高等学校信息安全类专业系列教材

信息安全数学基础

贾春福　钟安鸣　赵源超　编著

清 华 大 学 出 版 社
北京交通大学出版社
·北京·

内 容 简 介

本书系统地介绍了信息安全理论与技术所涉及的数论、代数、椭圆曲线等数学理论基础。全书共分为6章：第1章是预备知识，介绍了书中后面几章所涉及的基础知识；第2章和第3章是数论基础，包括整数的因子分解、同余式、原根、二次剩余、数论的应用等内容；第4章是代数系统，包括群、环、域的概念，一元多项式环和有限域理论初步等内容；第5章是椭圆曲线，包括椭圆曲线的预备知识、椭圆曲线、椭圆曲线上的离散对数等内容；第6章是线性反馈移位寄存器，包括反馈移位寄存器、分圆多项式和本原多项式、m序列等内容。书中每章末都配有适量习题，以供学生学习和复习巩固书中所学内容。

本书是高等学校信息安全专业本科生的教材，也可作为信息科学技术类专业（如计算机科学技术、通信工程和电子科学技术等）本科生和研究生的教材，同时，也可以供从事信息安全和其他信息技术工作的人员参考。

图书在版编目(CIP)数据

信息安全数学基础/贾春福，钟安鸣，赵源超编著. —北京：清华大学出版社；北京交通大学出版社，2010.3

ISBN 978－7－5121－0056－5

Ⅰ.①信…　Ⅱ.①贾…　②钟…　③赵…　Ⅲ.①信息系统－安全技术－应用数学　Ⅳ.①TP309　②O29

中国版本图书馆CIP数据核字(2010)第011503号

责任编辑：谭文芳
出版发行：清华大学出版社　邮编：100084　电话：010－62776969　http://www.tup.com.cn
　　北京交通大学出版社　邮编：100044　电话：010－51686414　http://press.bjtu.edu.cn
印 刷 者：
经　　销：全国新华书店
开　　本：185×260　印张：14　字数：358千字
版　　次：2010年3月第1版　2010年3月第1次印刷
书　　号：ISBN 978－7－5121－0056－5/TP・581
印　　数：1～3 000册　定价：24.00元

前　言

计算机与网络技术的飞速发展和广泛应用，极大地促进了社会的发展，也极大地改变了人们的生活和工作方式。与此同时，信息安全问题也更多地受到关注；信息安全理论与技术已经成为信息科学与技术中极为重要的研究领域；信息安全专门人才的培养受到了社会空前的重视。

“信息安全数学基础”是新兴的信息安全专业本科的专业基础课，对信息安全理论和技术的深入学习具有重要的意义。本书是在南开大学信息安全专业“信息安全数学基础”课程授课讲义的基础上整理而成的。全书共分为 6 章：第 1 章是预备知识，介绍了书中所涉及的基础知识；第 2 章和第 3 章是数论基础，包括整数的因子分解、同余式、原根、二次剩余和数论的应用等内容；第 4 章是代数系统，包括群、环、域的概念，一元多项式环和有限域理论初步等内容；第 5 章是椭圆曲线，包括椭圆曲线的预备知识、椭圆曲线、椭圆曲线上的离散对数等内容；第 6 章是线性反馈移位寄存器，包括反馈移位寄存器、分圆多项式和本原多项式、m 序列等内容。书中每章末都配有适量的习题，供学生在学习和复习巩固书中所学内容时使用。

本书内容的选取，我们参照了“信息安全类专业指导性专业规范”中对“信息安全数学基础”相关教学内容和要求的阐述；并将多年来积累的实际教学经验融入其中，力求知识系统化、较好地覆盖信息安全领域所涉及的数学基础知识。对书中内容所涉及的基础预备知识作了简明扼要的介绍；书中所涉及的数学结论都给出了详细的证明；习题的配置着力于帮助学生巩固所学的内容和能力拓展。本书适合高等学校信息安全、计算机科学技术和通信工程等专业本科生和研究生使用，也可供相关领域的科研人员和技术人员参考。

本书由贾春福、钟安鸣、赵源超等编写，最后由贾春福统稿。王冬、刘昕海等对书中的内容进行了校对，在此表示感谢。另外，本书是南开大学教材资助项目，在此也表示衷心的感谢。

由于时间仓促，书中难免有疏漏和不当之处，敬请读者批评指正。

编　者

2009 年 12 月于南开园

符　号　表

符　号	含　义
$\mathbf{N}$	自然数集
$\mathbf{Z}$	整数集
$\mathbf{Q}$	有理数集
$\mathbf{R}$	实数集
$\mathbf{C}$	复数集
$\mathbf{Z}_m$	整数的模 m 剩余类集
$\mathbf{F}_p$	当 p 为素数时，$\mathbf{Z}_p$ 的专有表示形式
$m\mathbf{Z}$	整数 m 的整数倍构成的集合
$\mathbf{Z}[x]$	整数环$(\mathbf{Z},+,\times)$上的一元多项式集
$\mathbf{Z}_m[x]$	$\mathbf{Z}_m$ 上的一元多项式集
$\mathbf{GF}(p)$	元素数为 p 的有限域
$\bmod n$	模 n 运算
$\gcd(m,n)$	整数 m,n 的最大公因子
$\mathrm{lcm}(m,n)$	整数 m,n 的最小公倍数
$\deg(f(x))$	多项式 $f(x)$的次数

目　　录

第 1 章　预备知识 …… 1
1.1　集合、关系和函数 …… 1
1.1.1　集合 …… 1
1.1.2　关系 …… 6
1.1.3　函数 …… 13
1.2　组合数学初步知识 …… 19
1.2.1　排列与组合 …… 19
1.2.2　生成函数 …… 26
习题 …… 33
第 2 章　数论基础(一) …… 35
2.1　整除 …… 35
2.1.1　整除与带余除法 …… 35
2.1.2　最大公因子与辗转相除法 …… 38
2.1.3　连分数 …… 43
2.1.4　算术基本定理 …… 50
2.1.5　梅森素数和费马素数 …… 53
2.2　同余 …… 55
2.2.1　同余的概念和性质 …… 55
2.2.2　剩余类和欧拉定理 …… 58
2.2.3　线性同余方程 …… 63
2.2.4　孙子定理与同余方程组 …… 67
2.2.5　高次同余方程 …… 74
习题 …… 79
第 3 章　数论基础(二) …… 82
3.1　原根 …… 82
3.1.1　整数的次数 …… 82
3.1.2　原根 …… 86
3.1.3　指数与 n 次剩余 …… 92
3.2　二次剩余 …… 96
3.2.1　二次剩余的概念和性质 …… 96
3.2.2　勒让德符号与二次互反律 …… 100
3.2.3　雅可比符号 …… 106
3.3　数论的典型应用 …… 109

3.3.1 素性检验算法 …… 109
3.3.2 因子分解算法 …… 115
习题 …… 117
第 4 章 代数系统基础 …… 119
4.1 群 …… 119
4.1.1 群及其基本性质 …… 119
4.1.2 子群 …… 123
4.1.3 循环群和群的生成 …… 125
4.1.4 陪集和拉格朗日定理 …… 128
4.1.5 同态与同构 …… 130
4.1.6 正规子群与商群 …… 134
4.1.7 循环群的分类 …… 137
4.1.8 置换群 …… 138
4.2 交换环和域 …… 141
4.2.1 交换环及其基本性质 …… 141
4.2.2 域及其基本性质 …… 147
4.2.3 同态与同构 …… 148
4.2.4 一元多项式环 …… 150
4.2.5 理想和商环 …… 151
4.3 域上的一元多项式环 …… 156
4.3.1 一元多项式的整除 …… 157
4.3.2 一元多项式环的理想 …… 160
4.3.3 域上一元多项式唯一分解定理 …… 161
4.3.4 多项式不可约性检验 …… 162
4.3.5 一元多项式的同余与商环 …… 164
4.4 有限域理论初步 …… 165
习题 …… 169
第 5 章 椭圆曲线 …… 171
5.1 椭圆曲线的预备知识 …… 171
5.1.1 仿射平面和射影平面 …… 171
5.1.2 判别式、结式和代数不变量 …… 173
5.1.3 一元三次方程的公式解——Cartan 公式 …… 177
5.2 椭圆曲线 …… 178
5.2.1 Weierstrass 方程 …… 178
5.2.2 椭圆曲线 …… 181
5.2.3 椭圆曲线上点的加法群(Mordell-Weil 群) …… 183
5.2.4 有限域上的椭圆曲线 …… 187
5.3 离散对数初步 …… 191
5.3.1 有限域上的离散对数 …… 191

5.3.2 椭圆曲线上的离散对数 …… 193
习题 …… 194
第 6 章 线性反馈移位寄存器(LFSR) …… 196
6.1 反馈移位寄存器 …… 196
6.1.1 反馈移位寄存器 …… 196
6.1.2 线性反馈移位寄存器(LFSR) …… 197
6.1.3 非线性组合移位寄存器简介 …… 198
6.2 分圆多项式和本原多项式 …… 198
6.2.1 分圆多项式 …… 198
6.2.2 本原多项式 …… 202
6.3 m 序列 …… 205
6.3.1 LFSR 的特征多项式 …… 205
6.3.2 m 序列的产生条件 …… 207
6.3.3 m 序列的特点 …… 208
6.3.4 m 序列的破译 …… 210
习题 …… 212
参考文献 …… 213

第 1 章　预 备 知 识

在当前的信息安全专业的课程体系中，由于“信息安全数学基础”课程涉及的一些数学基础知识在前期的“高等数学”等课程中介绍得较少，本书将对相关的这部分内容进行一些补充，以便读者能够顺利地阅读书中后续的各个章节.

本章是与书中后面几章内容相关的预备知识的介绍，包括集合、关系和函数的基本概念、排列与组合及生成函数等内容.

1.1　集合、关系和函数

集合论是德国著名数学家康托尔(Cantor)于 19 世纪末创立的，康托尔当时建立的集合论称为朴素集合论. 20 世纪初，策梅罗(Zermelo)给出了第一个集合论的公理系统，并在此基础上逐步形成了公理化集合论和抽象集合论，使该学科成为在数学中发展最快的一个分支.

集合论是现代数学的基础，通俗地讲，数学所研究的一切概念都可以用集合来定义，甚至包括很多已经非常熟悉的概念，如整数、实数和函数等，都可以用集合加以表示. 此外，集合概念的引入，也使得我们能够摆脱具体数系的束缚，建立和研究很多抽象的数学概念和对象，从而得到很多抽象层次上的具有更多普遍含义的结论，这一点将在本书的第 4 章得到较多的体现. 现在，集合论观点已经渗透到了古典分析、泛函、概率和信息论等各个领域. 本节将介绍集合论的基础知识，包括集合与关系、集合运算、函数和等势的概念和规则.

1.1.1　集合

1. 集合的概念

集合的概念是现代数学中最基本的概念之一. 一般来讲，把具有共同性质的一些事物汇集成一个整体，就形成一个**集合**. 这些事物称为**元素**或**成员**. 例如，所有 0 和 1 之间的实数，教室里的所有椅子，图书馆里的所有藏书都构成一个集合.

通常用大写英文字母 A，B，…表示集合，小写英文字母 a，b，…表示集合中的元素. 若元素 a 是集合 S 中的元素，则记作 $a\in S$，读作 a 属于 S，或 a 在 S 之中. 若元素 a 不是集合 S 中的元素，记作 $a\notin S$，读作 a 不属于 S，或 a 不在 S 之中.

对于一个集合 S，如果它是由有限个元素组成的，称 S 为**有限集**；否则称 S 为**无限集**.

集合通常有两种表示方法. 第一种方法是把集合中的元素列举出来，称作**列举法**. 例如

$$A=\{a,b,c,d\},\quad B=\{1,2,3,\cdots\}.$$

第二种方法称为**叙述法**，即用一种规则来限定某个元素是否属于该集合. 例如

$$S_1=\{x\mid x\text{ 是正整数}\},\ S_2=\{x\mid x\in\mathbf{N}\wedge x\leqslant 9\},\ S_3=\{x\mid x\in\mathbf{R}\wedge 5x^2-1=0\},$$

其中“$\wedge$”表示“并且”.

定义 1.1.1 设 A，B 是任意两个集合，假如 A 的每一个元素都是 B 的成员，则称 A 为 B 的**子集**，记作 $A\subseteq B$ 或 $B\supseteq A$，读作 A **包含于** B，或 B **包含** A. 符号化表示为

$$A\subseteq B\Leftrightarrow(\forall x)(x\in A\rightarrow x\in B),$$

其中“$\forall$”表示“任意”，“$\Leftrightarrow$”表示命题“等价”，“$\rightarrow$”表示“蕴涵”(命题内).

例如，设 $\mathbf{N}$ 为自然数集，$\mathbf{Q}$ 为有理数集，$A=\{1,2,3\}$，$B=\{1\}$，则

$$A\subseteq\mathbf{N},\ B\subseteq A,\ B\subseteq\mathbf{N},\ \mathbf{N}\subseteq\mathbf{Q}.$$

定义 1.1.2 如果集合 A 的每一个元素都属于 B，但集合 B 中至少有一个元素不属于 A，则称 A 为 B 的**真子集**，记作 $A\subset B$，读作 A **真包含于** B，或 B **真包含** A. 符号化表示为

$$A\subset B\Leftrightarrow A\subseteq B\wedge A\neq B,$$

或

$$A\subset B\Leftrightarrow(\forall x)(x\in A\rightarrow x\in B)\wedge(\exists x)(x\in B\wedge x\notin A).$$

例如，整数集是有理数集的真子集.

定义 1.1.3 设 A，B 是任意给定的两个集合，如果 $A\subseteq B$ 且 $B\subseteq A$，则称集合 A 和集合 B **相等**，记作 $A=B$. 符号化表示为

$$A=B\Leftrightarrow A\subseteq B\wedge B\subseteq A,$$

否则，称 A 与 B **不相等**，记作 $A\neq B$.

数学的所有分支中都会经常遇到需要证明两个集合相等的问题，注意，这个定义就是证明两个集合相等的关键所在，一般的证明步骤总结如下：

第一步，从集合 A 中任意选择一个元素，都能够证明这个元素也属于集合 B，根据定义 1.1.1，可以推得 $A\subseteq B$；

第二步，从集合 B 中任意选择一个元素，都能够证明这个元素也属于集合 A，根据定义 1.1.1，可以推得 $B\subseteq A$；

第三步，根据定义 1.1.3，可以推得 $A=B$.

例如，若 $A=\{3,6,9\}$，$B=\{6,9,3\}$，$C=\{3,9\}$，则可知 $A=B$，$A\neq C$.

从这个例子中可以看出，集合中元素的排列顺序是无关紧要的.

定义 1.1.4 不含任何元素的集合称为**空集**，记作$\varnothing$. 符号化表示为

$$\varnothing=\{x\mid p(x)\wedge\sim p(x)\},$$

其中，$p(x)$是任意谓词(谓词是用来描述客体的性质或关系的语句)，“$\sim$”表示“否”.

定理 1.1.1 对于任意一个集合 A，$\varnothing\subseteq A$.

证明 假设$\varnothing\subseteq A$ 是假，则至少存在一个元素 x，使 $x\in\varnothing$且 $x\notin A$. 因为空集$\varnothing$不包含任何元素，所以假设不成立，产生矛盾. 定理得证.

由空集和子集的定义可知，对于每个非空集合 A，至少有两个不同的子集 A 和$\varnothing$. 称 A 和$\varnothing$是 A 的**平凡子集**.

定理 1.1.2 空集是唯一的.

证明 用反证法. 假设存在两个空集$\varnothing_1$ 和$\varnothing_2$. 因为空集被包含于每一个集合中，于是有

$$\varnothing_1\subseteq\varnothing_2$$

且

$$\varnothing_2\subseteq\varnothing_1,$$

故$\varnothing_1=\varnothing_2$，即空集是唯一的.

定义 1.1.5　给定集合 A，由集合 A 的所有子集组成的集合称为集合 A 的**幂集**，记作 $\rho(A)$或 2^A，

$$\rho(A)=\{B\mid B\subseteq A\}.$$

例如，对于 $A=\{a,b,c\}$，有 $\rho(A)=\{\varnothing,\{a\},\{b\},\{c\},\{a,b\},\{b,c\},\{a,c\},\{a,b,c\}\}$.

定义 1.1.6　在一定范围内，如果所有集合均为某一集合的子集，则称该集合为**全集**，记作 E.

对于任一 $x\in A$，因为 $A\subseteq E$，故 $x\in E$. 符号化表示为

$$E=\{x\mid p(x)\vee\sim p(x)\},$$

其中，$p(x)$是任意谓词，“$\vee$”表示“或”.

全集是一个相对的概念，研究的问题不同，所取的全集也往往不同.

2. 集合运算

集合的运算就是以给定的集合为对象，按照确定的规则得到另外一些集合. **文氏图**(Venn Diagram)可以直观、形象地表示集合间的关系及运算结果. 在文氏图中，通常用一个矩形表示全集 E，然后在矩形的内部画一些圆(或其他封闭的曲线)，圆的内部代表集合，不同的圆代表不同的集合.

定义 1.1.7　设任意两个集合 A 和 B，由集合 A 和 B 的所有共同元素组成的集合 S，称为 A 和 B 的**交集**，记作 $A\cap B$. 显然

$$S=A\cap B=\{x\mid x\in A\wedge x\in B\}.$$

其文氏图如图 1.1.1 所示.

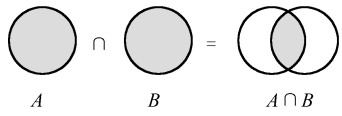

图 1.1.1　集合的交集

[例 1.1.1]　设 $A=\{0,2,a,7,c\}$，$B=\{r,m,0,c,2\}$，求 $A\cap B$.

解　$A\cap B=\{0,2,c\}$.

[例 1.1.2]　设 $A\subseteq B$，C 是任意集合，求证 $A\cap C\subseteq B\cap C$.

证明　由 $A\subseteq B$ 可知，若 $x\in A$，则 $x\in B$. 对于任意的 $x\in A\cap C$，由“$\cap$”的定义，有 $x\in A$ 且 $x\in C$，即 $x\in B$ 且 $x\in C$，故 $x\in B\cap C$. 因此，$A\cap C\subseteq B\cap C$.

定义 1.1.8　设任意两个集合 A 和 B，所有属于 A 或属于 B 的元素组成的集合 S，称为 A 和 B 的**并集**，记作 $A\cup B$. 显然

$$S=A\cup B=\{x\mid x\in A\vee x\in B\}.$$

文氏图表示如图 1.1.2 所示.

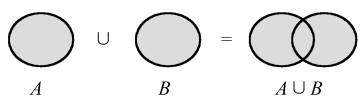

图 1.1.2　集合的并集

[例 1.1.3]　设 $A=\{a, 2\}$, $B=\{2, m\}$, 求 $A\cup B$.

解　$A\cup B=\{a, 2, m\}$.

[例 1.1.4]　设 $A\subseteq B$, $C\subseteq D$, 求证 $A\cup C\subseteq B\cup D$.

证明　对任意 $x\in A\cup C$, 有 $x\in A$ 或 $x\in C$. 若 $x\in A$, 则由 $A\subseteq B$, 有 $x\in B$, 故 $x\in B\cup D$. 若$x\in C$, 则由 $C\subseteq D$, 有 $x\in D$, 故 $x\in B\cup D$. 因此, $A\cup C\subseteq B\cup D$.

[例 1.1.5]　求证下列命题.

(1) $A\subseteq B$, 当且仅当 $A\cup B=B$;

(2) $A\subseteq B$, 当且仅当 $A\cap B=A$.

证明　(1) 若 $A\subseteq B$, 则对任意的 $x\in A$, 必有 $x\in B$. 又由于对任意的 $x\in A\cup B$, 有 $x\in A$ 或 $x\in B$, 故 $x\in B$, 所以 $A\cup B\subseteq B$. 又 $B\subseteq A\cup B$, 于是得到 $A\cup B=B$. 反之, 若 $A\cup B=B$, 因为 $A\subseteq A\cup B$, 所以 $A\subseteq B$.

(2) 其证明过程与(1)类似.

定义 1.1.9　设任意两个集合 A 和 B, 所有属于 A 而不属于 B 的一切元素组成的集合 S, 称为 B 对 A 的**补集**, 或称**对称补**, 记作 $A-B$. 显然

$$S=A-B=\{x\mid x\in A\wedge x\notin B\}=\{x\mid x\in A\wedge\sim(x\in B)\}.$$

$A-B$ 也称为集合 A 和 B 的**差**. 文氏图表示如图 1.1.3 所示.

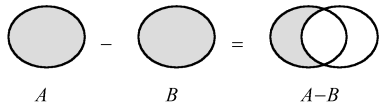

图 1.1.3　集合的对称补集

[例 1.1.6]　设 $A=\{a, 7, c\}$, $B=\{m, c, 2\}$, 求 $A-B$.

解　$A-B=\{a, 7\}$.

定义 1.1.10　设 E 为全集, 对任一集合 A 关于 E 的补集 $E-A$, 称为集合 A 的**绝对补**, 记作 $\sim A$ 或者 $\overline{A}$. 显然

$$\sim A=E-A=\{x\mid x\in E\wedge x\notin A\}.$$

[例 1.1.7]　设 A, B 为任意两个集合, 则 $A-B=A\cap\sim B$.

证明　对于任意的 x, 有

$$x\in A-B\Leftrightarrow x\in A\wedge x\notin B\Leftrightarrow x\in A\wedge x\in\sim B\Leftrightarrow x\in A\cap\sim B,$$

所以 $A-B=A\cap\sim B$.

定义 1.1.11　设任意两个集合 A 和 B, A 和 B 的**对称差**为集合 S, 其元素或属于 A, 或属于 B, 但不能既属于 A 又属于 B, 记作 $A\oplus B$. 显然

$$S=A\oplus B=(A-B)\cup(B-A)=\{x\mid(x\in A\wedge x\notin B)\vee(x\in B\wedge x\notin A)\}.$$

文氏图表示如图 1.1.4 所示.

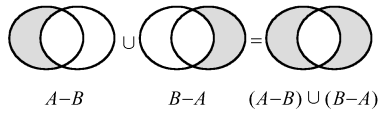

图 1.1.4　集合的对称差集

[例 1.1.8] 设 $A=\{4,6,8\}$，$B=\{1,4,8\}$，求 $A\oplus B$.

解 $A\oplus B=\{1,6\}$.

下面给出集合运算性质中最主要的几条定律.

定理 1.1.3 设 A，B，C 是全集 E 的任意子集.

(1) 幂等律 $A\cup A=A$

$A\cap A=A$

(2) 交换律 $A\cup B=B\cup A$

$A\cap B=B\cap A$

$A\oplus B=B\oplus A$

(3) 结合律 $(A\cup B)\cup C=A\cup(B\cup C)$

$(A\cap B)\cap C=A\cap(B\cap C)$

$(A\oplus B)\oplus C=A\oplus(B\oplus C)$

(4) 分配律 $A\cup(B\cap C)=(A\cup B)\cap(A\cup C)$

$A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$

$A\cap(B-C)=(A\cap B)-(A\cap C)$

(5) 同一律 $A\cup\varnothing=A$

$A\cap E=A$

$A-\varnothing=A$

$A\oplus\varnothing=A$

(6) 零律 $A\cup E=E$

$A\cap\varnothing=\varnothing$

(7) 互补律 $A\cup\sim A=E$

$A\cap\sim A=\varnothing$

$\sim E=\varnothing$

$\sim\varnothing=E$

(8) 吸收律 $A\cup(A\cap B)=A$

$A\cap(A\cup B)=A$

(9) 摩根定律 $\sim(A\cup B)=\sim A\cap\sim B$

$\sim(A\cap B)=\sim A\cup\sim B$

$A-(B\cup C)=(A-B)\cap(A-C)$

$A-(B\cap C)=(A-B)\cup(A-C)$

(10) 双重否定律 $\sim(\sim A)=A$

(11) $A\oplus A=\varnothing$ $A-A=\varnothing$ $A\cap B\subseteq A$ $A\cap B\subseteq B$

(12) $A\subseteq A\cup B$ $B\subseteq A\cup B$ $A-B\subseteq A$ $A-B=A\cap\sim B$

(13) $A\oplus B=(A-B)\cup(B-A)=(A\cup B)-(A\cap B)=(A\cap\sim B)\cup(\sim A\cap B)$

对于上面的集合基本定律，下面以例题的形式证明其中的一部分，其余留给读者作为习题完成.

[例 1.1.9] 证明幂等律 $A\cup A=A$.

证明 对于任意的 x，有

$$x\in A\cup A\Leftrightarrow x\in A\vee x\in A\Leftrightarrow x\in A,$$

所以 $A\cup A=A$.

[例 1.1.10]　证明交换律 $A\cap B=B\cap A$.

证明　对于任意的 x，有

$$x\in A\cap B\Leftrightarrow x\in A\wedge x\in B\Leftrightarrow x\in B\wedge x\in A\Leftrightarrow x\in B\cap A,$$

所以 $A\cap B=B\cap A$.

[例 1.1.11]　证明分配律 $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$.

证明　对于任意给定的 x，有

$$\begin{aligned}x\in A\cap(B\cap C)&\Leftrightarrow x\in A\wedge x\in(B\cup C)\\&\Leftrightarrow x\in A\wedge(x\in B\vee x\in C)\\&\Leftrightarrow(x\in A\wedge x\in B)\vee(x\in A\wedge x\in C)\\&\Leftrightarrow(x\in A\cap B)\vee(x\in A\cap C)\\&\Leftrightarrow x\in(A\cap B)\cup(A\cap C)\end{aligned}$$

所以 $A\cap(B\cup C)=(A\cap B)\cup(A\cap C)$.

[例 1.1.12]　证明吸收律，即 $A\cup(A\cap B)=A$.

证明　$A\cup(A\cap B)=(A\cap E)\cup(A\cap B)=A\cap(E\cup B)=A\cap E=A$.

[例 1.1.13]　证明摩根定律 $\sim(A\cup B)=\sim A\cap\sim B$.

证明

$$\begin{aligned}\sim(A\cup B)&=\{x\mid x\in\sim(A\cup B)\}=\{x\mid x\notin A\cup B\}=\{x\mid x\notin A\wedge x\notin B\}\\&=\{x\mid(x\in\sim A)\wedge(x\in\sim B)\}=\sim A\cap\sim B\end{aligned}$$

[例 1.1.14]　证明分配律 $A\cap(B-C)=(A\cap B)-(A\cap C)$.

证明　由于

$$A\cap(B-C)=A\cap(B\cap\sim C)=A\cap B\cap\sim C,$$

又

$$\begin{aligned}(A\cap B)-(A\cap C)&=(A\cap B)\cap\sim(A\cap C)\\&=(A\cap B)\cap(\sim A\cup\sim C)\\&=(A\cap B\cap\sim A)\cup(A\cap B\cap\sim C)\\&=\varnothing\cup(A\cap B\cap\sim C)\\&=A\cap B\cap\sim C\end{aligned}$$

故可知 $A\cap(B-C)=(A\cap B)-(A\cap C)$.

1.1.2　关系

关系的概念在日常生活中是普遍存在的，如师生关系、朋友关系、同学关系，等等. 在数学上，关系可以表达集合中元素间的联系. 在介绍关系的概念以前，首先介绍序偶和笛卡儿积的概念.

定义 1.1.12　由两个具有给定次序的个体 x 和 y(允许 $x=y$)所组成的序列，称为**序偶**，记作$<x, y>$. 其中 x 称为**第一分量**，y 称为**第二分量**.

序偶可以看作是具有两个元素的集合，但它与一般集合不同的是，序偶具有确定的次序. 例如，在集合中，有$\{a, b\}=\{b, a\}$，但对于序偶$<a, b>\neq<b, a>$.

定义 1.1.13 设$<a, b>$，$<x, y>$是两个序偶，则$<a, b>=<x, y>$当且仅当$a=x$且$b=y$.

注意，这个定义告诉我们证明两个序偶相等的关键在于，分别证明两个位置上的对应元素分别相等.

定义 1.1.14 由n个具有给定次序的个体a_1，a_2，…，a_n组成的序列，称为**有序 n 元组**，记作$<a_1, a_2, \cdots, a_n>$.

有序n元组的实质依然是序偶，可将其表示为

$$<a_1,a_2,\cdots,a_n>=\ll a_1,a_2,\cdots,a_{n-1}>,a_n>=\cdots=\lll\cdots\ll a_1,a_2>,a_3>,\cdots>,a_{n-1}>,a_n>$$

其中，a_i称为第i个分量. $<a_1, a_2, \cdots, a_n>=<b_1, b_2, \cdots, b_n>$当且仅当$a_i=b_i(i=1, 2, \cdots, n)$.

定义 1.1.15 设A_1，A_2，…，A_n是任意给定的n个集合，若有序n元组$<a_1, a_2, \cdots, a_n>$的第一个分量是取自集合A_1里的元素，第二个分量是取自集合A_2里的元素，…，第n个分量是取自集合A_n里的元素，则由所有这样的有序n元组所组成的集合称为集合A_1，A_2，…，A_n的**笛卡儿积**，并用$A_1\times A_2\times\cdots\times A_n$表示，即

$$A_1\times A_2\times\cdots\times A_n=\{<a_1, a_2, \cdots, a_n>\mid a_i\in A_i, i=1, 2, \cdots, n\}.$$

特别地，两个集合的笛卡儿积可以叙述为：任意给定两个集合A和B，若序偶的第一个分量是A的元素，第二个分量是B的元素，则所有这样的序偶的集合称为A和B的笛卡儿积或直积，记作$A\times B$，即

$$A\times B=\{<x, y>\mid x\in A\wedge y\in B\}.$$

[例 1.1.15] 设$A=\{0, 1\}$，$B=\{a, b\}$，$C=\varnothing$，则

$$A\times B=\{<0, a>, <0, b>, <1, a>, <1, b>\},$$
$$B\times A=\{<a, 0>, <a, 1>, <b, 0>, <b, 1>\},$$
$$A\times A=\{<0, 0>, <1, 1>, <1, 0>, <0, 1>\},$$
$$B\times B=\{<a, a>, <b, b>, <a, b>, <b, a>\},$$
$$A\times C=\varnothing,$$
$$C\times A=\varnothing.$$

显然，$A\times B\neq B\times A$，即笛卡儿积不满足交换律.

[例 1.1.16] 设$A=\{1\}$，$B=\{a, b\}$，$C=\{x, y\}$，则

$$(A\times B)\times C=\{\ll 1, a>, x>, \ll 1, a>, y>, \ll 1, b>, x>, \ll 1, b>, y>\},$$
$$A\times(B\times C)=\{<1, <a, x\gg, <1, <a, y\gg, <1, <b, x\gg, <1, <b, y\gg\}.$$

显然，$(A\times B)\times C\neq A\times(B\times C)$，即笛卡儿积不满足结合律.

定理 1.1.4 笛卡儿积的性质如下：

(1) 交换律不成立，即当$A\neq B$时，$A\times B\neq B\times A$.

(2) 结合律不成立，即$(A\times B)\times C\neq A\times(B\times C)$.

(3) 下列分配律是成立的：

① $A\times(B\cup C)=(A\times B)\cup(A\times C)$；

② $A\times(B\cap C)=(A\times B)\cap(A\times C)$；

③ $(A\cup B)\times C=(A\times C)\cup(B\times C)$；

④ $(A\cap B)\times C=(A\times C)\cap(B\times C)$；

⑤ $A\times(B-C)=(A\times B)-(A\times C)$；

⑥ $(A-B)\times C=(A\times C)-(B\times C)$.

(4) 若 $C\neq\varnothing$，则 $A\subseteq B\Leftrightarrow(A\times C\subseteq B\times C)\Leftrightarrow(C\times A\subseteq C\times B)$.

(5) 设 A，B，C，D 是四个非空集合，则 $A\times B\subseteq C\times D$ 当且仅当 $A\subseteq C$ 且 $B\subseteq D$.

[例 1.1.17] 证明分配率 $A\times(B\cup C)=(A\times B)\cup(A\times C)$.

证明 对任意的$<x,y>$，有

$$\begin{aligned}<x,y>\in A\times(B\cup C)&\Leftrightarrow x\in A\wedge y\in(B\cup C)\\&\Leftrightarrow x\in A\wedge(y\in B\vee y\in C)\\&\Leftrightarrow(x\in A\wedge y\in B)\vee(x\in A\wedge y\in C)\\&\Leftrightarrow<x,y>\in A\times B\ \vee<x,y>\in A\times C\\&\Leftrightarrow<x,y>\in(A\times B)\cup(A\times C)\end{aligned}$$

所以

$$A\times(B\cup C)=(A\times B)\cup(A\times C).$$

[例 1.1.18] 证明分配律$(A\cap B)\times C=(A\times C)\cap(B\times C)$.

证明 对任意的$<x,y>$，有

$$\begin{aligned}<x,y>\in(A\cap B)\times C&\Leftrightarrow x\in(A\cap B)\wedge y\in C\\&\Leftrightarrow(x\in A\wedge x\in B)\wedge y\in C\\&\Leftrightarrow(x\in A\wedge y\in C)\wedge(x\in B\wedge y\in C)\\&\Leftrightarrow<x,y>\in A\times C\wedge<x,y>\in B\times C\\&\Leftrightarrow<x,y>\in(A\times C)\cap(B\times C)\end{aligned}$$

所以

$$(A\cap B)\times C=(A\times C)\cap(B\times C).$$

[例 1.1.19] 设 A，B，C 是三个任意集合，且 $C\neq\varnothing$，则 $A\subseteq B$ 当且仅当 $A\times C\subseteq B\times C$.

证明 先证必要性. 设 $A\subseteq B$ 成立，则对任意的 x，若 $x\in A$，则必有 $x\in B$. 现对任意的$<x,y>$，有

$$<x,y>\in A\times C\Leftrightarrow x\in A\wedge y\in C\Rightarrow x\in B\wedge y\in C\Leftrightarrow<x,y>\in B\times C,$$

所以 $A\times C\subseteq B\times C$.

再证充分性. 设 $A\times C\subseteq B\times C$ 成立，因为 $C\neq\varnothing$，故存在 $y\in C$. 对于任意的 x，有

$$x\in A\Rightarrow x\in A\wedge y\in C\Leftrightarrow<x,y>\in A\times C\Rightarrow<x,y>\in B\times C\Leftrightarrow x\in B\wedge y\in C\Rightarrow x\in B,$$

其中“⇒”表示“蕴涵”(命题间)，所以 $A\subseteq B$. 证毕.

[例 1.1.20] 设 A，B，C，D 是四个非空集合，则 $A\times B\subseteq C\times D$ 当且仅当 $A\subseteq C$ 且$B\subseteq D$.

证明 先证必要性. 设 $A\times B\subseteq C\times D$ 成立，则对任意的 $x\in A$ 和 $y\in B$，有

$$x\in A\wedge y\in B\Leftrightarrow<x,y>\in A\times B\Rightarrow<x,y>\in C\times D\Leftrightarrow x\in C\wedge y\in D,$$

所以 $A\subseteq C$ 且 $B\subseteq D$.

再证充分性. 设 $A\subseteq C$ 且 $B\subseteq D$ 成立，则对任意的 $x\in A$ 和 $y\in B$，有

$$<x,y>\in A\times B\Leftrightarrow x\in A\wedge y\in B\Rightarrow x\in C\wedge y\in D\Leftrightarrow<x,y>\in C\times D,$$

所以 $A\times B\subseteq C\times D$. 证毕.

定义 1.1.16 设 A_1，A_2，…，A_n 是任意给定的集合，笛卡儿积 $A_1\times A_2\times\cdots\times A_n$ 的任

何一个子集 R 称为 A_1，A_2，…，A_n 上的一个 ***n* 元关系**.

特别地，设 A，B 是任意两个集合，则笛卡儿积 $A\times B$ 的任意一个子集 R 称为从集合 A 到集合 B 的一个**二元关系**，$<a, b>\in R$ 也可表示为 aRb. 如果一个二元关系是从集合 A 到其自身的关系，则这样的二元关系称为集合 A 上的关系.

例如，设 $A=\{1, 2, 3\}$，$B=\{a, b\}$，则

$$A\times B=\{<1, a>, <1, b>, <2, a>, <2, b>, <3, a>, <3, b>\},$$

$$B\times B=\{<a, a>, <a, b>, <b, a>, <b, b>\}$$

$A\times B$ 的任意一个子集都是一个关系，如 $R_1=\{<1, a>\}$，$R_2=\{<2, a>, <3, b>\}$ 等都是从 A 到 B 的关系；$R_3=\{<a, a>, <a, b>, <b, b>\}$ 是集合 B 上的一个二元关系.

对于有限集合上的二元关系 R 除了可以用序偶集合表示外，还可以用矩阵(通常称作**关系矩阵**)表示. 设 $A=\{a_1, a_2, \cdots, a_m\}$，$B=\{b_1, b_2, \cdots, b_n\}$，$R$ 为从 A 到 B 的一个二元关系，则对应于关系 R 的关系矩阵为 $\boldsymbol{M}_R=[r_{ij}]_{m\times n}$，其中

$$r_{ij}=\begin{cases}1, & \text{当}<a_i, b_j>\in R\\ 0, & \text{当}<a_i, b_j>\notin R\end{cases}\quad (i=1, 2, \cdots, m;\quad j=1, 2, \cdots, n).$$

例如，在上例中，R_1 和 R_2 对应的关系矩阵分别为

$$\boldsymbol{M}_{R_1}=\begin{bmatrix}1 & 0\\ 0 & 0\\ 0 & 0\end{bmatrix}$$

和

$$\boldsymbol{M}_{R_2}=\begin{bmatrix}0 & 0\\ 1 & 0\\ 0 & 1\end{bmatrix}.$$

定义 1.1.17 设 R 是从集合 A 到集合 B 的一个二元关系，则由 R 中所有序偶的第一个分量组成的集合称为关系 R 的**定义域**，记作 $D(R)$，由 R 中所有序偶的第二个分量组成的集合称为关系 R 的**值域**，记作 $V(R)$，即

$$D(R)=\{a\mid a\in A\wedge(\exists b)(<a, b>\in R)\},$$

$$V(R)=\{b\mid b\in B\wedge(\exists a)(<a, b>\in R)\}.$$

显然，$D(R)\subseteq A$，$V(R)\subseteq B$.

[例 1.1.21] 设 $A=\{1, 2, 3, 4\}$，求 A 上的整除关系，并求相应的定义域和值域.

解 整除关系 $R=\{<1, 1>, <1, 2>, <1, 3>, <1, 4>, <2, 4>, <4, 4>\}$，对应此整除关系 R 的定义域

$$D(R)=\{1, 2, 4\},$$

值域

$$V(R)=\{1, 2, 3, 4\}.$$

定义 1.1.18 设 R 是从集合 A 到集合 B 的一个二元关系，若 $R=\varnothing$，则称 R 为**空关系**，若 $R=A\times B$，则称 R 为**全域关系**.

定义 1.1.19 设 I_X 是集合 X 上的二元关系，如果 $I_X=\{<x, x>\mid x\in X\}$，则称 I_X 为 X 中的**恒等关系**.

例如，设 $A=\{1, 2, a\}$，则 A 中的恒等关系为

$$I_X=\{<1,1>,<2,2>,<a,a>\}.$$

注意：空关系、全域关系和恒等关系都是唯一的.

有了表达关系的各种方法，下面就可以对关系做进一步的讨论. 我们应该特别注意的是在集合 X 上的二元关系 R 的一些特殊性质.

定义 1.1.20　设 R 是集合 X 上的二元关系，如果对于任意的 $x\in X$，有 xRx，则称 R 是**自反的**，即

$$R\text{ 在 }X\text{ 上自反}\Leftrightarrow(\forall x)(x\in X\rightarrow xRx).$$

定义 1.1.21　设 R 是集合 X 上的二元关系，如果对于任意的 $x\in X$，都有 $<x,x>\notin R$，则称 R 为**反自反的**，即

$$R\text{ 在 }X\text{ 上反自反}\Leftrightarrow(\forall x)(x\in X\Rightarrow<x,x>\notin R).$$

[**例 1.1.22**]　设 $X=\{1,2,3\}$，给出 X 上的几个自反关系和反自反关系.

解　由定义可知

$$R_1=\{<1,1>,<2,2>,<3,3>,<1,3>,<2,3>\},$$
$$R_2=\{<1,1>,<2,2>,<3,3>,<1,2>\},$$
$$R_3=\{<1,1>,<2,2>,<3,3>,<3,2>\}$$

都是 X 上的自反关系. 另外，我们还注意到，X 上的全域关系和恒等关系也都是自反关系. 而

$$R_4=\{<1,2>,<1,3>,<2,3>\},$$
$$R_5=\{<3,2>\}$$

都是反自反关系.

$$R_6=\{<1,1>,<1,3>,<3,2>\},$$
$$R_7=\{<3,2>,<3,3>\}$$

既不自反，也不反自反的.

定义 1.1.22　设 R 是集合 X 上的二元关系，对于任意的 x，$y\in X$，若有 xRy 时，就有 yRx，则称 R 是**对称的**，即

$$R\text{ 在 }X\text{ 上对称}\Leftrightarrow(\forall x)(\forall y)(x\in X\wedge y\in X\wedge xRy\rightarrow yRx).$$

例如，在同一班级学习的同学关系是对称的，平面上三角形的相似关系是对称的，即若三角形 A 和三角形 B 相似，则 B 就相似于 A.

[**例 1.1.23**]　设 $A=\{2,3,5,7\}$，$R=\{<x,y>\mid\frac{x-y}{2}\text{是整数}\}$，证明 R 在 A 上是自反和对称的.

证明　因为对任意 $x\in A$，$\frac{x-x}{2}=0$，即 $<x,x>\in R$，所以 R 是自反的. 又设 x，$y\in A$，如果 $<x,y>\in R$，即 $\frac{x-y}{2}$ 是整数，则 $\frac{y-x}{2}$ 也必是整数，即 $<y,x>\in R$，因此 R 是对称的.

定义 1.1.23　设 R 是集合 X 上的二元关系，对于任意的 x，$y\in X$，若有 xRy，yRx，就有 $x=y$，则称 R 是**反对称的**，即

$$R\text{ 在 }X\text{ 上反对称}\Leftrightarrow(\forall x)(\forall y)(x\in X\wedge y\in X\wedge xRy\wedge yRx\rightarrow x=y).$$

例如，$A=\{1,2,3\}$，$S=\{<1,1>,<2,2>,<3,3>\}$，则 S 在 A 上是对称的也是反

对称的. 若 $S=\{<1,2>,<1,3>,<3,1>\}$，则 S 既不是对称关系，也不是反对称关系.

定义 1.1.24 设 R 是集合 X 上的二元关系，对于任意的 $x,y,z\in X$，若有 xRy，yRz，就有 xRz，则称 R 是**传递的**，即

$$R\text{ 在 }X\text{ 上传递}\Leftrightarrow(\forall x)(\forall y)(\forall z)(x\in X\wedge y\in X\wedge z\in X\wedge xRy\wedge yRz\rightarrow xRz).$$

例如，在实数集合中关系$<$，$>$和$=$关系，都是传递的. 又如，设 A 是人的集合，R 是 A 上的二元关系，若$<a,b>\in R$ 当且仅当 a 是 b 的祖先，则显然祖先关系 R 是传递的.

定义 1.1.25 设 R 是集合 X 上的二元关系，若 R 是自反、对称和传递的，则称 R 为 X 上的**等价关系**.

常见的等价关系有同一班级中的同学关系、直线间的平行关系等. 其主要意义在于它证实了应用抽象的一般原理的正确性，即在某些性质等价的个体中产生等价类，对全体的等价类进行分析往往比对全体本身进行分析更简单. 等价关系在后面章节中具有重要的应用.

二元关系是以序偶为元素的集合，所以它们也可以进行集合的运算，如交、并、补等而产生新的集合. 当然关系也可以进行一些其他的运算，如复合运算、逆运算、幂运算等.

定义 1.1.26 设 R 为 X 到 Y 的关系，S 为 Y 到 Z 的关系，则 $S\circ R$ 称为 R 和 S 的**复合关系**，即

$$S\circ R=\{<x,z>\mid x\in X\wedge z\in Z\wedge(\exists y)(y\in Y\wedge<x,y>\in R\wedge<y,z>\in S)\}.$$

所谓关系的复合运算或合成运算就是求 R 和 S 的复合关系 $S\circ R$. 例如，设 R 是人群中的父子关系，则 R 与 R 的复合关系就是祖孙关系.

复合运算是关系的二元运算，它能够由两个关系生成一个新关系，并且可以依次类推. 例如，R 是从 X 到 Y 的关系，S 是从 Y 到 Z 的关系，P 是从 Z 到 W 的关系，于是$(R\circ S)\circ P$ 和$R\circ(S\circ P)$都是从 X 到 W 的关系. 容易证明$(R\circ S)\circ P=R\circ(S\circ P)$，因此关系的复合运算是满足结合律的.

[例 1.1.24] 设集合 $X=\{1,2,3\}$，$Y=\{a,b,c\}$，$Z=\{\alpha,\beta,\chi\}$，R 是从 X 到 Y 的关系，$R=\{<1,a>,<2,b>,<2,c>\}$，S 是从 Y 到 Z 的关系，$S=\{<a,\beta>,<b,\alpha>,<c,\beta>,<c,\chi>\}$，求 $S\circ R$.

解 $S\circ R=\{<1,\beta>,<2,\alpha>,<2,\beta>,<2,\chi>\}$.

复合关系的关系矩阵也可以通过关系矩阵的逻辑乘法来求得. 设关系 R_1 的关系矩阵 $\boldsymbol{A}$ 和关系 R_2 的关系矩阵 $\boldsymbol{B}$ 分别为

$$\boldsymbol{A}=\begin{bmatrix}a_{11}&a_{12}&\cdots&a_{1n}\\a_{21}&a_{22}&\cdots&a_{2n}\\\vdots&\vdots&&\vdots\\a_{m1}&a_{m2}&\cdots&a_{mn}\end{bmatrix}$$

和

$$\boldsymbol{B}=\begin{bmatrix}b_{11}&b_{12}&\cdots&b_{1p}\\b_{21}&b_{22}&\cdots&b_{2p}\\\vdots&\vdots&&\vdots\\b_{n1}&b_{n2}&\cdots&b_{np}\end{bmatrix},$$

则复合关系 $R_2\circ R_1$ 的关系矩阵 $\boldsymbol{C}=\boldsymbol{A}\times\boldsymbol{B}$ 定义为

$$\boldsymbol{C}=\boldsymbol{A}\times\boldsymbol{B}=\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1p} \\ c_{21} & c_{22} & \cdots & c_{2p} \\ \vdots & \vdots & & \vdots \\ c_{m1} & c_{m2} & \cdots & c_{mp} \end{bmatrix},$$

其中 $c_{ik}=\bigvee\limits_{j=1}^{n}(a_{ij}\wedge b_{jk})$，$a_{ij}$ 和 b_{jk} 都只取 0 或 1，这里的运算是逻辑乘和逻辑加. 例如，例 1.1.24 中，R 和 S 的关系矩阵分别为

$$\boldsymbol{M}_R=\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

和

$$\boldsymbol{M}_S=\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{bmatrix},$$

而 $S\circ R$ 的关系矩阵为

$$\boldsymbol{M}_{S\circ R}=\begin{bmatrix} 0 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

定理 1.1.5　复合运算的性质如下：

(1) 满足结合律，即$(R\circ S)\circ P=R\circ(S\circ P)$；

(2) 不满足交换律，即 $R\circ S\neq S\circ R$；

(3) 复合运算对并运算满足分配律，即

$$R\circ(S\cup P)=(R\circ S)\cup(R\circ P),$$

$$(S\cup P)\circ R=(S\circ R)\cup(P\circ R);$$

(4) 复合运算对交运算满足下面的包含关系，即

$$R\circ(S\cap P)\subseteq(R\circ S)\cap(R\circ P),$$

$$(S\cap P)\circ R\subseteq(S\circ R)\cap(P\circ R);$$

(5) 设 R 是 X 到 Y 的关系，I_X 是 X 中的恒等关系，I_Y 是 Y 中的恒等关系，则

$$I_X\circ R=R\circ I_Y=R.$$

定义 1.1.27　设 R 是集合 A 上的二元关系，n 为自然数，则 R 的 n 次幂记作 R^n，并且规定

(1) $R^0=I_A$；

(2) $R^{n+1}=R^n\circ R$.

由定义可以看出，对于集合 A 上任意一个关系 R，都有 $R^0=I_A$，$R^1=R$.

[例 1.1.25]　设 $X=\{1,2,3\}$，X 中的二元关系 $R=\{<1,2>,<2,3>,<3,2>\}$，求 R 的各次幂.

解　由题意可知

$R^0=I_X=\{<1,1>,<2,2>,<3,3>\}$

$R^1=R=\{<1,2>,<2,3>,<3,2>\}$

$R^2=R\circ R=\{<1,3>,<2,2>,<3,3>\}$

$$R^3=R^2\circ R=\{<1,3>,<2,2>,<3,3>\}\circ\{<1,2>,<2,3>,<3,2>\}$$
$$=\{<1,2>,<2,3>,<3,2>\}=R$$
$$R^4=R^3\circ R=\{<1,2>,<2,3>,<3,2>\}\circ\{<1,2>,<2,3>,<3,2>\}=R^2$$
$$\vdots$$
$$R^{2n+1}=R=\{<1,2>,<2,3>,<3,2>\}$$
$$R^{2n+2}=R^2=\{<1,3>,<2,2>,<3,3>\}$$

幂运算有如下的性质.

定理 1.1.6 设 R 是集合 X 中的二元关系，$m,n\in\mathbf{N}$，则

(1) $R^m\circ R^n=R^{m+n}$；

(2) $(R^m)^n=R^{m\times n}$.

证明 (1) 对于任意给定的 $m\in\mathbf{N}$，对 n 用数学归纳法. 当 $n=0$ 时，

$$R^m\circ R^0=R^m\circ I_X=R^m=R^{m+0}.$$

假设当 $n=k$ 时，有 $R^m\circ R^k=R^{m+k}$. 当 $n=k+1$ 时，

$$R^m\circ R^{k+1}=R^m\circ(R^k\circ R)=(R^m\circ R^k)\circ R=R^{m+k}\circ R=R^{m+k+1}.$$

所以，对所有 $m,n\in\mathbf{N}$，都有 $R^m\circ R^n=R^{m+n}$.

(2) 读者可以模仿(1)进行证明.

定义 1.1.28 设 R 是 X 到 Y 的二元关系，若将 R 中每一序偶的元素顺序互换，则所得到的集合称为 R 的**逆关系**，记作 R^c，即

$$R^c=\{<y,x>|<x,y>\in R\}.$$

定理 1.1.7 设 R_1,R_2,R_3 都是从 A 到 B 的二元关系，则下列各式成立.

(1) $(R_1\cup R_2)^c=R_1^c\cup R_2^c$；

(2) $(R_1\cap R_2)^c=R_1^c\cap R_2^c$；

(3) $(A\times B)^c=B\times A$；

(4) $(\overline{R})^c=\overline{R^c}$(或$(\sim R)^c=\sim R^c$)；

(5) $(R_1-R_2)^c=R_1^c-R_2^c$.

证明 (1) 可知

$$<x,y>\in(R_1\cup R_2)^c\Leftrightarrow<y,x>\in R_1\cup R_2$$
$$\Leftrightarrow<y,x>\in R_1\vee<y,x>\in R_2$$
$$\Leftrightarrow<x,y>\in R_1^c\vee<x,y>\in R_2^c$$
$$\Leftrightarrow<x,y>\in R_1^c\cup R_2^c$$

类似地，读者可以自己证明(2)和(3).

(4) $<x,y>\in(\overline{R})^c\Leftrightarrow<y,x>\in\overline{R}\Leftrightarrow<y,x>\notin R\Leftrightarrow<x,y>\notin R^c\Leftrightarrow<x,y>\in\overline{R^c}$.

(5) 因为 $R_1-R_2=R_1\cap\overline{R_2}$，故

$$(R_1-R_2)^c=(R_1\cap\overline{R_2})^c=R_1^c\cap(\overline{R_2})^c=R_1^c\cap\overline{R_2^c}=R_1^c-R_2^c.$$

1.1.3 函数

1. 函数的定义

这里定义函数的概念为一组特殊的关系，函数的定义域和值域都是集合.

定义 1.1.29　设 X 和 Y 是任意两个集合，而 f 是 X 到 Y 的一个关系，若对每一个 $x\in X$，都有唯一的 $y\in Y$，使得$<x, y>\in f$，则称关系 f 为**函数**，记作 $f: X\to Y$ 或 $X\xrightarrow{f}Y$. 若$<x, y>\in f$，则 x 称为**自变元**，y 称为在 f 作用下 x 的**象**. $<x, y>\in f$ 也可以记作 $y=f(x)$，且记$f(X)=\{f(x)\mid x\in X\}$.

从函数的定义可以知道它与关系有以下两点区别：(1) 函数的定义域是 X，而不能是 X 的某个真子集；(2) 一个 $x\in X$ 只能对应于唯一的一个 y. 即如果 $f(x)=y$ 且 $f(x)=z$，那么 $y=z$，这称为函数的“单值性”. 从 X 到 Y 的函数也叫作从 X 到 Y 的**映射**.

在$<x, y>\in f$ 中，f 的第一个分量组成的集合就是函数 $y=f(x)$的**定义域**，记作 dom $f=X$，f 的**值域** ran$f\subseteq Y$，有时也记作 R_f，即

$$R_f=\{y\mid(\exists x)(x\in X\wedge y=f(x))\}.$$

集合 Y 称为 f 的**共域**，ran f 亦称为函数的**象集合**，很明显，ran $f=f(X)$.

注意：根据函数的记法 $f: X\to Y$，我们可以看出，即使两个函数的定义域和值域相同且它们包含的有序对也完全相同，但是如果它们的共域不同，那么这两个函数就是不同的函数，二者不等，这与中学学习的初等数学中的函数概念有略微的区别. 如此定义的原因在于，对于很多复杂的函数，我们很难一下就能够确定它的值域，但是确定它的共域要容易得多.

[**例 1.1.26**]　设 $X=\{1, 2, 3, 4\}$，$Y=\{a, b, c, d\}$，下列关系中哪些是函数？哪些不是函数？

$$f_1=\{<1, a>, <2, c>, <3, b>, <4, d>\},$$
$$f_2=\{<1, a>, <2, b>, <3, d>, <4, b>\},$$
$$f_3=\{<1, a>, <3, b>, <4, d>\},$$
$$f_4=\{<1, a>, <1, b>, <2, b>, <3, c>, <4, d>\}.$$

解　f_1 和 f_2 都是函数. f_3 不是函数，因为 $2\in X$，但没有对应的 y 值. f_4 也不是函数，因为 $x=1$ 对应了两个不同的值.

定义 1.1.30　设 f，g 都是从 A 到 B 的函数，若它们有相同的定义域和值域，并且对任意的 $x\in A$ 都有 $f(x)=g(x)$，则称函数 f 与 g **相等**，记作 $f=g$.

从函数的定义可以知道，$X\times Y$ 的子集并不都能成为 X 到 Y 的函数.

例如，设 $X=\{a, b, c\}$，$Y=\{0, 1\}$，则

$$X\times Y=\{<a, 0>, <b, 0>, <c, 0>, <a, 1>, <b, 1>, <c, 1>\}$$

有 2^6 个可能的子集，但其中只有 2^3 个子集定义为从 X 到 Y 的函数，即

$$f_0=\{<a, 0>, <b, 0>, <c, 0>\},\quad f_1=\{<a, 0>, <b, 0>, <c, 1>\},$$
$$f_2=\{<a, 0>, <b, 1>, <c, 0>\},\quad f_3=\{<a, 0>, <b, 1>, <c, 1>\},$$
$$f_4=\{<a, 1>, <b, 0>, <c, 0>\},\quad f_5=\{<a, 1>, <b, 0>, <c, 1>\},$$
$$f_6=\{<a, 1>, <b, 1>, <c, 0>\},\quad f_7=\{<a, 1>, <b, 1>, <c, 1>\}.$$

设 X 和 Y 都为有限集，分别有 m 个和 n 个不同元素，由于从 X 到 Y 任意一个函数的定义域是 X，在这些函数中每一个恰有 m 个序偶. 另外，任何元素 $x\in X$，可以有 Y 的 n 个元素中的任何一个作为它的象，故共有 n^m 个不同的函数. 在上例中 $n=2$，$m=3$，故应有 2^3 个不同的函数. 今后用符号 Y^X 表示从 X 到 Y 的所有函数的集合，甚至当 X 和 Y 是无限集时，也用这个符号.

下面讨论函数的几类特殊情况.

定义 1.1.31 对于 $X \xrightarrow{f} Y$ 的映射中，如果 ran $f=Y$，即 Y 的每一个元素是 X 中一个或多个元素的象，则称这个映射为**满射**.

设 $f: X \to Y$ 是满射，即对于任意 $y \in Y$，必存在 $x \in X$，使得 $f(x)=y$ 成立. 例如，$A=\{a, b, c, d\}$，$B=\{1, 2, 3\}$，若 $A \xrightarrow{f} B$ 为 $f(a)=1$，$f(b)=1$，$f(c)=3$，$f(d)=2$，则 f 是满射的.

定义 1.1.32 从 X 到 Y 的映射中，若 X 中没有两个元素有相同的象，则称这个映射为**入射**.

设 $f: X \to Y$ 是入射，即对于任意 x_1，$x_2 \in X$，$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$ 或者 $f(x_1)=f(x_2) \Rightarrow x_1=x_2$. 例如，函数 $f: \{a, b\} \to \{2, 4, 6\}$ 为 $f(a)=2$，$f(b)=6$，则这个函数是入射，但不是满射.

定义 1.1.33 从 X 到 Y 的映射，若既是满射又是入射，则称这个映射是**双射**的，也称这样的映射是一一对应的.

例如，令 $[a, b]$ 表示实数的闭区间，即 $[a, b]=\{x \mid a \leqslant x \leqslant b\}$，令 $f: [0, 1] \to [a, b]$，这里 $f(x)=(b-a)x+a$，这个函数是双射的.

定理 1.1.8 令 X 和 Y 为有限集，若 X 和 Y 的元素个数相同，记为 $|X|=|Y|$，则 $f: X \to Y$ 是入射的，当且仅当它是一个满射.

证明 若 f 是入射，则 $|X|=|f(X)|$，$|f(Y)|=|Y|$. 从 f 的定义有 $f(X) \subseteq Y$，而 $|f(Y)|=|Y|$，又因为 $|Y|$ 是有限的，故 $f(X)=Y$，因此，f 是满射.

若 f 是一个满射，根据满射定义，$f(X)=Y$，于是 $|X|=|Y|=|f(X)|$. 因为 $|X|=|f(X)|$，又 $|X|$ 是有限的，故 f 是一个入射，因此 f 是入射.

这个定理必须在有限集情况下才能成立，在无限集上不一定有效. 如 $f: I \to I$，这里 $f(x)=2x$，在这种情况下整数映射到偶整数，显然这是一个入射，但不是满射.

2. 逆函数和复合函数

在讨论二元关系时我们定义了逆关系，即从 X 到 Y 的关系 R，其逆关系 R^c 是 Y 到 X 的关系，符号化表示为 $<y, x> \in R^c \Leftrightarrow <x, y> \in R$. 但是对于函数就不能用简单的交换序偶的元素而得到逆函数，这是因为若有函数 $f: X \to Y$，但 f 的值域 R_f 可能只是 y 的一个真子集，即 $R_f \subset Y$，这不符合函数定义域的要求. 此外，若 $X \xrightarrow{f} Y$ 的映射是多对一的映射，即对于 $x_1 \neq x_2$，有 $<x_1, y> \in f$，$<x_2, y> \in f$，其逆关系将有 $<y, x_1> \in f^c$，$<y, x_2> \in f^c$，这就违反了函数值唯一性的要求. 为此，对函数求逆需要规定一些条件.

定理 1.1.9 设 $f: X \to Y$ 是一双射函数，那么 f^c 是 $Y \to X$ 的双射函数.

证明 设 $f=\{<x, y> \mid x \in X \wedge y \in Y \wedge f(x)=y\}$，$f^c=\{<y, x> \mid <x, y> \in f\}$. 因为 f 是满射，故每一个 $y \in Y$ 必存在 $<x, y> \in f$，因此必有 $<y, x> \in f^c$，即 f^c 第一个分量的集合为 Y. 又因为 f 是入射，对每一个 $y \in Y$ 恰有一个 $x \in X$，使 $<x, y> \in f$，因此仅有一个 $x \in X$，使 $<y, x> \in f^c$，即 y 对应唯一的 x，故 f^c 是函数.

又因为 ran $f^c=$ dom $f=X$，故 f^c 是满射. 又若 $y_1 \neq y_2$，有 $f^c(y_1) \neq f^c(y_2)$，因为 $f^c(y_1)=x_1$，$f^c(y_2)=x_2$，即 $x_1=x_2$，故 $f(x_1)=f(x_2)$，即 $y_1=y_2$，得出矛盾结果. 因此 f^c 是一个双射函数.

定义 1.1.34　设 $f: X \to Y$ 是一双射函数，称 $Y \to X$ 的双射函数 f^c 为 f 的**逆函数**，记作 f^{-1}.

例如，设 $A=\{1, 2, 3\}$，$B=\{a, b, c\}$，若 $f: A \to B$ 为 $f=\{<1, a>, <2, c>, <3, b>\}$，则 $f^{-1}=\{<a, 1>, <c, 2>, <b, 3>\}$；若 $f=\{<1, a>, <2, b>, <3, b>\}$，则 $f^{-1}=\{<a, 1>, <b, 2>, <b, 3>\}$就不是一个函数.

定义 1.1.35　设函数 $f: X \to Y$，$g: W \to Z$，若 $f(X) \subseteq W$，则

$$g \circ f=\{<x, z> \mid x \in X \wedge z \in Z \wedge (\exists y)(y \in Y \wedge y=f(X) \wedge z=g(y))\},$$

称 g 在函数 f 的左边可复合.

定理 1.1.10　两个函数的复合是一个函数.

证明　设 $g: W \to Z$，$f: X \to Y$ 为左复合，即 $f(X) \subseteq W$. 对于任意 $x \in X$，因为 f 为函数，故必有唯一的序偶$<x, y>$使 $y=f(x)$成立，而 $f(x) \in f(X)$，即 $f(x) \in W$. 又因为 y 是函数，故必有唯一序偶$<y, z>$使 $z=g(y)$成立. 根据复合定义，$<x, z> \in g \circ f$，即 X 中每个 x 对应 Z 中某个 z.

假定 $g \circ f$ 中包含序偶$<x, z_1>$和$<x, z_2>$，且 $z_1 \neq z_2$，这样在 Y 中必存在 y_1 和 y_2，使得在 f 中有$<x, y_1>$和$<x, y_2>$，在 g 中有$<y_1, z_1>$和$<y_2, z_2>$. 因为 f 是一个函数，故 $y_1=y_2$. 于是 g 中有$<y, z_1>$和$<y, z_2>$，但 g 是一个函数，故 $z_1=z_2$，即每个 $x \in X$ 只能有唯一的$<x, z> \in g \circ f$. 因此，$g \circ f$ 是一个函数.

[例 1.1.27]　设 $X=\{1, 2, 3\}$，$Y=\{p, q\}$，$Z=\{a, b\}$，而 $f=\{<1, p>, <2, p>, <3, q>\}$，$g=\{<p, b>, <q, b>\}$，求 $g \circ f$.

解　$g \circ f=\{<1, b>, <2, b>, <3, b>\}$.

定理 1.1.11　令 $g \circ f$ 是一个复合函数.

(1) 若 g 和 f 是满射的，则 $g \circ f$ 是满射的；

(2) 若 g 和 f 是入射的，则 $g \circ f$ 是入射的；

(3) 若 g 和 f 是双射的，则 $g \circ f$ 是双射的.

证明　(1) 设 $f: X \to Y$，$g: Y \to Z$，令 z 为 Z 的任意一个元素，因 g 是满射，故必有某个元素 $y \in Y$，使得 $g(y)=z$. 又因为 f 是满射，故必有某个元素 $x \in X$，使得 $f(x)=y$，故

$$g \circ f(x)=g(f(x))=g(y)=z.$$

因此 $R_{g \circ f}=Z$，$g \circ f$ 是满射的.

(2) 令 x_1，x_2 为 X 的元素，假定 $x_1 \neq x_2$，因为 f 是入射的，故 $f(x_1) \neq f(x_2)$. 又因为 g 是入射的且 $f(x_1) \neq f(x_2)$，故 $g(f(x_1)) \neq g(f(x_2))$，于是 $x_1 \neq x_2 \Rightarrow g \circ f(x_1) \neq g \circ f(x_2)$，因此，$g \circ f$ 是入射的.

(3) 因为 g 和 f 是双射的，根据(1)和(2)，$g \circ f$ 是满射和入射的，即双射.

由于函数的复合仍然是一个函数，故可求三个函数的复合.

[例 1.1.28]　设 $\mathbf{R}$ 为实数集合，对 $x \in \mathbf{R}$，有 $f(x)=x+2$，$g(x)=x-2$，$h(x)=3x$，求 $g \circ f$ 与 $h \circ (g \circ f)$.

解　由已知可得

$$g \circ f=\{<x, x> \mid x \in \mathbf{R}\}, \quad h \circ (g \circ f)=\{<x, 3x> \mid x \in \mathbf{R}\}.$$

一般地，有 $h \circ (g \circ f)=(h \circ g) \circ f$. 函数的复合函数是可结合的，故可以去掉上式中的括号.

定义 1.1.36 函数 $f: X\to Y$ 叫作**常函数**，如果存在某个 $y_0\in Y$，则对于每个 $x\in X$ 都有 $f(x)=y_0$，即 $f(X)=y_0$.

定义 1.1.37 如果 $I_X=\{<x, x>|x\in X\}$，则称 $I_X: X\to X$ 为**恒等函数**.

定理 1.1.12 设函数 $f: X\to Y$，则 $f=f\circ I_X=I_Y\circ f$.

这个定理的证明可以由定义直接得到.

定理 1.1.13 若 $f: X\to Y$ 是双射函数，则 $(f^{-1})^{-1}=f$.

证明 因为 $f: X\to Y$ 是双射函数，故 $f^{-1}: Y\to X$ 也是双射函数，因此

$$(f^{-1})^{-1}: X\to Y$$

也为双射函数，显然 $\mathrm{dom}\ f=\mathrm{dom}(f^{-1})^{-1}=X$. 又由

$$x\in X\Rightarrow f: x\to f(x)\Rightarrow f^{-1}: f(x)\to x\Rightarrow (f^{-1})^{-1}: x\to f(x),$$

可知 $(f^{-1})^{-1}=f$.

定理 1.1.14 若 $f: X\to Y$，$g: Y\to Z$ 都是双射函数，则 $(g\circ f)^{-1}=f^{-1}\circ g^{-1}$.

证明 因为 $f: X\to Y$，$g: Y\to Z$ 都是双射函数，所以 f^{-1} 和 g^{-1} 都是双射函数，即 $f^{-1}: Y\to X$，$g^{-1}: Z\to Y$，所以 $f^{-1}\circ g^{-1}: Z\to X$. 又 f，g 是双射函数，故 $g\circ f$ 是双射函数，所以 $(g\circ f)^{-1}$ 是双射的，即 $(g\circ f)^{-1}: Z\to X$. 因此

$$\mathrm{dom}(f^{-1}\circ g^{-1})=\mathrm{dom}(g\circ f)^{-1}=Z.$$

由 $\forall z\in Z\Rightarrow$ 存在唯一 $y\in Y$，使得 $g(y)=z\Rightarrow$ 存在唯一 $x\in X$，使得 $f(x)=y$，故

$$(f^{-1}\circ g^{-1})(z)=f^{-1}(g^{-1}(z))=f^{-1}(y)=x,$$

又 $(g\circ f)(x)=g(f(x))=g(y)=z$，故 $(g\circ f)^{-1}(z)=x$，因此对 $\forall z\in Z$ 有

$$(g\circ f)^{-1}(z)=(f^{-1}\circ g^{-1})(z).$$

因此 $(g\circ f)^{-1}=f^{-1}\circ g^{-1}$.

3. 集合的等势

有了双射函数的概念，就可以比较两个集合的“大小”是否相等，确定有限集和无限集的概念. 下面首先需要引进自然数集合.

定义 1.1.38 给定集合 A 的**后继集**定义为集合 $A^+=A\cup\{A\}$. 若 A 为空集$\varnothing$，则后继集为$\varnothing^+$，$(\varnothing^+)^+$，…，这些集合可写成如下形式：

$\varnothing\cup\{\varnothing\}$，$\varnothing\cup\{\varnothing\}\cup\{\varnothing\cup\{\varnothing\}\}$，$\varnothing\cup\{\varnothing\}\cup\{\varnothing\cup\{\varnothing\}\}\cup\{\varnothing\cup\{\varnothing\}\cup\{\varnothing\cup\{\varnothing\}\}\}$，…，可简化为$\{\varnothing\}$，$\{\varnothing,\{\varnothing\}\}$，$\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}$，….

若命名集合$\varnothing$为 0，那么，

$\varnothing^+=0^+=\{\varnothing\}\triangleq 1$，

$1^+=\{\varnothing,\{\varnothing\}\}\triangleq 2$，

$2^+=\{\varnothing,\{\varnothing\},\{\varnothing,\{\varnothing\}\}\}\triangleq 3$，

⋮

这样就得到了自然数集合$\{0, 1, 2, 3, \cdots\}$，这个集合亦能概括为如下公理形式(即 G. Peano 公理).

(1) $0\in\mathbf{N}$(其中 $0\triangleq\varnothing$)；

(2) 如果 $n\in\mathbf{N}$，那么 $n^+\in\mathbf{N}$(其中 $n^+=n\cup\{n\}$)；

(3) 如果一个子集 $S\subseteq\mathbf{N}$ 具有性质：

① $0\in S$，

② 如果 $n\in S$，有 $n^+\in S$，

则 $S=\mathbf{N}$.

性质(3)称极小性质，它指明了自然数系统的最小性，即自然数系统是满足公理(1)和(2)的最小集合.

当然，自然数集亦可不从 0 开始，这只需定义∅为 1，则自然数集就从 1 开始.

从上述定义可以看到，任意一个自然数可看作是某个集合的名字. 此外，从实际生活中我们知道，任意自然数，例如 3 这个概念是从观察许多只含三个元素的集合的共同特点而抽象概括出来的，这个共同特点就是体现于这些被观察的任意一个集合的元素都可与集合{∅，{∅}，{∅，{∅}}}中元素存在一一对应，且其任意两个集合的元素之间也存在一一对应. 由此可见，“对应”是集合之间进行比较的一个非常重要的概念.

例如，{2，4，6，…，$2n$，…}与{1，3，5，…，$2n-1$，…}之间存在着一一对应.

定义 1.1.39　当且仅当在集合 A 与集合 B 之间存在一一对应的函数时，集合 A 与集合 B 称为**等势的**(或称**同浓的**)，记作 $A\sim B$.

[例 1.1.29]　验证自然数集 $\mathbf{N}$ 与非负偶数集合 M 是等势的.

证明　由于 $\mathbf{N}$ 与 M 的元素之间可作一一对应的映射，即 $f(n)=2n$，故 $\mathbf{N}$ 与 M 是等势的.

[例 1.1.30]　设 P 为实数集合，S 是 P 的子集，即 $S\subseteq P$，且 $S=\{x\mid x\in P\wedge 0<x<1\}$，求证 $S\sim P$.

证明　令 $f: P\to S$，且 $f(x)=\dfrac{1}{\pi}\arctan x+\dfrac{1}{2}(-\infty<x<+\infty)$. 显然 f 的值域是 S，且 f 是双射函数.

[例 1.1.31]　证明区间[0，1]与(0，1)等势.

证明　设集合 $A=\{0,1,\dfrac{1}{2},\cdots,\dfrac{1}{n},\cdots\}$，显然 $A\subseteq[0,1]$. 定义 $f:[0,1]\to(0,1)$，使得

$$\begin{cases} f(0)=\dfrac{1}{2} & \\ f\left(\dfrac{1}{n}\right)=\dfrac{1}{n+2} & (\text{对 } n\geqslant 1) \\ f(x)=x & (\text{对 } x\in[0,1]-A) \end{cases}$$

可知 f 是双射函数. 命题得证.

定理 1.1.15　在集合族上等势关系是一个等价关系.

证明　设集合族为 S.

(1) 对任意 $A\in S$，必有 $A\sim A$.

(2) 若 A，$B\in S$，如果 $A\sim B$，必有 $B\sim A$.

(3) 若 A，B，$C\in S$，如果 $A\sim B$ 且 $B\sim C$，必有 $A\sim C$.

定义 1.1.40　如果有一个从集合{0，1，…，$n-1$}到 A 的双射集合，那么称集合 A 是**有限的**；如果集合 A 不是有限的，则它是**无限的**.

定理 1.1.16　自然数集合 $\mathbf{N}$ 是无限的.

证明　设 n 是 $\mathbf{N}$ 的任意元素，f 是任意的从{0，1，…，$n-1$}到 $\mathbf{N}$ 的函数. 设

$$k=1+\max\{f(0),f(1),\cdots,f(n-1)\},$$

那么 $k\in\mathbf{N}$，但对每一个 $x\in\{0,1,\cdots,n-1\}$，有 $f(x)\neq k$. 因此 f 不能是满射函数，即 f 也不是双射函数. 因为 n 和 f 都是任意的，故 $\mathbf{N}$ 是无限的.

对于有限集的大小概念很容易理解，**基数**是无限集的大小的核心概念，由于这个概念比较复杂，请有兴趣的读者参考有关集合论书籍.

1.2 组合数学初步知识

1.2.1 排列与组合

1. 基本计数原理

加法原理与乘法原理是两个最基本的计数原理. 在介绍它们之前，先引入集合的划分这个概念.

定义 1.2.1 设 $S_1,S_2,\cdots,S_m$ 是集合 S 的子集，且

$$S=S_1\cup S_2\cup\cdots\cup S_m,\text{ 其中 } S_i\cap S_j=\varnothing,\quad i\neq j,$$

则称 $S_1,S_2,\cdots,S_m$ 是 S 的一个**划分**.

将集合 S 中所含元素的个数记为 $|S|$. 下面的加法原理实质就是全体等于各部分之和这一原理的公式化.

定理 1.2.1(加法原理) 设 $S_1,S_2,\cdots,S_m$ 是集合 S 的一个划分，则

$$|S|=|S_1|+|S_2|+\cdots+|S_m|.$$

加法原理理解起来很容易，它还可以叙述为：若完成一件事情有 m 个方案，第 i ($i=1,2,\cdots,m$)个方案有 n_i 种方法可以实现，只要选择任一方案中的任一方法，就可以将这件事情完成，并且这些方法两两互不相同，则完成这件事情共有 $n_1+n_2+\cdots+n_m$ 种方法. 也可以这样叙述加法原理：如果共有 m 堆物体，且有 n_i 种方法能够从第 i ($i=1,2,\cdots,m$) 堆物体中选择一个物体，那么从这 m 堆物体中选择一个物体的方法共有 $n_1+n_2+\cdots+n_m$ 种.

[例 1.2.1] 某班有男生 30 人，女生 7 人，现在需要从该班任选一学生作为代表去参加会议，问共有多少种选法？

解 令

$$S=\{\text{班内所有学生}\},$$
$$A=\{\text{班内所有男生}\},$$
$$B=\{\text{班内所有女生}\}.$$

显然 A,B 是 S 的一个划分，则

$$|S|=|A|+|B|=30+7=37,$$

即全班共有 37 个学生，故有 37 种选法.

定理 1.2.2(乘法原理) 设 $S_1,S_2,\cdots,S_m$ 是 m 个有限集，则

$$|S_1\times S_2\times\cdots\times S_m|=|S_1|\cdot|S_2|\cdot\cdots\cdot|S_m|.$$

乘法原理还可以叙述为：若完成一件事情需要 m 个步骤，第 i ($i=1,2,\cdots,m$)个步骤有 n_i 种方法可以实现，每个步骤中方法的选取均与前面的步骤无关，则完成这件事情共有 $n_1\cdot n_2\cdot\cdots\cdot n_m$ 种方法.

[例 1.2.2]　某班有男生 30 人，女生 7 人，现在需要从该班分别选出男女生各一名作为代表去参加会议，问共有多少种选法？

解　令

$$A=\{\text{班内所有男生}\},$$

$$B=\{\text{班内所有女生}\}.$$

根据乘法原理，有$|A\times B|=|A|\cdot|B|=30\times 7=210$，即共有 210 种选法.

[例 1.2.3]　求整数 $3^7\times 5^8\times 11^9$ 的正整数因子的个数.

解　整数 $3^7\times 5^8\times 11^9$ 的正整数因子只能具有 $3^i\times 5^j\times 11^k$ 这样的形式，其中，$0\leqslant i\leqslant 7$，$0\leqslant j\leqslant 8$，$0\leqslant k\leqslant 9$. 显然，i 有 8 个选择，j 有 9 个选择，k 有 10 个选择，由乘法原理可知，正因子数目为 $8\times 9\times 10=720$.

相应地，还可以得到上面两个原理的“补”原理.

定理 1.2.3(减法原理)　设 E 为全集，则

$$|A|=|E|-|E-A|.$$

当应用减法原理的时候，通常选择某个自然的容易得到的集合 E 作为全集，它包含 A 的全部元素，且对 E 的元素和对 $E-A$ 的元素进行计数要比对 A 的元素进行计数容易的时候，才应用这个原理.

[例 1.2.4]　求比整数 $3^3\times 5^2\times 11^1$ 小且不是其因子的正整数个数.

解　由乘法原理可知，整数 $3^3\times 5^2\times 11^1$ 的正整数因子共有 $4\times 3\times 2=24$ 个. 则由减法原理可知，所求正整数个数为 $3^3\times 5^2\times 11^1-24=7\,401$ 个.

定理 1.2.4(除法原理)　设 S 为有限集合，它被划分为 m 个部分，且每个部分所含元素数量相同，则

$$m=\frac{|S|}{\text{单独一个部分中的元素数量}}.$$

2. 排列

为了今后表达方便，将具有 n $(n>0)$个元素的有限集称为 n 集，其具有 $r(0\leqslant r\leqslant n)$个元素的子集称为 r 子集. 0 子集是空集$\varnothing$.

定义 1.2.2　设 S 是一个 n 集，r 是正整数，则笛卡儿积

$$\underbrace{S\times S\times\cdots\times S}_{r}=\{<a_1,a_2,\cdots,a_r>\mid a_i\in S,\ i=1,2,\cdots,r\}$$

的元素$<a_1,a_2,\cdots,a_r>$称为 S 的一个 ***r* 重排列**，其个数记为 $RP(\infty,r)$.

定理 1.2.5　n 集 S 的 r 重排列的个数为 $RP(\infty,r)=n^r$.

证明　根据乘法原理，有

$$RP(\infty,r)=\left|\underbrace{S\times S\times\cdots\times S}_{r}\right|=\prod_{i=1}^{r}|S|=\prod_{i=1}^{r}n=n^r.$$

[例 1.2.5]　一个盒子中装有红、白、蓝三色球各一个. 每次从盒子中任取一个球，记录下此球的颜色后再将此球放回，如此进行两次，问共可能出现多少种颜色序列记录？若如此取球 5 次又如何？

解　先用枚举的方法来解决取球两次的问题. 显然，可能出现的颜色序列记录有：

红红，　红白，　红蓝，

白红，　白白，　白蓝，

蓝红， 蓝白， 蓝蓝.

于是，共有 9 种可能. 而这是一个很典型的二重排列问题，可以直接利用上面的公式求解，即

$$RP(\infty, 2)=3^2=9.$$

这与用枚举法所得的结果是相同的. 同理，对于取球 5 次的情况，有

$$RP(\infty, 5)=3^5=243,$$

即共可能出现 243 种颜色序列记录，这个结果如果用枚举法求解将会是十分痛苦的.

定义 1.2.3 若 n 集 S 的 r 重排列$<a_1, a_2, \cdots, a_r>$中的分量互不相同，则称这个 r 重排列为 S 的***r* 排列**，或称 n 集的 r 排列. n 集中所有不同 r 排列的个数记为 $P(n, r)$. 当 $n=r$ 时，n 集的 r 排列简称为 n 集的**全排列**.

今后，更多地是讨论这种无重复的排列.

排列的概念又可以用另一种更直观的方式叙述：设 $A=\{a_1, a_2, \cdots, a_n\}$是一 n 集，从 A 中任取 r 个不同元素按顺序排成一列，称为从 n 中取 r 的一个排列(即 A 的 r 排列).

显然，若 $r>n$，则 $P(n, r)=0$. 我们还约定 $P(n, 0)=1$.

定理 1.2.6 $P(n, r)=n(n-1)\cdots(n-r+1)$.

证明 当 $r>n$ 或 $r=0$ 时，定理显然成立. 下面讨论 $0<r\leqslant n$ 时的情形.

n 集的一个 r 排列设为$<a_1, a_2, \cdots, a_r>$，它由 r 个不同的元素组成，可依次分步取 a_1，a_2，…，a_r. 其中 a_1 可有 n 种取法，那么 a_2 则有 $n-1$ 种取法，依此类推，a_r 有 $n-(r-1)$种取法. 根据乘法原理，n 集的 r 排列数为

$$P(n, r)=n(n-1)\cdots(n-r+1).$$

当 $r\leqslant n$ 时，上面的式子也可写为

$$P(n, r)=\frac{n!}{(n-r)!}.$$

今后在讨论排列时，若无特别说明，一般均默认 $r\leqslant n$.

对于 n 集的全排列数，有如下推论.

推论 $P(n, n)=n!$.

[例 1.2.6] 求 20 000 到 70 000 间的偶数中由不同数字组成的五位数的个数.

解 设所求五位数为 $a_5a_4a_3a_2a_1$，其中 $2\leqslant a_5\leqslant 6$，$a_1\in\{0, 2, 4, 6, 8\}$. 下面分两种情况进行讨论.

若 $a_5\in\{2, 4, 6\}$，则 a_1 有 4 种取法，$a_4a_3a_2$ 有 $P(10-2, 3)=P(8, 3)$种取法. 根据乘法原理，可知此情况有

$$3\times 4\times P(8, 3)=4\,032$$

种取法.

若 $a_5\in\{3, 5\}$，则 a_1 有 5 种取法，$a_4a_3a_2$ 仍有 $P(8, 3)$种取法. 根据乘法原理，可知此情况有

$$2\times 5\times P(8, 3)=3\,360$$

种取法.

再根据加法原理，可知总的个数为

$$4\,032+3\,360=7\,392.$$

[例 1.2.7]　英文单表置换密码中，加密时将明文中的同样字母同时替换为另一个字母形成密文，求共有多少种可能的密钥.

解　由于英文中共有 26 个字母，因此要为它们搭配 26 个密文字母，那么一种固定的搭配方式就是一个密钥. 此外，由于加密以后的密文必须能够正常解密，就要求搭配的 26 个密文字母之间不能有任意二者相同，因此一个密钥实质上就是 26 个字母的一个全排列，所以密钥的总数为 26！种.

上面所讨论的这些排列均是将元素排成一行，可以称其为**线排列**. 若将这种线排列首尾相接，也就是使元素排列成一个圆圈，则称这种排列为**圆排列**.

在 r 个相异元素组成的圆排列中，将围成圆圈的 r 个元素同时按一个方向旋转，即每个元素都向左或向右转动一个位置，虽然使元素的绝对位置发生了变化，但相对位置不变，也就是说元素之间的相邻关系不变，这样的圆排列认为是同一种，否则就是不同的圆排列.

从 n 集中取 r 个相异的元素组成圆排列，其所有不同的圆排列的总数记为 $CP(n, r)$. 可以显然得到下面的定理.

定理 1.2.7　$CP(n, r)=\dfrac{P(n, r)}{r}=\dfrac{n!}{r(n-r)!}$.

[例 1.2.8]　将 n 个有标号的珠子排成一个圆圈，问共有多少种不同的排法？

解　根据上面的定理，有

$$CP(n, n)=\frac{P(n, n)}{n}=(n-1)!,$$

即共有$(n-1)!$种不同的排法.

3. 组合

定义 1.2.4　设 S 是一个 n 集，r 是一个非负整数，则称$[b_1, b_2, \cdots, b_r]$为 S 的***r* 可重组合**，其中 $b_1, b_2, \cdots, b_r \in S$，但未必互不相同. S 中的元素 b 在 r 可重组合$[b_1, b_2, \cdots, b_r]$中出现的次数称为 b 在这个可重组合中的**重数**.

例如，若 $S=\{x, y, z\}$，则$[x, y, x]$是 S 的 3 可重组合，且在此 3 可重组合中，x 的重数为 2，y 的重数为 1，z 的重数为 0.

n 集 $S=\{a_1, a_2, \cdots, a_n\}$的两个 r 可重组合称为**相等**，如果 S 中的每个元素在这两个 r 可重组合中的重数相等. 因此，有时也将 S 的 r 可重组合记为

$$[r_1a_1, r_2a_2, \cdots, r_na_n],$$

其中 $r_i(i=1, 2, \cdots, n)$是非负整数，表示 a_i 在此 r 可重组合中的重数，且满足

$$r_1+r_2+\cdots+r_n=r.$$

例如，对集合 $S=\{x, y, z\}$，其 3 可重组合$[x, y, x]$，$[x, x, y]$和$[y, x, x]$都是相等的，均可记为$[2x, 1y, 0z]$.

S 的所有不同 r 可重组合的个数记为 $RC(\infty, r)$.

定义 1.2.5　设 S 是一个 n 集，$[b_1, b_2, \cdots, b_r]$为 S 的 r 可重组合，若 S 中的每个元素在此 r 可重组合中的重数均为 0 或 1，即 $b_1, b_2, \cdots, b_r$ 互不相同，则称$[b_1, b_2, \cdots, b_r]$为 S 的***r* 组合**，或称 n 集的 r 组合. n 集中所有不同 r 组合的个数记为 $C(n, r)$或$\dbinom{n}{r}$.

我们发现，n 集的 r 组合其实就是 n 集的 r 子集，所以 n 集的 r 组合常用集合的记号$\{b_1,$

b_2，…，b_r}来表示.

0 可重组合就是 0 组合，也是 0 子集.

在实际中，还是这种无重复的组合应用得最多.

组合的概念又可以用另一种更直观的方式叙述：设 $A=\{a_1, a_2, \cdots, a_n\}$是一 n 集，从 A 中任取 r 个不同元素构成一组，称为从 n 中取 r 的一个组合(即 A 的 r 组合).

显然，若 $r>n$，则 $C(n, r)=0$. 我们还约定 $C(n, 0)=1$.

定理 1.2.8 $C(n, r)=\dfrac{P(n, r)}{r!}=\dfrac{n(n-1)\cdots(n-r+1)}{r!}$.

证明 当 $r>n$ 或 $r=0$ 时，定理显然成立. 下面讨论 $0<r\leqslant n$ 时的情形.

设 S 是一个 n 集，由于 S 的每个 r 组合可产生 S 的 $r!$ 个不同的 r 排列，而如此产生的 S 的 $r!C(n, r)$个不同的 r 排列恰好就是 S 的所有 r 排列，故可知

$$r!C(n, r)=P(n, r)=n(n-1)\cdots(n-r+1),$$

所以

$$C(n, r)=\frac{P(n, r)}{r!}=\frac{n(n-1)\cdots(n-r+1)}{r!}.$$

当 $r\leqslant n$ 时，上面的式子也可写为

$$C(n, r)=\frac{n!}{r!\ (n-r)!}.$$

今后在讨论组合时，若无特别说明，一般均默认 $r\leqslant n$.

[例 1.2.9] 某研究所有甲乙两个实验室，其中甲实验室有 7 名成员，乙实验室有 4 名成员，现欲从两个实验室共抽调 5 名成员组成联合技术小组进行项目攻关，若要求其中至少有 2 名成员来自乙实验室，问共有多少种抽调方案?

解 技术小组中来自乙实验室的人数可以是 2，3，4，于是分不同情况讨论.

若乙实验室 2 人，甲实验室 3 人，则方案数为

$$C(4, 2)C(7, 3)=\frac{4!}{2!\times 2!}\times\frac{7!}{3!\times 4!}=210.$$

若乙实验室 3 人，甲实验室 2 人，则方案数为

$$C(4, 3)C(7, 2)=\frac{4!}{3!}\times\frac{7!}{2!\times 5!}=84.$$

若乙实验室 4 人，甲实验室 1 人，则方案数为

$$C(4, 4)C(7, 1)=7.$$

最后，根据加法原理，可知总的抽调方案数为

$$210+84+7=301.$$

[例 1.2.10] 计算字符串“AAABBCCCC”的重合指数(即随机从字符串中取出两个字符时，二者有相同的概率，重合指数的计算在破译维吉尼亚密码中起到关键作用).

解 根据古典概率模型可知，所求概率为

$$p=\frac{\text{选到的两个字母相同的选择种数}}{\text{所有可能选择的种数}},$$

由于选择中没有顺序的要求，所以这里的选择是组合问题，而不是排列问题. 其中，所有可能选择的种数为 $C(9, 2)=36$，选到的两个字母同时为 A 的选择种数为 $C(3, 2)=3$，选到的

两个字母同时为 B 的选择种数为 $C(2,2)=1$，选到的两个字母同时为 C 的选择种数为 $C(4,2)=6$，所以

$$p=\frac{3+1+6}{36}=\frac{5}{18}.$$

推论　对于 $r\leqslant n$，$C(n,r)=C(n,n-r)$.

证明　$C(n,r)=\dfrac{n!}{r!(n-r)!}=C(n,n-r)$.

其实这个推论很好理解，因为若从 n 个元素中取走 r 个，必然余下 $n-r$ 个，故从 n 取 r 的组合与从 n 取 $n-r$ 的组合一一对应.

定理 1.2.9　n 集中所有不同的 r 可重组合的个数为

$$RC(\infty,r)=C(n+r-1,r).$$

证明　取 n 集 $S=\{1,2,\cdots,n\}$，于是 S 的每个 r 可重组合均可唯一地写成

$$[a_1,a_2,\cdots,a_r],\quad a_1\leqslant a_2\leqslant\cdots\leqslant a_r.$$

可设 S 的所有 r 可重组合(写成上述形式)所组成的集合为 G.

