A Monte Carlo Factoring Algorithm
With Linear Storage
By C. P. Schnorr* and H. W. Lenstra, Jr

Abstract We present an algorithm which will factor an integer \( n \) quite efficiently if the class number \( h(-n) \) is free of large prime divisors. The running time \( T(n) \) (number of compositions in the class group) satisfies prob[\( T(m) \leq o(n^{1/2}) \)] \( \geq (r - 2)^{-1} \) for random \( m \in [n/2, n] \) and \( r > 2 \). So far it is unpredictable which numbers will be factored fast. Running the algorithm on all discriminants \( -n \) with \( 2 \leq r \) and \( r = \sqrt{\ln n / \ln \ln n} \) every composite integer \( n \) will be factored in \( o(\exp(\sqrt{\ln \ln n}) \) bit operations. The method requires an amount of storage space which is proportional to the length of the input \( n \). In our analysis we assume a lower bound on the frequency of class numbers \( h(-m), m \leq n \) which are free of large prime divisors.

1. Introduction. The problem of factoring an integer \( n \) into its prime power divisors is computationally equivalent to determining all ambiguous, reduced positive forms \( ax^2 + bxy + cy^2 \) (notation \( (a, b, c) \), \( a, b, c \in \mathbb{Z} \) with discriminant \( b^2 - 4ac = -n, b^2 - 4ac = +n \), resp.) In fact these ambiguous forms correspond to the relatively prime factorizations of \( n \), i.e., to the pairs \( (n_1, n_2) \) with \( n = n_1n_2 \), \( \gcd(n_1, n_2) = 1 \).

According to Gauss [5] the equivalence classes of forms with fixed discriminant \( \Delta \) form a group under composition, the class group \( G(\Delta) \). The order \( h(\Delta) \) of this group is the class number. Multiplication in \( G(\Delta) \) can be done efficiently working with representatives of classes. The ambiguous classes are the classes \( H \) with \( H^2 = 1 \).

In case of negative discriminants \( \Delta < 0 \) there is a unique reduced form in each class, and this form can be efficiently calculated from any other class representative. Therefore, factoring \( n \) is computationally equivalent to determining representatives of all ambiguous classes in \( G(-n) \). The reduced forms of these classes correspond to the relatively prime factorizations \( n_1n_2 = n \) of \( n \).

In case of positive discriminants \( \Delta > 0 \), under a different concept of reduction, there are \( O(\sqrt{\Delta} \ln \Delta) \) reduced forms in each class. They form a cycle under the reduction operation. Composition of forms yields a group-like structure on the principal cycle. A reduced form \( (a, b, c) \) is ambiguous if its square under composition yields the unit form. Again, the ambiguous, reduced forms with discriminant \( n > 0 \) correspond to the relatively prime factorizations \( n_1n_2 = n \) of \( n \).

Several factoring algorithms have been developed on this basis. Here we are only concerned with negative discriminants. For positive discriminants see the algorithms of Shanks as described in Monier [11] and Wagstaff-Wunderlich [22].
discriminant Shanks [17] after guessing generators for $G(-n)$ computes the class number $h(-n)$ by exploiting the group structure. Then Shanks computes $H(k) = H^{1/2}$ for the smallest $k$ such that $H(k) \neq 1$ with $H \in G(-n)$ chosen arbitrarily. Clearly $H(k)$ is ambiguous. Under reasonable assumptions it takes $O(n^{1/2})$ steps to factor $n$ in this way. This method can be speeded up to an $O(n^{1/3})$-algorithm by approximating $h(-n)$ via the class formula (let $(\cdot | \cdot)$ denote the Kronecker symbol).

$$\lim_{n \to \infty} \frac{\sqrt{n}}{\pi} \prod_{p \mid m} \frac{p - (-1)}{(p - 1)^2} = h(-n),$$

which by the generalized Riemann hypothesis has an error term $O((\ln(\ln n))^{1/2})$. For this algorithm the amount of storage will be proportional to the running time.

In Schnorr (19) a method was proposed to generate ambiguous forms which is similar to the Morrison-Brillhart factoring algorithm. We collect equations

$$H_r^2 - \prod_{p \mid m} H_p^k \equiv a_p \pmod{r},$$

with $H_r \in G(-n)$ chosen at random and $H_r = ([p \ b \ c \ d])$ for small primes $p$. By combining these equations one obtains

$$H = \prod_{p} H_p^k \quad F = \prod_{p} H_p^{k \cdot a_p \pmod{r}}$$

such that $H^{-1} \equiv F^2 \pmod{F}$. Then $HF^{-1}$ is ambiguous. Under reasonable assumptions $n$ will be factored with $O(\exp \frac{1}{\ln \ln n})$ steps and $O(\exp \frac{1}{\ln \ln \ln n})$ storage.

The new algorithm, given the first $t$ primes $p_1 < p_2 < \ldots < p_t = n^{1/2}$, needs only to store a fixed number of forms which takes $O(t \log n)$ bits. Let $e_1 = \max \{r, p_1 \}$. Then Stage 1 of the new algorithm computes

$$H = H_{n, r}^{1/2}$$

for an arbitrarily chosen $H_{n, r} \in G(-n)$. Then compute $H^{1/2}$ for the smallest $k \leq \log \ln n$ such that $H^{1/2} = 1$. Clearly $H^{1/2}$ is ambiguous $n$ will be factored by Stage 1 if $h(-n)$ divides $2^{2k}H^{1/2}$ for some $k$. If Stage 1 fails then Stage 2 does a random walk through the group generated by $H$.

Stage 2 will factor $n$ if ord $H \cdot q \leq p_1$, for some $k \leq v$ if $h(-n)$ divides $2^{2k}P_1 \cdot q$, for some $q < p^2$. With $p_1 = n^{1/2}$, Stage 1 of the algorithm takes $O(p_1)$ compositions and for random composite $m = [0 | 0]$ with probability $\geq t^{-1}$ detects a proper divisor of $n$. Stage 2 also takes $O(p_1)$ compositions and with probability $\geq (t - 2)^{2r}$ $r \geq 2$ detects a proper divisor of $m$. Running Stage 1 on the integers $m$ for $1 \leq r$, $r = \sqrt{\ln n}$, every composite integer $n$ will be factored within $O(\exp \sqrt{\ln \ln n})$ bit operations. The latter bound already takes into account the cost of the arithmetic. The cost for a composition in $G(-n)$ is proportional to the cost of the extended Euclidean algorithm which given integers $a, r \leq \sqrt{n}$ computes $r = \gcd(a, v)$ using standard algorithms for multiplication and division this takes $O(\ln n)$ bit operations or binary Boolean operations see Knuth [17, 4.5.2, exercise 30, and algorithm X].

The particular features of the new factoring algorithm are:

1. It can easily be operated with $O(\log n)$ bit storage.
(2) it is Monte Carlo in the sense that every 1000th integer will be factored about 1000 times faster than average time,

(3) the integers which will be factored very fast are randomly distributed, there is no way to predict whether a given \( n \) will be factored fast,

(4) the algorithm is of the parallel type, e.g. 1000 processors will factor 1000 times faster

Properties (2) (3) seem to endanger the RSA-cryptoscheme, see Rivest et al [15]

In particular no methods are known that generate class numbers with large prime divisors

Stages 1 and 2 of the algorithm are presented in Sections 2 and 3. The main algorithm which factors arbitrary integers is given in Section 4. Some computational experience with the factoring algorithm is reported in Section 5. In Appendix I we collect basic theorems and algorithms on quadratic forms. Appendix II contains various tables which demonstrate the performance of the algorithm. We exemplify the distribution of class numbers and integers which are free of large prime divisors, the frequency of class numbers divisible by small primes, and the performance of various pseudo-random functions used in Stage 2 of the algorithm.

2 Stage 1 of the Algorithm

Let \( n \) be the integer to be factored \( -n \) is the discriminant of some quadratic form \( \phi \) if and only if \( -n = 1 \mod 4 \) or \( -n = 0 \mod 4 \)

The purpose of Stages 1, 2 is to find a nontrivial divisor of \( n \) provided \( -n \) is a discriminant and \( h(-n) \) is a product of small primes. In order to factor general integers \( n \) the main algorithm in Section 4 applies Stages 1, 2 to multiples \( n' \) with \( -n' = 0 \mod 4 \). If \( -n \) is a discriminant we can easily construct forms \( \langle a, b, c \rangle \) with discriminant \( -n \) choose a small odd prime \( p \) with \( (p) \neq 1 \) and solve \( b' = -n \mod 4p \) which yields \( b' = -n + 4pc \) for some \( c \in \mathbb{Z} \). Hence \( \langle p, b', c \rangle \) has discriminant \( -n \).

Throughout Sections 2, 3 we restrict ourselves to the case \( -n = 1 \mod 4 \) consult Theorem III, Appendix I for the case \( -n = 0 \mod 4 \). Then the unit \( \tau \in \mathbb{Z}(-n) \) is represented by the form \( (1, (1 + n)/4) \). This ambiguous class yields the improper factorization \( 1 = n \). The other ambiguous classes correspond in a 1:1 way to the relatively prime factorizations of \( n \) with nontrivial divisors.

