MATHEMATICS OF COMPUTATION
VOLUME 61, NUMBER 203
JULY 1993, PAGES 319–349

# THE FACTORIZATION OF THE NINTH FERMAT NUMBER

A. K. LENSTRA, H. W. LENSTRA, JR., M. S. MANASSE, AND J. M. POLLARD

*Dedicated to the memory of D. H. Lehmer*

ABSTRACT. In this paper we exhibit the full prime factorization of the ninth Fermat number $F_9 = 2^{512} + 1$. It is the product of three prime factors that have 7, 49, and 99 decimal digits. We found the two largest prime factors by means of the number field sieve, which is a factoring algorithm that depends on arithmetic in an algebraic number field. In the present case, the number field used was $\mathbf{Q}(\sqrt[5]{2})$. The calculations were done on approximately 700 workstations scattered around the world, and in one of the final stages a supercomputer was used. The entire factorization took four months.

## INTRODUCTION

For a nonnegative integer $k$, the *kth Fermat number* $F_k$ is defined by $F_k = 2^{2^k} + 1$. The ninth Fermat number $F_9 = 2^{512} + 1$ has 155 decimal digits:

$$\begin{aligned} F_9 = {} & 13407\,807929\,942597\,099574\,024998\,205846\,127479\,365820\,592393 \\ & 377723\,561443\,721764\,030073\,546976\,801874\,298166\,903427\,690031 \\ & 858186\,486050\,853753\,882811\,946569\,946433\,649006\,084097. \end{aligned}$$

It is the product of three prime numbers:

$$F_9 = p_7 \cdot p_{49} \cdot p_{99},$$

where $p_7$, $p_{49}$, and $p_{99}$ have 7, 49, and 99 decimal digits:

$$\begin{aligned} p_7 &= 2424833, \\ p_{49} &= 7455602\,825647\,884208\,337395\,736200\,454918\,783366\,342657, \\ p_{99} &= 741\,640062\,627530\,801524\,787141\,901937\,474059\,940781\,097519 \\ &\quad\; 023905\,821316\,144415\,759504\,705008\,092818\,711693\,940737. \end{aligned}$$

In binary, $p_7$, $p_{49}$, and $p_{99}$ have 22, 163, and 329 digits:

Received by the editor March 4, 1991 and, in revised form, August 3, 1992.
1991 *Mathematics Subject Classification*. Primary 11Y05, 11Y40.
*Key words and phrases*. Fermat number, factoring algorithm.

$$\begin{aligned}
p_7 &= 1001\,010000\,000000\,000001\,,\\
p_{49} &= 1010001\,100111\,110000\,110010\,110001\,010011\,001111\,001101\\
&\phantom{=}\ 101100\,111111\,001101\,101001\,111101\,000010\,001111\,101010\,110010\\
&\phantom{=}\ 101101\,010111\,100000\,110001\,010011\,001001\,010101\,000010\,100000\\
&\phantom{=}\ 000001\,,\\
p_{99} &= 10101\,101100\,110110\,001111\,010110\,100000\,010011\,100101\,010000\\
&\phantom{=}\ 101110\,011110\,100011\,001010\,111000\,110001\,111001\,100101\,110011\\
&\phantom{=}\ 010011\,000110\,111110\,011000\,100110\,010101\,001011\,000101\,100110\\
&\phantom{=}\ 011110\,000110\,110010\,000110\,111011\,001010\,010110\,001100\,001011\\
&\phantom{=}\ 111111\,111001\,001000\,101010\,101001\,111010\,100011\,001001\,111010\\
&\phantom{=}\ 010100\,000000\,101101\,101010\,111001\,000100\,110001\,101101\,100000\\
&\phantom{=}\ 000001\,.
\end{aligned}$$

The binary representation of $F_9$ itself consists of 511 zeros surrounded by 2 ones.

In this paper we discuss several aspects of the factorization of the ninth Fermat number. Section 1 is devoted to Fermat numbers and their place in number theory and its history. In §2 we address the general problem of factoring integers, and we describe the basic technique that many modern factoring methods rely on. In §3 we return to the ninth Fermat number, and we explain why previous factoring attempts of $F_9$ failed. We factored the number by means of the *number field sieve.* This method depends on a few basic facts from algebraic number theory, which are reviewed in §4. Our account of the number field sieve, in §5, can be read as an introduction to the more complete descriptions that are found in [28] and [10]. The actual sieving forms the subject of §6. The final stage of the factorization of $F_9$, which involved the solution of a huge linear system, is recounted in §7.

## 1. Fermat numbers

Fermat numbers were first considered in 1640 by the French mathematician Pierre de Fermat (1601–1665), whose interest in the problem of factoring integers of the form $2^m \pm 1$ arose from their connection with "perfect", "amicable", and "submultiple" numbers [47; 48, Chapter II, §IV]. He remarked that a number of the form $2^m + 1$, where $m$ is a positive integer, can be prime only if $m$ is a power of 2, which makes $2^m + 1$ a Fermat number. A Fermat number that is prime is called a *Fermat prime.* Fermat repeatedly expressed his strong belief that all Fermat numbers were prime. Apparently, this belief was based on his observation that each prime divisor $p$ of $F_k$ must satisfy a strong condition, namely $p \equiv 1 \bmod 2^{k+1}$. In present-day language, one would formulate his proof of this as follows. If $2^{2^k} \equiv -1 \bmod p$, then $(2 \bmod p)$ has multiplicative order $2^{k+1}$, and so $2^{k+1}$ divides $p-1$, by Fermat's own "little" theorem, which also dates from 1640. It is not clear whether Fermat was aware of the stronger condition $p \equiv 1 \bmod 2^{k+2}$ for prime divisors $p$ of $F_k$, $k \geq 2$. To prove this, it suffices to replace $(2 \bmod p)$, in the argument above, by its square root $(2^{2^{k-2}}(2^{2^{k-1}} - 1) \bmod p)$, which has order $2^{k+2}$. (It is amusing to note that also $(F_{k-1} \bmod p)$ has order $2^{k+2}$, because its square is an odd power of $(2 \bmod p)$.) Incidentally, from the binary representations of the prime factors of $F_9$ we see that

$$\operatorname{ord}_2(p_7 - 1) = 16, \quad \operatorname{ord}_2(p_{49} - 1) = 11, \quad \operatorname{ord}_2(p_{99} - 1) = 11,$$

where $\operatorname{ord}_2$ counts the number of factors 2.

The first five Fermat numbers $F_0 = 3$, $F_1 = 5$, $F_2 = 17$, $F_3 = 257$, and $F_4 = 65537$ are indeed prime, but to this day these remain the only known Fermat primes. Nowadays it is considered more likely, on loose probabilistic grounds, that there are only finitely many Fermat primes. It may well be that $F_0$ through $F_4$ are the only ones. On similar grounds, it is considered likely that all Fermat numbers are squarefree, with perhaps finitely many exceptions.

As for $F_5$, Fermat knew that any prime divisor of $F_5$ must be among 193, 449, 577, 641, $769, \ldots$, which is the sequence of primes that are $1 \bmod 2^6$, with $F_3 = 257$ omitted (distinct Fermat numbers are clearly relatively prime). Thus it is difficult to understand how he missed the factor 641, which is only the fourth one to try; among those that are $1 \bmod 2^7$, it is the first! One is led to believe that Fermat did not seriously attempt to verify his conjecture numerically, or that he made a computational error if he did. The factor 641 of $F_5$ was found by Euler in 1732, who thereby refuted Fermat's belief [18]. The cofactor $F_5/641 = 6700417$ is also prime.

Gauss showed in 1801 that Fermat primes are of importance in elementary geometry: a regular $n$-gon can be constructed with ruler and compasses if and only if $n$ is the product of a power of 2 and a set of distinct Fermat primes [19].

Since the second half of the nineteenth century, many mathematicians have been intrigued by the problem of finding prime factors of Fermat numbers and, more generally, numbers of the form $2^m \pm 1$. Somewhat later, this interest was extended to the larger class of *Cunningham numbers* $b^m \pm 1$ (with $b$ small and $m$ large) [16, 7]. The best factoring algorithms were usually applied to these numbers, so that the progress made in the general area of factoring large integers was reflected in the factorization of Fermat and Cunningham numbers.

The effort required for the complete prime factorization of a Fermat number may be expected to be substantially larger than for the preceding one, since the latter has only half as many digits (rounded upwards) as the former. In several cases the factorization could be accomplished only by means of a newly invented method. In 1880, Landry factored $F_6$, but his method was never published (see [25; 17, Chapter XV, p. 377; 20; 50]). In 1970, Morrison and Brillhart found the factorization of $F_7$ with the continued fraction method [36]. Brent and the fourth author factored $F_8$ in 1980 by means of a modified version of Pollard's rho method [6]. In 1988, Brent used the elliptic curve method to factor $F_{11}$ (see [4, 5]). Most recently, $F_9$ was factored in 1990 by means of the *number field sieve.*

Unlike methods previously used, the number field sieve is far more effective on Fermat and Cunningham numbers than on general numbers. Factoring general numbers of the order of magnitude of $F_9$ with the number field sieve—or with any other known method—requires currently substantially more time and financial resources than were spent on $F_9$; and factoring general numbers of the order of magnitude of $10^{15}F_9$ is not yet practically feasible.

The fact that the number field sieve performs abnormally well on Fermat and Cunningham numbers implies that these numbers are losing their value as a yardstick to measure progress in factoring. One wonders which class of numbers will take their place. Good test numbers for factoring algorithms should meet several conditions. They should be defined a priori, to avoid the impression that the factored numbers were generated by multiplying known factors. They should be easy to compute. They should not have known arithmetic properties that might be exploited by a special factorization algorithm. For any size range, there should be enough test numbers so that one does not quickly run out, but few enough to spark competition for them. They should have some mathematical significance, so that factoring them is a respectable activity. The last condition is perhaps a controversial one; but do we want to factor numbers that are obtained from a pseudorandom number generator, or from the digits of $\pi$ (see [2, 44])? The values of the partition function [1] meet the conditions above reasonably well, although they appear to be too highly divisible by small primes. In addition, their factorization is financially attractive (see [42]). We offer them to future factorers as test numbers. Nonetheless, factoring Fermat numbers remains a challenging problem, and it is likely to exercise a special fascination for a long time to come.

In addition to the more or less general methods mentioned above, a very special method has been used to search for factors of Fermat numbers. It proceeds not by fixing $k$ and searching for numbers $p$ dividing $F_k$, but by fixing $p$ and searching for numbers $k$ with $F_k \equiv 0 \bmod p$. To do this, one first chooses a number $p = u \cdot 2^l + 1$, with $u$ odd and $l$ relatively large, that is free of small prime factors; one can do this by fixing one of $u$, $l$ and sieving over the other. Next one determines, by repeated squarings modulo $p$, the residue classes $(2^{2^k} \bmod p)$, $k = 2, 3, \ldots$. From what we proved above about prime factors of Fermat numbers it follows that if no value $k \le l - 2$ is found with $2^{2^k} \equiv -1 \bmod p$, then $p$ does not divide *any* $F_k$, $k \ge 2$; in this case $p$ is discarded. If a value of $k$ is found with $2^{2^k} \equiv -1 \bmod p$—which one expects, loosely, to happen with probability $1/u$, if $p$ is prime—then $p$ is a factor of $F_k$. The primality of $p$ is then usually automatic from knowledge that one may have about smaller prime factors of $F_k$ or, if $p$ is sufficiently small, from the fact that all its divisors are $1 \bmod 2^{k+2}$.

Many factors of Fermat numbers have been found by the method just sketched. In 1903, A. E. Western [15] found the prime factor $p_7 = 2424833 = 37 \cdot 2^{16} + 1$ of $F_9$. In 1984, Keller found the prime factor $5 \cdot 2^{23473} + 1$ of $F_{23471}$; the latter number is the largest Fermat number known to be composite.

If no factor of $F_k$ can be found, one can apply a primality test that is essentially due to Pepin [37]: for $k \ge 1$, the number $F_k$ is prime if and only if $3^{(F_k-1)/2} \equiv -1 \bmod F_k$. This congruence can be checked in time $O((\log F_k)^3)$, and in time $O((\log F_k)^{2+\varepsilon})$ (for any positive $\varepsilon$) if one uses fast multiplication techniques. One should not view Pepin's test as a polynomial-time algorithm, however. In fact, the input is $k$, and from $\log F_k \approx 2^k \log 2$ we see that the time that the test takes is a *doubly exponential* function of the length $(\log k)/\log 2$ of the input. Pepin's test has indeed been applied only for a very limited collection of values of $k$.

Known factors of $F_k$ can be investigated for primality by means of general primality tests. In this way, Brillhart [22, p. 110] found in 1967 that the number $F_9/2424833$, which has 148 decimal digits, is composite. In 1988, Brent and Morain found that $F_{11}$ divided by the product of four relatively small prime factors is a prime number of 564 decimal digits, thereby completing the prime factorization of $F_{11}$.

The many results on factors of Fermat numbers that have been obtained by the methods above, as well as bibliographic information, can be found in [17, Chapter XV; 16, 7, 41, 23]. For up-to-date information one should consult the current issues of *Mathematics of Computation*, as well as the updates to [7] that are regularly published by S. S. Wagstaff, Jr. We give a brief summary of the present state of knowledge.

The complete prime factorization of $F_k$ is known for $k \leq 9$, for $k = 11$, and for no other $k$. One or more prime factors of $F_k$ are known for all $k \leq 32$ except $k = 14, 20, 22, 24, 28$, and $31$, as well as for 76 larger values of $k$, the largest being $k = 23471$. For $k = 10, 12, 13, 15, 16, 17$, and $18$ the cofactor is known to be composite. No nontrivial factor is known of $F_{14}$ or $F_{20}$, but it is known that these numbers are composite. For $k = 22, 24, 28, 31$, and all except 76 values of $k > 32$, it is unknown whether $F_k$ is prime or composite.

