A Brief History of Factoring and Primality Testing B. C. (Before Computers)

RICHARD A. MOLLIN
University of Calgary
Calgary, Alberta
Canada T2N 1N4
ramollin@math.ucalgary.ca

Factoring and primality testing have become increasingly important in today's information based society, since they both have produced techniques used in the secure transmission of data. However, often lost in the modern-day shuffle of information are the contributions of the pioneers whose ideas ushered in the computer age and, as we shall see, some of whose ideas are still used today as the underpinnings of powerful algorithms for factoring and primality testing. We offer this brief history to help readers know more about these contributions and appreciate their significance.

Virtually everyone who has graduated from high school knows the definition of a prime number, namely a $p \in \mathbb{N} = \{1, 2, 3, 4, \ldots\}$ such that $p > 1$ and if $p = \ell m$ where $\ell, m \in \mathbb{N}$, then either $\ell = 1$ or $m = 1$. (If $n \in \mathbb{N}$ and $n > 1$ is *not* prime, then n is called *composite*.) Although we cannot be certain, the concept of primality probably arose with the ancient Greeks over two and one-half millennia ago. The first *recorded* definition of prime numbers was given by Euclid around 300 BCE in his *Elements*. However, there is some indirect evidence that the concept of primality might have been known far earlier, for instance, to Pythagoras and his followers.

The Greeks of antiquity used the term *arithmetic* to mean what today we would call *number theory*, namely the study of the properties of the natural numbers and the relationships between them. The Greeks reserved the word *logistics* for the study of ordinary computations using the standard operations of addition/subtraction and multiplication/division, which we now call arithmetic. The Pythagoreans introduced the term *mathematics*, which to them meant the study of arithmetic, astronomy, geometry, and music. This curriculum became known as the *quadrivium* in the Middle Ages.

Although we have enjoyed the notion of a prime for millennia, only very recently have we developed *efficient* tests for primality. This seemingly trivial task is in fact much more difficult than it appears.

A *primality test* is an algorithm (a methodology following a set of rules to achieve a goal), the steps of which verify that given some integer n, we may conclude "n is a prime number." A *primality proof* is a successful application of a primality test.

Such tests are typically called *true primality tests* to distinguish them from *probabilistic primality tests* (which can only conclude that "n is prime" up to a specified likelihood). We will not discuss such algorithms here (see [**9**] for these).

A concept used frequently in primality testing is the notion of a *sieve*. A "sieve" is a process to find numbers with particular characteristics (for instance primes) by searching among *all* integers up to a prescribed bound, and eliminating invalid candidates until only the desired numbers remain. Eratosthenes (ca. 284–204 BCE) proposed the first sieve for finding primes. The following example illustrates the *Sieve of Eratosthenes*.

EXAMPLE 1. *Suppose that we want to find all primes less than* 30. *First, we write down all natural numbers less than* 30 *and bigger than* 1. *The first uncrossed number,*

2, is a prime. We now cross out all numbers (bigger than 2*) that are multiples of* 2 *(and hence composite).*

$$\{2, 3, \not{4}, 5, \not{6}, 7, \not{8}, 9, \not{10}, 11, \not{12}, 13, \not{14}, 15, \not{16}, 17, \not{18}, 19, \not{20}, 21, \not{22},$$
$$23, \not{24}, 25, \not{26}, 27, \not{28}, 29, \not{30}\}.$$

The next uncrossed number, 3*, must be a prime, so we cross out all numbers (bigger than* 3*) that are (composite) multiples of* 3*.*

$$\{2, 3, 5, 7, \not{9}, 11, 13, \not{15}, 17, 19, \not{21}, 23, 25, \not{27}, 29\}.$$

Then 5 *is the next uncrossed number, so we conclude it is prime, and we cross out all numbers (bigger than* 5*) that are multiples of* 5*.*

$$\{2, 3, 5, 7, 11, 13, 17, 19, 23, \not{25}, 29\}.$$

(We need not check any primes bigger than 5 *since such primes are larger than* $\sqrt{30}$*. An historical description of this fact follows.)*

The set of primes less than 30 *is what remains:*

$$\{2, 3, 5, 7, 11, 13, 17, 19, 23, 29\}.$$

The Sieve of Eratosthenes represents the only known algorithm from antiquity that we would call a primality test, but it is highly inefficient and it could not come close to verifying some of the primes known today. The number $2^{6972593} - 1$, shown to be prime on June 1, 1999, has 2,098,960 decimal digits (see the discussion of Mersenne primes below). Using the Sieve of Eratosthenes to verify its primality would take longer than the life expectancy of our sun using the fastest computers known today. The modern techniques that yield such a spectacular primality proof as this one are based on the ideas of later pioneers, whose contributions we highlight in this article.

Arabs and Italians

Arabic scholars were primarily responsible for preserving much of the mathematics from antiquity, and they extended many ancient results. Indeed, it was said that Caliph al-Mamun (809–833) experienced a vision, which included a visit from Aristotle; after this epiphany, al-Mamun was driven to have all of the Greek classics translated into Arabic, including Euclid's *Elements*.

