BULLETIN (New Series) OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 42, Number 1, Pages 3–38
S 0273-0979(04)01037-7
Article electronically published on September 30, 2004

IT IS EASY TO DETERMINE WHETHER A GIVEN INTEGER IS PRIME

ANDREW GRANVILLE

Dedicated to the memory of W. 'Red' Alford, friend and colleague

ABSTRACT. "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 ... It frequently happens that the trained calculator will be sufficiently rewarded by reducing large numbers to their factors so that it will compensate for the time spent. 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* ... It is in the nature of the problem that *any* method will become more complicated as the numbers get larger. Nevertheless, in the following methods the difficulties increase rather slowly ... The techniques that were previously known would require intolerable labor even for the most indefatigable calculator."

—from article 329 of *Disquisitiones Arithmeticae* (1801) by C. F. Gauss

There are few better known or more easily understood problems in pure mathematics than the question of rapidly determining whether a given integer is prime. As we read above, the young Gauss in his first book *Disquisitiones Arithmeticae* regarded this as a problem that needs to be explored for "the dignity" of our subject. However it was not until the modern era, when questions about primality testing and factoring became a central part of applied mathematics,[1] that there was a large group of researchers endeavoring to solve these questions. As we shall see, most of the key ideas in recent work can be traced back to Gauss, Fermat and other mathematicians from times long gone by, and yet there is also a modern spin: With the growth of computer science and a need to understand the true difficulty of a computation, Gauss's vague assessment "intolerable labor" was only recently

Received by the editors January 27, 2004, and, in revised form, August 19, 2004.

2000 *Mathematics Subject Classification.* Primary 11A51, 11Y11; Secondary 11A07, 11A41, 11B50, 11N25, 11T06.

L'auteur est partiellement soutenu par une bourse du Conseil de recherches en sciences naturelles et en génie du Canada.

[1]Because of their use in the data encryption employed by public key cryptographic schemes; see section 3a.

clarified by "running time estimates", and of course our desktop computers are today's "indefatigable calculators".

Fast factoring remains a difficult problem. Although we can now factor an arbitrary large integer far faster than in Gauss's day, still a 400 digit integer which is the product of two 200 digit primes is typically beyond practical reach.[2] This is just as well since the safety of electronic business transactions, such as when you use your ATM card or purchase something with your credit card over the Web depends on the intractability of such problems!

On the other hand we have been able to rapidly determine whether quite large numbers are prime for some time now. For instance recent algorithms can test an arbitrary integer with several thousand digits for primality within a month on a Pentium IV,[3] which allows us to easily create the cryptographic schemes referred to above. However the modern interpretation of Gauss's dream was not realized until August 2002, when three Indian computer scientists—Manindra Agrawal, Neeraj Kayal and Nitin Saxena—constructed a "polynomial time deterministic primality test", a much sought-after but elusive goal of researchers in the algorithmic number theory world. Most shocking was the simplicity and originality of their test ... whereas the "experts" had made complicated modifications on existing tests to gain improvements (often involving great ingenuity), these authors rethought the direction in which to push the usual ideas with stunning success.[4] Their algorithm is based on the following elegant characterization of prime numbers.

Agrawal, Kayal and Saxena. *For given integer $n \geq 2$, let r be a positive integer $< n$, for which n has order $> (\log n)^2$ modulo r. Then n is prime if and only if*

- *n is not a perfect power,*
- *n does not have any prime factor $\leq r$,*
- *$(x+a)^n \equiv x^n + a \mod (n, x^r - 1)$ for each integer $a, 1 \leq a \leq \sqrt{r} \log n$.*

We will discuss the meaning of technical notions like "order" and "$\equiv$" a little later. For the reader who has encountered the terminology before, one word of warning: The given congruences here are between polynomials in x, and *not* between the two sides evaluated at each integer x.

At first sight this might seem to be a rather complicated characterization of the prime numbers. However we shall see that this fits naturally into the historical progression of ideas in this subject, is not so complicated (compared to some other ideas in use), and has the great advantage that it is straightforward to develop into a fast algorithm for proving the primality of large primes.

Perhaps now is the moment to clarify our goals:

[2]My definition will stretch the usual use of the word "practical" by erring on the cautious side: I mean a calculation that can be done using all computers that have or will be built for the next century, assuming that improvements in technology do not happen much faster than we have seen in the last couple of decades (during which time computer technology has, by any standards, developed spectacularly rapidly).

[3]This is not to be confused with algorithms that test the primality of integers of a special shape. For example, the GIMPS project (the Great Internet Mersenne Prime Search) routinely tests Mersenne numbers, numbers of the form $2^p - 1$, for primality which have millions of digits. However these algorithms have very limited applicability.

[4]Though some experts had actually developed many of the same ideas; see section 6.5 for example.

1.1. Our objective is to find a "quick" foolproof algorithm to determine whether a given integer is prime. Everyone knows *trial division*, in which we try to divide n by every integer m in the range $2 \le m \le \sqrt{n}$. The number of steps in this algorithm will be at least the number of integers m we consider, which is something like $\sqrt{n}$ in the worst case (when n is prime). Note that $\sqrt{n}$ is roughly $2^{d/2}$ where d is the number of digits of n when written in binary (and d is roughly $\log n$ where, here and throughout, we will take logarithms in base 2).

The *objective* in this area has been to come up with an algorithm which works in no more than cd^A steps in the worst case, where c and A are some fixed positive constants, that is, an algorithm which works in *Polynomial Time* (which is often abbreviated as P). With such an algorithm one expects that one can rapidly determine whether any "reasonably sized" integer is prime. This is the modern interpretation of Gauss's dream.

Before the work of Agrawal, Kayal and Saxena the fastest that we could prove any primality testing algorithm worked was something like $d^{c \log\log d}$ steps,[5] for some constant $c > 0$. At first sight this might seem far away from a polynomial time algorithm, since $\log\log d$ goes to infinity as $d \to \infty$. However $\log\log d$ goes to infinity "with great dignity" (as Dan Shanks put it), and in fact never gets larger than 7 in practice![6] By January 2004 such algorithms were able[7] to prove the primality of 10,000 digit primes (in base 10), an extraordinary achievement.

The algorithm of Agrawal, Kayal and Saxena works in about $d^{7.5}$ steps;[8] and a modification by Lenstra and Pomerance in about d^6 steps. This realizes Gauss's

"SEVEN AND A HALF LOGS SHOULD DO IT!"

[5]Though it is believed, but unproven, that some of these tests always work in polynomial time.

[6]By this I mean that if all computers on earth were to focus simply on writing down as many digits as possible for the next century, they would write down far less than 2^{2^7} digits.

[7]Using a version of the *Elliptic Curve Primality Proving* routine of François Morain and his collaborators.

[8]And thus the legend "Seven and a half logs should do it!" on the Larry Gonick cartoon above. Printed with permission.

dream, or, in modern language, implies that "Primes are in P". No one has yet written a computer program implementing these algorithms which comes close to proving primality of primes as large as those 10,000 digit primes discussed above.

Building on a clever idea of Berrizbeitia, Bernstein (and, independently, Mihailescu and Avanzi) gave a modification that will "almost certainly" run in around d^4 bit operations. There is hope that this can be used to write a computer program which will prove the primality of 10,000 digit primes rapidly. This is a challenge for the near future.

1.2. This article is an elaboration of a lecture given at the "Current Events" special session during the 2004 annual meeting of the American Mathematical Society. The purpose is to explain the AKS[9] primality test, with complete proofs, and to put the result and ideas in appropriate historical context.

To start with I would like to discuss simpler ideas from the subject of primality testing, focusing on some that are closely related to the AKS algorithm. In the process we will discuss the notions of complexity classes from theoretical computer science[10] and in particular introduce the "P≠NP" problem, one of the great challenges of mathematics for the new millennium.

We will see that most of the key ideas used to prove the theorem above were already in broad circulation, and so it is surprising that such an approach was not successfully completed earlier. I believe that there were two reasons for this — first, the way in which these classical ideas were combined was clever and original in several aspects. Second the authors are not number theorists and came at it from a little bit of a different angle; in particular not being so aware of what was supposedly too difficult, they trod where number theorists fear to tread.

In the third section we will discuss "running time" of algorithms in some detail and how they are determined, and so analyze the AKS algorithm.

In the fourth, and perhaps most interesting, section, we give the proof of the main results. In fact Agrawal et al. have produced two manuscripts, the second giving an even easier proof than the first, and we shall discuss both these proofs and relevant background information.

To prove the best running times for the algorithm it is necessary to employ tools of analytic number theory. In section 5 we introduce the reader to some beautiful theorems about the distribution of primes that should be better known and use them to prove the claimed running times.

In section 6, we discuss the modified AKS algorithm of Berrizbeitia and Bernstein, as well as Lenstra's finite field primality test,[11] and then, in section 7, the AKS-inspired algorithm of Lenstra and Pomerance.

Since the first announcement of this result in August 2002 there have been more than a dozen preprints circulating containing interesting ideas concerning the AKS algorithm, though none have yet appeared in print. I have thus succumbed to the temptation to include several of these ideas in the final section, in part because they are quite accessible, and in part because they are too elegant to leave out.

[9]An abbreviation for Agrawal, Kayal and Saxena.

[10]The cost of conveying the essence, rather than the details, of these notions is that our definitions will be a little awry, but not in a way that effects the key considerations in our context.

[11]Since this twenty-year-old test has much in common with the AKS test and has a running time that is not far from polynomial.

Bernstein reckons that these and other ideas for improving the AKS algorithm result in a speed up by a factor of about two million, although, he cautions, "two million times faster is not necessarily fast."

1.3. Undergraduate research experiences. Manindra Agrawal is a faculty member in the Computer Science Department at the Indian Institute of Technology in Kanpur, India. The fundamental approach taken here to primality testing was developed by Agrawal in conjunction with two bachelor's theses which we will discuss in section 8.5, the first completed by Pashant Pandey and Rajat Bhattacharjee in 2001, the second completed by Neeraj Kayal and Nitin Saxena in 2002. Later that summer they developed what they had done into a first version of the characterization of primes given above. There can have been few undergraduate research experiences with such a successful outcome!

2. Primality testing and the Child's Binomial Theorem

2.1. Recognizing primes. To find a primality test that works faster than trial division we look for other simple characterizations of prime numbers which might be used in a more efficient algorithm. If you have studied a little number theory, then a simple characterization of the primes that comes to mind is:

Wilson's Theorem (1770)**.** *Integer $n \geq 2$ is prime if and only if n divides $(n-1)!+1$.*

In trying to convert this elegant characterization into a fast algorithm we run into the problem that there is no obvious way to compute $(n-1)!$ rapidly (or even $(n-1)! \pmod n$).

Another idea is to use Matijasevič's unbelievable polynomial (1970) which was essential in resolving Hilbert's Tenth Problem. He showed how to construct a polynomial with integer coefficients (in several variables) such that whenever one substitutes in integers for the variables and gets a positive value, then that value is a prime number; moreover, every prime number will be such a value of the polynomial. (In fact one can construct such a polynomial of degree 10 in 26 variables.) However it is far from evident how to quickly determine whether a given integer is a value taken by this polynomial (at least no one to date has found a nice way to do so), and so this seems to be a hopeless approach.

Primes come up in many different places in the mathematical literature, and some of these suggest ways to distinguish primes from composites. Those of us who are interested in primality testing always look at anything new with one eye open to this application, and yet finding a fast primality testing algorithm has remained remarkably elusive. The advent of the AKS algorithm makes me wonder whether we have missed some such algorithm, something that one could perform in a few minutes, by hand, on any enormous number.

Such speculation brings me to a passage from Oliver Sacks' *The man who mistook his wife for a hat*, in which he tells us of a pair of severely autistic twins with a phenomenal memory for numbers and a surprising aesthetic. Sacks discovered the twins holding a purely numerical conversation, in which one would mention a six-digit number, the other would listen, think for a moment and then beam a smile of contented pleasure before responding with another six-digit number for his brother. After listening for a while, Sacks wrote the numbers down and, following a hunch, determined that all of the numbers exchanged were primes. The next day, armed

with a table of primes, Sacks butted into their conversation, venturing an eight-digit prime and eliciting, after a short pause, enthusiastic smiles from the twins. Now the twins kept on going, increasing the number of digits at each turn, until they were trading (as far as Sacks could tell) twenty-digit prime numbers. So how did the twins do it? Perhaps we will never know, since the twins were eventually separated, became "socialized" and forgot their amazing algorithm!

Since we do not know of any such shortcuts, I wish to move on to a property of prime numbers that has fascinated mathematicians since antiquity and has been developed into several key approaches to primality testing.

2.2. The Child's Binomial Theorem. The binomial theorem gives a formula for expanding $(x+y)^n$ as a sum of multiples of terms $x^i y^{n-i}$, namely

$$(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^i y^{n-i} \tag{2.1}$$

where $\binom{n}{i} = \frac{n!}{i!(n-i)!}$ and $m! = m(m-1)(m-2)\dots 3.2.1$. This is one of the first significant formulas that students learn in high school, though, perhaps due to its complicated structure, all too often we find university undergraduates who are unable to recall this formula and indeed who write down the (generally) incorrect

$$(x+y)^n = x^n + y^n. \tag{2.2}$$

This is called, by some, the *Child's Binomial Theorem*. Despite (2.2) being wrong in general, we shall be interested in those (unlikely) circumstances in which this formula is actually correct!

One of the most amazing properties of prime numbers, discovered[12] by Fermat around 1637, is that if n is prime, then n divides $a^n - a$ for all integers a. This may be rewritten[13] as

$$a^n \equiv a \mod n \tag{2.3}$$

for all integers a and primes n; and thus

$$(x+y)^n \equiv x+y \equiv x^n + y^n \pmod n \tag{2.4}$$

for all integers x, y and primes n.

The integers form equivalence classes mod n, and this set of equivalence classes forms the ring denoted $\mathbb{Z}/n$ (or $\mathbb{Z}/n\mathbb{Z}$). If n is prime, then $\mathbb{Z}/n$ is a field, and so the last equation may be rewritten as (2.2) in $\mathbb{Z}/n$.

Actually (2.4) holds also for *variables* x and y when n is prime, as we shall see later in section 2.8, a true Child's Binomial Theorem. It is easy to deduce Fermat's Little Theorem from the Child's Binomial Theorem by induction: Evidently (2.3) holds for $a = 1$, and if (2.3) holds for all $a < A$, then taking $x = A-1$, $y = 1$ in (2.4) gives $A^n \equiv (A-1)^n + 1^n \equiv (A-1) + 1 = A \pmod n$ by the induction hypothesis.

2.3. Composite numbers may sometimes be recognized when they do not have a particular property that primes have. For example, as we noted above

[12]There is evidence that this was known for $a = 2$ far earlier.

[13]For the uninitiated, we say that $a \equiv b \pmod m$ if and only if m divides $b - a$; the main advantage of this notation is that we can do most regular arithmetic operations $\pmod m$.

Fermat's Little Theorem (1637)**.** *If n is prime, then n divides $a^n - a$ for all integers a.*

Therefore conversely, if integer n *does not divide* $a^n - a$ for some integer a, then n is composite. For example, taking $a = 2$ we calculate that

$$2^{1001} \equiv 123 \pmod{1001},$$

so we know that 1001 is composite.

We might ask whether this always works. In other words,

Is it true that *if n is composite, then n does not divide* $2^n - 2$?

For, if so, we have a very nice way to distinguish primes from composites. Unfortunately the answer is "no" since, for example,

$$2^{341} \equiv 2 \pmod{341},$$

but $341 = 11 \times 31$. Note though that by taking $a = 3$ above we get

$$3^{341} \equiv 168 \pmod{341},$$

so we can still use these ideas to prove that 341 is composite.

But then we might ask whether this always works, whether there is always *some* value of a that helps us prove a composite n is indeed composite.

