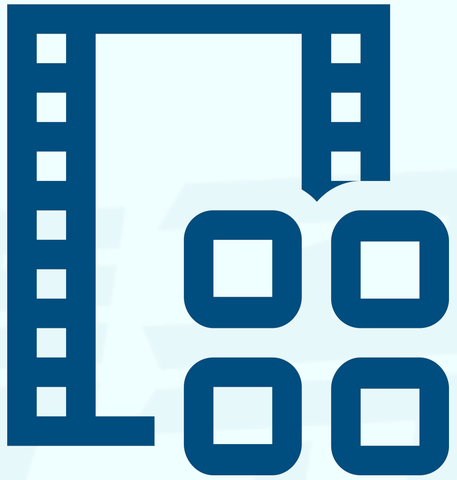
信道的纠错编码

信息论与编码原理不挂科 第八讲

信道的纠错编码

信息论与编码原理不挂科 第八讲



4大模块

5道题目

信息论与编码原理不挂科 第六讲

信道的纠错编码

模块1 纠错编码的基本概念

模块2 线性分组码

模块3 汉明码

模块4 循环码

纠错编码的基本概念

纠错编码的基本概念

纠错编码概述

香农第二定理表明，当$R < C$时，$P_E \to 0$的码存在。

香农第二定理的证明过程采用的是随机编码的方法:

- 随机编码所得的码集很大，通过搜索得到好码的方法在实际上很难实现;
- 即使找到了好码，这种码的码字也没有规律，不便于译码。

真正实用的信道编码方法还需要通过各种数学工具来构造，使码具有好的结构性以便于译码。

纠错编码概述

纠错编码的基本思路是，根据一定的规律，在待发送的信息码元中**人为的加入一些冗余码元（即监督码元）**，这些冗余码元与信息码元之间**以某种确定的规则相互关联与约束**。

在接收端按照既定的规则检验信息码元与监督码元之间的关系。如果传输过程出错，那么信息码元与监督码元之间的关系将受到破坏，从而可以发现错误乃至纠正错误。

纠错编码的基本概念

三种常用差错控制方式

1 反馈重传（ARQ）

发送端经编码后发出能够发现错误的码，接收端收到后经检验，如果发现传输中有错误，则通过反馈系统把这一判断结果反馈回发端，然后发送端把前面发出的信息重新传送一次， 直到接收端认为正确地收到信息为止。

2 前向纠错（FEC）

发送端发出的是具有纠错能力的纠错码，接收端根据译码规则进行译码。当误码个数在码的纠错能力范围内时，译码器可以自动纠正错误。

3 混合纠错

对发送端进行适当的编码。当错误不严重，在码的纠错能力范围之内时，采用自动纠错；当产生的差错超 出码的纠错能力范围时，通过反馈系统要求发端重发。混合纠错在实时性和译码复杂性方面是前向纠错和检错重发的折衷。

纠错编码的基本概念

纠错编码的分类

按码的功能分类

- 检错码：仅能检测误码
- 纠错码：可纠正误码
- 纠删码：纠错和恢复删除掉的数据能力，要求删除丢失的数据要有位置信息。

按信息码元与监督码元之间的检验关系分类

- 线性码：满足线性关系，满足一组线性方程
- 非线性码：不存在线性关系

纠错编码的分类

按信息码元与监督码元之间的约束方式不同分类

- 分组码：本码组的监督码元仅和本码组的信息元相关。
- 卷积码：本码组的监督码元不仅和本码组的信息元相关，而且与前面码组的信息码元有关。

按信息码元在编码后是否保持原形式不变分类

- 系统码：信息码元与监督码元在分组内有确定位置，编码后的信息码元保持不变；
- 非系统码：信息位打乱，与编码前不同。对观察和译码都带来麻烦，较少使用。

按纠正差错的类型可分类

- 纠随机错误码：纠正随机噪声引起的差错，差错的出现互不相关，彼此独立；
- 纠突发错误码：纠正脉冲干扰引起的差错，错误之间存在相关性；
- 纠随机和突发错误码：多种错误并存

线性分组码

线性分组码

线性分组码的基本概念

假设二元信息码组由k个信息码元组成，那么一共有2^k个不同的信息码组，即可以表示2^k个消息；

而编码器输出长度为n的码字，其中附加$r=n-k$个校验码元；

每个校验码元都是该信息码组某些信息码元的**模2和**；

则称这2^k个码字的集合为(n,k)**分组码**。

通过分析我们可以知道该分组码的编码信息率为$R=\frac{\log 2^k}{n}=\frac{k}{n}$，指码字中包含实际信息的多少。

我们不难发现二进制(n,k)分组码有2^k个合法码字，而整个码符号空间中有2^n个码字；

而任意一种2^k信息集合到二进制序列集合2^n的映射都是一种(n,k)分组码，因此一共有$C_{2^n}^{2^k}$种方案。

在这众多的编码方案之中，**线性分组码**是最具有实用价值的一类码，本讲我们将介绍两种主要的线性分组码，分别是**汉明码和循环码**。

线性分组码

校验矩阵

我们以(7,3)线性分组码为例，我们用一个码字矢量来表示其码字为

$$C = [c_1 , c_2 , c_3 , c_4 , c_5 , c_6 , c_7]$$

其中信息码元为c_1，c_2，c_3，校验码元为c_4，c_5，c_6，c_7；

其中校验码元我们可以根据下列方程得到（加法为模2和）

$$\begin{cases} c_4 = c_1 + c_3 \\ c_5 = c_1 + c_2 + c_3 \\ c_6 = c_1 + c_2 \\ c_7 = c_2 + c_3 \end{cases} \xrightarrow{\text{模2和}} \begin{cases} c_1 + c_3 + c_4 = 0 \\ c_1 + c_2 + c_3 + c_5 = 0 \\ c_1 + c_2 + c_6 = 0 \\ c_2 + c_3 + c_7 = 0 \end{cases}$$

我们可以得到该线性分组码的所有码字如右表所示。

信息位	对应码字
000	0000000
001	0011101
010	0100111
011	0111010
100	1001110
101	1010011
110	1101001
111	1110100

校验矩阵

$$\begin{cases} c_1 + c_3 + c_4 = 0 \\ c_1 + c_2 + c_3 + c_5 = 0 \\ c_1 + c_2 + c_6 = 0 \\ c_2 + c_3 + c_7 = 0 \end{cases} \rightarrow \underbrace{\left[\begin{array}{ccc:cccc} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]}_{\text{信息位}\quad\quad\text{校验位}} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow HC^T = 0^T$$

我们称矩阵H为校验矩阵，它表征的是监督码元与信息码元之间的关系；

由于所有码字都按同一个规则确定，因此该矩阵又称为一致监督方程或一致校验方程；

该矩阵的特点是校验码元的个数与矩阵的行数相等，(n, k) 线性分组码校验矩阵为$(n-k) \times n$ 维矩阵；

该矩阵**只可进行行变换，不可进行列变换**，否则会打乱监督码元与信息码元之间的约束关系！

校验矩阵

$$\begin{cases} c_1 + c_3 + c_4 = 0 \\ c_1 + c_2 + c_3 + c_5 = 0 \\ c_1 + c_2 + c_6 = 0 \\ c_2 + c_3 + c_7 = 0 \end{cases} \rightarrow \underbrace{\left[\begin{array}{ccc|cccc} 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 & 1 \end{array}\right]}_{\text{信息位}\quad\text{校验位}} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \rightarrow HC^T = 0^T$$

$$HC^T = 0^T \rightarrow (HC^T)^T = CH^T = (0^T)^T = 0$$

对于**系统码**，也就是信息码元与监督码元在分组内有确定位置，编码后的信息码元保持不变的码，我们可以将校验矩阵表示为

$$H = [Q \quad I]$$ 信息位在前，校验位在后

其中矩阵Q为$(n-k)\times k$ 维矩阵，I 为$(n-k)\times(n-k)$维单位阵，这样的矩阵为标准/典型校验矩阵。

生成矩阵

$$\begin{cases} c_1 + c_3 + c_4 = 0 \\ c_1 + c_2 + c_3 + c_5 = 0 \\ c_1 + c_2 + c_6 = 0 \\ c_2 + c_3 + c_7 = 0 \end{cases} \rightarrow \begin{cases} c_1 = c_1 \\ c_2 = c_2 \\ c_3 = c_3 \\ c_4 = c_1 + c_3 \\ c_5 = c_1 + c_2 + c_3 \\ c_6 = c_1 + c_2 \\ c_7 = c_2 + c_3 \end{cases} \rightarrow [c_1\ c_2\ c_3] \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix} = [c_1\ c_2\ c_3\ c_4\ c_5\ c_6\ c_7]$$

