Computational number theory

C. Pomerance

Department of Mathematics, Dartmouth College, Hanover, NH 03755, USA

1 Introduction

Historically mathematics has always been interested in computation. From the Egyptians who invented geometry so as to measure fields, and the Greeks whose Ptolemaic epicycles spurred trigonometry so they could predict the positions of the planets, to modern researchers whose wavelets allow CAT scans, ancient and not-so-ancient mathematicians have computed so as to better understand our natural world. Pure mathematics as well has historically benefitted from computation, with many of our great theorems and conjectures motivated at root by computational experience. It is said that Gauss, who was an excellent computationalist, needed only to work out a concrete example or two to discover, and then prove, the underlying theorem. While perhaps some branches of mathematics have lost contact with their computational roots, the advance of cheap computational power and convenient mathematical software have helped to reverse this trend.

Our topic here is computational number theory, which may be argued to be in the vanguard for the revisiting of pure mathematics to its computational roots. A prescient call-to-arms was issued by Gauss already in 1801:

"The problem of distinguishing prime numbers from composite numbers, and of resolving the latter into their prime factors, is known to be one of the most important and useful in arithmetic. It has engaged the industry and wisdom of ancient and modern geometers to such an extent that it would be superfluous to discuss the problem at length. Nevertheless we must confess that all methods that have been proposed thus far are either restricted to very special cases or are so laborious and difficult that even for numbers that do not exceed the limits of tables constructed by estimable men, they try the patience of even the practiced calculator. And these methods do not apply at all to larger numbers... Further, the dignity of the science itself seems to require that every possible means be explored for the solution of a problem so elegant and so celebrated."

In addition to this most basic issue of factorization into primes, essentially every branch of number theory has its computational component. And in some areas, there is such a robust computational literature, that we discuss the algorithms involved as mathematically interesting objects in their own right. In this article we shall briefly visit some examples of the computational spirit in analytic number theory (the distribution of primes and the Riemann Hypothesis), in diophantine equations (the abc conjecture), and in elementary number theory, as so poignantly called for by Gauss. With this rather broad brush, much interesting computational work in number theory must be left out, but it is my hope that this introduction will spur further interest.

2 Distinguishing prime numbers from composite numbers

The problem is simple to state. Given an integer $n > 1$, decide if n is prime or composite. And we all know an algorithm. Given n, trial divide n by the consecutive integers, either finding a proper factor and so deciding that n is composite, or not, and so deciding that n is prime. Since a composite number n has a proper factor d with $d \le \sqrt{n}$, one can give up on the trial dividing, and decide that n is prime, once one passes $\sqrt{n}$.

This straightforward method is excellent for mental computation with small numbers, and for machine computation for somewhat larger numbers. But it scales poorly, in that if you double the number of digits of n, the time for the worst case is squared. Actually, the algorithm does more than answer our question of whether n is prime or composite, it provides a nontrivial factorization in the latter case. There is nothing wrong with getting more for your money, but if time is money, this algorithm is quite costly.

What we would like is a simply computed criterion that primes satisfy and composites do not, or vice versa. How about Wilson's theorem? This asserts that if n is prime, then $(n-1)! \equiv -1 \pmod{n}$. (For a proof, try to pair residues with their multiplicative inverses, and find that the only residues not paired are 1 and $n-1$.) It is trivial to see that this congruence cannot hold for n composite. However, the Wilson criterion does not meet the standard of being simply computed; we know no especially rapid way of computing factorials modulo another number.

How about Fermat's "little theorem"? Namely, if n is prime and a is any integer, then $a^n \equiv a \pmod{n}$. (A proof is easily fashioned by induction on a, using the binomial theorem for the inductive step.) If computing a large factorial modulo something is hard, maybe it is also hard to compute a large power modulo something. Actually, one of our strongest tools in this subject is the "powermod" algorithm which allows us to compute modular powers very rapidly. The idea is to build up to the desired exponent by a sequence starting with 1 and where each term is the double or the double plus 1 of the preceding term.

Let us illustrate this computation with Fermat's little theorem, taking $a = 2$ and $n = 91$. We write 91 in base 2, getting 1011011. The key sequence then in binary can be read off from the initial base-2 digits, namely, 1, 10, 101, 1011, etc. These are the numbers

$$1,\ 2,\ 5,\ 11,\ 22,\ 45,\ 91,$$

and one can see that indeed we have each term either the double of the prior one or 1 more than this. Our sequence of powers is then

$$2^1 \equiv 2,\ 2^2 \equiv 4,\ 2^5 \equiv 32,\ 2^{11} \equiv 46,$$
$$2^{22} \equiv 23,\ 2^{45} \equiv 57,\ 2^{91} \equiv 37,$$

where each congruence is modulo 91, and each term in the sequence is found by squaring the prior one mod 91 or squaring and multiplying by 2 mod 91.

Wait a second, doesn't Fermat's little theorem say we are supposed to get 2 for the final residue? Well yes, but this is guaranteed only if n is prime. And as you have probably already noticed, 91 is composite. In fact, the computation proves this.

So, here is an example of a computation that proves that n is composite, yet it does not reveal any nontrivial factorization. The powermod algorithm illustrated above scales in an excellent fashion. Indeed, the number of elementary steps is $O((\log n)^3)$, so if the number of digits of n is doubled, the run time of the algorithm is multiplied by about 8. It is an example of a polynomial time algorithm, since the run time is a constant power of the input length.

The reader is invited to try out the powermod algorithm as above, but changing the base of the power from 2 to 3. The answer you should come to is that $3^{91} \equiv 3 \pmod{91}$, that is, the congruence for Fermat's little theorem holds. Since you already know that 91 is composite, I am sure you would not jump to the false conclusion that n is prime! So, as it stands, Fermat's little theorem can be used to sometimes recognize composites, but it cannot be used to recognize primes.

There are two interesting further points to be made regarding Fermat's little theorem. First, on the negative side, there are some composites, such as $n = 561$, where the Fermat congruence holds for *every* integer a. These numbers n are called Carmichael numbers, and unfortunately (from the point of view of testing primality) there are infinitely many, a result of Alford, Granville, and myself. But, on the positive side, if one were to choose randomly among all pairs a, n ranging up to a large bound x, for which $a^n \equiv a \pmod{n}$, almost certainly you would choose a pair with n prime, a result of Erdős and myself.

