



离散数学

Discrete Mathematics

第九讲：数论初步

吴楠

南京大学计算机科学与技术系



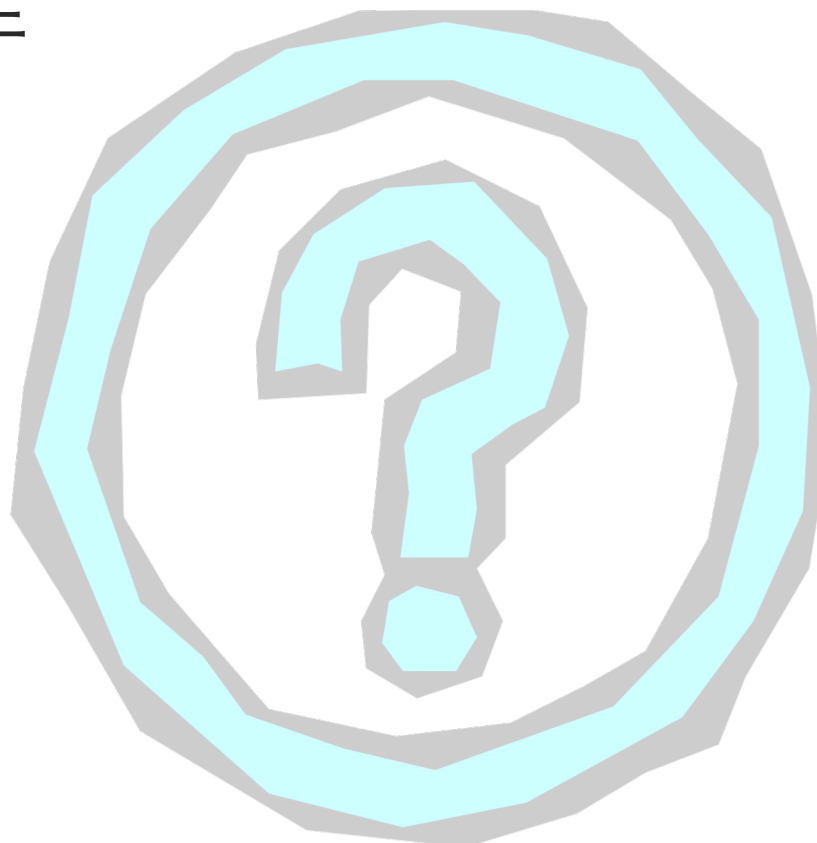
2025 年 3 月 24 日



前情提要



- 自然数与无穷公理
- 有限集与无穷集
- 集合的基数
- 集合的等势关系
- Cantor 定理
- 集合的优势关系





本讲主要内容



- 整数的性质
- 整数的基本运算
- 素 数
- Euler 函数与 Euler 定理



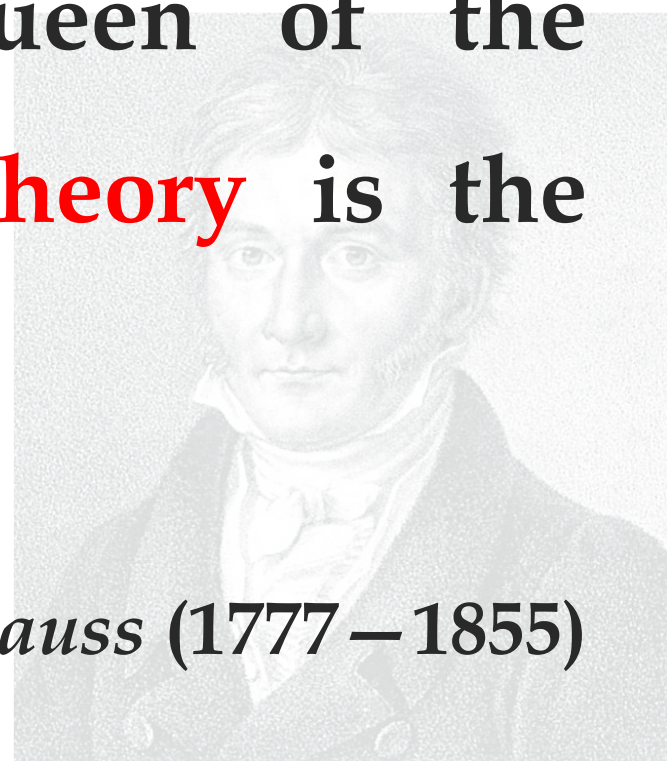


什么是数论？



Mathematics is the queen of the sciences and **number theory** is the queen of mathematics.

