

第六章：离散时间系统结构

- ◆ 6.1 线性常系数差分方程的方框图表示
- ◆ 6.2 线性常系数差分方程的信号流图表示
- ◆ 6.3 IIR系统的基本结构
- ◆ 6.4 转置形式
- ◆ 6.5 FIR系统的基本网络结构
- ◆ 6.6 有限精度数值效应概述
- ◆ 6.7 系数量化效应
- ◆ 6.8 数字滤波器中舍入噪声效应
- ◆ 6.9 IIR数字滤波器定点实现中的零输入极限环

6.5.1 直接型FIR系统结构

◆直接型FIR系统

FIR系统的差分方程可简化为

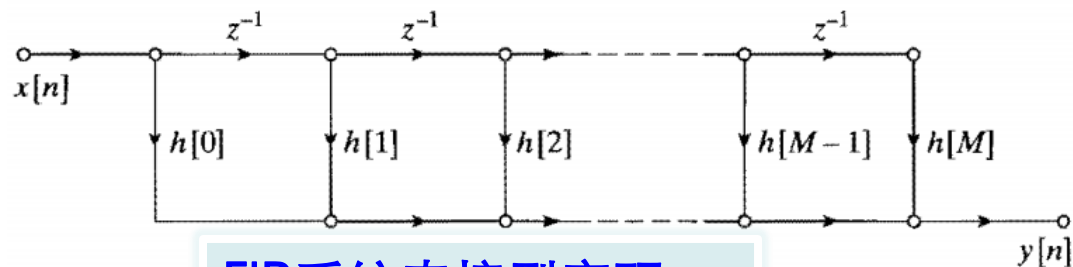
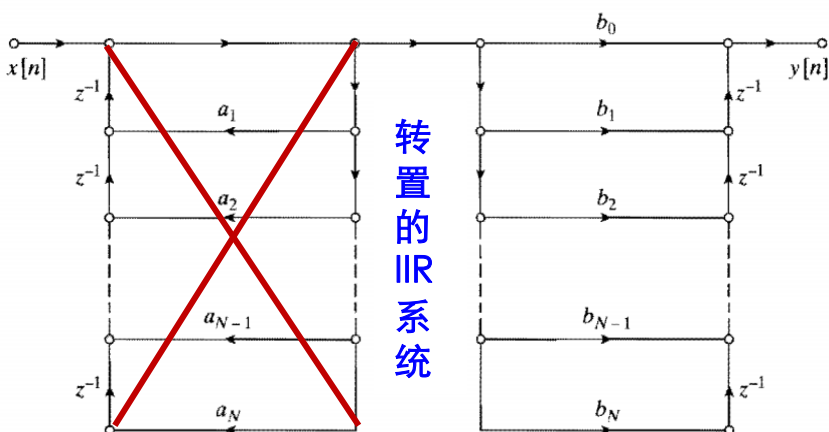
$$y[n] = \sum_{k=0}^M b_k x[n-k]$$

上式等效为

$$y[n] = h[n] * x[n]$$

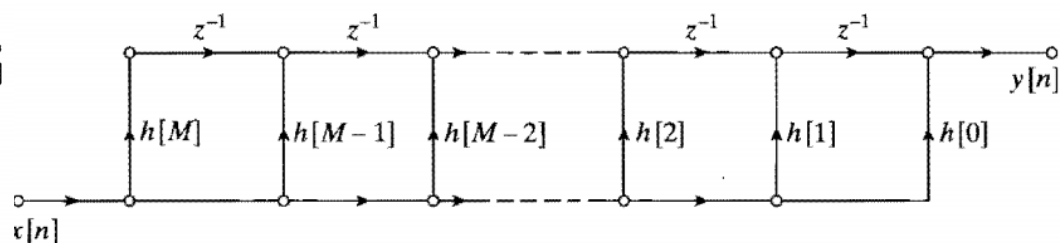
其中

$$h[n] = \begin{cases} b_n, & n = 0, 1, \dots, M \\ 0, & \text{其他} \end{cases}$$



FIR系统直接型实现

抽头延迟线结构或横向滤波器结构



FIR系统转置型实现

6.5.2 级联型FIR系统结构

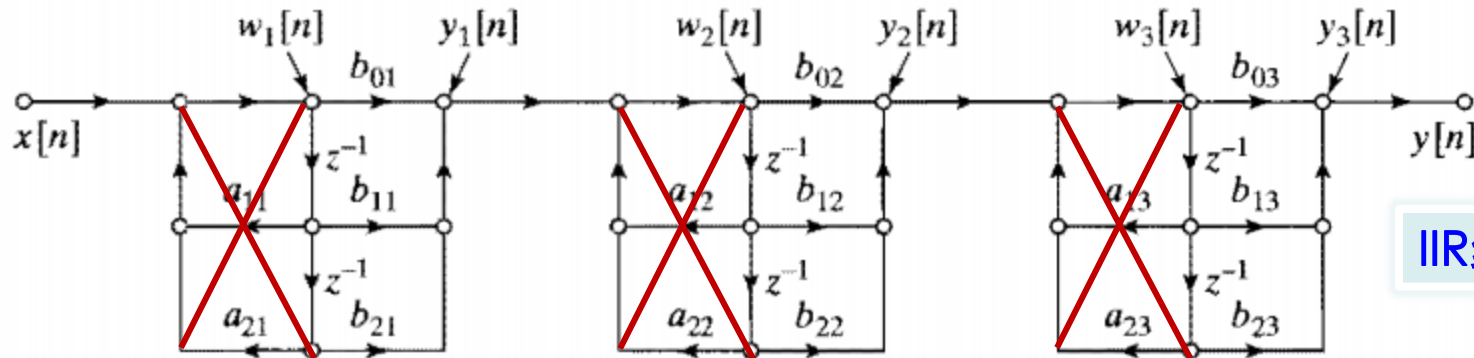
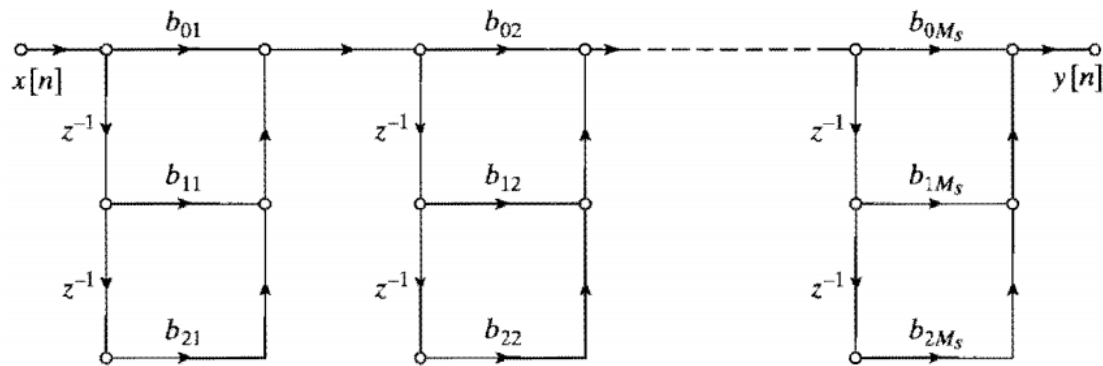
◆级联型FIR系统

将多项式系统函数因式化即可获得FIR系统的级联型

$$H(z) = \sum_{n=0}^M h[n] z^{-n} = \prod_{k=1}^{M_s} (b_{0k} + b_{1k} z^{-1} + b_{2k} z^{-2})$$

FIR系统级联型实现

其中 $M_s = \lfloor (M+1)/2 \rfloor$



IIR级联型

6.5.3 线性相位FIR系统结构

◆ 广义线性相位系统实现结构

若：单位脉冲响应满足如下条件的**广义线性相位**系统

$$h[n] = h[M - n], \quad 0 \leq n \leq M$$

$$\text{或} \quad h[n] = -h[M - n], \quad 0 \leq n \leq M$$

则：其实现结构中**系数乘法器的数目**基本上能减半。

□ 对于**I类**和**III类**FIR系统（ M 为偶数，**奇数**个系数）

$$y[n] = \sum_{k=0}^M h[k] x[n-k]$$

有 **$M+1$** 个系数乘法器

$$= \sum_{k=0}^{M/2-1} h[k] x[n-k] + h[M/2] x[n-M/2] + \sum_{k=M/2+1}^M h[k] x[n-k]$$

$$= \sum_{k=0}^{M/2-1} h[k] x[n-k] + h[M/2] x[n-M/2] + \sum_{k=0}^{M/2-1} h[M-k] x[n-M+k]$$

