

⑤ 快速指令周期：显然，指令周期越短，实时处理能力越强。TMS320C30 的指令周期仅为 60ns~75ns，而 TMS320C64 的指令周期仅为 0.91ns~1.66ns。

### 一、FFT 的软件实现

下面，我们以按时间抽选的基-2 FFT 算法为例，讨论离散傅里叶变换的 FFT 算法的计算机软件实现。这里采用输入倒位序、输出自然顺序的流程图，见图 4-5 ( $N=8$ )。

FFT 程序包括变址(倒位序)和  $L$  级递推计算( $N=2^L$ ,  $L$  为正整数)两大部分，下面分别加以讨论。

#### 1. 变址

FFT 变址(倒位序)流程图如图 4-32 所示。倒位序的规律在 4.3 节已经讨论过了。其中表 4-2 以  $N=8$  为例列出了自然顺序数与其倒位序数的关系，图 4-9 表示了倒位序的变址处理，设  $A(I)$  表示存放原自然顺序输入数据的内存单元， $A(J)$  表示存放倒位序数的内存单元， $I, J=0, 1, \dots, N-1$ 。按倒位序规律，当  $I=J$  时，不用变址；当  $I \neq J$  时，需要变址；但是当  $I < J$  时，进行变址在先，故在  $I > J$  时，就不需要再变址了，否则变址两次等于不变址。下面来讨论实现倒位序的一种方法——雷德(Rader)算法。

由表 4-2 可见，按自然顺序排列的二进制数，其下一个数总是比其上面一个数大 1，即下一个数是上面一个数在最低位加 1 并向高位进位而得到的。而倒位序二进制数的下一个数是上面一个数在最高位加 1 并由高位向低位进位而得到的。下面讨论如何实现这一“反向进位加法”。

$I, J$  都是从 0 开始。若已知某个倒位序数  $J$ ，要求下一个倒位序数，则应先判断  $J$  的最高位是否为 0，这可与  $k=N/2$  相比较，因为  $N/2$  总是等于 100…的。如果  $k > J$ ，则  $J$  的最高位为 0，只要将该位变为 1( $J$  与  $k-N/2$  相加即可)，就得到了下一个倒位序数；如果  $k \leq J$ ，则  $J$  的最高位为 1，可将最高位变为 0( $J$  减去  $N/2$  即可)。然后还需判断次高位，这可与  $k=N/4$  相比较，若次高位是 0，则将它改成 1( $J$  与  $N/4$  相加即可)，其他位不变，即得到下一个倒位序数；若次高位是 1，则需将它也变为 0( $J$  减去  $N/4$  即可)。然后还需判断再下一位，这可与  $k=N/8$  相比较……依次进行，总会碰到某位为 0(除非最后一个数  $N-1$  的各位全是 1，而这个数不需倒位序，不属于倒位序程序内的数)，这时把这个 0 改成 1，就得到下一个倒位序数。求出新的倒位序数  $J$  以后，当  $I < J$  时，进行变址交换。注意，在倒位序中， $x(0)$  和  $x(N-1)$  总是不需要交换的，因为 0 与  $N-1$  的倒位序数与原自然顺序数是一样的。

#### 2. $L$ 级递推计算

在表 4-3 中已将四种原位运算 FFT 的特点作了归纳。其中蝶形结对偶节点的距离与蝶形结的计算公式是最重要的，是不管哪种实现方法都要遵循的。但是， $W_N^k$  的求法则可有各不相同的方法，表 4-3 给出的求  $W_N^k$  的方法只是其中的一种。在下面讨论的  $L$  级递推计算方法中，求  $W_N^k$  就没有采用表 4-3 给出的方法。

由图 4-5 可以归纳出输入倒位序、输出自然顺序的按时间抽选的 FFT 流程图的一些特点。设  $N=2^L$ ，则有① 每级有  $N/2$  个蝶形结，第一级(列)的  $N/2$  个蝶形结都是同一种



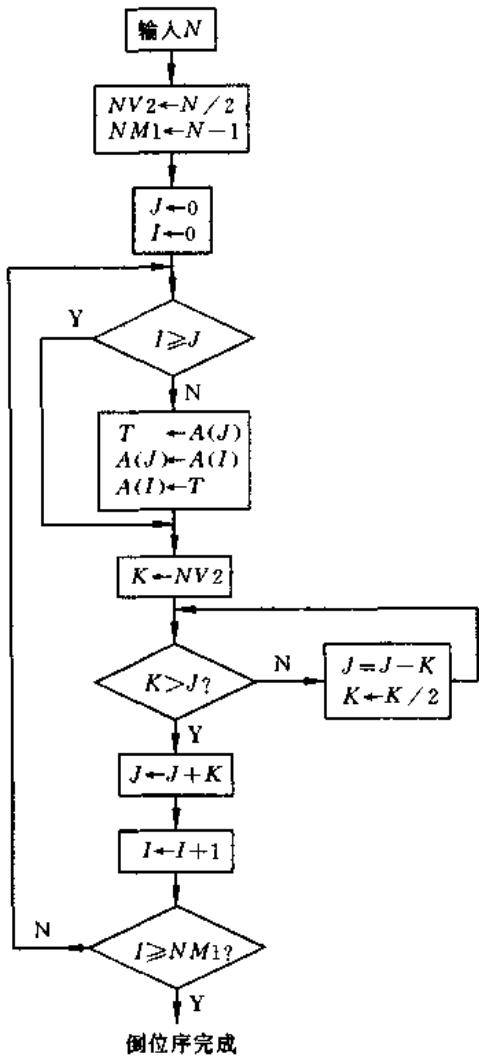


图 4-32 FFT 数据倒位序流程图

蝶形运算,也就是说,其系数都相同,为  $W_N^0$ 。第二级的  $N/2$  个蝶形结共有两种蝶形运算,一种系数为  $W_N^0$ ,一种系数为  $W_N^2$ ,每种各有  $N/4$  个蝶形结。这样可看出,第  $L$  级的  $N/2$  个蝶形结共有  $2^{L-1} = N/2$  种蝶形运算,即有  $N/2$  个不同的系数,分别为  $W_N^0, W_N^1, \dots, W_N^{\frac{N}{2}-1}$ ,也就是说,第  $M$  级有  $2^{M-1}$  种蝶形运算 ( $M=1, 2, \dots, L$ )。② 由最后一级向前每推进一级,则系数取后级系数中偶数序号那一半(例如在图 4-5 中,第三级系数为  $W_N^0, W_N^1, W_N^2, W_N^3$ ,而第二级系数为  $W_N^0, W_N^2$ )。③ 蝶形结两个节点的间距为  $2^{M-1}$ , $M$  为所在的级数,也就是每向前推进一级,间距就变成原间距的  $1/2$ 。

图 4-33 为按时间抽选 FFT 流程图,其中整个  $L$  级递推过程由三个嵌套循环构成。外层的一个循环控制  $L$  ( $L=\log_2 N$ ) 级的顺序运算;内层的两个循环控制同一级 ( $M$  相同) 各蝶形结的运算,其中最内一层循环控制同一种(即  $W_N^r$  中的  $r$  相同)蝶形结的运算,而中间一层循环则控制不同种(即  $W_N^r$  中的  $r$  不同)蝶形结的运算。 $I$  和  $IP$  是一个蝶形结的两个节点。

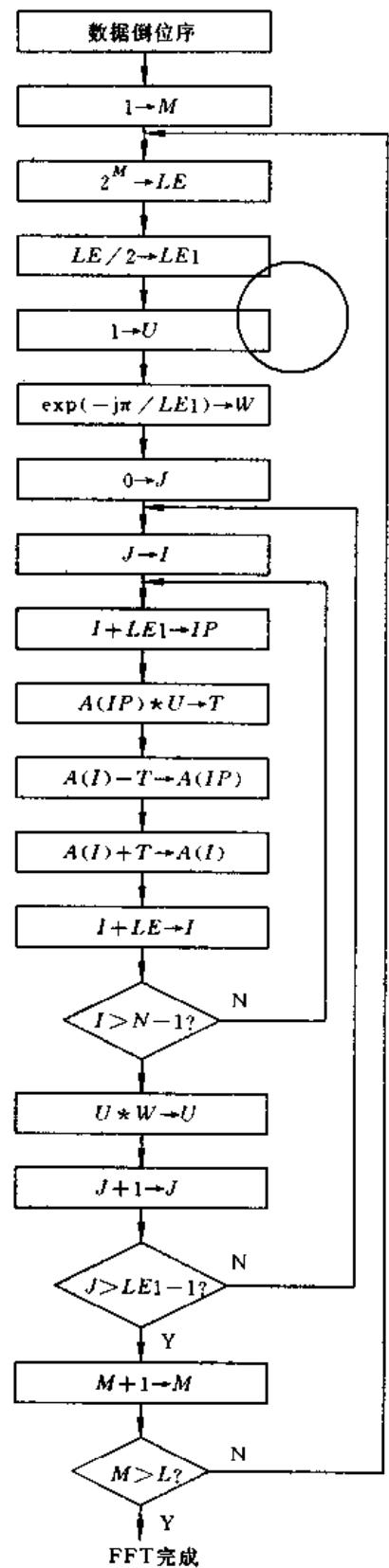


图 4-33 基-2 按时间抽选 FFT 流程图

设  $N=2^L$ ,  $M$  表示第  $M$  级,  $M=1, 2, \dots, L$ , 共有  $L=\log_2 N$  级, 每级有  $N/2$  个蝶形结, 在第  $M$  级中有  $2^{M-1}$  个不同的系数, 它们是  $W_{2^M}^k$  ( $k=0, 1, \dots, 2^{M-1}-1$ ), 例如, 在图 4-5 中, 当  $M=2$  (第二级) 时, 系数有  $2^{M-1}=2$  个, 分别是  $W_4^0=W_8^0=W_N^0$  与  $W_4^1=W_8^1=W_N^1$ 。在图 4-33 的  $U$  中放有  $W_{2^M}^k$ , 其起始值为  $U=W_{2^M}^0=1$ 。

在每一级中, 同一种系数对应的蝶形结有  $2^{L-M}$  个 (例如, 第一级  $M=1$ , 系数都是  $W_N^0$ , 对应的蝶形结有  $2^{3-1}=4$  个)。各蝶形结依次相距  $LE=2^M$  点 (例如, 第二级  $M=2$ , 蝶形结依次相距  $LE=4$  点)。图 4-33 的最内层循环, 其循环变量为  $I$ ,  $I$  用来控制同一种蝶形结运算。显然, 其步进值为蝶形结间距值  $LE=2^M$ , 同一种蝶形结中参加运算的两节点的间距为  $LE1=2^{M-1}$  点。图 4-33 的第二层循环, 其循环变量  $J$  用来控制计算不同种 (系数不同) 的蝶形结,  $J$  的步进值为 1。实际上也可看出, 最内层循环完成每级的蝶形结运算, 第二层 (中间层) 循环则完成系数  $W_N^k$  的运算。图 4-33 的最外层循环, 用循环变量  $M$  来控制运算的级数,  $M$  为 1 到  $L$ , 步进值为 1。当  $M$  改变时, 则  $LE1$ ,  $LE$  和系数  $U$  都会改变。

图 4-33 中的  $U, W, J$  为存放复数单元, 其乘法为复数乘法, 系数采用以下递推公式计算:

$$W_N^k = W_N^{(k-1)l} \cdot W_N^l$$

图 4-34 写出了按时间抽选、输入倒位序、输出自然顺序、 $N=2^L$  的基-2 FFT 算法的 C++ 程序。

下面我们针对表 4-3, 给出另一种按时间抽选、输入自然顺序、输出倒位序的基-2 FFT 流程图, 图 4-35, 读者可结合表 4-3 对此流程图加以分析。

## 二、FFT 的专用硬件实现

采用 FFT 专用硬件可满足一些实时运算的要求。例如, 雷达及声呐信号的处理与分析就要求实时完成。

几种基-2 FFT 硬件实现形式

1. 顺序处理

按时间抽取基-2 FFT 流图的蝶形结运算如(4-22)式所示, 可把它简写为

$$C = A + BW_N^r$$

$$D = A - BW_N^r$$

这一运算表示在图 4-36 的左上角 (其中  $r=0$ )。

顺序处理是按图 4-36 所标的次序来执行蝶形结运算的, 即先算第(1)个蝶形结, 再算第(2)个, 第(3)个……蝶形结, 直至第  $\frac{N}{2} \log_2 N$  个蝶形结 (对于图 4-36,  $N=8$ ,  $\frac{N}{2} \log_2 N=12$ ), 就完成了  $N$  点 FFT 的全部运算。由于输入数据是按自然顺序排列的, 如果不做处理, 则输出数据是倒位序排列的。这种顺序处理 FFT 的特点是:

- (1) 只用一个运算单元;
- (2) 输入量、中间量、输出量均使用同一存储器;

```

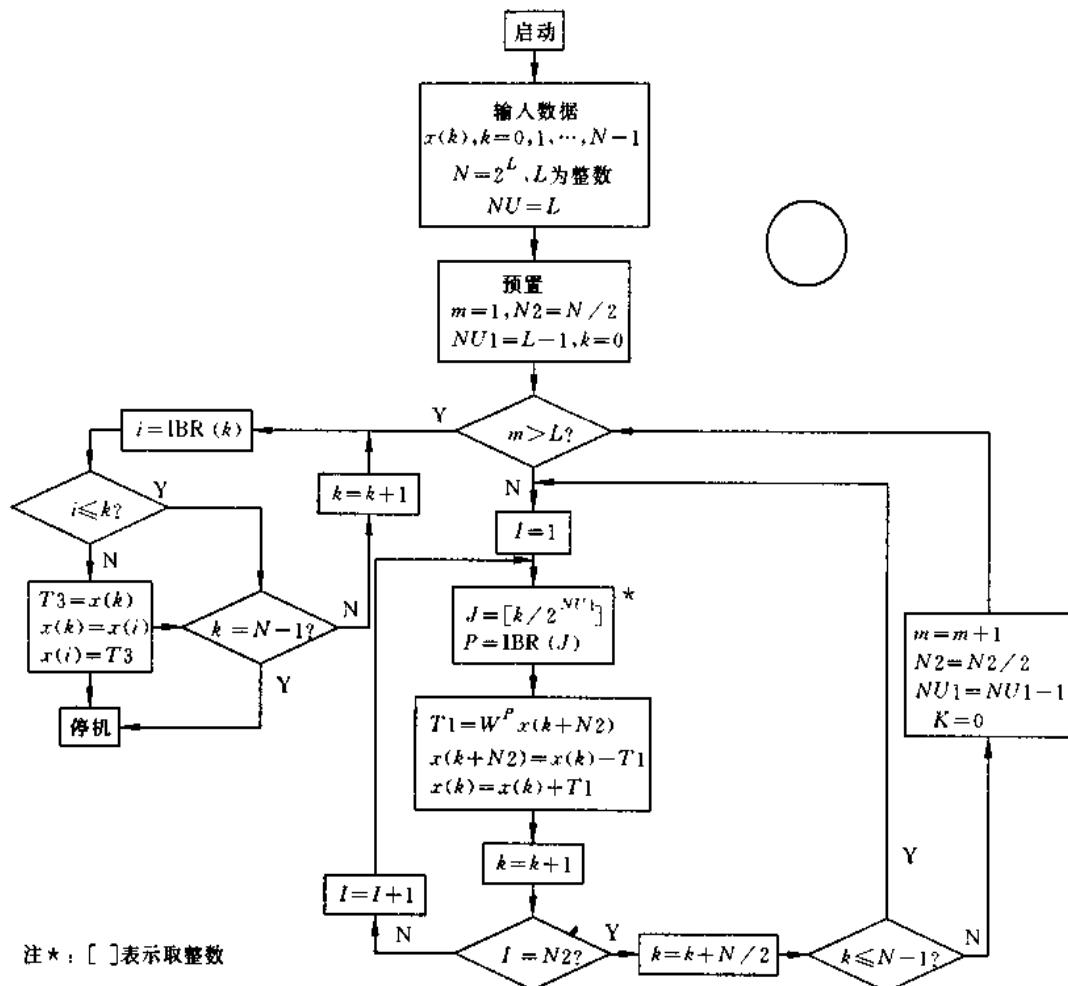
#include<complex.h>

int fft(complex * a,int l)
{
    const double pai=3.141592653589793;
    complex u,w,t;
    unsigned n=l,nv2,nml,k,le,lei,ip;
    unsigned i,j,m;
    double tmp;

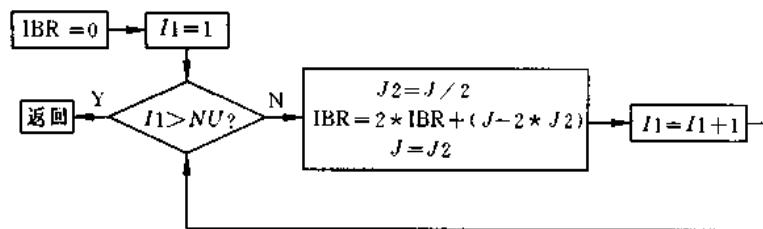
    n<<=1;
    nv2=n>>1;
    nml=n-1;
    i=0;
    for(i=0;i<nml;i++)
    {
        if(i<j)
        {
            t=a[j];
            a[j]=a[i];
            a[i]=t;
        }
        k=nv2;
        while(k<=j)
        {
            j-=k;
            k>>=1
        }
        j+=k
    }
    le=1
    for(m=1;m<=l;m++)
    {
        lei=le
        le<<=1;
        u=complex(1,0);
        tmp=pai/lei;
        w=complex(cos(tmp),-sin(tmp));
        for(j=0;j<lei;j++)
        {
            for(i=j;i<n;i+=lei)
            {
                ip=i+lei;
                t=a[ip]*u;
                a[ip]=a[i]-t;
                a[i]+=t;
            }
            u*=w;
        }
    }
    return 0;
}

```

图 4-34 基-2 FFT 算法的 C++ 程序  
(按时间抽选、输入倒位序、输出自然顺序,  $N=2^l$ )



(a) FFT流程图



(b) 倒位序流程图

图 4-35 按时间抽选、输入自然顺序、输出倒位序 FFT 流程图(见表 4-3)

(3) 顺序执行  $\frac{N}{2} \log_2 N$  次蝶形结运算, 如果一次蝶形结运算时间为  $T_B \mu s$ , 则总的运算时间为  $\frac{T_B N}{2} \log_2 N \mu s$ .

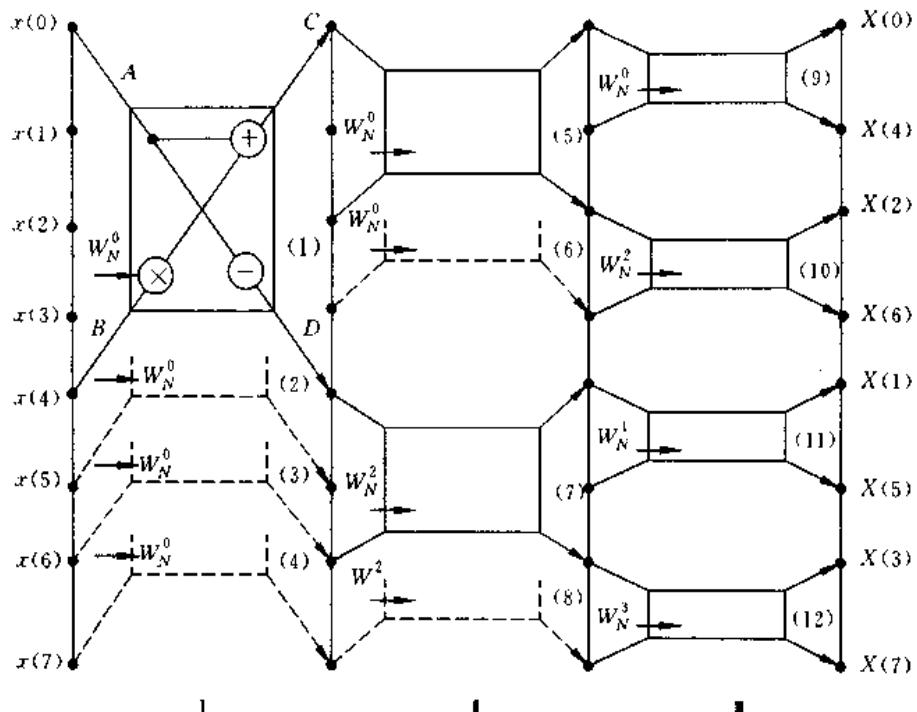


图 4-36  $N=8$  顺序处理的 FFT 流程图

### 2. 级联处理

若  $N$  点 FFT 的每一级(列)的  $N/2$  个蝶形结用一个独立运算单元来处理。就构成级联处理,以图 4-36 为例,第一个运算单元处理第一级(1)到(4)四个蝶形结,第二个运算单元处理第二级(5)到(8)四个蝶形结,第三个运算单元处理第三级(9)到(12)四个蝶形结。这样,数据的流量可以增加到原流量的  $L = \log_2 N$ (即为所需(列)级数)倍。级联处理 FFT 的特点是:

- (1) 用  $L = \log_2 N$  个运算单元并行运算;
- (2) 每个运算单元顺序执行  $N/2$  次蝶形结运算;
- (3) 每级数据执行运算时间为  $\frac{T_B N}{2} \mu\text{s}$ .

以  $N=8$  的 FFT 的为例,由于三个运算单元并行运算,由图 4-36 可看出,只有在第一个运算单元算完了三对数据以后,第二个运算单元才能开始运算,而只有在第二个运算单元算完了两对数据后,第三个运算单元才能开始运算,因而完成 FFT 所需时间,还应包括上述的等待时间,从而每个运算单元要含有延时用的缓冲存储。

### 3. 并行迭代处理

并行迭代处理是对一级中的  $N/2$  个蝶形结用  $N/2$  个运算单元并行运算。对于  $N=8$ ,采用四个运算单元,先同时算第一级的  $N/2=4$  个蝶形结,再依次算第二级和第三级的 4 个蝶形结,也就是说,每列中的运算是并行的,一列到另一列的迭代是顺序的,故名并行迭代运算。并行迭代处理 FFT 的特点是:

- (1) 有  $N/2$  个处理单元;
- (2) 并行执行  $N/2$  个蝶形结运算;

- (3) 顺序执行  $L = \log_2 N$  次迭代；  
 (4) 执行 FFT 的总时间为  $T_B \log_2 N \mu s$ 。

这种方法处理速度快,但设备量大,要求采用大规模集成电路。

#### 4. 阵列处理

阵列处理是并行实现全部蝶形结运算。例如  $N=8$ ,可并行运算 12 个蝶形结,此时,处理机输入端可以有三个不同的待处理的数据相继通过此处理机。阵列处理的特点是：

- (1) 有  $\frac{N}{2} \log_2 N$  个处理单元；  
 (2) 并行执行  $\frac{N}{2} \log_2 N$  个蝶形结运算；

(3) 执行运算的时间是  $T_B \mu s$ ,但是由于 FFT 各列蝶形结运算的先后次序决定了总运算时间仍为  $T_B \log_2 N \mu s$ ,与并行迭代处理时间相同；

(4) 单位时间内所能处理的数据组数比并行迭代处理时要大  $\log_2 N$  倍,也就是说,数据输入的流通量增大  $\log_2 N$  倍。

这种处理方法处理容量最大,但是设备量也最大,因而成本费用高。

下面讨论两种具体处理机——FFT 顺序处理机和基-2 流水线方式 FFT 处理机。

#### FFT 顺序处理机

按时间抽选(DIT)和按频率抽选(DIF)的 FFT 的蝶形结分别表示在(4-22)式和(4-28)式中,可将这两个公式分别简化成以下表示式：

$$\left. \begin{array}{l} C = A + W_N^k B \\ D = A - W_N^k B \end{array} \right\} \text{DIT} \quad (4-117)$$

$$\left. \begin{array}{l} C = A + B \\ D = (A - B)W_N^k \end{array} \right\} \text{DIF} \quad (4-118)$$

其中  $A, B$  表示输入的复数数据, $C, D$  表示蝶形结的输出结果, $W_N^k$  表示系数。如果将各复数用实部、虚部表示,则为

$$\begin{aligned} A &= a_r + j a_i & B &= b_r + j b_i \\ C &= c_r + j c_i & D &= d_r + j d_i \\ W_N^k &= e^{-j \frac{2\pi k}{N}} = \cos\left(\frac{2\pi k}{N}\right) - j \sin\left(\frac{2\pi k}{N}\right) \end{aligned}$$

则有

$$\left. \begin{array}{l} c_r = a_r + \left[ b_r \cos\left(\frac{2\pi k}{N}\right) + b_i \sin\left(\frac{2\pi k}{N}\right) \right] \\ d_r = a_r - \left[ b_r \cos\left(\frac{2\pi k}{N}\right) + b_i \sin\left(\frac{2\pi k}{N}\right) \right] \\ c_i = a_i - \left[ b_r \sin\left(\frac{2\pi k}{N}\right) - b_i \cos\left(\frac{2\pi k}{N}\right) \right] \\ d_i = a_i + \left[ b_r \sin\left(\frac{2\pi k}{N}\right) - b_i \cos\left(\frac{2\pi k}{N}\right) \right] \end{array} \right\} \text{DIT} \quad (4-119)$$

$$\left. \begin{aligned} c_r &= a_r + b_r \\ c_i &= a_i + b_i \\ d_r &= (a_r - b_r) \cos\left(\frac{2\pi k}{N}\right) + (a_i - b_i) \sin\left(\frac{2\pi k}{N}\right) \\ d_i &= (a_i - b_i) \cos\left(\frac{2\pi k}{N}\right) - (a_r - b_r) \sin\left(\frac{2\pi k}{N}\right) \end{aligned} \right\} \text{DIF} \quad (4-120)$$

由(4-119)式、(4-120)式看出,不管是 DIT 算法或 DIF 算法,都必须执行 4 次实数乘法及 6 次实数加(减)法。对于 DIT 算法,其运算的实现框图如图 4-37 所示,它的运算器包括一个复加(减)器和一个复乘器。

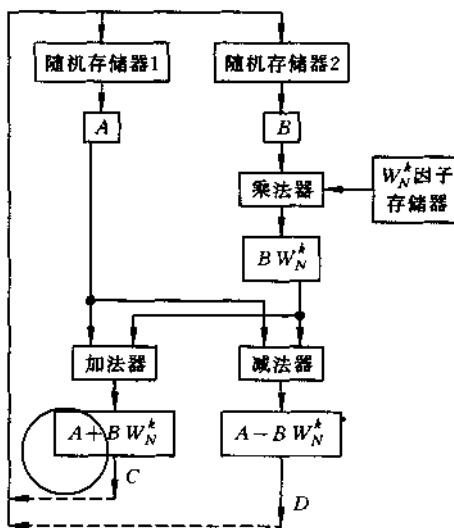


图 4-37 DIT 的 FFT 运算实现框图

当然,只有运算单元是不够的,构成 FFT 处理机的基本部分应包括运算单元、控制单元和存储单元,包括存储系数  $W_N^k$  的只读存储器(ROM)和存放数据和计算结果的随机存储器(RAM)。图 4-38 为 FFT 顺序处理机的基本形式,其他的高级 FFT 处理机只是在部件和程序上进行了某些改变,本质并没有变化。

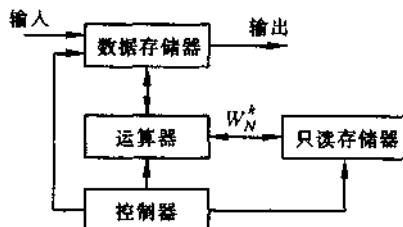


图 4-38 顺序处理机原理框图

由上面指出的 FFT 的一个蝶形结运算需 4 次实乘、6 次实加,可估算出执行一次蝶形运算所需时间。假定存储字只包含实部(或虚部),运算单元中只有单一的实数乘法器

和实数加法器,则读出  $a_r, a_i, b_r, b_i, \cos\left(\frac{2\pi k}{N}\right)$  及  $\sin\left(\frac{2\pi k}{N}\right)$  需要 6 个时钟节拍,写回  $c_r, c_i, d_r, d_i$  需要 4 个节拍,6 次加法需 6 拍,进行一次乘法的时间与乘法器的结构有关,若假定为 3 拍,则 4 次乘法共需 12 拍,于是进行一次蝶形结运算的时间  $T_B = 28$  个节拍,如果每个节拍为 100ns,则完成一次蝶形结运算的时间为  $T_B = 2.8\mu s$ ,如果  $N=1024$ ,则完成  $N$  点 FFT 所需的总运算时间为  $\frac{1}{2} T_B N \log_2 N = 2.8\mu s \times 5120 = 14.336\mu s$ 。

顺序处理方法简单,但处理速度低。为了提高速度,除了提高逻辑部件速度外,还可从处理结构上想些办法,例如:

(1) 使运算单元和存储器在时间安排上处于“匹配”状态,使运算器和存储器都不断地工作。图 4-39 所示为一种运算时间和读写时间的最佳匹配关系。

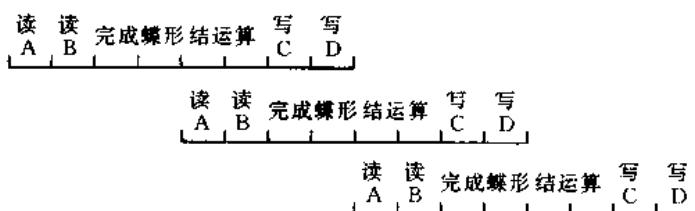


图 4-39 “最佳匹配”基-2 FFT 的读、写及蝶形结运算的计算机时序图

可以看出,当蝶形结运算进行到一半时,下一个蝶形结运算的输入数据被送入到运算单元,这要求运算单元要有缓冲存储器,一旦一次蝶形结运算结束后,立即把结果送回存储器,运算单元则不停地进行下一个蝶形结运算。由此看出一次蝶形结运算时间应等于 4 个读写存储时间,从而无论是运算单元还是存储单元,都在全部时间工作,任何一种单元速度的降低,都会使整个系统速度降低。反过来说,如果只提高某种单元的速度,并不能使总的速度提高。

(2) 把少量高速暂存存储器与低速主存存储器结合使用。系统若要处于“匹配”状态,则 RAM 必须严格按照节拍脉冲工作,在很多情况下,这就要求 RAM 必须是高速的存储器,而大量使用高速存储器,成本就会比较昂贵,特别是当  $N$  值很大,又有多批数据待处理时,更加如此。如果采用低价的低速存储器,又不能满足系统工作速度的要求。为了既能满足系统高速运算的要求,又可不需要大量的高速存储器,以节省费用,则可采用附加高速暂存存储器与较低速的主存储器配合使用的结构,如图 4-40 所示。

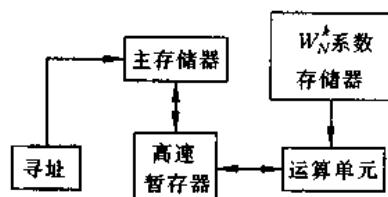


图 4-40 用高速暂存存储器提高 FFT 的速度

用普通的运算法完成  $N=16$  的处理,至少必须有 16 个数据的高速缓存单元。然而,如果把 FFT 流图作一些变化,如图 4-41 所示那样,它使 FFT 前两级运算的流程图有相

同的形式,该图是按  $N=16$  频率抽选、输入为自然顺序、输出为倒位序的情况下,把结构加以改变后的流程图。从图中看出,0,4,8,12 等四个数据的流程,在前两级运算中是一样的,且并不与其他数据点发生关系,同样,1,5,9,13 四个数据等的流程也有此特点。于是,可以从主存储器中一次只取出这四个数据,进行两级运算后,再放回主存储器,而不必像通常那样,每进行一次运算,就必须立即放回主存储器。因此,只需要 8 个数据的高速缓冲存储器(4 个数据单元在进行运算时,另 4 个正从主存储器中存取数据,然后交替工作)。在数据 0,4,8,12 做完两级蝶形结运算之后,可以用 1,5,9,13 重复这一过程,然后依此类推。这样便节省了一半高速存储器,且降低了对主存储器读写速度的要求。

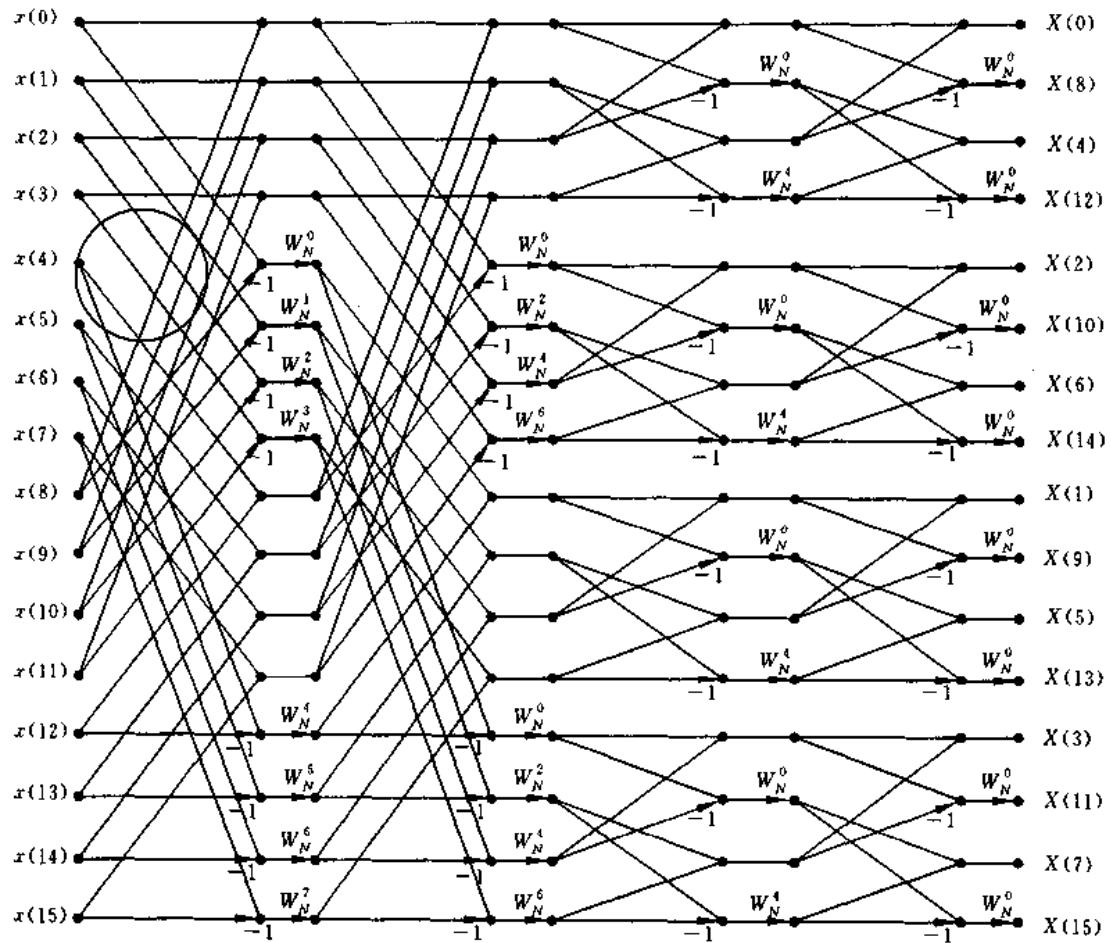


图 4-41 可以应用高速缓冲存储器的一种 FFT 流程图  
( $N=16$ , DIF, 输入自然顺序, 输出倒位序, 每两级运算结构相同)

### 基-2 流水线方式 FFT 处理机

为了提高 FFT 处理速度,特别是为了适用于多批数据处理,需要有比顺序处理还要快的办法。流水线 FFT 处理机就是一种适合上述要求的实用处理方案,这也就是前面讨论的级联处理方案,我们采用运算速度较快的方法来实现。

以图 4-10 的 DIT 算法的 FFT 为例,可以看出,后一级(列)的运算不一定要等前一级

运算完全执行完后才开始进行。前面已谈到,这种方案,每级安排一个运算器,当  $N=8$ ,共有  $L=\log_2 N=3$  个运算器,因而可以提高运算速度。它们的基本运算模块如图 4-42 所示,用此图作为  $N=8$  的 FFT 的第一级蝶形结运算模块。因为 FFT 第一级蝶形结的两个节点的间距为  $N/2=4$ ,所以只需 4 级移位寄存器,输入数据经双刀双掷开关  $K_1$  串行送入, $K_1$  先掷于位置 1,当移位寄存器中送入  $N/2=4$  个数据后,开关  $K_1$  接到 2 的位置。当第五个数据到来时,就与寄存中的第一个数据在运算单元中进行蝶形结运算,其结果  $X_0(0)+X_0(4)W_N^k$  作为第一级输出的第一个数据通过触点 2 串行送出后一级输入,而另一结果  $X_0(0)-X_0(4)W_N^k$  则作为第一级输出的第五个数据再存入移位寄存器中,去“挤掉”寄存器中的第一个数据(此数据不再应用了),当第一级的第六个数据输入时,正好与已在移位寄存器中的第二个数据在运算单元中进行蝶形结运算,其相加结果作为第二个输出数据,串行送到后一级,而其相减结果作为第六个输出数据再存入移位寄存器,把寄存器中的第二个数据“挤掉”。如此,当第一批  $N(N=8)$  个数据串行输入后,右边的节点 2 串行输出四个数据,这时, $K_1$  又将接到 1 位置,随后,移位寄存器向右输出第一级结果的第五到第八个数据,而其左方则输入并存储下一批数据的前四个,以备下一批数据的蝶形结运算。由以上讨论可见,这一级运算模块的输出比输入延迟四个抽样周期,移位寄存器移位速率与抽样频率相同,而蝶形结运算时间应小于一个抽样周期,此运算单元用于蝶形结运算的时间只占一半时间,即只在  $K$  位于 2 的位置时才进行运算。

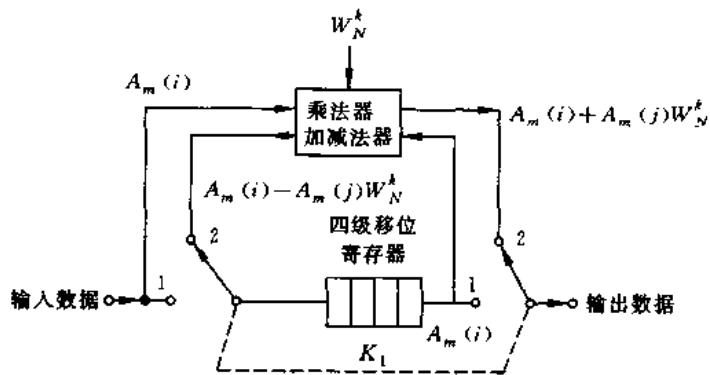


图 4-42  $N=8$ ,流水线 FFT 第一级运算模块

知道了第一级运算单元的工作原理之后,按  $N=8$  的算法流程,整个流水线 FFT 处理机的框图结构可见图 4-43 所示,其中运算单元的标号与所运算的 FFT 的级数相对应(即第一个运算单元运算第一级等等)。

现在来看第二级模块的工作。由图 4-5 可知第二级每个蝶形结的两个节点间距离为 2,故只需两级移位寄存器。开始时,“双刀双掷开关” $K_2$  掷于 1 的位置,当第一级输出的前两个数据已进入第二级移位寄存器,而第三个数据到来时, $K_2$  已从 1 转到 2,第三个数据直接进入运算单元 2 与已进入寄存器中的第一个数据进行蝶形结运算,和第一级一样,相加结果进入下一级,相减结果仍暂存到本级的第一个移位寄存器中。第四个数据到达此级时,和已存入移位寄存器的第二个数据进行蝶形结“加”、“减”运算,其结果进行与上述相同的处理,此时对第一级送来的前四个数据运算完毕, $K_2$  又转到 1 的位置,使移位寄

存器向右输出另两个结果(相减结果),再重复进行以上过程,完成第二级中的后两次蝶形结运算。从以上讨论可见,第二级的输出比输入只延迟两个抽样周期,开关  $K_2$  位于 1(或位于 2)的持续时间为两个抽样周期,只有开关位于 2 的位置时才进行运算。

对第三级可同样讨论,只不过这一级蝶形结两节点间距离为 1,故只需一级移位寄存器,每送入一个数据,开关就倒位一次,输出比输入延迟一个抽样周期,开关  $K_3$  位于 1(或 2)位置上的持续时间为一个抽样周期,同样,只在开关位于 2 位置时才进行运算。

按上述三级开关  $K_1, K_2, K_3$  在位置 1,2 处停留的规律,我们可以用一个三位二进制计数器各位的“0”或“1”的状态来控制这三级运算单元的开关,八个抽样周期上,计数器的状态如表 4-6 所示。

表 4-6 三位计数器在八个抽样周期上的状态表

输入数据		1	2	3	4	5	6	7	8
计数器各位状态	第 1 位	0	0	0	0	1	1	1	1
	第 2 位	0	0	1	1	0	0	1	1
	第 3 位	0	1	0	1	0	1	0	1

可以看出,要用计数器第 1 位控制开关  $K_1$ ,第 2 位控制开关  $K_2$ ,第 3 位控制开关  $K_3$ 。无论计数器哪一位为“0”状态时,与之相对应的运算单元的开关掷到 1 位置;为“1”状态时,相应的运算单元的开关掷于 2 位置。图 4-43 的开关状态是在计数器码为“1,0,1”时的情况。

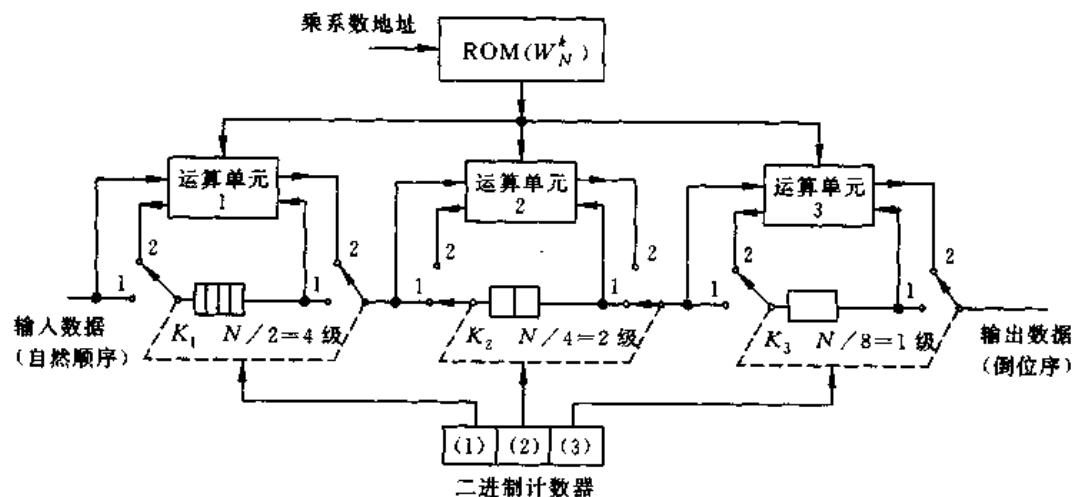


图 4-43  $N=8$ , 流水线 FFT 框图

同样,可以根据计数器的状态确定各运算单元的乘系数,从而可定出乘系数的地址码。例如,对运算单元 1,要求计数器第 1 位出现“1”时提供  $W_N^0$ ;对运算单元 2,要求计数器第 2 位第一次连续出现“1”时都提供  $W_N^0$ ,第二次连续出现“1”时都提供  $W_N^1$ ;对运算单元 3,要求计数器第 3 位第一次出现“1”时提供  $W_N^0$ ,第二、三、四次出现“1”时,则分别提供  $W_N^2, W_N^1, W_N^3$ 。

图 4-43 流水线式 FFT 处理机运算单元的利用率只有 50%，可以使运算器用于两路输入数据（双通道）的情况下，以充分发挥其作用，这就是双路流水线 FFT 处理机。若运算模块在单通道时能进行  $N$  点 FFT 变换，那么，在双通道时，FFT 变换点数将下降到  $N/2$ ，如图 4-44 所示，图中开关  $K$  以抽样率的速度轮流转换。

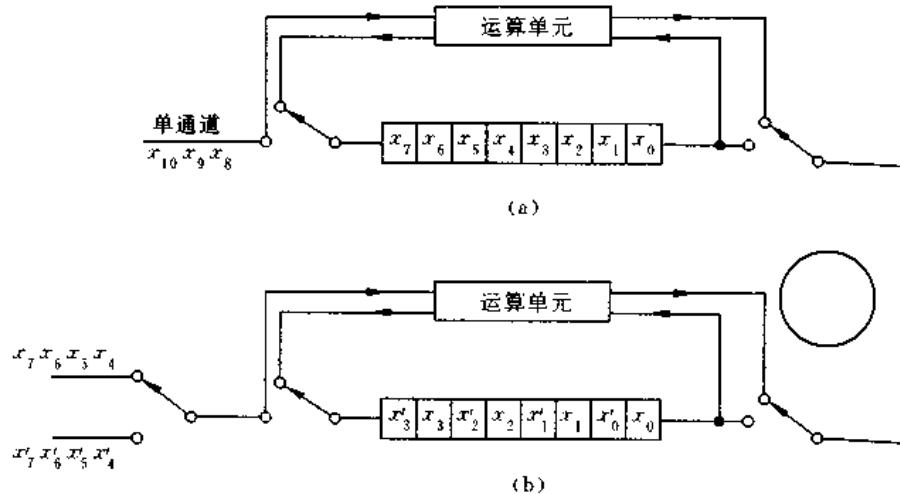


图 4-44 单、双通道工作时的原理图

(a) 单通道时，输出  $N=16$  点的 FFT 变换结果；(b) 双通道时，输出两组  $N=8$  点的 FFT 变换结果

单通道应用时，先把图 4-43 改为图 4-45，图中虚线方框内就是蝶形结运算单元。此图与图 4-43 的不同之处只是相减的结果采用单独的移位寄存器来做延迟存储。

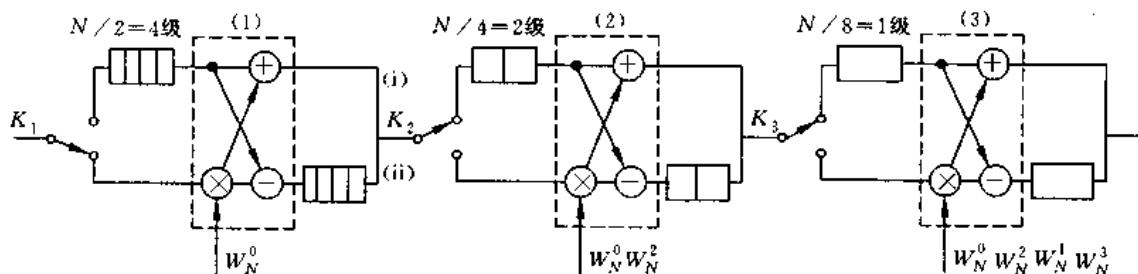


图 4-45 流水线 FFT 框图 ( $N=8$ )

图 4-45 的第一个运算单元输出的中间结果的次序为

输出(i)  $A_1(0) A_1(1) A_1(2) A_1(3)$

输出(ii)  $A_1(4) A_1(5) A_1(6) A_1(7)$

经过下路的移位寄存器延迟四位后的输出次序为  $A_1(0) A_1(1) A_1(2) A_1(3) A_1(4) A_1(5) A_1(6) A_1(7)$ 。实际上这种输出格式，对于实现下一级的运算并不方便，这是因为第二级运算中，蝶形结两节点的间距为 2，即第二个运算单元所要求的输入数据对的次序应为

上路  $A_1(0) A_1(1) A_1(4) A_1(5)$

下路  $A_1(2) A_1(3) A_1(6) A_1(7)$

要实现这样的输入次序，可采用以下方法：

(1) 先将(ii)对(i)延迟两个抽样周期(而不是图 4-45 所示的四个抽样周期延迟)得到

上路  $A_1(0) A_1(1) A_1(2) A_1(3)$

下路  $A_1(4) A_1(5) A_1(6) A_1(7)$

(2) 通过双刀双掷电子开关, 将  $A_1(2), A_1(3)$  分别与  $A_1(4), A_1(5)$  对调, 得到

上路  $A_1(0) A_1(1) A_1(4) A_1(5)$

下路  $A_1(2) A_1(3) A_1(6) A_1(7)$

(3) 将上路数据延迟两个抽样周期, 即得上面所列的第二个运算单元所要求的上、下两路输入数据对的次序。

同样道理, 对第 2 个运算单元的输出结果作类似地安排, 可得到第 3 个运算单元所需要的上、下两路数据对的以下排列次序:

上路  $A_2(0) A_2(2) A_2(4) A_2(6)$

下路  $A_2(1) A_2(3) A_2(5) A_2(7)$

为了达到上述目的, 应将图 4-45 模件中各级的下路延迟减去一半, 再将单刀双掷电子开关换成双刀双掷电子开关, 这就得到图 4-46 的双路流水线 FFT 处理机的原理方框图。此图的结构, 可以使两批输入数据依次进入, 作 FFT 处理。

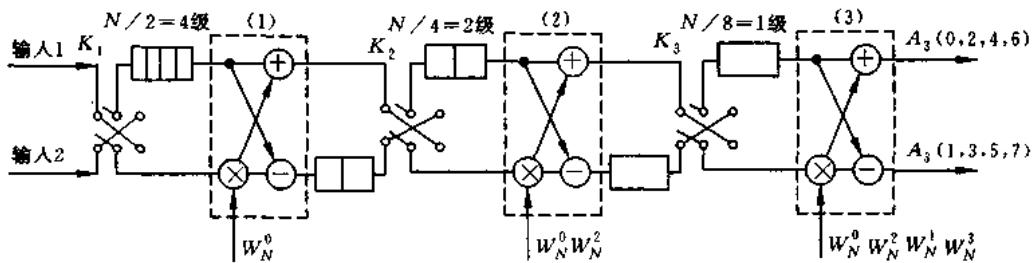


图 4-46 双路输入, 流水线 FFT 处理机原理框图

图 4-46 要求输入数据为图 4-47 所示的时序, 处理机输出分上、下两部分, 上部分输出  $N/2$  个结果为  $A_3(0), A_3(2), A_3(4), A_3(6)$  [相当于 DFT 的  $X(0), X(2), X(1), X(3)$ ], 下部分输出的  $N/2$  个结果为  $A_3(1), A_3(3), A_3(5), A_3(7)$  [相当于 DFT 的  $X(4), X(6), X(5), X(7)$ ]。这种处理机, 增加了几乎 50% ( $N$  越大, 越接近此百分数) 的移位寄存器作延时用。但它能同时处理两路输入数据, 使运算单元的利用率可达到 100%。要注意, 这里的双刀双掷开关  $K_1, K_2, K_3$  在位置 1 或 2 上停留时间分别为  $\frac{N}{2} = 4, \frac{N}{4} = 2, \frac{N}{8} = 1$  个抽样周期。

根据 4.5 节的讨论, 同一结构可用来实现流水线 FFT 的反变换(即 IFFT), 只需将乘系数用其共轭值代替, 将输出值再乘一个比例因子  $1/N$  即可, 双路流水线 FFT 正、反变换的处理方法, 在实现 FIR 滤波器和求卷积等方面都很有用处。

上面讨论的只是基-2 的流水线 FFT 结构, 高基数时, 模块比较复杂, 用得较多的是基-4 流水线结构。这里不再作介绍。

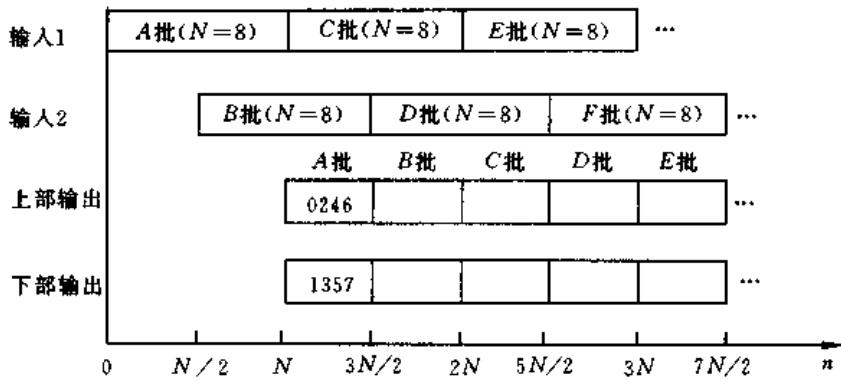


图 4-47 两路输入数据时,流水线处理机的输入输出数据时序

### 习 题

- 如果一台通用计算机的速度为平均每次复乘  $5\mu s$ , 每次复加  $0.5\mu s$ , 用它来计算  $512$  点的 DFT [ $x(n)$ ], 问直接计算需要多少时间, 用 FFT 运算需要多少时间。
- 已知  $X(k), Y(k)$  是两个  $N$  点实序列  $x(n), y(n)$  的 DFT 值, 今需要从  $X(k), Y(k)$  求  $x(n), y(n)$  值, 为了提高运算效率, 试用一个  $N$  点 IFFT 运算一次完成。
- $N=16$  时, 画出基-2 按时间抽选法及按频率抽选法的 FFT 流图(按时间抽选采用输入倒位序, 输出自然顺序, 按频率抽选采用输入自然顺序, 输出倒位序)。
- $N=16$ , 导出基-4 FFT 公式、画出流图, 并就运算量与基-2 的 FFT 相比较(不计乘土  $\pm 1$  及乘土  $\pm j$  的运算量)。
- 试用  $N$  为组合数时的 FFT 算法求  $N=12$  结果(采用混合基-3×4), 并画出流图。
- 同上题导出  $N=30=3\times 2\times 5$  的结果, 并画出流图。
- 研究一个  $M$  点的有限长序列  $x(n)$ .

$$x(n) = \begin{cases} x(n), & 0 \leq n \leq M-1 \\ 0, & \text{其他 } n \end{cases}$$

我们希望计算求  $z$  变换  $X(z) = \sum_{n=0}^{M-1} x(n)z^{-n}$  在单位圆上  $N$  个等间隔点上的抽样, 即在  $z_k = e^{j\frac{2\pi}{N}k}, k = 0, 1, \dots, N-1$  上的抽样。试对下列情况, 找出只用一个  $N$  点 DFT 就能计算  $X(z)$  的  $N$  个抽样的方法, 并证明之。

(a)  $N \leq M$

(b)  $N > M$

- 已知一个 8 点序列  $x(n)$

$$x(n) = \begin{cases} 1, & 0 \leq n \leq 7 \\ 0, & \text{其他 } n \end{cases}$$

试用 CZT 法求其前 10 点的复频谱  $X(z_k)$ 。已知  $z$  平面路径为  $A_0 = 0.8, \theta_0 = \frac{\pi}{3}, W_0 = 1, 2, \phi_0 = \frac{2\pi}{20}$ , 画出  $z_k$  的路径及 CZT 实现过程示意图。

- 在下列说法中选择正确的结论。线性调频  $z$  变换(CZT)可以用来计算一个有限长序列  $h(n)$  在  $z$  平面上的实轴上各  $(z_k)$  点的  $z$  变换  $H(z)$ , 使

(a)  $z_k = a^k, k = 0, 1, \dots, N-1, a$  为实数,  $a \neq \pm 1$ 。

- (b)  $z_k = ak$ ,  $k=0,1,\dots,N-1$ ,  $a$  为实数,  $a \neq 0$ 。  
(c) (a)和(b)两者都行。  
(d) (a)和(b)两者都不行。即线性调频  $z$  变换不能计算  $H(z)$  在  $z$  为实数时的抽样。

10. 当实现按时间抽选快速傅立叶变换算法时, 基本的蝶形计算

$$X_{m+1}(p) = X_m(p) + W_N X_m(q)$$

$$X_{m+1}(q) = X_m(p) - W_N X_m(q)$$

利用定点算术运算实现该蝶形计算时, 通常假设所有数字都已按一定比例因子化为小于 1。因此在蝶形计算的过程中还必须关心溢出问题。

(a) 证明如果我们要求  $|X_m(p)| < 1/2$  和  $|X_m(q)| < 1/2$ , 则在蝶形计算中不可能出现溢出, 即

$$\begin{aligned} \operatorname{Re}[X_{m+1}(p)] &< 1, & \operatorname{Im}[X_{m+1}(p)] &< 1 \\ \operatorname{Re}[X_{m+1}(q)] &< 1, & \operatorname{Im}[X_{m+1}(q)] &< 1 \end{aligned}$$

(b) 实际上要求:

$$\begin{aligned} |\operatorname{Re}[X_{m+1}(p)]| &< 1/2, & |\operatorname{Im}[X_{m+1}(p)]| &< 1/2 \\ |\operatorname{Re}[X_{m+1}(q)]| &< 1/2, & |\operatorname{Im}[X_{m+1}(q)]| &< 1/2 \end{aligned}$$

更容易些, 也更适合些。这些条件是否足以保证在蝶形计算中不会出现溢出? 请证明你的回答。

11.  $X(e^{j\omega})$  表示点数为 10 的有限长序列  $x(n)$  的傅立叶变换。我们希望计算  $X(e^{j\omega})$  在频率  $\omega_k = (2\pi k^2 / 100)$  ( $k=0,1,\dots,9$ ) 时的 10 个抽样。计算时不能采用先算出比要求数多的抽样然后再丢掉一些的办法。讨论采用下列各方法的可能性:

(a) 直接利用 10 点快速傅立叶变换算法。(b) 利用线性调频  $z$  变换算法。

12. 请用 C 语言编写程序

(1) 按频率抽选的 FFT 算法; (2) 分裂基 FFT 算法。

13. 我们希望利用一个单位抽样响应的点数  $N=50$  的有限长冲激响应滤波器来过滤一串很长的数据。要求利用重叠保留法通过快速傅立叶变换来实现这种滤波器, 为了做到这一点, (1) 输入各段必须重叠  $P$  个抽样点; (2) 必须从每一段产生的输出中取出  $Q$  个抽样点, 使这些从每一段得到的抽样连接在一起时, 得到的序列就是所要求的滤波输出。假设输入的各段为 100 个抽样点, 而离散傅立叶变换为 128 点。进一步假设, 圆周卷积的输出序列标号从  $n=0$  点到第  $n=127$  点。

(a) 求  $P$ ;

(b) 求  $Q$ ;

(c) 求取出来的  $Q$  个点之起点和终点的标号( $n$ ), 即确定从圆周卷积的 128 点中要取出哪些点, 去和前一段取出的点衔接起来。

## 第五章 数字滤波器的基本结构

### 5.1 数字滤波器结构的表示方法

在第二章中已经讲过,一个数字滤波器可以用系数函数表示为(见(2-91)式)

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} = \frac{Y(z)}{X(z)} \quad (5-1)$$

直接由此式可得出表示输入输出关系的常系数线性差分方程为

$$y(n) = \sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k) \quad (5-2)$$

可以看出,数字滤波器的功能就是把输入序列通过一定的运算(如(5-2)式)变换成为输出序列。可以用以下两种方法来实现数字滤波器:一种方法是把滤波器所要完成的运算编成程序并让计算机执行,也就是采用计算机软件来实现;另一种方法是设计专用的数字硬件、专用的数字信号处理器或采用通用的数字信号处理器来实现。

由(5-2)式看出,实现一个数字滤波器需要几种基本的运算单元——加法器、单位延时和常数乘法器。这些基本的单元可以有两种表示法——方框图法和信号流图法,因而一个数字滤波器的运算结构也有这样两种表示法,如图 5-1 所示。用方框图表示较明显直观,用流图表示则更加简单方便。以二阶数字滤波器

$$y(n) = a_1 y(n-1) + a_2 y(n-2) + b_0 x(n)$$

为例,其方框图结构如图 5-2 所示,这是读者较熟悉的。

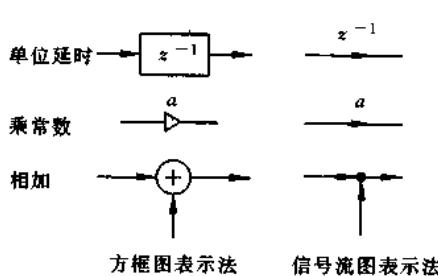


图 5-1 基本运算的方框图表示及流图表示

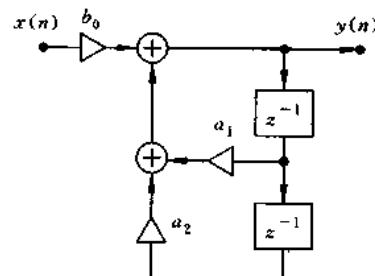


图 5-2 二阶数字滤波器的方框图结构

线性信号流图本质上与方框图表示法等效,只是符号上有差异。图 5-2 的二阶数字滤波器的等效信号流图结构如图 5-3 所示,图中 1,2,3,4,5 称为网络节点,  $x(n)$  处为输入节点或称源节点,表示注入流图的外部输入或信号源,  $y(n)$  处为输出节点或称阱节点。

节点之间用有向支路相连接,每个节点可以有几条输入支路和几条输出支路,任一节点的节点值等于它的所有输入支路的信号值之和。而输入支路的信号值等于这一支路起点处节点信号值乘以支路上的传输系数。如果支路上不标传输系数值,则认为其传输系数为1,而延迟支路则用延迟算子 $z^{-1}$ 表示,它表示单位延时。由此可得图5-3的各节点值为

$$w_2(n) = y(n)$$

$$w_3(n) = w_2(n-1) = y(n-1)$$

$$w_4(n) = w_3(n-1) = y(n-2)$$

$$w_5(n) = a_1 w_3(n) + a_2 w_4(n) = a_1 y(n-1) + a_2 y(n-2)$$

$$w_1(n) = b_0 x(n) + w_5(n) = b_0 x(n) + a_1 y(n-1) + a_2 y(n-2)$$

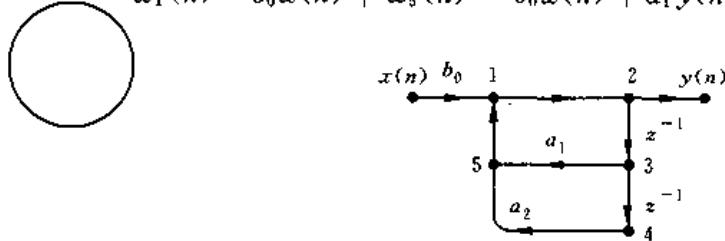


图5-3 图5-2的二阶数字滤波器的信号流图结构

源节点没有输入支路,阱节点没有输出支路。如果某节点有一个输入、一个或多个输出,则此节点相当于分支节点;如果某节点有两个或两个以上的输入,则此节点相当于相加器。因而节点2,3,4相当于分支节点,节点1,5相当于相加器。由此知道,对分支节点2有 $y(n)=w_2(n)=w_1(n)$ ,从而得出

$$y(n) = b_0 x(n) + a_1 y(n-1) + a_2 y(n-2)$$

这样我们就能清楚地看出其运算步骤和运算结构。下面我们都只采用信号流图来分析数字滤波器结构,前面我们已经用流图法分析了快速傅里叶变换的运算过程。

运算结构是重要的,不同结构所需的存储单元及乘法次数是不同的,前者影响复杂性,后者影响运算速度。此外,在有限精度(有限字长)情况下,不同运算结构的误差、稳定性是不同的。

由于无限长单位冲激响应(IIR)滤波器与有限长单位冲激响应(FIR)滤波器在结构上各有不同的特点,所以我们将分别对它们加以讨论。

## 5.2 无限长单位冲激响应(IIR)滤波器的基本结构

无限长单位冲激响应(IIR)滤波器有以下几个特点:

- (1) 系统的单位冲激响应 $h(n)$ 是无限长的;
- (2) 系统函数 $H(z)$ 在有限 $z$ 平面( $0 < |z| < \infty$ )上有极点存在;
- (3) 结构上存在着输出到输入的反馈,也就是结构上是递归型的。

但是,同一种系统函数 $H(z)$ 可以有多种不同的结构,它的基本网络结构有直接I

型、直接Ⅱ型、级联型和并联型四种。

### 一、直接Ⅰ型

一个IIR滤波器的有理系统函数为

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} = \frac{Y(z)}{X(z)} \quad (5-3)$$

表示这一系统输入输出关系的N阶差分方程为

$$y(n) = \sum_{k=1}^N a_k y(n-k) + \sum_{k=0}^M b_k x(n-k) \quad (5-4)$$

这就表示了一种计算方法。 $\sum_{k=0}^M b_k x(n-k)$ 表示将输入及延时后的输入,组成M节的延时网络,把每节延时抽头后加权(加权系数是 $b_k$ ),然后把结果相加,这就是一个横向结构网络。 $\sum_{k=1}^N a_k y(n-k)$ 表示将输出加以延时,组成N节的延时网络,再将每节延时抽头后加权(加权系数是 $a_k$ ),然后把结果相加。最后的输出 $y(n)$ 是把这两个和式相加而构成。由于包含了输出的延时部分,故它是个有反馈的网络。显然由(5-4)式右端的第一个和式构成了反馈网络。这种结构称为直接Ⅰ型结构,其结构流图如图5-4所示。由图可看出,总的网络是由上面讨论的两部分网络级联组成,第一个网络实现零点,第二个网络实现极点,从图中又可看出,直接Ⅰ型结构需要 $N+M$ 级延时单元。

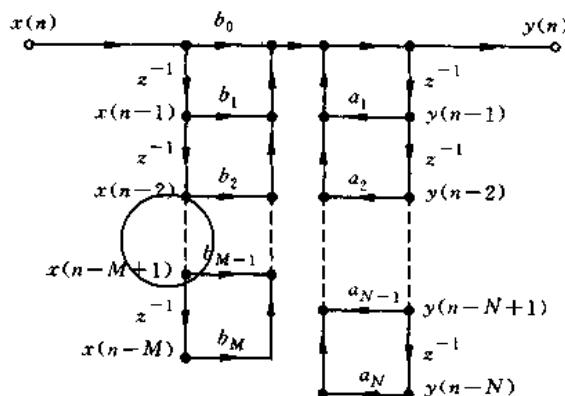


图5-4 实现N阶差分方程的直接Ⅰ型结构

### 二、直接Ⅱ型(典范型)

我们知道,一个线性移不变系统,若交换其级联子系统的次序,系统函数是不变的,也就是总的输入输出关系不改变。这样我们就得到另一种结构,如图5-5所示,它有两个级联子网络,第一个实现系统函数的极点,第二个实现系统函数的零点。可以看出,两行串行延时支路有相同的输入,因而可以把它们合并,则得到图5-6的结构,称为直接Ⅱ型结构或典范型结构。

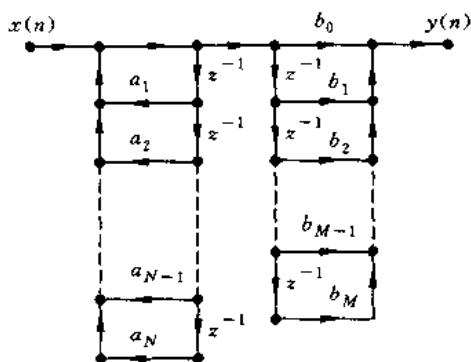


图 5-5 直接 I 型的变型, 将图 5-4 网络的零点与极点的级联次序互换

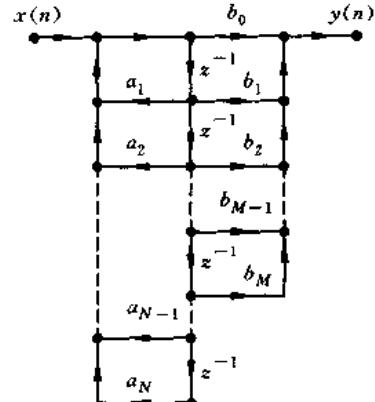


图 5-6 直接 II 型结构(典型型结构)

这种结构, 对于  $N$  阶差分方程只需  $N$  个延时单元(一般满足  $N \geq M$ ), 因而比直接 I 型延时单元要少, 这也是实现  $N$  阶滤波器所需的最少延时单元, 因而又称典型型。它可以节省存储单元(软件实现), 或节省寄存器(硬件实现), 比直接 I 型为好。但是, 它们都是直接型的实现方法, 其共同的缺点是系数  $a_k, b_k$  对滤波器的性能控制作用不明显, 这是因为它们与系统函数的零、极点关系不明显, 因而调整困难; 此外, 这种结构极点对系数的变化过于灵敏, 从而使系统频率响应对系数的变化过于灵敏, 也就是对有限精度(有限字长)运算过于灵敏, 容易出现不稳定或产生较大误差。

### 三、级联型

把(5-3)式的系统函数按零、极点进行因式分解, 则可表示成

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} = A \frac{\prod_{k=1}^{M_1} (1 - p_k z^{-1}) \prod_{k=1}^{M_2} (1 - q_k z^{-1}) (1 - q_k^* z^{-1})}{\prod_{k=1}^{N_1} (1 - c_k z^{-1}) \prod_{k=1}^{N_2} (1 - d_k z^{-1}) (1 - d_k^* z^{-1})}$$

式中  $M = M_1 + 2M_2$ ,  $N = N_1 + 2N_2$ 。一阶因式表示实根,  $p_k$  为实零点,  $c_k$  为实极点。二阶因式表示复共轭根,  $q_k, q_k^*$  表示复共轭零点,  $d_k, d_k^*$  表示复共轭极点。当  $a_k, b_k$  为实系数时, 上式就是最一般的零、极点分布表示法。把共轭因子组合成实系数的二阶因子, 则有

$$H(z) = A \frac{\prod_{k=1}^{M_1} (1 - p_k z^{-1}) \prod_{k=1}^{M_2} (1 + \beta_{1k} z^{-1} + \beta_{2k} z^{-2})}{\prod_{k=1}^{N_1} (1 - c_k z^{-1}) \prod_{k=1}^{N_2} (1 - \alpha_{1k} z^{-1} - \alpha_{2k} z^{-2})} \quad (5-5)$$

为了简化级联形式, 特别是在时分多路复用时, 采用相同形式的子网络结构就更有意义, 因而将实系数的两个一阶因子组合成二阶因子, 则整个  $H(z)$  就可以完全分解成实系数的二阶因子的形式

$$H(z) = A \prod_k \frac{1 + \beta_{1k} z^{-1} + \beta_{2k} z^{-2}}{1 - \alpha_{1k} z^{-1} - \alpha_{2k} z^{-2}} = A \prod_k H_k(z) \quad (5-6)$$

级联的节数视具体情况而定。当  $M=N$  时, 共有  $\left[\frac{N+1}{2}\right]$  节 ( $\left[\frac{N+1}{2}\right]$  表示  $\frac{N+1}{2}$  的整数)。如果有奇数个实零点, 则有一个  $\beta_{2k}$  等于零; 同样, 如果有奇数个实极点, 则有一个系数  $\alpha_{2k}$  等于零。每一个一阶、二阶子系统  $H_k(z)$  被称为一阶、二阶基本节,  $H_k(z)$  是用典型结构来实现的, 如图 5-7 所示。整个滤波器则是  $H_k(z)$  的级联如图 5-8 所示。一个六阶节系统的级联实现如图 5-9 所示。

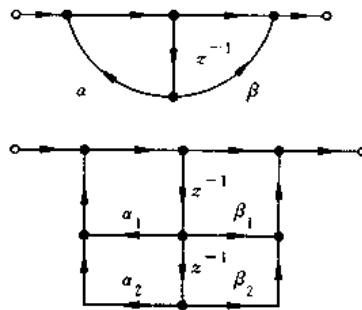


图 5-7 级联结构的一阶基本节和二阶基本节结构

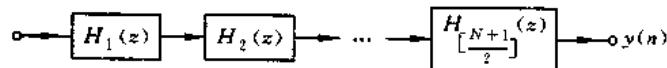


图 5-8 级联结构( $M=N$ )

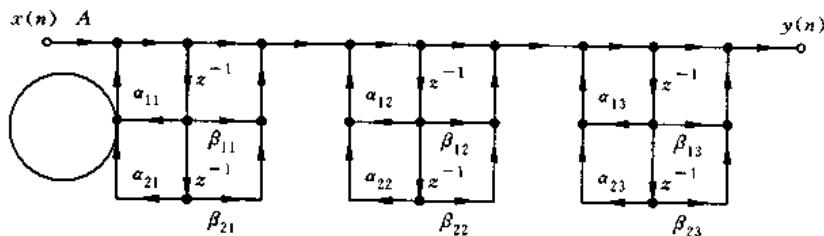


图 5-9 六阶 IIR 滤波器的级联结构

级联的特点是调整系数  $\beta_{1k}, \beta_{2k}$  就能单独调整滤波器的第  $k$  对零点, 而不影响其他零、极点, 同样, 调整系数  $\alpha_{1k}, \alpha_{2k}$  就能单独调整滤波器第  $k$  对极点, 而不影响其他零、极点。所以这种结构, 便于准确实现滤波器零、极点, 因而便于调整滤波器频率响应性能。

这种结构, 当  $M=N$  时, 分子、分母中二阶因子配合成基本二阶节可以有  $\left(\left[\frac{N+1}{2}\right]\right)!$  种, 而各二阶基本节的排列次序, 也可以有  $\left(\left[\frac{N+1}{2}\right]\right)!$  种, 它们都代表同一个系统函数  $H(z)$ 。但是, 当用二进制表示时, 只能采用有限位字长, 其所带来的误差, 对各种实现方案, 是不一样的, 因而对于配合与排列次序, 就存在着最优化的课题。

另外, 级联各节之间, 要有电平的放大和缩小, 以使变量值不会太大或太小。不能太

大是为了避免在定点制运算中产生溢出现象,不能太小是为了防止信号与噪声的比值太小,这将在以后讨论。级联结构具有最少的存储器。

#### 四、并联型

将因式分解的  $H(z)$  展成部分分式的形式,就得到并联型的 IIR 滤波器的基本结构

$$\begin{aligned} H(z) &= \frac{\sum_{k=0}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} \\ &= \sum_{k=1}^{N_1} \frac{A_k}{1 - c_k z^{-1}} + \sum_{k=1}^{N_2} \frac{B_k(1 - g_k z^{-1})}{(1 - d_k z^{-1})(1 - \bar{d}_k z^{-1})} + \sum_{k=0}^{M-N} G_k z^{-k} \quad (5-7) \end{aligned}$$

这一公式是最一般的表达式。式中  $N = N_1 + 2N_2$ , 由于系数  $a_k, b_k$  是实数, 故  $A_k, B_k, g_k, c_k, G_k$  都是实数,  $d_k$  是  $\bar{d}_k$  的共轭复数。当  $M < N$  时, 则(5-7)式中不包含  $\sum_{k=0}^{M-N} G_k z^{-k}$  项; 如果  $M = N$ , 则  $\sum_{k=0}^{M-N} G_k z^{-k}$  项变成  $G_0$  一项。一般 IIR 滤波器皆满足  $M \leq N$  的条件。(5-7)式表示系统是由  $N_1$  个一阶系统、 $N_2$  个二阶系统以及延时加权单元并联组合而成的。其结构实现如图 5-10 所示。而这些一阶和二阶系统都采用典范型结构实现。当  $M = N$  时,  $H(z)$  可表示为

$$H(z) = G_0 + \sum_{k=1}^{N_1} \frac{A_k}{1 - c_k z^{-1}} + \sum_{k=1}^{N_2} \frac{\gamma_{0k} + \gamma_{1k} z^{-1}}{1 - \alpha_{1k} z^{-1} - \alpha_{2k} z^{-2}} \quad (5-8)$$

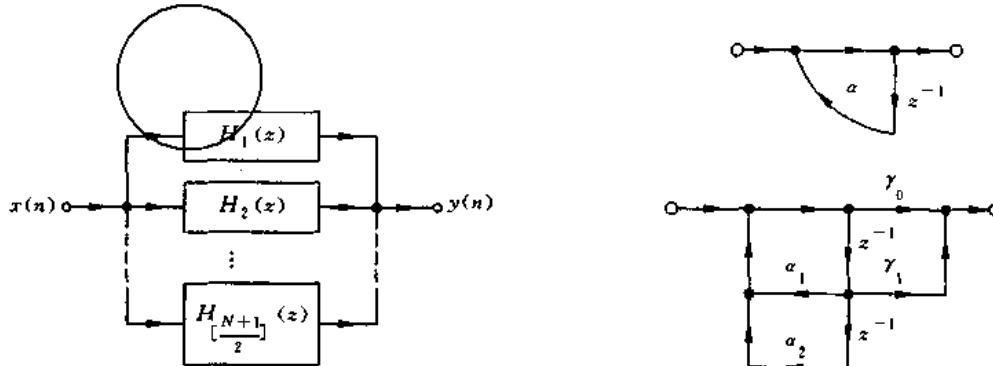


图 5-10 并联结构( $M=N$ )

图 5-11 并联结构的一阶基本节和二阶基本节结构

这里并联结构的一阶基本节、二阶基本节的结构如图 5-11 所示。

为了结构上的一致性,以便多路复用,一般将一阶实极点也组合成实系数二阶多项式,并将共轭极点对也化成实系数二阶多项式,当  $M=N$  时,有

$$H(z) = G_0 + \sum_{k=1}^{\left[\frac{N+1}{2}\right]} \frac{\gamma_{0k} + \gamma_{1k} z^{-1}}{1 - \alpha_{1k} z^{-1} - \alpha_{2k} z^{-2}} \quad (5-9)$$

可表示成

$$H(z) = G_0 + \sum_{k=1}^{\left[\frac{N+1}{2}\right]} H_k(z)$$

式中  $\left[\frac{N+1}{2}\right]$  表示取  $\frac{N+1}{2}$  的整数部分。当  $N$  为奇数时, 包含有一个一阶节, 即有一节的  $a_{2k} = \gamma_{1k} = 0$ , 当然这里并联的二阶基本节仍用典范型结构。图 5-12 画出了  $M=N=3$  时的并联型实现。

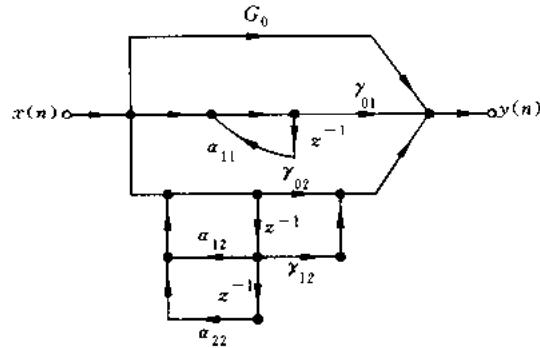


图 5-12 三阶 IIR 滤波器的并联型结构

并联型可以用调整  $a_{1k}, a_{2k}$  的办法来单独调整一对极点的位置, 但是不能像级联型那样单独调整零点的位置。此外, 并联结构中, 各并联基本节的误差互相没有影响, 所以比级联型的误差一般来说要稍小一些。因此在要求准确的传输零点的场合下, 宜采用级联型结构。

除了以上三种基本结构外, 还有一些其他的结构, 这取决于线性信号流图理论中的多种运算处理方法。当然各种流图都保持输入到输出的传输关系不变, 即  $H(z)$  不变。这其中有一种方法称为流图的转置, 它利用的是流图的转置定理。

**转置定理** 如果将原网络中所有支路方向倒转, 并将输入  $x(n)$  和输出  $y(n)$  相互交换, 则其系统函数  $H(z)$  不改变。

这里我们不去证明它了。利用转置定理, 可将上面讨论的各种结构加以转置而得到

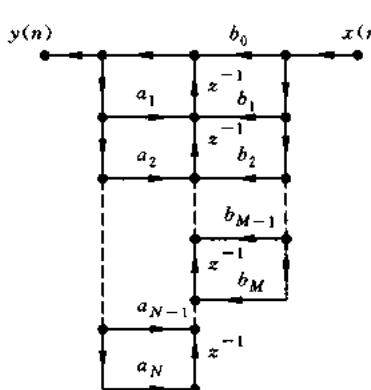


图 5-13 典范型结构的转置

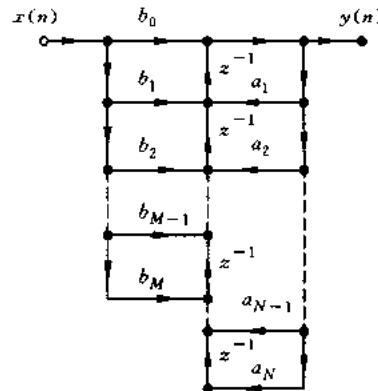


图 5-14 将图 5-13 画成输入在左, 输出在右的习惯形式

各种新的网络结构。例如,对图 5-6 的典范型结构,转置后的网络如图 5-13 所示,画成输入在左方、输出在右方的习惯形式,则如图 5-14 所示。

### 5.3 有限长单位冲激响应(FIR)滤波器的基本结构

有限长单位冲激响应滤波器有以下几个特点:

- (1) 系统的单位冲激响应  $h(n)$  在有限个  $n$  值处不为零;
- (2) 系统函数  $H(z)$  在  $|z|>0$  处收敛,在  $|z|>0$  处只有零点,有限  $z$  平面只有零点,而全部极点都在  $z=0$  处(因果系统);
- (3) 结构上主要是非递归结构,没有输出到输入的反馈,但有些结构中(例如频率抽样结构)也包含有反馈的递归部分。

设 FIR 滤波器的单位冲激响应  $h(n)$  为一个  $N$  点序列,  $0 \leq n \leq N-1$ , 则滤波器的系统函数为

$$H(z) = \sum_{n=0}^{N-1} h(n) z^{-n} \quad (5-10)$$

就是说,它有  $(N-1)$  阶极点在  $z=0$  处,有  $(N-1)$  个零点位于有限  $z$  平面的任何位置。FIR 滤波器有以下几种基本结构。

#### 一、横截型(卷积型、直接型)

(5-10)式的系统的差分方程表达式为

$$y(n) = \sum_{m=0}^{N-1} h(m) x(n-m) \quad (5-11)$$

很明显,这就是线性移不变系统的卷积和公式,也是  $x(n)$  的延时链的横向结构,如图 5-15 所示,称为横截型结构或卷积型结构,也可称为直接型结构。将转置定理用于图 5-15,可得到图 5-16 的转置直接型结构。

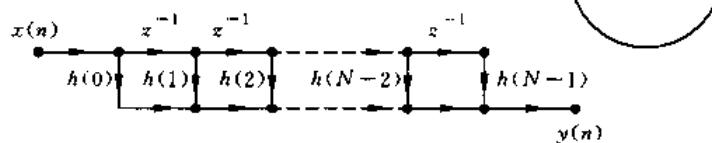


图 5-15 FIR 滤波器的横截型结构

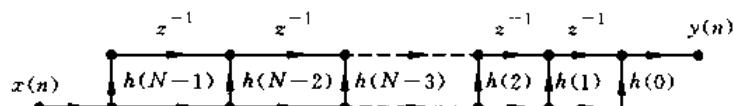


图 5-16 图 5-15 的转置结构

#### 二、级联型

将  $H(z)$  分解成实系数二阶因子的乘积形式

$$H(z) = \sum_{n=0}^{N-1} h(n)z^{-n} = \prod_{k=1}^{\lceil \frac{N}{2} \rceil} (\beta_{0k} + \beta_{1k}z^{-1} + \beta_{2k}z^{-2}) \quad (5-12)$$

其中  $\lceil N/2 \rceil$  表示取  $N/2$  的整数部分。若  $N$  为偶数，则  $N-1$  为奇数，故系数  $\beta_{2k}$  中有一个为零。这是因为，这时有奇数个根，其中复数根成共轭对，必为偶数，必然有奇数个实根。图 5-17 画出了  $N$  为奇数时 FIR 滤波器的级联结构，其中每一个二阶因子用图 5-15 的横截型结构。

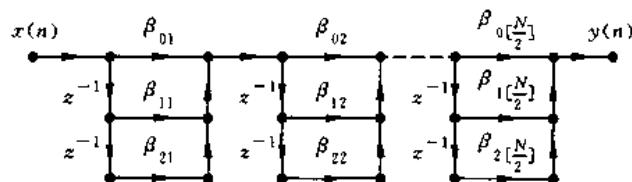


图 5-17 FIR 滤波器的级联型结构( $N$  为奇数)

这种结构的每一节控制一对零点，因而在需要控制传输零点时，可以采用它。但是这种结构所需要的系数  $\beta_{ik}$  ( $i=0,1,2; k=1,2,\dots,\lceil N/2 \rceil$ ) 比卷积型的系数  $h(n)$  要多，因而所需的乘法次数也比卷积型的要多。

### 三、频率抽样型

在第三章中已经说过，把一个  $N$  点有限长序列的  $z$  变换  $H(z)$  在单位圆上作  $N$  等分抽样，就得到  $\tilde{H}(k)$ ，其主值序列就等于  $h(n)$  的离散傅里叶变换  $H(k)$ 。那里也说到用  $H(k)$  表示  $H(z)$  的内插公式为

$$H(z) = (1 - z^{-N}) \frac{1}{N} \sum_{k=0}^{N-1} \frac{H(k)}{1 - W_N^{-k} z^{-1}} \quad (5-13)$$

这个公式就为 FIR 滤波器提供了另外一种结构，这种结构由两部分级联组成。

$$H(z) = \frac{1}{N} H_c(z) \sum_{k=0}^{N-1} H'_k(z) \quad (5-14)$$

其中级联的第一部分为

$$H_c(z) = 1 - z^{-N} \quad (5-15)$$

这是一个 FIR 子系统，是由  $N$  节延时单元构成的梳状滤波器，令

$$H_c(z) = 1 - z^{-N} = 0$$

则有

$$\begin{aligned} z_i^N &= 1 = e^{j2\pi i} \\ z_i &= e^{j\frac{2\pi}{N}i}, \quad i = 0, 1, \dots, N-1 \end{aligned}$$

即  $H_c(z)$  在单位圆上有  $N$  个等间隔角度的零点，它的频率响应为

$$H_c(e^{j\omega}) = 1 - e^{-j\omega N} = 2j e^{-j\frac{\omega N}{2}} \sin\left(\frac{\omega N}{2}\right) \quad (5-16)$$

