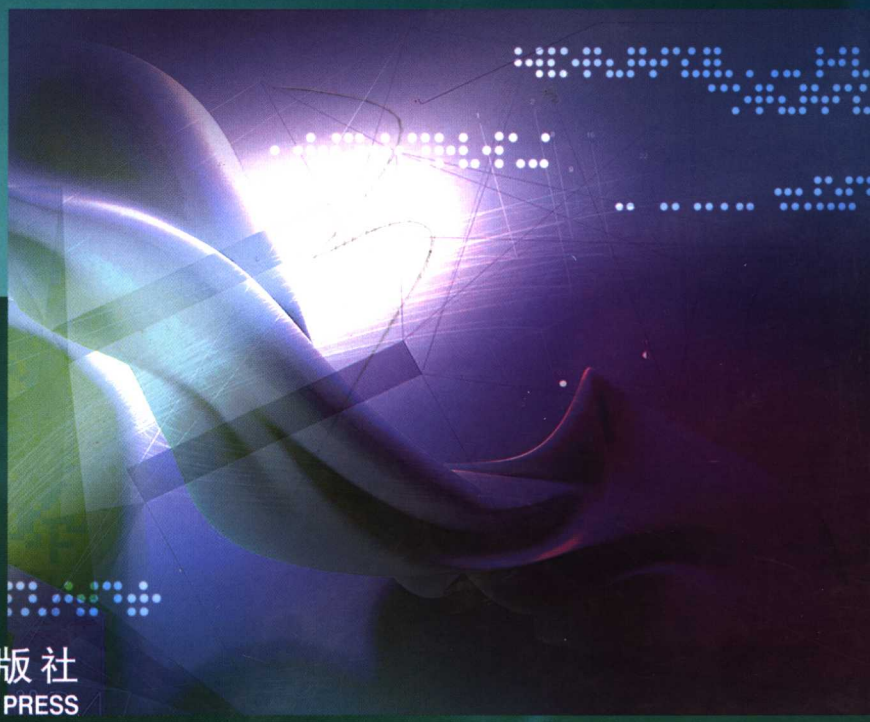
普通高等学校信息与计算科学专业系列丛书

教育科学“十五”国家规划课题研究成果

编码理论基础

陈鲁生　沈世镒　编



高等教育出版社
HIGHER EDUCATION PRESS

普通高等学校信息与计算科学专业系列教材

教育科学“十五”国家规划课题研究成果

编码理论基础

陈鲁生　沈世镒　编

高等教育出版社
HIGHER EDUCATION PRESS

内容提要

本书是关于编码理论的一本教材，主要介绍编码理论的基本知识。全书共十二章，可以分为两部分。第一部分是第二章至第四章，主要介绍编码理论中用到的代数基本知识，特别是有限域的基本知识。第二部分是第五章至第十二章，主要介绍编码理论的基本知识，包括线性码、Hamming 码、Golay 码、循环码、BCH 码、Reed－Muller 码以及线性码的重量分布等。

本书适合高等院校的信息科学、计算机科学以及通信等专业的本科生作为教材使用，也可供相关领域的科研人员和工程技术人员参考。

图书在版编目(CIP)数据

编码理论基础／陈鲁生，沈世镒编．—北京：高等教育出版社，2005.1
ISBN 7-04-016138-9

Ⅰ.编… Ⅱ.①陈…②沈 Ⅲ.编码理论—教材
Ⅳ.0157.4

中国版本图书馆CIP数据核字(2004)第143340号

策划编辑 王 瑜 **责任编辑** 赵天夫
封面设计 王凌波 **责任印制** 韩 刚

出版发行	高等教育出版社	**购书热线**	010－64054588
社 址	北京市西城区德外大街4号	**免费咨询**	800－810－0598
邮政编码	100011	**网 址**	http://www.hep.edu.cn
总 机	010－58581000		http://www.hep.com.cn
经 销	新华书店北京发行所		
印 刷	廊坊市文峰档案文化用品有限公司		
开 本	787×960 1/16	**版 次**	2005年1月第1版
印 张	14.25	**印 次**	2005年1月第1次印刷
字 数	260 000	**定 价**	18.20元

