These he would manipulate with great flair, only rarely paying attention to the question of convergence. In his *Introductio in Analysin Infinitorum* of 1748 are to be found many examples of his skill in this regard. One of his most famous results obtained by the formal manipulation of infinite series relates the exponential and logarithmic functions through complex numbers and is known as the *Euler identity*, namely,

$$e^{ix} = \cos x + i \sin x. \tag{4}$$

This memorable equation may be derived by noting that, if we substitute ix for x in the series expansion (1) for e^x and use the series expansions (2) and (3) for sin x and cos x, we get

$$\begin{aligned} e^{ix} &= 1 + ix - \frac{x^2}{2!} - \frac{ix^3}{3!} + \frac{x^4}{4!} + \frac{ix^5}{5!} + \ldots \\ &= (1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots) + i(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots) \\ &= \cos x + i \sin x. \end{aligned}$$

If we substitute π for x in (4), and note that $\cos \pi = -1$, $\sin \pi = 0$, we obtain one of the most remarkable equations in the whole of mathematics, namely,

$$e^{i\pi} + 1 = 0.$$

Of course, equations such as this only become meaningful when quantities such as $e^{i\pi}$ are properly defined: this was not carried out until the nineteenth century.

Using his freewheeling methods, Euler was able to achieve results which baffled his contemporaries. One of his most remarkable feats was the summing of the series $\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\ldots$. To effect this, Euler began with the equation (2) for sin x; the equation $\sin x = 0$ can then be thought of—after dividing through by x—as the infinite polynomial equation

$$0 = 1-\frac{x^2}{3!}+\frac{x^4}{5!}-\ldots$$

or, replacing x^2 by y,

$$0 = 1-\frac{y}{3!}+\frac{y^2}{5!}-\ldots$$

From the theory of algebraic equations it is known (and not difficult to show) that if the constant term in a polynomial equation is 1, then the sum of the reciprocals of the roots is the negative of the coefficient of the linear term, which in this case is $\frac{1}{3!}$. Moreover, the roots of the equation $\sin x = 0$ are known to be π, 2π, 3π,…; accordingly, the roots of the equation in y are π^2, $(2\pi)^2$, $(3\pi)^2$, … . Therefore

$$\frac{1}{6}=\frac{1}{\pi^2}+\frac{1}{(2\pi)^2}+\frac{1}{(3\pi)^2}+\ldots$$

or

$$\frac{\pi^2}{6}=\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\ldots \tag{5}$$

Using the cosine series (4) in place of the sine series, Euler similarly obtained the result

$$\frac{\pi^2}{8}=\frac{1}{1^2}+\frac{1}{3^2}+\frac{1}{5^2}+\ldots \tag{6}$$

Subtracting equation (5) from twice equation (6) gives

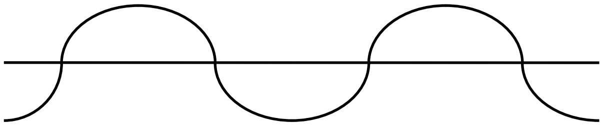
$$\frac{\pi^2}{12}=\frac{1}{1^2}-\frac{1}{2^2}+\frac{1}{3^2}-\frac{1}{4^2}+\ldots$$

These arguments, while hardly rigorous, are nothing less than spectacular. In the nineteenth century, the great majority of Euler's results were derived by rigorous, but less entertaining methods.

An important class of infinite series are the so-called *trigonometric series*, which had first arisen in connection with the analysis of planetary orbits. By a trigonometric series is meant any series of the form

$$\tfrac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos nx + b_n \sin nx)\,.$$

Each of the sine and cosine functions in this series has a wave-like or "sinusoidal" graph of the form



Thus the function represented by such a series is a "superposition" of such waves.

