20 Years of ECM

Paul Zimmermann, INRIA LORRAINE Loria

Bruce Dodson, LEHIGH UNIVERSITY

ANTS 7, July 2006

History of ECM (1/2)

ebruary 1985: invented by H. W. Lenstra, Jr.

nd of 1985 (Brent, Montgomery):

stage 2

Brent-Suyama's extension

fast polynomial evaluation

986: ECM routinely finds factors of 20 digits, Montgomery finds a 36-digit ctor of L_{464}

ent: "*we can forsee that p around 10^{50} may be accessible in a few years ne*"

History of ECM (2/2)

- 987 (Montgomery) unified description of ECM, P+1, P-1
- 988 (Brent) 21-digit and 22-digit factors of F_{11}
- 992 (Montgomery) “FFT extension”
- 995: Brent finds a 40-digit factor of F_{10} (291-digit input)
- 998: Curry finds a 53-digit factor of $2^{677} - 1$ with MPRIME
- 005: Dodson finds a 66-digit factor of $3^{466} + 1$

Notations

number to be factored

(unknown) prime factor of n

a prime

$M(d)$: cost of multiplying two d-bit integers, or two degree-d polynomials

Elliptic Curve

field of characteristic $\neq 2, 3$

ontgomery form:

$$E_{a,b} = \{(x : y) \in K^2, by^2 = x^3 + ax^2 + x\} \cup \{O_E\}$$

omogeneous form:

$$by^2z = x^3 + ax^2z + xz^2$$

here $(x : y : z)$ represents $(x/z : y/z)$.

The ECM algorithm

put: a number n, integer bounds $B_1 \leqslant B_2$

utput: a factor of n, or FAIL

hoose a random elliptic curve $E_{a,b} \bmod n$ and $P_0 = (x_0 : y_0 : z_0)$ on it

tage 1] Compute $Q := \prod_{\pi \leqslant B_1} \pi^{\lfloor (\log B_1)/(\log \pi) \rfloor} P_0$ on $E_{a,b}$

tage 2] for each prime π, $B_1 < \pi \leqslant B_2$:

compute $(x_\pi : y_\pi : z_\pi) = \pi Q$ on $E_{a,b}$

$g \leftarrow \gcd(n, z_\pi)$

if $g \neq 1$, output g and exit

utput FAIL.

Suyama's parametrization

hoose $\sigma > 5$

$= \sigma^2 - 5, \qquad v = 4\sigma$

$_0 = u^3, \qquad z_0 = v^3$

$= (v-u)^3(3u+v)/(4u^3v) - 2$

e then have

$$by^2z = x^3 + ax^2z + xz^2$$

th $b = u/z_0$ and $y_0 = (\sigma^2 - 1)(\sigma^2 - 25)(\sigma^4 - 25)$

and y useless: identify $P = (x : y : z)$ and $-P = (x : -y : z)$

idely used (Brent, Montgomery, Woltman)

Why does ECM work?

asse's theorem:

$$|g - (p + 1)| < 2\sqrt{p}$$

"random" integer in $[p + 1 - 2\sqrt{p}, p + 1 + 2\sqrt{p}]$

uyama's parametrization: 12 divides g

ontgomery's form: 4 divides g

is found if g is (B_1, B_2) smooth, in stage 1 if B_1-smooth

Complexity of ECM

$$O(L(p)^{\sqrt{2}+o(1)}M(\log n))$$

here $L(p) = e^{\sqrt{\log p \log \log p}}$

ontgomery (1992): stage 2 saves a factor of $\log p$

$(p)^{\sqrt{2}+o(1)}$: mathematical and algorithmic improvements

$I(\log n)$: arithmetic improvements

Stage 1

asic operations: curve addition and duplication.

$P, Q, P-Q) \longrightarrow P+Q$ in 6 multiplications mod n:

$$
\begin{array}{ll}
u \leftarrow (x_P + z_P)(x_Q - z_Q) & v \leftarrow (x_P - z_P)(x_Q + z_Q) \\
w \leftarrow (u+v)^2 & t \leftarrow (u-v)^2 \\
x_{P+Q} \leftarrow z_{P-Q} \cdot w & z_{P+Q} \leftarrow x_{P-Q} \cdot t.
\end{array}
$$