Stage 1

1. \( n \in \mathbb{N}, -n = 1 \mod 4 \) given
2. For some \( t \in \mathbb{N} \) compute the \( t \) first primes \( p_1 = 2, p_2 = 3, \ldots, p_t \)
3. Choose \( H_0 \in \mathbb{Z}(-n) \) arbitrarily
4. \( H = H_{1'} - r \) with \( c_1 = \max \{ r \in \mathbb{Z} : p' \in p_r \} \) if \( H = 1 \), then stop (in this case \( \text{ord}(H_0) \) is odd and another \( H_0 \) must be chosen)
5. \( H = H, e_* = \left[ \log \sqrt{n} \right] \)
6. go to Stage 2
7. (at this point \( s \) is ambiguous and yields some divisor \( d \) of \( n \))

Stage 1 by itself is the core of the new factoring algorithm. The improvements resulting from Stage 2 are important for practical applications but they scarcely influence the asymptotic time bound of the main algorithm.

\[ \text{Fact 1: Suppose } \text{ord}(H_1') = p \text{ and } \text{ord}(H_0) \text{ is even then Stage 1 generates an ambiguous class } S 
eq 1 \]
In case $-n \equiv 1 \mod 4$ every ambiguous class $S \neq 1$ yields a proper divisor of $n$. In particular, when $n$ has $d$ odd prime divisors, then $2^{d-1} h(-n)$ and there are exactly $2^{d-1}$ ambiguous classes corresponding to the $2^{d-1}$ pairs $(n_1, n_2)$ with $n_1 n_2 = n$, $n_1 \leq n_2$, $\gcd(n_1, n_2) = 1$. Moreover, when $n$ is composite and $H_0 \in G(-n)$ is chosen at random, then $\text{prob} [\text{ord}(H_0) \text{ even}] \geq 1/2$. Hence Stage 1 has a chance $\geq 1/2$ to find a proper divisor of $n$, provided $h(-n) | \Pi'_{i=1} p_i$. A few repetitions of Stage 1 almost surely generate a proper divisor of $n$, provided $h(-n) | \Pi'_{i=1} p_i$ and $n$ is composite.

**Fact 2.** Suppose $h(-n) | \Pi'_{i=1} p_i$ and $n$ is composite. If Stage 1 is passed with $H_0$ chosen independently $k$ times, then with probability $\geq 1 - 2^{-k}$ a proper divisor of $n$ has been found.

Next consider the chance that for random $m < n$

$$h(-m) | \prod_{i=1}^{t} p_i^r, \quad e_i = \max \{v \mid p_i^v < p_i^r\}$$

Siegel [21] proved

$$\forall \epsilon \exists n, \forall m \neq n, h(-m) \in [m^{1/2 - \epsilon}, m^{1/2 + \epsilon}]$$

We will base the analysis of Stage 1 on the following hypothesis:

For all $n$ and $t$

$$\# \left\{ m < n \mid h(-m) | \prod_{i=1}^{t} p_i^r \right\} / (0.5 n)$$

(21)

$$\geq \# \left\{ m < \sqrt{n} \mid m | \prod_{i=1}^{t} p_i^r \right\} / \sqrt{n}$$

H.G. Franke has tested this hypothesis experimentally, see Tables 1–3, Appendix II for $n \approx 4.7 \times 10^8$. In general the first term in (21) is considerably larger than the second. Note that the frequency of class numbers $h(-n)$ of fundamental discriminants (and a fortiori of general discriminants) which are divisible by $p$ is larger than $1/p$ and is close to $1/((p - 1)$ provided $p$ is small with respect to $n$. These experimental data and some recent calculations of Cohen and H.W. Lenstra, Jr indicate that class groups $G(-n)$ of fundamental discriminants are distributed like random Abelian groups of order $O(\sqrt{n} \log n)$. Cohen and Lenstra have calculated $\text{prob} [m \text{ divides } |G|]$ for $m = 2, 3, \ldots$ and for random Abelian groups, where the probability weight of $G$ is proportional to $1/|\text{Aut}(G)|$. The values of Cohen and Lenstra completely match with our experimental data. A corresponding observation with respect to prime discriminants has been made by Leopoldt as cited in Zimmer [23].

Recently Canfield, Erdos, and Pomerance improved the theoretical lower bound on the second term in (21). We refer in particular to the proof in Pomerance [14].

**Theorem 3** Let $\Psi(n, \omega) = \# \{x < n \mid \text{free of primes } > \omega\}$. For every $\epsilon > 0$ there exists $c$, such that for all $n \geq 10$ and all $r$ with $n^{1/2 - \epsilon} \geq \Psi(n, n^{1/2 - \epsilon}) / n \geq (c, r \ln r$)

In practice, however, $\Psi(n, n^{1/2}) / n$ is larger than the bound stated in Theorem 3. From experimental data, see Table 2, Appendix II, we conclude

(22) for all $n$ and $r < \sqrt{n} / \ln \ln n$, $\Psi(n, n^{1/2}) / n \geq r$.
A MONTE CARLO FACTORING ALGORITHM WITH LINEAR STORAGE

COROLLARY 4 Assume (2 2) and (2 1) Then for all \( n \geq 10^{20} \) and all \( p_i = n^{1/2^r} \) with 
\[ r \leq \sqrt{\ln n} / \ln \ln n \]

\[ \# \left\{ m \leq n \mid h(-m) \prod_{i=1}^{t} p_i^{e_i} \right\} / (0.5n) \geq 0.83 r^{-r} \]

Proof

\[ \# \left\{ m \leq n \mid h(-m) \prod_{i=1}^{t} p_i^{e_i} \right\} / (0.5n) \]

(2 1)

\[ \geq \Psi(\sqrt{n}, n^{1/2^r}) / \sqrt{n} - \sum_{i=1}^{t} p_i^{e_i - 1} \]

(2 2)

\[ \geq r^{-r} - n^{-1/r} \]

\[ \geq r^{-r} - 1/\ln n^{1/2^r} / \ln n^{1/2^r} \]

(since \( t = \pi(n^{1/2^r}) \leq 1/\ln n^{1/2^r} / \ln n^{1/2^r} \))

\[ \geq r^{-r} - 2/\ln n \]

\[ \geq r^{-r}(1 - 2/\ln n) \]

(in fact \( r \leq \sqrt{\ln n} / \ln \ln n \) implies \( r \geq n^{-1/2^r} \))

\[ \geq 0.83 r^{-r} \quad \text{for} \quad n > 10^{20}, r \leq \sqrt{\ln n} / \ln \ln n \]

Runtime of Stage 1 If \( H^p \) is computed by the binary method (see Knuth [7, §4 63]), this takes

\[ 2 \log_2 p_i^{e_i} \leq 2 \log_2 p_i^{2^r} \leq 4 \log_2 p_i \]

compositions in \( G(-n) \) Since there are about

\[ t \leq p_i / \ln p_i \]

primes \( \leq p_i \), this yields a worst case bound of

\[ \frac{4}{\ln 2} p_i = 5.8 p_i \]

compositions in total

On the average, the binary method is somewhat more efficient It takes about 1.5 \( \log p_i^{e_i} \) compositions to compute \( H^p \) and therefore Stage 1 will only take about 4.4 \( p_i \) compositions in total All together we have proved the following

THEOREM 5 Assume (2 1), (2 2) and that for every discriminant \( m \leq n \) (a single)

\( H_0 \in G(-n) \) in Stage 1 is chosen at random Then for all \( n \geq 10^{20} \) and all \( p_i = n^{1/2^r} \) with \( r \leq \sqrt{\ln n} / \ln \ln n \) Stage 1 factors at least a 0.83nr \# fraction of the discriminants \( \leq n \) and takes about 4.4 \( p_i \) compositions in \( G(-n) \)

Remark The discriminants which will be factored are "randomly" distributed in \([0, n]\)

For practical applications we advise to choose somewhat smaller exponents \( e_i' \) instead of the \( e_i \)

\[ e_i' = \max \{ r \mid p_i^{e_i'} \leq p_i \} \quad \text{with} \quad p_i = n^{1/2^r} \]
We used the larger \( e \) for proving Corollary 4 by a crude argument. Assuming
\[
\psi(n^{1/2})/n = O(r^{-r})
\]
one obtains Corollary 4 for the \( e \):
\[
\approx \left( \frac{m \leq n}{n} \frac{|h(-m)|}{\prod p_i} \right)/(0.5n)
\]
(assuming that (2.1) holds for the \( e \))
\[
> r^{-r} - \sum_{p_i} e_{p_i}^r \frac{\psi(n^{1/2-1/r})}{n^{1/2}}
\]
\[
> r^{-r} - O\left( \frac{n^{1/2-1/r}}{\ln(n^{1/2})} (r - 1)^{-r} \right)
\]
\[
> r^{-r} - O\left( \frac{2r}{\ln n} (r - 1)^{-r} \right)
\]
\[
r^{-r} - O\left( (r - 1)^{-r}/(r \ln n) \right) = r^{-r} (1 - O(1/\ln \ln n))
\]
(since \( r \leq \sqrt{\ln n/\ln \ln n} \))

The choice of the \( e \) is justified by our data in Appendix II. Tables 1, 2 show that there are only a few discriminants \( h(-m) = m \leq 4 \times 10^8 \) such that \( h(-m) = \prod_{i=1}^l p_i \) with \( \Delta \leq e \), for some \( i \geq 2 \).

