3.5 信道容量C

信息传输系统的衡量指标：

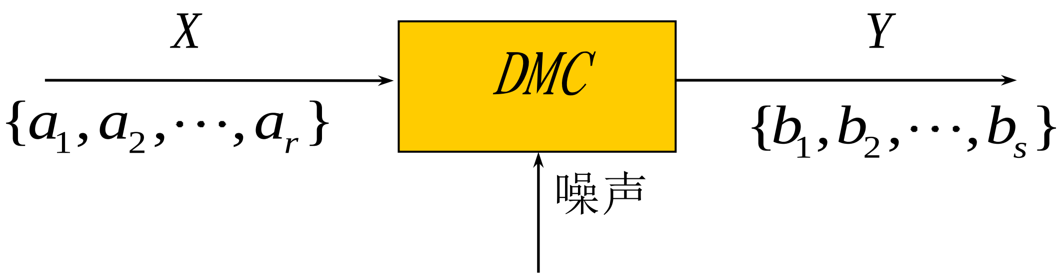
（1）数量（速度）指标：信息（传输）率R，即信道中平均每个符号传递的信息量；

（2）质量指标：平均差错率P_e，即对信道输出符号进行译码的平均错误概率。

愿望：信道传输信息时速度快、错误少，即R尽量大而P_e尽量小。

信息率能大到什么程度？这就是信道容量问题。

3.5.1 信道容量的定义



信息率R，就是信道的平均互信息量：

$$R = I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \quad \text{bit/符号}$$

信息传输速率R_t：
$$R_t = \frac{1}{t} I(X;Y) \quad \text{bit/秒}$$

信道容量C：信道的最大的信息率，即

$$C = \max_{P_X} R = \max_{P_X} I(X;Y) \quad \text{bit/符号}$$

信道容量也可以定义为信道的最大的信息速率，记为C_t：

$$C_t = \max_{P_X} R_t = \max_{P_X} \left\{ \frac{1}{t} I(X;Y) \right\} \quad \text{bit/秒}$$

关于信道容量的几点注释

$$C = \max_{P_X} R = \max_{P_X} I(X;Y) \quad \text{bit/符号}$$

(1) 信道容量C 是信道信息率的上限，定量描述了信道（信息的）最大通过能力；

(2) 使得给定信道的$I(X;Y)$ 达到最大值(即信道容量)的输入概率分布，称为最佳输入(概率)分布，记为 P_X^*；

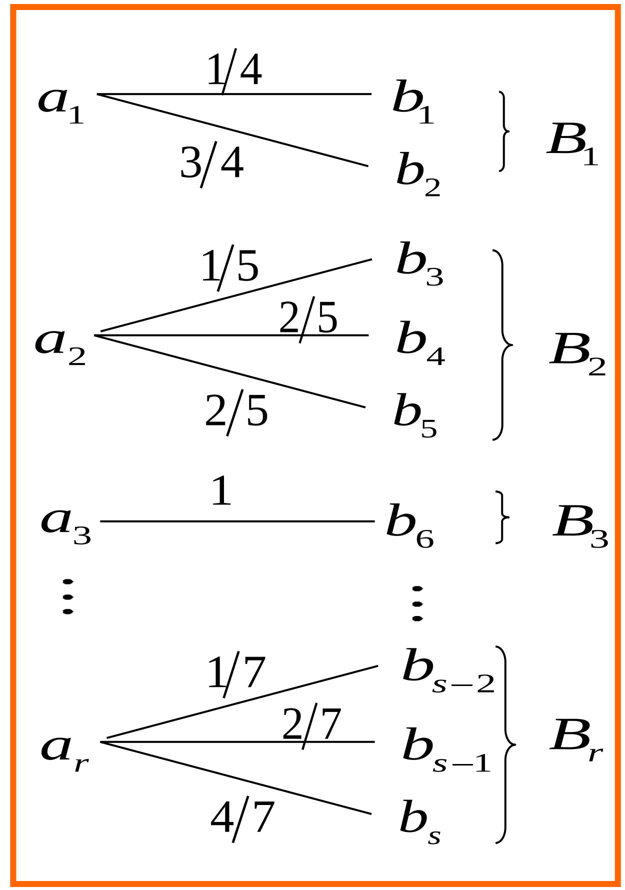
(3) 信道的$I(X;Y)$ 与输入概率分布P_X和转移概率分布$P_{Y|X}$两者有关，但信道容量是信道的固有参数，只与信道转移概率$P_{Y|X}$有关。

3.5.2 离散无噪信道的信道容量

无噪信道：无损信道、确定信道以及无损确定信道的统称。

（1）无损信道：损失熵为零的信道。

$$r=3\text{ 时，}\quad [P_{Y|X}]=\begin{bmatrix}1/4 & 3/4 & 0 & 0 & 0 & 0\\ 0 & 0 & 1/5 & 2/5 & 2/5 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}$$



转移矩阵特征：每列只有一个非零元素。

令 $[P_X]=\begin{bmatrix}p_1 & p_2 & p_3\end{bmatrix}$

则

$$[P_{XY}]=\begin{bmatrix}p_1/4 & 3p_1/4 & 0 & 0 & 0 & 0\\ 0 & 0 & p_2/5 & 2p_2/5 & 2p_2/5 & 0\\ 0 & 0 & 0 & 0 & 0 & p_3\end{bmatrix}$$

$$[P_{XY}]=\begin{bmatrix} p_1/4 & 3p_1/4 & 0 & 0 & 0 & 0 \\ 0 & 0 & p_2/5 & 2p_2/5 & 2p_2/5 & 0 \\ 0 & 0 & 0 & 0 & 0 & p_3 \end{bmatrix}$$

$$\Downarrow$$

$$[P_Y]=\begin{bmatrix} \frac{1}{4}p_1 & \frac{3}{4}p_1 & \frac{1}{5}p_2 & \frac{2}{5}p_2 & \frac{2}{5}p_2 & p_3 \end{bmatrix}$$

$$\Downarrow$$

$$[P_{X|Y}]=\begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

后验概率矩阵特征：每列只有一个元素为1，其余均为0，各列后验概率均组成确定性概率分布。

$$[P_{X|Y}]=\begin{bmatrix}1 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 1 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}$$

后验概率矩阵特征：每列只有一个元素为1，其余均为0，各列后验概率均组成确定性概率分布。

$$H(X|Y=b_j)=H[P(a_1|b_j),P(a_2|b_j),\cdots,P(a_r|b_j)]=0\quad;\quad j=1,2,\cdots,s$$

损失熵： $H(X|Y)=\sum_{j=1}^{s}P(b_j)H(X|Y=b_j)=0$

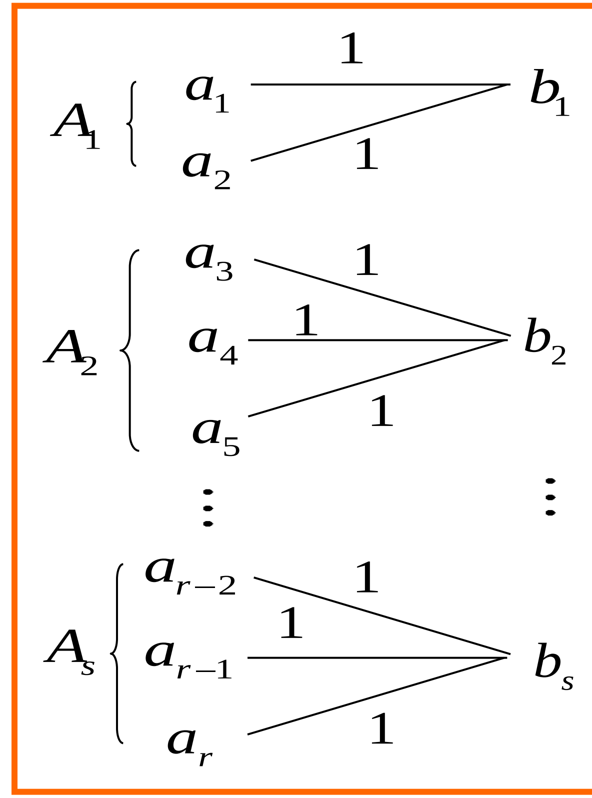
无损信道的信道容量： $C=\max_{P_X}I(X;Y)=\max_{P_X}\{H(X)-H(X|Y)\}=\max_{P_X}H(X)$

$$C=H(X)|_{P(a_i)=1/r}=\log r$$

(2) 确定信道：噪声熵为0的信道

当s=2时，

$$[P_{Y|X}]=\begin{bmatrix}1 & 0\\ 1 & 0\\ 0 & 1\\ 0 & 1\\ 0 & 1\end{bmatrix}$$



$$H(Y\mid X)=0$$

$$I(X;Y)=H(Y)-H(Y\mid X)=H(Y)$$

$$C=\max_{P_X} I(X;Y)=\max_{P_X} H(Y)$$

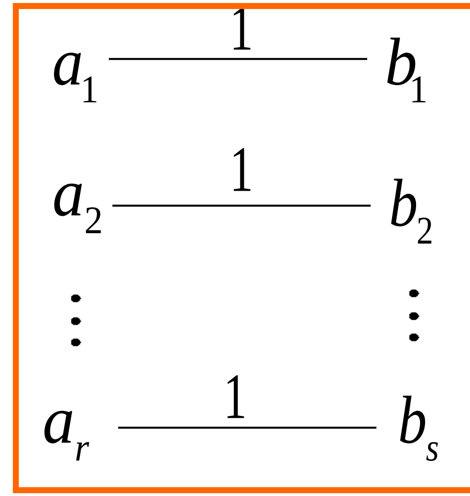
至少能够找到一种输入分布使输出的取值符号达到等概率分布，甚至有多个或无穷多个输入分布满足条件。

$$C=\max_{P_X} I(X;Y)=\max_{P_X} H(Y)=H(Y)\Big|_{P(b_j)=1/s}=\log s$$

(3)无损确定信道：损失熵和噪声熵均为0的信道

转移概率矩阵为单位阵：

$$[P_{Y|X}]=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$



令 $[P_X]=\begin{bmatrix}p_1 & p_2 & \cdots & p_r\end{bmatrix}$ 则

$$[P_{XY}]=\begin{bmatrix}p_1 & 0 & \cdots & 0\\ 0 & p_2 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & p_r\end{bmatrix}$$

$$[P_Y]=\begin{bmatrix}p_1 & p_2 & \cdots & p_r\end{bmatrix}$$

$$[P_{X|Y}]=\begin{bmatrix}1 & 0 & \cdots & 0\\ 0 & 1 & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & 1\end{bmatrix}$$

$$H(Y|X)=0$$

$$H(X|Y)=0$$

$$I(X;Y)=H(X)=H(Y)$$

$$C=\max_{P_X} I(X;Y)=\max_{P_X} H(X)=H(X)\big|_{P(a_i)=1/r}=\log r$$

3.5.3 离散对称信道

（1）几种对称信道的定义

定义1：信道r×s转移矩阵$[P_{Y|X}]$每一行s个元素，都由同一组元素 $\{p_1', p_2', \cdots, p_s'\}$的不同排列组成，则称为行排列阵，此类信道称为离散输入对称信道。

