Basic Math for competitive programing

Duy Huynh 2016

Contents

Chapter 1

Number theory

Not to be confused with Numerology.

Number theory or, in older usage,[note 1] **arithmetic** is a branch of pure mathematics devoted primarily to the study of the integers. It is sometimes called "The Queen of Mathematics" because of its foundational place in the discipline.[1] Number theorists study prime numbers as well as the properties of objects made out of integers (e.g., rational numbers) or defined as generalizations of the integers (e.g., algebraic integers).

Integers can be considered either in themselves or as solutions to equations (Diophantine geometry). Questions in number theory are often best understood through the study of analytical objects (e.g., the Riemann zeta function) that encode properties of the integers, primes or other number-theoretic objects in some fashion (analytic number theory). One may also study real numbers in relation to rational numbers, e.g., as approximated by the latter (Diophantine approximation).

The older term for number theory is *arithmetic*. By the early twentieth century, it had been superseded by "number theory".[note 2] (The word "arithmetic" is used by the general public to mean "elementary calculations"; it has also acquired other meanings in mathematical logic, as in *Peano arithmetic*, and computer science, as in *floating point arithmetic*.) The use of the term *arithmetic* for *number theory* regained some ground in the second half of the 20th century, arguably in part due to French influence.[note 3] In particular, *arithmetical* is preferred as an adjective to *number-theoretic*.

1.1 History

1.1.1 Origins

Dawn of arithmetic

The first historical find of an arithmetical nature is a fragment of a table: the broken clay tablet Plimpton 322 (Larsa, Mesopotamia, ca. 1800 BCE) contains a list of "Pythagorean triples", i.e., integers (a, b, c) such that $a^2 + b^2 = c^2$. The triples are too many and too large to have been obtained by brute force. The heading over the first column reads: "The *takiltum* of the diagonal which has been subtracted such that the width..."[2]

The table's layout suggests[3] that it was constructed by means of what amounts, in modern language, to the identity

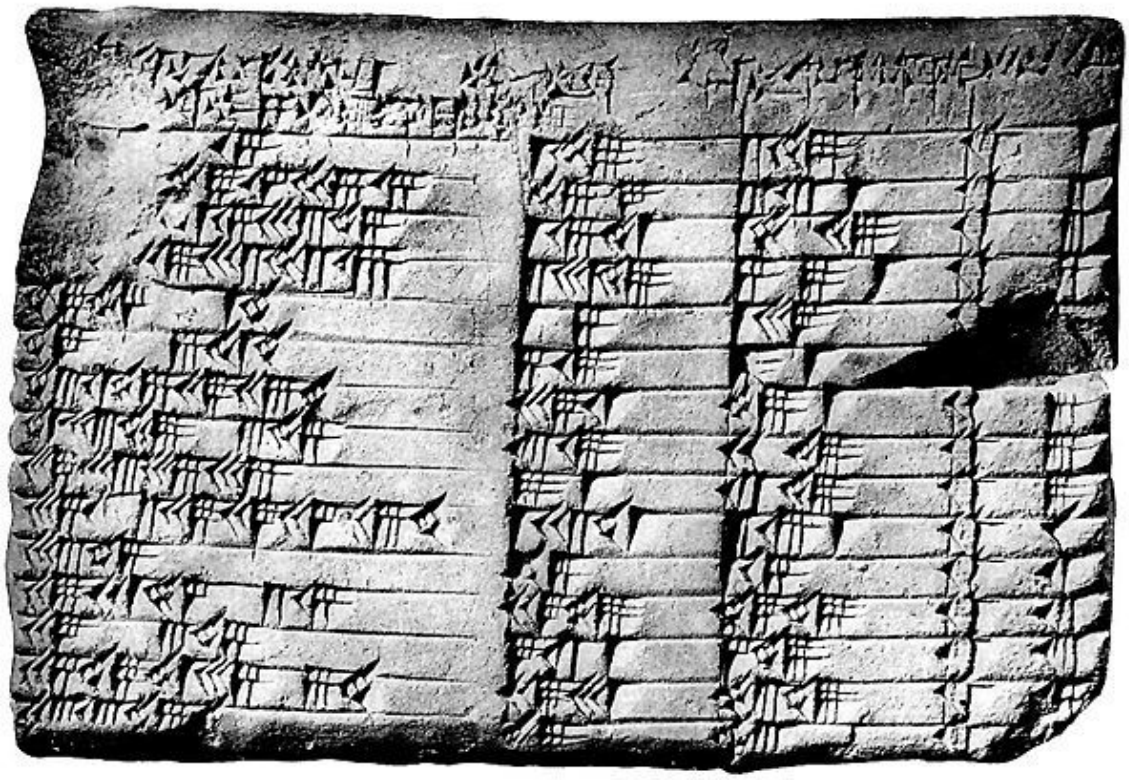
$$\left(\frac{1}{2}\left(x - \frac{1}{x}\right)\right)^2 + 1 = \left(\frac{1}{2}\left(x + \frac{1}{x}\right)\right)^2,$$

which is implicit in routine Old Babylonian exercises.[4] If some other method was used,[5] the triples were first constructed and then reordered by c/a , presumably for actual use as a "table", i.e., with a view to applications.

It is not known what these applications may have been, or whether there could have been any; Babylonian astronomy, for example, truly flowered only later. It has been suggested instead that the table was a source of numerical examples for school problems.[6][note 4]

While Babylonian number theory—or what survives of Babylonian mathematics that can be called thus—consists of

A Lehmer sieve, which is a primitive digital computer once used for finding primes and solving simple Diophantine equations.



The Plimpton 322 tablet

this single, striking fragment, Babylonian algebra (in the secondary-school sense of "algebra") was exceptionally well developed.[7] Late Neoplatonic sources[8] state that Pythagoras learned mathematics from the Babylonians. Much earlier sources[9] state that Thales and Pythagoras traveled and studied in Egypt.

Euclid IX 21—34 is very probably Pythagorean;[10] it is very simple material ("odd times even is even", "if an odd number measures [= divides] an even number, then it also measures [= divides] half of it"), but it is all that is needed to prove that $\sqrt{2}$ is irrational.[11] Pythagorean mystics gave great importance to the odd and the even.[12] The discovery that $\sqrt{2}$ is irrational is credited to the early Pythagoreans (pre-Theodorus).[13] By revealing (in modern terms) that numbers could be irrational, this discovery seems to have provoked the first foundational crisis in mathematical history; its proof or its divulgation are sometimes credited to Hippasus, who was expelled or split from the Pythagorean sect.[14] This forced a distinction between *numbers* (integers and the rationals—the subjects of arithmetic), on the one hand, and *lengths* and *proportions* (which we would identify with real numbers, whether rational or not), on the other hand.

The Pythagorean tradition spoke also of so-called polygonal or figurate numbers.[15] While square numbers, cubic numbers, etc., are seen now as more natural than triangular numbers, pentagonal numbers, etc., the study of the sums of triangular and pentagonal numbers would prove fruitful in the early modern period (17th to early 19th century).

We know of no clearly arithmetical material in ancient Egyptian or Vedic sources, though there is some algebra in both. The Chinese remainder theorem appears as an exercise [16] in Sun Zi's *Suan Ching*, also known as *The Mathematical Classic of Sun Zi* (3rd, 4th or 5th century CE.)[17] (There is one important step glossed over in Sun Zi's solution:[note 5] it is the problem that was later solved by Āryabhaṭa's Kuṭṭaka – see below.)

There is also some numerical mysticism in Chinese mathematics,[note 6] but, unlike that of the Pythagoreans, it seems to have led nowhere. Like the Pythagoreans' perfect numbers, magic squares have passed from superstition into recreation.

Classical Greece and the early Hellenistic period

Further information: Ancient Greek mathematics

Aside from a few fragments, the mathematics of Classical Greece is known to us either through the reports of contemporary non-mathematicians or through mathematical works from the early Hellenistic period.[18] In the case of number theory, this means, by and large, *Plato* and *Euclid*, respectively.

Plato had a keen interest in mathematics, and distinguished clearly between arithmetic and calculation. (By *arithmetic* he meant, in part, theorising on number, rather than what *arithmetic* or *number theory* have come to mean.) It is through one of Plato's dialogues—namely, *Theaetetus*—that we know that Theodorus had proven that $\sqrt{3}, \sqrt{5}, \ldots, \sqrt{17}$ are irrational. Theaetetus was, like Plato, a disciple of Theodorus's; he worked on distinguishing different kinds of incommensurables, and was thus arguably a pioneer in the study of number systems. (Book X of Euclid's Elements is described by Pappus as being largely based on Theaetetus's work.)

Euclid devoted part of his *Elements* to prime numbers and divisibility, topics that belong unambiguously to number theory and are basic to it (Books VII to IX of Euclid's Elements). In particular, he gave an algorithm for computing the greatest common divisor of two numbers (the Euclidean algorithm; *Elements*, Prop. VII.2) and the first known proof of the infinitude of primes (*Elements*, Prop. IX.20).

In 1773, Lessing published an epigram he had found in a manuscript during his work as a librarian; it claimed to be a letter sent by Archimedes to Eratosthenes.[19][20] The epigram proposed what has become known as Archimedes' cattle problem; its solution (absent from the manuscript) requires solving an indeterminate quadratic equation (which reduces to what would later be misnamed Pell's equation). As far as we know, such equations were first successfully treated by the Indian school. It is not known whether Archimedes himself had a method of solution.

Diophantus

Very little is known about Diophantus of Alexandria; he probably lived in the third century CE, that is, about five hundred years after Euclid. Six out of the thirteen books of Diophantus's *Arithmetica* survive in the original Greek; four more books survive in an Arabic translation. The *Arithmetica* is a collection of worked-out problems where the task is invariably to find rational solutions to a system of polynomial equations, usually of the form $f(x, y) = z^2$ or $f(x, y, z) = w^2$. Thus, nowadays, we speak of *Diophantine equations* when we speak of polynomial equations to which rational or integer solutions must be found.

One may say that Diophantus was studying rational points — i.e., points whose coordinates are rational — on curves and algebraic varieties; however, unlike the Greeks of the Classical period, who did what we would now call basic algebra in geometrical terms, Diophantus did what we would now call basic algebraic geometry in purely algebraic terms. In modern language, what Diophantus did was to find rational parametrizations of varieties; that is, given an equation of the form (say) $f(x_1, x_2, x_3) = 0$, his aim was to find (in essence) three rational functions g_1, g_2, g_3 such that, for all values of r and s , setting $x_i = g_i(r, s)$ for $i = 1, 2, 3$ gives a solution to $f(x_1, x_2, x_3) = 0$.

Diophantus also studied the equations of some non-rational curves, for which no rational parametrisation is possible. He managed to find some rational points on these curves (elliptic curves, as it happens, in what seems to be their first known occurrence) by means of what amounts to a tangent construction: translated into coordinate geometry (which did not exist in Diophantus's time), his method would be visualised as drawing a tangent to a curve at a known rational point, and then finding the other point of intersection of the tangent with the curve; that other point is a new rational point. (Diophantus also resorted to what could be called a special case of a secant construction.)

While Diophantus was concerned largely with rational solutions, he assumed some results on integer numbers, in particular that every integer is the sum of four squares (though he never stated as much explicitly).

Āryabhaṭa, Brahmagupta, Bhāskara

While Greek astronomy probably influenced Indian learning, to the point of introducing trigonometry,[21] it seems to be the case that Indian mathematics is otherwise an indigenous tradition;[22] in particular, there is no evidence that Euclid's Elements reached India before the 18th century.[23]

Āryabhaṭa (476–550 CE) showed that pairs of simultaneous congruences $n \equiv a_1 \bmod m_1$, $n \equiv a_2 \bmod m_2$ could be solved by a method he called *kuṭṭaka*, or *pulveriser*;[24] this is a procedure close to (a generalisation of) the Euclidean algorithm, which was probably discovered independently in India.[25] Āryabhaṭa seems to have had in mind applications to astronomical calculations.[21]

Brahmagupta (628 CE) started the systematic study of indefinite quadratic equations—in particular, the misnamed Pell equation, in which Archimedes may have first been interested, and which did not start to be solved in the West until the time of Fermat and Euler. Later Sanskrit authors would follow, using Brahmagupta's technical terminology.

A general procedure (the chakravala, or "cyclic method") for solving Pell's equation was finally found by Jayadeva (cited in the eleventh century; his work is otherwise lost); the earliest surviving exposition appears in Bhāskara II's Bīja-gaṇita (twelfth century).[26]

Indian mathematics remained largely unknown in Europe until the late eighteenth century;[27] Brahmagupta and Bhāskara's work was translated into English in 1817 by Henry Colebrooke.[28]

Arithmetic in the Islamic golden age

Further information: Mathematics in medieval Islam

In the early ninth century, the caliph Al-Ma'mun ordered translations of many Greek mathematical works and at least one Sanskrit work (the *Sindhind*, which may [29] or may not[30] be Brahmagupta's Brāhmasphuţasiddhānta). Diophantus's main work, the *Arithmetica*, was translated into Arabic by Qusta ibn Luqa (820–912). Part of the treatise *al-Fakhri* (by al-Karajī, 953 – ca. 1029) builds on it to some extent. According to Rashed Roshdi, Al-Karajī's contemporary Ibn al-Haytham knew[31] what would later be called Wilson's theorem.

Western Europe in the Middle Ages

Other than a treatise on squares in arithmetic progression by Fibonacci — who lived and studied in north Africa and Constantinople during his formative years, ca. 1175–1200 — no number theory to speak of was done in western Europe during the Middle Ages. Matters started to change in Europe in the late Renaissance, thanks to a renewed study of the works of Greek antiquity. A catalyst was the textual emendation and translation into Latin of Diophantus's *Arithmetica* (Bachet, 1621, following a first attempt by Xylander, 1575).

1.1.2 Early modern number theory

Fermat

Pierre de Fermat (1601–1665) never published his writings; in particular, his work on number theory is contained almost entirely in letters to mathematicians and in private marginal notes.[32] He wrote down nearly no proofs in number theory; he had no models in the area.[33] He did make repeated use of mathematical induction, introducing the method of infinite descent.

One of Fermat's first interests was perfect numbers (which appear in Euclid, *Elements* IX) and amicable numbers;[note 7] this led him to work on integer divisors, which were from the beginning among the subjects of the correspondence (1636 onwards) that put him in touch with the mathematical community of the day.[34] He had already studied Bachet's edition of Diophantus carefully;[35] by 1643, his interests had shifted largely to Diophantine problems and sums of squares[36] (also treated by Diophantus).

Fermat's achievements in arithmetic include:

- Fermat's little theorem (1640),[37] stating that, if a is not divisible by a prime p, then $a^{p-1} \equiv 1 \mathrm{mod} p$. [note 8]
- If a and b are coprime, then a^2+b^2 is not divisible by any prime congruent to −1 modulo 4;[38] and every prime congruent to 1 modulo 4 can be written in the form $a^2 + b^2$.[39] These two statements also date from 1640; in 1659, Fermat stated to Huygens that he had proven the latter statement by the method of infinite descent.[40] Fermat and Frenicle also did some work (some of it erroneous)[41] on other quadratic forms.
- Fermat posed the problem of solving $x^2 - Ny^2 = 1$ as a challenge to English mathematicians (1657). The problem was solved in a few months by Wallis and Brouncker.[42] Fermat considered their solution valid, but pointed out they had provided an algorithm without a proof (as had Jayadeva and Bhaskara, though Fermat would never know this.) He states that a proof can be found by descent.
- Fermat developed methods for (doing what in our terms amounts to) finding points on curves of genus 0 and 1. As in Diophantus, there are many special procedures and what amounts to a tangent construction, but no use of a secant construction.[43]
- Fermat states and proves (by descent) in the appendix to *Observations on Diophantus* (Obs. XLV)[44] that $x^4 + y^4 = z^4$ has no non-trivial solutions in the integers. Fermat also mentioned to his correspondents that

$x^3 + y^3 = z^3$ has no non-trivial solutions, and that this could be proven by descent.[45] The first known proof is due to Euler (1753; indeed by descent).[46]

Fermat's claim ("Fermat's last theorem") to have shown there are no solutions to $x^n + y^n = z^n$ for all $n \geq 3$ (the only known proof of which is beyond his methods) appears only in his annotations on the margin of his copy of Diophantus; he never claimed this to others[47] and thus would have had no need to retract it if he found any mistake in his supposed proof.

Euler

The interest of Leonhard Euler (1707–1783) in number theory was first spurred in 1729, when a friend of his, the amateur[note 9] Goldbach, pointed him towards some of Fermat's work on the subject.[48][49] This has been called the "rebirth" of modern number theory,[35] after Fermat's relative lack of success in getting his contemporaries' attention for the subject.[50] Euler's work on number theory includes the following:[51]

- *Proofs for Fermat's statements.* This includes Fermat's little theorem (generalised by Euler to non-prime moduli); the fact that $p = x^2 + y^2$ if and only if $p \equiv 1 \bmod 4$; initial work towards a proof that every integer is the sum of four squares (the first complete proof is by Joseph-Louis Lagrange (1770), soon improved by Euler himself[52]); the lack of non-zero integer solutions to $x^4 + y^4 = z^2$ (implying the case *n=4* of Fermat's last theorem, the case *n=3* of which Euler also proved by a related method).
- *Pell's equation*, first misnamed by Euler.[53] He wrote on the link between continued fractions and Pell's equation.[54]
- *First steps towards analytic number theory.* In his work of sums of four squares, partitions, pentagonal numbers, and the distribution of prime numbers, Euler pioneered the use of what can be seen as analysis (in particular, infinite series) in number theory. Since he lived before the development of complex analysis, most of his work is restricted to the formal manipulation of power series. He did, however, do some very notable (though not fully rigorous) early work on what would later be called the Riemann zeta function.[55]
- *Quadratic forms.* Following Fermat's lead, Euler did further research on the question of which primes can be expressed in the form $x^2 + Ny^2$, some of it prefiguring quadratic reciprocity.[56] [57][58]
- *Diophantine equations.* Euler worked on some Diophantine equations of genus 0 and 1.[59][60] In particular, he studied Diophantus's work; he tried to systematise it, but the time was not yet ripe for such an endeavour – algebraic geometry was still in its infancy.[61] He did notice there was a connection between Diophantine problems and elliptic integrals,[61] whose study he had himself initiated.

Lagrange, Legendre and Gauss

Joseph-Louis Lagrange (1736–1813) was the first to give full proofs of some of Fermat's and Euler's work and observations – for instance, the four-square theorem and the basic theory of the misnamed "Pell's equation" (for which an algorithmic solution was found by Fermat and his contemporaries, and also by Jayadeva and Bhaskara II before them.) He also studied quadratic forms in full generality (as opposed to $mX^2 + nY^2$) — defining their equivalence relation, showing how to put them in reduced form, etc.

Adrien-Marie Legendre (1752–1833) was the first to state the law of quadratic reciprocity. He also conjectured what amounts to the prime number theorem and Dirichlet's theorem on arithmetic progressions. He gave a full treatment of the equation $ax^2 + by^2 + cz^2 = 0$ [62] and worked on quadratic forms along the lines later developed fully by Gauss.[63] In his old age, he was the first to prove "Fermat's last theorem" for $n = 5$ (completing work by Peter Gustav Lejeune Dirichlet, and crediting both him and Sophie Germain).[64]

In his *Disquisitiones Arithmeticae* (1798), Carl Friedrich Gauss (1777–1855) proved the law of quadratic reciprocity and developed the theory of quadratic forms (in particular, defining their composition). He also introduced some basic notation (congruences) and devoted a section to computational matters, including primality tests.[65] The last section of the *Disquisitiones* established a link between roots of unity and number theory:

> The theory of the division of the circle...which is treated in sec. 7 does not belong by itself to arithmetic, but its principles can only be drawn from higher arithmetic.[66]

In this way, Gauss arguably made a first foray towards both Évariste Galois's work and algebraic number theory.

1.1.3 Maturity and division into subfields

Starting early in the nineteenth century, the following developments gradually took place:

- The rise to self-consciousness of number theory (or *higher arithmetic*) as a field of study.[67]
- The development of much of modern mathematics necessary for basic modern number theory: complex analysis, group theory, Galois theory—accompanied by greater rigor in analysis and abstraction in algebra.
- The rough subdivision of number theory into its modern subfields—in particular, analytic and algebraic number theory.

Algebraic number theory may be said to start with the study of reciprocity and cyclotomy, but truly came into its own with the development of abstract algebra and early ideal theory and valuation theory; see below. A conventional starting point for analytic number theory is Dirichlet's theorem on arithmetic progressions (1837),[68] [69] whose proof introduced L-functions and involved some asymptotic analysis and a limiting process on a real variable.[70] The first use of analytic ideas in number theory actually goes back to Euler (1730s),[71] [72] who used formal power series and non-rigorous (or implicit) limiting arguments. The use of *complex* analysis in number theory comes later: the work of Bernhard Riemann (1859) on the zeta function is the canonical starting point;[73] Jacobi's four-square theorem (1839), which predates it, belongs to an initially different strand that has by now taken a leading role in analytic number theory (modular forms).[74]

The history of each subfield is briefly addressed in its own section below; see the main article of each subfield for fuller treatments. Many of the most interesting questions in each area remain open and are being actively worked on.

1.2 Main subdivisions

1.2.1 Elementary tools

The term *elementary* generally denotes a method that does not use complex analysis. For example, the prime number theorem was first proven using complex analysis in 1896, but an elementary proof was found only in 1949 by Erdős and Selberg.[75] The term is somewhat ambiguous: for example, proofs based on complex Tauberian theorems (e.g. Wiener–Ikehara) are often seen as quite enlightening but not elementary, in spite of using Fourier analysis, rather than complex analysis as such. Here as elsewhere, an *elementary* proof may be longer and more difficult for most readers than a non-elementary one.

Number theory has the reputation of being a field many of whose results can be stated to the layperson. At the same time, the proofs of these results are not particularly accessible, in part because the range of tools they use is, if anything, unusually broad within mathematics.[76]

1.2.2 Analytic number theory

Main article: Analytic number theory
Analytic number theory may be defined

- in terms of its tools, as the study of the integers by means of tools from real and complex analysis;[68] or
- in terms of its concerns, as the study within number theory of estimates on size and density, as opposed to identities.[77]

Some subjects generally considered to be part of analytic number theory, e.g., sieve theory,[note 10] are better covered by the second rather than the first definition: some of sieve theory, for instance, uses little analysis,[note 11] yet it does belong to analytic number theory.

The following are examples of problems in analytic number theory: the prime number theorem, the Goldbach conjecture (or the twin prime conjecture, or the Hardy–Littlewood conjectures), the Waring problem and the Riemann Hypothesis. Some of the most important tools of analytic number theory are the circle method, sieve methods and L-functions (or, rather, the study of their properties). The theory of modular forms (and, more generally, automorphic forms) also occupies an increasingly central place in the toolbox of analytic number theory.[78]

One may ask analytic questions about algebraic numbers, and use analytic means to answer such questions; it is thus that algebraic and analytic number theory intersect. For example, one may define prime ideals (generalizations of prime numbers in the field of algebraic numbers) and ask how many prime ideals there are up to a certain size. This question can be answered by means of an examination of Dedekind zeta functions, which are generalizations of the Riemann zeta function, a key analytic object at the roots of the subject.[79] This is an example of a general procedure in analytic number theory: deriving information about the distribution of a sequence (here, prime ideals or prime numbers) from the analytic behavior of an appropriately constructed complex-valued function.[80]

1.2.3 Algebraic number theory

Main article: Algebraic number theory

An *algebraic number* is any complex number that is a solution to some polynomial equation $f(x) = 0$ with rational coefficients; for example, every solution x of $x^5 + (11/2)x^3 - 7x^2 + 9 = 0$ (say) is an algebraic number. Fields of algebraic numbers are also called *algebraic number fields*, or shortly *number fields*. Algebraic number theory studies algebraic number fields.[81] Thus, analytic and algebraic number theory can and do overlap: the former is defined by its methods, the latter by its objects of study.

It could be argued that the simplest kind of number fields (viz., quadratic fields) were already studied by Gauss, as the discussion of quadratic forms in *Disquisitiones arithmeticae* can be restated in terms of ideals and norms in quadratic fields. (A *quadratic field* consists of all numbers of the form $a + b\sqrt{d}$, where a and b are rational numbers and d is a fixed rational number whose square root is not rational.) For that matter, the 11th-century chakravala method amounts—in modern terms—to an algorithm for finding the units of a real quadratic number field. However, neither Bhāskara nor Gauss knew of number fields as such.

The grounds of the subject as we know it were set in the late nineteenth century, when *ideal numbers*, the *theory of ideals* and *valuation theory* were developed; these are three complementary ways of dealing with the lack of unique factorisation in algebraic number fields. (For example, in the field generated by the rationals and $\sqrt{-5}$, the number 6 can be factorised both as $6 = 2 \cdot 3$ and $6 = (1 + \sqrt{-5})(1 - \sqrt{-5})$; all of 2 , 3 , $1 + \sqrt{-5}$ and $1 - \sqrt{-5}$ are irreducible, and thus, in a naïve sense, analogous to primes among the integers.) The initial impetus for the development of ideal numbers (by Kummer) seems to have come from the study of higher reciprocity laws,[82]i.e., generalisations of quadratic reciprocity.

Number fields are often studied as extensions of smaller number fields: a field L is said to be an *extension* of a field K if L contains K. (For example, the complex numbers C are an extension of the reals R, and the reals R are an extension of the rationals Q.) Classifying the possible extensions of a given number field is a difficult and partially open problem. Abelian extensions—that is, extensions L of K such that the Galois group[note 12] $\mathrm{Gal}(L/K)$ of L over K is an abelian group—are relatively well understood. Their classification was the object of the programme of class field theory, which was initiated in the late 19th century (partly by Kronecker and Eisenstein) and carried out largely in 1900—1950.

An example of an active area of research in algebraic number theory is Iwasawa theory. The Langlands program, one of the main current large-scale research plans in mathematics, is sometimes described as an attempt to generalise class field theory to non-abelian extensions of number fields.

1.2.4 Diophantine geometry

Main article: Diophantine geometry

The central problem of *Diophantine geometry* is to determine when a Diophantine equation has solutions, and if it does, how many. The approach taken is to think of the solutions of an equation as a geometric object.

For example, an equation in two variables defines a curve in the plane. More generally, an equation, or system of equations, in two or more variables defines a curve, a surface or some other such object in n-dimensional space. In Diophantine geometry, one asks whether there are any *rational points* (points all of whose coordinates are rationals) or *integral points* (points all of whose coordinates are integers) on the curve or surface. If there are any such points, the next step is to ask how many there are and how they are distributed. A basic question in this direction is: are there finitely or infinitely many rational points on a given curve (or surface)? What about integer points?

An example here may be helpful. Consider the Pythagorean equation $x^2+y^2=1$; we would like to study its rational solutions, i.e., its solutions (x, y) such that *x* and *y* are both rational. This is the same as asking for all integer solutions to $a^2+b^2=c^2$; any solution to the latter equation gives us a solution $x=a/c$, $y=b/c$ to the former. It is also the same as asking for all points with rational coordinates on the curve described by $x^2+y^2=1$. (This curve happens to be a circle of radius 1 around the origin.)

The rephrasing of questions on equations in terms of points on curves turns out to be felicitous. The finiteness or not of the number of rational or integer points on an algebraic curve—that is, rational or integer solutions to an equation $f(x, y)=0$, where f is a polynomial in two variables—turns out to depend crucially on the *genus* of the curve. The *genus* can be defined as follows:[note 13] allow the variables in $f(x, y)=0$ to be complex numbers; then $f(x, y)=0$ defines a 2-dimensional surface in (projective) 4-dimensional space (since two complex variables can be decomposed into four real variables, i.e., four dimensions). Count the number of (doughnut) holes in the surface; call this number the *genus* of $f(x, y)=0$. Other geometrical notions turn out to be just as crucial.

There is also the closely linked area of Diophantine approximations: given a number x , how well can it be approximated by rationals? (We are looking for approximations that are good relative to the amount of space that it takes to write the rational: call a/q (with $\gcd(a, q)=1$) a good approximation to x if $|x-a/q|<\frac{1}{q^c}$, where c is large.) This question is of special interest if x is an algebraic number. If x cannot be well approximated, then some equations do not have integer or rational solutions. Moreover, several concepts (especially that of height) turn out to be crucial both in Diophantine geometry and in the study of Diophantine approximations. This question is also of special interest in transcendental number theory: if a number can be better approximated than any algebraic number, then it is a transcendental number. It is by this argument that π and e have been shown to be transcendental.

Diophantine geometry should not be confused with the geometry of numbers, which is a collection of graphical methods for answering certain questions in algebraic number theory. *Arithmetic geometry*, on the other hand, is a contemporary term for much the same domain as that covered by the term *Diophantine geometry*. The term *arithmetic geometry* is arguably used most often when one wishes to emphasise the connections to modern algebraic geometry (as in, for instance, Faltings' theorem) rather than to techniques in Diophantine approximations.

1.3 Recent approaches and subfields

The areas below date as such from no earlier than the mid-twentieth century, even if they are based on older material. For example, as is explained below, the matter of algorithms in number theory is very old, in some sense older than the concept of proof; at the same time, the modern study of computability dates only from the 1930s and 1940s, and computational complexity theory from the 1970s.

1.3.1 Probabilistic number theory

Main article: Probabilistic number theory

Take a number at random between one and a million. How likely is it to be prime? This is just another way of asking how many primes there are between one and a million. Further: how many prime divisors will it have, on average? How many divisors will it have altogether, and with what likelihood? What is the probability that it will have many more or many fewer divisors or prime divisors than the average?

Much of probabilistic number theory can be seen as an important special case of the study of variables that are almost, but not quite, mutually independent. For example, the event that a random integer between one and a million be divisible by two and the event that it be divisible by three are almost independent, but not quite.

It is sometimes said that probabilistic combinatorics uses the fact that whatever happens with probability greater than 0 must happen sometimes; one may say with equal justice that many applications of probabilistic number theory hinge on the fact that whatever is unusual must be rare. If certain algebraic objects (say, rational or integer solutions to certain equations) can be shown to be in the tail of certain sensibly defined distributions, it follows that there must be few of them; this is a very concrete non-probabilistic statement following from a probabilistic one.

At times, a non-rigorous, probabilistic approach leads to a number of heuristic algorithms and open problems, notably Cramér's conjecture.

1.3.2 Arithmetic combinatorics

Main articles: Arithmetic combinatorics and Additive number theory

Let A be a set of N integers. Consider the set $A + A = \{ m + n \mid m, n \in A \}$ consisting of all sums of two elements of A. Is $A + A$ much larger than A? Barely larger? If $A + A$ is barely larger than A, must A have plenty of arithmetic structure, for example, does A resemble an arithmetic progression?

If we begin from a fairly "thick" infinite set A , does it contain many elements in arithmetic progression: a , $a + b, a + 2b, a + 3b, \ldots, a + 10b$, say? Should it be possible to write large integers as sums of elements of A ?

These questions are characteristic of *arithmetic combinatorics*. This is a presently coalescing field; it subsumes *additive number theory* (which concerns itself with certain very specific sets A of arithmetic significance, such as the primes or the squares) and, arguably, some of the *geometry of numbers*, together with some rapidly developing new material. Its focus on issues of growth and distribution accounts in part for its developing links with ergodic theory, finite group theory, model theory, and other fields. The term *additive combinatorics* is also used; however, the sets A being studied need not be sets of integers, but rather subsets of non-commutative groups, for which the multiplication symbol, not the addition symbol, is traditionally used; they can also be subsets of rings, in which case the growth of $A + A$ and $A \cdot A$ may be compared.

1.3.3 Computations in number theory

Main article: Computational number theory

While the word *algorithm* goes back only to certain readers of al-Khwārizmī, careful descriptions of methods of solution are older than proofs: such methods (that is, algorithms) are as old as any recognisable mathematics—ancient Egyptian, Babylonian, Vedic, Chinese—whereas proofs appeared only with the Greeks of the classical period. An interesting early case is that of what we now call the Euclidean algorithm. In its basic form (namely, as an algorithm for computing the greatest common divisor) it appears as Proposition 2 of Book VII in *Elements*, together with a proof of correctness. However, in the form that is often used in number theory (namely, as an algorithm for finding integer solutions to an equation $ax + by = c$, or, what is the same, for finding the quantities whose existence is assured by the Chinese remainder theorem) it first appears in the works of Āryabhaṭa (5th–6th century CE) as an algorithm called *kuṭṭaka* ("pulveriser"), without a proof of correctness.

There are two main questions: "can we compute this?" and "can we compute it rapidly?". Anybody can test whether a number is prime or, if it is not, split it into prime factors; doing so rapidly is another matter. We now know fast algorithms for testing primality, but, in spite of much work (both theoretical and practical), no truly fast algorithm for factoring.

The difficulty of a computation can be useful: modern protocols for encrypting messages (e.g., RSA) depend on functions that are known to all, but whose inverses (a) are known only to a chosen few, and (b) would take one too long a time to figure out on one's own. For example, these functions can be such that their inverses can be computed only if certain large integers are factorized. While many difficult computational problems outside number theory are known, most working encryption protocols nowadays are based on the difficulty of a few number-theoretical problems.

On a different note — some things may not be computable at all; in fact, this can be proven in some instances. For instance, in 1970, it was proven, as a solution to Hilbert's 10th problem, that there is no Turing machine which can solve all Diophantine equations.[83] In particular, this means that, given a computably enumerable set of axioms, there are Diophantine equations for which there is no proof, starting from the axioms, of whether the set of equations has or does not have integer solutions. (We would necessarily be speaking of Diophantine equations for which there are no integer solutions, since, given a Diophantine equation with at least one solution, the solution itself provides a proof of the fact that a solution exists. We cannot prove, of course, that a particular Diophantine equation is of this kind, since this would imply that it has no solutions.)

1.4 Applications

The number-theorist Leonard Dickson (1874–1954) said "Thank God that number theory is unsullied by any application". Such a view is no longer applicable to number theory.[84] In 1974, Donald Knuth said "...virtually every theorem in elementary number theory arises in a natural, motivated way in connection with the problem of making computers do high-speed numerical calculations".[85] Elementary number theory is taught in discrete mathematics courses for computer scientists; and, on the other hand, number theory also has applications to the continuous in numerical analysis.[86] As well as the well-known applications to cryptography, there are also applications to many other areas of mathematics.[87][88]

1.5 Literature

Two of the most popular introductions to the subject are:

- G. H. Hardy; E. M. Wright (2008) [1938]. *An introduction to the theory of numbers* (rev. by D. R. Heath-Brown and J. H. Silverman, 6th ed.). Oxford University Press. ISBN 978-0-19-921986-5. Retrieved 2016-03-02.
- Vinogradov, I. M. (2003) [1954]. *Elements of Number Theory* (reprint of the 1954 ed.). Mineola, NY: Dover Publications.

Hardy and Wright's book is a comprehensive classic, though its clarity sometimes suffers due to the authors' insistence on elementary methods.[89] Vinogradov's main attraction consists in its set of problems, which quickly lead to Vinogradov's own research interests; the text itself is very basic and close to minimal. Other popular first introductions are:

- Ivan M. Niven; Herbert S. Zuckerman; Hugh L. Montgomery (2008) [1960]. *An introduction to the theory of numbers* (reprint of the 5th edition 1991 ed.). John Wiley & Sons. ISBN 978-8-12-651811-1. Retrieved 2016-02-28.
- Kenneth H. Rosen (2010). *Elementary Number Theory* (6th ed.). Pearson Education. ISBN 978-0-32-171775-7. Retrieved 2016-02-28.

Popular choices for a second textbook include:

- Borevich, A. I.; Shafarevich, Igor R. (1966). *Number theory*. Pure and Applied Mathematics. **20**. Boston, MA: Academic Press. ISBN 978-0-12-117850-5. MR 0195803.
- Serre, Jean-Pierre (1996) [1973]. *A course in arithmetic*. Graduate texts in mathematics. **7**. Springer. ISBN 978-0-387-90040-7.

1.6 Prizes

The American Mathematical Society awards the *Cole Prize in Number Theory*. Moreover number theory is one of the three mathematical subdisciplines rewarded by the *Fermat Prize*.

1.7 See also

- Algebraic function field
- Finite field
- p-adic number

1.8 Notes

[1] See two following notes.

[2] Already in 1921, T. L. Heath had to explain: "By arithmetic, Plato meant, not arithmetic in our sense, but the science which considers numbers in themselves, in other words, what we mean by the Theory of Numbers." (Heath 1921, p. 13)

[3] Take, e.g. Serre 1973. In 1952, Davenport still had to specify that he meant *The Higher Arithmetic*. Hardy and Wright wrote in the introduction to *An Introduction to the Theory of Numbers* (1938): "We proposed at one time to change [the title] to *An introduction to arithmetic*, a more novel and in some ways a more appropriate title; but it was pointed out that this might lead to misunderstandings about the content of the book." (Hardy & Wright 2008)

[4] Robson 2001, p. 201. This is controversial. See Plimpton 322. Robson's article is written polemically (Robson 2001, p. 202) with a view to "perhaps [...] knocking [Plimpton 322] off its pedestal" (Robson 2001, p. 167); at the same time, it settles to the conclusion that

> [...] the question "how was the tablet calculated?" does not have to have the same answer as the question "what problems does the tablet set?" The first can be answered most satisfactorily by reciprocal pairs, as first suggested half a century ago, and the second by some sort of right-triangle problems (Robson 2001, p. 202).

Robson takes issue with the notion that the scribe who produced Plimpton 322 (who had to "work for a living", and would not have belonged to a "leisured middle class") could have been motivated by his own "idle curiosity" in the absence of a "market for new mathematics".(Robson 2001, pp. 199–200)

[5] Sun Zi, *Suan Ching*, Ch. 3, Problem 26, in Lam & Ang 2004, pp. 219–220:

> [26] Now there are an unknown number of things. If we count by threes, there is a remainder 2; if we count by fives, there is a remainder 3; if we count by sevens, there is a remainder 2. Find the number of things.
> *Answer*: 23.
> *Method*: If we count by threes and there is a remainder 2, put down 140. If we count by fives and there is a remainder 3, put down 63. If we count by sevens and there is a remainder 2, put down 30. Add them to obtain 233 and subtract 210 to get the answer. If we count by threes and there is a remainder 1, put down 70. If we count by fives and there is a remainder 1, put down 21. If we count by sevens and there is a remainder 1, put down 15. When [a number] exceeds 106, the result is obtained by subtracting 105.

[6] See, e.g., Sun Zi, *Suan Ching*, Ch. 3, Problem 36, in Lam & Ang 2004, pp. 223–224:

> [36] Now there is a pregnant woman whose age is 29. If the gestation period is 9 months, determine the sex of the unborn child. *Answer*: Male.
> *Method*: Put down 49, add the gestation period and subtract the age. From the remainder take away 1 representing the heaven, 2 the earth, 3 the man, 4 the four seasons, 5 the five phases, 6 the six pitch-pipes, 7 the seven stars [of the Dipper], 8 the eight winds, and 9 the nine divisions [of China under Yu the Great]. If the remainder is odd, [the sex] is male and if the remainder is even, [the sex] is female.

This is the last problem in Sun Zi's otherwise matter-of-fact treatise.

[7] Perfect and especially amicable numbers are of little or no interest nowadays. The same was not true in medieval times – whether in the West or the Arab-speaking world – due in part to the importance given to them by the Neopythagorean (and hence mystical) Nicomachus (ca. 100 CE), who wrote a primitive but influential "Introduction to Arithmetic". See van der Waerden 1961, Ch. IV.

[8] Here, as usual, given two integers a and b and a non-zero integer m, we write $a \equiv b \bmod m$ (read "a is congruent to b modulo m") to mean that m divides $a - b$, or, what is the same, a and b leave the same residue when divided by m. This notation is actually much later than Fermat's; it first appears in section 1 of Gauss's Disquisitiones Arithmeticae. Fermat's little theorem is a consequence of the fact that the order of an element of a group divides the order of the group. The modern proof would have been within Fermat's means (and was indeed given later by Euler), even though the modern concept of a group came long after Fermat or Euler. (It helps to know that inverses exist modulo p (i.e., given a not divisible by a prime p, there is an integer x such that $xa \equiv 1 \bmod p$); this fact (which, in modern language, makes the residues mod p into a group, and which was already known to Āryabhaṭa; see above) was familiar to Fermat thanks to its rediscovery by Bachet (Weil 1984, p. 7). Weil goes on to say that Fermat would have recognised that Bachet's argument is essentially Euclid's algorithm.

[9] Up to the second half of the seventeenth century, academic positions were very rare, and most mathematicians and scientists earned their living in some other way (Weil 1984, pp. 159, 161). (There were already some recognisable features of professional *practice*, viz., seeking correspondents, visiting foreign colleagues, building private libraries (Weil 1984, pp. 160–161). Matters started to shift in the late 17th century (Weil 1984, p. 161); scientific academies were founded in England (the Royal Society, 1662) and France (the Académie des sciences, 1666) and Russia (1724). Euler was offered a

position at this last one in 1726; he accepted, arriving in St. Petersburg in 1727 (Weil 1984, p. 163 and Varadarajan 2006, p. 7). In this context, the term *amateur* usually applied to Goldbach is well-defined and makes some sense: he has been described as a man of letters who earned a living as a spy (Truesdell 1984, p. xv); cited in Varadarajan 2006, p. 9). Notice, however, that Goldbach published some works on mathematics and sometimes held academic positions.

[10] Sieve theory figures as one of the main subareas of analytic number theory in many standard treatments; see, for instance, Iwaniec & Kowalski 2004 or Montgomery & Vaughan 2007

[11] This is the case for small sieves (in particular, some combinatorial sieves such as the Brun sieve) rather than for large sieves; the study of the latter now includes ideas from harmonic and functional analysis.

[12] The Galois group of an extension *K/L* consists of the operations (isomorphisms) that send elements of L to other elements of L while leaving all elements of K fixed. Thus, for instance, *Gal(C/R)* consists of two elements: the identity element (taking every element $x + iy$ of C to itself) and complex conjugation (the map taking each element $x + iy$ to $x - iy$). The Galois group of an extension tells us many of its crucial properties. The study of Galois groups started with Évariste Galois; in modern language, the main outcome of his work is that an equation $f(x) = 0$ can be solved by radicals (that is, x can be expressed in terms of the four basic operations together with square roots, cubic roots, etc.) if and only if the extension of the rationals by the roots of the equation $f(x) = 0$ has a Galois group that is solvable in the sense of group theory. ("Solvable", in the sense of group theory, is a simple property that can be checked easily for finite groups.)

[13] It may be useful to look at an example here. Say we want to study the curve $y^2 = x^3 + 7$. We allow x and y to be complex numbers: $(a+bi)^2 = (c+di)^3 + 7$. This is, in effect, a set of two equations on four variables, since both the real and the imaginary part on each side must match. As a result, we get a surface (two-dimensional) in four-dimensional space. After we choose a convenient hyperplane on which to project the surface (meaning that, say, we choose to ignore the coordinate a), we can plot the resulting projection, which is a surface in ordinary three-dimensional space. It then becomes clear that the result is a torus, i.e., the surface of a doughnut (somewhat stretched). A doughnut has one hole; hence the genus is 1.

1.9 References

[1] Long 1972, p. 1.

[2] Neugebauer & Sachs 1945, p. 40. The term *takiltum* is problematic. Robson prefers the rendering "The holding-square of the diagonal from which 1 is torn out, so that the short side comes up...".Robson 2001, p. 192

[3] Robson 2001, p. 189. Other sources give the modern formula $(p^2 - q^2, 2pq, p^2 + q^2)$. Van der Waerden gives both the modern formula and what amounts to the form preferred by Robson.(van der Waerden 1961, p. 79)

[4] van der Waerden 1961, p. 184.

[5] Neugebauer (Neugebauer 1969, pp. 36–40) discusses the table in detail and mentions in passing Euclid's method in modern notation (Neugebauer 1969, p. 39).

[6] Friberg 1981, p. 302.

[7] van der Waerden 1961, p. 43.

[8] Iamblichus, *Life of Pythagoras*,(trans. e.g. Guthrie 1987) cited in van der Waerden 1961, p. 108. See also Porphyry, *Life of Pythagoras*, paragraph 6, in Guthrie 1987 Van der Waerden (van der Waerden 1961, pp. 87–90) sustains the view that Thales knew Babylonian mathematics.

[9] Herodotus (II. 81) and Isocrates (*Busiris* 28), cited in: Huffman 2011. On Thales, see Eudemus ap. Proclus, 65.7, (e.g. Morrow 1992, p. 52) cited in: O'Grady 2004, p. 1. Proclus was using a work by Eudemus of Rhodes (now lost), the *Catalogue of Geometers*. See also introduction, Morrow 1992, p. xxx on Proclus' reliability.

[10] Becker 1936, p. 533, cited in: van der Waerden 1961, p. 108.

[11] Becker 1936.

[12] van der Waerden 1961, p. 109.

[13] Plato, *Theaetetus*, p. 147 B, (e.g. Jowett 1871), cited in von Fritz 2004, p. 212: "Theodorus was writing out for us something about roots, such as the roots of three or five, showing that they are incommensurable by the unit;..." *See also* Spiral of Theodorus.

[14] von Fritz 2004.

[15] Heath 1921, p. 76.

[16] Sun Zi, *Suan Ching*, Chapter 3, Problem 26. This can be found in Lam & Ang 2004, pp. 219–220, which contains a full translation of the *Suan Ching* (based on Qian 1963). See also the discussion in Lam & Ang 2004, pp. 138–140.

[17] The date of the text has been narrowed down to 220–420 AD (Yan Dunjie) or 280–473 AD (Wang Ling) through internal evidence (= taxation systems assumed in the text). See Lam & Ang 2004, pp. 27–28.

[18] Boyer & Merzbach 1991, p. 82.

[19] Vardi 1998, p. 305-319.

[20] Weil 1984, pp. 17–24.

[21] Plofker 2008, p. 119.

[22] Any early contact between Babylonian and Indian mathematics remains conjectural (Plofker 2008, p. 42).

[23] Mumford 2010, p. 387.

[24] Āryabhaṭa, Āryabhatīya, Chapter 2, verses 32–33, cited in: Plofker 2008, pp. 134–140. See also Clark 1930, pp. 42–50. A slightly more explicit description of the kuṭṭaka was later given in Brahmagupta, *Brāhmasphuṭasiddhānta*, XVIII, 3–5 (in Colebrooke 1817, p. 325, cited in Clark 1930, p. 42).

[25] Mumford 2010, p. 388.

[26] Plofker 2008, p. 194.

[27] Plofker 2008, p. 283.

[28] Colebrooke 1817.

[29] Colebrooke 1817, p. lxv, cited in Hopkins 1990, p. 302. See also the preface in Sachau 1888 cited in Smith 1958, pp. 168

[30] Pingree 1968, pp. 97–125, and Pingree 1970, pp. 103–123, cited in Plofker 2008, p. 256.

[31] Rashed 1980, p. 305–321.

[32] Weil 1984, pp. 45–46.

[33] Weil 1984, p. 118. This was more so in number theory than in other areas (remark in Mahoney 1994, p. 284). Bachet's own proofs were "ludicrously clumsy" (Weil 1984, p. 33).

[34] Mahoney 1994, pp. 48, 53–54. The initial subjects of Fermat's correspondence included divisors ("aliquot parts") and many subjects outside number theory; see the list in the letter from Fermat to Roberval, 22.IX.1636, Tannery & Henry 1891, Vol. II, pp. 72, 74, cited in Mahoney 1994, p. 54.

[35] Weil 1984, pp. 1–2.

[36] Weil 1984, p. 53.

[37] Tannery & Henry 1891, Vol. II, p. 209, Letter XLVI from Fermat to Frenicle, 1640, cited in Weil 1984, p. 56

[38] Tannery & Henry 1891, Vol. II, p. 204, cited in Weil 1984, p. 63. All of the following citations from Fermat's *Varia Opera* are taken from Weil 1984, Chap. II. The standard Tannery & Henry work includes a revision of Fermat's posthumous *Varia Opera Mathematica* originally prepared by his son (Fermat 1679).

[39] Tannery & Henry 1891, Vol. II, p. 213.

[40] Tannery & Henry 1891, Vol. II, p. 423.

[41] Weil 1984, pp. 80, 91–92.

[42] Weil 1984, p. 92.

[43] Weil 1984, Ch. II, sect. XV and XVI.

[44] Tannery & Henry 1891, Vol. I, pp. 340–341.

[45] Weil 1984, p. 115.

[46] Weil 1984, pp. 115–116.

[47] Weil 1984, p. 104.

[48] Weil 1984, pp. 2, 172.

[49] Varadarajan 2006, p. 9.

[50] Weil 1984, p. 2 and Varadarajan 2006, p. 37

[51] Varadarajan 2006, p. 39 and Weil 1984, pp. 176–189

[52] Weil 1984, pp. 178–179.

[53] Weil 1984, p. 174. Euler was generous in giving credit to others (Varadarajan 2006, p. 14), not always correctly.

[54] Weil 1984, p. 183.

[55] Varadarajan 2006, pp. 45–55; see also chapter III.

[56] Varadarajan 2006, pp. 44–47.

[57] Weil 1984, pp. 177–179.

[58] Edwards 1983, pp. 285–291.

[59] Varadarajan 2006, pp. 55–56.

[60] Weil 1984, pp. 179–181.

[61] Weil 1984, p. 181.

[62] Weil 1984, pp. 327–328.

[63] Weil 1984, pp. 332–334.

[64] Weil 1984, pp. 337–338.

[65] Goldstein & Schappacher 2007, p. 14.

[66] From the preface of *Disquisitiones Arithmeticae*; the translation is taken from Goldstein & Schappacher 2007, p. 16

[67] See the discussion in section 5 of Goldstein & Schappacher 2007. Early signs of self-consciousness are present already in letters by Fermat: thus his remarks on what number theory is, and how "Diophantus's work [...] does not really belong to [it]" (quoted in Weil 1984, p. 25).

[68] Apostol 1976, p. 7.

[69] Davenport & Montgomery 2000, p. 1.

[70] See the proof in Davenport & Montgomery 2000, section 1

[71] Iwaniec & Kowalski 2004, p. 1.

[72] Varadarajan 2006, sections 2.5, 3.1 and 6.1.

[73] Granville 2008, pp. 322–348.

[74] See the comment on the importance of modularity in Iwaniec & Kowalski 2004, p. 1

[75] Goldfeld 2003.

[76] See, e.g., the initial comment in Iwaniec & Kowalski 2004, p. 1.

[77] Granville 2008, section 1: "The main difference is that in algebraic number theory [...] one typically considers questions with answers that are given by exact formulas, whereas in analytic number theory [...] one looks for *good approximations*."

[78] See the remarks in the introduction to Iwaniec & Kowalski 2004, p. 1: "However much stronger...".

[79] Granville 2008, section 3: "[Riemann] defined what we now call the Riemann zeta function [...] Riemann's deep work gave birth to our subject [...]"

[80] See, e.g., Montgomery & Vaughan 2007, p. 1.

[81] CITEREFMilne2014, p. 2.

[82] Edwards 2000, p. 79.

[83] Davis, Martin; Matiyasevich, Yuri; Robinson, Julia (1976). "Hilbert's Tenth Problem: Diophantine Equations: Positive Aspects of a Negative Solution". In Felix E. Browder. *Mathematical Developments Arising from Hilbert Problems*. Proceedings of Symposia in Pure Mathematics. XXVIII.2. American Mathematical Society. pp. 323–378. ISBN 0-8218-1428-1. Zbl 0346.02026. Reprinted in *The Collected Works of Julia Robinson*, Solomon Feferman, editor, pp.269–378, American Mathematical Society 1996.

[84] "The Unreasonable Effectiveness of Number Theory", Stefan Andrus Burr, George E. Andrews, American Mathematical Soc., 1992, ISBN 9780821855010

[85] Computer science and its relation to mathematics" DE Knuth – The American Mathematical Monthly, 1974

[86] "Applications of number theory to numerical analysis", Lo-keng Hua, Luogeng Hua, Yuan Wang, Springer-Verlag, 1981, ISBN 978-3-540-10382-0

[87] "Practical applications of algebraic number theory". Mathoverflow.net. Retrieved 2012-05-18.

[88] "Where is number theory used in the rest of mathematics?". Mathoverflow.net. 2008-09-23. Retrieved 2012-05-18.

[89] Apostol n.d.

1.10 Sources

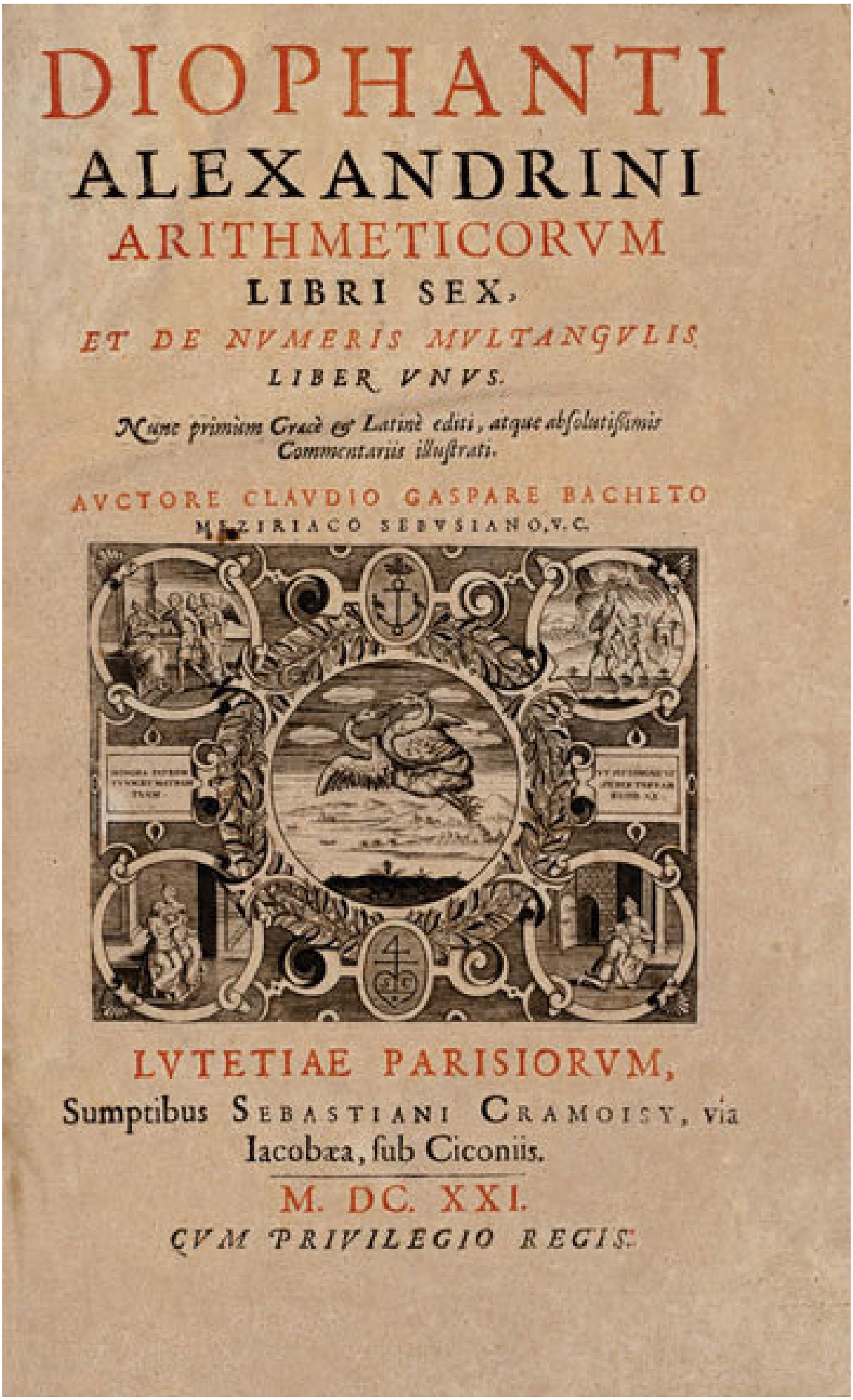
- Apostol, Tom M. (1976). *Introduction to analytic number theory*. Undergraduate Texts in Mathematics. Springer. ISBN 978-0-387-90163-3. Retrieved 2016-02-28.
- Apostol, Tom M. (n.d.). "An Introduction to the Theory of Numbers". (Review of Hardy & Wright.) Mathematical Reviews (MathSciNet) MR0568909. American Mathematical Society. Retrieved 2016-02-28. (Subscription needed)
- Becker, Oskar (1936). "Die Lehre von Geraden und Ungeraden im neunten Buch der euklidischen Elemente". *Quellen und Studien zur Geschichte der Mathematik, Astronomie und Physik*. Abteilung B:Studien (in German). Berlin: J. Springer Verlag. **3**: 533–53.
- Boyer, Carl Benjamin; Merzbach, Uta C. (1991) [1968]. *A History of Mathematics* (2nd ed.). New York: Wiley. ISBN 978-0-471-54397-8. 1968 edition at archive.org
- Clark, Walter Eugene (trans.) (1930). *The Āryabhaṭīya of Āryabhaṭa: An ancient Indian work on Mathematics and Astronomy*. University of Chicago Press. Retrieved 2016-02-28.
- Colebrooke, Henry Thomas (1817). *Algebra, with Arithmetic and Mensuration, from the Sanscrit of Brahmegupta and Bháscara.* London: J. Murray. Retrieved 2016-02-28.
- Davenport, Harold; Montgomery, Hugh L. (2000). *Multiplicative Number Theory*. Graduate texts in mathematics. **74** (revised 3rd ed.). Springer. ISBN 978-0-387-95097-6.
- Edwards, Harold M. (November 1983). "Euler and Quadratic Reciprocity". *Mathematics Magazine*. Mathematical Association of America. **56** (5): 285–291. doi:10.2307/2690368. JSTOR 2690368.
- Edwards, Harold M. (2000) [1977]. *Fermat's Last Theorem: a Genetic Introduction to Algebraic Number Theory*. Graduate Texts in Mathematics. **50** (reprint of 1977 ed.). Springer Verlag. ISBN 978-0-387-95002-0.
- Fermat, Pierre de (1679). *Varia Opera Mathematica* (in French and Latin). Toulouse: Joannis Pech. Retrieved 2016-02-28.
- Friberg, Jöran (August 1981). "Methods and Traditions of Babylonian Mathematics: Plimpton 322, Pythagorean Triples and the Babylonian Triangle Parameter Equations". *Historia Mathematica*. Elsevier. **8** (3): 277–318. doi:10.1016/0315-0860(81)90069-0.
- von Fritz, Kurt (2004). "The Discovery of Incommensurability by Hippasus of Metapontum". In Christianidis, J. *Classics in the History of Greek Mathematics*. Berlin: Kluwer (Springer). ISBN 978-1-4020-0081-2.
- Gauss, Carl Friedrich; Waterhouse, William C. (trans.) (1966) [1801]. *Disquisitiones Arithmeticae*. Springer. ISBN 978-0-387-96254-2.

- Goldfeld, Dorian M. (2003). "Elementary Proof of the Prime Number Theorem: a Historical Perspective" (PDF). Retrieved 2016-02-28.
- Goldstein, Catherine; Schappacher, Norbert (2007). "A book in search of a discipline". In Goldstein, C.; Schappacher, N.; Schwermer, Joachim. *The Shaping of Arithmetic after Gauss' "Disquisitiones Arithmeticae"*. Berlin & Heidelberg: Springer. pp. 3–66. ISBN 978-3-540-20441-1. Retrieved 2016-02-28.
- Granville, Andrew (2008). "Analytic number theory". In Gowers, Timothy; Barrow-Green, June; Leader, Imre. *The Princeton Companion to Mathematics*. Princeton University Press. ISBN 978-0-691-11880-2. Retrieved 2016-02-28.
- Porphyry; Guthrie, K. S. (trans.) (1920). *Life of Pythagoras*. Alpine, New Jersey: Platonist Press.
- Guthrie, Kenneth Sylvan (1987). *The Pythagorean Sourcebook and Library*. Grand Rapids, Michigan: Phanes Press. ISBN 978-0-933999-51-0.
- Hardy, Godfrey Harold; Wright, E. M. (2008) [1938]. *An Introduction to the Theory of Numbers* (Sixth ed.). Oxford University Press. ISBN 978-0-19-921986-5. MR 2445243.
- Heath, Thomas L. (1921). *A History of Greek Mathematics, Volume 1: From Thales to Euclid*. Oxford: Clarendon Press. Retrieved 2016-02-28.
- Hopkins, J. F. P. (1990). "Geographical and Navigational Literature". In Young, M. J. L.; Latham, J. D.; Serjeant, R. B. *Religion, Learning and Science in the `Abbasid Period*. The Cambridge history of Arabic literature. Cambridge University Press. ISBN 978-0-521-32763-3.
- Huffman, Carl A. (8 August 2011). Zalta, Edward N., ed. "Pythagoras". *Stanford Encyclopaedia of Philosophy* (Fall 2011 ed.). Retrieved 7 February 2012.
- Iwaniec, Henryk; Kowalski, Emmanuel (2004). *Analytic Number Theory*. American Mathematical Society Colloquium Publications. **53**. Providence, RI,: American Mathematical Society. ISBN 0-8218-3633-1.
- Plato; Jowett, Benjamin (trans.) (1871). *Theaetetus*.
- Lam, Lay Yong; Ang, Tian Se (2004). *Fleeting Footsteps: Tracing the Conception of Arithmetic and Algebra in Ancient China* (revised ed.). Singapore: World Scientific. ISBN 978-981-238-696-0. Retrieved 2016-02-28.
- Long, Calvin T. (1972). *Elementary Introduction to Number Theory* (2nd ed.). Lexington, VA: D. C. Heath and Company. LCCN 77171950.
- Mahoney, M. S. (1994). *The Mathematical Career of Pierre de Fermat, 1601–1665* (Reprint, 2nd ed.). Princeton University Press. ISBN 978-0-691-03666-3. Retrieved 2016-02-28.
- Milne, J. S. (2014). "Algebraic Number Theory". Available at www.jmilne.org/math.
- Montgomery, Hugh L.; Vaughan, Robert C. (2007). *Multiplicative Number Theory: I, Classical Theory,*. Cambridge University Press. ISBN 978-0-521-84903-6. Retrieved 2016-02-28.
- Morrow, Glenn Raymond (trans., ed.); Proclus (1992). *A Commentary on Book 1 of Euclid's Elements*. Princeton University Press. ISBN 978-0-691-02090-7.
- Mumford, David (March 2010). "Mathematics in India: reviewed by David Mumford" (PDF). *Notices of the American Mathematical Society*. **57** (3): 387. ISSN 1088-9477.
- Neugebauer, Otto E. (1969). *The Exact Sciences in Antiquity* (corrected reprint of the 1957 ed.). New York: Dover Publications. ISBN 978-0-486-22332-2. Retrieved 2016-03-02.
- Neugebauer, Otto E.; Sachs, Abraham Joseph; Götze, Albrecht (1945). *Mathematical Cuneiform Texts*. American Oriental Series. **29**. American Oriental Society etc.
- O'Grady, Patricia (September 2004). "Thales of Miletus". The Internet Encyclopaedia of Philosophy. Retrieved 7 February 2012.
- Pingree, David; Ya'qub, ibn Tariq (1968). "The Fragments of the Works of Ya'qub ibn Tariq". *Journal of Near Eastern Studies*. University of Chicago Press. **26**.

