Integer Factorization Methods

Trial division, Pollard's \( p - 1 \),
Pollard's \( \rho \), and Fermat's method

Christopher Koch\(^1\)

\(^1\)Department of Computer Science and Engineering
CSE489/589 Algorithms in CS & IT
New Mexico Tech

April 8, 2014
Overview

- Intro to modular arithmetic
- Euler’s theorem and Fermat’s little theorem
- Trial division
- Pollard’s $p-1$ method
- Cycles in $\mathbb{Z}/n\mathbb{Z}$
- Floyd’s cycle-finding algorithm
- Pollard’s $\rho$ method (Monte Carlo factorization)
- Birthday paradox
- Fermat’s method

Convention

$a, b, c, d, m, n$ are integers, $p, q$ are primes

Monte Carlo method: dependent on some kind of random number/function/process
Modular Arithmetic

- $a | b$ (a divides b) if $b$ is a multiple of $a$.
- Quotient and remainder unique in integer division
- Congruence modulo $n$:

$$a \equiv b \pmod{n} \text{ iff } n | (a - b).$$

Division Algorithm

Division Algorithm: Given $a, b$ there exist unique $q, r$

$$a = bq + r \quad \text{where} \quad 0 \leq r < b.$$

For example, $13 \equiv 8 \equiv 3 \pmod{5}$.

Intuition: $a$ and $b$ have same remainder when divided by $n$. 

Cost of Multiplication and GCD

Integer Factorization

- Trial Division
- Pollard’s $\rho - 1$
- Cycles in $\mathbb{Z}/n\mathbb{Z}$
- Floyd’s cycle-finding
- Pollard’s $\rho$
- Birthday paradox
- Fermat’s method
Residue classes

- Congruence modulo $n$ is an equivalence relation on integers.
- Equivalence classes: one for each remainder
  $$[a]_n = \{ x : x \equiv a \pmod{n} \}.$$
- Called residue classes mod $n$
Integers modulo \( n \)

- Integers modulo \( n \): set of residue classes mod \( n \):
  \[
  \mathbb{Z}/n\mathbb{Z} = \{ [r]_n : r \in \mathbb{Z} \} .
  \]

- How to do arithmetic in mod \( n \)? What is \([3]_4 + [1]_4\)?

For example, \( \mathbb{Z}/4\mathbb{Z} = \{ [0]_4, [1]_4, [2]_4, [3]_4 \} \)

- 24-hour clock
- Intuition: notion of “finite discreteness”
- Like \( b \)-bit (unsigned) integers – \( \mathbb{Z}/2^b\mathbb{Z} \) – overflow “wraps” back around (most general-purpose architectures don’t do saturation arithmetic, so it wraps around like modular arithmetic)
- But also, \([5]_4 \in \mathbb{Z}/4\mathbb{Z}\). Why?
- Because \([5]_4 = [1]_4\), since 5 \( \equiv 1 \) (mod 4).
Arithmetic mod $n$

**Definition**

Let $n \in \mathbb{Z}^+$ and $a, b \in \mathbb{Z}$. Then,

$$ [a]_n + [b]_n = [a + b]_n $$

$$ [a]_n \times [b]_n = [a \times b]_n $$

- Similarly,

$$ [a]_n - [b]_n = [a]_n + [-b]_n = [a - b]_n. $$

- Example here: 24-hour clock, adding 12 and 25, and multiplying 13 by 2
- Proof that well-defined: Pick $x \in [a]_n$, $y \in [b]_n$ different representatives, show $x + y \in [a + b]_n$ since $x \equiv a \pmod{n}$ and $y \equiv b \pmod{n}$ implies $x + y \equiv a + b \pmod{n}.$
GCD and Totatives