—— *Carl F. Gauss* (1777 – 1855)

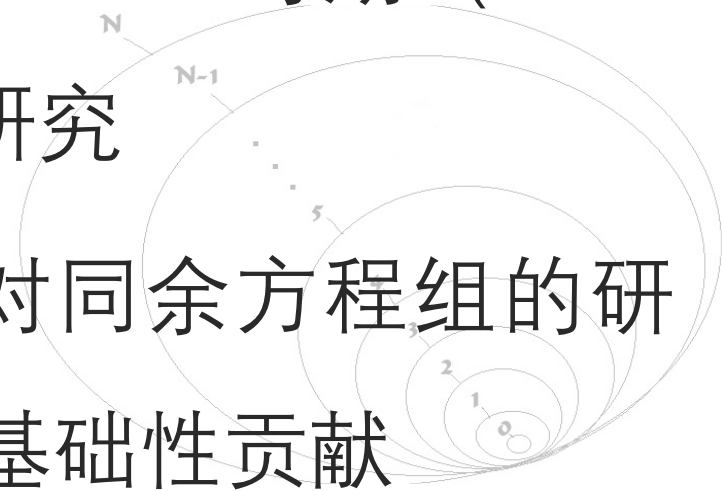




什么是数论 (续)



- 数论是纯数学的一个分支，也是纯数学的代表，它主要研究**整数**的性质
- 数论的早期研究可追溯至Euclid时期 (~300 B.C.)：对素数和整除的研究
- 中国古代 (~420 A.D.) 对同余方程组的研究为现代数论作出了部分基础性贡献





现代数论的早期铺垫



- 证明素数无穷

——Euclid: *Elements* (~300 B.C.)

- 筛法寻找素数

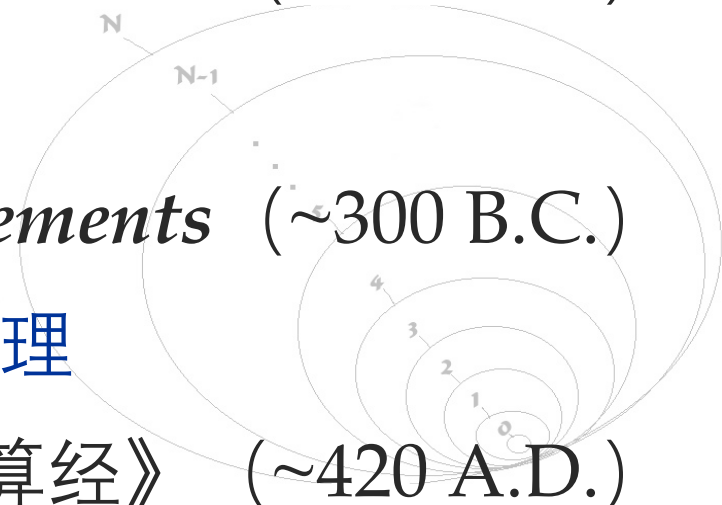
——Eratosthenes (~250 B.C.)

- 辗转相除法求最大公约数

——Euclid: *Elements* (~300 B.C.)

- 求解同余方程组的中国剩余定理

——《孙子算经》 (~420 A.D.)





整数的构造*



- **整数** (integer) 是由Peano公理衍生出来的；关于整数的算术系统也由Peano算术 (**PA**) $\langle \mathbb{N}, 0, + \rangle$ 推广而来
- 整数由 **PA** 定义的自然数构成的有序对 $(a, b) (a, b \in \mathbb{N})$ 和一个等价关系 $(a, b) \sim (c, d): a + d = b + c$ 来构造，整数集可视为上述等价关系关于 $\mathbb{N} \times \mathbb{N}$ 的一个等价类 (第十四讲详述)



整数集



- 整数集一般记为“ \mathbb{Z} ”（来源于德语“数”：*Zahlen* 的首字母），同时用 \mathbb{Z}^+ 表示正整数集（ $\mathbb{N} - \{0\}$ ），用 \mathbb{Z}^- 表示负整数集（ $\mathbb{Z} - \mathbb{N}$ ）
- \mathbb{Z} 为可列集： $\mathbb{Z} \approx \mathbb{N}$ ，基数为 \aleph_0
- \mathbb{Z} 是全序集（第十六讲详述），无上界和下界
- \mathbb{Z} 同加法运算构成一个循环群（第十七讲详述），同加法和乘法运算构成一个环（整数环^{*}）



