Fermat's Little Thm

RSA → Euler thm

gcd

power

m^e

Multiplicative inverse

01204211 Discrete Mathematics
Lecture 8a: Integers and GCD

Jittat Fakcharoenphol

October 1, 2024

Number theory: integers and divisibility

In the third part of the course, we study number theory, a once-thought-to-be "useless" branch of mathematics.

Number theory: integers and divisibility

In the third part of the course, we study number theory, a once-thought-to-be "useless" branch of mathematics.

Why?

Number theory: integers and divisibility

In the third part of the course, we study number theory, a once-thought-to-be "useless" branch of mathematics.

Why?

- The topic itself is very very beautiful.
- It has many applications in cryptography and error correcting codes.

Number theory: integers and divisibility

In the third part of the course, we study number theory, a once-thought-to-be "useless" branch of mathematics.

Why?

- The topic itself is very very beautiful.
- It has many applications in cryptography and error correcting codes.

We will cover:

- Basic concepts of divisibility, prime numbers, and congruence.
- How to quickly check if a number is prime.
- How to essentially perform "division" with integers, allowing us to work with important and useful objects like polynomials using only integers.

$\boxed{GF(2)}$ GF(p)

Number theory: integers and divisibility

In the third part of the course, we study number theory, a once-thought-to-be "useless" branch of mathematics.

Why?

- The topic itself is very very beautiful.
- It has many applications in cryptography and error correcting codes.

We will cover:

- Basic concepts of divisibility, prime numbers, and congruence.
- How to quickly check if a number is prime.
- How to essentially perform "division" with integers, allowing us to work with important and useful objects like polynomials using only integers.
- Applications like cryptography (RSA), secret sharing, erasure codes and error correcting codes.

Definitions

Definition (divisibility)

We say that an integer a **divides** b or b **is divisible by** a if there exist an integer k such that

$$b = ak.$$

If it is the case, we also write $a|b$. We also say that a is a **divisor** (or a **factor**) of b. On the other hand if a does not divide b, we write $a \not| b$.

Examples

If $a|b$ and $a|c$, prove that $a|(b+c)$.

Planning:

HAVE

$a|b \to \exists k \in \mathbb{Z}$ s.t. $b = a \cdot k$

$a|c \to \exists \ell \in \mathbb{Z}$ $c = a \cdot \ell$

Goal: $a|(b+c)$

WANT

$\exists x \in \mathbb{Z}$ s.t. $a \cdot x = b+c$

Proof: Because $a|b$, we know that there exists an integer k such that $b = k \cdot a$.

Also, since $a|c$, there exists an integer ℓ s.t. $c = \ell \cdot a$.

Therefore, we have that

$$b + c = k \cdot a + \ell \cdot a = a(\underbrace{k+\ell}_{\in \mathbb{Z}}).$$

However, k & ℓ are integers; thus $k+\ell$ is also an integer.

By the definition of divisibility, we can conclude that

$$a | b+c.$$

because $\exists$ an integer $x = k+\ell$, such that $ax = b+c$.

Examples

1. If $a|b$ and $a|c$, prove that $a|(b+c)$.

2. If $a|b$ and $b|c$, prove that $a|c$.

Proof: Since $a|b$, $\exists$ an integer k such that $b = k \cdot a$.

Also since $b|c$, $\exists$ an integer ℓ, s.t. $c = b \cdot \ell$. Thus

$$c = b \cdot \ell = k \cdot a \cdot \ell = (k \cdot \ell) \cdot a.$$

Since k & ℓ are integer $k\ell$ is also an integer, therefore, we know that

$$a | c.$$

$\blacksquare$

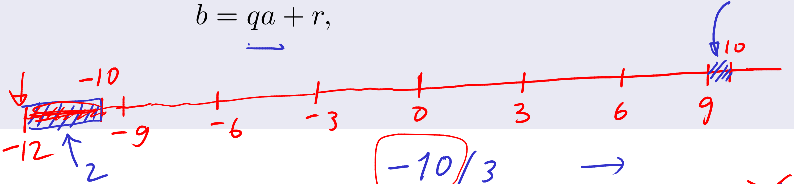
Remainder

Defintion (remainder) — positive

The **remainder** of the division of b with a is an integer r when there exists an integer q such that

$$b = qa + r,$$

where $0 \leq r < a$.



$-10/3$ → remainder ~~1~~ 2

$10 \bmod 3 = 1$

$25 \bmod 7 = 4$

$-10 \bmod 3 = 2$

$-25 \bmod 7 = 3$

Remainder

Defintion (remainder)

The **remainder** of the division of b with a is an integer r when there exists an integer q such that

$$b = qa + r,$$

where $0 \leq r < a$.
We refer to q as the **quotient** of the division.