In other words,

Is it true that *if n is composite, then there is some integer a for which n does not divide $a^n - a$?*

Again the answer is "no" since 561 divides $a^{561} - a$ for all integers a, yet $561 = 3 \times 11 \times 17$. Composite integers n which divide $a^n - a$ for all integers a are called *Carmichael numbers*, 561, 1105 and 1729 being the smallest three examples. Carmichael numbers are a nuisance, masquerading as primes like this, though computationally they only appear rarely. Unfortunately it was recently proved that there are infinitely many of them and that when we go out far enough they are not so rare as it first appears.

2.4. Square roots of 1. In a field, a non-zero polynomial of degree d has at most d roots. For the particular example $x^2 - 1$ this implies that 1 has just two square roots mod p, a prime > 2, namely 1 and -1.

What about mod composite n? For the smallest odd composite n with more than one prime factor, that is $n = 15$, we find $1^2 \equiv 4^2 \equiv 11^2 \equiv 14^2 \pmod{15}$; that is, there are four square roots of 1 (mod 15). And this is true in general: There are *at least* four distinct square roots of 1 (mod n) for any odd n which is divisible by two distinct primes. Thus we might try to prove n is composite by finding a square root of 1 (mod n) which is neither 1 nor -1, though the question becomes, how do we efficiently search for a square root of 1?

Our trick is to again use Fermat's Little Theorem, since if p is prime > 2, then $p - 1$ is even, and so a^{p-1} is a square. Hence $(a^{\frac{p-1}{2}})^2 = a^{p-1} \equiv 1 \pmod{p}$, so $a^{\frac{p-1}{2}} \pmod{p}$ is a square root of 1 mod p and must be 1 or -1. Therefore if $a^{\frac{n-1}{2}} \pmod{n}$ is neither 1 nor -1, then n is composite. Let's try an example: We have $64^{948} \equiv 1 \pmod{949}$, and the square root $64^{474} \equiv 1 \pmod{949}$. Hmmmm, we failed to prove 949 is composite like this, but, wait a moment, since 474 is even, we can take the square root again, and a calculation reveals that $64^{237} \equiv 220 \pmod{949}$, so 949 is composite. More generally, using this trick of repeatedly taking

square roots (as often as 2 divides $n-1$), we call integer a a *witness* to n being composite if the finite sequence

$$a^{n-1} \pmod n,\ a^{(n-1)/2} \pmod n, \dots,\ a^{(n-1)/2^u} \pmod n$$

(where $n-1=2^u v$ with v odd) is not equal to either $1,1,\dots,1$ or $1,1,\dots,1,-1$, $*,\dots,*$ (which are the only two possibilities were n a prime). One can compute high powers mod n very rapidly using "fast exponentiation", a technique we will discuss in section 3b.

It is easy to show that at least one-half of the integers a, $1 \le a \le n$, are witnesses for n, for each odd composite n. So can we find a witness "quickly" if n is composite?

• The most obvious idea is to try $a = 2, 3, 4, \dots$ consecutively until we find a witness. It is believed that there is a witness $\le 2(\log n)^2$, but we cannot prove this (though we can deduce this from a famous conjecture, the Generalized Riemann Hypothesis).[14]

• Pick integers $a_1, a_2, \dots, a_k, \dots$ from $\{1, 2, 3, \dots, n-1\}$ at random until we find a witness. By what we wrote above, if n is composite, then the probability that none of $a_1, a_2, \dots, a_k$ are witnesses for n is $\le 1/2^k$. Thus with a hundred or so such tests we get a probability that is so small that it is inconceivable that it could occur in practice, so we believe that any integer n for which none of a hundred randomly chosen a's is a witness is prime. We call such n *industrial strength primes*.

In practice the witness test allows us to accomplish Gauss's dream of quickly distinguishing between primes and composites, for either we will quickly get a witness to n being composite or, if not, we can be almost certain that our industrial strength prime is indeed prime. Although this solves the problem in practice, we cannot be absolutely certain that we have distinguished correctly when we claim that n is prime since we have no proof, and mathematicians like proof. Indeed if you claim that industrial strength primes are prime without proof, then a cynic might not believe that your randomly chosen a are so random or that you are unlucky or ... No, what we need is a proof that a number is prime when we think that it is.

2.5. Proofs and the complexity class NP. At the 1903 meeting of the American Mathematical Society, F.N. Cole came to the blackboard and, without saying a word, wrote down

$$2^{67}-1 = 147573952589676412927 = 193707721 \times 761838257287,$$

long-multiplying the numbers out on the right side of the equation to prove that he was indeed correct. Afterwards he said that figuring this out had taken him "three years of Sundays." The moral of this tale is that although it took Cole a great deal of work and perseverance to find these factors, it did not take him long to justify his result to a room full of mathematicians (and, indeed, to give a proof that he was correct). Thus we see that one can provide a short proof, even if finding that proof takes a long time.

In general one can exhibit factors of a given integer n to give a short proof that n is composite (such proofs are called *certificates*). By "short" we mean that the proof

[14]We will not discuss the Riemann Hypothesis, or its generalizations, here. Suffice to say that this is one of the most famous and difficult open problems of mathematics, so much so that the Clay Mathematics Insitute has now offered one million dollars for its resolution (see `http://www.claymath.org/millennium/`).

can be verified in polynomial time, and we say that such problems are in class NP (*non-deterministic polynomial time*).[15] We are not suggesting that the proof can be found in polynomial time, only that the proof can be checked in polynomial time; indeed we have no idea whether it is possible to factor numbers in polynomial time, and this is now the outstanding problem of this area.

What about primality testing? If someone gives you an integer and asserts that it is prime, can you check that this is so in polynomial time? Can they give you better evidence than their say-so that it is a prime number? Can they provide some sort of "certificate" that gives you all the information you need to verify that the number is indeed a prime? It is not, as far as I can see, obvious how to do so, certainly not so obvious as with the factoring problem. It turns out that some old remarks of Lucas from the 1870's can be modified for this purpose:

First note that n is prime if there are precisely $n-1$ integers a in the range $1 \le a \le n-1$ which are coprime to n. Therefore if we can show the existence of $n-1$ distinct values mod n which are coprime to n, then we have a proof that n is prime. In fact if n is prime, then these values form a cyclic group under multiplication and so have a generator g; that is, there exists an integer g for which $1, g, g^2, \dots, g^{n-2}$ are all coprime to n and distinct mod n, so these are the $n-1$ distinct values mod n that we are looking for (note that $g^{n-1} \equiv 1 \pmod n$ by Fermat's little theorem). Thus to show that n is prime we need simply exhibit g and prove that g has order[16] $n-1 \pmod n$. It can be shown that any such order must divide $n-1$, and so one can show that if g is not a generator, then $g^{(n-1)/q} \equiv 1 \pmod n$ for some prime q dividing $n-1$. Thus a "certificate" to show that n is prime would consist of g and $\{q \text{ prime} : q \text{ divides } n-1\}$, and the checker would need to verify that $g^{n-1} \equiv 1 \pmod n$ whereas $g^{(n-1)/q} \not\equiv 1 \pmod n$ for all primes q dividing $n-1$, something that can be accomplished in polynomial time using fast exponentiation.

There is a problem though: One needs certification that each such q is prime. The solution is to iterate the above algorithm, and one can show that no more than $\log n$ odd primes need to be certified prime in the process of proving that n is prime. Thus we have a polynomial time certificate (short proof) that n is prime, and so primality testing is in the class NP.

But isn't this the algorithm we seek? Doesn't this give a polynomial time algorithm for determining whether a given integer n is prime? The answer is "no", because along the way we would have had to factor $n-1$ quickly, something no one knows how to do in general.

2.6. Is P≠NP? The set of problems that are in the complexity class P are those for which one can find the solution, with proof, in polynomial time, while the set of problems that are in the complexity class NP are those for which one can check the proof of the solution in polynomial time. By definition P⊆NP, and of course we believe that there are problems, for example the factoring problem, which are in NP but not in P; however *this has not been proved*, and it is now perhaps the outstanding unresolved question of theoretical computer science. This is another of the Clay Mathematics Institute's million dollar problems and perhaps the most

[15]Note that NP is **not** "non-polynomial time", a common source of confusion. In fact it is "non-deterministic", because the method for discovering the proof is not necessarily determined.

[16] The *order* of $h \pmod n$ is the least positive integer k for which $h^k \equiv 1 \pmod n$.

likely to be resolved by someone with less formal training, since the experts seem to have few plausible ideas for attacking this question.

It had better be the case that P≠NP, else there is little chance that one can have safe public key cryptography (see section 3a) or that one could build a highly unpredictable (pseudo-)random number generator[17] or that we could have any one of several other necessary software tools for computers. Notice that one implication of the "P≠NP" question remaining unresolved is that no fast public key cryptographic protocol is, as yet, provably safe!

2.7. Random polynomial time algorithms. In section 2.4 we saw that if n is composite, then there is a probability of at least $1/2$ that a random integer a is a witness for the compositeness of n, and if so, then it provides a short certificate verifying that n is composite. Such a test is called a *random polynomial time* test for compositeness (denoted RP). As noted, if n is composite, then the randomized witness test is almost certain to provide a short proof of that fact in 100 runs of the test. If 100 runs of the test do not produce a witness, then we can be almost certain that n is prime, but we cannot be *absolutely* certain since no proof is provided.

Short of finding a polynomial time test for primality, we might try to find a *random* polynomial time test for primality (in addition to the one we already have for compositeness). This was achieved by Adleman and Huang in 1992 using a method of counting points on elliptic and hyperelliptic curves over finite fields (based on ideas of Goldwasser and Kilian). Although beautiful in structure, their test is very complicated and almost certainly impractical, as well as being rather difficult to justify theoretically in all its details. It does however provide a short certificate verifying that a given prime is prime and proves that primality testing is also in complexity class RP.

If this last test were practical, then you could program your computer to run the witness test by day and the Adleman-Huang test by night and expect that you would not only quickly distinguish whether given integer n is prime or composite, but also rapidly obtain a proof of that fact. However you could not be certain that this would work—you might after all be very unlucky—so mathematicians would still wish to find a polynomial time test that would always work no matter how unlucky you are!

2.8. The new work of Agrawal, Kayal and Saxena starts from an old beginning, the Child's Binomial Theorem, in fact from the following result, which is a good exercise for an elementary number theory course.

Theorem 1. *Integer n is prime if and only if $(x+1)^n \equiv x^n + 1 \pmod{n}$ in $\mathbb{Z}[x]$.*

Proof. Since $(x+1)^n - (x^n+1) = \sum_{1\le j\le n-1} \binom{n}{j} x^j$, we have that $x^n+1 \equiv (x+1)^n \pmod{n}$ if and only if n divides $\binom{n}{j}$ for all j in the range $1 \le j \le n-1$.

If $n = p$ is prime, then p appears in the numerator of $\binom{p}{j}$ but is larger than, and so does not divide, any term in the denominator, and hence p divides $\binom{p}{j}$ for $1 \le j \le p-1$.

[17]So-called "random number generators" written in computer software are not random since they need to work on a computer where everything is designed to be determined! Thus what are called "random numbers" are typically a sequence of numbers, determined in a totally predictable manner but which appear to be random when subjected to "randomness tests" in which the tester does not know how the sequence was generated.

If n is composite let p be a prime dividing n. In the expansion

$$\binom{n}{p} = \frac{n(n-1)(n-2)\dots(n-(p-1))}{p(p-1)\dots 1}$$

we see that the only terms p divides are the n in the numerator and the p in the denominator; and so if p^k is the largest power of p dividing n, then p^{k-1} is the largest power of p dividing $\binom{n}{p}$, and therefore n does not divide $\binom{n}{p}$. □

This simple theorem is the basis of the new primality test. In fact, why don't we simply compute $(x+1)^n - (x^n+1) \pmod n$ and determine whether or not n divides each coefficient? This is a valid primality test, but computing $(x+1)^n \pmod n$ is obviously slow since it will involve storing n coefficients!

Since the difficulty here is that the answer involves so many coefficients (as the degree is so high), one idea is to compute mod some small degree polynomial as well as mod n so that neither the coefficients nor the degree get large. The simplest polynomial of degree r is perhaps $x^r - 1$. So why not verify whether[18]

$$(x+1)^n \equiv x^n + 1 \mod (n, x^r - 1)? \tag{2.5}$$

This can be computed rapidly (as we will discuss in section 3b.2), and it is true for any prime n (as a consequence of the theorem above), but it is unclear whether this fails to hold for all composite n and thus provides a true primality test. However, the main theorem of Agrawal, Kayal and Saxena provides a modification of this congruence, which can be shown to succeed for primes and fail for composites, thus providing a polynomial time primality test. In section 4 we shall show that this is so, but first we discuss various computational issues.

3a. Computational issues: Factoring and primality testing as applied to cryptography

In cryptography we seek to transmit a secret message m from Alice to Bob in such a way that Oscar, who intercepts the transmission, cannot read the message. The idea is to come up with an encryption key Φ, an easily described mathematical function, which transforms m into $r := \Phi(m)$ for transmission. The number r should be a seemingly meaningless jumble of symbols that Oscar cannot interpret and yet Bob can decipher by computing $\Psi(r)$, where $\Psi = \Phi^{-1}$. Up until recently, knowledge of the encryption key Φ would allow the astute Oscar to determine the decryption key Ψ, and thus it was extremely important to keep the encryption key Φ secret, often a difficult task.

It seems obvious that if Oscar is given an encryption key, then it should be easy for him to determine the decryption key by simply reversing what was done to encrypt. However, in 1976 Diffie and Hellman postulated the seemingly impossible idea of creating a *public key* Φ, which Oscar can see yet which gives no hint in and of itself as to how to determine $\Psi = \Phi^{-1}$. If feasible this would rid Alice of the difficulty of keeping her key secret.

In modern public key cryptosystems the difficulty of determining Ψ from Φ tends to be based on an unsolved deep mathematical problem, preference being given to problems that have withstood the onslaught of the finest minds from Gauss

[18]The ring of polynomials with integer coefficients is denoted $\mathbb{Z}[x]$. Then $f(x) \equiv g(x) \mod (n, h(x))$ for $f(x), g(x), h(x) \in \mathbb{Z}[x]$ if and only if there exist polynomials $u(x), v(x) \in \mathbb{Z}[x]$ for which $f(x) - g(x) = nu(x) + h(x)v(x)$.

onwards, like the factoring problem. We now discuss the most famous of these public key cryptosystems.

3a.1. The RSA cryptosystem. In 1982, Ron Rivest, Adi Shamir and Len Adleman proposed a public key cryptosystem which is at the heart of much computer security today and yet is simple enough to teach to high school students. The idea is that Bob takes two large primes $p < q$, and their product $n = pq$, and determines two integers d and e for which $de \equiv 1 \pmod{(p-1)(q-1)}$ (which is easy). The public key, which Alice uses, will consist of the numbers n and the encryption key e, whereas Bob keeps the decryption key d secret (as well as p and q). We will suppose that the message m is an integer[19] in $[1, n-1]$. To encrypt m, Alice computes $r := \Phi(m) \equiv m^e \pmod{n}$, with $\Phi(m) \in [1, n-1]$, which can be done rapidly (see section 3b.2). To decrypt Bob computes $\Psi(r) := r^d \pmod{n}$, with $\Psi(r) \in [1, n-1]$. Using Fermat's Little Theorem (for p and q) the reader can easily verify that $m^{de} \equiv m \pmod{n}$, and thus $\Psi = \Phi^{-1}$.

Oscar knows n, and if he could factor n, then he could easily determine Ψ; thus the RSA cryptosystem's security depends on the difficulty of factoring n. As noted above, this is far beyond what is feasible today if we take p and q to be primes that contain more than 200 digits. Finding such large primes, however, is easy using the methods discussed in this article!

Thus the ability to factor n gives Oscar the ability to break the RSA cryptosystem, though it is unclear whether the RSA cryptosystem might be broken much more easily. It makes sense then to try to come up with a public key cryptosystem whose security is essentially as strong as the difficulty of the factoring problem.