因而幅度响应为

$$|H_c(e^{j\omega})| = 2 \left| \sin\left(\frac{\omega N}{2}\right) \right|$$

相角为

$$\arg[H_c(e^{j\omega})] = \frac{\pi}{2} - \frac{\omega N}{2} + m\pi, \quad \begin{cases} m = 0, \omega = 0 \text{ 到 } \frac{2\pi}{N} \\ m = 1, \omega = \frac{2\pi}{N} \text{ 到 } \frac{4\pi}{N} \\ \vdots \\ m = m, \omega = \frac{2m\pi}{N} \text{ 到 } \frac{(m+1)2\pi}{N} \end{cases}$$

其子网络结构及频率响应幅度如图 5-18 所示。

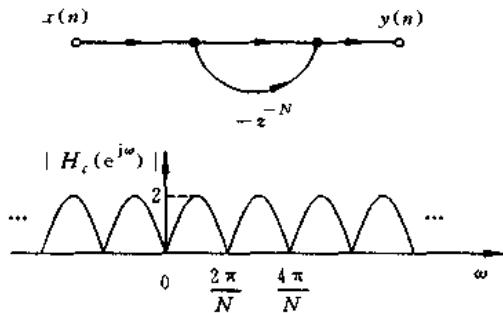


图 5-18 梳状滤波器结构及频率响应幅度

级联的第二部分为

$$\sum_{k=0}^{N-1} H'_k(z) = \sum_{k=0}^{N-1} \frac{H(k)}{1 - W_N^{-k} z^{-1}}$$

它是由  $N$  个一阶网络并联组成, 而这每一个一阶网络都是一个谐振器

$$H'_k(z) = \frac{H(k)}{1 - W_N^{-k} z^{-1}} \quad (5-17)$$

令  $H'_k(z)$  的分母为零, 即令

$$1 - W_N^{-k} z^{-1} = 0$$

可得到此一阶网络在单位圆上有一个极点

$$z_k = W_N^{-k} = e^{j\frac{2\pi k}{N}}$$

也就是说, 此一阶网络在频率为

$$\omega = \frac{2\pi}{N}k$$

处响应为无穷大, 故等效于谐振频率为  $\frac{2\pi}{N}k$  的无损耗谐振器。这个谐振器的极点正好与

梳状滤波器的一个零点 ( $i=k$ ) 相抵消, 从而使这个频率 ( $\omega=\frac{2\pi}{N}k$ ) 上的频率响应等于

$H(k)$ 。这样,  $N$  个谐振器的  $N$  个极点就和梳状滤波器的  $N$  个零点相互抵消, 从而在  $N$  个

频率抽样点 ( $\omega=\frac{2\pi}{N}k, k=0, 1, \dots, N-1$ ) 的频率响应就分别等于  $N$  个  $H(k)$  值。

$N$  个并联谐振器与梳状滤波器级联后, 就得到图 5-19 的频率抽样结构。频率抽样结构的特点是它的系数  $H(k)$  就是滤波器在  $\omega=\frac{2\pi}{N}k$  处的响应, 因此控制滤波器的频率响应

很方便。但是结构中所乘的系数  $H(k)$  及  $W_N^{-k}$  都是复数, 增加了乘法次数和存储量, 而且所有极点都在单位圆上, 由系数  $W_N^{-k}$  决定, 这样, 当系数量化时, 这些极点会移动, 有些极点就不能被梳状滤波器的零点所抵消(零点由延时单元决定, 不受量化的影响)。如果极点移到  $z$  平面单位圆外, 系统就不稳定了。

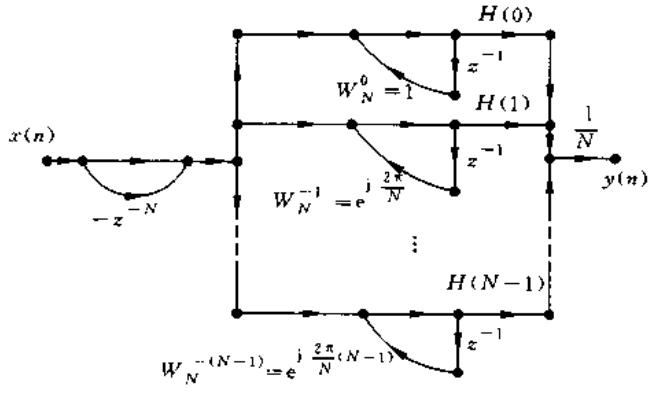


图 5-19 FIR 滤波器的频率抽样型结构

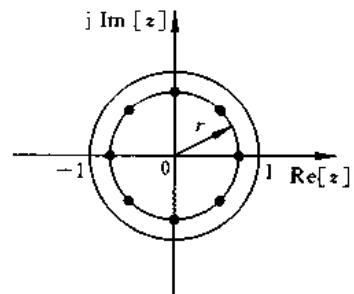


图 5-20 抽样点改到  $r \leq 1$  的圆上

为了克服系数量化后可能不稳定的缺点, 可以将频率抽样结构做一点修正, 即将所有零、极点都移到单位圆内某一靠近单位圆、半径为  $r$  ( $r$  小于或近似等于 1) 的圆上( $r$  为正实数), 如图 5-20 所示。这时

$$H(z) = \frac{1 - r^N z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H_r(k)}{1 - r W_N^{-k} z^{-1}}$$

$H_r(k)$  为新抽样点上的抽样值, 但是由于  $r \approx 1$ , 因此有

$$H_r(k) \approx H(k)$$

即

$$H_r(k) = H(z) |_{z=r W_N^{-k}} \approx H(z) |_{z=W_N^{-k}} = H(k)$$

所以

$$H(z) \approx \frac{1 - r^N z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H(k)}{1 - r W_N^{-k} z^{-1}} \quad (5-18)$$

下面, 我们化简这一公式。首先, 谐振器的各个根( $H(z)$  的极点)为

$$z_k = r e^{j \frac{2\pi}{N} k}, \quad k = 0, 1, \dots, N-1$$

为了使系数为实数, 可将共轭根合并, 在  $z$  平面上这些共轭根在半径为  $r$  的圆周上以实轴为轴成对称分布, 如图 5-21 所示, 满足

$$z_{N-k} = z_k^*$$

也就是

$$r W_N^{-(N-k)} = r e^{j \frac{2\pi}{N} (N-k)} = r \left( e^{j \frac{2\pi}{N} k} \right)^* = r W_N^{-k}$$

其次, 由于  $h(n)$  是实数, 故  $H(k) = \text{DFT}[h(n)]$  也是共轭对称的, 即

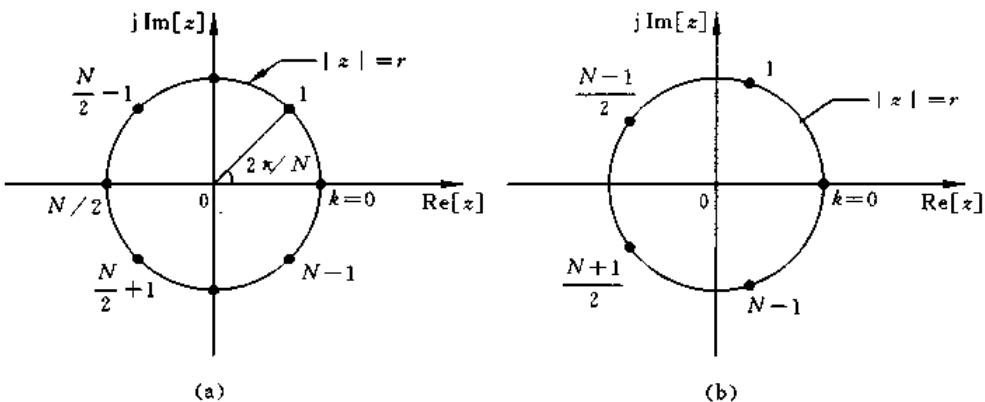


图 5-21 谐振器各个根的位置

(a)  $N$  为偶数; (b)  $N$  为奇数

$$H(k) = H^*((N-k))_N R_N(k), \begin{cases} k = 1, 2, \dots, \frac{N-1}{2}, & \text{当 } N \text{ 为奇数时} \\ k = 1, 2, \dots, \frac{N}{2}-1, & \text{当 } N \text{ 为偶数时} \end{cases}$$

因此,可以将第  $k$  个与第  $(N-k)$  个谐振器合并为一个实系数的二阶网络,以  $H_k(k)$  表示为

$$\begin{aligned} H_k(z) &= \frac{H(k)}{1 - rW_N^{-k}z^{-1}} + \frac{H(N-k)}{1 - rW_N^{-(N-k)}z^{-1}} = \frac{H(k)}{1 - rW_N^{-k}z^{-1}} + \frac{H^*(k)}{1 - rW_N^{*-k}z^{-1}} \\ &= \frac{\beta_{0k} + \beta_{1k}z^{-1}}{1 - z^{-1}2r\cos\left(\frac{2\pi k}{N}\right) + r^2z^{-2}}, \begin{cases} k = 1, 2, \dots, \frac{N-1}{2}, & N \text{ 为奇数} \\ k = 1, 2, \dots, \frac{N}{2}-1, & N \text{ 为偶数} \end{cases} \end{aligned} \quad (5-19)$$

其中

$$\begin{aligned} \beta_{0k} &= 2\operatorname{Re}[H(k)] \\ \beta_{1k} &= -2r\operatorname{Re}[H(k)W_N^k] \end{aligned} \quad (5-20)$$

由于这个二阶网络的极点在单位圆内,而不是在单位圆上,因而从频率响应的几何解释知道,它相当于一个有限  $Q$ (品质因数)的谐振器,谐振频率为

$$\omega_k = \frac{2\pi k}{N}$$

其结构如图 5-22 所示。

除共轭复根外,尚有实根。

当  $N$  为偶数时,如图 5-21(a)所示,有一对实根(相当于  $k=0$  及  $k=N/2$  两点)

$$z = \pm r$$

因而对应的一阶网络为

$$H_0(z) = \frac{H(0)}{1 - rz^{-1}} \quad (5-21)$$

$$H_{N/2}(z) = \frac{H(N/2)}{1 + rz^{-1}} \quad (5-22)$$

其结构如图 5-23 所示。

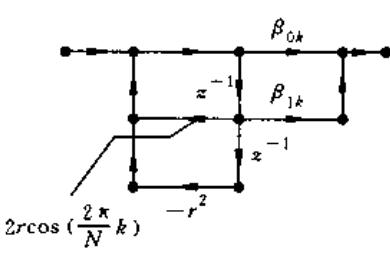


图 5-22 二阶谐振器

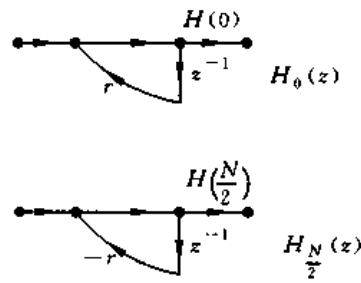


图 5-23 一阶网络(实根)

当  $N$  为奇数时, 如图 5-21(b) 所示, 只有一个实根(相当于  $k=0$  的点)

$$z = r$$

因而只有一个网络  $H_0(z)$  而没有  $H_{N/2}(z)$ 。

将谐振器的实根、复根以及梳状滤波器合起来得到修正后的频率抽样型总结构。当  $N$  为偶数时

$$\begin{aligned} H(z) &= (1 - r^N z^{-N}) \frac{1}{N} \left[ \frac{H(0)}{1 - rz^{-1}} + \frac{H(N/2)}{1 + rz^{-1}} + \sum_{k=1}^{N/2-1} \frac{\beta_{0k} + \beta_{1k} z^{-1}}{1 - z^{-1} 2r \cos\left(\frac{2\pi}{N} k\right) + r^2 z^{-2}} \right] \\ &= (1 - r^N z^{-N}) \frac{1}{N} \left[ H_0(z) + H_{N/2}(z) + \sum_{k=1}^{N/2-1} H_k(z) \right] \end{aligned} \quad (5-23)$$

当  $N$  为奇数时

$$\begin{aligned} H(z) &= (1 - r^N z^{-N}) \frac{1}{N} \left[ \frac{H(0)}{1 - rz^{-1}} + \sum_{k=1}^{(N-1)/2} \frac{\beta_{0k} + \beta_{1k} z^{-1}}{1 - z^{-1} 2r \cos\left(\frac{2\pi}{N} k\right) + r^2 z^{-2}} \right] \\ &= (1 - r^N z^{-N}) \frac{1}{N} \left[ H_0(z) + \sum_{k=1}^{(N-1)/2} H_k(z) \right] \end{aligned} \quad (5-24)$$

当  $N$  为偶数时, 其结构如图 5-24 所示。图中第一个  $H_0(z)$  及最后一个  $H_{N/2}(z)$  是一阶的, 其具体结构如图 5-23 所示, 当  $N$  为奇数时, 没有  $H_{N/2}(z)$ 。其他各  $H_k(z)$  都是二阶的, 其具体结构见图 5-22 所示。

频率抽样结构的另一个特点是它的零、极点数目只取决于单位抽样响应的点数, 因而, 只要单位冲激响应点数相同, 利用同一梳状滤波器、同一结构而只有加权系数  $\beta_{0k}, \beta_{1k}$ ,  $H(0), H(N/2)$  不同的谐振器, 就能得到各种不同的滤波器, 因而图 5-24 结构是高度模块化的, 适用于时分复用。

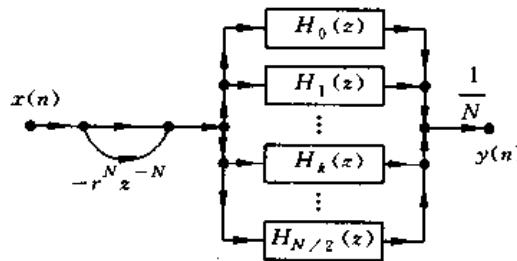


图 5-24 FIR 滤波器修正后的频率抽样结构

#### 四、快速卷积结构

在第三章中已经讲过,只要将两个有限长序列补上一定的零值点,就可以用两序列的圆周卷积代替两序列的线性卷积。也就是将  $x(n)$  和  $h(n)$  都变成  $L$  点序列,即将  $N_1$  点输入  $x(n)$  ( $0 \leq n \leq N_1 - 1$ ) 补  $L - N_1$  个零值点,将  $N_2$  点冲激响应  $h(n)$  ( $0 \leq n \leq N_2 - 1$ ) 补  $L - N_2$  个零值点,只要满足

$$L \geq N_1 + N_2 - 1$$

则  $x(n)$  与  $h(n)$  的  $L$  点圆周卷积就代表它们的线性卷积。

实际上,我们并不是在时域做圆周卷积,而是利用“时域序列的圆周卷积等效于频域的离散频谱的乘积”这一性质先在离散频域进行。具体表示方法如下。

1. 将  $x(n)$  和  $h(n)$  变成  $L$  点序列,  $L \geq N_1 + N_2 - 1$

$$x(n) = \begin{cases} x(n), & 0 \leq n \leq N_1 - 1 \\ 0, & N_1 \leq n \leq L - 1 \end{cases}$$

$$h(n) = \begin{cases} h(n), & 0 \leq n \leq N_2 - 1 \\ 0, & N_2 \leq n \leq L - 1 \end{cases}$$

2. 求  $x(n)$  与  $h(n)$  各自的  $L$  点 DFT

$$X(k) = \text{DFT}[x(n)], \quad L \text{ 点}$$

$$H(k) = \text{DFT}[h(n)], \quad L \text{ 点}$$

3. 将  $X(k)$  与  $H(k)$  相乘得  $Y(k)$

$$Y(k) = X(k)H(k), \quad L \text{ 点}$$

4. 求  $Y(k)$  的  $L$  点 IDFT, 得  $y(n)$ ,  $y(n)$  的前  $N_1 + N_2 - 1$  个点就等于  $x(n)$  与  $h(n)$  的线性卷积

$$y(n) = \text{IDFT}[Y(k)] = \text{IDFT}[X(k)H(k)] = x(n) \circledast h(n)$$

则  $L$  点的圆周卷积就能代表线性卷积, 即

$$y(n) = x(n) \circledast h(n) = x(n) * h(n), \quad 0 \leq n \leq N_1 + N_2 - 2$$

这样,我们就得到图 5-25 的快速卷积结构。当  $N_1, N_2$  足够长时,用这种结构计算线性卷积要快得多。实际上,这里的 DFT 和 IDFT 都是采用第四章的快速傅里叶变换计算方法(取  $L=2^p$ ,  $p$  为正整数)。

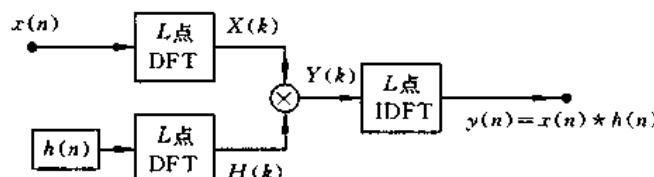


图 5-25 FIR 滤波器的快速卷积结构

#### 五、线性相位 FIR 滤波器的结构

FIR 滤波器的线性相位是非常重要的,因为数据传输以及图像处理都要求系统具有线性相位,而 FIR 滤波器由于它的冲激响应是有限长的,因而有可能做成严格线性相

位的。

我们知道,如果 FIR 滤波器单位冲激响应  $h(n)$  为实数,  $0 \leq n \leq N-1$ , 且满足以下的条件:

$$b(\alpha) = b(N-1-\alpha) \quad (5-25a)$$

$$f(x) = -f(N+1-x) \quad (5-25b)$$

也就是说,其对称中心在  $n=(N-1)/2$  处,则这种 FIR 滤波器就具有严格线性相位,下面我们来导出这种滤波器的结构。

设 FIR 滤波器的单位冲激响应为  $h(n), 0 \leq n \leq N-1$ , 且  $h(n)$  满足以上任一种对称条件. 其系统函数为

$$H(z) = \sum_{n=1}^{N-1} h(n)z^{-n} \quad (5-26)$$

下面对  $N$  为奇数及  $N$  为偶数两种情况分别加以讨论。

当  $N$  为奇数时

$$H(z) = \sum_{n=0}^{N-1} h(n)z^{-n} = \sum_{n=0}^{\frac{N-1}{2}-1} h(n)z^{-n} + h\left(\frac{N-1}{2}\right)z^{-\frac{N-1}{2}} + \sum_{n=\frac{N-1}{2}+1}^{N-1} h(n)z^{-n}$$

在第二个 $\Sigma$ 式中,令 $n=N=1=m$ ,再将 $m$ 换成 $n$ ,可得

$$H(z) = \sum_{n=0}^{\frac{N-1}{2}-1} h(n)z^{-n} + h\left(\frac{N-1}{2}\right)z^{-\frac{N-1}{2}} + \sum_{n=0}^{\frac{N-1}{2}-1} h(N-1-n)z^{-(N-1-n)}$$

代入线性相位奇偶对称条件  $h(N-1-n) = \pm h(n)$ , 可得

$$H(z) = \sum_{n=0}^{\frac{N-1}{2}-1} h(n)[z^{-n} \pm z^{-(N-1-n)}] + h\left(\frac{N-1}{2}\right)z^{-\frac{N-1}{2}} \quad (5-27)$$

其中,方括号内的“+”号表示  $h(n)$  是偶对称,“-”表示  $h(n)$  呈奇对称。 $h(n)$  奇对称时,必有  $h[(N-1)/2]=0$ (见(5-25b)式,由(5-27)式可画出  $N$  为奇数时,线性相位 FIR 滤波器的直接结构的流图如图 5-26 所示。

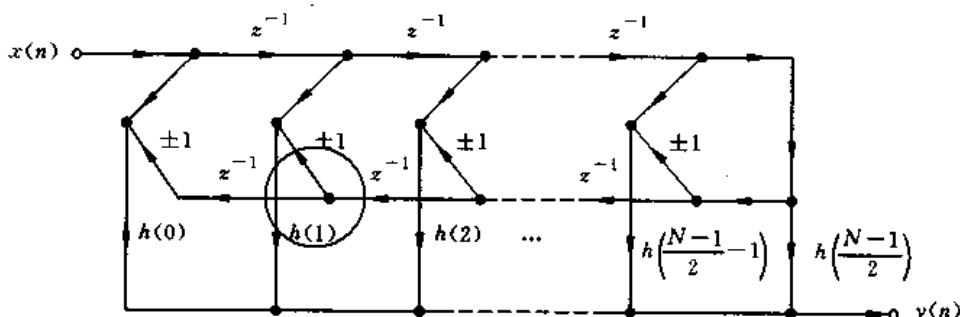


图 5-26  $N$  为奇数时线性相位 FIR 滤波器的直接型结构

$[h(n) \text{偶对称时} \pm 1 \text{取} +1,$   
 $h(n) \text{奇对称时} \pm 1 \text{取} -1, \text{且 } h\left(\frac{N-1}{2}\right) = 0 \text{ 即 } h\left(\frac{N-1}{2}\right) \text{ 处的连线断开}]$

当  $N$  为偶数时

$$H(z) = \sum_{n=0}^{N-1} h(n)z^{-n} = \sum_{n=0}^{(N/2)-1} h(n)z^{-n} + \sum_{n=N/2}^{N-1} h(n)z^{-n}$$

在第二个 $\Sigma$ 式中,令 $n=N-1-m$ ,再将 $m$ 换成 $n$ ,可得

$$H(z) = \sum_{n=0}^{(N/2)-1} h(n)z^{-n} + \sum_{n=0}^{(N/2)-1} h(N-1-n)z^{-(N-1-n)}$$

代入线性相位奇偶对称条件 $h(n)=\pm h(N-1-n)$ ,可得

$$H(z) = \sum_{n=0}^{(N/2)-1} h(n)[z^{-n} \pm z^{-(N-1-n)}] \quad (5-28)$$

其中,方括号内的“+”表示 $h(n)$ 呈偶对称,“-”表示 $h(n)$ 呈奇对称。由(5-28)式可画出 $N$ 为偶数时,线性相位 FIR 滤波器的直接结构的流图如图 5-27 所示。

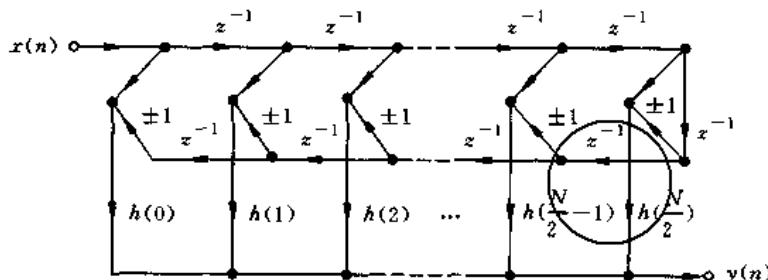


图 5-27  $N$  为偶数时,线性相位 FIR 滤波器的直接型结构  
( $h(n)$  偶对称时 $\pm 1$  取 $+1$ , $h(n)$  奇对称时 $\pm 1$  取 $-1$ )

由以上流图看出线性相位 FIR 滤波器结构比一般直接型结构可以节省一半数量的乘法次数。

我们在第七章 FIR 数字滤波器设计一章中还要讨论到线性相位 FIR 数字滤波器的级联结构。

## 5.4 数字滤波器的格型结构

上面我们讨论了 IIR 滤波器和 FIR 滤波器的各种结构,下面我们来讨论一种新的结构形式,即格型(Lattice)结构。事实证明:①由于它的模块化结构便于实现高速并行处理;②一个 $m$  阶格型滤波器可以产生从 1 阶到 $m$  阶的 $m$  个横向滤波器的输出性能;③它对有限字长的舍入误差不灵敏。由于这些优点,使得这种结构在现代谱估计、语音信号处理、自适应滤波等方面得到了广泛的应用。我们分别就全零点系统、全极点系统及极、零点系统的格型结构进行以下讨论。

### 一、全零点系统(FIR 系统)的格型结构

一个 $M$  阶的 FIR 滤波器的横向结构的系统函数 $H(z)$ 可写成

$$H(z) = B(z) = \sum_{i=0}^M b_i^{(M)} z^{-i} = 1 + \sum_{i=1}^M b_i^{(M)} z^{-i} \quad (5-29)$$

系数 $b_i^{(M)}$  表示 $M$  阶 FIR 系统的第 $i$  个系数,式中假定 $H(z)$  的首项系数 $h(0)=1$ 。而此

全零点 FIR 系统的格型结构则如图 5-28 所示。

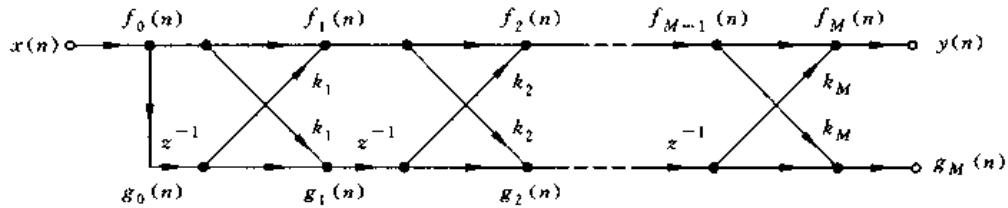


图 5-28 全零点系统(FIR 系统)的格型结构

下面我们来分析这一格型结构,讨论如何由横向结构(卷积结构或者直接结构)的参数导出格型结构的参数,或者由格型结构的参数如何导出横向结构的参数。

在 FIR 的横向结构中有  $M$  个参数  $b_i^{(M)}$  [或  $h(i)$ ]  $i=1,2,\dots,M$ ,共需  $M$  次乘法, $M$  次延迟;在 FIR 格型结构中也有  $M$  个参数  $k_i$  ( $i=1,2,\dots,M$ ), $k_i$  称之为反射系数,共需  $2M$  次乘法, $M$  次延迟。

此格型结构的信号只有正馈通路,没有反馈通路,所以是一个典型的 FIR 系统。

格型滤波器结构中的基本传输单元如图 5-29 所示,有如下的表达式:

$$f_m(n) = f_{m-1}(n) + k_m g_{m-1}(n-1), \quad m = 1, 2, \dots, M \quad (5-30)$$

$$g_m(n) = k_m f_{m-1}(n) + g_{m-1}(n-1), \quad m = 1, 2, \dots, M \quad (5-31)$$

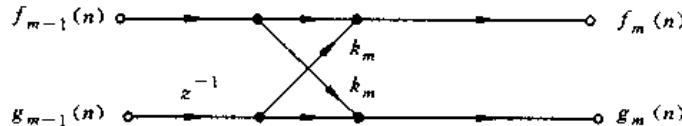


图 5-29 全零点(FIR 系统)格型结构基本传输单元

并且有

$$f_0(n) = g_0(n) = x(n) \quad (5-32)$$

$$f_M(n) = y(n) \quad (5-33)$$

若定义  $B_m(z), \bar{B}_m(z)$  分别是由输入端到第  $m$  个基本传输单元上端和下端所对应的系统函数

$$B_m(z) = \frac{F_m(z)}{F_o(z)} = 1 + \sum_{i=1}^m b_i^{(m)} z^{-i}, \quad m = 1, 2, \dots, M \quad (5-34a)$$

$$\bar{B}_m(z) = \frac{G_m(z)}{G_o(z)}, \quad m = 1, 2, \dots, M \quad (5-34b)$$

可以看出,当  $m=M$  时,  $B_M(z)=B(z)$ 。同样可以看出,  $m-1$  级的  $B_{m-1}(z)$  与图 5-29 的基本单元级联即得到  $m$  级的  $B_m(z)$ ,因此格型结构有着模块化的结构形式。

1. 首先,我们来看从高阶  $B_m(z)$  到低一阶的  $B_{m-1}(z)$  以及反过来从  $B_{m-1}(z)$  到  $B_m(z)$  的递推关系,这也就隐含了格型结构的  $k_i$  ( $i=1,2,\dots,M$ ) 和横向结构的各系数  $b_i^{(m)}$  ( $i=1,2,\dots,m; m=1,2,\dots,M$ ) 的递推关系。对(5-30)式、(5-31)式取  $z$  变换,可得

$$F_m(z) = F_{m-1}(z) + k_m z^{-1} G_{m-1}(z) \quad (5-35)$$

$$G_m(z) = k_m F_{m-1}(z) + z^{-1} G_{m-1}(z) \quad (5-36)$$

将(5-35)式除以  $F_o(z)$ , (5-36)式除以  $G_o(z)$ ,考虑到(5-34)式的表示方法,可得

$$B_m(z) = B_{m-1}(z) + k_m z^{-1} \bar{B}_{m-1}(z) \quad (5-37)$$

$$\bar{B}_m(z) = k_m B_{m-1}(z) + z^{-1} \bar{B}_{m-1}(z) \quad (5-38)$$

或者,反过来得到

$$B_{m-1}(z) = \frac{1}{1 - k_m^2} [B_m(z) - k_m \bar{B}_m(z)] \quad (5-39)$$

$$\bar{B}_{m-1}(z) = \frac{1}{1 - k_m^2} [-z k_m B_m(z) + z \bar{B}_m(z)] \quad (5-40)$$

这四个式子给出了格型结构中从高阶到低一阶或从低阶到高一阶的系统函数的递推关系。应当注意,  $B_M(z) = B(z)$ , 即包含有  $B(z)$  在内。

下面再将这四个关系式加以推导, 得出  $B_m(z)$  与  $B_{m-1}(z)$  的互相递推关系。由 (5-34) 式知

$$B_0(z) = \bar{B}_0(z) = 1 \quad (5-41)$$

因而将它代入(5-37)式、(5-38)式, 令  $m=1$ , 可得

$$B_1(z) = B_0(z) + k_1 z^{-1} \bar{B}_0(z) = 1 + k_1 z^{-1}$$

$$\bar{B}_1(z) = k_1 B_0(z) + z^{-1} \bar{B}_0(z) = k_1 + z^{-1}$$

也就是满足

$$\bar{B}_1(z) = z^{-1} B_1(z^{-1}) \quad (5-42)$$

同样, 令  $m=2, 3, \dots, M$ , 代入(5-37)式、(5-38)式, 就不难推出

$$\bar{B}_m(z) = z^{-m} B_m(z^{-1}) \quad (5-43)$$

将(5-43)式分别代入(5-37)式、(5-39)式, 可得

$$B_m(z) = B_{m-1}(z) + k_m z^{-m} B_{m-1}(z^{-1}) \quad (5-44)$$

$$B_{m-1}(z) = \frac{1}{1 - k_m^2} [B_m(z) - k_m z^{-m} B_m(z^{-1})] \quad (5-45)$$

这是两个重要的从低阶到高阶或从高阶到低阶的递推关系。注意, 这里有  $M$  阶 FIR 系统的  $B(z)$ 。

2. 其次, 我们来直接给出格型结构的反射系数与横式滤波器各系数的关系。将 (5-34a) 式的  $B_m(z) = 1 + \sum_{i=1}^m b_i^{(m)} z^{-i}$  以及  $B_{m-1}(z) = 1 + \sum_{i=1}^{m-1} b_i^{(m-1)} z^{-i}$  分别代入(5-44)式及(5-45)式, 利用待定系数法, 可分别得两组递推关系

$$\left. \begin{aligned} b_m^{(m)} &= k_m \\ b_i^{(m)} &= b_i^{(m-1)} + k_m b_{m-i}^{(m-1)} \end{aligned} \right\} \quad (5-46)$$

$$\left. \begin{aligned} k_m &= b_m^{(m)} \\ b_i^{(m-1)} &= \frac{1}{1 - k_m^2} [b_i^{(m)} - k_m b_{m-i}^{(m)}] \end{aligned} \right\} \quad (5-47)$$

以上两式中  $i=1, 2, \dots, (m-1)$ ;  $m=2, \dots, M$ 。

3. 综上, 当给出  $H(z)=B(z)=B_M(z)$  时, 可按以下步骤求出  $k_1, k_2, \dots, k_M$ :

(1) 由(5-46)式求出  $k_M=b_M^{(M)}$ ;

(2) 从(5-47)式, 由  $k_M$  及系数  $b_1^{(M)}, b_2^{(M)}, \dots, b_M^{(M)}$  求出  $B_{M-1}(z)$  的系数  $b_1^{(M-1)}, b_2^{(M-1)}, \dots, b_{M-1}^{(M-1)}=k_{M-1}$ , 或者由(5-45)式直接求出  $B_{M-1}(z)$ , 则  $k_{M-1}=b_{M-1}^{(M-1)}$ ;

(3) 重复(2), 可全部求出  $k_M, k_{M-1}, \dots, k_1, B_{M-1}(z), \dots, B_1(z)$ 。

[例 5-1] 一个 FIR 系统的系统函数为

$$H(z) = (1 - 0.8e^{j\frac{\pi}{4}}z^{-1})(1 - 0.8e^{-j\frac{\pi}{4}}z^{-1})(1 - 0.7z^{-1})$$

试求其格型结构。

解

$$\begin{aligned} H(z) &= (1 - 1.1313708z^{-1} + 0.64z^{-2})(1 - 0.7z^{-1}) \\ &= 1 - 1.8313708z^{-1} + 1.4319595z^{-2} - 0.448z^{-3} \end{aligned}$$

这是一个三阶系统,因而

$$b_1^{(3)} = -1.8313708, b_2^{(3)} = 1.4319595, b_3^{(3)} = -0.448, k_3 = b_3^{(3)} = -0.448$$

按照(5-47)式,可知

$$\begin{aligned} b_1^{(2)} &= \frac{1}{1-k_3^2}[b_1^{(3)} - k_3 b_2^{(3)}] = \frac{1}{0.799296}[-1.1898529] = -1.4886262 \\ b_2^{(2)} &= \frac{1}{1-k_3^2}[b_2^{(3)} - k_3 b_1^{(3)}] = \frac{0.6115053}{0.799296} = 0.7650549 \\ k_2 &= b_2^{(2)} = 0.7650549 \end{aligned}$$

因而

$$B_2(z) = 1 - 1.4886262z^{-1} + 0.7650549z^{-2}$$

同样可得

$$\begin{aligned} b_1^{(1)} &= \frac{1}{1-k_2^2}[b_1^{(2)} - k_2 b_2^{(2)}] = \frac{-0.3497454}{0.414691} = -0.8433879 \\ k_1 &= -0.8433879 \end{aligned}$$

因而

$$B_1(z) = 1 - 0.8433879z^{-1}$$

图 5-30 给出了此例题的格型结构。

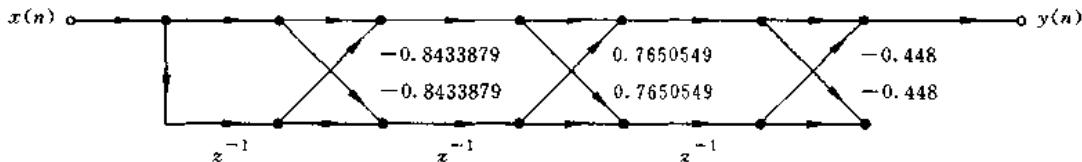


图 5-30 例 5-1 中的全零点 FIR 滤波器的格型结构

## 二、全极点系统(IIR 系统)的格型结构

全极点 IIR 滤波器的系统函数  $H(z)$  可表示为(递归结构)

$$H(z) = \frac{1}{A(z)} = \frac{1}{1 + \sum_{i=1}^M a_i^{(M)} z^{-i}} \quad (5-48)$$

其中  $a_i^{(M)}$  表示  $M$  阶全极点系统的第  $i$  个系数。下面来讨论格型结构和  $a_i^{(M)}$  的关系。

将(5-30)式加以变化后,把(5-30)式、(5-31)式重写如下:

$$f_{m-1}(n) = f_m(n) - k_m g_{m-1}(n-1), \quad m = 1, 2, \dots, M \quad (5-49)$$

$$g_m(n) = k_m f_{m-1}(n) + g_{m-1}(n-1), \quad m = 1, 2, \dots, M \quad (5-50)$$

这就是全极点 IIR 系统格型结构的基本单元,可用图 5-31 表示,这时  $f_m(n)$  是上支路的输入信号,  $f_{m-1}(n)$  是上支路的输出信号,而  $g_m(n)$  是下支路的输入信号,  $g_{m-1}(n)$  是下支路的输出信号。假定所给系统是  $M$  阶系统,并令  $x(n) = f_M(n)$ ,  $f_0(n) = g_0(n) = y(n)$ ,由图 5-31 作为基本单元所构成的全极点格型结构为图 5-32 所示。

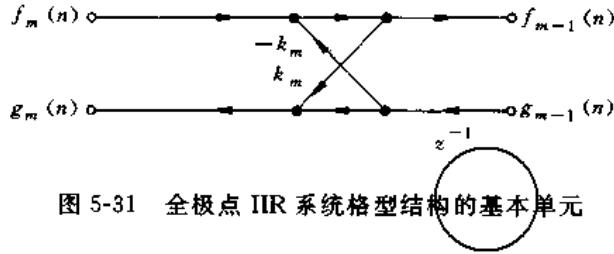


图 5-31 全极点 IIR 系统格型结构的基本单元

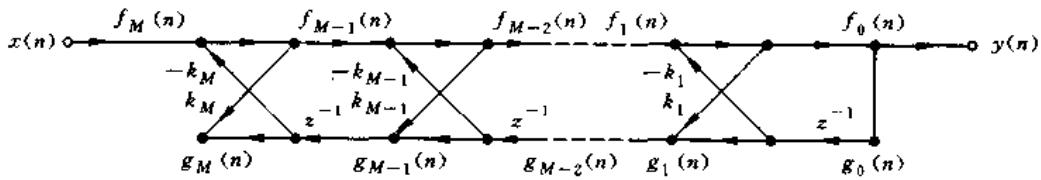


图 5-32 全极点 IIR 系统的格型结构

下面,我们利用与全零点格型滤波器推导中的相同方法,来导出图 5-32 的全极点格型滤波器的系统函数,并导出利用 IIR 滤波器的递归结构的系数  $a_i^{(M)}$ (见(5-48)式)求得参数  $k_1, k_2, \dots, k_M$  的方法。我们采用的是递推的方法。

(1) 令图 5-32 中  $M=1$ ,即对应于一阶的全极点格型结构,由(5-49)式及(5-50)式( $m=M=1$ )可知

$$f_0(n) = f_1(n) - k_1 g_0(n-1) \quad (5-51)$$

$$g_1(n) = k_1 f_0(n) + g_0(n-1) \quad (5-52)$$

由于一阶情况下(见图 5-32, $M=1$ )有

$$f_0(n) = g_0(n) = y(n)$$

$$f_1(n) = x(n)$$

则(5-51)式及(5-52)式可写成

$$y(n) = f_1(n) - k_1 y(n-1) = x(n) - k_1 y(n-1) \quad (5-53)$$

$$g_1(n) = k_1 y(n) + y(n-1) \quad (5-54)$$

可以看出,(5-53)式表示  $x(n)$  为输入,  $y(n)$  为输出的一阶 IIR 系统,(5-54)式则表示  $y(n)$  [ $g_0(n)$ ] 为输入,  $g_1(n)$  为输出的一阶 FIR 系统,由(5-53)式取  $z$  变换可得

$$\frac{Y(z)}{F_1(z)} = \frac{1}{1 + k_1 z^{-1}}$$

令

$$1 + k_1 z^{-1} = A_1(z)$$

则

$$\frac{Y(z)}{F_1(z)} = \frac{1}{1 + k_1 z^{-1}} = \frac{1}{A_1(z)}$$

同样,将(5-54)式取  $z$  变换可得

$$\frac{G_1(z)}{Y(z)} = k_1 + z^{-1} = z^{-1}(1 + k_1 z) = z^{-1}A_1(z^{-1})$$

令

$$z^{-1}A_1(z^{-1}) = \bar{A}_1(z)$$

则

$$\frac{G_1(z)}{Y(z)} = \bar{A}_1(z) = z^{-1}A_1(z^{-1})$$

(2) 上面是一阶系统的推导。下面讨论二阶全极点格型结构。在图 5-32 中令  $M=2$ , 则由(5-49)式及(5-50)式可得

$$f_1(n) = f_2(n) - k_2 g_1(n-1) \quad (5-55)$$

$$g_2(n) = k_2 f_1(n) + g_1(n-1) \quad (5-56)$$

在此二阶情况下有

$$\begin{aligned} f_0(n) &= g_0(n) = y(n) \\ f_2(n) &= x(n) \end{aligned}$$

考虑到(5-51)式及(5-52)式, 则可将(5-55)式及(5-56)式变成

$$\begin{aligned} y(n) &= -k_1(1+k_2)y(n-1) - k_2y(n-2) + f_2(n) \\ &= -k_1(1+k_2)y(n-1) - k_2y(n-2) + x(n) \end{aligned} \quad (5-57)$$

$$g_2(n) = k_2y(n) + k_1(1+k_2)y(n-1) + y(n-2) \quad (5-58)$$

这里可明显看出, (5-57)式表示一个二阶 IIR 系统, 输入是  $x(n)$  [ $f_2(n)$ ], 输出是  $y(n)$ ; 而(5-58)式表示一个二阶 FIR 系统, 输入是  $y(n)$ , 输出是  $g_2(n)$ 。对(5-57)式取  $z$  变换, 可得

$$\frac{Y(z)}{F_2(z)} = \frac{1}{1 + k_1(1+k_2)z^{-1} + k_2z^{-2}}$$

令

$$1 + k_1(1+k_2)z^{-1} + k_2z^{-2} = A_2(z) \quad (5-59)$$

则有

$$\frac{Y(z)}{F_2(z)} = \frac{1}{1 + k_1(1+k_2)z^{-1} + k_2z^{-2}} = \frac{1}{A_2(z)} \quad (5-60)$$

同样, 对(5-58)式取  $z$  变换, 可得

$$\frac{G_2(z)}{Y(z)} = k_2 + k_1(1+k_2)z^{-1} + z^{-2}$$

令

$$k_2 + k_1(1+k_2)z^{-1} + z^{-2} = \bar{A}_2(z)$$

则有

$$z^{-2}A_2(z^{-1}) = \bar{A}_2(z) \quad (5-61)$$

从而可得

$$\frac{G_2(z)}{Y(z)} = z^{-2}A_2(z^{-1}) = \bar{A}_2(z) \quad (5-62)$$

(3) 由此类推, 若定义

$$\frac{Y(z)}{F_m(z)} = \frac{1}{A_m(z)}, \quad \frac{G_m(z)}{Y(z)} = \bar{A}_m(z) \quad (5-63)$$

则有

$$\bar{A}_m(z) = z^{-m}A_m(z^{-1}) \quad (5-64)$$

且有

$$H(z) = \frac{Y(z)}{X(z)} = \frac{Y(z)}{F_M(z)} = \frac{1}{A_M(z)} = \frac{1}{1 + \sum_{i=1}^M a_i^{(M)} z^{-i}} \quad (5-65)$$

由此看出,图 5-32 对应的是一个全极点的 IIR 系统的格型结构。和全零点 FIR 系统的格型结构图的图 5-28 相比较,由于两个结构的基本差分方程((5-30)式、(5-31)式与(5-49)式、(5-50)式)是一样的,所以系数  $k_1, k_2, \dots, k_M$  以及  $a_i^{(m)}$  ( $i=1, 2, \dots, m; m=1, 2, \dots, M$ ) 与 FIR 系统的格型结构的计算方法是一样的,不同之处是这里(全极点系统)的系数是  $a_i^{(m)}$ ,代替了全零点系统中的系数  $b_i^{(m)}$ 。

[例 5-2] 一个全极点系统的系统函数为

$$H(z) = \frac{1}{1 - 1.8313708z^{-1} + 1.4319595z^{-2} - 0.448z^{-3}}$$

求此全极点系统的格型结构。

解 这个例子的分母多项式和全零点系统的多项式是完全相同的,因而求解方法完全和上一个例子相同。

即  $k_3 = -0.448, k_2 = 0.7650549, k_1 = -0.8433879$

其格型结构如图 5-33 所示。

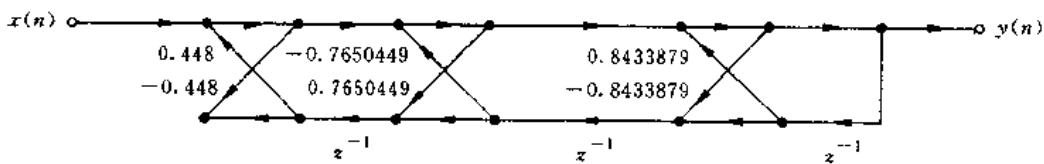


图 5-33 例 5-2 的全极点 IIR 滤波器的格型结构

### 三、零、极点系统(IIR 系统)的格型结构

我们知道,一个在有限  $z$  平面( $0 < |z| < \infty$ )既有极点又有零点的 IIR 系统的系统函数  $H(z)$  可表示为

$$H(z) = \frac{B(z)}{A(z)} = \frac{\sum_{i=0}^N b_i^{(N)} z^{-i}}{1 + \sum_{k=1}^M a_k^{(N)} z^{-k}} \quad (5-66)$$

这一系统的格型结构如图 5-34 所示。由图 5-34 可看出:

(1) 若  $k_1 = k_2 = \dots = k_N = 0$ , 即所有乘  $k$ (或  $-k$ ) 处的联线全断开, 则图 5-34 就变成一个  $N$  阶 FIR 系统的横向结构。

(2) 若  $c_1 = c_2 = \dots = c_N = 0$ , 即含  $c_1 \dots c_N$  的联线都断开,  $c_0 = 1$ , 那么图 5-34 就变成全极点 IIR 格型滤波器结构。(此时,若  $M=N$ , 则图 5-32 和图 5-34 完全相同)。

(3) 因此,图 5-34 的上半部分对应于全极点系统  $\frac{1}{A(z)} = \frac{F_0(z)}{X(z)}$  下半部分对应于全零点系统  $B(z)$ , 但下半部分无任何反馈, 故参数  $k_1, k_2, \dots, k_N$  仍可按全极点系统的方法求出。但上半部分对下半部分有影响, 所以这里的  $c_i$  和全零点系统的  $b_i$  不会相同, 我们的任务是想办法求出各个  $c_i, i=0, 1, \dots, N$ 。

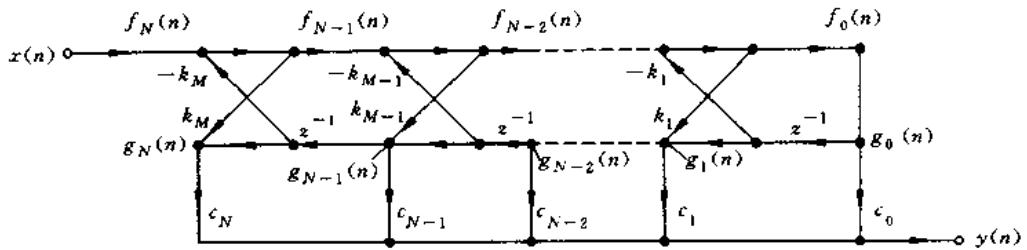


图 5-34 零-极点 IIR 系统的格型结构

由(5-63)式,有

$$\bar{A}_m(z) = \frac{G_m(z)}{G_0(z)} \quad (\text{注(5-63)式中的 } Y(z) \text{ 相当于图 5-34 的 } G_0(z))$$

所以

$$G_m(z) = G_0(z)\bar{A}_m(z), \quad m = 1, 2, \dots, N \quad (5-67)$$

$\bar{A}_m(z)$ 是由  $g_0(n)$ 至  $g_m(n)$ 之间的系统函数,如果令  $\bar{H}_m(z)$ 为由  $x(n)$ 至  $g_m(n)$ 之间的系统函数,考虑到(5-67)式,则

$$\bar{H}_m(z) = \frac{G_m(z)}{X(z)} = \frac{G_0(z)\bar{A}_m(z)}{X(z)} \quad (5-68)$$

因为

$$\frac{F_0(z)}{X(z)} = \frac{G_0(z)}{X(z)} = \frac{1}{A(z)} \quad (5-69)$$

将(5-69)式代入(5-68)式,可得

$$\bar{H}_m(z) = \frac{\bar{A}_m(z)}{A(z)} \quad (5-70)$$

从图 5-34 看出,整个系统的系统函数  $H(z) = \frac{B(z)}{A(z)}$  应是  $\bar{H}_0(z), \bar{H}_1(z), \dots, \bar{H}_N(z)$  分别用  $c_0, c_1, \dots, c_N$  加权后的相加(并联)

$$H(z) = \sum_{m=0}^N c_m \bar{H}_m(z) = \sum_{m=0}^N \frac{c_m \bar{A}_m(z)}{A(z)} = \sum_{m=0}^N \frac{c_m z^{-m} A_m(z^{-1})}{A(z)} = \frac{B(z)}{A(z)} \quad (5-71)$$

这里已代入(5-64)式。

注意,在求解  $k_1, k_2, \dots, k_N$  时,将同时产生出  $A_m(z)$  和  $\bar{A}_m(z)$ 。

(5-71)式给我们提供了以下两种求解参数  $c_1, c_2, \dots, c_N$  的方法。

第一种方法,以  $N=2$  为例,则有

$$B(z) = b_0^{(2)} + b_1^{(2)} z^{-1} + b_2^{(2)} z^{-2}$$

那么考虑到(5-71)式,代入  $N=2$ ,则有

$$B(z) = c_0 A_0(z^{-1}) + c_1 z^{-1} A_1(z^{-1}) + c_2 z^{-2} A_2(z^{-2}) \quad (5-72)$$

其中

$$A_0(z^{-1}) = 1, \quad A_1(z^{-1}) = 1 + a_1^{(1)} z, \quad A_2(z^{-1}) = 1 + a_1^{(2)} z + a_2^{(2)} z^2$$

令(5-72)等式两边的同次幂的系数相等,可得

$$c_0 + c_1 a_1^{(1)} + c_2 a_2^{(2)} = b_0^{(2)}$$

$$c_1 + c_2 a_1^{(2)} = b_1^{(2)}$$

$$c_2 = b_2^{(2)}$$

因此,由下到上可依次求得  $c_2, c_1, c_0$ 。一般情况下(任意  $N$  时)有

$$c_k = b_k^{(N)} - \sum_{m=k+1}^N c_m a_{m-k}^{(m)}, \quad k = 0, 1, \dots, N \quad (5-73)$$

第二种方法,由(5-71)式可得

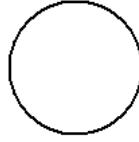
$$B(z) = B_N(z) = \sum_{m=0}^N c_m z^{-m} A_m(z^{-1}) \quad (5-74)$$

再定义

$$B_{N-1}(z) = \sum_{m=0}^{N-1} c_m z^{-m} A_m(z^{-1})$$

则有

$$B_{N-1}(z) = B_N(z) - c_N z^{-N} A_N(z^{-1}) \quad (5-75)$$



对一般项  $m$  来说,可有

$$B_{m-1}(z) = B_m(z) - c_m z^{-m} A_m(z^{-1}) \quad (5-76)$$

由此得出

$$c_m = b_m^{(m)}, \quad m = 0, 1, \dots, N \quad (5-77)$$

起始条件为  $c_N = b_N^{(N)}$ , 其中  $b_N^{(N)}$  是已给定的(见(5-66)式)。

式中  $b_m^{(m)}$  是多项式  $B_m(z)$  的最高阶次  $z^{-m}$  的系数。由于在求  $k_1, k_2, \dots, k_N$  时  $A_m(z)$  已逐次产生,因此,用(5-76)式求系数  $c_m$  是很方便的。

**[例 5-3]** 已知

$$H(z) = \frac{1 - 0.5z^{-1} + 0.2z^{-2} + 0.7z^{-3}}{1 - 1.8313708z^{-1} + 1.4319595z^{-2} - 0.448z^{-3}}$$

试求这个零极点 IIR 滤波器的格型结构。

解 这个例子中的极点部分与全极点型的例 5-2 一样,按上面的说明,这里的  $k_1, k_2, \dots, k_N$  可按全极点型的方法求出,实际上就是按例 5-1 的求全零点型的方法求解,把那里的  $b_i^{(M)}$  换成  $a_i^{(M)}$  即可。例 5-1 求出的  $k_1, k_2, k_3$  及  $a_1^{(2)}, a_2^{(2)}, a_3^{(1)}$ (在那里是  $b_1^{(2)}, b_2^{(2)}, b_3^{(1)}$ ) 为

$$k_1 = -0.8433879, \quad k_2 = 0.7650549, \quad k_3 = -0.448$$

$$a_1^{(2)} = -1.4886262, \quad a_2^{(2)} = 0.7650549, \quad a_3^{(1)} = -0.8433879$$

由  $H(z)$  可知

$$a_1^{(3)} = -1.8313708, \quad a_2^{(3)} = 1.4319595, \quad a_3^{(3)} = -0.448$$

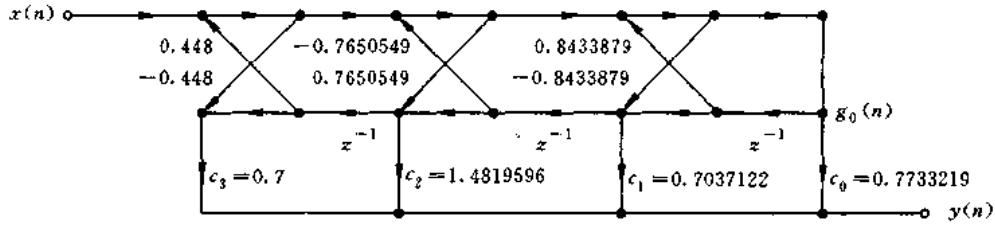


图 5-35 例 5-3 的零极点 IIR 滤波器的格型结构

利用(5-73)式可求出各  $c_i$  为

$$c_3 = b_3^{(3)} = 0.7$$

$$c_2 = b_2^{(3)} - c_3 a_1^{(3)} = 1.4819596$$

$$c_1 = b_1^{(3)} - c_2 a_1^{(2)} - c_3 a_2^{(3)} = 0.7037122$$

$$c_0 = b_0^{(3)} - c_1 a_1^{(1)} - c_2 a_2^{(2)} - c_3 a_3^{(3)} = 0.7733219$$

其格型结构如图 5-35 所示。如果用(5-76)式、(5-77)式来求各系数  $c_i (i=0,1,2,3)$ , 结果是一样的, 读者可自行求解。

## 习 题

1. 用直接 I 型及典范型结构实现以下系统函数:

$$H(z) = \frac{3 + 4.2z^{-1} + 0.8z^{-2}}{2 + 0.6z^{-1} - 0.4z^{-2}}$$

2. 用级联型结构实现以下系统函数:

$$H(z) = \frac{4(z+1)(z^2 - 1.4z + 1)}{(z - 0.5)(z^2 + 0.9z + 0.8)}$$

试问一共能构成几种级联型网络。

3. 给出以下系统函数的并联型实现:

$$H(z) = \frac{5.2 + 1.58z^{-1} + 1.41z^{-2} - 1.6z^{-3}}{(1 - 0.5z^{-1})(1 + 0.9z^{-1} + 0.8z^{-2})}$$

4. 用横截型结构实现以下系统函数:

$$H(z) = \left(1 - \frac{1}{2}z^{-1}\right)\left(1 + 6z^{-1}\right)\left(1 - 2z^{-1}\right)\left(1 + \frac{1}{6}z^{-1}\right)\left(1 - z^{-1}\right)$$

5. 已知 FIR 滤波器的单位冲激响应为

$$h(n) = \delta(n) + 0.3\delta(n-1) + 0.72\delta(n-2) + 0.11\delta(n-3) + 0.12\delta(n-4)$$

试画出其级联型结构实现。

6. 用频率抽样结构实现以下系统函数:

$$H(z) = \frac{5 - 2z^{-3} - 3z^{-6}}{1 - z^{-1}}$$

抽样点数  $N=6$ , 修正半径  $r=0.9$ 。

7. 设某 FIR 数字滤波器的系统函数为

$$H(z) = \frac{1}{5}(1 + 3z^{-1} + 5z^{-2} + 3z^{-3} + z^{-4})$$

试画出此滤波器的线性相位结构。

8. 设滤波器差分方程为

$$y(n) = x(n) + x(n-1) + \frac{1}{3}y(n-1) + \frac{1}{4}y(n-2)$$

- (1) 试用直接 I 型、典范型及一阶节的级联型、一阶节的并联型结构实现此差分方程。

- (2) 求系统的频率响应(幅度及相位)。

- (3) 设抽样频率为 10kHz, 输入正弦波幅度为 5, 频率为 1kHz, 试求稳态输出。

9. 写出图 P5-9 所示结构的系统函数及差分方程。

10. 已知

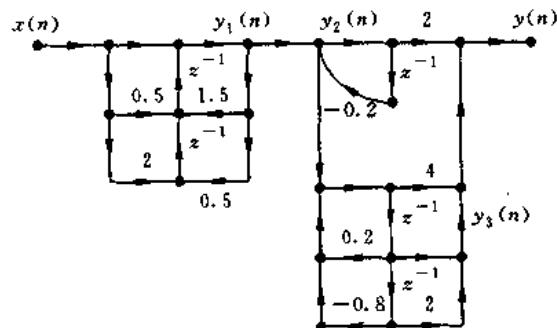
$$H(z) = 1 - 0.4z^{-1} - 0.8z^{-2} + 0.86z^{-3}$$

试求此滤波器的格型结构各系数, 并画出信号流图。

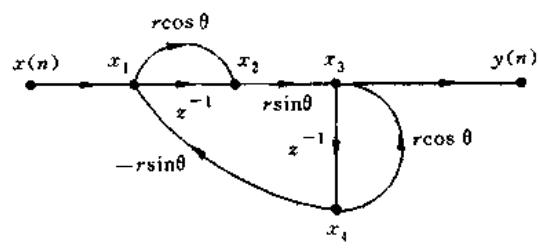
11. 已知

$$H(z) = \frac{1 + 0.85z^{-1} - 0.42z^{-2} + 0.34z^{-3}}{1 - 0.6z^{-1} - 0.78z^{-2} + 0.48z^{-3}}$$

试求此系统的格型结构。



(a)



(b)

图 P5-9

12. 已知全零 FIR 格型滤波器各系数为  $k_1 = 0.6125, k_2 = 0.6722, k_3 = -0.5822$ , 试求三阶 FIR 滤波器直接型结构的各系数  $b_i^{(3)}, i=1,2,3 (b_0=1)$ 。

13. 已知零极点格型滤波器各系数为  $k_1 = -0.7251, k_2 = 0.8142, k_3 = -0.6582, c_0 = 1.2543, c_1 = -0.7284, c_2 = -2.15, c_3 = -0.692$ , 试求三阶 IIR 滤波器直接型结构

$$H(z) = \frac{\sum_{i=0}^3 b_i z^{-i}}{1 + \sum_{i=1}^3 a_i z^{-i}}$$

的各系数  $a_i, i=1,2,3$

$b_i, i=0,1,2,3$



# 第六章 无限长单位冲激响应(IIR)数字滤波器的设计方法

## 6.1 引言

一般情况下,数字滤波器是一个线性移不变离散时间系统,利用有限精度算法来实现。数字滤波器的设计一般包括:

- (1) 按照任务的要求,确定滤波器的性能要求;
- (2) 用一个因果稳定的离散线性移不变系统的系统函数去逼近这一性能要求。系统函数有无限长单位冲激响应(IIR)系统函数及有限长单位冲激响应(FIR)系统函数两种;
- (3) 利用有限精度算法来实现这个系统函数。这里包括选择运算结构(如第五章的各种基本结构)、选择合适的字长(包括系数量化及输入变量、中间变量和输出变量的量化)以及有效数字的处理方法(舍入、截尾)等(这将在第八章中讨论);
- (4) 实际的技术实现,包括采用通用计算机软件或专用数字滤波器硬件来实现,或采用专用的或通用的数字信号处理器来实现。

本章和下一章讨论第(2)项内容,即逼近性能要求问题或系统函数的设计问题。

一般来说,滤波器的性能要求往往以频率响应的幅度特性的允许误差来表征。以低通滤波器为例,如图 6-1 所示,频率响应有通带、过渡带及阻带三个范围(而不是理想的陡截止的通带、阻带两个范围)。在通带内,幅度响应以误差  $\alpha_1$  逼近于 1, 即

$$1 - \alpha_1 \leq |H(e^{j\omega})| \leq 1, \quad |\omega| \leq \omega_c \quad (6-1a)$$

在阻带中,幅度响应以误差小于  $\alpha_2$  而逼近于零,即

$$|H(e^{j\omega})| \leq \alpha_2, \quad \omega_n \leq |\omega| \leq \pi \quad (6-1b)$$

其中  $\omega_c, \omega_n$  分别为通带截止频率和阻带截止频率,它们都是数字域频率。为了逼近理想低通滤波器特性,还必须有一个非零宽度( $\omega_n - \omega_c$ )的过渡带,在这个过渡带内的频率响应平滑地从通带下降到阻带。

虽然给出了通带的容限  $\alpha_1$  及阻带的容限  $\alpha_2$ ,但是,在具体技术指标中往往使用通带允许的最大衰减(波纹) $\delta_1$  及阻带应达到的最小衰减  $\delta_2$ 。 $\delta_1, \delta_2$  的定义为

$$\delta_1 = 20 \lg \frac{|H(e^{j0})|}{|H(e^{j\omega_c})|} = -20 \lg |H(e^{j\omega_c})| = -20 \lg(1 - \alpha_1) \quad (6-2a)$$

$$\delta_2 = 20 \lg \frac{|H(e^{j0})|}{|H(e^{j\omega_n})|} = -20 \lg |H(e^{j\omega_n})| = -20 \lg \alpha_2 \quad (6-2b)$$

式中,假定  $|H(e^{j0})| = 1$  已被归一化了。例如  $|H(e^{j\omega})|$  在  $\omega_c$  处满足  $|H(e^{j\omega_c})| = 0.707$ , 则  $\delta_1 = 3\text{dB}$ ; 在  $\omega_n$  处满足  $|H(e^{j\omega_n})| = 0.001$ , 则  $\delta_2 = 60\text{dB}$ 。

当然,也可给出相位的逼近要求或给出时域冲激响应的逼近要求。

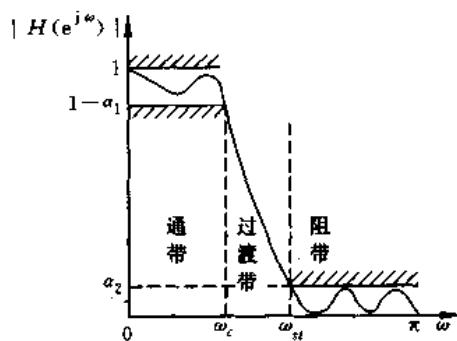


图 6-1 理想低通滤波器逼近的误差容限

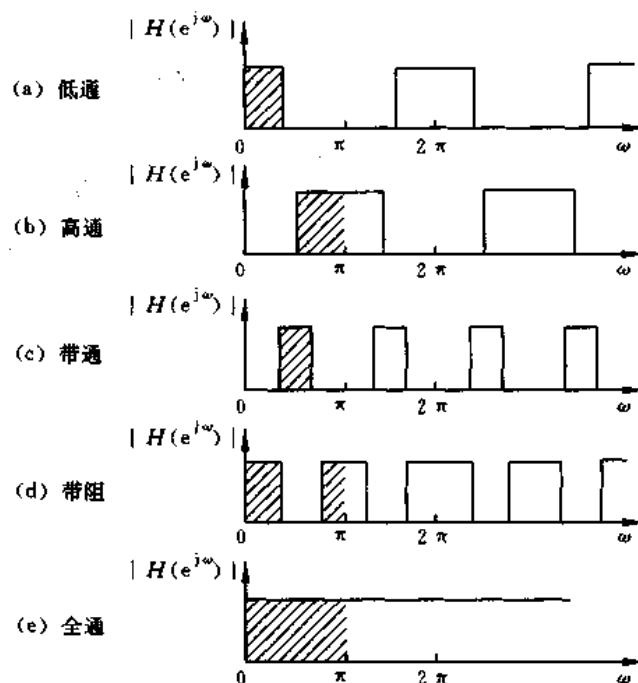


图 6-2 各种数字滤波器的理想幅度频率响应

与模拟滤波器类似，数字滤波器按频率特性划分也有低通、高通、带通、带阻、全通等类型。由于频率响应的周期性，频率变量以数字频率  $\omega$  来表示 ( $\omega = \Omega T = \Omega / f_s$ ,  $\Omega$  为模拟角频率,  $T$  为抽样时间间隔,  $f_s$  为抽样频率)，所以数字滤波器设计中必须给出抽样频率。图 6-2 为各种数字滤波器理想幅度频率响应(只表示了正频率部分)，这样的理想频率响应是不可能实现的，原因是频带之间幅度响应是突变的，因而其单位抽样响应是非因果的。因此要给出如图 6-1 的实际逼近容限。在图 6-2 中， $2\pi$  等于数字域抽样频率  $\omega_s$ ，即

$$\omega_s = \Omega_s T = 2\pi f_s T = 2\pi \frac{f_s}{f_s} = 2\pi, \quad (f_s = \frac{1}{T})$$

$\omega_s/2 = \pi$  是折叠频率。

按照奈奎斯特抽样定理，频率特性只能限于  $|\omega| < \omega_s/2 = \pi$  范围。

数字滤波器的系统函数为  $H(z)$ , 它在  $z$  平面单位圆上的值为滤波器频率响应  $H(e^{j\omega})$ , 表征数字滤波器频率响应特性的三个参量是幅度平方响应、相位响应和群延时响应。这三个参量之所以重要, 是因为一般逼近问题都涉及  $\omega$  的复频率响应  $H(e^{j\omega})$ 。

### 1. 幅度平方响应

当只需要逼近幅度响应而不考虑相位时, 例如标准的低通、高通、带通、带阻滤波器的逼近就是这样, 这时根据幅度平方响应来进行设计是很方便的。幅度平方响应定义为

$$\begin{aligned} |H(e^{j\omega})|^2 &= H(e^{j\omega})H^*(e^{j\omega}) = H(e^{j\omega})H(e^{-j\omega}) \\ &= |H(z)H(z^{-1})|_{z=e^{j\omega}} \end{aligned} \quad (6-3)$$

这里由于冲激响应为实函数, 故满足  $H^*(e^{j\omega}) = H(e^{-j\omega})$ , 也就是满足共轭对称条件。若  $z=re^{j\omega}$  是  $H(z)$  的极点, 则  $z=\frac{1}{r}e^{-j\omega}$  是  $H(z^{-1})$  的极点, 又由于  $H(z)$  的有理表达式中各系数为实数, 因而, 零极点必然都以共轭对形式出现, 故必有  $z=re^{-j\omega}$  和  $z=\frac{1}{r}e^{j\omega}$  两极点存在, 所以  $H(z)H(z^{-1})$  的极点既是共轭的, 又是以单位圆镜像对称的。为了使  $H(z)$  成为可实现的系统, 故只取单位圆内的那些极点作为  $H(z)$  的极点, 单位圆外的极点作为  $H(z^{-1})$  的极点。 $H(z)$  的零点一般不是唯一确定的, 可在  $z$  平面上的任意位置。如果我们选  $H(z)H(z^{-1})$  在  $z$  平面单位圆内的零点作为  $H(z)$  的零点, 则所得到的是最小相位延迟滤波器。

### 2. 相位响应

由于  $H(e^{j\omega})$  是复数, 可表示成

$$H(e^{j\omega}) = |H(e^{j\omega})| e^{j\beta(e^{j\omega})} = \text{Re}[H(e^{j\omega})] + j\text{Im}[H(e^{j\omega})] \quad (6-4)$$

所以

$$\beta(e^{j\omega}) = \arctan \left\{ \frac{\text{Im}[H(e^{j\omega})]}{\text{Re}[H(e^{j\omega})]} \right\} \quad (6-5)$$

由于

$$H^*(e^{j\omega}) = |H(e^{j\omega})| e^{-j\beta(e^{j\omega})}$$

所以又有

$$\begin{aligned} \beta(e^{j\omega}) &= \frac{1}{2j} \ln \left[ \frac{H(e^{j\omega})}{H^*(e^{j\omega})} \right] = \frac{1}{2j} \ln \left[ \frac{H(e^{j\omega})}{H(e^{-j\omega})} \right] \\ &= \frac{1}{2j} \ln \left[ \frac{H(z)}{H(z^{-1})} \right]_{z=e^{j\omega}} \end{aligned} \quad (6-6)$$

### 3. 群延迟响应

它是滤波器平均延迟的一个度量, 定义为相位对角频率的导数的负值, 即

$$\tau(e^{j\omega}) = -\frac{d\beta(e^{j\omega})}{d\omega} \quad (6-7)$$

可以化为

$$\tau(e^{j\omega}) = - \frac{d\beta(z)}{dz} \frac{dz}{d\omega} \Big|_{z=e^{j\omega}} = -jz \frac{d\beta(z)}{dz} \Big|_{z=e^{j\omega}} \quad (6-8)$$

由于

$$\ln[H(e^{j\omega})] = \ln|H(e^{j\omega})| + j\beta(e^{j\omega})$$

所以

$$\beta(e^{j\omega}) = \operatorname{Im}\{\ln[H(e^{j\omega})]\}$$

因而又有

$$\tau(e^{j\omega}) = -\operatorname{Im}\left[\frac{d}{d\omega}\{\ln[H(e^{j\omega})]\}\right] \quad (6-9)$$

同样可化为

$$\begin{aligned} \tau(e^{j\omega}) &= -\operatorname{Im}\left[\frac{d\{\ln[H(z)]\}}{dz} \frac{dz}{d\omega}\right]_{z=e^{j\omega}} \\ &= -\operatorname{Im}\left[jz \frac{d\{\ln[H(z)]\}}{dz}\right]_{z=e^{j\omega}} = -\operatorname{Re}\left[z \frac{d}{dz}\{\ln[H(z)]\}\right]_{z=e^{j\omega}} \\ &= -\operatorname{Re}\left[z \frac{dH(z)}{dz} \frac{1}{H(z)}\right]_{z=e^{j\omega}} \end{aligned} \quad (6-10)$$

当要求滤波器为线性相位响应特性时，则通带内群延迟特性就应是常数。

前面已说到，本章及下一章讨论滤波器设计中的第二个步骤，即用一个因果稳定的离散线性移不变系统的系统函数去逼近给定的性能要求。IIR 滤波器的系统函数为  $z^{-1}$ （或  $z$ ）的有理分式，即

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} \quad (6-11)$$

一般满足  $M \leq N$ ，这类系统称为  $N$  阶系统，当  $M > N$  时， $H(z)$  可看成是一个  $N$  阶 IIR 子系统与一个  $(M-N)$  阶的 FIR 子系统（多项式）的级数。以下讨论都假定  $M \leq N$ 。

IIR 滤波器的逼近问题就是去求出滤波器的各系数  $a_i$  和  $b_k$ ，使得在规定的意义上，例如通带起伏及阻带衰减的要求或采用最优化准则（最小均方误差要求或最大误差最小要求）逼近所要求的特性。这就是数学上的逼近问题。如果在  $s$  平面上去逼近，就得到模拟滤波器，如果在  $z$  平面上去逼近，则得到数字滤波器。

设计 IIR 数字滤波器一般有以下两种方法：

(1) 先设计一个合适的模拟滤波器，然后转换成满足预定指标的数字滤波器。这种方法很方便，这是因为模拟滤波器已经具有很多简单而又现成的设计公式，并且设计参数已经表格化了，设计起来既方便又准确。

(2) 计算机辅助设计法。这是一种最优化设计法。先确定一种最优准则，例如设计

出的实际频率响应幅度  $|H(e^{j\omega})|$  与所要求的理想频率响应幅度  $|H_d(e^{j\omega})|$  的均方误差最小准则, 或它们的最大误差最小准则等, 然后求在此最佳准则下滤波器系统函数的系数  $a_k, b_k$ 。这种设计一般得不到滤波器系数作为所要求的理想频率响应的闭合形式的函数表达式, 而是需要进行大量的迭代运算, 故离不开计算机。

## 6.2 最小与最大相位延时系统, 最小与最大相位超前系统

在第二章中已说到, 任一个线性移不变系统, 可用系统函数表示为(见(2-76)式、(2-77)式)

$$H(z) = \frac{\sum_{m=0}^M b_m z^{-m}}{\sum_{k=0}^N a_k z^{-k}} = K \frac{\prod_{m=1}^M (1 - c_m z^{-1})}{\prod_{k=1}^N (1 - d_k z^{-1})} = K z^{N-M} \frac{\prod_{m=1}^M (z - c_m)}{\prod_{k=1}^N (z - d_k)} \quad (6-12)$$

在第二章中, 还导出了系统频率响应的表达式(见(2-86)式、(2-87)式及(2-88)式)

$$H(e^{j\omega}) = K e^{j\omega(N-M)} \frac{\prod_{m=1}^M (e^{j\omega} - c_m)}{\prod_{k=1}^N (e^{j\omega} - d_k)} = |H(e^{j\omega})| e^{j\arg[H(e^{j\omega})]} \quad (6-13)$$

由于冲激响应为实数, 故  $K$  只能是实数(正数或负数), 它对幅角只引入固定值(0 或  $\pi$  弧度)。我们来研究  $\frac{H(e^{j\omega})}{K}$ 。

$\frac{H(e^{j\omega})}{K}$  的模  $\left| \frac{H(e^{j\omega})}{K} \right|$  为

$$\left| \frac{H(e^{j\omega})}{K} \right| = \frac{\prod_{m=1}^M |e^{j\omega} - c_m|}{\prod_{k=1}^N |e^{j\omega} - d_k|} = \frac{\text{各零矢量模的连乘积}}{\text{各极矢量模的连乘积}} \quad (6-14)$$

其中, “零矢”(极矢)是指零点(极点)指向  $z$  平面单位圆上要研究的频率点( $\omega$  幅角)的矢量。下面我们着重讨论相位的影响。

$\frac{H(e^{j\omega})}{K}$  的相角  $\arg\left[\frac{H(e^{j\omega})}{K}\right]$  为

$$\begin{aligned} \arg\left[\frac{H(e^{j\omega})}{K}\right] &= \sum_{m=1}^M \arg[e^{j\omega} - c_m] - \sum_{k=1}^N \arg[e^{j\omega} - d_k] + (N - M)\omega \\ &= \text{各零矢量辐角之和} - \text{各极矢量辐角之和} + (N - M)\omega \end{aligned} \quad (6-15)$$

若某一零点(或极点)位于单位圆内, 当  $\omega$  从 0 变到  $2\pi$  时, 即在  $z$  平面单位圆上正向(逆时针)旋转一周时, 零矢(或极矢)变化为  $2\pi$  弧度。

若某一零点(或极点)位于单位圆外, 当  $\omega$  从 0 变到  $2\pi$  时, 即在  $z$  平面单位圆上正向(逆时针)旋转一周时, 零矢(或极矢)变化为零。所以当  $\omega$  从 0 变到  $2\pi$  时, 只有单位圆内

的零点、极点对  $\arg\left[\frac{H(e^{j\omega})}{K}\right]$  有影响。我们若用  $m_i, m_o$  分别表示单位圆内与单位圆外的零点数, 以  $p_i, p_o$  分别表示单位圆内与单位圆外的极点数, 由(6-12)式可看出

$$M = m_i + m_o, \quad N = p_i + p_o$$

下面具体讨论零点、极点的分布对系统相角的影响。

(1) 对因果稳定系统、系统函数  $H(z)$  的全部极点在单位圆内, 收敛域是半径为  $r(r < 1)$  的某个圆的外部, 且满足  $n < 0$  时,  $h(n) = 0$ 。此时

$$p_o = 0, \quad p_i = N$$

当  $\omega$  从 0 变到  $2\pi$  时,  $\Delta\omega = 2\pi$ , 则  $\frac{H(e^{j\omega})}{K}$  的辐角的变化量为

$$\begin{aligned} \Delta\arg\left[\frac{H(e^{j\omega})}{K}\right] \Big|_{\Delta\omega=2\pi} &= 2\pi[m_i - p_i] + 2\pi(N - M) \\ &= 2\pi m_i - 2\pi M = -2\pi m_o \end{aligned}$$

这种系统当  $\omega$  由 0 而增加时, 辐角变化为负, 故称为相位“延时”(或相位“滞后”)系统。又可分为以下两种情况:

① 当全部零点在单位圆内时, 即  $m_i = M, (m_o = 0)$ , 则

$$\Delta\arg\left[\frac{H(e^{j\omega})}{K}\right] \Big|_{\Delta\omega=2\pi} = 0$$


这时, 相位变化最小。我们把这种系统称为最小相位延时系统。当然, 最小相位延时系统一定是因果稳定系统。

② 当全部零点在单位圆外时, 即  $m_i = 0, (m_o = M)$ , 则

$$\Delta\arg\left[\frac{H(e^{j\omega})}{K}\right] \Big|_{\Delta\omega=2\pi} = -2\pi M$$

这时, 相位变化最大, 又是负数。故我们把它称为最大相位延时系统。当然, 它也一定是因果稳定系统。

(2) 对逆因果稳定系统, 系统函数的全部极点在单位圆外, 收敛域是半径为  $r(r > 1)$  的某个圆的内部, 且满足  $n > 0$  时,  $h(n) = 0$ 。此时

$$p_i = 0, \quad p_o = N$$

当  $\omega$  从 0 变到  $2\pi$  时,  $\Delta\omega = 2\pi$ , 则  $\frac{H(e^{j\omega})}{K}$  的辐角的变化量为

$$\Delta\arg\left[\frac{H(e^{j\omega})}{K}\right] \Big|_{\Delta\omega=2\pi} = 2\pi m_i + 2\pi(N - M)$$

一般来说, 系统总满足  $N > M$ (也就是  $H(z)$  的分子  $z^{-1}$  的阶次小于分母  $z^{-1}$  的阶次), 因而这种系统当  $\omega$  由 0 而增加时, 辐角变化为正, 故称为相位“超前”(或相位“领先”)系统。又可分为以下两种情况:

① 当全部零点在单位圆内时, 即  $m_i = M, (m_o = 0)$ , 则

$$\Delta\arg\left[\frac{H(e^{j\omega})}{K}\right] \Big|_{\Delta\omega=2\pi} = 2\pi N = 2\pi p_o$$

这时, 相位变化最大。称为最大相位超前系统。当然, 它也一定是逆因果稳定系统。

② 当全部零点在单位圆外时, 即  $m_o = M, (m_i = 0)$ , 则

$$\Delta \arg \left[ \frac{H(e^{j\omega})}{K} \right] \Big|_{\Delta\omega=2\pi} = 2\pi(N - M) = 2\pi(p_o - m_o)$$

这时, 相位超前量最小, 故称为最小相位超前系统。当然, 它一定也是逆因果稳定系统。

与以上四种系统相对应的单位冲激响应分别有四种相应的序列。例如, 最小相位延时系统的单位冲激响应称为最小相位延时序列。

我们把以上四种系统及其因果性、稳定性、零点、极点的关系归纳在表 6-1 中。

表 6-1 四种系统的归纳

系 统	因 果 性	稳 定 性	零 点	极 点
最小相位延时系统	因 果	稳 定	单位圆内	单位圆内
最大相位延时系统	因 果	稳 定	单位圆外	单位圆内
最小相位超前系统	反 因 果	稳 定	单位圆外	单位圆外
最大相位超前系统	反 因 果	稳 定	单位圆内	单位圆外

由于最小相位延时系统(有时简称为最小相位系统)在通信中有重要的地位, 因而把它的一些重要性质归纳如下(其中有的性质的证明可见本章的习题):

(1) 在傅里叶变换  $H(e^{j\omega})$  相同的所有系统中, 最小相位系统具有最小的相位滞后, 即它的负的相位最小(相位绝对值最小)。

(2) 按照帕塞瓦定理由于傅里叶变换幅度相同的各系统的总能量应当相同, 但最小相位延时系统  $h_{\min}(n)$  的能量集中在  $n=0$  附近, 一般系统  $h(n)$  的能量则集中在  $n>0$  处, 也就是说, 如果  $h_{\min}(n), h(n)$  是  $N$  点有限长序列( $n=0, 1, \dots, N-1$ ), 则有

$$\sum_{n=0}^m |h(n)|^2 < \sum_{n=0}^N |h_{\min}(n)|^2, \quad m < N-1$$

$$\sum_{n=0}^{N-1} |h(n)|^2 = \sum_{n=0}^{N-1} |h_{\min}(n)|^2$$

(3) 由上一条关系可得出, 对相同傅里叶变换幅度的各序列, 最小相位序列的  $h_{\min}(0)$  最大(可用初值定理加以证明):

$$h_{\min}(0) > h(0)$$

(4) 在幅度响应  $|H(e^{j\omega})|$  相同的系统中, 只有唯一的一个最小相位延时系统。

(5) 利用级联全通函数的办法, 可将最小相位系统的零点反射到单位圆外, 而构成幅度响应相同的非最小相位延时系统(参见 6.3 节“全通系统”)。

### 6.3 全通系统

全通系统是指系统频率响应的幅度在所有频率  $\omega$  下均为 1 或某一常数的系统。如果我们令  $H_{ap}(z)$  表示全通系统的系统函数, 则对于所有  $\omega$ , 此系统的频率响应  $H_{ap}(e^{j\omega}) = H_{ap}(z) \Big|_{z=e^{j\omega}}$  都应满足

$$|H_{ap}(e^{j\omega})| = 1$$

简单的一阶全通系统的系统函数为

$$H_{ap}(z) = \frac{z^{-1} - a}{1 - az^{-1}}, \quad a \text{ 实数}, 0 < |a| < 1 \quad (6-16)$$

这一系统所对应的零点-极点位置如图 6-3 所示。

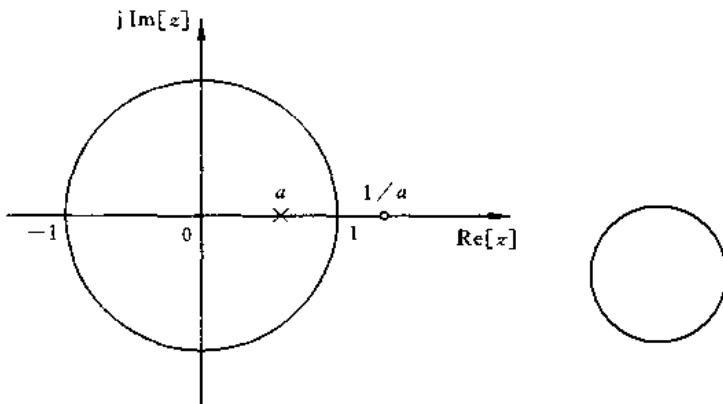


图 6-3  $a$  为实数,  $0 < |a| < 1$  时, 一阶全通系统的零点-极点位置

高阶有理全通系统是由一串一阶系统组成的, 这些一阶系统可以包括如(6-16)式的实零点-实极点的一阶系统, 还可以包括复数零点-复数极点对, 多个这两类(或只有一类)系统的级联就组成一个高阶全通网络系统。复数零点-极点的全通节的系统函数为

$$\frac{z^{-1} - a^*}{1 - az^{-1}}, \quad a \text{ 复数}, 0 < |a| < 1 \quad (6-17)$$

这一全通节系统函数的零点-极点出现在共轭反演位置上或说在镜像位置上(以单位圆为“镜子”), 如图 6-4 所示。

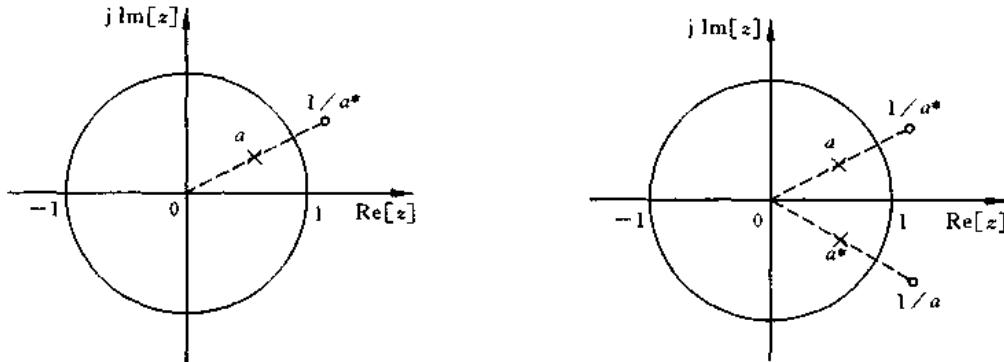


图 6-4  $a$  为复数,  $0 < |a| < 1$  时, 一阶全通节的零点-极点位置

图 6-5 二阶全通系统的零点-极点位置

当然,  $h(n)$  是实函数, 则  $H(z)$  是实系数有理系统函数, 其系数都应为实数, 因而其系统函数的复数极点(零点)必须以共轭对形式出现。例如, 一个实系数有理二阶全通系统函数应为

$$H_{ap}(z) = \frac{z^{-1} - a^*}{1 - az^{-1}} \cdot \frac{z^{-1} - a}{1 - a^* z^{-1}}, \quad |a| < 1 \quad (6-18)$$

也就是说, 它是由两个一阶全通节(其极点(零点)成共轭对)的级联组成的, 如图 6-5 所示。

可以证明(6-16)式、(6-17)式的一阶全通节在任何频率上, 其频率响应的模都为 1。例如, 对(6-17)式, 设  $a=re^{j\theta}$  为某一复数, 在任一频率  $\omega$  处研究, 令  $z=e^{j\omega}$ , 则有

$$\begin{aligned} \left| \frac{z^{-1} - a^*}{1 - az^{-1}} \right|_{z=e^{j\omega}} &= \left| \frac{e^{-j\omega} - re^{-j\theta}}{1 - re^{j\theta} \cdot e^{-j\omega}} \right| = |e^{-j\omega}| \cdot \left| \frac{1 - re^{j(\omega-\theta)}}{1 - re^{-j(\omega-\theta)}} \right| \\ &= \left| \frac{1 - r\cos(\omega - \theta) - j\sin(\omega - \theta)}{1 - r\cos(\omega - \theta) + j\sin(\omega - \theta)} \right| = 1 \end{aligned}$$

