The Elliptic Curve Method for Factoring

Paul Zimmermann, INRIA Nancy - Grand Est
Paul.Zimmermann@inria.fr

17 September, 2009, revised 25 September, 2010

Synonyms. ECM.


Definition. The Elliptic Curve Method (ECM for short) was invented in 1985 by H. W. Lenstra, Jr. [5]. It is suited to find small — say 10 to 40 digits — prime factors of large numbers. Among the different factorization algorithms whose complexity mainly depends on the size of the factor searched for (trial division, Pollard rho, Pollard $p - 1$, Williams $p + 1$), it is asymptotically the best method known. ECM can be viewed as a generalization of Pollard’s $p - 1$ method, just like ECPP generalizes the $n - 1$ primality test. ECM relies on Hasse’s theorem: if $p$ is prime, then an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$ has group order $p + 1 - t$ with $|t| \leq 2\sqrt{p}$, where $t$ depends on the curve. If $p + 1 - t$ is a smooth number (see smoothness), then ECM will — most probably — succeed and reveal the unknown factor $p$.

Background. Since 1985, many improvements have been proposed to ECM. Lenstra’s original algorithm had no second phase. Brent proposes in [2] a “birthday paradox” second phase, and further more technical refinements. In [7], Montgomery presents different variants of phase two of ECM and Pollard $p - 1$, and introduces a parameterization with homogeneous coordinates, which avoids inversions modulo $n$, with only 6 and 5 modular multiplications per addition and duplication on $E$, respectively. It is also possible to choose elliptic curves with a group order divisible by 12 or 16 [1, 7, 8].
Phase one of ECM works as follows. Let \( n \) be the number to factor. An elliptic curve is \( E(\mathbb{Z}/n\mathbb{Z}) = \{(x : y : z) \in \mathbb{P}^2(\mathbb{Z}/n\mathbb{Z}), y^2z \equiv x^3 + axz^2 + bz^3 \mod n\} \), where \( a, b \) are two parameters from \( \mathbb{Z}/n\mathbb{Z} \), and \( \mathbb{P}^2(\mathbb{Z}/n\mathbb{Z}) \) is the projective plane over \( \mathbb{Z}/n\mathbb{Z} \). The neutral element is \( \mathcal{O} = (0 : 1 : 0) \), also called point at infinity. The key idea is that computations in \( E(\mathbb{Z}/n\mathbb{Z}) \) project to \( E(\mathbb{Z}/p\mathbb{Z}) \) for any prime divisor \( p \) of \( n \), with the important particular case of quantities which are zero in \( E(\mathbb{Z}/p\mathbb{Z}) \) but not in \( E(\mathbb{Z}/n\mathbb{Z}) \). Pick at random a curve \( E \) and a point \( P \) on it. Then compute \( Q = k \cdot P \) where \( k \) is the product of all prime powers less than a bound \( B_1 \). Let \( p \) be a prime divisor of \( n \): if the order of \( E \) over \( \mathbb{Z}/p\mathbb{Z} \) divides \( k \), then \( Q \) will be the neutral element of \( E(\mathbb{Z}/p\mathbb{Z}) \), thus its \( z \)-coordinate will be zero modulo \( p \), hence gcd(\( z, n \)) will reveal the factor \( p \) (unless \( z \) is zero modulo another factor of \( n \), which is unlikely).

Phase one succeeds when all prime factors of \( g = \#E(\mathbb{Z}/p\mathbb{Z}) \) are less than \( B_1 \); phase two allows one prime factor \( g_1 \) of \( g \) to be as large as another bound \( B_2 \). The idea is to consider two families \( (a_iQ) \) and \( (b_jQ) \) of points on \( E \), and check whether two such points are equal over \( E(\mathbb{Z}/p\mathbb{Z}) \). If \( a_iQ = (x_i : y_i : z_i) \) and \( b_jQ = (x'_j : y'_j : z'_j) \), then gcd(\( x_i z'_j - x'_j z_i, n \)) will be non-trivial. This will succeed when \( g_1 \) divides a non-trivial \( a_i - b_j \). Two variants of phase two exist: the birthday paradox continuation chooses the \( a_i \)'s and \( b_j \)'s randomly, expecting that the differences \( a_i - b_j \) will cover most primes up to \( B_2 \), while the standard continuation chooses the \( a_i \)'s and \( b_j \)'s so that every prime up to \( B_2 \) divides at least one \( a_i - b_j \). Both continuations may benefit from the use of fast polynomial arithmetic, and are then called “FFT extensions” [8].

**Theory.** The expected running time of ECM is conjectured to be \( \mathcal{O}(L(p)^{\sqrt{2+o(1)}} M(\log n)) \) to find one factor of \( n \), where \( p \) is the (unknown) smallest prime divisor of \( n \), \( L(x) = e^{\sqrt{\log x \log \log x}} \) [cf. \( L \)-notation], \( M(\log n) \) represents the complexity of arithmetic modulo \( n \), and the \( o(1) \) in the exponent is for \( p \) tending to infinity. The second phase decreases the expected running time by a factor \( \log p \). Optimal bounds \( B_1 \) and \( B_2 \) may be estimated from the (usually unknown) size of the smallest factor of \( n \), using Dickman’s function [9]. For RSA moduli, where \( n \) is the product of two primes of roughly the same size, the running time of ECM is comparable to that of the Quadratic Sieve.

**Applications.** ECM has been used to find factors of Cunningham numbers \((a^n \pm 1 \text{ for } a = 2, 3, 5, 6, 7, 10, 11, 12)\). In particular Fermat numbers
$F_n = 2^{2^n} + 1$ are very good candidates for $n \geq 10$, since they are too large for general purpose factorization methods. Brent completed the factorization of $F_{10}$ and $F_{11}$ using ECM, after finding a 40-digit factor of $F_{10}$ in 1995, and two factors of 21 and 22 digits of $F_{11}$ in 1988 [3]. Brent, Crandall, Dilcher and Van Halewyn found a 27-digit factor of $F_{13}$ in 1995, a (different) 27-digit factor of $F_{16}$ in 1996, and a 33-digit factor of $F_{15}$ in 1997. In 2009, Bessel found a 35-digit factor of $F_{19}$.

Some applications of ECM are less obvious. The factors found by the Cunningham project [4] help to find primitive polynomials over $\text{GF}(q)$. They are also used in the Jacobi sum and cyclotomy tests for primality proving [6].

**Experimental Results.** Brent maintains a list of the ten largest factors found by ECM (http://wwwmaths.anu.edu.au/~brent/ftp/champs.txt); his extrapolation from previous data would give an ECM record of 85 digits in year 2018, and 100 digits in year 2025. As of September 2010, the ECM record is a factor of 73 digits.

**Open Problems.** It is not known whether the expected running time of ECM can be improved — either in phase 1 or in phase 2 — nor whether there exists a method with better asymptotic complexity depending only on the size $\log p$ of the smallest prime factor, apart from polynomial terms in $\log n$.

**Recommended Readings**


