Primality Tests

• PRIMES asks if a number $N$ is a prime.
• The classic algorithm tests if $k \mid N$ for $k = 2, 3, \ldots, \sqrt{N}$.
• But it runs in $\Omega\left(2^{\log_2 N}/2\right)$ steps.
The Fermat Test for Primality

Fermat’s “little” theorem (p. 493) suggests the following primality test for any given number $N$:

1: Pick a number $a$ randomly from $\{1, 2, \ldots, N - 1\}$;
2: if $a^{N-1} \not\equiv 1 \mod N$ then
3: return “$N$ is composite”;
4: else
5: return “$N$ is (probably) a prime”;
6: end if
The Fermat Test for Primality (concluded)

- **Carmichael numbers** are composite numbers that will pass the Fermat test for all \( a \in \{1, 2, \ldots, N - 1\} \).\(^a\)
  - The Fermat test will return “\( N \) is a prime” for all Carmichael numbers \( N \).

- Unfortunately, there are infinitely many Carmichael numbers.\(^b\)

- In fact, the number of Carmichael numbers less than \( N \) exceeds \( N^{2/7} \) for \( N \) large enough.

- So the Fermat test is an incorrect algorithm for PRIMES.

\(^a\)Carmichael (1910). Lo (1994) mentions an investment strategy based on such numbers!

\(^b\)Alford, Granville, & Pomerance (1992).
Square Roots Modulo a Prime

- Equation $x^2 \equiv a \mod p$ has at most two (distinct) roots by Lemma 64 (p. 498).
  - The roots are called square roots.
  - Numbers $a$ with square roots and $\gcd(a, p) = 1$ are called quadratic residues.
    * They are
      $$1^2 \mod p, 2^2 \mod p, \ldots, (p - 1)^2 \mod p.$$  

- We shall show that a number either has two roots or has none, and testing which is the case is trivial.\(^a\)

\(^a\)But no efficient deterministic general-purpose square-root-extracting algorithms are known yet.
Euler’s Test

Lemma 69 (Euler) Let $p$ be an odd prime and $a \neq 0 \mod p$.

1. If

$$a^{(p-1)/2} \equiv 1 \mod p,$$

then $x^2 \equiv a \mod p$ has two roots.

2. If

$$a^{(p-1)/2} \not\equiv 1 \mod p,$$

then

$$a^{(p-1)/2} \equiv -1 \mod p$$

and $x^2 \equiv a \mod p$ has no roots.
The Proof (continued)

- Let $r$ be a primitive root of $p$.
- Fermat’s “little” theorem says $r^{p-1} \equiv 1 \mod p$, so
  $$r^{(p-1)/2}$$
  is a square root of 1.
- In particular,
  $$r^{(p-1)/2} \equiv 1 \text{ or } -1 \mod p.$$ 
- But as $r$ is a primitive root, $r^{(p-1)/2} \not\equiv 1 \mod p$.
- Hence $r^{(p-1)/2} \equiv -1 \mod p$. 

©2020 Prof. Yuh-Dauh Lyuu, National Taiwan University
The Proof (continued)

• Let $a = r^k \mod p$ for some $k$.

• Suppose $a^{(p-1)/2} \equiv 1 \mod p$.

• Then

$$1 \equiv a^{(p-1)/2} \equiv r^{k(p-1)/2} \equiv \left[r^{(p-1)/2}\right]^k \equiv (-1)^k \mod p.$$ 

• So $k$ must be even.
The Proof (continued)

- Suppose $a = r^{2j} \mod p$ for some $1 \leq j \leq (p-1)/2$.
- Then
  \[ a^{(p-1)/2} \equiv r^{j(p-1)} \equiv 1 \mod p. \]
- The two distinct roots of $a$ are
  \[ r^j, -r^j (\equiv r^{j+(p-1)/2} \mod p). \]
  - If $r^j \equiv -r^j \mod p$, then $2r^j \equiv 0 \mod p$, which implies $r^j \equiv 0 \mod p$, a contradiction as $r$ is a primitive root.
The Proof (continued)

- As $1 \leq j \leq (p - 1)/2$, there are $(p - 1)/2$ such $a$'s.
- Each such $a \equiv r^{2j} \mod p$ has 2 distinct square roots.
- The square roots of all these $a$'s are distinct.
  - The square roots of different $a$'s must be different.
- Hence the set of square roots is $\{1, 2, \ldots, p - 1\}$.
- As a result,
  \[ a = r^{2j} \mod p, \quad 1 \leq j \leq (p - 1)/2, \]
  exhaust all the quadratic residues.
The Proof (concluded)

- Suppose $a = r^{2j+1} \mod p$ now.