3a.2. The ability to take square roots $\pmod{n}$ is not as benign as it sounds. In section 2.4 we saw that if we could find a square root b of 1 mod n which is neither 1 nor -1, then this proves that n is composite. In fact b yields a partial factorization of odd n, for

$$\gcd(b-1, n)\ \gcd(b+1, n) = \gcd(b^2 - 1, n) = n, \tag{3.1}$$

(as $b^2 \equiv 1 \pmod{n}$) whereas $\gcd(b-1, n)$, $\gcd(b+1, n) < n$ (since $b \not\equiv 1$ or $-1 \pmod{n}$), which together imply that $1 < \gcd(b-1, n)$, $\gcd(b+1, n) < n$, and hence (3.1) provides a non-trivial factorization of n.

More generally let us suppose that for a given odd, composite integer n with at least two distinct prime factors, Oscar has a function f_n such that if a is a square mod n, then $f_n(a)^2 \equiv a \pmod{n}$. Using f_n, Oscar can easily factor n (in random polynomial time), for if he picks integers b in $[1, n-1]$ at random, then $\gcd(f_n(b^2) - b, n)$ is a non-trivial factor of n provided that $b \not\equiv f_n(b^2)$ or $-f_n(b^2) \pmod{n}$; since there are at least four square roots of $b^2 \pmod{n}$, the probability that this provides a partial factorization of n is $\geq 1/2$.

Using this idea, Rabin constructed a public key cryptosystem which is essentially as hard to break as it is difficult to factor n.

[19]Any alphabetic message m can be transformed into numbers by replacing "A" by "01", "B" by "02", etc., and then into several such integers by cutting the digits (in binary representation) into blocks of length $< \log n$.

3a.3. On the difficulty of finding non-squares $\pmod p$**.** For a given odd prime p it is easy to find a square mod p: take 1 or 4 or 9, or indeed any $a^2 \pmod p$. Exactly $(p-1)/2$ of the non-zero values mod p are squares mod p, and so exactly $(p-1)/2$ are not squares mod p. One might guess that they would also be easy to find, but we do not know a surefire way to quickly find such a value for each prime p (though we do know a quick way to identify a non-square once we have one).

Much as in the search for witnesses discussed in section 2.4, the most obvious idea is to try $a = 2, 3, 4, \dots$ consecutively until we find a non-square. It is believed that there is a non-square $\le 2(\log p)^2$, but we cannot prove this (though we can also deduce this from the Generalized Riemann Hypothesis).

Another way to proceed is to pick integers $a_1, a_2, \dots, a_k, \dots$ from $\{1, 2, 3, \dots, n-1\}$ at random until we find a non-square. The probability that none of $a_1, a_2, \dots, a_k$ are non-squares mod p is $\le 1/2^k$, so with a hundred or so such choices it is inconceivable that we could fail!

3B. Computational issues: Running times of calculations

3b.1. Arithmetic on a computer. Suppose a and b are two positive integers, each with no more than ℓ digits when written in binary. We are interested in the number of bit operations a computer takes to perform various calculations. Both addition and subtraction can obviously be performed in $O(\ell)$ bit operations.[20] The most efficient method for multiplication (using Fast Fourier Transforms) takes time[21] $O(\ell \log \ell \log\log \ell)$. The precise "log" and "log log" powers in these estimates are more or less irrelevant to our analysis, so to simplify the writing we define $\tilde{O}(y)$ to be $O(y(\log y)^{O(1)})$. Then division of a by b and reducing $a \pmod b$ also take time $\tilde{O}(\ell)$.

Now suppose a and b are two polynomials, with integer coefficients, of degree less than r whose coefficients have no more than ℓ binary digits. Adding or subtracting will take $O(\ell r)$ operations. To multiply $a(x)$ and $b(x)$ we use the method of "single point evaluation" which is so well exploited in MAPLE. The idea comes from the observation that there is a natural bijection

$$\left\{ a(x) = \sum_{i=0}^{r-1} a_i x^i \in \mathbb{Z}[x] : -A < a_i \le A \text{ for all } i \right\} \xrightarrow{\Psi} \mathbb{Z}/(2A)^r,$$

where $\Psi(a) = a(2A)$. To recover $a_0, a_1, \dots a_{r-1}$ successively from this value, note that $a_0 \equiv a(2A) \pmod{2A}$ and $-A < a_0 \le A$ so a_0 is uniquely determined. Then $a_1 \equiv (a(2A) - a_0)/(2A) \pmod{2A}$ and $-A < a_1 \le A$, so a_1 is uniquely determined, and we continue like this. In an algorithm to determine $c(x) := a(x)b(x)$, we first note that the absolute values of the coefficients of $a(x)b(x)$ are all $< A := r2^{2\ell}$. Then we evaluate $a(2A)$ and $b(2A)$ and multiply these integers together to get $a(2A)b(2A)$, and then recover $c(2A)$ (by a single point evaluation). Unsurprisingly the most expensive task is the multiplication (of two integers which are each $< (2A)^r$) and so this algorithm takes time $\tilde{O}(r(\ell + \log r))$.

In our application we also need to reduce polynomials $a(x) \bmod (n, x^r - 1)$, where the coefficients of a have $O(\ell + \log r)$ digits and a has degree $< 2r$. Replacing each

[20] The notation "$O(*)$" can be read as "bounded by a fixed multiple of $*$".

[21] In this context, read "time" as a synonym for "bit operations".

x^{r+j} by x^j and then reducing the coefficients of the resulting polynomial mod n will take time $\tilde{O}(r(\ell + \log r))$ by the running times for integer arithmetic given above.

3b.2. Fast exponentiation. An astute reader might ask how we can raise something to the nth power "quickly", a problem which was beautifully solved by computer scientists long ago:[22]

We wish to compute $(x+a)^n \mod (n, x^r-1)$ quickly. Define $f_0(x) = (x+a)$ and then $f_{j+1}(x) \equiv f_j(x)^2 \mod (n, x^r-1)$ for $j \geq 0$ (at each step we determine $f_j(x)^2$ and then reduce mod $x^r - 1$ so the degree of the resulting polynomial is $< r$, and then reduce mod n to obtain f_{j+1}). Note that $f_j(x) \equiv (x+a)^{2^j} \mod (n, x^r-1)$.

Writing n in binary, say as $n = 2^{a_1} + 2^{a_2} + \cdots + 2^{a_\ell}$ with $a_1 > a_2 > \cdots > a_\ell \geq 0$, let $g_1(x) = f_{a_1}(x)$ and then $g_j(x) \equiv g_{j-1}(x) f_{a_j}(x) \mod (n, x^r-1)$ for $j = 1, 2, \ldots, \ell$. Therefore

$$g_\ell(x) \equiv (x+a)^{2^{a_1}+2^{a_2}+\cdots+2^{a_\ell}} = (x+a)^n \mod (n, x^r-1).$$

Thus we have computed $(x+a)^n \mod (n, x^r-1)$ in $a_1 + \ell \leq 2\log n$ such steps, where a step involves multiplying two polynomials of degree $< r$ with coefficients in $\{0, 1, \ldots, n-1\}$ and reducing $\mod (n, x^r-1)$, so has running time $\tilde{O}(r\ell(\ell + \log r))$.

3b.3. The AKS algorithm runs in $\tilde{O}(r^{3/2}(\log n)^3)$ bit operations, as we will now show. To transform the theorem of Agrawal, Kayal and Saxena into an algorithm we proceed as follows:

• Determine whether n is a perfect power.

We leave this challenge to our inventive reader, while noting that this can be done in no more than $\tilde{O}((\log n)^3)$ bit operations.

If n is not a perfect power, then we

• Find an integer r for which the order of $n \pmod r$ is $> (\log n)^2$.

The obvious way to do this is to compute $n^j \pmod q$ for $j = 1, \ldots, [(\log n)^2]$ and each integer $q > [(\log n)^2]$ until we find the first value of q for which none of these residues is equal to $1 \pmod q$. Then we can take $r = q$. This stage of the algorithm takes $\tilde{O}(r(\log n)^2)$ bit operations.[23]

• Determine whether $\gcd(a, n) > 1$ for some $a \leq r$, which will take $\tilde{O}(r(\log n)^2)$ bit operations using the Euclidean algorithm, provided $r < n$.

Finally we verify whether the Child's Binomial Theorem holds:

• Determine whether $(x+a)^n \equiv x^n + a \mod (n, x^r-1)$ for $a = 1, 2, \ldots, [\sqrt{r}\log n]$; each such congruence takes $\tilde{O}(r(\log n)^2)$ bit operations to verify by the previous section, and so a total running time of $\tilde{O}(r^{3/2}(\log n)^3)$.

Adding these times up shows that the running time of the whole algorithm is as claimed.

3b.4. The size of r. Since r is greater than the order of $n \pmod r$ which is $> (\log n)^2$, therefore $r > (\log n)^2$. This implies that the AKS algorithm cannot run in fewer than $\tilde{O}((\log n)^6)$ bit operations. On the other hand we expect to be able to find many such r which are $< 2(\log n)^2$ (in which case r will be prime and

[22] Legendre computed high powers mod p by similar methods in 1785!

[23] A subtlety that arises here and elsewhere is that $(\log n)^2$ could be so close to an integer J that it would take too many bit operations to determine whether $[(\log n)^2]$ equals $J-1$ or J. However, if we allow j to run up to J here, and in the AKS theorem if we allow a to run up to the smallest integer that is "obviously" $> \sqrt{r}\log n$, then we avoid this difficulty and do not significantly increase the running time.

the powers of n will generate all of the $r-1$ non-zero residues mod r), and this is borne out in computation though we cannot prove that this will always be true. Evidently, if there are such r then the AKS algorithm will run in $\tilde{O}((\log n)^6)$ bit operations which, as we have explained, is as good as we can hope for.

We can show unconditionally that there are integers r for which the order of n (mod r) is $> (\log n)^2$ and with r not too big. In section 4.3 we will use elementary estimates about prime numbers to show that such r exist around $(\log n)^5$, which leads to a running time of $\tilde{O}((\log n)^{10\frac{1}{2}})$ (since $10\frac{1}{2} = \frac{3}{2} \times 5 + 3$).

In section 6 we will use basic tools of analytic number theory to show that such r exist which are a little less than $(\log n)^{50/11}$ (using an old argument of Goldfeld), which leads to a running time of $O((\log n)^{9\frac{9}{11}})$.

It is important to note that the two upper bounds on the running time given above can both be made absolutely explicit from the proofs; in other words all the constants and functions implicit in the notation can be given precisely, and these bounds on the running time work for all n.

Using much deeper tools, a result of Fouvry[24] implies that such r exist around $(\log n)^3$, which leads to a running time of $O((\log n)^{7\frac{1}{2}})$. This can be improved using a recent result of Baker and Harman to $O((\log n)^{7.49})$. However the constants implicit in the "$O(.)$" notation here cannot be given explicitly by the very nature of the proof, as we will explain in section 5.4.

4. Proof of the theorem of Agrawal, Kayal and Saxena

We will assume that we are given an odd integer $n > 1$ which we know is not a perfect power, which has no prime factor $\leq r$, which has order $d > (\log n)^2$ modulo r, and such that

$$(x+a)^n \equiv x^n + a \mod (n, x^r - 1) \tag{4.1}$$

for each integer $a,\ 1 \leq a \leq A$ where we take $A = \sqrt{r}\log n$. By Theorem 1 we know that these hypotheses hold if n is prime, so we must show that they cannot hold if n is composite.

Let p be a prime dividing n so that

$$(x+a)^n \equiv x^n + a \mod (p, x^r - 1) \tag{4.2}$$

for each integer $a,\ 1 \leq a \leq A$. We can factor $x^r - 1$ into irreducibles in $\mathbb{Z}[x]$, as $\prod_{d|r} \Phi_d(x)$, where $\Phi_d(x)$ is the dth cyclotomic polynomial, whose roots are the primitive dth roots of unity. Each $\Phi_r(x)$ is irreducible in $\mathbb{Z}[x]$, but may not be irreducible in $(\mathbb{Z}/p\mathbb{Z})[x]$, so let $h(x)$ be an irreducible factor of $\Phi_r(x) \pmod p$. Then (4.2) implies that

$$(x+a)^n \equiv x^n + a \mod (p, h(x)) \tag{4.3}$$

for each integer $a, 1 \leq a \leq A$, since $(p, h(x))$ divides $(p, x^r - 1)$.

The congruence classes mod $(p, h(x))$ can be viewed as the elements of the ring $\mathbb{F} :\equiv \mathbb{Z}[x]/(p, h(x))$, which is isomorphic to the field of p^m elements (where m is the degree of h). In particular the non-zero elements of $\mathbb{F}$ form a cyclic group of order $p^m - 1$; moreover, $\mathbb{F}$ contains x, an element of order r, thus r divides $p^m - 1$. Since $\mathbb{F}$ is isomorphic to a field, the congruences (4.3) are much easier to work with than (4.1), where the congruences do not correspond to a field.

[24]Fouvry's 1984 result was at the time applied to prove a result about Fermat's Last Theorem.

Let H be the elements mod $(p, x^r - 1)$ generated multiplicatively by $x, x+1$, $x+2, \dots, x+[A]$. Let G be the (cyclic) subgroup of $\mathbb{F}$ generated multiplicatively by $x, x+1, x+2, \dots, x+[A]$; in other words G is the reduction of H mod $(p, h(x))$. All of the elements of G are non-zero, for if $x + a = 0$ in $\mathbb{F}$, then $x^n + a = (x + a)^n = 0$ in $\mathbb{F}$ by (4.3), so that $x^n = -a = x$ in $\mathbb{F}$, which would imply that $n \equiv 1 \pmod r$ and so $d = 1$, contradicting the hypothesis.

Note that if $g(x) = \prod_{0 \le a \le A}(x + a)^{e_a} \in H$, then

$$g(x)^n = \prod_a ((x + a)^n)^{e_a} \equiv \prod_a (x^n + a)^{e_a} = g(x^n) \mod (p, x^r - 1)$$

by (4.2). Define S to be the set of positive integers k for which $g(x^k) \equiv g(x)^k$ mod $(p, x^r - 1)$ for all $g \in H$. Then $g(x^k) \equiv g(x)^k$ in $\mathbb{F}$ for each $k \in S$, so that the Child's Binomial Theorem holds for elements of G in this field for the set of exponents S! Note that $p, n \in S$.

Our plan is to give upper and lower bounds on the size of G to establish a contradiction.

4.1. Upper bounds on $|G|$.

Lemma 4.1. *If $a, b \in S$, then $ab \in S$.*

Proof. If $g(x) \in H$, then $g(x^b) \equiv g(x)^b \mod (p, x^r - 1)$; and so, replacing x by x^a, we get $g((x^a)^b) \equiv g(x^a)^b \mod (p, (x^a)^r - 1)$, and therefore $\mod (p, x^r - 1)$ since $x^r - 1$ divides $x^{ar} - 1$. Therefore

$$g(x)^{ab} = (g(x)^a)^b \equiv g(x^a)^b \equiv g((x^a)^b) = g(x^{ab}) \mod (p, x^r - 1)$$

as desired. □

Lemma 4.2. *If $a, b \in S$ and $a \equiv b \mod r$, then $a \equiv b \mod |G|$.*

Proof. For any $g(x) \in \mathbb{Z}[x]$ we have that $u - v$ divides $g(u) - g(v)$. Therefore $x^r - 1$ divides $x^{a-b} - 1$, which divides $x^a - x^b$, which divides $g(x^a) - g(x^b)$; and so we deduce that if $g(x) \in H$, then $g(x)^a \equiv g(x^a) \equiv g(x^b) \equiv g(x)^b \mod (p, x^r - 1)$. Thus if $g(x) \in G$, then $g(x)^{a-b} \equiv 1$ in $\mathbb{F}$; but G is a cyclic group, so taking g to be a generator of G we deduce that $|G|$ divides $a - b$. □

Let R be the subgroup of $(\mathbb{Z}/r\mathbb{Z})^*$ generated by n and p. Since n is not a power of p, the integers $n^i p^j$ with $i, j \ge 0$ are distinct. There are $> |R|$ such integers with $0 \le i, j \le \sqrt{|R|}$ and so two must be congruent $\pmod r$, say

$$n^i p^j \equiv n^I p^J \pmod r.$$

By Lemma 4.1 these integers are both in S. By Lemma 4.2 their difference is divisible by $|G|$, and therefore

$$|G| \le |n^i p^j - n^I p^J| \le (np)^{\sqrt{|R|}} - 1 < n^{2\sqrt{|R|}} - 1.$$

(Note that $n^i p^j - n^I p^J$ is non-zero since n is neither a prime nor a perfect power.) We will improve this by showing that $n/p \in S$ and then replacing n by n/p in the argument above to get

$$|G| \le n^{\sqrt{|R|}} - 1. \tag{4.4}$$

Since n has order d mod r, $n^d \equiv 1 \pmod r$ and so $x^{n^d} \equiv x \pmod{x^r - 1}$. Suppose that $a \in S$ and $b \equiv a \pmod{n^d - 1}$. Then $x^r - 1$ divides $x^{n^d} - x$, which

divides $x^b - x^a$, which divides $g(x^b) - g(x^a)$ for any $g(x) \in \mathbb{Z}[x]$. If $g(x) \in H$, then $g(x)^{n^d} \equiv g(x^{n^d}) \mod (p, x^r - 1)$ by Lemma 4.1 since $n \in S$, and $g(x^{n^d}) \equiv g(x) \mod (p, x^r - 1)$ (as $x^r - 1$ divides $x^{n^d} - x$) so that $g(x)^{n^d} \equiv g(x) \mod (p, x^r - 1)$. But then $g(x)^b \equiv g(x)^a \mod (p, x^r - 1)$ since $n^d - 1$ divides $b - a$. Therefore

$$g(x^b) \equiv g(x^a) \equiv g(x)^a \equiv g(x)^b \mod (p, x^r - 1)$$

since $a \in S$, which implies that $b \in S$. Now let $b = n/p$ and $a = np^{\phi(n^d-1)-1}$, so that $a \in S$ by Lemma 4.1 since $p, n \in S$. Also $b \equiv a \pmod{n^d - 1}$ so $b = n/p \in S$ by the above.

