Continued Fractions and Quadratic Forms: A Novel Factorization Method

Giulia Salvatori
Politecnico di Torino

joint work with Nadir Murru (Università di Trento)

One World Numeration Seminar
February 4, 2025

Goal and bibliography

Goal

Develop a novel factorization method based on two works by Michele Elia.

Bibliography

Michele Elia (2019)
"Continued Fractions and Factoring"

Michele Elia (2021)
"Continued Fractions, Quadratic Fields, and Factoring: Some Computational Aspects"

Nadir Murru and Giulia Salvatori (2025)
"Integer Factorization via Continued Fractions and Quadratic Forms"

Table of Contents

Table of Contents

Factorization

Integer factorization is in $\mathcal{NP} \cap \text{co-}\mathcal{NP}$.

The most efficient known factorization algorithm is the **General Number Field Sieve**, with a heuristic running time of

$$\exp\left(\left(\sqrt[3]{\frac{64}{9}} + o(1)\right)(\ln N)^{1/3}(\ln \ln N)^{2/3}\right).$$

Used in public-key cryptography:

- RSA (1976)
- Rabin cryptosystem (1979)
- Goldwasser–Micali cryptosystem (1982)
- Paillier cryptosystem (1999)

Continued fractions

Definition

A simple continued fraction is an expression of the form

$$a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \cdots}}} = [a_0, a_1, a_2, a_3, \ldots],$$

where $a_0 \in \mathbb{Z}$ and $a_i \in \mathbb{N}^{>0}$, for $i > 0$. The terms $a_0, a_1, a_2, \ldots$ are called partial quotients of the continued fraction.

Continued fraction of a real number

Given $x_0 \in \mathbb{R}$, we have $x_0 = [a_0, a_1, a_2, \ldots]$, where the partial quotients are computed as follows

$$\begin{cases} a_i = \lfloor x_i \rfloor \\ x_{i+1} = \frac{1}{x_i - a_i} \end{cases}, \quad \text{if } a_i \neq x_i, \quad \forall i \geq 0.$$

Definition

A periodic continued fraction is an infinite continued fraction such that

$$a_i = a_{i+\tau}$$

for a fixed positive integer τ and all $i \geq \ell$, for some $\ell \geq 0$. The minimum τ that satisfies the equality is called period.

The expansion of $\sqrt{N}$

Theorem

Given $N \in \mathbb{N}^{>0}$, the continued fraction expansion of $\sqrt{N}$ is of the form

$$\sqrt{N} = [a_0, \overline{a_1, a_2, \ldots, a_2, a_1, 2a_0}].$$

We denote by π the length of period of the expansion of $\sqrt{N}$.

- (Stanton, Sudler, Williams)
  We have $\pi < 0.72\sqrt{N} \ln N$ for $N > 7$.
- (Littlewood)
  Under the GRH, if N is square-free we have $\pi \ll \sqrt{N} \ln \ln N$.
- (Vijayaraghavan)
  For any $\epsilon > 0$, infinitely many N satisfy $\pi > N^{1/2-\epsilon}$.

The sequences $\{Q_k\}_{k\geq 0}$ and $\{P_k\}_{k\geq 0}$

Definition (Lagrange's algorithm)

Let $N \in \mathbb{N}^{>0}$ non-square and $x_0 = \sqrt{N}$. We define P_k and Q_k, for $k \geq 0$, as

$$\begin{cases} a_k = \lfloor x_k \rfloor \\ P_{k+1} = a_k Q_k - P_k \\ Q_{k+1} = (N - P_{k+1}^2)/Q_k \end{cases}, \quad \text{where} \quad x_k = \frac{P_k + \sqrt{N}}{Q_k}, \quad k \geq 0.$$

- $0 < Q_k < \frac{2}{a_{k+1}}\sqrt{N}$ and $0 \leq P_k < \sqrt{N}$ for all $k \geq 0$.
- $\{P_k\}_{k\geq 1}$ and $\{Q_k\}_{k\geq 0}$ are periodic of period π.

How to find a factor

We will assume $N \in \mathbb{N}^{>0}$ to be composite and non-square.

Theorem (Elia)

If π is even, then $Q_{\pi/2}$ is a non-trivial factor of $2N$.

Theorem

If π is even, then $Q_{\pi/2} \neq 2$ if and only if the integer equations

$$X^2 - NY^2 = 2 \quad \textit{and} \quad X^2 - NY^2 = -2$$

are both unsolvable.

Proposition

If π is odd and $\left(\frac{-1}{N}\right) = -1$, then $Q_{(\pi+1)/2}$ contains a nontrivial factor of N.