- \( \gcd(a, b) \) is the greatest common divisor of \( a \) and \( b \)
- \( a, b \) are called coprime or relatively prime if \( \gcd(a, b) = 1 \). \( a \) is called a totative of \( b \) and vice versa.
- Bézout’s identity: If \( \gcd(n, m) = d \), then there exist \( k, l \) s.t. \( nk + ml = d \).
- \( \varphi(n) \) counts the number totatives less than \( n \):
  \[ \varphi(n) = |\{ c : 1 \leq c < n \text{ and } \gcd(c, n) = 1 \}|. \]
- We have \( \varphi(mn) = \varphi(n)\varphi(m) \).
Inverses mod $n$

- Notice: no division in mod $n$!
- Division is usually defined as multiplication by the multiplicative inverse.
- Multiplicative inverse of $[a]_n$ is $[b]_n$ such that $[a]_n[b]_n = [1]_n$; i.e. $ab \equiv 1 \pmod{n}$.

Example here: $2 \in \mathbb{Z}/4\mathbb{Z}$ and its inverse? $4 \in \mathbb{Z}/7\mathbb{Z}$ and its inverse?
Theorem

\([a]_n \in \mathbb{Z}/n\mathbb{Z} \text{ has a multiplicative inverse if and only if} \]
\[\gcd(a, n) = 1.\]

- Drawing from previous example: \(\gcd(4, 2) = 2\), while \(\gcd(4, 7) = 1\).
- That means that every element except 0 in \(\mathbb{Z}/p\mathbb{Z}\) has an inverse, since a prime is coprime to every element below it.
- Bézout’s identity again: \(\gcd(m, n) = 1\), then \(m[m^{-1}]_n + n[n^{-1}]_m = 1\).

- Proof: see write-up of notes on intro to mod arithmetic
Euler’s and Fermat’s Theorems

Theorem (Euler, Euler totient, Euler-Fermat)

Let $a, n$ be coprime. Then,

$$a^\varphi(n) \equiv 1 \pmod{n}.$$

Corollary (Fermat)

Unless $a$ is a multiple of $p$,

$$a^{p-1} \equiv 1 \pmod{p}.$$

- Leonhard Euler, 1707-1783. Euler-Fermat Theorem. Fermat’s little theorem was the origin, Euler gave proof for it in 1736 and kept working until Euler-Fermat was in its final form in 1761.
- E54 (Theorematum Quorundam ad Numeros Primos Spectantium Demonstratio) proves Fermat’s – by induction on binomial series expansion $(1 + 1)^p - 1$ etc / induction as we know it today (he bashes on Fermat’s notion of induction and how he disproved Fermat’s conjecture that $2^{2^n} + 1$ is prime for any $n$)
- Modern proof: using LaGrange’s theorem (subgroup $H$ of $G$), $o(H) \mid o(G)$, let $H = (\mathbb{Z}/n\mathbb{Z})^*$ and $G = \mathbb{Z}/n\mathbb{Z}$; and since $a^{o(H)} \equiv 1$ for $a \in H$
- Carmichael’s theorem: $a^{\lambda(n)} \equiv 1 \pmod{n}$ where $\lambda(n) = 0.5\varphi(n)$ if $n$ is a power of 2 and all other $\lambda(n) = \varphi(n)$.
- Pierre de Fermat, 1601-1665, Fermat’s last theorem $a^n + b^n \neq c^n$ for $n \geq 2$.
- Application to RSA: Let $M$ be a message. $M^{m\varphi(n)+1} \equiv M \pmod{n}$. So, let’s find $ed = m\varphi(n) + 1$, i.e. $ed \equiv 1 \pmod{\varphi(n)}$.
- $C(M) = M^e \pmod{n}$, $D(M) = M^d \pmod{n}$. Then, $C(D(M)) = M^{de} \equiv M^{m\varphi(n)+1} \equiv M \pmod{n}$
- Choose twoprime $n = pq$ for easy $\varphi$ computation
Cost of Multiplication and GCD

Convection
We will denote the cost of multiplication by $M(n)$ and the cost of the GCD by $G(n)$ for $n$-digit numbers.

