Factoring RSA-240 and computing discrete logarithms in a 240-digit prime field with the same software and hardware

Fabrice Boudot, Pierrick Gaudry, Aurore Guillevic, Nadia Heninger,
Emmanuel Thomé, Paul Zimmermann

December 10, 2020, Aarhus, Denmark

`https://members.loria.fr/AGuillevic/files/talks/20201210-Aarhus.pdf`

Outline

1. Introduction
2. Quadratic Sieve
3. Factorization with the Number Field Sieve
4. Our NFS record computation

Plan

1. Introduction
2. Quadratic Sieve
3. Factorization with the Number Field Sieve
4. Our NFS record computation

Introduction: public-key cryptography

1976 (Diffie–Hellman, DH) and 1977 (Rivert–Shamir–Adleman, RSA)
Asymmetric means distinct public and private keys

- encryption with a public key
- decryption with a private key
- deducing the private key from the public key is a very hard problem

Two hard problems:

- Integer factorization (for RSA)
- Discrete logarithm computation in a finite cyclic group (for Diffie–Hellman)

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice

Bob

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice **Bob**

0. chooses public parameters:
   modulus $N = p \cdot q$
   p, q distinct large safe primes
   encryption key $e = 3$ or $2^{16} + 1$
   private decryption key
   $d = e^{-1} \bmod \varphi(N) = (p-1)(q-1)$

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice **Bob**

0. chooses public parameters:
   modulus $N = p \cdot q$
   p, q distinct large safe primes
   encryption key $e = 3$ or $2^{16} + 1$ $\xrightarrow{N, e}$
   private decryption key
   $d = e^{-1} \bmod \varphi(N) = (p-1)(q-1)$

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice

0. chooses public parameters:
   modulus $N = p \cdot q$
   p, q distinct large safe primes
   encryption key $e = 3$ or $2^{16} + 1$
   private decryption key
   $d = e^{-1} \bmod \varphi(N) = (p-1)(q-1)$

$$\xrightarrow{N, e}$$

Bob

1. gets Alice's public key (N, e)
2. encodes m as integer in $[0, N-1]$
3. ciphertext $c = m^e \bmod N$
4. sends c to Alice

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice

0. chooses public parameters:
   modulus $N = p \cdot q$
   p, q distinct large safe primes
   encryption key $e = 3$ or $2^{16} + 1$
   private decryption key
   $d = e^{-1} \bmod \varphi(N) = (p-1)(q-1)$

$\xrightarrow{N, e}$

Bob

1. gets Alice's public key (N, e)
2. encodes m as integer in $[0, N-1]$
3. ciphertext $c = m^e \bmod N$
4. sends c to Alice

$\xleftarrow{c}$

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice

0. chooses public parameters:
   modulus $N = p \cdot q$
   p, q distinct large safe primes
   encryption key $e = 3$ or $2^{16} + 1$
   private decryption key
   $d = e^{-1} \bmod \varphi(N) = (p-1)(q-1)$

$\xrightarrow{N, e}$

Bob

1. gets Alice's public key (N, e)
2. encodes m as integer in $[0, N-1]$
3. ciphertext $c = m^e \bmod N$
4. sends c to Alice

$\xleftarrow{c}$

5. gets c from Bob
6. computes $m = c^d \bmod N$

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice

0. chooses public parameters:
   modulus $N = p \cdot q$
   p, q distinct large safe primes
   encryption key $e = 3$ or $2^{16} + 1$
   private decryption key
   $d = e^{-1} \bmod \varphi(N) = (p-1)(q-1)$

$\xrightarrow{N, e}$

Bob

1. gets Alice's public key (N, e)
2. encodes m as integer in $[0, N-1]$
3. ciphertext $c = m^e \bmod N$
4. sends c to Alice

$\xleftarrow{c}$

5. gets c from Bob
6. computes $m = c^d \bmod N$

It works: $m^{ed} \equiv m \bmod N$
because $ed = 1 \bmod (p-1)(q-1)$

Public-key encryption: 1977, Rivest, Shamir, Adleman (RSA)

Alice

0. chooses public parameters:
   modulus $N = p \cdot q$
   p, q distinct large safe primes
   encryption key $e = 3$ or $2^{16} + 1$
   private decryption key
   $d = e^{-1} \bmod \varphi(N) = (p-1)(q-1)$

$\xrightarrow{N, e}$

Bob

1. gets Alice's public key (N, e)
2. encodes m as integer in $[0, N-1]$
3. ciphertext $c = m^e \bmod N$
4. sends c to Alice

$\xleftarrow{c}$

5. gets c from Bob
6. computes $m = c^d \bmod N$

It works: $m^{ed} \equiv m \bmod N$
because $ed = 1 \bmod (p-1)(q-1)$

Hard tasks without knowing p, q if N is large enough:

- computing $(p-1)(q-1)$,
- computing a square root $\sqrt{x} = x^{1/2} \bmod N$,
- computing an e-th root $x^{1/e} \bmod N$.

RSA, security, attacks

The mathematical security relies on the hardness of computing d from N, e.
p, q are required to compute $\varphi(N)$
$\rightarrow$ security relies on the hardness of **integer factorization**.
Usecases:
`ssh-keygen` (linux), PGP: Enigmails on Thunderbird, Protonmail.
Note that short keys are not allowed:

```
ssh-keygen -b 512 -t rsa
Invalid RSA key length: minimum is 1024 bits
```

Survey by Dan Boneh in 1999 on many attacks because of wrong parameters or usage:

Dan Boneh.
Twenty years of attacks on the RSA cryptosystem.
Notices of the AMS, 46(2):203–213, February 1999.

Diffie–Hellman key exchange, discrete logarithm problem

Alice

Bob

Diffie–Hellman key exchange, discrete logarithm problem

Alice public parameters **Bob**

primes r, $\ell = (r-1)/2$, gen g, $\#\langle g \rangle = \ell$

Diffie–Hellman key exchange, discrete logarithm problem

Alice	public parameters	**Bob**
	primes r, $\ell = (r-1)/2$, gen g, $\#\langle g \rangle = \ell$	
$a \overset{\text{random}}{\longleftarrow} [1, \ell - 1]$	secret key	$b \overset{\text{random}}{\longleftarrow} [1, \ell - 1]$
$g^{a} \bmod r$	public value	$g^{b} \bmod r$

Diffie–Hellman key exchange, discrete logarithm problem

Alice public parameters **Bob**

primes r, $\ell = (r-1)/2$, gen g, $\#\langle g \rangle = \ell$

$a \overset{\text{random}}{\longleftarrow} [1, \ell - 1]$ secret key $b \overset{\text{random}}{\longleftarrow} [1, \ell - 1]$

g^a mod r public value g^b mod r

$$\xrightarrow{g^a}$$

$$\xleftarrow{g^b}$$

Diffie–Hellman key exchange, discrete logarithm problem

Alice public parameters **Bob**

primes r, $\ell = (r-1)/2$, gen g, $\#\langle g \rangle = \ell$

$a \xleftarrow{\text{random}} [1, \ell - 1]$ secret key $b \xleftarrow{\text{random}} [1, \ell - 1]$

$g^a \bmod r$ public value $g^b \bmod r$

$$\xrightarrow{g^a}$$

$$\xleftarrow{g^b}$$

gets Bob's public key g^b gets Alice's public key g^a

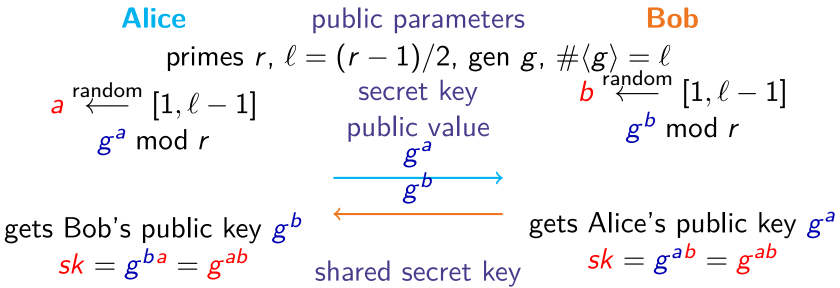
$sk = g^{b\,a} = g^{ab}$ shared secret key $sk = g^{a\,b} = g^{ab}$

