

# Computing Murphy-alpha in the special tower number field sieve algorithm, and applications

Aurore Guillevic

Inria Nancy, France

Séminaire LFANT, May 18, 2021

Joint work with Shashank Singh, IISER Bhopal, India

The Inria logo is written in a stylized, cursive red font.

# Outline

Discrete logarithm

Pairing-friendly curves, DL in  $\mathbb{F}_{p^n}^*$

Key-sizes for pairing-based crypto

# Outline

Discrete logarithm

Pairing-friendly curves, DL in  $\mathbb{F}_{p^n}^*$

Key-sizes for pairing-based crypto

# Discrete logarithm problem

**G** multiplicative group of order  $\ell$

$g$  generator,  $\mathbf{G} = \{1, g, g^2, g^3, \dots, g^{\ell-2}, g^{\ell-1}\}$

Given  $h \in \mathbf{G}$ , find integer  $x \in \{0, 1, \dots, \ell - 1\}$  such that  $h = g^x$ .

Exponentiation easy:  $(g, x) \mapsto g^x$

Discrete logarithm hard in well-chosen groups **G**

Common choices of **G**:

- prime finite field  $\mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$  (1976)
- characteristic 2 field  $\mathbb{F}_{2^n}$  ( $\approx$  1979)
- elliptic curve  $E(\mathbb{F}_p)$  (1985)

## Choosing key sizes

**Symmetric ciphers** (AES): key sizes are 128, 192 or 256 bits.

Perfect symmetric cipher: trying all keys of size  $n$  bits takes  $2^n$  tests

→ **brute-force search**

perfect symmetric cipher with secret key in  $[0, 2^n - 1]$ , of  $n$  bits  $\leftrightarrow n$  bits of security

For DL-based key exchange with  $p, \ell$  of  $\text{length}(p)$ ,  $\text{length}(\ell)$  bits:

$n$  bits of security  $\leftrightarrow$  the best (mathematical) attack should take at least  $2^n$  steps

- what is the fastest attack?
- how much time does it take with respect to  $\text{length}(p)$ ,  $\text{length}(\ell)$ ?

RSA and Diffie–Hellman keys are much larger.

*Cipher suite*: a pair of symmetric and asymmetric ciphers offering the same level of security.

## Discrete log problem

How fast can we invert the exponentiation function  $(g, x) \mapsto g^x$ ?

- $g \in G$  generator, exists always a preimage  $x \in \{1, \dots, \#G\}$
  - naive search, try them all:  $\#G$  tests
  - $O(\sqrt{\#G})$  generic algorithms
  - independent search in each distinct subgroup  
+ Chinese remainder theorem (Pohlig-Hellman)
- choose  $G$  of large prime order (no subgroup)
- complexity of inverting exponentiation in  $O(\sqrt{\#G})$
- **security level 128 bits** means  $\sqrt{\#G} \geq 2^{128}$   
take  $\#G = 2^{256}$   
analogy with symmetric crypto, keylength 128 bits (16 bytes)

## Discrete log problem

How fast can we invert the exponentiation function  $(g, x) \mapsto g^x$ ?

- $g \in G$  generator, exists always a preimage  $x \in \{1, \dots, \#G\}$
  - naive search, try them all:  $\#G$  tests
  - $O(\sqrt{\#G})$  generic algorithms
  - independent search in each distinct subgroup  
+ Chinese remainder theorem (Pohlig-Hellman)
- choose  $G$  of large prime order (no subgroup)
- complexity of inverting exponentiation in  $O(\sqrt{\#G})$
- **security level 128 bits** means  $\sqrt{\#G} \geq 2^{128}$   
take  $\#G = 2^{256}$   
analogy with symmetric crypto, keylength 128 bits (16 bytes)

Use additional structure of  $G$  if any.

## Discrete log problem when $\mathbf{G} = (\mathbb{Z}/p\mathbb{Z})^*$

Index calculus algorithm [Western–Miller 68, Adleman 79],  
prequel of the Number Field Sieve algorithm (NFS)

- $p$  prime,  $(p - 1)/2$  prime,  $\mathbf{G} = (\mathbb{Z}/p\mathbb{Z})^*$ , gen.  $g$ , target  $h$
- get many multiplicative relations in  $\mathbf{G}$
- get one multiplicative relation involving the target  $h$
- take logarithms: linear relations *in the exponents*
- solve a linear system to get discrete logarithms
- get  $x = \log h$



## Index calculus in $(\mathbb{Z}/p\mathbb{Z})^*$ : example

$p = 2039$ ,  $\ell = (p - 1)/2 = 1019$  prime

Smoothness bound  $B = 17$

$\mathcal{F}_{17} = \{2, 3, 5, 7, 11, 13, 17\}$  small primes up to  $B$ ,  $\#\mathcal{F} = 7$

$B$ -smooth integer:  $n = 2^{e_1} 3^{e_2} \dots 17^{e_7}$ ,  $e_i \geq 0$

is  $g^s \bmod p$  smooth?  $1 \leq s \leq 59$  is enough

$$\begin{array}{l}
 g^1 \equiv 7 = 7 \\
 g^5 \equiv 495 = 3^2 \cdot 5 \cdot 11 \\
 g^{30} \equiv 204 = 2^2 \cdot 3 \cdot 17 \\
 g^{44} \equiv 255 = 3 \cdot 5 \cdot 17 \\
 g^{52} \equiv 1088 = 2^6 \cdot 17 \\
 g^{57} \equiv 264 = 2^3 \cdot 3 \cdot 11 \\
 g^{59} \equiv 702 = 2 \cdot 3^3 \cdot 13
 \end{array}
 \rightarrow
 \begin{array}{ccccccc}
 & 2 & 3 & 5 & 7 & 11 & 13 & 17 \\
 \left[ \begin{array}{ccccccc}
 & & & & 1 & & & \\
 & & 2 & 1 & & 1 & & \\
 2 & 1 & & & & & 1 & \\
 & 1 & 1 & & & & 1 & \\
 6 & & & & & & 1 & \\
 3 & 1 & & & 1 & & & \\
 1 & 3 & & & & & 1 & 
 \end{array} \right] \cdot \mathbf{x} = \begin{bmatrix} 1 \\ 5 \\ 30 \\ 44 \\ 52 \\ 57 \\ 59 \end{bmatrix}
 \end{array}$$

$$\mathbf{x} = [325, 259, 664, 1, 861, 995, 140] \bmod 1019$$

$$7^{325} = -2, 7^{259} = -3, 7^{664} = 5, 7^1 = 7, 7^{861} = 11, 7^{995} = 13, 7^{140} = 17$$