- Then it has no square roots because all the square roots have been taken.

- Finally,

$$a^{(p-1)/2} \equiv \left[ r^{(p-1)/2} \right]^{2j+1} \equiv (-1)^{2j+1} \equiv -1 \mod p.$$
The Legendre Symbol and Quadratic Residuacity Test

- By Lemma 69 (p. 560),
  \[ a^{(p-1)/2} \mod p = \pm1 \]
  for \( a \not\equiv 0 \mod p \).

- For odd prime \( p \), define the Legendre symbol \((a \mid p)\) as
  \[
  (a \mid p) \overset{\Delta}{=} \begin{cases} 
  0, & \text{if } p \mid a, \\
  1, & \text{if } a \text{ is a quadratic residue modulo } p, \\
  -1, & \text{if } a \text{ is a quadratic nonresidue modulo } p.
  \end{cases}
  \]

- It is sometimes pronounced “a over p.”

\(^a\)Andrien-Marie Legendre (1752–1833).
The Legendre Symbol and Quadratic Residuacity Test (concluded)

- Euler’s test (p. 560) implies

\[ a^{(p-1)/2} \equiv (a | p) \mod p \]

for any odd prime \( p \) and any integer \( a \).

- Note that \((ab | p) = (a | p)(b | p)\).
Gauss’s Lemma

Lemma 70 (Gauss) Let $p$ and $q$ be two distinct odd primes. Then $(q \mid p) = (-1)^m$, where $m$ is the number of residues in $R \triangleq \{ iq \mod p : 1 \leq i \leq (p - 1)/2 \}$ that are greater than $(p - 1)/2$.

- All residues in $R$ are distinct.
  - If $iq = jq \mod p$, then $p \mid (j - i)$ or $p \mid q$.
  - But neither is possible.

- No two elements of $R$ add up to $p$.
  - If $iq + jq \equiv 0 \mod p$, then $p \mid (i + j)$ or $p \mid q$.
  - But neither is possible.
The Proof (continued)

- Replace each of the $m$ elements $a \in R$ such that $a > (p - 1)/2$ by $p - a$.
  - This is equivalent to performing $-a \mod p$.

- Call the resulting set of residues $R'$.

- All numbers in $R'$ are at most $(p - 1)/2$.

- In fact, $R' = \{1, 2, \ldots, (p - 1)/2\}$ (see illustration next page).
  - Otherwise, two elements of $R$ would add up to $p$,\(^a\) which has been shown to be impossible.

\(^a\)Because then $iq \equiv -jq \mod p$ for some $i \neq j$. 
$p = 7$ and $q = 5$. 
The Proof (concluded)

• Alternatively, \( R' = \{ \pm iq \mod p : 1 \leq i \leq (p - 1)/2 \} \), where exactly \( m \) of the elements have the minus sign.

• Take the product of all elements in the two representations of \( R' \).

• So

\[
[(p - 1)/2]! \equiv (-1)^m q^{(p-1)/2}[(p - 1)/2]! \mod p.
\]

• Because \( \gcd([(p - 1)/2]!, p) = 1 \), the above implies

\[
1 = (-1)^m q^{(p-1)/2} \mod p.
\]
Legendre’s Law of Quadratic Reciprocity\textsuperscript{a}

- Let \( p \) and \( q \) be two distinct odd primes.
- The next result says \((p \mid q)\) and \((q \mid p)\) are distinct if and only if both \( p \) and \( q \) are 3 mod 4.

**Lemma 71 (Legendre, 1785; Gauss)**

\[
(p \mid q)(q \mid p) = (-1)^{\frac{p-1}{2} \frac{q-1}{2}}.
\]

\textsuperscript{a}First stated by Euler in 1751. Legendre (1785) did not give a correct proof. Gauss proved the theorem when he was 19. He gave at least 8 different proofs during his life. The 152nd proof appeared in 1963. A computer-generated formal proof was given in Russinoff (1990). As of 2008, there had been 4 such proofs. Wiedijk (2008), “the Law of Quadratic Reciprocity is the first nontrivial theorem that a student encounters in the mathematics curriculum.”
The Proof (continued)