例：$$[P_{Y|X}] = \begin{bmatrix} 0.1 & 0.2 & 0.3 & 0.4 \\ 0.2 & 0.3 & 0.4 & 0.1 \end{bmatrix}$$

例： $[P_{Y|X}]=\begin{bmatrix}0.1 & 0.2 & 0.3 & 0.4\\ 0.2 & 0.3 & 0.4 & 0.1\end{bmatrix}$

$$H(Y|a_i)=H(p_1',p_2',\cdots,p_s') \qquad i=1,2,\cdots,r$$

$$H(Y|X)=\sum_{i=1}^{r}P(a_i)H(Y|a_i)=H(Y|a_i)\sum_{i=1}^{r}P(a_i)=H(p_1',p_2',\cdots,p_s')$$

$$\begin{aligned}C&=\max_{P_X} I(X;Y)\\&=\max_{P_X}\{H(Y)-H(Y|X)\}\\&=\max_{P_X}\{H(Y)\}-H(p_1',p_2',\cdots,p_s')\end{aligned}$$

将平均互信息量的最大值问题转换成输出熵的最大值问题。

定义2：信道r×s转移矩阵$[P_{Y|X}]$每一列r个元素，都由同一组元素 $\{q_1', q_2', \cdots, q_r'\}$的不同排列组成，则称为列排列阵，此类信道称为离散输出对称信道。

例： $$[P_{Y|X}] = \begin{bmatrix} 0.4 & 0.6 \\ 0.6 & 0.4 \\ 0.5 & 0.5 \end{bmatrix}$$

定义3：若信道转移矩阵$[P_{Y|X}]$既是行排列阵又是列排列阵，则称为行列排列阵，此类信道称为离散对称信道。

例： $$[P_{Y|X}] = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3} \end{bmatrix}$$

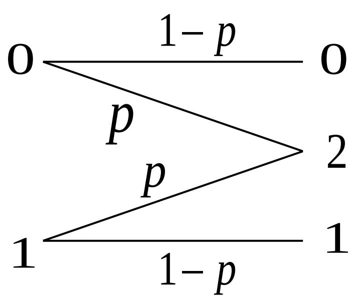
定义4：若信道转移矩阵$[P_{Y|X}]$的列可被划分成若干个互不相交的子集，且每个子集所组成的子阵是行列排列阵，则称此类信道称为离散准对称信道。

例：

$$[P_{Y|X}] = \begin{bmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6} \\ \frac{1}{3} & \frac{1}{6} & \frac{1}{3} & \frac{1}{6} \end{bmatrix}$$

- 要判断一个信道是否为离散准对称信道，必须对该信道的转移矩阵进行适当的调整，即按列重排再按列分块；
- 对转移矩阵调整的过程，就是定义中所说的将转移矩阵的列划分成子集再组成子阵的过程；
- 转移矩阵的列与输出符号对应，把列划分成互不相交的子集，相当于把信道的输出符号划分成互不相交的子集。

*BEC*的转移矩阵的重排和分块过程



（2）离散对称信道的信道容量

引理：离散对称信道输入等概率分布时，输出也等概率分布。

定理（对称*DMC*的信道容量）：对于对称*DMC*，当输入等概时达到信道容量，且

$$C = \log s - H(p_1', p_2', \cdots, p_s') \qquad \text{bit/符号}$$

其中，s ：信道输出符号个数或转移矩阵的列数。

$\{p_1', p_2', \cdots, p_s'\}$ ：转移矩阵任一行的s个转移概率。

例1：

$$[P_{Y|X}]=\begin{bmatrix}\frac{1}{3} & \frac{1}{3} & \frac{1}{6} & \frac{1}{6}\\ \frac{1}{6} & \frac{1}{6} & \frac{1}{3} & \frac{1}{3}\end{bmatrix}$$

解：

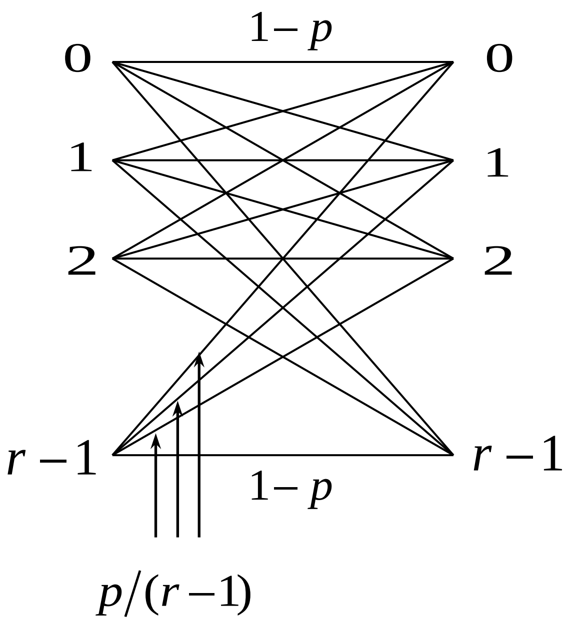
$$C=\max_{P_X}\{H(Y)\}-H(p_1',p_2',\cdots,p_s')$$

$$\xrightarrow{\text{引理}}\quad C=\log s-H(p_1',p_2',\cdots,p_s')$$

$$C=\log 4-H(\frac{1}{3},\frac{1}{3},\frac{1}{6},\frac{1}{6})$$

$$[P_X^*]=\begin{bmatrix}\frac{1}{2} & \frac{1}{2}\end{bmatrix}$$

例2：r进制均匀信道的信道容量



$$[P_{Y|X}]=\begin{bmatrix} \bar{p} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \frac{p}{r-1} & \bar{p} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{p}{r-1} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \bar{p} \end{bmatrix}$$

$$p+\bar{p}=1$$

r=2时，退化成为*BSC* 。

均匀信道的转移矩阵是一个对称方阵，也是行列排列阵，因此是对称信道。

均匀信道的转移矩阵是一个对称方阵，也是行列排列阵，因此是对称信道。由定理可知，其最佳输入分布为等概率分布：

$$[P_{Y|X}]=\begin{bmatrix} \bar{p} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \frac{p}{r-1} & \bar{p} & \frac{p}{r-1} & \cdots & \frac{p}{r-1} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{p}{r-1} & \frac{p}{r-1} & \frac{p}{r-1} & \cdots & \bar{p} \end{bmatrix}$$

$$P^*(a_i)=1/r \qquad i=1,2,\cdots,r$$

信道容量：
$$C=\log s-H(p_1',p_2',\cdots,p_s')$$
$$=\log r-H(\bar{p},\frac{p}{r-1},\cdots,\frac{p}{r-1})$$
$$=\log r+p\log\frac{p}{r-1}+(1-p)\log(1-p) \qquad \text{bit/符号}$$

r=2时，*BSC*的信道容量：
$$C_{BSC}=1+p\log(p)+(1-p)\log(1-p)\ =1-h_2(p)$$

bit/符号

（3）离散准对称信道的信道容量

定理（准对称*DMC*的信道容量）：对于准对称*DMC*，当输入等概时达到信道容量。

如果把准对称*DMC*的$[P_{Y|X}]$分块成 n 个行列排列子阵$\{Q_1, \ldots, Q_n\}$，再根据以上定理，可得准对称*DMC*的信道容量计算公式：

$$C = -\sum_{k=1}^{n} s_k \left(\frac{M_k}{r}\right) \log\left(\frac{M_k}{r}\right) - H(p_1', p_2', \cdots, p_s')$$

其中， r：转移矩阵$[P_{Y|X}]$的行数；

s_k：子阵Q_k的列数；

M_k：子阵Q_k的任一列元素之和；

$\{p_1', p_2', \cdots, p_s'\}$：转移矩阵$[P_{Y|X}]$任一行元素。

$$C=-\sum_{k=1}^{n}s_k\left(\frac{M_k}{r}\right)\log\left(\frac{M_k}{r}\right)-H(p_1',p_2',\cdots,p_s')$$

或 $$\boxed{C=\log r-H(p_1',p_2',\cdots,p_s')-\sum_{k=1}^{r}N_k\log M_k}$$

其中， r：转移矩阵$[P_{Y|X}]$的行数；

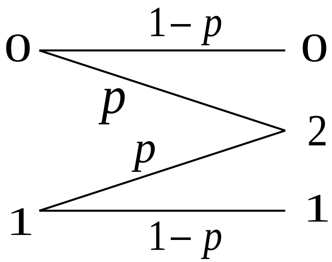
s_k：子阵Q_k的列数；

M_k：子阵Q_k的任一列元素之和；

N_k：子阵Q_k的任一行元素之和；

$\{p_1',p_2',\cdots,p_s'\}$：转移矩阵$[P_{Y|X}]$任一行元素。

例：求2进制删除信道(BEC)的信道容量



$X \backslash Y$	0	2	1
0	$1-p$	p	0
1	0	p	$1-p$

$X \backslash Y$	0	1	2
0	$1-p$	0	p
1	0	$1-p$	p
	Q_1		Q_2

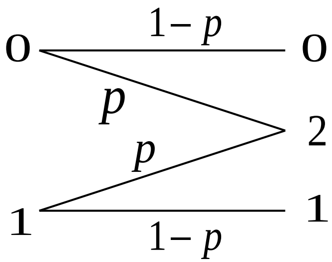
解1：$\begin{cases} r=2 \\ \{p_1', p_2', \cdots, p_s'\}:\{1-p, p, 0\} \end{cases}$ $\begin{cases} s_1=2 \\ M_1=1-p \end{cases}$ $\begin{cases} s_2=1 \\ M_2=2p \end{cases}$

$$C=-\sum_{k=1}^{n} s_k\left(\frac{M_k}{r}\right)\log\left(\frac{M_k}{r}\right)-H(p_1', p_2', \cdots, p_s')$$

$$C=-2\times\left(\frac{1-p}{2}\right)\log\left(\frac{1-p}{2}\right)-1\times\left(\frac{2p}{2}\right)\log\left(\frac{2p}{2}\right)-H(1-p, p, 0)=1-p \text{ bit/符号}$$