## Index calculus in $(\mathbb{Z}/p\mathbb{Z})^*$ : example

$$\mathbf{x} = [325, 259, 664, 1, 861, 995, 140] \bmod 1019$$

$$\mathbf{v} = [1344, 1278, 664, 1, 861, 995, 140] \bmod p - 1 = 2038$$

$$7^{1344} = 2, 7^{1278} = 3, 7^{664} = 5, 7^1 = 7, 7^{861} = 11, 7^{995} = 13, 7^{140} = 17$$

$$\text{Target } h = 314, g^7 \cdot 314 = 405 = 3^4 \cdot 5 \bmod p$$

$$\log_7 314 = -7 + 4 \log_7 3 + \log_7 5 = 1693 \bmod p - 1$$

$$7^{1693} = 314 \bmod p$$

# Index calculus in $(\mathbb{Z}/p\mathbb{Z})^*$

## Multiplicative relations over the **integers**

Smooth integers  $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$ ,  $p_i \leq B$  are quite common  $\rightarrow$  it works

Complexity  $e^{\sqrt{(2+o(1))(\log p)(\log \log p)}}$  (Pomerance 87)

# Index calculus in $(\mathbb{Z}/p\mathbb{Z})^*$

## Multiplicative relations over the **integers**

Smooth integers  $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$ ,  $p_i \leq B$  are quite common  $\rightarrow$  it works

Complexity  $e^{\sqrt{(2+o(1))(\log p)(\log \log p)}}$  (Pomerance 87)

## Improvements in the 80's, 90's:

- Sieve (faster relation collection)
- Smaller integers to factor
- Multiplicative relations in **number fields**
- Better **sparse linear algebra**
- Independent targets  $h$

## Number Field

- 1985: ElGamal, DL in  $GF(p^2)$  with two quadratic number fields
- 1986: Coppersmith–Odlyzko–Schroeppel, DL algorithm in  $GF(p)$
- 1995: Weber–Denny, record computation 85 dd with  $\mathbb{Q}[\sqrt{-2}]$

## Number Field

1985: ElGamal, DL in  $GF(p^2)$  with two quadratic number fields

1986: Coppersmith–Odlyzko–Schroeppel, DL algorithm in  $GF(p)$

1995: Weber–Denny, record computation 85 dd with  $\mathbb{Q}[\sqrt{-2}]$

- If  $p = 1 \pmod{4}$ , exists  $u, v$  s.t.  $p = u^2 + v^2$ ,  $\theta = \sqrt{-1}$
- If  $p = 3 \pmod{8}$ , exists  $u, v$  s.t.  $p = u^2 + 2v^2$ ,  $\theta = \sqrt{-2}$
- If  $p = 7 \pmod{8}$ , exists  $u, v$  s.t.  $p = u^2 - 2v^2$ ,  $\theta = \sqrt{2}$

and  $|u|, |v| < \sqrt{p}$

$u/v \equiv m \pmod{p}$  and  $m^2 + s = 0 \pmod{p}$

## Number Field

1985: ElGamal, DL in  $\text{GF}(p^2)$  with two quadratic number fields

1986: Coppersmith–Odlyzko–Schroepel, DL algorithm in  $\text{GF}(p)$

1995: Weber–Denny, record computation 85 dd with  $\mathbb{Q}[\sqrt{-2}]$

- If  $p \equiv 1 \pmod{4}$ , exists  $u, v$  s.t.  $p = u^2 + v^2$ ,  $\theta = \sqrt{-1}$
- If  $p \equiv 3 \pmod{8}$ , exists  $u, v$  s.t.  $p = u^2 + 2v^2$ ,  $\theta = \sqrt{-2}$
- If  $p \equiv 7 \pmod{8}$ , exists  $u, v$  s.t.  $p = u^2 - 2v^2$ ,  $\theta = \sqrt{2}$

and  $|u|, |v| < \sqrt{p}$

$u/v \equiv m \pmod{p}$  and  $m^2 + s = 0 \pmod{p}$

Define a map from  $\mathbb{Z}[\theta]$  to  $\mathbb{Z}/p\mathbb{Z}$

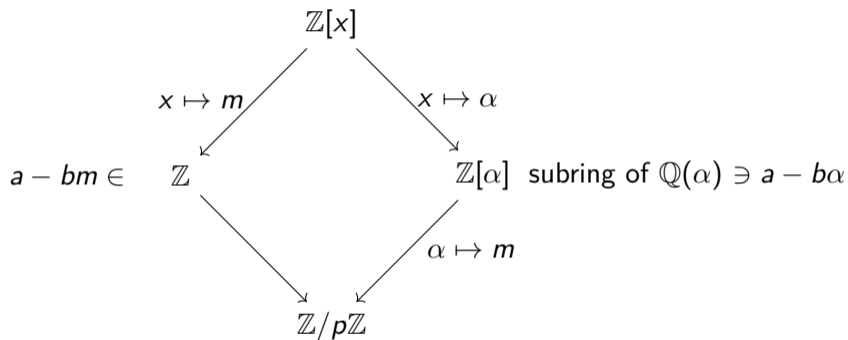
$$\phi: \mathbb{Z}[\theta] \rightarrow \mathbb{Z}/p\mathbb{Z}$$

$$\theta \mapsto m \pmod{p} \text{ where } m = u/v, \quad m^2 + s = 0 \pmod{p}$$

ring homomorphism  $\phi(a + b\theta) = a + bm$

$$\underbrace{\phi(a + b\theta)}_{\substack{\text{factor in} \\ \mathbb{Z}[\theta]}} = a + bm = (a + b \underbrace{u/v}_{=m}) = (\underbrace{av + bu}_{\text{factor in } \mathbb{Z}})v^{-1} \pmod{p}$$

## Commutative diagram for NFS





## Example in $\mathbb{Z}[\sqrt{2}]$

$p = 2039 = 7 \pmod{8}$ ,  $\ell = (p - 1)/2 = 1019$  prime

$$p = 3^2 - 2 \cdot 32^2$$

$\max(|a|, |b|) = A = 13$ ,  $B = 13$  smoothness bound

## Example in $\mathbb{Z}[\sqrt{2}]$

$p = 2039 = 7 \pmod{8}$ ,  $\ell = (p - 1)/2 = 1019$  prime

$$p = 3^2 - 2 \cdot 32^2$$

$\max(|a|, |b|) = A = 13$ ,  $B = 13$  smoothness bound

### Rational side

$\mathcal{F}_0 = \{2, 3, 5, 7, 11, 13\}$  small primes up to 13,  $\#\mathcal{F}_0 = 6$

$$g(x) = 32x - 3$$

## Example in $\mathbb{Z}[\sqrt{2}]$

$$p = 2039 = 7 \pmod{8}, \ell = (p - 1)/2 = 1019 \text{ prime}$$

$$p = 3^2 - 2 \cdot 32^2$$

$$\max(|a|, |b|) = A = 13, B = 13 \text{ smoothness bound}$$

### Rational side

$$\mathcal{F}_0 = \{2, 3, 5, 7, 11, 13\} \text{ small primes up to } 13, \#\mathcal{F}_0 = 6$$

$$g(x) = 32x - 3$$

### Algebraic side: think about the complex numbers in $\mathbb{C}$

$$-i(1+i)^2 = 2, (2+i)(2-i) = 5, (2+3i)(2-3i) = 13$$

$$2 = (\theta)^2, 7 = (\theta+3)(\theta-3)$$

$$\mathcal{F}_1 = \{(\theta), (\theta+3), (\theta-3)\} \text{ small prime ideals above } 2, 7, \#\mathcal{F}_1 = 3$$