整数的代数性质



- 下表给出 $\forall a, b, c \in \mathbb{Z}$ 关于加法和乘法的性质：

性质	加法	乘法
封闭性	$a + b$ 是整数	$a \times b$ 是整数
结合律	$a + (b + c) = (a + b) + c$	$a \times (b \times c) = (a \times b) \times c$
交换律	$a + b = b + a$	$a \times b = b \times a$
存在单位元	$a + 0 = a$	$a \times 1 = a$
存在逆元	$a + (-a) = 0$	在整数集中，只有1或 -1关于乘法存在整数逆元，其余整数 a 关于乘法的逆元为 $\frac{1}{a}$ ，都不为整数。
分配律	$a \times (b + c) = (a \times b) + (a \times c)$	

- 以下介绍数论中的一些有关整数的重要研究对象



整除



- **整除** (divisible) 是定义在 \mathbb{Z} 上的二元关系：
设 $a, b \in \mathbb{Z}, a \neq 0$, $a|b \Leftrightarrow (\exists c \in \mathbb{Z})(b = a \times c)$
- $a|b$ 读作“ a 整除 b ”
- 设 $a, b, c \in \mathbb{Z}$ 且 $a \neq 0$, 有:
 - $(a|b) \wedge (a|c) \rightarrow a|(b + c)$
 - $a|b \rightarrow a|(b \times c)$
 - $(a|b) \wedge (b|c) \rightarrow a|c$





余数



- 余数 (remainder) 来源于带余除法
- 定义 (带余除法) : 令 $a \in \mathbb{Z}, d \in \mathbb{Z}^+$, 则:
 $(\exists! q, r \in \mathbb{Z} \wedge 0 \leq r < d)(a = d \times q + r)$
 - 其中, a 称为被除数 (dividend), d 称为除数 (divisor), q 称为商 (quotient), r 称为余数
 - 记: $q = a \operatorname{div} d$, $r = a \bmod d$, 后者读作 “ a 模 d ”
- 例: $\because -11 = 3 \times (-4) + 1, \therefore -11 \bmod 3 = 1$

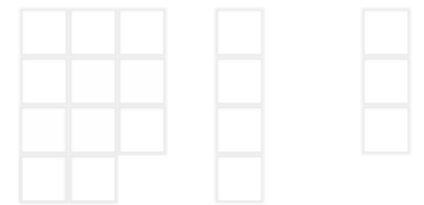


余数 (续)



- 模运算的基本性质：令 $a, b \in \mathbb{Z}, d \in \mathbb{Z}^+$ ，则：
 - $(a + b) \bmod d = (a \bmod d + b \bmod d) \bmod d$
 - $(a \times b) \bmod d = [(a \bmod d)(b \bmod d)] \bmod d$

Modulo operation



$$11 \bmod 4 = 3$$

ComputerHope.com



同余



- **同余** (congruence modulo) 是定义在 \mathbb{Z} 上的二元关系：设 $a, b \in \mathbb{Z}$,

$$a \equiv b(\text{mod } m) \Leftrightarrow (\exists m \in \mathbb{Z}^+)(m|(a - b))$$

- 上式读作“ a 与 b 模 m 同余 (a is congruent to b modulo m)”，称 m 为上述“同余的模 (modulus of the congruent)”

- 同余关系及符号“ \equiv ”由 C. Gauss 于1801年引入

- **例**： $2 \equiv 14(\text{mod } 12)$, $-5 \equiv 13(\text{mod } 6)$



素数



- 仅含2个平凡正因子（1和自身）的大于1的整数称为**素数**或者**质数**（prime number），大于1的非素数整数称为**合数**（composite number）
- **定理（算术基本定理）**：每个大于1的整数皆可分解为有限个素数的乘积（这些素数称为**素因子**）；且若不考虑顺序，则分解唯一
 - $n = p_1^{\alpha_1} p_2^{\alpha_2} \cdots p_k^{\alpha_k} \ (n > 1, p_1 < p_2 < \cdots < p_k, n, \alpha_i \in \mathbb{Z}^+)$



Eratosthenes 筛法寻找素数



■ 用Eratosthenes筛法求素数 (25以内)

2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

[2] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

[3] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

[5] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25



Eratosthenes 筛法寻找素数



■ 用Eratosthenes筛法求素数演示 (120以内)

	2	3	4	5	6	7	8	9	10	Prime numbers
11	12	13	14	15	16	17	18	19	20	
21	22	23	24	25	26	27	28	29	30	
31	32	33	34	35	36	37	38	39	40	
41	42	43	44	45	46	47	48	49	50	
51	52	53	54	55	56	57	58	59	60	
61	62	63	64	65	66	67	68	69	70	
71	72	73	74	75	76	77	78	79	80	
81	82	83	84	85	86	87	88	89	90	
91	92	93	94	95	96	97	98	99	100	
101	102	103	104	105	106	107	108	109	110	
111	112	113	114	115	116	117	118	119	120	



素数 (续)