我们令信息位矢量为 $m = [c_1\ c_2\ c_3]$，矩阵 $G = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$，则上述矩阵方程可以表示为

$$C = mG$$

也就是说，如果已知矩阵G，我们根据信息位就可以获得包含校验位的整个对应码字。

我们称矩阵G为生成矩阵

生成矩阵

$$\begin{cases} c_1 + c_3 + c_4 = 0 \\ c_1 + c_2 + c_3 + c_5 = 0 \\ c_1 + c_2 + c_6 = 0 \\ c_2 + c_3 + c_7 = 0 \end{cases} \rightarrow \begin{cases} c_1 = c_1 \\ c_2 = c_2 \\ c_3 = c_3 \\ c_4 = c_1 + c_3 \\ c_5 = c_1 + c_2 + c_3 \\ c_6 = c_1 + c_2 \\ c_7 = c_2 + c_3 \end{cases} \rightarrow [c_1\ c_2\ c_3] \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix} = [c_1\ c_2\ c_3\ c_4\ c_5\ c_6\ c_7]$$

$$G = \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix}$$ (n, k) 线性分组码生成矩阵为 $k \times n$ 维矩阵

对于**系统码**，生成矩阵可以表示为 $C = mG$ ，且

$$G = [I \quad P]$$ 信息位在前，校验位在后

其中矩阵 P 为 $k \times (n-k)$ 维矩阵，I 为 $k \times k$ 维单位阵，这样的矩阵为标准/典型生成矩阵。

生成矩阵

把生成矩阵的每一行用一个行向量 $G_i\ (i=1,2,...,k)$ 表示，则生成矩阵可以表示为 $G=\begin{bmatrix}G_1\\G_2\\\cdots\\G_k\end{bmatrix}$

设信息位 $m=[m_1\ m_2\ \cdots\ m_k]$，则信息位，生成矩阵与码字的关系为

$$C=mG=\sum_{i=1}^{k}m_iG_i$$

例如：$m=[0\ 1\ 1]$ $G=\begin{bmatrix}1&0&0&1&1&1&0\\0&1&0&0&1&1&1\\0&0&1&1&1&0&1\end{bmatrix}=\begin{bmatrix}G_1\\G_2\\G_3\end{bmatrix}$

$$C=mG=[0\ 1\ 1]\begin{bmatrix}1&0&0&1&1&1&0\\0&1&0&0&1&1&1\\0&0&1&1&1&0&1\end{bmatrix}=0\cdot G_1+1\cdot G_2+1\cdot G_3=[0100111]+[0011101]=[0111010]$$

生成矩阵

生成矩阵有一个重要的性质，那就是**生成矩阵的每一行都是一个合法的码字**。

当信息码组 $m = [m_1\ m_2\ \cdots\ m_k]$ 中仅有一个非零元素时，得到的码字即为**典型生成矩阵的某一行**。

$$C = mG = \sum_{i=1}^{3} m_i G_i = [m_1\ m_2\ m_3] \begin{bmatrix} G_1 \\ G_2 \\ G_3 \end{bmatrix}$$

$m = [1\ 0\ 0]$ 时，所得码字对应典型生成矩阵中的第一行。

$m = [0\ 1\ 0]$ 时，所得码字对应典型生成矩阵中的第二行。

$m = [0\ 0\ 1]$ 时，所得码字对应典型生成矩阵中的第三行。

由k个不相同的（线性无关的）非零码字可以构成一个生成矩阵，而由这k个不同的码字的不同线性组合可以生成整个码。（不同的线性组合根据信息位确定）

校验矩阵与生成矩阵的关系

生成矩阵的每一行都是一个合法的码字，因此对于生成矩阵 $G=\begin{bmatrix}G_1\\G_2\\\cdots\\G_k\end{bmatrix}$

存在 $HG_i^T=0^T$，因此对于整个生成矩阵，有 $HG^T=0^T$ 。

对于标准形式的校验矩阵与生成矩阵，有

$$HG^T=[Q\ I][I\ P]^T=[Q\ I]\begin{bmatrix}I\\P^T\end{bmatrix}=Q+P^T=0^T \quad \Rightarrow \quad Q=P^T$$

因此根据上述关系，**系统线性码**的校验矩阵与生成矩阵可以互换

$$H=[Q\ I]=[P^T\ I] \quad \Leftrightarrow \quad G=[I\ P]=[I\ Q^T]$$

例题8-1 已知（7，4）线性分组码的典型校验矩阵为 $H=\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$，试直接写出对应的生成矩阵。

解析8-1 $H=[Q\ I]=[P^T\ I] \quad \Leftrightarrow \quad G=[I\ P]=[I\ Q^T]$

$$Q=\begin{bmatrix}1&1&1&0\\0&1&1&1\\1&1&0&1\end{bmatrix} \rightarrow Q^T=\begin{bmatrix}1&0&1\\1&1&1\\1&1&0\\0&1&1\end{bmatrix} \qquad G=[I\ Q^T]=\begin{bmatrix}1&0&0&0&1&0&1\\0&1&0&0&1&1&1\\0&0&1&0&1&1&0\\0&0&0&1&0&1&1\end{bmatrix}$$

线性分组码

线性分组码的伴随式

线性分组码的伴随式，是指伴随接受码字的n-k维向量，它反映了**“信道对码字造成的干扰”**。

假设发送码字为C，接收到的码符号序列为Y，令伴随式为

$$S = YH^T \qquad 或 \qquad S^T = HY^T$$

若伴随式$S = 0$，则符合校验方程，此时Y是一个合法的码字；

若伴随式$S \neq 0$，则不符合校验方程，此时Y不是一个合法的码字，传输过程中产生了误码；

若令接受码字 $Y = C + E$，C为合法码字，E可视为附加在合法码字导致出错的**错误图样**；

此时，接受码字满足方程 $S = YH^T = (C + E)H^T = CH^T + EH^T = EH^T$ 。

或满足方程 $S^T = HY^T = H(C + E)^T = HC^T + HE^T = HE^T$。（两式互为转置关系），即伴随式仅与错误图样有关。

线性分组码的伴随式

此时，接受码字满足方程 $S = YH^T = (C+E)H^T = CH^T + EH^T = EH^T$ 。

或满足方程 $S^T = HY^T = H(C+E)^T = HC^T + HE^T = HE^T$ 。（两式互为转置关系），即伴随式仅与错误图样有关。

设 $E = [e_1, e_2, \cdots, e_n]$, $H = [H_1, H_2, \cdots, H_n]$ ，即将校验矩阵划分成列向量的形式；

那么伴随式 $S^T = HE^T = [H_1, H_2, \cdots, H_n]\begin{bmatrix} e_1 \\ e_2 \\ \cdots \\ e_n \end{bmatrix} = \sum_{i=1}^{N} e_i H_i$

