数论基础

回顾

问题1：什么是集合的基数?

- 有限集合的基数就是一个自然数，表示集合的元素个数
- 无限集合中最小的那一个就是自然数集合，它的基数是$\aleph_0$
- 我们也可以判断一个集合是有限还是无限

问题2：(基数)最大的集合有多大?

- 实数集比自然数集大
- 任意集合的幂集比自身大

问题3：如何比较两个(无限)集合的基数?

- 利用双射函数证明等势
- 利用Bernstein定理证明等势

本节提要

问题1：什么是(初等)数论?
问题2：素数有哪些性质?

现代数论的早期铺垫

- 证明质数无穷

——Euclid：*Elements* （~300 A.D.）

- 筛法寻找质数

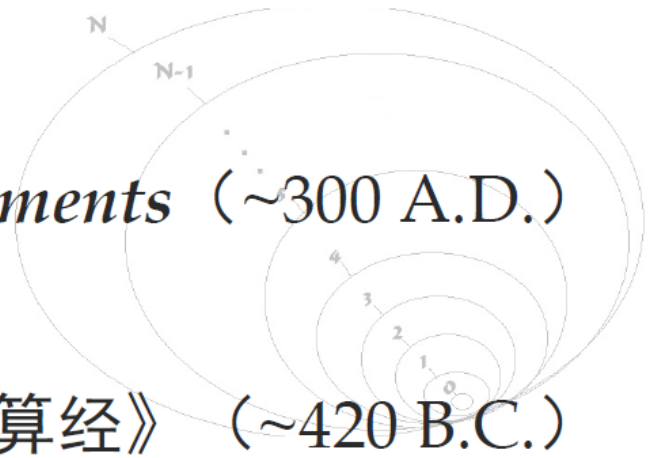
—— Eratosthenes （~250 A.D.）

- 辗转相除法求最大公约数

——Euclid： *Elements*（~300 A.D.）

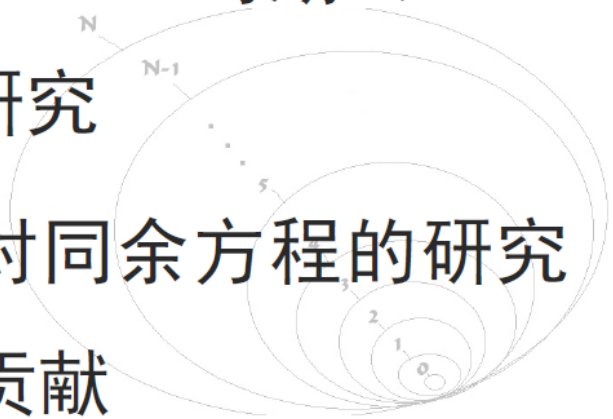
- 求解同余方程的中国剩余定理

——《孙子算经》 （~420 B.C.）



什么是数论

- 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.）：对质数和整除的研究
- 中国古代（~400 A.D.）对同余方程的研究为现代数论作出了基础性贡献



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（未来课程详述），无上界和下界
- $\mathbb{Z}$和加法运算形成一个循环群（未来课程详述）；和加法运算及乘法运算形成一个环（参见抽象代数资料*）

整数的代数性质

	Addition	multiplication
Closure Property	a +b=an integer example 6+2=8	ax b=an integer example 3x4=12
Associative	a+(b +c)=(a +b)+c example 9+(2+4)=(9+2)+4=15	ax(b x c)=(ax b)x c example 3x[-2)x4]= [3x(-2)]x4=-24
Distributive	ax(b +c)=(ax b)+(a x c), (a +b) x c=(a x c)+(b x c) example5x[2+(-3)]=[5x2]+[5x(-3)]=10-15=-5	
Commutative	a+ b=b +a example 3+(-2)=(-2)+3=1	ax b=b x a example 3x(-2)=(-2)x3=-6
Identity	a+0=a example 6+0=6	ax1=a example (-6)x1=-6
Inverse element	a+(-a)=0 example 6+(-6)=0	No inverse element
Zero product property		If a x b=0,then either a=0, or b=0 or both=0

整除

- 对任意整数a和b, $a \neq 0$, 我们说a整除b (记作$a|b$) , 如果存在整数c使得 $b = a\,c$.