Diffie–Hellman key exchange, discrete logarithm problem

Alice	public parameters	**Bob**
	primes r, $\ell = (r-1)/2$, gen g, $\#\langle g \rangle = \ell$	
$a \xleftarrow{\text{random}} [1, \ell - 1]$	secret key	$b \xleftarrow{\text{random}} [1, \ell - 1]$
g^a mod r	public value	g^b mod r
	$\xrightarrow{\quad g^a \quad}$	
	$\xleftarrow{\quad g^b \quad}$	
gets Bob's public key g^b		gets Alice's public key g^a
$sk = g^{b\,a} = g^{ab}$	shared secret key	$sk = g^{a\,b} = g^{ab}$

it works because $(g^a)^b = (g^b)^a = g^{ab}$

Diffie–Hellman key exchange, discrete logarithm problem



Alice public parameters **Bob**

primes r, $\ell = (r-1)/2$, gen g, $\#\langle g \rangle = \ell$

$a \xleftarrow{\text{random}} [1, \ell-1]$ secret key $b \xleftarrow{\text{random}} [1, \ell-1]$

$g^a \bmod r$ public value $g^b \bmod r$

$$\xrightarrow{g^a}$$

$$\xleftarrow{g^b}$$

gets Bob's public key g^b gets Alice's public key g^a

$sk = g^{b\,a} = g^{ab}$ shared secret key $sk = g^{a\,b} = g^{ab}$

it works because $(g^a)^b = (g^b)^a = g^{ab}$

Diffie-Hellman Problem

Given $\mathbb{G} = \langle g \rangle$, g, g^a, g^b, computes g^{ab}.

Discrete Logarithm Problem

Given $\mathbb{G} = \langle g \rangle$, g, g^a, computes a.

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 ↔ n bits of security

Choosing key sizes

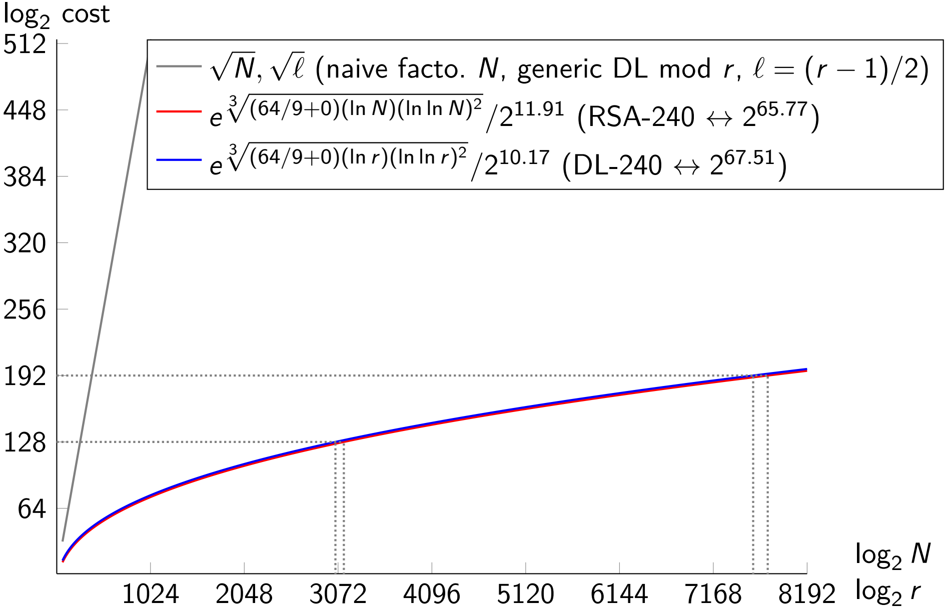
Symmetric ciphers (AES): key sizes are 128, 192 or 256 bits.
Perfect symmetric cipher: trying all keys of size n bits takes 2^n tests
$\rightarrow$ **brute-force search**

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

RSA with modulus N, DH with primes r and $\ell = (r-1)/2$:
n bits of security $\longleftrightarrow$ 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 length(N), resp. length(r) and length(ℓ)?

RSA and DH keys are much larger.
Cipher suite: a pair of symmetric and asymmetric ciphers offering the same level of security.


$\log_2$ cost
512
448
384
320
256
192
128
64
$\sqrt{N}, \sqrt{\ell}$ (naive facto. N, generic DL mod r, $\ell = (r-1)/2$)
$e^{\sqrt[3]{(64/9+0)(\ln N)(\ln\ln N)^2}}/2^{11.91}$ (RSA-240 $\leftrightarrow 2^{65.77}$)
$e^{\sqrt[3]{(64/9+0)(\ln r)(\ln\ln r)^2}}/2^{10.17}$ (DL-240 $\leftrightarrow 2^{67.51}$)
1024
2048
3072
4096
5120
6144
7168
8192
$\log_2 N$
$\log_2 r$

Factorization, Discrete Log Computation

Factoring RSA modulus N of 240 decimal digits (795 bits)
$N =$
124620366781718784065835044608106590434820374651678805754818
788883289666801188210855036039570272508747509864768438458621
054865537970253930571891217684318286362846948405301614416430
468066875699415246993185704183030512549594371372159029236099

Computing discrete logarithms in $\mathbb{Z}/r\mathbb{Z}$, $r = N + 49204$, $\ell = (r-1)/2$ prime

hardware:
Intel Xeon Gold 6130 processors, 2 CPUs, 16 physical cores/CPU, at 2.10 GHz

Factorization

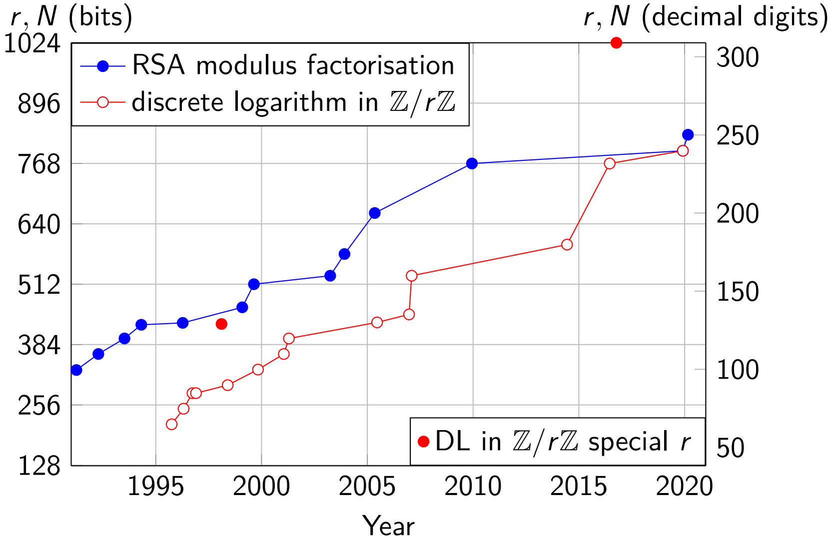
Integer factorization algorithms:

- trial division: try all prime numbers up to e.g. 10^7
- ECM (Elliptic Curve Method, Lenstra 87): find medium-size factors
- Quadratic sieve: N up to 100 decimal digits (dd)
- Number Field Sieve: N larger than 100 dd

Historical steps in integer factorization

- 1975, Morrison, Brillhard, continued fraction method CFRAC, factorization of $2^{2^7} + 1 = 2^{128} + 1$, see the *Cunningham project* `https://homes.cerias.purdue.edu/~ssw/cun/` $2^{128} + 1 = 340282366920938463463374607431768211457 = 59649589127497217 \times 5704689200685129054721$
- 1981, Dixon, random squares method
- 70's, unpublished: Schroeppel, Linear Sieve
- 1982, Pomerance, Quadratic Sieve
- 1987, Lenstra, Elliptic Curve Method (ECM)
- 1993, Buhler, Lenstra, Pomerance, General Number Field Sieve

Factorization and Discrete Log Records with NFS



Plan

1. Introduction
2. Quadratic Sieve
3. Factorization with the Number Field Sieve
4. Our NFS record computation

Square roots modulo N

In a field $\mathbb{R}$ or $\mathbb{C}$ or $\mathbb{F}_p$, if x is a square, it has two square roots $\sqrt{x}$ and $-\sqrt{x}$.
But in $\mathbb{Z}/N\mathbb{Z}$ with $N = pq$: **four** square roots.

