Irregular Primes and Cyclotomic Invariants to 12 Million

JOE BUHLER†, RICHARD CRANDALL‡, REIJO ERNVALL§, TAUNO METSÄNKYLÄ¶ and M. AMIN SHOKROLLAHI∥

†Department of Mathematics, Reed College, Portland, Oregon, U.S.A.
‡Center for Advanced Computation, Reed College, Portland, Oregon, U.S.A.
§Forssa Institute of Technology, Forssa, Finland
¶Department of Mathematics, University of Turku, Turku, Finland
∥Bell-L abs, Murray Hill, NJ, U.S.A.

Computations of irregular primes and associated cyclotomic invariants were extended to all primes up to 12 million using multisectioning/convolution methods and a novel approach which originated in the study of Stickelberger codes (Shokrollahi, 1996). The latter idea reduces the problem to that of finding zeros of a polynomial over \( \mathbb{F}_p \) of degree less than \((p - 1)/2\) among the quadratic nonresidues mod \( p \). Use of fast polynomial gcd-algorithms gives an \( O(p \log^2 p \log \log p) \)-algorithm for this task. A more efficient algorithm, with comparable asymptotic running time, can be obtained by using Schönhage–Strassen integer multiplication techniques and fast multiple polynomial evaluation algorithms; this approach is particularly efficient when run on primes \( p \) for which \( p - 1 \) has small prime factors. We also give some improvements on previous implementations for verifying the Kummer–Vandiver conjecture and for computing the cyclotomic invariants of a prime.

© 2001 Academic Press

1. Introduction

In what follows \( p \) always denotes an odd prime.

The Bernoulli numbers \( B_t \) are rational numbers defined by the formal identity

\[
\frac{x}{e^x - 1} = \sum_{t=0}^{\infty} B_t \frac{x^t}{t!}.
\]

The odd-index Bernoulli numbers \( B_{2n+1} \) are zero except for \( B_1 \). A pair \((p, 2t)\), \( 0 < t < (p - 1)/2 \) is called irregular, if and only if \( p \) divides \( B_{2t} \). The number of irregular pairs for \( p \) is called the index of irregularity of \( p \), and is denoted by \( i(p) \); \( p \) is called regular if \( i(p) = 0 \), and is called irregular otherwise.

The irregular pairs were computed for primes less than 125 000 by Wagstaff (1978) and then to 150 000 by Tanner and Wagstaff (1987). These algorithms were quadratic in the sense that the running time for a fixed prime \( p \) is proportional to \( p^2 \). Buhler et al. (1992) were the first to invent an \( O(p^{1+\varepsilon}) \)-algorithm for this task and used their method to extend the computations of irregular pairs to 1 million. Their approach—which we shall call the power series method—is based on the inversion of the power series \((e^x - 1)/x = \sum_{k=1}^{\infty} x^{k-1}/k! \) modulo \( x^{p-1} \) over \( \mathbb{F}_p \). This method, combined with...
further memory-saving strategies, was taken up again by Buhler et al. (1993) to extend previous computations to 4 million.

Associated to each irregular prime \( p \) there are the so-called cyclotomic invariants which describe the class groups of the cyclotomic fields with \( p \)-power conductor. These invariants are known up to 4 million as well, thanks to Wagstaff (1978), Ernvall and Metsänkylä (1991), Ernvall and Metsänkylä (1992), Buhler et al. (1993). In addition, the computations of irregular pairs are the starting point for a verification of the Kummer–Vandiver conjecture.

We report on an extension of all of the above computations to primes between 4 and 12 million.

2. Algorithms

We used two entirely different algorithms: the first was based on the power series method, combined with enhanced multisectioning and convolution algorithms, see Buhler et al. (1993) and Crandall (1996). The second method, which we call the root finding method, is a novel approach which originated in the study of Stickelberger codes initiated in Shokrollahi (1996). It is asymptotically comparable to Buhler et al.’s approach, but allows for great memory and running time savings if \( p-1 \) has only “small” prime divisors, as explained below. Fix a prime \( p \) and let \( w \) be a fixed primitive root modulo \( p \). For \( a \in \mathbb{Z} \) let

\[
h_c(x) := \sum_{j=0}^{p-2} \left\lfloor \frac{cR(w^{-j})}{p} \right\rfloor x^j \in \mathbb{F}_p[x].
\]

This polynomial arises as a generator polynomial for the Stickelberger code. The root finding method is based on the following property of these polynomials.

