UMSS

大学数学科学丛书 —32

公钥密码学的数学基础

王小云　王明强　孟宪萌 著

科学出版社

大学数学科学丛书 32

公钥密码学的数学基础

王小云　王明强　孟宪萌　著

科学出版社

北京

内 容 简 介

本书是根据作者多年的教学经验，在原有讲义的基础上经过修改、补充而成的. 书中介绍了公钥密码学中涵盖的数论代数基本知识与理论体系：第1章至第6章分别介绍了初等数论基础知识，主要包括同余、剩余类、原根和连分数的基本理论以及在公钥密码中的应用等；第7章至第9章描述了群、环、域三个基本的代数结构及其性质；第10章介绍了与密码学相关的计算复杂性理论及基本数学算法；第11章简单介绍了格理论及格密码分析的基本方法.

本书适合信息安全专业本科生、研究生使用，也适合从事信息安全的工程技术人员和教师参考.

图书在版编目(CIP)数据

公钥密码学的数学基础/王小云，王明强，孟宪萌著. —北京：科学出版社，2013

(大学数学科学丛书；32)

ISBN 978-7-03-035136-4

Ⅰ. ①数… Ⅱ. ①王… ②王… ③孟… Ⅲ. ①数论-高等学校-教材

Ⅳ. ①O156

中国版本图书馆 CIP 数据核字 (2012) 第 161428 号

责任编辑：陈玉琢／责任校对：郑金红

责任印制：钱玉芬／封面设计：陈 敬

科学出版社出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

北京佳艺恒彩印刷有限公司 印刷

科学出版社发行 各地新华书店经销

*

2013 年 1 月第 一 版 开本：B5(720×1000)

2013 年 1 月第一次印刷 印张：10 3/4

字数：195 000

定价：48.00 元

(如有印装质量问题，我社负责调换)

作 者 简 介

王小云, 教授, 1966 年出生, 1983 年至 1993 年就读于山东大学数学系, 先后获得学士、硕士和博士学位, 博士生导师潘承洞教授. 1993 年毕业后留校任教. 现为清华大学杨振宁讲座教授, 中国密码学会副理事长. 2005 年国家杰出青年基金获得者, 2006 年被聘为清华大学 "长江学者特聘教授". 主要研究方向是密码理论研究. 在密码分析领域, 给出了多个重要 Hash 函数算法 MD5 与 SHA-1 等的碰撞攻击.

王明强, 博士, 1970 年生, 2004 于山东大学数学系获得博士学位, 导师展涛教授. 现为山东大学副教授, 中国密码学会会员. 主要研究方向是数论、算术几何, 在可证明安全密码体质研究及椭圆曲线密码快速实现方面取得多个重要研究成果.

孟宪萌, 博士, 1971 年生, 1989 年起先后就读于吉林大学数学系和山东大学数学系获学士、硕士和博士学位, 攻读硕士博士学位期间的导师为展涛教授. 毕业后从事教学与科研工作, 现为山东财经大学教授, 中国密码学会会员. 主要研究方向是数论与密码, 在数论中的加性问题研究以及公钥密码算法 RSA 的安全性分析方面取得多个重要研究成果.

《大学数学科学丛书》序

按照恩格斯的说法, 数学是研究现实世界中数量关系和空间形式的科学. 从恩格斯那时到现在, 尽管数学的内涵已经大大拓展了, 人们对现实世界中的数量关系和空间形式的认识和理解已今非昔比, 数学科学已构成包括纯粹数学及应用数学内含的众多分支学科和许多新兴交叉学科的庞大的科学体系, 但恩格斯的这一说法仍然是对数学的一个中肯而又相对来说易于为公众了解和接受的概括, 科学地反映了数学这一学科的内涵. 正由于忽略了物质的具体形态和属性、纯粹从数量关系和空间形式的角度来研究现实世界, 数学表现出高度抽象性和应用广泛性的特点, 具有特殊的公共基础地位, 其重要性得到普遍的认同.