Under the caliphate of al-Mamun lived Mohammed ibn Musa al-Khowarizmi (Mohammed, son of Moses of Kharezm, now Khiva), who was one of those to whom Europe owes the introduction of the Hindu-Arabic number system. Around 825 CE, he completed a book on arithmetic, which was later translated into Latin in the twelfth century under the title *Algorithmi de numero Indorum*. This book is one of the best-known works which introduced to Europe the Hindu-Arabic number system. This may account for the widespread, although mistaken, belief that our numerals are Arabic in origin. Not long after Latin translations of al-Khowarizmi's book were available in Europe, readers began to attribute the new numerals to him, and began contracting his name, in connection with these numerals, to *algorism*, and ultimately to *algorithm*.

Al-Khowarizmi also wrote a book on algebra, *Hisab al-jabr wa'lmuqābala*. The word *algebra* is derived from *al-jabr* or *restoration*. The term referred to the operation of removing a quantity that is subtracted on one side of an equation and "restoring" it

on the other side as an added quantity. In the Spanish work *Don Quixote*, which came much later, the term *algebrist* is used for a *bone-setter* or *restorer*.

As we observed, Eratosthenes did not discuss the issue of when his algorithm would terminate. However, Ibn al-Banna (ca. 1258–1339) appears to have been the first to observe that, in order to find the primes less than n using the sieve of Eratosthenes, one can restrict attention to prime divisors less than $\sqrt{n}$.

Fibonacci The resurrection of mathematical interest in Europe during the thirteenth century is perhaps best exemplified by the work of Leonardo of Pisa (ca. 1170–1250), better known as Fibonacci. While living in North Africa, where his father served as consul, Fibonacci was tutored by an Arab scholar. Thus, Fibonacci was well-educated in the mathematics known to the Arabs. Fibonacci's first book, and certainly his best known, is *Liber Abaci* or *Book of Calculation* first published in 1202, which continued to promote the use of the Hindu-Arabic number system in Europe. However, only the second edition, published in 1228 has survived.

In this work, Fibonacci gave an algorithm to determine if n is prime by dividing n by natural numbers up to $\sqrt{n}$. This represents the first recorded instance of a *deterministic algorithm* for primality testing, where *deterministic* means that the algorithm always terminates with either a *yes* answer or a *no* answer. (A deterministic algorithm may also be viewed as an algorithm that follows the same sequence of operations each time it is executed with the same input. This is in contrast to *randomized algorithms* that make random decisions at certain points in the execution, so that the execution paths may differ each time the algorithm is invoked with the same input. See [**9**] for a discussion of some randomized algorithms, which we will not discuss here.)

Fibonacci also discussed the well-known class of *Fibonacci numbers*, $\{F_n\}$, defined by the sequence

$$F_1 = F_2 = 1, \quad F_n = F_{n-1} + F_{n-2} \quad (n \geq 3).$$

In addition to being one of Fibonacci's memorable accomplishments, this sequence later played a surprising role in primality testing, as we shall see.

Perfect numbers

Another distinguished set of numbers that had a deep influence on the development of primality testing was the set of perfect numbers. A *perfect number* is an integer $n \in \mathbb{N}$ equal to the sum of its *proper divisors* (those $m \in \mathbb{N}$ where $m \mid n$ but $m \neq n$). For example, 6 is a perfect number, since $6 = 3 + 2 + 1$. The Pythagoreans probably knew about perfect numbers; the idea is founded in mysticism, which was their venue. Perfect numbers appear in Euclid's *Elements*, so we know the concept had been around for some time. The number $2^{n-1}(2^n - 1)$ is perfect for $n = 2, 3, 5, 7$, and these are the first four perfect numbers: 6, 28, 496, and 8128.

The ancient Greeks attributed mystical properties to perfect numbers. St. Augustine (354–430 CE) is purported to have said that God created the earth in six days since the perfection of the work is signified by the perfect number 6. Also, the moon orbits the earth every twenty-eight days, and 28 is the second perfect number.

Pietro Antonio Cataldi (1548–1626) developed an algorithmic approach to primality testing, but is probably best known for his work on continued fractions. In particular, his work *Trattato del modo brevissimo di trovar la radice quadra delli numeri* published in 1613, represents a significant contribution to the development of continued fractions. His work on perfect numbers was also considerable. Among his thirty

books, he wrote on military applications of algebra, and even published an edition of Euclid's *Elements*. Cataldi proved that the fifth, sixth, and seventh perfect numbers are:

$$33550336 = 2^{12}(2^{13} - 1),$$
$$8589869056 = 2^{16}(2^{17} - 1),$$

and

$$137438691328 = 2^{18}(2^{19} - 1).$$

It is uncertain whether Cataldi was the first to discover these perfect numbers, but his are the first known proofs of these facts. Cataldi was also the first to observe that if $2^n - 1$ is prime then n must be prime. In fact, the following result was known to Cataldi, though proved by Fermat.