设 $(n+r-1)$ 集 $S^*=\{1,2,\cdots,n+r-1\}$，于是 S^* 的每个 r 组合均可唯一地写成

$$\{b_1,b_2,\cdots,b_r\},\quad b_1<b_2<\cdots<b_r.$$

可设 S^* 的所有 r 组合(写成上述形式)所组成的集合为 G^*. 因为

$$1\leqslant a_1<a_2+1<a_3+2<\cdots<a_r+r-1\leqslant n+r-1,$$

所以

$$\{a_1,a_2+1,a_3+2,\cdots,a_r+r-1\}\in G^*.$$

由此可知，G 与 G^* 的元素间存在着一一对应. 于是

$$RC(\infty,r)=|G|=|G^*|=C(n+r-1,r).$$

定理得证.

[例 1.2.11]　问 $(x+y+z)^4$ 有多少项?

解　由于 $(x+y+z)^4=(x+y+z)(x+y+z)(x+y+z)(x+y+z)$，显然其展开式的每一项次数均为 4，且是 $\{x,y,z\}$ 的 4 可重组合. 因此，这实际上是一个 $n=3$，$r=4$ 的可重组合问题，其组合数即项数为

$$C(3+4-1,4)=C(6,4)=15.$$

也可枚举出 $(x+y+z)^4$ 的各项：x^4，y^4，z^4，x^2y^2，x^2z^2，y^2z^2，x^2yz，xy^2z，xyz^2，xy^3，xz^3，x^3y，x^3z，y^3z，z^3y. 它们恰好是 15 项，这也验证了上面的结果是正确的.

在排列组合的诸多公式中，有许多有着很明显的实际意义，理解起来也并不困难. 下面就介绍一些与组合数有关的公式，并说明其实际的组合意义.

定理 1.2.10(加法公式)　$C(n,r)=C(n-1,r)+C(n-1,r-1)$.

我们可以讨论一下它的组合意义. 设 n 集 $S=\{a_1,a_2,\cdots,a_n\}$，其所有的 r 组合可分为两类：

(1) 包含元素 a_1 的 r 组合. 此类组合可视为先确定地选出 a_1，再从剩下的 $n-1$ 个元素中任取 $r-1$ 个元素而构成的组合，其组合数为 $C(n-1,r-1)$.

(2) 不包含元素 a_1 的 r 组合. 此类组合可视为从除 a_1 以外的 $n-1$ 个元素中任取 r 个元素而构成的组合，其组合数为 $C(n-1,r)$.

根据加法原理，便可得到上面的加法公式.

定理 1.2.11(乘法公式) $C(n,k)C(k,r)=C(n,r)C(n-r,k-r)$.

我们可以通过讨论乘法公式的等价形式

$$C(n,n-k)C(k,k-r)C(r,r)=C(n,r)C(n-r,n-k)C(k-r,k-r)$$

来得到它的组合意义. 等式左端可认为是组合问题“将 n 个元素分为 3 组，要求第一组有 $n-k$ 个元素，第二组有 $k-r$ 个元素，第三组有 r 个元素”的组合数. 等式右端是另一个类似的组合问题“将 n 个元素分为 3 组，要求第一组有 r 个元素，第二组有 $n-k$ 个元素，第三组有 $k-r$ 个元素”的组合数. 这两个组合问题显然是等价的，所以其组合数必然相等. 于是，可知乘法公式的等价形式成立，则乘法公式也成立.

定理 1.2.12

$$\begin{aligned}C(n+r+1,r)&=\sum_{i=0}^{r}C(n+i,i)\\&=C(n+r,r)+C(n+r-1,r-1)+C(n+r-2,r-2)+\cdots+C(n,0)\end{aligned}$$

当然，上面这个公式也可以写成

$$\begin{aligned}C(n+r+1,r)&=\sum_{i=0}^{r}C(n+i,n)\\&=C(n+r,n)+C(n+r-1,n)+C(n+r-2,n)+\cdots+C(n,n)\end{aligned}$$

我们再来讨论它的组合意义. 设$(n+r+1)$集 $S=\{a_1,a_2,\cdots,a_{n+r+1}\}$，对于 a_1 可将所有的 r 组合分为两类，即包含元素 a_1 的 r 组合和不包含元素 a_1 的 r 组合. 我们先考虑不包含元素 a_1 的 r 组合，它相当于从$(n+r)$集$\{a_2,a_3,\cdots,a_{n+r+1}\}$中取 r 个的组合，其组合数为 $C(n+r,r)$. 类似地，再将所有包含元素 a_1 的 r 组合对于 a_2 分为两类，即包含元素 a_2 的 r 组合和不包含元素 a_2 的 r 组合. 同样考虑不包含元素 a_2 的 r 组合，它相当于从$(n+r-1)$集$\{a_3,a_4,\cdots,a_{n+r+1}\}$中取 $r-1$ 个，再加上 a_1 而构成的 r 组合，其组合数为 $C(n+r-1,r-1)$. 依此类推，再利用加法原理，即可得到上面的公式.

定理 1.2.13(二项式定理) 设 n 为正整数，则

$$(a+b)^n=\sum_{r=0}^{n}C(n,r)a^rb^{n-r}.$$

二项式定理的组合意义是：若将 n 个相异的球放入 a，b 两个不同的盒子，要求其中 a 盒放入 r 个，b 盒放入 $n-r$ 个，且同盒之球不分次序，则方案数为 $C(n,r)$，即 a^rb^{n-r} 项的系数为 $C(n,r)$.

我们再来考虑这样一个问题. 对 n 集$\{a_1,a_2,\cdots,a_n\}$来说，其中每个元素都有“取”与“不取”两种可能，并由此构成所有状态，根据乘法原理，其总数为 2^n，而这恰好就是 n 集的 0 组合，1 组合，…，n 组合的组合数之和. 于是，有下面定理.

定理 1.2.14(和式公式) $C(n,0)+C(n,1)+C(n,2)+\cdots+C(n,n)=2^n$.

这个公式也可以根据二项式定理，令 $a=b=1$ 而得到证明.

定理 1.2.15(范德蒙恒等式)

$$\begin{aligned}\binom{m+n}{r}&=\binom{m}{0}\binom{n}{r}+\binom{m}{1}\binom{n}{r-1}+\cdots+\binom{m}{r}\binom{n}{0}\\&=\sum_{i=0}^{r}\binom{m}{i}\binom{n}{r-i}\end{aligned}$$

其中 $r\leqslant\min\{m,n\}$.

我们再来讨论它的组合意义. 假设有 m 个相异的红球，n 个相异的蓝球，从这 $m+n$ 个球中取 r 个的组合，其结果必是下列情况之一：有 i 个红球($i=0,1,2,\cdots,n$)，$r-i$ 个蓝球. 由于对每个固定的 i 有 $\binom{m}{i}\binom{n}{r-i}$ 种选法，故根据加法原理对 i 求和即得范德蒙恒等式.

若使范德蒙恒等式中 $m\leqslant n$，$r=m$，再利用 $\binom{m}{i}=\binom{m}{m-i}$，则有如下推论.

推论　若 $m\leqslant n$，则

$$\begin{aligned}\binom{m+n}{m}&=\binom{m}{0}\binom{n}{0}+\binom{m}{1}\binom{n}{1}+\cdots+\binom{m}{m}\binom{n}{m}\\&=\sum_{i=0}^{m}\binom{m}{i}\binom{n}{i}\end{aligned}$$

1.2.2　生成函数

1. 生成函数的概念

生成函数方法是计数研究中的一种重要方法，而且应用十分广泛. 它的基本思想是把离散的数列同多项式或幂级数一一对应起来，从而把离散数列间的结合关系转化为多项式或幂级数之间的运算.

我们先试着从一个例子入手，进而引入生成函数的概念.

$$\begin{aligned}(1+a_1x)(1+a_2x)\cdots(1+a_nx)=1&+(a_1+a_2+\cdots+a_n)x+(a_1a_2+a_1a_3+\cdots+a_{n-1}a_n)x^2\\&+\cdots+a_1a_2a_3\cdots a_nx^n\end{aligned}$$

其中 x 项的系数为 $a_1+a_2+\cdots+a_n$，它包含了 n 集 $\{a_1,a_2,\cdots,a_n\}$ 的全部 1 组合. x^2 项的系数为 $a_1a_2+a_1a_3+\cdots+a_{n-1}a_n$，它包含了 n 集 $\{a_1,a_2,\cdots,a_n\}$ 的全部 2 组合. x^3 项的系数为 $a_1a_2a_3+a_1a_2a_4+\cdots+a_{n-2}a_{n-1}a_n$，它包含了 n 集 $\{a_1,a_2,\cdots,a_n\}$ 的全部 3 组合. 依此类推，可知 x^r($r\leqslant n$)项的系数包含了 n 集 $\{a_1,a_2,\cdots,a_n\}$ 的全部 r 组合.

若令 $a_1=a_2=\cdots=a_n=1$，则 x^r 项的系数等于 $C(n,r)$，可得

$$(1+x)^n=C(n,0)+C(n,1)x+C(n,2)x^2+\cdots+C(n,n)x^n.$$

我们对上式可以做进一步的分析. 考虑

$$\begin{aligned}(1+x)^{m+n}&=(1+x)^m(1+x)^n\\&=[C(m,0)+C(m,1)x+\cdots+C(m,m)x^m]\cdot[C(n,0)+C(n,1)x+\cdots+C(n,n)x^n]\\&=C(m,0)C(n,0)+[C(m,0)C(n,1)+C(m,1)C(n,0)]x+\\&\quad[C(m,0)C(n,2)+C(m,1)C(n,1)+C(m,2)C(n,0)]x^2+\cdots+\\&\quad C(m,m)C(n,n)x^{m+n}\end{aligned}$$

而将上面等式左端展开，得

$$(1+x)^{m+n}=C(m+n,0)+C(m+n,1)x+C(m+n,2)x^2+\cdots+C(m+n,m+n)x^{m+n}.$$

比较 x 的对应项系数，可得

$$C(m+n,0)=C(m,0)C(n,0),$$

$$C(m+n,1)=C(m,0)C(n,1)+C(m,1)C(n,0),$$

$$C(m+n,2)=C(m,0)C(n,2)+C(m,1)C(n,1)+C(m,2)C(n,0),$$
$$\vdots$$

一般地，有

$$C(m+n,r)=C(m,0)C(n,r)+C(m,1)C(n,r-1)+\cdots+C(m,r)C(n,0),$$

这其实就是前面所讲过的范德蒙恒等式.

由于

$$(1+x)^n\left(1+\frac{1}{x}\right)^m=x^{-m}(1+x)^{m+n},$$

即

$$[C(n,0)+C(n,1)x+\cdots+C(n,n)x^n]\cdot[C(m,0)+C(m,1)x^{-1}+\cdots+C(m,m)x^{-m}]$$
$$=x^{-m}[C(m+n,0)+C(m+n,1)x+\cdots+C(m+n,m+n)x^{m+n}]$$

比较等号两端常数项可知，若 $m\leqslant n$，则有

$$C(m+n,m)=C(n,0)C(m,0)+C(n,1)C(m,1)+\cdots+C(n,m)C(m,m),$$

这就是前面所讲过的范德蒙恒等式的推论.

对于等式

$$(1+x)^n=C(n,0)+C(n,1)x+C(n,2)x^2+\cdots+C(n,n)x^n,$$

若令 $x=1$，则会得到前面所讲过的和式公式，即

$$C(n,0)+C(n,1)+C(n,2)+\cdots+C(n,n)=2^n.$$

另外，对等式

$$(1+x)^n=C(n,0)+C(n,1)x+C(n,2)x^2+\cdots+C(n,n)x^n$$

两端求导，可得

$$n(1+x)^{n-1}=C(n,1)+2C(n,2)x+\cdots+nC(n,n)x^{n-1}.$$

再令 $x=1$，则有

$$C(n,1)+2C(n,2)+3C(n,3)+\cdots+nC(n,n)=n2^{n-1}.$$

类似地，还可以从等式

$$(1+x)^n=C(n,0)+C(n,1)x+C(n,2)x^2+\cdots+C(n,n)x^n$$

推出一系列的结论. 可见，函数 $(1+x)^n$ 在对数列

$$C(n,0),C(n,1),C(n,2),\cdots,C(n,n)$$

的研究中具有十分重要的意义.

定义 1.2.6 对于数列 $\{a_n\}$，称

$$G(x)=a_0+a_1x+a_2x^2+\cdots=\sum_{n=0}^{\infty}a_nx^n$$

为该数列的**生成函数**. 同时，称 $\{a_n\}$ 为 $G(x)$ 的**生成数列**.

显然，有限数列 $C(n,r)$ $(r=0,1,2,\cdots,n)$ 的生成函数是 $(1+x)^n$.

数列 $\{a_n\}$ 与其生成函数是一一对应的，即给定数列便可根据定义得到对应的生成函数 $G(x)$，反之，若求得了生成函数，则其生成数列也就随之确定了.

这里将生成函数只看作一个形式上的函数，目的是利用其有关运算性质完成计数问题，所以不必考虑其是否收敛，即始终认为它是收敛的，并且是逐项可微和逐项可积的.

生成函数可以用来从数列的递推公式发现数列的通项公式.

［例 1.2.12］　一个数列满足如下递推公式，

$$a_{n+1}=2a_n+1(n\geqslant 0;\ a_0=0),$$

求它的通项公式.

解　设该数列的生成函数为 $G(x)$，对上面的公式两边同时乘上 x^n，且对 $n\geqslant 0$ 求和，则

$$\sum_{n\geqslant 0}a_{n+1}x^n=\sum_{n\geqslant 0}(2a_n+1)x^n,$$

因为左边

$$\begin{aligned}\sum_{n\geqslant 0}a_{n+1}x^n&=a_1+a_2x+a_3x^2+\cdots\\&=\{(a_0+a_1x+a_2x^2+a_3x^3+\cdots)-a_0\}/x\\&=\frac{G(x)}{x},\end{aligned}$$

且右边

$$\sum_{n\geqslant 0}(2a_n+1)x^n=2G(x)+\sum_{n\geqslant 0}x^n=2G(x)+\frac{1}{1-x},$$

所以

$$\frac{G(x)}{x}=2G(x)+\frac{1}{1-x},$$

解得

$$G(x)=\frac{x}{(1-x)(1-2x)}.$$

进一步计算得到

$$\begin{aligned}\frac{x}{(1-x)(1-2x)}&=x\left(\frac{2}{1-2x}-\frac{1}{1-x}\right)\\&=\{2x+2^2x^2+2^3x^3+\cdots\}-\{x+x^2+x^3+\cdots\}\\&=(2-1)x+(2^2-1)x^2+(2^3-1)x^3+\cdots,\end{aligned}$$

很明显，

$$a_n=2^n-1(n\geqslant 0).$$

［例 1.2.13］　一个数列满足如下递推公式，

$$a_{n+1}=2a_n+n(n\geqslant 0;\ a_0=1),$$

求它的通项公式.

解　设该数列的生成函数为 $G(x)$，对上面的公式两边同时乘上 x^n，且对 $n\geqslant 0$ 求和，则

$$\sum_{n\geqslant 0}a_{n+1}x^n=\sum_{n\geqslant 0}(2a_n+n)x^n,$$

因为左边

$$\begin{aligned}\sum_{n\geqslant 0}^{\infty}a_{n+1}x^n&=a_1+a_2x+a_3x^2+\cdots\\&=\{(a_0+a_1x+a_2x^2+a_3x^3+\cdots)-a_0\}/x=\frac{G(x)-a_0}{x}=\frac{G(x)-1}{x},\end{aligned}$$

且右边

$$\sum_{n\geqslant 0}(2a_n+n)x^n=2G(x)+\sum_{n\geqslant 0}nx^n,$$

其中，

$$\sum_{n\geqslant 0} nx^n = \sum_{n\geqslant 0} x\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)x^n = x\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)\sum_{n\geqslant 0}^{\infty} x^n = x\left(\frac{\mathrm{d}}{\mathrm{d}x}\right)\frac{1}{1-x} = \frac{x}{(1-x)^2}$$

所以

$$\frac{G(x)-1}{x} = 2G(x) + \frac{x}{(1-x)^2},$$

解得

$$G(x) = \frac{1-2x+2x^2}{(1-x)^2(1-2x)}.$$

利用部分分式展开的待定系数法，可以得到

$$G(x) = \frac{-1}{(1-x)^2} + \frac{2}{(1-2x)}.$$

通过对上式展开得到，

$$a_n = 2^{n+1} - n - 1 (n\geqslant 0).$$

表 1.2.1 列出了一些常用的生成函数，它们的证明只要利用函数展开成幂级数的方法即可得到.

表 1.2.1 常见的生成函数表

$\{a_k\}$，$k=0$，1，…	$G(x)$	$\{a_k\}$，$k=0$，1，…	$G(x)$
$a_k=1$	$\frac{1}{1-x}$	$a_k=a^k$	$\frac{1}{1-ax}$
$a_k=k$	$\frac{x}{(1-x)^2}$	$a_k=k+1$	$\frac{1}{(1-x)^2}$
$a_k=k(k+1)$	$\frac{2x}{(1-x)^3}$	$a_k=k^2$	$\frac{x(1+x)}{(1-x)^3}$
$a_k=k(k+1)(k+2)$	$\frac{6x}{(1-x)^4}$	$a_k=\binom{\alpha}{k}$，α 任意	$(1+x)^\alpha$
$a_0=0$，$a_k=\frac{\alpha^k}{k}$	$-\ln(1-ax)$	$a_k=\frac{\alpha^k}{k!}$，α 任意	$\mathrm{e}^{\alpha x}$
$a_k=\frac{(-1)^k}{(2k)!}$	$\cos\sqrt{x}$	$a_k=\frac{(-1)^k}{(2k+1)!}$	$\frac{1}{\sqrt{x}}\sin\sqrt{x}$
$a_k=\frac{(-1)^k}{2k+1}$	$\frac{1}{\sqrt{x}}\arctan\sqrt{x}$	$a_k=\binom{n+k}{k}$	$(1-x)^{-(n+1)}$

[例 1.2.14] 设有 2 个红球，1 个白球，1 个黄球，试求有多少种不同的组合方案？

解 设 r，w，y 分别代表红、白、黄三种球. 两个红球的取法与 r^0，r^1，r^2 对应起来，即红球的可能取法与 $1+r+r^2$ 中的各次幂一一对应，也就是说，$r^0=1$ 表示不取，r 表示取一个红球，r^2 表示取两个. 另外两种球的选取的表示方法与此类似. 于是

$$\begin{aligned}&(1+r+r^2)(1+w)(1+y)\\ =&1+(r+w+y)+(r^2+rw+ry+wy)+(r^2w+r^2y+rwy)+r^2wy\end{aligned}$$

由此可见，共有五种情况，我们可将其枚举出来，即

(1) 一个球都不取的组合数为 1；

(2) 取一个球的组合数为 3：红、白、黄三种球各取 1 个；

(3) 取两个球的组合数为 4：2 红，1 红 1 黄，1 红 1 白，1 白 1 黄；

(4) 取三个球的组合数为 3：2 红 1 白，2 红 1 黄，1 红 1 白 1 黄；

(5) 取四个球的组合数为 1，即全取.

设取 k 个球的组合数为 c_k，则数列 c_0，c_1，c_2，c_3，c_4 的生成函数为

$$\begin{aligned}G(x)&=(1+x+x^2)(1+x)^2\\&=1+3x+4x^2+3x^3+x^4\end{aligned}$$

于是，共有 $1+3+4+3+1=12$ 种组合方案.

2. 生成函数的性质

设数列$\{a_k\}$的生成函数为 $A(x)$，数列$\{b_k\}$的生成函数为 $B(x)$，数列$\{c_k\}$的生成函数为 $C(x)$，并且 $A(x)$，$B(x)$，$C(x)$均是收敛的，并且是逐项可微和逐项可积的.

根据生成函数的定义，我们不难得出以下性质.

定理 1.2.16　(1) $A(x)=B(x)$，当且仅当 $a_k=b_k$，$k=0$，1，2，…；

(2) $A(x)+B(x)=C(x)$，当且仅当 $a_k+b_k=c_k$，$k=0$，1，2，…；

(3) $A(x)B(x)=C(x)$，当且仅当

$$\begin{aligned}&c_0=a_0b_0,\\&c_1=a_0b_1+a_1b_0,\\&c_2=a_0b_2+a_1b_1+a_2b_0,\\&\quad\vdots\\&c_k=a_0b_k+a_1b_{k-1}+a_2b_{k-2}+\cdots+a_{k-1}b_1+a_kb_0,\\&\quad\vdots\end{aligned}$$

定理 1.2.17　若 $b_k=\begin{cases}0, & k<r\\ a_{k-r}, & k\geqslant r\end{cases}$，则 $B(x)=x^rA(x)$.

证明

$$\begin{aligned}B(x)&=b_0+b_1x+b_2x^2+\cdots+b_{r-1}x^{r-1}+b_rx^r+b_{r+1}x^{r+1}+\cdots\\&=\underbrace{0+0+\cdots+0}_{r}+a_0x^r+a_1x^{r+1}+\cdots\\&=x^r(a_0+a_1x+a_2x^2+\cdots)\\&=x^rA(x)\end{aligned}$$

定理 1.2.18　若 $b_k=a_{k+r}$，则

$$B(x)=\frac{A(x)-\sum\limits_{i=0}^{r-1}a_ix^i}{x^r}.$$

证明

$$\begin{aligned}B(x)&=b_0+b_1x+b_2x^2+\cdots\\&=a_r+a_{r+1}x+a_{r+2}x^2+\cdots\\&=\frac{1}{x^r}(a_rx^r+a_{r+1}x^{r+1}+a_{r+2}x^{r+2}+\cdots)\\&=\frac{1}{x^r}[A(x)-(a_0+a_1x+\cdots+a_{r-1}x^{r-1})]\\&=\frac{A(x)-\sum\limits_{i=0}^{r-1}a_ix^i}{x^r}\end{aligned}$$

定理 1.2.19 若 $b_k=\sum_{i=0}^{k}a_i$，则

$$B(x)=\frac{A(x)}{1-x}.$$

证明 先将等式 $b_k=\sum_{i=0}^{k}a_i$ 的两边都乘以 x^k，得

$$\begin{aligned}
b_0&=a_0,\\
b_1x&=a_0x+a_1x,\\
b_2x^2&=a_0x^2+a_1x^2+a_2x^2,\\
&\vdots\\
b_kx^k&=a_0x^k+a_1x^k+a_2x^k+\cdots+a_kx^k,\\
&\vdots
\end{aligned}$$

再将以上各式左右两边分别相加可得

$$\begin{aligned}
B(x)&=a_0(1+x+x^2+\cdots)+a_1x(1+x+x^2+\cdots)+a_2x^2(1+x+x^2+\cdots)+\cdots\\
&=(1+x+x^2+\cdots)(a_0+a_1x+a_2x^2+\cdots)\\
&=\frac{A(x)}{1-x}
\end{aligned}$$

定理 1.2.20 若 $\sum_{i=0}^{\infty}a_i$ 收敛，且 $b_k=\sum_{i=k}^{\infty}a_i$，则

$$B(x)=\frac{A(1)-xA(x)}{1-x}.$$

证明 由于

$$\begin{aligned}
b_0&=a_0+a_1+a_2+\cdots=A(1),\\
b_1&=a_1+a_2+a_3+\cdots=A(1)-a_0,\\
&\vdots\\
b_k&=a_k+a_{k+1}+a_{k+2}+\cdots=A(1)-a_0-a_1-a_2-\cdots-a_{k-1},\\
&\vdots
\end{aligned}$$

将上面等式的两边都乘以 x^k 并分别相加，可得

$$\begin{aligned}
B(x)&=A(1)(1+x+x^2+\cdots)-a_0x(1+x+x^2+\cdots)-a_1x^2(1+x+x^2+\cdots)-\cdots\\
&=(1+x+x^2+\cdots)[A(1)-x(a_0+a_1x+a_2x^2+\cdots)]\\
&=\frac{A(1)-xA(x)}{1-x}
\end{aligned}$$

定理 1.2.21 若 $b_k=ka_k$，则

$$B(x)=xA'(x).$$

证明

$$\begin{aligned}
B(x)&=\sum_{k=0}^{\infty}b_kx^k=\sum_{k=0}^{\infty}ka_kx^k=x\sum_{k=1}^{\infty}ka_kx^{k-1}=x\sum_{k=1}^{\infty}(a_kx^k)'\\
&=x\Big(\sum_{k=1}^{\infty}a_kx^k\Big)'=x[A(x)-a_0]'=xA'(x)
\end{aligned}$$

定理 1.2.22　若 $b_k=\dfrac{a_k}{1+k}$，则 $B(x)=\dfrac{1}{x}\int_0^x A(x)\mathrm{d}x$.

证明

$$B(x)=\sum_{k=0}^{\infty}b_kx^k=\sum_{k=0}^{\infty}\frac{a_k}{1+k}x^k=\sum_{k=0}^{\infty}a_k\,\frac{1}{x}\int_0^x x^k\mathrm{d}x$$

$$=\frac{1}{x}\int_0^x\Big(\sum_{k=0}^{\infty}a_kx^k\Big)\mathrm{d}x=\frac{1}{x}\int_0^x A(x)\mathrm{d}x$$

3. 生成函数的一个应用——整数拆分

所谓整数的拆分，就是把正整数 n 分解成若干正整数之和，下面给出它的定义.

定义 1.2.7　将一个正整数 n 分解成 k 个正整数之和，即

$$\begin{cases}n=n_1+n_2+\cdots+n_k,\ k\geqslant 1\\ \ \ n_i\geqslant 1,\ i=1,2,\cdots,k\end{cases}$$

我们称该分解是 n 的一个 ***k* 拆分**，并称 n_i 为**分项**.

根据是否需要考虑分项 n_i 之间的顺序，可将拆分分为两类. 例如，

$$6=2+2+1+1\quad 和\quad 6=1+2+1+2$$

都是 6 的 4 拆分. 若考虑 n_i 之间的顺序，则这两个拆分被认为是不同的，我们称这样的拆分为**有序拆分**. 否则，不考虑 n_i 之间的顺序，则这两个拆分被认为是相同的，我们称这样的拆分为**无序拆分**.

先来考虑有序拆分. 不妨对每个分项 n_i 加以条件限制，例如 $1\leqslant n_i\leqslant r_i(i=1,2,\cdots,k)$，当然 $r_i\leqslant n$，于是可得如下定理.

定理 1.2.23　对于 n 的 k 有序拆分

$$\begin{cases}n=n_1+n_2+\cdots+n_k,\ k\geqslant 1\\ 1\leqslant n_i\leqslant r_i,\ i=1,2,\cdots,k\end{cases}$$

其 k 有序拆分的个数的数列$\{q_k(n)\}$的生成函数为

$$G(x)=(x+x^2+\cdots+x^{r_1})(x+x^2+\cdots+x^{r_2})\cdots(x+x^2+\cdots+x^{r_k}).$$

n 的 k 有序拆分的组合意义相当于把 n 个相同的球放入 k 个不同的盒子里，第 i 个盒子的容量为 r_i，且使每盒非空. 显然，k 有序拆分的个数等于上述分配问题的方案数. 这样一来，上面的定理就不难理解了.

今后，更多考虑的是无序拆分. 根据前面的定义，在 n 的拆分中，若不考虑各分项 n_i 之间的顺序，不妨将各分项 n_i 按从大到小的顺序排列，即

$$\begin{cases}n=n_1+n_2+\cdots+n_k,\ k\geqslant 1\\ n_1\geqslant n_2\geqslant\cdots\geqslant n_k\geqslant 1\end{cases}$$

通常，可把 k 无序拆分简称为拆分，其拆分的个数记作 $p_k(n)$，n_1 称为**最大分项**.

n 的 k 无序拆分的组合意义相当于把 n 个相同的球放入 k 个相同的盒子里，且使每盒非空. 显然，k 无序拆分的个数 $p_k(n)$等于上述分配问题的方案数.

定义 1.2.8　正整数 n 的所有(无序)拆分的个数称为 n 的拆分数，记作 $p(n)$，即

$$p(n)=\sum_{k=1}^{n}p_k(n).$$

习题

1.(1) 设 $A=\{a,\{a\}\}$，下列各式成立吗?

$\{a\}\in\rho(A)$；$\{a\}\subseteq\rho(A)$；$\{\{a\}\}\in\rho(A)$；$\{\{a\}\}\subseteq\rho(A)$.

(2) 若 $A=\{a,\{b\}\}$，(1)中的各式成立吗？

2. 对于任意三个集合 A，B，C，试证明：若 $A\times B=A\times C$，且 $A\neq\varnothing$，则 $B=C$.

3. 证明若 S 为集合 X 上的二元关系，则

(1) S 是传递的，当且仅当$(S\circ S)\subseteq S$；

(2) S 是自反的，当且仅当 $I_X\subseteq S$.

4. 设 $E=\{1,2,3,4,5\}$，$A=\{1,4\}$，$B=\{1,2,5\}$，$C=\{2,4\}$，求

(1) $A\cap\sim B$；

(2) $(A\cup B)\cap(A\cup C)$；

(3) $\sim(A\cup B)$；

(4) $\rho(A)-\rho(C)$.

5. 证明定理 1.1.5 中的(2).

6. 设 A，B，C 是任意三个集合，证明下列各式.

(1) $A\oplus(B\oplus C)=(A\oplus B)\oplus C$；

(2) $A\cap(B\oplus C)=(A\cap B)\oplus(A\cap C)$.

7. 证明集合定律中其余部分.

8. 下列关系中哪些是函数？哪些是满射？哪些是单射？对于每一个函数写出它的逆函数.

(1) f_1：$\mathbf{N}\to\mathbf{N}$，$f_1(x)=x^2+1$；

(2) f_2：$\mathbf{N}\to\mathbf{Q}$，$f_2(x)=\dfrac{1}{x}$；

(3) f_3：$\{1,2,3\}\to\{\alpha,\beta,\chi\}$，$f_3=\{<1,\alpha>,<2,\beta>,<3,\chi>\}$.

9. 试证明：

(1) $f(A\cup B)=f(A)\cup f(B)$；

(2) $f(A\cap B)\subseteq f(A)\cap f(B)$.

10. 将-2，-1，0，1，2，3 共 6 个数选作二次函数 $y=ax^2+bx+c$ 的系数 a，b，c，使得抛物线 $y=ax^2+bx+c$ 的开口方向向下，问共可作出多少这样的二次函数？

11. 从 9 颗颜色不同的珠子中选出 7 颗用细绳穿成一环，问共有多少种不同的穿法？

12. 某糖果厂生产 12 种糖果，按照客户要求装袋出售，若要求每袋装 20 块糖果，问有多少种不同的装袋方式？若再要求每袋中每种糖果至少要有 1 块，那么会有多少种不同的装袋方式？

13. 设 n 为大于 3 的整数，若把 n 集 S 划分成 $n-3$ 个非空子集，问共有多少种分法？

14. 某专业有 8 名男生，5 名女生，现要组织一个由偶数名男生和至少两名女生组成的小组，试用生成函数方法求组合方式的总数.

15. 求以下数列的生成函数：

(1) 1，2，3，4，…，n，…；

(2) 1，3，6，10，…，$C(n+2,2)$，….

16. 试求生成函数 $G(x)=\dfrac{3+78x}{1-3x-54x^2}$的生成数列$\{a_n\}$.

17. 试求 5，6 和 7 的拆分数.

18. 求 a，b，c，d，e，f 六个字母的全排列中不允许出现 ace 和 bf 的情形的排列数.

19. 计算字符串“ABCABCACC”的重合指数.

20. 计算字符串“AAABBCCCC”和“ABCABCACCD”的交互重合指数(即分别随机从两个字符串中各取出一个字符时，二者相同的概率；破译维吉尼亚密码中也需要计算交互重合指数).

21. 英文的换位密码中，将明文每 5 个字母分成一组，一次特定的加密时，对每个组进行如下操作：组中的第 1 个字母放到密文相应分组的第 5 个位置，第 2 个字母放到第 1 个位置，第 3 个字母放到第 2 个位置，第 4 个字母放到第 3 个位置，这样一个具体的换位就对应一个密钥. 那么一共有多少个密钥？如果分组长度为 6 呢？

22. Fibonacci 数列的递推公式为 $F_{n+1}=F_n+F_{n-1}(n\geqslant 1;\ F_0=0;\ F_1=1)$，求它的通项公式.

23. 一个函数满足 $f(1)=1;\ f(2n)=f(n);\ f(2n+1)=f(n)+f(n+1)$，令 $F(x)=\sum\limits_{n\geqslant 0} f(n)x^{n-1}$ 为序列的生成函数，试证明 $F(x)=(1+x+x^2)F(x^2)$.

第 2 章　数论基础(一)

2.1　整除

在整数集合中，整除是一种重要的二元关系，相关的概念和性质包括素数、公因数和公倍数、辗转相除、算术基本定理等，这些概念和性质又是整数集合中另一种重要的二元关系——同余关系的基础。本节将对整数整除的相关概念和性质作详细的介绍。

2.1.1　整除与带余除法

数论主要是指关于整数性质的理论，在数学世界中占有独特的地位，它的许多问题概念上很容易理解，但是解决起来却非常困难. 伟大的数学家高斯曾经说过“数学是科学的女王，数论是数学的女王”，这代表了许多年以来人们将数论看作纯粹的理论数学而不是应用数学的普遍观点. 然而，正是这个数学领域却在当今的互联网络时代中发挥了巨大作用，它与信息安全尤其是保密通信领域紧密地结合在一起，凸现了它在实际问题中的可应用性.

我们通常用 $\mathbf{N}$ 表示正整数(自然数)集合，即 $\mathbf{N}=\{1,2,3,\cdots\}$，用 $\mathbf{Z}$ 表示整数集合，即 $\mathbf{Z}=\{\cdots,-3,-2,-1,0,1,2,3,\cdots\}$. 我们知道，两个整数的和、差、积仍然是整数，但如果用一个非零整数去除另一个整数，所得的商还会是整数吗？下面就开始研究这个问题.

定义 2.1.1　设 $a,b\in\mathbf{Z}$，$b\neq0$. 如果存在 $q\in\mathbf{Z}$ 使得 $a=qb$，那么就称 a 可被 b **整除**或者 b 整除 a，记为 $b|a$，且称 a 是 b 的**倍数**，b 是 a 的**因子**(也可称为**约数**或**除数**). 若 a 不能被 b 整除，则记为 $b\nmid a$.

需要注意的是，符号 $b|a$ 本身就包含了条件 $b\neq0$，不过 $a=0$ 是允许的. 同时，需要记住这个定义的关键在于整除关系是通过整数的乘法定义的，而不是通过除法定义的.

定义 2.1.2　若 b 为 a 的因子，且 $b\neq\pm1$，$b\neq\pm a$，则称 b 为 a 的**真因子**.

例如 2 和 7 是 14 的真因子，而 1 和 14 虽然是 14 的因子，但不是真因子.

定理 2.1.1　设 $a,b\in\mathbf{Z}$，则有

(1) $b|a\Leftrightarrow-b|a\Leftrightarrow b|-a\Leftrightarrow|b|||a|$；

(2) 设 $a\neq0$，如果 $b|a$，那么 $|b|\leqslant|a|$.

证明　(1) 可由以下各式两两等价推出：$a=qb$，$a=(-q)(-b)$，$-a=(-q)b$，$|a|=|q||b|$，其中 $q\in\mathbf{Z}$.

(2) 由(1)知 $|a|=|q||b|$. 当 $a\neq0$ 时，$|q|\geqslant1$. 定理得证.

定理 2.1.2　设 $a,b,c\in\mathbf{Z}$，

(1) 若 $b|a$ 且 $c|b$，则 $c|a$；

(2) 若 $b|a$，则 $b|ac$；

(3) 设 $c\neq0$，则 $b|a\Leftrightarrow bc|ac$；

(4) $b|a$ 且 $b|c \Leftrightarrow$ 对任意的 m，$n \in \mathbf{Z}$ 有 $b|am+cn$.

证明 (1) 因为 $b|a$ 且 $c|b$，则存在整数 q_1 和 q_2 使得 $a=q_1b$，$b=q_2c$，从而推出 $a=(q_1\ q_2)c$.

(2) 因为 $b|a$，则存在整数 q 使得 $a=qb$，从而 $ac=(qc)b$.

(3) 由于 $c\neq 0$，故 $a=qb$ 与 $ac=q(bc)$ 等价，其中 q 为整数.

(4) 因为 $b|a$ 且 $b|c$，则存在整数 q_1 和 q_2，使得 $a=q_1b$，$c=q_2b$，从而

$$am+cn=q_1bm+q_2bn=(q_1m+q_2n)b.$$

必要性得证. 取 $m=1$，$n=0$ 及 $m=0$，$n=1$ 就可推出充分性. 我们可以将 $am+cn(m,\ n\in\mathbf{Z})$ 形式的整数称为整数 a 和 c 的线性组合. 尽管，在读者看来这一条也像前面几条那样很直观，然而，下面的一些证明中会反复应用整除的这个性质，即如果一个整数整除另外两个整数，那么必然整除它们的任意线性组合.

[例 2.1.1] 证明：若 $b|a$ 且 $a|b$，则 $b=\pm a$.

证明 因为 $b|a$ 且 $a|b$，则存在整数 q_1 和 q_2 使得

$$a=q_1b,\quad b=q_2a,$$

可得 $a=q_1q_2a$. 由于 $a\neq 0$，所以 $q_1q_2=1$，$q_2=\pm 1$. 从而 $b=\pm a$.

[例 2.1.2] 设 $a=2t-1$. 若 $a|2n$，则 $a|n$.

证明 由 $a|2n$ 知 $a|2tn$，又 $a|an$，根据定理 2.1.2(4)，则 $a|2tn-an$. 由 $a=2t-1$ 知 $2tn-an=2tn-2tn+n=n$. 代入即得 $a|n$.

[例 2.1.3] 设 a，b 是两个给定的非零整数，且有整数 x，y，使得 $ax+by=1$. 证明：若 $a|n$ 且 $b|n$，则 $ab|n$.

证明 由

$$n=n(ax+by)=(na)x+(nb)y,$$

又 $ab|na$，$ab|nb$，根据定理 2.1.2(4)即得.

上面几个定理的证明中只需要利用整数的加法、减法和乘法的性质，这三个整数运算的结果都没有超出整数的范围，一旦考虑整数的除法，就需要用到整数的另一个性质，即所谓的良序原理.

良序原理 每一个由非负整数组成的非空集合 S 必定含有一个最小元素，也就是说，S 中存在一个元素 a，对任意 $b\in S$，都有 $a<b$ 成立.

很显然，良序原理符合我们对整数的直观感受和日常使用要求，但是这个原理是无法证明的，只能够作为公理写出，它是证明下面一些定理的关键基础和依据.

定理 2.1.3 设 a 和 b 为任意整数，$b>0$，则存在唯一的一对整数 q 和 r，使

$$a=qb+r,\quad 0\leqslant r<b. \tag{2.1.1}$$

其中 a 称为**被除数**，q 称为**商**，r 称为**余数**(或**非负最小剩余**).

证明 先证明存在性. 令集合

$$S=\{a-xb\,|\,x\in\mathbf{Z}\text{ 且 }0\leqslant a-xb\ \}$$

这个集合是非负整数的集合，且不是空集：因为 $b\geqslant 1$，所以 $|a|b\geqslant|a|$，那么 $a-(-|a|)b=a+|a|\ b\geqslant a+|a|\geqslant 0$. 即当 $x=-|a|$ 时，$a-xb\in S$. 根据良序原理，S 含有一个最小的元素 r，由集合 S 的定义可知，存在整数 q 满足

$$r=a-qb,\ r\geqslant 0.$$

我们也能够证明 $r<b$：因为如果 $r<b$ 不成立，则 $r\geqslant b$，那么

$$a-(q+1)b=(a-qb)-b=r-b\geqslant 0.$$

所以 $a-(q+1)b\in S$. 但是 $a-(q+1)b=(a-qb)-b=r-b<r$，这与 r 是 S 的最小元素相矛盾，于是 $r<b$，存在性得证.

下面来证明唯一性. 假设存在另一对整数 q_1 和 r_1 满足式(2.1.1)，即

$$a=q_1b+r_1,\ 0\leqslant r_1<b. \tag{2.1.2}$$

设 $r<r_1$，则 $0<r_1-r<b$. 将式(2.1.1)和式(2.1.2)相减，得$(q-q_1)b=r_1-r$，故 $b\mid(r_1-r)$. 这个结果与 $0<r_1-r<b$ 矛盾(定理 2.1.1(2)). 同理如果设 $r_1<r$，也会导出矛盾. 所以 $r=r_1$，进而 $q=q_1$. 证毕.

这个定理也被称作带余除法. 如果 $b<0$：由于 $-b>0$，这个定理意味着存在整数 q_1 与 r 使得 $a=q_1(-b)+r$，$0\leqslant r<-b$. 此时，令 $q=-q_1$，就能得到这个定理的推广，只要 $b\neq 0$，就存在唯一的一对整数 q 和 r，使 $a=qb+r$，　$0\leqslant r<|b|$.

显然，$b\mid a$ 的充要条件是 $r=0$. 注意，当 a 是负数时的余数可以不同(但两个余数之和为 $|b|$). 比较一个简单的例子中的两个余数的不同，如果用 7 去除 60 和 -60，得到 $60=7\times 8+4$ 和 $-60=7\times(-9)+3$.

定义 2.1.3　设 a，q，$r\in\mathbf{Z}$，满足 $a=2q+r$，$0\leqslant r<2$. 若 $r=0$，称 a 为**偶数**；若 $r=1$，称 a 为**奇数**.

定义 2.1.4　一个大于 1 的整数 p，若仅以 1 和自身 p 为其正因子，则称 p 为**素数**(或**质数**). 除 1 以外非素的正整数则称为**合数**(或**复合数**).

素数具有许多特殊而又美妙的性质，并且在数论科学的发展中起着十分重要的作用. 历史上的许多数学家都不禁为之倾倒. 下面介绍几个关于素数的基本定理.

定理 2.1.4　素数有无穷多个.

证明　用反证法. 假定只有有限个素数 p_1，p_2，…，p_k，考虑 $a=p_1p_2\cdots p_k+1$. 由于 a 是合数，所以它必有素因子，不妨假定这个素因子为 $p_j(1\leqslant j\leqslant k)$，显然 $p_j\mid a$. 因为

$$a-p_1p_2\cdots p_k=1,$$

又 $p_j\mid(p_1p_2\cdots p_k)$，故 $p_j\mid 1$. 但是素数 $p_j\geqslant 2$，所以 $p_j\mid 1$ 是不可能的. 因此推出矛盾，假设错误. 定理得证.

定理 2.1.5　对任意正整数 n，存在素数 p 满足 $n<p\leqslant n!+1$.

证明　考虑正整数 $a=n!+1$. 如果 a 是素数，则可取 $p=a$. 如果 a 是合数，则必有某个素因子 p. 先假定 $p\leqslant n$，那么必有 $p\mid n!$，所以 $p\mid(a-n!)$，即 $p\mid 1$，出现了矛盾，因此 $p>n$. 定理得证.

定理 2.1.6　如果整数 $n\geqslant 2$，那么在 $n!+2$ 与 $n!+n$ 之间必没有素数.

证明　由于 $n!$ 是从 1 到 n 的所有整数的连乘积，所以有

$$2\mid(n!+2),\ 3\mid(n!+3),\ \cdots,\ n\mid(n!+n).$$

证毕.

定理 2.1.7　若 n 为合数，则 n 必有素因子 p 满足 $p\leqslant\sqrt{n}$.

证明　不妨设 p 为 n 的最小素因子. 如果有 $n=rs$，其中 r 和 s 均为 n 的真因子，那么 p

$\leqslant r$ 且 $p\leqslant s$. 所以 $p^2\leqslant rs=n$，即 $p\leqslant\sqrt{n}$.

定理 2.1.7 给出了一种寻找素数的有效方法. 为了求出不超过给定正整数 $x(>1)$的所有素数，只要把从 2 到 x 的所有合数都删去即可. 因为不超过 x 的合数 a 必有一个素因子 $p\leqslant\sqrt{a}\leqslant\sqrt{x}$，所以只要先求出$\sqrt{x}$以内的全部素数$\{p_i,\ 1\leqslant i\leqslant k\}$(其中 k 为$\sqrt{x}$以内的素数个数)，然后把不超过 x 的 p_i 的倍数(p_i 本身除外)全部删去，剩下的就正好是不超过 x 的全部素数. 这种寻找素数的方法称为 Eratosthenes 筛法. 下面是一个具体应用的实例.

[例 2.1.4] 求出不超过 64 的所有素数.

解 先求出不超过$\sqrt{64}=8$的全部素数，依次为 2，3，5，7. 然后从 2 到 64 的所有整数中依次删去除了 2，3，5，7 以外的 2 的倍数，3 的倍数，5 的倍数和 7 的倍数，剩下的即为所求. 具体过程如下所示.

	2	3	~~4~~	5	~~6~~	7	~~8~~	~~9~~	~~10~~	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

可以看出，没有删去的数是

2，3，5，7，11，13，17，19，23，29，31，37，41，43，47，53，59，61.

它们就是不超过 64 的所有素数.

2.1.2 最大公因子与辗转相除法

定义 2.1.5 设 a_1，a_2，…，a_n 是 n 个不全为零的整数. 若整数 d 是它们之中每一个数的因子，那么 d 就称为 a_1，a_2，…，a_n 的一个**公因子**. 在整数 a_1，a_2，…，a_n 的所有公因子中最大的一个称为**最大公因子**，记作$(a_1,a_2,\cdots,a_n)$或者 $\gcd(a_1,a_2,\cdots,a_n)$. 特别地，若$(a_1,a_2,\cdots,a_n)=1$，我们称 a_1，a_2，…，a_n **互素**(或**互质**).

例如，12 和 −18 的公因子为$\{\pm1,\pm2,\pm3,\pm6\}$，它们的最大公因子$(12,-18)=6$. 而$(12,-18,35)=1$，于是我们说 12，−18 和 35 这三个整数是互素的. 需要注意的是，符号$(a_1,a_2,\cdots,a_n)$本身就包含了条件 a_1，a_2，…，a_n 不全为零.

定理 2.1.8 设 a，b，c 是任意三个不全为零的整数，且 $a=bq+c$，其中 q 是整数，则$(a,b)=(b,c)$.

证明 因为$(a,b)\mid a$，$(a,b)\mid b$，又 $c=a-bq$，所以$(a,b)\mid c$，即(a,b)是 b 和 c 的公因子，因而$(a,b)\leqslant(b,c)$. 同理可证$(b,c)\leqslant(a,b)$，于是$(a,b)=(b,c)$. 结合带余除法，可以得到这样一句话：被除数与除数的最大公因子等于除数与余数的最大公因子.

由最大公因子的定义，不难得到$(a_1,a_2,\cdots,a_n)=(|a_1|,|a_2|,\cdots,|a_n|)$. 另外，欲求一组不全为零的整数的最大公因子，只要求出其中全体非零整数的最大公因子即可，因为它们是相等的. 于是，先讨论两个正整数的最大公因子的求法. 当然，可以运用最大公因子的定义，先分别求出这两个数的所有因子，再从中挑出它们的最大公因子. 在这两个数比较小的情况下，这种方法是可行的，但若这两个数相对较大，那么分解其因子是十分困难的，只能另想办法. 下面介绍一种**“辗转相除法”**，它可以很好地解决求两个正整数的最大公因子的

问题，而且就目前来讲，这也是能在计算机上实现的解决此问题最好的算法. 这个算法也被称作**“欧几里得除法”**.

任给两个正整数 a 和 b，不妨设 $a\geqslant b$，由定理 2.1.3，有下列等式：

$$\begin{aligned}
a&=bq_1+r_1,\ 0<r_1<b,\\
b&=r_1q_2+r_2,\ 0<r_2<r_1,\\
&\vdots\\
r_{n-2}&=r_{n-1}q_n+r_n,\ 0<r_n<r_{n-1},\\
r_{n-1}&=r_nq_{n+1}+r_{n+1},\quad r_{n+1}=0.
\end{aligned}\tag{2.1.3}$$

因为 $b>r_1>r_2>\cdots>r_n>r_{n+1}=0$，所以经过有限步后，总可以得到一个余数是零，即式(2.1.3)中 $r_{n+1}=0$.

定理 2.1.9　若任给两个正整数 a 和 b，则 (a,b)就是式(2.1.3)中最后一个不等于零的余数，即 $(a,b)=r_n$.

证明　由定理 2.1.8 可知

$$r_n=(0,r_n)=(r_n,r_{n-1})=\cdots=(r_2,r_1)=(r_1,b)=(a,b).$$

定理得证.

定理 2.1.10　若任给两个正整数 a 和 b，则存在两个整数 m，n，使得

$$(a,b)=ma+nb.$$

即 (a,b)是 a 和 b 的线性组合.

证明　由式(2.1.3)可知

$$r_n=r_{n-2}-r_{n-1}q_n,$$

即 r_n 是 r_{n-2}和 r_{n-1}的线性组合，将 $r_{n-1}=r_{n-3}-r_{n-2}q_{n-1}$代入得 $r_n=r_{n-2}(1+q_nq_{n-1})-r_{n-3}q_n$，即 r_n 是 r_{n-3}和 r_{n-2}的线性组合，再将 $r_{n-2}=r_{n-4}-r_{n-3}q_{n-2}$代入，那么 r_n 也是 r_{n-4}和 r_{n-3}的线性组合，如此继续下去，直到将式(2.1.3)的最开始的两个式子代入完毕，最终可得 r_n 也是 a 和 b 的线性组合，即存在两个整数 m，n，使得 $r_n=ma+nb$. 又根据定理 2.1.9 知 $(a,b)=r_n$，定理得证.

显然，定理 2.1.10 中 a 和 b 的取值可推广到全部整数范围.

[例 2.1.5]　设 $a=8\ 656$，$b=-7\ 780$，求 (a,b)和整数 m，n，使 $ma+nb=(a,b)$.

解　$(a,b)=(8\ 656,-7\ 780)=(8\ 656,7\ 780)$. 运用辗转相除法，有

$$\begin{aligned}
8\ 656&=7\ 780\times1+876,\\
7\ 780&=876\times8+772,\\
876&=772\times1+104,\\
772&=104\times7+44,\\
104&=44\times2+16,\\
44&=16\times2+12,\\
16&=12\times1+4,\\
12&=4\times3+0.
\end{aligned}$$

因此，$(a,b)=4$. 再由

$$\begin{aligned}
4&=16-12\times1 && \text{初始步骤}\\
&=16-(44-16\times2) && \text{回代步骤}
\end{aligned}$$

$=16\times3-44$ 整理步骤

$=(104-44\times2)\times3-44$ 回代步骤

$=104\times3-44\times7$ 整理步骤

$=104\times3-(772-104\times7)\times7$ 回代步骤

$=104\times52-772\times7$ 整理步骤

$=(876-772\times1)\times52-772\times7$ 回代步骤

$=876\times52-772\times59$ 整理步骤

$=876\times52-(7\ 780-876\times8)\times59$ 回代步骤

$=876\times524-7\ 780\times59$ 整理步骤

$=(8\ 656-7\ 780\times1)\times524-7\ 780\times59$ 回代步骤

$=8\ 656\times524-7\ 780\times583$ 整理步骤

$=8\ 656\times524+(-7\ 780)\times583$ 规范步骤

因此,整数 $m=524$, $n=583$, 使 $ma+nb=(a,b)$.

从上面这个例子中,我们看到在求整数 m, n 的时候,初始步骤是将 $r_{n-2}=r_{n-1}q_n+r_n$ 写成 $r_n=r_{n-2}-r_{n-1}q_n$ 的形式,然后需要交替使用回代步骤和整理步骤直到 a 和 b 出现在等式的右边,最后的规范步骤是为了得到具有正确的正负号的 m, n,需要注意的是该步骤的关键在于,中间的符号必须成为'+'.

定理 2.1.11 设整数 a, b, c 满足 $c|a$ 且 $c|b$, 则 $c|(a,b)$.

证明 由定理 2.1.10 可知,存在两个整数 m, n, 使得

$$(a,b)=ma+nb.$$

因为 $c|a$ 且 $c|b$, 故 $c|ma+nb$, 即 $c|(a,b)$(即公因子整除最大公因子).

定理 2.1.12 设有整数 a, b, c, 其中 $c>0$, 则 $(ac,bc)=(a,b)c$.

证明 由定理 2.1.10 可知,存在两个整数 m, n 使得

$$(a,b)=ma+nb.$$

将等式左右两端同乘 c, 得

$$(a,b)c=m(ac)+n(bc).$$

因为 $(ac,bc)|m(ac)+n(bc)$, 所以 $(ac,bc)|(a,b)c$.

又显然有 $(a,b)c|\ ac$, $(a,b)c|\ bc$, 由定理 2.1.11 可知 $(a,b)c|(ac,bc)$.

因此, $(ac,bc)=(a,b)c$, 定理得证.

[例 2.1.6] 设 $a=16\times2\ 350$, $b=27\times2\ 350$, 求 (a,b).

解 $(a,b)=(16,27)\times2\ 350=1\times2\ 350=2\ 350$.

[例 2.1.7] 证明:若整数 a, b, d 满足 $d|a$, $d|b$, 则

$$\left(\frac{a}{d},\frac{b}{d}\right)=\frac{(a,b)}{|d|}.$$

特别地, $\left(\dfrac{a}{(a,b)},\dfrac{b}{(a,b)}\right)=1$.

证明 因为 $d|a$, $d|b$, 则由定理 2.1.12 有

$$(a,b)=\left(\frac{a}{|d|}\cdot|d|,\frac{b}{|d|}\cdot|d|\right)$$

$$=\left(\frac{a}{|d|},\frac{b}{|d|}\right)|d|$$

$$=\left(\frac{a}{d},\frac{b}{d}\right)|d|,$$

所以$\left(\frac{a}{d},\frac{b}{d}\right)=\frac{(a,b)}{|d|}$. 特别地，当$d=(a,b)$时，有

$$\left(\frac{a}{(a,b)},\frac{b}{(a,b)}\right)=1.$$

定理 2.1.13　整数a，b互素的充分必要条件是存在整数x，y，使得

$$xa+yb=1.$$

证明　由互素的定义及定理 2.1.10，必要性显然得证.

再证充分性. 不妨设$d=(a,b)$，则$d|a$且$d|b$. 若存在整数x，y使得

$$xa+yb=1,$$

则有$d|xa+yb$，即$d|1$，所以$d=1$，a，b互素. 证毕.

定理 2.1.14　设有整数a，b，c，若$a|bc$且$(a,b)=1$，则$a|c$.

证明　若$c=0$，结论显然成立. 下面不妨假定$c\neq 0$.

因为$(a,b)=1$，由定理 2.1.13 可知，存在整数m，n使得

$$ma+nb=1.$$

将等式左右两端同乘c，得

$$mac+nbc=c.$$

因为$a|ac$，$a|bc$，所以$a|mac+nbc$，即$a|c$. 定理得证.

以上讨论了两个整数的最大公因子的求解问题，那么对于两个以上的整数，如何才能求出其最大公因子呢?

定理 2.1.15　设$a_1,a_2,\cdots,a_n$是n个整数，其中$a_1\neq 0$. 令

$$(a_1,a_2)=d_2,(d_2,a_3)=d_3,\cdots,(d_{n-1},a_n)=d_n,$$

则

$$(a_1,a_2,\cdots,a_n)=d_n.$$

证明　由$d_n|a_n$，$d_n|d_{n-1}$，$d_{n-1}|a_{n-1}$，$d_{n-1}|d_{n-2}$，可知$d_n|a_{n-1}$，$d_n|d_{n-2}$. 依此类推，可得

$$d_n|a_n,d_n|a_{n-1},\cdots,d_n|a_1,$$

故d_n是a_1，a_2，…，a_n的公因子.

不妨设d为a_1，a_2，…，a_n的任意公因子. 因为$d|a_1$，$d|a_2$，则由定理 2.1.11 有$d|d_2$，又$d|a_3$，则$d|d_3$. 依此类推，可得$d|d_n$. 故$d\leqslant d_n$.

根据最大公因子的定义，可知

$$(a_1,a_2,\cdots,a_n)=d_n.$$

证毕.

[例 2.1.8]　计算(90，30，114，42，81).

解　因为

$$(90,30)=30,$$

$$(30,114)=6,$$

$$(6,42)=6,$$
$$(6,81)=3,$$

所以$(90,30,114,42,81)=3$.

定义 2.1.6 设$a_1,a_2,\cdots,a_n$是n个整数，若m是这n个数中每一个数的倍数，则m就称为这n个数的一个**公倍数**. 在$a_1,a_2,\cdots,a_n$的所有公倍数中最小的正整数称为**最小公倍数**，记作$[a_1,a_2,\cdots,a_n]$或者$\mathrm{lcm}(a_1,a_2,\cdots,a_n)$.

例如，12和-18的公倍数为$\{\pm 36,\pm 72,\cdots\}$，它们的最小公倍数$[12,-18]=36$. 由于任何正整数都不是零的倍数，故讨论整数的最小公倍数时，总假定这些整数都不为零. 类似于最大公因子，有$[a_1,a_2,\cdots,a_n]=[|a_1|,|a_2|,\cdots,|a_n|]$. 于是，先讨论两个正整数的最小公倍数的求法.

定理 2.1.16 设a和b为任意两个互素正整数，则其乘积即为最小公倍数.

证明 设m是a,b的任一公倍数，即$a|m,b|m$. 则有$m=ak$，即$b|ak$，其中k为正整数. 又$(a,b)=1$，则$b|k$. 于是存在正整数t使$k=bt$，$m=abt$，即$ab|m$，故$ab\leqslant m$. 由于ab显然是a,b的公倍数，且不大于a,b的任一公倍数m，所以它就是最小公倍数. 定理得证.

定理 2.1.17 设a和b为任意正整数，则

(1) 若m是a,b的任一公倍数，则$[a,b]|m$；

(2) $[a,b]=\dfrac{ab}{(a,b)}$.

证明 设正整数x,y满足$m=ax=by$，令$a=a_1(a,b)$，$b=b_1(a,b)$，则$a_1x=b_1y$. 因为$(a_1,b_1)=1$，所以$b_1|x$，即存在正整数t使$x=b_1t$. 于是，有

$$m=ax=ab_1t=\frac{ab}{(a,b)}t.$$

根据定理2.1.16，有$[a_1,b_1]=a_1b_1$，即

$$\left[\frac{a}{(a,b)},\frac{b}{(a,b)}\right]=\frac{ab}{(a,b)^2}.$$

将等式两端同乘(a,b)，得

$$[a,b]=\frac{ab}{(a,b)},$$

于是(2)得证. 所以有$m=[a,b]t$，即$[a,b]|m$，(1)也得证.

现在开始讨论两个以上整数的最小公倍数，给出下面的定理.

定理 2.1.18 设$a_1,a_2,\cdots,a_n$是n个整数，令

$$[a_1,a_2]=m_2,[m_2,a_3]=m_3,\cdots,[m_{n-1},a_n]=m_n,$$

则

$$[a_1,a_2,\cdots,a_n]=m_n.$$

证明 因为$m_i|m_{i+1}(i=2,3,\cdots,n-1)$，且$a_1|m_2$，$a_i|m_i(i=2,3,\cdots,n)$，所以$m_n$是$a_1,a_2,\cdots,a_n$的公倍数. 又设$m$是$a_1,a_2,\cdots,a_n$的任一公倍数，则由$a_1|m$，$a_2|m$，可知$m_2|m$，又$a_3|m$，可得$m_3|m$. 依此类推，最后得$m_n|m$，因此$m_n\leqslant|m|$. 所以$m_n=[a_1,a_2,\cdots,a_n]$.

[例 2.1.9] 计算$[90,30,114,42,81]$.

解 因为

$$[90,30]=\frac{90\times30}{(90,30)}=\frac{90\times30}{30}=90,$$

$$[90,114]=\frac{90\times114}{(90,114)}=\frac{90\times114}{6}=1\ 710,$$

$$[1\ 710,42]=\frac{1\ 710\times42}{(1\ 710,42)}=\frac{1\ 710\times42}{6}=11\ 970,$$

$$[11\ 970,81]=\frac{11\ 970\times81}{(11\ 970,81)}=\frac{11\ 970\times81}{9}=107\ 730,$$

所以$[90,30,114,42,81]=107\ 730$.

定理 2.1.19　设 $a_1,a_2,\cdots,a_n$ 是 n 个正整数，如果 $a_1\mid m,a_2\mid m,\cdots,a_n\mid m$，则

$$[a_1,a_2,\cdots,a_n]\mid m.$$

证明　对 n 用数学归纳法.

当 $n=2$ 时，由定理 2.1.17 可知.

假设 $n=k$ $(2<k<n)$时，命题成立，即 $m_k\mid m$，其中 $m_k=[a_1,a_2,\cdots,a_k]$.

当 $n=k+1$ 时，由

$$[m_k,a_{k+1}]=[a_1,a_2,\cdots,a_k,a_{k+1}],$$

可得$[a_1,a_2,\cdots,a_{k+1}]\mid m$. 于是定理得证.

2.1.3　连分数

这一节的内容与 2.1.2 节的辗转相除法有密切关系，我们可以利用辗转除法来求有理数的连分数表示形式. 下面给出连分数的定义.

定义 2.1.7　设 $a_0,a_1,a_2,\cdots,a_n$是一个实数列，除 a_0 以外都大于 0. 对于整数 $n\geqslant0$，将分数

$$a_0+\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+}}}\\ \vdots\\ +\frac{1}{a_n}\tag{2.1.4}$$

叫作 n 阶**有限连分数**. 当 a_0 是整数，$a_1,a_2,\cdots,a_n$ 都是正整数时，分数(2.1.4)叫作 n 阶**有限简单连分数**. 为了书写方便，将式(2.1.4)简记为

$$[a_0,a_1,\cdots,a_n].\tag{2.1.5}$$

将有限连分数

$$[a_0,a_1,\cdots,a_k],\ 0\leqslant k\leqslant n\tag{2.1.6}$$

叫作有限连分数式(2.1.4)的第 k 个**渐进分数**.上述简记方式与最小公倍数表示方式相同，阅读时，注意不同环境中的不同意义.

当式(2.1.4)中的 $n\to\infty$时，则分数

$$a_0+\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+}}}\\ \vdots\tag{2.1.7}$$

叫作**无限连分数**，可简记为

$$[a_0, a_1, a_2, \cdots]. \tag{2.1.8}$$

当 a_0 是整数，$a_1, \cdots, a_n$ 都是正整数时，分数(2.1.7)叫作**无限简单连分数**. 将有限连分数

$$[a_0, a_1, \cdots, a_k], k \geqslant 0 \tag{2.1.9}$$

叫作无限连分数(2.1.7)的第 k 个**渐进分数**.

对于无限连分数，有时也将其表示为

$$[a_0, a_1, \cdots, a_n],$$

但这里 $n \to \infty$.

定理 2.1.20 若使连分数$[a_0, a_1, \cdots, a_n]$的渐进分数分别为

$$[a_0, a_1, \cdots, a_i] = \frac{p_i}{q_i}, \quad 0 \leqslant i \leqslant n,$$

则这些渐进分数间有关系

$$\begin{aligned} &p_0 = a_0, \quad p_1 = a_1 a_0 + 1, \quad \cdots, \quad p_k = a_k p_{k-1} + p_{k-2}, \\ &q_0 = 1, \quad q_1 = a_1, \quad \cdots, \quad q_k = a_k q_{k-1} + q_{k-2}, \end{aligned}$$

其中 $2 \leqslant k \leqslant n$.

证明 用数学归纳法.

因为

$$\frac{p_0}{q_0} = a_0, \quad \frac{p_1}{q_1} = a_0 + \frac{1}{a_1} = \frac{a_1 a_0 + 1}{a_1},$$

$$\frac{p_2}{q_2} = a_0 + \frac{1}{a_1 + \dfrac{1}{a_2}} = \frac{a_2 a_1 a_0 + a_2 + a_0}{a_2 a_1 + 1},$$

所以当 $k = 0, 1, 2$ 时可直接验证.

假设当 $k = m$ $(2 \leqslant m < n)$时，命题成立，即

$$[a_0, a_1, \cdots, a_m] = \frac{p_m}{q_m} = \frac{a_m p_{m-1} + p_{m-2}}{a_m q_{m-1} + q_{m-2}},$$

则当 $k = m+1$ 时，有

$$\begin{aligned} [a_0, a_1, \cdots, a_m, a_{m+1}] &= [a_0, a_1, \cdots, a_m + \frac{1}{a_{m+1}}] \\ &= \frac{(a_m + \dfrac{1}{a_{m+1}}) p_{m-1} + p_{m-2}}{(a_m + \dfrac{1}{a_{m+1}}) q_{m-1} + q_{m-2}} \\ &= \frac{a_{m+1}(a_m p_{m-1} + p_{m-2}) + p_{m-1}}{a_{m+1}(a_m q_{m-1} + q_{m-2}) + q_{m-1}} \\ &= \frac{a_{m+1} p_m + p_{m-1}}{a_{m+1} q_m + q_{m-1}} \\ &= \frac{p_{m+1}}{q_{m+1}} \end{aligned}$$

于是定理得证.

定理 2.1.21 若连分数$[a_0, a_1, \cdots, a_n]$的渐进分数分别为

$$[a_0, a_1, \cdots, a_k]=\frac{p_k}{q_k},\ 0\leqslant k\leqslant n,$$

则 p_k 和 q_k 满足

$$p_kq_{k-1}-p_{k-1}q_k=(-1)^{k-1},\quad 1\leqslant k\leqslant n, \tag{2.1.10}$$

$$p_kq_{k-2}-p_{k-2}q_k=(-1)^ka_k,\quad 2\leqslant k\leqslant n. \tag{2.1.11}$$

证明　用数学归纳法.

当 $k=1$ 时，式(2.1.10)显然成立.

假设当 $k=m-1$ $(1<m\leqslant n)$时，命题成立，即

$$p_{m-1}q_{m-2}-p_{m-2}q_{m-1}=(-1)^{m-2}=(-1)^m,$$

则当 $k=m$ 时，由定理 2.1.20，有

$$\begin{aligned}p_mq_{m-1}-p_{m-1}q_m&=(a_mp_{m-1}+p_{m-2})q_{m-1}-p_{m-1}(a_mq_{m-1}+q_{m-2})\\&=-(p_{m-1}q_{m-2}-p_{m-2}q_{m-1})\\&=(-1)^{m-1}\end{aligned}$$

于是式(2.1.10)成立.

由式(2.1.10)和定理 2.1.20 可得

$$\begin{aligned}p_kq_{k-2}-p_{k-2}q_k&=(a_kp_{k-1}+p_{k-2})q_{k-2}-p_{k-2}(a_kq_{k-1}+q_{k-2})\\&=a_k(p_{k-1}q_{k-2}-p_{k-2}q_{k-1})\\&=(-1)^ka_k\end{aligned}$$

于是定理得证.

定理 2.1.22　对于简单连分数，有

(1) 当 $k\geqslant 2$ 时，$q_k\geqslant q_{k-1}+1$，因而对任何 k 来说，$q_k\geqslant k$；

(2) $\dfrac{p_{2k+1}}{q_{2k+1}}<\dfrac{p_{2k-1}}{q_{2k-1}}$，　$\dfrac{p_{2k}}{q_{2k}}>\dfrac{p_{2k-2}}{q_{2k-2}}$，　$\dfrac{p_{2k}}{q_{2k}}<\dfrac{p_{2k+1}}{q_{2k+1}}$；

(3) $\dfrac{p_k}{q_k}$为既约分数，即 p_k 与 q_k 互素.

证明　(1) 根据定理 2.1.20，显然有 $q_k\geqslant 1$，又因为 $a_k\geqslant 1$，所以当 $k\geqslant 2$ 时，有

$$q_k=a_kq_{k-1}+q_{k-2}\geqslant q_{k-1}+1.$$

又由 $q_0=1>0$，$q_1=a_1\geqslant 1$，故用数学归纳法，假设当 $k\geqslant 2$ 时，

$$q_{k-1}\geqslant k-1,$$

则应用上面的结论可得

$$q_k\geqslant q_{k-1}+1\geqslant(k-1)+1=k.$$

于是(1)得证.

(2) 根据定理 2.1.21，由

$$p_kq_{k-2}-p_{k-2}q_k=(-1)^ka_k,$$

即

$$\frac{p_k}{q_k}-\frac{p_{k-2}}{q_{k-2}}=\frac{(-1)^ka_k}{q_kq_{k-2}},$$

可知

$$\frac{p_{2k}}{q_{2k}}-\frac{p_{2k-2}}{q_{2k-2}}=\frac{(-1)^{2k}a_{2k}}{q_{2k}q_{2k-2}}=\frac{a_{2k}}{q_{2k}q_{2k-2}}>0,$$

即

$$\frac{p_{2k}}{q_{2k}}>\frac{p_{2k-2}}{q_{2k-2}}.$$

同理，有

$$\frac{p_{2k+1}}{q_{2k+1}}-\frac{p_{2k-1}}{q_{2k-1}}=\frac{(-1)^{2k+1}a_{2k+1}}{q_{2k+1}q_{2k-1}}=\frac{-a_{2k}}{q_{2k}q_{2k-2}}<0,$$

即

$$\frac{p_{2k+1}}{q_{2k+1}}<\frac{p_{2k-1}}{q_{2k-1}}.$$

由

$$p_kq_{k-1}-p_{k-1}q_k=(-1)^{k-1},$$

即

$$\frac{p_k}{q_k}-\frac{p_{k-1}}{q_{k-1}}=\frac{(-1)^{k-1}}{q_kq_{k-1}},$$

可知

$$\frac{p_{2k+1}}{q_{2k+1}}-\frac{p_{2k}}{q_{2k}}=\frac{(-1)^{2k}}{q_{2k+1}q_{2k}}=\frac{1}{q_{2k+1}q_{2k}}>0,$$

即

$$\frac{p_{2k}}{q_{2k}}<\frac{p_{2k+1}}{q_{2k+1}}.$$

(3) 根据定理 2.1.21，有

$$p_kq_{k-1}-p_{k-1}q_k=(-1)^{k-1}.$$

可知当 k 为奇数时，存在整数 $x=q_{k-1}$，$y=-p_{k-1}$ 使

$$p_kx+q_ky=1,$$

当 k 为偶数时，存在整数 $x=-q_{k-1}$，$y=p_{k-1}$ 使

$$p_kx+q_ky=1,$$

再根据定理 2.1.13，可知 p_k 与 q_k 互素，即$\frac{p_k}{q_k}$为既约分数.

定理 2.1.23 每一个简单连分数表示一个实数.

证明 每一个有限简单连分数显然表示一个有理数. 我们考虑无限简单连分数

$$[a_0, a_1, \cdots, a_k, \cdots],$$

$\frac{p_k}{q_k}$，$k\geqslant 0$ 是它的渐进分数. 由定理 2.1.22 可知

$$\frac{p_0}{q_0}, \frac{p_2}{q_2}, \cdots, \frac{p_{2k}}{q_{2k}}, \cdots$$

是一个单调递增数列，而

$$\frac{p_1}{q_1}, \frac{p_3}{q_3}, \cdots, \frac{p_{2k+1}}{q_{2k+1}}, \cdots$$

是一个单调递减数列，且

$$\frac{p_1}{q_1}>\frac{p_{2k+1}}{q_{2k+1}}>\frac{p_{2k}}{q_{2k}}>\frac{p_0}{q_0},$$

所以这两个数列也是有界的. 又因为

$$0<\frac{p_{2k+1}}{q_{2k+1}}-\frac{p_{2k}}{q_{2k}}=\frac{1}{q_{2k+1}q_{2k}}\leqslant\frac{1}{2k(2k+1)},$$

而

$$\lim_{k\to\infty}\frac{1}{2k(2k+1)}=0,$$

所以$\left[\frac{p_{2k}}{q_{2k}},\ \frac{p_{2k+1}}{q_{2k+1}}\right](k=0,1,2,\cdots)$作成一个区间套，则$\lim\limits_{k\to\infty}\frac{p_k}{q_k}$存在且唯一，于是定理得证.

上面证明了任一简单连分数表示一个唯一的实数，那么任一实数能否唯一地表示成简单连分数呢？下面就来讨论这个问题.

首先，给出一个直观的理解，将一个有理数写成$\frac{\text{分子}}{\text{分母}}$的形式，当然也可以把它看作另一个的等效形式$\frac{\text{被除数}}{\text{除数}}$，由带余除法得到

$$\frac{\text{被除数}}{\text{除数}}=\text{商}+\frac{\text{余数}}{\text{除数}}=\text{商}+\frac{1}{\frac{\text{除数}}{\text{余数}}}=\text{商}+\frac{1}{\frac{\text{新的被除数}}{\text{新的除数}}}=\text{商}+\frac{1}{\text{新的商}+\frac{\text{新的余数}}{\text{新的除数}}},$$

那么，反复利用这个过程，直到最新的余数为 0，就会得到该有理数的连分数形式，显然利用辗转相除法很快可以得到上式中的各个商. 上面过程的数学形式如下：

设 α 是一给定实数，若 α 是有理数，则 $\alpha=\frac{p}{q}$，其中 p，q 为整数，且 $q>0$. 由辗转相除法可得

$$\begin{aligned}
p&=a_0q+r_1, && 0<r_1<q,\\
q&=a_1r_1+r_2, && 0<r_2<r_1,\\
&\vdots\\
r_{n-2}&=a_{n-1}r_{n-1}+r_n, && 0<r_n<r_{n-1},\\
r_{n-1}&=a_nr_n+r_{n+1}, && r_{n+1}=0.
\end{aligned}$$

于是，有理数 $\alpha=\frac{p}{q}$可以表示为如下的有限简单连分数：

$$a_0+\cfrac{1}{a_1+\cfrac{1}{a_2+\cfrac{1}{a_3+\\ \vdots\\ +\frac{1}{a_n}}}}$$

即 $\alpha=\frac{p}{q}=[a_0,a_1,\cdots,a_n]$. 因此，任一有理数均可表示成有限简单连分数 .

若 α 是无理数，则由 $\alpha=[\alpha]+\{\alpha\}$，$0<\{\alpha\}<1$ 可得

$$\alpha=a_0+\frac{1}{\alpha_1},\quad a_0=[\alpha],\quad \alpha_1=\frac{1}{\{\alpha\}}>1,$$

$$\alpha_1=a_1+\frac{1}{\alpha_2},\quad a_1=[\alpha_1],\quad \alpha_2=\frac{1}{\{\alpha_1\}}>1,$$

$$\vdots$$

$$\alpha_k=a_k+\frac{1}{\alpha_{k+1}},\ a_k=[\alpha_k],\alpha_{k+1}=\frac{1}{\{\alpha_k\}}>1,$$

$$\vdots$$

于是，有 $\alpha=[a_0,a_1,\cdots,a_k,\alpha_{k+1}]$，显然 $\alpha_{k+1}=[a_{k+1},a_{k+2},\cdots]$.

定理 2.1.24 任一无理数可表示成无限简单连分数.

证明 对于无理数 α，通过上述步骤，可知当 $k\geqslant 1$ 时，$a_k=[\alpha_k]\geqslant 1$，于是只要证明

$$\lim_{k\to\infty}[a_0,a_1,\cdots,a_k]=\alpha,$$

即

$$\lim_{k\to\infty}\frac{p_k}{q_k}=\alpha.$$

由于

$$\frac{p_k}{q_k}=\frac{a_kp_{k-1}+p_{k-2}}{a_kq_{k-1}+q_{k-2}},$$

故

$$\begin{aligned}\alpha&=[a_0,a_1,\cdots,a_k+\frac{1}{\alpha_{k+1}}]\\&=\frac{(a_k+\frac{1}{\alpha_{k+1}})p_{k-1}+p_{k-2}}{(a_k+\frac{1}{\alpha_{k+1}})q_{k-1}+q_{k-2}}\\&=\frac{\alpha_{k+1}(a_kp_{k-1}+p_{k-2})+p_{k-1}}{\alpha_{k+1}(a_kq_{k-1}+q_{k-2})+q_{k-1}}\\&=\frac{\alpha_{k+1}p_k+p_{k-1}}{\alpha_{k+1}q_k+q_{k-1}}\end{aligned}$$

因此，再根据定理 2.1.21，有

$$\alpha-\frac{p_k}{q_k}=\frac{q_kp_{k-1}-q_{k-1}p_k}{q_k(\alpha_{k+1}q_k+q_{k-1})}=\frac{(-1)^k}{q_k(\alpha_{k+1}q_k+q_{k-1})}.$$

因为 $\alpha_k>a_k$，所以 $\alpha_{k+1}q_k+q_{k-1}>q_{k+1}$，又由定理 2.1.22，可知

$$\left|\alpha-\frac{p_k}{q_k}\right|<\frac{1}{q_kq_{k+1}}\leqslant\frac{1}{k(k+1)}.$$

于是，由 $\lim\limits_{k\to\infty}\frac{1}{k(k+1)}=0$，可知 $\lim\limits_{k\to\infty}\frac{p_k}{q_k}=\alpha$. 定理得证.

我们已经知道了任一无理数可表示成无限简单连分数，下面来证明其表示的唯一性.

定理 2.1.25 任一无理数只可表示成唯一的无限简单连分数.

证明 我们只需证明如果两个无限简单连分数

$$\alpha_0=[a_0,a_1,\cdots,a_k,\cdots],\quad \beta_0=[b_0,b_1,\cdots,b_k,\cdots]$$

相等，则 $a_k=b_k$，$k=0,1,2,\cdots$.

令 $\alpha_k=[a_k,a_{k+1},\cdots]$，$\beta_k=[b_k,b_{k+1},\cdots]$，则有

$$\alpha_k=a_k+\frac{1}{\alpha_{k+1}},\ \alpha_{k+1}>1,$$

$$\beta_k=b_k+\frac{1}{\beta_{k+1}},\ \beta_{k+1}>1,$$

于是

$$a_k=[\alpha_k],\quad b_k=[\beta_k].$$

利用数学归纳法. 因为 $\alpha_0=\beta_0$，所以 $a_0=[\alpha_0]=[\beta_0]=b_0$，且有 $\alpha_1=\beta_1$.

假设对于 k，有 $a_i=b_i$，$\alpha_{i+1}=\beta_{i+1}$，其中 $i=1,2,\cdots,k$，则对于 $k+1$，有

$$a_{k+1}=[\alpha_{k+1}]=[\beta_{k+1}]=b_{k+1},$$

且 $\alpha_{k+2}=\beta_{k+2}$. 于是定理得证.

而有理数的情况有些特殊. 我们知道任一有限简单连分数表示一个有理数，任一有理数均可表示成有限简单连分数，但这种表示不是唯一的. 类似于无理数表示成无限简单连分数的唯一性的证明，可得出以下结论.

定理 2.1.26　(1) 若有理分数 $\alpha=[a_0,a_1,\cdots,a_n]=[b_0,b_1,\cdots,b_m]$，且 $a_n>1$，$b_m>1$，则有

$$m=n,\quad a_i=b_i(i=0,1,\cdots,n).$$

(2) 任一有理分数 α 有且仅有两种有限简单连分数表示式，即

$$\alpha=[a_0,a_1,\cdots,a_n]=[a_0,a_1,\cdots,a_n-1,1],$$

其中 $a_n>1$.

[例 2.1.10]　将 $\frac{547}{263}$ 表示成简单连分数.

解　由辗转相除法可得

$$\begin{aligned}547&=2\times263+21\\263&=12\times21+11\\21&=1\times11+10\\11&=1\times10+1\\10&=10\times1\end{aligned}$$

于是 $\frac{547}{263}=[2,12,1,1,10]=[2,12,1,1,9,1]$.

[例 2.1.11]　将 $\sqrt{3}$ 表示成简单连分数.

解　对于无理数 $\alpha=\sqrt{3}$，有

$$a_0=[\sqrt{3}]=1,\quad \alpha_1=\frac{1}{\sqrt{3}-1}=\frac{\sqrt{3}+1}{2},$$

$$a_1=[\alpha_1]=1,\quad \alpha_2=\frac{2}{\sqrt{3}-1}=\sqrt{3}+1,$$

$$a_2=[\alpha_2]=2,\quad \alpha_3=\frac{1}{\sqrt{3}-1}=\alpha_1,$$

于是 $\sqrt{3}=[1,1,2,1,2,\cdots]$.

定义 2.1.8　对于无限简单连分数 $[a_0,a_1,a_2,\cdots]$，如果存在整数 $m\geqslant0$，且对于 m 存在正整数 k 使得对于所有 $n\geqslant m$，有

$$a_{n+k}=a_k,$$

那么，把这个无限简单连分数叫作**循环简单连分数**，简称**循环连分数**，记为

$$[a_0, a_1, \cdots, a_{m-1}, \overline{a_m, \cdots, a_{m+k-1}}].$$

显然，$\sqrt{3}=[1, \overline{1, 2}]$是循环连分数.

2.1.4 算术基本定理

前面讨论了一些有关素数和整数分解的问题，知道任意一个大于 1 的整数都至少有两个正因子，即 1 和它本身，且必有素因子. 那么是否每个整数一定可以唯一表示成若干素数的乘积呢？接下来的部分就讨论这个问题.

定理 2.1.27 设 p 为素数且 $p|ab$，则 $p|a$ 或 $p|b$.

证明 若 a 能被 p 整除，则定理显然得证. 若 a 不能被 p 整除，则$(a, p)=1$，可知存在整数 m，n，使得

$$ma+np=1,$$

所以

$$mab+npb=b.$$

由于 $p|ab$，所以 $p|b$. 定理得证.

推论 设 p 为素数，若 $p|a_1a_2\cdots a_n$，其中 a_1，a_2，…，a_n 是 n 个整数，则 $p|a_1$，$p|a_2$，…，$p|a_n$ 至少有一个成立.

证明 用数学归纳法.

当 $n=2$ 时，根据定理 2.1.27，显然成立.

假设 $n-1$ 时命题成立，即若 $p|a_1a_2\cdots a_{n-1}$，则 $p|a_1$，$p|a_2$，…，$p|a_{n-1}$至少有一个成立.

对于 n，由于 $p|(a_1a_2\cdots a_{n-1})\ a_n$，所以 $p|a_1a_2\cdots a_{n-1}$或 $p|a_n$. 再根据归纳假设，可知 $p|a_1$，$p|a_2$，…，$p|a_{n-1}$，$p|a_n$ 至少有一个成立. 命题得证.

上面的定理及其推论非常重要，因为它们给出了素数最重要的特点之一. 如果 p 不是素数，上面的结果不一定成立. 例如，$6|12$ 且 $12=3\times4$，但是 $6\nmid3$ 且 $6\nmid4$.

定理 2.1.28 设 a_1，a_2，…，a_n，c 是整数，如果$(a_i, c)=1$，$1\leqslant i\leqslant n$，则$(a_1a_2\cdots a_n, c)=1$.

证明 用反证法. 假设存在大于 1 的整数 m 满足$(a_1a_2\cdots a_n, c)=m$，则必存在素数 p，使 $p|m$，于是 $p|a_1a_2\cdots a_n$ 且 $p|c$. 由上面的推论，可知 $p|a_1$，$p|a_2$，…，$p|a_n$ 至少有一个成立，则$(a_i, c)=p$ 至少有一个成立. 这与命题中$(a_i, c)=1$ 矛盾，于是假设不成立. 定理得证.

定理 2.1.29 任一大于 1 的整数都可以表示成素数的乘积，且在不考虑乘积顺序的情况下，该表达式是唯一的. 即

$$n=p_1p_2\cdots p_s,\ p_1\leqslant p_2\leqslant\cdots\leqslant p_s, \tag{2.1.12}$$

其中 p_1，p_2，…，p_s 是素数，并且若

$$n=q_1q_2\cdots q_t,\ q_1\leqslant q_2\leqslant\cdots\leqslant q_t, \tag{2.1.13}$$

其中 q_1，q_2，…，q_t 是素数，则 $s=t$，$p_i=q_i(i=1, 2, \cdots, s)$.

证明 首先，用数学归纳法证明式(2.1.12)成立.

当 $n=2$ 时，式(2.1.12)显然成立.

假设对于一切大于 1 且小于 n 的正整数，式(2.1.12)都成立.

对于正整数 n，若 n 是素数，则式(2.1.12)对 n 成立.

若 n 是合数，则存在正整数 b，c 满足条件

$$n=bc,\ 1<b\leqslant c<a,$$

由归纳法假设，b 和 c 分别能表示成素数的乘积，故 n 能表示成素数的乘积，即式(2.1.12)成立.

下面证明唯一性.

假设对 n 同时有式(2.1.12)和式(2.1.13)成立，则

$$p_1p_2\cdots p_s=q_1q_2\cdots q_t. \tag{2.1.14}$$

由定理 2.1.27 可知，$\exists\, p_k$，q_j 使得 $p_1\mid q_j$，$q_1\mid p_k$，但由于 p_k 和 q_j 均为素数，故 $p_1=q_j$，$q_1=p_k$. 又 $p_1\leqslant p_k$，$q_1\leqslant q_j$，故同时有 $p_1\leqslant q_1$，$q_1\leqslant p_1$，因此 $p_1=q_1$，由式(2.1.14)得

$$p_2\cdots p_s=q_2\cdots q_t.$$

同理可得 $p_2=q_2$，$p_3=q_3$，依此类推，可知 $s=t$ 时，$p_s=q_s$. 唯一性得证.

以上定理被称为**算术基本定理**，也叫作整数的**唯一分解定理**，它反映了整数的本质. 将式(2.1.13)中相同的素数乘积写成素数幂的形式，可得以下推论.

推论　任一大于 1 的整数都能够唯一地表示成

$$n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},\ \alpha_i>0,\ i=1,2,\cdots,s, \tag{2.1.15}$$

其中 $p_i<p_j(i<j)$是素数.

式(2.1.15)称为 n 的**标准分解式**.

定理 2.1.30　设 n 是大于 1 的任一整数，其标准分解式由式(2.1.15)给出，那么 d 是 n 的正因子的充要条件是

$$d=p_1^{\beta_1}p_2^{\beta_2}\cdots p_s^{\beta_s},\ \alpha_i\geqslant\beta_i\geqslant 0,\ i=1,2,\cdots,s. \tag{2.1.16}$$

证明　先证充分性. 若式(2.1.16)成立，则存在正整数

$$c=p_1^{\alpha_1-\beta_1}p_2^{\alpha_2-\beta_2}\cdots p_s^{\alpha_s-\beta_s},$$

显然有 $n=cd$，所以 $d\mid n$.

再证必要性.

设 $d\mid n$，且 d 有素因子分解式

$$d=p_1^{\beta_1}p_2^{\beta_2}\cdots p_s^{\beta_s},\ \beta_i\geqslant 0,\ i=1,2,\cdots,s,$$

则必有

$$\alpha_i\geqslant\beta_i,\ i=1,2,\cdots,s.$$

否则，至少存在一个 i 满足 $1\leqslant i\leqslant s$，使 $\alpha_i<\beta_i$. 不妨设 $\alpha_1<\beta_1$. 由于 $d\mid n$ 及 $p_1^{\beta_1}\mid d$，所以

$$p_1^{\beta_1}\mid p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s}.$$

又由 $p_1^{\alpha_1}>0$，可得

$$p_1^{\beta_1-\alpha_1}\mid p_2^{\alpha_2}\cdots p_s^{\alpha_s}.$$

根据定理 2.1.27 的推论，可知存在 k 满足 $2\leqslant k\leqslant s$，使 $p_1\mid p_k$，这是不可能的. 于是必要性亦得证.

由以上定理可知，只要知道了正整数 n 的标准分解式，那么其所有的正因子也就都可以知道了，且可以由式(2.1.16)给出. 不难得出以下定理.

定理 2.1.31　设正整数 n 的标准分解式为

$$n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},\ \alpha_i>0,\ i=1,2,\cdots,s,$$

$\tau(n)$表示 n 的所有正因子的个数，则

$$\tau(n)=\tau(p_1^{\alpha_1})\tau(p_2^{\alpha_2})\cdots\tau(p_s^{\alpha_s})=(\alpha_1+1)(\alpha_2+1)\cdots(\alpha_s+1).$$

[例 2.1.11] 计算 360 的所有正因子的个数.

解 因为 $360=2^3\times3^2\times5$，所以

$$\tau(360)=(3+1)(2+1)(1+1)=24.$$

再根据最大公因子和最小公倍数的定义，显然得到如下结论.

定理 2.1.32 设 a，b 为两个正整数，其素因子分解式分别为

$$a=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},\ \alpha_i\geqslant0,\ i=1,2,\cdots,s,$$
$$b=p_1^{\beta_1}p_2^{\beta_2}\cdots p_s^{\beta_s},\ \beta_i\geqslant0,\ i=1,2,\cdots,s,$$

那么

$$(a,b)=p_1^{\gamma_1}p_2^{\gamma_2}\cdots p_s^{\gamma_s},\ \gamma_i=\min\{\alpha_i,\beta_i\},\ i=1,2,\cdots,s,$$
$$[a,b]=p_1^{\delta_1}p_2^{\delta_2}\cdots p_s^{\delta_s},\ \delta_i=\max\{\alpha_i,\beta_i\},\ i=1,2,\cdots,s.$$

对于任意的整数 α，β，显然有

$$\min\{\alpha,\beta\}+\max\{\alpha,\beta\}=\alpha+\beta,$$

由此可得

$$(a,b)[a,b]=ab,$$

由于(a,b)不可能为零，所以这个结果和定理 2.1.17 中已经证明过的结果相同.

[例 2.1.12] 计算整数 90，30，114，42，81 的最大公因子与最小公倍数.

解 先写出这些整数的标准分解式，即

$$90=2\times3^2\times5,$$
$$30=2\times3\times5,$$
$$114=2\times3\times19,$$
$$42=2\times3\times7,$$
$$81=3^4,$$

于是

$$(90,30)=2\times3\times5=30,$$
$$(30,114)=2\times3=6,$$
$$(6,42)=2\times3=6,$$
$$(6,81)=3,$$

所以整数 90，30，114，42，81 的最大公因子是 3.

由于

$$[90,30]=2\times3^2\times5=90,$$
$$[90,114]=2\times3^2\times5\times19=1\ 710,$$
$$[1\ 710,42]=2\times3^2\times5\times7\times19=11\ 970,$$
$$[11\ 970,81]=2\times3^4\times5\times7\times19=107\ 730,$$

所以整数 90，30，114，42，81 的最小公倍数是 107 730.

[例 2.1.13] 证明对于正整数 a，b，c，有$(a,[b,c])=[(a,b),(a,c)]$.

