THE NUMBER FIELD SIEVE ON MANY COMPUTERS

R.-M. Elkenbracht-Huizing, Peter L. Montgomery, R. D. Silverman,
R. K. Wackerbarth and S. S. Wagstaff, Jr.

Abstract. We briefly describe the special number field sieve integer factoring algorithm, emphasizing the polynomial selection, and tell how we have used it to factor large integers on many workstations.

1. The Number Field Sieve Algorithm

We could factor the odd positive integer n which is not a prime power if we could find integers x, y so that $x^2 \equiv y^2 \bmod n$ but $x \not\equiv \pm y \bmod n$. The first congruence implies that n divides $(x-y)(x+y)$. The second congruence implies that n does not divide $x-y$ or $x+y$. Hence, at least one prime factor of n divides $x-y$ and at least one prime factor of n does not divide $x-y$. Therefore, $\gcd(n, x-y)$ is a proper factor of n.

In the Continued Fraction Method and the Quadratic Sieve Method, many congruences (relations) of the form $a^2 \equiv q \bmod n$ are produced with q factored completely. Linear algebra (over GF(2)) is used to match up the prime factors of q to find a subset of the relations in which the product of the q's is a square, y^2, say. Let x be the product of the a's in these relations. Then $x^2 \equiv y^2 \bmod n$.

The (Special) Number Field Sieve (NFS) factors numbers of the form $n = r^e - s$, where r and $|s|$ are small positive integers. Actually, the Special NFS can be applied to numbers of the form $ar^e + bs^j$, and not just to $r^e - s$. Let $n = ar^e + bs^j$. If e and j are about equal, then $s^{-j}n = ar^{e-j}(r/s)^j + b$. This is a polynomial in (r/s) and can be treated in somewhat the same way as $r^e - s$.

The basic form works for $n = r^e - s$ as follows:

Choose a small positive integer d, the degree of an extension field. Let k be the least positive integer for which $kd \ge e$. Let $t = s \cdot r^{kd-e}$. Let f be the polynomial $X^d - t$. Let $m = r^k$. Then $f(m) = r^{kd} - s \cdot r^{kd-e} = r^{kd-e}n$ is a multiple of n.

Let α be a zero of f. Let $K = \mathbb{Q}(\alpha)$. We assume f is irreducible, else we could use its factorization to factor n. The degree of K over $\mathbb{Q}$ is d. Let Q_n denote the ring of rational numbers with denominator coprime to n. The subring $Q_n[\alpha]$ of K consists of expressions $\sum_{i=0}^{d-1}(s_i/t_i)\alpha^i$ with $s_i, t_i \in \mathbb{Z}$ and $\gcd(n, t_i) = 1$. Define a

1991 *Mathematics Subject Classification.* Primary 11A51, 11Y05; Secondary 11R21, 11Y40.

Key words and phrases. Number Field Sieve, factoring integers.

The computing reported in this work was partially supported by ARPA grant F30602-96-1-0334.

Typeset by $\mathcal{A}\mathcal{M}\mathcal{S}$-TeX

ring homomorphism $\phi : Q_n[\alpha] \to \mathbb{Z}/n\mathbb{Z}$ by the formula $\phi(\alpha) = (m \bmod n)$, so that $\phi(\sum_{i=0}^{d-1}(s_i/t_i)\alpha^i) = (\sum_{i=0}^{d-1} s_i t_i^{-1} m^i \bmod n)$ for $s_i, t_i \in \mathbb{Z}$ and $\gcd(n, t_i) = 1$.

In variations of the basic form, other polynomials may be used. The key properties required of the polynomial are that it is irreducible, that it has moderately small coefficients so that the norm of α is small, and that we know a non-trivial root m modulo n of it. The optimal degree for f is $((3+o(1))\log n/(2\log\log n))^{1/3}$ as $e \to \infty$ uniformly for r, s in a finite set. (See [4].) This optimal degree is about 5 for n of the size we have considered, which is between 100 and 160 digits.

For $0 < a \le A$ and $-B \le b \le B$, NFS uses a sieve to find factors of $a + bm$ and the norm of $a + b\alpha$ in $\mathbb{Z}$.

A pair (a, b) is saved in a file if a and b are relatively prime, $a + bm$ is smooth (has only relatively small prime factors), and the norm of $a + b\alpha$ is smooth. The norm of $a + b\alpha$ is $(-b)^d f(-a/b)$.

Using linear algebra, one finds a non-empty set S of pairs (a, b) of relatively prime integers such that

$$\prod_{(a,b)\in S} (a + bm) \text{ is a square in } \mathbb{Z},$$

and

$$\prod_{(a,b)\in S} (a + b\alpha) \text{ is a square in } Q_n[\alpha].$$

Let the integer x be a square root of the first product. Let $\beta \in Q_n[\alpha]$ be a square root of the second product. We have $\phi(\beta^2) \equiv x^2 \bmod n$ since $\phi(a + b\alpha) \equiv a + bm \bmod n$.