- Pingree, D.; al-Fazari (1970). "The Fragments of the Works of al-Fazari". *Journal of Near Eastern Studies*. University of Chicago Press. **28**.
- Plofker, Kim (2008). *Mathematics in India*. Princeton University Press. ISBN 978-0-691-12067-6.
- Qian, Baocong, ed. (1963). *Suanjing shi shu (Ten Mathematical Classics)* (in Chinese). Beijing: Zhonghua shuju. Retrieved 2016-02-28.
- Rashed, Roshdi (1980). "Ibn al-Haytham et le théorème de Wilson". *Archive for History of Exact Sciences*. **22** (4): 305–321. doi:10.1007/BF00717654.
- Robson, Eleanor (2001). "Neither Sherlock Holmes nor Babylon: a Reassessment of Plimpton 322" (PDF). *Historia Mathematica*. Elsevier. **28** (28): 167–206. doi:10.1006/hmat.2001.2317. Archived from the original (PDF) on 2014-10-21.
- Sachau, Eduard; Bīrūni,¯Muḥammad ibn Aḥmad (1888). *Alberuni's India: An Account of the Religion, Philosophy, Literature, Geography, Chronology, Astronomy and Astrology of India, Vol. 1*. London: Kegan, Paul, Trench, Trübner & Co. Retrieved 2016-02-28.
- Serre, Jean-Pierre (1996) [1973]. *A Course in Arithmetic*. Graduate texts in mathematics. **7**. Springer. ISBN 978-0-387-90040-7.
- Smith, D. E. (1958). *History of Mathematics, Vol I*. New York: Dover Publications.
- Tannery, Paul; Henry, Charles (eds.); Fermat, Pierre de (1891). *Oeuvres de Fermat*. (4 Vols.) (in French and Latin). Paris: Imprimerie Gauthier-Villars et Fils. Volume 1 Volume 2 Volume 3 Volume 4 (1912)
- Iamblichus; Taylor, Thomas (trans.) (1818). *Life of Pythagoras or, Pythagoric Life*. London: J. M. Watkins. For other editions, see Iamblichus#List of editions and translations
- Truesdell, C. A. (1984). "Leonard Euler, Supreme Geometer". In Hewlett, John (trans.). *Leonard Euler, Elements of Algebra* (reprint of 1840 5th ed.). New York: Springer-Verlag. ISBN 978-0-387-96014-2. This Google books preview of *Elements of algebra* lacks Truesdell's intro, which is reprinted (slightly abridged) in the following book:
- Truesdell, C. A. (2007). "Leonard Euler, Supreme Geometer". In Dunham, William. *The Genius of Euler: reflections on his life and work*. Volume 2 of MAA tercentenary Euler celebration. New York: Mathematical Association of America. ISBN 978-0-88385-558-4. Retrieved 2016-02-28.
- Varadarajan, V. S. (2006). *Euler Through Time: A New Look at Old Themes*. American Mathematical Society. ISBN 978-0-8218-3580-7. Retrieved 2016-02-28.
- Vardi, Ilan (April 1998). "Archimedes' Cattle Problem" (PDF). *American Mathematical Monthly*. **105** (4): 305–319. doi:10.2307/2589706.
- van der Waerden, Bartel L.; Dresden, Arnold (trans) (1961). *Science Awakening*. Vol. 1 or Vol 2. New York: Oxford University Press.
- Weil, André (1984). *Number Theory: an Approach Through History – from Hammurapi to Legendre*. Boston: Birkhäuser. ISBN 978-0-8176-3141-3. Retrieved 2016-02-28.

1.11 External links

- Hazewinkel, Michiel, ed. (2001), "Number theory", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Quotations related to Number theory at Wikiquote
- Number Theory Web



Title page of the 1621 edition of Diophantus' Arithmetica, *translated into Latin by Claude Gaspard Bachet de Méziriac.*

Al-Haytham seen by the West: frontispice of Selenographia, *showing Alhasen [sic] representing knowledge through reason, and Galileo representing knowledge through the senses.*

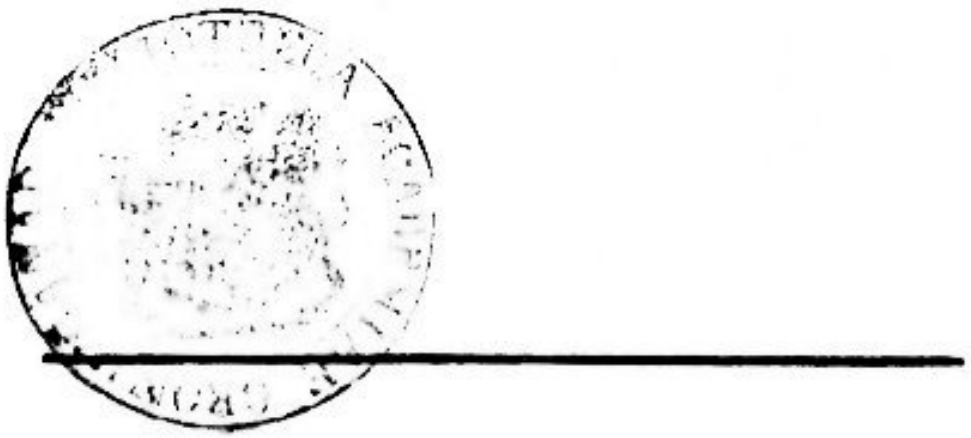
Leonhard Euler

DISQVISITIONES

ARITHMETICAE

AVCTORE

D. CAROLO FRIDERICO GAVSS



LIPSIAE

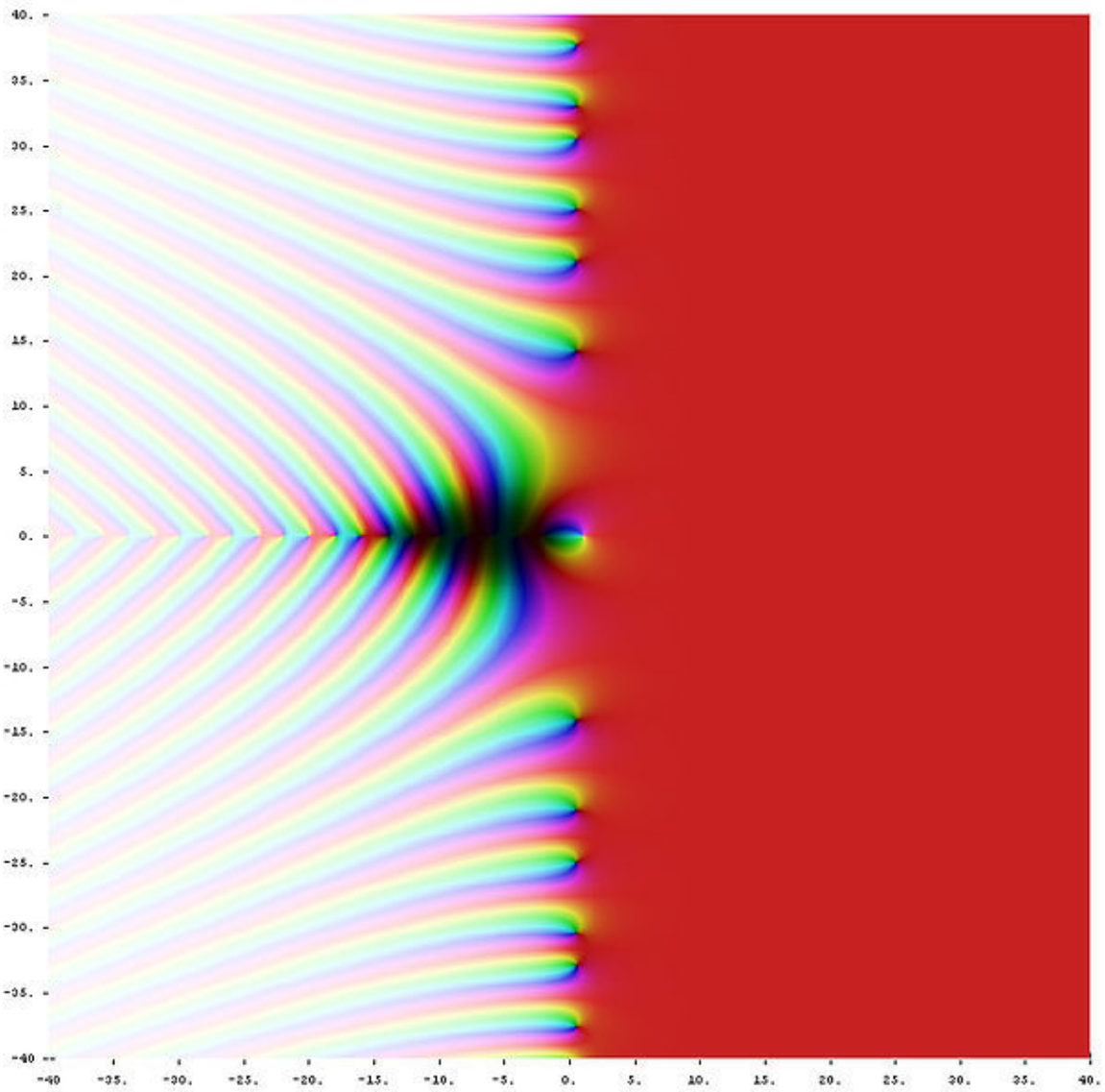
IN COMMISSIS APVD GERH. FLEISCHER, Jun.

1801.

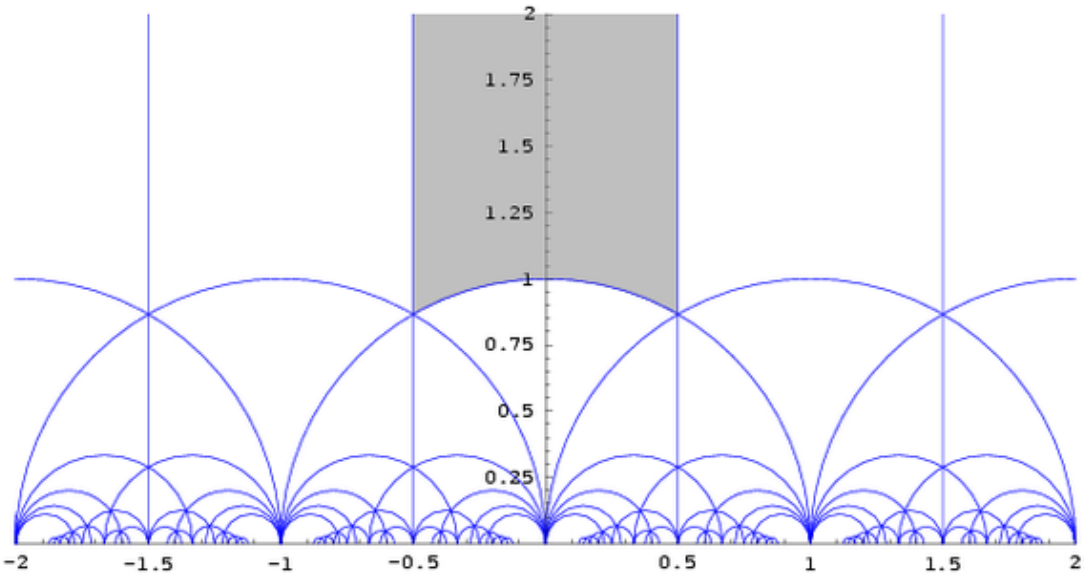
Carl Friedrich Gauss

Ernst Kummer

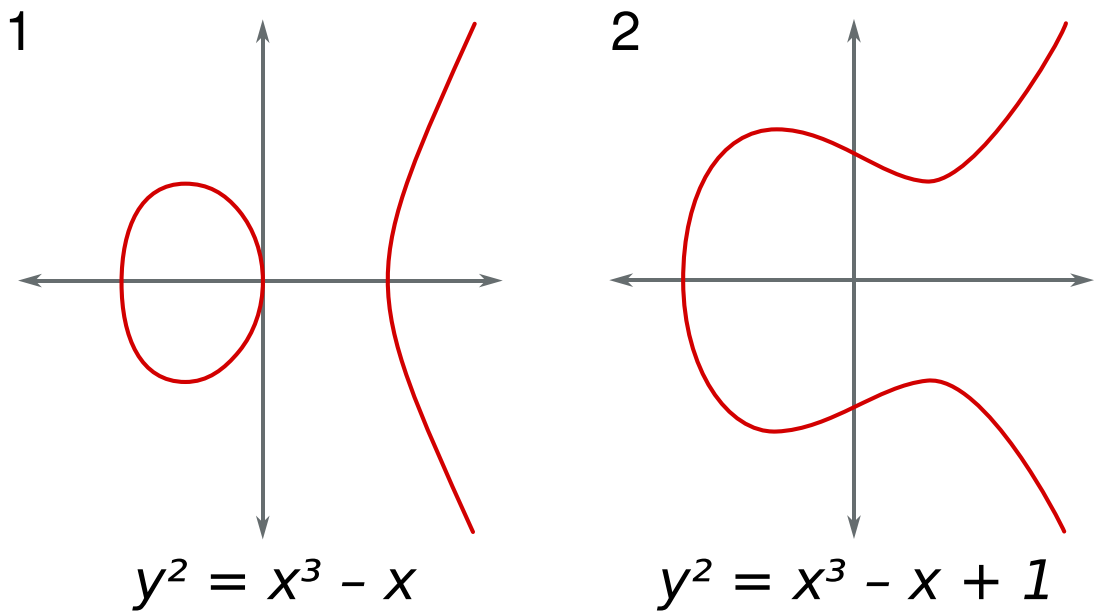
Peter Gustav Lejeune Dirichlet



Riemann zeta function $\zeta(s)$ *in the complex plane. The color of a point* s *gives the value of* $\zeta(s)$*: dark colors denote values close to zero and hue gives the value's argument.*



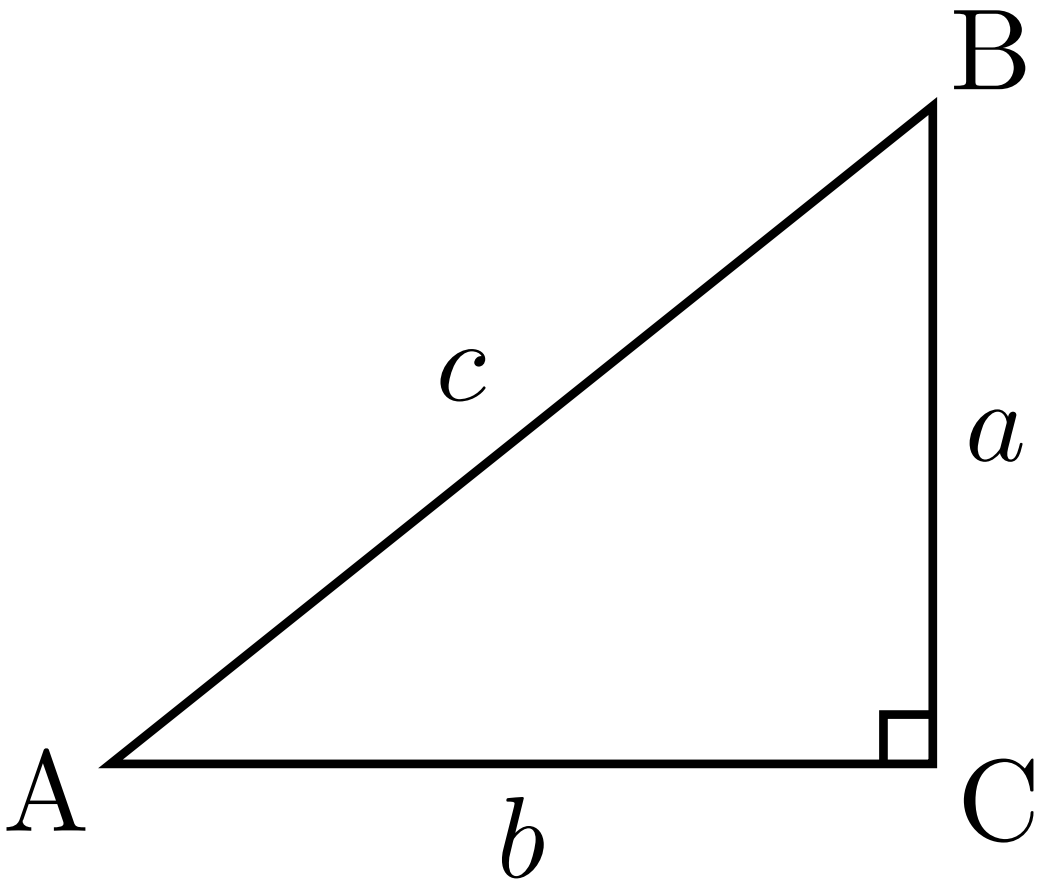
The action of the modular group on the upper half plane. The region in grey is the standard fundamental domain.



Two examples of an elliptic curve, i.e., a curve of genus 1 having at least one rational point. (Either graph can be seen as a slice of a torus in four-dimensional space.)

Chapter 2

Diophantine equation



Finding all right triangles with integer side-lengths is equivalent to solving the Diophantine equation $a^2 + b^2 = c^2$.

In mathematics, a **Diophantine equation** is a polynomial equation, usually in two or more unknowns, such that only the integer solutions are sought or studied (an integer solution is a solution such that all the unknowns take integer values). A **linear Diophantine equation** is an equation between two sums of monomials of degree zero or one. An **exponential Diophantine equation** is one in which exponents on terms can be unknowns.

Diophantine problems have fewer equations than unknown variables and involve finding integers that work correctly for all equations. In more technical language, they define an algebraic curve, algebraic surface, or more general object, and ask about the lattice points on it.

The word *Diophantine* refers to the Hellenistic mathematician of the 3rd century, Diophantus of Alexandria, who made a study of such equations and was one of the first mathematicians to introduce symbolism into algebra. The mathematical study of Diophantine problems that Diophantus initiated is now called **Diophantine analysis**.

While individual equations present a kind of puzzle and have been considered throughout history, the formulation of general theories of Diophantine equations (beyond the theory of quadratic forms) was an achievement of the twentieth century.

2.1 Examples

In the following Diophantine equations, w, x, y, and z are the unknowns and the other letters are given constants:

2.2 Linear Diophantine equations

2.2.1 One equation

The simplest linear Diophantine equation takes the form $ax + by = c$, where a, b and c are given integers. The solutions are described by the following theorem:

> *This Diophantine equation has a solution* (where x and y are integers) *if and only if* c *is a multiple of the greatest common divisor of* a *and* b. *Moreover, if* (x, y) *is a solution, then the other solutions have the form* $(x + kv, y - ku)$, *where* k *is an arbitrary integer, and* u *and* v *are the quotients of* a *and* b *(respectively) by the greatest common divisor of* a *and* b.

Proof: If d is this greatest common divisor, Bézout's identity asserts the existence of integers e and f such that $ae + bf = d$. If c is a multiple of d, then $c = dh$ for some integer h, and (eh, fh) is a solution. On the other hand, for every pair of integers x and y, the greatest common divisor d of a and b divides $ax + by$. Thus, if the equation has a solution, then c must be a multiple of d. If $a = ud$ and $b = vd$, then for every solution (x, y), we have

$$a(x + kv) + b(y - ku) = ax + by + k(av - bu) = ax + by + k(udv - vdu) = ax + by,$$

showing that $(x + kv, y - ku)$ is another solution. Finally, given two solutions such that $ax_1 + by_1 = ax_2 + by_2 = c$, one deduces that $u(x_2 - x_1) + v(y_2 - y_1) = 0$. As u and v are coprime, Euclid's lemma shows that there exists an integer k such that $x_2 - x_1 = kv$ and $y_2 - y_1 = -ku$. Therefore, $x_2 = x_1 + kv$ and $y_2 = y_1 - ku$, which completes the proof.

2.2.2 Chinese remainder theorem

The Chinese remainder theorem describes an important class of linear Diophantine systems of equations: let n_1, …, nk be k pairwise coprime integers greater than one, a_1, …, ak be k arbitrary integers, and N be the product $n_1 \cdots nk$. The Chinese remainder theorem asserts that the following linear Diophantine system has exactly one solution $(x, x_1$, …, $xk)$ such that $0 \le x < N$, and that the other solutions are obtained by adding to x a multiple of N:

$$\begin{aligned} x &= a_1 + n_1\, x_1 \\ &\vdots \\ x &= a_k + n_k\, x_k \end{aligned}$$

2.2.3 System of linear Diophantine equations

More generally, every system of linear Diophantine equations may be solved by computing the Smith normal form of its matrix, in a way that is similar to the use of the reduced row echelon form to solve a system of linear equations over a field. Using matrix notation every system of linear Diophantine equations may be written

$$A\,X = C,$$

where A is an $m \times n$ matrix of integers, X is an $n \times 1$ column matrix of unknowns and C is an $m \times 1$ column matrix of integers.

The computation of the Smith normal form of A provides two unimodular matrices (that is matrices that are invertible over the integers and have ±1 as determinant) U and V of respective dimensions $m \times m$ and $n \times n$, such that the matrix

$$B = [bi,j] = UAV$$

is such that bi,i is not zero for i not greater than some integer k, and all the other entries are zero. The system to be solved may thus be rewritten as

$$B\,(V^{-1}X) = UC.$$

Calling yi the entries of $V^{-1}X$ and di those of $D = UC$, this leads to the system

$$bi,i\ yi = di \text{ for } 1 \le i \le k,$$
$$0\ yi = di \text{ for } k < i \le n.$$

This system is equivalent to the given one in the following sense: A column matrix of integers x is a solution of the given system if and only if $x = Vy$ for some column matrix of integers y such that $By = D$.

It follows that the system has a solution if and only if bi,i divides di for $i \le k$ and $di = 0$ for $i > k$. If this condition is fulfilled, the solutions of the given system are

$$V \begin{bmatrix} \frac{d_1}{b_{1,1}} \\ \vdots \\ \frac{d_k}{b_{k,k}} \\ h_{k+1} \\ \vdots \\ h_n \end{bmatrix},$$

where hk_{+1}, ..., hn are arbitrary integers.

Hermite normal form may also be used for solving systems of linear Diophantine equations. However, Hermite normal form does not directly provide the solutions; to get the solutions from the Hermite normal form, one has to successively solve several linear equations. Nevertheless, Richard Zippel wrote that the Smith normal form "is somewhat more than is actually needed to solve linear diophantine equations. Instead of reducing the equation to diagonal form, we only need to make it triangular, which is called the Hermite normal form. The Hermite normal form is substantially easier to compute than the Smith normal form."[5]

Integer linear programming amounts to finding some integer solutions (optimal in some sense) of linear systems that include also inequations. Thus systems of linear Diophantine equations are basic in this context, and textbooks on integer programming usually have a treatment of systems of linear Diophantine equations.[6]

2.3 Diophantine analysis

2.3.1 Typical questions

The questions asked in Diophantine analysis include:

1. Are there any solutions?
2. Are there any solutions beyond some that are easily found by inspection?

3. Are there finitely or infinitely many solutions?
4. Can all solutions be found in theory?
5. Can one in practice compute a full list of solutions?

These traditional problems often lay unsolved for centuries, and mathematicians gradually came to understand their depth (in some cases), rather than treat them as puzzles.

2.3.2 Typical problem

The given information is that a father's age is 1 less than twice that of his son, and that the digits AB making up the father's age are reversed in the son's age (i.e. BA). This leads to the equation $10A + B = 2(10B + A) - 1$, thus $19B - 8A = 1$. Inspection gives the result $A = 7$, $B = 3$, and thus AB equals 73 years and BA equals 37 years. One may easily show that there is not any other solution with A and B positive integers less than 10.

2.3.3 17th and 18th centuries

In 1637, Pierre de Fermat scribbled on the margin of his copy of *Arithmetica*: "It is impossible to separate a cube into two cubes, or a fourth power into two fourth powers, or in general, any power higher than the second into two like powers." Stated in more modern language, "The equation $a^n + b^n = c^n$ has no solutions for any n higher than 2." And then he wrote, intriguingly: "I have discovered a truly marvelous proof of this proposition, which this margin is too narrow to contain." Such a proof eluded mathematicians for centuries, however, and as such his statement became famous as Fermat's Last Theorem. It wasn't until 1995 that it was proven by the British mathematician Andrew Wiles.

In 1657, Fermat attempted to solve the Diophantine equation $61x^2 + 1 = y^2$ (solved by Brahmagupta over 1000 years earlier). The equation was eventually solved by Euler in the early 18th century, who also solved a number of other Diophantine equations. The smallest solution of this equation in positive integers is $x = 226153980$, $y = 1766319049$ (see Chakravala method).

2.3.4 Hilbert's tenth problem

Main article: Hilbert's tenth problem

In 1900, David Hilbert proposed the solvability of all Diophantine equations as the tenth of his fundamental problems. In 1970, Yuri Matiyasevich solved it negatively, by proving that a general algorithm for solving all Diophantine equations cannot exist.

2.3.5 Diophantine geometry

Diophantine geometry, which is the application of techniques from algebraic geometry in this field, has continued to grow as a result; since treating arbitrary equations is a dead end, attention turns to equations that also have a geometric meaning. The central idea of Diophantine geometry is that of a rational point, namely a solution to a polynomial equation or a system of polynomial equations, which is a vector in a prescribed field K, when K is *not* algebraically closed.

2.3.6 Modern research

One of the few general approaches is through the Hasse principle. Infinite descent is the traditional method, and has been pushed a long way.

The depth of the study of general Diophantine equations is shown by the characterisation of Diophantine sets as equivalently described as recursively enumerable. In other words, the general problem of Diophantine analysis is blessed or cursed with universality, and in any case is not something that will be solved except by re-expressing it in other terms.

The field of Diophantine approximation deals with the cases of *Diophantine inequalities*. Here variables are still supposed to be integral, but some coefficients may be irrational numbers, and the equality sign is replaced by upper and lower bounds.

The most celebrated single question in the field, the conjecture known as Fermat's Last Theorem, was solved by Andrew Wiles[7] but using tools from algebraic geometry developed during the last century rather than within number theory where the conjecture was originally formulated. Other major results, such as Faltings' theorem, have disposed of old conjectures.

2.3.7 Infinite Diophantine equations

An example of an infinite diophantine equation is:

$$n = a^2 + 2b^2 + 3c^2 + 4d^2 + 5e^2 + \ldots,$$

which can be expressed as "How many ways can a given integer n be written as the sum of a square plus twice a square plus thrice a square and so on?" The number of ways this can be done for each n forms an integer sequence. Infinite Diophantine equations are related to theta functions and infinite dimensional lattices. This equation always has a solution for any positive n. Compare this to:

$$n = a^2 + 4b^2 + 9c^2 + 16d^2 + 25e^2 + \ldots,$$

which does not always have a solution for positive n.

2.4 Exponential Diophantine equations

If a Diophantine equation has as an additional variable or variables occurring as exponents, it is an exponential Diophantine equation. Examples include the Ramanujan–Nagell equation, $2^n - 7 = x^2$, and the equation of the Fermat-Catalan conjecture and Beal's conjecture, $a^m + b^n = c^k$ with inequality restrictions on the exponents. A general theory for such equations is not available; particular cases such as Catalan's conjecture have been tackled. However, the majority are solved via ad hoc methods such as Størmer's theorem or even trial and error.

2.5 See also

- Kuṭṭaka, Aryabhata's algorithm for solving linear Diophantine equations in two unknowns

2.6 Notes

[1] "Quotations by Hardy". Gap.dcs.st-and.ac.uk. Retrieved 20 November 2012.

[2] Everest, G.; Ward, Thomas (2006), *An Introduction to Number Theory*, Graduate Texts in Mathematics, **232**, Springer, p. 117, ISBN 9781846280443.

[3] Wiles, Andrew (1995). "Modular elliptic curves and Fermat's Last Theorem" (PDF). *Annals of Mathematics*. Annals of Mathematics. **141** (3): 443–551. doi:10.2307/2118559. JSTOR 2118559. OCLC 37032255.

[4] Noam Elkies (1988). "On $A^4 + B^4 + C^4 = D^4$". *Mathematics of Computation*. **51** (184): 825–835. doi:10.2307/2008781. JSTOR 2008781. MR 0930224.

[5] Richard Zippel (1993). *Effective Polynomial Computation*. Springer Science & Business Media. p. 50. ISBN 978-0-7923-9375-7.

[6] Alexander Bockmayr, Volker Weispfenning (2001). "Solving Numerical Constraints". In John Alan Robinson and Andrei Voronkov. *Handbook of Automated Reasoning Volume I*. Elsevier and MIT Press. p. 779. ISBN 0-444-82949-0 (Elsevier) ISBN 0-262-18221-1 (MIT Press).

[7] Solving Fermat: Andrew Wiles

2.7 References

- Mordell, L. J. (1969). *Diophantine equations*. Pure and Applied Mathematics. **30**. Academic Press. ISBN 0-12-506250-8. Zbl 0188.34503.
- Schmidt, Wolfgang M. (1991). *Diophantine approximations and Diophantine equations*. Lecture Notes in Mathematics. **1467**. Berlin: Springer-Verlag. ISBN 3-540-54058-X. Zbl 0754.11020.
- Shorey, T. N.; Tijdeman, R. (1986). *Exponential Diophantine equations*. Cambridge Tracts in Mathematics. **87**. Cambridge University Press. ISBN 0-521-26826-5. Zbl 0606.10011.
- Smart, Nigel P. (1998). *The algorithmic resolution of Diophantine equations*. London Mathematical Society Student Texts. **41**. Cambridge University Press. ISBN 0-521-64156-X. Zbl 0907.11001.
- Stillwell, John (2004). *Mathematics and its History* (Second ed.). Springer Science + Business Media Inc. ISBN 0-387-95336-1.

2.8 Further reading

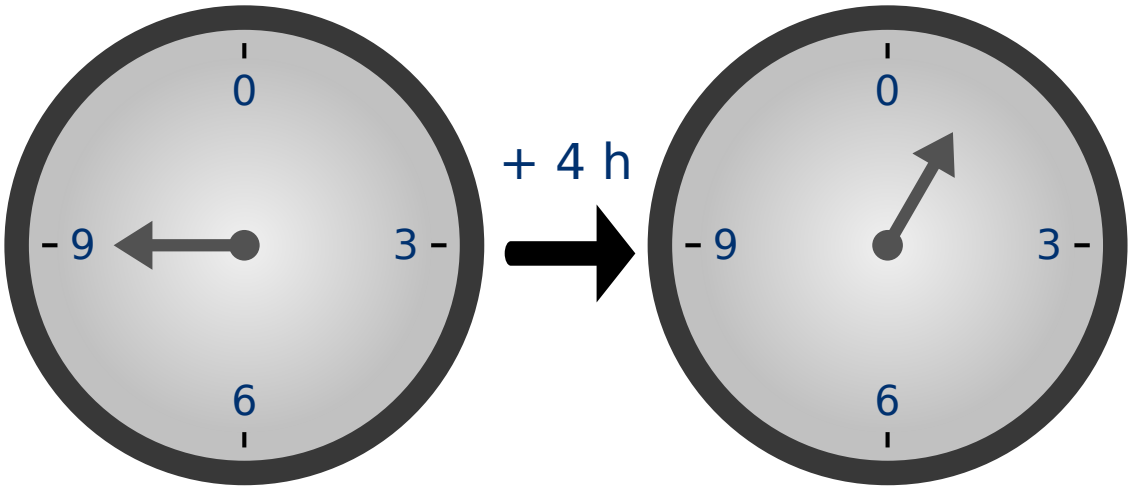
- Dickson, Leonard Eugene (2005) [1920]. *History of the Theory of Numbers. Volume II: Diophantine analysis*. Mineola, NY: Dover Publications. ISBN 978-0-486-44233-4. MR 0245500. Zbl 1214.11002.

2.9 External links

- Diophantine Equation. From MathWorld at Wolfram Research.
- Diophantine Equation. From PlanetMath.
- Hazewinkel, Michiel, ed. (2001), "Diophantine equations", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Dario Alpern's Online Calculator. Retrieved 18 March 2009

Chapter 3

Modular arithmetic



Time-keeping on this clock uses arithmetic modulo 12.

In mathematics, **modular arithmetic** is a system of arithmetic for integers, where numbers "wrap around" upon reaching a certain value—the **modulus** (plural **moduli**). The modern approach to modular arithmetic was developed by Carl Friedrich Gauss in his book *Disquisitiones Arithmeticae*, published in 1801.

A familiar use of modular arithmetic is in the 12-hour clock, in which the day is divided into two 12-hour periods. If the time is 7:00 now, then 8 hours later it will be 3:00. Usual addition would suggest that the later time should be 7 + 8 = 15, but this is not the answer because clock time "wraps around" every 12 hours; in 12-hour time, there is no "15 o'clock". Likewise, if the clock starts at 12:00 (noon) and 21 hours elapse, then the time will be 9:00 the next day, rather than 33:00. Because the hour number starts over after it reaches 12, this is arithmetic *modulo* 12. According to the definition below, 12 is congruent not only to 12 itself, but also to 0, so the time called "12:00" could also be called "0:00", since 12 is congruent to 0 modulo 12.

3.1 Congruence relation

This section is about the *(mod n)* notation. For the binary *mod* operation, see modulo operation.

Modular arithmetic can be handled mathematically by introducing a congruence relation on the integers that is compatible with the operations on integers: addition, subtraction, and multiplication. For a positive integer n, two integers a and b are said to be **congruent modulo *n***, written:

$$a \equiv b \pmod{n},$$

if their difference $a - b$ is an integer multiple of n (or n divides $a - b$). The number n is called the *modulus* of the congruence.

For example,

$$38 \equiv 14 \pmod{12}$$

because 38 − 14 = 24, which is a multiple of 12.

The same rule holds for negative values:

$$\begin{aligned} -8 &\equiv 7 \pmod{5} \\ 2 &\equiv -3 \pmod{5} \\ -3 &\equiv -8 \pmod{5}. \end{aligned}$$

Equivalently, $a \equiv b \bmod n$ can also be thought of as asserting that the remainders of the division of both a and b by n are the same. For instance:

$$38 \equiv 14 \pmod{12}$$

because both 38 and 14 have the same remainder 2 when divided by 12. It is also the case that 38 − 14 = 24 is an integer multiple of 12, which agrees with the prior definition of the congruence relation.

A remark on the notation: Because it is common to consider several congruence relations for different moduli at the same time, the modulus is incorporated in the notation. In spite of the ternary notation, the congruence relation for a given modulus is binary. This would have been clearer if the notation $a \equiv_n b$ had been used, instead of the common traditional notation.

The properties that make this relation a congruence relation (respecting addition, subtraction, and multiplication) are the following.

If

$$a_1 \equiv b_1 \pmod{n}$$

and

$$a_2 \equiv b_2 \pmod{n},$$

then:

- $a_1 + a_2 \equiv b_1 + b_2 \pmod{n}$
- $a_1 - a_2 \equiv b_1 - b_2 \pmod{n}$.

The above two properties would still hold if the theory were expanded to include all real numbers, that is if a_1, a_2, b_1, b_2, n were not necessarily all integers. The next property, however, would fail if these variables were not all integers:

- $a_1 a_2 \equiv b_1 b_2 \pmod{n}$.

3.2 Remainders

Main article: Modulo operation

The notion of modular arithmetic is related to that of the remainder in Euclidean division. The operation of finding the remainder is sometimes referred to as the modulo operation, and denoted with "mod" used as an infix operator. For example, the remainder of the division of 14 by 12 is denoted by 14 mod 12; as this remainder is 2, we have 14 mod 12 = 2.

The congruence, indicated by "≡" followed by "mod" between parentheses, means that the operator "mod", applied to both members, gives the same result. That is

$$A \equiv B \pmod{n}$$

is equivalent to

$$A \bmod n = B \bmod n.$$

The fundamental property of multiplication in modular arithmetic may thus be written

$$(a \bmod n)\,(b \bmod n) \equiv ab \pmod{n},$$

or, equivalently,

$$((a \bmod n)\,(b \bmod n)) \bmod n = (ab) \bmod n.$$

In computer science, it is the remainder operator that is usually indicated by either "%" (for example, in C, C++, Java, JavaScript, Perl, Python and Scala) or "mod" (for example, in Pascal, BASIC, SQL, Haskell, ABAP, and MATLAB), with exceptions (for example, Excel). These operators are commonly pronounced as "mod", but it is specifically a remainder that is computed (since in C++ a negative number will be returned if the first argument is negative, and in Python a negative number will be returned if the second argument is negative). The function *modulo* instead of *mod*, like 38 ≡ 14 (modulo 12) is sometimes used to indicate the common residue rather than a remainder (for example, in Ruby). For details of the specific operations defined in different languages, see the modulo operation page.

3.3 Residue systems

Each residue class modulo *n* may be represented by any one of its members, although we usually represent each residue class by the smallest nonnegative integer which belongs to that class (since this is the proper remainder which results from division). Any two members of different residue classes modulo *n* are incongruent modulo *n*. Furthermore, every integer belongs to one and only one residue class modulo *n*.[1]

The set of integers {0, 1, 2, …, *n* − 1} is called the **least residue system modulo *n***. Any set of *n* integers, no two of which are congruent modulo *n*, is called a **complete residue system modulo *n***.

It is clear that the least residue system is a complete residue system, and that a complete residue system is simply a set containing precisely one representative of each residue class modulo *n*.[2] The least residue system modulo 4 is {0, 1, 2, 3}. Some other complete residue systems modulo 4 are:

- {1, 2, 3, 4}
- {13, 14, 15, 16}
- {−2, −1, 0, 1}
- {−13, 4, 17, 18}
- {−5, 0, 6, 21}
- {27, 32, 37, 42}

Some sets which are *not* complete residue systems modulo 4 are:

- {−5, 0, 6, 22} since 6 is congruent to 22 modulo 4.
- {5, 15} since a complete residue system modulo 4 must have exactly 4 incongruent residue classes.

3.3.1 Reduced residue systems

Main article: Reduced residue system

Any set of $\varphi(n)$ integers that are relatively prime to n and that are mutually incongruent modulo n, where $\varphi(n)$ denotes Euler's totient function, is called a **reduced residue system modulo *n***.[3] The example above, {5,15} is an example of a reduced residue system modulo 4.

3.4 Congruence classes

Like any congruence relation, congruence modulo n is an equivalence relation, and the equivalence class of the integer a, denoted by an, is the set $\{\dots\ a - 2n, a - n, a, a + n, a + 2n, \dots\}$. This set, consisting of the integers congruent to a modulo n, is called the **congruence class** or **residue class** or simply **residue** of the integer a, modulo n. When the modulus n is known from the context, that **residue** may also be denoted $[a]$.

3.5 Integers modulo *n*

The set of all congruence classes of the integers for a modulus n is called the **ring of integers modulo *n***,[4] and is denoted $\mathbb{Z}/n\mathbb{Z}$, $\mathbb{Z}/n$, or $\mathbb{Z}_n$. The notation $\mathbb{Z}_n$ is, however, not recommended because it can be confused with the set of n-adic integers. The set is defined as follows.

$$\mathbb{Z}/n\mathbb{Z} = \{\overline{a}_n | a \in \mathbb{Z}\}\,.$$

When $n \neq 0$, $\mathbb{Z}/n\mathbb{Z}$ has n elements, and can be written as:

$$\mathbb{Z}/n\mathbb{Z} = \left\{\overline{0}_n, \overline{1}_n, \overline{2}_n, \dots, \overline{n-1}_n\right\}.$$

When $n = 0$, $\mathbb{Z}/n\mathbb{Z}$ does not have zero elements; rather, it is isomorphic to $\mathbb{Z}$, since $a_0 = \{a\}$.

We can define addition, subtraction, and multiplication on $\mathbb{Z}/n\mathbb{Z}$ by the following rules:

- $\overline{a}_n + \overline{b}_n = \overline{(a+b)}_n$
- $\overline{a}_n - \overline{b}_n = \overline{(a-b)}_n$
- $\overline{a}_n \overline{b}_n = \overline{(ab)}_n$.

The verification that this is a proper definition uses the properties given before.

In this way, $\mathbb{Z}/n\mathbb{Z}$ becomes a commutative ring. For example, in the ring $\mathbb{Z}/24\mathbb{Z}$, we have

$$\overline{12}_{24} + \overline{21}_{24} = \overline{9}_{24}$$

as in the arithmetic for the 24-hour clock.

The notation $\mathbb{Z}/n\mathbb{Z}$ is used, because it is the quotient ring of $\mathbb{Z}$ by the ideal $n\mathbb{Z}$ containing all integers divisible by n, where $0\mathbb{Z}$ is the singleton set $\{0\}$. Thus $\mathbb{Z}/n\mathbb{Z}$ is a field when $n\mathbb{Z}$ is a maximal ideal, that is, when n is prime.

In terms of groups, the residue class an is the coset of a in the quotient group $\mathbb{Z}/n\mathbb{Z}$, a cyclic group.[5]

The set $\mathbb{Z}/n\mathbb{Z}$ has a number of important mathematical properties that are foundational to various branches of mathematics.

Rather than excluding the special case $n = 0$, it is more useful to include $\mathbb{Z}/0\mathbb{Z}$ (which, as mentioned before, is isomorphic to the ring $\mathbb{Z}$ of integers), for example, when discussing the characteristic of a ring.

The ring of integers modulo n is a finite field if and only if n is prime. If n is a non-prime prime power, there exists a unique (up to isomorphism) finite field GF(n) with n elements, which must not be confused with the ring of integers modulo n, although they have the same number of elements.

3.6 Modular exponentiation

It is a very useful fact that the value $y \equiv a^x \pmod{n}$ can be computed very efficiently even when the number a^x is too large to compute (see modular exponentiation for details.) y can be computed without at any time dealing with numbers bigger than n^2.

3.7 Applications

Modular arithmetic is referenced in number theory, group theory, ring theory, knot theory, abstract algebra, computer algebra, cryptography, computer science, chemistry and the visual and musical arts.

It is one of the foundations of number theory, touching on almost every aspect of its study, and provides key examples for group theory, ring theory and abstract algebra.

Modular arithmetic is often used to calculate checksums that are used within identifiers. International Bank Account Numbers (IBANs), for example, make use of modulo 97 arithmetic to trap user input errors in bank account numbers.

In cryptography, modular arithmetic directly underpins public key systems such as RSA and Diffie–Hellman, and provides finite fields which underlie elliptic curves, and is used in a variety of symmetric key algorithms including Advanced Encryption Standard (AES), International Data Encryption Algorithm (IDEA), and RC4. RSA and Diffie–Hellman use modular exponentiation.

In computer algebra, modular arithmetic is commonly used to limit the size of integer coefficients in intermediate calculations and data. It is used in polynomial factorization, a problem for which all known efficient algorithms use modular arithmetic. It is used by the most efficient implementations of polynomial greatest common divisor, exact linear algebra and Gröbner basis algorithms over the integers and the rational numbers.

In computer science, modular arithmetic is often applied in bitwise operations and other operations involving fixed-width, cyclic data structures. The modulo operation, as implemented in many programming languages and calculators, is an application of modular arithmetic that is often used in this context. XOR is the sum of 2 bits, modulo 2.

In chemistry, the last digit of the CAS registry number (a number which is unique for each chemical compound) is a check digit, which is calculated by taking the last digit of the first two parts of the CAS registry number times 1, the previous digit times 2, the previous digit times 3 etc., adding all these up and computing the sum modulo 10.

In music, arithmetic modulo 12 is used in the consideration of the system of twelve-tone equal temperament, where octave and enharmonic equivalency occurs (that is, pitches in a 1:2 or 2:1 ratio are equivalent, and C-sharp is considered the same as D-flat).

The method of casting out nines offers a quick check of decimal arithmetic computations performed by hand. It is based on modular arithmetic modulo 9, and specifically on the crucial property that $10 \equiv 1 \pmod{9}$.

Arithmetic modulo 7 is used in algorithms that determine the day of the week for a given date. In particular, Zeller's congruence and the doomsday algorithm make heavy use of modulo-7 arithmetic.

More generally, modular arithmetic also has application in disciplines such as law (see for example, apportionment), economics, (see for example, game theory) and other areas of the social sciences, where proportional division and allocation of resources plays a central part of the analysis.

3.8 Computational complexity

Since modular arithmetic has such a wide range of applications, it is important to know how hard it is to solve a system of congruences. A linear system of congruences can be solved in polynomial time with a form of Gaussian elimination, for details see linear congruence theorem. Algorithms, such as Montgomery reduction, also exist to allow simple arithmetic operations, such as multiplication and exponentiation modulo n, to be performed efficiently on large numbers.

Some operations, like finding a Discrete logarithm or a Quadratic congruence appear to be as hard as Integer factorization and thus are a starting point for Cryptographic alogrithms and Encryption. These problems might be NP-intermediate.

Solving a system of non-linear modular arithmetic equations is NP-complete.[6]

3.9 Example implementations

Below are two reasonably fast C functions for performing modular multiplication on unsigned integers not larger than 63 bits, without overflow of the transient operations. An algorithmic way to compute $a \times b \pmod{m}$:

```
uint64_t mul_mod(uint64_t a, uint64_t b, uint64_t m) { uint64_t d = 0, mp2 = m >> 1; int i; if (a >= m) a %= m; if (b >= m) b %= m; for (i = 0; i < 64; ++i) { d = (d > mp2) ? (d << 1) - m : d << 1; if (a & 0x8000000000000000ULL) d += b; if (d > m) d -= m; a <<= 1; } return d; }
```

On computer architectures where an extended precision format with at least 64 bits of mantissa is available (such as the long double type of most x86 C compilers), the following routine is faster than any algorithmic solution, by employing the trick that, by hardware, floating-point multiplication results in the most significant bits of the product kept, while integer multiplication results in the least significant bits kept:

```
uint64_t mul_mod(uint64_t a, uint64_t b, uint64_t m) { long double x; uint64_t c; int64_t r; if (a >= m) a %= m; if (b >= m) b %= m; x = a; c = x * b / m; r = (int64_t)(a * b - c * m) % (int64_t)m; return r < 0 ? r + m : r; }
```

However, for both routines to work, m must not exceed 63 bits.

3.10 See also

- Boolean ring
- Circular buffer
- Congruence relation
- Division (mathematics)
- Finite field
- Legendre symbol
- Modular exponentiation
- Modular multiplicative inverse
- Modulo operation
- Number theory
- Pisano period (Fibonacci sequences modulo n)
- Primitive root modulo n
- Quadratic reciprocity
- Quadratic residue
- Rational reconstruction (mathematics)
- Reduced residue system
- Serial number arithmetic (a special case of modular arithmetic)
- Two-element Boolean algebra
- Topics relating to the group theory behind modular arithmetic:
  - Cyclic group
  - Multiplicative group of integers modulo n
- Other important theorems relating to modular arithmetic:

- Carmichael's theorem
- Chinese remainder theorem
- Euler's theorem
- Fermat's little theorem (a special case of Euler's theorem)
- Lagrange's theorem
- Thue's lemma

3.11 Notes

[1] Pettofrezzo & Byrkit (1970, p. 90)

[2] Long (1972, p. 78)

[3] Long (1972, p. 85)

[4] It is a ring, as shown below.

[5] Sengadir T., *Discrete Mathematics and Combinatorics*, p. 293, at Google Books

[6] Garey, M. R.; Johnson, D. S. (1979). *Computers and Intractability, a Guide to the Theory of NP-Completeness*. W. H. Freeman. ISBN 0716710447.

3.12 References

- John L. Berggren. "modular arithmetic". Encyclopædia Britannica.
- Apostol, Tom M. (1976), *Introduction to analytic number theory*, Undergraduate Texts in Mathematics, New York-Heidelberg: Springer-Verlag, ISBN 978-0-387-90163-3, MR 0434929, Zbl 0335.10001. See in particular chapters 5 and 6 for a review of basic modular arithmetic.
- Maarten Bullynck "Modular Arithmetic before C.F. Gauss. Systematisations and discussions on remainder problems in 18th-century Germany"
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Section 31.3: Modular arithmetic, pp. 862–868.
- Anthony Gioia, *Number Theory, an Introduction* Reprint (2001) Dover. ISBN 0-486-41449-3
- Long, Calvin T. (1972), *Elementary Introduction to Number Theory* (2nd ed.), Lexington: D. C. Heath and Company, LCCN 77171950
- Pettofrezzo, Anthony J.; Byrkit, Donald R. (1970), *Elements of Number Theory*, Englewood Cliffs: Prentice Hall, LCCN 71081766
- Sengadir, T. (2009). *Discrete Mathematics and Combinatorics*. Chennai, India: Pearson Education India. ISBN 978-81-317-1405-8. OCLC 778356123.

3.13 External links

- Hazewinkel, Michiel, ed. (2001), "Congruence", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- In this modular art article, one can learn more about applications of modular arithmetic in art.
- Weisstein, Eric W. "Modular Arithmetic". *MathWorld*.
- An article on modular arithmetic on the GIMPS wiki
- Modular Arithmetic and patterns in addition and multiplication tables
- Whitney Music Box—an audio/video demonstration of integer modular math

Chapter 4

Burnside's lemma

Burnside's lemma, sometimes also called **Burnside's counting theorem**, the **Cauchy–Frobenius lemma** or the **orbit-counting theorem**, is a result in group theory which is often useful in taking account of symmetry when counting mathematical objects. Its various eponyms are based on William Burnside, George Pólya, Augustin Louis Cauchy, and Ferdinand Georg Frobenius. The result is not due to Burnside himself, who merely quotes it in his book 'On the Theory of Groups of Finite Order', attributing it instead to Frobenius (1887).[1]

In the following, let G be a finite group that acts on a set X. For each g in G let X^g denote the set of elements in X that are fixed by g (also said to be left invariant by g), i.e. $X^g = \{ x \in X \mid g.x = x \}$. Burnside's lemma asserts the following formula for the number of orbits, denoted $|X/G|$:[2]

$$|X/G| = \frac{1}{|G|} \sum_{g \in G} |X^g|.$$

Thus the number of orbits (a natural number or $+\infty$) is equal to the average number of points fixed by an element of G (which is also a natural number or infinity). If G is infinite, the division by $|G|$ may not be well-defined; in this case the following statement in cardinal arithmetic holds:

$$|G||X/G| = \sum_{g \in G} |X^g|.$$

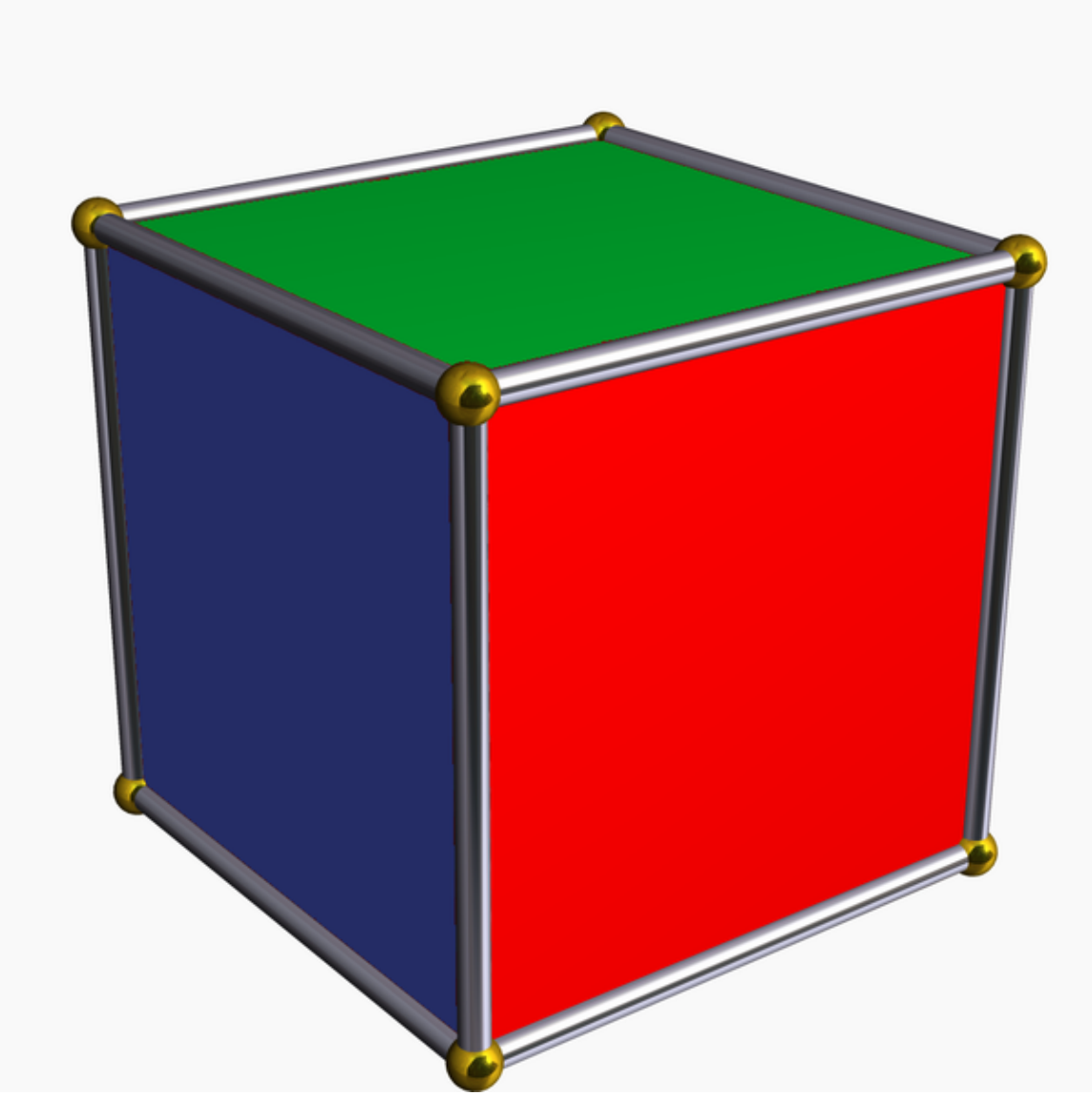
4.1 Example application

The number of rotationally distinct colourings of the faces of a cube using three colours can be determined from this formula as follows.

Let X be the set of 3^6 possible face colour combinations that can be applied to a cube in one particular orientation, and let the rotation group G of the cube act on X in the natural manner. Then two elements of X belong to the same orbit precisely when one is simply a rotation of the other. The number of rotationally distinct colourings is thus the same as the number of orbits and can be found by counting the sizes of the fixed sets for the 24 elements of G.

- one identity element which leaves all 3^6 elements of X unchanged
- six 90-degree face rotations, each of which leaves 3^3 of the elements of X unchanged
- three 180-degree face rotations, each of which leaves 3^4 of the elements of X unchanged
- eight 120-degree vertex rotations, each of which leaves 3^2 of the elements of X unchanged
- six 180-degree edge rotations, each of which leaves 3^3 of the elements of X unchanged

A detailed examination of these automorphisms may be found here.



Cube with coloured faces

The average fix size is thus

$$\frac{1}{24}\left(3^6 + 6 \cdot 3^3 + 3 \cdot 3^4 + 8 \cdot 3^2 + 6 \cdot 3^3\right) = 57.$$

Hence there are 57 rotationally distinct colourings of the faces of a cube in three colours. In general, the number of rotationally distinct colorings of the faces of a cube in n colors is given by

$$\frac{1}{24}\left(n^6 + 3n^4 + 12n^3 + 8n^2\right).$$

4.2 Proof

The first step in the proof of the lemma is to re-express the sum over the group elements $g \in G$ as an equivalent sum over the set elements $x \in X$:

$$\sum_{g\in G}|X^g| = |\{(g,x)\in G\times X \mid g.x = x\}| = \sum_{x\in X}|G_x|.$$

(Here $X^g = \{x \in X \mid g.x = x\}$ is the subset of all points of X fixed by $g \in G$, whereas $Gx = \{g \in G \mid g.x = x\}$ is the stabilizer subgroup of G that fixes the point $x \in X$.)

The orbit-stabilizer theorem says that there is a natural bijection for each $x \in X$ between the orbit of x, $G.x = \{g.x \mid g \in G\} \subseteq X$, and the set of left cosets G/Gx of its stabilizer subgroup Gx. With Lagrange's theorem this implies

$$|G.x| = [G : G_x] = |G|/|G_x|.$$

Our sum over the set X may therefore be rewritten as

$$\sum_{x\in X}|G_x| = \sum_{x\in X}\frac{|G|}{|G.x|} = |G|\sum_{x\in X}\frac{1}{|G.x|}.$$

Finally, notice that X is the disjoint union of all its orbits in X/G, which means the sum over X may be broken up into separate sums over each individual orbit.

$$\sum_{x\in X}\frac{1}{|G.x|} = \sum_{A\in X/G}\sum_{x\in A}\frac{1}{|A|} = \sum_{A\in X/G}1 = |X/G|.$$

Putting everything together gives the desired result:

$$\sum_{g\in G}|X^g| = |G|\cdot|X/G|.$$

This proof is essentially also the proof of the class equation formula, simply by taking the action of G on itself ($X = G$) to be by conjugation, $g.x = gxg^{-1}$, in which case Gx instantiates to the centralizer of x in G.

4.3 History: the lemma that is not Burnside's

William Burnside stated and proved this lemma, attributing it to Frobenius 1887, in his 1897 book on finite groups. But, even prior to Frobenius, the formula was known to Cauchy in 1845. In fact, the lemma was apparently so well known that Burnside simply omitted to attribute it to Cauchy. Consequently, this lemma is sometimes referred to as **the lemma that is not Burnside's**[3] (see also Stigler's law of eponymy). This is less ambiguous than it may seem: Burnside contributed many lemmas to this field.

4.4 See also

- Pólya enumeration theorem

4.5 Notes

[1] Burnside 1897, §119

[2] Rotman 1995, Chapter 3

[3] Neumann 1979

4.6 References

- Burnside, William (1897) *Theory of Groups of Finite Order*, Cambridge University Press, at Project Gutenberg and here at Archive.org. (This is the first edition; the introduction to the second edition contains Burnside's famous *volte face* regarding the utility of representation theory.)
- Frobenius, Ferdinand Georg (1887), "Ueber die Congruenz nach einem aus zwei endlichen Gruppen gebildeten Doppelmodul", *Crelle*, **CI**: 288.
- Neumann, Peter M. (1979), "A lemma that is not Burnside's", *The Mathematical Scientist*, **4** (2): 133–141, ISSN 0312-3685, MR 562002.
- Rotman, Joseph (1995), *An introduction to the theory of groups*, Springer-Verlag, ISBN 0-387-94285-8.

Chapter 5

Gaussian elimination

In linear algebra, **Gaussian elimination** (also known as **row reduction**) is an algorithm for solving systems of linear equations. It is usually understood as a sequence of operations performed on the corresponding matrix of coefficients. This method can also be used to find the rank of a matrix, to calculate the determinant of a matrix, and to calculate the inverse of an invertible square matrix. The method is named after Carl Friedrich Gauss (1777–1855), although it was known to Chinese mathematicians as early as 179 CE (see History section).

To perform row reduction on a matrix, one uses a sequence of elementary row operations to modify the matrix until the lower left-hand corner of the matrix is filled with zeros, as much as possible. There are three types of elementary row operations: 1) Swapping two rows, 2) Multiplying a row by a non-zero number, 3) Adding a multiple of one row to another row. Using these operations, a matrix can always be transformed into an upper triangular matrix, and in fact one that is in row echelon form. Once all of the leading coefficients (the left-most non-zero entry in each row) are 1, and every column containing a leading coefficient has zeros elsewhere, the matrix is said to be in reduced row echelon form. This final form is unique; in other words, it is independent of the sequence of row operations used. For example, in the following sequence of row operations (where multiple elementary operations might be done at each step), the third and fourth matrices are the ones in row echelon form, and the final matrix is the unique reduced row echelon form.

$$\left[\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 1 & 1 & -1 & 1 \\ 3 & 11 & 5 & 35 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 2 & 2 & 8 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 3 & 1 & 9 \\ 0 & -2 & -2 & -8 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & -2 & -3 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Using row operations to convert a matrix into reduced row echelon form is sometimes called **Gauss–Jordan elimination**. Some authors use the term Gaussian elimination to refer to the process until it has reached its upper triangular, or (non-reduced) row echelon form. For computational reasons, when solving systems of linear equations, it is sometimes preferable to stop row operations before the matrix is completely reduced.

5.1 Definitions and example of algorithm

The process of row reduction makes use of elementary row operations, and can be divided into two parts. The first part (sometimes called Forward Elimination) reduces a given system to *row echelon form*, from which one can tell whether there are no solutions, a unique solution, or infinitely many solutions. The second part (sometimes called back substitution) continues to use row operations until the solution is found; in other words, it puts the matrix into *reduced* row echelon form.

Another point of view, which turns out to be very useful to analyze the algorithm, is that row reduction produces a matrix decomposition of the original matrix. The elementary row operations may be viewed as the multiplication on the left of the original matrix by elementary matrices. Alternatively, a sequence of elementary operations that reduces a single row may be viewed as multiplication by a Frobenius matrix. Then the first part of the algorithm computes an LU decomposition, while the second part writes the original matrix as the product of a uniquely determined invertible matrix and a uniquely determined reduced row echelon matrix.

5.1.1 Row operations

See also: Elementary matrix

There are three types of **elementary row operations** which may be performed on the rows of a matrix:

Type 1: Swap the positions of two rows.

Type 2: Multiply a row by a nonzero scalar.

Type 3: Add to one row a scalar multiple of another.

If the matrix is associated to a system of linear equations, then these operations do not change the solution set. Therefore, if one's goal is to solve a system of linear equations, then using these row operations could make the problem easier.

5.1.2 Echelon form

Main article: Row echelon form

For each row in a matrix, if the row does not consist of only zeros, then the left-most non-zero entry is called the *leading coefficient* (or *pivot*) of that row. So if two leading coefficients are in the same column, then a row operation of type 3 (see above) could be used to make one of those coefficients zero. Then by using the row swapping operation, one can always order the rows so that for every non-zero row, the leading coefficient is to the right of the leading coefficient of the row above. If this is the case, then matrix is said to be in **row echelon form**. So the lower left part of the matrix contains only zeros, and all of the zero rows are below the non-zero rows. The word "echelon" is used here because one can roughly think of the rows being ranked by their size, with the largest being at the top and the smallest being at the bottom.

For example, the following matrix is in row echelon form, and its leading coefficients are shown in red.

$$\begin{bmatrix} 0 & \mathbf{2} & 1 & -1 \\ 0 & 0 & \mathbf{3} & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

It is in echelon form because the zero row is at the bottom, and the leading coefficient of the second row (in the third column), is to the right of the leading coefficient of the first row (in the second column).

A matrix is said to be in **reduced row echelon form** if furthermore all of the leading coefficients are equal to 1 (which can be achieved by using the elementary row operation of type 2), and in every column containing a leading coefficient, all of the other entries in that column are zero (which can be achieved by using elementary row operations of type 3).

5.1.3 Example of the algorithm

Suppose the goal is to find and describe the set of solutions to the following system of linear equations:

$$\begin{aligned} 2x + y - z &= 8 && (L_1) \\ -3x - y + 2z &= -11 && (L_2) \\ -2x + y + 2z &= -3 && (L_3) \end{aligned}$$

The table below is the row reduction process applied simultaneously to the system of equations, and its associated augmented matrix. In practice, one does not usually deal with the systems in terms of equations but instead makes use of the augmented matrix, which is more suitable for computer manipulations. The row reduction procedure may be summarized as follows: eliminate x from all equations below L_1 , and then eliminate y from all equations below L_2 . This will put the system into triangular form. Then, using back-substitution, each unknown can be solved for.

The second column describes which row operations have just been performed. So for the first step, the *x* is eliminated from L_2 by adding $\frac{3}{2}L_1$ to L_2 . Next *x* is eliminated from L_3 by adding L_1 to L_3 . These row operations are labelled in the table as

$$L_2 + \frac{3}{2}L_1 \to L_2$$

$$L_3 + L_1 \to L_3.$$

Once *y* is also eliminated from the third row, the result is a system of linear equations in triangular form, and so the first part of the algorithm is complete. From a computational point of view, it is faster to solve the variables in reverse order, a process known as back-substitution. One sees the solution is z = −1, y = 3, and x = 2. So there is a unique solution to the original system of equations.

Instead of stopping once the matrix is in echelon form, one could continue until the matrix is in *reduced* row echelon form, as it is done in the table. The process of row reducing until the matrix is reduced is sometimes referred to as **Gauss-Jordan elimination**, to distinguish it from stopping after reaching echelon form.

5.2 History

The method of Gaussian elimination appears in the Chinese mathematical text Chapter Eight *Rectangular Arrays* of *The Nine Chapters on the Mathematical Art*. Its use is illustrated in eighteen problems, with two to five equations. The first reference to the book by this title is dated to 179 CE, but parts of it were written as early as approximately 150 BCE.[1][2] It was commented on by Liu Hui in the 3rd century.