THEOREM 1. (PERFECT NUMBERS) *If* $2^n - 1$ *is prime, then* n *is prime and* $2^{n-1}(2^n - 1)$ *is perfect.*

Proof. Since $(2^m - 1) \mid (2^n - 1)$ whenever $m \mid n$, then n must be prime whenever $2^n - 1$ is prime. (Note that, in general, if $n = \ell m$, then for any $b \in \mathbb{N}$, $b^n - 1 = (b^m - 1)\sum_{j=1}^{\ell} b^{m(\ell-j)}$.)

Let S_1 be the sum of all divisors of 2^{n-1} and let S_2 be the sum of all the divisors of the prime $2^n - 1$. Then the sum S of all divisors of $2^{n-1}(2^n - 1)$ is given by:

$$S = \sum_{\ell \mid 2^{n-1}(2^n-1)} \ell = \sum_{\ell \mid 2^{n-1}, \ell' \mid (2^n-1)} \ell\ell' = \sum_{\ell \mid 2^{n-1}} \ell \sum_{\ell' \mid (2^n-1)} \ell' = S_1 S_2.$$

Also, $S_1 = \sum_{j=0}^{n-1} 2^j$, so as a geometric series, we know that

$$S_1 = 2^n - 1.$$

Finally, since $2^n - 1$ is prime, then $S_2 = 2^n$. Hence,

$$S = 2^n(2^n - 1),$$

so $2^{n-1}(2^n - 1)$ is perfect. ∎

Long after Cataldi, Euler showed that every even perfect number has the form given in Theorem 1. It is unknown whether there are any odd perfect numbers and the search for them has exceeded the bound 10^{300}. Moreover, if such a beast exists, then it is known that it must have at least twenty-nine (not necessarily distinct) prime factors (see Guy [**6**, B1, p. 44]).

The French enter the fray

Theorem 1 tells us that the search for even perfect numbers is essentially the search for primes of the form

$$M_p = 2^p - 1, \text{ where } p \text{ is prime.}$$

Such primes are called *Mersenne primes*, the largest known of which is given above (see: http://www.utm.edu/research/primes/largest.html). These are named after the mendicant monk, Marin Mersenne (1588–1648). Although Mersenne was not a for-

mally trained mathematician, he had great enthusiasm for number theory. Among his contributions were his multifarious communications with many of the outstanding scholars of the day, including Descartes, Fermat, Frénicle de Bessy, and Pascal. He also published *Cognitata Physica-Mathematica* in 1644 in which he claimed that of all the primes $p \leq 257$, the only Mersenne primes M_p that occur are when

$$p = 2, 3, 5, 7, 13, 17, 19, 31, 67, 127, 257.$$

It was not until the twentieth century that Mersenne's claims were completely checked. We now know that Mersenne made five mistakes. For example, M_p is *not* prime for $p = 67$ and $p = 257$, but M_p *is* prime for $p = 61$, $p = 89$, and $p = 107$. It is for this list, and the impact which it had, that these primes were named after him.

Pierre de Fermat (1607–1665) kept Mersenne informed of the progress that he, too, was making in number theory. In particular, Fermat informed him that he had proved

$$223 \mid (2^{37} - 1) = M_{37}.$$

Fermat was able to do this by using a series of results, that began with his well-known "little theorem."

THEOREM 2. (FERMAT'S LITTLE THEOREM) *If q is a prime not dividing $b \in \mathbb{N}$, then $q \mid (b^{q-1} - 1)$.*

Proof. The result is obvious if $q = 2$, so we assume that $q > 2$. We now use the Binomial Theorem. First, we establish that $q \mid \binom{q}{j}$ for any natural number $j < q$. Since $q > 2$ is prime, then neither j nor $q - j$ divides q for any j with $1 \leq j \leq q - 1$. Therefore, the integer

$$\binom{q}{j} = \frac{q!}{(q-j)!\, j!}$$

is a multiple of q. Now, the Binomial Theorem in conjunction with this fact tells us that

$$b^q = (b - 1 + 1)^q = \sum_{j=0}^{q} \binom{q}{j} (b-1)^{q-j} 1^j = (b-1)^q + 1 + qa_1$$

for some $a_1 \in \mathbb{N}$. Applying this same argument to $(b-1)^q$, we get

$$(b-1)^q = (b-2)^q + 1 + qa_2$$

for some $a_2 \in \mathbb{N}$. Continuing in this fashion, for each $(b-i)^q$ with $1 \leq i < b$, we ultimately get that

$$b^q = b + q \sum_{j=1}^{b} a_j.$$

Hence, $q \mid (b^q - b) = b(b^{q-1} - 1)$, but $q \nmid b$ so $q \mid (b^{q-1} - 1)$. ∎

Fermat was actually interested in a result slightly different from the "little theorem" stated above. It is trivial that the little theorem implies the following theorem:

