Number theory, ancient and modern

John Coates

1 Introduction

Number theory is the branch of mathematics concerned with the study of the mysterious and hidden properties of the most basic mathematical objects, namely the integers

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \cdots\},$$

and the rational numbers

$$\mathbb{Q} = \{m/n \, : \, m, \, n \in \mathbb{Z} \text{ and } n \neq 0\}.$$

It is the oldest part of mathematics, having its origins somewhere in Asia long before Greek mathematics (e.g. triples of integers, which are the side lengths of right-angled triangles, occur in Babylonian Cuneiform texts dating from 1900-1600 BC, and in Indian sutras dating from about 800 BC). Since the earliest time until the present day, it has been an experimental science. Number theorists look for the appearance of unexpected patterns and laws in numerical data, and then try to formulate general conjectures. Some of the many unproven conjectures are very old, including one we shall discuss, which can be traced back to Arab manuscripts a thousand years ago. The hardest part of number theory is to find proofs of conjectures, or more usually proofs of partial results in support of these conjectures. When proofs have been found in the past, they have nearly always relied on major conceptual advances in seemingly unrelated frontiers of mathematical knowledge. It should also be stressed that there are central parts of number theory, notably the study of the continued fraction algorithm, where there has been no important progress since the 18th century. For example, we know essentially nothing about the continued fraction expansion of real algebraic numbers of degree > 2. Also, even though the trascendence of π was proven by Hermite and Lindemann in the latter part of the 19th century, its continued fraction expansion remains a mystery. Finally, there is one genuinely new feature of research in number theory today, namely the ever growing links with the algorithmic side of

computer science. Indeed, computer science provides almost unlimited possibilities for numerical experimentation by researchers in number theory. On the other hand, number theory provides the basis of the principal cryptosystems which are secure against attack by modern computers.

Modern number theory is a vast subject, which overlaps many other areas of mathematics. In my lecture today, I want to at least touch on four major areas of number theory. All have aspects which are very old, but they are also important fields of current research. They are:-

1. Primality testing and factorization.
2. Diophantine equations
3. Distribution of prime numbers
4. Connexions between L-functions and arithmetic.

In conclusion, I wish to thank Sujatha and Carl Pomerance for their help in preparing this lecture.

2 Primality testing and Factorization

We only consider positive integers in this section. By a *factorization* of a positive integer N, we mean an expression

$$N = MR \tag{1}$$

where M and R are also integers (we then say that M and R are *divisors* of N). We say N is a *prime* if N has no non-trivial factorization i.e. 1 and N are the only divisors of N. The series of primes begins with

$$2,\ 3,\ 5,\ 7,\ 11,\ 13,\ 17,\ 19,\ 23, \cdots,$$

and it is easy to see that there are infinitely many primes (Euclid wrote down the first formal proof of this). The largest explicitly known prime at present is

$$2^{32582657} - 1,$$

which has over 9.8 million decimal digits.

By successively refining (1), it is clear that every N has a factorization

$$N = p_1 \dots p_t, \tag{2}$$

where the p_i are not necessarily distinct primes. The fundamental theorem of arithmetic (which curiously is not stated explicitly in Euclid), asserts that (2) is

unique up to the order of factors. From a naive viewpoint, the theory of factorization in $\mathbf{Z}$ is complete once we have this theorem. Moreover, there is an obvious algorithm (namely trial division by 2, then by all odd integers $\leq \sqrt{N}$) which either produces a factorization of N or tells us that N is prime.

