

# Notes on Factoring

MA 206

## The General Approach

Suppose I hand you  $n$ , a 200 digit integer and tell you that  $n$  is composite, with smallest prime factor around 50 digits. Finding a nontrivial factor of  $n$  will be extremely difficult. Now suppose I give you another number  $m$ , also 200 digits, and ask you to find a factor. Looks like I've doubled your work. But if I tell you that  $m$  and  $n$  share a common factor, your task is almost trivial: just compute  $(m, n)$  with Euclid's Algorithm, which is very fast. If  $m$  and  $n$  share a divisor then  $(m, n) > 1$  and you've got a proper divisor of  $n$  (unless  $n|m$ ). This is the idea of most factorization methods—to find a factor of  $n$  we try to cook up a second number  $m$  such that  $(m, n) > 1$ .

One common strategy for finding such a factor is to look for solutions to the congruence

$$x^2 \equiv y^2 \pmod{n}. \quad (1)$$

Suppose we find such a solution. Suppose also (as is likely—see below) it turns out that  $x \not\equiv \pm y \pmod{n}$ . Then  $(x-y, n) > 1$  and  $(x+y, n) > 1$ , and so we've found a proper factor of  $n$ , for equation (1) really says that  $n|(x^2 - y^2)$ , or equivalently  $n|(x-y)(x+y)$ . But if  $x \not\equiv \pm y \pmod{n}$  then  $n$  doesn't divide  $x - y$  or  $x + y$ . Therefore  $n = n'n''$  where  $n'|(x - y)$  and  $n''|(x + y)$ . Clearly  $n', n'' > 1$ , and so  $(x - y, n) \geq n' > 1$  and  $(x + y, n) \geq n'' > 1$ .

The methods for factoring described below all look for solutions to equation (1). They don't specifically enforce  $x \not\equiv \pm y \pmod{n}$ , but rather hope that this is true so that  $(x-y, n)$  and/or  $(x+y, n)$  will yield nontrivial factors of  $n$ . What are the odds that this will be true?

**Lemma:** Let  $n$  be odd and composite, with at least two distinct prime factors. For any fixed  $y$ , at least 1/2 of the  $x$  which satisfy  $x^2 \equiv y^2 \pmod{n}$  for  $y \neq 0$  also satisfy  $x \not\equiv \pm y \pmod{n}$ .

**Proof:** We already essentially proved this, when we did “coin flipping by telephone.” But I'll give a quick recap.

If  $n$  is odd and composite with at least two distinct prime factors then we can write  $n = n'n''$  where  $(n', n'') = 1$  and of course both are odd and greater than 1. You can easily check that

$$x^2 \equiv y^2 \pmod{n} \iff x^2 \equiv y^2 \pmod{n'}, \quad x^2 \equiv y^2 \pmod{n''}.$$

(This doesn't rely on  $n$  being odd, just  $(n', n'') = 1$ ). Now consider the congruence  $x^2 \equiv y^2 \pmod{n'}$ . This has AT LEAST two solutions for  $x$ , namely  $x \equiv \pm y$ . Note that because  $n'$  is odd,  $y$  and  $-y$  must be distinct  $\pmod{n'}$ , so this really is two solutions. There may be more. Let  $x_1 \equiv y \pmod{n'}$ . Similar remarks apply to  $x^2 \equiv y^2 \pmod{n''}$ ; let  $x_2 \equiv y \pmod{n''}$ . Now choose integers  $r$  and  $s$  so that  $rn' + sn'' = 1$ . Again, we're using  $(n', n'') = 1$ . Consider the four possible values of  $x$  defined by

$$x = \pm sn''x_1 \pm rn'x_2 \pmod{n}$$

obtained by taking all possible  $\pm$  combinations. You can easily check that all four values are distinct  $\pmod{n}$ , and all satisfy both  $x^2 \equiv y^2 \pmod{n'}$  and  $x^2 \equiv y^2 \pmod{n''}$ , and hence  $x^2 \equiv y^2 \pmod{n}$ . We conclude that  $x^2 \equiv y^2 \pmod{n}$  has AT LEAST 4 solutions, only two of which are  $x \equiv \pm y \pmod{n}$ . This proves the assertion. In the case that  $n = pq$  you can check that there are exactly four solutions (this is exactly what we did in coin flipping by telephone). ■