- 关于素数的命题可追溯至Euclid时期，最著名的命题之一为《几何原本》所提：若 $2^p - 1$ 为素数，则 $2^{p-1}(2^p - 1)$ 为完全数*（本身为其所有真因子之和的数）
- 对 $n \in \mathbb{Z}^+$ ，整数 $M_n = 2^n - 1$ 被称为Mersenne数，当 n 为合数时 M_n 必为合数，但当 n 为素数时 M_n 未必——甚至极少——为素数。对某素数 p ，若 M_p 为素数，则称 M_p 为Mersenne素数*





素数 (续)



- 截至今日，人类共发现52个Mersenne素数
 - M_2, M_3, M_5, M_7 于公元前被发现
 - 前12个Mersenne素数发现于手算时代
 - 在1952—1994年的计算机时代，发现了第13—34个Mersenne素数
 - 在1996年至今，互联网时代的分布式大规模计算（如GIMPS项目）发现了第35—51个Mersenne素数（但不知道第45到第51个之间是否还有其它Mersenne素数）
 - 目前（发现于2024年10月21日）已知最大的第52个Mersenne素数是 $2^{136279841} - 1$ ，它有41024320位





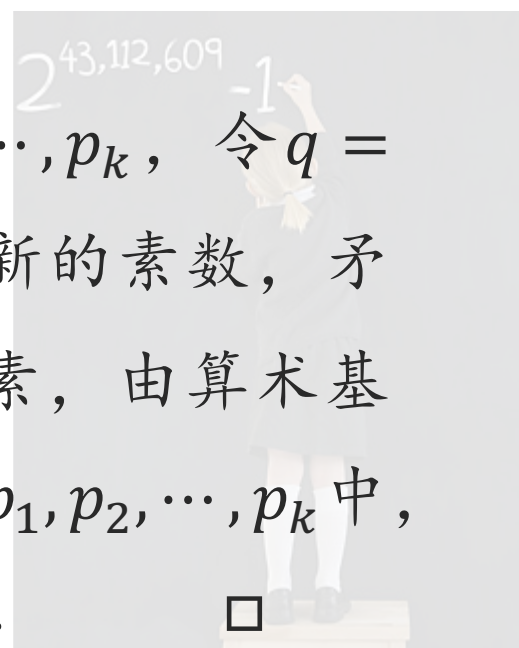
素数的性质



■ **命题：** 若 n 为合数，则其必含有不大于 \sqrt{n} 的素因子

■ **命题 (Euclid)：** 素数非有穷

○ **证明：** 反设素数有穷，列为 p_1, p_2, \dots, p_k ，令 $q = \prod_{i=1}^k p_i + 1$ ，则若 q 为素数，则其为新的素数，矛盾；若 q 为合数，因为 $\prod_{i=1}^k p_i$ 与 q 互素，由算术基本定理， q 的分解式中的素数均不在 p_1, p_2, \dots, p_k 中，为新的素数，亦矛盾。故原命题成立。 □





素数的性质 (续)



■ **定理*** (**素数定理**) : 设 $x \in]$

计数函数 (*i.e.* 不大于 x 的素数

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x / \ln x} = 1$$

○ 素数定理表明从不大于 n 的自然数中随
的概率约为 $1/\ln n$

○ 素数的分布随着 n 的增大 **逐渐稀疏**

○ 孪生素数猜想 (twin prime conjecture,

$$\liminf_{n \rightarrow \infty} (p_{n+1} - p_n) = 2$$

x	$\pi(x)$ ^[1]	$\pi(x) - \frac{x}{\ln x}$ ^[2]	$\frac{\pi(x)}{\frac{x}{\ln x}}$
10	4	-0.3	0.921
10^2	25	3.3	1.151
10^3	168	23	1.161
10^4	1,229	143	1.132
10^5	9,592	906	1.104
10^6	78,498	6,116	1.084
10^7	664,579	44,158	1.071
10^8	5,761,455	332,774	1.061
10^9	50,847,534	2,592,592	1.054
10^{10}	455,052,511	20,758,029	1.048
10^{11}	4,118,054,813	169,923,159	1.043
10^{12}	37,607,912,018	1,416,705,193	1.039
10^{13}	346,065,536,839	11,992,858,452	1.034
10^{14}	3,204,941,750,802	102,838,308,636	1.033
10^{15}	29,844,570,422,669	891,604,962,452	1.031
10^{16}	279,238,341,033,925	7,804,289,844,393	1.029
10^{17}	2,623,557,157,654,233	68,883,734,693,281	1.027
10^{18}	24,739,954,287,740,860	612,483,070,893,536	1.025
10^{19}	234,057,667,276,344,607	5,481,624,169,369,960	1.024
10^{20}	2,220,819,602,560,918,840	49,347,193,044,659,701	1.023
10^{21}	21,127,269,486,018,731,928	446,579,871,578,168,707	1.022
10^{22}	201,467,286,689,315,906,290	4,060,704,006,019,620,994	1.021
10^{23}	1,925,320,391,606,803,968,923	37,083,513,766,578,631,309	1.020