However, the theory does not end here, because in our age of high speed computers, most systems of cryptography are devised around the fact that it is fast to multiply two large integers, but slow to factor a large integer. We do not attempt to make this completely precise, but the essential point is clear. One sees easily that the number of "elementary operations" (by such an elementary operation, we mean the addition or multiplication of two digits from the set $0, \cdots, 9$, with the remainder being carried, if necessary) to multiply two large positive integers M and R is at most $2mr$, where m (respectively r) = number of digits of M (resp. of R). But $m-1 \leq \log_{10} M$ and $r-1 \leq \log_{10} R$. Hence, assuming $R \leq M$, we see that multiplication takes at most

$$2(\log_{10} M + 1)^2$$

elementary operations. In general, a number theoretic algorithm is said to be *polynomial*, if when it is applied to one or several integers $\leq M$, it takes at most

$$c(\log\ M)^w$$

elementary operations to complete the algorithm, where c and w are positive constants not depending on M. Clearly trial division by 2 and all odd integers $\leq \sqrt{N}$ does not provide a polynomial algorithm for factoring N. This immediately suggests the following:-

Fundamental Problem 1. Is there a polynomial algorithm for factoring?

Many better algorithms than trial division have been developed by number theorists, starting with Fermat and especially over the last forty years (see the excellent book [4] by R. Crandall and C. Pomerance). But the answer to the above problem still seems far away, and could be positive or negative.

Primality testing algorithms are those which test whether or not N is prime (but do not necessarily produce a factorization of N if N is composite). Here much more is known, culminating in the important recent result:-

Theorem 2.1 *(Agarwal, Kayal, Saxena)[1]. There exists a polynomial algorithm for primality testing.*

The proof of the above theorem, and a beautiful earlier conditional result, all rely on congruence techniques. Let m be a positive integer. We say two integers a

and b are *congruent* modulo m, written $a \equiv b \bmod m$, if m divides $a - b$. It is very fast to carry out multiplication and addition modulo m. Moreover, unlike exponentiation in $\mathbb{Z}$ itself, exponentiation modulo m can easily be shown to be a polynomial algorithm. The following remarkable, but conditional, polynomial primality testing algorithm using congruences modulo N is due to Miller and Rabin. We can suppose N is odd, and write

$$N - 1 = 2^s t \quad (s \geq 1,\, t \text{ odd}).$$

Let w be any integer in $\{2, \cdots, N-1\}$. We say w is a *witness* for N if at least one of the following congruences does not hold:-

$$\begin{aligned} w^t &\equiv 1 \mod N \\ w^{2^i t} &\equiv -1 \mod N \text{ for some } i \text{ with } 0 \leq i \leq s-1. \end{aligned}$$

It is easy to see that N is composite if and only if there exists a witness for N in the set $\{2, \cdots, N-1\}$. If N is composite, let $W(N)$ denote the smallest witness for N in this set. What is deep and remarkable is the following theorem.

Theorem 2.2 *(Ankeny, Bach)[2]. Assume the Riemann hypothesis for all Dirichlet L-functions. Then, for all odd composite N, we have*

$$W(N) < 2 \log^2 N.$$

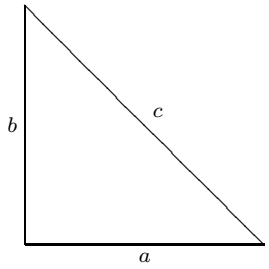
Here the Dirichlet L-functions, which are functions of a complex variable s, enter mysteriously into the study of this elementary arithmetic problem. We shall see this phenomenon occurring for other basic problems.

3 Diophantine equations

Diophantine equations refer to the study of the solutions of polynomial equations in two or more variables in either integers or rational numbers, rather than in real or complex numbers. A beautiful example of the subject, which we now briefly describe, turns out to be the oldest unsolved major problem in number theory, and possibly in the whole of mathematics.