$\longrightarrow 2P$ in 5 multiplications mod n:

$$
\begin{array}{lll}
u \leftarrow (x_P + z_P)^2 & v \leftarrow (x_P - z_P)^2 & t \leftarrow d(u-v) + v \\
x_{2P} \leftarrow uv & z_{2P} \leftarrow (u-v)t. &
\end{array}
$$

Stage Two

age 1 computes Q on E

age 2 succeeds when $\pi Q = O_E \bmod p$ for $B_1 \leqslant \pi \leqslant B_2$

$$\pi = \sigma + \tau, \quad \sigma \in S, \tau \in T$$

et $\sigma Q = (x_\sigma : y_\sigma)$, $\tau Q = (x_\tau : y_\tau)$

hen $x_\sigma = x_\tau \bmod P$: simply compute $\gcd(x_\sigma - x_\tau, n)$

andard continuation: S and T are arithmetic progressions.

$$S = \{id, 0 \leqslant id < B_2\}$$

$$T = \{j, 0 < j < d, \gcd(j, d) = 1\}$$

Example: $B_2 = 960, 162$ primes

$= 60$:

$= \{0, 60, 120, 180, 240, 300, 360, 420, 480, 540, 600, 660, 720, 780, 840, 900\}$

$$T = \{1, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 49, 53, 59\}$$

$jQ = (x : y)$, $-jQ = (x : -y)$, thus the prime $67 = 60 + 7$ will also
e hit when computing $\gcd(x_{120} - x_{53}, n)$

eep only $j = 1 \bmod 6$, or $j < d/2$ (Montgomery)

$= 90$:

$$S = \{0, 90, 180, 270, 360, 450, 540, 630, 720, 810, 900\}$$

$$T = \{1, 7, 13, 19, 31, 37, 43, 49, 61, 67, 73, 79\}$$

Fast Polynomial Arithmetic

$$h = \prod_{\sigma \in S} \prod_{\tau \in T} (x_\sigma - x_\tau) \bmod n$$

$$F(X) = \prod_{\tau \in T} (X - x_\tau), \qquad G(X) = \prod_{\sigma \in S} (X - x_\sigma)$$

(X) and $G(X)$ can be computed in $O(M(d) \log d)$ operations (product
ee) if $\deg(F), \deg(G) \leqslant d$.

$$h = \pm \prod_{\tau \in T} G(x_\tau) \bmod n$$

ultipoint polynomial evaluation: $O(M(d) \log d)$. Best constant with a
caled remainder tree" (Bostan, Lecerf, Schost, Bernstein)

Scaled Remainder Tree

Convert $n = 3586334585$ in hexadecimal

$$x_1 = n/16^8 \approx 0.83500858978368 34$$

$$= \lfloor 16^4 x_1 \rfloor / 16^4 \approx 0.83500859 \quad x_3 = 16^4 x_1 \bmod 1 \approx 0.12294006$$

$$0.8350 \quad 0.7622 \quad 0.1229 \quad 0.4727$$

$$0.84 \quad 0.36 \quad 0.76 \quad 0.20 \quad 0.12 \quad 0.97 \quad 0.47 \quad 0.56$$

$$D \quad 5 \quad C \quad 3 \quad 1 \quad F \quad 7 \quad 9$$

Other Improvements

ore technical details in the proceedings (pages 525-542):

efficient arithmetic mod n

evaluation of Lucas chains (Montgomery)

Kronecker-Schönhage's trick

stage 2 blocks

Brent-Suyama's extension

Montgomery's $d_1 d_2$ improvement

GMP-ECM

- toy-project started in 1999, to try how efficient GMP was
- several people used it, some of them for their research
- major improvements in version 5 (Kruppa) and 6 (Newman)
- now quite stable, used in Magma, distributed within Debian
- unified stage 2 for ECM, P-1, P+1

Dodson's 66-digit record

und on April 6, 2005, with GMP-ECM 6.0.1:

$$3^{466}+1 = 2 \times 5 \times 3733008450772109$$

$$\times 32403444713283317229486590 9 \times \underbrace{180 \cdots 513}_{180 \text{ digits}}$$

$$709601635082267320966424084955776789770864725643996885415676682297 \times p114$$