- Schoolbook multiplication: $M(n) \in O(n^2)$.
- Schönhage-Strassen: $M(n) \in O(n \lg n \lg \lg n)$.
- Euclidean GCD: $G(n) \in O(n^2)$.
- Schönhage’s GCD: $G(n) \in O(M(n) \lg n)$.
- Modular exponentiation ($a^k \mod b$): $O(M(c) \lg k)$, where $c = \max(\lg a, \lg b)$.
Theorem (Fundamental Theorem of Arithmetic)

Let $n$ be an integer. Then there exist unique primes $p_1, p_2, \ldots, p_k$ not necessarily distinct such that

$$n = p_1 \times p_2 \times \cdots \times p_k.$$

- In essence, every integer can be factored uniquely into primes. For example, $20 = 2 \times 2 \times 5$.
- FTA guarantees existence of that factorization, but how do you find it?

Convention

In the following slides, every big $O$ is given in terms of input values instead of input length.

FTA first written down and proved by Euclid in his treatise Elements, 300BC (differently, of course)

Proof: existence by induction, base case $n = 2$ / uniqueness: Euclid’s lemma or elementary using the smallest integer that is product of two different prime factorizations

Euclid’s lemma: $p | ab$ implies $p | a$ or $p | b$.

Erdős-Kac Theorem: $\omega(n)$ number of distinct prime factors of $n$, then the probability distribution of

$$\frac{\omega(n) - \log \log n}{\sqrt{\log \log n}}$$

is the standard normal distribution.
Trial Division

1: TrialDivision\(n\)
2: \(D \leftarrow ()\)
3: \textbf{for all} \(p\) \textbf{in} PRIMES\((\sqrt{n})\) \textbf{do}
4: \quad \textbf{while} \(n \mod p = 0\) \textbf{do}
5: \quad \quad \textbf{APPEND}(D, p)
6: \quad \textbf{n} \leftarrow n/p
7: \quad \textbf{if} \(n > 1\) \textbf{then}
8: \quad \quad \textbf{APPEND}(D, n)
9: \textbf{return} \(D\)

- How often does for-loop execute?
- Prime-counting function \(\pi(m)\).
- How often does while execute? In total, at most \(\log_p(n) \leq \lg n\) (since \(\lg 2 \leq \lg p\) for all \(p \geq 1\)).

Loop executes \(\pi(\sqrt{n})\).
Trial Division: Analysis

Theorem (Prime number theorem)

\[
\lim_{x \to \infty} \frac{\pi(x)}{x / \ln(x)} = 1.
\]

This implies \( \pi(x) \in O\left(\frac{x}{\ln x}\right) \).

Then, for an integer \( n \) to be factored, trial division is

\[
O\left(\pi\left(\sqrt{n}\right) \log(n) M(\log(n))\right) = O\left(\sqrt{n} M(\log(n))\right).
\]

- Landau notation (Bachman) in terms of limits: \( f(n) \in O(g(n)) \) if
  \[
  \lim_{n \to \infty} \frac{|f(n)|}{g(n)} < \infty.
  \]
- PNT conjectured by Gauß in 1792 by his own account at 15 years old (Gauß, Hubert Mania)
- PNT first proved by Hadamard and Vallée-Poussin (1896)
- TODO / MAYBE: talk about distribution of prime factors, the \( O \) of the first prime factor being found and last factor being found. Might be useful for following sections.
Pollard's $p - 1$ method

1: \textsc{PollardP-1}(n, B)
2: \hspace{1cm} K \leftarrow \prod_{p \leq B} p^{[\log_p(n)]}
3: \hspace{1cm} m \leftarrow (2^K - 1) \mod n \triangleright \text{modular exponentiation}
4: \hspace{1cm} g \leftarrow \gcd(m, n)
5: \hspace{1cm} \textbf{if } g = 1 \textbf{ then}
6: \hspace{1.5cm} \textbf{either increase } B \textbf{ and}
7: \hspace{1.5cm} \textbf{return } \textsc{PollardP-1}(n, B)
8: \hspace{1cm} \textbf{or return } \text{failure}
9: \hspace{1cm} \textbf{else}
10: \hspace{1.5cm} \textbf{return } g \triangleright g \text{ must be a divisor of } n

