# Introduction to Modern Cryptography

Benny Chor

The Prime Number Theorem
Primality Testing
Integer Multiplication and Factoring
as a One Way Function

Lecture 6

Tel-Aviv University

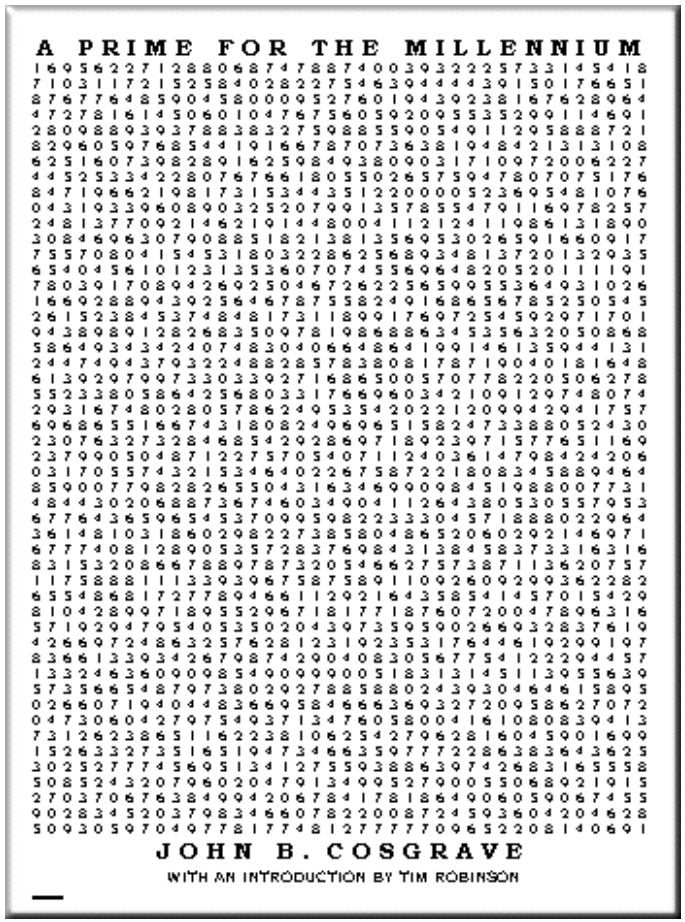
Posted November 21, 2009

# Remark: Multiplicative Generators in $Z_m^*$, $m$ composite

- If $m$ has two odd prime factors, then there are no multiplicative generators in $Z_m^*$. In other words, the order of all elements is smaller than $\phi(m)$.
- This leaves a rather limited repertoire for $m$. Either $m = 2^k$, $m = p^\ell$, or $m = 2^k \cdot p^\ell$ (where $k, \ell \geq 1$).
- A necessary and sufficient condition* for the existence of a primitive element in $Z_m^*$ is $m = 2, 4, p^\ell$ or $2p^\ell$, where $p$ is an odd prime.
- Examples (easily verified using Sage)
  - For $m = 25 = 5^2$, $\phi(m) = 5^2 - 5 = 20$. $3$ is a primitive element of $Z_{25}^*$.
  - For $m = 16 = 2^4$, $\phi(m) = 2^4 - 2^3 = 8$. The ring $Z_{16}^*$ has no primitive element.

*thanks to Shoni Dar for getting this straight.

# Prime Numbers and Primality Testing




http://www.iol.ie/∼tandmfl/mprime.htm

Published in 2000: A prime number with 2000 digits (40-by-50 table).
By John Cosgrave, Math Dept, St. Patrick's College, Dublin, Ireland.

