# A FACTORIZATION ALGORITHM BASED ON COHN'S IRREDUCIBILITY CRITERION

EOIN MACKALL

Abstract. We present an algorithm which, for an integer $n > 1$ as input, outputs either a proper divisor $d$ of $n$ or determines that $n$ is prime. On the one hand, this algorithm iterates through all integers from 2 to the square root of $n$, similar to trial division, and so requires $\mathcal{O}(\sqrt{n} \cdot \text{Poly}(\log(n)))$ bit operations in a worst case. On the other hand, for composite $n$, this algorithm is likely to terminate in less steps than needed in trial division.

## 1. Introduction

The problem of either finding an efficient factorization algorithm for large integers or, showing such an algorithm doesn't exist, is difficult. The large computational cost of all currently known factorization algorithms forms the basis for the security of the RSA cryptographic primitive, which is in widespread use today. Some of the best known algorithms for finding a proper divisor of a composite integer, e.g. the Elliptic Curve Method [Len87] or the General Number Field Sieve, are known to run with subexponential time complexity.

Trial division of an integer $n$, the process of dividing $n$ by $2, 3, ...$ until reaching a proper divisor, runs with exponential time complexity in the number of bits of $n$ in a worst case. For a small random integer $n$, trial division is often still one of the fastest methods to factor $n$, due to the speed with which division can be accomplished. However, when the smallest prime divisor of $n$ is large, trial division is unbearably slow on most modern commercially available computers.

We introduce in this paper an algorithm for finding a proper divisor of an integer $n$ which works in a similar way to trial division. For each integer $b = 2, 3, ...$, we describe a process which can potentially produce a proper divisor of $n$ at step $b$ and which will certainly terminate at a proper divisor $b$ of $n$. The idea is based on Cohn's Irreducibility Criterion for integer polynomials. Namely, for each integer $2 \leq b < \sqrt{n}$

*Date*: April 1, 2025.

2020 *Mathematics Subject Classification.* 11Y16.

*Key words and phrases.* integer factorization; polynomial factorization.

we form a polynomial whose coefficients are the digits of the base $b$ representation of $n$. We explain how ideas due to Murty [RM02] show that if this polynomial factors, then we have produced a divisor of $n$.

Our algorithm frequently terminates at an integer $b$ much less than the smallest prime factor of $n$. In order to analyze the effective difference between our algorithm and trial division, we introduce arithmetic functions which essentially count the number of "factoring bases" for a given integer. We then compare one of these functions to the standard divisor counting function on various sets of integers. We conjecture that this latter arithmetic functions of interest has an average order significantly greater than the divisor function (see Conjecture 3.5 for a precise statement) and we show that this arithmetic function is unbounded on the class of semiprime integers.

## 2. Cohn's Irreducibility Criterion

Let $n > 0$ be an integer. Given a second integer $b > 0$, we write $n = (b_r b_{r-1} \cdots b_1 b_0)_b$ to denote the base $b$ representation of $n$, so there is an equality

$$n = b_r \cdot b^r + b_{r-1} \cdot b^{r-1} + \cdots + b_1 \cdot b + b_0$$

for integers $0 \leq b_0, ..., b_r < b$.

**Definition 2.1.** Let $\Phi : \mathbb{N} \times \mathbb{Z}_{>1} \to \mathbb{Z}[x]$ be the function defined as follows: for any tuple $(n, b)$ we have

$$\Phi(n, b)(x) = b_r x^r + b_{r-1} x^{r-1} + \cdots + b_1 x + b_0$$

where $n = (b_r b_{r-1} \cdots b_1 b_0)_b$ is the base $b$ representation of $n$.

Cohn's Irreducibility Criterion, in its original form, is the statement that $\Phi(n, 10)(x)$ is irreducible if $n$ is prime. More recently, the phrase *Cohn's Irreducibility Criterion* may also refer to the following theorem, generalizing this original statement, due to Murty [RM02, Theorem 2].

