Jiu cuo bian ma
Ji sh he ying yong

归绍升 编著

纠错编码技术和应用

上海交通大学出版社出版

纠错编码技术和应用

归绍升　编著

上海交通大学出版社

内 容 简 介

本书分两大部分。前面六章为第一部分，主要介绍：纠错编码技术和近代代数的基本概念、代数码(包括分组码和卷积码)、几何码(包括欧氏几何码和射影几何码)、算术码，属于纠错编码技术的内容。后面四章为第二部分，主要介绍：几种自动请求重发方案、分组码和卷积码在计算机和通信系统中的应用、编码在扩频通信系统中的应用，属于纠错码应用的内容。

本书理论与实际并重，可用作通信、计算机等专业的研究生或高年级大学生的教材，也可供这些专业的工程技术人员参考。

纠错编码技术和应用

上海交通大学出版社出版

(淮海中路1984弄19号)

新华书店上海发行所发行

常熟市梅李印刷厂印装

开本787×1092毫米 1/16 印张11.75 字数281.000

1988年3月第1版 1988年4月第1次印刷

印数：1—3000

ISBN 7-313-00110-x/TN914 科技书目：168—299

定价：1.95元

前　言

自香农(C.E.Shannon)于1948年和1949年发表了两篇经典论文之后，戈莱(M.J.E.Golay)和汉明(R.W.Hamming)分别于1949年及1950年发表了有关纠错码的论文，纠错码才被提出，并于60年代开始发展。

纠错编码之所以日益受到人们的重视并得到迅速的发展，主要原因是由于数字通信的发展和电子计算机的广泛应用。我们知道，无论计算机还是数字通信，都是籍数字化和编码等技术来进行工作的。如用了纠错码，就可降低数据传输的错误概率，即增加传输系统的可靠性。对通信来说，常常用信道编码来增大所传信号在信道中的抗干扰能力，因此信道编码又可称为纠错编码。对计算机来说，要求籍编码技术来控制计算错误和数据传输错误。如采用算术码：还易于籍通用计算机来编码和译码。总之，纠错编码得到迅速发展的主要原因有三：(1) 数字通信(包括计算机)的发展促进了对纠错编码的研究；(2) 近代代数的理论为代数编码提供了理论基础；(3) 大规模集成电路为纠错码的发展提供了物质基础。因此，目下无论通信或计算机方面的工程技术人员，或者数学研究人员，都对编码技术很感兴趣，致使编码理论的研究蓬勃发展，异常活跃。

作者在讲授为研究生开设的“编码理论”课时，对这方面的丰富资料很为满意，但对偏重理论还是结合实际，广泛介绍还是系统概括，就感到有些无所适从。经过教学实践，初步认为采用数学分类法，即将内容分为(1) 代数码，(2) 几何码，(3) 算术码三类来概括，较为妥当。为了加强实践方面的内容，特在本书最后四章讲一些主要的应用。

对于上述三类码的涵义，先在此作一简要说明：所谓代数码，以线性分组码为例，是指“能利用校验矩阵H来校验接收矢量有无错误，和生成矩阵G来构造码”的一些码。所谓几何码，目下是指：用射影几何(Projective Geometry)及欧氏几何(Euclidean Geometry)来构成的一些大数逻辑可译码。所谓算术码，是指：所有编码和译码运算均属普通算术，能控制计算错误和数传错误，易于在通用计算机中完成编码和译码的一些码。当然，代数码的理论已较成熟，但几何码和算术码的研究还应加速进行。

本书共分十章。第一章绪论，主要介绍有关纠错编码技术的一些基本概念。第二章近代代数的基本概念，只介绍阅读本书时用得到的一些内容。第三章及第四章介绍代数码，主要为分组码和卷积码。第五章介绍几何码，第六章介绍算术码。第七章介绍有关自动请求重发的几种传输错误控制法。第八章介绍分组码在计算机数据存贮系统中的应用。第九章主要介绍卷积码在卫星通信系统及自动请求重发系统中的应用。第十章介绍在扩频通信系统中用编码来抗人为干扰。为了帮助读者掌握本书内容，各章都附有习题。因此，本书既可选作研究生或高年级大学生用的教材，又可供通信、计算机等专业人员参考。

由于作者水平有限，书中缺点和错误在所难免，敬请读者批评指正。

作者

1986年11月

纠错编码技术和应用

目　录

第一章　绪论

1.1 纠错编码的历史与发展概况…… 1
1.2 通信系统简介…… 2
1.3 错误的种类和有关名词的解释…… 3
1.4 纠错码分类…… 4
1.5 基本概念介绍…… 5
1.6 编码问题与研究方向…… 12
习题…… 13
第二章　近代代数的基本概念…… 15
2.1 群…… 15
2.2 环…… 17
2.3 域…… 19
2.4 矢量空间…… 19
2.5 矩阵…… 20
习题…… 23
第三章　代数码Ⅰ——分组码…… 24
3.1 汉明完备码…… 24
3.2 汉明循环码…… 29
3.3 BCH码…… 36
3.4 里德-索洛蒙码(简称 RS 码)…… 47
3.5 戈莱码…… 52
3.6 二次剩余码…… 58
习题…… 59
第四章　代数码Ⅱ——卷积码…… 61
4.1 编码方法…… 61
4.2 译码方法…… 68
习题…… 81
第五章　几何码…… 82
5.1 欧氏几何的基本概念…… 82
5.2 欧氏几何循环码…… 83
5.3 里德-马勒码 …… 87
5.4 射影几何码…… 92
5.5 差集循环码…… 96