分析这个方程，我们可以发现：

1. 当传输过程没有错误，即 $E = [0,0,\cdots,0]$ 时，$S^T = 0$ ；
2. 当传输过程有一位错误时，S^T 为校验矩阵中的某一列；
3. 当传输过程有多位错误，S^T 为校验矩阵中若干列的模2和。

给出了已知错误图样后伴随式的情形，初步明确了纠错的方向！

线性分组码的伴随式

根据 $S^T = HY^T = H(C+E)^T = HC^T + HE^T = HE^T$，我们发现，伴随式可以通过接受码字与校验矩阵之间的运算得到，且此时错误图样E也就随之确认，即 $S^T = HE^T$ 或 $S = EH^T$。

上式为一个线性方程组，它的解不唯一，也就是说，对于一个伴随式，错误图样E不唯一。

假设错误图样的其中一个解为 E_0，即 $E_0H^T = S$，那么对于任意一个码组集合中的合法码字C，恒有

$$(E_0 + C)H^T = E_0H^T + CH^T = E_0H^T = S$$

这个方程组一共有 2^k 个解，即 2^k 个错误图样与一个伴随式对应，但是真正的错误图样只有一种。

译码器必须从这些错误图样中选择正确的错误图样，才可以保证译码正确，否则会出现错误。

若为二进制对称信道传输，我们**选择错误图样中非零元素最少**的那一个作为最终的错误图样。

错误图样非零元素（码重）最少，也就是需要纠错的位数最少。

此时正确的码字为 $C = Y - E = Y + E$。

线性分组码的伴随式

对于已给线性分组码，一个伴随式对应一种错误图样E， 即可纠正对应的错误。

则n-k位的伴随式有2^{n-k}种，其与可纠正错误数目u的关系为：$2^{n-k} \geq \sum_{i=0}^{u} C_n^i$

当等号成立时，符合**完备码**的条件，即**一个伴随式对应一个错误图样**，它从整体角度提供了纠错性能的必要条件。

那么，我们将线性分组码的纠错步骤列写如下：

1. 根据校验方程，确定校验矩阵；
2. 对于接受码字，求解对应的伴随式；
3. 根据伴随式，若为二进制对称信道，则选择码重最小的错误图样（也就是出错位数最少），从而确定出错码元的位置；
4. 找到出错码元的位置以后，根据接受码字和错误图样进行纠错。

例题8-2 已知（7，3）线性分组码的校验矩阵为 $H=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}$，试判断接收到的如下码组是否有误码，并尝试纠错。

(1) $Y=(1010011)$

(2) $Y=(1110011)$

(3) $Y=(0011011)$

解析8-2 (1) $Y=(1010011)$时，求解此时的伴随式 $S^T=HY^T=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}\begin{bmatrix}1\\0\\1\\0\\0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\0\\0\\0\end{bmatrix}$

因此传输过程中没有误码；

例题8-2 已知（7，3）线性分组码的校验矩阵为 $H=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}$，试判断接收到的如下码组是否有误码，并尝试纠错。

(1) $Y=(1010011)$

(2) $Y=(1110011)$

(3) $Y=(0011011)$

解析8-2 (2) $Y=(1110011)$时，求解此时的伴随式 $S^T=HY^T=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}\begin{bmatrix}1\\0\\1\\0\\0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\1\\1\\1\end{bmatrix}$

$$Y=C+E \quad \Rightarrow \quad S=HY^T=H(C+E)^T=HC^T+HE^T=HE^T$$

$$S^T=HE^T=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}\begin{bmatrix}0\\1\\0\\0\\0\\0\\0\end{bmatrix}=\begin{bmatrix}0\\1\\1\\1\end{bmatrix}$$

对应校验矩阵的第二列，故第二位错，错误图样为 $E=(0100000)$

则纠错后正确的码组为 $C=Y+E=(1010011)$

（也有可能为后三位同时出错，但我们现阶段一般考虑错较少的位数）

例题8-2 已知（7，3）线性分组码的校验矩阵为 $H=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}$，试判断接收到的如下码组是否有误码，并尝试纠错。

(1) $Y=(1010011)$

(2) $Y=(1110011)$

(3) $Y=(0011011)$

解析8-2 (3) $Y=(0011011)$时，求解此时的伴随式 $S^T=HY^T=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix}\begin{bmatrix}1\\0\\1\\0\\0\\1\\1\end{bmatrix}=\begin{bmatrix}0\\1\\1\\0\end{bmatrix}$

$\begin{bmatrix}0\\1\\1\\0\end{bmatrix}=\begin{bmatrix}1\\1\\1\\0\end{bmatrix}+\begin{bmatrix}1\\0\\0\\0\end{bmatrix}=\begin{bmatrix}0\\1\\1\\1\end{bmatrix}+\begin{bmatrix}0\\0\\0\\1\end{bmatrix}=\begin{bmatrix}0\\1\\0\\0\end{bmatrix}+\begin{bmatrix}0\\0\\1\\0\end{bmatrix}$ 无法确定具体哪两位出错，因此有误码但无法纠错。

线性分组码

线性分组码的封闭性

线性分组码的封闭性指的是：线性分组码中任意两个码字之和依然是该码的合法码字。

也就是说，假设C_1和C_2分别是线性分组码C中的两个合法码字，那么满足$HC_1^T = 0^T$，$HC_2^T = 0^T$；

那么$H(C_1 + C_2)^T = HC_1^T + HC_2^T = 0^T$；

也就是说这两个码字之和所构成的新序列也满足校验方程，因此也为合法码字。

线性分组码的检纠错能力

- **重要定理**

  线性分组码的最小汉明距离，等于该码中非零码字的最小重量（码字重量即码元1的个数）。

  证明：

  设线性分组码码集为C，且码字 $U \in C$, $V \in C$ ；设码字 $Z = U + V$ ，根据线性分组码的封闭性，我们有 $Z \in C$ ；因此，如果U与V的汉明距离为d，即两个码字有d个码符号不一样，因此，两个码字的模2和有d个1（不一样的码符号模2和为1），也就是两码字模2和后码字重量为d。

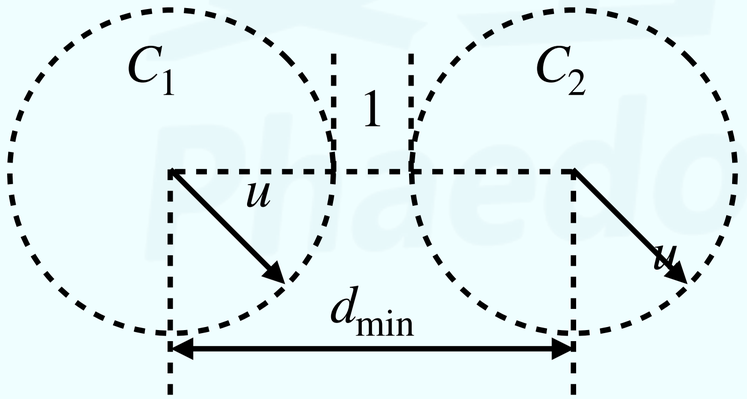
  因此我们可证得线性分组码的最小汉明距离等于该码中非零码字的最小重量。

线性分组码的检纠错能力

设线性分组码的最小汉明距离为 $d_{\min}$

- **线性分组码的纠错能力**

  该线性分组码具备纠正u个及以内的错误的充分必要条件为 $d_{\min} = 2u + 1$。



例：三元重复编码 $d_{\min} = 3$，可纠一位错。

可纠错

$$000 \begin{cases} 001 \\ 010 \\ 100 \end{cases}$$

可纠错

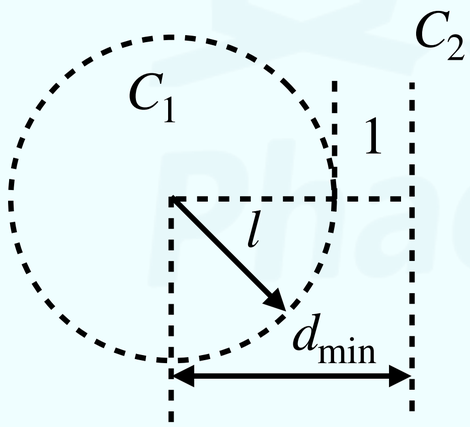
$$111 \begin{cases} 011 \\ 101 \\ 110 \end{cases}$$