A positive integer D is defined to be a *congruent number* if it is the area of a right angled triangle, all of whose sides have lengths in $\mathbb{Q}$ (the problem is non-trivial only because we insist that all three sides have rational lengths). For example, 5, 6, and 7 are all congruent numbers, because of the existence of the right

angled triangles



with (a, b, c) given respectively by $(40/6, 9/6, 41/6)$, $(3, 4, 5)$, and $(288/60, 175/60, 337/60)$. Note that, because of similarity considerations, we need only consider square free positive integers in determining whether a number is congruent or not. The tenth century Arab manuscript of al-Kazin in the Bibliothéque Nationale de Paris has tables showing that

5, 6, 14, 15, 21, 30, 34, 65, 70, 110, 154, 190, 210, 221, 231, 286, 330, 390, 429, 546, 1155, 1254, 1785, 1995, 2730, 3570, 4290, 5610, 7854, 10374

are all congruent. The first known European manuscript related to congruent numbers is Fibonacci's book *Liber Quadratorum* published in 1225, in which he points out that both 5 and 7 are congruent, and claims without proper justification that 1 is not. Of course, today vast tables of congruent numbers exist, and what remains of great interest are the theoretical problems.

Fundamental Problem 2. Prove the existence of an algorithm which decides in a finite number of steps whether or not a given positive integer is congruent.

In fact, a simple algorithm is conjectured, coming from the mysterious connexion of this purely arithmetic problem with L-functions (cf. the last part of this lecture). But it seems very difficult to prove this algorithm at present. Perhaps even more surprising from a naive point of view, but again simply explained via the conjectural connexion with L-functions, is:-

Fundamental Problem 3. Prove that every positive integer which is of the form $8n + 5$, $8n + 6$, or $8n + 7$ is congruent.

These problems are linked with diophantine equations, thanks to the following elementary lemma:-

Lemma 3.1 . *Let D be an integer ≥ 1. Then D is congruent if and only if there exists a point (x, y), with x, y in $\mathbb{Q}$ and $y \neq 0$ on the elliptic curve*

$$E_D \ : \ y^2 = x^3 - D^2 x. \tag{3}$$

We owe to Fermat, in the middle of the 17th century, a marvellous proof that 1 is not congruent. Not only did this proof introduce ideas that had a vast development in the 20th century, but Fermat noted that his proof also showed that there are no integers $x,\ y,\ z$ with $xyz \neq 0$ such that

$$x^4 + y^4 = z^4.$$

He subsequently went on to state, without proof, that there are no integers $x,\ y,\ z$ with $xyz \neq 0$ such that

$$x^n + y^n = z^n \tag{4}$$

when n is any integer ≥ 3.

Two major successes of recent research in number theory have been on diophantine equations.

Theorem 3.2 *(A. Wiles)[9]. There are no solutions in integers $x,\ y,\ z$ with $xyz \neq 0$ of* (4) *when $n \geq 3$.*

Theorem 3.3 *(P. Mihailescu)[7][8]. The only solution of the equation*

$$x^m - y^n = 1,$$

in positive integers $x,\ y,\ m,\ n$, with $m,\ n \geq 2$ is

$$x = 3,\ y = 2,\ m = 2,\ n = 3.$$

Both proofs make heavy use of ideas which have their origins in the great German school of number theory of the 19th century. In particular, Wiles' work used the theory of modular functions and forms which began with this school. Mihailescu's proof is based on Kummer's beautiful ideas about the arithmetic of cyclotomic fields, which Kummer had developed to study, with only limited success, the Fermat equation (4).

4 Distribution of prime numbers

If x is a real number ≥ 2, we define

$$\pi(x) = \text{number of primes } p \text{ with } p \leq x.$$

Thanks to the help of high speed computers, we know today the value of $\pi(x)$ for $x \leq 10^{21}$. In the late 18th century, Gauss, by examining tables of prime numbers, conjectured that a good approximation to $\pi(x)$ should be given by the function

$$\mathrm{li}(x) = \int_2^x dt/\log\ t.$$

All subsequent numerical work has shown what an extraordinarily good approximate estimate this is! Numerically $\pi(x) < \mathrm{li}(x)$ for all $x \leq 10^{21}$, but Littlewood proved that the function $\pi(x) - \mathrm{li}(x)$ changes sign infinitely often.