Square roots modulo N

In a field $\mathbb{R}$ or $\mathbb{C}$ or $\mathbb{F}_p$, if x is a square, it has two square roots $\sqrt{x}$ and $-\sqrt{x}$.
But in $\mathbb{Z}/N\mathbb{Z}$ with $N = pq$: **four** square roots.

```
N = 2021
for i in range(-N//2, N//2):
    if (i**2 % N) == 1:
        print(i)
```

Two pairs of square roots of $x = 1$: $(1, -1)$ and $(-988, 988)$

Square roots modulo N

In a field $\mathbb{R}$ or $\mathbb{C}$ or $\mathbb{F}_p$, if x is a square, it has two square roots $\sqrt{x}$ and $-\sqrt{x}$.
But in $\mathbb{Z}/N\mathbb{Z}$ with $N = pq$: **four** square roots.

```
N = 2021
for i in range(-N//2, N//2):
    if (i**2 % N) == 1:
        print(i)
```

Two pairs of square roots of $x = 1$: $(1, -1)$ and $(-988, 988)$

$$\begin{aligned} 988^2 &= 1^2 \bmod 2021 \\ \iff 988^2 - 1^2 &= 0 \bmod 2021 \\ \iff (988-1) \times (988+1) &= 0 \bmod 2021 \end{aligned}$$

Square roots modulo N

In a field $\mathbb{R}$ or $\mathbb{C}$ or $\mathbb{F}_p$, if x is a square, it has two square roots $\sqrt{x}$ and $-\sqrt{x}$.
But in $\mathbb{Z}/N\mathbb{Z}$ with $N = pq$: **four** square roots.

```
N = 2021
for i in range(-N//2, N//2):
    if (i**2 % N) == 1:
        print(i)
```

Two pairs of square roots of $x = 1$: $(1, -1)$ and $(-988, 988)$

$$\begin{aligned} & 988^2 = 1^2 \bmod 2021 \\ \iff & 988^2 - 1^2 = 0 \bmod 2021 \\ \iff & (988 - 1) \times (988 + 1) = 0 \bmod 2021 \end{aligned}$$

Compute a gcd (greatest common divisor):
$\gcd(988 - 1, 2021) = 47$, $\gcd(988 + 1, 2021) = 43$.
$N = 43 \times 47$

Factorization with the Quadratic Sieve

Input: N to be factored

If $X^2 \equiv Y^2 \bmod N$ and $X \neq \pm Y \bmod N$, then $\gcd(X \pm Y, N)$ gives a factor of N.

Find such X, Y.

Factorization with the Quadratic Sieve

Input: N to be factored

If $X^2 \equiv Y^2 \bmod N$ and $X \neq \pm Y \bmod N$, then $\gcd(X \pm Y, N)$ gives a factor of N.

Find such X, Y.

Set $m = \lfloor \sqrt{N} \rfloor$, set bounds A, B

Factorization with the Quadratic Sieve

Input: N to be factored

If $X^2 \equiv Y^2 \bmod N$ and $X \neq \pm Y \bmod N$, then $\gcd(X \pm Y, N)$ gives a factor of N.

Find such X, Y.

Set $m = \lfloor \sqrt{N} \rfloor$, set bounds A, B

For many small $a \leq A$, computes $n_a = (a+m)^2 - N$

if n_a is B-smooth, store the relation $n_a = p_1^{e_1} p_2^{e_2} \cdots p_j^{e_j}$ with all primes $p_i \leq B$

Factorization with the Quadratic Sieve

Input: N to be factored

If $X^2 \equiv Y^2 \bmod N$ and $X \neq \pm Y \bmod N$, then $\gcd(X \pm Y, N)$ gives a factor of N.

Find such X, Y.

Set $m = \lfloor \sqrt{N} \rfloor$, set bounds A, B
For many small $a \leq A$, computes $n_a = (a+m)^2 - N$

if n_a is B-smooth, store the relation $n_a = p_1^{e_1} p_2^{e_2} \cdots p_j^{e_j}$ with all primes $p_i \leq B$

Find a combination s.t. $n_{a_1} n_{a_2} \cdots n_{a_i} = p_1^{e_1} p_2^{e_2} \cdots p_k^{e_k}$ and all e_i even

$X = (a_1 + m)(a_2 + m) \cdots (a_i + m) \bmod N$, $Y = \sqrt{n_{a_1} n_{a_2} \cdots n_{a_i}} \bmod N$
If $X \neq \pm Y \bmod N$, computes $\gcd(X - Y, N)$.

Factorization with the Quadratic Sieve: example

$N = 2021$, $m = \lfloor \sqrt{N} \rfloor = 44$

Smoothness bound $B = 19$

$\mathcal{F} = \{2, 3, 5, 7, 11, 13, 17, 19\}$ small primes up to B, $i = \#\mathcal{F} = 8$

B-smooth integer: $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$, all $p_i \leq B$ primes

Factorization with the Quadratic Sieve: example

$N = 2021$, $m = \lfloor \sqrt{N} \rfloor = 44$

Smoothness bound $B = 19$

$\mathcal{F} = \{2, 3, 5, 7, 11, 13, 17, 19\}$ small primes up to B, $i = \#\mathcal{F} = 8$

B-smooth integer: $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$, all $p_i \leq B$ primes

is $n = (a+m)^2 - N$ smooth for small a?

$$\begin{aligned}(2+m)^2 - N &= 95 = 5 \cdot 19\\ (5+m)^2 - N &= 380 = 2^2 \cdot 5 \cdot 19\\ (17+m)^2 - N &= 1700 = 2^2 \cdot 5^2 \cdot 17\end{aligned}$$

Factorization with the Quadratic Sieve: example

$N = 2021$, $m = \lfloor \sqrt{N} \rfloor = 44$

Smoothness bound $B = 19$

$\mathcal{F} = \{2, 3, 5, 7, 11, 13, 17, 19\}$ small primes up to B, $i = \#\mathcal{F} = 8$

B-smooth integer: $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$, all $p_i \leq B$ primes

is $n = (a+m)^2 - N$ smooth for small a?

$$\begin{array}{rcl} (2+m)^2 - N = & 95 = & 5 \cdot 19 \\ (5+m)^2 - N = & 380 = & 2^2 \cdot 5 \cdot 19 \\ (17+m)^2 - N = & 1700 = & 2^2 \cdot 5^2 \cdot 17 \end{array} \rightarrow \begin{array}{c} \begin{array}{cccc} 2 & 5 & 17 & 19 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 2 & 1 & 0 & 1 \\ 2 & 2 & 1 & 0 \end{bmatrix} \end{array} \text{ exponents}$$

Factorization with the Quadratic Sieve: example

$N = 2021$, $m = \lfloor \sqrt{N} \rfloor = 44$

Smoothness bound $B = 19$

$\mathcal{F} = \{2, 3, 5, 7, 11, 13, 17, 19\}$ small primes up to B, $i = \#\mathcal{F} = 8$

B-smooth integer: $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$, all $p_i \le B$ primes

is $n = (a+m)^2 - N$ smooth for small a?

$$\begin{aligned} (2+m)^2 - N &= 95 = 5 \cdot 19 \\ (5+m)^2 - N &= 380 = 2^2 \cdot 5 \cdot 19 \quad \to \\ (17+m)^2 - N &= 1700 = 2^2 \cdot 5^2 \cdot 17 \end{aligned} \qquad \begin{array}{c} \begin{array}{cccc} 2 & 5 & 17 & 19 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

exponents
mod 2

Factorization with the Quadratic Sieve: example

$N = 2021$, $m = \lfloor \sqrt{N} \rfloor = 44$

Smoothness bound $B = 19$

$\mathcal{F} = \{2, 3, 5, 7, 11, 13, 17, 19\}$ small primes up to B, $i = \#\mathcal{F} = 8$

B-smooth integer: $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$, all $p_i \leq B$ primes

is $n = (a+m)^2 - N$ smooth for small a?