It is possible to combine Fermat's little theorem with another elementary property of (odd) prime numbers. If n is an odd prime, there are exactly 2 solutions to the congruence $x^2 \equiv 1 \pmod{n}$, namely ± 1. Actually, some composites have this property as well, but composites divisible by two different odd primes do not. Suppose n is an odd prime and say a is an integer not divisible by n. From the little theorem we get $a^{n-1} \equiv 1 \pmod{n}$, so that n divides $a^{n-1} - 1$. We now factor this last expression algebraically using difference-of-squares, and continue to do so as long as we can. Write $n - 1 = 2^s t$, where t is odd. Then

$$a^{n-1} - 1 = (a^{2^{s-1}t} + 1)(a^{2^{s-2}t} + 1) \cdots (a^t + 1)(a^t - 1)$$

so that, since the prime n must divide a factor,

$$\text{either } a^t \equiv 1 \pmod{n} \text{ or } a^{2^i t} \equiv -1 \pmod{n}$$

for some $i = 0, 1, \ldots, s-1$. Call this the "strong Fermat congruence". It takes no longer to try and verify the strong Fermat congruence for a given pair a, n with n odd, then it does to verify the ordinary Fermat congruence. Moreover, as was proved by Monier and Rabin, if n is an odd composite, the strong Fermat congruence fails for at least 3/4 of the choices for a in $[1, n-1]$. That is, there is no analogue of Carmichael numbers for this stronger test.

If you want only to be able to distinguish between primes and composites in practice, and you do not insist on proof, then you have read enough. Namely, given a large odd number n, choose 20 values of a at random from $[1, n-1]$, and begin trying to verify the strong Fermat congruence with these bases a. If ever it should fail, you may stop, the number n must be composite. And if the strong Fermat congruence holds, we might surmise that n is actually prime. Indeed, if n were composite, the Monier–Rabin theorem says the chance that the strong Fermat congruence would hold for 20 random bases is at most 4^{-20}, which is less than one chance in a trillion. However, if the number n is itself chosen by a random process, one also has to figure in the chance that we choose a prime: if the process chooses primes with probability smaller than 4^{-20}, a survivor of the test might be more likely composite than prime! In fact, primes do thin out, and if we are choosing randomly at a very high point, it would indeed be that our chance for choosing a prime is vanishingly small. However, the probability argument can be saved, since the fraction 3/4 in the Monier–Rabin theorem is only a worst case, and on average it is much better. This was worked out by Burthe and others.

If 3/4 of the numbers a in $[1, n-1]$ provide an easily checkable proof that the odd composite number n is indeed composite, surely it shouldn't be so hard to find just one! An idea of Miller is to sequentially check small numbers a until one is found. Excellent, but when do we stop? It can be heuristically argued that there is a constant c such that a stopping point of $c \log n \log\log n$ is adequate. Miller was able to prove the slightly weaker result that a stopping point of $c(\log n)^2$ is adequate, but the proof assumes a generalization of the Riemann Hypothesis. (Namely, one assumes the Riemann Hypothesis for L-functions with quadratic Dirichlet characters.) In further work, Bach was able to show that we may take $c = 2$ in this last result. Summarizing, if this generalized Riemann Hypotheis holds, and if the strong Fermat congruence holds for every integer a in $[1, 2(\log n)^2]$, then n is prime. So, modulo a famous unproved hypothesis in another field of mathematics, one can decide via a deterministic algorithm whether n is prime or composite in polynomial time.

The question is if we can make this decision without assuming unproved hypotheses. In 2002, Agrawal, Kayal, and Saxena answered this question with a resounding yes. In their paper [1], one can find the following beautiful theorem:

Theorem. *Suppose n is an integer with $n \geq 2$, r is a prime not dividing n such that the multiplicative order of $n \pmod r$ is larger than $(\log_2 n)^2$, and*

$$(x+a)^n \equiv x^n + a \pmod{x^r - 1, n} \qquad (1)$$

holds for each positive integer $a \leq \sqrt{r}\log_2 n$. If n has a prime factor greater than $\sqrt{r}\log_2 n$, then n is a power of this prime. In particular, if n has no prime factors in $[1, \sqrt{r}\log_2 n]$ and n is not a proper power, then n is prime.

Some words of explanation are in order. Since $n^{r-1} \equiv 1 \pmod r$, there is a least positive exponent that works here, we call it the multiplicative order of $n \pmod r$. The congruence in (1) has a double modulus, and it signifies that high-degree polynomials are to be reduced modulo $x^r - 1$ and their coefficients are to be reduced modulo n. Further, the notation $\log_2$ is the base-2 logarithm.

It is easy to convert this theorem into a test for primality. Note that if n is prime, then Fermat's little theorem and the binomial theorem imply that (1) holds. It is easy to check if n has a small prime factor or if n is a proper power (the latter is done say with Newton's method to approximate roots, and then checking if an integer approximation to the root, when raised to the appropriate power, actually gives n). Each congruence in (1) can be checked via the powermod method. Space does not allow an illustration, but the reader is invited to try a small example, say $n = 101$, $r = 3$, and $a = 2$: both sides should reduce to $x^2 + 2$. Assuming some fast subroutines for polynomial and integer arithmetic, the verification of (1) takes about $r(\log n)^2$ elementary steps. So the total time for the algorithm is then about $r^{3/2}(\log n)^3$. Clearly the hypothesis about the order implies that r cannot be taken lower than $(\log n)^2$, so $(\log n)^6$ is a lower bound for the run time.

In [1], the authors give an elementary upper bound for r of about $(\log n)^5$, and so attain a time complexity of about $(\log n)^{10.5}$ for the algorithm. Using a complex and numerically ineffective tool, they can bring down the exponent to 7.5. Recently, in [9], Lenstra and I present a complex, but numerically effective method of bringing the exponent down to 6. We do this by expanding the set of polynomials used beyond those of the form $x^r - 1$, and in particular we use polynomials that are constructed from Gaussian periods. Gauss first studied these periods (which are certain sums of complex roots of unity) in his discussion of straightedge and compass constructions of regular n-gons.

One can ask if the new primality test of Agrawal, Kayal, and Saxena, and its variants, are numerically practical. As of yet, the answer is these methods are not too competitive with prior methods. For example, using the arithmetic of elliptic curves, we have a heuristic algorithm that can produce bona fide proofs of primality. (When we say the algorithm is heuristic, we mean that though we believe it will do what we want, it is not guaranteed beforehand to do so.) This method, pioneered by Atkin and Morain, has recently proved the primality of a number with no special form that has over 15,000 decimal digits. The record for the new breed of tests is a measly 300 digits.

Note that for numbers of certain special forms, there are much faster primality tests. The most famous of these special forms is the Mersenne primes, namely primes that are 1 less than a power of 2. The record here is the prime $2^{25,964,951} - 1$, which has over 7.8 million decimal digits.

For much more on primality testing, and references to various other sources, see [5].

3 Factoring composite numbers

Compared to testing primality our ability to factor large numbers is still in the dark ages. In fact this imbalance between the two problems forms the bulwark for the security of electronic commerce on the Internet. Yes, it is indeed an odd form of application, and not something to brag about, since it depends on the inability of mathematicians to efficiently solve a basic problem!

Nevertheless, we do have our tricks. Part of the landscape is an algorithm found in Euclid for computing the greatest common divisor. One might naively think that given two positive integers m and n, to find their gcd, one should find all of their divisors, and pick the largest one common to the two. But Euclid's algorithm is much more efficient: if $m \geq n$, then replace m with its remainder upon division by n, and iterate, with the first division giving the remainder 0 indicating we have found the gcd.