Table 4, Appendix II by Odlyzko considers large integers. This table shows that for \( r \leq \sqrt{\ln n/\ln \ln n} \)
\[
\# \{ m \leq n, m \text{ even} \text{ } | \text{lcm}(2, 3, p_i) \} / n > r^{-r}
\]
Note that \( m \text{lcm}(2, 3, p_i) = m = \prod_{i=1}^l p_i \) with \( \Delta \leq e \), for all \( r \). Hence in practice Corollary 4 even holds when the \( e \) are taken for the \( e \), and the constant 0.83 can be replaced by some constant > 1. If we use the \( e \) then Stage 1 will only take about 2.2 \( r \) compositions.

3. Using a Pollard-Brent Recursion in Stage 2 If \( h(-n) = 2^r \prod_{i=1}^l p_i \) with \( e_i = \lceil \log n/\log p_i \rceil \), then Stage 1 fails to factor \( n \) and computes
\[
\overline{H} = H^{11} \quad H = \overline{H}^{11}
\]
Stage 2 uses \( H \) and will most likely find a proper divisor of \( n \) within \( O(r) \) steps, provided that \( \text{ord}(H) \leq p_i \) and \( \text{ord}(H_0) \) is even.

Stage 2 generates a random walk through the cyclic group \( \langle H \rangle \) with generator \( H \). With some function \( f \), \( \langle H \rangle \rightarrow \langle H \rangle \) let
\[
H_i = H \quad H_{i+1} = f(H_i)
\]
The function \( f \) must be chosen such that

(3 1) \( f \) is easy to compute
(3 2) \( f \) is sufficiently random
(3 3) every relation \( H_j - H_k \) with \( j \neq k \) yields an ambiguous class \( s \) depending on \( H \) if \( j \neq k \)

It is known (see Knuth [7, Exercise 31.12]) that \( \Delta < k < \sqrt{\pi/2} p_i \), and \( H_j - H_k \) can be expected if \( f \) is sufficiently random and \( \text{ord}(H) \leq p_i \).
We have two methods to design/and to associate the ambiguous class \( S \) to \( H, f, j \), \( k \). Both methods will produce ambiguous classes \( S \) with \( S \neq 1 \) whenever \( \text{ord}(H) \) is even. Experience must decide which of the methods is more efficient.

**Method 1** For some \( q \in \mathbb{N} \) choose random integers \( a_i \in \{ p^1, 2 p^2 \} \) for \( i = 1, \ldots, q \). Precompute \( F_i = H^{a_i} = 1, \ldots, q \). For some random function \( g \) \( \langle H \rangle \rightarrow \{ 1, \ldots, q \} \) and recursively compute

\[
H_1 = H, \quad H_{i+1} = H \cdot F_{g(H_i)}
\]

Use the procedure `search` below in order to find some \( j < k \) with \( H_j = H_k \). Then

\[
H_i H_j = H_k \quad \text{with} \quad T = \sum_{i<j} g(H_i)
\]

Most likely we will have \( k \leq 2p \), which implies \( T \leq 4p^3 \).

Now suppose that \( \text{ord}(H) \equiv 2 \mod 2^r \). We can easily compute \( \zeta \) with \( T \equiv 2^r \mod 2^{r+1} \). Then \( H^{T^2} \) has order \( 2^r \) and yields an ambiguous class

\[
S = H^{T^2} \quad \text{with} \quad S \neq 1 \quad \text{provided} \quad r \geq 3
\]

**Comment** A theoretical analysis of this method has been done by Sattler and Schnorr [16]. For small values of \( q \), e.g., \( q = 2 \times 3 \times 4 \), the commutativity of the recursion steps increases the number of recursion steps as compared with a pure random recursion \( f \) \( \langle H \rangle \rightarrow \langle H \rangle \). By experience this slowdown is negligible as soon as \( q \gg 16 \). We have tested this recursion scheme in class groups (see Table 5, Appendix II) and in cyclic groups \( \langle H \rangle = \mathbb{Z}/n\mathbb{Z} \), in particular with \( n \) prime. Method 1 even works well for nonrandom \( a_i \), like \( a_i = \epsilon^i \) with \( \epsilon \) fixed. The advantage of Method 1 over Method 2 is that it explicitly yields a multiple \( T \) of the order of \( H \). Also, Method 1 only takes a single group operation (i.e., composition in the case of class groups) per recursion step.

**Method 2** Choose a random function \( g \) \( \langle H \rangle \rightarrow \{ 1, \ldots, q \} \), choose random values \( a_1, a_2 \in \{ p^1, 2 p^2 \} \) and precompute \( F_i = H^{a_i} = 1, \ldots, q \).

**Recursion on \( H \)** (We compute \( H = H^1 \) and \( d_i = a_i \mod 2^{12} \))

\[
H_1 = H, \quad d_1 = 1
\]

for \( i = 1, 2, \ldots \) till `search` finds some \( j < k \) with \( H_j = H_k \).

\[
(H_{i+1}, d_{i+1}) = \begin{cases} (H^{a+3d} \mod 2^{12}), & \text{if} \ g(H_i) \leq q/2 \\ (H_1 H_{g(H_i)} d_i + a_{g(H_i)} \mod 2^{12}), & \text{otherwise} \end{cases}
\]

Use the procedure `search` below in order to find some \( j < k \) with \( H_j = H_k \). Since \( H_1 = H^1 \) and \( H = H^{2^r+1} \), it follows that \( H^{T+1} \equiv 1 \). We compute \( i \) such that \( d_i - d_{i-1} = 2 \mod 2^{r+1} \). Almost surely \( i \) will be less than \( 32 \) and this implies \( \epsilon^i = \epsilon^{2m} \) for some odd \( m \). It remains to compute \( H^m \), since \( H^{m^2} \) is ambiguous for some \( r < i \). We do not compute \( m \) explicitly, but we retrace the above recursion on \( H \) in the following way as soon as \( i > 1 \) and \( t = i \).[1] Then \( H^m \) can easily be computed from the \( g(H_i) \).

**Recursion on \( H \)** (We compute \( H_{i+1} = H_i^{T+1} \) and \( t = i \mod 2^{i} \) for \( i > 1 \))
for $i = 1, 2, \ldots, k$ do

$$(H_{i+1}, r_{i+1}) = \begin{cases} (H_i H_i, 3r_i - s 2^i) & \text{if } g(H_i) < q/2, \\ (H_i H_i, r_i + a_{g(H_i)} - s 2^i) & \text{with } s = \left(\frac{r_i + a_{g(H_i)}}{2^i}\right) \text{ if } g(H_i) > q/2 \end{cases}$$

It can easily be verified that $H_i = H_i^{\lceil r_i/2 \rceil}$. Hence $H_k H_k^{-1} = H_k^{\lceil r_i/2 \rceil/2} = H_k^m$ with $m$ odd. This yields

**Fact 6** Let $\text{ord}(H) = 2^i \mod 2^{i+1}$, $e < 32$, and $H_k H_k^{-1} = 1$, then $\text{ord}(H H_k^{-1}) = 2^e.$

Therefore $S = (H_k H_k^{1-1})^{2^{e-1}}$ is an ambiguous class with $S \neq 1$ whenever $e \neq 0$.

**Comment** We have tested Method 2 in class groups and in cyclic groups $\langle H \rangle = \mathbb{Z}/n\mathbb{Z}$, in particular with $n$ prime. For random functions $g : \langle H \rangle \rightarrow \{1, \ldots, q\}$ we obtained average values of about $\sqrt{2}/\sqrt{n}$ for the smallest index $k$ such that there exists some $j < k$ with $H_j = H_k$. See Table 6, Appendix II. On the average, Method 2 takes 1.5 group operations (i.e., compositions in the case of the class group) per recursion step. A recursion step takes 2 compositions if $H_{i+1} = H_i^3$ and 1 composition if $H_{i+1} = H_i P_{g(H_i)}$. By reducing the frequency of the $H_{i+1} = H_i^3$ steps, the average number of compositions per recursion step can still be reduced.

Method 2 also works well with nonrandom $a$, like $a_i = c^i$, $i = 1, \ldots, q$, with $c$ fixed. Because of the noncommutativity of the recursion steps, Method 2 works with a smaller number $q$ of multipliers $F_i = H^a$ than Method 1. We successfully applied Method 2 with $q = 4$

The following pseudo-random function $g : \langle H \rangle \rightarrow \{1, \ldots, q\}$ works well for both methods (let $(a, b, c)$ be the reduced form in $H$)

$$g(H) = \left\lfloor (b^2 \mod p) q/p \right\rfloor + 1$$

with $p$ a prime, $q < p < \sqrt{2}$, see Tables 5, 6, Appendix II for $p = 2^{13} - 1$.

The search for $H_k = H$, with $j < k$. Let $H_0 = H, H_{i+1} = f(H_i)$. We follow an idea of Brent [1] and do not store all the $H$ but only a fixed number of them. When computing $H_i$, the stored classes $H_{a^i(1)}$ for $i = 1, 2, \ldots, 7$. For sufficiently large $i$, will be such that $a^i(1) < i < 1^3 a^1(1) = 2 \cdot 14 a^1(1)$.