素数的性质 (续)



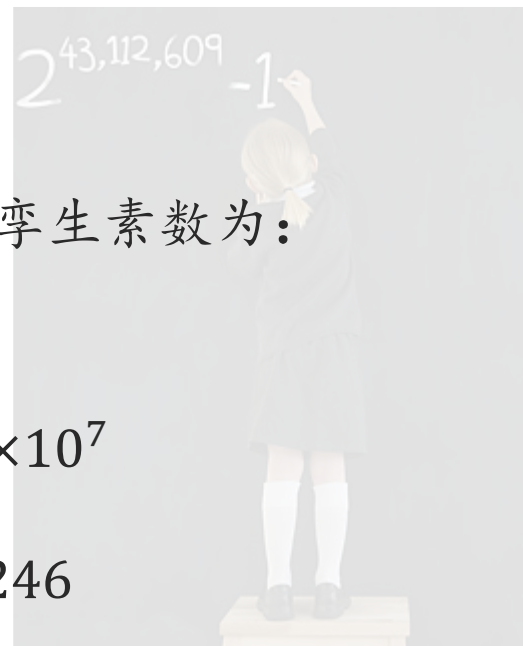
- **定理*** (**素数定理**) : 设 $x \in \mathbb{R}^+$, $\pi(x)$ 为素数计数函数 (*i.e.* 不大于 x 的素数的计数), 有

$$\lim_{x \rightarrow \infty} \frac{\pi(x)}{x / \ln x} = 1$$

- 截至目前 (发现于2016年9月), 发现的最大孪生素数为:

$$2996863034895 \times 2^{1290000} \pm 1$$

- 张益唐 (2013年5月): $\liminf_{n \rightarrow \infty} (p_{n+1} - p_n) \leq 7 \times 10^7$
- 陶哲轩 (2014年10月): $\liminf_{n \rightarrow \infty} (p_{n+1} - p_n) \leq 246$





最大公约数



- 设 $a, b \in \mathbb{Z}^+$ 且 $a \neq 0$ 或者 $b \neq 0$ ，可同时整除 a, b 的最大正整数称为 a 与 b 的 **最大公约数**（greatest common divisor, GCD），记为：

$$\gcd(a, b) = \max\{d \in \mathbb{Z}^+ \mid (d \mid a) \wedge (d \mid b)\}$$

- 称 $a, b \in \mathbb{Z}^+$ **互素**（mutually prime, coprime） \Leftrightarrow

$$\gcd(a, b) = 1 \quad (\text{通常简记为 } (a, b) = 1)$$

$$\begin{array}{l} 2 \mid 3, 6, 12, 8 \\ 2 \mid 3, 5, 4 \\ 3 \mid 3, 6, 2 \\ 2 \mid 1, 2 \\ \hline \text{LCM}(3, 6, 12, 8) \\ = 2 \times 2 \times 3 \times 1 \times 1 \times 1 \times 2 = 24 \end{array}$$





最大公约数的性质



- 定理 (**Bézout's identity**) : 设 $a, b \in \mathbb{Z}^+$, 则:

$$(\exists s, t \in \mathbb{Z})(\gcd(a, b) = sa + tb)$$

- 定理 (**辗转相减算法**) : 设 $a, b \in \mathbb{Z}^+, a < b$, 则:

$$\gcd(a, b) = \gcd(a, b - a)$$

- 定理 (**辗转相除算法**) : 设 $a, b \in \mathbb{Z}^+, a > b$, 则:

$$\gcd(a, b) = \gcd(b, a \bmod b)$$



求最大公约数的Euclid算法



```
function gcd(a, b) // a>0, b>0
  while a ≠ b
    if a > b
      a := a - b
    else
      b := b - a
  return a
```

```
function gcd(a, b) // 非全0正整数
  while b ≠ 0
    t := b
    b := a mod b
    a := t
  return a
```

```
function gcd(a, b) // a≥b≥0, a>0
  if b=0
    return a
  else
    return gcd(b, a mod b)
```



“欧几里得算法是所有算法的鼻祖，因为它是现存最古老的非凡算法。”

——高德纳，《计算机程序设计艺术，第二卷：半数值算法》，第二版（1981），p. 318.