The method in Europe stems from the notes of Isaac Newton.[3][4] In 1670, he wrote that all the algebra books known to him lacked a lesson for solving simultaneous equations, which Newton then supplied. Cambridge University eventually published the notes as *Arithmetica Universalis* in 1707 long after Newton left academic life. The notes were widely imitated, which made (what is now called) Gaussian elimination a standard lesson in algebra textbooks by the end of the 18th century. Carl Friedrich Gauss in 1810 devised a notation for symmetric elimination that was adopted in the 19th century by professional hand computers to solve the normal equations of least-squares problems.[5] The algorithm that is taught in high school was named for Gauss only in the 1950s as a result of confusion over the history of the subject.[6]

Some authors use the term *Gaussian elimination* to refer only to the procedure until the matrix is in echelon form, and use the term **Gauss-Jordan elimination** to refer to the procedure which ends in reduced echelon form. The name is used because it is a variation of Gaussian elimination as described by Wilhelm Jordan in 1888. However, the method also appears in an article by Clasen published in the same year. Jordan and Clasen probably discovered Gauss–Jordan elimination independently.[7]

5.3 Applications

The historically first application of the row reduction method is for solving systems of linear equations. Here are some other important applications of the algorithm.

5.3.1 Computing determinants

To explain how Gaussian elimination allows the computation of the determinant of a square matrix, we have to recall how the elementary row operations change the determinant:

- Swapping two rows multiplies the determinant by −1
- Multiplying a row by a nonzero scalar multiplies the determinant by the same scalar
- Adding to one row a scalar multiple of another does not change the determinant.

If the Gaussian elimination applied to a square matrix A produces a row echelon matrix B, let d be the product of the scalars by which the determinant has been multiplied, using above rules. Then the determinant of A is the quotient by d of the product of the elements of the diagonal of B: $\det(A) = \prod \text{diag}(B) \,/\, d$.

Computationally, for a $n \times n$ matrix, this method needs only $O(n^3)$ arithmetic operations, while solving by elementary methods requires $O(2^n)$ or $O(n!)$ operations. Even on the fastest computers, the elementary methods are impractical for n above 20.

5.3.2 Finding the inverse of a matrix

See also: Invertible matrix

A variant of Gaussian elimination called Gauss–Jordan elimination can be used for finding the inverse of a matrix, if it exists. If A is a n by n square matrix, then one can use row reduction to compute its inverse matrix, if it exists. First, the n by n identity matrix is augmented to the right of A, forming a n by $2n$ block matrix $[A \mid I]$. Now through application of elementary row operations, find the reduced echelon form of this n by $2n$ matrix. The matrix A is invertible if and only if the left block can be reduced to the identity matrix I; in this case the right block of the final matrix is A^{-1}. If the algorithm is unable to reduce the left block to I, then A is not invertible.

For example, consider the following matrix

$$A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}.$$

To find the inverse of this matrix, one takes the following matrix augmented by the identity, and row reduces it as a 3 by 6 matrix:

$$[A|I] = \left[\begin{array}{ccc|ccc} 2 & -1 & 0 & 1 & 0 & 0 \\ -1 & 2 & -1 & 0 & 1 & 0 \\ 0 & -1 & 2 & 0 & 0 & 1 \end{array}\right].$$

By performing row operations, one can check that the reduced row echelon form of this augmented matrix is:

$$[I|B] = \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{array}\right].$$

One can think of each row operation as the left product by an elementary matrix. Denoting by B the product of these elementary matrices, we showed, on the left, that $BA = I$, and therefore, $B = A^{-1}$. On the right, we kept a record of $BI = B$, which we know is the inverse desired. This procedure for finding the inverse works for square matrices of any size.

5.3.3 Computing ranks and bases

The Gaussian elimination algorithm can be applied to any $m \times n$ matrix A . In this way, for example, some 6×9 matrices can be transformed to a matrix that has a row echelon form like

$$T = \begin{bmatrix} a & * & * & * & * & * & * & * & * \\ 0 & 0 & b & * & * & * & * & * & * \\ 0 & 0 & 0 & c & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & d & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & e \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where the *s are arbitrary entries and *a, b, c, d, e* are nonzero entries. This echelon matrix T contains a wealth of information about A : the rank of A is 5 since there are 5 non-zero rows in T ; the vector space spanned by the columns of A has a basis consisting of the first, third, fourth, seventh and ninth column of A (the columns of *a, b, c, d, e* in T), and the *s tell you how the other columns of A can be written as linear combinations of the basis columns. This is a consequence of the distributivity of the dot product in the expression of a linear map as a matrix.

All of this applies also to the reduced row echelon form, which is a particular row echelon form.

5.4 Computational efficiency

The number of arithmetic operations required to perform row reduction is one way of measuring the algorithm's computational efficiency. For example, to solve a system of n equations for n unknowns by performing row operations on the matrix until it is in echelon form, and then solving for each unknown in reverse order, requires $n(n+1)$ / 2 divisions, $(2n^3 + 3n^2 - 5n)/6$ multiplications, and $(2n^3 + 3n^2 - 5n)/6$ subtractions,[8] for a total of approximately $2n^3$ / 3 operations. Thus it has arithmetic complexity of $O(n^3)$; see Big O notation. This arithmetic complexity is a good measure of the time needed for the whole computation when the time for each arithmetic operation is approximately constant. This is the case when the coefficients are represented by floating point numbers or when they belong to a finite field. If the coefficients are integers or rational numbers exactly represented, the intermediate entries can grow exponentially large, so the bit complexity is exponential.[9] However, there is a variant of Gaussian elimination, called Bareiss algorithm that avoids this exponential growth of the intermediate entries, and, with the same arithmetic complexity of $O(n^3)$, has a bit complexity of $O(n^5)$.

This algorithm can be used on a computer for systems with thousands of equations and unknowns. However, the cost becomes prohibitive for systems with millions of equations. These large systems are generally solved using iterative methods. Specific methods exist for systems whose coefficients follow a regular pattern (see system of linear equations).

To put an n by n matrix into reduced echelon form by row operations, one needs n^3 arithmetic operations; which is approximately 50% more computation steps.[10]

One possible problem is numerical instability, caused by the possibility of dividing by very small numbers. If, for example, the leading coefficient of one of the rows is very close to zero, then to row reduce the matrix one would need to divide by that number so the leading coefficient is 1. This means any error that existed for the number which was close to zero would be amplified. Gaussian elimination is numerically stable for diagonally dominant or positive-definite matrices. For general matrices, Gaussian elimination is usually considered to be stable, when using partial pivoting, even though there are examples of stable matrices for which it is unstable.[11]

5.4.1 Generalizations

The Gaussian elimination can be performed over any field, not just the real numbers.

Gaussian elimination does not generalize in any simple way to higher order tensors (matrices are array representations of order 2 tensors); even computing the rank of a tensor of order greater than 2 is a difficult problem.

5.5 Pseudocode

As explained above, Gaussian elimination writes a given $m \times n$ matrix A uniquely as a product of an invertible $m \times m$ matrix S and a row-echelon matrix T . Here, S is the product of the matrices corresponding to the row operations performed.

The formal algorithm to compute T from A follows. We write $A[i, j]$ for the entry in row i , column j in matrix A with 1 being the first index. The transformation is performed *in place*, meaning that the original matrix A is lost and successively replaced by T .

for k = 1 ... min(m,n): *Find the k-th pivot:* i_max := argmax (i = k ... m, abs(A[i, k])) **if** A[i_max, k] = 0 **error** "Matrix is singular!" **swap rows**(k, i_max) *Do for all rows below pivot:* **for** i = k + 1 ... m: f := A[i, k] / A[k, k] *Do for all remaining elements in current row:* **for** j = k + 1 ... n: A[i, j] := A[i, j] - A[k, j] * f *Fill lower triangular matrix with zeros:* A[i, k] := 0

This algorithm differs slightly from the one discussed earlier, because before eliminating a variable, it first exchanges rows to move the entry with the largest absolute value to the pivot position. Such *partial pivoting* improves the numerical stability of the algorithm; some other variants are used.

Upon completion of this procedure the augmented matrix will be in row-echelon form and may be solved by back-substitution.

With modern computers, Gaussian elimination is not always the fastest algorithm to compute the row echelon form of matrix. There are computer libraries, like BLAS, that exploit the specifics of the computer hardware and of the structure of the matrix to choose the best algorithm automatically.

5.6 Notes

[1] Calinger (1999), pp. 234–236

[2] Timothy Gowers; June Barrow-Green; Imre Leader (8 September 2008). *The Princeton Companion to Mathematics*. Princeton University Press. p. 607. ISBN 978-0-691-11880-2.

[3] Grcar (2011a), pp. 169-172

[4] Grcar (2011b), pp. 783-785

[5] Lauritzen (),P.3

[6] Grcar (2011b), p. 789

[7] Althoen, Steven C.; McLaughlin, Renate (1987), "Gauss–Jordan reduction: a brief history", *The American Mathematical Monthly*, Mathematical Association of America, **94** (2): 130–142, doi:10.2307/2322413, ISSN 0002-9890, JSTOR 2322413

[8] Farebrother (1988), p. 12

[9] Fang, Xin Gui; Havas, George (1997). "On the worst-case complexity of integer Gaussian elimination" (PDF). *Proceedings of the 1997 international symposium on Symbolic and algebraic computation*. ISSAC '97. Kihei, Maui, Hawaii, United States: ACM. pp. 28–31. doi:10.1145/258726.258740. ISBN 0-89791-875-4.

[10] J. B. Fraleigh and R. A. Beauregard, Linear Algebra. Addison-Wesley Publishing Company, 1995, Chapter 10

[11] Golub & Van Loan (1996), §3.4.6

5.7 References

- Atkinson, Kendall A. (1989), *An Introduction to Numerical Analysis* (2nd ed.), New York: John Wiley & Sons, ISBN 978-0-471-50023-0.
- Bolch, Gunter; Greiner, Stefan; de Meer, Hermann; Trivedi, Kishor S. (2006), *Queueing Networks and Markov Chains: Modeling and Performance Evaluation with Computer Science Applications* (2nd ed.), Wiley-Interscience, ISBN 978-0-471-79156-0.
- Calinger, Ronald (1999), *A Contextual History of Mathematics*, Prentice Hall, ISBN 978-0-02-318285-3.
- Farebrother, R.W. (1988), *Linear Least Squares Computations*, STATISTICS: Textbooks and Monographs, Marcel Dekker, ISBN 978-0-8247-7661-9.
- Lauritzen, Niels, *Undergraduate Convexity: From Fourier and Motzkin to Kuhn and Tucker*.
- Golub, Gene H.; Van Loan, Charles F. (1996), *Matrix Computations* (3rd ed.), Johns Hopkins, ISBN 978-0-8018-5414-9.
- Grcar, Joseph F. (2011a), "How ordinary elimination became Gaussian elimination", *Historia Mathematica*, **38** (2): 163–218, arXiv:0907.2397, doi:10.1016/j.hm.2010.06.003
- Grcar, Joseph F. (2011b), "Mathematicians of Gaussian elimination" (PDF), *Notices of the American Mathematical Society*, **58** (6): 782–792

- Higham, Nicholas (2002), *Accuracy and Stability of Numerical Algorithms* (2nd ed.), SIAM, ISBN 978-0-89871-521-7.
- Katz, Victor J. (2004), *A History of Mathematics, Brief Version*, Addison-Wesley, ISBN 978-0-321-16193-2.
- Kaw, Autar; Kalu, Egwu (2010). "Numerical Methods with Applications" (1st ed.). . External link in |publisher= (help); , Chapter 5 deals with Gaussian elimination.
- Lipson, Marc; Lipschutz, Seymour (2001), *Schaum's outline of theory and problems of linear algebra*, New York: McGraw-Hill, pp. 69–80, ISBN 978-0-07-136200-9.
- Press, WH; Teukolsky, SA; Vetterling, WT; Flannery, BP (2007), "Section 2.2", *Numerical Recipes: The Art of Scientific Computing* (3rd ed.), New York: Cambridge University Press, ISBN 978-0-521-88068-8

5.8 External links

- An online tool for performing elementary operations with rows and columns
- WebApp descriptively solving systems of linear equations with Gaussian Elimination
- A program that performs Gaussian elimination similarly to a human working on paper Exact solutions to systems with rational coefficients.
- Gaussian elimination www.math-linux.com.
- Gaussian elimination calculator with Linear Algebra tutorial.
- Gaussian elimination at Holistic Numerical Methods Institute
- LinearEquations.c Gaussian elimination implemented using C language
- Gauss–Jordan elimination Step by step solution of 3 equations with 3 unknowns using the All-Integer Echelon Method
- Gauss-Jordan Elimination calculator for n by m matrices, giving intermediate steps
- WildLinAlg13: Solving a system of linear equations on YouTube provides a very clear, elementary presentation of the method of row reduction.

Chapter 6

Matrix exponential

In mathematics, the **matrix exponential** is a matrix function on square matrices analogous to the ordinary exponential function. Abstractly, the matrix exponential gives the connection between a matrix Lie algebra and the corresponding Lie group.

Let X be an $n \times n$ real or complex matrix. The exponential of X, denoted by e^X or $\exp(X)$, is the $n \times n$ matrix given by the power series

$$e^X = \sum_{k=0}^{\infty} \frac{1}{k!} X^k.$$

where X^0 is defined to be the identity matrix I with the same dimensions as X . [1]

The above series always converges, so the exponential of X is well-defined. If X is a 1×1 matrix the matrix exponential of X is a 1×1 matrix whose single element is the ordinary exponential of the single element of X.

6.1 Properties

Let X and Y be $n \times n$ complex matrices and let a and b be arbitrary complex numbers. We denote the $n \times n$ identity matrix by I and the zero matrix by 0. The matrix exponential satisfies the following properties:[2]

- $e^0 = I$
- $e^{aX} e^{bX} = e^{(a+b)X}$
- $e^X e^{-X} = I$
- If $XY = YX$ then $e^X e^Y = e^Y e^X = e^{(X+Y)}$.
- If Y is invertible then $e^{YXY^{-1}} = Ye^X Y^{-1}$.
- $\exp(X^{\mathrm{T}}) = (\exp X)^{\mathrm{T}}$, where X^{T} denotes the transpose of X. It follows that if X is symmetric then e^X is also symmetric, and that if X is skew-symmetric then e^X is orthogonal.
- $\exp(X^*) = (\exp X)^*$, where X^* denotes the conjugate transpose of X. It follows that if X is Hermitian then e^X is also Hermitian, and that if X is skew-Hermitian then e^X is unitary.
- A Laplace transform of matrix exponentials amounts to the resolvent, $\int_0^\infty dt\, e^{-t}\, e^{tX} = I / (I - X)$.

6.1.1 Linear differential equation systems

Main article: matrix differential equation

One of the reasons for the importance of the matrix exponential is that it can be used to solve systems of linear ordinary differential equations. The solution of

$$\frac{d}{dt}y(t) = Ay(t), \quad y(0) = y_0,$$

where A is a constant matrix, is given by

$$y(t) = e^{At}y_0.$$

The matrix exponential can also be used to solve the inhomogeneous equation

$$\frac{d}{dt}y(t) = Ay(t) + z(t), \quad y(0) = y_0.$$

See the section on applications below for examples.

There is no closed-form solution for differential equations of the form

$$\frac{d}{dt}y(t) = A(t)\,y(t), \quad y(0) = y_0,$$

where A is not constant, but the Magnus series gives the solution as an infinite sum.

6.1.2 The exponential of sums

For any real numbers (scalars) x and y we know that the exponential function satisfies $e^{x+y} = e^x\,e^y$. The same is true for commuting matrices. If matrices X and Y commute (meaning that $XY = YX$), then

$$e^{X+Y} = e^X e^Y\ .$$

However, for matrices that do not commute the above equality does not necessarily hold. In this case the Baker–Campbell–Hausdorff formula can be used to calculate e^{X+Y}.

The converse is not true in general. The equation $e^{X+Y} = e^X\,e^Y$ does not imply that X and Y commute.

For Hermitian matrices there are two notable theorems related to the trace of matrix exponentials.

Golden–Thompson inequality

Main article: Golden–Thompson inequality

If A and H are Hermitian matrices, then

$$\operatorname{tr}\exp(A+H) \leq \operatorname{tr}(\exp(A)\exp(H)).\ [3]$$

Note that there is no requirement of commutativity. There are counterexamples to show that the Golden–Thompson inequality cannot be extended to three matrices – and, in any event, tr(exp(*A*)exp(*B*)exp(*C*)) is not guaranteed to be real for Hermitian *A*, *B*, *C*. However, the next theorem accomplishes this in one sense.

Lieb's theorem

The **Lieb's theorem**, named after Elliott H. Lieb, states that, for a fixed Hermitian matrix H, the function

$$f(A) = \operatorname{tr}\,\exp\left(H + \log A\right)$$

is concave on the cone of positive-definite matrices.[4]

6.1.3 The exponential map

Note that the exponential of a matrix is always an invertible matrix. The inverse matrix of e^X is given by e^{-X}. This is analogous to the fact that the exponential of a complex number is always nonzero. The matrix exponential then gives us a map

$$\exp\colon M_n(\mathbb{C}) \to \mathrm{GL}(n,\mathbb{C})$$

from the space of all *n*×*n* matrices to the general linear group of degree n, i.e. the group of all *n*×*n* invertible matrices. In fact, this map is surjective which means that every invertible matrix can be written as the exponential of some other matrix[5] (for this, it is essential to consider the field **C** of complex numbers and not **R**).

For any two matrices X and Y,

$$\|e^{X+Y} - e^X\| \leq \|Y\| e^{\|X\|} e^{\|Y\|},$$

where $\|\cdot\|$ denotes an arbitrary matrix norm. It follows that the exponential map is continuous and Lipschitz continuous on compact subsets of $Mn(\mathbf{C})$.

The map

$$t \mapsto e^{tX}, \qquad t \in \mathbb{R}$$

defines a smooth curve in the general linear group which passes through the identity element at $t = 0$.

In fact, this gives a one-parameter subgroup of the general linear group since

$$e^{tX} e^{sX} = e^{(t+s)X}.$$

The derivative of this curve (or tangent vector) at a point t is given by

$$\frac{d}{dt} e^{tX} = X e^{tX} = e^{tX} X. \qquad (1)$$

The derivative at $t = 0$ is just the matrix X, which is to say that X generates this one-parameter subgroup.

More generally,[6] for a generic t-dependent exponent, *X(t)*,

Taking the above expression $e^{X(t)}$ outside the integral sign and expanding the integrand with the help of the Hadamard lemma one can obtain the following useful expression for the derivative of the matrix exponent,

$$\left(\frac{d}{dt} e^{X(t)}\right) e^{-X(t)} = \frac{d}{dt} X(t) + \frac{1}{2!}[X(t), \frac{d}{dt} X(t)] + \frac{1}{3!}[X(t), [X(t), \frac{d}{dt} X(t)]] + \cdots$$

Note that the coefficients in the expression above are different from what appears in the exponential. For a closed form, see derivative of the exponential map.

6.1.4 The determinant of the matrix exponential

By Jacobi's formula, for any complex square matrix the following trace identity holds:[7]

In addition to providing a computational tool, this formula demonstrates that a matrix exponential is always an invertible matrix. This follows from the fact that the right hand side of the above equation is always non-zero, and so $\det(e^A) \neq 0$, which implies that e^A must be invertible.

In the real-valued case, the formula also exhibits the map

$$\exp\colon M_n(\mathbb{R}) \to \mathrm{GL}(n,\mathbb{R})$$

to not be surjective, in contrast to the complex case mentioned earlier. This follows from the fact that, for real-valued matrices, the right-hand side of the formula is always positive, while there exist invertible matrices with a negative determinant.

6.2 Computing the matrix exponential

Finding reliable and accurate methods to compute the matrix exponential is difficult, and this is still a topic of considerable current research in mathematics and numerical analysis. Matlab, GNU Octave, and SciPy all use the Padé approximant.[8][9][10]

6.2.1 Diagonalizable case

If a matrix is diagonal:

$$A = \begin{bmatrix} a_1 & 0 & \dots & 0 \\ 0 & a_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & a_n \end{bmatrix}$$

then its exponential can be obtained by exponentiating each entry on the main diagonal:

$$e^A = \begin{bmatrix} e^{a_1} & 0 & \dots & 0 \\ 0 & e^{a_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{a_n} \end{bmatrix}$$

This also allows one to exponentiate diagonalizable matrices. If $A = UDU^{-1}$ and D is diagonal, then $e^A = Ue^DU^{-1}$. Application of Sylvester's formula yields the same result. (To see this, note that addition and multiplication, hence also exponentiation, of diagonal matrices is equivalent to element-wise addition and multiplication, and hence exponentiation; in particular, the "one-dimensional" exponentiation is felt element-wise for the diagonal case.)

6.2.2 Projection case

If P is a projection matrix (i.e. is idempotent), its matrix exponential is $e^P = I + (e-1)P$. This may be derived by expansion of the definition of the exponential function and by use of the idempotency of P:

$$e^P = \sum_{k=0}^{\infty} \frac{P^k}{k!} = I + \left(\sum_{k=1}^{\infty} \frac{1}{k!}\right) P = I + (e-1)P\,.$$

6.2.3 Rotation case

For a simple rotation in which the perpendicular unit vectors a and b specify a plane,[11] the rotation matrix R can be expressed in terms of a similar exponential function involving a generator G and angle θ.[12][13]

$$G = ba^{\mathsf{T}} - ab^{\mathsf{T}} \qquad a^{\mathsf{T}}b = 0$$

$$-G^2 = aa^{\mathsf{T}} + bb^{\mathsf{T}} = P \qquad P^2 = P \qquad PG = GP = G\,,$$

$$\begin{aligned} R(\theta) = e^{G\theta} &= I + G\sin(\theta) + G^2(1-\cos(\theta)) \\ &= I - P + P\cos(\theta) + G\sin(\theta)\,. \end{aligned}$$

The formula for the exponential results from reducing the powers of G in the series expansion and identifying the respective series coefficients of G^2 and G with $-\cos(\theta)$ and $\sin(\theta)$ respectively. The second expression here for $e^{G\theta}$ is the same as the expression for $R(\theta)$ in the article containing the derivation of the generator, $R(\theta) = e^{G\theta}$.

In two dimensions, if $a = \left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $b = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, then $G = \left(\begin{smallmatrix}0 & -1\\1 & 0\end{smallmatrix}\right)$, $G^2 = \left(\begin{smallmatrix}-1 & 0\\0 & -1\end{smallmatrix}\right)$, and

$$R(\theta) = \begin{pmatrix}\cos(\theta) & -\sin(\theta)\\ \sin(\theta) & \cos(\theta)\end{pmatrix} = I\cos(\theta) + G\sin(\theta)$$

reduces to the standard matrix for a plane rotation.

The matrix $P = -G^2$ projects a vector onto the ab-plane and the rotation only affects this part of the vector. An example illustrating this is a rotation of 30° = π/6 in the plane spanned by a and b,

$$a = \begin{pmatrix}1\\0\\0\end{pmatrix} \qquad b = \frac{1}{\sqrt{5}}\begin{pmatrix}0\\1\\2\end{pmatrix}$$

$$I = \begin{pmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{pmatrix} \qquad G = \frac{1}{\sqrt{5}}\begin{pmatrix}0 & -1 & -2\\1 & 0 & 0\\2 & 0 & 0\end{pmatrix}$$

$$P = -G^2 = \frac{1}{5}\begin{pmatrix}5 & 0 & 0\\0 & 1 & 2\\0 & 2 & 4\end{pmatrix} \qquad P\begin{pmatrix}1\\2\\3\end{pmatrix} = \frac{1}{5}\begin{pmatrix}5\\8\\16\end{pmatrix} = a + \frac{8}{\sqrt{5}}b$$

$$\theta = \frac{\pi}{6} \quad \Rightarrow \quad R = \frac{1}{10}\begin{pmatrix}5\sqrt{3} & -\sqrt{5} & -2\sqrt{5}\\ \sqrt{5} & 8+\sqrt{3} & -4+2\sqrt{3}\\ 2\sqrt{5} & -4+2\sqrt{3} & 2+4\sqrt{3}\end{pmatrix}$$

Let $N = I - P$, so $N^2 = N$ and its products with P and G are zero. This will allow us to evaluate powers of R.

$$R\left(\frac{\pi}{6}\right) = N + P\frac{\sqrt{3}}{2} + G\frac{1}{2} \qquad R\left(\frac{\pi}{6}\right)^2 = N + P\frac{1}{2} + G\frac{\sqrt{3}}{2}$$

$$R\left(\frac{\pi}{6}\right)^3 = N + G \qquad R\left(\frac{\pi}{6}\right)^6 = N - P \qquad R\left(\frac{\pi}{6}\right)^{12} = N + P = I$$

Further information: Rodrigues' rotation formula and Axis–angle representation § Exponential map from so(3) to SO(3)

6.2.4 Nilpotent case

A matrix N is nilpotent if $N^q = 0$ for some integer q. In this case, the matrix exponential e^N can be computed directly from the series expansion, as the series terminates after a finite number of terms:

$$e^N = I + N + \frac{1}{2}N^2 + \frac{1}{6}N^3 + \cdots + \frac{1}{(q-1)!}N^{q-1}\,.$$

6.2.5 Generalization

When the minimal polynomial of a matrix X can be factored into a product of first degree polynomials, it can be expressed as a sum

$$X = A + N$$

where

- A is diagonalizable
- N is nilpotent
- A commutes with N (i.e. $AN = NA$)

This is the Jordan–Chevalley decomposition.

This means that we can compute the exponential of X by reducing to the previous two cases:

$$e^X = e^{A+N} = e^A e^N.$$

Note that we need the commutativity of A and N for the last step to work.

Another (closely related) method if the field is algebraically closed is to work with the Jordan form of X. Suppose that $X = PJP^{-1}$ where J is the Jordan form of X. Then

$$e^X = Pe^J P^{-1}.$$

Also, since

$$J = J_{a_1}(\lambda_1) \oplus J_{a_2}(\lambda_2) \oplus \cdots \oplus J_{a_n}(\lambda_n),$$

$$\begin{aligned} e^J &= \exp\left(J_{a_1}(\lambda_1) \oplus J_{a_2}(\lambda_2) \oplus \cdots \oplus J_{a_n}(\lambda_n)\right) \\ &= \exp\left(J_{a_1}(\lambda_1)\right) \oplus \exp\left(J_{a_2}(\lambda_2)\right) \oplus \cdots \oplus \exp\left(J_{a_k}(\lambda_k)\right). \end{aligned}$$

Therefore, we need only know how to compute the matrix exponential of a Jordan block. But each Jordan block is of the form

$$J_a(\lambda) = \lambda I + N$$

where N is a special nilpotent matrix. The matrix exponential of this block is given by

$$e^{\lambda I + N} = e^\lambda e^N.$$

6.2.6 Evaluation by Laurent series

By virtue of the Cayley–Hamilton theorem the matrix exponential is expressible as a polynomial of order n−1.

If P and Qt are nonzero polynomials in one variable, such that $P(A) = 0$, and if the meromorphic function

$$f(z) = \frac{e^{tz} - Q_t(z)}{P(z)}$$

is entire, then

$$e^{tA} = Q_t(A)$$

To prove this, multiply the first of the two above equalities by $P(z)$ and replace z by A.

Such a polynomial *Qt(z)* can be found as follows−−see Sylvester's formula. Letting a be a root of P, *Qa,t(z)* is solved from the product of P by the principal part of the Laurent series of f at a: It is proportional to the relevant Frobenius covariant. Then the sum *St* of the *Qa,t*, where a runs over all the roots of P, can be taken as a particular *Qt*. All the other *Qt* will be obtained by adding a multiple of P to *St(z)*. In particular, *St(z)*, the Lagrange-Sylvester polynomial, is the only *Qt* whose degree is less than that of P.

Example: Consider the case of an arbitrary 2-by-2 matrix,

$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix}.$$

The exponential matrix e^{tA}, by virtue of the Cayley–Hamilton theorem, must be of the form

$$e^{tA} = s_0(t)\, I + s_1(t)\, A$$

(For any complex number z and any $\boldsymbol{C}$-algebra B, we denote again by z the product of z by the unit of B.)

Let α and β be the roots of the characteristic polynomial of A,

$$P(z) = z^2 - (a+d)\, z + ad - bc = (z-\alpha)(z-\beta)\ .$$

Then we have

$$S_t(z) = e^{\alpha t}\frac{z-\beta}{\alpha-\beta} + e^{\beta t}\frac{z-\alpha}{\beta-\alpha}\ ,$$

and hence

$$s_0(t) = \frac{\alpha\, e^{\beta t} - \beta\, e^{\alpha t}}{\alpha-\beta}, \quad s_1(t) = \frac{e^{\alpha t} - e^{\beta t}}{\alpha-\beta}$$

if $\alpha \neq \beta$; while, if $\alpha = \beta$,

$$S_t(z) = e^{\alpha t}(1 + t(z-\alpha))\ ,$$

so that

$$s_0(t) = (1 - \alpha\, t)\, e^{\alpha t}, \quad s_1(t) = t\, e^{\alpha t}\ .$$

Defining

$$s \equiv \frac{\alpha+\beta}{2} = \frac{\operatorname{tr} A}{2}\ , \qquad q \equiv \frac{\alpha-\beta}{2} = \pm\sqrt{-\det\,(A - sI)},$$

we have

$$s_0(t) = e^{st}\left(\cosh(qt) - s\frac{\sinh(qt)}{q}\right), \qquad s_1(t) = e^{st}\frac{\sinh(qt)}{q},$$

where $\sin(qt)/q$ is 0 if t = 0, and t if q = 0. Thus,

Thus, as indicated above, the matrix A having decomposed into the sum of two mutually commuting pieces, the traceful piece and the traceless piece,

$$A = sI + (A - sI) \ ,$$

the matrix exponential reduces to a plain product of the exponentials of the two respective pieces. This is a formula often used in physics, as it amounts to the analog of Euler's formula for Pauli spin matrices, that is rotations of the doublet representation of the group SU(2).

The polynomial *St* can also be given the following "interpolation" characterization. Define $et(z) \equiv e^{tz}$, and n ≡ degP. Then $St(z)$ is the unique degree < *n* polynomial which satisfies $St^{(k)}(a) = et^{(k)}(a)$ whenever k is less than the multiplicity of a as a root of P. We assume, as we obviously can, that P is the minimal polynomial of A. We further assume that A is a diagonalizable matrix. In particular, the roots of P are simple, and the "interpolation" characterization indicates that *St* is given by the Lagrange interpolation formula, so it is the Lagrange–Sylvester polynomial .

At the other extreme, if $P = (z-a)^n$, then

$$S_t = e^{at} \sum_{k=0}^{n-1} \frac{t^k}{k!} (z-a)^k \ .$$

The simplest case not covered by the above observations is when $P = (z-a)^2 (z-b)$ with $a \neq b$, which yields

$$S_t = e^{at} \frac{z-b}{a-b} \left(1 + \left(t + \frac{1}{b-a} \right) (z-a) \right) + e^{bt} \frac{(z-a)^2}{(b-a)^2} \quad .$$

6.2.7 Evaluation by implementation of Sylvester's formula

A practical, expedited computation of the above reduces to the following rapid steps. Recall from above that an *n*×*n* matrix exp(*tA*) amounts to a linear combination of the first n−1 powers of A by the Cayley-Hamilton theorem. For diagonalizable matrices, as illustrated above, e.g. in the 2×2 case, Sylvester's formula yields exp(*tA*) = *Bα* exp(*tα*)+*Bβ* exp(*tβ*), where the Bs are the Frobenius covariants of A.

It is easiest, however, to simply solve for these Bs directly, by evaluating this expression and its first derivative at t=0, in terms of A and I, to find the same answer as above.

But this simple procedure also works for defective matrices, in a generalization due to Buchheim.[14] This is illustrated here for a 4×4 example of a matrix which is *not diagonalizable*, and the Bs are not projection matrices.

Consider

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & -1/8 \\ 0 & 0 & 1/2 & 1/2 \end{pmatrix} ,$$

with eigenvalues λ_1=3/4 and λ_2=1, each with a multiplicity of two.

Consider the exponential of each eigenvalue multiplied by t, exp(*λit*). Multiply each such by the corresponding undetermined coefficient matrix *Bi*. If the eigenvalues have an algebraic multiplicity greater than 1, then repeat the process, but now multiplying by an extra factor of t for each repetition, to ensure linear independence.

(If one eigenvalue had a multiplicity of three, then there would be the three terms: $B_{i_1} e^{\lambda_i t}$, $B_{i_2} t e^{\lambda_i t}$, $B_{i_3} t^2 e^{\lambda_i t}$. By contrast, when all eigenvalues are distinct, the Bs are just the Frobenius covariants, and solving for them as below just amounts to the inversion of the Vandermonde matrix of these 4 eigenvalues.)

Sum all such terms, here four such,

$$e^{At} = B_{1_1} e^{\lambda_1 t} + B_{1_2} t e^{\lambda_1 t} + B_{2_1} e^{\lambda_2 t} + B_{2_2} t e^{\lambda_2 t},$$

$$e^{At} = B_{1_1} e^{3/4t} + B_{1_2} t e^{3/4t} + B_{2_1} e^{1t} + B_{2_2} t e^{1t}$$

To solve for all of the unknown matrices B in terms of the first three powers of A and the identity, one needs four equations, the above one providing one such at t =0. Further, differentiate it with respect to t,

$$Ae^{At} = 3/4 B_{1_1} e^{3/4t} + (3/4t + 1)\, B_{1_2} e^{3/4t} + 1 B_{2_1} e^{1t} + (1t + 1)\, B_{2_2} e^{1t}\ ,$$

and again,

$$\begin{aligned} A^2 e^{At} =& (3/4)^2 B_{1_1} e^{3/4t} + \left((3/4)^2 t + (3/4 + 1 \cdot 3/4)\right) B_{1_2} e^{3/4t} + B_{2_1} e^{1t} \\ &+ \left(1^2 t + (1 + 1 \cdot 1)\right) B_{2_2} e^{1t} \\ =& (3/4)^2 B_{1_1} e^{3/4t} + \left((3/4)^2 t + 3/2\right) B_{1_2} e^{3/4t} + B_{2_1} e^{t} + (t + 2)\, B_{2_2} e^{t}\ , \end{aligned}$$

and once more,

$$\begin{aligned} A^3 e^{At} =& (3/4)^3 B_{1_1} e^{3/4t} + \left((3/4)^3 t + ((3/4)^2 + (3/2) \cdot 3/4))\right) B_{1_2} e^{3/4t} \\ &+ B_{2_1} e^{1t} + \left(1^3 t + (1 + 2) \cdot 1\right) B_{2_2} e^{1t} \\ =& (3/4)^3 B_{1_1} e^{3/4t} + \left((3/4)^3 t + 27/16)\right) B_{1_2} e^{3/4t} + B_{2_1} e^{t} + (t + 3 \cdot 1)\, B_{2_2} e^{t} \end{aligned}$$

(In the general case, n−1 derivatives need be taken.)

Setting t=0 in these four equations, the four coefficient matrices Bs may now be solved for,

$$\begin{aligned} I =& B_{1_1} + B_{2_1} \\ A =& 3/4 B_{1_1} + B_{1_2} + B_{2_1} + B_{2_2} \\ A^2 =& (3/4)^2 B_{1_1} + (3/2) B_{1_2} + B_{2_1} + 2 B_{2_2} \\ A^3 =& (3/4)^3 B_{1_1} + (27/16) B_{1_2} + B_{2_1} + 3 B_{2_2} \end{aligned}$$

to yield

$$\begin{aligned} B_{1_1} =& 128A^3 - 366A^2 + 288A - 80I \\ B_{1_2} =& 16A^3 - 44A^2 + 40A - 12I \\ B_{2_1} =& -128A^3 + 366A^2 - 288A + 80I \\ B_{2_2} =& 16A^3 - 40A^2 + 33A - 9I \end{aligned}$$

Substituting with the value for A yields the coefficient matrices

$$B_{1_1} = \begin{pmatrix} 0 & 0 & 48 & -16 \\ 0 & 0 & -8 & 2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$B_{1_2} = \begin{pmatrix} 0 & 0 & 4 & -2 \\ 0 & 0 & -1 & 1/2 \\ 0 & 0 & 1/4 & -1/8 \\ 0 & 0 & 1/2 & -1/4 \end{pmatrix}$$
$$B_{2_1} = \begin{pmatrix} 1 & 0 & -48 & 16 \\ 0 & 1 & 8 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$
$$B_{2_2} = \begin{pmatrix} 0 & 1 & 8 & -2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

so the final answer is

$$e^{tA} = \begin{pmatrix} e^t & te^t & (8t-48)\,e^t + (4t+48)\,e^{3t/4} & (16-2\,t)\,e^t + (-2t-16)\,e^{3t/4} \\ 0 & e^t & 8e^t + (-t-8)\,e^{3t/4} & -\frac{4e^t+(-t-4)e^{3t/4}}{2} \\ 0 & 0 & \frac{(t+4)e^{3t/4}}{4} & -\frac{te^{3t/4}}{8} \\ 0 & 0 & \frac{te^{3t/4}}{2} & -\frac{(t-4)e^{3t/4}}{4} \end{pmatrix}$$

The procedure is much shorter than Putzer's algorithm sometimes utilized in such cases.

See also: Derivative of the exponential map

6.3 Illustrations

Suppose that we want to compute the exponential of

$$B = \begin{bmatrix} 21 & 17 & 6 \\ -5 & -1 & -6 \\ 4 & 4 & 16 \end{bmatrix}.$$

Its Jordan form is

$$J = P^{-1}BP = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 16 & 1 \\ 0 & 0 & 16 \end{bmatrix},$$

where the matrix P is given by

$$P = \begin{bmatrix} -\frac{1}{4} & 2 & \frac{5}{4} \\ \frac{1}{4} & -2 & -\frac{1}{4} \\ 0 & 4 & 0 \end{bmatrix}.$$

Let us first calculate exp(J). We have

$$J = J_1(4) \oplus J_2(16)$$

The exponential of a 1×1 matrix is just the exponential of the one entry of the matrix, so $\exp(J_1(4)) = [e^4]$. The exponential of $J_2(16)$ can be calculated by the formula $e^{(\lambda I + N)} = e^{\lambda} e^{N}$ mentioned above; this yields[15]

$$\exp\left(\begin{bmatrix}16 & 1\\ 0 & 16\end{bmatrix}\right) = e^{16}\exp\left(\begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix}\right)$$
$$= e^{16}\left(\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} + \begin{bmatrix}0 & 1\\ 0 & 0\end{bmatrix} + \frac{1}{2!}\begin{bmatrix}0 & 0\\ 0 & 0\end{bmatrix} + \cdots\right) = \begin{bmatrix}e^{16} & e^{16}\\ 0 & e^{16}\end{bmatrix}.$$

Therefore, the exponential of the original matrix B is

$$\exp(B) = P\exp(J)P^{-1} = P\begin{bmatrix}e^4 & 0 & 0\\ 0 & e^{16} & e^{16}\\ 0 & 0 & e^{16}\end{bmatrix}P^{-1}$$
$$= \frac{1}{4}\begin{bmatrix}13e^{16}-e^4 & 13e^{16}-5e^4 & 2e^{16}-2e^4\\ -9e^{16}+e^4 & -9e^{16}+5e^4 & -2e^{16}+2e^4\\ 16e^{16} & 16e^{16} & 4e^{16}\end{bmatrix}.$$

6.4 Applications

6.4.1 Linear differential equations

The matrix exponential has applications to systems of linear differential equations. (See also matrix differential equation.) Recall from earlier in this article that a *homogeneous* differential equation of the form

$$\mathbf{y}' = A\mathbf{y}$$

has solution $e^{At}\,\mathbf{y}(0)$.

If we consider the vector

$$\mathbf{y}(t) = \begin{pmatrix}y_1(t)\\ \vdots\\ y_n(t)\end{pmatrix},$$

we can express a system of *inhomogeneous* coupled linear differential equations as

$$\mathbf{y}'(t) = A\mathbf{y}(t) + \mathbf{b}(t).$$

Making an ansatz to use an integrating factor of e^{-At} and multiplying throughout, yields

$$e^{-At}\mathbf{y}' - e^{-At}A\mathbf{y} = e^{-At}\mathbf{b}$$

$$e^{-At}\mathbf{y}' - Ae^{-At}\mathbf{y} = e^{-At}\mathbf{b}$$

$$\frac{d}{dt}(e^{-At}\mathbf{y}) = e^{-At}\mathbf{b}\,.$$

The second step is possible due to the fact that, if $AB = BA$, then $e^{At}B = Be^{At}$. So, calculating e^{At} leads to the solution to the system, by simply integrating the third step in ts.

Example (homogeneous)

Consider the system

$$\begin{array}{rcrrr} x' &=& 2x & -y & +z \\ y' &=& & 3y & -1z \\ z' &=& 2x & +y & +3z\,. \end{array}$$

The associated defective matrix is

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix}.$$

The matrix exponential is

$$e^{tA} = \frac{1}{2}\begin{bmatrix} e^{2t}(1+e^{2t}-2t) & -2te^{2t} & e^{2t}(-1+e^{2t}) \\ -e^{2t}(-1+e^{2t}-2t) & 2(t+1)e^{2t} & -e^{2t}(-1+e^{2t}) \\ e^{2t}(-1+e^{2t}+2t) & 2te^{2t} & e^{2t}(1+e^{2t}) \end{bmatrix},$$

so that the general solution of the homogeneous system is

$$\begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{x(0)}{2}\begin{bmatrix} e^{2t}(1+e^{2t}-2t) \\ -e^{2t}(-1+e^{2t}-2t) \\ e^{2t}(-1+e^{2t}+2t) \end{bmatrix} + \frac{y(0)}{2}\begin{bmatrix} -2te^{2t} \\ 2(t+1)e^{2t} \\ 2te^{2t} \end{bmatrix} + \frac{z(0)}{2}\begin{bmatrix} e^{2t}(-1+e^{2t}) \\ -e^{2t}(-1+e^{2t}) \\ e^{2t}(1+e^{2t}) \end{bmatrix},$$

amounting to

$$\begin{aligned} 2x &= x(0)(e^{2t}(1+e^{2t}-2t)) + y(0)(-2te^{2t}) + z(0)(e^{2t}(-1+e^{2t})) \\ 2y &= x(0)(-e^{2t}(-1+e^{2t}-2t)) + y(0)(2(t+1)e^{2t}) + z(0)(-e^{2t}(-1+e^{2t})) \\ 2z &= x(0)(e^{2t}(-1+e^{2t}+2t)) + y(0)(2te^{2t}) + z(0)(e^{2t}(1+e^{2t}))\,. \end{aligned}$$

Example (inhomogeneous)

Consider now the inhomogeneous system

$$\begin{array}{rcrcrcrl} x' &=& 2x & - & y & + & z & + & e^{2t} \\ y' &=& & & 3y & - & z & & \\ z' &=& 2x & + & y & + & 3z & + & e^{2t} \end{array}.$$

We again have

$$A = \begin{bmatrix} 2 & -1 & 1 \\ 0 & 3 & -1 \\ 2 & 1 & 3 \end{bmatrix},$$

and

$$\mathbf{b} = e^{2t}\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

From before, we already have the general solution to the homogeneous equation. Since the sum of the homogeneous and particular solutions give the general solution to the inhomogeneous problem, we now only need find the particular solution.

We have, by above,

$$\mathbf{y}_p = e^{tA}\int_0^t e^{(-u)A}\begin{bmatrix} e^{2u} \\ 0 \\ e^{2u} \end{bmatrix} du + e^{tA}\mathbf{c}$$

$$\mathbf{y}_p = e^{tA}\int_0^t \begin{bmatrix} 2e^u - 2ue^{2u} & -2ue^{2u} & 0 \\ -2e^u + 2(u+1)e^{2u} & 2(u+1)e^{2u} & 0 \\ 2ue^{2u} & 2ue^{2u} & 2e^u \end{bmatrix}\begin{bmatrix} e^{2u} \\ 0 \\ e^{2u} \end{bmatrix} du + e^{tA}\mathbf{c}$$

$$\mathbf{y}_p = e^{tA}\int_0^t \begin{bmatrix} e^{2u}(2e^u - 2ue^{2u}) \\ e^{2u}(-2e^u + 2(1+u)e^{2u}) \\ 2e^{3u} + 2ue^{4u} \end{bmatrix} du + e^{tA}\mathbf{c}$$

$$\mathbf{y}_p = e^{tA}\begin{bmatrix} -\frac{1}{24}e^{3t}(3e^t(4t-1)-16) \\ \frac{1}{24}e^{3t}(3e^t(4t+4)-16) \\ \frac{1}{24}e^{3t}(3e^t(4t-1)-16) \end{bmatrix} + \begin{bmatrix} 2e^t - 2te^{2t} & -2te^{2t} & 0 \\ -2e^t + 2(t+1)e^{2t} & 2(t+1)e^{2t} & 0 \\ 2te^{2t} & 2te^{2t} & 2e^t \end{bmatrix}\begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix},$$

which could be further simplified to get the requisite particular solution determined through variation of parameters. Note **c** = **y***p*(0). For more rigor, see the following generalization.

6.4.2 Inhomogeneous case generalization: variation of parameters

For the inhomogeneous case, we can use integrating factors (a method akin to variation of parameters). We seek a particular solution of the form $\mathbf{y}_p(t) = \exp(tA)\,\mathbf{z}\,(t)$,

$$\begin{aligned}\mathbf{y}_p'(t) &= (e^{tA})'\mathbf{z}(t) + e^{tA}\mathbf{z}'(t) \\ &= Ae^{tA}\mathbf{z}(t) + e^{tA}\mathbf{z}'(t) \\ &= A\mathbf{y}_p(t) + e^{tA}\mathbf{z}'(t)\ .\end{aligned}$$

For ***y**p* to be a solution,

$$\begin{aligned}e^{tA}\mathbf{z}'(t) &= \mathbf{b}(t) \\ \mathbf{z}'(t) &= (e^{tA})^{-1}\mathbf{b}(t) \\ \mathbf{z}(t) &= \int_0^t e^{-uA}\mathbf{b}(u)\,du + \mathbf{c}\ .\end{aligned}$$

Thus,

$$\begin{aligned}\mathbf{y}_p(t) &= e^{tA}\int_0^t e^{-uA}\mathbf{b}(u)\,du + e^{tA}\mathbf{c} \\ &= \int_0^t e^{(t-u)A}\mathbf{b}(u)\,du + e^{tA}\mathbf{c}\end{aligned},$$

where ***c*** is determined by the initial conditions of the problem.

More precisely, consider the equation

$$Y' - A\,Y = F(t)$$

with the initial condition $Y(t_0) = Y_0$, where A is an n by n complex matrix,

F is a continuous function from some open interval I to $\mathbb{C}^n$,

t_0 is a point of I, and

Y_0 is a vector of $\mathbb{C}^n$.

Left-multiplying the above displayed equality by e^{-tA} yields

$$Y(t) = e^{(t-t_0)A}\,Y_0 + \int_{t_0}^{t} e^{(t-x)A}\,F(x)\,dx\,.$$

We claim that the solution to the equation

$$P(d/dt)\,y = f(t)$$

with the initial conditions $y^{(k)}(t_0) = y_k$ for $0 \le k < n$ is

$$y(t) = \sum_{k=0}^{n-1} y_k\,s_k(t-t_0) + \int_{t_0}^{t} s_{n-1}(t-x)\,f(x)\,dx\,,$$

where the notation is as follows:

$P \in \mathbb{C}[X]$ is a monic polynomial of degree $n > 0$,

f is a continuous complex valued function defined on some open interval I,

t_0 is a point of I,

y_k is a complex number, and

sk(t) is the coefficient of X^k in the polynomial denoted by $S_t \in \mathbb{C}[X]$ in Subsection Evaluation by Laurent series above.

To justify this claim, we transform our order n scalar equation into an order one vector equation by the usual reduction to a first order system. Our vector equation takes the form

$$\frac{dY}{dt} - A\,Y = F(t), \quad Y(t_0) = Y_0,$$

where A is the transpose companion matrix of P. We solve this equation as explained above, computing the matrix exponentials by the observation made in Subsection Alternative above.

In the case n = 2 we get the following statement. The solution to

$$y'' - (\alpha+\beta)\,y' + \alpha\,\beta\,y = f(t), \quad y(t_0) = y_0, \quad y'(t_0) = y_1$$

is

$$y(t) = y_0\,s_0(t-t_0) + y_1\,s_1(t-t_0) + \int_{t_0}^{t} s_1(t-x)\,f(x)\,dx,$$

where the functions s_0 and s_1 are as in Subsection Evaluation by Laurent series above.

6.5 Matrix-matrix exponentials

The matrix exponential of another matrix (matrix-matrix exponential),[16] is defined as

$$X^Y = e^{\log(X)\cdot Y}$$

$${}^Y X = e^{Y\cdot\log(X)}$$

for X any normal and non-singular *n*×*n* matrix, and Y any complex *n*×*n* matrix.

For matrix-matrix exponentials, there is a distinction between the left exponential ${}^{\mathrm{Y}}\mathrm{X}$ and the right exponential X^{Y}, because the multiplication operator for matrix-to-matrix is not commutative. Moreover,

- If X is normal and non-singular, then X^{Y} and ${}^{\mathrm{Y}}\mathrm{X}$ have the same set of eigenvalues.
- If X is normal and non-singular, Y is normal, and $XY = YX$, then $X^Y = {}^Y X$.
- If X is normal and non-singular, and X, Y, Z commute with each other, then $X^{Y+Z} = X^Y \cdot X^Z$ and ${}^{Y+Z}X = {}^Y X \cdot {}^Z X$.

6.6 See also

- Matrix function
- Matrix logarithm
- Exponential function
- Exponential map (Lie theory)
- Magnus expansion
- Derivative of the exponential map
- Vector flow
- Golden–Thompson inequality
- Phase-type distribution
- Lie product formula
- Baker–Campbell–Hausdorff formula
- Frobenius covariant
- Sylvester's formula
- Trigonometric functions of matrices

6.7 References

[1] Hall 2015 Equation 2.1

[2] Hall 2015 Proposition 2.3

[3] Bhatia, R. (1997). *Matrix Analysis*. Graduate Texts in Mathematics. **169**. Springer. ISBN 978-0-387-94846-1.

[4] E. H. Lieb (1973). "Convex trace functions and the Wigner–Yanase–Dyson conjecture". *Adv. Math.* **11** (3): 267–288. doi:10.1016/0001-8708(73)90011-X. H. Epstein (1973). "Remarks on two theorems of E. Lieb". *Commun Math. Phys.* **31** (4): 317–325. doi:10.1007/BF01646492.

[5] Hall 2015 Exercises 2.9 and 2.10

[6] R. M. Wilcox (1967). "Exponential Operators and Parameter Differentiation in Quantum Physics". *Journal of Mathematical Physics*. **8** (4): 962–982. doi:10.1063/1.1705306.

[7] Hall 2015 Theorem 2.12

[8] "Matrix exponential - MATLAB expm - MathWorks Deutschland". Mathworks.de. 2011-04-30. Retrieved 2013-06-05.

[9] "GNU Octave - Functions of a Matrix". Network-theory.co.uk. 2007-01-11. Retrieved 2013-06-05.

[10] "scipy.linalg.expm function documentation". The SciPy Community. 2015-01-18. Retrieved 2015-05-29.

[11] in a Euclidean space

[12] Weyl, Hermann (1952). *Space Time Matter*. Dover. p. 142. ISBN 0-486-60267-2.

[13] Bjorken, James D.; Drell, Sidney D. (1964). *Relativistic Quantum Mechanics*. McGraw-Hill. p. 22.

[14] Rinehart, R. F. (1955). "The equivalence of definitions of a matric function". *The American Mathematical Monthly*, **62** (6), 395-414.

[15] This can be generalized; in general, the exponential of $Jn(a)$ is an upper triangular matrix with $e^a/0!$ on the main diagonal, $e^a/1!$ on the one above, $e^a/2!$ on the next one, and so on.

[16] Ignacio Barradas and Joel E. Cohen (1994). "Iterated Exponentiation, Matrix-Matrix Exponentiation, and Entropy" (PDF). Academic Press, Inc.

- Hall, Brian C. (2015), *Lie groups, Lie algebras, and representations: An elementary introduction*, Graduate Texts in Mathematics, **222** (2nd ed.), Springer
- Horn, Roger A.; Johnson, Charles R. (1991). *Topics in Matrix Analysis*. Cambridge University Press. ISBN 978-0-521-46713-1..
- Moler, Cleve; Van Loan, Charles F. (2003). "Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years Later" (PDF). *SIAM Review*. **45** (1): 3–49. doi:10.1137/S00361445024180. ISSN 1095-7200..
- Suzuki, Masuo (1985). "Decomposition formulas of exponential operators and Lie exponentials with some applications to quantum mechanics and statistical physics". *Journal of Mathematical Physics*. **26**: 601. doi:10.1063/1.526596.
- Curtright, T L; Fairlie, D B; Zachos, C K (2014). "A compact formula for rotations as spin matrix polynomials". *SIGMA*. **10**: 084. doi:10.3842/SIGMA.2014.084.
- Householder, Alston S. (2006). *The Theory of Matrices in Numerical Analysis*. Dover Books on Mathematics. ISBN 0486449726.
- Van Kortryk, T. S. (2016). "Matrix exponentials, SU(N) group elements, and real polynomial roots". *Journal of Mathematical Physics*. **57**: 021701. doi:10.1063/1.4938418.

6.8 External links

- Weisstein, Eric W. "Matrix Exponential". *MathWorld*.
- Module for the Matrix Exponential

Chapter 7

Prime number theorem

In number theory, the **prime number theorem** (**PNT**) describes the asymptotic distribution of the prime numbers among the positive integers. It formalizes the intuitive idea that primes become less common as they become larger by precisely quantifying the rate at which this occurs. The theorem was proved independently by Jacques Hadamard and Charles Jean de la Vallée-Poussin in 1896 using ideas introduced by Bernhard Riemann (in particular, the Riemann zeta function).

The first such distribution found is $\pi(N) \sim N / \log(N)$, where $\pi(N)$ is the prime-counting function and $\log(N)$ is the natural logarithm of N. This means that for large enough N, the probability that a random integer not greater than N is prime is very close to $1 / \log(N)$. Consequently, a random integer with at most $2n$ digits (for large enough n) is about half as likely to be prime as a random integer with at most n digits. For example, among the positive integers of at most 1000 digits, about one in 2300 is prime ($\log(10^{1000}) \approx 2302.6$), whereas among positive integers of at most 2000 digits, about one in 4600 is prime ($\log(10^{2000}) \approx 4605.2$). In other words, the average gap between consecutive prime numbers among the first N integers is roughly $\log(N)$.[1]

7.1 Statement

Let $\pi(x)$ be the prime-counting function that gives the number of primes less than or equal to x, for any real number x. For example, $\pi(10) = 4$ because there are four prime numbers (2, 3, 5 and 7) less than or equal to 10. The prime number theorem then states that $x / \log(x)$ is a good approximation to $\pi(x)$, in the sense that the limit of the *quotient* of the two functions $\pi(x)$ and $x / \log(x)$ as x increases without bound is 1:

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\log(x)} = 1,$$

known as **the asymptotic law of distribution of prime numbers**. Using asymptotic notation this result can be restated as

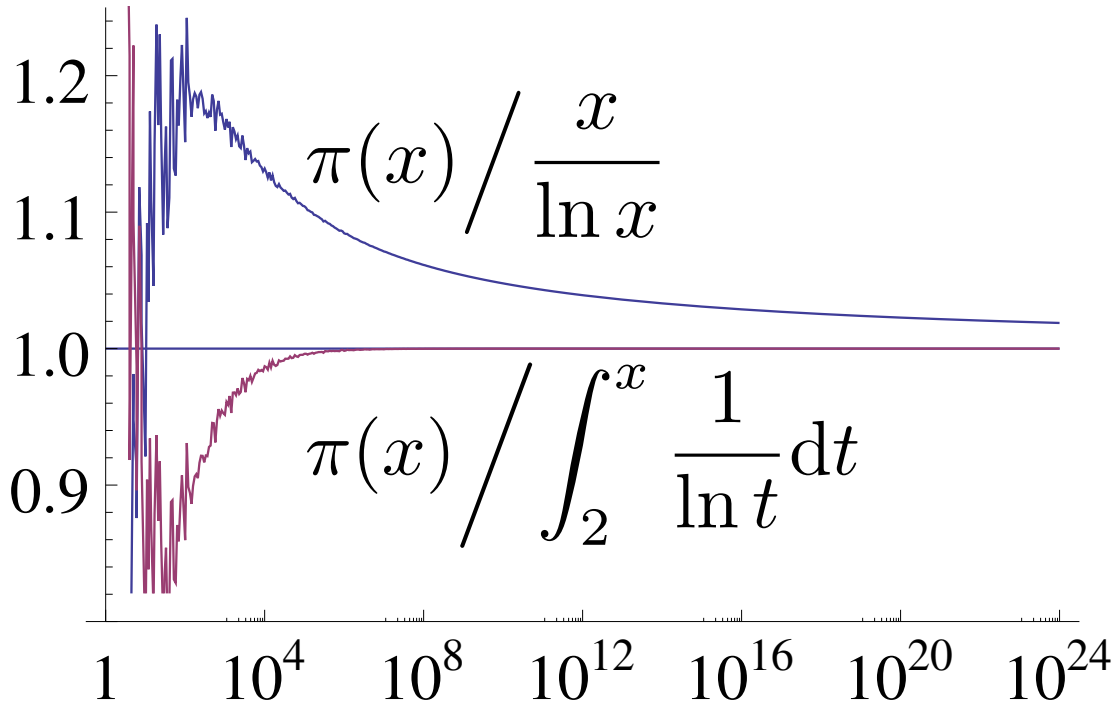
$$\pi(x) \sim \frac{x}{\log x}.$$

This notation (and the theorem) does *not* say anything about the limit of the *difference* of the two functions as x increases without bound. Instead, the theorem states that $x/\log(x)$ approximates $\pi(x)$ in the sense that the relative error of this approximation approaches 0 as x increases without bound.

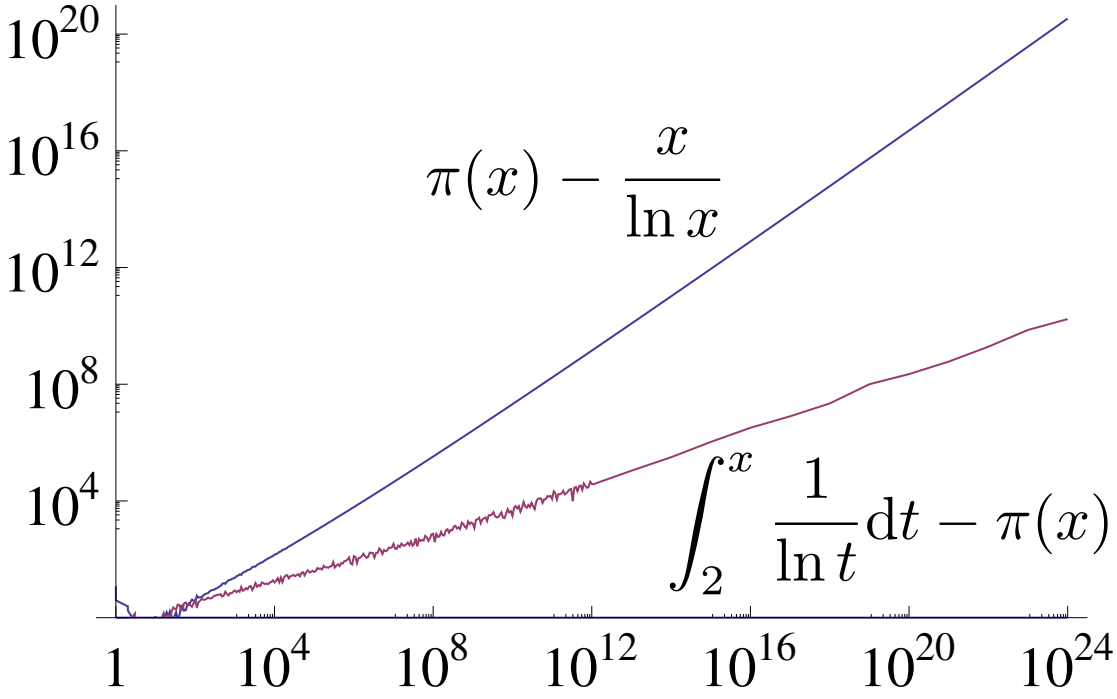
The prime number theorem is equivalent to the statement that the nth prime number p_n satisfies

$$p_n \sim n \log(n)$$

the asymptotic notation meaning, again, that the relative error of this approximation approaches 0 as n increases without bound. For example, the $200 \cdot 10^{15}$th prime number is 8512677386048191063,[2] and $(200 \cdot 10^{15})\log(200 \cdot 10^{15})$ rounds to 7967418752291744388, a relative error of about 6.4%.



Graph showing ratio of the prime-counting function $\pi(\mathrm{x})$ to two of its approximations, x*/log* x *and Li*(x)*. As* x *increases (note* x *axis is logarithmic), both ratios tend towards 1. The ratio for* x*/log* x *converges from above very slowly, while the ratio for Li*(x) *converges more quickly from below.*



Log-log plot showing absolute error of x*/log* x *and Li*(x)*, two approximations to the prime-counting function $\pi(\mathrm{x})$. Unlike the ratio, the difference between $\pi(\mathrm{x})$ and* x*/log* x *increases without bound as* x *increases. On the other hand, Li*(x) $- \pi(\mathrm{x})$ *switches sign infinitely many times.*

The prime number theorem is also equivalent to $\lim_{x\to\infty} \frac{\vartheta(x)}{x} = 1$, and $\lim_{x\to\infty} \frac{\psi(x)}{x} = 1$, where ϑ and ψ are the first and the second Chebyshev functions respectively.

7.2 History of the asymptotic law of distribution of prime numbers and its proof

Based on the tables by Anton Felkel and Jurij Vega, Adrien-Marie Legendre conjectured in 1797 or 1798 that $\pi(a)$ is approximated by the function $a/(\mathrm{A}\log(a) + B)$, where A and B are unspecified constants. In the second edition of his book on number theory (1808) he then made a more precise conjecture, with $A = 1$ and $B = -1.08366$. Carl Friedrich Gauss considered the same question at age 15 or 16 "ins Jahr 1792 oder 1793", according to his own recollection in 1849.[3] In 1838 Peter Gustav Lejeune Dirichlet came up with his own approximating function, the logarithmic integral $\mathrm{li}(x)$ (under the slightly different form of a series, which he communicated to Gauss). Both Legendre's and Dirichlet's formulas imply the same conjectured asymptotic equivalence of $\pi(x)$ and $x / \log(x)$ stated above, although it turned out that Dirichlet's approximation is considerably better if one considers the differences instead of quotients.

In two papers from 1848 and 1850, the Russian mathematician Pafnuty L'vovich Chebyshev attempted to prove the asymptotic law of distribution of prime numbers. His work is notable for the use of the zeta function $\zeta(s)$ (for real values of the argument "s", as are works of Leonhard Euler, as early as 1737) predating Riemann's celebrated memoir of 1859, and he succeeded in proving a slightly weaker form of the asymptotic law, namely, that if the limit of $\pi(x)/(x/\log(x))$ as x goes to infinity exists at all, then it is necessarily equal to one.[4] He was able to prove unconditionally that this ratio is bounded above and below by two explicitly given constants near 1, for all sufficiently large x.[5] Although Chebyshev's paper did not prove the Prime Number Theorem, his estimates for $\pi(x)$ were strong enough for him to prove Bertrand's postulate that there exists a prime number between n and $2n$ for any integer $n \geq 2$.

An important paper concerning the distribution of prime numbers was Riemann's 1859 memoir *On the Number of Primes Less Than a Given Magnitude*, the only paper he ever wrote on the subject. Riemann introduced new ideas into the subject, the chief of them being that the distribution of prime numbers is intimately connected with the zeros of the analytically extended Riemann zeta function of a complex variable. In particular, it is in this paper of Riemann that the idea to apply methods of complex analysis to the study of the real function $\pi(x)$ originates. Extending the ideas of Riemann, two proofs of the asymptotic law of the distribution of prime numbers were obtained independently by Jacques Hadamard and Charles Jean de la Vallée-Poussin and appeared in the same year (1896). Both proofs used methods from complex analysis, establishing as a main step of the proof that the Riemann zeta function $\zeta(s)$ is non-zero for all complex values of the variable s that have the form $s = 1 + it$ with $t > 0$.[6]

During the 20th century, the theorem of Hadamard and de la Vallée-Poussin also became known as the Prime Number Theorem. Several different proofs of it were found, including the "elementary" proofs of Atle Selberg and Paul Erdős (1949). While the original proofs of Hadamard and de la Vallée-Poussin are long and elaborate, later proofs introduced various simplifications through the use of Tauberian theorems but remained difficult to digest. A short proof was discovered in 1980 by American mathematician Donald J. Newman.[7][8] Newman's proof is arguably the simplest known proof of the theorem, although it is non-elementary in the sense that it uses Cauchy's integral theorem from complex analysis.

