# Prime Factorization and Factor Range Estimation

Henry Samuelson
hes227@cornell.edu

May 2019-March 2020

## 1 Introduction

Let: $q, p \in$ prime. Let: $N = qp$. Fermat's Factorization states:

$$N = (a + b)(a - b) \tag{1}$$

Unless $\sqrt{N} \in \mathbb{N}$, then $q > \lceil \sqrt{N} \rceil$, and $p < \lceil \sqrt{N} \rceil$. Hence we define:

$$\begin{aligned} a &= \lceil \sqrt{N} \rceil \\ N &= (\lceil \sqrt{N} \rceil + b)(\lceil \sqrt{N} \rceil - b) \end{aligned} \tag{2}$$

This works if the difference of the perfect square above $N$, $\lceil \sqrt{N} \rceil^2$, and $N$ is also square.

$$\sqrt{\lceil \sqrt{N} \rceil^2 - N} \in \mathbb{N} \tag{3}$$

To account for situations where the difference isn't square we can add an $k$ to make this always true.

$$\sqrt{(\lceil \sqrt{N} \rceil + k)^2 - N} \in \mathbb{N} \tag{4}$$

The $k$ insures that the square root will be in $\mathbb{Z}$. Hence we can adapt the earlier equation too:

$$N = (\lceil \sqrt{N} \rceil + k + b)(\lceil \sqrt{N} \rceil + k - b) \tag{5}$$

For all cases where the difference between $\sqrt{\lceil \sqrt{N} \rceil^2 - N} \in \mathbb{N}$ we assume $k = 0$. For non-zero $k$'s, the complexity is $NP$ hard, whereas when $k = 0$ the equation can be solved with basic algebra.

## 2 Determining b

We first define some rules for k and b. It is clear that $b > 0$ as $b = 0$ would mean $q = p$. We make the assumption that $k < b$. Then we can determine a relation between $q, p, b$.

$$\frac{q-p}{2} = b \tag{6}$$

This equation can be shown to be objectively true if the purpose of $b$ is thought of correctly. If $b$ is the distance from some middle point $(\lceil\sqrt{N}\rceil + k)^2$ between $q$ and $p$ then $b$ must be half of $q - p$, allowing $b$ to be added and subtracted in either direction, to find $q$ and $p$.

This next definition of $b$ is less obvious but crucial to defining a range for $b$.

$$b = \sqrt{(\lceil\sqrt{N}\rceil + k)^2 - N} \tag{7}$$

It turns out that equation (4), the one we want to solve for an integer to determine the correct $k$ is actually $b$. Now that we have to equations for $b$ we can eliminate b, and derive a direct relationship between $k, q, p$.

$$q - p = 2 * \sqrt{(\lceil\sqrt{N}\rceil + k)^2 - N} \tag{8}$$

This equation tells allot about the relation between $q, p$ and $k$, as we now have a solid equation for determining spacing which is very helpful in deriving the bounds of $b, k, q$ and $p$.

## 3 Determining Variable Bounds

These variable bounds are only true if we assume $k \neq 0$, as we can assume if $k = 0$, then $q - p$ must $= 2$, and it would be very algebraically simple. We can first work to determine bounds for k. As stated earlier $k < b$, this can function as our top bound, $k_{max}$. The top bound of $b_{max} = \lceil\sqrt{N}\rceil$ There is a very import relationship between the growth rate of $b$ and $k$. $b$ grows at a faster rate than $k$, which is given by eq (7). If we assume $b_{max}$, we can determine $k_{min}$.We know from the definition of $p$, that $p = (\lceil\sqrt{N}\rceil + k - b)$. Plugging in $b_{max}$ yields $k = 3$.

$$\begin{aligned}
&p = (\lceil\sqrt{N}\rceil + k_{min} - b_{max}) \geq 3 \\
&p = (\lceil\sqrt{N}\rceil + k_{min} - \lceil\sqrt{N}\rceil) \geq 3 \\
&k_{min} \geq 3
\end{aligned} \tag{9}$$

If $b_{max} = \lceil\sqrt{N}\rceil$ and $k < b$, we can use the relationship between $k$ and $b$ to calculate $k_{max}$. The larger the $b$ the larger the $k$. According to eq (7) we can solve using $b_{max}$ to yield $k_{max}$.