The smallest Fermat number that has not been completely factored is $F_{10}$. Its known prime factors are

$$11131 \cdot 2^{12} + 1 = 45\,592577,$$
$$395937 \cdot 2^{14} + 1 = 6487\,031809.$$

The cofactor has 291 decimal digits. Unless it has a relatively small factor, it is not likely to be factored soon.

The factorization of Fermat numbers is of possible interest in the theory of finite fields. Let $m$ be a nonnegative integer, and let the field $K$ be obtained by $m$ successive quadratic extensions of the two-element field, so that $\# K = 2^{2^m}$; an elegant explicit description of $K$ was given by Conway [14, Chapter 6] and another by Wiedemann [49]. It is easy to see that the multiplicative group of $K$ is a direct sum of $m$ cyclic groups of orders $F_0, F_1, \ldots, F_{m-1}$. Therefore, knowledge of the prime factors of Fermat numbers is useful if one wishes to determine the multiplicative order of a given nonzero element of $K$, or if one searches for a primitive root of $K$.

## 2. Factoring integers

In this section, $n$ is an odd integer greater than 1. It should be thought of as an integer that we want to factor into primes. We denote by $\mathbf{Z}$ the ring of integers, by $\mathbf{Z}/n\mathbf{Z}$ the ring of integers modulo $n$, and by $(\mathbf{Z}/n\mathbf{Z})^*$ the group of units (i.e., invertible elements) of $\mathbf{Z}/n\mathbf{Z}$.

2.1. **Factoring with square roots of** 1. The subgroup $\{x \in \mathbf{Z}/n\mathbf{Z} : x^2 = 1\}$ of $(\mathbf{Z}/n\mathbf{Z})^*$ may be viewed as a vector space over the two-element field $\mathbf{F}_2 = \mathbf{Z}/2\mathbf{Z}$, the vector addition being given by multiplication. Many factoring algorithms depend on the elementary fact that the dimension of this vector space is equal to the number of distinct prime factors of $n$. In particular, if $n$ is not a power of a prime number, then there is an element $x \in \mathbf{Z}/n\mathbf{Z}$, $x \neq \pm 1$, such that $x^2 = 1$. Moreover, explicit knowledge of such an element $x$, say $x = (y \bmod n)$, leads to a nontrivial factorization of $n$. Namely, from $y^2 \equiv 1 \bmod n$, $y \not\equiv \pm 1 \bmod n$, it follows that $n$ divides the product of $y - 1$ and $y + 1$ without dividing the factors, so that $\gcd(y - 1, n)$ and $\gcd(y + 1, n)$ are nontrivial divisors of $n$. They are in fact complementary divisors, so that only one of the gcd's needs to be calculated; this can be done with Euclid's algorithm. We conclude that, to factor $n$, it suffices to find $x \in \mathbf{Z}/n\mathbf{Z}$ with $x^2 = 1$, $x \neq \pm 1$.

2.2. **Repeated prime factors.** The procedure just sketched will fail if $n$ is a prime power, so it is wise to rule out that possibility before attempting to factor $n$ in this way. To do this, one can begin by subjecting $n$ to a primality test, as in [27, §5]. If $n$ is prime, the factorization is finished. Suppose that $n$ is not prime. One still needs to check that $n$ is not a prime *power*. This check is often omitted, since in many cases it is considered highly unlikely that $n$ is a prime power if it is not prime; it may even be considered highly likely that $n$ is squarefree, that is, not divisible by the square of a prime number. For example, suppose that $n$ is the unfactored portion of some randomly drawn integer, and one is certain that it has no prime factor below a certain bound $B$. Then the probability for $n$ not to be squarefree is $O(1/(B \log B))$, in a sense that can be made precise, and the probability that $n$ is a proper power of a prime number is even smaller. A similar statement may be true if $n$ is the unfactored portion of a Cunningham number, since, to our knowledge, no such number has been found to be divisible by the square of a prime factor that was difficult to find. Whether other classes of test numbers that one may propose behave similarly remains to be seen; if the number $n$ to be factored is provided by a "friend", or by a colleague who does not yet have sufficient understanding of the arithmetical properties of the numbers that his computations produce, it may be unwise to ignore the possibility of repeated prime factors.

2.3. **Squarefreeness tests.** No squarefreeness tests for integers are known that are essentially faster than factoring (see [9, §7]). This is often contrasted with the case of polynomials in one variable over a field $K$, in which case it suffices to take the gcd with the derivative. This illustrates that for many algorithmic questions the well-known analogy between $\mathbf{Z}$ and $K[X]$ appears to break down. Note also that for many fields $K$, including finite fields and algebraic number fields, there exist excellent practical factoring algorithms for $K[X]$ (see [26]), which have no known analogue in $\mathbf{Z}$.

There do exist factoring methods that become a little faster if one wishes only to test squarefreeness; for example, if $n$ is not a square—which can easily be tested—then to determine whether or not $n$ is squarefree it suffices to do trial division up to $n^{1/3}$ instead of $n^{1/2}$.

There is also a factoring method that has great difficulties with numbers $n$ that are not squarefree. Suppose, for example, that $p$ is a large prime for which $p - 1$ and $p + 1$ both have a large prime factor, and that $n$ has exactly two factors $p$. The factoring method described in [43], which depends on the use of "random class groups", does not have a reasonable chance of finding *any* nontrivial factor of $n$, at least not within the time that is conjectured in [43] (see [32, §11]).

2.4. **Recognizing powers.** Ruling out that $n$ is a prime power is much easier than testing $n$ for squarefreeness. One way to proceed is by testing that $n$ is not a proper power. Namely, if $n = m^l$, where $m$, $l$ are integers and $l > 1$, then $m \geq 3$, $2 \leq l \leq [(\log n)/\log 3]$, and one may assume that $l$ is prime. Hence, the number of values to be considered for $l$ is quite small, and this number can be further reduced if a better lower bound for $m$ is known, such as a number $B$ as in §2.2. For each value of $l$, one can calculate an integer $m_0$ for which $|m_0 - n^{1/l}| < 1$, using Newton's method, and test whether $n = m_0^l$; this is the case if and only if $n$ is an $l$th power. One can often save time by calculating $m_0$ only if $n$ satisfies the conditions

$$n^{l-1} \equiv 1 \mod l^2 \qquad (\text{mod } 8 \text{ if } l = 2)$$

and

$$n^{(q-1)/l} \equiv 1 \mod q$$

for several small primes $q$ with $q \equiv 1 \mod l$. These are necessary conditions for a number $n$ that is free of small prime factors to be an $l$th power, if $l$ is prime.

2.5. **Ruling our prime powers.** There is a second, less well-known way to proceed, which tests only that $n$ is not a *prime* power. It assumes that one has already proved that $n$ is composite by means of Fermat's theorem, which states that $a^n \equiv a \bmod n$ for every integer $a$, if $n$ is prime. Hence, if an integer $a$ has been found for which $a^n \not\equiv a \bmod n$, then one is sure that $n$ is composite. If $n$ is a prime power, say $n = p^k$, then Fermat's theorem implies that $a^p \equiv a \bmod p$ and hence also that $a^n = a^{p^k} \equiv a \bmod p$; that is, $p$ divides $a^n - a$, so it also divides $\gcd(a^n - a, n)$. This suggests the following approach. Having found an integer $a$ for which $(a^n - a \bmod n)$ is nonzero, we calculate the gcd of that number with $n$. If the gcd is 1, we can conclude that $n$ is not a prime power. If the gcd is not 1, then the gcd is a nontrivial factor of $n$, which is usually more valuable than the information that $n$ is or is not a prime power.

Nowadays one often proves compositeness by using a variant of Fermat's theorem that depends on the splitting

$$a^n - a = a \cdot (a^u - 1) \cdot \prod_{i=0}^{t-1} (a^{u \cdot 2^i} + 1),$$

where $n - 1 = u \cdot 2^t$, with $u$ odd and $t = \text{ord}_2(n-1)$. Hence, if $n$ is prime, then for any integer $a$ one of the $t + 2$ factors on the right is divisible by $n$. This variant has the advantage that the converse is true in a strong sense: if $n$ is not prime, then *most* integers $a$ have the property that *none* of the factors on the right is $0 \bmod n$ (see [40] for a precise statement and proof); such integers $a$ are called *witnesses* to the compositeness of $n$. Currently, if one is sure that the number $n$ to be factored is composite, it is usually because one has found such a witness. Just as above, a witness $a$ can be used to check that $n$ is in fact not a prime power: calculate $a^n - a \pmod n$, which one does most easily by first squaring the number $a^{u \cdot 2^{t-1}} \pmod n$ that was last calculated; if it is nonzero, one verifies as before that $\gcd(a^n - a, n) = 1$, and if it is zero then one of the $t + 2$ factors on the right has a nontrivial factor in common with $n$, which can readily be found. (In the latter case, $n$ is in fact not a prime power, since the odd parts of the $t + 2$ factors are pairwise relatively prime.)

As we mentioned in §1, the number $F_9/2424833$ was proved to be composite by Brillhart in 1967. We do not know whether he or anybody else proved that it is not a prime power until this fact became plain from its prime factorization. We did not, not because we thought it was not worth our time, but simply because we did not think of it. If it *had* been a prime power, our method would have failed completely, and we would have felt greatly embarrassed towards the many people who helped us in this project. One may believe that the risk that we were unconsciously taking was extremely small, but until the number was factored this was indeed nothing more than a belief. In any case, it would be wise to include, in the witness test described above, the few extra lines that prove that the number is not a prime power, and to explicitly publish this information about a number rather than just saying that it is composite.

2.6. **A general scheme.** For the rest of this section we assume that $n$, besides being odd and greater than 1, is not a prime power. We wish to factor $n$ into primes. As we have seen, each $x \in \mathbf{Z}/n\mathbf{Z}$ with $x^2 = 1$, $x \neq \pm 1$ gives rise to a nontrivial factor of $n$. In fact, it is not difficult to see that the full factorization of $n$ into powers of distinct prime numbers can be obtained from a set of generators of the $\mathbf{F}_2$-vector space $\{x \in \mathbf{Z}/n\mathbf{Z}: x^2 = 1\}$. (If we make this vector space into a *Boolean ring* with $x * y = (1+x+y-xy)/2$ as multiplication, then a set of ring generators also suffices.) The question is how to determine such a set of generators. Several algorithms have been proposed to do this, most of them following some refinement of the following scheme.

*Step* 1. *Selecting the factor base.* Select a collection of nonzero elements $a_p \in \mathbf{Z}/n\mathbf{Z}$, with $p$ ranging over some finite index set $\mathrm{P}$. *How* this selection takes place depends on the particular algorithm; it is usually not done randomly, but in such a way that Step 2 below can be performed in an efficient manner. The collection $(a_p)_{p \in \mathrm{P}}$ is called the *factor base.* We shall assume that all $a_p$ are units of $\mathbf{Z}/n\mathbf{Z}$. In practice, this is likely to be true, since if $n$ is difficult to factor, one does not expect one of its prime factors to show up in one of the $a_p$'s; one can verify the assumption, or find a nontrivial factor of $n$, by means of a gcd computation. Denote by $\mathbf{Z}^{\mathrm{P}}$ the additive abelian group consisting of all vectors $(v_p)_{p \in \mathrm{P}}$ with $v_p \in \mathbf{Z}$, and let $f: \mathbf{Z}^{\mathrm{P}} \to (\mathbf{Z}/n\mathbf{Z})^*$ be the group homomorphism (from an additively to a multiplicatively written group) that sends $(v_p)_{p \in \mathrm{P}}$ to $\prod_{p \in \mathrm{P}} a_p^{v_p}$. This map is surjective if and only if the elements $a_p$ generate $(\mathbf{Z}/n\mathbf{Z})^*$. For the choices of $a_p$ that are made in practice that is usually the case, although we are currently unable to prove this. (In general, hardly anything has been rigorously proved about practical factoring algorithms.)

*Step* 2. *Collecting relations.* Each element $v = (v_p)_{p \in \mathrm{P}}$ of the kernel of $f$ is a *relation* between the $a_p$, in the sense that $\prod_{p \in \mathrm{P}} a_p^{v_p} = 1$. In the second step, one looks for such relations by a method that depends on the algorithm. One stops as soon as the collection $V$ of relations that have been found has slightly more than $\#\mathrm{P}$ elements. One hopes that $V$ generates the kernel of $f$, although this is again typically beyond proof. Note that the kernel of $f$ is of finite index in $\mathbf{Z}^{\mathrm{P}}$, so that by a well-known theorem from algebra it is freely generated by $\#\mathrm{P}$ elements; therefore, the hope is not entirely unreasonable.

*Step* 3. *Finding dependencies.* For each $v \in V$, let $\overline{v} \in (\mathbf{Z}/2\mathbf{Z})^{\mathrm{P}} = \mathbf{F}_2^{\mathrm{P}}$ be the vector that one obtains from $v$ by reducing its coordinates modulo 2. Since $\# V > \#\mathrm{P}$, the vectors $\overline{v}$ are linearly dependent over $\mathbf{F}_2$. In Step 3, one finds explicit dependencies by solving a linear system. The matrix that describes the system tends to be *huge* and *sparse*, which implies that special methods can be applied (see [24]). Nevertheless, one usually employs ordinary Gaussian elimination. The size of the matrices may make it desirable to modify Gaussian elimination somewhat; see §7. Each dependency that is found can be written in the form $\sum_{v \in W} \overline{v} = 0$ for some subset $W \subset V$, and each such subset gives rise to a vector $w = (\sum_{v \in W} v)/2 \in \mathbf{Z}^{\mathrm{P}}$ for which $2 \cdot w$ belongs to the kernel of $f$. Each such $w$, in turn, gives rise to an element $x = f(w) \in (\mathbf{Z}/n\mathbf{Z})^*$ satisfying $x^2 = f(2 \cdot w) = 1$, and therefore possibly to a decomposition of $n$ into two nontrivial factors. If the factorization is trivial (because $x = \pm 1$), or, more generally, if the factors that are found are themselves not prime powers, then one repeats the same procedure starting from a different dependency between the vectors $\overline{v}$. Note that it is useless to use a dependency that is a linear combination of dependencies that have been used earlier. Also, if several factorizations of $n$ into two factors are obtained, they should be combined into one factorization of $n$ into several factors by a few gcd calculations. One stops when all factors are prime powers; if indeed $f$ is surjective and $V$ generates the kernel of $f$, this is guaranteed to happen before all dependencies between the $\overline{v}$ are exhausted.