7.3 Proof methodology

In a lecture on prime numbers for a general audience, Fields medalist Terence Tao described one approach to proving the prime number theorem in poetic terms: listening to the "music" of the primes. We start with a "sound wave" that is "noisy" at the prime numbers and silent at other numbers; this is the von Mangoldt function. Then we analyze its notes or frequencies by subjecting it to a process akin to Fourier transform; this is the Mellin transform. The next and most difficult step is to prove that certain "notes" cannot occur in this music. This exclusion of certain notes leads to the statement of the prime number theorem. According to Tao, this proof yields much deeper insights into the distribution of the primes than the "elementary" proofs.[9]

7.4 Proof sketch

Here is a sketch of the proof referred to in Tao's lecture mentioned above. Like most proofs of the PNT, it starts out by reformulating the problem in terms of a less intuitive, but better-behaved, prime-counting function. The idea is to count the primes (or a related set such as the set of prime powers) with *weights* to arrive at a function with smoother asymptotic behavior. The most common such generalized counting function is the Chebyshev function $\psi(x)$, defined

by

$$\psi(x) = \sum_{\substack{p^k \leq x, \\ p \text{ prime is}}} \log p.$$

This is sometimes written as $\psi(x) = \sum_{n \leq x} \Lambda(n)$, where $\Lambda(n)$ is the von Mangoldt function, namely

$$\Lambda(n) = \begin{cases} \log p & \text{if} n = p^k \text{ prime some for } p \text{ integer and } k \geq 1, \\ 0 & \text{otherwise.} \end{cases}$$

It is now relatively easy to check that the PNT is equivalent to the claim that $\lim_{x \to \infty} \psi(x)/x = 1$. Indeed, this follows from the easy estimates

$$\psi(x) = \sum_{p \leq x} \log p \left\lfloor \frac{\log x}{\log p} \right\rfloor \leq \sum_{p \leq x} \log x = \pi(x) \log x$$

and (using big O notation) for any $\varepsilon > 0$,

$$\psi(x) \geq \sum_{x^{1-\varepsilon} \leq p \leq x} \log p \geq \sum_{x^{1-\varepsilon} \leq p \leq x} (1-\varepsilon) \log x = (1-\varepsilon)(\pi(x) + O(x^{1-\varepsilon})) \log x.$$

The next step is to find a useful representation for $\psi(x)$. Let $\zeta(s)$ be the Riemann zeta function. It can be shown that $\zeta(s)$ is related to the von Mangoldt function $\Lambda(n)$, and hence to $\psi(x)$, via the relation

$$-\frac{\zeta'(s)}{\zeta(s)} = \sum_{n=1}^{\infty} \Lambda(n) n^{-s}.$$

A delicate analysis of this equation and related properties of the zeta function, using the Mellin transform and Perron's formula, shows that for non-integer *x* the equation

$$\psi(x) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \log(2\pi)$$

holds, where the sum is over all zeros (trivial and non-trivial) of the zeta function. This striking formula is one of the so-called explicit formulas of number theory, and is already suggestive of the result we wish to prove, since the term *x* (claimed to be the correct asymptotic order of $\psi(x)$) appears on the right-hand side, followed by (presumably) lower-order asymptotic terms.

The next step in the proof involves a study of the zeros of the zeta function. The trivial zeros –2, –4, –6, –8, ... can be handled separately:

$$\sum_{n=1}^{\infty} \frac{1}{2n\, x^{2n}} = -\frac{1}{2} \log \left(1 - \frac{1}{x^2} \right),$$

which vanishes for a large *x*. The nontrivial zeros, namely those on the critical strip $0 \leq \Re(s) \leq 1$, can potentially be of an asymptotic order comparable to the main term *x* if $\Re(\rho) = 1$, so we need to show that all zeros have real part strictly less than 1.

To do this, we take for granted that $\zeta(s)$ is meromorphic in the half-plane $\Re(s) > 0$, and is analytic there except for a simple pole at $s = 1$, and that there is a product formula $\zeta(s) = \prod_p (1 - p^{-s})^{-1}$ for $\Re(s) > 1$. This product formula follows from the existence of unique prime factorization of integers, and shows that $\zeta(s)$ is never zero in this

region, so that its logarithm is defined there and $\log \zeta(s) = -\sum_p \log(1-p^{-s}) = \sum_{p,n} p^{-ns}/n$. Write $s = x + iy$; then

$$|\zeta(x+iy)| = \exp(\sum_{n,p} \frac{\cos ny \log p}{np^{nx}}).$$

Now observe the identity $3 + 4\cos\phi + \cos 2\phi = 2(1+\cos\phi)^2 \geq 0$, so that

$$|\zeta(x)^3 \zeta(x+iy)^4 \zeta(x+2iy)| = \exp \sum_{n,p} \frac{3 + 4\cos(ny\log p) + \cos(2ny \log p)}{np^{nx}} \geq 1$$

for all $x > 1$. Suppose now that $\zeta(1+iy) = 0$. Certainly y is not zero, since $\zeta(s)$ has a simple pole at $s = 1$. Suppose that $x > 1$ and let x tend to 1 from above. Since $\zeta(s)$ has a simple pole at $s = 1$ and $\zeta(x+2iy)$ stays analytic, the left hand side in the previous inequality tends to 0 , a contradiction.

Finally, we can conclude that the PNT is "morally" true. To rigorously complete the proof there are still serious technicalities to overcome, due to the fact that the summation over zeta zeros in the explicit formula for $\psi(x)$ does not converge absolutely but only conditionally and in a "principal value" sense. There are several ways around this problem but many of them require rather delicate complex-analytic estimates that are beyond the scope of this article. Edwards's book[10] provides the details. Another method is to use Ikehara's Tauberian theorem, though this theorem is itself quite hard to prove. D. J. Newman observed that the full strength of Ikehara's theorem is not needed for the prime number theorem, and one can get away with a special case that is much easier to prove.

7.5 Prime-counting function in terms of the logarithmic integral

In a handwritten note on a reprint of his 1838 paper "Sur l'usage des séries infinies dans la théorie des nombres", which he mailed to Carl Friedrich Gauss, Peter Gustav Lejeune Dirichlet conjectured (under a slightly different form appealing to a series rather than an integral) that an even better approximation to π(*x*) is given by the offset logarithmic integral function Li(*x*), defined by

$$\mathrm{Li}(x) = \int_2^x \frac{1}{\log t}\,\mathrm{d}t = \mathrm{li}(x) - \mathrm{li}(2).$$

Indeed, this integral is strongly suggestive of the notion that the 'density' of primes around *t* should be 1/log *t*. This function is related to the logarithm by the asymptotic expansion

$$\mathrm{Li}(x) \sim \frac{x}{\log x} \sum_{k=0}^{\infty} \frac{k!}{(\log x)^k} = \frac{x}{\log x} + \frac{x}{(\log x)^2} + \frac{2x}{(\log x)^3} + \cdots .$$

So, the prime number theorem can also be written as π(*x*) ~ Li(*x*). In fact, in another paper in 1899 La Vallée Poussin proved that

$$\pi(x) = \mathrm{Li}(x) + O\left(x\mathrm{e}^{-a\sqrt{\log x}}\right) \qquad \text{as} x \to \infty$$

for some positive constant *a*, where *O*(...) is the big O notation. This has been improved to

$$\pi(x) = \mathrm{Li}(x) + O\left(x\,\exp\left(-\frac{A(\log x)^{3/5}}{(\log\log x)^{1/5}}\right)\right).$$

Because of the connection between the Riemann zeta function and π(*x*), the Riemann hypothesis has considerable importance in number theory: if established, it would yield a far better estimate of the error involved in the prime

number theorem than is available today. More specifically, Helge von Koch showed in 1901[11] that, if and only if the Riemann hypothesis is true, the error term in the above relation can be improved to

$$\pi(x) = \mathrm{Li}(x) + O\left(\sqrt{x}\log x\right).$$

The constant involved in the big O notation was estimated in 1976 by Lowell Schoenfeld:[12] assuming the Riemann hypothesis,

$$|\pi(x) - \mathrm{li}(x)| < \frac{\sqrt{x}\,\log x}{8\pi}$$

for all $x \geq 2657$. He also derived a similar bound for the Chebyshev prime-counting function ψ:

$$|\psi(x) - x| < \frac{\sqrt{x}\,\log^2 x}{8\pi}$$

for all $x \geq 73.2$. This latter bound has been shown to express a variance to mean power law (when regarded as a random function over the integers), $1/f$ noise and to also correspond to the Tweedie compound Poisson distribution.[13] Parenthetically, the Tweedie distributions represent a family of scale invariant distributions that serve as foci of convergence for a generalization of the central limit theorem.[14]

The logarithmic integral $\mathrm{Li}(x)$ is larger than $\pi(x)$ for "small" values of x. This is because it is (in some sense) counting not primes, but prime powers, where a power p^n of a prime p is counted as $1/n$ of a prime. This suggests that $\mathrm{Li}(x)$ should usually be larger than $\pi(x)$ by roughly $\mathrm{Li}(x^{1/2})/2$, and in particular should usually be larger than $\pi(x)$. However, in 1914, J. E. Littlewood proved that this is not always the case. The first value of x where $\pi(x)$ exceeds $\mathrm{Li}(x)$ is probably around $x = 10^{316}$; see the article on Skewes' number for more details.

7.6 Elementary proofs

In the first half of the twentieth century, some mathematicians (notably G. H. Hardy) believed that there exists a hierarchy of proof methods in mathematics depending on what sorts of numbers (integers, reals, complex) a proof requires, and that the prime number theorem (PNT) is a "deep" theorem by virtue of requiring complex analysis.[15] This belief was somewhat shaken by a proof of the PNT based on Wiener's tauberian theorem, though this could be set aside if Wiener's theorem were deemed to have a "depth" equivalent to that of complex variable methods. There is no rigorous and widely accepted definition of the notion of elementary proof in number theory. One definition is "a proof that can be carried out in first order Peano arithmetic." There are number-theoretic statements (for example, the Paris–Harrington theorem) provable using second order but not first order methods, but such theorems are rare to date.

In March 1948, Atle Selberg established, by elementary means, the asymptotic formula

$$\vartheta(x)\log(x) + \sum_{p\leq x}\log(p)\,\vartheta\left(\frac{x}{p}\right) = 2x\log(x) + O(x)$$

where

$$\vartheta(x) = \sum_{p\leq x}\log(p)$$

for primes p .[16] By July of that year, Selberg and Paul Erdős had each obtained elementary proofs of the PNT, both using Selberg's asymptotic formula as a starting point.[15][17] These proofs effectively laid to rest the notion that the PNT was "deep," and showed that technically "elementary" methods (in other words Peano arithmetic) were more powerful than had been believed to be the case. In 1994, Charalambos Cornaros and Costas Dimitracopoulos proved the PNT using only $I\Delta_0 + \exp$,[18] a formal system far weaker than Peano arithmetic. On the history of the elementary proofs of the PNT, including the Erdős–Selberg priority dispute, see an article by Dorian Goldfeld.[15]

7.7 Computer verifications

In 2005, Avigad *et al.* employed the Isabelle theorem prover to devise a computer-verified variant of the Erdős–Selberg proof of the PNT.[19] This was the first machine-verified proof of the PNT. Avigad chose to formalize the Erdős–Selberg proof rather than an analytic one because while Isabelle's library at the time could implement the notions of limit, derivative, and transcendental function, it had almost no theory of integration to speak of (Avigad et al. p. 19).

In 2009, John Harrison employed HOL Light to formalize a proof employing complex analysis.[20] By developing the necessary analytic machinery, including the Cauchy integral formula, Harrison was able to formalize "a direct, modern and elegant proof instead of the more involved 'elementary' Erdős–Selberg argument".

7.8 Prime number theorem for arithmetic progressions

Let $\pi_{n,a}(x)$ denote the number of primes in the arithmetic progression a, $a + n$, $a + 2n$, $a + 3n$, ... less than x. Lejeune Dirichlet and Legendre conjectured, and Vallée-Poussin proved, that, if a and n are coprime, then

$$\pi_{n,a}(x) \sim \frac{1}{\varphi(n)} \operatorname{Li}(x),$$

where φ is the Euler's totient function. In other words, the primes are distributed evenly among the residue classes $[a]$ modulo n with $\gcd(a, n) = 1$. This is stronger than Dirichlet's theorem on arithmetic progressions (which only states that there is an infinity of primes in each class) and can be proved using similar methods used by Newman for his proof of the prime number theorem.[21]

The Siegel–Walfisz theorem gives a good estimate for the distribution of primes in residue classes.

7.8.1 Prime number race

Although we have in particular

$$\pi_{4,1}(x) \sim \pi_{4,3}(x),$$

empirically the primes congruent to 3 are more numerous and are nearly always ahead in this "prime number race"; the first reversal occurs at $x = 26{,}861$.[22]:1–2 However Littlewood showed in 1914[22]:2 that there are infinitely many sign changes for the function

$$\pi_{4,1}(x) - \pi_{4,3}(x),$$

so the lead in the race switches back and forth infinitely many times. The phenomenon that $\pi_{4,3}(x)$ is ahead most of the time is called Chebyshev's bias. The prime number race generalizes to other moduli and is the subject of much research; Pál Turán asked whether it is always the case that $\pi(x;a,c)$ and $\pi(x;b,c)$ change places when a and b are coprime to c.[23] Granville and Martin give a thorough exposition and survey.[22]

7.9 Bounds on the prime-counting function

The prime number theorem is an *asymptotic* result. It gives an ineffective bound on $\pi(x)$ as a direct consequence of the definition of the limit: for all $\varepsilon > 0$, there is an S such that for all $x > S$,

$$(1-\varepsilon)\frac{x}{\log x} < \pi(x) < (1+\varepsilon)\frac{x}{\log x}.$$

However, better bounds on $\pi(x)$ are known, for instance Pierre Dusart's

$$\frac{x}{\log x}\left(1+\frac{1}{\log x}\right) < \pi(x) < \frac{x}{\log x}\left(1+\frac{1}{\log x}+\frac{2.51}{(\log x)^2}\right).$$

The first inequality holds for all $x \geq 599$ and the second one for $x \geq 355991$.[24]

A weaker but sometimes useful bound for $x \geq 55$ is

$$\tfrac{x}{\log x+2} < \pi(x) < \tfrac{x}{\log x-4}.$$ [25]

In Pierre Dusart's thesis there are stronger versions of this type of inequality that are valid for larger x. Later in 2010, Dusart proved:

$$\tfrac{x}{\log x-1} < \pi(x) \text{ for } x \geq 5393 \text{ , and}$$

$$\pi(x) < \tfrac{x}{\log x-1.1} \text{ for } x \geq 60184.$$ [26]

The proof by de la Vallée-Poussin implies the following. For every ε > 0, there is an S such that for all $x > S$,

$$\frac{x}{\log x-(1-\varepsilon)} < \pi(x) < \frac{x}{\log x-(1+\varepsilon)}.$$

7.10 Approximations for the *n*th prime number

As a consequence of the prime number theorem, one gets an asymptotic expression for the *n*th prime number, denoted by *pn*:

$$p_n \sim n\log n.$$

A better approximation is

$$\tfrac{p_n}{n} = \log n + \log\log n - 1 + \tfrac{\log\log n-2}{\log n} - \tfrac{(\log\log n)^2-6\log\log n+11}{2(\log n)^2} + o\left(\tfrac{1}{(\log n)^2}\right).$$ [27]

Again considering the $200 \cdot 10^{15}$ prime number 8512677386048191063, this gives an estimate of 8512681315554715386; the first 5 digits match and relative error is about 0.00005%.

Rosser's theorem states that *pn* is larger than *n* log *n*. This can be improved by the following pair of bounds:[28][29]

$$\log n + \log\log n - 1 < \frac{p_n}{n} < \log n + \log\log n \quad \text{for} n \geq 6.$$

7.11 Table of π(*x*), *x* / log *x*, and li(*x*)

The table compares exact values of π(*x*) to the two approximations *x* / log *x* and li(*x*). The last column, *x* / π(*x*), is the average prime gap below *x*.

The value for $\pi(10^{24})$ was originally computed assuming the Riemann hypothesis;[30] it has since been verified unconditionally.[31]

7.12 Analogue for irreducible polynomials over a finite field

There is an analogue of the prime number theorem that describes the "distribution" of irreducible polynomials over a finite field; the form it takes is strikingly similar to the case of the classical prime number theorem.

To state it precisely, let $F = \mathrm{GF}(q)$ be the finite field with q elements, for some fixed q, and let Nn be the number of monic *irreducible* polynomials over F whose degree is equal to n. That is, we are looking at polynomials with coefficients chosen from F, which cannot be written as products of polynomials of smaller degree. In this setting, these polynomials play the role of the prime numbers, since all other monic polynomials are built up of products of them. One can then prove that

$$N_n \sim \frac{q^n}{n}.$$

If we make the substitution $x = q^n$, then the right hand side is just

$$\frac{x}{\log_q x},$$

which makes the analogy clearer. Since there are precisely q^n monic polynomials of degree n (including the reducible ones), this can be rephrased as follows: if a monic polynomial of degree n is selected randomly, then the probability of it being irreducible is about $1/n$.

One can even prove an analogue of the Riemann hypothesis, namely that

$$N_n = \frac{q^n}{n} + O\left(\frac{q^{n/2}}{n}\right).$$

The proofs of these statements are far simpler than in the classical case. It involves a short combinatorial argument,[32] summarised as follows. Every element of the degree n extension of F is a root of some irreducible polynomial whose degree d divides n; by counting these roots in two different ways one establishes that

$$q^n = \sum_{d|n} dN_d,$$

where the sum is over all divisors d of n. Möbius inversion then yields

$$N_n = \frac{1}{n}\sum_{d|n} \mu(n/d)q^d,$$

where $\mu(k)$ is the Möbius function. (This formula was known to Gauss.) The main term occurs for $d = n$, and it is not difficult to bound the remaining terms. The "Riemann hypothesis" statement depends on the fact that the largest proper divisor of n can be no larger than $n/2$.

7.13 See also

- Abstract analytic number theory for information about generalizations of the theorem.
- Landau prime ideal theorem for a generalization to prime ideals in algebraic number fields.
- Riemann hypothesis

7.14 Notes

[1] Hoffman, Paul (1998). *The Man Who Loved Only Numbers*. New York: Hyperion Books. p. 227. ISBN 0-7868-8406-1. MR 1666054.

[2] "Prime Curios!: 8512677386048191063". *Prime Curios!*. University of Tennessee at Martin. 2011-10-09.

[3] C. F. Gauss. *Werke*, Bd 2, 1st ed, 444–447. Göttingen 1863.

[4] N. Costa Pereira (August–September 1985). "A Short Proof of Chebyshev's Theorem". *American Mathematical Monthly*. **92** (7): 494–495. doi:10.2307/2322510. JSTOR 2322510.

[5] M. Nair (February 1982). "On Chebyshev-Type Inequalities for Primes". *American Mathematical Monthly*. **89** (2): 126–129. doi:10.2307/2320934. JSTOR 2320934.

[6] Ingham, A. E. (1990). *The Distribution of Prime Numbers*. Cambridge University Press. pp. 2–5. ISBN 0-521-39789-8.

[7] Newman, Donald J. (1980). "Simple analytic proof of the prime number theorem". *American Mathematical Monthly*. **87** (9): 693–696. doi:10.2307/2321853. JSTOR 2321853. MR 0602825.

[8] Zagier, Don (1997). "Newman's short proof of the prime number theorem". *American Mathematical Monthly*. **104** (8): 705–708. doi:10.2307/2975232. JSTOR 2975232. MR 1476753.

[9] Video and slides of Tao's lecture on primes, UCLA January 2007.

[10] Edwards, Harold M. (2001). *Riemann's zeta function*. Courier Dover Publications. ISBN 0-486-41740-9.

[11] Von Koch, Helge (1901). "Sur la distribution des nombres premiers" [On the distribution of prime numbers]. *Acta Mathematica* (in French). **24** (1): 159–182. doi:10.1007/BF02403071. MR 1554926.

[12] Schoenfeld, Lowell (1976). "Sharper Bounds for the Chebyshev Functions $\theta(x)$ and $\psi(x)$. II". *Mathematics of Computation*. **30** (134): 337–360. doi:10.2307/2005976. JSTOR 2005976. MR 0457374..

[13] Kendal, WS (2013). "Fluctuation scaling and 1/*f* noise: shared origins from the Tweedie family of statistical distributions". *J Basic Appl Phys*. **2**: 40–49.

[14] Jørgensen, Bent; Martínez, José Raúl; Tsao, Min (1994). "Asymptotic behaviour of the variance function". *Scandinavian Journal of Statistics*. **21**: 223–243. JSTOR 4616314. MR 1292637.

[15] Goldfeld, Dorian (2004). "The elementary proof of the prime number theorem: an historical perspective" (PDF). In Chudnovsky, David; Chudnovsky, Gregory; Nathanson, Melvyn. *Number theory (New York, 2003)*. New York: Springer-Verlag. pp. 179–192. doi:10.1007/978-1-4419-9060-0_10. ISBN 0-387-40655-7. MR 2044518.

[16] Selberg, Atle (1949). "An Elementary Proof of the Prime-Number Theorem". *Annals of Mathematics*. **50** (2): 305–313. doi:10.2307/1969455. MR 0029410.

[17] Baas, Nils A.; Skau, Christian F. (2008). "The lord of the numbers, Atle Selberg. On his life and mathematics" (PDF). *Bull. Amer. Math. Soc*. **45** (4): 617–649. doi:10.1090/S0273-0979-08-01223-8. MR 2434348.

[18] Cornaros, Charalambos; Dimitracopoulos, Costas (1994). "The prime number theorem and fragments of *PA*" (PDF). *Archive for Mathematical Logic*. **33** (4): 265–281. doi:10.1007/BF01270626. MR 1294272.

[19] Avigad, Jeremy; Donnelly, Kevin; Gray, David; Raff, Paul (2008). "A formally verified proof of the prime number theorem". *ACM Transactions on Computational Logic*. **9** (1). arXiv:cs/0509025. doi:10.1145/1297658.1297660. MR 2371488.

[20] Harrison, John (2009). "Formalizing an analytic proof of the Prime Number Theorem". *Journal of Automated Reasoning*. **43** (3): 243–261. doi:10.1007/s10817-009-9145-6. MR 2544285.

[21] Soprounov, Ivan (1998). "A short proof of the Prime Number Theorem for arithmetic progressions" (PDF).

[22] Granville, Andrew; Martin, Greg (2006). "Prime Number Races" (PDF). *American Mathematical Monthly*. **113** (1): 1–33. doi:10.2307/27641834. JSTOR 27641834. MR 2202918.

[23] Guy, Richard K. (2004). *Unsolved problems in number theory* (3rd ed.). Springer-Verlag. A4. ISBN 978-0-387-20860-2. Zbl 1058.11001.

[24] Dusart, Pierre (1998). *Autour de la fonction qui compte le nombre de nombres premiers* (PhD thesis) (in French).

[25] Rosser, Barkley (1941). "Explicit Bounds for Some Functions of Prime Numbers". *American Journal of Mathematics*. **63** (1): 211–232. doi:10.2307/2371291. JSTOR 2371291. MR 0003018.

[26] Dusart, Pierre. "Estimates of Some Functions Over Primes without R.H." (PDF). arxiv.org. Retrieved 22 April 2014.

[27] Cesàro, Ernesto (1894). "Sur une formule empirique de M. Pervouchine". *Comptes rendus hebdomadaires des séances de l'Académie des sciences* (in French). **119**: 848–849.

[28] Bach, Eric; Shallit, Jeffrey (1996). *Algorithmic Number Theory*. Foundations of Computing Series. **1**. Cambridge: MIT Press. p. 233. ISBN 0-262-02405-5. MR 1406794.

[29] Dusart, Pierre (1999). "The kth prime is greater than $k(\log k + \log\log k - 1)$ for $k \geq 2$". *Mathematics of Computation*. **68** (225): 411–415. doi:10.1090/S0025-5718-99-01037-6. MR 1620223.

[30] "Conditional Calculation of pi(10^{24})". Chris K. Caldwell. Retrieved 2010-08-03.

[31] Platt, David (2015). "Computing $\pi(x)$ analytically". *Mathematics of Computation*. **84** (293): 1521–1535. arXiv:1203.5712. doi:10.1090/S0025-5718-2014-02884-6. MR 3315519.

[32] Chebolu, Sunil; Ján Mináč (December 2011). "Counting Irreducible Polynomials over Finite Fields Using the InclusionπExclusion Principle". *Mathematics Magazine*. **84** (5): 369–371. doi:10.4169/math.mag.84.5.369.

7.15 References

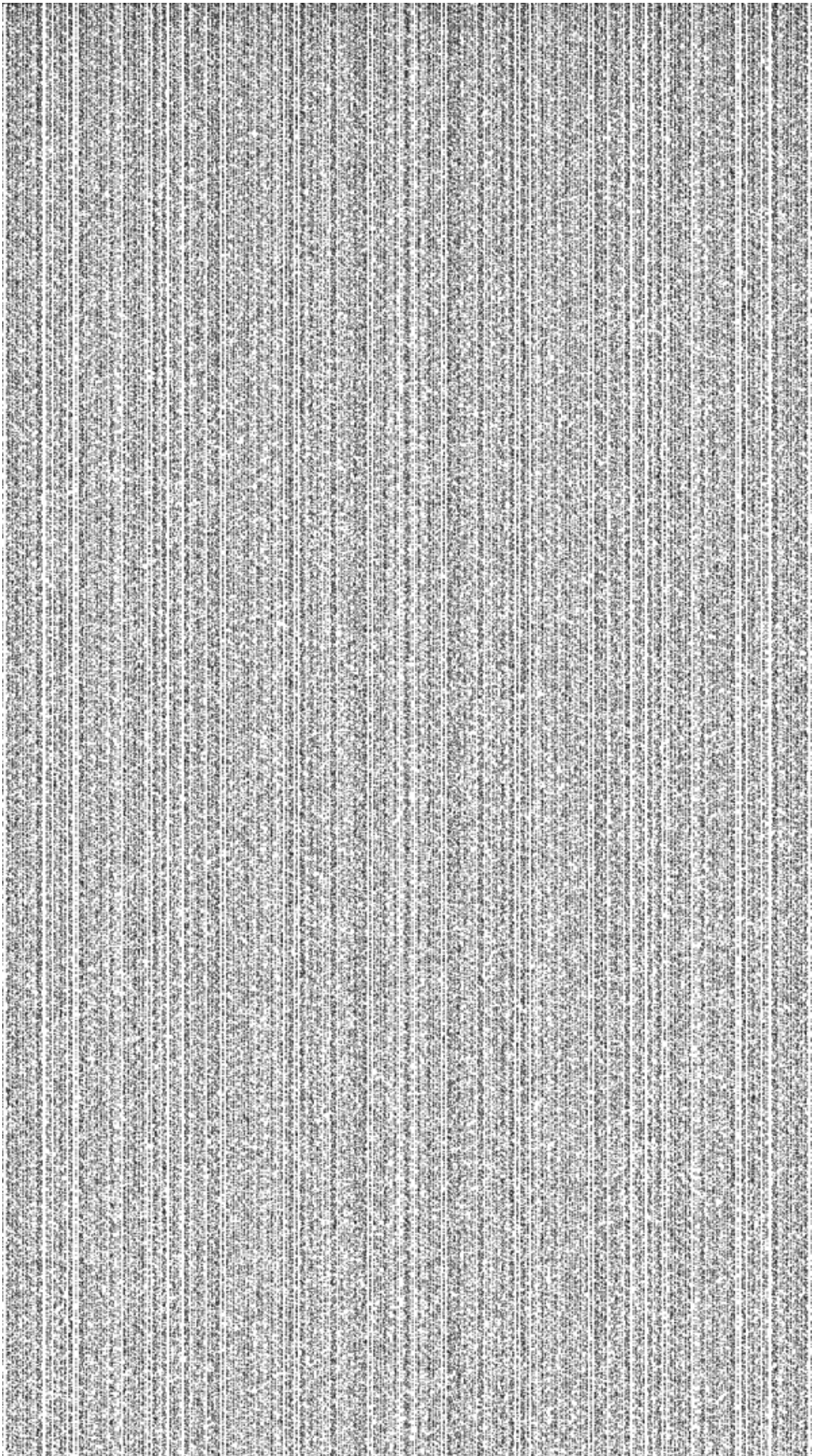
- Hardy, G. H. & Littlewood, J. E. (1916). "Contributions to the Theory of the Riemann Zeta-Function and the Theory of the Distribution of Primes". *Acta Mathematica*. **41**: 119–196. doi:10.1007/BF02422942.
- Granville, Andrew (1995). "Harald Cramér and the distribution of prime numbers" (PDF). *Scandinavian Actuarial Journal*. **1**: 12–28. doi:10.1080/03461238.1995.10413946.

7.16 External links

- Hazewinkel, Michiel, ed. (2001), "Distribution of prime numbers", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Table of Primes by Anton Felkel.
- Short video visualizing the Prime Number Theorem.
- Prime formulas and Prime number theorem at MathWorld.
-

"Prime number theorem". *PlanetMath*.

- How Many Primes Are There? and The Gaps between Primes by Chris Caldwell, University of Tennessee at Martin.
- Tables of prime-counting functions by Tomás Oliveira e Silva



Chapter 8

Cycle detection

This article is about iterated functions. For another use, see Cycle detection (graph theory).

In computer science, **cycle detection** or **cycle finding** is the algorithmic problem of finding a cycle in a sequence of iterated function values.

For any function f that maps a finite set S to itself, and any initial value x_0 in S, the sequence of iterated function values

$$x_0,\ x_1 = f(x_0),\ x_2 = f(x_1),\ \ldots,\ x_i = f(x_{i-1}),\ \ldots$$

must eventually use the same value twice: there must be some pair of distinct indices i and j such that $xi = xj$. Once this happens, the sequence must continue periodically, by repeating the same sequence of values from xi to xj_{-1}. Cycle detection is the problem of finding i and j, given f and x_0.

Several algorithms for finding cycles quickly and with little memory are known. Floyd's tortoise and the hare algorithm moves two pointers at different speeds through the sequence of values until they both point to equal values. Alternatively, Brent's algorithm is based on the idea of exponential search. Both Floyd's and Brent's algorithms use only a constant number of memory cells, and take a number of function evaluations that is proportional to the distance from the start of the sequence to the first repetition. Several other algorithms trade off larger amounts of memory for fewer function evaluations.

The applications of cycle detection include testing the quality of pseudorandom number generators and cryptographic hash functions, computational number theory algorithms, detection of infinite loops in computer programs and periodic configurations in cellular automata, and the automated shape analysis of linked list data structures.

8.1 Example

The figure shows a function f that maps the set $S = \{0,1,2,3,4,5,6,7,8\}$ to itself. If one starts from $x_0 = 2$ and repeatedly applies f, one sees the sequence of values

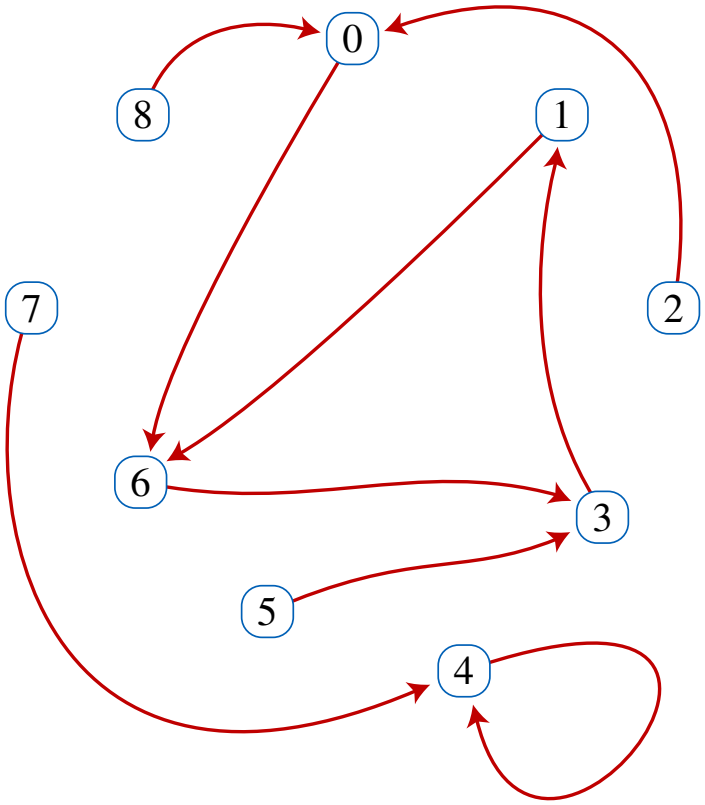
2, 0, 6, 3, 1, 6, 3, 1, 6, 3, 1, ...

The cycle in this value sequence is 6, 3, 1.

8.2 Definitions

Let S be any finite set, f be any function from S to itself, and x_0 be any element of S. For any $i > 0$, let $xi = f(xi_{-1})$. Let μ be the smallest index such that the value $x\mu$ reappears infinitely often within the sequence of values xi, and let λ (the loop length) be the smallest positive integer such that $x\mu = x\lambda_{+}\mu$. The cycle detection problem is the task of finding λ and μ.[1]

x	$f(x)$
0	6
1	6
2	0
3	1
4	4
5	3
6	3
7	4
8	0



A function from and to the set {0,1,2,3,4,5,6,7,8} and the corresponding functional graph

One can view the same problem graph-theoretically, by constructing a functional graph (that is, a directed graph in which each vertex has a single outgoing edge) the vertices of which are the elements of S and the edges of which map an element to the corresponding function value, as shown in the figure. The set of vertices reachable from starting vertex x_0 form a subgraph with a shape resembling the Greek letter rho (ρ): a path of length μ from x_0 to a cycle of λ vertices.[2]

8.3 Computer representation

Generally, f will not be specified as a table of values, the way it is shown in the figure above. Rather, a cycle detection algorithm may be given access either to the sequence of values *xi*, or to a subroutine for calculating f. The task is to find λ and μ while examining as few values from the sequence or performing as few subroutine calls as possible. Typically, also, the space complexity of an algorithm for the cycle detection problem is of importance: we wish to solve the problem while using an amount of memory significantly smaller than it would take to store the entire sequence.

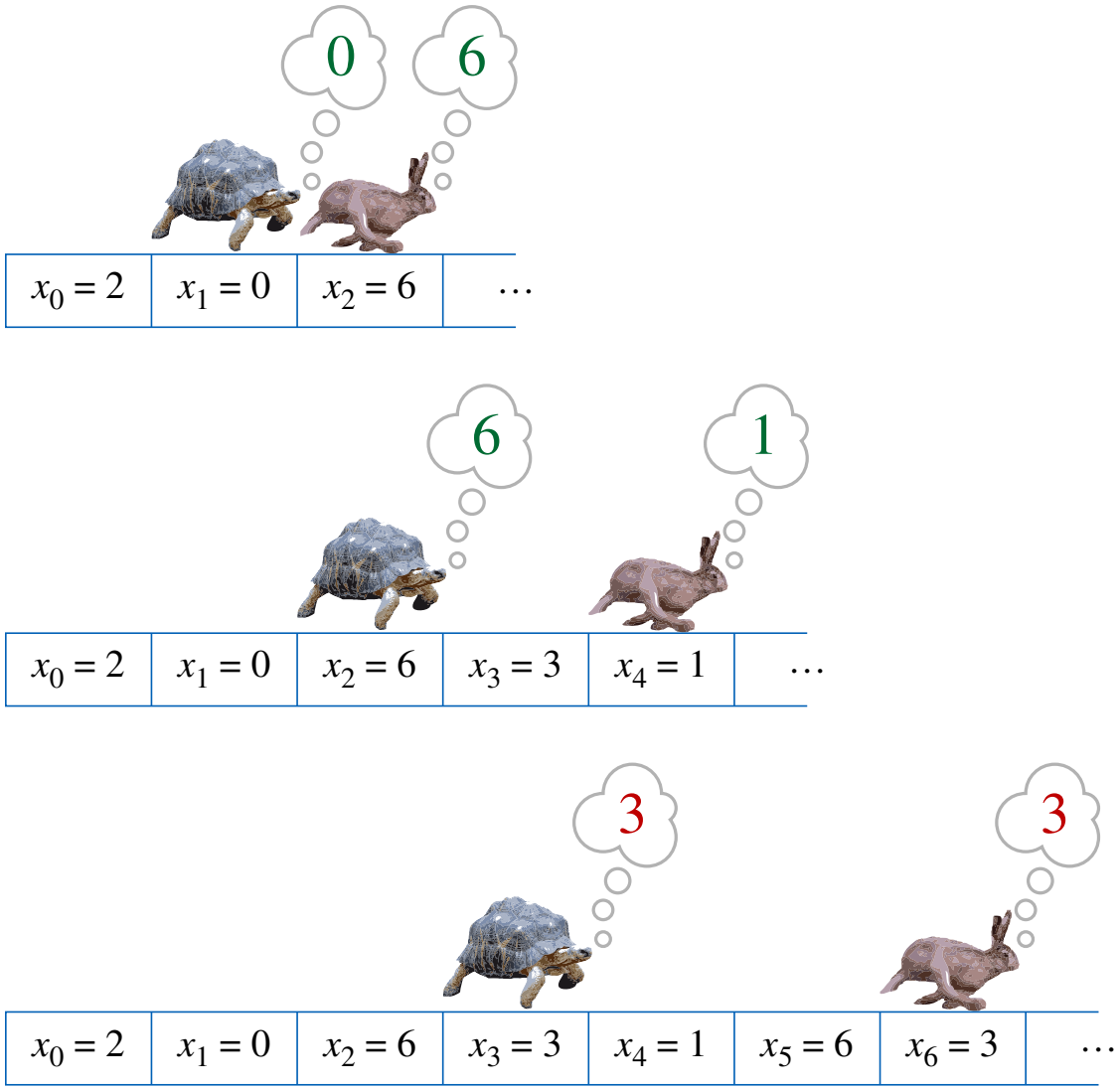
In some applications, and in particular in Pollard's rho algorithm for integer factorization, the algorithm has much more limited access to S and to f. In Pollard's rho algorithm, for instance, S is the set of integers modulo an unknown prime factor of the number to be factorized, so even the size of S is unknown to the algorithm. To allow cycle detection algorithms to be used with such limited knowledge, they may be designed based on the following capabilities. Initially, the algorithm is assumed to have in its memory an object representing a pointer to the starting value x_0. At any step, it may perform one of three actions: it may copy any pointer it has to another object in memory, it may apply f and replace any of its pointers by a pointer to the next object in the sequence, or it may apply a subroutine for determining whether two of its pointers represent equal values in the sequence. The equality test action may involve some nontrivial computation: for instance, in Pollard's rho algorithm, it is implemented by testing whether the difference between two stored values has a nontrivial greatest common divisor with the number to be factored.[2] In this context, by analogy to the pointer machine model of computation, an algorithm that only uses pointer copying, advancement within the

sequence, and equality tests may be called a pointer algorithm.

8.4 Algorithms

If the input is given as a subroutine for calculating f, the cycle detection problem may be trivially solved using only $\lambda + \mu$ function applications, simply by computing the sequence of values *xi* and using a data structure such as a hash table to store these values and test whether each subsequent value has already been stored. However, the space complexity of this algorithm is proportional to $\lambda + \mu$, unnecessarily large. Additionally, to implement this method as a pointer algorithm would require applying the equality test to each pair of values, resulting in quadratic time overall. Thus, research in this area has concentrated on two goals: using less space than this naive algorithm, and finding pointer algorithms that use fewer equality tests.

8.4.1 Tortoise and hare



Floyd's "tortoise and hare" cycle detection algorithm, applied to the sequence 2, 0, 6, 3, 1, 6, 3, 1, ...

Floyd's cycle-finding algorithm is a pointer algorithm that uses only two pointers, which move through the sequence at different speeds. It is also called the "tortoise and the hare algorithm", alluding to Aesop's fable of The Tortoise and the Hare.

The algorithm is named after Robert W. Floyd, who was credited with its invention by Donald Knuth.[3][4] However, the algorithm does not appear in Floyd's published work, and this may be a misattribution: Floyd describes algorithms for listing all simple cycles in a directed graph in a 1967 paper,[5] but this paper does not describe the cycle-finding problem in functional graphs that is the subject of this article. In fact, Knuth's statement (in 1969), attributing it to Floyd, without citation, is the first known appearance in print, and it thus may be a folk theorem, not attributable to a single individual.[6]

The key insight in the algorithm is that, for any integers $i \geq \mu$ and $k \geq 0$, $xi = xi + k\lambda$, where λ is the length of the loop to be found and μ is the index of the first element of the cycle. In particular, $i = k\lambda \geq \mu$, if and only if $xi = x_2i$. Thus, the algorithm only needs to check for repeated values of this special form, one twice as far from the start of the sequence as the other, to find a period ν of a repetition that is a multiple of λ. Once ν is found, the algorithm retraces the sequence from its start to find the first repeated value $x\mu$ in the sequence, using the fact that λ divides ν and therefore that $x\mu = x\mu + \nu$. Finally, once the value of μ is known it is trivial to find the length λ of the shortest repeating cycle, by searching for the first position $\mu + \lambda$ for which $x\mu + \lambda = x\mu$.

The algorithm thus maintains two pointers into the given sequence, one (the tortoise) at xi, and the other (the hare) at x_2i. At each step of the algorithm, it increases i by one, moving the tortoise one step forward and the hare two steps forward in the sequence, and then compares the sequence values at these two pointers. The smallest value of $i > 0$ for which the tortoise and hare point to equal values is the desired value ν.

The following Python code shows how this idea may be implemented as an algorithm.

def floyd(f, x0): # Main phase of algorithm: finding a repetition x_i = x_2i. # The hare moves twice as quickly as the tortoise and # the distance between them increases by 1 at each step. # Eventually they will both be inside the cycle and then, # at some point, the distance between them will be # divisible by the period λ. tortoise = f(x0) # f(x0) is the element/node next to x0. hare = f(f(x0)) while tortoise != hare: tortoise = f(tortoise) hare = f(f(hare)) # At this point the tortoise position, ν, which is also equal # to the distance between hare and tortoise, is divisible by # the period λ. So hare moving in circle one step at a time, # and tortoise (reset to x0) moving towards the circle, will # intersect at the beginning of the circle. Because the # distance between them is constant at 2ν, a multiple of λ, # they will agree as soon as the tortoise reaches index μ. # Find the position μ of first repetition. mu = 0 tortoise = x0 while tortoise != hare: tortoise = f(tortoise) hare = f(hare) # Hare and tortoise move at same speed mu += 1 # Find the length of the shortest cycle starting from x_μ # The hare moves one step at a time while tortoise is still. # lam is incremented until λ is found. lam = 1 hare = f(tortoise) while tortoise != hare: hare = f(hare) lam += 1 return lam, mu

This code only accesses the sequence by storing and copying pointers, function evaluations, and equality tests; therefore, it qualifies as a pointer algorithm. The algorithm uses $O(\lambda + \mu)$ operations of these types, and $O(1)$ storage space.[7]

8.4.2 Brent's algorithm

Richard P. Brent described an alternative cycle detection algorithm that, like the tortoise and hare algorithm, requires only two pointers into the sequence.[8] However, it is based on a different principle: searching for the smallest power of two 2^i that is larger than both λ and μ. For $i = 0, 1, 2, ...$, the algorithm compares x_{2^i-1} with each subsequent sequence value up to the next power of two, stopping when it finds a match. It has two advantages compared to the tortoise and hare algorithm: it finds the correct length λ of the cycle directly, rather than needing to search for it in a subsequent stage, and its steps involve only one evaluation of f rather than three.[9]

The following Python code shows how this technique works in more detail.

def brent(f, x0): # main phase: search successive powers of two power = lam = 1 tortoise = x0 hare = f(x0) # f(x0) is the element/node next to x0. while tortoise != hare: if power == lam: # time to start a new power of two? tortoise = hare power *= 2 lam = 0 hare = f(hare) lam += 1 # Find the position of the first repetition of length λ mu = 0 tortoise = hare = x0 for i in range(lam): # range(lam) produces a list with the values 0, 1, ... , lam-1 hare = f(hare) # The distance between the hare and tortoise is now λ. # Next, the hare and tortoise move at same speed until they agree while tortoise != hare: tortoise = f(tortoise) hare = f(hare) mu += 1 return lam, mu

Like the tortoise and hare algorithm, this is a pointer algorithm that uses $O(\lambda + \mu)$ tests and function evaluations and $O(1)$ storage space. It is not difficult to show that the number of function evaluations can never be higher than for Floyd's algorithm. Brent claims that, on average, his cycle finding algorithm runs around 36% more quickly than Floyd's and that it speeds up the Pollard rho algorithm by around 24%. He also performs an average case analysis for

a randomized version of the algorithm in which the sequence of indices traced by the slower of the two pointers is not the powers of two themselves, but rather a randomized multiple of the powers of two. Although his main intended application was in integer factorization algorithms, Brent also discusses applications in testing pseudorandom number generators.[8]

8.4.3 Time–space tradeoffs

A number of authors have studied techniques for cycle detection that use more memory than Floyd's and Brent's methods, but detect cycles more quickly. In general these methods store several previously-computed sequence values, and test whether each new value equals one of the previously-computed values. In order to do so quickly, they typically use a hash table or similar data structure for storing the previously-computed values, and therefore are not pointer algorithms: in particular, they usually cannot be applied to Pollard's rho algorithm. Where these methods differ is in how they determine which values to store. Following Nivasch,[10] we survey these techniques briefly.

- Brent[8] already describes variations of his technique in which the indices of saved sequence values are powers of a number R other than two. By choosing R to be a number close to one, and storing the sequence values at indices that are near a sequence of consecutive powers of R, a cycle detection algorithm can use a number of function evaluations that is within an arbitrarily small factor of the optimum $\lambda + \mu$.[11][12]

- Sedgewick, Szymanski, and Yao[13] provide a method that uses M memory cells and requires in the worst case only $(\lambda + \mu)(1 + cM^{-1/2})$ function evaluations, for some constant c, which they show to be optimal. The technique involves maintaining a numerical parameter d, storing in a table only those positions in the sequence that are multiples of d, and clearing the table and doubling d whenever too many values have been stored.

- Several authors have described *distinguished point* methods that store function values in a table based on a criterion involving the values, rather than (as in the method of Sedgewick et al.) based on their positions. For instance, values equal to zero modulo some value d might be stored.[14][15] More simply, Nivasch[10] credits D. P. Woodruff with the suggestion of storing a random sample of previously seen values, making an appropriate random choice at each step so that the sample remains random.

- Nivasch[10] describes an algorithm that does not use a fixed amount of memory, but for which the expected amount of memory used (under the assumption that the input function is random) is logarithmic in the sequence length. An item is stored in the memory table, with this technique, when no later item has a smaller value. As Nivasch shows, the items with this technique can be maintained using a stack data structure, and each successive sequence value need be compared only to the top of the stack. The algorithm terminates when the repeated sequence element with smallest value is found. Running the same algorithm with multiple stacks, using random permutations of the values to reorder the values within each stack, allows a time–space tradeoff similar to the previous algorithms. However, even the version of this algorithm with a single stack is not a pointer algorithm, due to the comparisons needed to determine which of two values is smaller.

Any cycle detection algorithm that stores at most M values from the input sequence must perform at least $(\lambda+\mu)(1+\frac{1}{M-1})$ function evaluations.[16][17]

8.5 Applications

Cycle detection has been used in many applications.

- Determining the cycle length of a pseudorandom number generator is one measure of its strength. This is the application cited by Knuth in describing Floyd's method.[3] Brent[8] describes the results of testing a linear congruential generator in this fashion; its period turned out to be significantly smaller than advertised. For more complex generators, the sequence of values in which the cycle is to be found may not represent the output of the generator, but rather its internal state.

- Several number-theoretic algorithms are based on cycle detection, including Pollard's rho algorithm for integer factorization[18] and his related kangaroo algorithm for the discrete logarithm problem.[19]

- In cryptographic applications, the ability to find two distinct values $x_{\mu-1}$ and $x_{\lambda+\mu-1}$ mapped by some cryptographic function f to the same value x_μ may indicate a weakness in f. For instance, Quisquater and Delescaille[15] apply cycle detection algorithms in the search for a message and a pair of Data Encryption Standard keys that map that message to the same encrypted value; Kaliski, Rivest, and Sherman[20] also use cycle detection algorithms to attack DES. The technique may also be used to find a collision in a cryptographic hash function.[21]
- Cycle detection may be helpful as a way of discovering infinite loops in certain types of computer programs.[22]
- Periodic configurations in cellular automaton simulations may be found by applying cycle detection algorithms to the sequence of automaton states.[10]
- Shape analysis of linked list data structures is a technique for verifying the correctness of an algorithm using those structures. If a node in the list incorrectly points to an earlier node in the same list, the structure will form a cycle that can be detected by these algorithms.[23] In Common Lisp, the S-expression printer, under control of the *print-circle* variable, detects circular list structure and prints it compactly.
- Teske[12] describes applications in computational group theory: determining the structure of an Abelian group from a set of its generators. The cryptographic algorithms of Kaliski et al.[20] may also be viewed as attempting to infer the structure of an unknown group.
- Fich (1981) briefly mentions an application to computer simulation of celestial mechanics, which she attributes to William Kahan. In this application, cycle detection in the phase space of an orbital system may be used to determine whether the system is periodic to within the accuracy of the simulation.[16]

8.6 References

[1] Joux, Antoine (2009), *Algorithmic Cryptanalysis*, CRC Press, p. 223, ISBN 9781420070033.

[2] Joux (2009), p. 224.

[3] Knuth, Donald E. (1969), *The Art of Computer Programming, vol. II: Seminumerical Algorithms*, Addison-Wesley, p. 7, exercises 6 and 7

[4] *Handbook of Applied Cryptography,* by Alfred J. Menezes, Paul C. van Oorschot, Scott A. Vanstone, p. 125, describes this algorithm and others

[5] Floyd, R.W. (1967), "Non-deterministic Algorithms", *J. ACM*, **14** (4): 636–644, doi:10.1145/321420.321422

[6] *The Hash Function BLAKE,* by Jean-Philippe Aumasson, Willi Meier, Raphael C.-W. Phan, Luca Henzen (2015), p. 21, footnote 8

[7] Joux (2009), Section 7.1.1, Floyd's cycle-finding algorithm, pp. 225–226.

[8] Brent, R. P. (1980), "An improved Monte Carlo factorization algorithm" (PDF), *BIT Numerical Mathematics* , **20** (2): 176–184, doi:10.1007/BF01933190.

[9] Joux (2009), Section 7.1.2, Brent's cycle-finding algorithm, pp. 226–227.

[10] Nivasch, Gabriel (2004), "Cycle detection using a stack", *Information Processing Letters*, **90** (3): 135–140, doi:10.1016/j.ipl.2004.01.016.

[11] Schnorr, Claus P.; Lenstra, Hendrik W. (1984), "A Monte Carlo factoring algorithm with linear storage", *Mathematics of Computation*, **43** (167): 289–311, doi:10.2307/2007414, JSTOR 2007414.

[12] Teske, Edlyn (1998), "A space-efficient algorithm for group structure computation", *Mathematics of Computation*, **67** (224): 1637–1663, doi:10.1090/S0025-5718-98-00968-5.

[13] Sedgewick, Robert; Szymanski, Thomas G.; Yao, Andrew C.-C. (1982), "The complexity of finding cycles in periodic functions", *SIAM Journal on Computing*, **11** (2): 376–390, doi:10.1137/0211030.

[14] van Oorschot, Paul C.; Wiener, Michael J. (1999), "Parallel collision search with cryptanalytic applications", *Journal of Cryptology*, **12** (1): 1–28, doi:10.1007/PL00003816.

[15] Quisquater, J.-J.; Delescaille, J.-P., "How easy is collision search? Application to DES", *Advances in Cryptology – EUROCRYPT '89, Workshop on the Theory and Application of Cryptographic Techniques*, Lecture Notes in Computer Science, **434**, Springer-Verlag, pp. 429–434.

[16] Fich, Faith Ellen (1981), "Lower bounds for the cycle detection problem", *Proc. 13th ACM Symposium on Theory of Computing*, pp. 96–105, doi:10.1145/800076.802462.

[17] Allender, Eric W.; Klawe, Maria M. (1985), "Improved lower bounds for the cycle detection problem", *Theoretical Computer Science*, **36** (2–3): 231–237, doi:10.1016/0304-3975(85)90044-1.

[18] Pollard, J. M. (1975), "A Monte Carlo method for factorization", *BIT*, **15** (3): 331–334, doi:10.1007/BF01933667.

[19] Pollard, J. M. (1978), "Monte Carlo methods for index computation (mod p)", *Mathematics of Computation*, American Mathematical Society, **32** (143): 918–924, doi:10.2307/2006496, JSTOR 2006496.

[20] Kaliski, Burton S., Jr.; Rivest, Ronald L.; Sherman, Alan T. (1988), "Is the Data Encryption Standard a group? (Results of cycling experiments on DES)", *Journal of Cryptology*, **1** (1): 3–36, doi:10.1007/BF00206323.

[21] Joux (2009), Section 7.5, Collisions in hash functions, pp. 242–245.

[22] Van Gelder, Allen (1987), "Efficient loop detection in Prolog using the tortoise-and-hare technique", *Journal of Logic Programming*, **4** (1): 23–31, doi:10.1016/0743-1066(87)90020-3.

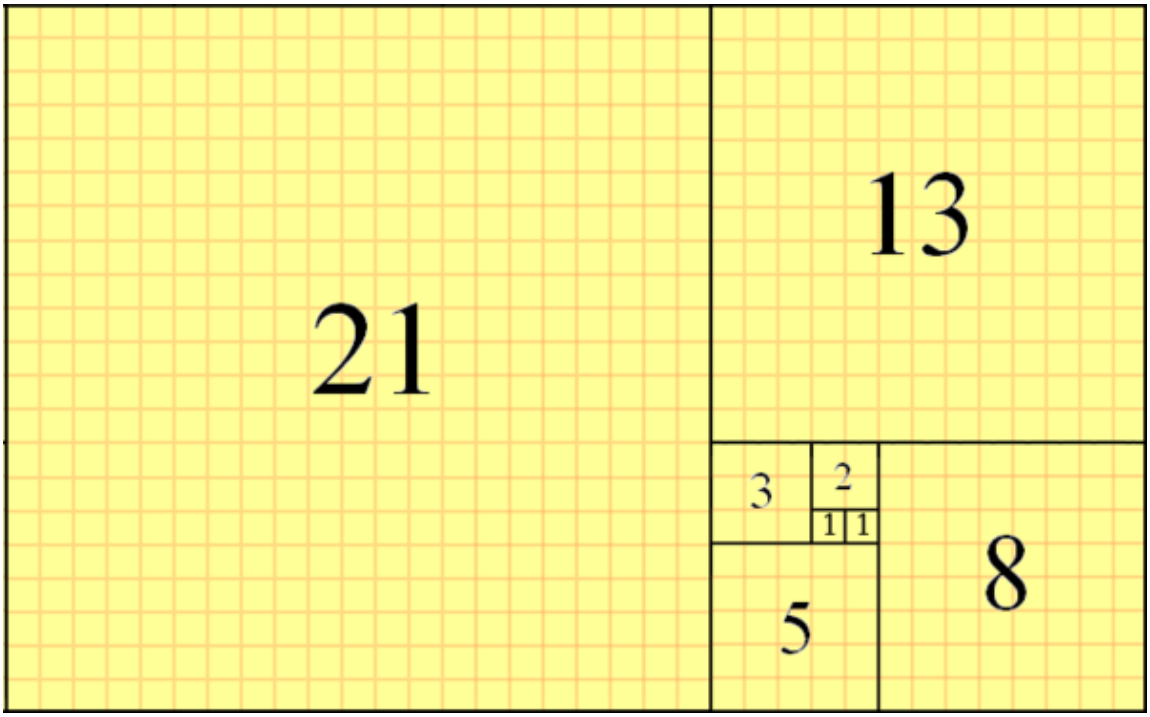
[23] Auguston, Mikhail; Hon, Miu Har (1997), "Assertions for Dynamic Shape Analysis of List Data Structures", *AADEBUG '97, Proceedings of the Third International Workshop on Automatic Debugging*, Linköping Electronic Articles in Computer and Information Science, Linköping University, pp. 37–42.

8.7 External links

- Gabriel Nivasch, The Cycle Detection Problem and the Stack Algorithm
- Tortoise and Hare, Portland Pattern Repository
- Floyd's Cycle Detection Algorithm (The Tortoise and the Hare)
- Brent's Cycle Detection Algorithm (The Teleporting Turtle)

Chapter 9

Fibonacci number



A tiling with squares whose side lengths are successive Fibonacci numbers

In mathematics, the **Fibonacci numbers** are the numbers in the following integer sequence, called the **Fibonacci sequence**, and characterized by the fact that every number in it is the sum of the two preceding ones:[1][2]

$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ \ldots$

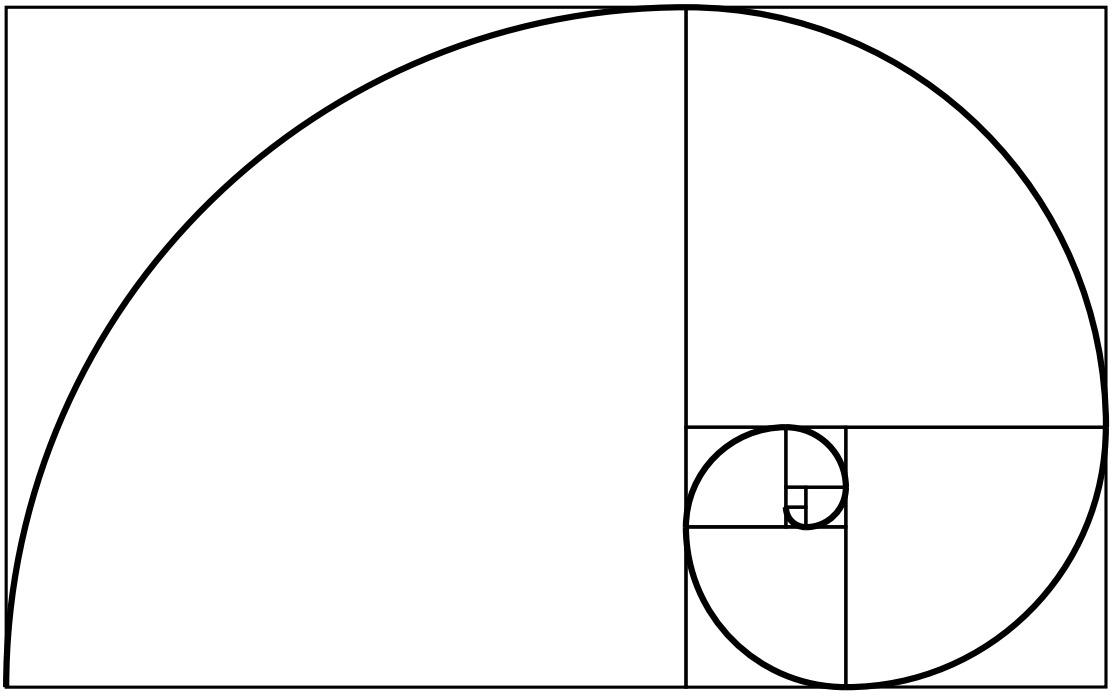
Often, especially in modern usage, the sequence is extended by one more initial term:

$0,\ 1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144,\ \ldots$.[3]

By definition, the first two numbers in the Fibonacci sequence are either 1 and 1, or 0 and 1, depending on the chosen starting point of the sequence, and each subsequent number is the sum of the previous two.

In mathematical terms, the sequence *Fn* of Fibonacci numbers is defined by the recurrence relation

$F_n = F_{n-1} + F_{n-2},$



The Fibonacci spiral: an approximation of the golden spiral created by drawing circular arcs connecting the opposite corners of squares in the Fibonacci tiling;[4] this one uses squares of sizes 1, 1, 2, 3, 5, 8, 13, 21, and 34.

with seed values[1][2]

$$F_1 = 1,\ F_2 = 1$$

or[5]

$$F_0 = 0,\ F_1 = 1.$$

The Fibonacci sequence is named after Italian mathematician Leonardo of Pisa, known as Fibonacci. His 1202 book *Liber Abaci* introduced the sequence to Western European mathematics,[6] although the sequence had been described earlier as Virahanka numbers in Indian mathematics.[7][8][9] By modern convention, the sequence begins either with $F_0 = 0$ or with $F_1 = 1$. The sequence described in *Liber Abaci* began with $F_1 = 1$.

Fibonacci numbers are closely related to Lucas numbers L_n in that they form a complementary pair of Lucas sequences $U_n(1,-1) = F_n$ and $V_n(1,-1) = L_n$. They are intimately connected with the golden ratio; for example, the closest rational approximations to the ratio are 2/1, 3/2, 5/3, 8/5, ...

Fibonacci numbers appear unexpectedly often in mathematics, so much so that there is an entire journal dedicated to their study, the Fibonacci Quarterly. Applications of Fibonacci numbers include computer algorithms such as the Fibonacci search technique and the Fibonacci heap data structure, and graphs called Fibonacci cubes used for interconnecting parallel and distributed systems. They also appear in biological settings,[10] such as branching in trees, phyllotaxis (the arrangement of leaves on a stem), the fruit sprouts of a pineapple,[11] the flowering of an artichoke, an uncurling fern and the arrangement of a pine cone's bracts.[12]

9.1 Origins

The Fibonacci sequence appears in Indian mathematics, in connection with Sanskrit prosody.[8][13] In the Sanskrit tradition of prosody, there was interest in enumerating all patterns of long (L) syllables that are 2 units of duration, and short (S) syllables that are 1 unit of duration. Counting the different patterns of L and S of a given duration results in the Fibonacci numbers: the number of patterns that are m short syllables long is the Fibonacci number Fm_{+1}.[9]

Susantha Goonatilake writes that the development of the Fibonacci sequence "is attributed in part to Pingala (200 BC), later being associated with Virahanka (c. 700 AD), Gopāla (c. 1135), and Hemachandra (c. 1150)".[7] Parmanand Singh cites Pingala's cryptic formula *misrau cha* ("the two are mixed") and cites scholars who interpret it in context as saying that the cases for *m* beats (*Fm+1*) is obtained by adding a [S] to *Fm* cases and [L] to the Fm_{-1} cases. He dates Pingala before 450 BC.[14]

However, the clearest exposition of the sequence arises in the work of Virahanka (c. 700 AD), whose own work is lost, but is available in a quotation by Gopala (c. 1135):

> Variations of two earlier meters [is the variation]... For example, for [a meter of length] four, variations of meters of two [and] three being mixed, five happens. [works out examples 8, 13, 21]... In this way, the process should be followed in all *mātrā-vṛttas* [prosodic combinations].[15]

The sequence is also discussed by Gopala (before 1135 AD) and by the Jain scholar Hemachandra (c. 1150).

Outside of India, the Fibonacci sequence first appears in the book *Liber Abaci* (1202) by Fibonacci.[6] Fibonacci considers the growth of an idealized (biologically unrealistic) rabbit population, assuming that: a newly born pair of rabbits, one male, one female, are put in a field; rabbits are able to mate at the age of one month so that at the end of its second month a female can produce another pair of rabbits; rabbits never die and a mating pair always produces one new pair (one male, one female) every month from the second month on. The puzzle that Fibonacci posed was: how many pairs will there be in one year?

- At the end of the first month, they mate, but there is still only 1 pair.
- At the end of the second month the female produces a new pair, so now there are 2 pairs of rabbits in the field.
- At the end of the third month, the original female produces a second pair, making 3 pairs in all in the field.
- At the end of the fourth month, the original female has produced yet another new pair, the female born two months ago produces her first pair also, making 5 pairs.