- Sum the elements of $R'$ in the previous proof in mod2.
- On one hand, this is just $\sum_{i=1}^{(p-1)/2} i \mod 2$.
- On the other hand, the sum equals

$$m p + \sum_{i=1}^{(p-1)/2} \left( i q - p \left\lfloor \frac{i q}{p} \right\rfloor \right) \mod 2$$

$$= m p + \left( q \sum_{i=1}^{(p-1)/2} i - p \sum_{i=1}^{(p-1)/2} \left\lfloor \frac{i q}{p} \right\rfloor \right) \mod 2.$$

- $m$ of the $iq \mod p$ are replaced by $p - iq \mod p$.
- But signs are irrelevant under mod2.
- $m$ is as in Lemma 70 (p. 568).
The Proof (continued)

• Ignore odd multipliers to make the sum equal

\[
m + \left( \sum_{i=1}^{(p-1)/2} i - \sum_{i=1}^{(p-1)/2} \left\lfloor \frac{iq}{p} \right\rfloor \right) \mod 2.
\]

• Equate the above with \( \sum_{i=1}^{(p-1)/2} i \) modulo 2.

• Now simplify to obtain

\[
m \equiv \sum_{i=1}^{(p-1)/2} \left\lfloor \frac{iq}{p} \right\rfloor \mod 2.
\]
The Proof (continued)

• \( \sum_{i=1}^{(p-1)/2} \left\lfloor \frac{iq}{p} \right\rfloor \) is the number of integral points below the line

\[ y = \left(\frac{q}{p}\right)x \]

for \( 1 \leq x \leq (p - 1)/2 \).

• Gauss’s lemma (p. 568) says \((q \mid p) = (-1)^m\).

• Repeat the proof with \(p\) and \(q\) reversed.

• Then \((p \mid q) = (-1)^{m'}\), where \(m'\) is the number of integral points above the line \(y = \left(\frac{q}{p}\right)x\) for \(1 \leq y \leq (q - 1)/2\).
The Proof (concluded)

• As a result,

\[(p \mid q)(q \mid p) = (-1)^{m+m'}.
\]

• But \(m + m'\) is the total number of integral points in the

\([1, \frac{p-1}{2}] \times [1, \frac{q-1}{2}]\) rectangle, which is

\[
\frac{p - 1}{2} \cdot \frac{q - 1}{2}.
\]
Above, $p = 11$, $q = 7$, $m = 7$, $m' = 8$. 
The Jacobi Symbol

- The Legendre symbol only works for odd prime moduli.
- The Jacobi symbol \((a \mid m)\) extends it to cases where \(m\) is not prime.
  - \(a\) is sometimes called the numerator and \(m\) the denominator.
- Trivially, \((1 \mid m) = 1\).
- Define \((a \mid 1) = 1\).

\(^a\)Carl Jacobi (1804–1851).
The Jacobi Symbol (concluded)

- Let $m = p_1p_2\cdots p_k$ be the prime factorization of $m$.

- When $m > 1$ is odd and $\gcd(a, m) = 1$, then

$$\left(\frac{a}{m}\right) \triangleq \prod_{i=1}^{k} \left(\frac{a}{p_i}\right).$$

  - Note that the Jacobi symbol equals $\pm 1$.
  - It reduces to the Legendre symbol when $m$ is a prime.
Properties of the Jacobi Symbol

The Jacobi symbol has the following properties when it is defined.

1. \((ab \mid m) = (a \mid m)(b \mid m)\).

2. \((a \mid mn) = (a \mid m)(a \mid n)\).

3. If \(a \equiv b \mod m\), then \((a \mid m) = (b \mid m)\).

4. \((-1 \mid m) = (-1)^{(m-1)/2}\) (by Lemma 70 on p. 568).

5. \((2 \mid m) = (-1)^{(m^2-1)/8}\).

6. If \(a\) and \(m\) are both odd, then
\[
(a \mid m)(m \mid a) = (-1)^{(a-1)(m-1)/4}.
\]

\(^a\)By Lemma 70 (p. 568) and some parity arguments.
Properties of the Jacobi Symbol (concluded)

• Properties 3–6 allow us to calculate the Jacobi symbol without factorization.
  – It will also yield the same result as Euler’s test\(^a\) when \(m\) is an odd prime.

• This situation is similar to the Euclidean algorithm.