**Theorem 2.2.** *Let $n > 0$ be an integer and let $b \geq 2$ be a second integer. If $n$ is prime, then $\Phi(n, b)(x)$ is irreducible.*

While the above theorem immediately implies the statement that if $\Phi(n, b)(x)$ is reducible then $n$ is composite, a careful reading of the proof of the theorem actually shows the stronger statement that any proper polynomial factor of $\Phi(n, b)(x)$ gives a proper divisor of $n$ on evaluation at $b$. Stated precisely:

**Theorem 2.3.** *Let $n > 0$ be an integer and let $b \geq 2$ be a second integer. Suppose that there is a factorization*

$$\Phi(n, b)(x) = f(x)g(x) \qquad \deg(f), \deg(g) < \deg \Phi(n, b)(x)$$

*for integer polynomials $f(x)$ and $g(x)$. Then $1 < f(b) < n$ is a proper divisor of $n$.*

*Proof.* We explain how the results of [RM02] allow one to deduce this result directly. Since $\Phi(n,b)(x)$ is nonnegative for any $x \geq 0$ (and in fact $\Phi(n,b)(x) = 0$ if and only if both $x = 0$ and $b$ divides $n$), we can assume that both $f(x)$ and $g(x)$ are also nonnegative for any $x \geq 0$. Since also $f(b)g(b) = \Phi(n,b)(b) = n$, we're reduced to showing that an equality $f(b) = 1$, for $f(x)$ as above, is impossible.

By [RM02, Lemma 2], any root $\alpha \in \mathbb{C}$ of $\Phi(n,b)(x)$ satisfies

$$|\alpha| < \frac{1 + \sqrt{1 + 4(b-1)}}{2}.$$

We can write

$$f(x) = c\prod_{i=1}^{m}(x - \alpha_i)$$

for the leading coefficient $c \in \mathbb{Z}$ and $\alpha_i$ a subset of the roots of $\Phi(n,b)(x)$. For any integer $b > 2$, we have

$$|\alpha_i| < \frac{1 + \sqrt{1 + 4(b-1)}}{2} \leq b - 1$$

so that $|b - \alpha_i| > 1$. In particular, for all $b > 2$, this implies $f(b) > 1$.

For the case when $b = 2$, a more precise bound and a more accurate argument is needed. In [RM02, Lemma 3] it's shown that any root $\alpha \in \mathbb{C}$ of a polynomial such as $\Phi(n,2)(x)$ satisfies $\mathrm{Re}(\alpha) < 3/2$. Thus

$$|f(2)| = c\prod_{i=1}^{m}|2 - \alpha_i| > c\prod_{i=1}^{m}|1 - \alpha_i| = |f(1)|.$$

But $f(1) \in \mathbb{Z}$ is positive, so $|f(2)| > |f(1)| \geq 1$. □

## 3. An Algorithm For Integer Factorization

Theorem 2.3 suggests the following Algorithm 1 for either finding a proper divisor of an integer $n > 1$ or concluding that $n$ is prime.

For each integer $2 \leq b < \lfloor\sqrt{n}\rfloor$, the coefficients of $\Phi(n,b)(x)$ can be calculated using at most $\mathcal{O}(\log(n)^3)$ bit operations. Further, since the polynomial $\Phi(n,b)(x)$ has degree $\lfloor\log_b(n)\rfloor$, it can be decided whether or not $\Phi(n,b)(x)$ is irreducible, and if $\Phi(n,b)(x)$ is found to be reducible then $\Phi(n,b)(x)$ can subsequently be factored, in at most $\mathcal{O}(\mathrm{Poly}(\log(n)))$ bit operations, see [LLL82]. Thus, the simple algorithm described has time complexity as stated in the abstract.

**Algorithm 1** A factorization algorithm for $n \in \mathbb{Z}$ and $n > 1$

```
function Is_irreducible(f(x))
    if f(x) is irreducible then
        return true
    else
        return false
    end if
end function

function Poly_factor(f(x))
    if Is_irreducible(f(x))== false then
        return g(x) dividing f(x) with deg(g) < deg(f)
    else
        return f(x)
    end if
end function

for b = 2 to b = ⌊√n⌋ do
    if Is_irreducible(Φ(n, b)(x))==false then
        g(x) ← Poly_factor(Φ(n, b)(x))
        return g(b)
    end if
end for
return n is prime
```

**Corollary 3.1.** *The above algorithm correctly returns either a divisor of $n > 1$ or returns that $n$ is prime using at most $\mathcal{O}(\sqrt{n} \cdot \mathrm{Poly}(\log(n)))$ bit operations.*

*Proof.* The algorithm runs at most $\sqrt{n}$ iterations. Each iteration uses $\mathcal{O}(\mathrm{Poly}(\log(n)))$ bit operations. To see that it terminates correctly, note that a composite integer $n$ has a divisor $d \leq \sqrt{n}$ and $\Phi(n, d)(x)$ is reducible for such $d$. □

**Remark 3.2.** The algorithm above can be improved in efficiency for the range between $\lfloor \sqrt[5]{n} \rfloor + 1$ and $\lfloor \sqrt{n} \rfloor$ using explicit formulas for the roots of such polynomials. For example, if $n > 1$ and $b \geq 2$ are integers such that $\Phi(n, b)(x) = cx^2 + dx + e$ for integers $c, d, e > 0$ then $\Phi(n, b)(x)$ factors in $\mathbb{Z}$ if and only if both $d^2 - ce = z^2$ is a square of an integer $z \geq 0$ and if $d \equiv z \pmod 2$.