一般来说,  $N$  阶数字全通滤波器的系统函数可表示为

$$\begin{aligned} H(z) &= \pm \prod_{k=1}^N \frac{z^{-1} - a_k^*}{1 - a_k z^{-1}} = \pm \frac{d_N + d_{N-1} z^{-1} + \dots + d_1 z^{-(N-1)} + z^{-N}}{1 + d_1 z^{-1} + \dots + d_{N-1} z^{-(N-1)} + d_N z^{-N}} \\ &= \pm \frac{z^{-N} D(z^{-1})}{D(z)} \end{aligned} \quad (6-19)$$

式中

$$D(z) = 1 + d_1 z^{-1} + \dots + d_{N-1} z^{-(N-1)} + d_N z^{-N} \quad (6-20)$$

$D(z)$  为具有实系数的多项式, 其根全在单位圆内, 当  $z=e^{j\omega}$  时, 满足

$$\text{○} \quad D(e^{j\omega}) = D^*(e^{-j\omega}) \quad (6-21)$$

所以有

$$|H(e^{j\omega})| = 1 \quad (6-22)$$

$H(z)$  满足全通系统的要求。同时  $D(z)$  的根( $H(z)$  的极点)应是单位圆内的共轭复数  $z_p = r_1 e^{\pm j\omega_1}$ , ( $r_1 < 1$ ), 而  $D(z^{-1})$  的根( $H(z)$  的零点)则在单位圆在外, 是它必然  $D(z)$  的根的“镜像”, 即  $z_0 = \frac{1}{r_1} e^{\pm j\omega_1}$ 。

下面讨论全通系统的应用。

(1) 任何一个因果稳定的(非最小相位延时)系统  $H(z)$  都可以表示成全通系统  $H_{ap}(z)$  和最小相位延时系统  $H_{min}(z)$  的级联, 即

$$H(z) = H_{min}(z) \cdot H_{ap}(z) \quad (6-23)$$

这一结论可用以下例子加以证明。设有一个因果稳定的非最小相位延时系统  $H(z)$ , 它的某一对共轭零点位于单位圆外  $z = \frac{1}{z_0}$  及  $z = \frac{1}{z_0^*}$  处,  $|z_0| < 1$ , 而其余的零点位于单位圆内。当然, 这一系统的极点必然全在单位圆内。我们可将  $H(z)$  表示成

$$H(z) = H_1(z) \cdot (z^{-1} - z_0) \cdot (z^{-1} - z_0^*) \quad (6-24)$$

其中  $H_1(z)$  是最小相位延时的, 另两个乘因子代表了单位圆外的一对共轭零点。我们也可将(6-24)式表示成

$$H(z) = H_1(z) \cdot (z^{-1} - z_0) \cdot (z^{-1} - z_0^*) \frac{1 - z_0^* z^{-1}}{1 - z_0^* z^{-1}} \cdot \frac{1 - z_0 z^{-1}}{1 - z_0 z^{-1}}$$

$$= H_1(z) \cdot (1 - z_0^* z^{-1}) \cdot (1 - z_0 z^{-1}) \cdot \frac{z^{-1} - z_0}{1 - z_0^* z^{-1}} \cdot \frac{z^{-1} - z_0^*}{1 - z_0 z^{-1}} \quad (6-25)$$

由于  $|z_0| < 1$ , 所以因子  $H_1(z) \cdot (1 - z_0^* z^{-1}) \cdot (1 - z_0 z^{-1})$  是最小相位延时的, 而因子  $\frac{z^{-1} - z_0}{1 - z_0^* z^{-1}} \cdot \frac{z^{-1} - z_0^*}{1 - z_0 z^{-1}}$  是两个全通函数的级联, 一定是一个全通系统, 所以可将(6-13)式表示成

$$H(z) = H_{\min}(z) \cdot H_{ap}(z)$$

这就是(6-23)式。其中

$$H_{\min}(z) = H_1(z) \cdot (1 - z_0^* z^{-1}) \cdot (1 - z_0 z^{-1})$$

$$H_{ap}(z) = \frac{z^{-1} - z_0}{1 - z_0^* z^{-1}} \cdot \frac{z^{-1} - z_0^*}{1 - z_0 z^{-1}}$$

在这里,  $H(z)$  和  $H_{\min}(z)$  的差别在于把  $H(z)$  的单位圆外的一对零点  $z = \frac{1}{z_0}$ ,  $z = \frac{1}{z_0^*}$  分别反射到单位圆内的“镜像”位置  $z = z_0^*$ ,  $z = z_0$  上, 就构成了  $H_{\min}(z)$  的零点, 可以用图 6-6 来说明这一反射情况。但是, 可以看出  $H(z)$  和  $H_{\min}(z)$  的频率响应的幅度是相同的, 即

$$|H(e^{j\omega})| = |H_{\min}(e^{j\omega})| \cdot |H_{ap}(e^{j\omega})| = |H_{\min}(e^{j\omega})| \quad (6-26)$$

它们之间的差别只是频率响应的相位不同而已。

从以上讨论可以看出, 单位圆外的零点反射到单位圆内镜像点(共轭反演点)上, 是关于单位圆反射的。同样可将最小相位延时系统的一个零点反射到单位圆外而构成另一个幅度函数相同, 相位函数不同的非最小相位延时系统。如果将单位圆内的全部零点反射到单位圆外, 则构成了幅度函数相同的最大相位延时系统。

(2) 如果设计出的滤波器是非稳定的, 则可用级联全通函数的办法将它变成一个稳定的滤波器, 例如原滤波器有一对极点在单位圆外  $z = \frac{1}{r} e^{\pm j\theta}$  处, 则可将此滤波器级联一个全通系统

$$H_{ap}(z) = \frac{z^{-1} - r e^{j\theta}}{1 - r e^{-j\theta} z^{-1}} \cdot \frac{z^{-1} - r e^{-j\theta}}{1 - r e^{j\theta} z^{-1}}$$

这样可以将单位圆外的一对极点加以抵消, 同时又不改变滤波器的幅度特性。

(3) 可以作为相位均衡器(群延时均衡器)用。IIR 滤波器其相位特性是非线性的, 因而群延时不为常数, 而在视频信号的传输中希望系统具有线性相位, 因而采用全通滤波器作为相位均衡器, 来校正系统的非线性相位, 以得到线性相位, 同时又不改变系统的幅度特性。

设全通滤波器为  $H_{ap}(z)$ , 系统(例如 IIR 滤波器)为  $H_d(z)$ , 则级联后系统  $H(z)$  为

$$H(z) = H_{ap}(z) H_d(z)$$

即

$$H(e^{j\omega}) = H_{ap}(e^{j\omega}) \cdot H_d(e^{j\omega}) = |H_{ap}(e^{j\omega})| \cdot |H_d(e^{j\omega})| \cdot e^{j[\varphi_{ap}(\omega) + \varphi_d(\omega)]}$$

相位关系为

$$\gamma(\omega) = \varphi_{ap}(\omega) + \varphi_d(\omega)$$

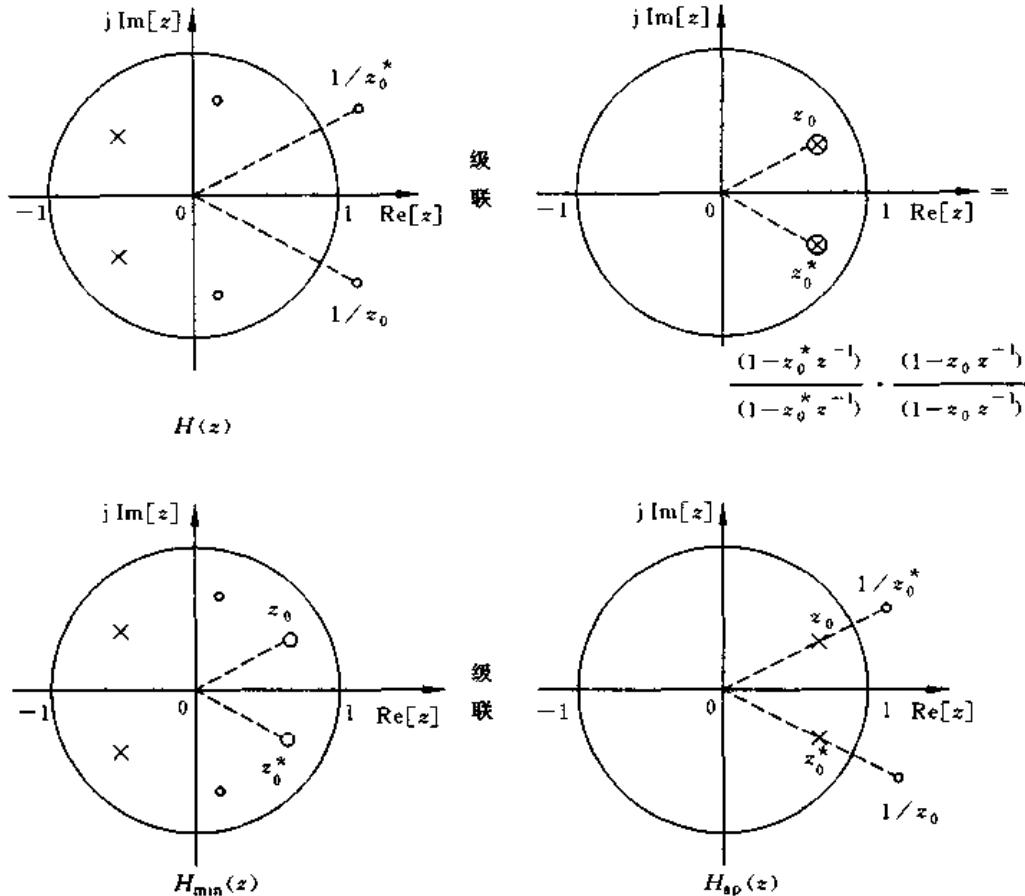


图 6-6 系统的等效变换

按  $\tau(\omega) = -\frac{d\varphi(\omega)}{d\omega}$  可得群延时关系为

$$\tau(\omega) = \tau_{ap}(\omega) + \tau_d(\omega)$$

我们希望, 在通带中满足  $\tau(\omega) = \tau_{ap}(\omega) + \tau_d(\omega) = \tau_0$ ,  $\tau_0$  是常数(不随  $\omega$  而变化)。则逼近误差的平方值为

$$e^2 = [\tau(\omega) - \tau_0]^2 = [\tau_{ap}(\omega) + \tau_d(\omega) - \tau_0]^2$$

$e^2$  是频率  $\omega$ 、全通函数极点和系数的函数( $\tau_d(\omega)$  为已知), 利用均方误差最小准则,(见第六章的最优化设计)可求得均衡器(全通函数)的有关参数。

#### 6.4 用模拟滤波器设计 IIR 数字滤波器

利用模拟滤波器设计数字滤波器就是要把  $s$  平面映射到  $z$  平面, 使模拟系统函数  $H_a(s)$  变换成所需的数字滤波器的系统函数  $H(z)$ 。这种由复变量  $s$  到复变量  $z$  之间的映射(变换)关系, 必须满足以下两条基本要求: 第一,  $H(z)$  的频率响应要能模仿  $H_a(s)$  的频率响应, 即  $s$  平面的虚轴  $j\omega$  必须映射到  $z$  平面上的单位圆  $e^{j\omega}$  上, 也就是频率轴要对应。第二, 因果稳定的  $H_a(s)$  应能映射成因果稳定的  $H(z)$ 。也就是  $s$  平面的左半平面

$\operatorname{Re}[s] < 0$  必须映射到  $z$  平面单位圆的内部  $|z| < 1$ 。

我们知道，“模拟原型”滤波器有多种设计方法，例如巴特沃思型滤波器、切贝雪夫型滤波器、椭圆函数型（考尔型）滤波器等。

从模拟滤波器映射成数字滤波器，也就是使数字滤波器能模仿模拟滤波器的特性，主要的有以下几种映射方法：冲激响应不变法、阶跃响应不变法、双线性变换法。

## 6.5 冲激响应不变法

### 一、变换原理

冲激响应不变法是使数字滤波器的单位冲激响应序列  $h(n)$  模仿模拟滤波器的单位冲激响应  $h_a(t)$ 。将模拟滤波器的单位冲激响应加以等间隔的抽样，使  $h(n)$  正好等于  $h_a(t)$  的抽样值，即满足

$$h(n) = h_a(nT) \quad (6-27)$$

其中  $T$  是抽样周期。

如果令  $H_a(s)$  是  $h_a(t)$  的拉普拉斯变换， $H(z)$  为  $h(n)$  的  $z$  变换，利用第二章 2.5 节抽样序列的  $z$  变换与模拟信号的拉普拉斯变换的关系，即利用(2-53)式，得

$$H(z)|_{s=e^{j\omega}} = \frac{1}{T} \sum_{k=-\infty}^{\infty} H_a\left(s - j\frac{2\pi k}{T}\right) \quad (6-28)$$

则可看出，冲激响应不变法将模拟滤波器的  $s$  平面变换成数字滤波器的  $z$  平面上，这个从  $s$  到  $z$  的变换  $z = e^{j\omega}$  正是第二章 2.5 节中从  $s$  平面变换成  $z$  平面上的变换关系(2-51)式。

如图 6-7 所示， $s$  平面上每一条宽度为  $2\pi/T$  的横条都将重叠地映射到整个  $z$  平面上，而每一横条的左半边映射到  $z$  平面单位圆以内，右半边映射到  $z$  平面单位圆以外，而  $s$  平面虚轴 ( $j\Omega$  轴) 映射到  $z$  平面单位圆上，虚轴上每一段长为  $2\pi/T$  的线段都映射到  $z$  平面单位圆上一周。由于  $s$  平面每一横条都要重叠地映射到  $z$  平面上，这正好反映了  $H(z)$  是和  $H_a(s)$  的周期延拓序列之间有变换关系  $z = e^{j\omega}$ ，故冲激响应不变法并不相当于从  $s$  平面到  $z$  平面上的简单代数映射关系。

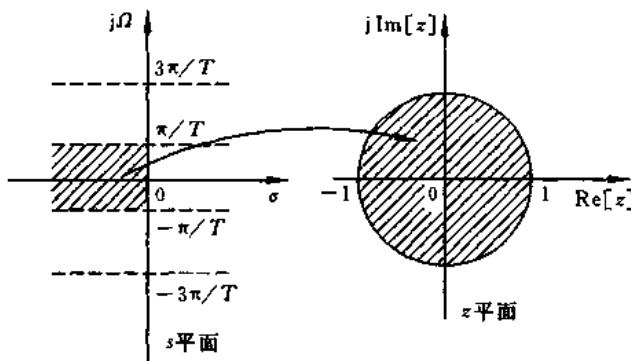


图 6-7 冲激响应不变法映射关系

## 二、混叠失真

由(6-28)式知,数字滤波器的频率响应和模拟滤波器的频率响应间的关系为

$$H(e^{j\omega}) = \frac{1}{T} \sum_{k=-\infty}^{\infty} H_a\left(j \frac{\omega - 2\pi k}{T}\right) \quad (6-29)$$

这就是说,数字滤波器的频率响应是模拟滤波器频率响应的周期延拓。因而正如第一章1.4节抽样定理所讨论的,只有当模拟滤波器的频率响应是限带的,且带限于折叠频率以内时,即

$$h_a(j\Omega) = 0, \quad |\Omega| \geq \frac{\pi}{T} = \frac{\Omega_c}{2} \quad (6-30)$$

才能使数字滤波器的频率响应在折叠频率以内重现模拟滤波器的频率响应而不产生混叠失真,即

$$H(e^{j\omega}) = \frac{1}{T} H_a\left(j \frac{\omega}{T}\right), \quad |\omega| < \pi \quad (6-31)$$

但是,任何一个实际的模拟滤波器频率响应都不是严格限带的,变换后就会产生周期延拓分量的频谱交叠,即产生频率响应的混叠失真,因而模拟滤波器的频率响应在折叠频率以上处衰减越大、越快,变换后频率响应混叠失真就越小。

对某一频率响应的系统由单位冲激响应进行抽样,抽样频率为  $f_s$ ,若使  $f_s$  增加,即令抽样时间间隔( $T=1/f_s$ )减小,则系统频率响应各周期延拓分量之间相距更远,因而可减小频率响应的混叠效应。

我们还应注意,在冲激响应不变法设计中,当滤波器的指标用数字域频率  $\omega$  给定时,若  $\omega_c$  不变,用减小  $T$  的方法就不能解决混叠问题。例如设计某一截止频率为  $\omega_c$  的低通滤波器,则要求与之相对应的模拟滤波器的截止频率为

$$\Omega_c = \frac{\omega_c}{T}$$

因而模拟折叠角频率  $\Omega$  的带域是  $[-\frac{\pi}{T}, \frac{\pi}{T}]$ ,随着  $T$  的减小,它会增加。而为了  $\omega_c$  不变,  
 $T$  减小时,  $\Omega_c$  也应增加,所以如果原来  $H_a(s)$  的截止频率  $\Omega_c > \frac{\pi}{T}$ ,即在  $[-\frac{\pi}{T}, \frac{\pi}{T}]$  域外  
 $H_a(s)$  的值不为零,则不论如何减小  $T$ ,由于要求  $\Omega_c$  与  $T$  有同样倍数的变化(以使  $\omega_c$  不变),故总有  $\Omega_c > \pi/T$ ,不能解决混叠问题。

## 三、模拟滤波器的数字化方法

由于冲激响应不变法要由模拟系统函数  $H_a(s)$  求拉普拉斯反变换得到模拟的冲激响应  $h_a(t)$ ,然后抽样后得到  $h(n) = h_a(nT)$ ,再取  $z$  变换得  $H(z)$ ,过程较复杂。而且由这一变换过程看出,它对部分分式表达的模拟系统函数更为方便。下面我们来讨论冲激响应不变法所造成的  $s$  平面和  $z$  平面的对应关系。

设模拟滤波器的系统函数  $H_a(s)$  只有单阶极点,且假定分母的阶次大于分子的阶次(一般都满足这一要求,因为只有这样才相当于一个稳定的模拟系统)。因此可将  $H_a(s)$

展开成部分分式表示式

$$H_a(s) = \sum_{k=1}^N \frac{A_k}{s - s_k} \quad (6-32)$$

其相应的冲激响应  $h_a(t)$  是  $H_a(s)$  的拉普拉斯反变换, 即

$$h_a(t) = \mathcal{L}^{-1}[H_a(s)] = \sum_{k=1}^N A_k e^{s_k t} u(t)$$

其中  $u(t)$  是连续时间的单位阶跃函数。在冲激响应不变法中, 要求数字滤波器的单位抽样响应等于  $h_a(t)$  的抽样, 即

$$h(n) = h_a(nT) = \sum_{k=1}^N A_k e^{s_k nT} u(n) = \sum_{k=1}^N A_k (e^{s_k T})^n u(n) \quad (6-33)$$

对  $h(n)$  求  $z$  变换, 即得数字滤波器的系统函数

$$\begin{aligned} H(z) &= \sum_{n=-\infty}^{\infty} h(n) z^{-n} = \sum_{n=0}^{\infty} \sum_{k=1}^N A_k (e^{s_k T} z^{-1})^n = \sum_{k=1}^N A_k \sum_{n=0}^{\infty} (e^{s_k T} z^{-1})^n \\ &= \sum_{k=1}^N \frac{A_k}{1 - e^{s_k T} z^{-1}} \end{aligned} \quad (6-34)$$

将(6-32)式的  $H_a(s)$  和(6-34)式的  $H(z)$  加以比较, 可以看出:

- (1)  $s$  平面上的单极点  $s = s_k$  变换到  $z$  平面上  $z = e^{s_k T}$  处的单极点;
- (2)  $H_a(s)$  与  $H(z)$  的部分分式的系数是相同的, 都是  $A_k$ ;
- (3) 如果模拟滤波器是稳定的, 即所有极点  $s_k$  位于  $s$  平面的左半平面, 即极点的实部小于零  $\operatorname{Re}[s_k] < 0$ , 则变换后的数字滤波器的全部极点在单位圆内, 即模小于 1,  $|e^{s_k T}| = e^{\operatorname{Re}(s_k)T} < 1$ , 因此数字滤波器也是稳定的;
- (4) 虽然冲激响应不变法能保证  $s$  平面极点与  $z$  平面极点有这种代数对应关系, 但是并不等于整个  $s$  平面与  $z$  平面有这种代数对应关系, 特别是数字滤波器的零点位置就与模拟滤波器零点位置没有这种代数对应关系, 而是随  $H_a(s)$  的极点  $s_k$  以及系数  $A_k$  两者而变化。

从(6-31)式看出, 数字滤波器频率响应还与抽样间隔  $T$  成反比, 如果抽样频率很高, 即  $T$  很小, 则滤波器增益会太高, 这很不好, 因而希望数字滤波器的频率响应不随抽样频率而变化, 故作以下修正, 令

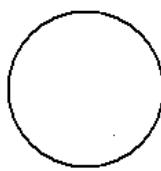
$$h(n) = Th_a(nT) \quad (6-35)$$

则有

$$H(z) = \sum_{k=1}^N \frac{TA_k}{1 - e^{s_k T} z^{-1}} \quad (6-36)$$

及

$$\begin{aligned} H(e^{j\omega}) &= \sum_{k=-\infty}^{\infty} H_a\left(j \frac{\omega}{T} - j \frac{2\pi k}{T}\right) \\ &\approx H_a\left(j \frac{\omega}{T}\right), \quad |\omega| < \pi \end{aligned} \quad (6-37)$$



由于  $h_a(t)$  是实数, 因而  $H_a(s)$  的极点必成共轭对存在, 即若  $s = s_k$  为极点, 其留数为  $A_k$ , 则必有  $s = s_k^*$  亦为极点, 且其留数为  $A_k^*$ 。因而这样一对共轭极点, 其  $H_a(s)$  变成

$H(z)$  关系为

$$\frac{A_k}{s - s_k} \rightarrow \frac{A_k}{1 - z^{-1} e^{j\omega T}}, \quad \frac{A_k^*}{s - s_k^*} \rightarrow \frac{A_k^*}{1 - z^{-1} e^{j\omega T}}$$

[例 6-1] 设模拟滤波器的系统函数为

$$H_a(s) = \frac{2}{s^2 + 4s + 3} = \frac{1}{s+1} - \frac{1}{s+3}$$

试利用冲激响应不变法,设计 IIR 数字滤波器。

解 直接利用(6-36)式可得到数字滤波器的系统函数为

$$H(z) = \frac{T}{1 - z^{-1} e^{-T}} - \frac{T}{1 - z^{-1} e^{-3T}} = \frac{Tz^{-1}(e^{-T} - e^{-3T})}{1 - z^{-1}(e^{-T} + e^{-3T}) + z^{-2} e^{-4T}}$$

设  $T=1$ , 则有

$$H(z) = \frac{0.318z^{-1}}{1 - 0.4177z^{-1} + 0.01831z^{-2}}$$

模拟滤波器的频率响应  $H_a(j\Omega)$  以及数字滤波器的频率响应  $H(e^{j\omega})$  分别为

$$H_a(j\Omega) = \frac{2}{(3 - \Omega^2) + j4\Omega}$$

$$H(e^{j\omega}) = \frac{0.3181e^{-j\omega}}{1 - 0.4177e^{-j\omega} + 0.01831e^{-j2\omega}}$$

$|H_a(j\Omega)|$  和  $|H(e^{j\omega})|$  画在图 6-8 上,由该图可看出,由于  $H_a(j\Omega)$  不是充分限带的,所以  $H(e^{j\omega})$  产生了很大的频谱混叠失真。

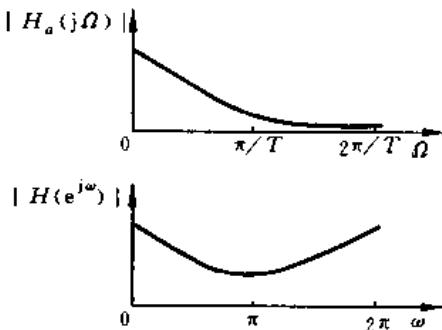


图 6-8 冲激响应不变法的频率响应幅度

#### 四、优缺点

从以上讨论可以看出,冲激响应不变法使得数字滤波器的冲激响应完全模仿模拟滤波器的冲激响应,也就是时域逼近良好,而且模拟频率  $\Omega$  和数字频率  $\omega$  之间呈线性关系  $\omega = \Omega T$ 。因而一个线性相位的模拟滤波器(例如贝塞尔滤波器)可以映射成一个线性相位的数字滤波器。但是,因为有频率响应混叠效应,所以冲激响应不变法只适用于限带的模拟滤波器,所以高通和带阻滤波器不宜采用冲激响应不变法,否则要加保护滤波器,滤掉高于折叠频率以上的频率。对于带通和低通滤波器,需充分限带,若阻带衰减越大,则

混叠效应越小。

## 6.6 阶跃响应不变法

同样,阶跃响应不变法是使数字滤波器的阶跃响应  $g(n)$  模仿模拟滤波器的阶跃响应  $g_a(t)$ ,即将模拟滤波器的阶跃响应加以等间隔的抽样,使  $g(n)$  正好等于  $g_a(t)$  的抽样值,满足

$$g(n) = g_a(t) |_{t=nT} = g_a(nT) \quad (6-38)$$

其中  $T$  是抽样周期。

设数字滤波器的系统函数为  $H(z)$  ( $H(z) = \mathcal{Z}[h(n)]$ ),如果其输入端作用一个阶跃函数  $u(n)$ ,则其输出端即为阶跃响应  $g(n)$ ,如图 6-9 所示。因而满足

$$g(n) = u(n) * h(n)$$

将此式两端取  $z$  变换可得

$$G(z) = \mathcal{Z}[g(n)] = \frac{z}{z-1} H(z) \quad (6-39)$$

所以

$$H(z) = \frac{z-1}{z} G(z) \quad (6-40)$$

对于模拟滤波器,设其系统函数为  $H_a(s)$  ( $H_a(s) = \mathcal{L}[h_a(t)]$ ),如果输入端作用一个阶跃函数  $u(t)$ ,则其输出端即为阶跃响应  $g_a(t)$ ,如图 6-10 所示。因而满足

$$g_a(t) = u(t) * h_a(t)$$

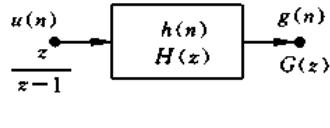


图 6-9 数字滤波器在单位阶跃序列  
作用下的情况

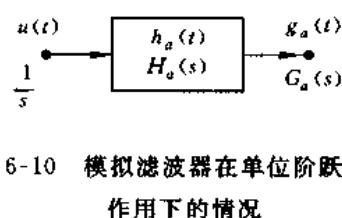


图 6-10 模拟滤波器在单位阶跃信号  
作用下的情况

将此式两端取拉普拉斯变换可得

$$G_a(s) = \mathcal{L}[g_a(t)] = \frac{1}{s} H_a(s) \quad (6-41)$$

即

$$g_a(t) = \mathcal{L}^{-1} \left[ \frac{1}{s} H_a(s) \right] \quad (6-42)$$

要满足阶跃响应不变,则应有

$$g(n) = g_a(nT) = \left[ \mathcal{L}^{-1} \left[ \frac{1}{s} H_a(s) \right] \right]_{t=nT} \quad (6-43)$$

将(6-43)式取  $z$  变换即得  $G(z)$ ,将它代入(6-40)式即得数字滤波器系统函数

$$H(z) = \frac{z-1}{z} \mathcal{Z}\left\{\left[\mathcal{L}^{-1}\left[\frac{1}{s} H_a(s)\right]\right]_{t=nT}\right\} \quad (6-44)$$

这就是阶跃响应不变法由模拟系统函数  $H_a(s)$  映射成数字系统系统函数  $H(z)$  的公式。

下面我们来看阶跃响应不变法是否有频率响应的周期延拓及混叠现象。

与(2-49)式相同, 模拟阶跃响应的理想抽样信号  $\hat{g}_a(t) = g_a(t)\delta_T(t)$  的拉普拉斯变换为

$$\hat{G}_a(s) = \mathcal{L}[\hat{g}_a(t)] = \sum_{n=-\infty}^{\infty} g_a(nT) e^{-snT} \quad (6-45)$$

而阶跃响应序列的  $z$  变换为

$$G(z) = \sum_{n=-\infty}^{\infty} g_a(n) z^{-n} \quad (6-46)$$

比较(6-45)式与(6-46)式, 可知

$$G(z)|_{z=e^{j\omega T}} = \hat{G}_a(s) \quad (6-47)$$

同样, 与(2-52)式相似, 模拟阶跃响应  $g_a(t)$  的拉普拉斯变换  $G_a(s)$  与  $g_a(t)$  的理想抽样信号  $\hat{g}_a(t)$  的拉普拉斯变换  $\hat{G}_a(s)$  的关系为

$$\hat{G}_a(s) = \frac{1}{T} \sum_{k=-\infty}^{\infty} G_a\left(s - jk \frac{2\pi}{T}\right) \quad (6-48)$$

所以, 由(6-47)式和(6-48)式, 并将(6-39)式、(6-41)式代入, 可得

$$\frac{z}{z-1} H(z)|_{z=e^{j\omega T}} = \frac{1}{T} \sum_{k=-\infty}^{\infty} G_a\left(s - jk \frac{2\pi}{T}\right) = \frac{1}{T} \sum_{k=-\infty}^{\infty} \frac{H_a\left(s - jk \frac{2\pi}{T}\right)}{s - jk \frac{2\pi}{T}} \quad (6-49)$$

由(6-49)式看出, 阶跃响应不变法仍然有频率响应的周期延拓现象, 因而如果  $H_a(s)$  不是严格限带的(一般都是如此), 必然会有频率响应的混叠失真, 但是由于有  $1/s$  因子(在频率轴上即为  $1/j\Omega$  因子), 使得频率响应幅度与频率  $\Omega$  成反比, 因而, 随着  $\Omega$  的增加, 频率响应的混叠现象一定比冲激响应不变法的要小。

**[例 6-2]** 已知模拟二阶巴特沃思低通滤波器的归一化(3dB 截止频率为 1) 模拟系统函数  $H'_a(s)$  为

$$H'_a(s) = \frac{1}{1 + 1.4142136s + s^2}$$

而 3dB 截止频率为 50Hz 的模拟滤波器, 需将归一化  $H'_a(s)$  中的  $s$  变量用  $\frac{s}{2\pi \times 50}$  来代替, 即

$$H_a(s) = H'_a\left(\frac{s}{100\pi}\right) = \frac{9.8696044 \times 10^4}{s^2 + 444.28830s + 9.8696044 \times 10^4}$$

设系统抽样频率为  $f_s = 500\text{Hz}$ , 要求从这一低通模拟滤波器设计一个低通数字滤波器, 采用阶跃响应不变法。

解 按(6-41)式, 模拟滤波器阶跃响应的拉普拉斯变换为

$$G_a(s) = \frac{1}{s} H_a(s) = \frac{9.8696044 \times 10^4}{s(s^2 + 444.28830s + 9.8696044 \times 10^4)}$$

$$= \frac{1}{s} - \frac{(s + 222.14415) + 222.14415}{(s + 222.14415)^2 + (222.14415)^2}$$

由于

$$\mathcal{L}[e^{-at}(\sin\Omega_0 t)u(t)] = \frac{\Omega_0}{(s+a)^2 + \Omega_0^2}$$

$$\mathcal{L}[e^{-at}(\cos\Omega_0 t)u(t)] = \frac{s+a}{(s+a^2) + \Omega_0^2}$$

$$\mathcal{L}[u(t)] = \frac{1}{s}$$

故

$$g_a(t) = \mathcal{L}^{-1}[G_a(s)] = \{1 - e^{-222.14415t}[\sin(222.14415t) + \cos(222.14415t)]\}u(t)$$

则

$$g(n) = g_a(nT) = \{1 - e^{-222.14415nT}[\sin(222.14415nT) + \cos(222.14415nT)]\}u(n)$$

利用以下的  $z$  变换关系：

$$\mathcal{Z}[x(n)] = X(z)$$

$$\mathcal{Z}[e^{-naT}x(n)] = X(e^{aT}z)$$

$$\mathcal{Z}[(\sin naT)u(n)] = \frac{z \sin aT}{z^2 - 2z \cos aT + 1}$$

$$\mathcal{Z}[(\cos naT)u(n)] = \frac{z^2 - z \cos aT}{z^2 - 2z \cos aT + 1}$$

$$\mathcal{Z}[u(n)] = \frac{z}{z - 1}$$

且代入  $a = 222.14415$ ,  $T = \frac{1}{f_s} = \frac{1}{500} = 2 \times 10^{-3}s$ , 可得阶跃响应  $g(n)$  的  $z$  变换

$$G(z) = \mathcal{Z}[g(n)] = \frac{z}{z - 1} - \frac{z^2 - 0.30339071z}{z^2 - 1.1580459z + 0.41124070}$$

$$= \frac{0.14534481z^2 + 0.10784999z}{(z - 1)(z^2 - 1.1580459z + 0.41124070)}$$

再由(6-40)式可得数字低通滤波器的系统函数为(用  $z^{-1}$  表示)

$$H(z) = \frac{z - 1}{z} G(z) = \frac{0.14534481z^{-1} + 0.10784999z^{-2}}{1 - 1.1580459z^{-1} + 0.41124070z^{-2}}$$

## 6.7 双线性变换法

冲激响应不变法与阶跃响应不变法是使数字滤波器在时域上模仿模拟滤波器,但是它们的缺点是产生频率响应的混叠失真,这是因为从  $s$  平面到  $z$  平面是多值的映射关系

所造成的。为了克服这一缺点,我们采用双线性变换法。

### 一、变换原理

双线性变换法是使数字滤波器的频率响应与模拟滤波器的频率响应相似的一种变换方法,为了克服多值映射这一缺点,我们首先把整个  $s$  平面压缩变换到某一中介的  $s_1$  平面上的一条横带里(宽度为  $\frac{2\pi}{T}$ ,即从  $-\frac{\pi}{T}$  到  $\frac{\pi}{T}$ ),其次再通过上面讨论过的标准变换关系  $z = e^{j\Omega T}$  将此横带变换到整个  $z$  平面上去,这样就使  $s$  平面与  $z$  平面是一一对应的关系,消除了多值变换性,也就消除了频谱混叠现象,如图 6-11 所示。

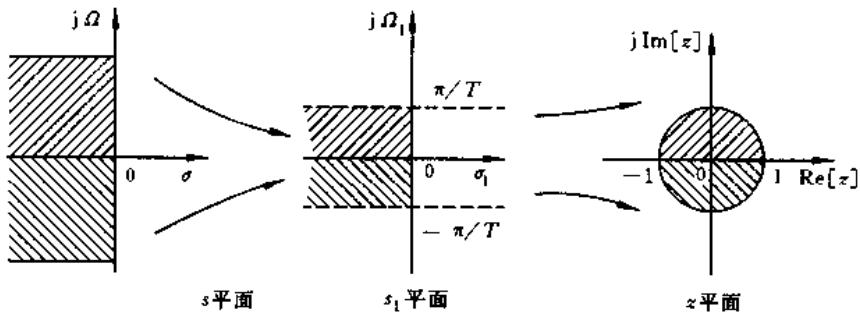


图 6-11 双线性变换法的映射关系

将  $s$  平面整个  $j\Omega$  轴压缩变换到  $s_1$  平面  $j\Omega_1$  轴上的  $-\frac{\pi}{T}$  到  $\frac{\pi}{T}$  一段,可以采用以下变换关系:

$$\Omega = \tan\left(\frac{\Omega_1 T}{2}\right) \quad (6-50)$$

这样,  $\Omega = \pm\infty$  变到  $\Omega_1 = \pm\frac{\pi}{T}$ ,  $\Omega = 0$  变到  $\Omega_1 = 0$ , 可将(6-50)式写成

$$j\Omega = \frac{e^{\frac{j\Omega_1 T}{2}} - e^{-\frac{j\Omega_1 T}{2}}}{e^{\frac{j\Omega_1 T}{2}} + e^{-\frac{j\Omega_1 T}{2}}}$$

解析延拓到整个  $s$  平面和  $s_1$  平面,令  $j\Omega = s$ ,  $j\Omega_1 = s_1$ , 则得

$$s = \frac{e^{\frac{s_1 T}{2}} - e^{-\frac{s_1 T}{2}}}{e^{\frac{s_1 T}{2}} + e^{-\frac{s_1 T}{2}}} = \tanh\left(\frac{s_1 T}{2}\right) = \frac{1 - e^{-s_1 T}}{1 + e^{-s_1 T}} \quad (6-51)$$

再将  $s_1$  平面通过以下标准变换关系映射到  $z$  平面:

$$z = e^{s_1 T} \quad (6-52)$$

从而得到  $s$  平面和  $z$  平面的单值映射关系为

$$s = \frac{1 - z^{-1}}{1 + z^{-1}} \quad (6-53)$$

$$z = \frac{1 + s}{1 - s} \quad (6-54)$$

一般来说,为了使模拟滤波器的某一频率与数字滤波器的任一频率有对应的关系,可

引入待定常数  $c$ ,使(6-50)式和(6-51)式变成

$$\Omega = c \tan\left(\frac{\Omega_1 T}{2}\right) \quad (6-55)$$

$$s = c \operatorname{th}\left(\frac{s_1 T}{2}\right) = c \frac{1 - e^{-s_1 T}}{1 + e^{-s_1 T}} \quad (6-56)$$

仍将  $z = e^{s_1 T}$  代入到(6-56)式,可得

$$s = c \frac{1 - z^{-1}}{1 + z^{-1}} \quad (6-57)$$

$$z = \frac{c + s}{c - s} \quad (6-58)$$

(6-57)式与(6-58)式是  $s$  平面与  $z$  平面之间的单值映射关系,这种变换就称为双线性变换。

## 二、变换常数 $c$ 的选择

用不同的方法选择  $c$  可使模拟滤波器频率特性与数字滤波器频率特性在不同频率处有对应的关系,也就是可以调节频带间的对应关系。选择常数  $c$  的方法有以下两种:

(1) 采用使模拟滤波器与数字滤波器在低频处有较确切的对应关系,即在低频处有  $\Omega \approx \Omega_1$ 。当  $\Omega_1$  较小时有

$$\tan\left(\frac{\Omega_1 T}{2}\right) \approx \frac{\Omega_1 T}{2}$$

由(6-55)式及  $\Omega \approx \Omega_1$  式可得

$$\Omega \approx \Omega_1 \approx c \frac{\Omega_1 T}{2}$$

因而得到

$$c = \frac{2}{T} \quad (6-59)$$

此时,模拟原型滤波器的低频特性近似等于数字滤波器的低频特性。

(2) 采用数字滤波器的某一特定频率(例如截止频率  $\omega_c = \Omega_{ic} T$ )与模拟原形滤波器的一个特定频率  $\Omega_c$  严格相对应,即

$$\Omega_c = c \tan\left(\frac{\Omega_{ic} T}{2}\right) = c \tan\left(\frac{\omega_c}{2}\right)$$

则有

$$c = \Omega_c \cot \frac{\omega_c}{2} \quad (6-60)$$

这一方法的主要优点是在特定的模拟频率和特定的数字频率处,频率响应是严格相等的,因而可以较准确的控制截止频率的位置。

## 三、逼近的情况

(6-57)式、(6-58)式的双线性变换确实符合我们在 6.4 节中提出的变换关系应满足

的两点要求,即

(1) 首先把  $z = e^s$  代入(6-41)式,可得

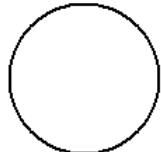
$$s = c \frac{1 - e^{-j\omega}}{1 + e^{-j\omega}} = j\omega \tan\left(\frac{\omega}{2}\right) = j\Omega \quad (6-61)$$

即  $s$  平面的虚轴确实与  $z$  平面的单位圆相对应。

(2) 其次,将  $s = \sigma + j\Omega$  代入(6-42)式中,得

$$z = \frac{c+s}{c-s} = \frac{(c+\sigma)+j\Omega}{(c-\sigma)-j\Omega}$$

因此



$$|z| = \sqrt{\frac{(c+\sigma)^2 + \Omega^2}{(c-\sigma)^2 + \Omega^2}}$$

由此看出,当  $\sigma < 0$  时,  $|z| < 1$ ; 当  $\sigma > 0$  时,  $|z| > 1$ ; 当  $\sigma = 0$  时,  $|z| = 1$ 。也就是说,  $s$  平面的左半平面映射到  $z$  平面的单位圆内,  $s$  平面的右半平面映射到  $z$  平面的单位圆外,  $s$  平面的虚轴映射到  $z$  平面的单位圆上。因此,稳定的模拟滤波器经双线性变换后所得的数字滤波器也一定是稳定的。

#### 四、优缺点

除了以上两点满足要求外,双线性变换最大的优点是避免了频率响应的混叠现象。讨论如下:

(6-61)式表示了模拟角频率  $\Omega$  与数字频率  $\omega$  之间的变换关系,重写如下:

$$\Omega = \omega \tan\left(\frac{\omega}{2}\right) \quad (6-62)$$

它表明  $s$  平面与  $z$  平面是单值的一一对应关系,  $s$  平面整个  $j\Omega$  轴单值对应于  $z$  平面单位圆的一周,即频率轴是单值变换关系。这个关系表示在图 6-12 上。

由图 6-12 看出,  $s$  平面的正、负虚轴分别映射成  $z$  平面单位圆的上半部分(辐角  $\omega$  为正)及下半部分(辐角  $\omega$  为负),频率轴是单值变换关系,且  $\Omega \rightarrow \infty$  时,  $\omega = \pi$  为折叠频率,故不会有高于折叠频率的分量,这就避免了冲激响应不变法的频率响应混叠现象。

但是又产生了新的问题,除了在零频率附近,(6-62)式的频率变换关系接近于线性关系外,当  $\Omega$  增加时,变换关系就是非线性的了,也就是频率  $\Omega$  与  $\omega$  之间存在着严重的非线性关系,如(6-62)式及图 6-12 所示。

由于这种频率之间的非线性变换关系,就产生了问题,首先,一个线性相位的模拟滤波器经双线性变换后就得到非线性相位的数字滤波器,不再保持原有的线性相位了;其次,这种非线性关系就要求模拟滤波器的幅频响应必须是分段常数型的,即某一频率段的幅频响应近似等于某一常数(这正是一般典型的低通、高通、带通、带阻型滤波器的响应特性),不然变换所产生的数字滤波器幅频响应相对于原模拟滤波器的幅频响应会有畸变,例如一个模拟微分器将不能转换成数字微分器,如图 6-13 所示。

对于分段常数的滤波器,双线性变换后,仍得到幅频特性为分段常数的滤波器,但是各个分段边缘的临界频率点产生了畸变,这种频率的畸变,可以通过频率的预畸来加以校

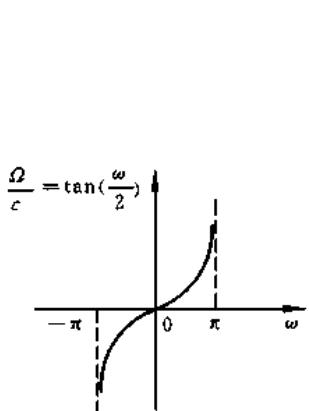


图 6-12 双线性变换的频率间非线性关系

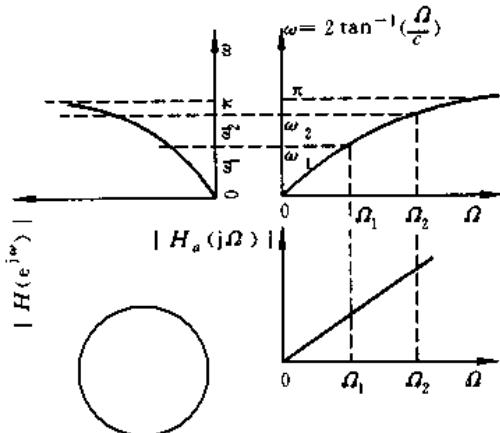


图 6-13 理想微分器经双线性变换后幅频响应产生畸变

正,也就是将临界频率事先加以畸变;然后经变换后正好映射到所需要的频率。例如要求数字带通滤波器的四个截止频率(即临界频率)为  $\omega_1, \omega_2, \omega_3, \omega_4$ ,如按线性变换所对应的模拟滤波器的四个截止频率应分别为

$$\Omega_1 = \frac{\omega_1}{T}, \quad \Omega_2 = \frac{\omega_2}{T}, \quad \Omega_3 = \frac{\omega_3}{T}, \quad \Omega_4 = \frac{\omega_4}{T}$$

但是模拟滤波器的这四个频率经双线性变换后(利用非线性的频率变换关系  $\omega = 2 \tan^{-1}(\frac{\Omega}{c})$ ),所得到的数字滤波器截止频率显然就不等于原来要求的  $\omega_1, \omega_2, \omega_3, \omega_4$  了。因而要将频率加以预畸,即利用

$$\Omega = c \tan \frac{\omega}{2}$$

的关系将这组数字频率  $\omega_i (i=1,2,3,4)$  变换成一组模拟频率  $\Omega_i (i=1,2,3,4)$ ,利用这组模拟频率来设计模拟带通滤波器,这就是我们所要求的模拟原型。对此模拟原形滤波器采用双线性变换,即可得到所需要的数字滤波器,它的截止频率正是我们原先所要求的一组  $\omega_i (i=1,2,3,4)$ 。这一预畸过程可见图 6-14。

## 五、模拟滤波器数字化方法

下面讨论利用双线性变换将模拟滤波器数字化的办法。

由于双线性变换法中,  $s$  到  $z$  之间的变换是简单的代数关系,见(6-57)式、(6-58)式,故可以由模拟系统函数通过(6-57)式的关系,用代数置换得到数字滤波器的系统函数,即

$$H(z) = H_a(s) \Bigg|_{s=\frac{1-z^{-1}}{1+z^{-1}}} = H_a \left[ c \frac{1-z^{-1}}{1+z^{-1}} \right] \quad (6-63)$$

也可以先将模拟系统函数分解成并联的子系统函数(子系统函数相加)或级联的子系统函数(子系统函数相乘),使每个子系统函数都变成低阶的(例如一、二阶的),然后再对

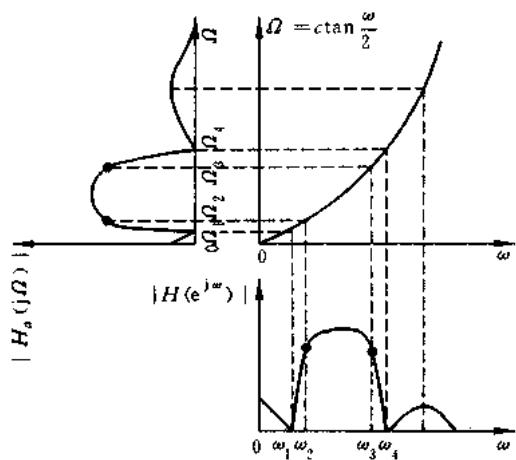


图 6-14 双线性变换的频率非线性预畸

每个子系统函数分别采用双线性变换。也就是说，分解为低阶的办法是在模拟系统函数上进行的，而模拟系统函数的分解已有大量的图表可以利用，分解起来比较方便。例如，设模拟系统函数分解为级联子系统

$$H_a(s) = H_{a_1}(s)H_{a_2}(s)\cdots H_{a_m}(s) \quad (6-64)$$

经双线性变换后，离散系统函数可表示为

$$H(z) = H_1(z)H_2(z)\cdots H_m(z) \quad (6-65)$$

其中

$$H_i(z) = H_{a_i}(s) \Bigg|_{s=\frac{1-z^{-1}}{1+z^{-1}}} , \quad i = 1, 2, \dots, m \quad (6-66)$$

如果模拟系统函数分解为并联子系统

$$H_a(s) = \bar{H}_{a_1}(s) + \bar{H}_{a_2}(s) + \cdots + \bar{H}_{a_n}(s) \quad (6-67)$$

经双线性变换后，离散系统函数为

$$H(z) = \bar{H}_1(z) + \bar{H}_2(z) + \cdots + \bar{H}_n(z) \quad (6-68)$$

其中

$$\bar{H}_i(z) = \bar{H}_{a_i}(s) \Bigg|_{s=\frac{1-z^{-1}}{1+z^{-1}}} , \quad i = 1, 2, \dots, n \quad (6-69)$$

应该注意，对于冲激响应不变法和阶跃响应不变法，则不能将模拟系统函数先分解成级联型子系统，这是因为  $s$  与  $z$  之间是超越函数的变换关系  $z = e^{j\omega}$ 。

## 六、用列表格的办法来完成双线性变换设计

(6-63)式的变换在概念上是很清楚的，但在实际应用时可能比较麻烦。因此，可以预先求出双线性变换法中离散系统函数的系数与模拟系统函数的系数之间的关系式，并列成表格，便可利用表格进行设计了。

设模拟系统函数的表达式为

$$H_a(s) = \frac{\sum_{i=0}^N d_i s^i}{\sum_{i=0}^N e_i s^i} = \frac{d_0 + d_1 s + d_2 s^2 + \dots + d_N s^N}{e_0 + e_1 s + e_2 s^2 + \dots + e_N s^N} \quad (6-70)$$

这里我们假定分子与分母多项式的阶次都为  $N$ 。若分子多项式的阶次小于分母的阶次，则可令分子多项式的最高几个阶次的系数  $d_i$  为零即可。

又设  $H_a(s)$  经双线性变换后得到  $H(z)$  的表达式为

$$H(z) = \frac{\sum_{i=0}^N A_i z^{-i}}{1 + \sum_{i=1}^N B_i z^{-i}} = \frac{A_0 + A_1 z^{-1} + A_2 z^{-2} + \dots + A_N z^{-N}}{1 + B_1 z^{-1} + B_2 z^{-2} + \dots + B_N z^{-N}} \quad (6-71)$$

表 6-2 给出了一阶到三阶的  $H(z)$  系数用  $H_a(s)$  系数以及变换常数  $c$  (由(6-59)式或(6-60)式确定的)表示的结果。对于更高阶次的滤波器，可利用低阶滤波器的级联或并联来实现。

表 6-2 双线性变换法中，用模拟滤波器  $H_a(s)$  的系数表示数字滤波器  $H(z)$  的系数

(c 是变换常数)	
一阶 $N=1$	
$A_0$	$(d_0 + d_1 c)/R$
$A_1$	$(d_0 - d_1 c)/R$
$B_1$	$(e_0 - e_1 c)/R$
$R$	$(e_0 + e_1 c)$
二阶 $N=2$	
$A_0$	$(d_0 + d_1 c + d_2 c^2)/R$
$A_1$	$(2d_0 - 2d_2 c^2)/R$
$A_2$	$(d_0 - d_1 c + d_2 c^2)/R$
$B_1$	$(2e_0 - 2e_2 c^2)/R$
$B_2$	$(e_0 - e_1 c + e_2 c^2)/R$
$R$	$(e_0 + e_1 c + e_2 c^2)$
三阶 $N=3$	
$A_0$	$(d_0 + d_1 c + d_2 c^2 + d_3 c^3)/R$
$A_1$	$(3d_0 + d_1 c - d_2 c^2 - 3d_3 c^3)/R$
$A_2$	$(3d_0 - d_1 c - d_2 c^2 + 3d_3 c^3)/R$
$A_3$	$(d_0 - d_1 c + d_2 c^2 - d_3 c^3)/R$
$B_1$	$(3e_0 + e_1 c - e_2 c^2 - 3e_3 c^3)/R$
$B_2$	$(3e_0 - e_1 c - e_2 c^2 + 3e_3 c^3)/R$
$B_3$	$(e_0 - e_1 c + e_2 c^2 - e_3 c^3)/R$
$R$	$(e_0 + e_1 c + e_2 c^2 + e_3 c^3)$

## \* 6.8 常用模拟低通滤波器特性

为了从模拟滤波器设计 IIR 数字滤波器, 必须先设计一个满足技术指标的模拟原型滤波器, 也就是要把数字滤波器的指标转变成模拟原型滤波器的指标, 因此我们就要设计这个模拟原型滤波器。

模拟滤波器的设计(逼近)不属于本课程的范围, 但是为了读者方便起见, 这里简单介绍几种最常用的模拟滤波器设计的基本方法。

模拟低通滤波器逼近有巴特沃思型、切贝雪夫型和考尔型(又称椭圆函数型)滤波器。而高通、带通、带阻等滤波器则可以利用变量变换方法, 由低通滤波器变换得到。

设计模拟滤波器是根据一组设计规范来设计模拟系统函数  $H_a(s)$ , 使其逼近某个理想滤波器特性。例如逼近图 6-15 所示理想的低通滤波器幅度特性, 这是根据幅度平方函数来逼近的。当然, 视应用情况不同, 也有根据相位特性或群延迟特性来进行逼近的。

### 一、由幅度平方函数确定系统函数

模拟滤波器幅度响应常用幅度平方函数  $|H_a(j\Omega)|^2$  来表示, 即

$$|H_a(j\Omega)|^2 = H_a(j\Omega)H_a^*(j\Omega)$$

由于滤波器冲激响应  $h(t)$  是实函数, 因而  $H(j\Omega)$  满足

$$H^*(j\Omega) = H(-j\Omega)$$

所以

$$|H_a(j\Omega)|^2 = H_a(j\Omega)H_a(-j\Omega) = H_a(s)H_a(-s) \Big|_{s=j\Omega} \quad (6-72)$$

其中  $H_a(s)$  是模拟滤波器的系统函数, 它是  $s$  的有理函数,  $H_a(j\Omega)$  是滤波器稳态响应即频率特性,  $|H_a(j\Omega)|$  是滤波器的稳态幅度特性。

现在问题是要求由已知的  $|H_a(j\Omega)|^2$  求得  $H_a(s)$ 。回到(6-72)式, 设  $H_a(s)$  有一个极点(或零点)位于  $s=s_0$  处, 由于  $h_a(t)$  即冲激响应为实数, 则极点(或零点)必以共轭对形式出现, 因而  $s=s_0^*$  处也一定有一极点(或零点), 所以与之对应  $H_a(-s)$  在  $s=-s_0$  和  $-s_0^*$  处必有极点(或零点), 因而  $H_a(s)$  的极点(零点)以及与之对应的  $H_a(-s)$  的极点(零点)可列表如下:

$H_a(s)$ 的极点(零点)	$-\sigma_1 \pm j\omega_1$	$-\sigma_0$	$-j\omega_0$
$H_a(-s)$ 的极点(零点)	$\sigma_1 \mp j\omega_1$	$\sigma_0$	$j\omega_0$

$H_a(s)H_a(-s)$  在虚轴上的零点(或极点), 我们知道, 稳定系统, 虚轴上是没有极点的, 对临

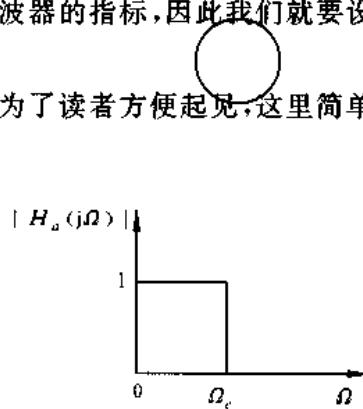


图 6-15 理想模拟低通滤波器  
幅度特性

界稳定情况,才会出现虚轴的极点)一定是二阶的,这是因为冲激响应  $h_a(t)$  是实的,因而  $H_a(s)H_a(-s)$  的极点(或零点)必成共轭对存在。 $H_a(s)H_a(-s)$  的极点、零点分布如图 6-16 所示,是成象限对称的。

我们知道,任何实际可实现的滤波器都是稳定的,因此其系统函数  $H_a(s)$  的极点一定落于  $s$  的左半平面,所以左半平面的极点一定属于  $H_a(s)$ ,则右半平面的极点必属于  $H_a(-s)$ 。

零点的分布则无此限制,只和滤波器的相位特性有关,如果要求是最小相位延时特性,则  $H_a(s)$  应取左半平面零点,如果有特殊要求,则按这种要求来考虑零点的分配,如无特殊要求,则可将对称零点的任一半(应为共轭对)取为  $H_a(s)$  的零点。

由此看出,由  $|H_a(j\Omega)|^2$  确定  $H_a(s)$  的方法如下:

(1) 由  $|H_a(j\Omega)|^2 \Big|_{\Omega^2 = -s^2} = H_a(s)H_a(-s)$  得到象限对称的  $s$  平面函数;

(2) 将  $H_a(s)H_a(-s)$  因式分解,得到各零点和极点。将左半平面的极点归于  $H_a(s)$ ,如无特殊要求,可取  $H_a(s)H_a(-s)$  以虚轴为对称轴的对称零点的任一半(应是共轭对)作为  $H_a(s)$  的零点,如要求是最小相位延时滤波器,则应取左半平面零点作为  $H_a(s)$  的零点。 $j\Omega$  轴上的零点或极点都是偶次的,其中一半(应为共轭对)属于  $H_a(s)$ 。

(3) 按照  $H_a(j\Omega)$  与  $H_a(s)$  的低频特性或高频特性的对比就可确定出增益常数。

(4) 由求出的  $H_a(s)$  的零点、极点及增益常数,则可完全确定系统函数  $H_a(s)$ 。

**[例 6-3]** 根据以下幅度平方函数  $|H_a(j\Omega)|^2$  确定系统函数  $H_a(s)$ :

$$|H_a(j\Omega)|^2 = \frac{16(25-\Omega^2)^2}{(49+\Omega^2)(36+\Omega^2)}$$

**解** 因为  $|H_a(j\Omega)|^2$  是  $\Omega$  的非负有理函数,它在  $j\Omega$  轴上的零点是偶次的,所以满足幅度平方函数的条件,先求

$$H_a(s)H_a(-s) = |H_a(j\Omega)|^2 \Big|_{\Omega^2 = -s^2} = \frac{16(25+s^2)^2}{(49-s^2)(36-s^2)}$$

其极点为  $s = \pm 7, s = \pm 6$ , 零点为  $s = \pm j5$ (为二阶)。

我们选出左半平面极点  $s = -6, s = -7$  及一对虚轴共轭零点  $s = \pm j5$  为  $H_a(s)$  的零、极点,并设增益常数为  $K_0$ ,则得  $H_a(s)$  为

$$H_a(s) = \frac{K_0(s^2 + 25)}{(s+7)(s+6)}$$

由  $H_a(s)|_{s=0} = H_a(j\Omega)|_{\Omega=0}$  的条件可得增益常数  $K_0$  为

$$K_0 = 4$$

最后得到  $H_a(s)$  为

$$H_a(s) = \frac{4(s^2 + 25)}{(s+7)(s+6)} = \frac{4s^2 + 100}{s^2 + 13s + 42}$$

下面介绍两种常用的低通滤波器特性。一般都是以低通滤波器为基础来讨论逼近函

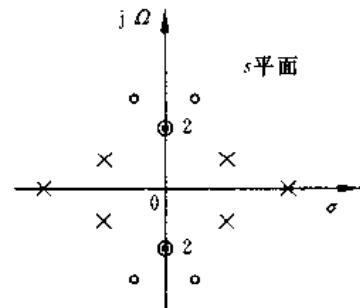


图 6-16  $H_a(s)H_a(-s)$  的零点、极点分布(成象限对称,虚轴零点上的“2”字表示是二阶零点)

数,而带通、带阻、高通滤波器则可用变换方法由低通滤波器映射而得到,这将在 IIR 滤波器的频率变换这部分中去介绍。

## 二、巴特沃思低通逼近

巴特沃思逼近又称最平幅度逼近。巴特沃思低通滤波器幅度平方函数定义为

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \left(\frac{\Omega}{\Omega_c}\right)^{2N}} \quad (6-73)$$

式中  $N$  为正整数,代表滤波器的阶次, $\Omega_c$  称为截止频率。当  $\Omega=\Omega_c$  时,有

$$|H_a(j\Omega_c)|^2 = \frac{1}{2}$$

即

$$|H_a(j\Omega_c)| = \frac{1}{\sqrt{2}}, \quad \delta_1 = 20\lg \left| \frac{H_a(j0)}{H_a(j\Omega_c)} \right| = 3\text{dB}$$

所以又称  $\Omega_c$  为巴特沃思低通滤波器的 3 分贝带宽。

巴特沃思低通滤波器的特点如下:

(1) 当  $\Omega=0$  时,  $|H_a(j0)|^2=1$ , 即在  $\Omega=0$  处无衰减。

(2) 当  $\Omega=\Omega_c$  时,  $|H_a(j\Omega_c)|^2=\frac{1}{2}$ ,  $|H_a(j\Omega_c)|=\frac{1}{\sqrt{2}}=0.707$ , 或  $\delta_1=-20\lg$

$|H_a(j\Omega_c)|=3\text{dB}$ ,  $\delta_1$  为通带最大衰减。即不管  $N$  为多少,所有的特性曲线都通过  $-3\text{dB}$  点,或说衰减  $3\text{dB}$ ,这就是  $3\text{dB}$  不变性。

(3) 在  $\Omega<\Omega_c$  的通带内  $|H_a(j\Omega)|^2$  有最大平坦的幅度特性,即  $N$  阶巴特沃思低通滤波器在  $\Omega=0$  处,  $|H_a(j\Omega)|_{\text{max}}^2$  的前  $(2N-1)$  阶导数为零,因而巴特沃思滤波器又称最平幅度特性滤波器。随着  $\Omega$  由 0 变到  $\Omega_c$ ,  $|H_a(j\Omega)|^2$  单调减小,  $N$  越大,减小得越慢,也就是通内特性越平坦。

(4) 当  $\Omega>\Omega_c$ ,即在过渡带及阻带中,  $|H_a(j\Omega)|^2$  也随  $\Omega$  增加而单调减小,但是  $\Omega/\Omega_c > 1$ ,故比通带内衰减的速度要快得多,  $N$  越大,衰减速度越大。当  $\Omega=\Omega_{st}$ ,即频率为阻带截止频率时,衰为  $\delta_2=-20\lg|H_a(j\Omega_{st})|$ ,  $\delta_2$  为阻带最小衰减。

巴特沃思低通滤波器的幅度特性如图 6-17 所示。

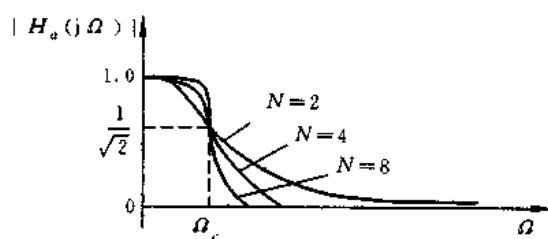


图 6-17 巴特沃思滤波器幅度特性及其与  $N$  的关系

在(6-73)式中,代入  $\Omega=s/j$ ,可得

$$|H_a(j\Omega)|_{\Omega=\frac{1}{j}}^2 = H_a(s)H_a(-s) = \frac{1}{1 + \left(\frac{s}{\Omega_c}\right)^{2N}} \quad (6-74)$$

所以巴特沃思滤波器的零点全部在  $s=\infty$  处，在有限  $s$  平面上只有极点，因而属于所谓“全极点型”滤波器。 $H_a(s)H_a(-s)$  (即幅度平方函数在  $s$  平面的解析延拓) 的极点为

$$s_k = (-1)^{\frac{k}{2N}}(j\Omega_c) = \Omega_c e^{j[\frac{1}{2} + \frac{2k-1}{2N}]\pi}, \quad k = 1, 2, \dots, 2N \quad (6-75)$$

由此看出， $H_a(s)H_a(-s)$  的极点分布的特点是：

(1) 极点在  $s$  平面上是象限对称的，分布在半径为  $\Omega_c$  的圆 (称巴特沃思圆) 上，共有  $2N$  个极点。

(2) 极点间的角度间隔为  $\frac{\pi}{N}$  rad。

(3) 极点绝不会落在虚轴上，因而滤波器才有可能是稳定的。

(4)  $N$  为奇数时，实轴上有极点； $N$  为偶数时，实轴上没有极点。

例如， $N=3$  及  $N=4$  时， $H_a(s)H_a(-s)$  的极点分布分别如图 6-18(a) 和 (b) 所示。

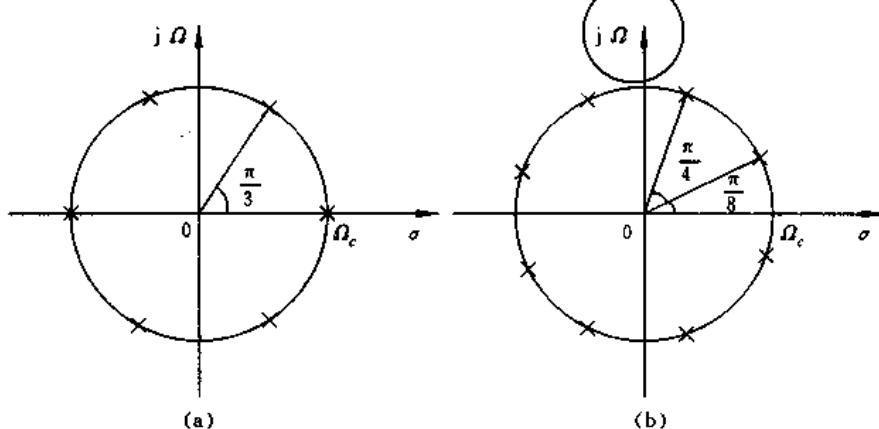


图 6-18 巴特沃思滤波器  $H_a(s)H_a(-s)$  在  $s$  平面上的极点位置

(a)  $N=3$ (三阶); (b)  $N=4$ (四阶)

$H_a(s)H_a(-s)$  在左半平面的极点即为  $H_a(s)$  的极点，因而

$$H_a(s) = \frac{\Omega_c^N}{\prod_{k=1}^N (s - s_k)} \quad (6-76)$$

这里分子系数为  $\Omega_c^N$ ，可由  $H_a(s)$  的低频特性决定，(代入  $H_a(0)=1$ ，可求得分子系数为  $\Omega_c^N$ )，而  $s_k$  为

$$s_k = \Omega_c e^{j[\frac{1}{2} + \frac{2k-1}{2N}]\pi}, \quad k = 1, 2, \dots, N \quad (6-77)$$

在一般设计中，都先把(6-74)式中的  $\Omega_c$  选为 1rad/s，这样使频率得到归一化，归一化后巴特沃思滤波器的极点分布以及相应的系统函数、分母多项式的系数都有现成的表格可查。

如果  $\Omega_c$  表示归一化频率响应中的参考角频率(一般取为 1rad/s，也可以是其他任意

数值),而所需的实际滤波器幅度响应中的参考角频率为  $\Omega_c$ (一般为截止频率或称3dB截止频率,也可以是其他衰减分贝处的频率)。令  $H_{an}(s)$  代表归一化系统的系统函数,  $H_a(s)$  代表所需的参考角频率为  $\Omega_c$  的系统的系统函数,那么把原归一化系统函数中的变量  $s$  用  $\frac{\Omega_c s}{\Omega_c}$  代替后,就得到所需系统的系统函数,即

$$s \rightarrow \frac{\Omega_c s}{\Omega_c} \quad (6-78)$$

$$H_a(s) = H_{an}\left(\frac{\Omega_c s}{\Omega_c}\right) \quad (6-79)$$

[例 6-4] 导出三阶巴特沃思低通滤波器的系统函数,设  $\Omega_c = 2 \text{ rad/s}$ 。

解 幅度平方函数是

$$|H(j\Omega)|^2 = \frac{1}{1 + \left(\frac{\Omega}{2}\right)^6}$$

令  $\Omega^2 = -s^2$ , 则有

$$H_a(s)H_a(-s) = \frac{1}{1 - \frac{s^6}{2^6}}$$

各极点满足(6-75)式

$$s_k = 2e^{j\left[\frac{\pi}{6} + \frac{2k-1}{6}\pi\right]}, \quad k = 1, 2, \dots, 6$$

而按(6-77)式,前面3个  $s_k$  ( $k=1, 2, 3$ )就是  $H_a(s)$  的极点。

所得出的六个  $s_k$  为

$$\begin{aligned} s_1 &= 2e^{j\frac{\pi}{3}} = -1 + j\sqrt{3} \\ s_2 &= 2e^{j\pi} = -2 \\ s_3 &= 2e^{j\frac{4}{3}\pi} = -1 - j\sqrt{3} \\ s_4 &= 2e^{j\frac{5}{3}\pi} = 1 - j\sqrt{3} \\ s_5 &= 2e^{j0} = 2 \\ s_6 &= 2e^{j\frac{1}{3}\pi} = 1 + j\sqrt{3} \end{aligned}$$

由  $s_1, s_2, s_3$  三个极点构成的系统函数为

$$H_a(s) = \frac{\Omega_c^3}{(s-s_1)(s-s_2)(s-s_3)} = \frac{8}{s^3 + 4s^2 + 8s + 8}$$

### 三、切贝雪夫低通逼近

巴特沃思滤波器的频率特性无论在通带与阻带都随频率而单调变化,因而如果在通带边缘满足指标,则在通带内肯定会有富裕量,也就是会超过指标的要求,因而并不经济。所以,更有效的办法是将指标的精度要求均匀地分布在通带内,或均匀分布在阻带内,或同时均匀分布在通带与阻带内,这样在同样通带、阻带性能要求下,就可设计出阶数较低的滤波器。这种精度均匀分布的办法可通过选择具有等波纹特性的逼近函数来完成。

切贝雪夫滤波器的幅度特性就在一个频带中(通带或阻带)具有这种等波纹特性。一

种是在通带中是等波纹的,在阻带中是单调的,称为切贝雪夫 I 型;一种是在通带内是单调的,在阻带内是等波纹的,称为切贝雪夫 II 型。由应用的要求来确定采用哪种型式的切贝雪夫滤波器。图 6-19、图 6-20 分别画出了  $N$  为奇数与  $N$  为偶数时的切贝雪夫 I, II 型滤波器的幅度特性。

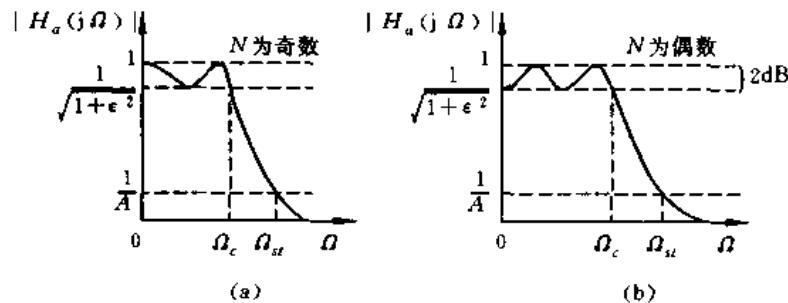


图 6-19 切贝雪夫 I 型滤波器的幅度特性(通带波纹 2dB)

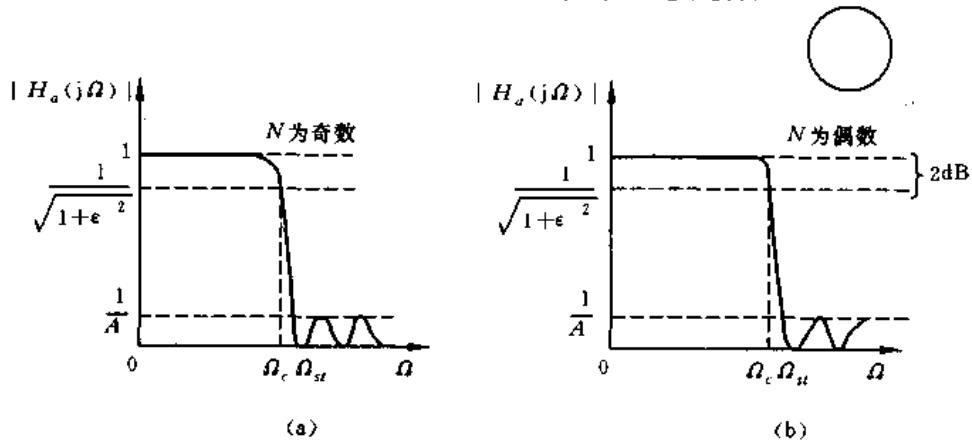


图 6-20 切贝雪夫 II 型滤波器的幅度特性

我们以切贝雪夫 I 型滤波器为例来讨论这种逼近。

切贝雪夫 I 型滤波器的幅度平方函数为

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \epsilon^2 C_N^2 \left(\frac{\Omega}{\Omega_c}\right)} \quad (6-80)$$

其中,  $\epsilon$  为小于 1 的正数, 它是表示通带波纹大小的一个参数,  $\epsilon$  越大, 波纹也越大。 $\frac{\Omega}{\Omega_c}$  为  $\Omega$  对  $\Omega_c$  的归一化频率,  $\Omega_c$  为截止频率, 也是滤波器的某一衰减分贝处的通带宽度(这一分贝数不一定是 3 分贝, 也就是说, 在切贝雪夫滤波器中,  $\Omega_c$  不一定是 3dB 的带宽)。 $C_N(x)$  是  $N$  阶切贝雪夫多项式, 定义为

$$C_N(x) = \begin{cases} \cos(N \cos^{-1} x), & |x| \leq 1 \\ \operatorname{ch}(N \operatorname{ch}^{-1} x), & |x| > 1 \end{cases} \quad (6-81)$$

(6-81)式可展开成多项式, 见表 6-3。从表 6-3 看出,  $N$  阶切贝雪夫多项式首项( $x^N$ )的系数为  $2^{N-1}$ 。由此可归纳出高阶( $N \geq 1$ )切贝雪夫多项式的递推公式为

$$C_{N+1}(x) = 2x C_N(x) - C_{N-1}(x) \quad (6-82)$$

表 6-3 切贝雪夫多项式

N	$C_N(x)$
0	1
1	$x$
2	$2x^2 - 1$
3	$4x^3 - 3x$
4	$8x^4 - 8x^2 + 1$
5	$16x^5 - 20x^3 + 5x$
6	$32x^6 - 48x^4 + 18x^2 - 1$

图 6-21 画出了  $N=0,1,2,3,4,5$  时的切贝雪夫多项式特性。由图可知, 切贝雪夫多项式的零值点在  $|x| \leq 1$  间隔内。当  $|x| \leq 1$  时,  $C_N(x)$  是余弦函数, 故

$$|C_N(x)| \leq 1$$

且多项式  $C_N(x)$  在  $|x| \leq 1$  内具有等波纹幅度特性; 当  $|x| > 1$  时,  $C_N(x)$  是双曲余弦函数, 它随  $x$  而单调地增加。

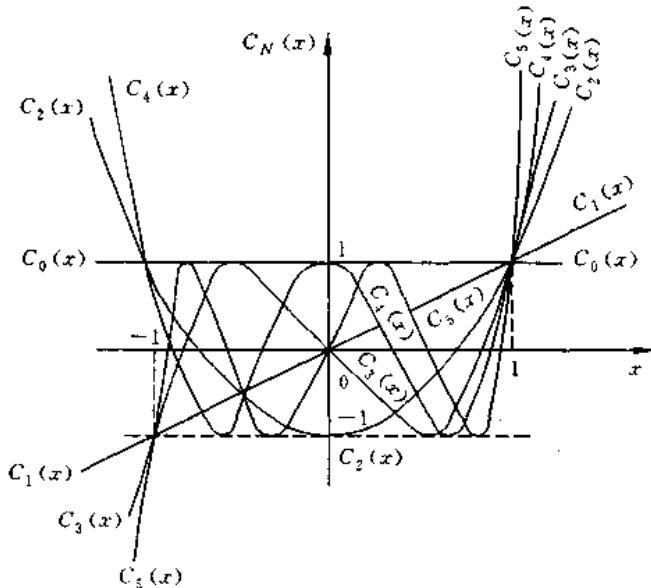


图 6-21  $N=0,1,2,3,4,5$  等各阶切贝雪夫多项式  $C_N(x)$  曲线

显然, 切贝雪夫滤波器的幅度函数为  $|H_a(j\Omega)| = \frac{1}{\sqrt{1+\epsilon^2 C_N^2(\frac{\Omega}{\Omega_c})}}$ 。 $|H_a(j\Omega)|$  的特点

如下:

(1) 当  $\Omega=0$ ,  $N$  为偶数时,  $H_a(j0) = \frac{1}{\sqrt{1+\epsilon^2}}$ ; 当  $N$  为奇数时,  $H_a(j0) = 1$ 。

(2) 当  $\Omega=\Omega_c$  时

$$|H_a(j\Omega)| = \frac{1}{\sqrt{1+\epsilon^2}}$$

即所有幅度函数曲线都通过  $\frac{1}{\sqrt{1+\epsilon^2}}$  点, 所以把  $\Omega_c$  定义为切贝雪夫滤波器的截止频率。在这个截止频率下, 幅度函数不一定下降 3dB, 可以是下降其他分贝值, 例如 1dB 等, 这是与巴特沃思滤波器不同之处。

(3) 在通带内, 即当  $\Omega < \Omega_c$  时, 则  $\frac{\Omega}{\Omega_c} < 1$ ,  $|H_a(j\Omega)|$  在  $1 \sim \frac{1}{\sqrt{1+\epsilon^2}}$  之间等波纹地起伏。

(4) 在通带之外, 即当  $\Omega > \Omega_c$  时, 随着  $\Omega$  的增大

$$\epsilon^2 C_N^2 \left( \frac{\Omega}{\Omega_c} \right) \gg 1$$

使  $|H_a(j\Omega)|$  迅速单调地趋近于零。

由幅度平方函数的(6-80)式看出, 切贝雪夫滤波器有三个参数:  $\epsilon$ ,  $\Omega_c$  和  $N$ 。 $\Omega_c$  是通带宽度, 一般是预先给定的。 $\epsilon$  是与通带波纹有关的一个参数, 通带波纹  $\delta_1$  表示成

$$\delta_1 = 10 \log_{10} \frac{|H_a(j\Omega)|_{\max}^2}{|H_a(j\Omega)|_{\min}^2} = 20 \log_{10} \frac{|H_a(j\Omega)|_{\max}}{|H_a(j\Omega)|_{\min}} (\text{dB}) \quad (6-83)$$

这里  $|H_a(j\Omega)|_{\max} = 1$ , 表示通带幅度响应的最大值。 $|H_a(j\Omega)|_{\min} = \frac{1}{\sqrt{1+\epsilon^2}}$ , 表示通带幅度响应的最小值, 故

$$\delta_1 = 10 \log_{10} (1 + \epsilon^2) \quad (6-84)$$

因而

$$\epsilon^2 = 10^{\delta_1/10} - 1 \quad (6-85)$$

可以看出, 给定通带波纹值  $\delta_1$  (dB) 后, 就能求得  $\epsilon^2$ , 这里应注意, 通带波纹值不一定是 3dB, 也可以是其他值, 例如 0.1dB 等等。

滤波器阶数  $N$  等于通带内最大值和最小值的总数。前面已经说过,  $N$  为奇数时,  $\Omega=0$  处  $|H_a(j\Omega)|$  为最大值 1,  $N$  为偶数时,  $\Omega=0$  处  $|H_a(j\Omega)|$  为最小值  $\frac{1}{\sqrt{1+\epsilon^2}}$  (见图 6-19)。 $N$  的数值可由阻带衰减来确定。设阻带起始点频率为  $\Omega_{st}$ , 此时阻带幅度平方函数值满足

$$|H_a(j\Omega_{st})|^2 \leq \frac{1}{A^2}$$

$A$  是常数(见图 6-19), 如果用误差的分贝数  $\delta_2$  表示, 则有

$$\delta_2 = 20 \log \frac{1}{1/A} = 20 \log A$$

所以

$$A = 10^{\delta_2/20} = 10^{0.05\delta_2} \quad (6-86)$$

设  $\Omega_{st}$  为阻带截止频率, 即当  $\Omega=\Omega_{st}$  时, 将上面的  $|H_a(j\Omega_{st})|^2$  的表达式代入(6-80)式, 可得

$$|H_a(j\Omega_{st})|^2 = \frac{1}{1 + \epsilon^2 C_N^2 \left( \frac{\Omega_{st}}{\Omega_c} \right)} \leq \frac{1}{A^2}$$

由此得出

$$C_N \left( \frac{\Omega_u}{\Omega_c} \right) \geq \frac{1}{\epsilon} \sqrt{A^2 - 1} \quad (6-87)$$

由于  $\frac{\Omega_u}{\Omega_c} > 1$ , 所以, 由(6-81)式的第二式有

$$C_N \left( \frac{\Omega_u}{\Omega_c} \right) = \operatorname{ch} \left[ N \operatorname{ch}^{-1} \left( \frac{\Omega_u}{\Omega_c} \right) \right]$$

再将(6-87)式代入, 可得

$$C_N \left( \frac{\Omega_u}{\Omega_c} \right) = \operatorname{ch} \left[ N \operatorname{ch}^{-1} \left( \frac{\Omega_u}{\Omega_c} \right) \right] \geq \frac{1}{\epsilon} \sqrt{A^2 - 1}$$

由此, 并考虑(6-86)式, 可得

$$N \geq \frac{\operatorname{ch}^{-1} \left[ \frac{1}{\epsilon} \sqrt{A^2 - 1} \right]}{\operatorname{ch}^{-1} \left( \frac{\Omega_u}{\Omega_c} \right)} = \frac{\operatorname{ch}^{-1} \left[ \frac{1}{\epsilon} \sqrt{10^{0.1B_2} - 1} \right]}{\operatorname{ch}^{-1} \left( \frac{\Omega_u}{\Omega_c} \right)} \quad (6-88)$$

如果要求阻带边界频率上衰减越大(即  $A$  越大), 也就是过渡带内幅度特性越陡, 则所需的阶数  $N$  越高。

或者对  $\Omega_u$  求解, 可得

$$\Omega_u = \Omega_c \operatorname{ch} \left\{ \frac{1}{N} \operatorname{ch}^{-1} \left[ \frac{1}{\epsilon} \sqrt{A^2 - 1} \right] \right\} = \Omega_c \operatorname{ch} \left\{ \frac{1}{N} \operatorname{ch}^{-1} \left[ \frac{1}{\epsilon} \sqrt{10^{0.1B_2} - 1} \right] \right\} \quad (6-89)$$

这里  $\Omega_c$  是切贝雪夫滤波器的通带宽度, 但不是 3dB 带宽, 可以求出 3dB 带宽为 ( $A=\sqrt{2}$ )

$$\Omega_{3dB} = \Omega_c \operatorname{ch} \left[ \frac{1}{N} \operatorname{ch}^{-1} \left( \frac{1}{\epsilon} \right) \right] \quad (6-90)$$

注意, 只有当  $\Omega_c < \Omega_{3dB}$  时, 才能采用(6-90)式来求解  $\Omega_{3dB}$  (因为满足  $\Omega_{3dB}/\Omega_c > 1$ )。

$N, \Omega_c, \epsilon$  给定后, 就可求得滤波器  $H_a(s)$ 。

当然, 已知  $N, \epsilon$  后, 也可先求出  $H_a(s)$  的极点分布, 即可求得  $H_a(s)$ 。我们由幅度平方函数来求解, 将  $\Omega = s/j$  代入  $|H_a(j\Omega)|^2$ , 再考虑(6-80)式, 可得

$$\left| H_a(j\Omega) \right|^2 \Big|_{\Omega=j\frac{s}{j\Omega_c}} = H_a(s) H_a(-s) = \frac{1}{1 + \epsilon^2 C_N^2 \left( \frac{s}{j\Omega_c} \right)}$$

令其分母多项式为零, 即

$$1 + \epsilon^2 C_N^2 \left( \frac{s}{j\Omega_c} \right) = 0$$

即

$$C_N \left( \frac{s}{j\Omega_c} \right) = \pm j \frac{1}{\epsilon} \quad (6-91)$$

令

$$\frac{s}{j\Omega_c} = \cos(\alpha + j\beta) = \cos\alpha \cdot \cos(j\beta) - \sin\alpha \cdot \sin(j\beta)$$

$$= \cos\alpha \cdot \operatorname{ch}\beta - j \sin\alpha \cdot \operatorname{sh}\beta \quad (6-92)$$

$\operatorname{ch}(\cdot)$  和  $\operatorname{sh}(\cdot)$  分别为双曲余弦函数与双曲正弦函数。由(6-92)式可得

$$s = \Omega_c \sin\alpha \cdot \operatorname{sh}\beta + j \Omega_c \cos\alpha \cdot \operatorname{ch}\beta \quad (6-93)$$

为了导出  $\alpha, \beta$  与  $N, \epsilon$  的关系, 把(6-92)式代入(6-91)式, 且考虑到  $C_N(x)$  的定义(见(6-81)式)

$$\begin{aligned} C_N\left(\frac{s}{j\Omega_c}\right) &\triangleq \cos\left[N \cos^{-1}\left(\frac{s}{j\Omega_c}\right)\right] = \cos[N(\alpha + j\beta)] \\ &= \cos(N\alpha) \cdot \operatorname{ch}(N\beta) - j \sin(N\alpha) \cdot \operatorname{sh}(N\beta) \\ &= \pm j \frac{1}{\epsilon} \end{aligned}$$

由此得出

$$\cos(N\alpha) \cdot \operatorname{ch}(N\beta) = 0 \quad (6-94a)$$

$$\sin(N\alpha) \cdot \operatorname{sh}(N\beta) = \mp \frac{1}{\epsilon} \quad (6-94b)$$

由(6-94a)式, 且考虑  $\operatorname{ch}(N\beta) \neq 0$ , 可解出

$$\cos(N\alpha) = 0$$

则有

$$\alpha = \frac{2k-1}{N} \cdot \frac{\pi}{2}, \quad k = 1, 2, \dots, 2N \quad (6-95)$$

把  $\alpha$  代入(6-94b)式, 可得

$$\operatorname{sh}(N\beta) = \mp \frac{1}{\epsilon}$$

因而

$$\beta = \frac{1}{N} \operatorname{sh}^{-1}\left(\mp \frac{1}{\epsilon}\right) = \mp \frac{1}{N} \operatorname{sh}^{-1} \frac{1}{\epsilon} \quad (6-96)$$