The recursion for $H_i$ is continued until some $H_k = H_{a^i(1)}$ has been found. The corresponding program looks like

**Search** $H_1 = H, a(1) = 1$ for $i = 1, \ldots, 7,$ for $i = 2,$ do

- compute $H_2$ from $H_1$
- if $\exists \nu \ H_{a^i(1)} = H$, then $[j = a(\nu) \ k = i \ \text{stop}]$
- if $1^3 a^1(1) < i < 1^3 a^1(1) = 2 \cdot 14 a^1(1)$

store $H_i$ instead of $H_{a^i(1)}$

$$[\begin{array}{c} \sigma(\nu) = (a(\nu + 1)) \text{ for } \nu \neq 7 \\ \text{for } \nu = 7 \end{array}]$$

C P SCHNORR AND H W LENSTRA JR

$H_1 = 1$ (the unit class), $r_1 = 1.$
Let $\lambda$ be the period and $\mu$ the length of the nonperiodic segment of the sequence $H_i$.

e.g.

$$H_n = H_n + \lambda, \quad H_i \neq H_i \quad \text{for } i < i < \mu + \lambda$$

Fact 7 The procedure search finds some $j < k$ with $H_k = H_j$ within $\leq 11m + \lambda$ recursion steps, $m = \max(\lambda, \mu)$.

Proof Since $\sigma$ increases by the factor 11, $\sigma$ will take some value $\sigma(\nu)$ with $m \leq \sigma(\nu) \leq 11m$. Hence the for-loop stops, at the latest, with

$$k = \sigma(\nu) + \lambda \leq 11m + \lambda, \quad j = \sigma(\nu),$$

and finds the equality $H_k = H_j$. $\square$

Under the assumption that each of the $\text{ord}(H)^{\text{ord}(H)}$ functions $f(H) \rightarrow <H>$ has probability $\text{ord}(H)^{-\text{ord}(H)}$, the stochastic behavior of $\mu, \lambda$ has been well analyzed (see Knuth [7, Exercise 3 112]).

The expected values of $\mu$ and $\lambda$ are

$$E(\mu) = 1 \pm \sqrt{\frac{\sigma(\nu)}{8}} + 1/3,$$

$$E(\mu + \lambda) = 1.25 \frac{\sigma(\nu)}{\text{ord}(H)} - 1/3,$$

$$\text{Prob} \left[ \mu + \lambda \leq \sqrt{\frac{\sigma(\nu)}{2 \text{ord}(H)}} \right] = e^{-\nu/4} = 0.46$$

We conclude from Fact 7 that the number of recursion steps in search will be about

$$11 \left( E(\mu) + E(\lambda) \right) = 1.32 \frac{\sigma(\nu)}{\text{ord}(H)}$$

provided that $f(H) \rightarrow <H>$ is sufficiently random.

If in Stage 2 we compute the $H_i$ for $i \leq 132 p$, then most likely some relation $H_j = H_k, j < k$, will be found, provided $\text{ord}(H) \leq p_i^2$. It remains to analyze the chance that $\text{ord}(H) \leq p_i^2$. For each prime $p, p_i < p < p_i$, we assume that the frequency of class numbers $h(-m), m \leq n$, which are divisible by $p$, is $\geq p^{-1}$, and we assume that $h(-m)/p$ factors like random integers of size $\sqrt{n}/p$. By retracing the proof of Corollary 4, we conclude from the assumptions (21) (22)

$$\text{For all } r, n, i \text{ with } n \equiv n_0, p_i = n^{1/2r}, r \leq \sqrt{\ln n}/\ln n$$

and for all primes $p = n^{1/2r} < p_i^2$

\begin{equation}
\begin{aligned}
&\# \left\{ m \leq n \mid h(-m) \mid \prod_{p \mid \sigma} p^{\sigma(p)} \right\} / (0 \sigma n) \\
&\geq 0.33 p^{r-1}(r - r/s)^{(r-s)/s}
\end{aligned}
\end{equation}

Summing over all $p, n^{1/(1+\epsilon)} < p < n^{1/r}$, this yields

$$\begin{aligned}
&\# \left\{ m \leq n \mid h(-m) \mid \prod_{p \mid \sigma} p^{\sigma(p)} \text{ with } p < p_i^2 \right\} / (0 \sigma n) \\
&\geq 0.83 \sum_{nm} \left( \frac{p}{n} \right)^{\sigma(r - 2/(1 + \epsilon))} \left( \frac{\text{ord}(H)}{n} \right)^{(1/s) \text{ ord}(H)}
\end{aligned}
\end{equation}

(Using Theorem 4.27 in Hardy and Wright [6] there follows)

$$\geq 0.83 \ln(1 + \epsilon)(r - 2/(1 + \epsilon)) \left( \frac{\text{ord}(H)}{n} \right)^{(1/s) \text{ ord}(H)}$$
Conclusion. Assume (21), (22) and that for every discriminant \( m \leq n \), \( H \), in \( G(-m) \) is chosen at random. Then \( \forall e > 0 \): \( \exists s > 0 \) \( \forall u > n, \) and all \( p_i = n^{1/2r}, r = \sqrt{\ln n}/\ln n \). Stages 1 and 2 with \( O(p_i) \) compositions, factor at least \( \epsilon_n (r - 2 + \ell) 'n \) discriminants \( \leq n \).

If one assumes that very large class numbers \( h(-n) \) factor like even integers of size \( n \), then we can compare the efficiency of Stages 1 and 2 by Odlyzko's Table 4. Appendix II. The table indicates that for class numbers \( h(-m) \equiv 10^4 \), \( l = 15, 20, 25, 30 \) the success frequency of Stages 1 and 2 is at least \( r^2 \) and is at most \( \sigma_2 \). The success frequency of Stage 1. Note that \( (r - 2)/r \) approaches \( \sigma_2 \) for large \( r \).

Remark. There is a well-known deterministic method for doing Stage 2 within \( \sqrt{2} \) \( p_i \) compositions and with \( O(p_i) \) storage. The method is explained in Shanks [17, p. 419] in terms of "baby" and "giant" steps. In our situation we can even speed this method by a constant factor if we exploit the fact that \( \text{ord}(H) \) will most likely have no prime divisor \( \equiv \). 4. The Main Algorithm. The new algorithm can be used for factoring any composite integer \( n \). We apply Stage 1 to multiples \( m \) of \( n \) such that \( -m \) is a discriminant. Here we exploit the observation that class numbers \( h(-m) \) of fundamental discriminants \( -m \) are uncorrelated for distinct values of \( s \). The nonfundamental discriminants \( -m \) should be discarded as far as possible. The discriminant \( \Delta \) is fundamental if

\(-\exists w \in \mathbb{N} \ w \neq 1 \ \Delta/w^2 \) is a discriminant

In fact, the class number formula (see Dirichlet [8]),

\[ h(-m) = \frac{\sqrt{m}}{\pi} \prod_{\text{prime } \rho} \left( 1 - \frac{1}{\rho} \left( \frac{-m}{\rho} \right) \right)^{\frac{1}{2}} \text{ for } m < 4 \]

implies for \( \gcd(w, m) = 1 \) and \( w \) square free

\[ h(-mw^2) = \frac{h(-m)}{w} \prod_{\text{prime } \rho} \left( 1 - \frac{1}{\rho} \left( \frac{-m}{\rho} \right) \right) \prod_{\text{prime } \rho} \left( \rho - \left( \frac{-m}{\rho} \right) \right) \]

Hence for small \( w \), \( h(-m) \) and \( h(-mw^2) \) have the same large prime divisors provided \( \gcd(m, w) = 1 \).

Main Algorithm. Let \( n \) be the number to be factored and \( p_1 = 2, p_2 = 3 \). \( p_i \) the first \( r \) primes \( p_i = n^{1/2r} \) (the appropriate choice of \( r \) will be determined in the subsequent analysis)

1. \( s = 0 \)
2. take the next \( s \) with \( \gcd(n, s) = 1 \), \( m = 0 \mod 4 \) and \( \{ \exists w \in \mathbb{N} \ w^2 \} \)
3. run Stage 1 on \( m, \) which takes \( O(p_i) \) compositions. If Stage 1 yields an ambiguous class \( S \) then go to 4, otherwise return to 2 and take the next \( s \)
4. if \( S \) yields a factorization of \( n \) then stop; otherwise go to 5
5. return to 3 and repeat Stage 1 on \( m, \) with independently chosen classes \( H_m \leq G(-n) \) until some factorization of \( n \) has been found. In order to prevent that merely useless ambiguous classes are generated, continue to build up the 2 Sylow groups \( S_i \) of \( G(-m) \). Use Stage 1 to generate classes \( m \)
A MONTE CARLO FACTORIZATION ALGORITHM WITH LINEAR STORAGE

$S_{2}(-ns)$ Apply the recursion step of the method below whenever a new $A \in S_{2}(-ns)$ has been found

The 2-Sylow group $S_{2}(-ns)$ is a direct product of cyclic groups of order $2^n$, $m > 0$, $S_{2}(-ns) \cong \oplus_{i=1}^{\lambda} \mathbb{Z}/2^{m_i} \mathbb{Z}$. Let $\lambda$ be the number of cyclic components, then $S_{2}(-ns)$ has $2^{\lambda}$ ambiguous classes. Let $d$ be the number of odd prime factors of $ns$. Then by Theorem III, Appendix I, $\lambda$ is $d - 1$, $d$ or $d + 1$ depending on the maximal power of 2 which divides $ns$. The ambiguous classes that do not yield a factorization of $n$ form a subgroup $S_{2}(-ns)$ of $S_{2}(-ns)$. Let $d_{a}, d_{s}$ be the numbers of distinct odd prime factors of $n$ and $s$. Since $\gcd(n, s) = 1$, we have $d = d_{a} + d_{s}$. It follows immediately from Theorem III, Appendix I, that the number $\lambda$ of cyclic components of $S_{2}(-ns)$ is $\lambda = \lambda_{a} - d_{a} + d_{s} + 1$.