Even period and non-trivial factorization

Let N be an RSA modulus: $N = pq$, where $p \neq q$ are primes.

$p \pmod 8$	$q \pmod 8$	$\pi \pmod 2$	$Q_{\pi/2}$
3	3	0	$\neq 2$
3	7		
7	7		
5	7		
5	3		
1	7	0	$= 2$ or $\neq 2$
1	3		
1	1	0 or 1	If π even, then $\neq 2$
1	5		
5	5		

Starting idea

We want to reach $Q_{\pi/2}$ or $Q_{(\pi+1)/2}$ as fast as possible.

We can compute $Q_0, Q_1, Q_2, Q_3, \ldots$ until we find a factor $\ldots$

$\ldots$ but π can be too large!

We need a way to move through $\{Q_k\}_{k\geq 0}$ making long jumps.

Table of Contents

Quadratic forms

Definition

Let $\Upsilon = \{\mathbf{F}_k\}_{k\geq 0}$, where

$$\mathbf{F}_k(x,y) = (-1)^k Q_k x^2 + 2P_{k+1}xy + (-1)^{k+1} Q_{k+1} y^2.$$

Definition

A quadratic form $F(x,y) = ax^2 + bxy + cy^2$ is reduced if

$$\left|\sqrt{\Delta} - 2\,|a|\right| < b < \sqrt{\Delta},$$

where $\Delta = b^2 - 4ac$ is the discriminant of F.

- The discriminant of $\mathbf{F}_k$ is $\Delta = 4N$ for all $k \geq 0$.
- $\mathbf{F}_k$ reduced for all $k \geq 0$.

Quadratic forms: periodicity

Definition

Let $\boldsymbol{\Upsilon} = \{\mathbf{F}_k\}_{k\geq 0}$, where

$$\mathbf{F}_k(x,y) = (-1)^k Q_k x^2 + 2P_{k+1}xy + (-1)^{k+1} Q_{k+1} y^2.$$

$$\mathbf{F}_k = \begin{cases} \mathbf{F}_{k+\pi} & \text{if } \pi \text{ even} \\ \mathbf{F}_{k+2\pi} & \text{if } \pi \text{ odd} \end{cases} \quad \text{for all } k \geq 0.$$

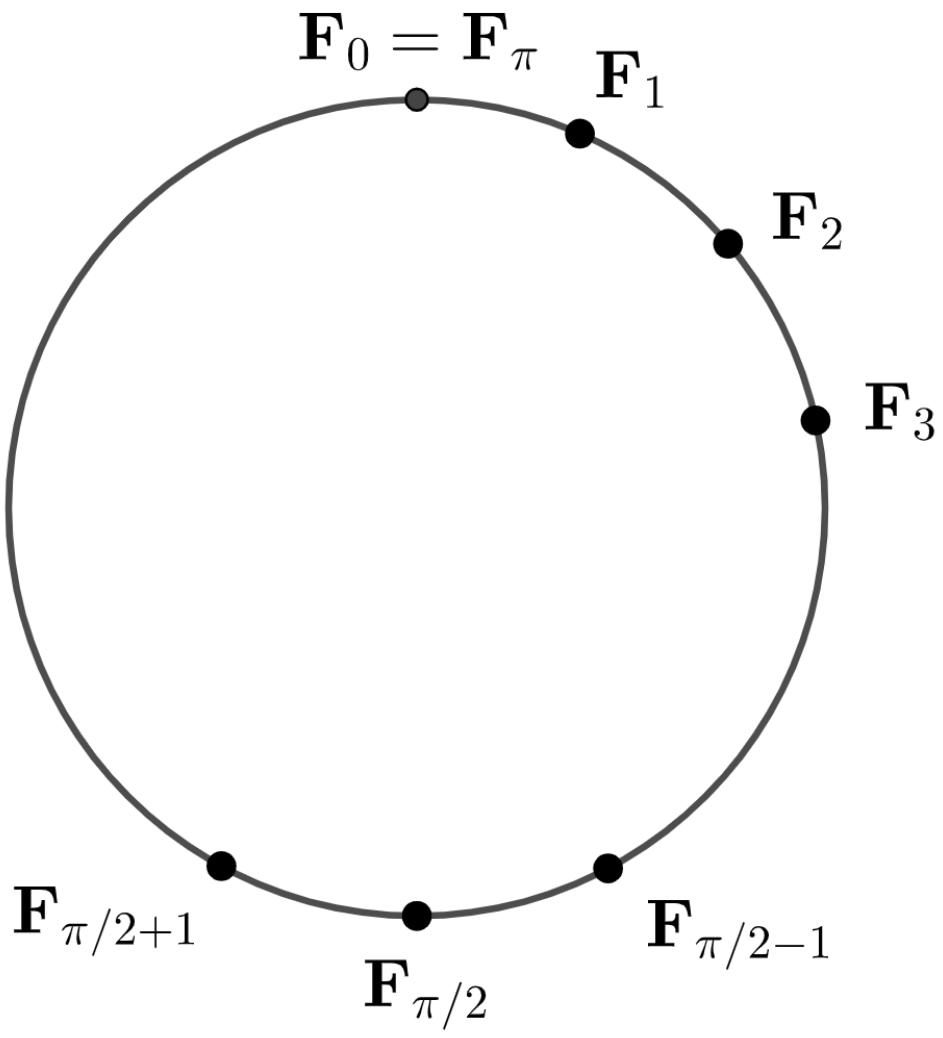
We want to reach $\mathbf{F}_{\pi/2}$ or $\mathbf{F}_{(\pi-1)/2}$ in a fast way.