最佳输入分布为等概率分布： $P_X^*=\{1/2, 1/2\}$

例：求2进制删除信道(BEC)的信道容量



$X \backslash Y$	0	2	1
0	$1-p$	p	0
1	0	p	$1-p$

$X \backslash Y$	0	1	2
0	$1-p$	0	p
1	0	$1-p$	p
	Q_1		Q_2

解2：$\begin{cases} r = 2 \\ \{p_1', p_2', \cdots, p_s'\} : \{1-p, p, 0\} \end{cases}$ $\begin{cases} M_1 = 1-p \\ N_1 = 1-p \end{cases}$ $\begin{cases} M_2 = 2p \\ N_2 = p \end{cases}$

$$C = \log r - H(p_1', p_2', \cdots, p_s') - \sum_{k=1}^{n} N_k \log M_k$$

$$C = 1 - H(1-p, p, 0) - (1-p)\log(1-p) - p\log 2p = 1-p \quad \text{bit/符号}$$

最佳输入分布为等概率分布：$P_X^* = \{1/2, 1/2\}$

3.5.4 一般DMC信道容量解的充要条件

定理(一般*DMC*信道容量解的充要条件)：一般*DMC* $\{X, P_{Y|X}, Y\}$ 其平均互信息量 $I(X;Y)$ 在输入分布为 $P_X^* = \{P^*(a_1), P^*(a_2), \cdots, P^*(a_r)\}$ 时取最大值的充要条件是

$$I(a_i;Y)\big|_{P_X=P_X^*} = C \quad \text{当 } P^*(a_i) > 0 \text{ 时}$$

$$I(a_i;Y)\big|_{P_X=P_X^*} \le C \quad \text{当 } P^*(a_i) = 0 \text{ 时}$$

式中：

$$I(a_i;Y) = \sum_{j=1}^{s} P(b_j \mid a_i) \log \frac{P(b_j \mid a_i)}{P(b_j)}$$

$$\begin{aligned} I(X;Y) &= \sum_{i=1}^{r} \sum_{j=1}^{s} P(a_i) P(b_j \mid a_i) \log \frac{P(b_j \mid a_i)}{P(b_j)} \\ &= \sum_{i=1}^{r} P(a_i) \sum_{j=1}^{s} P(b_j \mid a_i) \log \frac{P(b_j \mid a_i)}{P(b_j)} \\ &= \sum_{i=1}^{r} P(a_i) I(a_i;Y) \end{aligned}$$

例：设信道转移矩阵如下，求新的容量和最佳输入分布。

$$[P_{Y|X}]=\begin{bmatrix}1 & 0 & 0\\ 0 & 1-\delta & \delta\\ 0 & \delta & 1-\delta\end{bmatrix}$$

解：

$$I(a_i;Y)=\sum_{j=1}^{s}P(b_j|a_i)\log\frac{P(b_j|a_i)}{P(b_j)}=C$$

$$\begin{cases}I(a_1;Y)=\log\dfrac{1}{P(b_1)}=C\\ I(a_2;Y)=(1-\delta)\log\dfrac{1-\delta}{P(b_2)}+\delta\log\dfrac{\delta}{P(b_3)}=C\\ I(a_3;Y)=\delta\log\dfrac{\delta}{P(b_2)}+(1-\delta)\log\dfrac{1-\delta}{P(b_3)}=C\\ P(b_1)+P(b_2)+P(b_3)=1\end{cases}\Rightarrow\begin{cases}C=\log(1+2^{[1-h_2(\delta)]})=\log[1+2\delta^{\delta}(1-\delta)^{(1-\delta)}]\\ P(b_1)=2^{-C}=\dfrac{1}{1+2\delta^{\delta}(1-\delta)^{(1-\delta)}}\\ P(b_2)=P(b_3)=2^{-[C+h_2(\delta)]}=\dfrac{\delta^{\delta}(1-\delta)^{(1-\delta)}}{1+2\delta^{\delta}(1-\delta)^{(1-\delta)}}\end{cases}$$

$$[P_{Y|X}]=\begin{bmatrix}1 & 0 & 0\\ 0 & 1-\delta & \delta\\ 0 & \delta & 1-\delta\end{bmatrix}$$

设最佳输入分布为 $P_X^*=\{P^*(a_1),P^*(a_2),P^*(a_3)\}$

则

$$\begin{cases}P(b_1)=P^*(a_1)\\ P(b_2)=(1-\delta)P^*(a_2)+\delta P^*(a_3)\\ P(b_3)=\delta P^*(a_2)+(1-\delta)P^*(a_3)\end{cases}$$

$$\Downarrow$$

$$\begin{cases}P^*(a_1)=\dfrac{1}{1+2\delta^{\delta}(1-\delta)^{(1-\delta)}}\\ P^*(a_2)=P^*(a_3)=\dfrac{\delta^{\delta}(1-\delta)^{(1-\delta)}}{1+2\delta^{\delta}(1-\delta)^{(1-\delta)}}\end{cases}$$

3.5.5 信道容量的迭代算法

算法：

1.初始化信源分布 $p^{(0)}=(p_1,p_2,\cdots,p_i,\cdots,p_r)$（一般初始化为均匀分布），置迭代计数器k=0,设信道容量相对误差门限为 $\delta,\delta>0$；

2. $$\varphi_{ji}^{(k)}=\frac{p_{ij}p_i^{(k)}}{\sum_i p_{ij}p_i^{(k)}}$$

3. $$p_i^{(k+1)}=\frac{\exp\left[\sum_j p_{ij}\ln\varphi_{ji}^{(k)}\right]}{\sum_i\left\{\exp\left[\sum_j p_{ij}\ln\varphi_{ji}^{(k)}\right]\right\}}$$

4. $C^{(k+1)} = \ln \sum_i \left\{ \exp\left[\sum_j p_{ij} \ln \varphi_{ji}^{(k)} \right] \right\}$

5.如果$\Delta C^{(k+1)} = \left| \dfrac{C^{(k+1)} - C^{(k)}}{C^{(k+1)}} \right| \le \delta$,转向7

6.置迭代序号$k+1=k$,转向2

7.输出$p_i^{(k+1)}$和$C^{(k+1)}$的结果

8.停止

习题

• 信道转移矩阵如下：

(1)
$$[P_{Y|X}]=\begin{bmatrix}1 & 0 & 0\\ 0 & 0 & 1\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}$$

(2)
$$[P_{Y|X}]=\begin{bmatrix}1/2 & 1/3 & 1/6\\ 1/6 & 1/2 & 1/3\\ 1/3 & 1/6 & 1/2\end{bmatrix}$$

(3)
$$[P_{Y|X}]=\begin{bmatrix}1/4 & 3/4 & 0 & 0 & 0 & 0\\ 0 & 0 & 1/5 & 2/5 & 2/5 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}$$

(4)
$$[P_{Y|X}]=\begin{bmatrix}0.7 & 0.1 & 0.2\\ 0.2 & 0.1 & 0.7\end{bmatrix}$$

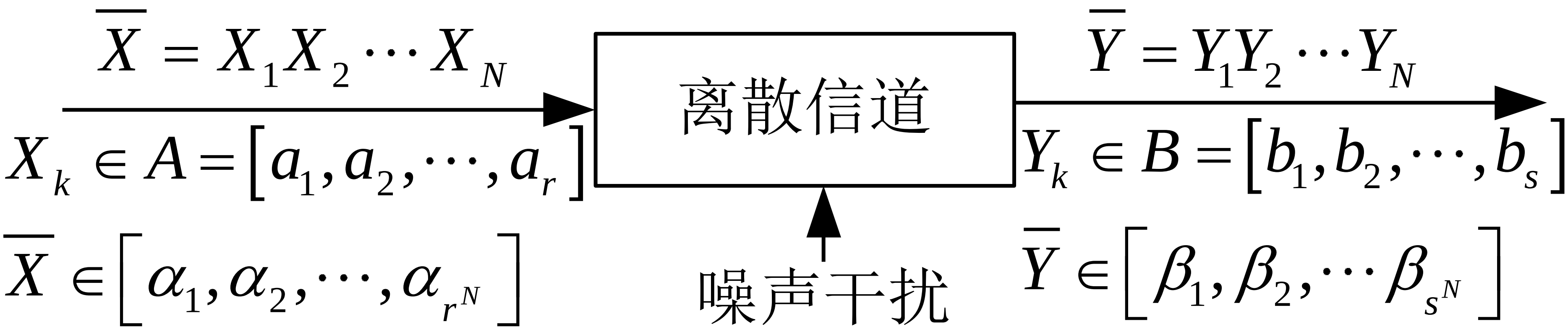
(5)
$$[P_{Y|X}]=\begin{bmatrix}1 & 0 & 0\\ 0 & 3/4 & 1/4\\ 0 & 1/4 & 3/4\end{bmatrix}$$

分别判断各信道属何种信道（无损、确定、无损确定、对称、准对称、一般），求出各信道的最佳输入分布和信道容量。

（注：$\log 3=1.585$，$\log 5=2.322$ ，$\log 6=2.585$ ）

3.6 扩展信道及其信道容量

3.6.1 扩展信道的数学模型



<u>N次扩展信道</u>的模型↑

把输入$\overline{X}$（也记为X^N)和输出$\overline{Y}$(也记为Y^N）

都分别当作一个新的随机变量——

联合随机变量，它们的取值集合分别为A^N和B^N:

$$\overline{\mathrm{X}} \in A^{N} = \left\{\alpha_1, \alpha_2, \cdots, \alpha_{r^N}\right\}$$

$$\alpha_h = \left(a_{h1}, a_{h2}, \cdots a_{hN}\right) \quad a_{hi} \in A = \left\{a_1, a_2, \cdots, a_r\right\}$$

$$\overline{Y} \in B^{N} = \left\{\beta_1, \beta_2, \cdots, \beta_{s^N}\right\}$$

$$\beta_l = \left(b_{l1}, b_{l2}, \cdots, b_{lN}\right) \quad b_{lj} \in B = \left\{b_1, b_2, \cdots b_s\right\}$$