$$\begin{aligned} b_{max} &= \left\lceil \sqrt{(\lceil\sqrt{N}\rceil + k_{max})^2 - N} \right\rceil = \lceil\sqrt{N}\rceil \\ k_{max} &= \left\lceil \sqrt{b_{max}^2 + N} \right\rceil - \lceil\sqrt{N}\rceil \\ k_{max} &= \left\lceil \sqrt{\lceil\sqrt{N}\rceil^2 + N} \right\rceil - \lceil\sqrt{N}\rceil \end{aligned} \tag{10}$$

Now we have $k_{max}$ directly in terms of $N$. This is what we need to determine a final bound. We can use $k_{min} \geq 3$ to help us solve the bottom bound for $b$, $b_{min}$.

$$\begin{aligned} b_{min} &= \left\lceil \sqrt{(\lceil\sqrt{N}\rceil + k_{min})^2 - N} \right\rceil \\ b_{min} &= \left\lceil \sqrt{(\lceil\sqrt{N}\rceil + 3)^2 - N} \right\rceil \end{aligned} \tag{11}$$

Now we have the top and bottom bounds for both $b$ and $k$ we can can rewrite $b$ and $k$ as,

$$\begin{gathered} 3 \leq k \leq \left\lceil \sqrt{\lceil\sqrt{N}\rceil^2 + N} \right\rceil - \lceil\sqrt{N}\rceil \\ \left\lceil \sqrt{(\lceil\sqrt{N}\rceil + 3)^2 - N} \right\rceil \leq b \leq \lceil\sqrt{N}\rceil \end{gathered} \tag{12}$$

Solid $b$ and $k$ bounds allow us to now determine bounds for $q$ and $p$. We will acknowledge the obvious but important relationships,

$$\frac{N}{q_{min}} = p_{max}, \ \frac{N}{p_{min}} = q_{max} \tag{13}$$

This is helpful, because calculating $q_{max}$ and $q_{min}$ is easy, whereas one cannot calculate $p_{min}$ and $p_{max}$ using the standard definitions of $q$ and $p$, due to the definitions of p including a $-$ sign.

$$\begin{aligned} &q = (\lceil\sqrt{N}\rceil + k + b) \\ &q_{max} = (\lceil\sqrt{N}\rceil + k_{min} + b_{max}) = (2\lceil\sqrt{N}\rceil + 3) \\ &q_{min} = (\lceil\sqrt{N}\rceil + k_{min} + b_{min}) = (\lceil\sqrt{N}\rceil + 3 + \left\lceil \sqrt{(\lceil\sqrt{N}\rceil + 3)^2 - N} \right\rceil) \end{aligned} \tag{14}$$

The bounds for $q$ are complete along with the bounds of $p$ using eq (13),

$$\begin{gathered}\left\lceil\sqrt{N}\right\rceil+3+\left\lceil\sqrt{(\left\lceil\sqrt{N}\right\rceil+3)^2-N}\right\rceil \leq q \leq 2\left\lceil\sqrt{N}\right\rceil+3 \\ \frac{N}{2\left\lceil\sqrt{N}\right\rceil+3} \leq p \leq \frac{N}{\left\lceil\sqrt{N}\right\rceil+3+\left\lceil\sqrt{(\left\lceil\sqrt{N}\right\rceil+3)^2-N}\right\rceil}\end{gathered} \tag{15}$$

For example the for $N = 2231 = qp = (97)(23)$, $k = 12, b = 37$, the estimated bounds are as follows:

$$\begin{aligned} 3 &\leq k \leq 20 \\ 20 &\leq b \leq 48 \\ 71 &\leq q \leq 99 \\ 23 &\leq p \leq 32 \end{aligned} \tag{16}$$

These bounds are quite good.