### Kraitchik's Algorithm

This is best done with an example. Let's factor  $n = 18601$ . Define the polynomial  $Q(x) = x^2 - n$  and let  $x_0 = \lfloor \sqrt{n} \rfloor$ . In this case  $x_0 = 136$ . Let  $x_k = x_0 + k$  and compute  $Q(x_k)$  for a few values of  $k$ . You obtain the following

$k$	$x_k$	$Q(x_k)$
1	137	$168 = (2^3)(3)(7)$
2	138	$443 = (443)$
3	139	$720 = (2^4)(3^2)(5)$
4	140	$999 = (3^3)(37)$
5	141	$1280 = (2^8)(5)$

Now notice that  $Q(139)Q(141) = 2^{12}3^25^2 = (2^6 \cdot 3 \cdot 5)^2 = 960^2$  is a perfect square. Put another way,

$$(139^2 - n)(141^2 - n) = 960^2.$$

Modulo  $n$  this is simply

$$(139 \cdot 141)^2 \equiv 960^2$$

and we have a solution to  $x^2 \equiv y^2 \pmod{n}$ , with  $x \equiv (139)(141) \pmod{n} = 998$  and  $y \equiv 960 \pmod{n} = 960$ . Now compute

$$(18601, 998 + 960) = 979, \quad (18601, 998 - 960) = 19$$

and we have nontrivial factors of  $n$ .

The general procedure for Kraitchik's algorithm is this. Given an integer  $n$  to factor, we let  $Q(x) = x^2 - n$  and  $x_0 = \lfloor \sqrt{n} \rfloor$ . We then consider  $Q(x_k)$  with  $x_k = x_0 + k$  for  $k = 1, 2, \dots$ . We look for products of the  $Q(x_k)$  which form perfect squares. If we find that

$$Q(x_{k_1})Q(x_{k_2}) \cdots Q(x_{k_m}) = y^2$$

then, noting that this is just  $(x_{k_1}^2 - n) \cdots (x_{k_m}^2 - n)$ , it follows that modulo  $n$  we have

$$x_{k_1}^2 \cdots x_{k_m}^2 \equiv y^2,$$

so we have a solution to  $x^2 \equiv y^2 \pmod{n}$  with  $x = x_{k_1} \cdots x_{k_m}$  and  $y = \sqrt{Q(x_{k_1}) \cdots Q(x_{k_m})}$ . We then compute  $(x \pm y, n)$  and hope it provides a nontrivial factor of  $n$ . If it doesn't we look for another combination of the  $Q$ 's that form a perfect square and try again.

Note that there is nothing absolutely necessary about using  $x_k = \lfloor \sqrt{n} \rfloor + k$ ; any choice for integers  $x_k$  can work, but this choice for  $x_k$  has the advantage that  $Q(x_k)$  will be relatively small (compared to  $n$ ) if  $k$  is small (think about it). As a result  $Q(x_k)$  is more likely to factor into smaller primes, making the task of combining the  $Q$ 's to form perfect squares easier.

### The Morrison-Brillhart Algorithm

I'll explain this in two steps. The first obvious difficulty for the above algorithm is that forming a perfect square from the  $Q$ 's appears to be a trial and error method. For a computer program it would be nice to have a more systematic approach. Here is one such approach.

Let us choose a "factor base"  $B$  consisting of small primes. As an example, for  $n = 18601$  we'll choose  $B = [2, 3, 5, 7, 11]$ , the first 5 primes. As before, we let  $x_k = \lfloor \sqrt{n} \rfloor + k$  and we compute  $Q(x_k)$  for  $k = 1, 2, \dots$ . We find  $Q(137) = 168$ . Note that 168 factors over  $B$ , i.e., 168 can be expressed entirely in terms of the primes in  $B$ , as  $168 = (2^3)(3)(7)$ . However,  $Q(138) = 443$  does not factor over  $B$ , so discard it. But  $Q(139) = 720 = (2^4)(3^2)(5)$  factors over

B. Do this for a few values of  $k$  and collect the results in the table below, listing with each value of  $Q(x_k)$  that splits over  $B$  the corresponding power to which each prime appears in the factorization of  $Q(x_k)$ :

$k$	$x_k$	$Q(x_k)$	2	3	5	7	11
1	137	168	3	1	0	1	0
3	139	720	4	2	1	0	0
5	141	1280	8	0	1	0	0
7	143	1848	3	1	0	1	1
13	149	3600	4	2	2	0	0
15	151	4200	3	1	2	1	0