证明 由于其素因子分解式可分别写为

$$a=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},$$

$$b=p_1^{\beta_1}p_2^{\beta_2}\cdots p_s^{\beta_s},$$

$$c=p_1^{\gamma_1}p_2^{\gamma_2}\cdots p_s^{\gamma_s},$$

其中 $\alpha_i,\beta_i,\gamma_i\geqslant 0$，$i=1,2,\cdots,s$，则

$$(a,[b,c])=p_1^{\eta_1}p_2^{\eta_2}\cdots p_s^{\eta_s},$$

其中 $\eta_i=\min\{\alpha_i,\max\{\beta_i,\gamma_i\}\}$，$i=1,2,\cdots,s$.

$$[(a,b),(a,c)]=p_1^{\tau_1}p_2^{\tau_2}\cdots p_s^{\tau_s},$$

其中 $\tau_i=\max\{\min\{\alpha_i,\beta_i\},\min\{\alpha_i,\gamma_i\}\}$，$i=1,2,\cdots,s$.

不难验证，对于 $i=1,2,\cdots,s$，无论 $\alpha_i,\beta_i,\gamma_i$ 有怎样的大小关系，$\eta_i=\tau_i$ 总是成立的. 于是命题得证.

[例 2.1.14]　设 a,b 是两个正整数，则存在整数 c,d，满足 $c\mid a$，$d\mid b$，使得

$$cd=[a,b],\quad (c,d)=1.$$

证明　设 a,b 可写成如下的因子分解式

$$a=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},\ b=p_1^{\beta_1}p_2^{\beta_2}\cdots p_s^{\beta_s},$$

其中当 $i=1,2,\cdots,t$ 时，$\alpha_i\geqslant\beta_i\geqslant 0$；当 $i=t+1,\cdots,s$ 时，$\beta_i>\alpha_i\geqslant 0$.

取

$$c=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_t^{\alpha_t},\quad d=p_{t+1}^{\beta_{t+1}}\cdots p_s^{\beta_s}$$

即为所求.

2.1.5　梅森素数和费马素数

由于素数在数论中占有最重要的地位，数学家一直希望能够找到能够描述素数的简单规律，尽管这样的规律到目前为止还没有找到，但是在这个过程中提出的一些问题，尤其是关于具有一些特定形式的素数的问题及相关概念，对密码学等具有比较重要的应用价值.

定义 2.1.9　若正整数 n 的所有正因子之和等于 $2n$，则 n 称为**完全数**.

今后以 $\sigma(n)$ 表示正整数 n 的所有正因子之和，于是，若 n 为完全数，则有 $\sigma(n)=2n$.

[例 2.1.15]　判断 6 和 28 是否为完全数.

解　由于 6 的所有正因子为 1，2，3，6，则

$$\sigma(6)=1+2+3+6=2\times 6,$$

所以 6 是完全数. 又 28 的所有正因子为 1，2，4，7，14，28，则

$$\sigma(28)=1+2+4+7+14+28=2\times 28,$$

所以 28 也是完全数.

定理 2.1.33　若正整数 n 的标准分解式为

$$n=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},$$

则

$$\sigma(n)=\frac{p_1^{\alpha_1+1}-1}{p_1-1}\cdots\frac{p_s^{\alpha_s+1}-1}{p_s-1}.$$

证明　由定理 2.1.30 可知，n 的所有因子可表示为

$$p_1^{x_1}p_2^{x_2}\cdots p_s^{x_s},\ 0\leqslant x_i\leqslant\alpha_i,\ i=1,2,\cdots,s,$$

故

$$\begin{aligned}\sigma(n)&=\sum_{x_1=0}^{\alpha_1}\cdots\sum_{x_s=0}^{\alpha_s}p_1^{x_1}\cdots p_s^{x_s}\\&=\sum_{x_1=0}^{\alpha_1}p_1^{x_1}\cdots\sum_{x_s=0}^{\alpha_s}p_s^{x_s}\\&=\frac{p_1^{\alpha_1+1}-1}{p_1-1}\cdots\frac{p_s^{\alpha_s+1}-1}{p_s-1}\end{aligned}$$

证毕.

定理 2.1.34 若 2^n-1 为素数，则 $2^{n-1}(2^n-1)$为偶完全数，且无其他偶完全数存在.

证明 令 $P=2^n-1$，则

$$2^{n-1}(2^n-1)=2^{n-1}p,$$

于是

$$\sigma(2^{n-1}p)=\frac{2^n-1}{2-1}\,\frac{p^2-1}{p-1}=p(p+1)=2^np.$$

所以 $2^{n-1}(2^n-1)$为完全数. 又显然有 $n\geqslant 2$，故 $2^{n-1}(2^n-1)$为偶完全数.

若 a 为一偶完全数，不妨令 $a=2^{n-1}q$，其中 $n\geqslant 2$，q 为奇数，则有

$$\sigma(a)=2a=2^nq=\frac{2^n-1}{2-1}\cdot\sigma(q),$$

故

$$\sigma(q)=\frac{2^nq}{2^n-1}=q+\frac{q}{2^n-1},$$

可知 $2^n-1\mid q$，则 q 和$\frac{q}{2^n-1}$均为 q 的因子，又 $\sigma(q)$为 q 的所有正因子之和，故 q 只有两个正因子，由整数的唯一分解定理和素数定义知 q 为素数，且

$$\frac{q}{2^n-1}=1.$$

所以 $q=2^n-1$，即 $a=2^{n-1}(2^n-1)$. 证毕.

于是，寻找偶完全数的问题就可化为寻找形如 2^n-1 的素数的问题. 而“2^n-1 是素数”与“n 是素数”之间是否存在着一定的联系呢？当 n 等于 2，3，5，7 时，2^n-1 毫无疑问是素数，但 $2^{11}-1=2\,047=23\times 89$. 由此可见，若 n 是素数，2^n-1 不一定是素数.

定理 2.1.35 若 2^n-1 为素数，则 n 必为素数.

证明 对于 $n>1$，假设 n 为合数，即 $n=bc$，其中 b，c 均为大于 1 的整数，则 $2^b-1\mid 2^n-1$，所以 2^n-1 为合数，于是定理得证.

定义 2.1.10 设 p 是一个素数，形如 2^p-1 的数叫作**梅森数**，记为 $M_p=2^p-1$. 当 M_p 为素数时，则称其为**梅森素数**.

梅森(Marin Mersenne，1588—1648)是法国的修道士，也是一位数学家. 当时，他无证明地提出，对不大于 257 的素数 p，当且仅当 $p=2$，3，5，7，13，17，19，31，67，127，257 时，M_p 为素数. 当然，他的这个结论是有错误的. 其中，M_{67} 和 M_{257} 是合数，而 M_{61}，M_{89}，M_{107} 也应该是素数，但这些结果的全部给出却是在三百多年以后的 1947 年. 至今已经知道有 40 个梅森素数，下面列出了它们所对应的 40 个 p.

2　　3　　5　　7　　13　　17　　19　　31

61	89	107	127	521	607	1 279	2 203
2 281	3 217	4 253	4 423	9 689	9 941	11 213	19 937
21 701	23 209	44 497	86 243	110 503	132 049	216 091	756 839
859 433	1 257 787	1 398 269	2 976 221	3 021 377			
6 972 593	13 466 917	20 996 011					

我们知道，每发现一个梅森素数，就可以相应地得到一个偶完全数. 是否存在无穷多个 p 使 M_p 为素数，进而得到无穷多个偶完全数，这是至今尚未解决的数论难题. 那么是否存在奇完全数呢？尽管几百年来许多数学家对此问题进行了大量的研究，但至今仍未解决.

定理 2.1.36　若 2^m+1 为素数，则 $m=2^n$.

证明　假设 m 有一个奇因子 q，令 $m=qr$，则

$$2^{qr}+1=(2^r)^q+1=(2^r+1)(2^{r(q-1)}-\cdots+1),$$

又 $1<2^r+1<2^{qr}+1$，故 2^m+1 非素数，与已知条件矛盾. 所以 m 没有奇因子，定理得证.

定义 2.1.11　若 n 为非负整数，则称 $F_n=2^{2^n}+1$ 为**费马数**. 当 F_n 为素数时，则称其为**费马素数**.

最前 5 个费马数分别为 $F_0=3$，$F_1=5$，$F_2=17$，$F_3=257$，$F_4=65\ 537$，它们都是素数. 据此，1640 年，法国数学家费马(Pierre de Fermat，1601—1665)猜想凡 F_n 皆为素数. 1732 年，瑞士数学家欧拉(Leonhard Euler，1707—1783)发现 $F_5=641\times 6\ 700\ 417$，故费马猜想并不正确，并且到目前为止，也只发现了这 5 个费马素数，因此有人推测仅存在有限个费马素数.

定理 2.1.37　任给两个费马数 F_a，F_b，$a\neq b$，则 F_a，F_b 互素.

证明　不妨设 $a>b\geqslant 0$，$a=b+c$，$c>0$，存在正整数 n，满足 $n\mid F_b$ 且 $n\mid F_{b+c}$，显然 n 必为奇数. 令 $t=2^{2^b}$，则有

$$\frac{F_{b+c}-2}{F_b}=\frac{2^{2^{b+c}}-1}{2^{2^b}+1}=\frac{t^{2^c}-1}{t+1}=t^{2^c-1}-t^{2^c-2}+\cdots-1,$$

故 $F_b\mid F_{b+c}-2$，又由 $n\mid F_b$ 且 $n\mid F_{b+c}$，可知 $n\mid 2$. 因为 n 是奇数，所以 $n=1$，即 F_a，F_b 互素.

2.2　同余

2.2.1　同余的概念和性质

在人们最开始学习两个整数除法的时候，可能比较关注于计算得到的商. 但是，从这一节开始，我们将要把视角变化一下，就是要关注于计算得到的余数. 如果两个整数 a 和 b 同时为奇数或者同时为偶数，那么我们早就知道称它们具有相同的奇偶性，其充分必要条件是：$a-b$ 是偶数，即 $2\mid(a-b)$，换个说法就是 a 和 b 被 2 除的时候具有相同的余数. 同余的理论就是从推广奇偶性这个概念开始的，只不过是奇偶性中整数 2 的角色被某个任意指定的正整数所替代. 为此，先引入同余与同余式的概念.

定义 2.2.1　给定一个正整数 m，如果用 m 去除两个整数 a 和 b 所得的余数相同，则称 a 和 b 模 m **同余**，记作

$$a\equiv b\pmod m;\tag{2.2.1}$$

否则，称 a 和 b 模 m **不同余**，记作

$$a \not\equiv b \pmod{m}.$$

关系式(2.2.1)称为模 m 的同余式，或简称**同余式**.

例如，$26 \equiv 2 \pmod 3$，$63 \equiv 3 \pmod 5$，$23 \equiv -5 \pmod 7$.

定理 2.2.1 整数 a 和 b 模 m 同余的充要条件是 $m \mid a-b$.

证明 先证必要性. 由 $a \equiv b \pmod m$，可设

$$a = mq_1 + r,\ b = mq_2 + r,\ 0 \leqslant r < m,$$

则 $a-b = m(q_1 - q_2)$，即 $m \mid a-b$.

再证充分性. 设

$$a = mq_1 + r_1,\ 0 \leqslant r_1 < m,$$
$$b = mq_2 + r_2,\ 0 \leqslant r_2 < m,$$

则 $a-b = m(q_1-q_2) + r_1 - r_2$. 由 $m \mid a-b$，可知 $m \mid r_1 - r_2$，则 $m \mid \mid r_1 - r_2 \mid$. 又因 $0 \leqslant r_2 < m$，所以 $-m < -r_2 \leqslant 0$，与 $0 \leqslant r_1 < m$ 两个不等式相加，得到 $-m \leqslant r_1 - r_2 \leqslant m$，即 $\mid r_1 - r_2 \mid < m$，故 $\mid r_1 - r_2 \mid = 0$，所以 $r_1 = r_2$. 定理得证.

于是，同余又可以定义如下，即若 $m \mid a-b$，则称 a 和 b 模 m 同余. 根据整除的定义，可以很直观地给出另一个判别同余的充要条件.

定理 2.2.2 整数 a 和 b 模 m 同余的充要条件是存在一个整数 k 使得

$$a = b + km.$$

由同余的定义，可以得到整数之间的同余具有等价关系的性质，利用它可以快捷地判断两个整数 a 和 b 是否模 m 同余.

定理 2.2.3 同余具有等价关系，即

(1) 自反性：$a \equiv a \pmod m$；

(2) 对称性：若 $a \equiv b \pmod m$，则 $b \equiv a \pmod m$；

(3) 传递性：若 $a \equiv b \pmod m$，$b \equiv c \pmod m$，则 $a \equiv c \pmod m$.

证明 (1)和(2)的证明略.

(3) 由 $m \mid a-b$ 和 $m \mid b-c$，得到 $m \mid [(a-b)+(b-c)]$，即 $m \mid a-c$.

定理 2.2.4 设 a_1，a_2，b_1，b_2 为四个整数，如果

$$a_1 \equiv b_1 \pmod m,\quad a_2 \equiv b_2 \pmod m,$$

则有

(1) $a_1 x + a_2 y \equiv b_1 x + b_2 y \pmod m$，其中 x，y 为任意整数；

(2) $a_1 a_2 \equiv b_1 b_2 \pmod m$；

(3) $a_1^n \equiv b_1^n \pmod m$，其中 $n > 0$.

证明 (1) 由于 $m \mid a_1 - b_1$，$m \mid a_2 - b_2$，故 $m \mid x(a_1-b_1) + y(a_2-b_2)$，又

$$x(a_1-b_1) + y(a_2-b_2) = (a_1x + a_2y) - (b_1x + b_2y),$$

则 $m \mid (a_1x + a_2y) - (b_1x + b_2y)$，即 $a_1x + a_2y \equiv b_1x + b_2y \pmod m$.

(2) 由于 $m \mid a_1 - b_1$，$m \mid a_2 - b_2$，故 $m \mid a_2(a_1-b_1) + b_1(a_2-b_2)$，又

$$a_2(a_1-b_1) + b_1(a_2-b_2) = a_1a_2 - b_1b_2,$$

则 $m \mid a_1a_2 - b_1b_2$，即 $a_1a_2 \equiv b_1b_2 \pmod m$.

(3) 由(2)可证.

[例 2.2.1]　求 $3^{2\,006}$，$3^{2\,009}$ 写成十进制数时的个位数.

解　由于

$$3^4\equiv 1 \pmod{10},$$

可得 $3^{4\times 501}\equiv 1 \pmod{10}$. 又 $3^2\equiv 9 \pmod{10}$，$2\,006=4\times 501+2$，故此可得 $3^{2\,006}\equiv 9 \pmod{10}$. 所以 $3^{2\,006}$ 写成十进制数时的个位数是 9.

同样地，由于

$$3^1\equiv 3 \pmod{10},\ 3^4\equiv 1 \pmod{10},$$

故可得 $3^{4\times 502}\equiv 1 \pmod{10}$. 又 $2\,009=4\times 502+1$，因此可得 $3^{2\,009}\equiv 3 \pmod{10}$. 所以 $3^{2\,009}$ 写成十进制数时的个位数是 3.

[例 2.2.2]　已知 2009 年 3 月 9 日是星期一，问之后第 2^{100} 天是星期几？2^{200} 天之后呢？

解　由于

$$2^1\equiv 2 \pmod{7},\ 2^3\equiv 1 \pmod{7},$$

可得 $2^{3\times 33}\equiv 1 \pmod{7}$. 又 $100=3\times 33+1$，则 $2^{100}\equiv 2 \pmod{7}$. 所以之后第 2^{100} 天是星期三.

同样地，由于

$$2^2\equiv 4 \pmod{7},\ 2^3\equiv 1 \pmod{7},$$

可得 $2^{3\times 66}\equiv 1 \pmod{7}$. 又 $200=3\times 66+2$，则 $2^{200}\equiv 4 \pmod{7}$. 所以之后第 2^{200} 天是星期五.

定理 2.2.5　设 $f(t)=a_nt^n+a_{n-1}t^{n-1}+\cdots+a_1t+a_0$ 与 $g(t)=b_nt^n+b_{n-1}t^{n-1}+\cdots+b_1t+b_0$ 是两个整系数多项式，满足

$$a_i\equiv b_i \pmod{m},\quad 0\leqslant i\leqslant n,$$

那么，若 $x\equiv y \pmod{m}$，则

$$f(x)\equiv g(y) \pmod{m}.$$

证明　由 $x\equiv y \pmod{m}$，可得

$$x^i\equiv y^i \pmod{m},\quad 0\leqslant i\leqslant n,$$

又 $a_i\equiv b_i \pmod{m}$，$0\leqslant i\leqslant n$，将它们对应相乘，则有

$$a_ix^i\equiv b_iy^i \pmod{m},\quad 0\leqslant i\leqslant n,$$

将这些同余式左右对应相加，可得

$$a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\equiv b_ny^n+b_{n-1}y^{n-1}+\cdots+b_1y+b_0 \pmod{m},$$

即 $f(x)\equiv g(y) \pmod{m}$.

[例 2.2.3]　证明正整数 n(十进制)能被 9 整除的充要条件是将 n 的各位数字相加所得之和能被 9 整除.

证明　n 可写为十进制表示式：

$$n=10^ka_k+10^{k-1}a_{k-1}+\cdots+10a_1+a_0,\ 0\leqslant a_i<10.$$

因为 $10^i\equiv 1 \pmod{9}$，$0\leqslant i\leqslant k$，所以

$$10^ka_k+10^{k-1}a_{k-1}+\cdots+10a_1+a_0\equiv a_k+a_{k-1}+\cdots+a_1+a_0 \pmod{9}.$$

因此，

$$10^ka_k+10^{k-1}a_{k-1}+\cdots+10\ a_1+a_0\equiv 0 \pmod{9}$$

的充要条件是

$$a_k+a_{k-1}+\cdots+a_1+a_0\equiv 0 \pmod{9}.$$

命题得证.

[例 2.2.4] 证明：当 n 是奇数时，2^n+1 能被 3 整除；当 n 是偶数时，2^n+1 不能被 3 整除.

证明 因为 $2\equiv -1 \pmod 3$，故 $2^n\equiv(-1)^n \pmod 3$，于是

$$2^n+1\equiv(-1)^n+1 \pmod 3.$$

因此，当 n 是奇数时，

$$2^n+1\equiv 0 \pmod 3,$$

即 2^n+1 能被 3 整除；当 n 是偶数时，

$$2^n+1\equiv 2 \pmod 3,$$

即 2^n+1 不能被 3 整除.

定理 2.2.6 若 $ac\equiv bc \pmod m$，且$(c, m)=d$，则 $a\equiv b \pmod{\frac{m}{d}}$.

证明 由 $m\mid c(a-b)$，可知$\frac{m}{d}\mid\frac{c}{d}(a-b)$，又$(\frac{m}{d}, \frac{c}{d})=1$，于是$\frac{m}{d}\mid a-b$，即

$$a\equiv b \pmod{\frac{m}{d}}.$$

例如，通过 $260\equiv 20 \pmod{30}$，$(10, 30)=10$，可得 $26\equiv 2 \pmod 3$.

定理 2.2.7 若 $a\equiv b \pmod m$，则有 $ak\equiv bk \pmod{mk}$，其中 k 为正整数.

证明 由 $m\mid a-b$，可知 $mk\mid ak-bk$，即 $ak\equiv bk \pmod{mk}$.

例如，通过 $26\equiv 2 \pmod 3$，可得 $260\equiv 20 \pmod{30}$.

定理 2.2.8 若 $a\equiv b \pmod m$，且有正整数 d 满足 $d\mid m$，则 $a\equiv b \pmod d$.

证明 由 $m\mid a-b$，$d\mid m$，可知 $d\mid a-b$，即 $a\equiv b \pmod d$.

例如，通过 $260\equiv 20 \pmod{30}$，可得 $260\equiv 20 \pmod 3$.

定理 2.2.9 若 $a\equiv b \pmod{m_i}$，$i=1, 2, \cdots, n$，则

$$a\equiv b \pmod{[m_1, m_2, \cdots, m_n]}.$$

证明 由 $m_i\mid a-b$，$i=1, 2, \cdots, n$，可知$[m_1, m_2, \cdots, m_n]\mid a-b$，即 $a\equiv b \pmod{[m_1, m_2, \cdots, m_n]}$.

例如，通过 $260\equiv 20 \pmod{30}$，$260\equiv 20 \pmod{80}$，又$[30, 80]=240$，可得 $260\equiv 20 \pmod{240}$.

定理 2.2.10 若 $a\equiv b \pmod m$，则$(a, m)=(b, m)$.

证明 由 $a\equiv b \pmod m$，可知存在整数 k 使得 $a=b+mk$，于是$(a, m)=(b, m)$.

以上介绍了同余的一些基本性质. 同余是数论中一个十分重要的概念，并且应用领域十分广泛，尤其是随着近代密码学的发展，同余及其相关理论的重要性越发显现出来.

2.2.2 剩余类和欧拉定理

因为同余是一种整数集合上的等价关系，所以可利用同余关系把全体整数划分成若干个等价类，并将每个等价类中的整数作为一个整体来考虑，进而可以得到一些相关的性质.

定义 2.2.2 设 m 是一给定正整数，令 C_r 表示所有与整数 r 模 m 同余的整数所组成的集合，则任意一个这样的 C_r 叫作模 m 的一个**剩余类**. 一个剩余类中的任一数叫作该类的**代表元**.

我们可以用集合的形式来描述剩余类的定义，即

$$C_r=\{a\mid a\in\mathbf{Z},\ a\equiv r\ (\mathrm{mod}\ m)\}=\{\cdots,\ r-2m,\ r-m,\ r,\ r+m,\ r+2m,\ \cdots\}.$$

显然 C_r 非空，因为 $r\in C_r$. 很多书中也使用$[r]$来表示 C_r.

下面的定理将考察整数与剩余类的关系和剩余类之间的关系，尽管整数有无限多个，然而剩余类的个数是有限的.

定理 2.2.11　设 m 为一正整数，C_0，C_1，…，C_{m-1}是模 m 的剩余类，则

(1) 任一整数恰包含在一个 C_r 中，这里 $0\leqslant r\leqslant m-1$；

(2) $C_a=C_b$ 的充要条件是 $a\equiv b\ (\mathrm{mod}\ m)$；

(3) C_a 与 C_b 的交集为空集的充要条件是 a 和 b 模 m 不同余.

证明　(1) 设 a 是任一整数，则存在唯一的整数 q，r 使得

$$a=qm+r,\quad 0\leqslant r<m,$$

于是有 $a\equiv r\ (\mathrm{mod}\ m)$，故 a 恰包含在 C_r 中.

(2) 先证必要性. 由于 $a\in C_a$，$b\in C_b$，又 $C_a=C_b$，显然有

$$a\equiv b\ (\mathrm{mod}\ m).$$

再证充分性. 对任意整数 $c\in C_a$，有

$$a\equiv c\ (\mathrm{mod}\ m).$$

又因为

$$b\equiv a\ (\mathrm{mod}\ m),$$

故 $b\equiv c\ (\mathrm{mod}\ m)$，即 $c\in C_b$，可见 $C_a\subseteq C_b$.

同理，对任意整数 $c\in C_b$，可证 $a\equiv c\ (\mathrm{mod}\ m)$，即 $c\in C_a$，可见 $C_b\subseteq C_a$.

于是，$C_a=C_b$.

(3) 由(2)可知必要性成立. 下面证明充分性.

用反证法. 假设 C_a 与 C_b 的交集非空，即存在整数 c 满足 $c\in C_a$ 且 $c\in C_b$，则有

$$a\equiv c\ (\mathrm{mod}\ m),$$

$$b\equiv c\ (\mathrm{mod}\ m).$$

于是，得到 $a\equiv b\ (\mathrm{mod}\ m)$，与假设矛盾. 因此 C_a 与 C_b 的交集为空集.

由上面的定理可以看到，尽管在剩余类的定义中 C_r 的下标可以在整数范围内任意取值，但是 C_r 本身必然与C_0，C_1，…，C_{m-1}中的某一个集合实际上是同一个集合，只不过是给集合取的名字不同而已，换句话说，一共就存在 m 个不同的剩余类. 例如，

$$C_m=\{\cdots,\ -m,\ 0,\ m,\ 2m,\ 3m,\ \cdots\}=C_0,$$

因此，在考察剩余类的时候，往往只需要用到 C_0，C_1，…，C_{m-1}这 m 个名字指称这 m 个集合就可以了.

定义 2.2.3　在模 m 的剩余类 C_0，C_1，…，C_{m-1}中各取一代表元 $a_i\in C_i$，$i=0$，1，…，$m-1$，则此 m 个数 a_0，a_1，…，a_{m-1}称为模 m 的一个**完全剩余系**.

由此定义和定理 2.2.11 显然可得到如下定理.

定理 2.2.12　m 个整数a_0，a_1，…，a_{m-1}为模 m 的一个完全剩余系的充要条件是它们两两模 m 不同余.

例如，以下是几个模 10 的完全剩余系：

0，1，2，3，4，5，6，7，8，9；

1，2，3，4，5，6，7，8，9，10；

10，21，22，23，34，45，46，67，78，99；

−9，−8，−7，−6，−5，−4，−3，−2，−1，0.

定义 2.2.4 对于正整数 m，

(1) 0，1，…，$m-1$ 为模 m 的一个完全剩余系，叫作模 m 的**最小非负完全剩余系**；

(2) 1，2，…，$m-1$，m 为模 m 的一个完全剩余系，叫作模 m 的**最小正完全剩余系**；

(3) $-(m-1)$，…，-1，0 为模 m 的一个完全剩余系，叫作模 m 的**最大非正完全剩余系**；

(4) $-m$，$-(m-1)$，…，-1 为模 m 的一个完全剩余系，叫作模 m 的**最大负完全剩余系**.

定理 2.2.13 设 k 是满足$(k, m)=1$ 的整数，b 是任意整数，若 a_0，a_1，…，a_{m-1}是模 m 的一个完全剩余系，则 ka_0+b，ka_1+b，…，$ka_{m-1}+b$ 也是模 m 的一个完全剩余系. 即若 x 遍历模 m 的一个完全剩余系，则 $kx+b$ 也遍历模 m 的一个完全剩余系.

证明 由定理 2.2.12，只需要证明当 a_0，a_1，…，a_{m-1}是模 m 的一个完全剩余系时，m 个整数

$$ka_0+b, ka_1+b, \cdots, ka_{m-1}+b$$

模 m 两两不同余. 用反证法，假设存在 a_i 和 $a_j(i\neq j)$使得

$$ka_i+b\equiv ka_j+b \pmod m,$$

则 $m|k(a_i-a_j)$. 由于$(k, m)=1$，所以 $m|a_i-a_j$，即 $a_i\equiv a_j(\bmod m)$，推出了矛盾，假设不成立. 于是，ka_0+b，ka_1+b，…，$ka_{m-1}+b$ 两两不同余，所以是模 m 的一个完全剩余系.

例如，0，1，2，3，4 为模 5 的一个完全剩余系，若令 $k=7$，$b=3$，则可以得到模 5 的另一个完全剩余系，即 3，10，17，24，31.

定理 2.2.14 若 $x_i(i=0, 1, \cdots, m_1-1)$是模 m_1 的完全剩余系，$y_j(j=0, 1, \cdots, m_2-1)$是模 m_2 的完全剩余系，其中$(m_1, m_2)=1$，则 $m_2x_i+m_1y_j(i=0, 1, \cdots, m_1-1, j=0, 1, \cdots, m_2-1)$是模 m_1m_2 的完全剩余系.

证明 同样由定理 2.2.12，只需要证明 $m_2x_i+m_1y_j(i=0, 1, \cdots, m_1-1, j=0, 1, \cdots, m_2-1)$这 m_1m_2 个整数模 m_1m_2 两两不同余. 用反证法，假设存在有序对(x_a, y_c)和(x_b, y_d)，且$(x_a, y_c)\neq(x_b, y_d)$，且使得

$$m_2x_a+m_1y_c\equiv m_2x_b+m_1y_d \pmod{m_1m_2},$$

由定理 2.2.8，有

$$m_2x_a+m_1y_c\equiv m_2x_b+m_1y_d \pmod{m_1},$$

即

$$m_2x_a\equiv m_2x_b \pmod{m_1}.$$

于是 $m_1|m_2(x_a-x_b)$，又$(m_1, m_2)=1$，则 $m_1|x_a-x_b$，即 $x_a\equiv x_b(\bmod m_1)$，由于它们来自于同一个模 m_1 的完全剩余系，所以 $x_a=x_b$. 同理可证，$y_c=y_d$. 说明$(x_a, y_c)=(x_b, y_d)$，与我们的假设矛盾. 所以假设不成立，定理得证.

在模 m 的一个剩余类中，若有一个数与 m 互素，则该剩余类中所有数都与 m 互素，此时称该剩余类与模 m 互素.

定义 2.2.5　与模 m 互素的剩余类的个数记为 $\varphi(m)$，$\varphi(m)$称为**欧拉函数**.

也可以说，欧拉函数 $\varphi(m)$是在序列 0，1，…，$m-1$ 中与模 m 互素的整数的个数，显然，$\varphi(m)$是一个定义在正整数集上的函数.

例如，由于 0，1，2，3，4，5 中与 6 互素的只有 1，5，因此有 $\varphi(6)=2$.

定义 2.2.6　在与模 m 互素的 $\varphi(m)$个剩余类中，各取一个代表元

$$a_1, a_2, \cdots, a_{\varphi(m)},$$

它们所组成的集合叫作模 m 的一个缩剩余系，简称为缩系.

例如，模 6 的缩系为 1 和 5. 当 $m=p$ 为素数时，显然有 $\varphi(p)=p-1$，并且 1，2，…，$p-1$ 是模 p 的缩系. 比如，有 $\varphi(7)=6$，而 1，2，3，4，5，6 是模 7 的缩系.

根据缩系的定义，不难得出以下定理.

定理 2.2.15　若 $a_1, a_2, \cdots, a_{\varphi(m)}$ 是 $\varphi(m)$个与 m 互素的整数，则 $a_1, a_2, \cdots, a_{\varphi(m)}$ 是模 m 的一个缩系的充要条件是它们两两模 m 不同余.

定理 2.2.16　若 a 是满足$(a, m)=1$ 的整数，$a_1, a_2, \cdots, a_{\varphi(m)}$ 是模 m 的一个缩系，则 $aa_1, aa_2, \cdots, aa_{\varphi(m)}$ 也是模 m 的一个缩系. 即若 x 遍历模 m 的一个缩系，则 ax 也遍历模 m 的一个缩系.

证明　由于$(a, m)=1$ 且 $(a_i, m)=1(i=1, 2, \cdots, \varphi(m))$，故$(aa_i, m)=1(i=1, 2, \cdots, \varphi(m))$. 若存在 a_k 和 $a_l(k\neq l)$使得 $aa_k\equiv aa_l \pmod m$，由于$(a, m)=1$，可得 $a_k\equiv a_l \pmod m$，这与条件 a_k 和 a_l 来自于模 m 的一个缩系是矛盾的. 所以假设不成立，$aa_1, aa_2, \cdots, aa_{\varphi(m)}$ 两两模 m 不同余，且它们是 $\varphi(m)$个不同的整数. 于是，$aa_1, aa_2, \cdots, aa_{\varphi(m)}$ 是模 m 的一个缩系.

利用这个定理，还可以得到其他的一些结论.

定理 2.2.17　若 a 是满足$(a, m)=1$ 的整数，则存在整数 c，$1\leqslant c<m$ 且$(c, m)=1$，使得

$$ac\equiv 1 \pmod m.$$

证明　因为$(a, m)=1$，由定理 2.2.16，当 x 遍历模 m 的非负最小剩余系中的缩系时，ax 也遍历模 m 的一个缩系. 而数 1 是非负最小剩余系中的缩系中的一个元素，于是，存在整数 c，$1\leqslant c<m$，使得 ac 和数 1 在同一个剩余类中，即 $ac\equiv 1 \pmod m$. 到这里，我们已经知道缩系里存在这样的整数 c. 由 $ac\equiv 1 \pmod m$进一步得到，存在整数 k，使得 $ac-1=km$，因此 $ac+(-k)m=1$，这正是$(c, m)=1$ 的充要条件，得证.

定理 2.2.18　设 m 是大于 1 的整数，若 a 是满足$(a, m)=1$ 的整数，则

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

证明　设 $r_1, r_2, \cdots, r_{\varphi(m)}$ 是模 m 的一个缩系，则由定理 2.2.16 可知 $ar_1, ar_2, \cdots, ar_{\varphi(m)}$ 也是模 m 的一个缩系，所以对于第一个缩系的每一个元素，都在第二个缩系中存在唯一的元素与之在同一个剩余系中，就是按对具有同余关系，所以

$$(ar_1)(ar_2)\cdots(ar_{\varphi(m)})\equiv r_1r_2\cdots r_{\varphi(m)} \pmod m,$$

即

$$a^{\varphi(m)}r_1r_2\cdots r_{\varphi(m)}\equiv r_1r_2\cdots r_{\varphi(m)} \pmod m.$$

由于

$$(r_i, m)=1, \quad i=1, 2, \cdots, \varphi(m),$$

故

$$(r_1 r_2 \cdots r_{\varphi(m)}, m)=1.$$

于是，根据定理 2.2.6 可得

$$a^{\varphi(m)} \equiv 1 (\mathrm{mod}\ m).$$

定理 2.2.18 又称作**欧拉定理**，通过这个定理可推出著名的**费马小定理**，即定理 2.2.19.

定理 2.2.19 若 p 是素数，则对任意整数 a，有

$$a^p \equiv a (\mathrm{mod}\ p).$$

证明 若 a 不能被 p 整除，即$(a, p)=1$，由欧拉定理，有

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

两端同乘 a 即得

$$a^p \equiv a (\mathrm{mod}\ p).$$

若 a 能被 p 整除，则

$$a \equiv 0 \ (\mathrm{mod}\ p), \quad a^p \equiv 0 \ (\mathrm{mod}\ p),$$

于是

$$a^p \equiv a (\mathrm{mod}\ p).$$

定理得证.

定理 2.2.20 设 m_1，m_2 为互素的两个正整数，若 x_1，x_2 分别遍历模 m_1 和模 m_2 的缩系，则 $m_2 x_1 + m_1 x_2$ 遍历模 $m_1 m_2$ 的缩系.

证明 由$(m_1, m_2)=1$，$(x_1, m_1)=1$，$(x_2, m_2)=1$，可知$(m_2 x_1, m_1)=1$，进而

$$(m_2 x_1 + m_1 x_2, m_1)=1.$$

同理，

$$(m_2 x_1 + m_1 x_2, m_2)=1.$$

于是，有

$$(m_2 x_1 + m_1 x_2, m_1 m_2)=1.$$

下面证明凡是与 $m_1 m_2$ 互素的数 a，必有

$$a \equiv m_2 x_1 + m_1 x_2 (\mathrm{mod}\ m_1 m_2), (x_1, m_1)=1, (x_2, m_2)=1.$$

由定理 2.2.14 可知有 x_1，x_2 使 $a \equiv m_2 x_1 + m_1 x_2 (\mathrm{mod}\ m_1 m_2)$，故只需证明当$(a, m_1 m_2)=1$时，$(x_1, m_1)=(x_2, m_2)=1$. 假设$(x_1, m_1)>1$，则存在素数 p，使 $p \mid x_1$，$p \mid m_1$，又因为

$$a \equiv m_2 x_1 + m_1 x_2 (\mathrm{mod}\ m_1 m_2),$$

于是 $p \mid a$，故$(a, m_1 m_2)>1$，推出了矛盾. 所以$(x_1, m_1)=1$，同理可证$(x_2, m_2)=1$.

最后，由定理 2.2.14 可知，所有的 $m_2 x_1 + m_1 x_2$ 两两模 $m_1 m_2$ 不同余. 于是定理得证.

由定理 2.2.20，可推出以下定理，它反映了欧拉函数 $\varphi(m)$的性质，即 $\varphi(m)$为一积性函数.

定理 2.2.21 设 m_1，m_2 为互素的两个正整数，则

$$\varphi(m_1 m_2)=\varphi(m_1)\varphi(m_2).$$

证明 当 x_1 遍历模 m_1 的缩系时，其遍历的整数个数为 $\varphi(m_1)$. 当 x_2 遍历模 m_2 的缩系时，其遍历的整数个数为 $\varphi(m_2)$. 由定理 2.2.20，$m_2 x_1 + m_1 x_2$ 遍历模 $m_1 m_2$ 的缩系，其遍历的整数个数为 $\varphi(m_1)\varphi(m_2)$. 又因为模 $m_1 m_2$ 的缩系的代表元个数为 $\varphi(m_1 m_2)$，所以

$$\varphi(m_1 m_2)=\varphi(m_1)\varphi(m_2).$$

定理得证.

以上定理很大程度地简化了求解欧拉函数值的过程. 例如，如果求 $\varphi(55)$的值，以前需要列出所有小于 55 且与 55 互素的正整数，而利用定理 2.2.21，有

$$\varphi(55)=\varphi(5)\varphi(11)=4\times10=40.$$

定理 2.2.22　设 m 有标准分解式

$$m=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},\ \alpha_i>0,\ i=1,2,\cdots,s,$$

则

$$\varphi(m)=m\prod_{i=1}^{s}\left(1-\frac{1}{p_i}\right).$$

证明　由 $\varphi(m)$的定义可知，$\varphi(p^\alpha)$等于从 p^α 减去在 $1,2,\cdots,p^\alpha$ 中与 p 不互素的数的个数. 又由于 p 是素数，故 $\varphi(p^\alpha)$等于从 p^α 减去在 $1,2,\cdots,p^\alpha$ 中被 p 整除的数的个数. 在

$$1,2,\cdots,p,p+1,\cdots,2p,\cdots,p^{\alpha-1}\cdot p$$

中，被 p 整除的数共有 $p^{\alpha-1}$个，故 $\varphi(p^\alpha)=p^\alpha-p^{\alpha-1}$. 由此，有

$$\varphi(m)=\prod_{i=1}^{s}\varphi(p_i^{\alpha_i})=\prod_{i=1}^{s}(p_i^{\alpha_i}-p_i^{\alpha_i-1})=\prod_{i=1}^{s}p_i^{\alpha_i}\left(1-\frac{1}{p_i}\right)=m\prod_{i=1}^{s}\left(1-\frac{1}{p_i}\right).$$

[**例 2.2.5**]　设正整数 n 是两个不同素数的乘积，如果已知 n 和欧拉函数 $\varphi(n)$的值，则可求出 n 的因子分解式.

证明　设此两个不同的素因子为 p 和 q，由于

$$\varphi(n)=\varphi(pq)=\varphi(p)\varphi(q)=(p-1)(q-1)=pq-p-q+1,$$

有关于 p 和 q 的方程组：

$$\begin{cases}p+q=n+1-\varphi(n)\\ p\cdot q=n\end{cases}$$

于是，p 和 q 可由二次方程

$$x^2-(n+1-\varphi(n))x+n=0$$

求出.

2.2.3　线性同余方程

前面研究了同余的概念和一些性质，现在我们开始讨论在模 m 的情况下多项式方程的求解问题.

定义 2.2.7　设多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0,$$

其中 $n>0$，$a_i(i=0,1,\cdots,n)$是整数，又设 $m>0$，则同余式

$$f(x)\equiv0\pmod m \tag{2.2.2}$$

称为模 m 的**同余方程**. 若 a_n 不能被 m 整除，则 n 称为 $f(x)$的**次数**，记为 $\deg f(x)$.

若 x_0 满足

$$f(x_0)\equiv0\pmod m,$$

则

$$x\equiv x_0\pmod m$$

叫作同余方程(2.2.2)的**解**. 如果 $y_0\equiv x_0\pmod m$，那么必然有 $f(y_0)\equiv f(x_0)\equiv0\pmod$

m)，所以不同的解是指互不同余的解.

由定义可知，求解同余方程(2.2.2)，只要将 0，1，…，$m-1$ 逐个代入式(2.2.2)中进行验算即可，但当 m 较大时，巨大的计算量难以令人满意.

[例 2.2.6] 求解同余方程 $x^4+3x^2-2x+1\equiv 0 \pmod 5$.

解 求解此模 5 的 4 次同余方程，可将 0，1，2，3，4 逐个代入，由于

$$2^4+3\times 2^2-2\times 2+1=25\equiv 0 \pmod 5,$$

故 $x\equiv 2 \pmod 5$是该同余方程的解.

[例 2.2.7] 求解同余方程 $x^2+1\equiv 0 \pmod 7$.

解 这是一个模 7 的 2 次同余方程，由于将 0，1，…，6 逐个代入方程中均不满足，故此同余方程无解.

下面讨论线性同余方程(一次同余方程)的求解问题.

定理 2.2.23 设$(a, m)=1$，则同余方程

$$ax\equiv b \pmod m \tag{2.2.3}$$

有且仅有一个解 $x\equiv ba^{\varphi(m)-1} \pmod m$.

证明 由于 1，2，…，m 组成一个模 m 的完全剩余系，又$(a, m)=1$，故 a，$2a$，…，ma 也组成一个模 m 的完全剩余系. 所以，其中有且仅有一个数设为 aj，满足

$$aj\equiv b \pmod m,$$

于是 $x\equiv j \pmod m$就是式(2.2.3)的唯一解.

因为

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

所以，有

$$a^{\varphi(m)}b\equiv b \pmod m,$$

即

$$a\cdot a^{\varphi(m)-1}b\equiv b \pmod m,$$

故 $x\equiv a^{\varphi(m)-1}b \pmod m$是(2.2.3)的唯一解.

定理 2.2.24 设$(a, m)=d$，则同余方程(2.2.3)有解的充要条件是 $d\mid b$. 并且在式(2.2.3)有解时，它的解数为 d，以及若 $x\equiv x_0 \pmod m$是式(2.2.3)的特解，则它的 d 个解为

$$x\equiv x_0+\frac{m}{d}t \pmod m,$$

其中 $t=0$，…，$d-1$.

证明 当 $d=1$ 时，即为定理 2.2.23，故此处可假定 $d>1$.

先证必要性. 如果式(2.2.3)有解 $x\equiv x_0 \pmod m$，则有 $m\mid ax_0-b$，又 $d\mid m$，故 $d\mid ax_0-b$. 又因为 $d\mid a$，所以有 $d\mid b$.

再证充分性. 如果 $d\mid b$，则$\frac{b}{d}$为整数，又$(\frac{a}{d}, \frac{m}{d})=1$，根据定理 2.2.23，同余方程

$$\frac{a}{d}x\equiv\frac{b}{d} \pmod{\frac{m}{d}}$$

有唯一解，由定理 2.2.7，这个解必满足同余方程(2.2.3)，故式(2.2.3)有解.

若 $x\equiv x_0(\mathrm{mod}\ \frac{m}{d})$是同余方程

$$\frac{a}{d}x\equiv\frac{b}{d}\quad(\mathrm{mod}\ \frac{m}{d})$$

的唯一解，则有 d 个模 m 不同余的整数

$$x_0,\ x_0+\frac{m}{d},\ x_0+2\frac{m}{d},\ \cdots,\ x_0+(d-1)\frac{m}{d},$$

是式(2.2.3)的解. 由于

$$ax_0\equiv b\ (\mathrm{mod}\ m),$$

且显然有

$$at\frac{m}{d}\equiv 0(\mathrm{mod}\ m),\ t=0,\ \cdots,\ d-1,$$

故

$$a(x_0+t\frac{m}{d})\equiv b\ (\mathrm{mod}\ m),$$

于是 $x\equiv x_0+\frac{m}{d}t\ (\mathrm{mod}\ m)$是式(2.2.3)的解. 又

$$x_0,\ x_0+\frac{m}{d},\ x_0+2\frac{m}{d},\ \cdots,\ x_0+(d-1)\frac{m}{d}$$

两两模 m 不同余，且对于其他解，均可在以上这 d 个解中找到一数与之模 m 同余，即式(2.2.3)只有 d 个解. 证毕.

[例 2.2.8]　求解一次同余方程 $28x\equiv 21\ (\mathrm{mod}\ 35)$.

解　由于 $d=(28,\ 35)=7$，且显然 21 能被 7 整除，故此同余方程有解.

先求出同余方程

$$4x\equiv 3(\mathrm{mod}\ 5)$$

的解为 $x\equiv 2(\mathrm{mod}\ 5)$，所以原同余方程

$$28x\equiv 21(\mathrm{mod}\ 35)$$

的一个特解为 $x_0\equiv 2(\mathrm{mod}\ 35)$.

于是原同余方程的全部解为

$$x\equiv 2+5t\ (\mathrm{mod}\ 35),\ t=0,\ 1,\ \cdots,\ 4,\ 5,\ 6,$$

即 $x\equiv 2,\ 7,\ 12,\ 17,\ 22,\ 27,\ 32\ (\mathrm{mod}\ 35)$.

根据定理 2.2.24 可知，当$(a,\ m)=1$ 时，一次同余方程

$$ax\equiv 1(\mathrm{mod}\ m)$$

有唯一解 $x\equiv a'\ (\mathrm{mod}\ m)$. 因此，给出如下定义.

定义 2.2.8　对于正整数 m 和整数 a，满足$(a,\ m)=1$，则存在唯一一个剩余类，从中任意选择的元素整数 a'，都会使

$$aa'\equiv 1\ (\mathrm{mod}\ m)$$

成立，此时称 a'为 a 的模 m **逆元**，记作 $a^{-1}(\mathrm{mod}\ m)$.

利用这个定义，再由定理 2.2.24 可得如下推论.

推论　满足定理 2.2.24 条件的一次同余方程

$$ax\equiv b\ (\mathrm{mod}\ m)$$

的全部解为

$$x\equiv\frac{b}{d}\cdot\left(\left(\frac{a}{d}\right)^{-1}(\bmod\frac{m}{d})\right)+\frac{m}{d}t(\bmod m),$$

其中 $t=0,1,\cdots,d-1$.

这里有一个非常重要的问题需要解决，即如何求得 a 的逆元. 我们当然可以选择一个完全剩余系，然后尝试其中的每一个元素，看该元素与 a 相乘以后再除以 m 得到的余数是否等于 1. 另一种做法是直接计算 $a^{\varphi(m)-1}$ 的数值作为 a 的逆元. 当 m 的数值很小时，这样的方法可行而且可能计算得比较快，但是，当 m 的数值很大时，就不可能在合理的时间内完成这个任务. 最可行的方法就是利用辗转相除法. 因为 $(a,m)=1$，所以存在两个整数 p,n，使得

$$1=(a,m)=pa+nm,$$

整理后得到 $pa-1=-nm$，即 $m|(pa-1)$，也就是

$$pa\equiv1(\bmod m),$$

显然 p 是 a 的逆元. 因此，只要利用辗转相除法及其回代过程计算出整数 p 就得到了 a 的逆元 a^{-1}.

例如设 $a=97$，$m=1\ 001$，求 a 的逆元.

解 运用辗转相除法，有

$$\begin{aligned}1\ 001&=97\times10+31,\\97&=31\times3+4,\\31&=4\times7+3,\\4&=3\times1+1,\\3&=1\times3+0.\end{aligned}$$

因此，$(a,m)=1$ 得到验证. 再由

$$\begin{aligned}1&=4-3\times1 && \text{初始步骤}\\&=4-(31-4\times7)\times1 && \text{回代步骤}\\&=4\times8-31 && \text{整理步骤}\\&=(97-31\times3)\times8-31 && \text{回代步骤}\\&=97\times8-31\times25 && \text{整理步骤}\\&=97\times8-(1\ 001-97\times10)\times25 && \text{回代步骤}\\&=97\times258-1\ 001\times25 && \text{整理步骤}\\&=97\times258+(-25)\times1\ 001 && \text{规范步骤}\end{aligned}$$

因此，整数 $p=258$ 是 a 的逆元. 验算一下，$pa=258\times97=25\ 026\equiv1\ (\bmod\ 1\ 001)$的确成立. $a^{-1}=258$，当然 a^{-1}的取值也可以是 $258+1\ 001$ 或者 $258-1\ 001$，等等. 求一个大整数的逆元是著名的密码学算法 RSA 的主要计算步骤之一，以一个大整数的欧拉函数作为模数，从已经选择的加密密钥开始，求得其逆元作为解密密钥.

2.2.4 孙子定理与同余方程组

我国古代的一部优秀数学著作《孙子算经》中，有一类叫作“物不知数”的问题，原文如下：

“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?”

这个问题可以表达如下：现有一未知数，被 3 除余 2，被 5 除余 3，被 7 除余 2，求此未知数. 我国明代数学家程大位(字汝思，号宾渠，1533—1606)在《算法统宗》这部著作中，把解法用一首优美的诗来总结：

三人同行七十稀,五树梅花廿一枝，
七子团圆整半月,除百零五便得知.

这首诗的意思是，将此未知数被 3 除所得的余数乘 70，被 5 除所得的余数乘 21，被 7 除所得的余数乘 15，再将它们求和，将和除以 105，得到的余数即为所求未知数. 于是，以上“物不知数”问题可求解如下：

$$2\times 70+3\times 21+2\times 15=233,$$

将 233 除以 105，余数 23 即为所求.

这个问题为什么可以这样求解？这是不是一种巧合？在这个问题中，我们遇到的是 3 除，5 除，7 除，如果用其他的数代替 3，5，7，能否有同样类似的解法？著名的“孙子定理”就是用来解决这类问题的.

这其实就是一个求一次同余方程组的问题，此同余方程组表示如下，注意其中每一行的模数各不相同而且两两互素：

$$\begin{cases} x\equiv 2(\bmod 3)\\ x\equiv 3(\bmod 5)\\ x\equiv 2(\bmod 7) \end{cases}$$

我们先来直观的看一下这个问题的解法，这个问题看上去是不好解的，但是如果换一个类似的问题，就会感觉好解了，如下：

$$\begin{cases} x\equiv 0(\bmod 3)\\ x\equiv 0(\bmod 5)\\ x\equiv 0(\bmod 7) \end{cases}$$

由同余的概念马上可知道，$3\mid x$，$5\mid x$，$7\mid x$，所以 $3\times 5\times 7\mid x$，即 $105\mid x$，因此方程组的解必为

$$x\equiv 0(\bmod 105).$$

再换一个稍微复杂的问题

$$\begin{cases} a\equiv 1(\bmod 3)\\ a\equiv 0(\bmod 5)\\ a\equiv 0(\bmod 7) \end{cases}$$

类似于上面问题的思考思路，由第二和第三式知道 $5\times 7\mid a$，即 $35\mid a$，也就是 a 为 35 的倍数，那么接下来要看 35 的倍数中哪些除以 3 余 1，也就是看 35 的倍数中哪些具有如下的性质

$$35\times n\equiv 1(\bmod 3),$$

很明显 35 与 n 互相为模 3 的逆元，35 本身不行，但是 70 就行了(注意这个时候 $n=2$)，从而 70+105 的倍数也行，所以方程组的解必为

$$a\equiv 70(\bmod 105).$$

同理，我们对方程组

$$\begin{cases} b\equiv 0 \pmod 3 \\ b\equiv 1 \pmod 5 \\ b\equiv 0 \pmod 7 \end{cases},$$

得到解为

$$b\equiv 21 \pmod{105}.$$

对方程组

$$\begin{cases} c\equiv 0 \pmod 3 \\ c\equiv 0 \pmod 5 \\ c\equiv 1 \pmod 7 \end{cases},$$

得到解为

$$c\equiv 15 \pmod{105}.$$

另外，我们很容易观察到：

$$\begin{cases} 2a\equiv 2 \pmod 3 \\ 2a\equiv 0 \pmod 5 \\ 2a\equiv 0 \pmod 7 \end{cases},$$

$$\begin{cases} 3b\equiv 0 \pmod 3 \\ 3b\equiv 3 \pmod 5 \\ 3b\equiv 0 \pmod 7 \end{cases}$$

和

$$\begin{cases} 2c\equiv 0 \pmod 3 \\ 2c\equiv 0 \pmod 5 \\ 2c\equiv 2 \pmod 7 \end{cases}.$$

所以，原来方程的解必为

$$x\equiv 2a+3b+2c \pmod{105}.$$

前面提及的实际数值解答为

$$x\equiv 2\times 70+3\times 21+2\times 15=233\equiv 23 \pmod{105}.$$

将此问题推广，可给出下面定理.

定理 2.2.25 设 m_1，m_2，…，m_k 是 k 个两两互素的正整数，若令

$$m=m_1m_2\cdots m_k,\ m=m_iM_i,\ i=1,2,\cdots,k,$$

则对任意的整数 b_1，$b_2\cdots$，b_k，同余方程组

$$\begin{cases} x\equiv b_1 \pmod{m_1} \\ x\equiv b_2 \pmod{m_2} \\ \quad\vdots \\ x\equiv b_k \pmod{m_k} \end{cases} \tag{2.2.4}$$

有唯一解

$$x\equiv M_1'M_1b_1+M_2'M_2b_2+\cdots+M_k'M_kb_k \pmod m, \tag{2.2.5}$$

其中

$$M_i'M_i\equiv 1 \pmod{m_i},\ i=1,2,\cdots,k.$$

证明 由于对任意给定的 i 和 j，若满足 $1\leqslant i,j\leqslant k$ 且 $i\neq j$，则有

$$(m_i, m_j)=1,$$

故

$$(m_i, M_i)=1.$$

于是对每一 M_i，存在一个唯一的 M_i'，$i=1, 2, \cdots, k$，使得

$$M_i'M_i\equiv 1 \pmod{m_i},\ i=1, 2, \cdots, k.$$

又由 $m=m_iM_i$，得 $m_i \mid M_j$，$i\neq j$，因此

$$M_1'M_1b_1+M_2'M_2b_2+\cdots+M_k'M_kb_k\equiv M_i'M_ib_i\equiv b_i \pmod{m_i},\ i=1, 2, \cdots, k,$$

即式(2.2.5)是同余方程组(2.2.4)的解.

再证明这个解的唯一性. 设 x_1，x_2 是满足同余方程组(2.2.4)的任意两个整数，则

$$x_1\equiv x_2\equiv b_i \pmod{m_i},\ i=1, 2, \cdots, k.$$

因为 m_1，m_2，…，m_k 是 k 个两两互素的正整数，由定理 2.2.9 可知

$$x_1\equiv x_2 \pmod{m},$$

即解是唯一的.

这个定理就是著名的**孙子定理**，也称**中国剩余定理**.

[例 2.2.9]　求解同余方程组

$$\begin{cases} x\equiv 1 \pmod 3 \\ x\equiv 2 \pmod 5 \\ x\equiv 4 \pmod 7 \\ x\equiv 6 \pmod{13} \end{cases}$$

解　利用定理 2.2.25，其中 $m_1=3$，$m_2=5$，$m_3=7$，$m_4=13$. 令 $m=m_1m_2m_3m_4=1\ 365$，则

$$M_1=m_2m_3m_4=455,\quad M_2=m_1m_3m_4=273,$$
$$M_3=m_1m_2m_4=195,\quad M_4=m_1m_2m_3=105,$$

分别求解同余方程

$$M_i'M_i\equiv 1 \pmod{m_i},\ i=1, 2, 3, 4,$$

得

$$M_1'=2,\ M_2'=2,\ M_3'=6,\ M_4'=1,$$

故此同余方程组的解为

$$x\equiv 2\times 455\times 1+2\times 273\times 2+6\times 195\times 4+1\times 105\times 6\equiv 7\ 312\equiv 487 \pmod{1\ 365}.$$

定理 2.2.26　设 m_1，m_2，…，m_k 是 k 个两两互素的正整数，令

$$m=m_1m_2\cdots m_k,$$
$$m=m_iM_i,$$
$$M_i'M_i\equiv 1 \pmod{m_i},\ i=1, 2, \cdots, k,$$

若 b_1，b_2，…，b_k 分别遍历模 m_1，m_2，…，m_k 的完全剩余系，则

$$M_1'M_1b_1+M_2'M_2b_2+\cdots+M_k'M_kb_k$$

遍历模 m 的完全剩余系.

证明　令

$$x_0\equiv M_1'M_1b_1+M_2'M_2b_2+\cdots+M_k'M_kb_k,$$

则当 b_1，b_2，…，b_k 分别遍历模 m_1，m_2，…，m_k 的完全剩余系时，x_0 遍历 m 个整数. 下面证

明这 m 个整数两两模 m 不同余. 若

$$M_1'M_1b_1+\cdots+M_k'M_kb_k\equiv M_1'M_1b'_1+\cdots+M_k'M_kb_k'\pmod m,$$

其中 b_i 和 b_i' 在同一个模 m_i 的完全剩余系中取值，由于 $m_i|m$，$m_i|M_j$，$i\neq j$，故

$$M_i'M_ib_i\equiv M_i'M_ib_i'\pmod{m_i},\ i=1,2,\cdots,k,$$

又因为

$$M_i'M_i\equiv 1\pmod{m_i},\ i=1,2,\cdots,k,$$

所以

$$b_i\equiv b_i'\pmod{m_i},\ i=1,2,\cdots,k.$$

由于 b_i 和 b_i' 在同一个模 m_i 的完全剩余系中取值，故只能有

$$b_i=b_i',\ i=1,2,\cdots,k.$$

定理得证.

以上定理可以看作是定理 2.2.14 的推广.

定理 2.2.27 同余方程组

$$\begin{cases}x\equiv b_1\pmod{m_1}\\ x\equiv b_2\pmod{m_2}\end{cases}$$

有解的充要条件是 $(m_1,m_2)|b_1-b_2$. 如果上述条件成立，则同余方程组模 $[m_1,m_2]$ 有唯一解.

证明 设 $(m_1,m_2)=d$，先证必要性. 若 x_0 为同余方程组的解，则有

$$x_0\equiv b_1\pmod d,\quad x_0\equiv b_2\pmod d,$$

两式相减得 $b_1-b_2\equiv 0\pmod d$，因此 $d|b_1-b_2$.

再证充分性. 若 $d|b_1-b_2$，则因 $x\equiv b_1\pmod{m_1}$ 的解可写为

$$x=b_1+m_1y,$$

将其代入 $x\equiv b_2\pmod{m_2}$ 得

$$m_1y\equiv b_2-b_1\pmod{m_2}.$$

因为 $(m_1,m_2)=d$，$d|b_2-b_1$，故上式有解，即原同余方程组有解.

设原同余方程组有两个解分别为 x_1 和 x_2，则

$$x_1-x_2\equiv 0\pmod{m_1},\quad x_1-x_2\equiv 0\pmod{m_2},$$

于是有 $x_1-x_2\equiv 0\pmod{[m_1,m_2]}$，即同余方程组模 $[m_1,m_2]$ 有唯一解. 证毕.

通过以上述定理可知，对于一次同余方程组

$$\begin{cases}x\equiv b_1\pmod{m_1}\\ x\equiv b_2\pmod{m_2}\\ \quad\vdots\\ x\equiv b_k\pmod{m_k}\end{cases}$$

其中 $k\geqslant 3$，若 $(m_1,m_2)|b_1-b_2$，可先解前面两个方程得

$$x\equiv b'_2\pmod{[m_1,m_2]}.$$

若 $([m_1,m_2],m_3)|b_2'-b_3$，则可再与后面的 $x\equiv b_3\pmod{m_3}$ 联立解出

$$x\equiv b_3'\pmod{[m_1,m_2,m_3]}.$$

依此类推，最后可得唯一解

$$x\equiv b_k'\pmod{[m_1,m_2,\cdots,m_k]}.$$

如果中间有一步出现无解，则原同余方程组无解.

[例 2.2.10]　判断方程组

$$\begin{cases} x\equiv 11 \pmod{36} \\ x\equiv 7 \pmod{40} \\ x\equiv 32 \pmod{75} \end{cases}$$

是否有解.

解　$(36, 40)=4$，$(36, 75)=3$，$(40, 75)=5$.

$$b_1-b_2=11-7=4,$$
$$b_1-b_3=11-32=-21,$$
$$b_2-b_3=7-32=-25.$$

因此方程组肯定有解，因为方程组有解条件都满足，即 $4|4$，$3|-21$，$5|-25$. 且解的模数是 $[36, 40, 75]=1\ 800$. 这个方程的解为 $x\equiv 407 \pmod{1\ 800}$. 有兴趣的读者可以自行练习写出全部求解过程.

定理 2.2.28　设 $m_1, m_2, \cdots, m_k$ 是 k 个两两互素的正整数，令 $m=m_1m_2\cdots m_k$，则同余方程

$$f(x)\equiv 0 \pmod{m} \tag{2.2.6}$$

与同余方程组

$$\begin{cases} f(x)\equiv 0 \pmod{m_1} \\ f(x)\equiv 0 \pmod{m_2} \\ \quad\vdots \\ f(x)\equiv 0 \pmod{m_k} \end{cases} \tag{2.2.7}$$

等价. 若用 T_i 表示同余方程

$$f(x)\equiv 0 \pmod{m_i}$$

的解数(即解的个数)，$i=1, 2, \cdots, k$，用 T 表示同余方程(2.2.6)的解数，则

$$T=T_1T_2\cdots T_k.$$

证明　设 x_0 为同余方程(2.2.6)的解，则

$$f(x_0)\equiv 0 \pmod{m}.$$

由定理 2.2.8 可知

$$f(x_0)\equiv 0 \pmod{m_i},\ i=1, 2, \cdots, k,$$

即 x_0 亦为同余方程组(2.2.7)的解.

若 x_0 为同余方程组(2.2.7)的解，即

$$f(x_0)\equiv 0 \pmod{m_i},\ i=1, 2, \cdots, k.$$

由定理 2.2.9 可知

$$f(x_0)\equiv 0 \pmod{m},$$

即 x_0 亦为同余方程(2.2.6)的解.

设同余方程 $f(x)\equiv 0 \pmod{m_i}$ 的解为 b_i，$i=1, 2, \cdots, k$. 由孙子定理可知同余方程组

$$\begin{cases} x\equiv b_1 \pmod{m_1} \\ x\equiv b_2 \pmod{m_2} \\ \quad\vdots \\ x\equiv b_k \pmod{m_k} \end{cases}$$

的解为

$$x\equiv M_1'M_1b_1+M_2'M_2b_2+\cdots+M_k'M_kb_k\pmod m.$$

由于

$$f(x)\equiv f(b_i)\equiv 0\pmod{m_i},\ i=1,2,\cdots,k,$$

故 x 亦为同余方程(2.2.6)的解. 于是当 b_i 遍历 $f(x)\equiv 0\pmod{m_i}$ 的所有解时，x 遍历同余方程(2.2.6)的所有解. 于是，有 $T=T_1T_2\cdots T_k$.

[例 2.2.11] 求解同余方程

$$x^4+2x^3+8x+9\equiv 0\pmod{35}.$$

解 设 $f(x)=x^4+2x^3+8x+9$，由定理 2.2.28 知同余方程 $f(x)\equiv 0\pmod{35}$ 等价于同余方程组

$$\begin{cases}f(x)\equiv 0\pmod 5\\ f(x)\equiv 0\pmod 7\end{cases}$$

用直接验算的方法容易得到 $f(x)\equiv 0\pmod 5$ 的解为

$$x\equiv 1,4\pmod 5,$$

$f(x)\equiv 0\pmod 7$ 的解为

$$x\equiv 3,5,6\pmod 7.$$

由孙子定理，可求出同余方程组

$$\begin{cases}x\equiv b_1\pmod 5\\ x\equiv b_2\pmod 7\end{cases}$$

当 (b_1,b_2) 分别取(1，3)，(1，5)，(1，6)，(4，3)，(4，5)，(4，6)时的解为

$$x\equiv 21b_1+15b_2\equiv 31,26,6,24,19,34\pmod{35}.$$

这 6 个解即为原同余方程的解.

这个定理也可以使我们能够利用孙子定理来解单个的具有较大模数的线性同余方程，这种方法可能计算上更有效率.

[例 2.2.12] 求解 $13x\equiv 71\pmod{380}$.

解 因为 $380=4\times5\times19$，所以它等价于如下方程组

$$\begin{cases}13x\equiv 71\pmod 4\\ 13x\equiv 71\pmod 5\\ 13x\equiv 71\pmod{19}\end{cases}$$

$$\Rightarrow\begin{cases}(4+4+4+1)x\equiv 71\pmod 4\\ (5+5+3)x\equiv 71\pmod 5\\ 13x\equiv 71\pmod{19}\end{cases}$$

$$\Rightarrow\begin{cases}x\equiv 71\pmod 4\\ 3x\equiv 71\pmod 5\\ 13x\equiv 71\pmod{19}\end{cases}$$

$$\Rightarrow\begin{cases}x\equiv 3\pmod 4\\ 3x\equiv 1\pmod 5\\ 13x\equiv 14\pmod{19}\end{cases}.$$

利用单同余方程式的解法可得到

$$\begin{cases} x \equiv 3 \pmod 4 \\ x \equiv 2 \pmod 5 \\ x \equiv 4 \pmod{19} \end{cases}.$$

接着应用孙子定理求解即可，最后得到的解为

$$x \equiv 327 \pmod{380}.$$

前面讨论的同余方程组问题中，我们注意到方程组中的每一行的模数都不相同，而且只有一个待解的未知元. 还有另一类重要的多元线性同余方程组问题，不同之处在于这类问题中的模数都相同，而且具有两个或者两个以上的未知元. 这样的问题与我们在线性代数中学过的关于实数和复数的方程组问题非常相像，而且可以使用很多线性代数中的向量和矩阵的表示及运算方法. 因此，下面主要通过实例来加深读者对此的理解.

[例 2.2.13] 在古典的 Hill 密码中，如果按对加密，则每一对明文组成的行向量用$[x_1, x_2]$来表示，加密后的密文对形成的行向量用$[y_1, y_2]$来表示，y_1，y_2 是由 x_1，x_2 的线性组合计算而来

$$\begin{cases} y_1 \equiv 11x_1 + 3x_2 \pmod{26} \\ y_2 \equiv 8x_1 + 7x_2 \pmod{26} \end{cases},$$

使用矩阵表达为

$$[y_1, y_2] \equiv [x_1, x_2] \begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix} \pmod{26}.$$

其中的 2 乘 2 阶矩阵被称作密钥，那么如何解密呢，即如何由(y_1, y_2)来计算得到(x_1, x_2)呢？实际上，我们可以采用消元方法来解，先消去未知元 x_2 解得 x_1，然后同样的方法，先消去未知元 x_1 解得 x_2. 还可以利用逆矩阵的方法，即

$$[x_1, x_2] \equiv [y_1, y_2] \begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix}^{-1} \pmod{26},$$

其中

$$\begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix}^{-1} = \begin{bmatrix} 7 & 18 \\ 23 & 11 \end{bmatrix}.$$

可以验证一下这个逆矩阵的正确性，如下：

$$\begin{aligned} &\begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix} \begin{bmatrix} 7 & 18 \\ 23 & 11 \end{bmatrix} \\ =&\begin{bmatrix} 11\times 7 + 8\times 23 & 11\times 18 + 8\times 11 \\ 3\times 7 + 7\times 23 & 3\times 18 + 7\times 11 \end{bmatrix} \\ =&\begin{bmatrix} 261 & 286 \\ 182 & 131 \end{bmatrix} \\ \equiv&\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \pmod{26} \end{aligned}$$

两者的乘积是单位矩阵，说明它们互为逆矩阵.

如果明文是$[x_1, x_2] = [9, 20]$，计算过程如下：

$$[9, 20] \begin{bmatrix} 11 & 8 \\ 3 & 7 \end{bmatrix} = [99 + 60, 72 + 140] \equiv [3, 4] \pmod{26}.$$

则密文为[3，4]. 反过来，接收方收到密文[3，4]后，希望恢复明文，计算过程如下：

$$[3,4]\begin{bmatrix}7 & 18\\ 23 & 11\end{bmatrix}=[21+92,54+44]\equiv[9,20](\mathrm{mod}\ 26),$$

的确正确地恢复了明文[9，20].

那么，在模 26 运算下，如何判断矩阵是否可逆？又如何计算可逆矩阵的逆矩阵呢？下面我们不加证明地给出有关定理.

定理 2.2.29 矩阵 $\boldsymbol{K}$ 在模 26 运算下存在可逆矩阵的充分必要条件是[det $\boldsymbol{K}$，26]=1 (det $\boldsymbol{K}$ 表示矩阵 $\boldsymbol{K}$ 的行列式的值).

定理 2.2.30 如果二阶矩阵

$$\boldsymbol{K}=\begin{bmatrix}k_{11} & k_{12}\\ k_{21} & k_{22}\end{bmatrix}$$

可逆，则其逆矩阵为

$$\boldsymbol{K}^{-1}=(\det \boldsymbol{K})^{-1}\begin{bmatrix}k_{22} & -k_{12}\\ -k_{21} & k_{11}\end{bmatrix}(\mathrm{mod}\ 26).$$

2.2.5 高次同余方程

我们知道，任一大于 1 的整数 m 均有标准分解式

$$m=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},\ \alpha_i>0,\ i=1,2,\cdots,s,$$

其中 $p_i<p_j(i<j)$是素数. 于是，由定理 2.2.28 可知，欲解 $f(x)\equiv 0\ (\mathrm{mod}\ m)$，只需求解同余方程组

$$\begin{cases}f(x)\equiv 0\ (\mathrm{mod}\ p_1^{\alpha_1})\\ f(x)\equiv 0\ (\mathrm{mod}\ p_2^{\alpha_2})\\ \quad\vdots\\ f(x)\equiv 0\ (\mathrm{mod}\ p_s^{\alpha_s})\end{cases}$$

所以，先来讨论 p 为素数时，同余方程

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\equiv 0(\mathrm{mod}\ p^{\alpha}) \tag{2.2.8}$$

的求解方法，其中 α 为正整数，且 a_n 不能被 p^{α} 整除.

定理 2.2.31 设 $x\equiv x_1(\mathrm{mod}\ p)$是同余方程

$$f(x)\equiv 0\ (\mathrm{mod}\ p) \tag{2.2.9}$$

的一个解，且满足$(f'(x_1),\ p)=1$，则同余方程(2.2.8)有解

$$x\equiv x_{\alpha}\ (\mathrm{mod}\ p^{\alpha}).$$

其中 x_{α} 由以下关系式递归得到：

$$\begin{cases}x_i\equiv x_{i-1}+p^{i-1}t_{i-1} & (\mathrm{mod}\ p^i)\\ t_{i-1}\equiv -\dfrac{f(x_{i-1})}{p^{i-1}}((f'(x_1))^{-1}\ \ (\mathrm{mod}\ p))\ (\mathrm{mod}\ p)\end{cases}$$

$i=2,3,\cdots,\alpha$. 这里，$f'(x)=\sum\limits_{i=1}^{n}ia_ix^{i-1}$ 表示 $f(x)$ 的导函数.

证明 用数学归纳法.

(1) 当 $\alpha=2$ 时，根据假设条件，同余方程(2.2.9)的所有解为

$$x=x_1+pt_1,\quad t_1=0,\pm1,\pm2,\cdots.$$

于是，考虑关于 t_1 的同余方程

$$f(x_1+pt_1)\equiv0\pmod{p^2}.$$

由泰勒公式，有

$$f(x_1)+pt_1f'(x_1)\equiv0\pmod{p^2},$$

又因为 $f(x_1)\equiv0\pmod{p}$，所以上述同余方程可写为

$$t_1f'(x_1)\equiv-\frac{f(x_1)}{p}\pmod{p}.$$

由 $(f'(x_1),p)=1$，根据定理 2.2.24 及其推论，此同余方程的唯一解为

$$t_1\equiv-\frac{f(x_1)}{p}((f'(x_1))^{-1}(\bmod p))\pmod{p}.$$

故

$$x\equiv x_2\equiv x_1+pt_1(\bmod p^2)$$

是同余方程 $f(x)\equiv0\pmod{p^2}$的解.

(2) 当 $\alpha\geqslant3$ 时，假设对 $i-1(3\leqslant i\leqslant\alpha)$成立，即同余方程

$$f(x)\equiv0\pmod{p^{i-1}}$$

有解

$$x=x_{i-1}+p^{i-1}t_{i-1},\quad t_{i-1}=0,\pm1,\pm2,\cdots.$$

于是，考虑关于 t_{i-1}的同余方程

$$f(x_{i-1}+p^{i-1}t_{i-1})\equiv0\pmod{p^i}.$$

由泰勒公式及 $p^{2(i-1)}\geqslant p^i$，可知

$$f(x_{i-1})+p^{i-1}t_{i-1}f'(x_{i-1})\equiv0\pmod{p^i},$$

因为 $f(x_{i-1})\equiv0\pmod{p^{i-1}}$，所以上述同余方程可写为

$$t_{i-1}f'(x_{i-1})\equiv-\frac{f(x_{i-1})}{p^{i-1}}\pmod{p}.$$

又 $f'(x_{i-1})\equiv f'(x_{i-2})\equiv\cdots\equiv f'(x_1)\pmod{p}$，进而有

$$(f'(x_{i-1}),p)=\cdots=(f'(x_1),p)=1,$$

再根据定理 2.2.24 及其推论，此同余方程的唯一解为

$$\begin{aligned}t_{i-1}&\equiv-\frac{f(x_{i-1})}{p^{i-1}}((f'(x_{i-1}))^{-1}\pmod{p})\\&\equiv-\frac{f(x_{i-1})}{p^{i-1}}((f'(x_1))^{-1}\pmod{p})\pmod{p}\end{aligned}$$

故

$$x\equiv x_i\equiv x_{i-1}+p^{i-1}t_{i-1}\pmod{p^i}$$

是同余方程 $f(x)\equiv0\pmod{p^i}$的解.

于是，根据数学归纳法，定理得证.

[例 2.2.14]　求解同余方程

$$f(x)=x^4+7x+4\equiv0\pmod{27}.$$

解　写出 $f(x)$的导函数，即

$$f'(x)=4x^3+7.$$

通过直接验算，可知同余方程

$$f(x)\equiv 0 \pmod 3$$

有一解

$$x_1\equiv 1 \pmod 3.$$

于是，有

$$f'(x_1)\equiv -1 \pmod 3,$$

进而

$$(f'(x_1))^{-1}\equiv -1 \pmod 3.$$

依次计算如下：

$$\begin{cases} t_1\equiv -\dfrac{f(x_1)}{3}((f'(x_1))^{-1} \pmod 3)\equiv 1 & \pmod 3 \\ x_2\equiv x_1+3t_1\equiv 4 & \pmod 9 \end{cases}$$

$$\begin{cases} t_2\equiv -\dfrac{f(x_2)}{3^2}((f'(x_1))^{-1} \pmod 3)\equiv 2 & \pmod 3 \\ x_3\equiv x_2+3^2t_2\equiv 22 & \pmod{27} \end{cases}$$

所以，原同余方程的解为

$$x_3\equiv 22 \pmod{27}.$$

现在重点讨论模 p 的同余方程

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\equiv 0 \pmod p \tag{2.2.10}$$

的求解方法，其中 a_n 不能被 p 整除.

在此之前，先引入多项式的辗转相除法，或称多项式的欧几里得除法.

定理 2.2.32 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

为 n 次整系数多项式，

$$g(x)=x^m+b_{m-1}x^{n-1}+\cdots+b_1x+b_0$$

为 m 次首一(最高项系数为 1)整系数多项式，其中 $m\geqslant 1$，则存在整系数多项式 $q(x)$ 和 $r(x)$ 使得

$$f(x)=g(x)q(x)+r(x),$$

其中 $\deg(r(x))<\deg(g(x))$.

证明 可分两种情况讨论.

(1) 若 $n<m$，可取 $q(x)=0$，$r(x)=f(x)$ 使结论成立.

(2) 若 $n\geqslant m$，可对 $f(x)$ 的次数 n 作数学归纳法.

当 $n=m$ 时，有

$$f(x)-a_ng(x)=(a_{n-1}-a_nb_{m-1})x^{n-1}+\cdots+(a_1-a_nb_0)x+a_0,$$

因此，取 $q(x)=a_n$，$r(x)=f(x)-a_ng(x)$ 可使结论成立.

假设当 $n=k-1$ 时，结论成立，其中 $k-1\geqslant m$.

当 $n=k$ 时，则有

$$f(x)-a_nx^{n-m}g(x)=(a_{n-1}-a_nb_{m-1})x^{n-1}+\cdots+(a_{n-m}-a_nb_0)x^{n-m}+a_{n-m-1}x^{n-m-1}+\cdots+a_0.$$

显然 $f(x)-a_nx^{n-m}g(x)$ 是次数小于等于 $n-1$ 的多项式，对其运用归纳假设或情况(1)，可

知存在整系数多项式 $q_1(x)$和 $r_1(x)$使得

$$f(x)-a_n x^{n-m}g(x)=g(x)q_1(x)+r_1(x),$$

其中 $\deg(r_1(x))<\deg(g(x))$. 因此，取 $q(x)=a_n x^{n-m}+q_1(x)$，$r(x)=r_1(x)$可使结论成立.

根据数学归纳法原理，可知结论成立，于是定理得证.

定理 2.2.33　同余方程(2.2.10)与一个次数小于 p 的模 p 的同余方程等价.

证明　由定理 2.2.32 可知，存在整系数多项式 $q(x)$和 $r(x)$使得

$$f(x)=(x^p-x)q(x)+r(x),$$

其中 $\deg r(x)<p$. 根据费马小定理，对任意整数 x 都有

$$x^p-x\equiv 0 \pmod p.$$

于是同余方程

$$f(x)\equiv 0 \pmod p$$

等价于同余方程

$$r(x)\equiv 0 \pmod p.$$

定理 2.2.34　同余方程(2.2.10)最多有 n 个解.

证明　可对 $f(x)$的次数 n 作数学归纳法.

当 $n=1$ 时，一次同余方程为

$$a_1 x+a_0\equiv 0 \pmod p,$$

由于 a_1 不能被 p 整除，即$(a_1, p)=1$，故同余方程恰有一个解，结论成立.

假设定理对次数为 $n-1$ $(n\geqslant 2)$的同余方程成立，即次数为 $n-1$ 的同余方程最多有 $n-1$ 个解. 下面证明同余方程(2.2.10)最多有 n 个解.

根据定理 2.2.33 可知，同余方程(2.2.10)与一个次数小于 p 的模 p 的同余方程等价，所以不妨设 $n\leqslant p-1$. 用反证法，假设同余方程(2.2.10) 有 $n+1$ 个解，设它们为

$$x\equiv x_i \pmod p,\ i=0,1,\cdots,n.$$

由于

$$f(x)-f(x_0)=\sum_{k=1}^{n}a_k(x^k-x_0^k)=(x-x_0)g(x),$$

显然，$g(x)$是首项系数为 a_n 的 $n-1$ 次整系数多项式，根据归纳假设，可知

$$g(x)\equiv 0 \pmod p$$

是 $n-1$ 次同余方程，至多有 $n-1$ 个解. 而由于

$$f(x_k)-f(x_0)=(x_k-x_0)g(x_k)\equiv 0 \pmod p$$

当 $k>0$ 时，$x_k-x_0\equiv 0 \pmod p$不成立，故 $n-1$ 次同余方程 $g(x)\equiv 0 \pmod p$有 n 个解，推出了矛盾. 于是假设不成立，定理得证.

定理 2.2.34 通常被称为**拉格朗日(Lagrange)定理**.

定理 2.2.35　如果同余方程

$$f(x)=a_n x^n+a_{n-1}x^{n-1}+\cdots+a_1 x+a_0\equiv 0 \pmod p$$

的解的个数大于 n，则 $p\mid a_i$，$i=0,1,\cdots,n$.

证明　用反证法. 假设存在某些系数不能被 p 整除，若这些系数的下标最大的为 k，$k\leqslant n$，则原同余方程可写为

$$f(x)\equiv a_k x^k+a_{k-1}x^{k-1}+\cdots+a_1 x+a_0\equiv 0 \pmod p.$$

根据上面的定理可知，此同余方程最多有 k 个解，与所给条件矛盾，故假设不成立. 定理得证.

定理 2.2.36 如果同余方程(2.2.10)有 k 个不同的解

$$x\equiv x_i \pmod{p},\ i=1,2,\cdots,k,\ 1\leqslant k\leqslant n,$$

则对任意整数 x，均有

$$f(x)\equiv(x-x_1)(x-x_2)\cdots(x-x_k)f_k(x) \pmod{p},$$

其中 $f_k(x)$是首项系数为 a_n 的 $n-k$ 次多项式.

证明 由定理 2.2.32 可知，存在整系数多项式 $f_1(x)$和 $r(x)$使得

$$f(x)=(x-x_1)f_1(x)+r(x),\ \deg r(x)<\deg(x-x_1).$$

显然，$f_1(x)$是首项系数为 a_n 的 $n-1$ 次多项式. 由于 $\deg(x-x_1)=1$，故 $r(x)=r$ 为整数，又因为 $f(x_1)\equiv 0 \pmod{p}$，所以有 $r\equiv 0 \pmod{p}$，即

$$f(x)=(x-x_1)f_1(x) \pmod{p}.$$

又因为 $f(x_i)\equiv 0 \pmod{p}$，并且 x_i 与 x_1 模 p 不同余，其中 $i=2,3,\cdots,k$，于是可知

$$f_1(x_i)\equiv 0 \pmod{p},\ i=2,3,\cdots,k.$$

同理，对多项式 $f_1(x)$可找到多项式 $f_2(x)$使得

$$\begin{cases} f_1(x)\equiv(x-x_2)f_2(x) & \pmod{p} \\ f_2(x_i)\equiv 0 & \pmod{p} \end{cases}$$

其中 $i=3,4,\cdots,k$. 依此类推，可得

$$f_{k-1}(x)\equiv(x-x_k)f_k(x) \pmod{p}.$$

于是，有

$$f(x)\equiv(x-x_1)\cdots(x-x_k)f_k(x) \pmod{p},$$

定理得证.

定理 2.2.37 对于素数 p 与正整数 n，$n\leqslant p$，同余方程

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\equiv 0 \pmod{p}$$

有 n 个解的充要条件是 x^p-x 被 $f(x)$除所得余式的所有系数均能被 p 整除.

证明 由定理 2.2.32 可知，存在整系数多项式 $q(x)$和 $r(x)$，使得

$$x^p-x=f(x)q(x)+r(x),$$

其中 $r(x)$的次数小于 n，$q(x)$的次数为 $p-n$.

现在证明必要性. 若原同余方程有 n 个解，则根据费马小定理，这 n 个解都是

$$x^p-x\equiv 0 \pmod{p}$$

的解，显然这 n 个解也都是

$$r(x)\equiv 0 \pmod{p}$$

的解. 由于 $r(x)$的次数小于 n，故由定理 2.2.35 可知，$r(x)$的所有系数均能被 p 整除.

再来证明充分性. 若 $r(x)$的所有系数均能被 p 整除，则显然有

$$r(x)\equiv 0 \pmod{p}.$$

又由费马小定理，可知对任意整数有

$$x^p-x\equiv 0 \pmod{p}.$$

因此，对任意整数有

$$f(x)q(x)\equiv 0 \pmod{p},$$

即它有 p 个不同的解

$$x\equiv 0,\ 1,\ \cdots,\ p-1\ (\mathrm{mod}\ p).$$

假设 $f(x)\equiv 0\ (\mathrm{mod}\ p)$的解数小于 n，则 $q(x)\equiv 0\ (\mathrm{mod}\ p)$的解数小于等于 $p-n$，故

$$f(x)q(x)\equiv 0\ (\mathrm{mod}\ p)$$

的解数小于 p，推出了矛盾. 所以 $f(x)\equiv 0\ (\mathrm{mod}\ p)$的解数为 n. 证毕.

[例 2.2.15]　判断同余方程

$$2x^3+5x^2+6x+1\equiv 0\ (\mathrm{mod}\ 7)$$

是否有 3 个解.

解　先将多项式化为首项系数为 1. 由于 $4\times 2\equiv 1\ (\mathrm{mod}\ 7)$，故有

$$4(2x^3+5x^2+6x+1)\equiv x^3-x^2+3x-3\equiv 0\ (\mathrm{mod}\ 7).$$

根据多项式的辗转相除法，可得

$$x^7-x=x(x^3+x^2-2x-2)(x^3-x^2+3x-3)+7x(x^2-1).$$

由上面定理可知原同余方程有 3 个解.

习 题

1. 证明：若 a 是整数，则 a^3-a 能被 3 整除.
2. 设$(u,\ v)=1$，试证$(u+v,\ u-v)=1$ 或 2.
3. 对于 $n\geqslant 1$ 利用数学归纳法证明：$8\mid 5^{2n}+7$；$21\mid 4^{n+1}+5^{2n-1}$.
4. 试证连续两个整数的立方之差不可能被 2 整除.
5. 设$(u,\ v)=1$，试证$(u+v,\ u^2+v^2)=1$ 或 2.
6. 设 $A=\{d_1,\ d_2,\ \cdots,\ d_k\}$为非零整数 a 的全体因数的集合，证明 $B=\{a/d_1,\ a/d_2,\ \cdots,\ a/d_k\}$也是 a 的全体因数的集合.
7. 证明对于任意整数 m 和 n，证明
   ① $8\nmid m^2-n^2-2$；
   ② 若 $2\nmid mn$，则 m^2+n^2 不能表示为一个整数平方的形式；
   ③ 若 $3\nmid mn$，则 m^2+n^2 也不能表示为一个整数平方的形式；
   ④ 若 m^2+n^2 能够表示为一个整数平方的形式，则 $6\mid mn$.
8. 设

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0$$

   是整系数多项式，证明若 $d\mid b-c$，则 $d\mid f(b)-f(c)$.
9. 用 Eratosthenes 筛法求出 200 以内的所有素数.
10. 设 $a>2$ 是奇数，证明
    (1) 一定存在正整数 $d\leqslant a-1$，使得 $a\mid 2^d-1$；
    (2) 若 d_0 是满足(1)的最小正整数 d，那么 $a\mid 2^h-1$ 的充要条件是 $d_0\mid h$.
11. 设 a 是奇数，证明一定存在正整数 d 使 2^d-3 与 a 互素.
12. 求下列各数的最大公因子和最小公倍数：
    (1) 621，437　　(2) 289，377
    (3) 20 785，44 350　　(4) 20 041，37 516

(5) 108，144，264，420，780

13. 计算整数 x 和 y，使得 (56，72)＝56x＋72y；(1 769，2 378)＝1 769x＋2 378y.

14. 计算下列各数的正因子的个数，并求其所有正因子之和：

(1) 675　　(2) 4 704　　(3) 5 544

15. 将$\frac{7\ 700}{2\ 145}$表示成简单连分数.

16. 将$\sqrt{5}$表示成简单连分数.

17. (1) 求 $7^{2\,046}$ 写成十进制数时的个位数；

(2) 求 $2^{1\,000}$ 的十进制表示中的末尾两位数字.

18. 证明：正整数 n(十进制)能被 3 整除的充要条件是将 n 的各位数字相加所得之和能被 3 整除.

19. 证明 $641 \mid 2^{2^5}+1$.

20. 试证：如果 $u \equiv v \pmod{n}$，那么$(u, n)=(v, n)$.

21. 试求 $1^5+2^5+3^5+\cdots+99^5$ 之和被 4 除余几?

22. 计算 555^{555} 被 7 除的余数.

23. 分别用模 5 和模 6 的完全剩余系与缩系，表示模 30 的完全剩余系与缩系.

24. 证明对任意整数 m，有 $\sum\limits_{d \mid m,\ d>0} \varphi(d)=m$.

25. 计算 20 以内的正整数的欧拉函数值.

26. 利用同余理论为 N 值球队安排一个单循环比赛表.

27. 求解下列一次同余方程：

(1) $27x \equiv 12 \pmod{15}$

(2) $24x \equiv 6 \pmod{81}$

(3) $91x \equiv 26 \pmod{169}$

(4) $71x \equiv 32 \pmod{3\,441}$

28. 利用孙子定理求解下列同余方程组：

(1) $\begin{cases} x \equiv 9 \pmod{12} \\ x \equiv 6 \pmod{25} \end{cases}$；　(2) $\begin{cases} x \equiv 5 \pmod{7} \\ x \equiv 12 \pmod{15} \\ x \equiv 18 \pmod{22} \end{cases}$；　(3) $\begin{cases} x \equiv 2 \pmod{9} \\ 3x \equiv 4 \pmod{5} \\ 4x \equiv 3 \pmod{7} \end{cases}$

29. 求解下列同余方程组(注意不只一个解)：

$$\begin{cases} 2x \equiv 3 \pmod{5} \\ 4x \equiv 2 \pmod{6} \\ 3x \equiv 2 \pmod{7} \end{cases}$$

30. 如果$(a, 133)=(b, 133)=1$，试证明 $133 \mid a^{18}-b^{18}$.

31. 举例说明对模数为合数的情况，拉格朗日定理一般不成立.

32. 求模 11 的一组完全剩余系$\{r_1, r_2, \cdots, r_{11}\}$，使得

$r_i \equiv -1 \pmod{2}$；$r_i \equiv 1 \pmod{3}$；$r_i \equiv 0 \pmod{7}$；

$r_i \equiv 1 \pmod{5}$，$1 \leqslant i \leqslant 11$.

33. 证明 2，2^2，2^3，…，2^{18}是模 27 的一个缩系.

34. 求相邻的 4 个正整数，它们依次可被 2^2，3^2，5^2 及 7^2 整除.

35. 求解同余方程

(1) $3x^{14}+4x^{13}+2x^{11}+x^9+x^6+x^3+12x^2+x\equiv 0 \pmod 5$；

(2) $x^3+5x^2+9\equiv 0 \pmod{27}$.

36. 证明同余方程

$$2x^3-x^2+3x+11\equiv 0 \pmod 5$$

有 3 个解.

37. 利用转化成联立方程组的方法解

$$91x\equiv 419 \pmod{440}.$$

38. 如下各个方程有几个解？

$$x^2-1\equiv 0 \pmod{168};$$
$$x^2+1\equiv 0 \pmod{70};$$
$$x^2+x+1\equiv 0 \pmod{91};$$
$$x^3+1\equiv 0 \pmod{140}.$$

39. 求 3 在模 97 下的逆元.

40. 求 13 的倍数，使得该数被 3,5,7,11 除所得的余数为 2.

41. 证明不存在整数 n 使得它的欧拉函数值为 14.

42. 如果一个密码系统中，明文 x 被加密成密文 y，使得 $y\equiv 7x+3 \pmod{26}$，那么由密文 y 解密得到明文 x 的公式是什么？

43. 已知 Hill 密码中的明文分组长度是 2，密钥 $\boldsymbol{K}$ 是一个 2 阶可逆方阵. 假设明文 3，14，2，19 所对应的密文是 1，14，11，21，试求密钥 $\boldsymbol{K}$.

第 3 章　数论基础(二)

3.1　原根

3.1.1　整数的次数

设 m 是大于 1 的整数，a 是与 m 互素的整数，我们考虑 a 的正整数次幂

$$a,a^2,a^3,\cdots,$$

由欧拉定理可知，有

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

然而，对很多 m 来说，往往存在比 $\varphi(m)$ 还小的 a 的幂 k，就已经使得 a^k 模 m 与 1 同余. 这就提示我们先来研究使

$$a^l\equiv 1 \pmod m$$

成立的最小正整数 l，并进一步讨论关于 l 的一些性质.