线性分组码的检纠错能力

设线性分组码的最小汉明距离为 $d_{\min}$

- **线性分组码的检错能力**

  该线性分组码具备检测l及以内个错误的充分必要条件为 $d_{\min} = l + 1$ 。



例：三元重复编码 $d_{\min} = 3$，可检两位错。

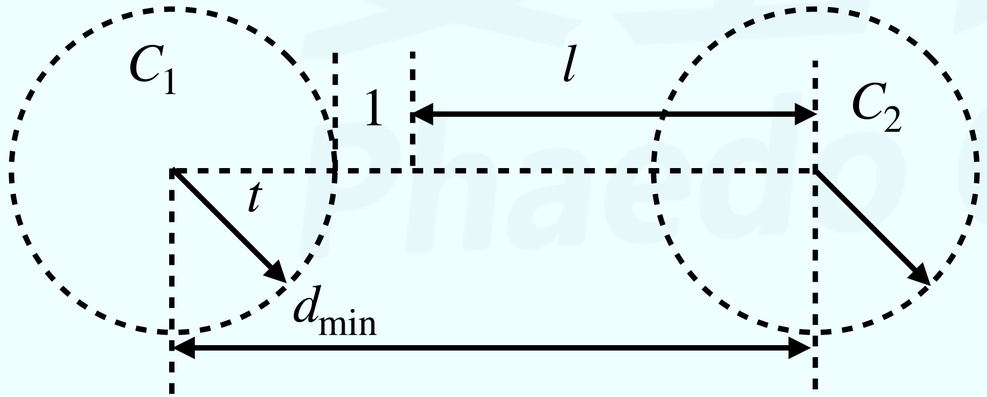
	可检错	
000/111	001	100
	010	101
	011	110

线性分组码的检纠错能力

设线性分组码的最小汉明距离为 $d_{\min}$

- **线性分组码的同时检、纠错能力**

  该线性分组码具备纠正t个错误，同时还可以发现l个错误($l > t$)的充分必要条件为 $d_{\min} = t + l + 1$。



例：四元重复编码 $d_{\min} = 4$，可检两位错，纠一位错。

可纠错

$$0000 \begin{cases} 0001 & 0100 \\ 0010 & 1000 \end{cases}$$

可纠错

$$1111 \begin{cases} 0111 & 1101 \\ 1011 & 1110 \end{cases}$$

可检错

$$0000/1111 \begin{cases} 0011 & 0110 & 1100 \\ 1001 & 1010 & 0101 \end{cases}$$

线性分组码

校验矩阵与最小距离的关系

对于线性分组码，设校验矩阵为H。

若校验矩阵H 中任意 t 列线性无关，存在 $t+1$ 列线性相关，则该码的最小汉明距离为$t+1$。

反之，若该码的最小汉明距离为$t+1$，则校验矩阵H 中任意 t 列线性无关，存在 $t+1$ 列线性相关。

若校验矩阵H 中任意 t 列线性无关，则$[H_1\ H_2\ \cdots\ H_n]\begin{bmatrix}1\\1\\\cdots\\1\end{bmatrix}_{\text{含}t\text{个}1}=0$的情况不存在，即没有码字有$t$个1；

若校验矩阵H 中存在 $t+1$列线性相关，则$[H_1\ H_2\ \cdots\ H_n]\begin{bmatrix}1\\1\\\cdots\\1\end{bmatrix}_{\text{含}t+1\text{个}1}=0$的情况存在，码字最少有$t+1$个1。

