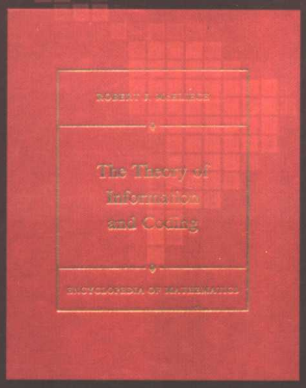
国外电子与通信教材系列

信息论与编码理论

（第二版）

The Theory of Information and Coding

Second Edition



[美] Robert J. McEliece 著
李 斗 殷 悦 罗 燕 等译
项海格 审校

電子工業出版社
Publishing House of Electronics Industry
http://www.phei.com.cn

国外电子与通信教材系列

信息论与编码理论

（第二版）

The Theory of Information and Coding

Second Edition

[美] Robert J. McEliece 著

李 斗 殷 悦 罗 燕 等译

项海格 审校

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内容简介

本书全面系统地介绍了由香农于1948年提出的信息论与编码理论的主要内容，以及近几十年来该领域的一些重要研究成果。作者首先在引言中向读者简单介绍了信息论与编码理论的基本思想；第一部分讲解了香农信息论与编码理论的主要内容，如熵和信息量的基本概念与性质，以及信道编码定理和信源编码定理；第二部分介绍了一些基于香农编码理论的信道和信源编码方法，具体包括线性码、循环玛、BCH和RS码、卷积码等信道纠错编码，以及变长信源编码等。本书内容丰富翔实，对基本概念和基础理论的阐述清晰明了，同时也充分反映了相关领域的研究进展情况。

本书适合作为高等院校信息与通信工程专业研究生或本科生的教材或参考书。书中提供的几十道例题和几百道习题也有助于具有一定概率论和线性代数知识的人自学。

版权贸易合同登记号　图字：01-2003-1042

图书在版编目（CIP）数据

信息论与编码理论：第2版/（美）麦克伊利斯（McEliece, R. J.）著；李斗等译. -北京：电子工业出版社，2004.2

（国外电子与通信教材系列）

书名原文：The Theory of Information and Coding, Second Edition

ISBN 7-5053-9337-5

Ⅰ.信… Ⅱ.①麦… ②李… Ⅲ.①信息论-教材②信源编码-编码理论-教材③信道编码-编码理论-教材 Ⅳ. TN911.2

中国版本图书馆CIP数据核字（2004）004555号

责任编辑：陶淑毅

印　　刷：北京兴华印刷厂

出版发行：电子工业出版社

北京市海淀区万寿路173信箱　邮编：100036

经　　销：各地新华书店

开　　本：787 × 1092　1/16　　印张：19　字数：486千字

印　　次：2004年2月第1次印刷

定　　价：29.00元

凡购买电子工业出版社的图书，如有缺损问题，请向购买书店调换；若书店售缺，请与本社发行部联系。联系电话：（010）68279077。质量投诉请发邮件至zlts@phei.com.cn，盗版侵权举报请发邮件至dbqq@phei.com.cn。

序

2001年7月间，电子工业出版社的领导同志邀请各高校十几位通信领域方面的老师，商量引进国外教材问题。与会同志对出版社提出的计划十分赞同，大家认为，这对我国通信事业、特别是对高等院校通信学科的教学工作会很有好处。

教材建设是高校教学建设的主要内容之一。编写、出版一本好的教材，意味着开设了一门好的课程，甚至可能预示着一个崭新学科的诞生。20世纪40年代MIT林肯实验室出版的一套28本雷达丛书，对近代电子学科、特别是对雷达技术的推动作用，就是一个很好的例子。