THEOREM 3. *Let $b \in \mathbb{N}$ and q a prime such that q does not divide b. Then there exists an $n \in \mathbb{N}$ such that $n \mid (q-1)$ and $q \mid (b^{(q-1)/n} - 1)$.*

This is trivially implied by the little theorem by setting $n = 1$. However, cases where larger values of n occur were of special interest to Fermat, as we shall see. The following result shows that in some cases we may actually find them:

COROLLARY 1. *If $p > 2$ is prime, then any prime divisor q of $2^p - 1$ must be of the form $q = 2mp + 1$ for some $m \in \mathbb{N}$. Also, if m is the smallest natural number for which $q \mid (b^m - 1)$, then $q \mid (b^t - 1)$ whenever $m \mid t$.*

Here, the number "$2m$" takes the role of n in the statement of Theorem 2. In particular, not only may we assume that $n > 1$, but that it is even.

Proof. First we prove the second assertion, which follows from the fact that if $t = ms$ for some $s \in \mathbb{N}$, then $(b^t - 1) = (b^m - 1)\sum_{j=1}^{s} b^{m(s-j)}$.

Now we establish the first assertion. Let q be a prime dividing $2^p - 1$. Then by the "little theorem," $q|(2^{q-1} - 1)$.

CLAIM. $\gcd(2^p - 1, 2^{q-1} - 1) = 2^{\gcd(p,q-1)} - 1$.

Let $g = \gcd(p, q - 1)$ and $g_1 = \gcd(2^p - 1, 2^{q-1} - 1)$. Thus, by the second assertion $(2^g - 1)|g_1$. It remains to show that $g_1|(2^g - 1)$. By the Euclidean Algorithm there exist $x, y \in \mathbb{N}$ such that $g = xp - y(q - 1)$. Since $g_1|(2^p - 1)$, then $g_1|(2^{px} - 1)$ by the second assertion, and similarly $g_1|2^{(q-1)y} - 1$. Thus, g_1 divides

$$2^{px} - 2^{(q-1)y} = 2^{(q-1)y}(2^{px-(q-1)y} - 1) = 2^{(q-1)y}(2^g - 1).$$

However, since g_1 is odd, then $g_1|(2^g - 1)$. This proves the Claim.

Since $q|(2^{q-1} - 1)$ and $q|(2^p - 1)$, then $g > 1$. However, since p is prime, then we must have that $g = p$, so $p|(q - 1)$. In other words, there exists an $n \in \mathbb{N}$ such that $q - 1 = np$. Since $p, q > 2$, then $n = 2m$ for some $m \in \mathbb{N}$. This completes the proof. ∎

From these results Fermat sought a number $n > 1$ as in Corollary 1 to use in trial divisions to test Mersenne numbers for primality. He saw that Corollary 1 could be useful to detect possible primes q such that $q \mid (2^{37} - 1)$. For example, $q = 37n + 1$ may be tested for low, even values of n until we find, when $n = 6$, that 223 divides $(2^{37} - 1) = (2^{(q-1)/n} - 1)$. Fermat also discovered that

$$q = 47 \text{ divides } (2^{23} - 1) = (2^{(q-1)/2} - 1)$$

using this method. We note that it takes only two trial divisions using this method to prove that

$$q = 233 \text{ divides } (2^{29} - 1) = (2^{(233-1)/8} - 1).$$

The reason is that by Corollary 1 any prime divisor of $2^{29} - 1$ must be of the form $58m + 1$ so by testing for $m = 1, 2, 3, 4$ of which only two are prime, 59 and 233, we get the nontrivial prime divisor $q = 233$.

Some of Fermat's most famous results were found in his correspondence with an excellent amateur mathematician, Bernard Frénicle de Bessy (1605–1675). In Fermat's letter dated October 8, 1640, Fermat's Little Theorem, in certain special cases, makes its first recorded appearance. Frénicle de Bessy also corresponded with Descartes, Huygens, and Mersenne. He actually solved several problems posed by Fermat, and posed further problems himself.

After Fermat's earlier success with Mersenne primes, he suggested to Frénicle de Bessy that numbers of the form

$$2^{2^n}+1$$

should be prime. Today such numbers are called *Fermat numbers*, denoted by $\mathfrak{F}_n$. Fermat knew that $\mathfrak{F}_n$ for $n = 0, 1, 2, 3, 4$ were prime, called *Fermat primes*, but could not prove primality for $n = 5$. Today we know that $\mathfrak{F}_n$ is composite for $5 \le n \le 24$, and it is suspected that $\mathfrak{F}_n$ is composite for all $n > 24$ as well. (On July 25, 1999, F_{382447}, which has over 10^{10^5} decimal digits, was shown to be composite by John Cosgrave (see http://www.spd.dcu.ie/johnbcos/fermat.htm).)