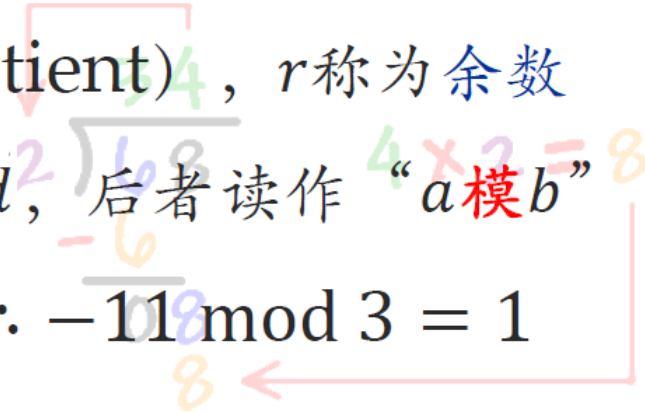
- 设a, b和c是整数， $a \neq 0$,
  - 若$a|b$，且 $a|c$，则 $a|(b+c)$
  - 若$a|b$，则 $a|(b\,c)$
  - 若$a|b$，且 $b|c$，则 $a|c$

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$b$”
- 例：$\because -11 = 3 \times (-4) + 1, \therefore -11 \bmod 3 = 1$



同余算术 (高斯, Gauss)

- 设a和b为整数，m为正整数，如果m整除$(b-a)$，就说a模m同余b. 记作 $a \equiv b \pmod{m}$.
- $a \equiv b \pmod{m}$ *iff* $a \pmod{m} = b \pmod{m}$
- $a \equiv b \pmod{m}$ *iff* $\exists k \in \mathbf{Z}.\ a = b + km$
- 举例 $-1 \equiv 5 \pmod 6$, $-2 \equiv 4 \pmod 6$, …, $-5 \equiv 1 \pmod 6$
  - $[0] = \{\ldots, -6, 0, 6, \ldots\}$
  - $[1] = \{\ldots -5, 1, 7, \ldots\}$
  - $[2] = \{\ldots, -4, 2, 8, \ldots\}$
  - …

同余算术

- 令a和b为整数，d为正整数，则
  - $(a + b) \bmod d = (a \bmod d + b \bmod d) \bmod d$.
  - $(a\ b) \bmod d = ((a \bmod d)\ (b \bmod d)) \bmod d$.

- 模m算术：令Z_m表示小于m的非负整数集合，定义$+_m$ / $\cdot_m$为模m加法/乘法
  - $a +_m b = (a+b) \bmod m$
  - $a \cdot_m b = (a \cdot b) \bmod m$

本节提要

问题1：什么是(初等)数论?

- 研究整数的性质：整除、余数、同余算术

问题2：素数有哪些性质?

素数（Prime）

- 大于1的正整数p称为素数，如果p仅有的正因子是1和p。大于1又不是素数的正整数称为合数。
- 正整数n是合数 iff $\exists a \in \mathbf{N}.\ 1 < a < n$, 且 $a \mid n$.
- 算术基本定理：每个大于1的正整数都可以唯一地写为一个素数或者若干个素数的乘积，其中素数因子以非递减序出现。
  - $n = p_1^{\alpha 1} p_2^{\alpha 2} \ldots p_k^{\alpha k}$
- 素数举例：2, 3, 5, 7, 11, 13, 17, 19, …
- 合数举例：$100 = 2^2 5^2$. $999 = 3^3 37$, $1024 = 2^{10}$.