5.6 极长码…… 99
习题…… 103
第六章 算术码…… 104
6.1 算术重量和算术距离…… 104
6.2 AN码 …… 105
6.3 BN码 …… 107
6.4 AN+B码…… 108
6.5 代数循环码和算术循环码的比较…… 109
习题…… 110
第七章 自动请求重发的错误控制法…… 111
7.1 三种基本的ARQ方案和两种性能的量度…… 111
7.2 SR 自动请求重发系统 …… 115
7.3 (SR+GBN)混合型ARQ方案 …… 118
7.4 (FEC+ARQ)两法混合型方案…… 119
7.5 可逆码、逆向法电路和Ⅱ型两法混合方案举例…… 121
习题…… 123
第八章 数据存贮系统中错误控制用的分组码…… 124
8.1 计算机主存贮器和控制存贮器用的(SEC-DED)码 …… 124
8.2 磁带用的最优矩形码…… 127
8.3 在IBM3850大容量存贮系统中用的(15,13)BCH码 …… 135
8.4 磁盘用的法尔码和里德-索洛蒙码 …… 139
习题…… 144
第九章 卷积码的具体应用…… 145
9.1 卷积码在空间和卫星通信系统中的应用…… 145
9.2 卷积码在受衰减和干扰的信道中纠正突发错误的应用…… 153
9.3 卷积码在自动请求重发系统中的应用…… 155
习题…… 156
第十章 编码在扩频通信系统中的应用…… 157
10.1 伪随机码 …… 157
10.2 直扩技术 …… 161
10.3 跳频技术 …… 162
10.4 扩频系统用的编码 …… 165
习题…… 171
附录Ⅰ…… 172
附录Ⅱ…… 174
附录Ⅲ…… 175
附录Ⅳ…… 176
附录Ⅴ…… 178
参考文献…… 179

第一章 绪 论

纠错编码又称信道编码，是提高数字传输可靠性的一种技术。本章主要介绍与纠错编码有关的基本问题。

1.1 纠错编码的历史与发展概况

香农(C.E.Shannon)在1948年发表的《通信的数学理论》和1957年发表的《适用于有扰信道的编码理论某些成果》中提出了关于有扰信道中传输消息的重要理论——香农第二定理。该定理指出：设信道具有确定的容量C bit/s，并已知传输消息的速率为R bits/s，那么只要$R<C$，就存在速率为R的纠错码。若用最大似然译码，其错误译码概率Pe就可以任意小。

但该定理并未明确指出如何将拟传输的消息进行纠错编码，也未提出这种具有纠错能力的传输系统的具体实现方法。

由于这种纠错码能提高通信的可靠性，故日益受到科技人员的重视，自60年代以来这方面的研究活动十分活跃。近年来，由于数字通信的发展，促进了纠错编码的研究；由于近代代数理论的应用，为代数编码提供了理论基础；由于大规模集成电路的进展，为纠错编码提供了物质基础；由于编、译码器可用计算机模拟来进行研究和开发，也促进了这方面研究的发展。自戈莱(M.J.E.Golay)于1949年发表“评论数字编码”和汉明(R.W.Hamming)于1950年发表“检错和纠错码”两篇文章以来，纠错码的发展大致可分下列三个阶段①：

第一阶段(1949—60年代初)重要成果有：

(1) 奠定了线性分组码的理论基础。

(2) 提出了纠正多个随机错误的BCH码。

(3) 提出了卷积码的序列译码。

(4) 彼得森(W.W.Peterson)所著《纠错码》一书于1961年出版(第一版)。

第二阶段(60年代初——60年代末)重要成果有：

(1) 代数编码理论日趋完善。伯利坎普(E.R.Berlekamp)所著《代数编码理论》一书于1968年出版。

(2) 提出了门限译码(即二进制情况的大数逻辑译码)。

(3) 提出了BCH码用的迭代译码算法。

(4) 提出了卷积码用的序列译码和维特比(Viterbi)译码算法。

第三阶段(70年代初——现在)重要成果有：

(1) 迅速发展实用的编、译码技术，例如：快速译码、分组码用软判决译码、多址信道编码及信道模化，编、译码器的计算机模拟等。

(2) 发现了一类渐近性能很好的分组码：戈帕(Goppa)码和贾斯特西(Justesen)码。

(3) 出版的著作较多，例如：

注① 下面讲到一些名词的具体意义，在以后各章中会详细说明。

(i) 美国贝尔实验室的麦克威廉斯(F.J.Macwilliams)和斯隆(N.J.A.Sloane)所著《纠错码理论》于1977年出版。

(ii) 中国科学院的万哲先所著《代数与编码》于1976年出版。

1.2 通信系统简介

图1.1是描述数字通信过程的方框图，也就是一个通信系统。

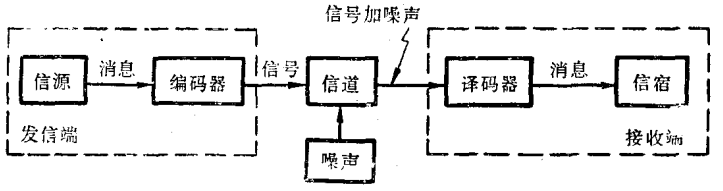

图 1.1 通信系统

图上几个名词的含义和各部分的功能简述如下：

消息 是载荷信息的有序符号序列(包括状态、字母、数字等)，或是载荷信息的连续时间函数。前者如书信、电报、数据等，称为离散消息；后者如语声、图片、活动图像等，称为连续消息。