At the end of the *n*th month, the number of pairs of rabbits is equal to the number of new pairs (which is the number of pairs in month $n - 2$) plus the number of pairs alive last month ($n - 1$). This is the *n*th Fibonacci number.[16]

The name "Fibonacci sequence" was first used by the 19th-century number theorist Édouard Lucas.[17]

9.2 List of Fibonacci numbers

The first 21 Fibonacci numbers *Fn* for $n = 0, 1, 2, \ldots, 20$ are:[18]

The sequence can also be extended to negative index *n* using the re-arranged recurrence relation

$$F_{n-2} = F_n - F_{n-1},$$

which yields the sequence of "negafibonacci" numbers[19] satisfying

$$F_{-n} = (-1)^{n+1} F_n.$$

Thus the bidirectional sequence is

9.3 Use in mathematics

The Fibonacci numbers occur in the sums of "shallow" diagonals in Pascal's triangle (see Binomial coefficient).[20]

$$F_n = \sum_{k=0}^{\left\lfloor \frac{n-1}{2} \right\rfloor} \binom{n-k-1}{k}$$

These numbers also give the solution to certain enumerative problems.[21] The most common such problem is that of counting the number of compositions of 1s and 2s that sum to a given total n: there are Fn_{+1} ways to do this.

For example, if $n = 5$, then $Fn_{+1} = F_6 = 8$ counts the eight compositions:

$$1+1+1+1+1 = 1+1+1+2 = 1+1+2+1 = 1+2+1+1 = 2+1+1+1 = 2+2+1 = 2+1+2 = 1+2+2,$$

all of which sum to $n = 5 = 6-1$.

The Fibonacci numbers can be found in different ways among the set of binary strings, or equivalently, among the subsets of a given set.

- The number of binary strings of length n without consecutive 1s is the Fibonacci number Fn_{+2}. For example, out of the 16 binary strings of length 4, there are $F_6 = 8$ without consecutive 1s – they are 0000, 0001, 0010, 0100, 0101, 1000, 1001 and 1010. By symmetry, the number of strings of length n without consecutive 0s is also Fn_{+2}. Equivalently, Fn_{+2} is the number of subsets $S \subset \{1,...,n\}$ without consecutive integers: $\{i, i+1\} \not\subset S$ for every i. The symmetric statement is: Fn_{+2} is the number of subsets $S \subset \{1,...,n\}$ without two consecutive skipped integers: that is, $S = \{a_1 < \ldots < a_k\}$ with $a_{i+1} \leq a_i + 2$.
- The number of binary strings of length n without an odd number of consecutive 1s is the Fibonacci number $Fn+1$. For example, out of the 16 binary strings of length 4, there are $F_5 = 5$ without an odd number of consecutive 1s – they are 0000, 0011, 0110, 1100, 1111. Equivalently, the number of subsets $S \subset \{1,...,n\}$ without an odd number of consecutive integers is $Fn+1$.
- The number of binary strings of length n without an even number of consecutive 0s or 1s is $2Fn$. For example, out of the 16 binary strings of length 4, there are $2F_4 = 6$ without an even number of consecutive 0s or 1s – they are 0001, 0111, 0101, 1000, 1010, 1110. There is an equivalent statement about subsets.

9.4 Relation to the golden ratio

9.4.1 Closed-form expression

Like every sequence defined by a linear recurrence with constant coefficients, the Fibonacci numbers have a closed-form solution. It has become known as "Binet's formula", even though it was already known by Abraham de Moivre:[22]

$$F_n = \frac{\varphi^n - \psi^n}{\varphi - \psi} = \frac{\varphi^n - \psi^n}{\sqrt{5}}$$

where

$$\varphi = \frac{1+\sqrt{5}}{2} \approx 1.61803\,39887\cdots$$

is the golden ratio (OEIS A001622), and

$$\psi = \tfrac{1-\sqrt{5}}{2} = 1 - \varphi = -\tfrac{1}{\varphi} \approx -0.61803\,39887\cdots$$ [23]

Since $\psi = -\frac{1}{\varphi}$, this formula can also be written as

$F_n = \frac{\varphi^n - (-\varphi)^{-n}}{\sqrt{5}} = \frac{\varphi^n - (-\varphi)^{-n}}{2\varphi - 1}$

To see this,[24] note that φ and ψ are both solutions of the equations

$$x^2 = x + 1,\ x^n = x^{n-1} + x^{n-2},$$

so the powers of φ and ψ satisfy the Fibonacci recursion. In other words,

$$\varphi^n = \varphi^{n-1} + \varphi^{n-2}$$

and

$$\psi^n = \psi^{n-1} + \psi^{n-2}\,.$$

It follows that for any values *a* and *b*, the sequence defined by

$$U_n = a\varphi^n + b\psi^n$$

satisfies the same recurrence

$$U_n = a\varphi^{n-1} + b\psi^{n-1} + a\varphi^{n-2} + b\psi^{n-2} = U_{n-1} + U_{n-2}.$$

If *a* and *b* are chosen so that U_0 = 0 and U_1 = 1 then the resulting sequence *Un* must be the Fibonacci sequence. This is the same as requiring *a* and *b* satisfy the system of equations:

$$\begin{cases} a + b = 0 \\ \varphi a + \psi b = 1 \end{cases}$$

which has solution

$$a = \frac{1}{\varphi - \psi} = \frac{1}{\sqrt{5}},\ b = -a$$

producing the required formula.

Taking U_0 and U_1 to be variables, a more general solution can be found for any starting values:

$$U_n = a\varphi^n + b\psi^n$$

where

$$a = \frac{U_1 - U_0\psi}{\sqrt{5}}$$

$$b = \frac{U_0\varphi - U_1}{\sqrt{5}}$$

9.4.2 Computation by rounding

Since

$$\left|\frac{\psi^n}{\sqrt{5}}\right| < \frac{1}{2}$$

for all $n \geq 0$, the number Fn is the closest integer to $\frac{\varphi^n}{\sqrt{5}}$. Therefore, it can be found by rounding, that is by the use of the nearest integer function:

$$F_n = \left[\frac{\varphi^n}{\sqrt{5}}\right],\ n \geq 0,$$

or in terms of the floor function:

$$F_n = \left\lfloor\frac{\varphi^n}{\sqrt{5}} + \frac{1}{2}\right\rfloor,\ n \geq 0.$$

Similarly, if we already know that the number $F > 1$ is a Fibonacci number, we can determine its index within the sequence by

$$n(F) = \left\lfloor \log_\varphi\left(F \cdot \sqrt{5} + \frac{1}{2}\right)\right\rfloor$$

9.4.3 Limit of consecutive quotients

Johannes Kepler observed that the ratio of consecutive Fibonacci numbers converges. He wrote that "as 5 is to 8 so is 8 to 13, practically, and as 8 is to 13, so is 13 to 21 almost", and concluded that the limit approaches the golden ratio φ.[25] [26]

$$\lim_{n\to\infty}\frac{F_{n+1}}{F_n} = \varphi$$

This convergence holds regardless of the starting values, excluding 0, 0. This can be derived from Binet's formula. For example, the initial values 3 and 2 generate the sequence 3, 2, 5, 7, 12, 19, 31, 50, 81, 131, 212, 343, 555, …, etc. The ratio of consecutive terms in this sequence shows the same convergence towards the golden ratio.

Another consequence is that the limit of the ratio of two Fibonacci numbers offset by a particular finite deviation in index corresponds to the golden ratio raised by that deviation. Or, in other words:

$$\lim_{n\to\infty}\frac{F_{n+\alpha}}{F_n} = \varphi^\alpha$$

9.4.4 Decomposition of powers of the golden ratio

Since the golden ratio satisfies the equation

$$\varphi^2 = \varphi + 1,$$

this expression can be used to decompose higher powers φ^n as a linear function of lower powers, which in turn can be decomposed all the way down to a linear combination of φ and 1. The resulting recurrence relationships yield Fibonacci numbers as the linear coefficients:

$\varphi^n = F_n\varphi + F_{n-1}$.

This equation can be proved by induction on n.

This expression is also true for $n < 1$ if the Fibonacci sequence Fn is extended to negative integers using the Fibonacci rule $F_n = F_{n-1} + F_{n-2}$.

9.5 Matrix form

A 2-dimensional system of linear difference equations that describes the Fibonacci sequence is

$$\begin{aligned}\begin{pmatrix} F_{k+2} \\ F_{k+1} \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} F_{k+1} \\ F_k \end{pmatrix} \\ \vec{F}_{k+1} &= \mathbf{A}\vec{F}_k \ ,\end{aligned}$$

which yields $\vec{F}_n = \mathbf{A}^n \vec{F}_0$. As the eigenvalues of the matrix $\boldsymbol{A}$ are $\varphi = \frac{1}{2}(1+\sqrt{5})$ and $-\varphi^{-1} = \frac{1}{2}(1-\sqrt{5})$, for the respective eigenvectors $\vec{\mu} = \binom{\varphi}{1}$ and $\vec{\nu} = \binom{-\varphi^{-1}}{1}$, and the initial value $\vec{F}_0 = \binom{1}{0} = \frac{1}{\sqrt{5}}\vec{\mu} - \frac{1}{\sqrt{5}}\vec{\nu}$, the nth term is

$$\begin{aligned}\vec{F}_n &= \frac{1}{\sqrt{5}} A^n \vec{\mu} - \frac{1}{\sqrt{5}} A^n \vec{\nu} \\ &= \frac{1}{\sqrt{5}} \varphi^n \vec{\mu} - \frac{1}{\sqrt{5}} (-\varphi)^{-n} \vec{\nu} \\ &= \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2} \right)^n \begin{pmatrix} \varphi \\ 1 \end{pmatrix} - \frac{1}{\sqrt{5}} \cdot \left(\frac{1-\sqrt{5}}{2} \right)^n \begin{pmatrix} -\varphi^{-1} \\ 1 \end{pmatrix},\end{aligned}$$

from which the nth element in the Fibonacci series as an analytic function of n is now read off directly:

$$F_n = \frac{1}{\sqrt{5}} \cdot \left(\frac{1+\sqrt{5}}{2} \right)^n - \frac{1}{\sqrt{5}} \cdot \left(\frac{1-\sqrt{5}}{2} \right)^n .$$

Equivalently, the same computation is performed by diagonalization of $\boldsymbol{A}$ through use of its eigendecomposition:

$$\begin{aligned}A &= S\Lambda S^{-1}, \\ A^n &= S\Lambda^n S^{-1},\end{aligned}$$

where $\Lambda = \begin{pmatrix} \varphi & 0 \\ 0 & -\varphi^{-1} \end{pmatrix}$ and $S = \begin{pmatrix} \varphi & -\varphi^{-1} \\ 1 & 1 \end{pmatrix}$. The closed-form expression for the nth element in the Fibonacci series is therefore given by

$$\begin{aligned}\begin{pmatrix} F_{n+1} \\ F_n \end{pmatrix} &= A^n \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} \\ &= S\Lambda^n S^{-1} \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} \\ &= S \begin{pmatrix} \varphi^n & 0 \\ 0 & (-\varphi)^{-n} \end{pmatrix} S^{-1} \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} \\ &= \begin{pmatrix} \varphi & -\varphi^{-1} \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \varphi^n & 0 \\ 0 & (-\varphi)^{-n} \end{pmatrix} \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & \varphi^{-1} \\ -1 & \varphi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix},\end{aligned}$$

which again yields

$$F_n = \frac{\varphi^n - (-\varphi)^{-n}}{\sqrt{5}}.$$

The matrix $\boldsymbol{A}$ has a determinant of −1, and thus it is a 2×2 unimodular matrix.

This property can be understood in terms of the continued fraction representation for the golden ratio:

$$\varphi = 1 + \cfrac{1}{1 + \cfrac{1}{1 + \cfrac{1}{1 + \ddots}}}$$

The Fibonacci numbers occur as the ratio of successive convergents of the continued fraction for φ, and the matrix formed from successive convergents of any continued fraction has a determinant of +1 or −1. The matrix representation gives the following closed expression for the Fibonacci numbers:

$$\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^n = \begin{pmatrix} F_{n+1} & F_n \\ F_n & F_{n-1} \end{pmatrix}.$$

Taking the determinant of both sides of this equation yields Cassini's identity,

$$(-1)^n = F_{n+1}F_{n-1} - F_n^2\,.$$

Moreover, since $\boldsymbol{A}^n\,\boldsymbol{A}^m = \boldsymbol{A}^{n+m}$ for any square matrix $\boldsymbol{A}$, the following identities can be derived (they are obtained form two different coefficients of the matrix product, and one may easily deduce the second one from the first one by changing n into $n + 1$),

$$\begin{aligned} F_m F_n + F_{m-1} F_{n-1} &= F_{m+n-1}, \\ F_m F_{n+1} + F_{m-1} F_n &= F_{m+n}. \end{aligned}$$

In particular, with $m = n$,

$$\begin{aligned} F_{2n-1} &= F_n^2 + F_{n-1}^2 \\ F_{2n} &= (F_{n-1} + F_{n+1})F_n \\ &= (2F_{n-1} + F_n)F_n. \end{aligned}$$

These last two identities provide a way to compute Fibonacci numbers recursively in O(log(n)) arithmetic operations and in time O($M(n)$ log(n)), where $M(n)$ is the time for the multiplication of two numbers of n digits. This matches the time for computing the nth Fibonacci number from the closed-form matrix formula, but with fewer redundant steps if one avoids recomputing an already computed Fibonacci number (recursion with memoization).[27]

9.6 Recognizing Fibonacci numbers

The question may arise whether a positive integer x is a Fibonacci number. This is true if and only if one or both of $5x^2 + 4$ or $5x^2 - 4$ is a perfect square.[28] This is because Binet's formula above can be rearranged to give

$$n = \log_\varphi\left(\frac{F_n\sqrt{5} + \sqrt{5F_n^2 \pm 4}}{2}\right)$$

which allows one to find the position in the sequence of a given Fibonacci number.

This formula must return an integer for all n, so the radical expression must be an integer (otherwise the logarithm does not even return a rational number).

9.7 Combinatorial identities

Most identities involving Fibonacci numbers can be proved using combinatorial arguments using the fact that Fn can be interpreted as the number of sequences of 1s and 2s that sum to $n - 1$. This can be taken as the definition of Fn, with the convention that $F_0 = 0$, meaning no sum adds up to −1, and that $F_1 = 1$, meaning the empty sum "adds up" to 0. Here, the order of the summand matters. For example, 1 + 2 and 2 + 1 are considered two different sums.

For example, the recurrence relation

$$F_n = F_{n-1} + F_{n-2},$$

or in words, the nth Fibonacci number is the sum of the previous two Fibonacci numbers, may be shown by dividing the Fn sums of 1s and 2s that add to $n - 1$ into two non-overlapping groups. One group contains those sums whose first term is 1 and the other those sums whose first term is 2. In the first group the remaining terms add to $n - 2$, so it has $Fn-_1$ sums, and in the second group the remaining terms add to $n - 3$, so there are Fn_{-2} sums. So there are a total of $Fn_{-1} + Fn_{-2}$ sums altogether, showing this is equal to Fn.

Similarly, it may be shown that the sum of the first Fibonacci numbers up to the nth is equal to the $(n + 2)$-nd Fibonacci number minus 1.[29] In symbols:

$$\sum_{i=1}^{n} F_i = F_{n+2} - 1$$

This is done by dividing the sums adding to $n + 1$ in a different way, this time by the location of the first 2. Specifically, the first group consists of those sums that start with 2, the second group those that start 1 + 2, the third 1 + 1 + 2, and so on, until the last group, which consists of the single sum where only 1's are used. The number of sums in the first group is $F(n)$, $F(n - 1)$ in the second group, and so on, with 1 sum in the last group. So the total number of sums is $F(n) + F(n - 1) + ... + F(1) + 1$ and therefore this quantity is equal to $F(n + 2)$.

A similar argument, grouping the sums by the position of the first 1 rather than the first 2, gives two more identities:

$$\sum_{i=0}^{n-1} F_{2i+1} = F_{2n}$$

and

$$\sum_{i=1}^{n} F_{2i} = F_{2n+1} - 1.$$

In words, the sum of the first Fibonacci numbers with odd index up to F_2n_{-1} is the $(2n)$th Fibonacci number, and the sum of the first Fibonacci numbers with even index up to F_2n is the $(2n + 1)$th Fibonacci number minus 1.[30]

A different trick may be used to prove

$$\sum_{i=1}^{n} {F_i}^2 = F_n F_{n+1},$$

or in words, the sum of the squares of the first Fibonacci numbers up to Fn is the product of the nth and $(n + 1)$th Fibonacci numbers. In this case note that Fibonacci rectangle of size Fn by $F(n + 1)$ can be decomposed into squares of size Fn, Fn_{-1}, and so on to $F_1 = 1$, from which the identity follows by comparing areas.

9.8 Other identities

Numerous other identities can be derived using various methods. Some of the most noteworthy are:[31]

9.8.1 Cassini and Catalan's identities

Main article: Cassini and Catalan identities

Cassini's identity states that

$$F_n^2 - F_{n+1}F_{n-1} = (-1)^{n-1}$$

Catalan's identity is a generalization:

$$F_n^2 - F_{n+r}F_{n-r} = (-1)^{n-r}F_r^2$$

9.8.2 d'Ocagne's identity

$$F_mF_{n+1} - F_{m+1}F_n = (-1)^nF_{m-n}$$
$$F_{2n} = F_{n+1}^2 - F_{n-1}^2 = F_n\left(F_{n+1} + F_{n-1}\right) = F_nL_n$$

where *Ln* is the *n*'th Lucas number. The last is an identity for doubling *n*; other identities of this type are

$$F_{3n} = 2F_n^3 + 3F_nF_{n+1}F_{n-1} = 5F_n^3 + 3(-1)^nF_n$$

by Cassini's identity.

$$F_{3n+1} = F_{n+1}^3 + 3F_{n+1}F_n^2 - F_n^3$$
$$F_{3n+2} = F_{n+1}^3 + 3F_{n+1}^2F_n + F_n^3$$
$$F_{4n} = 4F_nF_{n+1}\left(F_{n+1}^2 + 2F_n^2\right) - 3F_n^2\left(F_n^2 + 2F_{n+1}^2\right)$$

These can be found experimentally using lattice reduction, and are useful in setting up the special number field sieve to factorize a Fibonacci number.

More generally,[31]

$$F_{kn+c} = \sum_{i=0}^{k}\binom{k}{i}F_{c-i}F_n^iF_{n+1}^{k-i}.$$

Putting *k* = 2 in this formula, one gets again the formulas of the end of above section Matrix form.

9.9 Power series

The generating function of the Fibonacci sequence is the power series

$$s(x) = \sum_{k=0}^{\infty}F_kx^k.$$

This series is convergent for $|x| < \frac{1}{\varphi}$, and its sum has a simple closed-form:[32]

$$s(x) = \frac{x}{1 - x - x^2}$$

This can be proved by using the Fibonacci recurrence to expand each coefficient in the infinite sum:

$$\begin{aligned} s(x) &= \sum_{k=0}^{\infty} F_k x^k \\ &= F_0 + F_1 x + \sum_{k=2}^{\infty} \left(F_{k-1} + F_{k-2}\right) x^k \\ &= x + \sum_{k=2}^{\infty} F_{k-1} x^k + \sum_{k=2}^{\infty} F_{k-2} x^k \\ &= x + x \sum_{k=0}^{\infty} F_k x^k + x^2 \sum_{k=0}^{\infty} F_k x^k \\ &= x + x s(x) + x^2 s(x). \end{aligned}$$

Solving the equation

$$s(x) = x + x s(x) + x^2 s(x)$$

for $s(x)$ results in the above closed form.

If x is the reciprocal of an integer k that is greater than 1, the closed form of the series becomes

$$\sum_{n=0}^{\infty} \frac{F_n}{k^n} = \frac{k}{k^2 - k - 1}.$$

In particular,

$$\sum_{n=1}^{\infty} \frac{F_n}{10^{m(n+1)}} = \frac{1}{10^{2m} - 10^m - 1}$$

for all positive integers m.

Some math puzzle-books present as curious the particular value that comes from m=1, which is $\frac{s(\frac{1}{10})}{10} = \frac{1}{89} = .011235\ldots$.[33] Similarly, m=2 gives $\frac{s(\frac{1}{100})}{100} = \frac{1}{9899} = .000101020305081321 34\ldots$.

9.10 Reciprocal sums

Infinite sums over reciprocal Fibonacci numbers can sometimes be evaluated in terms of theta functions. For example, we can write the sum of every odd-indexed reciprocal Fibonacci number as

$$\sum_{k=0}^{\infty} \frac{1}{F_{2k+1}} = \frac{\sqrt{5}}{4} \vartheta_2^2 \left(0, \frac{3-\sqrt{5}}{2}\right),$$

and the sum of squared reciprocal Fibonacci numbers as

$$\sum_{k=1}^{\infty} \frac{1}{F_k^2} = \frac{5}{24} \left(\vartheta_2^4 \left(0, \frac{3-\sqrt{5}}{2}\right) - \vartheta_4^4 \left(0, \frac{3-\sqrt{5}}{2}\right) + 1\right).$$

If we add 1 to each Fibonacci number in the first sum, there is also the closed form

$$\sum_{k=0}^{\infty} \frac{1}{1+F_{2k+1}} = \frac{\sqrt{5}}{2},$$

and there is a *nested* sum of squared Fibonacci numbers giving the reciprocal of the golden ratio,

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{\sum_{j=1}^{k} {F_j}^2} = \frac{\sqrt{5}-1}{2}.$$

No closed formula for the reciprocal Fibonacci constant

$$\psi = \sum_{k=1}^{\infty} \frac{1}{F_k} = 3.359885666243\ldots$$

is known, but the number has been proved irrational by Richard André-Jeannin.[34]

The **Millin series** gives the identity[35]

$$\sum_{n=0}^{\infty} \frac{1}{F_{2^n}} = \frac{7-\sqrt{5}}{2},$$

which follows from the closed form for its partial sums as N tends to infinity:

$$\sum_{n=0}^{N} \frac{1}{F_{2^n}} = 3 - \frac{F_{2^N-1}}{F_{2^N}}.$$

9.11 Primes and divisibility

9.11.1 Divisibility properties

Every 3rd number of the sequence is even and more generally, every kth number of the sequence is a multiple of Fk. Thus the Fibonacci sequence is an example of a divisibility sequence. In fact, the Fibonacci sequence satisfies the stronger divisibility property[36][37]

$$\gcd(F_m, F_n) = F_{\gcd(m,n)}.$$

Any three consecutive Fibonacci numbers are pairwise coprime, which means that, for every n,

$$\gcd(Fn, Fn_{+1}) = \gcd(Fn, Fn_{+2}) = \gcd(Fn_{+1}, Fn_{+2}) = 1.$$

Every prime number p divides a Fibonacci number that can be determined by the value of p modulo 5. If p is congruent to 1 or 4 (mod 5), then p divides Fp_{-1}, and if p is congruent to 2 or 3 (mod 5), then, p divides Fp_{+1}. The remaining case is that $p = 5$, and in this case p divides F_p. These cases can be combined into a single formula, using the Legendre symbol:[38]

$$p \mid F_{p-\left(\frac{5}{p}\right)}.$$

9.11.2 Primality testing

The above formula can be used as a primality test in the sense that if

$n \mid F_{n-\left(\frac{5}{n}\right)}$, where the Legendre symbol has been replaced by the Jacobi symbol, then this is evidence that n is a prime, and if it fails to hold, then n is definitely not a prime. If n is composite and satisfies the formula, then n is a *Fibonacci pseudoprime*.

When m is large—say a 500-bit number—then we can calculate F_m (mod n) efficiently using the matrix form. Thus

$$\begin{pmatrix} F_{m+1} & F_m \\ F_m & F_{m-1} \end{pmatrix} \equiv \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^m \pmod{n}.$$

Here the matrix power A^m is calculated using Modular exponentiation, which can be adapted to matrices--modular exponentiation for matrices[39]

9.11.3 Fibonacci primes

Main article: Fibonacci prime

A *Fibonacci prime* is a Fibonacci number that is prime. The first few are:

2, 3, 5, 13, 89, 233, 1597, 28657, 514229, … OEIS A005478.

Fibonacci primes with thousands of digits have been found, but it is not known whether there are infinitely many.[40]

Fkn is divisible by *Fn*, so, apart from $F_4 = 3$, any Fibonacci prime must have a prime index. As there are arbitrarily long runs of composite numbers, there are therefore also arbitrarily long runs of composite Fibonacci numbers.

No Fibonacci number greater than $F_6 = 8$ is one greater or one less than a prime number.[41]

The only nontrivial square Fibonacci number is 144.[42] Attila Pethő proved in 2001 that there is only a finite number of perfect power Fibonacci numbers.[43] In 2006, Y. Bugeaud, M. Mignotte, and S. Siksek proved that 8 and 144 are the only such non-trivial perfect powers.[44]

9.11.4 Prime divisors of Fibonacci numbers

With the exceptions of 1, 8 and 144 ($F_1 = F_2$, F_6 and F_{12}) every Fibonacci number has a prime factor that is not a factor of any smaller Fibonacci number (Carmichael's theorem).[45] As a result, 8 and 144 (F_6 and F_{12}) are the only Fibonacci numbers that are the product of other Fibonacci numbers OEIS A235383.

The divisibility of Fibonacci numbers by a prime p is related to the Legendre symbol $\left(\frac{p}{5}\right)$ which is evaluated as follows:

$$\left(\frac{p}{5}\right) = \begin{cases} 0 & \text{if } p = 5 \\ 1 & \text{if } p \equiv \pm 1 \pmod 5 \\ -1 & \text{if } p \equiv \pm 2 \pmod 5. \end{cases}$$

If p is a prime number then

$$F_p \equiv \left(\tfrac{p}{5}\right) \pmod p \quad \text{and} \quad F_{p-\left(\frac{p}{5}\right)} \equiv 0 \pmod p.$$ [46][47]

For example,

$$\begin{aligned}
(\tfrac{2}{5}) &= -1, \quad F_3 = 2, \quad F_2 = 1,\\
(\tfrac{3}{5}) &= -1, \quad F_4 = 3, \quad F_3 = 2,\\
(\tfrac{5}{5}) &= 0, \quad F_5 = 5,\\
(\tfrac{7}{5}) &= -1, \quad F_8 = 21, \quad F_7 = 13,\\
(\tfrac{11}{5}) &= +1, \quad F_{10} = 55, \quad F_{11} = 89.
\end{aligned}$$

It is not known whether there exists a prime p such that

$$F_{p-(\frac{p}{5})} \equiv 0 \pmod{p^2}.$$

Such primes (if there are any) would be called Wall–Sun–Sun primes.

Also, if $p \neq 5$ is an odd prime number then:[47]

$$5F_{\frac{p\pm1}{2}}^2 \equiv \begin{cases} \frac{1}{2}\left(5\left(\frac{p}{5}\right) \pm 5\right) \pmod{p} & \text{if } p \equiv 1 \pmod 4 \\ \frac{1}{2}\left(5\left(\frac{p}{5}\right) \mp 3\right) \pmod{p} & \text{if } p \equiv 3 \pmod 4. \end{cases}$$

Example 1. $p = 7$, in this case $p \equiv 3 \pmod 4$ and we have:

$$(\tfrac{7}{5}) = -1: \qquad \tfrac{1}{2}\left(5(\tfrac{7}{5}) + 3\right) = -1, \quad \tfrac{1}{2}\left(5(\tfrac{7}{5}) - 3\right) = -4.$$

$F_3 = 2$ and $F_4 = 3$.

$5F_3^2 = 20 \equiv -1 \pmod 7$ and $5F_4^2 = 45 \equiv -4 \pmod 7$

Example 2. $p = 11$, in this case $p \equiv 3 \pmod 4$ and we have:

$$(\tfrac{11}{5}) = +1: \qquad \tfrac{1}{2}\left(5(\tfrac{11}{5}) + 3\right) = 4, \quad \tfrac{1}{2}\left(5(\tfrac{11}{5}) - 3\right) = 1.$$

$F_5 = 5$ and $F_6 = 8$.

$5F_5^2 = 125 \equiv 4 \pmod{11}$ and $5F_6^2 = 320 \equiv 1 \pmod{11}$

Example 3. $p = 13$, in this case $p \equiv 1 \pmod 4$ and we have:

$$(\tfrac{13}{5}) = -1: \qquad \tfrac{1}{2}\left(5(\tfrac{13}{5}) - 5\right) = -5, \quad \tfrac{1}{2}\left(5(\tfrac{13}{5}) + 5\right) = 0.$$

$F_6 = 8$ and $F_7 = 13$.

$5F_6^2 = 320 \equiv -5 \pmod{13}$ and $5F_7^2 = 845 \equiv 0 \pmod{13}$

Example 4. $p = 29$, in this case $p \equiv 1 \pmod 4$ and we have:

$$(\tfrac{29}{5}) = +1: \qquad \tfrac{1}{2}\left(5(\tfrac{29}{5}) - 5\right) = 0, \quad \tfrac{1}{2}\left(5(\tfrac{29}{5}) + 5\right) = 5.$$

$F_{14} = 377$ and $F_{15} = 610$.

$5F_{14}^2 = 710645 \equiv 0 \pmod{29}$ and $5F_{15}^2 = 1860500 \equiv 5 \pmod{29}$

For odd n, all odd prime divisors of Fn are congruent to 1 modulo 4, implying that all odd divisors of Fn (as the products of odd prime divisors) are congruent to 1 modulo 4.[48]

For example,

$$F_1 = 1, F_3 = 2, F_5 = 5, F_7 = 13, F_9 = 34 = 2 \cdot 17, F_{11} = 89, F_{13} = 233, F_{15} = 610 = 2 \cdot 5 \cdot 61.$$

All known factors of Fibonacci numbers $F(i)$ for all $i < 50000$ are collected at the relevant repositories.[49][50]

9.11.5 Periodicity modulo *n*

Main article: Pisano period

If the members of the Fibonacci sequence are taken mod *n*, the resulting sequence is periodic with period at most *6n*.[51] The lengths of the periods for various *n* form the so-called Pisano periods OEIS A001175. Determining a general formula for the Pisano periods is an open problem, which includes as a subproblem a special instance of the problem of finding the multiplicative order of a modular integer or of an element in a finite field. However, for any particular *n*, the Pisano period may be found as an instance of cycle detection.

9.12 Right triangles

Starting with 5, every second Fibonacci number is the length of the hypotenuse of a right triangle with integer sides, or in other words, the largest number in a Pythagorean triple. The length of the longer leg of this triangle is equal to the sum of the three sides of the preceding triangle in this series of triangles, and the shorter leg is equal to the difference between the preceding bypassed Fibonacci number and the shorter leg of the preceding triangle.

The first triangle in this series has sides of length 5, 4, and 3. Skipping 8, the next triangle has sides of length 13, 12 (5 + 4 + 3), and 5 (8 − 3). Skipping 21, the next triangle has sides of length 34, 30 (13 + 12 + 5), and 16 (21 − 5). This series continues indefinitely. The triangle sides *a*, *b*, *c* can be calculated directly:

$$a_n = F_{2n-1}$$
$$b_n = 2F_nF_{n-1}$$
$$c_n = F_n^2 - F_{n-1}^2.$$

These formulas satisfy $a_n^2 = b_n^2 + c_n^2$ for all *n*, but they only represent triangle sides when $n > 2$.

Any four consecutive Fibonacci numbers *Fn*, *Fn*+1, *Fn*+2 and *Fn*+3 can also be used to generate a Pythagorean triple in a different way:[52]

$$a = F_nF_{n+3}\,;\; b = 2F_{n+1}F_{n+2}\,;\; c = F_{n+1}^2 + F_{n+2}^2\,;\; a^2 + b^2 = c^2\,.$$

Example 1: let the Fibonacci numbers be 1, 2, 3 and 5. Then:

$$a = 1 \times 5 = 5$$
$$b = 2 \times 2 \times 3 = 12$$
$$c = 2^2 + 3^2 = 13$$
$$5^2 + 12^2 = 13^2$$

9.13 Magnitude

Since *Fn* is asymptotic to $\varphi^n/\sqrt{5}$, the number of digits in *Fn* is asymptotic to $n \log_{10} \varphi \approx 0.2090\, n$. As a consequence, for every integer $d > 1$ there are either 4 or 5 Fibonacci numbers with *d* decimal digits.

More generally, in the base *b* representation, the number of digits in *Fn* is asymptotic to $n \log_b \varphi$.

9.14 Applications

The Fibonacci numbers are important in the computational run-time analysis of Euclid's algorithm to determine the greatest common divisor of two integers: the worst case input for this algorithm is a pair of consecutive Fibonacci numbers.[53]

Brasch et al. 2012 show how a generalised Fibonacci sequence also can be connected to the field of economics.[54] In particular, it is shown how a generalised Fibonacci sequence enters the control function of finite-horizon dynamic optimisation problems with one state and one control variable. The procedure is illustrated in an example often referred to as the Brock–Mirman economic growth model.

Yuri Matiyasevich was able to show that the Fibonacci numbers can be defined by a Diophantine equation, which led to his original solution of Hilbert's tenth problem.[55]

The Fibonacci numbers are also an example of a complete sequence. This means that every positive integer can be written as a sum of Fibonacci numbers, where any one number is used once at most.

Moreover, every positive integer can be written in a unique way as the sum of *one or more* distinct Fibonacci numbers in such a way that the sum does not include any two consecutive Fibonacci numbers. This is known as Zeckendorf's theorem, and a sum of Fibonacci numbers that satisfies these conditions is called a Zeckendorf representation. The Zeckendorf representation of a number can be used to derive its Fibonacci coding.

Fibonacci numbers are used by some pseudorandom number generators.

They are also used in planning poker, which is a step in estimating in software development projects that use the Scrum (software development) methodology.

Fibonacci numbers are used in a polyphase version of the merge sort algorithm in which an unsorted list is divided into two lists whose lengths correspond to sequential Fibonacci numbers – by dividing the list so that the two parts have lengths in the approximate proportion φ. A tape-drive implementation of the polyphase merge sort was described in *The Art of Computer Programming*.

Fibonacci numbers arise in the analysis of the Fibonacci heap data structure.

The Fibonacci cube is an undirected graph with a Fibonacci number of nodes that has been proposed as a network topology for parallel computing.

A one-dimensional optimization method, called the Fibonacci search technique, uses Fibonacci numbers.[56]

The Fibonacci number series is used for optional lossy compression in the IFF 8SVX audio file format used on Amiga computers. The number series compands the original audio wave similar to logarithmic methods such as μ-law.[57][58]

Since the conversion factor 1.609344 for miles to kilometers is close to the golden ratio (denoted φ), the decomposition of distance in miles into a sum of Fibonacci numbers becomes nearly the kilometer sum when the Fibonacci numbers are replaced by their successors. This method amounts to a radix 2 number register in golden ratio base φ being shifted. To convert from kilometers to miles, shift the register down the Fibonacci sequence instead.[59]

9.15 In nature

Further information: Patterns in nature and Phyllotaxis

Fibonacci sequences appear in biological settings,[10] in two consecutive Fibonacci numbers, such as branching in trees, arrangement of leaves on a stem, the fruitlets of a pineapple,[11] the flowering of artichoke, an uncurling fern and the arrangement of a pine cone,[12] and the family tree of honeybees.[60] However, numerous poorly substantiated claims of Fibonacci numbers or golden sections in nature are found in popular sources, e.g., relating to the breeding of rabbits in Fibonacci's own unrealistic example, the seeds on a sunflower, the spirals of shells, and the curve of waves.[61]

Przemysław Prusinkiewicz advanced the idea that real instances can in part be understood as the expression of certain algebraic constraints on free groups, specifically as certain Lindenmayer grammars.[62]

A model for the pattern of florets in the head of a sunflower was proposed by H. Vogel in 1979.[63] This has the form

$$\theta = \frac{2\pi}{\phi^2} n,\ r = c\sqrt{n}$$

where n is the index number of the floret and c is a constant scaling factor; the florets thus lie on Fermat's spiral. The divergence angle, approximately 137.51°, is the golden angle, dividing the circle in the golden ratio. Because this ratio is irrational, no floret has a neighbor at exactly the same angle from the center, so the florets pack efficiently. Because the rational approximations to the golden ratio are of the form $F(j)$:$F(j + 1)$, the nearest neighbors of floret number n are those at $n \pm F(j)$ for some index j, which depends on r, the distance from the center. It is often said

that sunflowers and similar arrangements have 55 spirals in one direction and 89 in the other (or some other pair of adjacent Fibonacci numbers), but this is true only of one range of radii, typically the outermost and thus most conspicuous.[64]

9.15.1 The bee ancestry code

Fibonacci numbers also appear in the pedigrees of idealized honeybees, according to the following rules:

- If an egg is laid by an unmated female, it hatches a male or drone bee.
- If, however, an egg was fertilized by a male, it hatches a female.

Thus, a male bee always has one parent, and a female bee has two.

If one traces the pedigree of any male bee (1 bee), he has 1 parent (1 bee), 2 grandparents, 3 great-grandparents, 5 great-great-grandparents, and so on. This sequence of numbers of parents is the Fibonacci sequence. The number of ancestors at each level, Fn, is the number of female ancestors, which is Fn_{-1}, plus the number of male ancestors, which is Fn_{-2}.[65] This is under the unrealistic assumption that the ancestors at each level are otherwise unrelated.

9.15.2 The human X chromosome inheritance tree

Luke Hutchison noticed that number of possible ancestors on the X chromosome inheritance line at a given ancestral generation also follows the Fibonacci sequence.[66] A male individual has an X chromosome, which he received from his mother, and a Y chromosome, which he received from his father. The male counts as the "origin" of his own X chromosome ($F_1 = 1$), and at his parents' generation, his X chromosome came from a single parent ($F_2 = 1$). The male's mother received one X chromosome from her mother (the son's maternal grandmother), and one her father (the son's maternal grandfather), so two grandparents contributed to the male descendant's X chromosome ($F_3 = 2$). The maternal grandfather received his X chromosome from his mother, and the maternal grandmother received X chromosomes from both of her parents, so three great-grandparents contributed to the male descendant's X chromosome ($F_4 = 3$). Five great-great-grandparents contributed to the male descendant's X chromosome ($F_5 = 5$), etc. (Note that this assumes that all ancestors of a given descendant are independent, but if any genealogy is traced far enough back in time, ancestors begin to appear on multiple lines of the genealogy, until eventually a population founder appears on all lines of the genealogy.)

9.16 Generalizations

Main article: Generalizations of Fibonacci numbers

The Fibonacci sequence has been generalized in many ways. These include:

- Generalizing the index to negative integers to produce the negafibonacci numbers.
- Generalizing the index to real numbers using a modification of Binet's formula.[31]
- Starting with other integers. Lucas numbers have $L_1 = 1$, $L_2 = 3$, and $Ln = Ln_{-1} + Ln_{-2}$. Primefree sequences use the Fibonacci recursion with other starting points to generate sequences in which all numbers are composite.
- Letting a number be a linear function (other than the sum) of the 2 preceding numbers. The Pell numbers have $Pn = 2Pn_{-1} + Pn_{-2}$.
- Not adding the immediately preceding numbers. The Padovan sequence and Perrin numbers have $P(n) = P(n - 2) + P(n - 3)$.
- Generating the next number by adding 3 numbers (tribonacci numbers), 4 numbers (tetranacci numbers), or more. The resulting sequences are known as *n-Step Fibonacci numbers*.[67]
- Adding other objects than integers, for example functions or strings – one essential example is Fibonacci polynomials.

9.17 See also

- Elliott wave principle
- Fibonacci word
- Recursion (computer science)#Fibonacci
- The Fibonacci Association
- Verner Emil Hoggatt, Jr.
- Fibonacci numbers in popular culture
- Wythoff array

9.18 Notes

[1] Beck & Geoghegan 2010.

[2] Bóna 2011, p. 180.

[3] OEIS A000045

[4] John Hudson Tiner (200). *Exploring the World of Mathematics: From Ancient Record Keeping to the Latest Advances in Computers*. New Leaf Publishing Group. ISBN 978-1-61458-155-0.

[5] Lucas 1891, p. 3.

[6] Pisano 2002, pp. 404–5.

[7] Goonatilake, Susantha (1998), *Toward a Global Science*, Indiana University Press, p. 126, ISBN 978-0-253-33388-9

[8] Singh, Parmanand (1985), "The So-called Fibonacci numbers in ancient and medieval India", *Historia Mathematica*, **12** (3): 229–44, doi:10.1016/0315-0860(85)90021-7

[9] Knuth, Donald (2006), *The Art of Computer Programming*, 4. Generating All Trees – History of Combinatorial Generation, Addison–Wesley, p. 50, ISBN 978-0-321-33570-8, it was natural to consider the set of all sequences of [L] and [S] that have exactly m beats. ...there are exactly Fm+1 of them. For example the 21 sequences when $m = 7$ are: [gives list]. In this way Indian prosodists were led to discover the Fibonacci sequence, as we have observed in Section 1.2.8 (from v.1)

[10] Douady, S; Couder, Y (1996), "Phyllotaxis as a Dynamical Self Organizing Process" (PDF), *Journal of Theoretical Biology*, **178** (178): 255–74, doi:10.1006/jtbi.1996.0026

[11] Jones, Judy; Wilson, William (2006), "Science", *An Incomplete Education*, Ballantine Books, p. 544, ISBN 978-0-7394-7582-9

[12] Brousseau, A (1969), "Fibonacci Statistics in Conifers", *Fibonacci Quarterly* (7): 525–32

[13] Knuth, Donald (1968), *The Art of Computer Programming*, **1**, Addison Wesley, ISBN 81-7758-754-4, Before Fibonacci wrote his work, the sequence Fn had already been discussed by Indian scholars, who had long been interested in rhythmic patterns... both Gopala (before 1135 AD) and Hemachandra (c. 1150) mentioned the numbers 1,2,3,5,8,13,21 explicitly [see P. Singh Historia Math 12 (1985) 229–44]" p. 100 (3d ed)...

[14] Agrawala, VS (1969), Pāṇinikālīna Bhāratavarṣa *(Hn.). Varanasi-I: TheChowkhamba Vidyabhawan*, SadgurushiShya writes that Pingala was a younger brother of Pāṇini [Agrawala 1969, lb]. There is an alternative opinion that he was a maternal uncle of Pāṇini [Vinayasagar 1965, Preface, 121. … Agrawala [1969, 463–76], after a careful investigation, in which he considered the views of earlier scholars, has concluded that Pāṇini lived between 480 and 410 BC

[15] Velankar, HD (1962), *'Vṛttajātisamuccaya' of kavi Virahanka*, Jodhpur: Rajasthan Oriental Research Institute, p. 101, "For four, variations of meters of two [and] three being mixed, five happens. For five, variations of two earlier – three [and] four, being mixed, eight is obtained. In this way, for six, [variations] of four [and] of five being mixed, thirteen happens. And like that, variations of two earlier meters being mixed, seven morae [is] twenty-one. In this way, the process should be followed in all mātrā-vṛttas

[16] Knott, Ron. "Fibonacci's Rabbits". University of Surrey Faculty of Engineering and Physical Sciences.

[17] Gardner, Martin (1996), *Mathematical Circus*, The Mathematical Association of America, p. 153, ISBN 0-88385-506-2, It is ironic that Leonardo, who made valuable contributions to mathematics, is remembered today mainly because a 19th-century French number theorist, Édouard Lucas... attached the name Fibonacci to a number sequence that appears in a trivial problem in Liber abaci

[18] Knott, R, "Fib table", *Fibonacci*, UK: Surrey has the first 300 F*n* factored into primes and links to more extensive tables.

[19] Knuth, Donald (2008-12-11), "Negafibonacci Numbers and the Hyperbolic Plane", *Annual meeting*, The Fairmont Hotel, San Jose, CA: The Mathematical Association of America

[20] Lucas 1891, p. 7.

[21] Stanley, Richard (2011). *Enumerative Combinatorics I (2nd ed.)*. Cambridge Univ. Press. p. "121, Ex 1.35". ISBN 978-1-107-60262-5.

[22] Weisstein, Eric W. "Binet's Fibonacci Number Formula". *MathWorld*.

[23] Ball 2003, p. 156.

[24] Ball 2003, pp. 155–6.

[25] Kepler, Johannes (1966), *A New Year Gift: On Hexagonal Snow*, Oxford University Press, p. 92, ISBN 0-19-858120-3

[26] *Strena seu de Nive Sexangula*, 1611

[27] Dijkstra, Edsger W. (1978), *In honour of Fibonacci* (PDF)

[28] Gessel, Ira (October 1972), "Fibonacci is a Square" (PDF), *The Fibonacci Quarterly*, **10** (4): 417–19, retrieved April 11, 2012

[29] Lucas 1891, p. 4.

[30] Vorobiev, Nikolaĭ Nikolaevich; Martin, Mircea (2002), "Chapter 1", *Fibonacci Numbers*, Birkhäuser, pp. 5–6, ISBN 3-7643-6135-2

[31] Weisstein, Eric W. "Fibonacci Number". *MathWorld*.

[32] Glaister, P (1995), "Fibonacci power series", *The Mathematical Gazette*, **79** (486): 521, doi:10.2307/3618079, JSTOR 3618079

[33] Köhler, Günter (February 1985), "Generating functions of Fibonacci-like sequences and decimal expansions of some fractions" (PDF), *The Fibonacci Quarterly*, **23** (1): 29–35, retrieved December 31, 2011

[34] André-Jeannin, Richard (1989), "Irrationalité de la somme des inverses de certaines suites récurrentes", *C. R. Acad. Sci. Paris Sér. I Math.*, **308** (19): 539–541, MR 0999451

[35] Weisstein, Eric W. "Millin Series". *MathWorld*.

[36] Ribenboim, Paulo (2000), *My Numbers, My Friends*, Springer-Verlag

[37] Su, Francis E (2000), "Fibonacci GCD's, please", *Mudd Math Fun Facts*, et al, HMC

[38] Williams, H. C. (1982), "A note on the Fibonacci quotient $F_{p-\varepsilon}/p$ ", *Canadian Mathematical Bulletin*, **25** (3): 366–370, doi:10.4153/CMB-1982-053-0, MR 668957. Williams calls this property "well known".

[39] *Prime Numbers*, Richard Crandall, Carl Pomerance, Springer, second edition, 2005, p.142.

[40] Weisstein, Eric W. "Fibonacci Prime". *MathWorld*.

[41] Honsberger, Ross (1985), "Mathematical Gems III", *AMS Dolciani Mathematical Expositions* (9): 133, ISBN 0-88385-318-3

[42] Cohn, JHE (1964), "Square Fibonacci Numbers etc", *Fibonacci Quarterly*, **2**: 109–13

[43] Pethő, Attila (2001), "Diophantine properties of linear recursive sequences II", *Acta Math. Paedagogicae Nyíregyháziensis*, **17**: 81–96

[44] Bugeaud, Y; Mignotte, M; Siksek, S (2006), "Classical and modular approaches to exponential Diophantine equations. I. Fibonacci and Lucas perfect powers", *Ann. Math.*, **2** (163): 969–1018, arXiv:math/0403046, Bibcode:2004math.....3046B, doi:10.4007/annals.2006.163.969

[45] Knott, Ron, *The Fibonacci numbers*, UK: Surrey

[46] Ribenboim, Paulo (1996), *The New Book of Prime Number Records*, New York: Springer, p. 64, ISBN 0-387-94457-5

[47] Lemmermeyer 2000, pp. 73–4.

[48] Lemmermeyer 2000, p. 73.

[49] *Fibonacci and Lucas factorizations*, Mersennus collects all known factors of $F(i)$ with $i < 10000$.

[50] *Factors of Fibonacci and Lucas numbers*, Red golpe collects all known factors of $F(i)$ with $10000 < i < 50000$.

[51] Freyd, Peter; Brown, Kevin S. (1993), "Problems and Solutions: Solutions: E3410", *The American Mathematical Monthly*, **99** (3): 278–279, doi:10.2307/2325076

[52] Koshy, Thomas (2007), *Elementary number theory with applications*, Academic Press, p. 581, ISBN 0-12-372487-2

[53] Knuth, Donald E (1997), *The Art of Computer Programming*, 1: Fundamental Algorithms (3rd ed.), Addison–Wesley, p. 343, ISBN 0-201-89683-4

[54] Brasch, T. von; Byström, J.; Lystad, L.P. (2012), "Optimal Control and the Fibonacci Sequence" (PDF), *Journal of Optimization Theory and Applications*, **154** (3): 857–78, doi:10.1007/s10957-012-0061-2

[55] Harizanov, Valentina (1995), "Review of Yuri V. Matiyasevich, *Hibert's Tenth Problem*", *Modern Logic*, **5** (3): 345–355.

[56] Avriel, M; Wilde, DJ (1966), "Optimality of the Symmetric Fibonacci Search Technique", *Fibonacci Quarterly* (3): 265–9

[57] *Amiga ROM Kernel Reference Manual*, Addison–Wesley, 1991

[58] "IFF", *Multimedia Wiki*

[59] "Zeckendorf representation", *Encyclopedia of Math*

[60] "Marks for the da Vinci Code: B–". *Maths.* Computer Science For Fun: CS4FN.

[61] Simanek, D. "Fibonacci Flim-Flam". LHUP.

[62] Prusinkiewicz, Przemyslaw; Hanan, James (1989), *Lindenmayer Systems, Fractals, and Plants (Lecture Notes in Biomathematics)*, Springer-Verlag, ISBN 0-387-97092-4

[63] Vogel, H (1979), "A better way to construct the sunflower head", *Mathematical Biosciences*, **44** (44): 179–89, doi:10.1016/0025-5564(79)90080-4

[64] Prusinkiewicz, Przemyslaw; Lindenmayer, Aristid (1990), *The Algorithmic Beauty of Plants*, Springer-Verlag, pp. 101–7, ISBN 978-0-387-97297-8

[65] "The Fibonacci sequence as it appears in nature" (PDF), *The Fibonacci Quarterly*, **1** (1): 53–56, 1963

[66] Hutchison, Luke (September 2004). "Growing the Family Tree: The Power of DNA in Reconstructing Family Relationships" (PDF). *Proceedings of the First Symposium on Bioinformatics and Biotechnology (BIOT-04)*. Retrieved 2016-09-03.

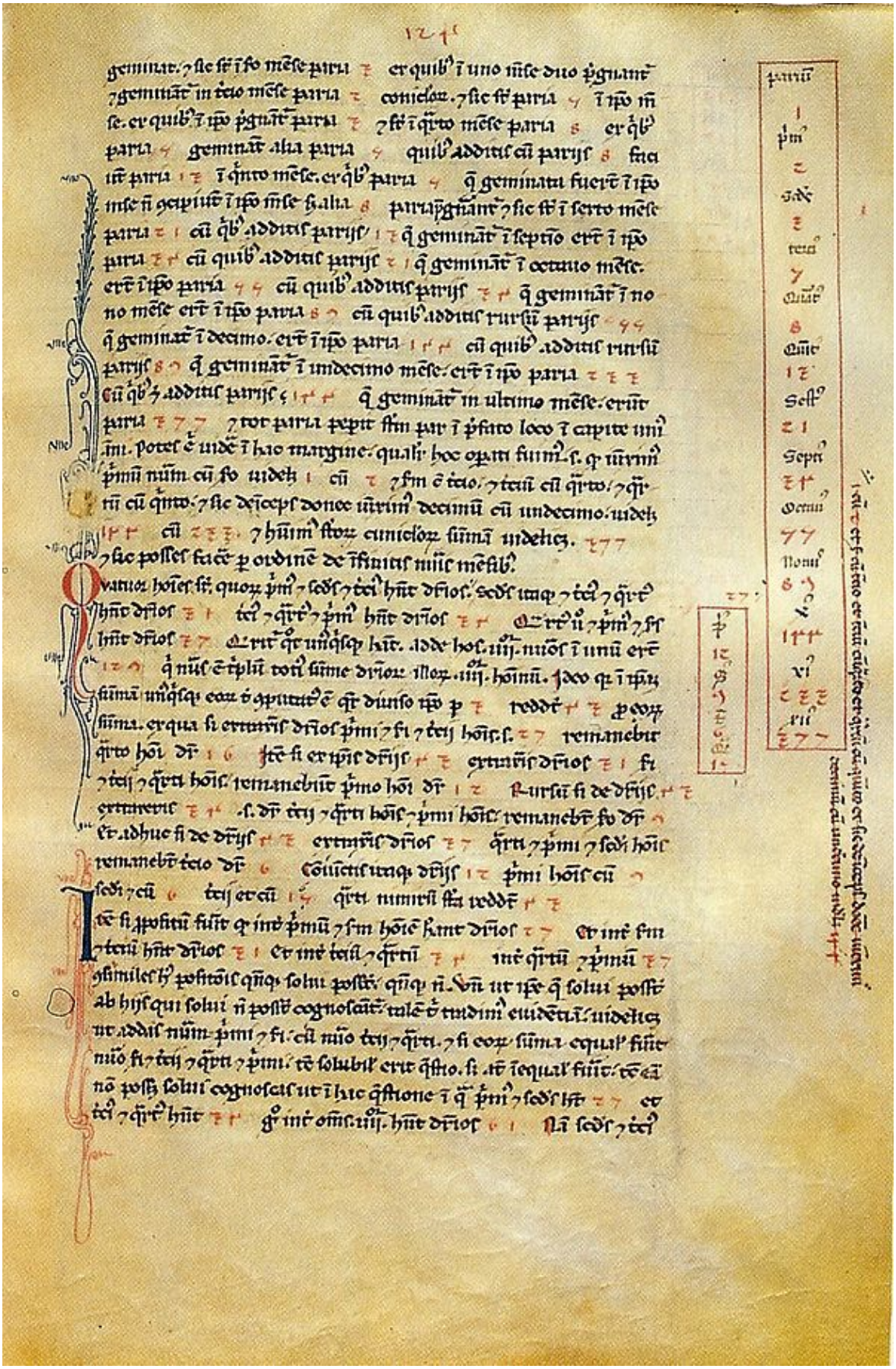
[67] Weisstein, Eric W. "Fibonacci *n*-Step Number". *MathWorld*.

9.19 References

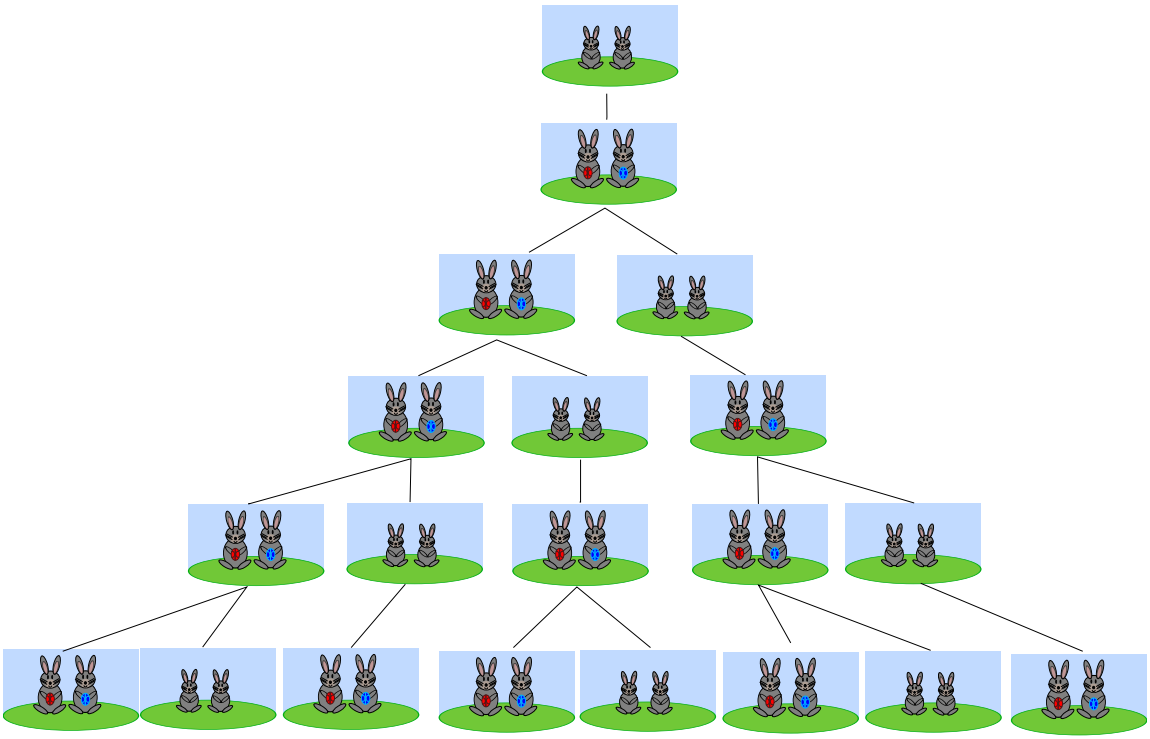
- Ball, Keith M (2003), "8: Fibonacci's Rabbits Revisited", *Strange Curves, Counting Rabbits, and Other Mathematical Explorations*, Princeton, NJ: Princeton University Press, ISBN 0-691-11321-1.
- Beck, Matthias; Geoghegan, Ross (2010), *The Art of Proof: Basic Training for Deeper Mathematics*, New York: Springer.
- Bóna, Miklós (2011), *A Walk Through Combinatorics* (3rd ed.), New Jersey: World Scientific.
- Lemmermeyer, Franz (2000), *Reciprocity Laws*, New York: Springer, ISBN 3-540-66957-4.
- Lucas, Édouard (1891), *Théorie des nombres* (in French), **1**, Gauthier-Villars.
- Pisano, Leonardo (2002), *Fibonacci's Liber Abaci: A Translation into Modern English of the Book of Calculation*, Sources and Studies in the History of Mathematics and Physical Sciences, Sigler, Laurence E, trans, Springer, ISBN 0-387-95419-8

9.20 External links

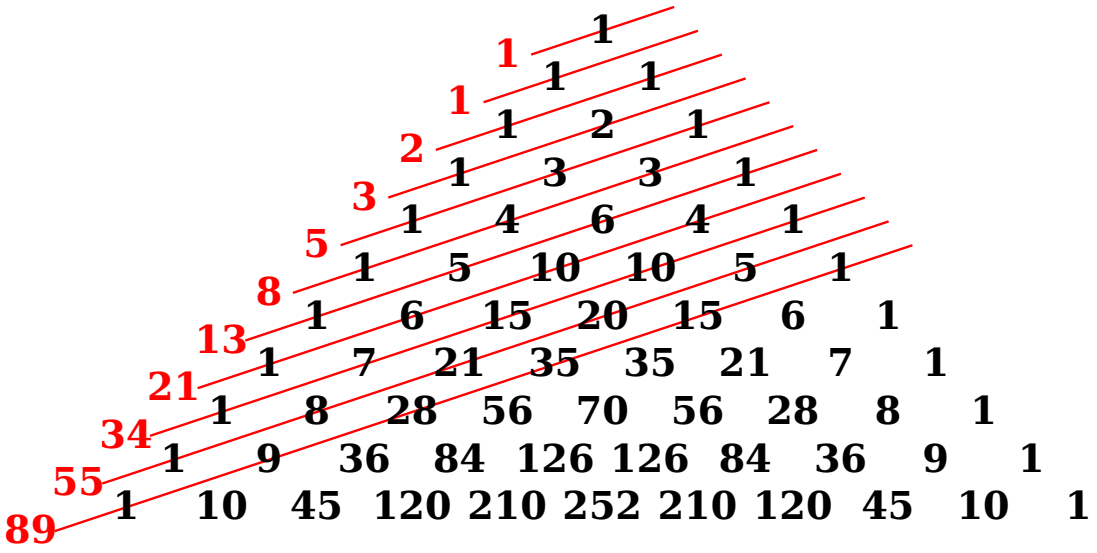
- Periods of Fibonacci Sequences Mod m at MathPages
- Scientists find clues to the formation of Fibonacci spirals in nature
-
- Fibonacci Sequence on *In Our Time* at the BBC. (listen now)
- Hazewinkel, Michiel, ed. (2001), "Fibonacci numbers", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- "Sloane's A000045 : Fibonacci Numbers". *The On-Line Encyclopedia of Integer Sequences*. OEIS Foundation.



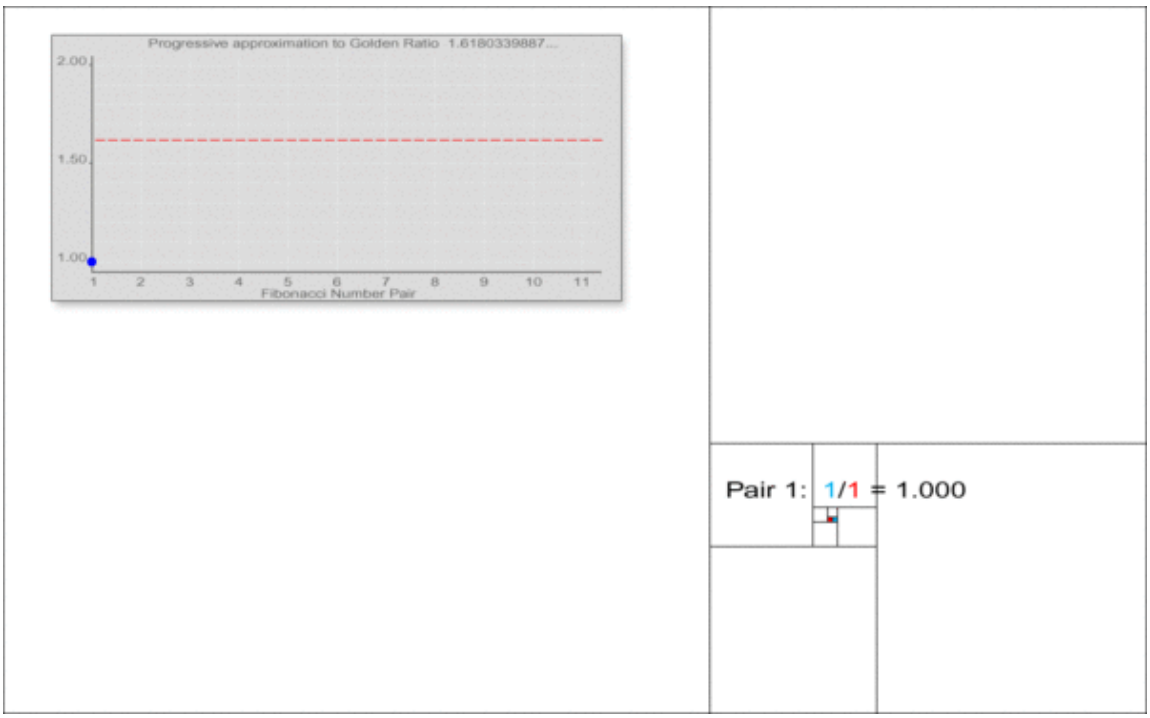
A page of Fibonacci's Liber Abaci *from the Biblioteca Nazionale di Firenze showing (in box on right) the Fibonacci sequence with the position in the sequence labeled in Latin and Roman numerals and the value in Hindu-Arabic numerals.*



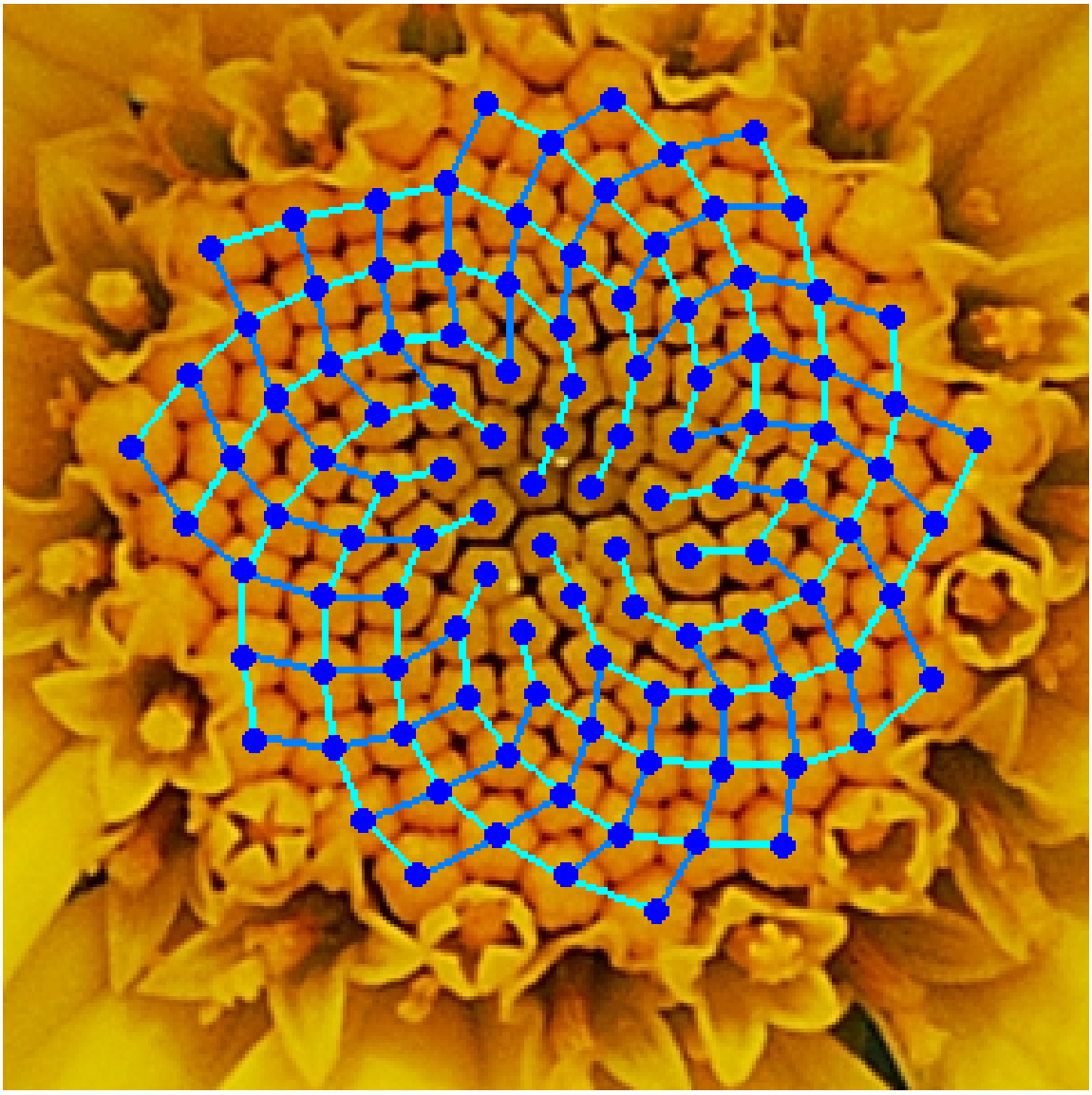
The number of rabbits form the Fibonacci sequence



The Fibonacci numbers are the sums of the "shallow" diagonals (shown in red) of Pascal's triangle.



