Fast square-free decomposition of integers using class groups

Erik Mulder

17 July 2024

Thanks to the organizers

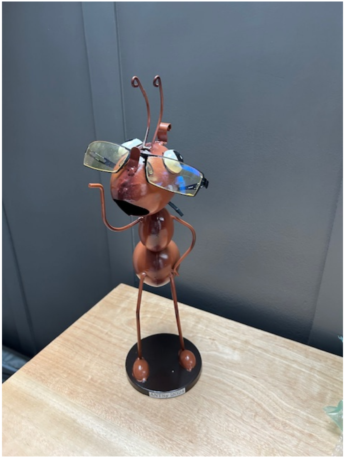

Figure: Cool guy

Introduction

Definition

Square-free decomposition problem: Let $n = a^2 b$ with $a, b \in \mathbb{N}$ and b square-free. Find a and b.

- Who doesn't like integer factorization?
- Cryptographic systems that use $n = p^2 q$ or $n = p^k q$ rely on the hardness of factoring of n.
- Computing ring of integers of number field
- Computing endomorphism ring of an elliptic curve over a finite field

Main result

Theorem

Assume some heuristic assumptions. Fix $0 \leq \alpha \leq 1$*. Then for all integers* $n = a^2 b$ *with* $b = n^\alpha$*, there is an algorithm that finds the square-free decomposition of* n *in expected time:*

$$\mathcal{O}(L_b[1/2, 1]) = \mathcal{O}(L_n[1/2, \sqrt{\alpha}]),$$

where

$$L_b[\alpha, c] = e^{(c+o(1)) \ln(b)^\alpha (\ln \ln(b))^{1-\alpha}}.$$

If a, b are distinct primes of roughly the same cryptographic size, then this is the current fastest method.

Binary quadratic forms

- $f(x, y) = ax^2 + bxy + cy^2 = (a, b, c)$, where $a, b, c \in \mathbb{Z}$.
- $D = D_f = b^2 - 4ac$ is the *discriminant* of f. We always have $D < 0$.
- f is *primitive* if $\gcd(a, b, c) = 1$.
- f, g are *equivalent* if there exists $A = \left(\begin{smallmatrix} p & q \\ r & s \end{smallmatrix}\right) \in \Gamma$ such that $f(x, y) = g(A \cdot (x, y)^T) = g(px + qy, rx + sy)$. Then $D_f = D_g$.
- If $|b| < a < c$, then f is *reduced*. Every form is equivalent to a unique reduced form.
- $C(D)$ is the *class group* of forms of discriminant D. With composition as its group operation. If $D = 0 \bmod 4$, then $e_D = (1, 0, \frac{-D}{4})$.
- The *class number* $h(D) \approx \sqrt{D}$ is the order of $C(D)$.
- Why not use ideals instead of forms?

Schnorr-Lenstra factorization

- Let $n \in \mathbb{N}$, possibly square-free
- Take a random $f \in C(-4n)$
- Construct k, consisting of powers of all primes up to some B
- Pray that $f^k = e_{-4n}$, i.e. $h(-4n)$ is smooth
- If so, try to construct g of order 2
- Then g is of the form $(d, 0, \frac{n}{d})$
- The complete factorization of n can be found this way
- If $h(-4n)$ is not smooth, try $C(-4ns)$ instead for some small square-free s
- Quite similar to the ECM

Derived forms I

Definition

Let p be a prime and D a discriminant. We say $f \in C(Dp^2)$ is *derived from* $g \in C(D)$ if there exists a 2×2 integer matrix A with $\det(A) = p$ such that $f(x, y) = g(A \cdot (x, y)^T)$.

Proposition

- *For every form $f \in C(Dp^2)$ there exists a unique $g \in C(D)$ up to equivalence such that f is derived from g.*
- *For each $g = (a, b, c) \in C(D)$ there are exactly $p - \left(\frac{D}{p}\right)$ inequivalent primitive forms in $C(Dp^2)$ that are derived from g.*
- *These forms are:*

$$(ap^2, p(b + 2ah), ah^2 + bh + c) \quad \textit{for } \ 0 \le h \le p - 1$$

$$\textit{and } (a, bp, cp^2)$$

Derived forms II

Corollary

$$h(Dp^2) = h(D) \cdot (p - \left(\frac{D}{p}\right)).$$

Proposition

If $f_1, f_2 \in C(Dp^2)$ are derived from $g_1, g_2 \in C(D)$ respectively, then $f_1 \cdot f_2$ is derived from $g_1 \cdot g_2$.