整个数学的发展史是和人类物质文明和精神文明的发展史交融在一起的. 作为一种先进的文化, 数学不仅在人类文明的进程中一直起着积极的推动作用, 而且是人类文明的一个重要的支柱. 数学教育对于启迪心智、增进素质、提高全人类文明程度的必要性和重要性已得到空前普遍的重视. 数学教育本质是一种素质教育; 学习数学, 不仅要学到许多重要的数学概念、方法和结论, 更要着重领会数学的精神实质和思想方法. 在大学学习高等数学的阶段, 更应该自觉地去意识并努力体现这一点.

作为面向大学本科生和研究生以及有关教师的教材, 教学参考书或课外读物的系列, 本丛书将努力贯彻加强基础、面向前沿、突出思想、关注应用和方便阅读的原则, 力求为各专业的大学本科生或研究生(包括硕士生及博士生)走近数学科学、理解数学科学以及应用数学科学提供必要的指引和有力的帮助, 并欢迎其中相当一些能被广大学校选用为教材, 相信并希望在各方面的支持及帮助下, 本丛书将会愈出愈好.

李大潜

2003 年 12 月 27 日

序

密码学作为信息安全的支撑学科, 是一门起源十分古老而在近代得到极其广泛有效应用、正在蓬勃发展的新兴学科. 密码学的两个基本问题是: 一是对信息加以保密, 要使第三方获得了加密的信息后, 也不知道信息的真实内容, 二是与此相对立的, 当第三方获得了加密的信息后, 可以设法破解从中获得信息的真实内容. 许许多多的方法都可以用于密码学, 对信息加以保密和破解, 数学一直是密码学的重要工具. 特别是提出公钥密码系统以来, 数学在密码学中重要性取得了无可争辩的地位, 也可以说密码学从此成为了数学中的一个独特学科.

王小云教授一直十分重视信息安全专业的人才培养和基础课的教材建设. 《数论与代数结构》是密码学的一门重要基础课. 早在 2003 年, 她就开始编写《数论与代数结构》讲义, 经过在山东大学和清华大学的教学试用, 不断充实改进, 取得很好的效果. 本书就是在这基础上, 与王明强、孟宪萌一起合作完成的. 本书内容作者在前言中作了详细介绍, 这里就不多说了. 我想特别指出, 本书的特点是: 将密码学中的算法及其复杂性理论与数论、代数的一些基本理论有机的密切结合在一起、贯穿全书. 这对培养密码学工作者的独特思维方式具有极为重要作用. 我觉得, 数学中十分有效的 '形式化' 思维在密码学中很难直接有效, 而 '形式化' 思维方式似乎并不能抓住密码学问题的本质. 这是数学工作者转向密码学研究时需要特别注意的一个关键问题. 因此, 本书对提高信息安全专业基础课的教学质量将有十分积极的作用.

王小云教授是家兄承洞的博士生. 承洞从学生时代开始就一直重视和提倡数学在科学技术中的应用, 他曾经参与过渗流理论、薄壳理论、样条理论及其应用和定向爆破等的初步研究和实践. 特别是在 1990 年前后, 他在山东大学开始成立密码研究小组并参与密码学理论及其应用的研究和人才培养, 取得了良好成果, 王小云教授就是最优秀的一位. 她在密码分析领域, 提出并建立了 Hash 函数碰撞攻击的理论与技术, 成功破解了由图灵奖获得者 Rivest 设计的 MD5, 和由美国国家标准与技术研究所 (NIST) 与安全局 (NSA) 设计的 SHA-1 等国际主要的一类基础密码算法 —Hash 函数, 它们是常用的电子签名与数字证书的核心技术. 这一杰出成就震动了密码学界, 迫使美国国家标准与技术研究所于 2007 年启动新 Hash 函数标准 SHA-3 的五年设计工程. 她在消息认证码分析领域和密码设计领域都取得重大成果, 她主持设计的 Hash 函数算法 SM3, 被采纳为国家实用化算法.