• Note also that \((a \mid m) = 1/(a \mid m)\) because \((a \mid m) = \pm 1\).\(^b\)

\(^a\)Recall p. 560.
\(^b\)Contributed by Mr. Huang, Kuan-Lin (B96902079, R00922018) on December 6, 2011.
Calculation of $(2200 | 999)$

\[
(2200 | 999) = (202 | 999) \\
= (2 | 999)(101 | 999) \\
= (-1)^{(999^2-1)/8}(101 | 999) \\
= (-1)^{124750}(101 | 999) = (101 | 999) \\
= (-1)^{(100)(998)/4}(999 | 101) = (-1)^{24950}(999 | 101) \\
= (999 | 101) = (90 | 101) = (-1)^{(101^2-1)/8}(45 | 101) \\
= (-1)^{1275}(45 | 101) = -(45 | 101) \\
= -(1)^{(44)(100)/4}(101 | 45) = -(101 | 45) = -(11 | 45) \\
= -(1)^{(10)(44)/4}(45 | 11) = -(45 | 11) \\
= -(1 | 11) = -1.
\]
A Result Generalizing Proposition 10.3 in the Textbook

**Theorem 72** The group of set $\Phi(n)$ under multiplication mod $n$ has a primitive root if and only if $n$ is either 1, 2, 4, $p^k$, or $2p^k$ for some nonnegative integer $k$ and an odd prime $p$.

This result is essential in the proof of the next lemma.
The Jacobi Symbol and Primality Test

Lemma 73  If \((M \mid N) \equiv M^{(N-1)/2} \mod N\) for all \(M \in \Phi(N)\), then \(N\) is a prime. (Assume \(N\) is odd.)

- Assume \(N = mp\), where \(p\) is an odd prime, \(\gcd(m, p) = 1\), and \(m > 1\) (not necessarily prime).
- Let \(r \in \Phi(p)\) such that \((r \mid p) = -1\).
- The Chinese remainder theorem says that there is an \(M \in \Phi(N)\) such that
  \[
  \begin{align*}
  M &= r \mod p, \\
  M &= 1 \mod m.
  \end{align*}
  \]

Mr. Clement Hsiao (B4506061, R88526067) pointed out that the textbook’s proof for Lemma 11.8 is incorrect in January 1999 while he was a senior.
The Proof (continued)

• By the hypothesis,

\[ M^{(N-1)/2} = (M \mid N) = (M \mid p)(M \mid m) = -1 \mod N. \]

• Hence

\[ M^{(N-1)/2} = -1 \mod m. \]

• But because \( M = 1 \mod m, \)

\[ M^{(N-1)/2} = 1 \mod m, \]

a contradiction.
The Proof (continued)

• Second, assume that $N = p^a$, where $p$ is an odd prime and $a \geq 2$.

• By Theorem 72 (p. 583), there exists a primitive root $r$ modulo $p^a$.

• From the assumption,

$$M^{N-1} = \left[M^{(N-1)/2}\right]^2 = (M|N)^2 = 1 \mod N$$

for all $M \in \Phi(N)$. 
The Proof (continued)

• As $r \in \Phi(N)$ (prove it), we have

$$r^{N-1} = 1 \text{ mod } N.$$

• As $r$’s exponent modulo $N = p^a$ is $\phi(N) = p^{a-1}(p - 1)$,

$$p^{a-1}(p - 1) | (N - 1),$$

which implies that $p | (N - 1)$.

• But this is impossible given that $p | N$. 
The Proof (continued)

• Third, assume that $N = mp^a$, where $p$ is an odd prime, $\gcd(m, p) = 1$, $m > 1$ (not necessarily prime), and $a$ is even.

• The proof mimics that of the second case.

• By Theorem 72 (p. 583), there exists a primitive root $r$ modulo $p^a$.

• From the assumption,

$$M^{N-1} = \left[M^{(N-1)/2}\right]^2 = (M|N)^2 = 1 \mod N$$

for all $M \in \Phi(N)$.
The Proof (continued)

• In particular,

\[ M^{N-1} = 1 \mod p^a \]  \hspace{1cm} (15)

for all \( M \in \Phi(N) \).

• The Chinese remainder theorem says that there is an \( M \in \Phi(N) \) such that

\[ M = r \mod p^a, \]

\[ M = 1 \mod m. \]

• Because \( M = r \mod p^a \) and Eq. (15),

\[ r^{N-1} = 1 \mod p^a. \]
The Proof (concluded)

- As $r$’s exponent modulo $N = p^a$ is $\phi(N) = p^{a-1}(p - 1)$,

$$p^{a-1}(p - 1) \mid (N - 1),$$

which implies that $p \mid (N - 1)$.

- But this is impossible given that $p \mid N$.  