$$\begin{array}{rcl} \to (2+m)^2 - N = & 95 = & 5 \cdot 19 \\ \to (5+m)^2 - N = & 380 = & 2^2 \cdot 5 \cdot 19 \quad \to \\ (17+m)^2 - N = & 1700 = & 2^2 \cdot 5^2 \cdot 17 \end{array} \qquad \begin{array}{cccc} 2 & 5 & 17 & 19 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array}$$

exponents
mod 2

Left kernel: $\begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$

$$(2+m)^2(5+m)^2 \equiv 2^2 \cdot 5^2 \cdot 19^2 \bmod N$$

Factorization with the Quadratic Sieve: example

$N = 2021$, $m = \lfloor\sqrt{N}\rfloor = 44$

Smoothness bound $B = 19$

$\mathcal{F} = \{2, 3, 5, 7, 11, 13, 17, 19\}$ small primes up to B, $i = \#\mathcal{F} = 8$

B-smooth integer: $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$, all $p_i \leq B$ primes

is $n = (a+m)^2 - N$ smooth for small a?

$$\begin{array}{rcl} (2+m)^2 - N = & 95 = 5 \cdot 19 & \\ (5+m)^2 - N = & 380 = 2^2 \cdot 5 \cdot 19 & \rightarrow \\ (17+m)^2 - N = & 1700 = 2^2 \cdot 5^2 \cdot 17 & \end{array} \quad \begin{array}{c} \begin{array}{cccc} 2 & 5 & 17 & 19 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

exponents mod 2

Left kernel: $\begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$

$$\begin{aligned} (2+m)^2(5+m)^2 &\equiv 2^2 \cdot 5^2 \cdot 19^2 \bmod N \\ (\underbrace{46 \cdot 49}_{X})^2 &\equiv (\underbrace{2 \cdot 5 \cdot 19}_{Y})^2 \bmod N \end{aligned}$$

Factorization with the Quadratic Sieve: example

$N = 2021$, $m = \lfloor\sqrt{N}\rfloor = 44$

Smoothness bound $B = 19$

$\mathcal{F} = \{2, 3, 5, 7, 11, 13, 17, 19\}$ small primes up to B, $i = \#\mathcal{F} = 8$

B-smooth integer: $n = p_1^{e_1} p_2^{e_2} \cdots p_i^{e_i}$, all $p_i \leq B$ primes

is $n = (a+m)^2 - N$ smooth for small a?

$$\begin{aligned} (2+m)^2 - N &= 95 = 5 \cdot 19 \\ (5+m)^2 - N &= 380 = 2^2 \cdot 5 \cdot 19 \\ (17+m)^2 - N &= 1700 = 2^2 \cdot 5^2 \cdot 17 \end{aligned} \rightarrow \begin{array}{c} \begin{array}{cccc} 2 & 5 & 17 & 19 \end{array} \\ \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix} \end{array}$$

exponents mod 2

Left kernel: $\begin{bmatrix} 1 & 1 & 0 \end{bmatrix}$

$$\begin{aligned} (2+m)^2(5+m)^2 &\equiv 2^2 \cdot 5^2 \cdot 19^2 \bmod N \\ (\underbrace{46 \cdot 49}_{X})^2 &\equiv (\underbrace{2 \cdot 5 \cdot 19}_{Y})^2 \bmod N \end{aligned}$$

$X = 2254 \equiv 233 \bmod N$, $Y = 190 \bmod N$

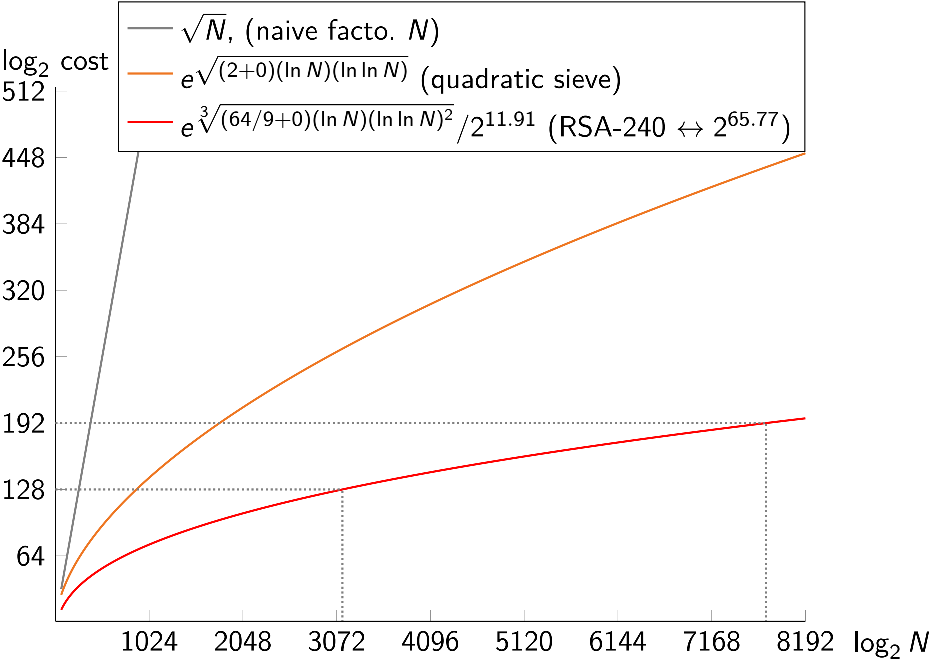
$\gcd(X - Y, N) = 43$, $\gcd(X + Y, N) = 47$

$N = 43 \cdot 47$

Quadratic Sieve: limitations for large numbers

Complexity: $e^{\sqrt{(2+o(1))\ln N \ln\ln N}}$

- $n = (a+m)^2 - N \approx 2A\sqrt{N}$
  Factor integers of size $\approx 2A\sqrt{N}$
- $\#\mathcal{F} = \#\{$ primes up to $B\} \approx B/\ln B$
- Computes left kernel of huge linear system modulo 2


$\sqrt{N}$, (naive facto. N)
$e^{\sqrt{(2+0)(\ln N)(\ln\ln N)}}$ (quadratic sieve)
$e^{\sqrt[3]{(64/9+0)(\ln N)(\ln\ln N)^2}}/2^{11.91}$ (RSA-240 $\leftrightarrow$ $2^{65.77}$)
$\log_2$ cost
512
448
384
320
256
192
128
64
1024
2048
3072
4096
5120
6144
7168
8192
$\log_2 N$

Plan

1. Introduction
2. Quadratic Sieve
3. Factorization with the Number Field Sieve
4. Our NFS record computation

Nowadays' method: the Number Field Sieve

- developed in the 80's and 90's
- reduce the size of the numbers to be factored from $A_0\sqrt{N}$ to $A^d\sqrt[d]{N}$ for a smaller $A < A_0$ and $d \in \{3, 4, 5, 6\}$
- two huge steps: collecting relations, solving a large sparse system

Nowadays' method: the Number Field Sieve

- developed in the 80's and 90's
- reduce the size of the numbers to be factored from $A_0\sqrt{N}$ to $A^d\sqrt[d]{N}$ for a smaller $A < A_0$ and $d \in \{3, 4, 5, 6\}$
- two huge steps: collecting relations, solving a large sparse system
- ElGamal 1985: discrete logarithms in $\mathbb{F}_{p^2}$ with a quadratic number field
- Coppersmith Odlysko Schroepel 1986 : discrete logarithms in $\mathbb{F}_p$ with a quadratic number field

Factorization with NFS: key idea

Reduce further the size of the integers to factor

Choose integer $m \approx \sqrt[d]{N}$

Write N in basis m: $N = c_0 + c_1 m + \ldots + c_d m^d$

Set $f_1(x) = c_0 + c_1 x + \ldots + c_d x^d \implies f_1(m) = 0$, set $f_0 = x - m \implies f_0(m) = 0$

Polynomials f_0, f_1 share a common root m modulo N

If f_1 is irreducible, define $\alpha \in \mathbb{C}$ a root of f_1

Factorization with NFS: key idea

Reduce further the size of the integers to factor

Choose integer $m \approx \sqrt[d]{N}$

Write N in basis m: $N = c_0 + c_1 m + \ldots + c_d m^d$

Set $f_1(x) = c_0 + c_1 x + \ldots + c_d x^d \implies f_1(m) = 0$, set $f_0 = x - m \implies f_0(m) = 0$