Animated GIF file showing successive tilings of the plane, and a graph of approximations to the Golden Ratio calculated by dividing successive pairs of Fibonacci numbers, one by the other. Uses the Fibonacci numbers 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144



Yellow Chamomile head showing the arrangement in 21 (blue) and 13 (aqua) spirals. Such arrangements involving consecutive Fibonacci numbers appear in a wide variety of plants.

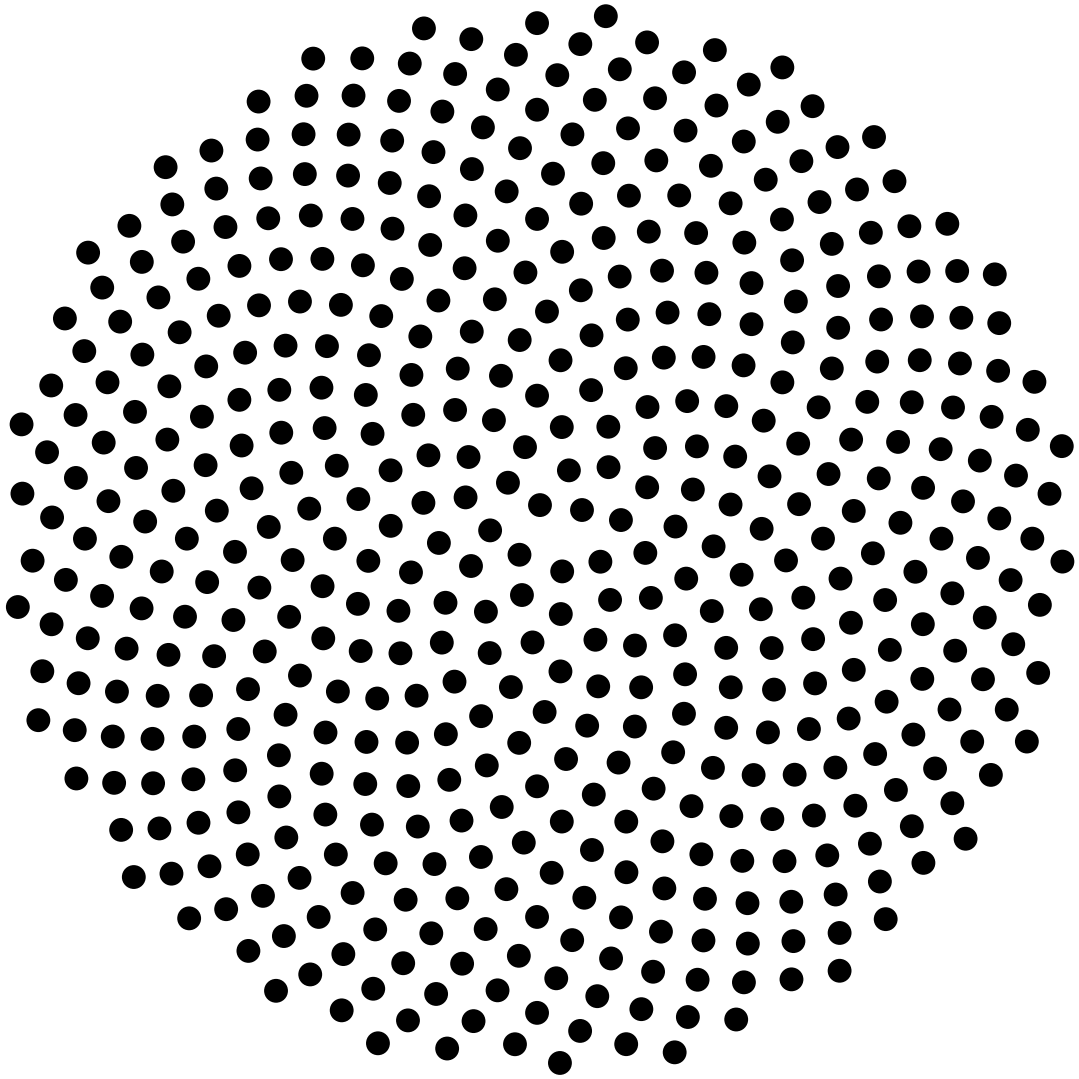
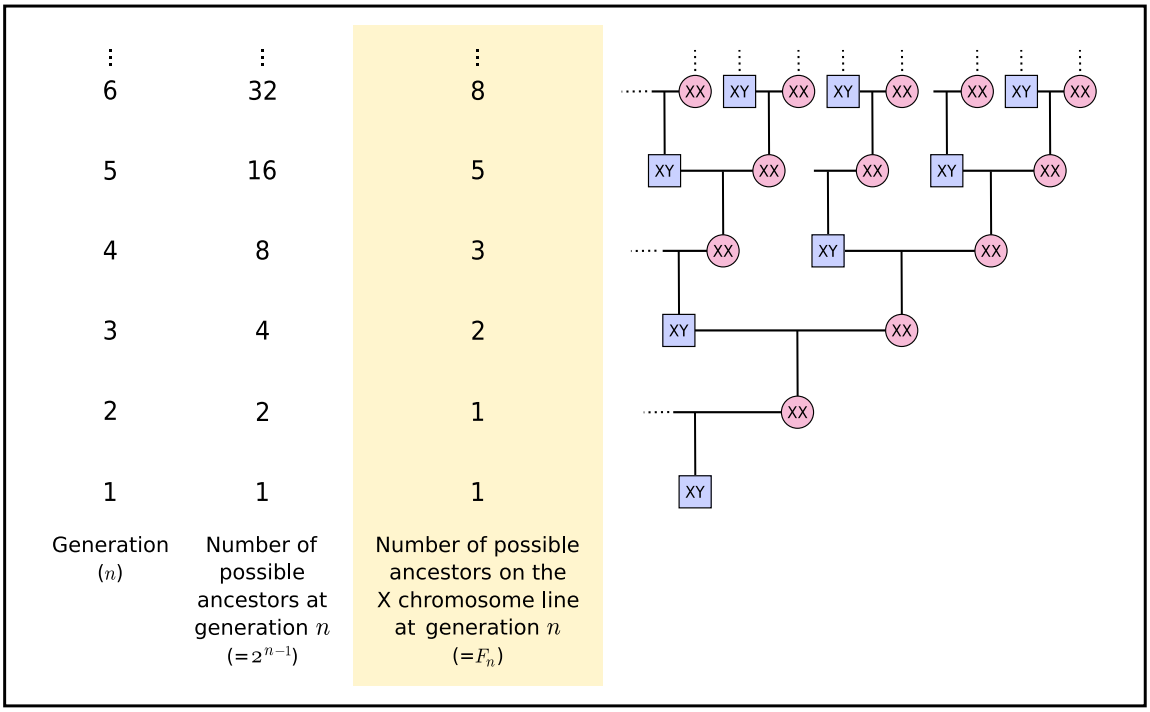


Illustration of Vogel's model for n = *1 … 500*



The number of possible ancestors on the X chromosome inheritance line at a given ancestral generation follows the Fibonacci sequence. (After Hutchison, L. "Growing the Family Tree: The Power of DNA in Reconstructing Family Relationships".[66])

Chapter 10

Catalan number

For names of numbers in Catalan, see List of numbers in various languages § Occitano-Romance.

In combinatorial mathematics, the **Catalan numbers** form a sequence of natural numbers that occur in various counting problems, often involving recursively-defined objects. They are named after the Belgian mathematician Eugène Charles Catalan (1814–1894).

Using zero-based numbering, the *n*th Catalan number is given directly in terms of binomial coefficients by

$$C_n = \frac{1}{n+1}\binom{2n}{n} = \frac{(2n)!}{(n+1)!\,n!} = \prod_{k=2}^{n} \frac{n+k}{k} \qquad \text{for } n \geq 0.$$

The first Catalan numbers for *n* = 0, 1, 2, 3, … are

> 1, 1, 2, 5, 14, 42, 132, 429, 1430, 4862, 16796, 58786, 208012, 742900, 2674440, 9694845, 35357670, 129644790, 477638700, 1767263190, 6564120420, 24466267020, 91482563640, 343059613650, 1289904147324, 4861946401452, … (sequence A000108 in the OEIS).

10.1 Properties

An alternative expression for *Cn* is

$$C_n = \binom{2n}{n} - \binom{2n}{n+1} = \frac{1}{n+1}\binom{2n}{n} \quad \text{for } n \geq 0,$$

which is equivalent to the expression given above because $\binom{2n}{n+1} = \frac{n}{n+1}\binom{2n}{n}$. This shows that *Cn* is an integer, which is not immediately obvious from the first formula given. This expression forms the basis for a proof of the correctness of the formula.

The Catalan numbers satisfy the recurrence relation

$$C_0 = 1 \quad \text{and} \quad C_{n+1} = \sum_{i=0}^{n} C_i\, C_{n-i} \quad \text{for} n \geq 0;$$

moreover,

$$C_n = \frac{1}{n+1}\sum_{i=0}^{n} \binom{n}{i}^2.$$

This is because $\binom{2n}{n} = \sum_{i=0}^{n} \binom{n}{i}^2$, since choosing *n* numbers from a 2*n* set of numbers can be uniquely divided into 2 parts: choosing *i* numbers out of the first *n* numbers and then choosing *n-i* numbers from the remaining *n* numbers.

They also satisfy:

$$C_0 = 1 \quad \text{and} \quad C_{n+1} = \frac{2(2n+1)}{n+2} C_n,$$

which can be a more efficient way to calculate them.

Asymptotically, the Catalan numbers grow as

$$C_n \sim \frac{4^n}{n^{3/2}\sqrt{\pi}}$$

in the sense that the quotient of the *n*th Catalan number and the expression on the right tends towards 1 as $n \to +\infty$. Some sources use just $C_n \approx \frac{4^n}{n^{3/2}}$.[1] (This can be proved by using Stirling's approximation for *n*!.)

The only Catalan numbers *Cn* that are odd are those for which $n = 2^k - 1$. All others are even.

The only prime Catalan numbers are *C2* = 2 and *C3* = 5.[2]

The Catalan numbers have an integral representation

$$C_n = \int_0^4 x^n \rho(x) dx$$

where $\rho(x) = \frac{1}{2\pi}\sqrt{\frac{4-x}{x}}$. This means that the Catalan numbers are a solution of the Hausdorff moment problem on the interval [0, 4] instead of [0, 1]. The orthogonal polynomials having the weight function $\rho(x)$ on $[0, 4]$ are

$$H_n(x) = \sum_{k=0}^{n} \binom{n+k}{n-k} (-x)^k.$$

10.2 Applications in combinatorics

There are many counting problems in combinatorics whose solution is given by the Catalan numbers. The book *Enumerative Combinatorics: Volume 2* by combinatorialist Richard P. Stanley contains a set of exercises which describe 66 different interpretations of the Catalan numbers. Following are some examples, with illustrations of the cases C_3 = 5 and C_4 = 14.

- *Cn* is the number of Dyck words[3] of length 2*n*. A Dyck word is a string consisting of *n* X's and *n* Y's such that no initial segment of the string has more Y's than X's. For example, the following are the Dyck words of length 6:

XXXYYY XYXXYY XYXYXY XXYYXY XXYXYY.

- Re-interpreting the symbol X as an open parenthesis and Y as a close parenthesis, *Cn* counts the number of expressions containing *n* pairs of parentheses which are correctly matched:

((())) ()(()) ()()() (())() (()())

- *Cn* is the number of different ways *n* + 1 factors can be completely parenthesized (or the number of ways of associating *n* applications of a binary operator). For *n* = 3, for example, we have the following five different parenthesizations of four factors:

((ab)c)d (a(bc))d (ab)(cd) a((bc)d) a(b(cd))

- Successive applications of a binary operator can be represented in terms of a full binary tree. (A rooted binary tree is *full* if every vertex has either two children or no children.) It follows that C_n is the number of full binary trees with $n + 1$ leaves:

- C_n is the number of non-isomorphic ordered trees with n vertices. (An ordered tree is a rooted tree in which the children of each vertex are given a fixed left-to-right order.)[4]

- C_n is the number of monotonic lattice paths along the edges of a grid with $n \times n$ square cells, which do not pass above the diagonal. A monotonic path is one which starts in the lower left corner, finishes in the upper right corner, and consists entirely of edges pointing rightwards or upwards. Counting such paths is equivalent to counting Dyck words: X stands for "move right" and Y stands for "move up".

The following diagrams show the case $n = 4$:

This can be succinctly represented by listing the Catalan elements by column height:[5]

[0,0,0,0][0,0,0,1][0,0,0,2][0,0,1,1]
[0,1,1,1] [0,0,1,2] [0,0,0,3] [0,1,1,2][0,0,2,2][0,0,1,3]
[0,0,2,3][0,1,1,3] [0,1,2,2][0,1,2,3]

- C_n is the number of different ways a convex polygon with $n + 2$ sides can be cut into triangles by connecting vertices with straight lines (a form of Polygon triangulation). The following hexagons illustrate the case $n = 4$:

- C_n is the number of stack-sortable permutations of $\{1, ..., n\}$. A permutation w is called stack-sortable if $S(w) = (1, ..., n)$, where $S(w)$ is defined recursively as follows: write $w = unv$ where n is the largest element in w and u and v are shorter sequences, and set $S(w) = S(u)S(v)n$, with S being the identity for one-element sequences. These are the permutations that avoid the pattern 231.

- C_n is the number of permutations of $\{1, ..., n\}$ that avoid the pattern 123 (or any of the other patterns of length 3); that is, the number of permutations with no three-term increasing subsequence. For $n = 3$, these permutations are 132, 213, 231, 312 and 321. For $n = 4$, they are 1432, 2143, 2413, 2431, 3142, 3214, 3241, 3412, 3421, 4132, 4213, 4231, 4312 and 4321.

- C_n is the number of noncrossing partitions of the set $\{1, ..., n\}$. *A fortiori*, C_n never exceeds the nth Bell number. C_n is also the number of noncrossing partitions of the set $\{1, ..., 2n\}$ in which every block is of size 2. The conjunction of these two facts may be used in a proof by mathematical induction that all of the *free* cumulants of degree more than 2 of the Wigner semicircle law are zero. This law is important in free probability theory and the theory of random matrices.

- C_n is the number of ways to tile a stairstep shape of height n with n rectangles. The following figure illustrates the case $n = 4$:

- C_n is the number of rooted binary trees with n internal nodes ($n + 1$ leaves or external nodes). Illustrated in following Figure are the trees corresponding to $n = 0,1,2$ and 3. There are 1, 1, 2, and 5 respectively. Here, we consider as binary trees those in which each node has zero or two children, and the internal nodes are those that have children.

- C_n is the number of ways to form a "mountain ranges" with n upstrokes and n down-strokes that all stay above the original line.The mountain range interpretation is that the mountains will never go below the horizon.

- C_n is the number of standard Young tableaux whose diagram is a 2-by-n rectangle. In other words, it is the number of ways the numbers 1, 2, ..., $2n$ can be arranged in a 2-by-n rectangle so that each row and each column is increasing. As such, the formula can be derived as a special case of the hook-length formula.

- C_n is the number of ways that the vertices of a convex $2n$-gon can be paired so that the line segments joining paired vertices do not intersect. This is precisely the condition that guarantees that the paired edges can be identified (sewn together) to form a closed surface of genus zero (a topological 2-sphere).

- *Cn* is the number of semiorders on *n* unlabeled items.[6]
- In chemical engineering *Cn* is the number of possible separation sequences which can separate a mixture of *n* components.[7]

10.3 Proof of the formula

There are several ways of explaining why the formula

$$C_n = \frac{1}{n+1}\binom{2n}{n}$$

solves the combinatorial problems listed above. The first proof below uses a generating function. The other proofs are examples of bijective proofs; they involve literally counting a collection of some kind of object to arrive at the correct formula.

10.3.1 First proof

We first observe that all of the combinatorial problems listed above satisfy Segner's[8] recurrence relation

$$C_0 = 1 \quad \text{and} \quad C_{n+1} = \sum_{i=0}^{n} C_i\, C_{n-i} \quad \text{for} n \geq 0.$$

For example, every Dyck word *w* of length ≥ 2 can be written in a unique way in the form

$$w = \mathrm{X}w_1\mathrm{Y}w_2$$

with (possibly empty) Dyck words w_1 and w_2.

The generating function for the Catalan numbers is defined by

$$c(x) = \sum_{n=0}^{\infty} C_n x^n.$$

The two recurrence relations together can then be summarized in generating function form by the relation

$$c(x) = 1 + xc(x)^2;$$

in other words, this equation follows from the recurrence relations by expanding both sides into power series. On the one hand, the recurrence relations uniquely determine the Catalan numbers; on the other hand, the generating function solution

$$c(x) = \frac{1-\sqrt{1-4x}}{2x} = \frac{2}{1+\sqrt{1-4x}}$$

has a power series at 0 and its coefficients must therefore be the Catalan numbers. The chosen solution satisfies the following condition.

$$\lim_{x\to 0^+} c(x) = C_0 = 1$$

The other solution has a pole at 0 and this reasoning doesn't apply to it.

The square root term can be expanded as a power series using the identity

$$\sqrt{1+y} = \sum_{n=0}^{\infty} \binom{\frac{1}{2}}{n} y^n = \sum_{n=0}^{\infty} \frac{(-1)^{n+1}}{4^n(2n-1)} \binom{2n}{n} y^n = 1 + \frac{1}{2}y - \frac{1}{8}y^2 + \cdots.$$

This is a special case of Newton's generalized binomial theorem; as with the general theorem, it can be proved by computing derivatives to produce its Taylor series. Setting $y = -4x$ and substituting this power series into the expression for $c(x)$ and shifting the summation index n by 1, the expansion simplifies to

$$c(x) = \sum_{n=0}^{\infty} \binom{2n}{n} \frac{x^n}{n+1}.$$

The coefficients are now the desired formula for C_n.

Another way to get $c(x)$ is to solve for $xc(x)$ and observe that $\int_0^x t^n\,dt$ appears in each term of the power series.

10.3.2 Second proof

This proof depends on a trick known as André's reflection method, which was originally used in connection with Bertrand's ballot theorem. (The reflection principle has been widely attributed to Désiré André, but his method did not actually use reflections; and the reflection method is a variation due to Aebly and Mirimanoff.[9]) We count the paths which start and end on the diagonal of the $n \times n$ grid. All such paths have n rightward and n upward steps. Since we can choose which of the $2n$ steps are upward (or, equivalently, rightward) ones, there are $\binom{2n}{n}$ total monotonic paths of this type. A *bad* path will cross the main diagonal and touch the next higher (*fatal*) diagonal (depicted red in the illustration). We flip the portion of the path occurring after that touch about that fatal diagonal, as illustrated; this geometric operation amounts to interchanging all the rightward and upward steps after that touch. In the section of the path that is not reflected, there is one more upward step than rightward steps, so the remaining section of the bad path has one more rightward than upward step (because it ends on the main diagonal). When this portion of the path is reflected, it will also have one more upward step than rightward steps. Since there are still $2n$ steps, there must now be $n + 1$ upward steps and $n - 1$ rightward steps. So, instead of reaching the target (n,n), all bad paths (after the portion of the path is reflected) will end in location $(n - 1, n + 1)$. As any monotonic path in the $n - 1 \times n + 1$ grid must meet the fatal diagonal, this reflection process sets up a bijection between the bad paths of the original grid and the monotonic paths of this new grid because the reflection process is reversible. The number of bad paths is therefore,

$$\binom{n-1+n+1}{n-1} = \binom{2n}{n-1} = \binom{2n}{n+1}$$

and the number of Catalan paths (i.e., good paths) is obtained by removing the number of bad paths from the total number of monotonic paths of the original grid,

$$C_n = \binom{2n}{n} - \binom{2n}{n+1}.$$

In terms of Dyck words, we start with a (non-Dyck) sequence of n X's and n Y's and interchange all X's and Y's after the first Y that violates the Dyck condition. At that first Y, there are $k + 1$ Y's and k X's for some k between 1 and $n - 1$.

10.3.3 Third proof

The following bijective proof, while being more involved than the previous one, provides a more natural explanation for the term $n + 1$ appearing in the denominator of the formula for C_n. A generalized version of this proof can be found in a paper of Rukavicka Josef (2011).[10]

Suppose we are given a monotonic path, which may happen to cross the diagonal. The **exceedance** of the path is defined to be the number of vertical edges which lie *above* the diagonal. For example, in Figure 2, the edges lying above the diagonal are marked in red, so the exceedance of the path is 5.

Now, if we are given a monotonic path whose exceedance is not zero, then we may apply the following algorithm to construct a new path whose exceedance is one less than the one we started with.

- Starting from the bottom left, follow the path until it first travels above the diagonal.
- Continue to follow the path until it *touches* the diagonal again. Denote by X the first such edge that is reached.
- Swap the portion of the path occurring before X with the portion occurring after X.

The following example should make this clearer. In Figure 3, the black dot indicates the point where the path first crosses the diagonal. The black edge is X, and we swap the red portion with the green portion to make a new path, shown in the second diagram.

Notice that the exceedance has dropped from three to two. In fact, the algorithm will cause the exceedance to decrease by one, for any path that we feed it, because the first vertical step starting on the diagonal (at the point marked with a black dot) is the unique vertical edge that under the operation passes from above the diagonal to below it; all other vertical edges stay on the same side of the diagonal.

It is also not difficult to see that this process is *reversible*: given any path P whose exceedance is less than n, there is exactly one path which yields P when the algorithm is applied to it. Indeed, the (black) edge X, which originally was the first horizontal step ending on the diagonal, has become the *last* horizontal step *starting* on the diagonal.

This implies that the number of paths of exceedance n is equal to the number of paths of exceedance $n - 1$, which is equal to the number of paths of exceedance $n - 2$, and so on, down to zero. In other words, we have split up the set of *all* monotonic paths into $n + 1$ equally sized classes, corresponding to the possible exceedances between 0 and n. Since there are

$$\binom{2n}{n}$$

monotonic paths, we obtain the desired formula

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

Figure 4 illustrates the situation for $n = 3$. Each of the 20 possible monotonic paths appears somewhere in the table. The first column shows all paths of exceedance three, which lie entirely above the diagonal. The columns to the right show the result of successive applications of the algorithm, with the exceedance decreasing one unit at a time. There are five rows, that is, $C_3 = 5$.

10.3.4 Fourth proof

This proof uses the triangulation definition of Catalan numbers to establish a relation between C_n and C_{n+1}. Given a polygon P with $n+2$ sides, first mark one of its sides as the base. If P is then triangulated, we can further choose and orient one of its $2n+1$ edges. There are $(4n+2)C_n$ such decorated triangulations. Now given a polygon Q with $n+3$ sides, again mark one of its sides as the base. If Q is triangulated, we can further mark one of the sides other than the base side. There are $(n+2)C_{n+1}$ such decorated triangulations. Then there is a simple bijection between these two kinds of decorated triangulations: We can either collapse the triangle in Q whose side is marked, or in reverse expand the oriented edge in P to a triangle and mark its new side. Thus

$$(4n + 2)C_n = (n + 2)C_{n+1}.$$

The binomial formula for C_n follows immediately from this relation and the initial condition $C_1 = 1$.

10.3.5 Fifth proof

This proof is based on the Dyck words interpretation of the Catalan numbers, so *Cn* is the number of ways to correctly match *n* pairs of brackets. We denote a (possibly empty) *correct* string with *c* and its inverse (where "[" and "]" are exchanged) with c^+. Since any *c* can be uniquely decomposed into *c* = [*c1*] *c2*, summing over the possible spots to place the closing bracket immediately gives the recursive definition

$$C_0 = 1 \quad \text{and} \quad C_{n+1} = \sum_{i=0}^{n} C_i \, C_{n-i} \quad \text{for} n \geq 0.$$

Now let *b* stand for a *balanced* string of length 2*n*—that is, containing an equal number of "[" and "]"—and $B_n = \binom{2n}{n} = d_n C_n$ with some factor *dn* ≥ 1. As above, any balanced string can be uniquely decomposed into either [*c*] *b* or] c^+ [*b*, so

$$B_{n+1} = 2\sum_{i=0}^{n} B_i C_{n-i}.$$

Also, any incorrect balanced string starts with *c*], so

$$B_{n+1} - C_{n+1} = \sum_{i=0}^{n} \binom{2i+1}{i} C_{n-i} = \sum_{i=0}^{n} \frac{2i+1}{i+1} B_i C_{n-i}.$$

Subtracting the above equations and using *Bi* = *di Ci* gives

$$C_{n+1} = 2\sum_{i=0}^{n} d_i C_i C_{n-i} - \sum_{i=0}^{n} \frac{2i+1}{i+1} d_i C_i C_{n-i} = \sum_{i=0}^{n} \frac{d_i}{i+1} C_i C_{n-i}.$$

Comparing coefficients with the original recursion formula for *Cn* gives *di* = *i* + 1, so

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

10.4 Hankel matrix

The *n*×*n* Hankel matrix whose (*i*, *j*) entry is the Catalan number C_{i+j-2} has determinant 1, regardless of the value of *n*. For example, for *n* = 4 we have

$$\det \begin{bmatrix} 1 & 1 & 2 & 5 \\ 1 & 2 & 5 & 14 \\ 2 & 5 & 14 & 42 \\ 5 & 14 & 42 & 132 \end{bmatrix} = 1.$$

Moreover, if the indexing is "shifted" so that the (*i*, *j*) entry is filled with the Catalan number C_{i+j-1} then the determinant is still 1, regardless of the value of *n*. For example, for *n* = 4 we have

$$\det \begin{bmatrix} 1 & 2 & 5 & 14 \\ 2 & 5 & 14 & 42 \\ 5 & 14 & 42 & 132 \\ 14 & 42 & 132 & 429 \end{bmatrix} = 1.$$

Taken together, these two conditions uniquely define the Catalan numbers.

10.5 History

The Catalan sequence was described in the 18th century by Leonhard Euler, who was interested in the number of different ways of dividing a polygon into triangles. The sequence is named after Eugène Charles Catalan, who discovered the connection to parenthesized expressions during his exploration of the Towers of Hanoi puzzle. The counting trick for Dyck words was found by D. André in 1887.

In 1988, it came to light that the Catalan number sequence had been used in China by the Mongolian mathematician Mingantu by 1730.[11][12] That is when he started to write his book *Ge Yuan Mi Lu Jie Fa*, which was completed by his student Chen Jixin in 1774 but published sixty years later. P.J. Larcombe (1999) sketched some of the features of the work of Mingantu, including the stimulus of Pierre Jartoux, who brought three infinite series to China early in the 1700s.

For instance, Ming used the Catalan sequence to express series expansions of $\sin(2\alpha)$ and $\sin(4\alpha)$ in terms of $\sin(\alpha)$.

10.6 Generalizations

The two-parameter sequence of non-negative integers $\frac{(2m)!(2n)!}{(m+n)!m!n!}$ is a generalization of the Catalan numbers. These are named **super-Catalan numbers**, by Ira Gessel. These number should not confused with the Schröder–Hipparchus numbers, which sometimes are also called super-Catalan numbers.

For $m = 1$, this is just two times the ordinary Catalan numbers, and for $m = n$, the numbers have an easy combinatorial description. However, other combinatorial descriptions are only known[13] for $m = 2$ and $m = 3$, and it is an open problem to find a general combinatorial interpretation.

10.7 See also

10.8 Notes

[1] Cormen, Thomas H.; Leiserson, Charles E.; Rivest, Ronald L. (1990). "Dynamic Programming". *Introduction to Algorithms*. Cambridge, Massachusetts: The MIT Press. p. 304. ISBN 0262031418.

[2] Koshy, Thomas; Salmassi, Mohammad (2006). *Parity and primality of Catalan numbers*. The Mathematical Association of America.

[3] Equivalent definitions of Dyck paths

[4] Stanley p.221 example (e)

[5] Črepinšek, Matej; Mernik, Luka (2009). "An efficient representation for solving Catalan number related problems" (PDF). *International Journal of Pure and Applied Mathematics*. **56** (4): 589–604.

[6] Kim, K. H.; Roush, F. W. (1978), "Enumeration of isomorphism classes of semiorders", *Journal of Combinatorics, Information &System Sciences*, **3** (2): 58–61, MR 538212.

[7] Thompson, R. W.; King, C. J. (1972), "Systematic synthesis of separation schemes", *American Institution of Chemical Engineers Journal*, **18** (5): 941–948.

[8] A. de Segner, Enumeratio modorum, quibus figurae planae rectilineae per diagonales dividuntur in triangula. *Novi commentarii academiae scientiarum Petropolitanae* **7** (1758/59) 203–209.

[9] Renault, Marc, Lost (and found) in translation: André's actual method and its application to the generalized ballot problem. Amer. Math. Monthly 115 (2008), no. 4, 358-−363.

[10] Rukavicka Josef (2011), *On Generalized Dyck Paths, Electronic Journal of Combinatorics online*

[11] The 18th century Chinese discovery of the Catalan numbers

[12] Ming Antu, the First Inventor of Catalan Numbers in the World

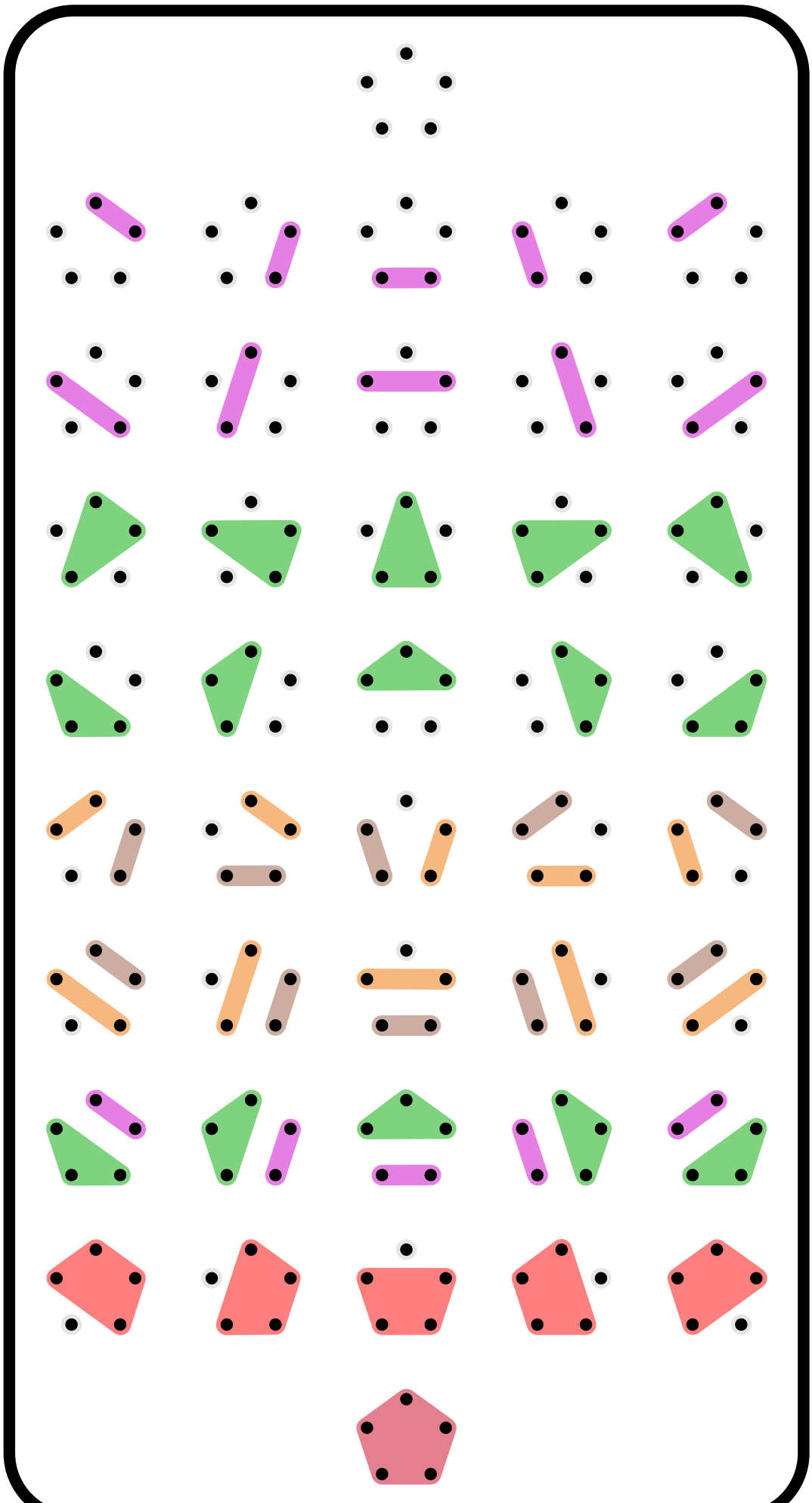
[13] Chen, Xin. "The super Catalan numbers S(m,m+s) for s<=3 and some integer factorial ratios" (PDF). *www-users.math.umn.edu/~{}reiner/REU/ChenWang2012.pdf*. Retrieved 26 September 2014.

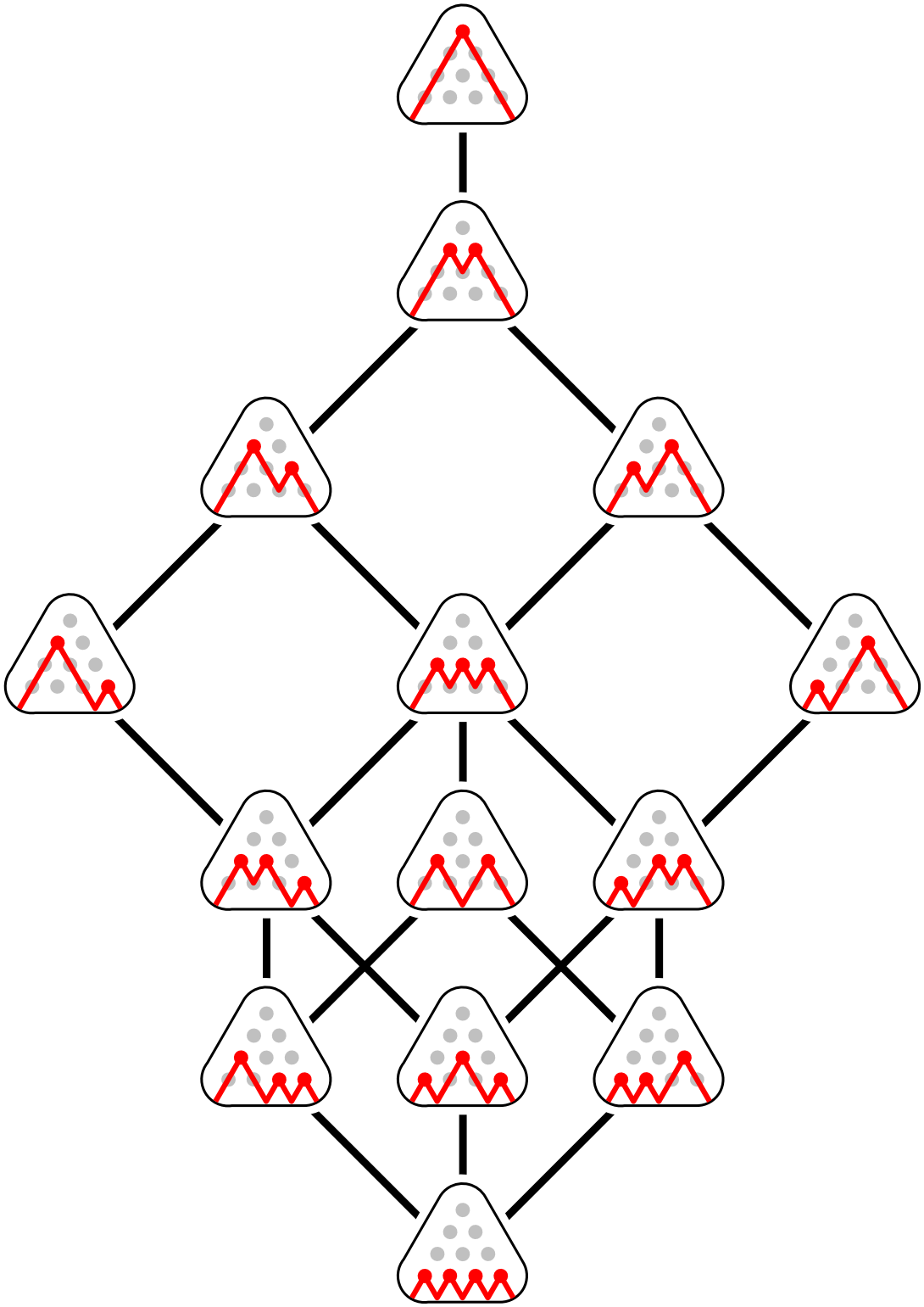
10.9 References

- Conway and Guy (1996) *The Book of Numbers*. New York: Copernicus, pp. 96–106.
- Gardner, Martin (1988), *Time Travel and Other Mathematical Bewilderments*, New York: W.H. Freeman and Company, pp. 253–266 (Ch. 20), ISBN 0-7167-1924-X
- Koshy, Thomas (2008), *Catalan Numbers with Applications*, Oxford University Press, ISBN 0-19-533454-X
- Koshy, Thomas & Zhenguang Gao (2011) "Some divisibility properties of Catalan numbers", Mathematical Gazette 95:96–102.
- Larcombe, P.J. (1999) "The 18th century Chinese discovery of the Catalan numbers", *Mathematical Spectrum* 32:5–7.
- Stanley, Richard P. (1999), *Enumerative combinatorics. Vol. 2*, Cambridge Studies in Advanced Mathematics, **62**, Cambridge University Press, ISBN 978-0-521-56069-6, MR 1676282
- Egecioglu, Omer (2009), *A Catalan-Hankel Determinant Evaluation* (PDF)

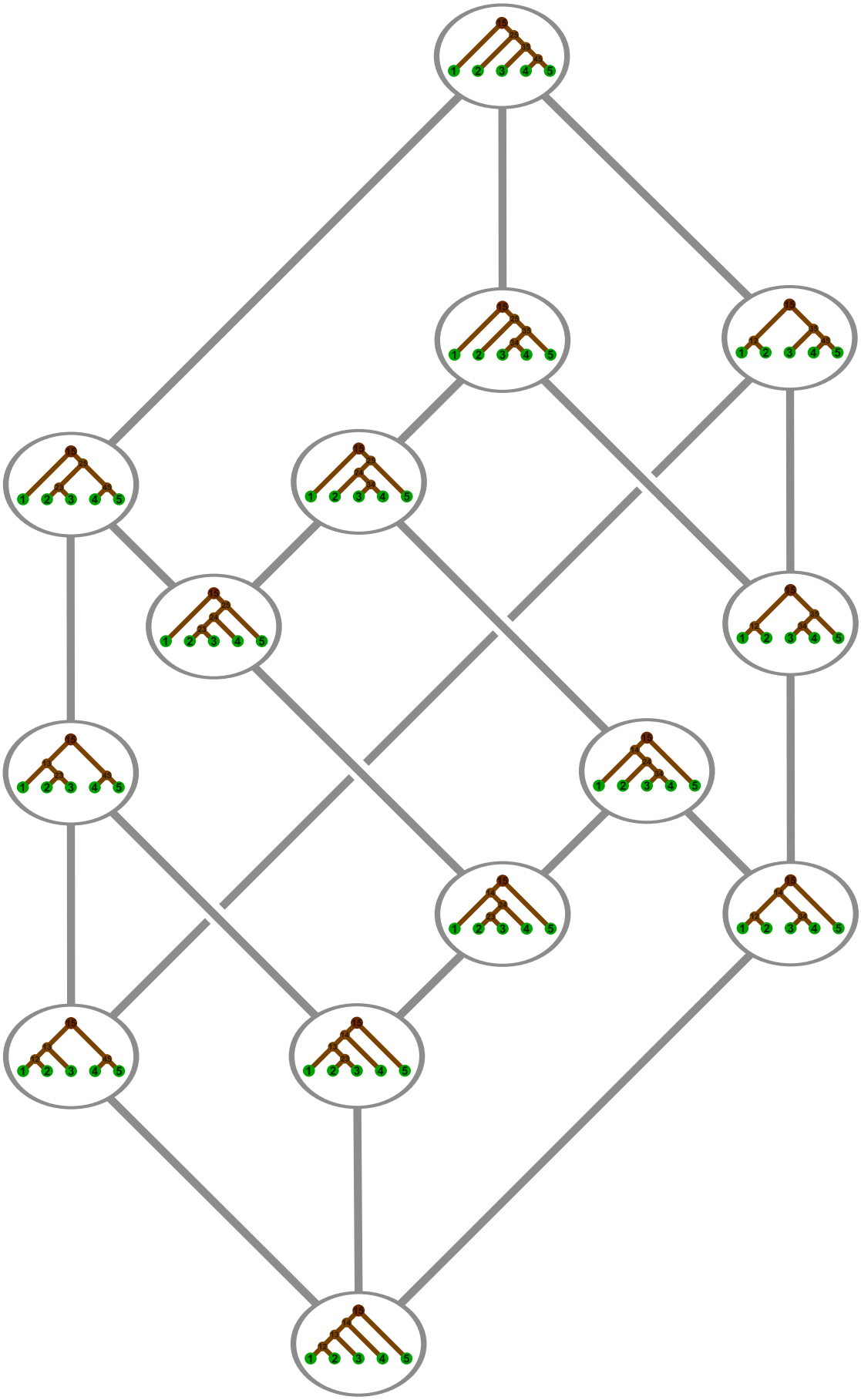
10.10 External links

- Stanley, Richard P. (1998), *Catalan addendum to Enumerative Combinatorics, Volume 2* (PDF)
- Weisstein, Eric W. "Catalan Number". *MathWorld*.
- Dickau, Robert M.: Catalan numbers Further examples.
- Davis, Tom: Catalan numbers. Still more examples.
- Schmidthammer, Jürgen: Catalan-Zahlen Zulassungsarbeit zum Staatsexamen (PDF-File; 7,05 MB)
- "Equivalence of Three Catalan Number Interpretations" from The Wolfram Demonstrations Project
- Dyck paths and binary trees in the FindStat database
- https://www.youtube.com/watch?v=pEJo0DJhYvU (Movie: generating all correct strings of parentheses with backtracking method in C language)

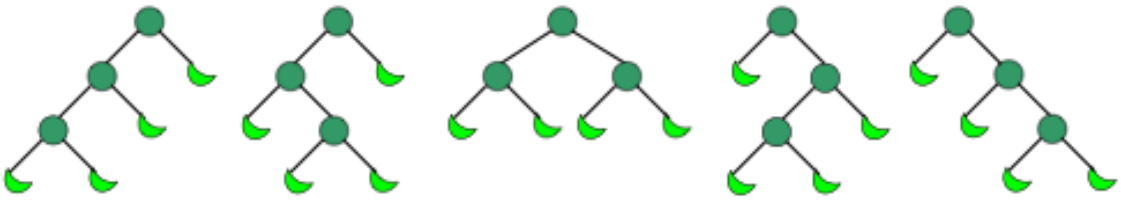


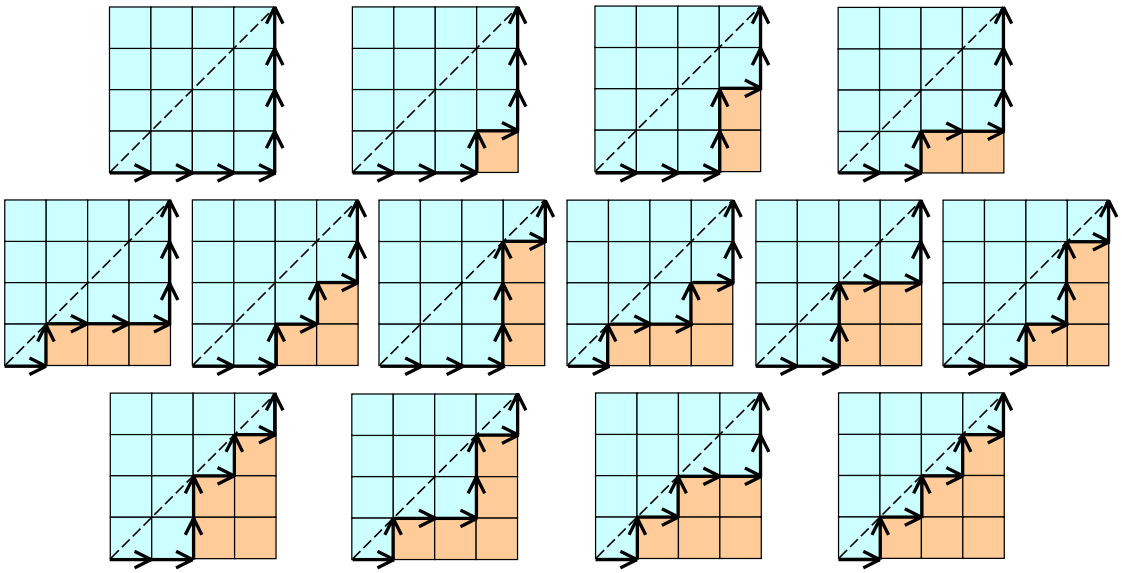


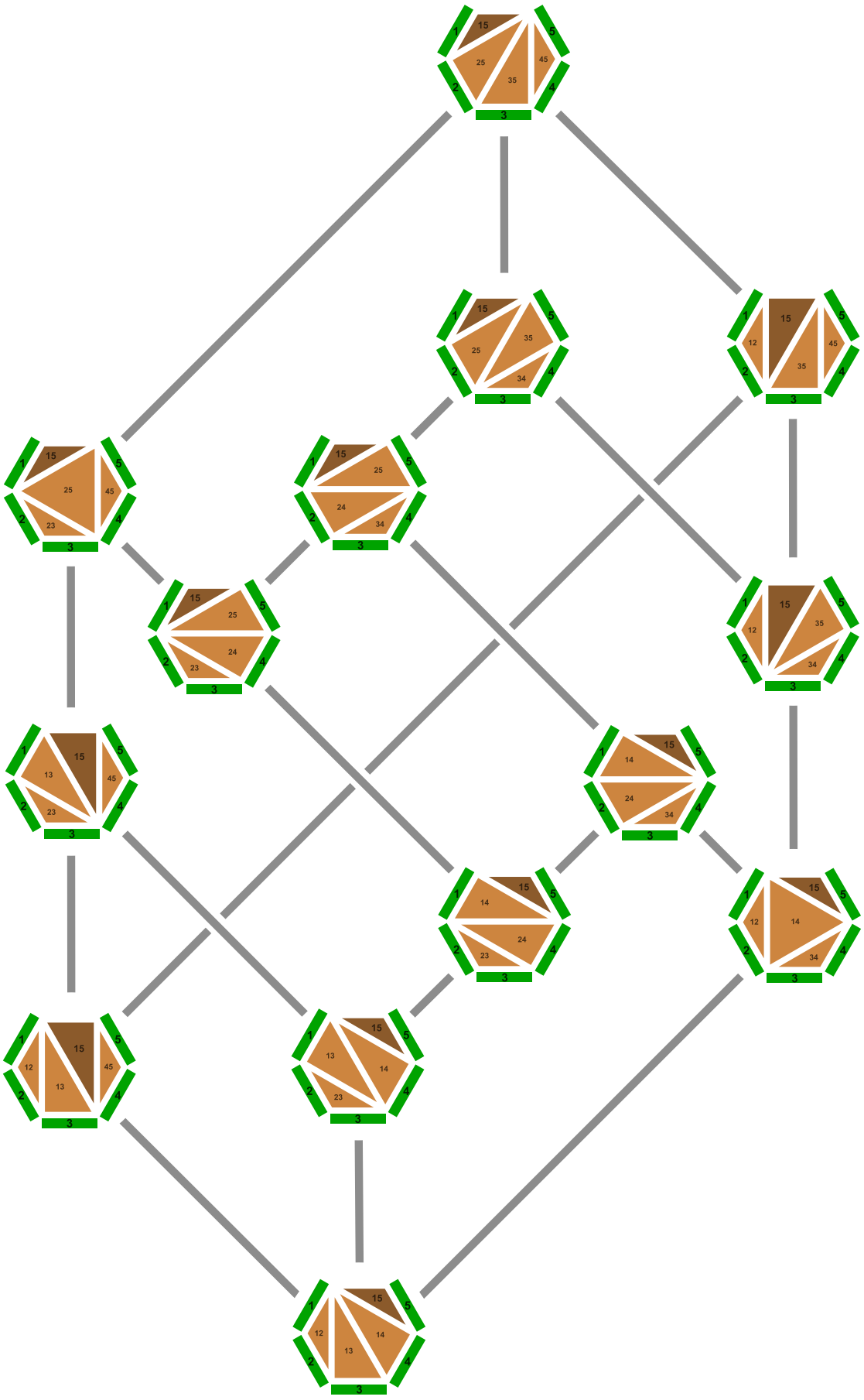
Lattice of the 14 Dyck words of length 8 - (*and*) *interpreted as* up *and* down



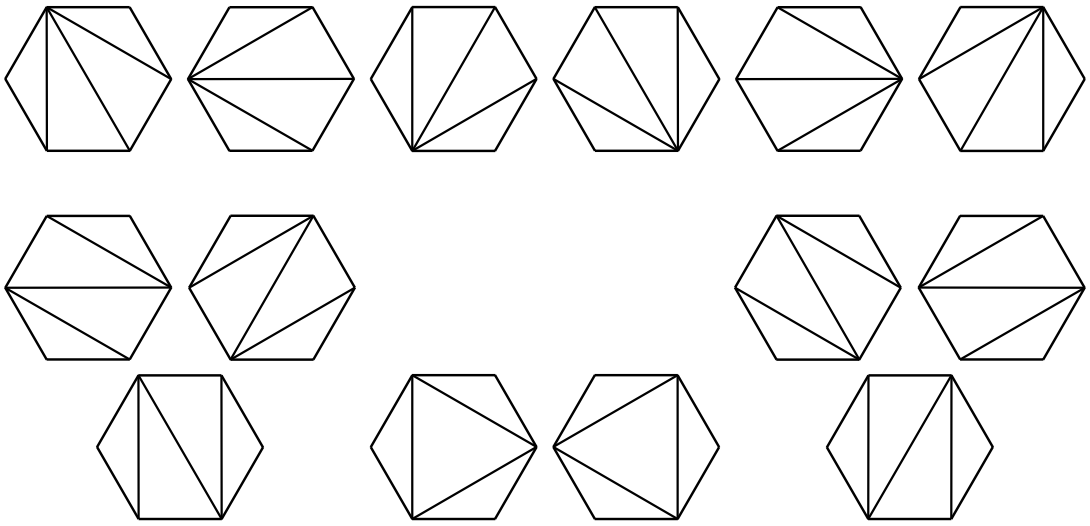
The associahedron of order 4 with the C_4=14 full binary trees with 5 leaves

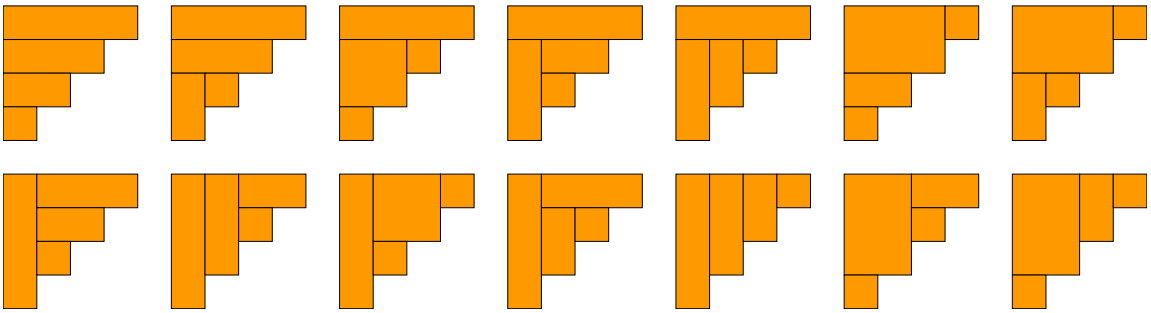


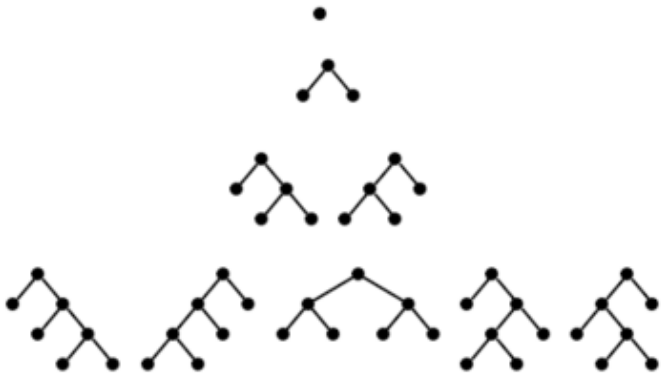




The triangles correspond to nodes of the binary trees.







Binary Trees

$n = 0$:	*	1 way
$n = 1$:	/\	1 way
$n = 2$:	/\ /\/\, / \	2 ways
$n = 3$:	/\ /\ /\ /\/\ / \ /\/\/\, /\/ \, / \/\, / \, / \	5 ways

Mountain Ranges

Catalan Number

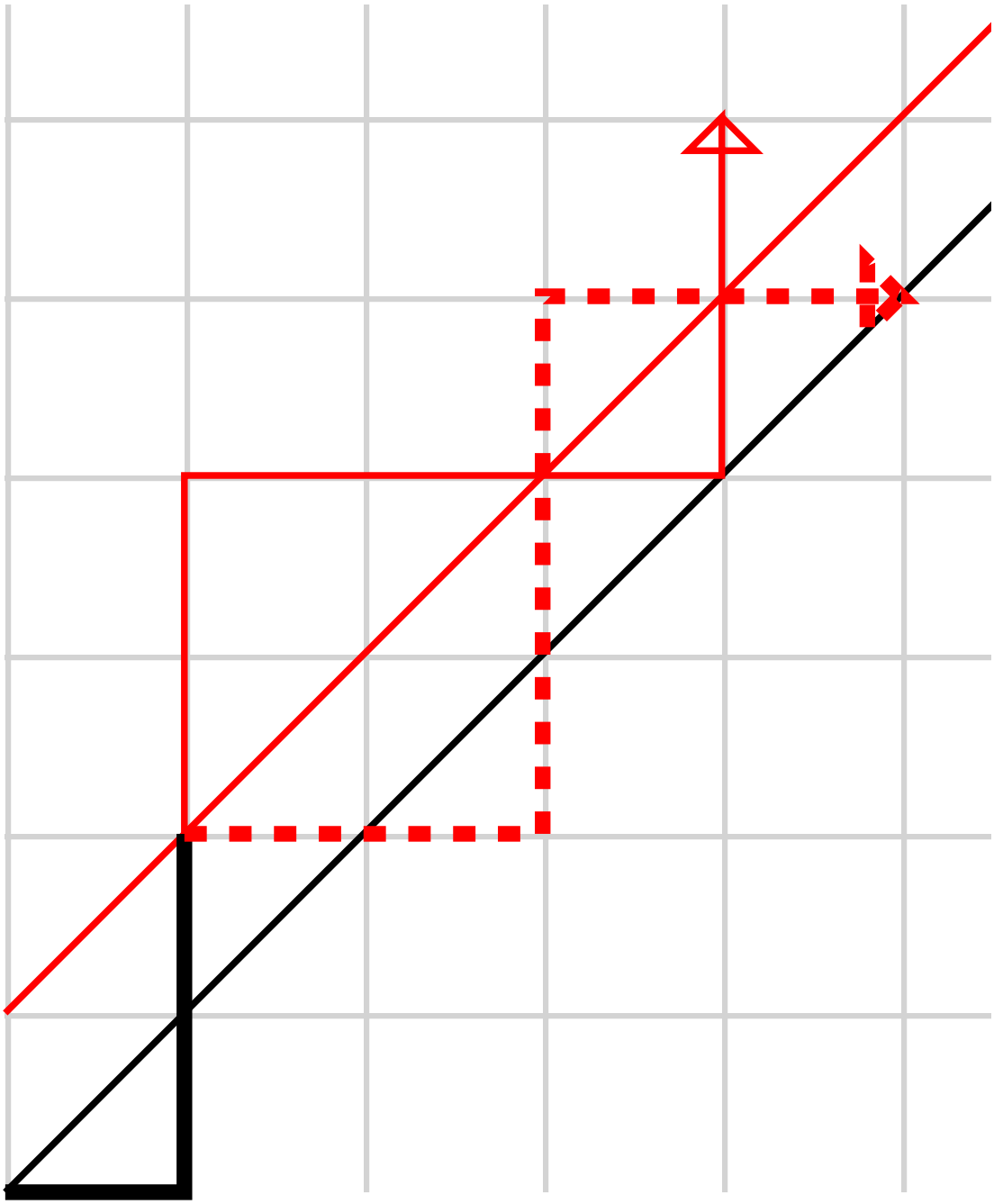


Figure 1. The invalid portion of the path is flipped. Bad paths reach (n - *1*, n + *1) instead of (n,n).*

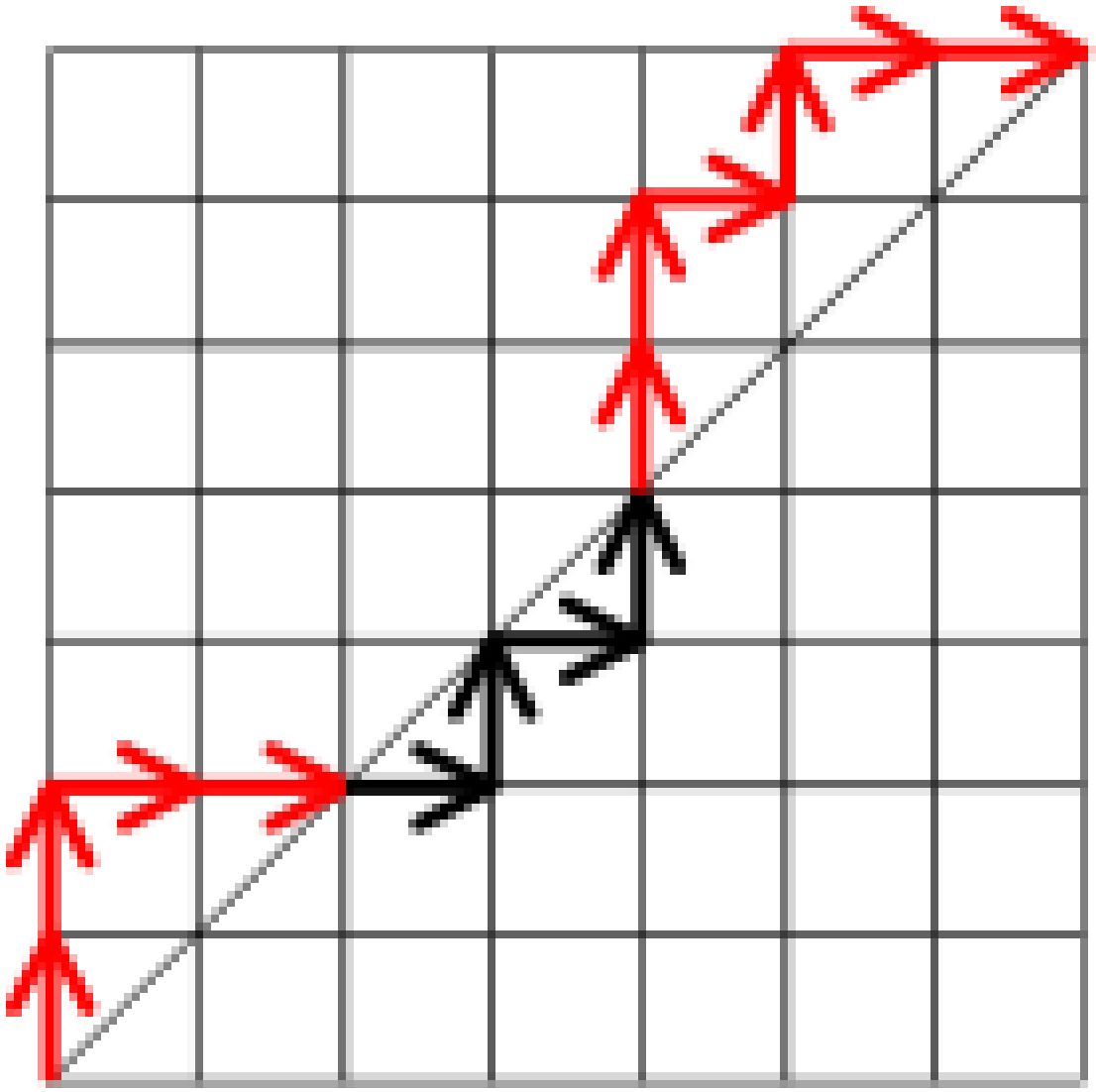


Figure 2. A path with exceedance 5.