算术基本定理的证明(存在性)

- 大于1的自然数必可写成素数之积
- 用反证法：
  - 假设存在大于1的自然数不能写成质数的乘积，把最小的那个称为n。
  - 大于1的自然数可以根据其可除性（是否能表示成两个不是自身的自然数的乘积）分成2类：质数、合数。
  - n不是质数：质数p可以写成质数乘积：p=p，这与假设不相符合。
  - n只能是合数：每个合数都可以分解成两个严格小于自身而大于1的自然数的积。设n = a * b，其中 a 和 b 都是介于1和n之间的自然数，因此，按照n的定义，a和b都可以写成质数的乘积。从而n = a * b也可以写成质数的乘积。由此产生矛盾。

算术基本定理的证明(唯一性)

- 引理：若质数p|ab，则不是 p|a，就是p|b。

证明：

(1)若p|a 则证明完毕。

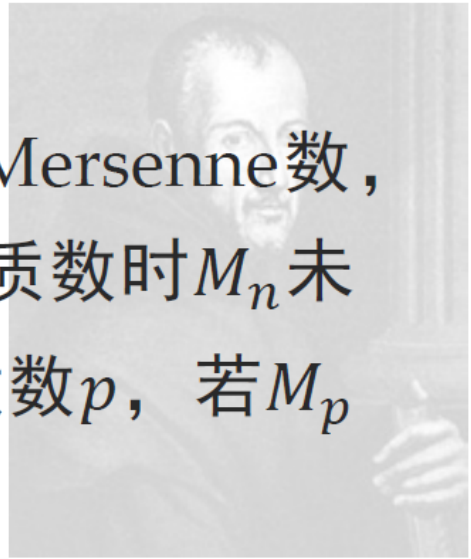
(2)若非p|a，那么两者的最大公约数为1。根据裴蜀(Bézout)定理，存在(m, n) 使得ma + np =1。于是b = b(ma + np) = abm + bnp。 由于p|ab，上式右边两项都可以被p整除。所以p|b。

算术基本定理的证明(唯一性)

- 再用反证法：
  - 假设有些大于1的自然数可以以多于一种的方式写成多个质数的乘积，那么假设n是其中最小的一个。
  - 首先n不是质数。将n用两种方法写出：
    - $n = p_1 * p_2 * p_3 * \cdots * p_r$
    - $n = q_1 * q_2 * q_3 * \cdots * q_s$
  - 根据引理，质数$p_1 | q_1 * q_2 * q_3 * \cdots * q_s$，所以$q_i$中有一个能被$p_1$整除，不妨设为$q_1$。但$q_1$也是质数，因此$q_1 = p_1$。
  - 所以，比n小的正整数n'$= p_2 * p_3 * \cdots * p_r$也可以写成$q_2 * q_3 * \cdots * q_s$。这与n的最小性矛盾！

梅森素数

- 关于质数的命题可追溯到Euclid时期，最著名的命题之一为《几何原本》所提之：若2^p-1为质数，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为质数时M_n未必——甚至极少——为质数。对某质数p，若M_p为质数，则称M_p为Mersenne质数



例：n是合数，则M_n也是合数

若n不是质数,则存在大于1,小于n的两个正整数a,b满足 $n=ab$.

于是

$2\text{^}n\text{-}1$

$=2\text{^}(ab)\text{-}1$

$=(2\text{^}a)\text{^}b\text{-}1$ (令$y=2\text{^}a$)

$=y\text{^}b\text{-}1$

$=(y\text{-}1)(y\text{^}(b\text{-}1)+y\text{^}(b\text{-}2)+\ldots+y+1)$

梅森素数

- 前四个梅森素数M2、M3、M5、M7在公元前就已经知道
- 前12个梅森素数在手算时代发现
- 1952-1996年发现了前34个梅森素数
- 往后的梅森素数都是由因特网梅森素数大搜索（GIMPS）分布式计算项目发现
- 目前，2018年12月发现第51个：$M_{82589933}$

埃拉托色尼筛选法(Eratosthenes, BC276–195)

- 用筛选法求质数（以25以内的为例）

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

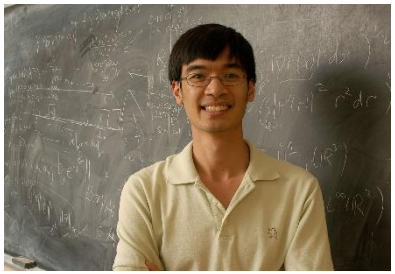
[5] 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25