定义 3.1.1　设 m 是大于 1 的整数，a 是与 m 互素的整数，使

$$a^l\equiv 1 \pmod m$$

成立的最小正整数 l 叫作 a 对模 m 的**次数**，记作 $\mathrm{ord}_m(a)$.

[例 3.1.1]　求 $\mathrm{ord}_{11}(a)$，其中 $a=1,2,\cdots,10$.

解　分别求 $a^i \pmod{11}$，$i=1,2,\cdots,10$，直至出现 $a^i\equiv 1 \pmod{11}$ 为止，可得

$$\mathrm{ord}_{11}(1)=1;$$

$$\mathrm{ord}_{11}(2)=\mathrm{ord}_{11}(6)=\mathrm{ord}_{11}(7)=\mathrm{ord}_{11}(8)=10;$$

$$\mathrm{ord}_{11}(3)=\mathrm{ord}_{11}(4)=\mathrm{ord}_{11}(5)=\mathrm{ord}_{11}(9)=5;$$

$$\mathrm{ord}_{11}(10)=2.$$

需要注意的是，在上面的定义里面只考虑那些与 m 互素的整数 a，对于 $\gcd(a,m)>1$ 的情况，是不可能存在一个正整数 l，使得关系式

$$a^l\equiv 1 \pmod m \quad (l\geqslant 1)$$

成立. 于是，当谈到“a 对模 m 的次数”的时候，即使没有明确地陈述条件 $\gcd(a,m)=1$，我们也是暗含地假设这个条件成立(同时，也隐含地认为 $m>1$ 的条件成立)，这样会使许多定理和问题的陈述变得简洁和容易记忆.

定理 3.1.1　设 a 对模 m 的次数是 $\mathrm{ord}_m(a)$，则非负整数 n 使得

$$a^n\equiv 1 \pmod m$$

的充要条件是 $\mathrm{ord}_m(a)\mid n$.

证明　先证必要性. 用反证法，假设 $\mathrm{ord}_m(a)\mid n$ 不成立，则存在整数 q,r 使得

$$n=\mathrm{ord}_m(a)q+r,\quad 0<r<\mathrm{ord}_m(a),$$

于是

$$a^r \equiv a^r (a^{\mathrm{ord}_m(a)})^q = a^n \equiv 1 (\mathrm{mod}\ m),$$

这与次数的定义中 $\mathrm{ord}_m(a)$的“最小”性质矛盾，故假设不成立，必要性得证.

再证充分性. 由于 $\mathrm{ord}_m(a) | n$，故存在整数 k 使得 $n = k\ \mathrm{ord}_m(a)$，于是

$$a^n = (a^{\mathrm{ord}_m(a)})^k \equiv 1 (\mathrm{mod}\ m).$$

定理得证.

根据定理 3.1.1，可以得到关于次数的一些性质.

定理 3.1.2　设 a 对模 m 的次数是 $\mathrm{ord}_m(a)$，则有

(1) $\mathrm{ord}_m(a) | \varphi(m)$；

(2) 若 $b \equiv a\ (\mathrm{mod}\ m)$，则 $\mathrm{ord}_m(b) = \mathrm{ord}_m(a)$.

证明　(1) 根据欧拉定理，有

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

再根据定理 3.1.1，定理显然得证.

(2) 由同余的基本性质，如果 $b \equiv a\ (\mathrm{mod}\ m)$，则 b 和 a 的任意相同次幂都同余，故显然二者次数相同，即 $\mathrm{ord}_m(b) = \mathrm{ord}_m(a)$.

由这个定理，知道 a 对模 m 的次数 $\mathrm{ord}_m(a)$必然是 $\varphi(m)$的因子，但是要注意，对于$\varphi(m)$的任意一个选定的因子 d，未必存在整数 a，使得 $\mathrm{ord}_m(a) = d$.

[例 3.1.2]　对于 $m = 12$，有 $\varphi(12) = 4$，但是不存在整数，它对模 12 的次数是 4. 因为通过计算可以得到

$$1^1 \equiv 5^2 \equiv 7^2 \equiv 11^2 \equiv 1 (\mathrm{mod}\ 12).$$

因此，任意整数的对模 12 的次数只能是 1 或者 2，而不可能是 4.

定理 3.1.3　设 a 对模 m 的次数是 $\mathrm{ord}_m(a)$，则对任意非负整数 s,t，

$$a^s \equiv a^t (\mathrm{mod}\ m)$$

的充要条件是

$$s \equiv t\ (\mathrm{mod}\ \mathrm{ord}_m(a)).$$

证明　先证必要性. 不妨设 $s \geqslant t$，若

$$a^s \equiv a^t (\mathrm{mod}\ m),$$

即 $m | (a^s - a^t)$，则有 $m | a^t (a^{s-t} - 1)$，因为 $\gcd(m, a) = 1$，所以 $m | (a^{s-t} - 1)$，即

$$a^{s-t} \equiv 1 (\mathrm{mod}\ m),$$

由定理 3.1.1 可知，$\mathrm{ord}_m(a) | s - t$，即

$$s \equiv t\ (\mathrm{mod}\ \mathrm{ord}_m(a)).$$

再证充分性. 若

$$s \equiv t\ (\mathrm{mod}\ \mathrm{ord}_m(a)),$$

则存在整数 q，使得 $s = t + q\ \mathrm{ord}_m(a)$，于是

$$a^s = a^{t + q\mathrm{ord}_m(a)} = a^t (a^{\mathrm{ord}_m(a)})^q \equiv a^t (1)^q = a^t (\mathrm{mod}\ m),$$

即

$$a^s \equiv a^t (\mathrm{mod}\ m).$$

定理得证.

定理 3.1.4　设 a 对模 m 的次数是 $\mathrm{ord}_m(a)$，则

$$1,a,a^2,\cdots,a^{\mathrm{ord}_m(a)-1}$$

两两模 m 不同余.

证明 假设存在整数 $s,t,0\leqslant s\leqslant t\leqslant \mathrm{ord}_m(a)-1$，使得

$$a^s\equiv a^t(\mathrm{mod}\ m).$$

则由定理 3.1.3 可知

$$s\equiv t\ (\mathrm{mod}\ \mathrm{ord}_m(a)),$$

显然在 $0\leqslant s\leqslant t\leqslant \mathrm{ord}_m(a)-1$ 这个范围中，$s=t$ 是唯一的可能，定理得证.

定理 3.1.5 设 a 对模 m 的次数是 $\mathrm{ord}_m(a)$，对任意非负整数 n，有

$$\mathrm{ord}_m(a^n)=\frac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),n)}.$$

证明 由于

$$a^{n\cdot \mathrm{ord}_m(a^n)}=(a^n)^{\mathrm{ord}_m(a^n)}\equiv 1(\mathrm{mod}\ m),$$

根据定理 3.1.1，有 $\mathrm{ord}_m(a)\mid n\ \mathrm{ord}_m(a^n)$，于是

$$\frac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),n)}\ \Big|\ \mathrm{ord}_m(a^n)\frac{n}{(\mathrm{ord}_m(a),n)}.$$

又因为$\left(\dfrac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),n)},\dfrac{n}{(\mathrm{ord}_m(a),n)}\right)=1$，所以

$$\frac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),n)}\ \Big|\ \mathrm{ord}_m(a^n).$$

另一方面，由于

$$(a^n)^{\frac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),n)}}=(a^{\mathrm{ord}_m(a)})^{\frac{n}{(\mathrm{ord}_m(a),n)}}\equiv 1(\mathrm{mod}\ m),$$

根据定理 3.1.1，有 $\mathrm{ord}_m(a^n)\mid\dfrac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),n)}$.

所以，

$$\mathrm{ord}_m(a^n)=\frac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),n)}.$$

定理 3.1.6 设 a 对模 m 的次数是 $\mathrm{ord}_m(a)$，对非负整数 n，使得

$$\mathrm{ord}_m(a)=\mathrm{ord}_m(a^n)$$

的充要条件是$(\mathrm{ord}_m(a),n)=1$.

证明 略. 这个定理实际上就是定理 3.1.5 的推论. 后面讨论原根时需要用到.

[例 3.1.3] 表 3.1.1 给出了整数 1 到 12 对 13 的次数，其中可以看到 $\mathrm{ord}_{13}(2)=12$，$\mathrm{ord}_{13}(4)=\mathrm{ord}_{13}(2^2)=6$，$\mathrm{ord}_{13}(8)=\mathrm{ord}_{13}(2^3)=4$，很容易验证 $6=12/\gcd(2,12)$和 $4=12/\gcd(3,12)$，这正是定理 3.1.5 给出的结论；次数与 $\mathrm{ord}_{13}(2)=12$ 相同的整数是 $6\equiv 2^5$，$7\equiv 2^{11}$，$11\equiv 2^7(\mathrm{mod}13)$，显然 5，11，7 都与 12 互素.

表 3.1.1 模 13 的次数表

整　数	1	2	3	4	5	6	7	8	9	10	11	12
次数	1	12	3	6	4	12	12	4	3	6	12	2

为了帮助读者更好地理解整数次数的概念并掌握其应用，下面给出了整数次数的几个性

质及其证明，读者可以根据需要阅读.

性质 1　设 m 和 n 都是大于 1 的整数，a 是与 m 和 n 互素的正整数，则

(1) 若 $n|m$，则 $\mathrm{ord}_n(a)|\mathrm{ord}_m(a)$.

(2) 若 $(m, n)=1$，则 $\mathrm{ord}_{mn}(a)=[\mathrm{ord}_m(a), \mathrm{ord}_n(a)]$.

证明　(1) 由于

$$a^{\mathrm{ord}_m(a)}\equiv 1(\mathrm{mod}\ m),$$

又 $n|m$，故可知

$$a^{\mathrm{ord}_m(a)}\equiv 1(\mathrm{mod}\ n).$$

于是，我们有 $\mathrm{ord}_n(a)|\mathrm{ord}_m(a)$.

(2) 由(1)可知

$$\mathrm{ord}_m(a)|\mathrm{ord}_{mn}(a),$$
$$\mathrm{ord}_n(a)|\mathrm{ord}_{mn}(a),$$

于是

$$[\mathrm{ord}_m(a), \mathrm{ord}_n(a)]|\mathrm{ord}_{mn}(a).$$

又因为

$$a^{[\mathrm{ord}_m(a),\ \mathrm{ord}_n(a)]}\equiv 1(\mathrm{mod}\ m),$$
$$a^{[\mathrm{ord}_m(a),\ \mathrm{ord}_n(a)]}\equiv 1(\mathrm{mod}\ n),$$

所以

$$a^{[\mathrm{ord}_m(a),\ \mathrm{ord}_n(a)]}\equiv 1(\mathrm{mod}\ mn).$$

于是

$$\mathrm{ord}_{mn}(a)|[\mathrm{ord}_m(a), \mathrm{ord}_n(a)].$$

因此，得到

$$\mathrm{ord}_{mn}(a)=[\mathrm{ord}_m(a), \mathrm{ord}_n(a)].$$

由性质 1 的(2)可直接得到下面性质.

性质 2　设 m 是大于 1 的整数，a 是与 m 互素的正整数，则当 m 的标准分解式为

$$m=p_1^{\alpha_1}p_2^{\alpha_2}\cdots p_s^{\alpha_s},\quad \alpha_i>0, i=1,2,\cdots,s$$

时，有

$$\mathrm{ord}_m(a)=[\mathrm{ord}_{p_1^{\alpha_1}}(a), \mathrm{ord}_{p_2^{\alpha_2}}(a), \cdots, \mathrm{ord}_{p_s^{\alpha_s}}(a)].$$

性质 3　设 m 和 n 都是大于 1 的整数，且 $(m, n)=1$，则对与 mn 互素的任意正整数 a，b，存在正整数 c，使得

$$\mathrm{ord}_{mn}(c)=[\mathrm{ord}_m(a), \mathrm{ord}_n(b)].$$

证明　考虑同余方程组

$$\begin{cases} x\equiv a(\mathrm{mod}\ m) \\ x\equiv b(\mathrm{mod}\ n) \end{cases},$$

由孙子定理可知，此同余方程组有唯一解

$$x\equiv c\ (\mathrm{mod}\ mn).$$

由定理 3.1.2(2)可知

$$\mathrm{ord}_m(c)=\mathrm{ord}_m(a),\quad \mathrm{ord}_n(c)=\mathrm{ord}_n(b),$$

于是，根据性质 1 中(2)，有

$$\mathrm{ord}_{mn}(c)=[\mathrm{ord}_m(c),\ \mathrm{ord}_n(c)]=[\mathrm{ord}_m(a),\ \mathrm{ord}_n(b)].$$

3.1.2 原根

定义 3.1.2 设 m 是大于 1 的整数，a 是与 m 互素的整数，若

$$\mathrm{ord}_m(a)=\varphi(m),$$

则 a 叫作 m 的**原根**.

在例 3.1.1 中，由于 $\varphi(11)=10$，故 2，6，7，8 是 11 的原根.

[例 3.1.4] 5 是否为 6 的原根？是否为 8 的原根？

解 由于 5 与 6 互素，$\varphi(6)=2$，又

$$5^1\equiv 5\ ,\quad 5^2\equiv 1\ (\mathrm{mod}\ 6),$$

故 $\mathrm{ord}_6(5)=\varphi(6)$，即 5 是 6 的原根.

由于 5 与 8 互素，$\varphi(8)=4$，又

$$5^1\equiv 5\ ,\quad 5^2\equiv 1\ (\mathrm{mod}\ 8),$$

故 $\mathrm{ord}_8(5)=2\neq\varphi(8)$，即 5 不是 8 的原根.

在下面的讨论中，基于与前一小节同样的道理，当谈到“a 是否为 m 的原根”的问题时，即使没有明确陈述定义中的条件“m 是大于 1 的整数，a 是与 m 互素的整数”，我们仍然是暗含地假设这个条件成立，这样陈述将变得简洁且易记忆.

定理 3.1.7 a 是 m 的原根的充要条件是

$$1,a,a^2,\cdots,a^{\varphi(m)-1}$$

是模 m 的一个缩系.

证明 先证必要性. 若 a 是 m 的原根，则

$$\mathrm{ord}_m(a)=\varphi(m),$$

根据定理 3.1.4，可知

$$1,a,a^2,\cdots,a^{\varphi(m)-1}$$

两两模 m 不同余. 又因为 a 与 m 互素，所以 a 的任意非负整数次幂都与 m 互素，因此这 $\varphi(m)$ 个数组成模 m 的一个缩系.

再证充分性. 若

$$1,a,a^2,\cdots,a^{\varphi(m)-1}$$

这 $\varphi(m)$个数是模 m 的一个缩系，则 a 与 m 互素，而且这 $\varphi(m)$个数之间两两不同余. 所以这 $\varphi(m)$个数中，除了 1 以外，其他 $\varphi(m)-1$ 个数都与 1 不同余，即对任一整数 s，$1\leqslant s\leqslant\varphi(m)-1$，$a^s$ 与 1 模 m 不同余. 根据欧拉定理，可知

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

所以由次数的定义可知

$$\mathrm{ord}_m(a)=\varphi(m),$$

即 a 是 m 的原根. 证毕.

定理 3.1.8 设 a 是 m 的一个原根，t 是非负整数，则 a^t 也是 m 的原根的充要条件是 $(t,\ \varphi(m))=1$.

证明 因为 $\mathrm{ord}_m(a)=\varphi(m)$，所以由定理 3.1.6 可知，$\mathrm{ord}_m(a^t)=\mathrm{ord}_m(a)=\varphi(m)$的充

要条件是$(t, \operatorname{ord}_m(a))=(t, \varphi(m))=1$. 即 a^t 是 m 的原根的充要条件是$(t, \varphi(m))=1$.

定理 3.1.9　设 a 是 m 的一个原根，则 m 恰有 $\varphi(\varphi(m))$个模 m 不同余的原根.

证明　由于 a 是 m 的原根，故 $\varphi(m)$个整数

$$1, a, a^2, \cdots, a^{\varphi(m)-1}$$

构成模 m 的一个缩系. 根据定理 3.1.8，a^t 是 m 的原根当且仅当$(t, \varphi(m))=1$. 因为这样的 t 共有 $\varphi(\varphi(m))$个，所以 m 恰有 $\varphi(\varphi(m))$个模 m 不同余的原根.

[例 3.1.5]　在例 3.1.3 中，模 13 的原根是 2，6，7，11，共 4 个原根，易验证

$$\varphi(\varphi(13))=\varphi(12)=\varphi(3\times 2^2)=2\times 2=4.$$

[例 3.1.6]　试求 8 的原根.

解　先求出 $\varphi(8)=4$. 易知

$$\operatorname{ord}_8(1)=1,\quad \operatorname{ord}_8(3)=2,\quad \operatorname{ord}_8(5)=2,\quad \operatorname{ord}_8(7)=2,$$

因此不存在 8 的原根.

由这个例子看出，对任意模数 m 来说，不一定存在原根，下面重点讨论一些原根的存在性问题. 正式讨论之前，作为预备，先来证明两个定理.

定理 3.1.10　设 a 和 b 对模 m 的次数分别是 $\operatorname{ord}_m(a)$和 $\operatorname{ord}_m(b)$，则

$$(\operatorname{ord}_m(a), \operatorname{ord}_m(b))=1$$

的充要条件是

$$\operatorname{ord}_m(ab)=\operatorname{ord}_m(a)\operatorname{ord}_m(b).$$

证明　由于$(a,m)=1,(b,m)=1$,故$(ab,m)=1$，且存在 $\operatorname{ord}_m(ab)$.

先证必要性. 由

$$a^{\operatorname{ord}_m(b)\operatorname{ord}_m(ab)}\equiv(a^{\operatorname{ord}_m(b)})^{\operatorname{ord}_m(ab)}(b^{\operatorname{ord}_m(b)})^{\operatorname{ord}_m(ab)}\equiv((ab)^{\operatorname{ord}_m(ab)})^{\operatorname{ord}_m(b)}\equiv 1 \pmod m$$

可知 $\operatorname{ord}_m(a)\mid\operatorname{ord}_m(b)\operatorname{ord}_m(ab)$，又$(\operatorname{ord}_m(a), \operatorname{ord}_m(b))=1$，所以 $\operatorname{ord}_m(a)\mid\operatorname{ord}_m(ab)$.

同理可证 $\operatorname{ord}_m(b)\mid\operatorname{ord}_m(ab)$. 由于$(\operatorname{ord}_m(a), \operatorname{ord}_m(b))=1$，所以 $\operatorname{ord}_m(a)\operatorname{ord}_m(b)\mid\operatorname{ord}_m(ab)$.

另一方面，由

$$(ab)^{\operatorname{ord}_m(a)\operatorname{ord}_m(b)}\equiv(a^{\operatorname{ord}_m(a)})^{\operatorname{ord}_m(b)}(b^{\operatorname{ord}_m(b)})^{\operatorname{ord}_m(a)}\equiv 1 \pmod m$$

可知 $\operatorname{ord}_m(ab)\mid\operatorname{ord}_m(a)\operatorname{ord}_m(b)$. 所以

$$\operatorname{ord}_m(ab)=\operatorname{ord}_m(a)\operatorname{ord}_m(b).$$

再证充分性. 由

$$(ab)^{[\operatorname{ord}_m(a), \operatorname{ord}_m(b)]}\equiv a^{[\operatorname{ord}_m(a), \operatorname{ord}_m(b)]}b^{[\operatorname{ord}_m(a), \operatorname{ord}_m(b)]}\equiv 1 \pmod m$$

可知 $\operatorname{ord}_m(ab)\mid[\operatorname{ord}_m(a), \operatorname{ord}_m(b)]$. 又

$$\operatorname{ord}_m(ab)=\operatorname{ord}_m(a)\operatorname{ord}_m(b),$$

于是 $\operatorname{ord}_m(a)\operatorname{ord}_m(b)\mid[\operatorname{ord}_m(a), \operatorname{ord}_m(b)]$. 所以

$$(\operatorname{ord}_m(a), \operatorname{ord}_m(b))=1.$$

定理得证.

定理 3.1.11　设 a 和 b 对模 m 的次数分别是 $\operatorname{ord}_m(a)$和 $\operatorname{ord}_m(b)$，则存在整数 c，使得

$$\operatorname{ord}_m(c)=[\operatorname{ord}_m(a), \operatorname{ord}_m(b)].$$

证明　因为对于整数 $\operatorname{ord}_m(a)$和 $\operatorname{ord}_m(b)$，存在整数 u，v 满足

$$u\mid\operatorname{ord}_m(a),\quad v\mid\operatorname{ord}_m(b),$$

并使得

$$(u,v)=1,\quad uv=[\mathrm{ord}_m(a),\mathrm{ord}_m(b)].$$

令

$$s=\frac{\mathrm{ord}_m(a)}{u},\quad t=\frac{\mathrm{ord}_m(b)}{v},$$

则

$$\mathrm{ord}_m(a^s)=\frac{\mathrm{ord}_m(a)}{(\mathrm{ord}_m(a),s)}=\frac{\mathrm{ord}_m(a)}{s}=u.$$ 同理，$\mathrm{ord}_m(b^t)=v$.

又由定理 3.1.10 可知

$$\mathrm{ord}_m(a^sb^t)=\mathrm{ord}_m(a^s)\mathrm{ord}_m(b^t)=uv=[\mathrm{ord}_m(a),\mathrm{ord}_m(b)].$$

于是，取 $c\equiv a^sb^t\pmod m$即可. 定理得证.

到这里，就可以得到如下的第一个原根存在性定理.

定理 3.1.12 设 p 是奇素数，则 p 的原根存在.

证明 在模 p 的缩系 $1,2,\cdots,p-1$ 中，记

$$u_r=\mathrm{ord}_p(r),\quad 1\leqslant r\leqslant p-1,$$

令 $u=[u_1,u_2,\cdots,u_{p-1}]$. 反复应用定理 3.1.11 可知，存在整数 g，使得

$$\mathrm{ord}_p(g)=u.$$

根据定理 3.1.2(1)，可知 $u\mid\varphi(p)$，即 $u\mid p-1$，所以 $u\leqslant p-1$.

由于

$$r^{u_r}\equiv 1\pmod p,\quad 1\leqslant r\leqslant p-1,$$

又 $u_r\mid u$，故

$$r^{u}\equiv 1\pmod p,\quad 1\leqslant r\leqslant p-1,$$

即同余方程

$$x^u\equiv 1\pmod p,$$

至少有 $p-1$ 个解

$$x\equiv 1,2,\cdots,p-1\pmod p.$$

又根据拉格朗日关于同余方程解的数量的定理可知，该方程至多有 u 个解，所以 $p-1\leqslant u$.

因此，我们有 $u=p-1$，即 $\mathrm{ord}_p(g)=u=p-1=\varphi(p)$. 所以 g 是 p 的原根. 定理得证.

定理 3.1.13 设 g 是奇素数 p 的一个原根，且满足

$$g^{p-1}\not\equiv 1\pmod{p^2},$$

则对每一个 $l\geqslant 2$，有

$$g^{\varphi(p^{l-1})}\not\equiv 1\pmod{p^l}.$$

证明 对 l 用数学归纳法. 当 $l=2$ 时，即为题设 $g^{p-1}\not\equiv 1\pmod{p^2}$，显然成立.

假设定理对 $l(l\geqslant 2)$成立，即

$$g^{\varphi(p^{l-1})}\not\equiv 1\pmod{p^l}$$

由欧拉定理可知

$$g^{\varphi(p^{l-1})}\equiv 1\pmod{p^{l-1}}.$$

所以存在整数 k 使得

$$g^{\varphi(p^{l-1})}=1+kp^{l-1},$$

由归纳假设可知，其中 k 不能被 p 整除(否则，如果 $k=k_1p$，那么 $g^{\varphi(p^{l-1})}=1+k_1pp^{l-1}=k_1p^l$，即 $g^{\varphi(p^{l-1})}\equiv 1\pmod{p^l}$，这与归纳假设矛盾). 将上式两端分别取 p 次方，可得

$$(g^{\varphi(p^{l-1})})^p=(g^{p^{l-1}-p^{l-2}})^p=g^{p^l-p^{l-1}}$$

$$=g^{\varphi(p^l)}=(1+kp^{l-1})^p=1+kp^l+k^2\frac{p(p-1)}{2}p^{2(l-1)}+rp^{3(l-1)},$$

其中 r 是一个整数. 由于 $2l-1\geqslant l+1$，$3(l-1)\geqslant l+1$，所以上式最右端从第三项起，都能够被 p^{l+1} 整除，因此

$$g^{\varphi(p^l)}\equiv 1+kp^l\pmod{p^{l+1}}.$$

因为 k 不能被 p 整除，所以有

$$g^{\varphi(p^l)}\not\equiv 1\pmod{p^{l+1}},$$

于是，定理对 $l+1$ 成立. 证毕.

定理 3.1.14　设 p 是一个奇素数，则对任意正整数 l，存在 p^l 的原根.

证明　当 $l=1$ 时，定理成立，可设 g 为 p 的原根，则有

$$g^{p-1}\equiv 1\pmod{p}.$$

若

$$g^{p-1}-1\not\equiv 0\pmod{p^2},$$

我们取 $r=g$. 反之，若

$$g^{p-1}-1\equiv 0\pmod{p^2},$$

我们取 $r=g+p$，由于 $r\equiv g\pmod{p}$，所以 r 也是 p 的原根，且

$r^{p-1}-1=(g+p)^{p-1}-1=g^{p-1}+(p-1)pg^{p-2}+$ p^2 的倍数项$-1\equiv -pg^{p-2}\not\equiv 0\pmod{p^2}$.

即总能够找到模 p 的原根 r，满足 $r^{p-1}\not\equiv 1\pmod{p^2}$.

下面开始证明 r 即为 $p^l(l\geqslant 2)$的原根. 设

$$t=\mathrm{ord}_{p^l}(r),$$

则有

$$r^t\equiv 1\pmod{p^l},$$

显然也有

$$r^t\equiv 1\pmod{p}.$$

因为 r 是 p 的原根，所以有 $\varphi(p)\mid t$，于是可记

$$t=\varphi(p)q.$$

由于 $t\mid\varphi(p^l)$，即 $\varphi(p)q\mid\varphi(p^l)$，又

$$\varphi(p^l)=p^{l-1}(p-1),\quad \varphi(p)=p-1,$$

故有 $q\mid p^{l-1}$.

不妨设 $q=p^k$，其中 $k\leqslant l-1$. 若这个不等式严格成立 $k<l-1$，则 $k+1\leqslant l-1$，即

$$l-k-2\geqslant 0$$

由

$$t=\varphi(p)p^k=(p-1)\,p^k=p^{k+1}-p^k,\quad \varphi(p^{l-1})=p^{l-1}-p^{l-2}=(p^{k+1}-p^k)\,p^{l-k-2}=t\,p^{l-k-2},$$

可知

$$t\mid\varphi(p^{l-1}),$$

因此

$$r^{\varphi(p^{l-1})}\equiv 1 \pmod{p^l}.$$

但这个结果显然与定理 3.1.13 矛盾，于是只能 $k=l-1$，即 $t=\varphi(p^l)$. 所以 r 是 p^l 的一个原根，定理得证.

从该定理看出，素数 p 的原根不一定是 p^2 的原根.

[例 3.1.7] 8 是 3 的原根，但不是 3^2 的原根，因为 $8^2\equiv 1 \pmod{3^2}$.

定理 3.1.15 设 p 是一个奇素数，则对任意正整数 l，存在 $2p^l$ 的原根.

证明 设 g 是 p^l 的一个原根，先证当 g 是奇数时，g 也是 $2p^l$ 的一个原根.

因为 $(g,p^l)=1$ 且 $(g,2)=1$，所以 $(g,2p^l)=1$，因此由欧拉定理可知

$$g^{\varphi(2p^l)}\equiv 1 \pmod{2p^l}.$$

设 $t=\mathrm{ord}_{2p^l}(g)$，又因为 $\varphi(2p^l)=\varphi(2)\varphi(p^l)=\varphi(p^l)$，故有 $t\mid\varphi(p^l)$.

由

$$g^t\equiv 1 \pmod{2p^l},$$

可知

$$g^t\equiv 1 \pmod{p^l}.$$

又因为 g 是 p^l 的一个原根，所以 $\varphi(p^l)\mid t$.

于是，有 $t=\varphi(p^l)=\varphi(2p^l)$，即 g 是 $2p^l$ 的一个原根.

当 g 是偶数时，则 $g+p^l$ 是奇数且为 p^l 的一个原根(因为 $g\equiv g+p^l \pmod{p^l}$)，可类似地按以上证明得出结论. 证毕.

上面讨论了具有原根的一些整数的特征，为了完整地给出具有原根的所有整数的特征，还需要排除那些没有原根的整数. 首先下面的定理将说明例 3.1.6 中的整数 8 为什么没有原根.

定理 3.1.16 设 a 是一个奇数，则对任意整数 $k\geqslant 3$，有

$$a^{\frac{1}{2}\varphi(2^k)}\equiv a^{2^{k-2}}\equiv 1 \pmod{2^k}.$$

即 $2^k(k\geqslant 3)$ 没有原根.

证明 用数学归纳法. 不妨设 $a=2b+1$，则有

$$a^2=4b(b+1)+1\equiv 1 \pmod{2^3},$$

注意其中 $2\mid b(b+1)$，而 $\varphi(2^3)=4$，所以结论对 $k=3$ 成立.

假设结论对 $k-1$ $(k>3)$成立，则有

$$a^{2^{(k-1)-2}}\equiv 1 \pmod{2^{k-1}},$$

即存在整数 q 使得

$$a^{2^{(k-1)-2}}=1+q2^{k-1}.$$

将等式两端分别平方，可得

$$a^{2^{k-2}}=(1+q2^{k-1})^2=1+(q+2^{k-2}q^2)2^k,$$

故

$$a^{2^{k-2}}\equiv 1 \pmod{2^k},$$

即结论对 k 成立. 于是定理得证.

有了前面的这些定理，就不难推出原根存在的充要条件了.

定理 3.1.17 设 m 是大于 1 的整数，则 m 的原根存在的充要条件是 m 为 $2,4,p^l,2p^l$ 之一，其中 $l\geqslant 1$，p 是奇素数.

证明　先证必要性. 设 m 的标准分解式为

$$m=p_1^{l_1}p_2^{l_2}\cdots p_s^{l_s},$$

其中 $p_i<p_j(i<j)$. 又设 a 为一与 m 互素的正整数，则必满足

$$(a,\ p_i^{l_i})=1,i=1,2,\cdots,s.$$

由欧拉定理，可知

$$a^{\varphi(p_i^{l_i})}\equiv 1\ (\mathrm{mod}\ p_i^{l_i}),\quad i=1,2,\cdots,s.$$

令 $h=[\varphi(p_1^{l_1}),\varphi(p_2^{l_2}),\cdots,\varphi(p_s^{l_s})]$，则

$$a^h\equiv 1\ (\mathrm{mod}\ p_i^{l_i}),\quad i=1,2,\cdots,s.$$

由于 $p_i^{l_i}(i=1,2,\cdots,s)$两两互素，于是$[p_1^{l_1},p_2^{l_2},\cdots,p_s^{l_s}]=m$，故有

$$a^h\equiv 1\ (\mathrm{mod}\ m).$$

因为 $h\leqslant\varphi(m)$，而当 $h<\varphi(m)$时，m 无原根存在，所以，若 m 有原根，则必须

$$h=\varphi(m),$$

即 $\varphi(p_i^{l_i})(i=1,2,\cdots,s)$两两互素.

因为 $\varphi(p^l)=p^{l-1}(p-1)$，当 p 为奇素数时，$\varphi(p^l)$必为偶数，所以当 m 有两个或两个以上的奇素数因子时，m 无原根. 于是，若使 m 有原根，m 只能具有 2^k，p^l，2^tp^l 三种形式之一，其中 k，t，l 均为正整数.

若 $t>1$，则 $\varphi(2^t)=2^{t-1}$与 $\varphi(p^l)$不互素，故只能 $t=1$.

若 $k\geqslant 3$，由定理 3.1.16 显然可知 2^k 无原根存在，故只能 $k=1$ 或 $k=2$.

综上所述，若 m 有原根，则 m 只能是 2，4，p^l，$2p^l$ 之一，必要性成立.

再证充分性.

当 $m=2$ 时，$\varphi(2)=1$，1 即为 2 的原根.

当 $m=4$ 时，$\varphi(4)=2$，3 即为 4 的原根.

当 $m=p^l$ 时，由定理 3.1.14 可知 m 的原根存在.

当 $m=2p^l$ 时，由定理 3.1.15 可知 m 的原根存在.

于是充分性也成立，定理得证.

下面再给出一种寻找原根的方法.

定理 3.1.18　设 m 是大于 2 的整数，$\varphi(m)$的所有不同的素因子是 $q_1,q_2,\cdots,q_s$，则与 m 互素的正整数 g 是 m 的一个原根的充要条件是

$$g^{\frac{\varphi(m)}{q_i}}\not\equiv 1\ (\mathrm{mod}\ m),\quad i=1,2,\cdots,s.$$

证明　先证必要性. 若 g 是 m 的一个原根，则有

$$\mathrm{ord}_m(g)=\varphi(m).$$

而

$$0<\frac{\varphi(m)}{q_i}<\varphi(m),i=1,2,\cdots,s,$$

所以

$$g^{\frac{\varphi(m)}{q_i}}\not\equiv 1\ (\mathrm{mod}\ m),\quad i=1,2,\cdots,s.$$

再证充分性. 用反证法，设

$$\mathrm{ord}_m(g)=v,$$

假定 g 不是 m 的一个原根，则 $v<\varphi(m)$，从而 $v\mid\varphi(m)$. 于是存在一个素数 q，使得

$$q \mid \frac{\varphi(m)}{v},$$

故又存在一个整数 u，使得

$$\frac{\varphi(m)}{v}=qu,$$

即

$$\frac{\varphi(m)}{q}=uv.$$

于是，有

$$g^{\frac{\varphi(m)}{q}}=(g^v)^u\equiv 1 \pmod m,$$

这与所给条件是矛盾的. 所以假设不成立，充分性得证. 证毕.

当 m 数值比较小时，可以利用这个定理很快发现 m 的原根. 但是，当 m 数值比较大时，可能很难找到 $\varphi(m)$的所有素数因子，这个时候就很难应用该定理了. 到目前为止，即使知道 m 有原根，人们也没有找到一个具有普遍性的容易的方法来发现 m 的原根. 然而，如果已知一个原根，那么其他的所有原根就可以比较容易地计算出来，该方法的根据就是定理 3.1.8.

[例 3.1.8] 求 41 的原根.

解 因为 $\varphi(m)=\varphi(41)=40=2^3\times 5$，所以 $\varphi(m)$的素因子是 $q_1=2, q_2=5$，进而

$$\frac{\varphi(m)}{q_1}=20,\quad \frac{\varphi(m)}{q_2}=8.$$

对 $g=2, 3, \cdots$ 逐个验算 g^{20} 和 g^8 是否与 1 模 m 同余，得

$$2^8\equiv 10,\quad 2^{20}\equiv 1,\quad 3^8\equiv 1,\quad 3^{20}\equiv 40,\quad 4^8\equiv 18,\quad 4^{20}\equiv 1,$$
$$5^8\equiv 18,\quad 5^{20}\equiv 1,\quad 6^8\equiv 10,\quad 6^{20}\equiv 40 \pmod{41},$$

可知 6 是 41 的最小原根.

根据定理 3.1.8，可知当 t 遍历 $\varphi(m)=40$ 的缩系

$$1, 3, 7, 9, 11, 13, 17, 19, 21, 23, 27, 29, 31, 33, 37, 39$$

时，6^t 遍历 41 的原根，即

$$6^1\equiv 6,\quad 6^3\equiv 11,\quad 6^7\equiv 29,\quad 6^9\equiv 19,\quad 6^{11}\equiv 28,\quad 6^{13}\equiv 24,$$
$$6^{17}\equiv 26,\quad 6^{19}\equiv 34,\quad 6^{21}\equiv 35,\quad 6^{23}\equiv 30,\quad 6^{27}\equiv 12,\quad 6^{29}\equiv 22,$$
$$6^{31}\equiv 13,\quad 6^{33}\equiv 17,\quad 6^{37}\equiv 15,\quad 6^{39}\equiv 7 \pmod{41}.$$

3.1.3 指数与 n 次剩余

如果 m 有一个原根 g，则根据定理 3.1.7 可知，

$$1, g, g^2, \cdots, g^{\varphi(m)-1}$$

是模 m 的一个缩系. 因此，对任一与 m 互素的整数 a，存在唯一的非负整数 r，$0\leqslant r<\varphi(m)$，使得

$$g^r\equiv a \pmod m.$$

由于原根具有上述性质，可以给出下面的定义.

定义 3.1.3 设 m 是大于 1 的整数，g 是 m 的一个原根，a 是与 m 互素的整数，则存在唯一的非负整数 r，$0\leqslant r<\varphi(m)$，满足

$$a\equiv g^r \pmod m,$$

于是，把 r 叫作以 g 为底 a 对模 m 的**指数**，记作 $\mathrm{ind}_g a$. 在不易引起混淆的情况下，可把 $\mathrm{ind}_g a$ 简写成 $\mathrm{ind}\, a$.

显然，根据定义有

$$a\equiv g^{\mathrm{ind}_g a} \pmod m.$$

有时，也把指数叫作**离散对数**，记作 $\log_g a$，于是

$$a\equiv g^{\log_g a} \pmod m.$$

定理 3.1.19　g 是 m 的一个原根，a 是与 m 互素的整数，如果非负整数 k 使得同余式

$$g^k\equiv a \pmod m$$

成立，则 k 满足

$$k\equiv \mathrm{ind}_g a \pmod{\varphi(m)}.$$

证明　因为

$$g^k\equiv a\equiv g^{\mathrm{ind}_g a} \pmod m,$$

根据定理 3.1.3 可知

$$k\equiv \mathrm{ind}_g a \pmod{\mathrm{ord}_m(g)}.$$

又因为 g 是 m 的一个原根，所以

$$\mathrm{ord}_m(g)=\varphi(m),$$

所以

$$k\equiv \mathrm{ind}_g a \pmod{\varphi(m)}.$$

定理 3.1.20　g 是 m 的一个原根，则

$$g^x\equiv g^y \pmod m$$

成立的充要条件是

$$x\equiv y \pmod{\varphi(m)}$$

成立.

证明　直接应用定理 3.1.3 和 $\mathrm{ord}_m(g)=\varphi(m)$，即得证.

下面的两个定理给出指数的最重要的性质.

定理 3.1.21　g 是 m 的一个原根，整数 a 和 b 均与 m 互素，则

(1) $\mathrm{ind}_g 1\equiv 0 \pmod{\varphi(m)}$；$\mathrm{ind}_g g\equiv 1 \pmod{\varphi(m)}$；

(2) $\mathrm{ind}_g(ab)\equiv \mathrm{ind}_g a+\mathrm{ind}_g b \pmod{\varphi(m)}$；

(3) $\mathrm{ind}_g a^k\equiv k\,\mathrm{ind}_g a \pmod{\varphi(m)}$，其中 k 为非负整数.

证明　(1) 因为

$$g^0\equiv 1 \pmod m,$$

根据定理 3.1.19，可知

$$\mathrm{ind}_g 1\equiv 0 \pmod{\varphi(m)}.$$

因为

$$g^1\equiv g \pmod m,$$

根据定理 3.1.19，可知

$$\mathrm{ind}_g g\equiv 1 \pmod{\varphi(m)}.$$

(2) 因为

$$ab\equiv g^{\mathrm{ind}_g(ab)}\pmod m,a\equiv g^{\mathrm{ind}_g a}\pmod m,b\equiv g^{\mathrm{ind}_g b}\pmod m,$$

所以

$$g^{\mathrm{ind}_g(ab)}\equiv g^{\mathrm{ind}_g a+\mathrm{ind}_g b}\pmod m.$$

根据定理 3.1.20，可知

$$\mathrm{ind}_g(ab)\equiv\mathrm{ind}_g a+\mathrm{ind}_g b\pmod{\varphi(m)}.$$

(3) 因为

$$a^k\equiv g^{\mathrm{ind}_g a^k}\pmod m,\ a\equiv g^{\mathrm{ind}_g a}\pmod m,$$

所以

$$g^{\mathrm{ind}_g a^k}\equiv a^k\equiv(g^{\mathrm{ind}_g a})^k\equiv g^{k\mathrm{ind}_g a}\pmod m.$$

根据定理 3.1.20，可知

$$\mathrm{ind}_g a^k\equiv k\mathrm{ind}_g a\pmod{\varphi(m)}.$$

定理 3.1.22 g 是 m 的一个原根，整数 a 和 b 均与 m 互素，则 $a\equiv b\pmod m$的充要条件是 $\mathrm{ind}_g a\equiv\mathrm{ind}_g b\pmod{\varphi(m)}$.

证明 先证必要性. 因为 $a\equiv b\pmod m$，所以

$$g^{\mathrm{ind}_g a}\equiv g^{\mathrm{ind}_g b}\pmod m,$$

所以由定理 3.1.20 可知

$$\mathrm{ind}_g a\equiv\mathrm{ind}_g b\pmod{\varphi(m)}.$$

再证充分性. 因为 $\mathrm{ind}_g a\equiv\mathrm{ind}_g b\pmod{\varphi(m)}$，所以存在整数 k 使得

$$\mathrm{ind}_g a=\mathrm{ind}_g b+k\varphi(m),$$

所以

$$g^{\mathrm{ind}_g a}=g^{\mathrm{ind}_g b+k\varphi(m)}=g^{\mathrm{ind}_g b}(g^{\varphi(m)})^k\equiv g^{\mathrm{ind}_g b}(1)^k=g^{\mathrm{ind}_g b}\pmod m$$

即

$$a\equiv b\pmod m.$$

上面两个定理使我们清楚地认识到指数的性质和实数的对数的性质具有非常相似的特点，因此可以利用原根，做出指数表.

[例 3.1.9] 做模 41 的指数表.

解 已知 $g=6$ 是 41 的原根，且 $\varphi(41)=40$，直接计算 $g^r\pmod m$，$0\leqslant r\leqslant 39$，即

$6^0\equiv 1$，$6^1\equiv 6$，$6^2\equiv 36$，$6^3\equiv 11$，$6^4\equiv 25$，$6^5\equiv 27$，
$6^6\equiv 39$，$6^7\equiv 29$，$6^8\equiv 10$，$6^9\equiv 19$，$6^{10}\equiv 32$，$6^{11}\equiv 28$，
$6^{12}\equiv 4$，$6^{13}\equiv 24$，$6^{14}\equiv 21$，$6^{15}\equiv 3$，$6^{16}\equiv 18$，$6^{17}\equiv 26$，
$6^{18}\equiv 33$，$6^{19}\equiv 34$，$6^{20}\equiv 40$，$6^{21}\equiv 35$，$6^{22}\equiv 5$，$6^{23}\equiv 30$，
$6^{24}\equiv 16$，$6^{25}\equiv 14$，$6^{26}\equiv 2$，$6^{27}\equiv 12$，$6^{28}\equiv 31$，$6^{29}\equiv 22$，
$6^{30}\equiv 9$，$6^{31}\equiv 13$，$6^{32}\equiv 37$，$6^{33}\equiv 17$，$6^{34}\equiv 20$，$6^{35}\equiv 38$，
$6^{36}\equiv 23$，$6^{37}\equiv 15$，$6^{38}\equiv 8$，$6^{39}\equiv 7$ $\pmod{41}$.

下面做出模 41 的指数表(见表 3.1.2)，第一行表示 $g^r\pmod m$的个位数，第一列表示 $g^r\pmod m$ 的十位数，交叉位置即为 r.

表 3.1.2 模 41 的指数表

	0	1	2	3	4	5	6	7	8	9

0		0	26	15	12	22	1	39	38	30
1	8	3	27	31	25	37	24	33	16	9
2	34	14	29	36	13	4	17	5	11	7
3	23	28	10	18	19	21	2	32	35	6
4	20									

我们知道，如果从已知整数 r 来计算 $a\equiv g^r \pmod m$很容易，而从已知整数 a 求整数 r 使得 $g^r\equiv a \pmod m$有时是很困难的. 指数表对解决此类问题有一定的帮助. 例如，通过查表可以很快地知道以 6 为底 28 对模 41 的指数是 11.

指数表可以用来解一些特殊类型的(高次)同余方程，下面开始讨论这个问题.

定义 3.1.4　设 m 是大于 1 的整数，a 是与 m 互素的整数，若 n ($n\geqslant 2$)次同余方程

$$x^n\equiv a \pmod m$$

有解，则 a 叫作模 m 的***n* 次剩余**. 否则，a 叫作模 m 的***n* 次非剩余**.

定理 3.1.23　g 是 m 的一个原根，a 是与 m 互素的整数，则同余方程

$$x^n\equiv a \pmod m \tag{3.1.1}$$

有解的充要条件是$(n,\varphi(m))\mid \mathrm{ind}_g a$. 并且，若此同余方程有解，则解数恰为$(n,\varphi(m))$.

证明　先证明同余方程(3.1.1)与同余方程

$$n\mathrm{ind}_g x\equiv \mathrm{ind}_g a \pmod{\varphi(m)} \tag{3.1.2}$$

等价. 若同余方程(3.1.1)有解，设为

$$x\equiv x_0 \pmod m,$$

则

$$x_0^n\equiv a \pmod m,$$

即

$$g^{\mathrm{ind}_g x_0^n}\equiv g^{n\mathrm{ind}_g x_0}\equiv g^{\mathrm{ind}_g a} \pmod m.$$

由定理 3.1.20 可知

$$n\mathrm{ind}_g x_0\equiv \mathrm{ind}_g a \pmod{\varphi(m)}.$$

反过来，若同余方程(3.1.2)有解，设为

$$x\equiv x_0 \pmod{\varphi(m)},$$

使得

$$n\mathrm{ind}_g x_0\equiv \mathrm{ind}_g a \pmod{\varphi(m)}.$$

由定理 3.1.20 可知

$$g^{n\mathrm{ind}_g x_0}\equiv g^{\mathrm{ind}_g x_0^n}\equiv g^{\mathrm{ind}_g a} \pmod m,$$

即

$$x_0^n\equiv a \pmod m.$$

因此，同余方程(3.1.1)与同余方程(3.1.2)等价.

由于对任一给定整数 X，同余方程

$$X\equiv \mathrm{ind}_g x \pmod{\varphi(m)}$$

总有解，故同余方程(3.1.2)有解的充要条件是

$$nX\equiv \mathrm{ind}_g a \pmod{\varphi(m)}$$

有解. 又根据定理 2.2.24，可知同余方程(3.1.1)有解的充要条件是$(n,\varphi(m))|\mathrm{ind}_g a$. 并且，若此同余方程有解，则解数恰为$(n,\varphi(m))$.

定理 3.1.24 g 是 m 的一个原根，a 是与 m 互素的整数，则 a 是模 m 的 n 次剩余的充要条件是

$$a^{\frac{\varphi(m)}{d}}\equiv 1 \pmod{m},\ d=(n,\varphi(m)).$$

证明 根据定理 3.1.23 可知

$$x^n\equiv a \pmod{m}$$

有解的充要条件是 $d|\mathrm{ind}_g a$，即

$$\mathrm{ind}_g a\equiv 0 \pmod{d}.$$

而这个式子的一个等价式(充要条件)为

$$\frac{\varphi(m)}{d}\mathrm{ind}_g a\equiv 0(\mathrm{mod}\varphi(m)).$$

由定理 3.1.20，可得其充要条件为

$$g^{\frac{\varphi(m)}{d}\mathrm{ind}_g a}\equiv a^{\frac{\varphi(m)}{d}}\equiv g^0\equiv 1(\mathrm{mod}\ m),$$

于是定理得证.

定理 3.1.25 a 是与素数 p 互素的整数，则 a 是模 p 的 2 次剩余的充要条件是

$$a^{\frac{p-1}{2}}\equiv 1(\mathrm{mod}\ p).$$

证明 略. 这个定理就是上面定理的推论.

[例 3.1.10] 求解同余方程

$$x^{12}\equiv 37 \pmod{41}.$$

解 因为 $\varphi(41)=40$，$d=(12,40)=4$，查模 41 的指数表得到 $\mathrm{ind}_g 37=32$，所以根据 $4|32$ 可知同余方程有解. 由于原同余方程与

$$12\mathrm{ind}_g x\equiv \mathrm{ind}_g 37=32 \pmod{40}$$

等价，即

$$3\mathrm{ind}_g x\equiv 8 \pmod{10},$$

由于 3 的逆元是 7，所以两边同时乘以 7 得到

$$\mathrm{ind}_g x\equiv 56\equiv 6 \pmod{10},$$

可解得

$$\mathrm{ind}_g x\equiv 6,16,26,36 \pmod{40},$$

故通过查模 41 的指数表可得到原同余方程的解为

$$x\equiv 39,18,2,23 \pmod{41}.$$

3.2 二次剩余

3.2.1 二次剩余的概念和性质

到目前为止，人们还没有找到具有普遍性的有效方法来求解一般的多项式同余方程. 除了求根方法的问题以外，还有一个与此有关的问题，即在没有求出方程的根的时候，是否存

在一个有效的方法来判断方程的可解性，也就是说判断方程有没有解. 二次同余方程在后面这个问题上有比较丰富的理论，其核心就是本节的重点——二次剩余和二次互反律.

在 3.1 节中，我们给出了 n 次剩余的定义. 其中当 $n=2$ 时，就得到二次剩余的定义. 显然，设 m 是大于 1 的整数，a 是与 m 互素的整数，若

$$x^2\equiv a \pmod m \tag{3.2.1}$$

有解，则 a 叫作模 m 的**二次剩余**，或**平方剩余**. 否则，a 叫作模 m 的**二次非剩余**，或**平方非剩余**.

下面关于一般形式的二次同余方程的讨论将使我们看到二次同余方程的可解性与二次剩余的概念是紧密联系在一起的.

考虑下面的二次同余方程

$$ax^2+bx+c\equiv 0 \pmod p \tag{3.2.2}$$

其中 p 是一个奇素数且 $a\not\equiv 0 \pmod p$，即 $(a,p)=1$. 所以 $(4a,p)=1$. 因此方程(3.2.2)与下面的方程等价

$$4a(ax^2+bx+c)\equiv 0 \pmod p,$$

即

$$(2ax+b)^2-(b^2-4ac)\equiv 0 \pmod p,$$

移项后得到

$$(2ax+b)^2\equiv (b^2-4ac) \pmod p.$$

现在，令 $y=2ax+b$，$d=b^2-4ac$，则得到

$$y^2\equiv d \pmod p \tag{3.2.3}$$

如果 $x\equiv x_0 \pmod p$ 是方程(3.2.2)的一个解，那么任意整数 $y_0\equiv 2ax_0+b \pmod p$ 就是方程(3.2.3)的解. 反过来，如果 $y\equiv y_0 \pmod p$ 是方程(3.2.3)的一个解，那么下面的线性同余方程

$$2ax\equiv y_0-b \pmod p$$

的解

$$x\equiv x_0=(2a)^{-1}(y_0-b) \pmod p$$

就是原方程(3.2.2)的一个解.

[例 3.2.1]　求解二次同余方程 $5x^2-6x+2\equiv 0 \pmod{13}$.

解　$d=b^2-4ac=36-40=-4$，因此需要先解如下的具有简单形式的二次同余方程

$$y^2\equiv -4\equiv 9 \pmod{13},$$

它的解是 $y\equiv 3, 10 \pmod{13}$. 接着需要分别求解两个线性同余方程

$$10x\equiv 9 \pmod{13},$$

和

$$10x\equiv 16 \pmod{13}.$$

由于 10 的逆元是 4，所以这两个方程的解分别为 $x\equiv 10, 12 \pmod{13}$. 这两个解就是原方程的解.

上面的讨论说明模数为奇素数的一般形式的二次同余方程(3.2.2)的可解性与 b^2-4ac 是否为二次剩余的问题是等价的. 根据第 2 章的高次同余方程的内容可知，对于一般的模数来说，总可以将方程化为模数为素数幂的联立方程组，同时模数为素数幂的方程的解可以通

过模数为素数的方程的解求得，此外模数为 2 的二次同余方程求解非常简单，因此，讨论模数为奇素数的方程(3.2.2)的可解性是至关重要的. 相应地，我们将着重讨论模数为奇素数的二次剩余问题，即

$$x^2 \equiv a \pmod{p}, \tag{3.2.4}$$

其中 p 是奇素数.

[例 3.2.2] 求模 13 的二次剩余和二次非剩余.

解 首先，如果 $a \equiv b \pmod{13}$，那么 a 是模 13 的二次剩余当且仅当 b 是模 13 的二次剩余. 因此，只需要在 1 到 12 的范围内找模 13 的二次剩余. 通过计算得到

$$1^2 \equiv 12^2 \equiv 1 \pmod{13},$$
$$2^2 \equiv 11^2 \equiv 4 \pmod{13},$$
$$3^2 \equiv 10^2 \equiv 9 \pmod{13},$$
$$4^2 \equiv 9^2 \equiv 3 \pmod{13},$$
$$5^2 \equiv 8^2 \equiv 12 \pmod{13},$$
$$6^2 \equiv 7^2 \equiv 10 \pmod{13},$$

所以，模 13 的二次剩余是 1,3,4,9,10,12. 当然，模 13 的二次非剩余是 2,5,6,7,8,11.

下面，给出二次剩余的**欧拉判别条件**，即定理 3.2.1.

定理 3.2.1 设 p 是奇素数，$(a, p)=1$，则

(1) a 是模 p 的二次剩余的充要条件是

$$a^{\frac{p-1}{2}} \equiv 1 \pmod{p};$$

(2) a 是模 p 的二次非剩余的充要条件是

$$a^{\frac{p-1}{2}} \equiv -1 \pmod{p}.$$

并且当 a 是模 p 的二次剩余时，同余方程(3.2.4)恰有二解.

证明 (1) 先证必要性. 若 a 是模 p 的二次剩余，则有整数 x 满足

$$x^2 \equiv a \pmod{p}.$$

因为$(a,p)=1$，所以$(x,p)=1$，

应用欧拉定理，可知

$$a^{\frac{p-1}{2}} \equiv (x^2)^{\frac{p-1}{2}} \equiv x^{p-1} \equiv 1 \pmod{p}.$$

再证充分性. 为了确定方程(3.2.4)的解的数量，需要应用定理 2.2.37，即需要考察多项式 $x^p - x$ 除以多项式 $x^2 - a$ 所得余多项式的系数.

因为

$$\begin{aligned} x^p - x &= x((x^2)^{\frac{p-1}{2}} - a^{\frac{p-1}{2}}) + (a^{\frac{p-1}{2}} - 1)x \\ &= (x^2 - a)xq(x) + (a^{\frac{p-1}{2}} - 1)x \end{aligned}$$

其中 $q(x)$是 x 的整系数多项式，所以待求的余多项式的系数为 $a^{\frac{p-1}{2}} - 1$. 由题设知道

$$a^{\frac{p-1}{2}} \equiv 1 \pmod{p},$$

即该余多项式的唯一系数 $a^{\frac{p-1}{2}} - 1$ 能被 p 整除，根据定理 2.2.37，可知

$$x^2 - a \equiv 0 \pmod{p},$$

也就是

$$x^2 \equiv a \pmod{p}$$

有二解，于是 a 是模 p 的二次剩余.

(2) 由于 a 与 p 互素，根据欧拉定理，可知

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

即 $p \mid a^{p-1} - 1$. 又因为

$$a^{p-1} - 1 = (a^{\frac{p-1}{2}} + 1)(a^{\frac{p-1}{2}} - 1),$$

所以

$$p \mid a^{\frac{p-1}{2}} - 1 \text{ 或 } p \mid a^{\frac{p-1}{2}} + 1.$$

根据(1)的证明，可知 a 是模 p 的二次非剩余的充要条件是

$$p \mid a^{\frac{p-1}{2}} + 1,$$

即

$$a^{\frac{p-1}{2}} \equiv -1 \pmod{p}.$$

证毕.

[例 3.2.3]　利用欧拉判别条件判断 2 和 3 是否为模 13 的二次剩余或者二次非剩余.

解　由于 $2^{\frac{(13-1)}{2}} = 2^6 = 64 \equiv 12 \equiv -1 \pmod{13}$，所以 2 是模 13 的二次非剩余. 而 $3^{\frac{(13-1)}{2}} = 3^6 = 27^2 \equiv 1^2 \equiv 1 \pmod{13}$，所以 3 是模 13 的二次剩余. 此时，$x^2 \equiv 3 \pmod{13}$必有两个解，在例 3.2.2 中我们已经知道解为 4 和 9.

定理 3.2.2　设 p 是奇素数，则模 p 的缩系中二次剩余与非二次剩余的个数各为$\frac{p-1}{2}$，且$\frac{p-1}{2}$个二次剩余分别与序列

$$1^2, 2^2, \cdots, \left(\frac{p-1}{2}\right)^2 \tag{3.2.5}$$

中的一个数同余，且仅与一个数同余.

证明　由定理 3.2.1 可知二次剩余的个数等于同余方程

$$x^{\frac{p-1}{2}} \equiv 1 \pmod{p}$$

的解数. 由于

$$x^{\frac{p-1}{2}} - 1 \mid x^p - x,$$

根据定理 2.2.37，可知二次剩余的个数是$\frac{p-1}{2}$. 于是二次非剩余的个数是

$$(p-1) - \frac{p-1}{2} = \frac{p-1}{2}.$$

显然式(3.2.5)中的数都是二次剩余，现证明它们互不同余. 用反证法，假设存在

$$k^2 \equiv l^2 \pmod{p},$$

其中 $1 \leqslant k < l \leqslant \frac{p-1}{2}$，则 $\pm k$，$\pm l$ 显然都满足下面的同余方程

$$x^2 \equiv l^2 \pmod{p}.$$

因此，整数 k，l，$p-l$，$p-k$ 也必然同时满足上面的同余方程. 而它们全都处在一个完全剩余系中，即 $1 \leqslant k < l \leqslant \frac{p-1}{2} < p-l < p-k \leqslant p-1$，说明上面的方程有四个互不同余的解，这

与定理 3.2.1 矛盾，于是假设不成立. 定理得证.

例 3.2.2 很好地验证了这个定理.

3.2.2 勒让德符号与二次互反律

以上虽然给出了模 p 的二次剩余的欧拉判别条件，但是当 p 比较大时，很难实际应用. 现在引入由大数学家勒让德发明的勒让德符号，以此给出一个比较便于实际计算的判别方法.

定义 3.2.1 设 p 是奇素数，$(a, p)=1$，定义**勒让德(Legendre)符号**如下：

$$\left(\frac{a}{p}\right)=\begin{cases}1, & \text{若 } a \text{ 是模 } p \text{ 的二次剩余；}\\ -1, & \text{若 } a \text{ 是模 } p \text{ 的二次非剩余.}\end{cases}$$

注：$\left(\frac{a}{p}\right)$读作 a 对 p 的勒让德符号.

[例 3.2.4] 利用例 3.2.2 写出对 13 的勒让德符号.

解

$$\left(\frac{1}{13}\right)=\left(\frac{3}{13}\right)=\left(\frac{4}{13}\right)=\left(\frac{9}{13}\right)=\left(\frac{10}{13}\right)=\left(\frac{12}{13}\right)=1,$$

$$\left(\frac{2}{13}\right)=\left(\frac{5}{13}\right)=\left(\frac{6}{13}\right)=\left(\frac{7}{13}\right)=\left(\frac{8}{13}\right)=\left(\frac{11}{13}\right)=-1.$$

利用勒让德符号，可以将定理 3.2.1 改写如下.

定理 3.2.1* 设 p 是奇素数，a 是与 p 互素的整数，则

$$\left(\frac{a}{p}\right)\equiv a^{\frac{p-1}{2}}\pmod p.$$

显然，我们有$\left(\frac{1}{p}\right)=1$.

进一步，可以得出有关勒让德符号的一些性质.

定理 3.2.3 设 p 是奇素数，a，b 都是与 p 互素的整数，有

(1) 若 $a\equiv b \pmod p$，则$\left(\frac{a}{p}\right)=\left(\frac{b}{p}\right)$；

(2) $\left(\frac{ab}{p}\right)=\left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$；

(3) $\left(\frac{a^2}{p}\right)=1$.

证明 (1) 因为 $a\equiv b \pmod p$，所以同余方程

$$x^2\equiv a \pmod p$$

等价于同余方程

$$x^2\equiv b \pmod p.$$

因此

$$\left(\frac{a}{p}\right)=\left(\frac{b}{p}\right).$$

(2) 根据欧拉判别条件，有

$$\left(\frac{a}{p}\right)\equiv a^{\frac{p-1}{2}}\pmod p,\quad \left(\frac{b}{p}\right)\equiv b^{\frac{p-1}{2}}\pmod p,\quad \left(\frac{ab}{p}\right)\equiv (ab)^{\frac{p-1}{2}}\pmod p.$$

因此

$$\left(\frac{ab}{p}\right)\equiv(ab)^{\frac{p-1}{2}}=a^{\frac{p-1}{2}}b^{\frac{p-1}{2}}\equiv\left(\frac{a}{p}\right)\left(\frac{b}{p}\right)\pmod{p}.$$

由于勒让德符号取值只有±1，且 p 是奇素数，故

$$\left(\frac{ab}{p}\right)=\left(\frac{a}{p}\right)\left(\frac{b}{p}\right).$$

(3) 显然，a^2 是模 p 的二次剩余，所以必有

$$\left(\frac{a^2}{p}\right)=1.$$

当 $a=\pm 2^k q_1^{l_1} q_2^{l_2}\cdots q_s^{l_s}$，其中 $q_i(i=1,2,\cdots,s)$为不同的奇素数，根据上面的定理有

$$\left(\frac{a}{p}\right)=\left(\frac{\pm 1}{p}\right)\left(\frac{2}{p}\right)^k\left(\frac{q_1}{p}\right)^{l_1}\cdots\left(\frac{q_s}{p}\right)^{l_s}.$$

因为$\left(\frac{1}{p}\right)=1$，所以任给一个与 p 互素的整数 a，计算$\left(\frac{a}{p}\right)$时，只需算出以下三种值：

$$\left(\frac{-1}{p}\right),\quad\left(\frac{2}{p}\right),\quad\left(\frac{q}{p}\right)(q\text{ 为奇素数}).$$

需要注意的是，这种计算方法依赖于对 a 的因子分解，而目前还没有找到高效的因子分解方法，因此这里的勒让德符号的计算方法对大的模数 p 和整数 a 来说不切实际. 更好的方法将在 3.2.3 节中介绍.

根据欧拉判别条件，显然可得出以下定理.

定理 3.2.4　设 p 是奇素数，有

$$\left(\frac{-1}{p}\right)=(-1)^{\frac{p-1}{2}}=\begin{cases}1,\text{若 } p\equiv 1\pmod 4,\\-1,\text{若 } p\equiv 3\pmod 4.\end{cases}$$

[例 3.2.5]　判断 $x^2\equiv -46\pmod{17}$是否有解.

解　$$\left(\frac{-46}{17}\right)=\left(\frac{-1}{17}\right)\left(\frac{46}{17}\right)=\left(\frac{46}{17}\right)=\left(\frac{17\times 2+12}{17}\right)=\left(\frac{12}{17}\right)=\left(\frac{3}{17}\right)\left(\frac{2^2}{17}\right)=\left(\frac{3}{17}\right),$$

而$\left(\frac{3}{17}\right)\equiv 3^{\frac{17-1}{2}}=3^8=81^2\equiv -1\pmod{17}$，所以原方程无解.

德国大数学家高斯(Carl Friedrich Gauss，1777—1855)证明了以下结果，称为**高斯引理**. 它是证明二次互反律的基础.

引理 3.2.1　设 p 是奇素数，a 是与 p 互素的整数，如果下列$\frac{p-1}{2}$个整数

$$a\cdot 1,a\cdot 2,a\cdot 3,\cdots,a\cdot\frac{p-1}{2}$$

对模 p 化简后得到的最小正剩余中大于$\frac{p}{2}$的个数是 m，则

$$\left(\frac{a}{p}\right)=(-1)^m.$$

证明　设 $a_1,a_2,\cdots,a_l$ 是整数

$$a\cdot 1,a\cdot 2,a\cdot 3,\cdots,a\cdot\frac{p-1}{2}$$

中对模 p 化简后小于$\frac{p}{2}$的最小正剩余，b_1，b_2，…，b_m 是这些整数中对模 p 化简后大于$\frac{p}{2}$的

最小正剩余，显然

$$l+m=\frac{p-1}{2},$$

则原来的$\frac{p-1}{2}$个整数之积和相应的最小正剩余之间具有如下关系

$$a^{\frac{p-1}{2}}\left(\frac{p-1}{2}\right)!=\prod_{k=1}^{\frac{p-1}{2}}ak\equiv\prod_{i=1}^{l}a_i\prod_{j=1}^{m}b_j\equiv(-1)^m\prod_{i=1}^{l}a_i\prod_{j=1}^{m}(p-b_j)\pmod p.$$

下面证明 $a_1, a_2, \cdots, a_l, p-b_1, p-b_2, \cdots, p-b_m$ 两两互不相等，这只需证明

$$a_s\neq p-b_t,\ s=1,2,\cdots,l,\ t=1,2,\cdots,m.$$

用反证法，假设存在

$$a_s=p-b_t,$$

则有

$$ak_i\equiv p-ak_j\pmod p,$$

即

$$ak_i+ak_j\equiv0\pmod p,$$

于是

$$k_i+k_j\equiv0\pmod p,$$

即有 $p\mid k_i+k_j$.

因为

$$1\leqslant k_i\leqslant\frac{p-1}{2},\ i=1,2,\cdots,\frac{p-1}{2},$$

所以

$$1\leqslant k_i+k_j\leqslant\frac{p-1}{2}+\frac{p-1}{2}<p,$$

这与 $p\mid k_i+k_j$ 矛盾，故假设不成立. 因此，$a_1, a_2, \cdots, a_l, p-b_1, p-b_2, \cdots, p-b_m$ 这$\frac{p-1}{2}$个整数两两互不相等.

由于

$$1\leqslant a_s\leqslant\frac{p-1}{2},s=1,2,\cdots,l,\quad 1\leqslant p-b_t\leqslant\frac{p-1}{2},t=1,2,\cdots,m,$$

故 $a_1,a_2,\cdots,a_l,p-b_1,p-b_2,\cdots,p-b_m$ 这$\frac{p-1}{2}$个整数是 $1,2,\cdots,\frac{p-1}{2}$的一个排列，于是

$$a^{\frac{p-1}{2}}\left(\frac{p-1}{2}\right)!\equiv(-1)^m\prod_{i=1}^{l}a_i\prod_{j=1}^{m}(p-b_j)=(-1)^m\left(\frac{p-1}{2}\right)!\pmod p,$$

则

$$a^{\frac{p-1}{2}}\equiv(-1)^m\pmod p.$$

再根据欧拉判别条件，有

$$\left(\frac{a}{p}\right)=(-1)^m.$$

证毕.

［例 3.2.6］ 利用高斯引理判断 5 是否为模 13 的二次剩余.

解　按照高斯引理，首先得到(13−1)/2=6 个整数，即 5，10，15，20，25，30，模 13 化简得到的最小正剩余为 5，10，2，7，12，4，其中三个大于 13/2，所以

$$\left(\frac{5}{13}\right)=(-1)^3=-1,$$

即 5 不是模 13 的二次剩余.

定理 3.2.5　设 p 是奇素数，有

(1) $\left(\dfrac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}=\begin{cases}1,\text{若 } p\equiv\pm1(\bmod 8),\\-1,\text{若 } p\equiv\pm3(\bmod 8).\end{cases}$

(2) 若$(a,2p)=1$，则 $\left(\dfrac{a}{p}\right)=(-1)^{\sum\limits_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]}$.

证明　由于当$(a,p)=1$ 时，

$$ak=p\left[\frac{ak}{p}\right]+r_k,\ 0<r_k<p,\ k=1,2,\cdots,\frac{p-1}{2},$$

我们对 $k=1,2,\cdots,\dfrac{p-1}{2}$求和，并利用高斯引理的证明，有

$$\begin{aligned}a\frac{p^2-1}{8}&=p\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]+\sum_{i=1}^{l}a_i+\sum_{j=1}^{m}b_j\\&=p\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]+\sum_{i=1}^{l}a_i+\sum_{j=1}^{m}(p-b_j)+2\sum_{j=1}^{m}b_j-mp\\&=p\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]+\frac{p^2-1}{8}-mp+2\sum_{j=1}^{m}b_j\end{aligned}$$

于是，

$$(a-1)\frac{p^2-1}{8}=p\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]-mp+2\sum_{j=1}^{m}b_j.$$

因为对每个奇素数 p，都有正整数 d 使

$$p=2d+1,$$

则有

$$(a-1)\frac{p^2-1}{8}=\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]+m+2\left(\sum_{j=1}^{m}b_j+d\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]-(d+1)m\right),$$

因此有

$$(a-1)\frac{p^2-1}{8}\equiv\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]+m\ (\bmod 2).$$

(1) 若取 $a=2$，则有 $0\leqslant\left[\dfrac{ak}{p}\right]\leqslant\left[\dfrac{p-1}{p}\right]=0$，因此

$$m\equiv\frac{p^2-1}{8}\ (\bmod 2),$$

即存在整数 v 使得

$$m=2v+\frac{p^2-1}{8}.$$

再根据高斯引理，可知

$$\left(\frac{2}{p}\right)=(-1)^m=(-1)^{\frac{p^2-1}{8}}.$$

若 $p\equiv\pm1\pmod 8$，则存在正整数 d 使得 $p=8d\pm1$，从而

$$\left(\frac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}=(-1)^{2(4d^2\pm d)}=1.$$

若 $p\equiv\pm3\pmod 8$，则存在正整数 d 使得 $p=8d\pm3$，从而

$$\left(\frac{2}{p}\right)=(-1)^{\frac{p^2-1}{8}}=(-1)^{2(4d^2\pm3d)+1}=-1.$$

(2) 若取 a 为奇数，即$(a,2p)=1$，则 $a-1\equiv0\pmod 2$，因此有

$$\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]+m\equiv0\pmod 2,$$

所以上式中两个加数必然同为奇数或者偶数，即

$$m\equiv\sum_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]\pmod 2.$$

再根据高斯引理，可知

$$\left(\frac{a}{p}\right)=(-1)^m=(-1)^{\sum\limits_{k=1}^{\frac{p-1}{2}}\left[\frac{ak}{p}\right]}.$$

[例 3.2.7] 利用上面定理判断 5 是否为模 13 的二次剩余.

解 首先计算$(13-1)/2=6$ 个整数，即[5/13],[10/13],[15/13],[20/13],[25/13],[30/13]，分别为 0,0,1,1,1,2，所以

$$\left(\frac{5}{13}\right)=(-1)^{0+0+1+1+1+2}=-1,$$

即 5 不是模 13 的二次剩余.

下面给出著名的**二次互反律**.

定理 3.2.6 设 p,q 是奇素数，$p\neq q$，则

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}.$$

证明 因为 p,q 是奇素数，所以

$$(q,2p)=1,\quad (p,2q)=1,$$

于是分别有

$$\left(\frac{q}{p}\right)=(-1)^{\sum\limits_{h=1}^{\frac{p-1}{2}}\left[\frac{qh}{p}\right]},\quad \left(\frac{p}{q}\right)=(-1)^{\sum\limits_{k=1}^{\frac{q-1}{2}}\left[\frac{pk}{q}\right]},$$

因此只需证明

$$\sum_{h=1}^{\frac{p-1}{2}}\left[\frac{qh}{p}\right]+\sum_{k=1}^{\frac{q-1}{2}}\left[\frac{pk}{q}\right]=\frac{p-1}{2}\cdot\frac{q-1}{2}$$

即可.

考察长为$\frac{p}{2}$、宽为$\frac{q}{2}$的长方形内的整数点个数，如图 3.2.1 所示.

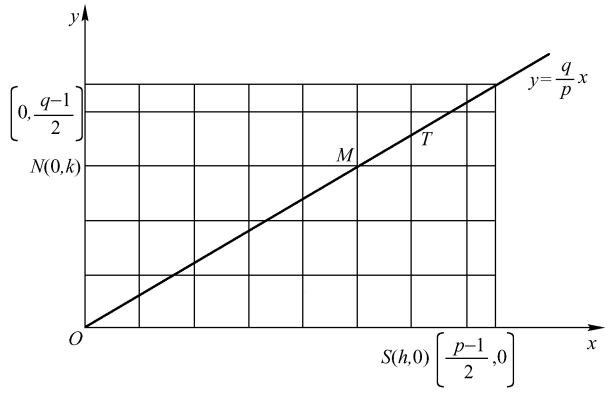

图 3.2.1　长为$\frac{p}{2}$，宽为$\frac{q}{2}$的长方形内的整数点个数

设点 S 的坐标为$(h,0)$，点 T 是直线 $x=h$ 与直线 $y=\frac{q}{p}x$ 的交点，其中 h 为整数，且

$$0\leqslant h\leqslant\frac{p-1}{2}.$$

设点 N 的坐标为$(0,k)$，点 M 是直线 $y=k$ 与直线 $y=\frac{q}{p}x$ 的交点，其中 k 为整数，且

$$0\leqslant k\leqslant\frac{q-1}{2}.$$

在垂直直线 ST 上，整数点个数为$\left[\frac{qh}{p}\right]$，于是，下三角形内的整数点个数为 $\sum\limits_{h=1}^{\frac{p-1}{2}}\left[\frac{qh}{p}\right]$. 在水平直线 NM 上，整数点个数为$\left[\frac{pk}{q}\right]$，于是，上三角形内的整数点个数为 $\sum\limits_{k=1}^{\frac{p-1}{2}}\left[\frac{pk}{q}\right]$. 因为对角线上除原点外无整数点，所以长方形内整数点个数为

$$\sum_{h=1}^{\frac{p-1}{2}}\left[\frac{qh}{p}\right]+\sum_{k=1}^{\frac{q-1}{2}}\left[\frac{pk}{q}\right]=\frac{p-1}{2}\cdot\frac{q-1}{2}.$$

证毕.

在实际应用中，有时也把二次互反律写为如下形式：

$$\left(\frac{q}{p}\right)=(-1)^{\frac{p-1}{2}\cdot\frac{q-1}{2}}\left(\frac{p}{q}\right).$$

[例 3.2.8]　2 和 3 是否模 17 的二次剩余？

解　由于

$$\left(\frac{2}{17}\right)=(-1)^{\frac{17^2-1}{8}}=(-1)^{2\cdot 18}=1,$$

故 2 是模 17 的二次剩余.

由二次互反律，有

$$\left(\frac{3}{17}\right)=(-1)^{\frac{3-1}{2}\cdot\frac{17-1}{2}}\left(\frac{17}{3}\right)=\left(\frac{17}{3}\right)=\left(\frac{-1}{3}\right)=(-1)^{\frac{3-1}{2}}=-1,$$

故 3 是模 17 的二次非剩余.

[例 3.2.9] 同余方程

$$x^2\equiv 137\ (\mathrm{mod}\ 227)$$

是否有解?

解 因为 227 为素数,则

$$\left(\frac{137}{227}\right)=\left(\frac{-90}{227}\right)=\left(\frac{-1}{227}\right)\left(\frac{2\cdot 3^2\cdot 5}{227}\right)=-\left(\frac{2}{227}\right)\left(\frac{5}{227}\right),$$

而

$$\left(\frac{2}{227}\right)=(-1)^{\frac{227^2-1}{8}}=(-1)^{\frac{226\cdot 228}{8}}=-1,$$

又由二次互反律,有

$$\left(\frac{5}{227}\right)=(-1)^{\frac{5-1}{2}\cdot\frac{227-1}{2}}\left(\frac{227}{5}\right)=\left(\frac{227}{5}\right)=\left(\frac{2}{5}\right)=(-1)^{\frac{5^2-1}{8}}=-1,$$

因此,

$$\left(\frac{137}{227}\right)=-1,$$

即原同余方程无解.

3.2.3 雅可比符号

定义 3.2.2 设正奇数 $m=p_1\ p_2\cdots p_r$ 是奇素数 $p_i(i=1,2,\cdots,r)$的乘积,a 是与 m 互素的整数,定义**雅可比(Jacobi)符号**如下:

$$\left(\frac{a}{m}\right)=\left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right).$$

从形式上看,雅可比符号只是将勒让德符号中的素数 p 推广到了正奇数 m,但其意义就不相同了. 我们知道,若 a 对 p 的勒让德符号为 1,则可知 a 是模 p 的二次剩余,但当 a 对 m 的雅可比符号为 1 时,却不能判断 a 是模 m 的二次剩余. 例如,3 是模 119 的二次非剩余,但

$$\left(\frac{3}{119}\right)=\left(\frac{3}{7}\right)\left(\frac{3}{17}\right)=-\left(\frac{1}{3}\right)\left(\frac{-1}{3}\right)=(-1)(-1)=1.$$

下面来分析雅可比符号的一些性质.

显然,有$\left(\frac{1}{m}\right)=\left(\frac{1}{p_1}\right)\left(\frac{1}{p_2}\right)\cdots\left(\frac{1}{p_r}\right)=1.$

定理 3.2.7 设 m 是正奇数,a,b 都是与 m 互素的整数,有

(1) 若 $a\equiv b(\mathrm{mod}\ m)$,则$\left(\frac{a}{m}\right)=\left(\frac{b}{m}\right)$;

(2) $\left(\frac{ab}{m}\right)=\left(\frac{a}{m}\right)\left(\frac{b}{m}\right)$;

(3) $\left(\frac{a^2}{m}\right)=1.$

证明 设 $m=p_1p_2\cdots p_r$,其中 $p_i(i=1,2,\cdots,r)$是奇素数.

(1) 因为 $a\equiv b(\bmod p)$，所以

$$\left(\frac{a}{m}\right)=\left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right)=\left(\frac{b}{p_1}\right)\left(\frac{b}{p_2}\right)\cdots\left(\frac{b}{p_r}\right)=\left(\frac{b}{m}\right).$$

(2)

$$\begin{aligned}\left(\frac{ab}{m}\right)&=\left(\frac{ab}{p_1}\right)\left(\frac{ab}{p_2}\right)\cdots\left(\frac{ab}{p_r}\right)\\&=\left(\frac{a}{p_1}\right)\left(\frac{b}{p_1}\right)\left(\frac{a}{p_2}\right)\left(\frac{b}{p_2}\right)\cdots\left(\frac{a}{p_r}\right)\left(\frac{b}{p_r}\right)\\&=\left(\frac{a}{p_1}\right)\left(\frac{a}{p_2}\right)\cdots\left(\frac{a}{p_r}\right)\left(\frac{b}{p_1}\right)\left(\frac{b}{p_2}\right)\cdots\left(\frac{b}{p_r}\right)\\&=\left(\frac{a}{m}\right)\left(\frac{b}{m}\right)\end{aligned}$$

(3)

$$\left(\frac{a^2}{m}\right)=\left(\frac{a^2}{p_1}\right)\left(\frac{a^2}{p_2}\right)\cdots\left(\frac{a^2}{p_r}\right)=1.$$

定理 3.2.8　设 m 是正奇数，有

(1) $\left(\dfrac{-1}{m}\right)=(-1)^{\frac{m-1}{2}}$；

(2) $\left(\dfrac{2}{m}\right)=(-1)^{\frac{m^2-1}{8}}$.

证明　设 $m=p_1p_2\cdots p_r$，其中 $p_i(i=1,2,\cdots,r)$是奇素数.

(1) 因为

$$m=\prod_{i=1}^{r}p_i=\prod_{i=1}^{r}(1+p_i-1)\equiv 1+\sum_{i=1}^{r}(p_i-1)(\bmod 4),$$

则有

$$\frac{m-1}{2}\equiv\sum_{i=1}^{r}\frac{p_i-1}{2}(\bmod 2),$$

于是

$$\left(\frac{-1}{m}\right)=\prod_{i=1}^{r}\left(\frac{-1}{p_i}\right)=(-1)^{\sum\limits_{i=1}^{r}\frac{p_i-1}{2}}=(-1)^{\frac{m-1}{2}}.$$

(2) 因为

$$m^2=\prod_{i=1}^{r}p_i^2=\prod_{i=1}^{r}(1+p_i^2-1)\equiv 1+\sum_{i=1}^{r}(p_i^2-1)\ (\bmod 16),$$

则有

$$\frac{m^2-1}{8}\equiv\sum_{i=1}^{r}\frac{p_i^2-1}{8}\ (\bmod 2),$$

于是

$$\left(\frac{2}{m}\right)=\prod_{i=1}^{r}\left(\frac{2}{p_i}\right)=(-1)^{\sum\limits_{i=1}^{r}\frac{p_i^2-1}{8}}=(-1)^{\frac{m^2-1}{8}}.$$

定理 3.2.9　设 m，n 是互素的正奇数，则

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right)=(-1)^{\frac{m-1}{2}\cdot\frac{n-1}{2}}.$$

证明 设 $m=p_1p_2\cdots p_r$，$n=q_1q_2\cdots q_s$，其中 $p_i(i=1,2,\cdots,r)$，$q_j(j=1,2,\cdots,s)$都是奇素数，则

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right)=\prod_{j=1}^{s}\left(\frac{m}{q_j}\right)\prod_{i=1}^{r}\left(\frac{n}{p_i}\right)=\prod_{i=1}^{r}\prod_{j=1}^{s}\left(\frac{q_j}{p_i}\right)\left(\frac{p_i}{q_j}\right)=(-1)^{\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{p_i-1}{2}\cdot\frac{q_j-1}{2}}$$

由定理 3.2.8 中的证明可知

$$\sum_{i=1}^{r}\frac{p_i-1}{2}\equiv\frac{m-1}{2}\pmod 2,$$

则

$$\sum_{i=1}^{r}\sum_{j=1}^{s}\frac{p_i-1}{2}\cdot\frac{q_j-1}{2}=\sum_{i=1}^{r}\frac{p_i-1}{2}\sum_{j=1}^{s}\frac{q_j-1}{2}\equiv\frac{m-1}{2}\cdot\frac{n-1}{2}\pmod 2,$$

所以

$$\left(\frac{m}{n}\right)\left(\frac{n}{m}\right)=(-1)^{\frac{m-1}{2}\cdot\frac{n-1}{2}}.$$

在实际应用中，有时也可把上式写为如下形式：

$$\left(\frac{n}{m}\right)=(-1)^{\frac{m-1}{2}\cdot\frac{n-1}{2}}\left(\frac{m}{n}\right).$$

通过上面这些定理，我们发现雅可比符号具有和勒让德符号一样的计算法则，于是当 m 为正奇数时，不必再把 m 分解成素因子的乘积，所以计算起来更方便.

[例 3.2.10] 同余方程

$$x^2\equiv 286\pmod{563}$$

是否有解？

解 我们用辗转相除法求得$(286,563)=1$，于是不必考虑 563 是否为素数即可计算雅可比符号，即

$$\left(\frac{286}{563}\right)=\left(\frac{2}{563}\right)\left(\frac{143}{563}\right)=(-1)^{\frac{563^2-1}{8}}(-1)^{\frac{143-1}{2}\cdot\frac{563-1}{2}}\left(\frac{563}{143}\right)=\left(\frac{-9}{143}\right)=\left(\frac{-1}{143}\right)=-1,$$

所以原同余方程无解.

实际上，由雅可比符号的定义，很容易证明，当 a 是模 m 的二次剩余时，则有$\left(\frac{a}{m}\right)=1$必然成立，所以，当$\left(\frac{a}{m}\right)=-1$ 时，a 一定是模 m 的二次非剩余. 但是，正如前面所述，$\left(\frac{a}{m}\right)=1$ 不一定说明 a 是模 m 的二次剩余.

通俗地讲，前面的讨论都是关于如何判断一个整数是否具有模 p(或者 m) 的平方根问题的，在这一节的最后我们针对一种特殊情况给出明确的求平方根的计算公式.

定理 3.2.10 素数 $p\equiv 3\pmod 4$，且 a 为模 p 的二次剩余，则 $\pm a^{\frac{p+1}{4}}$ 为 a 的模 p 平方根.

证明 由欧拉判别条件可以推得

$$(\pm a^{\frac{p+1}{4}})^2=a^{\frac{p+1}{2}}=a^{\frac{p-1}{2}}a\equiv 1a=a\pmod p$$

且 $\pm a^{\frac{p+1}{4}}$ 是仅有的两个解，即 $\pm a^{\frac{p+1}{4}}$ 为 a 的模 p 平方根.

[例 3.2.11] Rabin 密码体制中，由明文 x 按下式计算密文

$$y=x^2 \bmod 77,$$

相应地，借用平方根符号可以将解密过程表示为

$$x=\sqrt{y} \bmod 77.$$

如果密文为 $y=23$，为了解密需要先求 23 对模 7 和模 11 的平方根. 因为 7 和 11 都是符合上面定理题设的素数，所以，利用公式得到这两个平方根(此处只写出“正”解)。

$$23^{\frac{7+1}{4}}=23^2\equiv 2^2\equiv 4(\bmod 7),$$

$$23^{\frac{11+1}{4}}=23^3\equiv 1^3\equiv 1(\bmod 11).$$

再利用中国剩余定理计算得到明文的四个可能值，$x=10$，32，45，67.

注：由于该密码体制的加密过程本身是一个多对一的函数，所以解密过程必然得到多个解，因此，在实际使用的时候，需要额外的冗余来保证恢复正确的那一个明文.

3.3　数论的典型应用

3.3.1　素性检验算法

由于素数在数论中占有特别重要的地位，因此如何将素数与合数区分开来的方法，以及如何将一个合数分解为素因子之积的方法一直是数论的重要的基本问题. 所谓素性检验，就是要判断给定的正整数 n 是不是素数. 尽管素性检验是数论中一个古老的问题，随着近代密码学的发展，对这个问题的研究不再仅仅局限在理论上，而是具有了实实在在的应用价值，其中最典型的应用就是为许多重要的密码学算法生成具有十进制百位以上长度的大素数.

目前的素性检验算法大体可以分为以下三种不同的类型。

(1) 能够对所有素数进行检验，而且给出的答案一定是严格正确的，即“确定性检验”. 例如，在 2.1 节中，基于定理 2.1.7，即若 n 为合数，则 n 必有素因子 p 满足 $p\leqslant\sqrt{n}$，我们介绍了 Eratosthenes 筛法. 于是，要判断一个大于 1 的整数 n 是否素数，只要用不大于$\sqrt{n}$的所有素数去试除 n，如果其中有一个素数能整除 n，则 n 是合数，否则 n 是素数. 这种素性检验的方法称为试除法，显然它的计算复杂度是很高的，尤其当 n 很大时，其计算量是难以接受的. 计算量太大是这一类型算法的共同问题. 因此，需要寻找一些其他效率更高的素性检验方法.

(2) 能够对所有素数进行检验，但是给出的答案具有一定的错误概率，即“概率性检验”. 这种类型的算法一般具有合理长度的运行时间，而且不论测试哪个数，错误概率都能够控制到充分小的范围.

(3) 只能够对具有特殊形式的数进行检验，如梅森数和费马数等. 这种类型的算法能够高效率地得到正确的答案，在发现已知的最大素数的实际应用中取得了巨大的成功，但是，由于缺少通用性，不适合在密码学环境中使用.

这里，主要讨论前两种类型的算法.

定理 3.3.1　(Lucas 检验) 如果存在整数 a，使得 $a^{n-1}\equiv 1 \pmod n$，并且对 $n-1$ 的任意素数因子 p，$a^{\frac{n-1}{p}}\not\equiv 1 \pmod n$成立，那么 n 是素数.

证明　由 $a^{n-1}\equiv 1 \pmod n$，可知 $\mathrm{ord}(a)\mid n-1$. 设 $n-1=j\ \mathrm{ord}(a)$，其中 j 为正整数. 假设 $j>1$，那么 j 必然有一个素数因子 q，从而存在整数 h 使得 $j=qh$. 则

$$a^{\frac{n-1}{q}}=a^{\frac{j\mathrm{ord}(a)}{q}}=a^{h\mathrm{ord}(a)}=(a^{\mathrm{ord}(a)})^h\equiv 1\ (\mathrm{mod}\ n),$$

这与定理的条件矛盾，所以假设不成立，即只能 $j=1$，所以 $n-1=\mathrm{ord}(a)$. 我们已经知道，$\mathrm{ord}(a)\leqslant\varphi(n)\leqslant n-1$，因此，$\varphi(n)=n-1$. 这说明 n 必为素数.

[例 3.3.1] 使用 Lucas 检验法，判断 997 是否为素数.

解 选择 $a=7$，验证得到 $7^{996}\equiv 1\ (\mathrm{mod}\ 997)$. 因为 $996=2^2\times 3\times 83$，计算得到

$$7^{996/2}=7^{498}\equiv 996\ (\mathrm{mod}\ 997),$$
$$7^{996/3}=7^{332}\equiv 304\ (\mathrm{mod}\ 997),$$
$$7^{996/83}=7^{12}\equiv 9\ (\mathrm{mod}\ 997),$$

所以 997 是素数.

可以对 Lucas 检验法进行如下改进.

定理 3.3.2 （Lehmer 检验）如果对 $n-1$ 的任意素数因子 p_i，都存在一个整数 a_i，使得 $a_i^{n-1}\equiv 1\ (\mathrm{mod}\ n)$和 $a_i^{\frac{n-1}{p_i}}\not\equiv 1\ (\mathrm{mod}\ n)$成立，那么 n 是素数.

证明 设有如下标准分解式，$n-1=p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r}$. 由 $a_i^{n-1}\equiv 1\ (\mathrm{mod}\ n)$，可知 $\mathrm{ord}(a_i)\mid n-1$，由 $a_i^{\frac{n-1}{p_i}}\not\equiv 1(\mathrm{mod}\ n)$，可知 $\mathrm{ord}(a_i)\nmid\frac{n-1}{p_i}$，所以，$p_i^{k_i}\mid\mathrm{ord}(a_i)$. 由于 $\mathrm{ord}(a_i)\mid\varphi(n)$，所以 $p_i^{k_i}\mid\varphi(n)$，因此 $n-1\mid\varphi(n)$. 我们可以断定 $n-1=\varphi(n)$，即 n 是素数.

[例 3.3.2] 使用 Lehmer 检验法，判断 997 是否为素数.

