Factoring Algorithms
1. Pollard’s $p-1$ algorithm
   - Pollard’s rho algorithm (Omitted)
2. Dixon’s random squares algorithm
   - Morrison and Brillhart’s continued fraction method (Omitted)
   - Lentra’s elliptic curve method (ECM) (Omitted)
3. Quadratic sieve (QS)
4. Factoring algorithms in practice
1. Pollard's p-1 algorithm

input: an integer \( n \), and a prespecified “bound” \( B \)

output: factors of \( n \)

\( a \leftarrow 2 \)

for \( j \leftarrow 2 \) to \( B \)

\( \text{do } a \leftarrow a^j \mod n \)

\( d \leftarrow \gcd(a - 1, n) \)

if \( 1 < d < n \)

then return\((d)\)

else return\("failure\)
Suppose \( p \) is a prime divisor of \( n \), and suppose that \( q \leq B \) for every prime power \( q \mid (p-1) \). Then
\[ (p-1) \mid B! \]
At the end of for loop, we have
\[ a = 2^B! \mod n \]
Now
\[ 2^{p-1} = 1 \mod p \quad \text{(by Fermat’s little Thm)} \]
Since \( (p-1) \mid B! \), it follows
\[ a = 2^B! = 1 \mod p \]
and hence \( p \mid (a-1) \). Since we also have \( p \mid n \),
d\( = \gcd(a-1, n) \) will be a non-trivial divisor of \( n \)
(unless \( a = 1 \)).
E.g. $n=15770708441$, $B=180$

\[ a = 2^{180}! = 11620221425 \]
\[ D = \gcd(a-1, n) = 135979 \]

In fact, the complete factorization of $n$ into primes is
\[ 15770708441 = 135979 \times 115979 \]

The factorization succeeds because 135978 has only “small” prime factors:
\[ 135978 = 2 \times 3 \times 131 \times 173 \]
2. Dixon’s random squares algorithm

The idea is to locate \( x, y \in \mathbb{Z}_n \) with \( x^2 \equiv y^2 (\mod n) \); if \( x \neq \pm y (\mod n) \), then \( \gcd(x+y, n) \) is a nontrivial factor of \( n \).

(Why?) since \( n \mid (x-y)(x+y) \) but neither of \( x-y \) or \( x+y \) is divisible by \( n \).

- Eg. \( n=15, x=2, y=7 \) \( (2^2=7^2 \mod 15) \Rightarrow \gcd(2+7, 15)=3 \) is a nontrivial factor of \( n \).
- Eg. \( n=77, x=10, y=32 \) \( (10^2=32^2 \mod 77) \Rightarrow \gcd(10+32, 77)=7 \) is a nontrivial factor of \( n \).
factor base and $p_t$-smooth

- A factor base $B=\{p_1, p_2, \ldots, p_t\}$ consisting of the first $t$ primes is selected. If $b$ factors over $B$, $b$ is said to be $p_t$-smooth.

- Eg: $B=\{2, 3, 5\}$, $b=2^3 \cdot 5^6$ is 5-smooth; $b=2^3 \cdot 7^6$ is not 5-smooth.

- We may include -1 in $B$ to handle the negative $b$

$B=\{p_0, p_1, p_2, \ldots, p_t\}$, with $p_0=-1$. 
Algorithm

input: a composite integer \( n \) and factor base \( B = \{p_1, p_2, \ldots, p_t\} \)
output: factors of \( n \)

(1) Suppose \( t+1 \) pairs \((a_i, b_i = a_i^2 \text{ mod } n)\) are obtained, where

\[ b_i \text{ is } p_t\text{-smooth over } B \text{ and the factorizations are given by} \]

\[ b_i = \prod_{j=1}^{t} p_j^{e_{ij}}, \quad 1 \leq i \leq t+1. \]

(2) A set \( S \) is to be selected so that \( \prod_{i \in S} b_i \) has only even powers of primes appearing.

(3) Let \( x = \prod_{i \in S} a_i \) and \( y = \sqrt[\prod_{i \in S} b_i] \), and do the following compare

3.1 If \( x = \pm y \text{ (mod } n) \), then return "not factoring".

3.2 If \( x \neq \pm y \text{ (mod } n) \), then return \( \gcd(x + y, n) \).
Eg: $n=10057$, $t=5$, $B=\{2,3,5,7,11\}$

<table>
<thead>
<tr>
<th>$i$</th>
<th>$a_i$</th>
<th>$b_i = a_i^2 \mod n$</th>
<th>factorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>231</td>
<td>1018</td>
<td>$2*509$ (discard!)</td>
</tr>
<tr>
<td>1</td>
<td>105</td>
<td>968</td>
<td>$2^3*11^2$</td>
</tr>
<tr>
<td>2</td>
<td>115</td>
<td>3168</td>
<td>$2^5<em>3^2</em>11$</td>
</tr>
<tr>
<td>3</td>
<td>1006</td>
<td>6336</td>
<td>$2^6<em>3^2</em>11$</td>
</tr>
<tr>
<td>4</td>
<td>3010</td>
<td>8800</td>
<td>$2^5<em>5^2</em>11$</td>
</tr>
<tr>
<td>5</td>
<td>4014</td>
<td>882</td>
<td>$2<em>3^2</em>7^2$</td>
</tr>
<tr>
<td>6</td>
<td>4023</td>
<td>2816</td>
<td>$2^8*11$</td>
</tr>
</tbody>
</table>

If $S=\{4,5,6\}$, then $x=3010*4014*4023 \mod n=2748$

$y=2^7*3*5*7*11 \mod n=7042$

Since $2748 \neq \pm 7042 (\mod n)$, we obtain a nontrivial factor $\gcd(x+y,n)=89$, and $10057=89*113$.