- Periodicals of the Cambridge Phil Society, 1974, Theorems on Primality Testing and Factorization
- and Lenstra's ECM paper and MIT Elliptic Curves Spring 2013 course
- $n$ to be factored, finds \textbf{non-trivial divisor} of $n$
- $B$ smoothness bound
- Since $m < n$, $g < n$
- Original method detailed a second step to be taken for deterministic factorization: $L < M < n$, $M < L^2$.
- $b \equiv a^m \mod n$ where $m$ is product of primes less than or equal $L$ to some power. Find $d = \gcd(b - 1, n)$. If $d = n$, decrease $L$ and repeat.
- Step 2: For primes $L < p < M$, $F_p = (b^p - 1) \mod n$, find $\gcd(F_p, n)$. 

\begin{enumerate}
\item \hspace{1cm} John M. Pollard, 1974.
\item \hspace{1cm} Periodicals of the Cambridge Phil Society, 1974, Theorems on Primality Testing and Factorization
\item \hspace{1cm} and Lenstra's ECM paper and MIT Elliptic Curves Spring 2013 course
\item \hspace{1cm} $n$ to be factored, finds \textbf{non-trivial divisor} of $n$
\item \hspace{1cm} $B$ smoothness bound
\item \hspace{1cm} Since $m < n$, $g < n$
\item \hspace{1cm} Original method detailed a second step to be taken for deterministic factorization: $L < M < n$, $M < L^2$.
\item \hspace{1cm} $b \equiv a^m \mod n$ where $m$ is product of primes less than or equal $L$ to some power. Find $d = \gcd(b - 1, n)$. If $d = n$, decrease $L$ and repeat.
\item \hspace{1cm} Step 2: For primes $L < p < M$, $F_p = (b^p - 1) \mod n$, find $\gcd(F_p, n)$.
\end{enumerate}
Pollard’s $p - 1$: Why does it work?

Corollary (Fermat’s little theorem)

For $a < p$, $a^{p-1} \equiv 1 \pmod{p}$. That is, $p | (a^{p-1} - 1)$.

- Assume $p$ is a prime divisor of $n$.
- That means that $\gcd(a^{p-1} - 1, n) \geq p$.
- The preceding also works if the exponent is a multiple of some $p - 1$, i.e. $a^K - 1$ where $K$ is a multiple of $p - 1$.
- Goal: choose $K$ such that it is likely to be the multiple of some $p - 1$ for a prime divisor $p$.

I.e. we have found a non-trivial divisor.

$p - 1$ only finds factors $p$ that are of the form

$$p - 1 = A$$

where $A$ is some product of primes less than $B$. ($p - 1$ must be $B$-(power)smooth.)

- Original version: $p - 1 = Ag$ for some $L < q < M$.
- Example 1: $n = 540143$, choose $B = 8$ and $K = \text{lcm}(2, \ldots B) = 840$ (easier). Then, $2^K \mod n = 53047$ and $\gcd(2^K - 1, n) = 421$. Then, $540143 = 421 \times 1283$.
- Example 2: $n = 491389 = 383 \times 1283$. Because $383 - 1 = 2 \times 191$, we have $191|K$ and $B \geq 191$ (for lcm). INFEASIBLE.
Pollard’s $p - 1$: Analysis

The exp and modular exp can be combined:

1: $K \leftarrow 2$
2: for all $p$ in PRIMES($B$) do
3: \hspace{1em} $pc \leftarrow p$
4: \hspace{1em} while $pc < n$ do
5: \hspace{2em} $K \leftarrow KP \pmod{n}$
6: \hspace{1em} $pc \leftarrow pc \cdot p$
7: $g \leftarrow \gcd(K - 1, n)$

- $\sum_p \left\lfloor \log_p(n) \right\rfloor$ multiplications and mod exps.
- Each mod exp is $O(\lg(p)M(\lg n))$
- Each mult $M(\lg n)$.
- Then, $\sum_p \log_p(n) \log(p)M(\lg n) = \sum_p \lg(n)M(\lg n)$
Pollard’s $p - 1$: Analysis