我国领导部门对教材建设一直非常重视。20世纪80年代，在原教委教材编审委员会的领导下，汇集了高等院校几百位富有教学经验的专家，编写、出版了一大批教材；很多院校还根据学校的特点和需要，陆续编写了大量的讲义和参考书。这些教材对高校的教学工作发挥了极好的作用。近年来，随着教学改革不断深入和科学技术的飞速进步，有的教材内容已比较陈旧、落后，难以适应教学的要求，特别是在电子学和通信技术发展神速、可以讲是日新月异的今天，如何适应这种情况，更是一个必须认真考虑的问题。解决这个问题，除了依靠高校的老师和专家撰写新的符合要求的教科书外，引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，是会有好处的。

一年多来，电子工业出版社为此做了很多工作。他们成立了一个“国外电子与通信教材系列”项目组，选派了富有经验的业务骨干负责有关工作，收集了230余种通信教材和参考书的详细资料，调来了100余种原版教材样书，依靠由20余位专家组成的出版委员会，从中精选了40多种，内容丰富，覆盖了电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等方面，既可作为通信专业本科生和研究生的教学用书，也可作为有关专业人员的参考材料。此外，这批教材，有的翻译为中文，还有部分教材直接影印出版，以供教师用英语直接授课。希望这些教材的引进和出版对高校通信教学和教材改革能起一定作用。

在这里，我还要感谢参加工作的各位教授、专家、老师与参加翻译、编辑和出版的同志们。各位专家认真负责、严谨细致、不辞辛劳、不怕琐碎和精益求精的态度，充分体现了中国教育工作者和出版工作者的良好美德。

随着我国经济建设的发展和科学技术的不断进步，对高校教学工作会不断提出新的要求和希望。我想，无论如何，要做好引进国外教材的工作，一定要联系我国的实际。教材和学术专著不同，既要注意科学性、学术性，也要重视可读性，要深入浅出，便于读者自学；引进的教材要适应高校教学改革的需要，针对目前一些教材内容较为陈旧的问题，有目的地引进一些先进的和正在发展中的交叉学科的参考书；要与国内出版的教材相配套，安排好出版英文原版教材和翻译教材的比例。我们努力使这套教材能尽量满足上述要求，希望它们能放在学生们的课桌上，发挥一定的作用。

最后，预祝“国外电子与通信教材系列”项目取得成功，为我国电子与通信教学和通信产业的发展培土施肥。也恳切希望读者能对这些书籍的不足之处、特别是翻译中存在的问题，提出意见和建议，以便再版时更正。

吴佑寿

中国工程院院士、清华大学教授

“国外电子与通信教材系列”出版委员会主任

出 版 说 明

进入21世纪以来，我国信息产业在生产和科研方面都大大加快了发展速度，并已成为国民经济发展的支柱产业之一。但是，与世界上其他信息产业发达的国家相比，我国在技术开发、教育培训等方面都还存在着较大的差距。特别是在加入WTO后的今天，我国信息产业面临着国外竞争对手的严峻挑战。

作为我国信息产业的专业科技出版社，我们始终关注着全球电子信息技术的发展方向，始终把引进国外优秀电子与通信信息技术教材和专业书籍放在我们工作的重要位置上。在2000年至2001年间，我社先后从世界著名出版公司引进出版了40余种教材，形成了一套“国外计算机科学教材系列”，在全国高校以及科研部门中受到了欢迎和好评，得到了计算机领域的广大教师与科研工作者的充分肯定。

引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，将有助于我国信息产业培养具有国际竞争能力的技术人才，也将有助于我国国内在电子与通信教学工作中掌握和跟踪国际发展水平。根据国内信息产业的现状、教育部《关于“十五”期间普通高等教育教材建设与改革的意见》的指示精神以及高等院校老师们反映的各种意见，我们决定引进“国外电子与通信教材系列”，并随后开展了大量准备工作。此次引进的国外电子与通信教材均来自国际著名出版商，其中影印教材约占一半。教材内容涉及的学科方向包括电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等，其中既有本科专业课程教材，也有研究生课程教材，以适应不同院系、不同专业、不同层次的师生对教材的需求，广大师生可自由选择和自由组合使用。我们还将与国外出版商一起，陆续推出一些教材的教学支持资料，为授课教师提供帮助。