Polynomials f_0, f_1 share a common root m modulo N

If f_1 is irreducible, define $\alpha \in \mathbb{C}$ a root of f_1

Define a map from $\mathbb{Z}[\alpha]$ to $\mathbb{Z}/N\mathbb{Z}$

$$\begin{array}{rcl} \phi\colon \mathbb{Z}[\alpha] & \to & \mathbb{Z}/N\mathbb{Z} \\ \alpha & \mapsto & m \bmod N \text{ where } f_1(m) = 0 \bmod N \end{array}$$

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

$$\phi \quad \underbrace{(a + b\alpha)}_{\text{factor in } \mathbb{Z}[\alpha]} \quad = \quad \underbrace{a + bm}_{\text{factor in } \mathbb{Z}} \quad \bmod N$$

Factorization with NFS: key idea

Reduce further the size of the integers to factor
Choose integer $m \approx \sqrt[d]{N}$
Write N in basis m: $N = c_0 + c_1 m + \ldots + c_d m^d$
Set $f_1(x) = c_0 + c_1 x + \ldots + c_d x^d \implies f_1(m) = 0$, set $f_0 = x - m \implies f_0(m) = 0$
Polynomials f_0, f_1 share a common root m modulo N
If f_1 is irreducible, define $\alpha \in \mathbb{C}$ a root of f_1

Define a map from $\mathbb{Z}[\alpha]$ to $\mathbb{Z}/N\mathbb{Z}$

$$\begin{array}{rcl} \phi\colon \mathbb{Z}[\alpha] & \to & \mathbb{Z}/N\mathbb{Z} \\ \alpha & \mapsto & m \bmod N \text{ where } f_1(m) = 0 \bmod N \end{array}$$

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

$$\phi \quad \underbrace{(a + b\alpha)}_{\substack{\text{factor in } \mathbb{Z}[\alpha] \\ \text{size } A^d N^{1/d}}} \quad = \quad \underbrace{a + bm}_{\substack{\text{factor in } \mathbb{Z} \\ \text{size } AN^{1/d}}} \qquad \bmod N$$

Factorization in $\mathbb{Z}[\alpha]$

Factor $N = 2021$

$m = 38,\ 7 + 15m + m^2 = N,\ f_1(x) = x^2 + 15x + 7,\ f_0 = x - m$

Factorization in $\mathbb{Z}[\alpha]$

Factor $N = 2021$
$m = 38,\ 7 + 15m + m^2 = N,\ f_1(x) = x^2 + 15x + 7,\ f_0 = x - m$
$\alpha \in \mathbb{C}$ a root of f_1, factorization in $\mathbb{Z}[\alpha]$:
$7 = (7^+)(7^-), 19 = (19^+)(19^-), 23 = (23^+)(23^-)$

Factorization in $\mathbb{Z}[\alpha]$

Factor $N = 2021$

$m = 38$, $7 + 15m + m^2 = N$, $f_1(x) = x^2 + 15x + 7$, $f_0 = x - m$

$\alpha \in \mathbb{C}$ a root of f_1, factorization in $\mathbb{Z}[\alpha]$:

$7 = (7^+)(7^-)$,$19 = (19^+)(19^-)$,$23 = (23^+)(23^-)$

Factorization in $\mathbb{Z}[i]$, $i \in \mathbb{C}$, $i^2 = -1$:

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

Factorization in $\mathbb{Z}[\alpha]$

Factor $N = 2021$

$m = 38$, $7 + 15m + m^2 = N$, $f_1(x) = x^2 + 15x + 7$, $f_0 = x - m$

$\alpha \in \mathbb{C}$ a root of f_1, factorization in $\mathbb{Z}[\alpha]$:

$7 = (7^+)(7^-)$,$19 = (19^+)(19^-)$,$23 = (23^+)(23^-)$

Factorization in $\mathbb{Z}[i]$, $i \in \mathbb{C}$, $i^2 = -1$:

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

Fondamental Unit: $u = 2\alpha + 1$ and $\text{Norm}(u) = 1$

Norm

Factorization in $\mathbb{Z}[\alpha]$

Factor $N = 2021$
$m = 38,\ 7 + 15m + m^2 = N,\ f_1(x) = x^2 + 15x + 7,\ f_0 = x - m$
$\alpha \in \mathbb{C}$ a root of f_1, factorization in $\mathbb{Z}[\alpha]$:
$7 = (7^+)(7^-), 19 = (19^+)(19^-), 23 = (23^+)(23^-)$
Factorization in $\mathbb{Z}[i]$, $i \in \mathbb{C}$, $i^2 = -1$:
$(1 + i)(1 - i) = 2,\ (2 + i)(2 - i) = 5,\ (2 + 3i)(2 - 3i) = 13$
Fondamental Unit: $u = 2\alpha + 1$ and $\mathrm{Norm}(u) = 1$

Norm

The norm of $a - bi$ in $\mathbb{Z}[i]$ is $\mathrm{Norm}(a - bi) = a^2 + b^2$

Factorization in $\mathbb{Z}[\alpha]$

Factor $N = 2021$

$m = 38$, $7 + 15m + m^2 = N$, $f_1(x) = x^2 + 15x + 7$, $f_0 = x - m$

$\alpha \in \mathbb{C}$ a root of f_1, factorization in $\mathbb{Z}[\alpha]$:

$7 = (7^+)(7^-)$,$19 = (19^+)(19^-)$,$23 = (23^+)(23^-)$

Factorization in $\mathbb{Z}[i]$, $i \in \mathbb{C}$, $i^2 = -1$:

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

Fondamental Unit: $u = 2\alpha + 1$ and $\mathrm{Norm}(u) = 1$

Norm

The norm of $a - bi$ in $\mathbb{Z}[i]$ is $\mathrm{Norm}(a - bi) = a^2 + b^2$

The norm of $a - b\alpha$ in $\mathbb{Z}[\alpha]$ is

$$\mathrm{Norm}(a - b\alpha) = b^2 f(a/b) = a^2 + 15ab + 7b^2$$

Factorization in $\mathbb{Z}[\alpha]$

Factor $N = 2021$

$m = 38$, $7 + 15m + m^2 = N$, $f_1(x) = x^2 + 15x + 7$, $f_0 = x - m$

$\alpha \in \mathbb{C}$ a root of f_1, factorization in $\mathbb{Z}[\alpha]$:

$7 = (7^+)(7^-)$,$19 = (19^+)(19^-)$,$23 = (23^+)(23^-)$

Factorization in $\mathbb{Z}[i]$, $i \in \mathbb{C}$, $i^2 = -1$:

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

Fondamental Unit: $u = 2\alpha + 1$ and $\text{Norm}(u) = 1$

Norm

The norm of $a - bi$ in $\mathbb{Z}[i]$ is $\text{Norm}(a - bi) = a^2 + b^2$

The norm of $a - b\alpha$ in $\mathbb{Z}[\alpha]$ is

$$\text{Norm}(a - b\alpha) = b^2 f(a/b) = a^2 + 15ab + 7b^2$$

To factor $a - b\alpha \in \mathbb{Z}[\alpha]$,
compute $\text{Norm}(a - b\alpha) \in \mathbb{Z}$ and factor in $\mathbb{Z}$
→ To factor N, factor many smaller integers.

a,b	$a - bm =$ factor in $\mathbb{Z}$	$a^2 + 15ab + 7b^2$	factor in $\mathbb{Z}[\alpha]$
$-23,2$	$-99 = -3^2 \cdot 11$	$-133 = -7 \cdot 19$	$(7^+)(19^+)$
$-22,1$	$-60 = -2^2 \cdot 3 \cdot 5$	$161 = 7 \cdot 23$	$(7^+)(23^+)$
$-16,1$	$-54 = -2 \cdot 3^3$	$23 = 23$	(23^-)
$-14,1$	$-52 = -2^2 \cdot 13$	$-7 = -7$	(7^-)
$-13,1$	$-51 = -3 \cdot 17$	$-19 = -19$	(19^-)
$-9,2$	$-85 = -5 \cdot 17$	$-161 = -7 \cdot 23$	$(7^+)(23^-)$
$-8,5$	$-198 = -2 \cdot 3^2 \cdot 11$	$-361 = -19^2$	$(19^-)^2$
$-8,15$	$-578 = -2 \cdot 17^2$	$-161 = -7 \cdot 23$	$(7^+)(23^+)$
$-7,1$	$-45 = -3^2 \cdot 5$	$-49 = -7^2$	$(7^-)^2$
$-6,13$	$-500 = -2^2 \cdot 5^3$	$49 = 7^2$	$(7^+)^2$
$-2,1$	$-40 = -2^3 \cdot 5$	$-19 = -19$	(19^+)
$-1,1$	$-39 = -3 \cdot 13$	$-7 = -7$	(7^+)
$-1,2$	$-77 = -7 \cdot 11$	$-1 = -1$	
$5,4$	$-147 = -3 \cdot 7^2$	$437 = 19 \cdot 23$	$(19^-)(23^-)$
$6,1$	$-32 = -2^5$	$133 = 7 \cdot 19$	$(7^+)(19^-)$
$7,6$	$-221 = -13 \cdot 17$	$931 = 7^2 \cdot 19$	$(7^-)^2(19^+)$