In the second half of the 19th century, Hadamard and de la Vallée Poussin proved the celebrated prime number theorem asserting that

$$\lim_{x\to\infty} \pi(x)/\mathrm{li}(x) = 1.$$

They studied the Riemann zeta function, which is the analytic continuation of the series

$$\zeta(s) = \sum_{n=0}^{\infty} 1/n^s \ \ (\mathrm{Re}(s) > 1),$$

and proved that $\zeta(s)$ has no zero on the line $\mathrm{Re}(s) = 1$. Their work was heavily influenced by Riemann's great memoir of 1859. In this memoir, Riemann made the celebrated conjecture:-

Fundamental Problem 4 (Riemann Hypothesis). All zeroes of $\zeta(s)$ in the half plane $\mathrm{Re}(s) > 0$ lie on the line $\mathrm{Re}(s) = 1/2$.

Thanks to a blend of theory and modern computing power, it is now known that the first 10^{13} zeroes of $\zeta(s)$ in $\mathrm{Re}(s) > 0$ satisfy $\mathrm{Re}(s) = 1/2$. Morevover, it is known that the Riemann Hypothesis is equivalent to the assertion that

$$\mid \pi(x) - \mathrm{li}(x) \mid \leq x^{1/2} \log x \ \ \ (x \geq 3).$$

Despite intensive efforts by many mathematicians, a proof of the Riemann Hypothesis seems as far away as ever today.

Numerical examples of finite arithmetic progressions of primes have long been noticed, for example

$$7,\ 37,\ 67,\ 97,\ 127,\ 157.$$

However, it was only very recently that the following striking theoretical result was established.

Theorem 4.1 *(Green and Tao)[6]. There exist arbitrarily long arithmetic progressions of prime numbers.*

Many other old conjectures about primes still resist theoretical attack, even though the numerical and theoretical evidence in support of them seems overwhelming. For example, it is unknown whether there are infinitely many prime pairs $(p, p+2)$, or whether there are infinitely many primes of the form $n^2 + 1$.

5 Connection between L-functions and arithmetic

We have already seen deep unproven assertions about primality testing, congruent numbers, and the difference $\pi(x) - \mathrm{li}(x)$, which are implied by unproven assertions about L-functions. If the L-functions were not there, we would still have noticed the same phenomena from numerical data, but there would be no theoretical reason why we would expect them to be always true. The simplest proven example of this mysterious phenomenon was discovered by Dirichlet in 1837, as an unexpected by-product of his proof of the theorem:-

Theorem 5.1 *(Dirichlet). Let a, b be relatively prime positive integers. Then there are infinitely many primes of the form $an + b$, $(n = 1, 2, \cdots)$.*

Here is a curious consequence of his beautiful proof. Let p be any odd prime, and consider the set

$$P = \{1, \cdots, p-1\}.$$

If w is any element of P, we say that w is a *quadratic residue* mod p if there exists an integer x such that $x^2 \equiv w \mod p$; if no such x exists, we say that w is a *quadratic non-residue* mod p. For example, if $p = 19$, the quadratic residues mod 19 are given by

$$1,\ 4,\ 5,\ 6,\ 7,\ 9,\ 11,\ 16,\ 17.$$

In general, it is easy to see that P always as exactly $(p-1)/2$ quadratic residues and $(p-1)/2$ quadratic non-residues. As p varies, we would expect the residues and non-residues to be distributed at random in the set P. But Dirichlet noticed that his work on L-functions proved that this is not the case for the primes p with $p \equiv 3 \mod 4$. Let $\mathfrak{P}$ be the subset of P given by

$$\mathfrak{P} = \{1, \cdots, (p-1)/2\}.$$

Theorem 5.2 *(Dirichlet). For every prime p with $p \equiv 3 \mod 4$, the set $\mathfrak{P}$ contains more quadratic residues than non-residues mod p.*

No proof of this simple statement, which does not use L-functions, has ever been found!