So, if we can build up a special number m that may be likely to have a nontrivial factor in common with n, we might then use Euclid's algorithm to discover this factor. For example, Pollard and Strassen (independently) used this idea, together with fast subroutines for multiplication and polynomial evaluation, to enhance the trial division method discussed in the last section. Somewhat miraculously, one can take the integers up to $n^{1/2}$, break them into $n^{1/4}$ subintervals of length $n^{1/4}$, and for each subinterval, figure the gcd of n with the product of all the integers in the subinterval, spending only about $n^{1/4}$ elementary steps total. If n is composite, at least one gcd will be larger than 1, and then a search over the first such subinterval will locate a nontrivial factor of n. In fact, this algorithm stands as the fastest rigorous and deterministic method to factor that we know.

Most practical factoring algorithms are based on heuristics. We cannot prove rigorously that the method works relatively quickly, or even works at all, but in practice it does.

One factoring idea, which is quite rigorous, goes back to Fermat. Namely, every odd composite is the difference of two squares, which when factored algebraically gives a nontrivial numerical factorization. Indeed, if n has the nontrivial factorization ab, then let $u = (a+b)/2$ and $v = (a-b)/2$, so that $n = u^2 - v^2$, and $a = u + v$, $b = u - v$. I often mention a contest problem from my high school years: factor 8051. The trick is to notice that $8051 = 90^2 - 7^2$, from which the factorization $83 \cdot 97$ can be read off. Though it works very well if n has a divisor very close to $n^{1/2}$, in the worst case, the Fermat method is slower than ordinary trial division.

My quadratic sieve method (which follows work of Kraitchik, Brillhart–Morrison, and Schroeppel) tries to efficiently extend Fermat's idea to all odd composites. For example, take $n = 1649$. We start just above $n^{1/2}$ with $m = 41$, and consider the numbers $m^2 - 1649$. As m runs, we will eventually hit a value where $m^2 - 1649$ is a square, and so be able to use Fermat's method. Instead, we search for a nonempty set $\mathcal{M}$ of numbers m such that

$$\prod_{m \in \mathcal{M}} (m^2 - 1649)$$

is a square. Say it is u^2, and let $v = \prod_{m \in \mathcal{M}} m$. Since $m^2 - 1649 \equiv m^2 \pmod{1649}$, we get

$$u^2 \equiv v^2 \pmod{1649}.$$

This is not quite the same as having $u^2 - v^2 = 1649$, but we can nevertheless compute the gcd of $u - v$ (or $u + v$) with 1649, and, with luck, it will be nontrivial.

An obvious difficulty is finding such a set $\mathcal{M}$. Let us look at the first few numbers:

$$\begin{aligned} 41^2 - 1649 &= 32, \\ 42^2 - 1649 &= 115, \\ 43^2 - 1649 &= 200. \end{aligned}$$

It is easy to see that $32 \cdot 200 = 80^2$, so that we may take $\mathcal{M} = \{41, 43\}$. Thus $u = 80$, and $v = 114 \equiv 41 \cdot 43 \pmod{1649}$. We have $u - v = 80 - 114 = -34$, and the gcd of 34 with 1649 is 17. Done: a nontrivial factor of 1649 has been found.

Of course we will not always choose the first and third numbers to form $\mathcal{M}$. One small lesson from the above example is that the first and third numbers are "smooth", meaning that all of their prime factors are small, while the second number is divisible by the relatively large prime 23. Since it will be difficult to use large primes in assembling a square (a mate for each must be found), we might then just forget about any number that is not smooth. And for the smooth ones, if there are sufficiently many of them, we can use linear algebra to find our square. For the numbers 32 and 200 above, their prime factorizations are $2^5 3^0 5^0$ and $2^3 3^0 5^2$, which correspond to the "exponent vectors" $(5, 0, 0)$ and $(3, 0, 2)$. Since squares only depend on the parity of the exponents, we reduce these exponent vectors mod 2, getting $(1, 0, 0)$ and $(1, 0, 0)$. Notice that they are linearly dependent over the finite field $\mathbf{F}_2$ of two elements; their sum is the 0-vector.

In general with the quadratic sieve, one finds smooth numbers in the sequence $m^2 - n$, forms the exponent vectors mod 2, and then uses linear algebra to find a linear dependency. Since the only scalars in $\mathbf{F}_2$ are 0 and 1, a linear dependency corresponds to a subset sum being the 0-vector, which then corresponds to a set $\mathcal{M}$ for which we search.

In addition, the "sieve" in the quadratic sieve comes in with the search for smooth values $m^2 - n$. These numbers are the consecutive values of a quadratic polynomial (in the variable m), so those divisible by a given prime can be found in regular places in the sequence. E.g., in our illustration, $m^2 - 1649$ is divisible by 5 precisely when $m \equiv 2$ or $3 \pmod 5$. A sieve, very much like the sieve of Eratosthenes, can then be used to efficiently find the special numbers m where $m^2 - n$ is smooth.

Finally note that if the final gcd yields only a trivial factor with n, one can continue just a bit longer and find more linear dependencies, each with a fresh chance at splitting n.

Assuming that the polynomial values $m^2 - n$ are as likely to be smooth as random numbers of the same size, and that there is no conspiracy forcing only trivial factorizations upon us, the time the quadratic sieve takes to factor n is

$$\exp\left((1 + o(1))\sqrt{\log n \, \log\log n}\right).$$

This time complexity is somewhat intermediate between polynomial time, which would look like $\exp(C \log\log n)$, and exponential time, which would look like $\exp(C \log n)$.

Lenstra and I actually have a rigorous, random factoring method with the same time complexity as that above for the quadratic sieve. The method uses class groups of imaginary quadratic number fields, and is based on work of Shanks, Seysen, and A. Lenstra. However, the method is not so computer practical, and, if you had to choose in practice between the two, you should go with the quadratic sieve. One great landmark for the quadratic sieve was the 1994 factorization of the 129-digit RSA cryptographic challenge first published in Martin Gardner's column in *Scientific American* in 1977.

The number field sieve, which is another sieve-based factoring algorithm, was discovered in the late 1980s by Pollard for integers of special form, and later developed for general integers by Buhler, Lenstra, and myself. The method is similar in spirit to the quadratic sieve, but assembles its squares from the product of certain sets of algebraic integers. The number field sieve has the heuristic time complexity of the shape

$$\exp\left(c(\log n)^{1/3}(\log\log n)^{2/3}\right),$$

for a value of c slightly below 2. For composite numbers beyond 100 digits or so and which have no small prime factor, it is the method of choice, with the current record being 200 decimal digits.

If the number n has a small prime factor, it is of course easy to factor n via trial division. If the prime factor is beyond the range of convenient trials, but still considerably below $\sqrt{n}$, one should use Lenstra's elliptic curve method. Since we probably don't know beforehand how large the primes are in a given number, it may be prudent to start with something like trial division, and then the elliptic curve method, before attempting the quadratic or number field sieves.