If $S=\{1,5\}$, then $x=105*4014 \mod n=9133$ and $y=2^2*3*7*11=924$.

Unfortunately, $9133 \equiv -924 (\mod n)$, and no useful information is obtained.
Eg : n=15770708441, t=6, B={2,3,5,7,11,13}

\[8340934156^2 = 3\times7 \pmod{n}\]
\[12044942944^2 = 2\times7\times13 \pmod{n}\]
\[2773700011^2 = 2\times3\times13 \pmod{n}\]

\[(8340934156\times12044942944\times2773700011)^2 = (2\times3\times7\times13)^2 \pmod{n}\]

\[9503435785^2 = 546^2 \pmod{n}\]

gcd(9503435785–546, 15770708441)=115759

to find the factor 115759 of n
3. Quadratic sieve algorithm (simple version)

input: a composite integer \( n \)
output: factors of \( n \)

(1) Construct a factor base with \(-1\)

(2) Define \( m = \left\lfloor \sqrt{n} \right\rfloor \), \( q(z) = (z + m)^2 - n \)

(3) Let \( a_i = z + m \) and \( b_i = q(z) = a_i^2 - n \) for \( z = 0, 1, -1, 2, -2, \ldots \). A set \( S \) is to be selected so that \( \prod_{i \in S} b_i \) has only even powers of primes appearing.

(4) Let \( x = \prod_{i \in S} a_i \) and \( y = \sqrt{\prod_{i \in S} b_i} \), and do the following

\[ x \neq \pm y \pmod{n}, \text{ then return } \gcd(x + y, n). \]
- **Improvements:**

- We may include -1 in B to handle the negative b
  \[ B = \{ p_0, p_1, p_2, \ldots, p_t \}, \text{ with } p_0 = -1. \]

- Define
  \[ m = \left\lfloor \sqrt{kn} \right\rfloor, \quad q(z) = (z + m)^2 - kn \]

Let \( a_i = z + m \) and \( b_i = q(z) = a_i^2 - kn \)
for \( z = 0, 1, -1, 2, -2, \ldots \) \( \text{ and } k = 1, 2, \ldots \).
Eg: n=10057

\[ m = \left\lfloor \sqrt{n} \right\rfloor = 100 \]

\[ q(z) = (z + 100)^2 - 10057 \]

\[ B = \{2,3,11,19\} \cup \{-1\} \]

<table>
<thead>
<tr>
<th>( z )</th>
<th>( a = z + m )</th>
<th>( b = q(z) )</th>
<th>factorization</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>100</td>
<td>-57</td>
<td>-3*19</td>
</tr>
<tr>
<td>-1</td>
<td>99</td>
<td>-256</td>
<td>-2^8</td>
</tr>
<tr>
<td>1</td>
<td>101</td>
<td>144</td>
<td>2^4*3^2</td>
</tr>
<tr>
<td>-3</td>
<td>97</td>
<td>-648</td>
<td>-2^3*3^4</td>
</tr>
<tr>
<td>5</td>
<td>105</td>
<td>968</td>
<td>2^3*11^2</td>
</tr>
</tbody>
</table>

If \( S = \{1\} \), then \( x = 101 \) and \( y = 2^2\*3 \).

Since \( x \not\equiv y \pmod{n} \), we obtain a nontrivial factor \( \gcd(x+y,n) = 113 \), and \( 10057 = 89\*113 \).

If \( S = \{-1,-3, 5\} \), then \( x = 99\*97\*105 \) and \( y = 2^7\*3^2\*11 \).

Unfortunately, \( x \equiv y \pmod{n} \), and no useful information is obtained.
4. Factoring algorithms in practice

(Asymptotic running times)

1. Quadratic sieve
   \[ O(\exp((1 + o(1)) \sqrt{\ln n \ln \ln n})) \]

2. Elliptic curve (p is the smallest prime factor of n)
   \[ O(\exp((1 + o(1)) \sqrt{2 \ln p \ln \ln p})) \]

3. Number field sieve
   \[ O(\exp((1.92 + o(1)) (\ln n)^{\frac{1}{3}} (\ln \ln n)^{\frac{2}{3}})) \]
The RSA Challenge
http://www.rsa.com/rsalabs/node.asp?id=2092

- **RSA-640 is factored!** (11/2/2005) (193 digits)
- **RSA-200 is factored!** (5/9/2005) (663 bits)
- **RSA-576 is factored!** (12/3/2003) (174 digits)
- **RSA-160 is factored!** (1/6/2003) (530 bits)
- **RSA-155 is factored!** (8/22/1999) (512 bit)
- **RSA-140 is factored!** (2/2/1999)