©2020 Prof. Yuh-Dauh Lyuu, National Taiwan University  
Page 590
The Number of Witnesses to Compositeness

Theorem 74 (Solovay & Strassen, 1977) If \(N\) is an odd composite, then \((M \mid N) \equiv M^{(N-1)/2} \mod N\) for at most half of \(M \in \Phi(N)\).

- By Lemma 73 (p. 584) there is at least one \(a \in \Phi(N)\) such that \((a \mid N) \neq a^{(N-1)/2} \mod N\).
- Let \(B \triangleq \{ b_1, b_2, \ldots, b_k \} \subseteq \Phi(N)\) be the set of all distinct residues such that \((b_i \mid N) \equiv b_i^{(N-1)/2} \mod N\).
- Let \(aB \triangleq \{ ab_i \mod N : i = 1, 2, \ldots, k \}\).
- Clearly, \(aB \subseteq \Phi(N)\), too.
The Proof (concluded)

- \(| aB | = k. \)
  - \( ab_i \equiv ab_j \mod N \) implies \( N | a(b_i - b_j) \), which is impossible because \( \gcd(a, N) = 1 \) and \( N > |b_i - b_j| \).

- \( aB \cap B = \emptyset \) because
  \[
  (ab_i)^{(N-1)/2} \equiv a^{(N-1)/2}b_i^{(N-1)/2} \not\equiv (a | N)(b_i | N) \equiv (ab_i | N).
  \]

- Combining the above two results, we know
  \[
  \frac{|B|}{\phi(N)} \leq \frac{|B|}{|B \cup aB|} = 0.5.
  \]
1: if $N$ is even but $N \neq 2$ then
2:   return “$N$ is composite”;
3: else if $N = 2$ then
4:   return “$N$ is a prime”;  
5: end if
6: Pick $M \in \{2, 3, \ldots, N - 1\}$ randomly;
7: if gcd($M, N$) > 1 then
8:   return “$N$ is composite”;
9: else
10:   if $(M \mid N) \equiv M^{(N-1)/2} \mod N$ then
11:      return “$N$ is (probably) a prime”;  
12: else
13:      return “$N$ is composite”;  
14: end if
15: end if
Analysis

- The algorithm certainly runs in polynomial time.
- There are no false positives (for COMPOSITENESS).
  - When the algorithm says the number is composite, it is always correct.
Analysis (concluded)

- The probability of a false negative (again, for COMPOSITENESS) is at most one half.
  - Suppose the input is composite.
  - By Theorem 74 (p. 591),
    \[
    \text{prob[algorithm answers “no” | } N \text{ is composite}] \leq 0.5.
    \]
  - Note that we are not referring to the probability that \( N \) is composite when the algorithm says “no.”

- So it is a Monte Carlo algorithm for COMPOSITENESS\(^a\) by the definition on p. 539.

\(^a\)Not PRIMES.
The Improved Density Attack for \textit{COMPOSITENESS}

- All numbers $< N$
- Witnesses to compositeness of $N$ via common factor
- Witnesses to compositeness of $N$ via Jacobi
Randomized Complexity Classes; RP

• Let $N$ be a polynomial-time precise NTM that runs in time $p(n)$ and has 2 nondeterministic choices at each step.

• $N$ is a **polynomial Monte Carlo Turing machine** for a language $L$ if the following conditions hold:
  – If $x \in L$, then at least half of the $2^{p(n)}$ computation paths of $N$ on $x$ halt with “yes” where $n = |x|$.
  – If $x \not\in L$, then all computation paths halt with “no.”

• The class of all languages with polynomial Monte Carlo TMs is denoted **RP** (randomized polynomial time).\(^a\)

\(^a\)Adleman & Manders (1977).
Comments on RP

- In analogy to Proposition 41 (p. 335), a “yes” instance of an RP problem has many certificates (witnesses).
- There are no false positives.
- If we associate nondeterministic steps with flipping fair coins, then we can phrase RP in the language of probability.
  - If \( x \in L \), then \( N(x) \) halts with “yes” with probability at least 0.5.
  - If \( x \notin L \), then \( N(x) \) halts with “no.”
Comments on RP (concluded)

- The probability of false negatives is $\leq 0.5$.

- But any constant $\epsilon$ between 0 and 1 can replace 0.5.
  - Repeat the algorithm
    $$k \triangleq \left\lceil -\frac{1}{\log_2 \epsilon} \right\rceil$$
    times.
  - Answer “no” only if all the runs answer “no.”
  - The probability of false negatives becomes $\epsilon^k \leq 0.5$. 
Where RP Fits

• P ⊆ RP ⊆ NP.
  – A deterministic TM is like a Monte Carlo TM except that all the coin flips are ignored.
  – A Monte Carlo TM is an NTM with more demands on the number of accepting paths.