6.5.3 线性相位FIR系统结构

□对于**I类**FIR线性相位系统（ M 为偶数），利用 $h[n]$ 的**对称性**

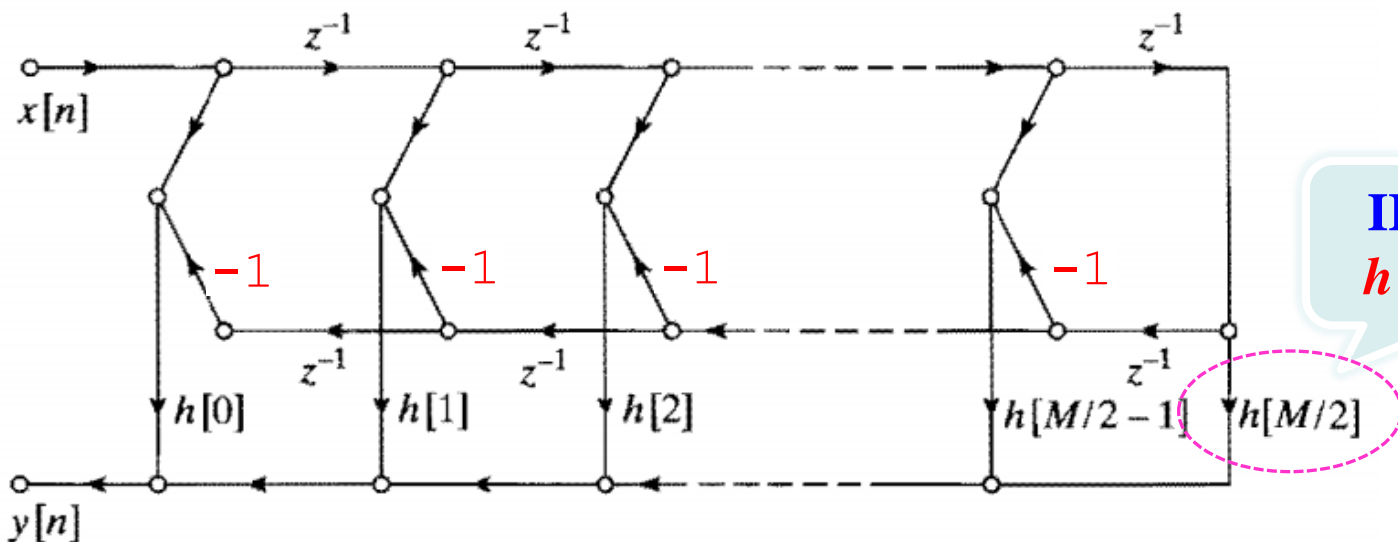
$$y[n] = \sum_{k=0}^{M/2-1} h[k] (x[n-k] + x[n-M+k]) + h[M/2] x[n-M/2]$$

只需 $M/2+1$ 个
系数乘法器

□对于**III类**FIR线性相位系统（ M 为偶数），利用 $h[n]$ 的**反对称性**

$$y[n] = \sum_{k=0}^{M/2-1} h[k] (x[n-k] - x[n-M+k])$$

只需 $M/2$ 个
系数乘法器



6.5.3 线性相位FIR系统结构

□对于**II类**FIR线性相位系统（ M 为奇数），利用 $h[n]$ 的**对称性**

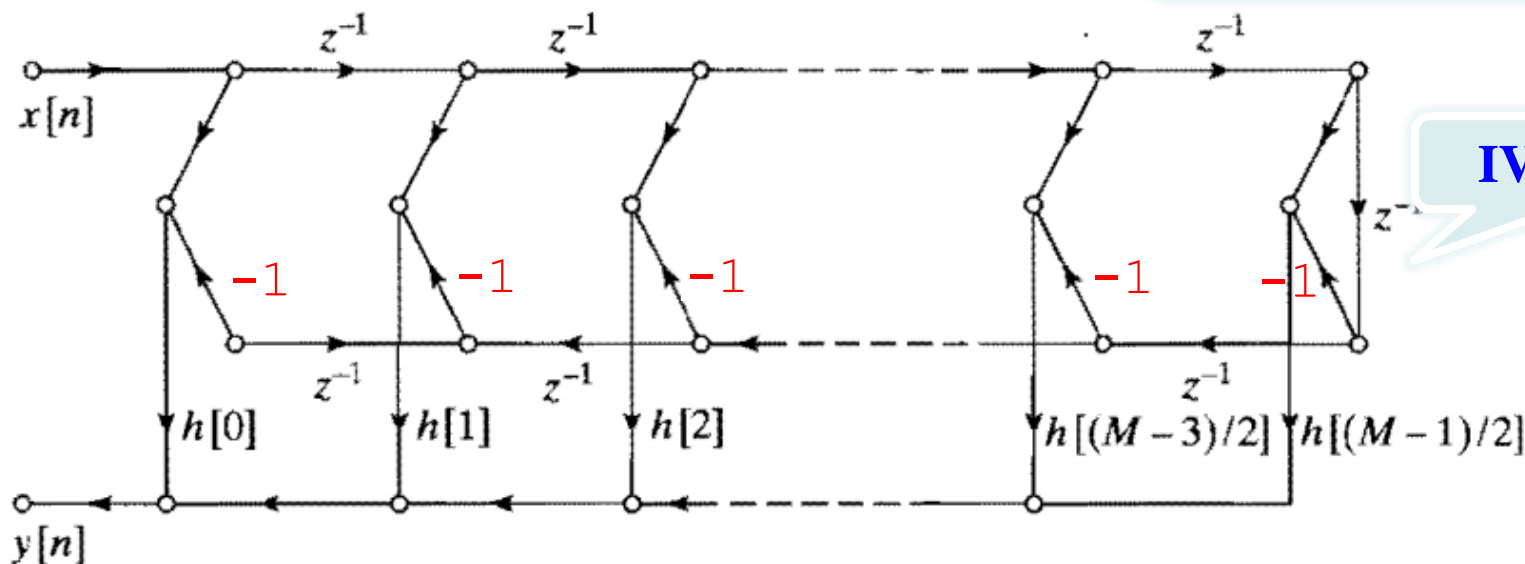
$$y[n] = \sum_{k=0}^{(M-1)/2} h[k] (x[n-k] + x[n-M+k])$$

需 $(M+1)/2$
个系数乘法器

□对于**IV类**FIR线性相位系统（ M 为奇数），利用 $h[n]$ 的**反对称性**

$$y[n] = \sum_{k=0}^{(M-1)/2} h[k] (x[n-k] - x[n-M+k])$$

需 $(M+1)/2$
个系数乘法器



IV类系统

6.5.3 线性相位FIR系统结构

◆由FIR线性相位系统零点获得实现结构分析

由FIR线性相位系统零点特性，系统有4组零

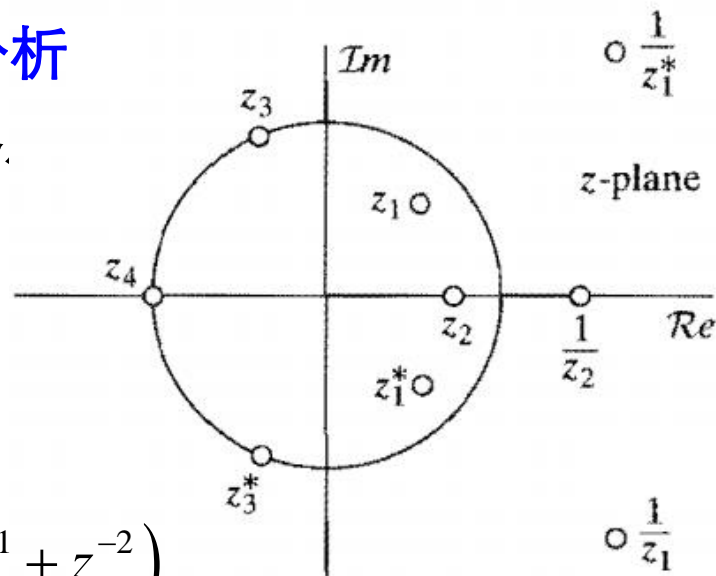
$$(z_4); (z_2 \quad 1/z_2); (z_3 \quad z_3^*);$$

$$(z_1 \quad 1/z_1 \quad z_1^* \quad (z_1^*)^{-1})$$

系统函数可分解成一阶、二阶和四阶
因子的乘积

$$H(z) = h[0] \left(1 + z^{-1}\right) \left(1 + az^{-1} + z^{-2}\right) \left(1 + bz^{-1} + z^{-2}\right) \\ \left(1 + cz^{-1} + dz^{-2} + cz^{-3} + z^{-4}\right)$$

$$\text{式中 } a = (z_2 + 1/z_2), b = 2\operatorname{Re}\{z_3\}, c = -2\operatorname{Re}\{z_1 + 1/z_1\}, d = 2 + |z_1 + 1/z_1|^2$$