“primes” of norm up to  $B$

$$f(x) = x^2 - 2$$

## Example in $\mathbb{Z}[\sqrt{2}]$

$$p = 2039 = 7 \pmod{8}, \ell = (p - 1)/2 = 1019 \text{ prime}$$

$$p = 3^2 - 2 \cdot 32^2$$

$$\max(|a|, |b|) = A = 13, B = 13 \text{ smoothness bound}$$

### Rational side

$$\mathcal{F}_0 = \{2, 3, 5, 7, 11, 13\} \text{ small primes up to } 13, \#\mathcal{F}_0 = 6$$

$$g(x) = 32x - 3$$

### Algebraic side: think about the complex numbers in $\mathbb{C}$

$$-i(1+i)^2 = 2, (2+i)(2-i) = 5, (2+3i)(2-3i) = 13$$

$$2 = (\theta)^2, 7 = (\theta+3)(\theta-3)$$

$$\mathcal{F}_1 = \{(\theta), (\theta+3), (\theta-3)\} \text{ small prime ideals above } 2, 7, \#\mathcal{F}_1 = 3$$

“primes” of norm up to  $B$

$$f(x) = x^2 - 2$$

### Units

$$\mathcal{U}_1 = \{(\theta+1)\} \text{ fundamental unit}$$

## Example in $\mathbb{Z}[\sqrt{-2}]$

$a, b$	$av + bu = \text{factor in } \mathbb{Z}$	$a^2 + 2b^2$	factor in $\mathbb{Z}[\sqrt{-2}]$	units
-9, 4	$-300 = -2^2 \cdot 3 \cdot 5^2$	$49 = 7^2$	$(\theta - 3)^2$	$-(\theta + 1)^2$
-5, 3	$-169 = -13^2$	$7 = 7$	$(\theta - 3)$	$(\theta + 1)^2$
-3, 1	$-99 = -3^2 \cdot 11$	$7 = 7$	$(\theta + 3)$	-1
-1, 0	$-32 = -2^5$	$1 =$		-1
-1, 1	$-35 = -5 \cdot 7$	$-1 =$		$-(\theta + 1)$
0, 1	$-3 = -3$	$-2 = -2$	$(\theta)$	-1
1, 0	$32 = 2^5$	$1 =$		1
1, 2	$26 = 2 \cdot 13$	$-7 = -7$	$(\theta + 3)$	$-(\theta + 1)^{-1}$
2, 3	$55 = 5 \cdot 11$	$-14 = -2 \cdot 7$	$(\theta)(\theta - 3)$	1
3, 2	$90 = 2 \cdot 3^2 \cdot 5$	$1 =$		$(\theta + 1)^{-2}$
4, 1	$125 = 5^3$	$14 = 2 \cdot 7$	$(\theta)(\theta + 3)$	$(\theta + 1)^{-1}$
12, 11	$351 = 3^3 \cdot 13$	$-98 = -2 \cdot 7^2$	$(\theta)(\theta - 3)^2$	-1

## Example in $\mathbb{Z}[i]$ : Matrix

Build the matrix of relations:

- one row per  $(a, b)$  pair s.t. both norms are smooth
- one column per prime of  $\mathcal{F}_{\text{rat}}$
- one column for  $1/V$
- one column per prime ideal of  $\mathcal{F}_{\text{alg}}$
- one column per unit  $(-1, i)$
- store the exponents

## Example in $\mathbb{Z}[\sqrt{-2}]$

$$M = \begin{array}{cccccccc} & 2 & 3 & 5 & 7 & 11 & 13 & 1/v & \theta & \theta + 3 & \theta - 3 & \theta + 1 \\ \left[ \begin{array}{cccccccc} 2 & 1 & 2 & & & & & -1 & & & -2 & -2 \\ & & & & & & 2 & -1 & & & -1 & -2 \\ & & 2 & & & 1 & & -1 & & -1 & & \\ 5 & & & & & & & -1 & & & & \\ & & & 1 & 1 & & & -1 & & & & -1 \\ & 1 & & & & & & -1 & -1 & & & \\ 5 & & & & & & & -1 & & & & \\ 1 & & & & & & 1 & -1 & & -1 & & 1 \\ & & & 1 & & 1 & & -1 & -1 & & -1 & \\ 1 & 2 & 1 & & & & & -1 & & & & 2 \\ & & & 3 & & & & -1 & -1 & -1 & & 1 \\ & & 3 & & & & 1 & -1 & -1 & & -2 & \end{array} \right] \end{array}$$

$$\mathbf{x} = [1, 515, 140, 301, 335, 928, 5, 510, 341, 979, 436]$$

$$\mathbf{x}/301 = [325, 259, 664, 1, 861, 995, 606, 672, 773, 247, 59]$$

We had before:  $\mathbf{x} = [325, 259, 664, 1, 861, 995, 140] \pmod{1019}$

# Number Field Sieve

Since 1993 (Gordon, Schirokauer):

$$L_p(1/3, c) = e^{(c+o(1))(\log p)^{1/3}(\log \log p)^{2/3}}$$

- polynomial selection
- **relation collection**  $L_p(1/3, 1.923)$   
sieve to enumerate efficiently  $(a, b)$  pairs
- **sparse linear algebra**  $L_p(1/3, 1.923)$   
compute right kernel mod prime  $\ell$ , block-Wiedemann alg.
- individual discrete logarithm

Latest record computation: 795-bit (240dd)  $p = \text{RSA-240} + 49204$  prime,  
 $\ell = (p - 1)/2$  prime

Boudot, Gaudry, Guillevic, Heninger, Thomé, Zimmermann, Crypto'2020



## Variants of NFS: Complexities

large characteristic  $p = L_{p^n}(\alpha)$ ,  $\alpha > 2/3$ :

---

$(64/9)^{1/3} \simeq 1.923$  NFS

special  $p$ :

$(32/9)^{1/3} \simeq 1.526$  SNFS

medium characteristic  $p = L_{p^n}(\alpha)$ ,  $1/3 < \alpha < 2/3$ :