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

Build the matrix of relations:

- one row per (a, b) pair s.t. both sides are smooth
- one column per prime $\{2, 3, 5, 7, 11, 13, 17\}$
- one column per prime ideal $(7^+), (7^-), (19^+), (19^-), (23^+), (23^-)$
- store the exponents mod 2

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

$$M = \begin{array}{ccccccccccccc}
2 & 3 & 5 & 7 & 11 & 13 & 17 & (7^+) & (7^-) & (19^+) & (19^-) & (23^+) & (23^-) \\
\hline
0 & 2 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\
2 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\
1 & 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
2 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\
1 & 2 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 2 & 1 & 0 & 0 & 0 & 1 & 0 \\
0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 \\
2 & 0 & 3 & 0 & 0 & 0 & 0 & 2 & 0 & 0 & 0 & 0 & 0 \\
3 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 2 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\
5 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 2 & 1 & 0 & 0 & 0
\end{array}$$

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

$$M = \begin{array}{c} \begin{array}{ccccccccccccc} 2 & 3 & 5 & 7 & 11 & 13 & 17 & (7^{+}) & (7^{-}) & (19^{+}) & (19^{-}) & (23^{+}) & (23^{-}) \end{array} \\ \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{array} \bmod 2$$

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

$$M = \left[\begin{array}{ccccccccccccc}
2 & 3 & 5 & 7 & 11 & 13 & 17 & (7^+) & (7^-) & (19^+) & (19^-) & (23^+) & (23^-) \\
 & & & & 1 & & & 1 & & 1 & & & \\
 & 1 & 1 & & & & & 1 & & & & 1 & \\
1 & 1 & & & & & & & & & & & 1 \\
 & & & & & 1 & & & 1 & & & & \\
 & 1 & & & & & 1 & & & & 1 & & \\
 & & 1 & & & & 1 & 1 & & & & & 1 \\
1 & & & & 1 & & & & & & & & \\
1 & & & & & & & 1 & & & & 1 & \\
 & & 1 & & & & & & & & & & \\
 & & 1 & & & & & & & & & & \\
1 & & 1 & & & & & & & 1 & & & \\
 & 1 & & & & 1 & & 1 & & & & & \\
 & & & 1 & 1 & & & & & & & & \\
 & 1 & & & & & & & & & 1 & & 1 \\
1 & & & & & & & 1 & & & 1 & & \\
 & & & & & 1 & 1 & & & 1 & & &
\end{array}\right]$$

sparse

Example: from left kernel in GF(2) to factorization

$$\ker M \qquad \begin{array}{c} (\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ ) \\ (\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 1\ ) \\ (\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ ) \\ (\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ ) \end{array}$$

Example: from left kernel in GF(2) to factorization

ker M

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 1 & 0 & 1 & 0 & 1 & 1 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 1 & 1 \end{pmatrix}$$
$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Relations #9 and #10:

$$(-7-m) = -45 = -3^2 \cdot 5 \qquad -7-\alpha = (7^-)^2$$
$$(-6-13m) = -500 = -2^2 \cdot 5^3 \qquad -6-13\alpha = (7^+)^2$$

Example: from left kernel in GF(2) to factorization

$$\ker M \quad \begin{array}{c} (\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ ) \\ (\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 1\ ) \\ (\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ ) \\ (\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ ) \end{array}$$

Relations #9 and #10: $\left| \begin{array}{l} (-7-m) \ = \ -45 = -3^2 \cdot 5 \\ (-6-13m) = -500 = -2^2 \cdot 5^3 \end{array} \right| \begin{array}{l} -7-\alpha \ = (7^-)^2 \\ -6-13\alpha = (7^+)^2 \end{array}$

$(-7-m)(-6-13m) = (2 \cdot 3 \cdot 5^2)^2$, but $(-7-\alpha)(-6-13\alpha) = -49-98\alpha$ not square

because of the units

Example: from left kernel in GF(2) to factorization

$$\ker M \quad \begin{array}{c} (\,1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\,) \\ (\,0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 1\,) \\ (\,0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\,) \\ (\,0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\,) \end{array}$$

Relations #9 and #10: $\left| \begin{array}{l} (-7-m) \quad = \; -45 = -3^2 \cdot 5 \\ (-6-13m) = -500 = -2^2 \cdot 5^3 \end{array} \right| \quad \begin{array}{l} -7-\alpha \quad = (7^-)^2 \\ -6-13\alpha = (7^+)^2 \end{array}$

$(-7-m)(-6-13m) = (2 \cdot 3 \cdot 5^2)^2$, but $(-7-\alpha)(-6-13\alpha) = -49-98\alpha$ not square

because of the units

Relations # $\{5, 10, 11, 12, 15, 16\}$:

$(-13-m)(-6-13m)(-2-m)(-1-m)(6-m)(7-6m) = 530400^2$

$(-13-\alpha)(-6-13\alpha)(-2-\alpha)(-1-\alpha)(6-\alpha)(7-6\alpha) = -3113264-6456485\alpha$ not square

Example: from left kernel in GF(2) to factorization

$$\text{ker } M \quad \begin{array}{c}
(\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 1\ 0\ 1\ 0\ 1\ 1\ 1\ ) \\
(\ 0\ 0\ 1\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 1\ ) \\
(\ 0\ 0\ 0\ 0\ 1\ 0\ 0\ 0\ 0\ 1\ 1\ 1\ 0\ 0\ 1\ 1\ ) \\
(\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 0\ 1\ 1\ 0\ 0\ 0\ 0\ 0\ 0\ )
\end{array}$$

Relations #9 and #10: $\left| \begin{array}{l} (-7-m) \quad = \ -45 = -3^2 \cdot 5 \\ (-6-13m) = -500 = -2^2 \cdot 5^3 \end{array} \right| \begin{array}{l} -7-\alpha \quad = (7^-)^2 \\ -6-13\alpha = (7^+)^2 \end{array}$

$(-7-m)(-6-13m) = (2 \cdot 3 \cdot 5^2)^2$, but $(-7-\alpha)(-6-13\alpha) = -49-98\alpha$ not square because of the units

Relations # $\{5, 10, 11, 12, 15, 16\}$:

$(-13-m)(-6-13m)(-2-m)(-1-m)(6-m)(7-6m) = 530400^2$

$(-13-\alpha)(-6-13\alpha)(-2-\alpha)(-1-\alpha)(6-\alpha)(7-6\alpha) = -3113264-6456485\alpha$ not square

But $(-49-98\alpha)(-3113264-6456485\alpha) = (-12103-25137\alpha)^2$

$X = 150 \cdot 530400 = 1314 \bmod N$ $\qquad Y = (-12103-25137m) = 750 \bmod N$

$\gcd(X-Y, N) = 47$, $\gcd(X+Y, N) = 43$ $\qquad N = 43 \cdot 47$

Factorization with NFS: recap

1. Polynomial selection: find two irreducible polynomials in $\mathbb{Z}[x]$ sharing a common root m modulo N
2. Relation collection: computes many smooth relations
3. Linear algebra: takes logarithms mod 2 of the relations: large sparse matrix over $\mathbb{F}_2$, computes left kernel
4. Characters: find a combination of the vectors of the kernel so that $X^2 \equiv Y^2 \bmod N$
5. Square root: computes X, Y
6. Factor N: computes $\gcd(X - Y, N)$

Plan

1. Introduction
2. Quadratic Sieve
3. Factorization with the Number Field Sieve
4. **Our NFS record computation**

Latest record computations

Fabrice Boudot, Pierrick Gaudry, Aurore Guillevic, Nadia Heninger, Emmanuel Thomé, and Paul Zimmermann.
Comparing the difficulty of factorization and discrete logarithm: A 240-digit experiment.
In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part II*, volume 12171 of *LNCS*, pages 62–91. Springer, Heidelberg, August 2020.