系统函数表示为一种线性相位单元级联结构，其多项式的阶数 $M=9$ ；

不同系数乘法器的数目为5，与I类FIR实现结构系数个数 $(M+1)/2$ 相同。

第六章：离散时间系统结构

- ◆ 6.1 线性常系数差分方程的方框图表示
- ◆ 6.2 线性常系数差分方程的信号流图表示
- ◆ 6.3 IIR系统的基本结构
- ◆ 6.4 转置形式
- ◆ 6.5 FIR系统的基本网络结构
- ◆ 6.6 有限精度数值效应概述
- ◆ 6.7 系数量化效应
- ◆ 6.8 数字滤波器中舍入噪声效应
- ◆ 6.9 IIR数字滤波器定点实现中的零输入极限环

6.6.1 数的表示法

一个实数用**无限精度的补码**可表示为

$$x = X_m \left(-b_0 + \sum_{i=1}^{\infty} b_i 2^{-i} \right)$$

X_m 是任意幅度**加权因子**, b_i 取值0或1, b_0 为**符号位**。

采用**B比特量化**可表示为

$$\hat{x} = Q_B[x] = X_m \left(-b_0 + \sum_{i=1}^B b_i 2^{-i} \right) = X_m \hat{x}_B$$

其中 $-X_m \leq \hat{x} \leq X_m$, $\hat{x}_B = -b_0 2^0 + b_1 2^{-1} + \cdots + b_B 2^{-B}$

B比特量化阶 $\Delta = X_m 2^{-B}$

\hat{x} 小数部分 (也称**尾数**) 用定位符号**补码表示**为

$$\hat{x}_B = b_0 \diamond b_1 b_2 \cdots b_B$$

其中 b_0 与 b_1 之间的 \diamond 代表二进制**小数点**

B比特量化误差可表示为 $e = \hat{x} - x = Q_B[x] - x$

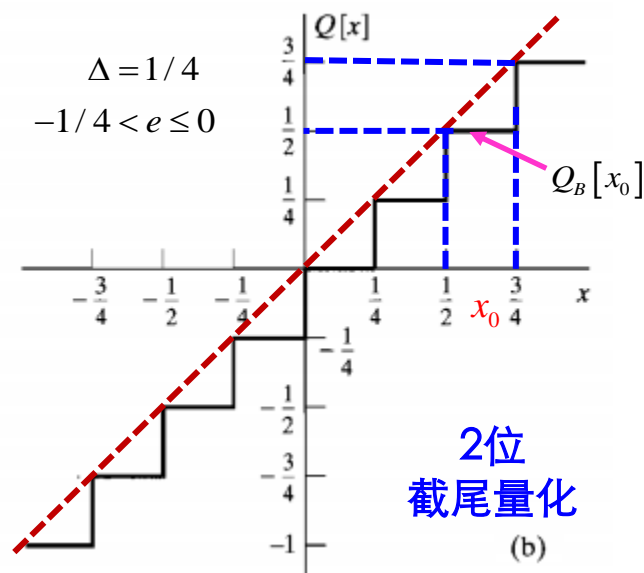
6.6.1 数的表示法

◆ 量化方法

- 截尾量化：保留B位，抛弃余下的尾数

$$\hat{x} = Q_B[x] = X_m \left(-b_0 + \sum_{i=1}^B b_i 2^{-i} \right)$$

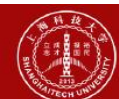
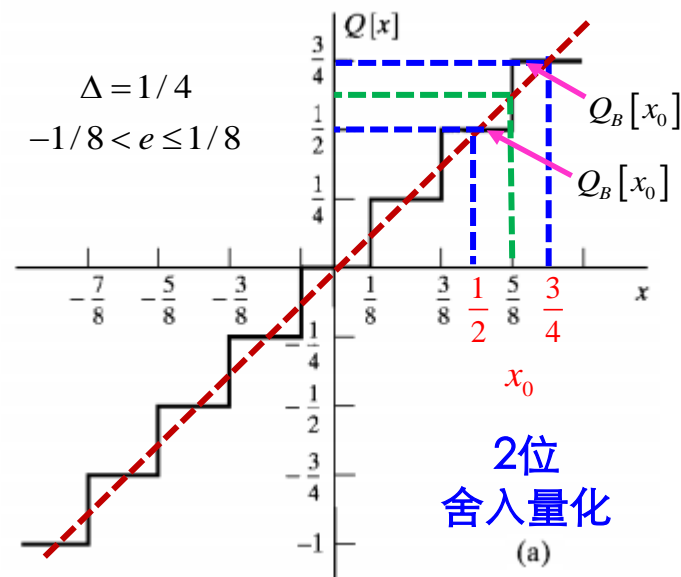
截尾量化误差为 $-\Delta < e \leq 0$



- 舍入量化：按最接近的值取B位码，等效为原数增加 $\Delta/2$ 后按B位截尾

$$\hat{x} = Q_B[x + \Delta/2] = X_m \left(-b_0 + \sum_{i=1}^B b_i 2^{-i} \right)$$

舍入量化误差为 $-\Delta/2 < e \leq \Delta/2$

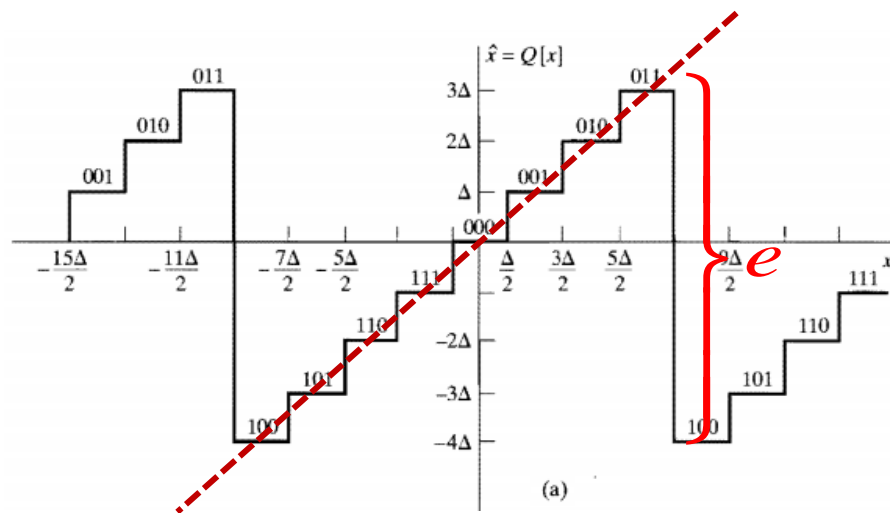


6.6.1 数的表示法

◆ **量化溢出**：量化数值（绝对值）大于量化幅度加权因子 X_m 。

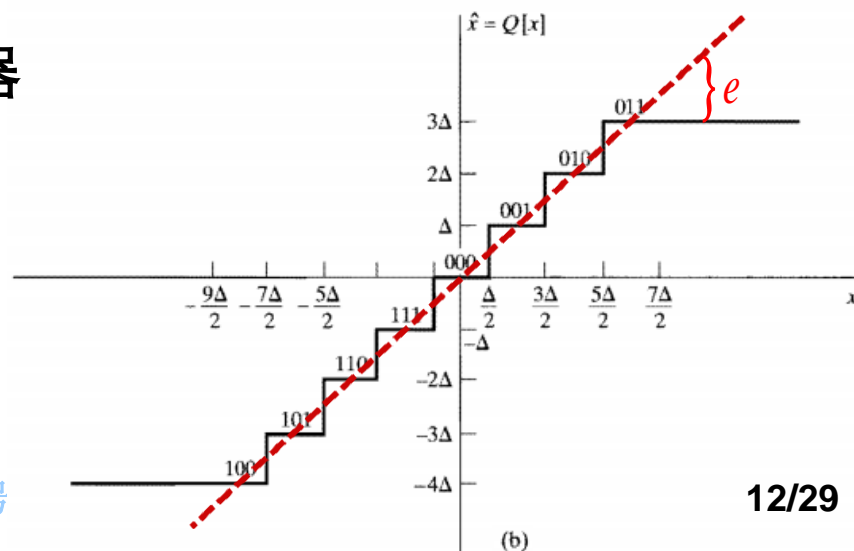
➤ **算术溢出补码舍入量化器**

算术量化溢出导致量化值**正负极性改变**，量化误差突然增加。



➤ **饱和溢出（嵌位）舍入量化器**

饱和量化溢出等效为**限制量化器输出的最大绝对值**，避免了量化误差突然增加。



6.6.1 数的表示法

数的定点表示：小数点 \diamond 位置固定不变，量化范围 $[-1,1)$

$$\hat{x}[n] = X_m \hat{x}_B[n]$$

量化值 $\hat{x}_B[n]$ 的 $B+1$ 比特量化编码的2的补码分数表示：

$$b_0 \diamond b_1 b_2 \cdots b_b \cdots b_B \quad \text{其中, } b_b \text{ 取值为0或1}$$