Remainder

Defintion (remainder)

The **remainder** of the division of b with a is an integer r when there exists an integer q such that

$$b = qa + r,$$

where $0 \leq r < a$.
We refer to q as the **quotient** of the division.

Examples:

Remainder

Defintion (remainder)

The **remainder** of the division of b with a is an integer r when there exists an integer q such that

$$b = qa + r,$$

where $0 \leq r < a$.
We refer to q as the **quotient** of the division.

Examples:

We use operator `mod` to denote an operation for finding the remainder of a division.
I.e., $a \bmod b$ is the remainder of dividing a with b.

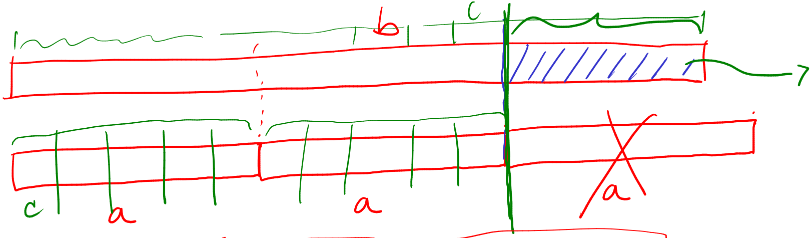
Examples

$= r = b \bmod a$.

Let r be the remainder of the division of b by a. Assume that $c|a$ and $c|b$. Prove that $c|r$.

a $\qquad$ b $\qquad$ $b \bmod a = r$

$c|a$ $\qquad$ $c|b$ $\qquad \Longrightarrow \qquad$ $c \mid b \bmod a$



$\boxed{b \bmod a}$

$= b - x \cdot a$

where $x = \left\lfloor \frac{b}{a} \right\rfloor$

Proof: Recall that $\boxed{b \bmod a} = \boxed{b - \left\lfloor \frac{b}{a} \right\rfloor \cdot a}$.

[Use the fact from exercise 5 to conclude that c divides $b - \left\lfloor \frac{b}{a} \right\rfloor \cdot a$.]

More examples

$a^2 - 1 = (a-1)\boxed{(a+1)}$

For every integer a, $a - 1 | a^2 - 1$.

Primes

$\cancel{1}$ 2

Definition (primes)

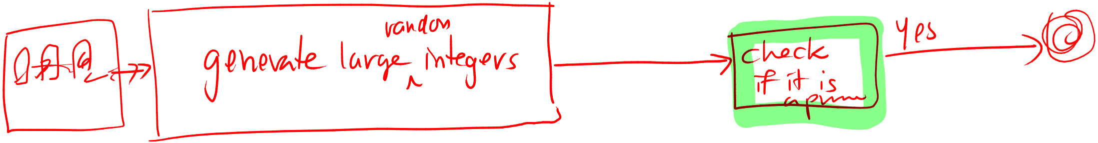
- An integer $p > 1$ is a **prime** if its divisors are only p, $-p$, 1, and -1.
- If an integer $n > 1$ is not a prime, it is called a **composite**.
- Note: 1 is not a prime and also not a composite.

Fundamental theorem of arithmetic

$$100 = 10 \cdot 10 = \underbrace{5 \cdot 2 \cdot 5 \cdot 2} = \underbrace{2 \cdot 5 \cdot 5 \cdot 2}$$

Unique factorization

Every integer greater than 1 can be represented uniquely as a product of prime numbers, up to the order of the factors.



Algorithm for testing primes

$\underbrace{2, 3, \ldots\ldots, \boxed{\sqrt{n}}}$

Recall our CheckPrime2 algorithm

```
Algorithm CheckPrime2(n):  // Input: an integer n
    if n <= 1:
        return False
    let s = square root of n
    i = 2
    while i <= s:
        if n is divisible by i:
            return False
        i = i + 1
    return True
```

$\boxed{n} = p\,q$

$\sqrt{n}$ divisions

How fast can it run?

Algorithm for testing primes

Recall our `CheckPrime2` algorithm

```
Algorithm CheckPrime2(n):  // Input: an integer n
    if n <= 1:
        return False
    let s = square root of n
    i = 2
    while i <= s:
        if n is divisible by i:
            return False
        i = i + 1
    return True
```

How fast can it run? Note that $s = \sqrt{n}$; therefore, it takes time $O(\sqrt{n})$ to run.

Efficient algorithms

RSA 2000 bits

$\sqrt{2^{2000}} \sim 2^{1000}$

Is $O(\sqrt{n})$ for checking a prime number efficient?