2.7. **The rational sieve and smoothness.** A typical example is the *rational sieve.* In this factoring algorithm the factor base is selected to be

$$\mathrm{P} = \{p \colon p \text{ is prime}, p \le B\},$$
$$a_p = (p \bmod n) \qquad (p \in \mathrm{P}),$$

where $B$ is a suitably chosen bound. Collecting relations between the $a_p$ is done as follows. Using a sieve, one searches for positive integers $b$ with the property that both $b$ and $n + b$ are *B-smooth*, that is, have all their prime factors smaller than or equal to $B$. Replacing both sides in the congruence $b \equiv n + b \bmod n$ by their prime factorizations, we see that each such $b$ gives rise to a multiplicative relation between the $a_p$. The main merit of the resulting factoring algorithm—which is, essentially, the number field sieve, with the number field chosen to be the field of rational numbers—is that it illustrates the scheme above concisely. The rational sieve is not recommended for practical use, not because it is inefficient in itself, but because other methods are much faster.

The choice of the "smoothness bound" $B$ is very important: if $B$, and hence $\#\mathrm{P}$, is chosen too large, one needs to generate *many* relations, and one may end up with a matrix that is larger than one can handle in Step 3. On the other hand, if $B$ is chosen too small, then not enough integers $b$ will be found for which both $b$ and $n + b$ are $B$-smooth. The same remarks apply to the other algorithms that satisfy our schematic description.

In practice, the optimal value for $B$ is determined empirically. In theory, one makes use of results that have been proved about the function $\psi$ defined by

$$\psi(x, y) = \#\{m \in \mathbf{Z} \colon 0 < m \le x,\ \ m \text{ is } y\text{-smooth}\};$$

so $\psi(x, y)/[x]$ is equal to the probability that a random positive integer $\le x$ has all its prime factors $\le y$. Brief summaries of these results, which are adequate for the purposes of factoring, can be found in [38, §2; 27, §2.A and (3.16)].

Not surprisingly, one finds that both from a practical and a theoretical point of view the optimal choice of the smoothness bound and the performance of the factoring algorithm depend mainly on the size of the numbers that one wishes to be smooth. The smaller these numbers are, the more likely are they to be smooth, the smaller the smoothness bound that can be taken, and the faster the algorithm. For a fuller discussion of this we refer to [10, §10].

In the rational sieve, one wishes the numbers $b(n+b)$ to be smooth, and since $b$ is small, these numbers may be expected to be $n^{1+o(1)}$ (for $n \to \infty$). The theory of the $\psi$-function then suggests that the optimal choice for $B$ is

$$B = \exp((\sqrt{2}/2 + o(1))(\log n)^{1/2}(\log\log n)^{1/2}) \qquad (n \to \infty),$$

and that the running time of the entire algorithm is

$$\exp((\sqrt{2} + o(1))(\log n)^{1/2}(\log\log n)^{1/2}) \qquad (n \to \infty).$$

(This assumes that the dependencies in Step 3 are found by a method that is faster than Gaussian elimination.)

2.8. **Other factoring algorithms.** A big improvement is brought about by the *continued fraction* method [36] and by the *quadratic sieve* algorithm [38, 45], which belong to the same family. In these algorithms the numbers that one wishes to be smooth are only $n^{1/2+o(1)}$. This leads to the conjectured running time

$$\exp((1 + o(1))(\log n)^{1/2}(\log\log n)^{1/2}) \qquad (n \to \infty),$$

the smoothness bound being approximately the square root of this. Although the quadratic sieve never had the honor of factoring a Fermat number, it is still considered to be the best practical algorithm for factoring numbers without small prime factors.

In the *number field sieve* [28, 10], the numbers that one wishes to be smooth are $n^{o(1)}$, or more precisely

$$\exp(O((\log n)^{2/3}(\log\log n)^{1/3})),$$

and both the smoothness bound and the running time are conjecturally of the form

$$\exp(O((\log n)^{1/3}(\log\log n)^{2/3})).$$

This leads one to expect that the number field sieve is asymptotically the fastest factoring algorithm that is known. It remains to be tested whether for numbers in realistic ranges the number field sieve beats the quadratic sieve, if one does not restrict to special classes of numbers like Fermat numbers and Cunningham numbers.

It is to be noted that the running time estimates that we just gave depend only on the number to be factored, and not on the size of the factor that is found. Thus, the quadratic sieve algorithm needs just as much time to find a small prime factor as to find a large one. There exist other factoring algorithms, not satisfying our schematic description, that are especially good at finding small prime factors of a number. These include trial division, Pollard's $p \pm 1$ method, Pollard's rho method, and the elliptic curve method (see [27, 31, 3, 34]).

## 3. The ninth Fermat number

As we mentioned in §1, A. E. Western discovered in 1903 the factor 2424833 of $F_9$, and Brillhart proved in 1967 that $F_9/2424833$ is composite. In this section we let $n$ be the number $F_9/2424833$, which has 148 decimal digits:

$$\begin{aligned} n = {} & 5529\,373746\,539492\,451469\,451709\,955220\,061537\,996975\,706118 \\ & 061624\,681552\,800446\,063738\,635599\,565773\,930892\,108210\,210778 \\ & 168305\,399196\,915314\,944498\,011438\,291393\,118209\,. \end{aligned}$$

We review the attempts that have been made to factor $n$.

We do not believe that the possibility of factoring $n$ by means of the quadratic sieve algorithm was ever seriously considered. It would not have been beyond human resources, but it would have presented considerable financial and organizational difficulties.

Several factoring algorithms that are good at finding small prime factors had been applied to $n$. Richard Brent tried Pollard's $p \pm 1$ method and a modified version of Pollard's rho method (see [27]), both without success. He estimates that if there had been a prime factor less than $10^{20}$, it would probably have been found by the rho method. The failure of the rho method is simply due to the size of the least prime factor $p_{49}$ of $n$. The $p \pm 1$ method would have been successful if at least one of the four numbers $p_{49} \pm 1$, $p_{99} \pm 1$ had been built from small prime factors. The failure of this method is explained by the factorizations

$$\begin{aligned} p_{49} - 1 &= 2^{11} \cdot 19 \cdot 47 \cdot 82\,488781 \cdot 1143\,290228\,161321 \\ &\quad \cdot 43\,226490\,359557\,706629\,, \\ p_{49} + 1 &= 2 \cdot 3 \cdot 167\,982422\,287027 \\ &\quad \cdot 7397\,205338\,652138\,126604\,651761\,133609\,, \\ p_{99} - 1 &= 2^{11} \cdot 1129 \cdot 26813 \cdot 40\,644377 \cdot 17\,338437\,577121 \\ &\quad \cdot 16\,975143\,302271\,505426\,897585\,653131\,126520 \\ &\quad 182328\,037821\,729720\,833840\,187223\,, \\ p_{99} + 1 &= 2 \cdot 3^2 \cdot 83 \\ &\quad \cdot 496412\,357849\,752879\,199991\,393508\,659621\,191392\,758432 \\ &\quad 074313\,189974\,107191\,710682\,399400\,942498\,539967\,666627\,. \end{aligned}$$

These factorizations were found by Richard Crandall with the $p-1$ method and the elliptic curve method. (He used a special second phase that he developed in collaboration with Joe Buhler, that is similar to the second phase given in [3].)

Several people, including Richard Brent, Robert Silverman, Peter Montgomery, Sam Wagstaff, and ourselves, attempted to factor $n$ using the elliptic curve method, supplemented with a second phase. Brent tried 5000 elliptic curves, his first-phase bound (i.e., the bound $B_1$ from [34]) ranging from 240000 to 400000. This took 200 hours on a Fujitsu VP 100. Robert Silverman and Peter Montgomery tried 500 elliptic curves each, with a first-phase bound equal to $1\,000000$. We tried approximately 2000 elliptic curves, with first-phase bounds ranging from 300000 to 1 000000, during a one-week run on a network of approximately 75 Firefly workstations at Digital Equipment Corporation Systems Research Center (DEC SRC). The elliptic curve method did not succeed in finding a factor. Our experience indicates that *if* there had been a prime factor less than $10^{30}$, it would almost certainly have been found. If there had been a factor less than $10^{40}$ we should probably have continued with the elliptic curve method. Our decision to stop was justified by the final factorization, which the elliptic curve method did not have a reasonable chance of finding without major technological or algorithmic improvements.

The best published lower bound for the prime factors of $n$ that had been rigorously established before $n$ was completely factored is $2^{47} \approx 1.4 \cdot 10^{14}$ (see [21, Table 2]). We have been informed by Robert Silverman that the work leading to [35] implied a lower bound $2048 \cdot 10^{10}$, and that he later improved this to $2048 \cdot 10^{12}$. The best unpublished lower bound that we are aware of is $2^{51} \approx 2.25 \cdot 10^{15}$, due to Gary Gostin (1987).

If we had been certain—which we were not—that $n$ had no prime factor less than $10^{30}$, then we would have known that $n$ is a product of either two, three, or four prime factors. Among all composite numbers of 148 digits that have no prime factor less than $10^{30}$, about 15.8% are products of three primes, about 0.5% are products of four primes, and the others are products of two primes. We expected—rightly, as it turned out—to find two prime factors, but some of us would have been more excited with three large ones.

## 4. Algebraic number theory

We factored $F_9$ by means of the *number field sieve*, which is a factoring algorithm that makes use of rings of algebraic integers. The number field sieve was introduced in [28] as a method for factoring Cunningham numbers. Meanwhile, a variant of the number field sieve has been invented that can, in principle, factor general numbers, but it has not yet proved to be of practical value (see [10]).

In this section we review the basic properties of the ring $\mathbf{Z}[\sqrt[5]{2}]$, which is the ring that was used in the case of $F_9$. A more general account of algebraic number theory can be found in [46], and for computational techniques we refer to [11].

**4.1. The number field $\mathbf{Q}(\sqrt[5]{2})$ and the norm map.** The elements of the field $\mathbf{Q}(\sqrt[5]{2})$ can be written uniquely as expressions $\sum_{i=0}^{4} q_i \sqrt[5]{2}^i$, with $q_i$ belonging to the field $\mathbf{Q}$ of rational numbers. For computational purposes we identify these elements with vectors consisting of five rational components $q_0$, $q_1$, $q_2$, $q_3$, $q_4$, and addition and subtraction in the field are then just vector addition and subtraction. From the rule $\sqrt[5]{2}^5 = 2$ one readily deduces how elements of the field are to be multiplied. Explicitly, multiplying an element of the field by $\beta = \sum_{i=0}^{4} q_i \sqrt[5]{2}^i$ amounts to multiplying the corresponding column vector by the matrix

$$\begin{pmatrix} q_0 & 2q_4 & 2q_3 & 2q_2 & 2q_1 \\ q_1 & q_0 & 2q_4 & 2q_3 & 2q_2 \\ q_2 & q_1 & q_0 & 2q_4 & 2q_3 \\ q_3 & q_2 & q_1 & q_0 & 2q_4 \\ q_4 & q_3 & q_2 & q_1 & q_0 \end{pmatrix}.$$

The *norm* $N(\beta)$ of $\beta$ is defined to be the determinant of this matrix, which is a rational number. Note that the norm can be written as a homogeneous fifth-degree polynomial in the $q_i$, with integer coefficients. We have

$$N(\beta\gamma) = N(\beta)N(\gamma) \quad \text{for } \beta, \gamma \in \mathbf{Q}(\sqrt[5]{2}),$$

because the matrix belonging to $\beta\gamma$ is the product of the two matrices belonging to $\beta$ and $\gamma$. Applying this to $\gamma = \beta^{-1}$, and using that $N(1) = 1$, we find that $N(\beta) \neq 0$ whenever $\beta \neq 0$.

The norm is one of the principal tools for studying the multiplicative structure of the field, and almost all that the number field sieve needs to know about multiplication is obtained from the norm map. In particular, for the purposes of the number field sieve no multiplication routine is needed.

Below it will be useful to know that

$$\text{(4.2)} \qquad N(a - b\sqrt[5]{2}^{\,i}) = a^5 - 2^i b^5 \quad \text{for } a, b \in \mathbf{Q}, \ 1 \le i \le 4.$$

One proves this by evaluating the determinant of the corresponding matrices.

Division in the field can be done by means of linear algebra, since finding $\gamma/\beta$ is the same as solving the equation $\beta \cdot x = \gamma$, which can be written as a system of five linear equations in five unknowns. There exist better methods, but we do not discuss these, since the number field sieve needs division just as little as it needs multiplication.