**Theorem 2.1.** If \( p \) is an odd prime, \( c \) is an integer such that \( 1 \leq c \leq p-1 \), and \( k \) is an integer such that \( 1 < k < p-1 \) then

\[
h_c(w^k) \equiv (c - k) \frac{B_{p-k}}{p-k} \mod p.
\]

**Proof.** Our proof is taken from Shokrollahi (1995). The elements of the galois group \( G \) of the \( p \)th cyclotomic field over \( \mathbb{Q} \) are the \( \sigma_c: \zeta \mapsto \zeta^c \), \( 1 \leq c \leq p-1 \), where \( \zeta = e^{2\pi i/p} \). The group ring \( \mathbb{F}_p[G] \) is isomorphic to \( \mathbb{F}_p[x]/(x^{p-1} - 1) \) via the morphism \( \varphi \) of \( \mathbb{F}_p \)-algebras sending \( \sigma_w \) to \( x \mod (x^{p-1} - 1) \).

Let \( \theta := \frac{1}{p} \sum_d d \sigma_d^{-1} \in \mathbb{Q}[G] \) be the Stickelberger element. Then

\[
(c - \sigma_c)\theta = \sum_{d=1}^{p-1} \left( \frac{cd}{p} - \frac{R(cd)}{p} \right) \sigma_d^{-1}
\]

\[
= \sum_{d=1}^{p-1} \left( \frac{cd}{p} \right) \sigma_d^{-1}
\]

\[
= \sum_{j=0}^{p-2} \left( \frac{cR(w^{-j})}{p} \right) \sigma_w^j.
\]
We thus deduce that \((c - \sigma_c)\theta \mod p\) is a well-defined element of \(F_p[G]\) and
\[
\varphi((c - \sigma_c)\theta \mod p) = h_c \mod (x^{p-1} - 1).
\]
Let \(\omega\) be the Teichmüller character of \(G\), i.e. \(\omega(\sigma_c) \equiv c \mod p\), and consider the central primitive idempotents \(\varepsilon_{\omega^k} := \frac{1}{p-1} \sum_{c=1}^{p-1} \omega^k(\sigma_c)s_{c^{-1}}\) of \(Z_p[G]\) corresponding to \(\omega^k\) for \(0 \leq k \leq p-2\). Note that
\[
\varepsilon_{\omega^k}s_c = \omega^k(\sigma_c)\varepsilon_{\omega^k}.
\]
We deduce that
\[
\varepsilon_{\omega^k}(c - \sigma_c)\theta = (c - \omega^k(\sigma_c))B_{1,\omega^{-k}\varepsilon_{\omega^k}},
\]
where \(B_{1,\omega^{-k}} := \frac{1}{p} \sum_{c=1}^{p-1} \omega^{-k}(\sigma_c)\) is the first generalized Bernoulli number corresponding to \(\omega^{-k}\). We remark that \(B_{1,\omega^{-k}} \in Z_p\) for \(k \neq 1\), and \(B_{1,\omega^{-k}} = 0\) if \(k\) is nonzero and even (cf. Washington, 1982, p. 31). We further deduce from (2.1) that for all \(\gamma \in Z_p[G]\) and all \(k \neq 1\) we have:
\[
\varphi(\varepsilon_{\omega^k}\gamma \mod p) \equiv \varphi(\gamma \mod p)(w^k)\varphi(\varepsilon_{\omega^k} \mod p) \mod pZ_p[x]/(x^{p-1} - 1),
\]
where \(\varphi(\gamma \mod p)(w^k) \in Z_p\) is the value of \(\varphi(\gamma \mod p)\) (regarded as a polynomial over \(Z_p\)) at \(w^k\). (We have to exclude \(k = 1\) since \(B_{1,\omega^{-1}} \notin Z_p\).) Altogether this gives
\[
h_c(w^k) \equiv (c - \epsilon^k)B_{1,\omega^{-k}}
\]
for \(0 \leq k \leq p-2\), \(k \neq 1\). Now we use the congruence
\[
B_{1,\omega^{-n}} \equiv \frac{B_{n+1}}{n+1} \mod p
\]
holding for all odd \(n\) satisfying \(n \neq -1 \mod p-1\), (see Washington, 1982, Corollary 5.15). This gives the assertion of the theorem for odd \(k\) with \(1 < k \leq p-2\). For even \(k\) in this range the congruence is trivially valid since both sides vanish \(\mod p\). □