- $\sum_p \lceil \log_p(n) \rceil$ multiplications and mod exps.
- Each mod exp is $O(\lg(p) M(\lg n))$
- Each mult $M(\lg n)$.
- Then, $\sum_p \log_p(n) \lg(p) M(\lg n) = \sum_p \lg(n) M(\lg n)$
- Then, we have

$$O(G(\lg n) + \pi(B) \lg(n) M(\lg n)).$$

Then, complexity of one iteration of Pollard’s $p - 1$ is

$$O(\pi(B) \lg(n) M(\lg n)).$$
Cycles in $\mathbb{Z}/n\mathbb{Z}$

Definition

A sequence $\{X_i\}_{i \geq 0}$ is considered periodic if there exists $a$ such that $X_{m+a} = X_m$ for all $m \geq 0$

- Ultimately periodic if for all $m \geq M$ (some starting value)

- Periodic: $1, 2, 3, 1, 2, 3, 1, 2, 3, \ldots$
- Ultimately periodic: $3, 2, 4, 5, 1, 4, 5, 1, 4, 5, 1, \ldots$
- Since $\mathbb{Z}/n\mathbb{Z}$ “wraps around” in a sense, we can find cycles in it.
- $a$ is called period.
Overview

Modular Arithmetic

Division Algorithm and Congruence

Residue classes mod n

Integers modulo n

Arithmetic with integers mod n

GCD and Totatives

Inverses mod n

Euler's Theorem

Cost of Multiplication and GCD

Integer Factorization

Trial Division

Pollard's $p - 1$

Cycles in $\mathbb{Z}/n\mathbb{Z}$

Floyd's cycle-finding

Pollard's $\rho$

Birthday Paradox

Fermat's method

For example, let $n = 5$, $f(x) = (x + 4) \mod 5$ and $x_0 = 1$.

$1, 0, 4, 3, 2, 1, 0, \ldots$

Base: $n = \mu$. Then, $X_{n+\lambda} = X_n = X_\mu$.

Assume $X_{n+\lambda} = X_n$ for all some $n$

Then, $X_{n+1+\lambda} = f(X_{n+\lambda}) = f(X_n) = X_{n+1}$.

Example: $n = 19$, $f(x) = (x^2 - 1) \mod n$, $x_0 = 2$.

Sequence: 2, 3, 8, 6, 16 -> 8 DRAW RHO

Then, $\exists n$ s.t. $X_{2n} = X_n$ (let $n = \lambda$)
Floyd’s cycle-finding algorithm

**Input:** function $f$ and start-value $x_0$

1: `FLOYD_CYCLE(f, x_0)`
2: \[ x \leftarrow f(x_0), y \leftarrow f(f(x_0)) \]
3: \[ \text{while } x \neq y \text{ do} \]
4: \[ x \leftarrow f(x) \]
5: \[ y \leftarrow f(f(y)) \]

- Think of tortoise and hare going in circles: one slow, one fast; eventually they cross
- same with hour-hand and minute-hand of a clock
Pollard's $\rho$ method

1: PollardRho($f, n$)
2: \hspace{1em} $x \leftarrow 2, y \leftarrow 2, g \leftarrow 1$
3: \hspace{1em} while $g = 1$ do
4: \hspace{2em} $x \leftarrow f(x)$  \hspace{0.5em} $\triangleright$ Pollard used $f(x) = x^2 - 1 \pmod{n}$
5: \hspace{2em} $y \leftarrow f(f(y))$
6: \hspace{2em} $g \leftarrow \gcd(|x - y|, n)$
7: \hspace{1em} if $g = n$ then
8: \hspace{2em} \hspace{1em} return failure
9: \hspace{1em} else
10: \hspace{2em} \hspace{1em} return $g$  \hspace{0.5em} $\triangleright$ $g$ must be a divisor of $n$
Pollard’s $\rho$: Why does it work?