本书如有缺页、倒页、脱页等质量问题，请到所购图书销售部门联系调换。

物料号：16138－00

信息与计算科学专业系列教材编委会

总 序

根据教育部 1998 年颁布的普通高等学校专业目录，“信息与计算科学”专业被列为数学类下的一个新专业 (它覆盖原有的计算数学及其应用软件、信息科学与运筹控制等专业). 这一新专业的设置很好地适应了新世纪以信息技术为核心的全球经济发展格局下的数学人才培养与专业发展的需要. 然而，作为一个新专业，对其专业内涵、专业规范、教学内容与课程体系等有一个自然的认识与探索过程. 教育部数学与统计学教学指导委员会数学类专业教学指导分委员会 (下称教指委) 经过过去两年艰苦细致的工作，对这些问题现在已有了比较明确的指导意见，发表了《关于信息与计算科学专业办学现状与专业建设相关问题的调查报告》及《信息与计算科学专业教学规范》 (讨论稿)(见《大学数学》第 19 卷 1 期 (2003)). 为此，全国高等学校教学研究中心在承担全国教育科学“十五”国家级规划课题 ——“21 世纪中国高等教育人才培养体系的创新与实践”研究工作的基础上，根据教指委所颁布的新的教学规范，组织国内各高校的专家教授，进行其子项目课题“21 世纪中国高等学校信息与计算科学专业教学内容与课程体系的创新与实践”的研究与探索. 为推动本专业的教材建设，该项目课题小组与高等教育出版社联合成立了“信息与计算科学专业系列教材编委会”，邀请有多年教学和科研经验的教师编写系列教材，由高等教育出版社独家出版，并冠以教育科学“十五”国家规划课题研究成果.

按照新的《信息与计算科学专业教学规范》 (讨论稿), 信息与计算科学专业是以信息技术和计算技术的数学基础为研究对象的理科类专业. 其目标是培养学生具有良好的数学基础和数学思维能力，掌握信息与计算科学基础理论、方法与技能，受到科学研究的训练，能解决信息技术和科学与工程计算中的实际问题的高级专门人才. 毕业生能在科技、教育、信息产业、经济与金融等部门从事研究、教学、应用开发和管理工作，能继续攻读研究生学位. 根据这一专业目标定位和落实“强基础、宽口径、重实际、有侧重、创特色”的办学指导思想，我们认为，本专业在数学基础、计算机基础、专业基础方面应该得到加强，而各学校在这三个基础方面可大体一致，但专业课 (含选修课) 允许各校自主选择、体现各自特点. 考虑到已有大量比较成熟的数学基础与计算机基础课程教材，本次教材编写主要侧重于专业基础课与专业课 (含选修课) 方面.

信息与计算科学，就其范畴与研究内容而言，是数学、计算机科学和信息

工程等学科的交叉，已远远超出数学学科的范畴．但作为数学学科下的一个理科专业，信息与计算科学专业则主要研究信息技术的核心基础与运用现代计算工具高效求解科学与工程问题的数学理论与方法（或更简明地说，研究定向于信息技术与计算技术的数学基础），这一专业定位明显地与计算机科学与信息工程专业构成区别．基于这一定位，信息与计算科学专业可包括信息科学与科学计算（计算数学）两个大的方向．科学计算方向在我国已有长期的办学经验，通常被划分为偏微分方程数值解、最优化理论与方法、数值逼近与数值代数、计算基础等学科子方向．然而，对于信息科学，它到底应该怎样划分学科子方向？应该怎样设置专业与专业基础课？所有这些都仍是正在探索的问题．