Constructing $S_{2}(-ns)$ till a factorization of $n$ is found. Given a procedure that generates elements of $S_{2}(-ns)$ (this will be done by Stage I) we recursively construct subsets $\{A_{1}, \ldots, A_{\lambda}\} \subset S_{2}(-ns)$, $\mathcal{X} \leq \lambda$ such that $\lambda$.

(41) $|\langle A_{1}, A_{2}\rangle| = \prod_{i=1}^{\lambda} \text{ord}(A_{i})$

Let $\text{ord}(A_{i}) = 2^{m_{i}}$, $B_{i} = A_{i}^{m_{i}}$, $\langle A_{1}, A_{2}\rangle = S_{2}(-ns)$. Then $S_{2}(-ns) \cong \oplus_{i=1}^{\lambda} \mathbb{Z}/2^{m_{i}} \mathbb{Z}$ and $B_{1}, \ldots, B_{\lambda}$ generate the subgroups of ambiguous classes of $S_{2}(-ns)$. After each recursion step either a factorization of $n$ has been found or the new group $S_{2}(-ns)$ will be the subgroup of $S_{2}(-ns)$ which is generated by the previous $S_{2}(-ns)$ and the element $A \in S_{2}(-ns)$ obtained in step I.

(42) Algorithm for $S_{2}(-ns)$

0 $\mathcal{X} = 0$ $A_{1} = 1$ ($= \text{the unit class}$)
1 generate another $A \in S_{2}(-ns)$ $A \neq 1$
2 compute $A, A^{2}, \ldots, A^{|\mathcal{X}|} = 1$ and put $B = A^{2}$
3 if $B$ yields a factorization of $n$ then stop
4 test whether $B \subseteq \langle B_{1}, B_{2}\rangle$, if $B \not\subseteq \langle B_{1}, B_{2}\rangle$ then go to 5 else compute
5 $A_{\lambda+1} = A$ $B_{\lambda+1} = B$ $\mathcal{X} = \lambda + 1$ return $\lambda$
6 if $\exists j \in J$ $m_{j} < m$ then select $j \in J$ with $m_{j}$ minimal and interchange $A$ with $A_{j}$
7 (we have $A^{2^{m-1}} = \prod_{j \in J} A_{j}^{2^{m_{j}} - 1}$, $m \leq m_{j}$ for $j \in J$) Put $A = A \prod_{j \in J} A_{j}^{2^{m_{j}} - 1}$, if $A = 1$ go to 1 else go to 2 (the new $m$ to be computed in 2 will be smaller than the present $m$ since $A^{2^{m-1}} = 1$ holds for the new $A$)

Run Time Analysis of the Main Algorithm. We separately bound

1 the number $T(n)$ of bit operations to be done till some $i$ has been reached with $h(-ns) \prod_{i=1}^{m} p_{i}^{e_{i}}$ $e_{i} = \max \{ p_{i}, p_{j} \leq p_{i}^{e_{i}} \}$

2. The number $\overline{T}(n)$ of bit operations for building up the 2-Sylow group $S_{2}(-ns)$ of $G(-ns)$ till a factorization of $n$ is found.
We will assume that Corollary 4 extends to multiples of \( n \), with \( r > S \) and \( \sum \frac{1}{p_i} < 1 \). Our experimental data in fact confirm the lower bound \( r^+ \). The assumption (4.3) implies
\[
\forall n \geq n_0 \quad \forall p_i = \left( n^{3r/c} \right)^{1/2r} \quad r \leq \frac{\ln n}{\ln \ln n}
\]
Since Stage 1 takes \( O(p_i) \) compositions, we have
\[
T(n) = O\left( p_i r^+(\ln n)^2 \right) = O\left( n^{1/2r} r^+ (\ln n)^2 \right)
\]
Here \( O(\ln n)^2 \) takes into account the costs for the arithmetic. We choose \( r = \sqrt{\ln n/\ln \ln n} \), \( p_i = \left( n^{3r/c} \right)^{1/2r} = O(n^{1/2r} \sqrt{r}) \). Then all together (4.3) implies
\[
T(n) = o\left( \exp \frac{\ln n}{\ln \ln n} \right)
\]
2. \( T(n) \) In order to factor \( n \) we need only to find at most \( d_i + 2 \) cyclic components of \( S_j(-ns) \). If the passes through Stage 1 generate independent elements of \( S_j(-ns) \) then \( k \) passes of Stage 1 with probability \( \geq 1 - 2^{-k} \) detect a new cyclic component of \( S_j(-ns) \). Hence almost surely we need at most \( O(d_i) \) passes through Stage 1, and each pass takes \( O(p_i) \) compositions. The number of steps for updating the information on \( S_j(-ns) \) can be bounded as \( O(s) \), the most costly operation in Algorithm 4.2 is to check whether \( B \in \{ B_1, B_2 \} \) (step 4). Since \( \lambda \leq \lambda' \leq d_i + 1 \) this can be done in a crude way by comparing \( B \) with each of the \( 2^k \leq 2^{d_i + 1} = O(s) \) elements of \( \{ B_1, B_2 \} \). This takes \( O(s) \) steps and is sufficient for our purposes. We obtain
\[
T(n) = O\left( d_i (p_i + s) (\ln n)^2 \right) = O\left( \log s (n^{1/2r} \sqrt{r} + s) (\ln n)^2 \right)
\]
with \( s \leq r^+ \), \( r \leq \sqrt{\ln n/\ln \ln n} \). Here again \( O(\ln n)^2 \) bounds the cost for the arithmetic. It follows immediately that \( T(n) = o(T(n)) \).

Conclusion If (4.3) holds, then the Main Algorithm, using only Stage 1 takes \( o(\exp \frac{\ln n}{\ln \ln n}) \) bit operations to factor arbitrary, composite integers \( n \).

If we also apply Stage 2, then \( s \) will be bounded as \( O((r - 2)^{r^+}) \), and this will save a time factor of about \( r^2 = \ln n/\ln \ln n \).

5. Some Computational Experience The new factoring method has been programmed in Fortran on a DEC-1091 at Frankfurt University. The core of the algorithm is a subroutine for composition of quadratic forms written in machine language, and based on the improved composition method proposed by Seysen [20]. The arithmetic operations and the gcd-calculations have been programmed for two-word integers, i.e. for integers \( \leq 2^{30} \). This means that the program can factor integers \( n \leq 2^{310} \approx 10^{36} \) using multipliers \( \approx 2^{20} \). Stage 1 uses the exponents \( c_i = \max\{ r, p_i \leq p_i \} \). Hence the number of compositions per multiplier for Stage 1 is about \( 2^2 p_i \). Both methods for Stage 2 have been tested and they are comparable in efficiency.
The integers $n$ which have been factored for Table 0 are products of two distinct primes $p_1$, $p_2$ of nearly the same size. It turns out that the median of the factoring time is considerably smaller than the average factoring time. This is due to a small fraction of integers $n$ which take extremely many multipliers. On the other hand, there is a considerable fraction of integers which only take very few multipliers. For instance, the seventh Fermat number $F_7 = 2^{2^7} + 1 = 163$ only took 7 multipliers and was factored in about 7 minutes. Here we used $p_1 = 163$, but we ran Stage 2 for only 7500 compositions, hence each multiplier took about 1 minute. The multiplier 15 has been successful.

We observed that the factoring method is somewhat faster for integers with more than 2 prime divisors. By our observation class numbers of discriminants with many prime divisors tend to have fewer large prime divisors compared with class numbers of discriminants which are prime or products of two primes. For instance, for $n = 10^{10}$ a product of 5 primes $p = 10^5$, the algorithm on the average only took 8.7 multipliers. The median of the number of multipliers has been 5 compared with 14.4 and 8 in Table 0. We have factored a sample of 200 of these integers $n$.

Appendix I on Quadratic Forms. We report classical theorems and algorithms on quadratic forms see Gauss [5], Mathews [10]. A quadratic form $ax^2 + bx + c$ with $a, b, c \in \mathbb{Z}$ is denoted as $(a, b, c)$. Its discriminant is $\Delta = b^2 - 4ac$. $(a, b, c)$ is primitive if $a > 0$, primitive if $gcd(a, b, c) = 1$. Two forms $(a, b, c)$, $(a', b', c')$ are equivalent if there exists a linear transformation with integer coefficients and determinant 1 transforming the one form into the other, i.e.,

$$I \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} I = \begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix} $$
for some integer matrix $T$ with $\det T = 1$. Let $[(a, b, c)]$ be the class represented by $(a, b, c)$. For negative discriminants we always restrict to positive forms.

