Number-Theoretic Algorithms

• What are the factors of $326,818,261,539,809,441,763,169$?

• There is no known efficient algorithm.

• What is the greatest common divisor of $835,751,544,820$ and $391,047,152,188$?

• Euclid’s algorithm solves this efficiently.

• These two facts are the basis for the RSA public-key cryptosystem.
Basic Number Theory

- **Divisibility**
  - $3 | 12$ “$3$ divides $12$”, “$12$ is a multiple of $3$”

- **Factors**
  - Factors (non-trivial divisors) of $20$ are $2, 4, 5, 10$

- **Primes**
  - $2, 3, 5, 7, 11, 13, 17, 19, 23, 29, \ldots$
  - $1$ is not prime
  - There are infinitely many primes.
Unique Factorization

- Divisibility by a prime
  - If p is prime and p \mid ab, then p \mid a or p \mid b.
- Unique factorization
  - Every integer has a unique factorization as a product of primes.
  - 5280 = 2^5 \cdot 3^1 \cdot 5^1 \cdot 11^1
Division Theorem

- For any integer \( a \) and any positive integer \( n \), there are unique integers \( q \) and \( r \), such that \( 0 \leq r < n \) and \( a = qn + r \).
- Quotient \( q \) and remainder \( r \)
- Notation: \( r = a \mod n \)
Greatest Common Divisors

• Any two integers, not both 0, have a greatest common divisor (gcd).
• \( \text{gcd}(24,30)=6 \)
• \( a, b \) are relatively prime if \( \text{gcd}(a,b)=1 \).
Euclid’s Algorithm

• For any nonnegative integer \( a \) and any positive integer \( b \),
  \[ \text{gcd}(a,b) = \text{gcd} \left( b, a \mod b \right) \]

• Euclid’s algorithm (ca. 300 B.C.)

\[
\text{EUCLID}(a,b) \\
\{ \\
\quad \text{if} \ (b = 0) \ \text{then return} \ a \\
\quad \text{else return} \ \text{EUCLID}(b, a \mod b) \\
\}
\]
Example

EUCLID(120, 23)
= EUCLID(23, 5)
= EUCLID(5, 3)
= EUCLID(3, 2)
= EUCLID(2, 1)
= EUCLID(1, 0)
= 1

So 120 and 23 are relatively prime.
Extended Euclid’s Algorithm

• Theorem 31.2: \( \gcd(a,b) \) is the smallest positive integer in the set \( \{ax+by : x,y \in \mathbb{Z} \} \).

• Euclid’s Algorithm can calculate \( x \) and \( y \) such that \( ax+by = \gcd(a,b) \).
Example

• $120 / 23 = 5 \text{ r } 5$
  - So $5 = 120 - 5 \cdot 23$

• $23 / 5 = 4 \text{ r } 3$
  - So $3 = 23 - 4 \cdot 5 = 23 - 4 \cdot (120 - 5 \cdot 23) = -4 \cdot 120 + 21 \cdot 23$

• $5 / 3 = 1 \text{ r } 2$
  - So $2 = 5 - 1 \cdot 3 = (120 - 5 \cdot 23) - 1 \cdot (-4 \cdot 120 + 21 \cdot 23)$
    = $5 \cdot 120 - 26 \cdot 23$

• $3 / 2 = 1 \text{ r } 1$
  - So $1 = 3 - 1 \cdot 2 = (-4 \cdot 120 + 21 \cdot 23) - 1 \cdot (5 \cdot 120 - 26 \cdot 23)$
    = $-9 \cdot 120 + 47 \cdot 23$
Modular Arithmetic

• We do all arithmetic modulo n.
• Powers of 3
  - 1, 3, 9, 27, 81, 243, ...
• Powers of 3 modulo 7
  - 1, 3, 2, 6, 4, 5, 1, 3, 2, 6, 4, 5, ...
• Fermat’s Theorem:
  - If p is prime and 1 ≤ a < p, then \( a^{p-1} = 1 \) (mod p).
Multiplicative Inverses