任何技术都可以认为是延伸与扩展人的某种功能的方式与方法，所以信息技术可以认为是扩展人的信息器官功能的技术．人的信息器官主要包括感觉器官、传导器官（传导神经网络）、思维器官和效应器官四大类型，其功能则主要是信息获取、信息传输、信息处理和信息应用（控制），因而感测技术、通信技术、智能技术与控制技术通常被认为是最基本的信息技术（常称之为信息技术的四基元），其它信息技术可认为是这四种基本技术的高阶逻辑综合或分解衍生．所以可以把信息科学理解为是“有关信息获取、信息传输、信息处理与信息控制基础的科学”．从这个意义上，我们认为：信息处理（包括图像处理、信号分析等）、信息编码与信息安全、计算智能（人工智能、模式识别等）、自动控制等可构成信息科学的主要学科子方向．这一认识也是教指委设置信息与计算科学专业信息科学方向课程的基本依据．

本系列教材正是基于以上认识，为落实新的《信息与计算科学专业教学规范》（讨论稿）而组织编写的．我们相信，该系列教材的出版对缓解本专业教材的紧缺局面，对推动信息与计算科学专业的快速与健康发展会大有裨益．

从长远的角度看，为适应不同类型院校和不同层次要求的课程需求，本系列教材编委会还将不断组织教材的修订和编写新的教材，从而使本专业的教学用书做到逐步充实、完善和多样化．我们诚恳希望采用本系列教材的教师、同学们及广大读者对书中存在的问题及时指正并提出修改意见和建议．

信息与计算科学专业系列教材编委会

2003 年 8 月 31 日

前　言

在通信系统和计算机系统中，提高信息传输的可靠性，始终是我们追求的目标．纠错编码是提高信息传输可靠性的一种重要方法．

多年来，我们一直在南开大学数学科学学院为信息科学专业的本科生讲授“代数与编码”课程．本书就是在此基础上编写而成的．纠错编码理论从创始至今已经发展成为一个具有丰富成果和重要应用价值的研究方向．作为一本大学本科生的教材，本书讲授的只是纠错编码理论的一些最基本的内容，不可能包含纠错编码理论的全部内容，实际上我们也没有打算这样做．本书只是纠错编码理论的一个入门．

本书共十二章，可以分为两部分．第一部分是第二章至第四章，主要讲授纠错编码理论中用到的一些基本的代数知识，特别是有限域上的代数知识．第二部分是第五章至第十二章，主要讲授纠错编码理论的基本知识．

在本书的编写过程中，我们力求简明扼要，容易理解．关于本书中的定理，绝大多数我们都给出了证明．没有给出证明的定理分为两种情况．第一种情况为，根据已有的定义和结论，定理的证明是显然的．第二种情况为，定理的证明比较繁琐，或者证明过程需要更多的基础知识．对于第一种情况，读者可以作为练习自行证明；对于第二种情况，对证明感兴趣的读者可以参阅有关的参考文献．

在本书各章的后面，我们都给出了少量的习题．在本书的最后我们给出了习题的解答．对于习题，我们强烈建议初次学习纠错编码理论的读者自己思考给出解答．我们给出的习题解答只是提供一个参考．

众所周知，要想掌握编码理论，需要有比较全面的数学基础知识．许多不同的数学领域都在编码理论中发挥了它们的作用．我们假定本书的读者具备了高等代数、概率论以及简单的组合数学等基本知识．

本书适合高等院校的信息科学、计算机科学以及通信等专业的本科生作为教材使用，也可供相关领域的科研人员和工程技术人员参考．

由于时间仓促，本书难免有疏漏和不当之处，敬请读者批评指正．

郑重声明

目 录

第一章 引言

编码理论 (Coding Theory) 始于 1948 年 Claude Shannon 在 Bell System Technical Journal 上发表的著名论文 "A Mathematical Theory of Communication". Shannon 指出，对于任意给定的一个通信信道，一定存在码率 R 不大于信道容量 C 并且尽可能接近 C 的纠错码，使得在信道接收端译码错误的概率任意小. 遗憾的是，Shannon 并没有告诉我们如何寻找这样的纠错码. Shannon 的成就在于证明了这样的纠错码是存在的. 这就很自然地促使人们试图来构造这样的纠错码.