信号 是由消息经过编码和调制转换过来的、适于信道上传输的电量或光量(适于光通信系统中光缆上传输)；光量是由信号电流驱动光源发光、即对光源进行光强调制而得的。如载荷信息的信号物理量(如幅度等)的改变，在时间上是离散的，就称为离散信号；如在时间和取值两方面都离散，就称为数字信号；如信号为时间的连续函数，就称为连续信号或模拟信号。

信息 是通信系统传输和处理的对象，被载荷于消息、信号之中。另一种说法是：消息和信号中所蕴含的意义就称信息。

接着介绍通信系统各组成部分的功能。

信源产生消息。信源可以是人或计算机，输出可以是连续波形或是离散的符号(或字母)序列。

编码器包括信源编码器、信道编码器和调制器。信源编码器主要用来将消息编码，并根据香农第一定理中关于信源编码后码字平均长度L与信源每个符号的平均信息量$H(s)$间内在规律：$L \geqslant H(S)$，将码长减短(即去除多余码元)。故信源编码的目的是提高传输的有效性，即用来克服高的数码率(每秒比特数)需要宽频带，因而对传输、存贮、处理带来的困难。信道编码器主要用来进行提高数字传输抗干扰能力的纠错编码。即在香农第二定理中信息率(R)<信道容量(C)的条件下将码长增长，目的在于提高传输的可靠性。调制器用脉冲序列(即编码器输出)来调制载波。例如，“1”可示成时宽(即持续时间)为T的正脉冲，“0”可示成时宽为T的负脉冲(或空号)。

如为光通信系统，则在紧靠信道的左边应加入一台把电转换成光的光发端机，在紧靠信道的右边也应加入一台把光转换成电的光收端机。该两台光端机的详细情况从略。

信道是传输信号的媒介或途径。除各种通信信道(例如，电缆、波导、空间等)之外，还包括例如具有磁头(有写入和读出功能)的磁带或磁盘这一种存贮媒介，这时的图1.1应改称为信息存储系统。

噪声(包括外来干扰)本散布在系统各点，为了方便，才把它集中表示。因此，我们现在可讲："信号进入信道，受到噪声的干扰"。图上信道输出是"信号加噪声"，就是说这个接收信号已发生了错误。

译码器包括解调器和译码器。解调器对每个时宽为T的接收信号进行判决，以确定发送的是"1"，还是"0"(即恢复脉冲序列，并将传输数字估值)。解调器输出称为接收序列，由于信道噪声的干扰，可能与信道编码器输出的码字不一致，它们间不同的码元就表示发生了错误。译码器中的信道译码器根据信道编码规则和信道统计特性，完成以下任务：1)纠正传输错误，产生发送码字的估值；2)将码字估值变换成信源编码器输出序列的估值。信源译码器根据信源编码规则，将信源编码器输出序列的估值变换成信源输出(即消息)的估值，并送至信宿(即用户)。

1.3 错误的种类和有关名词的解释

在通信系统的接收端，如接收矢量$\boldsymbol{r}$和发送的原码字$\boldsymbol{v}$不一样，例如$\boldsymbol{v}=(11000)$，而$\boldsymbol{r}=(10001)$，则$\boldsymbol{r}$中有两处出现错误，即有2个错误。这种错误是由信道中噪声的干扰所引起的。

编码书刊中常提到信道错误图样，在此介绍一下：

发送的码字：	$\boldsymbol{v}=(v_1,v_2,v_3,\cdots,v_n)$	例如：$\boldsymbol{v}=(11000)$
接收矢量：	$\boldsymbol{r}=(r_1,r_2,r_3,\cdots,r_n)$	$\boldsymbol{r}=(10001)$
信道错误图样：	$\boldsymbol{e}=\boldsymbol{v}+\boldsymbol{r}=(v_1+r_1,v_2+r_2,\cdots,v_n+r_n)$ $=(e_1,e_2,e_3,\cdots,e_n)$	$\boldsymbol{e}=(1+1,1+0,0+0,0+0,0+1)$ $=(01001)$

从$\boldsymbol{e}=(01001)$可以看出：如$e_i=v_i+r_i=1$，则表示码字中第i位受到干扰，因而出现错误。上面具体例子是：从左端起的第2位和第5位是错误的。