Efficient algorithms

Is $O(\sqrt{n})$ for checking a prime number efficient?
What is the "size" of the input to the problem?

Efficient algorithms

Is $O(\sqrt{n})$ for checking a prime number efficient?
What is the "size" of the input to the problem? The input contains one integer n; is the size of the input just 1?

Efficient algorithms

Is $O(\sqrt{n})$ for checking a prime number efficient?
What is the "size" of the input to the problem? The input contains one integer n; is the size of the input just 1?
When working with input consisting only a few numbers, we typically use the number of bits. For integer n, the number of bits of n is $\lceil \log_2 n \rceil$.

Efficient algorithms

Is $O(\sqrt{n})$ for checking a prime number efficient?
What is the "size" of the input to the problem? The input contains one integer n; is the size of the input just 1?
When working with input consisting only a few numbers, we typically use the number of bits. For integer n, the number of bits of n is $\lceil \log_2 n \rceil$.

n	number of bits of n	$\sqrt{n}$
2	1	1.414
4	2	2
16	4	4
1,024	10	32
1,048,576	20	1,024
1,125,899,906,842,624	50	33,554,432
1,267,650,600,228,229,401,496,703,205,376	100	1,125,899,906,842,624

Efficient algorithms

Is $O(\sqrt{n})$ for checking a prime number efficient?
What is the "size" of the input to the problem? The input contains one integer n; is the size of the input just 1?
When working with input consisting only a few numbers, we typically use the number of bits. For integer n, the number of bits of n is $\lceil \log_2 n \rceil$.

n	number of bits of n	$\sqrt{n}$
2	1	1.414
4	2	2
16	4	4
1,024	10	32
1,048,576	20	1,024
1,125,899,906,842,624	50	33,554,432
1,267,650,600,228,229,401,496,703,205,376	100	1,125,899,906,842,624

Side note: Recall that the first step in RSA is to find a pair of large primes. Typically we want them to be of size in the *thousand* bits.

Greatest Common Divisors (GCD)

Definition (GCD)

For integers x and y, the **greatest common divisor** (or GCD) of x and y is the largest integer g such that $g|x$ and $g|y$. We refer to it as $gcd(x, y)$.

Greatest Common Divisors (GCD)

Definition (GCD)

For integers x and y, the **greatest common divisor** (or GCD) of x and y is the largest integer g such that $g|x$ and $g|y$. We refer to it as $gcd(x, y)$.

A simple way to find $gcd(x, y)$:

```
g = min(x,y)
while (x mod g != 0) or (y mod g != 0):
  g -= 1
return g
```

Greatest Common Divisors (GCD)

Definition (GCD)

For integers x and y, the **greatest common divisor** (or GCD) of x and y is the largest integer g such that $g|x$ and $g|y$. We refer to it as $gcd(x, y)$.

A simple way to find $gcd(x, y)$:

```
g = min(x,y)
while (x mod g != 0) or (y mod g != 0):
  g -= 1
return g
```

What is the running time of this algorithm?

Greatest Common Divisors (GCD)

Definition (GCD)

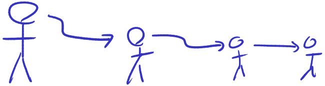
For integers x and y, the **greatest common divisor** (or GCD) of x and y is the largest integer g such that $g|x$ and $g|y$. We refer to it as $gcd(x, y)$.

A simple way to find $gcd(x, y)$:

```
g = min(x,y)
while (x mod g != 0) or (y mod g != 0):
  g -= 1
return g
```

What is the running time of this algorithm? Does it run in polynomial time on the size of the input?

Euclid's algorithm



```
Algorithm Euclid(x,y):
  if x mod y == 0:
    return y
  else:
    return Euclid(y, x mod y)
```

// $y \mid x$

x	y
12311	24324

x	y
12311	24324
24324^x	12311^y
12311	12013
12013	298
298	93
93	19
19	17
17	2
2	1

Euclid's algorithm

$gcd(x,y) = gcd(y, x \bmod y)$

```
Algorithm Euclid(x,y):
  if x mod y == 0:
    return y
  else:
    return Euclid(y, x mod y)
```

$g = gcd(x,y)$

g also divides $x \bmod y$.

57,888,139

Let's see how it works with $Euclid(12311, 24324)$:

Euclid(12311, 24324)
Euclid(24324, 12311)
Euclid(12311, 12013)
Euclid(12013, 298)
Euclid(298, 93)
Euclid(93, 19)
Euclid(19, 17)
Euclid(17, 2)
Euclid(2, 1)

275,790,979

179,638,933

93,810,803

18,037,259

$x = k \cdot y + (x \bmod y)$

Proofs

We have to prove two properties:

- For any integers x and y, $\mathrm{Euclid}(x, y) = gcd(x, y)$.
- The running time of Euclid.