4.2. Lower bounds on $|G|$. We wish to show that there are many distinct elements of G. If $f(x), g(x) \in \mathbb{Z}[x]$ with $f(x) \equiv g(x) \mod (p, h(x))$, then we can write $f(x) - g(x) \equiv h(x)k(x) \mod p$ for some polynomial $k(x) \in \mathbb{Z}[x]$. Thus if f and g both have smaller degree than h, then $k(x) \equiv 0 \pmod p$ and so $f(x) \equiv g(x) \pmod p$. Hence all polynomials of the form $\prod_{1 \le a \le A}(x + a)^{e_a}$ of degree $< m$ (the degree of $h(x)$) are distinct elements of G. Therefore if m, the order of $p \pmod r$, is large, then we can get good lower bounds on $|G|$.

This was what Agrawal, Kayal and Saxena did in their first preprint, and to prove such r exist they needed to use non-trivial tools of analytic number theory. In their second preprint, inspired by remarks of Hendrik Lenstra, they were able to replace m by $|R|$ in this result, which allows them to give an entirely elementary proof of their theorem and to get a stronger result when they do invoke the deeper estimates.

Lemma 4.3. *Suppose that $f(x), g(x) \in \mathbb{Z}[x]$ with $f(x) \equiv g(x) \mod (p, h(x))$ and that the reductions of f and g in $\mathbb{F}$ both belong to G. If f and g both have degree $< |R|$, then $f(x) \equiv g(x) \pmod p$.*

Proof. Consider $\Delta(y) := f(y) - g(y) \in \mathbb{Z}[y]$ as reduced in $\mathbb{F}$. If $k \in S$, then

$$\Delta(x^k) = f(x^k) - g(x^k) \equiv f(x)^k - g(x)^k \equiv 0 \mod (p, h(x)).$$

It can be shown that x has order r in $\mathbb{F}$ so that $\{x^k : k \in R\}$ are all distinct roots of $\Delta(y) \bmod (p, h(x))$. Now, $\Delta(y)$ has degree $< |R|$, but $\ge |R|$ distinct roots mod $(p, h(x))$, and so $\Delta(y) \equiv 0 \bmod (p, h(x))$, which implies that $\Delta(y) \equiv 0 \pmod p$ since its coefficients are independent of x. □

By definition R contains all the elements generated by $n \pmod r$, and so R is at least as large as d, the order of $n \pmod r$, which is $> (\log n)^2$ by assumption. Therefore A, $|R| > B$, where $B := [\sqrt{|R|} \log n]$. Lemma 4.3 implies that the products $\prod_{a \in T}(x + a)$ give distinct elements of G for every proper subset T of $\{0, 1, 2, \ldots, B\}$, and so

$$|G| \ge 2^{B+1} - 1 > n^{\sqrt{|R|}} - 1,$$

which contradicts (4.3). This completes the proof of the theorem of Agrawal, Kayal and Saxena.

4.3. Large orders mod r. The prime number theorem can be paraphrased[25] as: *The product of the primes up to x is roughly e^x.* A weak explicit version states that the product of the primes between N and $2N$ is $\ge 2^N$ for all $N \ge 1$.

[25]See section 5.1 for a more precise version, or the book [20].

Lemma 4.4. *If $n \ge 6$, then there is a prime $r \in [(\log n)^5, 2(\log n)^5]$ for which the order of n mod r is $> (\log n)^2$.*

Proof. If not, then the order of n mod r is $\le I := (\log n)^2$ for every prime $r \in [N, 2N]$ with $N := (\log n)^5$, so that their product divides $\prod_{i \le I}(n^i - 1)$. But then

$$2^N \le \prod_{\substack{N \le r \le 2N \\ r \text{ prime}}} r \le \prod_{i \le I}(n^i - 1) < n^{\sum_{i \le I} i} < 2^{(\log n)^5},$$

for $n \ge 6$, giving a contradiction. □

The bound on r here holds for all $n \ge 6$, and thus using this bound our running time analysis of AKS is *effective*; that is, one can explicitly bound the running time of the algorithm for all $n \ge 6$. In some of the better bounds on r discussed in section 3.4, the proofs are not effective in that they do not imply how large n must be for the given upper bound for r to hold.

4.4. Large prime factors of the order of $n \pmod r$. The other estimates mentioned in section 3b.4 all follow from using deeper results of analytic number theory which show that there are many primes r for which $r - 1$ has a large prime factor $q = q_r > (\log n)^2$. By showing that this large prime q divides the order of $n \pmod r$ for all but a small set of exceptional r, we deduce that the order of $n \pmod r$ is $\ge q > (\log n)^2$. (In the first version of the AKS paper they needed m, the order of $p \pmod r$, to be $> (\log n)^2$ and obtained this through the same argument, since the order of $n \pmod r$ divides the product of the orders of $p \pmod r$, where the product is taken over the prime divisors p of n, and thus q must divide the order of $p \pmod r$ for some prime p dividing n.) We now describe our argument a little more explicitly in terms of a well-believed

Conjecture. *For any given θ in the range $0 < \theta < 1/2$ there exists $c = c(\theta) > 0$ such that there are at least $2cR/\log R$ primes r in $[R, 2R]$ for which $r - 1$ has a prime factor $q > r^{1/2+\theta}$, provided R is sufficiently large.*

Lemma 4.5. *Assume the conjecture for some θ, $0 < \theta < 1/2$. Suppose n is a sufficiently large integer and that $c(\theta)R^{2\theta} \ge \log n$. There are at least $c(\theta)R/\log R$ primes r in $[R, 2R]$ for which the order of $n \pmod r$ is $> r^{1/2+\theta}$.*

Proof. We will show that the number N of primes r given in the conjecture for which q does not divide the order of $n \pmod r$ is $< c(\theta)R/\log R$. Now, if r is such a prime, then the order of $n \pmod r$ divides $(r-1)/q$, and so is $< r/q \le r^{1/2-\theta} \le (2R)^{1/2-\theta}$. Therefore the product of such r divides $\prod_{m \le (2R)^{1/2-\theta}} (n^m - 1)$, which implies that

$$R^N \leqq \prod_{m \le (2R)^{1/2-\theta}} n^m < n^{R^{1-2\theta}}$$

and thus $N < R^{1-2\theta}(\log n)/(\log R) \le c(\theta)R/\log R$ as desired. □

One easily deduces

Corollary 4.6. *Assume the conjecture for some θ, $0 < \theta < 1/2$, and define $\rho(\theta) := \max\{1/(2\theta), 4/(1+2\theta)\}$. There exists a constant $c'(\theta)$ such that if n is a sufficiently large integer, then there exists a prime $r < c'(\theta)(\log n)^{\rho(\theta)}$ for which the order of n mod r is $> (\log n)^2$.*

In section 5.3 we will prove the conjecture for some value of $\theta > .11$ using basic tools of analytic number theory, so that Corollary 4.6 shows us that we can take $r < (\log n)^{50/11}$. Fouvry proved a version of the conjecture for some $\theta > 1/6$, and so a modification of Corollary 4.6 shows us that we can take $r < (\log n)^3$. These estimates were used in section 3b.4.

5. Results about counting primes that everyone should know

5.1. Primes and sums over primes. The prime number theorem tells us that $\pi(x)$, the number of primes up to x, is[26] $\sim x/\ln x$ as $x \to \infty$. This can be shown to be equivalent to the statement

$$\sum_{\substack{p\le x\\ p \text{ prime}}} \ln p \ \sim \ x \tag{5.1}$$

mentioned in section 4.3. (Here and throughout section 5 we will use p to denote only prime numbers, and so if we rewrite the above sum as $\sum_{p\le x} \ln p$, then the reader should assume it is a sum over primes p in this range.) There is no easy proof of the prime number theorem, though there are several that avoid any deep machinery.

In the proofs below we will deduce several estimates from (5.1).

There will also be estimates of less depth. For example we encounter several sums, over primes of positive summands, which converge when extended to a sum over all integers ≥ 2, so that the original sum converges.

We now give a well-known elementary estimate. First observe that since the primes in $(m, 2m]$ always divide $\binom{2m}{m}$, which is $\le 2^{2m}$, thus $\sum_{m<p\le 2m} \log p \le 2m$; and summing over $m = [x/2^i] + 1$ for $i = 1, 2, \dots$ we deduce that $\sum_{p\le x} \log p \le 2(x + \log x)$, a weak but explicit version of (5.1). Now

$$\sum_{n\le x} \ln n = \sum_{n\le x} \left(\sum_{p^a | n} \ln p \right) = \sum_{p^a \le x} \ln p \sum_{\substack{n\le x\\ p^a | n}} 1 = \sum_{p^a \le x} \left[\frac{x}{p^a} \right] \ln p$$

$$= \sum_{p^a \le x} \left\{ \frac{x}{p^a} + O(1) \right\} \ln p = x \sum_{p\le x} \frac{\ln p}{p-1} + O(x),$$

since

$$\sum_{p\le x<p^a} (\ln p)/p^a \le (2/x) \sum_{p\le x} \ln p = O(1)$$

and

$$\sum_{p^a \le x} \ln p = \sum_{a\ge 1} \sum_{p\le x^{1/a}} \ln p = O(x),$$

using our weak form of (5.1). One can use elementary calculus to show that $\sum_{n\le x} \ln n = x \ln x + O(x)$ and thus deduce from the above that

$$\sum_{p\le x} \frac{\ln p}{p-1} = \ln x + O(1). \tag{5.2}$$

[26] We write $u(x) \sim v(x)$ if $u(x)/v(x) \to 1$ as $x \to \infty$.

5.2. Primes in arithmetic progressions. We are interested in estimating $\pi(x;q,a)$, the number of primes up to x that are $\equiv a \pmod q$. We know that

$$\pi(x;q,a) \sim \pi(x)/\phi(q) \text{ as } x \to \infty \text{ if } (a,q)=1,$$

where $\phi(q) = \#\{a \pmod q : (a,q)=1\}$. We wish to know whether something like this[27] holds for "small" x. Unconditionally we can only prove this when $x > \exp(q^\delta)$ for any fixed $\delta > 0$ once q is sufficiently large, but "$x > \exp(q^\delta)$" is far too big to use in this (and many other) applications. On the other hand we believe that this is so for $x > q^{1+\delta}$ for any fixed $\delta > 0$ once q is sufficiently large and can prove it when $x > q^{2+\delta}$ assuming the Generalized Riemann Hypothesis.

However there is an unconditional result that this holds for "almost all" q in this range, where "almost all" is given in the following form:

The Bombieri-Vinogradov Theorem. *For every $A > 0$ there exists $B = B(A) > 0$ such that if x is sufficiently large and $Q = \sqrt{x}/(\ln x)^B$, then*

$$\sum_{q\le Q} \max_{(a,q)=1} \left|\pi(x;q,a) - \frac{\pi(x)}{\phi(q)}\right| \le \frac{x}{(\ln x)^A}.$$

This is believed[28] to be true when $Q = x^{1-\delta}$ for any $\delta > 0$.

We also know good bounds on the number of primes in a segment of an arithmetic progression.

Selberg's version[29] of the Brun-Titchmarsh Theorem. *For all positive X and $x > q$, with a and q coprime integers, we have*

$$\pi(X+x;q,a) - \pi(X;q,a) \le \frac{2x}{\phi(q)\ln(x/q)}.$$

5.3. Proof of the conjecture (of section 4.4) for $\theta \le .11$. (This proof is essentially due to Goldfeld.) First note the identity

$$\begin{aligned}\sum_{\substack{q^a\le 2R\\ q \text{ prime},\ a\ge 1}} \ln q\,\{\pi(2R;q^a,1) - \pi(R;q^a,1)\} &= \sum_{\substack{q^a\le 2R\\ q \text{ prime},\ a\ge 1}} \ln q \sum_{\substack{R<p\le 2R\\ p\equiv 1 \pmod{q^a}}} 1\\ &= \sum_{R<p\le 2R}\ \sum_{q^a\mid p-1} \ln q\\ &= \sum_{R<p\le 2R} \ln(p-1) \sim R, \qquad (5.3)\end{aligned}$$

in which the last estimate may easily be deduced from (5.1). The contribution of the prime powers q^a, with $a \ge 2$ and $q^a \le \sqrt{R}$, to the first sum in (5.3) is

$$\le 4 \sum_{\substack{q^a\le\sqrt{R}\\ a\ge 2}} \frac{\ln q}{q^{a-1}(q-1)} \cdot \frac{R}{\ln R} = O\left(\frac{R}{\ln R}\right),$$

by the Brun-Titchmarsh Theorem for $q^a \le \sqrt{R}$, and then by extending the sum to all integers $q \ge 2$ and noting that this new sum is bounded. For prime powers q^a,

[27] That is, whether $1-\epsilon < \pi(x;q,a)/(\pi(x)/\phi(q)) < 1+\epsilon$ for given small $\epsilon > 0$.

[28] And if so we could prove the conjecture of section 4.4 for any $\theta < 1/2$ and then deduce that the AKS algorithm runs in $O((\log n)^{6+o(1)})$ bit operations by Corollary 4.6.

[29] Though this first appeared in a paper of Montgomery and Vaughan.

with $a \ge 2$ and $q^a \ge \sqrt{R}$, in the first sum in (5.3) we bound the number of primes by the number of terms in the arithmetic progression, namely $\le 2R/q^a$. Now for given q the number of such $a \ge 2$ is $\le (\ln R)/(\ln q)$, and so their contribution to the sum is $\le 2\sqrt{R}(\ln R)$, which totals to $O(R/\ln R)$ when summing over all $q \le (2R)^{1/3}$. For the remaining q we only have the $a = 2$ term which thus contributes $\le 2R\sum_{q>(2R)^{1/3}} \ln q/q^2 = O(R/\ln R)$. Thus the total contribution of all these "$q^a, a \ge 2$" terms in (5.3) is $O(R/\ln R)$.