我一点也不懂密码学, 但我非常高兴能为王小云的书写上几句不一定确切的外

行话, 我觉得这是我的责任与义务. 我希望她安心搞科研、热心搞教学, 不计名利, 为密码学的科研和人才培养作出更大贡献.

潘承彪

2012 年 8 月 3 日

前　言

自 1976 年 Diffie 和 Hellman 提出公钥密码的思想以来, 密码学家设计了多个具有代表性的公钥密码算法. 这些密码算法的安全性均基于一些经典数学难题求解的困难性, 如因子分解问题、离散对数问题、背包问题以及格中的最短向量问题等. 而公钥密码算法分析的核心就是研究这些数学难题的快速求解算法. 为了更好地让信息安全专业的学生顺利学习、掌握现代密码学的基本理论, 深刻领会密码学与数学领域的学科交叉特点, 特编写了《公钥密码学的数学基础》作为信息安全专业的数学基础课教材. 本书所涉及的理论知识都是现代密码学特别是公钥密码学所需要的数学基础知识, 不仅可以作为信息安全专业本科生教学的教材, 也是密码科技工作者必要的专业参考书.

2003 年, 山东大学信息安全专业设立之初, 作者就着手撰写《数论与代数结构》的讲义是现代密码学特别是公钥密码学所需要的数学基础知识, 本书不是初等数论和抽象代数的简单组合, 而是反映信息安全学科交叉特点, 并体现数学理论与密码应用相结合的教材. 本书的内容主要有以下三方面的特色: 一是数论与代数基本理论涵盖了一些重要的密码基础数学理论. 如我们既介绍辗转相除法、Euler 定理、孙子定理、原根等初等数论基本理论, 也讲述了在密码学中广泛使用的利用辗转相除法求最大公因子、求模逆模幂运算、离散对数、因子分解等密码基础数学理论. 二是注重理论与实践的紧密结合, 并突出实践. 在讲到比较重要的算法时, 我们都配备一定数量的实践题目, 使学生能体会到理论在实践中的应用. 三是将算法复杂性理论贯穿全书, 介绍与数论代数基本理论相关的算法及其复杂性, 让读者初步体会数学理论在密码算法中的应用.

全书分为 11 章. 第 1 章至第 6 章分别介绍了初等数论的基本理论和工具: 同余、原根、剩余类、连分数等. 原根的理论是 Diffie 和 Hellman 公钥密码算法的理论基础, 连分数在 RSA 公钥算法的分析和因子分解问题中都有重要的应用. 第 7 章至第 9 章介绍了抽象代数的基本概念, 给出了群、环、域三个基本的代数结构及其性质, 重点介绍了在大数乘法及密码快速实现方面有重要应用的中国剩余定理. 第 10 章介绍了计算复杂度的基本理论及密码学相关的基本数学算法: 素判定问题、离散对数问题、因子分解问题. 第 11 章是格理论的简单介绍及格基约化算法 ——LLL 算法在公钥密码算法 RSA 分析中的应用.

本书是经过山东大学信息安全专业教学中多次使用并反馈修改的结果, 这对本书的最终完成具有十分重要的意义. 在本书的写作过程中, 初等数论部分重点参

考了潘承洞与潘承彪两位教授出版的《初等数论》教材; 抽象代数部分参考了吴品三、张禾瑞以及刘绍学三位教授分别出版的《近世代数》教材. 本书的出版得到了教育部信息安全特色专业建设项目以及国家自然科学基金重点项目 (No.61133013) 的资助; 郑世慧、冯骐、魏普文等提出许多宝贵的修改意见, 在此表示衷心的感谢!
限于作者水平, 本书难免存在不足之处, 敬请读者批评指正.

