Integer factorization,
part 1: the Q sieve

D. J. Bernstein
Sieving small integers $i > 0$ using primes 2, 3, 5, 7:

<p>| | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>7</td>
<td></td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>9</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>10</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>13</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>14</td>
<td>2</td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td>3</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

etc.
Sieving $i$ and $611 + i$ for small $i$
using primes 2, 3, 5, 7:

|    | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 | 11 | 12 | 13 | 14 | 15 | 16 | 17 | 18 | 19 | 20 |
|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|----|
|    | 2  | 2  | 3  | 22 | 5  | 2  | 3  | 22 | 33 | 2  | 5  | 22 | 3  | 2  | 7  | 22 | 2  | 33 | 2  | 5  |
|    | 2  | 3  | 5  | 22 | 33 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 | 22 |

<table>
<thead>
<tr>
<th></th>
<th>612</th>
<th>613</th>
<th>614</th>
<th>615</th>
<th>616</th>
<th>617</th>
<th>618</th>
<th>619</th>
<th>620</th>
<th>621</th>
<th>622</th>
<th>623</th>
<th>624</th>
<th>625</th>
<th>626</th>
<th>627</th>
<th>628</th>
<th>629</th>
<th>630</th>
<th>631</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>2 2</td>
<td>3 3</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
<td>2 2</td>
<td>3 5</td>
</tr>
</tbody>
</table>

etc.
Have complete factorization of the “congruences” \( i(611 + i) \) for some \( i \)'s.

\[
14 \cdot 625 = 2^1 3^0 5^4 7^1.
64 \cdot 675 = 2^6 3^3 5^2 7^0.
75 \cdot 686 = 2^1 3^1 5^2 7^3.
\]

\[
14 \cdot 64 \cdot 75 \cdot 625 \cdot 675 \cdot 686 \equiv 2^8 3^4 5^8 7^4 = (2^4 3^2 5^4 7^2)^2.
\]

\[
\gcd\{611, 14 \cdot 64 \cdot 75 - 2^4 3^2 5^4 7^2\} = 47.
\]

\[
611 = 47 \cdot 13.
\]
Why did this find a factor of 611? Was it just blind luck:
\[ \gcd\{611, \text{random}\} = 47? \]

No.

By construction \( n \) divides \( s^2 - t^2 \) where \( s = 14 \cdot 64 \cdot 75 \) and \( t = 2^4 3^2 5^4 7^2 \).

So each prime \( \geq 7 \) dividing \( n \) divides either \( s - t \) or \( s + t \).

Not terribly surprising (but not guaranteed in advance!) that one prime divided \( s - t \) and the other divided \( s + t \).
Why did the first three completely factored congruences have square product? Was it just blind luck?

Yes. The exponent vectors \((1, 0, 4, 1), (6, 3, 2, 0), (1, 1, 2, 3)\) happened to have sum 0 mod 2.

But we didn’t need this luck! Given long sequence of vectors, quickly find nonempty subsequence with sum 0 mod 2.
This is linear algebra over $\mathbb{F}_2$. Guaranteed to find subsequence if number of vectors exceeds length of each vector.

e.g. for $n = 671$:
\[
1(n + 1) = 2^53^15^07^1;
4(n + 4) = 2^23^35^27^0;
15(n + 15) = 2^13^15^17^3;
49(n + 49) = 2^43^25^17^2;
64(n + 64) = 2^63^15^17^2.
\]

$\mathbb{F}_2$-kernel of exponent matrix is gen by $(0 1 0 1 1)$ and $(1 0 1 1 0)$; e.g., $1(n + 1)15(n + 15)49(n + 49)$ is a square.
Plausible conjecture: Q sieve can separate the odd prime divisors of any $n$, not just 611.

Given $n$ and parameter $y$:

1. Try to completely factor $i(n + i)$ for $i \in \{1, 2, 3, \ldots, y^2\}$ into products of primes $\leq y$.

2. Look for nonempty set of $i$’s with $i(n + i)$ completely factored and with $\prod_i i(n + i)$ square.

3. Compute $\gcd\{n, s - t\}$ where $s = \prod_i i$ and $t = \sqrt{\prod_i i(n + i)}$. 
How large does $y$ have to be for this to find a square?