# The Prime Number Theorem

- The fact that there are infinitely many primes was proved already by Euclid, in his Elements (Book IX, Proposition 20).
- The proof is by contradiction: Suppose there are finitely many primes $p_1, p_2, \ldots, p_k$. Then $p_1 \cdot p_2 \cdot \ldots \cdot p_k + 1$ cannot be divisible by any of the $p_i$, so its prime factors are none of the $p_i$s. (Note that $p_1 \cdot p_2 \cdot \ldots \cdot p_k + 1$ need not be a prime itself, *e.g.* $2 \cdot 3 \cdot 5 \cdot 7 \cdot 11 \cdot 13 + 1 = 30,031 = 59 \cdot 509$.)
- Once we know there are infinitely many primes, we may wonder how many are there up to an integer $x$.
- Let $\pi(x)$ denote the number of primes, $p$, up to $x$. For example, $\pi(30) = 4 + 4 + 2 = 10$.
- The prime number theorem: $\pi(x) \approx \frac{x}{\ln x}$.
  Furthermore, for $x \geq 55$,

$$\frac{x}{\ln x + 2} \leq \pi(x) \leq \frac{x}{\ln x - 4} \ .$$

# The Prime Number Theorem (cont.)

- Denote by $p_n$ the $n$-th prime number. As a consequence of the prime number theorem, we have (asymptotically) $p_n \approx n \ln n$. Furthermore, for $n \geq 6$,

$$n \ln n + n(\ln \ln n - 1) < p_n < n \ln n + n \ln \ln n \ .$$

- For univariate polynomials $f(x)$, the analog notion to primality is irreducibility. Over finite fields $GF(p)$, the analog question to estimating $\pi(x)$ is estimating $N_k$, the number of irreducible polynomials of degree $k$. It is known that $N_k \approx p^k/k$ .
- Thus there are many prime numbers and many irreducible polynomials (one in length of number or deg of polynomial).
- Finally, unrelated but fascinating, is Goldbach's conjecture: Every even integer greater than 2 can be written as the sum of two primes (*e.g.* $64 = 17 + 47$, $130 = 41 + 89$).

# Testing Primality/Compositeness

- Now that we know there are heaps of primes, we would like to efficiently test if a given integer is prime.
- Given an $n$ bits integer $m$, $2^{n-1} \leq m < 2^n$, we want to determine if $m$ is composite.
- The decision problem is certainly in NP (guess a factor and verify).
- The search problem, "given $m$ , find all its factors" is believed to be intractable. So search and decision are seemingly not equivalent here.

- Determining if if $m$ is prime turns out to be in NP as well (slightly more complicated, but by now you got all necessary tools, and will see this in recitation).

# Primality (Actually Compositeness) Testing

Question: Is there a better way to solve the decision problem (test if $m$ is composite) than by solving the search problem (factor $m$)?

Basic Idea [Solovay-Strassen, 1977]: To show that $m$ is composite, enough to find evidence that $m$ does not behave like a prime.
Such evidence need not include any prime factor of $m$.

# Primality Testing: Fermat Little Theorem

By Fermat little theorem, if $p$ is a prime and $a$ is in the range $1 \leq a \leq p-1$, then $a^{p-1} = 1 \pmod{p}$.

Suppose that if $m$ is an integer, and for some $a$ in in the range $2 \leq a \leq m-1$, $a^{m-1} \neq 1 \pmod{m}$. Such $a$ supplies a concrete evidence that $m$ is composite (but says nothing about $m$'s factorization).

Example: A proof, courtesy of Sage, that $(2^{271}+855)(2^{273}+5)$ is composite. Try the raw product (below) without this prior info:

$$57586096570152913699974892898380567793532123114264532903689671329431521032595057735 47$$

$$62127218213418370600635751564409932087528242170854099597452360087788392189830 91$$

```
m=(2^271+855)*(2^273+5)
a=53
mod(a,m)^(m-1)
```

evaluate

```
476199945079664649117280271145429334681695979085954845597307977858423779\
598860900382463448088082123974525462218348730693218810327954349851403109\
93675494451804665046
```

This proof gives no clue on $m$'s factorization (and Sage's `factor` was of no help – at least within my span of patience...).

# Applicability of Fermat Test

Question: Given a composite number, $m$, is there always a Fermat witness, $a$, $2 \leq a \leq m-1$? Furthermore, are there enough of them so that if we pick many $a$'s at random, we will hit at least one Fermat witness (with high probability)?