经常遇到的错误有两种：

(1) 随机错误——由随机噪声所引起。由于随机噪声的特性，该错误的特点是：各码元是否发生错误是互相独立的，因而通常不会成片地出现错误。

(2) 突发错误——由突发噪声所引起。由于突发噪声的特性，使"各个码元是否错误"存在相关性，因此，该错误是成片出现的。在一个突发错误持续长度内，开头和最末的码元总是错的，中间一些码元有的错，有的不错，但错的码元相对地比较多。具体例子如：闪电或开关瞬态属于突发噪声，会引起电话线受扰；磁带缺陷也是突发噪声。这些噪声都会使错误成群出现，形成突发错误。

我们知道：为了提高接收端收到消息的可靠性，应采用纠错编码技术。但如要纠正错误，首先应能检查出错误，然后才能进行纠正。因此，对于错误，应有检错和纠错这两个概念。现用检错码和纠错码的定义来加以说明：

(1) 检错码——有发现错误能力的码。通常用检错码的通信系统必须具有反馈信道。当发现收到的消息有错误时，接收端通过反馈信道发出一个信号，要求发信端把该消息再次发送

（即重发），直到接收端认为准确为止。由于错误是随机的，重发可能会收到准确的消息。

（2）纠错码——有发现并纠正错误能力的码，是一种重要的抗干扰码。本书以后各章将专门论述这种码。

1.4 纠错码分类

现在先列出纠错码分类框图，然后再逐一加以解释。

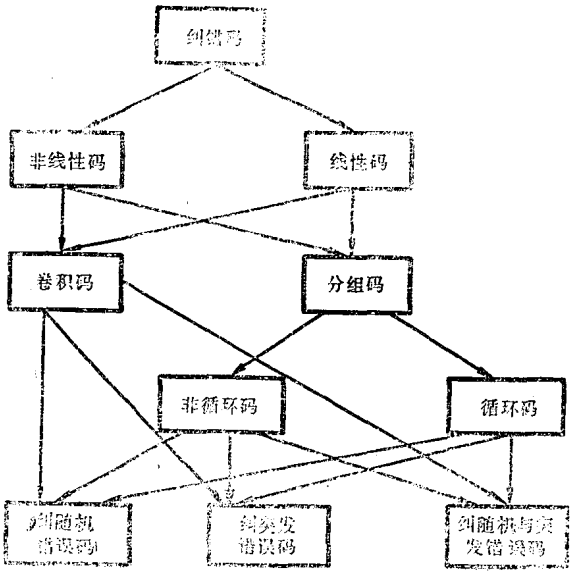

图1.2 纠错码分类

1．按校验位与信息位间的关系来分类，纠错码有下列两类：

（1）线性码——校验位与信息位之间呈线性关系，即可把校验规则列成线性方程组。

（2）非线性码——校验位与信息位之间不存在线性关系。

关于校验位和信息位的意义，简述如下：在一个二进制码的码字中，含有信息的若干位二进制数称为信息位；供校验接收矢量的各位有无错误之用的若干二进制数多余位称为校验位。

2．按对信息位的处理方法来分类，纠错码分为：

（1）分组码——如码组长度为n位，信息位数为k，则每一码组的$(n-k)$个校验位仅与本码组的k个信息位有关，而与别组的信息位无关。

（2）卷积码——如码段长度为n_0位，该段有k_0个信息位，则该码一个码段的(n_0-k_0)个校验位不仅与本段的k_0个信息位有关，且也与前m段的信息位有关。该码的约束长度为$(m+1)n_0$位。

3．按码字之间的关系来分类，纠错码可分为：

（1）循环码——该码特点是：如将码字循环右移一位，就得这一组码字中的另一码字；即

同一组的所有码字均可由循环移位获得。例如第三章中表3-6所示的(7,4)循环码，在$2^4=16$个码字中除去了一个全零码字和一个全一码字之外，还有二组码字，每组各有7个码字。

(2) 非循环码——不能从循环移位法获得全部码字。

4. 按被纠错误的类型来分类，纠错码可分：

(1) 纠随机错误码——设码字间的最小距离为d_{min}，则该码的纠错能力t(即具有纠t个随机错误的能力)可用下式表示：

$$t \leqslant \left[\frac{d_{min}-1}{2}\right], \tag{1-1}$$

式中$\left[\frac{d_{min}-1}{2}\right]$表示不大于$\frac{d_{min}-1}{2}$的整数。通常$t$取上限。可增大$d_{min}$来增大纠错能力$t$，但有一定的限度。

为了说明最小距离d_{min}，应先介绍汉明距离。两个码字u和v间的汉明距离$d(u,v)$是它们间的不同符号(即0或1)数。例如$u=10010110001$，$v=11001010101$，$d(u,v)=5$(即有5位的符号不相同)。而最小距离d_{min}则是在某码的范围内，所有可能的汉明距离中的最小数值。

(2) 纠突发错误码——设n为码字的位数，k为信息位数，该码的纠错能力(能纠正长l位的突发错误)可用下式表示：

$$l \leqslant \frac{n-k}{2}, \tag{1-2}$$

通常l取上限，可增加码多余度(即$r=n-k$)来增大纠错能力l(位)。

(3) 纠随机与突发错误码——以交错码为例，如把能纠t个随机错误的(n,k)码中λ个码字作为矩阵的λ行，然后按列发送，就可构成$(\lambda n,\lambda k)$交错码。即构成既能纠t个随机错误、又能纠t个长为λ的突发错误的码。