$\mathbf{F}_{\pi/2} = ((-1)^{\pi/2}\mathbf{Q}_{\pi/\mathbf{2}}, 2P_{\pi/2+1}, (-1)^{\pi/2+1} Q_{\pi/2+1})$

$\mathbf{F}_{(\pi-1)/2} = ((-1)^{(\pi-1)/2} Q_{(\pi-1)/2}, 2P_{(\pi+1)/2}, (-1)^{(\pi+1)/2}\mathbf{Q}_{(\pi+\mathbf{1})/\mathbf{2}})$

We will refer to $\boldsymbol{\Upsilon}$ as the cycle.

The cycle $\boldsymbol{\Upsilon}$ if π even



Reduction operator ρ

Definition

Let $F = (a, b, c)$ be a form with $ac \neq 0$. If its discriminant $\Delta > 0$ is a non-square integer, we define the reduction operator ρ as

$$\rho(a, b, c) = \left(c, r(-b, c), \frac{r(-b, c)^2 - \Delta}{4c} \right),$$

where $r(-b, c)$ is the unique r such that $r + b \equiv 0 \pmod{2c}$ and

$$-|c| < r \leq |c| \quad \text{if} \quad \sqrt{\Delta} < |c|,$$

$$\sqrt{\Delta} - 2|c| < r < \sqrt{\Delta} \quad \text{if} \quad |c| < \sqrt{\Delta}.$$

- There exists $n \geq 0$ such that $\rho^n(F)$ is reduced.
- $\rho(\mathbf{F}_k) = \mathbf{F}_{k+1}$ for all $k \geq 0$.