The new square-free decomposition algorithm I

From now on, $n = p^2 b$, with p prime.

Proposition

Suppose $g \in C(-4n)$ is derived from $e_{-4b} = (1, 0, b) \in C(-4b)$ and $g \not\sim e_{-4n} = (1, 0, n)$. Furthermore, suppose that g is reduced and $b > p^2$. Then

$$g = (p^2, 2pk, k^2 + b)$$

for some $-p/2 \leq k \leq p/2$.

Lemma

Suppose $g \in C(-4n)$ is derived from e_{-4b}. Let r be a prime with $r \neq p$. Lift g to some $h \in C(-4nr^2)$. Then $I = h^{r - \left(\frac{-4n}{r}\right)}$ is not only derived from e_{-4b}, but also from e_{-4br^2}.

The condition of the proposition now becomes: $br^2 > p^2$, so take $r > \sqrt{n}$.

The new square-free decomposition algorithm II

New factorization plan:

- Take random $f \in C(-4n)$, compute $g = f^k$
- If $g = e_{-4n}$, then continue as Schnorr-Lenstra
- Otherwise, pray that g is derived from e_{-4b}, i.e. $h(-4b)$ is smooth
- Let $r > \sqrt{n}$ be a prime.
- Lift g to some $h \in C(-4nr^2)$ and compute $I = h^{r - \left(\frac{-4n}{r}\right)}$.
- Reduce I and read off factor p^2 of n
- If $h(-4b)$ is not smooth, try $C(-4ns)$ instead for some small square-free s (then $h(-4bs)$ needs to be smooth)

Main theorem

Theorem

Assume some heuristic assumptions. Fix $0 \leq \alpha \leq 1$. Then for all integers $n = a^2 b$ with $b = n^\alpha$, there is an algorithm that finds the square-free decomposition of n in expected time:

$$\mathcal{O}(L_b[1/2, 1]) = \mathcal{O}(e^{(1+o(1))\sqrt{\ln(b)\ln\ln(b)}}) = \mathcal{O}(L_n[1/2, \sqrt{\alpha}]).$$

- The runtime follows from optimizing the size of the exponent k.

Stage 2: Electric Boogaloo

- In stage 1, our k consisted of all primes up to some B.
- We then computed $g = f^k$ and hoped that g is derived from e_{-4b}.
- If not, then quite often we are missing just one prime factor of $h(-4b)$.
- In stage 2, we take some $B_2 > B$ and check for each prime $q \in [B, B_2]$ separately if g^q is derived from e_{-4b}.
- Using a generic method, we can take $B_2 = B\ln(B)$ without increasing the asymptotic runtime per $C(-4ns)$ that we try.
- Roughly a factor $\ln(\ln(b))$ fewer groups have to be tried this way.

Stage 2 jealousy

- Algorithms like ECM have a much better stage 2
- Completely different, with FFT or Pollard rho method
- Roughly $B_2 = B^2$ can be used there (instead of $B_2 = B \ln(B)$)
- Roughly a factor $\ln(b)$ fewer groups would have to be tried that way.
- Please help!

Timings I

Let $n = p^2 q$ with $p \approx q$.

	$q \approx 10^{15}$	$q \approx 10^{20}$	$q \approx 10^{25}$	$q \approx 10^{30}$	$q \approx 10^{35}$
Mean time with stage 2	0.11	0.80	5.16	30.30	135.91
Mean time only stage 1	0.16	1.63	11.65	66.71	357.74
Mean ECM time	0.07	1.01	14.15	141.78	1549.16

Table: Comparison between factorization algorithms, in seconds

Timings II

	$q \approx 10^{20}$	$q \approx 10^{30}$	$q \approx 10^{40}$	$q \approx 10^{50}$
Mean time NFS	61.17 s	22.82 m	572.51 m	~ 8 d
Mean time with stage 2	0.72 s	33.63 s	9.74 m	174.47 m
Mean number of groups	6.19	26.80	63.54	206.36

Table: Comparison with the NFS

Thank you!

Please find a better stage 2 :)