解 选择 $a_1=5$，$a_2=7$，$a_3=3$. 计算得到

$$5^{996}\equiv 1\ (\mathrm{mod}\ 997),$$
$$5^{996/2}=5^{498}\equiv 996\ (\mathrm{mod}\ 997),$$
$$7^{996}\equiv 1\ (\mathrm{mod}\ 997),$$
$$7^{996/3}=7^{332}\equiv 304\ (\mathrm{mod}\ 997),$$
$$3^{996}\equiv 1\ (\mathrm{mod}\ 997),$$
$$3^{996/83}=3^{12}\equiv 40\ (\mathrm{mod}\ 997),$$

所以 997 是素数.

定理 3.3.3 （Pocklington 检验）对 $n-1$ 做不完全因子分解，得到 $n-1=mj$，其中有标准分解式 $m=p_1^{k_1}p_2^{k_2}\cdots p_r^{k_r}$，$m\geqslant\sqrt{n}$且$(m,j)=1$. 如果对每个 $p_i(1\leqslant i\leqslant r)$，都存在一个整数 a_i，使得 $a_i^{n-1}\equiv 1\ (\mathrm{mod}\ n)$和$(a_i^{\frac{n-1}{p_i}}-1,n)=1$ 成立，那么 n 是素数.

证明 令 p 为 n 的任意素因子，因为 $a_i^{n-1}\equiv 1\ (\mathrm{mod}\ n)$，所以 $a_i^{n-1}\equiv 1\ (\mathrm{mod}\ p)$，因此 $\mathrm{ord}_p(a_i)\mid p-1$. 由 $a_i^{n-1}\equiv 1\ (\mathrm{mod}\ p)$，也可知 $\mathrm{ord}_p(a_i)\mid n-1$. 因为$(a_i^{\frac{n-1}{p_i}}-1,n)=1$，所以 $p\nmid a_i^{\frac{n-1}{p_i}}-1$，即 $a_i^{\frac{n-1}{p_i}}\not\equiv 1\ (\mathrm{mod}\ p)$，从而 $\mathrm{ord}_p(a_i)\nmid\frac{n-1}{p_i}$，由此得到 $p_i^{k_i}\mid\mathrm{ord}_p(a_i)$，也即得 $p_i^{k_i}\mid p-1$. 因为这个关系对任意 i 都成立，所以 $m\mid p-1$，必有 $p>m$. 又 $m\geqslant\sqrt{n}$，则 $p>\sqrt{n}$. 由 p 的任意性可知，n 的素因子全都比$\sqrt{n}$大，这只有在 n 为素数的情况下是正确的. 否则，如果 n 为合数，那么必有比$\sqrt{n}$小的素因子. 总之，n 为素数.

[例 3.3.3] 使用 Pocklington 检验法，判断 997 是否为素数.

解 对 $n-1=996$ 做不完全因子分解得到 $996=12\times 83$，其中 $83>\sqrt{997}$. 选择 $a_1=2$，计算得到

$$2^{996}\equiv 1 \pmod{997},$$

$$(2^{996/83}-1,\ 997)=(4\,095,\ 997)=1,$$

所以 997 是素数.

上面三个定理决定的素性检验算法是确定性的算法，但是它们的效率低. 为了提高检验效率，往往采用非确定性的算法，这些算法中的大多数都与费马小定理有关. 根据费马小定理，我们知道，若 n 是一个素数，则对任意整数 b，有

$$b^n\equiv b \pmod{n}.$$

于是，它的逆否命题也成立，即对大于 1 的整数 n，若存在整数 b，使

$$b^n\not\equiv b \pmod{n},$$

则 n 是一个合数. 那么，能否说对大于 1 的整数 n，若存在整数 b，使

$$b^n\equiv b \pmod{n},$$

则 n 就是一个素数呢？结论是否定的. 例如，

$$2^{341}\equiv 2 \pmod{341},$$

但 $341=11\times 12$.

定义 3.3.1　设 b 是大于 1 的整数，如果一个合数 n 满足

$$b^n\equiv b \pmod{n},$$

则 n 叫作基 b 的**伪素数**.

显然，341 就是基 2 的伪素数.

引理 3.3.1　设 n，d 为正整数，若 $d\mid n$，则 $2^d-1\mid 2^n-1$.

证明　由于 $d\mid n$，故存在正整数 q 使得 $n=dq$. 于是，有

$$2^n-1=(2^d)^q-1=(2^d-1)((2^d)^{q-1}+(2^d)^{q-2}+\cdots+2^d+1),$$

所以 $2^d-1\mid 2^n-1$.

定理 3.3.4　若 n 是一个基 2 的伪素数，则 2^n-1 也是一个基 2 的伪素数. 因此，存在无穷多个基 2 的伪素数.

证明　设 $m=2^n-1$，因为 $2^n\equiv 2 \pmod{n}$，所以 $n\mid m-1$，根据上面引理，可知

$$2^n-1\mid 2^{m-1}-1,$$

即 $m\mid 2^{m-1}-1$. 因此，有

$$2^m\equiv 2 \pmod{m}.$$

因为 n 是合数，所以存在素因子 p，使 $p\mid n$，根据引理可知

$$2^p-1\mid 2^n-1,$$

于是 $m=2^n-1$ 也是合数. 定理得证.

基 2 的伪素数虽然有无穷多个，但在某一个区间内，比素数少得多. 例如，小于 10^9 的素数有 50 847 534 个，但仅有 5 597 个基 2 的伪素数. 因此，对小于 10^9 的正整数 n，若满足

$$2^n\equiv 2 \pmod{n},$$

则它是素数的概率非常大，也就是说，n 是素数这一断言出错的概率很小，仅为

$$\frac{5\,597}{50\,847\,534+5\,597}\approx 0.000\,1.$$

但是，由于采用固定的基(如，上面采用的基为 2)这种方法毕竟会出错，所以它在实际的素性检验工作中用处不大. 那么，是否采用更多的不同的基对同一个大数进行测试就可以解决

问题呢?

下面给出费马素性检验算法,它采用更多的不同的基.

(1) 给定大于 2 的奇数 n 和安全参数 k. 置 $i=0$.

(2) 若 $i<k$,在区间 $1<b<n-1$ 中随机选取整数 b;否则,n 很有可能是素数,算法结束.

(3) 若 $b^n\equiv b(\mathrm{mod}\ n)$,则置 $i=i+1$,转(2).

(4) n 必为合数,算法结束.

但是,费马素性检验算法在一些情况下,注定会给出错误的判断.

定义 3.3.2 设 n 是一个合数,如果对所有与 n 互素的正整数 b,都满足

$$b^{n-1}\equiv 1(\mathrm{mod}\ n),$$

则 n 叫作**卡米歇尔(Carmichael)数或者绝对伪素数.**

[例 3.3.4] 证明 561 是一个卡米歇尔数.

证明 易知 $561=3\times 11\times 17$. 如果正整数 b 满足 $(b,561)=1$,则有

$$(b,3)=(b,11)=(b,17)=1.$$

于是,

$$b^2\equiv 1(\mathrm{mod}\ 3),\quad b^{10}\equiv 1(\mathrm{mod}\ 11),\quad b^{16}\equiv 1(\mathrm{mod}\ 17),$$

进而

$$b^{560}\equiv (b^2)^{280}\equiv 1(\mathrm{mod}\ 3),$$
$$b^{560}\equiv (b^{10})^{56}\equiv 1(\mathrm{mod}\ 11),$$
$$b^{560}\equiv (b^{16})^{35}\equiv 1(\mathrm{mod}\ 17),$$

因此,

$$b^{560}\equiv 1(\mathrm{mod}\ 561),$$

即 561 是一个卡米歇尔数.

下面列出了前 10 个卡米歇尔数:

561,1 105,1 729,2 465,2 821,6 601,8 911,10 585,15 841,29 341.

卡米歇尔曾在 1912 年猜想:存在无穷多个卡米歇尔数. 但直到 1992 年,这个猜想才被 W. Alford, G. Granville 和 C. Pomerance 证明.

我们直接给出 n 是一个卡米歇尔数的充要条件.

定理 3.3.5 n 是一个卡米歇尔数的充要条件是:

$$n=\prod_{i=1}^{k} p_i,\quad k\geqslant 3,\quad p_i\neq p_j(i\neq j),$$

p_i 为奇素数,且对所有的 $1\leqslant i\leqslant k$ 有 $p_i-1\mid n-1$.

由于卡米歇尔数的存在,我们不能使用费马小定理的逆命题来证明一个数是素数,即使采用很多不同的基进行检验.

定理 3.3.6 设 p 为素数,则

$$x^2\equiv 1(\mathrm{mod}\ p)$$

的充要条件是

$$x\equiv \pm 1(\mathrm{mod}\ p).$$

证明 先证必要性. 由于

$$x^2\equiv 1(\bmod p),$$

则有

$$(x-1)(x+1)\equiv 0(\bmod p),$$

即 $p|(x-1)(x+1)$，于是 $p|x-1$ 或 $p|x+1$，故

$$x-1\equiv 0(\bmod p)\quad 或\quad x+1\equiv 0\ \ (\bmod p),$$

即

$$x\equiv \pm 1(\bmod p).$$

反过来，如果 $x-1\equiv 0(\bmod p)$或 $x+1\equiv 0(\bmod p)$，那么 $x^2\equiv 1(\bmod p)$. 证毕.

显然我们可以得到以下推论.

推论　对于大于 2 的整数 n，若存在整数 x 满足

$$x^2\equiv 1(\bmod n),$$

但 $x\not\equiv 1(\bmod n)$，则 n 是合数.

例如，取 $x=6$，则

$$x^2=6^2\equiv 1(\bmod 35),$$

但 $x=6\not\equiv \pm 1(\bmod 35)$，上面的推论说明 35 必定是合数.

根据上面的定理和推论，来研究一种新的素性检验的方法.

设 n 是正奇数，且有 $n=2^td+1$，其中 d 为奇素数，则有

$$b^{n-1}-1=b^{2^td}-1=(b^d-1)(b^d+1)(b^{2d}+1)\cdots(b^{2^{t-1}d}+1),$$

其中$(b,\ n)=1$.

我们知道，如果

$$b^{n-1}=b^{2^td}\not\equiv 1\ (\bmod n),$$

那么 n 必为合数. 所以，若要使 n 为素数，必须

$$b^{n-1}=b^{2^td}\equiv 1\ (\bmod n),$$

并且，根据定理 3.3.6 和推论，序列

$$b^{2^{t-1}d},\ b^{2^{t-2}d},\cdots,b^{2d},b^d$$

中第一个模 n 不等于 1 的数必模 n 等于 -1. 若 n 不满足上述条件，则必为合数；若 n 满足上述条件，则很有可能是素数.

定义 3.3.3　设 n 是一个奇合数，且有 $n=2^td+1$，其中 d 为奇素数，设整数 b 与 n 互素，如果 n 和 b 满足

$$b^d\equiv 1\ (\bmod n),$$

或存在一个整数 r，$0\leqslant r<t$，使得

$$b^{2^rd}\equiv -1\ (\bmod n),$$

则 n 叫作基 b 的**强伪素数**.

我们直接给出以下定理.

定理 3.3.7　设 n 是一个奇合数，则在区间 $0<b<n$ 中，最多有$(n-1)/4$ 个数 b 使 n 是基 b 的强伪素数.

下面介绍一下强伪素性检验算法，又称为 Miller-Rabin 素性检验.

(1) 给定大于 2 的奇数 n，令 $n-1=2^td$，其中 d 为奇数. 在区间 $1<b<n$ 中随机选取整数 b.

(2) 置 $i=0$，$y\equiv b^d \pmod n$.

(3) 若 $i=0$ 且 $y=1$，或者 $y=n-1$，则 n 可能是素数，算法结束. 若 $i>0$ 且 $y=1$，则转(5).

(4) 置 $i=i+1$. 若 $i<t$，则置 $y\equiv y^2 \pmod n$，转(3).

(5) n 必为合数，算法结束.

如果对同一个合数 n，k 次执行上面的测试，并且每次的结果都为“n 可能是素数”的概率只为 0.25^k，因此可以通过增加执行测试次数将发生错误判断的概率降低到任何预先给定的范围内.

还有一种与强伪素性检验法类似的素性检验方法，即欧拉伪素性检验法，但它不如前者效果好. 它利用了雅可比符号，并且以欧拉判别条件作为理论依据.

设 n 是一个奇素数，则有

$$b^{\frac{n-1}{2}}\equiv\left(\frac{b}{n}\right) \pmod n$$

对任意整数 b 成立. 因此，对于大于 2 的正整数 n，若存在与 n 互素的整数 b，使得

$$b^{\frac{n-1}{2}}\not\equiv\left(\frac{b}{n}\right) \pmod n,$$

则 n 是合数.

定义 3.3.4 设 n 是一个奇合数，整数 b 与 n 互素，若 n 和 b 满足

$$b^{\frac{n-1}{2}}\equiv\left(\frac{b}{n}\right) \pmod n$$

则 n 叫作基 b 的**欧拉伪素数**.

例如，取 $n=1\,105=5\times13\times17$，$b=2$，可得

$$2^{552}\equiv1 \pmod{1\,105},$$

$$\left(\frac{2}{1\,105}\right)=(-1)^{\frac{1\,105^2-1}{8}}=1,$$

于是，有

$$2^{552}\equiv\left(\frac{2}{1\,105}\right) \pmod{1\,105},$$

故 1 105 是一个基 2 的欧拉伪素数.

利用这种方法进行检验的算法被称为 Solovay-Strassen 检验. 类似于 Miller-Rabin 素性检验，也可以通过多次随机选择不同基的方法将发生错误判断的概率降低到任何预先给定的范围内. 已经证明 k 次执行这种测试，并且每次的结果都为“n 可能是素数”的概率只为 0.5^k.

定理 3.3.8 若 n 是基 b 的欧拉伪素数，则 n 也是基 b 的伪素数.

证明 若 n 是基 b 的欧拉伪素数，则有

$$b^{\frac{n-1}{2}}\equiv\left(\frac{b}{n}\right) \pmod n,$$

将上式两端平方，得

$$b^{n-1}\equiv1 \pmod n,$$

于是，n 是基 b 的伪素数.

此外，还有椭圆曲线素性检验算法和 AKS 素性检验算法，等等，由于它们都需要更加高

深的数论知识，这里不再讨论，有兴趣的读者可以参考有关文献.

当密码学算法中需要生成大的素数的时候，一般采用的方法是：首先随机生成一个具有指定位数的大整数然后对该数进行概率性的素性检验，如果检验的结果是否定的，则重新进行随机生成和检验步骤，直到找到具有规定长度的素数为止.

3.3.2　因子分解算法

因子分解的困难性要比素性检验的困难性大得多，因为素性检验结果的正确性不需要通过因子分解来证明. 对于密码学算法来说，因子分解的困难性正是有利之处. 反过来说，在密码学应用环境中，因子分解算法研究的目的主要在于破译密码. 最简单的也是最古老的分解因子方法就是试除法，但是前面已经提到这种方法效率最低. 第一个在因子分解问题上真正的进步是费马提出的算法. 下面的算法都是针对奇素数进行的，因为一个数中的 2 的因子是非常容易分解出来的.

1. 费马分解算法

对于奇整数 n，当能够获得如下方程

$$n=x^2-y^2$$

的整数解时，也就获得了 n 的两个因子，因为

$$n=(x-y)(x+y).$$

反过来，当获得如下因子分解

$$n=ab(a\geqslant b\geqslant 1)$$

的时候，也就获得了上面方程的整数解，因为

$$n=\left(\frac{a+b}{2}\right)^2-\left(\frac{a-b}{2}\right)^2.$$

该方法过程为，首先确定最小的整数 k，使得 $k^2\geqslant n$. 然后，对下面的数列

$$k^2-n,(k+1)^2-n,(k+2)^2-n,(k+3)^2-n,\cdots,((n+1)/2)^2-n$$

按顺序进行测试，直到找到一个整数 $m\geqslant\sqrt{n}$ 使得 m^2-n 是一个平方整数，从而也就找到了一对因子，如果一直运行到上面数列最后一个数才找到平方整数，那么 n 就是素数没有非平凡因子.

[例 3.3.5]　对 $n=119\ 143$ 进行因子分解.

解　因为 $345^2<119\ 143<346^2$，所以 $k=346$，

$$346^2-119\ 143=573,$$
$$347^2-119\ 143=1\ 266,$$
$$348^2-119\ 143=1\ 961,$$
$$349^2-119\ 143=2\ 658,$$
$$350^2-119\ 143=3\ 357,$$
$$351^2-119\ 143=4\ 058,$$
$$352^2-119\ 143=4\ 761=69^2,$$

所以，$n=(352-69)(352+69)=283\times 421$.

2. Pollard ρ 分解算法

首先，确定一个简单的二次以上整系数多项式，例如 $f(x)=x^2+a$，$a\neq -2,0$. 然后，

从一个初始值 x_0 开始，利用迭代公式

$$x_{k+1}\equiv f(x_k)\pmod n$$

计算一个序列 $x_1,x_2,x_3,\cdots$. 另 d 为 n 的一个非平凡因子，因为模 d 的剩余类个数比模 n 的剩余类个数少很多，所以很可能存在某个 x_k 和 x_j 是属于同一个模 d 的剩余类的同时又是属于不同的模 n 的剩余类的. 因为 $d\mid x_k-x_j$ 而 $n\nmid x_k-x_j$，所以 $\gcd(x_k-x_j,n)$ 是 n 的非平凡因子.

[例 3.3.6] 求一个 $n=2\,189$ 的非平凡因子.

解 选择 $x_0=1$ 和 $f(x)=x^2+1$，得到序列 $x_1=2,x_2=5,x_3=26,x_4=677,x_5=829,\cdots$.

因为 $\gcd(x_5-x_3,n)=11$. 所以 11 是一个因子.

如果在该算法中，对序列中的任意两个量的差都进行计算的话，那么计算的开销就太大了. 改进方法是只对 $x_{2k}-x_k$ 进行计算.

[例 3.3.7] 求一个 $n=30\,623$ 的非平凡因子.

解 选择 $x_0=3$ 和 $f(x)=x^2-1$，得到序列 8，63，3 968，4 801，21 104，28 526，18 319，18 926，….

$$\gcd(x_2-x_1,n)=1,$$
$$\gcd(x_4-x_2,n)=1,$$
$$\gcd(x_6-x_3,n)=1,$$
$$\gcd(x_8-x_4,n)=113,$$

所以 113 是一个因子.

如果算法在运行预先规定的步数后不成功，可以重新选择 x_0 或者 $f(x)$再开始处理.

3. Pollard $p-1$ 分解算法

这个方法对如下的情况起作用：奇合数 n 有一个素因子 p，而且 $p-1$ 是小素数之积.

该算法需要预先选择一个整数 k，只要 k 充分大，就可以保证 $(p-1)\mid k!$，接着选择一个整数 a，使得 $1<a<p-1$，计算 $a^{k!}\equiv m\pmod n$. 因为存在整数 j，使 $k!=j(p-1)$，所以

$$m\equiv a^{k!}\equiv a^{j(p-1)}\equiv(a^{p-1})^j\equiv 1^j=1\pmod p.$$

也就是 $p\mid(m-1)$. 因此 $\gcd(m-1,n)>1$，只要 $m\not\equiv 1\pmod n$，$\gcd(m-1,n)$必为 n 的非平凡因子.

[例 3.3.8] 求一个 $n=2\,987$ 的非平凡因子.

解 选 $a=2$ 和 $k=7$.

$$2^{7!}\equiv 755\pmod{2\,987},$$
$$\gcd(755-1,2\,987)=29,$$

所以，29 是一个非平凡因子.

此外，还有 Dixon 的随机平方算法、连分数分解算法、二次筛法、数域筛法、椭圆曲线因子分解算法，等等. 尽管不再详细讨论这些算法，但是下面指出它们中的大多数都基于如下的思路.

如果能够找到两个整数 x 和 y，满足

$$x^2\equiv y^2\pmod n,$$

且

$$x \not\equiv \pm y \pmod{n},$$

那么 $\gcd(x-y, n)$和 $\gcd(x+y, n)$是 n 的非平凡因子. 例如，可以验证 $10^2 \equiv 32^2 \pmod{77}$，那么，根据上面的分析就知道 $\gcd(32-10, 77)=11$ 和 $\gcd(32+10, 77)=7$ 都是 77 的非平凡因子.

Dixon 的随机平方算法、连分数分解算法、二次筛法和数域筛法之间的不同主要在于如何构造满足上面同余条件的两个整数 x 和 y.

习题

1. 34 对模 37 的次数是多少？
2. 2^{12}对模 37 的次数是多少？
3. 2 是模 61 的一个原根，利用这个事实，在小于 61 的正整数中，找到所有次数为 4 的整数.
4. 设 $ab \equiv 1 \pmod{m}$，求证 $\mathrm{ord}_m(a) = \mathrm{ord}_m(b)$.
5. 2 是 19 的原根，造出 19 的指数表，求出如下各方程的最小正剩余解：
   (1)$8x^4 \equiv 3 \pmod{19}$；
   (2)$5x^3 \equiv 2 \pmod{19}$；
   (3)$x^7 \equiv 1 \pmod{19}$.
6. 3 是 17 的原根，造出 17 的指数表，求出满足如下各方程的整数 x：
   (1)$3^x \equiv 7 \pmod{17}$；
   (2)$3^x \equiv x \pmod{17}$.
7. 求证 3 是一个 17 的原根，找到最小正剩余系中的所有原根.
8. 求证如果 g^k 是 m 的原根，那么 g 也是 m 的原根.
9. 试问 47,55,59 的原根是否存在？若存在则求出其所有的原根.
10. 求解同余方程
$$x^{22} \equiv 5 \pmod{41}.$$
11. 求出同余方程 $x^2 \equiv 8 \pmod{287}$的所有解.
12. 在最小正剩余系中，求出 29 的所有二次剩余.
13. 下列各方程有几个解？
    $x^2 \equiv 19 \pmod{170}$；$x^2 \equiv 38 \pmod{79}$；$x^2 \equiv 76 \pmod{165}$.
14. 1 999 是一个素数，运用欧拉判别条件确定是否 $1\,999 \mid 2^{999}-1$.
15. 判断同余方程
$$x^2 \equiv 191 \pmod{397}$$
    是否有解.
16. 判断同余方程
$$x^2 \equiv 11 \pmod{511}$$
    是否有解.
17. 求所有奇素数 p，它以 3 为其二次剩余.
18. 求所有奇素数 p，它以 -3 为其二次剩余.
19. 不用求解，证明 $x^6 \equiv 8 \pmod{89}$有解，有几个解？3 是 89 的一个原根，找到该方程的最小正剩余解.

20. 求解同余方程 $x^5 \equiv 2 \pmod{73}$.

21. $(a, 71)=1$，求证 $x^{26} \equiv a \pmod{71}$ 和 $x^{26} \equiv a \pmod{71}$不可能同时有解.

22. 使用高斯引理计算 $\left(\frac{14}{23}\right)$.

23. 是否存在正整数 n 使得 n^2-3 是 313 的倍数？

24. 试证 $1\,105=5\times13\times17$ 和 $2\,821=7\times13\times31$ 是卡米歇尔数.

25. 试证 341 不是卡米歇尔数.

26. 对于 $n=(6k+1)(12k+1)(18k+1)$形式的整数，如果其中的三个因子是素数，那么 n 是卡米歇尔数. 请找到满足条件的最小数.

27. 以 3 为基，对 1 709 进行 Lucas 检验.

28. 对 907 进行 Pocklington 检验.

29. 判断 25 是否为对基 7 的强伪素数.

30. 判断 2 047 是否为对基 3 的强伪素数.

31. 运用费马分解算法求 2 279 和 278 153 的非平凡因子.

32. 运用 Pollard ρ 分解算法分解算法求 1 133 和 8 051 的非平凡因子.

33. 运用 Pollard $p-1$ 分解算法求 1 711 的非平凡因子.

34. 试设计一种类似于 Pollard $p-1$ 分解算法的方法，用来对如下情况的整数进行分解，奇合数 n 有一个素因子 p，而且 $p+1$ 是小素数之积.

35. 利用 $12^2 \equiv 5^2 \pmod{119}$求 119 的两个非平凡因子.

第 4 章　代数系统基础

代数是由算术演变而来的，为了对数学问题进行求解，往往需要首先用一种方法根据问题中的已知量和未知量简洁地描述出问题本身，从而产生用字母等符号代替数的思想方法，即建立方程或者方程组，然后对方程进行求解. 随着这种方法使用得越来越多也越来越熟练，人们发展到以符号代替各种事物，乃至于概念、规则和映射，等等. 而且，人们发现各种数系的运算规则具有很多相似性，从而将这种共同的规则进行提纯，对抽象的集合赋予运算规则，然后进行研究. 在这样的思想方法下，代数学从 19 世纪开始有了惊人的发展，不仅使得我们对数有了更好的理解，而且也使我们对其他很多表面看起来不是数的概念有了深入的研究，带动了数学整体的发展. 代数的基础是研究各种代数系统，即定义了运算的抽象集合. 主要的代数系统有群、环、域等. 代数系统理论对信息科学，尤其是编码理论和密码学，有很大的实用意义，能够对众多算法进行本质性的更加精炼的描述.

4.1　群

4.1.1　群及其基本性质

群是一种重要的代数系统，关于群的理论研究是由法国的天才数学家伽罗华(1811—1832)开创的. 当时，这个理论的出现主要是为了解决如下的问题：对于一般的高次代数方程是否存在像解二次方程那样的求根公式，其中只需要用到四则运算和开(高次)方运算. 群论有相当广泛的应用，如可以给出数论中很多定理的更加直截了当的证明.

群的概念离不开运算，而运算的实质就是将两个同类型的事物组合在一起形成第三个与它们同样类型的事物. 比如，我们熟悉的普通加减乘除运算都是将两个数组合后形成第三个数，还有函数的复合是将两个函数组合后形成第三个函数. 二元运算的正式定义如下.

定义 4.1.1　集合 G 上的**二元运算**是一个如下的函数

$$*:G\times G\to G.$$

更详细地说，集合 G 上的二元运算是为每一个 G 上的有序对(a,b)分配一个 G 中的确定元素 $*(a,b)$与之对应，通常将 $*(a,b)$用更加自然的方式写成 $a*b$，当然，也可以用“$\cdot$”、“$\odot$”、“$\oplus$”、“$\times$”、“$*$”、“$+$”等符号来表示这样的二元运算. 有时为了简化书写，在不引起歧义的情况下，可以用 ab 来表示 $a*b$.

[例 4.1.1]　加法运算是函数 $+(a,b)=a+b$，减法运算是函数 $-(a,b)=a-b$，函数的复合是函数 $\circ(f,g)=g\circ f$. 然而，请注意，向量的点积运算就不是这里定义的二元运算，因为参与点积运算的两个对象是向量，而计算结果是一个数，它们属于不同的集合.

定义 4.1.2　$*$ 是集合 G 上的二元运算，如果对任意 $a,b,c\in G$，都有

$$a*(b*c)=(a*b)*c,$$

那么就称该二元运算 $*$ 满足**结合律**.

定义 4.1.3 $*$ 是集合 G 上的二元运算，且 $a, b\in G$，满足

$$a*b=b*a,$$

那么就称这**两个元素可交换**；如果对任意 $a, b\in G$，都有

$$a*b=b*a,$$

那么就称该二元运算 $*$ 满足**交换律**.

[例 4.1.2] 实数的加法和乘法很显然同时满足交换律和结合律. 而减法既不满足结合律也不满足交换律.

定义 4.1.4 $*$ 是集合 G 上的二元运算，H 是 G 的子集，如果对任意 $a, b\in H$，都有

$$a*b\in H,$$

那么就称子集 H 在二元运算 $*$ 下是**封闭的**. 此时，很明显，如果二元运算 $*$ 在 G 上满足结合律，那么，它必然也在 H 上满足结合律；如果二元运算 $*$ 在 G 上满足交换律，那么，它必然也在 H 上满足交换律.

定义 4.1.5 **群**是一个有序对$(G, *)$，其中 G 是一个集合，$*$ 是集合 G 上的一个二元运算，它们满足如下条件：

(群条件 1) $*$ 满足结合律；

(群条件 2)存在 $e\in G$，对任意 $a\in G$ 有 $e*a=a$，元素 e 称为 G 的**单位元**；

(群条件 3)对任意 $a\in G$，存在 $a'\in G$ 使得 $a'*a=e$.

注意，上面的“群条件 2”保证群总是非空的. 如果$(G, *)$是群，我们经常会说“G 在运算 $*$ 下是一个群”，或者在不引起歧义的情况下，更简单地称“G 是一个群”. 如果群 G 是一个有限集合，那么我们称它为有限群.

下面将看到，实际上，群具有更好的性质.

引理 4.1.1 G 是一个群，$a\in G$ 满足 $a*a=a$，那么 $a=e$.

证明 因为 G 是一个群，所以存在 $a'\in G$ 使得 $a'*a=e$；因为 $a*a=a$，所以

$$a'*(a*a)=a'*a,$$

上式左边$=a'*(a*a)=(a'*a)*a=e*a=a$，而上式右边$=a'*a=e$. 所以 $a=e$.

定理 4.1.1 G 是一个群，那么

(1) 对任意 $a\in G$ 都有 $a*a'=e$.

(2) 对任意 $a\in G$ 都有 $a*e=a$.

(3) 如果 $e_0\in G$ 满足：对任意 $a\in G$ 有 $e_0*a=a$，那么 $e_0=e$.

(4) 如果 $b\in G$ 使得 $b*a=e$，那么 $b=a'$.

证明 (1) 因为

$$\begin{aligned}(a*a')*(a*a')&=[(a*a')*a]*a'=[a*(a'*a)]*a'\\&=[a*e]*a'=a*[e*a']=a*a',\end{aligned}$$

所以由引理 4.1.1 可知，$a*a'=e$.

(2) $a*e=a*(a'*a)=(a*a')*a=e*a=a$.

(3) 由题设知 $e_0*e_0=e_0$，所以由引理 4.1.1 可知，$e_0=e$. 这个定理告诉我们，群只有一个唯一的单位元.

(4) $b=b*e=b*(a*a')=(b*a)*a'=e*a'=a'$. 这个定理告诉我们，对任意 $a\in G$，存在唯一一个 $a'\in G$ 使得 $a'*a=e$.

定理 4.1.1(4)使得我们可以做如下的定义.

定义 4.1.6　G 是一个群，$a\in G$，存在唯一一个 $a'\in G$ 使得 $a'*a=e$，我们称 a' 为元素 a 的**逆元**，而且用符号 a^{-1} 表示 a 的**逆元**.

[例 4.1.3]　$\mathbf{Z}$ 在普通加法下是一个群，我们早已熟知两个整数的和还是整数，而且普通加法满足结合律，其中 $e=0$，对任意 a，$a^{-1}=-a$. 同理，$\mathbf{Q}$，$\mathbf{R}$，$\mathbf{C}$ 三个集合在普通加法下也都是群.

[例 4.1.4]　$\mathbf{Q}-\{0\}$ 在普通乘法下是群，我们熟知两个非零有理数的积还是有理数，而且普通乘法满足结合律，其中 $e=1$，对任意 $a\in\mathbf{Q}-\{0\}$，$a^{-1}=\dfrac{1}{a}$. 同理，$\mathbf{R}-\{0\}$ 和 $\mathbf{C}-\{0\}$ 也在普通乘法下是群. 然而，$\mathbf{Z}-\{0\}$ 在普通乘法下不是群，因为这个集合中的任意不等于 1 或者 -1 的元素在这个集合中没有逆元.

[例 4.1.5]　线性代数课程中向量空间 $\mathbf{R}^n$ 在向量加法下是群，由向量空间的定义可知，任意两个向量的和还是向量，而且向量加法满足结合律(因为向量的坐标是实数，实数满足加法结合律)，其中 e 为零向量，对任意向量 $a\in\mathbf{R}^n$，$a^{-1}=-a$.

对于群$(G,*)$中的任意 $n(n\geqslant 2)$个元素 $a_1,a_2,a_3,\cdots,a_n$，我们在计算它们的“积”的时候，先写成如下的表达式 $a_1*a_2*a_3*\cdots*a_n$，然后在保证正确配对的情况下，向该表达式任意加入成对的括号，这样就会以很多种不同的计算次序得到表达式的“积”. 例如 $n=2$ 时，只有一种计算次序“$a_1\cdot a_2$”，$n=3$ 时，有两种计算次序$(a_1\cdot a_2)\cdot a_3$和 $a_1\cdot(a_2\cdot a_3)$，这样我们自然关心一个如下的问题：在群中由不同计算次序得到的结果是否都相等？或者说，“积”是否不受计算次序的影响. 对这个问题，由归纳法可以很容易得到肯定的答案，总结在下面的定理，我们称其为“推广的结合律”.

定理 4.1.2　设 $a_1,a_2,a_3,\cdots,a_n$ 是群$(G,*)$中的任意 $n(n\geqslant 2)$个元素，对任意 $1\leqslant i_1<i_2<\cdots<i_k\leqslant n$，有

$$(a_1\cdot a_2\cdot\cdots\cdot a_{i_1})\cdot(a_{i_1+1}\cdot a_{i_1+2}\cdot\cdots\cdot a_{i_2})\cdot\cdots\cdot(a_{i_k+1}\cdot a_{i_k+2}\cdot\cdots\cdot a_n)$$
$$=a_1\cdot a_2\cdot\cdots\cdot a_n.$$

群还具有以下一些基本性质.

定理 4.1.3　G 是一个群，则

(1) 对任意 $a\in G$ 有$(a^{-1})^{-1}=a$.

(2) 对任意 $a,b\in G$ 有$(a*b)^{-1}=b^{-1}*a^{-1}$.

证明　(1)因为 $a*a^{-1}=e$，且$(a^{-1})^{-1}*a^{-1}=e$，所以 a 和$(a^{-1})^{-1}$都是 a^{-1}的逆元，由逆元的唯一性可知$(a^{-1})^{-1}=a$.

(2) $(b^{-1}*a^{-1})*(a*b)=b^{-1}*(a^{-1}*a)*b=b^{-1}*(a^{-1}*a)*b=b^{-1}*e*b=b^{-1}*b=e$. 由逆元的唯一性可知$(b^{-1}*a^{-1})=(a*b)^{-1}$.

定理 4.1.4　G 是一个群，则**消去律**成立，即对 $a,b,x\in G$.

(1) 如果 $ax=bx$，那么 $a=b$.

(2) 如果 $xa=xb$，那么 $a=b$.

证明　(1) 因为 $ax=bx$，所以$(ax)x^{-1}=(bx)x^{-1}$，由结合律可知 $a(x\,x^{-1})=b(x\,x^{-1})$，

所以 $a\,e=b\,e$，得 $a=b$. 同理(2)得证.

定理 4.1.5 G 是一个群，a，$b\in G$，则方程 $ax=b$ 和 $ya=b$ 有唯一的解.

证明 对 $ax=b$ 两边同时左乘以 a^{-1}，则 $a^{-1}(ax)=a^{-1}\,b$，所以得到解为 $x=a^{-1}\,b$，a^{-1} 是唯一的，因此该解是唯一解. 同理，第二个方程的唯一解为 $y=b\,a^{-1}$.

注意，在前面的各个证明中，在计算元素的“积”时，没有交换任意参加运算的元素的顺序，因为一般的群不一定满足交换律.

[例 4.1.6] 不满足交换律的群：

(1) 在矩阵乘法下，所有 2×2 的可逆实数矩阵组成一个群. 其中，矩阵乘法满足结合律，而且两个可逆矩阵的乘积还是可逆矩阵，单位元是单位矩阵，任意元素的逆元是其逆矩阵. 但是，矩阵乘法不满足交换律.

(2) 在函数的复合运算下，所有 **R** 到 **R** 上的可逆函数组成一个群. 其中，函数的复合满足结合律，而且两个可逆函数的复合还是可逆函数，单位元是恒等函数，任意元素的逆元是其逆函数. 但是，函数的复合不满足交换律.

定义 4.1.7 如果群 G 中的二元运算 $*$ 还额外满足交换律，就称 G 是一个**交换群**或**阿贝尔群**.

[例 4.1.7] (1) 普通加法下的 **Z**，**Q**，**R**，**C** 和普通乘法下的 $\mathbf{Q}-\{0\}$，$\mathbf{R}-\{0\}$，$\mathbf{C}-\{0\}$ 都是阿贝尔群.

(2) 复平面上的单位圆，即复数集合 $\{z\in\mathbf{C}:|z|=1\}$，在复数乘法下也是一个阿贝尔群. 其中，z 的逆元为 z 的共轭，且 $|z$ 的共轭$|=1$.

(3) 1 在复数范围内的所有 n 次根，即集合 $\{(e^{\frac{2\pi i}{n}})^k:0\leqslant k<n\}$，在复数乘法下也是一个阿贝尔群. 其中，元素的逆元也为它的共轭.

为了符号使用的方便，我们可以用“1”来表示群的单位元.

定义 4.1.8 对正整数 n，群中元素 a 的**幂**为：$a^n=\underbrace{a\,a\cdots a}_{\text{共}n\text{项}}$ 和 $a^{-n}=\underbrace{a^{-1}\,a^{-1}\cdots a^{-1}}_{\text{共}n\text{项}}$；$a^0=1$.

定理 4.1.6 对正整数 m 和 n，群中元素 a 的幂满足：

(1) $(a^{-1})^n=(a^n)^{-1}$；

(2) $a^{n+m}=a^n a^m$；

(3) $(a^n)^m=a^{nm}$.

证明 略(很容易，读者可自行作为练习).

对于一般整数(不一定是正整数)m 和 n，上面的定理也是成立的，虽然很直观，不过证明非常繁琐，读者直接使用就可以了.

另外，当不是抽象地讨论一些特定的群时，对群上的运算和元素使用自然的表示方法. 如对于普通加法下的群，使用“+”表示运算符；用 0 表示单位元；用 $-a$ 表示 a 的逆元；对正整数 n，用 na 表示 $\underbrace{a+a+\cdots+a}_{\text{共}n\text{项}}$，用 $-na$ 表示 $\underbrace{-a-a-\cdots-a}_{\text{共}n\text{项}}$，且 $0a=0$. 此时，上面定理中的三个表达式需要改写如下：

(1) $n(-a)=-(n\,a)$.

(2) $(n+m)a=na+ma$.

(3) $m(na)=(mn)a$.

定义 4.1.9 在群 G 中，对元素 a 来说，使 $a^n=1$ 的最小正整数 n 称为元素 a 的**阶**，记

为 **ord**(a). 如果不存在这样的正整数，那么称 a 为无限阶元素.

[例 4.1.8]　(1) 在任何群中，ord(1)=1，且只有单位元的阶为 1.

(2) 普通加法下的 **Z**，**Q**，**R**，**C** 中，每个非零数都是无限阶的.

(3) 普通乘法下的 **Q**−{0}和 **R**−{0}中，ord(−1)=1，其他不等于 1 和−1 的数都是无限阶的.

定理 4.1.7　群 G 中元素 a 的阶为 k，如果 $a^n=1$，那么 $k|n$.

证明　由带余除法可知，$n=qk+r$，其中 $0\leqslant r<k$. 所以

$$1=a^n=a^{qk+r}=(a^k)^q a^r=a^r,$$

由阶的定义中的“最小”性质可知，只能 $r=0$，即 $k|n$.

定理 4.1.8　有限群 G 中元素 a 的阶必为有限数.

证明　观察如下序列

$$1,\ a,\ a^2,\ \cdots,\ a^n,\ \cdots$$

由于以上序列中的元素都属于 G，且 G 是有限群，所以该无限长序列中必然存在重复的元素，设重复的元素为 a^m 和 a^n，其中 $m>n$，即

$$a^m=a^n,$$

所以

$$1=a^m a^{-n}=a^{m-n},$$

其中 $m-n>0$ 说明 a 的阶必为有限数.

下面给出几个稍微复杂一些的群的例子.

[例 4.1.9]　设 n 为大于等于 1 的整数，集合 $\mathbf{Z}_n=\{0,1,\cdots,n-1\}$，我们如下定义其上的二元运算⊕，对于任意 $a,b\in\mathbf{Z}_n$，有 $a\oplus b=(a+b)\%n$，其中“$\%n$”表示计算除以 n 以后得到的余数 mod n，则($\mathbf{Z}_n$，⊕)构成一个交换群，其单位元为 0，a 的逆元为 $n-a$.

[例 4.1.10]　参考上一个例子，p 是一个素数，集合 $\mathbf{Z}_p^*=\{1,\cdots,P-1\}$，即从 $\mathbf{Z}_p$ 里去掉元素 0，我们如下定义其上的二元运算⊗，对于任意 $a,b\in\mathbf{Z}_p^*$，有 $a\otimes b=a\times b\%n$，则($\mathbf{Z}_p^*$，⊗)是交换群，其中单位元为 1，a 的逆元与同余中定义的逆元相同.

[例 4.1.11]　Klein 四元群为集合 $G=\{a,b,c,e\}$，其上二元运算·的定义如表 4.1.1 所示.

表 4.1.1　Klein 四元群的定义

·	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

观察表 4.1.1，$\forall x\in G$，都有 $x\cdot e=e\cdot x=e$，所以 e 是单位元；$\forall x\in G$，都有 $x\cdot x=e$，所以 x 的逆元就是 x 自身. 尽管通过验证 $\forall x,y,z\in G$，都有 $(x\cdot y)\cdot z=x\cdot(y\cdot z)$，我们可以证明($G$，·)满足结合律，但是这种方法相当繁琐. 后面讲到的知识可以用来提供更好的证明方法.

4.1.2　子群

定义 4.1.10　(G，$*$)是一个群，子集 $H\subset G$，如果下面三条同时满足

(1) $1\in H$；

(2) 子集 H 在二元运算 $*$ 下是封闭的；

(3) 如果 $x\in H$，那么 $x^{-1}\in H$，

则称 H 是 G 的**子群**，记为 $H\leqslant G$. 由于$\{1\}$和 G 本身必然是 G 的子群，所以为了与其他子群进行区分，称$\{1\}$为**平凡子群**，否则为**非平凡子群**；如果子群 $H\neq G$，称 H 为**真子群**，记为 $H<G$.

定理 4.1.9 定义 4.1.10 中的子群必然是群.

证明 定义 4.1.10 中的条件(2)说明子群 H 上有一个二元运算，该二元运算是 G 上二元运算 $*$ 的限制，因此当然满足结合律(群条件 1). 定义 4.1.10 中的条件(1)即为(群条件 2)，定义 4.1.10 中的条件(3)即为(群条件 3). 所以，子群必然是群.

[**例 4.1.12**] 在普通加法下，$\mathbf{Z}\leqslant\mathbf{Q}$，$\mathbf{Q}\leqslant\mathbf{R}$，$\mathbf{R}\leqslant\mathbf{C}$.

[**例 4.1.13**] 群的子集是群但不是子群的例子：

在普通加法下 $\mathbf{R}$ 是群，在普通乘法下 $\mathbf{Q}-\{0\}$是群，$\mathbf{Q}-\{0\}$明显是 $\mathbf{R}$ 的子集，但是，$\mathbf{Q}-\{0\}$不是 $\mathbf{R}$ 的子群. 从这个例子可看到子群定义中要求子群和群的二元运算具有一致性的重要性.

[**例 4.1.14**] 在向量加法下，一个向量空间的任意子空间是它的子群.

定理 4.1.10 (子群判别标准)G 是一个群，它的子集 H 是子群当且仅当如下两个条件同时成立：

(1) $H\neq\varnothing$；

(2) $\forall x, y\in H$，$xy^{-1}\in H$.

证明 先证必要性. 如果 H 是子群，则 $1\in H$，所以 $H\neq\varnothing$. $\forall x, y\in H$，有 $y^{-1}\in H$，再由子群上运算的封闭性可知 $xy^{-1}\in H$.

再证充分性. 由(1)知，存在元素 $x\in H$，再由(2)知，$xx^{-1}\in H$，即 $1\in H$. $\forall x\in H$，由题设知 $1x^{-1}\in H$，即 $x^{-1}\in H$. 因此，$\forall y\in H$，$y^{-1}\in H$，由题设知，$\forall x$，$y^{-1}\in H$，$x(y^{-1})^{-1}\in H$，即 $xy\in H$，也即子集 H 在二元运算下是封闭的. 因此，H 满足子群的定义中的所有三个条件.

如果使用加法记号来表示群上的二元运算，则上面定理中条件(2)应该写为，$\forall x, y\in H$，$x-y\in H$.

[**例 4.1.15**] 在普通加法下，偶数集合是 $\mathbf{Z}$ 的子群；所有 3 的倍数组成的集合是 $\mathbf{Z}$ 的子群；更一般地，对固定的整数 n，所有 n 的倍数组成的集合是 $\mathbf{Z}$ 的子群. 命题正式陈述如下：

设 $n\in\mathbf{Z}$，令 $n\mathbf{Z}=\{n\times k\,|\,k\in\mathbf{Z}\}$，则$(n\mathbf{Z},+)$是$(\mathbf{Z},+)$的子群.

证明 显然 $n\mathbf{Z}$ 是 $\mathbf{Z}$ 的一个非空子集，且 $\forall a, b\in n\mathbf{Z}$，存在 $i, j\in\mathbf{Z}$ 使得

$$a=n\times i,\quad b=n\times j,$$

因此

$$a-b=n\times i-n\times j=n\times(i-j)\in n\mathbf{Z}.$$

由“子群判别标准”知$(n\mathbf{Z},+)$是$(\mathbf{Z},+)$的子群.

定理 4.1.11 G 是一个有限群，它的非空子集 H 是子群当且仅当子集 H 在 G 的二元运算下是封闭的.

证明 先证必要性. 由子群的定义可知.

再证充分性. 子集 H 非空，所以存在 $x\in H$，由于子集 H 在 G 的二元运算下是封闭的，

所以对任意正整数次幂有 $x^n \in H$. 因为 G 是一个有限群，所以 x 的阶必然有限，即存在正整数 m，使得 $x^m=1$，所以 $1 \in H$. $\forall y \in H$，由于 y 的阶必然有限，即存在正整数 k，使得 $y^k=1$，所以由逆元的定义可知，$y^{k-1}=y^{-1}$，因为 $y^{k-1} \in H$，所以 $y^{-1} \in H$. 因此，H 满足子群的定义中的所有三个条件.

上面定理中，如果允许 G 是一个无限群，那么结论不一定成立.

[例 4.1.16]　在普通加法下，$\mathbf{Z}$ 是一个无限群，考虑自然数集合 $\mathbf{N}$，非空且在普通加法下封闭，但是 $\mathbf{N}$ 不是 $\mathbf{Z}$ 的子群.

4.1.3　循环群和群的生成

定义 4.1.11　G 是一个群，且 $a \in G$，令集合

$$<a>=\{a^n \mid n \in \mathbf{Z}\},$$

则称集合 $<a>$ 为**由元素 a 生成的 G 的循环子群**.

定理 4.1.12　$<a>$ 是一个群，且是 G 的子群.

证明　由循环子群的定义可知，$1=a^0 \in <a>$；$(a^n)^{-1}=a^{-n} \in <a>$；$a^n a^m=a^{n+m} \in <a>$ 且结合律显然成立，满足群的定义中的条件，所以 $<a>$ 是一个群. 又由于 $<a>$ 显然是 G 的子集，所以 $<a>$ 是 G 的子群.

定义 4.1.12　G 是一个群，如果存在 $a \in G$ 使得，

$$G=<a>,$$

则称 G 为**循环群**，而且称 a 为 G 的**生成元**.

由前面的定义和定理显然可知任何群的循环子群必定是循环群. 另外，一个循环群可以有不只一个生成元，例如，如果集合

$$G=<a>=\{a^n \mid n \in \mathbf{Z}\},$$

则因为

$$\{a^n \mid n \in \mathbf{Z}\}=\{(a^{-1})^n \mid n \in \mathbf{Z}\}=<a^{-1}>,$$

所以有 $G=<a^{-1}>$.

[例 4.1.17]　$(\mathbf{Z}, +)$ 是交换群，任取 $a \in \mathbf{Z}$，则 $<a>=\{n \times a \mid n \in \mathbf{Z}\}$，则 $(<a>, +)$ 是 $(\mathbf{Z}, +)$ 的循环子群. 当 $a=0$ 时，这个子群仅由一个元素 0 组成；当 $a=1$ 时 $<1>=\mathbf{Z}$，所以 $(\mathbf{Z}, +)$ 是循环群，1 是 $\mathbf{Z}$ 的生成元. 当然也有当 $a=-1$ 时 $<-1>=\mathbf{Z}$，所以 -1 也是 $\mathbf{Z}$ 的生成元. 当 $a=2$ 或者 -2 时 $<2>=<-2>=$ 偶数集合，所以 2 和 -2 都是偶数集合的生成元. 注意，奇数集合不是 $\mathbf{Z}$ 的循环子群，其实奇数集合根本不是群，因为不含单位元 1.

[例 4.1.18]　集合 $\mathbf{Z}_6=\{0, 1, \cdots, 5\}$，1 是 $\mathbf{Z}_6$ 的生成元，另一个明显的生成元是 5. 它的子集 $\{0, 3\}$ 是一个循环子群，该子群的生成元只有一个是 3. 它的另一个子集 $\{0, 2, 4\}$ 也是一个循环子群，该子群的生成元是 2 和 4.

[例 4.1.19]　集合 $\mathbf{Z}_5^*=\{1, \cdots, 4\}$，即从 $\mathbf{Z}_5$ 里去掉元素 0，则 2 和 3 是它的生成元. 该集合的子集 $\{1, 4\}$ 是一个循环子群，生成元是 4. 我们可以举出很多这样的例子，详细内容，读者可以参考前面数论中有关原根的部分.

定义 4.1.13　如果 G 为有限群，则我们称它的元素数量为该**群的阶**，记为 $|G|$.

定理 4.1.13　如果 $G=<a>$ 是一个循环群，且 $|G|=n$，则当且仅当 $(k, n)=1$ 时，a^k

是G的生成元.

证明 先证必要性. 如果a^k是G的生成元，则$a\in<a^k>$，即存在整数s，使得$(a^k)^s=a$. 两边同时乘以a^{-1}得到$a^{ks-1}=1$. 由定理4.1.7可知，$n|ks-1$，即存在整数t使得$tn=ks-1$，所以$tn=ks-1$，就是$ks-tn=1$，由数论中的定理可知$(k,n)=1$.

再证充分性. 如果$(k,n)=1$，则存在整数s和t，使得$ks+tn=1$. 则$a=a^{ks+tn}=a^{ks}(a^n)^t=a^{ks}(1)^t=a^{ks}$，所以$a\in<a^k>$，因此$G=<a>\leqslant<a^k>$，但是由于明显有$<a^k>\leqslant G$，所以$G=<a^k>$，即$a^k$是$G$的生成元.

我们马上可以得到如下的推论.

推论 n阶循环群共有$\varphi(n)$个生成元.

[例4.1.20] 考虑集合$\mathbf{Z}_{12}=\{0,1,\cdots,11\}$，注意到$\varphi(12)=4$，所以共有4个生成元，最明显的一个是1. 与12互素的4个k为1，5，7，11. 所以其他3个生成元为5，7，11.

定理4.1.14 如果$G=<a>$是一个循环群，且$S\leqslant G$，则S必定是循环群，且如果k是使得$a^k\in S$的最小正整数，则a^k是S的生成元.

证明 当$S=\{1\}$时，命题显然成立. 下面设$S\neq\{1\}$. 因为$G=<a>$且$S\leqslant G$，所以S中的元素必然是a的幂. 如果$a^k\in S$，那么由于S是群，则$a^{-k}\in S$，因此S中必然存在a的正整数次幂. 设k是使得$a^k\in S$的最小正整数，则对任意$a^m\in S$，因为$m=tk+r$，其中$0\leqslant r<k$，则因为S是群，所以$a^{-k}\in S$，所以$a^r=a^m(a^{-k})^t$，所以$a^r\in S$，注意到r是非负数和k是使得$a^k\in S$的最小正整数，我们马上知道$r=0$，即$k|m$，由a^m的任意性可知S中的元素都是a^k的幂，且由S是群，可知a^k的任意幂都在S中，所以

$$S=\{(a^k)^n\,|\,n\in\mathbf{Z}\},$$

即S必定是循环群，它的生成元是a^k.

定理4.1.15 G是有限群，且$a\in G$，则$\mathrm{ord}(a)=|<a>|$.

证明 因为G是有限群，所以$\mathrm{ord}(a)$一定为有限数. 令$k=\mathrm{ord}(a)$，则$1,a,a^2,\cdots,a^{k-1}$这k个元素必然互不相同. 因为，假设存在重复，既存在$0\leqslant i<j\leqslant k-1$使得$a^i=a^j$，则$a^{j-i}=1$，其中$j-i$显然是小于$k$的正整数，这与$k=\mathrm{ord}(a)$的“最小”性质矛盾.

如果另集合$H=\{1,a,a^2,\cdots,a^{k-1}\}$，则$|H|=k$. 很明显，$H\subseteq<a>$. 对任意$a^i\in<a>$，由带余除法得到$i=qk+r$，其中$0\leqslant r<k$，则$a^i=a^{qk+r}=a^{qk}a^r=(a^k)^qa^r=a^r\in H$，所以$<a>\subseteq H$，所以$<a>=H$. 所以$|<a>|=|H|=k=\mathrm{ord}(a)$.

定理4.1.16 $G=<a>$是有限循环群，且$|G|=n$，则对任意整除n的正整数d，一定存在一个唯一的阶为d的循环子群，该循环子群为$<a^{n/d}>$.

证明 因$(a^{n/d})^d=a^n=1$，所以$\mathrm{ord}(a^{n/d})|d$. 又因为$(a^{n/d})^{\mathrm{ord}(a^{n/d})}=1$，所以$n|n/d\mathrm{ord}(a^{n/d})$，所以必然有$\mathrm{ord}(a^{n/d})/d$为整数，即$d|\mathrm{ord}(a^{n/d})$. 因此$d=\mathrm{ord}(a^{n/d})$. 由定理4.1.15可知

$$|<a^{n/d}>|=\mathrm{ord}(a^{n/d})=d,$$

即存在性得证.

下面证明唯一性. 令H是G的一个阶为d的循环子群，则H必然是循环子群，当然可以写成如下形式$H=<x>$. 因为$x\in G$，所以必然存在整数m使得$x=a^m$，所以$(a^m)^d=1$，所以$n|md$，所以存在整数k使得$md=nk$. 因此，$x=(a^{n/d})^k$，所以$H=<x>\subseteq<a^{n/d}>$，由于这两个子群的阶相同，所以必然有$H=<a^{n/d}>$.

接下来讨论由给定的子群构造新的子群的问题.

定理 4.1.17　$(G,\cdot)$是群，$\{(H_i,\cdot)\mid i\in I\}$是$(G,\cdot)$的一族子群，其中 I 是某个索引集合，则$(\bigcap_{i\in I}H_i,\cdot)$是$(G,\cdot)$的一个子群. 也就是说，子群的交集还是子群.

证明　对任意 $i\in I$，因为 H_i 是子群，所以 $1\in H_i$，所以 $1\in\bigcap_{i\in I}H_i$，即$\bigcap_{i\in I}H_i\neq\varnothing$；对任意 $a,b\in\bigcap_{i\in I}H_i$，都有 $a,b\in H_i(i\in I)$，即 a,b 属于这一族子群中的每一个子群，由定理 4.1.10 知 $a\cdot b^{-1}\in H_i(i\in I)$. 所以有 $a\cdot b^{-1}\in\bigcap_{i\in I}H_i$. 由定理 4.1.10 知$(\bigcap_{i\in I}H_i,\cdot)$是$(G,\cdot)$的一个子群. 证毕.

设$(G,\cdot)$为群，S 是G 的子集，则 S 对运算“$\cdot$”不一定封闭，即使 S 对运算“$\cdot$”封闭，S 也不一定是 G 的子群，那么给定子集 S 时，如何由 S 得到一个子群呢？

定义 4.1.14　设$(G,\cdot)$为群，S 是 G 的子集，设$(H_i\mid i\in I,\cdot)$是$(G,\cdot)$的所有包含集合 S 的子群，即 $S\subseteq H_i(i\in I)$，则$(\bigcap_{i\in I}H_i,\cdot)$叫作**由 S 生成的子群**，记为$(<S>,\cdot)$，S 中的元素叫子群$(<S>,\cdot)$的**生成元**.

容易证明，

$$<S>=\{b_{k_1}\cdot b_{k_2}\cdot\cdots\cdot b_{k_r}\mid b_{k_j}\in S\text{ 或 }b_{k_j}^{-1}\in S,\ j=1,\cdots,r,\ r\in\mathbf{N}\}.$$

特别是当 S 由一个元素 a 组成时，

$$<S>=<a>=\{a^n\mid n\in\mathbf{Z}\}.$$

当$(X,\cdot)$为交换群，且 S 由有限个元素 $a_1,a_2,\cdots,a_m$ 组成时，

$$<S>=<a_1,a_2,\cdots,a_m>=\{a_1^{n_1}\cdot\cdots\cdot a_m^{n_m}\mid n_1,\cdots,n_m\in\mathbf{Z}\}.$$

[例 4.1.21]　$(G,\cdot)$为群，$S=\varnothing$是 G 的子集，求$<\varnothing>$.

解　因为$\varnothing$包含于 G 的任意子群 H，所以$<\varnothing>$是 G 的所有子群的交集，而$\{1\}$是一个子群且包含于 G 的任意子群 H，所以 G 的所有子群的交集就是$\{1\}$，从而$<\varnothing>=\{1\}$.

[例 4.1.22]　循环群 $G=<a>$是由子集$\{a\}$生成的，在书写时使用记号$<a>$，而不使用记号$<\{a\}>$.

定义 4.1.15　设$(G,\cdot)$为群，如果存在 G 的子集 S，使得 $G=<S>$，且对 S 的任一真子集 S'，必有 $G\neq<S'>$，那么 S 称为群$(G,\cdot)$的**极小生成集**，也称为群$(G,\cdot)$的一组**基**. 当 S 是有限集时就说$(G,\cdot)$是**有限生成群**.

[例 4.1.23]　对任一 $a\in\mathbf{Z}$，令$<a>=\{n\times a\mid n\in\mathbf{Z}\}$，我们知道$(<a>,+)$是$(\mathbf{Z},+)$的子群. 当 $a=1$ 时$<1>=\mathbf{Z}$，所以$(\mathbf{Z},+)$是循环群，1 是它的生成元. 此外，$\{2,3\}$是它的一个极小生成集，即 $\mathbf{Z}=<1>=<\{2,3\}>$，且明显 $\mathbf{Z}\neq<2>=$偶数集合以及$\mathbf{Z}\neq<3>=3$的倍数集合. 参见如下定理.

定理 4.1.18　设 $a,b\in\mathbf{Z}$，$A=<a>$，$B=<b>$，$C=\{sa+tb\mid s,t\in\mathbf{Z}\}$，则

(1) $C=<d>$，其中 $d=(a,b)$；

(2) $A\cap B=<m>$，其中 $m=[a,b]$.

证明　(1) 由集合 C 的元素的形式，易知 C 是 $\mathbf{Z}$ 的子群，且 $C=<\{a,b\}>$. 由于 $\mathbf{Z}$ 是循环群，所以根据定理 4.1.14 可知，它的子群 C 也是循环群，即存在某个正整数 d，使得 $C=<d>$，其中 $d=C$ 中的最小正整数. 因为(a,b)是 a,b 的线性组合，所以$(a,b)\in C$，且其他 C 中的元素因为都是 a,b 的线性组合，所以必然都是(a,b)的倍数，因此(a,b)是 C 中的最小正整数. 所以 $d=(a,b)$，即 $C=<(a,b)>$.

(2) 因为 A 中的元素一定是 a 的倍数且 B 中的元素一定是 b 的倍数，所以 $A\cap B$ 中的元素一定是 a 和 b 的公倍数. 反过来，a 和 b 的任意公倍数一定属于 $A\cap B$. 因为 A 和 B 都是 $\mathbf{Z}$ 的子群，所以由定理 4.1.17 可知，$A\cap B$ 也一定是 $\mathbf{Z}$ 的子群，又由于 $\mathbf{Z}$ 是循环群，所以根据定理 4.1.14 可知，$A\cap B$ 也是循环群，即存在某个正整数 m，使得 $A\cap B=<m>$，其中 $m=A\cap B$ 中的最小正整数. 显然，$A\cap B$ 中的最小正整数为 $[a,b]$，所以 $A\cap B=<[a,b]>$.

4.1.4 陪集和拉格朗日定理

群的阶和它的子群的阶之间有一定的关系，这就是本节拉格朗日定理所揭示的内容.

定义 4.1.16 $(G,\cdot)$为群，$H\leqslant G$，$a\in G$，用符号 aH 表示如下的 G 的子集

$$aH=\{ah\mid h\in H\},$$

并且称这样的子集为子群 H 的**左陪集**.

明显，$a=a\cdot 1\in aH$. 如果 $a\notin H$，那么 $1\notin aH$，否则存在 $h\in H$，使得 $1=ah$，即 $a=h^{-1}\in H$，导出矛盾，这说明当 $a\notin H$ 时，左陪集 aH 不是群.

注意，如果对群上的二元运算采用“+”记号，则左陪集应该如下表示：

$$a+H=\{a+h\mid h\in H\}.$$

[例 4.1.24] 令$<3>=\{n\times 3\mid n\in\mathbf{Z}\}$，我们知道$(<3>,+)$是$(\mathbf{Z},+)$的子群，求出$<3>$的所有左陪集.

解 另 $a=0$，则相应的左陪集为 $A=\{0+n\times 3\mid n\in\mathbf{Z}\}=\{k\mid k\equiv 0(\mathrm{mod}3)\}$；另 $a=1$，则相应的左陪集为 $B=\{1+n\times 3\mid n\in\mathbf{Z}\}=\{k\mid k\equiv 1(\mathrm{mod}3)\}$；另 $a=2$，则相应的左陪集为 $C=\{2+n\times 3\mid n\in\mathbf{Z}\}=\{k\mid k\equiv 2(\mathrm{mod}3)\}$；当然我们还可以另 a 取其他数值，然后求相应的左陪集，经过计算不难发现，当 $a\equiv 0$ 时，得到的左陪集就是 A；当 $a\equiv 1$ 时，得到的左陪集就是 B；当 $a\equiv 2$ 时，得到的左陪集就是 C. 因此所有三个左陪集是仅有的解.

[例 4.1.25] 考虑向量空间 $\mathbf{R}^2$，则过原点的一条直线是它的子空间，因此必然是它的子群，那么任何一条与该直线平行的直线都是它的陪集.

另一方面，我们来研究利用群$(G,\cdot)$的一个子群$(H,\cdot)$来对 G 进行分类的问题，为了说明这一点，我们先看一个例子.

对于群$(\mathbf{Z},+)$，可以利用模 3 同余关系(注意，同余关系是等价关系)将 $\mathbf{Z}$ 划分成三个剩余类 C_0，C_1，C_2. 我们尝试用另一种方法来表达这种同余关系：令$<3>=\{n\times 3\mid n\in\mathbf{Z}\}$，则$(<3>,+)$是$(\mathbf{Z},+)$的子群. 再来看 $\mathbf{Z}$ 上的模 3 同余关系，由同余的定义知，若 $a\equiv b(\mathrm{mod}\ 3)$，则有 $3\mid(a-b)$，也有 $3\mid(b-a)$，也就是必存在某个整数 k，使得 $-a+b=3\times k$，而 $3\times k\in H$，故有 $a\equiv b(\mathrm{mod}\ 3)$等价于$-a+b\in<3>$.

因此，$\mathbf{Z}$ 上的任两个元素 a 与 b 模 3 同余的充要条件是$-a+b\in<3>$. 而同余关系是等价关系，因此，此例中 a 与 b 等价相当于$-a+b\in<3>$. 故可以用 $-a+b\in H$ 来确定等价关系，从而对 $\mathbf{Z}$ 进行分类. 这样，也可以说 $\mathbf{Z}$ 的剩余类是利用 $\mathbf{Z}$ 的子群$<3>$来划分的.

我们将这种情况推广至一般的群.

定义 4.1.17 设群$(H,\cdot)$为群$(G,\cdot)$的子群，确定 G 上的一个关系$\equiv$(注意该符号与同余符号的区别)，$a\equiv b$ 当且仅当 $a^{-1}\cdot b\in H$. 这个关系叫 G 上关于 H 的**左陪集关系**.

定理 4.1.19　设群$(H, \cdot)$为群$(G, \cdot)$的子群，则G上关于H的左陪集关系$\equiv$是等价关系.

证明

(1) 自反性：因为$a^{-1} \cdot a=1 \in H$，所以有$a \equiv a$.

(2) 传递性：设$a \equiv b$，$b \equiv c$，则有

$$a^{-1} \cdot b \in H, \quad b^{-1} \cdot c \in H,$$

$$a^{-1} \cdot c=a^{-1} \cdot (b \cdot b^{-1}) \cdot c=(a^{-1} \cdot b) \cdot (b^{-1} \cdot c) \in H,$$

故有$a \equiv c$.

(3) 对称性：若$a \equiv b$，则有$a^{-1} \cdot b \in H$，$(a^{-1} \cdot b)^{-1}=b^{-1} \cdot a \in H$，故有$b \equiv a$.

由于左陪集关系$\equiv$同时满足以上三个性质，所以它是等价关系.

因为左陪集关系是一个等价关系，我们可以利用左陪集关系对G进行分类.

定义 4.1.18　群$(G, \cdot)$的子群$(H, \cdot)$所确定的左陪集关系对G划分等价类，将下面的等价类叫作**以 a 为代表元的等价类**.

$$[a]=\{x \mid x \in G \text{ 且 } a \equiv x\}.$$

定理 4.1.20　设群$(H, \cdot)$为群$(G, \cdot)$的子群，则$[a]=aH$.

证明　对任意$x \in [a]$，因为$a \equiv x$，所以存在$h \in H$，使得$a^{-1} \cdot x=h$，因此$x=a\,h \in aH$，所以可知$[a] \subseteq aH$.

反过来，对任意$x \in aH$，存在$h \in H$，使得$x=ah$，所以$a^{-1} \cdot x=a^{-1}\,ah=h \in H$，即$a \equiv x$，所以$x \in [a]$，因此可知$aH \subseteq [a]$.

所以，$[a]=aH$. 证毕.

定理 4.1.21　设群$(H, \cdot)$为群$(G, \cdot)$的子群，a，$b \in G$，则

(1) $aH=bH$ 当且仅当$b^{-1} \cdot a \in H$. 特别地，$aH=H$ 当且仅当$a \in H$.

(2) 如果$aH \cap bH \neq \varnothing$，那么$aH=bH$.

(3) 对任意$a \in G$，$|aH|=|H|$.

证明　(1) 如果$aH=bH$，则对任意$x \in aH=bH$，存在h，$h' \in H$，使得$x=ah=bh'$，所以$b^{-1} \cdot a=h' \cdot h^{-1}$，因为$H$是子群，所以$h' \cdot h^{-1} \in H$，即$b^{-1} \cdot a \in H$.

反过来，如果$b^{-1} \cdot a \in H$，即$b \equiv a$，则对任意$x \in [b]$有$b \equiv x$，由$\equiv$的对称性和传递性可知，$a \equiv x$，所以$x \in [a]$，即$[b] \subseteq [a]$；又由$\equiv$的对称性可知，$a \equiv b$，则同理可得$[a] \subseteq [b]$. 所以$[a]=[b]$. 由定理 4.1.20 立即可知$aH=bH$.

因为$eH=H$，所以$aH=H=eH$ 当且仅当$e^{-1} \cdot a \in H$，即$a \in H$.

(2) 如果$aH \cap bH \neq \varnothing$，则存在$x \in aH$，$bH$，那么必然存在$h$，$h' \in H$，使得$x=ah=bh'$，所以$b^{-1} \cdot a=h' \cdot h^{-1}$，因为$H$是子群，所以$h' \cdot h^{-1} \in H$，即$b^{-1} \cdot a \in H$. 由(1)立即可知，$aH=bH$.

(3) 令函数$f: H \to aH$为$f(h)=ah$，则很明显这是一个双射，所以aH和H等势，即$|aH|=|H|$.

证毕.

定理 4.1.22　**(拉格朗日定理)** 设群$(H, \cdot)$为有限群$(G, \cdot)$的子群，则$|H|$是$|G|$的因子.

证明 设 $a_iH(1\leqslant i\leqslant n)$为 H 的所有共 n 个不同的左陪集，则

$$G=a_1H\cup a_2H\cup\cdots\cup a_nH,$$

由定理 4.1.21 可知 $a_iH(1\leqslant i\leqslant n)$两两不相交，所以

$$|G|=|a_1H|+|a_2H|+\cdots+|a_nH|,$$

由定理 4.1.21 可知$|a_iH|=|H|(1\leqslant i\leqslant n)$，所以$|G|=n|H|$，即$|H|$是$|G|$的因子. 证毕.

定义 4.1.19 设群$(G,\cdot)$有一个子群$(H,\cdot)$，则 H 在G 中的两两不相交左陪集组成的集合$\{aH|a\in G\}$叫作 H 在 G 中的**商集**，记为 G/H；G/H 中两两不相交的左陪集的个数叫作 H 在 G 中的**指标**，记为$[G:H]$.

如果$(G,\cdot)$为有限群，其阶为$|G|$，此时$[G:H]$就是定理 4.1.22 证明中的变量 n，所以

$$|G|=[G:H]|H|,$$

很明显，指标$[G:H]$也是$|G|$的因子.

[例 4.1.26] $(n\mathbf{Z},+)(n\in\mathbf{Z})$在整数加法群$(\mathbf{Z},+)$的中的商集为

$$\mathbf{Z}/n\mathbf{Z}=\{\{a+(n\mathbf{Z})\}|a\in\mathbf{Z}\}=\{\{a+nh|h\in\mathbf{Z}\}|a\in\mathbf{Z}\}=\{[0],[1],\cdots,[n-1]\}.$$

定理 4.1.23 设$(G,\cdot)$是有限群，且 $a\in G$，则 $\mathrm{ord}(a)$是$|G|$的因子.

证明 由定理 4.1.15 可知，$\mathrm{ord}(a)=|<a>|$. 而$<a>$是 G 的子群，则由拉格朗日定理可知，$|<a>|$是$|G|$的因子，即 $\mathrm{ord}(a)$是$|G|$的因子.

定理 4.1.24 设$(G,\cdot)$是有限群，则对任意 $a\in G$ 有，$a^{|G|}=1$.

证明 由定理 4.1.23 可知，$\mathrm{ord}(a)$是$|G|$的因子，即存在整数 k，使得$|G|=k\ \mathrm{ord}(a)$. 所以，$a^{|G|}=a^{k\,\mathrm{ord}(a)}=(a^{\mathrm{ord}(a)})^k=(1)^k=1$.

定理 4.1.25 设 p 是一个素数，群$(G,\cdot)$的阶为 p，则 G 必为循环群.

证明 选择一个元素 $a\in G$，且 $a\neq1$，则由拉格朗日定理可知$|<a>|$是 p 的因子，由于$|<a>|>1$，且 p 是一个素数，所以$|<a>|=P=|G|$，所以$<a>=G$.

如果模仿左陪集概念的定义方法，也可以定义所谓的右陪集，很明显使用右陪集的概念，一样能够推导出本节中的类似定理.

4.1.5 同态与同构

有些群虽然从表面上看似乎不同，但经仔细分析后发现它们是“相同”的，例如有两个群$(\{1,-1\},\times)$与$(\{0,1\},\oplus)$，运算$\times$和$\oplus$的定义分别见表 4.1.1 和表 4.1.2. 我们仔细考察这两个群后可以发现，如果将第二个群中的元素 0，1 分别换成第一个群中元素 1，-1，将第二个群中的运算符$\oplus$换成第一个群中的$\times$，那么得到的运算法则与第一个群的运算法则完全一样. 这说明，这两个群仅仅是元素和运算符的表示方式不同，而它们的实质是一样的，只要将表示形式统一后，这两个群完全可以看成一个群. 这种表示形式不同而实质上相同的群称为同构的群. 比较两个群在多大程度上类似是一个很重要的问题，这里关键的概念就是同态与同构.

表 4.1.1　代数系统($\{1,-1\},\times$)

$\times$	1	-1
1	1	-1
-1	-1	1

表 4.1.2　代数系统($\{0,1\},\oplus$)

$\oplus$	0	1
0	0	1
1	1	0

通过上面的例子可以看出，两个群同构必满足以下两个条件：

(1) 具有相同的阶；

(2) 运算法则对应相同，即一个群中的两个元素经过运算后的结果元素与另一个群中对应的两个元素经运算后所得的结果元素互相对应.

将这些性质进一步抽象得到群同构的概念.

定义 4.1.20　$(X,\cdot)$与$(Y,*)$是两个群，如果存在一个一一对应的映射(即双射)$f:X\to Y$，使得对任意 $x_1,x_2\in X$，都有

$$f(x_1\cdot x_2)=f(x_1)*f(x_2),$$

则称 f 是一个从$(X,\cdot)$到$(Y,*)$的**同构映射**，亦称群$(X,\cdot)$与$(Y,*)$**同构**，记作

$$(X,\cdot)\cong(Y,*).$$

一个群与自身的同构叫作**自同构**.

[**例 4.1.27**]　群 X 上的恒等映射 $1_X:X\to X$ 明显是一个**自同构**.

从同构的定义可以看出，要证明两个群$(X,\cdot)$与$(Y,*)$同构，关键是找到一个从$(X,\cdot)$到$(Y,*)$的同构映射.

[**例 4.1.28**]　群$(\mathbf{R},+)$和$(\mathbf{R}^+,\times)$同构

证明　对$(\mathbf{R},+)$与$(\mathbf{R}^+,\times)$有一个一一对应的映射 $f:\mathbf{R}\to\mathbf{R}^+$，

$$f(x)=e^x,$$

且对任意 $a,b\in\mathbf{R}$，有

$$f(a+b)=e^{a+b}=e^a\times e^b=f(a)\times f(b).$$

这个例子中的映射 f 的逆函数 $g:\mathbf{R}^+\to\mathbf{R}$ 为

$$g(x)=\ln(x),$$

它也是一个同构映射，因为

$$g(a\times b)=\ln(a\times b)=\ln a+\ln b=g(a)+g(b).$$

[**例 4.1.29**]　群$(\mathbf{C},+)$和$(\mathbf{R}^2,+)$同构

证明　对$(\mathbf{C},+)$与$(\mathbf{R}^2,+)$有一个一一对应的映射 $f:\mathbf{C}\to\mathbf{R}^2$，

$$f(a+ib)=(a,b),$$

且对任意 $a+ib,c+id\in\mathbf{C}$，有

$$\begin{aligned}f([a+ib]+[c+id])&=f([a+c]+i[b+d])=(a+c,b+d)\\&=(a,b)+(c,d)=f(a+ib)+f(c+id).\end{aligned}$$

实际上，不难证明同构是任意一族群上的一种等价关系，特别地，我们知道同构满足对称性，即如果$(X,\cdot)\cong(Y,*)$，那么$(Y,*)\cong(X,\cdot)$. 我们有如下定理.

定理 4.1.26　任意一族群上的同构关系是一种等价关系.

此定理的证明留作练习.

定理 4.1.27　如果群$(X,\cdot)$满足交换律，且$(X,\cdot)\cong(Y,*)$，则$(Y,*)$也满足交

换律.

证明 因为$(X, \cdot) \cong (Y, *)$，所以存在一个一一对应的映射 $f: X \to Y$. 因此，对任意 $a', b' \in Y$，存在 $a, b \in X$，使得 $f(a)=a'$ 和 $f(b)=b'$. 所以

$$a' * b' = f(a) * f(b) = f(a \cdot b) = f(b \cdot a) = f(b) * f(a) = b' * a'.$$

得证.

从这个定理可知，如果一个群满足交换律，而另一个群不满足交换律，则这两个群不可能同构. 另外，这个定理再一次说明同构的群之间的高度相似性，即同构的群有很多类似的性质，其他的例子包括：同构的群必然同时是或者不是循环群，同构的群中具有同样阶数的元素的数量必然相同，等等. 当对一个群了解很多信息时，如果能够证明另一个群与它同构，那么也就能够很好地了解这个新的群的相应信息. 尽管这种方法能够提供很大的好处，但是需要明确的是，一般来说确定两个群是否同构不是一件容易的事情.

群的同构条件是非常严格的，同构的两个群不仅要有相同的阶，而且要存在一一对应的同构映射，如果将这些条件放宽一点，将得到群之间的另外一种关系——同态.

定义 4.1.21 $(X, \cdot)$与$(Y, *)$是两个群，如果存在一个映射 $f: X \to Y$，使得对任意 $x_1, x_2 \in X$，都有

$$f(x_1 \cdot x_2) = f(x_1) * f(x_2),$$

则称 f 是一个从$(X, \cdot)$到$(Y, *)$的**同态映射**或称群$(X, \cdot)$与$(Y, *)$**同态**，记作

$$(X, \cdot) \backsim (Y, *).$$

如果 f 是单射，则称此同态为**单同态**，如果 f 是满射，则称此同态为**满同态**，如果 f 是双射，则此同态就是同构. 一个群与自身的同态叫**自同态**.

[例 4.1.30] $(X, \cdot)$与$(Y, *)$是两个群，我们把映射 $f: X \to Y$ 定义为，对任意 $x \in X$，都有

$$f(x) = e_Y,$$

其中 e_Y 是$(Y, *)$的单位元，那么 f 明显是一个简单的同态映射.

[例 4.1.31] 群$(\mathbf{Z}, +)$与群$(\mathbf{R}-\{0\}, \times)$同态，且是单同态，因为存在一个从 $\mathbf{Z}$ 到 $\mathbf{R}-\{0\}$的(单射)同态映射

$$f(x) = \mathrm{e}^x,$$

对任意 $x_1, x_2 \in \mathbf{Z}$，有

$$f(x_1 + x_2) = \mathrm{e}^{x_1 + x_2} = \mathrm{e}^{x_1} \times \mathrm{e}^{x_2} = f(x_1) \times f(x_2).$$

[例 4.1.32] 群$(\mathbf{Z}, +)$到群$(\mathbf{Z}_n, \oplus)$的映射 $f: \mathbf{Z} \to \mathbf{Z}_n$ 是

$$f(a) = a \% n$$

是一个同态映射，其中 $\%n$ 意义同例 4.1.9.

证明 对任意 $a, b \in \mathbf{Z}$，有

$$f(a+b) = (a+b) \% n,$$

$$f(a) \oplus f(b) = (a \% n) \oplus (b \% n) = ((a \% n) + (b \% n)) \% n = (a+b) \% n,$$

所以 $f(a+b) = f(a) \oplus f(b)$. 而且这个同态明显是一个满同态.

下面我们讨论一些关于同态的性质.

定理 4.1.28 设两个群满足$(S, \cdot) \backsim (G, \odot)$，$e$ 和 e' 分别为它们的单位元，同态映射

为 $f:S\to G$，则有

(1) $f(e)=e'$；

(2) 对任意 $a\in S$，$f(a^{-1})=f(a)^{-1}$；

(3) 对任意 $n\in\mathbf{Z}$ 和 $a\in S$，$f(a^n)=f(a)^n$.

证明(1) 因为群$(S,\cdot)$和群$(G,\odot)$同态，所以

$$f(e)=f(e\cdot e)=f(e)\odot f(e),$$

所以

$$e'=f(e)\odot f(e)^{-1}=(f(e)\odot f(e))\odot f(e)^{-1}=f(e)\odot(f(e)\odot f(e)^{-1})=f(e)\odot e'=f(e).$$

(2) $e'=f(e)=f(a^{-1}\cdot a)=f(a^{-1})\odot f(a)$，由逆元的定义有

$$f(a)^{-1}=f(a^{-1}).$$

(3) 对于 $n\geqslant 0$，我们能够利用数学归纳法很容易证明 $f(a^n)=f(a)^n$. 对于 $n<0$，有

$$f(a^n)=f(a^{-(-n)})=f((a^{-1})^{(-n)})=f((a^{-1}))^{(-n)}=(f(a)^{-1})^{(-n)}=f(a)^{-(-n)}=f(a)^n.$$

定义 4.1.22　设两个群满足$(S,\cdot)\backsim(G,\odot)$，$e$ 和 e'分别为它们的单位元，同态映射为 $f:S\to G$，令集合

$$\ker f=\{a\mid a\in S\text{ 且 }f(a)=e'\},$$

我们称该集合为同态 f 的**核**，且另集合

$$\operatorname{im} f=f(S)=\{f(a)\mid a\in S\},$$

我们称该集合为同态 f 的**像**.

[例 4.1.33]　$(X,\cdot)$与$(Y,*)$是两个群，我们把映射 $f:X\to Y$ 定义为，对任意 $x\in X$，都有

$$f(x)=e_Y,$$

那么 f 是一个同态映射，其中 $\ker f=X$ 和 $\operatorname{im} f=\{e_Y\}$.

[例 4.1.34]　群$(\mathbf{Z},+)$到群$(\mathbf{Z}_n,\oplus)$的映射 $f:\mathbf{Z}\to\mathbf{Z}_n$ 是 $f(a)=a\%n$，则

$$\ker f=n\mathbf{Z},\ \operatorname{im} f=\mathbf{Z}_n.$$

定理 4.1.29　设两个群满足$(S,\cdot)\backsim(G,\odot)$，$e$ 和 e'分别为它们的单位元，同态映射为 $f:S\to G$，则有

(1) $\ker f\leqslant S$(将 $\ker f$ 称为同态 f 的**核子群**)，且 f 是单同态的充要条件是 $\ker f=\{e\}$；

(2) $\operatorname{im} f\leqslant G$(将 $\operatorname{im} f$ 称为同态 f 的**像子群**)，且 f 是满同态的充要条件是 $f(S)=G$；

(3) 如果 $G'\leqslant G$，$f^{-1}(G')=\{a\mid a\in S\text{ 且 }f(a)\in G'\}$，则 $f^{-1}(G')\leqslant S$.

证明　(1)由定理 4.1.28 知 $f(e)=e'$，所以 $e\in\ker f$，即 $\ker f$ 不是空集.

对任意 $a,b\in\ker f$，有 $f(a)=e'$，$f(b)=e'$，

$$f(a\cdot b^{-1})=f(a)\odot f(b^{-1})=e'\odot f(b)^{-1}=e'\odot(e')^{-1}=e',$$

因此有 $a\cdot b^{-1}\in\ker f$，由定理 4.1.10 的子群判别标准知 $\ker f\leqslant S$.

设 f 为单同态，则 S 中满足 $f(a)=e'=f(e)$的元素只有 $a=e$，因此 $\ker f=\{e\}$.

反过来，设 $\ker f=\{e\}$，对任意满足 $a,b\in S$，若 $f(a)=f(b)$，则必有

$$f(a\cdot b^{-1})=f(a)\odot f(b^{-1})=f(a)\odot f(b)^{-1}=f(a)\odot f(a)^{-1}=e',$$

因此 $a\cdot b^{-1}\in\ker f$，$a\cdot b^{-1}=e$，所以 $a=b$. 因此，f 是单同态.

(2) 由定理 4.1.28 可知 $f(e)=e'$，所以 $e'\in\operatorname{im} f$，即 $\operatorname{im} f$ 不是空集.

对任意 x，$y\in \mathrm{im}\ f$，必存在 a，$b\in S$，使得 $f(a)=x$，$f(b)=y$，显然 $a\cdot b^{-1}\in S$(S 是群)，

$$x\odot y^{-1}=f(a)\odot f(b)^{-1}=f(a)\odot f(b^{-1})=f(a\cdot b^{-1}).$$

因为 $a\cdot b^{-1}\in S$，所以 $f(a\cdot b^{-1})\in \mathrm{im}\ f$，即 $x\odot y^{-1}\in \mathrm{im}\ f$，由定理 4.1.10 的子群判别标准知 $\mathrm{im}\ f\leqslant G$.

由满同态的定义知 f 为满同态的充要条件为 $f(S)=G$.

(3) 因为 $G'\leqslant G$，所以 $e'\in G'$，由定理 4.1.28 知 $f(e)=e'$，所以 $e\in f^{-1}(G')$，即 $f^{-1}(G')$不是空集. 对任意 a，$b\in f^{-1}(G')$，必存在 x，$y\in G'$，满足 $f(a)=x$，$f(b)=y$，因为 $G'\leqslant G$，所以有

$$x\odot y^{-1}\in G',$$

即

$$x\odot y^{-1}=f(a)\odot f(b)^{-1}=f(a)\odot f(b^{-1})=f(a\cdot b^{-1})\in G',$$

所以 $a\cdot b^{-1}\in f^{-1}(G')$，由定理 4.1.10 的子群判别标准知 $f^{-1}(G')\leqslant S$.

4.1.6 正规子群与商群

任何同态的核子群都具有一个特殊的性质，即核子群一定是所谓的正规子群，下面就来正式给出正规子群的定义.

定义 4.1.23 设两个群满足 $K\leqslant G$，如果对任意 $k\in K$ 和 $g\in G$ 都有 $gkg^{-1}\in K$，则称 K 为 G 的**正规子群**，记为 $K\lhd G$.

定理 4.1.30 设 f 是群$(S,\cdot)$到群$(G,\odot)$的同态映射，e 和 e'分别是$(S,\cdot)$和$(G,\odot)$的单位元，则 $\ker f\lhd S$；

证明 对任意 $a\in S$，$b\in \ker f$，有

$$f(a\cdot b\cdot a^{-1})=f(a)\odot f(b)\odot f(a^{-1})=f(a)\odot e'\odot f(a)^{-1}=f(a)\odot f(a)^{-1}=e',$$

所以 $a\cdot b\cdot a^{-1}\in \ker f$，由正规子群的定义知 $\ker f\lhd S$. 证毕.

定理 4.1.31 任意交换群 G 的每个子群 K 都是正规子群.

证明 对任意 $g\in G$，$k\in K$，有

$$gkg^{-1}=kgg^{-1}=ke=k,$$

所以 $gkg^{-1}\in K$，由正规子群的定义知 $K\lhd G$. 证毕.

[**例 4.1.35**] 设 $n>1$ 是整数，则$(n\mathbf{Z},+)$是$(\mathbf{Z},+)$的正规子群. 因为对任意 $g\in\mathbf{Z}$，$k\in n\mathbf{Z}$，由“+”运算满足交换律，所以有

$$g+k-g=k\in n\mathbf{Z},$$

因此 $n\mathbf{Z}$ 是 $\mathbf{Z}$ 的一个正规子群. 注意，$(n\mathbf{Z},+)=(<n>,+)$.

4.1.4 节的最后曾经提到，如果模仿左陪集概念的定义方法，也可以定义所谓的右陪集，为了继续关于正规子群的讨论，给出右陪集相关概念的定义和不加证明的有关定理.

定义 4.1.24 $(G,\cdot)$为群，$H\leqslant G$，$a\in G$，我们用符号 Ha 表示如下的 G 的子集

$$Ha=\{ha\,|\,h\in H\},$$

并且称这样的子集为子群 H 的**右陪集**.

定义 4.1.25　$(G,\cdot)$为群，$H\leqslant G$，我们确定 G 上的一个关系$\equiv$，$a\equiv b$ 当且仅当 $a\cdot b^{-1}\in H$. 这个关系叫 G 上关于 H 的**右陪集关系**.

定理 4.1.32　设群$(H,\cdot)$为群$(G,\cdot)$的子群，则 G 上关于 H 的右陪集关系$\equiv$是等价关系.

定义 4.1.26　群$(G,\cdot)$的子群$(H,\cdot)$所确定的右陪集关系对 G 划分等价类，我们将下面的等价类叫**作以 a 为代表元的等价类**.

$$[a]=\{x\mid x\in G\text{ 且 }a\equiv x\}.$$

定理 4.1.33　设群$(H,\cdot)$为群$(G,\cdot)$的子群，则$[a]=Ha$.

定理 4.1.34　群$(G,\cdot)$的子群$(N,\cdot)$是正规子群的充要条件是：对任意 $a\in G$ 有，

$$aN=Na.$$

证明　充分性. 对任意 $a\in G$，因为 $aN=Na$，所以对任意 $n\in N$ 必存在一个 $s\in N$，使得

$$a\cdot n=s\cdot a,$$

故有

$$a\cdot n\cdot a^{-1}=s\cdot a\cdot a^{-1}=s\in N,$$

即$(N,\cdot)$是正规子群.

必要性. 因为$(N,\cdot)$是正规子群，所以对任意 $a\in G$ 和任意 $n\in N$ 都有，$a\cdot n\cdot a^{-1}\in N$，故存在一个 $s\in N$ 使得

$$a\cdot n\cdot a^{-1}=s,$$

于是

$$a\cdot n=s\cdot a,$$

因此对任意 $a\cdot n\in aN$ 必存在 $s\in N$，使得 $a\cdot n=s\cdot a\in Na$，即 $a\cdot n\in Na$，所以 $aN\subseteq Na$.

反过来对任意 $a\in G$ 和任意 $n\in N$ 都有，$a^{-1}\cdot n\cdot(a^{-1})^{-1}\in N$，故存在一个 $s\in N$ 使得

$$a^{-1}\cdot n\cdot(a^{-1})^{-1}=s,$$

于是

$$n\cdot a=a\cdot s,$$

因此对任意 $n\cdot a\in Na$ 必存在 $s\in N$，使得 $n\cdot a=a\cdot s\in aN$，即 $n\cdot a\in aN$，所以 $Na\subseteq aN$.

综上得 $aN=Na$. 证毕.

从这个定理可知，由正规子群形成的陪集没有左右之分，此时就能够使用陪集这个术语，相应地，由正规子群形成的商集 G/N 是由没有左右之分的陪集组成的.

定理 4.1.35　设群$(G,\cdot)$有一个正规子群$(N,\cdot)$，$T=G/N$ 是 N 在 G 中的商集，在商集 T 上定义二元运算“$\odot$”为：对任意 $aN,bN\in T(a,b\in G)$，

$$aN\odot bN=(a\cdot b)N,$$

则$(T,\odot)$构成群.

证明　首先证明运算“$\odot$”的定义中两个任意元素(陪集)的计算结果不依赖于陪集代表元的选择. 即要证明对任意 $aN=a'N$，$bN=b'N$，都有$(ab)N=(a'b')N$ 成立. 事实上，由 G 中的结合律和正规子群的性质，有

$$(ab)N=a(bN)=a(b'N)=a(Nb')=(aN)b'=(a'N)b'=a'(Nb')=a'(b'N)=(a'b')N.$$

首先，“$\odot$”满足结合律，因为

$$(aN\odot bN)\odot cN=(a\cdot b)N\odot cN=((a\cdot b)\cdot c)N=(a\cdot(b\cdot c))N$$
$$=aN\odot(b\cdot c)N=aN\odot(bN\odot cN).$$

设 e 是$(G,\cdot)$的单位元，则 $eN=N$ 是$(T,\odot)$的单位元. 这是因为对任意 $a\in G$，

$$aN\odot N=aN\odot eN=(a\cdot e)N=aN,$$
$$N\odot aN=eN\odot aN=(e\cdot a)N=aN.$$

对任意 $a\in G$，aN 存在逆元 $a^{-1}N$. 事实上，

$$aN\odot a^{-1}N=(a\cdot a^{-1})N=eN=N,$$
$$a^{-1}N\odot aN=(a^{-1}\cdot a)N=eN=N.$$

综上所述可知，$(T,\odot)$构成群. 证毕.

