The factoring of large semi-primes continues to be of major interest in connection with public key cryptography. Such numbers have the form $N=pq$, where $p$ and $q$ are prime numbers to be determined. Our purpose here is to look at an alternate new way to factor such semi-primes when $N$ becomes large.

As we have already shown in several earlier notes, all primes $p$ and $q$, whenever both are greater than 3, must have the form-

$$p=6n\pm1 \quad \text{and} \quad q=6m\pm1$$

The choice of sign is dictated by the form of $N$ which must also always have the form $N=6k\pm1$. When $N \mod(6)=1$, we must have-

$$p=6n+1 \quad \text{and} \quad q=6m+1 \quad \text{or} \quad p=6n-1 \quad \text{and} \quad q=6m-1$$

When $N \mod(6)=5$, we must have –

$$p=6n+1 \quad \text{and} \quad q=6m-1 \quad \text{or} \quad p=6n-1 \quad \text{and} \quad q=6m+1$$

Any semi-prime will have primes $p$ and $q$ satisfying one of these four possibilities.

Let us work out in more detail these four possible cases-

1. $N \mod(6)=1$ with $p=6n+1$ and $q=6m+1$: Multiplying things together produces-

$$N=36nm+6(n+m)+1 = 9[(m+n)^2-(m-n)^2]+6(n+m)+1$$

On letting $A=n+m$ and $B=m-n$ we have-

$$B = \frac{\sqrt{(3A+1)^2 - N}}{3}$$

From this last equation one sees at once that $A>\lfloor\sqrt{N}-1\rfloor/3$ and that the right hand side RHS must be a positive integer since $B$ is required to be such.

2. $N \mod(6)=1$ with $p=6n-1$ and $q=6m-1$: Some manipulations here produce-

$$B = \frac{\sqrt{(3A-1)^2 - N}}{3}$$

3. $N \mod(6)=5$ and $p=6n-1$ with $q=6m+1$: We find-
\[ B = \frac{-1 + \sqrt{9A^2 - N}}{3} \]

(4) \( N \mod(6) = 5 \) with \( p = 6n + 1 \) and \( q = 6m - 1 \): This time we have-

\[ B = \frac{1 + \sqrt{9A^2 - N}}{3} \]

Since \( A = n + m \) and \( B = n - m \) will always be integers, it is necessary that the right hand sides (RHS) of the above four possible values of \( B \) all have the form of real positive integers.

Let us now consider some specific examples. Take first the simple case of \( N = 403 \) where \( \sqrt{N} = 20.0748 \) and \( N \mod(6) = 1 \). This means \( N = 6(67) + 1 \) and we can try \( p = 6n + 1 \) and \( q = 6m + 1 \) which represents form (1) above. Accordingly \( \text{RHS} = \sqrt{(3A + 1)^2 - 403} / 3 \). For \( \text{RHS} \) to remain real it is necessary that \( A > \left[ \sqrt{403} - 1 \right] / 3 = 6.358 \). So starting with \( A = 6 \) we get the following table-

<table>
<thead>
<tr>
<th>( A = n + m )</th>
<th>( B = \sqrt{(3A + 1)^2 - N} / 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>( \sqrt{-42} / 3 )</td>
</tr>
<tr>
<td>7</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>( \sqrt{222} / 3 )</td>
</tr>
</tbody>
</table>

So after just two trials we have found \( A = n + m = 7 \) and \( B = m - n = 3 \). We get \( n = 2 \) and \( m = 5 \). This in turn yields the factored result \( p = 31 \) and \( q = 13 \). Since this form already gave a correct answer it is not necessary to go on to the second possibility (2).

Let us next take the historically interesting semi-prime \( N = 455839 \) which is often used to demonstrate the Lenstra elliptic curve factorization method. Here \( N = 6(6n - 1) + 1 \) and \( N \mod(6) = 1 \). Let us try case (2) above. This means \( \text{RHS} = \sqrt{(3A - 1)^2 - N} / 3 \) and we should start the search near \( A = \left[ 1 + \sqrt{N} \right] / 3 = 225.386 \). Doing so we get the following table-

<table>
<thead>
<tr>
<th>( A )</th>
<th>( B = \sqrt{(3A - 1)^2 - N} / 3 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>225</td>
<td>( \sqrt{-1563} / 3 )</td>
</tr>
<tr>
<td>226</td>
<td>( \sqrt{2490} / 3 )</td>
</tr>
<tr>
<td>227</td>
<td>27</td>
</tr>
</tbody>
</table>

So after just three trials we find the integer solutions \( A = 227 \) and \( B = 27 \). This is equivalent to \( p = 599 \) and \( q = 761 \) and leads at once to the factored result-

\[ 455839 = 599 \times 761 \]

The procedure is clearly much shorter (in this case) than the elliptic curve factorization method itself.