Fermat's work also had consequences for factoring. We define a *factorization algorithm* as one that solves the problem of determining the complete factorization of an integer $n > 1$, as guaranteed by the Fundamental Theorem of Arithmetic. In other words, the algorithm should find distinct primes p_j and $a_j \in \mathbb{N}$ such that $n = \prod_{j=1}^{k} p_j^{a_j}$. We observe that it suffices for such algorithms to merely find $r, s \in \mathbb{N}$ such that $1 < r \le s < n$ with $n = rs$ (called *splitting* n), since we can then apply the algorithm to r and to s, thereby recursively splitting each composite number until a complete factorization is found. Furthermore, since *deciding* whether a given $n > 1$ is composite or prime is easier, in general, than factoring, one should always check first whether n is composite (primality test) before applying a factorization algorithm.

In 1643, Fermat developed a method for factoring that was based on a simple observation. If $n = rs$ is an odd natural number with $r < \sqrt{n}$, then

$$n = a^2 - b^2 \text{ where } a = (s+r)/2 \text{ and } b = (s-r)/2.$$

Hence, in order to find a factor of n, we look through the various quantities $a^2 - n$ as a ranges among the values $a = \lfloor\sqrt{n}\rfloor + 1, \lfloor\sqrt{n}\rfloor + 2, \ldots, (n-1)/2$ until we find a perfect square, which will play the role of b^2. (Here $\lfloor z \rfloor$ is the *greatest integer function* or *floor*, namely the greatest integer less than or equal to z.) Once the appropriate values of a and b have been determined, we may solve for the factors r and s. This is called the *difference of squares method* of factoring, and it has been rediscovered numerous times.

From Euler to Gauss

The Swiss mathematician, Leonhard Euler (1707–1783), became interested in Fermat's work in 1730. He found that

$$641 \mid \mathfrak{F}_5,$$

thereby refuting Fermat's conjecture. Euler's method was to generalize a result of Fermat:

THEOREM 4. (EULER'S RESULT ON FERMAT NUMBERS) *If* $\mathfrak{F}_n = 2^{2^n} + 1$, *then every prime divisor of* $\mathfrak{F}_n$ *is of the form* $2^{n+1}r + 1$ *for some* $r \in \mathbb{N}$.

Proof. Let p be a prime divisor of $\mathfrak{F}_n$. Suppose that $m \in \mathbb{N}$ is the smallest value such that $p \mid (2^m - 1)$, and set $2^m = pw + 1$ for some integer w. Then since $p \mid (2^{2^n} + 1)$ and $p \mid (2^{2^{n+1}} - 1)$, we must have $2^{n+1} \ge m > 2^n$. By the Division Algorithm, there must exist $k \in \mathbb{N}$ and a nonnegative integer $\ell < m$, such that $2^{n+1} = mk + \ell$. Since $m > 2^n$, then $k = 1$ must be true. Also, p divides

$$(2^{2^{n+1}} - 1) = 2^{m+\ell} - 1 = (2^m)2^\ell - 1 = (pw+1)2^\ell - 1,$$

so we have shown that $p \mid (2^\ell - 1)$. By the minimality of m, we must have $\ell = 0$. Hence, $m = 2^{n+1}$.

Now, by Theorem 2, $p \mid (2^{p-1} - 1)$, so $p - 1 \geq 2^{n+1}$. Again, by the Division Algorithm, there must exist $r \in \mathbb{N}$ and a nonnegative integer $\ell_1 < 2^{n+1}$ such that $p - 1 = 2^{n+1}r + \ell_1$. Since p divides

$$(2^{p-1} - 1) = 2^{2^{n+1}r+\ell_1} - 1 = (2^{2^{n+1}})^r 2^{\ell_1} - 1 = (pw + 1)^r 2^{\ell_1} - 1,$$

a quick application of the Binomial Theorem shows that this equals $(pv_1 + 1)2^{\ell_1} - 1$ for some integer v_1. This means that $p \mid (2^{\ell_1} - 1)$, forcing $\ell_1 = 0$ by the minimality of 2^{n+1}. Hence, $p = 2^{n+1}r + 1$. ∎

In particular, we know from Theorem 4 that all divisors of $\mathfrak{F}_5$ must be of the form $64k + 1$. Thus, Euler only needed five trial divisions to find the factor 641, namely for $k = 3, 4, 7, 9, 10$, since the values $64k + 1$ for $k = 2, 5, 8$ are divisible by 3, and those for $k = 1, 6$ are divisible by 5.

Euler also knew of the seven perfect numbers

$$2^{n-1}(2^n - 1) \text{ for } n = 2, 3, 5, 7, 13, 17, 19.$$

By 1771, he had determined that M_{31} is also prime (using a methodology we outline below), the largest known prime to that date, a record that held until 1851.

In 1830, a valuable technique for factoring any odd integer n was discovered by Adrien-Marie Legendre (1752–1833) using the theory of *quadratic residues*. This theory, studied since the time of Euler and greatly advanced by Gauss was applied by Legendre to develop a new sieve method. An integer c is called a *quadratic residue* modulo $n \in \mathbb{N}$ if there is an integer x such that

$$c \equiv x^2 (\mathrm{mod}\, n)$$

(meaning that $n \mid (c - x^2)$).