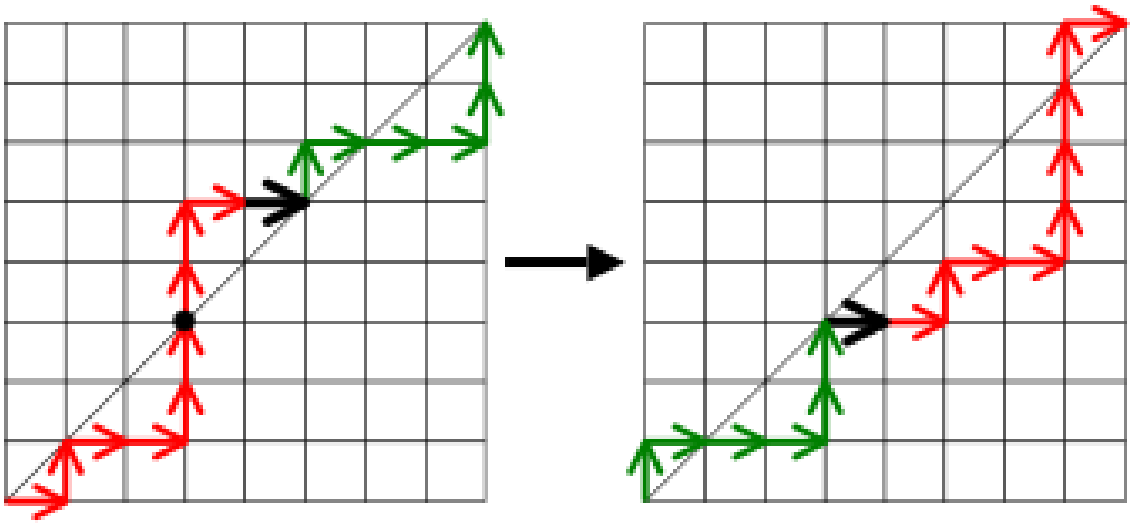


Figure 3. The green and red portions are being exchanged.

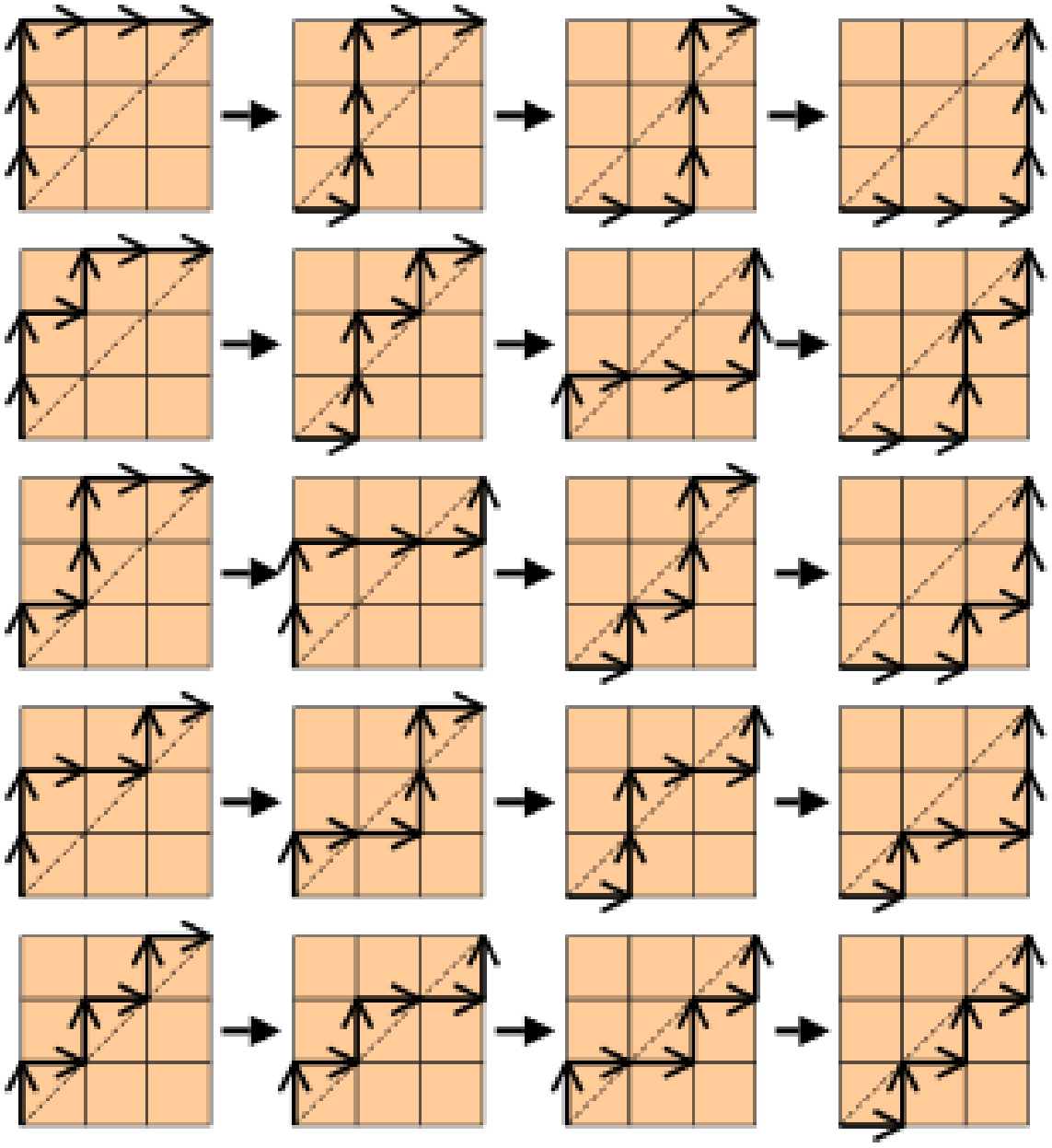
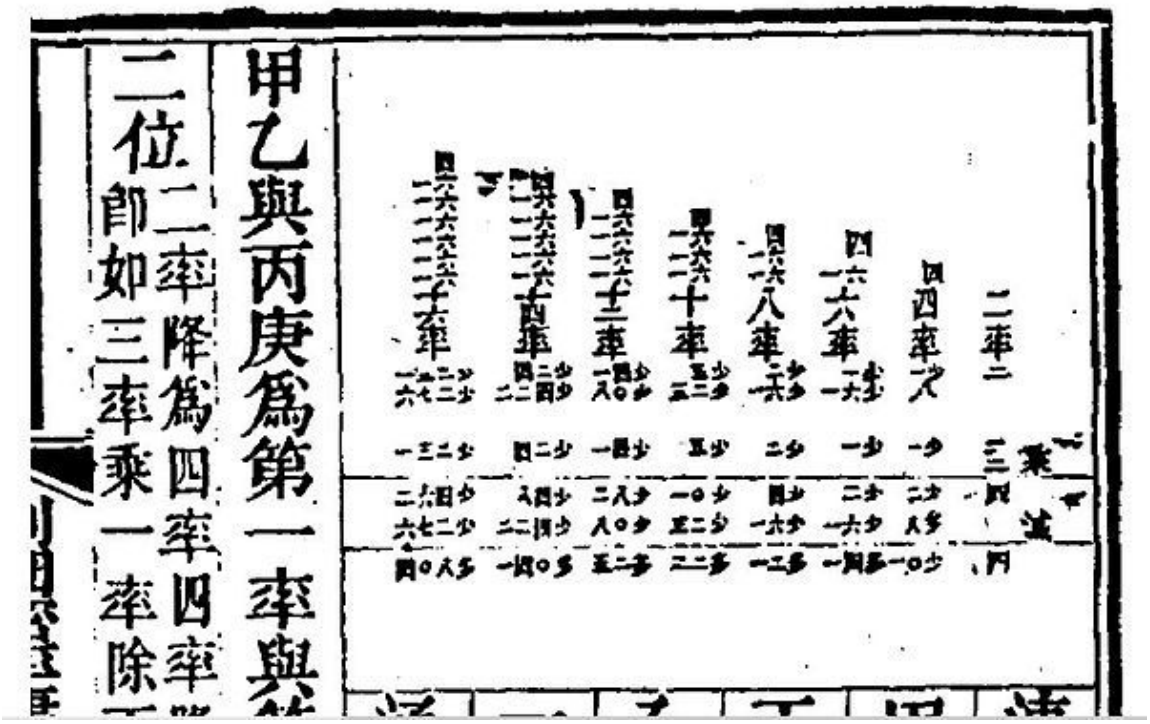


Figure 4. All monotonic paths in a 3×3 grid, illustrating the exceedance-decreasing algorithm.



Catalan numbers in Mingantu's book The Quick Method for Obtaining the Precise Ratio of Division of a Circle *volume III*

Chapter 11

Extended Euclidean algorithm

In arithmetic and computer programming, the **extended Euclidean algorithm** is an extension to the Euclidean algorithm, which computes, besides the greatest common divisor of integers a and b, the coefficients of Bézout's identity, that is integers x and y such that

$$ax + by = \gcd(a, b).$$

This is a certifying algorithm, because the gcd is the only number that can simultaneously satisfy this equation and divide the inputs. It allows one to compute also, with almost no extra cost, the quotients of a and b by their greatest common divisor.

Extended Euclidean algorithm also refers to a very similar algorithm for computing the polynomial greatest common divisor and the coefficients of Bézout's identity of two univariate polynomials.

The extended Euclidean algorithm is particularly useful when a and b are coprime, since x is the modular multiplicative inverse of a modulo b, and y is the modular multiplicative inverse of b modulo a. Similarly, the polynomial extended Euclidean algorithm allows one to compute the multiplicative inverse in algebraic field extensions and, in particular in finite fields of non prime order. It follows that both extended Euclidean algorithms are widely used in cryptography. In particular, the computation of the modular multiplicative inverse is an essential step in RSA public-key encryption method.

11.1 Description

The standard Euclidean algorithm proceeds by a succession of Euclidean divisions whose quotients are not used, only the *remainders* are kept. For the extended algorithm, the successive quotients are used. More precisely, the standard Euclidean algorithm with a and b as input, consists of computing a sequence $q_1, \ldots, q_k$ of quotients and a sequence $r_0, \ldots, r_{k+1}$ of remainders such that

$$\begin{aligned}
r_0 &= a \\
r_1 &= b \\
&\vdots \\
r_{i+1} &= r_{i-1} - q_i r_i \quad \text{and} \quad 0 \leq r_{i+1} < |r_i| \\
&\vdots
\end{aligned}$$

It is the main property of Euclidean division that the inequalities on the right define uniquely q_i and r_{i+1} from r_{i-1} and r_i.

The computation stops when one reaches a remainder r_{k+1} which is zero; the greatest common divisor is then the last non zero remainder r_k.

The extended Euclidean algorithm proceeds similarly, but adds two other sequences, as follows

$$
\begin{aligned}
r_0 &= a & r_1 &= b \\
s_0 &= 1 & s_1 &= 0 \\
t_0 &= 0 & t_1 &= 1 \\
&\vdots & &\vdots \\
r_{i+1} &= r_{i-1} - q_i r_i \quad \text{and} 0 \leq r_{i+1} < |r_i| \quad \text{defines (this} q_i) \\
s_{i+1} &= s_{i-1} - q_i s_i \\
t_{i+1} &= t_{i-1} - q_i t_i \\
&\vdots
\end{aligned}
$$

The computation also stops when $r_{k+1} = 0$ and gives

- r_k is the greatest common divisor of the input $a = r_0$ and $b = r_1$.
- The Bézout coefficients are s_k and t_k, that is $\gcd(a, b) = r_k = as_k + bt_k$
- The quotients of *a* and *b* by their greatest common divisor are given by $s_{k+1} = \pm\frac{b}{\gcd(a,b)}$ and $t_{k+1} = \pm\frac{a}{\gcd(a,b)}$

Moreover, if *a* and *b* are both positive and $\gcd(a, b) \neq \min(a, b)$, then

$$|s_i| < \frac{b}{\gcd(a, b)} \quad \text{and} \quad |t_i| < \frac{a}{\gcd(a, b)}$$

for $0 \leq i \leq k$. This means that the pair of Bézout's coefficients provided by the extended Euclidean algorithm is one of the two minimal pairs of Bézout coefficients. In addition it means that the algorithm can be done without integer overflow when *a* and *b* are representable integers.

11.1.1 Example

The following table shows how the extended Euclidean algorithm proceeds with input 240 and 46. The greatest common divisor is the last non zero entry, 2 in the column "remainder". The computation stops at row 6, because the remainder in it is 0. Bézout coefficients appear in the last two entries of the second-to-last row. In fact, it is easy to verify that −9 × 240 + 47 × 46 = 2. Finally the last two entries 23 and −120 of the last row are, up to the sign, the quotients of the input 46 and 240 by the greatest common divisor 2.

11.1.2 Proof

As $0 \leq r_{i+1} < |r_i|$, the sequence of the r_i is a decreasing sequence of nonnegative integers (from *i* = 2 on). Thus it must stop with some $r_{k+1} = 0$. This proves that the algorithm stops eventually.

As $r_{i+1} = r_{i-1} - r_i q_i$, the greatest common divisor is the same for (r_{i-1}, r_i) and (r_i, r_{i+1}). This shows that the greatest common divisor of the input $a = r_0, b = r_1$ is the same as that of $r_k, r_{k+1} = 0$. This proves that r_k is the greatest common divisor of *a* and *b*. (Until this point, the proof is the same as that of the classical Euclidean algorithm.)

As $a = r_0$ and $b = r_1$, we have $as_i + bt_i = r_i$ for *i* = 0 and 1. The relation follows by induction for all $i > 1$: $r_{i+1} = r_{i-1} - r_i q_i = (as_{i-1} + bt_{i-1}) - (as_i + bt_i)q_i = (as_{i-1} - as_i q_i) + (bt_{i-1} - bt_i q_i) = as_{i+1} + bt_{i+1}$. Thus s_k and t_k are Bézout coefficients.

Let us consider the matrix

$$A_i = \begin{pmatrix} s_{i-1} & s_i \\ t_{i-1} & t_i \end{pmatrix}.$$

The recurrence relation may be rewritten in matrix form

$$A_{i+1} = A_i \cdot \begin{pmatrix} 0 & 1 \\ 1 & -q_i \end{pmatrix}.$$

The matrix A_1 is the identity matrix and its determinant is one. The determinant of the rightmost matrix in the preceding formula is -1. It follows that the determinant of A_i is $(-1)^{i-1}$. In particular, for *i* = *k* + 1, we have $s_k t_{k+1} - t_k s_{k+1} = (-1)^k$. Viewing this as a Bézout's identity, this shows that s_{k+1} and t_{k+1} are coprime. The relation $as_{k+1} + bt_{k+1} = 0$ that has been proved above and Euclid's lemma shows that s_{k+1} divides *b* and t_{k+1} divides *a*. As they are coprime, they are, up to their sign the quotients of *b* and *a* by their greatest common divisor.

To prove the last assertion, assume that *a* and *b* are both positive and $\gcd(a, b) \neq \min(a, b)$. Then, $a \neq b$, and if $a < b$, it can be seen that the *s* and *t* sequences for (*a*,*b*) under the EEA are, up to initial 0s and 1s, the *t* and *s* sequences for (*b*,*a*). The definitions then show that the (*a*,*b*) case reduces to the (*b*,*a*) case. So assume that $a > b$ WLOG.

It can be seen that s_2 is 1 and s_3 (which exists by $\gcd(a, b) \neq \min(a, b)$) is a negative integer. Thereafter, the s_i alternate in sign and strictly increase in magnitude, which follows inductively from the definitions and the fact that $q_i \geq 1$ for $1 \leq i \leq k$, the case i=1 holding because $a > b$. The same is true for the t_i after the first few terms, for the same reason. Thus, $|s\ \ k+1\ \ | = |\frac{b}{\gcd(a,b)}|$ and $|t\ \ k+1\ \ | = |\frac{a}{\gcd(a,b)}|$ are strictly larger in absolute value than any previous s_i or t_i , respectively.

11.2 Polynomial extended Euclidean algorithm

See also: Polynomial greatest common divisor § Bézout's identity and extended GCD algorithm

For univariate polynomials with coefficients in a field, everything works in a similar way, Euclidean division, Bézout's identity and extended Euclidean algorithm. The first difference is that, in the Euclidean division and the algorithm, the inequality $0 \leq r_{i+1} < |r_i|$ has to be replaced by an inequality on the degrees $\deg r_{i+1} < \deg r_i$. Otherwise, everything which precedes in this article remains the same, simply by replacing integers by polynomials.

A second difference lies in the bound on the size of the Bézout coefficients provided by the extended Euclidean algorithm, which is more accurate in the polynomial case, leading to the following theorem.

If a and b are two nonzero polynomials, then the extended Euclidean algorithm produces the unique pair of polynomials (*s*, *t*) *such that*

$$as + bt = \gcd(a, b)$$

and

$$\deg s < \deg b - \deg(\gcd(a, b)), \quad \deg t < \deg a - \deg(\gcd(a, b)).$$

A third difference is that, in the polynomial case, the greatest common divisor is defined only up to the multiplication by a non zero constant. There are several ways to define the greatest common divisor unambiguously.

In mathematics, it is common to require that the greatest common divisor be a monic polynomial. To get this, it suffices to divide every element of the output by the leading coefficient of r_k. This allows that, if *a* and *b* are coprime, one gets 1 in the right-hand side of Bézout's inequality. Otherwise, one may get any non-zero constant. In computer algebra, the polynomials have commonly integers coefficients, and this way of normalizing the greatest common divisor introduces too many fractions to be convenient.

The second way to normalize the greatest common divisor in the case of polynomials with integers coefficients is to divide every output by the content of r_k, to get a primitive greatest common divisor. If the input polynomials are coprime, this normalization provides also a greatest common divisor equal to 1. The drawback of this approach is that a lot of fractions should be computed and simplified during the computation.

A third approach consists in extending the algorithm of subresultant pseudo-remainder sequences in a way that is similar to the extension of the Euclidean algorithm to the extended Euclidean algorithm. This allows that, when starting

with polynomials with integer coefficients, all polynomials that are computed have integer coefficients. Moreover, every computed remainder r_i is a subresultant polynomial. In particular, if the input polynomials are coprime, then the Bézout's identity becomes

$$as + bt = \text{Res}(a, b),$$

where $\text{Res}(a, b)$ denotes the resultant of a and b. In this form of Bézout's identity there is no denominator in the formula. If one divides everything by the resultant one gets the classical Bézout's identity, with an explicit common denominator for the rational numbers that appear in it.

11.3 Pseudocode

To implement the algorithm that is described above, one should first remark that only the two last values of the indexed variables are needed at each step. Thus, for saving memory, each indexed variable must be replaced by only two variables.

For simplicity, the following algorithm (and the other algorithms in this article) uses parallel assignments. In a programming language which does not have this feature, the parallel assignments need to be simulated with an auxiliary variable. For example, the first one,

(old_r, r) := (r, old_r - quotient *r)

is equivalent to

prov := r; r := old_r - quotient * prov; old_r := prov;

and similarly for the other parallel assignments. This leads to the following code:

function extended_gcd(a, b) s := 0; old_s := 1 t := 1; old_t := 0 r := b; old_r := a **while** r ≠ 0 quotient := old_r **div** r (old_r, r) := (r, old_r - quotient * r) (old_s, s) := (s, old_s - quotient * s) (old_t, t) := (t, old_t - quotient * t) **output** "Bézout coefficients:", (old_s, old_t) **output** "greatest common divisor:", old_r **output** "quotients by the gcd:", (t, s)

The quotients of a and b by their greatest common divisor, which are output, may have an incorrect sign. This is easy to correct at the end of the computation, but has not been done here for simplifying the code. Similarly, if either a or b is zero and the other is negative, the greatest common divisor that is output is negative, and all the signs of the output must be changed.

11.4 Simplification of fractions

A fraction a/b is in canonical simplified form if a and b are coprime and b is positive. This canonical simplified form can be obtained by replacing the three **output** lines of the preceding pseudo code by

if $s = 0$ **then output** "Division by zero" **if** $s < 0$ **then** $s := -s$; $t := -t$ (*for avoiding negative denominators*) **if** $s = 1$ **then output** $-t$ (*for avoiding denominators equal to* 1) **output** $-t/s$

The proof of this algorithm relies on the fact that s and t are two coprime integers such that $as + bt = 0$, and thus $\frac{a}{b} = -\frac{t}{s}$. To get the canonical simplified form, it suffices to move the minus sign for having a positive denominator.

If b divides a evenly, the algorithm executes only one iteration, and we have $s = 1$ at the end of the algorithm. It is the only case where the output is an integer.

11.5 Computing multiplicative inverses in modular structures

The extended Euclidean algorithm is the basic tool for computing multiplicative inverses in modular structures, typically the modular integers and the algebraic field extensions. An important instance of the latter case are the finite fields of non-prime order.

11.5.1 Modular integers

Main article: Modular arithmetic

If n is a positive integer, the ring $\mathbf{Z}/n\mathbf{Z}$ may be identified with the set $\{0, 1, ..., n-1\}$ of the remainders of Euclidean division by n, the addition and the multiplication consisting in taking the remainder by n of the result of the addition and the multiplication of integers. An element a of $\mathbf{Z}/n\mathbf{Z}$ has a multiplicative inverse (that is, it is a unit) if it is coprime to n. In particular, if n is prime, a has a multiplicative inverse if it is not zero (modulo n). Thus $\mathbf{Z}/n\mathbf{Z}$ is a field if and only if n is prime.

Bézout's identity asserts that a and n are coprime if and only if there exist integers s and t such that

$$ns + at = 1$$

Reducing this identity modulo n gives

$$at = 1 \mod n.$$

Thus t, or, more exactly, the remainder of the division of t by n, is the multiplicative inverse of a modulo n.

To adapt the extended Euclidean algorithm to this problem, one should remark that the Bézout coefficient of n is not needed, and thus does not need to be computed. Also, for getting a result which is positive and lower than n, one may use the fact that the integer t provided by the algorithm satisfies $|t| < n$. That is, if $t < 0$, one must add n to it at the end. This results in the pseudocode, in which the input n is an integer larger than 1.

function inverse(a, n) t := 0; newt := 1; r := n; newr := a; **while** newr ≠ 0 quotient := r **div** newr (t, newt) := (newt, t - quotient * newt) (r, newr) := (newr, r - quotient * newr) **if** r > 1 **then return** "a is not invertible" **if** t < 0 **then** t := t + n **return** t

11.5.2 Simple algebraic field extensions

Extended Euclidean algorithm is also the main tool for computing multiplicative inverses in simple algebraic field extensions. An important case, widely used in cryptography and coding theory is that of finite fields of non-prime order. In fact, if p is a prime number, and $q = p^d$, the field of order q is a simple algebraic extension of the prime field of p elements, generated by a root of an irreducible polynomial of degree d.

A simple algebraic extension L of a field K, generated by the root of an irreducible polynomial p of degree d may be identified to the quotient ring $K[X]/\langle p\rangle$, , and its elements are in bijective correspondence with the polynomials of degree less than d. The addition in L is the addition of polynomials. The multiplication in L is the remainder of the Euclidean division by p of the product of polynomials. Thus, to complete the arithmetic in L, it remains only to define how to compute multiplicative inverses. This is done by the extended Euclidean algorithm.

The algorithm is very similar to that provided above for computing the modular multiplicative inverse. There are two main differences: firstly the last but one line is not needed, because the Bézout coefficient that is provided has always a degree less than d. Secondly, the greatest common divisor which is provided, when the input polynomials are coprime, may be any non zero element of K; this Bézout coefficient (a polynomial generally of positive degree) has thus to be multiplied by the inverse of this element of K. In the pseudocode which follows, p is a polynomial of degree greater than one, and a is a polynomial. Moreover, **div** is an auxiliary function that computes the quotient of the Euclidean division.

function inverse(a, p) t := 0; newt := 1; r := p; newr := a; **while** newr ≠ 0 quotient := r **div** newr (r, newr) := (newr, r - quotient * newr) (t, newt) := (newt, t - quotient * newt) **if** degree(r) > 0 then **return** "Either p is not irreducible or a is a multiple of p" **return** (1/r) * t

Example

For example, if the polynomial used to define the finite field GF(2^8) is $p = x^8 + x^4 + x^3 + x + 1$, and $a = x^6 + x^4 + x + 1$ is the element whose inverse is desired, then performing the algorithm results in the computation described in the

following table. Let us recall that in fields of order 2^n, one has -z = z and z + z = 0 for every element z in the field). Note also that 1 being the only nonzero element of GF(2), the adjustment in the last line of the pseudocode is not needed.

Thus, the inverse is $x^7 + x^6 + x^3 + x$, as can be confirmed by multiplying the two elements together, and taking the remainder by p of the result.

11.6 The case of more than two numbers

One can handle the case of more than two numbers iteratively. First we show that $\gcd(a, b, c) = \gcd(\gcd(a, b), c)$. To prove this let $d = \gcd(a, b, c)$. By definition of gcd d is a divisor of a and b . Thus $\gcd(a, b) = kd$ for some k . Similarly d is a divisor of c so $c = jd$ for some j . Let $u = \gcd(k, j)$. By our construction of u , $ud|a, b, c$ but since d is the greatest divisor u is a unit. And since $ud = \gcd(\gcd(a, b), c)$ the result is proven.

So if $na + mb = \gcd(a, b)$ then there are x and y such that $x \gcd(a, b) + yc = \gcd(a, b, c)$ so the final equation will be

$$x(na + mb) + yc = (xn)a + (xm)b + yc = \gcd(a, b, c).$$

So then to apply to n numbers we use induction

$$\gcd(a_1, a_2, \ldots, a_n) = \gcd(a_1, \gcd(a_2, \gcd(a_3, \ldots, \gcd(a_{n-1}, a_n))), \ldots),$$

with the equations following directly.

11.7 See also

- Euclidean domain
- Linear congruence theorem
- Kuṭṭaka

11.8 References

- Knuth, Donald. *The Art of Computer Programming*. Addison-Wesley. Volume 2, Chapter 4.
- Thomas H. Cormen, Charles E. Leiserson, Ronald L. Rivest, and Clifford Stein. *Introduction to Algorithms*, Second Edition. MIT Press and McGraw-Hill, 2001. ISBN 0-262-03293-7. Pages 859–861 of section 31.2: Greatest common divisor.

11.9 External links

- Source for the form of the algorithm used to determine the multiplicative inverse in GF(2^8)

Chapter 12

Factorial

In mathematics, the **factorial** of a non-negative integer n, denoted by $n!$, is the product of all positive integers less than or equal to n. For example,

$$5! = 5 \times 4 \times 3 \times 2 \times 1 = 120.$$

The value of $0!$ is 1, according to the convention for an empty product.[1]

The factorial operation is encountered in many areas of mathematics, notably in combinatorics, algebra, and mathematical analysis. Its most basic occurrence is the fact that there are $n!$ ways to arrange n distinct objects into a sequence (i.e., permutations of the set of objects). This fact was known at least as early as the 12th century, to Indian scholars.[2] Fabian Stedman, in 1677, described factorials as applied to change ringing.[3] After describing a recursive approach, Stedman gives a statement of a factorial (using the language of the original):

> Now the nature of these methods is such, that the changes on one number comprehends [includes] the changes on all lesser numbers, ... insomuch that a compleat Peal of changes on one number seemeth to be formed by uniting of the compleat Peals on all lesser numbers into one entire body;[4]

The notation $n!$ was introduced by Christian Kramp in 1808.[5]

The definition of the factorial function can also be extended to non-integer arguments, while retaining its most important properties; this involves more advanced mathematics, notably techniques from mathematical analysis.

12.1 Definition

The factorial function is formally defined by the product

$$n! = \prod_{k=1}^{n} k,$$

or by the recurrence relation

$$n! = \begin{cases} 1 & \text{if} n = 0, \\ (n-1)! \times n & \text{if} n > 0 \end{cases}.$$

The factorial function can also be defined by using the power rule as

$$n! = D^n x^n. \quad [6]$$

All of the above definitions incorporate the instance

$$0! = 1,$$

in the first case by the convention that the product of no numbers at all is 1. This is convenient because:

- There is exactly one permutation of zero objects (with nothing to permute, "everything" is left in place).
- The recurrence relation $(n + 1)! = n! \times (n + 1)$, valid for $n > 0$, extends to $n = 0$.
- It allows for the expression of many formulae, such as the exponential function, as a power series:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

- It makes many identities in combinatorics valid for all applicable sizes. The number of ways to choose 0 elements from the empty set is $\binom{0}{0} = \frac{0!}{0!0!} = 1$. More generally, the number of ways to choose (all) n elements among a set of n is $\binom{n}{n} = \frac{n!}{n!0!} = 1$.

The factorial function can also be defined for non-integer values using more advanced mathematics, detailed in the section below. This more generalized definition is used by advanced calculators and mathematical software such as Maple or Mathematica.

12.2 Applications

Although the factorial function has its roots in combinatorics, formulas involving factorials occur in many areas of mathematics.

- There are $n!$ different ways of arranging n distinct objects into a sequence, the permutations of those objects.
- Often factorials appear in the denominator of a formula to account for the fact that ordering is to be ignored. A classical example is counting k-combinations (subsets of k elements) from a set with n elements. One can obtain such a combination by choosing a k-permutation: successively selecting and removing an element of the set, k times, for a total of

  $$n^{\underline{k}} = n(n-1)(n-2)\cdots(n-k+1)$$

  possibilities. This however produces the k-combinations in a particular order that one wishes to ignore; since each k-combination is obtained in $k!$ different ways, the correct number of k-combinations is

  $$\frac{n^{\underline{k}}}{k!} = \frac{n(n-1)(n-2)\cdots(n-k+1)}{k(k-1)(k-2)\cdots 1}.$$

  This number is known as the binomial coefficient $\binom{n}{k}$, because it is also the coefficient of X^k in $(1 + X)^n$.

- Factorials occur in algebra for various reasons, such as via the already mentioned coefficients of the binomial formula, or through averaging over permutations for symmetrization of certain operations.
- Factorials also turn up in calculus; for example they occur in the denominators of the terms of Taylor's formula, where they are used as compensation terms due to the n-th derivative of x^n being equivalent to $n!$.
- Factorials are also used extensively in probability theory.

- Factorials can be useful to facilitate expression manipulation. For instance the number of k-permutations of n can be written as

$$n^{\underline{k}} = \frac{n!}{(n-k)!};$$

while this is inefficient as a means to compute that number, it may serve to prove a symmetry property of binomial coefficients:

$$\binom{n}{k} = \frac{n^{\underline{k}}}{k!} = \frac{n!}{(n-k)!k!} = \frac{n^{\underline{n-k}}}{(n-k)!} = \binom{n}{n-k}.$$

12.3 Number theory

Factorials have many applications in number theory. In particular, $n!$ is necessarily divisible by all prime numbers up to and including n. As a consequence, $n > 5$ is a composite number if and only if

$$(n-1)! \equiv 0 \pmod{n}.$$

A stronger result is Wilson's theorem, which states that

$$(p-1)! \equiv -1 \pmod{p}$$

if and only if p is prime.

Legendre's formula gives the multiplicity of the prime p occurring in the prime factorization of $n!$ as

$$\sum_{i=1}^{\infty} \left\lfloor \frac{n}{p^i} \right\rfloor$$

or, equivalently,

$$\frac{n - s_p(n)}{p-1},$$

where $s_p(n)$ denotes the sum of the standard base-p digits of n.

The only factorial that is also a prime number is 2, but there are many primes of the form $n! \pm 1$, called factorial primes.

All factorials greater than $1!$ are even, as they are all multiples of 2. Also, all factorials from $5!$ upwards are multiples of 10 (and hence have a trailing zero as their final digit), because they are multiples of 5 and 2.

12.4 Series of reciprocals

The reciprocals of factorials produce a convergent series: (see e)

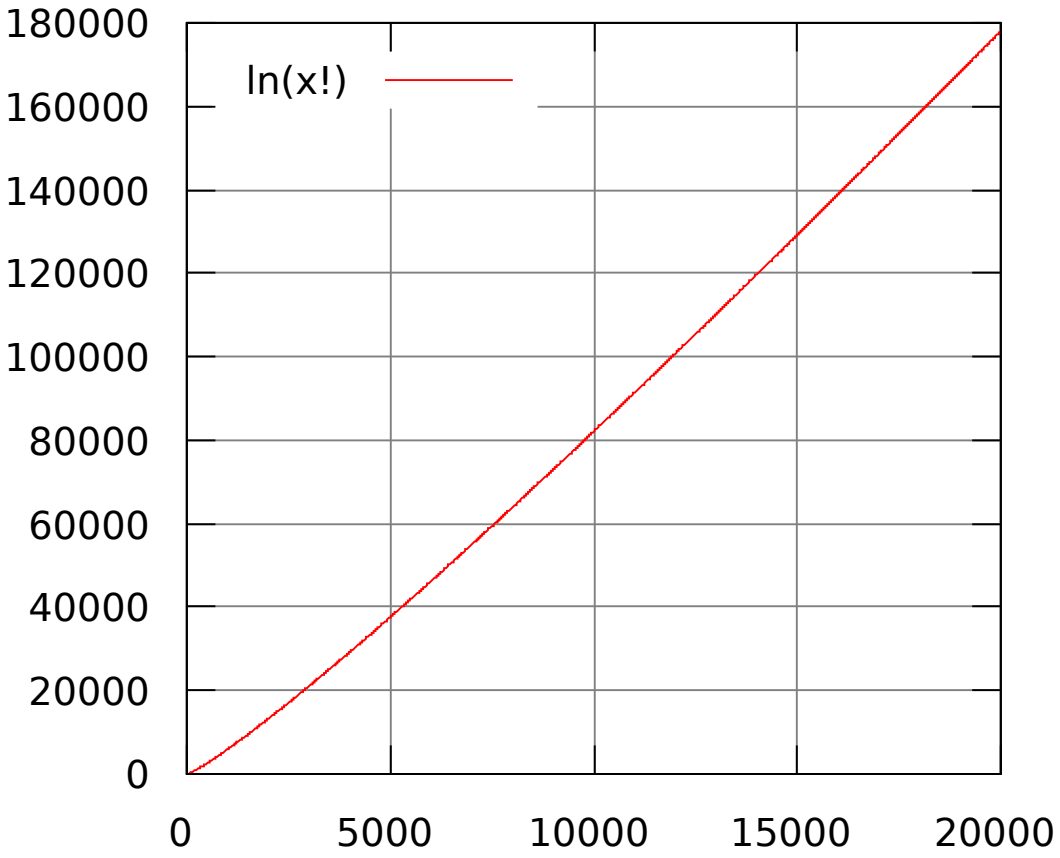
$$\sum_{n=0}^{\infty} \frac{1}{n!} = \frac{1}{1} + \frac{1}{1} + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \ldots = e\,.$$

Although the sum of this series is an irrational number, it is possible to multiply the factorials by positive integers to produce a convergent series with a rational sum:

$$\sum_{n=0}^{\infty} \frac{1}{(n+2)n!} = \frac{1}{2} + \frac{1}{3} + \frac{1}{8} + \frac{1}{30} + \frac{1}{144} \ldots = 1\,.$$

The convergence of this series to 1 can be seen from the fact that its partial sums are less than one by an inverse factorial. Therefore, the factorials do not form an irrationality sequence.[7]

12.5 Rate of growth and approximations for large n



Plot of the natural logarithm of the factorial

As n grows, the factorial $n!$ increases faster than all polynomials and exponential functions (but slower than double exponential functions) in n.

Most approximations for $n!$ are based on approximating its natural logarithm

$$\ln n! = \sum_{x=1}^{n} \ln x.$$

The graph of the function $f(n) = \ln n!$ is shown in the figure on the right. It looks approximately linear for all reasonable values of n, but this intuition is false. We get one of the simplest approximations for $\ln n!$ by bounding the sum with an integral from above and below as follows:

$$\int_1^n \ln x\, dx \leq \sum_{x=1}^{n} \ln x \leq \int_0^n \ln(x+1)\, dx$$

which gives us the estimate

$$n \ln\left(\frac{n}{e}\right) + 1 \leq \ln n! \leq (n+1) \ln\left(\frac{n+1}{e}\right) + 1.$$

Hence $\ln n! \sim n \ln n$ (see Big *O* notation). This result plays a key role in the analysis of the computational complexity of sorting algorithms (see comparison sort). From the bounds on $\ln n!$ deduced above we get that

$$e\left(\frac{n}{e}\right)^n \leq n! \leq e\left(\frac{n+1}{e}\right)^{n+1}.$$

It is sometimes practical to use weaker but simpler estimates. Using the above formula it is easily shown that for all n we have $(n/3)^n < n!$, and for all $n \geq 6$ we have $n! < (n/2)^n$.

For large n we get a better estimate for the number $n!$ using Stirling's approximation:

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

This in fact comes from an asymptotic series for the logarithm, and n factorial lies between this and the next approximation:

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}.$$

Another approximation for $\ln n!$ is given by Srinivasa Ramanujan (Ramanujan 1988)

$$\ln n! \approx n \ln n - n + \frac{\ln(n(1+4n(1+2n)))}{6} + \frac{\ln(\pi)}{2}$$

or

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n [1 + 1/(2n) + 1/(8n^2)]^{1/6}.$$

Both this and $\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}$ give a relative error on the order of $1/n^3$, but Ramanujan's is about four times more accurate. However, if we use *two* correction terms (as in Ramanujan's approximation) the relative error will be of order $1/n^5$:

$$n! \approx \sqrt{2\pi n}\left(\frac{n}{e}\right)^n \exp\left(\frac{1}{12n} - \frac{1}{360n^3}\right)$$

12.6 Computation

If efficiency is not a concern, computing factorials is trivial from an algorithmic point of view: successively multiplying a variable initialized to 1 by the integers 2 up to n (if any) will compute $n!$, provided the result fits in the variable. In functional languages, the recursive definition is often implemented directly to illustrate recursive functions.

The main practical difficulty in computing factorials is the size of the result. To assure that the exact result will fit for all legal values of even the smallest commonly used integral type (8-bit signed integers) would require more than 700 bits, so no reasonable specification of a factorial function using fixed-size types can avoid questions of overflow. The values 12! and 20! are the largest factorials that can be stored in, respectively, the 32-bit and 64-bit integers commonly used in personal computers. Floating-point representation of an approximated result allows going a bit further, but this also remains quite limited by possible overflow. Most calculators use scientific notation with 2-digit

decimal exponents, and the largest factorial that fits is then 69!, because $69! < 10^{100} < 70!$. Other implementations (e.g., computer software such as spreadsheet programs) can often handle larger values.

Most software applications will compute small factorials by direct multiplication or table lookup. Larger factorial values can be approximated using Stirling's formula. Wolfram Alpha can calculate exact results for the ceiling function and floor function applied to the binary, natural and common logarithm of $n!$ for values of n up to 249999, and up to 20,000,000! for the integers.

If the exact values of large factorials are needed, they can be computed using arbitrary-precision arithmetic. Instead of doing the sequential multiplications $((1 \times 2) \times 3) \times 4 \dots$, a program can partition the sequence into two parts, whose products are roughly the same size, and multiply them using a divide-and-conquer method. This is often more efficient.[8]

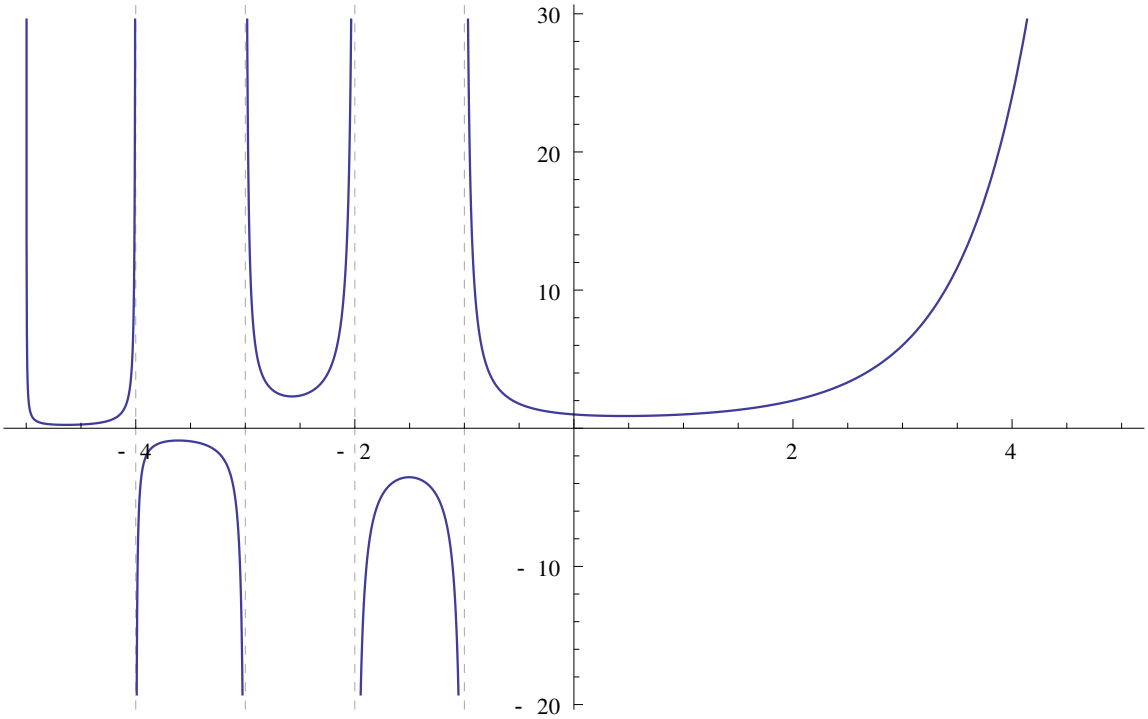
The asymptotically best efficiency is obtained by computing $n!$ from its prime factorization. As documented by Peter Borwein, prime factorization allows $n!$ to be computed in time $O(n(\log n \log\log n)^2)$, provided that a fast multiplication algorithm is used (for example, the Schönhage–Strassen algorithm).[9] Peter Luschny presents source code and benchmarks for several efficient factorial algorithms, with or without the use of a prime sieve.[10]

12.7 Extension of factorial to non-integer values of argument

12.7.1 The Gamma and Pi functions

Main article: Gamma function

Besides nonnegative integers, the factorial function can also be defined for non-integer values, but this requires more



The factorial function, generalized to all real numbers except negative integers. For example, 0! = 1! = 1, (−0.5)! = √π, (0.5)! = √π/2.

advanced tools from mathematical analysis. One function that "fills in" the values of the factorial (but with a shift of 1 in the argument) is called the Gamma function, denoted $\Gamma(z)$, defined for all complex numbers z except the non-positive integers, and given when the real part of z is positive by

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

Its relation to the factorials is that for any natural number n

$$n! = \Gamma(n+1).$$

Euler's original formula for the Gamma function was

$$\Gamma(z) = \lim_{n\to\infty} \frac{n^z n!}{\prod_{k=0}^{n}(z+k)}.$$

An alternative notation, originally introduced by Gauss, is sometimes used. The **Pi function**, denoted $\Pi(z)$ for real numbers z no less than 0, is defined by

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

In terms of the Gamma function it is

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

It truly extends the factorial in that

$$\Pi(n) = n! \text{ for } n \in \mathbf{N}\,.$$

In addition to this, the Pi function satisfies the same recurrence as factorials do, but at every complex value z where it is defined

$$\Pi(z) = z\Pi(z-1)\,.$$

In fact, this is no longer a recurrence relation but a functional equation. Expressed in terms of the Gamma function this functional equation takes the form

$$\Gamma(n+1) = n\Gamma(n)\,.$$

Since the factorial is extended by the Pi function, for every complex value z where it is defined, we can write:

$$z! = \Pi(z)$$

The values of these functions at half-integer values is therefore determined by a single one of them; one has

$$\Gamma\left(\frac{1}{2}\right) = \left(-\frac{1}{2}\right)! = \Pi\left(-\frac{1}{2}\right) = \sqrt{\pi},$$

from which it follows that for $n \in \mathbf{N}$,

$$\Gamma\left(\frac{1}{2}+n\right) = \left(-\frac{1}{2}+n\right)! = \Pi\left(-\frac{1}{2}+n\right) = \sqrt{\pi}\prod_{k=1}^{n}\frac{2k-1}{2} = \frac{(2n)!}{4^n n!}\sqrt{\pi} = \frac{(2n-1)!}{2^{2n-1}(n-1)!}\sqrt{\pi}.$$

For example,

$$\Gamma(4.5) = 3.5! = \Pi(3.5) = \frac{1}{2}\cdot\frac{3}{2}\cdot\frac{5}{2}\cdot\frac{7}{2}\sqrt{\pi} = \frac{8!}{4^4 4!}\sqrt{\pi} = \frac{7!}{2^7 3!}\sqrt{\pi} = \frac{105}{16}\sqrt{\pi} \approx 11.63.$$

It also follows that for $n \in \mathbf{N}$,

$$\Gamma\left(\frac{1}{2}-n\right) = \left(-\frac{1}{2}-n\right)! = \Pi\left(-\frac{1}{2}-n\right) = \sqrt{\pi}\prod_{k=1}^{n}\frac{2}{1-2k} = \frac{(-4)^n n!}{(2n)!}\sqrt{\pi}.$$

For example,

$$\Gamma(-2.5) = (-3.5)! = \Pi(-3.5) = \frac{2}{-1}\cdot\frac{2}{-3}\cdot\frac{2}{-5}\sqrt{\pi} = \frac{(-4)^3 3!}{6!}\sqrt{\pi} = -\frac{8}{15}\sqrt{\pi} \approx -0.9453.$$

The Pi function is certainly not the only way to extend factorials to a function defined at almost all complex values, and not even the only one that is analytic wherever it is defined. Nonetheless it is usually considered the most natural way to extend the values of the factorials to a complex function. For instance, the Bohr–Mollerup theorem states that the Gamma function is the only function that takes the value 1 at 1, satisfies the functional equation $\Gamma(n+1) = n\Gamma(n)$, is meromorphic on the complex numbers, and is log-convex on the positive real axis. A similar statement holds for the Pi function as well, using the $\Pi(n) = n\Pi(n-1)$ functional equation.

However, there exist complex functions that are probably simpler in the sense of analytic function theory and which interpolate the factorial values. For example, Hadamard's 'Gamma'-function (Hadamard 1894) which, unlike the Gamma function, is an entire function.[11]

Euler also developed a convergent product approximation for the non-integer factorials, which can be seen to be equivalent to the formula for the Gamma function above:

$$\begin{aligned} n! = \Pi(n) &= \prod_{k=1}^{\infty}\left(\frac{k+1}{k}\right)^n \frac{k}{n+k} \\ &= \left[\left(\frac{2}{1}\right)^n \frac{1}{n+1}\right]\left[\left(\frac{3}{2}\right)^n \frac{2}{n+2}\right]\left[\left(\frac{4}{3}\right)^n \frac{3}{n+3}\right]\cdots. \end{aligned}$$

However, this formula does not provide a practical means of computing the Pi or Gamma function, as its rate of convergence is slow.

12.7.2 Applications of the Gamma function

The volume of an n-dimensional hypersphere of radius R is

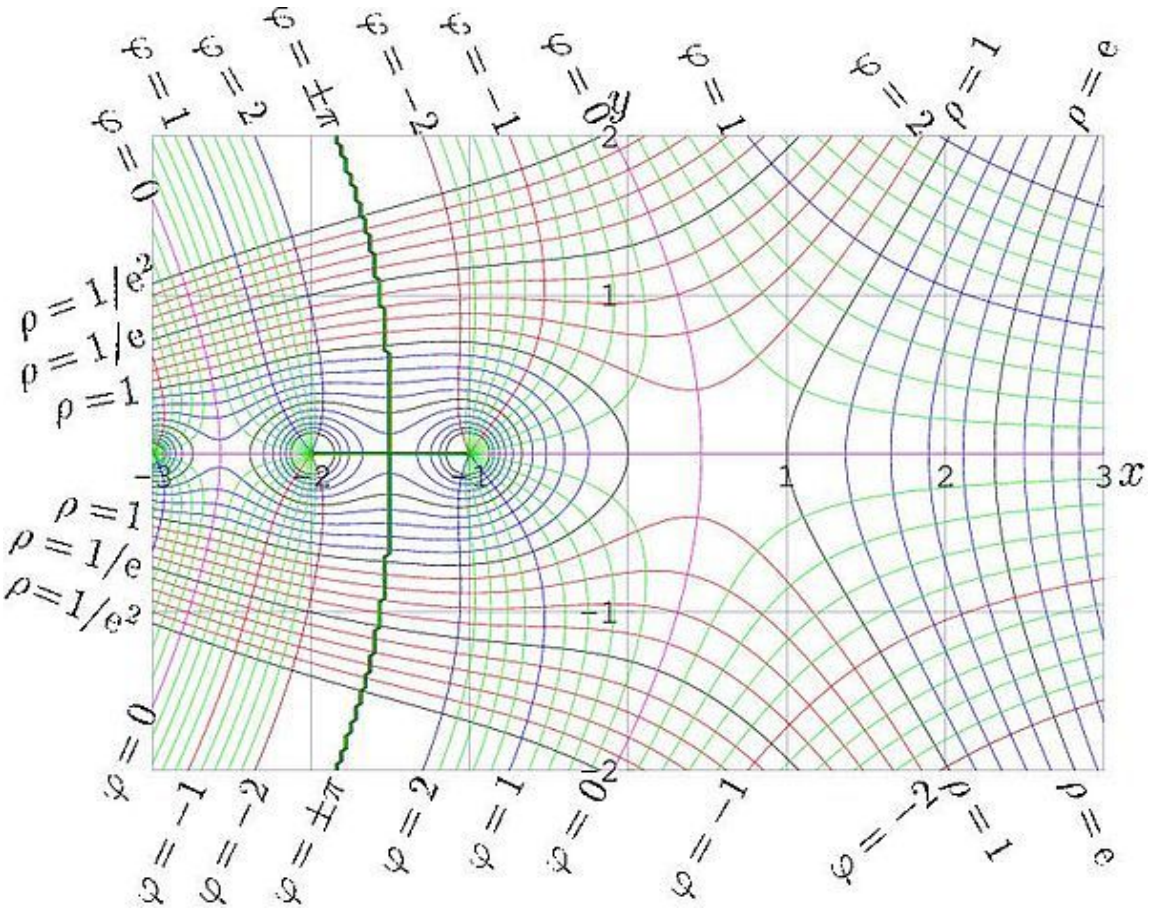
$$V_n = \frac{\pi^{n/2}}{\Gamma((n/2)+1)}R^n.$$

12.7.3 Factorial at the complex plane

Representation through the Gamma-function allows evaluation of factorial of complex argument. Equilines of amplitude and phase of factorial are shown in figure. Let $f = \rho\exp(\mathrm{i}\varphi) = (x+\mathrm{i}y)! = \Gamma(x+\mathrm{i}y+1)$. Several levels of constant modulus (amplitude) $\rho =$ const and constant phase $\varphi =$ const are shown. The grid covers range $-3 \le x \le 3$, $-2 \le y \le 2$ with unit step. The scratched line shows the level $\varphi = \pm\pi$.

Thin lines show intermediate levels of constant modulus and constant phase. At poles $x+\mathrm{i}y \in$ (negative integers) , phase and amplitude are not defined. Equilines are dense in vicinity of singularities along negative integer values of the argument.

For $|z| < 1$, the Taylor expansions can be used:



Amplitude and phase of factorial of complex argument

$$z! = \sum_{n=0}^{\infty} g_n z^n.$$

The first coefficients of this expansion are

where γ is the Euler constant and ζ is the Riemann zeta function. Computer algebra systems such as SageMath can generate many terms of this expansion.

12.7.4 Approximations of factorial

For the large values of the argument, factorial can be approximated through the integral of the digamma function, using the continued fraction representation. This approach is due to T. J. Stieltjes (1894). Writing $z!$ = exp(P(z)) where P(z) is

$$P(z) = p(z) + \log(2\pi)/2 - z + \left(z + \frac{1}{2}\right)\log(z),$$

Stieltjes gave a continued fraction for p(z)

$$p(z) = \cfrac{a_0}{z + \cfrac{a_1}{z + \cfrac{a_2}{z + \cfrac{a_3}{z + \ddots}}}}$$

The first few coefficients a_n are[12]

There is a misconception that $\log(z!) = P(z)$ or $\log(\Gamma(z+1)) = P(z)$ for any complex $z \neq 0$. Indeed, the relation through the logarithm is valid only for specific range of values of z in vicinity of the real axis, while $|\Im(\Gamma(z+1))| < \pi$. The larger is the real part of the argument, the smaller should be the imaginary part. However, the inverse relation, $z!$ = exp($P(z)$), is valid for the whole complex plane apart from zero. The convergence is poor in vicinity of the negative part of the real axis. (It is difficult to have good convergence of any approximation in vicinity of the singularities). While $|\Im(z)| > 2$ or $\Re(z) > 2$, the 6 coefficients above are sufficient for the evaluation of the factorial with the complex<double> precision. For higher precision more coefficients can be computed by a rational QD-scheme (H. Rutishauser's QD algorithm).[13]

12.7.5 Non-extendability to negative integers

The relation $n! = n \times (n-1)!$ allows one to compute the factorial for an integer given the factorial for a *smaller* integer. The relation can be inverted so that one can compute the factorial for an integer given the factorial for a *larger* integer:

$$(n-1)! = \frac{n!}{n}.$$

Note, however, that this recursion does not permit us to compute the factorial of a negative integer; use of the formula to compute $(-1)!$ would require a division by zero, and thus blocks us from computing a factorial value for every negative integer. (Similarly, the Gamma function is not defined for non-positive integers, though it is defined for all other complex numbers.)

12.8 Factorial-like products and functions

There are several other integer sequences similar to the factorial that are used in mathematics:

12.8.1 Double factorial

Main article: Double factorial

The product of all the odd integers up to some odd positive integer n is called the **double factorial** of n, and denoted by $n!!$.[14] That is,

$$(2k-1)!! = \prod_{i=1}^{k}(2i-1) = \frac{(2k)!}{2^k k!} = \frac{{}_{2k}P_k}{2^k} = \frac{(2k)^{\underline{k}}}{2^k}.$$

For example, 9!! = 1 × 3 × 5 × 7 × 9 = 945.

The sequence of double factorials for n = 1, 3, 5, 7, ... starts as

1, 3, 15, 105, 945, 10395, 135135, ... (sequence A001147 in the OEIS)

Double factorial notation may be used to simplify the expression of certain trigonometric integrals,[15] to provide an expression for the values of the Gamma function at half-integer arguments and the volume of hyperspheres,[16] and to solve many counting problems in combinatorics including counting binary trees with labeled leaves and perfect matchings in complete graphs.[14][17]

12.8.2 Multifactorials

A common related notation is to use multiple exclamation points to denote a **multifactorial**, the product of integers in steps of two ($n!!$), three ($n!!!$), or more. The double factorial is the most commonly used variant, but one can similarly define the triple factorial ($n!!!$) and so on. One can define the k-th factorial, denoted by $n!^{(k)}$, recursively for positive integers as

$$n!^{(k)} = \begin{cases} n & \text{if} 0 < n \leq k \\ n((n-k)!^{(k)}) & \text{if} n > k \end{cases}$$

though see the alternative definition below. In addition, similarly to 0! = 1!/1 = 1, one can define:

$$n!^{(k)} = 1 \text{ if } -k < n \leq 0$$

Some mathematicians have suggested an alternative notation of $n!_2$ for the double factorial and similarly $n!_k$ for other multifactorials, but this has not come into general use.

In the same way that $n!$ is not defined for negative integers, and $n!!$ is not defined for negative even integers, $n!^{(k)}$ is not defined for negative integers divisible by k .

Alternative extension of the multifactorial

Alternatively, the multifactorial $z!^{(k)}$ can be extended to most real and complex numbers z by noting that when z is one more than a positive multiple of k then

$$z!^{(k)} = z(z-k)\cdots(k+1) = k^{(z-1)/k}\left(\frac{z}{k}\right)\left(\frac{z-k}{k}\right)\cdots\left(\frac{k+1}{k}\right) = k^{(z-1)/k}\frac{\Gamma\left(\frac{z}{k}+1\right)}{\Gamma\left(\frac{1}{k}+1\right)}.$$

This last expression is defined much more broadly than the original; with this definition, $z!^{(k)}$ is defined for all complex numbers except the negative real numbers divisible by k. This definition is consistent with the earlier definition only for those integers z satisfying $z \equiv 1 \bmod k$.

In addition to extending $z!^{(k)}$ to most complex numbers z, this definition has the feature of working for all positive real values of k. Furthermore, when $k = 1$, this definition is mathematically equivalent to the $\Pi(z)$ function, described above. Also, when $k = 2$, this definition is mathematically equivalent to the alternative extension of the double factorial.

12.8.3 Primorial

The primorial (sequence A002110 in the OEIS) is similar to the factorial, but with the product taken only over the prime numbers.

12.8.4 Quadruple factorial

Contrary to consistency with the above definition of the term *double factorial*, the quadruple factorial is not the multifactorial $n!^{(4)}$; it is a much larger number given by $(2n)!/n!$, starting as

1, 2, 12, 120, 1680, 30240, 665280, ... (sequence A001813 in the OEIS).

It is also equal to

$$2^n \frac{(2n)!}{n!2^n} = 2^n \frac{(2 \cdot 4 \cdots 2n)(1 \cdot 3 \cdots (2n-1))}{2 \cdot 4 \cdots 2n}$$

$$= (1 \cdot 2) \cdot (3 \cdot 2) \cdots ((2n-1) \cdot 2) = (4n-2)!^{(4)}.$$

12.8.5 Superfactorial

Main article: Large numbers
"N$" redirects here. For the currency, see Namibian dollar.

Neil Sloane and Simon Plouffe defined a **superfactorial** in The Encyclopedia of Integer Sequences (Academic Press, 1995) to be the product of the first n factorials. So the superfactorial of 4 is

$$\mathrm{sf}(4) = 1! \times 2! \times 3! \times 4! = 288.$$

In general

$$\mathrm{sf}(n) = \prod_{k=1}^{n} k! = \prod_{k=1}^{n} k^{n-k+1} = 1^n \cdot 2^{n-1} \cdot 3^{n-2} \cdots (n-1)^2 \cdot n^1.$$

Equivalently, the superfactorial is given by the formula

$$\mathrm{sf}(n) = \prod_{0 \le i < j \le n} (j-i)$$

which is the determinant of a Vandermonde matrix.

The sequence of superfactorials starts (from $n = 0$) as

1, 1, 2, 12, 288, 34560, 24883200, 125411328000, ... (sequence A000178 in the OEIS)

Alternative definition

See also: Tetration

Clifford Pickover in his 1995 book *Keys to Infinity* used a new notation, *n$*, to define the superfactorial

$$n\$ \equiv \underbrace{n!^{n!^{\cdot^{\cdot^{\cdot^{n!}}}}}}_{n!},$$

or as,

$$n\$ = n![4]n!$$

where the [4] notation denotes the hyper4 operator, or using Knuth's up-arrow notation,

$$n\$ = (n!) \uparrow\uparrow (n!).$$

This sequence of superfactorials starts:

$1\$ = 1$

$2\$ = 2^2 = 4$

$3\$ = 6[4]6 = {}^{6}6 = 6^{6^{6^{6^{6}}}}$.

Here, as is usual for compound exponentiation, the grouping is understood to be from right to left:

$a^{b^c} = a^{(b^c)}$.

12.8.6 Hyperfactorial

Occasionally the **hyperfactorial** of n is considered. It is written as $H(n)$ and defined by

$$H(n) = \prod_{k=1}^{n} k^k = 1^1 \cdot 2^2 \cdot 3^3 \cdots (n-1)^{n-1} \cdot n^n.$$

For n = 1, 2, 3, 4, ... the values $H(n)$ are 1, 4, 108, 27648,... (sequence A002109 in the OEIS).

The asymptotic growth rate is

$$H(n) \sim An^{(6n^2+6n+1)/12} e^{-n^2/4}$$

where A = 1.2824... is the Glaisher–Kinkelin constant.[18] $H(14)$ = 1.8474...×10^{99} is already almost equal to a googol, and $H(15)$ = 8.0896...×10^{116} is almost of the same magnitude as the Shannon number, the theoretical number of possible chess games. Compared to the Pickover definition of the superfactorial, the hyperfactorial grows relatively slowly.

The hyperfactorial function can be generalized to complex numbers in a similar way as the factorial function. The resulting function is called the K-function.

12.9 See also

- Alternating factorial
- Bhargava factorial
- Digamma function
- Exponential factorial
- Factorial number system
- Factorial prime
- Factorion
- List of factorial and binomial topics
- Pochhammer symbol, which gives the falling or rising factorial
- Subfactorial
- Trailing zeros of factorial
- Triangular number, the additive analogue of factorial

12.10 Notes

[1] Ronald L. Graham, Donald E. Knuth, Oren Patashnik (1988) *Concrete Mathematics*, Addison-Wesley, Reading MA. ISBN 0-201-14236-8, p. 111

[2] N. L. Biggs, *The roots of combinatorics*, Historia Math. 6 (1979) 109–136

[3] Stedman, Fabian (1677), *Campanalogia*, London, pp. 6–9 The publisher is given as "W.S." who may have been William Smith, possibly acting as agent for the Society of College Youths, to which society the "Dedicatory" is addressed.

[4] Stedman 1677, p. 8.

[5] Higgins, Peter (2008), *Number Story: From Counting to Cryptography*, New York: Copernicus, p. 12, ISBN 978-1-84800-000-1 says Krempe though.

[6] http://ocw.mit.edu/courses/mathematics/18-01-single-variable-calculus-fall-2006/lecture-notes/lec4.pdf

[7] Guy, Richard K. (2004), "E24 Irrationality sequences", *Unsolved problems in number theory* (3rd ed.), Springer-Verlag, p. 346, ISBN 0-387-20860-7, Zbl 1058.11001.

[8] GNU MP software manual, "Factorial Algorithm" (retrieved 22 January 2013).

[9] Peter Borwein. "On the Complexity of Calculating Factorials". *Journal of Algorithms* 6, 376–380 (1985)

[10] Peter Luschny, *Fast-Factorial-Functions: The Homepage of Factorial Algorithms*.

[11] Peter Luschny, *Hadamard versus Euler - Who found the better Gamma function?*.

[12] Digital Library of Mathematical Functions, http://dlmf.nist.gov/5.10

[13] Peter Luschny, *On Stieltjes' Continued Fraction for the Gamma Function.*.

[14] Callan, David (2009), *A combinatorial survey of identities for the double factorial*, arXiv:0906.1317.

[15] Meserve, B. E. (1948), "Classroom Notes: Double Factorials", *The American Mathematical Monthly*, **55** (7): 425–426, doi:10.2307/2306136, MR 1527019

[16] Mezey, Paul G. (2009), "Some dimension problems in molecular databases", *Journal of Mathematical Chemistry*, **45** (1): 1–6, doi:10.1007/s10910-008-9365-8.

[17] Dale, M. R. T.; Moon, J. W. (1993), "The permuted analogues of three Catalan sets", *Journal of Statistical Planning and Inference*, **34** (1): 75–87, doi:10.1016/0378-3758(93)90035-5, MR 1209991.

[18] Weisstein, Eric W. "Glaisher–Kinkelin Constant". *MathWorld*.

12.11 References

- Hadamard, M. J. (1894), *Sur L'Expression Du Produit $1 \cdot 2 \cdot 3 \cdot \cdots \cdot (n-1)$ Par Une Fonction Entière* (PDF) (in French), *OEuvres de Jacques Hadamard*, Centre National de la Recherche Scientifiques, Paris, 1968
- Ramanujan, Srinivasa (1988), *The lost notebook and other unpublished papers*, Springer Berlin, p. 339, ISBN 3-540-18726-X

12.12 External links

- Hazewinkel, Michiel, ed. (2001), "Factorial", *Encyclopedia of Mathematics*, Springer, ISBN 978-1-55608-010-4
- Weisstein, Eric W. "Factorial". *MathWorld*.
- Factorial at PlanetMath.org.