素数的性质

- 下列自然数哪些是素数？
  - 101
  - 2^2-1，2^3-1，2^5-1，…，2^p-1,…
  - 2^{11}-1=2047=23·89 （大数的因数分解有点难）
  - （搜寻尽可能大的梅森素数）
- 如果n是合数，那么n必有不大于$\sqrt{n}$的一个素因子.
- 存在无限多个素数
  - 证明. 反证法，假设只有有限个素数，$p_1, p_2 \ldots, p_k$
  - 令$q=1+p_1p_2 \ldots p_k$，q的素因子是新的素数，矛盾。

素数的性质

- 任意给定K，存在K个成等差级数的素数(陶哲轩,格林, 2004)
  - 举例：当K=3时，我们有3, 7, 11。



- 任一大于2的偶数都可以写成2个素数之和?
  - 1+1（哥德巴赫猜想，1742）
  - 1+2（陈景润证明，1966）

- 素数的分布?
  - 无穷多个“特殊形式的素数”，比如：搜寻尽可能大的梅森素数。
  - 素数定理：不超过n的素数有多少个？接近于$n / \ln n$ (n充分大时)

张益唐 与 孪生素数猜想

生于1955-

庾信平生最萧瑟，
暮年诗赋动江关

2013：存在无穷多个素数对相差都小于**7000**万

$$\liminf_{n\to\infty}(p_{n+1}-p_n)<7\times10^7$$

最大公约数

- 能整除两个（正）整数的最大正整数称为这两个正整数的最大公约数。记法：$\gcd(a, b)$
  - $\gcd(a, b) = max\{ d \in N^+ \mid d|a, d|b\}$，$a \neq 0$ 或者 $b \neq 0$
  - 我们称a和b是互素的，如果$\gcd(a, b) = 1$

- 若$a = p_1^{\alpha 1} p_2^{\alpha 2} \ldots p_k^{\alpha k}$，$b = p_1^{\beta 1} p_2^{\beta 2} \ldots p_k^{\beta k}$，

  则 $\gcd(a, b) = p_1^{\gamma 1} p_2^{\gamma 2} \ldots p_k^{\gamma k}$，$\gamma_i = \min\{\alpha_i, \beta_i\}$

最大公约数的性质

- 定理（线性合成）：设$a, b \in \mathbb{Z}^+$，则：

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

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

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

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

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

欧几里德算法（求最大公约数）

```
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)  // 不全为0的自然数
  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)
```

最大公约数的性质（续）

- **gcd(a, b)一定是a和b的线性组合，即：**

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

- If d is $GCD(a, b)$, then $d = sa + tb$ for some integer s and t.
  - Let x be the smallest positive integer that can be written as $sa + tb$. For any common divisor c of a, b, $c|(sa + tb)$, which means that x is no less than any common divisor of a, b.
  - Let $a = qx + r$ $(0 \leq r < x)$, then $r = a - q(sa + tb) = (1 - qs)a - qtb$. Since r is also of the form of $sa + tb$, r can not be positive, and must be 0. So, $a = qx$, that is, $x|a$. Similarly, $x|b$.
  - Conclusion: $x = sa + tb$ is the largest common divisor of a and b. And it is a multiple of any other common divisors.

最大公约数的性质（续）

- **gcd(a, b)**一定是a和b的线性组合，即：

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

- 裴蜀(Bézout)定理：非零整数a和b是互素的 *iff* $\exists s, t \in \mathbf{Z}.\ sa + tb = 1$
  - 必要性显然。
  - 以下证明其充分性。假设$\exists s, t \in \mathbf{Z}.\ sa + tb = 1$.
    - 假设$\gcd(a, b) = d$，$\exists a_1, b_1 \in \mathbf{Z}.\ a = a_1 d,\ b = b_1 d$.
    - 我们有$sa_1 d + tb_1 d = 1$. 即 $(sa_1 + tb_1)d = 1$.
    - 因此$d = 1$. 即$\gcd(a, b) = 1$。

例

If a, b, and c are positive integers such that $\gcd(a, b) = 1$ and $a \mid bc$, then $a \mid c$.

