Using Lehmer's factor stencils

Chris Staecker

June 22, 2018

This is a companion to my short video about the factor stencils:

`https://www.youtube.com/watch?v=DiTql8maTiE`

The method that I explain here is described in Lehmer's booklet [1]. I am not a number theorist, so I won't really try to explain why this procedure works, I'll just describe in detail how to do it.

Let N be some big number that we wish to factor. The basic idea is that we do some repeated computations which will tell us which stencils to choose from our set- we continue these computations for as long as it takes to get enough stencils. The stencils are lined up, and the holes which show all the way through indicate possible prime factors of N. If we carry out the computations far enough, we will accumulate enough stencils so that only a few holes are punched through, and these can be checked by hand to determine the factors.

The stencils that I use in the video have holes for the first 200 primes, that is, up to 541. Thus these stencils will suffice to factor any number having at least one prime factor less than or equal to 541, which would include every number up to $541^2 = 292681$. If some but not all factors of N are bigger than 541, the stencils will detect the small ones, and then we can just divide N be the ones we find. Lehmer's original set goes up to 48593, which is the 5000th prime.[1] So Lehmer's original set can be used to factor every number up to $48593^2 = 2361279649$. In practice it will succeed for most larger numbers too, since any large number chosen at random is very unlikely to have all of its factors greater than 48593.

The stencils

The holes punched on the stencils have no discernible pattern when you look at them. Each stencil is labeled with an R value, and the holes are punched according to this rule: a hole at a prime p is punched when R is a *quadratic residue modulo* p. This means that there is some whole number x with $x^2 = R$ mod p, that is, $x^2 - R$ is a multiple of p. Because of the way modular arithmetic

[1] This is by Lehmer's counting. In Lehmer's time, 1 was considered to be a prime number- today we would say that 48593 is the 4999th prime.

works, you can check this formula by brute force by trying all values of x between 0 and p. Square them and subtract R and check if the answer is a multiple of p.

If R is itself a square, then $x^2 = R \mod p$ is satisfied for any p, and so if R is a square then every hole will be punched. This is useless for the stencils, so there are no stencils for square values of R.

My set has a stencil for each (nonsquare) R value from -50 to 50. Lehmer's original set went from -238 to 238. I'm not sure why he chose those values. The fact that I have fewer R values means that somebody using my set may have to do more steps to produce an answer. Any R value greater than 50 won't produce a usable stencil, so we'll just have to skip that step. This won't end the computation, it may just make it go on longer.

The continued fraction of $\sqrt{N}$

The raw data needed to begin the computation is the continued fraction expansion of $\sqrt{N}$. We need to write something like this:

$$\sqrt{N} = q_1 + \cfrac{1}{q_2 + \cfrac{1}{q_3 + \cfrac{1}{q_4 + \ddots}}}$$

where all q_i are whole numbers. This seems complicated but for number theorists in Lehmer's time (and now) it was regarded as a standard procedure. We will demonstrate with an example:

Let's find the continued fraction expansion of $\sqrt{48}$. If you're using the internet, Wolfram Alpha can compute this like so:

`http://www.wolframalpha.com/input/?i=continued+fraction+of+sqrt(48)`

If you have a calculator with a square root button, many steps will be easier. We'll explain how to do it all by hand though, that is, only using addition, subtraction, multiplication, and division.

First we compute q_1, which is the whole number part of $\sqrt{48}$. If you have a calculator, just type $\sqrt{48}$ and you'll see it's $6.928\ldots$, and so $q_1 = 6$. Without a calculator, we observe that $6 < \sqrt{48} < 7$ because we know $6^2 = 36$ and $7^2 = 49$. If N were a very large number we would guess based on the most significant digits, and then check our guesses by squaring.