转移概率集合为

$$P_{\overline{Y}|\overline{X}} = \left\{P(\beta_l \mid \alpha_h) \mid h = 1, 2, \cdots, r^N; l = 1, 2, \cdots, s^N\right\}$$

数学模型可记为： $\left\{\overline{X}, P_{\overline{Y}|\overline{X}}, \overline{Y}\right\}$

信道是DMC的充要条件：

$$P(\beta_l \mid \alpha_h) = P(b_{l1}b_{l2}\cdots b_{lN} \mid a_{h1}a_{h2}\cdots a_{hN})$$

$$= \prod_{k=1}^{N} P(b_{lk} \mid a_{hk})$$ 对于任意N均成立。

例3.11 求BSC的2次扩展信道数学模型

解：单符号BSC的输入和输出符号集分别为：

$\boldsymbol{A=\{a_1,a_2\}}$； $\boldsymbol{B=\{b_1,b_2\}}$

2次扩展信道的输入和输出符号集分别为

$A^2=\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}=\{a_1a_1, a_1a_2, a_2a_1, a_2a_2\}$

$B^2=\{\beta_1, \beta_2, \beta_3, \beta_4\}=\{b_1b_1, b_1b_2, b_2b_1, b_2b_2\}$

计算转移概率，

$$P(\beta_1 | \alpha_1) = P(b_1b_1 | a_1a_1)$$

$$= P(b_1 | a_1)P(b_1 | a_1) = \bar{p}^2$$

$$P(\beta_2 | \alpha_1) = P(b_1b_2 | a_1a_1)$$

$$= P(b_1 | a_1)P(b_2 | a_1) = \bar{p}p$$

$$P(\beta_4 \mid \alpha_1) = P(b_2 b_2 \mid a_1 a_1)$$
$$= P(b_2 \mid a_1) P(b_2 \mid a_1) = p^2$$
$$\vdots \qquad \vdots \qquad \vdots$$

可以得到2次扩展信道的转移概率矩阵：

$$\left[P_{\overline{Y}|\overline{X}} \right] = \begin{bmatrix} \overline{p}^2 & \overline{p}p & \overline{p}p & p^2 \\ \overline{p}p & \overline{p}^2 & p^2 & \overline{p}p \\ \overline{p}p & p^2 & \overline{p}^2 & \overline{p}p \\ p^2 & \overline{p}p & \overline{p}p & \overline{p}^2 \end{bmatrix}$$

3.6.2 扩展信道的平均互信息量和信道容量

一、基本的表达式

1、扩展信道的平均互信息量

$$I(\overline{X};\overline{Y}) = I(X^N, Y^N)$$

$$= \sum_{h=1}^{N}\sum_{l=1}^{N} P(\alpha_h, \beta_l)\log\frac{P(\alpha_h, \beta_l)}{P(\alpha_h)P(\beta_l)}$$

$$= \sum_{h=1}^{N}\sum_{l=1}^{N} P(\alpha_h)P(\beta_l \mid \alpha_h)\log\frac{P(\beta_l \mid \alpha_h)}{P(\beta_l)}$$

2、平均互信息量与各类熵的关系

$$I(\overline{X};\overline{Y}) = H(\overline{X}) - H(\overline{X}|\overline{Y})$$

$$= H(\overline{Y}) - H(\overline{Y}|\overline{X})$$

定理3.8　　信源发出的N元随机变量序列

$\overline{X} = X_1 X_2 \cdots X_N$，

通过信道传送，输出N元随机变量

$\overline{Y} = Y_1 Y_2 \cdots Y_N$。若信道无记忆，

则有$I(\overline{X};\overline{Y}) \le \sum_{k=1}^{N} I(X_k;Y_k)$

证明：

$$H(\overline{Y}\,|\,\overline{X}) = -\sum_{h=1}^{r^N}\sum_{l=1}^{s^N} P(\alpha_h, \beta_l)\log P(\beta_l \,|\, \alpha_h)$$

$$= -\sum_{h_1,h_2,\cdots h_N}\;\sum_{l_1,l_2,\cdots l_N} P(a_{h1}, a_{h2}, \cdots a_{hN}, b_{l1}, b_{l2}, \cdots, b_{lN}) \bullet \log\left[\prod_{k=1}^{N} P(b_{l_k} \,|\, a_{h_k})\right]$$

$$= -\sum_{k=1}^{N}\sum_{h_k}\sum_{l_k} P(a_{h_k}, b_{l_k}) \log P(b_{l_k} \,|\, a_{h_k})$$

$$= \sum_{k=1}^{N} H(Y_k \,|\, X_k)$$（信道无记忆）

且$H(\overline{Y}) = H(Y_1 Y_2 \cdots Y_n)$

$$= H(Y_1) + H(Y_2 \mid Y_1) + H(Y_3 \mid Y_2 Y_1) + \cdots + H(Y_N \mid Y_{N-1} \cdots Y_2 Y_1)$$

$$\le \sum_{k=1}^{N} H(Y_k)$$

故

$$I(\overline{X};\overline{Y}) = H(\overline{Y}) - H(\overline{Y} \mid \overline{X})$$

$$\le \sum_{k=1}^{N} H(Y_k) - \sum_{k=1}^{N} H(Y_k \mid X_k)$$

$$= \sum_{k=1}^{N} I(X_k;Y_k)$$

定理3.9

信源发出的N元随机变量序列

$$\overline{X} = X_1 X_2 \cdots X_N,$$

通过信道传送，输出N元随机变量序列

$$\overline{Y} = Y_1 Y_2 \cdots Y_N。$$

若信源无记忆，则有

$$I(\overline{X};\overline{Y}) \geq \sum_{k=1}^{N} I(X_k;Y_k)$$

证明：

因信源无记忆，

$$故 H(\overline{X}) = H(X_1 X_2 \cdots X_N) = \sum_{k=1}^{N} H(X_k)$$

$$且 H(\overline{X} \mid \overline{Y}) = H(X_1 X_2 \cdots X_N \mid Y_1 Y_2 \cdots Y_N)$$

$$= H(X_1 \mid Y_1 Y_2 \cdots Y_N) + H(X_2 \mid Y_1 Y_2 \cdots Y_N)$$

$$+ \cdots + H(X_N \mid Y_1 Y_2 \cdots Y_N)$$

$$\leq \sum_{k=1}^{N} H(X_k \mid Y_k)$$

推论：信源发出的N元随机变量序列

$\overline{X} = X_1 X_2 \cdots X_N$，

通过信道传送，输出 N元随机变量序列

$\overline{Y} = Y_1 Y_2 \cdots Y_N$。

若信道和信源均无记忆，则有

$$I(\overline{X};\overline{Y}) = \sum_{k=1}^{N} I(X_k;Y_k)$$

信源信道均无记忆时的信道容量：

1、N元随机变量序列X_1，X_2，…，X_N中各X_k独立同分布，即所有X_k都是完全相同的随机变量，记这些相同的输入随机变量为X。

2、无记忆信源的N元序列加到无记忆信道，得到的N元输出序列Y_1，Y_2，…，Y_N中各Y_k必然是独立同分布的，所有Y_k都是完全相同的随机变量，记这些相同的随机变量为Y。

3、$I(X_1;Y_1)=I(X_2;Y_2)=\ldots=I(X_k;Y_k)=\ldots=I(X;Y)$

$\therefore$ 若信源信道均无记忆，则有$I(\overline{X};\overline{Y})=NI(X;Y)$

离散无记忆信道*N*次扩展信道的信道容量为

$$\boldsymbol{C}^{N}=\max_{\boldsymbol{P}_X}\boldsymbol{I}(\overline{\boldsymbol{X}};\overline{\boldsymbol{Y}})=\max\sum_{\boldsymbol{k}=1}^{\boldsymbol{N}}\boldsymbol{I}(\boldsymbol{X}_{\boldsymbol{k}};\boldsymbol{Y}_{\boldsymbol{k}})$$

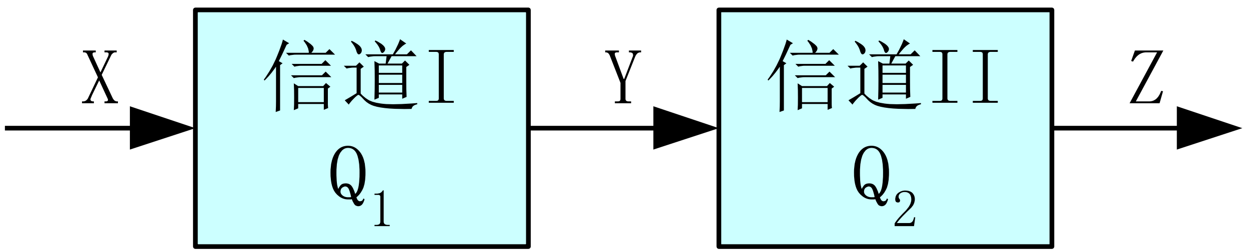
$$=\sum_{\boldsymbol{k}=1}^{\boldsymbol{N}}\max\boldsymbol{I}(\boldsymbol{X}_{\boldsymbol{k}};\boldsymbol{Y}_{\boldsymbol{k}})=\sum_{\boldsymbol{k}=1}^{\boldsymbol{N}}\boldsymbol{C}_{\boldsymbol{k}}$$

因信道容量是信道的固有参数，只与信道自身的统计特性有关，对同一信道，所有C_k均相同，等于单符号信道容量C，因此，离散无记忆信道的*N*次扩展信道的信道容量为

$$C^{N}=NC$$

3.7 信道的组合

3.7.1 串连信道：前一信道的输出符号集与后一信道的输入符号集一致。



Q_1、Q_2为两个信道的转移矩阵

记串联信道中3个随机变量X、Y、Z的取值符号集分别为

$A_X=\{a_1,a_2,\ldots,a_r\}$; $A_Y=\{b_1,b_2,\ldots,b_s\}$; $_Z=\{c_1,c_2,\ldots,c_t\}$;