Now look for linear combinations of the rows of the table above which add up to even powers for each prime. For example, the  $k = 3$  and  $k = 5$  rows add to give the “exponent vector”  $[12\ 2\ 2\ 0\ 0]$ , corresponding to a perfect square  $2^{12}3^25^2$ . In fact, the table above encodes more information than needed to look for perfect squares. What we should really do is take the exponents modulo 2, so 0 denotes an even exponent and 1 and odd. In this case we have

$k$	$x_k$	$Q(x_k)$	2	3	5	7	11
1	137	168	1	1	0	1	0
3	139	720	0	0	1	0	0
5	141	1280	0	0	1	0	0
7	143	1848	1	1	0	1	1
13	149	3600	0	0	0	0	0
15	151	4200	1	1	0	1	0

Now we’re looking for linear combinations of the rows which add up to zero (we’re doing all exponent arithmetic mod 2 now). Of course, the  $k = 3$  and  $k = 5$  row add to all zeros. You can also see that the  $k = 13$  row is already a perfect square, for it is all zeros (and 3600 is a perfect square). Are there other combinations of rows which add to all zeros mod 2?

There’s a better way to pose the problem of finding rows which sum to zero. To do so we’re going to appeal to linear algebra. If you’ve studied any linear or matrix algebra, what follows should look familiar. If you haven’t, consider this a motivation to study some! From the more practical point of view, if you haven’t done any linear algebra, you can simply accept the fact that there are systematic ways to find all possible combinations of the rows in the table above which sum to zero mod 2.

Let  $M$  denote the matrix corresponding to the exponent vectors in the table above,

$$M = \begin{bmatrix} 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

We're looking for combinations of the rows which add to zero. In matrix notation, we want to find a row vector  $\mathbf{b} = [b_1 \ b_2 \ b_3 \ b_4 \ b_5 \ b_6]$  (mod 2, so all  $b_j$  are zeros and ones) such that

$$\mathbf{b}M = 0$$

where 0 denotes the appropriate dimensional zero vector. If we take the transpose of the above equation, what we want to find is a column vector  $\mathbf{b}$  so that

$$M^T \mathbf{b} = 0$$

where  $T$  is transpose. In linear algebra terms, we want to find vectors in the nullspace of  $M^T$ . This is a standard linear algebra computation, and we won't go into it here. I'll simply mention that one can use something like Gaussian elimination to find all vectors in (or a basis for) the nullspace.

For our matrix we can compute that the vectors

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

are a basis for the nullspace of  $M^T$ , and so the corresponding products of the appropriate  $Q(k)$  should be perfect squares. The first vector, for example, says that row 5 of the table above, corresponding to 3600, provides a perfect square all by itself. The second vector corresponds to the product of the  $Q$ 's in the second and third rows, and the last vector corresponds to the product of the first and last rows,  $Q(137)Q(151) = 705600 = 840^2$ . From this last combination we conclude that  $(137 \cdot 151)^2 \equiv 840^2 \pmod{n}$  and so compute

$$(137 \cdot 151 + 840, n) = 209$$

to find a proper factor of  $n$ .

Suppose we have  $m$  primes in our factor base. How many  $Q$  that split over the base do we need to accumulate before we can form a perfect square? In other words, how many rows do we need to accumulate in  $M$  (or columns in  $M^T$ ) before  $M^T$  is guaranteed to have a nonempty nullspace? It's a basic fact that this must happen once  $M$  has more rows than columns. In general, if  $M$  has  $k$  more rows than columns then we can expect a minimum of  $k$  vectors in the nullspace, and so we can form  $k$  different perfect squares.

There are a few other issues to consider. We compute  $x_k^2 - n$  for various integers  $x_k$  in the hopes that for some  $p$  in our factor base we have  $p|(x_k^2 - n)$ . This can be rephrased as

$$n \equiv x_k^2 \pmod{p}$$

so that  $n$  is a quadratic residue mod  $p$ . This need not be true. For example, suppose that the last digit of  $n$  is a 3, and that  $p = 5$ . It is impossible to find any integer  $x$  such that  $3 \equiv x^2 \pmod{5}$ , i.e., 3 is not a quadratic residue mod 5. In this case it would be pointless to use 5 in the factor base for  $n$ . More generally, the only primes worth including in the factor base are those for which  $n$  is a quadratic residue. Recall also that we have a simple test for quadratic residues:  $n$  is a quadratic residue mod  $p$  if and only if  $n^{(p-1)/2} \equiv 1 \pmod{p}$ . Thus we construct the factor base by choosing the smallest primes possible subject to this restriction.