---

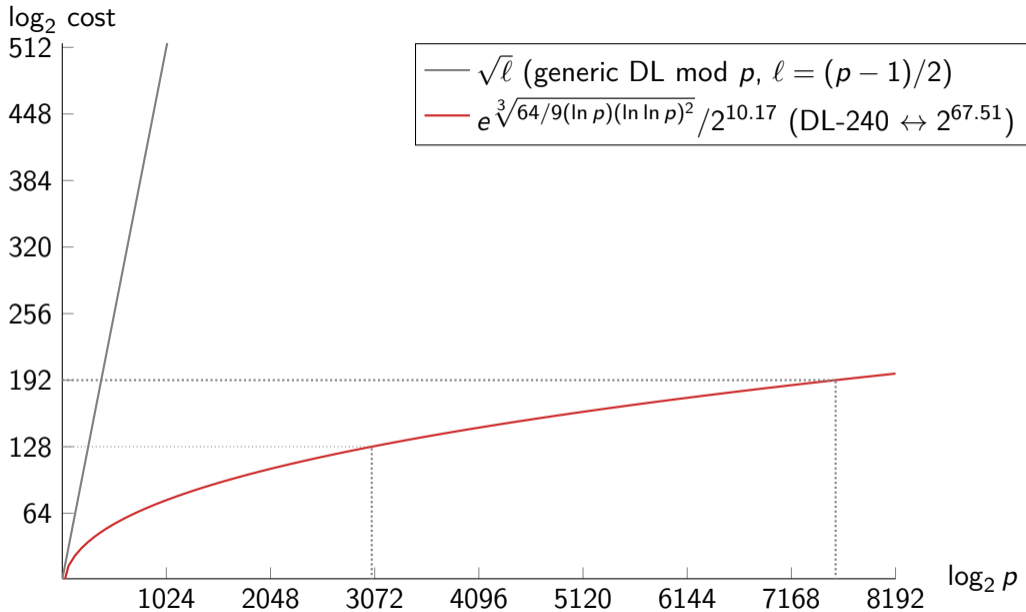
$(96/9)^{1/3} \simeq 2.201$  prime  $n$  NFS-HD (Conjugation)

$(48/9)^{1/3} \simeq 1.747$  composite  $n$  (Kim-Barbulescu 2016),  
best case of TNFS: when parameters fit perfectly

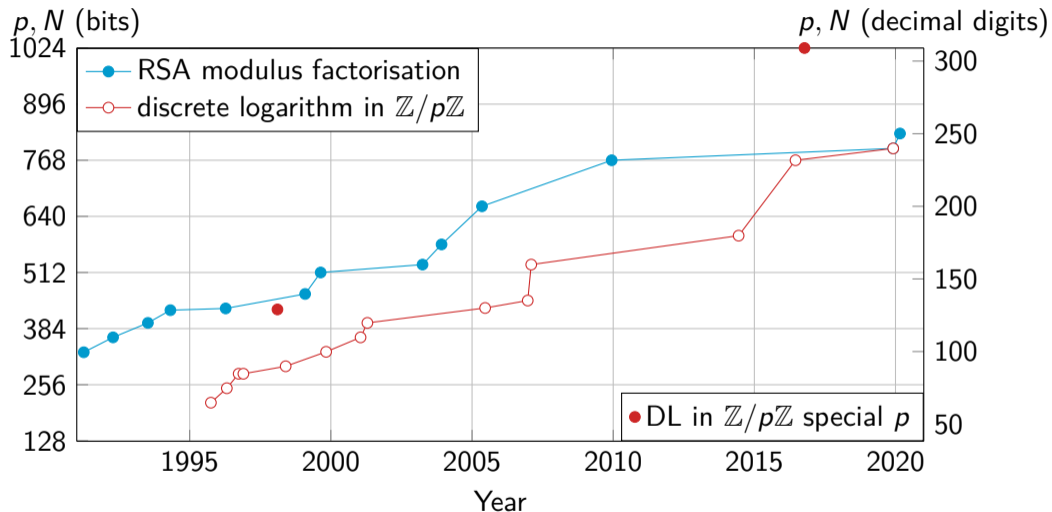
special  $p$ :

$(64/9)^{1/3} \simeq 1.923$  NFS-HD+Joux-Pierrot'13

$(32/9)^{1/3} \simeq 1.526$  composite  $n$ , best case of STNFS (Kim-Barbulescu 2016)



# Record computations



# Outline

Discrete logarithm

Pairing-friendly curves, DL in  $\mathbb{F}_{p^n}^*$

Key-sizes for pairing-based crypto

## Cryptographic pairing: black-box properties

$(\mathbf{G}_1, +)$ ,  $(\mathbf{G}_2, +)$ ,  $(\mathbf{G}_T, \cdot)$  three cyclic groups of large prime order  $r$

Bilinear Pairing: map  $e : \mathbf{G}_1 \times \mathbf{G}_2 \rightarrow \mathbf{G}_T$

1. bilinear:  $e(P_1 + P_2, Q) = e(P_1, Q) \cdot e(P_2, Q)$ ,  $e(P, Q_1 + Q_2) = e(P, Q_1) \cdot e(P, Q_2)$
2. non-degenerate:  $e(g_1, g_2) \neq 1$  for  $\langle g_1 \rangle = \mathbf{G}_1$ ,  $\langle g_2 \rangle = \mathbf{G}_2$
3. efficiently computable.

Mostly used in practice:

$$e([a]P, [b]Q) = e([b]P, [a]Q) = e(P, Q)^{ab} .$$

$\leadsto$  Many applications in asymmetric cryptography

(identity-based encryption, short signatures, NIZK, ZK-SNARK...)

# Pairing-based cryptography

## Weil or Tate pairing on an elliptic curve


Discrete logarithm problem with one more dimension.

$$e: E(\mathbb{F}_p)[\ell] \times E(\mathbb{F}_{p^n})/\ell E(\mathbb{F}_{p^n}) \longrightarrow \mathbb{F}_{p^n}^*, \quad e([a]P, [b]Q) = e(P, Q)^{ab}$$

# Pairing-based cryptography

## Weil or Tate pairing on an elliptic curve

Discrete logarithm problem with one more dimension.

$$e: E(\mathbb{F}_p)[\ell] \times E(\mathbb{F}_{p^n})/\ell E(\mathbb{F}_{p^n}) \longrightarrow \mathbb{F}_{p^n}^*, \quad e([a]P, [b]Q) = e(P, Q)^{ab}$$


### Attacks

- discrete logarithm computation in  $E(\mathbb{F}_p)$  : hard problem (exponential, in  $O(\sqrt{\ell})$ )
- inversion of  $e$  : hard problem (exponential)
- discrete logarithm computation in  $\mathbb{F}_{p^n}^*$  : **easier, subexponential**  $\rightarrow$  take a large enough field

## Pairing-friendly curves are special

$\ell \mid p^n - 1$ ,  $\mathbf{G}_T \subset \mathbb{F}_{p^n}$ ,  $n$  is minimal : **embedding degree**

Tate Pairing:  $e : \mathbf{G}_1 \times \mathbf{G}_2 \rightarrow \mathbf{G}_T$

When  $n$  is small, the curve is *pairing-friendly*.

This is very rare: usually  $\log n \sim \log \ell$  ([Balasubramanian Koblitz]).



## Pairing-friendly curves are special

$\ell \mid p^n - 1$ ,  $\mathbf{G}_T \subset \mathbb{F}_{p^n}$ ,  $n$  is minimal : **embedding degree**

Tate Pairing:  $e : \mathbf{G}_1 \times \mathbf{G}_2 \rightarrow \mathbf{G}_T$

When  $n$  is small, the curve is *pairing-friendly*.

This is very rare: usually  $\log n \sim \log \ell$  ([Balasubramanian Koblitz]).

