A BOX METHOD FOR FACTORING SEMI-PRIMES

INTRODUCTION:

In a recent note on this web page we have shown that any semi-prime \( N = pq \) satisfies the equality:

\[
(\sqrt{N})^2 + y^2 = x^2
\]

where:

\[
x = \frac{(q + p)}{2} \quad \text{and} \quad y = \frac{(q - p)}{2}
\]

Both \( x \) and \( y \) must be integers with the basic requirement that:

\[
y = \sqrt{x^2 - N}
\]

So, for example, one has \( N = 377 \) producing \( x = 21 \) and \( y = 8 \) since \( 377 + 64 = 441 \). It is convenient to cast the relation between \( \sqrt{N} \), \( x \) and \( y \) into a Pythagorean triangle with the sides marked as shown:

By dividing each side of this triangle by \( \sqrt{N} \) one arrives at the equivalent form given by the triplet:

\[[1, k/r, k]\]
with the hypotenuse \( k = \frac{x}{\sqrt{N}} \) and \( r = \frac{x}{y} \). we can plot this last result as the curve-

\[
k = \frac{r}{\sqrt{r^2 - 1}}
\]

which looks as follows-

Also I have added to this continuous curve (shown in red) the values for eight larger semi-primes. The graph clearly shows that all semi-primes regardless of size fall along this curve but one usually does not know at exactly what point a given semi-prime \( N \) lies before solving the problem.

It is our purpose here to introduce a new box method to recover the integer value of both \( x \) and \( y \) for a given \( N = pq \) when \( p < \sqrt{N} < q \).

**ESSENCE OF THE BOX METHOD:**

In our previous note we found that for many semi-primes have values of \( k \) not far removed from \( k = 1.05 \). This suggests one conduct a search for integer \( y \) by evaluating the radical-

\[
y = \sqrt{(x_{\text{APPROX}} + n)^2 - N} =
\]
over the range \(-b<n<b\), with \(b\) being no larger than 10% of \(k\sqrt{N}\). Here \(x_{\text{APPROX}}\) is the closest integer to \(k\sqrt{N}\). If we consider \(N=1211111\) with 
\(\sqrt{N}=348.01\) one gets \(x_{\text{APPROX}}=365\) when \(k=1.05\). Applying the search program-

\[
\text{for } n \text{ from } -36 \text{ to } 36 \text{ do } \{n+365, \text{evalf}(\sqrt{(365+n)^2-N})\} \text{od;}
\]

one finds after nine trials that \(x=356\) and \(y=75\). Hence \(p=356-75=281\) and 
\(q=356+75=431\). From this last solution we see that \(k\) is actually \(k=1.0229\). So a choice of \(k=1.02\) would produce an integer \(y\) even faster.

Now the problem is that in factoring large \(N\)s, such as those of importance in public key cryptography, is that we don’t know beforehand the precise value of \(k\) and so we do not know where along the above \(k-r\) curve the particular solution falls. To hone in on this solution with our present search method we start with a search using 
\(k=1.05\) where \(x_{\text{APPROX}}=1.05\sqrt{N}\)and extend the search from \(n=0\) down to 
The nearest integer form of \(n= -\sqrt{N}\). This calculation is easiest to visualize by drawing a box over the 2D area \(-\sqrt{N}<n<0\), and \(v\) from that existing at the starting point indicated in the following graph out to very large \(r\). The following graph indicates the resultant box covering this search-

If no integer solution for \(x\) and \(y\) is found in searching over this first box, a second calculation is initiated over the second box shown and so on. The second box extends over \(1.05<k<1.10\).

**SAMPLE CALCULATIONS USING THE BOX METHOD:**
Let us demonstrate the box method by factoring several examples of large semi-primes. Consider first-

\[ N=598523 \quad \text{where} \quad \sqrt{N}=773.64 \text{ and } x_{\text{APPROX}}=812 \text{ when } k=1.05. \]

A search extending from \( n=0 \) down to \( n=-38 \) produces no integer results. Going on to the second box where \( k=1.10 \) and \( x_{\text{APPROX}}=851 \) and extending the search from \( n=0 \) to \(-39\) again produces no integer result. However for the third box where \( k=1.15 \) and \( x_{\text{APPROX}}=890 \), one finds the integer result \([x,y]=[858, 371]\). Hence the point on the k-r curve falls at \( k=x/\sqrt{N}=1.1103 \) and \( r=x/y=2.315 \). The values of \( p \) and \( q \) follow-

\[ p=858-371=487 \quad \text{and} \quad q=858+371=1229 \]

The next semi-prime we factor by the box method is-

\[ N=744451 \quad \text{where} \quad \sqrt{N}=2728.46 \text{ and } x_{\text{APPROX}}=1.05 \sqrt{N}=2865 \]

Running the search extending from \( n=0 \) down to \( n=-136 \) produces the integer solution-

\[ [x,y]=[2781, 538] \]

This corresponds to \( p=2243 \) and \( q=3319 \). The point on the k-r curve this time lies within the first box at \( k=2781/2728.46=1.1926 \) and \( r=2781/538=5.1691 \).

As a third example consider the semi-prime –

\[ N=185451 \quad \text{where} \quad \sqrt{N}=442.098 \]

Starting with the first box we get no integer answer. Then onto the second and third box with still no integer answer. Finally in the fourth box whose upper corner lies at \( k=1.2 \) we search over \(-26<n<0\) and get an answer-

\[ [x,y]=[526, 285] \]

This means that \( p=526-285=241 \) and \( q=526+285=811 \). Also \( k=x/\sqrt{N}=1.1898 \) and \( r=526/285=1.8456 \).

CONCLUDING REMARKS:

The box method outlined above appears to work and will yield quick integer answers for \( x \) and \( y \) when the semi-prime is confined to one of the lower number boxes. Any \( N \) which has \( r>3.3 \) can be handled by using \( k=1.05 \). To be in this first box the requirement is that \( p \) and \( q \) do not differ by a large fraction from each other.
The larger the p and q difference the higher the box number will be and so will require more effort. Since one does not know beforehand in which box the integer solution for x and y lies, it is necessary to start with box one and work oneself up to box two and higher until a solution is found. This procedure can be automated. An alternate approach is to just guess a value of k and run a search about the corresponding $x_{\text{APPROX}}$. If no answer is found then pick another k and repeat the procedure. Eventually, if one leaves no gaps in the k-r curve, an integer value for x and y will be found. To make a secure public key in cryptography it will be a good idea to have $q >> p$ as this will require an adversary to go through hundreds of boxes before the typical 100 digit long semi-prime is factored. I leave you with an example of an unsecure 12 digit long semi-prime-

$$N=159772014601 \text{ where } \sqrt{N}=399714.9126$$

By looking at the bottom of the first box one has the search program-

```
for n from 0 to 40 do {399715+n, evalf(sqrt((n+399715)^2-N))}od;
```

It produces $[x,y]=[399749, 5220]$ at $k=1.000085269$ and requires only 34 trial calculations to get the answer. Here r has the rather large value of $r=x/y=76.58$.

U.H.Kurzweg
Feb.1, 2018