因此该线性分组码的最小码重为$t+1$，因此最小汉明距离为$t+1$。

校验矩阵与最小距离的关系

对于线性分组码，设校验矩阵为H。

若校验矩阵H 中任意 t 列线性无关，存在 $t+1$ 列线性相关，则该码的最小汉明距离为$t+1$。

反之，若该码的最小汉明距离为$t+1$，则校验矩阵H 中任意 t 列线性无关，存在 $t+1$ 列线性相关。

如果校验矩阵中任何一列不为零矢量，且任何两列都不相等（即任何两列线性无关），则码字的最小距离至少是3。

汉明码

小节1 汉明码的定义

小节2 汉明码的构造

汉明码

小节1 汉明码的定义

小节2 汉明码的构造

汉明码

汉明码的定义

汉明码是一种**能够纠正单个错误的线性分组码**，其最小汉明距离为3。

对于(n,k)汉明码，假设校验位一共有 $m = n - k$ 位，那么其总码长与信息位个数满足：

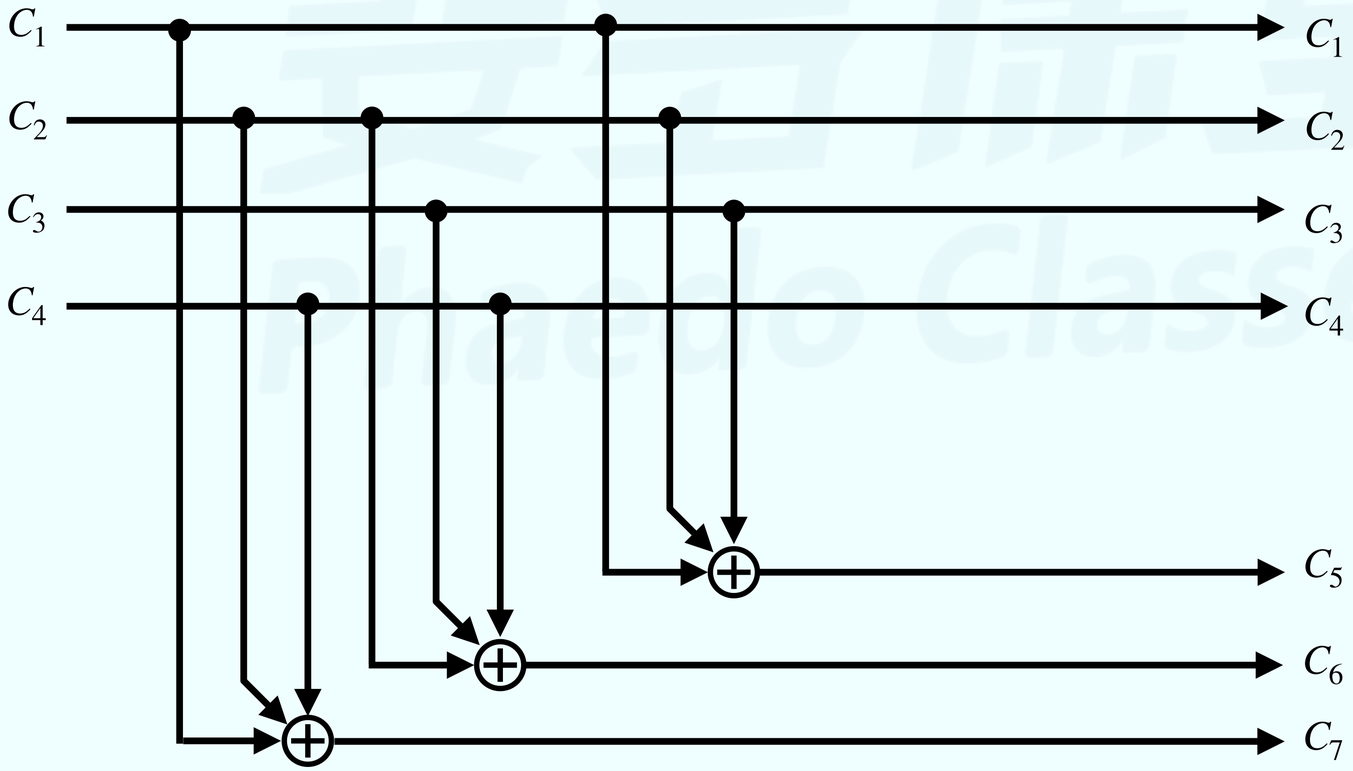
$n = 2^m - 1 \qquad k = 2^m - 1 - m$

比如$(7,4)$汉明码，$(3,1)$汉明码。

汉明码的编码电路与译码电路

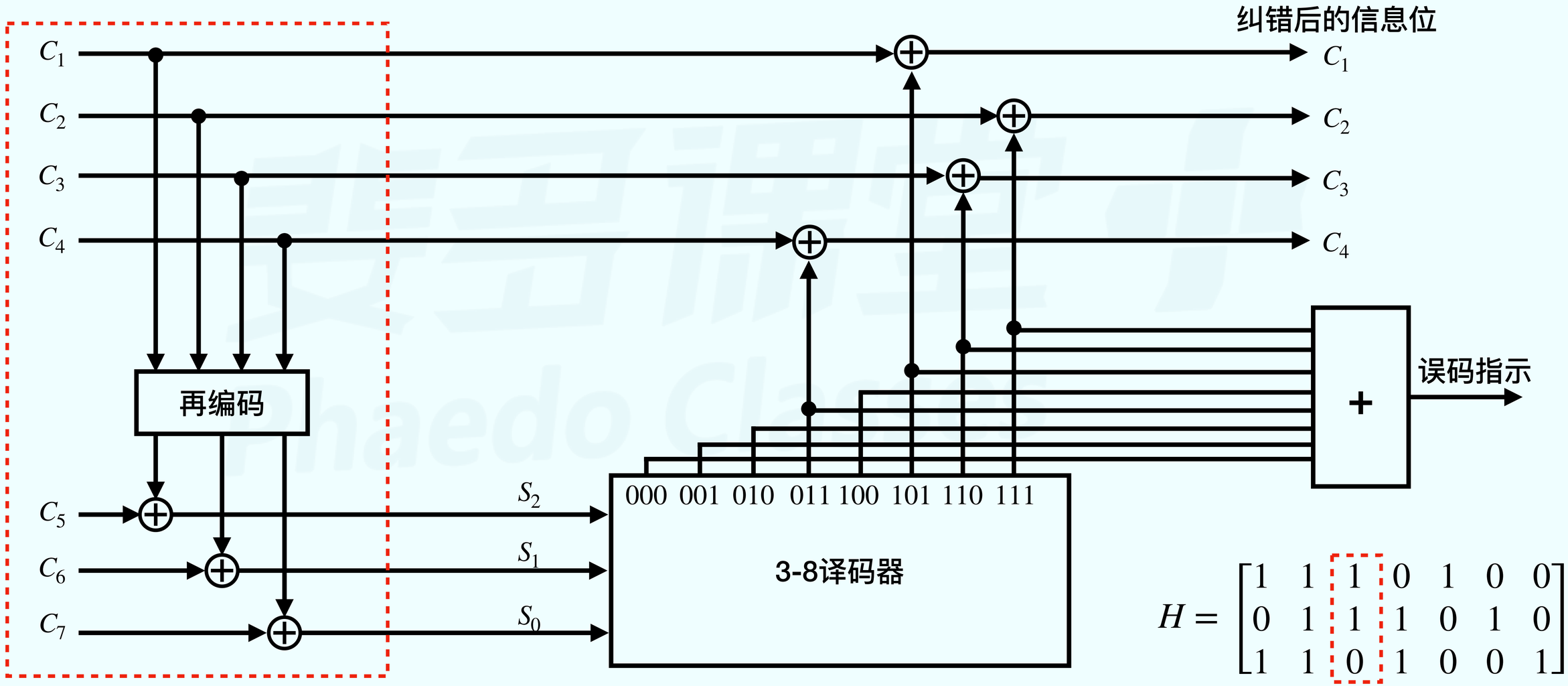
$$H=\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$$