对应 $\hat{x}_B[n]$ 取值为

$$\hat{x}_B[n] = -b_0 2^0 + b_1 2^{-1} + b_2 2^{-2} \cdots b_B 2^{-B}$$

缺点：在给定位数的条件下较难同时解决大数值的溢出和小数值的量化误差问题。

6.6.1 数的表示法

数的浮点表示：小数点 \diamond 位置可变，量化范围 $[-2^c, 2^c)$

量化值 $\hat{x}_B[n]$ 的 $B+1$ 比特量化编码的2的补码分数表示：

$$b_0 b_1 \cdots b_c \diamond b_{c+1} \cdots b_B \quad \text{其中, } b_b \text{ 为取值为0或1}$$

其中 c 称为阶，为可变整数，指明小数点的位置； b_0 仍为符号位

对应 $\hat{x}_B[n]$ 取值为

$$\hat{x}_B[n] = -b_0 2^c + b_1 2^{c-1} + \cdots + b_c 2^0 + b_{c+1} 2^{-1} + \cdots + b_B 2^{-B+c}$$

例如： $c = 2$ ，量化范围 $[-4, 4)$ ，

$$\hat{x}_B = -b_0 2^2 + b_1 2^1 + b_2 2^0 + b_3 2^{-1} + \cdots + b_B 2^{-B+c}$$