Next we apply the Bombieri-Vinogradov Theorem with $A = 2$, $Q = \sqrt{R}/(\ln R)^{B(2)}$ and $x = R$ and $2R$ to obtain

$$\sum_{\substack{q\le Q\\ q \text{ prime}}} \ln q(\pi(2R;q,1) - \pi(R;q,1)) \sim \frac{R}{\ln R}\left\{\sum_{\substack{q\le Q\\ q \text{ prime}}} \frac{\ln q}{q-1} + O(1)\right\} \sim \frac{R}{2}$$

by (5.2). Finally we apply the Brun-Titchmarsh Theorem to note that

$$\sum_{\substack{Q<q\le(2R)^{1/2+\theta}\\ q \text{ prime}}} \ln q\ \{\pi(2R;q,1) - \pi(R;q,1)\} \le \sum_{Q<q\le(2R)^{1/2+\theta}} \frac{\ln q}{(q-1)} \cdot \frac{2R}{\ln(R/q)}.$$

To estimate this last sum we break the sum over q into intervals $(QL^i, QL^{i+1}]$ for $i = 0, 1, 2, \dots, I-1$. Here L is chosen so that $QL^I = (2R)^{1/2+\theta}$, and we shall require that $L \to \infty$ as R does, but with $L = R^{o(1)}$. In such an interval we find that each $\ln(R/q) \sim \ln(R^{1/2}/L^i)$, and so the sum over the interval is $\sim 2R\ln L/\ln(R^{1/2}/L^i)$ by (5.2). Thus the quantity above is $\sim 2R\sum_{i=0}^{I-1}\left(\frac{1}{2}\frac{\ln R}{\ln L} - i\right)^{-1}$. Approximating this sum by an integral and then taking $u = t\ln L/\ln R$ we get

$$2R\int_{t=0}^{I-1}\left(\frac{1}{2}\frac{\ln R}{\ln L} - t\right)^{-1} dt \sim 2R\int_0^{\theta}(1/2-u)^{-1}du = 2R\ln(1/(1-2\theta)).$$

Combining the above estimates for the various q and using the fact that $\log q \le \log 2R$ give

$$\sum_{\substack{(2R)^{1/2+\theta}<q\\ q \text{ prime}}} \{\pi(2R;q,1) - \pi(R;q,1)\} \ge \{2c(\theta) + o(1)\}\frac{R}{\ln R}, \tag{5.4}$$

where $c(\theta) := 1/4 - \ln(1/(1-2\theta))$, which is > 0 provided $\theta < \frac{1-e^{-1/4}}{2} \simeq .1105996084$. Note that (5.4) implies the conjecture since $r \le 2R$.

5.4. How can it be so hard to make some results explicit? In section 5.2 we stated that one can unconditionally prove $\pi(x;q,a) \sim \pi(x)/\phi(q)$ when $(a,q) = 1$ where $\phi(q) = \#\{a \pmod q : (a,q) = 1\}$ when $x > \exp(q^\delta)$ for any fixed $\delta > 0$ once q is sufficiently large, but there is a catch. If $\delta \le 1/2$, we do not know how large we mean when we write "q is sufficiently large" (though we do for $\delta > 1/2$). The reader might suppose that this fault exposes a lack of calculation ability on the part of analytic number theorists,[30] but this is not the case. It is in the very nature of the proof itself that, at least for now, an explicit version is impossible, for the proof comes in two parts. The first, under the supposition that the Generalized Riemann

[30]I know I did when I first encountered this statement.

Hypothesis is true, gives an explicit version of the result. The second, under the supposition that the Generalized Riemann Hypothesis is false, gives an explicit version of the result *in terms of the first counterexample*. We strongly believe that the Generalized Riemann Hypothesis is true, but it remains an unproved conjecture, indeed one of the great open problems of mathematics. So, as long as it remains unproved, we seem to be stuck with this situation, for how can we put numbers into the second case in the proof when we believe that there are no counterexamples?

This result of Siegel underpins many of the key results of analytic number theory; hence many of the results inherit this property of being inexplicit. Make Siegel's result explicit and you change the face of analytic number theory, but for now there is no sign that this will happen and so we are lumbered with this considerable burden, particularly when trying to apply analytic number theory results to determine complexity of algorithms.

6. A randomized algorithm with running time $O((\log n)^{4+o(1)})$

Next we describe an algorithm that distinguishes primes from composites and provides a proof within $O((\log n)^{4+o(1)})$ steps. The only drawback is that this is not guaranteed to work. Each time one runs the algorithm the probability that it reports back is $\geq 1/2$, but each run is independent, so after 100 runs the probability that one has not yet distinguished whether the given integer is prime or composite is $< 1/2^{100}$, which is negligible. This is arguably our best solution to at least part of the problem that Gauss described as "so elegant and celebrated." As we discussed in section 1.1, this does not yet work in practice on quite such large numbers as certain tests which are not yet proven but believed to run in polynomial time, but I believe that it is only a matter of time before this situation is rectified.

The algorithm is a modification of AKS given by Bernstein, following up on ideas of Berrizbeitia, as developed by Qi Cheng (and a similar modification was also given by Mihailescu and Avanzi). This is an RP algorithm for primality testing which is faster, easier and more elegant than that of Adleman and Huang. In practice this makes the original AKS algorithm irrelevant, for if we run the "witness" test, which is an RP algorithm for compositeness, half of the time and run the AKS-Berrizbeitia-Cheng-Bernstein-Mihailescu-Avanzi RP algorithm for primality the other half, then a number n is, in practice, certain to yield its secrets faster (in around $O((\log n)^{4+o(1)})$ steps) than by the original AKS algorithm!

6.1. Yet another characterization of the primes. For a given monic polynomial $f(x)$ with integer coefficients of degree $d \geq 1$ and positive integer n, we say that $\mathbb{Z}[x]/(n, f(x))$ is an *almostfield* with parameters $(e, v(x))$ if

(a) Positive integer e divides $n^d - 1$,
(b) $v(x)^{n^d-1} \equiv 1 \bmod (n, f(x))$, and
(c) $v(x)^{(n^d-1)/q} - 1$ is a unit in $\mathbb{Z}[x]/(n, f(x))$ for all primes q dividing e.

If n is prime and $f(x) \pmod n$ is irreducible, then $\mathbb{Z}[x]/(n, f(x))$ is a field; moreover any generator $v(x)$ of the multiplicative group of elements of this field satisfies (b) and (c) for any e satisfying (a).

Bernstein. *For given integer $n \geq 2$, suppose that $\mathbb{Z}[x]/(n, f(x))$ is an almostfield with parameters $(e, v(x))$ where $e > (2d \log n)^2$. Then n is prime if and only if*

- *n is not a perfect power,*

- $(t-1)^{n^d} \equiv t^{n^d} - 1 \mod (n, f(x), t^e - v(x))$ *in* $\mathbb{Z}[x,t]$ *(that is, we work with polynomials with integer coefficients in independent variables* t *and* x*).*

Proof. Write $N = n^d$ and $v = v(x)$. If n is a perfect power, then n is composite. If n is prime, then the second condition holds by the Child's Binomial Theorem. So we may henceforth assume that n is not a perfect power and is not prime, and we wish to show that $(t-1)^{n^d} \not\equiv t^{n^d} - 1 \mod (n, f(x), t^e - v(x))$. Let p be a prime dividing n and $h(x)$ an irreducible factor of $f(x) \pmod p$, so that $\mathbb{F} = \mathbb{Z}[x]/(p, h(x))$ is isomorphic to a finite field. Let $P = |\mathbb{F}| = p^{\deg h}$, and note that since $p < n$ and $\deg h \le \deg f$, hence $P < N$.

Let $\zeta \equiv v^{(N-1)/e} \mod (p, h(x))$ so that ζ is an element of order e in $\mathbb{F}$: To see this, note that $\zeta^e \equiv v^{N-1} \equiv 1 \mod (p, h(x))$ by (b); whereas if ζ had order m, a proper divisor of e, then let q be a prime divisor of e/m so that $1 \equiv \zeta^{e/q} \equiv v^{(N-1)/q} \mod (p, h(x))$, contradicting (c).

The polynomials of the form $\prod_{i=0}^{e-1} (\zeta^i t - 1)^{a_i}$ in $\mathbb{F}[t]$ are distinct, and so those of degree $\le e-1$ are distinct in $\mathbb{F}[t]/(t^e - v)$.

Now $t^N = t^{N-1} t \equiv v^{(N-1)/e} t \pmod{t^e - v}$, so that $t^N \equiv \zeta t \mod (p, h(x), t^e - v)$. Thus our second criterion implies that $(t-1)^N \equiv \zeta t - 1 \mod (p, h(x), t^e - v)$. Moreover replacing t by $\zeta^i t$ gives $(\zeta^i t - 1)^N \equiv \zeta^{i+1} t - 1 \mod (p, h(x), t^e - v)$ for any integer $i \ge 0$ (since $(\zeta^i t)^e - v = t^e - v$), and thus $(t-1)^{N^i} \equiv \zeta^i t - 1 \mod (p, h(x), t^e - v)$ by a suitable induction argument. Note that $(t-1)^{N^e} \equiv (t-1) \mod (p, h(x), t^e - v)$.

Therefore for proper subsets I of $\{0, 1, \dots, e-1\}$ the powers $(t-1)^{\sum_{i \in I} N^i} \equiv \prod_{i \in I} (\zeta^i t - 1) \mod (p, h(x), t^e - v)$ all have degree $\le e-1$ and so are distinct polynomials, and thus there are at least $2^e - 1$ distinct powers of $(t-1) \mod (p, h(x), t^e - v)$.

Now e is the order of an element of $\mathbb{F}^*$, which is a cyclic group of order $P-1$, and so $P-1$ is a multiple of e. Therefore $v^{(P-1)/e}$ is an eth root of 1 in $\mathbb{F}$, so must be a power of ζ, say ζ^ℓ. Arguing as in two paragraphs above, but now with N and ζ replaced by P and ζ^ℓ, we see that $(t-1)^{P^j} \equiv \zeta^{j\ell} t - 1 \mod (p, h(x), t^e - v)$. Combining these results we obtain that $(t-1)^{N^i P^j} \equiv \zeta^{i + j\ell} t - 1 \mod (p, h(x), t^e - v)$ for all integers i, j.

There are more than e pairs of integers (i, j) with $0 \le i, j \le [\sqrt{e}]$, and so there exist two numbers of the form $i + j\ell$ (with i and j in this range) that are congruent (mod e), say $i + j\ell \equiv I + J\ell \pmod e$. Therefore if $u := N^i P^j$ and $U := N^I P^J$, then $(t-1)^u \equiv \zeta^{i+j\ell} t - 1 = \zeta^{I + J\ell} t - 1 \equiv (t-1)^U \mod (p, h(x), t^e - v)$. We will show that $t - 1$ is a unit $\mod (p, h(x), t^e - v)$ so that we can deduce that there are no more than $|U - u| < (NP)^{\sqrt{e}} - 1 < N^{2\sqrt{e}} - 1$ distinct powers of $(t-1) \mod (p, h(x), t^e - v)$, and thus $2^e < N^{2\sqrt{e}}$, contradicting the hypothesis.

Now $v(x) \ne 1$ in $\mathbb{F}$ by (c), so that $t - 1$ is not a factor of $t^e - v(x)$ in $\mathbb{F}[t]$; in other words $t - 1$ is a unit in the ring $\mathbb{F}[t]/(t^e - v(x)$, that is $\mod (p, h(x), t^e - v)$. □

6.2. Running this primality test in practice. We will show that if n is prime, then an almostfield may be found rapidly in random polynomial time. The primality test given by the statement of the theorem may evidently be implemented by similar methods to those we discussed previously.

Assume that n is prime. By the inclusion-exclusion formula one can prove that there are[31] $(1/d)\sum_{\ell|d}\mu(d/\ell)n^\ell$ irreducible polynomials mod n of degree d. The biggest term here is the one with $\ell = d$; that is, roughly $1/d$ of the polynomials of degree d are irreducible. Thus selecting degree d polynomials at random we should expect to find an irreducible one in $O(d)$ selections. Verifying f is irreducible can be done by checking, via the Euclidean algorithm, that $x^{n^d} - x \equiv 0 \bmod (n, f(x))$ and $x^{n^{d/q}} - x$ is a unit in $\mathbb{Z}[x]/(n, f(x))$ for all primes q dividing d. Once we have found f we know that $\mathbb{Z}[x]/(n, f(x))$ is a field. The elements of $\mathbb{Z}[x]/(n, f(x))$ can be represented by the polynomials $v(x) \bmod n$ of degree $< d$. The proportion of these that satisfy (b) and (c) is $\prod_{p|e}(1 - 1/p) > 1/2 \ln\ln e$, and so selecting such $v(x)$ at random we should expect to find $v(x)$ satisfying (b) and (c) in $O(d)$ selections.

The main part of the running time comes in verifying that $(t-1)^{n^d} \equiv t^{n^d} - 1 \bmod (n, f(x), t^e - v(x))$, which will take $d \log n$ steps, each of which will cost $O(de(\log n)^{1+o(1)})$ bit operations, giving a total time of $O(d^2 e(\log n)^{2+o(1)})$ bit operations. The conditions $d \geq 1,\ e > (2\log n)^2$ imply that the running time cannot be better than $O((\log n)^{4+o(1)})$, and we will indicate in the next section how to find d and e so that we obtain this running time.

6.3. More analytic number theory. To find an almostfield when n is prime we need to find d and e for which e divides $n^d - 1$ and with d and e satisfying certain conditions. Constructions typically give e as a product of primes p which do not divide n and for which $p-1$ divides d, since then p divides $n^d - 1$ by Fermat's Little Theorem, and thus e divides $n^d - 1$.

However, to ensure that e is large, for instance $e > (2d \log n)^2$ as required in the hypothesis of Bernstein's result, we need to use the ideas of analytic number theory. Our general construction looks as follows: For given $z < y$, with $z \geq \epsilon y$ for fixed $\epsilon > 0$, let d be the least common multiple of the integers up to z and e be the product of all primes $p \leq y$ such that all prime power divisors q^a of $p - 1$ are $\leq z$. Note that $d = \exp(z + o(z))$ by the prime number theorem and $e = \prod_{p\leq y} p / \prod_{p\in\mathcal{P}} p$, where $\mathcal{P}$ is the set of primes $p \leq y$ for which $p - 1$ has a prime power divisor q^a which is $> z$.

If $p \in \mathcal{P}$ write $p - 1 = kq^a$ with $q^a > z$, so that $k < y/z \leq 1/\epsilon$. Now the number of $q^a \in (z, y)$ with $a \geq 2$ is $O(\sqrt{y})$, and so there are $O(\sqrt{y}/\epsilon)$ values of $p \in \mathcal{P}$ for which $p - 1$ has a prime power divisor q^a with $a \geq 2$.

In our first construction we take $y = 4z$, so that if $a = 1$, then we have a prime pair of the form $q, kq + 1$ with $k < 4$, and so $k = 2$. For this we use a well-known bound on the number of prime pairs of the form $q, 2q + 1$:

Lemma 6.1. *There exists a constant $c > 0$ such that there are $\leq 2cx/(\log x)^2$ primes $q \leq x$ for which $2q + 1$ is also prime, for all $x \geq 2$.*

Therefore $|P| = O(y/(\log y)^2)$ and so $e = \exp(y + o(y))$ by the prime number theorem. If we take $y = (4+\epsilon)\log\log n$, then we get $e > (2d \log n)^2$ as required, and the values of d and e can, in practice, be found quickly. However, by the remarks of the previous section the running time will be $O((\log n)^{8+O(\epsilon)})$, so we need to choose d and e slightly differently.

[31]Here $\mu(m)$, the Mobius function, is defined to be $(-1)^k$ if m is squarefree and has k distinct prime factors, and 0 otherwise.

This time we take $z = \epsilon y$ with $y = (2+3\epsilon)\log\log n$. We need the generalization of Lemma 6.1 to prime pairs of the form $q, kq+1$:

Lemma 6.2. *There exists an absolute constant $c > 0$ such that there are $\le c(k/\phi(k))(x/(\log x)^2)$ primes $q \le x$ for which $kq+1$ is also prime, for all even integers k and all $x \ge 2$.*

In this case, corresponding to each prime $p \in \mathcal{P}$ with $a = 1$, we have a prime pair $q, kq+1$ with $k \le 1/\epsilon$ and $q \le y/k$. For given $k \le 1/\epsilon$ there are $\le cy/(\log(\epsilon y))^2$ such prime pairs, by Lemma 6.2 with $x = y/k$, since $\phi(k) \ge 1$. Therefore $|\mathcal{P}| = O(y/(\epsilon(\log y)^2) + \sqrt{y}/\epsilon) = o(y/\log y)$, so the product of the primes in $\mathcal{P}$ is $\le y^{|\mathcal{P}|} = \exp(o(y))$. Thus $e = \exp(y + o(y))$ by the prime number theorem, and so $e > (2d\log n)^2$ as required; but now the running time will be $O((\log n)^{4+O(\epsilon)})$, and letting $\epsilon \to 0$ we get the desired result.

