MORE ON FACTORING LARGE SEMI-PRIMES

We have discussed the factoring of semi-primes N=pq in several earlier notes. Our purpose here is to look some more at such factoring techniques using an approach we recently employed for Q=12n±1 primes. Our starting point is to note that with the exception of 2 all primes are odd integers and hence we can write-

$$\mathbf{N=(2n-1)(2m-1)=4U-2V+1 \quad with \quad U=nm \text{ and } V=n+m}$$

Letting K=(N-1)/2 we can write things as the Diophantine Equation-

$$\mathbf{2U-V=K}$$

For a given integer K this equation has multiple integer solutions given by-

$$\mathbf{V=-K \bmod(2)+2s}$$

with s=0,1,2,3 .. . V and K must both be even or both be odd since 2U is always even. Eliminating m , using the definitions for U and V, allows one to write-

$$\mathbf{p=(V-1)+\mathrm{sqrt}[(V-1)^2-N]}$$

With A=-Kmod(2) we have V=A+2s and p may be written as-

$$\mathbf{p=(A-1+2s)+\mathrm{sqrt}[(A-1+2s)^2-N]}$$

Note that had we used a minus sign in front of the radical, the other factor q<p would have been produced. To avoid imaginary results it is necessary that the radical in this last expression be positive. Hence the lowest allowed value for A must have-

$$\mathbf{A>1+\mathrm{sqrt}(N)}$$

Picking the value of A to be the nearest integer to 1+sqrt(N) and also one where the symmetry of A and K are the same, we can now run the integer values of s from zero to several hundred in the equation for p. Eventually a point will be reached where p becomes an integer. This will be the desired solution p. The second factor then follows from q=N/p . If the number of required s evaluations using different integer s exceeds several hundred, then one should start the evaluation again at another point V_1>A and run the calculations again. If that still does not yield an integer p then one can try the still larger starting point V_2 >V_1>A. With enough effort a solution (following the above approach)is always possible.

Let us carry out the calculation procedure starting with the small semi-prime-

$$\mathbf{N=2479}$$

Here A≈1+sqrt(2479)=1+49.7795…=50.7795... and K=(N-1)/2=1239. So we can take A=51 and carry out the following evaluation using the MAPLE command –

for s from 0 to 10 do {s, (50+2*s)+sqrt((50+2*s)^2-2479) }od;

It at once produces the answer –

p=67 at s=1 with **q=N/p=37**

Note that if we had not matched the odd character of both A and K we would have missed the solution for p completely. That is using A=52 gives no integer solution for the radical.

Consider next the larger semi-prime-

N=455839

This happens to be the number which often is used in the literature to demonstrate the Lenstra Elliptic Curve Factorization Technique. We have –

A=677 ≈1+sqrt(455839) and **K=(N-1)/2=227919**

and we solve-

$$\mathbf{p=(676+2s)+sqrt[(676+2s)^2-455839]}$$

starting with s=0. After just three calculations we arrive at the solution –

s=2 with p=761 and **q=N/p=599**

This result was thus obtained with considerably less effort than required by the Elliptic Curve Factorization Method.

Let us next attack the ten digit semi-prime

N=2177724221

Here we have –

2 U - V = K= 1088862110

We start our evaluation with A=46668 ≈1+sqrt(N)=46667.0928 and look for an integer p solution of the equation-

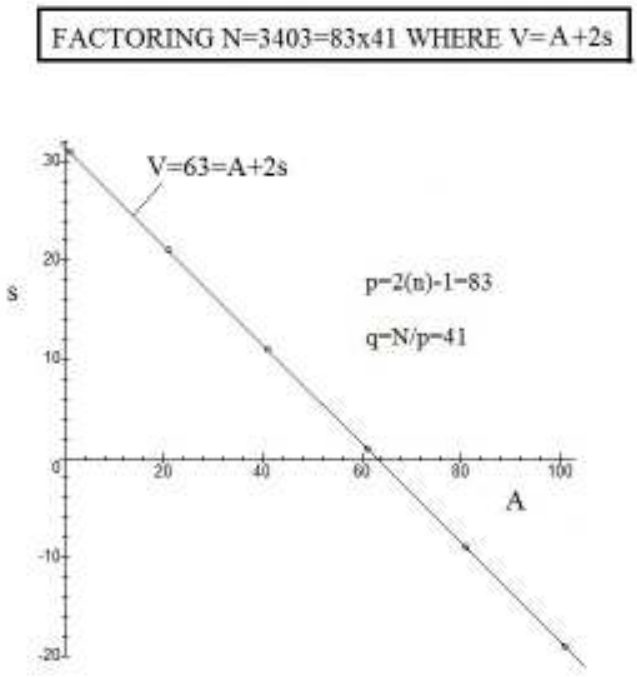
$$\mathbf{p=(46667+2s)+sqrt[(46667+2s)^2-2177724221]}$$

The first integer solution found is-

p=69383 at **s=1859**

This time it took 1859 operations to find the solution. To reduce this number of calculations one could change the starting point to V_1=50000. This would then require only 193 trial calculations involving s from 0 to 200. It is clear from this last result that when the semi-primes approach hundred digit length, as one encounters in public key cryptography, the search can become rather time consuming although, in theory at least, it will always work.

If one goes back and looks at the solution **V=A+2s** to the Diophantine Equation , it is clear that p is a function of both A and s. The integer solution for p is thus generated by one of many points along a straight line A=const-2s in the A-s plane. We demonstrate this point by looking at the following graph-



The points in this graph correspond to the same solution of p=83 for the semi-prime N=3403. It shows that one can vary either A or s to find the factors of N. It also indicates the advantage to trying several different starting values of A hopefully finding a value for which a solution to p is found requiring only small number of different s trials. A=61 would be a good choice although we do not really know this beforehand.

Summary of Factoring Procedure:

FACTORING N=pq=(2n-1)(2m-1)

(1)-Expand into a DIOPHANTINE EQUATION-

$$U-2V=K \text{ with } U=nm,\ V=n+m,\ \text{and } K=(N-1)/2$$

(2)-Solve for p and q in terms of V and N-

$$p=(V-1)+\text{sqrt}[(V-1)^2-N] \text{ and } q=(V-1)-\text{sqrt}[(V-1)^2-N]$$

(3)-Diophantine equation has general solution-

$$U=B+s \text{ and } V=A+2s \text{ where } A>1+\text{sqrt}(N) \text{ and } s=0,1,2,3,\ldots$$

(4)-Pick A to have same even or odd character as K and then run Evaluation of p until an s is found leading to an integer value for p. This will be the solution with q =N/p also occurring for the same s.

U.H.Kurzweg
October 25, 2012