Factorization of RSA-240 (795 bits) in December 2019
and RSA-250 (829 bits) in February 2020
Video at Crypto: `https://youtube.com/watch?v=Qk207A4H7kU`

Latest record computations

$$\begin{aligned}
\text{RSA-240} &= 124620366781718784065835044608106590434820374651678805754818\\
&\quad 788883289666801188210855036039570272508747509864768438458621\\
&\quad 054865537970253930571891217684318286362846948405301614416430\\
&\quad 468066875699415246993185704183030512549594371372159029236099,\\
p &= 509435952285839914555051023580843714132648382024111473186660\\
&\quad 296521821206469746700620316443478873837606252372049619334517,\\
q &= 244624208838318150567813139024002896653802092578931401452041\\
&\quad 221336558477095178155258218897735030590669041302045908071447.
\end{aligned}$$

Latest record computations

$$\begin{aligned}
\text{RSA-250} &= 214032465024074496126442307283933356300861471514475501779775492\\
&\ 088141802344714013664334551909580467961099285187247091458768739\\
&\ 626192155736304745477052080511905649310668769159001975940569345\\
&\ 7452230589325976697471681738069364894699871578494975937497937,\\
p &= 641352894770715802787901901705773890848250147429434472081168596\\
&\ 32024532344630238623598752668347708737661925585694639798853367,\\
q &= 333720275949781565562260106053551142279407603447675546667845209\\
&\ 87023841729210037080257448673296881877565718986258036932062711
\end{aligned}$$

Breaking the previous record: Why?

- Record computations needed for key-size recommendations
- Open-source software Cado-NFS
- Motivation to improve all the steps
- Testing folklore ideas competitive only for huge sizes
- Exploits improvements of ECM (Bouvier–Imbert PKC'2020)
- Scaling the code for larger sizes improves the running-time on smaller sizes

The CADO-NFS software

Record computations with the CADO-NFS software.

- Important software development effort since 2007.
- 250k lines of C/C++ code, 60k for relation collection only.
- Significant improvements since 2016.
  - improved parallelism: strive to get rid of scheduling bubbles;
  - versatility: large freedom in parameter selection;
  - prediction of behaviour and yield: essential for tuning.
- Open source (LGPL), open development model (gitlab).
  Our results can be reproduced.

Factorization 240 dd

$N = \text{RSA-240}$

Polynomial selection

$$
\begin{aligned}
m &= 105487753732969860223795041295860517380/17780390513045005995253 \\
f_1 &= 10853204947200x^6 \\
&\quad -4763683724115259920x^5 \\
&\quad -6381744461279867941961670x^4 \\
&\quad +974448934853864807690675067037x^3 \\
&\quad +179200573533665721310210640738061170x^2 \\
&\quad +1595712553369335430496125795083146688523x \\
&\quad -221175588842299117590564542609977016567191860 \\
f_0 &= 17780390513045005995253x \\
&\quad -105487753732969860223795041295860517380
\end{aligned}
$$

$\mathrm{Res}(f_0, f_1) = 120N$

Integers $(a - bm)$ much smaller than $\mathrm{Norm}(a - b\alpha)$.

Relation collection with lattice sieving

Most time-consuming part.
How to enumerate (a, b), and detect smooth $a - b\alpha$, $a - bm$?

Special-$\mathfrak{q}$ (spq) Sieving

ideal $\mathfrak{q} = (q, \alpha - r)$ in $\mathbb{Z}[\alpha]$ s.t. $\mathrm{Norm}(\mathfrak{q}) = q$
Basis $(q, \alpha - r) \to$ reduced basis $(\boldsymbol{u}, \boldsymbol{v}) = (u_0 + u_1\alpha, v_0 + v_1\alpha)$
$(a - b\alpha) \to i\boldsymbol{u} + j\boldsymbol{v} = s - t\alpha$, and $q \mid \mathrm{Norm}(s - t\alpha)$

Allow Parallelization

Consider all primes $q \in$[0.8G, 7.4G] (G=10^9) s.t. $\exists\ \mathfrak{q}$

- for $q \in$[0.8G, 2.1G]: **Lattice Sieve** on both sides
- for $q \in$[2.1G, 7.4G]: **Lattice Sieve** for f_1 (large norms) and **Factorization Tree** for f_0 (much smaller norms)

$\#$ spq $\approx 3.0e8 \approx 2^{28}$
Sieve area per spq: $\mathcal{A} = [-2^{15}, 2^{15}] \times [0, 2^{16}]$, $\#\mathcal{A} = 2^{32}$

Relations look like

small primes, special-q, large primes

✓	$5^2 \cdot 11 \cdot 23 \cdot 287093 \cdot 870953 \cdot 20179693 \cdot 28306698811 \cdot 47988583469$	$2^3 \cdot 5 \cdot 7 \cdot 13 \cdot 31 \cdot 61 \cdot 14407 \cdot 26563253 \cdot 86800081 \cdot 269845309 \cdot 802234039 \cdot 1041872869 \cdot 5552238917 \cdot 12144939971 \cdot 15856830239$
✓	$3 \cdot 1609 \cdot 77699 \cdot 235586599 \cdot 347727169 \cdot 369575231 \cdot 9087872491$	$2^3 \cdot 3 \cdot 5 \cdot 13 \cdot 19 \cdot 23 \cdot 31 \cdot 59 \cdot 239 \cdot 3989 \cdot 7951 \cdot 2829403 \cdot 31455623 \cdot 225623753 \cdot 811073867 \cdot 1304127157 \cdot 78955382651 \cdot 129320018741$
✓	$5 \cdot 1381 \cdot 877027 \cdot 15060047 \cdot 19042511 \cdot 11542780393 \cdot 13192388543$	$2^4 \cdot 5 \cdot 13 \cdot 31 \cdot 59 \cdot 823 \cdot 2801 \cdot 26539 \cdot 2944817 \cdot 3066253 \cdot 87271397 \cdot 108272617 \cdot 386616343 \cdot 815320151 \cdot 1361785079 \cdot 12322934353$
✓	$2^3 \cdot 5^2 \cdot 173 \cdot 971 \cdot 613909489 \cdot 929507779 \cdot 1319454803 \cdot 2101983503$	$2^7 \cdot 3^2 \cdot 5 \cdot 29 \cdot 1021 \cdot 42589 \cdot 190507 \cdot 473287 \cdot 31555663 \cdot 654820381 \cdot 802234039 \cdot 19147596953 \cdot 23912934131 \cdot 52023180217$
✗	$2^2 \cdot 15193 \cdot 232891 \cdot 19514983 \cdot 139295419 \cdot 540260173 \cdot 606335449$	$2^2 \cdot 3^4 \cdot 13 \cdot 19 \cdot 74897 \cdot 1377667 \cdot 55828453 \cdot 282012013 \cdot 802234039 \cdot 3350122463 \cdot 35787642311 \cdot 37023373909 \cdot 128377293101$
✗	$2^2 \cdot 5^4 \cdot 439 \cdot 1483 \cdot 13121 \cdot 21383 \cdot 67751 \cdot 452059523 \cdot 33099515051$	$2^2 \cdot 3^3 \cdot 11 \cdot 13 \cdot 19 \cdot 5023 \cdot 3683209 \cdot 98660459 \cdot 802234039 \cdot 1506372871 \cdot 4564625921 \cdot 27735876911 \cdot 32612130959 \cdot 45729461779$

small primes: abundant → dense column in the matrix

large primes: rare → sparse colum, limit to 2 or 3 on each side.