此外，“国外电子与通信教材系列”的引进和出版工作得到了教育部高等教育司的大力支持和帮助，其中的部分引进教材已通过“教育部高等学校电子信息科学与工程类专业教学指导委员会”的审核，并得到教育部高等教育司的批准，纳入了“教育部高等教育司推荐——国外优秀信息科学与技术系列教学用书”。

为做好该系列教材的翻译工作，我们聘请了清华大学、北京大学、北京邮电大学、东南大学、西安交通大学、天津大学、西安电子科技大学、电子科技大学等著名高校的教授和骨干教师参与教材的翻译和审校工作。许多教授在国内电子与通信专业领域享有较高的声望，具有丰富的教学经验，他们的渊博学识从根本上保证了教材的翻译质量和专业学术方面的严格与准确。我们在此对他们的辛勤工作与贡献表示衷心的感谢。此外，对于编辑的选择，我们达到了专业对口；对于从英文原书中发现的错误，我们通过与作者联络、从网上下载勘误表等方式，逐一进行了修订；同时，我们对审校、排版、印制质量进行了严格把关。

今后，我们将进一步加强同各高校教师的密切关系，努力引进更多的国外优秀教材和教学参考书，为我国电子与通信教材达到世界先进水平而努力。由于我们对国内外电子与通信教育的发展仍存在一些认识上的不足，在选题、翻译、出版等方面的工作中还有许多需要改进的地方，恳请广大师生和读者提出批评及建议。

电子工业出版社

教材出版委员会

译 者 序

随着信息时代的来临，信息论与编码理论在许多领域得到了广泛的应用，并产生了深远的影响。本书全面系统地介绍了由香农在 1948 年提出的信息论与编码理论的主要内容，以及近几十年来的一些重要研究成果。

本书作者 Robert J. McEliece 是美国加州理工学院电子工程系的知名教授，并且是美国国家工程院院士，IEEE、美国工程教育协会和美国数学协会会员。他长期从事信息论与编码理论的研究和教学工作，因成绩卓越而多次获奖，曾因在纠错编码领域的突出贡献而获 NASA 团体成就奖，以及 IEEE 成立百年杰出贡献奖章。

本书是剑桥大学出版社出版的“Encyclopedia of Mathematics and Its Applications”（《数学及其应用百科全书》）系列丛书中的一卷。它内容丰富翔实，对基本概念和基础理论的阐述清晰明了，同时也充分反映了相关领域的研究进展情况，适合作为高等院校信息与通信工程专业研究生或本科生的教材或参考书。书中提供的几十道例题和几百道习题也有助于具有一定概率论和线性代数知识的人自学。多年以来，北京大学信息科学技术学院一直将本书原版作为研究生课程“信息与编码理论”的主要参考用书。这次非常感谢电子工业出版社给予了我们翻译本书的机会。

本书的引言和第一部分的第 1 章至第 6 章的正文由李斗翻译，引言和第一部分的习题及注释由殷悦翻译；第二部分第 7 章至第 9 章的正文由殷悦翻译，第 10 章至第 12 章的正文由罗燕翻译，第二部分的习题及注释也由罗燕翻译；附录部分由方东翻译。全书译文由李斗统校，项海格教授审校。赵玉萍、刘志敏和陈江等教授对本书的翻译工作也做出了贡献。在翻译过程中，我们对原书中的一些错误做了更正。由于译者水平有限，译文中的错误和不妥之处，希望读者批评指正。

第一版序

本书旨在自成体系地介绍信息论与编码理论中的基本结论。它写于 1972 年至 1976 年，当时我在美国加州理工学院讲授这门课程。我的学生中有一半是电子工程系的研究生；其他的则来自各个领域(数学、物理、生物，甚至有一个是英语专业的)。因此这门课程面向的既是专业人员也是非专业人员，本书也相应如此。