4.3. **The number ring $\mathbf{Z}[\sqrt[5]{2}]$ and smoothness.** The elements $\sum_{i=0}^{4} r_i \sqrt[5]{2}^{\,i}$ of $\mathbf{Q}(\sqrt[5]{2})$ for which all $r_i$ belong to $\mathbf{Z}$ form a subring, which is denoted by $\mathbf{Z}[\sqrt[5]{2}]$. If $\beta$ belongs to $\mathbf{Z}[\sqrt[5]{2}]$, then the matrix associated with $\beta$ has integer entries, so its determinant $N(\beta)$ belongs to $\mathbf{Z}$. If $B$ is a positive real number, then a nonzero element $\beta$ of $\mathbf{Z}[\sqrt[5]{2}]$ will be called *$B$-smooth* if the absolute value $|N(\beta)|$ of its norm is $B$-smooth in the sense of §2.7. We note that $|N(\beta)|$ can be interpreted as the index of the subgroup $\beta\mathbf{Z}[\sqrt[5]{2}] = \{\beta\gamma : \gamma \in \mathbf{Z}[\sqrt[5]{2}]\}$ of $\mathbf{Z}[\sqrt[5]{2}]$:

$$\text{(4.4)} \qquad |N(\beta)| = \#(\mathbf{Z}[\sqrt[5]{2}]/\beta\mathbf{Z}[\sqrt[5]{2}]) \quad \text{for } \beta \in \mathbf{Z}[\sqrt[5]{2}], \ \beta \neq 0.$$

This follows from the following well-known lemma in linear algebra: *if $A$ is a $k \times k$ matrix with integer entries and nonzero determinant, and we view $A$ as a map $\mathbf{Z}^k \to \mathbf{Z}^k$, then the index of $A\mathbf{Z}^k$ in $\mathbf{Z}^k$ is finite and equal to $|\det A|$.*

4.5. **Ring homomorphisms.** We will need to know a little about ring homomorphisms defined on $\mathbf{Z}[\sqrt[5]{2}]$. Let $R$ be a commutative ring with 1. If $\psi: \mathbf{Z}[\sqrt[5]{2}] \to R$ is a ring homomorphism, then the element $c = \psi(\sqrt[5]{2})$ of $R$ clearly satisfies $c^5 = 2$, where 2 now denotes the element $1 + 1$ of $R$. Conversely, if $c \in R$ satisfies $c^5 = 2$, then there is a unique ring homomorphism $\psi: \mathbf{Z}[\sqrt[5]{2}] \to R$ satisfying $\psi(\sqrt[5]{2}) = c$, namely the map defined by

$$\psi\left(\sum_{i=0}^{4} r_i \sqrt[5]{2}^{\,i}\right) = \sum_{i=0}^{4} r_i c^i \qquad (r_i \in \mathbf{Z});$$

here the $r_i$ on the right are interpreted as elements of $R$, just as we put $2 = 1+1$ above. We conclude that giving a ring homomorphism from $\mathbf{Z}[\sqrt[5]{2}]$ to $R$ is the same as giving an element $c$ of $R$ that satisfies $c^5 = 2$.

**Example.** Let $n=(2^{512}+1)/2424833$, and put $R=\mathbf{Z}/n\mathbf{Z}$ and $c=(2^{205} \bmod n)$. We have $2^{512} \equiv -1 \bmod n$, and therefore

$$c^5 = (2^{1025} \bmod n) = (2 \cdot (2^{512})^2 \bmod n) = (2 \bmod n).$$

Hence, there is a ring homomorphism $\varphi\colon \mathbf{Z}[\sqrt[5]{2}] \to \mathbf{Z}/n\mathbf{Z}$ with $\varphi(\sqrt[5]{2}) = (2^{205} \bmod n)$. This ring homomorphism will play an important role in the following section.

4.6. **Fifth roots of 2 in finite fields.** One of the first things to do if one wishes to understand the arithmetic of a ring like $\mathbf{Z}[\sqrt[5]{2}]$ is to find ring homomorphisms to *finite fields* of small cardinality. As we just saw, this comes down to finding, for several small prime numbers $p$, an element $c$ that lies in a finite extension of the field $\mathbf{F}_p = \mathbf{Z}/p\mathbf{Z}$ and that satisfies $c^5 = 2$. First we consider the case that $c$ lies in $\mathbf{F}_p$ itself. Each such $c$ gives rise to a ring homomorphism $\mathbf{Z}[\sqrt[5]{2}] \to \mathbf{F}_p$, which will be denoted by $\psi_{p,c}$. The first seven examples of such pairs $(p, c)$ are

$$(4.7) \qquad (2, 0), (3, 2), (5, 2), (7, 4), (13, 6), (17, 15), (19, 15).$$

For example, the presence of the pair $(17, 15)$ on this list means that $15^5 \equiv 2 \bmod 17$; and the absence of other pairs $(17, c)$ means that $(15 \bmod 17)$ is the only zero of $X^5-2$ in $\mathbf{F}_{17}$. Note that the prime $p = 11$ is skipped, and that all other primes less than 20 occur exactly once on the list. In general, each prime $p$ that is not congruent to $1 \bmod 5$ occurs exactly once. To prove this, let $p$ be such a prime and let $k$ be a positive integer satisfying $5k \equiv 1 \bmod (p-1)$. Then the two maps $f, g\colon \mathbf{F}_p \to \mathbf{F}_p$ defined by $f(x) = x^5$, $g(x) = x^k$ are inverse to each other. Hence, there is a unique fifth root of 2 in $\mathbf{F}_p$, and it is given by $(2^k \bmod p)$. For a prime $p$ with $p \equiv 1 \bmod 5$ the fifth-power map is five-to-one. Therefore, such a prime either does not occur at all, or it occurs five times. For example, $p = 11$ does not occur, and $p = 151$ gives rise to the five pairs

$$(4.8) \qquad (151, 22), (151, 25), (151, 49), (151, 90), (151, 116).$$

Asymptotically, one out of every five primes that are $1 \bmod 5$ is of the second sort.

The case that $c$ lies in a proper extension of $\mathbf{F}_p$ is fortunately not needed in the number field sieve. It is good to keep in mind that such $c$'s nevertheless exist. For example, in a field $\mathbf{F}_{81}$ of order 81 the polynomial $(X^5-2)/(X-2) = X^4 + 2X^3 + X^2 + 2X + 1$ has four zeros; these zeros are conjugate over $\mathbf{F}_3$, and they are fifth roots of 2. In the field $\mathbf{F}_{361} = \mathbf{F}_{19}(i)$ (with $i^2 = -1$), the polynomial $X^5 - 2$ has, in addition to the zero $(15 \bmod 19)$ from (4.7), two pairs of conjugate zeros, namely $11 \pm 3i$ and $10 \pm 7i$.

4.9. **Ideals and prime ideals.** We recall from algebra that an *ideal* of $\mathbf{Z}[\sqrt[5]{2}]$ is an additive subgroup $\mathbf{b} \subset \mathbf{Z}[\sqrt[5]{2}]$ with the property that $\beta\gamma \in \mathbf{b}$ for all $\beta \in \mathbf{b}$ and all $\gamma \in \mathbf{Z}[\sqrt[5]{2}]$. The zero ideal $\{0\}$ will not be of any interest to us. The *norm* $\mathrm{N}\mathbf{b}$ of a nonzero ideal $\mathbf{b} \subset \mathbf{Z}[\sqrt[5]{2}]$ is defined to be the index of $\mathbf{b}$ in $\mathbf{Z}[\sqrt[5]{2}]$, that is, $\mathrm{N}\mathbf{b} = \#(\mathbf{Z}[\sqrt[5]{2}]/\mathbf{b})$; this is finite, since $\mathbf{b}$ contains $\beta\mathbf{Z}[\sqrt[5]{2}]$ for some nonzero $\beta$, and $\beta\mathbf{Z}[\sqrt[5]{2}]$ has already finite index (see (4.4)).

We also recall from algebra that a subset of $\mathbf{Z}[\sqrt[5]{2}]$ is an ideal if and only if it is the kernel of some ring homomorphism that is defined on $\mathbf{Z}[\sqrt[5]{2}]$. We call a nonzero ideal a *prime ideal*, or briefly a *prime* of $\mathbf{Z}[\sqrt[5]{2}]$, if it is equal to the kernel of a ring homomorphism from $\mathbf{Z}[\sqrt[5]{2}]$ to some *finite field*; and if that finite field can be taken to be a prime field $\mathbf{F}_p$, then the ideal is called a *first-degree prime*. Thus (4.7) can be viewed as a table of the "small" first-degree primes of $\mathbf{Z}[\sqrt[5]{2}]$.

If $\mathbf{p}$ is a first-degree prime, corresponding to a pair $(p, c)$, then the map $\psi_{p,c}$ induces an isomorphism $\mathbf{Z}[\sqrt[5]{2}]/\mathbf{p} \cong \mathbf{F}_p$, and therefore $\mathbf{Np}$ is equal to the prime number $p$. Conversely, if $\mathbf{p}$ is a nonzero ideal of prime norm $p$, then $\mathbf{p}$ is a first-degree prime; this is because $\mathbf{Z}[\sqrt[5]{2}]/\mathbf{p}$ is a ring with $p$ elements, and therefore isomorphic to $\mathbf{F}_p$.

In general, the norm of a prime $\mathbf{p}$ is a power $p^f$ of a prime number $p$, and $f$ is called the *degree* of $\mathbf{p}$. For example, the conjugacy classes of fifth roots of 2 in $\mathbf{F}_{81}$ and $\mathbf{F}_{361}$ indicated above give rise to one fourth-degree prime of norm 81 and two second-degree primes of norm 361. These are the smallest norms of primes of $\mathbf{Z}[\sqrt[5]{2}]$ that are of degree greater than 1.

4.10. **Generators of ideals.** Most of what we said so far about the ring $\mathbf{Z}[\sqrt[5]{2}]$ is, with appropriate changes, valid for any ring that one obtains by adjoining to $\mathbf{Z}$ a zero of an irreducible polynomial with integer coefficients and leading coefficient 1. At this point, however, we come to a property that does not hold in this generality. Namely,

$$\mathbf{Z}[\sqrt[5]{2}] \text{ is a principal ideal domain}, \tag{4.11}$$

which means that every ideal $\mathbf{b}$ of $\mathbf{Z}[\sqrt[5]{2}]$ is a *principal* ideal, that is, an ideal of the form $\beta\mathbf{Z}[\sqrt[5]{2}]$, with $\beta \in \mathbf{Z}[\sqrt[5]{2}]$. If $\mathbf{b} = \beta\mathbf{Z}[\sqrt[5]{2}]$, then $\beta$ is called a *generator* of $\mathbf{b}$.

For the proof of (4.11) we need a basic result from algebraic number theory (cf. [46, §10.2]). It implies that there is a positive constant $M$, the *Minkowski constant*, which can be explicitly calculated in terms of the ring, and which has the following property: if each prime ideal of norm at most $M$ is principal, then *every* ideal of the ring is principal. In the case of the ring $\mathbf{Z}[\sqrt[5]{2}]$ one finds that $M \doteq 13.92$, so only the primes of norm at most 13 need to be looked at. From $13 < 81$ we see that all these primes are first-degree primes.

We conclude that to prove (4.11) it suffices to show that the first-degree primes corresponding to the pairs $(2, 0)$, $(3, 2)$, $(5, 2)$, $(7, 4)$, and $(13, 6)$ are principal. This can be done without the help of an electronic computer, as follows. Trying a few values for $a$, $b$, and $i$ in (4.2), one finds that the element $1 - \sqrt[5]{2}^3$ has norm $-7$. By (4.4), the ideal $(1 - \sqrt[5]{2}^3)\mathbf{Z}[\sqrt[5]{2}]$ has norm 7, so it is a first-degree prime, corresponding to a pair $(p, c)$ with $p = 7$. But there is only one such pair, namely the pair $(7, 4)$. We conclude that the prime corresponding to the pair $(7, 4)$ is equal to $(1 - \sqrt[5]{2}^3)\mathbf{Z}[\sqrt[5]{2}]$, and therefore principal. The argument obviously generalizes to any prime number $p$ that occurs exactly once as the norm of a prime; in other words, if $p$ is a prime number with $p \not\equiv 1 \bmod 5$, and $\mathbf{p}$ is the unique prime of norm $p$, then for $\pi \in \mathbf{Z}[\sqrt[5]{2}]$ we have

$$\mathbf{p} = \pi\mathbf{Z}[\sqrt[5]{2}] \Leftrightarrow |N(\pi)| = p. \tag{4.12}$$

Applying this to $\pi = \sqrt[5]{2}$, $p = 2$, we find that the prime corresponding to $(2, 0)$ is principal. The prime of norm 3 is taken care of by $\pi = 1 + \sqrt[5]{2}$, the prime of norm 5 by $\pi = 1 + \sqrt[5]{2}^2$, and the prime of norm 13 by $\pi = 3 - 2\sqrt[5]{2}^3$. This proves (4.11).

It will be useful to have a version of (4.12) that is also valid for primes that are 1 mod 5. Let $\mathbf{p}$ be a first-degree prime of $\mathbf{Z}[\sqrt[5]{2}]$, corresponding to a pair $(p, c)$, and let $\pi \in \mathbf{Z}[\sqrt[5]{2}]$. Then we have

$$\mathbf{p} = \pi\mathbf{Z}[\sqrt[5]{2}] \Leftrightarrow \psi_{p,c}(\pi) = 0 \quad \text{and} \quad |N(\pi)| = p. \tag{4.13}$$

To prove $\Rightarrow$, suppose that $\mathbf{p} = \pi\mathbf{Z}[\sqrt[5]{2}]$. Then we have $\pi \in \mathbf{p}$, and $\mathbf{p}$ is the kernel of $\psi_{p,c}$, so $\psi_{p,c}(\pi) = 0$. Also, from (4.4) we see that $|N(\pi)| = \mathbf{Np} = p$. To prove $\Leftarrow$, suppose that $\psi_{p,c}(\pi) = 0$ and $|N(\pi)| = p$. Then $\pi$ belongs to the kernel $\mathbf{p}$ of $\psi_{p,c}$, so $\pi\mathbf{Z}[\sqrt[5]{2}]$ is contained in $\mathbf{p}$. Since they both have index $p$ in $\mathbf{Z}[\sqrt[5]{2}]$, they must be equal. This proves (4.13).

