F12ECM

A program for finding
the factors of
the twelfth Fermat number

Elliptic Curve Method and Probabilities

Yves Gallot

February 15, 2021

F_{12}ECM is free source code, under the MIT license.

Chapter 1

Fermat numbers

> Si je puis une fois tenir la raison fondamentale que 3, 5, 17, etc. sont nombres premiers, il me semble que je trouverai de très belles choses en cette matière,
>
> *Fermat à Mersenne*
> *25 décembre 1640*

Pierre de Fermat conjectured that every number of the form $F_n = 2^{2^n} + 1$, where n is a non-negative integer, is prime [6]. Today these positive integers are named Fermat numbers. The first five Fermat numbers are prime, but Leonhard Euler proved in 1732 that 641 divides F_5.

F_6 was completely factored by T Clausen, F. Landry and H. Le Lasseur in 1855. In 1970, M. A. Morrison and J. Brillhart cracked F_7 by the Continued Fraction method [10]. In 1980, R. P. Brent and J. M. Pollard used a modification of Pollard's rho method to factor F_8. R. P. Brent completely factored F_{11} in 1988 by ECM [3]. In 1990, A. K. Lenstra, H. W. Lenstra, M. S. Manasse and J. M. Pollard organized a distributed computation on approximately 700 workstations around the world and factored F_9 by the Number Field Sieve [8]. Finally R. P. Brent completely factored F_{10} in 1995 by ECM [3].

The smallest Fermat number which is not completely factored is F_{12}. Six prime factors are known, the 54-digit factor was found by Michael Vang in 2010 using GMP–ECM [12].

$$
\begin{aligned}
F_5 &= 641 \cdot 6700417 \\
F_6 &= 274177 \cdot 67280421310721 \\
F_7 &= 59649589127497217 \cdot 5704689200685129054721 \\
F_8 &= 1238926361552897 \cdot P_{62} \\
F_9 &= 2424833 \cdot 7455602825647884208337395736200454918783366342657 \cdot P_{99} \\
F_{10} &= 45592577 \cdot 6487031809 \cdot 4659775785220018543264560743076778192897 \cdot P_{252} \\
F_{11} &= 319489 \cdot 974849 \cdot 167988556341760475137 \cdot 3560841906445833920513 \cdot P_{564} \\
F_{12} &= 114689 \cdot 26017793 \cdot 63766529 \cdot 190274191361 \cdot 1256132134125569 \cdot \\
&\quad 5686306475353569551690334109408678048393607420608184 33 \cdot C_{1133}
\end{aligned}
$$

Chapter 2

Elliptic curves modulo p

An elliptic curve over $\mathbb{Z}/p\mathbb{Z}$ is a set of points $P = (x, y) \in (\mathbb{Z}/p\mathbb{Z})^2$ such that

$$y^2 = x^3 + ax + b$$

with the condition $4a^3 + 27b^2 \neq 0$ and an additional point at infinity O.

The point $Q = (x, -y)$ is on the curve: $Q = -P$ is the point opposite of P. The curve has the property that if a non-vertical line intersects it at two points P and Q, then it will also have a third point R of intersection. The addition law is defined by $P + Q = -R$. If $P = Q$, the tangent of the curve at P is considered. If the line is vertical, we have $P + -P = O$. The points on an elliptic curve and the addition form an abelian group. O is the identity of the group: we have $P + O = O + P = P$.

Hasse's theorem on elliptic curves over finite fields provides an estimate of the number of points. If $\#E$ is the order of the group of an elliptic curve E over $\mathbb{Z}/p\mathbb{Z}$ then

$$|\#E - (p + 1)| \leq 2\sqrt{p}.$$

Given a prime $p > 3$ and any integer n such that $|n - (p + 1)| \leq 2\sqrt{p}$, there exists a and b such that $|\#E(a, b)| = n$. Furthermore, the numbers of points are uniformly distributed over the interval $[p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p}]$.

To avoid the time-consuming inversion over $\mathbb{Z}/p\mathbb{Z}$, projective coordinates are preferred for computations. X, Y, Z are integers such that $x = X/Z$ and $y = Y/Z$. If $P \neq O$ then $Z \neq 0$ and the coordinates of the identity O are $(0, Y, 0)$.