Let y be the integer for which $\phi(\beta) \equiv y \bmod n$. Then $x^2 \equiv y^2 \bmod n$, which gives us a chance to factor n.

See [4] and the articles in [3] for a more complete description of the number field sieve algorithm.

2. Some Examples

These examples illustrate the NFS algorithm applied to numbers of the form $r^e \pm 1$ from the book [1]. The individuals performing these computations are called NFSNET.

Example 1: The number n to factor is a divisor of $6^{199} - 1$. We assume $n = 6^{199} - 1$. Let $d = 5$, $f(X) = X^5 - 6$ and $m = 6^{40}$. Then $f(m) = 6n$ and $\alpha = 6^{1/5}$ is a zero of f. The number field is $K = \mathbb{Q}(\alpha)$. The degree $[K : \mathbb{Q}] = 5$ since f is irreducible over $\mathbb{Q}$.

The next four examples show how one can choose better polynomials in certain cases.

Example 2: The number n to factor is a divisor of $10^{158}+1$. Assume $n = 10^{158}+1$. Note that $100n = 10^{160} + 100 = (2^5)(5 \cdot 10^{31})^5 + 100$. Hence $25n = 8m^5 + 25$, where $m = 5 \cdot 10^{31}$. Let $f(X) = 8X^5 + 25$. Then $\alpha = (-25/8)^{1/5}$ is a zero of f and m is a zero of f modulo n. The number field is $K = \mathbb{Q}(\alpha)$. The degree $[K : \mathbb{Q}] = 5$ since f is irreducible over $\mathbb{Q}$. The $m = 5 \cdot 10^{31}$ is half what it would have been if we had used the polynomial $f(X) = X^5 + 100$. The coefficients of $8X^5 + 25$ are smaller than those of $X^5 + 100$ and consequently $a + b\alpha$ has a slightly smaller norm. Thus $a + bm$ and the norm of $a + b\alpha$ are smaller and more likely to be smooth than if we had used the obvious polynomial. This trick can be used for any composite base r, such as 6, 10 or 12, but clearly not for a prime base. It was not used in Example 1 because $X^5 - 6$ already has small coefficients.

Example 3: The number n to factor is a divisor of $(7^{187} - 1)/(7^{17} - 1)$. Note that $187 = 11 \cdot 17$. If we assumed that $n = 7^{187} - 1$, then n would be larger and we would have to do more sieving than if we assumed that $n = (7^{187} - 1)/(7^{17} - 1)$. But this n does not have the form $r^e - s$ and we must work harder to find a suitable polynomial and root. Write $k = 7^{17}$. Then $n = (k^{11} - 1)/(k - 1) = g(k)$, where $g(X) = X^{10} + X^9 + \cdots + X + 1$. Degree 10 is too large. We attempt to reduce the degree to about 5. Factor an X^5 from g:

$$g(X) = X^5(X^5 + X^4 + X^3 + X^2 + X + 1 + X^{-1} + X^{-2} + X^{-3} + X^{-4} + X^{-5}).$$

Since the expression in parentheses is unchanged when X is replaced by X^{-1}, it can be written as a polynomial in $X + X^{-1}$. One computes that this polynomial is

$$f(X) = X^5 + X^4 - 4X^3 - 3X^2 + 3X + 1.$$

Then $X^5 f(X + X^{-1}) = g(X)$. Let $m = k + k^{-1} \bmod n$. The numerator and denominator of m are small and $k^5 f(m) \equiv g(k) \equiv 0 \bmod n$. Since $\gcd(k, n) = 1$, we have $f(m) \equiv 0 \bmod n$. Let α be a zero of f. As f is irreducible over $\mathbb{Q}$, the degree $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 5$. When a and b are small, the norm of $a + b\alpha$ is near 7^{34} which is 7^3 smaller (and more likely to be smooth) than if we had used $n = 7^{187} - 1$, $f(X) = 49X^5 - 1$ and $m = 7^{37}$.

Example 4: The number n to factor is a divisor of $(2^{559} - 1)/(2^{43} - 1)$. Note that $559 = 13 \cdot 43$. If we assumed that $n = 2^{559} - 1$, then n would be larger and we would have to do more sieving than if we assumed that $n = (2^{559} - 1)/(2^{43} - 1)$. To find a suitable polynomial, write $k = 2^{43}$. Then $n = (k^{13} - 1)/(k - 1) = g(k)$, where $g(X) = X^{12} + X^{11} + \cdots + X + 1$. Since $X^{-6} g(X)$ is unchanged when X is replaced by X^{-1}, it can be written as a polynomial in $X + X^{-1}$. One finds that this polynomial is

$$f(X) = X^6 + X^5 - 5X^4 - 4X^3 + 6X^2 + 3X - 1.$$

Then $X^6 f(X + X^{-1}) = g(X)$. Let $m = k + k^{-1} \bmod n$. Then $k^6 f(m) \equiv g(k) \equiv 0 \bmod n$. Since $\gcd(k, n) = 1$, we have $f(m) \equiv 0 \bmod n$. Let α be a zero of f. As f is irreducible over $\mathbb{Q}$, the degree $[\mathbb{Q}(\alpha) : \mathbb{Q}] = 6$. When a and b are small, the norm of $a + b\alpha$ is near 2^{86} which is 2^{25} smaller (and more likely to be smooth) than if we had used $n = 2^{559} - 1$, $f(X) = 16X^5 - 1$ and $m = 2^{111}$. The f in this example was the first sextic used by NFSNET.