全书由三部分组成：引言、第一部分(信息论)和第二部分(编码理论)。读者首先应该阅读引言部分，因为它给出了本书所涉及内容的一个概述。在第一部分中，第 1 章是基础，但是并不需要首先阅读这一章，因为它实际上只是关于熵和互信息量的一个总结，最好将它作为参考，以便在学习第 2 章至第 5 章时查阅。第 6 章是对前沿成果的一个综述，可以独立阅读。在第二部分中，第 7 章是基础，应该在阅读第 8 章和第 9 章之前阅读。但是第 10 章的大部分，以及第 11 章则是完全独立于第 7 章的。第 12 章是独立于其他所有章节的另一个综述。

每一章后面的习题非常重要，其中介绍了正文中忽略的很多细节，以及没有提到的许多重要结论。建议读者至少要阅读一下这些习题。

本书包含四个附录。附录 A 概括介绍了第一部分所需的概率论知识。附录 B 讨论的是凸函数和 Jensen 不等式。在第一部分中经常引用 Jensen 不等式，不熟悉该不等式的读者应该首先阅读附录 B。附录 C 简要介绍了第 9 章所用到的有限域的主要结论。附录 D 则讲解了一种在方向图中计算路径的算法，第 10 章中便用到了这种算法。

这里有必要说明一下标号排序：节、插图、例题、定理、公式，以及习题都是依据各章的编号，采用两部分数字标记。因此“2.3 节”、“定理 3.4”与“习题 4.17”分别表示第 2 章的第 3 节、第 3 章的第 4 个定理、第 4 章的第 17 道习题。附录按字母索引，所以“式(B.4)”表示附录 B 中的第 4 个公式。

下面是一些需要解释的特殊符号：$\lfloor x \rfloor$表示小于或等于 x 的最大整数；而$\lceil x \rceil$表示大于或等于 x 的最小整数。

最后，我很高兴在这里表达我的谢意。感谢 Gus Solomon，我这门学科的启蒙老师；感谢 John Pierce，给了我在加州理工学院教学的机会；感谢 Gian-Carlo Rota，鼓励我写成此书；感谢 Len Baumert，Stan Butman，Gene Rodemich 和 Howard Rumsey，我在书中引用了他们的研究成果；感谢 Jim Lesh 和 Jerry Heller，他们提供了图 6.7 和图 12.2 的数据；感谢 Bob Hall 绘制插图；感谢我的打字员 Ruth Stratton，Lillian Johnson，特别是 Dian Rapchak；并感谢 Ruth Flohn 进行了审稿。

Robert J. McEliece

第二版序

本版的主要修订是在第二部分。这一版对上一版的第 8 章进行了修改，扩展为新的两章，即第 8 章和第 9 章。原来的第 9 章至第 11 章相应地变为第 10 章至第 12 章。新版的第 8 章全面地介绍了循环码的数学原理，以及它们的移位寄存器电路的实现，这一章的最后讨论了如何利用循环码纠正突发错误。而新版的第 9 章与原来的第 8 章非常相似，只是对 Reed-Solomon 码的介绍更为详尽，体现了它们在实际应用中的重要性。新的两章都附加了几十道习题。

前　言

我们所谓的通信，其核心问题就是信息的传输。作为一个备受关注的领域，它所涉及的范围之广已引起了哲学家们的关注，并产生了一门蓬勃发展的技术学科。

我们将这一切归功于香农[①]，是他首先认识到可以采用一种系统的、有规可循的方法，解决信息的编码、传输和译码等一系列问题：他于 1948 年发表的经典论文，标志着数学领域一个新篇章的诞生。