把(6-95)式的  $\alpha$  及(6-96)式的  $\beta$  代入(6-93)式, 可得  $H_a(s)H_a(-s)$  的极点为

$$\begin{aligned} s_k &= \sigma_k + j\Omega_k = \mp \Omega_c \sin\left(\frac{2k-1}{2N}\pi\right) \cdot \operatorname{sh}\left(\frac{1}{N} \operatorname{sh}^{-1} \frac{1}{\epsilon}\right) \\ &\quad + j\Omega_c \cos\left(\frac{2k-1}{2N}\pi\right) \cdot \operatorname{ch}\left(\frac{1}{N} \operatorname{sh}^{-1} \frac{1}{\epsilon}\right), \quad k = 1, 2, \dots, 2N \end{aligned} \quad (6-97)$$

即

$$\sigma_k = \mp \Omega_c \sin\left[\frac{\pi}{2N}(2k-1)\right], \quad k = 1, 2, \dots, 2N \quad (6-98a)$$

$$\Omega_k = \Omega_c \cos\left[\frac{\pi}{2N}(2k-1)\right], \quad k = 1, 2, \dots, 2N \quad (6-98b)$$

其中

$$a = \operatorname{sh}\left[\frac{1}{N} \operatorname{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right] \quad (6-99a)$$

$$b = \operatorname{ch}\left[\frac{1}{N} \operatorname{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right] \quad (6-99b)$$

将(6-98a)式、(6-98b)式分别取平方再化简, 可得  $H_a(s)H_a(-s)$  在  $s$  平面上的极点分布满足的关系式

$$\frac{\sigma_k^2}{\Omega_c^2 a^2} + \frac{\Omega_k^2}{\Omega_c^2 b^2} = 1 \quad (6-100)$$

这是一个椭圆方程,由于双曲余弦  $\text{ch}(\cdot)$  大于双曲正弦  $\text{sh}(\cdot)$ ,故长轴为  $\Omega_c b$ (在虚轴上),短轴为  $\Omega_c a$ (在实轴上)。

取  $H_a(s) H_a(-s)$  在  $s$  左半平面的极点,即(6-97)式的左半平面的  $s_k$ ,就是  $H_a(s)$  的极点。

$$\sigma_k = -\Omega_c a \sin\left[\frac{\pi}{2N}(2k-1)\right], \quad k = 1, 2, \dots, N \quad (6-101a)$$

$$\Omega_k = \Omega_c b \cos\left[\frac{\pi}{2N}(2k-1)\right], \quad k = 1, 2, \dots, N \quad (6-101b)$$

实际上,用(6-99)式求解  $a, b$  并不方便,还可以再加简化,考虑到

$$a = \text{sh}\left[\frac{1}{N} \text{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right] = \frac{\exp\left[\frac{1}{N} \text{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right] - \exp\left[-\frac{1}{N} \text{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right]}{2} \\ = \frac{1}{2}\left(\gamma^{\frac{1}{N}} - \gamma^{-\frac{1}{N}}\right) \quad (6-102a)$$

$$b = \text{ch}\left[\frac{1}{N} \text{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right] = \frac{\exp\left[\frac{1}{N} \text{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right] + \exp\left[-\frac{1}{N} \text{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right]}{2} \\ = \frac{1}{2}\left(\gamma^{\frac{1}{N}} + \gamma^{-\frac{1}{N}}\right) \quad (6-102b)$$

其中

$$\gamma = \exp\left[\text{sh}^{-1}\left(\frac{1}{\epsilon}\right)\right]$$

查数学手册知,若

$$y = \text{sh}^{-1}\left(\frac{1}{\epsilon}\right)$$

则有

$$y = \ln\left(\frac{1}{\epsilon} + \sqrt{\frac{1}{\epsilon^2} + 1}\right)$$

将此式取反变换有

$$\exp(y) = \frac{1}{\epsilon} + \sqrt{\frac{1}{\epsilon^2} + 1}$$

$$\gamma = \frac{1}{\epsilon} + \sqrt{\frac{1}{\epsilon^2} + 1} \quad (6-103)$$

由此得出

$$a = \frac{1}{2}(\gamma^{\frac{1}{N}} - \gamma^{-\frac{1}{N}}) \quad (6-104a)$$

$$b = \frac{1}{2}(\gamma^{\frac{1}{N}} + \gamma^{-\frac{1}{N}}) \quad (6-104b)$$

这样,由(6-103)式可求得  $\gamma$ ,由  $\gamma$  可求得  $a, b$ ,从而利用(6-101)式求得  $H_a(s)$  的极点。

求出幅度平方函数的极点后, $H_a(s)$  的极点就是  $s$  平面左半平面的诸极点  $s_i$ ,从而得到切贝雪夫滤波器的系统函数为

$$H_a(s) = \frac{K}{\prod_{i=1}^N (s - s_i)} \quad (6-105a)$$

常数  $K$  可由  $A(\Omega)$  和  $H_a(s)$  的低频或高频特性对比如求得。

也可用下面的方法直接求出(6-105a)式的系数  $K$ : 将(6-80)式开平方, 并代入  $\Omega = s/j$ , 再考虑(6-105a)式, 则有

$$|H_a(s)| = \frac{1}{\sqrt{1 + \epsilon^2 C_N^2 \left(\frac{s}{j\Omega_c}\right)}} = \frac{K}{\left|\prod_{i=1}^N (s - s_i)\right|}$$

可以看出, 第一个等号后的公式, 其分母多项式的首项( $s^N$ )的系数不为 1, 这是因为  $C_N \left(\frac{s}{j\Omega_c}\right)$  的首项  $\left[\left(\frac{s}{j\Omega_c}\right)^N\right]$  系数为  $2^{N-1}$ , 因而其  $s^N$  项的系数为  $\frac{2^{N-1}}{\Omega_c^N}$ , 整个分母多项式  $s^N$  项的系数则为  $\frac{\epsilon \cdot 2^{N-1}}{\Omega_c^N}$ 。而第二个等号后的分式, 其分母的首项( $s^N$ )的系数为 1, 因而, 为使第二个等号两端的函数(皆化为首项( $s^N$ )的系数为 1)相等, 必须满足

$$K = \frac{\Omega_c^N}{\epsilon \cdot 2^{N-1}}$$

将此  $K$  值代入(6-105a)式, 可得

$$H_a(s) = \frac{\frac{1}{\epsilon} \cdot \frac{1}{2^{N-1}} \cdot \Omega_c^N}{\prod_{i=1}^N (s - s_i)} \quad (6-105b)$$

图 6-22 也画出了确定切贝雪夫 I 型滤波器极点在椭圆上的位置的方法。先求出大圆(半径为  $b\Omega_c$ )和小圆(半径为  $a\Omega_c$ )上按等间隔角  $\pi/N$  均分的各个点, 这些点是虚轴对

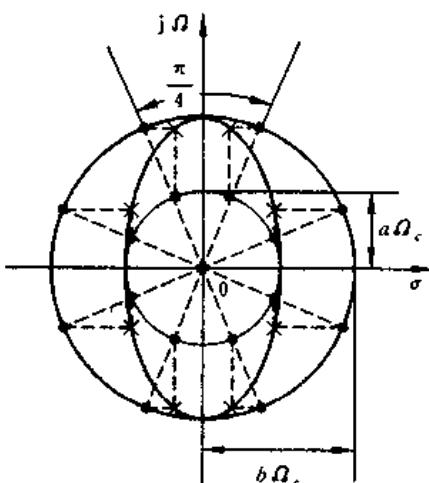


图 6-22  $N=4$  时模拟切贝雪夫 I 型滤波器极点位置图

称的,且一定都不落在虚轴上。 $N$ 为奇数时,有落在实轴上的点, $N$ 为偶数时,实轴上也没有。幅度平方函数的极点(在椭圆上)的位置是这样确定的:其垂直坐标由落在大圆上的各等间隔点规定,其水平坐标由落在小圆上的各等间隔点规定。

切贝雪夫Ⅱ型滤波器的幅度特性在通带内具有单调的特性(在 $\Omega=0$ 附近最平坦),而在阻带内具有等波纹特性,其幅度平方函数可表示成

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \epsilon^2 \left[ \frac{C_N(\Omega_u)}{C_N(\Omega_u/\Omega)} \right]^2} \quad (6-106)$$

式中 $\Omega_u$ 是阻带衰减达到规定数值的最低频率。

可以证明,切贝雪夫Ⅱ型滤波器既有极点又有零点,零点是虚数。由于篇幅有限,这里就不作讨论了。

还有一种通带阻带都具有等波纹幅度特性的滤波器,称为椭圆函数滤波器或考尔滤波器,因为比较专门,这里不作讨论。

**[例 6-5]** 试导出二阶切贝雪夫系统函数,已知通带波纹为 1dB,归一化截止频率为 $\Omega_c=1$  rad/s。

解 由于 $\delta_t=1$  dB。根据(6-85)式,可得

$$\epsilon^2 = 10^{t/10} - 1 = 10^{0.1} - 1 = 0.25892541$$

由于 $\Omega_c=1$ ,故 $x=\frac{\Omega}{\Omega_c}=\Omega$ ,由表 6-2,代入 $x=\Omega$  可得

$$C_2(\Omega) = 2\Omega^2 - 1$$

则

$$C_2^2(\Omega) = 4\Omega^4 - 4\Omega^2 + 1$$

代入 $C_2^2(\Omega)$ 及 $\epsilon^2$  到(6-80)式,可得

$$|H_a(j\Omega)|^2 = \frac{1}{1.0357016\Omega^4 + 1.0357016\Omega^2 + 1.25892541}$$

令 $s=j\Omega$ ,即 $s^2=-\Omega^2$ ,可得

$$H_a(s)H_a(-s) = |H_a(j\Omega)|^2 |_{\Omega^2=-s^2} = \frac{1}{1.0357016s^4 - 1.0357016s^2 + 1.25892541}$$

从分母多项式的根得出 $H_a(s)H_a(-s)$ 的极点为

$$s_1 = 1.0500049e^{j58.484569^\circ}$$

$$s_2 = 1.0500049e^{j121.51543^\circ}$$

$$s_3 = 1.0500049e^{-j121.51543^\circ}$$

$$s_4 = 1.0500049e^{-j58.484569^\circ}$$

幅度平方函数的这些极点在 $s$ 平面一个椭圆上。系统函数 $H_a(s)$ 由 $H_a(s)H_a(-s)$ 的左半平面极点( $s_2, s_3$ )确定,考虑到直流增益为 $1/\sqrt{1+\epsilon^2}$ (因为 $N$ 是偶数),最后可得

$$H_a(s) = \frac{0.9826133}{s^2 + 1.0977343s + 1.1025103}$$

利用(6-105b)式,  $H_a(s)$  的分子也可得到同样的系数。

下面再给一个例子, 说明如何利用给定的模拟滤波器性能要求, 来求模拟滤波器的阶次及系统函数。

[例 6-6] 给定模拟低通滤波器的性能指标为, 在通带内, 即在  $0 \leq \Omega \leq 2\pi \times 10^4 \text{ rad/sec}$  内, 幅度函数的波纹(起伏)  $\delta_1 \leq 1 \text{ dB}$ ; 在阻带内, 即在  $\Omega \geq 2\pi \times 1.5 \times 10^4 \text{ rad/sec}$  时, 幅度函数衰减  $\delta_2 \geq 15 \text{ dB}$ 。试求用切贝雪夫滤波器实现时, 所需阶次  $N$  以及滤波器系统函数  $H_a(s)$  的表达式。

解 由题意知, 通带波纹  $\delta_1 = 1 \text{ dB}$ , 阻带衰减  $\delta_2 = 15 \text{ dB}$ 。

(1) 求通带内波纹参数  $\epsilon$ 。由(6-85)式, 有

$$\epsilon = \sqrt{10^{0.1\delta_1} - 1} = \sqrt{10^{0.1} - 1} = 0.508847$$

(2) 求阶次  $N$ 。由(6-88)式, 可得

$$N \geq \frac{\operatorname{ch}^{-1} \left[ \frac{1}{\epsilon} \sqrt{10^{0.1\delta_2} - 1} \right]}{\operatorname{ch}^{-1} \left( \frac{\Omega_u}{\Omega_c} \right)} = 3.1977$$

取  $N=4$ 。

(3) 求  $H_a(s)$  的极点。由(6-103)式, 可得

$$\gamma = \frac{1}{\epsilon} + \sqrt{\frac{1}{\epsilon^2} + 1} = 4.1702477$$

由(6-104)式, 可求得

$$a = 1.064402, \quad b = 0.3646251$$

因而极点分布的椭圆的长轴( $\Omega_c b$ )及短轴( $\Omega_c a$ )分别为

$$(\Omega_c b) = 2\pi \times 10^4 \times 1.064402 = 6.6878347 \times 10^4$$

$$(\Omega_c a) = 2\pi \times 10^4 \times 0.3646251 = 2.2910071 \times 10^4$$

利用(6-101)式, 可求出极点分布为(考虑到极点成共轭对的关系, 如图 6-22 所示)

$$\begin{aligned} s_1 &= s_4^* = \sigma_1 + j\Omega_1 = -a\Omega_c \sin\left(\frac{\pi}{2N}\right) + jb\Omega_c \cos\left(\frac{\pi}{2N}\right) \\ &= -0.876303 \times 10^4 + j6.1787534 \times 10^4 = 6.241 \times 10^4 e^{j98.08^\circ} \\ s_2 &= s_3^* = -a\Omega_c \sin\left(\frac{3\pi}{2N}\right) + jb\Omega_c \cos\left(\frac{3\pi}{2N}\right) \\ &= -2.1166145 \times 10^4 + j2.5593233 \times 10^4 = 3.321 \times 10^4 e^{j129.59^\circ} \end{aligned}$$

(4) 求滤波器的系统函数  $H_a(s)$  为

$$\begin{aligned} H_a(s) &= \frac{A_0}{\prod_{k=1}^4 (s - s_k)} = \frac{A_0}{(s - s_1)(s - s_1^*)(s - s_2)(s - s_2^*)} \\ &= \frac{A_0}{(s^2 + 1.7535 \times 10^4 s + 3.895 \times 10^9)(s^2 + 4.233 \times 10^4 s + 11.029 \times 10^9)} \end{aligned}$$

现在要求系数  $A_0$ , 由于  $N$  为偶数时, 应满足(见图 6-19(b))

$$|H(j0)| = \frac{1}{\sqrt{1+\epsilon^2}} = 0.8912509$$

因而有

$$H_a(s) \Big|_{s=0} = \frac{A_0}{3.895 \times 10^9 \times 11.029 \times 10^9} = 0.8912509$$

由此求出

$$A_0 = 3.8286 \times 10^{18}$$

或用(6-105b)式直接求出  $A_0$ , 结果是一样的。由此得出

$$\begin{aligned} H_a(s) &= \frac{3.8286 \times 10^{18}}{(s^2 + 1.7535 \times 10^4 s + 3.895 \times 10^9)(s^2 + 4.233 \times 10^4 s + 11.029 \times 10^9)} \\ &= \frac{3.8286 \times 10^{18}}{(s^4 + 5.897 \times 10^4 s^3 + 5.741 \times 10^9 s^2 + 1.843 \times 10^4 s + 4.296 \times 10^{18})} \end{aligned}$$

#### 四、归一化原型滤波器设计数据

由变换法设计数字滤波器,首先要设计符合指标要求的模拟滤波器,而要设计模拟滤波器,一般都有现成的数据表格可查。但是,模拟滤波器有各种类型(低通、高通、带通、带阻、全通等),而且仅就一种类型而言,截止频率又各不相同,可是设计表格不能有那么多,因而一般只给出归一化原型滤波器的设计数据。归一化原型滤波器是指截止频率  $\Omega_c$  已经归一化成  $\Omega_c' = 1$  的低通滤波器,对于截止频率为某个  $\Omega_c$  的低通滤波器,则可用(6-78)式(用到这里时  $\Omega_c = 1$ ),令  $s/\Omega_c$  代替归一化原型滤波器系统函数中的  $s$ ,即

$$s \rightarrow \frac{s}{\Omega_c} \quad (6-107)$$

对于其他高通、带通、带阻滤波器,则可用后面讨论到的频带变换法,由归一化原型滤波器经频带变换得出。

归一化的巴特沃思低通滤波器、切比雪夫低通滤波器的极点分布及相应多项式的系数都已算出,并已列成表格。

归一化原型(低通)系统函数的一般形式是

$$H_m(s) = \frac{d_0}{a_0 + a_1 s + a_2 s^2 + \cdots + a_N s^N} \quad (6-108)$$

表 6-4 巴特沃思多项式  $s^N + a_{N-1}s^{N-1} + \cdots + a_2s^2 + a_1s + 1$  ( $a_0 = a_N = 1$ ) 的系数

$N$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$
1	1								
2	1.4142136								
3	2.0000000	2.0000000							
4	2.6131259	3.4142136	2.6131259						
5	3.2360680	5.2360680	5.2360680	3.2360680					
6	3.8637033	7.4641016	9.1416202	7.4641016	3.8637033				
7	4.4939592	10.0978347	14.5917939	14.5917939	10.0978347	4.4939592			
8	5.1258309	13.1370712	21.8461510	25.6883559	21.8461510	13.1370712	5.1258309		
9	5.7587705	16.5817187	31.1634375	41.9863857	41.9863857	31.1634375	16.5817187	5.7587705	
10	6.3924532	20.4317291	42.8020611	64.8823963	74.2334292	64.8823963	42.8020611	20.4317291	6.3924532

式中分母多项式的系数见表 6-4、表 6-5, 表中分别列出巴特沃思函数和四个通带波纹值(0.5dB, 1dB, 2dB, 3dB)的切贝雪夫函数, 滤波器的阶次由  $N=1$  到  $N=10$ ,  $d_0$  由低频或高频特性确定, 如果希望直流( $\Omega=0$ )增益为 1, 则  $d_0=a_0$ 。

表 6-5 切贝雪夫滤波器分母多项式  $s^N + a_{N-1}s^{N-1} + \dots + a_1s + a_0$  ( $a_N=1$ ) 的系数(见(6-108)式)

$N$	$a_0$	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$a_7$	$a_8$	$a_9$
a. 1/2-dB 波纹( $\epsilon=0.3493114, \epsilon^2=0.1220184$ ).										
1	2.8627752									
2	1.5162026	1.4256245								
3	0.7156938	1.5348954	1.2529130							
4	0.3790506	1.0254553	1.7168662	1.1973856						
5	0.1789234	0.7525181	1.3095747	1.9373675	1.1724909					
6	0.0947626	0.4323669	1.1718613	1.5897635	2.1718446	1.1591761				
7	0.0447309	0.2820722	0.7556511	1.6479029	1.8694079	2.4126510	1.1512176			
8	0.0236907	0.1525444	0.5735604	1.1485894	2.1840154	2.1492173	2.6567498	1.1460801		
9	0.0111827	0.0941198	0.3408193	0.9836199	1.6113880	2.7814990	2.4293297	2.9027337	1.1425705	
10	0.0059227	0.0492855	0.2372688	0.6269689	1.5274307	2.1442372	3.4409268	2.7097415	3.1498757	1.1400664
b. 1-dB 波纹( $\epsilon=0.5088471, \epsilon^2=0.2589254$ ).										
1	1.9652267									
2	1.1025103	1.0977343								
3	0.4913067	1.2384092	0.9883412							
4	0.2756276	0.7426194	1.4539248	0.9528114						
5	0.1228267	0.5805342	0.9743961	1.6888160	0.9368201					
6	0.0689069	0.3070808	0.9393461	1.2021409	1.9308256	0.9282510				
7	0.0307066	0.2136712	0.5486192	1.3575440	1.4287930	2.1760778	0.9231228			
8	0.0172267	0.1073447	0.4478257	0.8468243	1.8369024	1.6551557	2.4230264	0.9198113		
9	0.0076767	0.0706048	0.2441864	0.7863109	1.2016071	2.3781188	1.8814798	2.6709468	0.9175476	
10	0.0043067	0.0344971	0.1824512	0.4553892	1.2444914	1.6129856	2.9815094	2.1078524	2.9194657	0.9159320
c. 2-dB 波纹( $\epsilon=0.7647831, \epsilon^2=5848932$ ).										
1	1.3075603									
2	0.6367681	0.8038164								
3	0.3268901	1.0221903	0.7378216							
4	0.2057651	0.5167981	1.2564819	0.7162150						
5	0.0817225	0.4593491	0.6934770	1.4995433	0.7064606					
6	0.0514413	0.2102706	0.7714618	0.8670149	1.7458587	0.7012257				
7	0.0204228	0.1660920	0.3825056	1.1444390	1.0392203	1.9935272	0.6978929			
8	0.0128603	0.0729373	0.3587043	0.5982214	1.5795807	1.2117121	2.2422529	0.6960646		
9	0.0051076	0.0543756	0.1684473	0.6444677	0.8568648	2.0767479	1.3837464	2.4912897	0.6946793	
10	0.0032151	0.0233347	0.1440057	0.3177560	1.0389104	1.15825287	2.6362507	1.5557424	2.7406032	0.6936904
d. 3-dB 波纹( $\epsilon=0.9976283, \epsilon^2=0.9952623$ ).										
1	1.0023773									
2	0.7079478	0.6448996								
3	0.2505943	0.9283480	0.5972404							
4	0.1769869	0.4047679	1.1691176	0.5815799						
5	0.0626391	0.4079421	0.5488626	1.4149847	0.5744296					
6	0.0442467	0.1634299	0.6990977	0.6906098	1.6628481	0.5706979				
7	0.0156621	0.1461530	0.3000167	1.0518448	0.8314411	1.9115507	0.5684201			
8	0.0110617	0.0564813	0.3207646	0.4718990	1.4666990	0.9719473	2.1607148	0.5669476		
9	0.0039154	0.0475900	0.1313851	0.5834984	0.6789075	1.9438443	1.1122863	2.4101346	0.5659234	
10	0.0027654	0.0180313	0.1277560	0.2492043	0.9499208	0.9210659	2.4834205	1.2526467	2.6597378	0.5652218

表 6-6 巴特沃思多项式  $E(s)$  的根  $s_i$  (见(6-109)式)

$N=1$	$N=2$	$N=3$	$N=4$	$N=5$	$N=6$	$N=7$	$N=8$	$N=9$	$N=10$
-1.0000000	-0.7071068 $\pm j0.7071068$	-1.0000000 $\pm j0.9238795$	-0.3826834 $\pm j0.9238795$	-1.0000000 $\pm j0.9659258$	-0.2588190 $\pm j0.9659258$	-1.0000000 $\pm j0.9807853$	-0.1950903 $\pm j0.9807853$	-1.0000000 $\pm j0.9876883$	-0.156345 $\pm j0.9876883$
				-0.5000000 $\pm j0.8660254$	-0.3090170 $\pm j0.3826834$	-0.7071068 $\pm j0.7071068$	-0.2225208 $\pm j0.9749279$	-0.5555702 $\pm j0.8314696$	-0.1736382 $\pm j0.9848078$
					-0.8090170 $\pm j0.5877852$	-0.9659258 $\pm j0.2588190$	-0.6234898 $\pm j0.7818315$	-0.8314696 $\pm j0.5555702$	-0.5000000 $\pm j0.8660254$
						-0.9009689 $\pm j0.4338837$	-0.9807853 $\pm j0.1950903$	-0.7660444 $\pm j0.6427876$	-0.8910065 $\pm j0.4538905$
							-0.9396926 $\pm j0.3420201$	-0.9876883 $\pm j0.156345$	-0.9876883 $\pm j0.156345$

表 6-7 切贝雪夫滤波器分母多项式  $E(s)$  的根  $s_i$ (见(6-109)式)

$N=1$	$N=2$	$N=3$	$N=4$	$N=5$	$N=6$	$N=7$	$N=8$	$N=9$	$N=10$
a. $1/2-\text{dB}$ 波纹 ( $\epsilon = 0.3493114, \epsilon^2 = 0.1220184$ )									
-2.8627752	-0.7128122	-0.6264565	-0.1753531	-0.3623196	-0.0776501	-0.2561700	-0.0436201	-0.1984053	-0.0278994
$\pm j1.0049425$	$\pm j1.0162529$	$\pm j1.0084608$	$\pm j1.0119629$	$-0.2121440$	$-0.0570032$	$-0.1242195$	$\pm j1.0050021$	$\pm j1.0032732$	
$-0.3132282$	$-0.4233398$	$-0.1119629$	$\pm j1.0155574$	$\pm j1.0115574$	$\pm j1.0064085$	$\pm j0.8519996$	$-0.0344527$	$-0.0809672$	
$\pm j1.0219275$	$\pm j0.4209457$	$-0.2931227$	$-0.2897940$	$-0.1597194$	$-0.1859076$	$\pm j0.5692879$	$\pm j1.0040040$	$\pm j0.9050658$	
$\pm j0.6251768$	$\pm j0.2702162$	$\pm j0.8070770$	$\pm j0.8070770$	$-0.2308012$	$-0.2192929$	$-0.1519873$	$\pm j0.8829063$	$\pm j0.7182643$	
				$\pm j0.4478939$	$\pm j0.1999073$	$\pm j0.6553170$	$\pm j0.4611541$	$\pm j0.1864400$	$-0.1761499$
						$\pm j0.3486869$	$\pm j0.1589029$		
b. $1-\text{dB}$ 波纹 ( $\epsilon = 0.5088471, \epsilon^2 = 0.2589254$ )									
-1.9652267	-0.5488672	-0.4941706	-0.1395360	-0.2894933	-0.0621810	-0.2054141	-0.0350082	-0.1593305	-0.0224144
$\pm j0.8951286$	$\pm j0.9833792$	$\pm j0.9934115$	$-0.1698817$	$-0.0457089$	$\pm j0.9964513$	$\pm j0.9977755$			
$-0.2470853$	$-0.3368697$	$-0.0894584$	$\pm j0.7272225$	$\pm j0.9952839$	$-0.0996950$	$\pm j0.9972297$	$-0.0276674$	$\pm j0.9977755$	
$\pm j0.9659987$	$\pm j0.4073290$	$\pm j0.9901071$	$-0.2342050$	$-0.1280736$	$\pm j0.844706$	$\pm j0.9972297$	$-0.1013166$	$-0.1013166$	
$\pm j0.6119198$	$\pm j0.2661837$	$\pm j0.7981557$	$-0.1492041$	$\pm j0.5644443$	$\pm j0.8769490$	$\pm j0.7143284$	$-0.0796652$	$-0.0650493$	
			$-0.1850717$	$-0.1759983$	$-0.1220542$	$\pm j0.6508954$	$\pm j0.4586271$	$\pm j0.9001063$	
			$\pm j0.4429430$	$\pm j0.1982065$	$-0.1497217$	$\pm j0.3463342$	$\pm j0.1580321$	$-0.1415193$	

续表

$N=1$	$N=2$	$N=3$	$N=4$	$N=5$	$N=6$	$N=7$	$N=8$	$N=9$	$N=10$
c. 2-dB 滤波 ( $\epsilon = 0.7647831, \epsilon^2 = 0.5848932$ )									
-1.3075603	-0.4019082	-0.3689108	-0.1048872	-0.2163083	-0.0469732	-0.1552958	-0.0264924	-0.1206298	-0.0169758
$\pm j0.6893750$		$\pm j0.9579530$	$\pm j0.9817052$		$\pm j0.9817052$		$\pm j0.9897870$		$\pm j0.9934868$
	-0.1844554	-0.2532202	-0.0674610	-0.1283352	-0.0345566	-0.0754439	-0.0209471	-0.0767332	
	$\pm j0.9230771$	$\pm j0.3967971$	$\pm j0.9734557$	$\pm j0.7186581$	$\pm j0.9866139$	$\pm j0.8391009$	$\pm j0.9919471$	$\pm j0.7112580$	
			-0.1766151	-0.1753064	-0.0958253	-0.1129098	-0.0603149	-0.0492657	
			$\pm j0.6016287$	$\pm j0.2630471$	$\pm j0.7912029$	$\pm j0.5606693$	$\pm j0.8723036$	$\pm j0.8962374$	
					-0.1399167	-0.1331862	-0.0924078	-0.0966894	
					$\pm j0.4390845$	$\pm j0.1968809$	$\pm j0.6474475$	$\pm j0.4566558$	
							-0.1133549	-0.1071810	
							$\pm j0.3444996$	$\pm j0.1573528$	
d. 3-dB 滤波 ( $\epsilon = 0.9976283, \epsilon^2 = 0.9952623$ )									
-1.0023773	-0.3224498	-0.2986202	-0.0851704	-0.1775085	-0.038295	-0.1264854	-0.0215782	-0.0982716	-0.0138320
$\pm j0.7771576$		$\pm j0.9464844$	$\pm j0.9764060$		$\pm j0.9867664$		$\pm j0.9867664$		$\pm j0.9915418$
	-0.2056195	-0.0548531	-0.1044450	-0.0281456	-0.0614494	-0.0170647	-0.0401419		
	$\pm j0.9038144$	$\pm j0.3920467$	$\pm j0.9659238$	$\pm j0.7147788$	$\pm j0.9826957$	$\pm j0.8365401$	$\pm j0.9855156$	$\pm j0.8944827$	
			-0.1436074	-0.1426745	-0.0788623	-0.0919655	-0.0491358	-0.0625225	
			$\pm j0.5969738$	$\pm j0.2616272$	$\pm j0.7880608$	$\pm j0.5589382$	$\pm j0.8701971$	$\pm j0.7098655$	
					$\pm j0.4373407$	-0.1139594	-0.1084807	-0.0752804	-0.0787829
						$\pm j0.4373407$	$\pm j0.1962800$	$\pm j0.6458839$	$\pm j0.4557617$
							$\pm j0.3436677$	$\pm j0.3436677$	$\pm j0.1570448$

如将(6-108)式的分母表示成因式形式,令  $E(s)$  代表分母多项式的因式形式,则归一化原型(低通)系统函数可表示成

$$H_m(s) = \frac{d_0}{E(s)} = \frac{d_0}{(s - s_1)(s - s_2) \cdots (s - s_N)} \quad (6-109)$$

其中如有复根,则一定是共轭存在的;如有  $s - s_i$  复根,就一定有  $s - s_i^*$  复根。因式形式的  $E(s)$  的根见表 6-6、表 6-7。

再一次强调,这两个表格都是归一化  $\Omega_c = 1$  的结果,对于具体的(不为 1)的  $\Omega_c$ ,其系统函数  $H_a(s)$  要将归一化系统函数  $H_m(s)$  中的复变量  $s$  利用(6-107)式加以变换而得到。即

$$H_a(s) = H_m(s) \Big|_{s=\frac{s}{\Omega_c}} = H_m\left(\frac{s}{\Omega_c}\right) \quad (6-110)$$

下面利用查表的办法,由归一化的系统函数,导出所需的系统函数。

**[例 6-7]** 给定模拟滤波器指标与例 6-6 相同,即

$$\delta_1 \leq 1\text{dB}, \quad 0 \leq \Omega \leq 2\pi \times 10^4 \text{ rad/sec}$$

$$\delta_2 \geq 15\text{dB}, \quad \Omega \geq 2\pi \times 1.5 \times 10^4 \text{ rad/sec}$$

求用巴特沃思滤波器实现时,所需的阶次、滤波器的截止频率  $\Omega_c$  以及滤波器的系统函数  $H_a(s)$  的表达式。

**解** 滤波器的 1dB 截止频率为  $\Omega_p = 2\pi \times 10^4 \text{ rad/sec}$ , 15dB 的截止频率为  $\Omega_u = 2\pi \times 1.5 \times 10^4 \text{ rad/sec}$ , 而  $\Omega_c$  为待求的巴特沃思滤波器 3dB 的截止频率。

(1) 求  $N$ 。由题意可写出

$$\begin{aligned} 20 \log \left[ \frac{1}{|H_a(j2\pi \times 10^4)|} \right] &= 1 \\ 20 \log \left[ \frac{1}{|H_a(j2\pi \times 1.5 \times 10^4)|} \right] &= 15 \end{aligned}$$

由此得出

$$|H_a(j2\pi \times 10^4)| = 10^{-\frac{1}{20}}$$

$$|H_a(j2\pi \times 1.5 \times 10^4)| = 10^{-\frac{15}{20}}$$

将它们代入巴特沃思滤波器幅度平方函数的表达式(6-73)式中,得

$$\begin{aligned} |H_a(j2\pi \times 10^4)| &= \frac{1}{\sqrt{1 + \left(\frac{2\pi \times 10^4}{\Omega_c}\right)^{2N}}} = 10^{-\frac{1}{20}} \\ |H_a(j2\pi \times 1.5 \times 10^4)| &= \frac{1}{\sqrt{1 + \left(\frac{2\pi \times 1.5 \times 10^4}{\Omega_c}\right)^{2N}}} = 10^{-\frac{15}{20}} \end{aligned}$$

将此两式对  $N$  求解,可得

$$N = \log\left(\frac{10^{1.5} - 1}{10^{0.1} - 1}\right) / 2 \log\left(\frac{2\pi \times 1.5 \times 10^4}{2\pi \times 10^4}\right) = 5.885$$

取  $N=6$ 。

(2) 求 3dB 处的截止频率  $\Omega_c$ 。把  $N=6$  代入到阻带指标要求的公式中,阻带指标满

足要求,但是由于  $N=6 > 5.885$ ,因而通带指标一定超过性能要求。将  $N=6$  代入阻带条件,可得

$$|H_a(j2\pi \times 1.5 \times 10^4)| = \frac{1}{\sqrt{1 + \left(\frac{2\pi \times 1.5 \times 10^4}{\Omega_c}\right)^2}} = 10^{-\frac{15}{20}}$$

则有

$$\Omega_c = 2\pi \times 1.13 \times 10^4 \text{ rad/sec}$$

(3) 由于  $N$  和  $\Omega_c$  都已求出,可以用查表法来求解。查巴特沃思滤波器的表 6-4(分母为  $N$  阶多项式)或表 6-6(分母为因式分解形式),可得到归一化  $\Omega_c=1$  的系统函数为

$$\begin{aligned} H_{ar}(s) &= \frac{1}{s^6 + 3.8637033s^5 + 7.4641016s^4 + 9.1416202s^3 + 7.4641016s^2 + 3.8637033s + 1} \\ &= \frac{1}{(s^2 + 0.517638s + 1)(s^2 + 1.4142136s + 1)(s^2 + 1.9318516s + 1)} \end{aligned}$$

利用(6-110)式( $s \rightarrow \frac{s}{\Omega_c}$ )可得到截止频率为  $\Omega_c = 2\pi \times 1.13 \times 10^4 \text{ rad/sec}$  时的系统函数

$$\begin{aligned} H_a(s) &= \frac{\Omega_c^6}{s^6 + 3.8637033\Omega_c^2 s^5 + 7.4641016\Omega_c^4 s^4 + 9.1416202\Omega_c^3 s^3 + 7.4641016\Omega_c^2 s^2 + 3.8637033\Omega_c^4 s + \Omega_c^6} \\ &= \frac{\Omega_c^6}{(s^2 + 0.517638\Omega_c s + \Omega_c^2)(s^2 + 1.4142136\Omega_c^2 s + \Omega_c^2)(s^2 + 1.9318516\Omega_c s + \Omega_c^2)} \end{aligned}$$

代入  $\Omega_c$  的值,得

$$\begin{aligned} H_a(s) &= \frac{1.281 \times 10^{29}}{s^6 + 2.743 \times 10^5 s^5 + 3.763 \times 10^{10} s^4 + 3.272 \times 10^{15} s^3 + 1.897 \times 10^{20} s^2 + 6.971 \times 10^{24} s + 1.281 \times 10^{29}} \\ &= \frac{1.281 \times 10^{29}}{(s^2 + 3.675 \times 10^4 s + 5.041 \times 10^9)(s^2 + 10.041 \times 10^4 s + 5.041 \times 10^9)(s^2 + 13.716 \times 10^4 s + 5.041 \times 10^9)} \end{aligned}$$

与例 6-5 比较,在同样性能指标下,巴特沃思滤波器所用的阶次( $N=6$ )比切贝雪夫滤波器的阶次( $N=4$ )要多。

## 6.9 设计 IIR 滤波器的频率变换法

同样,实际应用中的数字滤波器也有低通、高通、带通、带阻等类型。这样设计各类数字滤波器可以有以下两种方法:

(1) 把一个归一化原型模拟低通滤波器经模拟频带变换成为所需要类型(包括高通、带通、带阻与另一截止频率的低通)的模拟滤波器,然后再通过冲激响应不变法或双线性变换法数字化为所需类型的数字滤波器,如图 6-23(a)所示。实际上是把这一方法中的两步合成一步来实现,即把模拟归一化低通原型变换到模拟低通、高通、带通、带阻等滤波器的公式与用双线性变换得到相应数字滤波器的公式合并,就可直接从模拟低通归一化原型通过一定的频率变换关系,一步完成各类型数字滤波器的设计,如图 6-23(b)所示。这里只谈了双线性变换法,因为冲激响应不变法是有频率响应混叠失真效应的,只对设计严格

能限带的数字低通、带通滤波器才能应用,对于数字高通、带阻滤波器,不能直接应用。

(2) 由模拟低通原型先利用冲激响应不变法或双线性变换法数字化成数字低通滤波器,然后利用数字频带变换法,将它转换成所需要的各型数字滤波器(另一截止频率的数字低通,数字高通、数字带通、数字带阻等),如图 6-23(c)所示。

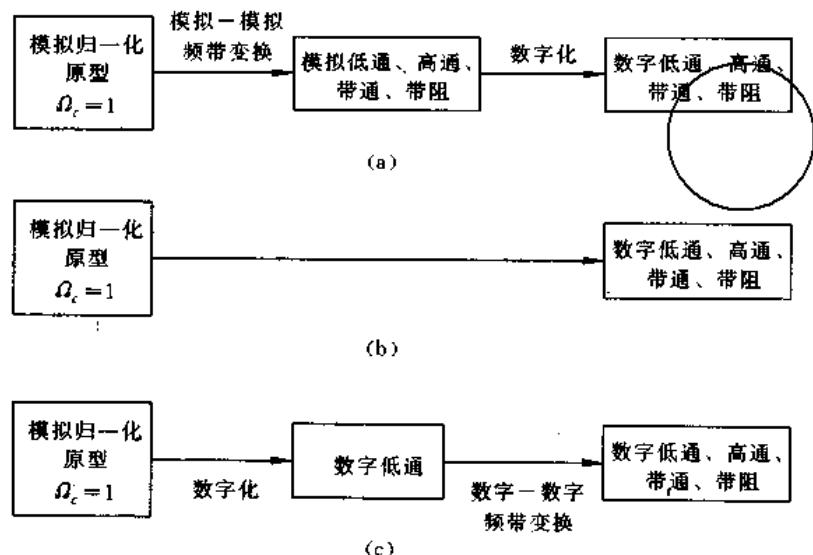


图 6-23 设计 IIR 数字滤波器的频率变换法

(a) 先模拟频带变换,再数字化; (b) 把(a)的两步合成一步直接设计;

(c) 先数字化,再进行数字频带变换

## 6.10 先利用模拟域频带变换法,再利用数字化法设计数字各型滤波器

### 一、模拟低通滤波器变成数字低通滤波器

首先把数字滤波器的性能要求转换为与之相应的作为“样本”的模拟滤波器的性能要求,根据此性能要求设计模拟滤波器,这可以用查表的办法,也可以用解析的方法。然后通过冲激响应不变法或双线性变换法,将此“样本”模拟低通滤波器  $H_a(s)$  数字化为所需的数字滤波器  $H(z)$ 。下面用实例来加以说明。

#### 1. 冲激响应不变法设计数字巴特沃思低通滤波器

给定抽样频率  $f_s = 10\text{kHz}$ , 要求在频率小于  $1\text{kHz}$  的通带内, 幅度特性下降小于  $1\text{dB}$ ; 在频率大于  $f_u = 1.5\text{kHz}$  的阻带内, 衰减大于  $15\text{dB}$ 。

**解** (1) 讨论  $f$  与  $\omega$  的关系及数字域性能的公式表示。模拟频率与数字频率之间为线性关系,  $\omega = \Omega T = 2\pi f T = 2\pi \frac{f}{f_s}, T = \frac{1}{f_s} = 10^{-4}$ ,  $T$  为抽样周期, 故

$$f_c = 1\text{kHz} \quad \text{对应于} \quad \omega_c = 2\pi \times 1 \times 10^3 \times \frac{1}{10 \times 10^3} = 0.2\pi$$

$$f_a = 1.5 \text{ kHz} \rightarrow \omega_a = 2\pi \times 1.5 \times 10^3 \times \frac{1}{10 \times 10^3} = 0.3\pi$$

按衰减的定义

$$20 \log_{10} \left| \frac{H(e^{j0})}{H(e^{j0.2\pi})} \right| \leq 1 \quad (6-111a)$$

$$20 \log_{10} \left| \frac{H(e^{j0})}{H(e^{j0.3\pi})} \right| \geq -15 \quad (6-111b)$$

设  $\omega=0$  处频率响应幅度归一化为 1, 即  $|H(e^{j0})|=1$ , 则上两式变成

$$20 \log_{10} |H(e^{j0.2\pi})| \geq -1 \quad (6-112a)$$

$$20 \log_{10} |H(e^{j0.3\pi})| \leq -15 \quad (6-112b)$$

这就是数字滤波器的性能指标的表达式。

(2) 下面把数字滤波器性能要求转变为“样本”模拟滤波器的性能要求。由于  $\omega=\Omega T$ , 按修正后的(6-37)式, 设没有混叠效应(即混叠效应设计完成后进行校验), 则有

$$H(e^{j\omega}) = H_a \left( j \frac{\omega}{T} \right) = H_a(j\Omega), \quad |\omega| \leq \pi \quad (6-113)$$

利用(6-113)式, 由(6-112)式可写出模拟滤波器的指标为

$$20 \log_{10} \left| H_a \left( j \frac{0.2\pi}{T} \right) \right| = 20 \log_{10} |H_a(j2\pi \times 10^3)| \geq -1 \quad (6-114a)$$

$$20 \log_{10} \left| H_a \left( j \frac{0.3\pi}{T} \right) \right| = 20 \log_{10} |H_a(j3\pi \times 10^3)| \leq -15 \quad (6-114b)$$

(3) 计算“样本”模拟滤波器所需的阶数  $N$  及 3dB 截止频率  $\Omega_c$ 。巴特沃思滤波器的幅度平方函数是

$$|H_a(j\Omega)|^2 = \frac{1}{1 + \left( \frac{\Omega}{\Omega_c} \right)^{2N}} \quad (6-115)$$

以分贝值形式表示(6-115)式, 即

$$20 \log_{10} |H_a(j\Omega)| = -10 \log_{10} \left[ 1 + \left( \frac{\Omega}{\Omega_c} \right)^{2N} \right] \quad (6-116)$$

把求出的性能指标关系(6-114a)式、(6-114b)式代入(6-116)式, 得

$$-10 \log_{10} \left[ 1 + \left( \frac{2\pi \times 10^3}{\Omega_c} \right)^{2N} \right] \geq -1 \quad (6-117a)$$

$$-10 \log_{10} \left[ 1 + \left( \frac{3\pi \times 10^3}{\Omega_c} \right)^{2N} \right] \leq -15 \quad (6-117b)$$

先用等号来满足指标, 可得

$$1 + \left( \frac{2\pi \times 10^3}{\Omega_c} \right)^{2N} = 10^{-1} \quad (6-118a)$$

$$1 + \left( \frac{3\pi \times 10^3}{\Omega_c} \right)^{2N} = 10^{-1.5} \quad (6-118b)$$

解此两方程,得  $N=5.8858$ ,  $\Omega_c=7.04743 \times 10^3$ 。 $N$  是滤波器阶次,必须取整数,为了满足或超过给定指标,故应选取比求出的  $N$  大一点的整数,取  $N=6$ ,代入通带条件(6-118a)式,得  $\Omega_c=7.0321 \times 10^3$ ,即

$$N=6, \quad \Omega_c=7.0321 \times 10^3 \quad (6-119)$$

显然,利用(6-119)式的  $N$  和  $\Omega_c$ ,通带指标正好满足,而阻带指标则可超过,这正好对减小冲激响应不变法造成的频谱混叠效应是有利的。

(4) 用查表法,查表 6-4、表 6-6。当  $N=6$  时,归一化原型模拟低通巴特沃思滤波器的频率响应为

$$\begin{aligned} H_a(s) &= \frac{1}{s^6 + 3.8637033s^5 + 7.4641016s^4 + 9.1416202s^3 + 7.4641016s^2 + 3.8637033s + 1} \\ &= \frac{1}{(s^2 + 0.51763809s + 1)(s^2 + 1.41421356s + 1)(s^2 + 1.931851652s + 1)} \end{aligned} \quad (6-120)$$

将  $s$  用  $\frac{s}{\Omega_c} = \frac{s}{7.0321 \times 10^3}$  代入,可得截止频率为  $\Omega_c$  的模拟低通滤波器

$$\begin{aligned} H_a(s) &= 120923.14 \times 10^{18} / [(s^2 + 3.64003 \times 10^3 s + 49.4504 \times 10^6) \\ &\quad \times (s^2 + 9.94475 \times 10^3 s + 49.4504 \times 10^6) \\ &\quad \times (s^2 + 13.58478 \times 10^3 s + 49.4504 \times 10^6)] \end{aligned} \quad (6-121)$$

另一种办法是,将  $\Omega_c=7.0321 \times 10^3 \text{ rad/s}$ ,  $N=6$  代入(6-77)式,得到  $s$  平面左半平面的三对极点为

$$\begin{aligned} \Omega_c \left[ \cos\left(\frac{7}{12}\pi\right) \pm j\sin\left(\frac{7}{12}\pi\right) \right] &= (-1.82001553 \pm j6.79239041) \times 10^3 \\ \Omega_c \left[ \cos\left(\frac{9}{12}\pi\right) \pm j\sin\left(\frac{9}{12}\pi\right) \right] &= (-4.97237489 \pm j4.97237489) \times 10^3 \quad (6-122) \\ \Omega_c \left[ \cos\left(\frac{11}{12}\pi\right) \pm j\sin\left(\frac{11}{12}\pi\right) \right] &= (-6.79239041 \pm j1.82001553) \times 10^3 \end{aligned}$$

利用(6-122)式,由(6-76)式即可得到模拟滤波器系统函数,结果与(6-121)式相同。

(5) 将  $H_a(s)$  展成部分分式,然后利用冲激响应不变法修正后的公式(6-36),可得所需数字滤波器的系统函数

$$\begin{aligned} H(z) &= \frac{0.2871 - 0.4466z^{-1}}{1 - 1.297z^{-1} + 0.6949z^{-2}} + \frac{-2.1428 + 1.1454z^{-1}}{1 - 1.0691z^{-1} + 0.3699z^{-2}} \\ &\quad + \frac{1.8558 - 0.6304z^{-1}}{1 - 0.9972z^{-1} + 0.2570z^{-2}} \end{aligned}$$

当  $z=e^{j\omega}$  时,即得到数字滤波器的频率响应,如图 6-24 所示。

可以看出,它在通带边缘( $\omega=0.2\pi$ )处恰好满足衰减小于 1dB 的要求,而在阻带边缘( $\omega=0.3\pi$ )处,衰减则大于 15dB,超过性能要求。这表明此滤波器是充分限带的,故混叠效应可以忽略。若得出的数字滤波器不满足指标要求,则可采用更高阶的滤波器,或调整滤波器参数而维持  $N$  不变,再行计算。

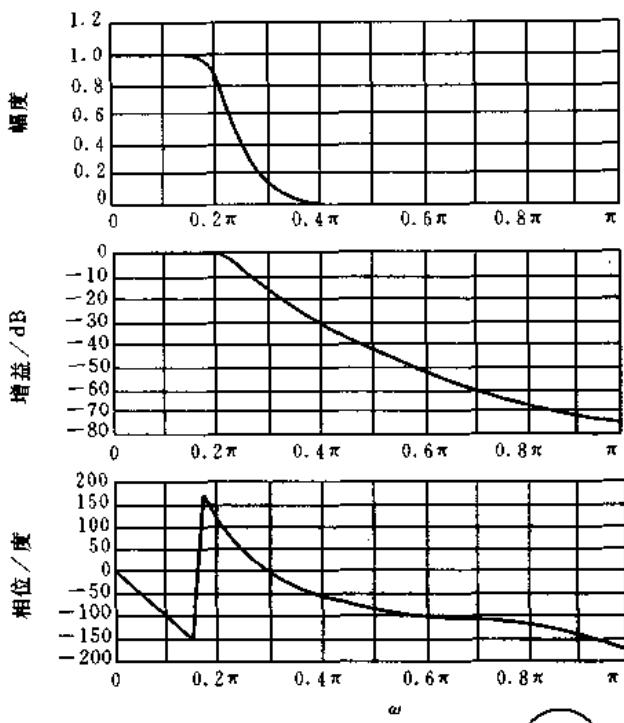


图 6-24 用冲激响应不变法设计出的六阶巴特沃思低通数字滤波器频率响应

## 2. 双线性变换法设计数字切贝雪夫低通滤波器

这里仍利用上一实例的指标,但是直接由数字域给定指标,即在  $\omega \leq 0.2\pi$  的通带范围内幅度特性下降小于 1dB,在  $0.3\pi \leq \omega \leq \pi$  的阻带范围内衰减大于 15dB。

解 (1) 首先导出指标的公式表达式,仍为(6-112a)式与(6-112b)式。

(2) 由于采用双线性变换,  $\omega$  和  $\Omega$  之间是非线性关系,因而将数字域指标变为模拟域指标时,首先要将数字截止频率按(6-62)式预畸为模拟“样本”滤波器的截止频率。

重写(6-62)式为

$$\Omega = c \tan \frac{\omega}{2} \quad (6-123)$$

前面已说过,  $c$  是调节模拟频带与数字频带间对应关系的一个常数,我们采用使模拟频率特性与数字频率特性在低频处有较确切对应关系的常数  $c=2/T$ ,则有

$$\Omega = \frac{2}{T} \tan \frac{\omega}{2} \quad (6-124)$$

利用此式,求出对应的模拟“样本”滤波器的有关频率。

这样,模拟系统函数可写成

$$H_a(j\Omega) = H_a\left(j \frac{2}{T} \tan \frac{\omega}{2}\right) \quad (6-125)$$

利用(6-112a)与(6-112b)两式,将数字域指标变成模拟“样本”的指标,用(6-125)式的  $H_a(j\Omega)$  代替两式中的  $H(e^{j\omega})$ ,且将  $\omega=0.2\pi, \omega=0.3\pi$  分别代入(6-125)式,为了简化,令  $T=1$ ,则得到对模拟滤波器的指标要求为

$$20 \log_{10} \left| H_a \left( j2 \tan \frac{0.2\pi}{2} \right) \right| \geq -1 \quad (6-126a)$$

$$20 \log_{10} \left| H_a \left( j2 \tan \frac{0.3\pi}{2} \right) \right| \leq -15 \quad (6-126b)$$

由此看出，在把数字滤波器的性能要求转化为模拟滤波器的性能要求中，已经考虑了频率预畸。

(3) 求  $\epsilon$ 。由于我们考虑的是波纹为  $\delta_1=1$ dB 的切贝雪夫等波纹模拟滤波器，而

$$\Omega_c = 2 \tan \frac{0.2\pi}{2} = 0.6498394$$

就是 1dB 的截止频率，根据(6-85)式可求得

$$\epsilon = \sqrt{10^{2/\text{10}} - 1} = \sqrt{10^{0.1} - 1} = 0.508847 \quad (6-127)$$

(4) 根据(6-88)式计算滤波器阶次  $N$ 。由于  $\delta_2=15$ dB，故  $10^{0.1\delta_2}=10^{1.5}$ ，故

$$\Omega_c = 2 \tan \left( \frac{0.2\pi}{2} \right), \quad \Omega_s = 2 \tan \left( \frac{0.3\pi}{2} \right) \quad (6-128)$$

代入(6-88)式，可得

$$\begin{aligned} N &\geq \frac{\operatorname{ch}^{-1} \left[ \frac{1}{\epsilon} \sqrt{10^{0.1\delta_2} - 1} \right]}{\operatorname{ch}^{-1} \left( \frac{\Omega_s}{\Omega_c} \right)} = \frac{\operatorname{ch}^{-1} \left[ \frac{1}{0.50885} \sqrt{10^{1.5} - 1} \right]}{\operatorname{ch}^{-1} \left[ \frac{2 \tan \left( \frac{0.3\pi}{2} \right)}{2 \tan \left( \frac{0.2\pi}{2} \right)} \right]} \\ &= \frac{\operatorname{ch}^{-1}(10.8750811)}{\operatorname{ch}^{-1}(1.568158088)} = \frac{3.0775}{1.02104} = 3.0141 \quad (6-129) \end{aligned}$$

选定  $N=4$ 。

(5) 计算  $\gamma$  及  $a, b$ 。由(6-103)式得

$$\gamma = \frac{1}{\epsilon} + \sqrt{\frac{1}{\epsilon^2} + 1} = 4.1702$$

由(6-104a)式得

$$a = \frac{1}{2} (\gamma^{\frac{1}{N}} - \gamma^{-\frac{1}{N}}) = 0.3646235$$

由(6-104b)式得

$$b = \frac{1}{2} (\gamma^{\frac{1}{N}} + \gamma^{-\frac{1}{N}}) = 1.0644015$$

从而求得

$$a\Omega_c = 0.2369322, \quad b\Omega_c = 0.691648$$

(6) 求左半平面的两对极点。由(6-101a)式及(6-101b)式，当  $k=1, 2, 3, 4$  时，可得左半平面两对极点

$$-a\Omega_c \sin \frac{\pi}{8} \pm jb\Omega_c \cos \frac{\pi}{8} = -0.0906699 \pm j0.6389997$$

$$-a\Omega_c \sin \frac{3\pi}{8} \pm jb\Omega_c \cos \frac{3\pi}{8} = -0.2188969 \pm j0.2646819$$

(7) 求模拟滤波器的系统函数  $H_a(s)$ 。按(6-105)式可得

$$H_a(s) = \frac{0.04381}{(s^2 + 0.4378s + 0.1180)(s^2 + 0.1814s + 0.4166)} \quad (6-130)$$

分子中的常数 0.04381 是这样求出的:由图 6-19 可知,  $N$  为偶数时, 在  $s=0$  处(即  $\Omega=0$ ),  $|H_a(j0)| = \frac{1}{\sqrt{1+\epsilon^2}} = 0.891250942$ , 它应与  $s=0$  处(即  $\Omega=0$ )(6-130)式的值相等, 从而求得分子中常数。或者直接用(6-105b)式求出(6-130)式分子中的常数, 结果是一样的。

实际上, 在求出第(4)步以后, 利用波纹分贝数  $\delta=1$ dB 及所求得的  $N=4$ , 可以直接查表 6-7, 得到归一化原型切贝雪夫模拟滤波器的系统函数

$$H_{an}(s) = \frac{0.2456533}{(s^2 + 0.2790720s + 0.9865049)(s^2 + 0.6737394s + 0.2793981)} \quad (6-131)$$

按(6-110)式, 把(6-131)式中的  $s$  换成  $s/\Omega_c$ , 即可求得所需的系统函数, 和(6-130)式完全相同。这样设计就方便多了。

(8) 求数字滤波器系统函数  $H(z)$ 。对  $H_a(s)$  作双线性变换,  $T=1$  时可得

$$\begin{aligned} H(z) &= H_a(s) \Big|_{s=\frac{2}{T} \frac{1-z^{-1}}{1+z^{-1}}} \\ &= \frac{0.001836(1+z^{-1})^4}{(1 - 1.4996z^{-1} + 0.8482z^{-2})(1 - 1.5548z^{-1} + 0.6493z^{-2})} \end{aligned} \quad (6-132)$$

代入  $z=e^{j\omega}$ , 可得所设计出的滤波器的频率响应  $H(e^{j\omega})$ , 其幅频特性  $20 \lg |H(e^{j\omega})|$  如图 6-25 所示。

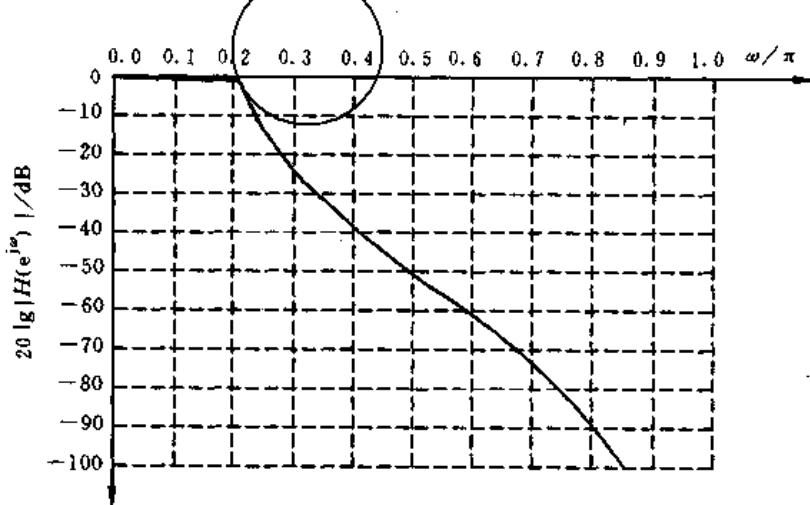


图 6-25 用双线性变换法设计出的四阶切贝雪夫低通数字滤波器幅频响应

当然, 也可以用冲激响应不变法设计切贝雪夫数字低通滤波器, 或用双线性变换法设计巴特沃思数字低通滤波器, 而其中有些步骤和这两个例子中的步骤是相同的。读者可以自行练习。

## 二、模拟低通滤波器转换成数字带通滤波器

### 1. 由模拟低通到模拟带通的变换

变换关系为

$$s = p + \frac{\bar{\Omega}_0^2}{p} \quad (6-133)$$

其中  $s$  是模拟低通拉普拉斯变量 ( $s = \sigma + j\omega$ )， $p$  是模拟带通拉普拉斯变量 ( $p = \bar{\sigma} + j\bar{\Omega}$ )， $\bar{\Omega}_0$  是模拟带通的几何中心频率。令  $p = j\bar{\Omega}$ ，代入 (6-133) 式，可得

$$s = j\bar{\Omega} - j \frac{\bar{\Omega}_0^2}{\bar{\Omega}} = j \left( \bar{\Omega} - \frac{\bar{\Omega}_0^2}{\bar{\Omega}} \right)$$

故  $p$  平面的虚轴与  $s$  平面的虚轴相对应，代入  $s = j\omega$ ，消去  $j$ ，可得

$$\Omega = \frac{\bar{\Omega}^2 - \bar{\Omega}_0^2}{\bar{\Omega}} \quad (6-134)$$

按 (6-134) 式画出的  $\bar{\Omega} \geq 0$  时的图形如图 6-26 所示。按 (6-134) 式及此图可看出，对应频率的关系为

$$\Omega = 0 \rightarrow \bar{\Omega} = \bar{\Omega}_0, \quad \Omega = \Omega_c \rightarrow \bar{\Omega} = \bar{\Omega}_2, \quad \Omega = -\Omega_c \rightarrow \bar{\Omega} = \bar{\Omega}_1$$

$\Omega_c$  为低通滤波器的通带截止频率，故低通的通带映射到带通的  $\bar{\Omega}_1$  到  $\bar{\Omega}_2$  之间，见图 6-27

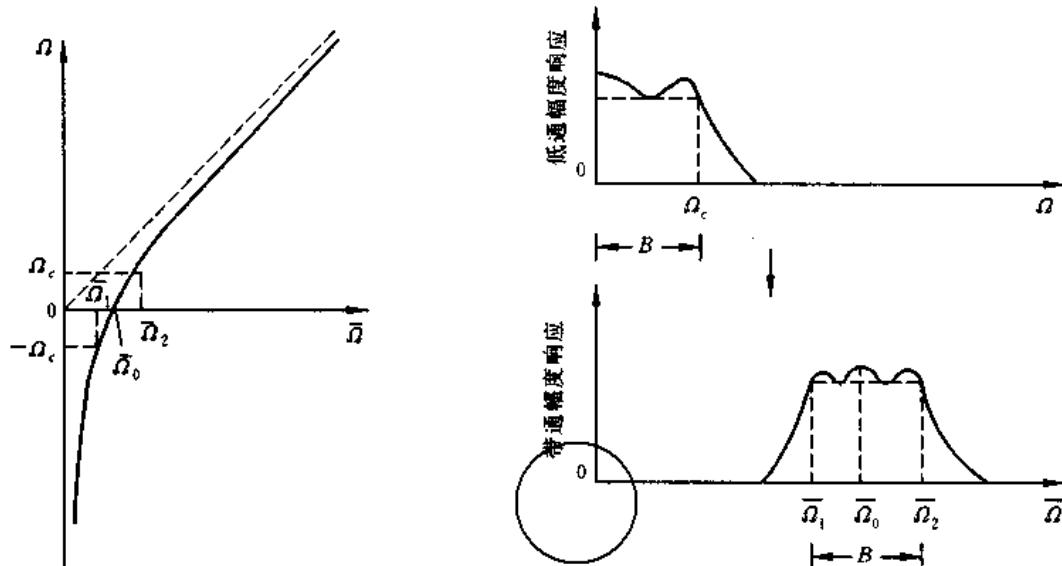


图 6-26 模拟低通滤波器到模拟带通滤波器频率的变换关系

图 6-27 低通幅度响应到带通幅度响应的变换

所示，它是带通滤波器的通带。把上面的映射结果代到 (6-134) 式中，可得

$$\Omega_c = \frac{\bar{\Omega}_2^2 - \bar{\Omega}_0^2}{\bar{\Omega}_2} \quad (6-135)$$

$$-\Omega_c = \frac{\bar{\Omega}_1^2 - \bar{\Omega}_0^2}{\bar{\Omega}_1} \quad (6-136)$$

由(6-135)式和(6-136)式的相加与相减，并化简后可得以下两个有用的关系式

$$\bar{\Omega}_0 = \sqrt{\bar{\Omega}_1 \bar{\Omega}_2} \quad (6-137)$$

$$B = \bar{\Omega}_2 - \bar{\Omega}_1 = \Omega_c \quad (6-138)$$

$B$  为带通滤波器的带宽，它也等于低通滤波器的带宽，带通滤波器在  $\bar{\Omega}_1$  与  $\bar{\Omega}_2$  处的幅度响应与低通滤波器在  $\Omega_c$  处的幅度响应相对应。由(6-137)式看出， $\bar{\Omega}_0$  是带通滤波器通带的几何对称中心角频率。

因此，由模拟低通滤波器来确定模拟带通滤波器的方法如下：

(1) 定出低通系统函数  $H_{LP}(s)$ ，其截止频率  $\Omega_c$  应等于所需模拟带通滤波器的带宽  $\Omega_c = B = \bar{\Omega}_2 - \bar{\Omega}_1$ ；

(2) 在  $H_{LP}(s)$  中代入低通到带通的变换关系式(6-133)，即得带通滤波器的系统函数

$$H_{BP}(s) = H_{LP}(s) \Big|_{s=\frac{p^2+\bar{\Omega}_0^2}{p}}$$

## 2. 由模拟带通到数字带通的变换

利用双线性变换法，即利用

$$p = c \frac{1-z^{-1}}{1+z^{-1}} \quad (6-139)$$

就可将模拟带通系统函数变成数字带通系统函数。

3. 把(6-133)式的变换与(6-139)式的变换结合起来，可得到直接从模拟低通原型的  $s$  平面变换成数字带通滤波器的  $z$  平面的表达式，也就是直接联系  $s$  与  $z$  之间的变换公式

$$s = \frac{p^2 + \bar{\Omega}_0^2}{p} = \left[ \frac{c^2(1-z^{-1})^2}{(1+z^{-1})^2} + \bar{\Omega}_0^2 \right] / \left[ \frac{c(1-z^{-1})}{1+z^{-1}} \right]$$

经推导后得

$$s = \frac{c^2 + \bar{\Omega}_0^2}{c} \cdot \frac{z^{-2} - 2 \left[ \frac{c^2 - \bar{\Omega}_0^2}{c^2 + \bar{\Omega}_0^2} \right] z^{-1} + 1}{1 - z^{-2}} \quad (6-140)$$

下面来化简(6-140)式。利用双线性变换的频率间关系

$$\bar{\Omega} = ctan \frac{\omega}{2} \quad (6-141)$$

将(6-141)式代入(6-137)式、(6-138)式，可得

$$\tan^2 \frac{\omega_0}{2} = \tan \frac{\omega_1}{2} \cdot \tan \frac{\omega_2}{2} \quad (6-142)$$

$$\tan \frac{\omega_2}{2} - \tan \frac{\omega_1}{2} = \frac{\Omega_c}{c} \quad (6-143)$$

考虑到模拟带通到数字带通是通带中心频率相对应的映射关系，设数字带通的中心频率

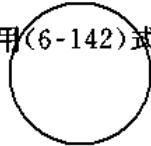
为  $\omega_0$ , 则有

$$\bar{\Omega}_0 = c \tan \frac{\omega_0}{2} \quad (6-144)$$

因而

$$c = \bar{\Omega}_0 \cot \frac{\omega_0}{2} \quad (6-145)$$

设  $D = \frac{c^2 + \bar{\Omega}_0^2}{c}$ , 用(6-142)式、(6-143)式、(6-145)式和一些标准三角恒等式可得



$$D = \bar{\Omega}_0 \cot\left(\frac{\omega_2 - \omega_1}{2}\right) \quad (6-146)$$

设  $E = 2 \frac{c^2 - \bar{\Omega}_0^2}{c^2 + \bar{\Omega}_0^2}$ , 同样可以得出

$$E = 2 \frac{\cos\left(\frac{\omega_2 + \omega_1}{2}\right)}{\cos\left(\frac{\omega_2 - \omega_1}{2}\right)} = \frac{2 \sin(\omega_1 + \omega_2)}{\sin\omega_1 + \sin\omega_2} = 2 \cos\omega_0 \quad (6-147)$$

这样,(6-140)式可以表示成以下更简练的形式:

$$s = D \frac{z^{-2} - Ez^{-1} + 1}{1 - z^{-2}} \quad (6-148)$$

所以,在设计时,要给定中心频率和带宽或者是中心频率和边带频率,利用(6-146)式和(6-147)式来确定  $D$  和  $E$  两常数,然后利用(6-148)式的变换,把模拟低通系统函数  $H_{LP}(s)$  一步变成数字带通系统函数

$$H(z) = H_{LP}(s) \Big|_{s=D \frac{z^{-2}-Ez^{-1}+1}{1-z^{-2}}} \quad (6-149)$$

可以看出,数字带通滤波器的极点数(或阶数)将是模拟低通滤波器极点数的两倍。

下面来讨论模拟低通滤波器和数字带通滤波器频率之间的关系。令

$$s = j\Omega, z = e^{j\omega}$$

代入(6-148)式,经推导后可得

$$\Omega = D \frac{\cos\omega_0 - \cos\omega}{\sin\omega} \quad (6-150)$$

其变换关系曲线如图 6-28 所示,其映射关系为

$$\Omega = 0 \rightarrow \omega = \omega_0$$

$$\Omega = \infty \rightarrow \omega = \pi$$

$$\Omega = -\infty \rightarrow \omega = 0$$

也就是说,低通滤波器的通带( $\Omega=0$ 附近)映射到带通滤波器的通带( $\omega=\omega_0$ 附近),低通的阻带( $\Omega=\pm\infty$ )映射到带通的阻带( $\omega=0, \pi$ )。

[例 6-8] 设计一个抽样频率为  $f_s = 1\text{kHz}$  的数字带通滤波器,其性能要求为:通带范围从 200Hz 到 250Hz,在此两频率处衰减不大于 3dB,

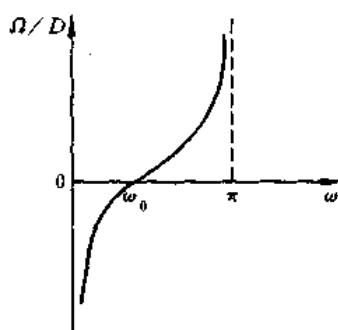


图 6-28 从模拟低通变换到数字带通时频率间关系的曲线

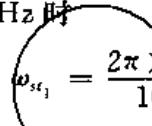
在 100Hz 和 400Hz 频率处衰减不得小于 20dB, 采用巴特沃思型滤波器。

解 先求出所需数字滤波器在数字域的各个临界频率。通带的上下边界频率为

$$\omega_2 = \Omega_2 T = \frac{\Omega_2}{f_s} = \frac{2\pi \times 250}{1000} = 0.5\pi$$

$$\omega_1 = \Omega_1 T = \frac{\Omega_1}{f_s} = \frac{2\pi \times 200}{1000} = 0.4\pi$$

阻带边界频率为, 当频率为 100Hz 时



$$\omega_{st_1} = \frac{2\pi \times 100}{1000} = 0.2\pi$$

当频率为 400Hz 时

$$\omega_{st_2} = \frac{2\pi \times 400}{1000} = 0.8\pi$$

由(6-146)式求得

$$D = \Omega_c \cot\left(\frac{\omega_2 - \omega_1}{2}\right) = \Omega_c \cot\left(\frac{0.5\pi - 0.4\pi}{2}\right) = \Omega_c \cot(0.05\pi)$$

取  $\Omega_c = 1 \text{ rad/s}$ , 则  $D = 6.313751515$ 。由(6-147)式可求得  $E$  为

$$E = 2\cos\omega_0 = \frac{\cos\left(\frac{\omega_2 + \omega_1}{2}\right)}{\cos\left(\frac{\omega_2 - \omega_1}{2}\right)} = 0.31676888$$

设  $\Omega_{st}$  为满足数字带通滤波器要求的“样本”模拟低通原型的阻带起始频率, 可按(6-150)式的变换关系求得

$$\Omega_{st1} = D \frac{\cos\omega_0 - \cos\omega_{st_1}}{\sin\omega_{st_1}} = -6.98883225$$

$$\Omega_{st2} = D \frac{\cos\omega_0 - \cos\omega_{st_2}}{\sin\omega_{st_2}} = 10.39143462$$

取  $\Omega_{st} = \min(|\Omega_{st1}|, |\Omega_{st2}|) = 6.98883225$ , 巴特沃思低通响应, 在此  $\Omega_{st}$  处, 衰减要大于 20dB, 故有

$$-10 \log_{10} \left[ 1 + \left( \frac{\Omega_{st}}{\Omega_c} \right)^{2N} \right] \leq -20$$

可解出

$$N \geq \frac{\log_{10}(10^2 - 1)}{2 \log_{10}(6.98883225)} = 1.181681836$$

选  $N=2$ , 查表 6-4 可得二阶巴特沃思滤波器的归一化原型系统函数为

$$H_{LP}(s) = \frac{1}{s^2 + 1.4142136s + 1}$$

将变换关系式(6-148)代入此式, 并将求出的  $D, E$  参量代入, 可得

$$H(z) = H_{LP}(s) \Big|_{s=\frac{1-z^{-1}+z^{-2}}{1-z^{-2}}}$$

$$\begin{aligned}
&= (1 - z^{-2})^2 \left[ (D^2 + \sqrt{2}D + 1) - (2D^2E + \sqrt{2}DE)z^{-1} + (D^2E^2 + 2D^2 - 2)z^{-2} \right. \\
&\quad \left. + (\sqrt{2}DE - 2ED^2)z^{-3} + (D^2 - \sqrt{2}D + 1)z^{-4} \right] \\
&= 0.02025069(1 - z^{-2})^2 / (1 - 0.5636968z^{-1} + 1.639294z^{-2} \\
&\quad - 0.4496743z^{-3} + 0.6400452z^{-4})
\end{aligned}$$

读者一定会问,这里为什么采用归一化低通原型(截止频率为1)?这是因为,对于某一具体  $\Omega_c$ (不等于1),设计滤波器时要用  $s/\Omega_c$  代替归一化原型中的  $s$ ,则实际模拟滤波器的系统函数为(以本题为例,一般情况下也是正确的)

$$H'_{LP}(s) = \frac{1}{\frac{s^2}{\Omega_c^2} + \frac{s^2}{\Omega_c^2} + 1}$$

变成数字滤波器

$$H(z) = H'_{LP}(s) \Big|_{s=D\frac{1-e^{-j\omega_c}}{1-e^{-j\omega_c}} \cdot \Omega_c \cos\left(\frac{\omega_2-\omega_1}{2}\right)} \left(\frac{1-e^{-j\omega_c}}{1-e^{-j\omega_c}}\right)$$

这样,  $s/\Omega_c$  中的  $\Omega_c$  就和  $D$  中的  $\Omega_c$  互相抵消,所以只须用  $\Omega_c=1$  的归一化原型  $H_{LP}(s)$  设计即可。

利用本书所录光盘的“辅助设计”子系统,所得设计结果附在本章最后面。

### 三、模拟低通滤波器变成数字带阻滤波器

#### 1. 由模拟低通到模拟带阻的变换

这一低通到带阻的变换关系为

$$s = \frac{\bar{\Omega}_0^2 p}{p^2 + \bar{\Omega}_0^2} \quad (6-151)$$

式中  $s$  为模拟低通原型拉普拉斯变量 ( $s=\sigma+j\omega$ ),  $p$  为模拟带阻的拉普拉斯变量 ( $p=\bar{\sigma}+j\bar{\Omega}$ ),  $\bar{\Omega}_0$  是模拟带阻滤波器的几何中心频率。令  $p=j\bar{\Omega}$ ,代入(6-151)式,可得

$$s = j \frac{\bar{\Omega}_0^2 \bar{\Omega}}{\bar{\Omega}_0^2 - \bar{\Omega}^2}$$

故  $p$  平面上的虚轴与  $s$  平面上的虚轴相对应,代入  $s=j\bar{\Omega}$ ,消去  $j$  可得

$$\bar{\Omega} = \frac{\bar{\Omega}_0^2 \bar{\Omega}}{\bar{\Omega}_0^2 - \bar{\Omega}^2} \quad (6-152)$$

按(6-152)式画出的  $\bar{\Omega} \geq 0$  时的图形如图 6-29 所示,由(6-152)式及此图可看出对应频率的关系为

$$\begin{aligned}
\bar{\Omega} = 0 &\rightarrow \bar{\Omega} = 0, \bar{\Omega} = \infty \quad \bar{\Omega} = \Omega_c \rightarrow \bar{\Omega} = \bar{\Omega}_0, \\
\bar{\Omega} = \pm \infty &\rightarrow \bar{\Omega} = \bar{\Omega}_0, \quad \bar{\Omega} = -\Omega_c \rightarrow \bar{\Omega} = -\bar{\Omega}_0
\end{aligned}$$

其中  $\Omega_c$  为低通滤波器的通带截止频率,故低通的阻带映射到带阻的阻带,如图 6-30 所示。

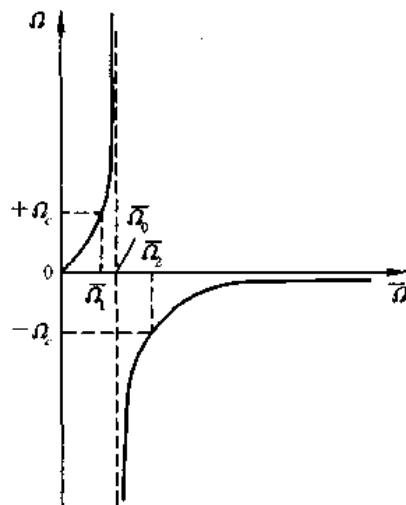


图 6-29 模拟低通滤波器到模拟带阻滤波器频率的变换关系

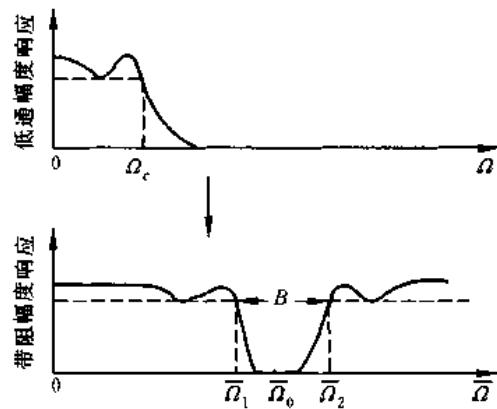


图 6-30 低通幅度响应到带阻幅度响应的变换

将上面的映射结果代入(6-152)式,可得

$$\Omega_c = \frac{\bar{\Omega}_0^2 \bar{\Omega}_1}{\bar{\Omega}_0^2 - \bar{\Omega}_1^2} \quad (6-153)$$

$$-\Omega_c = \frac{\bar{\Omega}_0^2 \bar{\Omega}_2}{\bar{\Omega}_0^2 - \bar{\Omega}_2^2} \quad (6-154)$$

将(6-153)式、(6-154)式相加、相减并化简,可得以下有用的关系式:

$$\bar{\Omega}_0 = \sqrt{\Omega_1 \Omega_2} \quad (6-155)$$

$$B = \bar{\Omega}_2 - \bar{\Omega}_1 = \frac{\bar{\Omega}_0^2}{\Omega_c} = \frac{\bar{\Omega}_1 \bar{\Omega}_2}{\Omega_c} \quad (6-156)$$

$B$  称带阻滤波器的阻带带宽。从(6-156)式看出,  $B$  与低通原型的带宽  $\Omega_c$  成反比, (6-155)式说明  $\bar{\Omega}_0$  是带阻滤波器阻带的几何对称中心角频率。

因此,由模拟低通滤波器来确定模拟带阻滤波器的方法如下:

(1) 定出低通系统函数  $H_{LP}(s)$ , 而低通的截止频率  $\Omega_c$  与带阻滤波器阻带带宽  $B$  成反比;

(2) 在  $H_{LP}(s)$  中代入(6-151)式的变换关系,可得带阻滤波器系统函数

$$H_{BR}(p) = H_{LP}(s) \Big|_{s=\frac{\bar{\Omega}_0^2 p}{p^2 + \bar{\Omega}_0^2}} \quad (6-157)$$

## 2. 由模拟带阻到数字带阻的变换

仍利用双线性变换

$$p = c \frac{1 - z^{-1}}{1 + z^{-1}} \quad (6-158)$$

3. 把变换式(6-151)与变换式(6-158)结合起来,可得到直接从模拟低通原型滤波器

的  $s$  平面变成数字带阻滤波器的  $z$  平面的表达式

$$s = \frac{\bar{\Omega}_0^2 p}{p^2 + \bar{\Omega}_0^2} = \frac{\bar{\Omega}_0^2 c \frac{1 - z^{-1}}{1 + z^{-1}}}{c^2 (1 - z^{-1})^2 + \bar{\Omega}_0^2}$$

经推导后得

$$s = \frac{c \bar{\Omega}_0^2 (1 - z^{-2})}{(c^2 + \bar{\Omega}_0^2) \left[ 1 - 2 \frac{c^2 - \bar{\Omega}_0^2}{c^2 + \bar{\Omega}_0^2} z^{-1} + z^{-2} \right]} \quad (6-159)$$

下面来化简(6-159)式。利用双线性变换的频率间关系

$$\bar{\Omega} = c \tan \frac{\omega}{2} \quad (6-160)$$

$\omega$  是数字带阻滤波器的数字频率。将(6-160)式代入(6-155)式及(6-156)式, 可得

$$\tan^2 \frac{\omega_0}{2} = \tan \frac{\omega_1}{2} \tan \frac{\omega_2}{2} \quad (6-161)$$

$$\tan \frac{\omega_2}{2} - \tan \frac{\omega_1}{2} = c^2 - \frac{\tan^2 \frac{\omega_0}{2}}{\Omega_c} = c^2 \frac{\tan \frac{\omega_1}{2} \tan \frac{\omega_2}{2}}{\Omega_c} \quad (6-162)$$

考虑到模拟带阻滤波器到数字带阻滤波器是阻带中心频率相对应的映射关系, 设数字带阻滤波器的阻带中心频率为  $\omega_0$ , 则有

$$\bar{\Omega}_0 = c \tan \frac{\omega_0}{2} \quad (6-163)$$

因而

$$c = \bar{\Omega}_0 \cot \frac{\omega_0}{2} \quad (6-164)$$

令

$$D_1 = \frac{c \bar{\Omega}_0^2}{c^2 + \bar{\Omega}_0^2}$$

用(6-161)式、(6-162)式、(6-164)式和有关的三角恒等式可得

$$D_1 = \Omega_c \tan \left( \frac{\omega_2 - \omega_1}{2} \right) \quad (6-165)$$

令

$$E_1 = 2 \frac{c^2 - \bar{\Omega}_0^2}{c^2 + \bar{\Omega}_0^2}$$

与带通滤波器一样, 可以导出

$$E_1 = 2 \cos \omega_0 = \frac{\cos \left( \frac{\omega_2 + \omega_1}{2} \right)}{\cos \left( \frac{\omega_2 - \omega_1}{2} \right)} \quad (6-166)$$

将(6-165)式、(6-166)式代入(6-159)式,可得

$$s = \frac{D_1(1-z^{-2})}{1-E_1z^{-1}+z^{-2}} \quad (6-167)$$

因此,从模拟低通系统函数  $H_{LP}(s)$ , 经过(6-167)式的一步变换,就可变成数字带阻系统函数

$$H(z) = H_{LP}(s) \Big|_{s=\frac{D_1(1-z^{-2})}{1-E_1z^{-1}+z^{-2}}} \quad (6-168)$$

可以看出,数字带阻滤波器的极点数(或阶数)等于模拟低通滤波器极点数的两倍。

下面讨论模拟低通滤波器和数字带阻滤波器频率之间的关系。令  $s=j\Omega, z=e^{j\omega}$  代入(6-167)式,经推导后可得

$$\Omega = D_1 \frac{\sin \omega}{\cos \omega - \cos \omega_0} \quad (6-169)$$

其变换关系曲线见图 6-31,其映射关系为

$$\Omega = 0 \rightarrow \omega = 0, \omega = \pi \quad \Omega = \pm \infty \rightarrow \omega = \omega_0$$

也就是说,确实低通滤波器的通带( $\Omega=0$ 附近)映射到带阻滤波器的阻带范围之外( $\omega=0$ 附近及  $\omega=\pi$ 附近),低通滤波器的阻带( $\Omega=\pm\infty$ 附近)映射到带阻滤波器的阻带上( $\omega=\omega_0$ 附近)。

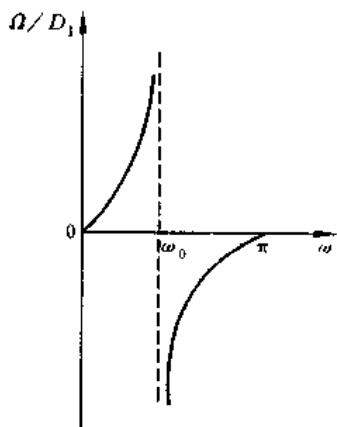


图6-31 从模拟低通变换到数字带阻时,频率间关系的曲线

**[例 6-9]** 要设计一个巴特沃思带阻数字滤波器,其频率特性如例 6-9 图所示,其  $-3\text{dB}$  衰减处的边带频率分别为  $f_1 = 10\text{kHz}$ ,  $f_2 = 35\text{kHz}$ , 其  $-14\text{dB}$  衰减处的频率分别为  $f_{u_1} = 18\text{kHz}$ ,  $f_{u_2} = 25\text{kHz}$ , 抽样频率为  $f_s = 100\text{kHz}$ 。

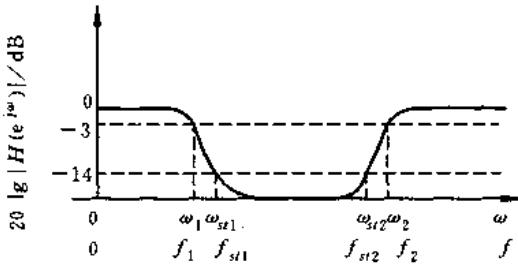
解 (1) 对应的各数字频率

$$\omega_1 = 2\pi \frac{f_1}{f_s} = \frac{2\pi \times 10 \times 10^3}{100 \times 10^3} = 0.2\pi$$

$$\omega_2 = 2\pi \frac{f_2}{f_s} = \frac{2\pi \times 35 \times 10^3}{100 \times 10^3} = 0.7\pi$$

$$\omega_{n_1} = 2\pi \frac{f_{n_1}}{f_s} = \frac{2\pi \times 18 \times 10^3}{100 \times 10^3} = 0.36\pi$$

$$\omega_{n_2} = 2\pi \frac{f_{n_2}}{f_s} = \frac{2\pi \times 25 \times 10^3}{100 \times 10^3} = 0.5\pi$$



例 6-9 所需带阻滤波器幅频特性

(2) 求  $D_1, E_1$ 。采用归一化原型  $\Omega_c=1$  的低通滤波器作为变换原型，则低通到带阻的变换中所需常数  $D_1, E_1$  分别为

$$D_1 = \Omega_c \tan\left(\frac{\omega_2 - \omega_1}{2}\right) = 1 \times \tan(0.25\pi) = 1$$

$$E_1 = 2\cos \omega_0 = \frac{2\cos\left(\frac{\omega_2 + \omega_1}{2}\right)}{\cos\left(\frac{\omega_2 - \omega_1}{2}\right)} = \frac{2\cos(0.45\pi)}{\cos(0.25\pi)} = 0.4424635$$

(3) 求  $\Omega_n$ 。它是满足数字带阻滤波器要求的归一化原型模拟低通滤波器的阻带起始频率。可按(6-169)式来求。对应于  $\omega_{n_1}$  有

$$\begin{aligned} \Omega_{n_1} &= D_1 \frac{\sin \omega_{n_1}}{\cos \omega_{n_1} - \cos \omega_0} = \frac{\sin(0.36\pi)}{\cos(0.36\pi) - \frac{\sin(0.45\pi)}{\cos(0.25\pi)}} \\ &= \frac{0.904827}{0.4257792 - 0.2212318} = 4.4235566 \end{aligned}$$

对应于  $\omega_{n_2}$  有

$$\Omega_{n_2} = D_1 \frac{\sin \omega_{n_2}}{\cos \omega_{n_2} - \cos \omega_0} = \frac{-1}{0.2212318} = -4.5201458$$