To prove these results, Dirichlet introduced a generalisation of the Riemann-zeta function, which we call the *Dirichlet L-function*. Here is one of the key examples of such an L-function. Let p be an odd prime number. We define $\left(\frac{n}{p}\right)$ for $n \in \mathbb{Z}$ to be zero if p divides n, $+1$ if $(n, p) = 1$ and n is a quadratic residue

mod p, and -1 if $(n,p) = 1$ and n is a quadratic non-residue mod p. Dirichlet then defined the L-function

$$L(s, \left(\frac{\cdot}{p}\right)) = \sum_{n=1}^{\infty} \left(\frac{n}{p}\right) / n^s \quad (\mathrm{Re}(s) > 1).$$

Unlike the Riemann-zeta function, this function does not have a pole at $s = 1$. In fact, Dirichlet found a beautiful closed formula for its value at $s = 1$, from which the above result follows easily. But in all other ways, this L-function seems to behave like $\zeta(s)$. It has an analytic continuation to the whole complex plane. The *generalised Riemann hypothesis* asserts that all of its zeroes in $\mathrm{Re}(s) > 0$ should lie on the line $\mathrm{Re}(s) = 1/2$. Needless to say, a proof of this assertion seems as distant as the classical Riemann hypothesis.

In the 20th century, it was realised that it was important to generalize the notion of zeta and L-functions to a very wide class of arithmetic objects, including the elliptic curve

$$E_D : y^2 = x^3 - D^2 x,$$

which occurred earlier in our discussion of the congruent number problem. The key to defining these new L-functions is via *Euler products*, so named because Euler in the 18th century proved the identity

$$\zeta(s) = \prod_q (1 - q^{-s})^{-1} \quad (\mathrm{Re}(s) > 1),$$

where the product is taken over all prime numbers q. We sketch the definition of the L-function in the special case of E_D, but in fact it can be made in vastly greater generality. Assume that D is square-free. Let q be a prime number which does not divide $2D$. Let N_q denote the numbers of pairs (x, y) where x and y run over the integers modulo q, which satisfy the congruence

$$y^2 \equiv x^3 - D^2 x \mod q.$$

Put

$$a_q = q - N_q.$$

We then define the L-function of E_D by the Euler product

$$L(E_D, s) = \prod_{(q,2D)=1} (1 - a_q q^{-s} + q^{1-2s})^{-1}$$

where the product is taken over all primes q which do not divide $2D$. The Euler product only converges over the half plane $\mathrm{Re}(s) > 3/2$, but it can be analytically

continued over the whole complex plane, a result which in this case is essentially due to Eisenstein and Kronecker. For this function, it is the vertical line $\mathrm{Re}(s) = 1$ which plays the analogue of the line $\mathrm{Re}(s) = 1/2$ for the Riemann zeta function and the Dirichlet L-functions. Of course, we believe that every zero of $L(E_D, s)$ in $\mathrm{Re}(s) > 0$ should lie on the line $\mathrm{Re}(s) = 1$. But a totally new phenomena can occur here, as was discovered about 1960 by Birch and Swinnerton-Dyer [3]. The following is a very special case of their celebrated conjecture.

Fundamental Problem 5. Prove that there exists a point (x, y) on E_D with $x,\ y \in \mathbb{Q}$ and $y \neq 0$ if and only if $L(E_D, s)$ vanishes at $s = 1$.

One direction of this problem (the existence of a point implies the existence of a zero) was proven by Wiles and myself [5]. The proof in the other direction still seems far away. The difficulty is that if $L(E_D, s)$ vanishes at $s = 1$, then a p^∞-descent on the curve shows that, for every prime p, there appears cohomologically to be a rational point (x, y) with $\neq 0$. But no way is known at present to conclude that there does indeed exist an actual rational point giving rise to these possibly phantom cohomological points.