Let’s aim for number of completely factored congruences to exceed length of each vector, guaranteeing a square. (This is somewhat pessimistic; smaller numbers usually work.)

Vector length $\approx y/\log y$.

Will there be $> y/\log y$ completely factored congruences out of $y^2$ congruences?
What’s chance of random \( i(n + i) \) being \( y\text{-smooth} \), i.e., completely factored into primes \( \leq y \)?

Consider, e.g., \( y = \lfloor n^{1/10} \rfloor \).

Uniform random integer in \([1, y^2]\) has \( y\)-smoothness chance \( \approx 0.306 \);
uniform random integer in \([1, n]\) has chance \( \approx 2.77 \cdot 10^{-11} \).

Plausible conjecture:
\( y\)-smoothness chance of \( i(n + i) \) is \( \approx 8.5 \cdot 10^{-12} \).

Find \( \approx 8.5 \cdot 10^{-12} y^2 \)
fully factored congruences.
If \( n \geq 2^{340} \) and \( y = \lfloor n^{1/10} \rfloor \) then
\[ 8.5 \cdot 10^{-12} y^2 > 3y / \log y, \]
and approximations seem fairly close, so conjecturally the \( \mathbb{Q} \) sieve will find a square.

Find many independent squares with negligible extra effort.
If \( \gcd \) turns out to be 1, try the next square.

Conjecturally always works: splits odd \( n \) into prime-power factors.
How about $y \approx n^{1/u}$ for larger $u$?

Uniform random integer in $[1, n]$ has $n^{1/u}$-smoothness chance roughly $u^{-u}$.

Plausible conjecture: $Q$ sieve succeeds with $y = \lfloor n^{1/u} \rfloor$ for all $n \geq u^{(1+o(1))u^2}$; here $o(1)$ is as $u \to \infty$. 
How about letting \( u \) grow with \( n \)?

Given \( n \), try sequence of \( y \)'s in geometric progression until \( Q \) sieve works;
e.g., increasing powers of 2.

Plausible conjecture: final \( y \in \exp \sqrt{\left( \frac{1}{2} + o(1) \right) \log n \log \log \log n} \),
\( u \in \sqrt{(2 + o(1)) \log n / \log \log \log n} \).

Cost of \( Q \) sieve is a power of \( y \), hence subexponential in \( n \).
More generally, if $y \in \exp \sqrt{\left( \frac{1}{2c} + o(1) \right) \log n \log \log n}$, conjectured $y$-smoothness chance is $1/y^{c+o(1)}$.

Find enough smooth congruences by changing the range of $i$'s: replace $y^2$ with $y^{c+1+o(1)} = \exp \sqrt{\left( \frac{(c+1)^2+o(1)}{2c} \right) \log n \log \log n}$.

Increasing $c$ past 1 increases number of $i$’s but reduces linear-algebra cost. So linear algebra never dominates when $y$ is chosen properly.
Improving smoothness chances

Smoothness chance of \(i(n + i)\) degrades as \(i\) grows.
Smaller for \(i \approx y^2\) than for \(i \approx y\).

Crude analysis: \(i(n + i)\) grows.
\(\approx yn\) if \(i \approx y\);
\(\approx y^2n\) if \(i \approx y^2\).

More careful analysis:
\(n + i\) doesn’t degrade, but
\(i\) is always smooth for \(i \leq y\),
only 30\% chance for \(i \approx y^2\).

Can we select congruences to avoid this degradation?
Choose \( q \), square of large prime. Choose a “\( q \)-sublattice” of \( i \)’s: arithmetic progression of \( i \)’s where \( q \) divides each \( i(n + i) \).

e.g. progression \( q - (n \mod q) \), \( 2q - (n \mod q) \), \( 3q - (n \mod q) \), etc.

Check smoothness of generalized congruence \( i(n + i)/q \) for \( i \)’s in this sublattice.

e.g. check whether \( i, (n + i)/q \) are smooth for \( i = q - (n \mod q) \) etc.

Try many large \( q \)’s.