Now we must compute q_2. This is the whole number part of the number x_2 which satisfies:

$$\sqrt{48} = 6 + \frac{1}{x_2}.$$

Solve this for x_2:

$$x_2 = \frac{1}{\sqrt{48} - 6}.$$

At this point if you have a calculator with a square root button, you can just compute the right side, and q_2 is the whole number part of the answer. To do it without using a calculator, “rationalize” by multiplying the fraction on top and bottom by $\sqrt{48}+6$:

$$x_2 = \frac{\sqrt{48}+6}{48-36} = \frac{\sqrt{48}+6}{12}.$$

Now $\sqrt{48}$ is more than 6 but less than 7, so the numerator here is more than 12 but less than 13. Since the denominator is 12, this means that x_2 is more than 1 but less than 2. So the whole number part of x_2 is 1, so $q_2 = 1$.

Now for q_3, this is the whole number part of the number x_3 which satisfies:

$$\sqrt{48} = 6 + \frac{1}{1+\frac{1}{x_3}}.$$

Actually using what we did already, we see that $x_2 = 1 + \frac{1}{x_3}$, so:

$$\frac{\sqrt{48}+6}{12} = 1 + \frac{1}{x_3},$$

and solving for x_3 we get:

$$x_3 = \frac{12}{\sqrt{48}-6}.$$

Rationalize again:

$$x_3 = \frac{12(\sqrt{48}+6)}{48-36} = \sqrt{48}+6$$

We got lucky here and the denominators canceled. Since $\sqrt{48}$ is between 6 and 7, the whole number part of x_3 is 12, so $q_3 = 12$.

So far we have computed:

$$\sqrt{48} = 6 + \cfrac{1}{1 + \cfrac{1}{12 + \cfrac{1}{?}}}$$

We can continue this procedure in the same way for as many steps as we need.

The recurrences

Once we have the terms of the continued fraction, we are ready to begin the real computation. We must compute terms of this system of recurrences:

$$\begin{aligned} P_n &= q_{n-1}Q_{n-1} - P_{n-1} \\ Q_n &= q_{n-1}(P_{n-1} - P_n) + Q_{n-2} \end{aligned}$$

using these initial values:

$$\begin{aligned} P_1 &= 0, & Q_1 &= 1 \\ P_2 &= q_1, & Q_2 &= N - q_1^2 \end{aligned}$$

This part is easy and boring. Let's do a few steps using the example from the video, $N = 189121$. For this example, the continued fraction expansion of $\sqrt{N}$ is:

$$\sqrt{189121} = [434; 1, 7, 2, 1, 2, 1, 13, \dots]$$

So $q_1 = 434$, $q_2 = 1$, $q_3 = 7$, etc.

We begin with the initial conditions $P_1 = 0$, $Q_1 = 1$, $P_2 = 434$, $Q_2 = 189121 - 434^2 = 765$.

For step $n = 3$ we plug in to the recurrences and we get:

$$\begin{aligned} P_3 &= q_2 Q_2 - P_2 = 1 \cdot 765 - 434 = 331 \\ Q_3 &= q_2(P_2 - P_3) + Q_1 = 1 \cdot (434 - 331) + 1 = 104 \end{aligned}$$

For further steps, just keep going. For $n = 4$ we get:

$$\begin{aligned} P_4 &= q_3 Q_3 - P_3 = 7 \cdot 104 - 331 = 397 \\ Q_4 &= q_3(P_3 - P_4) + Q_2 = 7(331 - 397) + 765 = 303 \end{aligned}$$

For $n = 5$ we get:

$$\begin{aligned} P_5 &= q_4 Q_4 - P_4 = 2 \cdot 303 - 397 = 209 \\ Q_5 &= q_4(P_4 - P_5) + Q_3 = 2(397 - 209) + 104 = 480 \end{aligned}$$

etc. You can continue here for as many steps as you need.

The values of the Q's are how you choose which stencils to line up (you don't use the P's for anything, but you need them to compute the Q's). The rule is: Take each Q value and factor it. Remove any squares from the factorization, and multiply by -1 when n is even, and the result is the R value.