Inverse reduction operator ρ^{-1}

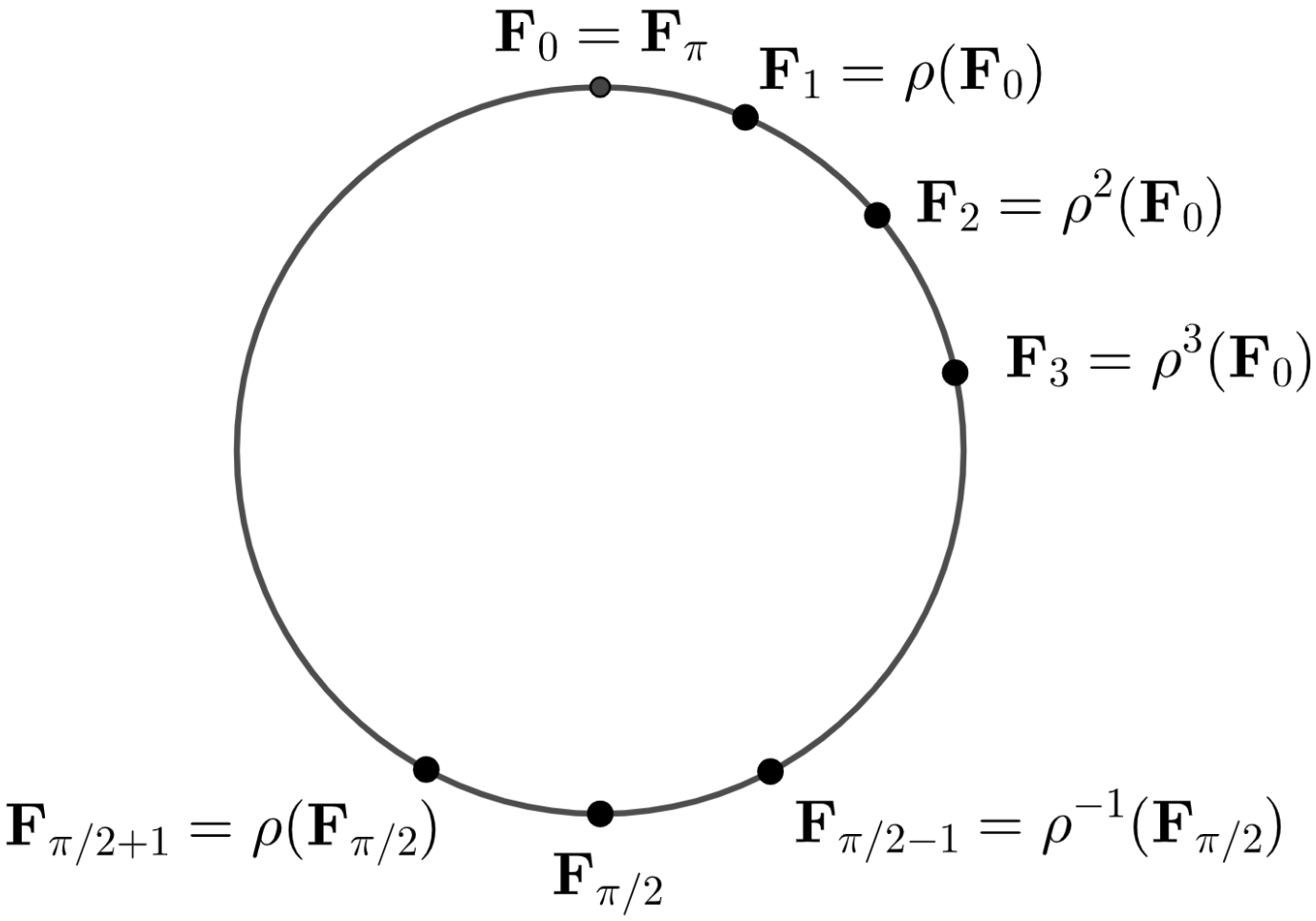
Definition

Let $F = (a, b, c)$ be a form with $ac \neq 0$. If its discriminant $\Delta > 0$ is a non-square integer, we define the ρ^{-1} as

$$\rho^{-1}(a, b, c) = \left(\frac{r(-b, a)^2 - \Delta}{4a}, r(-b, a), a \right).$$

- If F reduced, then $\rho(\rho^{-1}(F)) = \rho^{-1}(\rho(F)) = F$.
- $\rho^{-1}(\mathbf{F}_k) = \mathbf{F}_{k-1}$ for all $k > 0$.

The cycle $\boldsymbol{\Upsilon}$ if π even



Gauss composition

Definition

The Gauss composition $F \circ G$ of two quadratic forms $F = (a_1, b_1, c_1)$ and $G = (a_2, b_2, c_2)$, both having discriminant Δ, is

$$(a_3, b_3, c_3) = \left(d_0 \frac{a_1 a_2}{n^2}, b_1 + \frac{2a_1}{n} \left(\frac{t(b_2 - b_1)}{2} - c_1 v \right), \frac{b_3^2 - \Delta}{4a_3} \right),$$

where $\beta = (b_1 + b_2)/2$, $n = \gcd(a_1, a_2, \beta)$, $a_1 t + a_2 u + \beta v = n$, and $d_0 = \gcd(a_1, a_2, \beta, c_1, c_2, (b_1 - b_2)/2)$.

$$F = (a, b, c) \longleftrightarrow I_F = a\mathbb{Z} + \frac{-b+\sqrt{\Delta}}{2}\mathbb{Z} \text{ ideal of } \mathcal{O}_{\mathbb{Q}(\sqrt{\Delta})}$$

$$F \circ G \longleftrightarrow I_F I_G$$

Giant step

Definition

The giant step of $\mathbf{F}_n$ and $\mathbf{F}_m$ is the composition

$$\mathbf{F}_n \bullet \mathbf{F}_m = \rho^t(\mathbf{F}_n \circ \mathbf{F}_m),$$

realized through the Gauss composition $\mathbf{F}_n \circ \mathbf{F}_m$, followed by the minimum number t of reduction operations ρ to obtain a reduced form.

- $\mathbf{F}_n \circ \mathbf{F}_m$ may not be reduced.
- The number of applications of ρ needed to reduce $\mathbf{F}_n \circ \mathbf{F}_m$ is $O(\mathrm{ln}(N))$.
- For all $n, m \geq 0$ we have $\mathbf{F}_n \bullet \mathbf{F}_m \in \Upsilon$.