纠错码还可按构码用的数学方法(或称数学结构)来分类，这对学习编码理论和发展新的编码来说很有用。目下已可分为三类：1)代数码；2)几何码；3)算术码。详细情况可参阅以后各章。

1.5 基本概念介绍

本节采用具体例子来介绍纠错码的基本概念，目的在于使读者易于了解，并能引起对纠错编码的兴趣。主要列举重复码、线性分组码和线性卷积码。

1. 一般情况

为了检错和纠错，在编码时总是要在码字中引入多余度；即码组(也叫码字)长度n要比信息位数k大，增加了一些多余码元。

以重复码为例，该码用000来代表0，111来代表1，共有两个码字。显然，所增加的两位数字并不增多信息，是多余的，因而使传信效率降低。此外，除去传送信息的000和111两种组合(即两个码字)外，尚有：001，010，011，100，101，110六种组合未予采用。当信道上信噪比足够大时，我们可以肯定000和111中不会产生多于一个错误，如接收到001，010，

100，我们就可肯定实际上是000，即信息为“0”；同理，如接收到011，101，110，也可肯定为111，即信息为“1”。因此，多余码元可使我们检出一个错误，并且还可纠正这个错误。这就提高了传信的可靠性。

然而，错误是由随机噪声引起的，多余码元实际上并不能保证在任何情况下都可检出或纠正每个错误。在上例中，如果产生两个错误，那末按上述方法译码就会出现把信息“0”读作“1”或把信息“1”读作“0”的情况。如产生三个错误(虽然可能性不大)，则肯定会读错，而且会对所读出的结果毫不怀疑。因此，用编码方法来获得检错和纠错的能力很值得深入研究。

2．线性分组码

(1) 线性分组码的数学结构(一)——假设有M个等概出现的消息，它们的长度都是k位，存在着$M=2^k$的关系。今加上r个多余位，使每一个码字长度为$k+r\equiv n$位。这就使n位的二进制数字序列共有2^n个，但可能的消息数只有2^k个，故只能有2^k个码字〔即(2^n-2^k)个n位序列不是码字〕。此时码字的形式为：$m_1m_2\cdots m_k\ p_1p_2\cdots p_r$，在这里，$m_i$表示第$i$个信息位，$p_j$表示第$j$个多余位(也称校验位)，$m_i$和$p_j$都可能是0或1。

各个校验位都可从下列线性方程组求得：

$$\left.\begin{aligned}
&h_{11}m_1+h_{12}m_2+\cdots+h_{1k}m_k+1p_1+0p_2+\cdots+0p_r=0\\
&h_{21}m_1+h_{22}m_2+\cdots+h_{2k}m_k+0p_1+1p_2+\cdots+0p_r=0\\
&\cdots\cdots\\
&h_{r1}m_1+h_{r2}m_2+\cdots+h_{rk}m_k+0p_1+0p_2+\cdots+1p_r=0
\end{aligned}\right\}\qquad(1\text{-}3)$$

式中h_{ij}是常数，等于0或1。(1-3)式中每一个方程都代表多余位($p_1p_2\cdots p_r$)和信息位($m_1m_2\cdots m_k$)间的线性关系。以$k_{11}m_1+h_{12}m_2+\cdots+1p_1+0p_2+\cdots+0p_r=0$为例，$h_{11}m_1$，$h_{12}m_2$，$\cdots 1p_1$，$0p_2$，…都表示模2乘(有时用$h_{11}\odot m_1$表示)，$h_{11}m_1+h_{12}m_2+\cdots$表示模2加(有时用$h_{11}m_1\oplus h_{12}m_2\cdots$表示)。模2加法的运算规则为：$0\oplus1=1\oplus0=1$，$0\oplus0=1\oplus1=0$；模2乘法的运算规则为：$1\odot0=0\odot1=0\odot0=0$，$1\odot1=1$。

从校验方程组(1-3)式可写成校验矩阵：

$$H=\begin{bmatrix}
h_{11}h_{12}\cdots h_{1k} & 1 & 0 & 0\cdots0\\
h_{12}h_{22}\cdots h_{2k} & 0 & 1 & 0\cdots0\\
\cdots\cdots & & & \\
h_{r1}h_{r2}\cdots h_{rk} & 0 & 0 & 0\cdots1
\end{bmatrix}。\qquad(1\text{-}4)$$

该矩阵具有r行和n列，是码的一种数学结构。今将码字写成行矩阵：

$$v=(m_1m_2\cdots m_kp_1\cdots p_r)$$

故(1-3)式可写成：

$$vH^T=0,\qquad(1\text{-}5)$$

上式中v就是发送矢量(即码字)。如信道无扰，接收矢量r就是v(两者相同)；如信道有扰，r和v就不一定相同。设接收机中具有形成乘积rH^T的设备，则：

$rH^T\neq0$时，就知道已经发生了错误(至少为一个错误)；

$rH^T=0$时，就有两种可能性：1)未发生错误(即信道无扰)；2)已发生足够多的错误，使一个可能消息变成了另一个消息，无法检出这种错误(在信噪比较大时，此种情况不大会发生)。