[例 4.1.36] 在商集 $\mathbf{Z}/n\mathbf{Z}=\{\{a+(n\mathbf{Z})\}\mid a\in\mathbf{Z}\}=\{[0],[1],\cdots,[n-1]\}$上如下定义运算$\oplus$：

$$[a]\oplus[b]=[a+b],$$

则$(\mathbf{Z}/n\mathbf{Z},\oplus)$是一个群.

定义 4.1.27 定理 4.1.35 中的群$(T,\odot)$叫作群$(G,\cdot)$对正规子群$(N,\cdot)$的**商群**. 记为

$$(T,\odot)=(G,\cdot)/(N,\cdot)=(G/N,\odot).$$

实际上，不难证明例 4.1.9 中定义的群$(\mathbf{Z}_n,\oplus)$与例 4.1.36 中的商群是同构的，即

$$(\mathbf{Z}_n,\oplus)\cong(\mathbf{Z}/n\mathbf{Z},\oplus),$$

同构的群没有本质差别，因此，可以用 $\mathbf{Z}/n\mathbf{Z}$ 的定义来代替 $\mathbf{Z}_n$ 的定义，即将 $\mathbf{Z}_n$ 定义为

$$\mathbf{Z}_n=\{\{a+(n\mathbf{Z})\}\mid a\in\mathbf{Z}\}=\{[0],[1],\cdots,[n-1]\}.$$

下面来讨论正规子群与群同态的关系.

定理 4.1.36 如果两个群满足$(N,\cdot)\lhd(S,\cdot)$，按定理 4.1.35 构造商群$(S/N,\odot)$，且如下定义映射 $f:S\to S/N$，

$$f(a)=aN,$$

则 f 是一个同态映射，且 $\ker f=N$.

证明 映射 f 满足

$$f(a\cdot b)=(a\cdot b)N=aN\odot bN=f(a)\odot f(b),$$

所以 f 是从群$(S,\cdot)$到群$(S/N,\odot)$的同态映射. 而群$(S/N,\odot)$的单位元为 N，设 $a\in S$，如果

$$f(a)=N,$$

则有

$$aN=f(a)=N=eN,$$

由定理 4.1.21，$aN=eN$ 的充要条件是 $e^{-1}\cdot a\in N$，即 $a\in N$，所以 $\ker f=N$.

定义 4.1.28 $(N,\cdot)\lhd(S,\cdot)$，定义映射 $f:S\to S/N$，

$$f(a)=aN,$$

则 f 是群$(S,\cdot)$到其商群$(S/N,\odot)$的一个同态映射，由 f 建立的从群$(S,\cdot)$到群$(S/N,\odot)$的同态叫作**自然同态**.

定理 4.1.37 （基本同构定理）设 $f:S\to G$ 是群$(S,\cdot)$到群$(G,\times)$的同态映射，则存

在 S/ker f 到 im f 的映射 h :$S/\ker f \to \text{im } f$，使得

$$(S/\ker f, \odot) \cong (\text{im } f, \times).$$

证明　设群$(S, \cdot)$的单位元为 e，群$(G, \times)$的单位元为 e'. 由定理 4.1.30 可知

$$(\ker f, \cdot) \lhd (S, \cdot),$$

所以商群$(S/\ker f, \odot)$一定存在，由定理 4.1.35 可知商群$(S/\ker f, \odot)$的单位元为 $\ker f$. 如下定义映射 h :$S/\ker f \to \text{im } f$，对任意陪集 $x\ker f \in S/\ker f$，有

$$h(x\ker f) = f(x),$$

则对任意 $a\ker f$，$b\ker f \in S/\ker f$，有

$$\begin{aligned} h(a\ker f \odot b\ker f) &= h((a \cdot b)\ker f) \\ &= f(a \cdot b) \\ &= f(a) \times f(b) \\ &= h(a\ker f) \times h(b\ker f) \end{aligned}$$

所以 h 是从群$(S/\ker f, \odot)$到群$(\text{im } f, \times)$的同态映射.

其次，h 是单同态映射. 事实上，由定理 4.1.29 可知$(\text{im } f, \times)$是$(G, \times)$的子群，故$(\text{im } f, \times)$的单位元是 e'. 对任意 $a\ker f \in \ker h$，下面两式同时成立，

$$h(a\ker f) = f(a),$$

$$h(a\ker f) = e',$$

所以有

$$f(a) = e',$$

即 $a \in \ker f$，由定理 4.1.21 知当 $a \in \ker f$ 时，$a\ker f = \ker f$，即 $\ker h$ 中只有一个元素 $\ker f$，所以有

$$\ker h = \{\ker f\}.$$

而 $\ker f$ 是商群$(S/\ker f, \odot)$的单位元，由定理 4.1.29 可知 h 为单同态映射.

最后来证明 h 是满同态. 事实上，对任意 $c \in \text{im } f$，存在 $a \in S$，使得 $f(a) = c$，从而

$$h(a\ker f) = f(a) = c.$$

所以任意 im f 中的元素 c 在映射 h 下都有原像 $a\ker f$. 至此我们证明了 h 是从群$(S/\ker f, \odot)$到群$(\text{im } f, \times)$的单同态映射，又是满同态映射，所以 h 是从群$(S/\ker f, \odot)$到群$(\text{im } f, \times)$的同构映射.

4.1.7　循环群的分类

前面给出过循环群的定义，循环群是一类重要的群，在计算机密码学中有着重要的应用. 现在有了基本同构定理，就可以更方便地对循环群进行分类.

定理 4.1.38　整数加法群$(\mathbf{Z}, +)$的每个子群 H 都是循环群，并且有

$$H = \langle 0 \rangle \text{ 或 } H = \langle m \rangle = m\mathbf{Z},$$

其中 m 是 H 中的最小正整数. 如果 $H \neq \langle 0 \rangle$，则 H 是无限群.

证明　当 $H = (\{0\}, +)$时，显然 $H = \langle 0 \rangle$，而且 H 是有限群.

当 H 为$(\mathbf{Z}, +)$的非零子群时，H 中至少存在一个非零整数 a，因 H 是群，所以也有 $-a \in H$，所以 H 中一定存在正整数. 设 H 中的最小正整数为 m，对任意 $a \in H$，由带余除

法，一定存在整数 q，r 使得

$$a=q\times m+r \text{ 且 } 0\leqslant r<m.$$

所以由群的性质可知 $r=a-q\times m\in H$，如果 $r\neq 0$，则 $0<r<m$，这与 m 的最小性矛盾. 因此 $r=0$，$a=qm\in m\mathbf{Z}$，故 $H\subseteq m\mathbf{Z}$；当群 H 中含有正整数 m 时，由群的基本性质可知 m 的任意倍数也必然在群 H 中，即 $m\mathbf{Z}\subseteq H$，综上所述，所以有 $H=m\mathbf{Z}$.

下面是有关分类的定理.

定理 4.1.39 每个无限循环群同构于整数加法群$(\mathbf{Z}, +)$，每个阶为 m 的有限循环群同构于$(\mathbf{Z}/m\mathbf{Z}, \oplus)$.

证明 设循环群$(G, \cdot)$的生成元为 a，则有

$$G=<a>=\{a^n \mid n\in\mathbf{Z}\}.$$

构造从 $\mathbf{Z}$ 到 G 的映射 $f:\mathbf{Z}\to G$，使得对任意 $n\in\mathbf{Z}$，

$$f(n)=a^n,$$

则对任意 m，$n\in\mathbf{Z}$ 有

$$f(m+n)=a^{m+n}=a^m\cdot a^n=f(m)\cdot f(n).$$

且对每个 $a^n\in G$，都能在 $\mathbf{Z}$ 中找到一个 n 使得映射 f 成立，故 f 为从群$(\mathbf{Z}, +)$到群$(G, \cdot)$的满同态，即 $im\ f=G$. 由定理 4.1.37 可知

$$\mathbf{Z}/\ker f \cong (G, \cdot).$$

而 $\ker f$ 是 $\mathbf{Z}$ 的子群，由定理 4.1.38 可知，$\ker f=<0>$或 $\ker f=m\mathbf{Z}$，前者对应于无限循环群，后者对应于 m 阶的有限循环群.

这个定理告诉我们如下两点：

(1) 无限循环群同构于整数加法群. 也就是说，对无限循环群的研究可归结为对整数加法群的研究. 而整数加法群的各种性质在初等数学中已有深刻的分析揭示. 人们对整数加法群的认识已有数千年的历史，对它的性质几乎已全部了解，所以可以说，无限循环群的问题已基本解决. 阶为 m 的循环群同构于整数模 m 剩余类加法群，对它的研究可归结为对剩余类加法群的研究. 而对剩余类加法群的研究在数论中已有深刻的分析，此类循环群的问题也已基本解决.

(2) 同样阶数的两个循环群必然同构，因为它们都同构于某个 $\mathbf{Z}/m\mathbf{Z}$，且群之间的同构关系是一种邓加关系，具有传递性和对称性，因此它们必然同构.

[例 4.1.37] $\mathbf{Z}_n=\{[0], [1], \cdots, [n-1]\}$，$(\mathbf{Z}_n, \oplus)$是有限循环群. 其中$[1]$是生成元，与 n 互素的任一整数也是生成元.

4.1.8 置换群

置换群是一类重要的群，它们的元素是置换，实际上密码学上任何分组加密方法都可以看作将所有可能的明文进行置换得到所有可能的密文. 置换群不仅被人们看作一种需要加以深入研究的群，而且置换群还是最能说明群的各种概念的例子. 本节对置换群进行简要的讨论.

定义 4.1.29 给定非空集合 X，将任意一个双射 α：$X\to X$ 称作集合 X 的一个**置换**.

由上面的定义可知，置换实际上就是双射函数，如果把函数的复合“。”看作一种置换间

的二元运算，那么由非空集合 X 的所有置换组成的集合就是一个群，将这个群用符号记为 $(S_X, \circ)$，因为

(1) 封闭性：任意选择两个置换 $\alpha: X \to X$ 和 $\beta: X \to X$，因为它们都是双射，所以复合函数 $\alpha \circ \beta: X \to X$ 也是双射，即置换，因此“$\circ$”是 S_X 上的(封闭的)二元运算.

(2) 结合律：由于一般的函数的复合是满足结合性质的，所以“$\circ$”满足结合律.

(3) 单位元：我们定义恒等置换 $1_X: X \to X$ 为，对任意 $x \in X$，$1_X(x)=x$，则对任意置换 α，

$$\alpha \circ 1_X = 1_X \circ \alpha = \alpha,$$

所以 1_X 是单位元.

(4) 逆元：对任意置换 $\alpha: X \to X$，因为是双射，所以存在双射的逆函数 $\alpha^{-1}: X \to X$，满足

$$\alpha \circ \alpha^{-1} = \alpha^{-1} \circ \alpha = 1_X,$$

所以置换 α^{-1} 是 α 的逆元.

综上所述，$(S_X, \circ)$ 满足所有的群条件，所以的确是一个群.

定义 4.1.30　我们将上述的群 $(S_X, \circ)$ 称为集合 X 上的**对称群**. 当 $X=\{1, 2, \cdots, n\}$ 时，称 S_X 为 ***n* 次对称群**，记作 S_n.

可以用如下的两行记号来表达 S_n 中的置换 α：

$$\alpha = \begin{pmatrix} 1 & 2 & \cdots & n \\ \alpha(1) & \alpha(2) & \cdots & \alpha(n) \end{pmatrix}$$

利用排列组合的知识很容易得到 S_n 的元素数量是 $n!$. 下面讨论如何用另一种方式表达这么多的置换.

定义 4.1.31　设 $\alpha \in S_n$，$A=\{i_1, i_2, \cdots, i_r\} \subseteq \{1, 2, \cdots, n\}$，$B=\{1, 2, \cdots, n\}-A$，如果置换 α 满足，

(1) 对 A 中的元素有 $\alpha(i_1)=i_2$，$\alpha(i_2)=i_3$，$\cdots$，$\alpha(i_{r-1})=i_r$，$\alpha(i_r)=i_1$；

(2) 对任意 $i \in B$ 有，$\alpha(i)=i$；

则称置换 α 为一个 ***r*－轮换**，记为 $\alpha=(i_1, i_2, \cdots, i_r)$. 也把 2－轮换称为**对换**.

[**例 4.1.38**]　一个 3－轮换为 $\alpha=(2, 1, 3) \in S_5$，它的意思就是 $\alpha(1)=3$，$\alpha(2)=1$，$\alpha(3)=2$，$\alpha(4)=4$，$\alpha(5)=5$.

[**例 4.1.39**]　任意一个 1－轮换都等于恒等置换，比如 S_3 中共有 3 个 1－轮换，那么我们立即可知 $(1)=(2)=(3)$.

[**例 4.1.40**]　有如下的置换之间的等式

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 4 & 2 & 3 \end{pmatrix} = (1 \quad 5 \quad 3 \quad 4 \quad 2)$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 4 & 5 \end{pmatrix} = (1 \quad 2 \quad 3)$$

我们能够将任意置换分解为多个轮换的乘积(从现在开始，用“乘积”来称呼置换之间的复合，且在用符号表示两个置换复合的时候，省略“$\circ$”)，将此称为置换的轮换分解.

[**例 4.1.41**]　如下的置换分解为 2 个轮换之积

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 4 & 3 \end{pmatrix} = (1 \quad 2)\,(3 \quad 4)$$

[例 4.1.42] 如下的置换 $\alpha\in S_5$ 可以用两种不同的轮换的乘积进行表示

$$\alpha=(1\ \ 2)(1\ \ 3\ \ 4\ \ 2\ \ 5)(2\ \ 5\ \ 1\ \ 3)=(1\ \ 4)(3\ \ 5)(2).$$

观察例 4.1.42 可见，尽管该置换可以用两种不同的轮换的乘积进行表示，但是显然第二种方式更简洁明了，因此，引入如下的概念.

定义 4.1.32 设 $\alpha,\beta\in S_n$ 是两个轮换，且这两个轮换的记号中没有共同的数字，则称 α 和 β **不相交**. 如果一组轮换中任意两个轮换都不相交，则称该组轮换**不相交**.

[例 4.1.43] 如下的置换 $\alpha\in S_5$ 是 3 个不相交的轮换的乘积

$$\alpha=(1\ \ 4)(3\ \ 5)(2).$$

由于 1－轮换都等于恒等置换，所以下面的讨论中一律不再写出 1－轮换，对于恒等置换，写作 1_n. 例如：上例中的置换写作如下的轮换分解

$$\alpha=(1\ \ 4)(3\ \ 5).$$

对于任意 $\alpha\in S_n$，如果已知它的轮换分解，那么求出它的逆置换的方法为：将它的轮换分解中的每一个轮换中的数字倒排即可.

[例 4.1.44] $\alpha=(1\ \ 4)(3\ \ 5)$，求出 $\alpha^{-1}=(4\ \ 1)(5\ \ 3)$.

另外，由于 S_2 只有两个元素，明显是一个交换群. 注意，对于 $n\geqslant 3$，S_n 是非交换群.

[例 4.1.45] 对于任意 $S_n(n\geqslant 3)$ 都有

$$(1\ \ 2)(1\ \ 3)=(1\ \ 3\ \ 2)$$

和

$$(1\ \ 3)(1\ \ 2)=(1\ \ 2\ \ 3),$$

所以

$$(1\ \ 2)(1\ \ 3)\neq(1\ \ 3)(1\ \ 2).$$

尽管，对于 $n\geqslant 3$，S_n 是非交换群，然而，其中的很多元素是可交换的，特别是不相交的轮换是可交换的.

定理 4.1.40 不相交的轮换是可交换的.

证明 因为轮换只针对自身记号内的数字进行换位，而对自身记号外的数字的作用只是固定该值，所以两个不相交的轮换在执行上不论谁先谁后，总体效果是一样的，即不相交的轮换是可交换的.

因此对一个置换的不相交轮换分解来说，随意调整其中各轮换的次序不会改变该置换.

[例 4.1.46] (1 4)、(3 5)和(2 6)是 3 个不相交的 2－轮换，因此有

$$(1\ \ 4)(3\ \ 5)(2\ \ 6)=(3\ \ 5)(1\ \ 4)(2\ \ 6)=(2\ \ 6)(3\ \ 5)(1\ \ 4).$$

注意，实际上一个轮换有不同的记号方式，即：将表示该轮换的记号中的数字进行循环换位，则不会改变该轮换(但是，习惯上一般将轮换中最小的数写在第一个位置).

[例 4.1.47] (1 2 3)=(2 3 1)=(3 1 2). 其中，一般采用(1 2 3)这个记号.

如果遵守以下的规定，我们将不加证明地给出重要的定理 4.1.41：

(1) 对一个置换的不相交轮换分解，随意调整其中各轮换的次序，我们把这些不同的记法看作同一个轮换分解；

(2) 我们把一个轮换的不同的记号方式看作同一个轮换.

(3) 对一个置换的不相交轮换分解，一定去掉任何 1－轮换.

定理 4.1.41　S_n 中的任意置换一定能够分解为不相交轮换的乘积，且这种分解是唯一的.

下面讨论置换的阶.

一个 r－轮换 $\alpha=(i_1, i_2, \cdots, i_r)$ 的 k 次幂 α^k 就是连续施行 k 次该 r－轮换 α，当 $k\leqslant r-1$ 时，$\alpha^k(i_1)=i_{1+k}\neq i_1$，所以 $\alpha^k\neq 1_n$；当 $k=r$ 时，对任意 m 有，$\alpha^k(i_m)=i_m$，所以 $\alpha^k=1_n$. 综上所述可知，$\mathrm{ord}(\alpha)=r$. 对任意置换来说，为了求得它的阶，则首先将该置换进行不相交的轮换分解，那么该置换的阶就等于所有轮换因子的长度的最小公倍数.

[例 4.1.48]　观察 3－轮换$(1\quad 2\quad 3)\in S_3$，

$$(1\quad 2\quad 3)=\begin{pmatrix}1&2&3\\2&3&1\end{pmatrix},$$

$$(1\quad 2\quad 3)(1\quad 2\quad 3)=\begin{pmatrix}1&2&3\\3&1&2\end{pmatrix},$$

$$(1\quad 2\quad 3)(1\quad 2\quad 3)(1\quad 2\quad 3)=\begin{pmatrix}1&2&3\\1&2&3\end{pmatrix},$$

所以，$(1\quad 2\quad 3)$的阶是 3.

[例 4.1.49]　求 $\alpha=(1\quad 2\quad 3)(4\quad 5)\in S_5$ 的阶.

解　$\mathrm{ord}(\alpha)=[3, 2]=6$.

定义 4.1.33　我们将任意对称群的任意子群称为一个**置换群**.

[例 4.1.50]　S_4 的子集 $V=\{(1),(1\quad 2)(3\quad 4),(1\quad 3)(2\quad 4),(1\quad 4)(2\quad 3)\}$是一个子群，即 V 是一个置换群.

证明　恒等置换$(1)\in S_4$；对任意 $\alpha\in V$ 有 $\alpha^2=(1)$，所以 $\alpha^{-1}=\alpha\in V$. 另外，

$[(1\quad 2)(3\quad 4)][(1\quad 3)(2\quad 4)]=[(1\quad 3)(2\quad 4)][(1\quad 2)(3\quad 4)]=(1\quad 4)(2\quad 3)$，

$[(1\quad 2)(3\quad 4)][(1\quad 4)(2\quad 3)]=[(1\quad 4)(2\quad 3)][(1\quad 2)(3\quad 4)]=(1\quad 3)(2\quad 4)$，

$[(1\quad 3)(2\quad 4)][(1\quad 4)(2\quad 3)]=[(1\quad 4)(2\quad 3)][(1\quad 3)(2\quad 4)]=(1\quad 2)(3\quad 4)$，

即运算在 V 上封闭，所以 V 是一个 S_4 的子群，因此 V 是一个置换群. 注意，这里的 V 与例 4.1.11 中的群是同构的.

4.2　交换环和域

4.1 节讨论了群，它们都是某种定义了一个二元运算的集合组成的代数系统. 然而，整数集合 **Z**、有理数集合 **Q**、实数集合 **R**、复数集合 **C** 及整数集 **Z** 模 m 的剩余类集合，都是定义了两个二元运算(即“加法”和“乘法”)的代数系统. 下面来研究定义了两个二元运算的代数系统——交换环和域.

4.2.1　交换环及其基本性质

定义 4.2.1　设 R 是一个给定的集合，在其上定义了两种二元运算“＋”和“·”，如果满足以下条件：

(1) $(R, +)$是一个交换群；

(2) 二元运算“·”满足结合律；

(3) 二元运算“·”下存在单位元；

(4) 二元运算“·”满足交换律；

(5) 对于这两种运算有以下的分配律成立，即对任意 a，b，$c\in R$ 有，

$$a\cdot(b+c)=(a\cdot b)+(a\cdot c),$$

则将三元有序组$(R,+,\cdot)$称为**交换环**. 通常把运算“+”称为交换环中的“加法”，“·”称为交换环中的“乘法”. 运算“+”下的单位元称为交换环的**零元**，记为 0，且元素 a 在运算“+”下的逆元称为元素 a 的**负元**；运算“·”下的单位元称为交换环的**幺元**，记为 1，且如果元素 a 在运算“·”下存在其逆元，则称该逆元为元素 a 的**逆元**，满足这样条件的元素 a 称为**可逆元素**.

当然，如同群的讨论一样，也可以使用其他的符号来表示交换环上的两种二元运算，如“⊕”和“⊗”等等. 有时为了书写的简洁，在表示乘法的时候，也可以省略符号“·”. 另外，还规定乘法的运算优先级比加法高，例如，此时分配律可以写为

$$a\cdot(b+c)=a\cdot b+a\cdot c.$$

由于交换环的乘法也满足交换律，所以另一个分配律也必成立，即对任意 a，b，$c\in R$ 有，

$$(b+c)\cdot a=(b\cdot a)+(c\cdot a),$$

因为

$$(b+c)\cdot a=a\cdot(b+c)=(a\cdot b)+(a\cdot c)=(b\cdot a)+(c\cdot a).$$

[例 4.2.1] $(\mathbf{Z},+,\times)$是一个交换环，零元是 0，幺元是 1，可逆元素只有 -1 和 1. $(\mathbf{Q},+,\times)$、$(\mathbf{R},+,\times)$和$(\mathbf{C},+,\times)$都是交换环，零元都是 0，幺元都是 1，除了 0 以外，所有的其他元素都是可逆元素.

[例 4.2.2] 仅由一个元素 a 组成的交换环$(\{a\},\oplus,\otimes)$，称为**零环**. 由零元和幺元的定义易证，零环$(\{a\},\oplus,\otimes)$中的元素 a 就是它的零元和幺元. 该交换环中不存在可逆元素.

[例 4.2.3] 用 $\mathbf{Z}[i]$表示集合$\{a+bi\mid a,b\in\mathbf{Z}\}$，其中 i 为虚数单位，则 $\mathbf{Z}[i]$关于复数的加法和乘法构成交换环，称为高斯整数环. 零元是 0，幺元是 1，可逆元素只有 -1、1、i 和 $-i$.

[例 4.2.4] 令 $\mathbf{Q}(\sqrt{2})=\{a+b\sqrt{2}\mid a,b\in\mathbf{Q}\}$，则$(\mathbf{Q}(\sqrt{2}),+,\times)$是一个交换环，零元是 0，幺元是 1，除了 0 以外，所有的其他元素都是可逆元素.

[例 4.2.5] 令 $\mathbf{Z}(\sqrt{2})=\{a+b\sqrt{2}\mid a,b\in\mathbf{Z}\}$，则$(\mathbf{Z}(\sqrt{2}),+,\times)$是一个交换环，零元是 0，幺元是 1，都有哪些可逆元素呢？(留作习题)

[例 4.2.6] 当 $n\geqslant 2$ 时，$\mathbf{Z}$ 模 n 的剩余类的集合 $\mathbf{Z}_n=\{[0],[1],\cdots,[n-1]\}$，其中$[r]$也可以看作 $n\mathbf{Z}$ 的陪集，由商群的知识已知，当对任意$[a],[b]\in\mathbf{Z}_n$ 定义下面的加法时

$$[a]\oplus[b]=[a+b],$$

$(\mathbf{Z}_n,\oplus)$是群，且由于

$$[a]\oplus[b]=[a+b]=[b+a]=[b]\oplus[a],$$

可知$(\mathbf{Z}_n,\oplus)$是交换群.

现在，如果另外对任意$[a],[b]\in\mathbf{Z}_n$ 定义二元运算⊗为

$$[a]\otimes[b]=[a\times b],$$

则($\mathbf{Z}_n$, $\oplus$, $\otimes$)是一个交换环，称为**整数模 n 的剩余类环**，零元为$[0]$，幺元为$[1]$，$[a]$是可逆元素的充要条件是$(a, n)=1$. 所有非零元都可逆的充要条件是 n 是素数.

证明　(1) 由于($\mathbf{Z}_n$, $\oplus$)是交换群，零元为$[0]$，所以交换环定义中的条件(1)已经满足.

也很容易证明$\otimes$的定义不依赖于代表元的选择：对任意$[a]$，$[b]\in\mathbf{Z}_n$，如果$[a]=[c]$且$[b]=[d]$，那么必然有

$$[a]\otimes[b]=[a\times b]=[c\times d]=[c]\otimes[d].$$

(2) 设$[a]$，$[b]$，$[c]\in\mathbf{Z}_n$，则

$$([a]\otimes[b])\otimes[c]=[a\times b]\otimes[c]=[(a\times b)\times c]=[a\times b\times c],$$
$$[a]\otimes([b]\otimes[c])=[a]\otimes[b\times c]=[a\times(b\times c)]=[a\times b\times c],$$

所以$([a]\otimes[b])\otimes[c]=[a]\otimes([b]\otimes[c])$，因此$\otimes$满足结合律，即交换环定义中的条件(2)已经满足.

(3) 因 $n\geqslant 2$，故$[1]\in\mathbf{Z}_n$，对任意$[a]\in\mathbf{Z}_n$，有

$$[1]\otimes[a]=[1\times a]=[a],$$
$$[a]\otimes[1]=[a\times 1]=[a],$$

所以$[1]$是运算$\otimes$下的单位元，即交换环定义中的条件(3)已经满足.

(4) 设$[a]$，$[b]\in\mathbf{Z}_n$，则

$$[a]\otimes[b]=[a\times b],$$
$$[b]\otimes[a]=[b\times a],$$

所以$[a]\otimes[b]=[b]\otimes[a]$，因此$\otimes$满足交换律，即交换环定义中的条件(4)已经满足.

(5) 设$[a]$，$[b]$，$[c]\in\mathbf{Z}_n$，则

$$[a]\otimes([b]\oplus[c])=[a]\otimes[b+c]=[a\times(b+c)]=[a\times b+a\times c],$$
$$([a]\otimes[b])\oplus([a]\otimes[c])=[a\times b]\oplus[a\times c]=[a\times b+a\times c],$$

所以$[a]\otimes([b]\oplus[c])=([a]\otimes[b])\oplus([a]\otimes[c])$，因此两种运算满足分配律，即交换环定义中的条件(5)已经满足.

综上所述可知，($\mathbf{Z}_n$, $\oplus$, $\otimes$)是一个交换环.

(6) 如果$[a]\in\mathbf{Z}_n$ 是可逆元素，那么必然存在$[b]\in\mathbf{Z}_n$，使得下式成立，

$$[a]\otimes[b]=[1],$$
$$[a\times b]=[1],$$

所以

$$a\times b\equiv 1(\bmod n),$$

显然该式成立的充要条件是$(a, n)=1$.

因此所有非零元都可逆的充要条件是

$$(1, n)=1, (2, n)=1, \cdots, (n-1, n)=1,$$

即 n 是素数. 得证.

由于在交换环中有两个二元运算，为了不混淆起见，我们把元素 a 在加法群$(R, +)$中的逆元素及它的幂分别记为“$-a$”和“na”($n\in\mathbf{Z}$)，元素 a 在乘法下的幂仍记为“a^n”($n\in\mathbf{N}$)，当 a 是交换环的可逆元素时，它的逆元和幂可分别记为“a^{-1}”和“a^n”($n\in\mathbf{Z}$). 另外，可以将 $a+(-b)$简洁地记作 $a-b$，并且可以按习惯称其为“减法”. 在这样的约定下，通过以下几个定理

给出交换环的基本性质.

定理 4.2.1 $(R, +, \cdot)$为交换环，对任意 $a, b, c \in R$ 有，

(1) $0 \cdot a = a \cdot 0 = 0$. $a \cdot b$；

(2) $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$；

(3) $(-a) \cdot (-b) = a \cdot b$；

(4) $a \cdot (b-c) = a \cdot b - a \cdot c$；

(5) $(b-c) \cdot a = b \cdot a - c \cdot a$；

证明 (1)利用分配律可知 $0 \cdot a = (0+0) \cdot a = (0 \cdot a) + (0 \cdot a)$，由加法群$(R, +)$的消去律知

$$0 \cdot a = 0.$$

同理可证 $a \cdot 0 = 0$.

(2)至(5)的证明留作读者练习.

定理 4.2.2 $(R, +, \cdot)$为交换环，$a, b \in R$，$m, n \in \mathbf{Z}$，则

(1) $m(na) = (mn)a$；

(2) $ma + na = (m+n)a$；

(3) $(na) \cdot b = a \cdot (nb) = n(a \cdot b)$；

(4) $(ma) \cdot (nb) = (mn)(a \cdot b)$；

(5) $(ma^h) \cdot (na^k) = (mn)a^{h+k}$.(其中 $h, k \in \mathbf{N}$；当 a 是可逆元时，可取 $h, k \in \mathbf{Z}$.)

定理 4.2.3 一般交换环上的二项式定理.$(R, +, \cdot)$为交换环，$a, b \in R$，则

$$(a+b)^n = \sum_{k=0}^{n} \frac{n!}{k!\,(n-k)!} a^k \cdot b^{n-k}.$$

定理 4.2.1、定理 4.2.2 和定理 4.2.3 给出了交换环的一些性质，最重要的是，从这些基本性质上可以看到，在元素的运算规律上，一般的交换环与整数集合 $\mathbf{Z}$ 都是非常相像的. 这些定理的证明都很简单，留作练习.

定义 4.2.2 如果交换环 R 的一个子集 S 满足如下三个条件，

(1) $1 \in S$；

(2) 如果 $a, b \in S$，则 $a - b \in S$；

(3) 如果 $a, b \in S$，则 $ab \in S$；

则称 S 是 R 的**子环**.

与一个子群本身就是一个群类似，可以证明一个子环本身就是一个交换环.

定理 4.2.4 $(R, +, \cdot)$为交换环，S 是 R 的子环，则$(S, +, \cdot)$是一个交换环.

证明 因为 $1 \in S$，所以 $0 = 1 - 1 \in S$，由 $0 \in S$ 和子环定义中条件(2)，根据定理 4.1.10 可知$(S, +)$是$(R, +)$的子群，同时明显$(S, +)$继承了$(R, +)$的交换性，所以$(S, +)$是交换群，即交换环定义中的条件(1)已经满足.

因为 $1 \in S$，所以交换环定义中的条件(3)已经满足.

由子环定义中条件(3)可知，乘法运算在 S 上封闭，所以 S 继承了 R 在乘法下的结合律、交换律；又由于加法运算同时在 S 上封闭，所以 S 继承了 R 在这两种运算下的分配律，即交换环定义中的条件(2)、(4)和(5)已经满足.

综上所述，$(S, +, \cdot)$本身是一个交换环. 得证.

定义 4.2.3　设交换环$(S, +, \cdot)$是交换环$(R, +, \cdot)$的子环，则我们称 R 是 S 的**扩环**(或**扩张**). 如果 $S=R$ 或 $S=\{0\}$，那么显然 S 是 R 的子环，称为**平凡子环**，平凡子环以外的子环称为**真子环**.

[例 4.2.7]　$(\mathbf{Z}, +, \times)$、$(\mathbf{Q}, +, \times)$和$(\mathbf{R}, +, \times)$都是$(\mathbf{C}, +, \times)$的子环；$(\mathbf{Z}, +, \times)$和$(\mathbf{Q}, +, \times)$都是$(\mathbf{R}, +, \times)$的子环；$(\mathbf{Z}, +, \times)$是$(\mathbf{Q}, +, \times)$的子环.

[例 4.2.8]　证明 交换环$(\mathbf{Q}(\sqrt{2}), +, \times)$是交换环$(\mathbf{R}, +, \times)$的一个子环，其中 $\mathbf{Q}(\sqrt{2})=\{a+b\sqrt{2} \mid a, b\in\mathbf{Q}\}$.

证明　$1=0+0\sqrt{2}$，所以 $1\in\mathbf{Q}(\sqrt{2})$；

对任意 $x_1=a+b\sqrt{2}\in\mathbf{Q}(\sqrt{2})$，$x_2=c+d\sqrt{2}\in\mathbf{Q}(\sqrt{2})$，其中 $a, b, c, d\in\mathbf{Q}$，必有

$$x_1-x_2=(a-c)+(b-d)\sqrt{2}\in\mathbf{Q}(\sqrt{2}),$$

$$x_1\times x_2=(a\times c+2\times b\times c)+(a\times d+b\times c)\sqrt{2}\in\mathbf{Q}(\sqrt{2}),$$

且明显

$$\mathbf{Q}(\sqrt{2})\subset\mathbf{R},$$

由子环的定义可知，$(\mathbf{Q}(\sqrt{2}), +, \times)$是$(\mathbf{R}, +, \times)$的子环. 得证.

例 4.2.1 中的交换环 $\mathbf{Z}$ 和例 4.2.6 中的交换环 $\mathbf{Z}_n(n\geqslant 2)$有一点很大的不同. 这就是当 $a, b\in\mathbf{Z}$ 时，若 $a\neq 0$，$b\neq 0$ 则必有 $a\times b\neq 0$，即两个非零元之积一定是非零元. 但是这种性质在 $\mathbf{Z}_n$ 中不一定成立，当 n 为合数时，一定存在两个非零元之积等于零元的情况. 例如在 $\mathbf{Z}_{12}$ 中，

$$[2]\otimes[6]=[3]\otimes[4]=[12]=[0],$$

也就是说，交换环 $\mathbf{Z}$ 与交换环 $\mathbf{Z}_n(n\geqslant 2)$是两种本质上不同的交换环. 下面我们引入零因子的概念来说明这种现象.

定义 4.2.4　设$(R, +, \cdot)$是交换环，$a\in\mathbf{R}$ 且 $a\neq 0$，如果存在 $b\in\mathbf{R}$ 且 $b\neq 0$，使得 $a\cdot b=0$ 成立，则称 a 是交换环 R 的**零因子**.

[例 4.2.9]　$\mathbf{Z}$ 中没有零因子. 容易证明，$\mathbf{Z}_n$ 中没有零因子的充要条件是 n 为素数.

定理 4.2.5　设$(R, +, \cdot)$是交换环，$a\in R$ 且 $a\neq 0$，则 a 不是零因子的充要条件是对 a 有乘法消去律成立，即对任意 $b, c\in R$，如果 $a\cdot b=a\cdot c$，那么 $b=c$.

证明　充分性. 已知对任意 $b, c\in R$，若有 $a\cdot b=a\cdot c$，则必有 $b=c$. 当 $c=0$ 时，由定理 4.2.1 知 $a\cdot c=0$，故 $a\cdot b=0$，又因为 $b=c$，所以此时必有 $b=0$，于是得到结论：要使 $a\cdot b=0$ 成立，必有 $b=0$，所以 a 不是交换环 R 的零因子.

必要性. 已知 R 为交换环，$a\in R$ 不是零因子，b, c 是 R 中的任意元素且 $a\cdot b=a\cdot c$.

因为

$$a\cdot(c+(-b))=(a\cdot c)+(a\cdot(-b))=(a\cdot b)+(a\cdot(-b))$$
$$=a\cdot(b+(-b))=a\cdot 0=0,$$

即

$$a\cdot(c+(-b))=0.$$

又因为 a 不是零因子，所以必有 $c+(-b)=0$，即 $b=c$，对 a 有乘法消去律成立. 得证.

定义 4.2.5 交换环$(R, +, \cdot)$中至少有两个元素(其中一个是零元，另一个是幺元)，且R中没有零因子，则我们称这样的交换环为**整环**.

[例 4.2.10] $(\mathbf{Z}, +, \times)$、$(\mathbf{Q}, +, \times)$、$(\mathbf{R}, +, \times)$和$(\mathbf{C}, +, \times)$都是整环. $(\mathbf{Z}_n, \oplus, \otimes)$是整环的充要条件是$n$为素数. 所以，$\mathbf{Z}_2$、$\mathbf{Z}_3$和$\mathbf{Z}_5$等，都是整环；而$\mathbf{Z}_4$、$\mathbf{Z}_6$和$\mathbf{Z}_8$等，都不是整环.

我们可以将整数的整除概念推广到一般的交换环上.

定义 4.2.6 $(R, +, \cdot)$是交换环，$a, b \in R$，$b \neq 0$. 如果存在一个元素$c \in R$，使得$a = b \cdot c$，则称b**整除**a或者a**被**b**整除**，记为$b \mid a$. 把b称为a的**因子**，a称为b的**倍元**. 如果b，c都不可逆，则b称为a的**真因子**.

$b \mid a$是否成立，不仅依赖于a和b的值，而且还依赖于集合R. 比如，在$(\mathbf{Q}, +, \times)$中$6 \mid 5$，因为$5 = 6 \times \frac{5}{6}$，并且$\frac{5}{6} \in \mathbf{Q}$. 但是，在$(\mathbf{Z}, +, \times)$中$6 \nmid 5$，因为不存在$c \in \mathbf{Z}$，使得$5 = 6 \times c$.

定义 4.2.7 $(R, +, \cdot)$是交换环，$a, b \in R$，则R中任意形如$sa + tb$的元素称为a和b的一个**线性组合**，其中$s, t \in R$.

就像在整数集合里一样，很容易证明：交换环中任意两个元素的公因子必然整除这两个元素的任意线性组合.

定义 4.2.8 $(R, +, \cdot)$是交换环，$a, b \in R$，如果存在可逆元素$u \in R$，使得$a = bu$，则称a，b为**相伴元素**.

[例 4.2.11] 由于$(\mathbf{Z}, +, \times)$的可逆元素只有-1和1，所以对于任何$t \in \mathbf{Z}$，它的相伴元素只有$-t$和t. $(\mathbf{Q}, +, \times)$、$(\mathbf{R}, +, \times)$和$(\mathbf{C}, +, \times)$的任意两个非零元a，b都是相伴元素，因为$a = b\frac{a}{b}$，而且$\frac{a}{b}$是可逆元素.

定义 4.2.9 $(R, +, \cdot)$是交换环，$p \in R$且$p \neq 1$，如果p没有真因子，则称p为**不可约元素**，或**素元**，或**即约元素**.

由上面定义可知，一般的交换环中的不可约元素就相当于整数中的素数.

定义 4.2.10 $(R, +, \cdot)$是交换环，R的**特征**是使$n1 = 0$的最小正整数n，如果不存在这样的正整数，则我们称R的特征是0. R的特征记作$\mathrm{char}(R)$.

如果交换环$(R, +, \cdot)$的特征为n，则对任意$a \in R$，都有$na = 0$.

[例 4.2.12] $(\mathbf{Z}, +, \times)$、$(\mathbf{Q}, +, \times)$、$(\mathbf{R}, +, \times)$和$(\mathbf{C}, +, \times)$的特征都是0.

[例 4.2.13] $(\mathbf{Z}_n, \oplus, \otimes)$的特征是$n$.

在介绍了交换环的特征的概念后，引入一个在编码理论和密码学中都有重要价值的定理.

定理 4.2.6 $(R, +, \cdot)$是交换环，$\mathrm{char}(R) = p$为素数，则对任意$a, b \in R$，有

$$(a+b)^p = a^p + b^p.$$

证明 根据定理 4.2.3 有

$$(a+b)^p = a^p + \sum_{k=1}^{p-1} \frac{p!}{k!\,(p-k)!} a^k \cdot b^{p-k} + b^p,$$

其中 $a^k \cdot b^{p-k} \in R$. 对于整数 $\dfrac{p!}{k!(p-k)!}(1 \leqslant k \leqslant p-1)$，易知当 p 为素数时，有

$$p \left| \frac{p!}{k!(p-k)!} \quad (1 \leqslant k \leqslant p-1), \right.$$

于是

$$\frac{p!}{k!(p-k)!} a^k \cdot b^{p-k} = 0,$$

所以有 $(a+b)^p = a^p + b^p$ 成立.

4.2.2　域及其基本性质

从交换环的定义可以看出，交换环中要求“加法”下构成交换群，即每个元素都有其负元，然而全体非零元在“乘法”下不一定构成群，原因在于非零元不一定是可逆元. 例如，$(\mathbf{Z}, +, \times)$和$(\mathbf{Z}_n, \oplus, \otimes)$($n$ 是合数)就都是这种情形. 但有些交换环的非零元素皆为可逆元素，例如$(\mathbf{Q}, +, \times)$，$(\mathbf{R}, +, \times)$，$(\mathbf{C}, +, \times)$及$(\mathbf{Z}_p, \oplus, \otimes)$(其中 p 是素数). 后面这类交换环构成了一种新的代数系统——域.

定义 4.2.11　设$(F, +, \cdot)$是一个非零交换环，如果它的每个非零元都是可逆元素，那么它称为**域**.

我们注意，任意一个域首先是一个交换环，这时，$(F, +, \cdot)$作为交换环的特征 $\mathrm{char}(F)$也称为域$(F, +, \cdot)$的特征.

由域的定义知，在域$(F, +, \cdot)$中$(F, +)$构成交换群，$(F-\{0\}, \cdot)$也构成交换群. 因此交换环$(F, +, \cdot)$是一个域当且仅当$(F-\{0\}, \cdot)$是交换群.

[例 4.2.14]　$(\mathbf{Q}, +, \times)$，$(\mathbf{R}, +, \times)$，$(\mathbf{C}, +, \times)$及$(\mathbf{Z}_p, \oplus, \otimes)$(其中 p 是素数)都是域，也可以将$(\mathbf{Z}_p, \oplus, \otimes)$($p$ 是素数)记作$(\mathbf{F}_p, \oplus, \otimes)$.

定理 4.2.7　域是整环.

证明　设$(F, +, \cdot)$为域，由域的定义知 F 为交换环且 F 中一定存在幺元，只需证明 F 中不存在零因子就可以了. 假设 F 中存在零因子，即存在 $a, b \in F$，$a \neq 0$，$b \neq 0$ 满足

$$a \cdot b = 0.$$

由域的定义知，a，b 均为可逆元素，于是

$$a = a \cdot (b \cdot b^{-1}) = (a \cdot b) \cdot b^{-1} = 0 \cdot b^{-1} = 0,$$

可得 $a = 0$；

$$b = (a^{-1} \cdot a) \cdot b = a^{-1} \cdot (a \cdot b) = a^{-1} \cdot 0 = 0,$$

可得 $b = 0$.

这与 $a \neq 0$，$b \neq 0$ 的假设矛盾，故 F 中不存在零因子，即 F 是整环. 得证.

注意，这个定理的逆命题不一定成立.

[例 4.2.15]　$(\mathbf{Z}, +, \times)$是整环，但不是域.

定理 4.2.8　有限整环一定是域.

证明　设$(F, +, \cdot)$为有限整环，其阶为 $n(n \in \mathbf{N})$，则可设

$$F = \{a_1, a_2, \cdots, a_n\},$$

其中 $a_i \neq a_j (i \neq j)$. 对任意 $a \in F$ 且 $a \neq 0$，构造集合

$$F'=a\cdot F=\{a\cdot a_1,\ a\cdot a_2,\ \cdots,\ a\cdot a_n\}.$$

由交换环中乘法的封闭性，F'中的每个元素$a\cdot a_i(1\leqslant i\leqslant n)$都在$F$中，所以$F'\subseteq F$. 又由整环满足消去律，对任意$a\cdot a_i$，$a\cdot a_j\in F'(1\leqslant i,\ j\leqslant n$且$i\neq j)$都有$a\cdot a_j\neq a\cdot a_j$，否则与$a_i\neq a_j(i\neq j)$矛盾，所以$F'$和$F$等势，所以有$F=F'$. 所以一定存在$a\cdot a_i\in F'$，使得$a\cdot a_i=1$，即有限整环$F$中任意非零元素$a$一定存在逆元. 证毕.

定理 4.2.9 设n为域$(F,+,\cdot)$的特征，若n不为0，则n必为素数.

证明 用反证法，设n不是素数，则必存在整数$1<a,b<n$，使得$n=ab$. 则有

$$(a1)\cdot(b1)=(ab)1=n1=0.$$

因为域是整环，域中无零因子，所以$a1=0$或$b1=0$，这与特征n的最小性矛盾(见定义 4.2.10)，证毕.

定义 4.2.12 设$(F,+,\cdot)$是域，$(H,+,\cdot)$是$(F,+,\cdot)$的子环，且$(H,+,\cdot)$本身也是域，则$(H,+,\cdot)$是$(F,+,\cdot)$的**子域**；$(F,+,\cdot)$是$(H,+,\cdot)$的**扩域**(或**扩张**). 若$H=F$，$(H,+,\cdot)$显然是$(F,+,\cdot)$的子域，这种子域称为**平凡子域**，非平凡子域称为**真子域**. 需要注意的是，零环$(H,+,\cdot)=(\{0\},+,\cdot)$不是域，故$(\{0\},+,\cdot)$不是$(F,+,\cdot)$的子域，这一点与子环的性质不同.

4.2.3 同态与同构

为了讨论的方便，这里先给出交换环和域的同态、同构的基本概念，它们被用来比较两个交换环(或域)之间的相似程度.

定义 4.2.13 设$(X,+,\cdot)$和$(Y,\oplus,\otimes)$是两个交换环(或域)，如果存在一个从集合X到集合Y的映射f，使得对任意$a,b\in X$，有

$$f(1_X)=1_Y,$$
$$f(a+b)=f(a)\oplus f(b),$$
$$f(a\cdot b)=f(a)\odot f(b),$$

成立，则称$(X,+,\cdot)$与$(Y,\oplus,\odot)$**同态**，记作

$$(X,+,\cdot)\backsim(Y,\oplus,\odot).$$

f称为从$(X,+,\cdot)$到$(Y,\oplus,\odot)$的**同态映射**. 如果f是单射，则称此同态为**单同态**；如果f是满射，则称此同态为**满同态**；如果f是双射，则称此同态为**同构**. 记作

$$(X,+,\cdot)\cong(Y,\oplus,\odot).$$

两个同构的交换环(或域)没有本质差别，基本上可以看作是同一个事物，下面用整环的分式域来说明同构的概念.

实际上，能够很容易证明整环的子环本身必然是整环，因此域的子环本身必然是整环. 那么，反过来会怎么样呢？在初等数学中，为了在整数环$\mathbf{Z}$中使任意两个非零整数都能进行乘法的逆运算——除法，我们引入了分数，得到了有理数集合这样一个域. 这一方法可推广到任意整环中去.

定义 4.2.14 设$(M,+,\times)$是一个整环，令$M^*=M-\{0\}$，$E=M\times M^*$，在E上定义关系R，对任意$(a,b)\in E$，$(c,d)\in E$，$(a,b)R(c,d)$当且仅当

$$a\times d=b\times c.$$

实际上，R 是 E 上的等价关系. 验证如下：

(1) 自反性：因为 $a\times b=b\times a$，所以$(a, b)R(a, b)$.

(2) 对称性：如果$(a, b)R(c, d)$，则 $a\times d=b\times c$，即 $c\times b=a\times d$，所以$(c, d)R(a, b)$.

(3) 传递性：如果$(a, b)R(c, d)$且$(c, d)R(e, f)$，则 $a\times d=b\times c$ 且 $c\times f=d\times e$，

$$a\times d\times f=b\times c\times f \text{ 且 } b\times c\times f=b\times d\times e,$$

$$a\times d\times f=b\times d\times e,$$

因为$(M, +, \times)$是一个整环且 $d\neq 0$，则利用消去律可得 $a\times f=b\times e$，即$(a, b)R(e, f)$.

集合上的等价关系能够将集合划分为等价类，记 E 上(a, b)关于 R 的等价类为

$$\frac{a}{b}=[(a, b)]=\{(e, f)\mid (e, f)\in E \text{ 且 } (a, b)R(e, f)\}.$$

所有这样的等价类构成商集 E/R，我们定义 E/R 上的加法“$\oplus$”和乘法“$\otimes$”如下，对任意 $(a, b)\in E$，$(c, d)\in E$，有

$$\frac{a}{b}\oplus\frac{c}{d}=\frac{a\times d+b\times c}{b\times d},$$

$$\frac{a}{b}\otimes\frac{c}{d}=\frac{a\times c}{b\times d}.$$

因为 b，$d\in M^*$，所以 $b\neq 0$，$d\neq 0$，由 M 中无零因子，$b\times d\neq 0\in M^*$，所以这样定义的运算“$\oplus$”和“$\otimes$”满足封闭性. 此外，需要验证如上规定的运算结果与代表元无关：

如果$(a, b)=(a', b')$且$(c, d)=(c', d')$，那么 $ab'=a'b$ 且 $cd'=c'd$，则

$$(a\times d+b\times c)b'\times d'=ab'dd'+bcb'd'=a'bddd'+bc'dbb'=(a'\times d'+b'\times c')b\times d,$$

$$\frac{a}{b}\oplus\frac{c}{d}=\frac{a'}{b'}\oplus\frac{c'}{d'},$$

同理有

$$\frac{a}{b}\otimes\frac{c}{d}=\frac{a'}{b'}\otimes\frac{c'}{d'}.$$

容易验证，$(E/R, \oplus)$ 构成一个交换群，零元为 $\frac{0}{b}$，$\frac{a}{b}$ 的负元为 $\frac{-a}{b}$；$\left((E/R)-\left\{\frac{0}{b}\right\}, \otimes\right)$构成一个交换群，幺元为$\frac{b}{b}$，$\frac{a}{b}$的逆元为$\frac{b}{a}$. 因此，$(E/R, \oplus, \otimes)$构成一个域.

定义 4.2.15　上述由整环$(M, +, \times)$构造而来的域$(E/R, \oplus, \otimes)$称为整环 M 的**分式域**. 记作 **Frac(*M*)**.

定理 4.2.10　整环必和其分式域的一个子环同构.

证明　设$(M, +, \times)$是一个整环，$(Y, \oplus, \otimes)=\mathrm{Frac}(M)$. 考虑 Y 的子集

$$S=\left\{\frac{a}{1}\mid a\in X\right\},$$

容易验证$(S, \oplus, \otimes)$是$(Y, \oplus, \otimes)$的子环. 建立从 M 到 S 的映射 $f: M\to S$，

$$f(a)=\frac{a}{1},$$

显然这是一个一一对应的映射，且满足对任意 a，$b\in M$，有

$$f(a+b)=\frac{a+b}{1}=\frac{a\times 1+1\times b}{1\times 1}=\frac{a}{1}\oplus\frac{b}{1}=f(a)\oplus f(b),$$

$$f(a\times b)=\frac{a\times b}{1}=\frac{a\times b}{1\times 1}=\frac{a}{1}\otimes\frac{b}{1}=f(a)\otimes f(b),$$

所以

$$(S,\oplus,\otimes)\trianglelefteq(M,+,\times).$$

定理得证.

通常我们仍然用符号“+”和“×”表示整环$(M,+,\times)$的分式域中的“加法”和“乘法”，而不会引起混淆，正如初等数学中对整数运算和分数运算使用相同的加号和乘号一样. 这样就可以将整环$(M,+,\times)$看成它的分式域$(\mathrm{Frac}(M),+,\times)$的子环，将分式域$\mathrm{Frac}(M)$看成整环 X 的扩张或扩域.

[例 4.2.16] $(\mathbf{Z},+,\times)$的分式域是$(\mathbf{Q},+,\times)$，即 $\mathrm{Frac}(\mathbf{Z})=\mathbf{Q}$.

[例 4.2.17] 后面要讲到的域上的一元多项式集合是一个整环，所以它的分式域存在：设 K 是一个域，则 $K[x]$是域 K 上的一元多项式集合，为整环，它的分式域记为 $K(X)$，即

$$K(X)=\left\{\frac{f(x)}{g(x)}\mid f(x),\ g(x)\in K[X],\ g(x)\neq 0\right\}.$$

4.2.4 一元多项式环

在密码学和编码理论中经常用到一种称为多项式环的环，下面给出一元多项式环的定义.

定义 4.2.16 设$(R,+,\cdot)$是交换环，x 是一个变元，n 是非负整数，$a_0,a_1,\cdots,a_n\in R$，则

$$f(x)=a_0+a_1x+\cdots+a_nx^n,$$

称为**交换环 R 上的一元多项式**. 其中，$a_0,a_1,\cdots,a_n$ 称为该多项式的**系数**，a_0 还称为**常数项**. 如果一个多项式的所有系数都是 0，那么该多项式称为**零多项式**. 如果 $a_n\neq 0$，那么 a_n 称为**首项系数**，n 称为一元多项式 $f(x)$的**次数**，记作

$$\deg f(x)=n.$$

将 $a_n=1$ 的多项式称为**首一多项式**. 所有交换环 R 上的一元多项式组成的集合记为 $\boldsymbol{R}[\boldsymbol{x}]$.

注意，零多项式没有次数，因为零多项式就没有非零的系数.

需要注意的是，在这个定义中，符号“+”并不是 R 中的加法运算，a_nx^n 也不是 R 中的乘法运算，仅仅是一种符号.

设$(R,+,\cdot)$是交换环，定义在 $R[x]$上的二元运算加法“+”和乘法“×”如下：对任意两个一元多项式，

$$f(x)=a_0+a_1x+\cdots+a_nx^n\in R[x],$$
$$g(x)=b_0+b_1x+\cdots+b_mx^m\in R[x],$$

令

$$(f+g)(x)=(a_0+b_0)+(a_1+b_1)x+\cdots+(a_l+b_l)x^l,$$

其中 $l=\max\{m,n\}$. 当 $n<l$ 时，$a_j=0(n<j\leqslant l)$，当 $m<l$ 时，$b_j=0(m<j\leqslant l)$.

$$(f\times g)(x)=c_0+c_1x+\cdots+c_{n+m}x^{n+m},$$

其中

$$c_k=\sum_{i+j=k}a_ib_j\quad(0\leqslant i\leqslant n,\ 0\leqslant j\leqslant m,\ 0\leqslant k\leqslant n+m).$$

定理 4.2.11　$(R, +, \cdot)$是交换环，$f(x)$和 $g(x)$是 $R[x]$中两个非零多项式，则

(1) $f\times g=$零多项式或者 $\deg f\times g\leqslant \deg f+\deg g$.

(2) 如果$(R, +, \cdot)$是整环，那么 $f\times g\neq$零多项式且 $\deg f\times g=\deg f+\deg g$.

该定理的证明很容易，留作读者自行练习.

容易验证，当$(R, +, \cdot)$是交换环时，$(R[x], +, \times)$也构成一个交换环，其零元是零多项式，幺元为 $f(x)=1$，$f(x)=a_0+a_1x+\cdots+a_nx^n$ 的负元为 $f(x)=(-a_0)+(-a_1)x+\cdots+(-a_n)x^n$. 进一步，由定理 4.2.11(2)可知，当$(R, +, \cdot)$是整环时，$(R[x], +, \times)$也是整环. 另外，注意如下的 $R[x]$的子集(只含有常数项的多项式的集合，该集合里的元素称为**常多项式**)

$$S=\{f(x)\mid f(x)=r, r\in R\},$$

很明显$(S, +, \times)$是$(R[x], +, \times)$的子环. 在 R 和 S 之间建立如下的双射，

$$r\to f(x)=r,$$

也很明显，该双射是一个同构映射，因此 $R\cong S$. 因此，可以将 R 看作是$(R[x], +, \times)$的子环. 综上所述，有如下的定义.

定义 4.2.17　$(R, +, \cdot)$是交换环，则称$(R[x], +, \times)$为 R **上的一元多项式环**.

4.2.5　理想和商环

理想在交换环中的重要性如同正规子群在群中的作用一样. 我们在这里介绍理想和商环的概念，主要是为了后面导入有限域，而有限域在现代密码学中应用广泛. 先从一个例子谈起.

在例 4.2.6 中，由于加法群$(\mathbf{Z}, +)$的子群$(n\mathbf{Z}, +)$是正规子群，所以能够自然地得到商群$(\mathbf{Z}_n, \oplus)$. 在此基础上，又在 $\mathbf{Z}_n$ 上定义了运算$\otimes$，这样得到了交换环$(\mathbf{Z}_n, \oplus, \otimes)$，即整数模 n 的剩余类环. 注意，交换环$(\mathbf{Z}_n, \oplus, \otimes)$的运算性质，如交换律、结合律，实际上是从交换环$(\mathbf{Z}, +, \times)$“继承”来的. 因此，可以将交换环$(\mathbf{Z}_n, \oplus, \otimes)$看作是从交换环$(\mathbf{Z}, +, \times)$导出的. 这一方法能否推广到一般的交换环上去呢？下面具体讨论这一问题.

设$(R, +, \cdot)$是交换环，且$(H, +)$是其加法群$(R, +)$的子群，由陪集和商群理论知，由$(R, +)$关于其子群$(H, +)$的所有陪集组成的集合 $T=R/H$，对加法运算“$+$”导出的运算“$\oplus$”也构成一个群$(T, \oplus)$(此群称为商群)，在这里，运算“$\oplus$”定义为，对任意 $a+H$，$b+H\in T$，

$$(a+H)\oplus(b+H)=(a+b)+H.$$

如果在集合 T 中能相应地定义由乘法运算“$\cdot$”导出的运算“$\odot$”，那么就可得到一个新的交换环$(T, \oplus, \odot)$. 因此问题在于加法子群$(H, +)$要满足什么样的条件才能定义这样的乘法运算. 设 $a+H$，$b+H$ 是子群$(H, +)$的陪集，由定义知 $a+H$ 表示集合$\{a+h\mid h\in H\}$，$b+H$ 表示集合$(b+h'\mid h'\in H)$. 显然，要定义满足要求的乘法运算“$\odot$”，必须使陪集 $a+H$ 中的任何元素 $a+h$ 与陪集 $b+H$ 中的任何元素 $b+h'$作乘法运算“$\otimes$”后都属于同一陪集，把这个陪集定义为 $a+h$ 与 $b+H$ 作运算“$\odot$”的结果. 由交换环$(R, +, \cdot)$的运算满足分配律知，对任意 h，$h'\in H$，有

$$(a+h)\cdot(b+h')=(a\cdot b)+(a\cdot h'+b\cdot h+h\cdot h').$$

由 $a, b\in R$ 及 $h, h'\in H$ 的任意性可知，当 H 满足条件："对任意 $r\in R$ 及 $h\in H$ 一定有 $r\cdot h\in H$"时，有

$$a\cdot h'\in H,$$
$$b\cdot h\in H,$$
$$h\cdot h'\in H.$$

由加法群$(H, +)$的封闭性，有

$$a\cdot h'+b\cdot h+h\cdot h'\in H,$$

又因为 $a\cdot b\in R$，所以

$$(a\cdot b)+(a\cdot h'+b\cdot h+h\cdot h')\in R/H,$$

即计算结果属于陪集 $(a\cdot b)+H$.

这时可以这样定义运算"$\odot$"：

$$(a+H)\odot(b+H)=(a\cdot b)+H.$$

不难验证，加法子群$(H, +)$满足条件"对任意 $r\in R$ 及 $h\in H$ 一定有 $r\cdot h\in H$"时，如上定义的$(T, \oplus, \odot)$构成一个交换环. 符合这种特殊条件的加法子群$(H, +)$就是理想，我们有下面定义.

定义 4.2.18 $(R, +, \cdot)$是一个交换环，I 是 R 的子集，使得

(1) $0\in I$；

(2) 对任意的 $a, b\in I$，都有 $a+b\in I$；

(3) 对任意的 $a\in I$ 和 $r\in R$，都有 $ra\in I$；

则将 I 称为 R 的**理想**. 显然 R 的两个平凡子环$\{0\}$和 R 都是 R 的理想，称为 R 的**平凡理想**，非平凡理想称为**真理想**.

[例 4.2.18] $n\mathbf{Z}$ 是交换环$(\mathbf{Z}, +, \times)$的一个理想.

证明 $0=n\times 0\in n\mathbf{Z}$；对任意 $a, b\in n\mathbf{Z}$，存在整数 a'和 b'，使得 $a=n\,a'$和 $b=n\,b'$，则

$$a+b=n\,a'+n\,b'=n(a'+b')\in n\mathbf{Z};$$

对任意 $a\in n\mathbf{Z}$ 和任意 $r\in\mathbf{Z}$，存在整数 a'使得 $a=n\,a'$，则

$$ra=rna'=n(ra')\in n\mathbf{Z};$$

由理想的定义可知，$n\mathbf{Z}$ 是一个理想.

由两个交换环的同态映射 f 的定义可知，该同态映射 f 必然也是这两个交换环的加法群之间的同态映射，因此存在 $\ker f$ 和 $\mathrm{im}\ f$，我们能够证明 $\ker f$ 是一个理想.

定理 4.2.12 设 f 是交换环 S 到交换环 G 的同态映射，则 $\mathrm{im}\ f$ 是 G 的子环，$\ker f$ 是 S 的理想.

证明 由同态定义，$1=f(1)\in\mathrm{im}\ f$；对任意 $a, b\in\mathrm{im}\ f$，存在 $a', b'\in S$，使得 $a=f(a')$和 $b=f(b')$，则

$$a-b=f(a')-f(b')=f(a'-b')\in\mathrm{im}\ f,$$
$$ab=f(a')f(b')=f(a'b')\in\mathrm{im}\ f,$$

因此由子环定义知 $\mathrm{im}\ f$ 是 G 的子环.

由 $\ker f$ 是 S 的加法群的子群可知，$0\in\ker f$；对任意 $a, b\in\ker f$，必有 $a+b\in\ker f$.

对任意 $a\in\ker f$ 和任意 $r\in S$，必有 $f(ra)=f(r)f(a)=f(r)0=0$，即 $ra\in\ker f$. 因此，由理想定义可知，$\ker f$ 是 S 的理想.

定理 4.2.13　交换环 R 的任意一族理想的交是 R 的理想.

证明　参考定理 4.1.17 的证明方法，留作读者自行练习.

定义 4.2.19　$(R,+,\cdot)$是一个交换环，H 是 R 的非空子集，$(H_i\mid i\in\mathbf{N})$是 R 的所有包含集合 H 的理想，即 $H\subseteq H_i(i\in N)$，则 $\bigcap_{i\in N}H_i$ 叫作**由子集 H 生成的理想**，记为(H)，H 中的元素叫作理想(H)的**生成元**. 如果 $H=\{a_1,a_2,\cdots,a_n\}(n\in\mathbf{N})$，则理想$(H)$记为$(a_1,a_2,\cdots,a_n)$，并称为**有限生成的理想**，由一个元素生成的理想(a)叫作**主理想**.

定理 4.2.14　$(R,+,\cdot)$是一个交换环，$a\in R$，$H=\{a_1,a_2,\cdots,a_n\}\subset R$，则

(1) $(a)=\{x\cdot a\mid x\in R\}$，

(2) $(H)=(a_1,a_2,\cdots,a_n)=\{x_1\cdot a_1\oplus x_2\cdot a_2\oplus\cdots\oplus x_n\cdot a_n\mid x_i\in R,\ 1\leqslant i\leqslant n\}$.

证明　很明显，$a_1,a_2,\cdots,a_n$ 的所有线性组合组成的集合必然是一个理想，而且如果 $a_1,a_2,\cdots,a_n\in R$ 的某个理想，则该理想必然包含 $a_1,a_2,\cdots,a_n$ 的所有线性组合组成的集合，所以，根据定义 4.2.19 可知命题成立.

[例 4.2.19]　$(R,+,\cdot)$是任意一个交换环，则 R 必然是自身的主理想，因为 $R=(1)$.

[例 4.2.20]　$(R,+,\cdot)$是任意一个交换环，则零环$\{0\}$必然是 R 的主理想，因为$\{0\}=(0)$.

[例 4.2.21]　$n\mathbf{Z}$ 是交换环$(\mathbf{Z},+,\times)$的主理想，因为 $n\mathbf{Z}=\{k\times n\mid k\in\mathbf{Z}\}=(n)$. 典型地，偶数集合是主理想(2).

定义 4.2.20　如果交换环$(R,+,\cdot)$的所有理想都是主理想，则交换环 R 称为**主理想环**.

[例 4.2.22]　求证$(\mathbf{Z},+,\times)$是主理想环.

证明　设 H 是 $\mathbf{Z}$ 的非零理想，则至少存在一个非零整数 $a\in H$，由理想的性质，因为 $-1\in\mathbf{Z}$，所以有

$$-a=(-1)\times a\in H,$$

于是 H 中有正整数存在，设 d 为 H 中的最小正整数，则 $H=(d)=\{n\times d\mid n\in\mathbf{Z}\}$. 这是因为，对任意 $a\in H$，由欧几里得除法定理，一定存在整数 q，r 使得

$$a=q\times d+r,\ 0\leqslant r<d,$$

这样，由 $a\in H$ 及 $q\times d\in H$，有 $r=a-q\times d\in H$. 但由于 $0\leqslant r<d$，又 d 是 H 中的最小正整数，所以

$$r=0,$$

$$a=q\times d\in(d),$$

从而 $H\subseteq(d)$，又显然$(d)\subseteq H$，所以 $H=(d)$. 即 $\mathbf{Z}$ 的任意理想 H 都可以写成(d)的形式. 所以 $\mathbf{Z}$ 是主理想环.

定理 4.2.15　交换环 R 的子集 H 是 R 的理想的充要条件是：

(1) $0\in H$；

(2) 对任意的 $a,b\in H$，都有 $a-b\in H$；

(3) 对任意的 $r\in R$ 和 $h\in H$，都有 $rh\in H$.

证明 与理想的定义中 3 个条件对比可知，差别只在第(2)条，因此只需证明(2)的充分性与必要性即可.

必要性. 因为 $1\in R$ 且 R 是加法群，所以 $-1\in R$，因此，对任意的 a，$b\in H$，$(-1)b\in H$，所以 $a+(-1)b\in H$，即 $a-b\in H$.

充分性. 因为 $1\in R$ 且 R 是加法群，所以 $-1\in R$，因此，对任意的 a，$b\in H$，$(-1)b\in H$，所以 $a-(-1)b\in H$，即 $a+b\in H$. 得证.

这个定理的(1)和(2)实际上就是(加法)子群的判定定理，因此这个定理告诉我们，交换环 R 的理想必然是 R 的加法子群.

定理 4.2.16 设$(R,+,\cdot)$为交换环，H 是它的理想，再设 T 是加法群$(R,+)$关于其子群$(H,+)$的所有不同陪集组成的集合，即商群 $T=R/H=\{a+H\mid a\in R\}$，那么$(T,\oplus,\odot)$构成交换环. 其中运算“$\oplus$”，“$\odot$”的定义为：对任意 $a+H\in T$，$b+H\in T(a,b\in R)$，有

$$(a+H)\oplus(b+H)=(a+b)+H,$$
$$(a+H)\odot(b+H)=(a\cdot b)+H.$$

证明 由于$(H,+)$是交换群$(R,+)$的子群，所以也是交换子群，当然是正规子群. 由商群的理论知$(T,\oplus)$构成商群，所以本定理中关于加法的结论必然成立.

现在只需证明二元运算$\odot$满足结合律、交换律和存在幺元，以及两种运算满足分配律即可. 首先要证明运算“$\odot$”的定义不依赖于 T 中元素的代表元的选择. 即要证明：对任意 $a+H=a'+H$，$b+H=b'+H$，都有

$$(a\cdot b)+H=(a'\cdot b')+H.$$

由陪集的性质可知

$$a-a'=h_1,\ b-b'=h_2,\ \text{其中}\ h_1,h_2\in H,$$

从而

$$\begin{aligned}(a\cdot b)+H&=[(a'+h_1)\cdot(b'+h_2)]+H\\&=[(a'\cdot b')+(a'\cdot h_2)+(h_1\cdot b')+(h_1\cdot h_2)]+H\\&=(a'\cdot b')+[(a'\cdot h_2)+(h_1\cdot b')+(h_1\cdot h_2)]+H.\end{aligned}$$

又因为 H 是R 的理想，所以$(a'\cdot h_2)$，$(h_1\cdot b')$，$(h_1\cdot h_2)\in H$，因此

$$(a'\cdot h_2)+(h_1\cdot b')+(h_1\cdot h_2)\in H$$

于是

$$[(a'\cdot h_2)+(h_1\cdot b')+(h_1\cdot h_2)]+H=H,$$
$$(a\cdot b)+H=(a'\cdot b')+H.$$

由运算“$\odot$”的定义，显然$(a+H)\odot(b+H)=(a\cdot b)+H\in T$，即运算“$\odot$”对 T 满足封闭性，对任意 $a+H$，$b+H$，$c+H\in T(a,b,c\in R)$，则

$$[(a+H)\odot(b+H)]\odot(c+H)=[(a\cdot b)+H]\odot(c+H)=(a\cdot b\cdot c)+H,$$
$$(a+H)\odot[(b+H)\odot(c+H)]=(a+H)\odot[(b\cdot c)+H]=(a\cdot b\cdot c)+H,$$

所以运算“$\odot$”满足结合律.

$$(a+H)\odot(b+H)=(a\cdot b)+H=(b\cdot a)+H=(b+H)\odot(a+H),$$

所以运算“⊙”满足交换律.

$$(a+H)\odot(1+H)=(a\cdot 1)+H=a+H,$$
$$(1+H)\odot(a+H)=(1\cdot a)+H=a+H,$$

所以运算“⊙”的幺元是 $1+H$.

$$[(a+H)+(b+H)]\odot(c+H)=[(a+b)+H]\odot(c+H)=[(a+b)\cdot c]+H$$
$$=[(a\cdot c)+(b\cdot c)]+H,$$
$$(a+H)\odot(c+H)\odot(b+H)\odot(c+H)=[(a\cdot c)+H]\odot[(b\cdot c)+H]$$
$$=[(a\cdot c)+(b\cdot c)]+H,$$

所以两种运算满足分配律.

所以，综上所述，$(T,\oplus,\odot)$构成交换环. 得证.

定义 4.2.21　定理 4.2.16 中的交换环$(T,\oplus,\odot)=(R/H,\oplus,\odot)$称为 R 关于理想 H 的**商环**.

[例 4.2.23]　当 $n\geqslant 2$ 时，$\mathbf{Z}_n$ 为 $\mathbf{Z}$ 关于理想 $n\mathbf{Z}$ 的**商环**. 这是信息安全研究中最重要的交换环.

[例 4.2.24]　考虑定义在整数交换环 $\mathbf{Z}$ 上的一元多项式环$(\mathbf{Z}[x],+,\times)$，求 $\mathbf{Z}[x]/(x)$.

解　我们在多项式中省略运算符“×”. 则由多项式 x 生成的理想

$$(x)=\{xf(x)\mid f(x)\in\mathbf{Z}[x]\},$$

易知(x)是所有常数项为 0 的一元多项式.

对任意一元多项式 $z(x)\in\mathbf{Z}[x]$，设 $z(x)$的**常数项**为 $a(a\in\mathbf{Z})$，则集合(陪集)

$$[z(x)]=z(x)+(x)=\{z(x)+xf(x)\mid xf(x)\in(x)\}$$

是商环 $\mathbf{Z}[x]/(x)$中的一个元素，显然陪集$[z(x)]$由一系列 $\mathbf{Z}[x]$中的一元多项式组成，$[z(x)]$中各个多项式的共同特点是它们的常数项都是 a，即对任意多项式 $p(x)\in[z(x)]$，都有

$$p(x)-a\in(x),$$

所以

$$[z(x)]=[a]=a+(x).$$

即陪集$[z(x)]$和陪集$[a]$中的元素一样都是一元多项式且该一元多项式与整数 a 的差是理想(x)中的元素(常数项为 0 的一元多项式). 于是有商环

$$\mathbf{Z}[x]/(x)=\{[a]\mid a\in\mathbf{Z}\}.$$

[例 4.2.25]　考虑定义在整数交换环 $\mathbf{Z}$ 上的一元多项式环$(\mathbf{Z}[x],+,\times)$，求 $\mathbf{Z}[x]/(m)(2\leqslant m\in\mathbf{N})$.

解　由 m 生成的理想为

$$(m)=\{mf(x)\mid f(x)\in\mathbf{Z}[x]\},$$

它的元素是系数为 m 倍数的多项式. 因此，当求得两个一元多项式 $p_1(x)$和 $p_2(x)$的差 $p(x)=p_1(x)-p_2(x)$ 后，如果 $p(x)$的系数为 m 的倍数，那么下面的两个陪集相等

$$[p_1(x)]=[p_2(x)],$$

所以

$$\mathbf{Z}[x]/(m)=\{[c_nx^n+\cdots+c_0]\mid 0\leqslant c_i<m,\ 1\leqslant i\leqslant n,\ n\in\mathbf{N}\},$$

它同构于整数模 m 的剩余类环 $\mathbf{Z}_m$ 上的一元多项式环，即 $\mathbf{Z}[x]/((m))\cong\mathbf{Z}_m[x]$(读者自证).

下面这两个定理是群论中相应定理的延伸.

定理 4.2.17 如果 H 是交换环$(R,+,\cdot)$的理想，则如下定义的映射 $f:R\to R/H$，

$$f(a)=a+H$$

是核为 H 的同态映射(该同态称为**自然同态**).

证明 定理 4.2.16 已经证明 R/H 是一个交换环，下面只需验证同态的三个条件.

$f(1)=1+H$，所以同态定义中条件(1)满足；

$f(a)\oplus f(b)=(a+H)\oplus(b+H)=(a+b)+H=f(a+b)$，所以同态定义中条件(2)满足；

$f(a)\odot f(b)=(a+H)\odot(b+H)=(a\cdot b)+H=f(a\cdot b)$，所以同态定义中条件(3)满足；

因此，f 是同态映射.

由于 $\ker f=\{a\mid f(a)=0+H=H,\ a\in R\}$，所以对任意 $a\in H$，则 $f(a)=a+H=H$，得到$a\in\ker f$，从而 $H\subseteq\ker f$；反过来，对任意 $a\in\ker f$，则 $a+H=f(a)=H$，得到 $a\in H$，从而 $\ker f\subseteq H$，所以 $\ker f=H$.

定理 4.2.18 设 $f:S\to G$ 是交换环 S 到交换环 G 的同态映射，则存在 $S/\ker f$ 到 $\mathrm{im}\ f$ 的映射 $h:S/\ker f\to\mathrm{im}\ f$，使得 $S/\ker f\cong\mathrm{im}\ f$.

本定理的证明类似于定理 4.1.37，不再赘述，只要令 $h(a+\ker f)=f(a)$，读者即可自行证明.

4.3 域上的一元多项式环

4.2 节已经提到多项式环在密码学和编码理论中有重要应用，而且给出了一元多项式环的初步概念，这一节将深入研究一元多项式环. 从这一节开始不再讨论一般交换环上的一元多项式环，而是专门讨论域上的一元多项式环，因为后者比前者具有更多的应用. 我们用$(K[x],+,\cdot)$来表示定义在域$(K,+,\cdot)$上的多项式环，不再有意区分多项式之间的运算符号和域中元素之间的运算符号. 为了行文简洁，我们在不引起歧义的情况下省略多项式的乘法运算符.

除了多项式的两个运算以外，还可以从线性空间的角度来理解域上的一元多项式环. 设多项式环 $K[x]$中的多项式 $f(x)=a_0+a_1x+a_2x^2+\cdots+a_nx^n$，$c\in K$，定义 c 与 $f(x)$的**数乘**运算为：

$$cf(x)=ca_0+ca_1x+ca_2x^2+\cdots+ca_nx^n,$$

则 $cf(x)\in K[x]$.

有了域 K 上多项式的加法和数乘的定义，我们可以看出，多项式 $K[x]$对加法与数乘构成 ***K* 上的线性空间**，即除了 $K[x]$满足是加法交换群的条件以外，对任意 a，$b\in K$，$f(x)$，$g(x)\in K[x]$还满足如下性质：

(1) $(a+b)f(x)=af(x)+bf(x)$；

(2) $a(f(x)+g(x))=af(x)+ag(x)$；

(3) $(ab)f(x)=a(bf(x))$;

(4) $1f(x)=f(x)$.

因为读者以前接触的往往都只是实数域上的多项式，所以在这里有必要通过简单的例子，让读者熟悉一下其他域上的多项式，注意这里的计算与实数域的情况是不同的.

[例 4.3.1] 设 $f(x)=x+1$，对下面两种情况求 $f(x)+f(x)$和 $f(x)f(x)$.

(1) $f(x)\in\mathbf{R}[x]$，即 $f(x)$是 $\mathbf{R}$ 上的一个多项式.

(2) $f(x)\in\mathbf{Z}_2[x]$，即 $f(x)$是 $\mathbf{Z}_2$ 上的一个多项式.

解(1) $f(x)+f(x)=2x+2$，$f(x)f(x)=x^2+2x+1$.

(2) 与(1)有很大的不同：

$$f(x)+f(x)=x+1+x+1=0,$$

$$f(x)f(x)=x^2+x+x+1=x^2+1.$$

大量关于整数的定理在一元多项式环中都有类似的定理，而且相应的证明也是类似的.

4.3.1　一元多项式的整除

由于域$(K,+,\times)$上的一元多项式环 $K[x]$本身一定是交换环，所以定义 4.2.6 和定义 4.2.7 中的“整除”、“线性组合”的概念也适用于 $K[x]$，当然，由于 $K[x]$的元素一定是多项式，所以在实际使用这些术语时，一般将这些术语中涉及的“元”或“元素”用“式”或“多项式”直接指称. 如将“因子”改称“因式”，将“倍元”改称“倍式”，等等.

由定义 4.2.6 可以看出，对任意 $f(x)$，$g(x)$，$h(x)\in K[x]$，及常多项式 $c\in K(c\neq 0)$，多项式整除具有如下性质：

(1) $f(x)\mid 0$;

(2) 如果 $f(x)\neq 0$，则 $0\mid f(x)$;

(3) $c\mid f(x)$;

(4) 如果 $f(x)\mid g(x)$，则 $cf(x)\mid g(x)$;

(5) 如果 $f(x)\mid g(x)$，$g(x)\mid h(x)$，则 $f(x)\mid h(x)$;

(6) 如果 $f(x)\mid g(x)$，$f(x)\mid h(x)$，则 $f(x)$整除 $g(x)$和 $h(x)$的任意线性组合$u(x)g(x)+v(x)h(x)$;

(7) 如果 $f(x)\mid g(x)$，$g(x)\mid f(x)$，则存在 $c\in K(c\neq 0)$，使 $f(x)=cg(x)$.

[例 4.3.2] $\mathbf{Z}[x]$中有$(x+1)\mid(x^2-1)$，$(x-1)\mid(x^n-1)(n\in\mathbf{N})$.

[例 4.3.3] $\mathbf{R}[x]$中有$(x+1)\mid(x^2+1)$，但是，$\mathbf{Z}_2[x]$中有$(x+1)\mid(x^2+1)$，参见例 4.3.1.

定义 4.3.1　如果 $f(x)$，$g(x)\in K[x]$不全为零多项式，且 $d(x)\in K[x]$，$d(x)\neq 0$，若 $d(x)\mid f(x)$ 且 $d(x)\mid g(x)$，则称 $d(x)$为 $f(x)$，$g(x)$的一个**公因式**. 如果 $d(x)$还满足

(1) $d(x)$是首一多项式(即最高次项系数为 1 的多项式);

(2) 对 $f(x)$，$g(x)$的任意公因式 $c(x)$，都有 $c(x)\mid d(x)$，则称 $d(x)$为 $f(x)$，$g(x)$的**最大公因式**，记作$(f(x),g(x))$. 如果$(f(x),g(x))=1$，则称 $f(x)$，$g(x)$**互素**.

定义 4.3.2　设 $p(x)$是 $K[x]$内的一个多项式，$\deg p(x)\geqslant 1$，如果 $p(x)$在 $K[x]$内的因式仅有零次多项式及 $ap(x)$，这里 $a\in K$，$a\neq 0$，则称 $p(x)$是 $K[x]$内的一个**不可约多项式**，

否则称其为**可约多项式**.

多项式是否可约与所在的环有关，例如 x^2+1 在 $\mathbf{R}[x]$上是不可约的，但在 $\mathbf{Z}_2[x]$上是可约的. 参见例 4.3.1，在 $\mathbf{Z}_2[x]$有

$$(x+1)(x+1)=x^2+x+x+1=x^2+1.$$

[例 4.3.4] 对于二次多项式 $f(x)=x^2-2x+2$

(1) 在整数环 $\mathbf{Z}$ 上讨论它的可约性；

(2) 对于任意素数 p，在域 $\mathbf{Z}_p$ 上讨论它的可约性；

(3) 对 $p<10$，在 $\mathbf{Z}_p$ 上分解 $f(x)$.

解

(1) 根据不可约多项式的定义可知，如果 $f(x)$可约，那么 $f(x)$只能分解成如下形式

$$f(x)=(x-\alpha)(x-\beta),$$

然而，只有

$$\alpha=1+i,\ \beta=1-i,$$

是满足要求的 α 和 β 值.

由于 $i\notin\mathbf{Z}$，因此 $f(x)$在 $\mathbf{Z}$ 上不可约. 同理，$f(x)$在实数域 $\mathbf{R}$ 上也不可约，但在复数域 $\mathbf{C}$ 上可约：

$$f(x)=(x-1-i)(x-1+i).$$

(2) 上式中 $i^2=-1$，因此，当且仅当 -1 是模 p 的一个二次剩余时，$f(x)$在 $\mathbf{Z}_p$ 上是可约的. 即存在 $x\in\mathbf{Z}_p$，使得

$$x^2\equiv-1(\mathrm{mod}\ p)$$

成立. 由欧拉判别准则有，当素数 $p\neq2$ 时，判断 $\mathbf{Z}_p$ 中 -1 是否为模 p 的一个二次剩余，只须看下式

$$(-1)^{\frac{p-1}{2}}\equiv1(\mathrm{mod}\ p)$$

是否成立. 如果成立，则 -1 是为模 p 的一个二次剩余，$f(x)$在 $\mathbf{Z}_p$ 上可约；反之，如果不成立，则 -1 不是模 p 的一个二次剩余，$f(x)$在 $\mathbf{Z}_p$ 上不可约. 当然 $P=2$ 的情况要单独考虑.

当 $p=2$ 时，由域的特征的定义有

$$f(x)=x^2-2x+2=x^2=x\cdot x,$$

因此 $f(x)$在 $\mathbf{Z}_2$ 上是可约的.

(3) 在所有小于 10 的且不等于 2 的素数中，满足欧拉判别准则的数只有 5，所以 $f(x)$在 $\mathbf{Z}_5$ 上是可约的：

$$f(x)=x^2-2x+1+1=(x-1)^2-4=(x-1-2)(x-1+2)$$
$$=(x-3)(x+1)=(x+2)(x+1).$$

定理 4.3.1 **(带余除法)** 设 $f(x)$，$g(x)\in K[x]$，$f(x)\neq0$，则存在唯一的 $q(x)$，$r(x)\in K[x]$，使

$$g(x)=q(x)f(x)+r(x),$$

其中 $r(x)=0$ 或 $\deg r(x)<\deg f(x)$. $q(x)$和 $r(x)$分别称为用 $f(x)$去除 $g(x)$所得的**商式**和**余式**.

证明 存在性. 设

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n \quad (a_0\neq 0).$$

如果 $n=0$，则 $f(x)=a_0$，取 $q(x)=\dfrac{1}{a_0}g(x)$，$r(x)=0$ 即可.

下面假定 $n>0$. 对 $g(x)$的次数做数学归纳法. 如果 $g(x)=0$ 或 $\deg g(x)<n$，则令 $q(x)=0$，$r(x)=g(x)$即满足要求. 设 $\deg g(x)<m(m\geqslant n)$时，命题正确，则当 $\deg g(x)=m$ 时，有

$$g(x)=b_0x^m+b_1x^{m-1}+\cdots+b_m \quad (b_0\neq 0).$$

令

$$g_1(x)=g(x)-\frac{b_0}{a_0}x^{m-n}f(x).$$

若 $g_1(x)=0$，则取 $q(x)=\dfrac{b_0}{a_0}x^{m-n}$，$r(x)=0$. 否则，因 $\deg g_1(x)<m$，按归纳假设，存在 $q_1(x)$，$r_1(x)\in K[x]$，使得

$$g_1(x)=q_1(x)f(x)+r_1(x),$$

这里 $r_1(x)=0$ 或 $\deg r_1(x)<\deg f(x)$. 现令

$$q(x)=\frac{b_0}{a_0}x^{m-n}+q_1(x),\ r(x)=r_1(x),$$

则显然有 $g(x)=q(x)f(x)+r(x)$.

唯一性. 设 $\tilde{q}(x)$，$\tilde{r}(x)$也满足命题要求，那么

$$q(x)f(x)+r(x)=\tilde{q}(x)f(x)+\tilde{r}(x),$$
$$[q(x)-\tilde{q}(x)]f(x)=\tilde{r}(x)-r(x).$$

比较两边的次数，即可知 $\tilde{r}(x)-r(x)=0$，$q(x)-\tilde{q}(x)=0$.

定理 4.3.1 相当于初等数论中的整数的带余除法，又称为多项式的欧几里得除法. 我们知道，在初等数论中可以用辗转相除法求两个整数的最大公因子，这种方法同样可以用于求两个多项式的最大公因子.

给定 $f(x)$，$g(x)\in K[x]$，$f(x)\neq 0$，做带余除法

$$g(x)=q(x)f(x)+r(x) \quad (r(x)=0 \text{ 或 } \deg r(x)<\deg f(x)).$$

不难得到$(g(x),f(x))=(f(x),r(x))$. 现在如下做辗转相除法，直到某一步的余式为 0 停止：

$$g(x)=q(x)f(x)+r(x),$$
$$f(x)=q_1(x)r(x)+r_1(x)(\text{若 } r(x)\neq 0 \text{ 执行这一步}),$$
$$r(x)=q_2(x)r_1(x)+r_2(x)(\text{若 } r_1(x)\neq 0 \text{ 执行这一步}),$$
$$\vdots$$

因 $\deg f(x)>\deg r_1(x)>\deg r_2(x)>\cdots$，故必有 $r_{m+1}(x)=0$ 而 $r_m(x)\neq 0$，即

$$r_{m-1}(x)=q_{m+1}(x)r_m(x),$$

于是

$$\begin{aligned}&(g(x),f(x))\\=&(f(x),r(x))\\=&(r(x),r_1(x))\\=&(r_1(x),r_2(x))\end{aligned}$$

$$\begin{aligned}&\vdots\\&=(r_{m-1}(x),\ r_m(x))\\&=(r_m(x),\ r_{m+1}(x))\\&=(r_m(x),\ 0)\\&=ar_m(x),\end{aligned}$$

$ar_m(x)$为首一多项式(即最高次项系数为 1 的多项式). 这便求出了$(f(x),\ g(x))$.

这种求两个多项式最大公因式的方法称为**多项式的辗转相除法**或**广义欧几里得除法**. 类似于整数中的辗转相除法，可以利用回代过程将$(f(x),\ g(x))$表达成 $f(x)$和 $g(x)$的线性组合，即如下定理.

定理 4.3.2 给定不全为零的两个多项式 $f(x),\ g(x)\in K[x]$，则一定存在 $a(x),\ b(x)\in K[x]$ 使得，$(f(x),\ g(x))=a(x)f(x)+b(x)g(x)$.

4.3.2 一元多项式环的理想

由于域上的一元多项式环 $K[x]$本身就是交换环，所以交换环中理想的概念仍然适用于 $K[x]$中的理想. 对任意 $f(x)\in K[x]$，由定理 4.2.14 可知

$$(f(x))=\{u(x)f(x)\mid u(x)\in K[x]\}$$

是由 $f(x)$生成的主理想.

定理 4.3.3 主理想的简单性质：

(1) $(f(x))\subseteq(g(x))$且 $g(x)\neq0\Leftrightarrow g(x)\mid f(x)$.

(2) $(f(x))=(g(x))\Leftrightarrow g(x)=cf(x)$，其中 $c\in K,\ c\neq0$.

证明 (1)如果 $g(x)\mid f(x)$，则 $f(x)$的倍式必然也是 $g(x)$的倍式，即$(f(x))$的元素必然是$(g(x))$的元素，得到$(f(x))\subseteq(g(x))$.

反过来，设 $f(x)=q(x)g(x)+r(x)$，其中 $r(x)=0$ 或 $\deg r(x)<\deg g(x)$. 由理想的性质可知，$r(x)\in(g(x))$，所以只能 $r(x)=0$，即 $f(x)=q(x)g(x)$，因此 $g(x)\mid f(x)$.

(2) 由(1)可知，$(f(x))=(g(x))$的充要条件是 $g(x)\mid f(x)$且 $f(x)\mid g(x)$，即 $g(x)=cf(x)$，其中 $c\in K,\ c\neq0$(由多项式整除的基本性质(7)得来).

定理 4.3.4 域上的一元多项式环$[x]$是主理想环，即如果 I 是 $K[x]$的一个非零理想，则存在 $K[x]$内的首一多项式 $f(x)$，使 $I=(f(x))$.

证明 在 I 中选一个次数最低的多项式 $f(x)$，因对任意的 $a\in K$，有 $af(x)\in I$，故可设 $f(x)$为首一多项式. 由于 $f(x)\in I$，所以 $f(x)$的任意倍式$\in I$，即$(f(x))\subseteq I$. 现设 $g(x)$为 I 中任意元素，按带余除法，有 $q(x),\ r(x)\in K[x]$，使得

$$g(x)=q(x)f(x)+r(x),$$

其中 $r(x)=0$ 或 $\deg r(x)<\deg f(x)$，但 $r(x)=g(x)-q(x)f(x)$仍属于 I，由 $f(x)$是次数最低的多项式可知必定有 $r(x)=0$，于是 $g(x)=q(x)f(x)$，即 $g(x)\in(f(x))$，从而 $I\subseteq(f(x))$，综上所述可得 $I=(f(x))$.