How large should the factor base be? The smaller it is, the fewer  $Q$  we need to accumulate before we're guaranteed that we can form perfect squares, but with a smaller factor base it is far less likely that  $Q(x_k)$  will split. On the other hand, a larger base gives more splits, but requires that we find more squares. A relatively straightforward analysis (that I won't give here) shows that a good choice "on average" is to choose the base to have about  $\sqrt{\exp(\sqrt{\log(n) \log(\log(n))})}$  primes.

## Continued Fractions

Recall that if we have a real number  $x_0$  expressed as a continued fraction,

$$x_0 = [a_0, a_1, s_2, \dots]$$

then the successive convergents are rational numbers  $p_k/q_k$ , where

$$\frac{p_k}{q_k} = [a_0, a_1, \dots, a_k].$$

The numbers  $a_k$  can be computed by setting  $a_0 = [x_0]$  and then defining  $x_{k+1} = 1/(x_k - a_k)$ ,  $a_{k+1} = [x_{k+1}]$ . If  $x_0 = \sqrt{n}$  for some number positive integer  $n$  which is not perfect square then it turns out that the  $x_k$  are given by

$$x_k = \frac{P_k + \sqrt{n}}{Q_k}$$

for integers  $P_k$  and  $Q_k$ , where

$$P_{k+1} = a_k Q_k - P_k, \quad Q_{k+1} = \frac{n - P_{k+1}^2}{Q_k}.$$

Most importantly, recall Proposition 2.8 on page 185 of Gibling:  $p_k^2 - nq_k^2 = (-1)^{k+1}Q_{k+1}$ . Let's define  $Q_k^* = (-1)^k Q_k$  so that the proposition becomes  $p_k^2 - nq_k^2 = Q_{k+1}^*$ .

Here's how continued fractions work for factoring: Modulo  $n$  the last equation becomes

$$p_k^2 \equiv Q_{k+1}^* \pmod{n}.$$

Now if we can find some subset of the  $p_k$  such that  $Q_{k_1+1}^* Q_{k_2+1}^* \cdots Q_{k_m+1}^*$  is a perfect square, say  $y^2$ , then we'll have

$$(p_{k_1} p_{k_2} \cdots p_{k_m})^2 \equiv Q_{k_1+1}^* Q_{k_2+1}^* \cdots Q_{k_m+1}^* \pmod{n}$$

or  $x^2 \equiv y^2 \pmod{n}$  where  $x = p_{k_1} p_{k_2} \cdots p_{k_m}$  and  $y = \sqrt{Q_{k_1+1}^* Q_{k_2+1}^* \cdots Q_{k_m+1}^*}$ . We can then compute  $(x \pm y, n)$  and hope it's greater than 1. In essence, the numbers  $p_k$  are playing the role of the  $x_k$  in Kraitchik's algorithm and the  $Q_{k+1}^*$  are playing the role of  $Q(x_k)$ .

So why would we want to replace a simple quadratic polynomial like  $Q(x_k)$  with apparently more complicated continued fraction expansions? Because as was proved in class,  $Q_k < 2\sqrt{n}$  for all  $k$ . Contrast this to  $Q(x_k)$ , which rapidly exceeds  $\sqrt{n}$  and grows large very quickly. Because the  $Q_k$  in the continued fraction expansion stay (relatively) small, they are much more likely to split over the factor base, and so more rapidly provide us with the material we need to form perfect squares.

There are two additional details to consider before doing an example. It is the case that half of the numbers  $Q_k^*$  will be negative. So, while 9 is a perfect square,  $-9$  is not, and we can't just ignore the negative sign. So we don't—we just try to form perfect squares and account for the negative sign in the obvious way—perfect squares have to be positive and so any corresponding

product of  $Q$ 's must have an even number of negative signs. As you'll see below, this amounts to including  $-1$  as a prime in the factor base.

Also, just as with Kraitchik's method, we need should not include in our factor base any primes for which  $n$  is not a quadratic residue. If a prime  $p$  divides  $Q_{k+1}$  then from  $p_k^2 - nq_k^2 = (-1)^{k+1}Q_{k+1}$  we have  $p_k^2 \equiv nq_k^2 \pmod{p}$ . Now  $p$  cannot divide  $q_k$ , for then we have  $p_k^2 \equiv 0 \pmod{p}$ , implying  $p|p_k$ , i.e.,  $p$  would divide both  $p_k$  and  $q_k$ , contradicting the fact that  $(p_k, q_k) = 1$  (this is 1.3(b) on page 179, problem 6 on the test). Thus  $(q_k, p) = 1$  and there exists  $r$  such that  $rq_k \equiv 1 \pmod{p}$ . Multiply both sides of  $p_k^2 \equiv nq_k^2 \pmod{p}$  by  $r^2$  to obtain  $(rp_k)^2 \equiv n \pmod{p}$ , so  $n$  is a quadratic residue mod  $p$ . If  $n$  is not a quadratic residue mod  $p$  then there is no hope of  $p$  dividing  $Q_{k+1}$ .