We thus need to find the roots of \(f\) among the quadratic nonresidues of \(F_p\). Use of fast gcd-algorithms gives an \(O(p\log^2 p\log\log p)\) algorithm for this task (see, e.g. Bürgisser et al., 1996, Chapter 3). The first step towards a faster algorithm is the following: the polynomial \((x^N + 1)\), \(N = (p-1)/2\), whose roots are the quadratic nonresidues in \(F_p\), has a factorization
\[
x^N + 1 = \prod_{j=0}^{d-1} (x^q - w^{(2j+1)q}),
\]
where \(q\) is any divisor of \(N\) and \(d = N/q\). Letting \(q\) be the largest prime divisor of \(N\), we can efficiently compute the polynomials \(f_j(x) := f(x) \mod (x^q - w^{(2j+1)q})\) by using the same idea as in the multiple evaluation algorithm of Borodin and Moenck (1974) (see also Bürgisser et al., 1996). This strategy reduces the problem to that of finding the roots of a polynomial of degree \(< q\) among the \((p-1)/q\)th roots of unity in \(F_p\). Employing Bluestein’s trick (Bluestein, 1970, or Schönhage, 1982, equation 3.2), this problem is reduced to computing a (nega)cyclic convolution of two vectors of length \(q\). Following a suggestion of Schönhage (1995) we transformed the latter problem to an instance of integer multiplication using his method described in Schönhage (1982).

Further savings can be achieved by noting that whenever we compute the zeros of the polynomial \(f\) in a certain coset of a fixed subgroup of \(F_p^\times\), we are actually computing the polynomial product of two polynomials \(h\) and \(v\) one of which, say \(v\), is fixed for all the
different cosets. So we are actually dealing with the problem of multiplying one integer with several other integers. This can be solved efficiently in the following way: we use the Schönhage–Strassen algorithm for multiplying integers. In a first step we generate \(v\) and its Fourier transform and store it. Then, for each \(h\) encountered, we compute its transform, multiply it with the transform of \(v\), and transform the product back. This reduces the number of Fourier transforms per polynomial multiplication from three to two. Another improvement can be gained using an idea of Reischert (1995): If \(p\) is not a Fermat prime, then the largest prime factor of \(p - 1\) is odd, hence we have to perform a negacyclic convolution of \(h\) and \(v\). Translated into integer multiplication, this means that we are performing multiplication modulo \(2^m + 1\) for some \(m\). If \(m\) is such that \(32m = 2^k l < (2k - 1)2^{2k}\), then one can perform the Schönhage–Strassen multiplication algorithm to compute this integer product (see Schönhage et al., 1994, p. 32).

About 60% of the 256,631 primes between 4 and 8 million were processed by the root finding method on a cluster of 56 SPARC-stations of the Department of Computer Science, University of Bonn, and of the Gesellschaft für Mathematik und Datenverarbeitung, Bonn. The program was written in TP-code (Schönhage et al., 1994). TP is an invention of Schönhage and simulates a multitape Turing machine. It is a software implementation of a Turing Processor, which can be programmed via TPAL, the Turing Processor Assembly Language. Currently, there exists a substantial collection of algorithms written in TPAL, including the classical routines for computing with integers and many of the asymptotically fast algorithms for this domain. These clean and efficient implementations made TP the natural choice to implement the root finding method in. The different primes were distributed among the machines available using a resource management program written by Vetter (1995).

The remaining 40% of the primes less than 8 million were handled by the power series method run on a cluster of 100 workstations at NeXT Software, Inc. The convolution algorithms involved are described in detail (including C source code) in Crandall (1996). Using similar algorithms, almost all of the 248,283 primes between 8 and 12 million were handled by a cluster of Alpha processors at the Center for Communications Research in Princeton, New Jersey.

We used the three implementations to extensively cross-check the results of the different algorithms. Besides these checks, we incorporated internal check-sum identities: the power series method used the identity \(\sum_{n=0}^{p-3} 2^n (n + 1)B_n \equiv -4 \mod p\) derived in Buhler et al. (1992). The root finding method was checked with the identity \(\sum_t f(t) = (p - 1)f(0)/2\) holding for all polynomials \(f\) over \(\mathbb{F}_p\) of degree less than \((p - 1)/2\), where \(t\) runs over the quadratic nonresidues of \(\mathbb{F}_p\). These tests seem to be very stringent, and give us considerable confidence in the accuracy of our lists of irregular pairs.