**Example.** The number $\pi = 1 + \sqrt[5]{2}^2 - 2\sqrt[5]{2}^3$ is found to have norm $-151$. Substituting successively the values $c = 22, 25, 49, 90, 116$ listed in (4.8) for $\sqrt[5]{2}$, we find that only $c = 116$ gives rise to a number that is 0 mod 151. Hence, $\pi$ generates the prime corresponding to the pair $(151, 116)$. (Alternatively, one can determine the correct value of $c$ by calculating the gcd of $X^5 - 2$ and $1 + X^2 - 2X^3$ in $\mathbf{F}_{151}[X]$, which is found to be $X - 116$.)

4.14. **Unique factorization.** A basic theorem in algebra asserts that principal ideal domains are unique factorization domains. Thus (4.11) implies that the nonzero elements of $\mathbf{Z}[\sqrt[5]{2}]$ can be factored into prime elements in an essentially unique way. More precisely, let for every prime $\mathbf{p}$ of $\mathbf{Z}[\sqrt[5]{2}]$ an element $\pi_{\mathbf{p}}$ with $\mathbf{p} = \pi_{\mathbf{p}}\mathbf{Z}[\sqrt[5]{2}]$ be chosen. Then there exist for every nonzero $\beta \in \mathbf{Z}[\sqrt[5]{2}]$ uniquely determined nonnegative integers $m(\mathbf{p})$ such that $m(\mathbf{p}) = 0$ for all but finitely many $\mathbf{p}$, and such that

$$\beta = \varepsilon \cdot \prod_{\mathbf{p}} \pi_{\mathbf{p}}^{m(\mathbf{p})},$$

where $\varepsilon$ belongs to the group $\mathbf{Z}[\sqrt[5]{2}]^*$ of units of $\mathbf{Z}[\sqrt[5]{2}]$, and where the product ranges over all primes $\mathbf{p}$ of $\mathbf{Z}[\sqrt[5]{2}]$. We have $m(\mathbf{p}) > 0$ if and only if $\beta \in \mathbf{p}$, and in this case we say that $\mathbf{p}$ *occurs* in $\beta$. We shall call $m(\mathbf{p})$ the *number of factors* $\mathbf{p}$ *in* $\beta$. Note that we have

$$|N(\beta)| = \prod_{\mathbf{p}} \mathbf{Np}^{m(\mathbf{p})}, \tag{4.15}$$

because $|N(\pi_{\mathbf{p}})| = \mathbf{Np}$ and $|N(\varepsilon)| = 1$, both by (4.4).

**Examples.** First let $\beta = -1 + \sqrt[5]{2}^4$. The norm of $\beta$ is 15, so from (4.15) we see that only the primes of norms 3 and 5 occur in $\beta$, each with exponent 1. Using the generators $1 + \sqrt[5]{2}$ and $1 + \sqrt[5]{2}^2$ that we found above for these primes, we obtain the prime factorization

$$-1 + \sqrt[5]{2}^4 = \varepsilon_1 \cdot (1 + \sqrt[5]{2}) \cdot (1 + \sqrt[5]{2}^2),$$

where $\varepsilon_1 = -1 + \sqrt[5]{2}$. Note that $\varepsilon_1$ is indeed a unit, by $N(\varepsilon_1) = 1$ and (4.4). Similarly, one finds that the prime factorization of the element $1 + \sqrt[5]{2}^3$ of norm 9 is given by

$$1 + \sqrt[5]{2}^3 = \varepsilon_2 \cdot (1 + \sqrt[5]{2})^2,$$

where $\varepsilon_2 = -1 + \sqrt[5]{2}^2 - \sqrt[5]{2}^3 + \sqrt[5]{2}^4$. The factorization of the number 5 is quite special: it is given by

$$5 = \varepsilon_3 \cdot (1 + \sqrt[5]{2}^2)^5, \tag{4.16}$$

where $\varepsilon_3 = \varepsilon_1^2 \varepsilon_2^{-2}$.

4.17. **Units.** The *Dirichlet unit theorem* (see [46, §12.4]) describes the unit groups of general rings of algebraic integers. It implies that the group $\mathbf{Z}[\sqrt[5]{2}]^*$ of units of $\mathbf{Z}[\sqrt[5]{2}]$ is generated by two multiplicatively independent units of infinite order, together with the unit $\varepsilon_0 = -1$. We found that we could take these two units of infinite order to be the elements $\varepsilon_1$ and $\varepsilon_2$ from the examples just given, in the sense that every unit $\varepsilon$ that we ever encountered was of the form

$$\varepsilon = \varepsilon_0^{v(0)} \varepsilon_1^{v(1)} \varepsilon_2^{v(2)}, \quad \text{with } v(0), v(1), v(2) \in \mathbf{Z}.$$

We never attempted to prove formally that *every* unit is of this form, although this would probably have been easy from the material that we accumulated. There exist good algorithms that can be used to verify this (see [8]).

Given a unit $\varepsilon$, one can find the integers $v(i)$ in the following way. It is easily checked that $N(\varepsilon_0) = -1$ and that $N(\varepsilon_1) = N(\varepsilon_2) = 1$. Hence, $N(\varepsilon) = \varepsilon_0^{v(0)} = (-1)^{v(0)}$, and this determines $v(0) \pmod 2$. Next let $c_1 = \exp((\log 2)/5)$ and $c_2 = \exp((2\pi i + \log 2)/5)$; these are complex fifth roots of 2. Denote by $\psi_i$ the ring homomorphism from $\mathbf{Z}[\sqrt[5]{2}]$ to the field of complex numbers that maps $\sqrt[5]{2}$ to $c_i$, for $i = 1, 2$. Then we have

$$\begin{aligned} \log|\psi_1(\varepsilon)| &= v(1) \cdot \log|\psi_1(\varepsilon_1)| + v(2) \cdot \log|\psi_1(\varepsilon_2)|, \\ \log|\psi_2(\varepsilon)| &= v(1) \cdot \log|\psi_2(\varepsilon_1)| + v(2) \cdot \log|\psi_2(\varepsilon_2)|. \end{aligned}$$

A direct calculation shows that $\log|\psi_1(\varepsilon_1)| \log|\psi_2(\varepsilon_2)| - \log|\psi_1(\varepsilon_2)| \log|\psi_2(\varepsilon_1)| \neq 0$, so $v(1)$, $v(2)$ can be solved uniquely from a system of two linear equations. Since the $v(i)$ are expected to be integers, we can do the computation in limited precision and round the result to integers. The inverse of the coefficient matrix can be computed once and for all.

4.18. **A table of first-degree primes.** The table (4.7) of first-degree primes of norm up to 19 was, for the purpose of factoring $F_9$, extended up to 1294973; see §6 for the considerations leading to the choice of this limit. We made the table by treating all prime numbers $p \leq 1294973$ individually. For primes $p$ that are not 1 mod 5 we found $c$ with the formula $c \equiv 2^k \bmod p$ given in §4.6. For primes $p$ that are 1 mod 5 we first checked whether $2^{(p-1)/5} \equiv 1 \bmod p$, which is a necessary and sufficient condition for 2 to have a fifth root modulo $p$. If this condition was satisfied—which occurred for 4944 primes, ranging from 151 to 1294471—then the five values of $c \pmod p$ were found by means of a standard algorithm for finding zeros of polynomials over finite fields (see [26]). The entire calculation took only a few minutes on a DEC3100 workstation. We found that there are 99500 first-degree primes of norm up to 1294973, of which the last one is given by $(1294973, 1207394)$.

4.19. **A table of prime elements.** For each of the 99500 primes $\mathbf{p}$ in our table we also needed to know an explicit generator $\pi_{\mathbf{p}}$. These can be found by means of a brute-force search, as follows. Calculate the norms of all elements

$\sum_{i=0}^{4} r_i \sqrt[5]{2}^i \in \mathbf{Z}[\sqrt[5]{2}]$ for which the integers $|r_i|$ are below some large bound; since the norm is a polynomial of degree five in the $r_i$, one can use a difference scheme in this calculation. Whenever an element is found of which the absolute value of the norm is equal to $p$ for one of the pairs $(p, c)$ in the table, then one knows that a generator of a prime of norm $p$ has been found. If $p \not\equiv 1 \bmod 5$, then $c$ is uniquely determined by $p$, and the pair $(p, c)$ can be crossed off the list. If $p \equiv 1 \bmod 5$, then we use (4.13) to determine the correct value of $c$ for which $(p, c)$ can be crossed off the list.

What we actually did was slightly different. We did not search among the elements $\sum_{i=0}^{4} r_i \sqrt[5]{2}^i$ as just described, but only among the elements that belong to the subring $\mathbf{Z}[\alpha]$ of $\mathbf{Z}[\sqrt[5]{2}]$, where $\alpha = -\sqrt[5]{2}^3$. This enabled us to use a program that was written for a previous occasion. We considered all 1092846526 expressions $\sum_{i=0}^{4} s_i \alpha^i \in \mathbf{Z}[\alpha]$ for which the $s_i$ have no common factor, for which $s_i \geq 0$ if $s_{i+1}$ through $s_4$ are 0, and that lie in the "sphere" $\sum_{i=0}^{4} s_i^2 2^{6i/5} \leq 15000$. In this way we determined 49726 of the 99500 generators. For the other 49774 first-degree prime ideals $\mathbf{p}$ the same search produced generators for the ideals $\alpha\mathbf{p}$ of norm $8 \cdot \mathbf{Np}$, so that we could determine the proper generators by dividing out $\alpha$. The whole calculation took only a few hours on a single workstation.

We found it convenient to have $N(\pi_{\mathbf{p}}) > 0$ for all $\mathbf{p}$. To achieve this, one can replace $\pi_{\mathbf{p}}$ by $-\pi_{\mathbf{p}}$, if necessary.

## 5. The number field sieve

As in §3, we let $n$ be the number $F_9/2424833$. The account of the number field sieve that we give in this section is restricted to the specific case of the factorization of the number $n$.

To factor $n$ with the number field sieve, we made use of the ring $\mathbf{Z}[\sqrt[5]{2}]$ that was discussed in the previous section. As we saw in §4.5, there is a ring homomorphism $\varphi: \mathbf{Z}[\sqrt[5]{2}] \to \mathbf{Z}/n\mathbf{Z}$ that maps $\sqrt[5]{2}$ to $2^{205} \bmod n$. An important role is played by the element $\alpha = -\sqrt[5]{2}^3$, which has the property that $\varphi(\alpha) = (-2^{615} \bmod n) = (2^{103} \bmod n)$. What is important about this is that $2^{103}$ is very small with respect to $n$; it is not much bigger than $\sqrt[5]{n}$. Note that for any $a, b \in \mathbf{Z}$ we have

$$(5.1) \qquad \varphi(a + b\alpha) = \varphi(a + 2^{103} b) \quad (\text{in } \mathbf{Z}/n\mathbf{Z}).$$

This equality plays the role that the congruence $b \equiv n + b \bmod n$ played in the rational sieve from §2.7.

In the rational sieve, the factor base was formed by all prime numbers up to a certain limit $B$. In the present case the factor base was selected as follows. Let the set $\mathrm{P} \subset \mathbf{Z}[\sqrt[5]{2}]$ consist of: (i) the 99700 prime numbers $p \leq B_1 = 1295377$; (ii) the three generating units $\varepsilon_0$, $\varepsilon_1$, and $\varepsilon_2$ (see §4.17); (iii) the generators $\pi_{\mathbf{p}}$ of the 99500 first-degree primes $\mathbf{p}$ of $\mathbf{Z}[\sqrt[5]{2}]$ with $\mathbf{Np} \leq B_2 = 1294973$ (see §§4.18 and 4.19). For each $p \in \mathrm{P}$, let $a_p = \varphi(p) \in \mathbf{Z}/n\mathbf{Z}$. These formed the factor base.

Relations were found in several ways. In the first place, there are relations that are already valid in $\mathbf{Z}[\sqrt[5]{2}]$, before $\varphi$ is applied. Three such relations are given by $\varepsilon_0^2 = 1$, $2 = \sqrt[5]{2}^5$, and $5 = \varepsilon_1^2 \varepsilon_2^{-2}(1 + \sqrt[5]{2}^2)^5$ (see (4.16)), but we did not use these (the first one is in fact useless). In addition, there is one such relation for each of the 4944 prime numbers $p \equiv 1 \bmod 5$ that occur five times in the table of pairs $(p, c)$ from §4.18. Such a prime number $p$ factors in $\mathbf{Z}[\sqrt[5]{2}]$ as

$$p = \varepsilon \cdot \prod_{\mathbf{p}} \pi_{\mathbf{p}}, \tag{5.2}$$

where $\varepsilon$ is a unit and $\mathbf{p}$ ranges over the five primes of norm $p$. To see this, observe that from $\psi_{p,c}(p) = 0$ it follows that each of these $\mathbf{p}$'s occurs in $p$. Since this accounts for the full norm $p^5$ of $p$ (cf. (4.15)), we obtain (5.2). The unit $\varepsilon$ occurring in (5.2) can be expressed in $\varepsilon_1$ and $\varepsilon_2$ by means of the method explained in §4.17 (the unit $\varepsilon_0$ does not occur, since $p$ and the $\pi_{\mathbf{p}}$ are of positive norm). Note that for this method we do not need to know the unit $\varepsilon$ itself, but only the numbers $\log|\psi_i(\varepsilon)|$ for $i = 1, 2$, and these can by (5.2) be computed from the corresponding quantities for $p$ and $\pi_{\mathbf{p}}$. The 4944 relations found in this way constituted no more than 2.5% of the $\sim 200000$ relations that we needed.