过去的 30 年中，在这一新生领域涌现出了数量惊人的著作，其中的一些术语甚至已经成为我们日常用语的一部分。

这本专著（实际上是两本专著合二为一）系统全面地介绍了通信领域的两个方面：编码与传输。

第一个方面（本书第二部分的主要内容）是代数理论的力与美的完美例证；第二个方面则属于概率论领域，由香农以新奇而独创的方式揭开序幕。

Mark Kac

① C. E. Shannon, A Mathematical Theory of Communication, *Bell System Tech. J.* **27**(1948)(Introduction: 379 ~ 382; Part one: Discrete Noiseless Systems, 382 ~ 405; Part two: The Discrete Channel with Noise(and Appendixes), 406 ~ 423; Part III: Mathematical Preliminaries, 623 ~ 636; Part IV: The Continuous Channel(and Appendixes), 637 ~ 656).

目　录

第一部分　信　息　论

第二部分　编 码 理 论

引　言

1948 年，Claude Shannon[1]（克劳德·香农）在他的经典论文“A Mathematical Theory of Communication”（通信中的数学理论）的引言部分中写道：

“通信中的基本问题就是在某一点精确或近似地再生另一点选择的信息。”

为解决这一问题，他在该论文中提出了应用数学的一个全新分支，现在称之为信息理论和/或编码理论。本书的目的就是介绍这一理论提出后 30 年来的主要成果。

在引言这一章，将通过一对特殊的数学模型，二进制对称信源和二进制对称信道，来介绍信息理论的核心思想。

二进制对称信源（简称信源）是一个可以发出定义为“0”和“1”的两种特定符号的实体，速率为单位时间内 R 个符号。我们称这些符号为比特（bits）[bits 是 binary digits（二进制数据）的缩写]。信源随机地发出这些比特，“0”和“1”的发出概率相同。假设信源速率 R 是连续变化的，即 R 可以是任何非负的数值。

二进制对称信道（简称 BSC[2]）是一个在单位时间内可以传输 1 比特数据的实体。但是该信道并不是完全可靠的：存在一个固定的概率 p（称为原始误比特率[3]），满足 $0 \leqslant p \leqslant \frac{1}{2}$，使输出比特与输入比特不相同。

现在设想有两个人，一个是发送者，另一个是接收者。发送者必须使接收者尽可能准确地接收信源的输出，而他们之间惟一的通信链路就是前面描述过的 BSC（但是在启动信源之前，允许发送者和接收者在一起了解彼此将采用的数据处理策略）。这里还假设发送者和接收者都可以无任何限制地利用计算能力、存储容量、政府经费和其他资源。

现在要问，对于给定的一个信源速率 R，发送者和接收者之间通过 BSC 的通信可以达到多高的可靠度？我们最终将给出此问题的一个非常精确的通用结论，但现在先考虑一些特例。

假设 $R = 1/3$，这意味着信道传输数据的速率是信源产生数据速率的三倍，因此在传输之前，可以对信源输出的每一比特重复编码三次。例如，如果信源输出的前五个比特是 10100，编码流将是 111000111000000。对应每个信源比特，接收者将收到三个比特，但是由于存在信道“噪声”，这三个比特可能不完全相同。如果信道干扰了传输的第 2、第 5、第 6、第 12 和第 13 比特，接收者将收到 101011111001100。稍微思考一下你就会认为，此时接收者对信源比特进行译码的最佳策略是对接收的三个比特多票判决。在我们的例子中，这样译出的接收信息是 11100，第二比特出现了一个错误。一般情况下，如果一个信源比特的三个重复编码比特中有两个或三个被信道干扰了，接收就会出错。因此，如果用 P_e 表示误比特率，

$$\begin{aligned} P_e &= P\{2\text{ 个信道错误}\} + P\{3\text{ 个信道错误}\} \\ &= 3p^2(1-p) \quad + p^3 \\ &= 3p^2 - 2p^3 \end{aligned} \tag{0.1}$$