- Let $p | n$ prime.
- Want $p | (x - y)$ so that $\gcd(|x - y|, n) \geq p$.
- $p | (x - y)$ means $x \equiv y \pmod{p}$.
- When a cycle mod $p$ is found, we find a factor.
- When does that happen? Birthday paradox
- For the birthday paradox to work, we need to expect that $f$ is a uniform function: Every remainder has an equal probability of being chosen.
- This is a conjecture, but empirical data approximately supports it
Overview

Modular Arithmetic

Discrete Algorithm and Congruences

Residue classes mod n

Integers modulo n

Arithmetic with integers mod n

GCD and Totatives

Inverses mod n

Euler's Theorem

Cost of Multiplication and GCD

Integer Factorization

Trial Division

Pollard's p − 1

Cycles in Z/nZ

Floyd's cycle-finding

Pollard's ρ

Birthday paradox

• “How many people need to be in a room so that there is a probability of \( m \) that two of them have the same birthday?”

• “How many random variables do we need to draw from \( f \) such that two of them have the same remainder mod \( p \) with probability \( m \)?” \((X_i \equiv X_j \pmod{p})\)

• Of course, \( 0 < m < 1 \).

• Original birthday paradox: \( m = 0.5 \)
Assume every event equally likely.

\[ P(X_i \equiv r) = \frac{1}{p} \]

Assume the events are independent.

\[ P(X_i \equiv r \text{ and } X_j \equiv r) = P(X_i \equiv r)P(X_j \equiv r) = \frac{1}{p^2} \]

Probability that once \( X_i \) is chosen, \( X_j \) will have same birthday:

\[ P(X_i \equiv X_j) = \frac{1}{p} \]

Complement: probability that all remainders are different.

- Dropping \( \mod p \) for convenience
- Due to the functional iteration, the random variables \( X_i \) and \( X_j \) are NOT independent! But we assume so for the sake of the analysis (and so does Pollard in his initial paper.)
- Independent: Once \( X_i \) is there, the probability that \( X_j \) will be the same remainder is \( 1/p \).
Let $A_i$ be the event that $X_i \neq X_j$ for all $0 \leq j < i$. Then, the event that choosing $\lambda$ random variables yields distinct remainders is

$$B_{\lambda} = \bigcap_{i=0}^{\lambda-1} A_i = B_{\lambda-1} \cap A_{\lambda-1}$$

By defn of conditional probability:

$$P(B_{\lambda}) = P(B_{\lambda-1})P(A_{\lambda-1}|B_{\lambda-1})$$

Then,

$$P(A_i|B_i) = \frac{p-i}{p},$$

since for $A_i$, $i$ remainders are already "occupied" and $p-i$ remainders are "left."
Expanding, we have (since \( P(B_1) = P(A_0) = 1 \))

\[
P(B_\lambda) = \prod_{i=0}^{\lambda-1} P(A_i | B_i) = \prod_{i=0}^{\lambda-1} \frac{p-i}{p} \\
= \prod_{i=0}^{\lambda-1} \left(1 - \frac{i}{p}\right) = \frac{p!}{(p-\lambda)!p^\lambda}
\]

Using the approximation \( 1 - x \approx e^{-x} \) (Taylor series),

\[
P(B_\lambda) \approx 1 \times \prod_{i=1}^{\lambda-1} e^{-i/p} = e^{-\sum_{i=1}^{\lambda-1} i/p} = e^{-(\lambda^2-\lambda)/2p}
\]

Now, we want \( P(B_\lambda) \leq 1 - m \).

Notice that this gets us the median for \( m = 0.5 \! \).
Thus,
\[ e^{-\frac{\lambda^2 - \lambda}{2p}} \leq 1 - m \]
\[ \lambda^2 - \lambda + 2p \ln(1 - m) \geq 0 \]

Then,
\[ \lambda \geq \frac{1}{2} + \frac{1}{2} \sqrt{1 - 8p \ln(1 - m)} \]

- Then, in Pollard’s \( \rho \), we find a cycle mod \( p \) with probability \( \frac{1}{2} \) after approximately \( \frac{1}{2} \sqrt{8 \ln(2)} p \approx 1.177\sqrt{p} \) iterations.
- In fact, we always find a cycle mod \( p \) in \( \theta(\sqrt{p}) \) steps.