Relations look like

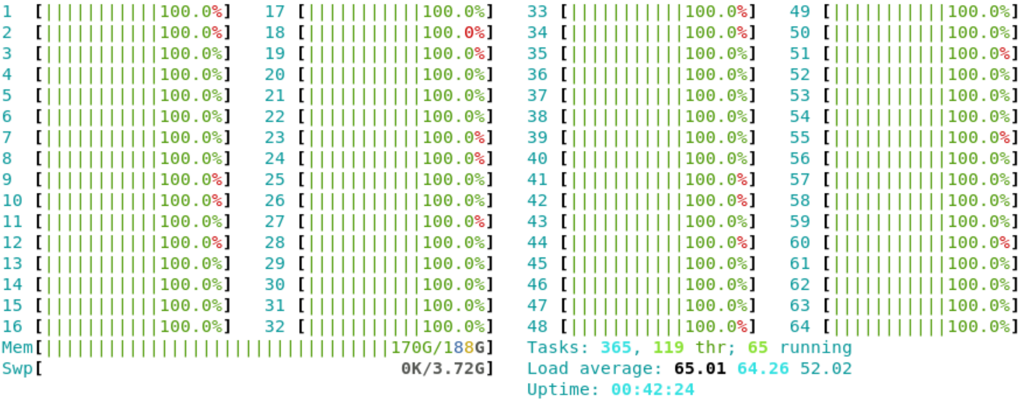
small primes, special-q, large primes

✓	$5^2 \cdot 11 \cdot 23 \cdot 287093 \cdot 870953 \cdot 20179693 \cdot 28306698811 \cdot 47988583469$	$2^3 \cdot 5 \cdot 7 \cdot 13 \cdot 31 \cdot 61 \cdot 14407 \cdot 26563253 \cdot 86800081 \cdot 269845309 \cdot 802234039 \cdot 1041872869 \cdot 5552238917 \cdot 12144939971 \cdot 15856830239$
✓	$3 \cdot 1609 \cdot 77699 \cdot 235586599 \cdot 347727169 \cdot 369575231 \cdot 9087872491$	$2^3 \cdot 3 \cdot 5 \cdot 13 \cdot 19 \cdot 23 \cdot 31 \cdot 59 \cdot 239 \cdot 3989 \cdot 7951 \cdot 2829403 \cdot 31455623 \cdot 225623753 \cdot 811073867 \cdot 1304127157 \cdot 78955382651 \cdot 129320018741$
✓	$5 \cdot 1381 \cdot 877027 \cdot 15060047 \cdot 19042511 \cdot 11542780393 \cdot 13192388543$	$2^4 \cdot 5 \cdot 13 \cdot 31 \cdot 59 \cdot 823 \cdot 2801 \cdot 26539 \cdot 2944817 \cdot 3066253 \cdot 87271397 \cdot 108272617 \cdot 386616343 \cdot 815320151 \cdot 1361785079 \cdot 12322934353$
✓	$2^3 \cdot 5^2 \cdot 173 \cdot 971 \cdot 613909489 \cdot 929507779 \cdot 1319454803 \cdot 2101983503$	$2^7 \cdot 3^2 \cdot 5 \cdot 29 \cdot 1021 \cdot 42589 \cdot 190507 \cdot 473287 \cdot 31555663 \cdot 654820381 \cdot 802234039 \cdot 19147596953 \cdot 23912934131 \cdot 52023180217$

small primes: abundant → dense column in the matrix
large primes: rare → sparse colum, limit to 2 or 3 on each side.

Before linear algebra: filtering step
as many cheap combinations as possible → smaller matrix

Relation collection looks like



```
1  [|||||||||||100.0%]   17 [|||||||||||100.0%]   33 [|||||||||||100.0%]   49 [|||||||||||100.0%]
2  [|||||||||||100.0%]   18 [|||||||||||100.0%]   34 [|||||||||||100.0%]   50 [|||||||||||100.0%]
3  [|||||||||||100.0%]   19 [|||||||||||100.0%]   35 [|||||||||||100.0%]   51 [|||||||||||100.0%]
4  [|||||||||||100.0%]   20 [|||||||||||100.0%]   36 [|||||||||||100.0%]   52 [|||||||||||100.0%]
5  [|||||||||||100.0%]   21 [|||||||||||100.0%]   37 [|||||||||||100.0%]   53 [|||||||||||100.0%]
6  [|||||||||||100.0%]   22 [|||||||||||100.0%]   38 [|||||||||||100.0%]   54 [|||||||||||100.0%]
7  [|||||||||||100.0%]   23 [|||||||||||100.0%]   39 [|||||||||||100.0%]   55 [|||||||||||100.0%]
8  [|||||||||||100.0%]   24 [|||||||||||100.0%]   40 [|||||||||||100.0%]   56 [|||||||||||100.0%]
9  [|||||||||||100.0%]   25 [|||||||||||100.0%]   41 [|||||||||||100.0%]   57 [|||||||||||100.0%]
10 [|||||||||||100.0%]   26 [|||||||||||100.0%]   42 [|||||||||||100.0%]   58 [|||||||||||100.0%]
11 [|||||||||||100.0%]   27 [|||||||||||100.0%]   43 [|||||||||||100.0%]   59 [|||||||||||100.0%]
12 [|||||||||||100.0%]   28 [|||||||||||100.0%]   44 [|||||||||||100.0%]   60 [|||||||||||100.0%]
13 [|||||||||||100.0%]   29 [|||||||||||100.0%]   45 [|||||||||||100.0%]   61 [|||||||||||100.0%]
14 [|||||||||||100.0%]   30 [|||||||||||100.0%]   46 [|||||||||||100.0%]   62 [|||||||||||100.0%]
15 [|||||||||||100.0%]   31 [|||||||||||100.0%]   47 [|||||||||||100.0%]   63 [|||||||||||100.0%]
16 [|||||||||||100.0%]   32 [|||||||||||100.0%]   48 [|||||||||||100.0%]   64 [|||||||||||100.0%]
Mem[|||||||||||||||||||||||||||||||||170G/188G]   Tasks: 365, 119 thr; 65 running
Swp[                                  0K/3.72G]   Load average: 65.01 64.26 52.02
                                                  Uptime: 00:42:24
```

Discrete Logarithm 240 dd

$r = N + 49204$, $\ell = (r-1)/2$ prime

$f_1 = 39x^4 + 126x^3 + x^2 + 62x + 120$

$$\begin{aligned}f_0 = {} & 286512172700675411986966846394359924874576536408786368056\,x^3 \\ & +249088203007157661364751159824397355165818886038172555398 90\,x^2 \\ & -187636975600130165644039539283271210355804094599448546527 37\,x \\ & -236610408827000256250190838220824122997878994595785432202599\end{aligned}$$

Res $f_0, f_1 = -540r$

More balanced integers

Smaller matrix but kernel modulo large prime ℓ

Relations, matrix size, core-years timings

	RSA-240	DLP-240
polynomial selection deg f_0, deg f_1	76 core-years 1, 6	152 core-years 3, 4
relation collection raw relations unique relations	794 core-years 8 936 812 502 6 011 911 051	2400 core-years 3 824 340 698 2 380 725 637
filtering after singleton removal after clique removal after merge	days 2 603 459 110 × 2 383 461 671 1 175 353 278 × 1 175 353 118 282M rows, density 200	days 1 304 822 186 × 1 000 258 769 149 898 095 × 149 898 092 36M rows, density 253
linear algebra	83 core-years	625 core-years
characters, sqrt, ind log	days	days

Conclusion

- Parameterization strategies
- Extensive simulation framework for parameter choices
- Implementation scales well

Conclusion

- Parameterization strategies
- Extensive simulation framework for parameter choices
- Implementation scales well

Comparisons:

- Comparison with previous record (DLP-768, 232 digits, 2016):
  On identical hardware, our DLP-240 computation would have taken less time than the 232-digits computation.
- Finite field DLP is not much harder than integer factoring.

Thank you

Fabrice Boudot, Pierrick Gaudry, Aurore Guillevic, Nadia Heninger, Emmanuel Thomé, and Paul Zimmermann.
Comparing the difficulty of factorization and discrete logarithm: A 240-digit experiment.
In Daniele Micciancio and Thomas Ristenpart, editors, *CRYPTO 2020, Part II*, volume 12171 of *LNCS*, pages 62–91. Springer, Heidelberg, August 2020.

Factorization of RSA-240 (795 bits) in December 2019
and RSA-250 (829 bits) in February 2020
Video at Crypto: `https://youtube.com/watch?v=Qk207A4H7kU`

RSA and the quantum computer

1994: Peter Shor, algorithm for integer factorization with a quantum computer

Factorization of a n-bit integer requires a perfect quantum computer with $2n$ qbits (quantum bits)

Quantum computer extremely hard to build

Record computation in 2018: $4\ 088\ 459 = 2017 \times 2027$

RSA-1024 (bits) will be factored before a quantum compter become competitive.