Suppose we wish to find prime divisors of an integer n. For different primes p, Legendre studied congruences of the form

$$x^2 \equiv \pm p (\mathrm{mod}\, n).$$

Suppose a solution to this congruence could be found. This would imply that $\pm p$ is a quadratic residue modulo all prime factors of n. This fact can be used to greatly reduce the search for prime divisors of n by only considering those primes q for which p is also a quadratic residue $(\mathrm{mod}\, q)$. For instance, suppose 2 is a quadratic residue $(\mathrm{mod}\, n)$. A result that follows from Fermat's Little Theorem states that 2 is a quadratic residue modulo a prime q if and only if $q \equiv \pm 1 (\mathrm{mod}\, 8)$. Thus, already we have halved the search for factors of n (by eliminating odd divisors whose remainders are $\pm 3 (\mathrm{mod}\, 8)$).

Legendre applied this method repeatedly for various primes p. This can be viewed as constructing a (quadratic) sieve by computing lots of residues modulo n, thereby eliminating potential prime divisors of n that sit in various linear sequences. He found that if you computed enough of them, then one could eliminate primes up to $\sqrt{n}$ as prime divisors and thus show n was prime.

Some results of Euler had actually anticipated Legendre's work. He considered two representations of n:

$$n = x^2 + ay^2 = z^2 + aw^2,$$

so

$$(xw)^2 \equiv (n - ay^2)w^2 \equiv nw^2 - ay^2w^2 \equiv -ay^2w^2 \equiv (z^2 - n)y^2 \equiv (zy)^2 (\mathrm{mod}\, n),$$

and we are back to a potential factor for n. The basic idea in the above, for a given $n \in \mathbb{N}$, is simply that if we can find integers x, y such that

$$x^2 \equiv y^2 (\mathrm{mod}\, n), \tag{1}$$

and $x \not\equiv \pm y (\mathrm{mod}\, n)$, then $\gcd(x - y, n)$ is a nontrivial factor of n. This idea is still exploited by numerous algorithms in current use: *Pollard's $p - 1$ algorithm, the continued fraction algorithm, the quadratic sieve*, and the powerful *number field sieve*. For a complete description of these methods and their applications to cryptography, see [**9**].

Legendre was only concerned with building the sieve on the prime factors of n, and so he was unable to *predict*, for a given prime p, a second residue to yield a square. In other words, if he found a solution to $x^2 \equiv py^2 (\mathrm{mod}\, n)$, he could not predict a different solution $w^2 \equiv pz^2 (\mathrm{mod}\, n)$. If he had been able to do this, then he would have been able to combine the two as

$$(xw)^2 \equiv (pzy)^2 (\mathrm{mod}\, n),$$

so if $xw \not\equiv \pm pzy (\mathrm{mod}\, n)$, then $\gcd(xw - pzy, n)$ would be a nontrivial factor of n, thereby putting us back in the situation given in (1).

The idea of trying to match the primes to create a square can be attributed to Maurice Borisovich Kraitchik (1882–1957). Kraitchik, in the early 1920s, reasoned that it might suffice to find a *multiple* of $n \in \mathbb{N}$ as a difference of squares. He chose a quadratic polynomial of the form $kn = ax^2 \pm by^2$ for some $k \in \mathbb{N}$. In its simplest form with $k = a = b = 1$, he would sieve over $x^2 - n$ for $x \geq \lfloor \sqrt{n} \rfloor$. This is the basic idea behind the quadratic sieve method mentioned above. Thus, what Kraitchik had done was to opt for "fast" generation of quadratic residues, and in so doing abandoned Legendre's Method (meaning that, generally, he did not have residues less than $2\sqrt{n}$), but gained control over finding of two distinct residues at a given prime to form a square (as described above), which Legendre was unable to do. Thus, Kraitchik could start at values bigger than $\sqrt{n}$ and sieve until "large" residues were found.

A version of Legendre's method for factoring was developed by one of the greatest mathematicians who ever lived, Carl Friederich Gauss (1777–1855), in his influential masterpiece *Disquisitiones Arithmeticae* [**5**]. Gauss recognized the importance of factoring [**5**, Art. 329, p. 396]: "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." Gauss also discussed another factoring method in [**5**], which may be described as follows.

Suppose that we want to factor $n \in \mathbb{N}$. We choose some $m \in \mathbb{N}$ such that $\gcd(m, n) = 1$. Suppose that $x = r, s \in \mathbb{N}$ are two solutions of

$$n \mid (x^2 - m). \tag{2}$$

Then, if $n \nmid (r \pm s)$, then $\gcd(r - s, n)$ must be a nontrivial divisor of n because $n \mid (r - s)(r + s)$, while $n \nmid (r - s)$ and $n \nmid (r + s)$.