应该取  $\Omega_n = \min(|\Omega_{n_1}|, |\Omega_{n_2}|) = 4.4235566$ ，这是因为取较小的频率值，如果此处满足衰减要求，则较大的频率处一定满足衰减要求。

(4) 求  $N$ 。它是归一化原型模拟低通滤波器的阶数，由巴特沃思低通滤波器频率响应的公式(6-73)式代入  $\Omega=\Omega_n$ ，且将该式变成分贝关系，即

$$20 \log |H_a(j\Omega_n)| = -10 \lg_{10} \left[ 1 + \left( \frac{\Omega_n}{\Omega_c} \right)^{2N} \right] \leq -14$$

代入  $\Omega_c=1$ ，得

$$N \geq \frac{\log(10^{1.4} - 1)}{2 \lg(4.4235566)} = \frac{1.3823569}{2 \times 0.6457715} = 1.0703141$$

取  $N=2$ 。

(5) 求  $H_a(s)$ 。查表 6-6,  $N=2$ , 可得归一化( $\Omega_c=1$ )的  $H_a(s)$  为

$$H_{LP}(s) = \frac{1}{(s + 0.7071068 + j0.7071068)(s + 0.7071068 - j0.7071068)} \\ = \frac{1}{s^2 + 1.4142136s + 1}$$

(6) 求  $H(z)$ 。

$$H(z) = H_{LP}(s) \Big|_{s=\frac{\Omega_c(1-z^{-2})}{1-\Omega_c z^{-1}+z^{-2}}} \\ = \frac{0.2942598 - 0.2603984z^{-1} + 0.6461279z^{-2} - 0.2603984z^{-3} + 0.2942598z^{-4}}{1 - 0.4436717z^{-1} + 0.06306938z^{-2} - 0.07712521z^{-3} + 0.1715782z^{-4}}$$

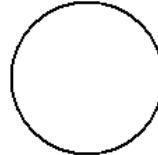
利用本书所附光盘的“辅助设计”子系统, 所得设计结果附在本章最后面。

#### 四、模拟低通滤波器变成数字高通滤波器

##### 1. 由模拟低通到模拟高通的变换

变换关系为

$$s = \frac{\Omega_c \bar{\Omega}_c}{p} \quad (6-170)$$



其中  $s$  为模拟低通原型拉普拉斯变量,  $\Omega_c$  为此低通原型的通带截止频率,  $p$  为模拟高通拉普拉斯变量,  $\bar{\Omega}_c$  为与  $\Omega_c$  相对应的高通滤波器的截止频率。

在(6-170)式中, 令  $p=j\bar{\Omega}_c$ , 则得

$$s = -j \frac{\Omega_c \bar{\Omega}_c}{\bar{\Omega}_c}$$

故  $p$  平面上的虚轴与  $s$  平面上的虚轴相对应, 代入  $s=j\bar{\Omega}_c$ , 消去  $j$  可得

$$\Omega = -\frac{\Omega_c \bar{\Omega}_c}{\bar{\Omega}_c}, \quad \text{或} \quad \frac{\Omega}{\Omega_c} = -\frac{\bar{\Omega}_c}{\bar{\Omega}} \quad (6-171)$$

按(6-171)式画出曲线如图 6-32 所示。

由(6-171)式及图 6-32 看出对应频率间关系为

$$\Omega = 0 \rightarrow \bar{\Omega} = \infty, \quad \Omega = \infty \rightarrow \bar{\Omega} = 0, \quad \Omega = \Omega_c \rightarrow \bar{\Omega} = -\bar{\Omega}_c$$

也就是说, 低通原型的通带, 以相反的关系平移到从  $\bar{\Omega}=\bar{\Omega}_c$  到  $\bar{\Omega}=\infty$  的高通的通带(幅度响应关系可见图 6-33 所示)。

若取  $\Omega_c = \bar{\Omega}_c$ , 则低通的通带宽度就等于高通的阻带宽度。

由模拟低通滤波器系统函数确定模拟高通滤波器系统函数的方法如下:

- (1) 确定低通系统函数  $H_{LP}(s)$ , 其参考角频率  $\Omega_c$  (一般为截止频率) 由高通参考角频率(一般为截止频率)  $\bar{\Omega}_c$  选定, 从下面例子可以看出, 为了方便, 一般都选  $\Omega_c=1$  的归一化原型低通滤波器;

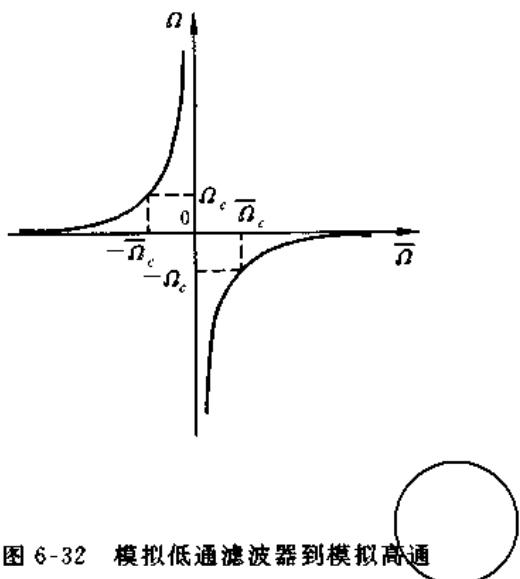


图 6-32 模拟低通滤波器到模拟高通滤波器频率的变换关系

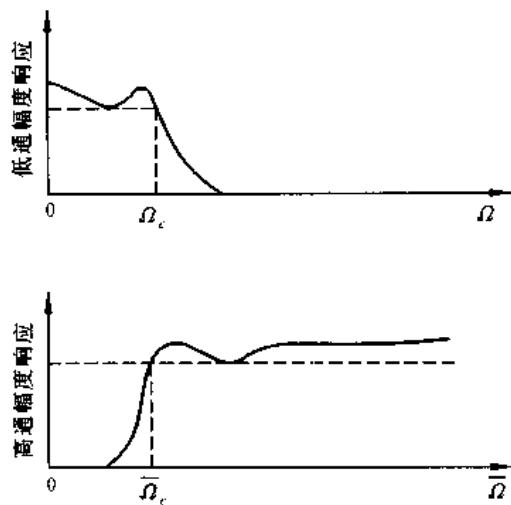


图 6-33 低通幅度响应到高通幅度响应的变换

(2) 在所得到的  $H_{LP}(s)$  中代入变换关系(6-170)式, 得到高通系统函数

$$H_{HP}(p) = H_{LP}(s) \Big|_{s=\frac{\Omega_c \bar{\Omega}_c}{p}} \quad (6-172)$$

## 2. 由模拟高通到数字高通的变换

仍利用双线性变换

$$p = c \frac{1 - z^{-1}}{z + z^{-1}} \quad (6-173)$$

3. 把变换式(6-170)和变换式(6-173)结合起来, 可得到直接从模拟低通原型变成数字高通滤波器的表达式, 也就是直接联系  $s$  与  $z$  之间的变换公式

$$s = \frac{\Omega_c \bar{\Omega}_c}{p} = \frac{\Omega_c \bar{\Omega}_c}{c \left( \frac{1 - z^{-1}}{1 + z^{-1}} \right)} = \frac{\Omega_c \bar{\Omega}_c}{c} \cdot \frac{1 + z^{-1}}{1 - z^{-1}} = C_1 \frac{1 + z^{-1}}{1 - z^{-1}} \quad (6-174)$$

由此得到数字高通系统函数为

$$H(z) = H_{LP}(s) \Big|_{s=C_1 \frac{1+z^{-1}}{1-z^{-1}}} \quad (6-175)$$

可以看出, 数字高通滤波器和模拟低通滤波器的极点数目(或阶次)是相同的。

根据双线性变换, 模拟高通频率与数字高通频率之间的关系为

$$\bar{\Omega} = c \tan \frac{\omega_c}{2}$$

则

$$\bar{\Omega}_c = c \tan \frac{\omega_c}{2}$$

又因  $C_1 = \frac{\Omega_c \bar{\Omega}_c}{c}$ , 故

$$C_1 = \Omega_c \tan \frac{\omega_c}{2} \quad (6-176)$$

下面讨论模拟低通滤波器和数字高通滤波器频率之间的关系。令  $s=j\Omega$ ,  $z=e^{j\omega}$ , 代入(6-174)式, 可得

$$\Omega = -C_1 \cot \frac{\omega}{2}$$

或

$$|\Omega| = C_1 \cot \frac{\omega}{2} \quad (6-177)$$

其变换关系曲线如图 6-34 所示。

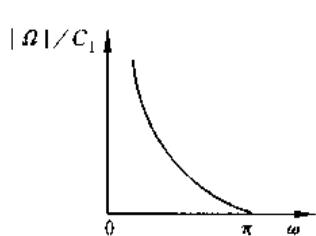
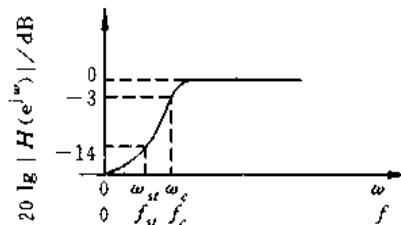


图 6-34 从模拟低通变换到数字高通时  
频率间关系的曲线



例 6-10 所需高通滤波器幅频特性

**[例 6-10]** 要设计一个巴特沃思高通滤波器, 其通带截止频率(-3dB 点处)为  $f_c = 3\text{kHz}$ , 阻带上限截止频率  $f_u = 2\text{kHz}$ , 通带衰减不大于 3dB, 阻带衰减不小于 14dB, 抽样频率  $f_s = 10\text{kHz}$ 。其幅频特性如例 6-10 图所示。

解 (1) 求对应的各数字频率

$$\omega_t = 2\pi \frac{f_c}{f_s} = \frac{2\pi \times 3 \times 10^3}{10 \times 10^3} = 0.6\pi$$

$$\omega_u = 2\pi \frac{f_u}{f_s} = \frac{2\pi \times 3 \times 10^3}{10 \times 10^3} = 0.4\pi$$

(2) 求常数  $C_1$ 。采用归一化( $\Omega_c = 1$ )原型低通滤波器作为变换的低通原型, 则低通到高通的变换中所需的  $C_1$  为(见表 6-8)

$$C_1 = \Omega_c \tan \frac{\omega_c}{2} = 1 \times \tan \left( \frac{0.6\pi}{2} \right) = 1.37638192$$

(3) 求低通原型  $\Omega_u$ 。设  $\Omega_u$  为满足数字高通滤波器的归一化原型模拟低通滤波器的阻带上限截止频率, 可按  $\Omega = C_1 \cot \frac{\omega}{2}$  的预畸变换关系来求, 得

$$\Omega_u = C_1 \cot \frac{\omega_u}{2} = 1.3763819 \times 1.3763819 = 1.8944272$$

(4) 求阶次  $N$ 。按阻带衰减求原型归一化模拟低通滤波器的阶次  $N$ , 由巴特沃思低通滤波器频率响应的公式(6-74)式取对数, 即

$$20 \log |H_u(j\Omega_u)| = -10 \log_{10} \left[ 1 + \left( \frac{\Omega_u}{\Omega_c} \right)^{2N} \right] \leq -14$$

其中  $\Omega_c = 1$ , 解得

$$N = \frac{\log(10^{1.4} - 1)}{2 \log(1.8944272)} = \frac{1.3823569}{0.5549558} = 2.4909314$$

取  $N=3$ 。

(5) 求归一化巴特沃思低通原型的  $H_o(s)$ 。取  $N=3$ , 查表 6-4 可得  $H_{LP}(s)$  为

$$H_{LP}(s) = \frac{1}{s^3 + 2s^2 + 2s + 1}$$

(6) 求数字高通滤波器的系统函数  $H(z)$ , 有

$$\begin{aligned} H(z) &= H_{LP}(s) \Big|_{s=C_1 \frac{1+z^{-1}}{1-z^{-1}}} \\ &= \frac{(1-z^{-1})^3}{C_1^3 (1+z^{-1})^3 + 2C_1^2 (1+z^{-1})^2 (1-z^{-1}) + 2C_1 (1+z^{-1})(1-z^{-1})^2 + (1-z^{-1})^3} \\ &= \frac{1-3z^{-1}+3z^{-2}-z^{-3}}{(C_1^3 + 2C_1^2 + 2C_1 + 1) + (3C_1^3 + 2C_1^2 - 2C_1 - 3)z^{-1} + (3C_1^3 - 2C_1^2 - 2C_1 + 3)z^{-2} + (C_1^3 - 2C_1^2 + 2C_1 - 1)z^{-3}} \\ &= \frac{1}{C_1^3 + 2C_1^2 + 2C_1 + 1} (1-3z^{-1}+3z^{-2}-z^{-3}) \\ &= \frac{1}{1 + \frac{3C_1^3 + 2C_1^2 - 2C_1 - 3}{C_1^3 + 2C_1^2 + 2C_1 + 1} z^{-1} + \frac{3C_1^3 - 2C_1^2 - 2C_1 + 3}{C_1^3 + 2C_1^2 + 2C_1 + 1} z^{-2} + \frac{C_1^3 - 2C_1^2 + 2C_1 - 1}{C_1^3 + 2C_1^2 + 2C_1 + 1} z^{-3}} \end{aligned}$$

将  $C_1$  代入, 可求得

$$H(Z) = \frac{0.09907984(1-3z^{-1}+3z^{-2}-z^{-3})}{1 + 0.5717848z^{-1} + 0.4201167z^{-2} + 0.05569325z^{-3}}$$

利用本书所附光盘的“辅助设计”子系统, 所得设计结果附在本章最后面。

我们把各种变换的必要设计公式归纳到表 6-8 上。这些变换都是代数式, 因而对任何形式表示(级联、并联、直接等)的模拟滤波器的变换都适用。

表 6-8 根据模拟低通原型设计各类数字滤波器的频率变换式及有关设计参数的表达式

数字滤波器类型	频率变换式	设计参数的表达式
高通	$s = C_1 \frac{1+z^{-1}}{1-z^{-1}}$ $\Omega = C_1 \cot \frac{\omega}{2}$	$C_1 = \Omega_c \tan \frac{\omega_c}{2}$
带通	$s = D \left[ \frac{1-Ez^{-1}+z^{-2}}{1-z^{-2}} \right]$ $\Omega = D \frac{\cos \omega_0 - \cos \omega}{\sin \omega}$	$D = \Omega_c \cot \left( \frac{\omega_2 - \omega_1}{2} \right)$ $E = \frac{2 \cos \left( \frac{\omega_2 + \omega_1}{2} \right)}{\cos \left( \frac{\omega_2 - \omega_1}{2} \right)} = 2 \cos \omega_0$
带阻	$s = D_1 \frac{1-z^{-2}}{1-E_1 z^{-1}+z^{-2}}$ $\Omega = D_1 \frac{\sin \omega}{\cos \omega - \cos \omega_0}$	$D_1 = \Omega_c \tan \left( \frac{\omega_2 - \omega_1}{2} \right)$ $E_1 = \frac{2 \cos \left( \frac{\omega_2 + \omega_1}{2} \right)}{\cos \left( \frac{\omega_2 - \omega_1}{2} \right)} = 2 \cos \omega_0$

## 6.11 先将模拟归一化低通原型数字化为数字低通,再利用数字域频带变换法设计数字各型滤波器

这里的第一步即将模拟归一化低通原型滤波器数字化为数字低通滤波器的办法前面已经讨论过了。本节讨论第二步,即直接由给定的数字低通滤波器转换成各种类型的数字滤波器的方法——数字域频带变换。

如果我们给定数字滤波器的低通原型系统函数  $H_L(z)$ ,同样可以通过一定的变换,来设计其他各种不同类型的数字滤波器系统函数  $H_d(Z)$ ,这种变换是将  $H_L(z)$  的  $z$  平面映射变换到  $H_d(Z)$  的  $Z$  平面,这一从  $z$  到  $Z$  的映射关系为

$$z^{-1} = G(Z^{-1}) \quad (6-178)$$

则有

$$H_d(Z) = H_L(z) \Big|_{z^{-1}=G(Z^{-1})} \quad (6-179)$$

下面我们来讨论对变换函数  $G(Z^{-1})$  的要求。总的要求是使一个因果稳定的数字低通有理系统函数  $H_L(z)$  变换成的  $H_d(Z)$  也必须是因果稳定的有理系统函数,为此

(1) 频率响应要满足一定的变换要求,因而频率轴应能对应起来,也就是说  $z$  平面的单位圆必须映射到  $Z$  平面的单位圆上;

(2) 从因果稳定性角度看,  $z$  平面的单位圆内部必须映射到  $Z$  平面的单位圆内部;

(3) 系统函数  $G(Z^{-1})$  必须是  $Z^{-1}$  的有理函数。

设  $\theta$  和  $\omega$  分别为  $z$  平面与  $Z$  平面的数字频率变量,即  $z = e^{j\theta}, Z = e^{j\omega}$ ,根据(6-178)式可得

$$e^{-j\theta} = G(e^{-j\omega}) = |G(e^{-j\omega})| e^{j\arg[G(e^{-j\omega})]} \quad (6-180)$$

这就要求

$$|G(e^{-j\omega})| = 1 \quad (6-181a)$$

$$\theta = -\arg[G(e^{-j\omega})] \quad (6-181b)$$

(6-181a)式表明函数  $G(Z^{-1})$  在其  $Z$  平面的单位圆上的幅度必须恒等于 1,这样的函数就是全通函数。将全通函数的表示式代入(6-178)式,可得

$$z^{-1} = G(Z^{-1}) = \pm \prod_{i=1}^N \frac{Z^{-1} - a_i}{1 - a_i Z^{-1}} \quad (6-182)$$

其中  $a_i$  是  $G(Z^{-1})$  的极点,可以是实数,也可以是共轭复数,但必须保证极点在单位圆内,即  $|a_i| < 1$ ,以保证变换的稳定性不改变,读者可以证明  $|z| < 1$  时,一定满足  $|Z| < 1$ 。

$G(Z^{-1})$  的所有零点都是其极点的共轭倒数  $\left(\frac{1}{a_i^*}\right)$ ,  $N$  称为全通函数的阶数,可以证明,当

$\omega$  由 0 变到  $\pi$  时,全通函数的相角  $\arg[G(e^{-j\omega})]$  的变化量为  $N\pi$ 。选择合适的  $N$  和  $a_i$ ,就可得到各类变换。

### 一、数字低通—数字低通

此时,  $H_L(e^{j\theta})$  和  $H_d(e^{j\omega})$  都是低通系统函数,只不过截止频率不相同,因而  $\theta$  从 0 变

到  $\pi$  时, 相应的  $\omega$  也应从 0 变到  $\pi$ , 按全通函数相角变化量为  $N\pi$ , 故可确定其阶数应为  $N=1$ , 故变换函数为一阶全通函数

$$z^{-1} = G(Z^{-1}) = \frac{Z^{-1} - \alpha}{1 - \alpha Z^{-1}} \quad (6-183)$$

其中  $\alpha$  为实数, 此式满足  $G(1)=1, G(-1)=-1$  的变换关系, 且需  $|\alpha| < 1$ 。变换关系式 (6-183) 中的变换边界条件为

$$\begin{aligned} z = 1 &\rightarrow Z = 1 \\ Z = e^{j\theta} &\rightarrow Z = e^{j\theta} \end{aligned} \quad (6-184)$$

由此得出数字低通—数字低通的频率变换关系以及与之对应的幅度响应关系如图 6-35 所示。

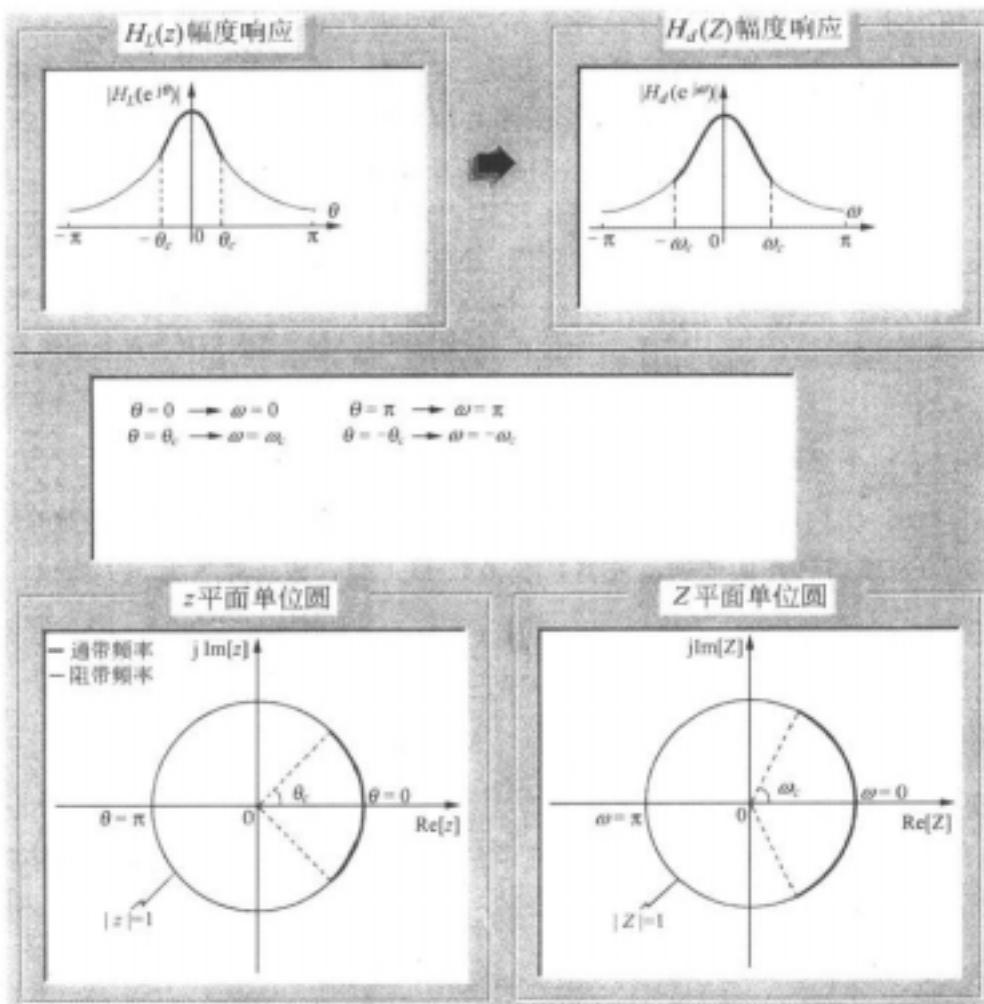


图 6-35 数字低通—数字低通的变换

概括起来,  $z$  域单位圆和  $Z$  域单位圆上各点的对应关系可表示为

$$\begin{aligned} z: \quad 1 &\rightarrow e^{j\theta_c} \rightarrow -1 \rightarrow e^{-j\theta_c} \rightarrow 1 \\ Z: \quad 1 &\rightarrow e^{j\omega_c} \rightarrow -1 \rightarrow e^{-j\omega_c} \rightarrow 1 \end{aligned}$$

在(6-183)式中,代入  $z=e^{j\theta}$ ,  $Z=e^{j\omega}$ , 则得频率间关系为

$$e^{-j\theta} = \frac{e^{-j\omega} - \alpha}{1 - \alpha e^{-j\omega}} \quad (6-185)$$

也可表示为

$$e^{-j\omega} = \frac{e^{-j\theta} + \alpha}{1 + \alpha e^{-j\theta}} = e^{-j\theta} \frac{1 + \alpha e^{j\theta}}{1 + \alpha e^{-j\theta}} = e^{-j\theta} \frac{1 + \alpha \cos \theta + j\alpha \sin \theta}{1 + \alpha \cos \theta - j\alpha \sin \theta} \quad (6-186)$$

由此求得

$$\omega = \arctan \left[ \frac{(1 - \alpha^2) \sin \theta}{2\alpha + (1 + \alpha^2) \cos \theta} \right] = \theta - 2 \arctan \left[ \frac{\alpha \sin \theta}{1 + \alpha \cos \theta} \right] \quad (6-187)$$

$\theta$  与  $\omega$  的关系如图 6-36 所示,除  $\alpha=0$  外(此时  $\omega=\theta$ ),在其他  $\alpha$  情况下,频率变换都是非线性关系。 $\alpha>0$  时表示频率压缩, $\alpha<0$  时表示频率扩张。但是,对于幅度响应为分段常数的滤波器,变换后仍可得类似的频率响应。设低通原型的截止频率为  $\theta_c$ , 变换后的对应截止频率为  $\omega_c$ , 代入(6-185)式,可以解得所需  $\alpha$  值为

$$\alpha = \frac{\sin \left( \frac{\theta_c - \omega_c}{2} \right)}{\sin \left( \frac{\theta_c + \omega_c}{2} \right)} \quad (6-188)$$

这样就确定了整个变换函数。

因此,由(6-183)式、(6-188)式,就可由已有的数字低通滤波器系统函数  $H_L(z)$ (截止频率为  $\theta_c$ )变换得到新的低通滤波器系统函数  $H_d(Z)$ , (截止频率为  $\omega_c$ ), 即

$$H_d(Z) = H_L(z) \Big|_{z^{-1} = \frac{Z^{-1} - \alpha}{1 - \alpha Z^{-1}}} \quad (6-189)$$

## 二、数字低通—数字高通

低通变换为高通只需将低通频率响应在单位圆上旋转  $180^\circ$ , 也就是将  $Z$  变换成  $-Z$  即可, 这就是旋转变换。只需将(6-183)式中的  $Z^{-1}$  用  $-Z^{-1}$  代替, 就完成了数字低通到数字高通( $Z$ )的变换, 即

$$z^{-1} = \frac{-Z^{-1} - \alpha}{1 - \alpha Z^{-1}} = - \left( \frac{Z^{-1} + \alpha}{1 + \alpha Z^{-1}} \right) \quad (6-190)$$

其中  $\alpha$  为实数, 此式应满足  $G(-1)=1$ ,  $G(1)=-1$  的关系, 且有  $|\alpha|<1$ 。显示这一变换对应的边界条件为

$$z = 1 \rightarrow Z = -1 \quad (6-191a)$$

$$z = e^{j\theta} \rightarrow Z = e^{-j\theta} \quad (6-191b)$$

由此得出低通—高通的频率变换关系以及与之对应的幅度响应关系, 如图 6-37 所示。

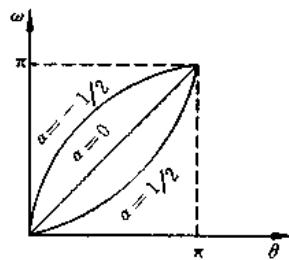


图 6-36 数字低通—数字低通  
变换的频率间非线性关系

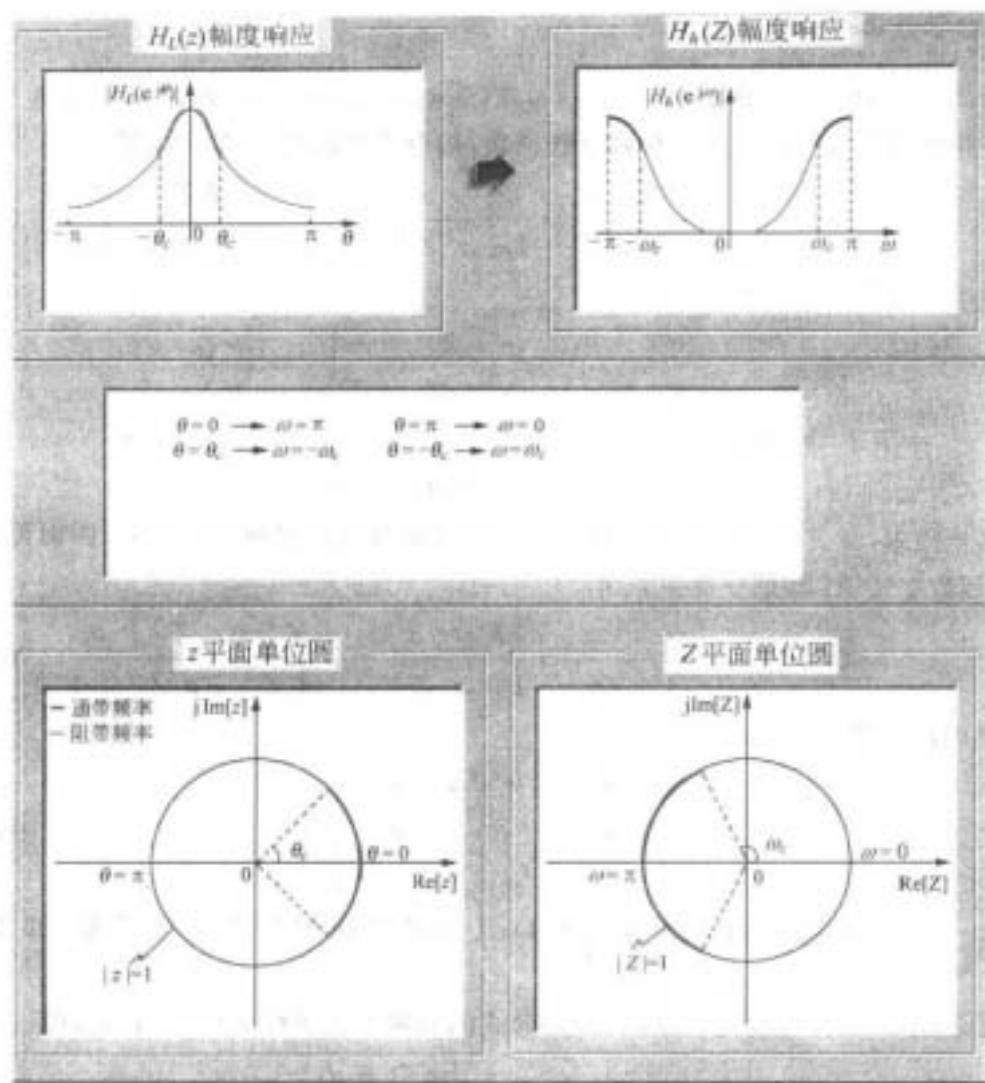


图 6-37 数字低通—数字高通的变换

为了求得  $\alpha$ , 可将(6-191b)代入(6-190)式, 可得

$$e^{-j\theta_c} = -\frac{e^{j\theta_c} + \alpha}{1 + \alpha e^{j\theta_c}}$$

由此可解出

$$\alpha = -\frac{\cos\left(\frac{\theta_c + \omega_c}{2}\right)}{\cos\left(\frac{\theta_c - \omega_c}{2}\right)} \quad (6-192)$$

概括起来, 数字低通—数字高通变换时,  $z$  域和  $Z$  域单位圆上各点的对应关系表示为

$$\begin{aligned} z: & \quad 1 \rightarrow e^{j\theta_c} \rightarrow -1 \rightarrow e^{-j\theta_c} \rightarrow 1 \\ Z: & -1 \rightarrow e^{-j\omega_c} \rightarrow 1 \rightarrow e^{j\omega_c} \rightarrow -1 \end{aligned}$$

### 三、数字低通—数字带通

对于低通到带通的变换,由于变换后 $\omega$ 在 $(-\pi, \pi)$ 之间或 $(0, 2\pi)$ 之间形成两个通带,也就是变换函数需要 $\theta$ 在本身单位圆上转两次,故令函数的阶次 $N=2$ ,则有

$$z^{-1} = G(Z^{-1}) = \pm \frac{Z^{-1} - a^*}{1 - aZ^{-1}} \cdot \frac{Z^{-1} - a}{1 - a^* Z^{-1}} = \pm \frac{Z^{-2} + d_1 Z^{-1} + d_2}{d_2 Z^{-2} + d_1 Z^{-1} + 1} \quad (6-193)$$

由于变换应满足下面的条件:

$$\begin{aligned} z^{-1} = -1 &\rightarrow Z^{-1} = -1 \\ z^{-1} = -1 &\rightarrow Z^{-1} = 1 \end{aligned}$$

因而,按 $z^{-1}=G(Z^{-1})$ 可知

$$G(-1) = -1, G(1) = -1$$

考虑到这个约束,加之 $d_1, d_2$ 都是实数,(6-193)式中分式前必须取“-”号,因而可得数字低通( $z$ )到数字带通( $Z$ )的变换关系为

$$z^{-1} = G(Z^{-1}) = - \frac{Z^{-1} - a^*}{1 - aZ^{-1}} \cdot \frac{Z^{-1} - a}{1 - a^* Z^{-1}} = - \frac{Z^{-2} + d_1 Z^{-1} + d_2}{d_2 Z^{-2} + d_1 Z^{-1} + 1} \quad (6-194)$$

按照变换的以下边界条件:

$$\left. \begin{aligned} z = e^{-j\theta_c} &\rightarrow Z = e^{j\omega_1} \\ z = e^{j\theta_c} &\rightarrow Z = e^{j\omega_2} \\ z = 1 &\rightarrow Z = e^{\pm j\omega_0} \end{aligned} \right\} \quad (6-195)$$

可以得出低通—带通的频率变换关系以及与之对应的幅度响应关系,如图 6-38 所示。

为了求出 $d_1, d_2$ ,可将边界条件(6-195)式的前两个式子代入(6-194)式,得

$$e^{j\theta_c} = - \frac{e^{-j2\omega_1} + d_1 e^{-j\omega_1} + d_2}{d_2 e^{-j2\omega_1} + d_1 e^{-j\omega_1} + 1} \quad (6-196)$$

$$e^{-j\theta_c} = - \frac{e^{-j2\omega_2} + d_1 e^{-j\omega_2} + d_2}{d_2 e^{-j2\omega_2} + d_1 e^{-j\omega_2} + 1} \quad (6-197)$$

由(6-196)式可得

$$d_2 e^{j(\theta_c - 2\omega_1)} + d_1 e^{j(\theta_c - \omega_1)} + e^{j\theta_c} + e^{-j2\omega_1} + d_1 e^{-j\omega_1} + d_2 = 0$$

$$\text{即 } e^{j\theta_c} + e^{-j2\omega_1} + d_1 (e^{j(\theta_c - \omega_1)} + e^{-j\omega_1}) + d_2 (1 + e^{j(\theta_c - 2\omega_1)}) = 0$$

或

$$e^{j(\frac{\theta_c}{2} - \omega_1)} \left\{ \left[ e^{j(\frac{\theta_c}{2} + \omega_1)} + e^{-j(\frac{\theta_c}{2} + \omega_1)} \right] + d_1 \left[ e^{j\frac{\theta_c}{2}} + e^{-j\frac{\theta_c}{2}} \right] + d_2 \left[ e^{j(\frac{\theta_c}{2} - \omega_1)} + e^{-j(\frac{\theta_c}{2} - \omega_1)} \right] \right\} = 0$$

$$\text{由此得出 } \cos\left(\frac{\theta_c}{2} + \omega_1\right) + d_1 \cos\left(\frac{\theta_c}{2}\right) + d_2 \cos\left(\frac{\theta_c}{2} - \omega_1\right) = 0 \quad (6-198)$$

同样由(6-197)式可得

$$\cos\left(\frac{\theta_c}{2} - \omega_2\right) + d_1 \cos\left(\frac{\theta_c}{2}\right) + d_2 \cos\left(\frac{\theta_c}{2} + \omega_2\right) = 0 \quad (6-199)$$

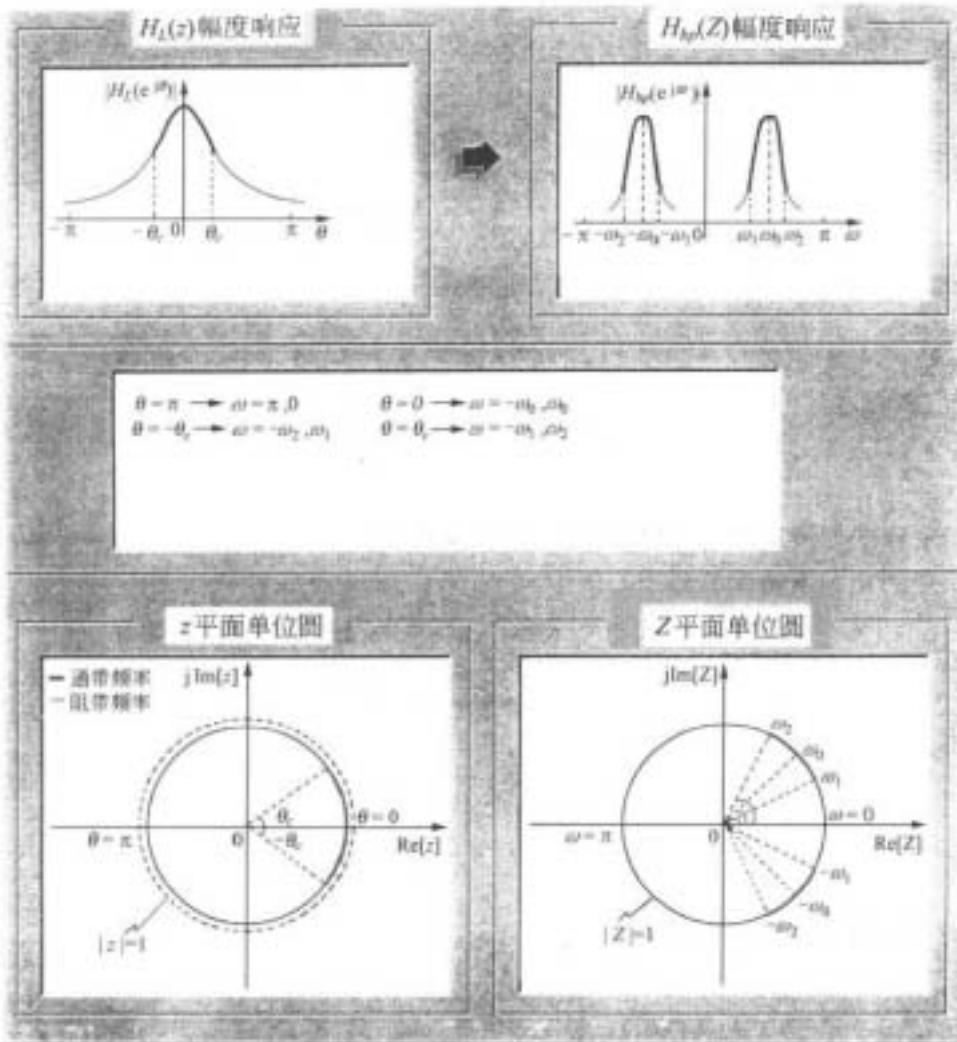


图 6-38 数字低通—数字带通的变换

由(6-198)式及(6-199)式可解出  $d_1, d_2$  的表达式:

(1) 求  $d_2$ 。由(6-199)式减去(6-198)式, 再利用三角恒等式, 可求得

$$d_2 = \frac{1 + \cot\left(\frac{\theta_c}{2}\right)\tan\left(\frac{\omega_2 - \omega_1}{2}\right)}{1 + \cot\left(\frac{\theta_c}{2}\right)\tan\left(\frac{\omega_2 - \omega_1}{2}\right)} = \frac{1 - \cot\left(\frac{\theta_c}{2}\right)\tan\left(\frac{\omega_2 - \omega_1}{2}\right)}{1 + \cot\left(\frac{\theta_c}{2}\right)\tan\left(\frac{\omega_2 - \omega_1}{2}\right)} \quad (6-200)$$

(2) 求  $d_1$ 。将(6-199)式乘以  $\cos\left(\frac{\theta_c}{2} - \omega_2\right)$ , 将(6-198)式乘以  $\cos\left(\frac{\theta_c}{2} + \omega_1\right)$ , 然后将二者相减, 再利用三角恒等式, 可得

$$d_1 = \frac{-2\cos\left(\frac{\omega_2 + \omega_1}{2}\right)/\cos\left(\frac{\omega_2 - \omega_1}{2}\right)}{1 + \cot\left(\frac{\theta_c}{2}\right)\tan\left(\frac{\omega_2 - \omega_1}{2}\right)} \quad (6-201)$$

在(6-200)式及(6-201)式中,令

$$\alpha = \frac{\cos\left(\frac{\omega_2 + \omega_1}{2}\right)}{\cos\left(\frac{\omega_2 - \omega_1}{2}\right)} \quad (6-202)$$

$$k = \tan\left(\frac{\theta_c}{2}\right) \cot\left(\frac{\omega_2 - \omega_1}{2}\right)$$

将(6-202)式代入(6-200)式、(6-201)式,可得

$$d_2 = \left(1 - \frac{1}{k}\right) / \left(1 + \frac{1}{k}\right) = \frac{k-1}{k+1} \quad (6-203)$$

$$d_1 = \frac{-2\alpha}{1 + \frac{1}{k}} = \frac{-2\alpha k}{k+1} \quad (6-204)$$

其中  $\omega_2, \omega_1$  分别为带通滤波器要求的上、下截止频率。

为了得到  $\alpha$  和带通滤波器中心频率  $\omega_0$  的关系,将以下条件代入(6-194)式:

$$z^{-1} = e^{-j\theta_c} = 1 \rightarrow Z^{-1} = e^{-j\omega_0}$$

则有

$$1 = -\frac{e^{-j2\omega_0} + d_1 e^{-j\omega_0} + d_2}{d_2 e^{-j2\omega_0} + d_1 e^{-j\omega_0} + 1}$$

化简可得

$$\cos\omega_0 = \frac{-d_1}{1+d_2} = \frac{\frac{2\alpha k}{k+1}}{1+\frac{k-1}{k+1}} = \alpha \quad (6-205)$$

由(6-202)式和(6-205)式可知

$$\alpha = \frac{\cos\left(\frac{\omega_2 + \omega_1}{2}\right)}{\cos\left(\frac{\omega_2 - \omega_1}{2}\right)} = \cos\omega_0$$

概括起来,数字低通—数字带通变换时,  $z$  域和  $Z$  域单位圆上各点的对应关系表示为

$$z: -1 \rightarrow e^{-j\theta_c} \rightarrow 1 \rightarrow e^{j\theta_c} \rightarrow -1 \rightarrow e^{-j\theta_c} \rightarrow 1 \rightarrow e^{j\theta_c} \rightarrow -1$$

$$Z: -1 \rightarrow e^{-j\omega_0} \rightarrow e^{-j\omega_1} \rightarrow 1 \rightarrow e^{j\omega_1} \rightarrow e^{j\omega_0} \rightarrow e^{-j\omega_0} \rightarrow -1$$

#### 四、数字低通—数字带阻

低通到带阻的变换,由于变换后  $\omega$  在  $(-\pi, \pi)$  之间或  $(0, 2\pi)$  之间形成两个阻带,也就是变换函数需要  $\theta$  在本身单位圆上转两次,故全通函数的阶次  $N=2$ ,则仍为(6-193)式的表达式,但需考虑全通函数前的正负符号的选择,为此,进一步考虑到边界条件为

$$z^{-1} = 1 \rightarrow Z^{-1} = 1$$

$$z^{-1} = 1 \rightarrow Z^{-1} = -1$$

因而,按  $z^{-1} = G(Z^{-1})$  可知

$$G(1) = 1, \quad G(-1) = 1$$

考虑到这个约束,并且  $d_1, d_2$  都是实数,(6-193)式分式前面必须取“+”号,所以可得数字

低通( $z$ )—数字带阻( $Z$ )的变换关系为

$$z^{-1} = G(Z^{-1}) = \frac{Z^{-1} - a}{1 - a^* Z^{-1}} \cdot \frac{Z^{-1} - a^*}{1 - aZ^{-1}} = \frac{Z^{-2} + d_1 Z^{-1} + d_2}{d_2 Z^{-2} + d_1 Z^{-1} + 1} \quad (6-206)$$

按照变换的以下边界条件：

$$\begin{aligned} z = e^{-j\theta} &\rightarrow Z = e^{j\theta} \\ z = e^{j\theta} &\rightarrow Z = e^{-j\theta} \\ z = 1 &\rightarrow Z = \pm 1 \end{aligned} \quad (6-207)$$

可以得出低通—带阻的频率变换关系以及与之对应的幅度响应关系，如图 6-39 所示。

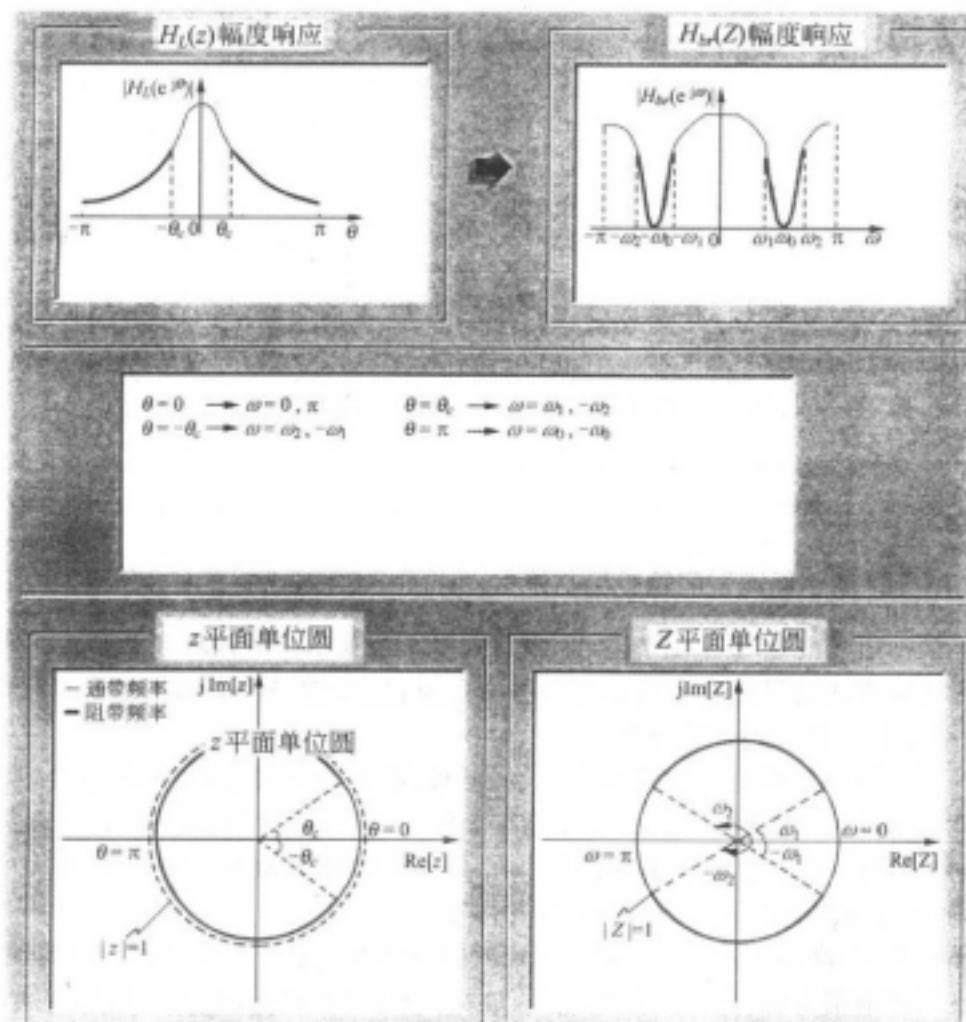


图 6-39 数字低通—数字带阻的变换

用与导出数字通带滤波器的相同方法，可解得

$$d_1 = \frac{-2\cos\left(\frac{\omega_2 + \omega_1}{2}\right) / \cos\left(\frac{\omega_2 - \omega_1}{2}\right)}{1 + \tan\left(\frac{\theta_c}{2}\right)\tan\left(\frac{\omega_2 - \omega_1}{2}\right)} \quad (6-208)$$

$$d_2 = \frac{1 - \tan\left(\frac{\theta_c}{2}\right) \tan\left(\frac{\omega_2 - \omega_1}{2}\right)}{1 + \tan\left(\frac{\theta_c}{2}\right) \tan\left(\frac{\omega_2 - \omega_1}{2}\right)} \quad (6-209)$$

令

$$\left. \begin{aligned} \alpha &= \cos\left(\frac{\omega_2 + \omega_1}{2}\right) / \cos\left(\frac{\omega_2 - \omega_1}{2}\right) \\ k &= \tan\left(\frac{\theta_c}{2}\right) \tan\left(\frac{\omega_2 - \omega_1}{2}\right) \end{aligned} \right\} \quad (6-210)$$

将(6-210)式代入(6-208)式、(6-209)式,可得

$$d_1 = \frac{-2\alpha}{1+k} \quad (6-211)$$

$$d_2 = \frac{1-k}{1+k} \quad (6-212)$$

同样可求得

$$\alpha = \cos\omega_0 \quad (6-213)$$

以上式子中  $\omega_2, \omega_1$  分别为带阻滤波器阻带边沿的上、下两个通带截止频率,见图 6-39,  $\omega_0$  为阻带中心频率。

概括起来,数字低通—数字带阻变换时,  $z$  域和  $Z$  域单位圆上各点的对应关系表示为

$$\begin{aligned} z: 1 \rightarrow e^{j\theta_c} \rightarrow -1 \rightarrow e^{-j\theta_c} \rightarrow 1 \rightarrow e^{j\theta_c} \rightarrow -1 \rightarrow e^{-j\theta_c} \rightarrow 1 \\ Z: 1 \rightarrow e^{j\omega_1} \rightarrow e^{j\omega_0} \rightarrow e^{j\omega_2} \rightarrow -1 \rightarrow e^{-j\omega_2} \rightarrow e^{-j\omega_0} \rightarrow e^{-j\omega_1} \rightarrow 1 \end{aligned}$$

## 五、数字低通—多通带的数字滤波器

这是最一般的情况,在特定情况下,就可得到前面所引出的四种基本变换。这里我们关心的是多通带滤波器的阶次  $N$  的选择方法以及它的全通变换函数前的正负号的选择。

### 1. 多通带滤波器阶数 $N$ 的选择

变换的全通函数的阶数  $N$  即为  $(-\pi, \pi)$  范围或  $(0, 2\pi)$  范围内通带的数目,我们知道频率响应的幅度响应是对  $\omega=0$  的纵轴成偶对称的。若实际给出通带数为 2 个,这一定是指  $(0, \pi)$  范围内的通带数,由于偶对称关系在  $(-\pi, \pi)$  范围内就有 4 个通带,因而  $N=4$ 。

### 2. 多通带滤波器的全通变换函数前的符号

上面已看到,高通和带通滤波器的全通变换函数前是取“-”号,而低通和带阻滤波器的全通变换函数前是取“+”号,推而广之,如果  $\omega=0$  时是阻带的情况(例如高通、带通滤波器),则全通变换函数前取“-”号,如果  $\omega=0$  时,是通带的情况(例如低通、带阻滤波器),则全通变换函数前取“+”号。下面分别加以讨论。

(1) 当  $\omega=0$  时是阻带的情况(例如高通、带通滤波器),即对于多带通滤波器,则全通变换函数为

$$z^{-1} = G(z^{-1}) = -\frac{Z^{-N} + d_1 Z^{-(N-1)} + d_2 Z^{-(N-2)} + \cdots + d_{N-1} Z^{-1} + d_N}{1 + d_1 Z^{-1} + d_2 Z^{-2} + \cdots + d_{N-1} Z^{-(N-1)} + d_N Z^{-N}} = \frac{-Z^{-N} D(z^{-1})}{D(z)} \quad (6-214)$$

这个关系式可见  $N$  阶全通滤波器的(6-19)式。与低通原型滤波器上、下截止频率  $\theta_c, -\theta_c$  相对应可找到变换后多通带滤波器的每一通带的对应的两个频率，在  $(0 \sim \pi)$  之间，共有  $N/2$  个通带，共有  $N$  个独立的频率点，可得到  $N$  个方程，这样可解出  $N$  个系数  $d_k (k=1, 2, \dots, N)$ ，设第  $i$  个通带上边带为  $\omega_{i1}$ ，下边带为  $\omega_{i2}$ ，与低通原型数字频率( $\theta$ )的对应关系为： $-\theta_c$  对应于  $\omega_{i1}$ ， $\theta_c$  对应于  $\omega_{i2}$ ，即按照变换的以下边界条件：

$$\begin{aligned} z^{-1} = e^{j\theta_c} &\rightarrow Z^{-1} = e^{-j\omega_{i1}} \\ z^{-1} = e^{-j\theta_c} &\rightarrow Z^{-1} = e^{-j\omega_{i2}} \end{aligned} \quad (6-215)$$

将(6-215)式代入(6-214)式中，经化简可得

$$\left. \begin{aligned} \cos\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1}\right) + \sum_{k=1}^N d_k \cos\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1} - k\omega_{i1}\right) &= 0 \\ \cos\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2}\right) + \sum_{k=1}^N d_k \cos\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2} + k\omega_{i2}\right) &= 0 \end{aligned} \right\} \quad i = 1, 2, \dots, \frac{N}{2} \quad (6-216)$$

从(6-216)式的  $N$  个方程，可联立解出  $N$  个系数  $d_k (k=1, 2, \dots, N)$ ，这样就可得到变换后多通带滤波器的系统函数。

(2) 当  $\omega=0$  时是通带的情况(例如低通、带阻滤波器)，即对于多带阻滤波器，此时变换函数为

$$z^{-1} = G(Z^{-1}) = \frac{Z^{-N} + d_1 Z^{-(N-1)} + d_2 Z^{-(N-2)} + \dots + d_{N-1} Z^{-1} + d_N}{1 + d_1 Z^{-1} + d_2 Z^{-2} + \dots + d_{N-1} Z^{-(N-1)} + d_N Z^{-N}} = \frac{D(Z^{-1})}{Z^{-N} D(Z)} \quad (6-217)$$

同理可得

$$\left. \begin{aligned} \sin\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1}\right) + \sum_{k=1}^N d_k \sin\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1} - k\omega_{i1}\right) &= 0 \\ \sin\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2}\right) + \sum_{k=1}^N d_k \sin\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2} + k\omega_{i2}\right) &= 0 \end{aligned} \right\} \quad i = 1, 2, \dots, \frac{N}{2} \quad (6-218)$$

解此  $N$  个联立方程后，即可求出要求的  $N$  个系数  $d_k (k=1, 2, \dots, N)$ ，这样就可得到变换后多通带滤波器的系统函数。

我们把这 6 种变换的变换公式及变换参数关系都列在表 6-9 中。

表 6-9 由截止频率为  $\theta_c$  的低通数字滤波器( $z$ )变换为各型数字滤波器( $Z$ )

变换类型	变换公式	变换参数的公式
低通   低通	$z^{-1} = \frac{Z^{-1} - \alpha}{1 - \alpha Z^{-1}}$	$\alpha = \sin\left(\frac{\theta_c - \omega_c}{2}\right) / \sin\left(\frac{\theta_c + \omega_c}{2}\right)$
低通   高通	$z^{-1} = -\left(\frac{Z^{-1} + \alpha}{1 + \alpha Z^{-1}}\right)$	$\alpha = -\cos\left(\frac{\theta_c + \omega_c}{2}\right) / \cos\left(\frac{\theta_c - \omega_c}{2}\right)$
低通   带通	$z^{-1} = -\frac{Z^{-2} - \frac{2ak}{k+1}Z^{-1} + \frac{k-1}{k+1}}{\frac{k-1}{k+1}Z^{-2} - \frac{2ak}{k+1}Z^{-1} + 1}$	$\alpha = \cos\left(\frac{\omega_2 + \omega_1}{2}\right) / \cos\left(\frac{\omega_2 - \omega_1}{2}\right) = \cos\omega_0$ $k = \cot\left(\frac{\omega_2 - \omega_1}{2}\right) \tan\frac{\theta_c}{2}$ $\omega_2, \omega_1$ 为要求的上、下截止频率 $\omega_0$ 为通带中心频率

续表

变换类型	变换公式	变换参数的公式
低通 + 带阻	$z^{-1} = \frac{Z^{-2} - \frac{2\alpha}{1+k} Z^{-1} + \frac{1-k}{1+k}}{1-k Z^{-2} - \frac{2\alpha}{1+k} Z^{-1} + 1}$	$\alpha = \frac{\cos\left(\frac{\omega_2 + \omega_1}{2}\right)}{\cos\left(\frac{\omega_2 - \omega_1}{2}\right)} = \cos\omega_0$ $k = \tan\left(\frac{\omega_2 - \omega_1}{2}\right) \tan\frac{\theta_c}{2}$ $\omega_2, \omega_1$ 为要求的上、下截止频率 $\omega_0$ 为阻带中心频率
低通 + 多带通	$z^{-1} = \frac{Z^{-N} + \sum_{k=1}^N d_k Z^{-(N-k)}}{1 + \sum_{k=1}^N d_k Z^{-k}}$	$\cos\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1}\right) + \sum_{k=1}^N d_k \cos\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1} - k\omega_{i1}\right) = 0$ $\cos\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2}\right) + \sum_{k=1}^N d_k \cos\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2} + k\omega_{i2}\right) = 0$ $\omega_{i2}, \omega_{i1}$ 分别为第 $i$ 个通带的上、下截止频率, $1 \leq i \leq N/2, N/2$ 为 $(0 \sim \pi)$ 之间的通带数
低通 + 多带阻	$z^{-1} = \frac{Z^{-N} + \sum_{k=1}^N d_k Z^{-(N-k)}}{1 + \sum_{k=1}^N d_k Z^{-k}}$	$\sin\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1}\right) + \sum_{k=1}^N d_k \sin\left(\frac{\theta_c}{2} + \frac{N}{2}\omega_{i1} - k\omega_{i1}\right) = 0$ $\sin\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2}\right) + \sum_{k=1}^N d_k \sin\left(\frac{\theta_c}{2} - \frac{N}{2}\omega_{i2} + k\omega_{i2}\right) = 0$ $\omega_{i2}, \omega_{i1}$ 分别为第 $i$ 个阻带的上、下截止频率, $1 \leq i \leq N/2, N/2$ 为 $(0 \sim \pi)$ 之间的阻带数

下面举两个例子, 来具体说明设计的方法。

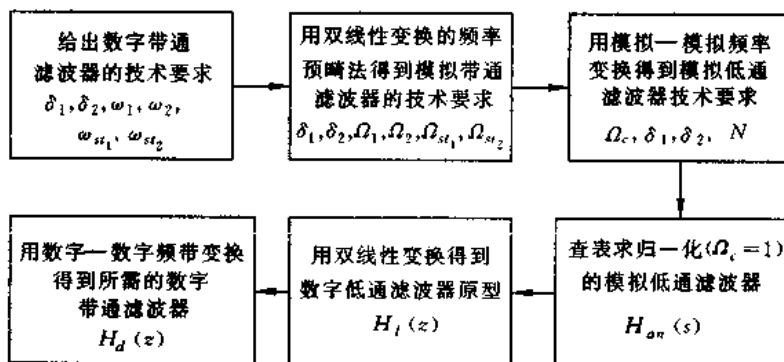
[例 6-11] 要求设计一个数字带通滤波器, 给定指标为

通带内允许起伏:  $\geq -1\text{dB}$ ,  $0.4\pi \leq \omega \leq 0.5\pi$

阻带衰减:  $\leq -15\text{dB}$ ,  $0 \leq \omega \leq 0.2\pi, 0.7\pi \leq \omega \leq \pi$

用切贝雪夫滤波器逼近, 求此滤波器的系统函数  $H_d(z)$ 。

解 整个设计步骤可用以下框图表示:



具体设计过程如下:

(1) 数字滤波器的各项指标为

$$\delta_1 = 1\text{dB}, \delta_2 = 15\text{dB}, \omega_1 = 0.4\pi, \omega_2 = 0.5\pi, \omega_{s1} = 0.2\pi, \omega_{s2} = 0.7\pi$$

(2) 采用双线性变换的频率预畸法(为了避免混叠),求相应的模拟带通滤波器的各项技术要求(令  $T=1$ )。

$$\Omega_1 = 2\tan\left(\frac{\omega_1}{2}\right) = 2\tan\left(\frac{0.4\pi}{2}\right) = 1.4531$$

$$\Omega_2 = 2\tan\left(\frac{\omega_2}{2}\right) = 2\tan\left(\frac{0.5\pi}{2}\right) = 2$$

$$\Omega_{s1} = 2\tan\left(\frac{\omega_{s1}}{2}\right) = 2\tan\left(\frac{0.2\pi}{2}\right) = 0.6498$$

$$\Omega_{s2} = 2\tan\left(\frac{\omega_{s2}}{2}\right) = 2\tan\left(\frac{0.7\pi}{2}\right) = 3.9252$$

(3) 求模拟低通滤波器技术要求,按模拟低通—模拟带通的变换关系(6-128)式,可得

$$\Omega_c = \Omega_2 - \Omega_1 = 0.5469$$

由(6-137)式可得带通滤波器通带中心频率为

$$\Omega_0 = \sqrt{\Omega_1 \Omega_2} = 1.7048$$

因而,由(6-134)式可知,低通滤波器的截止频率应由以下  $\Omega'_{s1}, \Omega'_{s2}$  中选一个:

$$\Omega'_{s2} = \frac{\Omega_{s2}^2 - \Omega_0^2}{\Omega_{s2}} = 3.1848$$

$$\Omega'_{s1} = \frac{\Omega_{s1}^2 - \Omega_0^2}{\Omega_{s1}} = -3.8223$$

取模拟低通滤波器的阻带截止频率为

$$\Omega_u = \min(|\Omega'_{s2}|, |\Omega'_{s1}|) = |\Omega'_{s2}| = 3.1848$$

这是因为较小频率值处衰减如能满足要求( $-15\text{dB}$ ),则频率值较大处衰减一定会超过要求( $<-15\text{dB}$ )。图 6-40 给出了此例中得到的模拟带通滤波器设计要求特性及与之对应的低通原型特性。

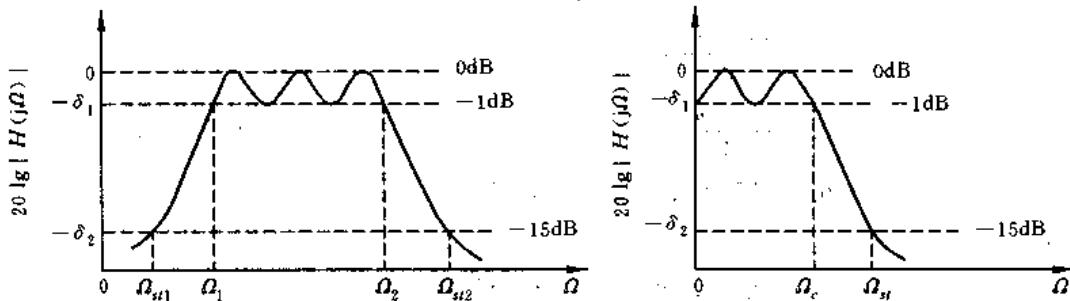


图 6-40 例 6-11 中模拟带通滤波器设计要求特性及低通原型特性

模拟低通滤波器的阶数为

$$N \geq \frac{\operatorname{ch}^{-1}\left[\frac{1}{\epsilon} \sqrt{10^{0.1\delta_2} - 1}\right]}{\operatorname{ch}^{-1}\left(\frac{\Omega_u}{\Omega_c}\right)}$$

将  $\epsilon = \sqrt{10^{k_1/10} - 1} = \sqrt{10^{0.1} - 1} = 0.508847$  代入此式, 可得

$$N \geq \frac{\operatorname{ch}^{-1}\left(\frac{1}{0.508847} \sqrt{10^{1.5} - 1}\right)}{\operatorname{ch}^{-1}\left(\frac{3.1848}{0.5469}\right)} = \frac{\operatorname{ch}^{-1}(10.8751)}{\operatorname{ch}^{-1}(5.82337)} = \frac{3.0775}{2.44756} = 1.257$$

取  $N=2$ 。

(4) 查表 6-5 求归一化切贝雪夫模拟低通滤波器系统函数  $H_m(s)$ ,  $N=2$ ,  $\delta_1=1\text{dB}$ , 可得( $N$  为偶数,  $H_m(0)=\frac{1}{\sqrt{1+\epsilon^2}}$ , 可求得  $H_m(s)$  分子系数)

$$H_m(s) = \frac{0.982613}{s^2 + 1.0977s + 1.1025}$$

(5) 用双线性变换法, 求数字低通原型滤波器的系统函数

$$\begin{aligned} H_t(z) &= \frac{0.9826}{\left(\frac{2(1-z^{-1})}{1+z^{-1}}\right)^2 + 1.0977\left(\frac{2(1-z^{-1})}{1+z^{-1}}\right) + 1.1025} \\ &= \frac{0.1346(1+z^{-1})^2}{1 - 0.7941z^{-1} + 0.3983z^{-2}} \end{aligned}$$

(6) 利用数字低通—数字带通变换求得数字带通滤波器的系统函数  $H_d(z)$ 。由于模拟归一化原型低通滤波器截止频率为  $\Omega_c=1$ , 双线性变换后, 数字低通原型低通滤波器的截止频率  $\theta_c$  应满足双线性变换频率间的关系式

$$\Omega_c = \frac{2}{T} \tan \frac{\theta_c}{2}$$

取  $T=1$ , 则有  $\theta_c = 2 \operatorname{arctan} \left( \frac{\Omega_c}{2} \right) = 2 \operatorname{arctan} \frac{1}{2} = 0.9273 = 0.2952\pi$ , 而按表 6-9 可知

$$\begin{aligned} \alpha &= \frac{\cos\left(\frac{\omega_2 + \omega_1}{2}\right)}{\cos\left(\frac{\omega_2 - \omega_1}{2}\right)} = \frac{\cos(0.45\pi)}{\cos(0.05\pi)} = \frac{0.1564}{0.9877} = 0.1583 \\ k &= \cot\left(\frac{\omega_2 - \omega_1}{2}\right) \tan \frac{\theta_c}{2} = \cot(0.05\pi) \cdot \tan(0.1476\pi) \\ &= 6.31375 \times 0.50006 = 3.1573 \end{aligned}$$

所以

$$\begin{aligned} d_1 &= \frac{-2ak}{k+1} = \frac{-0.9996}{4.1573} = -0.2404 \\ d_2 &= \frac{k-1}{k+1} = \frac{2.1569}{4.1573} = 0.5188 \end{aligned}$$

因而

$$z^{-1} = G(Z^{-1}) = -\frac{Z^{-2} + d_1 Z^{-1} + d_2}{d_2 Z^{-2} + d_1 Z^{-1} + 1} = -\frac{Z^{-2} - 0.2404 Z^{-1} + 0.5188}{0.5188 Z^{-2} - 0.2404 Z^{-1} + 1}$$

从而可得

$$\begin{aligned}
H_d(z) &= H_l(z)|_{z^{-1}=G(z^{-1})} = -(z^{-2} + d_1 z^{-1} + d_2)/(d_2 Z^{-2} + d_1 Z^{-1} + 1) \\
&= \frac{0.1346 \left[ 1 - (Z^{-2} - 0.2404Z^{-1} + 0.5188)/(0.5188Z^{-2} - 0.2404Z^{-1} + 1) \right]^2}{1 + 0.7941 \left( \frac{Z^{-2} - 0.2404Z^{-1} + 0.5188}{0.5188Z^{-2} - 0.2404Z^{-1} + 1} \right) + 0.3983 \left( \frac{Z^{-2} - 0.2404Z^{-1} + 0.5188}{0.5188Z^{-2} - 0.2404Z^{-1} + 1} \right)^2} \\
&= \frac{0.03117Z^{-4} - 0.06233Z^{-2} + 0.03117}{1.5192 - 0.8701Z^{-1} + 2.5854Z^{-2} - 0.7309Z^{-3} + 1.0794Z^{-4}} \\
&= \frac{0.02052Z^{-4} - 0.04103Z^{-2} + 0.02052}{1 - 0.5727Z^{-1} + 1.7078Z^{-2} - 0.4811Z^{-3} + 0.7105Z^{-4}}
\end{aligned}$$

图 6-41 表示了相应的带通滤波器的幅度特性。

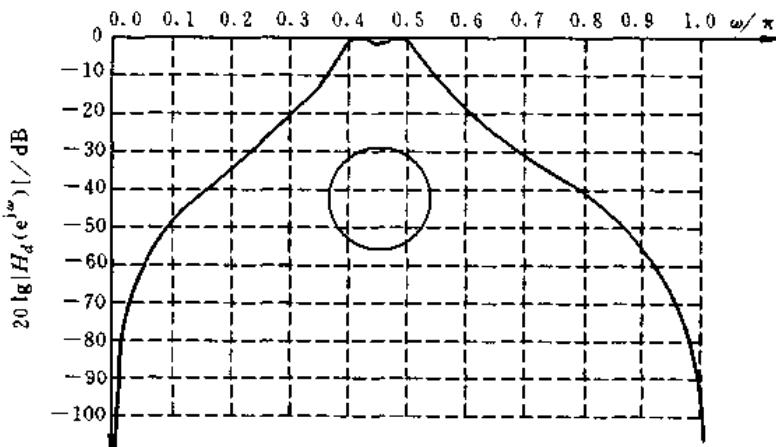


图 6-41 例 6-11 的带通滤波器的幅频特性

[例 6-12] 要求设计有两个通带的带通滤波器, 给定通带指标为

$$\frac{3\pi}{16} \leq \omega \leq \frac{5\pi}{16}, \quad \frac{5\pi}{8} \leq \omega \leq \frac{7\pi}{8}$$

如采用上例中求出的数字低通滤波器作为数字低通原型滤波器, 试求此双带通滤波器的系统函数  $H_d(Z)$ 。

解 这是多通带频率变换, 由于要求的是在  $(0, \pi)$  之间的双通带, 则在  $(-\pi, \pi)$  之间共有  $N=4$  个通带, 因而频率变换式应为 4 阶全通函数

$$z^{-1} = G(z^{-1}) = -\frac{Z^{-4} + d_1 Z^{-3} + d_2 Z^{-2} + d_3 Z^{-1} + d_4}{d_4 Z^{-4} + d_3 Z^{-3} + d_2 Z^{-2} + d_1 Z^{-1} + 1}$$

由上例知  $\theta_c = 0.2952\pi$

又第一个带通有  $\omega_{11} = \frac{3\pi}{16}$ ,  $\omega_{12} = \frac{5\pi}{16}$

第二个带通有  $\omega_{21} = \frac{5\pi}{8}$ ,  $\omega_{22} = \frac{7\pi}{8}$

把  $\theta_c$  和  $\omega_{ii}$ ,  $\omega_{ij}$  ( $i=1, 2$ ) 分别代入 (6-216) 式, 可得一组包含  $d_k$  ( $k=1, 2, 3, 4$ ) 的联立

方程

$$\begin{aligned}\cos\left(\frac{\theta_c}{2} + 2\omega_{11}\right) + d_1 \cos\left(\frac{\theta_c}{2} + \omega_{11}\right) + d_2 \cos\left(\frac{\theta_c}{2}\right) + d_3 \cos\left(\frac{\theta_c}{2} - \omega_{11}\right) + d_4 \cos\left(\frac{\theta_c}{2} - 2\omega_{11}\right) &= 0 \\ \cos\left(\frac{\theta_c}{2} + 2\omega_{21}\right) + d_1 \cos\left(\frac{\theta_c}{2} + \omega_{21}\right) + d_2 \cos\left(\frac{\theta_c}{2}\right) + d_3 \cos\left(\frac{\theta_c}{2} - \omega_{21}\right) + d_4 \cos\left(\frac{\theta_c}{2} - 2\omega_{21}\right) &= 0 \\ \cos\left(\frac{\theta_c}{2} - 2\omega_{12}\right) + d_1 \cos\left(\frac{\theta_c}{2} - \omega_{12}\right) + d_2 \cos\left(\frac{\theta_c}{2}\right) + d_3 \cos\left(\frac{\theta_c}{2} + \omega_{12}\right) + d_4 \cos\left(\frac{\theta_c}{2} + 2\omega_{12}\right) &= 0 \\ \cos\left(\frac{\theta_c}{2} - 2\omega_{22}\right) + d_1 \cos\left(\frac{\theta_c}{2} - \omega_{22}\right) + d_2 \cos\left(\frac{\theta_c}{2}\right) + d_3 \cos\left(\frac{\theta_c}{2} + \omega_{22}\right) + d_4 \cos\left(\frac{\theta_c}{2} + 2\omega_{22}\right) &= 0\end{aligned}$$

将各参数代入, 可求得以下联立方程:

$$\begin{aligned}0.8688d_1 + 0.8944d_2 + 0.1250d_3 - 0.7555d_4 &= -0.0709 \\ -0.6552d_1 + 0.8944d_2 - 0.9975d_3 + 0.9487d_4 &= -0.3162 \\ 0.4952d_1 + 0.8944d_2 + 0.9922d_3 + 0.7555d_4 &= 0.709 \\ -0.7555d_1 + 0.8944d_2 + 0.0709d_3 - 0.9487d_4 &= 0.3162\end{aligned}$$

解此联立方程可得

$$d_1 = -0.2317 \quad d_2 = -0.0198 \quad d_3 = 0.3145 \quad d_4 = -0.1439$$

从而所要求的数字频率变换式为

$$z^{-1} = G(z^{-1}) = -\frac{Z^{-4} - 0.2317Z^{-3} - 0.0198Z^{-2} + 0.3145Z^{-1} - 0.1439}{1 - 0.2317Z^{-1} - 0.0198Z^{-2} + 0.3145Z^{-3} - 0.1439Z^{-4}}$$

而从上例中  $N=2$  的低通原型为

$$H_t(z) = \frac{0.1346(1+z^{-1})^2}{1 - 0.7941z^{-1} + 0.3983z^{-2}}$$

将  $z^{-1}$  的表达式代入  $H_t(z)$  中, 经化简即可求得所需的双通带滤波器的系统函数

$$\begin{aligned}H_d(Z) &= (0.1970 - 0.1882Z^{-1} + 0.0449Z^{-2} + 0.1882Z^{-3} - 0.4838Z^{-4} + 0.1882Z^{-5} \\ &\quad + 0.0449Z^{-6} - 0.1882Z^{-7} + 0.1970Z^{-8}) / (1 - 0.2497Z^{-1} - 0.0173Z^{-2} \\ &\quad + 0.4907Z^{-3} + 0.3648Z^{-4} + 0.0976Z^{-5} + 0.0435Z^{-6} + 0.0012Z^{-7} + 0.3409Z^{-8})\end{aligned}$$

其对应的幅度特性如图 6-42 所示。

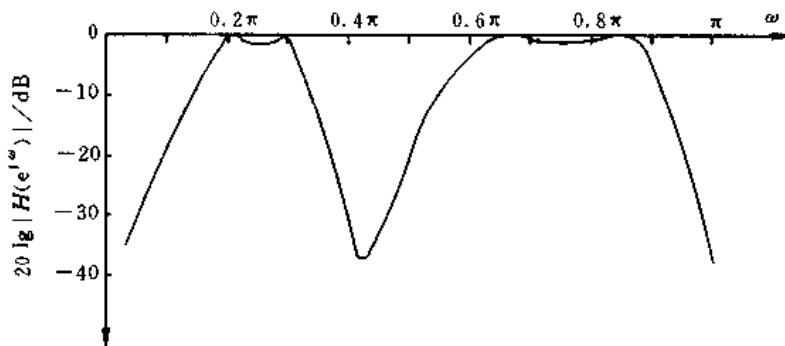


图 6-42 例 6-12 的双带通滤波器的幅频特性

## 6.12 直接在数字域设计 IIR 数字滤波器

这种设计法是直接在离散时域或离散频域设计 IIR 数字滤波器，在频域就是幅度平方函数法，在时域有帕德(Pade)逼近法以及在此法基础上的波形形成滤波器设计法。

### 一、IIR 数字滤波器的频域直接设计——幅度平方函数设计法

前面在讨论模拟巴特沃思滤波器以及模拟切贝雪夫滤波器时，是由其幅度平方函数  $|H_a(j\Omega)|^2$  来导出系统函数  $H_a(s)$  的。在数字域也可依照这一办法，利用幅度平方函数  $|H(e^{j\omega})|^2$  来导出系统函数  $H(z)$ 。问题是找到镜像多项式来描述  $|H(e^{j\omega})|^2$ ，镜像多项式就是其极点满足  $z=z_i$  及  $z=\frac{1}{z_i}$ ，即以单位圆为“镜子”而对称，同时为了是实系数的多项式，极点又必须成共轭对。为了使  $H(z)$  是  $z^{-1}$  的有理分式，而  $H(z)$  在单位圆上是  $e^{j\omega}$  的有理分式，因而幅度平方函数  $|H(e^{j\omega})|^2$  可以表示成  $\omega$  的两个三角函数的比值，这可很容易地导出。设数字滤波器的系统函数为

$$H(z) = \frac{\sum_{i=0}^M b_i z^{-i}}{1 + \sum_{k=1}^N a_k z^{-k}}$$

则滤波器幅度平方响应为

$$\begin{aligned} |H(e^{j\omega})|^2 &= H(z)H(z^{-1})|_{z=e^{j\omega}} = \frac{\sum_{i=0}^M b_i z^{-i} \sum_{k=0}^M b_k z^k}{\sum_{k=0}^N a_k z^{-k} \sum_{i=0}^N a_i z^i} \Big|_{z=e^{j\omega}} \\ &= \frac{\sum_{i=0}^M b_i e^{-j\omega i} \sum_{k=0}^M b_k e^{j\omega k}}{\sum_{k=0}^N a_k e^{-j\omega k} \sum_{i=0}^N a_i e^{j\omega i}} \quad (a_0 = 1) \end{aligned} \quad (6-219)$$

可写成

$$|H(e^{j\omega})|^2 = \frac{\sum_{k=0}^M c_k \cos(\omega k)}{\sum_{k=0}^N d_k \cos(\omega k)} \quad (6-220)$$

其中  $c_k, d_k$  是与  $a_k, b_k$  有关的量，利用三角恒等式

$$\cos \alpha = 2\cos^2\left(\frac{\alpha}{2}\right) - 1$$

可将(6-220)式写成

$$|H(e^{j\omega})|^2 = \frac{\sum_{k=0}^M e_k \cos^2\left(\frac{\omega k}{2}\right)}{\sum_{k=0}^N g_k \cos^2\left(\frac{\omega k}{2}\right)} \quad (6-221)$$

此式是许多综合数字滤波器方法的基础。

依照模拟滤波器情况,(6-221)式可用以下幅度平方函数代替:

$$|H(e^{j\omega})|^2 = \frac{1}{1 + A_N^2(\omega)} \quad (6-222)$$

对于低通滤波器,在巴特沃思滤波器情况下, $A_N^2(\omega)$ 可表示成

$$A_N^2(\omega) = \frac{\tan^{2N}\left(\frac{\omega}{2}\right)}{\tan^{2N}\left(\frac{\omega_c}{2}\right)} \quad (6-223)$$

其中 $\omega_c$ 是低通截止频率。同样对于低通切贝雪夫滤波器, $A_N^2(\omega)$ 可表示成

$$A_N^2(\omega) = \epsilon^2 C_N^2 \left[ \frac{\tan\left(\frac{\omega}{2}\right)}{\tan\left(\frac{\omega_c}{2}\right)} \right] \quad (6-224)$$

其中 $C_N(x)$ 是 $N$ 阶切贝雪夫多项式, $\epsilon$ 是波纹参数,可以证明,利用(6-223)式或(6-224)式,用 $z$ 代替 $e^{j\omega}$ 后,所得到的 $H(z)H(z^{-1})$ 一定是镜像多项式,满足设计要求。取其单位圆内的极点,即可得到 $H(z)$ 。

实际上可以导出

$$\tan^2\left(\frac{\omega}{2}\right) = \frac{\sin^2\left(\frac{\omega}{2}\right)}{\cos^2\left(\frac{\omega}{2}\right)} = \frac{1 - \cos\omega}{1 + \cos\omega} = \frac{1 - \frac{1}{2}(e^{j\omega} + e^{-j\omega})}{1 + \frac{1}{2}(e^{j\omega} + e^{-j\omega})}$$

代入 $z = e^{j\omega}$ ,可得

$$\tan^2\left(\frac{\omega}{2}\right) \Big|_{z=e^{j\omega}} = \frac{1 - \frac{1}{2}(z + z^{-1})}{1 + \frac{1}{2}(z + z^{-1})} = -\frac{(z - 1)^2}{(z + 1)^2} = -\frac{(1 - z^{-1})^2}{(1 + z^{-1})^2} \quad (6-225)$$

将(6-225)式代入(6-224)式,然后再代入(6-222)式,即可求得 $H(z)H(z^{-1})$ ,取其单位圆内的极点,即可组成 $H(z)$ 。

$A_N^2(\omega)$ 除了可以利用 $\frac{\tan^2\left(\frac{\omega}{2}\right)}{\tan^2\left(\frac{\omega_c}{2}\right)}$ 外,还可利用 $\frac{\sin^2\left(\frac{\omega}{2}\right)}{\sin^2\left(\frac{\omega_c}{2}\right)}$ (产生低通) $\frac{\cot^2\left(\frac{\omega}{2}\right)}{\cot^2\left(\frac{\omega_c}{2}\right)}$ (产生高通)。

可以证明,用幅度平方函数(6-223)式、(6-224)式得到的巴特沃思和切贝雪夫数字滤波器与同类型模拟滤波器的双线性变换相当。

## 二、IIR 数字滤波器的时域直接设计

### 1. 帕德逼近法

利用时域办法来设计一个 IIR 滤波器,使它的单位抽样响应逼近我们所要求的单位抽样响应。

设滤波器是因果性的(物理上可实现),其系统函数为

$$H(z) = \frac{\sum_{i=0}^M b_i z^{-i}}{\sum_{i=0}^N a_i z^{-i}} = \sum_{k=0}^{\infty} h(k) z^{-k} \quad (6-226)$$

帕德逼近法是找  $M+N+1$  个系数  $a_i, b_i$ (设  $a_0=1$ )以使得在  $0 \leq k \leq P-1$  的范围内,使  $h(k)$  逼近一个所希望的单位抽样响应  $h_d(k)$ 。

巴勒斯和帕克斯等人已证明,可以找到一组  $a_i, b_i$ ,使得以下的加权均方误差最小

$$\epsilon = \sum_{k=0}^{P-1} [h_d(k) - h(k)]^2 w(k) \quad (6-227)$$

其中  $w(k)$  是误差序列的正加权函数,因为  $h(k)$  是滤波器参数  $a_i, b_i$  的非线性函数,故只能用迭代法来得到  $\epsilon$  的最小值。当  $P=N+M+1$  时,解  $(N+M+1)$  个线性方程,可以得到使  $\epsilon$  为极小的滤波器参数。

设  $a_0=1$ , (6-226)式可写成

$$\sum_{k=0}^{P-1} h(k) z^{-k} \sum_{i=0}^N a_i z^{-i} = \sum_{i=0}^M b_i z^{-i}$$

代入  $P=N+M+1$ ,则有

$$\sum_{k=0}^{M+N} h(k) z^{-k} \sum_{i=0}^N a_i z^{-i} = \sum_{i=0}^M b_i z^{-i}$$

令等式两边  $z$  的同次幂的系数相等,即可得  $N+M+1$  个方程:

$$\begin{aligned} h(0) &= b_0 \\ h(0)a_1 + h(1) &= b_1 \\ h(0)a_2 + h(1)a_1 + h(2) &= b_2 \\ &\vdots \end{aligned}$$

由于  $i>M$  时,  $b_i=0$ ,因而一般表达式可分为以下两种表示(一般情况下  $M<N$ ):

$$\sum_{j=0}^k a_j h(k-j) = b_k, \quad 0 \leq k \leq M \quad (6-228)$$

$$\sum_{j=0}^N a_j h(k-j) = 0, \quad M < k \leq M+N = P-1 \quad (6-229)$$

因为我们希望  $h(k)$  充分逼近  $h_d(k)$ ,所以用  $h_d(k)$  代替  $h(k)$ ,即令  $h(k)=h_d(k)$  ( $k=0, 1, \dots, M+N$ ),这样来求解(6-228)式及(6-229)式,得到  $N$  个  $a_i$  及  $M+1$  个  $b_i$ ,即

(1) 令  $h_d(k)=h(k)$  ( $k=M+1, M+2, \dots, M+N$ ),利用(6-229)式求得  $N$  个系数  $a_i$  ( $i=1, 2, \dots, N$ );

(2) 求得  $a_i$  后, 令  $h_d(k)=h(k)$  ( $k=0, 1, \dots, M$ ), 利用(6-228)式求得  $M+1$  个系数  $b_i$  ( $i=0, 1, \dots, M$ )。由于是无限长冲激响应系统, 一定满足  $M \leq N$ , 故方程组可以求解, 各  $a_i, b_i$  得出后, 也就确定了系统函数  $H(z)$ 。

可以看出, 这里是取(6-226)式  $H(z)$  的幂级数表示的前  $M+N+1$  项 ( $k=0$  到  $N+M$ ), 令其单位抽样响应  $h(n)$  等于所要求的单位抽样响应  $h_d(n)$ , 而把  $k=M+N$  以后的项都不加考虑。

我们可以将(6-228)式和(6-229)式用矩阵形式来表示为

$$\begin{bmatrix} b_0 \\ b_1 \\ b_2 \\ \vdots \\ b_{M-1} \\ b_M \\ 0 \\ \vdots \\ \vdots \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h(0) & 0 & 0 & \cdots & \cdots & \cdots & 0 & 1 \\ h(1) & h(0) & 0 & \cdots & \cdots & \cdots & 0 & a_1 \\ h(2) & h(1) & h(0) & \cdots & \cdots & \cdots & 0 & a_2 \\ \cdots & \vdots \\ h(M-1) & h(M-2) & h(M-3) & \cdots & \cdots & \cdots & 0 & \vdots \\ h(M) & h(M-1) & h(M-2) & \cdots & \cdots & \cdots & 0 & \vdots \\ h(M+1) & h(M) & h(M-1) & \cdots & \cdots & \cdots & 0 & \vdots \\ \cdots & \vdots \\ h(N-1) & h(N-2) & h(N-3) & \cdots & \cdots & h(0) & 0 & \vdots \\ h(N) & h(N-1) & h(N-2) & \cdots & \cdots & h(1) & h(0) & \vdots \\ h(N+1) & h(N) & h(N-1) & \cdots & \cdots & h(2) & h(1) & \vdots \\ \cdots & \vdots \\ h(P-1) & h(P-2) & h(P-3) & \cdots & \cdots & h(P-N) & h(P-N-1) & a_N \end{bmatrix} \quad (6-230)$$

