MATHEMATICS OF COMPUTATION
VOLUME 59, NUMBER 200
OCTOBER 1992, PAGES 717–722

# IRREGULAR PRIMES TO ONE MILLION

J. P. BUHLER, R. E. CRANDALL, AND R. W. SOMPOLSKI

ABSTRACT. Using "fast" algorithms for power series inversion (based on the fast Fourier transform and multisectioning of power series), we have calculated all irregular primes up to one million, including their indices of irregularity and associated irregular pairs. Using this data, we verified that Fermat's "Last Theorem" and Vandiver's conjecture are true for these primes. Two primes with index of irregularity five were already known; we find that there are nine other primes less than one million with index five and that the prime 527377 is the unique prime less than one million with index six.

A pair of integers $(p, k)$ is said to be an irregular pair if $p$ is a prime, $k$ is an even integer satisfying $2 \leq k \leq p-3$, and $p$ divides the numerator of the Bernoulli number $B_k$. Irregular pairs have been computed by Vandiver, D. H. Lehmer, E. Lehmer, Selfridge, Nicol, Pollack, Johnson, Wada, Wagstaff, and Tanner (see [10] and references therein). The most recent tabulations described in [11] cover all primes $p < 150000$. The purpose of this paper is to describe the computation of all irregular pairs for $p < 10^6$.

These calculations have several well-known applications. In all cases known so far, the list of all irregular pairs of $p$ enables one to verify Fermat's "Last Theorem" (FLT) for the prime $p$. The technique for doing this originates with Vandiver (see, e.g., [12]); using these ideas, we find that FLT is true for all $p < 10^6$. Note also that recent ideas of Frey, Serre, and Ribet (see [6]) provide evidence for the truth of FLT; specifically, Ribet shows that the Taniyama-Weil conjectures on elliptic curves imply FLT.

The tabulation of the irregular pairs also allows one to verify Vandiver's conjecture, for which there seems to be little theoretical evidence one way or the other. Vandiver's conjecture asserts that $p$ does not divide the class number $h^+(p)$ of the totally real subfield of the cyclotomic field generated by the $p$th roots of unity. For small $p$ this conjecture is true for the trivial reason that $h^+(p) < p$; however, examples of $h^+(p) > p$ are known (see [8]). Our calculations show that Vandiver's conjecture is true for all $p < 10^6$.

The table of irregular pairs could also be used to calculate Iwasawa invariants for the corresponding primes. We did not do this as part of our calculations; see [9] or [4] for a discussion of this problem.

Previous computations of irregular pairs have used algorithms that take $O(p^2)$ arithmetic operations for each prime $p$. The Bernoulli numbers are defined, in

Received by the editor April 30, 1991 and, in revised form, September 10, 1991.
1991 *Mathematics Subject Classification.* Primary 11B68; Secondary 11D41, 65T20, 11Y99.
The first author was partially supported by National Science Foundation Grant DMS-9012989.

the even suffix notation, by the formal power series

$$\frac{x}{e^x - 1} = \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}.$$

To compute irregular pairs for a prime $p$, the coefficients of this power series must be computed modulo $p$ out to $k = p - 3$. Our basic idea was to use fast Fourier transform (FFT) algorithms to compute the power series inverse in time $O(p \log(p))$ (a related, though slightly less efficient, idea is described in [1]). In order to minimize running time and memory requirements, it was necessary to modify this initial idea by multisectioning the power series (as described below) and fine-tuning the underlying FFT algorithm. The computations were performed for $p < 10^6$ on a network of NeXT workstations [2, 3] and to $p < 400,000$ on an IBM 3090 at the Cornell National Supercomputing Facility (CNSF). In addition to checking the results by implementing several programs on different machines, we implemented some stringent internal consistency checks to verify the calculations of the Bernoulli numbers.

With our current implementation it is possible to find all irregular pairs for primes near $10^7$. For example, we find that $p = 8,388,019$ is regular, presumably the largest explicit regular prime known. Incidentally, the challenge of finding large irregular primes is qualitatively quite different. For example, the numerator of $B_{118}$ is of the form $59q$, where $q$ is a 100-digit prime, so $(q, 118)$ is an irregular pair.