# 4 Solving For k

We can rewrite the b relation equation– eq(8)– to be in terms of k to determine how many $k$'s we have to brute force directly in order to deterine the factors of $N$, $p$ and $q$.

$$k_{actual} = \frac{1}{2}(q + p - 2\left\lceil\sqrt{N}\right\rceil) \tag{17}$$

This means that the number of $k$'s we have to guess is directly dependent upon the distance between $p$ and $q$, and the actual value of $N$. Since we can calculate the bottom bound of $k$ we can subtract it from the number of steps it takes to solve to calculate a new complexity.

$$k_{guesses} = \frac{1}{2}(q + p - 2\left\lceil\sqrt{N}\right\rceil) - 3 \tag{18}$$

Given the equation it would appear to make $N$ as resistant as possible to brute forcing $k$'s, the best thing to do would be to maximize the first half of the equation by maximizing both $q$ and $p$. And then to minimize the second half of the equation by making $N$ or $q * p$ smaller. This means that there is an optimal ratio that exists that maximizes $q + p$ while minimizing $q * p$. This may seem counter intuitive at first, as it is commonly thought that a larger $N$ is better, but it is really only better when $q - p$ and $q + p$ are larger.

# 5 A Direct Function

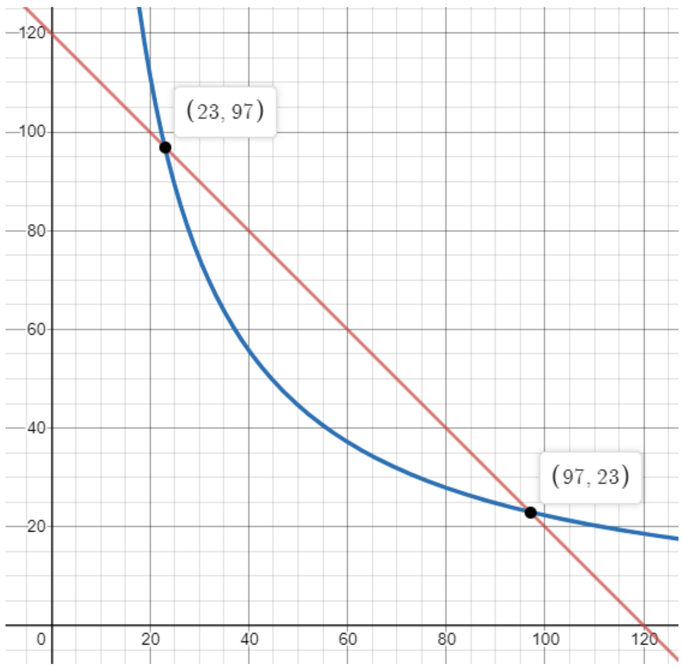
We have an equation for $N$, eq. (5), we have an equation for $b$ in terms of $k$ eq. (7). Now we also have an equation for k, ($k_{actual}$), eq. (17). Eq. 7 and 17 can be plugged into eq. 5.

$$
\begin{aligned}
&N = (\left\lceil\sqrt{N}\right\rceil + k + b)(\left\lceil\sqrt{N}\right\rceil + k - b)\\
&N = (\left\lceil\sqrt{N}\right\rceil + k + (\sqrt{(\left\lceil\sqrt{N}\right\rceil + k)^2 - N}))(\left\lceil\sqrt{N}\right\rceil + k - (\sqrt{(\left\lceil\sqrt{N}\right\rceil + k)^2 - N}))\\
&N = (\left\lceil\sqrt{N}\right\rceil + \frac{1}{2}(q + p - 2\left\lceil\sqrt{N}\right\rceil) + (\sqrt{(\left\lceil\sqrt{N}\right\rceil + \frac{1}{2}(q + p - 2\left\lceil\sqrt{N}\right\rceil))^2 - N}))\\
&(\left\lceil\sqrt{N}\right\rceil + \frac{1}{2}(q + p - 2\left\lceil\sqrt{N}\right\rceil) - (\sqrt{(\left\lceil\sqrt{N}\right\rceil + \frac{1}{2}(q + p - 2\left\lceil\sqrt{N}\right\rceil))^2 - N}))
\end{aligned} \tag{19}
$$

Simplify,

$$
\begin{aligned}
&N = (\frac{1}{2}(\lceil q\rceil + \lceil p\rceil) + \sqrt{(\frac{1}{2}(\lceil q\rceil + \lceil p\rceil))^2 - N})(\frac{1}{2}(\lceil q\rceil + \lceil p\rceil) - \sqrt{(\frac{1}{2}(\lceil q\rceil + \lceil p\rceil))^2 - N}))\\
&\text{Let: } J = \frac{1}{2}(\lceil q\rceil + \lceil p\rceil)\\
&N = J^2 - \left\lceil\sqrt{J^2 - N}\right\rceil^2
\end{aligned} \tag{20}
$$