• COMPOSITENESS ∈ RP;\(^a\) PRIMES ∈ coRP;
  PRIMES ∈ RP.\(^b\)
  – In fact, PRIMES ∈ P.\(^c\)

• RP ∪ coRP is an alternative “plausible” notion of efficient computation.

\(^a\)Rabin (1976); Solovay & Strassen (1977).
\(^b\)Adleman & Huang (1987).
\(^c\)Agrawal, Kayal, & Saxena (2002).
ZPP\textsuperscript{a} (Zero Probabilistic Polynomial)

- The class \textbf{ZPP} is defined as \textit{RP} \cap \textit{coRP}.
- A language in ZPP has \textit{two} Monte Carlo algorithms, one with no false positives (\textit{RP}) and the other with no false negatives (\textit{coRP}).
- If we repeatedly run both Monte Carlo algorithms, \textit{eventually} one definite answer will come (unlike \textit{RP}).
  - A \textit{positive} answer from the one without false positives.
  - A \textit{negative} answer from the one without false negatives.

\textsuperscript{a}Gill (1977).
The ZPP Algorithm (Las Vegas)

1: \{Suppose \( L \in \text{ZPP} \).\}
2: \{\( N_1 \) has no false positives, and \( N_2 \) has no false negatives.\}
3: while true do
4: \hspace{1em} if \( N_1(x) = \text{”yes”} \) then
5: \hspace{2em} return \text{”yes”};
6: \hspace{1em} end if
7: \hspace{1em} if \( N_2(x) = \text{”no”} \) then
8: \hspace{2em} return \text{”no”};
9: \hspace{1em} end if
10: end while
ZPP (concluded)

- The \textit{expected} running time for the correct answer to emerge is polynomial.
  - The probability that a run of the 2 algorithms does not generate a definite answer is 0.5 (why?).
  - Let $p(n)$ be the running time of each run of the while-loop.
  - The expected running time for a definite answer is
    \[
    \sum_{i=1}^{\infty} 0.5^i p(n) = 2p(n).
    \]

- Essentially, ZPP is the class of problems that can be solved, without errors, in expected polynomial time.
Large Deviations

• Suppose you have a biased coin.

• One side has probability $0.5 + \epsilon$ to appear and the other $0.5 - \epsilon$, for some $0 < \epsilon < 0.5$.

• But you do not know which is which.

• How to decide which side is the more likely side—with high confidence?

• Answer: Flip the coin many times and pick the side that appeared the most times.

• Question: Can you quantify your confidence?
The (Improved) Chernoff Bound$^a$

**Theorem 75 (Chernoff, 1952)** Suppose $x_1, x_2, \ldots, x_n$ are independent random variables taking the values 1 and 0 with probabilities $p$ and $1 - p$, respectively. Let $X = \sum_{i=1}^{n} x_i$.
Then for all $0 \leq \theta \leq 1$,

$$\text{prob}[X \geq (1 + \theta)pn] \leq e^{-\theta^2pn/3}.$$ 

- The probability that the deviate of a **binomial random variable** from its expected value $E[X] = E[\sum_{i=1}^{n} x_i] = pn$ decreases exponentially with the deviation.

$^a$Herman Chernoff (1923–). This bound is asymptotically optimal. The original bound is $e^{-2\theta^2p^2n}$ (McDiarmid, 1998).
The Proof

- Let $t$ be any positive real number.
- Then

$$\text{prob}[ X \geq (1 + \theta) pn ] = \text{prob}[ e^{tX} \geq e^{t(1+\theta) pn} ].$$

- Markov’s inequality (p. 542) generalized to real-valued random variables says that

$$\text{prob} \left[ e^{tX} \geq kE[e^{tX}] \right] \leq 1/k.$$

- With $k = e^{t(1+\theta) pn} / E[e^{tX}]$, we have\(^a\)

$$\text{prob}[ X \geq (1 + \theta) pn ] \leq e^{-t(1+\theta) pn E[e^{tX}]}.$$

\(^a\)Note that $X$ does not appear in $k$. Contributed by Mr. Ao Sun (R05922147) on December 20, 2016.
The Proof (continued)

- Because $X = \sum_{i=1}^{n} x_i$ and $x_i$’s are independent,
  $$E[e^{tX}] = (E[e^{tx_1}])^n = [1 + p(e^t - 1)]^n.$$