在给定Y之后，Z的取值与X无关，这意味着

$$P(c_k \mid a_i b_j) = P(c_k \mid b_j)$$

对所有i, j, k　(XYZ组成一个马尔可夫链)

$$\therefore P\left(c_k \mid a_i\right) = \sum_{j=1}^{s} P(b_j c_k \mid a_i)$$

$$= \sum_{j=1}^{s} P(b_j \mid a_i) P(c_k \mid a_i b_j) = \sum_{j=1}^{s} P(b_j \mid a_i) P(c_k \mid b_j)$$

这说明串联信道的转移概率矩阵是各单元信道的转移概率矩阵之积：

设N个单元信道的转移概率矩阵分布为$\boldsymbol{Q}_1$、$\boldsymbol{Q}_2$、$\cdots$、$\boldsymbol{Q}_N$，

则整个串联信道的转移概率矩阵为$\boldsymbol{Q} = \boldsymbol{Q}_1 \boldsymbol{Q}_2 \cdots \boldsymbol{Q}_N = \prod_{k=1}^{N} \boldsymbol{Q}_k$

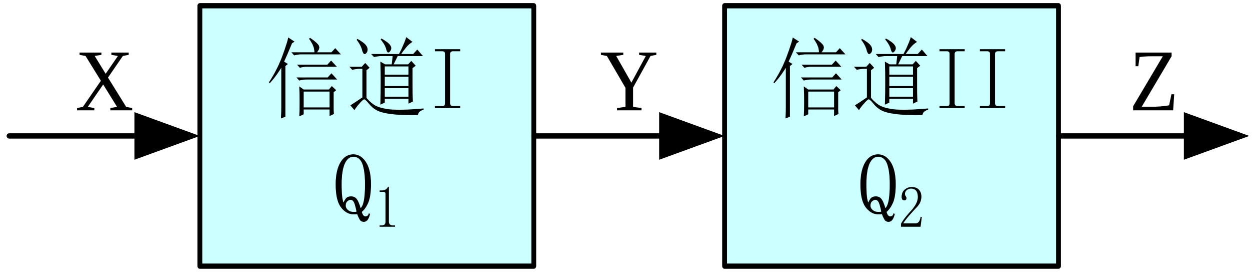
定理3.10 若随机变量XYZ组成马尔可夫链，则有

$$I(X;Z) \le I(X;Y)$$,等号成立的充要条件是

$$P(a_i|b_j c_k)=P(a_i|c_k);$$

$$I(X;Z) \le I(Y;Z)$$,等号成立的充要条件是

$$P(c_k|a_i b_j)=P(c_k|a_i)$$



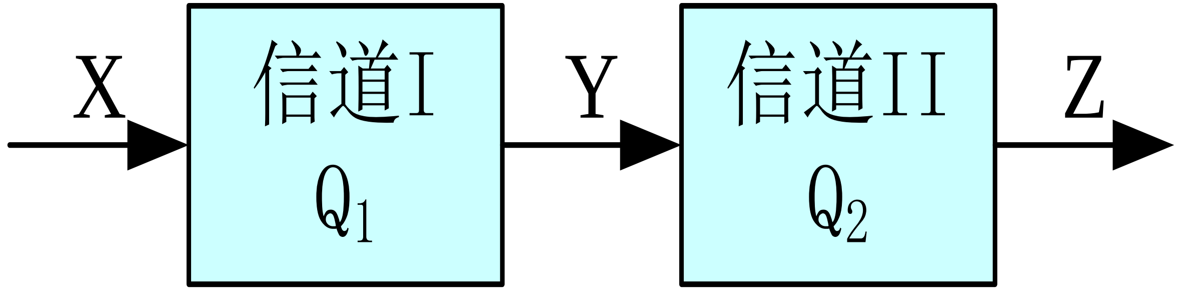
Q_1、Q_2为两个信道的转移矩阵

解释：

（1）从Z中所获得的X的信息不大于从Y中获得的X的信息，也就是说信道II对我们了解X的信息无任何帮助。

（2）信息不增原理：通过信道的信息不会增加。

（3）数据处理定理：数据经过处理之后，不会使信息增加，随着数据的不断处理，从处理后的数据中所得的原始信息会愈来愈少。



Q_1、Q_2为两个信道的转移矩阵

串联信道的信道容量:

与组成串联信道的各单元信道的信道容量之间无确切关系，**必须根据整个信道的数学模型进行求解**

例3.12 求2个相同二元对称信道（BSC）组成串联信道的信道容量

解：单个信道转移概率矩阵为：

$$Q_1 = Q_2 = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix}$$

串联信道转移概率矩阵为

$$Q = Q_1 Q_2 = \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} \begin{bmatrix} \bar{p} & p \\ p & \bar{p} \end{bmatrix} = \begin{bmatrix} 1-2p\bar{p} & 2p\bar{p} \\ 2p\bar{p} & 1-2p\bar{p} \end{bmatrix}$$

因串联信道仍然是二元对称信道，故

$$C = \log s - H(p_1', p_2', \ldots p_s') = 1 - h_2(2p\bar{p}) \quad bit/\text{符号}$$

推广1：若是**N**个**BSC**串联，则可以得到总的转移概率矩阵为：

$$Q=\begin{bmatrix}\bar{p} & p\\ p & \bar{p}\end{bmatrix}^N=\begin{bmatrix}1-\dfrac{1-(1-2p)^N}{2} & \dfrac{1-(1-2p)^N}{2}\\ \dfrac{1-(1-2p)^N}{2} & 1-\dfrac{1-(1-2p)^N}{2}\end{bmatrix}$$

仍然为对称信道，信道容量为

$$C_{(N)}=1-h_2\left(\frac{1-(1-2p)^N}{2}\right)\quad bit/\text{符号}$$

推广2：只要信道是有噪的，及**0<*p*<1**,则

$$\lim_{N\to\infty} Q = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix};$$

$$\lim_{N\to\infty} C_{(N)} = 1 - h_2\left(\frac{1}{2}\right) = 0 \quad bit/\text{符号}$$

可见，有噪**BSC**的串联信道，随着串联级数的增加，整个信道的容量趋于零。（因为随着有噪串联环节的增加，串联信道的平均互信息是递减的）

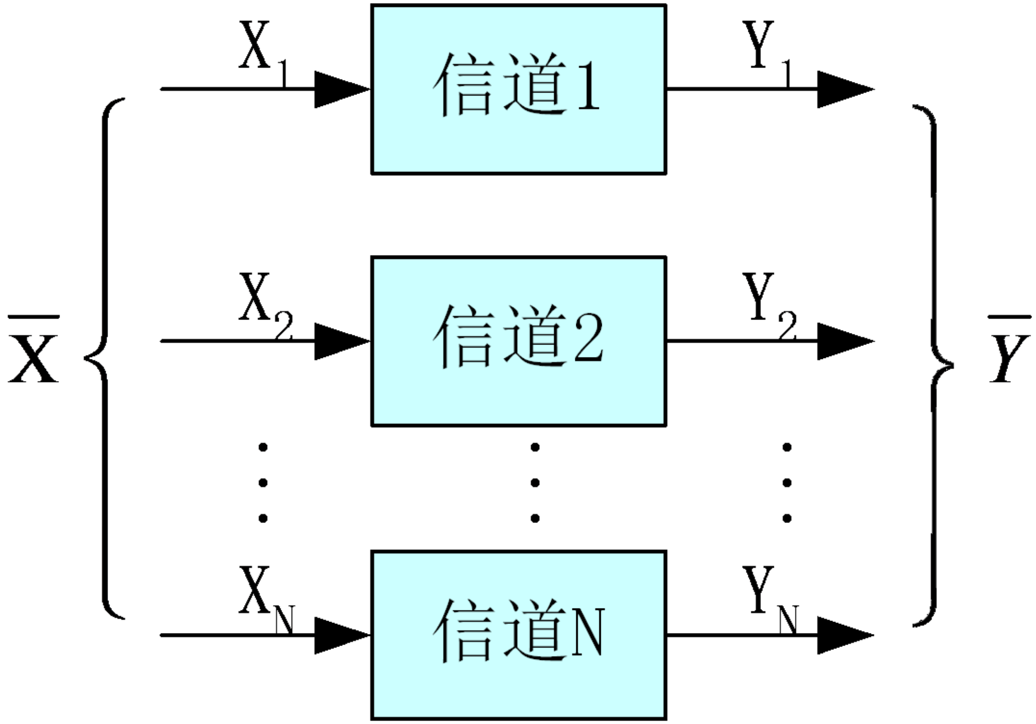
3.7.2 独立并联信道

N 维输入 $\bar{X}$ 的各分量分别送入 N 个独立信道，

各独立信道的输出组成 N 维输出 $\bar{Y}$。因此，

独立并联信道可等效成一个多符号信道——

N 次扩展信道 $\{\bar{X}, P_{\bar{X}|\bar{Y}}, \bar{Y}\}$，其平均互信息量为 $I(\bar{X};\bar{Y})$。



信道容量： 由独立并联信道的特点可知，Y_k只与X_k有关，即**等效信道是无记忆的**，根据定理3.8，平均互信息量满足

$$I(\overline{X};\overline{Y}) \le \sum_{k=1}^{N} I(X_k;Y_k)$$

所以，独立并联信道的信道容量为

$$C = \max_{P_X} I(\overline{X};\overline{Y}) = \sum_{k=1}^{N} I(X_k;Y_k) = \sum_{k=1}^{N} C_k$$

即独立并联信道的信道容量为各组成信道的信道容量之和

3.8 信源与信道的匹配

(1) 符号匹配:

信源输出的符号必须是信道能够传送的符号，即要求信源符号集就是信道的入口符号集或入口符号集的子集，这是实现信息传输的必要条件，可在信源与信道之间加入编码器予以实现

(2)信息匹配：信源与信道匹配的程度可用信道剩余度来衡量，

信道绝对剩余度＝$C-I$（$X;Y$）

信道相对剩余度

$$=\frac{C-I(X;Y)}{C}\times 100\%=1-\frac{I(X;Y)}{C}\times 100\%$$

剩余度	大	小	零
匹配程度	信源与信道匹配程度低	信源与信道匹配程度高	信源与信道完全匹配
信道的利用情况	信道的信息传递能力未得到充分利用	信道的信息传递能力得到较充分利用	信道的信息传递能力得到完全利用