The final simplification with $J$ can be confusing. Eq. (19) is a linear line that is equivalent to $y = -x + (q + p)$, where $x$ and $y$ are possible $q$, $p$ values, and where $(q + p)$ is some constant. This means that at all points on the line the $x$ and $y$ values add to $(q + p)$. It is also true that this line must intersect the line $y = x$, due to its slope being $-x$. Then one can substitute $x$ for $y$. Plugging in $x$ for $y$, yields just $J = x$, so we can consider $J$ as a variable with no definition except solving fo $J$ will be solving for the case in which $x = y$. This will not be the actual answer we need for $x$ and $y$, but it will tell us what $(q + p)$ is. If this is known we can factorize our number as we know $qp = N$ and $y = -x + (q + p)$, as there is only two free variables since $(q + p)$ and $N$ are known..There will be two intersections of the lines, with coordinates $(q, p)$, and $(p, q)$. As shown bellow, the red line is $2231 = J^2 - \left\lceil\sqrt{J^2 - N}\right\rceil^2$, where $J$ has a definition in terms of $q$ and $p$, and the blue line is, $qp = 2231$. Their intersections are the prime factors of $N$, in this case 2231.




## 6 Substitution With A Totient

The totient of the factor of two primes can be defined as $(p-1)(q-1) = \phi$. It can also be rewritten as $N - q - p + 1 = \phi$. This last identity is very powerful. Thinking about the totient in terms of $k$ and $b$ we can define it as,

$$\phi = (\left\lceil \sqrt{N} \right\rceil - 1 + k + b)(\left\lceil \sqrt{N} \right\rceil - 1 + k - b) \tag{21}$$

This is reasonable as $\phi = (p-1)(q-1)$. Plugging into the definition of $b$ from eq. (7) we get the powerful relation,

$$\begin{aligned} &b = \sqrt{(48+k)^2 - N} = \sqrt{(47+k)^2 - \phi} \\ &\text{Simplify,} \\ &N - \phi = 2\left\lceil \sqrt{N} \right\rceil - 1 + 2k \end{aligned} \tag{22}$$

Substituting for $N - \phi$ will lead to the previously derived definition of $k = (q+p)/2 - \left\lceil \sqrt{N} \right\rceil$. We now have the relation,

$$N - \phi = 2\left\lceil \sqrt{N} \right\rceil - 1 + 2k = q + p - 1 \tag{23}$$

We now write a solid definition of $k$ in terms of $\phi$,

$$k = \frac{N - \phi + 1 - 2\left\lceil \sqrt{N} \right\rceil}{2} \tag{24}$$

Plugging the definition of $k$ into $b$ will help us seek further insights.

$$\begin{aligned} b &= \sqrt{\left(\frac{N-\phi+1-2\left\lceil\sqrt{N}\right\rceil}{2}+\left\lceil\sqrt{N}\right\rceil\right)^2-N} \\ b &= \sqrt{\left(\frac{N-\phi}{2}+\frac{1}{2}\right)^2-N} \\ b &= \sqrt{\left(\frac{N-\phi+1}{2}\right)^2-N} \end{aligned} \tag{25}$$

$k$ can be rewritten,

$$\begin{aligned} k &= \frac{N-\phi+1-2\left\lceil\sqrt{N}\right\rceil}{2} \\ k &= \left(\frac{N-\phi+1}{2}\right)-\left\lceil\sqrt{N}\right\rceil \end{aligned} \tag{26}$$

Let $J$ be defined by, $J=\frac{N-\phi+1}{2}$. Then,

$$\begin{aligned} b &= \sqrt{J^2-N} \\ k &= J-\left\lceil\sqrt{N}\right\rceil \end{aligned} \tag{27}$$

This is the same $J$ that we found in 'A Direct Function'. We prove this by plugging definitions of $b$ and $k$ in terms of $J$ into eq.(5).

$$\begin{aligned} N &= (\left\lceil\sqrt{N}\right\rceil+k+b)(\left\lceil\sqrt{N}\right\rceil+k-b) \\ N &= (J+\left\lceil\sqrt{J^2-N}\right\rceil)(J-\left\lceil\sqrt{J^2-N}\right\rceil) \\ N &= J^2-\left\lceil\sqrt{J^2-N}\right\rceil^2 \end{aligned} \tag{28}$$