- Substituting, we obtain
  $$\text{prob}[X \geq (1 + \theta) pn] \leq e^{-t(1+\theta)pn} [1 + p(e^t - 1)]^n \leq e^{-t(1+\theta)pn} e^{pn(e^t-1)}$$
  as $(1 + a)^n \leq e^{an}$ for all $a > 0$. 
The Proof (concluded)

• With the choice of \( t = \ln(1 + \theta) \), the above becomes

\[
\text{prob}[X \geq (1 + \theta) pn] \leq e^{pn[\theta-(1+\theta)\ln(1+\theta)]}.
\]

• The exponent expands to\(^a\)

\[
-\frac{\theta^2}{2} + \frac{\theta^3}{6} - \frac{\theta^4}{12} + \cdots
\]

for \( 0 \leq \theta \leq 1 \).

• But it is less than

\[
-\frac{\theta^2}{2} + \frac{\theta^3}{6} \leq \theta^2 \left( -\frac{1}{2} + \frac{\theta}{6} \right) \leq \theta^2 \left( -\frac{1}{2} + \frac{1}{6} \right) = -\frac{\theta^2}{3}.
\]

\(^a\)Or McDiarmid (1998): \( x - (1 + x) \ln(1 + x) \leq -3x^2/(6 + 2x) \) for all \( x \geq 0 \).
How Good Is the Bound?

Chernoff bound

true probability

$n$
Other Variations of the Chernoff Bound

The following can be proved similarly (prove it).

**Theorem 76** *Given the same terms as Theorem 75 (p. 605),*

\[ \text{prob}[X \leq (1 - \theta) pn] \leq e^{-\theta^2 pn/2}. \]

The following slightly looser inequalities achieve symmetry.

**Theorem 77 (Karp, Luby, & Madras, 1989)** *Given the same terms as Theorem 75 (p. 605) except with \(0 \leq \theta \leq 2\),*

\[ \text{prob}[X \geq (1 + \theta) pn] \leq e^{-\theta^2 pn/4}, \]
\[ \text{prob}[X \leq (1 - \theta) pn] \leq e^{-\theta^2 pn/4}. \]
Power of the Majority Rule

The next result follows from Theorem 76 (p. 610).

**Corollary 78**  If \( p = (1/2) + \epsilon \) for some \( 0 \leq \epsilon \leq 1/2 \), then

\[
\text{prob } \left[ \sum_{i=1}^{n} x_i \leq n/2 \right] \leq e^{-\epsilon^2 n/2}.
\]

- The textbook’s corollary to Lemma 11.9 seems too loose, at \( e^{-\epsilon^2 n/6} \).

- Our original problem (p. 604) hence demands, e.g., \( n \approx 1.4k/\epsilon^2 \) independent coin flips to guarantee making an error with probability \( \leq 2^{-k} \) with the majority rule.

---

\(^{a}\)See Dubhashi & Panconesi (2012) for many Chernoff-type bounds.
BPP\textsuperscript{a} (Bounded Probabilistic Polynomial)

- The class \textbf{BPP} contains all languages \( L \) for which there is a precise polynomial-time NTM \( N \) such that:
  - If \( x \in L \), then at least \( 3/4 \) of the computation paths of \( N \) on \( x \) lead to “yes.”
  - If \( x \notin L \), then at least \( 3/4 \) of the computation paths of \( N \) on \( x \) lead to “no.”

- So \( N \) accepts or rejects by a \textit{clear} majority.

\textsuperscript{a}Gill (1977).
Magic 3/4?

- The number 3/4 bounds the probability (ratio) of a right answer away from 1/2.
- Any constant *strictly* between 1/2 and 1 can be used without affecting the class BPP.
- In fact, as with RP,
  \[
  \frac{1}{2} + \frac{1}{q(n)}
  \]
  for any polynomial \( q(n) \) can replace 3/4.
- The next algorithm shows why.
The Majority Vote Algorithm

Suppose $L$ is decided by $N$ by majority $(1/2) + \epsilon$.