Rare for \( i \)’s to overlap.
e.g. \( n = 314159265358979323 \): 

Original \( \mathbf{Q} \) sieve:

<table>
<thead>
<tr>
<th>( i )</th>
<th>( n + i )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>314159265358979324</td>
</tr>
<tr>
<td>2</td>
<td>314159265358979325</td>
</tr>
<tr>
<td>3</td>
<td>314159265358979326</td>
</tr>
</tbody>
</table>

Use \( 997^2 \)-sublattice, 

\( i \in 802458 + 994009 \mathbf{Z} \):

<table>
<thead>
<tr>
<th>( i )</th>
<th>( (n + i)/997^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>802458</td>
<td>316052737309</td>
</tr>
<tr>
<td>1796467</td>
<td>316052737310</td>
</tr>
<tr>
<td>2790476</td>
<td>316052737311</td>
</tr>
</tbody>
</table>
Crude analysis: Sublattices eliminate the growth problem. Have practically unlimited supply of generalized congruences
\[(q-(n \mod q)) \frac{n+q-(n \mod q)}{q}\]
between 0 and \(n\).

More careful analysis: Sublattices are even better than that! For \(q \approx n^{1/2}\) have
\[i \approx (n + i)/q \approx n^{1/2} \approx y^{u/2}\]
so smoothness chance is roughly
\[(u/2)^{-u/2}(u/2)^{-u/2} = 2^u/u^u,\]
\(2^u\) times larger than before.
Even larger improvements from changing polynomial \( i(n + i) \).

“Quadratic sieve” (QS) uses \( i^2 - n \) with \( i \approx \sqrt{n} \); have \( i^2 - n \approx n^{1/2 + o(1)} \), much smaller than \( n \).

“MPQS” improves \( o(1) \) using sublattices: \( (i^2 - n)/q \). But still \( \approx n^{1/2} \).

“Number-field sieve” (NFS) achieves \( n^{o(1)} \).
Fast linear algebra

Given \( y \times y \) matrix over \( \mathbb{F}_2 \) specifying linear \( M : \mathbb{F}_2^y \to \mathbb{F}_2^y \).

“Solving linear equations”: given \( w \in \mathbb{F}_2^y \),
find some \( v \in \mathbb{F}_2^y \) with \( Mv = w \).

Using an algorithm for that:
Choose uniform random \( r \in \mathbb{F}_2^y \);
compute \( w = Mr \); use algorithm to find \( v \) with \( Mv = w \).
This produces uniform random kernel element, namely \( v - r \).
“Elimination” solves linear equations using $O(y^3)$ bit operations.

“Series denominators” solve linear equations using $y^{2+o(1)}$ bit operations if the equations are sparse.

“Sparse”: can evaluate $M$ using $y^{1+o(1)}$ bit operations. Certainly true in $\mathbb{Q}$ sieve with usual choices of $y$. 
What’s the denominators method?

Consider nontrivial relation
\[ p_0 w + p_1 Mw + \cdots + p_y M^y w = 0. \]

I’ll assume \( p_0 = 1 \) for simplicity, so
\[ w = -p_1 Mw - \cdots - p_y M^y w = Mv \]
where \( v = -p_1 w - \cdots. \)

Consider series in \( \mathbf{F}_2^y[[t]] \):
\[ w + (Mw)t + (M^2w)t^2 + \cdots. \]

Multiplying series by poly \( p_0 t^y + p_1 t^{y-1} + \cdots + p_y t^0 \)
in \( \mathbf{F}_2[t] \) produces
poly in \( \mathbf{F}_2^y[t] \) of degree \( < y \).
Save time by projecting from $\mathbf{F}_2^y[[t]]$ to $\mathbf{F}_2[[t]]$.

Choose linear $r : \mathbf{F}_2^y \rightarrow \mathbf{F}_2$. Series $r(w) + r(Mw)t + \cdots$ has denominator dividing $p_0t^y + p_1t^{y-1} + \cdots + p_yt^0$.

Compute denominator of series from first $2y$ terms of series via continued fractions.

Repeat for three random $r$’s, compute lcm of denominators. Obtain $p_0, p_1, \ldots$ with probability close to 1.