$$\begin{cases}c_5=c_1+c_2+c_3\\c_6=c_2+c_3+c_4\\c_7=c_1+c_2+c_4\end{cases}$$



根据校验矩阵或者校验方程，作出(7,4)汉明码的编码电路如左图所示。

汉明码的编码电路与译码电路

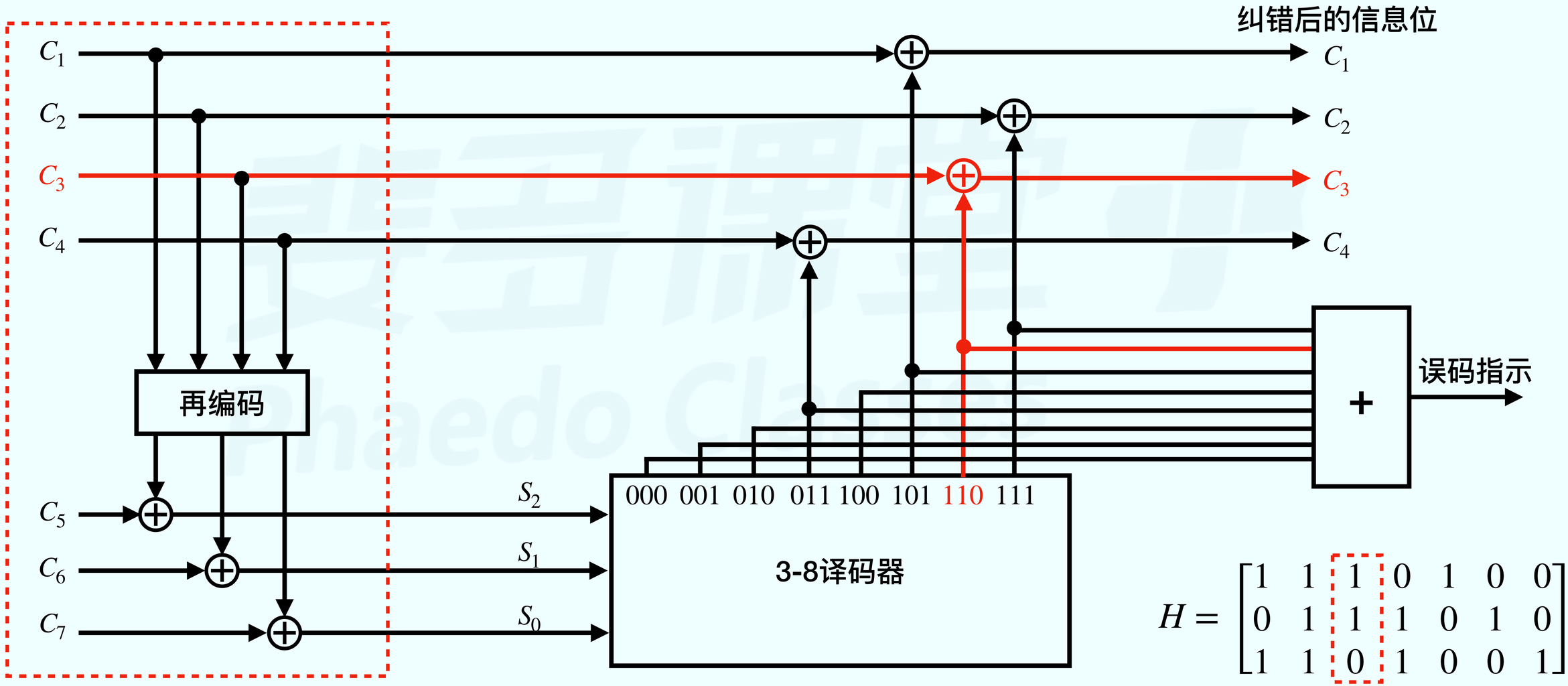


$$H=\begin{bmatrix}1&1&1&0&1&0&0\\0&1&1&1&0&1&0\\1&1&0&1&0&0&1\end{bmatrix}$$

求解伴随式的过程

根据伴随式确定出错位，进行错误图样的附加纠错

汉明码的编码电路与译码电路



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

求解伴随式的过程

根据伴随式确定出错位，进行错误图样的附加纠错

汉明码

小节1 汉明码的定义

小节2 汉明码的构造

汉明码生成矩阵的构造

构造方法

n列由除全0以外的$m=(n-k)$位码组构成，每个码组只在某列中出现一次。当发生可纠的单个错误时，伴随式即为为 H 阵中对应的列，译码比较方便。

根据上述方法得到的是非典型校验矩阵，**若对约束关系不作要求时，可以通过列置换将非典型生成矩阵变换为典型**校验矩阵；**若对约束关系有要求时，只能做初等行变换**。

通过列置换构造典型校验矩阵得到的码字与原先的约束关系会发生改变，因此生成的码会变化，但纠错能力不变，因此在构造只满足一定信息位与校验位数目条件的汉明码时，可以列置换。

例题8-3 构造符合如下要求的汉明码：

(1) $m=3$ 的系统汉明码； (2) $m=4$ 的汉明码。

解析8-3 (1) $m=3$ 时，$n=2^m-1=7$ ，$k=n-m=4$

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

(列置换法，此时只给出了码元数目的分配，可以进行列置换)

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

(行初等变换法)

(2) $m=4$ 时，$n=2^m-1=15$，$k=n-m=11$

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

(未强调系统码时可以不变换)

循环码

小节1 循环码多项式与循环移位

小节2 循环码的生成多项式与生成矩阵

循环码

小节1 循环码多项式与循环移位

小节2 循环码的生成多项式与生成矩阵

循环码的特点

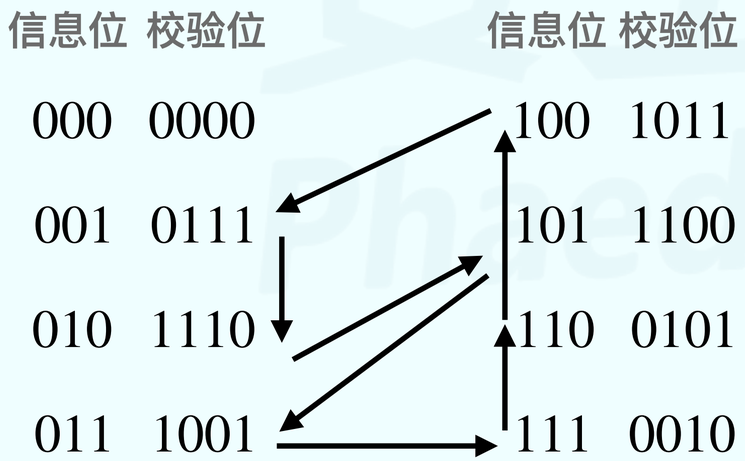
循环码是线性分组码的一个重要子集。

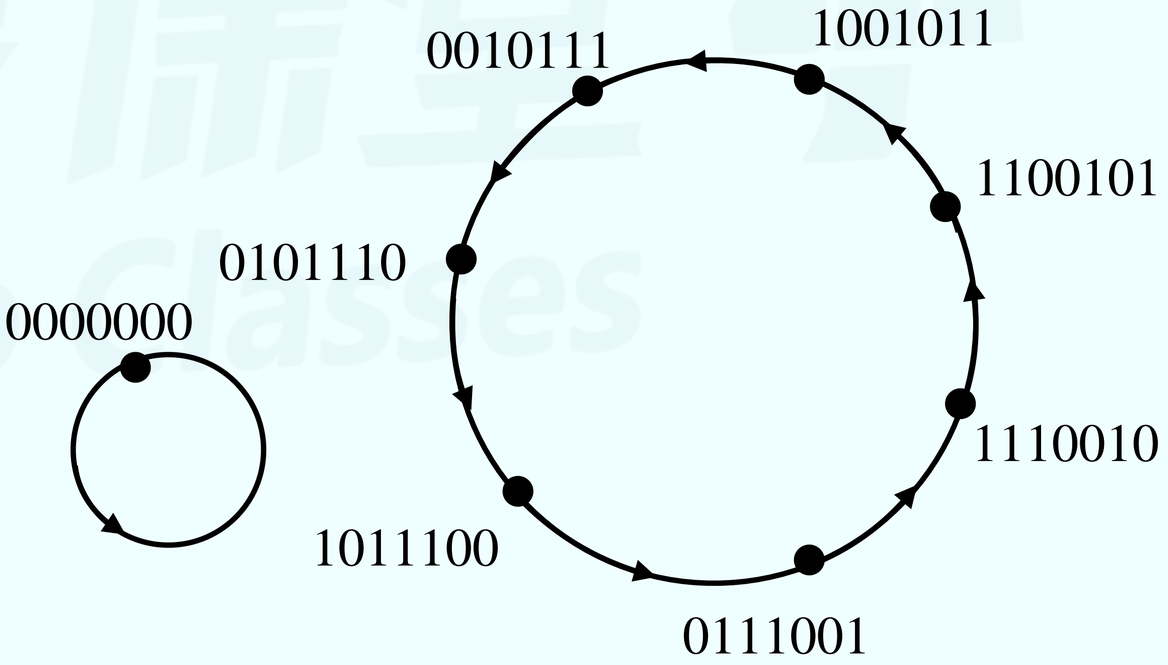
循环码有严密的代数学理论基础，检错和纠错能力较强，而且编码和解码设备都不太复杂。