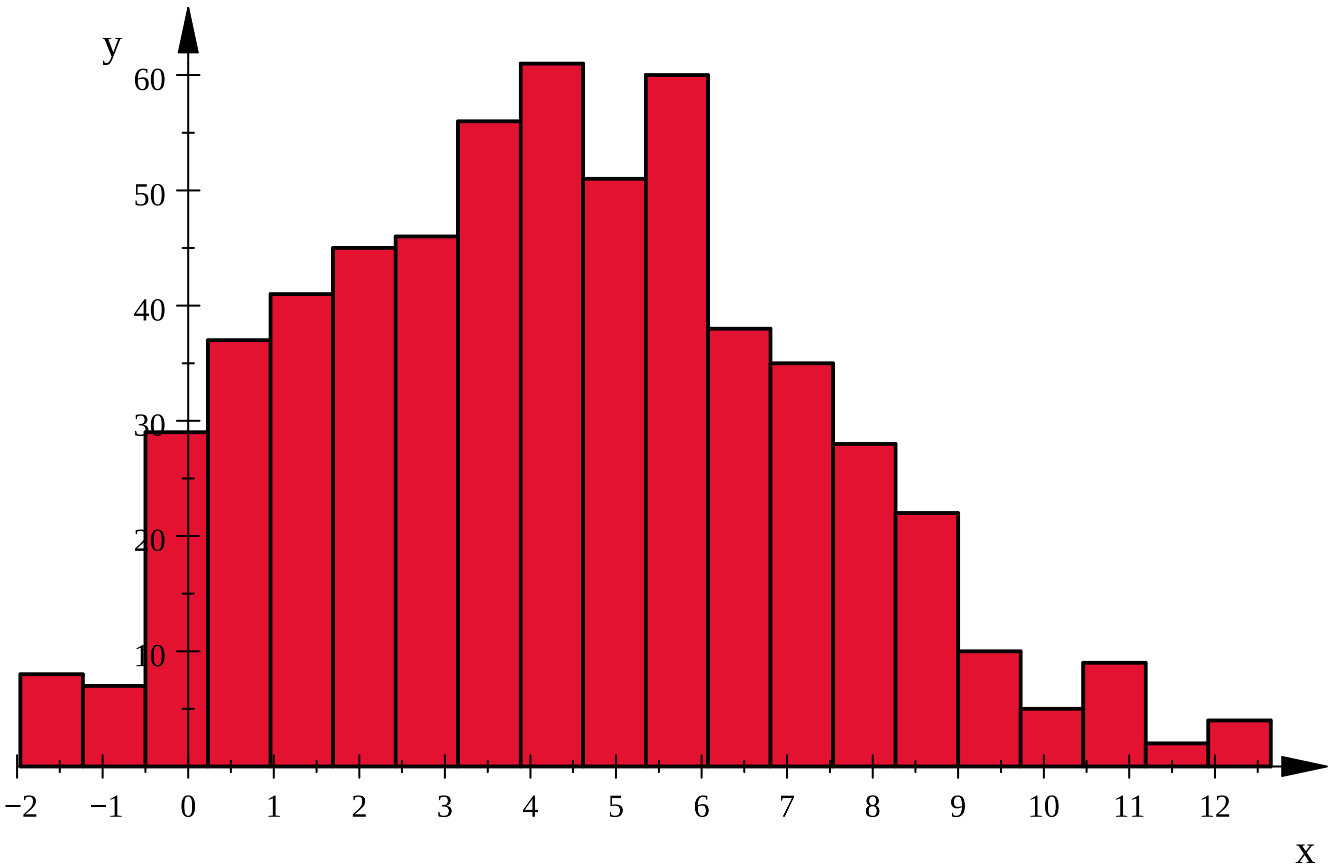
$$_1 = 1.1 \cdot 10^8, B_2 \approx 6.8 \cdot 10^{11}, \sigma = 1875377824$$

11s on 2.4Ghz Opteron (749 + 262).

Group Order : $2^2 \times 3 \times 11243 \times 336181 \times 844957 \times 1866679 \times 6062029$

$$\times 7600843 \times 8046121 \times 8154571 \times 13153633 \times 249436823$$

Efficiency of large B_2



stogram of $\log(g_1/B_1)$ for 594 Cunningham factors found since 2000

$\log 100 \approx 4.6$).

Current records

CM: 66 digits (Dodson, 2005)

1: 58 digits (Zimmermann, 2005)

-1: 48 digits (Kruppa, 2003)

Can you do better?