# 7 b division attack

If $b \mod k = 0$ or $k=\frac{b}{D}$ where, $D \in \mathbb{N}$ and is unknown; then the factorization of $N = qp$ is insecure, and can be exploited. Equation (5) can be written to have $k$ in terms of $b$.

$$\begin{aligned} N &= (\left\lceil\sqrt{N}\right\rceil+k+b)(\left\lceil\sqrt{N}\right\rceil+k-b) \\ N &= (\left\lceil\sqrt{N}\right\rceil+\frac{b}{D})+b)(\left\lceil\sqrt{N}\right\rceil+\frac{b}{D}-b) \end{aligned} \tag{29}$$

The equation can be solved for $b$ where $b \in \mathbb{N}$. The equation for $b$ in terms of $D$ is:

$$b = \frac{\sqrt{D^4\left\lceil\sqrt{N}\right\rceil^2-D^4N+D^2N}+D\left\lceil\sqrt{N}\right\rceil}{D^2-1} \tag{30}$$

This equation makes a lot of sense as $b > k$ which means $D > 1$. This holds true as seen in the denominator of (9). Solving for a $b \in \mathbb{Z}$, yields the correct solution for both $b$ and $D$. We can simplify the operations to guess the correct $D$. We can break up the definition of $b$ into three distinct integer parts, the numerator in the square root, the numerator, and the denominator. Assuming they are all integers we can determine a simplification for determining $b$.

$$\begin{aligned}
&A, B, C \in \mathbb{N} \\
&\frac{\sqrt{A} + B}{C} \\
&\sqrt{A} \notin \mathbb{N}, \text{then}, \\
&\sqrt{A} + B \notin \mathbb{N} \\
&\frac{(\sqrt{A} \notin \mathbb{N}) + B}{C} \notin \mathbb{N}
\end{aligned} \tag{31}$$

(10) shows that $b \in \mathbb{N}$ is entirely dependent upon, the contents of the square root being square. So we can now instead solve for an integer solution for:

$$\sqrt{D^4 \left\lceil \sqrt{N} \right\rceil^2 - D^4 N + D^2 N} \in \mathbb{N} \tag{32}$$

After finding an integer solution for (11), we can plug the values of $b$ and $D$ back into (8).

# 8 Example

$$\begin{aligned}
&N = qp = 101 * 23 = 2323 \\
&\text{Assume, } D = 2 \\
&\sqrt{D^4 \left\lceil \sqrt{N} \right\rceil^2 - D^4 N + D^2 N} = \\
&\sqrt{2^4 \left\lceil \sqrt{2323} \right\rceil^2 - 2^4 * 2323 + 2^2 * 2323} = \sqrt{10540} \\
&\sqrt{10540} \notin \mathbb{N} \text{ So, } D = D + 1 \\
&\sqrt{3^4 \left\lceil \sqrt{2323} \right\rceil^2 - 3^4 * 2323 + 3^2 * 2323} = \sqrt{27225} \\
&\sqrt{27225} = 165 \in \mathbb{N}
\end{aligned} \tag{33}$$

Though we know the contents of the square root are square, there is still a chance that given our estimate for $D$ that $b \notin \mathbb{N}$. So we must now calculate all of $b$ and confirm it is an integer using (9).

$$b = \frac{\sqrt{27225} + D\left\lceil\sqrt{N}\right\rceil}{D^2 - 1}$$
$$b = \frac{\sqrt{27225} + 3 * 49}{3^2 - 1}$$
$$b = 39 \in \mathbb{N}$$

We now plug $b$ and $D$ into (8).

$$N = (\left\lceil\sqrt{2323}\right\rceil + \frac{39}{3}) + 39)(\left\lceil\sqrt{2323}\right\rceil + \frac{39}{3} - 39)$$
$$N = (101)(23) \tag{34}$$

# 9 Conclusion

Though there are good rules put in place to insure that $\sqrt{(\left\lceil\sqrt{N}\right\rceil + k)^2 - N} \in \mathbb{Z}$, there isn't proper rules in place to insure that that $b \mod k \neq 0$, which allows for the b division attack.