Proofs

We have to prove two properties:

- For any integers x and y, $\mathrm{Euclid}(x, y) = gcd(x, y)$.
- The running time of Euclid.

Note that when $x < y$, $\mathrm{Euclid}(x, y)$ just calls itself with both arguments swapped, i.e., $\mathrm{Euclid}(y, x)$. After that, in each call, x is always larger than y. For simplicity of the analysis, we shall work only with the case that $x > y$.

Theorem 1

For any integers x and y such that $x > y$, $\mathrm{Euclid}(x, y) = gcd(x, y)$.

Proof.

We prove using strong induction. For the base case, note that when $y|x$, $gcd(x, y) = y$; therefore, the base case of the algorithm is correct.
Our induction hypothesis is: for any $x' < x$ and $y' < y$, $\mathrm{Euclid}(x', y') = gcd(x', y')$.
Now assume that $y \not| x$. The Euclid algorithm returns $\mathrm{Euclid}(y, x \bmod y)$ as the gcd. Note that $y < x$ and $x \bmod y < y$. Therefore, we can use the I.H. to claim that

$$\mathrm{Euclid}(y, x \bmod y) = gcd(y, x \bmod y).$$

Thus, we are left to show that

$$gcd(x, y) = gcd(y, x \bmod y).$$

□

What is $x \bmod y$?

What is $x \bmod y$?

Let $\lfloor a \rfloor$ be the largest integer a' such that $a' \leq \lfloor a \rfloor$.

What is $x \bmod y$?

Let $\lfloor a \rfloor$ be the largest integer a' such that $a' \leq \lfloor a \rfloor$.

$$x \bmod y = x - \left\lfloor \frac{x}{y} \right\rfloor \cdot y$$

Lemma 2

If $a|x$ *and* $a|y$, *then* $a|x \bmod y$.

Lemma 2

If $a|x$ *and* $a|y$, *then* $a|x \bmod y$.

Lemma 3

$gcd(x, y) = gcd(y, x \bmod y)$

How many recursive calls does Euclid's algorithm makes?

How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.

How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?

How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?
- When $y \leq x/2$, when Euclid(x, y) calls Euclid$(y, x \bmod y)$ the first argument decreases by half.

How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?
- When $y \leq x/2$, when Euclid(x, y) calls Euclid$(y, x \bmod y)$ the first argument decreases by half.
- How about when $y > x/2$?

How many recursive calls does Euclid's algorithm makes?

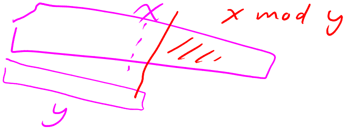
Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?
- When $y \leq x/2$, when Euclid(x, y) calls Euclid$(y, x \bmod y)$ the first argument decreases by half.
- How about when $y > x/2$?
  Euclid$(x, y) \Rightarrow$

How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?
- When $y \leq x/2$, when Euclid(x, y) calls Euclid$(y, x \bmod y)$ the first argument decreases by half.
- How about when $y > x/2$?
  Euclid$(x, y) \Rightarrow$ Euclid$(y, x \bmod y) \Rightarrow$



How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?
- When $y \leq x/2$, when Euclid(x, y) calls Euclid$(y, x \bmod y)$ the first argument decreases by half.
- How about when $y > x/2$?
  Euclid$(x, y) \Rightarrow$ Euclid$(y, x \bmod y) \Rightarrow$ Euclid$(x \bmod y, y \bmod (x \bmod y))$

$x \bmod y < x/2$

How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?
- When $y \leq x/2$, when Euclid(x, y) calls Euclid$(y, x \bmod y)$ the first argument decreases by half.
- How about when $y > x/2$?
  Euclid$(x, y) \Rightarrow$ Euclid$(y, x \bmod y) \Rightarrow$ Euclid$(x \bmod y, y \bmod (x \bmod y))$ Note that in this case, $x \bmod y = x - y \leq x/2$.

How many recursive calls does Euclid's algorithm makes?

Consider Euclid(x, y):

- If we start with $x < y$, the next calls will always have that $x > y$; so we have at most one call with $x < y$.
- When can we decrease the value of x or y in the calls?
- When $y \leq x/2$, when Euclid(x, y) calls Euclid$(y, x \bmod y)$ the first argument decreases by half.
- How about when $y > x/2$?
  Euclid$(x, y) \Rightarrow$ Euclid$(y, x \bmod y) \Rightarrow$ Euclid$(x \bmod y, y \bmod (x \bmod y))$ Note that in this case, $x \bmod y = x - y \leq x/2$. Thus, after two recursive calls, the first argument decreases by half.
- How many times can that happen?
- The first argument can decrease by a factor of two for at most $\log x$ times. Therefore, the Euclid algorithm runs in time $O(\log \max\{x, y\}) = O(\log x + \log y)$.