P+1 record

```
te: Sat, 29 Mar 2003 23:53:51 +0100
om: Alexander Kruppa <alexander.kruppa@stud.tu-muenchen.de>

found this factor of L1849 with gmp-ecm 5.0 P+1 today:

8 = 884764954216571039925598516362554326397028807829

8+1 = 2 5 19 2141 30983 32443 35963 117833 3063121 80105797 2080952771

is beats the previous record, a p39 found by Paul Leyland, by almost
 digits (leading digits 88.. vs 13..).

e cofactor has 330 digits and is probably prime.

st regards,
ex
```

Record Group Order Factor

odson, December 2005, 47-digit factor of $5^{430}+1$:

$$g = 2^2 \times 3 \times 13 \times 347 \times 659 \times 163481 \times 260753$$

$$\times 9520793 \times 25074457 \times 81325590104999$$

$_1 = 260M$, $g \approx 300000 \cdot B_1$

Save/Resume Interface

```
./ecm -save toto -pm1 -mpzmod -x0 2 5000000 < c71
P-ECM 6.1 [powered by GMP 4.2] [P-1]
put number is 131...487 (71 digits)
ing B1=5000000, B2=352526802, polynomial x^24, x0=2
ep 1 took 3116ms
ep 2 took 2316ms
cat toto
THOD=P-1; B1=5000000; N=131...487; X=0x125...19f; CHECKSUM=2287710189;
OGRAM=GMP-ECM 6.1; X0=0x2; WHO=zimmerma@macaron.loria.fr; TIME=...;
./ecm -resume toto 1e7
P-ECM 6.1 [powered by GMP 4.2] [ECM]
suming P-1 residue saved by zimmerma@macaron.loria.fr with GMP-ECM 6.1
put number is 131...487 (71 digits)
ing B1=5000000-10000000, B2=880276332, polynomial x^24
ep 1 took 3076ms
ep 2 took 4304ms
******** Factor found in step 2: 1448595612076564044790098185437
obable prime cofactor 908...651 has 40 digits
```

Library Interface

his provides a direct way to call ECM from a C program:

```
#include "ecm.h"

res = ecm_factor (mpz_t f, mpz_t n, double B1, NULL);
```

sed in Magma since version V2.12 (July 2005).

New P-1 record

ound by Takahiro Nohara (amateur mathematician, Tokyo Electron,
renoble) on June 29, 2006 on a standard PC.

put is 277-digit cofactor from $960^{119} - 1$, $B_1 = 3 \cdot 10^7$, $B_2 = 3 \cdot 10^{10}$:

```
P-ECM 6.0 [powered by GMP 4.1.4] [P-1]
put number is 453...679 (277 digits)
ing B1=30000000, B2=30000000000, polynomial x^60, x0=3595167554
ep 1 took 270553ms
ep 2 took 255414ms
******** Factor found in step 2: 672...541
und probable prime factor of 66 digits:
2038771836751227845696565342450315062141551559473564642434674541
mposite cofactor 674...419 has 211 digits
```

Stage 1 speedup

the duplication formula:

$$z_{2P} = (4x_P z_P)[(x_P - z_P)^2 + d(4x_P z_P)]$$

e has to multiply by $d = (a+2)/4$ where the initial curve in Montgomery rm is:

$$by^2 = x^3 + ax^2 + x.$$

ernstein suggest to use $a = 4d + 2$ and $b = 16d + 18$ with starting point $x = 2 : y = 1)$.

d is a small integer, the product by d costs $O(n)$ instead of $O(M(n))$.

point doubling then costs only 4 full multiplies, instead of 5.

Gaudry's assembly code for REDC

audry wrote assembly code for fused multiply and Montgomery reduction.

hlon, Pentium 4, Opteron.

ze 1 to 20 words.

o to 25% speedup in stage 1 (here 14% on Pentium M):

```
P-ECM 6.1 [powered by GMP 4.2] [ECM]
put number is 952...581 (155 digits)
ing B1=1000000, B2=1045563762, polynomial Dickson(6), sigma=1225316034
ep 1 took 74584ms
ep 2 took 31294ms

P-ECM 6.1.1 [powered by GMP 4.2] [ECM]
put number is 952...581 (155 digits)
ing B1=1000000, B2=1045563762, polynomial Dickson(6), sigma=52552621
ep 1 took 64152ms
ep 2 took 30821ms
```

GMP-ECM 6.1.1

eleased on July 19, 2006.

a few bug fixes wrt version 6.1

includes Gaudry's assembly code:

`configure -with-asm-redc`

ownload at `ecm.gforge.inria.fr`

Summary (1/2)

CM can benefit from state-of-the-art algorithms:

arithmetic modulo n

polynomial arithmetic

hese algorithms might be useful for other problems

everal tricks improve the success expectation

Summary (2/2)

ne main bottleneck remains stage 1.

onsider $B_1 = 10^6$ with $B_2 = 9.7 \cdot 10^8$: we need 957 curves to find a 5-digit number.

ssume stage 1 improves by a factor 2, i.e. we can perform in the same time $_1 = 2 \cdot 10^6$. Then 593 curves are enough, which gives a global gain of 3%.

ssume now that stage 2 improves by a factor 2. Then we can use $_2 = 2.9 \cdot 10^9$ instead, and need only 744 curves, which gives a global ain of 22% only!