纠错码是提高信息传输可靠性的一种重要手段. 纠错编码理论的早期研究热点是以 BCH 码和 Reed-Solomon 码为代表的分组码以及卷积码等. 20 世纪 80 年代的研究热点是代数几何码和网格编码. 现在的研究热点则是 Turbo 码和 LDPC 码 (Low-Density Parity-Check codes).

纠错编码理论的一个重要应用领域是通信. 早期的许多编码理论研究人员都在贝尔电话实验室 (Bell Telephone Laboratories) 工作过. 因此，大量编码理论的早期文献可以在 Bell System Technical Journal 上找到，这是毫不奇怪的. 通过查阅参考文献，不难发现，编码理论的许多最重要的结果都是发表在 IEEE Transactions on Information Theory 上. 这是有关信息论的一本最重要的学术期刊.

§1.1 通信系统

简单地说，一个数字通信系统主要由信源、信道、信道编码器以及信道译码器四个基本部分组成，如图 1.1 所示.

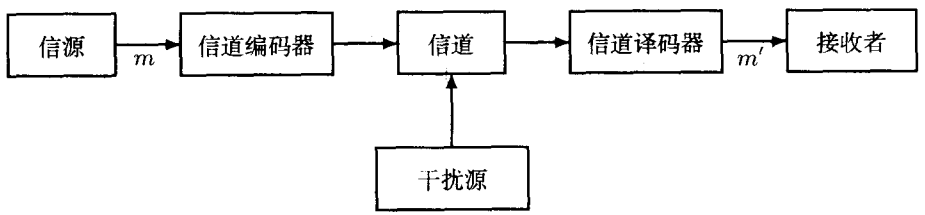


图 1.1 通信系统

信息的产生和发送者称为信源 (source). 我们一般称由信源产生的信息为消息 (message). 由信源输出的消息要经过某种通信渠道传送给称为信宿的接收者 (receiver). 譬如, 在一个电话通信系统中, 打电话的用户就是这个通信系统的信源, 而接听电话的用户就是这个通信系统的信宿.

所谓信道 (channel) 就是将信源消息传送给接收者的渠道. 譬如在有线电话通信系统中, 电话线就是信道. 实际的信道可以是电缆、光纤、高频无线电连接、卫星通信连接等等.

信源发出的消息在信道传输的过程中, 可能会遇到各种干扰, 从而使传输的信源消息产生失真. 对信道的干扰通常称为噪声 (noise). 噪声可以是人为的错误、闪电、热噪声、设备的不完善等等. 噪声会使信道上传输的数据发生错误.

信道编码器 (encoder) 将信源输出的消息编码成为一个码字 (codeword). 由于噪声的干扰, 码字在信道的传输过程中可能会产生错误. 信道译码器 (decoder) 接收信道的输出, 并试图纠正在信道传输过程中产生的错误, 然后从已经纠错的码字恢复原始的信源消息.

另外, 我们将存储系统也看做是一个特殊的通信系统. 对于存储系统我们也可以用图 1.1 来进行描述. 譬如计算机存储系统、录音机存储系统等等. 这时的信道就是存储介质, 信道的输入就是往存储介质上写入信息, 信道的输出就是从存储介质上读取信息.

§1.2 编码理论的主要目标

一个实用的通信系统要求在信道接收端能够尽可能准确地复制出信源发送的消息. 但在实际的信道中, 由于噪声的干扰, 信源消息在信道传输过程中会发生一些错误, 使得在信道接收端接收到的信息与信源发送的消息有一定的误差.

为了增强信源消息在信道传输过程中的抗干扰能力, 就需要增加信源消息的冗余度, 这是信道编码器的任务. 经过信道编码器增加了冗余度的信源消息, 在信道的传输过程中, 即使受到噪声的干扰产生了一些错误, 在信道接收端的信道译码器也可以利用增加的冗余信息来进行纠错, 从而尽可能正确地复制出信源消息.