循环码除了具有线性分组码的一般性质外，还具有循环性：**循环码中任一许用码组经过循环移位后，所得到的码组仍然是许用码组。**

循环码的特点

循环码除了具有线性分组码的一般性质外，还具有循环性：**循环码中任一许用码组经过循环移位后，所得到的码组仍然是许用码组。**





循环码多项式

设循环码的码字为 $C = [c_1\ c_2\ \cdots\ c_n]$，则定义该码字的码字多项式为

$$C(x) = c_1 x^{n-1} + c_2 x^{n-2} + \cdots + c_{n-1} x + c_n$$

设码字为1100101，则用码字多项式可以表示为

$$C(x) = 1 \times x^6 + 1 \times x^5 + 0 \times x^4 + 0 \times x^3 + 1 \times x^2 + 0 \times x^1 + 1$$

$$= x^6 + x^5 + x^2 + 1$$

对于二进制码，每个码多项式的系数不是0就是1；

x仅是码元符号位置的标记，并没有取值的意义。

按模取余运算的引入

循环码的循环移位特性可以用码多项式来证明。

我们在此引入模运算：

$1+1=2 \equiv 0 \pmod 2$　商1余0

$1+2=3 \equiv 1 \pmod 2$　商1余1

模运算是一个通过除法取余数的过程

模几就除以几

对于整数，有模n运算：若一个整数m可以表示为

$$\frac{m}{n}=Q+\frac{p}{n},p<n \quad \text{则} \ m \equiv p \pmod n$$

在模n运算下，若一个整数m等于其被n所除所得的余数。

按模取余运算的引入

在码多项式中也有类似的按模运算的规则。

若任意一个多项式 $F(x)$ 被一个 n 次多项式 $N(x)$ 除，得到商式 $Q(x)$ 与一个次数小于 n 的余式 $R(x)$，即：

$$\frac{F(x)}{N(x)} = Q(x) + \frac{R(x)}{N(x)}$$

则类似刚刚的定义方式可以定义多项式的模运算为 $F(x) \equiv R(x) \pmod{N(x)}$

例：x^3 被 x^3+1 除可得余式为1，则 $x^3 \equiv 1 \pmod{(x^3+1)}$

x^4+x^2+1 被 x^3+1 除可得余式为 x^2+x+1，则 $x^4+x^2+1 \equiv x^2+x+1 \pmod{(x^3+1)}$

$$\begin{array}{r} x \\ x^3+1 \;\overline{)\; x^4+x^2+1} \\ \underline{x^4+x} \\ x^2+x+1 \end{array}$$

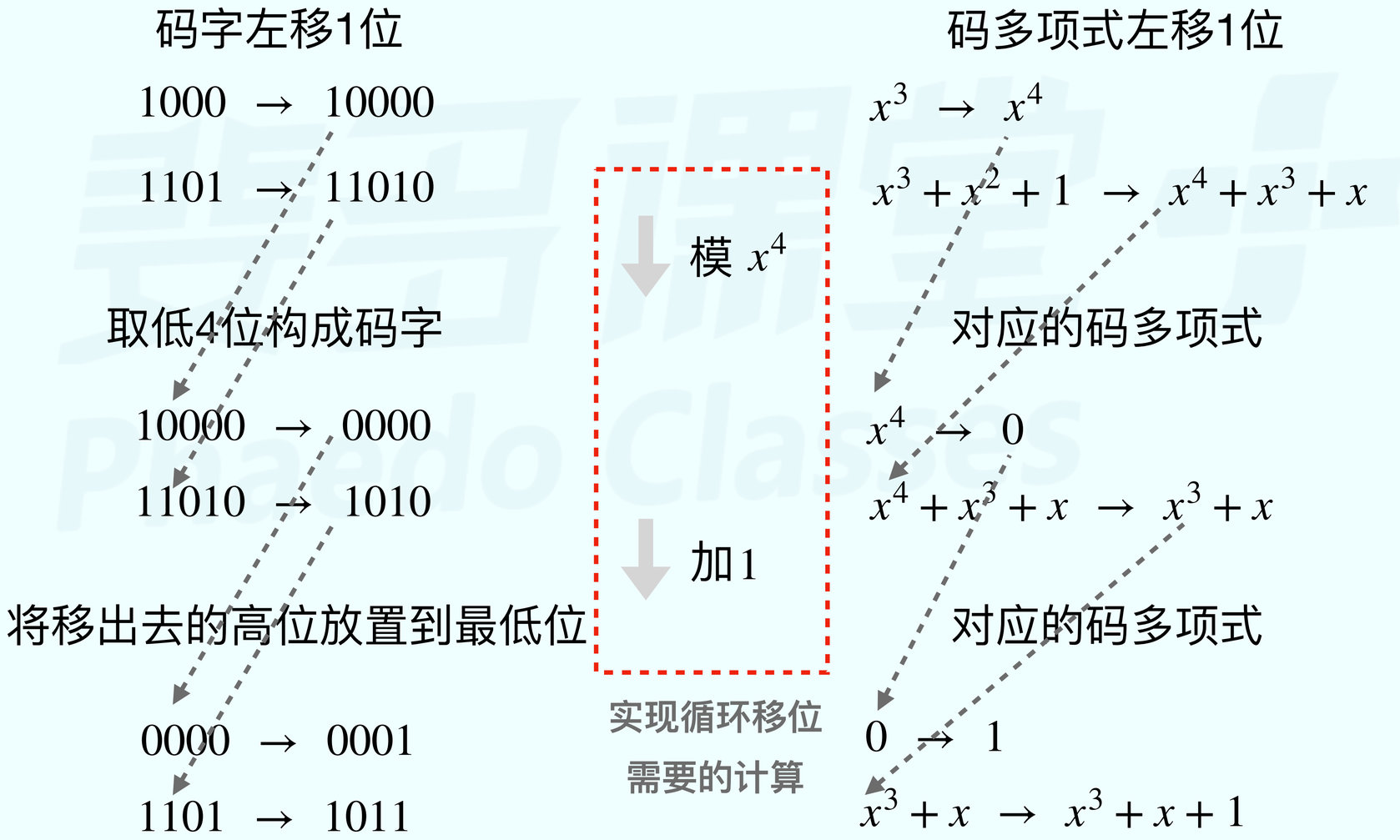
码多项式运算的含义

1. 移位运算：$x^iC(x)$，根据x次数与码元位置的关系可知，此运算将码字左移i位，右侧空位补零。
2. 加法运算：对于模2运算，**两个码多项式的相同指数项应该消去**，一般不考虑减法运算。
3. 乘法运算与除法运算：除法在列写竖式的过程中，对应位不是相减，而是模2加。

对于移位运算，我们举出如下例子：

码字左移1位	码多项式左移1位
$1000 \rightarrow 10000$	$x^3 \rightarrow x^4$
$1101 \rightarrow 11010$	$x^3+x^2+1 \rightarrow x^4+x^3+x$

循环移位的理解


码字左移1位
1000 → 10000
1101 → 11010
取低4位构成码字
10000 → 0000
11010 → 1010
将移出去的高位放置到最低位
0000 → 0001
1101 → 1011
模 x^4
加1
实现循环移位
需要的计算
码多项式左移1位
$x^3 \rightarrow x^4$
$x^3+x^2+1 \rightarrow x^4+x^3+x$
对应的码多项式
$x^4 \rightarrow 0$
$x^4+x^3+x \rightarrow x^3+x$
对应的码多项式
$0 \rightarrow 1$
$x^3+x \rightarrow x^3+x+1$

循环移位的运算意义

码字循环移位		对应的码多项式
$1000 \rightarrow 0001$	模 x^4+1	$x^3 \rightarrow 1$
$1101 \rightarrow 1011$		$x^3+x^2+1 \rightarrow x^3+x+1$

若$C(x)$是一个长为n的许用码字，则$x^iC(x)$在模(x^n+1)运算后依然是一个n长许用码字。

将码字左移i位，将移出去的高位循环移动到低位的0处右侧空位补零

循环移位！

例如：某循环码的一个码字为1100101，即$C(x)=x^6+x^5+x^2+1$

左移一位，码多项式为 $xC(x)=x^7+x^6+x^3+x$；

$x^7+x^6+x^3+x \mod (x^7+1)=x^6+x^3+x+1$ 对应循环移位的码字为1001011

循环码

小节1 循环码多项式与循环移位

小节2 循环码的生成多项式与生成矩阵

循环码的生成多项式与生成矩阵「引入」

关于循环码，我们有两个重要的结论：

- 在循环码中一个(n, k)码有2^k个不同的非零码组中，没有连续k位均为0的码组，即**连0的个数最多只能有k-1位**。

如果有连续k位均为0的码组，那么经若干次循环移位后，我们一定将得到一个信息位全为0，但监督位不全为0的码组，根据校验位和信息位的线性约束关系，这样的码字在线性码中不可能存在。