It would be nice, had it been the case. Unfortunately, it is not. There are some $m$ for which Fermat test always fails.

```
m=225593397919
a1=77665
a2=899086
a3=444444444
print(mod(a1,m)^(m-1))
print(mod(a2,m)^(m-1))
print(mod(a3,m)^(m-1))
```

```
1
1
1
```

We just saw three witnesses for $m = 225593397919$ that fail to prove it is a composite. You can try looking for others yourselves.
But hey, maybe the lecturer is pulling your leg and $m$ is prime. . . ?

Nope: $6619$ divides $m$ !

# Carmichael Numbers

These are composites $m$ where Fermat test fails, namely $a^{m-1} = 1 \pmod m$ for almost all $a$, $2 \leq a \leq m-1$.

Theorem: $m$ is a Carmichael number iff $m = p_1 \cdot p_2 \cdot p_3 \cdot \ldots \cdot p_k$, where $k \geq 3$, all $p_i$ are distinct primes, and for every $p_i$, $p_i - 1$ divides $m - 1$.

Example:

```
m=225593397919
factor(m)
```

```
2207 * 6619 * 15443
```

```
print((m-1) % 2206)
print((m-1) % 6618)
print((m-1) % 15442)
```

```
0
0
0
```

Carmichael numbers are rare, still there are infinitely many of them, and we'd like our compositeness test to "catch" them as well.

# Extended Evidence for Compositeness

Given an integer, $m$, we will say that $a$, $2 \leq a \leq m-1$ is an extended witness for $m$'s compositeness if either

1. $\gcd(m, a) > 1$ (non trivial factor).
2. $a^{m-1} \neq 1 \pmod{m}$ (Fermat test).
3. $a^2 = 1 \pmod{m}$ but $a \neq m-1$
   (implying $1$ has more than two square roots in $Z_m^*$).

# Back to Our Favorite $m = 225593397919$

With $m$ being a Carmichael number, we won't easily find an extended witness $a$ that is either a non trivial factor (type 1) or flunks the Fermat test (type 2).
Let $m-1 = 2r$. Suppose $b$ is not a witness of type 2, namely $b^{m-1} = (b^r)^2 = 1 \pmod m$. Denote $a = b^r$. If $a \neq \pm 1 \pmod m$ then $a$ is an extended witness of type (3).

Example:

```
m=225593397919; b1=899086; b2=444444444
print(gcd(b1,m),(mod(b1,m)^(m-1)),(mod(b1,m)^((m-1)/2)))
print(gcd(b2,m),(mod(b2,m)^(m-1)),(mod(b2,m)^((m-1)/2)))
```

```
(1, 1, 37615935855)
(1, 1, 187977462064)
```

Gotcha! In both cases, for $a_i = b_i^{(m-1)/2}$ we have $a_i^2 = 1 \pmod m$ but $a_i \neq \pm 1 \pmod m$. This proves that $m = 225593397919$ is composite.

# Extending to General $m$

- Let $m-1 = 2^k \cdot r$, with $r$ odd.
- For any $b$, $b^{m-1} = ((\ldots((b^r)^2)^2\ldots)^2)^2$ ($k$ squaring operations).
- If $b^{m-1} \neq 1 \pmod m$ then $b$ is a witness of type (2). $\sqrt{}$
- Otherwise, let $a_0 = b^r, a_1 = a_0^2, a_2 = a_1^2, \ldots, a_k = a_{k-1}^2$. Then $a_k = b^{m-1} \pmod m$.
- Let $j$ be the smallest index with $a_j = 1 \pmod m$. (There is always such $j$ since $a_k = 1 \pmod m$.)
- If $0 < j$ and $a_{j-1} \neq -1 \pmod m$, then $a_{j-1}$ is an extended witness of type (3), hence $m$ is composite. $\sqrt{}\sqrt{}$

Any $b$ satisfying either $\sqrt{}$ or $\sqrt{}\sqrt{}$ will be called a smart witness.

# Smart Witness – Sage example

```
m=45123394470901560 4501  # Carmichael number as well
print("m-1=",factor(m-1))
print
b=33333
print(mod(b,m)^((m-1)/4))
print(mod(b,m)^((m-1)/2))
print(mod(b,m)^(m-1))
```

```
('m-1=', 2^2 * 3^4 * 5^3 * 7 * 13 * 17 * 10427 * 690712081)

19618548344805160 1210
36098186654993601 5479
1
```

We have $a_1 \neq \pm 1 \pmod{m}$, but $a_1^2 = 1 \pmod{m}$.
So $b = 33333$ satisfies $\sqrt{\sqrt{}}$, and is therefore a smart witness for the compositeness of $m = 45123394470901560 4501$.