(2) 纠正错误的方法——设发送矢量为v，接收矢量为r，出现的信道错误图样为e，可写

成下列关系式：

$$\boldsymbol{r}=\boldsymbol{v}+\boldsymbol{e} \tag{1-6}$$

为了纠错，必须知道$\boldsymbol{r}$中哪些位上存在错误。这可由校正子(又称伴随式)$\boldsymbol{s}$来确定：

$$\boldsymbol{s}=\boldsymbol{r}H^T=\boldsymbol{v}H^T+\boldsymbol{e}H^T=\boldsymbol{e}H^T。 \tag{1-7}$$

今设第i位是错误的，因此

$$\boldsymbol{e}=(00\cdots010\cdots0), \tag{1-8}$$

↑——第i位有错误

(1-8)式$\boldsymbol{e}$中元素为“0”的各位不存在错误，而“1”的各位确实存在错误。从(1-8)和(1-7)式得：

$$\boldsymbol{s}=\boldsymbol{e}H^T=(00\cdots010\cdots0)\begin{pmatrix} h_{11} & h_{21}\cdots h_{r1} \\ h_{12} & h_{22} \quad h_{r2} \\ \vdots & \vdots \quad \vdots \\ h_{1k} & h_{2k} \quad h_{rk} \\ 1 & 0 \quad 0 \\ 0 & 1 \quad 0 \\ \vdots & \vdots \quad \vdots \\ 0 & 0 \quad 1 \end{pmatrix}$$

$$=(h_{1i}h_{2i}\cdots h_{ri})。$$

↑——由此矢量示出 i 位是出现错误的位置

(3) 具体例子——假设要把16个等概率出现的消息构码，即$16=2^k$，故$k=4$位。为了纠正一个错误，$r=2$不行，因为$n=4+2=6$，校验矩阵H只有2行($r=2$)、6列($n=6$)，无法排列出各不相同的6列，故$r=2$显然不够用。6列应各不相同，主要目的是：使校正子$\boldsymbol{s}$能定出错误位置，进行纠错。倘改为$r=3$，就可排出(7,4)系统码的H矩阵：

$$\begin{matrix} & m_1 & m_2 & m_3 & m_4 & p_1 & p_2 & p_3 & \\ H= & 1 & 1 & 1 & 0 & 1 & 0 & 0 & \text{行1} \\ & 1 & 1 & 0 & 1 & 0 & 1 & 0 & \text{行2} \\ & 1 & 0 & 1 & 1 & 0 & 0 & 1 & \text{行3} \end{matrix} \tag{1-9}$$

该码的编码方法如下：先列出码字序列$m_1m_2m_3m_4p_1p_2p_3$，然后根据H中1、2、3行的次序画出有关位的联系线：

行1　$m_1m_2m_3m_4p_1p_2p_3$

行2　$m_1m_2m_3m_4p_1p_2p_3$

行3　$m_1m_2m_3m_4p_1p_2p_3$

如消息为1010，则从上列关系算出：

行1　1 0 1 0 p_1 p_2 p_3　　$1+0+1=0=p_1$

行2　1 0 1 0 p_1 p_2 p_3　　$1+0+0=1=p_2$

行3　1 0 1 0 p_1 p_2 p_3　　$1+1+0=0=p_3$

即算得码字为：1010010，表1-1列出(7.4)码的$2^4=16$个码字。

表1-1　　(7,4)码的16个码字

	m_1 m_2 m_3 m_4 p_1 p_2 p_3	汉明距离
	0 0 0 0 0 0 0	4
u	1 0 0 0 1 1 1	
		3
u'	0 1 0 0 1 1 0	4
	1 1 0 0 0 0 1	
		4
u''	0 0 1 0 1 0 1	4
v'	1 0 1 0 0 1 0	
		3
	0 1 1 0 0 1 1	4
	1 1 1 0 1 0 0	
		7
u'''	0 0 0 1 0 1 1	4
	1 0 0 1 1 0 0	
		3
	0 1 0 1 1 0 1	4
	1 1 0 1 0 1 0	
		4
	0 0 1 1 1 1 0	4
	1 0 1 1 0 0 1	
		3
v'	0 1 1 1 0 0 0	4
	1 1 1 1 1 1 1	

对于纠错和译码，从表1-1看出：最小汉明距离$d_{\min}=3$，故该码能纠$t=\left[\frac{d_{\min}-1}{2}\right]=\left[\frac{3-1}{2}\right]=1$个错误。如接收矢量$r=(1010111)$，则

$$s=rH^T=(1010111)\begin{pmatrix}1&1&1\\1&1&0\\1&0&1\\0&1&1\\1&0&0\\0&1&0\\0&0&1\end{pmatrix}=(101)$$

左面开始第三位是错的

由s知，发送矢量应为1000111(即表1-1中的u)，这就是经纠错后的译码结果。

顺便讲一下汉明权和汉明距离间的关系。汉明重量是指一个码字中的非零位数。今从表1-1中任取两个码字如v和u)，则

$$\begin{array}{r}v=0111000\\u=1000111\\\hline v+u=1111111\end{array}\quad(+)$$

v和u间的汉明距离$d(v,u)=7$，而$(v+u)$的汉明重量为：$W(v+u)=7$，因此得下列关系式：

$$d(v,u)=W(v+u)。\tag{1-10}$$

(4) 线性分组码的数学结构(二)——生成矩阵G是码的另一种数学结构。它是一个k行、n列矩阵：

$$G=\begin{pmatrix}v_1\\v_2\\\vdots\\v_k\end{pmatrix}=\begin{pmatrix}v_{11}&v_{12}&v_{13}\cdots v_{1n}\\v_{21}&v_{22}&v_{23}\cdots v_{2n}\\\vdots&&\\v_{k1}&v_{k2}&v_{k3}\cdots v_{kn}\end{pmatrix}。$$