定义 4.3.3 (理想的和) 设 I_1 与 I_2 是 $K[x]$的理想，令

$$I_1+I_2=\{f(x)+g(x)\mid f(x)\in I_1,\ g(x)\in I_2\},$$

则 I_1+I_2 也是 $K[x]$的一个理想(读者自证)，称为 I_1 与 I_2 的**和**.

定理 4.3.5　域 K 上的一元多项式环 $K[x]$ 中二理想 $(f(x))$ 与 $(g(x))$ 的和等于由 $f(x)$ 与 $g(x)$ 的最大公因子生成的理想.

证明　不妨设 $f(x)$，$g(x)$ 不全为零，则

$$(f(x))+(g(x))\neq(0),$$

故可设

$$(f(x))+(g(x))=(d(x)),$$

$d(x)$ 为首一多项式. 因 $(f(x))\subseteq(d(x))$，故 $d(x)\mid f(x)$，同理 $d(x)\mid g(x)$，即 $d(x)$ 为 $f(x)$，$g(x)$ 的一个公因式. 若 $d_1(x)$ 为 $f(x)$ 和 $g(x)$ 的任一公因式，则由 $d_1(x)\mid f(x)$ 推知 $(f(x))\subseteq(d_1(x))$，同理 $(g(x))\subseteq(d_1(x))$，于是

$$(d(x))=(f(x))+(g(x))\subseteq(d_1(x)),$$

而这表明 $d_1(x)\mid d(x)$，所以 $d(x)=(f(x),g(x))$.

这个命题的直接推论如下.

推论 1　设 $f(x)$ 与 $g(x)$ 是域 K 上的一元多项式环 $K[x]$ 中二多项式，$f(x)$ 与 $g(x)$ 的最大公因子为 $d(x)$，则存在 $u(x)$，$v(x)\in K[x]$，使得 $d(x)=u(x)f(x)+v(x)g(x)$. 即定理 4.3.2.

基于这个推论，还可以得到两个重要的推论.

推论 2　设 $f(x)$，$g(x)$ 是 $K[x]$ 内两个不全为零的多项式，则下列命题等价：

(1) $f(x)$ 与 $g(x)$ 互素；

(2) 存在 $u(x)$，$v(x)\in K[x]$，使 $u(x)f(x)+v(x)g(x)=1$；

(3) $(f(x))+(g(x))=K[x]$.

推论 3　设 $f(x)$，$g(x)$，$h(x)\in K[x]$，并且 $f(x)\neq0$，如果 $f(x)\mid g(x)h(x)$ 且 $(f(x),g(x))=1$，则 $f(x)\mid h(x)$.

4.3.3　域上一元多项式唯一分解定理

根据上面定理 4.3.5 推论 3，可得下面的引理.

引理 4.3.1　设 $p(x)$ 为 $K[x]$ 内不可约多项式，设 $f_1(x)$，$f_2(x)$，…，$f_k(x)\in K[x]$. 若 $p(x)\mid\prod_{i=1}^{k}f_i(x)$，则 $p(x)$ 整除某个 $f_j(x)$.

定理 4.3.6　(因式分解唯一定理) 设 K 是一个域，给定多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n\quad(a_i\in K,\ a_0\neq0),$$

则 $f(x)$ 可以分解为

$$f(x)=a_0(p_1(x))^{k_1}(p_2(x))^{k_2}\cdots(p_r(x))^{k_r}(k_i>,\ i=1,2,\cdots,r),$$

其中 $p_1(x)$，…，$p_r(x)$ 是 $K[x]$ 内首项系数为 1 且两两不同的不可约多项式. 而且，除了不可约多项式的排列次序外，上面的分解式是由 $f(x)$ 唯一决定的.

证明　先证存在性，对 $\deg f(x)$ 做数学归纳法. 当 $\deg f(x)=0$ 时，命题显然成立.

设命题对 $\deg f(x)<n$ 的多项式 $f(x)$ 成立. 下面考察 $\deg f(x)=n$ 时的情况.

如果 $f(x)$ 本身是不可约的，则 $p_1(x)=\dfrac{1}{a_0}f(x)$ 仍为不可约多项式，而 $f(x)=a_0p_1(x)$，故命题成立.

如果 $f(x)$可约，那么它有一个非平凡因式 $g(x)$，故有分解式 $f(x)=g(x)h(x)$，这里 $0<\deg g(x)<\deg f(x)$，$0<\deg h(x)<\deg f(x)$，按照归纳假设，$g(x)$与 $h(x)$均可分解为互不相同的不可约多项式的幂的乘积，这样，$f(x)$显然也有这样的分解式.

再证唯一性. 对 $\deg f(x)$做数学归纳法. $\deg f(x)=0$ 时命题显然成立.

设命题对 $\deg f(x)<n$ 的多项式 $f(x)$成立. 现考察 $\deg f(x)=n$ 的情形. 设其有两个分解式. 因为 $a_0\neq 0$，约去 a_0 后得到

$$(p_1(x))^{k_1}(p_2(x))^{k_2}\cdots(p_r(x))^{k_r}=(q_1(x))^{l_1}(q_2(x))^{l_2}\cdots(q_s(x))^{l_s}, \tag{4.3.1}$$

从上式知 $p_1(x)\mid(q_1(x))^{l_1}(q_2(x))^{l_2}\cdots(q_s(x))^{l_s}$，因为 $p_1(x)$是不可约多项式，根据引理，$p_1(x)$整除某个 $q_i(x)$，不妨设 $p_1(x)\mid q_1(x)$. 但 $q_1(x)$也是不可约多项式，故只能有

$$p_1(x)=aq_1(x)(a\in K).$$

又因为 $p_1(x)$与 $q_1(x)$首项系数都是 1，故 $a=1$，即 $p_1(x)=q_1(x)$，从式(4.3.1)两边消去 $p_1(x)$，得

$$g(x)=(p_1(x))^{k_1-1}(p_2(x))^{k_2}\cdots(p_r(x))^{k_r}=(q_1(x))^{l_1-1}(q_2(x))^{l_2}\cdots(q_s(x))^{l_s}.$$

现在 $\deg g(x)=\deg f(x)-\deg p_1(x)<n$，按照归纳法，应有 $r=s$，且适当排列不可约多项式次序后，有 $p_i(x)=q_i(x)$，$k_i=l_i(i=1,2,\cdots,r)$. 由此可知，$f(x)$的分解式是唯一的.

4.3.4 多项式不可约性检验

由 4.3.3 节可知，不可约多项式就像整数当中的素数一样，因此不可约多项式的重要性也类似于素数的重要性. 本节讨论域上多项式的分解算法与不可约性检验问题.

定理 4.3.7 设多项式 $p(x)$，$q(x)\in K[x]$，则有

$$\deg(p(x)q(x))=\deg p(x)+\deg q(x).$$

证明 这个结论的成立依赖于一个基本事实：域中没有零因子. 令

$$p(x)=a_0+a_1x+a_2x^2+\cdots+a_{m-1}x^{m-1}+a_mx^m,$$

$$q(x)=b_0+b_1x+b_2x^2+\cdots+b_{n-1}x^{n-1}+b_nx^n.$$

显然这里 $p(x)$，$q(x)$的首项系数均不为零. 在乘积 $p(x)q(x)$中最高次项是 x^{m+n}，它只可能由 $p(x)$和 $q(x)$的首项相乘而来，因此它的系数为 a_mb_n. 又因为 $p(x)$和 $q(x)$的首项系数都不为零，而且域中非零元素乘积仍非零，所以 x^{m+n}的系数非零. 这就证明了多项式乘积的次数等于各自次数的和. 根据这一结论容易得出：如果 $K[x]$中的多项式 $d(x)$是多项式 $f(x)$的一个真因子，则有

$$0<\deg d(x)<\deg f(x).$$

根据定理 4.3.7，可以得到有关域上多项式分解的如下一些简单推论：

推论 1 每个线性多项式是不可约的，因为不存在次数介于 1 和 0 之间的多项式；

推论 2 如果一个二次多项式可以真分解，则它必定是两个线性因子的乘积；

推论 3 如果一个三次多项式可以真分解，则它必定至少有一个线性因子；

推论 4 如果一个四次或更高次多项式可以真分解，则它可能没有线性因子.

定理 4.3.8 $K[x]$上的多项式 $f(x)$有一个线性因子 $x-a(a\in K)$，当且仅当 $f(a)=0$.

证明 如果 $f(x)=(x-a)q(x)$，那么 $f(a)=(a-a)q(a)=0$ $q(a)=0$.

反过来，设 $f(x)=(x-a)q(x)+r(x)$，其中 $\deg r(x)=0$ 或者 $r(x)=0$，即

$$f(x)=(x-a)q(x)+c,$$

其中 $c\in K$. 因为 $f(a)=0$，所以

$$c=f(a)-(a-a)q(a)=0-0=0,$$

即 $f(x)=(x-a)q(x)$，从而 $x-a$ 是 $f(x)$的线性因子.

[例 4.3.5]　下面来讨论 $\mathbf{Z}_2[x]$中的不可约多项式和多项式的分解问题.

$\mathbf{Z}_2[x]$中只有 1 个零次多项式，那就是 1，一般认为零多项式的次数是无穷.

在 $\mathbf{Z}_2[x]$中刚好有两个线性多项式，即 x 和 $x+1$. 根据定理 4.3.7 推论 1，它们是不可约的.

$\mathbf{Z}_2[x]$中有 4 个二次多项式：x^2，x^2+x，x^2+1，x^2+x+1，下面分别来考察它们的不可约性，

$$x^2=x\,x;$$

$$x^2+x=x(x+1);$$

$$x^2+1=(x+1)^2;$$

对于 x^2+x+1，根据定理 4.3.8 容易验证这个多项式没有线性因子，$0^2+0+1=1$，$1^2+1+1=1$，所以 x^2+x+1 在 $\mathbf{Z}_2[x]$是不可约的，它也是 $\mathbf{Z}_2[x]$中唯一的二次不可约多项式.

$\mathbf{Z}_2[x]$中有 8 个三次多项式. 如果只考虑不可约的，则首先可以把常数项为 0 的多项式排除，因为它们必定有线性因子 x. 根据定理 4.3.8，$\mathbf{Z}_2[x]$中所有常数项非零且 x 取 1 得到 0 值的三次多项式，必定有一个因子 $x+1$. 因此 $\mathbf{Z}_2[x]$中只有两个不可约的三次多项式，它们是 x^3+x^2+1，x^3+x+1.

$\mathbf{Z}_2[x]$中有 16 个四次多项式. 如果常数项为 0，则它必有因子 x；如果非零系数的个数为偶数，则它必有因子 $x+1$，这样只有 4 个可能的四次不可约多项式：

$$x^4+x^3+x^2+x+1,$$

$$x^4+x^3+1,$$

$$x^4+x^2+1,$$

$$x^4+x+1.$$

容易看出，这四个多项式在 $\mathbf{Z}_2[x]$中均没有线性因子. 接下来寻找它们的不可约二次因子，由前面的讨论可知，$\mathbf{Z}_2[x]$上的二次不可约多项式为 x^2+x+1，这样就只有 $x^4+x^2+1=(x^2+x+1)^2$是可约的，余下的三个四次多项式 $x^4+x^3+x^2+x+1$，x^4+x^3+1，x^4+x+1 则为 $\mathbf{Z}_2[x]$上的四次不可约多项式.

$\mathbf{Z}_2[x]$中有 32 个五次多项式. 除去常数项为 0 的就只剩下 16 个，再排除那些有偶数个非零系数项的多项式(它们有线性因子 $x+1$)，就剩下 $C_4^3+C_4^1=8$ 个需要进一步判断，这 8 个多项式没有线性因子，所以它们只可能是如下形式的低次多项式的乘积：

不可约的二次多项式×不可约的三次多项式

我们已经知道 $\mathbf{Z}_2[x]$中只有一个二次不可约多项式，有两个三次不可约多项式. 这样就有两个没有线性因子的五次可约多项式，它们是

$$(x^2+x+1)(x^3+x^2+1)=x^5+x+1,$$

$$(x^2+x+1)(x^3+x+1)=x^5+x^4+1.$$

这样在 $\mathbf{Z}_2[x]$中就有 6 个五次不可约多项式. 除了上面排除的两个可约多项式以外，它

们必须是：具有常数项 1，非零系数个数为奇数. 下面给出这 6 个五次不可约多项式：

$$x^5+x^3+x^2+x+1,$$
$$x^5+x^4+x^2+x+1,$$
$$x^5+x^4+x^3+x+1,$$
$$x^5+x^4+x^3+x^2+1,$$
$$x^5+x^3+1,$$
$$x^5+x^2+1.$$

将 $\mathbf{Z}_2[x]$五次以内的不可约多项式总结如表 4.3.1 所示.

表 4.3.1 $Z_2[x]$上的低次不可约多项式

次数	不可约多项式
1	x，$x+1$
2	x^2+x+1
3	x^3+x^2+1，x^3+x+1
4	$x^4+x^3+x^2+x+1$，x^4+x^3+1，x^4+x+1
5	$x^5+x^3+x^2+x+1$，$x^5+x^4+x^2+x+1$，$x^5+x^4+x^3+x+1$，$x^5+x^4+x^3+x^2+1$，x^5+x^3+1，x^5+x^2+1

4.3.5 一元多项式的同余与商环

定义 4.3.4 给定 $K[x]$中的一个首一多项式 $m(x)$. 如果 $K[x]$中的两个多项式 $f(x)$，$g(x)$ 满足

$$m(x)\mid(f(x)-g(x)),$$

则我们称 $f(x)$和 $g(x)$**模** $m(x)$**同余**，记作

$$f(x)\equiv g(x)(\bmod m(x)),$$

否则，称 $f(x)$和 $g(x)$**模** $m(x)$**不同余**. 记作

$$f(x)\not\equiv g(x)(\bmod m(x)).$$

由多项式的带余除法知，任一多项式都与其被 $m(x)$除的余式 $r(x)$模 $m(x)$同余，其中 $\deg r(x)<\deg m(x)$或者 $r(x)=0$，该余式 $r(x)$叫作 $f(x)$模 $m(x)$的**最小余式**，记为 $f(x)\%m(x)$或者 $f(x)(\bmod m(x))$.

[例 4.3.6] 设

$$f(x)=x^5+x^4+x^3+x^2+x+1\in\mathbf{Z}_2[x],$$
$$g(x)=x^3+x+1\in\mathbf{Z}_2[x],$$

求 $f(x)\%\ g(x)$.

解 因为

$$
\begin{array}{r}
x^2+x \\
x^3+x+1 \overline{)\,x^5+x^4+x^3+x^2+x+1} \\
\underline{x^5\quad\ +x^3+x^2\quad\ \ } \\
x^4+\quad\ +x+1 \\
\underline{x^4+\quad\ +x^2+x\quad\ } \\
x^2+1
\end{array}
$$

所以 $f(x)\%\ g(x)=x^2+1$.

定义 4.3.5　设 $p(x)\in K[x]$，$(p(x))$是由 $p(x)$生成的理想，则 $K[x]/(p(x))$称一元多项式环 $K[x]$关于 $p(x)$的商环. 该商环中的“加法”和“乘法”运算法则定义为：对任意

$$[f(x)]=\{f(x)+u(x)p(x)\mid f(x),\ u(x)\in K[x]\}\ \in K[x]/(p(x)),$$
$$[g(x)]=\{g(x)+u(x)p(x)\mid g(x),\ u(x)\in K[x]\}\ \in K[x]/(p(x)),$$

有

$$[f(x)]+[g(x)]=[f(x)+g(x)],$$
$$[f(x)]\,[g(x)]=[f(x)g(x)].$$

这个定义就是前面商环讨论中的关于商环上两种运算定义的具体化，在这里明确写出只是为了让读者看得更清晰. 当然，上面两个运算也与下面两式等价

$$[f(x)]+[g(x)]=[(f(x)+g(x))\%p(x)],$$
$$[f(x)]\,[g(x)]=[(f(x)g(x))\%p(x)].$$

定理 4.3.9　设 $p(x)\in K[x]$且为不可约多项式，则商环 $K[x]/(p(x))$对定义 4.3.5 中定义的“加法”和“乘法”运算构成一个域.

证明　由商环的讨论可知，显然 $K[x]/(p(x))$是一个交换环，有幺元[1]，所以只要证明 $K[x]/(p(x))$中的非零元在 $K[x]/(p(x))$中都有乘法逆元即可. 因为 $p(x)$是不可约多项式，所以对任意$[f(x)]\in K[x]/(p(x))$且$[f(x)]\neq 0$，都有$(f(x),\ p(x))=1$. 根据定理 4.3.5 推论 2，存在多项式 $s(x)$，$t(x)\in K[x]$使得

$$s(x)f(x)+t(x)p(x)=1,$$

即 $s(x)f(x)\equiv 1(\mathrm{mod}\ p(x))$，这说明$[f(x)]$为可逆元素，$[s(x)]$为其逆元，从而 $K[x]/(p(x))$中的任意非零元素都为可逆元素，即 $K[x]/(p(x))$构成一个域.

[例 4.3.7]　设 $K=\mathbf{Z}/q\mathbf{Z}$，其中 q 是素数. 设 $p(x)$是 $K[x]$中的 n 次不可约多项式，则

$$K[x]/(p(x))=\{[a_{n-1}x^{n-1}+\cdots a_1x+a_0]\mid a_i\in K\},$$

可以将这个集合看作由所有次数小于 n，系数在 K 内的多项式组成. 这是一个元素个数有限的域，其元素个数为 p^n.

4.4　有限域理论初步

元素个数有限(阶数有限)的域称为有限域. 有限域中的运算可以使运算结果保持在有限的范围内，且不会有取整误差. 有限域理论在密码学和编码理论领域(如椭圆曲线、离散对

数、AES(Rijndael)加密体系、纠错码等等)有广泛的应用.

定义 4.4.1 阶数有限的域称为**有限域**,或者**伽罗华域**(Galois Field).

由定理 4.2.9 知,有限域的特征必为素数. 通过前面的学习已知,对于任一素数 p,一定存在特征为 p 的有限域 $\mathbf{Z}/p\mathbf{Z}=\mathbf{Z}_p$. 事实上,$\mathbf{Z}/p\mathbf{Z}$ 是最简单的有限域,它含有 p 个元素,但是,在应用中出于不同的原因,我们需要更多的有限域. 因为计算机都采用二进制,为方便机器实现,使用特征为 2 的域是最方便的,在 $\mathbf{Z}/p\mathbf{Z}$ 这样的域中,只有 $\mathbf{Z}/2\mathbf{Z}$ 满足这个条件,但在很多应用场合下,我们希望有限域中有较多的元素,如果仅局限于 $\mathbf{Z}/p\mathbf{Z}$ 上,就不可能同时满足"有限"和"特征为 2"这两方面的要求.

由例 4.3.7 可知,对于每一个素数 p 和每一个正整数 n,都存在一个含有 p^n 个元素的有限域. 这样的结论反过来说也成立,即有下面定理(证明略).

定理 4.4.1 对于任一素数 p 和任一正整数 n,必然存在阶为 p^n 的有限域,并且在同构意义下,这样的有限域是唯一的.

定理 4.4.2 有限域中元素的个数必为素数,或者素数的正整数幂.

元素个数为素数 p 的有限域记为 $\mathbf{GF}(p)$,元素个数为素数 p 的 n 次幂(n 正整数)的有限域记为 $\mathbf{GF}(p^n)$.

定义 4.4.2 有限域 $\mathbf{F}$ 所包含的最小子域称为 $\mathbf{F}$ 的**素域**,$\mathbf{F}$ 的素域的阶称为 $\mathbf{F}$ 的特征,设 p 是素数,有限域 $\mathbf{F}$ 的阶为 $q=p^n$,则 $\mathbf{F}$ 的素域的阶为 p,$\mathbf{F}$ 是其素域的扩域.

例 4.3.7 实际上给出了一种有限域的构造方法,即设 K 为有限域,$m(x)$ 是 $K[x]$ 中的 n 次不可约多项式,则商环

$$K[x]/(m(x))$$

可视为一个有限域. 当 $K=\mathbf{Z}/p\mathbf{Z}=\mathbf{Z}_p$(其中 p 为素数)时,这个有限域中元素的个数为 p^n,该有限域就是 $\mathbf{GF}(p^n)$,不可约多项式 $m(x)$ 称为有限域 $\mathbf{GF}(p^n)$ 的**生成多项式**. 事实上,对于任意素数 p 和正整数 n,一定存在 $\mathbf{Z}_p[x]$ 上的 n 次不可约多项式.

关于 $\mathbf{GF}(p^n)$ 与 $\mathbf{GF}(p)$ 的关系,直接给出如下结论:

(1) $\mathbf{GF}(p^n)$ 是 $\mathbf{GF}(p)$ 的扩域,实际上,更详细地说就是,$\mathbf{GF}(p)$ 与 $\mathbf{GF}(p^n)$ 的一个子域同构,该子域是集合$\{0,1,1+1,1+1+1,\cdots,\underbrace{1+1+\cdots+1}_{p-1\text{个}1}\}$.

(2) $\mathbf{GF}(p^n)$ 中包含了 $\mathbf{GF}(p)$ 上所有 n 次不可约多项式的全部 n 个根.

(3) 如果 $\mathbf{GF}(p^n)$ 的生成多项式 $m(x)$ 有一个根 α,则可把 $\mathbf{GF}(p^n)$ 中的每一个元素表示成系数在 $\mathbf{GF}(p)$ 上且次数低于 n 的 α 的多项式. 这种表示方法称为**有限域的多项式基**,我们可以用多项式基对有限域中的元素进行统一编码,从而方便用软件或硬件来实现有限域运算.

[例 4.4.1] 求 $\mathbf{GF}(2^3)$ 的元素和以不可约多项式 x^3+x+1 为生成多项式的加法表与乘法表.

解 设 $\mathbf{GF}(2^3)$ 中的元素为 $a_2x^2+a_1x+a_0$(a_2,a_1,$a_0\in\mathbf{Z}/2\mathbf{Z}$),根据 a_2,a_1,a_0 的不同取值,可得

a_2	a_1	a_0	$a_2x^2+a_1x+a_0$
0	0	0	0
0	0	1	1
0	1	0	x
0	1	1	$x+1$
1	0	0	x^2
1	0	1	x^2+1
1	1	0	x^2+x
1	1	1	x^2+x+1

所以 **GF**(2^3)的元素为

$$0，1，x，x+1，x^2，x^2+1，x^2+x，x^2+x+1.$$

以 x^3+x+1 为生成多项式的加法表与乘法表分别如表 4.4.1 和表 4.4.2 所示.

表 4.4.1　GF(2^3)的加法表

+	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
0	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
1	1	0	$x+1$	x	x^2+1	x^2	x^2+x+1	x^2+x
x	x	$x+1$	0	1	x^2+x	x^2+x+1	x^2	x^2+1
$x+1$	$x+1$	x	1	0	x^2+x+1	x^2+x	x^2+1	x^2
x^2	x^2	x^2+1	x^2+x	x^2+x+1	0	1	x	$x+1$
x^2+1	x^2+1	x^2	x^2+x+1	x^2+x	1	0	$x+1$	x
x^2+x	x^2+x	x^2+x+1	x^2	x^2+1	x	$x+1$	0	1
x^2+x+1	x^2+x+1	x^2+x	x^2+1	x^2	$x+1$	x	1	0

表 4.4.2　GF(2^3)的乘法表

×	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
0	0	0	0	0	0	0	0	0
1	0	1	x	$x+1$	x^2	x^2+1	x^2+x	x^2+x+1
x	0	x	x^2	x^2+x	$x+1$	1	x^2+x+1	x^2+1
$x+1$	0	$x+1$	x^2+x	x^2+1	x^2+x+1	x^2	1	x
x^2	0	x^2	$x+1$	x^2+x+1	x^2+x	x	x^2+1	1
x^2+1	0	x^2+1	1	x^2	x	x^2+x+1	$x+1$	x^2+x
x^2+x	0	x^2+x	x^2+x+1	1	x^2+1	$x+1$	x	x^2
x^2+x+1	0	x^2+x+1	x^2+1	x	1	x^2+x	x^2	$x+1$

[例 4.4.2] 给素域 $\mathbf{Z}/5\mathbf{Z}$ 添加一个 2 的平方根，可以构造有限域$(\mathbf{Z}/5\mathbf{Z})[x]/((x^2-2))$. 首先注意到在 $\mathbf{Z}/5\mathbf{Z}$ 中 2 不是二次剩余(不存在 2 的平方根)，因此二次多项式 x^2-2 在 $\mathbf{Z}/5\mathbf{Z}$ 上是不可约多项式，于是$(\mathbf{Z}/5\mathbf{Z})[x]/(x^2-2)$构成有限域，它实际上就是 $\mathbf{GF}(5^2)$. $\mathbf{GF}(5^2)$的 25 个元素为$[ax+b]$，$0\leqslant a\leqslant 4$，$0\leqslant b\leqslant 4$. $\mathbf{Z}/5\mathbf{Z}$ 是$(\mathbf{Z}/5\mathbf{Z})[x]/(x^2-2)$的子域，并且有 $x^2\equiv 2 \bmod(x^2-2)$，所以$[x]$就是 2 的平方根.

下面讨论有限域 $\mathbf{GF}(p^n)$中乘法逆元的求法. 设 n 次多项式 $m(x)$是 $\mathbf{GF}(p^n)$的生成多项式，则 $m(x)$是不可约多项式. 对于 $\mathbf{GF}(p^n)$中的任意可逆元素$[f(x)]$，必有 $f(x)\not\equiv 0 \bmod m(x)$，为了寻找 $f(x)$模 $m(x)$的逆元，可以使用扩展欧几里得算法在 $\mathbf{GF}(p^n)$中求多项式 $s(x)$和 $t(x)$，使得

$$s(x)f(x)+t(x)m(x)=1,$$

那么即有

$$s(x)f(x)-1=-t(x)m(x).$$

因此 $s(x)f(x)\equiv 1 \bmod m(x)$，即$[f(x)]$的逆元为$[s(x)]$. 因为 $f(x)\not\equiv 0 \bmod m(x)$且 $m(x)$是不可约的，$f(x)$和 $m(x)$的最大公因子为 1，所以这样的 $s(x)$和 $t(x)$一定是存在的.

[例 4.4.3] 在有限域 $\mathbf{Z}_2[x]/((x^2+x+1))$上求 x 的乘法逆元.

解 首先由欧几里得算法有

$$(x^2+x+1)-(x+1)x=1,$$

于是有

$$(x+1)x\equiv 1 \bmod(x^2+x+1),$$

即

$$x^{-1}\equiv(x+1)\bmod(x^2+x+1).$$

[例 4.4.4] 在有限域 $\mathbf{Z}_2[x]/((x^4+x+1))$上求 x^2+x+1 的乘法逆元.

解 由欧几里得算法有

$$(x^4+x+1)-(x^2+x)(x^2+x+1)=1,$$

于是有

$$(x^2+x)(x^2+x+1)\equiv 1 \bmod(x^4+x+1),$$

这样就有

$$(x^2+x+1)^{-1}\equiv(x^2+x)\bmod(x^4+x+1).$$

关于有限域的乘法，还有一个非常有用的性质. 我们注意到，对任意有限域 F，由于它的非零元素都有逆元，所以 $F-\{0\}$必然是一个(乘法)群，而且人们已经证明 $F-\{0\}$是一个循环群.

[例 4.4.5] 在 $\mathbf{GF}(2^3)$中，x 是一个 $\mathbf{GF}(2^3)-\{0\}$的生成元，即 $\mathbf{GF}(2^3)-\{0\}=<x>$，$\mathrm{ord}(x)=7$，下面通过计算来加以验证：

$$x^0\equiv 1(\bmod\ x^3+x+1),$$
$$x^1\equiv x(\bmod\ x^3+x+1),$$
$$x^2\equiv x^2(\bmod\ x^3+x+1),$$
$$x^3\equiv x+1(\bmod\ x^3+x+1),$$
$$x^4\equiv x^2+x(\bmod\ x^3+x+1),$$
$$x^5\equiv x^2+x+1(\bmod\ x^3+x+1),$$

$$x^6 \equiv x^2+1 (\bmod\ x^3+x+1),$$
$$x^7 \equiv 1 \equiv x^0 (\bmod\ x^3+x+1).$$

可以利用这些公式快速计算 **GF**(2^3)中元素的乘积. 比如，可以求 x^2+x+1 和 x^2+1 的乘积：

$$(x^2+x+1)(x^2+1)=x^5\ x^6=x^{11}=x^4\ x^7=x^4 1=x^4=x^2+x,$$

结果与表 4.4.2 中一致.

习题

1. 求证：正有理数集合 Q^+ 和正实数集合 R^+ 在普通乘法下是群.
2. 在矩阵乘法下，所有 2×2 的可逆矩阵组成群，验证：矩阵 $\begin{bmatrix}1 & 1\\ 0 & 1\end{bmatrix}$ 是该群的元素，且它是无限阶的.
3. 在整数集 **Z** 中定义二元运算“$*$”为 $n*m=-n-m$，n，$m\in\mathbf{Z}$. 证明这个二元运算是交换的，但不是结合的.
4. 在整数集 **Z** 中定义二元运算“$*$”为 $n*m=n+m-2$，n，$m\in\mathbf{Z}$. 证明(**Z**，$*$)是群.
5. 证明群的同构关系是等价关系.
6. 证明：若 G 为有限集且对运算“$\cdot$”满足结合律与消去律，则(G，$\cdot$)构成一个群.
7. 设(G，$\cdot$)是一个群，H 是 G 的一个有限子集，证明(H，$\cdot$)构成(G，$\cdot$)的子群的充要条件是：对任意 a，$b\in H$ 有 $a\cdot b\in H$.
8. 质数阶的群没有非平凡子群，且一定是循环群.
9. 求证循环群是交换群.
10. 证明群中元素与其逆元具有相同的阶.
11. 有限群(G，$\cdot$)中的任何元素 a 的阶可整除 $|G|$.
12. 将置换 $\begin{pmatrix}1&2&3&4&5&6&7&8\\2&5&6&8&1&4&7&3\end{pmatrix}$ 分解成不相交的轮换.
13. 将置换之积 $\begin{pmatrix}1&2&3&4&5&6&7&8\\2&5&6&8&1&4&7&3\end{pmatrix}\circ\begin{pmatrix}1&2&3&4&5&6&7&8\\2&5&6&4&1&8&7&3\end{pmatrix}$ 分解成不相交的轮换.
14. 求证：定义在整环上的多项式集能构成整环.
15. 求证高斯整数环的可逆元素只有-1、1、i 和$-i$.
16. 令 $\mathbf{Z}(\sqrt{2})=\{a+b\sqrt{2}\mid a,b\in\mathbf{Z}\}$，则($\mathbf{Z}(\sqrt{2})$，$+$，$\times$)是一个交换环，都有哪些可逆元素呢？
17. 求证($Q(\sqrt{2})$，$+$，$\times$)是整环也是域.
18. 求证环($\mathbf{Z}_m$，$+$，$\times$)是整环的充要条件是 m 为素数.
19. 找出 $\mathbf{Z}/6\mathbf{Z}$ 的所有理想.
20. 试证明商环($\mathbf{Z}[x]/((x))$，$+$，$\times$)与整数环(**Z**，$+$，$\times$)同构.
21. 设 **F** 是域，证明 **F** 的分式域是 **F** 自身.
22. 举例说明环上的多项式与域上的多项式有哪些不同点.
23. 求 $x^2+3x+2=0$ 在 $\mathbf{Z}/5\mathbf{Z}$ 中的两个根.
24. 验证 $\mathbf{Z}/8\mathbf{Z}$ 的乘法群不能由单个元素生成.

25. 找出 $\mathbf{Z}_3[x]$中所有 2 次和 3 次的不可约多项式.

26. 证明不存在元素 $x\in\mathbf{Z}_{13}$，使 $x^5=1$，除非 $x=1$.

27. 验证 x^5+x+1 在 $\mathbf{Z}_2[x]$中不可约.

28. 求 $\mathbf{Z}_3[x]$中的多项式 x^6+x^3+1 和 x^2+x+1 的最大公因式.

29. 将 $\mathbf{Z}_2[x]$中的多项式 $x^{12}+x^9+x^8+x^7+x^6+x^5+1$ 分解为不可约因式的乘积.

30. 求证：在群$(\mathbf{Z}/m\mathbf{Z},+)$中，$m$ 是使 $ma=1$ 成立的最小正整数.

31. 证明在环 $\mathbf{Z}/35\mathbf{Z}$ 中消去律不成立：即存在 a，b，$c\in\mathbf{Z}/35\mathbf{Z}$，使得 $ca=cb$ 但 $a\neq b$.

32. 编程实现求 $\mathbf{Z}_2[x]$上两个多项式的乘法算法.

33. 编程实现求 $\mathbf{Z}_2[x]$上两个多项式的最大公因式的算法.

34. 编程实现多项式的欧几里德除法.

35. 设有限域 $\mathbf{Z}$ 的特征为 p，证明对任意的 a，$b\in\mathbf{F}$，恒有 $a^{p^n}+b^{p^n}=(a+b)^{p^n}$.

36. $p(x)\in K[x]$，m 和 n 是正整数，求证$(p^n(x)-1,\ p^m(x)-1)=p^{(m,\ n)}(x)-1$.

第 5 章　椭 圆 曲 线

第六届国际密码学会议推荐了两种应用于公钥密码体系的加密算法：基于大整数因子分解问题(IFP)的 RSA 算法和基于椭圆曲线上离散对数计算问题(ECDLP)的 ECC 算法. 它们随后得到广泛的关注和应用. 本章主要介绍椭圆曲线密码体系的数学原理，内容包括椭圆曲线、离散对数和椭圆曲线上的离散对数等，为系统地学习现代密码学打下基础.

5.1　椭圆曲线的预备知识

5.1.1　仿射平面和射影平面

定义 5.1.1　域 K 上的集合 $K^2=\{(x, y) \mid x, y\in K\}$ 称为域 K 上的**仿射平面**，K^2 中的元素称为**仿射平面上的点**，可以用**仿射坐标**(x, y)表示. 例如中学里讲过的笛卡儿平面就是实数域 $\mathbf{R}$ 上的仿射平面，又称为欧氏平面.

定义 5.1.2　设$(K, +, \cdot)$为域，0为 K 中的零元，$K^*=K\backslash\{0\}$，R 是集合

$$M=\{(x, y, z) \mid x, y, z\in K\}\backslash\{0, 0, 0\}$$

上的一个等价关系，$(x_1, y_1, z_1)R(x_2, y_2, z_2)$当且仅当存在 $u\in K^*$，$x_1=u\cdot x_2$，$y_1=u\cdot y_2$，$z_1=u\cdot z_2$，则 M 由 R 划分的等价类集合(商集)M/R 称为**域 K 上的射影平面**. 通常把 M/R 中的等价类称为**射影平面上的点**，并可以用等价类中的任意一个元素作为代表元. 这个用来表示射影平面上的点的代表元称为射影平面的**齐次坐标**，记作(X, Y, Z).

由仿射平面和射影平面的定义知，对于定义在同一个域上的仿射平面和射影平面，当 $Z\neq\bar{0}$时，射影平面中的点(X, Y, Z)与仿射平面中的点(x, y)可以建立如下对应关系：

$$x=X/Z,$$
$$y=Y/Z.$$

下面在实数域 $\mathbf{R}$ 上给出仿射平面和射影平面的一个实例，为此，先引入无穷远点和无穷远直线的概念. 我们知道，在欧氏几何中，平面上任意两条相异直线位置关系只有两种——平行和相交，引入无穷远点后，可以将这两种关系统一起来.

如图 5.1.1 所示，$AB\perp L_1$，AB 交 L_1 于 A，$L_1/\!/L_2$，直线 AP 由 AB 起绕 A 点逆时针方向转动，P 为 AP 与 L_2 的交点. 当$\angle\theta\to 90^\circ$时，有 $AP\to L_1$，可设想 L_2 上有一点 P_∞，它是 L_1 与 L_2 的交点，称之为无穷远点.

定义 5.1.3　平面上两条平行线的交点称为**无穷远点**，记为 P_∞.

在无穷远点的定义中，有几点需要注意.

(1) 因为平行线是有方向的，所以无穷远点也是有方向的，不同方向平行直线的无穷远点不同.

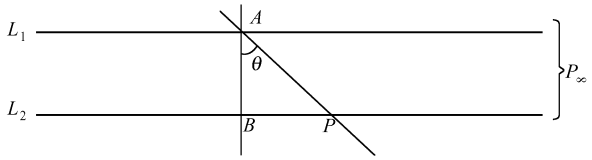

图 5.1.1 无穷远点

(2) 一条直线 L 上的无穷远点只能有一个(因为过定点且与已知直线平行的直线只能有一条，而两条直线的交点只有一个).

(3) 一组相互平行的直线有公共的无穷远点.

(4) 平面上任何两条相交的直线 L_1 和 L_2 有不同的无穷远点(设 L_1 和 L_2 交于点 A，且有相同的无穷远点 P_∞，则过两相异点 A 和 P_∞ 有两相异直线 L_1 和 L_2，与欧几里得第一公理矛盾).

(5) 一个平面上全体无穷远点(各个方向上的无穷远点组成的集合)构成一条无穷远直线.

有了无穷远点的概念后，我们来探讨实数域 **R** 上的欧氏平面与射影平面之间的关系，即平面直角坐标与齐次坐标之间的关系. 进而揭示无穷远点的数学本质.

平面上两相异直线 L_1 和 L_2 的平面直角坐标方程为

$$L_1: a_1x+b_1y+c_1=0,$$

$$L_2: a_2x+b_2y+c_2=0,$$

其中 a_1，b_1 不同时为 0，a_2，b_2 不同时为 0. 设

$$D=\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}, \quad D_x=\begin{vmatrix} b_1 & c_1 \\ b_2 & b_2 \end{vmatrix}, \quad D_y=\begin{vmatrix} c_1 & a_1 \\ c_2 & a_2 \end{vmatrix}.$$

当 $D\neq 0$ 时，L_1 和 L_2 相交于一平常点 $P(x, y)$(为与无穷远点相区别，把原来欧氏平面上的点叫作平常点)，由克莱姆法则，$P(x, y)$的坐标为

$$x=\frac{D_x}{D},$$

$$y=\frac{D_y}{D}.$$

可以看出，表征交点 $P(x, y)$位置信息的独立变量有三个，即 D_x，D_y，D，因此可以抽象地将 P 表示为(D_x, D_y, D). 这恰好是齐次坐标的形式. 由齐次坐标的性质有

$$\left(\frac{D_x}{D}, \frac{D_y}{D}, 1\right)=(x, y, 1).$$

当 $D=0$ 时，$L_1 /\!/ L_2$，可以说 L_1 和 L_2 相交于一个无穷远点 P_∞，用过原点且平行于 L_1 的一条直线的方向来指出它的方向，而这条直线的方程为 $a_1x+b_1y=0$(或 $a_2x+b_2y=0$). 由于 $D=0$，表征 P_∞ 方向信息的独立变量只有两个，即 D_x，D_y，也可以将 P_∞ 抽象地表示为 $(D_x, D_y, 0)$. 设此时 L_1 和 L_2 的斜率为 k，则有

$$k=-\frac{a_1}{b_1}=-\frac{a_2}{b_2},$$

$$(D_x, D_y, 0)=(D_x, kD_x, 0)\Rightarrow(1, k, 0)=(b_1, -a_1, 0)=(b_2, -a_2, 0).$$

如上所述，如果把平常点和无穷远点都用齐次坐标(X, Y, Z)的形式来表示，并规定：

(1) X，Y，Z 不能同时为 0；

(2) 当 $Z\neq 0$ 时，令 $x=\dfrac{X}{Z}$，$y=\dfrac{Y}{Z}$，则(x, y)表示欧氏平面上的点(平常点)；

(3) 当 $Z=0$ 时，(X, Y, Z)表示无穷远点，$-\dfrac{Y}{X}$表示此无穷远点对应的平行线的斜率($X=0$对应着与 x 轴垂直的一组平行线上的无穷远点).

这样就可以把平常点和无穷远点的坐标统一起来. 欧氏平面的点加上无穷远点后构成的平面称定义在实数域 **R** 上的**射影平面**. 从上面的例子可以看出，用齐次坐标可以很方便地表示射影平面上的点. 如果将满足等价关系 R 的所有齐次坐标看成是一个坐标，则齐次坐标与射影平面上的点一一对应. 由规定(3)还可以看出，射影平面上所有无穷远点构成的线——无穷远线的方程为

$$Z=0.$$

[例 5.1.1]　求欧氏平面上的点(1，2)在定义在实数域 **R** 上的射影平面上的齐次坐标.

解　点(1，2)为平常点，$Z\neq 0$，

$$x=X/Z=1,$$
$$y=Y/Z=2.$$

所以，有 $X=Z$，$Y=2Z$，齐次坐标为$(Z, 2Z, Z)$，$Z\neq 0$. 即形如$(Z, 2Z, Z)$，$Z\neq 0$ 的齐次坐标，例如(1，2，1)，(2，4，2)，(1.2，2.4，1.2)等都是欧氏平面上的点(1，2)的在射影平面上的齐次坐标.

[例 5.1.2]　求欧氏平面上的直线 $ax+by+c=0$ 在定义在实数域 **R** 上的射影平面上的齐次坐标方程.

解　将 $x=X/Z$，$y=Y/Z$ 代入 $ax+by+c=0$ 得齐次坐标方程为

$$aX+bY+cZ=0.$$

[例 5.1.3]　欧氏平面上的椭圆方程

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

化成实数域 **R** 上的射影平面上的齐次坐标方程为

$$\frac{X^2}{a^2}+\frac{Y^2}{b^2}=Z^2. \tag{5.1.1}$$

将无穷远直线 $Z=0$ 代入式(5.1.1)在实数域 **R** 上解此方程得

$$X=Y=0,$$

即 $X=Y=Z=0$，而(0，0，0)不是一个合法的齐次坐标，故椭圆与无穷远直线不相交.

与例 5.1.3 同理可证欧氏平面上的抛物线与无穷远直线相切，双曲线与无穷远直线相交于两个点.

5.1.2　判别式、结式和代数不变量

定义 5.1.4　设 K 为数域，其零元记为 0，多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0\in K[x]$$

对任意 $b\in K$，定义

$$f(b)=a_nb^n+a_{n-1}b^{n-1}+\cdots+a_0.$$

如果 $f(b)=0$，则称 b 为多项式 $f(x)$的一个**根**或**零点**.

由代数基本定理，复数域 **C** 上的一元 n 次多项式在复数域上恰好有 n 个根. 对于一般的数域，此性质不一定成立，我们给出如下定义：

定义 5.1.5　设 K 为一个数域，如果每个系数在 K 上的 n 次多项式恰有 n 个在 K 上的根，则称域 K 为**代数闭域**.

[**例** 5.1.4]　实数域 **R** 不是代数闭域，因为实系数多项式 $x^2+1=0$ 无实根，同理可证有理数域 **Q** 不是代数闭域. 此外，有限域 **F** 也不是代数闭域，设 a_1，…，a_n 是有限域 **F** 上的所有元素，则多项式 $(x-a_1)(x-a_2)\cdots(x-a_n)+1$ 在 **F** 上没有根.

对每个域 K，都存在一个代数闭域$\overline{K}$，使 K 包含在$\overline{K}$中(证明略)，此时称$\overline{K}$为 K 的代数闭域. 例如复数域 **C** 就是实数域 **R** 的代数闭域.

定义 5.1.6　设 K 为代数闭域，给定 $K[x]$上的一个 n 次多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0 \quad (a_n\neq 0).$$

设 x_1，x_2，…，x_n 是它的 n 个根，则

$$D(f)=a_n^{2n-2}\Delta(x_1,\cdots,x_n),$$

称为 $f(x)$的**判别式**. 这里 n 元式 $\Delta(x_1,\cdots,x_n)$定义为

$$\Delta(x_1,\cdots,x_n)=\prod_{1\leqslant j<i\leqslant n}(x_i-x_j)^2=\begin{vmatrix}1 & 1 & \cdots & 1\\ x_1 & x_2 & \cdots & x_n\\ \vdots & \vdots & & \vdots\\ x_1^{n-1} & x_2^{n-1} & \cdots & x_n^{n-1}\end{vmatrix}^2.$$

显然，方程 $f(x)=0$ 有重根的充分必要条件是 $D(f)=0$.

定义 5.1.7　设 m，n 为正整数，考虑代数闭域 K 上的多项式

$$f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_0,$$
$$g(x)=b_mx^m+b_{m-1}x^{m-1}+\cdots+b_0.$$

令

$$R(f,g)=\begin{vmatrix}a_n & a_{n-1} & a_{n-2} & \cdots & a_0 & 0 & \cdots & 0\\ 0 & a_n & a_{n-1} & \cdots & a_1 & a_0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \cdots & a_n & a_{n-1} & \cdots & a_0\\ b_m & b_{m-1} & b_{m-2} & \cdots & b_0 & 0 & \cdots & 0\\ 0 & b_m & b_{m-1} & \cdots & b_1 & b_0 & \cdots & 0\\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & \cdots & b_m & b_{m-1} & \cdots & b_0\end{vmatrix}\begin{matrix}\\ \\ \\ \\ \}\,m\text{ 行}\\ \\ \\ \\ \}\,n\text{ 行}\end{matrix},$$

称 $R(f,g)$为多项式 f，g 的**结式**.

例如，设 $f(x)=x^2-1$，$g(x)=x^3-4x$，则

$$R(f,g)=\begin{vmatrix}1&0&-1&0&0\\0&1&0&-1&0\\0&0&1&0&-1\\1&0&-4&0&0\\0&1&0&-4&0\end{vmatrix}=-9.$$

可以证明(证明过程略)，代数闭域 K 上的多项式 $f(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots a_1x+a_0$ $(a_n\neq0)$与其导数 $f'(x)=na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+a_1$ 的结式满足

$$R(f,f')=a_n^{n-1}\prod_{i=1}^{n}f'(x_i),$$

这里 $x_1,x_2,\cdots,x_n$ 是 $f(x)$在 K 上的 n 个根.

由定义 5.1.6，有

$$D(f)=a_n^{2n-2}\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)^2.$$

由定理 4.3.6，多项式 $f(x)$可唯一分解为

$$f(x)=a_n\prod_{i=1}^{n}(x-x_i),$$

故

$$f'(x)=a_n\sum_{i=1}^{n}(x-x_1)(x-x_2)\cdots(x-x_{i-1})(x-x_{i+1})\cdots(x-x_n).$$

以 $x=x_i$ 代入上式，得

$$f'(x_i)=a_n\prod_{\substack{j=1\\j\neq i}}^{n}(x_i-x_j),$$

于是有

$$\prod_{i=1}^{n}f'(x_i)=a_n^n\prod_{\substack{i,\,j=1\\i\neq j}}^{n}(x_i-x_j)=(-1)^{\frac{n(n-1)}{2}}a_n^n\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)^2.$$

从而

$$\begin{aligned}R(f,f')&=a_n^{n-1}\prod_{i=1}^{n}f'(x_i)\\&=(-1)^{\frac{n(n-1)}{2}}a_n^{2n-1}\prod_{1\leqslant i<j\leqslant n}(x_j-x_i)^2\\&=(-1)^{\frac{n(n-1)}{2}}a_nD(f)\end{aligned}$$

于是得到 f 的判别式与 $R(f,f')$之间的关系式(定理 5.1.1)。

定理 5.1.1　设代数闭域 K 上的 n 次多项式 $f(x)$首项系数为 a_n，则 $f(x)$的判别式满足

$$D(f)=(-1)^{\frac{n(n-1)}{2}}a_n^{-1}R(f,f')$$

这个定理指出了一种在不知道多项式根的情况下求多项式判别式的方法.

[例 5.1.5]　验证复数域 $\mathbf{C}$ 上的一元二次多项式 $f(x)=ax^2+bx+c$ 的判别式为 b^2-4ac.

解　由定理 5.1.1，$D(f)=(-1)^{\frac{2(2-1)}{2}}a^{-1}R(f,f')=-\dfrac{1}{a}\begin{vmatrix}a&b&c\\2a&b&0\\0&2a&b\end{vmatrix}=b^2-4ac.$

同理可证复数域 **C** 上的一元三次多项式 $f(x)=x^3+px+q$ 的判别式

$$D(f)=(-1)^{\frac{3(2-1)}{2}}R(f,f')=-\begin{vmatrix}1&0&p&q&0\\0&1&0&p&q\\3&0&p&0&0\\0&3&0&p&0\\0&0&3&0&p\end{vmatrix}=-4p^3-27q^3.$$

定理 5.1.2 （Vieta 定理） 设 K 是代数闭域，多项式

$$f(x)=x^n+a_{n-1}x^{n-1}+\cdots+a_1x+a_0\in K[x]$$

在 K 中的 n 个根为 x_1，x_2，…，x_n，则

$$\begin{aligned}\sum_{i=1}^{n}x_i&=-a_{n-1},\\ \sum_{1\leqslant i<j\leqslant n}x_ix_j&=a_{n-2},\\ \sum_{1\leqslant i<j<k\leqslant n}x_ix_jx_k&=-a_{n-3},\\ &\vdots\\ \sum x_1x_2\cdots x_n&=(-1)^na_0.\end{aligned}$$

可以看出，这一定理是初中学过的一元二次方程 Vieta 定理的推广，这一定理的证明可以在高等代数教材中找到.

最后简单介绍代数不变量的概念. 先看一个例子，将二次三项式

$$ax^2+bxy+cy^2$$

施以线性变换

$$\begin{aligned}x&=\alpha x'+\beta y',\\ y&=\gamma x'+\delta y',\\ \alpha\delta-\beta\gamma&=1,\end{aligned}$$

得到

$$a'x'^2+2b'x'y'+c'y'^2,$$

其中

$$\begin{aligned}a'&=a\alpha^2+b\alpha\gamma+c\gamma^2,\\ b'&=a\alpha\beta+b\frac{\alpha\delta+\beta\gamma}{2}+c\gamma\delta,\\ c'&=a\beta^2+b\beta\delta+c\delta^2.\end{aligned}$$

则判别式 $b^2-4ac(=b'^2-4a'c')$是一个**代数不变量**，简称**不变量**.

一般而言，n 个变元 x_1，x_2，…，x_n 的 m 次齐次多项式 $J(x_1,\cdots,x_n)$被称为 n 元 m 次代数形式. 设线性变换 T 将变元$(x_1,\cdots,x_n)$变为$(X_1,\cdots,X_n)$，此时多项式 $J(x_1,\cdots,x_n)$变为 $\overset{*}{J}(X_1,\cdots,X_n)$，$J$ 的系数 a_0，a_1，…，a_q 变为 $\overset{*}{J}$ 的系数 A_0，A_1，…，A_q. 若对全体线性变换 T 有 $J=\overset{*}{J}$，则称 J 为**不变式**，称在线性变换下保持不变的 J 的系数的任何函数 I 为 J 的一个**不变量**. 凯莱和西尔维斯特等人计算、构造了大量特殊的不变量，这也是 1840—1870 年间古典不变量理论研究的主要方向. 进一步的发展提出了更一般的问题——寻找不

变量的完备系，即对任意给定元数与次数的代数形式，求出最小可能个数的有理整不变量，使任何其他有理整不变量可以表成这个完备集合的具有数值系数的有理整函数. 这样的完备系亦叫代数形式的基. 1888 年，希尔伯特改变代数不变量问题的提法：给定了无限多个包含有限个变元的代数形式系，问在什么条件下存在一组有限的代数形式系，使所有其他的形式都可表成它们的线性组合？希尔伯特证明了这样的形式系是存在的，然后应用此结果于不变量而得到了不变量有限整基的存在定理. 他的不变量理论还把模、环、域的抽象理论带到了显著地位，从而引导了抽象代数学派.

5.1.3　一元三次方程的公式解 —— Cartan 公式

考虑复数域 $\mathbf{C}$ 上的一元三次方程式

$$f(x)=x^3+ax^2+bx+c, \tag{5.1.2}$$

作变换

$$x=y-\frac{1}{3}a,$$

化为缺二次项方程

$$y^3+py+q=0.$$

考虑如下方程的根

$$x^3+px+q=0 \tag{5.1.3}$$

(1) 若 $q=0$，则 $x_1=0$，$x_2=\sqrt{-p}$，$x_3=-\sqrt{-p}$ 是方程的根.

(2) 若 $p=0$，则 $x_1=\sqrt[3]{-q}$，$x_2=\sqrt[3]{-q\omega}$，$x_3=\sqrt[3]{-q\omega^2}$ 是方程的根，其中 $\omega=-\frac{1}{2}+\frac{\sqrt{3}}{2}i$.

(3) 若 $p\neq 0$，$q\neq 0$，令 $x=u+v$，得

$$x^3-3uvx-(u^3+v^3)=0. \tag{5.1.4}$$

比较方程(5.1.3)和方程(5.1.4)可得

$$\begin{cases} uv=-\dfrac{1}{3}p \\ u^3+v^3=-q \end{cases} \quad \text{即} \quad \begin{cases} u^3v^3=-\dfrac{1}{27}p^3 \\ u^3+v^3=-q \end{cases}$$

由 Vieta 定理知，u^3，v^3 是方程 $y^2+qy-\frac{p^3}{27}=0$ 的两个根. 所以

$$u^3=-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}, \quad v^3=-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}.$$

于是得到方程(5.1.4)的三个根(即其名的 Cartan 公式)

$$x_1=\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}+\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}},$$

$$x_2=\omega\left(\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}\right)+\omega^2\left(\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}\right),$$

$$x_3=\omega^2\left(\sqrt[3]{-\frac{q}{2}+\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}\right)+\omega\left(\sqrt[3]{-\frac{q}{2}-\sqrt{\frac{q^2}{4}+\frac{p^3}{27}}}\right).$$

由定义 5.1.6 知方程(5.1.3)的判别式

$$\begin{aligned}D&=(x_1-x_2)^2(x_2-x_3)^2(x_3-x_1)^2\\&=-4(x_1x_2+x_2x_3+x_3x_1)^3-27(x_1x_2x_3)^2\\&=-(4p^3+27q^2)=-108\left(\frac{q^2}{4}+\frac{p^3}{27}\right)\end{aligned}$$

记为$-(4p^3+27q^2)$为Δ，显然，

(a) $\Delta<0$ 时，方程(5.1.3)有一个实根和一对复根；

(b) $\Delta=0$ 时，方程(5.1.3)有三个实根；特别是当 $4p^3=-27q^2\neq 0$ 时，三个实根中有两个相等；$p=q=0$ 时，有三重零根；

(c) $\Delta>0$ 时，方程(5.1.3)有三个不等的实根.

5.2 椭圆曲线

椭圆曲线是由非奇异 Weierstrass 方程确定的一种曲线，学习椭圆曲线之前先简要介绍一下椭圆曲线的历史背景. 首先要明确：椭圆曲线的形状并不是椭圆. 椭圆曲线的研究源于椭圆积分，至于椭圆积分的背景知识和怎样由椭圆积分导出椭圆曲线，本书不作详细介绍，感兴趣的读者可以参考代数几何方面的教材. 本书从 Weierstrass 方程的角度直接给出椭圆曲线的定义.

5.2.1 Weierstrass 方程

定义 5.2.1 设 K 为域，a_1，a_2，a_3，a_4，$a_6\in K$，形如

$$y^2+a_1xy+a_3y=x^3+a_2x^2+a_4x+a_6 \tag{5.2.1}$$

的方程称为**域 K 上的 Weierstrass 方程**，通常记为 $E(K)$，a_1，a_2，a_3，a_4，a_6 称为 Weierstrass 方程的**系数**. 域 K 可以是实数域 **R**，也可以是复数域 **C**，还可以是有限域.

我们可以这样来记忆 Weierstrass 方程系数的下标，设 x，y，a_1，a_2，a_3，a_4，a_6 的权值如表 5.2.1 所示.

表 5.2.1 Weierstrass 方程系数权值表

x 的权值	2
y 的权值	3
a_i 的权值	i

则 Weierstrass 方程的每一项权值都是 6，这也是为什么没有系数 a_5 的原因.

如果域 K 的特征不为 2，令 $\eta=y+(a_1x+a_3)/2$，消去 y 的二次项，方程(5.2.1)可化为

$$\eta^2=x^3+\frac{b_2}{4}x^2+\frac{b_4}{2}x+\frac{b_6}{4}, \tag{5.2.2}$$

其中

$$\begin{cases}b_2=a_1^2+4a_2\\b_4=a_1a_3+2a_4\\b_6=a_3^2+4a_6\end{cases} \tag{5.2.3}$$

如果域 K 的特征不为 2、3，令 $\xi=x+b_2/12$，消去 x 的二次项，方程(5.2.2)可化为

$$\eta^2=\xi^3-\frac{c_4}{48}\xi-\frac{c_6}{864}, \tag{5.2.4}$$

其中

$$\begin{cases}c_4=b_2^2-24b_4\\ c_6=-b_2^3+36b_2b_4-216b_6\end{cases} \tag{5.2.5}$$

另定义

$$b_8=a_1^2a_6-a_1a_3a_4+4a_2a_6+a_2a_3^2-a_4^2, \tag{5.2.6}$$

$$\Delta=-b_2^2b_8-8b_4^3-27b_6^2+9b_2b_4b_6. \tag{5.2.7}$$

系数 b 和 c 的下标也可理解成它们的权值. 我们将式(5.2.1)，式(5.2.2)和式(5.2.4)分别称为 Weierstrass 方程的 a 形式、b 形式和 c 形式. 容易验证，系数定义式式(5.2.3)，式(5.2.5)，式(5.2.6)和式(5.2.7)对所有 Weierstrass 方程都成立，不管域 K 的特征是否为 2 或 3.

可以证明(留作练习)，当域 K 的特征不为 2 时，Δ 等于 b 形式的 Weierstrass 方程右边多项式的判别式的 16 倍，即

$$\Delta=16D\left(x^3+\frac{b_2}{4}x^2+\frac{b_4}{2}x+\frac{b_6}{4}\right).$$

所以当域 K 的特征不为 2 时，K 上的 Weierstrass 方程(5.2.2)有重根当且仅当 $\Delta=0$.

当 $\Delta\neq0$ 时，定义

$$j=c_4^3/\Delta, \tag{5.2.8}$$

称 j 为 Weierstrass 方程(5.2.1)的 j **不变量**，可记作 $j(E(k))$.

定义

$$\kappa=2y+a_1x+a_3, \tag{5.2.9}$$

当域 K 的特征不为 2 时，有 $\kappa=2\eta$(不恒为零). 当域 K 的特征为 2 时，有

$$\begin{aligned}\Delta\neq0&\Rightarrow b_2^2b_8+b_6^2+b_2b_4b_6\neq0\\ &\Rightarrow a_1^4(a_1^2a_6+a_1a_3a_4+a_2a_3^2+a_4^2)+a_3^4+a_1^3a_3^3\neq0\\ &\Rightarrow a_1^6a_6+a_1^5a_3a_4+a_1^4a_2a_3^2+a_1^4a_4^2+a_3^4+a_1^3a_3^3\neq0\\ &\Rightarrow a_1\text{ 和 }a_3\text{ 不能同时为 0.}\end{aligned}$$

因此 $\Delta\neq0$ 时，对任意域 K 上的 Weierstrass 方程都有 κ 不恒为零，于是有

$$\kappa^2=4x^3+b_2x^2+2b_4x+b_6.$$

通常称 c_4，c_6 和 Δ 为 Weierstrass 方程 $E(K)$的**三个基本参数**，它们的权值分别为 4、6 和 12. 例如方程 $y^2+y=x^3-x^2$ 的三个基本参数为 16，-152，-11，即 $c_4=16$，$c_6=-152$，$\Delta=11$(请读者自己验证).

另外，容易验证，等式

$$4b_8=b_2b_6-b_4^2, \tag{5.2.10}$$

$$1728\Delta=c_4^3-c_6^2, \tag{5.2.11}$$

$$j=\frac{c_4^3}{\Delta}=1\ 728+\frac{c_6^2}{\Delta} \tag{5.2.12}$$

成立(留作练习).

[**例 5.2.1**]　当域 K 的特征不为 2 和 3 时，对 Weierstrass 方程的 c 形式进行变换，令 $\eta'=6^3\eta$，$\xi'=6^2\xi$，得到

$$(\eta')^2=(\xi')^3-27\bar{c}_4\xi'-54\bar{c}_6. \tag{5.2.13}$$

注意，在参数 c_4，c_6 上加了上划线，这是因为，对于式(5.2.13)给出的 Weierstrass 方程，由式(5.2.5)可知

$$c_4=b_2^2-24b_4=0-24(2-27\bar{c}_4)=1\ 296\bar{c}_4,$$
$$c_6=-b_2^3+36b_2b_4-216b_6=0+0-216(4(-54\bar{c}_6))=46\ 656\bar{c}_6,$$

而对于式(5.2.4)给出的 Weierstrass 方程，可以验证由式(5.2.5)计算得到的参数 c_4，c_6 与式(5.2.4)中原来的参数 c_4，c_6 是一致的.

[例 5.2.2] 求下列 Weierstrass 方程

(1) $y^2=x^3+px+q$,

(2) $y^2=x^3+q$,

(3) $y^2=x^3+px$.

的三个基本参数 c_4，c_6，Δ.

解

(1) 对 $y^2=x^3+px+q$，有 $a_1=0$，$a_2=0$，$a_3=0$，$a_4=p$，$a_6=q$，

$$c_4=b_2^2-24b_4=0-24(2(p))=-48p,$$
$$c_6=-b_2^3+36b_2b_4-216b_6=0+0-216(4(q))=-864q,$$
$$\Delta=\frac{c_4^3-c_6^2}{1\ 728}=\frac{(-48p)^3-(-864q)^2}{1\ 728}=-16(4p^3+27q^2).$$

(2) 对 $y^2=x^3+q$，由(1)的结论，

$$c_4=0,\ c_6=-864q,\ \Delta=-432q^2\neq 0,$$

当 $\Delta=-432q^2\neq 0$ 时，有 $j=\frac{c_4^3}{\Delta}=0$.

(3) 对 $y^2=x^3+px$，由(1)的结论，

$$c_4=-48p,\ c_6=0,\ \Delta=-64p^3,$$

当 $\Delta=-64p^3\neq 0$ 时，有 $j=\frac{c_4^3}{\Delta}=\frac{(-48p)^3}{-64p^3}=1\ 728$.

以上用仿射坐标研究了 Weierstrass 方程. 令 $x=\frac{X}{Z}$，$y=\frac{Y}{Z}$，方程(5.2.1)可以化为

$$Y^2Z+a_1XYZ+a_3YZ^2=X^3+a_2X^2Z+a_4XZ^2+a_6Z^3, \tag{5.2.14}$$

其中 a_1，a_2，a_3，a_4，$a_6\in K$，方程(5.2.14)称为 Weierstrass 方程的**齐次坐标形式**. 同样，齐次坐标下的 Weierstrass 方程也有 a 形式(即方程(5.2.14))，b 形式和 c 形式，其 Δ，j 的定义与仿射坐标下的情况一致.

容易验证，凡是满足方程(5.2.1)的平常点$(x,\ y)$，其齐次坐标$(x,\ y,\ 1)$一定满足方程(5.2.14)，另外齐次坐标为(0，1，0)的无穷远点也满足方程(5.2.14)，且在齐次坐标等价意义下(0，1，0)是唯一满足方程(5.2.14)的无穷远点，将这个无穷远点记为 O，即

$$O=(0,\ 1,\ 0),$$

易知**无穷远点 O 在与 y 轴平行的直线上**. 由此可见，Weierstrass 方程的齐次坐标形式比仿射坐标形式多包含了一个无穷远点 O.

定义 5.2.2 设

$$F(X,\ Y,\ Z)=Y^2Z+a_1XYZ+a_3YZ^2-X^3-a_2X^2Z-a_4XZ^2-a_6Z^3,$$

如果对任意齐次坐标点(X, Y, Z)，偏导数$\frac{\partial F}{\partial X}$，$\frac{\partial F}{\partial Y}$，$\frac{\partial F}{\partial Z}$不同时为 0，则称 Weierstrass 方程(5.2.14)为**非奇异的**或**光滑的**，否则就称之为**奇异的**.

我们不加证明地给出：

定理 5.2.1　域 K 上的 Weierstrass 方程 $E(K)$为非奇异的充要条件是 $\Delta\neq 0$.

[例 5.2.3]　求证实数域 $\mathbf{R}$ 上的 Weierstrass 方程 $Y^2Z=X^3$ 为奇异的.

证明　设 $F(X, Y, Z)=Y^2Z-X^3$，则有

$$\frac{\partial F}{\partial X}=-3X^2,\ \frac{\partial F}{\partial Y}=2YZ,\ \frac{\partial F}{\partial Z}=Y^2$$

在点$(0, 0, 1)$处$\frac{\partial F}{\partial X}$，$\frac{\partial F}{\partial Y}$，$\frac{\partial F}{\partial Z}$同时为 0，所以域 $\mathbf{R}$ 上的 Weierstrass 方程 $Y^2Z=X^3$ 是奇异的. 也可以用定理 5.2.1 来证明：方程 $Y^2Z=X^3$ 的系数 a_1，a_2，a_3，a_4，a_6 全为 0，所以 $\Delta=0$，因此它是奇异的.

5.2.2　椭圆曲线

定义 5.2.3　设 K 为域，0为 K 的零元，$\overline{K}$为 K 的代数闭域，K 上的 Weierstrass 方程

$$Y^2Z+a_1XYZ+a_3YZ^2=X^3+a_2X^2Z+a_4XZ^2+a_6Z^3$$

为非奇异的(即满足 $\Delta\neq 0$)，M 为$\overline{K}$上的射影平面，则 M 上满足方程(5.2.14)的所有点(X, Y, Z)组成的集合称为**域 K 上的椭圆曲线**，记为 $E(\overline{K})$，即

$$M=\{(X, Y, Z)\mid X, Y, Z\in\overline{K}\}\backslash\{0, 0, 0\},$$

$$E(\overline{K})=\{(X, Y, Z)\in M\mid Y^2Z+a_1XYZ+a_3YZ^2=X^3+a_2X^2Z+a_4XZ^2+a_6Z^3\}.$$

定义 5.2.3 用 Weierstrass 方程的齐次坐标形式给出了椭圆曲线的定义，我们知道，Weierstrass 方程的齐次坐标形式只比 Weierstrass 方程的仿射坐标形式多包含一个无穷远点 O，因此还可以用 Weierstrass 方程的仿射坐标形式给出椭圆曲线的定义.

定义 5.2.4　设 K 为域，$\overline{0}$为 K 的零元，$\overline{K}$为 K 的代数闭域，K 上的 Weierstrass 方程

$$y^2+a_1xy+a_3y=x^3+a_2x^2+a_4x+a_6$$

为非奇异的(即满足 $\Delta\neq 0$)，则仿射平面$\overline{K}^2$ 上满足方程(5.2.14)的所有点(x, y)加上无穷远点 $O=(0, 1, 0)$组成的集合称为**域 K 上的椭圆曲线**，记为 $E(\overline{K})$，即

$$E(\overline{K})=\{(x, y)\in\overline{K}^2\mid y^2+a_1xy+a_3y=x^3+a_2x^2+a_4x+a_6\}\cup\{O\}.$$

5.1 节中已经指出，平常点的仿射坐标(x, y)与齐次坐标(X, Y, Z)(等价齐次坐标看成一个坐标)存在一一对应关系，因此椭圆曲线的两个定义 5.2.3 和定义 5.2.4 本质上是一致的. 为了方便，以后讨论椭圆曲线时主要采用定义 5.2.4 的形式.

注意，在上述定义中，$E(\overline{K})$是一系列点的集合，这些点的坐标值在$\overline{K}$内，这就意味着 $E(\overline{K})$是相应 Weierstrass 方程在$\overline{K}$(而不是 K)内的解集.

定义 5.2.5　设域 F 是域 K 的扩域，$E(\overline{K})$是域 K 上的椭圆曲线，点 $P\in E(\overline{K})$，如果 P 的所有坐标值都在 F 内或 P 为无穷点 O，则称点 P 为 $E(\overline{K})$上**关于 F 的有理点**. $E(\overline{K})$上所有关于 F 的有理点组成的集合记为 $E(F)$. 以后提到“椭圆曲线上的点”都是指椭圆曲线关于某一域的有理点.

注意，$E(\overline{K})$上关于域 F 的有理点的坐标值不一定是有理数，除非集合 F 为有理数集.

[例 5.2.4] 设实数域 **R** 上的 Weierstrass 方程 $E(\mathbf{R})$为 $Y^2Z=X^3-XZ^2$，求证 $E(\mathbf{R})$可以确定一条椭圆曲线，并画出其图像.

解 由 $E(\mathbf{R})$的定义可知 $a_1=0$，$a_2=0$，$a_3=0$，$a_6=0$，$a_4=-1$，

$$\Delta=-b_2^2b_8-8b_4^3-27b_6^2+9b_2b_4b_6=0-8(2(-1))^3-0+0=64>0,$$

所以 $E(\mathbf{R})$可以确定一条椭圆曲线. 将 $E(\mathbf{R})$上所有关于 **R** 的有理点画在平面直角坐标系内，可得到 $E(\mathbf{R})$的图像(注意图像上没有表示出无穷远点)，如图 5.2.1 所示.

同理可得到实数域 **R** 上的椭圆曲线 $Y^2Z=X^3+XZ^2+Z^3$(此椭圆曲线的判别式 $\Delta=-b_2^2b_8-8b_4^3-27b_6^2+9b_2b_4b_6=0-8(2(1))^3-27(4)^2+0<0$)的图像，如图 5.2.2 所示.

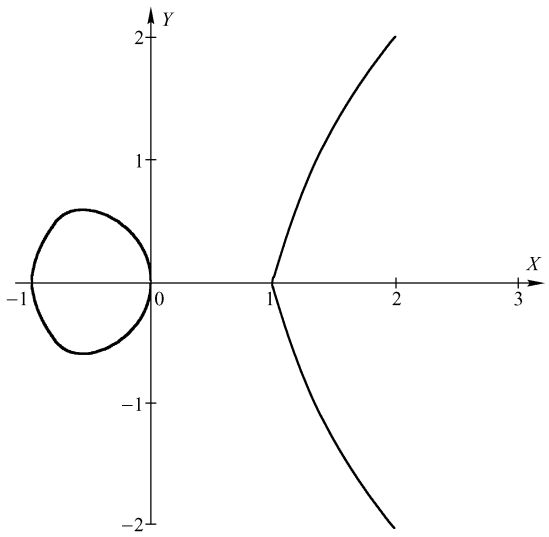

图 5.2.1 $Y^2Z=X^3-XZ^2$ 的图像

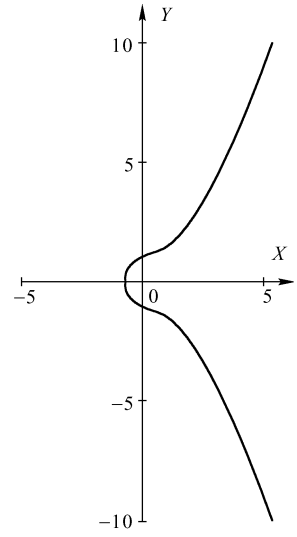

图 5.2.2 $Y^2Z=X^3+XZ^2+Z^3$ 的图像

注意到图 5.2.1 中的椭圆曲线与 X 轴有三个交点，其 $\Delta>0$；而图 5.2.2 中的椭圆曲线与 X 轴有一个交点，其 $\Delta<0$. 请读者思考：这一结论是否对所有实数域 **R** 上的椭圆曲线都成立？

最后，给出实数域 **R** 上几种形如 $y^2=x^3+ax+b$ 的椭圆曲线的图像，供读者参考，如图 5.2.3 所示.

5.2.3 椭圆曲线上点的加法群(Mordell-Weil 群)

在 5.2.2 节中看到了椭圆曲线的图像，这种图像上点与点之间究竟有什么联系呢？事实上，通过巧妙的定义，椭圆曲线上的点对特定的运算——我们称之为“加法”，可以构成 Abel群. 这就意味着，椭圆曲线上点与点之间有着非常深刻的内在联系. 正是因为椭圆曲线具有这一优良特性，才使它在密码学领域中有广泛的应用. 下面给出椭圆曲线上点的加法定义.

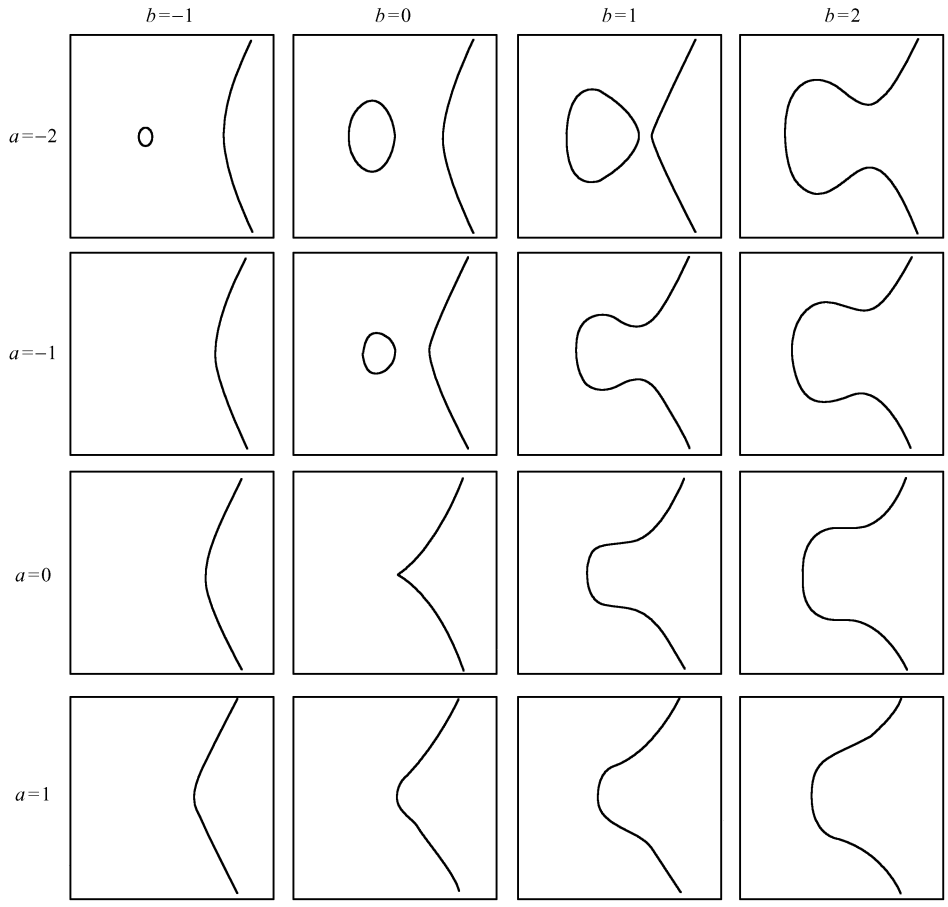

图 5.2.3　实数域 **R** 上几种形如 $y^2=x^3+ax+b$ 的椭圆曲线的图像
（注意：a 和 b 同时为 0 时，方程 $y^2=x^3$ 是奇异的，故不是椭圆曲线）

Bezout 曾经证明，设 E 是域 K 上的 Weierstrass 方程，点 P 和 Q 是 E 上的点，P 和 Q 的坐标值都在域 K 内，则过点 P 和 Q 的直线与 E 一定交于一点 R，R 的坐标值也在域 K 内，记为 $R=PQ$. 如图 5.2.4 所示，如果 $P=Q$，则 R 为 E 在 P 点的切线与 E 的另一个交点，如果过 P 和 Q 的直线在 Q 点与 E 相切，则 $R=PQ=Q$. 这就是代数几何里著名的 Bezout 定理.

定义 5.2.6　设 P，Q 是椭圆曲线 E 上的两点，$R=PQ$，O 为无穷远点. 如图 5.2.4 所示，点 P 与 Q 的加法定义为

$$P+Q=O(PQ)=OR, \tag{5.2.15}$$

即 R 是过 P 和 Q 的直线与 E 的第三个交点，$P+Q$ 是过 O 与 R 的直线与 E 的第三个交点，图 5.2.3描述了这种加法的几何意义，图 5.2.3(a)是 $P\neq Q$ 的情况，图 5.2.3(b)是 $P=Q$ 的情况. 容易验证，过点 O 和 R 的直线与 Y 轴平行.

注意：椭圆曲线上点 P 与 Q 的“加法”不是 P 与 Q 的相应坐标值之间的加法，要将椭圆曲线上点的加法与解析几何里的向量加法严格区分开来.

由于在平面解析几何的范围内考虑椭圆曲线问题，因此无穷远点在上面的图中表示不出来，便理解椭圆曲线上的加法定义时要默认它们存在. 比如定义中平行于 Y 轴的直线 OR 和

椭圆曲线的交点应该有三个，第三个点就在无穷远点. 这从另外一个方面说明，在定义椭圆曲线时“人为”引入的那个无穷远点是实际存在的，在引入齐次坐标和射影平面后，这个无穷远点就自然出现了.

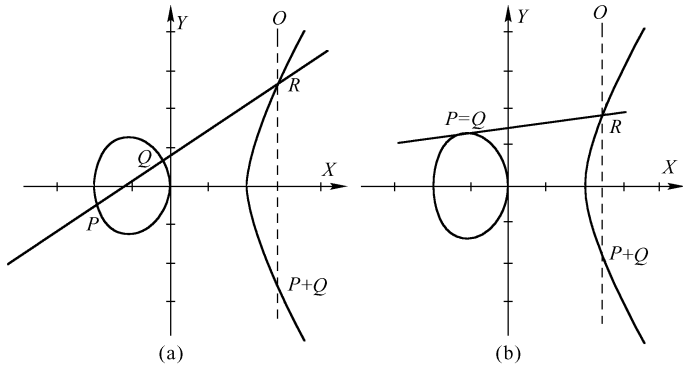

图 5.2.4 椭圆曲线上点的加法

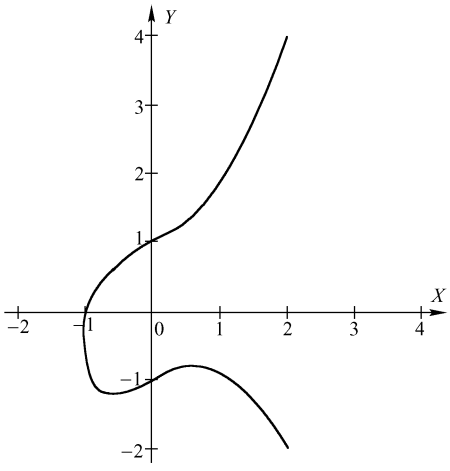

图 5.2.5 椭圆曲线 $y^2-xy=x^3+1$ 的图像

还有一点需要注意的是，如果 $E(\overline{K})$是定义在域 K 上的椭圆曲线，F 是 K 的扩域，P，Q 是 $E(\overline{K})$上的关于 F 的有理点，则过 P，Q 的直线与 $E(\overline{K})$的第三个交点一定是 $E(\overline{K})$上的关于 F 的有理点. 这是因为直线和曲线的联立方程的系数都取自域 F，而方程组有 3 个解，其中两个解就是 P 和 Q，第三个解必为关于 F 的有理点.

需要特别指出的是，点 $P+Q$ 与点 R 不一定关于 X 轴对称，原因是椭圆曲线的图像不一定关于 X 轴对称，例如 **R** 上椭圆曲线 $y^2-xy=x^3+1$ (如图 5.2.5所示).

定理 5.2.2 椭圆曲线 E 上点的加法具有如下性质：

(1) 若 P 和 Q 是 E 上任意两点，P 与 Q 的连线 L 交 E 于另一点 R，则：

(2) 对任意 $P\in E$ 有 $P+O=P$；

(3) 对任意 P，$Q\in E$ 有 $P+Q=Q+P$；

(4) 对任意 $P\in E$，E 上存在一点 $-P$，使得 $P+(-P)=O$；

(5) 对于 P 上的任意点 P，Q，R，有$(P+Q)+R=P+(Q+R)$.

证明

(1) 因为椭圆曲线与无穷远线只有一个交点 O，所以在椭圆曲线上有 $OO=O$ 成立，又由条件知 $R=PQ$，于是$(P+Q)+R=OR+R=O((OR)R)=OO=O$. 图 5.2.6 给出的 $P+Q+R=O$ 的几种不同情况.

(2) $P+O=O(PO)=P$.

(3) $P+Q=O(PQ)=O(QP)=Q+P$.

(4) 设 $P'=PO$，则 $PP'=P(PO)=O$，$O(PP')=OO=O$，由式(5.2.15)知 $P+P'=O$，于是 $-P=P'=PO$.

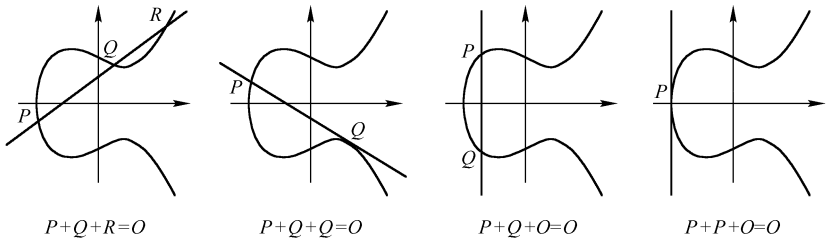

图 5.2.6　椭圆曲线上 $P+Q+R=O$ 的几种不同情况

(5) 可通过各种情况进行验证，过程比较繁琐，略去. 感兴趣的读者请参阅 Michigan 大学 J. S. Milne 著的 *Elliptic Curves*.

从定理 5.2.2 可知，椭圆曲线 E 上的点对式(5.2.15)定义的加法满足结合律和交换律，任意点 $P\in E$ 都有负元 $-P$ 存在，且可以将 O 点看作零元，于是椭圆曲线上的点对式(5.2.15)定义的加法构成 Abel 群. 为了简化起见，可将 $P+P$ 记为 $2P$，依此类推，

$$mP=P+P+\cdots+P(\text{共 } m \text{ 个 } P),$$

称为椭圆曲线上点 P 的 m **倍加**. 与此对应，当 $P\neq Q$ 时，$P+Q$ 称为椭圆曲线上点的**普加**或**点加**.

例如，图 5.2.7 给出了椭圆曲线 $y^2=x^3-x$ 上一点 P 的 3 倍加的几何意义.

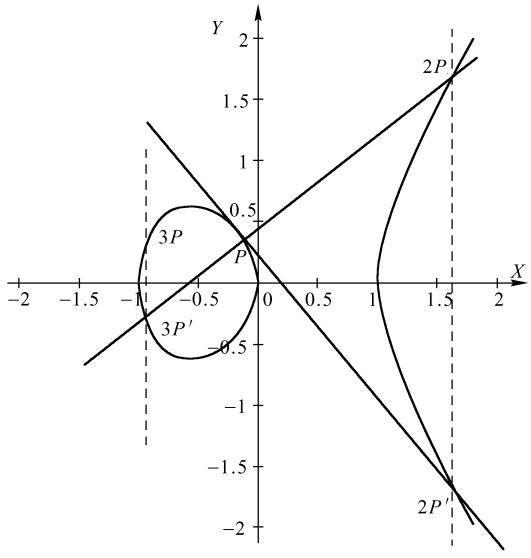

图 5.2.7　椭圆曲线上点 P 的倍加

下面给出椭圆曲线上点加运算的代数描述.

定理 5.2.3　设椭圆曲线 E 的一般 Weierstrass 方程为

$$y^2+a_1xy+a_3y=x^3+a_2x^2+a_4x+a_6 \quad (\Delta\neq 0).$$

定义 E 为

$$E=\{(x,\ y)\mid y^2+a_1xy+a_3y=x^3+a_2x^2+a_4x+a_6\}\cup\{O\} \tag{5.2.16}$$

设 $P_1=(x_1,\ y_1)$，$P_2=(x_2,\ y_2)$是 E 上的异于无穷远点 O 的两个点，则有

$$-P_1=(x_1,\ -y_1-a_1x_1-a_3), \tag{5.2.17}$$

如果 $P_2=-P_1$，则 $P_1+P_2=O$；如果 $P_2\neq-P_1$，设 $P_3=(x_3,\ y_3)=P_1+P_2$，则 x_3，y_3 可由下式给出

$$\begin{cases}x_3=k^2+a_1k-a_2-x_1-x_2\\ y_3=k(x_1-x_3)-a_1x_3-y_1-a_3\end{cases} \tag{5.2.18}$$

其中参数 k 定义为

$$k=\begin{cases}\dfrac{y_2-y_1}{x_2-x_1}, & \text{如果 } x_1\neq x_2\text{(对应于点加)}\\ \dfrac{3x_1^2+2a_2x_1+a_4-a_1y_1}{2y_1+a_1x_1+a_3}, & \text{如果 } x_1=x_2\text{(对应于倍加)}\end{cases} \tag{5.2.19}$$

证明

(1) 设 $P_1=(x_1,\ y_1)\neq O$，则其负元 $-P_1=(x_1',\ y_1')$为过 P_1 和 O 直线与 E 的第三个交点，显然 $x_1'=x_1$，过 P_1 和 O 直线方程为

$$x=x_1.$$

代入方程(5.2.16)得

$$y^2+(a_1x_1+a_3)y-(x_1^3+a_2x_1^2+a_4x_1+a_6)=0.$$

由 Vieta 定理知

$$y_1+y_1'=-(a_1x_1+a_3),$$
$$y_1'=-y_1-a_1x_1-a_3.$$

于是证明了式(5.2.17)

$$-P_1=(x_1,\ -y_1-a_1x_1-a_3).$$

(2) 如果 $P_2=-P_1$，则 $P_1+P_2=O$. 如果 $P_2\neq-P_1$，我们来求 $P_3=(x_3,\ y_3)=P_1+P_2$，设 $P_3'=(x_3',\ y_3')=P_1P_2$，即 P_3' 是过 P_1 和 P_2 的直线与 E 的第三个交点，显然 $x_3=x_3'$，考虑过 P_1 和 P_2 的直线 L：

$$y=kx+t,$$

当 $x_1\neq x_2$ 时，则直线 L 的斜率 k 为

$$k=\frac{y_2-y_1}{x_2-x_1},$$

当 $x_1=x_2$ 时，则直线 L 的斜率 k 为(可以用偏导数来计算，具体过程从略)

$$k=\frac{3x_1^2+2a_2x_1+a_4-a_1y_1}{2y_1+a_1x_1+a_3},$$

将 $y=kx+t$ 代入方程(5.2.16)得

$$x^3-(k^2+a_1k-a_2)x^2+(a_4-2kt-a_1t-a_3k)x+a_6-t^2-a_3t=0.$$

由 Vieta 定理知

$$x_1+x_2+x_3=k^2+a_1k-a_2,$$

即

$$x_3=k^2+a_1k-a_2-x_1-x_2,$$

代入方程 $y=kx+t$ 得

$$y'_3=-(kx_3+t)-a_1x_3-a_3.$$

根据式(5.2.17)

$$\begin{aligned}y_3&=-y'_3-a_1x_3-a_3\\&=-(kx_3+t)-a_1x_3-a_3\\&=-(kx_3+y_1-kx_1)-a_1x_3-a_3\\&=k(x_1-x_3)-a_1x_3-y_1-a_3,\end{aligned}$$

于是证明了式(5.2.18)

$$\begin{cases}x_3=k^2+a_1k-a_2-x_1-x_2\\y_3=k(x_1-x_3)-a_1x_3-y_1-a_3\end{cases}.$$

[例 5.2.5]　椭圆曲线 E：$y^2+xy=x^3+a_2x^2+a_6$($\Delta\neq0$，a_2，$a_6\in K$，K 的特征为 2)上点的加法. 设 $P_1=(x_1,\ y_1)$，$P_2=(x_2,\ y_2)$是 E 上的异于无穷远点 O 的两个点，则有

$$-P_1=(x_1,\ y_1+x_1),$$

如果 $P_2=-P_1$，则 $P_1+P_2=O$；如果 $P_2\neq-P_1$，设 $P_3=(x_3,\ y_3)=P_1+P_2$，则当 $x_1\neq x_2$ 时，x_3，y_3 可由下式给出

$$\begin{cases}x_3=\left(\dfrac{y_2+y_1}{x_2+x_1}\right)^2+\dfrac{y_2+y_1}{x_2+x_1}+a_2+x_1+x_2\\y_3=\dfrac{(y_2+y_1)(x_1+x_3)}{x_2+x_1}+x_3+y_1\end{cases},$$

当 $x_1=x_2$ 时，x_3，y_3 可由下式给出

$$\begin{cases}x_3=x_1^2+\dfrac{a_6}{x_1^2}\\y_3=x_1^2+\left(x_1+\dfrac{y_1}{x_1}\right)x_3+x_3\end{cases}$$

需要指出的是，本节定义的椭圆曲线上点的加法群称为 Mordell-Weil 群. 其实发现这个群结构的时间比 Mordell 和 Weil 要早，可以追溯到 Poincare 甚至 Abel. 这个群之所以被称为 Mordel-Weil 群，是因为 Mordell 最早给出了有理域上椭圆曲线 Mordell-Weil 群上有限生成定理，而 Weil 推广了他的结果.

对于密码工作者来说，用到的域一般都是有限域(在某些情况下也会涉及一些 p-adic 域和复数域). 有限域上的椭圆曲线和 p-adic 域上的椭圆曲线没有直观的图形表示，复数域上的椭圆曲线的图形是个三维空间中的环面. 在这些情况下思考问题会缺乏直观性，此时最好借用实数域上椭圆曲线的图形来支持我们的直觉.

5.2.4　有限域上的椭圆曲线

5.2.3 节中给出了任意域上椭圆曲线的加法运算公式，在现代密码体制中最常用的椭圆曲线是有限域上的椭圆曲线，特别是 $\mathbf{Z}_p$($p>3$ 且为素数)和 $\mathbf{Z}_{2^m}$($m\geqslant1$)上的椭圆曲线，它们在 ECC 密码体制中有着广泛的应用，下面重点介绍这两类椭圆曲线.

定义 5.2.7　设 $\mathbf{Z}_p$($p>3$ 且为素数)为特征大于 3 的素域，则 $\mathbf{Z}_p$ 上椭圆曲线 $E(\mathbf{Z}_p)$有类似于式(5.2.4)的形式，为简单起见，在密码学实践中，将 $E(\mathbf{Z}_p)$的 Weierstrass 方程定义为

$$E(\mathbf{Z}_p):\ y^2=x^3+ax+b,\tag{5.2.20}$$

其中 $\Delta=-16(4a^3+27b^2)\neq 0$. 有些文献将特征大于 3 的素域 $\mathbf{Z}_p$ 上的椭圆曲线方程记为

$$E:\ y^2\equiv x^3+ax+b\pmod p,\tag{5.2.20'}$$

简记为 $E_p(a,b)$.