Two classes $[(a, b, c)], [(\tilde{a}, \tilde{b}, c)]$ yield a new class $[(A, B, C)]$ by composition as follows (for explanation see Lenstra [9]):

$$d = \gcd(a, \tilde{a}, (b + \tilde{b})/2)$$

Let $\alpha, \beta, \gamma \in \mathbb{Z}$ be such that $\alpha a + \beta \tilde{a} + \gamma (b + \tilde{b}) - d$

$$A = a\alpha /d^2$$

$$B = \tilde{b} + 2\tilde{a} \left( \frac{\beta - \tilde{b}}{2} - \gamma c \right) \mod \frac{d}{a}$$

$$C = (\Delta + B^2)/(4A)$$

$[(A, B, C)]$ does not depend on the particular choice for $\alpha, \beta, \gamma, B$, and $C$ $(A, B, C)$ will be primitive, if $(a, b, c)$ and $(\tilde{a}, \tilde{b}, c)$ are primitive.

THEOREM I. The equivalence classes of primitive quadratic forms with discriminant $\Delta$ form an Abelian group $G(\Delta)$ under composition. Its order $h(\Delta)$ is the class number.

The unit class $1$ in $G(\Delta)$ is represented by the form

$$(1, 0, -\Delta/4) \quad \text{if } \Delta \equiv 0 \mod 4,$$

$$(1, 1, (1 - \Delta)/4) \quad \text{if } \Delta \equiv 1 \mod 4$$

The inverse of $[(a, b, c)]$ is $[(a, -b, c)]$.

A class $H \in G(\Delta)$ is ambiguous if $H^2 = 1$. The following assertions are equivalent:

1. $H$ is ambiguous.
2. every form $(a, b, c)$ in $H$ is equivalent to $(a, -b, c)$.
3. $T \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix} T^{-1} = \begin{pmatrix} a & b/2 \\ b/2 & c \end{pmatrix}$

for some integer matrix $T$ with $\det T = -1$.

4. there is a form $(a, b, c)$ in $H$ with $a|b$.

For negative discriminants $\Delta$, classes in $G(\Delta)$ correspond to reduced forms $(a, b, c)$ is reduced if

1. $|b| \leq a \leq c$ and
2. $b \neq 0$ if $|b| = a$ or $a = c$ (i.e. if $[(a, b, c)]$ is ambiguous).

THEOREM II (GAUSS [5, ART 172]). In each equivalence class with negative discriminant there is exactly one reduced form.

Gauss also gave a gcd-like reduction algorithm which transforms a given form $(a, b, c)$ into an equivalent reduced form.

(5.2) reduction process for $(a, b, c)$

1. find $v \in \mathbb{Z}$, $-|a| < b + 2va \leq |a|$
2. $b = b + 2va = (b^2 - a^2)/(4a)$
3. if $(a, b, c)$ is not reduced then replace $(a, b, c)$ by $(c, -b, a)$ and go to 1.

The reduced forms of ambiguous classes with $\Delta < 0$ are of either of the following types:

1. $b = 0$
2. $a = b$
3. $a = c$
These forms are called ambiguous. Every ambiguous form corresponds to a factorization of the discriminant as

\( \Delta = -4ac, \quad \Delta = b(b - 4c), \quad \Delta = (b - 2a)(b + 2a) \)

In order to describe this correspondence more precisely, let

\[ F(n) = \{(n_1, n_2) \in \mathbb{N}^2 \mid n = n_1n_2, \gcd(n_1, n_2) = 1\} \]

be the set of relatively prime, ordered factor pairs of \( n \in \mathbb{N} \). We have \( \#F(n) = 2^d \) where \( d \) is the number of distinct prime divisors of \( n \).

Let \( \Delta \equiv 2^e \mod 2^{e+1}, \Delta < 0 \), then the set \( A(\Delta) \) of ambiguous forms with discriminant \( \Delta \) is either in 1-1 or in 1-2-correspondence with \( F(-\Delta/2^e) \).

**Theorem III** Let the discriminant \( \Delta < 0 \) have \( d \) odd prime divisors. Then the set \( A(\Delta) \) of ambiguous (reduced, positive primitive) forms with discriminant \( \Delta = -2^en \) odd, is obtained from the \((n_1, n_2) \in F(n) \) with \( n_1 \leq n_2 \) as follows:

\[
\begin{align*}
\text{case } e &= 0 \ (e \Delta \equiv 1 \mod 4) \quad \#A(\Delta) = 1 \quad \#F(n) = 2^{d - 1} \\
(n_1, n_2) / &\quad \text{if } 3n_1 < n_2 \\
(2n_1, 2n_1, n_1 + n_2) / &\quad \text{if } 3n_1 < n_2 \\
(n_1 + n_2) / &\quad \text{if } 3n_1 < n_2 \\
\text{case } e &= 2, n \equiv -1 \mod 4 \quad \#A(\Delta) = 1 \quad \#F(n) = 2^{d - 1} \\
(n_1, 0, n_2) &
\end{align*}
\]

\[
\begin{align*}
\text{case } e &= 3, 4 \quad \#A(\Delta) = \#F(n) = 2^d \\
(\min(n_12^{e-2}, n_2), 0, \max(n_12^{e-2}, n_2)) / &\quad \text{if } 3n_1 < n_2 \\
(\min(n_22^{e-2}, n_1), 0, \max(n_22^{e-2}, n_1)) &
\end{align*}
\]

We have listed pairwise inequivalent forms corresponding to distinct positive ambiguous classes. They have been arranged according to their types (i), (ii), (iii) as introduced above.

Theorem III can easily be obtained from Gauss [5, Art 257–259]. Observe that our classes with discriminant \( \Delta \equiv 0 \mod 4 \ (\Delta \equiv 1 \mod 4 \text{ resp.}) \) correspond to primitive Gauss classes with determinant \( D = \Delta/4 \) (improper primitive Gauss classes with determinant \( D = \Delta \), resp.) The number of ambiguous classes has also been listed in Caves [2, p. 342].

The Efficiency of Composition An efficient composition algorithm is the main requirement for a satisfactory implementation of our factoring algorithm. All
calculations in $G(\Delta)$ are done with reduced forms. Composition consists of two parts:

1. Evaluation of (5.1) \((a, b, c), (\tilde{a}, \tilde{b}, c) \rightarrow (A, B, C)\) (this amounts to an extended gcd-calculation on integers of size $O(\sqrt{|\Delta|})$).
2. Reduction of \((A, B, C)\).

If the reduction is done as in (5.2), this corresponds to an extended gcd-calculation on integers of size $O(\sqrt{|\Delta|})$. However, M. Seysen [20] found a faster reduction algorithm for this particular situation. Reducing \((A, B, C)\) by this algorithm corresponds to only half an extended gcd-calculation on integers of size $O(\sqrt{|\Delta|})$.

Appendix II Statistical Tables

Table 1 shows the distribution of class numbers $h(-m)$ without large prime divisors for discriminants $-m$ in the interval $I = [-472,650,003, -472,600,000]$. There are 25,002 discriminants, the minimal, maximal, and average class numbers are 1518, 47,452, and 9,469,77. We put

$$\tilde{c}(m) = \max \left\{ v_p^m h(-m) \right\}$$

For every prime $p_i = 2, 3, \ldots, 89$ we record the percentage of those discriminants $-m \in I$ satisfying the following conditions:

<table>
<thead>
<tr>
<th>Column</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$h(-m)$ is free of primes $&gt; p_i$, if $h(-m) = \prod_{i=1}^{r} p_i^{m_i}$</td>
</tr>
<tr>
<td>2</td>
<td>for all $i \neq 2$, $\tilde{c}(m) \leq c_i = \max \left{ v_p p_i^m \right}$</td>
</tr>
<tr>
<td>3</td>
<td>for all $i \neq 2$, $\tilde{c}(m) \leq c_i' = \max \left{ v_p p_i^m \right}$</td>
</tr>
<tr>
<td>4</td>
<td>$h(-m) \prod_{i=1}^{r} p_i^{m_i} q$ for some $q &lt; p_i^m$, $q$ prime</td>
</tr>
<tr>
<td>5</td>
<td>$h(-m) \prod_{i=1}^{r} p_i^{m_i} q^{m_q}$ for some $q &lt; p_i^m$, $q$ prime</td>
</tr>
<tr>
<td>6</td>
<td>$h(-m) \prod_{i=1}^{r} p_i^{m_i} q$ for some $q &lt; p_i^m$, $q$ prime</td>
</tr>
<tr>
<td>7</td>
<td>$i = \ln n/(2 \ln p_i)$</td>
</tr>
<tr>
<td>8</td>
<td>$10^3 \cdot r$</td>
</tr>
</tbody>
</table>