(n,k)线性码的码字都可由G中各行的线性组合生成：

码字=(消息组)×(生成矩阵)

$$\begin{aligned}v&=m\times G\\&=(m_1m_2\cdots m_k)\begin{pmatrix}v_1\\v_2\\\vdots\\v_k\end{pmatrix}\\&=m_1v_1+m_2v_2+\cdots+m_kv_k,\end{aligned}\tag{1-11}$$

式中$m_i=0$或1，$i=1,2,\cdots,k$。故G称为生成矩阵。

(n,k)系统码的H矩阵见(1-4)式。所谓系统码，是指码字的k位消息组和$(n-k)$位校验组排列成次序分明、互不相混，如图1.3所示。且该码还有一个特点，可由H矩阵来求生成矩阵；如(1-12)和(1-13)式所示。

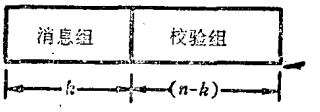

图 1.3 系统码示意图

$$H=\begin{pmatrix}h_{11}h_{12}\cdots h_{1k}10\cdots0\\h_{21}h_{22}\cdots h_{2k}01\cdots0\\\cdots\qquad\qquad\cdots\\h_{r1}h_{r2}\quad h_{rk}00\cdots1\end{pmatrix}_{r\times n}=[P_{r\times k}I_{r\times r}],\tag{1-12}$$

$$G=[I_{k\times k}(P^T)_{k\times r}],\tag{1-13}$$

式中P^T为P的转置矩阵，$I_{r\times r}$和$I_{k\times k}$为单位矩阵。今用(7,4)系统码的(1-9)式来说明H和G间的关系：

$$H=\begin{pmatrix}1&1&1&0&1&0&0\\1&1&0&1&0&1&0\\1&0&1&1&0&0&1\end{pmatrix}=[P_{3\times4}I_{3\times3}],$$

$$G=[I_{4\times4}(P^T)_{4\times3}]=\begin{pmatrix}1&0&0&0&1&1&1\\0&1&0&0&1&1&0\\0&0&1&0&1&0&1\\0&0&0&1&0&1&1\end{pmatrix}=\begin{pmatrix}u\\u'\\u''\\u'''\end{pmatrix}^{①}。$$

① 可从表1-1中找到u,u',u'',u'''

用G求码字，可按(1-11)式进行。如$m_1m_2m_3m_4=1010$，则码字

$$u'=(1010)\begin{pmatrix}1&0&0&0&1&1&1\\0&1&0&0&1&1&0\\0&0&1&0&1&0&1\\0&0&0&1&0&1&1\end{pmatrix}=(\,1\;0\;1\;1\;0\;0\;1\;0\,),$$

所得结果和从H算得的相同。

3. 线性卷积码

在1.4节中已讲到卷积码和分组码在编码方面不同之处。今以$n_0=2$，$k_0=1$，和$N=2$的(n_0,k_0,N)卷积码为例，图1.4示出该码的编码器。

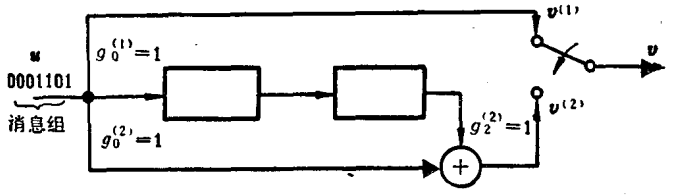

图1.4 (2,1,2)卷积码编码器

生成序列为：

$$g^{(1)}=(g_0^{(1)},\ g_1^{(1)},\ g_2^{(1)})=(100),$$

$$g^{(2)}=(g_0^{(2)},\ g_1^{(2)},\ g_2^{(2)})=(101),$$

输出序列为：

$$v^{(1)}=u*g^{(1)}=(v_0^{(1)},\ v_1^{(1)},\ v_2^{(1)},\cdots,v_6^{(1)}),\qquad(1\text{-}14)$$

$$v^{(2)}=u*g^{(2)}=(v_0^{(2)},\ v_1^{(2)},\ v_2^{(2)},\cdots,v_6^{(2)})。\qquad(1\text{-}15)$$

(1-14)和(1-15)两式为离散卷积，故称卷积码。该编码器输出为：

$$V=(v_0^{(1)}\ v_0^{(2)},\ v_1^{(1)}\ v_1^{(2)},\cdots,v_6^{(1)}\ v_6^{(2)})。\qquad\cdots(1\text{-}16)$$