Distance of forms

Definition

Given a quadratic form $F = (a, b, c)$ with discriminant Δ, and $n > 0$, we define the distance δ as follows:

$$\delta(F, \rho(F)) = \frac{1}{2} \ln \left| \frac{b + \sqrt{\Delta}}{b - \sqrt{\Delta}} \right|,$$

and

$$\delta(F, \rho^n(F)) = \sum_{i=1}^{n} \delta(\rho^{i-1}(F), \rho^i(F)).$$

Important

δ is not a metric distance!

δ-length of a cycle

Definition

We define

$$R^+(N) = \ln(\bar{x} + \bar{y}\sqrt{N}),$$

where $\bar{x} + \bar{y}\sqrt{N}$ is the minimal solution of the Pell's equation $X^2 - NY^2 = 1$.

Theorem

If π even, then $\delta(\mathbf{F}_0, \mathbf{F}_\pi) = R^+(N)$ and $\delta(\mathbf{F}_0, \mathbf{F}_{\pi/2}) = R^+(N)/2$.

Theorem

If π odd, then $\delta(\mathbf{F}_0, \mathbf{F}_{2\pi}) = R^+(N)$, $\delta(\mathbf{F}_0, \mathbf{F}_\pi) = R^+(N)/2$, and $\delta(\mathbf{F}_0, \mathbf{F}_{(\pi-1)/2}) = R^+(N)/4 + O(\ln N)$.

Properties of δ

Theorem (H. Lenstra)

Let $\mathbf{F}_n, \mathbf{F}_m \in \Upsilon$, *and* $\mathbf{F}_r = \mathbf{F}_n \bullet \mathbf{F}_m$. *Then,*

$$\delta(\mathbf{F}_0, \mathbf{F}_r) = \delta(\mathbf{F}_0, \mathbf{F}_n) + \delta(\mathbf{F}_0, \mathbf{F}_m) + \delta(\mathbf{F}_n \circ \mathbf{F}_m, \mathbf{F}_r),$$

and $|\delta(\mathbf{F}_n \circ \mathbf{F}_m, \mathbf{F}_r)| < 2\ln(4N)$.

- $\delta(\mathbf{F}_k, \mathbf{F}_{k+1}) < \frac{1}{2}\ln N$
- $\delta(\mathbf{F}_k, \mathbf{F}_{k+2}) > \ln 2$

Corollary

If $\delta(\mathbf{F}_i, \mathbf{F}_j) = D$, then $\frac{2D}{\ln(4N)} < |j - i| < \frac{2D}{\ln 2} + 1$.

