

Elementary Factoring methods

From Fermat's Little Theorem and variants, one can fairly quickly show that a composite number is, in fact composite. If we actually want to factor that number, to date, all we know is trial division. In this set of notes, we discuss several methods of factoring numbers other than by trial division. The first method dates back to Fermat, the others are less than 50 years old.

Fermat's method:

This is based on the trick that if $n = x^2 - y^2$, then $n = (x - y)(x + y)$. To try to express n as a difference of two squares, one lets $x_0 = \lfloor \sqrt{n} \rfloor$, and calculates

$$(x_0 + k)^2 - n,$$

for $k = 0, 1, 2, \dots$, stopping when a square is returned. Moreover, since

$$(x_0 + k + 1)^2 - n = [(x_0 + k)^2 - n] + 2(x_0 + k) + 1,$$

we only have to calculate one square. After that, we just add successive odd numbers to the total. For example, if $n = 3977$, $x_0 = 64$ and we have:

k	$x_0 + k$	$2(x_0 + k) + 1$	$(x_0 + k)^2 - n$
0	64	129	119
1	65	131	248
2	66	133	379
3	67	135	512
4	68	137	647
5	69	139	$784 = 28^2$

Thus, $3977 = 69^2 - 28^2 = (69 - 28)(69 + 28) = 41 \cdot 97$.

Fermat's method is good only for small numbers and numbers $n = pq$, where p and q are nearly equal. An example I've come across is $n = 10^{22} + 1$: trial division gives factors 89 and 101. We are left with the 19-digit number

1,112,470,797,641,561,909

to factor. Fermat's method fairly quickly finds

$$1,112,470,797,641,561,909 = (1056689261)(1052788969).$$

How good is Fermat's method? For small numbers, it is a reasonable thing to try. But it is actually worse than trial division in general! The worst case for Fermat's method is the case where n is prime. In this case, n factors as $n \cdot 1$, so we need $x + y = n$, $x - y = 1$. This means that $x = \frac{n+1}{2}$ and $y = \frac{n-1}{2}$. Starting with $x = \sqrt{n}$, this means that we would have to go through $\frac{n+1}{2} - \sqrt{n}$ steps before concluding that n is prime. With trial division, it only takes \sqrt{n} steps to conclude that n is prime. To see the difference, suppose that n is around 10^{10} . This is a very small number, as factoring goes. If n is prime, it will take approximately 10^5 steps with trial division to show this. With Fermat's method, it will take $\frac{1}{2}10^{10} - 10^5$ steps. Thus the difference in the methods is 100,000 steps vs. 4,999,900,000 steps.

On average, one expects a composite number n to have a prime divisor of size $n^{.63}$ and coprime part of size $n^{.37}$. If the coprime part is actually prime, then trial division will take roughly $n^{.37}$ steps to factor n . Fermat's method will take roughly $\frac{1}{2}n^{.63}$ steps. Again, trial division wins. In general, one should never use Fermat's method to completion: you might try for several thousand (several million?) steps hoping to get lucky, but then you should switch to something else.

The next two methods were both devised by a mathematician by the name of John Pollard. They are both considerably better than trial division. However, before using them one should probably check if $2^n \equiv 2 \pmod{n}$. If it is not, we know for sure that n is composite. If the congruence holds, then n is almost certainly prime, and one should probably invest effort into proving that it is prime rather than trying to factor it.

Pollard's rho method (1975):

This method uses an iterated functions approach. Let $f(x) = x^2 + 1$, and consider the sequence $f(0), f(f(0)), f(f(f(0))),$ etc. (mod p). This sequence will be eventually periodic. This means that after a while, it will settle down into a repeating pattern. For example, if $p = 23$, the sequence is 0, 1, 2, 5, 3, 10, 9, 13, 9, 13, \dots . Let $f^m(x)$

represent $f(f(\dots f(x)\dots))$ with m compositions. For any p , there is a k and an m for which $f^m(0) \equiv f^k(0) \pmod{p}$. Once this occurs, $f^{m+1}(0) \equiv f^{k+1}(0)$, $f^{m+2}(0) \equiv f^{k+2}(0) \pmod{p}$, etc. This means that if p is an unknown divisor of n , and if we could find m and k , then we might be able to find p , since p divides

$$\text{GCD}(f^m(0) - f^k(0), n).$$

How can we find m and k when we don't even know p ? We use something called Floyd's Cycle Finding Algorithm. Floyd's cycle finding algorithm works like this: suppose that a_0, a_1, a_2, \dots is eventually periodic. Then $a_m = a_{2m}$ for some m . We can use this to form a factoring algorithm: To factor n , for $k = 1, 2, \dots$ calculate $\text{GCD}(f^{2k}(0) - f^k(0), n)$. In fact, we need only calculate $f^k(0) \pmod{n}$, so the numbers do not get too large. To calculate both f^{2k} and f^k , we really only calculate f^k and keep track of the numbers we found. What we really do is for even k , calculate $f^k - f^{k/2}$. As an example, let $n = 1357$. We have:

f^k	$f^{k/2}$	difference	gcd
1			
2	1	1	1
5			
26	2	24	1
677			
1021	5	153	1
266			
193	26	754	1
611			
147	677	-530	1
1255			
906	1021	-115	23

and so, 23 is a divisor of 1357.

Pollard's p - 1 method (1974):

Recall Fermat's little theorem (again!): for any prime p , and any number a , $a^p \equiv a \pmod{p}$. If p does not divide a , then $a^{p-1} \equiv 1 \pmod{p}$. In particular, if $p > 2$, then $2^{p-1} \equiv 1 \pmod{p}$. If m is a multiple of $p - 1$, say $m = k(p-1)$, then $2^m = (2^{p-1})^k \equiv 1^k \equiv 1 \pmod{p}$. This means that $p|2^m - 1$ for any m where $(p-1)|m$. For example, when $p = 7$, then $p - 1 = 6$. By the above, if m is a multiple of 6, $2^m - 1$ is divisible by 7. So, for example, when $m = 12$, we have $2^{12} - 1 = 4095 = 7 \times 585$.

This can be turned into a factoring method as follows: take a sequence of m 's with lots of small factors (we will use the sequence $m_k = k!$, but other sequences would work.) For each term in the sequence, calculate $\text{GCD}(n, 2^{m_k} - 1)$. Stop when the gcd returns a number > 1 . This method will find a prime divisor p of n , if $p - 1$ divides one of the m_k . This method works very well if $p - 1$ has all small prime divisors.

For example, the maple command `ifactor(1037 - 1, easy);` returns $(3)^2 _c28$ (247629013). What this means is that it found 9 and 247,629,013 as factors of $10^{37} - 1$, leaving a 28-digit number which it knew to be composite. The factor $p = 247,629,013$ was found using the $p - 1$ method. It was successful because

$$p - 1 = (2)^2(3)(37)(41)(61)(223)$$

has all small divisors. In particular, it did NOT find the smaller prime divisor $q = 2,028,119$ because $q - 1 = (2)(37)(27407)$ has a larger prime divisor.

Example: factor $n = 3811$.

k	2	3	4	5	6
$2^{k!} \pmod{3811}$	4	64	1194	2172	3257
$\text{gcd}(2^{k!}-1, 3811)$	1	1	1	1	37

In a real life example, Pollard's method was used to show that $10^{53} - 1$ is divisible by $p = 1325815267337711173$. In fact, this prime was found quickly since

$$p - 1 = 2^2 \cdot 3^2 \cdot 11 \cdot 53 \cdot 1279 \cdot 1553 \cdot 3557 \cdot 8941$$

has all small divisors.