An elliptic curve can be modeled by an equation $y^2 = x^3 + ax + b$, where a, b come from a particular field (of characteristic not equal to 2 or 3), and $4a^3 + 27b^2 \neq 0$. If one considers the points (x, y) which satisfy the equation, and where x, y come from the coefficient field or a larger field, these points form a commutative group. The group law translates geometrically to the assertion that $P_1 + P_2 + P_3 = 0$ if and only if the three points are collinear. Here, "0" is the point at infinity, and it is visible when one considers the curve in projective space, that is, introducing another variable getting $y^2z = x^3 + axz^2 + bz^3$. The projective point $(0 : 1 : 0)$ is the point at infinity, and it is the only point that is added into the picture by introducing the third variable.

The inverse of a point (x, y) in the elliptic curve group is given simply by $(x, -y)$; that is, the line through them is vertical, and intersects the curve again at 0, the point at infinity. Yhe procedure for adding two points that do not sum to 0 happens to involve finding a multiplicative inverse in the field containing the coordinates of the points. This fact is crucial to the elliptic curve factoring method.

Suppose n is a number we are trying to factor, and we have integers a, b with $4a^3 + 27b^2$ coprime to n. And suppose we are trying to use the elliptic curve group law, where we are considering points with coordinates mod n. If p, q are two different primes dividing n, we are implicitly working simultaneously with elliptic curves over the two finite fields $\mathbf{F}_p$ and $\mathbf{F}_q$. It may well occur that we end up adding two points that sum to 0 in the elliptic curve over $\mathbf{F}_p$, but do not in the elliptic curve over $\mathbf{F}_q$. This mismatch would be detected by trying to find the multiplicative inverse of some residue mod n that is divisible by p, but not divisible by q. Such residues do not have inverses! In this case, the algorithm used to try to find an inverse (an elaboration of Euclid's gcd algorithm) actually finds a nontrivial divisor of n instead. And now our true motive is revealed—all we wanted was some nontrivial factor of n, and here it is.

The elliptic curve method chooses random elliptic curves each with a nontrivial point P on it. One then attempts to compute the point mP, where m has many divisors (say the least common multiple of the integers to some point). Our computation for mP is tried using arithmetic mod n and with an analogue of the powermod method discussed in the previous section. As mentioned, the process of adding two points that do not sum to 0 involves inverting, which in turn involves taking a gcd with n. With luck, one will encounter an example where the inversion doesn't work, which will then reveal a nontrivial factor of n. If m is divisible by the order of the elliptic curve group over $\mathbf{F}_p$, but not by the order of the elliptic curve group over $\mathbf{F}_q$, then there is a good chance for the method to work. So this is why one uses numbers m with lots of divisors. When one changes the curve, one also can change the group orders, and so have a fresh chance to factor, if prior chances haven't panned out.

Heuristically, the expected time for the elliptic curve method to find the least prime factor p of n is

$$\exp\left((1 + o(1))\sqrt{2\log p\,\log\log p}\right)$$

arithmetic operations mod n. So, unlike the sieve methods, the running time for the elliptic curve method depends strongly on the size of the prime to be discovered.

For much more on these factorization methods, many other methods, and original references, the reader is referred to [5].

4 The Riemann Hypothesis and the distribution of the primes

As a teenager looking at a modest table of primes, Gauss conjectured that their frequency decays logarithmically and that $\mathrm{li}(x) = \int_2^x \mathrm{d}t/\log t$ should be a good approximation for $\pi(x)$, the number of primes in $[1, x]$. Sixty years later, Riemann showed how Gauss's conjecture could be proved assuming that the Riemann zeta function $\zeta(s) = \sum_n n^{-s}$ has no zeros in the complex half-plane where the real part of s is greater than 1/2. The series for $\zeta(s)$ converges only for $\mathrm{Re}\, s > 1$, but it may be analytically continued to $\mathrm{Re}\, s > 0$, with a simple pole at $s = 1$. This continuation may be seen quite concretely via the identity $\zeta(s) = s/(s-1) - s\int_1^\infty \{x\}x^{-s-1}\,\mathrm{d}x$, with $\{x\}$ the fractional part of x (so that $\{x\} = x - [x]$), noting that this integral converges quite nicely in the half plane $\mathrm{Re}\, s > 0$. In fact, via Riemann's functional equation mentioned below, $\zeta(s)$ can be continued to a meromorphic function in the whole complex plain, with the single pole at $s = 1$.

The assertion that $\zeta(s) \neq 0$ for $\mathrm{Re}\, s > 1/2$ is known as the Riemann Hypothesis; arguably it is the most famous unsolved problem in mathematics. Though Hadamard and de la Vallee Poussin in 1896 (independently) were able to prove a weak form of Gauss's conjecture (known as the prime number theorem), the uncanny apparent strength of the approximation $\mathrm{li}(x)$ to $\pi(x)$ is breathtaking. For example, take $x = 10^{22}$. We have

$$\pi\left(10^{22}\right) = 201{,}467{,}286{,}689{,}315{,}906{,}290$$

exactly, and to the nearest integer, we have

$$\mathrm{li}\left(10^{22}\right) \approx 201{,}467{,}286{,}691{,}248{,}261{,}497.$$

As you can plainly see, Gauss's guess is right on the money!

The numerical computation of $\mathrm{li}(x)$ is simple via numerical methods for integration, and it is directly obtainable in various mathematics computing packages. However, the computation of $\pi(10^{22})$ above is far from trivial. It would be far too laborious to count these $\approx 2 \times 10^{20}$ primes one by one, so how are they counted? In fact, we have various combinatorial tricks to count without listing everything. For example, one does not need to count one by one to see that there are exactly $2[10^{22}/6] + 1$ integers in the interval $[1, 10^{22}]$ that are coprime to 6. Rather one thinks of these numbers grouped in blocks of 6, with 2 in each block coprime to 6. (The "+1" comes from the partial block at the end.) Building on early ideas of Meissel and Lehmer, Lagarias, Miller, and Odlyzko presented an elegant combinatorial method for computing $\pi(x)$ that takes about $x^{2/3}$ elementary steps. This method was refined by Deleglise and Rivat, and then Gourdon found a way of distributing the computation to many computers. The computation of $\pi(10^{22})$ above is due to him.

From work of von Koch, and later Schoenfeld, we know that the Riemann Hypothesis is *equivalent* to the assertion that

$$|\pi(x) - \mathrm{li}(x)| < \sqrt{x}\log x \qquad (2)$$

for all $x \geq 3$ (see [5], Exercise 1.37). Thus, the mammoth calculation of $\pi\left(10^{22}\right)$ might be viewed as computational evidence for the Riemann Hypothesis—in fact, if the count had turned out to violate (2), we would have had a disproof.