将矩阵分块, 得

$$\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_M \\ \vdots \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} h(0) & 0 & \cdots & \cdots & \cdots & 0 & 1 \\ \vdots & \vdots & \ddots & & & \vdots & a_1 \\ h(M) & \cdots & \cdots & \cdots & \cdots & 0 & a_2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h(M+1) & \cdots & \cdots & \cdots & \cdots & 0 & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ h(P-1) & \cdots & \cdots & \cdots & \cdots & h(P-N-1) & a_N \end{bmatrix} \quad (6-231)$$

可表示为

$$\begin{bmatrix} b \\ 0 \end{bmatrix} = \begin{bmatrix} h_1 \\ h_2 \end{bmatrix} [a] \quad (6-232)$$

在上面的式子中,  $P=M+N+1$ , 则  $h_2$  有  $N$  行和  $(N+1)$  列, 由于  $a_0=1$ , 故可以得到

$$[h^{(1)}] = -[h_2][a^{(1)}]$$

其中  $[h^{(1)}]$  是  $[h_2]$  的第一列, 即

$$[h^{(1)}] = \begin{bmatrix} h(M+1) \\ h(M+2) \\ \vdots \\ h(P-1) \end{bmatrix} \quad (6-233)$$

$[h_3]$ 是 $[h_2]$ 的除第一列外剩余的 $N \times N$ 方阵,即

$$[h_3] = \begin{bmatrix} h(M) & h(M-1) & \cdots & \cdots & 0 \\ h(M+1) & h(M) & \cdots & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ h(P-2) & h(P-3) & \cdots & \cdots & h(P-N-1) \end{bmatrix} \quad (6-234)$$

$[a^{(1)}]$ 是省略掉 $a_0=1$ 后的 $[a]$ ,即

$$[a^{(1)}] = \begin{bmatrix} a_1 \\ a_2 \\ \cdots \\ a_N \end{bmatrix} \quad (6-235)$$

可以用下面式子解出各个 $a_i, b_i$ 系数:

$$[a^{(1)}] = -[h_3]^{-1}[h^{(1)}] \quad (6-236)$$

$$[b] = [h_1][a] \quad (6-237)$$

由于帕德逼近法只考虑时域逼近,因而当滤波器频率特性阻带衰减要求过高时(例如大于40dB),这一方法就难于达到要求,此时,可用此法得到的系数,作为其他更好的最优算法的初始估计值。

## 2. 波形形成滤波器

这种滤波器是对某特定输入波形,可以得到一个确定的输出波形,它是在帕德逼近法基础上的时域直接设计。由于

$$H(z) = \frac{Y(z)}{X(z)} = \frac{\sum_{i=0}^M y(i)z^{-i}}{\sum_{i=0}^{N'} x(i)z^{-i}} = \sum_{k=0}^{P-1} h(k)z^{-k} \quad (6-238)$$

如果有 $(N'+1)$ 个输入波形抽样 $x(i)$ 和 $(M'+1)$ 个输出波形抽样 $y(i)$ ,则其相应的线性差分方程右端应有 $M'+N'+1$ 项,也就是当冲激响应正好是 $M'+N'+1$ 项时,则可求得 $(M'+N'+1)$ 个系数,如帕德逼近法所述,设计是方便的。但是,问题是设计少于 $M'+N'+1$ 项系数的滤波器,这时由于舍弃一些项,可能造成滤波器的不稳定,所以要对所求出的结果加以检验。

设计波形形成器可分为以下两个步骤:

第一步,由所给的输入 $x(n)$ 以及所需的输出 $y_d(n)$ 来求实际滤波器的最佳冲激响应。设 $x(0), x(1), \dots, x(N')$ 表示输入波形的抽样值,而 $y_d(0), y_d(1), \dots, y_d(M')$ 表示所需要的输出波形的抽样值, $y(0), y(1), \dots, y(M')$ 表示实际的输出波形的抽样值。设滤波器的单位抽样响应为 $K$ 项( $K < M' + N' + 1$ ),我们按照最小均方误差的准则来求单位抽样响应的最佳值。即要求在 $h(j)$ 为某最佳值时,以下公式成立:

$$E = \sum_{n=0}^{M'} [y(n) - y_d(n)]^2 = \min \quad (6-239)$$

由于  $x(n)$  与  $h(n)$  都是因果序列, 所以实际输出为

$$y(n) = \sum_{i=0}^n x(n-i)h(i)$$

因而

$$E = \sum_{n=0}^{M'} \left\{ \left[ \sum_{i=0}^n x(n-i)h(i) \right] - y_d(n) \right\}^2 \quad (6-240)$$

令

$$\frac{\partial E}{\partial h(j)} = 0, \quad j = 0, 1, \dots, K-1 \quad (6-241)$$

可得

$$\sum_{n=0}^{M'} \left\{ \left[ \sum_{i=0}^n x(n-i)h(i) \right] - y_d(n) \right\} x(n-j) = 0, \quad j = 0, 1, \dots, K-1 \quad (6-242)$$

求解  $K$  个联立方程, 即可解得  $K$  个最佳的  $h(n)$ 。

第二步, 求出  $K$  个  $h(n)$  后, 按照帕德逼近法, 就可求得滤波器的  $K$  个系数  $a_i, b_i$  ( $a_0 = 1$  是已知的)。举例说明如下:

**[例 6-13]** 已知波形形成滤波器的时域性能要求为

$$\text{所需输出波形 } y_d(n) = \{y_d(0), y_d(1), y_d(2), y_d(3)\} = \{1, 0.25, 0.1, 0.01\}$$

$$\text{给定输入波形 } x(n) = \{x(0), x(1)\} = \{3, 1\}$$

- 求 (1) 四项滤波器(四个系数  $a_1, a_2, b_0, b_1$ , 但  $a_0 = 1$  除外)的系统函数;  
 (2) 三项滤波器(三个系数  $b_0, b_1, a_1$ )的系统函数。

解 由所给  $x(n), y(n)$ , 可得滤波器的系统函数应为

$$H(z) = \frac{Y(z)}{X(z)} = \frac{1 + 0.25z^{-1} + 0.1z^{-2} + 0.01z^{-3}}{3 + z^{-1}} \quad (6-243)$$

其相应的差分方程为

$$y(n) = \frac{1}{3} [x(n) + 0.25x(n-1) + 0.1x(n-2) + 0.01x(n-3) - y(n-1)] \quad (6-244)$$

### (1) 四项滤波器

已知有 4 个输出  $M' + 1 = 4$ , 故  $M' = 3$ , 代入(6-240)式, 可得

$$\begin{aligned} E &= \sum_{n=0}^3 \left\{ \sum_{i=0}^n x(n-i)h(i) - y_d(n) \right\}^2 \\ &= [x(0)h(0) - y_d(0)]^2 + [x(1)h(0) + x(0)h(1) - y_d(1)]^2 \\ &\quad + [x(2)h(0) + x(1)h(1) + x(0)h(2) - y_d(2)]^2 \\ &\quad + [x(3)h(0) + x(2)h(1) + x(1)h(2) + x(0)h(3) - y_d(3)]^2 \end{aligned} \quad (6-245)$$

注意, 这里  $x(2) = x(3) = 0$ 。把各  $x(n)$  值, 各  $y_d(n)$  值代入(6-245)式, 可得

$$E = 10h^2(0) - 6.5h(0) + 6h(0)h(1) + 10h^2(1) - 1.7h(1) + 6h(1)h(2)$$

$$+ 10h^2(2) - 0.62h(2) + 6h(2)h(3) + 9h^2(3) - 0.06h(3) + 1.0726 \quad (6-246)$$

由于要求四项滤波器，故  $K-1=3$ ，利用(6-241)式，求最佳的  $h(j)$  ( $j=0, 1, \dots, K-1$ ) 值，可得

$$\begin{aligned} 20h(0) + 6h(1) &= 6.5 \\ 6h(0) + 20h(1) + 6h(2) &= 1.7 \\ 6h(1) + 20h(2) + 6h(3) &= 0.62 \\ 6h(2) + 18h(3) &= 0.06 \end{aligned}$$

联立求解此四个方程可得

$$\begin{aligned} h(0) &= 0.3333 \\ h(1) &= -0.0278 \\ h(2) &= 0.0426 \\ h(3) &= -0.0109 \end{aligned}$$

求出  $h(n)$  后，首先就是利用帕德逼近法求出滤波器系统函数  $H(z)$  的系数，由(6-231)式的分块矩阵，代入  $M=1, N=2$ （因为是四项系数），得

$$\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} h(0) & 0 & 0 \\ h(1) & h(0) & 0 \\ \vdots & \vdots & \vdots \\ h(2) & h(1) & h(0) \\ h(3) & h(2) & h(1) \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix}$$

由(6-236)式得

$$\begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = -\begin{bmatrix} h(1) & h(0) \\ h(2) & h(1) \end{bmatrix}^{-1} \begin{bmatrix} h(2) \\ h(3) \end{bmatrix}$$

解此矩阵方程可得

$$\begin{aligned} a_1 &= 0.1824 \\ a_2 &= -0.1126 \end{aligned}$$

由(6-237)式得

$$\begin{bmatrix} b_0 \\ b_1 \end{bmatrix} = \begin{bmatrix} h(0) & 0 & 0 \\ h(1) & h(0) & 0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ a_2 \end{bmatrix}$$

解此矩阵方程可得

$$\begin{aligned} b_0 &= 0.3333 \\ b_1 &= 0.0330 \end{aligned}$$

因而实际得到的系统函数是

$$H(z) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1} + a_2 z^{-2}} = \frac{0.3333 + 0.0330 z^{-1}}{1 + 0.1824 z^{-1} - 0.1126 z^{-2}} = \frac{Y(z)}{X(z)}$$

所以实际得到的差分方程为

$$y(n) = 0.3333x(n) + 0.0330x(n-1) - 0.1824y(n-1) + 0.1126y(n-2)$$

这样,对于给定输入  $x(0)=3, x(1)=1$ ,可求得实际输出  $y(n)$ 并列表表示如下。表中把所需的输出  $y_d(n)$ 也列入,以便比较。

$n$	0	1	2	3	4	5	6	7	8
$y(n)$	0.9999	0.2499	0.1	0.0099	0.0095	0.0006	0.0012	-0.0003	0.0002
$y_d(n)$	1	0.25	0.1	0.01	0	0	0	0	0

其次,我们来检验所设计的滤波器是否稳定,先将  $H(z)$ 化成

$$H(z) = \frac{0.3333(z+0.0990)z}{(z+0.4389)(z-0.2565)}$$

由此看出,所设计的滤波器极点为  $z=0.4389, z=-0.2565$ ,都位于单位圆内,滤波器是稳定的,而且实际输出接近于所要求的输出。

## (2) 三项滤波器

和四项滤波器的惟一差别是  $h(3)=0$ ,故(6-246)式变成

$$\begin{aligned} E = & 10h^2(0) - 6.5h(0) + 6h(0)h(1) + 10h^2(1) - 1.7h(1) + 6h(1)h(2) \\ & + 10h^2(2) - 0.62h(2) + 1.0726 \end{aligned}$$

求解

$$\frac{\partial E}{\partial h(j)} = 0, \quad j = 0, 1, 2$$

可得

$$\begin{aligned} 20h(0) + 6h(1) &= 6.5 \\ 6h(0) + 20h(1) + 6h(2) &= 1.7 \\ 6h(1) + 20h(2) &= 0.62 \end{aligned}$$

可得

$$\begin{aligned} h(0) &= 0.3330 \\ h(1) &= -0.0266 \\ h(2) &= 0.0390 \end{aligned}$$

利用公式(6-236)式求得

$$a_1 = -\frac{1}{h(1)} \cdot h(2) = \frac{0.0390}{0.0266} = 1.4662$$

利用公式(6-237)式求得

$$\begin{aligned} b_0 &= h(0) = 0.3330 \\ b_1 &= h(1) + a_1h(0) = -0.0266 + 0.3330 \times 1.4662 = 0.4616 \end{aligned}$$

从而得到系统函数

$$H(z) = \frac{b_0 + b_1 z^{-1}}{1 + a_1 z^{-1}} = \frac{0.3330 + 0.4616 z^{-1}}{1 + 1.4662 z^{-1}} = \frac{0.3330(z + 1.3862)}{z + 1.4662}$$

可以看出,  $H(z)$ 的极点在  $z=-1.4662$  处,位于  $z$  平面单位圆的外部,故此滤波器是不稳定的,因而不能得到所要求的输出波形。可以利用级联全通滤波器的办法,使极点反射到单位圆内,从而可得到有相同频率响应幅度的稳定滤波器。

## 6.13 设计 IIR 滤波器的最优化方法

当设计的是任意频率响应的 IIR 数字滤波器时,就不存在明显的解析方法,也就是说系数不存在显式表达式,其设计方法都是依靠计算机来解线性或非线性方程组的算法,以此来确定滤波器的系数。这就是最优化算法,而这一算法依据的是某一最优化准则。

采用这种计算机辅助设计技术来逼近任意频响特性,可以有多种方法。我们只限于导出各种方法的设计方程的公式,而不去讨论数值计算的细节。

### 一、最小均方误差设计法

根据频域均方误差最小准则,施泰利兹(K. Steiglitz)于 1970 年提出了这种设计方法。

我们已经知道,在一组离散频率点  $\omega_i (i=1, 2, \dots, M)$  上所要求的频率响应  $H_d(e^{j\omega})$  的值为  $H_d(e^{j\omega_i})$ , 假定实际求出的频率响应为  $H(e^{j\omega})$ , 那么, 在这些给定的离散频率点处, 所要求的频率响应的幅度与求出的实际频率响应的幅度的均方误差为

$$E = \sum_{i=1}^M [|H(e^{j\omega_i})| - |H_d(e^{j\omega_i})|]^2 \quad (6-247)$$

设计的目的是, 调整各  $H(e^{j\omega})$ , 即调整  $H(e^{j\omega})$  的系数, 使  $E$  为最小, 这样得到的  $H(e^{j\omega})$  作为  $H_d(e^{j\omega})$  的逼近值。

实际滤波器  $H(e^{j\omega})$  常采用二阶节的级联形式表示, 因为这种结构其频率响应对系数变化的灵敏度低(这使系数量化造成的误差减小), 便于调整频率响应, 而且在最优化过程中计算导数较为方便。设

$$H(z) = A \prod_{n=1}^k \frac{1 + a_n z^{-1} + b_n z^{-2}}{1 + c_n z^{-1} + d_n z^{-2}} = AP(z) \quad (6-248)$$

将从(6-248)式得到的  $H(e^{j\omega}) = H(z) \Big|_{z=e^{j\omega}}$  代入(6-247)式, 可以看出, 均方误差  $E$  是  $a_n$ ,  $b_n$ ,  $c_n$ ,  $d_n (n=1, 2, \dots, k)$  以及  $A$  的函数, 所以  $E$  是  $(4k+1)$  个未知参量的函数。

将  $A$  以外的  $4k$  个参量用矢量  $\Phi$  表示为

$$\Phi = (a_1, b_1, c_1, d_1, a_2, b_2, c_2, d_2, \dots, a_k, b_k, c_k, d_k)^T$$

其中角标 T 表示转置, 则(6-247)式可写成

$$E = E(\Phi, A) = \sum_{i=1}^M [|AP(e^{j\omega_i}, \Phi)| - |H_d(e^{j\omega_i})|]^2 \quad (6-249)$$

为了求出使  $E$  最小的各未知参数, 可以利用以下  $4k+1$  个方程:

$$\frac{\partial E(\Phi, A)}{\partial |A|} = 0 \quad (6-250)$$

$$\frac{\partial E(\Phi, A)}{\partial \phi_n} = 0, \quad n = 1, 2, \dots, 4k \quad (6-251)$$

式中  $\phi_n$  是  $\Phi$  的第  $n$  个分量。利用计算机就可以解出这  $4k+1$  个系数, 把系数值代入

(6-248)式,即求得所设计的实际  $H(z)$ 。下面来讨论(6-250)式和(6-251)式的求解。

由(6-249)式可得

$$\frac{\partial E(\Phi, A)}{\partial |A|} = \sum_{i=1}^M (2[|A| + |P(e^{j\omega_i}, \Phi)| + |H_d(e^{j\omega_i})|] + |P(e^{j\omega_i}, \Phi)|)$$

令  $\frac{\partial E(\Phi, A)}{\partial |A|} = 0$ , 可解得

$$|A_0| = \frac{\sum_{i=1}^M |P(e^{j\omega_i}, \Phi)| + |H_d(e^{j\omega_i})|}{\sum_{i=1}^M |P(e^{j\omega_i}, \Phi)|^2} \quad (6-252)$$

其中

$$P(e^{j\omega_i}, \Phi) = \prod_{n=1}^k \frac{1 + a_n e^{-j\omega_i} + b_n e^{-j2\omega_i}}{1 + c_n e^{-j\omega_i} + d_n e^{-j2\omega_i}} \quad (6-253)$$

式中  $|A_0|$  表示使  $E(\Phi, A)$  最小的最佳增益的模, 由于只需考虑幅度误差, 故最优化过程不考虑  $A_0$  的符号。

将  $|A_0|$  代入(6-249)式, 可得  $E(\Phi, A_0)$ , 它只是  $\Phi$  的函数, 可表示为

$$\hat{E}(\Phi) = E(\Phi, A_0) \quad (6-254)$$

下面来讨论(6-251)式。考虑到  $A = A_0$ , 且将(6-254)式代入, 可求得其他  $4k$  个未知数  $\Phi$  的最佳值:

$$\frac{\partial \hat{E}(\Phi)}{\partial \phi_n} = \frac{\partial E(\Phi, A_0)}{\partial \phi_n} + \frac{\partial E(\Phi, A_0)}{\partial A_0} \cdot \frac{\partial A_0}{\partial \phi_n}, \quad n = 1, 2, \dots, 4k \quad (6-255)$$

因为  $A_0$  是实常数, 故上式右端第二项为零, 因而

$$\begin{aligned} \frac{\partial \hat{E}(\Phi)}{\partial \phi_n} &= \frac{\partial E(\Phi, A_0)}{\partial \phi_n} \\ &= 2|A_0| \cdot \sum_{i=1}^M [|A_0| + |P(e^{j\omega_i}, \Phi)| + |H_d(e^{j\omega_i})|] \cdot \frac{\partial |P(e^{j\omega_i}, \Phi)|}{\partial \phi_n} \end{aligned} \quad (6-256)$$

由于  $|P(e^{j\omega_i}, \Phi)| = [P(e^{j\omega_i}, \Phi) \cdot P^*(e^{j\omega_i}, \Phi)]^{1/2}$  (6-257)

其中,  $P^*(e^{j\omega_i}, \Phi)$  表示  $P(e^{j\omega_i}, \Phi)$  的复共轭值。把(6-257)式对  $\phi_n$  求偏导数, 可得

$$\begin{aligned} \frac{\partial |P(e^{j\omega_i}, \Phi)|}{\partial \phi_n} &= \frac{1}{2} [P(e^{j\omega_i}, \Phi) \cdot P^*(e^{j\omega_i}, \Phi)]^{-\frac{1}{2}} \\ &\times \left[ P(e^{j\omega_i}, \Phi) \frac{\partial P^*(e^{j\omega_i}, \Phi)}{\partial \phi_n} + P^*(e^{j\omega_i}, \Phi) \cdot \frac{\partial P(e^{j\omega_i}, \Phi)}{\partial \phi_n} \right] \\ &= \frac{1}{|P(e^{j\omega_i}, \Phi)|} \cdot \text{Re} \left[ P^*(e^{j\omega_i}, \Phi) \cdot \frac{\partial P(e^{j\omega_i}, \Phi)}{\partial \phi_n} \right] \end{aligned} \quad (6-258)$$

其中  $\text{Re}[\ ]$  表示取实部。

我们知道,  $\phi_n$  为  $4k$  个系数  $a_n, b_n, c_n, d_n (n = 1, 2, \dots, k)$ 。把(6-253)式代入(6-258)式, 可得  $4k$  个方程为

$$\frac{\partial |P(e^{j\omega_i}, \Phi)|}{\partial a_n} = |P(e^{j\omega_i}, \Phi)| \cdot \text{Re} \left[ \frac{z_i^{-1}}{1 + a_n z_i^{-1} + b_n z_i^{-2}} \right]_{z_i = e^{j\omega_i}}$$

$$\begin{aligned}
\frac{\partial |P(e^{j\omega_i}, \Phi)|}{\partial b_n} &= |P(e^{j\omega_i}, \Phi)| \cdot \operatorname{Re} \left[ \frac{z_i^{-2}}{1 + a_n z_i^{-1} + b_n z_i^{-2}} \right]_{z_i = e^{j\omega_i}} \\
\frac{\partial |P(e^{j\omega_i}, \Phi)|}{\partial c_n} &= -|P(e^{j\omega_i}, \Phi)| \cdot \operatorname{Re} \left[ \frac{z_i^{-1}}{1 + c_n z_i^{-1} + d_n z_i^{-2}} \right]_{z_i = e^{j\omega_i}} \\
\frac{\partial |P(e^{j\omega_i}, \Phi)|}{\partial d_n} &= -|P(e^{j\omega_i}, \Phi)| \cdot \operatorname{Re} \left[ \frac{z_i^{-2}}{1 + c_n z_i^{-1} + d_n z_i^{-2}} \right]_{z_i = e^{j\omega_i}}
\end{aligned}
\quad n = 1, 2, \dots, k \quad (6-259)$$

只要把(6-259)式代入(6-256)式,就可得到均方误差函数  $\bar{E}(\Phi)$  对  $4k$  个参量  $\phi_n$  的偏导数,令其等于零,即令

$$\frac{\partial \bar{E}(\Phi)}{\partial \phi_n} = \frac{\partial E(\Phi, A_0)}{\partial \phi_n} = 0, \quad n = 1, 2, \dots, 4k \quad (6-260)$$

可得到  $4k$  个非线性方程的方程组,通常用迭代的方法求解,其中弗莱切-鲍威尔(Fletcher-Powell)的优化算法是效率比较高的一种算法,它是以最陡下降法的线性搜索为主的一种混合型算法,其主要方法如下:

设第  $j$  次迭代结果为

$$\Phi^{(j)} = (a_1^{(j)}, b_1^{(j)}, c_1^{(j)}, d_1^{(j)}, a_2^{(j)}, b_2^{(j)}, c_2^{(j)}, d_2^{(j)}, \dots, a_k^{(j)}, b_k^{(j)}, c_k^{(j)}, d_k^{(j)})^T$$

这里“ $T$ ”表示转置。而下一次迭代值  $\Phi^{(j+1)}$  等于本次迭代值  $\Phi^{(j)}$  加一个正比于  $E^{(j)}$  的梯度  $\nabla E^{(j)}$  的负值的变化量,即

$$\Phi^{(j+1)} = \Phi^{(j)} + \Delta \Phi^{(j)} = \Phi^{(j)} - \mu \cdot \nabla E^{(j)} \quad (6-261)$$

其中步长为

$$\Delta \Phi^{(j)} = -\mu \nabla E^{(j)} \quad (6-262)$$

而

$$\nabla E^{(j)} = \left[ \frac{\partial E^{(j)}}{\partial a_1}, \frac{\partial E^{(j)}}{\partial b_1}, \frac{\partial E^{(j)}}{\partial c_1}, \frac{\partial E^{(j)}}{\partial d_1}, \dots, \frac{\partial E^{(j)}}{\partial a_k}, \frac{\partial E^{(j)}}{\partial b_k}, \frac{\partial E^{(j)}}{\partial c_k}, \frac{\partial E^{(j)}}{\partial d_k} \right]^T \quad (6-263)$$

表示均方误差的梯度。由于某点的梯度方向是表示该点变化率最大的方向,因而  $-\nabla E^{(j)}$  表示  $E^{(j)}(\Phi, A_0)$  下降最快的方向,故是最陡下降法,  $\mu$  是控制稳定性和收敛性的一个参量。

随后,我们计算以下均方误  $E\left(\Phi^{(j)} - \mu \frac{\partial E^{(j)}}{\partial \Phi_n}\right)$ ,  $E\left(\Phi^{(j)} - 2\mu \frac{\partial E^{(j)}}{\partial \Phi_n}\right)$ ,  $E\left(\Phi^{(j)} - 3\mu \frac{\partial E^{(j)}}{\partial \Phi_n}\right)$ , … 直到使  $E(\Phi, A_0)$  值为最小,停止增加,此时的最佳点定义为第  $(j+1)$  次迭代的  $\Phi^{(j+1)}$  值。如此不断进行搜索,求得一组  $\Phi$  值,使  $E(\Phi, A_0)$  最小,这组  $\Phi = [a_1, b_1, c_1, d_1, \dots, a_k, b_k, c_k, d_k]^T$  即为要求的解。应当指出,这种最小化方法只涉及幅度函数。

由于对零点、极点位置没有任何限制,最优化算法的结果,得到的参量值可能相当于一个不稳定的滤波器。也就是可能有在单位圆外的极点,此时,可级联一个全通网络来将单位圆外的极点反射到单位圆内镜像位置上,即,如果有  $z = \frac{1}{r} e^{j\theta}$  ( $r < 1$ ) 的极点,则级联一个全通网络,其极点在  $r e^{j\theta}$  处,零点在  $\frac{1}{r} e^{j\theta}$  处,两网络级联后,在单位圆外  $z = \frac{1}{r} e^{j\theta}$  处的

零点、极点互相抵消, 只有单位圆内  $z=re^{j\theta}$  处有一极点, 这就好像将单位圆外的极点  $(\frac{1}{r}e^{j\theta})$  反射到单位圆内镜像处  $(re^{j\theta})$ 。这样处理后, 并不改变幅度响应特性, 但整个级联后的系统却变成了一个稳定的滤波器。

将所有单位圆外的极点反射到单位圆内后, 还可再次用此最佳化程序, 使均方误差进一步减小。如果要求滤波器是最小相位的, 可以把单位圆外的零点反射到单位圆内。

须注意, 上面所说的  $\omega$  值可以是均匀分布的, 也可以不是均匀分布的。这样, 设计起来, 就比较灵活, 视需要情况来决定不同滤波区域抽样点的间隔。

随着二阶节节数目  $k$  的增大, 实际设计出的滤波器对待求的滤波器的幅度特性的逼近程度就越好。

## 二、最小 $p$ 误差设计法

这一方法是最小均方误差法的推广, 在这里, 以频率响应误差的  $2p$  次幂的加权平均值最小化来代替上面频率响应幅度误差的二次幂的平均值最小化, 且这种方法对频率响应的幅度响应和相位响应(群延时响应)两者都适用。

把待设计的滤波器的级联型系统函数用极坐标形式表示零点、极点, 有

$$H(z) = A \prod_{n=1}^k \frac{(z - r_{0n} e^{j\theta_{0n}})(z - r_{0n} e^{-j\theta_{0n}})}{(z - r_{pn} e^{j\theta_{pn}})(z - r_{pn} e^{-j\theta_{pn}})} \quad (6-264)$$

频率响应可表示成

$$H(e^{j\omega}) = H(z) \Big|_{z=e^{j\omega}} = \left| H(e^{j\omega}) \right| e^{j\beta(e^{j\omega})} \quad (6-265)$$

它的幅度响应为

$$\left| H(e^{j\omega}) \right| = \left\{ |A| \prod_{n=1}^k \frac{\left| z - r_{0n} e^{j\theta_{0n}} \right| + \left| z - r_{0n} e^{-j\theta_{0n}} \right|}{\left| z - r_{pn} e^{j\theta_{pn}} \right| + \left| z - r_{pn} e^{-j\theta_{pn}} \right|} \right\}_{z=e^{j\omega}} \quad (6-266)$$

而它的群延时响应按(6-10)式可表示为

$$\tau(e^{j\omega}) = -\frac{d\beta(e^{j\omega})}{d\omega} = -\operatorname{Re} \left[ z \frac{dH(z)}{dz} \cdot \frac{1}{H(z)} \right]_{z=e^{j\omega}} \quad (6-267)$$

为了深入研究(6-266)式、(6-267)式的幅度响应及群延时响应, 要分别求它的每一个因子的幅度及群延时。

令  $\Phi$  表示  $4k+1$  个参量的矢量, 即

$$\Phi = (r_{01}, \theta_{01}, r_{02}, \theta_{02}, \dots, r_{0k}, \theta_{0k}, r_{p1}, \theta_{p1}, r_{p2}, \theta_{p2}, \dots, r_{pk}, \theta_{pk}, A)^T$$

则相乘的每一因子可表示成(以两零点项为例)

$$\begin{aligned} (z - r_{0n} e^{\pm j\theta_{0n}}) \Big|_{z=e^{j\omega}} &= e^{\pm j\theta_{0n}} (e^{j(\omega \mp \theta_{0n})} - r_{0n}) \\ &= e^{\pm j\theta_{0n}} [\cos(\omega \mp \theta_{0n}) - r_{0n} + j\sin(\omega \mp \theta_{0n})] \end{aligned} \quad (6-268)$$

它的幅度为

$$\left| z - r_{0n} e^{\pm j\theta_{0n}} \right|_{z=e^{j\omega}} = [1 - 2r_{0n}\cos(\omega \mp \theta_{0n}) + r_{0n}^2]^{1/2} \quad (6-269)$$

它的群延时为(利用(6-267)式, 在这里  $H(z)=z-r_{0n}e^{\pm j\theta_{0n}}$ )

$$\begin{aligned}
\tau(\Phi, \omega) &= -\frac{d}{d\omega} [\arg(z - r_{0n} e^{\pm j\theta_{0n}})]_{z=e^{j\omega}} = -\operatorname{Re}\left[\frac{e^{j\omega}}{e^{j\omega} - r_{0n} e^{\pm j\theta_{0n}}}\right] \\
&= -\operatorname{Re}\left[\frac{1}{1 - r_{0n} \cos(\omega \mp \theta_{0n}) + j r_{0n} \sin(\omega \mp \theta_{0n})}\right] \\
&= -\frac{1 - r_{0n} \cos(\omega \mp \theta_{0n})}{1 - 2r_{0n} \cos(\omega \mp \theta_{0n}) + r_{0n}^2} \quad (6-270)
\end{aligned}$$

同样,可得  $H(e^{j\omega})$  中两个极点乘因子的幅度及群延时的类似表示式。

这样,整个滤波器的幅度响应可表示成

$$\begin{aligned}
|H(e^{j\omega})| &= \alpha(\Phi, \omega) \\
&= |A| \prod_{n=1}^k \frac{[1 - 2r_{0n} \cos(\omega - \theta_{0n}) + r_{0n}^2]^{1/2} [1 - 2r_{0n} \cos(\omega + \theta_{0n}) + r_{0n}^2]^{1/2}}{[1 - 2r_{pn} \cos(\omega - \theta_{pn}) + r_{pn}^2]^{1/2} [1 - 2r_{pn} \cos(\omega + \theta_{pn}) + r_{pn}^2]^{1/2}} \quad (6-271)
\end{aligned}$$

整个滤波器的群延响应可表示为

$$\begin{aligned}
\tau(e^{j\omega}) &= -\frac{d}{d\omega} \{ \arg |H(e^{j\omega})| \} = \tau(\Phi, \omega) \\
&= \sum_{n=1}^k \left[ \frac{1 - r_{pn} \cos(\omega - \theta_{pn})}{1 - 2r_{pn} \cos(\omega - \theta_{pn}) + r_{pn}^2} + \frac{1 - r_{pn} \cos(\omega + \theta_{pn})}{1 - 2r_{pn} \cos(\omega + \theta_{pn}) + r_{pn}^2} \right. \\
&\quad \left. - \frac{1 - r_{0n} \cos(\omega - \theta_{0n})}{1 - 2r_{0n} \cos(\omega - \theta_{0n}) + r_{0n}^2} - \frac{1 - r_{0n} \cos(\omega + \theta_{0n})}{1 - 2r_{0n} \cos(\omega + \theta_{0n}) + r_{0n}^2} \right] \quad (6-272)
\end{aligned}$$

有了(6-271)式及(6-272)式,我们可以设计一个 IIR 滤波器(其幅度响应为  $\alpha(\Phi, \omega)$ , 群延时响应为  $\tau(\Phi, \omega)$ , 用它去逼近一个待求的滤波器(其幅度响应为  $\alpha_d(\omega)$ , 群延时响应为  $\tau_d(\omega)$ )所采用的  $p$  误差公式为

$$L_{2p}^*(\Phi) = \sum_{i=1}^J W_a(\omega_i) [\alpha(\Phi, \omega_i) - \alpha_d(\omega_i)]^{2p} \quad (6-273)$$

$$L_{2p}^t(\Phi) = \sum_{i=1}^J W_t(\omega_i) [\tau(\Phi, \omega_i) - \tau_d(\omega_i)]^{2p} \quad (6-274)$$

两式中  $W_a(\omega_i)$  和  $W_t(\omega_i)$  分别是幅度响应的加权函数及群延时响应的加权函数。当  $p=1$ , 且  $W_a(\omega_i)=1$ (对所有  $i$ )时, 幅度响应的  $p$  误差公式(6-273)式就是均方误差公式的(6-247)式。

由于  $L_{2p}^*(\Phi)$  和  $L_{2p}^t(\Phi)$  都是矢量  $\Phi$  的函数, 故最小  $p$  误差设计准则就是求  $L_{2p}^*(\Phi)$ ,  $L_{2p}^t(\Phi)$  相对于  $\Phi$  的最小值问题。可以证明, 如果从稳定的逼近解出发, 当  $2p \geq 2$ , 且当加权函数  $W_a(\omega)$ [或  $W_t(\omega)$ ]是正值时, 有  $L_{2p}^*(\Phi)$  或  $L_{2p}^t(\Phi)$  的局部最小值存在, 且其最佳参数对应于一个稳定滤波器的系统函数。

下面以幅度函数的最小  $p$  误差设计为例, 对设计公式推导如下:

先求

$$\frac{\partial L_{2p}^*(\Phi)}{\partial \phi_n} = 0, \quad n = 1, 2, \dots, 4k+1 \quad (6-275)$$

$\phi_n$  是矢量  $\Phi$  的每一个分量, 具体可表示为

$$\frac{\partial L_{2p}^n(\Phi)}{\partial r_{0n}} = \sum_{i=1}^j 2pW_a(\omega_i) [\alpha(\Phi, \omega_i) - \alpha_d(\omega_i)]^{2p-1} \cdot \frac{\partial \alpha(\Phi, \omega_i)}{\partial r_{0n}} = 0 \quad (6-276a)$$

$$\frac{\partial L_{2p}^n(\Phi)}{\partial \theta_{0n}} = \sum_{i=1}^j 2pW_a(\omega_i) [\alpha(\Phi, \omega_i) - \alpha_d(\omega_i)]^{2p-1} \cdot \frac{\partial \alpha(\Phi, \omega_i)}{\partial \theta_{0n}} = 0 \quad (6-276b)$$

以及

$$\frac{\partial L_{2p}^n(\Phi)}{\partial r_{pn}} = 0 \quad (6-276c)$$

$$\frac{\partial L_{2p}^n(\Phi)}{\partial \theta_{pn}} = 0 \quad (6-276d)$$

其中  $n=1, 2, \dots, k$ 。还有

$$\frac{\partial L_{2p}^n(\Phi)}{\partial |A|} = 0 \quad (6-276e)$$

从(2-276)的各式中, 需要求以下偏导数:

$$\begin{aligned} \frac{\partial \alpha(\Phi, \omega_i)}{\partial r_{0n}} &= \frac{\partial |H(e^{j\omega})|}{\partial r_{0n}} \\ &= |H(e^{j\omega})| \cdot \left[ \frac{r_{0n} - \cos(\omega_i - \theta_{0n})}{1 - 2r_{0n}\cos(\omega_i - \theta_{0n}) + r_{0n}^2} + \frac{r_{0n} - \cos(\omega_i + \theta_{0n})}{1 - 2r_{0n}\cos(\omega_i + \theta_{0n}) + r_{0n}^2} \right] \end{aligned}$$

$$\begin{aligned} \frac{\partial \alpha(\Phi, \omega_i)}{\partial \theta_{0n}} &= \frac{\partial |H(e^{j\omega})|}{\partial \theta_{0n}} \\ &= |H(e^{j\omega})| \cdot \left[ \frac{-r_{0n}\sin(\omega_i - \theta_{0n})}{1 - 2r_{0n}\cos(\omega_i - \theta_{0n}) + r_{0n}^2} + \frac{r_{0n}\sin(\omega_i + \theta_{0n})}{1 - 2r_{0n}\cos(\omega_i + \theta_{0n}) + r_{0n}^2} \right] \end{aligned}$$

另外还有  $\frac{\partial \alpha(\Phi, \omega_i)}{\partial r_{pn}}$  及  $\frac{\partial \alpha(\Phi, \omega_i)}{\partial \theta_{pn}}$  的表示式。

把这些公式代入(6-276)式的各式中, 即可得到  $4k+1$  个非线性方程。仍可用弗莱切-鲍威尔的优化算法来进行求解。

对相位响应可做同样处理。

### 三、线性规划设计法

对于一个有规定的幅度特性的 IIR 滤波器, 可以利用线性规划技术, 得到它的等波纹逼近(即最大误差最小化逼近)。如果系统函数表示为

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 + \sum_{k=1}^N a_k z^{-k}} \quad (6-277)$$

则按(6-220)式可写出

$$\left| H(e^{j\omega}) \right|^2 = \frac{\hat{N}(\omega)}{\hat{D}(\omega)} = \frac{\sum_{k=0}^M c_k \cos \omega k}{\sum_{k=0}^N d_k \cos \omega k} = \frac{c_0 + \sum_{k=1}^M c_k \cos \omega k}{d_0 + \sum_{k=1}^N d_k \cos \omega k} \quad (6-278)$$

由此式可看出,  $\hat{N}(\omega)$  和  $\hat{D}(\omega)$  对系数  $c_k, d_k$  来说是线性的, 因而可以用线性规划技术来确定各系数  $c_k, d_k$ , 以保证  $|H(e^{j\omega})|^2$  逼近一个已给定的幅度平方特性  $F(\omega) = |H_d(e^{j\omega})|^2$ , 且使逼近误差的峰值最小, 即要求是等波纹逼近。也就是说, 逼近问题就是使以下不等式成立:

$$-\epsilon(\omega) \leq \frac{\hat{N}(\omega)}{\hat{D}(\omega)} - F(\omega) \leq \epsilon(\omega) \quad (6-279)$$

其中,  $\epsilon$  是误差的容限, 一般在不同频段, 它的数值可以是不相等的。

由于  $F(\omega)$  和  $\epsilon(\omega)$  都是已知的, 故可将(6-279)式表示成  $c_k$  和  $d_k$  的一组线性不等式

$$\hat{N}(\omega) - \hat{D}(\omega)[F(\omega) + \epsilon(\omega)] \leq 0 \quad (6-280)$$

$$-\hat{N}(\omega) + \hat{D}(\omega)[F(\omega) - \epsilon(\omega)] \leq 0 \quad (6-281)$$

再加上两个附加不等式条件

$$-\hat{N}(\omega) \leq 0 \quad (6-282)$$

$$-\hat{D}(\omega) \leq 0 \quad (6-283)$$

从(6-280)式到(6-283)式就完全确定了逼近问题。用线性规划法求解这四个不等式, 可直接求出滤波器的系数  $c_k, d_k$ , 从而求得  $a_k, b_k$ 。

为了解以上四个线性不等式, 将不等式在左边减去一个辅助变量  $w$ , 若不等式仍满足, 则再减去一个  $w$  值, 直到所减的  $w$  值为零, 即不能再减了, 那么逼近问题有解, 可将得到的系数作为线性规划的输出; 如果减去  $w(w>0)$  后逼近问题没有解, 则要修改  $F(\omega)$  或  $\epsilon(\omega)$  (一般都是修改  $\epsilon(\omega)$ ), 或同时修改二者, 以便得到一个解答。

**[例 6-14]** 图 6-43(a) 表示逼近得到的幅度响应, 图 6-43(b) 表示此幅度响应的平方, 图 6-43(c) 表示所要求的(待逼近的)幅度平方响应, 图 6-43(d) 为最大逼近误差响应, 图中  $\delta$  是一个未知量, 需要加以调整, 以得到最佳值。显然,  $F(\omega) + \epsilon(\omega)$  是幅度平方响应的上限值,  $F(\omega) - \epsilon(\omega)$  是幅度平方响应的下限值。

为了保证  $0 < k\delta < 1$ , 则  $\delta$  的上限、下限应满足

$$0 < \delta < \frac{1}{k+1}$$

开始取一个在此范围内的  $\delta$  起始估计值, 就规定了  $F(\omega)$  及  $\epsilon(\omega)$ 。如果将此  $\delta$  代入后, 四个不等式组有解, 则减小  $\delta$  值再求(即减去一个  $w>0$  的值), 如果不等式组无解, 则增加  $\delta$  值再求解; 若  $\delta>0$  时无解, 则需改变  $\epsilon(\omega)$ , 即增加  $\delta$ 。如此重复地增加  $\delta$  的下限(如果解不存在)及降低  $\delta$  的上限(如果有解), 则能以理论上任意要求的准确度来得到  $\delta$  的极小值(也就是当  $\delta$  不能再减小了, 即  $w=0$  时, 就得到解答)。

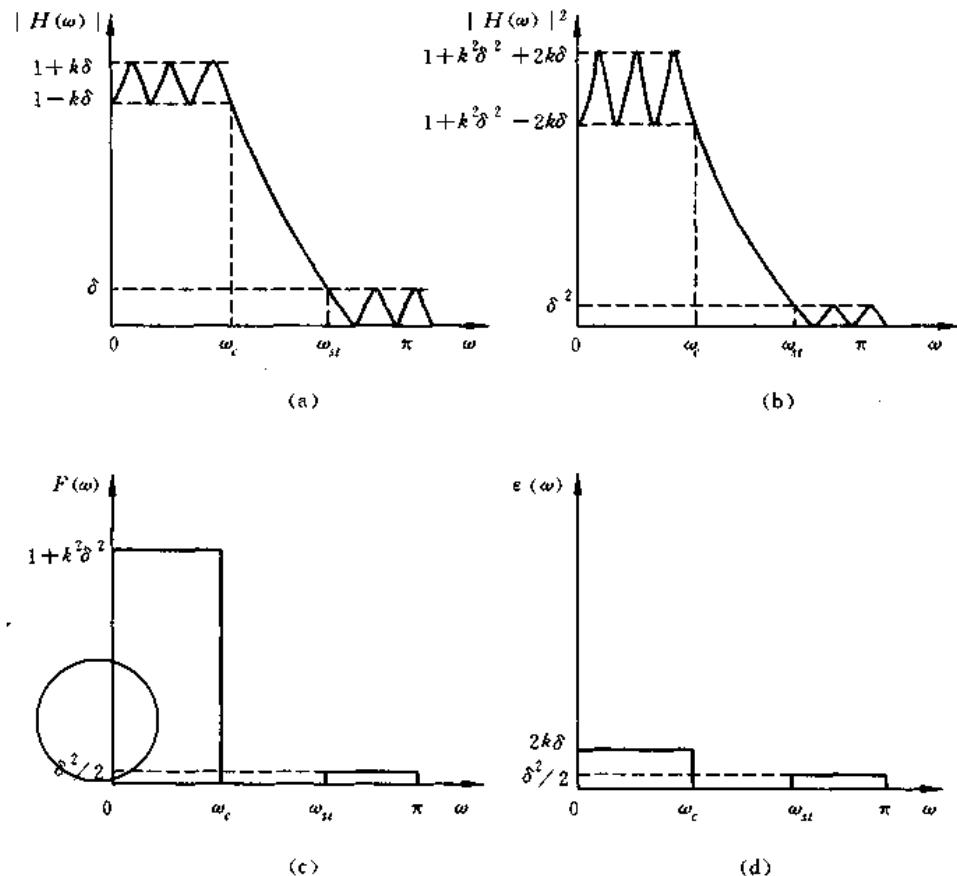


图 6-43 例 6-14 的各有关图

## 6.14 数字滤波器的实现

当数字滤波器系统函数  $H(z)$  已设计出,且由有限字长效应已选定了信号处理器的字长后,余下的问题是将  $H(z)$  所对应的线性差分方程加以实现,以求出在输入  $x(n)$  作用下,滤波器的输出值  $y(n)$ 。和 FFT 实现一样,数字滤波器也可以有两种实现办法:计算机软件实现,通用或专用硬件实现。

### 一、数字滤波器的软件实现

假定已设计出的六阶数字低通滤波器的系统函数  $H(z)$  为

$$H(z) = H_1(z) \cdot H_2(z) H_3(z) = 7.378 \times 10^{-4} \frac{1 + 2z^{-1} + z^{-2}}{1 - 1.2686z^{-1} + 0.7051z^{-2}} \\ \times \frac{1 + 2z^{-1} + z^{-2}}{1 - 1.0106z^{-1} + 0.3583z^{-2}} \cdot \frac{1 + 2z^{-1} + z^{-2}}{1 - 0.9044z^{-1} + 0.2155z^{-2}}$$

如果将此式表示成图 6-44 的框图,则可写出相应于  $H_1(z), H_2(z), H_3(z)$  的差分方程

$$\begin{aligned}
 w(n) &= 1.2686w(n-1) - 0.7051w(n-2) + x(n) + 2x(n-1) + x(n-2) \\
 v(n) &= 1.0106v(n-1) - 0.3583v(n-2) + w(n) + 2w(n-1) + w(n-2) \\
 z(n) &= 0.9044z(n-1) - 0.2155z(n-2) + v(n) + 2v(n-1) + v(n-2) \\
 y(n) &= 7.378 \times 10^{-4} z(n)
 \end{aligned}$$

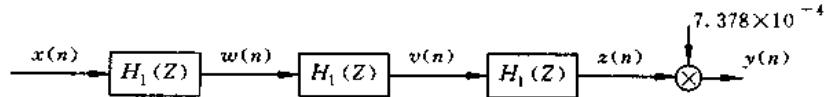


图 6-44 六阶级联滤波器的框图

此系统显然是级联系统，每节的输出就成为下一节的输入，其具体级联结构见图 6-45。按此结构，即可画出实现此数字滤波器的逻辑流程图如图 6-46 所示。按照此流程图就很容易编出程序，当给定输入  $x(n)$  时，就可得到输出  $y(n)$ 。对于并联结构或其他结构，亦可得到类似的流程图和程序，读者可自己去推求。

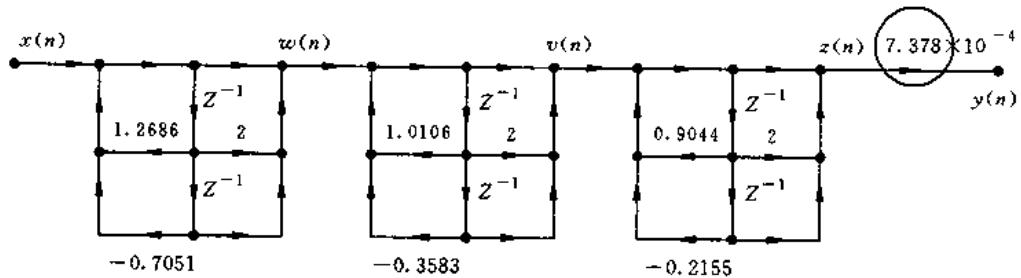


图 6-45 六阶数字滤波器的级联结构

## 二、数字滤波器的专用硬件实现

适合于实现数字滤波器的专用硬件应具有的主要特征如下：

- (1) 需要实时运算；
- (2) 需要实现某一特定的滤波运算，即运算的差分方程的系数  $a_i, b_i$  应是常数；
- (3) 一般采用定点二进制运算；
- (4) 要求所设计出的数字滤波器（它取代模拟滤波器）价格便宜，性能优良；
- (5) 要求数字滤波器是相对简单的，配套的独立数字系统。

除以上特征外，数字滤波器和其他数字信号处理器一样，都有一个公共特征，即它必须完成阵列乘法。实际实现时，是用加法器来做乘、加运算。下面来介绍这种用阵列乘法原理实现的数字滤波器专用硬件。阵列乘法可用以下形式表示：

$$A = \sum_{k=1}^I d_k B_k \quad (6-284)$$

其中  $d_k$  是滤波器的各个常系数 ( $k=1, 2, \dots, I$ )， $B_k$  是输入或输出的各抽样数据值  $x(n), x(n-1), \dots, y(n-1), y(n-2), \dots$ 。例如，以下线性差分方程所描述的滤波器就是 (6-284) 式的一种形式：

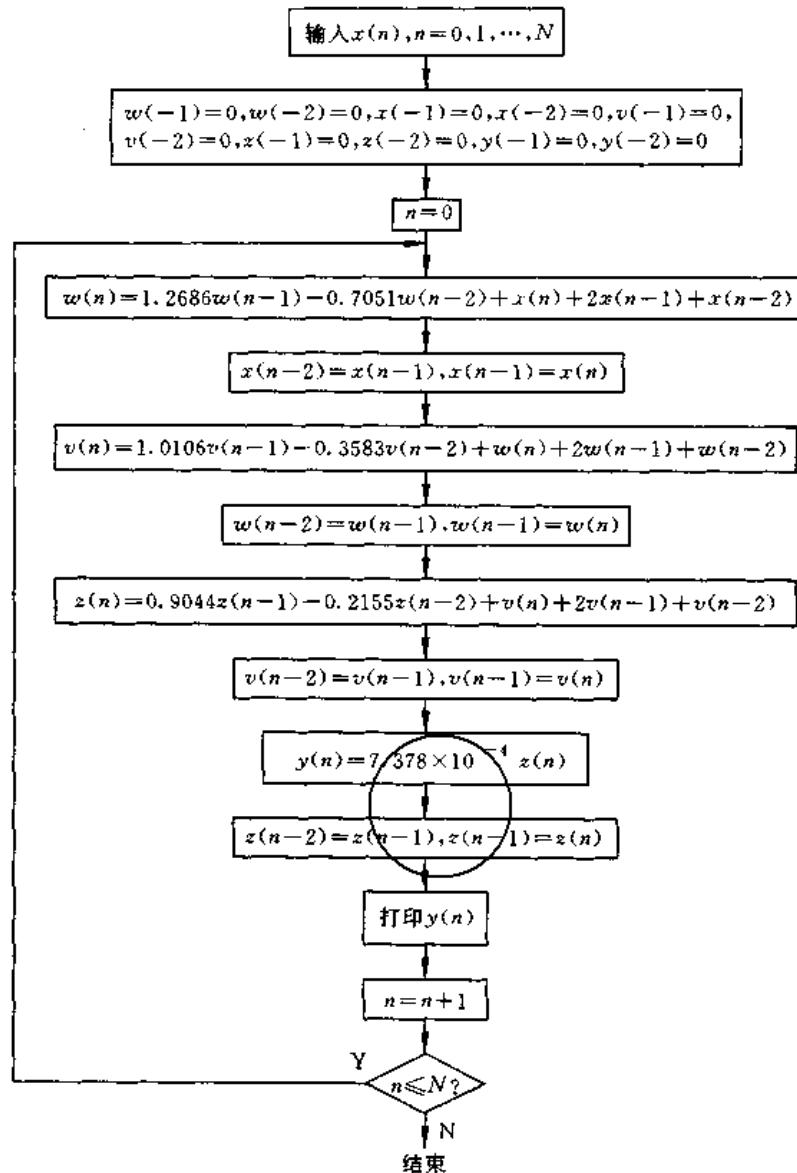


图 6-46 六阶数字滤波器级联结构的逻辑流程图

$$y(n) = 0.4123x(n) + 0.2321x(n-1) - 0.1525y(n-1) + 0.0251y(n-2)$$

我们假定滤波器的所有输入序列  $x(n), x(n-1), \dots$ , 所有输出序列  $y(n-1), y(n-2), \dots$  的幅度都被限制在  $|B_k| < 1$ 。如果用  $(b+1)$  位(含符号位)二进制补码来表示数据  $B_k$ , 即用  $c_0 \cdot c_1 c_2 \cdots c_b$  来表示, 其中  $c_i$  取 0, 1 两种可能中的一个 ( $i=0, 1, \dots, b$ ), 则  $B_k$  的十进制值(见公式(8-5))为

$$B_k = \left( -c_0 + \sum_{i=1}^b c_i 2^{-i} \right)_k$$

利用此式, 则(6-284)式给出的阵列乘法可重写为

$$A = \sum_{k=1}^l \left[ d_k \left( -c_0 + \sum_{i=1}^b c_i 2^{-i} \right)_k \right]$$

$$\begin{aligned}
&= \sum_{k=1}^I \left[ -d_k(c_0)_k + d_k \sum_{i=1}^b (c_i 2^{-i})_k \right] \\
&= - \sum_{k=1}^I d_k(c_0)_k + \left[ \sum_{k=1}^I d_k \sum_{i=1}^b 2^{-i} (c_i)_k \right] \\
&= - \sum_{k=1}^I d_k(c_0)_k + \left[ \sum_{i=1}^b 2^{-i} \sum_{k=1}^I d_k(c_i)_k \right]
\end{aligned} \tag{6-285}$$

式中,  $(c_0)_k$  表示滤波器运算中的第  $k$  个数据(即  $x(n), x(n-1), \dots, y(n-2), y(n-1)$  中的某个数据)的符号位, 而  $(c_i)_k$  ( $i = 1, 2, \dots, b$ ) 表示第  $k$  个数据的第  $i$  位二进制数, 共有  $I$  个数据 ( $k = 1, 2, \dots, I$ )。 $(c_0)_k, (c_i)_k$  二者取值不是 1 就是 0, 因此  $\sum_{k=1}^I d_k(c_0)_k$  和  $\sum_{k=1}^I d_k(c_i)_k$  实际上是受  $(c_0)_k$  和  $(c_i)_k$  所控制的系数  $d_k$  的可能组合, 很明显,  $\sum_{k=1}^I d_k(c_0)_k$  和  $\sum_{k=1}^I d_k(c_i)_k$  各自都有  $2^I$  个可能的值。表 6-10 列举了当  $I = 2$  时, 四种可能出现的值。对于第  $i$  位, 由全部  $I$  个数据  $(c_0)_1, (c_0)_2, \dots, (c_0)_I$  所决定的  $\sum_{k=1}^I d_k(c_0)_k$  是这  $2^I$  个可能值中的某一个。

表 6-10  $I=2$  的编址存储值

$(c_0)_1$ [或 $(c_i)_1$ ]	$(c_0)_2$ [或 $(c_i)_2$ ]	$\sum_{k=1}^2 d_k(c_0)_k$ [或 $\sum_{k=1}^2 d_k(c_i)_k$ ]
0	0	0
1	0	$d_1$
0	1	$d_2$
1	1	$d_1 + d_2$

上面的讨论提供了实现数字滤波器运算的一种方法, 就是将全部可能有的  $2^I$  个  $\sum_{k=1}^I d_k(c_0)_k$  和  $\sum_{k=1}^I d_k(c_i)_k$  的值预先存放在只读存储器(ROM)中,(由于  $I$  和  $d_k$  是已知的, 因而可能有的全部  $2^I$  个  $\sum_{k=1}^I d_k(c_0)_k$  和  $\sum_{k=1}^I d_k(c_i)_k$  也是已知的)。然后对全部的  $I$  个数据, 取第  $i$  位的  $(c_0)_1, (c_0)_2, \dots, (c_0)_I$  作为地址码, 从存储器中取出此第  $i$  位的编址存储值  $\sum_{k=1}^I d_k(c_0)_k$  和  $\sum_{k=1}^I d_k(c_i)_k$ 。共取  $(b+1)$  次 ( $i = 0, 1, \dots, b$ , 而  $i = 0$  表示符号位)。利用(6-285)式,  $A$  值(十进制值)的计算可简单地表示为

$$A = - \left[ \begin{array}{c} \text{编址} \\ \text{存储值} \end{array} \right] + \left[ \sum_{i=1}^{b+1} 2^{-i} \left( \begin{array}{c} \text{编址} \\ \text{存储值} \end{array} \right) \right] \tag{6-286}$$

由上面的讨论可知, 原来按(6-284)式计算  $A$  需要  $I$  次乘法和  $(I-1)$  次加法, 现在按(6-286)式计算  $A$  并不需要乘法器, 只要做  $b$  次加法就可完成。这种方法的缺点是, 当阵

列乘法求和项数  $I$  较大时,  $2^I$  很大, 因而只读存储器容量需要很大。

### IIR 数字滤波器的硬件实现——位并行方式实现

位并行运算系统与“位串行方式”(主要限于用在 FIR 滤波器上)相比较, 可提供更快的运算速度, 而其代价则是增加硬件量。因而, 位并行运算系统适合于用来实现简单的一阶 IIR 数字滤波器, 现举例来加以说明。

[例 6-15] IIR 数字滤波器的线性差分方程为

$$y(n) = 0.75x(n) - 0.5y(n-1) \quad (6-287)$$

假定

$$\begin{aligned} y(-1) &= 0 \\ x(n) &= 0, \quad \text{当 } n < 0 \\ x(n) &= 0.5, \quad \text{当 } n \geq 0 \end{aligned}$$

滤波器字长为  $b+1=6$  位, 采用位并行实现, 并说明系统如何计算出滤波器的输出值  $y(2)$ 。

解 利用已给的线性差分方程, 可算出

$$\begin{aligned} y(0) &= 0.75 \times 0.5 - 0.5 \times 0 = 0.375 \\ y(1) &= 0.75 \times 0.5 - 0.5 \times 0.375 = 0.1875 \\ y(2) &= 0.75 \times 0.5 - 0.5 \times 0.1875 = 0.28125 \end{aligned}$$

为了用并行运算方式算出  $A$ (即  $y(n)$ ), 可以采用加法树, 把保持在  $(b+1)$  个 ROM 中的编址存储值同时相加, 如图 6-47 所示。

根据表 6-10, 其基本 ROM 图如下:

ROM 地址		ROM 内容
$A_1$	$A_2$	
0	0	0.00000
1	0	0.11000
0	1	1.10000
1	1	0.01000

按照(6-286)式, 采用补码表示法, 由  $x(n)$  和  $y(n-1)$  的符号位规定的编址 ROM1 中, 存储着负的编址存储值, 即基本 ROM 编址存储值的 2 的补码; 由  $x(n)$  和  $y(n-1)$  的最高有效位(即小数点后的第一位)规定的编址 ROM2 中, 存储着基本 ROM 编址存储值乘  $2^{-1}$  后的值; 而由  $x(n)$  和  $y(n-1)$  的次高位规定的编址 ROM3 中, 存储着基本 ROM 编址存储值乘以  $2^{-2}$  后的值; 依此类推, 当沿 SR1, SR2 逐位右移时, 由  $x(n)$  和  $y(n-1)$  的第  $m$  高有效位规定的编址 ROM $m$  中, 存储着基本 ROM 编址存储值乘  $2^{-m}$  后的值( $1 \leq m \leq b$ )。

参照图 6-47, 计算过程如下, 最后得到的  $A$  值即为  $y(2)$ :

ROM1 输出 = 0.00000

ROM2 输出 = 0.01100

ROM3 输出 = 0.00000

ROM4 输出 = 1.11110

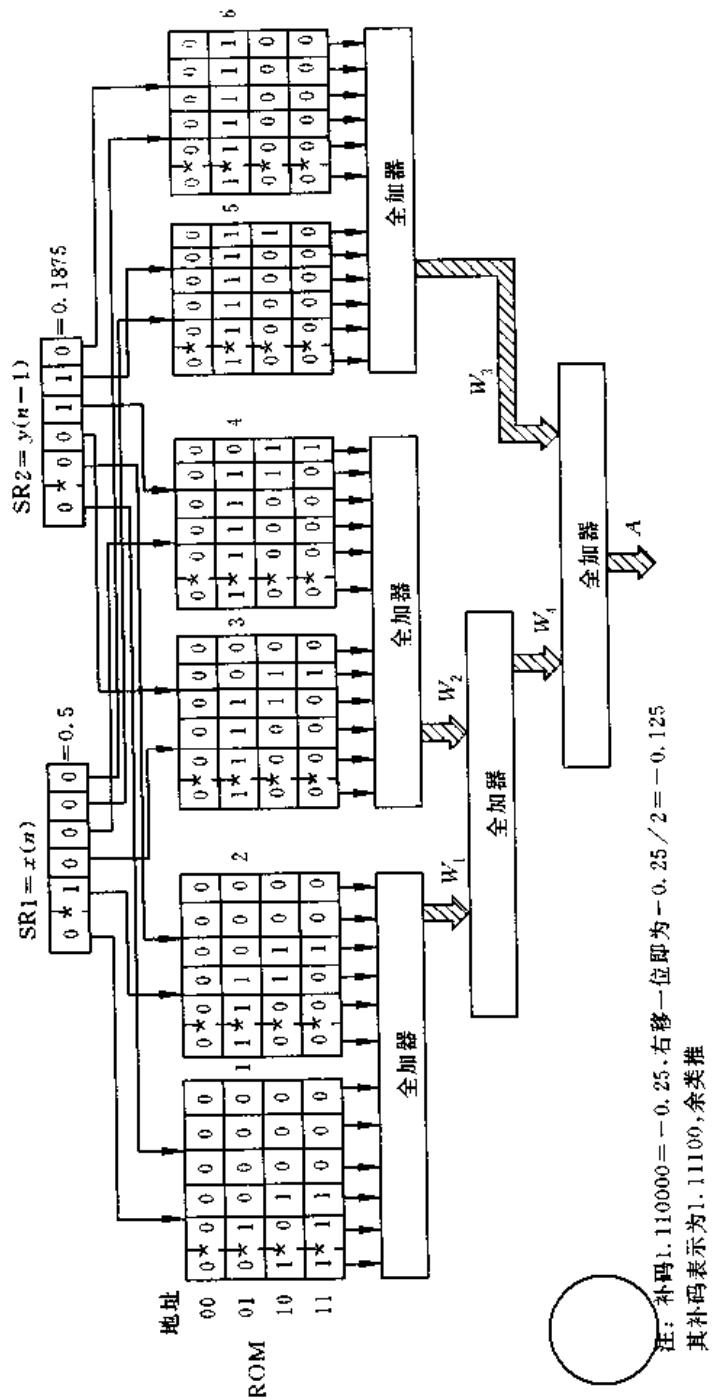


图6-47 位并行运算系统的框图

ROM5 输出 = 1.1111

ROM6 输出 = 0.00000

全加器输出  $w_1 = \text{ROM1} \text{ 输出} + \text{ROM2} \text{ 输出} = 0.01100$

全加器输出  $w_2 = \text{ROM3} \text{ 输出} + \text{ROM4} \text{ 输出} = 1.11110$

全加器输出  $w_3 = \text{ROM5} \text{ 输出} + \text{ROM6} \text{ 输出} = 1.11111$

全加器输出  $w_4 = w_1 + w_2 = 1 : 0.01010$

全加器输出  $A = w_3 + w_4 = 1 : 0.01001 = 0.28125$

很明显,位并行方式实现的处理速度快,而位串行方式实现(在第七章中讨论)的处理速度慢。一般来说,增加处理速度的代价是相应地增加硬件,因此,滤波器的设计者应在硬件的成本和滤波器速度之间作出适当的折衷选择,在满足处理速度要求下,使实现的成本最经济。

要知道,实际的系统比上面讨论的例子要复杂得多。上面我们只讨论了直接型结构的实现方法,并联型结构与级联型结构的基本单元可以用上面讨论的方法来实现。

[附] 例 6-8 的设计结果(采用本书所附 DSP-MCAI-2.0 光盘的“辅助设计”子系统)

滤波器类型: 带通滤波器

输入参数: 通带截止频率:  $0.4\pi$   $0.5\pi$

通带波动损耗: 3dB

阻带截止频率:  $0.2\pi$   $0.8\pi$

阻带波动损耗: 20dB

所用模拟原型低通滤波器: 巴特沃思模拟原型低通滤波器

输出参数:

滤波器阶数: 4 阶

滤波器系数( $a(i)$ 为分子, $b(i)$ 为分母,按  $z^{-1}$  的升幂排列):

$i$	$a(i)$	$b(i)$
0	2.0250690E-02	1.0000000E+00
1	0.0000000E+00	-5.6369680E-01
2	-4.0501380E-02	1.6392940E+00
3	0.0000000E+00	-4.4967430E-01
4	2.0250690E-02	6.4004520E-01

[附] 例 6-9 的设计结果(采用本书所附 DSP-MCAI-2.0 光盘的“辅助设计”子系统)

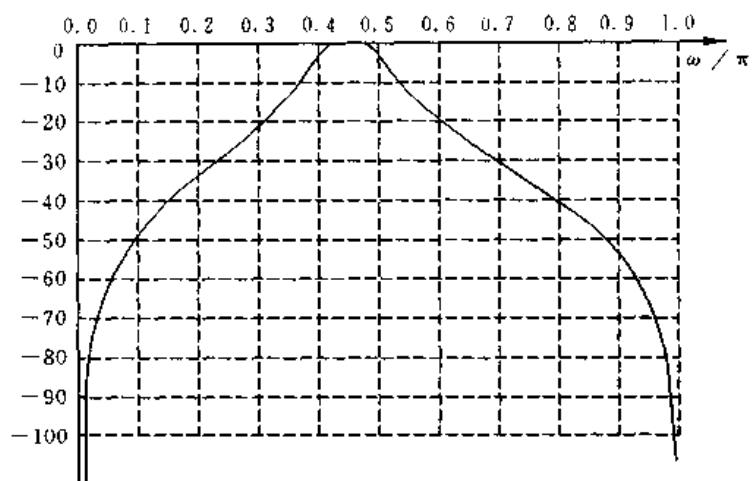
滤波器类型: 带阻滤波器

输入参数: 通带截止频率:  $0.2\pi$   $0.7\pi$

通带波动损耗: 3dB

阻带截止频率:  $0.36\pi$   $0.5\pi$

阻带波动损耗: 14dB



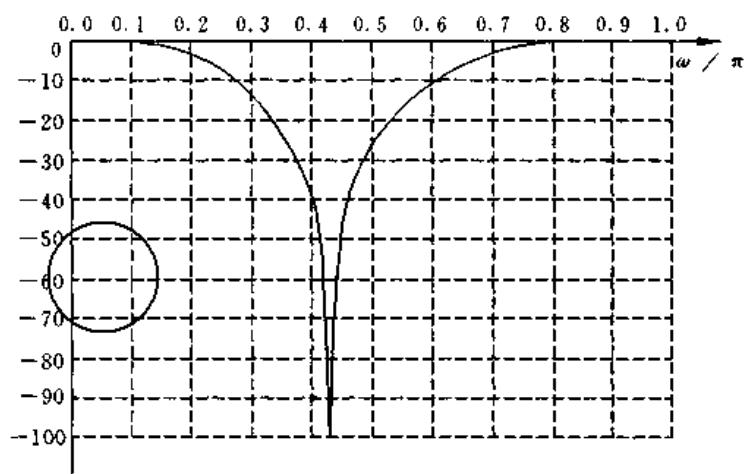
所用模拟原型低通滤波器：巴特沃思模拟原型低通滤波器

输出参数：

滤波器阶数：4 阶

滤波器系数( $a(i)$ 为分子,  $b(i)$ 为分母, 按  $z^{-1}$  的升幂排列)：

$i$	$a(i)$	$b(i)$
0	2.9425980E-01	1.0000000E+00
1	-2.6039840E-01	-4.4367170E-01
2	6.4612790E-01	6.3069380E-02
3	-2.6039840E-01	-7.7125210E-02
4	2.9425980E-01	1.7157820E-01



[附] 例 6-10 的设计结果(采用本书所附 DSP-MCAI-2.0 光盘的“辅助设计”子系统)

滤波器类型：高通滤波器

输入参数：通带截止频率： $0.6\pi$

通带波动损耗: 3dB  
 阻带截止频率:  $0.4\pi$   
 阻带波动损耗: 14dB

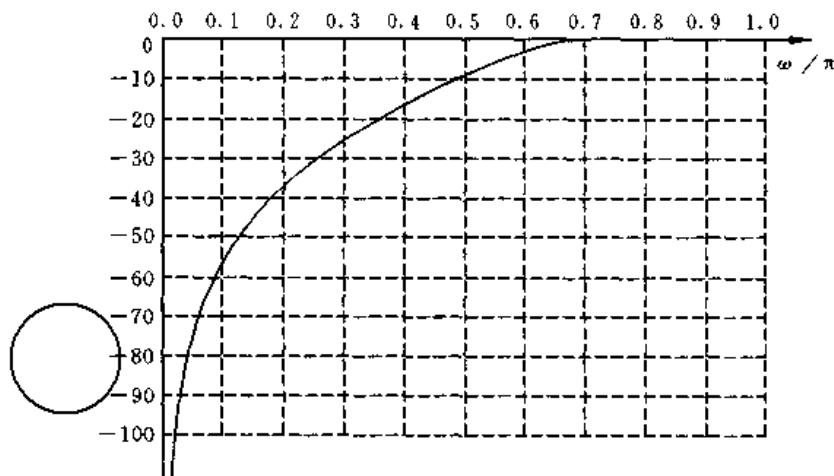
所用模拟原型低通滤波器: 巴特沃思模拟原型低通滤波器

输出参数:

滤波器阶数: 3 阶

滤波器系数( $a(i)$ 为分子,  $b(i)$ 为分母, 按  $z^{-1}$  的升幂排列):

$i$	$a(i)$	$b(i)$
0	9.9079840 E-02	1.0000000 E+00
1	-2.9723950 E-01	5.7178480 E-01
2	2.9723950 E-01	4.2011670 E-01
3	-9.9079840 E-02	5.5693250 E-02



### 习题

1. 用冲激响应不变法将以下  $H_a(s)$  变换为  $H(z)$ , 抽样周期为  $T$ 。

$$(1) H_a(s) = (s+a)/[(s+a)^2 + b^2]。$$

$$(2) H_a(s) = A/(s-s_0)^{n_0}, n_0 \text{ 为任意正整数。}$$

2. 已知模拟二阶巴特沃思低通滤波器的归一化系统函数为

$$H'_a(s) = 1/(1 + 1.4142136s + s^2)$$

而 3dB 截止频率为 50Hz 的模拟滤波器, 需将归一化的  $H'(s)$  中的  $s$  变量用  $\frac{s}{2\pi \times 50}$  来代替

$$H_a(s) = H'_a \left( \frac{s}{100\pi} \right) = \frac{9.8696044 \times 10^4}{s^2 + 444.28830s + 9.8696044 \times 10^4}$$

设系统抽样频率为  $f_s = 500\text{Hz}$ , 要求从这一低通模拟滤波器设计一个低通数字滤波器, 采用阶跃响应不变法。

3. 设有一模拟滤波器

$$H_a(s) = 1/(s^2 + s + 1)$$

抽样周期  $T=2$ , 试用双线性变换法将它转变为数字系统函数  $H(z)$ 。

4. 要求从二阶巴特沃思模拟滤波器用双线性变换导出一低通数字滤波器, 已知 3dB 截止频率为 100Hz, 系统抽样频率为 1kHz。

5. 试导出二阶巴特沃思低通滤波器的系统函数。设  $\Omega_c = 3\text{rad/s}$ 。

6. 试导出二阶切贝雪夫低通滤波器的系统函数。已知通带波纹为 2dB, 归一化截止频率为  $\Omega_c = 2 \text{ rad/s}$ (试用不同于书本的解法解答)。

7. 已知模拟滤波器有低通、高通、带通、带阻等类型, 而实际应用中的数字滤波器有低通、高通、带通、带阻等类型。则设计各类型数字滤波器可以有哪些方法? 试画出这些方法的结构表示图并注明其变换方法。

8. 某一低通滤波器的各种指标和参数要求如下:

- (1) 巴特沃思频率响应, 采用双线性变换法设计;
- (2) 当  $0 \leq f \leq 2.5\text{Hz}$  时, 衰减小于 3dB;
- (3) 当  $f \geq 50\text{Hz}$  时, 衰减大于或等于 40dB;
- (4) 抽样频率  $f_s = 200\text{Hz}$ 。

试确定系统函数  $H(z)$ , 并求每级阶数不超过二阶的级联系统函数。

9. 用双线性变换法设计一个三阶巴特沃思数字带通滤波器, 抽样频率为  $f_s = 500\text{Hz}$ , 上、下边带截止频率分别为  $f_2 = 150\text{Hz}$ ,  $f_1 = 30\text{Hz}$ 。

10. 要设计一个二阶巴特沃思带阻数字滤波器, 其阻带 3dB 的边带频率分别为 40kHz, 20kHz, 抽样频率  $f_s = 200\text{kHz}$ 。

11. 用双线性变换法设计一个三阶切贝雪夫数字高通滤波器, 抽样频率为  $f_s = 8\text{kHz}$ , 截止频率为  $f_c = 2\text{kHz}$ (不计 4kHz 以上的频率分量)。

12. 试导出从低通数字滤波器变为高通数字滤波器的设计公式。

13. 试导出从低通数字滤波器变为带通数字滤波器的设计公式。

14. 试导出从低通数字滤波器变为带阻数字滤波器的设计公式。

15. 令  $h_a(t), s_a(t)$  和  $H_a(s)$  分别表示一个时域连续的线性时不变滤波器的单位冲激响应, 单位阶跃响应和系统函数。令  $h(n), s(n)$  和  $H(z)$  分别表示时域离散线性移不变数字滤波器的单位抽样响应, 单位阶跃响应和系统函数。

(1) 如果  $h(n) = h_a(nT)$ , 是否  $s(n) = \sum_{k=-\infty}^{\infty} h_a(kT)$ ?

(2) 如果  $s(n) = s_a(nT)$ , 是否  $h(n) = h_a(nT)$ ?

16. 假设  $H_a(s)$  在  $s = s_0$  处有一个  $r$  阶极点, 则  $H_a(s)$  可以表示成

$$H_a(s) = \sum_{k=1}^r \frac{A_k}{(s - s_0)^k} + G_a(s)$$

式中  $G_a(s)$  只有一阶极点。

(1) 写出由  $H_a(s)$  计算常数  $A_k$  的公式。

(2) 求出用  $s_0$  及  $g_a(t)$  ( $G_a(s)$  的拉普拉斯反变换) 表示的冲激响应  $h_a(t)$  的表示式。

(3) 假设我们定义  $h(n) = h_a(nT)$  为某一数字滤波器的单位冲激响应。试利用(2)的结果写出系统函数  $H(z)$  的表示式。

(4) 导出直接从  $H_a(s)$  得到  $H(z)$  的方法。

17. 图 P6-17 表示一个数字滤波器的频率响应。

- (1) 用冲激响应不变法, 试求原型模拟滤波器频率响应。
- (2) 当采用双线性变换法时, 试求原型模拟滤波器频率响应。

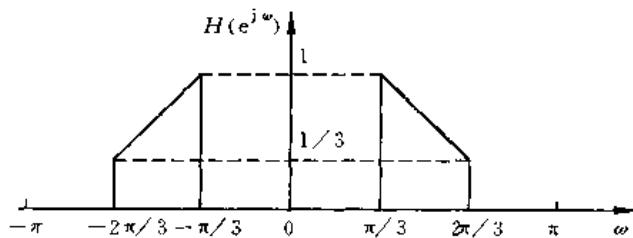


图 P6-17

18. 需设计一个数字低通滤波器, 通带内幅度特性在  $\omega \leq 0.3\pi$  处衰减在 0.75dB 内, 阻带在  $\omega = 0.5\pi$  到  $\pi$  之间的频率上衰减至少为 25dB。采用冲激响应不变法及双线性变换法, 试确定“样本”模拟系统函数及其极点, 并求所得到的数字滤波器的系统函数(设抽样周期  $T=1$ )。

19. 任何一个非最小相位系统均可表示成一个最小相位系统与一个全通系统的级联, 即

$$H(z) = H_{\min}(z)H_{ap}(z)$$

其中  $H_{ap}(z)$  是稳定的因果的全通滤波器,  $H_{\min}(z)$  是最小相位系统

令

$$\Phi(\omega) = \arg[H(e^{j\omega})]$$

$$\Phi_{\min}(\omega) = \arg[H_{\min}(e^{j\omega})]$$

试证明对于所有  $\omega$ , 有

$$-\frac{d\Phi(\omega)}{d\omega} > -\frac{d\Phi_{\min}(\omega)}{d\omega}$$

此不等式说明, 最小相位系统具有最小群延时, 所以也是最小时延系统。

20. 设  $h_{\min}(n)$  是最小相位序列,  $h(n)$  是非最小相位的因果序列,

且满足  $|H_{\min}(e^{j\omega})| = |H(e^{j\omega})|$ , 试证明  $|h(0)| < |h_{\min}(0)|$ 。

提示: 利用初值定理。

21. 如果最小相位延时的有限长序列为  $h_{\min}(n)$ ,  $0 \leq n \leq N-1$ 。

试求最大相位延时序列的  $h(n)$  为  $h(n) = h_{\min}(N-1-n)$ 。

22. 需设计一个数字切贝雪夫带通滤波器, 给定指标为

- (1) 波纹  $\delta_1 \leq 2\text{dB}$ , 当  $200\text{Hz} \leq f \leq 400\text{Hz}$ ,
- (2) 衰减  $\delta_2 \geq 20\text{dB}$ , 当  $f \leq 100\text{Hz}$ ,  $f \geq 600\text{Hz}$ ,
- (3) 抽样频率  $f_s = 2\text{kHz}$

试用: ① 冲激响应不变法; ② 双线性变换法进行设计, 最后写出  $H(z)$  的表达式, 并画出系统的幅频响应特性。

23. 需设计一个数字巴特沃思高通滤波器, 给定指标为

- (1) 衰减  $\delta_2 \geq 30\text{dB}$ , 当  $f \leq 3\text{kHz}$ ,
- (2) 衰减  $\delta_1 \leq 3\text{dB}$ , 当  $f \geq 5\text{kHz}$ ,
- (3) 抽样频率  $f_s = 20\text{kHz}$

试用: 双线性变换法进行设计, 最后写出  $H(z)$  的表达式, 并画出系统的幅频响应特性。

24. 需设计一个数字切贝雪夫带阻滤波器, 给定指标为

- (1) 衰减  $\delta_2 \geq 18\text{dB}$ , 当  $1\text{kHz} \leq f \leq 2\text{kHz}$ 。  
 (2) 波纹  $\delta_1 \leq 3\text{dB}$ , 当  $f \leq 500\text{Hz}, f \geq 3\text{kHz}$ 。  
 (3) 抽样频率  $f_s = 10\text{kHz}$ .

试用: 双线性变换法进行设计, 最后写出  $H(z)$  的表达式, 并画出系统的幅频响应特性。

25. 需设计一个数字切贝雪夫带阻滤波器, 给定指标为

- (1) 波纹  $\delta_2 \geq 18\text{dB}$ , 当  $10\text{kHz} \leq f \leq 20\text{kHz}$ 。  
 (2) 衰减  $\delta_1 \leq 3\text{dB}$ , 当  $f \leq 5\text{kHz}, f \geq 30\text{kHz}$ 。  
 (3) 抽样频率  $f_s = 100\text{kHz}$ .

试用: 双线性变换法进行设计, 最后写出  $H(z)$  的表达式, 并画出系统的幅频响应特性。

设计时请先想一想, 这一题和上一题有什么相似处, 由此应该得出什么结论。

26. 如图 P6-26 所示系统

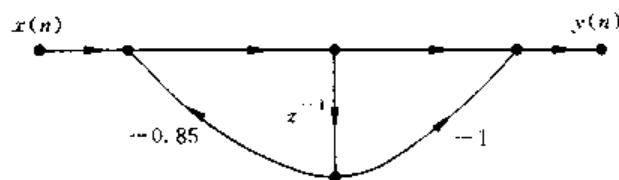


图 P6-26

- (1) 写出该系统的系统函数  $H(z)$ , 画出系统的幅频特性, 并问这一系统是哪一种通带滤波器?  
 (2) 在上述系统中, 用下列差分方程表示的网络代替它的  $z^{-1}$  延时单元

$$y(n) = x(n-2) - a[x(n-1) - y(n-1)], \quad 0 < a < 1$$

问变换后的数字网络是哪一种通带滤波器?