Credits

ur sponsors: INRIA Lorraine/LORIA and Lehigh University

enstra, Pollard, Williams for inventing ECM, P-1, P+1

ent, Montgomery for improving them

ranlund, the main author of GMP

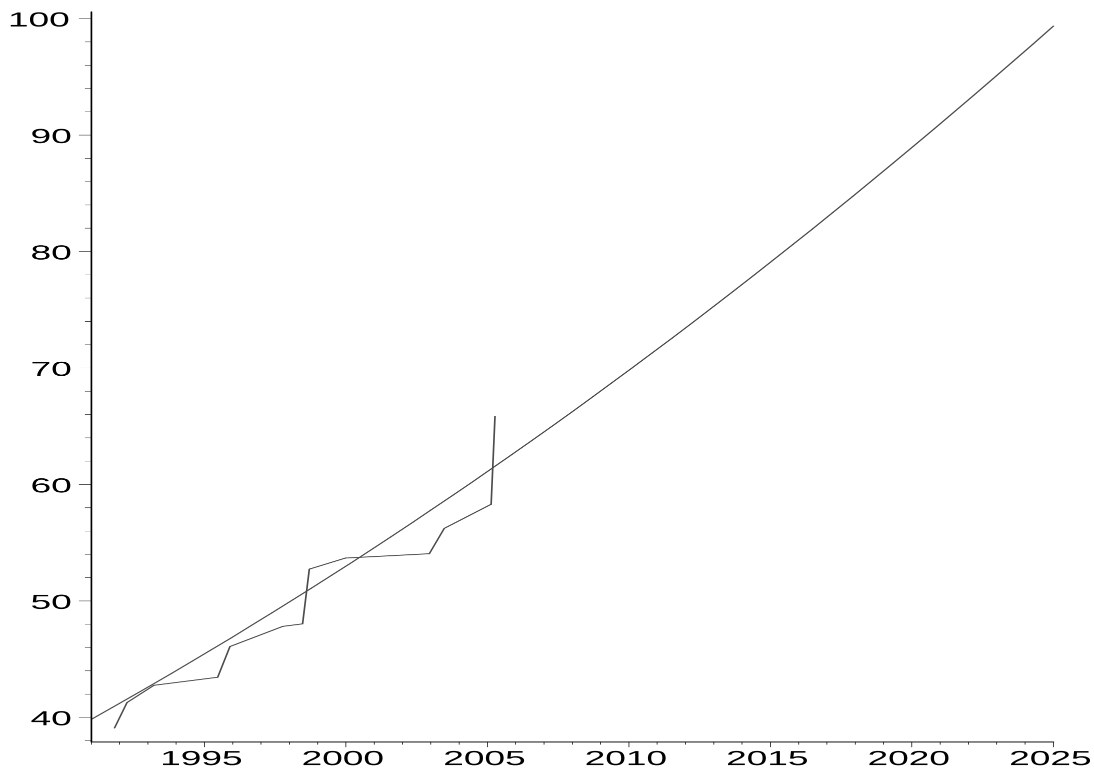
he GMP-ECM co-authors: Gaudry, Fougeron, Fousse, Kruppa, Newman

sers who did or will find nice factors!

ECM records since 1991

ent's formula (D = digits, Y = year):

$$\sqrt{D} = \frac{Y - 1932.3}{9.3}$$



100 digits at ANTS XVII?

Is ECM useful?

ate: Tue, 19 Apr 2005 15:12:29 +0200 (CEST)
rom: hwl@math.leidenuniv.nl (H.W. Lenstra)
ear Paul - Thanks for your message and for letting me
now about the 66 digit ecm record! I always love hearing
bout those records. When ecm was just invented, someone
hom I do not name predicted it would never (NEVER) find
factor of more than 30 digits! All the best - Hendrik