We found the remaining $\sim 195000$ relations between the $a_p$ by searching for pairs of integers $a$, $b$, with $b > 0$, satisfying the following conditions:

(5.3) $\gcd(a, b) = 1$;

(5.4) $|a + 2^{103}b|$ is built up from prime numbers $\leq B_1$ and at most one larger prime number $p_1$, which should satisfy $B_1 < p_1 < 10^8$;

(5.5) $|a^5 - 8b^5|$ is built up from prime numbers $\leq B_2$ and at most one larger prime number $p_2$, which should satisfy $B_2 < p_2 < 10^8$.

If the large prime $p_1$ in (5.4) does not occur, then we write $p_1 = 1$, and likewise for $p_2$ in (5.5). Pairs $a$, $b$ for which $p_1 = p_2 = 1$ will be called *full relations*, and the other pairs *partial relations*.

We note that the number $a^5 - 8b^5$ equals the norm of $a + b\alpha$, by (4.2). Hence, condition (5.5), with $p_2 = 1$, is equivalent to the requirement that $a + b\alpha$ be $B_2$-smooth, in the terminology of §4.3.

Before we describe, in §6, how the search for such pairs was performed, let us see how they give rise to relations between the $a_p$. We begin with a lemma concerning the prime factorization of elements of the form $a + b\alpha$.

**Lemma.** *Let* $a, b \in \mathbf{Z}$, $\gcd(a, b) = 1$. *Then all primes* $\mathbf{p}$ *that occur in* $a + b\alpha$ *are first-degree primes.*

*Proof.* Suppose that $\mathbf{p}$ occurs in $a + b\alpha$, and let $\psi$ be a ring homomorphism from $\mathbf{Z}[\sqrt[5]{2}]$ to a finite field $F$ such that $\mathbf{p}$ is the kernel of $\psi$. Let $p$ be the characteristic of $F$, so that $\mathbf{F}_p$ is a subfield of $F$. We have $a + b\alpha \in \mathbf{p}$, so $\psi(a + b\alpha) = 0$, and therefore

$$\psi(a) = -\psi(b)\psi(\alpha). \tag{5.6}$$

Note that $\psi(a)$ and $\psi(b)$ belong to $\mathbf{F}_p$, because $a, b \in \mathbf{Z}$. If $\psi(b) = 0$, then by (5.6) we have $\psi(a) = 0$ as well, so $b$ and $a$ are both divisible by $p$, which contradicts that $\gcd(a, b) = 1$. Hence, $\psi(b) \neq 0$, and from (5.6) we now see that $\psi(\alpha) = -\psi(a)/\psi(b)$ also belongs to $\mathbf{F}_p$. We claim that $\psi(\sqrt[5]{2})$ belongs to

$\mathbf{F}_p$ as well. If $p = 2$, we have $\psi(\sqrt[5]{2})^5 = \psi(2) = 0$, so $\psi(\sqrt[5]{2}) = 0$, which does belong to $\mathbf{F}_2$. If $p \neq 2$, then $\alpha^2 = 2\sqrt[5]{2}$ implies that $\psi(\sqrt[5]{2}) = \psi(\alpha)^2/\psi(2)$, which belongs to $\mathbf{F}_p$. From $\psi(\sqrt[5]{2}) \in \mathbf{F}_p$ it follows that $\psi$ maps all of $\mathbf{Z}[\sqrt[5]{2}]$ to $\mathbf{F}_p$. Hence, $\mathbf{p}$ is the kernel of a ring homomorphism from $\mathbf{Z}[\sqrt[5]{2}]$ to $\mathbf{F}_p$, which by definition means that it is a first-degree prime. This proves the lemma.

The lemma reduces the factorization of $a + b\alpha$, with $\gcd(a, b) = 1$, to the factorization of its norm $a^5 - 8b^5$, as follows. Let $p$ be a prime number dividing $a^5 - 8b^5$. If $p \not\equiv 1 \bmod 5$, then $p$ is the norm of a unique prime $\mathbf{p}$, and the number of factors $\mathbf{p}$ in $a + b\alpha$ must be equal to the number of factors $p$ in $a^5 - 8b^5$. If $p \equiv 1 \bmod 5$, then we have to determine which fifth root $c$ of 2 $(\bmod\, p)$ is involved. By (5.6), we must have $(c \bmod p)^3 = (a \bmod p)/(b \bmod p)$, and this uniquely determines $c$, since $c^3 \equiv c'^3 \bmod p$ gives $2c \equiv 2c' \bmod p$ upon squaring. Once we have determined $c$, we know which $\mathbf{p}$ occurs in $a + b\alpha$, and again the number of factors $\mathbf{p}$ in $a + b\alpha$ is equal to the number of factors $p$ in $a^5 - 8b^5$.

Let us now first consider the case that $a$, $b$ is a full relation. Then the factorization of $a + b\alpha$ has the form

$$a + b\alpha = \varepsilon \cdot \prod_{\mathbf{p}} \pi_{\mathbf{p}}^{u(\mathbf{p})},$$

where $\varepsilon$ is a unit and $\mathbf{p}$ ranges over the first-degree primes of norm at most $B_2$. We just explained how the exponents $u(\mathbf{p})$ can be determined from the prime factorization of $a^5 - 8b^5$. We can write

$$\varepsilon = \prod_{i=0}^{2} \varepsilon_i^{v(i)},$$

where the $v(i)$ are determined as in §4.17; just as with (5.2), it is not necessary to calculate $\varepsilon$ for this. Factoring $a + 2^{103}b$, we obtain an identity of the form

$$a + 2^{103}b = \prod_p p^{w(p)},$$

with $p$ ranging over the prime numbers $\leq B_1$ and $w(p) \in \mathbf{Z}_{\geq 0}$ (if $a + 2^{103}b < 0$, use $-a$, $-b$ instead of $a$, $b$). Now replace, in (5.1), both sides by their factorizations. Then we find that

$$\prod_{i=0}^{2} \varphi(\varepsilon_i)^{v(i)} \cdot \prod_{\mathbf{p}} \varphi(\pi_{\mathbf{p}})^{u(\mathbf{p})} = \prod_p \varphi(p)^{w(p)}.$$

In this way, each full relation $a$, $b$ gives rise to a relation between the $a_p$.

With partial relations the situation is a bit more complicated. They give rise to relations between the $a_p$ only if they are combined into *cycles*, as described in [30]. In each cycle, one takes an alternating product of relations $\varphi(a + b\alpha) = \varphi(a + 2^{103}b)$, in such a way that the large prime ideals and prime numbers cancel. This leads to a relation between the $a_p$, by a procedure that is completely similar to the one above. It is not necessary to know generators $\pi_{\mathbf{p}}$ for the large prime ideals, since these are divided out.

If, in (5.5), we have $p_2 > 1$, then the additional prime ideal corresponds to the pair $(p_2, c \bmod p_2)$, where $c = a^2/(2b^2)$; this is uniquely determined by $p_2$ unless $p_2 \equiv 1 \bmod 5$.

## 6. Sieving

The search for pairs $a$, $b$ satisfying conditions (5.3), (5.4), and (5.5) was performed by means of a standard sieving technique that is a familiar ingredient of the quadratic sieve algorithm (see [38]). For a description of this technique as it is used in the number field sieve, we refer to [28] and [10, §§4 and 5].

We used 2.2 million values of $b$, all satisfying $0 < b < 2.5 \cdot 10^6$. For each $b$, we sieved $|a + 2^{103}b|$ with the primes $\leq B_1$, and we sieved $|a^5 - 8b^5|$ with the primes $\leq B_2$, each over $10^8$ consecutive $a$-values centered roughly at $8^{1/5} \cdot b$.

The best values for $a$ are those that are close to $8^{1/5} \cdot b$. If we take for instance $b = 10^6$, then for such $a$'s we are asking for simultaneous smoothness of two numbers close to $10^{37}$ and $8 \cdot 10^{30}$; for $b = 10^7$ this becomes $10^{38}$ and $8 \cdot 10^{35}$. The quadratic sieve algorithm when applied to $n$ would depend on the smoothness of numbers close to $\sqrt{n}$ times the sieve length, which amounts to at least $10^{80}$. This is the main reason why the number field sieve performs better for this value of $n$ than the quadratic sieve. The comparison is still very favorable when $a$ is further removed from the center of its interval, although the numbers become larger. The tails of the interval are less important, so the fact that centering it at 0 would have been better did not bother us.

Smaller $b$-values are more likely to produce good pairs $a$, $b$ than larger ones. The best approach is therefore to process the $b$-values consecutively starting at 1, until the total number of full relations plus the number of independent cycles among the partial relations that have been found equals $\sim 195000$. One can only hope that this happens before $b$ assumes prohibitively large values. Of course, $B_1$ and $B_2$ must have been selected in such a way that one is reasonably confident that this approach will succeed. This is discussed below.

We started sieving in mid-February 1990 on approximately 35 workstations at Bellcore. On the workstations we were using (DEC3100's and SPARC's) each $b$ took approximately eight minutes to process. We had to split up the $a$-intervals of length $10^8$ into 200 intervals of length $5 \cdot 10^5$, in order to avoid undue interference with other programs. After a month of mostly night-time use of these workstations, the first range of $10^5$ $b$'s was covered. Mid-March, the network of Firefly workstations at DEC SRC was also put to work. This approximately tripled our computing power. With these forces we could have finished the sieving task within another seven months. However, at the time, we did not know this, since we did not know how far we would have to go with $b$.

Near the end of March it was rumored that we had a competitor. After attempts to join forces had failed, we decided to accelerate a little by following the strategy described in [29]. We posted messages on various electronic bulletin boards, such as sci.crypt and sci.math, soliciting help. A sieving program, plus auxiliary driver programs to run it, were made available at a central machine at DEC SRC in Palo Alto to anyone who expressed an interest in helping us. After contacting one of us personally, either by electronic mail or by telephone, a possible contributor was also provided with a unique range of consecutive

$b$-values. The size of the range assigned to a particular contributor depended on the amount of free computing time the contributor expected to be able to donate. Each range was sized to last for about one week, after which a new range was assigned. This allowed us to distribute the available $b$'s reasonably evenly over the contributors, so that the $b$'s were processed more or less consecutively.

It is difficult to estimate precisely how many workstations were enlisted in this way. Given that we had processed 2.2 million $b$'s by May 9, and assuming that we mostly got night-time cycles, we must have used the equivalent of approximately 700 DEC3100 workstations. We thus achieved a sustained performance of more than 3000 mips for a period of five weeks, at no cost. (Mips is a unit of speed of computing, 1 mips being one million instructions per second.) The total computational effort amounted to about 340 mips-years (1 mips-year is about $3.15 \cdot 10^{13}$ instructions). We refer to the acknowledgments at the end of this paper for the names of many of the people and institutions who responded to our request and donated computing time.

Each copy of the sieving program communicated the pairs $a$, $b$ that it found by electronic mail to DEC SRC, along with the corresponding pair $p_1$, $p_2$ and, in the case $p_2 > 1$, $p_2 \equiv 1 \bmod 5$, the residue class $(a/b \bmod p_2)$. In order not to overload the mail system at DEC SRC, the pairs were sent at regular intervals. At DEC SRC, these data were stored on disk. Notice that the corresponding two factorizations were *not* sent, due to storage limitations. These were later recomputed at DEC SRC, but only for the relations that turned out to be useful in producing cycles. The residue class $(a/b \bmod p_2)$ could also have been recomputed, but since it simplified the cycle counting we found it more convenient to send it along. Notice that $(a/b \bmod p_2)$ distinguishes between the five prime ideals of norm $p_2$.

When we ran the quadratic sieve factoring algorithm in a similar manner (see [29]), we could be wasteful with inputs: we made sure that different inputs were distributed to our contributors, but not that they were actually processed. We could afford this approach because we had millions of inputs, each of which was in principle capable of producing thousands of relations. For the number field sieve the situation is different: each $b$ produces only a small number of relations, if any, and the average yield decreases as $b$ increases. In order not to lose our rather scarce and valuable "good" inputs (i.e., the small $b$-values), we wanted to be able to monitor what happened to them after they were given out. For this reason, each copy of the sieving program also reported through electronic mail which $b$'s from its assigned range it had completed. This allowed us to check them off from the list of $b$'s we had distributed. Values that were not checked off within approximately ten days were redistributed. Occasionally this led to duplications, but these could easily be sorted out.

By May 7 we had used approximately 2.1 million $b$'s less than 2.5 million, and we had collected 44106 full relations and 2999903 partial relations. The latter gave rise to a total of 158105 cycles. Since $44106 + 158105$ is well over 195000, this was already more than we needed. Nevertheless, to facilitate finding the dependencies, we went on for two more days. By May 9, after approximately 2.2 million $b$'s, we had 45719 full relations and 176025 cycles among 3 114327 partial relations. Only about one fifth of these 3 114327 relations turned out to be useful, in the sense that they actually appeared in one of the 176025 cycles. It took a few hours on a single workstation to find the cycles in terms of the $a$, $b$, $p_1$, and $p_2$ involved, by means of an algorithm explained in [30]. The number of cycles of each length is given in Table 1.

TABLE 1

| cycle length | number of cycles | cycle length | number of cycles |
|---|---|---|---|
| 2 | 48289 | 11 | 473 |
| 3 | 43434 | 12 | 243 |
| 4 | 32827 | 13 | 100 |
| 5 | 22160 | 14 | 55 |
| 6 | 13444 | 15 | 14 |
| 7 | 7690 | 16 | 8 |
| 8 | 4192 | 17 | 2 |
| 9 | 2035 | 19 | 2 |
| 10 | 1055 | 20 | 2 |

This is what we hoped and more or less expected to happen, but there was no guarantee that our approach would work. For any choice of $B_1$ and $B_2$ (and size of $a$-interval) we could quite accurately predict how many full and partial relations we would find by processing all $b$'s up to a certain realistic limit. This made it immediately clear that values $B_1$ and $B_2$ for which full relations alone would suffice would be prohibitively large.

