We have shown in several earlier notes that a semi-prime \( N=\text{pq} \) can be factored by letting –

\[
p=2n+1 \quad \text{and} \quad q=2m+1
\]

with \( n \) and \( m \) being unknown integers. On setting \( U=nm \) and \( V=n+m \), we obtain the Diophantine Equation –

\[
2U+V=(N-1)/2=M
\]

This has the solution –

\[
U=nm=k \quad \text{and} \quad V=(n+m)=M-2k
\]

where the \( ks \) are positive integers. On eliminating \( m \) from these last two equations, one finds the quadratic form –

\[
n^2-(M-2k)n+k=0
\]

Recalling that \( n=(p-1)/2 \) and introducing the new transformations –

\[
x=M-2k+1 \quad \text{and} \quad y=p-x
\]

allows us to write things as a Brahmagupta like equation –

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

where \( x, y \) and \( N \) are always integers.

We have thus reduced the procedure for factoring a semi-prime \( N=\text{pq} \) to solving the above algebraic expression for \( x \) and \( y \). The value of \( x \) must be greater than the square root of \( N \) and equal to the first value above this point where the radical has an integer value.

Both sides of this algebraic equation may be squared and then plotted as the following hyperbola –
The solution to this equation will have just two integer solutions which we can designate as \([x_1, y_1]\) and \([(N+1)/2, (N-1)/2]\). The first integer pair is of interest in determining the factors \(p\) and \(q\) of the semi-prime \(N\). The second pair is of not much interest since it only states the obvious that \(p=N\) and \(q=1\) will always satisfy \(N=pq\).

The procedure for finding the solution pair \([x_1, y_1]\), indicated as the blue dot in the figure, is to first write down its symbolic form-

\[\begin{align*}
x_1 &= a + b \\
y_1 &= \sqrt{(a+b)^2 - N}
\end{align*}\]

where \(a\) and \(b\) are constants. Next one chooses a value for ‘\(a\)’ greater or equal to \(\sqrt{N}\) and then evaluates the radical by varying \(b\) over a limited range until a value of \(b\) is found for which \(y_1\) is an integer. If none is found, the procedure is repeated with a larger value for ‘\(a\)’. Eventually an integer value for \(b\) will be found and both \(x_1\) and \(y_1\) will have been determined.

Let us demonstrate the search procedure for the trivial case of \(N=77\). Here \(\sqrt{77}=8.77\) so that we can choose as our starting point the value \(a=9\). The solution \(y_1=2\) follows at once if \(b\) is set to zero. Thus we have as our desired solution pair the result \([x_1, y_1]=[9, 2]\). From it one finds \(p=9+2=11\) and \(q=9-2=7\). When the semi-primes become larger the search can become time consuming but will still work when use of computers is made. Take, for instance, the semi-prime-

\[N = 12063943 \text{ where } \sqrt{N} = 3473.318730\ldots\]
Here we can start the search by setting \( a = 3474 \) and running \( b \) over the range \( 0 < b < 200 \). In computer language (with our MAPLE program) one has the command:

\[
N := 12063943; \quad \text{for } b \text{ from 0 to 200 do } \{ b, \text{ evalf(sqrt((3474+b)^2-N))} \} \text{ od;}
\]

On running it, no integer value for \( b \) is found for which the radical is an integer. So we next try \( a = 3474 + 200 = 3674 \) and again run \( b \) over the same range as before. That is we try:

\[
\text{for } b \text{ from 0 to 200 do } \{ b, \text{ evalf(sqrt((3674+b)^2-N))} \} \text{ od;}
\]

This time we find \( y_1 = 1419 \) at \( b = 78 \) and \( x_1 = (3674+78) = (a+b) = 3752 \). Thus we have factored the semi-prime \( N \) into:

\[
p = x_1 + y_1 = 5171 \quad \text{and} \quad q = x_1 - y_1 = 2333
\]

It took about 278 divisions to obtain this result in a split second. The standard brute force approach of finding \( p \) and \( q \) by dividing \( N \) by every prime up to about \( \sqrt{N} \) would take about 487 divisions and require knowledge of the values of the first 487 primes.

Another approach to solving \( y = \sqrt{x^2 - N} \) is to write \( N \) as the difference of two perfect squares. That is \( N = A^2 - B^2 \) so that \( x = A \) and \( y = B \). For example the semi-prime \( 299 = 18^2 - 5^2 \) so that \( x = 18 \) and \( y = 5 \) yielding \( p = x + y = 23 \) and \( q = x - y = 13 \). One is aided in this approach to factoring by tables such as the following:

<table>
<thead>
<tr>
<th>p/q</th>
<th>3</th>
<th>5</th>
<th>7</th>
<th>11</th>
<th>13</th>
<th>17</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>9</td>
<td>15</td>
<td>21</td>
<td>33</td>
<td>39</td>
<td>51</td>
</tr>
<tr>
<td>5</td>
<td>25</td>
<td>35</td>
<td>49</td>
<td>77</td>
<td>91</td>
<td>119</td>
</tr>
<tr>
<td>7</td>
<td>6^2-1^2</td>
<td>8^2-3^2</td>
<td>9^2-2^2</td>
<td>121</td>
<td>143</td>
<td>187</td>
</tr>
<tr>
<td>11</td>
<td>8^2-5^2</td>
<td>9^2-4^2</td>
<td>10^2-3^2</td>
<td>12^2-1^2</td>
<td>169</td>
<td>221</td>
</tr>
<tr>
<td>13</td>
<td>10^2-7^2</td>
<td>11^2-6^2</td>
<td>12^2-5^2</td>
<td>14^2-3^2</td>
<td>15^2-2^2</td>
<td>289</td>
</tr>
</tbody>
</table>

What is clear from this table is that any semi-prime \( N = pq = qp \) can always be written as the difference of two perfect squares. This of course makes sense since one is dealing with the hyperbola \( x^2 - y^2 = N \). Once the squares are found one has the factors \( p = x + y = A + B \) and \( q = x - y = A - B \). For example, from the table, we have \( N = 187 \) yields \( x = 14 \) and \( y = 3 \) so \( p = 17 \) and \( q = 11 \). The drawback of this factoring approach is that one will have a difficult time actually finding the values of the integers \( A \) and \( B \) whenever \( N \) gets very large.

The above search procedures for finding the \( p \) and \( q \) factors can become extremely time consuming when the semi-primes have several hundred digit length as they do in public key cryptography. Even more advanced techniques such as the elliptic curve factorization method and the generalized grid procedure are at present
incapable of breaking public keys of several hundred digits or greater length. All that a public key author has to make sure of is that such semi-prime public keys do not have p and q lie too close to each other or that q is a small number. A public key which I constructed several years ago has some 800 digits in p and 600 digits in q. The combined 1400 digit long semi-prime N =pq should be unbreakable for many years to come.

Easily factorable semi-primes are found when p and q are close in value to each other as is the case for double primes. For example if

\[ N = 146313135077 \]

We can choose a=382509 which is close to sqrt(N). Doing so, one finds that b=0 produces the integer value \( y_1 = 2 \). Hence N factors into-

\[ p = 382511 \quad \text{and} \quad q = 382507 \]

with little effort. Another semi-prime which is very easy to factor is \( N = 2677577 \) since it is divisible by the small prime \( q = 7 \) leaving \( p = 382511 \). In public key cryptography one avoids the use of such vulnerable semi-primes.

February, 2013