Introduction to the Shank's SQUFOF Integer Factoring Algorithm

D. J. Guan *

May 26, 2010

Abstract

Shank's SQUFOF integer factoring algorithm works with integers which are at most $2\sqrt{n}$. It can be implemented very efficiently for factoring 62-bit integers in a 32-bit computer.

Theory of SQUFOF Algorithm

A *binary quadratic form* (a, b, c) is a polynomial in x and y

$$f(x, y) = ax^2 + bxy + cy^2$$

$$= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix},$$

where a, b, and c are integers.

Quadratic forms and ideals are related to quadratic numbers.

Let $D = b^2 - 4ac > 0$.

$\frac{-b + \sqrt{D}}{2|a|}$ is the quadratic number in $\mathbf{Q}(\sqrt{D})$ associated to the form (a, b, c)

Ideals are used in conceptual proofs.

Quadratic forms are used in computation.

Two quadratic forms f and g are *equivalent* if there exists an integer matrix $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ of determinant equal to 1 such that

$$g(x,y) = f(\alpha x + \beta y, \gamma x + \delta y)$$

$$= \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$

Theorem 1 *Equivalence preserves the discriminant* $D = b^2 - 4ac$.

Primitive and Reduced Quadratic Forms

A quadratic form (a, b, c) is *primitive* if $\gcd(a, b, c) = 1$.

A quadratic form (a, b, c) of discriminant $D > 0$ is *reduced* if

$$\left|\sqrt{D} - 2|a|\right| < b < \sqrt{D}.$$

Theorem 2 *If (a, b, c) is a reduced form, then*

1. *a and c are of opposite sign.*
2. *$|a| + |c|$ and b are less than $\sqrt{D}$.*

Theorem 3 *A form (a, b, c) is reduced if and only if*

$$\left|\sqrt{D} - 2|c|\right| < b < \sqrt{D}.$$

Theorem 4 *Let* $\tau = \dfrac{-b + \sqrt{D}}{2|a|}$ *be the quadratic number associated to the form* (a, b, c). (a, b, c) *is reduced if and only if* $0 < \tau < 1$ *and its conjugate* $-\sigma(\tau) > 1$.

Reduction Operator ρ

Let $D > 0$ be a discriminant, $a \neq 0$ and b are integers.

The reduction operator ρ is defined as

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

where $r(b, a)$ is the unique integer r satisfying

1. $r \equiv b \pmod{2a}$, and
2. if $|a| > \sqrt{D}$ then $-|a| < r \leq |a|$,
   if $|a| < \sqrt{D}$ then $\sqrt{D} - 2|a| < r \leq \sqrt{D}$.

The inverse of ρ is given by

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

Reduction of Indefinite Quadratic Forms

while (a, b, c) is not reduced **do**
$(a, b, c) = \rho(a, b, c)$

Theorem 5 *The number of iterations is at most* $2 + \left\lceil \log \frac{|c|}{\sqrt{D}} \right\rceil$.

Theorem 6 *If* (a, b, c) *is a reduced form then* $\rho(a, b, c)$ *is also a reduced form.*

Theorem 7 *The reduced forms equivalent to* (a, b, c) *are exactly the forms* $\rho^k(a, b, c)$ *for sufficiently large* k.

Composition of Forms

The composition of forms come from the product of ideals.

Let (a_1, b_1, c_1) and (a_2, b_2, c_2) be two forms of the same discriminant D, $s = (b1 + b_2)/2$, $t = (b_1 - b_2)/2$.

Let u, v, w, and d be such that $ua_1 + va_2 + ws = d = \gcd(a_1, a_2, s)$, $d_0 = \gcd(d, c_1, c_2, t)$.