将方程(5.2.20)与标准 Weierstrass 方程(5.2.1)比较，得到

$$a_1=a_2=a_3=0,\ a_4=a,\ a_6=b,$$

由此可导出椭圆曲线 $E(\mathbf{Z}_p)$上点的负元公式和加法公式.

设 $P_1=(x_1,y_1)$，$P_2=(x_2,y_2)$是 $E(\mathbf{Z}_p)$上的异于无穷远点 O 的两个点，由式(5.2.17)可得

$$-P_1=(x_1,\ -y_1-a_1x_1-a_3)=(x_1,\ -y_1),\tag{5.2.21}$$

如果 $P_2=-P_1$，则 $P_1+P_2=O$；如果 $P_2\neq -P_1$，设 $P_3=(x_3,y_3)=P_1+P_2$，由式(5.2.18)可得

$$\begin{cases}x_3=k^2+a_1k-a_2-x_1-x_2=k^2-x_1-x_2\\ y_3=k(x_1-x_3)-a_1x_3-y_1-a_3=k(x_1-x_3)-y_1\end{cases}\tag{5.2.22}$$

其中参数 k 定义为

$$k=\begin{cases}\dfrac{y_2-y_1}{x_2-x_1}, & \text{如果 } x_1\neq x_2\text{(对应于点加)}\\ \dfrac{3x_1^2+2a_2x_1+a_4-a_1y_1}{2y_1+a_1x_1+a_3}=\dfrac{3x_1^2+a}{2y_1}, & \text{如果 } x_1=x_2\text{(对应于倍加)}\end{cases}\tag{5.2.23}$$

[例 5.2.6] 求 $\mathbf{Z}_5$ 上椭圆曲线

$$E:\ y^2\equiv x^3+2x+3\pmod 5$$

上的所有点(即所有关于 $\mathbf{Z}_5$ 的有理点)，并在其中选两个点计算它们的加法.

解 $x\pmod 5$的所有可能值为 0，1，2，3，4. 将它们代入 E 的方程可以计算出相应 y 的值，从而得到 E 上点的坐标，计算过程如表 5.2.2 所示.

表 5.2.2 例 5.2.6 计算过程

$x\equiv 0$	$y^2\equiv 3\pmod 5$	y 无解
$x\equiv 1$	$y^2\equiv 6\equiv 1\pmod 5$	$y\equiv 1,4\pmod 5$
$x\equiv 2$	$y^2\equiv 15\equiv 0\pmod 5$	$y\equiv 0\pmod 5$
$x\equiv 3$	$y^2\equiv 36\equiv 1\pmod 5$	$y\equiv 1,4\pmod 5$
$x\equiv 4$	$y^2\equiv 75\equiv 0\pmod 5$	$y\equiv 0\pmod 5$

所以椭圆曲线 $E:\ y^2\equiv x^3+2x+3\pmod 5$上所有点为：

$$(1,1),(1,4),(2,0),(3,1),(3,4),(4,0),O.$$

下面在 $E:\ y^2\equiv x^3+2x+3\pmod 5$上计算$(1,4)+(3,1)$.

$$k\equiv\frac{y_2-y_1}{x_2-x_1}\equiv\frac{1-4}{3-1}\equiv 1\pmod 5,$$

$$\begin{cases}x_3\equiv k^2-x_1-x_2\equiv 1^2-1-3\equiv 2\pmod 5\\ y_3\equiv k(x_1-x_3)-y_1\equiv 1\ (1-2)-4\equiv 0\pmod 5\end{cases}$$

所以

$$(1,4)+(3,1)=(2,0).$$

定义 5.2.8　椭圆曲线 $E(K)$上点的个数称为**椭圆曲线 $E(K)$的阶**，记为 $\sharp E(K)$. 设 P 是椭圆曲线 $E(K)$上的一点，若存在最小的整数 n，使得 $nP=O$，则称 n 是**点 P 的阶**，记作

$$\text{ord } P=n.$$

如果椭圆曲线 $E(K)$上点 P 的阶存在，由群论中的拉格朗日定理知 ord $P \mid \sharp E(K)$.

事实上，有限域上的椭圆曲线上所有的点 P 的阶都是存在的(证明从略).

由例 5.2.6 可知椭圆曲线 E：$y^2 \equiv x^3+2x+3 \pmod 5$的阶为 7. 一般地，对于椭圆曲线 $E(\mathbf{Z}_p)$，因为 $\mathbf{Z}_p$ 中共有 p 个元素，根据 Weierstrass 方程 $y^2=x^3+ax+b$ 可知 $\sharp E(K) \leqslant 2p+1$(加 1 是因为有无穷远点).

更进一步，我们有 Hass 定理(证明略).

定理 5.2.4 (Hass 定理)　设 E 是定义在 $\mathbf{Z}_q$($q=p^n$，p 是素数)上的椭圆曲线，则 E 的阶 $\sharp E(\mathbf{Z}_p)$满足

$$\left| \sharp E(\mathbf{Z}_q)-(q+1) \right| \leqslant 2\sqrt{q} \tag{5.2.24}$$

由 Hass 定理知，当 q 足够大时，$\sharp E(\mathbf{Z}_p)$近似等于 q.

[例 5.2.7]　求椭圆曲线 $E_{23}(1, 1)$的上所有的点，并画出其图像.

解　$E_{23}(1, 1)$的方程为 $y^2=x^3+x+1 \pmod{23}$，表 5.2.3 列出了曲线上的所有点.

表 5.2.3　例 5.2.7 的所有点

(0, 1)	(6, 4)	(12, 19)
(0, 22)	(6, 19)	(13, 7)
(1, 7)	(7, 11)	(13, 16)
(1, 16)	(7, 12)	(17, 3)
(3, 10)	(9, 7)	(17, 20)
(3, 13)	(9, 16)	(18, 3)
(4, 0)	(11, 3)	(18, 20)
(5, 4)	(11, 20)	(19, 5)
(5, 19)	(12, 4)	(19, 18)

将表 5.2.3 中的点画在平面直角坐标系中，得到椭圆曲线 $E_{23}(1, 1)$的图像，如图 5.2.8 所示.

从图上可以看出，有限域上的椭圆曲线并不是一条连续的“曲线”，而是一系列离散的点，于是，前面定义椭圆曲线上点的加法时用到的“切线”、“斜率”等概念就不像对实数域上的椭圆曲线那样有明确的几何意义，只能从抽象的角度来理解.

从这个例题还可以看出，椭圆曲线 $E_{23}(1, 1)$上共有 28 个点(加上无穷远点 O)，因此有

$$\sharp E_{23}(1, 1)=28.$$

对于 p 较大时，求椭圆曲线 $E_p(a, b)$的阶是一件计算量很大的事情，如何快速有效地计算一条随机选取的椭圆曲线的阶是椭圆曲线理论中的一个有重要实用意义的问题，感兴趣的读者可参考有关资料.

[例 5.2.8]　已知 $P=(1, 3)$为 $\mathbf{Z}_{2\,773}$上椭圆曲线

$$E：y^2 \equiv x^3+4x+4 \pmod{2\,773}$$

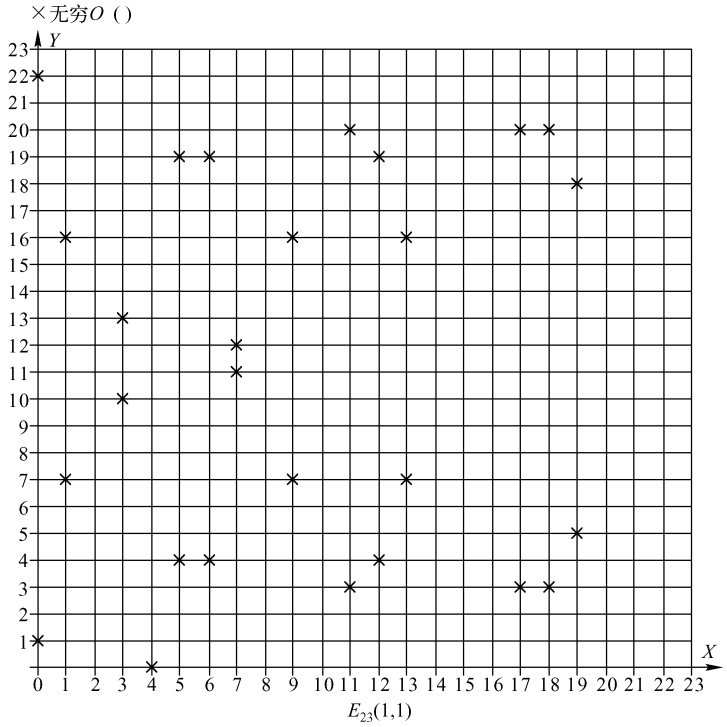

图 5.2.8 椭圆曲线 $E_{23}(1,1)$的图像

上的点，求 $2P$.

解 由广义欧几里得除法可知 $2\ 311\times 6\equiv 1 \pmod{2\ 773}$，所以

$$k\equiv\frac{3x_1^2+a}{2y_1}\equiv\frac{3\cdot 1^2+4}{2\cdot 3}\equiv\frac{7}{6}\equiv 7\times 2\ 311\equiv 2\ 312 \pmod{2\ 773},$$

$$\begin{cases} x_3\equiv k^2-x_1-x_2\equiv 2\ 312^2-1-1\equiv 1\ 771 \pmod{2\ 773} \\ y_3\equiv k(x_1-x_3)-y_1\equiv 2\ 312(1-1\ 771)-3\equiv 705 \pmod{2\ 773} \end{cases}$$

所以有

$$2P=P+P=(1\ 771,\ 705).$$

设域 K 的特征为 2，$E(K)$是定义在 K 上的椭圆曲线

$$E(K):\ y^2+\bar{a}_1xy+\bar{a}_3=x^3+\bar{a}_2x^2+\bar{a}_4x+\bar{a}_6,$$

由式(5.2.8)知

$$j(E(K))=(\bar{a}_1)^{12}/\Delta.$$

当 $j(E(K))\neq 0$，即 $\bar{a}_1\neq 0$ 时，作坐标变换

$$\begin{cases} x=(\bar{a}_1)^2x+\dfrac{\bar{a}_3}{\bar{a}_1} \\ y=(\bar{a}_1)^3y+\dfrac{(\bar{a}_1)^2\bar{a}_4+(\bar{a}_1)^2}{(\bar{a}_1)^3} \end{cases}$$

$E(K)$可化为

$$y^2+xy=x^3+ax^2+b.$$

定义 5.2.9 在密码学实践中，$\mathbf{Z}_{2^m}$ 上椭圆曲线 $E(\mathbf{Z}_{2^m})$的 Weierstrass 方程定义为

$$E(\mathbf{Z}_{2^m}):\ y^2+xy=x^3+ax^2+b, \tag{5.2.25}$$

其中 $b\neq 0$，这是因为式(5.2.25)给出的 Weierstrass 方程 $\Delta=b$(请读者自己验证). 特别地，当 $a=0$，$b=1$ 或 $a=1$，$b=1$ 时，式(5.2.25)称为 **Koblitz 曲线**，常记作 K－m，此类曲线在实现椭圆曲线密码体制时速度较快.

与 $E(\mathbf{Z}_p)$上类似，可以导出 $E(\mathbf{Z}_{2^m})$上点的负元公式和加法公式.

设 $P_1=(x_1,\ y_1)$，$P_2=(x_2,\ y_2)$是 $E(\mathbf{Z}_{2^m})$上的异于无穷远点 O 的两个点，则

$$-P_1=(x_1,\ -y_1-x_1)=(x_1,\ x_1+y_1), \tag{5.2.26}$$

如果 $P_2=-P_1$，则 $P_1+P_2=O$；如果 $P_2\neq -P_1$，设 $P_3=(x_3,\ y_3)=P_1+P_2$，则

$$\begin{cases} x_3=k^2+a_1k-a_2-x_1-x_2=k^2+k+a_2+x_1+x_2 \\ y_3=k(x_1-x_3)-a_1x_3-y_1-a_3=k(x_1+x_3)+x_3+y_1 \end{cases} \tag{5.2.27}$$

其中参数 k 定义为

$$k=\begin{cases} \dfrac{y_2-y_1}{x_2-x_1}=\dfrac{y_2+y_1}{x_2+x_1}, & \text{如果 } x_1\neq x_2\text{(对应于点加)} \\ \dfrac{3x_1^2+2a_2x_1+a_4-a_1y_1}{2y_1+a_1x_1+a_3}=\dfrac{x_1^2+y_1}{x_1}, & \text{如果 } x_1=x_2\text{(对应于倍加)} \end{cases} \tag{5.2.28}$$

[例 5.2.9]　以 $f(t)=t^2+t+1$ 为本原多项式(有关本原多项式的知识，参见 6.2 节)，可以得到有限域 $\mathbf{Z}_{2^2}$ 的一种表示形式，即

$$\mathbf{Z}_{2^2}=\{0,\ 1,\ t,\ 1+t\}(\bmod\ f(t)),$$

据此求椭圆曲线 $E(\mathbf{Z}_{2^2})$：$y^2+xy=x^3+ax^2+t$ 上的所有点.

解　$x\equiv 0\ (\bmod\ f(t))\Rightarrow y^2\equiv t\ (\bmod\ f(t))\Rightarrow y\equiv 1+t\ (\bmod\ f(t))$，

$x\equiv 1\ (\bmod\ f(t))\Rightarrow y^2+y\equiv 1+t\ (\bmod\ f(t))\Rightarrow y$ 无解，

$x\equiv t\ (\bmod\ f(t))\Rightarrow y^2+ty\equiv t^3+t\equiv t+1\ (\bmod\ f(t))\Rightarrow y\equiv 1,\ 1+t\ (\bmod\ f(t))$，

$x\equiv t+1\ (\bmod\ f(t))\Rightarrow y^2+(t+1)y\equiv t^3+t\equiv 1\ (\bmod\ f(t))\Rightarrow y$ 无解，

所以 $E(\mathbf{Z}_{2^2})$：$y^2+xy=x^3+ax^2+t$ 上的所有点为

$$(0,\ 1+t),\ (t,\ 1),\ (t,\ 1+t),\ O.$$

5.3　离散对数初步

本节在代数系统的背景上定义离散对数，并介绍密码学中最常用的两种离散对数体系.

5.3.1　有限域上的离散对数

定义 5.3.1　设群 H 是一个 n 阶有限乘法群，a，$y\in H$. 则定义**离散对数问题(DLP)**为：求解整数 x，使得 $y=a^x$，则 x 称为 H 上以 a 为底的 y 的**离散对数**，记为

$$x=L_a(y).$$

显然，如果 H 是一个 n 阶有限循环乘法群，a 是它的生成元，则任意 $y\in H$ 以 a 为底的离散对数 x 一定存在，且 $0\leqslant x\leqslant n-1$.

在密码学实践中，最常用的离散对数有两大类：基于有限域的离散对数和基于椭圆曲线的离散对数，本节重点讨论这两类离散对数.

设 p 为素数，如果 a 是 p 的原根，即 a 是有限域 $\mathbf{GF}(p)$的一个生成元，则

$$y\equiv a^x(\bmod\ p)(0\leqslant x\leqslant p-1) \tag{5.3.1}$$

遍历 $\mathbf{GF}(p)$中的所有元素，对任意的整数 y，可以找到唯一的整数 x 满足式(5.3.1). 此时 x

称为有限域 **GF**(p)上以 a 为底的 y 的离散对数.

[例 5.3.1] 2 是素数 13 的一个原根，则素域 **GF**(13)内关于 $y=1, 2, \cdots, 12$ 的离散对数表如表 5.3.1 所示.

表 5.3.1 例 5.3.1 离散对数表

y	1	2	3	4	5	6	7	8	9	10	11	12
$L_a(y)$	12	1	4	2	9	5	11	3	8	10	7	6

定理 5.3.1 设 p 为素数，a 是 p 的原根，有限域 **GF**(p)上的离散对数具有如下性质：

(1) $L_a(a)=1$,

(2) $L_a(1)=0$,

(3) 对任意 $b_1, b_2 \in \mathbf{Z}$, $L_a(b_1 b_2) \equiv L_a(b_1)+L_a(b_2) \pmod{p-1}$,

(4) 对任意 $b, r \in \mathbf{Z}$, $L_a(b^r) \equiv rL_a(b) \pmod{p-1}$.

这些性质可以很容易地利用初等数论的知识证明，留作练习.

设 p 为素数，a 是 p 的原根，已知 a，x，p 计算模指数 $a^x \pmod p$ 并不困难，将 x 化成二进制形式，即

$$x=b_0+b_1 2+b_2 2^2+\cdots+b_{n-1}2^{n-1} \quad (b_i=0, 1),$$

则

$$a^x \equiv a^{b_0+b_1 2+b_2 2^2+\cdots+b_{n-1}2^{n-1}} \equiv a^{b_0}(a^2)^{b_1}(a^{2^2})^{b_2}\cdots(a^{2^{n-1}})^{b_{n-1}} \pmod p,$$

而且有

$$(a^{2^i})^{b_i}=\begin{cases}1, & (b_i=0)\\ a^{2^i}, & (b_i=1)\end{cases}$$

这样就可以将对 $a^x \pmod p$ 的计算化为对 $a^{2^i} \pmod p$ 的计算，而对 $a^{2^i} \pmod p$ 的计算具有“平方模”的统一形式，即

$$a^2 \equiv aa \pmod p,$$
$$a^{2^2} \equiv a^2 a^2 \pmod p,$$
$$\vdots$$
$$a^{2^{n-1}} \equiv a^{2^{n-2}} a^{2^{n-2}} \pmod p,$$

于是可以用递归程序来计算 $a^x \pmod p$. 算法描述如下：

```
输入：整数 a，  x，  p 要求 a>0，x≥0，p>1
输出：a^x (mod p)
int mod_exp(int a,  int x,  int p)
{
if (x==0)  return (1);
if (!(x%2)) return(mod_exp((a*a)%p, x/2, p));
return(a*(mod_exp((a*a)%p, x/2, p))(mod p));
}
```

可见，求模指数的算法需要的递归次数为$\lfloor log_2 x \rfloor+1$($\lfloor y \rfloor$表示不大于数 y 的最大整数)，每次递归要执行一次模平方运算或一次模平方运算外加一次模乘运算，其时间复杂度为

$$O((\lfloor \log_2 x \rfloor+1)^2)=O((\log_2 p)^2)\ (\text{因为 } x<p).$$

在讨论了模指数的计算方法后，下面来介绍计算离散对数的常用方法——商克(Shank)法.

设 p 为素数，如果 a 是 p 的原根，已知 $y\equiv a^x \pmod p$ $(0\leqslant x\leqslant p-1)$，求 x，商克法基本步骤如下(所有的运算在 mod p 下进行).

(1) 选一正整数 $d\approx\sqrt{p}$，则存在 q，r，使得

$$x=qd+r,\ 0\leqslant r<d,\ q\leqslant\frac{x}{d}<\sqrt{p},\ r<d\approx\sqrt{p}.$$

(2) 建立表$(\lambda, L_a(\lambda))$，$L_a(\lambda)=0, 1, \cdots, d-1$，按 λ 顺序排列，以便于检索. 因为 $0\leqslant r<d$，所以(a^r, r)一定出现在这个表中.

(3) 由于 $y=a^x=a^{qd+r}$，如果知道了 q 的取值，因为

$$y\cdot(a^{-d})^q=a^{qd+r}\cdot(a^{-d})^q=a^r,$$

我们就可以查表$(\lambda, L_a(\lambda))$，通过 a^r 确定 r 的值. 因为 $q<\sqrt{p}$，可以从 $q=0$ 开始试算，最多经过$\sqrt{p}$步试算一定能找到 q 的值.

商克算法的时间复杂度为 $O(\sqrt{p})$，这是一个全指数级的复杂度. 到目前为止，最好的求解有限域上的离散对数算法的时间复杂度是亚指数级的，人们还没有找到计算复杂性在 $O((\log n)^k)$(k 为常数)内的算法. 当 p 非常大时，计算离散对数是非常困难的. 正是利用离散对数的难解性，人们设计了各种基于离散对数的密码体制. 这里简要介绍基于离散对数的狄菲－黑尔曼(Diffie-Hellman)密钥交换体制：假设用户 A 与用户 B 要交换密钥，他们首先共同约定一个大素数 p 和 p 的原根 g(p 和 g 不需要保密)，A 任选一个整数 a，$1\leqslant a\leqslant p-1$，将 a 保密，将 $g^a \pmod p$公开. 用户 B 任选一个整数 b，将 $1\leqslant b\leqslant p-1$ 保密. 将 $g^b \pmod p$公开. 则 A 与 B 之间可确定通信密钥为 $K_{AB}=g^{ab} \pmod p$. 这是因为，用户 A 知道 a 和 $g^b \pmod p$，可以计算

$$K_{AB}\equiv(g^b)^a\equiv g^{ab} \pmod p,$$

用户 B 知道 b 和 $g^a \pmod p$，可以计算

$$K_{AB}\equiv(g^a)^b\equiv g^{ab} \pmod p,$$

而窃听者只能知道 g^a 和 g^b(当然还有 p 和 g)，基于离散对数的难解性，窃听者不能从 $g^a \pmod p$或 $g^b \pmod p$计算 a 或 b，从而无法得到 $g^{ab} \pmod p$.

5.3.2　椭圆曲线上的离散对数

椭圆曲线上的离散对数是有限域上的离散对数在椭圆曲线上的一种类比物. 给定有限域 $\mathbf{Z}_q$($q=p^r$ 为素数幂)上的一条椭圆曲线 E，并给定这条曲线上的两点 P 和 Q，求正整数 k(如果存在的话)使之满足 $Q=kP$ 的问题，称为椭圆曲线上的离散对数问题(ECDLP). 当点 P 的阶为大素数时，普遍认为 ECDLP 是难解的. 反过来，由 5.2.4 节的知识，已知 E 上的一点 P 和正整数 k，求 E 上的另一点 $Q=kP$ 则是很容易的.

下面介绍基于椭圆曲线的 ElGamal 公钥密码体制[①].

通信双方 A 和 B 事先在公开的信道上选定有限域 $\mathbf{Z}_q$，其中 $q=p^r$，p 为大素数(例如 $p\approx 2^{180}$)上的一条椭圆曲线 E，并在 E 中选择一个基点 G(G 要能生成一个很大的子群，这个子群最好和椭圆曲线 E 本身所构成的群一样大或较接近，也就是说，G 的阶 n 要足够大).

A 与 B 之间的密钥交换可以按如下方式进行：

A 选择一个比 n 小的正整数 a(a 可以认为是 A 的私钥)，并计算出 aG(aG 可以认为是 A 之公钥)，且将 aG 传输给 B；

B 类似地选择一个私钥 b 并计算出其公钥 bG，将 bG 传输给 A；

现假定 A 要给 B 传输信息 M(明文). 首先 A 将明文 M 嵌入到 E 上点 P_m，并选定一个随机数 k，利用 B 的公钥 bG 计算出密文 $C=(kG,\ P_m+k(bG))$，将 C 传输给 B. 为了能够将 C 变换回 M，B 需要对 C 进行解密计算，但由于 B 知道 b，所以 B 可以很容易地计算出 $P_m=P_m+k(bG)-b(kG)=C-b(kG)$.

显然，为了破译密文 C，敌方需要从 bG 中求出 b，而这是一个计算椭圆曲线上离散对数的问题(ECDLP)，ECDLP 的难解性保证了这种密码体制的安全性.

习题

1. 求直线 $X+Y+Z=0$ 上无穷远点的坐标.
2. 判断直线 $aX+bY+cZ=0$ 上的无穷远点和无穷远直线与直线 $aX+bY=0$ 的交点，是否同一个点.
3. 设 m，n 为大于 0 的正整数，求证代数闭域 K 上的多项式

$$f(x)=a_0x^n+a_1x^{n-1}+\cdots+a_n,$$
$$g(x)=b_0x^m+b_1x^{m-1}+\cdots+b_m.$$

的结式是 $m+n-2$ 阶的.

4. 用结式求多项式 $f(x)=3x^3+2x^2+x+1$ 的判别式 $D(f(x))$.
5. 设域 K 的特征不为 2，K 上的 Weierstrass 方程 $E(K)$为

$$\eta^2=x^3+\frac{b_2}{4}x^2+\frac{b_4}{2}x+\frac{b_6}{4},$$

求证对 $E(K)$有

$$\Delta=16D\left(x^3+\frac{b_2}{4}x^2+\frac{b_4}{2}x+\frac{b_6}{4}\right).$$

6. 求证域 K 上的 Weierstrass 方程 $E(K)$有如下等式成立：

$$4b_8=b_2b_6-b_4^2,$$
$$1\ 728\Delta=c_4^3-c_6^2,$$
$$j=\frac{c_4^3}{\Delta}=1\ 728+\frac{c_6^2}{\Delta}.$$

---

① 公钥密码体制有两个相关的密钥，分别称为公钥和私钥，用户的公钥对外公开，不须保密，而私钥须严格保密. 用一个密钥对明文进行加密，而用另一个密钥对密文进行解密. 公钥密码体制必须满足：(1) 由加密密钥确定解密密钥是不可能的；(2) 两个相关密钥中的任何一个都可以用作加密而让另一个用作解密.

7. 设 $j\neq 0$，1 728，求 Weierstrass 方程

$$y^2+xy=x^3-\frac{36}{j-1\ 728}x-\frac{1}{j-1\ 728}$$

的 j 不变量.

8. 求出 $E_{11}(1,6)$上所有的点. 并在平面直角坐标系内画出它的图像.

9. 已知 $E_{11}(1,6)$上一点 $G(2,7)$，求 $2G$ 到 $13G$ 所有的值.

10. 求 $E_{17}(1,1)$上的所有点，并求这些点的阶.

11. 编程绘制实数域 **R** 上形如 $y^2=x^3+ax+b$ 的椭圆曲线图像. 要求输入系数 a，b，检查奇异性条件，并输出符合条件的椭圆曲线图像.

第 6 章　线性反馈移位寄存器(LFSR)

反馈移位寄存器是产生伪随机流(又称伪随机数序列)的一种数学模型，在密码学、编码理论、建模与仿真等领域有广泛的应用. 信息论指出，任何由确定过程生成的随机流——也就是输出完全由输入决定的随机流——从本质上看都不是随机的. 由于很难生成真正的随机数，在许多实际应用中就采用**伪随机数发生器**(pseudo-Random Number Generator，pRNG)，许多高质量的 pRNG 都是以反馈移位寄存器为基础构造的. 本章讨论反馈移位寄存器的基本理论和 m 序列的概念，分析 m 序列的产生方法，最后介绍伪随机性判定的基本理论. 此外，学习本章的知识，还有助于我们深入理解有限域和多项式理论.

6.1　反馈移位寄存器

6.1.1　反馈移位寄存器

定义 6.1.1　**GF**(2)上一个 n 位**反馈移位寄存器**是由 n 个二元存储器 a_1，a_2，…，a_n 和一个反馈函数 $f(a_1, a_2, \cdots, a_n)$组成的动态系统，如图 6.1.1 所示.

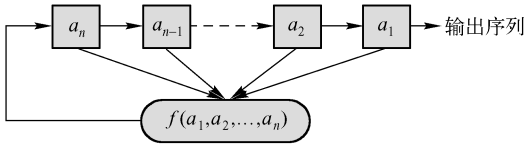

图 6.1.1　反馈移位寄存器

每个存储器称为反馈移位寄存器的一位，在任一时刻，这些位中的内容构成该反馈移位寄存器的状态，每一状态对应于 **GF**(2)上的一个 n 维向量，一个反馈移位寄存器共有 2^n 种可能的状态. 每一时刻的状态可用 n 长序列 a_1，a_2，…，a_n 或 n 维向量$(a_1, a_2, \cdots, a_n)$表示，其中 a_i 是第 i 位存储器的内容. 反馈函数 $f(a_1, a_2, \cdots, a_n)$是 n 元布尔函数，其 n 个变元 a_1，a_2，…，a_n 可以独立地取 0 和 1 这两个可能的值，反馈函数中的运算可以有与、或、非等逻辑运算，最后的函数值也为 0 或 1.

反馈移位寄存器的初始状态由用户确定，其运行由移位时钟脉冲来控制，当第 i 个移位时钟脉冲到来时，每一位存储器 a_i 都将其内容向下一位 a_{i-1} 传递，并根据寄存器此时的状态 a_1，a_2，…，a_n 计算 $f(a_1, a_2, \cdots, a_n)$，作为下一时刻的 a_n.

[例 6.1.1]　图 6.1.2 是一个 3 位反馈移位寄存器，其初始状态为

$$(a_1, a_2, a_3)=(1, 0, 1),$$

反馈函数为

$$f(a_1, a_2, a_3)=a_1 \oplus a_2 a_3,$$

其中⊕是模 2 加法，其运行状态和输出由表 6.1.1 给出.

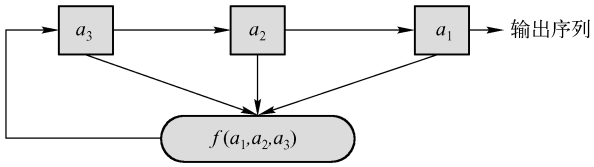

图 6.1.2　例 6.1.1 中的 3 位反馈移位寄存器

表 6.1.1　例 6.1.1 中的 3 位反馈移位寄存器的状态和输出

状态(a_3，a_2，a_1)			输　出
1	0	1	1
1	1	0	0
1	1	1	1
0	1	1	1
1	0	1	1
1	1	0	0
⋮	⋮	⋮	⋮

可以看出，这个反馈移位寄存器的输出序列为 101110111011…，周期为 4.

6.1.2　线性反馈移位寄存器(LFSR)

定义 6.1.2　如果移位寄存器的反馈函数 $f(a_1, a_2, \cdots, a_n)$是 $a_1, a_2, \cdots, a_n$ 的线性函数，则称之为**线性反馈移位寄存器**(Linear Feedback Shift Register，LFSR). 此时反馈函数可表示为

$$f(a_1, a_2, \cdots, a_n)=c_n a_1 \oplus c_{n-1} a_2 \oplus \cdots \oplus c_1 a_n, \tag{6.1.1}$$

其中，常数 $c_i=0$ 或 1，⊕是模 2 加法. $c_i=0$ 或 1 可用开关的断开和闭合来实现，如图 6.1.3 所示. 输出序列$\{a_i\}$满足

$$a_{n+k}=c_n a_k \oplus \cdots \oplus c_2 a_{n+k-2} \oplus c_1 a_{n+k-1}, \tag{6.1.2}$$

其中 k 为正整数.

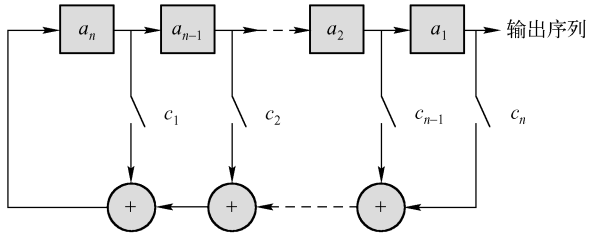

图 6.1.3　线性反馈移位寄存器

线性反馈移位寄存器因其实现简单、速度快、有较为成熟的理论等优点而成为构造随机流生成器的最重要的模型之一.

[**例 6.1.2**]　图 6.1.4 是一个 5 位线性反馈移位寄存器，其初始状态为

$$(a_1, a_2, a_3, a_4, a_5)=(1, 0, 0, 1, 1),$$

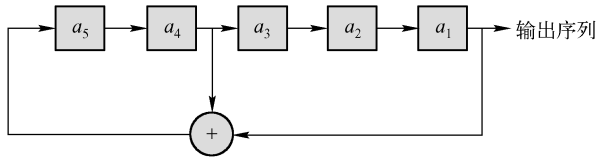

图 6.1.4　例 6.1.2 中的 5 位线性反馈移位寄存器

反馈函数为

$$f(a_1, a_2, a_3, a_4, a_5)=a_1 \oplus a_4$$

可求出输出序列为

$$\underbrace{1001101001000010101110110001111}_{31\text{位}}\underbrace{100110}\cdots$$

周期为 31.

在线性反馈移位寄存器中总是假定 $c_1, c_2, \cdots, c_n$ 中至少有一个不为 0，否则 $f(a_1, a_2, \cdots, a_n)$恒为 0，这样的话，在 n 个脉冲后状态必然是 00…0，且这个状态必将一直持续下去. 当只有一个系数不为 0，线性反馈移位寄存器就退化成一种延迟装置. 一般对于 n 位线性反馈移位寄存器，总是假定 $c_n=1$.

线性反馈移位寄存器输出序列的性质完全由其反馈函数决定. n 位线性反馈移位寄存器最多有 2^n 个不同的状态. 若其初始状态为 0，则其状态恒为 0. 若其初始状态非 0，则其后继状态不会恒为 0. 因此 n 位线性反馈移位寄存器的状态周期小于或等于 2^n-1. 其输出序列的周期与状态周期相等，也小于或等于 2^n-1，选择了合适的反馈函数，就可使序列的周期达到最大值 2^n-1，这样的序列具有最大的随机性.

定义 6.1.3 周期达到最大值 2^n-1 的 n 位线性反馈移位寄存器输出序列称为 **m 序列**.

例 6.1.2 中的线性反馈移位寄存器所输出的序列就是一个 m 序列.

为了从理论上深入分析线性反馈移位寄存器及 m 序列的产生条件，需要引入分圆多项式和本原多项式的概念. 6.2 节将深入讨论这一问题.

6.1.3 非线性组合移位寄存器简介

为了使伪随机流发生器 pRNG 产生的序列尽可能复杂，应保证其周期尽可能大、线性复杂度和不可预测性尽可能高，因此常使用多个线性反馈移位寄存器来构造 pRNG. 非线性组合移位寄存器(如图 6.1.5 所示)是最常见的 pRNG 之一. 它由 s 个线性反馈移位寄存器和一个非线性组合函数 f 构成，线性反馈移位寄存器提供良好的统计特性，非线性组合函数 f 负责提供高线性复杂度. 非线性组合移位寄存器的例子有 Geffe 序列生成器、J-K 触发器、Pless 生成器、钟控序列生成器等.

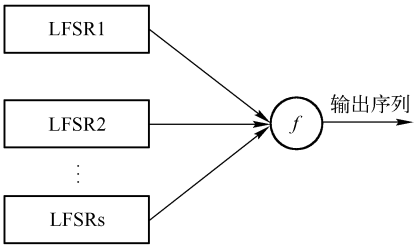

图 6.1.5 非线性组合移位寄存器示例

移位寄存器并不是构造伪随机流发生器的唯一途径，还存在许多其他生成伪随机数的方法，如同余发生器、Blum-Blum-Shub 发生器、Naor-Reingold 发生器等，感兴趣的读者可参考相关资料.

6.2 分圆多项式和本原多项式

为了从理论上深入分析线性反馈移位寄存器及 m 序列的产生条件，需要引入分圆多项式和本原多项式的概念. 这实际上是对第 4 章多项式理论的扩充.

6.2.1　分圆多项式

定义 6.2.1　设 K 是一个域，多项式 $f(x)\in K[x]$，$f(x)=c_nx^n+\cdots+c_1x+c_0$，则

$$f'(x)=nc_nx^{n-1}+(n-1)c_{n-1}x^{n-2}+\cdots+2c_2x+c_1.$$

这样，从纯代数意义上定义了多项式 $f(x)$的“导数”$f'(x)$，这个“导数”只是形式上的，并没有关于 $f(x)$的连续性或可导条件方面的假设. 容易证明，与微积分中的导数类似，对于这里定义的多项式“导数”，有如下性质成立.

定理 6.2.1　设 K 是一个域，对 $r\in K$ 和多项式环 $K[x]$上的两个多项式 $f(x)$，$g(x)$，有：

$$(1)\ (r\cdot f(x))'=r\cdot f'(x),\tag{6.2.1}$$

$$(2)\ (f(x)+g(x))'=f'(x)+g'(x),\tag{6.2.2}$$

$$(3)\ (f(x)g(x))'=f'(x)g(x)+f(x)g'(x).\tag{6.2.3}$$

这个定理的证明留作练习.

定理 6.2.2　设 K 是特征为 $p(p>0)$的域，$f(x)\in K[x]$，$m(x)$是 $K[x]$上的不可约多项式，则

(1) 如果 $m^2(x)\mid f(x)$，则 $m(x)\mid\gcd(f(x),f'(x))$.

(2) 如果 $m(x)\mid\gcd(f(x),f'(x))$且 K 中每个元素都有 p 次根，则 $m^2(x)\mid f(x)$.

在证明这一定理之前，先指出，对于素域 $K=\mathbf{Z}/p\mathbf{Z}$，条件“K 中每个元素都有 p 次根”是一定满足的，这是因为对任意 $a\in K$ 有

$$a^p \bmod p=(a^{p-1}a)\bmod p=a^{p-1}\bmod p\cdot a \bmod p=a \bmod p.$$

证明

(1) 因为 $m^2(x)\mid f(x)$，可设 $f(x)=m^2(x)g(x)$　$(g(x)\in K[x])$，由定理 6.2.1(3)知

$$f'(x)=2m(x)m'(x)g(x)+m^2(x)g'(x)=m(x)(2m'(x)g(x)+m(x)g'(x)),$$

于是 $f'(x)$为 $m(x)$的倍式，即 $m(x)\mid f'(x)$，又由 $m^2(x)\mid f(x)$可知 $m(x)\mid f(x)$，所以有

$$m(x)\mid\gcd(f(x),f'(x)).$$

(2) 根据多项式的欧几里得除法，用 $m(x)$除 $f(x)$得

$$f(x)/m(x)=q(x)\cdot m(x)+r(x),$$

这里 $\deg r(x)<\deg m(x)$，于是有

$$f(x)=q(x)\cdot m^2(x)+r(x)m(x),\tag{6.2.4}$$

两边求导有

$$f'(x)=q'(x)\cdot m^2(x)+2q(x)\cdot m(x)\cdot m'(x)+r'(x)m(x)+r(x)m'(x).\tag{6.2.5}$$

因为 $m(x)\mid\gcd(f(x),f'(x))$，所以 $m(x)\mid f'(x)$，于是式(6.2.5)右边的每一项都能被 $m(x)$整除，于是有 $m(x)\mid r(x)m'(x)$，又因为 $m(x)$是不可约的，所以只有 $m(x)\mid m'(x)$或 $m(x)\mid r(x)$两种可能. 用反证法来证明 $m(x)\mid m'(x)$是不可能的. 因为 $\deg m'(x)<\deg m(x)$，如果$m(x)\mid m'(x)$，则 $m'(x)$必为零多项式. 设

$$m(x)=a_nx^n+a_{n-1}x^{n-1}+\cdots+a_2x^2+a_1x+a_0,$$

那么

$$m'(x)=na_nx^{n-1}+(n-1)a_{n-1}x^{n-2}+\cdots+2a_2x^1+a_1,$$

因为 $m'(x)$为零多项式，所以必有 $la_l=0$ $(1\leqslant l\leqslant n)$，由于乘法满足消去律，对任意 $a_l\neq 0$，必有 $l\cdot 1=0$，于是有域 K 的特征 $p\mid l$，所以 $m(x)$可以写成如下形式

$$m(x)=a_{pm}x^{pm}+a_{p(m-1)}x^{p(m-1)}+\cdots+a_{2p}x^{2p}+a_px^p+a_0,$$

由条件 K 中每个元素都有 p 次根，设 $b_m^p=a_{pm}$，再根据定理 4.2.6 得

$$m(x)=(b_mx^m+b_{m-1}x^{m-1}+\cdots+b_2x^2+b_1x+b_0)^p.$$

这与 $m(x)$的不可约性矛盾，于是证明了 $m(x)\mid m'(x)$是不可能的，于是必有 $m(x)\mid r(x)$，由 $\deg r(x)<\deg m(x)$，得 $r(x)=0$，由式(6.2.4)知 $m^2(x)\mid f(x)$，证毕.

定理 6.2.2 实际上给出了判断素域上多项式是否有重因子的充要条件：

推论 设 K 为素域，$f(x)\in K[x]$，$f(x)$在 $K[x]$上无重因子当且仅当 $\gcd(f(x),f'(x))=1$.

定理 6.2.3 设 m，n 为两个整数，则有

$$\gcd(x^m-1,x^n-1)=x^{\gcd(m,n)}-1. \tag{6.2.6}$$

证明 当 $m=n$ 时，结论显然成立；当 $m\neq n$ 时，对 m 和 n 的最大值作归纳法，当 $m>n$ 时，作因子变形

$$x^m-1-x^{m-n}\cdot(x^n-1)=x^{m-n}-1,$$

因此，如果 $d(x)\mid x^m-1$ 且 $d(x)\mid x^n-1$，必有 $d(x)\mid x^{m-n}-1$，于是

$$\gcd(x^m-1,x^n-1)=\gcd(x^{m-n}-1,x^n-1).$$

由归纳假设

$$\gcd(x^{m-n}-1,x^n-1)=x^{\gcd(m-n,n)}-1,$$

所以

$$\gcd(x^m-1,x^n-1)=x^{\gcd(m-n,n)}-1=x^{\gcd(m,n)}-1.$$

当 $m<n$ 时，结论同理可证.

定理 6.2.4 设域 K 的特征为 p，如果 p 不能整除正整数 n，则多项式 x^n-1 没有重因式.

证明 因为 p 不能整除 n，所以 $n\cdot 1\neq 0$，即 $n\cdot 1$ 存在乘法逆元，设 $n\cdot 1$ 的乘法逆元为 t，有

$$(tx)\cdot(nx^{n-1})-(x^n-1)=(tn)\cdot x^n-(x^n-1)=x^n-(x^n-1)=1,$$

由多项式的欧几里得除法，有

$$\gcd(x^n-1,(x^n-1)')=\gcd(x^n-1,nx^{n-1})=1,$$

由定理 6.2.2 的推论知，x^n-1 没有重因式，证毕.

定义 6.2.2 设域 K 的特征为 p，p 不能整除正整数 n，多项式

$$\varphi_n(x)=\frac{x^n-1}{\operatorname{lcm}(\{x^d-1\mid 0<d<n,\ d\mid n\})} \tag{6.2.7}$$

称为 n **次分圆多项式**，分母里所取的最小公倍式为首一的.

容易验证，对于满足 $0<d<n$ 且 $d\mid n$ 的所有整数 d，一定有$(x^d-1)\mid(x^n-1)$，根据域上多项式的唯一分解定理，所有能整除 x^n-1 的因式的最小公倍式也能整除 x^n-1，所以，对任意正整数 n，分圆多项式 $\varphi_n(x)$一定存在.

分圆多项式具有如下性质：

定理 6.2.5 设 m，n 为任意两个不同的正整数，域 K 的特征 p 不能整除 m，n，则

(1) $\varphi_n(x)$是首一多项式；

(2) 不同分圆多项式是互素的，即 $\gcd(\varphi_n(x), \varphi_m(x))=1$；

(3) $x^n-1=\prod\limits_{0<d\leqslant n,\ d\mid n}\varphi_d(x)$，亦即 $\varphi_n(x)=\dfrac{x^n-1}{\prod\limits_{0<d<n,\ d\mid n}\varphi_d(x)}$；

(4) $\deg(\varphi_n(x))=\varphi(n)$.

证明

(1) 因为 $\varphi_n(x)$是首一多项式 x^n-1 与首一最小公倍式的商，所以它一定是首一的.

(2) 设 m，n 是不同的正整数，由分圆多项式的定义，$\varphi_n(x)\mid x^n-1$，$\varphi_m(x)\mid x^m-1$，因此

$$\gcd(\varphi_n(x), \varphi_m(x))\mid\gcd(x^n-1, x^m-1).$$

由定理 6.2.3 知

$$\gcd(x^m-1, x^n-1)=x^{\gcd(m, n)}-1,$$

于是有

$$\gcd(\varphi_n(x), \varphi_m(x))\mid(x^{\gcd(m, n)}-1),$$

可以设

$$(x^{\gcd(m, n)}-1)=q_1(x)\cdot\gcd(\varphi_n(x), \varphi_m(x)). \tag{6.2.8}$$

另一方面，由分圆多项式的定义，有

$$\varphi_n(x)\mid\left(\frac{x^n-1}{x^{\gcd(m, n)}-1}\right),$$

所以

$$\gcd(\varphi_n(x), \varphi_m(x))\mid\left(\frac{x^n-1}{x^{\gcd(m, n)}-1}\right).$$

可以设

$$\frac{x^n-1}{x^{\gcd(m, n)}-1}=q_2(x)\cdot\gcd(\varphi_n(x), \varphi_m(x)), \tag{6.2.9}$$

由式(6.2.9)和式(6.2.8)得

$$\begin{aligned}x^n-1&=q_2(x)\cdot\gcd(\varphi_n(x), \varphi_m(x))\cdot(x^{\gcd(m, n)}-1)\\&=q_2(x)\cdot\gcd(\varphi_n(x), \varphi_m(x))\cdot q_1(x)\cdot\gcd(\varphi_n(x), \varphi_m(x))\\&=q_2(x)\cdot q_1(x)\cdot(\gcd(\varphi_n(x), \varphi_m(x)))^2,\end{aligned}$$

因为 K 的特征不能整除 n，由定理 6.2.4 知 x^n-1 没有重因式，于是只能有

$$\gcd(\varphi_n(x), \varphi_m(x))=1.$$

(3) 用数学归纳法，当 $n=1$ 时，结论显然成立，假设 $d<n$ 时，

$$x^d-1=\prod_{0<e\leqslant d,\ e\mid d}\varphi_e(x), \tag{6.2.10}$$

由式(6.2.7)

$$x^n-1=\varphi_n(x)\cdot\operatorname{lcm}(\{x^d-1\mid 0<d<n,\ d\mid n\}), \tag{6.2.11}$$

在(2)中已经证明了当 $m\neq n$ 时，有 $\gcd(\varphi_n(x), \varphi_m(x))=1$，由式(6.2.10)得

$$\operatorname{lcm}(\{x^d-1\mid 0<d<n,\ d\mid n\})=\prod_{0<d<n,\ d\mid n}\varphi_d(x), \tag{6.2.12}$$

式(6.2.12)代入式(6.2.11)得

$$x^n-1=\varphi_n(x)\cdot\prod_{0<d<n,\ d\mid n}\varphi_d(x)=\prod_{0<d\leqslant n,\ d\mid n}\varphi_d(x),$$

亦即

$$\varphi_n(x)=\frac{x^n-1}{\prod\limits_{0<d<n,\ d\mid n}\varphi_d(x)}. \tag{6.2.13}$$

(4) 利用欧拉函数的性质 $\sum\limits_{0<d\leqslant n,\ d\mid n}\varphi(d)=n$ 很容易证明，具体证明过程留给读者.

式(6.2.13)实际上给出了一种用递归计算分圆多项式的方法.

[例 6.2.1] 求域 $\mathbf{Z}/2\mathbf{Z}$ 上的分圆多项式 $\varphi_1(x)$、$\varphi_3(x)$、$\varphi_5(x)$、$\varphi_7(x)$、$\varphi_{15}(x)$.

解 由式(6.2.13)

$$\varphi_1(x)=x-1,$$

$$\varphi_3(x)=\frac{x^3-1}{\varphi_1(x)}=\frac{x^3-1}{x-1}=x^2+x+1,$$

$$\varphi_5(x)=\frac{x^5-1}{\varphi_1(x)}=\frac{x^5-1}{x-1}=x^4+x^3+x^2+x+1,$$

$$\varphi_7(x)=\frac{x^7-1}{\varphi_1(x)}=\frac{x^7-1}{x-1}=x^6+x^5+x^4+x^3+x^2+x+1,$$

$$\begin{aligned}\varphi_{15}(x)&=\frac{x^{15}-1}{\varphi_1(x)\varphi_3(x)\varphi_5(x)}\\&=\frac{x^{15}-1}{\varphi_3(x)(\varphi_1(x)\varphi_5(x))}\\&=\frac{x^{15}-1}{(x^2+x+1)(x^5-1)}=x^8+x^7+x^5+x^4+x^3+x+1.\end{aligned}$$

6.2.2 本原多项式

定义 6.2.3 $\mathbf{Z}_p[x]$中的 n 次多项式 $p(x)$称为 n **次本原多项式**，如果 $p(x)\mid(x^{p^n-1}-1)$，且对 $0<t<p^n-1$ 的任意整数 t，$p(x)$不整除 x^t-1，本原多项式的另一个等价定义为：n 次多项式 $p(x)$为本原多项式，当且仅当 $x^{p^n-1}\equiv1\pmod{p(x)}$，且对 $0<t<p^n-1$ 的任意整数 t，$x^t\not\equiv1\pmod{p(x)}$.

关于本原多项式，有如下定理.

定理 6.2.6 $\mathbf{Z}_p[x]$中的 n 次多项式 $p(x)$为本原多项式当且仅当 $x^{p^n-1}\equiv1\pmod{p(x)}$，且对 $0<d<p^n-1$，$d\mid p^n-1$ 的任意整数 d，有 $x^d\not\equiv1\pmod{p(x)}$.

证明 为行文简洁，设 $N=p^n-1$. 只需证明，如果 $x^N\equiv1\pmod{p(x)}$，存在 $0<t<N$ 的整数 t 使得 $x^t\equiv1\pmod{p(x)}$，那么一定存在 N 的真因子 d，有 $x^d\equiv1\pmod{p(x)}$，设 a，b 为满足如下条件的整数

$$\gcd(t,N)=aN+bt,$$

则

$$x^{\gcd(N,t)}=x^{aN+bt}=(x^N)^a\cdot(x^t)^b=1^a\cdot1^b\equiv1\pmod{p(x)},$$

注意对负指数，我们需要知道 x 是模 $p(x)$可逆的. 这一点是一定满足的，因为由

$$x^N\equiv1\pmod{p(x)},$$

可得

$$x \cdot x^{N-1} \equiv 1 (\mathrm{mod}\ p(x)).$$

即 x^{N-1} 是 x 模 $p(x)$ 的逆元. 所以总是存在 N 的真因子 $d=\gcd(N, t)$，使得 $x^d \not\equiv 1 (\mathrm{mod}\ p(x))$. 证毕.

下面研究分圆多项式和本原多项式的关系，从而给出一种从分圆多项式的角度求本原多项式的算法. 为此，先补充一个关于有限域结构的定理.

定理 6.2.7　设 $K^{\times}$ 表示有限域 K 的乘法群，则 $K^{\times}$ 为循环群.

证明　设 N 是有限域 K 的阶(即 K 中元素的个数)，由域的性质，K 中任何非零元素都存在逆元，因此乘法群 $K^{\times}$ 的阶为 $N-1$，由拉格朗日定理，对任意 $\alpha \in K^{\times}$，都有

$$\alpha^{N-1}=1,$$

这就是说，$K^{\times}$ 中 $N-1$ 个元素都是多项式 $f(x)=x^{N-1}-1$ 的根. 另一方面，由代数基本定理，$N-1$ 次多项式 $f(x)$ 在 K 中至多有 $N-1$ 个根. 因此，$f(x)$ 在 K 中恰有 $N-1$ 个不同根，即 $f(x)$ 无重根.

设 K 的特征为 p，由定理 6.2.4，p 不整除 $N-1$，由分圆多项式的性质

$$f(x)=x^{N-1}-1=\prod_{0<d\leqslant N-1,\ d\mid(N-1)} \varphi_d(x),$$

即多项式 $\prod\limits_{0<d\leqslant N-1,\ d\mid(N-1)} \varphi_d(x)$ 在域 K 中有 $N-1$ 个根. 由定理 6.2.5，对于满足 $0<d\leqslant N-1$，

$d \mid (N-1)$ 的整数 d，因为 $\deg \varphi_d(x)=\varphi(d)$，所以 $\varphi_d(x)$ 在 K 中最多有 $\varphi(d)$ 个根；又因为所有 $\varphi_d(x)$ 两两互素，所以不同的 $\varphi_d(x)$ 没有公共根，再根据欧拉函数的性质

$$\sum_{0<d\leqslant N-1,\ d\mid(N-1)} \varphi(d)=N-1,$$

所以 $\varphi_d(x)$ 在 K 中有且仅有 $\varphi(d)$ 个根. 所以分圆多项式 $\varphi_{N-1}(x)$ 一定有 $\varphi(N-1)$ 个根，设 β 是这 $\varphi(N-1)$ 个根中的任意一个，则一定有 $\beta^{N-1}=1$(因为 $\varphi_{N-1}(x) \mid (x^{N-1}-1)$)，且不存在 $e<N-1$，使得 $\beta^e=1$(因为不同分圆多项式是互素的)，由循环群的定义，群 $K^{\times}$ 构成阶为 $N-1$ 的循环群，$\varphi_{N-1}(x)$ 的 $\varphi(N-1)$ 个根都是 $K^{\times}$ 的生成元. 证毕.

定义 4.2.15 给出了群中元素的阶的定义，从定理 6.2.7 的证明过程可以看出，分圆多项式 $\varphi_d(x)$ 的每个根的阶为 d.

定理 6.2.8　设 $\mathbf{Z}_p[x]$ 中多项式 $p(x)$ 是 n 次本原多项式，$N=p^n-1$，则 $p(x)$ 整除 N 次分圆多项式，即

$$p(x) \mid \varphi_N(x).$$

证明　因为多项式 $p(x)$ 是 n 次本原多项式，所以有

$$p(x) \mid (x^N-1),$$

且不存在 $d<N$，$d \mid N$ 使

$$p(x) \mid (x^d-1),$$

即 $p(x)$ 整除从 x^N-1 中约去 $\mathrm{lcm}(\{x^d-1 \mid 0<d<n,\ d \mid n\})$ 得到的多项式，由定义 6.2.2，从 x^N-1 中约去 $\mathrm{lcm}(\{x^d-1 \mid 0<d<n,\ d \mid n\})$ 恰好得到分圆多项式 $\varphi_N(x)$，所以有

$$p(x) \mid \varphi_N(x).$$

定理 6.2.9　$\mathbf{Z}_p[x]$ 中 p^n-1 次分圆多项式 $\varphi_{p^n-1}(x)$ 的每个不可约因式都是 n 次本原多

项式.

证明 不失一般性，仅考虑 $n>1$ 的情况，因为 $n=1$ 时的线性本原多项式很容易处理.

首先证明 $\varphi_{p^n-1}(x)$的每个不可约因式的次数为 n. 设 $q(x)$为 $\varphi_{p^n-1}(x)$的不可约因式，d 是 $q(x)$的次数，由定理 4.4.2，

$$L=\mathbf{Z}_p[x]/(q(x))$$

为一个 p^d 阶有限域，其乘法群 $L^\times$ 的阶为 p^d-1，设 x 在 $L^\times$ 中的像为 α，即

$$\alpha\equiv x(\bmod q(x)),$$

由分圆多项式的性质，α 在 $L^\times$ 中的阶为 p^n-1，由拉格朗日定理，必有$(p^n-1)|(p^d-1)$，所以 $d\geqslant n$；另一方面，因为

$$q(x)|\varphi_{p^n-1}(x),\ \varphi_{p^n-1}(x)|(x^{p^n-1}-1),$$

所以有 $q(x)|(x^{p^n-1}-1)$，即

$$x^{p^n-1}\equiv 1(\bmod q(x)). \tag{6.2.14}$$

设 $f(x)=\sum_{i=0}^{t}a_ix^i$，$a_i\in\mathbf{Z}_p$，则

$$(f(x))^{p^n}=\Big(\sum_{i=0}^{t}a_ix^i\Big)^{p^n}=\sum_{i=0}^{t}a_i^{p^n}(x^{p^n})^i=\sum_{i=0}^{t}a_i(x^{p^n})^i=f(x^{p^n}),$$

由式(6.2.14)有

$$(f(x))^{p^n}=f(x^{p^n})=f(x^{p^n-1}x)\equiv f(x)\ (\bmod q(x)).$$

对于模 $q(x)$，共有 p^d 个不同的多项式(即域 L 的阶为 p^d)，由于同余方程

$$X^{p^n}-X\equiv 0(\bmod q(x))$$

的根的个数不会超过 p^n，故有 $p^d\leqslant p^n$，所以 $d\leqslant n$；因此，$\varphi_{p^n-1}(x)$的每个不可约因式的次数为 n.

接着来证明 $\varphi_{p^n-1}(x)$的每个次数为 n 的不可约因式$q(x)$一定是本原多项式，式(6.2.14)已经指出，$x^{p^n-1}\equiv 1(\bmod q(x))$，因为 p 与 p^n-1 是互素的，由分圆多项式的定义，对任意整数 $t<p^n-1$，$\varphi_{p^n-1}(x)$与 x^t-1 没有公因子，因此不存在整数 $t<p^n-1$，使得 $x^t\equiv 1(\bmod q(x))$，因此 $q(x)$为 n 次本原多项式. 定理证毕.

由定理 6.2.7，p^n-1 阶分圆多项式 $\varphi_{p^n-1}(x)$的次数为 $\varphi(p^n-1)$，由定理 6.2.9，$\varphi_{p^n-1}(x)$的不可约因式恰为 n 次本原多项式，因为多项式积的次数等于各多项式次数的和，所以有以下定理.

定理 6.2.10 $\mathbf{Z}_p[x]$中 n 次本原多项式的数量为

$$\frac{\varphi(p^n-1)}{n}.$$

[例 6.2.2] 结合 4.3.5 节中例 4.3.5 关于 $\mathbf{Z}_2[x]$上不可约多项式的结论和例 6.2.1，求 $\mathbf{Z}_2[x]$上五次以内的本原多项式.

解 (1) $\mathbf{Z}_2[x]$上的线性多项式 $x+1$ 是本原多项式，因为它不可约且整除$(2^1-1)=1$ 次分圆多项式 $\varphi_1(x)=x-1=x+1$，线性多项式 x 虽然不可约，但它不整除 $\varphi_1(x)$，所以它不是本原多项式.

(2) $\mathbf{Z}_2[x]$上的二次多项式 x^2+x+1 是本原多项式，因为它不可约且整除$(2^2-1)=3$ 次分圆多项式 $\varphi_3(x)=x^2+x+1$.

(3) $\mathbf{Z}_2[x]$上有两个三次不可约多项式：x^3+x^2+1 和 x^3+x+1，由例 6.2.1 可知

$$\varphi_{2^3-1}(x)=\varphi_7(x)=x^6+x^5+x^4+x^3+x^2+x+1,$$

而

$$(x^3+x^2+1)\cdot(x^3+x+1)=x^6+x^5+x^4+x^3+x^2+x+1,$$

所以两个三次不可约多项式 x^3+x^2+1 和 x^3+x+1 都是本原多项式.

(4) $\mathbf{Z}_2[x]$上有三个四次不可约多项式：$x^4+x^3+x^2+x+1$，x^4+x^3+1 和 x^4+x+1，由例 6.2.1 可知

$$\varphi_{2^4-1}(x)=\varphi_{15}(x)=x^8+x^7+x^5+x^4+x^3+x+1,$$

因为

$$(x^4+x^3+1)\cdot(x^4+x+1)=x^8+x^7+x^5+x^4+x^3+x+1,$$

所以四次不可约多项式 x^4+x^3+1 和 x^4+x+1 都是本原多项式，但 $x^4+x^3+x^2+x+1$ 不是本原多项式.实际上，由例 6.2.1 可知

$$\varphi_5(x)=x^4+x^3+x^2+x+1,$$

由式(6.2.13)可知，$\varphi_5(x)$一定不是 $\varphi_{15}(x)$的因子.

(5) $\mathbf{Z}_2[x]$上有六个五次不可约多项式：$x^5+x^3+x^2+x+1$，$x^5+x^4+x^2+x+1$，$x^5+x^4+x^3+x+1$，$x^5+x^4+x^3+x^2+1$，x^5+x^3+1 和 x^5+x^2+1，根据分圆多项式的定义，

$$\varphi_{2^5-1}(x)=\varphi_{31}(x)=\frac{x^{31}-1}{x-1}.$$

这是一个 30 次的多项式. 由定理 6.2.9，它恰好是六个五次不可约多项式的乘积，于是五次不可约多项式 $x^5+x^3+x^2+x+1$，$x^5+x^4+x^2+x+1$，$x^5+x^4+x^3+x+1$，$x^5+x^4+x^3+x^2+1$，x^5+x^3+1 和 x^5+x^2+1 都是本原多项式，读者不妨自行验证这六个不可约多项式的积恰为 $\varphi_{31}(x)$.

本节讨论了有限域上分圆多项式和本原多项式的概念以及它们之间的关系，给出了通过分解分圆多项式求本原多项式的方法. 分圆多项式判定和分圆多项式分解的高效算法对密码学和编码理论具有十分重要的实际意义，目前这一域领的研究比较活跃，有兴趣的读者可参考相关资料.

6.3　*m* 序列

我们已经知道，*m* 序列是线性反馈移位寄存器 LFSR 所能产生的周期最长的随机序列，一个自然而又实际的问题就是：满足什么条件的 LFSR 能够产生 *m* 序列？本节从理论上深入分析这一问题.

6.3.1　LFSR 的特征多项式

由式(6.1.2)知 *n* 位 LFSR 的输出序列$\{a_i\}$满足递推关系，

$$a_{n+k}=c_1a_{n+k-1}\oplus c_2a_{n+k-2}\oplus\cdots\oplus c_na_k,$$

c_1，c_2，…，$c_n\in\mathbf{Z}_2$，我们定义矩阵 $\boldsymbol{L}$ 如下：

$$\boldsymbol{L}=\begin{bmatrix} c_1 & c_2 & c_3 & \cdots & c_{n-1} & c_n \\ 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & 0 \end{bmatrix}, \tag{6.3.1}$$

即矩阵的最上面一行是 LFSR 系数(注意它们的顺序)，一条次对角线上元素为 1，其余元素皆为 0. 利用 $\boldsymbol{L}$ 矩阵可以将 LFSR 的输出序列递推关系表示如下：

$$\boldsymbol{L}\cdot\begin{bmatrix} a_{n+k-1} \\ a_{n+k-2} \\ \vdots \\ a_{k+1} \\ a_k \end{bmatrix}=\begin{bmatrix} a_{n+k} \\ a_{n+k-1} \\ \vdots \\ a_{k+2} \\ a_{k+1} \end{bmatrix}, \tag{6.3.2}$$

设 LFSR 的初始状态

$$\boldsymbol{v}=\begin{bmatrix} a_n \\ a_{n-1} \\ \vdots \\ a_2 \\ a_1 \end{bmatrix}$$

可以表示为线性组合的形式，即有

$$\boldsymbol{v}=k_1\boldsymbol{v}_1+k_2\boldsymbol{v}_2+\cdots+k_n\boldsymbol{v}_n,$$

其中 k_1，k_2，…，k_n 为线性组合系数，$\boldsymbol{v}_1$，$\boldsymbol{v}_2$，…，$\boldsymbol{v}_n$ 为矩阵 $\boldsymbol{L}$ 的特征向量(事实上，任意初始状态列向量 $\boldsymbol{v}$ 都可以表示成 $\boldsymbol{v}_1$，$\boldsymbol{v}_2$，…，$\boldsymbol{v}_n$ 的线性组合的形式，这一结论请读者自证)，由式(6.3.2)可知，

$$\begin{bmatrix} a_{n+k} \\ a_{n+k-1} \\ \vdots \\ a_{k+2} \\ a_{k+1} \end{bmatrix}=\boldsymbol{L}\cdot\begin{bmatrix} a_{n+k-1} \\ a_{n+k-2} \\ \vdots \\ a_{k+1} \\ a_k \end{bmatrix}=\boldsymbol{L}^2\cdot\begin{bmatrix} a_{n+k-2} \\ a_{n+k-3} \\ \vdots \\ a_k \\ a_{k-1} \end{bmatrix}=\cdots=\boldsymbol{L}^k\cdot\boldsymbol{v}.$$

设与 v_i 对应的特征值为 $\lambda_i(1\leqslant i\leqslant n)$，结合矩阵理论可知

$$\begin{bmatrix} a_{n+k} \\ a_{n+k-1} \\ \vdots \\ a_{k+2} \\ a_{k+1} \end{bmatrix}=\boldsymbol{L}^k\cdot\boldsymbol{v}=\lambda_1^k k_1\boldsymbol{v}_1+\lambda_2^k k_2\boldsymbol{v}_2+\cdots+\lambda_n^k k_n\boldsymbol{v}_n, \tag{6.3.3}$$

容易得到，矩阵 $\boldsymbol{L}$ 的特征多项式为

$$p(x)=x^n-c_1x^{n-1}-c_2x^{n-2}-\cdots-c_{n-1}x-c_n. \tag{6.3.4a}$$

由线性代数的知识，矩阵 $\boldsymbol{L}$ 的特征多项式 $p(x)$ 的根就是矩阵 $\boldsymbol{L}$ 的特征值 $\lambda_i(1\leqslant i\leqslant n)$. 可以从

分析 $p(x)$入手来研究 LFSR 输出序列的特征. 因为 LFSR 是定义在 $\mathbf{Z}_2$ 上的，$p(x)\in\mathbf{Z}_2[x]$，所以式(6.3.4a)也可以写成

$$p(x)=x^n+c_1x^{n-1}+c_2x^{n-2}+\cdots+c_{n-1}x+c_n. \tag{6.3.4b}$$

定义 6.3.1　式(6.3.4a)或式(6.3.4b)称为 n 位 LFSR 的**特征多项式**.

6.3.2　m 序列的产生条件

定理 6.3.1　设 $p(x)$是 n 位 LFSR 的特征多项式，则此 LFSR 输出序列为 m 序列的充要条件是 $p(x)$为 n 次本原多项式.

我们分三个步骤来证明这一定理：第一步，n 次本原多项式 $p(x)$在有限域 $\mathbf{Z}_{2^n}$ 中一定有 n 个不同的根，即式(6.3.3)中的 n 个特征值 $\lambda_i(1\leqslant i\leqslant n)$是一定存在的；第二步，如果 LFSR 的特征多项式 $p(x)$为 n 次本原多项式，则 $p(x)$的每个根的阶为 2^n-1，即矩阵 $\boldsymbol{L}$ 的每个特征值的阶 2^n-1；第三步，如果矩阵 $\boldsymbol{L}$ 的每个特征值的阶为 2^n-1，则 LFSR 的输出序列为 m 序列.

证明　在 4.4 节介绍有限域理论时曾指出，$\mathbf{Z}_{2^n}$ 中包含了 $\mathbf{Z}_2$ 上所有 n 次不可约多项式的全部 n 个根，由定理 6.2.9，当 $p(x)$为 n 次本原多项式时，$p(x)$是 $\mathbf{Z}_2[x]$上的 n 次不可约多项式，所以 $p(x)$在 $\mathbf{Z}_{2^n}$ 中一定存在 n 个根，且这 n 个根是不同的.

由定理 6.2.7，任何有限域的乘法群都是循环群，所以 $\mathbf{Z}_{2^n}^{\times}$ 是阶为 2^n-1 的循环群，也就是说，$\mathbf{Z}_{2^n}$ 的 2^n-1 非零元素都满足

$$x^{2^n-1}-1=0,$$

而 2^n-1 阶分圆多项式 $\varphi_{2^n-1}(x)$就是从 $x^{2^n-1}-1$ 中约去所有 $x^d-1(0<d<2^n-1,\ d\mid(2^n-1))$的公因子后剩余的部分，所以 $\varphi_{2^n-1}(x)$在 $\mathbf{Z}_{2^n}^{\times}$ 中的每个根的阶都为 2^n-1. 根据定理 6.2.9，当 $p(x)$为 n 次本原多项式时，

$$p(x)\mid\varphi_{2^n-1}(x),$$

所以 $p(x)$的 n 个根都是 2^n-1 阶的，即矩阵 $\boldsymbol{L}$ 的 n 个特征值 $\lambda_i(1\leqslant i\leqslant n)$都是 2^n-1 阶的.

设ℓ为 LFSR 的周期，由式(6.3.3)，ℓ是使

$$\lambda_1^{i+\ell}k_1\boldsymbol{v}_1+\lambda_2^{i+\ell}k_2\boldsymbol{v}_2+\cdots+\lambda_n^{i+\ell}k_n\boldsymbol{v}_n=\lambda_1^ik_1\boldsymbol{v}_1+\lambda_2^ik_2\boldsymbol{v}_2+\cdots+\lambda_n^ik_n\boldsymbol{v}_n,$$

成立的最小正整数，上式可化简得到

$$(\lambda_1^{i+\ell}-\lambda_1^i)\boldsymbol{v}_1+(\lambda_2^{i+\ell}-\lambda_2^i)\boldsymbol{v}_2+\cdots+(\lambda_n^{i+\ell}-\lambda_n^i)\boldsymbol{v}_n=0,$$

即

$$(\lambda_1^{\ell}-1)\lambda_1^i\boldsymbol{v}_1+(\lambda_2^{\ell}-1)\lambda_2^i\boldsymbol{v}_2+\cdots+(\lambda_n^{\ell}-1)\lambda_n^i\boldsymbol{v}_n=0. \tag{6.3.5}$$

因为特征向量是线性无关的，即任何关于特征向量 $v_i(1\leqslant i\leqslant n)$的线性组合

$$a_1\boldsymbol{v}_1+a_2\boldsymbol{v}_2+\cdots+a_n\boldsymbol{v}_n=0$$

当且仅当线性组合系数 a_1，a_2，…，a_n 同时为零. 所以式(6.3.5)成立当且仅当对所有的 i 有系数$(\lambda_i^{\ell}-1)\lambda_i^i=0$，因为 λ_i 本身不为 0，所以这个充分必要条件就是对所有的 i，有 $\lambda_i^{\ell}-1=0$. 因为每个 λ_i 的阶为 2^n-1，ℓ是使所有 $\lambda_i^{\ell}-1=0$ 的最小正整数，所以这个充要条件最终化为$\ell=2^n-1$.

至此证明了 n 位 LFSR 的输出序列周期为 2^n-1 的充要条件是，其特征多项式 $p(x)$为 n 次本原多项式，再根据 m 序列的定义，得到：n 位 LFSR 的输出序列为 m 序列的充要条件是，

其特征多项式 $p(x)$为 n 次本原多项式. 定理得证.

[例 6.3.1] 根据例 6.2.2 的结论，$p(x)=x^4+x^3+1$ 是 4 次本原多项式，若 4 位 LFSR 以 $p(x)$以特征多项式，则输出序列的递推关系为

$$a_k=a_{k-1}\oplus a_{k-4},$$

若初始状态为 1001，则输出序列为

$$\underbrace{100100011110101}_{15位}\underbrace{100100011110101}_{15位}\cdots$$

周期为 $2^4-1=15$，即输出序列为 m 序列.

6.3.3 m 序列的特点

m 序列具有周期性，所以它本质上是伪随机序列，衡量一个伪随机序列好坏的标准有多种，比较通用的有著名的 Golomb 的三个公设条件，Rueppel 的线性复杂度随机走动条件，线性逼近及产生该序列的布尔函数满足的相关免疫条件等.

设$\{a_i\}=(a_1a_2a_3\cdots)$为 0、1 序列，例如 0001100111，其前三个数字是 000，称为 0 的 3 **游程**，11 是 1 的 2 **游程**，接下来 00 是 0 的 2 游程，再接下来 111 是 1 的 3 游程.

定义 6.3.2 $\mathbf{Z}_2$ 上周期为 T 的序列$\{a_i\}$的**自相关系数**定义为

$$R(\tau)=\frac{1}{T}\sum_{k=1}^{T}(-1)^{a_k}(-1)^{a_{k+\tau}},\ 0\leqslant\tau\leqslant T-1 \tag{6.3.6}$$

定义中的和式表示序列$\{a_i\}$与序列$\{a_{i+\tau}\}$(序列$\{a_i\}$向后平移 τ 位得到)在一个周期内对应位相同与对应位不同的位数之差，即序列$(a_ia_{i+1}\cdots a_{i+T-1})$与序列$(a_{i+\tau}a_{i+\tau+1}\cdots a_{i+\tau+T-1})$对应位相同与对应位不同的位数之差. 显然，$R(0)=1$，当 $\tau\neq0$ 时，称 $R(\tau)$为**异相自相关系数**.

例如，对于例 6.3.1 中 LFSR 的周期为 15 的输出序列，取 $i=1$，$\tau=1$，有

$$(a_1a_2\cdots a_{15})=(100100011110101),$$

$$(a_2a_3\cdots a_{16})=(001000111101011),$$

$(a_1a_2\cdots a_{15})$与$(a_2a_3\cdots a_{16})$的异相自相关系数为

$$R(1)=\frac{1}{15}(-1+1-1-1+1+1-1+1+1+1-1-1-1-1+1)=-\frac{1}{15}.$$

Golomb 对伪随机周期序列提出三个条件，称为 Golomb 随机性三公设：

(1) 在序列的一个周期内，0 与 1 的个数相差至多为 1；

(2) 在序列的一个周期内，长为 i 的游程占游程总数的$\frac{1}{2^i}$，且在等长的游程中 0 的游程个数与 1 的游程个数相等；

(3) 异相自相关函数是一个常数.

公设(1)意味着序列中 0 与 1 出现的概率基本上相等；公设(2)说明在序列的任何一位上出现 0 和 1 的概率相同；公设(3)意味着将序列与其平移后的序列做比较，不能给出其他任何信息.

下面的定理说明 m 序列满足 Golomb 的三个公设.

定理 6.3.2 $\mathbf{Z}_2$ 上的 n 位 LFSR 产生的 m 序列具有如下性质：

(1) 在一个周期内，0 出现的次数为 $2^{n-1}-1$，1 出现的次数为 2^{n-1}；

(2) 在一个周期内，总游程数为 2^{n-1}，对 $1\leqslant i\leqslant n-2$，长为 i 的游程有 2^{n-i-1} 个，且 0、1 的游程各半，长为 $n-1$ 的 0 游程有 1 个，长为 n 的 1 游程有 1 个；

(3) $\{a_i\}$的异相自相关函数 $R(\tau)=-\dfrac{1}{2^n-1}$，$(0<\tau\leqslant 2^n-2)$.

证明　(1) 6.1 节已经指出，n 位 LFSR 除了全 0 状态外，最多可能有 2^n-1 不同的状态，每个状态的第一位对应着 LFSR 输出序列中的一位. 在 n 位 LFSR 产生一个周期的 m 序列的过程中，LFSR 遍历 2^n-1 种不同状态各一次，由排列组合的知识可得，这些状态中有 2^{n-1} 个在第一位上取 1，有 $2^n-1-2^{n-1}=2^{n-1}-1$ 个状态在第一位上取 0，结论得证.

(2) 对 $n=1$，2，结论显然成立；考虑 $n>2$ 的情况，当 $1\leqslant i\leqslant n-2$ 时，n 位 LFSR 产生的 m 序列的一个周期中，长为 i 的 0 游程数目等于 LFSR 运行过程中如下形式状态的出现数目：

$$1\underbrace{0\cdots 0}_{i\text{个}0}1*\cdots *$$

其中 $n-i-2$ 个 $*$ 可任取 0 或 1. 这种状态共有 2^{n-i-2} 个，同理可得长为 i 的 1 游程的数目也为 2^{n-i-2} 个，所以长为 i 的游程的总数目为 2^{n-i-1} 个. 由于寄存器中不会出现全 0 状态，所以不会出现 0 的 n 游程，但必有一个 1 的 n 游程，而且不会出现更大的 1 的游程，因为若出现 1 的 $n+1$ 游程，寄存器中就必然有两个相邻的全 1 状态，这与“在产生 m 序列的一个周期的过程中 LFSR 遍历 2^n-1 种不同状态各一次”相矛盾. 于是 1 的 n 游程只能出现在如下的输出序列中：

$$\cdots 0\underbrace{11\cdots 1}_{n\text{个}1}0\cdots$$

而要产生上面的输出序列片段，寄存器必须连续经历以下三个状态：

$$\underbrace{11\cdots 1}_{n-1\text{个}1}0\quad\underbrace{11\cdots 1}_{n\text{个}1}\quad 0\underbrace{11\cdots 1}_{n-1\text{个}1}$$

由于$\underbrace{11\cdots 1}_{n-1\text{个}1}0$，$0\underbrace{11\cdots 1}_{n-1\text{个}1}$这两个状态只能出现一次，它们中间的状态是全 1 状态，所以不会产生 1 的 $n-1$ 游程. 最后，一旦寄存器进入了状态$\underbrace{00\cdots 0}_{n-1\text{个}0}1$，下一个状态必为 $1\underbrace{00\cdots 0}_{n-1\text{个}0}$(因为不能出现全 0 状态)，这两个状态会产生一个 0 的 $n-1$ 游程

$$\cdots 1\underbrace{00\cdots 0}_{n-1\text{个}0}1\cdots$$

综上可见，在 n 位 LFSR 产生的 m 序列的一个周期内，共有

$$1+1+\sum_{i=1}^{n-2}2^{n-i-1}=2^{n-1}$$

个游程.

(3) 设$\{a_i\}$是一周期为 2^n-1 的 m 序列，对于任意一正整数 τ $(0<\tau\leqslant 2^n-2)$，由二元域上的运算性质，$\{a_i\}+\{a_{i+\tau}\}$在一个周期内为 0 的数目正好就是序列$\{a_i\}$与$\{a_{i+\tau}\}$对应位相同的位的数目. 设序列$\{a_i\}$满足递推关系：

$$a_{n+k}=c_1a_{n+k-1}\oplus c_2a_{n+k-2}\oplus\cdots\oplus c_na_k,$$

故

$$a_{n+k+\tau}=c_1a_{n+k+\tau-1}\oplus c_2a_{n+k+\tau-2}\oplus\cdots\oplus c_na_{k+\tau},$$

所以

$$a_{n+k}\oplus a_{n+k+\tau}=c_1(a_{n+k-1}\oplus a_{n+k+\tau-1})\oplus c_2(a_{n+k-2}\oplus a_{n+k+\tau-2})\oplus\cdots\oplus c_n(a_k\oplus a_{k+\tau}).$$

设 $b_j=a_j\oplus a_{j+\tau}$，则由$\{a_i\}$的递推公式可得$\{b_i\}$的递推公式：

$$b_{n+k}=c_1b_{n+k-1}\oplus c_2b_{n+k-2}\oplus\cdots\oplus c_nb_k,$$

所以$\{b_i\}$也是 m 序列. 为了计算 $R(\tau)$，只要将$\{b_i\}$在一个周期内的 0 的个数减去 1 的个数，再除以周期 2^n-1 即可，由(2)的结论可知

$$R(\tau)=-\frac{1}{2^n-1}.$$

实际上，在证明结论(3)时，我们附带证明了如下结论：m 序列满足移位可加性，设 m 序列的周期为 T，从这个 m 序列中任意截取两段长度为 T 的子序列，模 2 相加后得到的子序列必出现在本 m 序列中.

6.3.4 m 序列的破译

我们知道，如果知道了 m 序列的反馈函数，就可以对 m 序列的后继内容进行预测. 所谓 m 序列的破译，是指已知 m 序列的一个子序列，从中求出 m 序列的反馈函数的问题. 设有 n 位 LFSR 产生的 m 序列$\{a_i\}$，已知该 m 序列的一个长度为 $2n$ 的子序列，不妨设该子序列为 $(a_1a_2\cdots a_{2n})$，由式(6.3.2)

$$\begin{bmatrix}a_{n+k}\\a_{n+k-1}\\\vdots\\a_{k+2}\\a_{k+1}\end{bmatrix}=\boldsymbol{L}\cdot\begin{bmatrix}a_{n+k-1}\\a_{n+k-2}\\\vdots\\a_{k+1}\\a_k\end{bmatrix},$$

令

$$\boldsymbol{S}_{k+1}=(a_{n+k}\quad a_{n+k-1}\quad\cdots\quad a_{k+2}\quad a_{k+1})^{\mathrm{T}},$$

则式(6.3.2)可表示为

$$\boldsymbol{S}_{k+1}=\boldsymbol{L}\cdot\boldsymbol{S}_k. \tag{6.3.7}$$

而

$$\begin{aligned}\boldsymbol{S}_1&=(a_1\quad a_2\quad\cdots\quad a_n)^{\mathrm{T}},\\\boldsymbol{S}_2&=(a_2\quad a_3\quad\cdots\quad a_{n+1})^{\mathrm{T}},\\&\vdots\\\boldsymbol{S}_{n+1}&=(a_{n+1}\quad a_{n+2}\quad\cdots\quad a_{2n})^{\mathrm{T}}.\end{aligned}$$

设矩阵

$$\boldsymbol{X}=[\boldsymbol{S}_1\quad\boldsymbol{S}_2\quad\cdots\quad\boldsymbol{S}_n],$$

则有

$$(a_{n+1}\quad a_{n+2}\quad\cdots\quad a_{2n})=(c_n\quad c_{n-1}\quad\cdots\quad c_1)\begin{bmatrix}a_1&a_2&\cdots&a_n\\a_2&a_3&\cdots&a_{n+1}\\\vdots&\vdots&&\vdots\\a_n&a_{n+1}&\cdots&a_{2n-1}\end{bmatrix}, \tag{6.3.8}$$

即

$$\boldsymbol{S}'_{n+1}=(c_n \quad c_{n-1} \quad \cdots \quad c_1)\boldsymbol{X}.$$

如果 $\boldsymbol{X}$ 可逆，则可求得

$$(c_n \quad c_{n-1} \quad \cdots \quad c_1)=\boldsymbol{S}'_{n+1}\boldsymbol{X}^{-1}. \tag{6.3.9}$$

下面证明 $\boldsymbol{X}$ 的确是可逆的.

由序列递推关系

$$a_{n+k}=c_1a_{n+k-1}\oplus c_2a_{n+k-2}\oplus\cdots\oplus c_na_k,$$

可给出向量的递推关系

$$\boldsymbol{S}_{n+k}=c_1\boldsymbol{S}_{n+k-1}\oplus c_2\boldsymbol{S}_{n+k-2}\oplus\cdots\oplus c_n\boldsymbol{S}_k.$$

设 $m(m\leqslant n+1)$是使 $\boldsymbol{S}_1$，$\boldsymbol{S}_2$，…，$\boldsymbol{S}_m$ 线性相关的最小整数，即存在不全为 0 的系数 λ_1，λ_2，…，λ_m，其中不妨设 $\lambda_1\neq0$，使得

$$\lambda_m\boldsymbol{S}_1\oplus\lambda_{m-1}\boldsymbol{S}_2\oplus\cdots\oplus\lambda_2\boldsymbol{S}_{m-1}\oplus\boldsymbol{S}_m=0,$$

成立，即

$$\boldsymbol{S}_m=\lambda_m\boldsymbol{S}_1\oplus\lambda_{m-1}\boldsymbol{S}_2\oplus\cdots\oplus\lambda_2\boldsymbol{S}_{m-1},$$

由式(6.3.7)，对于任一整数 i，有

$$\begin{aligned}\boldsymbol{S}_{m+i}&=\boldsymbol{L}^i\boldsymbol{S}_m=\boldsymbol{L}^i(\lambda_m\boldsymbol{S}_1\oplus\lambda_{m-1}\boldsymbol{S}_2\oplus\cdots\oplus\lambda_2\boldsymbol{S}_{m-1})\\&=\lambda_m\boldsymbol{L}^i\boldsymbol{S}_1\oplus\lambda_{m-1}\boldsymbol{L}^i\boldsymbol{S}_2\oplus\cdots\oplus\lambda_2\boldsymbol{L}^i\boldsymbol{S}_{m-1}\\&=\lambda_m\boldsymbol{S}_{i+1}\oplus\lambda_{m-1}\boldsymbol{S}_{i+2}\oplus\cdots\oplus\lambda_2\boldsymbol{S}_{i+m-1},\end{aligned}$$

由此又推出序列的递推关系

$$a_{m+i}=\lambda_2a_{m+i-1}\oplus\lambda_3a_{m+i-2}\oplus\cdots\oplus\lambda_ma_{i+1},$$

即产生该序列的 LFSR 的位数小于 m，因已知产生该序列的 LFSR 为 n 位，所以 m 只得取 $n+1$，所以 $\boldsymbol{S}_1$，$\boldsymbol{S}_2$，…，$\boldsymbol{S}_n$ 是线性无关的，即 $\boldsymbol{X}$ 可逆. 于是有如下定理.

定理 6.3.3　已知某 m 序列是由 n 位 LFSR 产生的，又知该 m 序列的长度为 $2n$ 的一个子序列，就可以求出该 m 序列的反馈函数.

[例 6.3.2]　已知由 5 位 LFSR 产生的 m 序列的一个子序列为 1101001000，求该 m 序列的反馈函数.

解　由式(6.3.8)

$$(a_6 \quad a_7 \quad a_8 \quad a_9 \quad a_{10})=(c_5 \quad c_4 \quad c_3 \quad c_2 \quad c_1)\begin{bmatrix}a_1 & a_2 & a_3 & a_4 & a_5\\ a_2 & a_3 & a_4 & a_5 & a_6\\ a_3 & a_4 & a_5 & a_6 & a_7\\ a_4 & a_5 & a_6 & a_7 & a_8\\ a_5 & a_6 & a_7 & a_8 & a_9\end{bmatrix},$$

即

$$(0 \quad 1 \quad 0 \quad 0 \quad 0)=(c_5 \quad c_4 \quad c_3 \quad c_2 \quad c_1)\begin{bmatrix}1 & 1 & 0 & 1 & 0\\ 1 & 0 & 1 & 0 & 0\\ 0 & 1 & 0 & 0 & 1\\ 1 & 0 & 0 & 1 & 0\\ 0 & 0 & 1 & 0 & 0\end{bmatrix}.$$

而

$$\begin{bmatrix}1&1&0&1&0\\1&0&1&0&0\\0&1&0&0&1\\1&0&0&1&0\\0&0&1&0&0\end{bmatrix}^{-1}=\begin{bmatrix}0&1&0&0&1\\1&0&0&1&0\\0&0&0&0&1\\0&1&0&1&1\\1&0&1&1&0\end{bmatrix},$$

由式(6.3.9)，有

$$(c_5\quad c_4\quad c_3\quad c_2\quad c_1)=(0\quad 1\quad 0\quad 0\quad 0)\begin{bmatrix}0&1&0&0&1\\1&0&0&1&0\\0&0&0&0&1\\0&1&0&1&1\\1&0&1&1&0\end{bmatrix}=(1\quad 0\quad 0\quad 1\quad 0).$$

该 m 序列的反馈函数为

$$f(a_1,a_2,a_3,a_4,a_5)=a_1\oplus a_4.$$

习题

1. 3 位 LFSR 在 $c_3=1$ 时可有 4 种线性反馈函数，设其初始状态为(1，0，1)，求各线性反馈函数的输出序列及周期.
2. 证明定理 6.2.1.
3. $\mathbf{Z}_2[x]$上的多项式 $x^6+x^4+x^2+1$ 有一个重因子，请求出这个重因子.
4. 证明 $\varphi_n(x)$的次数为 $\varphi(n)$，即 $\deg(\varphi_n(x))=\varphi(n)$.
5. 求域 $\mathbf{Z}_2$ 上的分圆多项式 $\varphi_9(x)$和 $\varphi_{21}(x)$.
6. 在域 $\mathbf{Z}_3$ 上找出两个首一且本原的二次多项式.
7. 求证：如果 2^d-1 是素数，则 $\mathbf{Z}_2[x]$中的任何 d 次不可约多项式必都是本原的.
8. 求证 m 序列满足移位可加性，即 m 序列同相移为任意值的同一 m 序列的模 2 和是另一相移的 m 序列.
9. 求证满足 Golomb 第二公设的伪随机序列在任何一位上出现 0 和 1 的概率相同.
10. (选做). 可以证明，每个 m 序列都存在一个与之对应的反商(或反码)，有时也称为镜像序列. 两者输出的 0、1 序列在周期内输出次序正好是镜像的，并且其特征多项式互为反商本原多项式. n 次本原多项式 $f(x)$的反商本原多项式定义为

    $$f^R(x)=x^n f\left(\frac{1}{x}\right).$$

    请以五次本原多项式为例，验证以上结论.
11. 编程实现由 n 位 LFSR 产生的 m 序列的破译，要求输入 m 序列的长度为 $2n$ 的一个子序列，输出该 m 序列的反馈函数.

参 考 文 献

[1] 柯召，孙琦. 数论讲义. 2 版. 北京：高等教育出版社，2003.

[2] GARRETT P. 密码学导引 . 吴世忠，等译. 北京：机械工业出版社，2003.

[3] 潘承洞，潘承彪. 代数数论. 济南：山东大学出版社，2001.

[4] 杨波. 现代密码学. 北京：清华大学出版社，2003.

[5] 陈恭亮. 信息安全数学基础. 北京：清华大学出版社，2004.

[6] TRAPPE W，WASHINGTON L C. Introduction to cryptography with coding theory. 影印版. 北京：科学出版社，2004.

[7] GOLDREICH O. Foundations of cryptography basic tools. 影印版. 北京：电子工业出版社，2003.

[8] 陈鲁生，沈世镒. 现代密码学. 北京：科学出版社，2002.

[9] 李盘林. 离散数学. 北京：高等教育出版社，1999.

[10] 姜丹. 信息论与编码. 北京：中国科学技术出版社，2002.

[11] MENEZES A. 应用密码学手册. 胡磊，等译. 北京：电子工业出版社，2005.

[12] 王育民、何大可. 保密学：基础和应用. 西安：西安电子科技大学出版社，1990.

[13] HUNGERFORD T W. 代数学. 冯克勤，译. 长沙：湖南教育出版社，1985.

[14] 李继国. 信息安全数学基础. 武汉：武汉大学出版社，2006.

[15] BLAKE I. Elliptic curves in cryptography. Cambridge Univ. Press，2000.

[16] CRANDALL R，POMERANCE C. Prime numbers：a computational perspective. Springer，2001.

[17] FERMIGIER S. Collection of links on research articles on elliptic curves and related topics. http：// www. fermigier. com/fermigier/elliptic. html. en. 2009.

[18] ENDERTON H B. 集合论基础. 影印版. 北京：人民邮电出版社，2006.

[19] BRUALDI R. 组合数学. 冯舜玺，等译. 北京：机械工业出版社，2005.

[20] SILVERMAN J H. 数论概论 . 影印版. 北京：机械工业出版社，2006.

[21] ROSEN K H. 初等数论及其应用. 影印版. 北京：机械工业出版社，2005.

[22] YAN S Y. 计算数论. 影印版. 北京：世界图书出版公司，2004.

[23] ROTMAN J. 抽象代数基础教程. 影印版. 北京：机械工业出版社，2004.