一般来说，实际信源的概率分布P_X未必是信道的最佳分布P^*_X,所以，

$I(X;Y) \le C$,剩余度不为零。

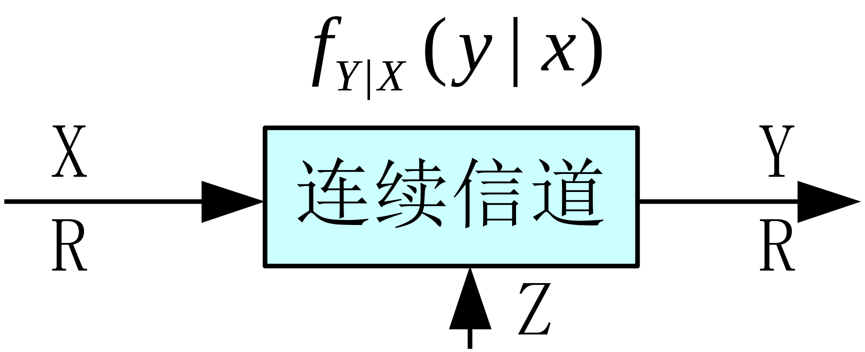
因此，要求信源与信道达到信息的完全匹配是不现实的，只要信道剩余度较小就可以了。

3.9 连续信道及其信道容量

连续信道：是时间离散、幅值连续信道的简称。连续信道的输入和输出都是定义在整个实域R或R的某个子集上的连续性随机变量。

3.9.1 连续信道的数学模型

一、单维连续信道：输入X、输出Y以及噪声Z都是取值于整个实域R的一维连续型随机变量。统计特性由转移概率密度函数$f_{Y|X}(y|x)$描述。数学模型记为：$\{X, f_{Y|X}(y|x), Y\}$



$$\int_R f_{Y|X}(y \mid x)dy = 1$$

二、多维连续信道：

输入X、输出Y以及噪声Z都是多维连续随机变量。

数学模型为：$\left\{\overline{X}, f_{\overline{Y}|\overline{X}}(\overline{y}\,|\,\overline{x}), \overline{Y}\right\}$

三、连续无记忆信道：

假设连续信道的N维输入

$$\overline{\boldsymbol{X}} = \boldsymbol{X}_1\boldsymbol{X}_2\cdots\boldsymbol{X}_N,\quad N\text{维输出}\ \overline{\boldsymbol{Y}} = \boldsymbol{Y}_1\boldsymbol{Y}_2\cdots\boldsymbol{Y}_N,$$

转移概率密度函数满足

$$f_{\overline{Y}|\overline{X}}(\overline{y}\,|\,\overline{x}) = f_{\overline{Y}|\overline{X}}(y_1 y_2 \cdots y_N \,|\, x_1 x_2 \cdots x_N)$$

$$= \prod_{k=1}^{N} f_{Y_k|X_k}(y_k \,|\, x_k)$$

四、连续信道的平均互信息量*I*(*X*;*Y*)

$$I(X;Y)=\iint_{R\ R} f_{XY}(x,y)\log\frac{f_{XY}(x,y)}{f_X(x)f_Y(y)}dxdy$$

$$=\iint_{R\ R} f_X(x)f_{Y|X}(y|x)\log\frac{f_{Y|X}(y|x)}{f_Y(y)}dxdy$$

$$I(X;Y)=h(Y)-h(Y|X)$$

$$=h(X)-h(X|Y)$$

五、连续信道的信息传输能力：信道容量

与离散信道不同，连续信道 的输入取值区间和概率密度不能完全描述实际输入的某些性质,如输入信号的幅值受限、功率受限等。

为了适当反映这种情况，可对连续信道的输入加一个限制条件$b(X)$，而连续信道的信道容量定义为该信道的$I(X;Y)$在条件$b(X)$下关于$f_X(x)$的最大值

$$C = \max_{f_X(x)} \{ I(X;Y); b(X) \}$$

$$= \max_{f_X(x)} \{ h(Y) - h(Y \mid X); b(X) \}$$

若最大值不存在，可取其最小上界（上确界）：

$$C = \sup_{f_X(x)} \{ I(X;Y); b(X) \}$$

$$= \sup_{f_X(x)} \{ h(Y) - h(Y \mid X); b(X) \}$$

实际应用中，信道输入信号的平均功率 $E(X^2)$ 总是限定在一定范围之内，假设输入信号的平均功率限定在 P_S 以内，这时，对信道输入信号的限制条件可描述为

$$b(X): E(X^2) \le P_S$$

此时，连续信道在输入平均功率受限时的信道容量为：

或

$$C(P_S) = \max_{f_X(x)} \left\{ I(X;Y); E(X^2) \le P_S \right\}$$

$$C(P_S) = \sup_{f_X(x)} \left\{ I(X;Y); E(X^2) \le P_S \right\}$$

3.9.2 加性高斯噪声信道的信道容量

求解一般信道的信道容量非常困难，往往只能得出数值解。只有一些特殊的连续信道如加性噪声信道，才能推出简明的信道容量表达式。幸运的是，实际使用的连续信道大部分可近似看成是加性噪声信道，研究这种信道在理论和实践两方面都有重大意义。

一、概念：信道的输入X、输出Y以及噪声Z 3个随机变量之间满足 $Y=X+Z$ 且输入X与干扰Z无关，则称该信道为加性噪声信道。

二、加性噪声信道的转移概率密度函数:

$$f_{Y|X}(y|x)=f(y-x)=f_Z(z)$$

给定x，出现y的概率取决于z（详细推导参见文献）

上式说明，转移概率密度是由噪声引起的，加性噪声信道的转移概率密度函数等于噪声的概率密度函数

三、加性高斯噪声信道的容量

1、噪声分布密度

$$f_Z(z)=\frac{1}{\sqrt{2\pi\sigma}}e^{-(z-\mu)^2/2\sigma^2}$$

2、转移概率密度

$$f_{Y|X}(y\,|\,x)=f_Z=\frac{1}{\sqrt{2\pi\sigma}}e^{-(y-x-\mu)^2/2\sigma^2}$$

3、条件熵（噪声熵）

$$h(Y|X)=\int\limits_R\int\limits_R f_X(x)f_{Y|X}(y|x)\log\frac{1}{f_{Y|X}(y|x)}dxdy$$

$$=\int\limits_R\int\limits_R f_X(x)f_Z(y-x)\log\frac{1}{f_Z(y-x)}dxdy$$

$$=\int\limits_R\int\limits_R f_X(x)f_Z(z)\log\frac{1}{f_Z(z)}dxdz$$

$$=h(Z)$$

4、平均互信息量

$$I(X;Y)=h(Y)-h(Y|X)=h(Y)-h(Z)$$

5、信道容量

$$C(P_S) = \max_{f_X(x)} \left\{ h(Y); E(X^2) \le P_S \right\} - h(Z)$$

$$\text{或} C(P_S) = \sup_{f_X(x)} \left\{ h(Y); E(X^2) \le P_S \right\} - h(Z)$$

把对*I*(*X*;*Y*)求最大值的问题转化为

对*h*(*Y*)求最大值的问题

6、噪声Z的微分熵

设加性噪声Z是均值为0，方差为σ_z^2的高斯分布，

即概率密度函数为 $$f_Z(z)=\frac{1}{\sqrt{2\pi}\sigma_z}e^{-z^2/2\sigma_z^2}$$

假设Z的均值为0并不失一般性。因为若Z`是均值为$\mu$ 的随机变量，通过坐标变换Z=Z`- μ,则Z是均值为0的随机变量，并且这种特殊的坐标变换不改变随机变量的熵

Z的平均功率：$\sigma_z^2=E\left[Z^2\right]-E^2[Z]=E\left[Z^2\right]=P_N$

Z的概率密度函数可写成： $f_Z(z)=\dfrac{1}{\sqrt{2\pi P_N}}e^{-z^2/2P_N}$

Z的微分熵*h*(Z):

$$h(Z)=\frac{1}{2}\log(2\pi eP_N)$$

信道容量:

$$C(P_S)=\max_{f_X(x)}\left\{h(Y);E(X^2)\le P_S\right\}-\frac{1}{2}\log(2\pi eP_N)$$

问题的关键转化为求信道输出熵关于概率密度$f_X(x)$的最大值,可以分3步来求信道容量

（1）与对待噪声Z的方式一样，设信道输入X的均值为0，则信道输出Y的均值为：

$$E(Y)=E(X+Z)=E(X)+E(Z)=0$$

平均功率为

$$E(Y^2)=E\left[\left(X+Z\right)^2\right]=E\left[X^2\right]+E\left[Z^2\right]+2E\left[XZ\right]$$

$$=E\left[X^2\right]+E\left[Z^2\right]=P_S+P_N\left(\text{X、Z统计独立}\right)$$

说明对于加性噪声信道，当输入X和噪声Z的均值都是0、平均功率受限与P_S和P_N时，输出Y的均值也为0，其平均功率受限于P_S＋P_N

（2）根据连续最大熵的已知结论，平均功率受限时，随机变量只有服从高斯分布才会使熵达到最大。现在，信道输出Y的均值为0，平均功率受限于P_S+P_N，因此，只有Y服从高斯分布时，其熵最大。

$$f_Y(y)=\frac{1}{\sqrt{2\pi(P_S+P_N)}}e^{-y^2/2(P_S+P_N)}$$

最佳输入分布：因为Y和Z均服从高斯分布，而$X=Y-Z$，所以X也服从高斯分布，而X的均值为0，平均功率为P_S，故最佳输入分布为

$$f_X(x)=\frac{1}{\sqrt{2\pi P_S}}e^{-x^2/2P_S}$$

（3）求出Y关于输入概率密度的最大熵

$$\max_{f_X(x)}\left\{h(Y);E(X^2)\le P_S\right\}$$

$$=\max_{f_X(x)}\left\{h(Y);E(Y^2)\le P_S+P_N\right\}$$

$$=\frac{1}{2}\log\left[2\pi e\left(P_S+P_N\right)\right]$$

加性高斯噪声信道的信道容量:

$$C(P_S)=\frac{1}{2}\log\left[2\pi e\left(P_S+P_N\right)\right]-\frac{1}{2}\log(2\pi eP_N)$$

$$=\frac{1}{2}\log\left(1+\frac{P_S}{P_N}\right)$$

$$C(P_S)=\frac{1}{2}\log\left[2\pi e\left(P_S+P_N\right)\right]-\frac{1}{2}\log(2\pi eP_N)$$

$$=\frac{1}{2}\log\left(1+\frac{P_S}{P_N}\right)$$

说明：

（1）在加性高斯噪声信道中传输信息，高斯分布的输入信号是最有效的；

（2）信道容量与信噪比P_S/P_N有关。

问题：对于加性噪声信道，若输入信号服从高斯分布，那么，什么性质的噪声最有害？

定理3.11 ——

对于无记忆加性噪声信道，假设输入信号服从高斯分布，且噪声的平均功率受限，则服从高斯分布的噪声使信道平均互信息量达到最小。

即$I(X_G;Y) \geq I(X_G;Y_G)$

3.9.3 一般加性噪声信道的信道容量的界

非高斯加性噪声信道的信道容量计算非常复杂，即使在平均功率受限条件下，也无法给出解析形式的解，只能对其上下限作出估计。

以下定理给出了平均功率受限条件下一般加性噪声信道的信道容量的上下界

定理3.12 对于一般的无记忆加性噪声信道，假设输入信号的平均功率受限于P_S，噪声的平均功率受限于P_N，则信道容量$C(P_S)$的上下界为

$$\frac{1}{2}\log\left(1+\frac{P_S}{P_N}\right)\le C\left(P_S\right)\le\frac{1}{2}\log\left(\frac{P_S+P_N}{P}\right),$$