# Miller Theorem (1977)

Let $m-1=2^k \cdot r$, with $r$ odd. If $m$ is composite, then[†] there is a small smart witness $b$, where small means $b < 3 \cdot (\log m)^2/2$ (the improved constant $3/2$ was shown by Wedeniwski in 2001).

---

[†] Assuming the extended Riemann hypothesis. The "regular" Riemann hypothesis, formulated by Bernhard Riemann in 1859, is one of the most famous and important unsolved problems in mathematics, dealing with the distribution of non-trivial zero of the Riemann zeta function. It is part of in Hilbert's eighth problem, together with the Goldbach conjecture. The Clay Institute has offered $1,000,000 for resolving it (a similar prize is offered for resolving P vs. NP). The extended conjecture deals with the distribution of zeroes not only for the Riemann zeta function, but for any Dirichlet L-series.

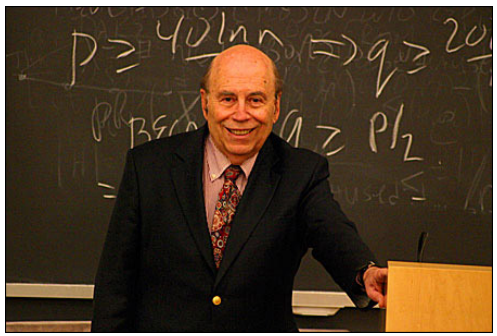
# Miller Theorem (1977)

Let $m-1=2^k \cdot r$, with $r$ odd. If $m$ is composite, then there is a small smart witness $b$, where small means $b < 3 \cdot (\log m)^2/2$ (the improved constant was shown by Wedeniwski in 2001).

- This means that going over all $b < 3 \cdot (\log m)^2/2$ and applying the extended test to each of them, we get a deterministic polynomial time algorithm for testing if $m$ is a prime.
- If $m$ passes all tests, it is a prime.
- The complexity is $O(\log^3 m)$ operations per $b$.
- There are $O(\log^2 m)$ numbers $b$ to test, so overall it is $O(\log^5 m)$ operations.
- The only caveat is the dependence on a very heavy, unproved conjecture.

# Rabin's Theorem (1980)

Let $m-1=2^k \cdot r$, with $r$ odd. If $m$ is composite, then at least $3m/4$ of all $b$ in the range $1<b<m$ are smart witnesses.



No assumptions required, and proof of statement employs only elementary arguments.

Each $b$ takes $O(\log^3 m)$ bit operations to test, so if we probe just $O(1)$ of them, complexity is $O(\log^3 m)$.

# Miller-Rabin Randomized Primality Testing

- The input is an odd integer $m$ with $n$ bits ($2^{n-1} < m < 2^n$)
- Repeat $100$ times
  - Pick $b$ in the range $1 < b < m$ at random and independently.
  - Check if $b$ is a smart witness.
- If one or more $b$ is a smart witness, output "$m$ is composite".
- If no smart witness found, output "$m$ is prime".

# Miller-Rabin Randomized Primality Testing

- The input is an odd integer $m$ with $n$ bits ($2^{n-1} < m < 2^n$)
- Repeat $100$ times
  - Pick $b$ in the range $1 < b < m$ at random and independently.
  - Check if $b$ is a smart witness.
- If one or more $b$ is a smart witness, output "$m$ is composite".
- If no smart witness found, output "$m$ is prime".

Remark: Solovay and Strassen have invented in 1977 a different, and slightly less efficient randomized primality testing algorithm.

# Properties of Miller-Rabin Primality Testing

- Randomized: uses coin flips to pick the $b$'s.
- Run time is polynomial in $n$, the length of $m$.
- If $m$ is prime, the algorithm always outputs "$m$ is prime".