中国剩余定理（孙子定理）

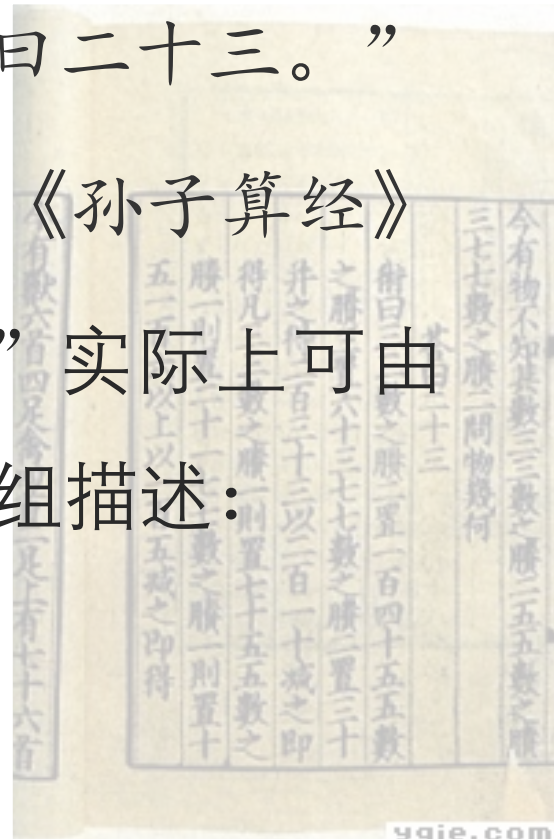


- “今有物不知其数：三三数之剩二，五五数之剩三，七七数之剩二。问物几何？答曰二十三。”

——《孙子算经》

- 上述问题中的三个“ $\times \times$ 数之剩几”实际上可由三个一元线性同余方程组成的方程组描述：

$$\begin{cases} x \equiv 2 \pmod{3} \\ x \equiv 3 \pmod{5} \\ x \equiv 2 \pmod{7} \end{cases}$$





中国剩余定理 (续)



- 一元线性同余方程组可写为：

$$\begin{cases} x \equiv a_1 \pmod{m_1} \\ x \equiv a_2 \pmod{m_2} \\ \vdots \\ x \equiv a_n \pmod{m_n} \end{cases} \quad \begin{cases} x \equiv 2 \pmod{3} \\ x \equiv 3 \pmod{5} \\ x \equiv 2 \pmod{7} \end{cases}$$

- 定理 (线性同余方程组的解存在定理)：设正整数 m_1, m_2, \dots, m_n 两两互素，则一元线性同余方程组有解 $x = \sum_{i=1}^n a_i t_i M_i$ ，且解在模 M 同余下唯一。其中 $M = \prod_{i=1}^n m_i$ ， $M_i = M/m_i$ ， $t_i M_i \equiv 1 \pmod{m_i}$ ， $i = 1, 2, \dots, n$ 。 t_i 称为 M_i 的“数论倒数（模逆元）”。该定理的证明请见【Rosen】p. 235



欧拉函数



- **定义 (欧拉函数, Euler's totient function)** : 对任意 $n \in \mathbb{Z}^+$,

$$\varphi(n) = |\{m \in \mathbb{Z}^+ | m \leq n \wedge (m, n) = 1\}|$$

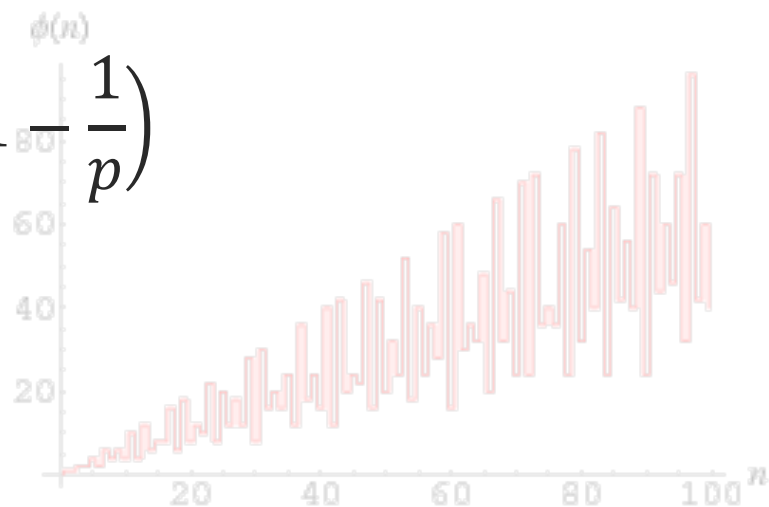
- **例:** $\varphi(3) = 2$, $\varphi(4) = 2$, $\varphi(12) = 4$