- 在码字中，最低阶次的多项式$g(x)$，必须是一个常数项不为0的$(n\text{-}k)$次 多项式，而且只有一个。

因为如果有两个常数项不为0的多项式相加。将会出现连续0的个数超过k-1，违反了封闭性原则。

循环码的生成多项式与生成矩阵

从(n,k)循环码中取出一个**前面k-1位都是0的码字**，定义这个码字的码多项式为**生成多项式$g(x)$**。

那么该多项式的次数为 $m = n - k$，即校验码元的位数。

为了保证构成生成矩阵中每行线性不相关，那么通常用生成多项式$g(x)$构造生成矩阵G。

生成矩阵G的构造方式为 $G = \begin{bmatrix} x^{k-1}g(x) \\ x^{k-2}g(x) \\ \cdots \\ xg(x) \\ g(x) \end{bmatrix}$ 矩阵每一行都是一个码字，行与行之间为循环移位关系。

循环码中合法的码多项式必定是生成多项式的倍式，即 $C(x) = V(x)g(x)$

利用该性质可以进行是否是合法循环码码字的校验。

但利用上述方法构造的生成矩阵不是系统码的生成矩阵，即它为非典型生成矩阵。

如果给定生成多项式且用此方法得到非典型生成矩阵，我们可以对矩阵进行**初等行变换**得到典型阵。

例题8-4 已知$(7,4)$系统循环码的生成多项式为$g(x)=x^3+x^2+1$，试构造该循环码的生成矩阵。

解析8-4 根据生成矩阵的构造方法，且n=7，k=4，可得

$$G=\begin{bmatrix} x^{k-1}g(x) \\ x^{k-2}g(x) \\ \cdots \\ xg(x) \\ g(x) \end{bmatrix}=\begin{bmatrix} x^3g(x) \\ x^2g(x) \\ xg(x) \\ g(x) \end{bmatrix}=\begin{bmatrix} x^6+x^5+x^3 \\ x^5+x^4+x^2 \\ x^4+x^3+x \\ x^3+x^2+1 \end{bmatrix}=\begin{bmatrix} 1&1&0&1&0&0&0 \\ 0&1&1&0&1&0&0 \\ 0&0&1&1&0&1&0 \\ 0&0&0&1&1&0&1 \end{bmatrix}$$

所得矩阵为非系统生成阵，我们对其进行初等行变换即可得到生成阵

$$G=\begin{bmatrix} 1&1&0&1&0&0&0 \\ 0&1&1&0&1&0&0 \\ 0&0&1&1&0&1&0 \\ 0&0&0&1&1&0&1 \end{bmatrix}\rightarrow\begin{bmatrix} 1&0&0&0&1&1&0 \\ 0&1&0&0&0&1&1 \\ 0&0&1&0&1&1&1 \\ 0&0&0&1&1&0&1 \end{bmatrix}$$

注意：给定生成多项式就给定了码元之间的约束关系，此时不能做列变换！！！

循环码的生成多项式与生成矩阵

下面我们介绍一种构造**典型生成矩阵**的方法。

对于系统码，其码字必然符合多项式

$$C(x) = m(x)x^{n-k} + r(x)$$

$$= \underbrace{m_1 x^{n-1} + \cdots + m_k x^{n-k}}_{\text{信息位}} + \underbrace{r_1 x^{n-k-1} + \cdots + r_{n-k}}_{\text{校验位}}$$

因此，校验位多项式满足 $r(x) = C(x) + m(x)x^{n-k} = V(x)g(x) + m(x)x^{n-k}$

则有 $\dfrac{r(x)}{g(x)} = V(x) + \dfrac{m(x)x^{n-k}}{g(x)}$，因此 $r(x) = m(x)x^{n-k} \mod g(x)$ 。

因此，我们已知信息位和生成多项式就可以求出校验位。

循环码的生成多项式与生成矩阵

$$r(x) = m(x)x^{n-k} \mod g(x)$$

- **根据此原理我们就可以得到一个关于循环码的编码方法：**

第一步：用 x^{n-k} 乘以信息位多项式 $m(x)$ （实际上是把信息码元后面填 $n-k$ 个0）

第二步：用 $x^{n-k}m(x)$除以生成多项式 $g(x)$，所得到的余式即为校验位对应的多项式 $r(x)$。

- **我们给出系统循环码生成多项式的一般表示形式如下，那么就可以因此确定典型生成矩阵：**

$$G = \begin{bmatrix} x^{n-1} + r_1(x) \\ x^{n-2} + r_2(x) \\ \cdots \\ x^{n-k} + r_k(x) \end{bmatrix}$$

信息位对应的单位矩阵

其中，
$$\begin{cases} r_1(x) = x^{n-1} \mod g(x) \\ r_2(x) = x^{n-2} \mod g(x) \\ \cdots\cdots \\ r_k(x) = x^{n-k} \mod g(x) \end{cases}$$

例题8-5 已知一循环码的生成多项式为 $g(x)=x^4+x^2+x+1$，输入信息位110，求码多项式，求典型生成矩阵和校验矩阵。

解析8-5 根据生成多项式的最高次数为n-k=4，且k=3，可得n=7，所求循环码为(7, 3)循环码

给定信息位110，即 $m(x)=x^2+x$，可得 $x^{n-k}m(x)=x^4m(x)=x^4(x^2+x)=x^6+x^5$

$$\frac{x^{n-k}m(x)}{g(x)}=\frac{x^6+x^5}{x^4+x^2+x+1}=(x^2+x+1)+\frac{x^2+1}{x^4+x^2+x+1}$$

因此监督位的码多项式 $r(x)=x^2+1$，对应0101，故编码结果为1100101，码多项式为$C(x)=x^6+x^5+x^2+1$

求典型生成矩阵：

$$r_1(x)=x^6 \mod g(x)=x^2+1$$
$$r_2(x)=x^5 \mod g(x)=x$$
$$r_3(x)=x^4 \mod g(x)=1$$

故所求典型生成矩阵：$G=\begin{bmatrix}x^6+r_1(x)\\x^5+r_2(x)\\x^4+r_3(x)\end{bmatrix}=\begin{bmatrix}1&0&0&0&1&0&1\\0&1&0&0&0&1&0\\0&0&1&0&0&0&1\end{bmatrix}$

$$G=[I\ P]=[I\ Q^T] \quad \Leftrightarrow \quad H=[Q\ I]=[P^T\ I]=\begin{bmatrix}0&0&0&1&0&0&0\\1&0&0&0&1&0&0\\0&1&0&0&0&1&0\\1&0&1&0&0&0&1\end{bmatrix}$$

循环码的生成多项式与生成矩阵

两种构造方法

一是求解校验多项式，即用除法求余，然后得到生成多项式的表示，并得到典型（系统）生成矩阵。

$$G=\begin{bmatrix} x^{n-1}+r_1(x) \\ x^{n-2}+r_2(x) \\ \cdots \\ x^{n-k}+r_k(x) \end{bmatrix}$$

二是直接根据生成多项式和循环码的定义，得到非系统码的生成矩阵，然后对该矩阵进行初等变换，变成系统码的生成矩阵。

$$G=\begin{bmatrix} x^{k-1}g(x) \\ x^{k-2}g(x) \\ \cdots \\ xg(x) \\ g(x) \end{bmatrix}$$