## 1. Algorithms

Fix a prime $p$. In order to compute the Bernoulli numbers $B_k$ for $k \leq p-3$, we need to compute the inverse of the power series $f = (e^x - 1)/x$ out to the $x^{p-3}$-term, or, as we shall say, out to $O(x^{p-2})$. All computations of the coefficients need to be done modulo $p$. If $g$ is an approximation to $f^{-1}$, then the standard Newton iteration for taking reciprocals (see [5]) says that $h = 2g - fg^2$ is a better approximation; more precisely, if $g - f^{-1} = O(x^n)$ then $h - f^{-1} = O(x^{2n})$. Thus, the inverse of $f$ to $O(x^{p-3})$ can be computed in $O(\log(p))$ Newton iterations, each requiring three polynomial multiplications with polynomials whose degree doubles at each iteration. Only the final iteration is done to full precision and it is easy to see that the total time is bounded by a constant times the cost of the final polynomial multiplication. (All of our times are measured in arithmetic operations, but the arithmetic is being done on integers of size $p$, so the bit complexity only differs from these by a power of $\log(p)$; in our actual computations all primes fit in a word, and so an arithmetic operation effectively took constant time.)

Two polynomials of degree $k$ were multiplied by the usual device of padding the polynomial with zeros to the next higher power of two, $K = 2^a$, and then using the FFT (see [5]). Both integer and floating-point FFT algorithms were implemented.

In the integer version, the FFT was done modulo two primes, each larger than $2^{30}$ and each congruent to 1 modulo $K$; the results were combined by Chinese remainder techniques. (Note that two 30-bit primes are sufficient to determine the product over the integers, since the coefficients of the product of two polynomials of degree $K$ with coefficients bounded by $p$ is bounded

by $pK^2$, which is bounded by $p^3$ and hence $2^{60}$.) The primes were chosen to be congruent to 1 modulo $K$, so that the appropriate roots of unity existed modulo those primes; no convolution algorithm working directly in the field of $p$ elements seemed competitive.

In the floating-point implementations of the FFT the results were rounded to the nearest integer and then reduced modulo $p$. Considerable accuracy was gained by using a "balanced" representation of the coefficients in which the coefficients were kept in the interval $[-(p-1)/2, (p-1)/2]$. This assured that the sequences being convolved were generally bipolar, so that the floating-point error of the FFT floating-point convolution is markedly less than when the coefficients are in the usual interval $[0, p-1]$. Careful estimates of the error together with experimental confirmation convinced us that IEEE standard 64-bit floating-point arithmetic gave sufficient precision to handle primes $p < 4 \cdot 10^6$. (The isolated calculations, referred to above, for primes up to $10^7$ were done by also computing the Bernoulli numbers modulo $2^{16}$ to compensate for the slightly insufficient precision in the floating-point computation; this doubled the running time.)

The inversion of the power series

$$f(x) = (e^x - 1)/x = 1 + x/2! + x^2/3! + \cdots + x^{p-3}/(p-2)! + O(x^{p-2})$$

modulo $p$ can be simplified by multiplying the coefficients through by $(p-2)!$ to avoid computing inverses modulo $p$. A more significant savings can be achieved by using the identity

$$\frac{x^2}{\cosh(x) - 1} = -2 + \sum_{n=0}^{\infty} \frac{(2n-1)B_{2n}}{(2n)!} x^{2n}$$

(which can be proved by differentiating the expression for $(f(x)^{-1} + f(-x)^{-1})/x$ in terms of the Bernoulli numbers). Since this is a power series in $x^2$, the necessary Bernoulli numbers can be computed by a power series merely of length about $p/2$. In order to run our programs on workstations with limited memory, it turned out to be important to extend this multisectioning idea [7, p. 132] even further. We ended up multisectioning by eighth roots of unity to find the identities

$$\frac{x}{\sqrt{2}} \frac{\cosh(x/\sqrt{2})}{\sinh(x/\sqrt{2})} = \sum_{n=0}^{\infty} \frac{2^n B_{2n}}{(2n)!} x^{2n} = \frac{A_0(x) + A_2(x) + A_4(x) + A_6(x)}{D(x)},$$

where the power series $A_i$ and $D$ have coefficients

$$A_k(x) = \sum_{n=0}^{\infty} \frac{A_{kn}}{(8n+k+3)!} x^{8n+k}, \qquad D(x) = \sum_{n=0}^{\infty} \frac{D_n}{(8n+4)!} x^{8n}$$

that are completely defined by their initial two terms and a recursion that applies to the coefficients $c_n$ of *all* five power series:

$$c_{n+1} = -136c_n - 16c_{n-1}, \qquad n \geq 1,$$
$$D_0 = 24, \qquad D_1 = -3168,$$
$$A_{00} = 6, \quad A_{01} = -792, \quad A_{20} = 20, \quad A_{21} = -2704,$$
$$A_{40} = -28, \quad A_{41} = 3824, \quad A_{60} = 96, \quad A_{61} = -13056.$$