The main interest in the above algorithm is that it often terminates at an iterate $2 \leq k \leq \lfloor \sqrt{n} \rfloor$ where $k$ is not a proper divisor of $n$.

**Example 3.3.** Let $n = 207314063849$. Then

$$n = 207314063849 = 323131 \cdot 641579$$

is a prime factorization of $n$. But

$$\Phi(207314063849, 40098) = 128x^2+37640x+25817 = (8x+347)(16x+11).$$

Evaluating at $x = 40098$ yields $323131 = 8(40098) + 347$.

To quantify the number of successes obtained through the above algorithm, we introduce the following arithmetic functions.

**Definition 3.4.** For any integer $n \geq 2$, define the following set

$$B(n) = \{b \in \mathbb{Z} : 1 < b \leq n,\ \Phi(n, b)(x) \text{ is not primitive or is reducible}\}.$$

Then define $\beta(n) = \#B(n)$ to be the cardinality of this set. Similarly, for any $k \in \mathbb{N}$ define $(\tau_{\geq k}\beta)(n)$ to be the function with

$$(\tau_{\geq k}\beta)(n) = \#\left(B(n) \cap \{1, ..., \lfloor \sqrt[k]{n} \rfloor\}\right).$$

It follows immediately from the definition that $\beta(n) \geq \tau(n) - 1$. We include a plot, in Figure 1 below, of points $(n, \tau_{\geq 2}\beta(n))$ ranging over all integers $2 \leq n < 771055$. For comparison we include a plot, Figure 2 below, of the points $(n, \tau(n))$ for the divisor counting function $\tau$ over the same interval of integers.

A line plot of the cumulative average for the function $\tau_{\geq 2}\beta$ on the integers $2 \leq n < 771055$ is given in Figure 3. The graphic is suggestive of the following conjecture:

**Conjecture 3.5.** *There exists constants $c, \varepsilon > 0$ such that the average order of $\tau_{\geq 2}\beta$ is given as*

$$\sum_{n \leq x} \tau_{\geq 2}\beta(n) = cx^{1+\varepsilon} + o(x^{1+\varepsilon}).$$

We don't know how one might go about proving Conjecture 3.5. Instead, we focus the remainder of this paper on analyzing the behavior of the function $\tau_{\geq 2}\beta$ on semiprimes, i.e. integers $n$ which factor $n = pq$ as a product of two primes $p, q$. These integers are interesting not only due to their practical use in RSA encryption schemes but, because they form the class of composite integers where $\tau$ takes the minimal value 4. In contrast, Figure 4, which displays points $(n, \tau_{\geq 2}\beta(n))$ over a sample of 16695 semiprimes $n$ of size no greater than $10^{12}$, shows that $\tau_{\geq 2}\beta$ is very much nonconstant on semiprimes.

Heuristically, the function $\tau_{\geq 2}\beta$ seems to take on a small value at a semiprime $n = pq$ whenever $p$ and $q$ are greatly differing in size. Conversely, experimental evidence suggests that $\tau_{\geq 2}\beta$ takes on a larger

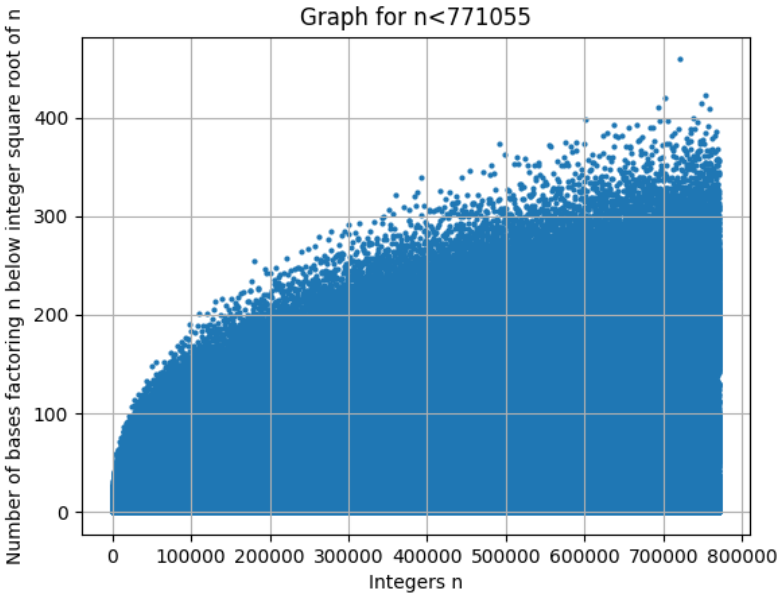


FIGURE 1. A plot of values for $(\tau_{\geq 2}\beta)(n)$ for all integers $n$ with $2 \leq n < 771055$.