Let's factor 18601 again as an example. We'll use a factor base consisting of  $B = [2, 3, 5, 7, 11]$ . You can easily check that 18601 is indeed a quadratic residue for each prime. Here's a table that summarizes the computation:

$k$	$p_k$	$Q_{k+1}^*$	$-1$	$2$	$3$	$5$	$7$	$11$
1	273	125	0	0	0	3	0	0
2	409	-128	1	7	0	0	0	0
3	682	99	0	0	2	0	0	1
4	1773	-40	1	3	0	1	0	0
6	5812	-72	1	3	2	0	0	0
7	10155	81	0	0	4	0	0	0
8	17676	-21	1	0	1	0	1	0

Note that  $k = 6$  is missing from the table, since  $Q_7^*$  did not split over the factor base. Note also how small the  $Q_k^*$  are compared to the  $Q(x_k)$  in Kraitchik's method. Amazingly, of the first 8 iterations all but one of the  $Q^*$  split over our base. If we now consider the table mod 2, and write it out as a matrix  $M$  we find that

$$M = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{bmatrix}.$$



Computing a basis for the nullspace of  $M^T$  gives vectors

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$$

The first vector says that the sixth row ( $k = 7$ ) is itself a perfect square, which you can easily see, for  $Q_8^* = 81$ . We then have  $x^2 \equiv y^2 \pmod{n}$  with  $x = 10155$  and  $y = 9$ . We find  $(10155 + 9, n) = 11$ , a proper factor of  $n$ . For the second vector we obtain  $Q_2^* Q_3^* Q_5^* = (125)(-128)(-40) = 640,000 = 800^2$  as a perfect square. This also leads to a nontrivial factorization by computing  $(p_1 p_2 p_4 + 800, n) = 11$ . Finally, the last vector gives  $Q_3^* Q_7^* = 9216 = 96^2$ , leading to  $(p_2 p_6 + 96, n) = 19$ , another factor of  $n$ .

One problem that you might encounter in the continued fraction method for factoring is when the continued fraction expansion of  $\sqrt{n}$  cycles very rapidly. For example, if  $n = d^2 + 1$  for some  $d$  then  $\sqrt{n} = [d, 2d, 2d, 2d, \dots]$ . In this case you'll only get two distinct  $Q^*$ , probably not enough to form any squares. The solution here is to use  $Q^*$ 's that come from the continued fraction expansion of  $\sqrt{\lambda n}$ , where  $\lambda$  is some suitable multiplier. Assuming we generate the  $Q_{k+1}^*$ ,  $p_k$ , and  $q_k$  from  $\sqrt{\lambda n}$ , we then have

$$p_k^2 - \lambda n q_k^2 = (-1)^{k+1} Q_{k+1}.$$

If we look at this mod  $n$  then we still have  $p_k^2 \equiv (-1)^{k+1} Q_{k+1} = Q_{k+1}^*$  and can proceed exactly as before.

## The Quadratic Sieve

The quadratic sieve does not use continued fractions to generate solutions to  $x^2 \equiv y^2 \pmod{n}$ , but rather returns to the polynomial  $Q(x) = x^2 - n$  (or variations of it). If you look back at Kraitchik's method with the addition of the factor base idea, you see that we compute  $Q(x_k)$  for various  $x_k$  and then attempt to split  $Q(x_k)$  over the factor base. The method for doing this is simply to try dividing  $Q(x_k)$  by the first prime in the base, then the second, then the third, etc. When we reach the last prime in the base, if  $Q(x_k)$  has

been reduced to 1 then we conclude that it factors over the base. For large values of  $n$ , however, very few of the  $Q(x_k)$  actually split and so we waste a great deal of time trial dividing  $Q(x_k)$  by primes only to discard all of our work when we reach the last prime in the base and find that  $Q(x_k)$  didn't split.