It may be a bit opaque what (2) has to do with the location of the zeros of $\zeta(s)$. There are zeta zeros at the negative even integers, the so-called trivial zeros. The nontrivial zeros ρ are known to be infinite in number, and as mentioned above, are conjectured to satisfy $\mathrm{Re}\,\rho \leq 1/2$. There is a certain symmetry among these zeros ρ, indeed $\bar{\rho}$, $1-\rho$, and $1-\bar{\rho}$ are also zeros, so the Riemann Hypothesis is the assertion that all of these nontrivial zeros satisfy $\mathrm{Re}\,\rho = 1/2$. (It is this symmetry with ρ and $1-\rho$, which follows from Riemann's functional equation $\zeta(1-s) = 2(2\pi)^{-s}\cos(\frac{1}{2}\pi s)\Gamma(s)\zeta(s)$, that perhaps lends some heuristic support for the Riemann Hypothesis.)

The connection to prime numbers begins with the fundamental theorem of arithmetic, which yields the identity

$$\begin{aligned}\zeta(s) \;=\; \sum_{n=1}^{\infty} n^{-s} \;&=\; \prod_{p \text{ prime}} \sum_{j=0}^{\infty} p^{-js} \\ &=\; \prod_{p \text{ prime}} \left(1 - p^{-s}\right)^{-1},\end{aligned}$$

where the product converges when $\mathrm{Re}\, s > 1$. Thus, taking the logarithmic derivative, we have

$$\frac{\zeta'(s)}{\zeta(s)} = -\sum_{p \text{ prime}} \frac{\log p}{p^s - 1} = -\sum_{p \text{ prime}} \sum_{j=1}^{\infty} \frac{\log p}{p^{js}}.$$

That is, if $\Lambda(n) = \log p$ if $n = p^j$, where p is prime and $j \geq 1$, and $\Lambda(n) = 0$ if n is not in this form, we have

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

Through various relatively routine calculations, one can then relate the function

$$\psi(x) = \sum_{n \leq x} \Lambda(n)$$

to the residues at the poles of ζ'/ζ, which correspond to the zeros (and single pole) of ζ. In fact, as Riemann showed, we have the following beautiful formula:

$$\psi(x) = x - \sum_{\rho} \frac{x^\rho}{\rho} - \log(2\pi) - \frac{1}{2} \log\left(1 - x^{-2}\right)$$

if x itself is not a prime or prime power, and where the sum over the nontrivial zeros ρ of ζ is to be understood in the symmetric sense where we sum over those ρ with $|\mathrm{Im}\,\rho| < T$ and let $T \to \infty$. An understanding of the function $\psi(x)$ gives readily, through elementary manipulations, an equal understanding of $\pi(x)$, and it should be clear now that $\psi(x)$ is intimately connected to the nontrivial zeros ρ of ζ.

The function $\psi(x)$ defined above has an extraordinarily simple interpretation. It is the logarithm of the least common multiple of the integers in the interval $[1, x]$. As with (2) we have an elementary translation of the Riemann Hypothesis: it is equivalent to the assertion that

$$|\psi(x) - x| < \sqrt{x} \log^2 x$$

for all $x \geq 3$. This inequality involves only the elementary concepts of least common multiple, natural logarithm, absolute value, and square root, yet it is an equivalent formulation of the celebrated Riemann Hypothesis.

We have actually calculated a number of nontrivial zeros ρ of $\zeta(s)$ and have verified that they lie on the line $\mathrm{Re}\, s = 1/2$. One might wonder how someone can computationally verify that a complex number ρ has $\mathrm{Re}\,\rho = 1/2$. For example, suppose that we are carrying calculations to (an unrealistically large) 10^{10} significant digits, and suppose we come across a zero with real part $1/2 + 10^{-10^{100}}$. It would be far beyond the precision of the calculation to be able to distinguish this number from $1/2$ itself. Nevertheless, we do have a method for seeing if particular zeros ρ satisfy $\mathrm{Re}\,\rho = 1/2$. There are two ideas involved. First, we can use the argument principle from complex analysis to count the number of zeros ρ with $0 < \mathrm{Im}\,\rho < T$ for various values of T. Next, we can construct a real function $g(t)$ which is 0 if and only if $\zeta(1/2 + it) = 0$. We can then calculate sign changes for $g(t)$ in the interval $(0, T)$, inferring a lower bound for the number of zeros of ζ with $0 < \mathrm{Im}\,\rho < T$ and $\mathrm{Re}\,\rho = 1/2$. If this lower bound is equal to the actual count, then the Riemann Hypothesis has been verified up to height T. If the counts do not match, it may be that the Riemann Hypothesis fails, or that a fleeting sign change was missed, or that some zero occurs with multiplicity larger than 1. So far none of these events has occurred.

The first few nontrivial zeros were computed by Riemann himself. The famous cryptographer and early computer scientist Turing also computed some zeta zeros. The current record for this kind of calculation is held by Gourdon, who has shown that the first 10^{13} zeta zeros with positive imaginary part, all have real part equal to $1/2$, as predicted by Riemann. Gourdon's method is a modification of that pioneered by Odlyzko and Schönhage [11], who ushered in the modern age of zeta-zero calculations.

Explicit calculations with the zeta function can lead to highly useful explicit estimates for prime numbers. If p_n is the nth prime, then the prime number theorem implies that $p_n \sim n \log n$ as $n \to \infty$. Actually, there is a secondary term of order $n \log \log n$, and so for all sufficiently large n, we have $p_n > n \log n$. By using explicit zeta estimates, Rosser was able to put a numerical bound on the "sufficiently large" in this statement, and then by checking small cases, was able to prove that in fact $p_n > n \log n$ in all cases. The paper [13] of Rosser and Schoenfeld is filled with this kind of numerically explicit and highly useful inequalities.

Let us assume for a moment that the Riemann Hypothesis has been proved. What would be the next question? Probably it would be the vertical distribution on the line $\operatorname{Im} s = 1/2$ of the nontrivial zeros ρ. As mentioned, we have a fairly concise understanding of how many zeros there should be up to a given height T. In fact, as already found by Riemann, this count is within $O(\log T)$ of $\frac{1}{2\pi}T \log\left(\frac{1}{2\pi e}T\right)$. Ignoring the error, the derivative of this function of T is $\frac{1}{2\pi}\log\left(\frac{1}{2\pi}T\right)$, so the reciprocal of this quantity should be the average zero spacing at height T. Thus, if the consecutive zeros are $1/2 + i\gamma_n$, where $0 < \gamma_1 < \gamma_2 < \dots$, then we can normalize the spacings by considering the statistic

$$\delta_n = (\gamma_{n+1} - \gamma_n)\frac{1}{2\pi}\log\left(\frac{1}{2\pi}\gamma_n\right).$$

The numbers δ_n thus give normalized spacings between consecutive zeta zeros. Though on average they are about 1, they may sometimes be small and they may sometimes be large. That is, they should have a distribution. Perhaps Poisson? Maybe Gaussian (or maybe $\log \delta_n$ is Gaussian)? Something else? Well, we can think and we can compute. Let's try thinking first.