The cycle $\boldsymbol{\Upsilon}$ if π even

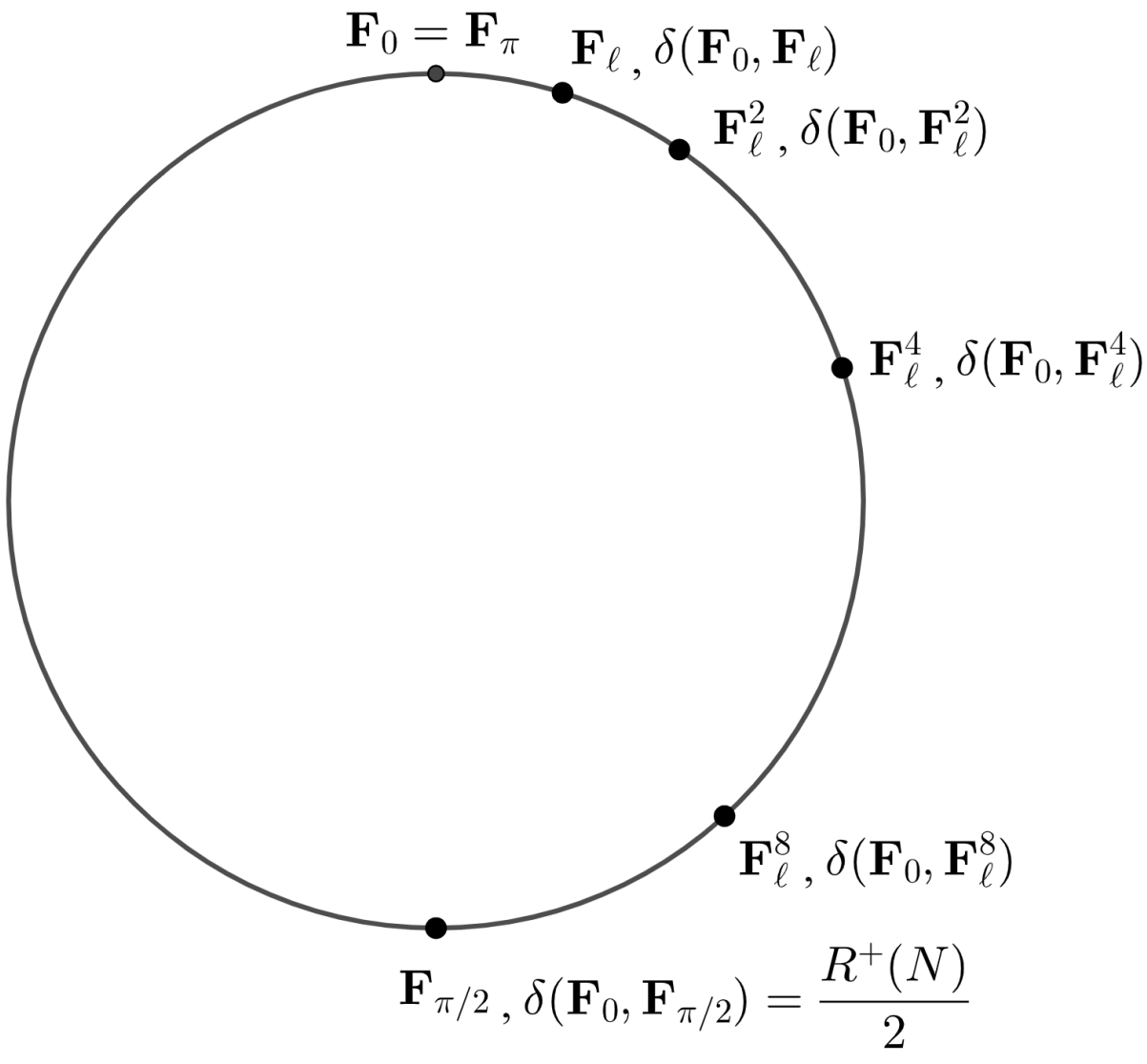


Table of Contents

Idea of our method for even period

We assume π even.

Operations

- ρ and ρ^{-1} steps: length-one jump in the cycle.
- Giant step: long jump in the cycle.

1. Use giant steps to compute $\bar{F}$, an approximation of $\mathbf{F}_{\pi/2}$.
2. Reach $\mathbf{F}_{\pi/2}$ computing the forms near $\bar{F}$.

Algorithm scheme, part (1)

Input

- $N > 0$ non-square such that $\pi \equiv 0 \pmod 2$
- $R^+(N)$

1. Starting from $\mathbf{F}_0$, compute the forms $\mathbf{F}_i$ for $i = 0, \ldots, \ell$, until $\delta(\mathbf{F}_0, \mathbf{F}_\ell) \geq 2\ln(4N) + 1$.