Computing power

- $E(m) = m^{e} \bmod n$
- $D(k) = k^{d} \bmod n$

$123^{10012132452 7}$

How fast can we compute x^y?

Computing power

How fast can we compute x^y?

```
Algorithm power(x,y):
  a = 1
  for i = 1,2,...,y:
    a *= x
  return a
```

y multiplications

Computing power

How fast can we compute x^y?

```
Algorithm power(x,y):
  a = 1
  for i = 1,2,...,y:
    a *= x
  return a
```

What is the running time?

Computing power

$x \cdot x = x^2$

$x^2 \cdot x^2 = x^4$

$x^4 \cdot x^4 = x^8$

$((((x)^2)^2)^2) \ldots$

How fast can we compute x^y?

```
Algorithm power(x,y):
  a = 1
  for i = 1,2,...,y:
    a *= x
  return a
```

What is the running time? Is it efficient?

Repeated squaring

If y is a power of two, we can find x^y using small number of multiplications using repeated squaring. E.g.,

$$x^{16} = (x^8)^2 = ((x^4)^2)^2 = (((x^2)^2)^2)^2.$$

$\log_2 16$

Repeated squaring

If y is a power of two, we can find x^y using small number of multiplications using repeated squaring. E.g.,

$$x^{16} = (x^8)^2 = ((x^4)^2)^2 = (((x^2)^2)^2)^2.$$

```
Algorithm power(x,y):   // for y=2^k
  if y == 0:
    return 1
  else:
    a = power(x, y / 2)
    return a*a
```

power(3, 64)
power(3, 32)
power(3, 16)
pw(3, 8)
pw(3, 4)
3, 2
3, 1

$(x^{y/2})^2 = x^y$

2^k

makes $\underline{k}$ calls

$\log_2 y$

Repeated squaring (general y)

```
Algorithm power(x,y):
  if y == 0:
    return 1
  else:
    a = power(x, floor(y / 2))
    if y mod 2 == 0:
      return a*a
    else
      return a*a*x
```

pow(3, 20)

$243 \cdot 243 =$ pow(3, 10)

$9 \cdot 9 \cdot 3 = 243$ pow(3, 5)

$3^2 = 9$ pow(3, 2)

3 pow(3, 1)

$3 \cdot 1 \cdot 1 = 3$

1 pow(3, 0)

3^5

$3^2 \cdot 3^2 \cdot 3$

Repeated squaring (general y)

```
Algorithm power(x,y):
  if y == 0:
    return 1
  else:
    a = power(x, floor(y / 2))
    if y mod 2 == 0:
      return a*a
    else
      return a*a*x
```

What is the number of recursive calls?

$\log_2 y$

Repeated squaring (general y)

```
Algorithm power(x,y):
  if y == 0:
    return 1
  else:
    a = power(x, floor(y / 2))
    if y mod 2 == 0:
      return a*a
    else
      return a*a*x
```

What is the number of recursive calls?
What is the running time?

Repeated squaring (general y)

x^y can be a really large number

```
Algorithm power(x,y):
  if y == 0:
    return 1
  else:
    a = power(x, floor(y / 2))
    if y mod 2 == 0:
      return a*a
    else
      return a*a*x
```

$\boxed{k^d}$ mod n

What is the number of recursive calls?
What is the running time?
While the number of multiplication is small, the numbers involved is huge as x^y has $y \log x$ bits. Computing x^y exactly definitely takes a long time.

Repeated squaring (general y, mod n)

$y = 112 \times 10 = 1{,}120$

$x = 72$

$\boxed{72^{1120} \bmod 101}$

Computing $\boxed{x^y \bmod n}$:

power(3, 1120)

pow(3, 560)

```
Algorithm power(x,y,n):
  if y == 0:
    return 1
  else:
    a = power(x, floor(y / 2)) mod n
    if y mod 2 == 0:
      return a*a mod n
    else
      return a*a*x mod n
```

$y / 2 \cdot 2 \cdot 2 \cdots$

	x^y
1120	
560	
280	
140	
70	
35	
17	$e^2 \cdot x =$
8	e
4	
2	
1	

$\frac{101}{n \ m}$

value of k such that $2^k = y$

POW(x, 1120, 101)

$\log_2 2^k = k = \log_2 y$