Trigonometric series were investigated in the eighteenth century by Euler, *Jean–le–Rond d'Alembert* (1717–1783), Clairaut, and Lagrange, but it was *Joseph Fourier* (1768–1831)—in his papers of 1807, 1811, and above all in his celebrated work *Théorie Analytique de la Chaleur* of 1822—who first advanced the claim that an *arbitrary* function $y = f(x)$ can be represented as a trigonometric series, thus introducing the branch of mathematics that has become known as *harmonic analysis.* Fourier obtains the expressions (in which he was anticipated by Euler)

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\, \mathrm{d}x \qquad \mathrm{b}_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\, \mathrm{d}x$$

for the coefficients of the trigonometric series expansion of $f(x)$. Fourier's work was so influential that the trigonometric series representing a function is called its *Fourier expansion* and the coefficients of this series its *Fourier coefficients.*

Fourier did not supply a rigorous proof that the trigonometric series associated with a function converges to the value of the function, and later work showed that for certain functions this was not in fact the case. It was Dirichlet who, in 1828, gave the first rigorous proof of convergence of the Fourier expansion of a function subject to certain restrictions. He showed that, if $y = f(x)$ is continuous at each point of its domain, periodic of period 2π, that is, $f(x + 2\pi) = f(x)$, and $\int_{-\pi}^{\pi} f(x)\mathrm{d}x$ is finite, then the Fourier expansion of f always converges to $f(x)$.

The mathematicians of the eighteenth century had been somewhat cavalier in their use of infinite series, and by the century's end certain dubious or downright absurd results obtained by their means provoked enquiries as to the validity of operating with them. In the forefront of these investigations was Cauchy, who in his *Cours d'Analyse* of 1821 gave the rigorous definition of convergence of an infinite series employed today, and who also formulated the purely internal criterion for convergence which has come to bear his name (although it was known earlier to Bolzano). This states that a necessary and sufficient condition that an infinite series $\sum a_n$ converge is that, for any given value of p, the difference between the partial sums S_n and S_{n+p} tends to zero as n tends to infinity.

While the injection of rigor into the theory of infinite series administered in the nineteenth century may have tamed the subject, it has remained colourful. The resulting precision also made possible the discovery of new and striking facts about infinite series. *Riemann's theorem* is an example. This stems from the observation that certain series containing positive and negative terms exist with the property that, while they converge, the series obtained by making all their terms positive do not. For instance,

$$1-\tfrac{1}{2}+\tfrac{1}{3}-\tfrac{1}{4}+\ldots$$

converges to log 2, but, as we have seen, the corresponding series of positive terms

$$1+\tfrac{1}{2}+\tfrac{1}{3}+\tfrac{1}{4}+\ldots$$

diverges to infinity. A series of this type is called *conditionally convergent.* Riemann's theorem asserts that, given any conditionally convergent series, its terms can be rearranged so that the resulting series converges to *any value whatever.* Thus, for example, it can be shown that

$$1+\tfrac{1}{3}-\tfrac{1}{2}+\tfrac{1}{5}+\tfrac{1}{7}-\tfrac{1}{4}+\ldots=\tfrac{3}{2}\log(2)$$

and

$$1-\tfrac{1}{2}-\tfrac{1}{4}+\tfrac{1}{3}-\tfrac{1}{6}+\tfrac{1}{8}+\ldots=\tfrac{1}{2}\log(2)\,.$$

Differential Equations

In the seventeenth and eighteenth centuries mathematicians applied the calculus to a broader and broader range of physical problems. The mathematical formulation of many of these problems involved what came to be known as *differential equations,* that is, equations involving derivatives of a function or functions, from which the functions themselves are to be determined.

A general problem of this sort is the following. Consider a material particle of mass m moving along a line, and let x denote its distance at time t from some given fixed point on the line. Let us assume that the motion is caused by some force F, the value of which depends on x, on the velocity $\mathrm{d}x/\mathrm{d}t$, and on the time t; to indicate this dependence we write $F = F(x, \mathrm{d}x/\mathrm{d}t, t)$. According to Newton's second law of motion, the action of the force F on the particle produces an acceleration $a = \mathrm{d}^2x/\mathrm{d}t^2$ with $F = ma$, and so we obtain the differential equation

$$M\mathrm{d}^2x/\mathrm{d}t^2 = F(x, \mathrm{d}x/\mathrm{d}t, t). \qquad (1)$$

More specifically, let us suppose that the particle is moving in a resisting medium, for example in a liquid or a gas, under the influence of the elastic force of two springs, acting in accordance with *Hooke's law* (after *Robert Hooke*, 1635–1703), which states that the elastic force of a spring acts towards the equilibrium position and is proportional to the deviation therefrom. Assuming that the equilibrium position is at $x = 0$, the elastic force is then proportional to $-bx$, where b is some positive constant. We also suppose that the resistance of the medium is proportional to the speed of the motion, that is, equal to $-a\mathrm{d}x/\mathrm{d}t$, where a is some positive constant. In that case the force F acting on the particle is $-bx - a\mathrm{d}x/\mathrm{d}t$ and so equation (1) assumes the form