Barreto-Naehrig (BN),  $n = 12$ :

$$r(x) = 36x^4 + 36x^3 + 18x^2 + 6x + 1, \quad p(x) = 36x^4 + 36x^3 + 24x^2 + 6x + 1$$

Barreto-Lynn-Scott (BLS),  $n = 12$ :

$$r(x) = x^4 - x^2 + 1, \quad p(x) = (x - 1)^2 r(x) / 3 + x$$

$$D = -3, \quad j = 0, \quad \mathbf{G}_T \subset \mathbb{F}_{p^{12}}$$

$p$  is special

Examples:

Ethereum blockchain has a 254-bit BN curve with  $x = 0x44e992b44a6909f1$ ,

Zcash cryptocurrency has a 381-bit BLS curve with  $x = -0xd201000000010000$ .

# Outline

Discrete logarithm

Pairing-friendly curves, DL in  $\mathbb{F}_{p^n}^*$

Key-sizes for pairing-based crypto

## Discrete Log in $\mathbb{F}_{p^n}$

$\mathbb{F}_{p^n}$  much less investigated than  $\mathbb{F}_p$  or integer factorization.  
Much better results in pairing-related fields

## Discrete Log in $\mathbb{F}_{p^n}$

$\mathbb{F}_{p^n}$  much less investigated than  $\mathbb{F}_p$  or integer factorization.

Much better results in pairing-related fields

- Special NFS in  $\mathbb{F}_{p^n}$ : Joux–Pierrot 2013
- Tower NFS (TNFS): Barbulescu Gaudry Kleinjung 2015
- Extended Tower NFS: Kim–Barbulescu, Kim–Jeong, Sarkar–Singh 2016
- Tower of number fields

Use more structure: subfields

## Special Tower NFS

$\mathbb{F}_{p^{12}}$ , subfield  $\mathbb{F}_{p^2}$  defined by  $y^2 + 1$

Idea:  $a + bx$  in NFS  $\rightarrow (a_0 + a_1y) + (b_0 + b_1y)x$  in TNFS

Integers to factor are **much smaller**

- factors integer  $\text{Norm}_f = \text{Res}(\text{Res}(\mathbf{a} + \mathbf{b}x, f_y(x)), y^2 + 1)$
- factors integer  $\text{Norm}_g = \text{Res}(\text{Res}(\mathbf{a} + \mathbf{b}x, g_y(x)), y^2 + 1)$

Res = resultant of polynomials

## Special Tower NFS

$\mathbb{F}_{p^{12}}$ , subfield  $\mathbb{F}_{p^2}$  defined by  $y^2 + 1$

Idea:  $a + bx$  in NFS  $\rightarrow (a_0 + a_1y) + (b_0 + b_1y)x$  in TNFS

Integers to factor are **much smaller**

- factors integer  $\text{Norm}_f = \text{Res}(\text{Res}(\mathbf{a} + \mathbf{b}x, f_y(x)), y^2 + 1)$
- factors integer  $\text{Norm}_g = \text{Res}(\text{Res}(\mathbf{a} + \mathbf{b}x, g_y(x)), y^2 + 1)$

Res = resultant of polynomials

$p = p(s)$  is special

## Special Tower NFS

$\mathbb{F}_{p^{12}}$ , subfield  $\mathbb{F}_{p^2}$  defined by  $y^2 + 1$

Idea:  $a + bx$  in NFS  $\rightarrow (a_0 + a_1y) + (b_0 + b_1y)x$  in TNFS

Integers to factor are **much smaller**

- factors integer  $\text{Norm}_f = \text{Res}(\text{Res}(\mathbf{a} + \mathbf{b}x, f_y(x)), y^2 + 1)$
- factors integer  $\text{Norm}_g = \text{Res}(\text{Res}(\mathbf{a} + \mathbf{b}x, g_y(x)), y^2 + 1)$

Res = resultant of polynomials

$p = p(s)$  is special

Index calculus in the 80's: implemented *before* complexity known

TNFS: complexity known, implementation just started for  $\text{GF}(p^6)$

https:

[//listserv.nodak.edu/cgi-bin/wa.exe?A2=1302&L=NMBRTHRY&D=0&P=3444089](https://listserv.nodak.edu/cgi-bin/wa.exe?A2=1302&L=NMBRTHRY&D=0&P=3444089)

Discrete logarithm in  $\text{GF}(p^6)$  of 521 bits with the Tower NFS algorithm

Gabrielle De Micheli, Pierrick Gaudry, Cécile Pierrot

## Special Tower NFS

1. Polynomial selection: choose 3 polynomials  $h, f, g$
2. Relation collection: obtain many smooth norms of  
 $\mathbf{a} + \mathbf{b}\theta_f = (a_0 + a_1\tau + \dots + a_i\tau^i) + (b_0 + b_1\tau + \dots + b_i\tau^i)\theta_f, \mathbf{a} + \mathbf{b}\theta_g$
3. Filtering step of the matrix (apply Galois automorphisms if any)
4. Linear algebra
5. Individual discrete logarithm

Are the norms as smooth as integers of the same size?



## Ranking polynomials: Murphy's $\alpha$ and $E$

B. A. Murphy, PhD thesis, 1999

Input: irreducible polynomials  $f, g$ , sharing root  $m \bmod p$  ( $p \mid \text{Res}(f, g)$ )

- $\alpha(f, B_0)$ : bias in smoothness between norms and integers (up to some bound  $B_0$ )  
 $\alpha(f, B_0), \alpha(g, B_0) < 0$  wanted
- $E(f, g, B_f, B_g, \text{area} = A^2/2)$ : estimation of the yield of polynomials  
 $B_f, B_g$  smoothness bounds of  $f, g$  sides  
How many relations would  $(f, g)$  produce?
- Rank many  $(f_i, g_j)$ , choose the best pair

## Ranking polynomials: Murphy's $\alpha$ and $E$

B. A. Murphy, PhD thesis, 1999

Input: irreducible polynomials  $f, g$ , sharing root  $m \bmod p$  ( $p \mid \text{Res}(f, g)$ )

- $\alpha(f, B_0)$ : bias in smoothness between norms and integers (up to some bound  $B_0$ )  
 $\alpha(f, B_0), \alpha(g, B_0) < 0$  wanted
- $E(f, g, B_f, B_g, \text{area} = A^2/2)$ : estimation of the yield of polynomials  
 $B_f, B_g$  smoothness bounds of  $f, g$  sides  
How many relations would  $(f, g)$  produce?
- Rank many  $(f_i, g_i)$ , choose the best pair

Generalization to the TNFS setting:

- $\alpha(h, f), \alpha(h, g)$   
SageMath & Magma code, generalization from `cado-nfs`  $\alpha$   
(Bai, Gaudry, Hanrot, Thomé, Zimmermann)
- Monte-Carlo simulation for Murphy's  $E$

## Polynomial selection

- 10% total time in a record computation
- Which criteria?