$y^2 = x^3 + ax + b$ is called the Weierstrass form but there exist some alternative representations of elliptic curves. Some of them are faster for computations. F_{12}ECM representation is the Montgomery curve

$$By^2 = x^3 + Ax^2 + x.$$

The Montgomery curves are a subset of Weierstrass curves.

Montgomery coordinates are the projective coordinates (X, Z). The computation of Y is not needed for the Elliptic Curve Method.

Chapter 3

Elliptic Curve Method

3.1 Algorithm

The elliptic curve factorization method (ECM) is an extension of Pollard's $p-1$ algorithm [11].

Pollard's $p-1$ algorithm finds the prime factors p such that $p-1$ is B-smooth.
If $p>2$ and e is a multiple of $p-1$ then by Fermat's little theorem we have $2^e \equiv 1 \pmod{p}$. For a fixed bound B, $M = \prod_{\substack{p \leq B \\ p \text{ prime}}} p^{\lfloor \log_p B \rfloor}$ is computed modulo n and finally $g = \gcd(2^M - 1, n)$.
If n is a composite integer, $p \mid n$, $p-1$ is B-powersmooth and $q \mid n$ but $q-1$ is not B-powersmooth then $1 < g < n$ and g is a multiple of p.
In practice, a two-stage variant of the algorithm is implemented: instead of requiring that $p-1$ has all its factors less than B, if all but one of them are less than B_1 and the remaining factor is less than B_2 then the range $]B_1; B_2]$ can be tested more quickly. M is computed for $B = B1$ and the second stage is $M' = \prod_{\substack{B_1 < p \leq B_2 \\ p \text{ prime}}} \left((2^M)^p - 1\right)$. If p_n and p_{n+1} are two consecutive prime numbers then $A^{p_{n+1}} = A^{p_n} \cdot A^{d_n}$ where $d_n = p_{n+1} - p_n$. The d_n are relatively small then the values of $A^2, A^4, A^6, \ldots$ can be precomputed. Then $(2^M)^p$ is calculated with a single multiplication.

Fermat's little theorem can be extended to a group G such that each element of G is invertible. Pollard's $p-1$ is based on $(\mathbb{Z}/p\mathbb{Z})^\times$. A finite field of order q exists if and only if $q = p^k$, where p is a prime number and k is a positive integer. The order of $\mathbf{F}_q^\times$ is $p^k - 1$. A factor of $p^k - 1$ is sufficient for Pollard's method and the factors of $p^k - 1$ are the cyclotomic polynomials [1]. But in practice this algorithm is slower than ECM except $p+1$.

Since the points on an elliptic curve over $\mathbb{Z}/p\mathbb{Z}$ forms an abelian group, Pollard's algorithm can be extended to this set. Because of Hasse's theorem, ECM is a "$p+1-a$" algorithm. a is unknown and $a \in [-2\sqrt{p}, 2\sqrt{p}]$. With different curves, we can expect that a is a random number in a large range.

If $\mathscr{P}$ is the probability that $p+1-a$ is B-smooth, with n curves the likelihood of success is $\mathscr{P}_n = 1-(1-\mathscr{P})^n$. If $\mathscr{P} \ll 1$ and $n \sim 1/\mathscr{P}$ then $\mathscr{P}_n = 1 - e^{-1} \approx 63.2\%$.
If $\mathscr{P}$ is 1% with Pollard's $p-1$, the likelihood of success with ECM is 63.2% with 100 curves and 99.99% with 1000 curves.

3.2 Largest prime factors

Dickman [5] proved that the probability that a large integer n has no prime factor exceeding n^α approaches a limit $F(\alpha)$ as $n \to \infty$, where

$$F(\alpha) = \begin{cases} 1 - \int_\alpha^1 F(\frac{t}{1-t})\frac{dt}{t} & \text{if } 0 \le \alpha < 1, \\ 1 & \text{if } \alpha \ge 1. \end{cases}$$

Let $u = 1/\alpha$. $F(1/u) = 1 - \int_{1/u}^1 F(\frac{t}{1-t})\frac{dt}{t}$. If $t' = 1/t$ we have $\frac{t}{1-t} = \frac{1}{t'-1}$ and $\frac{dt}{t} = -\frac{dt'}{t'}$ then $F(1/u) = 1 - \int_1^u F(\frac{1}{t'-1})\frac{dt'}{t'}$. The relation becomes

$$F(1/u) = \rho(u) = \begin{cases} 1 - \int_1^u \frac{\rho(t-1)}{t}\,dt & \text{if } u > 1, \\ 1 & \text{otherwise.} \end{cases}$$

ρ is the Dickman function used to estimate the proportion of smooth numbers up to a given bound.