$$m\,\mathrm{d}^2x/\mathrm{d}t^2 + a\,\mathrm{d}x/\mathrm{d}t + bx = 0.$$

This is the *differential equation of motion* of the particle. Its solution $x = x(t)$ gives the position x of the particle as a function of the time t.

Differential equations also arise in the determination of the *form of curves*. For example, in 1690 Jacques Bernoulli posed the problem of finding the curve assumed by a flexible inextensible cord hanging freely from two fixed points, the curve Leibniz named the *catenary.* This problem had been considered as far back as the fifteenth century by Leonardo da Vinci; Galileo mistakenly thought that the curve was a parabola. Huygens showed that this was not correct and that the curve would assume a parabolic form only when the combined weight of the cord and any objects suspended from it is uniform per horizontal run, as in a suspension bridge. In 1691 Leibniz, Huygens and Jean Bernoulli published independent solutions to the problem, the latter employing the methods of the calculus. Writing $y = f(x)$ for the equation of the curve, Bernoulli obtains the differential equation

$$1 + (\mathrm{d}y/\mathrm{d}x)^2 = a^2y,$$

and gives the correct solution

$$y = \tfrac{1}{2}a(e^{x/a} - e^{-x/a}).$$

Perhaps the most celebrated solution to a set of differential equations is Newton's derivation of Kepler's laws of planetary motion. Formulated in 1609 and 1619 by *Johannes Kepler* (1571–1630), these laws are the following three assertions:

1. *The planets move about the sun in elliptical orbits with the sun at one focus*[6].

---

[6] An ellipse may be defined as the locus of all points the sum of whose distances from two given points is constant: these two points are the *foci* of the ellipse.

2. *The radius vector joining a planet to the sun sweeps over equal areas in equal intervals of time.*

3. *The square of the time of one complete revolution of a planet about its orbit is proportional to the cube of the orbit's semimajor axis.*

In his *Philosophiae Naturalis Principia Mathematica*—universally known as the *Principia*—of 1687, Newton formulates his laws of motion and the inverse square law of gravitational attraction between two bodies, and from them deduces Kepler's laws, at the same time creating the science of *celestial mechanics.* Although Newton casts his work in the form of classical Greek geometry, it is almost certain that he obtained his results in the first instance by the use of the calculus, which means, in essence, that he solved the differential equations of motion of a body moving under a central force.

Partial differential equations are also frequently encountered in physics. These are equations involving *partial derivatives.* If $f(x, y, z, \ldots)$ is a function of the variables $x, y, z, \ldots,$ the partial derivative $\frac{\partial f}{\partial x}$ of f with respect to one of these variables, $x,$ say, is obtained by fixing the values of the remaining variables $y, z, \ldots,$ and then differentiating the resulting function of $x.$ Similarly one obtains higher partial derivatives $\frac{\partial^2 f}{\partial x^2}$ and mixed partial derivatives $\frac{\partial^2 f}{\partial x \partial y}$. It is a fundamental property of partial differentiation that the order in which the differentiations are made is irrelevant: thus, for example, $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$.

One of the first, and undoubtedly most important, partial differential equations—the *wave equation*—made its appearance in the eighteenth century in connection with the analysis of a vibrating string, a violin string, for example. In 1746 d'Alembert gave the correct form of the equation governing a flexible continuous vibrating string:

$$T\frac{\partial^2 u}{\partial x^2} = \rho\frac{\partial^2 u}{\partial t^2} \qquad (2)$$

where $u = u(x, t)$ is the displacement of the string from its equilibrium position at time t and position x along it, T is the tension in the string and ρ is its density. D'Alembert shows that (2) has a solution of the form

$$u(x, t) = \varphi(ct + x) + \psi(ct - x),$$

where $c = \sqrt{T/\rho}$ and φ, ψ are arbitrary twice differentiable functions. This solution may be considered as the superposition of two "standing waves", each moving in opposite directions along the string.