表1-2示出(1-14)和(1-15)式的计算结果：

表1-2 (1-14)式和(1-15)式的计算结果：

	$g_0^{(1)}$	$g_1^{(1)}$	$g_2^{(1)}$	$v^{(1)}=u*g^{(1)}$
	1	0	0	
$u_6\ u_5\ u_4\ u_3\ u_2\ u_1\ u_0$ 0 0 0 1 1 0 1 →	1			$v_0^{(1)}=u_0\,g_0^{(1)}=1$
	0	1		$v_1^{(1)}=u_1\,g_0^{(1)}+u_0\,g_1^{(1)}=0$
	1	0	1	$v_2^{(1)}=u_2\,g_0^{(1)}+u_1\,g_1^{(1)}+u_0\,g_2^{(1)}=1$
	1	1	0	$v_3^{(1)}=u_3\,g_0^{(1)}+u_2\,g_1^{(1)}+u_1\,g_2^{(1)}=1$
	0	1	1	$v_4^{(1)}=u_4\,g_0^{(1)}+u_3\,g_1^{(1)}+u_2\,g_2^{(1)}=0$
	0	0	1	$v_5^{(1)}=u_5\,g_0^{(1)}+u_4\,g_1^{(1)}+u_3\,g_2^{(1)}=0$
	0	0	0	$v_6^{(1)}=u_6\,g_0^{(1)}+u_5\,g_1^{(1)}+u_4\,g_2^{(1)}=0$
				$v^{(1)}=v_0^{(1)}\ v_1^{(1)}\ v_2^{(1)}\ v_3^{(1)}\ v_4^{(1)}\ v_5^{(1)}\ v_6^{(1)}$
				$=1\ \ 0\ \ 1\ \ 1\ \ 0\ \ 0\ \ 0$

	$g_0^{(2)}$	$g_1^{(2)}$	$g_2^{(2)}$	$v^{(2)}=u*g^{(2)}$
	1	0	1	
$u_6\ u_5\ u_4\ u_3\ u_2\ u_1\ u_0$ →				
0 0 0 1 1 0 1	1			$v_0^{(2)}=u_0 g_0^{(2)}=1$
	0	1		$v_1^{(2)}=u_1 g_0^{(2)}+u_0 g_1^{(2)}=0$
	1	0	1	$v_2^{(2)}=u_2 g_0^{(2)}+u_1 g_1^{(2)}+u_0 g_2^{(2)}=0$
	1	1	0	$v_3^{(2)}=u_3 g_0^{(2)}+u_2 g_1^{(2)}+u_1 g_2^{(2)}=1$
	0	1	1	$v_4^{(2)}=u_4 g_0^{(2)}+u_3 g_1^{(2)}+u_2 g_2^{(2)}=1$
	0	0	1	$v_5^{(2)}=u_5 g_0^{(2)}+u_4 g_1^{(2)}+u_3 g_2^{(2)}=1$
	0	0	0	$v_6^{(2)}=u_6 g_0^{(2)}+u_5 g_1^{(2)}+u_4 g_2^{(2)}=0$

$$v^{(2)}=v_0^{(2)}\ v_1^{(2)}\ v_2^{(2)}\ v_3^{(2)}\ v_4^{(2)}\ v_5^{(2)}\ v_6^{(2)}$$
$$=1\ 0\ 0\ 1\ 1\ 1\ 0$$

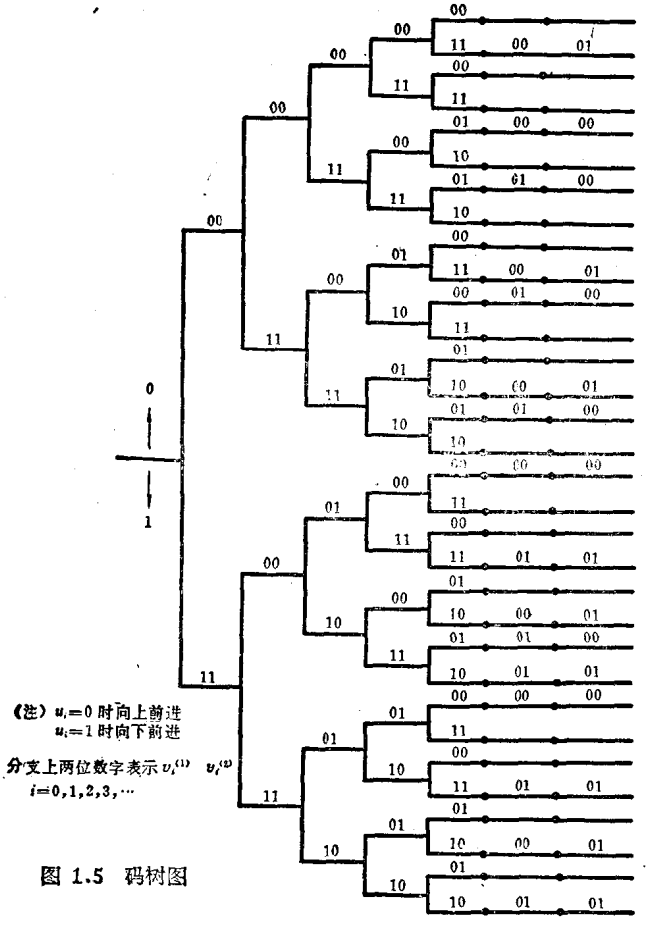

图 1.5 码树图