优点： 既保持较大的动态范围，又可控制量化误差。

缺点： 运算处理复杂度较高。

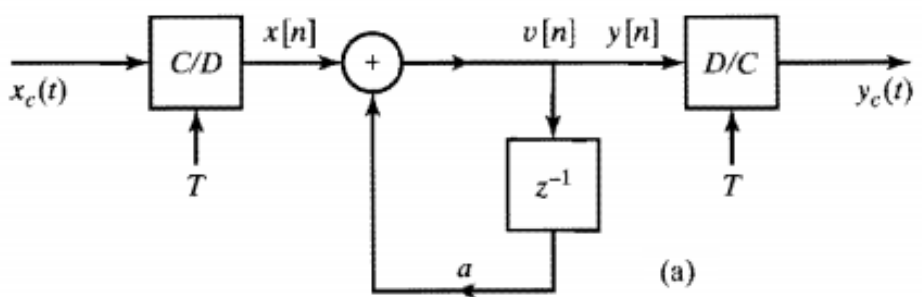
6.6.2 系统实现中的量化

◆连续时间信号离散时间处理系统中的量化

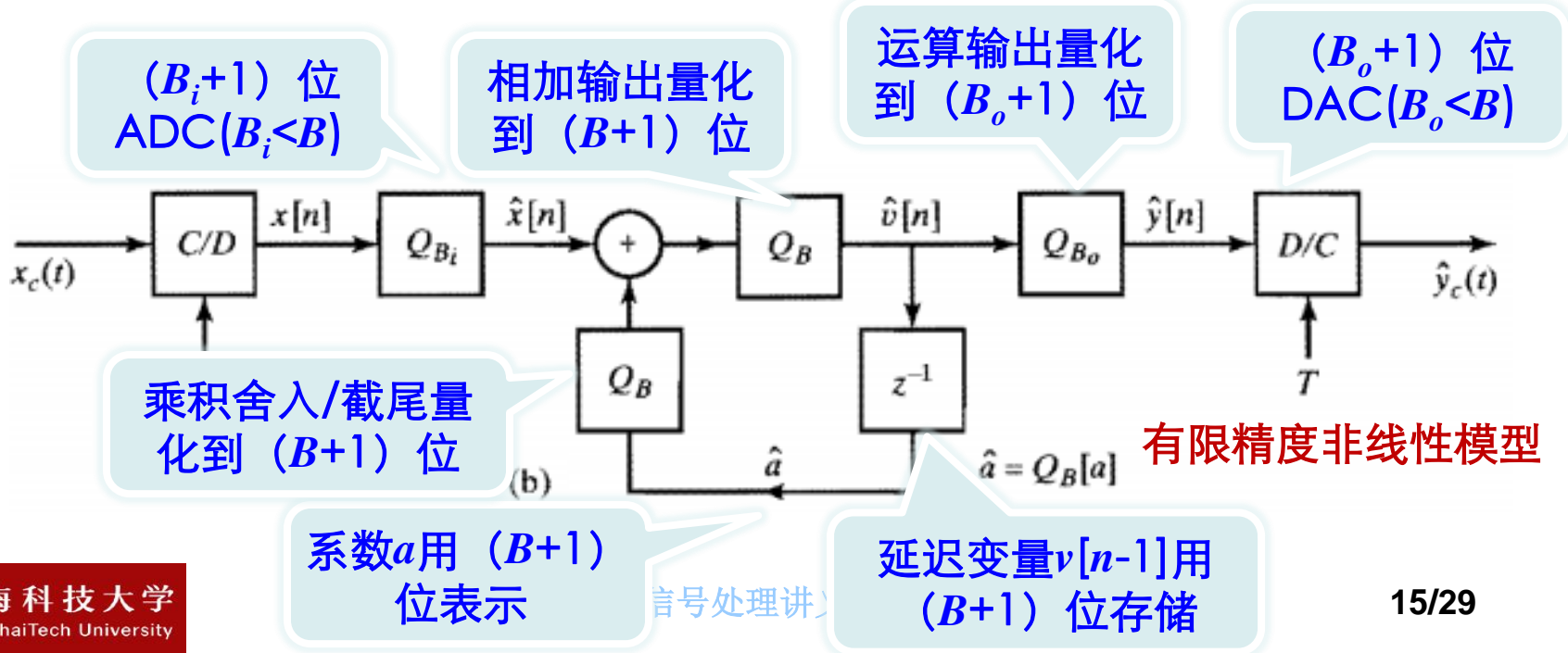
离散时间处理系统的系统函数为

$$H(z) = \frac{1}{1 - az^{-1}}$$

系统采用 $(B+1)$ 位运算实现



无限精度理想系统

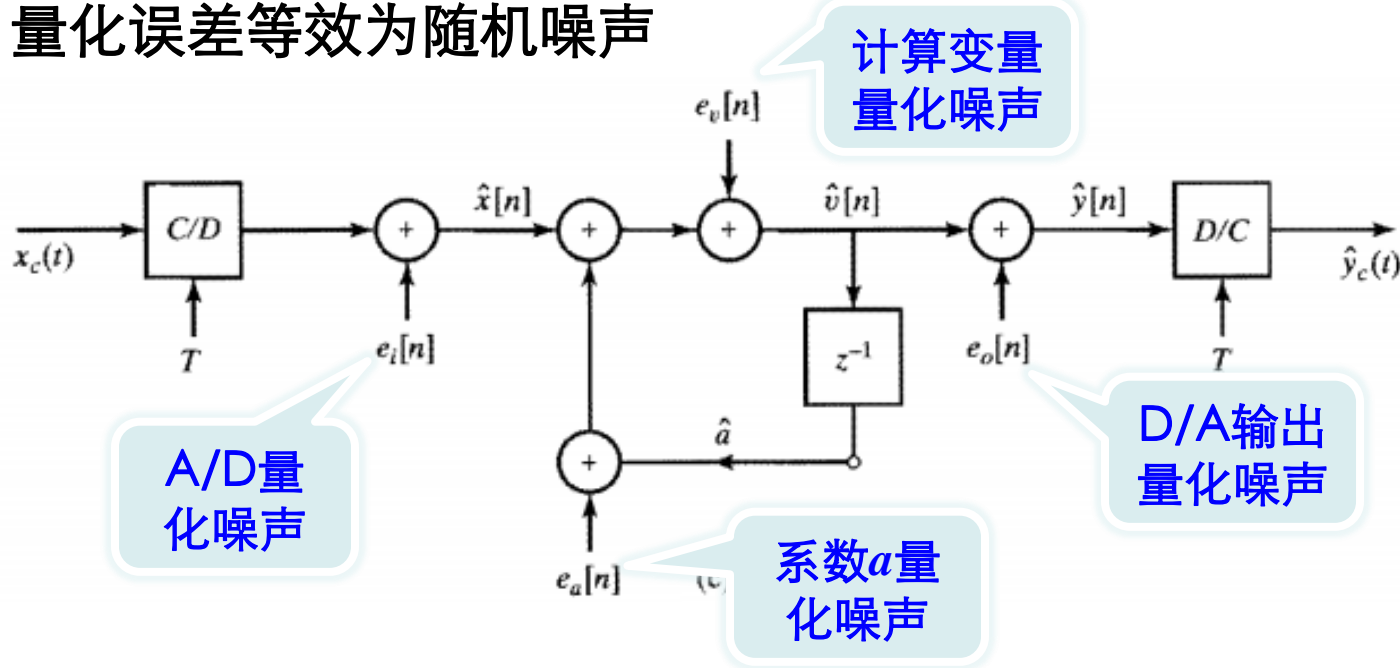


有限精度非线性模型

6.6.2 系统实现中的量化

◆ 系统量化误差线性化模型

- 非线性量化处理等效为线性叠加量化误差
- 量化误差等效为随机噪声



□ 有限精度量化对系统的影响

- 1) 改变系统的响应特性
- 2) 导致零输入极限环现象

第六章：离散时间系统结构

- ◆ 6.1 线性常系数差分方程的方框图表示
- ◆ 6.2 线性常系数差分方程的信号流图表示
- ◆ 6.3 IIR系统的基本结构
- ◆ 6.4 转置形式
- ◆ 6.5 FIR系统的基本网络结构
- ◆ 6.6 有限精度数值效应概述
- ◆ 6.7 系数量化效应
- ◆ 6.8 数字滤波器中舍入噪声效应
- ◆ 6.9 IIR数字滤波器定点实现中的零输入极限环

6.7.1 IIR系统系数量化效应

采用无限精度系数的直接型结构实现的IIR系统函数为

$$H(z) = \sum_{k=0}^M b_k z^{-k} / \left(1 - \sum_{k=1}^N a_k z^{-k} \right)$$

采用量化系数的直接型结构实现的IIR系统函数可表示为

$$\hat{H}(z) = \sum_{k=0}^M \hat{b}_k z^{-k} / \left(1 - \sum_{k=1}^N \hat{a}_k z^{-k} \right)$$

式中 $\hat{a}_k = a_k + \Delta a_k$, $\hat{b}_k = b_k + \Delta b_k$ 为量化系数, Δa_k 和 Δb_k 为量化误差

◆ 两个结论

若极点（或零点）是密集集中在一起，分母（分子）上小的（量化）误差会引起直接型结构极点（零点）有大的偏移

分解为多个小系统的级联型和并联型结构对系数量化的敏感度一般比直接型小得多

6.7.2 椭圆滤波器系数量化影响

◆例：系数量化对IIR椭圆滤波器的影响

带通IIR椭圆滤波器，幅度响应满足

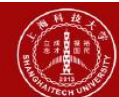
$$\begin{cases} 0.99 \leq |H(e^{j\omega})| \leq 1.01, & 0.3\pi \leq \omega \leq 0.4\pi \\ |H(e^{j\omega})| \leq 0.01(-40dB), & \omega \leq 0.29\pi \\ |H(e^{j\omega})| \leq 0.01(-40dB), & 0.41\pi \leq \omega \leq \pi \end{cases}$$

滤波器采用12阶实现，按级联型实现的系统函数和系数如下

$$H(z) = \prod_{k=1}^6 \frac{b_{0k} + b_{1k}z^{-1} + b_{2k}z^{-2}}{1 - a_{1k}z^{-1} + a_{2k}z^{-2}}$$

表 6.1 十二阶椭圆滤波器未量化的级联型系数

k	a_{1k}	a_{2k}	b_{0k}	b_{1k}	b_{2k}
1	0.738 409	-0.850 835	0.135 843	0.026 265	0.135 843
2	0.960 374	-0.860 000	0.278 901	-0.444 500	0.278 901
3	0.629 449	-0.931 460	0.535 773	-0.249 249	0.535 773
4	1.116 458	-0.940 429	0.697 447	-0.891 543	0.697 447
5	0.605 182	-0.983 693	0.773 093	-0.425 920	0.773 093
6	1.173 078	-0.986 166	0.917 937	-1.122 226	0.917 937



6.7.2 椭圆滤波器系数量化影响

◆ 级联型未量化（32比特）系数和16比特量化系数比较

表 6.1 十二阶椭圆滤波器未量化的级联型系数

k	a_{1k}	a_{2k}	b_{0k}	b_{1k}	b_{2k}
1	0.738 409	-0.850 835	0.135 843	0.026 265	0.135 843
2	0.960 374	-0.860 000	0.278 901	-0.444 500	0.278 901
3	0.629 449	-0.931 460	0.535 773	-0.249 249	0.535 773
4	1.116 458	-0.940 429	0.697 447	-0.891 543	0.697 447
5	0.605 182	-0.983 693	0.773 093	-0.425 920	0.773 093
6	1.173 078	-0.986 166	0.917 937	-1.122 226	0.917 937

0.738403

-0.850830

0.135841

0.026268

0.135841

表 6.2 十二阶椭圆滤波器 16 位量化的级联型系数

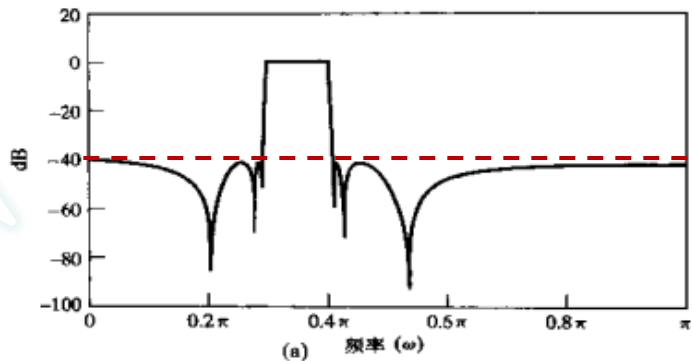
k	a_{1k}	a_{2k}	b_{0k}	b_{1k}	b_{2k}
1	$24\ 196 \times 2^{-15}$	$-27\ 880 \times 2^{-15}$	$17\ 805 \times 2^{-17}$	$3\ 443 \times 2^{-17}$	$17\ 805 \times 2^{-17}$
2	$31\ 470 \times 2^{-15}$	$-28\ 180 \times 2^{-15}$	$18\ 278 \times 2^{-16}$	$-29\ 131 \times 2^{-16}$	$18\ 278 \times 2^{-16}$
3	$20\ 626 \times 2^{-15}$	$-30\ 522 \times 2^{-15}$	$17\ 556 \times 2^{-15}$	$-8\ 167 \times 2^{-15}$	$17\ 556 \times 2^{-15}$
4	$18\ 292 \times 2^{-15}$	$-30\ 816 \times 2^{-15}$	$22\ 854 \times 2^{-15}$	$-29\ 214 \times 2^{-15}$	$22\ 854 \times 2^{-15}$
5	$19\ 831 \times 2^{-15}$	$-32\ 234 \times 2^{-15}$	$25\ 333 \times 2^{-15}$	$-13\ 957 \times 2^{-15}$	$25\ 333 \times 2^{-15}$
6	$19\ 220 \times 2^{-14}$	$-32\ 315 \times 2^{-15}$	$15\ 039 \times 2^{-14}$	$-18\ 387 \times 2^{-14}$	$15\ 039 \times 2^{-14}$