In 1748 Euler gave a solution to the wave equation in the form of a trigonometric series

$$u(x,\, t) = \sum A_n \sin n\pi x/\ell \ \cos n\pi ct/\ell.$$

This solution presents the characteristic vibrations of the string in terms of a superposition of infinitely many elementary wave motions of sinusoidal form.

Later the wave equation was extended to three spatial dimensions, taking the form

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = c^2 \frac{\partial^2 u}{\partial t^2}.$$

This equation, one of the most important in mathematical physics, governs the vibrations of solid bodies and the propagation of waves such as sound and light.

Another important partial differential equation in mathematical physics is the *potential equation,* also known as *Laplace's equation,* which takes the form

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} + \frac{\partial^2 V}{\partial z^2} = 0.$$

This equation, which was given in 1789 by *Pierre Simon de Laplace* (1749–1827), governs the *gravitational potential* $V(x, y, z)$ outside a gravitating body. The potential function $V(x, y, z)$ is characterized by the property that its partial derivatives with respect to the three spatial coordinates x, y and z are the three components of the gravitational force exerted by the body at the point (x, y, z).

Other important partial differential equations arise in connection with *fluid flow.* In a paper of 1755 entitled *General Principles of the Motion of Fluids,* Euler derives his famous equations of flow for perfect nonviscous fluids. If we write u, v, w for the components of the velocity of the fluid at time t in the x, y and z directions, p for the pressure and ρ for the density of the fluid, the first of Euler's equations for compressible flow takes the form

$$\frac{\partial p}{\partial x} = -\rho\left(u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} + \frac{\partial u}{\partial t} \right)$$

with analogous expressions for $\frac{\partial p}{\partial y}$ and $\frac{\partial p}{\partial z}$. Euler also obtains a general form of the so-called *equation of continuity* which had been derived previously by d'Alembert for incompressible flow, and which expresses the fact that matter is neither created nor destroyed in fluid motion. Euler's version of the equation is:

$$\frac{\partial p}{\partial t} + \frac{\partial(\rho u)}{\partial x} + \frac{\partial(\rho v)}{\partial y} + \frac{\partial(\rho w)}{\partial z} = 0.$$

For incompressible flow ρ is constant and we obtain d'Alembert's form of the equation

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0.$$

If we follow Euler by introducing a function S, called the *velocity potential*, whose partial derivatives with respect to x, y, and z are u, v and w, respectively, we see that S satisfies Laplace's equation. This equation is also characteristic of fluid flow.

Finally, we must mention the *heat equation* which was obtained by Fourier in his *Théorie Analytique de la Chaleur*. By physical arguments, Fourier showed that the temperature T at a point (x, y, z) at time t in a homogeneous and isotropic body satisfies the equation

$$\frac{\partial^2 T}{\partial x^2}+\frac{\partial^2 T}{\partial y^2}+\frac{\partial^2 T}{\partial z^2}=k^2\frac{\partial T}{\partial t},$$

where k is a constant whose value depends on the material of the body. When T is independent of the time, $\frac{\partial T}{\partial t} = 0$ and the equation reduces to Laplace's equation. In one dimension Fourier gave a solution to the equation in the form of a trigonometric series with exponential coefficients

$$T=\sum b_n \exp(-n^2\pi^2/k^2\,\ell^2)t\sin n\pi x/\ell.$$

The presence of the exponential terms here reflects the fact that, as a body cools, its temperature falls off as e^{-kt} with the time t.

Complex Analysis

One of the most significant and original mathematical creations of the nineteenth century was the *theory of functions of a complex variable,* or *complex analysis* for short. Now, strictly speaking, a function of a complex variable is just a function f: $\mathbb{C} \to \mathbb{C}$, so that, for each $z = x + iy$, $f(z)$ is a complex number whose real and imaginary parts are uniquely determined by x and y. In other words, there are functions $u(x, y)$ and $v(x, y)$ of two real variables x and y such that $f(x + iy) = u(x, y) + iv(x, y)$. Thus, according to this definition, a complex function is essentially nothing more than a pair of real functions.