value at a semiprime $n = pq$ whenever $p$ and $q$ are similar in size, $p \approx q$. In the latter case, we can make this evidence into a precise result.

**Example 3.6.** For the semiprime

$$n = 30674101 = 331 \cdot 92671$$

we have $\tau_{\geq 2}\beta(n) = 1$ so that $\Phi(n, b)(x)$ either factors or $\Phi(n, b)(x)$ is not primitive for $2 \leq b < \sqrt{n}$ if and only if $b = 331$.

On the other hand, for the semiprime

$$m = 250000070000004899 = 500000069 \cdot 500000071$$

we have $\tau_{\geq 2}\beta(m) = 56706$.

**Theorem 3.7.** *Let $\mathcal{S} \subset \mathbb{N}$ denote the set of semiprime integers. Then*

$$\sup\{\tau_{\geq 2}\beta(s) : s \in \mathcal{S}\} = \infty.$$

*Proof.* We use the fact [Zha14] that there exists an integer $k > 1$ and an infinite sequence of primes $p_1 < p_2, p_3 < p_4, \ldots$ satisfying the inequalities $p_{2i} - p_{2i-1} < k$ for all integers $i \geq 1$. We set $n_i = p_{2i-1}p_{2i}$ and we set $j_i = \lfloor p_{2i-1}/2 \rfloor + 1$. Note that for any integer $p_{2i-1} > b \geq j_i + \frac{k}{2}$ we have

$$p_{2i} = b + m_{2i}(b) \quad \text{and} \quad p_{2i-1} = b + m_{2i-1}(b)$$

for integers $1 \leq m_{2i}(b), m_{2i-1}(b) < b$ with $m_{2i}(b) - m_{2i-1}(b) < k$.

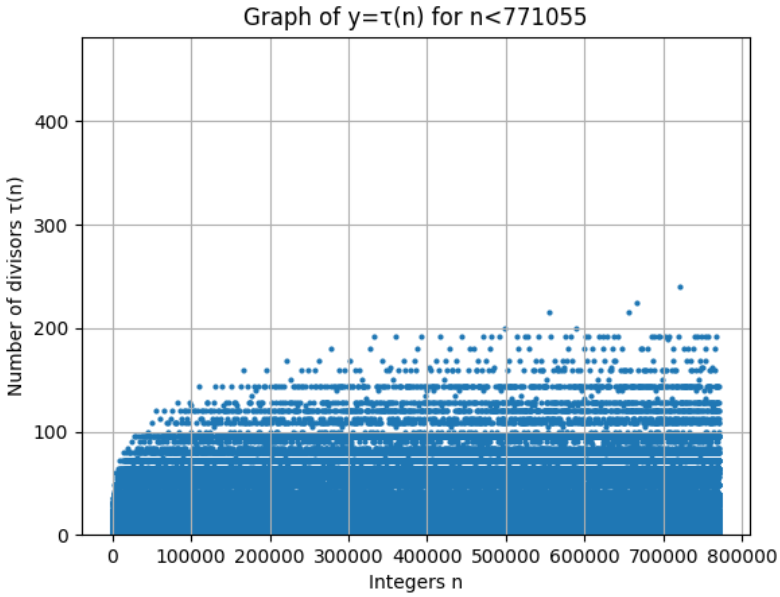


FIGURE 2. A plot of values for the divisor function $\tau(n)$ for integers $2 \leq n < 771055$.