Wanted: norms more often  $B$ -smooth than integers of the same size

Very fast choice

<https://gitlab.inria.fr/tnfs-alpha/alpha>

```
from tnfs.alpha.alpha2d import alpha2d
```

```
ZZx.<x> = ZZ[]
```

```
alpha2d(x2-2, 2000)
```

```
alpha2d(x2+1, 2000)
```

```
alpha2d(x2+2, 2000)
```

$f$	$\alpha(f, 2000)$
$x^2 - 2$	1.752860
$x^2 + 1$	1.366415
$x^2 + 2$	1.116395
...	
$8x^2 + 3x + 10$	-0.097068

## Smoothness probabilities: definition of $\alpha$

$$\alpha(f, B) = \sum_{\substack{\ell \text{ prime} \\ \ell \leq B}} \ln \ell \left( \mathbb{E}(\text{val}_\ell(n), n \in \mathbb{Z}) - \mathbb{E}(\text{val}_\ell(\text{Res}_x(f(x), a + bx), (a, b) \in I \times J)) \right)$$

$f$  with many roots mod many small primes  $\ell$  wanted

If  $B \rightarrow \infty$ ,  $\alpha$  converges (Th. Barbulescu–Lachand 2017)

## Smoothness probabilities: definition of $\alpha$

$$\alpha(f, B) = \sum_{\substack{\ell \text{ prime} \\ \ell \leq B}} \ln \ell \left( \mathbb{E}(\text{val}_\ell(n), n \in \mathbb{Z}) - \mathbb{E}(\text{val}_\ell(\text{Res}_x(f(x), a + bx), (a, b) \in I \times J)) \right)$$

$f$  with many roots mod many small primes  $\ell$  wanted

If  $B \rightarrow \infty$ ,  $\alpha$  converges (Th. Barbulescu–Lachand 2017)

Average valuation at  $\ell = 2$  for random integers  $n$  :

- $1/2$  chance to be even, if even:  $\text{val}_2 n \geq 1$
- $1/4$  chance to be multiple of 4, in this case:  $\text{val}_2 n \geq 2$
- $1/2^i$  chance to be multiple of  $2^i$ , in this case :  $\text{val}_2 n \geq i$

## Smoothness probabilities: definition of $\alpha$

$$\alpha(f, B) = \sum_{\substack{\ell \text{ prime} \\ \ell \leq B}} \ln \ell \left( \mathbb{E}(\text{val}_\ell(n), n \in \mathbb{Z}) - \mathbb{E}(\text{val}_\ell(\text{Res}_x(f(x), a + bx), (a, b) \in I \times J)) \right)$$

$f$  with many roots mod many small primes  $\ell$  wanted

If  $B \rightarrow \infty$ ,  $\alpha$  converges (Th. Barbulescu–Lachand 2017)

Average valuation at  $\ell = 2$  for random integers  $n$  :

- $1/2$  chance to be even, if even:  $\text{val}_2 n \geq 1$
- $1/4$  chance to be multiple of 4, in this case:  $\text{val}_2 n \geq 2$
- $1/2^i$  chance to be multiple of  $2^i$ , in this case :  $\text{val}_2 n \geq i$

With  $\text{val}_\ell(n)$  a random variable in  $\mathbb{N}$ ,

$$\mathbb{E}(\text{val}_\ell(n), n \in \mathbb{Z}) = \sum_{i=1}^{\infty} i \Pr[\text{val}_\ell(n) = i] = \sum_{i=1}^{\infty} \Pr[\text{val}_\ell(n) \geq i]$$

## Smoothness probabilities: definition of $\alpha$

$$\alpha(f, B) = \sum_{\substack{\ell \text{ prime} \\ \ell \leq B}} \ln \ell \left( \mathbb{E}(\text{val}_\ell(n), n \in \mathbb{Z}) - \mathbb{E}(\text{val}_\ell(\text{Res}_x(f(x), a + bx), (a, b) \in I \times J)) \right)$$

$f$  with many roots mod many small primes  $\ell$  wanted

If  $B \rightarrow \infty$ ,  $\alpha$  converges (Th. Barbulescu–Lachand 2017)

Average valuation at  $\ell = 2$  for random integers  $n$  :

- $1/2$  chance to be even, if even:  $\text{val}_2 n \geq 1$
- $1/4$  chance to be multiple of 4, in this case:  $\text{val}_2 n \geq 2$
- $1/2^i$  chance to be multiple of  $2^i$ , in this case :  $\text{val}_2 n \geq i$

With  $\text{val}_\ell(n)$  a random variable in  $\mathbb{N}$ ,

$$\mathbb{E}(\text{val}_\ell(n), n \in \mathbb{Z}) = \sum_{i=1}^{\infty} i \Pr[\text{val}_\ell(n) = i] = \sum_{i=1}^{\infty} \Pr[\text{val}_\ell(n) \geq i]$$

$$= \frac{1}{\ell} + \frac{1}{\ell^2} + \frac{1}{\ell^3} + \dots + \frac{1}{\ell^k} + \dots = \sum_{k=1}^{\infty} \frac{1}{\ell^k} = \lim_{k \rightarrow \infty} \frac{1 - 1/\ell^k}{1 - 1/\ell} - 1 = \frac{1}{\ell - 1} \text{ with } \ell \text{ prime}$$

## Computing $\alpha(f)$

$$\mathbb{E}(\text{val}_\ell(\text{Res}_x(f, a + bx)), (a, b) \in I \times J)$$

Let  $n_{\ell^k}$  the number of distinct roots of  $f \bmod \ell^k$

$$\text{val}_\ell(f) = \text{val}_\ell(\text{Res}_x(f, a + bx)) = \sum_{k=1}^{\infty} \frac{\ell}{\ell+1} \frac{n_{\ell^k}}{\ell^k}$$

### Lemma (Hensel)

Let  $f(x)$  a polynomial of degree  $\geq 2$ ,  $r$  a simple root mod  $\ell$  ( $f'(r) \not\equiv 0 \pmod{\ell}$ ).  
For all  $k > 1$ ,  $f(x)$  has a unique root  $r_k \in \mathbb{Z}/\ell^k\mathbb{Z}$  above  $r$ .

If  $n_{\ell^k} = n_\ell$  for all  $k \geq 1$ ,  $\text{val}_\ell(f) = \ell n_\ell / (\ell^2 - 1)$ .