由于 $p \leqslant \frac{1}{2}$,这个值一般比原始误比特率 p 小;这种简单的编码策略提高了信道的可靠度,并且 p 值越小,提高就越显著。

现在很容易看出,通过对每个比特重复编码更多次可以得到更高的可靠度。因此如果 $R=1/(2n+1)$,其中 n 为正整数,则在传输之前就可以对每个比特重复编码 $2n+1$ 次(见习题 0.2),并且采用前面提到的多票判决准则译码。现在很容易得出最终误比特率 $P_e^{(2n+1)}$ 的公式:

$$\begin{aligned} P_e^{(2n+1)} &= \sum_{k=n+1}^{2n+1} P\{2n+1\text{个传输比特中有 } k\text{ 个信道错误}\} \\ &= \sum_{k=n+1}^{2n+1} \binom{2n+1}{k} p^k (1-p)^{2n+1-k} \\ &= \binom{2n+1}{n+1} p^{n+1} + p \text{ 的高次项} \end{aligned} \tag{0.2}$$

如果 $n>1$,该值随着 $p \to 0$ 而趋近于 0 的速度要比前面考虑的 $n=1$ 的特殊情况快[4]。因此就很容易理解,为什么长重复序列比短重复序列更有效。这里需要进一步强调的是,对于原始误比特率 $p<\frac{1}{2}$ 的固定 BSC,当 $n \to \infty$ 时,$P_e^{(2n+1)} \to 0$,即通过这些重复编码方式,可以使信道达到理想的可靠度。研究 $P_e^{(2n+1)}$ 的公式(0.2)最终可以得出这样的结论。但也可以换一种方式,采用*弱大数定理*① 来进行分析,如果在信道中传输 N 个比特,则对于任意 $\varepsilon>0$,有:

$$\lim_{N\to\infty} P\left\{\left|\frac{\text{信道错误的数目}}{N} - p\right| > \varepsilon\right\} = 0 \tag{0.3}$$

也就是说,只要 N 的取值足够大,接收比特中出错的比例就不可能与 p 相差很多。因此可以对 $P_e^{(2n+1)}$ 做如下估计:

$$\begin{aligned} P_e^{(2n+1)} &= P\left\{\text{接收的传输比特中出错的比例} \right. \\ &\quad \left. \geqslant \frac{n+1}{2n+1} = \frac{1}{2} + \frac{1}{4n+2}\right\} \\ &\leqslant P\left\{\text{比例} > \tfrac{1}{2}\right\} \\ &\leqslant P\left\{|\text{比例} - p| > \tfrac{1}{2} - p\right\} \end{aligned}$$

根据式(0.3),当 $n \to \infty$ 时,$P_e^{(2n+1)}$ 确实趋近于 0。我们由此得出结论:如果 R 的值很小,即使信道本身的噪声很大,也可以使最终的错误概率很小。这个结论当然并不奇怪。

以上讨论的都是速率小于 1 的情况。当速率大于 1 时情况会怎么样呢? 在这种条件下,通信的可靠度又如何?

如果 $R>1$,不妨只传输信源比特的 $1/R$ 部分,并让接收者以抛硬币的方式猜测其余的部分。很容易计算出这种简单方式的最终误比特率为:

---

① 在附录 A 中讨论。

$$\begin{aligned} P_e &= \frac{1}{R}\times p + \frac{R-1}{R}\times\frac{1}{2} \\ &= \tfrac{1}{2} - (\tfrac{1}{2} - p)/R \end{aligned} \tag{0.4}$$