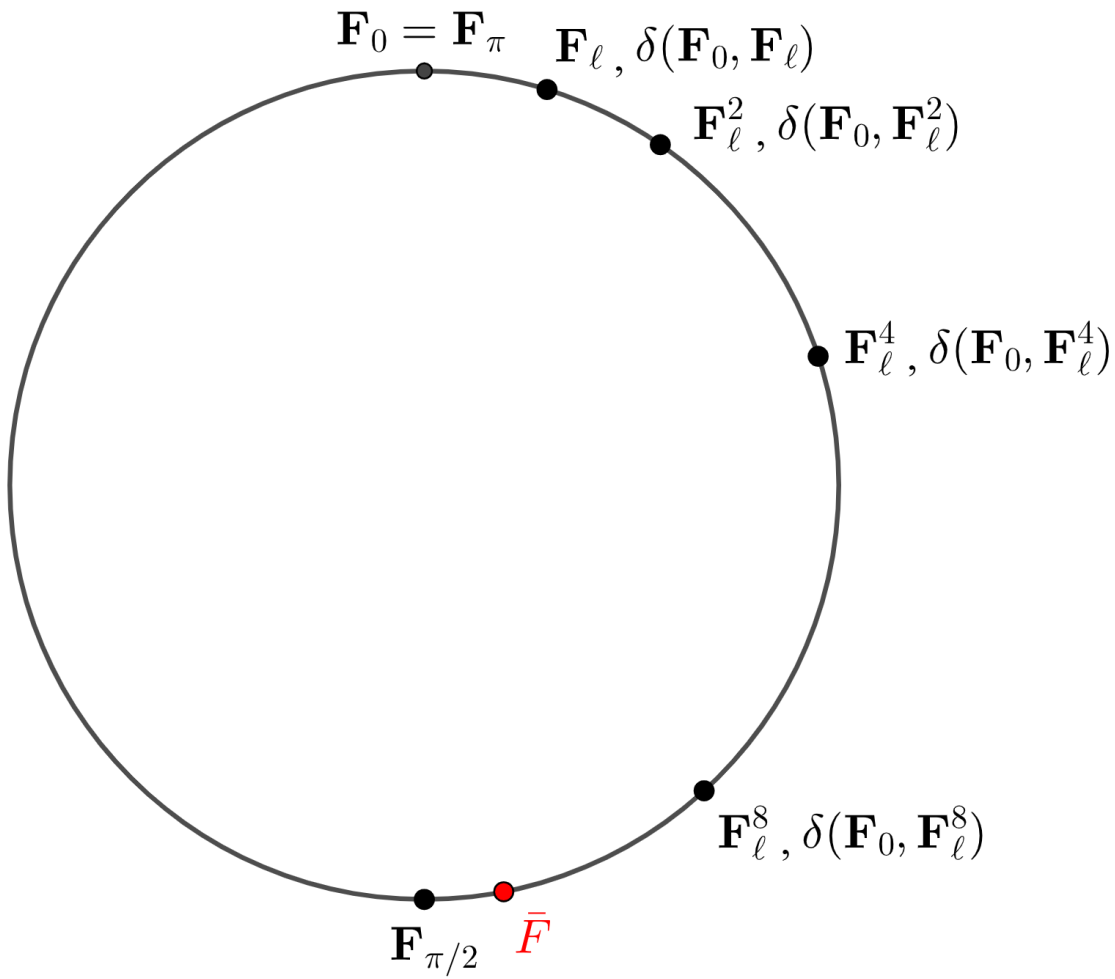
2. Compute the quadratic forms $\mathbf{F}_\ell^{2^i}$, using giant steps, and their distance $d_i \leftarrow \delta(\mathbf{F}_0, \mathbf{F}_\ell^{2^i})$, for $i = 1, \ldots, \tau$, with τ such that $d_{\tau-1} \leq R^+(N)/2 < d_\tau$.

Algorithm scheme, part (2)

3. Set $\bar{F} \leftarrow \mathbf{F}_\ell^{2^{\tau-1}}$ and $\bar{d} \leftarrow d_{\tau-1}$.
   For $i = \tau - 2, \ldots, 0$: if $\bar{d} + d_i < R^+(N)/2$, then update $\bar{F} \leftarrow \bar{F} \bullet \mathbf{F}_\ell^{2^i}$ and $\bar{d} \leftarrow \bar{d} + d_i$.

4. $\bar{F} \in \boldsymbol{\Upsilon}$. Compute $\rho(\bar{F}), \rho^2(\bar{F}), \rho^3(\bar{F}), \ldots$ and $\rho^{-1}(\bar{F}), \rho^{-2}(\bar{F}), \rho^{-3}(\bar{F}), \ldots$ until a non-trivial factor of $2N$ is found.

The cycle Υ if π even



$d_{k+1} = 2d_k + O(\ln N)$ for $k \geq 0$ and $\delta(\mathbf{F}_0, \bar{F}) = \bar{d} + O(\ln N)$.

Computational complexity

Theorem (Murru, S.)

The value of τ in the algorithm is at most $\left\lceil \log_2 \frac{R^+(N)}{2} \right\rceil = O(\ln N)$.

Theorem (Murru, S.)

The form $\bar{F}$ obtained at the end of the first phase of the method satisfies

$$\left| \delta(\mathbf{F}_0, \mathbf{F}_{\pi/2}) - \delta(\mathbf{F}_0, \bar{F}) \right| = O((\ln N)^2).$$

The computational complexity of our algorithm is $O((\ln N)^2)$.

Multiple of $R^+(N)$ and regulator

Since $\boldsymbol{\Upsilon}$ is periodic, and our goal is to find the quadratic form in the middle of some period, the method can be adapted to take in input $aR^+(N)$ for some $a \in \mathbb{N}^{>0}$.

1. If a odd, a factor of $2N$ is found in the position at distance $\frac{aR^+(N)}{2}$ from the beginning.
2. If a even, in a position at distance $\frac{aR^+(N)}{2}$ is found $\mathbf{F}_\pi$. In this case, the procedure is repeated with target the position at distance $\frac{aR^+(N)}{4}$. The process is iterated ℓ times until $\frac{a}{2^\ell}$ odd.