### 3. Results

Incorporating the earlier data on primes less than 4 million, we find that the number \(N_k\) of (odd) primes less than 12 million with index of irregularity \(i(p)\) equal to \(k\) is given in the following table.
Here $p_k$ is the proportion of primes that have index $k$, i.e. $N_k$ divided by the number of odd primes less than 12 million. The bottom line gives the natural heuristic for that fraction, noticed by Lehmer, Siegel, and others. Namely, if the numerators of the Bernoulli numbers are uniformly distributed mod $p$ then for each prime $p$ there are roughly $p/2$ events where the probability of success is $1/p$ so that one expects the distribution to be Poisson with mean $1/2$.

Several people have asked whether or not the proportions have any trend. To briefly answer that question we give the counts for each interval of length 1 million, including the corresponding proportion given in brackets for indices less than 5. The $n$th line gives the counts and proportions for primes $p$ between $(n-1) \times 10^6$ and $n \times 10^6$.

No statistically significant trend seems apparent; chi-square tests suggest that the observed deviation from the Poisson hypothesis is well within expected random variation.

The table itself can be obtained (at least, in 1999) by anonymous ftp from ftp.reed.edu in the directory reed/users/jpb.

The four primes of index 7 are 3 238 481, known from earlier computations, and three new ones: 5 216 111, 5 620 861, and 9 208 289.

As described in Buhler et al. (1993) and references therein, the calculation of irregular pairs is the first (and most time-consuming) step in calculating the so-called cyclotomic invariants and in verifying the Kummer–Vandiver conjecture. This conjecture asserts that the class number $h_p^+$ of the totally real subfield $\mathbf{Q}(\cos(2\pi/p))$ of the field of $p$th roots of unity is prime to $p$. There are reasons to believe that the conjecture is true and reasons to believe that it is false; it follows from our computations that any counterexample has to have $p$ larger than 12 million. Similarly, there were no surprising results in the calculation of the cyclotomic invariants. In particular, the class groups behave as expected (see Buhler et al., 1993) and the lambda-invariant is equal to the index of irregularity for all primes up to 12 million. On heuristic grounds there is a small probability that the lambda-invariant could be larger than the index of irregularity (see Washington, 1982, or Washington's appendix in Lang, 1990, for a thorough discussion), but so far this has not happened.
In both of the Kummer–Vandiver and cyclotomic invariant calculations it is interesting to assess the probability of an “exceptional event” (i.e. a counterexample to the Kummer–Vandiver conjecture or a lambda-invariant exceeding the index of irregularity). In both cases some tempting heuristics (along the lines of assuming that various integers that arise are uniformly distributed mod $p$) lead to a prediction that the number of exceptional events up to $x$ is asymptotic to \( \sum_{p < x} \frac{1}{p} \). This sum diverges very slowly, so perhaps it is not surprising that no such events have been observed. However, it is clear that this heuristic does not capture the whole story; for instance, the class number \( h^+_p \) is likely to be small for small $p$ and therefore $p$ can not divide $h^+_p$ for small $p$. Of course, one possible explanation for the lack of exceptional events is that none exist; i.e. these heuristics might well be wrong. For a more thorough discussion of some of these questions see Washington (1982), Lang (1990), or Banaszak et al. (1996).

4. Vandiver and Cyclotomic Tests

For primes under 125 000, Wagstaff noted that the calculation of the irregular indices dwarfed the time required to compute the cyclotomic invariants and the time required to check Kummer–Vandiver. However, naive implementations have complexity $O(p \log^{2+\epsilon} p)$, and this back-of-the-envelope asymptotic complexity has turned out to be accurate: as the primes increased the proportion of time required by these secondary tests became more and more noticeable.

Two simple devices reduced the complexity to $O(p \log^{1+\epsilon} p)$, and greatly expedited this part of the calculations.

First, recall from Ernvall and Metsänkylä (1992) that the computation of the cyclotomic invariants involves the computation of the sums

\[
S_1 := \sum_{x=1}^{(p-1)/2} x^{t-1} q_x \mod p, \quad S_3 := \sum_{x=1}^{(p-1)/2} x^{t-1} \mod p^2
\]

where \((p,t)\) is an irregular pair and \(q_x = (x^{p-1} - 1)/p \mod p\) is the Fermat quotient at $x$. Rather than computing these sums in a straightforward way, Ernvall and Metsänkylä (1992) found it convenient to pass from $x$ to $2x$ or $p - 2x$ according to whether $2x$ was less than $(p-1)/2$ or not. We keep track of the orbits of “multiplication by 2” by keeping an array of $(p-1)/2$ bits and mark each value of $x$ as it is visited. This process is expedited by the observation that \(q_{2x} = q_x + q_2\) if $2x < p/2$ and \(q_{p-2x} = q_2 + q_x + (2x)^{-1}\) if $p/2 < 2x$.