现在以 $R=3$ 为例，来介绍另外一种当 $R>1$ 时可以采用的略有创意的方式。如果 $R=3$，信道只能传输信源所产生比特的三分之一。因此发送者将信源输出分为三比特一组，并且只传输这三比特中占多数的比特。例如，如果信源输出 101110101000101，发送者将在信道中传输 11101。接收者只需将收到的每个比特重复三次。此时如果信道干扰了传输的第二个比特，接收者将收到 10101，并扩展为 111000111000111，于是就产生了五比特的错误。一般情况下，最终的误比特率为：

$$\begin{aligned} P_e &= \tfrac{1}{4}\times(1-p) + \tfrac{3}{4}\times p \\ &= \tfrac{1}{4} + p/2 \end{aligned} \tag{0.5}$$

注意这个结果比我们前面取 $R=3$ 时，通过“抛硬币”方式得出的 $\frac{1}{3}+p/3$ 要小。当 R 取其他的整数值时，采用该方式的通用结论作为习题(见习题 0.4)留给读者推导。

到目前为止我们所考虑的方式都是微不足道的，当然也并非毫无意义。现在介绍一个更重要的例子，事实上在 1948 年以前人们对这种方式还毫无了解。

这里假设 $R=4/7$，即对于信源产生的每四个比特，信道都有时间再传输三个附加比特。我们仔细地选择这些附加比特：如果四个信源比特表示为 x_0，x_1，x_2，x_3，则附加或者冗余(或者奇偶校验)比特表示为 x_4，x_5，x_6，它们由下列等式确定：

$$\begin{aligned} x_4 &\equiv x_1 + x_2 + x_3 \pmod 2 \\ x_5 &\equiv x_0 + x_2 + x_3 \pmod 2 \\ x_6 &\equiv x_0 + x_1 + x_3 \pmod 2 \end{aligned} \tag{0.6}$$

举例说明，如果 $(x_0, x_1, x_2, x_3) = (0110)$，则 $(x_4, x_5, x_6) = (011)$，而信道将传输的全部七比特码字为 0110011。

为了说明接收者如何通过被干扰了的七比特码字估计四个信源比特，即描述他的译码算法，我们按照下面的方式重写奇偶校验式(0.6)：

$$\begin{aligned} x_1 + x_2 + x_3 + x_4 \qquad\quad &= 0 \\ x_0 \quad + x_2 + x_3 \quad + x_5 \quad &= 0 \\ x_0 + x_1 \quad + x_3 \quad\quad + x_6 &= 0 \end{aligned} \tag{0.7}$$

[式(0.7)中采用的也是模 2 运算。]换一种略微不同的描述方式，如果定义二进制矩阵 H 为：

$$H = \begin{bmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix}$$

则我们看到 16 个可能的码字 $\mathbf{x} = (x_0, x_1, x_2, x_3, x_4, x_5, x_6)$ 都满足矩阵-矢量方程：

$$H\mathbf{x}^T = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \tag{0.8}$$

[式(0.8)中上角标 T 表示“转置”。]

可以设想 BSC 对传输的每一比特加(模 2)0 或者 1,加 0 表示这一比特接收时没有错误,加 1 则表示有错误。因此如果传输的是 $\mathbf{x}=(x_0,\ x_1,\cdots,\ x_6)$,则接收矢量是 $\mathbf{y}=(x_0+z_0,\ x_1+z_1,\cdots,\ x_6+z_6)$,其中如果信道在传输的第 i 比特产生了一个错误,则 $z_i=1$,否则 $z_i=0$。因此,若将 $\mathbf{z}=(z_0,\ z_1,\cdots,\ z_6)$定义为错误图案,则 $\mathbf{y}=\mathbf{x}+\mathbf{z}$。

接收者接收到 $\mathbf{y}$ 时希望确定 $\mathbf{x}$,为此他将计算下面的矢量 $\mathbf{s}=(s_0,\ s_1,\ s_2)$:

$$\begin{aligned}\mathbf{s}^T &= H\mathbf{y}^T\\ &= H(\mathbf{x}+\mathbf{z})^T\\ &= H\mathbf{x}^T+H\mathbf{z}^T\\ &= H\mathbf{z}^T \quad [\text{见式}(0.8)]\end{aligned} \tag{0.9}$$