Error analysis for this: [http://dx.doi.org/10.1137/1033051](http://dx.doi.org/10.1137/1033051)
Different analysis due to Knuth: mean instead of median.

\[ E[\lambda] = \sum_{\lambda=1}^{p+1} P(B_\lambda) = 1 + \sum_{\lambda=1}^{p} P(B_\lambda) = 1 + \sum_{\lambda=1}^{p} \frac{p!}{(p-\lambda)!p^\lambda} \]

Define the Ramanujan Q function:

\[ Q(n) = \sum_{k=1}^{n} \frac{n^k}{(n-k)!n^k} \]

Then,

\[ E[\lambda] = 1 + Q(p) \]

The Q function can be approximated by

\[ Q(p) \approx \sqrt{\frac{\pi p}{2}} \approx 1.2533\sqrt{p} \]

- summing the tail probabilities
- On Ramanujan’s Q function: http://algo.inria.fr/flajolet/Publications/FIGrKiPr95.pdf
Fermat's method

$n$ must be odd.

1. $\text{Fermat}(n)$
2. $a \leftarrow \sqrt{n}$
3. $b \leftarrow a^2 - n$
4. while $b$ is not a square do
   a. $a \leftarrow a + 1$
   b. $b \leftarrow a^2 - n$
5. return $a - \sqrt{b}$ ▷ or $a + \sqrt{b}$

- Pierre de Fermat, when? Original paper?
- $n$ must be odd
- finds non-trivial divisor
Fermat’s: Why does it work?

- Every odd integer is the difference of two squares
- \( n = a^2 - b^2 = (a + b)(a - b) \)
- We hope that \( 1 < a + b < n \) (or equivalently same for \( a - b \))
- Rearrange: \( b^2 = a^2 - n \).
- Try values for \( a \) until \( b^2 \) is a square.
- Worst case: \( n \) is prime. \( O(n) \) steps.
- Works best when prime factor is close to square-root of \( n \).
Fermat's: An Improvement

- Is there a way to know when values of $a$ make $b^2$ a square?
Fermat’s: An Improvement

- Is there a way to know when values of $a$ make $b^2$ a square?
- Bézout’s identity again: $\gcd(m, n) = 1$, then $m^{m-1}n + n^{n-1}m = 1$.

**Theorem (Chinese Remainder Theorem)**

Let $\gcd(n, m) = 1$. Then the following system has a solution and every solution is congruent mod $mn$:

$$x \equiv a \pmod{n} \quad x \equiv b \pmod{m}$$

Solutions are $x \equiv am^{m-1}n + bn^{n-1}m \pmod{mn}$.

- Want to factor $n = 2,345,678,917$.
- $\sqrt{n} = 48433$.
- $a^2 \equiv 0, 1, 4, 9 \pmod{16}$, $n \equiv 5 \pmod{16}$.
- $a^2 - n \equiv 11, 12, 15, 9 \pmod{16}$. Then, $a^2 \equiv 9 \pmod{16}$ only solution. That only happens $a \equiv \pm 3, \pm 5 \pmod{16}$.
- Also consider $\pmod{9}$. $n \equiv 7 \pmod{9}$.
- $a^2 \equiv 0, 1, 4, 7 \pmod{9}$. Then, $a^2 - n \equiv 2, 3, 6, 0 \pmod{9}$.
- Thus, $a^2 \equiv 7 \pmod{9}$ only solution. That’s when $a \equiv \pm 4 \pmod{9}$.
- $a = cn^{n-1} + dm^{m-1} = 16(4)c + (-7)9d$.
- $c \pmod{d} \equiv \pm 3 \pm 5 \pmod{13}$. Let $c$ be the mod $m$ values, $d$ the mod $n$.
- Then, $a \equiv \pm 5, \pm 13 \pmod{9 \times 16}$ (72).
- Only 4 out of 72 integers are possible – only 4/72 need to be checked.