6.4. Bernstein's construction was originally analyzed using a beautiful result of Prachar. If $p-1$ divides d for every prime p dividing e, as above, then how large can e be in terms of d? An obvious upper bound for e is $\prod_{m|d}(m+1)$ in the case that one more than each divisor of d is a prime. So if $\tau(d)$ is the number of divisors of d, then $e < d^{\tau(d)}$. Can we obtain e which is anywhere near this big? Prachar's idea is to look for primes of the form $mk+1$ for some small integer k, as m runs through the divisors of d. Under the assumption of the Generalized Riemann Hypothesis, for each fixed $m|d$, there are $> d^2/2\log d$ primes of the form $mk+1$ with $d^2 < k < 3d^2$, so their product is $> \exp(d^2)$. Therefore there exists some value of $k < 3d^2$ such that the product e of the primes of the form $km+1$ with $m|d$ is $> \exp(\tau(d)/2)$ (and we replace d above by $D := kd < 3d^3$). If we took the original d to be the product of the primes $\le z$, then, by the prime number theorem, $d = \exp(z + o(z))$ and $\log e > 2^{\pi(z)-1}$ so that $D < (\log e)^{c\log\log\log e}$, for some constant $c > 0$. This argument can be justified without assumption, since as we discussed in section 5.2, we "almost always" get the expected number of primes in an arithmetic progression in the relevant ranges (see [2] for the technical details).

6.5. Lenstra's 1985 finite field primality test uses many of the same ideas as the AKS test and these variants. It is surprising how close researchers were twenty years ago to obtaining a polynomial time primality test.

Lenstra. *For a given almostfield $\mathbb{Z}[x]/(n, f(x))$ with parameters $(e, v(x))$, if*

(d) $g(T) := \prod_{i=0}^{d-1} (T - v(x)^{n^i}) \in (\mathbb{Z}[x]/(n, f(x)))[T]$,

then $p \equiv n^j \pmod e$ for some j, $0 \le j \le d-1$, for each prime p dividing n.

Proof. Let $h(x)$ be an irreducible factor of f in $(\mathbb{Z}/p\mathbb{Z})[x]$, and define $\mathbb{F} := \mathbb{Z}[x]/(p, h(x))$. Now $g(v(x)) = 0$ by (d), so that $g(v(x)^p) = g(v(x))^p = 0$ in $\mathbb{F}$, by the Child's Binomial Theorem. Therefore $v(x)^p = v(x)^{n^i}$ in $\mathbb{F}$ for some i. This implies that $p \equiv n^i$ mod the order of $v(x)$ in $\mathbb{F}$. The result follows since the order of $v(x)$ in $\mathbb{F}$ is divisible by e, by (b) and by (c). □

We use this as a primality test by selecting d and e as in section 6.4 so that $n^2 > e > n$ and $D < (\log n)^{O(\log\log\log n)}$ and then finding an almostfield as described in section 6.2. If (d) does not hold, then n is composite (since our choice of $f(x)$ is irreducible if n is prime). If d does hold, then the "candidates" to be prime factors

of n by Lenstra's theorem are the least residues of $n, n^2, n^3, \dots, n^{d-1} \pmod e$, and these are easily checked by trial division. The running time of this random test for the primality of n is thus $O((\log n)^{O(\log\log\log n)})$, just a smidgin slower than polynomial time. (This is a simplification of Lenstra's test, which was actually a little more involved but compensated by being faster in terms of the constant implicit in the "O" in the exponent.)

7. Stop the press: $(\log n)^6$ achieved

7.1. Lenstra and Pomerance obtain the desired running time. Lenstra and Pomerance significantly modified the AKS algorithm so that it will, in theory, work as fast as can be hoped for, namely in $\tilde{O}((\log n)^6)$ bit operations. Moreover it is possible to give explicit constants in this statement, in other words a computable upper bound on the running time that holds for all n.

Their key idea is to replace the polynomial $\Phi_r(x)$ in AKS by a polynomial $f(x)$ with certain properties: For a given monic polynomial $f(x)$ with integer coefficients of degree d, and positive integer n, we say that $\mathbb{Z}[x]/(n, f(x))$ is a *pseudofield* if

(a) $f(x^n) \equiv 0 \bmod (n, f(x))$,
(b) $x^{n^d} - x \equiv 0 \bmod (n, f(x))$, and
(c) $x^{n^{d/q}} - x$ is a unit in $\mathbb{Z}[x]/(n, f(x))$ for all primes q dividing d.

When n is prime and $f(x)$ is irreducible mod n, then these criteria are all true and $\mathbb{Z}[x]/(n, f(x))$ is a field.

Lenstra and Pomerance. *For given integer $n \ge 2$ let d be an integer in $((\log n)^2, n)$ for which there exists a monic polynomial $f(x)$ of degree d with integer coefficients such that $\mathbb{Z}[x]/(n, f(x))$ is a pseudofield. Then n is prime if and only if*

- *n is not a perfect power,*
- *n does not have any prime factor $\le d$,*
- *$(x+a)^n \equiv x^n + a$ mod $(n, f(x))$ for each integer $a, 1 \le a \le A := \sqrt{d}\log n$.*

One can easily see that this theorem has its genesis in the work of Agrawal, Kayal and Saxena, yet is more general. This generality is what allows them to prove that it can be achieved in far fewer steps.

Evidently for a given f one can quickly determine whether one gets a pseudofield, and if so check the criteria of the theorem. Thus if we can quickly find an f which gives a pseudofield, this approach will lend itself to a quick primality test. Lenstra and Pomerance's construction of f comes back full circle to Gauss's *Disquisitiones* and his construction of regular n-gons, in particular what are now known as "Gaussian periods".

In the next subsection we will give Agrawal's proof of Lenstra and Pomerance's theorem. In section 7.3 we will introduce Gaussian periods and discuss how to construct f with the required properties. In sections 7.4 and 7.5 we sketch the proof that this can be done in the required number of steps.

7.2. Proof that this is a characterization of the primes. Suppose that n is composite and satisfies the three hypotheses. Let p be a prime dividing n and let $h(x)$ be an irreducible factor of $f(x) \pmod p$, so that $\mathbb{F} \equiv \mathbb{Z}[x]/(p, h(x))$ is isomorphic to a finite field.

As in section 4, let H be the elements $\pmod{p, f(x)}$ generated multiplicatively by $x, x+1, x+2, \dots, x+[A]$; let G be the (cyclic) subgroup of $\mathbb{F}$ generated multiplicatively by $x, x+1, x+2, \dots, x+[A]$; and let S be the set of positive integers of the form $p^i n^j$ with $i, j \geq 0$. Define r to be the order of $x \pmod{p, f(x)}$ so that d is the order of $n \bmod r$ by the definitions (b) and (c) of a pseudofield, and further $x^{n^0}, x^{n^1}, \dots, x^{n^{d-1}}$ are distinct $\pmod{p, f(x)}$ and even $\pmod{p, h(x)}$. Therefore the polynomial $g(T) := \prod_{i=0}^{d-1}(T - x^{n^i}) \in \mathbb{F}[T]$ has distinct roots; moreover $g(x^p) = g(x)^p = 0$ in $\mathbb{F}$ and therefore x^p must equal x^{n^j} in $\mathbb{F}$ for some j. This implies that $p \equiv n^j \pmod{r}$, and so if R is the subgroup of $(\mathbb{Z}/r\mathbb{Z})^*$ generated by n and p, then R is in fact generated by n alone and so has d elements.

The same proof works for Lemma 4.1 except for the observation that $x^r - 1$ divides $x^{kr} - 1$ for any $k \in S$. We replace this by the observation that $f(x^k) \equiv 0 \pmod{(p, f(x))}$ for all $k \in S$, which holds since $f(x^n) \equiv 0 \pmod{(p, f(x))}$ by (a), and $f(x^p) \equiv f(x)^p \pmod{p} \equiv 0 \pmod{(p, f(x))}$ by the Child's Remainder Theorem $\pmod{p}$.

Lemma 4.2 follows with much the same proof (with $(p, x^r - 1)$ replaced by $(p, f(x))$) since $(p, f(x))$ divides $(p, x^r - 1)$. The remarks after (4.4) follow simply by replacing the first sentence there by (b) and otherwise replacing $(p, x^r - 1)$ by $(p, f(x))$. Therefore (4.4) holds and Lemma 4.3 holds, so just as in section 4.2, we note that $|G| \geq 2^{B+1} - 1$ where $B := [A]$, so that $|G| > n^{\sqrt{|R|}} - 1$, contradicting (4.4).

7.3. The construction of f. Let $\zeta_r = e^{2i\pi/r}$ for prime r. If q divides $r-1$, we define the *Gaussian period* to be $\eta = \sum_{j \in J} \zeta_r^j$ where $J = J_{r,q} := \{j \pmod{r} : j^{(r-1)/q} \equiv 1 \pmod{r}\}$ is the set of residue classes $\pmod{r}$ which are coprime to r and which are qth powers [32] $\pmod{r}$. Now J is a subgroup of the cyclic group $(\mathbb{Z}/r\mathbb{Z})^*$, and so $J = \{g^{qi} : 0 \leq i < (r-1)/q\}$ for a generator g of $(\mathbb{Z}/r\mathbb{Z})^*$. Moreover J has q cosets in $(\mathbb{Z}/r\mathbb{Z})^*$, namely $g^i J$ for $0 \leq i \leq q-1$. Thus $\eta = \eta_0$ has conjugates[33] $\eta_i := \sum_{j \in J} \zeta_r^{g^i j}$ for $i = 0, 1, 2, \dots, q-1$. The minimum polynomial for η over $\mathbb{Q}$ is $f(x) = \prod_{i=0}^{q-1}(x - \eta_i)$ which has degree q, is monic, and has integer coefficients.

Let p be prime, different from r. In the field $\mathbb{Q}(\zeta_r)$, primes over $\mathbb{Q}$ (like p) might factor further[34] into prime ideals, so suppose $\mathcal{P}$ is a prime ideal factor of p in $\mathbb{Q}(\zeta_r)$. Also $f(x)$ might factor $\pmod{p}$, so let $g(x)$ be that factor of $f(x)$ $\pmod{p}$ for which $g(\eta) \equiv 0 \pmod{\mathcal{P}}$. Now, the Child's Binomial Theorem tells us that $g(x^p) \equiv g(x)^p \pmod{p}$, and hence mod $\mathcal{P}$, and therefore $g(\eta^p) \equiv g(\eta)^p \equiv 0 \pmod{\mathcal{P}}$. Thus $\eta^p \equiv \eta_k \pmod{p}$ where $\eta_k \in g^k J$ is a root of $g(x) \bmod \mathcal{P}$, and by similar arguments we find that $\eta_{ik \bmod q}$ are also, for $i = 0, 1, \dots, q-1$. These are all distinct exactly when the order of $p^{(r-1)/q} \pmod{r}$ is q. From this it is possible to deduce that $f(x)$ is irreducible $\pmod{p}$ if and only if the order of $p^{(r-1)/q} \pmod{r}$ is q.

In this way we can construct an irreducible polynomial of degree q over $\mathbb{F}_p$, the minimum polynomial of the Gaussian period η. Furthermore, if we have several

[32] That is, there exists $i \pmod{r}$ for which $i^q \equiv j \pmod{r}$.

[33] The automorphisms of the field $\mathbb{Q}(\zeta_r)$ are each induced by a map $\zeta_r \to \zeta_r^k$ for some k, $1 \leq k \leq r-1$.

[34] For example, $5 = (1 + 2\zeta_4)(1 - 2\zeta_4)$ in $\mathbb{Q}(\zeta_4)$.

primes $r_1, r_2, \dots, r_k$ and pairwise coprime positive integers $q_1, q_2, \dots, q_k$ for which q_i divides $r_i - 1$ for each i, then $f(x)$, the minimum polynomial of $\eta_1\eta_2\cdots\eta_k$ over $\mathbb{Q}$, has degree $q_1q_2\dots q_k$ and is irreducible $\pmod p$ if and only if the order of $p^{(r_i-1)/q_i} \pmod r$ is q_i for each i. This leads to our construction of f: For given n we look for q_i and r_i as above, though with p replaced by n. Note that if n is prime, then $\mathbb{Z}[x]/(n, f(x))$ is a pseudofield. If n is composite, then either $\mathbb{Z}[x]/(n, f(x))$ fails to be a pseudofield, in which case we have a proof that n is composite, or it is a pseudofield, in which case we can apply the theorem of Lenstra and Pomerance, provided $q_1q_2\dots q_k > (\log n)^2$, to test whether n is prime. In fact they prove that such f exist with $q_1q_2\dots q_k$ just a little bigger than $(\log n)^2$:

Proposition 7.1. *There exists a computable constant n_0 such that if $n \ge n_0$, then there exist primes $r_1, r_2, \dots, r_k < (\log n)^2$ and pairwise coprime positive integers $q_1, q_2, \dots, q_k$ with q_i dividing $r_i - 1$ and for which the order of $n^{(r_i-1)/q_i} \pmod r$ is q_i for each i, such that $(\log n)^2 < q_1q_2\dots q_k < 4(\log n)^2$.*

To determine whether the order of $n^{(r_i-1)/q_i} \pmod r$ is q_i we need only check that $n^{r_i-1} \equiv 1 \pmod{r_i}$ whereas $n^{(r_i-1)/q_i} \not\equiv 1 \pmod{r_i}$. Using Proposition 7.1 naively we can determine suitable values for r_i and q_i in $O((\log n)^3)$ steps. Moreover all of the results used in analyzing the steps of this algorithm can be written down with the constants explicitly given. In other words the running time of the Lenstra and Pomerance algorithm can be given totally explicitly.

7.4. The continuous postage stamp problem. Post offices issue stamps of different value to allow the sender to affix a suitable amount for postage. If S is the finite set of positive integers which gives the values of the different stamps (in cents), then to be able to make up any price (in cents) using an assortment of stamps, it is necessary that $\gcd\{s \in S\} = 1$. On the other hand if $\gcd\{s \in S\} = 1$, then we know that every sufficiently large integer can be represented as a non-negative integral linear combination of elements of S, but not necessarily every positive integer. The Frobenius postage problem is to determine the largest integer that cannot be so represented. If $|S| = 2$, then this is easy, but there is no known formula for arbitrary S; in fact this problem is `NP-hard`.[35]

Perhaps the most surprising part of Lenstra and Pomerance's proof of Proposition 7.1 is that, ultimately, their analytic argument depends on a continuous analogy to the Frobenius postage problem: Let S be an open subset of the positive reals that is closed under addition. One might expect that if S is sufficiently dense, then that forces every sufficiently large real number to be in S; here "sufficiently dense" is in terms of $\int_{0\le t\le x,\ t\in S} \frac{dt}{t}$. They conjectured the following criterion, which was proved by Daniel Bleichenbacher:

Lemma 7.2. *If S is an open subset of the positive reals that is closed under addition for which $\int_{0\le t\le u,\ t\in S} \frac{dt}{t} > u$ for some $u \in (0, 1]$, then $1 \in S$.*

7.5. The analytic part of the argument, the proof of Proposition 7.1, has several complicated details, so we will just sketch the main ideas. Let $x = (\log n)^{1+3\eta}$ where $\eta = 1/\log\log\log n$.