We end this lecture by indicating why a proof of the above conjecture would yield an answer to Fundamental Problems 2 and 3 discussed earlier. Firstly, the classical functional equation relating $L(E_D, s)$ and $L(E_D, 2 - s)$ shows that $L(E_D, s)$ has a zero of odd (resp. even) multiplicity at $s = 1$ when $D \equiv 5,\ 6,\ 7 \bmod 8$ (resp. when $D \equiv 1,\ 2,\ 3 \bmod 8$). This explains, in particular,the mysterious assertion given by Fundamental Problem 3, and we stress that there is no explanation of this assertion which does not rely on its connexion with L-functions. When $D \equiv 1,\ 2,\ 3 \bmod 8$, one has to use a closed formula for the value $L(E_D, 1)$ to decide whether or not it is zero, and fortunately at least two such closed formulae are known. We briefly explain one of these formulae, which is due to Tunnell [10]. Consider the formal power series in T given by

$$g(T) = T\prod_{n=1}^{\infty}(1 - T^{8n})(1 - T^{16n}),\ \ \theta_k(T) = 1 + 2\sum_{n=1}^{\infty} T^{2kn^2}\ (k = 1, 2).$$

By multiplying these two formal power series, we can define two sequences of integers $a(n), b(n)\ (n \geq 1)$ by the expansions

$$g(T)\theta_1(T) = \sum_{n=1}^{\infty} a(n)T^n,\ g(T)\theta_2(T) = \sum_{n=1}^{\infty} b(n)T^n.$$

Then it is proven in [10] that, for each odd positive square free integer N, we have

$$L(E_N, 1) = u(N)a(N)^2,\ \ L(E_{2N}, 1) = v(N)b(N)^2,$$

where $u(N)$ and $v(N)$ are certain explicit complex numbers which are both clearly non-zero. Thus this L-value vanishes if and only if the corresponding integer $a(N)$ or $b(N)$ vanishes. Since there is plainly a simple finite algorithm for computing $a(N)$ and $b(N)$, we have our desired algorithm for testing whether a square free positive integer is a congruent number or not. For example, $b(17) = 0$, showing that $L(E_{34}, 1) = 0$. In fact, $D = 34$ is the smallest congruent number satisfying $D \equiv 1,\ 2,\ 3 \bmod 8$.

References

[1] M. Agarwal, N. Kayal, N. Saxena, *PRIMES is in P*, Ann. of Math. **160** (2004), 781–793.

[2] E. Bach, *Explicit bounds for primality testing and related problems*, Math.Comp. **55** (1990), 355–390.

[3] B. Birch, H. Swinnerton-Dyer, *Notes on elliptic curves (II)*, Crelle **218** (1965), 79-108.

[4] R. Crandall, C. Pomerance, Prime numbers - A Computational Perspective, 2nd. Ed. Springer, (2005).

[5] J. Coates, A. Wiles, On the conjecture of Birch and Swinnerton-Dyer, Invent. Math. **39** (1977), 233-251.

[6] B. Green, T. Tao, *The primes contain arbitrarily long arithmetic progressions*, To appear in Ann. of Math.

[7] P. Mihailescu, *Primary cyclotomic units and a proof of Catalan's conjecture*, J. Reine Angew. Math. **572** (2004), 167–195.

[8] P. Mihailescu, *Reflection, Bernoulli numbers and the proof of Catalan's conjecture*, European Congress of Mathematics, 325–340, Eur. Math. Soc., Zürich, 2005.

[9] A. Wiles, *Modular elliptic curves and Fermat's last theorem*, Ann. of Math. **141** (1995), 443–551.

[10] J. Tunnell, *A classical Diophantine problem and modular forms of weight 3/2*, Invent. Math. **72** (1983), 323-334.

John Coates
Emmanuel College
Cambridge CB2 3AP,
England
J.H.Coates@dpmms.cam.ac.uk