As a third example consider the Mersenne Number \( N = 2^{11} - 1 = 2047 \) which has \( N \mod(6) = 1 \). Here we choose \( N = (6n - 1)(6m - 1) \) as covered by approach (2) above. It says \( A > \left[ \sqrt{N} + 1 \right] / 3 = 15.41 \). Trying \( A = 16, 17, 18, \) and
19, we find the integer solution $A=m+n=19$ and $B=m-n=11$. This means $n=4$ and $m=15$. Hence the number factors into-

$$2047 = [6(4)-1][6(15)-1] = 23 \times 89$$

So just four trials have produced a factoring.

Although all of the above examples gave rapid solutions requiring only a few trials, this will generally not be the case when $N$ gets very large and/or $A$ and $B$ are nearly equal. An example of such a number is the Fermat Number-

$$N=2^{32}+1=4294967297 = 641 \times 6700417$$

For this number $N \mod(6)=5$ and $641 \mod(6)=5$ and $6700417 \mod(6)=1$. It suggests $p=6n-1$ and $q=6m+1$ which leads to-

$$B = \frac{-1 + \sqrt{9A^2 - N}}{3}$$

with $A>\sqrt{N}/3=7415.953$. Since we already have given the factors $p$ and $q$, we know that $n=107$ and $m=1116736$. So we have $A=1116843$ and $B=1116629$. These two integers are almost equal and thus one would expect the need for many trials starting with the minimum of $A=7416$. Indeed it would require a little over a million trials $1116843-7416=1109427$. Clearly this is impractical and one must look for an alternative route. One way to sometimes get around the problem is to look at exponential forms of different $A$s and then multiply some of these together so that the radical becomes a perfect square. This is the approach used by most quadratic sieve approaches employed in factoring large semi-primes. The drawback of this approach is that it is often difficult to find product of several different $A$s so that they make $9A^2-N$ a perfect square. In theory it’s a good approach, but in practice becomes so cumbersome that no one has at this date managed to factor 500 digit long semi-primes by this method despite of intensive efforts including long runs on the world’s fastest super-computers. We suggest here an alternate (but brute force) approach using a new integer $C=A-B=2n$ and noting that $A+B=2m$. For the Fermat Number $N=2^{32}+1$ or other numbers of the type $N \mod(6)=5$, this produces the equations-

$$A = \frac{N + (3C - 1)^2}{6(3C-1)} \quad , \quad B = \frac{N + 1 - 9C^2}{6(3C-1)}$$

We can now pick a range of $Cs$ to find integer values of $A$ and $B$. A simple computer search using $N=2^{32}+1$ produces the integer solutions-

$$C=214 , \quad A=1116843 , \quad B=1116629$$

and requires just 214 trials to find. These integer values produce $n=107$ and $m=6699881$. Thus the Fermat number factors as-

$$2^{32}+1 = : = 4294967297 = 641 \times 6700417$$
This approach will work for any of the four cases defined above. They will each produce their own version of the following equations:

\[ A = f(C, N) \quad \text{and} \quad B = g(C, N) \]

One then simply searches for the integer value of \( C \) for which \( A \) and \( B \) are also simultaneously integers.

The following table summarizes the functional forms \( f(C, N) \) and \( g(C, N) \) for the four cases -

<table>
<thead>
<tr>
<th>Form of ( p ) and ( q )</th>
<th>( A = f(C, N) )</th>
<th>( B = g(C, N) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( p = 6n + 1, q = 6m + 1 )</td>
<td>( \frac{N - 1 + 9C^2}{6(3C + 1)} )</td>
<td>( \frac{N - (3C + 1)^2}{6(3C + 1)} )</td>
</tr>
<tr>
<td>( p = 6n - 1, q = 6m - 1 )</td>
<td>( \frac{N - 1 + 9C^2}{6(3C - 1)} )</td>
<td>( \frac{N - (3C - 1)^2}{6(3C - 1)} )</td>
</tr>
<tr>
<td>( p = 6n - 1, q = 6m + 1 )</td>
<td>( \frac{N + (3C - 1)^2}{6(3C - 1)} )</td>
<td>( \frac{N + 1 - 9C^2}{6(3C - 1)} )</td>
</tr>
<tr>
<td>( p = 6n + 1, q = 6m - 1 )</td>
<td>( \frac{N + (3C + 1)^2}{6(3C + 1)} )</td>
<td>( \frac{N + 1 - 9C^2}{6(3C + 1)} )</td>
</tr>
</tbody>
</table>

Notice that none of these involve the taking of radicals as in the earlier cases, but simply represent the ratio of two polynomials in \( C \).