纠错编码理论 (Error-Correcting Coding Theory) 的主要目标之一, 就是寻找能够纠正尽可能多的错误并且冗余度不大的好码. 这些好码还应该具有编码和译码速度快等特性.

§1.3 编码理论的应用

编码理论的一个重要应用领域当然是通信系统. 譬如卫星通信、电话通信、计算机网络中计算机之间的信息传输等等, 都可以借助纠错码来达到可靠的通信. 在军事通信中, 我们也常常使用纠错码来防止敌人的有意干扰.

许多通信系统对发射功率都有限制. 譬如在通信中继卫星中, 功率是非常昂贵的. 纠错码是降低所需功率的一种最好的方法, 因为在接收端接收到的微弱信号可以借助于纠错码来正确地得到恢复.

编码理论的另一个重要应用领域是计算机系统. 在计算机存储器、数字磁带、磁盘以及光盘中, 可以利用纠错码来保护数据. 在数字逻辑电路中, 利用纠错码可以纠正由于电路失常而造成的数据错误.

对于计算机系统中的数据传输, 因为一个差错就会破坏整个计算机程序, 所以一般只能允许极低的差错率. 因此, 纠错码在计算机系统中的应用显得越来越重要.

在商业领域中, 大家在商品包装上常见的条形码也是一种纠错码. 条形码通常称为图形码, 它利用图形来记录和表达信息. 条形码由一组黑白相间的条纹组成, 黑白条纹的不同宽度代表不同的信息. 条形码中包含了一定的纠错信息, 可以纠正由于条码的模糊不清等原因造成的读取错误. 条形码的主要优点是可以利用扫描仪实现非接触的快速阅读. 另外, 它还具有制作成本低廉, 识别操作简单等优点. 因此, 条形码在运输、仓储以及超级市场等物流管理行业中得到了广泛的应用.

第二章　抽象代数的基本知识

要想掌握编码理论，需要有比较全面的数学基础知识. 许多不同的数学领域都在编码理论中发挥了它们的作用. 代数自然是最重要的，但读者还必须了解初等数论、概率论以及组合论中的一些基本概念和基本结论.

本章介绍抽象代数中的群、环以及域的基本知识.

§2.1　半群

定义 2.1 设 S 是一个非空集合，$*$ 是 S 上的一个二元运算. 所谓 $*$ 是 S 上的一个二元运算，即对任意 $a,b\in S$，都有 $a*b\in S$. 如果运算 $*$ 满足结合律，即对任意 $a,b,c\in S$，都有

$$(a*b)*c=a*(b*c),$$

则称 S 关于运算 $*$ 构成一个半群 (semigroup)，记为 $\langle S,*\rangle$.

例 2.1　设 A 是一个非空集合. 令

$$2^A=\{S\mid S\subseteq A\}.$$

显然，集合的交运算 ($\cap$) 和并运算 ($\cup$) 都是 2^A 上的二元运算，并且都满足结合律. 因此，$\langle 2^A,\cap\rangle$ 和 $\langle 2^A,\cup\rangle$ 都是半群. □

对于一个半群 $\langle S,*\rangle$，其中的运算 $*$ 有时称为乘法，$a*b$ 称为 a 与 b 的积. 在不会引起歧义的情况下，我们有时会将 $a*b$ 简记为 ab.

在一个半群 $\langle S,*\rangle$ 中，对于任意正整数 n 和任意 $a\in S$，我们定义

$$a^n=\underbrace{a*a*\cdots*a}_{n}.$$

显然，对于任意正整数 m 和 n，$a^m*a^n=a^{m+n}$，$(a^m)^n=a^{mn}$.

定义 2.2 设 $\langle S,*\rangle$ 是一个半群. 如果对任意 $a,b\in S$，都有

$$a*b=b*a,$$

则称 $\langle S,*\rangle$ 是一个可交换半群 (commutative semigroup).

对于可交换半群，其中的运算有时称为加法，我们有时会用 $+$ 来表示可交换半群的运算. 设 $\langle S,+\rangle$ 是一个可交换半群. 对任意正整数 n 和任意 $a\in S$，我