One can prove that for most primes r, much of the size of $r - 1$ is made up of "large" primes, that is, that $\prod_{q|r-1,\ q>x^{\eta^2}} q > x^{1-\eta}$ for all but $O(x/(\log x)^3)$ of the

[35]That is, it is at least as difficult as any other `NP` problem.

primes $r \le x$. If we have such an r, then it is very likely that $n^{(r-1)/q} \pmod r$ has order q for some such q dividing $r-1$, for if not, then n has order $< x^\eta \pmod r$, but there are few such r as can be shown by an argument like that in the proof of Lemma 4.5.

Let $\mathcal{Q}$ be the set of primes $q \in (x^{\eta^2}, x^{1/2}]$ such that there exists a prime $r \le x$ which is $\equiv 1 \pmod q$ and for which q is the order of $n^{(r-1)/q} \pmod r$. Using the above observations, together with a sharpened form of the Brun-Titchmarsh theorem, one can show that $\sum_{q\in\mathcal{Q}} 1/q > 3/11 - o(1)$.

Let $\epsilon = \eta^2/2\log\log n$. For each $q \in \mathcal{Q}$ let $\tau_q := \log q/2\log\log n$ and define $S_0 = \cup_{q\in\mathcal{Q}}(\tau_q - \epsilon, \tau_q)$ and then S to be the closure of S_0 under addition. By the prime number theorem, $\sum_{x^\alpha<q<x^\beta,\ q \text{ prime}} 1/q \sim \ln(\beta/\alpha) = \int_\alpha^\beta \frac{dt}{t}$, from which we easily deduce that $\int_{0\le t\le 1/4+\eta,\ t\in S} \frac{dt}{t} > \{1+o(1)\}\sum_{q\in\mathcal{Q}} 1/q > 3/11 - o(1)$ by the above. Therefore $1 \in S$ by Lemma 7.2 so that there exists $V \subset S_0$ such that $\sum_{v\in V} v = 1$. Let U be the set of $q \in \mathcal{Q}$ for which $v \in (\tau_q - \epsilon, \tau_q)$ as we run through $v \in V$ (this simple argument may be modified to show that different v give rise to different q). Thus $0 < \sum_{q\in U} \log q - 2\log\log n < 2$ as required for Proposition 7.1.

8. Minor improvements and tempting ideas

In this section we begin with four areas in which there are elegant ideas for improving the proof of section 4, selected for the beauty of the mathematics involved. Next we discuss Agrawal and his students' quest for a polynomial time primality test, and then develop their ideas to rewrite the AKS test without polynomials.

8.1. Varying the elements of G. In section 4 we took G to be the subgroup of $\mathbb{F}$ generated by $V := \{x+1, \dots, x+[A]\}$ and verified that $n \in S = S_G$ (by checking (4.1) for $a = 1, 2, \dots, A$). If we can deduce from this knowledge that (4.1) holds for other values of a, then we can expand the size of V without changing the proof except for showing that G is larger than before. So for any $a,\ 1 \le a \le A$ verify that $a^n \equiv a \pmod n$ (else n is composite), and then select integer b so that $ba \equiv 1 \pmod n$. Now, (4.1) can be rewritten as $(x+a)^n - (x^n+a) = nu(x) + (x^r-1)v(x)$ for certain $u(x), v(x) \in \mathbb{Z}[x]$, and one can show that $\deg u < n,\ \deg v < n-r$ using the Euclidean algorithm in $\mathbb{Z}[x]$. Replacing x by $1/x$ and multiplying through by $(bx)^n$ we obtain $(b+abx)^n - (b^n + (ab^n)x^n) = nU(x) - (x^r-1)V(x)$ where $U(x) := b^n x^n u(1/x),\ V(x) := b^n x^{n-r} v(1/x) \in \mathbb{Z}[x]$. Now $(x+b)^n \equiv (abx+b)^n \pmod n$ and $x^n \equiv (ab^n)x^n \pmod n$, so we have proved that (4.1) also holds when we replace a by b. Thus we may replace V by $V \cup \{a^{-1} \pmod n :\ 1 \le a \le A\}$, a set that will be almost twice as large (and thus our lower bound on the size of G will be squared).

Just as we replaced x by x^{-1} in (4.1) to find another case of the Child's Binomial Theorem, so we may also replace x by x^k for any integer k since $x^r - 1$ divides $x^{kr} - 1$, and thus we may generate G using all polynomials of the form

$$\{x^k + a :\ k \in \mathbb{N},\ 1 \le a \le A\} \cup \{x^k + b :\ k \in \mathbb{N},\ 1 \le a \le A,\ b \equiv a^{-1} \pmod n\}.$$

From Lemma 4.3 we deduced, in section 4.2, that each polynomial generated as a product of the polynomials $\{x, x+1, \dots, x+[A]\}$ which has degree $< |R|$ belongs to a different element of G, and thus we obtained a lower bound on the size of G by counting the number of such polynomials without repeated roots. Voloch suggested a cleverer counting argument. Suppose that G contains k distinct linear

polynomials $x+a_1,\dots,x+a_k$. We wish to find the largest possible subset U of $\mathbb{Z}^k$ with the property that if $\mathbf{u},\mathbf{v}\in U$, then

$$\sum_{1\le i\le k}\max\{u_i-v_i,0\}\text{ and }\sum_{1\le j\le k}\max\{v_j-u_j,0\}\le \ell, \tag{8.1}$$

for $\ell=|R|-1$. Letting $m_i=\min_{\mathbf{u}\in U}u_i$, the polynomials

$$\{g_{\mathbf{u}}=\prod_{1\le i\le k}(x+a_i)^{u_i-m_i}:\ \mathbf{u}\in U\}$$

are all distinct elements of G, for if $g_{\mathbf{u}}=g_{\mathbf{v}}$ in $\mathbb{F}$, then

$$\prod_i (x+a_i)^{\max\{u_i-v_i,0\}}=\prod_j (x+a_j)^{\max\{v_j-u_j,0\}}$$

which are both of degree $<|R|$, and this contradicts Lemma 4.3. Thus $|G|\ge|U|$. Determining $u(k,\ell)$, the size of the largest such set U, is an elegant and open[36] combinatorial problem about which we know the following:

Define $U_{s,t}:=\{\mathbf{u}\in\mathbb{Z}^k:\ \sum_{1\le i\le k}\max\{u_i,0\}\le s$ and $\sum_{1\le i\le k}\max\{-u_i,0\}\le t\}$. Evidently $U_{\ell,0}=\{\mathbf{u}\in\mathbb{Z}^k:$ Each $u_i\ge 0$ and $\sum_{1\le i\le k}u_i\le\ell\}$ satisfies (8.1). Voloch noted that $U_{s,\ell-s}$ satisfies (8.1) for any $0\le s\le\ell$, for if $\mathbf{u},\mathbf{v}\in U_{s,\ell-s}$, then $\sum_{1\le i\le k}\max\{u_i-v_i,0\}\le\sum_{1\le i\le k}(\max\{u_i,0\}+\max\{-v_i,0\})\le s+(\ell-s)=\ell$. On the other hand if we fix $\mathbf{v}\in U$, then $U\subset\mathbf{v}+U_{\ell,\ell}$ by (8.1). Thus we deduce that $|U_{\ell,\ell}|\ge u(k,\ell)\ge\max_{0\le s\le\ell}|U_{s,\ell-s}|$. We can determine the number of elements of $U_{s,t}$ by counting those $\mathbf{n}\in U_{s,t}$ with exactly j values of i for which $u_i\ge 0$ to obtain

$$|U_{s,t}|=\sum_{j=0}^{k}\binom{k}{j}\binom{s+j}{j}\binom{t}{k-j}.$$

8.2. Upper bounds on $|G|$, revisited. In section 4.1 the "pigeonhole principle" was used to find two elements of R which are congruent mod r. Poonen and Lenstra replaced this by

Minkowski's convex body theorem. *Let Λ be a lattice which is a sublattice of $\mathbb{Z}^2$. Let U be a convex*[37] *body in $\mathbb{R}^2$ which is symmetric about the origin. If the area of U is at least 4 times the determinant*[38] *of Λ, then U contains a lattice point of Λ other than the origin.*

We also note the following lemma which is left as an elementary exercise for the reader:

Lemma 8.1. *For positive a,b,c define $T=\{(x,y)\in\mathbb{R}^2:\ x,y\ge 0$, and $ax+by\le c\}$ and $U=\{\mathbf{t}-\mathbf{t}':\mathbf{t},\mathbf{t}'\in T\}$. Then $U=T\cup -T\cup V\cup -V$, where $V=\{(x,-y)\in\mathbb{R}^2:\ 0\le ax,by\le c\}$, so U is a symmetric convex body in $\mathbb{R}^2$, whose area is 6 times the area of T.*

Take $T=\left\{(x,y)\in\mathbb{R}^2_{\ge 0}:p^x(n/p)^y\le n^{\sqrt{|R|/3}}\right\}$ in Lemma 8.1 with $a=\log p$, $b=\log(n/p)$ and $c=\sqrt{|R|/3}\log n$. Then $\mathrm{Area}(U)=6\mathrm{Area}(T)=3c^2/ab\ge 4|R|$, since $\log p\log(n/p)\le\frac14\log^2 n$. The lattice $\Lambda=\left\{(i,j)\in\mathbb{Z}^2:p^i(n/p)^j\equiv 1\pmod r\right\}$ has determinant $|R|$, so by Minkowski's convex body theorem, T contains two distinct

[36] As far as I know.

[37] That is, U contains the line segment connecting any two points in U.

[38] That is, the area of the smallest parallelogram of Λ.

points, call them (i, j) and (I, J), which are congruent mod r. As in section 4.1, $|G|$ divides $p^i(n/p)^j - p^I(n/p)^J$, so that $|G| \le n^{\sqrt{|R|/3}} - 1$, an improvement on (4.3).

8.3. Using polynomials of higher degree in G. Above we obtained lower bounds on the size of G by determining a large set of polynomials in G of degree $< |R|$, since Lemma 4.3 guarantees that they will be distinct. Voloch showed that few pairs of polynomials in G of slightly larger degree can be equal, and so improved the lower bounds on $|G|$. His method revolves around one of the most remarkable results in the arithmetic of function fields:

abc-theorem for polynomials. *If $a, b, c \in \mathbb{C}[t]$ and not all in $\mathbb{C}$, with $a + b = c$, where a and b have no common factors, then the degrees of a, b and c are each less than the total number of distinct roots of abc.*

For a proof, see [13]. By the example in Theorem 1, $t^p + 1 = (t+1)^p$, this result does not carry over to fields of characteristic p, for in this example the maximum degree is p and yet there are just 2 distinct roots. However it is not difficult to modify the proof in [13] to show that if $a + b = c$ where a, b, c are genuine coprime polynomials in a field of characteristic p, then either $a = A^p, b = B^p$ and $c = C^p$ where $A + B = C$ or the above conclusion holds. From this we deduce

Lemma 8.2. *Let K be a field of characteristic p. Suppose $a, b, c \in K[t]$ and have no common factor, and that no two of $u, v, w \in K[t]$ have a common factor, where $au + bv = cw$ and $g = (au, bv, cw)$. Then $g = (a, b)(a, c)(b, c)$ and either au, bv and cw each equal g times the p^{th} power of a polynomial in K, or*

$$2 \max\{\deg u, \deg v, \deg w\} > \max(\deg a, \deg b, \deg c) - \#\left\{\text{distinct roots of } abc/g^2\right\}.$$

Proof. Note that (a, b) divides $au + bv = cw$ and has no common factor with c, so divides w. Similarly (b, c) divides u, and (a, c) divides v. Therefore $g = (a, b)(a, c)(b, c)$ and g divides uvw.

Applying the above abc-theorem for polynomials in a field of characteristic p we find that either $au/g, bv/g$ and cw/g are each a p^{th} power or the degrees of $au/g, bv/g$ and cw/g are all less than the number of distinct roots of $(au/g)(bv/g)(cw/g)$, which is $\le \deg(uvw/g) + \#\left\{\text{of distinct roots of } abc/g^2\right\}$, and the result follows. □

Corollary 8.3. *Let K be a field of characteristic p. If $a, b, c \in K[t]$ and have no common factor, and $m \in K[t]$ with $a \equiv b \equiv c \mod m$, where a and m have no common factor, then*

$$\max\{\deg a, \deg b, \deg c\} \ge \min\left\{\frac{1}{2}(\deg m + p), 2\deg m - \#\{\text{distinct roots of } abc/g^2\}\right\}$$

where $g = (a, b)(a, c)(b, c)$.

Proof. Let $U = (b - c)/m$, $V = (c - a)/m$ and $W = (b - a)/m$. Let $h = (U, V, W)$ and $U = hu, V = hv, W = hw$, so that $au + bv = cw$. Let $G = (au, bv, cw)$. Note that $(a, b, c) = (u, v, w) = 1$ by definition. Also (u, v, c) divides $(b - c, c - a, c) = (b, a, c) = 1$, so $(u, v, c) = 1$ and similarly $(u, w, b) = (v, w, a) = 1$. Therefore the hypothesis of Lemma 8.2 holds with $g = G$. Note that, by definition,

$$\max\{\deg u, \deg v, \deg w\} \le \max\{\deg a, \deg b, \deg c\} - \deg(hm). \tag{8.2}$$

If $au/g, bv/g$ and cw/g are not all p^{th} powers (or all belong to K), then by substituting (8.2) into Lemma 8.2, we get the second lower bound here. If $au/g, bv/g$ and cw/g are all p^{th} powers but are not all in K, then

$$p \le \max\{\deg au, \deg bv, \deg cw\} \le 2\max\{\deg a, \deg b, \deg c\} - \deg(m)$$

by (8.2). □

Suppose that p is prime and $h(x)$ is irreducible mod p where $\deg h(x) \le p/3$. Suppose that V is a set of linear polynomials. Consider the set $\langle V\rangle$ of polynomials of degree $< 2\deg h - |V|$, which are products of elements of V. By Corollary 8.3 (with $K = \mathbb{F}_p, m = h$ and $a, b, c \in \langle V\rangle$) we deduce that there are at least $\frac{1}{2}|\langle V\rangle|$ distinct polynomials in $\langle V\rangle$ mod $h(x)$. This allows us to improve the lower bound on G that we achieved in section 4 and may be combined with the ideas in section 8.1.

8.4. Fermat's Last Theorem, again. Fouvry's result, mentioned in section 3b.4, was inspired by its application to the first case of Fermat's Last Theorem, that is, integer solutions x, y, z to $x^p + y^p = z^p$ where p does not divide xyz. In 1910 Wieferich proved that if this has a solution, then $2^{p-1} \equiv 1 \pmod{p^2}$, which seems to be true very rarely. It turns out that knowledge about such congruences can be used to reduce the value of r above.