Differentiating both sides of the definition of $\rho(u)$ for $u > 1$ gives $t\,\rho'(t) = -\rho(t-1)$. Integration by parts of $\rho(t)$ and t is $\int_1^u \rho(t)\,dt = [t\,\rho(t)]_1^u - \int_1^u t\,\rho'(t)\,dt$. Hence $\int_1^u \rho(t)\,dt = u\,\rho(u) - 1 + \int_0^{u-1} \rho(t)\,dt$. Since $\int_0^1 \rho(t)\,dt = 1$ we have

$$\rho(u) = \frac{1}{u}\int_{u-1}^u \rho(t)\,dt.$$

This relation can be used for the numerical computation of ρ by approximating the integral with the trapezoidal formula [9]. If $1 \le u < 2$, $\rho(u) = 1 - \int_1^u \frac{dt}{t} = 1 - \log u$.

Knuth and Trabb Pardo [7] extended Dickman's theorem and shown that the probability that the k^{th} largest prime factor of a number n is at most n^α tends to a limiting distribution $F_k(\alpha)$ as $n \to \infty$, where $F_0(\alpha) = 0$ for all α by convention and for $k \ge 1$

$$F_k(\alpha) = \begin{cases} 1 - \int_\alpha^1 \left(F_k(\frac{t}{1-t}) - F_{k-1}(\frac{t}{1-t})\right)\frac{dt}{t} & \text{if } 0 \le \alpha < 1, \\ 1 & \text{if } \alpha \ge 1. \end{cases}$$

We can define the generalized Dickman function $\rho_k(u) = F_k(1/u)$ and we have

$$\rho_k(u) = \begin{cases} 1 - \int_1^u (\rho_k(t-1) - \rho_{k-1}(t-1))\frac{dt}{t} & \text{if } u > 1 \text{ and } k \ge 1, \\ 1 & \text{if } 0 < u \le 1 \text{ and } k \ge 1, \\ 0 & \text{otherwise.} \end{cases}$$

The differential equation is $t\,\rho_k'(t) = -\rho_k(t-1) + \rho_{k-1}(t-1)$. Integrating by parts, we get

$$\rho_k(u) = \frac{1}{u}\left(\int_{u-1}^u \rho_k(t)\,dt + \int_0^{u-1} \rho_{k-1}(t)\,dt\right).$$

This relation can be used for the numerical computation of ρ_2.