其中$P=\frac{1}{2\pi e}e^{2h(Z)}$是

具有微分熵$h(Z)$随机变量Z的熵功率

证明：先证明下界。仍然假设输入和噪声均为0。

$$C(P_S)=\max_{f_X(x)}\left\{I(X;Y);E(X^2)\le P_S\right\}\ge I(X_G;Y)$$

(信道容量是平均互信息量的最大值)

由定理3.11可知，$I(X_G;Y)\ge I(X_G;Y_G)$

$I(X_G;Y)\ge I(X_G;Y_G)$即为加性高斯噪声的信道容量，故

$$C\left(P_S\right)\ge I(X_G;Y_G)\ge\frac{1}{2}\log\left(1+\frac{P_S}{P_N}\right)$$

下界得证

再证上界: 由以上分析可知，当输入信号和噪声的平均功率分别受限于P_S和P_N时，此时信道输出信号Y的平均功率受限于P_S+P_N，Y服从高斯分布时其熵最大，即

$$h(Y) \le \max_{f_X(x)} \left[h(Y); E(Y) \le P_S + P_N \right]$$

$$= \frac{1}{2} \log \left[2\pi e \left(P_S + P_N \right) \right]$$

于是

$$C(P_S) \le \frac{1}{2} \log \left[2\pi e \left(P_S + P_N \right) \right] - h(Z)$$

然后用熵功率将上式表示成统一形式，因为噪声Z的熵功率为：

$$P=\frac{1}{2\pi e}e^{2h(Z)}$$

则Z的熵为：

$$\therefore C(P_S)\le\frac{1}{2}\log\left[2\pi e\left(P_S+P_N\right)\right]-\frac{1}{2}\log 2\pi eP$$

$$=\frac{1}{2}\log\frac{P_S+P_N}{P}$$

$h(Z)=\frac{1}{2}\ln 2\pi eP$ 奈特/符号

或 $h(Z)=\frac{1}{2}\log 2\pi eP$ 比特/符号

上界得证

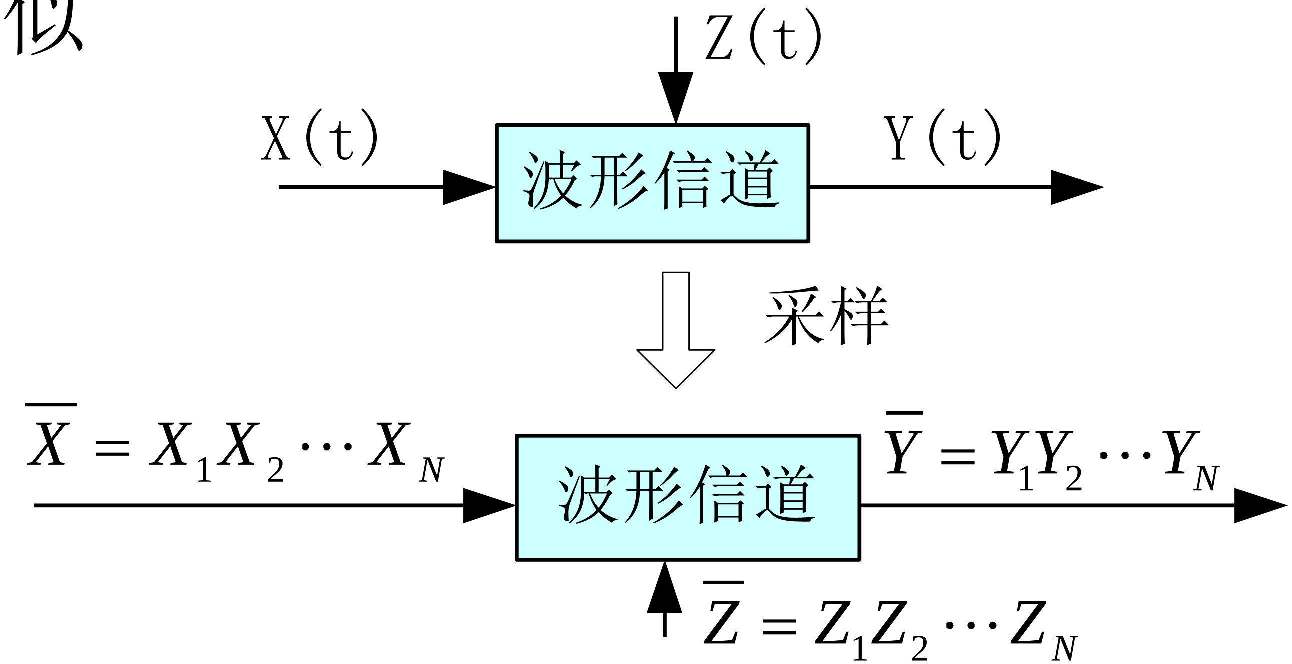
3.10 波形信道及其信道容量

一、波形信道（waveform channel）的概念：

输入/输出随时间连续取值、且取值集合是连续区间的信道，也称为模拟信道。在波形信道的输入和输出分别用随机过程$X(t)$和$Y(t)$来描述。

二、近似处理

在波形信道的持续时间*T*内对其输入和输出进行采样，采样所得的信道可以用一个*N*维连续信道来近似



一般来说，波形信道是无穷维连续信道，所以，当$N \to \infty$ 时，N维连续信道的平均互信息的极限就是波形信道的平均互信息

$$
\begin{aligned}
I(X(t);Y(t)) &= \lim_{N \to \infty} I(\overline{X};\overline{Y}) \\
&= \lim_{N \to \infty} I(X_1 X_2 \cdots X_N ; Y_1 Y_2 \cdots Y_N)
\end{aligned}
$$

三、信道容量：持续时间为*T*的波形信道的信道容量为

$$C_T(P_S) = \lim_{N\to\infty} \max_{f_{\overline{X}}(x)} \left\{ I(\overline{X};\overline{Y}); E(\overline{X^2}) \le P_S \right\}$$

波形信道是无穷维连续信道，信道容量的一般性研究在数学上存在相当大的困难。我们只讨论一种简单的波形信道————带限、加性高斯白噪声信道，对该信道的研究在理论与实用上都有很大意义。