Lemma 8.4. *Let ℓ be a given prime and let $L = \ell$ if ℓ is odd, $L = 4$ if $\ell = 2$. Let k be the order of n mod L. Suppose that ℓ^e is the highest power of ℓ dividing $n^k - 1$. Then $k\ell^j$ is the order of n mod ℓ^{e+j}.*

Proof. If $n^K - 1 = \ell^E s$ where ℓ does not divide integer s, and $\ell^E > 2$, then, by the binomial theorem, $n^{K\ell} = (1 + \ell^E s)^\ell \equiv 1 + \ell^{E+1}s \pmod{\ell^{E+2}}$. The result follows from a suitable induction hypothesis. □

Thus if $n \equiv \pm 3 \pmod 8$, then n has order 2^{i-2} mod 2^i, for all $i \ge 3$, so n has order $> \log^2 n$ for some $r < 8\log^2 n$, taking r to be an appropriate power of two. More generally if $n \not\equiv \pm 1 \pmod{2^s}$ for any $s \ge 2$, then n mod r has order $> \log^2 n$ for some $r < 2^s \log^2 n$. This argument may be extended to odd primes ℓ:

Corollary 8.5. *If $n^{\ell-1} \not\equiv 1 \pmod{\ell^2}$ for odd prime ℓ, then there exists $r \le \ell^2 \log^2 n$ for which n mod r has order $> \log^2 n$.*

Proof. We may assume ℓ is odd and so, in Lemma 8.4, we have $e = 1$. Select j as small as possible so that $k\ell^j > \log^2 n$, and thus $\ell^j \le k\ell^j \le \ell \log^2 n$. Then $r := \ell^{j+1} \le \ell^2 \log^2 n$. □

8.5. Undergraduate theses. Manindra Agrawal had worked for several years with students trying to find a criteria for primality of this sort. In April 2001, Pashant Pandey and Rajat Bhattacharjee's bachelor's thesis "Primality Testing" studied what would happen if

$$(x-1)^n \equiv x^n - 1 \mod (n, x^r - 1) \tag{8.3}$$

holds for odd composite n and odd prime r. Their hope was that if this happens, then $n \equiv -1, 0$ or $1 \pmod r$. However such a neat conclusion seems unlikely, even for $r = 5$: To prove that there are infinitely many Carmichael numbers, one constructs integers n with prime factors that have certain extraordinary divisibility properties. In a similar vein Lenstra and Pomerance conjecture that there are

infinitely many squarefree integers n with $4k+1$ prime factors (with $k \ge 1$) such that every prime p which divides n is $\equiv 3 \mod 80$ and has $p-1$ divides $n-1$, and $p+1$ divides $n+1$. Given such an n, note that $p^2 \equiv 1$ or $-1 \mod 5$, and so $(x-1)^{p^2} \equiv x^{p^2}-1 \equiv x-1$ or $-x^{-1}(x-1) \mod (p, x^5-1)$, respectively; either way $(x-1)^{10(p^2-1)} \equiv 1 \mod (p, x^5-1)$. But $n \equiv p \mod 10(p^2-1)$ and so $(x-1)^n \equiv (x-1)^p \equiv x^p - 1 \equiv x^n - 1 \mod (p, x^5-1)$, and thus (8.3) holds with $r=5$ by the Chinese Remainder Theorem.

The following summer, Kayal and Saxena noted that (8.3) implies some of the more classical criteria:

Proposition 8.6. *Assume that* $(x-1)^n \equiv x^n - 1 \mod (m, x^r-1)$ *where* m, n, r *are positive integers with* n *odd and* $(r,n)=1$. *Then* $r^n \equiv r \pmod m$.

Proof. Let $\zeta = \mathrm{e}^{2i\pi/r}$ so that $\zeta^r = 1$ but $\zeta \ne 1$. Now $\prod_{j=1}^{r-1}(t-\zeta^j) = (t^r-1)/(t-1)$ $= 1 + t + t^2 + \cdots + t^{r-1}$, so taking $t=1$ gives $\prod_{j=1}^{r-1}(1-\zeta^j) = r$. Taking $x = \zeta^j$ in the hypothesis implies that $(1-\zeta^j)^n = 1 - \zeta^{jn} \pmod m$. Multiplying these equations together for $j = 1, 2, \dots, r-1$ we obtain the result by noting that $\{jn \pmod r : \ 1 \le j \le r-1\} = \{j \pmod r : \ 1 \le j \le r-1\}$ since $(n,r)=1$. □

They also show that if $(r,m)=1$, then $r^{(n-1)/2}$ must be -1 or $1 \pmod m$, that is a square root[39] of 1 mod m. For $r=5$ they determine the value of $F_{(n-1)/2} \pmod n$ where F_k is the kth Fibonacci number. Moreover that if (8.3) holds for each $r \le 2(\ln n)^2$, then n is squarefree. There may yet be much more of interest to deduce from (8.3).

Taking $x=1$ in (4.1) we obtain that $(a+1)^n \equiv a+1 \mod n$. By taking $x = \zeta^j$ in (4.1) as in the proof of Proposition 8.6, we deduce $(\zeta^j + a)^n \equiv \zeta^{jn} + a \pmod n$. Now if d divides r, then $\Phi_d(x) = \prod (x - \zeta^j)$ where the product is over those $j \pmod r$ for which $(j,r) = r/d$; so multiplying the above congruence over such j we obtain $\Phi_d(-a)^n \equiv \Phi_d(-a) \pmod n$ assuming that n is odd and $(n,r)=1$. Thus we get many congruences of the form (2.3) and even an improvement of Proposition 8.6, since we get $p^n \equiv p \pmod n$ for each prime p dividing r, as $\Phi_p(-1) = p$.

8.6. An equivalent formulation: Recurrence sequences. It seems strange that a primality test should be formulated in terms of polynomials! Inspired by the results mentioned in section 8.5, we now show how to reformulate the AKS test entirely in terms of integers.

Lemma 8.7. *Suppose that* $(n,r)=1$. *We have* $(x+a)^n \equiv x^n + a \mod (n, x^r-1)$ *if and only if* $(\zeta + a)^n \equiv \zeta^n + a \mod n$ *for all* ζ *satisfying* $\zeta^r = 1$.

Proof. If $(x+a)^n \equiv x^n + a \mod (n, x^r-1)$, then we get $(\zeta+a)^n \equiv \zeta^n + a \mod n$ for each ζ satisfying $\zeta^r = 1$ simply by substituting $x = \zeta$. If $(\zeta + a)^n \equiv \zeta^n + a \mod n$, then we may write $(x+a)^n \equiv x^n + a \mod (n, x-\zeta)$. The ideals $(n, x-\zeta)$ are coprime (for the gcd of any two contains $(n, \zeta - \zeta')$ which divides $(n,r)=1$). The result follows by the Chinese Remainder Theorem in $\mathbb{Z}[\zeta, x]$. □

Thus the criteria is reformulated in terms of algebraic integers in a number field, but we desire criteria in terms of rational integers. To do this we introduce linear

[39] In fact, they determine which square root of 1 in terms of the Legendre-Jacobi symbol, one of many relevant things that we have decided not to discuss in this article.

recurrence sequences (the best known example being the Fibonacci numbers); we give, without proof, the properties we need in the following paragraph.

Suppose that $\prod_{j=1}^{\ell}(x-\alpha_j) = x^\ell - \sum_{i=0}^{\ell-1} u_i x^i \in \mathbb{Z}[x]$ has no repeated roots. Let $c(t) \in \mathbb{Q}(t)$. Define $f_k = \sum_{i=1}^{\ell} c_i \alpha_i^k$ for each $k \ge 0$, where $c_j = c(\alpha_j)$. An alternate way to define f_k is by the values for $k = 0, 1, 2, \dots, \ell-1$ and thereafter by $f_{m+\ell} = \sum_{i=0}^{\ell-1} u_i f_{m+i}$ for all $m \ge 0$. The Fibonacci numbers make up the example where our polynomial is $x^2 - x - 1$ and $c(t) = t/(t^2+1)$, with $F_0 = 0,\ F_1 = 1$.

We claim that if we know $\alpha_1, \alpha_2, \cdots, \alpha_\ell$ and $f_0, \dots, f_{\ell-1}$, then we can recover $c_1, \cdots, c_\ell$. For this information gives us the ℓ linear equations $\sum_{i=1}^{\ell} \alpha_i^j c_i = f_j$ for $j = 0, 1, \cdots, \ell-1$ in the ℓ unknowns c_i, and the matrix of coefficients is the Vandermonde $\{\alpha_i^j\}_{1\le i\le \ell,\ 0\le j\le \ell-1}$, which has determinant $\pm\prod_{1\le i<j\le \ell}(\alpha_j - \alpha_i)$, and this is non-zero since the α_i are distinct. In fact this argument works mod n provided the determinant is coprime to n.

Take $\ell = r$ and $\alpha_i = \zeta^i + a$ for each i where $\zeta = \mathrm{e}^{2i\pi/r}$ is a primitive rth root of unity. The Vandermonde determinant above is then a power of r (by the proof of Proposition 8.6) and so coprime to n. Choose $c(t)$ so that $(c(\alpha_i), n) = 1$ for all i, and define $C(t) = c(t)((t+a)^n - (t^n + a))$. If $(\zeta^i + a)^n \equiv \zeta^{in} + a \mod n$ for all i, then $F_j \equiv 0 \mod n$ for all j, evidently. Moreover if $F_j \equiv 0 \mod n$ for $0 \le j \le r-1$, then we can deduce, as above, that $C(\zeta^i) \equiv 0 \mod n$ for all i, and therefore $(\zeta^i + a)^n \equiv \zeta^{in} + a \mod n$. Now $F_j = f_{n+j} - af_j - g_j$ where $g_j := \sum_{i=1}^{r} c_i \zeta^{in}(\zeta^i + a)^j$. We have enormous freedom in choosing $c(t)$; we select $c(t) = (t-a)/r$. Given that $(1/r)\sum_{0\le i\le r-1} \zeta^{ij} = 0$ unless r divides j, in which case it equals 1, we deduce that $f_j = 0$ for $0 \le j \le r-2$ and $f_{r-1} = 1$; whereas $g_j = \binom{j}{k} a^{j-k}$ where k is the least non-negative residue of $-(n+1) \mod r$, and we define $\binom{j}{k} = 0$ if $j < k$. Thus we have proved the following:

Proposition 8.8. *Suppose that $(n, r) = 1$. Define the sequence $\{f_m\}_{m\ge 0}$ of integers by $f_0 = f_1 = \cdots = f_{r-2} = 0,\ f_{r-1} = 1$ and*

$$f_{m+r} = \sum_{j=1}^{r-1} \binom{r}{j} (-1)^j a^{r-j} f_{m+j} + (a^r + 1) f_m \quad \text{for all } m \ge 0.$$

Then $(x+a)^n \equiv x^n + a \mod (n, x^r - 1)$ if and only if

$$f_{n+j} \equiv \binom{j}{k} a^{j-k} + af_j \mod n \quad \text{for } 0 \le j \le r-1.$$

In the special case $a = -1$ we can work with a linear recurrence of order $\ell = (r-1)/2$, because $(1-\zeta^i)^n \equiv 1 - \zeta^{in} \mod n$ is trivial for $i = 0$ and since it gives $(1-\zeta^{-i})^n \equiv 1 - \zeta^{-in} \mod n$ by multiplying through by $-\zeta^{-in}$. So take $\alpha_i = (1-\zeta^{2i})(1-\zeta^{-2i})$ for $1 \le i \le \ell$, with $c(\alpha_i) = \zeta^i + \zeta^{-i}$ (which can be attained by an appropriate polynomial $c(t)$). We define $C(t)$ so that $C(\alpha_i) = c(\alpha_i)\gamma(\alpha_i)$ where

$$\begin{aligned}\gamma(\alpha_i) &= \alpha_i^{(n+1)/2} + (-1)^{(n-1)/2}(\zeta^i - \zeta^{-i})(\zeta^{in} - \zeta^{-in}) \\ &= (-1)^{(n-1)/2}\zeta^{-ni}(\zeta^i - \zeta^{-i})((1-\zeta^{2i})^n - 1 + \zeta^{2in}).\end{aligned}$$

Working through a similar argument to that above we then obtain the following.

Define

$$g_r(x) := \prod_{1\le i\le (r-1)/2} (x - (1-\zeta^{2i})(1-\zeta^{-2i})) = x^{(r-1)/2} - \sum_{0\le i\le (r-3)/2} g_{r,i} x^i.$$

Note that $g_r(2-y^2-y^{-2}) = (-1)^{(r-1)/2}(y^r-y^{-r})/(y-y^{-1})$.

Proposition 8.9. *Suppose that $(n,r)=1$ with r an odd prime, and let k be the least residue, in absolute value, of $n/2 \mod r$. Define the sequence $\{f_m\}_{m\geq 0}$ of integers by $f_0=1,\ f_1=\cdots=f_{(r-3)/2}=0$ and*

$$f_{m+(r-1)/2} = \sum_{j=0}^{(r-3)/2} g_{r,j} f_{m+j} \quad \text{for all } m \geq 0.$$

Then $(x-1)^n \equiv x^n - 1 \mod (n, x^r-1)$ if and only if

$$f_{(n+1)/2+j} \equiv (-1)^{\frac{n-1}{2}+k}\ r\frac{k}{j+1}\binom{2j+2}{j+k+1} \mod n \quad \text{for } 0 \leq j \leq (r-3)/2.$$

With different choices of $c(t)$ we will obtain different recurrence sequences with the same basic property.

Acknowledgements: I would like to thank Matt Baker, Dan Bernstein, Michael Guy, Eric Pine, François Dorais, Felipe Voloch, the referee and particularly Manindra Agrawal and Carl Pomerance for reading a draft version of this article and providing numerous suggestions. I am also indebted to Dan Bernstein's fine website `http://cr.yp.to/primetests.html`, which contains a wealth of information and ideas.

References

1. Leonard M. Adleman and Ming-Deh A. Huang, *Primality testing and abelian varieties over finite fields*, Lecture Notes in Mathematics, vol. 1512, Springer-Verlag, Berlin. MR1176511 (93g:11128)
2. Leonard M. Adleman, Carl Pomerance and Robert S. Rumely, *On distinguishing prime numbers from composite numbers*, Annals of Mathematics **117** (1983), 173–206. MR0683806 (84e:10008)
3. Manindra Agrawal, Neeraj Kayal and Nitin Saxena, *PRIMES is in P* (to appear).
4. W. R. Alford, Andrew Granville and Carl Pomerance, *There are infinitely many Carmichael numbers*, Annals of Mathematics **139** (1994), 703–722. MR1283874 (95k:11114)
5. Roger C. Baker and Glynn Harman, *The Brun-Titchmarsh Theorem on average*, Progr. Math. **138** (1996), 39-103. MR1399332 (97h:11096)
6. D. J. Bernstein, *Proving primality in essentially quartic random time* (to appear).
7. Pedro Berrizbeitia, *Sharpening "PRIMES is in P" for a large family of numbers* (to appear).
8. Dan Boneh, *Twenty years of attacks on the RSA Cryptosystem*, Notices Amer. Math. Soc. **46** (1999), 203-213. MR1673760
9. Richard Crandall and Carl Pomerance, *Prime numbers. A computational perspective*, Springer-Verlag, New York, 2001. MR1821158 (2002a:11007)
10. Whitfield Diffie and Martin E. Hellman, *New directions in cryptography*, IEEE Transactions on Information Theory **22** (1976), 644–654. MR0437208 (55:10141)
11. Etienne Fouvry, *Théorème de Brun-Titchmarsh; application au théorème de Fermat*, Invent. Math **79** (1985), 383–407. MR0778134 (86g:11052)
12. Dorian M. Goldfeld, *On the number of primes p for which $p+a$ has a large prime factor*, Mathematika **16** (1969), 23-27. MR0244176 (39:5493)
13. Andrew Granville and Thomas Tucker, *It's as easy as abc*, Notices Amer. Math. Soc. **49** (2002), 1224-1231. MR1930670 (2003f:11044)
14. Shafi Goldwasser and Joe Kilian, *Almost all primes can be quickly certified*, Proceedings of the 18th annual ACM symposium on theory of computing, Association for Computing Machinery, New York, 1986.
15. Donald E. Knuth, *The art of computer programming, volume 2: Seminumerical algorithms*, Addison-Wesley, Reading, 1969. MR0286318 (44:3531)

16. H.W. Lenstra, Jr., *Galois theory and primality testing*, Lecture Notes in Mathematics **1142** (1985), 169-189. MR0812498 (87g:11171)
17. H.W. Lenstra, Jr., and Carl Pomerance, *Primality testing with Gaussian periods* (to appear).
18. Yu. V. Matijasevich, *Hilbert's Tenth Problem*, MIT Press, Cambridge, MA, 1993. MR1244324 (94m:03002b)
19. Preda Mihailescu and Roberto Avanzi, *Efficient 'quasi - deterministic' primality test improving AKS* (to appear).
20. Gérald Tenenbaum and Michel Mendès France, *The Prime Numbers and Their Distribution*, Student Mathematical Library **6** (2000), Amer. Math. Soc. MR1756233 (2001j:11086)
21. Paulo Ribenboim, *The new book of prime number records*, Springer-Verlag, New York, 1995. MR1377060 (96k:11112)
22. José Felipe Voloch, *On some subgroups of the multiplicative group of finite rings* (to appear).

Département de Mathématiques et Statistique, Université de Montréal, CP 6128 succ. Centre-Ville, Montréal, QC H3C 3J7, Canada

E-mail address: andrew@dms.umontreal.ca