# 第七章 有限长单位冲激响应(FIR)数字滤波器的设计方法

## 7.1 引言

无限长单位冲激响应(IIR)数字滤波器的优点是可以利用模拟滤波器设计的结果,而模拟滤波器的设计有大量图表可查,方便简单。但是它也有明显的缺点,就是相位的非线性;若需线性相位,则要采用全通网络进行相位校正。我们知道,图像处理以及数据传输都要求信道具有线性相位特性。而有限长单位冲激响应(FIR)数字滤波器就可以做成具有严格的线性相位,同时又可以具有任意的幅度特性。此外,FIR滤波器的单位抽样响应是有限长的,因而滤波器一定是稳定的。再有,只要经过一定的延时,任何非因果有限长序列都能变成因果的有限长序列,因而总能用因果系统来实现。最后,FIR滤波器由于单位冲激响应是有限长的,因而可以用快速傅里叶变换(FFT)算法来实现过滤信号,从而可大大提高运算效率。但是,要取得很好的衰减特性,FIR滤波器  $H(z)$  的阶次比 IIR 滤波器的要高。

IIR 滤波器设计中的各种变换法对 FIR 滤波器设计是不适用的,这是因为那里是利用有理分式的系统函数,而 FIR 滤波器的系统函数只是  $z^{-1}$  的多项式。

从以上讨论看出,我们最感兴趣的是具有线性相位的 FIR 滤波器。对非线性相位的 FIR 滤波器,一般可以用 IIR 滤波器来代替,同样幅度特性,IIR 滤波器所需阶数比 FIR 滤波器的阶数要少得多,因此这里不去讨论它。

## 7.2 线性相位 FIR 滤波器的特点

FIR 滤波器的单位冲激响应  $h(n)$  是有限长的( $0 \leq n \leq N-1$ ),其  $z$  变换为

$$H(z) = \sum_{n=0}^{N-1} h(n) z^{-n}$$

这是  $z^{-1}$  的( $N-1$ )阶多项式,在有限  $z$  平面( $0 < |z| < \infty$ )有( $N-1$ )个零点,而位于  $z$  平面原点  $z=0$  处,则有( $N-1$ )阶极点。

### 一、线性相位条件

$h(n)$  的频率响应  $H(e^{j\omega})$  为

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n) e^{-j\omega n} \quad (7-1)$$

当  $h(n)$  为实序列时,可将  $H(e^{j\omega})$  表示成

$$H(e^{j\omega}) = \pm |H(e^{j\omega})| e^{j\theta(\omega)} = H(\omega) e^{j\theta(\omega)} \quad (7-2)$$

其中  $|H(e^{j\omega})|$  是真正的幅度响应,而  $H(\omega)$  是可正可负的实函数,有两类准确的线性相

位,分别要求满足

$$\theta(\omega) = -\tau\omega \quad (7-3a)$$

$$\theta(\omega) = \beta - \tau\omega \quad (7-3b)$$

其中  $\tau, \beta$  都是常数,表示相位是通过坐标原点  $\omega=0$  或是通过  $\theta(0)=\beta$  的斜直线,二者的群延时都是常数  $\tau = -\frac{d\theta(\omega)}{d\omega}$ 。

把(7-3a)式和(7-3b)式的关系分别代入(7-2)式中,并考虑(7-1)式,可得

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n} = \pm |H(e^{j\omega})| e^{-j\omega\tau} \quad (7-4)$$

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n} = \pm |H(e^{-j\omega})| e^{-j(\omega-\beta)} \quad (7-5)$$

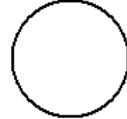
令(7-4)式等式两端实部虚部分别相等,可得到对(7-3a)式的一类线性相位必须要求

$$\pm |H(e^{j\omega})| \cos(\omega\tau) = \sum_{n=0}^{N-1} h(n) \cos(\omega n)$$

$$\pm |H(e^{j\omega})| \sin(\omega\tau) = \sum_{n=0}^{N-1} h(n) \sin(\omega n)$$

两式相除,可得

$$\tan(\omega\tau) = \frac{\sin(\omega\tau)}{\cos(\omega\tau)} = \frac{\sum_{n=0}^{N-1} h(n) \sin(\omega n)}{\sum_{n=0}^{N-1} h(n) \cos(\omega n)}$$



因而

$$\sum_{n=0}^{N-1} h(n) \sin(\omega\tau) \cos(\omega n) - \sum_{n=0}^{N-1} h(n) \cos(\omega\tau) \sin(\omega n) = 0$$

$$\text{即 } \sum_{n=0}^{N-1} h(n) \sin[(\tau - n)\omega] = 0 \quad (7-6)$$

要使(7-6)式成立,必须满足

$$\tau = \frac{N-1}{2} \quad (7-7)$$

$$h(n) = h(N-1-n), \quad 0 \leq n \leq N-1 \quad (7-8)$$

(7-8)式是 FIR 滤波器具有(7-3a)式的线性相位的必要且充分条件,它要求单位冲激响应的  $h(n)$  序列以  $n=(N-1)/2$  为偶对称中心,此时时间延时  $\tau$  等于  $h(n)$  长度  $N-1$  的一半,即为  $\tau=(N-1)/2$  个抽样周期。 $N$  为偶数时,延时为整数; $N$  为奇数时,延时为整数加半个抽样周期。不管  $N$  为奇偶,此时  $h(n)$  都应满足对  $n=(N-1)/2$  轴呈偶对称。

对(7-3b)式的另一类线性相位,将(7-5)式作同样推导可知,必须要求

$$\sum_{n=0}^{N-1} h(n) \sin[(\tau - n)\omega - \beta_0] = 0 \quad (7-9)$$

要使(7-9)式成立,必须满足

$$\tau = \frac{N-1}{2} \quad (7-10)$$

$$\beta = \pm \frac{\pi}{2} \quad (7-11)$$

$$h(n) = -h(N-1-n), \quad 0 \leq n \leq N-1 \quad (7-12)$$

(7-12)式是 FIR 滤波器具有(7-3b)式的线性相位的必要且充分条件, 它要求单位冲激响应序列  $h(n)$  以  $n=(N-1)/2$  为奇对称中心, 此时延时  $\tau$  等于  $(N-1)/2$  个抽样周期, 在  $h(n)$  的这种奇对称情况下, 满足  $h\left(\frac{N-1}{2}\right) = -h\left(\frac{N-1}{2}\right)$ , 因而  $h\left(\frac{N-1}{2}\right) = 0$ 。这种线性相位情况和前一种不同之处是, 除了产生线性相位外, 还有  $\pm\pi/2$  的固定相移。

由于  $h(n)$  有上述奇对称和偶对称两种, 而  $h(n)$  的点数  $N$  又有奇数、偶数两种情况, 因而  $h(n)$  可以有 4 种类型, 如图 7-1 和图 7-2 所示, 分别对应于 4 种线性相位 FIR 数字滤波器。

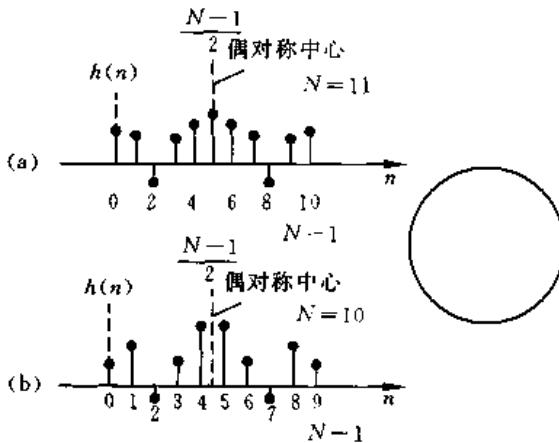


图 7-1  $h(n)$  偶对称  
(a)  $N$  为奇数; (b)  $N$  为偶数

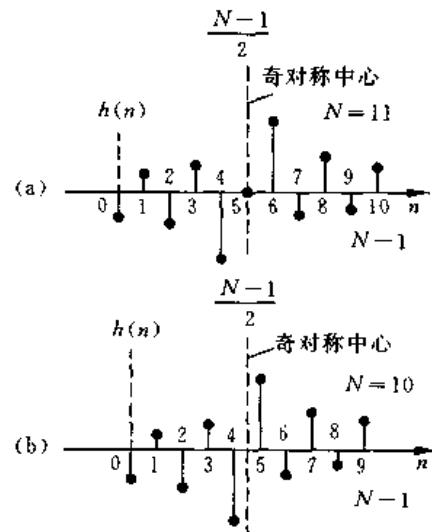


图 7-2  $h(n)$  奇对称  
(a)  $N$  为奇数; (b)  $N$  为偶数

## 二、线性相位 FIR 滤波器频率响应的特点

在以下的讨论中, 按(7-2)式把频率响应  $H(e^{j\omega})$  表示成

$$H(e^{j\omega}) = H(\omega)e^{j\theta(\omega)} \quad (7-13)$$

其中  $H(\omega)$  是幅度函数, 它是一个纯实数, 可以包括正值和负值, 即  $H(\omega) = \pm |H(e^{j\omega})|$ ,  $\theta(\omega)$  是相位函数。

已经知道, 线性相位 FIR 滤波器的冲激响应应该满足(7-8)式和(7-12)式, 即

$$h(n) = \pm h(N-1-n)$$

因而系统函数可表示为

$$\begin{aligned} H(z) &= \sum_{n=0}^{N-1} h(n)z^{-n} = \sum_{n=0}^{N-1} \pm h(N-1-n)z^{-n} \\ &= \sum_{m=0}^{N-1} \pm h(m)z^{-(N-1-m)} = \pm z^{-(N-1)} \sum_{m=0}^{N-1} h(m)z^m \end{aligned}$$

即

$$H(z) = \pm z^{-(N-1)} H(z^{-1}) \quad (7-14)$$

进一步写成

$$\begin{aligned} H(z) &= \frac{1}{2} [H(z) \pm z^{-(N-1)} H(z^{-1})] = \frac{1}{2} \sum_{n=0}^{N-1} h(n) [z^{-n} \pm z^{-(N-1-n)} z^n] \\ &= z^{-(\frac{N-1}{2})} \sum_{n=0}^{N-1} h(n) \left[ \frac{z^{(\frac{N-1}{2}-n)} \pm z^{-(\frac{N-1}{2}-n)}}{2} \right] \end{aligned} \quad (7-15)$$

在这一公式中,方括号内有“±”号。当取“+”号时,  $h(n)$  满足  $h(n) = h(N-1-n)$  偶对称; 当取“-”号时,  $h(n)$  满足  $h(n) = -h(N-1-n)$  奇对称。下面对应这两种情况分别讨论它们的频率响应。

### 1. $h(n)$ 偶对称

由(7-15)式可知, 频率响应为

$$H(e^{j\omega}) = H(z) |_{z=e^{j\omega}} = e^{-j(\frac{N-1}{2})\omega} \sum_{n=0}^{N-1} h(n) \cos \left[ \left( \frac{N-1}{2} - n \right) \omega \right] \quad (7-16)$$

将此式与(7-13)式对比, 可得幅度函数为

$$H(\omega) = \sum_{n=0}^{N-1} h(n) \cos \left[ \left( \frac{N-1}{2} - n \right) \omega \right] \quad (7-17)$$

相位函数为

$$\theta(\omega) = - \left( \frac{N-1}{2} \right) \omega \quad (7-18)$$

幅度函数  $H(\omega)$  可为正值或负值, 相位函数  $\theta(\omega)$  是严格的线性相位, 如图 7-3 所示, 可以看出,  $h(n)$  对  $n=(N-1)/2$  偶对称时, FIR 滤波器是具有准确的线性相位的滤波器。说明滤波器有  $(N-1)/2$  个抽样的延时, 它等于单位冲激响应  $h(n)$  的长度的一半。

### 2. $h(n)$ 奇对称

此时, 由(7-15)式可知, 频率响应为

$$\begin{aligned} H(e^{j\omega}) &= H(z) |_{z=e^{j\omega}} = j e^{-j(\frac{N-1}{2})\omega} \sum_{n=0}^{N-1} h(n) \sin \left[ \left( \frac{N-1}{2} - n \right) \omega \right] \\ &= e^{-j(\frac{N-1}{2})\omega + j\frac{\pi}{2}} \sum_{n=0}^{N-1} h(n) \sin \left[ \left( \frac{N-1}{2} - n \right) \omega \right] \end{aligned} \quad (7-19)$$

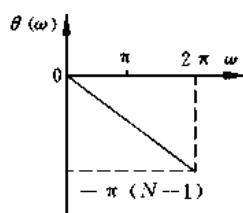


图 7-3  $h(n)$  偶对称时的  
线性相位特性

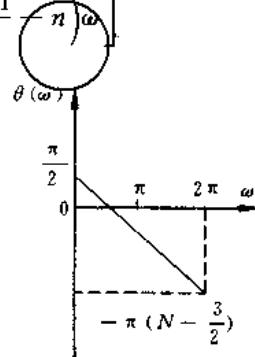


图 7-4  $h(n)$  奇对称时的 90°相移  
线性相位特性

将此式与(7-13)式比较,可得幅度函数为

$$H(\omega) = \sum_{n=0}^{N-1} h(n) \sin\left[\left(\frac{N-1}{2} - n\right)\omega\right] \quad (7-20)$$

相位函数为

$$\theta(\omega) = -\left(\frac{N-1}{2}\right)\omega + \frac{\pi}{2} \quad (7-21)$$

幅度函数可正可负,相位函数既是线性相位的,又包括 $\pi/2$ 的相移,如图7-4所示,看出不仅有 $(N-1)/2$ 个抽样间隔的延时,而且还产生一个 $90^\circ$ 的相移。这种使所有频率的相移皆为 $90^\circ$ 的网络,称为 $90^\circ$ 移相器,或称正交变换网络,它和理想低通滤波器、理想微分器一样,有着极重要的理论和实际意义。

因而 $h(n)$ 对 $n=(N-1)/2$ 奇对称时,FIR滤波器将是一个具有准确的线性相位的理想正交变换网络。

### 三、幅度函数的特点

下面分成四种情况分别讨论 $H(\omega)$ 的特点。

1.  $h(n)$ 偶对称, $N$ 为奇数

从 $h(n)$ 偶对称的幅度函数(7-17)式

$$H(\omega) = \sum_{n=0}^{N-1} h(n) \cos\left[\left(\frac{N-1}{2} - n\right)\omega\right]$$

看出,不但 $h(n)$ 对于 $(N-1)/2$ 呈偶对称,满足 $h(n) = h(N-1-n)$ ,而且 $\cos\left[\left(\frac{N-1}{2} - n\right)\omega\right]$ 也对 $(N-1)/2$ 呈偶对称,满足

$$\cos\left[\left(\frac{N-1}{2} - n\right)\omega\right] = \cos\left[\omega\left(n - \frac{N-1}{2}\right)\right] = \cos\left\{\omega\left[\frac{N-1}{2} - (N-1-n)\right]\right\}$$

因而,整个 $\sum$ 内各项之间满足第 $n$ 项与第 $(N-1-n)$ 项是相等的。所以可以把两两相等的项合并,即 $n=0$ 与 $n=N-1$ 项合并, $n=1$ 与 $n=N-2$ 项合并等等,由于 $N$ 是奇数,故余下中间一项( $n=(N-1)/2$ ),而其余各项组合后共有 $(N-1)/2$ 项,则幅度函数可表示成

$$H(\omega) = h\left(\frac{N-1}{2}\right) + \sum_{n=0}^{(N-3)/2} 2h(n) \cos\left[\left(\frac{N-1}{2} - n\right)\omega\right]$$

$$\xrightarrow{\frac{N-1}{2} - n = m} h\left(\frac{N-1}{2}\right) + \sum_{m=1}^{(N-1)/2} 2h\left(\frac{N-1}{2} - m\right) \cos(m\omega)$$

可表示成

$$H(\omega) = \sum_{n=0}^{(N-1)/2} a(n) \cos(\omega n) \quad (7-22)$$

$$a(0) = h\left(\frac{N-1}{2}\right) \quad (7-23)$$

$$a(n) = 2h\left(\frac{N-1}{2} - n\right), \quad n = 1, 2, \dots, \frac{N-1}{2}$$

由此看出,当  $h(n)$  为偶对称,  $N$  为奇数时,由于  $\cos(n\omega)$  对于  $\omega=0, \pi, 2\pi$  皆为偶对称,所以幅度函数  $H(\omega)$  对  $\omega=0, \pi, 2\pi$  也呈偶对称。

### 2. $h(n)$ 偶对称, $N$ 为偶数

和  $h(n)$  偶对称、 $N$  为奇数情况的讨论相同,不同点仅由于  $N$  是偶数,故(7-17)式中没有单独的项,皆可两两合并为  $N/2$  项,即

$$H(\omega) = \sum_{n=0}^{N/2-1} 2h(n)\cos\left[\omega\left(\frac{N-1}{2}-n\right)\right] \xrightarrow{n=N/2-m} \sum_{m=1}^{N/2} 2h\left(\frac{N}{2}-m\right)\cos\left[\omega\left(m-\frac{1}{2}\right)\right]$$

可表示为

$$H(\omega) = \sum_{n=1}^{N/2} b(n)\cos\left[\omega\left(n-\frac{1}{2}\right)\right] \quad (7-24)$$

式中

$$b(n) = 2h\left(\frac{N}{2}-n\right), \quad n = 1, 2, \dots, \frac{N}{2} \quad (7-25)$$

由此看出,当  $h(n)$  为偶对称,  $N$  为偶数时,  $H(\omega)$  有以下特点:

- (1) 当  $\omega=\pi$  时,  $\cos\left[\omega\left(n-\frac{1}{2}\right)\right]=\cos\left[\pi\left(n-\frac{1}{2}\right)\right]=0$ , 故  $H(\pi)=0$ , 也就是  $H(z)$  在  $z=-1$  处必然有一个零点;
- (2) 由于  $\cos\left[\omega\left(n-\frac{1}{2}\right)\right]$  对  $\omega=\pi$  奇对称, 所以  $H(\omega)$  对  $\omega=\pi$  呈奇对称, 对  $\omega=0, 2\pi$  呈偶对称;
- (3) 如果一个滤波器在  $\omega=\pi$  时,  $H(\omega)$  不为零(例如高通滤波器或带阻滤波器), 则不能用这种滤波器。

### 3. $h(n)$ 奇对称, $N$ 为奇数

利用  $h(n)$  奇对称的幅度函数(7-20)式

$$H(\omega) = \sum_{n=0}^{N-1} h(n)\sin\left[\omega\left(\frac{N-1}{2}-n\right)\right]$$

由于

$$h(n) = -h(N-1-n)$$

则有

$$h\left(\frac{N-1}{2}\right) = -h\left(N-1-\frac{N-1}{2}\right) = -h\left(\frac{N-1}{2}\right)$$

所以

$$h\left(\frac{N-1}{2}\right) = 0$$

也就是中间项  $h\left(\frac{N-1}{2}\right)$  一定为零。

$h(n)$  是奇对称的, 而  $\sin\left[\left(\frac{N-1}{2}-n\right)\omega\right]$  也是奇对称的, 即

$$\begin{aligned} \sin\left[\left(\frac{N-1}{2}-n\right)\omega\right] &= -\sin\left[\left(n-\frac{N-1}{2}\right)\omega\right] \\ &= -\sin\left\{\left[\frac{N-1}{2}-(N-1-n)\right]\omega\right\} \end{aligned}$$

这样相乘的结果,(7-20)式  $\sum$  中的第  $n$  项与第  $(N-1-n)$  项的数值是相等的, 可将两两

相等的项合并, 合并后为  $(N-1)/2$  项, 可得

$$H(\omega) = \sum_{n=0}^{(N-3)/2} 2h(n) \sin\left[\omega\left(\frac{N-1}{2}-n\right)\right] \xrightarrow{m=\frac{N-1}{2}-n} \sum_{m=1}^{(N-1)/2} 2h\left(\frac{N-1}{2}-m\right) \sin(\omega m)$$

可表示为

$$H(\omega) = \sum_{n=1}^{(N-1)/2} c(n) \sin(\omega n) \quad (7-26)$$

式中

$$c(n) = 2h\left(\frac{N-1}{2}-n\right), \quad n = 1, 2, \dots, \frac{N-1}{2} \quad (7-27)$$

由此看出, 当  $h(n)$  为奇对称,  $N$  为奇数时,  $H(\omega)$  有以下特点:

(1) 由于  $\sin(n\omega)$  在  $\omega=0, \pi, 2\pi$  处都为零, 因此  $H(\omega)$  在  $\omega=0, \pi, 2\pi$  处必为零, 也就是  $H(z)$  在  $z=\pm 1$  处都为零;

(2) 由于  $\sin(n\omega)$  在  $\omega=0, \pi, 2\pi$  处都呈奇对称, 故  $H(\omega)$  对  $\omega=0, \pi, 2\pi$  也呈奇对称。

#### 4. $h(n)$ 奇对称, $N$ 为偶数

此时和上面第 3 种的情况是一样的, 但两两合并后共有  $N/2$  项, 因而有

$$H(\omega) = \sum_{n=0}^{N/2-1} 2h(n) \sin\left[\omega\left(\frac{N-1}{2}-n\right)\right] \xrightarrow{n=\frac{N}{2}-m} \sum_{m=1}^{N/2} 2h\left(\frac{N}{2}-m\right) \sin\left[\omega\left(m-\frac{1}{2}\right)\right]$$

可表示为

$$H(\omega) = \sum_{n=1}^{N/2} d(n) \sin\left[\omega\left(n-\frac{1}{2}\right)\right] \quad (7-28)$$

式中

$$d(n) = 2h\left(\frac{N}{2}-n\right), \quad n = 1, 2, \dots, \frac{N}{2} \quad (7-29)$$

由此看出, 当  $h(n)$  为奇对称,  $N$  为偶数时,  $H(\omega)$  有以下特点:

(1) 由于  $\sin\left[\omega\left(n-\frac{1}{2}\right)\right]$  在  $\omega=0, 2\pi$  处为零, 所以  $H(\omega)$  在  $\omega=0, 2\pi$  处也为零。即  $H(z)$  在  $z=1$  处为零点;

(2) 由于  $\sin\left[\omega\left(n-\frac{1}{2}\right)\right]$  在  $\omega=0, 2\pi$  处呈奇对称, 在  $\omega=\pi$  处呈偶对称, 故  $H(\omega)$  在  $\omega=0, 2\pi$  处呈奇对称, 在  $\omega=\pi$  处呈偶对称。

第 3 种与第 4 种线性相位 FIR 滤波器适合于在微分器及  $90^\circ$  移相器(希尔伯特变换器)中应用。

由(7-18)式和(7-21)式可得任一种线性相位 FIR 滤波器的群延时都为

$$\tau(e^{j\omega}) = -\frac{d\theta}{d\omega} = \frac{N-1}{2} \quad (7-30)$$

可以看出, 当  $N$  为奇数时, 滤波器的群延时为整数个抽样间隔; 当  $N$  为偶数时, 滤波器的群延时为整数个抽样间隔加上  $1/2$  个抽样间隔。

我们把这四种线性相位 FIR 滤波器的特性归纳在表 7-1 中。

表 7-1 四种线性相位 FIR 滤波器特性

偶对称单位冲激响应 $h(n) = h(N-1-n)$		
情况 1	相位响应 $\theta(\omega) = -\omega \left( \frac{N-1}{2} \right)$ 	$N$ 为奇数 $h(n)$  $a(n)$  $H(\omega) = \sum_{n=0}^{(N-1)/2} a(n) \cos(n\omega)$ 
	$N$ 为偶数 $h(n)$  $b(n)$  $H(\omega) = \sum_{n=1}^{N/2} b(n) \cos \left[ \left( n - \frac{1}{2} \right) \omega \right]$ 	
奇对称单位冲激响应 $h(n) = -h(N-1-n)$		
情况 3	相位响应 $\theta(\omega) = -\omega \left( \frac{N-1}{2} \right) + \frac{\pi}{2}$ 	$N$ 为奇数 $h(n)$  $C(n)$  $H(\omega) = \sum_{n=1}^{(N-1)/2} C(n) \sin(n\omega)$ 
	$N$ 为偶数 $h(n)$  $d(n)$  $H(\omega) = \sum_{n=1}^{N/2} d(n) \sin \left[ \omega \left( n - \frac{1}{2} \right) \right]$ 	
四、零点位置		
由(7-14)式可知, $H(z)$ 和 $H(z^{-1})$ 两者只差 $N-1$ 个抽样的延时及 $\pm 1$ 的乘因子, 其他则完全相同。因为它们是有限长序列的 $z$ 变换, 因而都是 $z$ (或 $z^{-1}$ ) 的多项式。		

(1) 若  $z=z_i$  是  $H(z)$  的零点, 即  $H(z_i)=0$ , 则  $z=\frac{1}{z_i}=z_i^{-1}$  也一定是  $H(z)$  的零点,

因为由(7-14)式知  $H(z_i^{-1})=\pm z_i^{N-1}H(z_i)=0$ 。

(2) 由于  $h(n)$  是实数, 所以  $H(z)$  的零点必然是以共轭对存在的, 因而  $z=z_i^*$  及  $z=(z_i^{-1})^*=\frac{1}{z_i^*}$ , 也一定是  $H(z)$  的零点。

综合(1),(2)两点可知, 线性相位 FIR 数字滤波器的零点必是互为倒数的共轭对, 或者说是共轭镜像的。有以下四种情况:

① 零点  $z_i$  既不在实轴上, 也不在单位圆上, 即  $z_i=r_i e^{j\theta_i}$ ,  $r_i \neq 1$ ,  $\theta_i \neq 0$ , 零点是两组互为倒数的共轭对, 如图 7-5(a)所示, 因而它们的基本因子为

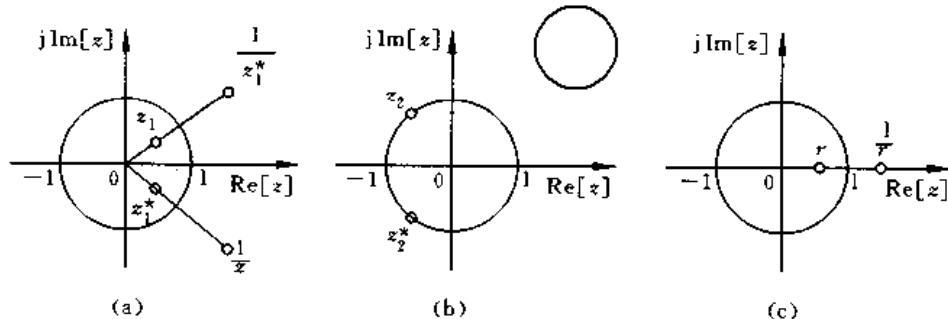
$$\begin{aligned} H_i(z) &= (1 - z^{-1} r_i e^{j\theta_i})(1 - z^{-1} r_i e^{-j\theta_i}) \left(1 - z^{-1} \frac{1}{r_i} e^{j\theta_i}\right) \left(1 - z^{-1} \frac{1}{r_i} e^{-j\theta_i}\right) \\ &= 1 - 2 \left(\frac{r_i^2 + 1}{r_i}\right) (\cos\theta_i) \cdot z^{-1} + \left(r_i^2 + \frac{1}{r_i^2} + 4\cos\theta_i\right) z^{-2} \\ &\quad - 2 \left(\frac{r_i^2 + 1}{r_i}\right) (\cos\theta_i) z^{-3} + z^{-4} \\ &= 1 + az^{-1} + bz^{-2} + az^{-3} + z^{-4} \end{aligned} \quad (7-31)$$

式中

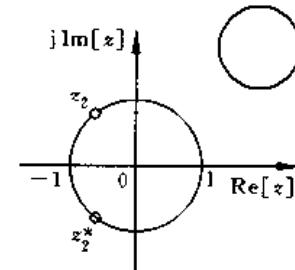
$$a = -2 \left(\frac{r_i^2 + 1}{r_i}\right) (\cos\theta_i), \quad b = r_i^2 + \frac{1}{r_i^2} + 4\cos\theta_i$$

如果化成两个实系数二阶多项式(把共轭对因子相乘), 则可表示为

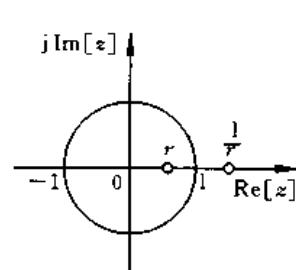
$$H_i(z) = \frac{1}{r_i^2} [1 - 2r_i(\cos\theta_i) \cdot z^{-1} + r_i^2 z^{-2}] [r_i^2 - 2r_i(\cos\theta_i) \cdot z^{-1} + z^{-2}] \quad (7-32)$$



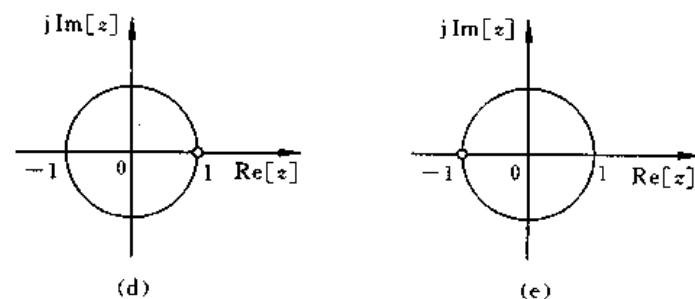
(a)



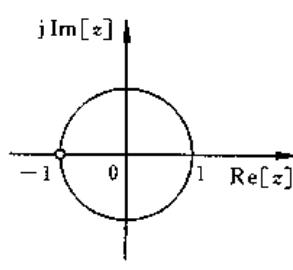
(b)



(c)



(d)



(e)

图 7-5 线性相位 FIR 滤波器的零点位置图

(7-31)式可用线性相位 FIR 滤波器的直接型结构实现,(7-32)式则可用图 7-6 的线性相位 FIR 滤波器的级联型结构实现,在这种情况下, $N=5, \tau=\frac{N-1}{2}=2$ 。

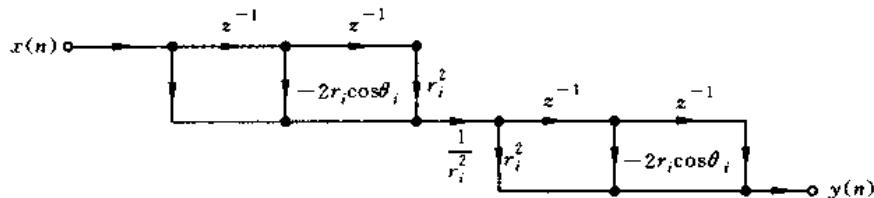


图 7-6 线性相位 FIR 滤波器的级联结构

② 零点  $z_i$  在单位圆上,但不在实轴上,即  $r_i=1, \theta_i \neq 0$  或  $\pi$ ,此时零点的共轭值就是它的倒数,这种零点情况,如图 7-5(b)所示,它们的基本因子为

$$H_i(z) = (1 - z^{-1} e^{j\theta_i})(1 - z^{-1} e^{-j\theta_i}) = 1 - 2r(\cos \theta_i)z^{-1} + z^{-2} \quad (7-33)$$

此时  $N=3, \tau=\frac{N-1}{2}=1$ 。

③ 零点  $z_i$  在实轴上,但不在单位圆上,即  $r_i \neq 1, \theta_i=0$  或  $\pi$ ,此时零点是实数,它没有复共轭部分,只有倒数,倒数也在实轴上,这种零点情况见图 7-5(c)所示,它们的基本因子为

$$H_i(z) = (1 \pm r_i z^{-1})\left(1 \pm \frac{1}{r_i} z^{-1}\right) = 1 \pm \left(r_i + \frac{1}{r_i}\right)z^{-1} + z^{-2} \quad (7-34)$$

式中“+”号相当于  $\theta_i=\pi$ ,零点在负实轴上,“-”号相当于  $\theta_i=0$ ,零点在正实轴上。此时  $N=3, \tau=\frac{N-1}{2}=1$ 。

④ 零点  $z_i$  既在实轴上也在单位圆上,即  $r_i=1, \theta_i=0$  或  $\pi$ ,这时零点只有两种可能情况,即  $z_i=1$  或  $z_i=-1$ ,分别如图 7-5(d)、图 7-5(e)所示,这时零点既是自己的复共轭,又是倒数,其基本因子为

$$H_i(z) = 1 \pm z^{-1} \quad (7-35)$$

式中“+”号表示零点在  $z=-1$  处,“-”号表示零点在  $z=1$  处。此时  $N=2, \tau=\frac{N-1}{2}=\frac{1}{2}$ ,即有半个抽样的延时。

我们从幅度响应的讨论中已经知道,对于第二种线性相位滤波器( $h(n)$ 偶对称, $N$ 为偶数), $H(\pi)=0$ ,故必有单根  $z=-1$ ,即包含有图 7-5(e)所示的第 4 类零点。而第四种( $h(n)$ 奇对称, $N$ 为偶数) $H(0)=0$ ,故必有单根  $z=1$ ,即包含有图 7-5(d)所示的第 4 类零点。第三种( $h(n)$ 奇对称, $N$ 为奇数), $H(0)=H(\pi)=0$ ,故必须  $z=1$  及  $z=-1$  都为零点。

显然,线性相位 FIR 滤波器的  $H(z)$  只可能是由以上这几种因子的组合而构成。

了解了线性相位 FIR 滤波器的各种特性,便可根据实际需要选择合适类型的 FIR 滤波器,同时设计时要遵循有关的约束条件。下面我们只讨论线性相位 FIR 滤波器的设计方法,因为它是用得最多的。

### 7.3 窗函数设计法

这种方法也称为傅里叶级数法。

#### 一、设计方法

一般是先给所要求的理想的滤波器频率响应  $H_d(e^{j\omega})$ , 要求设计一个 FIR 滤波器频率响应  $H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n)e^{-j\omega n}$  来逼近  $H_d(e^{j\omega})$ 。但是设计是在时域进行的, 因而先由  $H_d(e^{j\omega})$  的傅里叶反变换导出  $h_d(n)$ , 即

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega}) e^{j\omega n} d\omega \quad (7-36)$$

由于  $H_d(e^{j\omega})$  是矩形频率特性, 故  $h_d(n)$  一定是无限长的序列, 且是非因果的, 而我们要设计的是 FIR 滤波器, 其  $h(n)$  必然是有限长的, 所以要用有限长的  $h(n)$  来逼近无限长的  $h_d(n)$ , 最有效的方法是截断  $h_d(n)$ , 或者说用一个有限长度的窗口函数序列  $w(n)$  来截取  $h_d(n)$ , 即

$$h(n) = w(n)h_d(n) \quad (7-37)$$

因而窗函数序列的形状及长度的选择就很关键。

我们以一个截止频率为  $\omega_c$  的线性相位的理想矩形幅度特性的低通滤波器为例来加以讨论。设低通特性的群延时为  $\alpha$ , 即

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega\alpha}, & -\omega_c \leq \omega \leq \omega_c \\ 0, & \omega_c < \omega \leq \pi, -\pi \leq \omega < -\omega_c \end{cases} \quad (7-38)$$

这表明, 在通带  $|\omega| \leq \omega_c$  范围内,  $H_d(e^{j\omega})$  的幅度是均匀的, 其值为 1, 相位是  $-\omega\alpha$ 。利用(7-36)式可得

$$h_d(n) = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{-j\omega\alpha} e^{j\omega n} d\omega = \frac{\omega_c}{\pi} \frac{\sin[\omega_c(n - \alpha)]}{\omega_c(n - \alpha)} \quad (7-39)$$

$h_d(n)$  表示在图 7-7 上, 它是中心点在  $\alpha$  的偶对称无限长非因果序列, 要得到有限长的  $h(n)$ , 一种最简单的办法就是取矩形窗  $R_N(n)$ , 如图 7-7 所示, 即

$$w(n) = R_N(n)$$

但是按照线性相位滤波器的约束,  $h(n)$  必须是偶对称的, 对称中心应为长度的一半  $(N-1)/2$ , 因而必须  $\alpha = (N-1)/2$ , 所以有

$$\text{○} \quad \left\{ \begin{array}{l} h(n) = h_d(n)w(n) = \begin{cases} h_d(n), & 0 \leq n \leq N-1 \\ 0, & n \text{ 为其他值} \end{cases} \\ \alpha = \frac{N-1}{2} \end{array} \right. \quad (7-40)$$

将(7-39)式代入(7-40)式, 可得

$$h(n) = \begin{cases} \frac{\omega_c}{\pi} \frac{\sin[\omega_c(n - \frac{N-1}{2})]}{\omega_c(n - \frac{N-1}{2})}, & 0 \leq n \leq N-1 \\ 0, & n \text{ 为其他值} \end{cases} \quad (7-41)$$

此时,一定满足  $h(n)=h(N-1-n)$  这一线性相位特性的条件。图 7-7 示出了  $h_d(n)$ ,  $w(n)=R_N(n)$  以及它们的傅里叶变换的幅度图形。

下面求  $h(n)$  的傅里叶变换,也就是找出待求 FIR 滤波器的频率特性,以便能看出加窗处理后究竟对频率响应有何影响。

按照复卷积公式,在时域是相乘(见(7-37)式),则频域上是周期性卷积关系,即

$$H(e^{j\omega}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\theta}) W(e^{j(\omega-\theta)}) d\theta \quad (7-42)$$

因而  $H(e^{j\omega})$  逼近  $H_d(e^{j\omega})$  的好坏,完全取决于窗函数的频率特性  $W(e^{j\omega})$ 。

窗函数  $w(n)$  的频率特性  $W(e^{j\omega})$  为

$$W(e^{j\omega}) = \sum_{n=0}^{N-1} w(n) e^{-j\omega n} \quad (7-43)$$

对矩形窗  $R_N(n)$ ,则有

$$W_R(e^{j\omega}) = \sum_{n=0}^{N-1} e^{-j\omega n} = e^{-j\omega(\frac{N-1}{2})} \frac{\sin(\frac{\omega N}{2})}{\sin(\frac{\omega}{2})} \quad (7-44)$$

也可表示成幅度函数与相位函数

$$W_R(e^{j\omega}) = W_R(\omega) e^{-j(\frac{N-1}{2})\omega} \quad (7-45)$$

其中

$$W_R(\omega) = \frac{\sin(\frac{\omega N}{2})}{\sin(\frac{\omega}{2})} \quad (7-46)$$

$W_R(e^{j\omega})$  就是在第三章讨论过的频域抽样内插函数(差一个常数因子  $1/N$ ),其幅度函数  $W_R(\omega)$  在  $\omega=\pm 2\pi/N$  之内为一个主瓣,两侧形成许多衰减振荡的旁瓣,如图 7-7 所示  $W_R(\omega)$  是周期函数,这里只画了  $|\omega|<\pi$  的一部分图形。

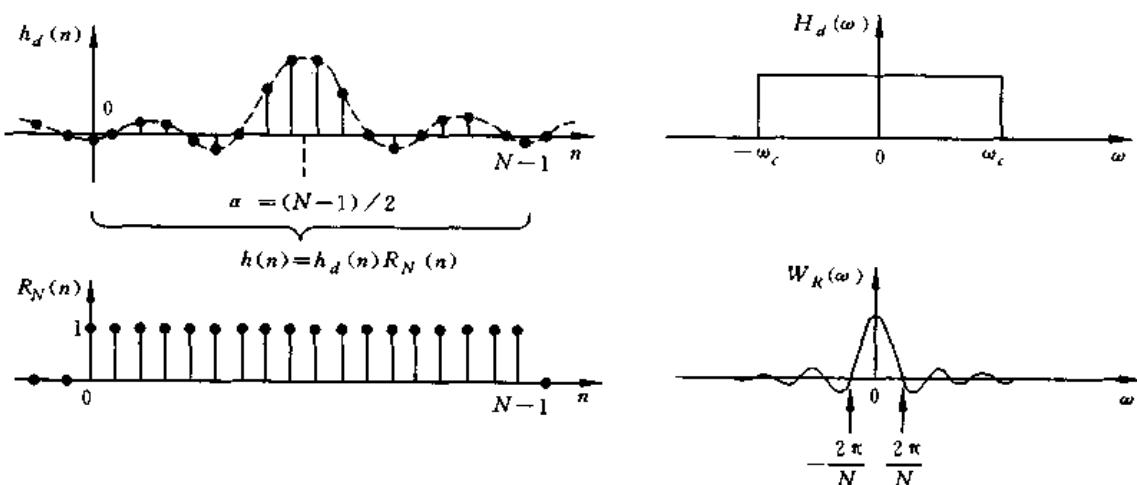


图 7-7 理想矩形幅频特性的  $h_d(n)$  和  $H_d(\omega)$  以及矩形窗函数序列的  $w(n)=R_N(n)$  及  $W_R(\omega)$

如果将理想频率响应也写成

$$H_d(e^{j\omega}) = H_d(\omega) e^{-j(\frac{N-1}{2})\omega} \quad (7-47)$$

则其幅度函数为

$$H_d(\omega) = \begin{cases} 1, & |\omega| \leq \omega_c \\ 0, & \omega_c < |\omega| \leq \pi \end{cases} \quad (7-48)$$

将(7-45)式和(7-47)式代入频域复卷积关系(7-42)式中,可得 FIR 滤波器的频率响应  $H(e^{j\omega})$  为

$$\begin{aligned} H(e^{j\omega}) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(\theta) e^{-j(\frac{N-1}{2})\theta} W_R(\omega - \theta) e^{-j(\frac{N-1}{2})(\omega - \theta)} d\theta \\ &= e^{-j(\frac{N-1}{2})\omega} \cdot \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(\theta) W_R(\omega - \theta) d\theta \end{aligned} \quad (7-49)$$

显然,这个频率响应也是线性相位的。

同样令

$$H(e^{j\omega}) = H(\omega) e^{-j(\frac{N-1}{2})\omega} \quad (7-50)$$

则实际求得的 FIR 数字滤波器的幅度函数  $H(\omega)$  为

$$H(\omega) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(\theta) W_R(\omega - \theta) d\theta \quad (7-51)$$

由此可见,对实际 FIR 滤波器频率响应的幅度函数  $H(\omega)$  起影响的是窗函数频率响应的幅度函数  $W_R(\omega)$ 。

(7-51)式卷积过程可用图 7-8 说明,只要看几个特殊的频率点,就可以看出  $H(\omega)$  的一般情况。特别要注意卷积过程给  $H(\omega)$  造成的起伏现象。

(1)  $\omega=0$  时的零频率处的响应值  $H(0)$  应该是图 7-8 中(a)图与(b)图两函数乘积的积分,也就是  $W_R(\theta)$  在  $\theta=-\omega_c$  到  $\theta=\omega_c$  一段内的积分面积。由于一般情况下都满足  $\omega_c \gg 2\pi/N$ ,所以  $H(0)$  可以近似看成是  $\theta$  从  $-\pi$  到  $\pi$  的  $W_R(\theta)$  的全部积分面积。

(2)  $\omega=\omega_c$  时,  $H_d(\theta)$  正好与  $W_R(\omega-\theta)$  的一半重叠,如图 7-8(c)所示,因此  $\frac{H(\omega_c)}{H(0)}=0.5$ 。

(3)  $\omega=\omega_c - \frac{2\pi}{N}$  时,  $W_R(\omega-\theta)$  的全部主瓣在  $H_d(\theta)$  的通带  $|\omega| \leq \omega_c$  之内,如图 7-8(d) 所示,因此卷积结果有最大值,即  $H\left(\omega_c - \frac{2\pi}{N}\right)$  为最大值,频率响应出现正肩峰。

(4)  $\omega=\omega_c + \frac{2\pi}{N}$  时,  $W_R(\omega-\theta)$  的全部主瓣都在  $H_d(\theta)$  的通带之外,如图 7-8(e)所示,而通带内的旁瓣负的面积大于正的面积,因而卷积结果达到最负值,出现负的肩峰。

(5) 当  $\omega > \omega_c + \frac{2\pi}{N}$  时,随着  $\omega$  的增加,  $W_R(\omega-\theta)$  左边旁瓣的起伏部分将扫过通带,卷积值也将随  $W_R(\omega-\theta)$  的旁瓣在通带内面积的变化而变化,故  $H(\omega)$  将围绕着零值而波动。当  $\omega$  由  $\omega_c - \frac{2\pi}{N}$  向通带内减小时,  $W_R(\omega-\theta)$  的右旁瓣将进入  $H_d(\omega)$  的通带,右旁瓣的起伏造成  $H(\omega)$  值将围绕  $H(0)$  值而摆动。卷积得到的  $H(\omega)$  如图 7-8(f)所示。

综上所述,加窗处理对理想矩形频率响应产生以下几点影响:

(1) 使理想频率特性不连续点处边沿加宽,形成一个过渡带,过渡带的宽度等于窗的

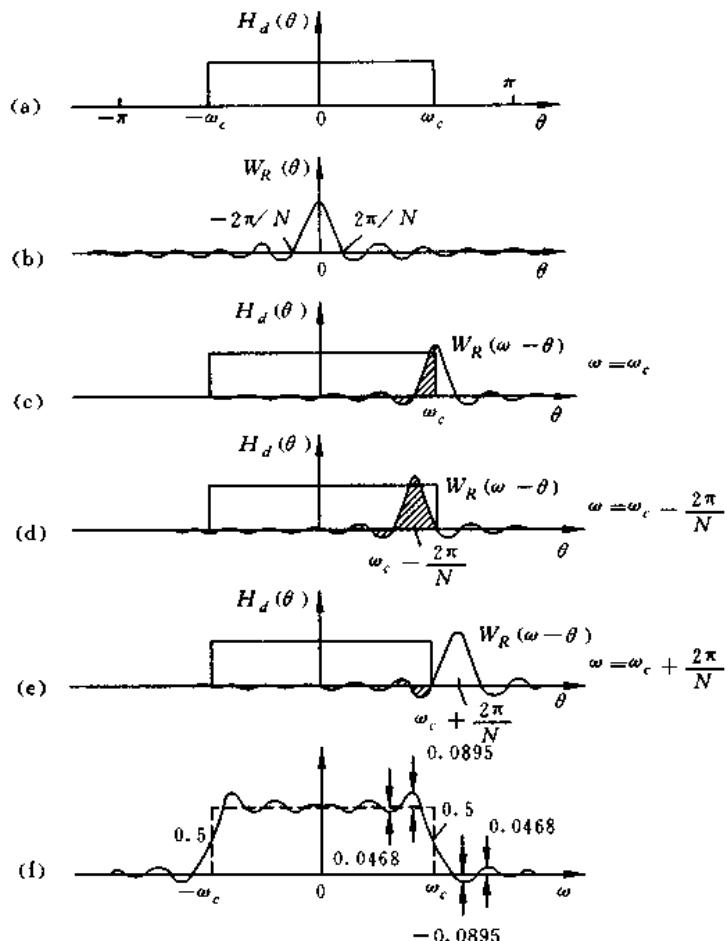


图 7-8 矩形窗的卷积过程

频率响应  $W_R(\omega)$  的主瓣宽度  $\Delta\omega = 4\pi/N$ 。注意，我们这里所说的过渡带是指两个肩峰之间的宽度，与滤波器的真正过渡带还有一些区别，也就是说，滤波器的过渡带比这个数值 ( $4\pi/N$ ) 要小。

(2) 在截止频率  $\omega_c$  的两边  $\omega = \omega_c \pm \frac{2\pi}{N}$  的地方（即过渡带的两边）， $H(\omega)$  出现最大的肩峰值，肩峰的两侧形成起伏振荡，其振荡幅度取决于旁瓣的相对幅度，而振荡的多少，则取决于旁瓣的多少。

(3) 增加截取长度  $N$ ，则在主瓣附近的窗的频率响应为

$$W_R(\omega) = \frac{\sin\left(\frac{N\omega}{2}\right)}{\sin\left(\frac{\omega}{2}\right)} \approx \frac{\sin\left(\frac{N\omega}{2}\right)}{\frac{\omega}{2}} = N \frac{\sin x}{x} \quad (7-52)$$

其中  $x = N\omega/2$ 。可见，改变  $N$ ，只能改变窗谱的主瓣宽度、改变  $\omega$  坐标的比例以及改变  $W_R(\omega)$  的绝对值大小，但是不能改变主瓣与旁瓣的相对比例（当然  $N$  太小时，会影响旁瓣的相对值），这个相对比例是由  $\sin x/x$  决定的，或者说只由窗函数的形状来决定的。因

而,当截取长度  $N$  增加时,只会减小过渡带宽  $4\pi/N$ ,而不会改变肩峰的相对值。例如在矩形窗情况下,最大相对肩峰值为 8.95%, $N$  增加时, $2\pi/N$  减小,故起伏振荡变密,最大肩峰则总是 8.95%,这种现象称为吉布斯(Gibbs)效应。我们知道,窗谱肩峰的大小,影响到  $H(\omega)$  通带的平稳和阻带的衰减,对滤波器性能影响很大。

## 二、各种窗函数

矩形窗截断造成肩峰为 8.95%,则阻带最小衰减为  $20 \log_{10}(8.95\%) = -21\text{dB}$ ,这个衰减量在工程上常常是不够的。为了加大阻带衰减,只能改善窗函数的形状。从(7-51)式的频域周期卷积式看出,只有当窗谱逼近冲激函数时,也就是绝大部分能量集中在频谱中点时, $H(\omega)$  才会逼近  $H_d(\omega)$ 。这相当于窗的宽度为无穷长,等于不加窗口截断,没有实际意义。

从以上讨论可看出,一般希望窗函数满足两项要求:① 窗谱主瓣尽可能地窄,以获得较陡的过渡带;② 尽量减少窗谱的最大旁瓣的相对幅度,也就是能量尽量集中于主瓣,这样使肩峰和波纹减小,就可增大阻带的衰减。但是这两项要求是不能同时得到满足的,往往是增加主瓣宽度以换取对旁瓣的抑制。因而选用不同形状的窗函数都是为了得到平坦的通带幅度响应和较小的阻带波纹(也就是加大阻带衰减)。因而所选用的窗函数,其频谱旁瓣电平要较小,而主瓣就会加宽。这就是说窗函数在边沿处( $n=0$  和  $n=N-1$  附近)比矩形窗变化要平滑而缓慢,以减小由陡峭的边缘所引起的旁瓣分量,使阻带衰减增大。但窗谱的主瓣宽度却比矩形窗的要宽,这造成滤波器幅度函数过渡带的加宽。

常用的窗函数有以下几种:

### 1. 矩形窗

前面已讨论过了。

$$w(n) = R_N(n)$$

$$W_R(e^{j\omega}) = W_R(\omega) e^{-j(\frac{N-1}{2})\omega}$$

$$W_R(\omega) = \frac{\sin\left(\frac{N\omega}{2}\right)}{\sin\left(\frac{\omega}{2}\right)}$$

### 2. 三角形(Bartlett)窗

$$w(n) = \begin{cases} \frac{2n}{N-1}, & 0 \leq n \leq \frac{N-1}{2} \\ 2 - \frac{2n}{N-1}, & \frac{N-1}{2} < n \leq N-1 \end{cases} \quad (7-53)$$

窗谱为

$$W(e^{j\omega}) = \frac{2}{N-1} \left[ \frac{\sin\left[\left(\frac{N-1}{4}\right)\omega\right]}{\sin\left(\frac{\omega}{2}\right)} \right]^2 e^{-j(\frac{N-1}{2})\omega}$$

$$\approx \frac{2}{N} \left[ \frac{\sin\left(\frac{N\omega}{4}\right)}{\sin\left(\frac{\omega}{2}\right)} \right]^2 e^{-j(\frac{N-1}{2})\omega} \quad (7-54)$$

“ $\approx$ ”在  $N \gg 1$  时成立。

此时主瓣宽度为  $8\pi/N$ 。

### 3. 汉宁(Hanning)窗,又称升余弦窗

$$w(n) = \frac{1}{2} \left[ 1 - \cos\left(\frac{2\pi n}{N-1}\right) \right] R_N(n) \quad (7-55)$$

利用傅里叶变换的调制特性,即利用

$$e^{j\omega_0 n} x(n) \Leftrightarrow X(e^{j(\omega-\omega_0)})$$

“ $\Leftrightarrow$ ”表示互为傅里叶变换对。再利用

$$\cos n\omega_0 = \frac{e^{jn\omega_0} + e^{-jn\omega_0}}{2}$$

考虑到  $R_N(n)$  的傅里叶变换为

$$W_R(e^{j\omega}) = W_R(\omega) e^{-j(\frac{N-1}{2})\omega}$$

则得

$$\begin{aligned} W(e^{j\omega}) &= \text{DTFT}[w(n)] \\ &= \left\{ 0.5W_R(\omega) + 0.25 \left[ W_R\left(\omega - \frac{2\pi}{N-1}\right) + W_R\left(\omega + \frac{2\pi}{N-1}\right) \right] \right\} e^{-j(\frac{N-1}{2})\omega} \\ &= W(\omega) e^{-j(\frac{N-1}{2})\omega} \end{aligned} \quad (7-56)$$

当  $N \gg 1$  时,  $N-1 \approx N$ , 所以窗谱的幅度函数为

$$W(\omega) \approx 0.5W_R(\omega) + 0.25 \left[ W_R\left(\omega - \frac{2\pi}{N}\right) + W_R\left(\omega + \frac{2\pi}{N}\right) \right] \quad (7-57)$$

这三部分之和使旁瓣互相抵消,能量更集中在主瓣,见图 7-9,但是代价是主瓣宽度比矩形窗的主瓣宽度增加一倍,即为  $8\pi/N$ 。

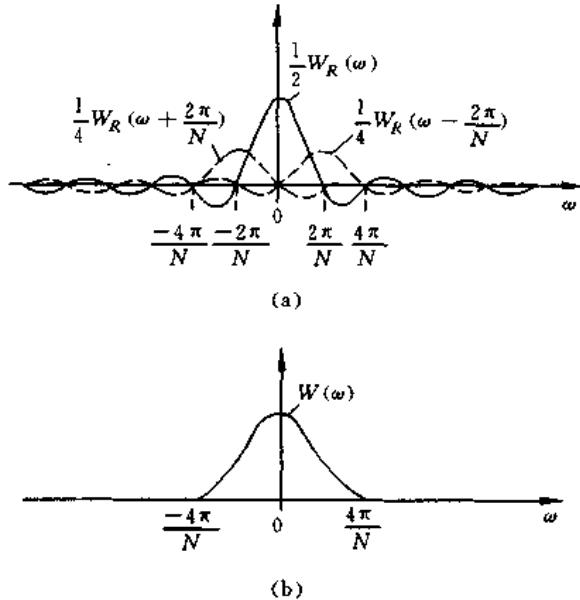


图 7-9 汉宁(Hanning)窗谱

汉宁窗是下面一类窗中的特例:

$$\left\{ \begin{array}{l} w(n) = \left[ \cos^{\alpha} \left( \frac{n\pi}{N-1} \right) \right] R_N(n) \\ w(n) = \left[ \sin^{\alpha} \left( \frac{n\pi}{N-1} \right) \right] R_N(n) \end{array} \right. \quad (7-58)$$

$$\left\{ \begin{array}{l} w(n) = \left[ \cos^{\alpha} \left( \frac{n\pi}{N-1} \right) \right] R_N(n) \\ w(n) = \left[ \sin^{\alpha} \left( \frac{n\pi}{N-1} \right) \right] R_N(n) \end{array} \right. \quad (7-59)$$

当  $\alpha=2$  时, (7-59)式就是汉宁窗。

#### 4. 海明(Hamming)窗,又称改进的升余弦窗

对升余弦加以改进,可以得到旁瓣更小的效果,窗形式为

$$w(n) = \left[ 0.54 - 0.46 \cos \left( \frac{2\pi n}{N-1} \right) \right] R_N(n) \quad (7-60)$$

其频率响应的幅度函数为

$$\begin{aligned} W(\omega) &= 0.54W_R(\omega) + 0.23 \left[ W_R \left( \omega - \frac{2\pi}{N-1} \right) + W_R \left( \omega + \frac{2\pi}{N-1} \right) \right] \\ &\approx 0.54W_R(\omega) + 0.23 \left[ W_R \left( \omega - \frac{2\pi}{N} \right) + W_R \left( \omega + \frac{2\pi}{N} \right) \right], \quad (\text{当 } N \gg 1) \end{aligned} \quad (7-61)$$

结果可将 99.963% 的能量集中在窗谱的主瓣内,与汉宁窗相比,主瓣宽度相同为  $8\pi/N$ ,但旁瓣幅度更小,旁瓣峰值小于主瓣峰值的 1%。

同样,海明窗是下面一类窗的特例( $\alpha=0.54$  时):

$$w(n) = \left[ \alpha - (1-\alpha) \cos \left( \frac{2\pi n}{N-1} \right) \right] R_N(n) \quad (7-62)$$

#### 5. 布拉克曼(Blackman)窗,又称二阶升余弦窗

为了更进一步抑制旁瓣,可再加上余弦的二次谐波分量,得到布拉克曼窗

$$w(n) = \left[ 0.42 - 0.5 \cos \left( \frac{2\pi n}{N-1} \right) + 0.08 \cos \left( \frac{4\pi n}{N-1} \right) \right] R_N(n) \quad (7-63)$$

其频谱的幅度函数为

$$\begin{aligned} W(\omega) &= 0.42W_R(\omega) + 0.25 \left[ W_R \left( \omega - \frac{2\pi}{N-1} \right) + W_R \left( \omega + \frac{2\pi}{N-1} \right) \right] \\ &\quad + 0.04 \left[ W_R \left( \omega - \frac{4\pi}{N-1} \right) + W_R \left( \omega + \frac{4\pi}{N-1} \right) \right] \end{aligned} \quad (7-64)$$

此时主瓣宽度为矩形窗谱主瓣宽度的三倍,即为  $\frac{12\pi}{N}$ 。

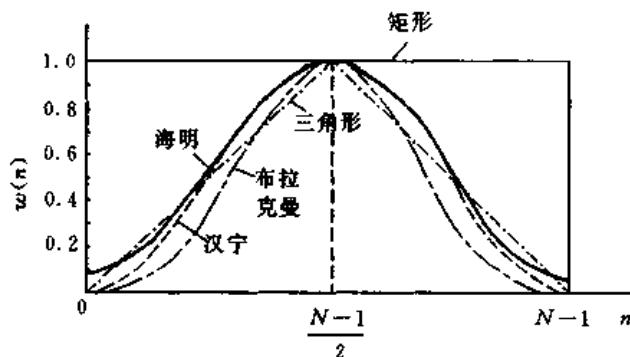


图 7-10 设计有限长单位冲激响应滤波器常用的几种窗函数

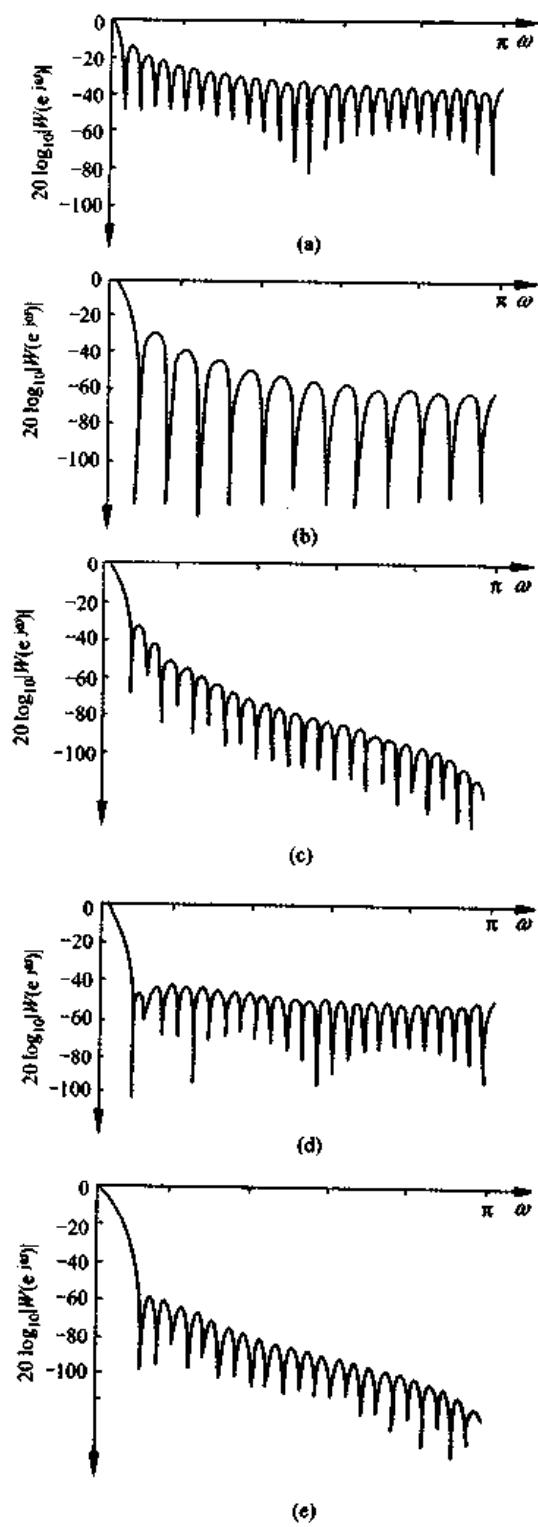


图 7-11 图 7-10 的各种窗函数的  
傅里叶变换 ( $N=51$ )

(a) 矩形窗; (b) 巴特列特窗(三角形窗);  
(c) 汉宁窗; (d) 海明窗; (e) 布拉克曼窗

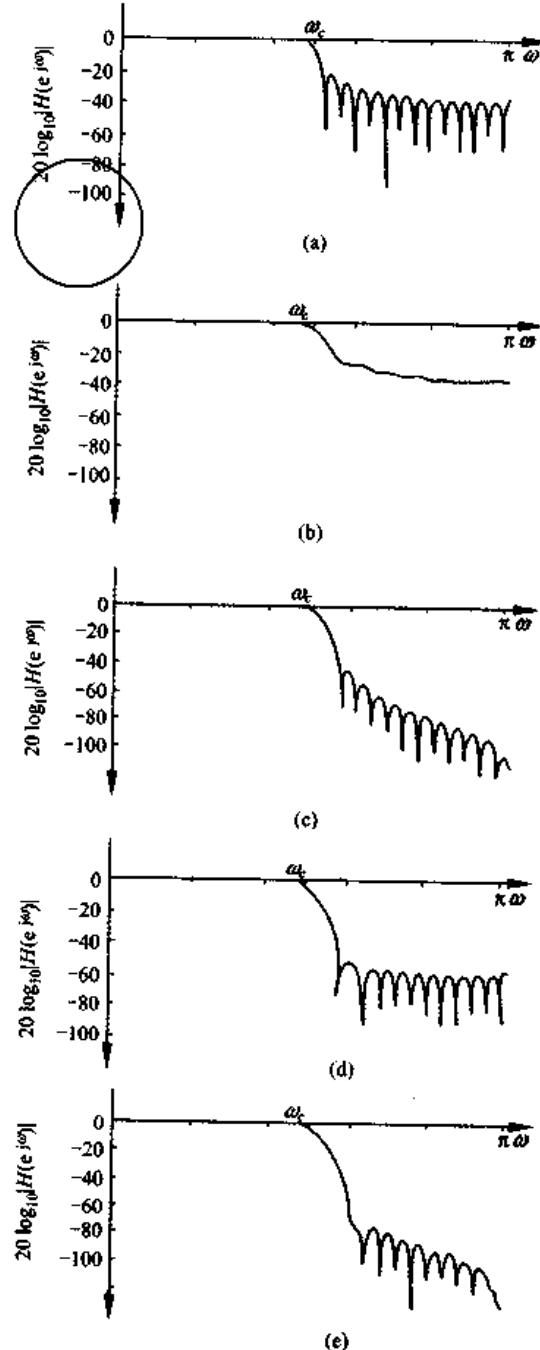


图 7-12 理想低通滤波器加窗后的  
幅度响应 ( $N=51$ )

(a) 矩形窗; (b) 巴特列特窗(三角形窗);  
(c) 汉宁窗; (d) 海明窗; (e) 布拉克曼窗

布拉克曼窗是下面一类窗的特例：

$$w(n) = \left[ \sum_{m=0}^M (-1)^m \alpha_m 0.5 \cos\left(\frac{2\pi n}{N-1} m\right) \right] R_N(n) \quad (7-65)$$

选  $M=2, \alpha_0=0.42, \alpha_1=0.50, \alpha_2=0.08$  时，即为布拉克曼窗。

图 7-10 给出了以上五种窗的窗函数，图 7-11 则是  $N=51$  时这五种窗函数的频谱。可以看出，随着窗形状的变化，旁瓣衰减加大，但主瓣宽度也相应的加宽了。

图 7-12 是利用这五种窗函数，对同一指标  $N=51$ ，截止频率  $\omega_c=0.5\pi$ ，设计的 FIR 线性相位低通滤波器频率特性。

#### 6. 凯泽窗(Kaiser)

这是一种适应性较强的窗，其窗函数的表示式为

$$w(n) = \frac{I_0\left(\beta \sqrt{1 - \left(1 - \frac{2n}{N-1}\right)^2}\right)}{I_0(\beta)}, \quad 0 \leq n \leq N-1 \quad (7-66)$$

其中  $I_0(\cdot)$  是第一类变形零阶贝塞尔函数， $\beta$  是一个可自由选择的参数，它可以同时调整主瓣宽度与旁瓣电平， $\beta$  越大，则  $w(n)$  窗越窄，而频谱的旁瓣越小，但主瓣宽度也相应增加。因而改变  $\beta$  值就可对主瓣宽度与旁瓣衰减进行选择，凯泽窗函数的曲线如图 7-13 所示。一般选择  $4 < \beta < 9$ ，这相当于旁瓣幅度与主瓣幅度的比值由 3.1% 变到 0.047%（-30dB 到 -67dB）。

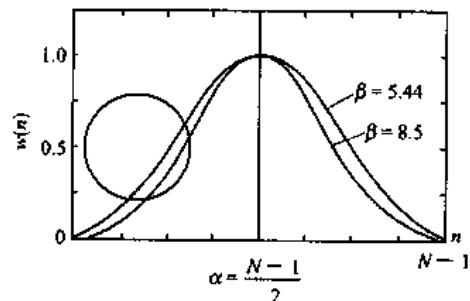


图 7-13 凯泽窗函数

凯泽窗在不同  $\beta$  值下的性能归纳在表 7-2 中。

表 7-2 凯泽窗参数对滤波器的性能的影响

$\beta$	过渡带	通带波纹/dB	阻带最小衰减/dB
2.120	$3.00\pi/N$	$\pm 0.27$	-30
3.384	$4.46\pi/N$	$\pm 0.0868$	-40
4.538	$5.86\pi/N$	$\pm 0.0274$	-50
5.658	$7.24\pi/N$	$\pm 0.00868$	-60
6.764	$8.64\pi/N$	$\pm 0.00275$	-70
7.865	$10.0\pi/N$	$\pm 0.000868$	-80
8.960	$11.4\pi/N$	$\pm 0.000275$	-90
10.056	$12.8\pi/N$	$\pm 0.000087$	-100

当  $\beta=0$  时，相当于矩形窗，这是因为  $I_0(0)=1$ ，故  $w(n)=1, 0 \leq n \leq N-1$ 。

当  $\beta=5.44$  时，相当于海明窗，但凯泽窗旁瓣频谱收敛得更快。海明窗除 0.037% 能量外，都在主瓣之内，而凯泽窗除 0.012% 能量外都在主瓣之内，因而能量更加集中在主瓣中。

当  $\beta=8.5$  时，相当于布拉克曼窗。

由(7-65)式看出，凯泽窗函数是以  $n=(N-1)/2$  为对称中心呈偶对称的，即

$$w(n) = w(N-1-n)$$

而

$$w\left(\frac{N-1}{2}\right) = \frac{I_0(\beta)}{I_0(\beta)} = 1$$

从  $n=(N-1)/2$  这一中点向两边变化时,  $w(n)$  逐渐减小, 最边上两点

$$w(0) = w(N-1) = \frac{1}{I_0(\beta)}$$

参数  $\beta$  越大,  $w(n)$  变化越快, 如图 7-13 所示。

凯泽窗设计中有经验公式可供使用, 给定过渡带宽  $\Delta\omega$  (rad), 阻带衰减  $\delta_2 = -20\log_{10} \alpha_2$  (dB), 则可求得凯泽窗 FIR 滤波器的阶数  $N$  和形状参数  $\beta$ , 即

$$N = \frac{\delta_2 - 7.95}{2.286\Delta\omega}$$

$$\beta = \begin{cases} 0.1102(\delta_2 - 8.7), & \delta_2 \geq 50\text{dB} \\ 0.5842(\delta_2 - 21)^{0.4} + 0.07886(\delta_2 - 21), & 21\text{dB} < \delta_2 < 50\text{dB} \\ 0, & \delta_2 \leq 21\text{dB} \end{cases}$$

表 7-3 归纳了以上提到的几种窗的主要性能, 供设计 FIR 滤波器时参考。

表 7-3 六种窗函数基本参数的比较

窗函数	窗谱性能指标		加窗后滤波器性能指标	
	旁瓣峰值 /dB	主瓣宽度 /(2\pi/N)	过渡带宽 \Delta\omega /(2\pi/N)	阻带最小衰减 /dB
矩形窗	-13	2	0.9	-21
三角形窗	-25	4	2.1	-25
汉宁窗	-31	4	3.1	-44
海明窗	-41	4	3.3	-53
布拉克曼窗	-57	6	5.5	-74
凯泽窗 (\beta=7.865)	-57		5	-80

从以上讨论可以看出, 最小阻带衰减只由窗形状决定, 不受  $N$  的影响, 而过渡带的宽度则随窗宽的增加而减小。

### 三、窗函数法的设计步骤

- (1) 首先是给定所要求的频率响应函数  $H_d(e^{j\omega})$ ;
- (2) 其次, 利用(7-22)式求  $h_d(n) = \text{IDTFT}[H_d(e^{j\omega})]$ ;
- (3) 再次, 由过渡带宽及阻带最小衰减的要求, 利用表 7-2、表 7-3, 可选定窗  $w(n)$  的形状及  $N$  的大小, 一般  $N$  要通过几次试探而最后确定;
- (4) 求得所设计的 FIR 滤波器的单位抽样响应

$$h(n) = h_d(n)w(n), \quad n = 0, 1, \dots, N-1$$

(5) 求  $H(e^{j\omega}) = \text{DTFT}[h(n)]$ , 检验是否满足设计要求, 如不满足, 则需重新设计。

### 四、窗函数法计算中的主要问题

首先当  $H_d(e^{j\omega})$  很复杂或不能按(7-36)式直接计算积分时, 则必须用求和代替积分, 以便在计算机上计算, 也就是要计算离散傅里叶反变换, 一般都采用 FFT 来计算。将积分限分成  $M$  段, 也就是令抽样频率为

$$\omega_k = \frac{2\pi}{M} k, \quad k = 1, 2, \dots, M-1$$

则有

$$h_M(n) = \frac{1}{M} \sum_{k=0}^{M-1} H_d(e^{\frac{j\pi k}{M}}) e^{j\frac{2\pi}{M}kn} \quad (7-67)$$

频域的抽样造成时域序列的周期延拓, 延拓周期是  $M$ , 即

$$h_M(n) = \sum_{r=-\infty}^{\infty} h_d(n+rM) \quad (7-68)$$

由于  $h_d(n)$  有可能是无限长的序列, 因而严格说, 必须  $M \rightarrow \infty$  时,  $h_M(n)$  才能等于  $h_d(n)$  而不产生混叠现象, 即

$$h_d(n) = \lim_{M \rightarrow \infty} h_M(n) \quad (7-69)$$

实际上, 由于  $h_d(n)$  随  $n$  增加衰减很快, 一般只要  $M$  足够大, 即  $M \gg N$ , 就足够了。其次, 窗函数设计法的另一个困难就是需要预先确定窗函数的形状和窗序列的点数  $N$ , 以满足给定的频率响应指标。这一困难可利用计算机采用累试法加以解决。

一般在设计凯泽窗时, 对零阶变形贝塞尔函数可采用无穷级数来表达

$$I_0(x) = \sum_{k=0}^{\infty} \left[ \frac{1}{k!} \left( \frac{x}{2} \right)^k \right]^2 = 1 + \sum_{k=1}^{\infty} \left[ \frac{1}{k!} \left( \frac{x}{2} \right)^k \right]^2 \quad (7-70)$$

这个无穷级数可用有限项级数去近似, 项数多少由要求的精度来确定。

窗函数法的优点是简单, 有闭合形式的公式可循, 因而很实用。其缺点是通带、阻带的截止频率不易控制。

## 五、线性相位 FIR 低通滤波器的设计

我们用一个例子来说明。

[例 7-1] 设计一个线性相位 FIR 低通滤波器, 给定抽样频率为  $\Omega_s = 2\pi \times 1.5 \times 10^4$  (rad/sec), 通带截止频率为  $\Omega_p = 2\pi \times 1.5 \times 10^3$  (rad/sec), 阻带起始频率为  $\Omega_u = 2\pi \times 3 \times 10^3$  (rad/sec), 阻带衰减不小于  $-50$  dB。幅度特性如图 7-14 所示。

解 (1) 求各对应的数字频率:

$$\text{通带截止频率为 } \omega_p = \frac{\Omega_p}{f_s} = 2\pi \frac{\Omega_p}{\Omega_s} = 0.2\pi$$

$$\text{阻带起始频率为 } \omega_u = \frac{\Omega_u}{f_s} = 2\pi \frac{\Omega_u}{\Omega_s} = 0.4\pi$$

阻带衰减相当于  $\delta_2 = 50$  dB

(2) 设  $H_d(e^{j\omega})$  为理想线性相位滤波器

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega r}, & |\omega| \leq \omega_c \\ 0, & \text{其他} \end{cases} \quad (7-71)$$

首先由所需低通滤波器的过渡带求理想低通滤波器的截止频率  $\Omega_c$  (由于  $\Omega_c$  为两个肩峰值处的频率的中点, 而由  $\Omega_p$  到  $\Omega_u$  之间的过渡带宽并非两个肩峰值间的频率差, 因而以下求出的  $\Omega_c$  有一定的近似。)

$$\Omega_c = \frac{1}{2}(\Omega_p + \Omega_u) = 2\pi \times 2.25 \times 10^3 \text{ (rad/sec)}$$

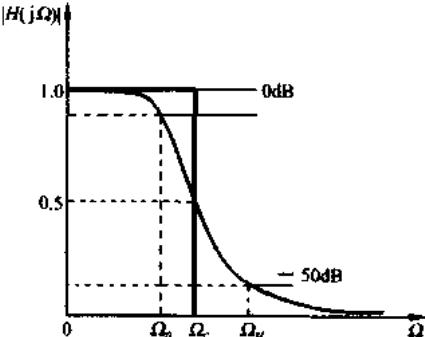


图 7-14 例 7-1 要求的模拟  
低通滤波器特性

其对应的数字频率为  $\omega_c = \frac{\Omega_c}{f_s} = 2\pi \frac{\Omega_c}{\Omega_s} = 0, 3\pi$

由此可得

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{-j\omega\tau} e^{j\omega n} d\omega = \frac{1}{2\pi} \int_{-\omega_c}^{\omega_c} e^{j\omega(n-\tau)} d\omega$$

$$= \begin{cases} \sin[\omega_c(n-\tau)], & n \neq \tau \\ \frac{\omega_c}{\pi}, & n = \tau \end{cases} \quad (7-72)$$

其中,  $\tau$  为线性相位所必须的移位, 我们已经知道应满足  $\tau = \frac{N-1}{2}$ 。

(3) 由阻带衰减  $\delta_2$  来确定窗形状, 由过渡带宽确定  $N$ 。

由于  $\delta_2 = 50$  dB, 查表 7-3 可选海明窗, 其阻带最小衰减  $-53$  dB 满足要求。

所要求的过渡带宽(数字频域)  $\Delta\omega = 2\pi \frac{\Omega_u - \Omega_p}{\Omega_s} = 2\pi \times \frac{2\pi \times 1.5 \times 10^3}{2\pi \times 1.5 \times 10^4} = 0.2\pi$

由于海明窗过渡带宽满足  $\Delta\omega = \frac{6.6\pi}{N}$ , 所以

$$N = \frac{6.6\pi}{\Delta\omega} = \frac{6.6\pi}{0.2\pi} = 33$$

$$\tau = \frac{N-1}{2} = 16$$

(4) 由海明窗表达式  $w(n)$  确定 FIR 滤波器的  $h(n)$ 。

海明窗  $w(n) = \left[ 0.54 - 0.46 \cos\left(\frac{2\pi n}{N-1}\right) \right] R_N(n)$

$$h_d(n) = \frac{\sin\left[\omega_c\left(n - \frac{N-1}{2}\right)\right]}{\pi\left(n - \frac{N-1}{2}\right)}$$

所以  $h(n) = h_d(n) \cdot w(n) = \frac{\sin[0.3\pi(n-16)]}{\pi(n-16)} \cdot \left[ 0.54 - 0.46 \cos\left(\frac{\pi n}{16}\right) \right] R_N(n)$

(5) 由  $h(n)$  求  $H(e^{j\omega})$  检验各项指标是否满足要求, 如不满足要求要改变  $N$ , 或改变窗形状(或两者都改变)来重新计算。

$H(e^{j\omega})$  的图形已画在图 7-15 上, 满足设计要求。

当然, 窗函数法也可设计其他类型的滤波器, 例如高通滤波器、带通滤波器、带阻滤波器。利用奇对称单位冲激响应的特点(见表 7-1)还可以设计  $90^\circ$  移相位(或称离散希尔伯特变换器)以及幅度响应与  $\omega$  成线性关系的线性差分器。

## 六、线性相位 FIR 高通滤波器的设计

按指标要求的理想线性相位高通滤波器的频率响应为

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega\tau}, & \omega_c \leq |\omega| \leq \pi \\ 0, & \text{其他 } \omega \end{cases} \quad (7-73)$$

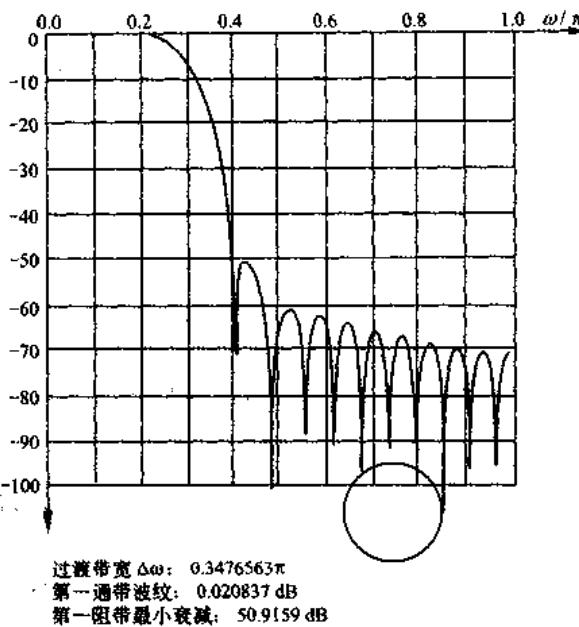


图 7-15 例 7-1 设计出的线性相位 FIR 低通滤波器幅频特性  
(海明窗,  $N=33$ )

其中  $\tau = \frac{N-1}{2}$ , 它的单位冲激响应为

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega}) e^{j\omega n} d\omega = \frac{1}{2\pi} \left[ \int_{-\pi}^{-\omega_c} e^{j\omega(n-\tau)} d\omega + \int_{\omega_c}^{\pi} e^{j\omega(n-\tau)} d\omega \right]$$

$$= \begin{cases} \frac{1}{\pi(n-\tau)} \{ \sin[(n-\tau)\pi] - \sin[(n-\tau)\omega_c] \}, & n \neq \tau \\ \frac{1}{\pi} (\pi - \omega_c) = 1 - \frac{\omega_c}{\pi}, & n = \tau \end{cases} \quad (7-74)$$

选定窗  $w(n)$  即可得所需线性相位 FIR 高通滤波器的单位冲激响应

$$h(n) = h_d(n) \cdot w(n)$$

选用哪一种窗函数和阻带衰减有关, 而时域窗的点数  $N$  则和过渡带宽有关, 但是由表 7-1 看出, 只能采用偶对称单位冲激响应 [ 这是一般通带滤波器 ( 包括低通、高通、带通、带阻 ) 所要求的单位冲激响应 ], 另外, 对高通滤波器来说  $N$  只能取奇数 ( $N=2m+1, m=0, 1, \dots$ ), 因为  $N$  为偶数时  $H(\omega)$  在  $\omega=\pi$  处  $H(\omega)|_{\omega=\pi}=H(\pi)=0$ , 不能做为高通滤波器。求出  $h(n)$  后, 可求  $H(e^{j\omega})=\text{DTFT}[h(n)]$ , 以此检验是否满足指标要求, 否则要重新设计, 这和低通滤波器的讨论一样。

## 七、线性相位 FIR 带通滤波器的设计

(1) 其理想线性相位带通滤波器的频率响应为

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega\tau}, & 0 < \omega_1 \leq |\omega| \leq \omega_2 < \pi \\ 0, & \text{其他 } \omega \end{cases} \quad (7-75)$$

其中  $\tau = \frac{N-1}{2}$ 。

(2) 求此滤波器的单位冲激响应  $h_d(n)$  为

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega}) e^{j\omega n} d\omega = \frac{1}{2\pi} \left[ \int_{-\omega_1}^{-\omega_2} e^{j\omega(n-\tau)} d\omega + \int_{\omega_1}^{\omega_2} e^{j\omega(n-\tau)} d\omega \right]$$

$$= \begin{cases} \frac{1}{\pi(n-\tau)} \{ \sin[(n-\tau)\omega_2] - \sin[(n-\tau)\omega_1] \}, & n \neq \tau \\ \frac{1}{\pi} (\omega_2 - \omega_1), & n = \tau \end{cases} \quad (7-76)$$

这里, 当  $\omega_1 = 0, \omega_2 = \omega_c$  时, 即为理想线性相位低通滤波器。其结果与(7-72)式相一致, 当  $\omega_2 = \pi, \omega_1 = \omega_c$  时, 即为理想线性相位高通滤波器。其结果与(7-74)式相一致。

(3) 根据对通带、阻带衰减的要求以及过渡带宽的要求, 可选定窗函数  $w(n)$ , 根据  $w(n)$ , 则可求得所需线性相位带通滤波器

$$h(n) = h_d(n) \cdot w(n)$$

(4) 由  $h(n)$  求实际  $H(e^{j\omega})$ , 即

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n) e^{-j\omega n}$$

(5) 检验是否满足设计要求, 否则重新选用窗形式和(或)  $N$ , 再进行计算。

## 八、线性相位 FIR 数字带阻滤波器的设计

步骤完全相同, 只是理想频率特性有所不同。

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega\tau}, & 0 \leq |\omega| \leq \omega_1, \omega_2 \leq |\omega| \leq \pi \\ 0, & \text{其他 } \omega \end{cases} \quad (7-77)$$

其中  $\tau = \frac{N-1}{2}$ 。同样可得

$$h_d(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega}) e^{j\omega n} d\omega$$

$$= \frac{1}{2\pi} \left[ \int_{-\omega_2}^{-\omega_1} e^{j\omega(n-\tau)} d\omega + \int_{-\omega_1}^{\omega_1} e^{j\omega(n-\tau)} d\omega + \int_{\omega_1}^{\omega_2} e^{j\omega(n-\tau)} d\omega \right]$$

$$= \begin{cases} \frac{1}{\pi(n-\tau)} \{ \sin[(n-\tau)\pi] + \sin[(n-\tau)\omega_1] - \sin[(n-\tau)\omega_2] \}, & n \neq \tau \\ \frac{1}{\pi} (\pi - \omega_1 + \omega_2), & n = \tau \end{cases} \quad (7-78)$$

由理想滤波器的低通公式(7-72), 高通公式(7-74)、带通公式(7-76)式, 以及带阻公式(7-78)可以看出:

- (1) 一个高通滤波器相当于一个全通滤波器减去一个低通滤波器;
- (2) 一个带通滤波器相当于两个低通滤波器相减, 其中一个截止频率为  $\omega_2$ , 另一个截止频率为  $\omega_1$ ;
- (3) 一个带阻滤波器相当于一个低通滤波器(截止频率为  $\omega_1$ )加上一个高通滤波器

(截止频率为  $\omega_2$ )。

同样,线性相位 FIR 带阻滤波器只能采用偶对称单位冲激响应,  $N$  等于奇数来设计,道理与讨论高通滤波器是一样的。这些结论请读者自己结合以上四个公式和基本概念加以分析。

当然,这里指的相加相减都是相当于并联结构。

### 九、线性相位数字线性差分器的设计

我们从表 7-1 看出,利用冲激响应为奇对称的情况可以设计幅度函数随  $\omega$  近似成线性变化的线性差分器。

(1) 设理想线性差分器的频率响应为

$$H_d(e^{j\omega}) = |H_d(\omega)| e^{j\varphi_d(\omega)} = j\omega, \quad |\omega| \leq \pi \quad (7-79)$$

(2) 求  $h_d(n)$ 。因为  $H_d(e^{j\omega})$  是  $\omega$  的虚奇函数,则  $h_d(n)$  应是  $\omega$  的实奇函数,由  $H_d(e^{j\omega})$  可求得  $h_d(n)$  为

$$\begin{aligned} h_d(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} j\omega e^{j\omega n} d\omega = \frac{j}{2\pi} \int_{-\pi}^{\pi} \omega \cos(\omega n) d\omega - \frac{1}{2\pi} \int_{-\pi}^{\pi} \omega \sin(\omega n) d\omega \\ &= -\frac{1}{2\pi} \int_{-\pi}^{\pi} \omega \sin(\omega n) d\omega = \frac{-1}{2\pi n^2} \int_{-\pi}^{\pi} x \sin x dx \end{aligned}$$

用分部积分可得

$$h_d(n) = \begin{cases} \frac{1}{n}(-1)^n, & n \neq 0 \\ 0, & n = 0 \end{cases} \quad (7-80)$$

$n=0$  时  $h_d(n)=0$  是因为  $h_d(n)$  应是一个实的奇对称序列,如果要用线性相位 FIR 滤波器来实现它,就必须有  $(N-1)/2$  个抽样的延时,所以可以得到实际线性差分器的  $h(n)$  为

$$h(n) = h_d\left(n - \frac{N-1}{2}\right) \cdot w(n) = \frac{(-1)^{n-\frac{N-1}{2}}}{n - \frac{N-1}{2}} \cdot w(n), \quad n = 0, 1, \dots, N-1 \quad (7-81)$$

$w(n)$  为窗函数时,窗的点数为  $N$ 。这里看出,  $(N-1)/2$  必须为整数,也就是  $N$  必须为  $N \geq 3$  的奇数。

由所得到的  $h(n)$  可求出线性差分器的频率响应  $H(e^{j\omega}) = |H(\omega)| e^{j\varphi(\omega)}$ 。

由  $H_d(e^{j\omega}) = j\omega$  看出,理想差分器的相位特性

$$\varphi_d(\omega) = \begin{cases} \frac{\pi}{2}, & 0 \leq \omega < \pi \\ -\frac{\pi}{2}, & -\pi \leq \omega < 0 \end{cases}$$

由(7-81)式可看出,设计出的实际线性差分器  $H(e^{j\omega})$  的相位特性  $\varphi(\omega)$  应该多一个线性相位项  $-(\frac{N-1}{2})\omega$ , 即

$$\varphi(\omega) = \begin{cases} \frac{\pi}{2} - \frac{(N-1)\omega}{2}, & 0 \leq \omega < \pi \\ -\frac{\pi}{2} - \frac{(N-1)\omega}{2}, & -\pi \leq \omega < 0 \end{cases} \quad (7-82)$$

由于线性差分器  $H(e^{j\omega})$  是包含  $90^\circ$  移相的线性相位 FIR 滤波器, 因而由表 7-1 知, 它的单位冲激响应必是奇对称的, 且  $N$  必须为奇数。

### 十、线性相位 $90^\circ$ 移相器(离散希尔伯特(Hilbert)变换器)的设计

从表 7-1 看出, 当单位冲激响应奇对称时, 显然能设计满足这种要求的变换器。

设理想  $90^\circ$  移相器为

$$H_d(e^{j\omega}) = \begin{cases} -j, & 0 \leq \omega < \pi \\ j, & -\pi \leq \omega < 0 \end{cases} \quad (7-83)$$

其单位冲激响应为

$$\begin{aligned} h_d(n) &= \frac{1}{2\pi} \int_{-\pi}^{\pi} H_d(e^{j\omega}) e^{jn\omega} d\omega = \frac{1}{2\pi} \int_{-\pi}^0 j e^{jn\omega} d\omega - \frac{1}{2\pi} \int_0^{\pi} j e^{jn\omega} d\omega \\ &= \frac{1 - (-1)^n}{n\pi} = \begin{cases} \frac{2}{n\pi}, & n = \text{奇数} \\ 0, & n = \text{偶数} \end{cases} \end{aligned} \quad (7-84)$$

这个单位冲激响应是奇对称的, 为得到线性相位的  $h_d(n)$ , 必须有  $(N-1)/2$  个抽样的延时, 所以, 所得到的实际的线性相位  $90^\circ$  移相器的单位冲激响应为

$$h(n) = h_d\left(n - \frac{N-1}{2}\right) \cdot w(n) = \frac{2}{\left(n - \frac{N-1}{2}\right)\pi} \cdot w(n), \quad 0 \leq n \leq N-1 \quad (7-85)$$

由(7-84)式知,  $n - \frac{N-1}{2}$  = 奇数时,  $h(n)$  才有非零数值, 且  $(N-1)/2$  必须是整数, 故  $N$  必须是  $\geq 3$  的奇数。按表 7-1, 必须选用单位冲激响应奇对称且  $N$  为奇数的情况来设计线性相位  $90^\circ$  移相器。

### 7.4 频率抽样设计法

窗函数法是从时域出发, 把理想的  $h_d(n)$  用一定形状的窗函数截取成有限长的  $h(n)$ , 以此  $h(n)$  来近似理想的  $h_d(n)$ 。这样得到的频率响应  $H(e^{j\omega})$  逼近于所要求的理想频率响应  $H_d(e^{j\omega})$ 。

频率抽样法则是从频域出发, 把给定的理想频率响应  $H_d(e^{j\omega})$  加以等间隔抽样, 即

$$H_d(e^{j\omega}) \Big|_{\omega=\frac{2\pi k}{N}} = H_d(k)$$

然后以此  $H_d(k)$  作为实际 FIR 数字滤波器的频率特性的抽样值  $H(k)$ , 即令

$$H(k) = H_d(k) = H_d(e^{j\omega}) \Big|_{\omega=\frac{2\pi k}{N}}, \quad k = 0, 1, \dots, N-1 \quad (7-86)$$

知道  $H(k)$  后, 由 DFT 定义, 可以用频域的这  $N$  个抽样值  $H(k)$  来惟一确定有限长序列  $h(n)$ , 而由 3.7 节  $X(z)$  的内插公式知道, 利用这  $N$  个频域抽样值  $H(k)$  同样可求得 FIR 滤波器的系统函数  $H(z)$  及频率响应  $H(e^{j\omega})$ 。这个  $H(z)$  或  $H(e^{j\omega})$  将逼近  $H_d(z)$  或  $H_d(e^{j\omega})$ ,  $H(z)$  和  $H(e^{j\omega})$  的内插公式为

$$H(z) = \frac{1-z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H(k)}{1-W_N^{-k}z^{-1}} \quad (7-87)$$

$$H(e^{j\omega}) = \sum_{k=0}^{N-1} H(k) \Phi\left(\omega - \frac{2\pi k}{N}\right) \quad (7-88)$$