We have that $R^+(N) = kR(N)$ for $k \leq 6$, where $R(N)$ is the regulator of $\mathbb{Q}(\sqrt{N})$.

Computing the regulator

We look for a fast algorithm that computes an integral multiple of $R(N)$, or a good approximation of it.

Vollmer's method to compute $R(N)$:

- Monte Carlo algorithm
- Cost (under GRH): $O\left(\exp\left(\frac{3}{\sqrt{8}}\sqrt{\ln N \ln\ln N}\right)\right)$

Main factorization algorithms that make use of continued fractions or quadratic forms:

- CFRAC: $O(\exp(\sqrt{2\ln N \ln\ln N}))$
- SQUFOF: $O(\sqrt[4]{N})$

Table of Contents

Future research: Class Number Formula

Theorem (Class Number Formula)

Let $N > 0$ square-free. The following holds:

$$h(N)R(N) = \frac{\sqrt{D}}{2} L(1,\chi) = \frac{\sqrt{D}}{2} \sum_{x=1}^{\infty} \left(\frac{D}{x}\right) \frac{1}{x},$$

where:

- *$h(N)$ is the class number,*
- *χ is the Kronecker symbol $\left(\frac{D}{\cdot}\right)$,*
- *$L(1,\chi)$ is the L-function associated to χ,*
- *$D = N$ if $N \equiv 1 \pmod 4$ and $D = 4N$ otherwise.*

$$L(1,\chi) = \prod_{\text{prime } p} \left(1 - \left(\frac{D}{p}\right) \frac{1}{p}\right)^{-1}$$ Euler Product

Future research: approximation of $h(N)R(N)$

Two methods to estimating $h(N)R(N)$ by approximating $L(1,\chi)$.

Bach's method:

- Cost: $O(N^{1/5+\epsilon})$,
- Error (assuming ERH): $O(N^{2/5+\epsilon})$,
- Idea: Weighted average of truncated Euler Products to compute $S(Q,N)$, an approximation of $L(1,\chi)$.

Bibliography

Eric Bach (1995)
"Improved Approximations for Euler Products"

Future research: approximation of $h(N)R(N)$

Two methods to estimating $h(N)R(N)$ by approximating $L(1,\chi)$.

Srinivasan's method:

- Expected time: $O(N^{1/5+\epsilon})$,
- Error: $O(N^{2/5+\epsilon})$,
- Idea: The **Random Summation Technique**, i.e., taking random terms in the Euler Product expansion of $L(1,\chi)$.

Bibliography

Anitha Srinivasan (1998)
"Computations of class numbers of real quadratic fields"

Srinivasan's method: key points

1. $L(1,\chi) = \left(1 - \left(\frac{D}{2}\right)\frac{1}{2}\right)^{-1} \sum_{x \text{ odd}} \left(\frac{D}{x}\right)\frac{1}{x}$,
2. Approximate $S = \sum_{x \text{ odd and } x \leq D^2} \left(\frac{D}{x}\right)\frac{1}{x}$,
3. $M = \lceil D^{1/5} \rceil$ i.i.v. Y_i taking any odd integer value $1 \leq n \leq D^2$

$$\mathbb{P}(Y_i = n) = \frac{\lambda}{n} \quad \text{for } 1 \leq n \leq D^2 \text{ and } n \text{ odd},$$

4. Let X_i be the random variable $\left(\frac{D}{Y_i}\right)$ for $1 \leq i \leq M$,
5. Then, $S = \frac{1}{\lambda M}\mathbb{E}(X_1 + \cdots + X_M)$,
6. Approximate S with $\frac{1}{\lambda M}\sum_{i=1}^{M} X_i$.

Future research: Analytic Formula

Proposition

The following holds:

$$h(N)R(N) = \frac{1}{2}\sum_{x\geq 1}\left(\frac{D}{x}\right)\left(\frac{\sqrt{D}}{x}\operatorname{erfc}\left(x\sqrt{\frac{\pi}{D}}\right) + E_1\left(\frac{\pi x^2}{D}\right)\right),$$

where

- $\operatorname{erfc}(z) = 1 - \frac{2}{\sqrt{\pi}}\sum_{n=0}^{\infty}\frac{(-1)^n z^{2n+1}}{n!(2n+1)}$,
- $E_1(z) = -\gamma - \ln z - \sum_{n=1}^{\infty}\frac{(-1)^n z^n}{n\cdot n!}$,
- $\gamma = \int_1^{\infty}\left(-\frac{1}{x} + \frac{1}{\lfloor x\rfloor}\right)$ is the Euler–Mascheroni constant.

Thank you for your attention!