The composition of (a_1, b_1, c_1) and (a_2, b_2, c_2) is defined as

$$(a_3, b_3, c_3) = \left(\frac{d_0 a_1 a_2}{d^2},\ b_2 + \frac{2a_2(v(s - b_2) - wc_2)}{d},\ \frac{b_3^2 - D}{4a_3} \right).$$

If $\gcd(a, b) = 1$ then $(a, b, c)^2 = \left(a^2,\ b,\ \frac{c}{a} \right)$

Ambiguous Forms

A quadratic form (a, b, c) for discriminant D is *ambiguous* if

$$(a, b, c)^2 \equiv (1, x, y) \bmod \Gamma_\infty$$

for some integers x and y, where $\Gamma_\infty = \left\{ \left(\begin{array}{cc} 1 & m \\ 0 & 1 \end{array} \right) \middle| \, m \in \mathbf{Z} \right\}$.

If (a, b, c) is ambiguous then $a \mid b$, which implies that

$$D = b^2 - 4ac = (ka)^2 - 4ac = a(k^2 a - 4c).$$

Therefore, a is a factor of D.

Find an Ambiguous Form

Start from the identity form $f = \left(1,\ b,\ \dfrac{b^2 - D}{4}\right)$.

Find a form $g = (x^2, y, z)$ by repeatedly applying the ρ function, i.e., compute $\rho(f), \rho^2(f), \ldots, \rho^k(f) = g$.

Let $h = (x, y, xz)$.

1. h is not primitive.
   Let p be a prime dividing x and y, then $p^2 \mid D = y^2 - 4x^2z$,
   i. e., p^2 is a factor of D.

2. h is primitive.
   $h^2 = g$. Let $f = h^{-1} = (x, -y, xz)$. f will be on an ambiguous cycle. We can find an ambigous form g by repeatedly applying the ρ function, i.e., compute

$$\rho(f), \rho^2(f), \ldots, \rho^l(f) = g.$$

Theorem 8 *There is an ambiguous from g in the cycle of f at $l = k/2$.*

Two problems with the above method:

1. Some ambiguous forms will correspond to a trivial factor of n.

2. No guarantee to find a square form other than the identity.

Solution to problem 1:

Trivial factors of n occurs when h lies in the priciple cycle itself.

This implies that $a^2 < \sqrt{D}$, which occurs quite rare.

We can store these values and avoid them in the search.

Problem 2 are more basic:

When the principle cycle is short, there may not be another square form. For example, when $n = (2k+1)^2 + 4$, the length of the cycle is 1.

The solution is to work with many D's, which is a small multiple of n. The chance of all the D's has short cycle is small.

Note that most implementation of the SQUFOF is a probabilistic algorithm. It may fail to find a nontrivial factor of n.

The SQUFOF Algorithm

input: n
output: a factor of n
algorithm:

if n is prime **then** return(1)
if n is a square **then** return($\sqrt{n}$)
if $n \equiv 1 \pmod 4$ **then**
  $D = n$, $d = \left\lfloor \sqrt{D} \right\rfloor$, $b = \lfloor (d-1)/2 + 1 \rfloor$
else
  $D = 4n$, $d = \left\lfloor \sqrt{D} \right\rfloor$, $b = \lfloor d/2 + 1 \rfloor$
end
$f = (a, b, c) = (1, b, (b^2 - D)/4)$, $Q = \emptyset$
for $i = 1, 2, \ldots$ **do**

$f = (a, b, c) = \rho(a, b, c)$
if $|a| \leq \left\lfloor \sqrt{d} \right\rfloor$ **then** $Q = Q \cup \{|a|\}$
if i is even **then**
if $a = s^2$ for some integer $s \notin Q$ **then**
if $t = \gcd(a, b, D) > 1$ **then**
return(t^2)
else
$f = (a, b, c) = (s, -b, sc)$
while f is not reduce **do** $f = (a, b, c) = \rho(a, b, c)$
do $s = b$; $f = (a, b, c) = \rho(a, b, c)$ **while** $s \neq b$
if a is even **then** $a = a/2$
return($|a|$)
end
end
end
end

References

1. Henri Cohen. *A Course in Computational Algebraic Number Theory*. Springer-Verlag, 1993.

2. Hans Riesel. *Prime Numbers and Computer Methods for Factorization*. Birkhäuser, 1994.

3. Robert D. Silverman. The multiple polynomial quadratic sieve. *Mathematics of Computation*, 48(117):329–339, January 1987.