Proof: Because $\gcd(a, b) = 1$, by Bézout's theorem there are integers s and t such that

$$sa + tb = 1.$$

Multiplying both sides of this equation by c, we obtain

$$sac + tbc = c.$$

We can now use Theorem 1 of Section 4.1 to show that $a \mid c$. By part (ii) of that theorem, $a \mid tbc$. Because $a \mid sac$ and $a \mid tbc$, by part (i) of that theorem, we conclude that a divides $sac + tbc$. Because $sac + tbc = c$, we conclude that $a \mid c$, completing the proof. ◁

If p is a prime and $p \mid a_1 a_2 \cdots a_n$, where each a_i is an integer, then $p \mid a_i$ for some i.

中国剩余定理（孙子算经，5世纪）

$$(S):\ \begin{cases} x \equiv a_1 \pmod{m_1} \\ x \equiv a_2 \pmod{m_2} \\ \quad \vdots \\ x \equiv a_n \pmod{m_n} \end{cases}$$

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’。

- 假设正整数$m_1, m_2, \ldots, m_n$两两互素，一元线性同余方程组（S）有解，在模M同余下是唯一的。

$$M = m_1 \times m_2 \times \cdots \times m_n = \prod_{i=1}^{n} m_i \qquad M_i = M/m_i,\ \forall i \in \{1,2,\cdots,n\}$$

$$t_i M_i \equiv 1 \pmod{m_i},\ \forall i \in \{1,2,\cdots,n\}.$$

$$x = \sum_{i=1}^{n} a_i t_i M_i.$$

中国剩余定理（孙子算经，5世纪）

$$(S): \quad \begin{cases} x \equiv 2 \pmod 3 \\ x \equiv 3 \pmod 5 \\ x \equiv 2 \pmod 7 \end{cases}$$

$$70 = 2\times 35 \equiv \begin{cases} 1 \pmod 3 \\ 0 \pmod 5 \\ 0 \pmod 7 \end{cases}, \quad 21 = 1\times 21 \equiv \begin{cases} 0 \pmod 3 \\ 1 \pmod 5 \\ 0 \pmod 7 \end{cases}, \quad 15 = 1\times 15 \equiv \begin{cases} 0 \pmod 3 \\ 0 \pmod 5 \\ 1 \pmod 7 \end{cases}$$

$$2\times 70 + 3\times 21 + 2\times 15 \equiv \begin{cases} 2\times 1 + 3\times 0 + 2\times 0 \equiv 2 \pmod 3 \\ 2\times 0 + 3\times 1 + 2\times 0 \equiv 3 \pmod 5 \\ 2\times 0 + 3\times 0 + 2\times 1 \equiv 2 \pmod 7 \end{cases}$$

$$x = 233 + k\times 105,\ k \in \mathbb{Z}.$$

欧拉函数(ϕ函数)

- 不大于n且与n互质的正整数的个数，记为 $\phi(n)$ 。
- $\phi(n) = |\{ k \mid 1 \le k \le n , \gcd(k, n) = 1\}|, n \in N^+$
  - $\phi(3) = 2, \phi(4) = 2, \phi(12) = 4$
- 设 $n = p_1^{\alpha 1} p_2^{\alpha 2} \ldots p_k^{\alpha k}$
- 令 $A_i = \{ x \mid 1 \le x \le n , p_i$整除$x\}$
- $\phi(n) = | \sim A_1 \cap \sim A_2 \cap \ldots \cap \sim A_k |$