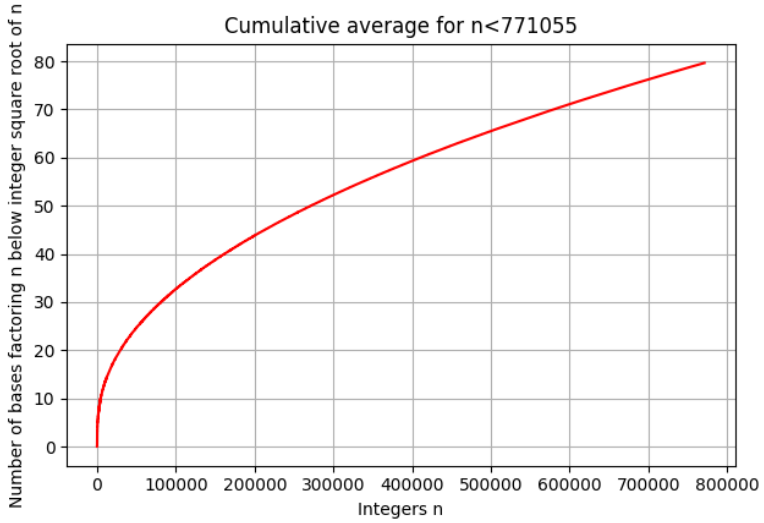


FIGURE 3. A plot of the cumulative average for the function $(\tau_{\geq 2}\beta)(n)$ for all integers $n$ with $2 \leq n < 771055$.

It follows from the above that there is an equality

$$\begin{aligned} n_i = p_{2i-1}p_{2i} &= (b + m_{2i-1}(b))(b + m_{2i}(b)) \\ &= b^2 + (m_{2i-1}(b) + m_{2i}(b))b + m_{2i-1}(b)m_{2i}(b). \end{aligned} \tag{1}$$

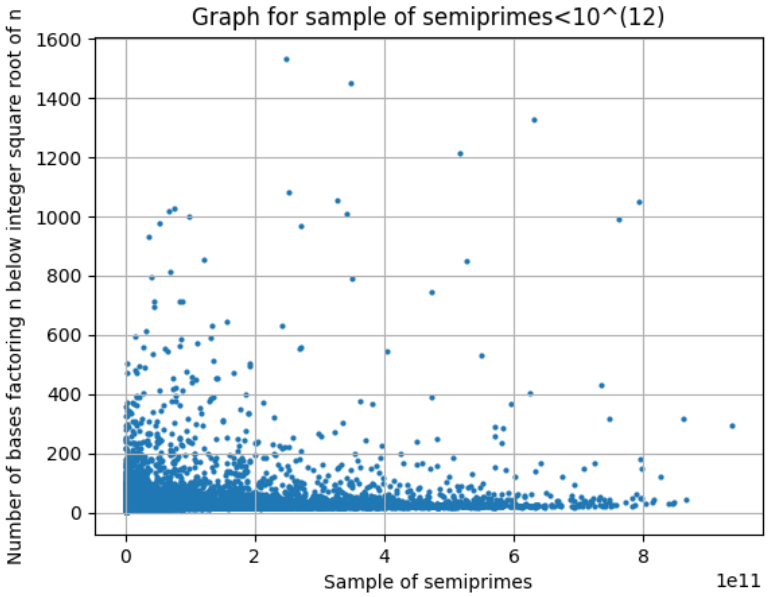


FIGURE 4. A plot of values for $(\tau_{\geq 2}\beta)(n)$ over a sample of 16695 semiprimes less than $10^{12}$.

For any $N \in \mathbb{N}$, we may choose $i \gg 0$ sufficiently large so that for all $b$ with $k+1 \leq m_{2i}(b) < N$ we have

$$m_{2i-1}(b) + m_{2i}(b) < j_i \quad \text{and} \quad m_{2i-1}(b)m_{2i}(b) < j_i.$$

Hence the equation (1) yields a factorization of $\Phi(n_i, b)(x)$ for all such integers $b$. It follows that $\tau_{\geq 2}\beta(n_i) \geq N - k - 2$ for $i \gg 0$ large enough. Since $N$ was arbitrary, this implies the claim. □

## REFERENCES

[Len87] H. W. Lenstra, Jr., *Factoring integers with elliptic curves*, Ann. of Math. (2) **126** (1987), no. 3, 649–673. MR 916721

[LLL82] A. K. Lenstra, H. W. Lenstra, Jr., and L. Lovász, *Factoring polynomials with rational coefficients*, Math. Ann. **261** (1982), no. 4, 515–534. MR 682664

[RM02] M. Ram Murty, *Prime numbers and irreducible polynomials*, Amer. Math. Monthly **109** (2002), no. 5, 452–458. MR 1901498

[Zha14] Yitang Zhang, *Bounded gaps between primes*, Ann. of Math. (2) **179** (2014), no. 3, 1121–1174. MR 3171761

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CALIFORNIA SANTA CRUZ, SANTA CRUZ, CA USA

*Email address*: eoinmackall@gmail.com