Thus we were faced with the problem of choosing $B_1$ and $B_2$ in such a way that the full relations plus the cycles among the partials would be likely to provide us with sufficiently many relations between the $a_p$. It is, however, hard to predict how many partials are needed to produce a given number of cycles. For instance, the average number of cycles of length 2 resulting from a given number of partials can be estimated quite accurately, but the variance is so large that for each particular collection of partials this estimate may turn out to be far too optimistic or pessimistic. An estimate that is too low is harmless, but an estimate that is too high has very serious consequences: once $b$ is sufficiently large, hardly any new fulls or partials will be found, and the only alternative is to start all over again with larger $B_1$ and $B_2$. As a consequence, we selected the values for $B_1$ and $B_2$ carefully and conservatively, we made sure that we did not skip many $b$-values, and we milked each $b$ for all it was worth by using an excessively long $a$-interval.

We decided to set the size of the factor base approximately equal to $2 \cdot 10^5$ only after experiments had ruled out $1.2 \cdot 10^5$, $1.4 \cdot 10^5$, and $1.6 \cdot 10^5$ as probably too small, and $1.8 \cdot 10^5$ as too risky. For $2 \cdot 10^5$ we predicted $\sim 50000$ full and at least 3 million partial relations after the first 2.5 million $b$'s. This prediction was based on Figure 1 (next page), where the results of some preliminary runs of the sieving program are presented. For $i$ ranging from 1 to 40 the total number of relations (fulls plus partials) found for the 300 consecutive $b$'s starting at $i \cdot 10^5$ is given as a function of $i$. The upper curve gives the yield for an $a$-interval of length $10^8$, the lower curve for length $2 \cdot 10^7$.

Our experience with other number field sieve factorizations made us hope that 3 million partials would produce 150000 cycles, which indeed turned out

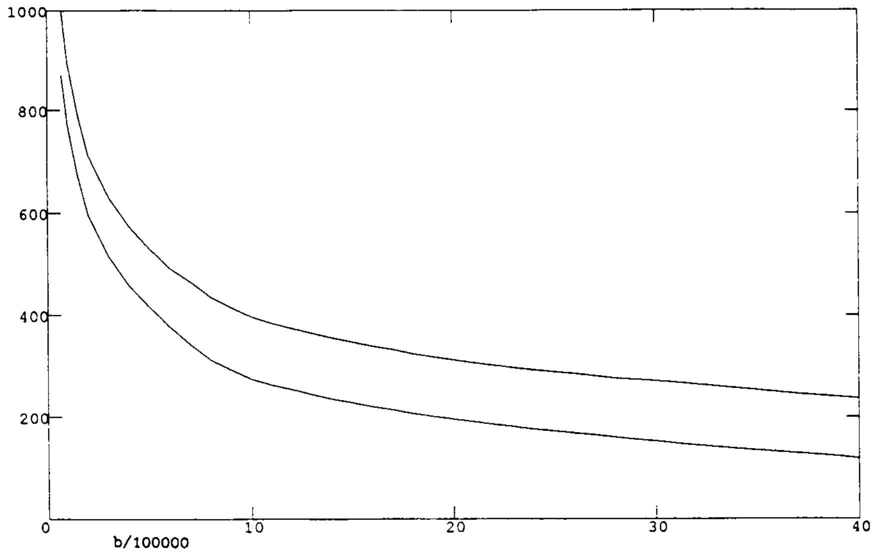

FIGURE 1. The number of full and partial relations found per 300 *b*'s, for $2.10^7$ and for $10^8$ *a*'s

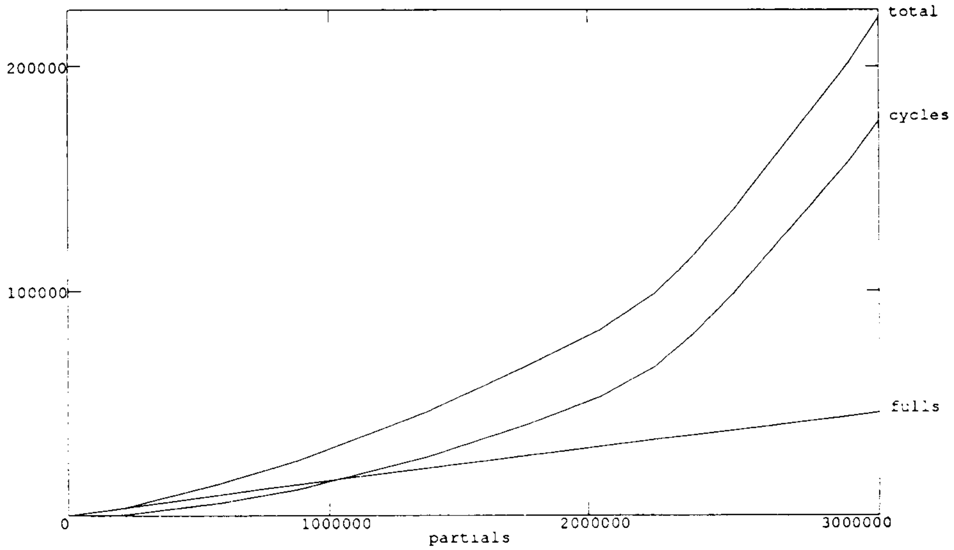

FIGURE 2. The number of cycles and full relations as a function of the number of partial relations

to be the case. But even if 3 million partials had not been enough, we knew that the *b*'s between 2.5 and 4 million would lead to at least another million partials, and a good chance to find enough cycles. In Figure 2 we give the number of cycles, the number of full relations, and their sum, that were obtained after we had found a given number of partial relations. This does not include the initial 4944 relations.

Now that we have seen how everything worked out in this particular case, we know that with the same $B_1$ and $B_2$ and a much *smaller* $a$-interval we could have produced 3 million partials in much less time after using *more* $b$'s. For example, halving the length of the $a$-interval would reduce the average yield per $b$-value by only 15%. It would probably have been optimal to use about $1.5 \cdot 10^7$ values of $a$ per $b$, with $b$ ranging up to about 5.5 million; this would have taken about 40% of the time that we actually spent. Still, we cannot be certain that this would have given rise to the same number of cycles.

We could have profited a little from the known factor 2424833 of $F_9$ by putting it in the factor base, along with the prime ideal corresponding to $(2424833, 2^{205} \bmod 2424833)$, since the prime appears on the right if and only if the prime ideal appears on the left. We realized this only after the third author had found seven "awfully suspicious" pairs $a$, $b$, namely pairs with $p_1 = p_2 = 2424833$, while generating the cycles.

To conclude the second step, the full relations and the cycles had to be transformed into relations between the $a_p$. To this end, we recomputed the $2 \cdot 722241$ factorizations corresponding to the 722241 (not all distinct) pairs $a$, $b$ involved, and determined the unit contributions. This work was divided over fifteen workstations at DEC SRC, and it took about sixteen hours.

## 7. Finding dependencies

As a result of the computations described in the previous section, we had $4944+45719+176025 = 226688$ relations between $3+99700+99500 = 199203$ different $a_p$'s. To finish the factorization of $n$, we had to determine a few dependencies between the 226688 rows of the 199203-column matrix over $\mathbf{F}_2$ that one obtains by taking the relations (i.e., the exponents of the $a_p$) modulo 2. A dense representation of this matrix would require more than 5 Gigabytes ($= 5 \cdot 2^{30}$ bytes) of storage, where one byte represents 8 bits. Fortunately, the matrix is sparse, because relatively few primes and prime ideals appear in the factorizations leading to the relations; this situation is slightly worsened by the fact that we obtained many relations by combining partial relations. In any case, there were only 11 264596 nonzero entries in the matrix, for an average of 49.7 nonzero entries per row. Thus, the entire matrix could easily be stored.

Finding dependencies was still a challenging task. The sieving step had posed no problems that had not already been solved for other numbers, except that an unusually large amount of computing time had to be arranged. The matrix step, however, presented a difficulty that had not been encountered in previous factorizations. Actually, the only reason that we had not embarked upon the factorization of $F_9$ earlier is that we did not know how to handle the matrix.

The largest matrices that we had ever dealt with in previous factorizations contained approximately 80000 columns, and a few more rows. Dependencies modulo 2 among the rows were found in an entirely straightforward fashion by means of ordinary Gaussian elimination, with pivot-search from the sparse side. In this way some profit could be gained from the sparseness, but not much: usually, the storage that one ultimately needs is about two thirds of what it would have been in the dense case. This fits in only 0.5 Gigabytes for an 80000 matrix, so that the elimination task for such a matrix is more or less trivial for someone with access to a large supercomputer. At DEC SRC, where the computations were carried out, the only machine with enough disk space that could be devoted entirely to the elimination task was a four-processor Firefly workstation. On this workstation, elimination of a sparse 80000-matrix takes approximately six weeks. Here we should note that for two of the three 80000-matrices we processed in this way, the resulting dependencies turned out to be faulty. In both instances a rerun (with another six-week wait!) was successful. We suspect that in both first runs an irreproducible cache read or write error had occurred. Clearly, a single bit error can render the entire computation worthless.

Extrapolation of these figures to a 200000-matrix did not look promising. Even if our workstation had enough disk space, $6 \cdot (2.5)^3 \approx 90$ weeks is unacceptably long, and the probability of a bit error occurring would be unacceptably large. On a supercomputer the figures still would have looked unattractive. Therefore, we investigated whether there was a better way to profit from the sparseness of the matrix.

Among the several existing techniques for dealing with sparse matrices, we decided to attempt *structured Gaussian elimination* [24, 39]. In structured Gaussian elimination the columns of the matrix are partitioned into *heavy* and *sparse* columns. Initially, all columns are considered sparse. Roughly speaking, one does eliminations with pivots in sparse columns that cause fill-in only in the heavy columns of the matrix, thereby removing the pivot rows and columns from the matrix. When this is impossible, one either moves some of the columns from the sparse to the heavy part, or one removes some excess rows, if there are any. Next, one tries again. This is repeated until no sparse columns are left. For reasons that are not yet understood it seems to be beneficial to have many excess rows initially.

During this process one does not keep track of what happens in the heavy columns, but one remembers only which eliminations have been carried out. This information can then be used to build the smaller but much denser matrix corresponding to the heavy columns, and to convert dependencies among its rows into dependencies among the rows of the original matrix. Dependencies in the smaller matrix can be found by means of ordinary Gaussian elimination.

It took us a few hours on a single workstation to reduce our 226688-row and 199203-column matrix to a 72413-row and 72213-column matrix. We kept 200 excess rows, to have a reasonable guarantee that one of the dependencies would be useful. It took slightly more than one day to actually build the small matrix and to verify that all entries in the sparse and eliminated part were indeed zero. The small matrix turned out to be entirely dense. In the small matrix we included at regular intervals rows that consisted of the sum (modulo 2) of all previous rows, thus creating several spurious but predictable dependencies.

We immediately set out to reduce this "small" matrix, using ordinary Gaussian elimination and our familiar set-up at DEC SRC. This time, however, we had some protection against bit errors: if one of the spurious dependencies failed to show up, something must have gone wrong recently. Then we could back up a few hundred rows, and restart the elimination from a point where we were confident that everything was still correct. We estimate that the entire elimination on this single workstation would have taken less than seven weeks.

While this process was making its slow progress, the third author, tired of keeping it alive and not too confident of its outcome, contacted Roger Frye and Mike McKenna at Thinking Machines, and explained the problem to them.

After a short while they had written a Gaussian elimination program for a Connection Machine. They estimated that their program, when executed on a 65536-processor Connection Machine, could handle our 72000-matrix within three hours. Jim Hudgens and George Marsaglia at the Supercomputer Computation Research Institute at Florida State University arranged the computer time we needed. We sent a box with ten tapes containing the data for the matrix by Federal Express to Florida. Jim Hudgens consolidated these ten tapes into one "exotape". During the evening of June 14 he mounted the exotape, so that Roger Frye and Mike McKenna, remotely logged in from Thinking Machines in Cambridge, Massachusetts, could read the data as one large sequential file, and execute the program. It solved the system in three hours, but then a crash occurred, due to a mistake in the output routine. The second run, which again took three hours, produced a few hundred dependencies among the rows of the dense 72000-matrix.

In the early morning of June 15, 1990, the dependencies were sent, electronically, to DEC SRC, where they were converted into dependencies of the original sparse 200000-matrix. At least, that is what we hoped that they would turn out to be. At 9:15 PDT we started our final program, the attempt to factor $n$ by processing the dependencies sequentially until the factorization was found. This led to the most exciting moment of the entire factorization of $F_9$: at 9:45 PDT the program concluded that the first alleged dependency among the rows of the sparse 200000-matrix was a true one. This moment of great relief could not be spoilt by the sobering message, displayed at 10:15 PDT, that the first dependency had just given rise to the trivial factorization of $n$. An hour later, at 11:15 PDT (18:15 GMT), the second dependency proved to be luckier by finding a 49-digit factor. Both this factor and the 99-digit cofactor were announced prime, because no witnesses to their compositeness could be found among five randomly chosen integers (see §2).

Five minutes later the backup Gaussian elimination process, still crunching along on a single workstation, was terminated, five days short of its goal. Still on June 15, Andrew Odlyzko used the first author's Cray X-MP implementation of the Jacobi sum primality test [12, 13] to prove that both factors were indeed prime.