The one line program used in the solution search is-

\[
\text{N:=given value ; A:=f(C,N); B:=g(C,N); for } C \text{ from a to b find } \{C,A,B\} \text{ od;}
\]

Here \( a < C < b \) is a range of \( C \) chosen. Typically \( a = 1 \) and \( b \) must be less than \( (-1 + \sqrt{N})/3 \) if we demand that \( p = 6n + 1 < \sqrt{N} \). One can place \( a \) and \( b \) (where \( a < b \)) anywhere in the region \( 1 < C < [-1 + \sqrt{N}]/3 \), so that smaller chunks of \( a \) to \( b \) can be tested separately.

Let us demonstrate this last technique in more detail for the relatively easy example of-

\[ N = 2623 = 43 \times 61 \]

Here \( N \mod(6) = 1 \), \( 43 \mod(6) = 1 \), and \( 61 \mod(6) = 1 \). This means \( p = 6n + 1 \) and \( q = 6m + 1 \) with \( n = 7 \) and \( m = 10 \). Also \( \sqrt{N} = 51.215 \). Let us pretend that we don’t know the values of \( p \) and \( q \). We then have from the above table-

\[
A = f(C, N) = \frac{9C^2 + N - 1}{6(3C + 1)} \quad \text{and} \quad B = g(C, N) = \frac{N - (3C + 1)^2}{6(3C + 1)}
\]

The three quantities \( A, B, \) and \( C \) must all be positive real integers. The upper limit on \( C \) is \( [-1 + \sqrt{N}]/3 = 16.738 \). We thus have our search routine given by-

\[
\text{for } C \text{ from 1 to 17 do } \{C,A,B\} \text{ od;}
\]

It yields the only possible all integer solution \( \{C,A,B\} = [14, 3, 17] \) in the given range. Hence \( m = (17 + 3)/2 = 10 \) and \( n = (17 - 3)/2 = 7 \). Thus the desired solution \( p = 43 \) and \( q = 61 \) has been obtained. We can also look at the graphical version of this solution as shown in the following figure-
It is only for the integer value of $C=14$ that integer values for $A$ and $B$ are found simultaneously. These values are $A=17$ and $B=3$.

The advantage of the present approach to factoring semi-primes is that we can look at calculation chunks $a<C<b$ anywhere in $1<C<\left[1+\sqrt{N}\right]/3$. For instance if we had chosen the chunk of $C$ in the range $13\leq C\leq 15$, the output containing the desired integer solution can be summarized as:

<table>
<thead>
<tr>
<th>C</th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>13</td>
<td>1381/80</td>
<td>341/80</td>
</tr>
<tr>
<td>14</td>
<td>17</td>
<td>3</td>
</tr>
<tr>
<td>15</td>
<td>1549/92</td>
<td>169/92</td>
</tr>
</tbody>
</table>

So only two trials are necessary to find the integer solutions compared to 14 trials in the initial calculation.

To get a better feel for the integer solutions $A$, $B$, and $C$ we can look at the following 1D graph-
We see that $C$ equals the distance between $A$ and $B$ and is always an even number. It can become large as the semi-prime $N=pq$ becomes large. However the fact that in the search process $A$ and $B$ can be evaluated over limited chunks for $C$ makes the process manageable. Remember that the full trial range is $1 < C < \left(-1 + \sqrt{N}\right)/3$. Probably a good starting point for factoring a typical large semi-prime is an integer $C$ close to $\left(-1 + \sqrt{N}\right)/6$. Only even values of $C$ need to be considered in the trials.

As one last example consider factoring $N=640081$ where $\sqrt{N}=2529.8$ and mod$(6)=1$. So trying $p=6n+1$ and $q=6m+1$ we know $1 < C < 842$. So we start our trials with a chunk $400 < C < 500$. It produces the integer solution $C=482$, $A=978$, and $B=496$. From this follows at once the factored result:

$$640081 = 1447 \times 4423$$

We point out that in the above procedure for finding $A$ and $B$ for a given $C$ requires a solution of just $A=f(C,N)$ or of $B=g(C,N)$ since, by the definition of $C=A-B$, we can always recover the other. It should also not be forgotten that sometimes the alternate approach of $B=F(A,N)$ involving square roots will work as well or even better than $C=g(B,N)$.