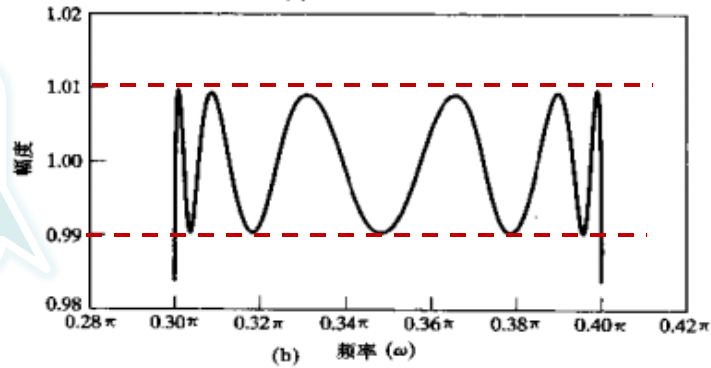
6.7.2 椭圆滤波器系数量化影响

量化系数对不同结构带通IIR椭圆滤波器幅度响应的影响

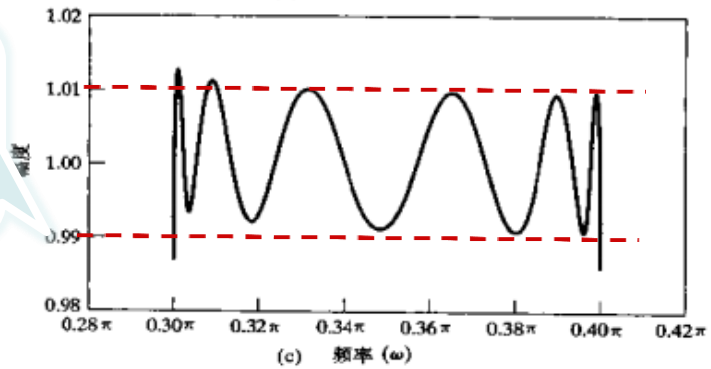
未量化系数



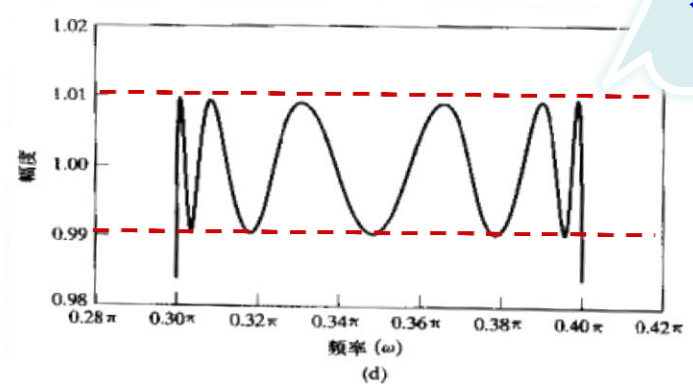
未量化系数级联型通带



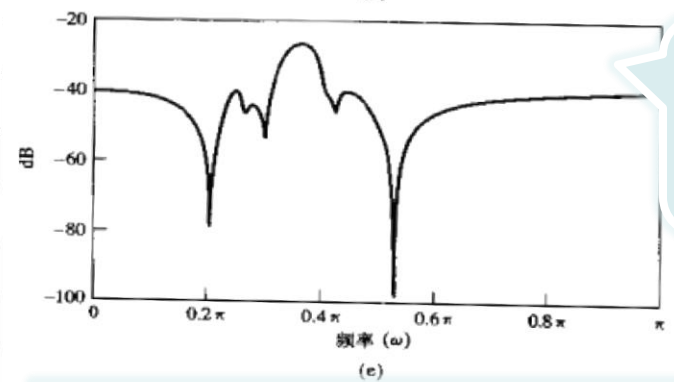
16位量化系数级联型通带



16位量化系数并联型通带



16位量化系数直接型



相同的系数量化位宽下
✓直接型对系数量化误差最敏感
✓并联型的敏感度略好于级联型

6.7.3 量化二阶网段的极点

◆ 系统系数量化对极点影响

➤ 具有复共轭双极点的直接型实现系统

系统函数为

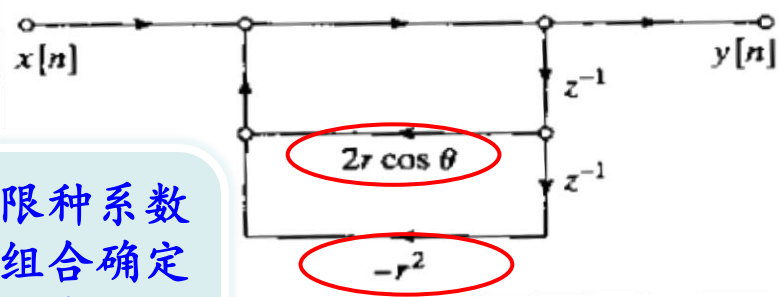
$$H(z) = \frac{1}{1 - (2r \cos \theta) z^{-1} + r^2 z^{-2}}$$

系统双极点为 $z = re^{\pm j\theta}$

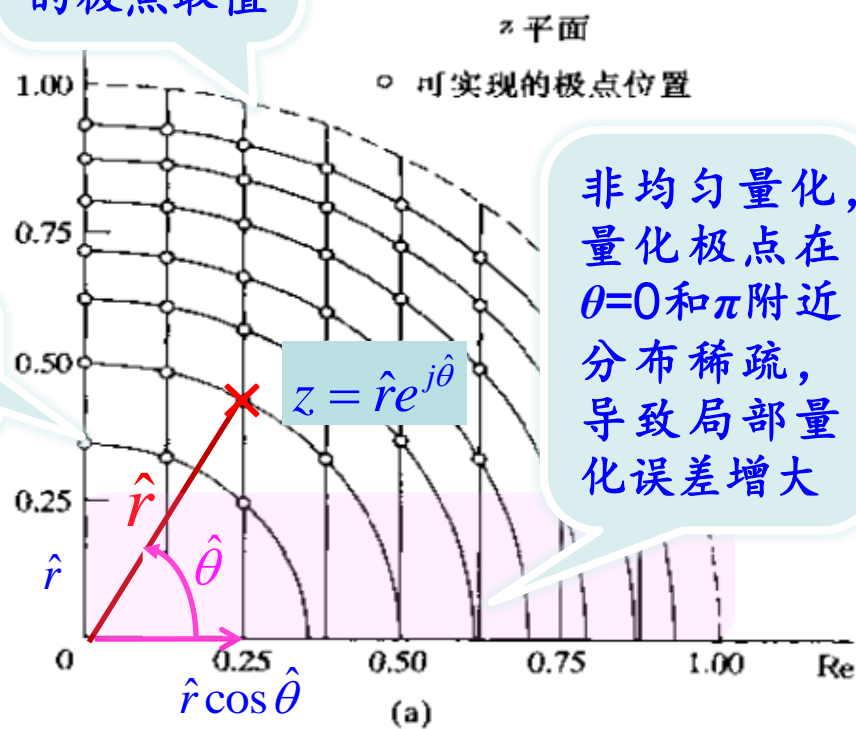
系统系数被量化后将导致极点仅处于z平面有限的几个网格点上

网格点由同心圆（对应于系数 $-r^2$ 的量化）与垂直线（对应于系数 $2r \cos \theta$ 的量化）的交点确定。

例：系数4比特（B=3）量化网格
 r^2 限在7个正值和零上， $2r \cos \theta$ 限在7个正值、零和8个负值上。



有限种系数的组合确定了所有可能的极点取值



z平面
○ 可实现的极点位置

非均匀量化，量化极点在 $\theta=0$ 和 π 附近分布稀疏，导致局部量化误差增大

6.7.3 量化二阶网段的极点

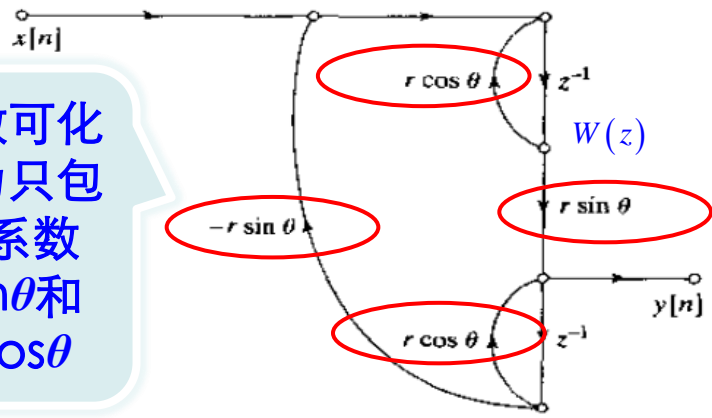
具有复共轭双极点的耦合型实现系统

系统函数为

$$H(z) = \frac{(r \sin \theta) z^{-1}}{1 - (2r \cos \theta) z^{-1} + r^2 z^{-2}}$$