作　者

2012 年 7 月

目　　录

第1章 整 除

整除是数论中的基本概念. 本章主要介绍与整除相关的一些基本概念及其性质. 这些基本概念如整除、因子、公因子、最小公倍数、分解因子等, 早在中学时期已被大家所熟悉. 在这里我们将给出这些概念的严格的数学定义. 通过掌握这些概念的数学定义及相关性质, 我们可以进一步解决许多初等数论里与整除相关的问题. 整除理论内容丰富, 解决问题方法灵活. 它不仅是数论、代数的基础, 而且在密码学中有很广泛的应用, 如整数的素分解、Euclid 算法求最大公因子等问题在密码学中都有极其重要的应用.

§1.1 整除的概念

我们用集合 $\mathbb{Z}$ 表示全体整数组成的集合, $\mathbb{N}$ 表示自然数的全体. 下面给出整除的定义.

定义 1.1 设 $a, b \in \mathbb{Z}$, 如果存在 $q \in \mathbb{Z}$, 使得 $b = aq$, 那么, 就说 b 可被 a 整除, 记作 $a|b$, 称 b 是 a 的倍数, a 是 b 的因子 (也可称为约数、除数). 否则就说 b 不能被 a 整除, 或 a 不整除 b, 记作 $a \nmid b$.

由定义及乘法的运算规律, 立即可得出整除的以下性质:

定理 1.2 设 a, b, $c \in \mathbb{Z}$, 则

(1) $a|b$ 且 $b|c \Rightarrow a|c$;

(2) $a|b$ 且 $a|c \Leftrightarrow$ 对任意的 $x, y \in \mathbb{Z}$ 有 $a|bx + cy$;

(3) 若 $m \in \mathbb{Z}$ 且 $m \neq 0$, 则 $a|b \Leftrightarrow ma|mb$;

(4) $a|b$ 且 $b|a \Rightarrow a = \pm b$;

(5) 若 $b \neq 0$, 则 $a|b \Rightarrow |a| \leqslant |b|$.

证明 (1) 由于 $a|b$, 根据整除的定义知存在 q_1, 使 $b = aq_1$. 同样存在 q_2 使得 $c = bq_2$, 从而

$$c = q_2 b = (q_1 q_2)a,$$

即 $a|c$.

(2) 由 $a|b$,$a|c$ 知, 存在 r, s 使得 $b = ar, c = as$. 对任意的 $x, y \in \mathbb{Z}$ 有

$$bx + cy = arx + asy = a(rx + sy),$$

故 $a|bx + cy$. □

性质 (3), (4), (5) 证明类似, 读者可以自己补出证明.

显然 $\pm 1, \pm b$ 是 b 的因子, 我们称其为 b 的显然因子, 其他因子称为 b 的非显然因子, 或真因子. 由此我们可引出素元的定义.

定义 1.3　*设整数 $p \neq 0, \pm 1$. 如果 p 除了显然因子 $\pm 1, \pm p$ 外没有其他的因子, 那么 p 就称为素元 (常称为素数), 若整数 $a \neq 0, \pm 1$, 且 a 除显然因子外还含有真因子, 则称 a 为合数.*

注　一般情况下, 素数我们只取正的.

下面我们介绍几个关于素数的定理.

定理 1.4　*若 a 为合数, 则 a 的最小真因子为素数.*

证明　由 a 为合数知 $a > 2$. 设 d 为 a 的最小真因子. 若 d 不为素数, 则存在 d 的真因子 d', 使 $d'|d$, 由性质 (1) 知 $d'|a$, 与 d 为最小真因子矛盾. 定理得证. □

定理 1.5　*素数有无穷多个.*

证明　假设只有有限个素数, 设为 $p_1, p_2, \cdots, p_k$. 考虑 $a = p_1 p_2 \cdots p_k + 1$, 由定理 1.4 知, 整除 a 的最小真因子一定为素数, 记为 p. 由于 p 为素数, 因而 p 必等于某个 p_i, 所以 $p|a$, $p|p_1 p_2 \cdots p_k$ 同时成立, 从而 $p|1$, 这与 p 是素数矛盾. 因此定理得证. □