Moreover, we note in

Observe that the entries in columns 1-3 of Table 1 are always greater than $10^2 r$, which confirms Corollary 4. For $r \leq \sqrt{n} / \ln n \ln n$ ($1 < r < 55, p_i = 53$) the entries in column 3 are only slightly smaller than those in columns 1, 2. This suggests that Stage 1 should be done with the smaller exponents $c_i$ instead of the $c_i'$. Table 2 has the same meaning as Table 1 but is restricted to fundamental discriminants in the same interval $I$. Minimal, maximal, and average class numbers are 1518, 47,425, and 10,073,9. There are 15,195 fundamental discriminants in $I$. 
### Table 1

<table>
<thead>
<tr>
<th>$p_i$</th>
<th>col 1</th>
<th>col 2</th>
<th>col 3</th>
<th>col 4</th>
<th>col 5</th>
<th>col 6</th>
<th>col 7</th>
<th>col 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>3</td>
<td>1.26</td>
<td>0.51</td>
<td>0.51</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>5</td>
<td>3.69</td>
<td>1.65</td>
<td>1.65</td>
<td>12.03</td>
<td>12.03</td>
<td>12.03</td>
<td>12.03</td>
<td>12.03</td>
</tr>
<tr>
<td>7</td>
<td>6.96</td>
<td>4.98</td>
<td>4.98</td>
<td>35.44</td>
<td>35.44</td>
<td>35.44</td>
<td>35.44</td>
<td>35.44</td>
</tr>
<tr>
<td>11</td>
<td>10.22</td>
<td>8.68</td>
<td>8.68</td>
<td>41.55</td>
<td>41.55</td>
<td>41.55</td>
<td>41.55</td>
<td>41.55</td>
</tr>
<tr>
<td>13</td>
<td>14.02</td>
<td>13.04</td>
<td>13.04</td>
<td>50.71</td>
<td>50.71</td>
<td>50.71</td>
<td>50.71</td>
<td>50.71</td>
</tr>
<tr>
<td>17</td>
<td>17.09</td>
<td>16.46</td>
<td>16.46</td>
<td>61.77</td>
<td>61.77</td>
<td>61.77</td>
<td>61.77</td>
<td>61.77</td>
</tr>
<tr>
<td>23</td>
<td>23.09</td>
<td>22.64</td>
<td>22.64</td>
<td>74.04</td>
<td>74.04</td>
<td>74.04</td>
<td>74.04</td>
<td>74.04</td>
</tr>
<tr>
<td>29</td>
<td>29.73</td>
<td>25.55</td>
<td>25.55</td>
<td>92.61</td>
<td>92.61</td>
<td>92.61</td>
<td>92.61</td>
<td>92.61</td>
</tr>
<tr>
<td>31</td>
<td>31.28</td>
<td>28.03</td>
<td>28.03</td>
<td>93.22</td>
<td>93.22</td>
<td>93.22</td>
<td>93.22</td>
<td>93.22</td>
</tr>
<tr>
<td>37</td>
<td>37.42</td>
<td>30.29</td>
<td>30.29</td>
<td>97.57</td>
<td>97.57</td>
<td>97.57</td>
<td>97.57</td>
<td>97.57</td>
</tr>
<tr>
<td>41</td>
<td>41.66</td>
<td>32.53</td>
<td>32.53</td>
<td>98.78</td>
<td>98.78</td>
<td>98.78</td>
<td>98.78</td>
<td>98.78</td>
</tr>
<tr>
<td>43</td>
<td>43.68</td>
<td>34.56</td>
<td>34.56</td>
<td>99.02</td>
<td>99.02</td>
<td>99.02</td>
<td>99.02</td>
<td>99.02</td>
</tr>
<tr>
<td>47</td>
<td>47.36</td>
<td>36.54</td>
<td>36.54</td>
<td>99.44</td>
<td>99.44</td>
<td>99.44</td>
<td>99.44</td>
<td>99.44</td>
</tr>
</tbody>
</table>

### Table 2

<table>
<thead>
<tr>
<th>$p_i$</th>
<th>col 1</th>
<th>col 2</th>
<th>col 3</th>
<th>col 4</th>
<th>col 5</th>
<th>col 6</th>
<th>col 7</th>
<th>col 8</th>
</tr>
</thead>
<tbody>
<tr>
<td>2</td>
<td>0.17</td>
<td>0.17</td>
<td>0.17</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
<td>0.31</td>
</tr>
<tr>
<td>3</td>
<td>1.26</td>
<td>0.51</td>
<td>0.51</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
<td>1.69</td>
</tr>
<tr>
<td>5</td>
<td>3.69</td>
<td>1.65</td>
<td>1.65</td>
<td>12.03</td>
<td>12.03</td>
<td>12.03</td>
<td>12.03</td>
<td>12.03</td>
</tr>
<tr>
<td>7</td>
<td>6.96</td>
<td>4.98</td>
<td>4.98</td>
<td>35.44</td>
<td>35.44</td>
<td>35.44</td>
<td>35.44</td>
<td>35.44</td>
</tr>
<tr>
<td>11</td>
<td>10.22</td>
<td>8.68</td>
<td>8.68</td>
<td>41.55</td>
<td>41.55</td>
<td>41.55</td>
<td>41.55</td>
<td>41.55</td>
</tr>
<tr>
<td>13</td>
<td>14.02</td>
<td>13.04</td>
<td>13.04</td>
<td>50.71</td>
<td>50.71</td>
<td>50.71</td>
<td>50.71</td>
<td>50.71</td>
</tr>
<tr>
<td>17</td>
<td>17.09</td>
<td>16.46</td>
<td>16.46</td>
<td>61.77</td>
<td>61.77</td>
<td>61.77</td>
<td>61.77</td>
<td>61.77</td>
</tr>
<tr>
<td>23</td>
<td>23.09</td>
<td>22.64</td>
<td>22.64</td>
<td>74.04</td>
<td>74.04</td>
<td>74.04</td>
<td>74.04</td>
<td>74.04</td>
</tr>
<tr>
<td>29</td>
<td>29.73</td>
<td>25.55</td>
<td>25.55</td>
<td>92.61</td>
<td>92.61</td>
<td>92.61</td>
<td>92.61</td>
<td>92.61</td>
</tr>
<tr>
<td>31</td>
<td>31.28</td>
<td>28.03</td>
<td>28.03</td>
<td>93.22</td>
<td>93.22</td>
<td>93.22</td>
<td>93.22</td>
<td>93.22</td>
</tr>
<tr>
<td>37</td>
<td>37.42</td>
<td>30.29</td>
<td>30.29</td>
<td>97.57</td>
<td>97.57</td>
<td>97.57</td>
<td>97.57</td>
<td>97.57</td>
</tr>
<tr>
<td>41</td>
<td>41.66</td>
<td>32.53</td>
<td>32.53</td>
<td>98.78</td>
<td>98.78</td>
<td>98.78</td>
<td>98.78</td>
<td>98.78</td>
</tr>
<tr>
<td>43</td>
<td>43.68</td>
<td>34.56</td>
<td>34.56</td>
<td>99.02</td>
<td>99.02</td>
<td>99.02</td>
<td>99.02</td>
<td>99.02</td>
</tr>
<tr>
<td>47</td>
<td>47.36</td>
<td>36.54</td>
<td>36.54</td>
<td>99.44</td>
<td>99.44</td>
<td>99.44</td>
<td>99.44</td>
<td>99.44</td>
</tr>
</tbody>
</table>
### Table 3

<table>
<thead>
<tr>
<th>$v$</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
<th>9</th>
<th>10</th>
</tr>
</thead>
<tbody>
<tr>
<td>all dissct</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>fund dissct</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Table 3 shows the percentages of discriminants (fundamental discriminants, resp.) $-m \in I$ such that $v$ divides $h(-m)$ for $v = 2, \ldots, 100$. These percentages are always greater than 100/$p$, which confirms hypothesis (2.1). For small primes $p$ these frequencies are close to 100/(p - 1).

Table 4 is due to A. Odlyzko. The entry $a_{k,v}$ in the line starting with $k \leq 1$ and column headed with $v (v = 2k, 3, 5, 7, 11, 13, 17, 19)$ is the number of integers $m$ from among the first 100,000 even integers $> 10^5$ which have the property that

$$m/\gcd(m, \text{lcm}(1, 2^k)) \leq 2^{\frac{k}{2}}.$$  

The last two columns record $r \sim \ln 10^5/\ln 2^k$ and $10^5 r$. 

---

**Table 3**

- **Table 3** shows the percentages of discriminants (fundamental discriminants, resp.) $-m \in I$ such that $v$ divides $h(-m)$ for $v = 2, \ldots, 100$. These percentages are always greater than 100/$p$, which confirms hypothesis (2.1). For small primes $p$ these frequencies are close to 100/(p - 1).

**Table 4** is due to A. Odlyzko. The entry $a_{k,v}$ in the line starting with $k \leq 1$ and column headed with $v (v = 2k, 3, 5, 7, 11, 13, 17, 19)$ is the number of integers $m$ from among the first 100,000 even integers $> 10^5$ which have the property that

$$m/\gcd(m, \text{lcm}(1, 2^k)) \leq 2^{\frac{k}{2}}.$$  

The last two columns record $r \sim \ln 10^5/\ln 2^k$ and $10^5 r$. 