(Note that $4A_{0n} = D_n$, which slightly simplifies the computations.) The point of multisectioning is that the only polynomial inversion required is the inversion of $D(x)$ to $O(x^{8m})$, where $m = [(p-5)/8]$. Since $D(x)$ is a power series in $x^8$, this means that we are inverting a polynomial of length approximately $p/8$. One curious feature of this multisectioning is that the factorial in the denominator of the term in $D(x)$ required to compute the last Bernoulli number $B_{p-3}$ is not relatively prime to $p$. Rather than trying to find other multisectioning identities, we opted for computing $B_{p-3} \bmod p$ using the Voronoi identity (relation (1) in [10]). This did not appreciably increase the running time, since this last Bernoulli number can be calculated modulo $p$ in time $O(p)$.

In the computations carried out at CNSF the multisectioning was done using ninth roots of unity. Thus, there were nine rational functions with a common denominator; the polynomials were more complicated than those given above, and the recursion relations were considerably more complicated (requiring extended-precision integer constants). The Bernoulli numbers $B_{p-7}$, $B_{p-5}$, and $B_{p-3}$ were all "inaccessible" as above, and had to be computed separately.

Multisectioning provided a considerable gain not only in efficiency, but also in memory requirements. In addition to requiring less memory for the series being inverted, the $A_i$ can be computed serially, so that space is only needed for one $A_i$ array. If this technique were to be extended, the time requirements for inversion could be decreased still further, although the recurrence relations would become more complicated and further inaccessible coefficients would be introduced at the tail end of the power series.

Any large-scale computation is vulnerable to all sorts of errors, both in software and hardware. We checked the Bernoulli numbers by empirically verifying, for each $p$, the identity

$$\sum_{n=0}^{p-3} 2^n (n+1) B_n \equiv -4 \pmod{p}.$$

(One way to prove the identity is to find the coefficient of $x^{p-2}$ in

$$\sum_{k=0}^{\infty} B_k \frac{x^k}{k!} = \frac{x}{e^x - 1} = \frac{2x}{e^{2x} - 1} \frac{e^x + 1}{2} = \left( \sum_{n=0}^{\infty} \frac{2^n B_n x^n}{n!} \right) \left( 1 + \frac{1}{2} \sum_{m=1}^{\infty} \frac{x^m}{m!} \right)$$

and to use Wilson's Theorem to find that, for $n$ even, $n!(p-2-n)! \equiv 1/(n+1) \bmod p$.) This identity turned out to be especially useful; it enabled us to catch bugs in our programs, and to detect hardware errors (faulty memory on computers in the distributed network) that had not been otherwise noticed.

## 2. Results

The computation of irregular pairs was done for all $p < 10^6$. On a single 68040-based NeXT station it takes our implementations about 200 seconds to compute all relevant $B_k$ for a prime near $10^6$. The Vandiver criterion [10, 12] was applied to all irregular pairs, and, as in the earlier computations up to 150000, FLT and Vandiver's conjecture were always found to be true.

The largest index of irregularity was six; the unique prime with irregularity six is $p = 527377$. In addition to the two primes of index five already known [11], nine further primes of index five were found. The indices for all primes of index greater than or equal to five are given in Table 1.

TABLE 1. Irregular primes of index 5 or 6

| $p$ | $k$ such that $p$ divides $B_k$ |
|---|---|
| 78233 | 10400, 32084, 46620, 47364, 64628 |
| 94693 | 11636, 54754, 76326, 80650, 84726 |
| 162791 | 5374, 55866, 91758, 113422, 148008 |
| 334183 | 71956, 147746, 185584, 249484, 269172 |
| 432749 | 58230, 106152, 118198, 226438, 381994 |
| 527377 | 45740, 121620, 275372, 329694, 405590, 427078 |
| 675823 | 41770, 240886, 303428, 407948, 532058 |
| 679519 | 3040, 114300, 305012, 442932, 526062 |
| 700643 | 57626, 77204, 272956, 349742, 367798 |
| 731593 | 11824, 110020, 161232, 195510, 303270 |
| 754969 | 107012, 200390, 444842, 629186, 708078 |
| 845309 | 74094, 169160, 339356, 351774, 628666 |