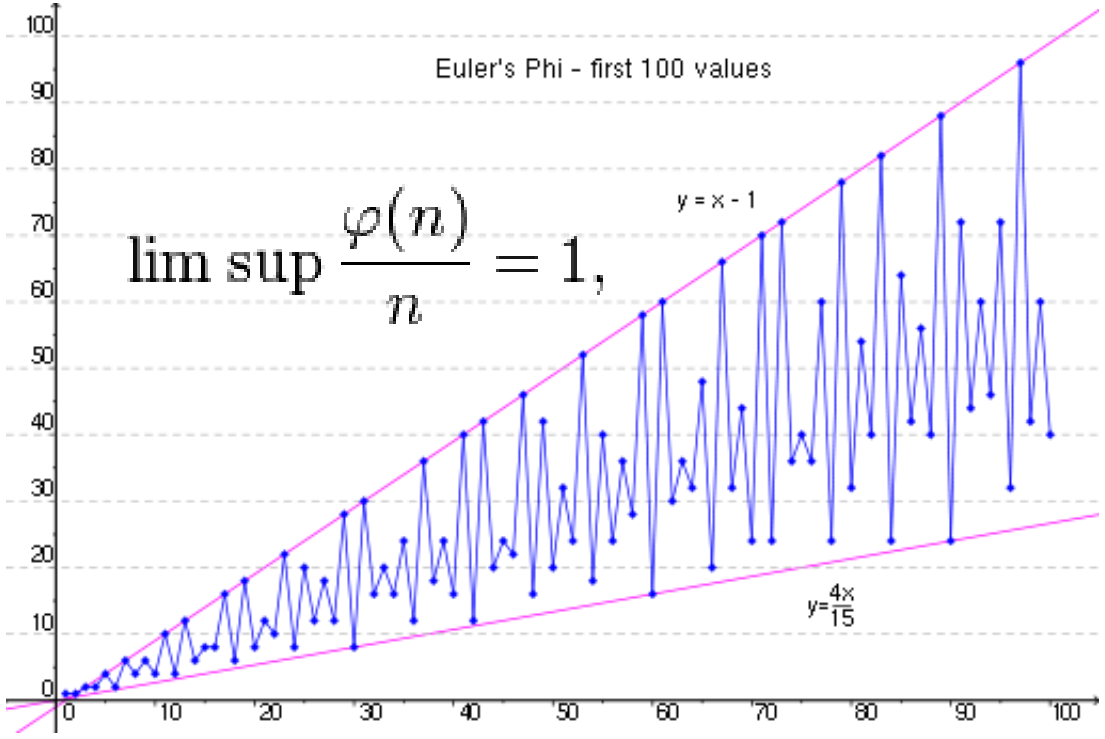
$$= n - (n/p_1 + \ldots + n/p_k) + (n/p_1p_2 + \ldots + n/p_{k-1}p_k)$$

$$- \ldots + (-1)^k n/p_1p_2 \ldots p_k$$

$$= n(1 - 1/p_1)(1 - 1/p_2) \ldots (1 - 1/p_k)$$

欧拉函数

$$\varphi(n) = \prod_{i=1}^{r} p_i^{k_i - 1}(p_i - 1) = \prod_{p|n} p^{\alpha_p - 1}(p-1) = n \prod_{p|n} \left(1 - \frac{1}{p}\right)$$



$$\limsup \frac{\varphi(n)}{n} = 1,$$

$$\liminf \frac{\varphi(n)}{n} \log\log n = e^{-\gamma}.$$

欧拉常数 γ = 0.577215665...

欧拉函数的性质

- **ϕ(*p*)=*p*-1，*p*是素数**
- **如果*m*与*n*互素，则φ(*mn*) = φ(*m*)φ(*n*).**
  - ***φ(mn)=φ(m)φ(n)·d/φ(d)*，其中*d=gcd(m,n)***