From $Q_1 = 1$ we factor and remove squares and there's nothing left, so we don't get an R value in that step. From $Q_2 = 765$ we factor it and get $Q_2 = 5 \cdot 17$. There are no squares here so we obtain $R = -765$ (remember multiply by -1 for even steps). We would now grab the stencil labeled $R = -765$, but this is too big for our set so we'll just have to continue without it.

From $Q_3 = 104 = 2^3 \cdot 13$ we remove squares and end up with $R = 2 \cdot 13 = 26$, so we get the stencil for $R = 26$.

From $Q_4 = 303 = 3 \cdot 101$ there are no squares so we get $R = -303$. This is also too big so we're out of luck in this step.

From $Q_5 = 480 = 2^3 \cdot 3 \cdot 5$ we get $R = 2 \cdot 3 \cdot 5 = 30$, so we grab the stencil labeled $R = 30$.

Lining up these stencils, every hole punched all the way through is a possible divisor of N. Since we've only found 2 stencils so far, we'll still see many punched holes. It is up to the computer (you) to decide when to stop. At any time, you can just start testing divisors by long division. At this point our two stencils still show 19 holes through, so if you stop now you'd have to do 19 different long divisions to test which are the true divisors. If that seems like too many, you can compute more Q's and try to find more stencils.

In the video, I continued the computation until there were only 2 divisors showing through the stencils. This required me to go out to Q_{10}. For this example, the Q's and P's work out as follows:

n	P	Q	R
1	0	1	
2	434	765	$-765 = -5 \cdot 17$
3	331	104	$26 = 2 \cdot 13$
4	397	303	$303 = 3 \cdot 101$
5	209	480	$30 = 2 \cdot 3 \cdot 5$
6	271	241	-241
7	211	600	$6 = 2 \cdot 3$
8	389	63	-7
9	430	67	67
10	374	735	$-15 = -3 \cdot 5$

Taking these values, with my stencil set we can line up the stencils for R values: 26, 30, 6, -7, and -15. If I had Lehmer's larger set of stencils, I could've also used the stencil for $R = 67$. One more helpful thing, which I mentioned in the video: when we find one R value which is a multiple of another, then we can also use the quotient as an R value. Thus after step 6 when we have $R = 30$ and $R = 6$, we can also use the stencil for $R = 5$, since $30 \div 6 = 5$. Once we have $R = 5$, we can also use $R = -17$ because step 2 gave us $R = -5 \cdot 17$. When we get to step 10 we get $R = -15$, and this gives us $R = -2$ for free because we already had $R = 30$. These extra R values give us more stencils to line up which will speed everything up significantly.

In this example, after these 10 steps, lining up all the stencils shows only 2 holes at 379 and 499, and these are the factors of N.

In conclusion...

Congratulations- you can now factor large numbers using the fastest procedure known in the 1930s. In [1], Lehmer writes:

> The computation of the successive Q's is very well adapted to a computing machine. With a little practice it is a simple matter to carry out the expansion for some fifty or one hundred terms in half an hour. For a number of the order of two billion or less, fifty terms in the expansion will usually supply one with a sufficient number of small residues to factor the number.

If I was working in the 1930s, I also would've tried to use a machine to do the arithmetic for computing the recurrences. Machines at the time were built for adding. A machine like the Monroe Model L[2] can do repeated additions to accomplish multiplication, but it is a procedure that must be learned (there

[2] See my video about the Monroe Model L: https://www.youtube.com/watch?v=OCNqWY_kNZ8

is no "multiply button"). A machine like the Felt & Tarrant Comptometer[3] can multiply even faster, but would require serious training to use reliably. For a non-expert operator, like me, performing these multiplications on a contemporary machine would introduce many chances for mistakes. A slide rule can multiply, but is not appropriate here because we require the answers to be exact in every step.

Lehmer concludes: "In fact after long experience with this method of finding quadratic residues I am confident that no better method has so far been found." Obviously, this is no longer true.

References

[1] Lehmer, Derrick Norman, Factor Stencils. Carnegie institution of Washington, 1939.

[3] My video about multiplying and dividing with a Comptometer: `https://www.youtube.com/watch?v=AlGbblf8deg`