系统仍具有双极点 $z = r e^{\pm j\theta}$

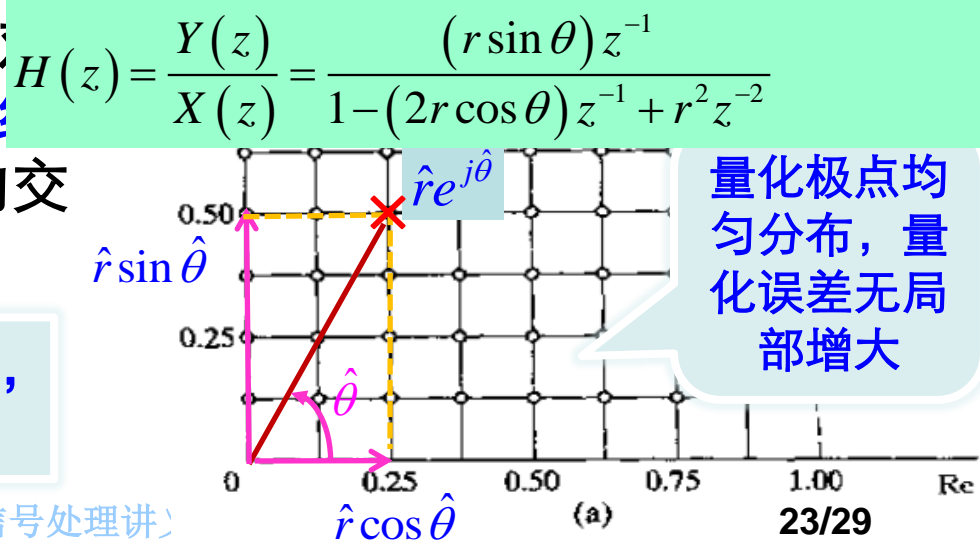
系数可化简为只包含系数 $r \sin \theta$ 和 $r \cos \theta$



量化系数 $r \cos \theta$ 和 $r \sin \theta$ 正好分别为极点 z 的实部和虚部。
 $W(z) = X(z) z^{-1} + W(z) r \cos \theta z^{-1} - r \sin \theta Y(z) z^{-2}$
 $Y(z) = W(z) r \sin \theta + Y(z) z^{-1} r \cos \theta$

量化后的极点位置由水平线（对应系数 $r \sin \theta$ 的量化）与垂直线（对应系数 $r \cos \theta$ 的量化）的交点确定。

缺点：耦合型实现复杂度增加，需要比直接型多2次乘法。



6.7.4 FIR系统系数量化效应

◆FIR系统的系数量化系统模型

FIR系统的系统函数表示为

$$H(z) = \sum_{n=0}^M h[n]z^{-n}$$

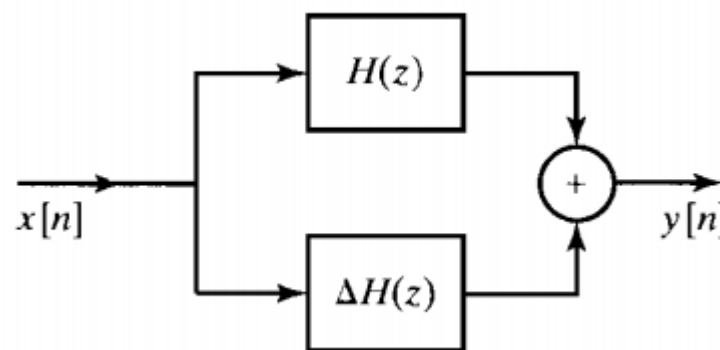
假设系数 $\{h[n]\}$ 被量化, 获得新系数 $\{\hat{h}[n] = h[n] + \Delta h[n]\}$ 。

量化系数系统的系统函数表示为

$$\hat{H}(z) = \sum_{n=0}^M \hat{h}[n]z^{-n} = H(z) + \Delta H(z)$$

式中
$$\Delta H(z) = \sum_{n=0}^M \Delta h[n]z^{-n}$$

为误差系统的系统函数



量化系统可等效为未量化系统与误差系统的并联（相加）

6.7.5 最优FIR滤波器量化示例

◆例：系数量化对线性相位低通FIR滤波器的幅度响应影响

线性相位低通FIR滤波器，幅度响应满足

$$\begin{cases} 0.99 \leq |H(e^{j\omega})| \leq 1.01, & 0 \leq \omega \leq 0.4\pi \\ |H(e^{j\omega})| \leq 0.001(-60dB), & 0.6\pi \leq \omega \leq \pi \end{cases}$$

滤波器采用27阶等波纹最佳一致逼近实现，未量化和量化系数如下

Coefficient	Unquantized	16 bits	14 bits	13 bits	8 bits
$h[0] = h[27]$	1.359657×10^{-3}	45×2^{-15}	11×2^{-13}	6×2^{-12}	0×2^{-7}
$h[1] = h[26]$	-1.616993×10^{-3}	-53×2^{-15}	-13×2^{-13}	-7×2^{-12}	0×2^{-7}
$h[2] = h[25]$	-7.738032×10^{-3}	-254×2^{-15}	-63×2^{-13}	-32×2^{-12}	-1×2^{-7}
$h[3] = h[24]$	-2.686841×10^{-3}	-88×2^{-15}	-22×2^{-13}	-11×2^{-12}	0×2^{-7}
$h[4] = h[23]$	1.255246×10^{-2}	411×2^{-15}	103×2^{-13}	51×2^{-12}	2×2^{-7}
$h[5] = h[22]$	6.591530×10^{-3}	216×2^{-15}	54×2^{-13}	27×2^{-12}	1×2^{-7}
$h[6] = h[21]$	-2.217952×10^{-2}	-727×2^{-15}	-182×2^{-13}	-91×2^{-12}	-3×2^{-7}
$h[7] = h[20]$	-1.524663×10^{-2}	-500×2^{-15}	-125×2^{-13}	-62×2^{-12}	-2×2^{-7}
$h[8] = h[19]$	3.720668×10^{-2}	1219×2^{-15}	305×2^{-13}	152×2^{-12}	5×2^{-7}
$h[9] = h[18]$	3.233332×10^{-2}	1059×2^{-15}	265×2^{-13}	132×2^{-12}	4×2^{-7}
$h[10] = h[17]$	-6.537057×10^{-2}	-2142×2^{-15}	-536×2^{-13}	-268×2^{-12}	-8×2^{-7}
$h[11] = h[16]$	-7.528754×10^{-2}	-2467×2^{-15}	-617×2^{-13}	-308×2^{-12}	-10×2^{-7}
$h[12] = h[15]$	1.560970×10^{-1}	5115×2^{-15}	1279×2^{-13}	639×2^{-12}	20×2^{-7}
$h[13] = h[14]$	4.394094×10^{-1}	14399×2^{-15}	3600×2^{-13}	1800×2^{-12}	56×2^{-7}