将素数从小到大排列, 假设 p_n 表示第 n 个素数, $\pi(x)$ 表示不超过 x 的素数个数. 虽然我们无法知道 p_n 的确切位置, 但是, 我们可以得到 p_n 的弱上界估计. 下面定理仅描述了 p_n 的一个弱上界估计与 $\pi(x)$ 的一个弱下界估计. 而对于 $\pi(x)$ 更为精确的估计, 超出本书的讨论范围, 有兴趣的读者可参考文献 [20].

定理 1.6　*将全体素数按从小到大的顺序排列, 则第 n 个素数 p_n 与 $\pi(x)$ 分别有以下结论:*

(1) $p_n \leqslant 2^{2^{n-1}}, n = 1, 2, \cdots$;

(2) $\pi(x) > \log_2 \log_2 x, x \geqslant 2$.

证明　(1) 我们用归纳法来证明定理的第一部分成立. 当 $n = 1$ 时, 命题显然成立. 假设对于 $n \leqslant k$ 时, 命题成立. 当 $n = k + 1$ 时, 由定理 1.4 知, $p_{k+1} \leqslant p_1 p_2 \cdots p_k + 1$, 由归纳假设

$$p_{k+1} \leqslant 2^{2^0} 2^{2^1} \cdots 2^{2^{k-1}} + 1 = 2^{2^k - 1} + 1 < 2^{2^k}.$$

于是命题 1 得证.

(2) 对于任意的 $x \geqslant 2$, 必存在唯一的整数 n, 使得 $2^{2^{n-1}} \leqslant x < 2^{2^n}$, 从而由第一部分结论得

$$\pi(x) \geqslant \pi(2^{2^{n-1}}) \geqslant n > \log_2 \log_2 x.$$

定理得证. □

初等数论还有一个最基本的结论：带余除法定理, 它是整除的一般情形, 也是 Euclid 算法的基础.

定理 1.7 *设 a, b 是两个给定的整数且 $a \neq 0$. 那么一定存在唯一的一对整数 q 与 r, 满足*

$$b = qa + r, \quad 0 \leqslant r < |a|,$$

其中, r 被称为 b 被 a 除后的最小非负余数. 此外 $a|b$ 的充要条件是 $r = 0$.

证明 唯一性：若还有整数 q' 与 r' 满足

$$b = aq' + r', \quad 0 \leqslant r' < |a|.$$

则有 $0 \leqslant |r' - r| < |a|$, 及 $r' - r = (q - q')a$. 由整除的性质立即可得 $r' - r = 0$, 所以唯一性成立.

存在性：当 $a|b$ 时, 可取 $q = \dfrac{b}{a}$, $r = 0$. 当 $a \nmid b$ 时, 考虑集合

$$T = \{b - ka | k = \pm 1, \pm 2, \pm 3, \cdots\},$$

容易看出, 集合 T 中必有正整数 (例如, 取 $k = \dfrac{-2|ab|}{a}$), T 中必有一个最小正整数, 记为

$$t_0 = b - k_0 a > 0.$$

我们来证明必有 $t_0 < |a|$. 因为 $a \nmid b$, 所以 $t_0 \neq |a|$. 若 $t_0 > |a|$, 则 $t_1 = t_0 - |a| > 0$. 显见 $t_1 \in T$, $t_1 < t_0$, 这和 t_0 的最小性矛盾. 取 $q = k_0$, $r = t_0$ 就满足要求. 显然, $a|b$ 的充要条件是 $r = 0$. 定理得证. □