Now we consider the test of the Kummer–Vandiver conjecture. Briefly (see Washington, 1982, for details), the class number $h^+_p$ is the index of the “cyclotomic units” inside the real units of the $p$th cyclotomic field. To check that this index is prime to $p$ it is useful to exploit the action of the galois group and check eigenspaces for characters of the galois group, which can be done by checking that an explicit cyclotomic unit lying in that eigenspace is not a $p$th power. It turns out that this is equivalent to verifying that for each irregular pair $(p,t)$ the unit

\[
\prod_{c=1}^{p-1} \left( \frac{\zeta^{wc/2} - \zeta^{-wc/2}}{\zeta^{c/2} - \zeta^{-c/2}} \right)^{c^{p-1-t}}
\]

is not a $p$th power (where, as above, $w$ is a primitive root mod $p$, and $\zeta = e^{2\pi i/p}$). It
Irregular Primes to Twelve Million

suffices to verify that this is not a \( p \)th power for some prime ideal of degree one. By some further algebra, this reduces to the following test: let \( q \) be a prime congruent to 1 modulo \( p \), \( z \) be a \( p \)th root of unity mod \( q \), and

\[
V_{p,t} = \prod_{c=1}^{(p-1)/2} (z^c - z^{-c})^{p-1-t} \mod q.
\]

Then to check that the above unit is not a \( p \)th power it suffices (Washington, 1982) to check that \( V_{p,t}^{(q-1)/p} \) is not congruent to 1 modulo \( q \). In all cases, we found that it sufficed to take \( q \) to be the smallest prime congruent to 1 mod \( p \), and \( z = 2^{(q-1)/p} \). Historically, this test was devised by Vandiver, building on Kummer’s results on cyclotomic fields, in connection with Fermat’s last theorem: Vandiver proved that Fermat’s last theorem is true for \( p \) if there is a prime \( q \) less than \( p^2 - p \) such that the above criterion is true for all irregular pairs \((p, t)\) for the given prime \( p \).

One observation that simplifies the calculation of \( V_{p,t} \) slightly is that we only need to calculate \( e^{p-1-t} \) modulo \( p \) (rather than \( q - 1 \)) since we later raise \( V_{p,t} \) to the power \((q - 1)/p\). However, a dramatic speed-up (for which we thank Peter Montgomery) is obtained by “caching” the exponentiations for evaluation all at once at the end.

Each exponent \( e = e^{p-1-t} \) in the product for \( V_{p,t} \) is written in the form \( e = e_0 + 4096e_1 \) with \( 0 \leq e_i < 4096 \). We keep arrays \texttt{low} and \texttt{high} that serve as storage bins for the postponed exponentiations. Rather than explicitly performing the exponentiation

\[
u^e = u^{e_0}u^{4096e_1} \mod q,
\]

where \( u = (z^c - z^{-c}) \mod q \), we merely multiply \texttt{low}[e_0] and \texttt{high}[e_1] by \( u \). At the end we need to compute

\[
V_{p,t} = \prod_{i=1}^{4095} \text{low}[i]^{\text{high}[i]}^{4096i}.
\]

This can be done by a simple loop; e.g. the product of the \texttt{low}[i] can be calculated by

1. \texttt{product} = 1; \texttt{terms} = 1
2. \texttt{for} \( i = 4095, 4094, \cdots, 2, 1 \)
   
   \texttt{terms} = \texttt{terms} * \texttt{low}[i]
   
   \texttt{product} = \texttt{product} * \texttt{terms}

The product of the high exponent terms can be calculated similarly. The savings over a direct evaluation of the product for \( V_{p,t} \) is considerable.

Acknowledgements

Many thanks go to D. Reischert, A. Schönhage, and E. Vetter for their help in implementing and enhancing the root finding method. Peter Montgomery provided some important insights at several points during the project. The authors are also grateful to T. Donahue and J. Doenias of NeXT Software, Inc., and to David Goldschmidt and Jeff Prisner of the Center for Communications Research; in addition we thank the Department of Computer Science of the University of Bonn, the Gesellschaft für Mathematik und Datenverarbeitung in Bonn, the Center for Scientific Computing of Finland,
the Center for Communications Research in Princeton, and the International Computer Science Institute in Berkeley for providing time on their computers.

References


Originally Received 3 September 1996
Accepted 6 April 1999