Perhaps the Riemann zeta function is an isolated example, or perhaps it can be viewed in some sort of broader context. In the early 20th century, Hilbert and Pólya suggested that the zeros of the zeta function might correspond to eigenvalues of some operator. Now this is provocative!, but also somewhat vague. What operator? What space? Some 50 years later in a now famous conversation between Dyson and Montgomery at the Institue for Advanced Study, it was conjectured that the nontrivial zeros behave like the eigenvalues of a random matrix from the Gaussian Unitary Ensemble. Now known as the GUE conjecture, it can be numerically tested in various ways. Odlyzko has done this, and he has found that the higher one looks for batches of zeros, the more persuasive is the correspondence to what the GUE conjecture predicts.

For example, consider the 1,041,417,089 numbers δ_n for n starting at 10^{23} + 17,368,588,794 (at about height 1.3×10^{22}). For each interval $(j/100, (j+1)/100]$ we can compute the proportion of these normalized gaps that lie in this interval, and plot it. If we were dealing with eigenvalues from a random matrix from the Gaussian Unitary Ensemble, we would expect these statistics to converge to the Gaudin distribution (for which there is no closed formula, but which is easily computable). Andrew Odlyzko has kindly supplied me with the graph in Figure 1, which plots the Gaudin distribution against the data just described (but leaving out every second data point to avoid clutter). Like pearls on a necklace! The fit is absolutely remarkable.

Figure 1 goes here.

It seems clear that we have been thinking correctly about these zeta zeros, but a close empirical fit to a curve does not a proof make. We await the next chapter in this development.

5 Diophantine equations and the abc conjecture

Let us move now from the Riemann Hypothesis to Fermat's Last Theorem. Until the last decade it too was considered one of the most famous unsolved problems in mathematics, once even having a mention on a *Star Trek* episode. The assertion is that the equation $x^n + y^n = z^n$ has no solutions in positive integers x, y, z, n, where $n \geq 3$. This conjecture had remained unproved for three-and-a-half centuries, till Andrew Wiles published a proof in 1995. In addition, perhaps more important than the solution of this particular diophantine equation, the centuries-long quest for a proof helped establish the field of algebraic number theory. And the proof itself established a long-sought and wonderful connection between modular forms and elliptic curves.

But do you know why Fermat's Last Theorem is true? That is, just in case you are not an expert on all of the intricacies of the proof, are you surprised that there are in fact no solutions? In fact, there is a fairly simple heuristic argument that supports the assertion. First note that the case $n = 3$, namely $x^3 + y^3 = z^3$, can be handled by elementary methods, and this in fact had already been done by Euler. So, let us focus on the cases when $n \geq 4$.[1] Let $\mathcal{S}_n$ be the set of positive nth powers of integers. How likely is it that the

[1] Actually, Fermat himself had a simple proof in the case $n = 4$, but we ignore this.

sum of two members of $\mathcal{S}_n$ is itself a member of $\mathcal{S}_n$? Well, not at all likely, since Wiles has proved that this never occurs! But, recall we are trying to think naively. First notice that if there is one solution then there are infinitely many. Indeed, if $x^n + y^n = z^n$, then for any positive integer a, we have $(ax)^n + (ay)^n = (az)^n$. So, let us only consider primitive solutions x, y, z which have no common factor larger than 1. Not seeing any special reason for or against the sum of two coprime members of $\mathcal{S}_n$ being itself a member of $\mathcal{S}_n$, let us assume that the chances for this occurring are as likely as a random set of integers with the same rough distribution. In particular, in the interval $[1, x]$ there are $[x^{1/n}]$ members of $\mathcal{S}_n$, so let us look at a random set $\mathcal{R}_n$ such that for each x, there are at most $x^{1/n}$ members in $[1, x]$.

The number of pairs $r_1, r_2 \in \mathcal{R}_n$ with $r_1 + r_2 \in (x/2, x]$ is at most $x^{2/n}$, and since any particular integer in $(x/2, x]$ is in $\mathcal{R}_n$ with probability at most $2x^{1/n-1}$, we therefore expect at most $2x^{3/n-1}$ solutions to $r_1 + r_2 = r_3$ with each $r_i \in \mathcal{R}_n$ and $r_3 \in (x/2, x]$. Now let x run through powers of 2, and let n run through the integers at least 4. We thus expect a total of at most

$$2\sum_{n\geq 4}\sum_{m\geq n} 2^{m(3/n-1)}$$

solutions to $r_1 + r_2 = r_3$ where r_1, r_2, r_3 all come from the same $\mathcal{R}_n$ for some n at least 4. (Fermat's Last Theorem is trivially true if one of the powers is 1, so I assume here that we only start looking at intervals $(2^{m-1}, 2^m]$ when $m \geq n$.) The inner sum is at most $2^{3-n}/(1 - 2^{-1/4})$, so the full expression is smaller than 13. Thus, we would expect only finitely many solutions to $r_1 + r_2 = r_3$. Further, if in our particular case we do not stumble on any small solutions, we might carry this over to the random situation, and start the m variable, or even better the n variable, at a high point indicating that we are only interested in large solutions. But the tail of a convergent series is tiny if taken sufficiently far out, so it would seem that probabilistically there really shouldn't be any solutions at all. An argument similar to this was actually published by two famous mathematicians: Erdős and Ulam [7].

One might ask how far Fermat's Last Theorem had been verified before the final proof by Wiles. In fact, the paper [2] reports a verification for all exponents n up to 4,000,000. This type of calculation, which is far from trivial, has its roots in 19th century work of Kummer, and early 20th century work of Vandiver. In fact, [2] also verifies in the same range a related conjecture of Vandiver dealing with cyclotomic fields, which conjecture may in fact be false in general.

The probabilistic thinking above, combined with computing for small cases, can carry us deeply into some very provocative conjectures. For example, the above argument carries over essentially intact to the diophantine equation $x^u + y^v = z^w$, where x, y, z are coprime positive integers and u, v, w are integers all at least 4. Here it is especially important to insist that x, y, z are coprime since there are many trivial solutions when they are not. For example, notice that $17^4 + 34^4 = 17^5$, and perhaps you can see how this generalizes. What about letting the exponents u, v, w dip down to 2 or 3? Of course we must be wary about cases where we already know plenty of solutions, such as $x^2 + y^2 = z^2$. But a moment's reflection should convince you that if the exponent trio u, v, w satisfies $1/u + 1/v + 1/w < 1$, then the Erdős–Ulam argument carries over, and there should be at most finitely many solutions in all.