If  $\ell \nmid \text{disc}(f)$ ,

$$\alpha_\ell(f) = \left( \frac{1}{\ell-1} - \frac{n_\ell}{\ell^2-1} \right) \log \ell$$



## Examples: $\alpha$ looks like computing the factor bases

$\ell$	$f = x^2 - 2 \pmod{\ell}$	$\ell$	$f = x^2 + 1 \pmod{\ell}$	$\ell$	$f = 8x^2 + 3x + 10 \pmod{\ell}$
2	$x^2$ bad prime	2	$(x + 1)^2$ bad prime	2	$x^2$ projective root
7	$(x - 3)(x + 3)$	5	$(x + 2)(x - 2)$	3	$(x + 1)(x - 1)$
17	$(x - 6)(x + 6)$	13	$(x + 5)(x - 5)$	5	$x(x - 4)$
23	$(x - 5)(x + 5)$	17	$(x + 4)(x - 4)$	7	$(x - 3)(x - 1)$
31	$(x - 8)(x + 8)$	29	$(x + 12)(x - 12)$	13	$(x - 8)(x - 3)$
41	$(x - 17)(x + 17)$	37	$(x + 6)(x - 6)$	47	$(x - 22)(x - 7)$
47	$(x - 7)(x + 7)$	41	$(x + 9)(x - 9)$	53	$(x - 31)(x - 15)$
71	$(x - 12)(x + 12)$	53	$(x + 23)(x - 23)$	67	$(x - 51)(x - 24)$
73	$(x - 32)(x + 32)$	61	$(x + 11)(x - 11)$	73	$(x - 70)(x - 30)$
79	$(x - 9)(x + 9)$	73	$(x + 27)(x - 27)$	79	$(x - 43)(x - 6)$
89	$(x - 25)(x + 25)$	89	$(x + 34)(x - 34)$	83	$(x - 75)(x - 18)$
97	$(x - 14)(x + 14)$	97	$(x + 22)(x - 22)$	89	$(x - 63)(x - 59)$
$\alpha(f, 100) = 1.5981$		$\alpha(f, 100) = 1.2752$		$\alpha(f, 100) = -0.1432$	

## $\alpha$ for bad primes (multiple roots)

If  $\ell \mid \text{disc}(f)$ ,  $\ell$  is a *bad prime*.

Exists  $e$  such that  $n_{\ell^k} = n_{\ell^e}$  for all  $k \geq e$ .

Computes such  $e$  and  $n_{\ell^k}$  for all  $k \leq e$ .

Reverse-engineering of `cado-nfs/polyselect/auxiliary.c`

Ugly details in the appendix of the paper.

## Simulation without sieving

Polynomial selection: for many pairs  $(f, g)$

- compute  $\alpha(h, f), \alpha(h, g)$  (w.r.t. subfield) **bias in smoothness**
- select polys  $f, g$  with negative bias  $\alpha(f), \alpha(g)$  if possible
- **Monte-Carlo** simulation with  $10^6$  random samples from  $\mathcal{S} = \{(a_0 + a_1y + \dots + a_dy^d) + (b_0 + b_1y + \dots + b_dy^d)x, |a_i|, |b_j| < A\}$   
For each sample:
  1. compute its algebraic norm  $N_f, N_g$  in each number field
  2. smoothness probability  $(N_f, \alpha_f), (N_g, \alpha_g)$  with Dickman- $\rho$
- Average smoothness probability of samples
  - estimation of the total number of possible relations in  $\mathcal{S}$
  - **Murphy's  $E$  for TNFS**

## Simulation without sieving

Polynomial selection: for many pairs  $(f, g)$

- compute  $\alpha(h, f), \alpha(h, g)$  (w.r.t. subfield) **bias in smoothness**
- select polys  $f, g$  with negative bias  $\alpha(f), \alpha(g)$  if possible
- **Monte-Carlo** simulation with  $10^6$  random samples from  $\mathcal{S} = \{(a_0 + a_1y + \dots + a_dy^d) + (b_0 + b_1y + \dots + b_dy^d)x, |a_i|, |b_j| < A\}$   
For each sample:
  1. compute its algebraic norm  $N_f, N_g$  in each number field
  2. smoothness probability  $(N_f, \alpha_f), (N_g, \alpha_g)$  with Dickman- $\rho$
- Average smoothness probability of samples
  - estimation of the total number of possible relations in  $\mathcal{S}$
  - **Murphy's  $E$  for TNFS**

dichotomy to approach the best balanced parameters

smoothness bound  $B$ , coefficient bound  $A$ .

→ refinement of Barbulescu–Duquesne technique [BD19 Journal of Cryptology]

# Murphy's $\alpha$ function

$\alpha(f)$  for NFS estimates the bias in smoothness

Algebraic norms in  $K_f = \mathbb{Q}[x]/(f(x))$  of  $\log_2 N_f$  bits have same smoothness proba as integers of  $\log_2 N_f + \alpha(f)/\log(2)$  bits

$\rightarrow \alpha(f) < 0$  wanted

$\alpha(f)$  computes the exact number of roots of  $f(x) \bmod q^k$

for all primes  $q < 2000$  (say)

Easy prime  $q \nmid \text{disc}(f)$ , tricky prime  $q \mid \text{disc}(f)$

## Implementation for TNFS

Reverse-engineering of `cado-nfs/polyselect/{auxiliary.c,alpha.sage}`

Magma and SageMath <https://gitlab.inria.fr/tnfs-alpha/alpha>

Same algorithm, prime  $q \rightarrow$  prime ideal of norm  $q$

## Example : Barreto-Naehrig curve, $p$ 254 bits

$$p = 36s^4 + 36s^3 + 24s^2 + 6s + 1 \text{ where } s = -(2^{62} + 2^{55} + 1)$$

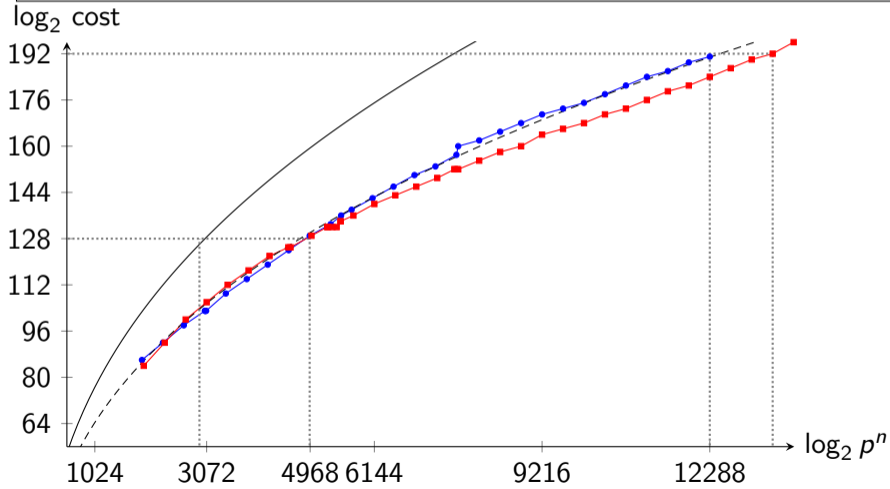
$$f = 36x^8 + 36yx^6 + 24y^2x^4 + 6y^3x^2 + y^4$$