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

欧拉定理

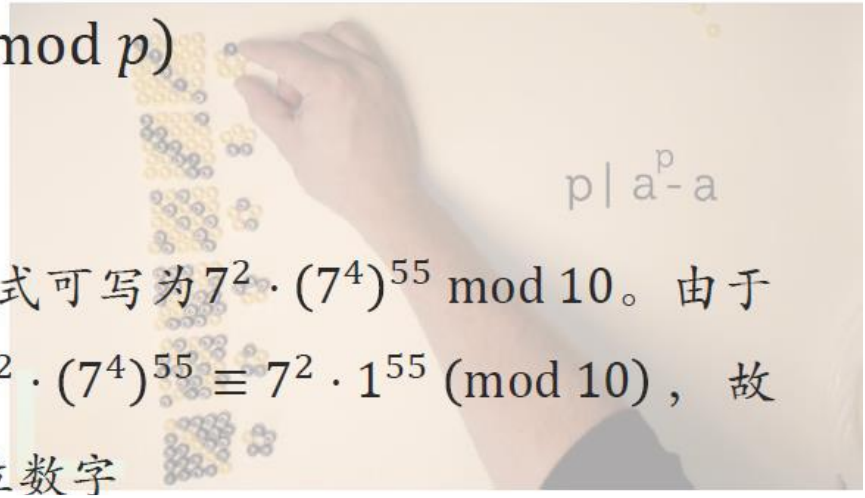
- 定理（Euler定理）：对$a,n \in \mathbb{Z}^+$，若$(a,n)=1$，则：

$$a^{\varphi(n)} \equiv 1 (\mathrm{mod}\, n)$$

- 若上述$n \in \mathbb{Z}^+$为质数，由欧拉函数的性质易得到：
- 定理（Fermat小定理）：设正整数a不是质数p之倍数，则：

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

- 例：求7^{222}的个位数字
  - 解：待求即为$7^{222} \bmod 10$，上式可写为$7^2 \cdot (7^4)^{55} \bmod 10$。由于$(7,10)=1$，由Euler定理，$7^2 \cdot (7^4)^{55} \equiv 7^2 \cdot 1^{55} (\mathrm{mod}\ 10)$，故$7^{222} \bmod 10 = 9$即为$7^{222}$之个位数字



RSA的数学基础*

- 若a与n互质，则 $a^{\varphi(n)} \equiv 1 \pmod{n}$，
  - 若$\alpha \equiv 1 \pmod{\varphi(n)}$，则 $a^{\alpha} \equiv a \pmod{n}$
- 若$n=pq$, $\alpha \equiv 1 \pmod{\varphi(n)}$, $0<m<n$, 则$m^{\alpha} \equiv m \pmod{n}$

- 选取大质素p, q: $n=pq$ (n难以分解成质素乘积）.
- 令 $k= \varphi(n)$ (不知道n的质因子，k难以求出）.
- 设e为公钥，d为私钥，满足 $ed \equiv 1 \pmod{k}$.
- 加密：$S = m^e \pmod{n}$.
- 解密：$t = S^d \pmod{n}$. ($t = m$，why?)

RSA的数学基础*

- 根据加密S可以写成$S = m^e - kn$，带入解密，即证$(m^e - kn)^d \equiv m \pmod n$
- 由$ed \equiv 1 \pmod{\varphi(n)}$，有$ed = h\varphi(n)+1$，再代入上式得$m^{h\varphi(n)+1} \equiv m \pmod n$
- 1- 若m与n互质就是欧拉定理
- 2- 若m与n不互质，不妨设m=kp。（m只能是kp或者kq）
  - 根据欧拉定理有$(kp)^{q-1} \equiv 1 \pmod q$（k与q互质）
  - =》$[(kp)^{q-1}]^{h(p-1)} \times kp \equiv kp \pmod q$
  - =》$(kp)^{ed} \equiv kp \pmod q$（根据$ed = h\varphi(n)+1=h(p-1)(q-1)+1$）
  - =》$(kp)^{ed} = tq + kp$（于是有：t必能被p整除, t=t'p）
  - =》$(kp)^{ed} = t'pq + kp$（m=kp，n=pq）
  - =》$m^{ed} \equiv m \pmod n$

本节小结

问题1：什么是(初等)数论?

- 研究整数的性质：整除、余数、同余算术

问题2：素数有哪些性质?

- 素数基本定理
- 最大公约数
- 中国剩余定理/同余方程
- 欧拉定理/费马小定理

作业

- 见课程网站