12.13 Text and image sources, contributors, and licenses

12.13.1 Text

- **Number theory** *Source:* https://en.wikipedia.org/wiki/Number_theory?oldid=741842027 *Contributors:* AxelBoldt, Derek Ross, Calypso, Brion VIBBER, Mav, Zundark, Tarquin, XJaM, Michael Shulman, Christian List, Miguel~enwiki, Twilsonb, Stevertigo, Teun-Spaans, Michael Hardy, Chris-martin, Ixfd64, Dcljr, GTBacchus, Delirium, Minesweeper, Ams80, Ahoerstemeier, Nikai, Rotem Dan, Iorsh, [212], Schneelocke, Hashar, Markb, Revolver, Charles Matthews, Timwi, Dcoetzee, Dysprosia, Jitse Niesen, The Anomebot, Xiaodai~enwiki, Tpbradbury, Sabbut, Jose Ramos, Qianfeng, Finlay McWalter, Bearcat, Robbot, Jaredwf, Fredrik, Romanm, Lowellian, Gandalf61, Fuelbottle, Lupo, PrimeFan, Jleedev, Ancheta Wis, Giftlite, Recentchanges, Pretzelpaws, Lethe, Fropuff, Everyking, Gubbubu, Gadfium, LiDaobing, Antandrus, Beland, Robert Brockway, Bob.v.R, Khaosworks, Karol Langner, APH, Stefan64, Tsemii, Xmlizer, Rich Farmbrough, Paul August, Bender235, ESkog, Ben Standeven, Appleboy, Tompw, El C, Jpgordon, Mysteronald, Maurreen, Hagerman, Pearle, Storm Rider, Msh210, Alansohn, Arthena, Neonumbers, Diego Moya, Velella, SidP, Evil Monkey, CloudNine, Dirac1933, Igorpak, HenryLi, Oleg Alexandrov, Mcsee, Richard Arthur Norton (1958-), Linas, Unixer, Jimbryho, Ruud Koot, Wikiklrsc, GregorB, Dionyziz, Graham87, Dpv, Chenxlee, Mayumashu, R.e.b., The wub, DoubleBlue, Sango123, Vuong Ngan Ha, FlaBot, Mathbot, Malhonen, Scythe33, Haonhien, Chobot, Digitalme, YurikBot, Wavelength, Lexi Marie, Lenthe, JabberWok, Stassats, Joth, Welsh, Ino5hiro, Ms2ger, Kompik, Rwxrwxrwx, Lt-wiki-bot, Arthur Rubin, Willtron, GrinBot~enwiki, That Guy, From That Show!, Marquez~enwiki, Sardanaphalus, SmackBot, Mmernex, Rebollo fr~enwiki, McGeddon, Jagged 85, Rouenpucelle, AustinKnight, Hmains, Anastasios~enwiki, Chris the speller, Bluebot, ChuckHG, PrimeHunter, MalafayaBot, Spellchecker, DHN-bot~enwiki, Colonies Chris, Sct72, Modest Genius, Ianmacm, Lwassink, Bidabadi~enwiki, Andrew Dalby, SashatoBot, Lambiam, ArglebargleIV, UberCryxic, Evildictaitor, Tdudkowski, Jim.belk, Gary13579, Stwalkerster, Childzy, Mets501, Mathsci, Kripkenstein, Joseph Solis in Australia, LDH, Zero sharp, Igoldste, Courcelles, Tawkerbot2, Stifynsemons, CRGreathouse, CmdrObot, Mikeliuk, Sdorrance, Chrisahn, Ken Gallager, Myasuda, Kronecker, Doctormatt, Gogo Dodo, Karl-H, Thijs!bot, Epbr123, Atmd, O, Bhowmickr, Marek69, Woody, RobHar, Sherbrooke, Escarbot, Luna Santin, Guy Macon, Marquess, Qwerty Binary, Normanzhang, JAnDbot, MER-C, Hut 8.5, Magioladitis, JamesBWatson, Wlod, Usien6, Kroposky, Bubba hotep, Systemlover, NJR ZA, Kope, DerHexer, Khalid Mahmood, TheRanger, Vandermude, Glrx, R'n'B, Nono64, J.delanoy, Trusilver, Numbo3, Maurice Carbonaro, Smeira, Tarotcards, Krishnachandranvn, Chiswick Chap, Policron, CompuChip, Milogardner, CombFan, Treisijs, VolkovBot, JohnBlackburne, Philip Trueman, TXiKiBoT, Oshwah, Hotjava, Anna Lincoln, Plclark, Magmi, Pleaseee, Broadbot, Kmhkmh, Blurpeace, Joseph A. Spadaro, Symane, GirasoleDE, SieBot, Calliopejen1, GrooveDog, Iames, KoenDelaere, S2000magician, Amahoney, ClueBot, MeowMeow163, Justin W Smith, John.D.Ward, Paulsavala, Drmies, Mild Bill Hiccup, DragonBot, PixelBot, Bercant, Cenarium, Arjayay, Jotterbot, H.Marxen, Crowsnest, XLinkBot, Marc van Leeuwen, Killthesteel, Ajcheema, Virginia-American, Addbot, Betterusername, Ronhjones, SpillingBot, Download, Uncia, Ausefi1900, Feketekave, Bluebusy, TeH nOmInAtOr, Luckas-bot, Yobot, MinorProphet, Zhouhaigang, Xylune, AnomieBOT, Rubinbot, Royote, Materialscientist, Citation bot, Maxis ftw, ArthurBot, Gypsydave5, Xqbot, Smk65536, Gilo1969, Anne Bauval, Tyrol5, Gap9551, Vaywatch, GrouchoBot, Omnipaedista, Charvest, Raulshc, Aprogrammer, Lexy-lou, Bekus, DanRawsthorne, FrescoBot, Imbalzanog, LucienBOT, Tobby72, BrideOfKripkenstein, Machine Elf 1735, Pinethicket, MarcelB612, Artorio, Garald, Jauhienij, Tim1357, FoxBot, Trappist the monk, Dinamik-bot, Vrenator, Reach Out to the Truth, Korepin, KurtSchwitters, Ccrazymann, EmausBot, Lollipopweare, Fly by Night, EleferenBot, Jmencisom, Wikipelli, Bethnim, ZéroBot, Knight1993, Usability, Midas02, D.Lazard, Ian Rastall, Bobdylan1234567, Donner60, Nobrook, Unga Khan, Anita5192, ClueBot NG, Wcherowi, Satellizer, Helpful Pixie Bot, PhnomPencil, JohnChrysostom, AvocatoBot, Solomon7968, Xosé Antonio, Brad7777, Weierstrass1, Ducknish, JYBot, Dexbot, Deltahedron, Apdenum, Jamesx12345, Telfordbuck, Vanamonde93, Prem nath singh, Jodosma, SakeUPenn, Syferion, YiFeiBot, Programmingcaffeine, ProfessorMoriarty1811, Canto55, IagoQnsi, Loraof, SoSivr, GeneralizationsAreBad, J Steed Huang, KasparBot, Sweepy, Qwertydeeznutz, Baking Soda, Marvellous Spider-Man and Anonymous: 303
- **Diophantine equation** *Source:* https://en.wikipedia.org/wiki/Diophantine_equation?oldid=738523367 *Contributors:* AxelBoldt, Magnus~enwiki, XJaM, Heron, Michael Hardy, Gnomon42, Cyde, Julesd, Ruhrjung, Charles Matthews, Timwi, Dysprosia, Jitse Niesen, Robbot, Fredrik, MathMartin, Davidl9999, Robinh, Tea2min, Matthew Stannard, Giftlite, Zigger, Everyking, Mckaysalisbury, Bobblewik, Vivero~enwiki, Gauss, Icairns, Dmr2, Jelammers, Chalst, Billymac00, La goutte de pluie, Haham hanuka, HasharBot~enwiki, Msh210, Burn, Hu, Dirac1933, Oleg Alexandrov, Linas, Ajb, M412k, Chenxlee, Staecker, FlaBot, RobertG, Chobot, Siddhant, YurikBot, Wavelength, Gaius Cornelius, Wimt, Scope creep, Reyk, Pred, Matikkapoika~enwiki, Dash77, Jsnx, SmackBot, InverseHypercube, Jagged 85, Flamarande, Richfife, Oli Filth, Tree Biting Conspiracy, Hooriaj, Nbarth, DHN-bot~enwiki, Tsca.bot, Cícero, Wen D House, Matt Whyndham, Bidabadi~enwiki, Lambiam, Don't fear the reaper, Asyndeton, Seqsea, Az1568, Albregis, CRGreathouse, CmdrObot, Nunquam Dormio, Myasuda, Yrodro, Fl, Sam Staton, Kazubon~enwiki, M a s, Chrislk02, Thijs!bot, LaGrange, Marek69, QuiteUnusual, Magioladitis, Vanish2, David Eppstein, Ricardv46, Clokr, Gargiaparna, Indeed123, Krishnachandranvn, STBotD, DorganBot, TXiKiBoT, Hqb, Nxavar, Wtt, EnJx, Drschawrz, Yintan, ClueBot, Justin W Smith, Plastikspork, ChandlerMapBot, Bender2k14, J.Gowers, Nilaish, Hatsoff, Rabbit67890, Addbot, Math1353, Ronhjones, Shirtwaist, LaaknorBot, AnnaFrance, Nfogravity, Howler200, Legobot, Luckas-bot, Yobot, AnomieBOT, Materialscientist, Onesius, Xqbot, Doulos Christos, Tobby72, Projectxanadu, BenzolBot, Kiefer.Wolfowitz, SkinnyPrude, LittleWink, Rohitphy, Logical Gentleman, Marksmith55, Archaicmath, Duoduoduo, EmausBot, KHamsun, Dcirovic, 1curtisom, Fred Gandt, D.Lazard, Wayne Slam, Chewings72, ClueBot NG, MelbourneStar, Baseball Watcher, Rezabot, Helpful Pixie Bot, Mokhtari34, Tekwani, GregorDS, BattyBot, MahdiBot, ChrisGualtieri, Deltahedron, BeaumontTaz, Jochen Burghardt, Brirush, Rrmath28, Trompedo, Asn22102, Tudor987, Ghulamabbass, Arpan Mathur, MNSMUPhysicist, NyanCatGirl, Loraof, KasparBot and Anonymous: 122
- **Modular arithmetic** *Source:* https://en.wikipedia.org/wiki/Modular_arithmetic?oldid=737184979 *Contributors:* AxelBoldt, CYD, Bryan Derksen, Zundark, The Anome, Toby Bartels, Imran, Mjb, Montrealais, Patrick, Michael Hardy, Wshun, Shellreef, Ixfd64, Takuya-Murata, Ahoerstemeier, Darkwind, Kragen, Julesd, Ciphergoth, Revolver, Charles Matthews, Timwi, Far neil, Dcoetzee, Joshuabowman, Dysprosia, Hyacinth, SirJective, Onebyone, Fredrik, Schutz, Gandalf61, Rholton, JackofOz, ElBenevolente, Tea2min, Connelly, Giftlite, Lunkwill, Bfinn, Fropuff, Peruvianllama, KuniShiro~enwiki, Ikitakoja, Chinasaur, Jason Quinn, Antandrus, Bob.v.R, Profvk, IOH, Mike Rosoft, Pyrop, Discospinster, Guanabot, Too Old, Notinasnaid, Paul August, Bender235, Sgeo, Superninja, Spoon!, Hurricane111, Jung dalglish, Boredzo, Tgr, Obradovic Goran, Haham hanuka, Msh210, Alexander Guy, Kmill, BRW, Count Iblis, Dirac1933, Bsadowski1, SteinbDJ, Oleg Alexandrov, Linas, LOL, StradivariusTV, Guardian of Light, Akruppa~enwiki, Graham87, Josh Parris, AySz88, VKokielov, Alexb@cut-the-knot.com, Margosbot~enwiki, Maxal, Mister Farkas, Siddhant, YurikBot, Wavelength, Ec-, Laurentius, Eraserhead1, Hairy Dude, Alpt, Van der Hoorn, Rsrikanth05, Lsdan, Trovatore, Dwragg, Sekelsenmat, EverettColdwell, Mikeblas, Nicholas Perkins, Ronyclau, Googl, Arthur Rubin, JLaTondre, Jsnx, SmackBot, Imz, BiT, Chris the speller, Ciacchi, Silly rabbit, Nbarth, AeroSpace, Gatherton, Nixeagle, Lesnail, Maksim-bot, CorbinSimpson, UU, Kazov, Akshaysrinivasan, Luís Felipe Braga, Vina-iwbot~enwiki, Lambiam, MvH, Makyen, BenRayfield, Waggers, Onionmon, Johnruble, Jerrybtaylor, JSoules, Powerslide, Vaughan

Pratt, CRGreathouse, Citrus538, Green caterpillar, GeneChase, Marek69, Pragmagician, RobHar, WinBot, Fireice, Rbb l181, JAnDbot, Asmeurer, Albmont, PeterStJohn, JoergenB, R'n'B, Touisiau, NewEnglandYankee, In Transit, Policron, Lukax, Diggory Hardy, Rémih, King Lopez, VolkovBot, Pleasantville, LokiClock, Rponamgi, Paulburnett, Francesco.fracaro, Anonymous Dissident, Wiae, Quindraco, Rjgodoy, Alexanderhowell, Dwandelt, Jerryobject, Frillystevens, Iain99, JackSchmidt, OKBot, Laurentseries, Anchor Link Bot, Jfromcanada, Dabomb87, ClueBot, The Thing That Should Not Be, DragonBot, He7d3r, Bender2k14, Peter.C, Jtle515, Johnuniq, Bill Arden, Virginia-American, Dsimic, Wyatt915, Addbot, Zahd, Aboctok, Msirbola, LinkFA-Bot, Zorrobot, Jarble, خالد حسني, Jordsan, KamikazeBot, AnomieBOT, Angry bee, Erel Segal, Rubinbot, 9258fahsflkh917fas, FreeRangeFrog, Xqbot, Acebulf, Nitrxgen, GrouchoBot, Adaniels85, Drdpw, Sonoluminesence, CopineDeLaMer, Mfwitten, DrilBot, Martinvl, Night Jaguar, Toolnut, Numericana, Nickyus, EmausBot, Tttfffkkk, Quondum, D.Lazard, DASHBotAV, Rmashhadi, Anita5192, ClueBot NG, Wikigold96, David.xia, Versatranitsonlywaytofly, Kranthi117, Manoharsingh23, Widr, MerlIwBot, Helpful Pixie Bot, Krearick, Dexbot, FoCuSandLeArN, Ictfac, Buspirtraz, M.ollivander, François Robere, Hamidrjahanjou, EvergreenFir, Barymar, Dillon128, Callmenorm, Dough34, LarsHugo, Irvin Miller, Zeuslightning125, Garfield Garfield, Dyott, Dorivaldo de C. M. dos Santos, JrRong, Warren Leywon, Jasonalaw, Mindotaur, NoToleranceForIntolerance and Anonymous: 261

- **Burnside's lemma** *Source:* https://en.wikipedia.org/wiki/Burnside'{}s_lemma?oldid=736348982 *Contributors:* AxelBoldt, Zundark, The Anome, Tarquin, Michael Hardy, Firebirth, Karada, Hawthorn, Charles Matthews, Giftlite, Fropuff, Waltpohl, Profvk, Gauge, Crisófilax, Oleg Alexandrov, Joriki, Simetrical, Stolee, Julien Tuerlinckx, Kesla, Salix alba, HappyCamper, Hillman, Michael Slone, Crasshopper, FF2010, Arthur Rubin, Charles^, SmackBot, Silly rabbit, Davcrav, Dreadstar, CRGreathouse, Zahlentheorie, OckhamTheFox, Headbomb, Dricherby, Sullivan.t.j, Sudleyplace, VolkovBot, Solitonic, Kmhkmh, Arcfrk, Hangman154, JackSchmidt, Justin W Smith, Addbot, Neumann7, Charleswallingford, Calle, AnomieBOT, Sjin.75~enwiki, Sz-iwbot, Kiefer.Wolfowitz, John of Reading, WikitanvirBot, Wishcow, ZéroBot, ClueBot NG, Joel B. Lewis, Episcophagus, I am One of Many, Some1Redirects4You, StenchWeasel and Anonymous: 26
- **Gaussian elimination** *Source:* https://en.wikipedia.org/wiki/Gaussian_elimination?oldid=740603892 *Contributors:* AxelBoldt, Eloquence, Zundark, Tarquin, Ap, AdamRetchless, Michael Hardy, Dominus, Nixdorf, Tongpoo, Zeno Gantner, TakuyaMurata, Dori, Eric119, Evercat, Dcoetzee, Joshuabowman, Dysprosia, Jitse Niesen, Fibonacci, David.Monniaux, Robbot, Fredrik, Mattblack82, Romanm, Lowellian, Babbage, MathMartin, Giftlite, BenFrantzDale, Dissident, FunnyMan3595, Kainaw, Raylu, Mh, Mormegil, Guanabot, BBUCommander, Paul August, Gauge, Bobo192, Smalljim, 3mta3, MPerel, Alansohn, Liao, ABCD, Galaxiaad, Oleg Alexandrov, LOL, Macaddct1984, Gisling, Qwertyus, Seyon, Rjwilmsi, Ohanian, AySz88, RexNL, Quuxplusone, Kri, Chobot, Jersey Devil, YurikBot, Tong~enwiki, Proidiot, Gareth Jones, Nothing1212, Crasshopper, Mike1024, HereToHelp, Cjfsyntropy, MagneticFlux, Aaron Will, SmackBot, InverseHypercube, Andres Agudelo, APW, Jfgrcar, Hmains, Nbarth, Fiziker, Lhf, Eliyak, Ni fr, Rohan Lean, Jim.belk, Applejuicefool, Denshade, CBM, Tac-Tics, Lee, Braveorca, Headbomb, Yettie0711, Urdutext, Salgueiro~enwiki, LiraNuna, JAnDbot, Thenub314, Richard Giuly, David Eppstein, ANONYMOUS COWARD0xC0DE, Glrx, Mathemaduenn~enwiki, K.menin, Fylwind, STBotD, JohnBlackburne, LokiClock, Constant314, Anonymous Dissident, Ejtttje, Wolfrock, Daviddoria, PericlesofAthens, SieBot, BoxTurtle, BotMultichill, Rinconsoleao, ClueBot, Adrianwn, JP.Martin-Flatin, Ladne2, He7d3r, Bender2k14, 3ICE, Flange45, Marc van Leeuwen, Marc CAT, D.M. from Ukraine, Addbot, Maclary, Matěj Grabovský, Arbitrarily0, Blablablob, Luckas-bot, Yobot, AnomieBOT, Cd1207, Rychlik, Sz-iwbot, Materialscientist, Bouron, Citation bot, Obersachsebot, Xqbot, Raffamaiden, 33rogers, FrescoBot, Rotideypoc41352, VI, HJ Mitchell, AstaBOTh15, I dream of horses, Kiefer.Wolfowitz, Tom.Reding, RedBot, MastiBot, Kanglipedia, Gryllida, TobeBot, NortyNort, Minimac, RjwilmsiBot, Thiridaz, Steve.wellons, Slawekb, D.Lazard, Tecknoize, Donner60, Liuthar, ClueBot NG, Ooz dot ie, Bakrnl, Snark1994, Rurik the Varangian, Mt daydream, ServiceAT, Robert the Devil, Dexbot, Mark L MacDonald, Cerabot~enwiki, MindAfterMath, Isarra (HG), L33torizor, B14709, Chip Wildon Forster, NeapleBerlina, Jeffrey Bosboom, Ben476, Mohit.del94, Atomicboyrt, Delcroip, Muyan93, Nkgsv, SPIKE SPIKE BAD, Akemdh, Latif0x0, Blurbleflurth and Anonymous: 205
- **Matrix exponential** *Source:* https://en.wikipedia.org/wiki/Matrix_exponential?oldid=740718326 *Contributors:* Tbackstr, XJaM, Jdpipe, Michael Hardy, TakuyaMurata, Cyp, Dysprosia, Jitse Niesen, Fredrik, MathMartin, Robinh, Tea2min, Giftlite, BenFrantzDale, Fropuff, Andris, Antandrus, Shahab, Piil, Gauge, Iamunknown, Linas, HappyCamper, R.e.b., John Deas, Mathbot, Jeff02, Algebraist, Wavelength, Goffrie, Maechler, Arthur Rubin, SmackBot, Diegotorquemada, Moogle001, Vanished User 0001, Chiao, Jbergquist, Paul Matthews, HenningThielemann, Xantharius, Thijs!bot, Epbr123, Sherbrooke, BigJohnHenry, Ben pcc, Jay Gatsby, Mah159, David Eppstein, Franp9am, Adavidb, Haseldon, BrianOfRugby, Buechen, Thom Tyrrell, Solian en, Maxzimet, Cuzkatzimhut, VolkovBot, Lechatjaune, Kfgauss, Curtdbz, Mr. Granger, Sfan00 IMG, Hblatz, Wenjiajing, Addbot, Glasscup7, WardenWalk, Balabiot, Luckas-bot, Yobot, Ptbotgourou, Xuancong, Anypodetos, AnomieBOT, Citation bot, Bdmy, Kensaii, Demihalf, RobinK, Milolance, Duoduoduo, EmausBot, Sashwattanay, GoingBatty, Iamfullofspam, ZéroBot, Quondum, AManWithNoPlan, Milad pourrahmani, Lhibbeler, Zfeinst, Zueignung, Sailsbystars, Aerthis, Boris Breuer, El Roih, Enopet, Helpful Pixie Bot, BG19bot, Odaniel1, Ece8950, KitchiRUs, Lukeje, Dexbot, Mogism, Lucie911, Robleroble, Promise her a definition, Mathphysman, Abitslow, ASCarretero and Anonymous: 66
- **Prime number theorem** *Source:* https://en.wikipedia.org/wiki/Prime_number_theorem?oldid=735222248 *Contributors:* AxelBoldt, Bryan Derksen, XJaM, Bernfarr, Chas zzz brown, Michael Hardy, Kidburla, TakuyaMurata, Looxix~enwiki, Wael Ellithy, Charles Matthews, Dcoetzee, Dysprosia, Jitse Niesen, Robbot, TMC1221, Gandalf61, Wereon, JensG~enwiki, JerryFriedman, Tea2min, Decrypt3, Giftlite, Dbenbenn, Lupin, Herbee, Python eggs, Fred Stober, Rich Farmbrough, Guanabot, Roybb95~enwiki, Paul August, Dmr2, Bender235, Jnestorius, EmilJ, Billymac00, John Vandenberg, OBryant, Sligocki, Skimaxpower, Mcsee, Linas, Lucienve, Xiong Chiamiov, Reddwarf2956, Ketiltrout, Rjwilmsi, HappyCamper, R.e.b., Bubba73, FlaBot, Mathbot, Maxal, Alexjohnc3, Enon, Sstrader, Diza, Sodin, Scythe33, Glenn L, Chobot, ScottAlanHill, YurikBot, Wavelength, Dmharvey, KSmrq, Anomalocaris, LMSchmitt, DYLAN LENNON~enwiki, Bbaumer, Kompik, Arthur Rubin, GrinBot~enwiki, RDBury, BeteNoir, Dahn, PrimeHunter, MalafayaBot, Gutworth, Viebel, Ck lostsword, Khazar, Ninjagecko, ZAB, Statsone, JoshuaZ, WAREL, CRGreathouse, CmdrObot, Myasuda, Gregbard, Doctormatt, Mon4, Robertinventor, Karl-H, Thijs!bot, Headbomb, Nuesken, Noel Bush, AbcXyz, Mdotley, Salgueiro~enwiki, Olaf, JamesBWatson, Zooloo, David Eppstein, Kope, Nono64, Fellwalker57, Rrostrom, Maurice Carbonaro, Tarotcards, Policron, Dessources, DavidCBryant, Inwind, Eric Ng, TXiKiBoT, Nxavar, Anonymous Dissident, Don4of4, Hagman, Arcfrk, GirasoleDE, SieBot, Jonlandrum, Murraythemathgeek, Justin W Smith, DragonBot, Watchduck, Jsondow, DumZiBoT, Katsushi, Nicolae Coman, Addbot, DOI bot, Frobitz, Uncia, LinkFA-Bot, SPat, Charleswallingford, II MusLiM HyBRiD II, Bab dz, AnomieBOT, Citation bot, ArthurBot, Xqbot, Gap9551, Shirik, Raulshc, D'ohBot, Motomuku, Citation bot 1, PrBeacon, Pinethicket, Primalbeing, Mikewarbz, Carel.jonkhout, TobeBot, Duoduoduo, Thomassteinke, RjwilmsiBot, EmausBot, John of Reading, Vincent Semeria, RA0808, Dcirovic, Slawekb, Endlessoblivion, Quondum, Toshio Yamaguchi, Maschen, Sapphorain, Pottermagic, ClueBot NG, Helpful Pixie Bot, Asmallworld, Brad7777, BattyBot, Boeing720, Dexbot, Deltahedron, AppliedMathematics, Mark viking, David9550, Monkbot, Twain Cleft, Teddyktchan, Whyes19, Npip99 and Anonymous: 123
- **Cycle detection** *Source:* https://en.wikipedia.org/wiki/Cycle_detection?oldid=736791266 *Contributors:* Michael Hardy, Charles Matthews, Zoicon5, Altenmann, MathMartin, Decrypt3, Giftlite, Mellum, Macrakis, Vadmium, BRW, Tabletop, Rjwilmsi, YurikBot, Borgx, Cate,

InverseHypercube, Nbarth, Allan McInnes, EdC~enwiki, BananaFiend, Gabn1, Chrisahn, Cydebot, Widefox, Salgueiro~enwiki, Colin Greene, David Eppstein, R'n'B, Cobi, Mohamedafattah, BradAustin2, Jamelan, Svick, ClueBot, Johnuniq, Rror, WikHead, Addbot, Tinus74, MrOllie, Yobot, KamikazeBot, Citation bot, Xqbot, Charvest, Citation bot 1, Corwinjoy, Paradoxolog, Sudozero, ClueBot NG, BG19bot, Themysteriousimmigrant, Mimibar, Phanikumarv, WEF brat sad, Sumit b130338cs, Satyaanveshee and Anonymous: 63

- **Fibonacci number** *Source:* https://en.wikipedia.org/wiki/Fibonacci_number?oldid=741741511 *Contributors:* AxelBoldt, The Anome, XJaM, Christian List, PierreAbbat, FvdP, Nonenmac, Heron, Karl Palmen, Camembert, Quercusrobur, Patrick, D, Michael Hardy, Palnatoke, Llywrch, Dominus, Shyamal, Kku, SGBailey, Ixfd64, Chinju, Seav, TakuyaMurata, GTBacchus, Ee79, Ahoerstemeier, Jimfbleak, Angela, Den fjättrade ankan~enwiki, Александър, Nikai, Jiang, Jacquerie27, Cherkash, Raven in Orbit, Hashar, Alex Vinokur~enwiki, Charles Matthews, Timwi, Dcoetzee, Dysprosia, Gutza, Tb, Zoicon5, Tpbradbury, Hyacinth, Gutsul, Fibonacci, McKay, AnonMoos, Shafei, Phil Boswell, Chuunen Baka, Nufy8, Robbot, Pigsonthewing, Cgranade, Fredrik, TMC1221, RedWolf, Gandalf61, Yarvin, Academic Challenger, Wikibot, Wereon, Isopropyl, PrimeFan, Jleedev, Tea2min, Adam78, Dave6, Tosha, Giftlite, Levin, Herbee, Numerao, Fropuff, Average Earthman, Snowdog, Pashute, Sunny256, Mellum, Gareth Wyn, DO'Neil, Guanaco, Dmmaus, Mboverload, Jackol, Pne, Mckaysalisbury, Gubbubu, Zeimusu, Noe, Antandrus, MarkSweep, Lesgles, Mukerjee, Gauss, EBB, Enzotib, Zfr, Sam Hocevar, Peter Kwok, ArthurDenture, Ukexpat, Myles Callum, McCart42, Kadambarid, Porges, Mike Rosoft, Mormegil, DanielCristofani, Diagonalfish, Discospinster, Guanabot, Satyadev, Francis Schonken, LindsayH, D-Notice, Antaeus Feldspar, Paul August, Bender235, Zaslav, Violetriga, Brian0918, El C, Rgdboer, Surachit, Crisófilax, Tom, Shamilton, Bobo192, Iamunknown, Alex.zeffertt, Billymac00, Smalljim, Srl, Jxn, Juzeris, Toh, Blotwell, Emhoo~enwiki, Manu.m, Nsaa, Eddideigel, Hyperdivision, Jumbuck, Schissel, Alansohn, Terrycojones, Cjthellama, RobertStar20, Esrob, Burn, Spangineer, Bart133, Caesura, Snowolf, Wtmitchell, Velella, M3tainfo, RJFJR, H2g2bob, Woodstone, Algocu, LunaticFringe, Sturmde, Oleg Alexandrov, Brookie, Velho, Simetrical, Reinoutr, Woohookitty, Linas, Lochaber, Blumpkin, RHaworth, Shreevatsa, Igny, David Haslam, Borb, Dustball, Unixer, SP-KP, Hbdragon88, Firien, I64s, GregorB, Cornince, Sam Coutu-Oughton, Ryan Reich, Prashanthns, Jacj, Reddwarf2956, Palica, Gerbrant, Graham87, Qwertyus, Zzedar, GrundyCamellia, Porcher, Jshadias, Crzrussian, Sjakkalle, Rjwilmsi, Lars T., Matt.whitby, MarSch, Orville Eastland, Bruce1ee, TheRingess, Salix alba, Supreme geek overlord, Bhadani, Bmenrigh, Fred Bradstadt, FlaBot, Fivemack, Mathbot, Maxal, TheDJ, TeaDrinker, Kri, Glenn L, Chobot, Heatherc, Krishnavedala, DVdm, EamonnPKeane, YurikBot, RobotE, Hairy Dude, Deeptrivia, Nmondal, Crackpottheorist, Phantomsteve, Petiatil, Sarranduin, Postglock, Loom91, KSmrq, DanielKO, Inaniae, Ukdragon37, Stephenb, Gaius Cornelius, Pseudomonas, Salsb, Wimt, NawlinWiki, Rick Norwood, Johann Wolfgang, Trovatore, LMSchmitt, RazorICE, JocK, Johantheghost, Fantusta, Xdenizen, Fixedd, Pixiequix, Mikeblas, Tony1, Merosonox, Aaron Schulz, RyanJones, CLW, DomenicDenicola, Wknight94, Pejhman, NormDor, Georgewilliamherbert, Zzuuzz, Closedmouth, Djkimmons, Redgolpe, Junkmonkey, Davidwt, QmunkE, Katieh5584, JDspeeder1, GrinBot~enwiki, Ffangs, DVD R W, Finell, Tom Morris, Eog1916, Robertd, A bit iffy, GrafZahl, SmackBot, RDBury, YellowMonkey, Ashenai, Roger Hui, InverseHypercube, Melchoir, Bwilliams, KocjoBot~enwiki, Allixpeeke, Senordingdong, Karl Stroetmann, ScaldingHotSoup, Mit-Mit, Mscuthbert, Lucas9000, Delldot, Rouenpucelle, Katsarts, Shai-kun, Srnec, Yamaguchi??, Robbjedi, Gilliam, Skizzik, ERcheck, Psiphiorg, Wutchamacallit27, Hitman012, Chris the speller, JCSantos, Agateller, Alan smithee, PrimeHunter, Miquonranger03, Silly rabbit, Timneu22, SchfiftyThree, Stevage, Liebeskind, Adun12, Octahedron80, Nbarth, Jxm, DHN-bot~enwiki, Beamrider, Zven, Konstable, Darth Panda, Zsinj, Can't sleep, clown will eat me, ZombieDance, Tamfang, Ahudson, Armend, Addshore, Pax85, Soosed, JustAnotherJoe, Cybercobra, Cubbi, Dreadstar, Jeffasinger, DMacks, Danielklein, Daniel.Cardenas, Jóna Þórunn, Pilotguy, Kukini, Desmond71, Zac67, Ohconfucius, SashatoBot, Eliyak, Krashlandon, Richard L. Peterson, Jambo961, MagnaMopus, Jrvz, Cronholm144, Gobonobo, Laursch, JoshuaZ, Jim.belk, Scetoaux, Aleenf1, IronGargoyle, Adamli9, Ko34, Smith609, Illythr, Slakr, TheHYPO, Dicklyon, Brainix, Waggers, Plattler01, Mets501, Openyourminds, Snezzy, Euphonisten~enwiki, Mkosto, MystRivenExile, Joachimheck, Iridescent, Abhinaba, Madmath789, Wfgiuliano, Shoeofdeath, Beno1000, CapitalR, Punnytno1, Bharatveer, Trialsanderrors, Courcelles, Dpeters11, Alex kraemer, Tawkerbot2, JRSpriggs, Harold f, Fvasconcellos, FatalError, Nyktos, JForget, CRGreathouse, Ahy1, Tanthalas39, Snorkelman, Patchouli, Ale jrb, Wafulz, Sir Vicious, Zarex, Dgw, Yarnalgo, DanielRigal, Lentower, Myasuda, Fph, Flying Saucer, Nilfanion, Mapletip, Divineprime, MC10, SunDog, Crossmr, DrunkenSmurf, Gogo Dodo, JFreeman, Jayen466, Soetermans, Gtalal, Tawkerbot4, DumbBOT, Pingku, Surturz, Planet guru, Ward3001, Protious, Mtpaley, Omicronpersei8, Dyanega, Gimmetrow, Thijs!bot, Epbr123, Mercury~enwiki, Curious George 57, TonyTheTiger, Timo3, Mojo Hand, Headbomb, Marek69, Dfrg.msc, My name, Jbl1975, Muaddeeb, CharlotteWebb, Kevin.gould, Gierszep, SusanLesch, Dawnseeker2000, CompositeFan, Escarbot, Jimmah, Stannered, Mentifisto, Hmrox, WikiSlasher, Vafthrudnir, AntiVandalBot, Majorly, Luna Santin, Asadlarik3, Seaphoto, Quintote, Paste, Dylan Lake, Lonestar662p3, Altamel, Wazinger, Storkk, Res2216firestar, HolyT, JAnDbot, RETROFUTURE, Kaobear, Bakasuprman, MER-C, Henkk78, Hut 8.5, Frankie816, PhilKnight, Four Dog Night, Johngouf85, GoodDamon, Magioladitis, Connormah, Secret Squïrrel, Markwalters79, Prof.rick, Bongwarrior, VoABot II, Sabamo, Hillgentleman, Swat510, JNW, Albmont, PeterStJohn, Demortes, Twsx, Nikevich, Ahabvihrea, SparrowsWing, Theroadislong, Animum, Dirac66, 28421u2232nfenfcenc, David Eppstein, Sil3ntr3con, Styrofoam1994, Martynas Patasius, Cpl Syx, RyanHLewis, Kope, DerHexer, DinoBot, Jonomacdrones, MichaelGood, S3000, MartinBot, GimliDotNet, Slyfoxx, Ravichandar84, Anaxial, Arithmonic, R'n'B, Olivier Danvy, Vox Rationis, Lilac Soul, Joemoser, Tgeairn, J.delanoy, Prokofiev2, AstroHurricane001, Cyanolinguophile, P n, Eliz81, Mike.lifeguard, B Gallagher, Hjsmithh, Vanished user 342562, Icseaturtles, Katalaveno, McSly, Ignatzmice, Bailo26, Evaristor, (jarbarf), GongYi, Chiswick Chap, Daniel5Ko, Exdejesus, NewEnglandYankee, Fibonaccimaster, SJP, Policron, Gfis, Juliancolton, ChanTab, PTP2009, Sarregouset, Pdcook, A-Doo, ImMAW, Halmstad, Dreamfoundry, CardinalDan, Xenonice, Cuzkatzimhut, VolkovBot, TreasuryTag, CWii, Thedjatclubrock, Pleasantville, JohnBlackburne, Nburden, Rocketman116, JoeDeRose, LokiClock, AlnoktaBOT, HeckXX, Jangles5150, Devileen, Larry R. Holmgren, Nousernamesleft, FerralMoonrender, LeilaniLad, Philip Trueman, TXiKiBoT, Rambo forever, BuickCenturyDriver, BillyNair, Jtrollheiser, Miranda, Nxavar, Rei-bot, Anonymous Dissident, Aymatth2, Clark Kimberling, Qxz, Stuza2, Aznfoo, Ocolon, Lradrama, Corvus cornix, BilinearWidder, Broadbot, Jackfork, Blackskilled, Berstff 27, Dangiankit, Jaswenso, Wenli, BroVic, Moshi1618, DrJohnG, Falcon8765, Lukehodgson, Insanity Incarnate, Dmcq, Monty845, Jamrb, MagMath, Munci, Ishboyfay, Sfmammamia, Steven Weston, Arjun024, GoonerDP, SieBot, Ivan Štambuk, Calliopejen1, Tresiden, Tiddly Tom, Work permit, Jauerback, Pogmothoin2486, Lemonflash, Winchelsea, Caltas, RJaguar3, Killerray, Triwbe, GrooveDog, Toddst1, Flyer22 Reborn, Tiptoety, Radon210, JetLover, SPACKlick, Oxymoron83, Faradayplank, Iceglow, Alex.muller, IdreamofJeanie, OKBot, Dstlascaux, MFSP, Sean.hoyland, WikiLaurent, Nic bor, Denisarona, Stephen.c.johnson, Escape Orbit, Renfield286, ClueBot, Avenged Eightfold, GorillaWarfare, Marius.bancila, Justin W Smith, F cooper 8472, The Thing That Should Not Be, Lockoom, Vacio, Nsk92, Bugtank, Maniac18, Objective3000, Polyamorph, SuperHamster, Xavexgoem, Blindmansays, Hdfkkfghvjmbb, LizardJr8, ChandlerMapBot, Arunsingh16, Puchiko, Manishearth, DragonBot, Ggea, Excirial, Ticmarc, Jusdafax, Watchduck, Tomeasy, Coralmizu, PixelBot, Tylerdmace, Dragflame383, ParisianBlade, NuclearWarfare, 7times7, Jotterbot, Razorflame, Dekisugi, Wikidsp, Saebjorn, Thingg, Sarindam7, Yonizilpa, Aitias, Tulcod, Ertemplin, Versus22, Johnuniq, DumZiBoT, SlaterDeterminant, Lefterov, Alchemist Jack, Jean-claude perez, Helixweb, Ryan8374, Mew 147258, XLinkBot, Spitfire, Tarquilu, Dark Mage, BodhisattvaBot, Duncan, DaL33T, Little Mountain 5, Avoided, NellieBly, Virginia-American, Kombiman, PL290, Manfi, Badgernet, NHJG, Jmaccaro, RyanCross, Wyatt915, Pyninja, Addbot, Somebody9973, Willking1979, Some jerk on the Internet, DOI bot, Tcncv, Landon1980, Theleftorium, GSMR, Jackwestjr, Jessepfrancis, Ronhjones, Gustavo José Meano Brito, Opland,

- **Catalan number** *Source:* https://en.wikipedia.org/wiki/Catalan_number?oldid=739296349 *Contributors:* Damian Yerrick, AxelBoldt, XJaM, Michael Hardy, Dominus, Charles Matthews, Xiaodai~enwiki, Robbot, Fredrik, Henrygb, Reiner Martin, Timrollpickering, JesseW, PrimeFan, Giftlite, Achurch, Sj, Herbee, Fropuff, Fudo, Profvk, Sam nead, Rama, ESkog, Rgdboer, Crisófilax, EmilJ, Jcreed, Blotwell, Caesura, Greg Kuperberg, Igorpak, Linas, Gisling, Sukolsak, Rjwilmsi, Joffan, R.e.b., Maxim Razin, FlaBot, John Baez, Mathbot, StephanCom, Maxal, Vince Vatter, ChongDae, Chobot, Dylan Thurston, YurikBot, Dmharvey, RussBot, JocK, Mikeblas, Zwobot, Kompik, Arthur Rubin, GrEp, Robertd, SmackBot, RDBury, Nihonjoe, Chodges, Unyoyega, Bluebot, MalafayaBot, Javalenok, Mhym, Armend, B jonas, MichaelBillington, Tanyakh, Lambiam, EdC~enwiki, A. Pichler, CRGreathouse, Arratax, Yrodro, Swagatata, Thijs!bot, Mojo Hand, Headbomb, Paxinum, Dan D. Ric, Magioladitis, Hans Lundmark, David Eppstein, JoergenB, Connor Behan, Dannyxyz22, Maproom, AndPud, Quantling, Vadik, Chinneeb, Amikake3, Synthebot, Brianga, Moltean, SieBot, Steven Crossin, ClueBot, Bob1960evens, Veltzerdoron, JuPitEer, Sabri76, Auntof6, Sambrow, DragonBot, Watchduck, LaosLos, Vonsolms, Marc van Leeuwen, Gonfer, SilvonenBot, MystBot, Tpeng, Addbot, Olli Niemitalo, Luckytoilet, Lightbot, مانی, Luckas-bot, Yobot, AnomieBOT, VanishedUser sdu9aya9fasdsopa, Ciphers, Hairhorn, Kingpin13, Ulric1313, Rar, Xqbot, Drilnoth, Nitrxgen, Kamitsaha, FrescoBot, Ikska, King Midas1234, MondalorBot, Jshimbi, Rabiddog51sb, The tree stump, EmausBot, John of Reading, Netwriter2, ZéroBot, Cartanalgebra, Quondum, Num Ref, Toshio Yamaguchi, Donner60, ClueBot NG, Wcherowi, Colapeninsula, Bernhard Oemer, Joel B. Lewis, Helpful Pixie Bot, Knightofairplanes, BG19bot, יהודה שמחה ולדמן, BattyBot, Sa145, Purdygb, ChrisGualtieri, Ys93, EuroCarGT, Enterprisey, Themysteriousimmigrant, Veryltdbeard, Cyrapas, Tango303, Julia Abril, Jackmcbarn, Monkbot, GeoffreyT2000, JMP EAX, Aman anand1994, TheMathCat, Ronklein, Jt omega and Anonymous: 116
- **Extended Euclidean algorithm** *Source:* https://en.wikipedia.org/wiki/Extended_Euclidean_algorithm?oldid=735617948 *Contributors:* AxelBoldt, Taral, Michael Hardy, Ixfd64, J-Wiki, Silverfish, Dcoetzee, Jogloran, Cintrom, Robbot, Rorro, Aisotton, JesseW, Jleedev, Nitishkorula, Tdgs, Connelly, Giftlite, Herbee, Pgan002, Finog, Lemontea, Paul August, Oleg Alexandrov, Eddy264, Ruud Koot, MFH, GregorB, MrSomeone, Graham87, NeoUrfahraner, Nguyen Thanh Quang, Quuxplusone, Algebraist, Hairy Dude, Dmharvey, Robost, Ronyclau, Arthur Rubin, Cedar101, Samuel Blanning, Cmglee, SmackBot, InverseHypercube, Teimu.tm, Chris the speller, Comfortably Paranoid, DHN-bot~enwiki, Tamfang, Contrasedative, Alca Isilon~enwiki, Jbonneau, Bill Malloy, Jbolden1517, CRGreathouse, Aherunar, Joshuamonkey, Oerjan, Courtjester555, Magioladitis, Stdazi, David Eppstein, Hbent, Glrx, Martinor, Krishnachandranvn, Robertgreer, Sigmundur, Goyston, R00723r0, Nxavar, Rei-bot, Josephholsten, Kmhkmh, Moonriddengirl, Zero2ninE, Mr. Stradivarius, PerryTachett, Rdhettinger, ClueBot, Jdgilbey, Estirabot, Marc van Leeuwen, Addbot, Bluebusy, Luckas-bot, Yobot, BlazerKnight, ArthurBot, X7q, Toolnut, Luckyblue1, Otrabajo, John of Reading, Skysmurf, WikitanvirBot, Josve05a, D.Lazard, Dsmithsmithy, JordiGH, Frietjes, Partmedia, David815, WikiBC, Purdygb, Lp.vitor, Dane Bouchie, Kyle1009, Solid Frog, Prappl93 and Anonymous: 177
- **Factorial** *Source:* https://en.wikipedia.org/wiki/Factorial?oldid=741043209 *Contributors:* AxelBoldt, Zundark, XJaM, Christian List, Matusz, Karl Palmen, Youandme, Bdesham, Patrick, Michael Hardy, Chris-martin, Dominus, SGBailey, Ixfd64, TakuyaMurata, Eric119, Pde, Minesweeper, Ahoerstemeier, Stevenj, JWSchmidt, Kevin Baas, Poor Yorick, Nikai, Andres, Jonik, Schneelocke, Agtx, Emperorbma, Dcoetzee, Rob.derosa, Dysprosia, Wik, Furrykef, Hyacinth, Taxman, Fibonacci, Sabbut, McKay, Denelson83, Robbot, Astronautics~enwiki, Fredrik, Altenmann, Sverdrup, Yarvin, Henrygb, JB82, Isopropyl, PrimeFan, Wile E. Heresiarch, Jleedev, Connelly, Giftlite, Ævar Arnfjörð Bjarmason, MSGJ, Herbee, Anton Mravcek, TomViza, Frencheigh, Mboverload, Xenoglossophobe, Knutux, Zeimusu, MarkSweep, Tels~enwiki, Maximaximax, Tomruen, Petershank, Jiel.B, Thorwald, Slady, Rich Farmbrough, Guanabot, ArnoldReinhold, Paul August, Bender235, Zaslav, Nabla, Iamunknown, Robotje, Phlake, Obradovic Goran, Wrs1864, Jonathunder, Quaoar, Jumbuck, Eric Kvaalen, JebeddiahSpringfield, Andrew Gray, Fritzpoll, PAR, Wtmitchell, Super-Magician, Evil Monkey, Deathphoenix, Oleg Alexandrov, Waabu, Arneth, Georgia guy, MattGiuca, Koolman2, MONGO, Eras-mus, Zzyzx11, Audiovideo, Palica, Stevey7788, Graham87, Seb-Gibbs, Eumeme~enwiki, Zzedar, Salix alba, Jwmcleod, NeonMerlin, Bubba73, Bryan H Bell, Yamamoto Ichiro, Nomet, Mathbot, Parerga, Glenn L, CiaPan, Chobot, Bgwhite, WriterHound, Wavelength, Ec-, Hairy Dude, Michael Slone, Arado, JabberWok, Rhythm, Thatoneguy, Darkmeerkat, Bota47, Ms2ger, Jezzabr, Arthur Rubin, Acer, Mrbowtie, Kier07, Gesslein, Carlosguitar, GrinBot~enwiki, Capitalist, Marquez~enwiki, SmackBot, InverseHypercube, Melchoir, TheArcher, PrimeHunter, Octahedron80, DHN-bot~enwiki, Colonies Chris, Nicolas.Wu, Chlewbot, Kaimiddleton, Jgoulden, Lhf, Agreatnate, Wybot, Blahm, Autopilot,

Lambiam, ArglebargleIV, Saraghav, Wholmestu, Jim.belk, IronGargoyle, Ckatz, Waggers, Eridani, Mets501, Dhp1080, EdC~enwiki, Necrid Master, Domitori, Happy-melon, A. Pichler, MoleculeUpload, Tawkerbot2, FatalError, CRGreathouse, ZICO, Arabic Pilot, Vvargoal, Tim1988, Mattbuck, FilipeS, Misof, Icek~enwiki, Chrispringle, Boltsman, Vanished user vjhsduheuiui4t5hjri, Mon4, Goldencako, Thijs!bot, King Bee, TXiKi, Curlytop999, AntiVandalBot, Ben pcc, Dcluett, Hannes Eder, Keith111, Db099221, H3llbringer, Alexandre Vassalotti, Magioladitis, Sushi Tax, Swpb, Stuart Morrow, Usien6, Garoto burns, Paul Niquette, Dirac66, David Eppstein, Connor Behan, Pek the Penguin, RaitisMath, MartinBot, Nishantsah, Jasonbosland, Tgeairn, Vorratt, Trusilver, BruceHodge, Keithcc, Lantonov, Indeed123, Alphapeta, Krishnachandranvn, Daniel5Ko, Quantling, Policron, Fishcorn, Izno, Pleasantville, JohnBlackburne, AlnoktaBOT, TXiKiBoT, Jobu0101, Anonymous Dissident, Agricola44, Elphion, Broadbot, Jackfork, Alparmarta, Dmcq, Burgercat, Kbrose, Vs49688, Ernest lk lam, Dogah, SieBot, Korax1214, Spyswimmer33, Quest for Truth, Pilover819, Aruton, Svick, Qwerty mac13, Anchor Link Bot, Ken123BOT, Nitrolicious, ClueBot, Justin W Smith, The Thing That Should Not Be, Jagun, Oldjackson, DragonBot, Excirial, Bender2k14, Aprock, Thingg, XLinkBot, Marc van Leeuwen, Dragohunter, Mitch Ames, Mradam2008, Addbot, Melab-1, Arphibagon, Eutactic, WFPM, MrOllie, Download, CheekierMonkey, Idbelange, Jasper Deng, Numbo3-bot, Xario, Lightbot, דוד שי, Wirkstoff, Jarble, Super Spider, Legobot, Luckas-bot, Yobot, Mech Aaron, Qonnec, Godden46, Kilom691, THEN WHO WAS PHONE?, AnomieBOT, Piano non troppo, ImperatorExercitus, Citation bot, Eric Rowland, Analphabot, Xqbot, Wperdue, Nitrxgen, Hassan210360, Gap9551, GrouchoBot, Jordaan12, Daniel Strobusch, Raulshc, A.amitkumar, Fredericksgary, Prari, Nniigeell, Sae1962, Haein45, Robo37, Citation bot 1, DrilBot, Pinethicket, Mktos532, Gruntler, Danielcorrea, Dude1818, Namaxwell, Ekul81, The Perfection, ThinkEnemies, MegaSloth, Mean as custard, Shafaet, NerdyScienceDude, Mathacw, Whywhenwhohow, EmausBot, GeneralCheese, Tommy2010, Wikipelli, EllaLM, Thewhyman, Ethaniel, Google Child, SporkBot, Num Ref, Chewings72, ClueBot NG, Wcherowi, Frietjes, Parcly Taxel, Joel B. Lewis, Nullzero, Acimarol, Vinicius Metal, Helpful Pixie Bot, Martin of Sheffield, Okidan, Mark Arsten, Vashistha.avinash, Thenewmathemagician, M'encarta, Hicksjo, ChrisGualtieri, Scribbles47, Mkhan3189, Stephan Kulla, Frosty, Initialfluctuation, Jamesx12345, Pokajanje, Jochen Burghardt, Pop-up casket, Burninthruthesky, GN7Original, Jan.cajthaml, Vikasdatta, Mat.wyszynski, Gmobyone11, Legaledits13, Xxad337, Minhaskamal, Prabhakar asthana, Nerd in Texas, GeoffreyT2000, Raycheng200, Ssaz 12, KH-1, JohnScott623, Jimmy7575, Esquivalience, MetazoanMarek, Jccarlosa, Mindotaur and Anonymous: 435

12.13.2 Images

- **File:10_DM_Serie4_Vorderseite.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/0/0d/10_DM_Serie4_Vorderseite.jpg *License:* Public domain *Contributors:* http://www.bundesbank.de/Redaktion/DE/Standardartikel/Kerngeschaeftsfelder/Bargeld/dm_banknoten.html#doc18118bodyText2 *Original artist:* Deutsche Bundesbank, Frankfurt am Main, Germany
- **File:34*21-FibonacciBlocks.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/db/34%2A21-FibonacciBlocks.png *License:* CC BY-SA 4.0 *Contributors:* Own work *Original artist:* ???
- **File:Binary_Tree.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/1/1e/Binary_Tree.png *License:* CC BY-SA 4.0 *Contributors:* http://mathcircle.berkeley.edu/BMC6/pdf0607/catalan.pdf *Original artist:* mathcircle
- **File:Catalan-Hexagons-example.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/a/a8/Catalan-Hexagons-example.svg *License:* Public domain *Contributors:* http://en.wikipedia.org/wiki/File:Catalan-Hexagons-example.svg *Original artist:* Dmharvey
- **File:Catalan_number-path_reflection.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/3/36/Catalan_number-path_reflection.svg *License:* CC BY-SA 4.0 *Contributors:* Own work *Original artist:* Javalenok
- **File:Catalan_number_4x4_grid_example.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/f/f4/Catalan_number_4x4_grid_example.svg *License:* Public domain *Contributors:* ? *Original artist:* ?
- **File:Catalan_number_algorithm_table.png** *Source:* https://upload.wikimedia.org/wikipedia/en/6/65/Catalan_number_algorithm_table.png *License:* PD *Contributors:*
  Own work
  Original artist:
  Dmharvey (talk) (Uploads)
- **File:Catalan_number_binary_tree_example.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/0/01/Catalan_number_binary_tree_example.png *License:* Public domain *Contributors:* ? *Original artist:* ?
- **File:Catalan_number_exceedance_example.png** *Source:* https://upload.wikimedia.org/wikipedia/en/a/aa/Catalan_number_exceedance_example.png *License:* PD *Contributors:*
  Own work
  Original artist:
  Dmharvey (talk) (Uploads)
- **File:Catalan_number_swapping_example.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/1/13/Catalan_number_swapping_example.png *License:* Public domain *Contributors:* ? *Original artist:* ?
- **File:Catalan_stairsteps_4.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/6/63/Catalan_stairsteps_4.svg *License:* CC BY-SA 3.0 *Contributors:* Self-made, using Mathematica and Ruby *Original artist:* Robert Dickau
- **File:Clock_group.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/a/a4/Clock_group.svg *License:* CC-BY-SA-3.0 *Contributors:* Transferred from en.wikipedia to Commons. *Original artist:* The original uploader was Spindled at English Wikipedia
- **File:Commons-logo.svg** *Source:* https://upload.wikimedia.org/wikipedia/en/4/4a/Commons-logo.svg *License:* CC-BY-SA-3.0 *Contributors:* ? *Original artist:* ?
- **File:Complex_zeta.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/1/1b/Complex_zeta.jpg *License:* Public domain *Contributors:* made with Mathematica, own work. See User:Jan_Homann/Mathematics for an explanation of how to generate images like these with Mathematica. *Original artist:* Jan Homann
- **File:Diophantus-cover.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/6/60/Diophantus-cover.jpg *License:* Public domain *Contributors:* ? *Original artist:* ?
- **File:Disqvisitiones-800.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/e/e3/Disqvisitiones-800.jpg *License:* Public domain *Contributors:* ? *Original artist:* ?

- **File:Dyck_lattice_D4.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/0/01/Dyck_lattice_D4.svg *License:* CC BY 3.0 *Contributors:* Own work *Original artist:* Watchduck (a.k.a. Tilman Piesk)
- **File:ECClines-3.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/d0/ECClines-3.svg *License:* CC BY-SA 3.0 *Contributors:* Own work based on Image:ECexamples01.png by en:User:Image:Dake and Image:ECClines-2.svg by SuperManu. *Original artist:* **G**YassineMrabetTalk✉
- **File:ErnstKummer.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/2/2f/ErnstKummer.jpg *License:* Public domain *Contributors:* http://www.math.uni-hamburg.de/home/grothkopf/fotos/math-ges/ *Original artist:* Unknown
- **File:Face_colored_cube.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/0/02/Face_colored_cube.png *License:* Attribution *Contributors:* I created this work myself, based on original image :Commons:Image:Hexahedron.png created by Tomruen which is also licensed as below *Original artist:* The Anome (talk)
- **File:Factorial05.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/c/c6/Factorial05.jpg *License:* Attribution *Contributors:* Own work *Original artist:* Domitori (Dmitrii Kouznetsov)
- **File:FibonacciChamomile.PNG** *Source:* https://upload.wikimedia.org/wikipedia/commons/5/5a/FibonacciChamomile.PNG *License:* CC BY 2.5 *Contributors:*
- Mother_and_daughter.jpg *Original artist:* User:Alvesgaspar:
- **File:FibonacciRabbit.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/7/7a/FibonacciRabbit.svg *License:* CC BY-SA 3.0 *Contributors:*
- Ein_Hase_mit_blauem_Ei.svg *Original artist:* Ein_Hase_mit_blauem_Ei.svg: MichaelFrey & Sundance Raphael
- **File:Fibonacci_spiral_34.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/9/93/Fibonacci_spiral_34.svg *License:* Public domain *Contributors:* Own work using: Inkscape *Original artist:* User:Dicklyon
- **File:Fibonacci_tiling_of_the_plane_and_approximation_to_Golden_Ratio.gif** *Source:* https://upload.wikimedia.org/wikipedia/commons/1/1e/Fibonacci_tiling_of_the_plane_and_approximation_to_Golden_Ratio.gif *License:* CC BY-SA 4.0 *Contributors:* Own work *Original artist:* RobThePoor
- **File:Folder_Hexagonal_Icon.svg** *Source:* https://upload.wikimedia.org/wikipedia/en/4/48/Folder_Hexagonal_Icon.svg *License:* Cc-by-sa-3.0 *Contributors:* ? *Original artist:* ?
- **File:Free-to-read_lock_75.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/8/80/Free-to-read_lock_75.svg *License:* CC0 *Contributors:* Adapted from 9px|Open_Access_logo_PLoS_white_green.svg *Original artist:* This version:Trappist_the_monk (talk) (Uploads)
- **File:Functional_graph.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/d7/Functional_graph.svg *License:* Public domain *Contributors:* Own work *Original artist:* David Eppstein
- **File:Generalized_factorial_function.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/7/73/Generalized_factorial_function.svg *License:* CC BY-SA 3.0 *Contributors:* Mathematica source code: *Original artist:* Self
- **File:Gold,_silver,_and_bronze_rectangles_vertical.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/c/c1/Gold%2C_silver%2C_and_bronze_rectangles_vertical.png *License:* CC BY-SA 4.0 *Contributors:* Own work *Original artist:* Hyacinth
- **File:Hevelius_Selenographia_frontispiece.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/3/32/Hevelius_Selenographia_frontispiece.png *License:* Public domain *Contributors:* Johannes Hevelius, Selenographia *Original artist:* file: myself; artwork drawn by Adolph Boÿ, engraved by Jeremias Falck
- **File:Internet_map_1024.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/d2/Internet_map_1024.jpg *License:* CC BY 2.5 *Contributors:* Originally from the English Wikipedia; description page is/was here. *Original artist:* The Opte Project
- **File:Lehmer_sieve.jpg** *Source:* https://upload.wikimedia.org/wikipedia/en/9/9d/Lehmer_sieve.jpg *License:* CC-BY-3.0 *Contributors:* ? *Original artist:* ?
- **File:Leonhard_Euler.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/d7/Leonhard_Euler.jpg *License:* Public domain *Contributors:*

  2. Kunstmuseum Basel
  Original artist: Jakob Emanuel Handmann
- **File:Liber_abbaci_magliab_f124r.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/0/04/Liber_abbaci_magliab_f124r.jpg *License:* Public domain *Contributors:* ? *Original artist:* ?
- **File:Log-factorial.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/d2/Log-factorial.svg *License:* Public domain *Contributors:* Transferred from en.wikipedia to Commons by Eric Bauman using CommonsHelper. *Original artist:* Ec- at English Wikipedia
- **File:Mingantu'{}s_Catalan_numbers.JPG** *Source:* https://upload.wikimedia.org/wikipedia/commons/4/47/Mingantu%27s_Catalan_numbers.JPG *License:* Public domain *Contributors:* mybook *Original artist:* Mingantu (1692-1763)
- **File:ModularGroup-FundamentalDomain-01.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/a/ad/ModularGroup-FundamentalDomain-01.png *License:* CC-BY-SA-3.0 *Contributors:* from en wikipedia *Original artist:* Fropuff
- **File:Mountain_Ranges.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/7/78/Mountain_Ranges.png *License:* CC BY-SA 4.0 *Contributors:* http://mathcircle.berkeley.edu/BMC6/pdf0607/catalan.pdf *Original artist:* mathcircle
- **File:Noncrossing_partitions_5.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/e/e7/Noncrossing_partitions_5.svg *License:* CC BY 3.0 *Contributors:* Own work *Original artist:* Watchduck (a.k.a. Tilman Piesk)

- **File:Nuvola_apps_edu_mathematics_blue-p.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/3/3e/Nuvola_apps_edu_mathematics_blue-p.svg *License:* GPL *Contributors:* Derivative work from Image:Nuvola apps edu mathematics.png and Image:Nuvola apps edu mathematics-p.svg *Original artist:* David Vignoni (original icon); Flamurai (SVG convertion); bayo (color)
- **File:OEISicon_light.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/d8/OEISicon_light.svg *License:* Public domain *Contributors:* Own work *Original artist:* Watchduck (a.k.a. Tilman Piesk)
- **File:PascalTriangleFibanacci.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/b/bf/PascalTriangleFibanacci.svg *License:* CC BY-SA 3.0 *Contributors:* Own work *Original artist:* RDBury
- **File:Peter_Gustav_Lejeune_Dirichlet.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/3/32/Peter_Gustav_Lejeune_Dirichlet.jpg *License:* Public domain *Contributors:* Unknown *Original artist:* Unknown
- **File:Pierre_de_Fermat.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/4/4b/Pierre_de_Fermat.png *License:* Public domain *Contributors:* ? *Original artist:* ?
- **File:Plimpton_322.jpg** *Source:* https://upload.wikimedia.org/wikipedia/commons/c/c2/Plimpton_322.jpg *License:* Public domain *Contributors:* image copied from http://www.math.ubc.ca/~{}cass/courses/m446-03/pl322/pl322.html *Original artist:* photo author unknown
- **File:Portal-puzzle.svg** *Source:* https://upload.wikimedia.org/wikipedia/en/f/fd/Portal-puzzle.svg *License:* Public domain *Contributors:* ? *Original artist:* ?
- **File:Prime_number_theorem_absolute_error.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/6/6e/Prime_number_theorem_absolute_error.svg *License:* CC0 *Contributors:* Own work *Original artist:* User:Dcoetzee
- **File:Prime_number_theorem_ratio_convergence.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/8/87/Prime_number_theorem_ratio_convergence.svg *License:* CC0 *Contributors:* Own work *Original artist:* User:Dcoetzee
- **File:Primes_-_distribution_-_up_to_19_primorial.png** *Source:* https://upload.wikimedia.org/wikipedia/commons/0/0a/Primes_-_distribution_-_up_to_19_primorial.png *License:* CC BY-SA 3.0 *Contributors:* Own work *Original artist:* Endlessoblivion
- **File:Question_book-new.svg** *Source:* https://upload.wikimedia.org/wikipedia/en/9/99/Question_book-new.svg *License:* Cc-by-sa-3.0 *Contributors:*

  Created from scratch in Adobe Illustrator. Based on Image:Question book.png created by User:Equazcion *Original artist:* Tkgd2007
- **File:Rtriangle.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/6/6f/Rtriangle.svg *License:* CC-BY-SA-3.0 *Contributors:* ? *Original artist:* ?
- **File:SunflowerModel.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/a/ad/SunflowerModel.svg *License:* CC-BY-SA-3.0 *Contributors:* Own work *Original artist:* Doron
- **File:Tamari_lattice,_hexagons.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/3/35/Tamari_lattice%2C_hexagons.svg *License:* Public domain *Contributors:* Own work *Original artist:* Watchduck (a.k.a. Tilman Piesk)
- **File:Tamari_lattice,_trees.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/f/ff/Tamari_lattice%2C_trees.svg *License:* Public domain *Contributors:* Own work *Original artist:* Watchduck (a.k.a. Tilman Piesk)
- **File:Tortoise_and_hare_algorithm.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/5/5f/Tortoise_and_hare_algorithm.svg *License:* CC BY 3.0 *Contributors:* self-made, based on tortoise photo by Aaron Logan and Hare photo by Malene Thyssen. Both photos are freely licensed but require attribution in any derivative works, hence I am using a similar license for this image. *Original artist:* David Eppstein
- **File:Wiki_letter_w_cropped.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/1/1c/Wiki_letter_w_cropped.svg *License:* CC-BY-SA-3.0 *Contributors:* This file was derived from Wiki letter w.svg:
  Original artist: Derivative work by Thumperward
- **File:Wikibooks-logo-en-noslogan.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/d/df/Wikibooks-logo-en-noslogan.svg *License:* CC BY-SA 3.0 *Contributors:* Own work *Original artist:* User:Bastique, User:Ramac et al.
- **File:Wikiquote-logo.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/f/fa/Wikiquote-logo.svg *License:* Public domain *Contributors:* Own work *Original artist:* Rei-artur
- **File:Wikiversity-logo-en.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/1/1b/Wikiversity-logo-en.svg *License:* CC BY-SA 3.0 *Contributors:* Own work *Original artist:* Snorky
- **File:X_chromosome_ancestral_line_Fibonacci_sequence.svg** *Source:* https://upload.wikimedia.org/wikipedia/commons/e/ed/X_chromosome_ancestral_line_Fibonacci_sequence.svg *License:* CC BY-SA 4.0 *Contributors:* Own work *Original artist:* Algorithmicist

12.13.3 Content license

- Creative Commons Attribution-Share Alike 3.0