$$g = x^2 + sy = x^2 + 4647714815446351873y$$

$$B = 2000$$

$h$	$1/\zeta_{K_h}(2)$	$\alpha(h, f, B)$	$\alpha(h, g, B)$	$\alpha_f + \alpha_g$
$y^6 + y^5 - y^2 - y - 1$	0.953	2.042	2.479	4.521
$y^6 - y^4 + y^3 + y^2 - 1$	0.917	1.288	1.740	3.028
$y^6 + y^3 + y^2 - y - 1$	0.917	2.419	2.876	5.295
$y^6 + y^5 - y^3 + y - 1$	0.909	0.278	2.357	2.636
$y^6 + y^5 + y^4 + y^3 + y^2 + y - 1$	0.883	2.341	2.033	4.374
$y^6 + y^4 + y^3 + y - 1$	0.867	0.899	2.526	3.425
$y^6 + y^4 + y^2 + y + 1$	0.836	1.955	1.141	3.095
$y^6 + y^5 + y^2 - y + 1$	0.763	0.891	1.264	2.155
$y^6 + y^5 - y^4 + y^3 + y^2 + y - 1$	0.756	0.956	1.177	2.133
$y^6 + y^5 + y - 1$	0.736	1.925	2.108	4.032
$y^6 + y^5 + y^3 - y^2 + y - 1$	0.732	1.729	2.099	3.828
$y^6 + y^3 + y - 1$	0.728	-0.250	1.191	0.941
$y^6 + y^3 - y + 1$	0.720	1.605	1.348	2.952
$y^6 + y^3 + y^2 + 1$	0.718	1.151	1.294	2.445
$y^6 - y^4 + y^3 - y^2 - y - 1$	0.710	0.406	2.278	2.684
$y^6 + y^5 - y^3 + y^2 - y + 1$	0.697	1.572	0.818	2.390
$y^6 + y^4 + y + 1$	0.679	1.319	1.683	3.002

- Simul. in  $\mathbb{F}_{p^{12}}$ , BN, STNFS deg  $h = 6, 4$
- Simul. in  $\mathbb{F}_{p^{12}}$ , BLS12, STNFS deg  $h = 12, 6$
- $L_{p^n}^0(1/3, 1.923)/2^{10.17}$  (DL theoretical re-scaled DL-240dd  $\leftrightarrow 2^{67.51}$ )
- - -  $L_{p^n}^0(1/3, 1.526)/2^{4.5}$  (SNFS theoretical re-scaled SDL-1024  $\leftrightarrow 2^{64.4}$ )



## Numerical example: BLS12-446 bits

$$p(x) = (x - 1)^2(x^4 - x^2 + 1)/3 + x$$

$$r(x) = x^4 - x^2 + 1$$

$$s = -(2^{74} + 2^{73} + 2^{63} + 2^{57} + 2^{50} + 2^{17} + 1)$$

seed with `enumerate_sparse_T.sage` [G. Masson Thomé]

<https://gitlab.inria.fr/smasson/cocks-pinch-variant>

$p = p(s)$  of 446 bits, twist-secure curve

$p^k$  5352 bits

$$h = Y^6 - Y^4 + Y^3 - Y + 1$$

$$f_y = X^{12} - 2yX^{10} + 2y^3X^6 + y^5X^2 + y^4 - y^3 + y - 1$$

$$g_y = X^2 - uy = X^2 + 28343567510342708887553y$$

$$A = 968, B = 2^{68.2}$$

Estimated cost:  $\approx 2^{132}$



## Key size for pairings


	DL cost $\sim 2^{128}$			DL cost $\sim 2^{192}$		
$\mathbb{F}_p$	3072–3200 bits			7400–8000 bits		
$\mathbb{F}_{p^n}$ , curve	$\log_2 r$	$\log_2 p$	$\log_2 p^n$	$\log_2 r$	$\log_2 p$	$\log_2 p^n$
$\mathbb{F}_{p^6}$ , MNT	660	660	3960	1647	1647	9882
$\mathbb{F}_{p^{12}}$ , BN	446	446	5352	1022	1022	12264
$\mathbb{F}_{p^{12}}$ , BLS	299	446	5352	768	1150	13800
$\mathbb{F}_{p^{16}}$ , KSS	256	330	5280	605	766	12256
$\mathbb{F}_{p^{18}}$ , KSS	256	348	6264	474	638	11484
$\mathbb{F}_{p^{24}}$ , BLS	256	318	7632	409	509	12216

- BN-254  $\approx 2^{103}$
- BN-382  $\approx 2^{123}$ , BLS12-381  $\approx 2^{126}$
- BN-446 and BLS12-446  $\approx 2^{132}$ , BN-462 and BLS12-461  $\approx 2^{135}$

Other curves:

- Fotiadis-Martindale [FM19]  $k = 12$  with  $r = r_{\text{BN}}$  like BLS12
- modified Cocks-Pinch with  $k = 8$  and  $\rho = 2.125$  [GMT19]

## References






-  Aurore Guillevic and Shashank Singh.  
On the alpha value of polynomials in the tower number field sieve algorithm.  
*Mathematical Cryptology*, 1(1), Feb. 2021.  
<https://journals.flvc.org/mathcryptology/article/view/125142>  
ePrint 2019/885

Source code: <https://gitlab.inria.fr/tnfs-alpha/alpha>

Webpage on pairing-friendly curves

<https://members.loria.fr/AGuillevic/pairing-friendly-curves/>

# Bibliography I

-  S. Bai, R. P. Brent, and E. Thomé.  
Root optimization of polynomials in the number field sieve.  
*Math. Comp.*, 84(295):2447–2457, 2015.
-  R. Barbulescu and S. Duquesne.  
Updating key size estimations for pairings.  
*Journal of Cryptology*, 32(4):1298–1336, Oct. 2019.
-  R. Barbulescu, P. Gaudry, and T. Kleinjung.  
The tower number field sieve.  
In T. Iwata and J. H. Cheon, editors, *ASIACRYPT 2015, Part II*, volume 9453 of *LNCS*, pages 31–55.  
Springer, Heidelberg, Nov. / Dec. 2015.
-  R. Barbulescu and A. Lachand.  
Some mathematical remarks on the polynomial selection in NFS.  
*Math. Comp.*, 86(303):397–418, 2017.
-  F. Boudot, P. Gaudry, A. Guillevic, N. Heninger, E. Thomé, and P. Zimmermann.  
Comparing the difficulty of factorization and discrete logarithm: A 240-digit experiment.  
In D. Micciancio and T. Ristenpart, editors, *CRYPTO 2020, Part II*, volume 12171 of *LNCS*, pages 62–91.  
Springer, Heidelberg, Aug. 2020.

# Bibliography II



A. Guillevic.

A short-list of pairing-friendly curves resistant to special TNFS at the 128-bit security level.

In A. Kiayias, M. Kohlweiss, P. Wallden, and V. Zikas, editors, *PKC 2020, Part II*, volume 12111 of *LNCS*, pages 535–564. Springer, Heidelberg, May 2020.



T. Kim and R. Barbulescu.

Extended tower number field sieve: A new complexity for the medium prime case.

In M. Robshaw and J. Katz, editors, *CRYPTO 2016, Part I*, volume 9814 of *LNCS*, pages 543–571. Springer, Heidelberg, Aug. 2016.



B. A. Murphy.

*Polynomial selection for the number field sieve integer factorisation algorithm.*

Phd thesis, Australian National University, Australia, 1999.

<http://maths-people.anu.edu.au/~brent/pd/Murphy-thesis.pdf>.