下面利用带余除法对整数进行分类. 设 $a \geqslant 2$ 是给定的正整数, $j = 0, 1, 2, \cdots, a - 1$. 对给定的 j, 被 a 除后余数等于 j 的全体整数是

$$ak + j, \quad k = \pm 1, \pm 2, \pm 3, \cdots,$$

这些整数组成的集合记为 $S_{a,j}$. 集合 $\{S_{a,j} : 0 \leqslant j \leqslant a - 1\}$ 满足以下两个性质：

(1) $\{S_{a,j} : 0 \leqslant j \leqslant a - 1\}$中的任两个 $S_{a,j}$ 两两不相交, 即

$$S_{a,j} \cap S_{a,j'} = \varnothing, \quad 0 \leqslant j \neq j' \leqslant a - 1.$$

(2) $\{S_{a,j} : 0 \leqslant j \leqslant a - 1\}$ 中所有子集的并等于 $\mathbb{Z}$, 即

$$\bigcup_{0 \leqslant j \leqslant a-1} S_{a,j} = \mathbb{Z}.$$

这样按被 a 除后所得不同的最小非负余数, 将全体整数分成了两两不相交的 a 个类. 这种分类对于一些数学问题的处理会带来很大的方便.

例 1.8 对于任意整数 x, x^3 被 9 除后所得的最小非负余数是 0, 1, 8.

证明　对于任意的整数 x, 存在 $0 \leqslant j \leqslant 8$ 使得 $x \in S_{9,j}$. 因此只需检验 0 至 8 之间的数即可,

$$\begin{aligned} 0^3 &= 0 \times 9 + 0, \quad 1^3 = 0 \times 9 + 1, \quad 2^3 = 0 \times 9 + 8, \\ 3^3 &= 3 \times 9 + 0, \quad 4^3 = 7 \times 9 + 1, \quad 5^3 = 13 \times 9 + 8, \\ 6^3 &= 24 \times 9 + 0, \quad 7^3 = 38 \times 9 + 1, \quad 8^3 = 56 \times 9 + 8. \end{aligned}$$

例题得证. □

例 1.9　设 $a \geqslant 2$ 是给定的正整数. 那么, 任一正整数 n 必可唯一表示为

$$n = r_k a^k + r_{k-1} a^{k-1} + \cdots + r_1 a^1 + r_0,$$

其中整数 $k \geqslant 0$, $0 \leqslant r_j \leqslant a - 1 (0 \leqslant j \leqslant k)$, $r_k \neq 0$. 这就是正整数 n 的 a 进制表示.

证明　对正整数 n 必有唯一的 $k \geqslant 0$, 使 $a^k \leqslant n < a^{k+1}$. 由带余除法知, 必有唯一的 q_0, r_0 满足

$$n = q_0 a + r_0, \quad 0 \leqslant r_0 < a.$$

若 $k = 0$, 则必有 $q_0 = 0, 1 \leqslant r_0 < a$, 所以结论成立. 假设 $k = m \geqslant 0$ 时结论成立. 那么当 $k = m + 1$ 时, 上式中的 q_0 必满足

$$a^m \leqslant q_0 < a^{m+1}.$$

由假设知

$$q_0 = s_m a^m + \cdots + s_0,$$

其中 $0 \leqslant s_j \leqslant a - 1 (0 \leqslant j \leqslant m - 1)$, $1 \leqslant s_m \leqslant a - 1$. 因而有

$$n = s_m a^{m+1} + \cdots + s_0 a + r_0.$$

即结论对 $m + 1$ 也成立. 定理得证. □

例 1.10　将 10 进制整数 27182 写成 12 进制的形式.

解　进行下面的算法

$$\begin{aligned} 27182 &= 12 \cdot 2265 + 2 \\ 2265 &= 12 \cdot 188 + 9 \\ 188 &= 12 \cdot 15 + 8 \\ 15 &= 12 \cdot 1 + 3 \\ 1 &= 12 \cdot 0 + 1 \end{aligned}$$

得

$$27182_{10} = 13892_{12}.$$