TABLE 2. Irregularity index densities

| $r$ | $\pi_r(10^6)$ | $\pi_r(10^6)/\pi(10^6)$ | $e^{-1/2}/(2^r r!)$ |
|---|---|---|---|
| 0 | 47627 | 0.60673656 | 0.60653065 |
| 1 | 23816 | 0.30340012 | 0.30326532 |
| 2 | 5954 | 0.07585003 | 0.07581633 |
| 3 | 956 | 0.01217880 | 0.01263605 |
| 4 | 132 | 0.00168159 | 0.00157950 |
| 5 | 11 | 0.00014013 | 0.00015795 |
| 6 | 1 | 0.00001273 | 0.00001316 |

As has been noted by several people (see [9]), if the numerators of the Bernoulli numbers are uniformly random modulo odd primes, then the index of irregularity should satisfy a Poisson distribution with mean $1/2$; as a special case, noted by Lehmer and Siegel, the density of the irregular primes should be $1 - e^{-1/2}$. More generally, if $\pi(n)$ is the number of primes less than or equal to $n$, and $\pi_r(n)$ denotes the number of odd primes less than or equal to $n$ with index of irregularity $r$, then this would predict that $\pi_r(n)/\pi(n)$ is approximately equal to $e^{-1/2}/2^r r!$. The predicted and observed densities are tabulated in Table 2 (although it should be clear that the densities for $r \geq 3$ are based on far too little data to have any significance). The values truncated at 125000 were also computed and checked against the table in [10]; we found complete agreement.

## ACKNOWLEDGMENTS

The authors are indebted to Morris Meyer and Avadis Tevanian, Jr., of the Operating Systems Group at NeXT Computer, Inc., for their support in implementing and maintaining the network that performed most of our calculations.

The authors are also grateful to the Cornell National Supercomputing Facility and the San Diego Supercomputer Center for providing time on their computers.

## BIBLIOGRAPHY

1. M. Chellali, *Accélération de calcul de nombres de Bernoulli*, J. Number Theory **28** (1988), 347–362.
2. R. E. Crandall, *The NeXT computer as physics machine*, Computers in Physics **4** (1990), 132–141.
3. ——, *Mathematica for the Sciences*, Addison-Wesley, Reading, MA, 1991.
4. R. Ernvall and T. Metsänkylä, *Cyclotomic invariants for primes between* 125000 *and* 150000, Math. Comp. **56** (1991), 851–858.
5. D. Knuth, *The art of computer programming*, vol. 2, Addison-Wesley, Reading, MA, 1981.
6. K. Ribet, *On modular representations of* $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ *arising from modular forms*, Invent. Math. **100** (1990), 431–476.
7. J. Riordan, *Combinatorial identities*, Wiley, New York, 1968.
8. R. Schoof and L. Washington, *Quintic polynomials and real cyclotomic fields with large class number*, Math. Comp. **50** (1989), 543–556.
9. R. Sompolski, *The second case of Fermat's last theorem for fixed irregular prime exponents*, Ph.D. thesis, University of Illinois at Chicago, 1991.
10. S. S. Wagstaff, Jr., *The irregular primes to* 125000, Math. Comp. **32** (1978), 583–591.
11. S. S. Wagstaff, Jr. and J. W. Tanner, *New congruences for Bernoulli numbers*, Math. Comp. **48** (1987), 340–350.
12. L. Washington, *Introduction to cyclotomic fields*, Springer-Verlag, New York, 1982.

DEPARTMENT OF MATHEMATICS, REED COLLEGE, PORTLAND, OREGON 97202
*E-mail address*: jpb@reed.edu

SCIENTIFIC COMPUTATION GROUP, NEXT COMPUTER, INC., 900 CHESAPEAKE DRIVE, REDWOOD CITY, CALIFORNIA 94063
*E-mail address*: Richard_Crandall@next.com

OAKTON COMMUNITY COLLEGE, DES PLAINES, ILLINOIS 60016
*E-mail address*: u26210@uicvm.cc.uic.edu