1: $\textbf{for } i = 1, 2, \ldots, 2k + 1 \textbf{ do}$
2: \hspace{1em} Run $N$ on input $x$;
3: $\textbf{end for}$
4: $\textbf{if}$ “yes” is the majority answer $\textbf{then}$
5: \hspace{1em} “yes”;
6: $\textbf{else}$
7: \hspace{1em} “no”;
8: $\textbf{end if}$
Analysis

- By Corollary 78 (p. 611), the probability of a false answer is at most $e^{-\epsilon^2 k}$.
- By taking $k = \lceil 2/\epsilon^2 \rceil$, the error probability is at most $1/4$.
- Even if $\epsilon$ is any inverse polynomial, $k$ remains a polynomial in $n$.
- The running time remains polynomial: $2k + 1$ times $N$’s running time.
Aspects of BPP

- BPP is the most comprehensive yet plausible notion of efficient computation.
  - If a problem is in BPP, we take it to mean that the problem can be solved efficiently.
  - In this aspect, BPP has effectively replaced P.
- \((\text{RP} \cup \text{coRP}) \subseteq (\text{NP} \cup \text{coNP})\).
- \((\text{RP} \cup \text{coRP}) \subseteq \text{BPP}\).
- Whether \(\text{BPP} \subseteq (\text{NP} \cup \text{coNP})\) is unknown.
- But it is unlikely that \(\text{NP} \subseteq \text{BPP}\).\(^a\)

\(^a\)See p. 628.
coBPP

- The definition of BPP is symmetric: acceptance by clear majority and rejection by clear majority.
- An algorithm for $L \in \text{BPP}$ becomes one for $\overline{L}$ by reversing the answer.
- So $\overline{L} \in \text{BPP}$ and $\text{BPP} \subseteq \text{coBPP}$.
- Similarly $\text{coBPP} \subseteq \text{BPP}$.
- Hence $\text{BPP} = \text{coBPP}$.
- This approach does not work for RP.\(^a\)

\(^a\)It did not work for NP either.
BPP and coBPP
“The Good, the Bad, and the Ugly”
Circuit Complexity

- Circuit complexity is based on boolean circuits instead of Turing machines.
- A boolean circuit with \( n \) inputs computes a boolean function of \( n \) variables.
- Now, identify \texttt{true}/1 with “yes” and \texttt{false}/0 with “no.”
- Then a boolean circuit with \( n \) inputs accepts certain strings in \( \{0, 1\}^n \).
- To relate circuits with an arbitrary language, we need one circuit for each possible input length \( n \).
Formal Definitions

• The **size** of a circuit is the number of *gates* in it.

• A **family of circuits** is an infinite sequence 
\( C = (C_0, C_1, \ldots) \) of boolean circuits, where \( C_n \) has \( n \) boolean inputs.

• For input \( x \in \{0, 1\}^* \), \( C_{|x|} \) outputs 1 if and only if \( x \in L \).

• In other words,

\[
C_n \text{ accepts } L \cap \{0, 1\}^n.
\]
Formal Definitions (concluded)

- $L \subseteq \{0, 1\}^*$ has **polynomial circuits** if there is a family of circuits $C$ such that:
  - The size of $C_n$ is at most $p(n)$ for some fixed polynomial $p$.
  - $C_n$ accepts $L \cap \{0, 1\}^n$. 
Exponential Circuits Suffice for All Languages

- Theorem 16 (p. 212) implies that there are languages that cannot be solved by circuits of size $2^n/(2n)$.

- But surprisingly, circuits of size $2^{n+2}$ can solve all problems, decidable or otherwise!
Exponential Circuits Suffice for All Languages (continued)

**Proposition 79** All decision problems (decidable or otherwise) can be solved by a circuit of size \(2^{n+2}\) and depth \(2n\).

- We will show that for any language \(L \subseteq \{0, 1\}^\ast\), \(L \cap \{0, 1\}^n\) can be decided by a circuit of size \(2^{n+2}\).

- Define boolean function \(f : \{0, 1\}^n \to \{0, 1\}\), where

\[
  f(x_1x_2\cdots x_n) = \begin{cases} 
  1, & x_1x_2\cdots x_n \in L, \\
  0, & x_1x_2\cdots x_n \notin L.
  \end{cases}
\]
The Proof (concluded)

• Clearly, any circuit that implements $f$ decides $L \cap \{0, 1\}^n$.

• Now,

$$f(x_1x_2\cdots x_n) = (x_1 \land f(1x_2\cdots x_n)) \lor (\neg x_1 \land f(0x_2\cdots x_n)).$$

• The circuit size $s(n)$ for $f(x_1x_2\cdots x_n)$ hence satisfies

$$s(n) = 4 + 2s(n-1)$$

with $s(1) = 1$.

• Solve it to obtain $s(n) = 5 \times 2^{n-1} - 4 \leq 2^{n+2}$.

• The longest path consists of an alternating sequence of $\lor$s and $\land$s.