Landry, Lucas, and Lehmer

Once we enter the nineteenth century, the work of several individuals stands out in the development of factoring and primality testing. Among these is C. G. Reuschle (1813–1875). Tables compiled by Reuschle included all known prime factors of $b^n - 1$ for

$$b = 2, 3, 5, 7, 11, \text{and } n \leq 42.$$

He also included partial factorizations of $2^n - 1$ for some values of $n \leq 156$, and as we have already seen, this information is useful for primality testing. In 1925, Cunningham and Woodall [**3**] published tables of factorizations of $b^n \pm 1$ for a small number of values $b \leq 12$, and some high powers of n. As a consequence, work on extending these tables has come to be known as the *Cunningham Project*. (Relatively recent work on the Cunningham Project and related problems is available [**1**].)

In the late 1860s a Parisian named Fortuné Landry worked on finding factors of M_n. He began by looking closely at Euler's proof of the primality of M_{31}, and tried to improve it. Euler had observed that any prime p dividing M_{31} must be of the form $p \equiv 1 \bmod 62$. Since primes dividing forms of the type $x^2 - 2y^2$ must satisfy $p \equiv \pm 1 (\bmod 8)$, Euler knew that if $p \mid M_{31}$, then $p \equiv 1, 63 (\bmod 248)$, given that $p \mid ((2^{16})^2 - 2)$. Since $\lfloor \sqrt{M_{31}} \rfloor = 46340$, Euler had to trial divide M_{31} by primes of the form $p = 248k + 1$ or $p = 248k + 63$, where $p \leq 46340$. Finding no such primes, he concluded that M_{31} must be prime. Landry extended Euler's ideas as follows: If a given $n \in \mathbb{N}$ is known to have only factors of the form $ax + b$ for some (known) integers a, b, and if $n = (ax_1 + b)(ax_2 + b)$ for some (unknown) integers x_1, x_2, he deduced that there exist integers q, h, r such that

$$x_1 x_2 = q - bh, \tag{3}$$

$$x_1 + x_2 = r + ah, \tag{4}$$

and

$$h = \frac{q - x_1(r - x_1)}{ax_1 + b}. \tag{5}$$

Landry also showed that if $h = gh' + k$ for some integers g, h', then there exists a bound B such that if $x_1 > B$, then $h' = 0$. Hence, his algorithm involved testing all possible values of $k = h$ to see whether Equations (3)–(4) have solutions when $x_1 > B$. If they do, then we have a factor of n. If they do not, then we test all possible values of $x_1 \leq B$ to see if Equation (5) has a solution. We either get a factor of n or show that n is prime. (Dickson [**4**, p. 371] mentions Landry's efforts in the case where $a = 6$ and $b = \pm 1$, for instance.)

Using these methods, Landry completely factored $2^n \pm 1$ for all $n \leq 64$ with four exceptions. He even found the largest known prime of the time, namely

$$(2^{53} + 1)/321 = 2805980762433.$$

Perhaps the most influential nineteenth-century individual in the area of primality testing was François Édouard Anatole Lucas (1842–1891). Lucas had interests in recreational mathematics, such as his invention of the well-known *Tower of Hanoi* problem. However, his serious interest was in number theory, especially Diophantine analysis. Although he spent only the years 1875–1878 on the problems of factoring and primality testing, his contribution was impressive. Some of the ideas developed by Lucas may be interpreted today as the beginnings of computer design. He studied Fibonacci numbers and by 1877 had completely factored the first sixty of them. This led him to develop results on the divisibility of Fibonacci numbers, and ultimately to a proof that M_{127} is prime (modulo a corrected proof of the theorem below). The significance of this feat is revealed by the fact that this number held the distinction of being the largest known prime for three-quarters of a century. A larger prime was not found until 1951 by Miller and Wheeler [**8**].

The influence that Lucas had on modern-day primality testing is well described in a recent book by Hugh Williams [**10**], devoted to a discussion of the work of Lucas and his influence on the history of primality testing.

To see how Lucas determined that M_{127} is prime, we state the following result that was known to Lucas, although it was not given a *valid* proof until 1913 by R. D. Carmichael [2].

THEOREM 5. *Suppose that F_k denotes the k^{th} Fibonacci number and $n \in \mathbb{N}$ is given. If $n \equiv \pm 3 (\mathrm{mod}\ 10)$ and $n \mid F_{n+1}$ but $n \nmid F_m$ for all divisors m of n with $1 \leq m \leq n$, then n is prime. Also, if $n \equiv \pm 1 (\mathrm{mod}\ 10)$ and $n \mid F_{n-1}$ but $n \nmid F_m$ for all divisors m of n with $1 \leq m \leq n-2$, then n is prime.*

Based upon this result, all Lucas had to establish was that $M_{127} \mid F_{2^{127}}$ but $M_{127} \nmid F_{2^n}$ for all natural numbers $n < 127$. He did this in 1876, using methods that led to a primality test, the last we include in our historical discussion.