Now our main topic of computing comes in. There actually *are* solutions! For example, since $1 + 8 = 9$, we have a solution to $x^7 + y^3 = z^2$, where $x = 1$, $y = 2$, and $z = 3$. (The exponent 7 is chosen to insure that the reciprocal sum of the exponents is less than 1. Actually any integer at least 7 works, but since in each case the power involved is the number 1, they should all together be considered as just one example.) The so-called Fermat–Catalan conjecture is that there are at most finitely many powers x^u, y^v, z^w, with $\gcd(x, y, z) = 1$, $1/u + 1/v + 1/w < 1$, and

$$x^u + y^v = z^w. \tag{3}$$

This generalization of Fermat's Last Theorem is also a generalization of the Catalan Conjecture (recently proved by Mihăilescu) that 8 and 9 are the only consecutive nontrivial powers. Other solutions to (3) are found in the following short list:

$$\begin{aligned}
2^5+7^2 &= 3^4,\\
13^2+7^3 &= 2^9,\\
2^7+17^3 &= 71^2,\\
3^5+11^4 &= 122^2,\\
33^8+1549034^2 &= 15613^3,\\
1414^3+2213459^2 &= 65^7,\\
9262^3+15312283^2 &= 113^7,\\
17^7+76271^3 &= 21063928^2,\\
43^8+96222^3 &= 30042907^2.
\end{aligned}$$

The larger members were found in an exhaustive computer search by Beukers and Zagier. Perhaps with $1+8=9$, this is the complete list of all solutions, or perhaps not—we have no proof.

However, for particular choices of u, v, w, more can be said. For example, using some of the tools developed by Wiles, Darmon and Granville [6] have shown that for any fixed choice of u, v, w with reciprocal sum at most 1, there are only finitely many coprime solutions to (3). Handling the case of a particular exponent triple, if it can be handled at all, often involves a delicate interplay between arithmetic geometry, effective methods in transcendental number theory, and good hard computing. In particular, the exponent triple sets $\{2,3,7\}$, $\{2,3,8\}$, $\{2,3,9\}$, and $\{2,4,5\}$ are known to have all their solutions in the above table. See [12] for the treatment of the case $\{2,3,7\}$ and links to other work.

The "abc conjecture" of Oesterlé and Masser is deceptively simple. It involves positive integer solutions to the equation $a+b=c$, hence the name. It is perhaps surprising that anything interesting can be said about this equation! But notice that the Fermat–Catalan equation (3) is a special case, and this equation already contains Fermat's Last Theorem within it. To put some meaning into $a+b=c$, we define the "radical" of a nonzero integer n as the product of the primes that divide n, denoting this as $\mathrm{rad}(n)$. So, for example, $\mathrm{rad}(-10)=10$, $\mathrm{rad}(72)=6$, and $\mathrm{rad}(65536)=2$. In particular, high powers have small radicals in comparison to the number itself, and so do many other numbers. Basically the abc conjecture asserts that if $a+b=c$, then the radical of abc cannot be too small. More specifically:

The abc conjecture. *For each $\epsilon > 0$ there are at most finitely many coprime positive integer triples a, b, c with $a+b=c$ and $\mathrm{rad}(abc) < c^{1-\epsilon}$.*

Note that the abc conjecture immediately solves the Fermat–Catalan problem. Indeed if u, v, w are positive integers with $1/u+1/v+1/w<1$, then it is easily found that we must have $1/u+1/v+1/w \le 41/42$. Suppose we have a coprime solution to (3). Then $x \le z^{w/u}$ and $y \le z^{w/v}$, so that

$$\mathrm{rad}(x^u y^v z^w) \le xyz \le (z^w)^{41/42}.$$

Thus, the abc conjecture with $\epsilon = 1/42$ implies that there are at most finitely many solutions.

The abc conjecture has many other marvelous consequences; for a delightful survey, see [8]. In fact, the abc conjecture and its generalizations can be used to prove so many things, that I have joked that it is beginning to resemble a false statement, since a false statement implies everything. But probably the abc conjecture is true. Indeed, though a bit harder to see, the same type of probabilistic argument as Erdős and Ulam gave for Fermat's Last Theorem works here as well.

We begin with a perfectly rigorous proposition that quantifies the distribution of integers with a small radical. Let $R(x,y)$ denote the number of positive integers $n \le x$ with $\mathrm{rad}(n) \le y$. Then for y sufficiently large and $x \ge y$, we have

$$R(x,y) \le y \exp\left(3\sqrt{\frac{\log x \, \log\log y}{\log\log x}}\right). \tag{4}$$

This rather forbidding expression arises quite naturally. One first fixes a particular integer $r \le y$, and asks how many $n \le x$ can have $\mathrm{rad}(n) = r$. Say r has the prime factorization $q_1 q_2 \cdots q_k$, where these prime factors are all different. Then $\mathrm{rad}(n) = r$ if and only if $n = q_1^{a_1} q_2^{a_2} \cdots q_k^{a_k}$ for positive integers $a_1, a_2, \ldots, a_k$. But $n \le x$, so we are counting the number of lattice points $(a_1, a_2, \ldots, a_k)$ strictly in the first orthant and lying in the simplex $\sum z_i \log q_i \le \log x$. This number is clearly majorized by the volume of the simplex, since each such lattice point is the "upper right" vertex of some unit hypercube lying totally within the simplex. This volume is

$$\frac{(\log x)^k}{k! \, \log q_1 \log q_2 \cdots \log q_k}.$$

By replacing each q_i with the ith prime, denote it p_i, we can only make this volume larger. Further, using the explicit inequalities in [13] mentioned in the preceding section, one can prove that

$$\log p_1 \log p_2 \cdots \log p_k \geq \frac{1}{4}(\log k)^k$$

for all $k \geq 3$. Summarizing, if r has k prime factors and $k \geq 3$, then the number of $n \leq x$ with $\operatorname{rad}(n) = r$ is at most $4(\log x/\log k)^k/k!$. Further, an elementary and easily proved inequality of Hardy and Ramanujan asserts that the number of positive integers $r \leq y$ with exactly k different prime factors is at most $c_1(y/\log y)(\log\log y + c_2)^{k-1}/(k-1)!$ for certain explicit constants c_1, c_2. Putting these inequalities together, it is now routine to get (4).

Now we connect (4) heuristically to the abc conjecture. Suppose i, j, k are positive integers, and consider coprime positive integers $a, b, c \leq x$ with $\operatorname{rad}(a) \leq e^i$, $\operatorname{rad}(b) \leq e^j$, $\operatorname{rad}(c) \leq e^k$. Using (4) we have quite rigorously that the number of such triples is at most

$$e^{i+j+k} \exp\left(l\left(\sqrt{\log i} + \sqrt{\log j} + \sqrt{\log k}\right)\right),$$

where $l = 3\sqrt{\log x/\log\log x}$. Say now we sum this expression over all integer triples i, j, k with $i+j+k \leq (1-\delta)\log x$. There are less than $(\log x)^3$ such triples, so the total number of choices for $a, b, c \leq x$ with $\operatorname{rad}(abc) \leq x^{1-\delta}$ is at most

$$x^{1-\delta}(\log x)^3 \exp\left(9\sqrt{\log x}\right).$$

For a random choice of $a, b, c \leq x$ and $c > x/2$, the chance that $a + b = c$ should be about $1/x$. Let $\delta = 10/\sqrt{\log x}$. Thus, we might expect at most

$$(\log x)^3 \exp\left(\sqrt{\log x}\right)$$

solutions to $a+b=c$ with these constraints. Here we have $c \in (x/2, x]$, and now letting x run through powers of 2, we cover all possibilities. Since the sum of the last displayed expression as x runs over powers of 2 is convergent, we thus should expect at most finitely many solutions. This argument suggests then that if $a+b=c$ are coprime positive integers and c is sufficiently large, then we have

$$\operatorname{rad}(abc) > c^{1-10/\sqrt{\log c}}, \qquad (5)$$

which implies the abc conjecture.