带限、加性高斯白噪声信道：频带限制在一定范围之内、受加性高斯白噪声干扰的波形信道

	白噪声	有色噪声
噪声功率谱密度	在整个频域内均匀分布	在整个频域内分布不均匀

带限白噪声：白噪声$Z(t)$的频带限制为B，即$f\leqslant |B|$,则称之为带限白噪声，其功率谱密度为

$$P_Z(f)=\frac{N_0}{2} \quad -B\leq f\leq B,$$

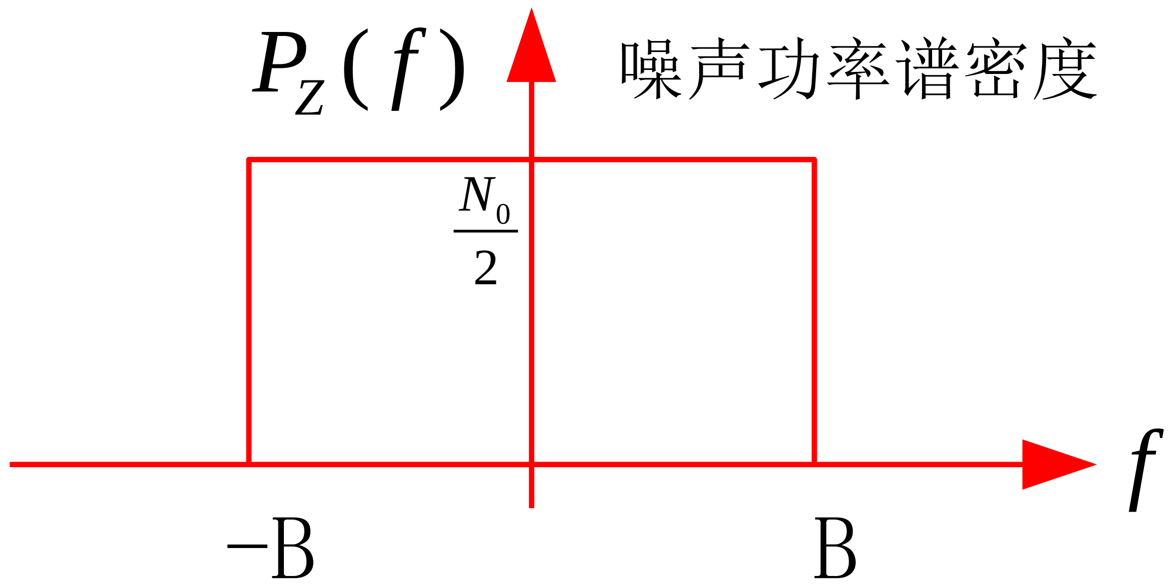
N_0为一个常数，单位是“W/Hz?（瓦/赫）

白噪声是一种理想噪声信号，实际并不存在，但如果噪声功率谱密度均匀分布的频率范围远大于所研究系统的带宽，则此噪声可认为是白噪声，

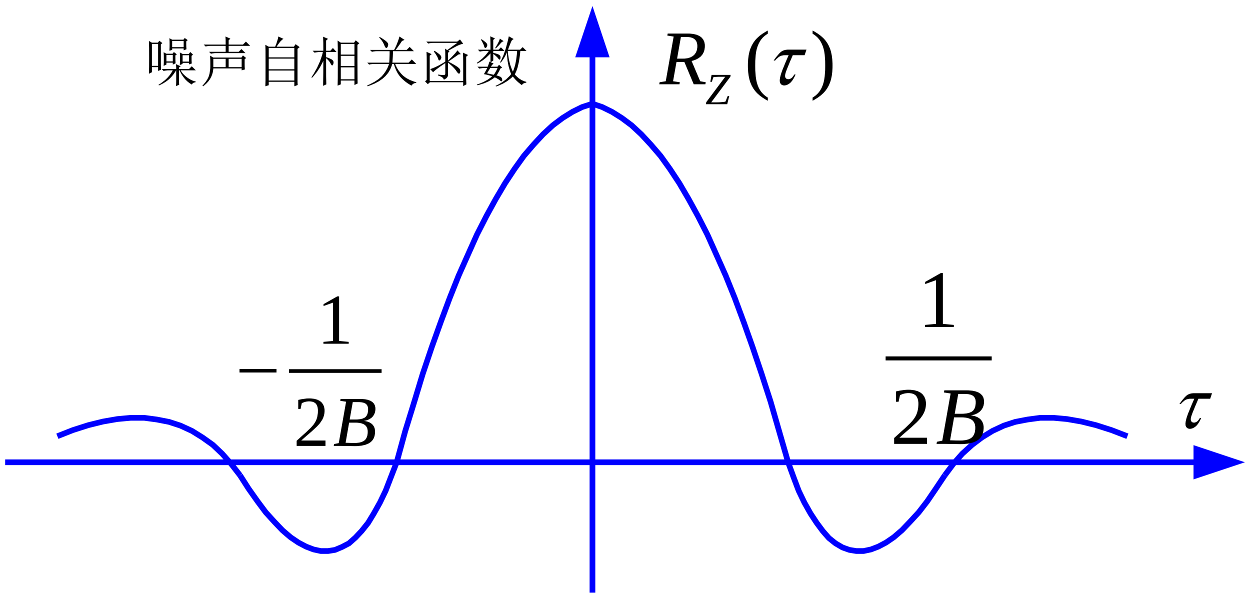
白噪声经滤波变为带限白噪声。服从高斯分布的带限白噪声称为带限高斯白噪声。带限白噪声的相关函数是其功率谱密度的傅立叶反变换：

$$R_Z(\tau)=\int_{-B}^{B}\frac{N_0}{2}e^{j2\pi f\tau}df=N_0B\frac{\sin(2\pi B\tau)}{2\pi B\tau}$$

其平均功率为 $\boldsymbol{P_N}=\boldsymbol{R_Z}(0)=\boldsymbol{N_0B}$



$$R_Z(\tau) = \int_{-B}^{B} \frac{N_0}{2} e^{j2\pi f\tau} df = N_0 B \frac{\sin(2\pi B\tau)}{2\pi B\tau}$$



当$\tau=k/2B$(k=1,2,…)时，自相关函数$R_Z(\tau)=0$,即

从这些点上得到的所有随机变量{$Z(k/2B)$, k=1,2,…}互不相关

根据采样定理，带限为B的随机信号$Z(t)$可用一组独立随机变量序列$\{Z(k/2B), k=1,2,\ldots\}$来近似表示。另外，工程中所遇到的信号通常是限时信号，即只在有限的时段起作用。

如果$Z(t)$的作用时段为$[0,T]$，在此时段内以频率$f=2B$进行采样，只能取$N=2BT$个点。若记$Z_k=Z(k/2B)$,则上述结论可归纳为：

带限为 B 、限时为 T 的随机信号 $Z(t)$ 可用 $N=2BT$ 个相互独立的随机变量组成的序列 $\bar{Z}=Z_1Z_2\cdots Z_N$来近似表示。

如果$Z(t)$服从均值为0的高斯分布，采样之后所得的$2BT$个独立分量也都服从均值为0的高斯分布，各独立分量的平均功率等于$[0,T]$时段内的总功率除以分量个数：

$$\mathrm{P}_{N_k} = \frac{P_N T}{N} = \frac{N_0 BT}{2BT} = \frac{N_0}{2}, \quad k = 1, 2, \cdots N$$

按如下步骤求出带限为B、限时为T的加性高斯白噪声信道的信道容量

（1）时间离散化：因为$Y(t)=X(t)+Z(t)$，按照采样定理，对信号和噪声都只需要$N=2BT$个采样点，因此波形信道可用 N维连续信道来近似：

$$\overline{Y} = \overline{X} + \overline{Z},$$

其中$\overline{X} = X_1X_2\cdots X_N$， $\overline{Y} = Y_1Y_2\cdots Y_N$， $\overline{Z} = Z_1Z_2\cdots Z_N$

（2）根据信道的无记忆特性和噪声分量的相互独立性质，上述*N*维连续信道又可看出是*N*个独立的一维连续加性噪声信道的并联

并且，

$$Y_k = X_k + Z_k, \quad k = 1, 2, \cdots, N$$

$$I\Big(X(t);Y(t)\Big) = I\Big(\overline{X};\overline{Y}\Big) \le \sum_{k=1}^{2BT} I(X_k;Y_k) \quad (\because \text{信道无记忆})$$

因为$Z(t)$服从均值为0的高斯分布，且平均功率为$P_N=N_0B$,

所以，采样所得的噪声分量Z_k也服从均值为0的高斯分布，平均功率为$P_{Zk}=N_0/2$，N个独立的一维连续加性噪声信道都是加性高斯信道。

各平均功率受限于P_{Sk}的子信道，在X_k服从均值为0的高斯分布时，达到信道容量：

$$C_k(P_{S_k}) = \max\left\{I(X_k;Y_k); E(X_k^2) \le P_{S_k}\right\}$$

$$= \frac{1}{2}\log(1+\frac{P_{S_k}}{P_{N_k}}) = \frac{1}{2}\log(1+\frac{P_{S_k}}{N_0/2})$$

因此有：

$$I\left(X(t);Y(t)\right)=I\left(\overline{X};\overline{Y}\right)\le\sum_{k=1}^{2BT}C_k=\frac{1}{2}\sum_{k=1}^{2BT}\log(1+\frac{P_{S_k}}{N_0/2})$$

（3）若输入$X(t)$的平均功率受限于P_s，必须适当分配各子信道输入的平均功率P_{Sk}，才能使$I(X(t);Y(t))$达到最大，这相当于形成求极值的约束条件：

$$P_S=\frac{1}{T}\int_0^T E[X^2(t)]dt=\frac{1}{T}\sum_{k=1}^{2BT}E[X_k^2]=\frac{1}{T}\sum_{k=1}^{2BT}P_{S_k}$$

然后在此条件下，求：

$$C_T(P_S)=\max I(\overline{X};\overline{Y})=\max\left[\frac{1}{2}\sum_{k=1}^{2BT}\log(1+\frac{P_{S_k}}{N_0/2})\right]$$

由拉各朗日乘数法可知：当所以输入分量的平均功率P_{Sk}都相等时，出现最大值。即

$$P_S=\frac{1}{T}\sum_{k=1}^{2BT}P_{S_k}=\frac{1}{T}2BTP_{S_k}=2BP_{S_k}$$

$$\Rightarrow P_{S_k}=P_S/2B$$

信道容量为（香农信道容量公式）：

$$C_T(P_S) = \frac{1}{2}\sum_{k=1}^{2BT}\log(1+\frac{P_S/2B}{N_0/2})$$

$$= BT\log\left(1+\frac{P_S}{N_0 B}\right) \quad bit$$

四、香农信道容量公式的启发意义：

（1）用频带换取信噪比，即采用扩频通信

在信噪比$P_S/P_N=P_S/(N_0B)$不变的前提下，增大频带B，可增大信道容量。这种方法对空间通信有很现实的意义，因为在这种情况下频率资源相对丰富，而能源则很珍贵。

但用扩频方法来增大信道容量，其作用是有限的，因为当$B\to\infty$时，信道容量C的极限是有限的：

$$\lim_{B\to\infty} C(P_S) = \lim_{B\to\infty} B\log\left(1+\frac{P_S}{N_0 B}\right)$$

$$= \lim_{B\to\infty} \frac{1}{\ln 2} B\ln\left(1+\frac{P_S}{N_0 B}\right)(\text{换底公式})$$

$$= \frac{1}{\ln 2}\ln\lim_{B\to\infty}\left[\left(1+\frac{P_S}{N_0 B}\right)^{N_0 B/P_S}\right]^{\frac{P_S}{N_0}}$$

$$= \frac{P_S}{N_0\ln 2}\ln\lim_{B\to\infty}\left[\left(1+\frac{P_S}{N_0 B}\right)^{N_0 B/P_S}\right]$$

$$= \frac{P_S}{N_0\ln 2}\approx 1.44\frac{P_S}{N_0}\quad bit/s$$

$$C_T(P_S)=\frac{1}{2}\sum_{k=1}^{2BT}\log(1+\frac{P_S/2B}{N_0/2})=BT\log\left(1+\frac{P_S}{N_0B}\right)\quad bit/s$$

（2）用信噪比换取频带。频带B不变时，增大信噪比P_S/P_N即可增大信道容量C。这种方法也有局限性，因为增大信噪比P_S/P_N是靠加大输入功率P_S来实现的，而

$$\lim_{P_S\to\infty}\frac{dC(P_S)}{dP_S}=\lim_{P_S\to\infty}\frac{d}{dP_S}B\log\left(1+\frac{P_S}{N_0B}\right)$$

$$=\lim_{P_S\to\infty}B\bullet\frac{1}{\ln 2}\bullet\frac{1}{1+\frac{P_S}{N_0B}}\bullet\frac{1}{N_0B}=\lim_{P_S\to\infty}\frac{1}{\ln 2}\bullet\left(\frac{1}{N_0+P_S/B}\right)=0$$

可见，随着P_S的增大，$C(P_S)$的增长率逐步变小，直至为零。

这意味着，当P_S大到一定程度之后，即使P_S增加很多，$C(P_S)$的增长幅度却很小，得不偿失。

作业：3.20（第二版）