- 由容斥原理 (容斥原理相关内容请根据课件自学) 可证:

$$\varphi(n) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$

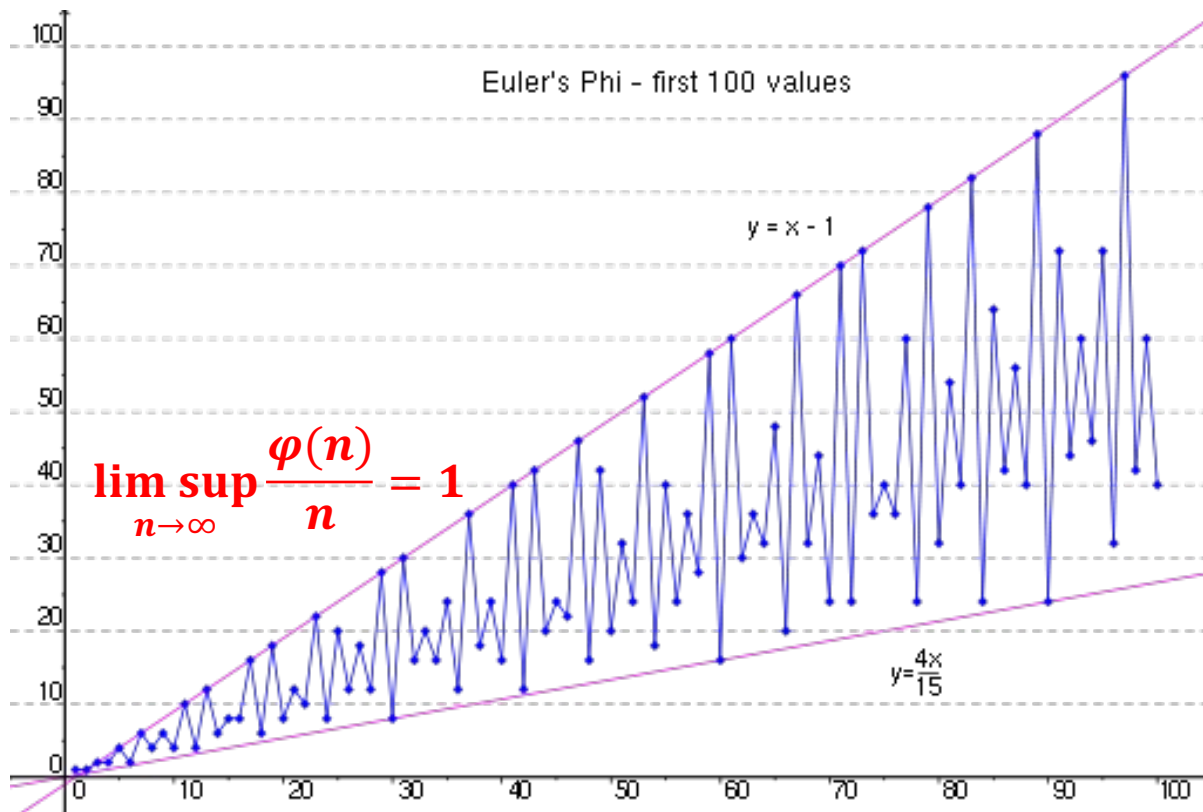
其中 $\{p\}$ 为 n 的所有素因子的集合

- $(m, n) = 1 \rightarrow \varphi(mn) = \varphi(m)\varphi(n)$
- p 为素数 $\rightarrow \varphi(p) = p - 1$





欧拉函数的界*



$$\liminf_{n \rightarrow \infty} \frac{\varphi(n)}{n} \log \log n = e^{-\gamma} \quad (\text{欧拉常数 } \gamma \approx 0.5772)$$

Eulers Totient

Eulers Totient

Number:

100

Generate Exit

1	2	3	7	8	9
4	5	6	0	del	

Eulers Totient 1.0



欧拉定理



- **定理 (Euler定理)** : 对 $a, n \in \mathbb{Z}^+$, 若 $(a, n) = 1$, 则:

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

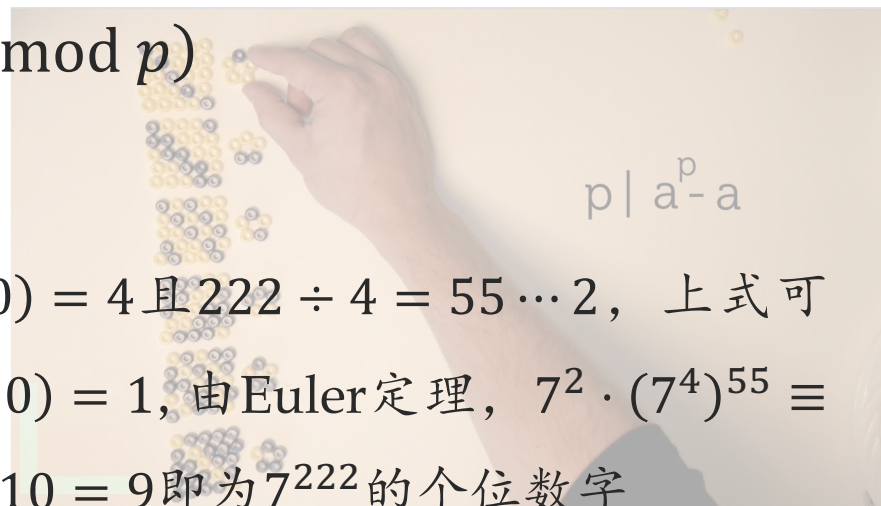
- 若上述 $n \in \mathbb{Z}^+$ 为素数, 由欧拉函数的性质易得到:

- **定理 (Fermat小定理)** : 设正整数 a 和素数 p , 且 $p \nmid a$, 则:

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

- **例**: 求 7^{222} 的个位数字

- **解**: 即求 $7^{222} \bmod 10$, 因 $\varphi(10) = 4$ 且 $222 \div 4 = 55 \cdots 2$, 上式可写为 $7^2 \cdot (7^4)^{55} \bmod 10$ 。因 $(7, 10) = 1$, 由Euler定理, $7^2 \cdot (7^4)^{55} \equiv 7^2 \cdot 1^{55} \pmod{10}$, 故 $7^{222} \bmod 10 = 9$ 即为 7^{222} 的个位数字





RSA算法的数论基础*



- RSA算法为目前最著名的公开密钥加密算法（属非对称加密算法）
- RSA算法于1977年被MIT的 R. Rivest, A. Shamir 和 L. Adelman 共同提出
- RSA算法被广泛应用于加密通信、金融信息安全、电子商务和数字签名等领域





RSA算法的数论基础*



■ RSA算法理论：

- **公钥与私钥的产生：** (1) 随意选择两个大的素数 p 和 q ($p \neq q$)，计算 $N = pq$ ；
(2) 根据欧拉函数，求得 $r = \varphi(p)\varphi(q) = (p-1)(q-1)$ ； (3) 选择一个小于 r 的整数 e ，求得 e 关于 r 的模逆元（数论倒数） $d \equiv e^{-1} \pmod{r}$ （模逆元存在 $\Leftrightarrow (e, r) = 1$ ）； (4) 将 p 和 q 的记录销毁。令 (N, e) 是公钥， (N, d) 是私钥。
Alice将她的公钥传给Bob，而将她的私钥隐藏
- **加密：** 将要加密传输的消息编码为 n （如采用Unicode），然后计算 $n^e \equiv c \pmod{N}$ ，Bob算出 c 后就可将它传递给Alice
- **解密：** Alice得到Bob的消息 c 后就可以利用她的密钥 d 来解密。她可以用以下公式来将 c 转换为 n ： $c^d \equiv n \pmod{N}$ ，得到 n 后，她便可以将消息复原

■ RSA算法的安全性源自factoring问题的时间复杂性 (BQP)



埃拉托色尼 (Eratosthenes 276-194, B.C.)



- It is known that Eratosthenes was born in Cyrene, a Greek colony west of Egypt, and spent time studying at Plato's Academy in Athens. We also know that King Ptolemy II invited Eratosthenes to Alexandria to tutor his son and that later Eratosthenes became chief librarian at the famous library at Alexandria, a central repository of ancient wisdom. Eratosthenes was an extremely versatile scholar, writing on mathematics, geography, astronomy, history, philosophy, and literary criticism.



—— *Rosen: Discrete Mathematics and Its Applications*

- 埃拉托色尼在人类历史上第一个对宇宙距离进行了科学测量。大约在公元前240年，他测的地球的直径约为12800千米，周长约为40000千米。可是这个几乎准确的数值没有被人们广泛接受。1700多年以后的哥伦布仍然相信比埃拉托色尼晚100多年的另一位希腊天文学家重测的错误数据——周长约28800千米。如果哥伦布知道地球的真实大小，也许就不敢如此冒险了。

——[美]阿西莫夫：《阿西莫夫最新科学指南》



Tips: 埃拉托色尼



本次课后作业



- 教材内容：[Rosen] 4.1节，4.3节，4.4节；8.5.2节（自学内容）
- 课后习题：
 - Problem Set 9
- 提交时间：4月1日 10:00 前