# Properties of Miller-Rabin Primality Testing

- Randomized: uses coin flips to pick the $b$'s.
- Run time is polynomial in $n$, the length of $m$.
- If $m$ is prime, the algorithm always outputs "$m$ is prime".
- 
- If $m$ is composite, the algorithm may err and outputs "$m$ is prime".
- However, to err, all random choices of $b$'s should yield non-witnesses. Therefore,

$$\text{Probability of error} < \left(\frac{1}{4}\right)^{100} \lll 1 \ .$$

# Primality Testing

In terms of complexity classes, the Miller-Rabin algorithm, as well as the Solovay-Strassen algorithm, imply

## Composites $\in$ RP

Where RP=Random Poly Time, one sided error.
Easy fact: RP is contained in NP.

# Primality Testing

In terms of complexity classes, the Miller-Rabin algorithm, as well as the Solovay-Strassen algorithm, imply

## Composites $\in$ RP

Where RP=Random Poly Time, one sided error.
Easy fact: RP is contained in NP.

For all practical purposes, the Miller-Rabin algorithm (and various optimizations thereof) supply a satisfactory solution for identifying primes.

Still the question whether Composites,primes $\in$ P remained open.

# Primality Testing in P

In summer 2002, Prof. Manindra Agrawal and his Ph.D. students Neeraj Kayal and Nitin Saxena, from the India Institute of Technology, Kanpur, finally found a deterministic polynomial time algorithm for determining primality. Initially, their algorithm ran in time $O(n^{12})$. In 2005, Carl Pomerance and H. W. Lenstra, Jr. improved this to running in time $O(n^6)$.

Agrawal, Kayal, and Saxena received the 2006 Fulkerson Prize and the 2006 Gödel Prize for their work.

# Primality Testing in P

Excerpts from the SIGACT Award citation:

"In August 2002 one of the most ancient computational problems was finally solved. Agrawal, Kayal, and Saxena presented an unconditional deterministic polynomial time algorithm that determines whether an input number is prime or composite. All previously known polynomial time primality tests were based on probabilistic methods or they relied on an unproven assumption, known as the generalized Riemann Hypothesis. The result obtained by Agrawal, Kayal, and Saxena can be seen as a crowning achievement of a long algorithmic and mathematical quest. A remarkable aspect of the article is that the final exposition itself turns out to be rather simple. The text as published in Annals of Mathematics is a masterpiece in mathematical reasoning. It has a high density of tricks and techniques, but the arguments come in a brilliantly simple manner; they remain completely elementary. . ."

# Integer Multiplication & Factoring as a One Way Function.

Multiplying two $n$ bit numbers takes time $O(n^2)$.

Factoring an $n$ bit number takes time $2^{c \cdot n^{1/3}}$ (using the currently best algorithm).

Easy: $p, q \longrightarrow m = p \cdot q$ (finding random primes & integer multiplication).

Hard: $m = p \cdot q \longrightarrow p, q$ (integer factorization).

Question: Can public key cryptosystem be based on this observation?

# The RSA Public Key Cryptosystem (1978)

- Bob's private information: two large primes $p, q$.
- Public information: Their product, $m = p \cdot q$. An integer $e$ that is relatively prime to $\phi(m) = (p-1) \cdot (q-1)$.
- More private information: An integer $d$ that is relatively prime to $\phi(m) = (p-1) \cdot (q-1)$ and satisfies $d \cdot e = 1 \pmod{\phi(m)}$.

# The RSA Public Key Cryptosystem (1978)

- Bob's private information: two large primes $p, q$.
- Public information: Their product, $m = p \cdot q$. An integer $e$ that is relatively prime to $\phi(m) = (p-1) \cdot (q-1)$.
- More private information: An integer $d$ that is relatively prime to $\phi(m) = (p-1) \cdot (q-1)$ and satisfies $d \cdot e = 1 \pmod{\phi(m)}$.

- Messages $A$ are elements in $Z_m$, namely numbers in $[1, \ldots, m-1]$.
- To encrypt $A$, compute $C = A^e \pmod{m}$, and send $C$ to Bob.
- To decrypt $C$, Bob computes $C^d = A^{d \cdot e} = A \pmod{m}$.