这里 $\mathbf{s}$ 称为 $\mathbf{y}$ 的伴随式[5](syndrome)。伴随式中的一个 0 分量表示 $\mathbf{y}$ 满足相应的奇偶校验方程,而 1 分量则表示不满足。根据式(0.9),伴随式不依赖于发送的码字,而只与错误图案 $\mathbf{z}$ 有关。但是由于 $\mathbf{x}=\mathbf{y}+\mathbf{z}$,如果接收者找到了 $\mathbf{z}$,也就知道了 $\mathbf{x}$,因此译码问题的重点是寻找 $\mathbf{z}$。等式 $\mathbf{s}^T=H\mathbf{z}^T$ 表明,$\mathbf{s}^T$ 是与 $\mathbf{z}$ 中 1 分量相对应的 H 矩阵中列矢量的(二进制)和,而 $\mathbf{z}$ 中 1 分量对应的比特是被信道干扰的:

$$\mathbf{s}^T=z_0\begin{bmatrix}0\\1\\1\end{bmatrix}+z_1\begin{bmatrix}1\\0\\1\end{bmatrix}+\cdots+z_6\begin{bmatrix}0\\0\\1\end{bmatrix} \tag{0.10}$$

接收者的任务是,一旦计算出 $\mathbf{s}$,就通过方程式 $\mathbf{s}^T=H\mathbf{z}^T$"求解"$\mathbf{z}$。遗憾的是,这里只有三个方程却有七个未知量,因此对应任一 $\mathbf{s}$,总有 16 个可能的 $\mathbf{z}$。这显然已经是一个进步了,因为 $\mathbf{z}$ 本来有 128 种可能的取值,但是接收者如何在剩下的 16 个值中进行选择呢? 例如,假设接收到 $\mathbf{y}=(0111001)$,则 $\mathbf{s}=(101)$,$\mathbf{z}$ 的 16 个候选值为:

0	1	0	0	0	0	0		0	0	1	0	0	1	1
1	1	0	0	0	1	1		0	0	0	1	0	1	0
0	0	0	0	1	0	1		0	1	1	1	0	0	1
0	1	1	0	1	1	0		1	0	1	0	0	0	0
0	1	0	1	1	1	1		1	0	0	1	0	0	1
1	0	0	0	1	1	0		1	1	1	1	0	1	0
1	1	1	0	1	0	1		0	0	1	1	1	0	0
1	1	0	1	1	0	0		1	0	1	1	1	1	1

面对这样一组可能的错误图案,显而易见的是:由于原始的误比特率 $p<\frac{1}{2}$,错误图案包含的 1 的个数(错误)越少,就越有可能是实际的错误图案。在这个例子中,幸运的是,重量最小的错误图案(0100000)只有一个,这里的重量代表错误图案中 1 的个数。此时接收者对 $\mathbf{z}$ 的最佳估计(同时依据伴随式和信道统计特性)是 $\mathbf{z}=(0100000)$;对传输码字的估计是 $\mathbf{x}=\mathbf{y}+\mathbf{z}=(0011001)$;而最终对四个信源比特的估计是(0011)。

当然在上面的例子中也并不是真正的幸运,因为可以看到对于任意伴随式 $\mathbf{s}$,方程 $H\mathbf{z}^T=\mathbf{s}^T$ 总是存在重量为 0 或 1 的惟一解。比如,如果 $\mathbf{s}=(000)$,则期望的结果是 $\mathbf{z}=(0000000)$。但是如果 $\mathbf{s}\neq(000)$,则 $\mathbf{s}^T$ 必然作为 H 的某一列出现;如果 $\mathbf{s}^T=H$ 的第 i 列,则第 i 位为 1、其余位都为 0 的错误图案 $\mathbf{z}$ 是 $H\mathbf{z}^T=\mathbf{s}^T$ 的惟一最小重量解。