It would be nice to find a means for more efficiently identifying those  $Q(x_k)$  which will split over our base. There is such method, and it's really a variation on the *sieve of Eratosthenes*. Here's how the traditional sieve of Eratosthenes works. Suppose we want to generate a list of all primes up to 1000. One way is trial division. Take each integer  $k$  in turn and try dividing by all integers less than  $\sqrt{k}$  (or all primes). This is analogous to what Kraitchik's algorithm with the factor base was doing, and it's a lot of work. The sieve of Eratosthenes is far more efficient and works as follows. Create a list or an array contained the integers from 1 up to 1000. Circle or mark 2 to indicate that it's prime. Then step through the array with a stepsize of 2, starting at 4, and strike those elements—they're all multiples of 2, and so not prime. Return to the start and find the next element which has not been eliminated, in this case 3. Circle 3 and then strike every third element from 6 onward—they're all multiples of 3. In general, after having struck all multiples of  $p_{k-1}$  we return to  $p_{k-1}$  and find the next element in the array—this is  $p_k$ . We circle  $p_k$  and then strike every  $p_k$ th element that follows. Notice that we don't do any trial division of the elements as we strike them. If an element appears in the 52nd position then it's a multiple of 13 automatically; we don't trial divide it by 13. Notice also that when we're done striking multiples of  $p = 31$ , our array contains only primes, for every composite less than 1000 has a prime factor less than  $\sqrt{1000}$ , i.e., a prime factor of 31 or less, and so has been struck.

Suppose instead of finding primes less than 1000 we want to find all integers less than 1000 which split over the factor base  $B = [2, 3, 5]$ , that is, integers with only 2's, 3's, and 5's in their prime power factorization. Again, construct an array containing integers from 1 to 1000. We know that only elements with an even index are divisible by 2. Step through the array and consider each such element in turn. For each, remove from it (divide) the highest possible power of 2. Repeat this procedure for 3, dividing every 3rd element by the highest possible power of 3. Repeat for 5. At the end of this procedure any number with only 2, 3 and 5 in its factorization has been reduced to 1; all other elements are greater than 1. The simple idea is that elements which are divisible by  $p$  appear at regular intervals in the array, so

that we need not bother trial dividing any other elements to see if they are divisible by  $p$ . We “sieve” the array by  $p$ . This proves to save an enormous amount of work.

This is the idea behind the quadratic sieve. It turns out that those values of  $Q(x_k) = x_k^2 - n$  which are divisible by a prime  $p$  in the factor base appear at regular intervals with respect to the index  $k$ . We can thus save a lot of work by forming an array of the  $Q(x_k)$  and “sieving” by the primes in the factor base.

Consider an example, with  $n = 57469$ . Suppose that  $p = 11$  is a prime in the factor base. Of course,  $n$  is a quadratic residue mod  $p$ . In fact,  $n \pmod{11} = 5$  and the equation  $x^2 \equiv 5 \pmod{11}$  has two solutions,  $x \equiv 4 \pmod{11}$  and  $x \equiv 7 \pmod{11}$ . Also, in this case  $m_0 = \lfloor \sqrt{n} \rfloor = 239$ , so  $x_k = 239 + k$ . Now if we’re looking for values of  $x$  such that  $p|Q(x)$ , that is,  $x^2 \equiv n \pmod{p}$ , we need only consider  $x$  such that  $x \equiv 4 \pmod{11}$  and  $x \equiv 7 \pmod{11}$ . For  $x_k = 239 + k$  this means we need only consider those values of  $k$  such that  $239 + k \equiv 4 \pmod{11}$  and  $239 + k \equiv 7 \pmod{11}$ , which boil down to  $k \equiv 7 \pmod{11}$  and  $k \equiv 10 \pmod{11}$ . If we form an array, indexed from 1, of the form

$$[Q(x_1) \quad Q(x_2) \quad Q(x_3) \quad Q(x_4) \quad \cdots]$$

then we can divide each element  $Q(x_k)$  with  $k \equiv 7, 10 \pmod{11}$  by the highest possible power of 11. Do the same for the other primes in the factor base. When you’re done, those elements which equal 1 are precisely those which split over the factor base.

The procedure above requires us to solve the equation  $x^2 \equiv n \pmod{p}$  for each prime in the factor base. How do we do this? If  $p$  is of the form  $p = 4k+3$  then we’ve already done it. In the “coin flipping by telephone” material we showed that in this case  $x = \pm n^{k+1} \pmod{p}$  solves  $x^2 \equiv n \pmod{p}$ . If  $p$  is of the form  $p = 4k + 1$  then things are slightly more difficult. We’ll talk about how to solve  $x^2 \equiv n \pmod{p}$  in this case in chapter 11.