In the 1930s a pioneering giant in the world of primality testing, Derrick Henry Lehmer (1905–1991), extended the ideas of Lucas to provide the following primality test. (A look at his collected works [7] is highly recommended.)

Lucas-Lehmer true primality test for Mersenne numbers The algorithm consists of the following steps performed on a Mersenne number $M_n = 2^n - 1$ with $n \geq 3$.

(1) Set $s_1 = 4$ and compute $s_j \equiv s_{j-1}^2 - 2 (\mathrm{mod}\ M_n)$ for $j = 1, 2, \ldots, n-1$.
(2) If $s_{n-1} \equiv 0 (\mathrm{mod}\ M_n)$, then conclude that M_n is prime. Otherwise, conclude that M_n is composite.

Lucas knew only that the test was sufficient for primality, and this only for certain restricted types of values of n. In 1930, Lehmer proved both that the condition is necessary and that the test holds for any $n \in \mathbb{N}$.

EXAMPLE 2. *Input $M_7 = 127$. Then we compute $\overline{s_j}$, the least nonnegative residue of s_j modulo M_7 as follows. $\overline{s_2} = 14$, $\overline{s_3} = 67$, $\overline{s_4} = 42$, $\overline{s_5} = 111$, and $\overline{s_6} = 0$. Thus, M_7 is prime by the Lucas-Lehmer Test.*

This celebrated test is a fine example of the efforts of the pioneers such as Lehmer whose work, it may reasonably be said, had a deep and lasting influence upon the development of *computational number theory*, an experimental science with its feet in both the mathematical and computer science camps. One aspect of computational number theory that has given it high profile is *cryptography*, the study of methods for sending messages in secret.

Our age is dominated by information, and the need for secrecy is paramount in industry, academe, and the military, not to mention our personal lives. As we send email messages and financial data, we hope they remain private. Factoring and primality testing play a dominant role in the development of modern cryptographic techniques. Though the ideas of the pioneers are ubiquitous in modern algorithms, credit for their work is often overlooked. We hope to have increased the readers interest, understanding, and appreciation for these ideas.

Acknowledgments. The author's research is supported by NSERC Canada grant # A8484. Thanks go to the two anonymous referees whose comments inspired a complete rewriting of the first draft of this article in order to make the article more accessible and focused. Also, thanks to Glenn Appleby for help with the final edition.

REFERENCES

1. J. Brillhart, D. H. Lehmer, J. L. Selfridge, B. Tuckerman, and S. S. Wagstaff Jr., Factorizations of $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$ up to High Powers, *Contemporary Math.* **22**, Amer. Math. Soc., Providence, R.I., Second Edition (1988).
2. R. D. Carmichael, On the numerical factors of the arithmetic forms $\alpha^n \pm \beta^n$, *Annals of Math.*, **15** (1913), 30–70.

3. A. J. C. Cunningham and H. J. Woodall, *Factorization of* $y^n \mp 1$, $y = 2, 3, 5, 6, 7, 10, 11, 12$ *Up to High Powers* (n), Hodgson, London, 1925.
4. L. Dickson, *Theory of Numbers*, Vol. I, Chelsea, New York, 1992.
5. C. F. Gauss, *Disquisitiones Arithmeticae* (English edition), Springer-Verlag, Berlin, 1985.
6. R. K. Guy, *Unsolved Problems in Number Theory*, Vol. **1**, Second Edition, Springer-Verlag, Berlin, 1994.
7. D. H. Lehmer, *Selected Papers of D. H. Lehmer*, Vol. I–III, D. McCarthy (Ed.), The Charles Babbage Research Centre, St. Pierre, Canada, 1981.
8. J. C. P. Miller, Large primes, *Eureka* **14** (1951), 10–11.
9. R. A. Mollin, *An Introduction to Cryptography*, Chapman and Hall/CRC Press, New York, 2001.
10. H. C. Williams, *Édouard Lucas and Primality Testing*, Canadian Mathematical Society Series of Monographs and Advanced Texts, Vol. **22**, Wiley-Interscience, New York, Toronto, 1998.

Doing Math

DONNA DAVIS
P.O. Box 23392
Billings, MT 59104

Au contraire, Three Dog Night, one
is not the loneliest number. Two,
however, is indeed the loneliest number
since the number one—is out of the running.
In a three-legged race, it helps to become one
with the two you're tied-to.

The Foursquare Gospel Church has truth
cornered. Five in the hive and before
you know it, honey, we're done for.
Six times six times six gets
you 200+ years
to plant apricots and pots of petunias.

Two calls me again—it's that double
helix, the doublecross, the double we
all are said to have somewhere in the world.
Mine was on a TV show once, *Jeopardy*, but
then again aren't we all in danger?

Seven's so lucky, we should prime the pump
with her before every bath and baptism. Eight
won't wait—ask any cat with only one more life left.
Nine is fine—faceted like a sparkle.
And ten lets you start again where

your number system's bass'ed
and cello'ed and violin'ed and viola'ed
and the way to heaven is charted for you.