## Acknowledgments

We would like to thank everyone who helped with collecting relations: Kannan Alagappan, Steve Albrecht, Sharon Allan, Dean Alvis, Jon M. Arnold, Thomas G. Arnold, Amy Asbury, Elena Aschkenasi, Bob Ayers, Pete Balkus, Joel Bartlett, David M. Bernardo, Doug Bonin, Anita Borg, Wieb Bosma, Patrick Boyle, Richard P. Brent, John Brownie, Mark Buda, Joe Buhler, Bruce Bullis, Toni Burks, Al Butler, Neil J. Calkin, Don Carignan, Jerome Chailloux, George Chaltas, Chran-Ham Chang, Bruce Chase, Scott Chase, Carol Cignotti, Ian Clements, Donn Coe, Henri Cohen, Leisa Condie, Tom Coppeto, Jeff Cormier, Ken Cowan, Richard Critz, Richard L. Curtis, Stu Davidson, Chris DeLise, Bhupendra Desai, Srinivas Desirazu, Bob Devine, Jeff Diewald, Jeff Dike, John Dillon, Jeremy Dion, Paul DiTommaso, Dan Donahue, Jeff Donovan, Mike Doody, David W. Dunkle, Charles B. Dunn, Gene Durso, John Dustin, Charles

Dyer, Alan Eustace, Marc Evans, Mike Ferrara, Ed Flecchia, John Forecast, Joel Foss, Chris Franklin, Sandy Fraser, Hania Gajewska, Sachin Galgalikar, Edward A. Gardner, Eran Gartner, Morrie Gasser, Eric Gauthier, Lance J. Gay, Carl Gerstle, David Gibson, Stephen Gildea, Andy Goldstein, Mike Greenfield, Tim Greenwood, Liz Guth, Ramsey Haddad, Kenneth J. Haduch, Lori Hagen, John C. Hallyburton, Jr., Gary Harding, Bob Harris, Charles Haynes, B. J. Herbison, Ted Hess, Max Hillson, Buren Hoffman, Jim Horning, Felix S. Hsu, Han Hsu, Scott Huddleston, Gary Huff, Peter Ilieve, Frank Infante, Philippe Jacquet, Jeff Janock, Jeff Jenkins, Mike Johnson, Kevin Jones, Dave Juitt, Bill Kalsow, Irving Kaplansky, Harsh Kapoor, Philip L. Karlton, Bill Katz, Christopher A. Kent, Jeff Kenyon, Alan Kirby, Manfred Koethe, John T. Kohl, David M. Kuchta, Rick Landau, Sundaram Laxman, Ted Lemon, Bill Licea-Kane, John Linn, Walter Lioen, Todd Little, Jim Lo, Mark Longo, Kevin Lynch, Pat Madden, Joseph A. Martin, Robert N. Mayo, Murray S. Mazer, James T. McCartney, Joel McCormack, Ellen McDermott, Randall V. Meyers, Michael J. Miano, Steven Miller, Thomas Mitchell, Jeffrey Mogul, Bruce Moore, François Morain, A. E. Mossberg, David Mostardi, Victoria Murphy, Gopal Nagarajan, Jeff E. Nelson, Chuck Newman, Marc Nozell, Vinay Nulkar, Paul E. Oppenheimer, Lisa Palermo, Bill Parke, Tom Patterson, Eileen Perez, Don Pettini, Nigel Poole, Eric D. Postpischil, Edward G. Prentice, Dennis Racca, Ramgopal Ramgiri, Ram Rao, Jon Reeves, Brian Reid, August G. Reinig, Herman te Riele, John Riley, Buzzy Ritter, John Robinson, D. C. Rocks, David M. Rosenberg, Eduardo Santiago, Sanjay Saxena, Richard Schedler, Jeffrey I. Schiller, Michael Sclafani, Jeff Sebring, Mike Sekurski, Shekhar Sengupta, Mark Shand, Robert Shen, our competitor, John Simakauskas, Al Simons, Michael Soha, Kiran Somalwar, Bill Sommerfeld, Bob Souza, Vivian Sovinsky, Jerry Stange, Alan Stanier, John T. Stapleton, Jorge Stolfi, Geof Stone, Steve Strange, Richard Swan, Ed Taranto, Pahtiban Thilagar, Benjamin J. Thomas III, Bob Thomas, Gary Thomas, Mathews Thomas, Dennis Ting, Ward Travis, Win Treese, Tom Truscott, Jim Turner, Percy Tzelnic, Bob Unnold, Srinivasa Uppugunduri, Mohan Vaghul, Virendra Verma, Brick Verser, Paul Vixie, David Wall, David F. Wall, Chuck Wan, Bradley M. Waters, Dave Weiss, Dan White, Bill Whitney, Dick Wilkins, Dik Winter, Ted Wobber, Frank M. Woodall, Jr., Tom Woodburn, John Wray, Frank Zereski, Paul Zimmermann, John Zornig, plus many people at Bellcore and MSRI, at DEC's research laboratories CRL, PRL, SRC, and WRL, and at WSE and WSL. This list is incomplete, since some of our contributors were known to us only as electronic addresses that were no longer in service when we tried to get their names. We apologize for any omissions and misspellings.

We are also grateful to the people who helped with Gaussian elimination: Andrew Odlyzko and Carl Pomerance for theoretical assistance, Jim Hudgens, George Marsaglia and the Supercomputer Computation Research Institute at Florida State University for computing time on the Connection Machine, and Roger Frye and Mike McKenna at Thinking Machines for writing and running a Gaussian elimination program on the Connection Machine.

The Mathematical Sciences Research Institute in Berkeley is gratefully acknowledged for giving us the opportunity to meet with various people involved in this project. We also thank our competitor for spurring us on, and Richard Brent, Richard Crandall, Wilfrid Keller, Robert Silverman, and Sam Wagstaff for providing information concerning previous $F_9$ factoring attempts and factors of Fermat numbers in general. We thank Cynthia Hibbard and Lyle Ramshaw for their comments on an earlier version of this paper.

The first author would like to thank DEC SRC and Bob Taylor for their hospitality and support in the summer of 1989, when most of the programming for the implementation of the number field sieve was done. The second author is grateful to the Institute for Advanced Study in Princeton and the Université de Franche-Comté in Besançon for their hospitality and support while this paper was being written. The second author was supported by NSF under Grant No. DMS 90-02939.

## Bibliography

1. G. E. Andrews, *The theory of partitions*, Addison-Wesley, Reading, MA, 1976.
2. A. O. L. Atkin and F. Morain, *Elliptic curves and primality proving*, Research report 1256, INRIA, June 1990.
3. R. P. Brent, *Some integer factorization algorithms using elliptic curves*, Austral. Comp. Sci. Comm. **8** (1986), 149–163.
4. ____, *Factorization of the eleventh Fermat number (preliminary report)*, Abstracts Amer. Math. Soc. **10** (1989), 89T-11-73.
5. ____, *Parallel algorithms for integer factorisation*, Number Theory and Cryptography (J. H. Loxton, ed.), London Math. Soc. Lecture Note Series, vol. 154, Cambridge Univ. Press, Cambridge, 1990, pp. 26–37.
6. R. P. Brent and J. M. Pollard, *Factorization of the eighth Fermat number*, Math. Comp. **36** (1981), 627–630.
7. 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*, 2nd ed., Contemp. Math., vol. 22, Amer. Math. Soc., Providence, RI, 1988.
8. J. A. Buchmann, *A generalization of Voronoĭ's unit algorithm*. II, J. Number Theory **20** (1985), 192–209.
9. J. A. Buchmann and H. W. Lenstra, Jr., *Approximating rings of integers in number fields*, in preparation.
10. J. P. Buhler, H. W. Lenstra, Jr., and C. Pomerance, *Factoring integers with the number field sieve* (to appear).
11. H. Cohen, *A course in computational algebraic number theory*, Springer-Verlag (to appear).
12. H. Cohen and A. K. Lenstra, *Implementation of a new primality test*, Math. Comp. **48** (1987), 103–121, S1–S4.
13. H. Cohen and H. W. Lenstra, Jr., *Primality testing and Jacobi sums*, Math. Comp. **42** (1984), 297–330.
14. J. H. Conway, *On numbers and games*, Academic Press, London, 1976.
15. A. Cunningham and A. E. Western, *On Fermat's numbers*, Proc. London Math. Soc. (2) **1** (1904), 175.
16. A. J. C. Cunningham and H. J. Woodall, *Factorisation of* $(y^n \mp 1)$, $y = 2, 3, 5, 6, 7, 10, 11, 12$ *up to high powers* $(n)$, Hodgson, London, 1925.
17. L. E. Dickson, *History of the theory of numbers*, vol. I, Carnegie Inst. of Washington, Washington, 1919.
18. L. Euler, *Observationes de theoremate quodam Fermatiano aliisque ad numeros primos spectantibus*, Comm. Acad. Sci. Petropol. **6** (1732/1733), 103–107; Leonhardi Euleri Opera Omnia, Ser. I, vol. II, Teubner, Leipzig, 1915, pp. 1–5.
19. C. F. Gauss, *Disquisitiones arithmeticae*, Fleischer, Leipzig, 1801.
20. A. Gérardin, *Méthodes de Landry*, L'intermédiaire des Mathématiciens **16** (1909), 199–201.

21. G. B. Gostin and P. B. McLaughlin, Jr., *Six new factors of Fermat numbers*, Math. Comp. **38** (1982), 645–649.

22. J. C. Hallyburton, Jr., and J. Brillhart, *Two new factors of Fermat numbers*, Math. Comp. **29** (1975), 109–112; corrigendum, ibid. **30** (1976), 198.

23. W. Keller, *Factors of Fermat numbers and large primes of the form* $k \cdot 2^n + 1$. II, submitted for publication.

24. B. A. LaMacchia and A. M. Odlyzko, *Solving large sparse linear systems over finite fields*, Advances in Cryptology, Proc. Crypto '90, Lecture Notes in Comput. Sci., vol. 537, Springer-Verlag, Berlin and New York, 1991, pp. 109–133.

25. F. Landry, *Note sur la décomposition du nombre* $2^{64} + 1$ (Extrait), C. R. Acad. Sci. Paris **91** (1880), 138.

26. A. K. Lenstra, *Factorization of polynomials*, pp. 169–198 in [33].

27. A. K. Lenstra and H. W. Lenstra, Jr., *Algorithms in number theory*, Handbook of Theoretical Computer Science, Volume A, Algorithms and Complexity (J. Van Leeuwen, ed.), Elsevier, Amsterdam, 1990, Chapter 12.

28. A. K. Lenstra, H. W. Lenstra, Jr., M. S. Manasse, and J. M. Pollard, *The number field sieve*, (to appear). Extended abstract: Proc. 22nd Annual ACM Sympos. Theory of Computing (STOC) (Baltimore, May 14–16, 1990), pp. 564-572.

29. A. K. Lenstra and M. S. Manasse, *Factoring by electronic mail*, Advances in Cryptology, Eurocrypt '89, Lecture Notes in Comput. Sci., vol. 434, Springer-Verlag, Berlin and New York, 1990, pp. 355–371.

30. ——, *Factoring with two large primes*, Math. Comp. (to appear).

31. H. W. Lenstra, Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), 649–673.

32. H. W. Lenstra, Jr., and C. Pomerance, *A rigorous time bound for factoring integers*, J. Amer. Math. Soc. **5** (1992), 483–516.

33. H. W. Lenstra, Jr., and R. Tijdeman (eds.), *Computational methods in number theory*, Math. Centre Tracts, no. 154/155, Mathematisch Centrum, Amsterdam, 1983.

34. P. L. Montgomery, *Speeding the Pollard and elliptic curve methods of factorization*, Math. Comp. **48** (1987), 243–264.

35. P. L. Montgomery and R. D. Silverman, *An FFT extension to the* $P-1$ *factoring algorithm*, Math. Comp. **54** (1990), 839–854.

36. M. A. Morrison and J. Brillhart, *A method of factoring and the factorization of* $F_7$, Math. Comp. **29** (1975), 183–205.

37. T. Pepin, *Sur la formule* $2^{2^n} + 1$, C. R. Acad. Sci. Paris **85** (1877), 329–331.

38. C. Pomerance, *Analysis and comparison of some integer factoring algorithms*, pp. 89–139 in [33].

39. C. Pomerance and J. W. Smith, *Reduction of huge, sparse matrices over finite fields via created catastrophes*, Experiment. Math. **1** (1992), 89–94.

40. M. O. Rabin, *Probabilistic algorithm for testing primality*, J. Number Theory **12** (1980), 128–138.

41. H. Riesel, *Prime numbers and computer methods for factorization*, Birkhäuser, Boston, 1985.

42. RSA Data Security Corporation, Inc., sci.crypt, May 18, 1991; information available by sending electronic mail to challenge-rsa-list@rsa.com.

43. C. P. Schnorr and H. W. Lenstra, Jr., *A Monte Carlo factoring algorithm with linear storage*, Math. Comp. **43** (1984), 289–311.

44. D. Shanks, *On Gauss and composition*. I, Number theory and applications, NATO Adv. Sci. Inst. Ser. C: Math. Phys. Sci. **265**, Kluwer, Dordrecht, 1989, see p. 174.

45. R. D. Silverman, *The multiple polynomial quadratic sieve*, Math. Comp. **48** (1987), 329–339.

46. I. N. Stewart and D. O. Tall, *Algebraic number theory*, Chapman and Hall, London, 1979.

47. P. Tannery and C. Henry (eds.), *Oeuvres de Fermat*, vol. II, *Correspondance*, Gauthier-Villars, Paris, 1894.

48. A. Weil, *Number theory: an approach through history*, Birkhäuser, Boston, 1983.

49. D. Wiedemann, *An iterated quadratic extension of* GF(2), Fibonacci Quart. **26** (1988), 290–295.

50. H. C. Williams, *How was $F_6$ factored?*, Math. Comp. **61** (1993), 463–474.

Bellcore, Room 2Q334, 445 South Street, Morristown, New Jersey 07960
*E-mail address*: lenstra@bellcore.com

Department of Mathematics, University of California, Berkeley, California 94720
*E-mail address*: hwl@math.berkeley.edu

DEC SRC, 130 Lytton Avenue, Palo Alto, California 94301
*E-mail address*: msm@src.dec.com

Tidmarsh Cottage, Manor Farm Lane, Tidmarsh, Reading, Berkshire, RG8 8EX, England