A stronger version of this heuristic argument appears in the thesis of van Frankenhuijsen [15].

One might wonder how the numerical evidence stacks up against (5). This inequality asserts that if a, b, c are coprime, $a + b = c$, and $\operatorname{rad}(abc) = r$, then $\log(c/r)/\sqrt{\log c} < 10$. So, let $T(a,b,c)$ denote the test statistic $\log(c/r)/\sqrt{\log c}$. A website maintained by Nitaj [10] contains a wealth of information about the abc conjecture. Checking the data there for large values of $T(a,b,c)$, the largest example we find is

$$\begin{aligned} a &= 7^2 \cdot 41^2 \cdot 311^3 = 2477678547239 \\ b &= 11^{16} \cdot 13^2 \cdot 79 = 613474843408551921511 \\ c &= 2 \cdot 3^3 \cdot 5^{23} \cdot 953 = 613474845886230468750, \end{aligned}$$

where

$$\begin{aligned} \operatorname{rad}(abc) &= 2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 \cdot 41 \cdot 79 \cdot 311 \cdot 953 \\ &= 28828335646110, \\ T(a,b,c) &= \frac{\log(c/\operatorname{rad}(abc))}{\sqrt{\log c}} = 2.43886\ldots. \end{aligned}$$

Is it always true that $T(a,b,c) < 10$?

One can get carried away with heuristics, forgetting that one is not actually proving a theorem, but making a guess. Heuristics are often based on the idea of randomness, and all bets are off if there is some underlying structure. But how do we know that there is no underlying structure? For example, the above heuristics for the abc conjecture work perfectly well if one considers more than 3 numbers. So let us consider an "abcde" conjecture. Here the heuristics suggest that if a, b, c, d, e are nonzero integers that are coprime in pairs, and $a+b+c+d+e = 0$, then for each $\epsilon > 0$, we have

$$\operatorname{rad}(abcde)^{1+\epsilon} < \max\{|a|, |b|, |c|, |d|, |e|\}$$

for at most finitely many cases. But consider the polynomial identity

$$(1-x)^5 + (1+x)^5 = 10(x^2+1)^2 - 8$$

(kindly supplied to me by Granville). Say $x = 7^{4k} - 1$ and write $8 = 9 - 1$, so that we have an abcde example where the 5 numbers are pairwise coprime, the largest in absolute value is 7^{20k}, and

$$\begin{aligned} \operatorname{rad}(abcde) &\leq 210\left(7^{4k} - 2\right)\left((7^{4k} - 1)^2 + 1\right) \\ &< 210 \cdot 7^{12k}. \end{aligned}$$

The heuristics are saying that this cannot be, yet here it is right before our eyes!

What is happening here is that the polynomial identity is supplying an underlying structure. Granville suggests that for a 4-term abcd conjecture, one should allow for small pairwise gcd's, since otherwise, e.g., one would be precluded from using even numbers. With this proviso, Granville conjectures that for each $\epsilon > 0$, all counterexamples come from at most finitely many polynomial families. And the number of polynomial families would grow to infinity as ϵ descends to 0.

There is also a generalization of the abc conjecture to many terms where one assumes only that the various terms as a group have no common prime factor, rather than looking at pairwise gcd's. See [10] for references to this and much other interesting work.

We have looked here at only a small portion of the field of diophantine equations, and then we have looked mainly at the interplay of heuristics and computational searches for small solutions. For much more on the subject of computational diophantine methods, see [14].

As mentioned at the start, this article is not meant to be encyclopedic. In particular the vast area of computational algebraic number theory has been omitted. For a wealth of material in this subject, see [3], [4].

Acknowledgments. I am grateful to the following people who generously shared their expertise: Xavier Gourdon, Andrew Granville, Andrew Odlyzko, Ed Schaefer, Cam Stewart, Rob Tijdeman, and Machiel van Frankenhuijsen. I am also grateful to David Pomerance for helpful suggestions with the exposition. I was supported in part by NSF grant DMS-0401422.

Bibliography

1. Agrawal, M., Kayal, N., and Saxena, N. 2004 PRIMES is in P. *Ann. of Math.* **160**, 781–793.
2. Buhler, J., Crandall, R., Ernvall, R., and Metsänkylä, T. 1993 Irregular primes and cyclotomic invariants to four million. *Math. Comp.* **61**, 151–153.
3. Cohen, H. 1993 A course in computational algebraic number theory. *Graduate Texts in Mathematics* **138**, Springer-Verlag, Berlin.
4. Cohen, H. 2000 Advanced topics in computational number theory. *Graduate Texts in Mathematics* **193**, Springer-Verlag, New York.
5. Crandall, R. and Pomerance, C. 2005 Prime Numbers: a computational perspective. Second edition. Springer–Verlag, New York.
6. Darmon, H. and Granville, A. 1995 On the equations $z^m = F(x,y)$ and $Ax^p + By^q = Cz^r$. *Bull. London Math. Soc.* **27**, 513–543.
7. Erdős, P. and Ulam, S. 1971 Some probabilistic remarks on Fermat's last theorem. *Rocky Mountain J. Math.* **1**, 613–616.
8. Granville, A. and Tucker T. J. 2002 It's as easy as *abc*. *Notices Amer. Math. Soc.* **49**, 1224–1231.
9. Lenstra, Jr., H. W. and Pomerance, C. (to appear) Primality testing with Gaussian periods.
10. Nitaj, A. 2005 The ABC conjecture home page. `http://www.math.unicaen.fr/~nitaj/abc.html`
11. Odlyzko, A. M. and Schönhage, A. 1988 Fast algorithms for multiple evaluations of the Riemann zeta function. *Trans. Amer. Math. Soc.* **309**, 797–809.
12. Poonen, B., Schaefer, E., and Stoll, M. (to appear) Twists of $X(7)$ and primitive solutions to $x^2 + y^3 = z^7$.
13. Rosser, J. B. and Schoenfeld, L. 1962 Approximate formulas for some functions of prime numbers. *Illinois J. Math.* **6**, 64–94.
14. Smart, N. 1998 The algorithmic resolution of Diophantine equations. *London Mathematical Society Student Texts* **41**, Cambridge University Press, Cambridge.
15. van Frankenhuijsen, M. 1995 Hyberbolic spaces and the abc conjecture. PhD thesis, Universiteit Nijmegen.