Both Cauchy and Riemann independently saw that this definition, while perfectly legitimate, was too broad to be of much practical value. They proposed instead that attention be confined to complex functions which are *differentiable* in the same sense as real functions. Thus a function f: $\mathbf{D} \to \mathbb{C}$, where $\mathbf{D}$ is a region in the complex plane is said to be *differentiable* at a point z_0 of $\mathbf{D}$ if

$$\frac{f(z)-f(z_0)}{z-z_0} \tag{1}$$

tends to a unique limit as $z \to z_0$. A function which is differentiable in this sense at every point of a region **D** is said to be *analytic* (also: *regular* or *holomorphic*) on **D**.

Since analyticity requires that (1) has a unique limit when $z - z_0$ tends to zero not merely through purely real or purely imaginary values, but *along any path whatsoever*, it is evidently a very strong condition. Cauchy and Riemann independently discovered that the real and imaginary parts $u(x, y)$ and $v(x, y)$ of an analytic function satisfy what are now known as the *Cauchy–Riemann equations*, viz.,

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

And conversely, if u and v are continuous and have continuous partial derivatives satisfying the Cauchy-Riemann equations throughout a region, then $f = u + \mathrm{i}v$ is analytic in that region.

From the Cauchy-Riemann equations we deduce, by partial differentiation,

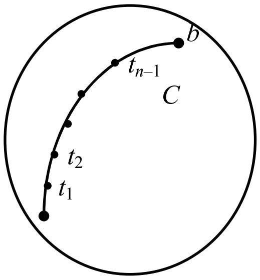
$$\frac{\partial^2 v}{\partial x \partial y} = \frac{\partial^2 u}{\partial x^2} = -\frac{\partial^2 u}{\partial y^2} \qquad \frac{\partial^2 u}{\partial x \partial y} = -\frac{\partial^2 v}{\partial x^2} = -\frac{\partial^2 v}{\partial y^2}.$$

Accordingly both u and v satisfy the two-dimensional version of Laplace's equation, that is

$$\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2} = 0.$$

It follows that, by separating any analytic function into its real and imaginary parts, two solutions of Laplace's equation are immediately obtained. It is this fact that makes complex analysis so useful in the solution of problems in mathematical physics, especially in the theory of fluid flow.

Cauchy proved what is perhaps the single most important result of complex analysis, now known as the *Cauchy integral theorem.* Given a curve C in a region of the complex plane, the *integral* of an analytic function f along C may be defined in a manner analogous to that of the definite integral of a real function. Namely, if *a, b* are



a

the endpoints of the curve, we subdivide it into n segments by means of the successive points $a = t_0, t_1, t_2, \ldots, t_n = b$ and form the sum

$$S_n = \sum_{i=1}^{n} f(t_i^*)(t_i - t_{i-1}),$$

where t_i^* denotes any point on C lying between t_{i-1} and t_i. If we now make the subdivision finer and finer by allowing the number of points to increase without limit in such a way that the greatest of the distances $|t_i - t_{i-1}|$ tends to zero, then S_n tends to a limit which is independent of the choice of the points t_i and t_i^*. This limit is the integral of $f(z)$ along C and is written

$$\int_C f(z)\mathrm{d}z.$$

Now the Cauchy integral theorem asserts that, if $f(z)$ is analytic in a simply connected region **D** of the complex plane, and C, C' are any two simple curves in **D** with the same end points, then

$$\int_C f(z)\mathrm{d}z = \int_{C'} f(z)\mathrm{d}z.$$

That is, *the integral is independent of the path.* In particular, if C is a closed curve, then

$$\int_C f(z)\mathrm{d}z = 0.$$

Cauchy's theorem has many consequences, one of the most important of which we mention in conclusion. This is that every analytic function can be expanded as a *power series*. Precisely, if $f(z)$ is analytic on and inside a simple closed curve C, then, for any given point a inside C, there are complex numbers $b_1, b_2, \ldots$ for which

$$f(z) = f(a) + b_1(z - a) + b_2(z - a)^2 + \ldots$$

whenever z is inside C. Notice that, since the coefficients $b_1, b_2, \ldots$ depend only on a, a function analytic in a region can be represented as a single power series, not just locally, but throughout the entire region. This shows once again what a powerful condition analyticity is.