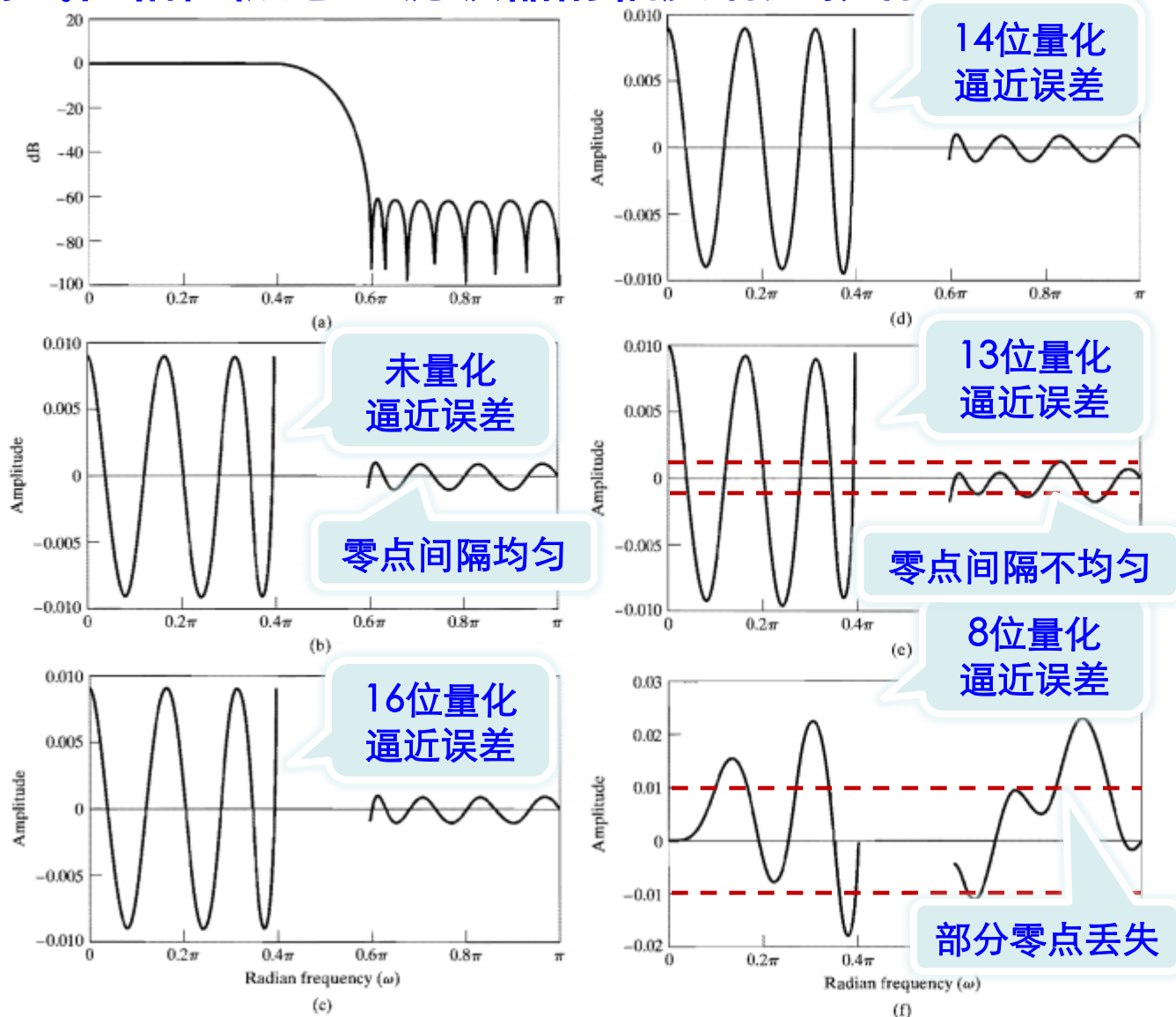
6.7.5 最优FIR滤波器量化示例

➤不同量化系数对线性相位低通FIR滤波器的幅度响应影响

□逼近误差:

✓通带逼近于1
的误差

✓阻带逼近于0
的误差

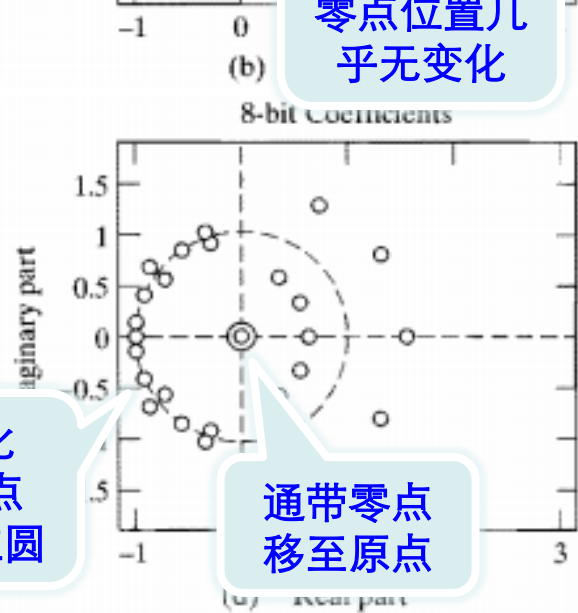
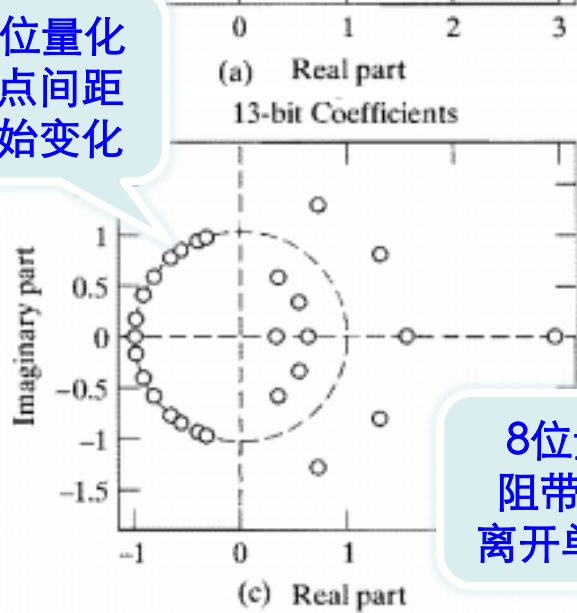
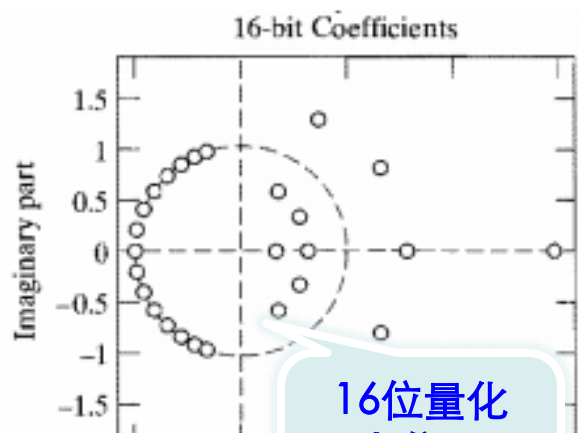
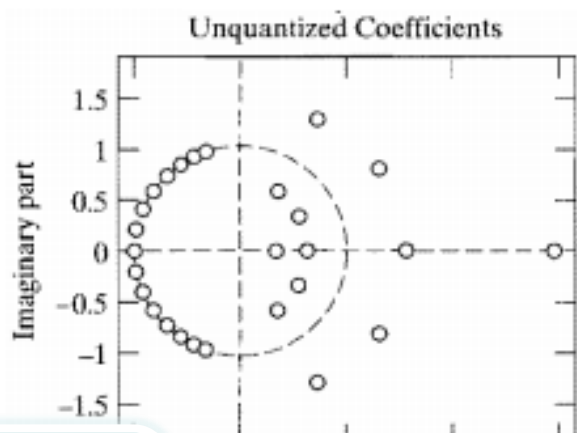
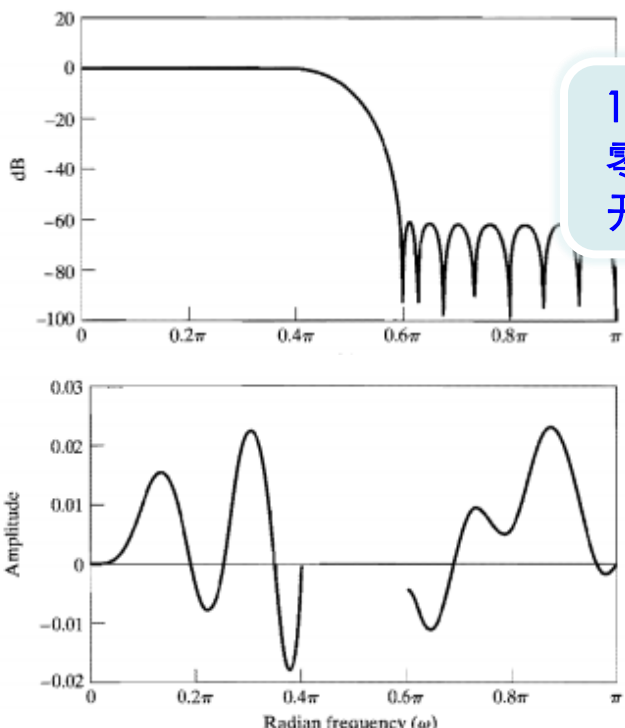


6.7.5 最优FIR滤波器量化示例

➤不同量化系数对线性相位低通FIR滤波器的零点影响

不在单位圆的零点
影响通带特性

在单位圆上的零点
影响阻带特性



13位量化
零点间距
开始变化

16位量化
零点位置
几乎无变化

8位量化
阻带零点
离开单位圆

通带零点
移至原点

6.7.6 线性相位特性的保持

◆ 系数量化误差对线性相位的影响

- 对于线性相位FIR系统，系数量化后，其单位脉冲响应仍然满足对称或反对称性，即系数量化直接型系统仍易保持广义线性相位。
- 由包含少量零点的低阶网段级联构成的高阶FIR系统，要保持线性相位，必须级联的每个低阶网段具有线性相位。

线性相位系统的零点必须满足单位圆上共轭，单位圆内/外互为共轭、倒数等特性，设计网段结构时必须使其零点具备这些特点。

例：用 $(1 + az^{-1} + z^{-2})$ 二阶网段实现单位圆上和圆内零点

当系数 a 被量化时，系统在单位圆上的互为共轭的2个复零点只会在单位圆上移动；

或者系统在单位圆内和单位圆外互为倒数的2个实零点也会保持为实的。