Example 5: The number to factor is a divisor of $5^{505} - 1$. Of course this is $X^5 - 1$ for $X = 5^{101}$. The polynomial factors as $X - 1$ times a biquadratic polynomial. Both factors are irreducible over the integers. However, when $X = 5^{5h}$, where h is odd, the value of the biquadratic polynomial splits into two nearly equal factors. This *Aurifeuillian* factorization (see III C 2 of [1]) is

$$5^{5h} - 1 = (5^h - 1)L_{5h}M_{5h},$$

where

$$L_{5h}, M_{5h} = 5^{2h} + 3 \cdot 5^h + 1 \mp 5^{(h+1)/2}(5^h + 1).$$

The particular number of this example is a divisor of

$$n = M_{505} = 5^{202} + 5^{152} + 3 \cdot 5^{101} + 5^{51} + 1.$$

This suggests using NFS with the biquadratic polynomial

$$f(X) = 25X^4 + 25X^3 + 15X^2 + 5X + 1$$

and root $m = 5^{50}$. Indeed $n = f(m)$. We used the same polynomial f with the root $m = 5^{51}$ to factor a divisor of M_{515}.

3. The NFSNET Project

NFSNET is a collaboration of the following people to factor numbers by the Number Field Sieve.

Bob Silverman wrote the core of the sieving code. Richard Wackerbarth wrote the task distribution and result collection code, makes the assignments, and collects the relations. The sieving process is divided into many small tasks, each consisting of sieving $a + bm$ and the norm of $a + b\alpha$ for a in a short interval and for all b. Each participating computer requests a task from a central computer, performs it and sends the relations it finds to the central computer when it requests the next task.

Sam Wagstaff counts, checks and filters (that is, condenses) the relations. When enough relations have been collected, he tells Richard to start sieving the next number. Marije Elkenbracht-Huizing and Peter Montgomery do the linear algebra to solve the large matrix of equations and extract the square root in the algebraic number field. The code for the filtering, linear algebra and square root stages, mainly written by Peter Montgomery, is described extensively in [2].

The volunteer sievers run the sieve program, which takes most of the time, on their computers. They include: Leo Broukhis, Ed Buzzi, Damien Doligez, Oyvind Eilertsen, Lamont Granquist, Bill Hodgeman, James Howe, Matthew Jackson, Michel Kern, John Reiser, Harry J. Smith, Gene Stark, Ray Van Tassle, Richard Wackerbarth and Paul Zimmermann.

We currently have about 100 computers on three continents. We need more volunteer sievers. If you would like to join this elite group, please contact Richard Wackerbarth at `rkw@dataplex.net` to get the sieve program.

So far we have factored more than a dozen numbers, including the ones in the examples above. All of the numbers we have factored come from [1].

References

1. J. Brillhart, D. H. Lehmer, J. L. Selfridge, B. Tuckerman and S. S. Wagstaff, Jr., *Factorizations of $b^n \pm 1$, $b = 2, 3, 5, 6, 7, 10, 11, 12$ up to high powers*, Second edition, Contemporary Mathematics **22**, Amer. Math. Soc., Providence, 1988.
2. R.-M. Huizing, *An implementation of the number field sieve*, Technical report NM-R9511, Centrum voor Wiskunde en Informatica, Amsterdam, 1995, to appear in *Experimental Mathematics.*
3. A. K. Lenstra and H. W. Lenstra, Jr., *The development of the number field sieve*, Lecture Notes in Mathematics 1554, Springer-Verlag, Berlin, New York, 1993.
4. A. K. Lenstra, H. W. Lenstra, Jr., M. S. Manasse and J. M. Pollard, *The number field sieve*, Proceedings 22nd Annual ACM Symposium on Theory of Computing (STOC), Baltimore, 1990, pp. 564–572.

CWI, Kruislaan 413, 1098 SJ Amsterdam, The Netherlands
E-mail address: marije@cwi.nl

780 Las Colindas Road, San Rafael, CA 94903-2346
E-mail address: pmontgom@cwi.nl

24 Standish Drive, Canton, MA 02021
E-mail address: bobs@sappho.rl.af.mil

8801 Camelia Lane, Austin, TX 78759-7510
E-mail address: rkw@dataplex.net

Department of Computer Sciences, Purdue University, West Lafayette, Indiana 47907-1398
E-mail address: ssw@cs.purdue.edu