- If $a$ is relatively prime to $n$, then there exists $x$ such that $ax = 1 \pmod{n}$.
- $x$ is the multiplicative inverse of $a \pmod{n}$.
- We can find $x$ using the Extended Euclid's Algorithm.
  - $ax + ny = 1$ implies that $ax = 1 \pmod{n}$
- Example
  - The multiplicative inverse of 23 (mod 120) is 47, since $1 = -9 \cdot 120 + 47 \cdot 23$. 
Public Key Cryptography

- **Goal**: Allow users to communicate securely even if they don’t share a secret key.
- Each user publishes a **public key** and also keeps a **private key** secret.
- Anyone can encrypt a message using Alice’s public key, but only she can decrypt it, using her private key.
- Also, Alice can “sign” a message by encrypting it with her **private key**.
The RSA Cryptosystem

• Randomly choose two large primes p and q.
  - p = 835,751,544,821  q = 391,047,152,189
  - (Really p and q should be about 150 digits long.)
• Let n = pq.
  - n = 326,818,261,539,809,441,763,169
• Idea: Factoring n is hard!
• Compute $\phi(n) = (p-1)(q-1)$.
  - $\phi(n) = 326,818,261,538,582,643,066,160$
  - ($\phi(n)$ gives the number of integers less than n that are relatively prime to n.)
RSA Cryptosystem, continued

- Choose $e$ relatively prime to $\varphi(n)$.
  - $e = 3$

- Use Extended Euclid’s Algorithm to compute $d$, the multiplicative inverse of $e \pmod{\varphi(n)}$.
  - $d = 217,878,841,025,721,762,044,107$

- $(e, n)$ is the RSA public key.
- $(d, n)$ is the RSA private key.
- Encryption: $E(M) = M^e \pmod{n}$.
- Decryption: $D(C) = C^d \pmod{n}$. 
Fast Exponentiation

• Since d is huge, $C^d \mod n$ cannot be computed naïvely.
• We can do it in $2\log d$ multiplications:
  
  ```
  fun exp(C, d, n) = 
    if d = 0 then 1
    else if even(d) then 
      exp(C*C mod n, d/2, n)
    else C*exp(C, d-1, n) mod n
  ```
Correctness of RSA

- Encrypting and decrypting $M$ gives $D(E(M)) = E(D(M)) = M^{ed} \pmod{n}$.
- By the choice of $e$ and $d$, we have $ed = 1 + k(p-1)(q-1)$, for some $k$.
- Calculating mod $p$, if $M \neq 0 \pmod{p}$, then $M^{ed} = M(M^{p-1})^{k(q-1)} = M(1)^{k(q-1)} = M \pmod{p}$ using Fermat's Theorem.
- And, of course, if $M = 0 \pmod{p}$, then again $M^{ed} = M \pmod{p}$. 
Correctness of RSA, Continued

• A similar calculation shows that
  \( M^{ed} = M \pmod{q} \).

• Hence we have
  \( p \mid M^{ed} - M \) and \( q \mid M^{ed} - M \)

• Because \( \gcd(p, q) = 1 \), this implies that
  \( pq \mid M^{ed} - M \)

• So \( M^{ed} = M \pmod{n} \).
Example

- $n = 326,818,261,539,809,441,763,169$
- $e = 3$
- $d = 217,878,841,025,721,762,044,107$
- $M = 12,345,678,901,234,567,890$
- Encryption: $E(M) = M^e \mod n$
- $E(M) = 268,102,434,874,902,796,719,062$
- Decryption: $D(C) = C^d \mod n$
- $D(E(M)) = 12,345,678,901,234,567,890$
Finding Big Primes

- **Prime Number Theorem**: the number of primes less than or equal to \( n \) is about \( n/\ln n \).
- Hence a random 512-bit number is prime with probability about \( 1/\ln 2^{512} \approx 1/355 \).
- So random search will work well, if we can test for primality.
- **Randomized tests**: For example, if \( a^{n-1} \not\equiv 1 \pmod{n} \), then \( n \) cannot be prime.
- Agrawal, Kayal and Saxena found a polynomial-time algorithm in 2002!
Factoring Big Integers

• Many very sophisticated algorithms have been developed.
• But all take exponential time.
• Today, factoring an arbitrary 300-digit integer remains infeasible (apparently).