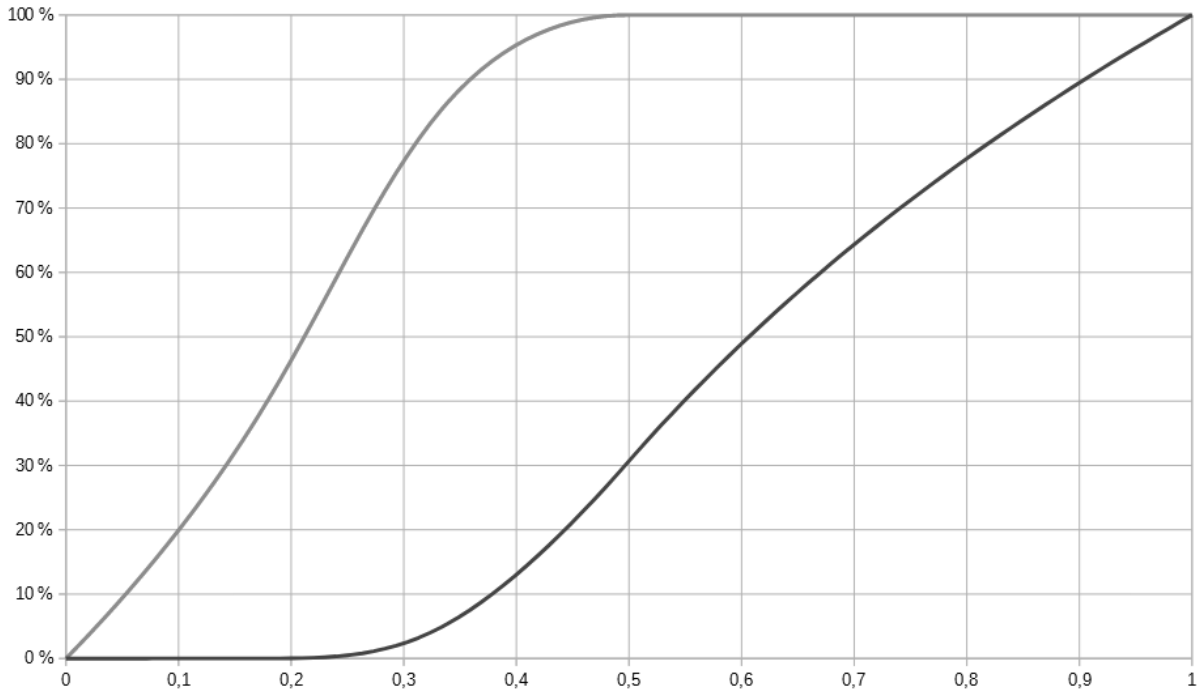

Figure 3.1: $F(\alpha)$ and $F_2(\alpha)$.

3.3 Probabilities

Chapter 4

Implementation

Bibliography

[1] Eric Bach and Jeffrey Shallit, *Factoring with cyclotomic polynomials*, Math. Comp. **52** (1989), 201–219, DOI: https://doi.org/10.1090/S0025-5718-1989-0947467-1.

[2] Richard P. Brent, *Factorization of the tenth Fermat number*, Math. Comp. **68** (1999), 429–451, DOI: https://doi.org/10.1090/S0025-5718-99-00992-8.

[3] Richard P. Brent, *Factorization of the tenth and eleventh Fermat numbers*, Report TR-CS-96-02, Computer Sciences Laboratory, Australian National Univ., Canberra, Feb. 1996, https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.70.6415&rep=rep1&type=pdf.

[4] Richard P. Brent and John M. Pollard, *Factorization of the eighth Fermat number*, Math. Comp. **36** (1981), 627–630, DOI: https://doi.org/10.1090/S0025-5718-1981-0606520-5.

[5] Karl Dickman, *On the Frequency of Numbers Containing Prime Factors of a Certain Relative Magnitude*, Arkiv för Mat., Astron. och Fys. 22A, 1-14, 1930.

[6] Pierre de Fermat, *Lettre à Marin Mersenne*, https://www.archive.org/stream/oeuvresdefermat942ferm#page/212/mode/2up.

[7] Donald E. Knuth and Luis Trabb Pardo, *Analysis of a simple factorization algorithm*, Theoretical Computer Science, Volume 3, Issue 3, December 1976, Pages 321–348, DOI: https://doi.org/10.1016/0304-3975(76)90050-5.

[8] A. K. Lenstra, H. W. Lenstra, M. S. Manasse and J. M. Pollard *The factorization of the ninth Fermat number*, Math. Comp. **61** (1993), 319–349, DOI: https://doi.org/10.1090/S0025-5718-1993-1182953-4.

[9] J. van de Lune and E. Wattel, *On the numerical solution of a differential-difference equation arising in analytic number theory*, Math. Comp. **23** (1969), 417–421, DOI: https://doi.org/10.1090/S0025-5718-1969-0247789-3.

[10] Michael A. Morrison and John Brillhart, *A method of factoring and the factorization of F_7*, Math. Comp. **29** (1975), 183–205, DOI: https://doi.org/10.1090/S0025-5718-1975-0371800-5.

[11] J. M. Pollard, *Theorems on factorization and primality testing*, Mathematical Proceedings of the Cambridge Philosophical Society, Volume **76**, Issue 3, November 1974, pp. 521 - 528, DOI: https://doi.org/10.1017/S0305004100049252.

[12] https://caramel.loria.fr/f12.txt.