Table 4 also confirms our assumption (2.2) that \( a_{x+k} \) is the number of integers \( m \) among the first 100,000 even integers greater than 10 such that

\[
m = \prod p_i^{a_i} \quad \text{with} \quad p_i^{a_i} < 2^x
\]

(which implies \( m \prod p_i^{a_i} \) for the first prime \( p_i > 2^x \)). The table shows

\[
a_{x+k} > 10^{0.5^{r-x}} \quad \text{for} \quad r = \ln 10 / \ln 2^x \leq 1 \ln 10 / \ln 10^x
\]
This suggests an even stronger assumption than (2.2)
\[ \left\{ m \in \mathbb{N} \mid \prod_{i=1}^{r} p_i^{e_i} / n > r \right\} \]
for all \( r \leq \sqrt{n} \ln n / \ln \ln n \) and \( p_i \leq n^{1/2} / \ln n \). Here \( e_i = \max \{ r, p_i^{1/2} \leq p_i \} \).

Table 4 can be used to balance Stages 1 and 2. If we factor a discriminant \( n = 10^{22} \), then \( h(-n) \) will be about \( 10^i \). We choose \( p_j = 2^k \). Then hypothesis (2.1) suggests that there is some \( s \leq 10^s / \ln 2 \), with
\[ h(-ns) \mid \prod_{i=1}^{r} p_i q^s \quad \text{and} \quad q \leq 2^{2k} r \]
Hence Stages 1 and 2 will run on at most \( 10^s / \ln 2 \) multiples \(-ns\). Stage 1 with the exponents \( e_j \) takes about \( 2^k p_j \) compositions. If we run Stage 2 with Method 2 for \( 2^{k/2} \) recursion steps, then Stages 1 and 2 will most likely factor this particular \( n \) (see Table 6, Appendix II, for the performance of Method 2 in Stage 2). In this way Stage 2 takes about \( 1.5 \cdot 2^{k/2} \) compositions. Therefore the total number of compositions of the Main Algorithm will be bounded by
\[ b_{L,K} = \frac{10^s}{\ln 2} \cdot 2^k (2^2 + 15 \cdot 2^{k/2}) \]

Examples: \( n = 10^{34} \). Choose \( k = 12 \), \( p = 0 \). \( a_{15120} = 7388 \). We have \( \{ b_{15120} \} = 205132 \), and \( n \) will be factored in about \( 2 \cdot 10^5 = 2 \times 10^5 \) compositions.

\( n = 10^{40} \). Choose \( k = 14 \), \( p = 0 \). \( a_{20144} = 2182 \). We have \( \{ b_{20144} \} = 2778221 \), and \( n \) will be factored in about \( 2 \cdot 8 \cdot 10^5 = 1.6 \times 10^5 \) compositions.

\( n = 10^{50} \). Choose \( k = 17 \), \( p = 0 \). \( a_{81180} = 1610 \). We have \( \{ b_{81180} \} = 30122136 \), and \( n \) will be factored in about \( 2 \cdot 9 \cdot 10^5 = 1.8 \times 10^5 \) compositions.

\( n = 10^{60} \). Choose \( k = 18 \), \( p = 0 \). \( a_{40180} = 375 \). We have \( \{ b_{40180} \} = 458648746 \), and \( n \) will be factored in about \( 2 \cdot 6 \cdot 10^5 = 1.2 \times 10^5 \) compositions.

The examples show that the number of compositions while factoring \( n \) is smaller than \( \exp n / \ln n \). For instance, for \( n = 10^{60} \) we have \( 2 \cdot 10^5 = 0.0010 \) \exp n / \ln n \). The examples indicate that our algorithm will be faster on integers \( n > 10^{60} \) than the Morrison-Brillhart algorithm. Wunderlich [22] reports that the Morrison-Brillhart algorithm for \( n = 10^{60} \) takes about \( 322 n^{3/2} = 3.8 \cdot 10^8 \) divisions of \( Q, Q = O(\sqrt{n}) \), by small primes \( p \). Meanwhile the above estimations have been verified by a program running on the DEC 1091 in Frankfurt, see Section 5.

Table 5 demonstrates the performance of Method 1 of Stage 2. We choose the pseudo-random function \( g(\Delta) \rightarrow \{1, 16\} \)
\[ g(\Delta) = \left\lfloor b_{i \mod (2^{13} - 1)} / 16 \right\rfloor \]
where \( (a, b, c) \) is the reducible form corresponding to \( \Delta \). We consider the method with 6 distinct samples of exponents \( a_i \) and three samples with \( a_i \), chosen at random and three samples with regular \( a_i \). \( a_i = -c^{1/2} \mod 2^{13} \). \( j = 1 \) with \( \epsilon = 2 \cdot 3 \cdot 5 \). We have the recursion
\[ H_{i-1} = H_i H_{i+1} \]
with \( H_0 \in \mathbb{G}(\Delta) \) chosen at random. We apply this recursion to the 50 largest discriminants \( \Delta < -10^m \) with \( \Delta \equiv 1 \mod 4 \) for \( m = 8, 10, 12, 14 \).

For every sample the 10\(^{th}\)-column records four values:

1. the average period length
2. the median of the period lengths
3. the average number of recursion steps until search finds some \( k \) with \( \exists j < k \leq H_j = H_k \)
4. the median of the number of recursion steps

The particular favorable performance of \( a_5 = 2^{11} \mod 2^{10} \) and of \( a_7 = 3^{11} \mod 70 \) can be explained by the fact that the order of most class groups is even and is a multiple of 3 for about half of the class groups. Despite this favorable performance for random class groups the choice of \( a_5 = 2^{11} \mod 2^{10} \) is unfavorable for the factoring algorithm since in the particular situation of Stage 2 the class numbers are free of small prime divisors.

Table 6 shows the performance of Method 2 in Stage 2 for the pseudo-random function \( g(\Delta) \rightarrow \{1, 4\} \)

\[
g(H) = \left\lfloor \left\lfloor b^7 \mod (2^{11} - 1) \right\rfloor 4/(2^{11} - 1) \right\rfloor + 1
\]

where \((a, b, c)\) is the reduced form corresponding to \( H \). We used \( a_j = d^i, i = 3, 4 \), with constants \( d = 2, 3, \ldots, 8 \). The recursion scheme is...
\[ H_{r+1} = \begin{cases} H_r^2, & \text{if } g(H_r) \leq 2, \\ H_r H_0^2 & \text{otherwise}, \end{cases} \]

\( H_0 \in G(\Delta) \) is chosen at random.

For every \( d = 2, 3, \ldots, 8 \) and \( m = 8, 14 \) we applied this recursion to the 50 largest discriminants \( \Delta \approx -10^m \) with \( \Delta = 1 \mod 4 \). For every \( d \) and \( m \) the table records four values:

1. the average period length
2. the median of the period lengths,
3. the average number of recursion steps till search finds some \( k \) with \( \exists f < k \)
4. the median of the number of recursion steps

| Table 6 |
|---|---|---|---|
| base \( d \) | \( 10^8 \) | \( 10^{10} \) | \( 10^{12} \) | \( 10^{14} \) |
| 2 | 21 | 63 | 178 | 684 |
| 4 | 12 | 42 | 112 | 299 |
| 8 | 51 | 158 | 395 | 1513 |
| 16 | 43 | 92 | 122 | 1099 |
| 3 | 28 | 116 | 231 | 783 |
| 6 | 17 | 53 | 86 | 555 |
| 12 | 63 | 219 | 503 | 1519 |
| 24 | 50 | 157 | 260 | 1118 |
| 4 | 24 | 76 | 185 | 780 |
| 8 | 15 | 41 | 83 | 277 |
| 16 | 48 | 165 | 360 | 1260 |
| 32 | 31 | 123 | 285 | 618 |
| 5 | 14 | 109 | 273 | 746 |
| 10 | 22 | 58 | 143 | 429 |
| 20 | 74 | 224 | 570 | 1613 |
| 40 | 50 | 161 | 357 | 1013 |
| 6 | 23 | 43 | 98 | 415 |
| 12 | 12 | 24 | 79 | 196 |
| 24 | 55 | 114 | 299 | 927 |
| 48 | 35 | 72 | 188 | 618 |
| 7 | 28 | 90 | 276 | 713 |
| 14 | 24 | 74 | 154 | 376 |
| 28 | 63 | 224 | 516 | 1533 |
| 56 | 41 | 159 | 386 | 925 |
| 9 | 26 | 76 | 209 | 783 |
| 18 | 13 | 47 | 129 | 445 |
| 36 | 53 | 144 | 371 | 1449 |
| 72 | 41 | 94 | 278 | 918 |

**Acknowledgement** Several members of the Frankfurt working group participated in implementing this algorithm. H. G. Franke produced Tables 1, 3 and implemented large integer arithmetic and an efficient composition algorithm in machine language. J. Saitter produced Tables 5, 6 which compare Methods 1, 2 of Stage 2.
Thanks are due to the computer centres of Frankfurt and Amsterdam university for providing computing time on the DEC-1091 in Frankfurt and the CDC-Cyber in Amsterdam. We also like to thank A. Odlyzko for the permission to include in this paper Table 4 which is part of a larger statistic made at Bell Laboratories.

Fachbereich Mathematik
Universität Frankfurt
6 Frankfurt am Main West Germany

Mathematisch Instituut
Universiteit Amsterdam
1018 WB Amsterdam The Netherlands