其中  $\Phi(\omega)$  是内插函数

$$\Phi(\omega) = \frac{1}{N} \cdot \frac{\sin\left(\frac{\omega N}{2}\right)}{\sin\left(\frac{\omega}{2}\right)} e^{-j\omega\left(\frac{N-1}{2}\right)} \quad (7-89)$$

将(7-89)式代入(7-88)式,化简后可得

$$H(e^{j\omega}) = \frac{1}{N} e^{-j\left(\frac{N-1}{2}\right)\omega} \sum_{k=0}^{N-1} H(k) e^{-j\frac{\pi k}{N}} \cdot \frac{\sin\left(\frac{\omega N}{2}\right)}{\sin\left(\frac{\omega}{2} - \frac{\pi k}{N}\right)} \quad (7-90a)$$

即

$$H(e^{j\omega}) = e^{-j\left(\frac{N-1}{2}\right)\omega} \sum_{k=0}^{N-1} H(k) \cdot \frac{1}{N} e^{j\frac{\pi k}{N}(N-1)} \cdot \frac{\sin\left[N\left(\frac{\omega}{2} - \frac{\pi k}{N}\right)\right]}{\sin\left(\frac{\omega}{2} - \frac{\pi k}{N}\right)} \quad (7-90b)$$

从内插公式(7-88)看到,在各频率抽样点上,滤波器的实际频率响应是严格地和理想频率响应数值相等,即  $H(e^{j\frac{2\pi k}{N}}) = H(k) = H_d(k) = H_d(e^{j\frac{2\pi k}{N}})$ 。但是在抽样点之间的频率响应则是由各抽样点的加权内插函数的延伸叠加而形成的,因而有一定的逼近误差,误差大小取决于理想频率响应曲线形状,理想频率响应特性变化越平缓,则内插值越接近理想值,逼近误差越小,如图 7-16(b)梯形理想频率特性所示。反之,如果抽样点之间的理想频率特性变化越陡,则内插值与理想值之误差就越大,因而在理想频率特性的不连续点附近,就会产生肩峰和波纹,如图 7-16(a)矩形理想频率特性所示。

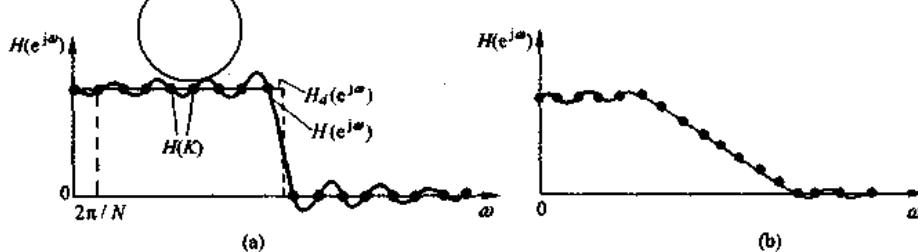


图 7-16 频率抽样的响应

### 一、线性相位的约束

如果我们设计的是线性相位的 FIR 数字滤波器,则其抽样值  $H(k)$  的幅度和相位一定要满足 7.2 节所讨论的约束条件,这些条件已归纳在表 7-1 之中。

对于第一类线性相位 FIR 滤波器,即  $h(n)$  偶对称、 $N$  为奇数时,由表 7-1 第一栏得

$$H(e^{j\omega}) = H(\omega) e^{-j\left(\frac{N-1}{2}\right)\omega} \quad (7-91)$$

其中幅度函数  $H(\omega)$  应为偶对称的,即

$$H(\omega) = H(2\pi - \omega) \quad (7-92)$$

如果抽样值  $H(k) = H(e^{j\frac{2\pi}{N}k})$  也用幅值  $H_k$  (纯标量)与相角  $\theta_k$  表示, 则为

$$H(k) = H(e^{j\frac{2\pi}{N}k}) = H\left(\frac{2\pi}{N}k\right)e^{j\theta_k} = H_k e^{j\theta_k}$$

由(7-91)式可知,  $\theta_k$  必须为

$$\theta_k = -\left(\frac{N-1}{2}\right)\frac{2\pi}{N}k = -k\pi\left(1 - \frac{1}{N}\right) \quad (7-93)$$

$H_k$  必须满足偶对称, 由(7-92)式关系可知

$$H_k = H_{N-k} \quad (7-94)$$

对于第二类线性相位 FIR 滤波器, 即  $h(n)$  偶对称,  $N$  为偶数, 由表 7-1 第二栏知, 其  $H(e^{j\omega})$  的表达式仍为(7-91)式, 但是, 其幅度函数是奇对称的, 即

$$H(\omega) = -H(2\pi - \omega) \quad (7-95)$$

所以, 这时的  $H_k$  也应满足奇对称要求, 即

$$H_k = -H_{N-k} \quad (7-96)$$

而  $\theta_k$  则与(7-93)完全一样。

对于第三类及第四类线性相位正交变换 FIR 滤波器, 幅度和相位同样应满足表 7-1 的约束关系。

## 二、频率抽样的两种方法

对  $H_d(e^{j\omega})$  进行频率抽样, 就是在  $z$  平面上单位圆上的  $N$  个等间隔点上抽取出频率响应值。在单位圆上可以有两种抽样方式, 第一种是第一个抽样点在  $\omega=0$  处(或在  $z=e^{j0}=1$  处), 第二种是第一个抽样点在  $\omega=\pi/N$  处(或在  $z=e^{j\frac{\pi}{N}}$  处), 每种方式可分为  $N$  是偶数与  $N$  是奇数两种, 如图 7-17 所示。

第一种频率抽样(I型)就是上面讨论的抽样, 即

$$H(k) = H_d(k) = H_d(e^{j\omega}) \Big|_{\omega=\frac{2\pi}{N}k}, \quad 0 \leq k \leq N-1 \quad (7-97)$$

其内插公式仍和(7-87)式、(7-90a)式相同。

第二种频率抽样(II型)满足

$$\begin{aligned} H(k) &= H_d(k) = H_d(e^{j\omega}) \Big|_{\omega=\frac{\pi}{N}+\frac{2\pi}{N}k=\frac{\pi}{N}(k+\frac{1}{2})} = \sum_{n=0}^{N-1} h(n) e^{-j\frac{2\pi}{N}(k+\frac{1}{2})n} \\ &= \sum_{n=0}^{N-1} [h(n) e^{-j\frac{\pi}{N}n}] e^{-j\frac{2\pi}{N}nk}, \quad 0 \leq k \leq N-1 \end{aligned} \quad (7-98)$$

求离散傅里叶反变换, 可得

$$h(n) e^{-j\frac{\pi}{N}n} = \frac{1}{N} \sum_{k=0}^{N-1} H(k) e^{j\frac{2\pi}{N}nk}$$

由此得出

$$h(n) = \frac{1}{N} \sum_{k=0}^{N-1} H(k) e^{j\frac{2\pi}{N}(k+\frac{1}{2})n} \quad (7-99)$$

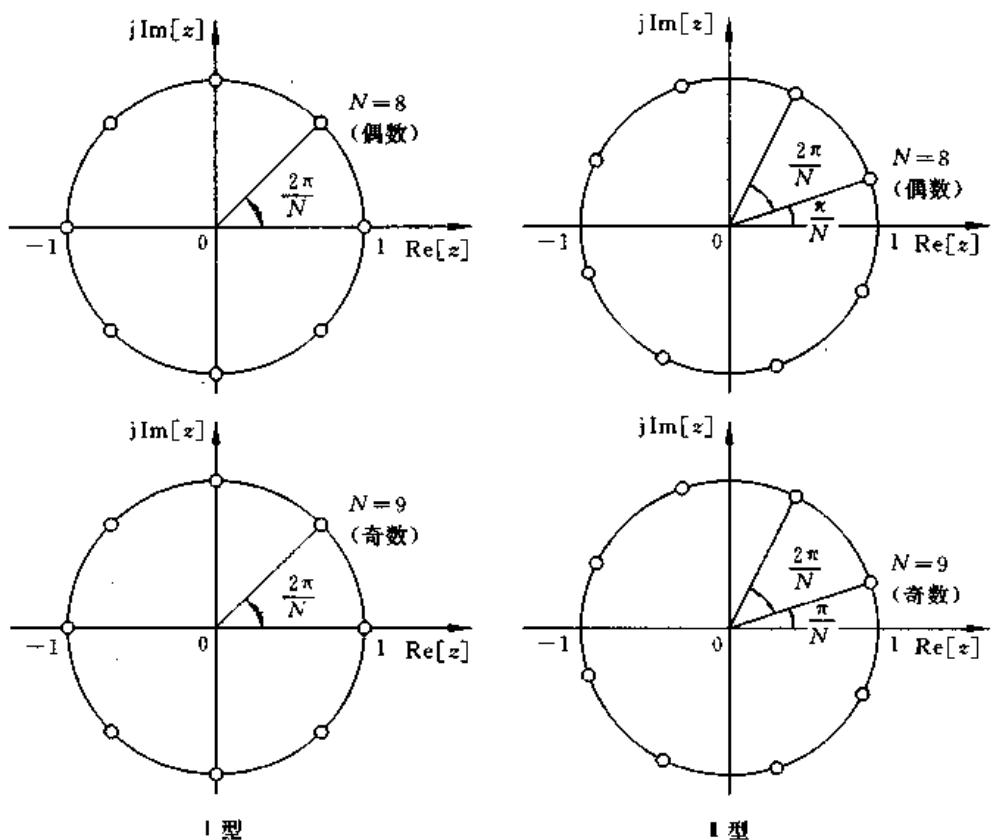


图 7-17 两种频率抽样(I型, II型)

求此  $h(n)$  的  $z$  变换, 可得

$$\begin{aligned} H(z) &= \sum_{n=0}^{N-1} \left[ \frac{1}{N} \sum_{k=0}^{N-1} H(k) e^{j\frac{2\pi}{N}(k+\frac{1}{2})n} \right] z^{-n} \\ &= \frac{1+z^{-N}}{N} \sum_{k=0}^{N-1} \frac{H(k)}{1 - e^{j\frac{2\pi}{N}(k+\frac{1}{2})} z^{-1}} \end{aligned} \quad (7-100)$$

在单位圆  $z=e^{j\omega}$  上计算, 即得频率响应

$$H(e^{j\omega}) = \frac{\cos\left(\frac{\omega N}{2}\right)}{N} e^{-j\left(\frac{N-1}{2}\right)\omega} \sum_{k=0}^{N-1} \frac{H(k) e^{-j\frac{\pi}{N}(k+\frac{1}{2})}}{j \sin\left[\frac{\omega}{2} - \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]} \quad (7-101)$$

(7-100)式及(7-101)式两式就是由第二种理想频率抽样得到 FIR 滤波器系统函数  $H(z)$  和频率响应  $H(e^{j\omega})$  的内插公式。

为了设计上的方便, 在讨论频率抽样法时, 我们可将频率抽样  $H(k)$  表示成幅度响应  $|H(k)|$  (永为正值) 及相位响应  $\theta(k)$  的形式, 即

$$H(k) = |H(k)| e^{j\theta(k)}, \quad k = 0, 1, \dots, N-1 \quad (7-102)$$

这样来导出第一种及第二种频率抽样的设计公式。

### 三、线性相位第一种频率抽样

由于

$$H(k) = \sum_{n=0}^{N-1} h(n) e^{-j\frac{2\pi}{N}nk}$$

当  $h(n)$  为实数时, 满足

$$H(k) = H^*((N-k))_N R_N(k) = H^*(N-k) \quad (7-103)$$

注意, 采用  $H^*((N-k))_N R_N(k) = H^*(N-k)$  的简化表达式时, 对  $k=0$  一点要作说明, 因为按定义  $X(k)$  只有  $N$  个值  $0 \leq k \leq N-1$ , 故  $k=0$  时应为  $H^*((N-k))_N = H^*(0)$ , 而不是  $H^*(N)$ 。因为已经习惯于把  $H(k)$  认为是分布在  $N$  等分圆周上, 故一定满足  $H(N)=H(0)$ , 所以可采用简化的表达式。以后凡遇到  $H(N)$  都应理解为  $H((N))_N = H(0)$ 。由此得出

$$\bigcirc |H(k)| = |H(N-k)| \quad (7-104)$$

$$\theta(k) = -\theta(N-k) \quad (7-105)$$

也就是说,  $H(k)$  的模  $|H(k)|$  以  $k=N/2$  为对称中心呈偶对称,  $H(k)$  的相角  $\theta(k)$  以  $k=N/2$  为对称中心呈奇对称。再利用线性相位的条件  $\theta(e^{jk\omega}) = -\frac{N-1}{2}\omega$ , 即可得到:

当  $N$  为奇数时, 有

$$\theta(k) = \begin{cases} -\frac{2\pi}{N}k\left(\frac{N-1}{2}\right), & k = 0, 1, \dots, \frac{N-1}{2} \\ \frac{2\pi}{N}(N-k)\left(\frac{N-1}{2}\right), & k = \frac{N+1}{2}, \dots, N-1 \end{cases} \quad (7-106)$$

当  $N$  为偶数时, 有

$$\theta(k) = \begin{cases} -\frac{2\pi}{N}k\left(\frac{N-1}{2}\right), & k = 0, 1, \dots, \left(\frac{N}{2}-1\right) \\ \frac{2\pi}{N}(N-k)\left(\frac{N-1}{2}\right), & k = \left(\frac{N}{2}+1\right), \dots, N-1 \\ 0, & k = \frac{N}{2} \end{cases} \quad (7-107)$$

$$H\left(\frac{N}{2}\right) = 0 \quad (7-108)$$

(7-108) 式是由表 7-1 情况 2 得来, 当  $N$  为偶数时的线性相位 FIR 滤波器在  $\omega=\pi$  (即  $k=N/2$ ) 处,  $H(e^{jk\omega})=0$ 。

由此可知, 当  $N$  为奇数时, 有

$$H(k) = \begin{cases} |H(k)| e^{-j\frac{2\pi}{N}k\left(\frac{N-1}{2}\right)}, & k = 0, 1, \dots, \frac{N-1}{2} \\ |H(N-k)| e^{j\frac{2\pi}{N}(N-k)\left(\frac{N-1}{2}\right)}, & k = \frac{N+1}{2}, \dots, N-1 \end{cases} \quad (7-109)$$

同样, 当  $N$  为偶数时, 有

$$H(k) = \begin{cases} |H(k)| e^{-j\frac{2\pi}{N}k\left(\frac{N-1}{2}\right)}, & k = 0, 1, \dots, \left(\frac{N}{2}-1\right) \\ 0, & k = \frac{N}{2} \\ |H(N-k)| e^{j\frac{2\pi}{N}(N-k)\left(\frac{N-1}{2}\right)}, & k = \left(\frac{N}{2}+1\right), \dots, N-1 \end{cases} \quad (7-110)$$

把这两关系式(7-109)式及(7-110)式分别代入(7-90b)式,经过化简,可得:

对线性相位第一种频率抽样,当  $N$  为奇数时

$$H(e^{j\omega}) = e^{-j(\frac{N-1}{2})\omega} \left\{ \frac{|H(0)| \sin\left(\frac{\omega N}{2}\right)}{N \sin\left(\frac{\omega}{2}\right)} + \sum_{k=1}^{\frac{N-1}{2}} \frac{|H(k)|}{N} \left[ \frac{\sin\left[N\left(\frac{\omega}{2} - \frac{\pi}{N}k\right)\right]}{\sin\left(\frac{\omega}{2} - \frac{\pi}{N}k\right)} + \frac{\sin\left[N\left(\frac{\omega}{2} + \frac{\pi}{N}k\right)\right]}{\sin\left(\frac{\omega}{2} + \frac{\pi}{N}k\right)} \right] \right\} \quad (7-111)$$

当  $N$  为偶数时

$$H(e^{j\omega}) = e^{-j(\frac{N-1}{2})\omega} \left\{ \frac{|H(0)| \sin\left(\frac{\omega N}{2}\right)}{N \sin\left(\frac{\omega}{2}\right)} + \sum_{k=1}^{\frac{N}{2}-1} \frac{|H(k)|}{N} \times \left[ \frac{\sin\left[N\left(\frac{\omega}{2} - \frac{\pi}{N}k\right)\right]}{\sin\left(\frac{\omega}{2} - \frac{\pi}{N}k\right)} + \frac{\sin\left[N\left(\frac{\omega}{2} + \frac{\pi}{N}k\right)\right]}{\sin\left(\frac{\omega}{2} + \frac{\pi}{N}k\right)} \right] \right\} \quad (7-112)$$

#### 四、线性相位第二种频率抽样

由于



$$H(k) = \sum_{n=0}^{N-1} h(n) e^{-j\frac{2\pi}{N}(k+\frac{1}{2})n}$$

当  $h(n)$  是实数时,满足

$$H(k) = H^*((N-1-k))_N R_N(k) = H^*(N-1-k) \quad (7-113)$$

也就是

$$|H(k)| = |H(N-1-k)| \quad (7-114)$$

$$\theta(k) = -\theta(N-1-k) \quad (7-115)$$

此时,频率抽样的幅度  $|H(k)|$  以  $k = \frac{N-1}{2}$  为对称中心呈偶对称,  $H(k)$  的相角  $\theta(k)$  以  $k = \frac{N-1}{2}$  为对称中心呈奇对称。再利用线性相位的条件  $\theta(e^{j\omega}) = -\frac{N-1}{2}\omega$ , 即可得到:

当  $N$  为奇数时

$$\theta(k) = \begin{cases} -\frac{2\pi}{N}\left(k + \frac{1}{2}\right)\left(\frac{N-1}{2}\right), & k = 0, 1, \dots, \frac{N-3}{2} \\ 0, & k = \frac{N-1}{2} \\ \frac{2\pi}{N}\left(N-k-\frac{1}{2}\right)\left(\frac{N-1}{2}\right), & k = \frac{N+1}{2}, \dots, N-1 \end{cases} \quad (7-116)$$

当  $N$  为偶数时

$$\theta(k) = \begin{cases} -\frac{2\pi}{N} \left( k + \frac{1}{2} \right) \left( \frac{N-1}{2} \right), & k = 0, 1, \dots, \left( \frac{N}{2} - 1 \right) \\ \frac{2\pi}{N} \left( N - k - \frac{1}{2} \right) \left( \frac{N-1}{2} \right), & k = \frac{N}{2}, \dots, N-1 \end{cases} \quad (7-117)$$

由此得出,当  $N$  为奇数时

$$H(k) = \begin{cases} |H(k)| e^{-j\frac{2\pi}{N}(k+\frac{1}{2})(\frac{N-1}{2})}, & k = 0, 1, \dots, \frac{N-3}{2} \\ |H(\frac{N-1}{2})|, & k = \frac{N-1}{2} \\ |H(N-1-k)| e^{j\frac{2\pi}{N}(N-k-\frac{1}{2})(\frac{N-1}{2})}, & k = \frac{N+1}{2}, \dots, N-1 \end{cases} \quad (7-118)$$

当  $N$  为偶数时

$$H(k) = \begin{cases} |H(k)| e^{-j\frac{2\pi}{N}(k+\frac{1}{2})(\frac{N-1}{2})}, & k = 0, 1, \dots, \left( \frac{N}{2} - 1 \right) \\ |H(N-1-k)| e^{j\frac{2\pi}{N}(N-k-\frac{1}{2})(\frac{N-1}{2})}, & k = \frac{N}{2}, \dots, N-1 \end{cases} \quad (7-119)$$

把这两关系式(7-118)式及(7-119)式分别代入(7-101)式,经过化简,则有:

线性相位第二种频率抽样,当  $N$  为奇数时

$$H(e^{j\omega}) = e^{-j(\frac{N-1}{2})\omega} \left\{ \frac{|H(\frac{N-1}{2})|}{N} \frac{\cos(\frac{\omega N}{2})}{\cos(\frac{\omega}{2})} + \sum_{k=0}^{\frac{N-3}{2}} \frac{|H(k)|}{N} \times \left( \frac{\sin\left\{N\left[\frac{\omega}{2} - \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]\right\}}{\sin\left[\frac{\omega}{2} - \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]} + \frac{\sin\left\{N\left[\frac{\omega}{2} + \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]\right\}}{\sin\left[\frac{\omega}{2} + \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]} \right) \right\} \quad (7-120)$$

当  $N$  为偶数时

$$H(e^{j\omega}) = e^{-j(\frac{N-1}{2})\omega} \left\{ \sum_{k=0}^{\frac{N-1}{2}} \frac{|H(k)|}{N} \left[ \frac{\sin\left\{N\left[\frac{\omega}{2} - \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]\right\}}{\sin\left[\frac{\omega}{2} - \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]} + \frac{\sin\left\{N\left[\frac{\omega}{2} + \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]\right\}}{\sin\left[\frac{\omega}{2} + \frac{\pi}{N}\left(k + \frac{1}{2}\right)\right]} \right] \right\} \quad (7-121)$$

(7-111)式、(7-112)式、(7-120)式、(7-121)式四个式子都可在线性相位 FIR 滤波器的最佳化过程中使用,  $N$  的奇、偶,用第一种还是第二种频率抽样,这要由使用者选择,并且主要取决于待设计的滤波器。

应该指出,这里的频率抽样法设计与第五章所讨论的频率抽样结构并不是一回事,当然二者的理论根据都是第三章 3.7 节的频域抽样理论。应用频域抽样理论建立的 FIR 滤波器结构,对任何 FIR 系统函数都能采用。而本节所讨论的频率抽样设计法,只涉及设计 FIR 滤波器的系统函数,并不涉及滤波器的结构,它可以用任何型的结构来实现,可

以用频率抽样结构,可以用横向结构,也可以用级联结构等来实现。

### 五、过渡带抽样的优化设计

为了提高逼近质量,使逼近误差更小,也就是减小在通带边缘由于抽样点的陡然变化而引起的起伏振荡(这种起伏振荡使阻带内最小衰减变小,例如从衰减30dB变小为衰减20dB)。和窗口法的平滑截断一样,这里是使理想频率响应的不连续点的边缘加上一些过渡的抽样点(在这些点上抽样的最佳值由计算机算出),从而增加过渡带,减小频带边缘的突变,也就减小了起伏振荡,增大了阻带最小衰减。这些抽样点上的取值不同,效果也就不同,因为由(7-88)式看出,每一个频率抽样值,都要产生一个与常数  $\sin\left(\frac{N\omega}{2}\right)/\sin\left(\frac{\omega}{2}\right)$  成正比且在频率上位移  $\frac{2\pi k}{N}$  的频率响应,而 FIR 滤波器的频率响应就是各  $H(k)$  与相应的内插函数  $\varphi\left(\omega - \frac{2\pi k}{N}\right)$  相乘后的线性组合。如果精心设计过渡带的抽样值,就有可能使它的有用频带(通带、阻带)的波纹得以减小。从而设计出较好的滤波器。一般过渡带取一、二、三点抽样值即可得到满意结果。在低通设计中,不加过渡抽样点时,阻带最小衰减为-20dB,一点过渡抽样的最优设计,阻带最小衰减可提高到-40dB 到-54dB 左右,二点过渡抽样的最优设计可达-60dB 到-75dB 左右,而加三点过渡抽样的最优设计则可达-80dB 到-95dB 左右。

加过渡抽样点的示意如图 7-18 所示。

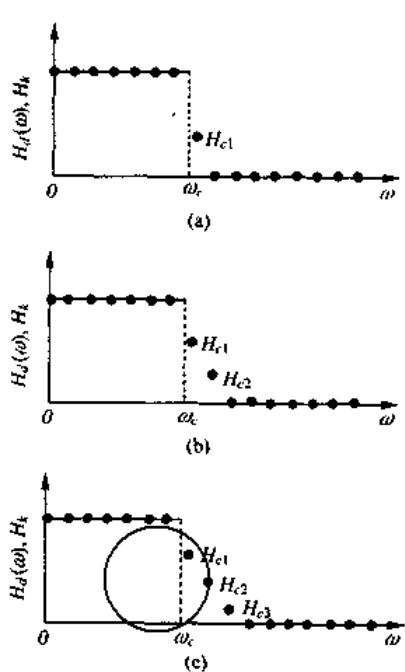


图 7-18 加过渡抽样点(过渡带)  
(a) 一点过渡带;(b) 二点过渡带;(c) 三点过渡带

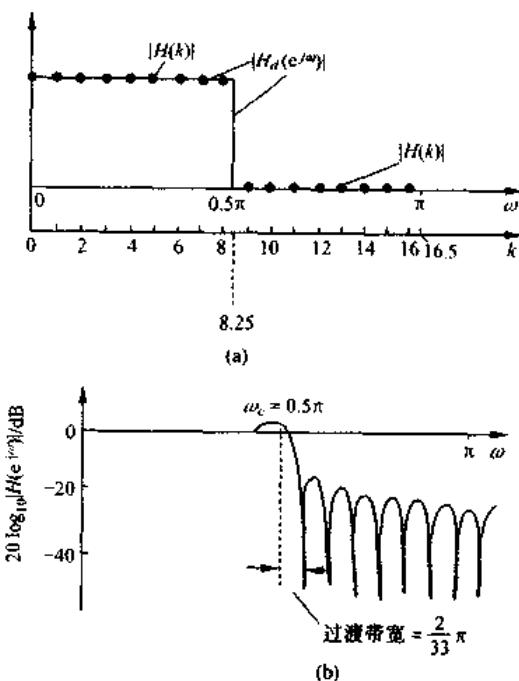


图 7-19 频率抽样法设计低通滤波器例子  
(a)  $|H_d(e^{j\omega})|$  和  $|H(k)|$  ( $N=33$ );(b) 设计结果

频率抽样法的优点是可以在频域直接设计,并且适合于最优化设计,缺点是抽样频率

只能等于  $2\pi/N$  的整数倍(第一种频率抽样)或等于  $2\pi/N$  的整数倍加上  $\pi/N$ (第二种频率抽样),因而不能确保截止频率  $\omega_c$  的自由取值。要想实现自由地选择截止频率,必须增加抽样点数  $N$ ,但这又使计算量加大。

[例 7-2] 利用频率抽样法,设计一个低通 FIR 数字滤波器,其理想频率特性是矩形的,即

$$|H_d(e^{j\omega})| = \begin{cases} 1, & 0 \leq \omega \leq \omega_c \\ 0, & \text{其他 } \omega \end{cases}$$

已知  $\omega_c = 0.5\pi$ ,抽样点数为  $N=33$ ,要求滤波器具有线性相位。

解 根据指标,可画出频率抽样后的  $H(k)$  序列,如图 7-19(a) 所示。由于  $|H(k)|$  是对称于  $\omega=\pi$  的,我们又只对  $0 \leq \omega \leq \pi$  即  $0 \leq k \leq 16$  的区间感兴趣,故可将  $\pi \leq \omega \leq 2\pi$  即  $17 \leq k \leq 32$  的图形略去不画。截止频率  $\omega_c = 0.5\pi$  满足  $\frac{16\pi}{33} \leq \omega_c \leq \frac{17\pi}{33}$ ,按第一种频率抽样方式来设计, $N=33$ ,则

$$|H(k)| = \begin{cases} 1, & 0 \leq k \leq \text{Int}\left[\frac{N\omega_c}{2\pi}\right] = \frac{N-1}{4} \\ 0, & \text{Int}\left[\frac{N\omega_c}{2\pi}\right] + 1 \leq k \leq \frac{N-1}{2} \end{cases}$$

$\text{Int}[\cdot]$  表示取整数部分。将这些值代入(7-111)式,可得

$$H(e^{j\omega}) = e^{-j\omega_0} \left\{ \frac{\sin\left(\frac{33}{2}\omega\right)}{33\sin\left(\frac{\omega}{2}\right)} + \sum_{k=1}^8 \left[ \frac{\sin\left[33\left(\frac{\omega}{2} - \frac{k\pi}{33}\right)\right]}{33\sin\left(\frac{\omega}{2} - \frac{k\pi}{33}\right)} + \frac{\sin\left[33\left(\frac{\omega}{2} + \frac{k\pi}{33}\right)\right]}{33\sin\left(\frac{\omega}{2} + \frac{k\pi}{33}\right)} \right] \right\}$$

按此式计算  $20\log_{10}|H(e^{j\omega})|$  的结果如图 7-19(b) 所示。由图看出,过渡带宽为  $2\pi/33$ ,而最小阻带衰减则约为  $-20$  dB。这一衰减在大多数情况下是不令人满意的。

为了改善频率特性,以满足指标要求,可在通带和阻带交界处安排一个或几个不等于零也不等于 1 的抽样值。本例中用优化算法算出在  $k=9$  处,  $|H(9)| = 0.5$ ,则得图 7-20(a) 所示结果。这相当于加宽过渡带,其宽度为  $2 \times 2\pi/N = 4\pi/33$ ,算出阻带最小衰减约为  $-40$  dB 左右,如图 7-20(b) 所示。

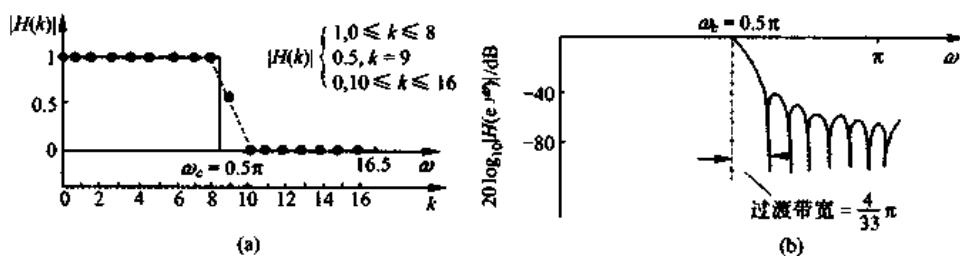


图 7-20 增加过渡带非零抽样值及其影响

(a) 要求的频率响应  $|H_d(e^{j\omega})|$  及其抽样  $|H(k)|$ ; (b) 设计结果

如果要进一步增加阻带衰减,可再添上第二个不等于 1 也不等于零的抽样,这样过渡带又加宽了,如果不允许再增大过渡带宽,而又希望增大阻带衰减,还可增加抽样点数  $N$ ,例如同样是  $\omega_c = 0.5\pi$ ,以  $N=65$  进行抽样,并在  $k=17, k=18$  处插进两个优化的抽样

值 $|H(17)|=0.5886$ , $|H(18)|=0.1065$ ,则其结果如图 7-21 所示,过渡带宽为 $6\pi/65$ ,并没有增加,而阻带最小衰减则可达到 $-60\text{dB}$ ( $N=65$ ),但阶次高了,因而运算量增大了。

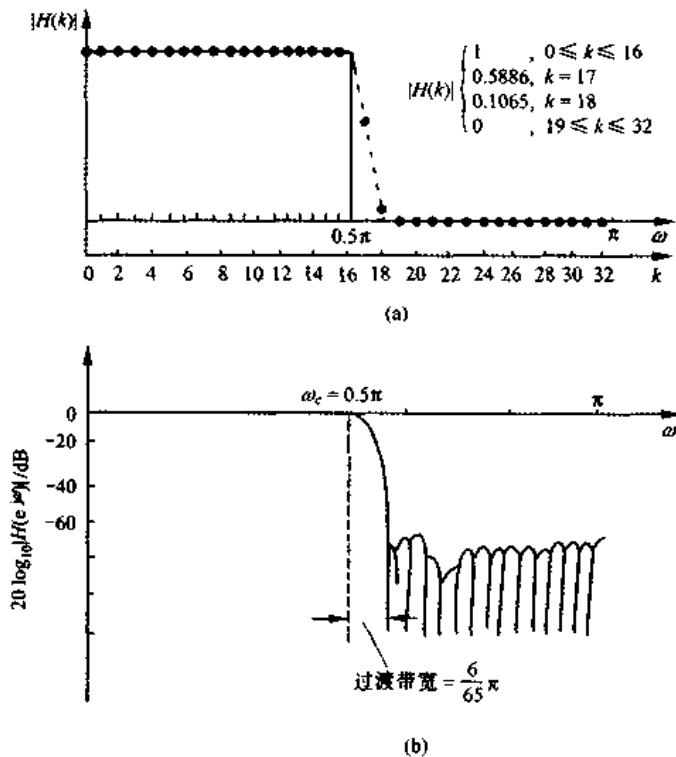


图 7-21 增加抽样点数  $N$  及过渡带频率抽样值的情况  
(a) 要求的频率响应  $|H_d(e^{j\omega})|$  及其抽样  $|H(k)|$ ; (b) 设计结果

## 7.5 设计 FIR 滤波器的最优化方法

在 7.4 节频率抽样设计法的讨论中,曾提到过对过渡带抽样的优化设计办法,这样得到的结果虽然很接近最优化,但却不是最优化设计,因为它只是将过渡带的几个抽样点作为变量,而通带、阻带的其他抽样点的值都是预先规定的常数。最优化设计则是将所有抽样值皆作为变量,在某一优化准则下,通过计算机进行迭代运算,以得到最优的结果。

设计 FIR 滤波器可以有两种最优化准则,即均方误差最小准则和最大误差最小化准则。后面将会看到,后一准则设计出的滤波器在同样阶数时性能更优越,故我们要着重讨论这一准则的设计方法。但是为了进行比较,前一准则也加以介绍。

### 一、均方误差最小准则

这一准则是使误差能量最小,若用  $H_d(e^{j\omega})$  表示要求的频率响应,用  $H(e^{j\omega})$  表示实际得到的滤波器频率响应,以  $E(e^{j\omega})$  表示频率响应误差,即

$$E(e^{j\omega}) = H_d(e^{j\omega}) - H(e^{j\omega}) \quad (7-122)$$

则均方误差为

$$e^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 d\omega = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_d(e^{j\omega}) - H(e^{j\omega})|^2 d\omega \quad (7-123)$$

设计的目的就是选择一组  $h(n) = \mathcal{F}^{-1}[H(e^{j\omega})]$  使得  $e^2$  最小。我们先将(7-123)式中的  $H_d(e^{j\omega})$  和  $H(e^{j\omega})$  分别用它们的冲激响应表示, 即

$$H_d(e^{j\omega}) = \sum_{n=-\infty}^{\infty} h_d(n) e^{-j\omega n}$$

$$H(e^{j\omega}) = \sum_{n=0}^{N-1} h(n) e^{-j\omega n}$$

由于用 FIR 滤波器来逼近, 故  $h(n)$  长度是有限长的。将它们代入(7-122)式, 可得

$$E(e^{j\omega}) = H_d(e^{j\omega}) - H(e^{j\omega}) = \sum_{n=0}^{N-1} [h_d(n) - h(n)] e^{-j\omega n} + \sum_{\text{其他 } n} h_d(n) e^{-j\omega n} \quad (7-124)$$

按照帕塞瓦公式[(2-48)式]有

$$e^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |E(e^{j\omega})|^2 d\omega = \sum_{n=0}^{N-1} |h_d(n) - h(n)|^2 + \sum_{\text{其他 } n} |h_d(n)|^2 \quad (7-125)$$

由此式看出, 等式右边第二个求和式只取决于给定的特性  $h_d(n)$ , 它和设计值  $h(n)$  无关, 故是一个常数, 要使  $e^2$  最小, 就必须使第一项求和式最小, 即希望

$$|h_d(n) - h(n)| = 0, \quad 0 \leq n \leq N-1$$

在这一条件下, 就有

$$e^2 = \min(e^2)$$

也就是说, 要满足

$$h(n) = \begin{cases} h_d(n), & 0 \leq n \leq N-1 \\ 0, & \text{其他 } n \end{cases} \quad (7-126)$$

此式恰好是矩形窗的结果, 所以我们说, 矩形窗设计结果一定满足最小均方误差准则。在 7.3 节的讨论中已看到, 矩形窗虽然过渡带最窄, 但是由于吉布斯(Gibbs)效应, 窗谱的肩峰过大, 造成所设计出的滤波器通带起伏不均匀且过大, 而阻带衰减则过小, 不能满足要求。

## 二、最大误差最小化准则

该准则也称加权切贝雪夫等波纹逼近。由于能得到严格线性相位是 FIR 滤波器的主要优点(有别于 IIR 滤波器), 故我们只讨论线性相位 FIR 滤波器的设计问题。

为了程序的通用性, 使其可用到带通(包括低通、高通、带通、带阻及多带通、多带阻等)滤波器及微分器、离散希尔伯特变换器等不同情况的线性相位 FIR 滤波器的设计中, 因此, 我们首先要将第七章讨论的线性相位 FIR 滤波器的四种情况的频率响应的幅度函数  $H(\omega)$  的表达式统一到一种公式上, 即利用三角恒等式把它们都表示成两项相乘的形式, 其中一项是  $\omega$  的固定函数, 记为  $Q(\omega)$ , 另一项为若干个余弦函数之和, 记为  $P(\omega)$ , 这样表达后, 再用一种算法来求各种情况的最佳逼近。

综合 7.2 节线性相位滤波器的讨论可知, 当  $h(n)$  为偶对称及奇对称,  $N$  为奇、偶的四种情况, 其频率响应表达式可表示为

$$H(e^{j\omega}) = e^{-j(\frac{N-1}{2})\omega} \cdot e^{j(\frac{\pi}{2})L} \cdot H(\omega) \quad (7-127)$$

其中  $H(\omega)$  是幅度函数, 是标量值, 可包括正、负值, 若不考虑线性相位项, 可将四种情况的  $L$  及  $H(\omega)$  列在表 7-4 中。

表 7-4 线性相位 FIR 滤波器四种情况

线性相位 FIR 滤波器四种情况	$L$	$H(\omega)$
I. $N$ 奇数, $h(n)$ 偶对称	0	$\sum_{n=0}^{\frac{(N-1)/2}{2}} a(n) \cos(\omega n)$
II. $N$ 偶数, $h(n)$ 偶对称	0	$\sum_{n=1}^{N/2} b(n) \cos\left[\omega\left(n - \frac{1}{2}\right)\right]$
III. $N$ 奇数, $h(n)$ 奇对称	1	$\sum_{n=1}^{\frac{(N-1)/2}{2}} c(n) \sin(\omega n)$
IV. $N$ 偶数, $h(n)$ 奇对称	1	$\sum_{n=1}^{N/2} d(n) \sin\left[\omega\left(n - \frac{1}{2}\right)\right]$

利用三角恒等式, 可以将  $H(\omega)$  表示成两项相乘的形式, 即

$$H(\omega) = Q(\omega) \cdot P(\omega) \quad (7-128)$$

(1) 对于第 I 种情况, 即  $h(n)$  偶对称,  $N$  为奇数时, 有

$$H(\omega) = \sum_{n=0}^{\frac{N-1}{2}} a(n) \cos(\omega n) = \sum_{n=0}^{\frac{N-1}{2}} \tilde{a}(n) \cos(\omega n) \quad (7-129)$$

由此知

$$\begin{aligned} Q(\omega) &= 1 \\ P(\omega) &= \sum_{n=0}^{\frac{N-1}{2}} \tilde{a}(n) \cos(\omega n) \end{aligned}$$

其中

$$\tilde{a}(n) = a(n) \quad n = 0, 1, \dots, \frac{N-1}{2} \quad (7-130)$$

(2) 对于第 II 种情况, 即  $h(n)$  偶对称,  $N$  为偶数时, 可以证明

$$H(\omega) = \sum_{n=1}^{\frac{N}{2}} b(n) \cos\left[\omega\left(n - \frac{1}{2}\right)\right] = \cos\left(\frac{\omega}{2}\right) \sum_{n=0}^{\frac{N-1}{2}} \tilde{b}(n) \cos(\omega n) \quad (7-131)$$

证: 将恒等式  $\cos A \cdot \cos B = \frac{1}{2} [\cos(A+B) + \cos(A-B)]$  代入上式的等式最右边可得

$$\begin{aligned} \sum_{n=0}^{\frac{N-1}{2}} \tilde{b}(n) \cos\left(\frac{\omega}{2}\right) \cos(\omega n) &= \frac{1}{2} \sum_{n=0}^{\frac{N-1}{2}} \tilde{b}(n) \left\{ \cos\left[\omega\left(n + \frac{1}{2}\right)\right] + \cos\left[\omega\left(n - \frac{1}{2}\right)\right] \right\} \\ &= \frac{1}{2} \sum_{n=1}^{\frac{N}{2}} \tilde{b}(n-1) \cos\left[\omega\left(n - \frac{1}{2}\right)\right] + \frac{1}{2} \sum_{n=0}^{\frac{N-1}{2}} \tilde{b}(n) \cos\left[\omega\left(n - \frac{1}{2}\right)\right] \\ &= \frac{1}{2} \tilde{b}(0) \cos\left[\omega\left(-\frac{1}{2}\right)\right] + \frac{1}{2} \sum_{n=1}^{\frac{N-1}{2}} [\tilde{b}(n) + \tilde{b}(n-1)] \cos\left[\omega\left(n - \frac{1}{2}\right)\right] \end{aligned}$$

$$+ \frac{1}{2} \tilde{b}\left(\frac{N}{2} - 1\right) \cos\left[\omega\left(\frac{N}{2} - \frac{1}{2}\right)\right]$$

将上式与原式(7-131)比较,可得出

$$\left. \begin{aligned} b(1) &= \tilde{b}(0) + \frac{1}{2} \tilde{b}(1) \\ b(n) &= \frac{1}{2} [\tilde{b}(n) + b(n-1)], \quad n = 2, 3, \dots, \frac{N}{2} - 1 \\ b\left(\frac{N}{2}\right) &= \frac{1}{2} \tilde{b}\left(\frac{N}{2} - 1\right) \end{aligned} \right\} \quad (7-132)$$

于是,利用这几个公式,由下向上可由  $b(n)$  ( $n=1, 2, \dots, \frac{N}{2}$ ) 递推求出  $\tilde{b}(n)$  ( $n=0, 1, \dots, \frac{N}{2} - 1$ )。因而有

$$Q(\omega) = \cos\left(\frac{\omega}{2}\right)$$

$$P(\omega) = \sum_{n=0}^{\frac{N-1}{2}} \tilde{b}(n) \cos(\omega n)$$

(3) 对于第Ⅲ种情况,即  $h(n)$  奇对称,  $N$  为奇数时,可证明

$$H(\omega) = \sum_{n=1}^{\frac{N-1}{2}} c(n) \sin(\omega n) = \sin \omega \sum_{n=0}^{\frac{N-3}{2}} \bar{c}(n) \cos(\omega n) \quad (7-133)$$

利用恒等式

$$\cos A \cdot \sin B = \frac{1}{2} [\sin(A+B) - \sin(A-B)]$$

同样可证明以上  $H(\omega)$  的等式((7-133)式),且可得到

$$\left. \begin{aligned} c(1) &= \bar{c}(0) - \frac{1}{2} \bar{c}(2) \\ c(n) &= \frac{1}{2} [\bar{c}(n-1) - \bar{c}(n+1)], \quad n = 2, 3, \dots, \frac{N-5}{2} \\ c(n) &= \frac{1}{2} \bar{c}(n-1), \quad n = \frac{N-3}{2}, \frac{N-1}{2} \end{aligned} \right\} \quad (7-134)$$

同样,利用这几个公式,由下向上可由  $c(n)$  ( $n=1, 2, \dots, \frac{N-1}{2}$ ) 求出  $\bar{c}(n)$  ( $n=0, 1, \dots, \frac{N-3}{2}$ )。因而有

$$\begin{aligned} Q(\omega) &= \sin \omega \\ P(\omega) &= \sum_{n=0}^{\frac{N-3}{2}} \bar{c}(n) \cos(\omega n) \end{aligned}$$

(4) 对于第Ⅳ种情况,即  $h(n)$  奇对称,  $N$  为偶数时,可证明

$$H(\omega) = \sum_{n=1}^{\frac{N}{2}} d(n) \sin\left[\omega\left(n - \frac{1}{2}\right)\right] = \sin\left(\frac{\omega}{2}\right) \sum_{n=0}^{\frac{N-1}{2}} \tilde{d}(n) \cos(\omega n) \quad (7-135)$$

利用恒等式

$$\cos A \cdot \sin B = \frac{1}{2} [\sin(A+B) - \sin(A-B)]$$

同样可证明以上  $H(\omega)$  的等式((7-135)式),且可得到

$$\left. \begin{aligned} d(1) &= \tilde{d}(0) - \frac{1}{2}\tilde{d}(1) \\ d(n) &= \frac{1}{2}[\tilde{d}(n-1) - \tilde{d}(n)], \quad n = 2, 3, \dots, \frac{N}{2}-1 \\ d\left(\frac{N}{2}\right) &= \frac{1}{2}\tilde{d}\left(\frac{N}{2}-1\right) \end{aligned} \right\} \quad (7-136)$$

利用这几个公式,由下向上可由  $d(n) \left( n=1, 2, \dots, \frac{N}{2} \right)$  求得  $\tilde{d}(n) \left( n=0, 1, \dots, \frac{N}{2}-1 \right)$ 。

因而有

$$Q(\omega) = \sin\left(\frac{\omega}{2}\right)$$

$$P(\omega) = \sum_{n=0}^{\frac{N}{2}-1} \tilde{d}(n) \cos(\omega n)$$

我们将以上四种情况的  $H(\omega) = Q(\omega) \cdot P(\omega)$  归纳在表 7-5 中。对于情况 II, III 和 IV,  $Q(\omega)$  或在  $\omega=0$  与  $\omega=\pi$  两处皆为零,或在其中一处为零,由其具体函数确定。

表 7-5 用  $H(\omega) = Q(\omega) \cdot P(\omega)$  表示四种线性相位 FIR 滤波器

线性相位 FIR 滤波器四种情况	$Q(\omega)$	$P(\omega)$
I. $N$ 奇数, $h(n)$ 偶对称	1	$\sum_{n=0}^{\frac{N-1}{2}} \tilde{d}(n) \cos(\omega n)$
II. $N$ 偶数, $h(n)$ 偶对称	$\cos\left(\frac{\omega}{2}\right)$	$\sum_{n=0}^{\frac{N-1}{2}} \tilde{d}(n) \cos(\omega n)$
III. $N$ 奇数, $h(n)$ 奇对称	$\sin\omega$	$\sum_{n=0}^{\frac{N-3}{2}} \tilde{d}(n) \cos(\omega n)$
IV. $N$ 偶数, $h(n)$ 奇对称	$\sin\left(\frac{\omega}{2}\right)$	$\sum_{n=0}^{\frac{N-1}{2}} \tilde{d}(n) \cos(\omega n)$

现在,我们要引出加权切贝雪夫等波纹逼近问题。首先,由于在滤波器设计中通带与阻带误差性能的要求是不一样的,为了统一使用最大误差最小化准则,因而采用误差函数加权的办法,使得不同频段(例如通带与阻带)的加权误差最大值是相等的。设所要求的(已给定)滤波器的频率响应的幅度函数为  $H_d(\omega)$ ,用线性相位四种 FIR 滤波器之一的幅度函数  $H(\omega)$  做逼近函数,设逼近误差的加权函数为  $W(\omega)$ ,则加权逼近误差函数定义为

$$E(\omega) = W(\omega)[H_d(\omega) - H(\omega)] \quad (7-137)$$

由于不同频带中误差函数  $[H_d(\omega) - H(\omega)]$  的最大值不一样,故不同频带中  $W(\omega)$  值可以不同,在公差要求严的频带上可以采用较大的加权值,而公差要求低的频带上,加权值可取较小值。这样使得在各频带上的加权误差  $E(\omega)$  要求一致(即最大值一样)。

将(7-128)式代入(7-137)式,得

$$E(\omega) = W(\omega)[H_d(\omega) - P(\omega)Q(\omega)] = W(\omega)Q(\omega)\left[\frac{H_d(\omega)}{Q(\omega)} - P(\omega)\right] \quad (7-138)$$

最后这一等式,除了在  $\omega=0$  和  $\omega=\pi$  的一处或同时在二处[视  $Q(\omega)$  的情况而定]外,对其他任何频率都是正确的。

令

$$\hat{H}_d(\omega) = \frac{H_d(\omega)}{Q(\omega)}, \quad \hat{W}(\omega) = W(\omega)Q(\omega) \quad (7-139)$$

则(7-138)式可化为

$$E(\omega) = \hat{W}(\omega)[\hat{H}_d(\omega) - P(\omega)] \quad (7-140)$$

这就是加权逼近误差函数的最终表达式。利用这一表达式,线性相位 FIR 滤波器的加权切贝雪夫等波纹逼近问题可看成是求一组系数  $a(n)$  [ $a(n)$  可表示  $\tilde{a}(n)$  或  $\tilde{b}(n)$  或  $\tilde{c}(n)$  或  $\tilde{d}(n)$ ],使其在完成逼近的各个频带上(这里只指通带或阻带,不包括过渡带), $E(\omega)$  的最大绝对值达到极小,如果用  $\|E(\omega)\|$  表示这个极小值,则

$$\|E(\omega)\| = \min_{\text{各系数}} [\max_{\omega \in A} |E(\omega)|] \quad (7-141)$$

其中  $A$  表示所研究的各通带和阻带。

对于线性相位 FIR 滤波器设计的切贝雪夫等波纹逼近法,帕克斯(Parks)和麦克莱伦(McClellan)引进逼近理论的一个定理,得出如下的交错定理。

**交错定理** 若  $P(\omega)$  是  $r$  个余弦函数的线性组合,即

$$P(\omega) = \sum_{n=0}^{r-1} a(n) \cos(\omega n) \quad (7-142)$$

$A$  是  $(0, \pi)$  内的一个闭区间(包括各通带和阻带,但不包括过渡带), $\hat{H}_d(\omega)$  是  $A$  上的一个连续函数,那么, $P(\omega)$  是  $\hat{H}_d(\omega)$  的唯一地和最佳的加权切贝雪夫逼近的充分必要条件是:加权逼近误差函数  $E(\omega)$  在  $A$  中至少有  $(r+1)$  个极值点,即  $A$  中至少有  $(r+1)$  个点  $\omega_i$ ,且  $\omega_1 < \omega_2 < \omega_3 < \dots < \omega_r < \omega_{r+1}$ ,使得

$$E(\omega_i) = -E(\omega_{i+1}), \quad i = 1, 2, \dots, r$$

并且

$$|E(\omega_i)| = \max_{\omega \in A} |E(\omega)|$$

这一逼近可以用图 7-22 来说明。

设所要求的滤波器频率响应为

$$H_d(e^{j\omega}) = \begin{cases} 1, & 0 \leq \omega \leq \omega_c \\ 0, & \omega_u \leq \omega \leq \pi \end{cases}$$

式中  $\omega_c$  为通带频率,  $\omega_u$  为阻带频率。现在的任务是,寻找一个  $H(e^{j\omega})$ ,使其在通带和阻带内最佳地一致逼近  $H_d(e^{j\omega})$ 。图 7-22 中,  $\delta_1$  为通带纹波峰值,  $\delta_2$  为阻带纹波峰值。这样,对设计的低通数字滤波器  $H(e^{j\omega})$ ,共有五个参数,即  $\omega_c, \omega_u, \delta_1, \delta_2$  和相应的单位抽样响应的长度  $N$ 。根据上述交错点组定理,如果  $H(e^{j\omega})$  是对  $H_d(e^{j\omega})$  的最佳一致逼近,那么  $H(e^{j\omega})$  在通带和阻带内应具有如图 7-22 的等纹波性质。所以最佳一致逼近有时又称等波纹逼近。

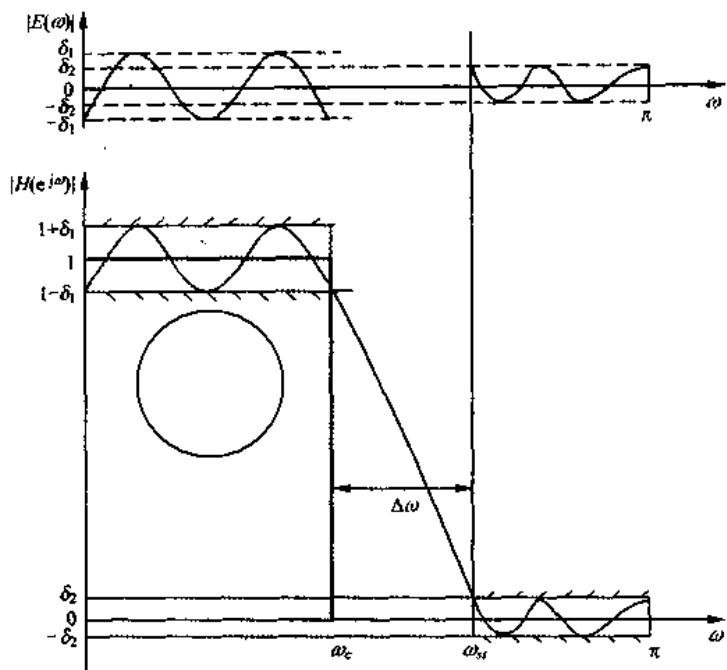


图 7-22 低通数字滤波器的一致逼近

交错定理非常重要,它是很多切贝雪夫最优逼近算法的基础。在讨论最优算法前,我们先引出线性相位 FIR 滤波器极值点数目的重要结果。

#### 线性相位 FIR 滤波器的频率响应极值点数目的限制

上面交错定理已指出,最优线性相位 FIR 滤波器的加权逼近误差函数  $E(\omega)$  至少有  $(r+1)$  个极值,而  $r$  是用于逼近的余弦函数的个数。

$E(\omega)$  的极值包括以下两种:

(1)  $H(\omega)$  的极值点(在大多数情况下,  $H(\omega)$  的极值也是  $E(\omega)$  的极值);

(2)  $E(\omega)$  单有的极值点(它不属于  $H(\omega)$  的)。

两种极点数目之和就是  $E(\omega)$  极值点的最大数目。

先来看  $H(\omega)$  的极值点数目。我们以第 I 种情况  $H(\omega)$  为例推导其极值点最大数目。

将  $H(\omega)$  表达式中的求和的上限用  $r-1$  表示,即表示成  $r$  个余弦之和(对于第 I 种情况,  $r=\frac{N+1}{2}$ ),即

$$H(\omega) = \sum_{n=0}^{r-1} \bar{a}(n) \cos(\omega n) \quad (7-143)$$

为了便于求  $H(\omega)$  的极值点( $0 \leq \omega \leq \pi$ )的最大数目,我们把余弦的倍角化为余弦的多项式形式,即

$$\cos(\omega n) = \sum_{m=0}^n a_{mn} (\cos \omega)^m, \quad 0 \leq \omega \leq \pi \quad (7-144)$$

将(7-144)式代入(7-143)式,可得

$$H(\omega) = \sum_{n=0}^{r-1} \bar{a}(n) \sum_{m=0}^n a_{mn} (\cos \omega)^m = \sum_{n=0}^{r-1} \bar{a}(n) (\cos \omega)^n \quad (7-145)$$

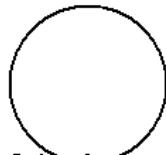
其中  $\bar{a}(n)$  是合并  $(\cos\omega)^n$  的同幂次项系数而得到的,为了求各极值点,取  $H(\omega)$  对  $\omega$  的导数,可得

$$\frac{dH(\omega)}{d\omega} = \sum_{n=1}^{r-1} n\bar{a}(n)(\cos\omega)^{n-1}(-\sin\omega) = \sin\omega \sum_{m=0}^{r-2} \bar{b}(m)(\cos\omega)^m \quad (7-146)$$

其中  $\bar{b}(m) = -(m+1)\bar{a}(m+1)$ 。为了求极值点的最大数目,令  $x = \cos\omega$ ,代入(7-146)式,可得

$$F(x) = \frac{dH(\omega)}{d\omega} \Big|_{\omega=\cos^{-1}x} = \sqrt{1-x^2} \cdot \sum_{m=0}^{r-2} \bar{b}(m)x^m = F_1(x)F_2(x) \quad (7-147)$$

其中



$$F_1(x) = \sqrt{1-x^2}$$

$$F_2(x) = \sum_{m=0}^{r-2} \bar{b}(m)x^m$$

可以看出,当  $x=1$  (相当于  $\omega=\cos^{-1}x=0$ ) 和  $x=-1$  (相当于  $\omega=\cos^{-1}x=\pi$ ) 两处,  $F_1(x)=0$ ; 由于  $F_2(x)$  是  $(r-2)$  阶多项式,故在  $-1 < x < 1$  (即  $0 < \omega < \pi$  范围内),  $F_2(x)$  最多有  $(r-2)$  个零值点。综合这两种零值点,可知在闭区间  $-1 \leq x \leq 1$  上 (即  $0 \leq \omega \leq \pi$  上),  $F(x)$  最多有  $r$  个零值点。所以对线性相位 FIR 滤波器第 I 种情况,  $H(\omega)$  的极值点数目  $N_c$  应满足  $N_c \leq r = \frac{N+1}{2}$ , 同样可求得其他三种情况极值点数目  $N_c$ , 归纳在表 7-6 中。

表 7-6 四种线性相位 FIR 滤波器余弦数目  $r$  及极值点数目  $N_c$  与  $N$  的关系

类 型	$P(\omega)$	$N_c \leq r$
I. $N$ 奇数, $h(n)$ 偶对称	$r = \frac{N+1}{2}$	$N_c \leq \frac{N+1}{2}$
II. $N$ 偶数, $h(n)$ 偶对称	$r = \frac{N}{2}$	$N_c \leq \frac{N}{2}$
III. $N$ 奇数, $h(n)$ 奇对称	$r = \frac{N-1}{2}$	$N_c \leq \frac{N-1}{2}$
IV. $N$ 偶数, $h(n)$ 奇对称	$r = \frac{N}{2}$	$N_c \leq \frac{N}{2}$

其次,我们来看  $E(\omega)$  单独有的极值点(不属于  $H(\omega)$  的)。可以看出,如果要在几个频带上求逼近问题,很显然在每个频率的端点上,误差函数会得到一个极值,而且这些点一般不是  $H(\omega)$  的极值点,但是  $\omega=0$  和  $\omega=\pi$  要除外,因为这两频率下,  $H(\omega)$  常可能有极值。例如,第 I 种情况的低通滤波器,有两个频带(一个通带,一个阻带),其加权误差函数最多能有  $r+2 = \frac{N+5}{2}$  个极值,即  $r = \frac{N+1}{2}$  个  $H(\omega)$  的极点加上位于通带和阻带边界的两个  $E(\omega)$  单有的极值。又如第 I 种情况的带通滤波器(一个通带、两个阻带)的加权误差函数最多能有  $r+4 = \frac{N+9}{2}$  个极值,即  $r = \frac{N+1}{2}$  个  $H(\omega)$  的极值加上位于通带阻带边界的四个  $E(\omega)$  单有的极值。

从以上分析看出,第 I 种情况低通滤波器的误差函数最多能有  $\frac{N+5}{2}$  个极值,而交错定理指出:最优逼近时,其误差函数至少要有  $r+1 = \frac{N+3}{2}$  个极值。这样一来,对所要求

的低通响应的惟一最好逼近来说,误差函数要有  $r+1 = \frac{N+3}{2}$  个或  $r+2 = \frac{N+1}{2} + 2 = \frac{N+5}{2}$  个极值。用  $r+2 = \frac{N+5}{2}$  个极值来实现的滤波器比交错定理所要求的最少的  $r+1 = \frac{N+3}{2}$  个极值还要多一个,把这类滤波器称为最多波纹滤波器,对于低通滤波器,这种最多波纹滤波器又称为超波纹滤波器。

预先知道极值点的最大数目很重要,因为有些设计方法只能设计极值点数目有最大可能值的最多波纹滤波器,而我们所讨论瑞米兹(Remez)算法则可设计任何最优线性相位 FIR 滤波器。它是一种最为实用的最优化算法。

#### 瑞米兹交换算法设计最优线性相位 FIR 滤波器

这一算法是给定单位抽样响应长度  $N$ 、通带截止频率  $\omega_c$  和阻带截止频率  $\omega_a$  来进行设计的,算法已由加权切贝雪夫逼近来描述,可见(7-140)式,此公式中逼近函数  $P(\omega)$  是  $r$  个独立的余弦函数之和,利用交错定理,可给出加权逼近误差函数  $E(\omega)$  应满足的一组必要且充分条件,这样就可得到最大误差最小化的最优逼近——等波纹逼近。

**基于交错定理,最优线性相位 FIR 滤波器的设计步骤如下:**

- (1) 输入部分,包括滤波器性能要求及滤波器类型,前者指的是所需的频率响应的幅度函数  $H_d(\omega)$ ,加权函数  $W(\omega)$  和滤波器单位抽样响应长度  $N$ ,后者是要指出所需设计的是带通(包括低通、带通、高通、带阻等)滤波器或是微分器或是希尔伯特变换器;
- (2) 用公式表示逼近问题,也就是用(7-140)式来表示加权逼近误差  $E(\omega)$ ;
- (3) 用瑞米兹多次交换算法,求逼近问题的解;
- (4) 计算滤波器的单位抽样响应。

这里步骤(1)表示设计者利用算法时必须具有的条件,步骤(2)在前面加权切贝雪夫逼近中已讨论到,而步骤(3)正是瑞米兹交换算法求逼近问题的解。在整个设计程序中,瑞米兹算法是以子程序形式出现的。

最佳线性相位 FIR 滤波器设计算法的框图如图 7-23 所示。它完成上面所说的四个设计步骤,更仔细地说,它包括:① 输入数据、滤波器性能要求及滤波器类型;② 根据滤波器的类型和单位抽样响应长度,确定逼近函数  $[\cos(\cdot)]$  的个数  $r$ ,例如,由表 7-6 知,第 I 种情况线性相位 FIR 滤波器的  $r = \frac{N+1}{2}$ ,  $N$  为单位抽样响应的长度;③ 在  $\omega$  从 0 到  $\pi$  的频率区间,用密集的格点来表示离散频率,两格点的距离可表示成

$$\frac{\pi}{\text{总格点数}} = \frac{\pi}{\text{格点密度} \times r}$$

将格点处频率用下标的频率来表示并赋以标称频率值;④ 用子程序 EFF 和 WATE 分别计算各格点频率上要求的函数值  $H_d(\omega)$  和加权函数值  $W(\omega)$ ;⑤ 用公式表达逼近问题,即将(7-137)式变成(7-140)式,将  $H_d(\omega)$  和  $W(\omega)$  分别变成  $\hat{H}_d(\omega)$  和  $\hat{W}(\omega)$ ;⑥ 瑞米兹交换算法求解逼近问题,这包括设定( $r+1$ )个极值频率的初始假设值(猜想值);⑦ 计算滤波器单位抽样响应;⑧ 把最佳误差和单位抽样响应打印出来。

下面分别讨论⑥⑦两个步骤,即瑞米兹交换算法以及求滤波器单位抽样响应两个框图。

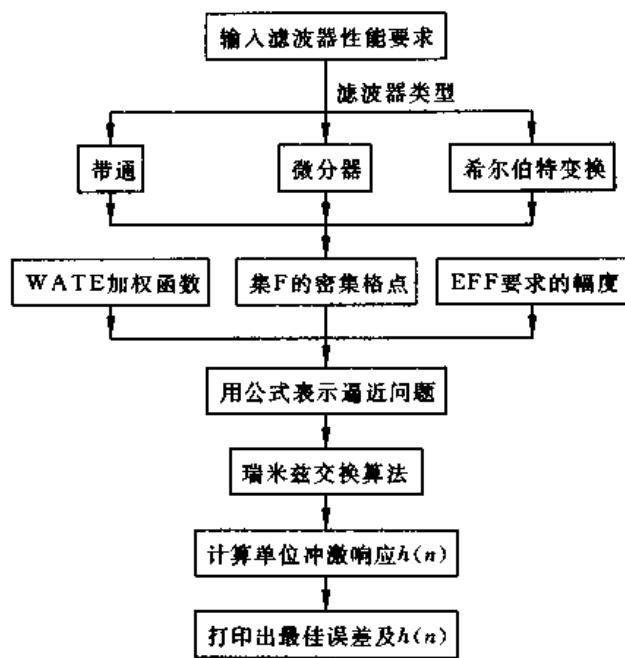


图 7-23 最佳线性相位 FIR 滤波器设计算法框图

### 瑞米兹交换算法

这一算法的流程图如图 7-24 所示。

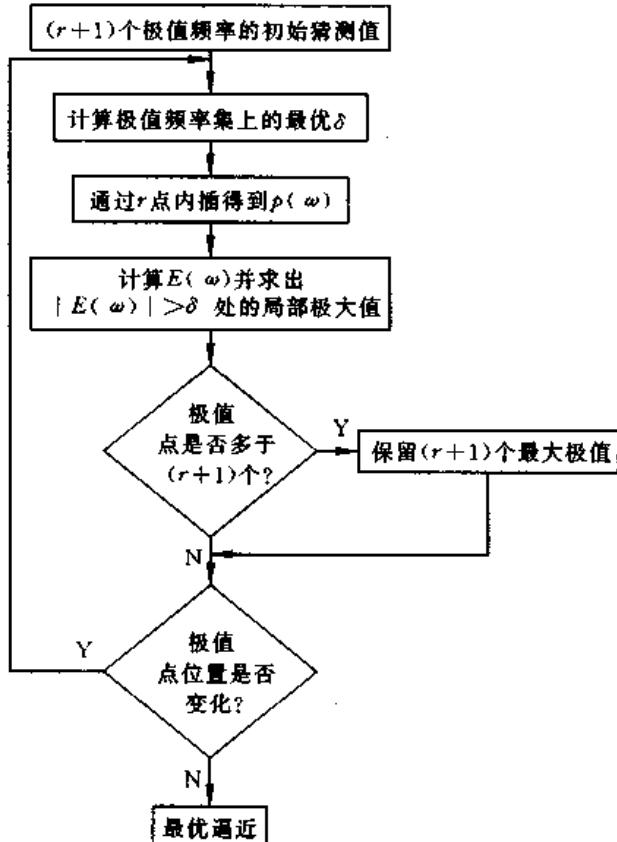


图 7-24 瑞米兹交换算法流程图

第一步,设定( $r+1$ )个极值点频率 $\omega_k(k=0,1,\dots,r)$ 的初始猜测值,它是按等间隔设定的,这些频率位于通带区间 $0 \leq \omega \leq \omega_c$ 和阻带区间 $\omega_a \leq \omega \leq \pi$ 内,由于通带截止频率 $\omega_c$ 及阻带截止频率 $\omega_a$ 是固定的,所以应满足

$$\omega_c = \omega_l, \quad \omega_a = \omega_{l+1}, \quad (0 < l < r)$$

假定这些频率点上的误差函数的值均为 $\delta$ ,其符号正负交错。就是说,按问题的要求,对于给定的一组极值点频率 $\{\omega_k\}(k=0,1,\dots,r)$ ,需要求解以下方程:

$$\hat{W}(\omega_k)[\hat{H}_d(\omega_k) - P(\omega_k)] = (-1)^k \delta, \quad k = 0, 1, \dots, r \quad (7-148)$$

其中

$$P(\omega_k) = \sum_{n=0}^{r-1} \alpha(n) \cos(\omega_k n) \quad (7-149)$$

将(7-148)式写成

$$\hat{H}_d(\omega_k) = \frac{(-1)^k \delta}{\hat{W}(\omega_k)} + P(\omega_k), \quad k = 0, 1, \dots, r \quad (7-150)$$

利用矩阵形式来表示(7-150)式(考虑到(7-149)式)为

$$\begin{bmatrix} 1 & \cos\omega_0 & \cos 2\omega_0 & \cdots & \cos[(r-1)\omega_0] & \frac{1}{\hat{W}(\omega_0)} \\ 1 & \cos\omega_1 & \cdots & \cdots & & \frac{-1}{\hat{W}(\omega_1)} \\ \vdots & \vdots & & & & \vdots \\ 1 & \cos\omega_r & \cdots & \cdots & \cos[(r-1)\omega_r] & \frac{(-1)^r}{\hat{W}(\omega_r)} \end{bmatrix} \cdot \begin{bmatrix} \alpha(0) \\ \alpha(1) \\ \vdots \\ \alpha(r-1) \\ \delta \end{bmatrix} = \begin{bmatrix} \hat{H}_d(\omega_0) \\ \hat{H}_d(\omega_1) \\ \vdots \\ \hat{H}_d(\omega_r) \end{bmatrix} \quad (7-151)$$

此时可对 $r+1$ 个未知数(包括 $r$ 个 $\alpha(n)(n=0,\dots,r-1)$ 及一个 $\delta$ )来求解。这样直接求解既难且慢,更有效办法是用下面的解析式先计算出 $\delta$ :

$$\delta = \frac{\sum_{i=0}^r a_i \hat{H}_d(\omega_i)}{\sum_{i=0}^r (-1)^i a_i / \hat{W}(\omega_i)} \quad (7-152)$$

式中

$$a_i = \prod_{\substack{k=0 \\ k \neq i}}^r \frac{1}{\cos\omega_i - \cos\omega_k} \quad (7-153)$$

第二步,利用(7-150)式算出在 $\omega_0, \omega_1, \dots, \omega_{r-1}$ 上的 $P(\omega_i)$ 值,用 $c_i$ 来表示它,即

$$c_i = \hat{H}_d(\omega_i) - (-1)^i \frac{\delta}{\hat{W}(\omega_i)}, \quad i = 0, 1, \dots, r-1 \quad (7-154)$$

利用重心形式的拉格朗日内插公式得到 $P(\omega)$ 为

$$P(\omega) = \frac{\sum_{i=0}^{r-1} \left( \frac{\beta_i}{\cos\omega - \cos\omega_i} \right) c_i}{\sum_{i=0}^{r-1} \left( \frac{\beta_i}{\cos\omega - \cos\omega_i} \right)} \quad (7-155)$$

式中

$$\beta_i = \prod_{\substack{k=0 \\ k \neq i}}^{r-1} \frac{1}{\cos \omega_i - \cos \omega_k} \quad (7-156)$$

当然,当  $i=r$  时,(7-154)式也成立,因为它也满足(7-150)式。

第三步,在求出  $P(\omega)$  的内插值后,根据(7-140)式,即根据

$$E(\omega) = \hat{W}(\omega)[\hat{H}_d(\omega) - P(\omega)]$$

在密集的频率组上来计算  $E(\omega)$  值。若在这组频率的所有频率上皆有  $|E(\omega)| \leq \delta$ , 则最佳逼近已经得到, 而  $\delta$  是波纹的极值, 且  $\omega_0, \omega_1, \dots, \omega_r$  的初始假设值是交错点组频率, 计算即可结束。但实际上第一次不会恰好得此结果, 在该频率组的某些频率处总会出现  $|E(\omega)| > \delta$ , 这时就找出误差曲线上的  $(r+1)$  个极值频率点, 以它们作为新的极值频率点代替原来的初始猜测值, 又得到新的一组交错点组频率, 然后利用(7-152)式重新计算  $\delta$  值, 并利用上面的有关式子计算  $P(\omega)$  及  $E(\omega)$  值, 这就构成新一轮的迭代。

由于每次迭代得到的新交错点组频率都是  $E(\omega)$  的局部极值点频率, 因此  $\delta$  是递增的, 最后收敛的  $\delta$  是自己的上限、即重复到  $\delta$  与前一次迭代的  $\delta$  值相同时迭代即中止, 此时就得到问题的求解。

为了求出峰值点, 应在通带和阻带上把频率分点取得更密一些, 以便在这些细分点上搜索出峰值点。如果在任一次迭代中  $E(\omega)$  的极值点多于  $(r+1)$  个, 就保留  $|E(\omega)|$  值最大的那  $(r+1)$  个频率作为下次迭代的初始猜测极值点。

迭代最后的结果必然是误差曲线上每个格点频率上的误差值都满足  $|E(\omega)| \leq \delta$ , 而在  $(r+1)$  个极值点频率处的  $|E(\omega)| = \delta$  为最大值, 并且  $E(\omega)$  具有正负交错的符号, 这就表明加权切贝雪夫等波纹逼近已经完成。

应再次指出, 在每次迭代中, 都把通带截止频率  $\omega_c$ 、阻截止频率  $\omega_u$  定为诸极值点频率中的两个频率。所以这一方法是已知  $N$  (即  $r$ ),  $\omega_c$  和  $\omega_u$ , 求最佳的  $\delta$  值。

由(7-137)式及对该式  $[E(\omega)]$  的讨论中知道, 若最后得到的  $|E(\omega)| = \delta = \delta_2$ , 即通带、阻带的加权误差是相同的, 为了使通带实际误差为  $\delta_1$ , 阻带实际误差为  $\delta_2$ , 那么, 就应规定加权函数  $W(\omega)$  为

$$W(\omega) = \begin{cases} \frac{\delta_2}{\delta_1}, & \text{当 } \omega \text{ 在通带中} \\ 1, & \text{当 } \omega \text{ 在阻带中} \end{cases} \quad (7-157)$$

这样可知,  $[\hat{H}_d(\omega) - H(\omega)]$  在通带中最大误差为  $\delta_1$ , 在阻带中最大误差为  $\delta_2$ , 则瑞米兹迭代算法所求得的极值  $|E(\omega)| = \delta$ , 正是要求的  $\delta_2$  的最小值。

如果要求的  $\delta_1$  和  $\delta_2$  值是已知的, 则计算滤波器时, 可固定  $\omega_c$  值, 改变  $\omega_u$  值, 重复进行以上迭代计算, 直到使求出的  $\delta_1, \delta_2$  满足已知数值为止。

#### 计算滤波器的单位抽样响应

由于已求得  $P(\omega)$ , 而  $P(\omega)$  满足

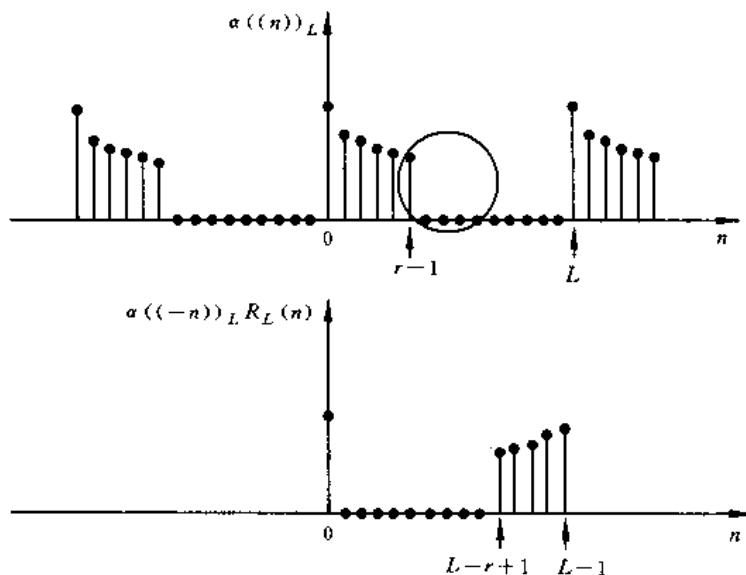
$$P(\omega) = \sum_{n=0}^{r-1} a(n) \cos(\omega n)$$

因为  $a(n)$  是实数, 故



$$P(\omega) = \operatorname{Re} \left[ \sum_{n=0}^{r-1} \alpha(n) e^{-j\omega n} \right] = \operatorname{Re} \{ \mathcal{F}[\alpha(n)] \} \quad (7-158)$$

花体字“ $\mathcal{F}$ ”表示傅里叶变换。想求出滤波器的单位抽样响应  $h(n)$ , 就要求得  $\alpha(n)$ , 可以利用离散傅里叶反变换来求解, 将  $P(\omega)$  在频域抽样, 设  $z$  平面上单位圆(频域)上抽样点数为  $L=2^M$ , (取 2 的幂次是为了 FFT 计算方便)。频域抽样, 时域就要按  $L$  点为周期产生周期延拓, 为了防止混叠效应, 应取  $L=2^M \geq N$ , 可见图 7-25(注意, 这时对四种情况, 线性相位 FIR 滤波器都可满足不混叠的要求), 由(7-158)式可得



$L=2^M$ , 若  $L-r+1 \geq r$  就不混叠, 即  $L \geq 2r-1$ , 第 I 种情况的  $r$  最大, 为  $r=(N+1)/2$  代入可得  $L \geq N$

图 7-25 不产生混叠现象的  $L$  值的示意图

$$P(k) = P(\omega) \Big|_{\omega=\frac{2\pi k}{L}} = \operatorname{Re} \left[ \sum_{n=0}^{r-1} \alpha(n) e^{-j\frac{2\pi k}{L} n} \right] = \operatorname{Re}[A(k)] \quad (7-159)$$

考虑到表 3-3(DFT 性质表)的第 12 条性质  $\alpha(n)$  的 DFT[即  $A(k)$ ]的实部的 IDFT 等于  $\alpha(n)$  的圆周共轭对称分量, 由于没有混叠, 在  $0 \leq n \leq r-1$  范围内, 它就等于  $\alpha(n)$ 。这样, 由(7-159)式求 IDFT[ $P(k)$ ], 即可求得  $\alpha(n)$ , 从而求得滤波器的单位抽样响应  $h(n)$ 。求出  $h(n)$  后, 即可将  $E(\omega)$  和  $h(n)$  打印输出。

## 7.6 FIR 数字滤波器的硬件实现——位串行方式实现

图 7-26 的方框图表示用阵列乘法实现 FIR 数字滤波器的硬件系统。参照此图, 利用(6-286)式  $A = -[\text{编址存储值}] + [\sum_{i=1}^b 2^{-i} (\text{编址存储值})]$ ,  $A$  的计算可概括为以下几步:

(1) 开始将  $(b+1)$  位寄存器 RA, RB, RC, SR1, SR2, …, SRI 都清零,  $2^b$  个编址存储值已存放在 ROM 中;

(2)  $x(n)$  的  $(b+1)$  位二进制表示形式是

$$(-c_0 + \sum_{i=1}^b c_i 2^{-i})_k \quad k = 1, 2, \dots, I$$

$x(n)$  是滤波器的输入值, 将它存放在  $(b+1)$  位的移位寄存器 SR1 中;

(3) 把每个移位寄存器 (SR1, SR2, …, SRI) 的最低有效位  $A_1, A_2, \dots, A_I$  (如图 7-26 所示) 用作 ROM 的地址编码 (长度为  $I$ ), 以便从 ROM 中取出这个地址所决定的编址存储值, 将其寄存到  $(b+1)$  位的寄存器 RA 中;

(4) 用算术逻辑部件 (ALU) 形成  $RC + RA$ , 其结果寄存到  $(b+1)$  位的寄存器 RB 中, 即  $RB = RC + RA$ , 这就产生了第一部分结果;

(5) 如图 7-26 所示的联接方法, 将寄存器 RB 的内容右移一位, 即乘上  $2^{-1}$  ((6-286) 式中的  $2^{-i}$ , 这里  $i=1$ ), 并且保存到寄存器 RC 中;

(6) 将移位寄存器 SR1, SR2, …, SRI 的内容右移一位。

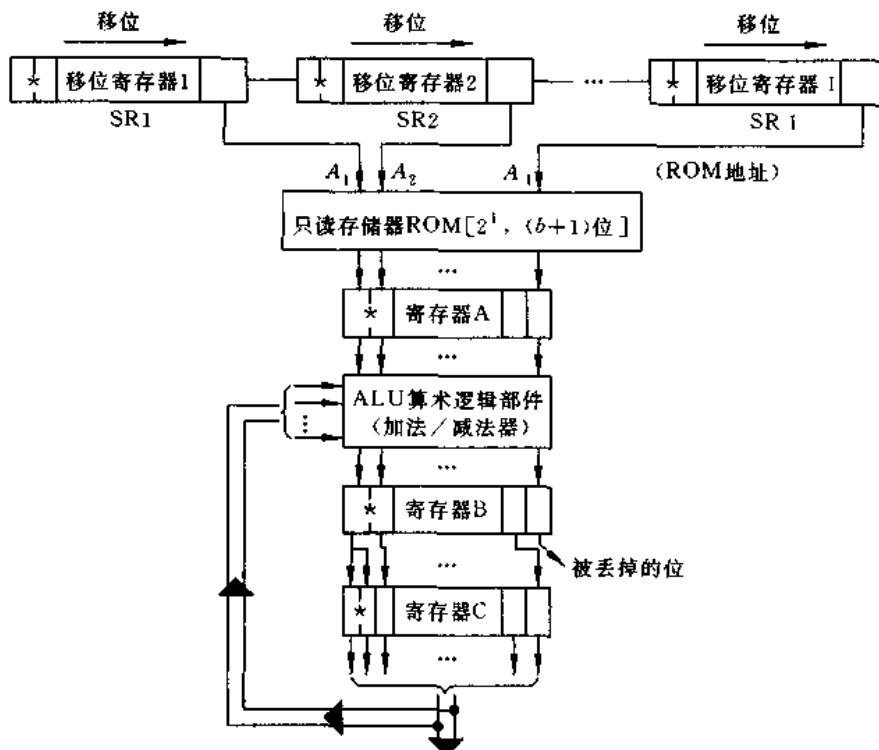


图 7-26 计算阵列乘法的硬件方框图

再重复步骤(3)到(6)的过程  $b$  次, 最后一次 (即第  $(b+1)$  次) 迭代的第(4)步是在 ALU 中形成  $RB = RC - RA$  (相当于  $i=0$  的符号位的运算), 这就是所需结果  $A$ 。因而, 经过第(3)步到第(6)步的  $(b+1)$  次迭代, 最后在寄存器 RB 中的值即为  $A$ 。系统运算到此时, 就将 RC 寄存器清零, 以准备下一次计算。 $A$  的二进制值可以并行地转移到一个锁存寄存器上, 以便去继续进行 D/A(数—模)转换, 从而形成抽样数据信号, 进而经过滤波器后, 得到连续输出信号  $y(t)$ 。

下一个通过模数转换而得到的输入值  $x(n)$  到来时, 又返回执行第(2)步, 然后重复第

(3)步到第(6)步,经 $(b+1)$ 次迭代,可得到新的A值。下面我们用实例来说明系统的运算过程。

[例 7-3] 设滤波器的线性差分方程为

$$y(n) = 0.75x(n) - 0.5x(n-1) \quad (7-160)$$

假定

$$x(n) = \begin{cases} 0, & n < 0 \\ 0.141, 0.181, 0.28125, 0.5, 0.3, \dots, & n = 0, 1, \dots \end{cases}$$

设处理器字长为 $b+1=6$ 位。试用图 7-26 所示系统,说明如何用硬件计算滤波器的输出值 $y(3)$ 。

解 将 $x(n)$ 的相应值代入(7-160)式,可得

$$y(3) = 0.75 \times 0.5 - 0.5 \times 0.28125 = 0.234375$$

用二进制数表示,则有

$$y(3) = 0.001111$$

由于硬件字长是 $b+1=6$ 位,所以抽样数据和滤波器系数可表示为(二进制数用补码表示, $d_1$  表示系数,可参见(2-285)式)

$$x(3) = 0.5 = 0.10000$$

$$x(2) = 0.28125 = 0.01001$$

$$d_1 = 0.75 = 0.11000$$

$$d_2 = -0.5 = 1.10000$$

$$d_1 + d_2 = 0.75 - 0.5 = 0.25 = 0.01000$$

根据表 6-2,这个例子对应于图 7-26 的 ROM 图应为

ROM 地址		ROM 内容
$A_1$	$A_2$	
0	0	0.00000
1	0	0.11000
0	1	1.10000
1	1	0.01000

第一次迭代过程如下:

第(1)步 寄存器清零,即

$$RA = RB = RC = 0.00000$$

第(2)步  $x(3)$ 送到 SR1 上,前一次迭代中的 $x(2)$ 值则寄存到 SR2 中,即

$$SR1 = 0.10000 \quad SR2 = 0.01001$$

第(3)步 ROM 的地址位 $A_1$ 等于 SR1 的最低有效位,ROM 的地址位 $A_2$ 等于 SR2 的最低有效位,也就是 $A_1=0, A_2=1$ 。从 ROM 中取出由地址 $A_1, A_2$ 所决定的编址存储值,并将它送到寄存器 RA 中,即 $RA=1.10000$ ,在这一步中,只有 RA 寄存器改变内容;

第(4)步 在 ALU 中,将 RA 与 RC 的内容相加,并将结果存储到 RB 中,即

$$RA = 1.10000 \text{ (见第(3)步)}$$

$$RC = 0.00000 \text{ (见第(1)步)}$$

$$RB = RA + RC = 1.10000$$

在这一步中,只有 RB 改变内容。

第(5)步 采用硬件斜移并行联接的办法(右移一位,见图 7-26),将 RB 的内容复制到 RC 中,在这一步中,最重要是将 RB 的符号位加以复制,以产生 RC 的符号位以及最高位,如图 7-26 所示,这是补码运算的要求。例如,补码负数  $-0.25 = 1.11000$ ,被 2 除可得  $-0.25/2 = -0.125 = 1.1110$ ,上述联接方法(即图 7-26 的联接方法),就能满足这一要求;

第(6)步 将移位寄存器 SR1 和 SR2 的内容右移一位,即

$$SR1 = 0.01000 \quad SR2 = 0.00100$$

应注意,SR1 的最低有效位右移一位后,就成为 SR2 的符号位。

至此,完成了第一次迭代。同样,余下的 5 次迭代可循环通过第(3)步到第(6)步来实现。表 7-7 概括了在每次迭代的起始处(第(3)步)及末尾处(第(6)步)各寄存器的内容。

表 7-7 [例 7-3]中寄存器内容的概括

迭代号	执行步骤	移位寄存器 1 的内容	移位寄存器 2 的内容	ROM 地址		寄存器 A 内容	寄存器 B 内容	寄存器 C 内容
				A <sub>1</sub>	A <sub>2</sub>			
起始	(1)	0.00000	0.00000	0	0	0.00000	0.00000	0.00000
	(2)	0.10000 <i>x(3)</i>	0.01001 <i>x(2)</i>	0	1	0.00000	0.00000	0.00000
1	(3)	0.10000	0.01001	0	1	1.10000	0.00000	0.00000
	(6)	0.01000	0.00100	0	0	1.10000	1.10000	1.11000
2	(3)	0.01000	0.00100	0	0	0.00000	1.10000	1.11000
	(6)	0.00100	0.00010	0	0	0.00000	1.11000	1.11100
3	(3)	0.00100	0.00010	0	0	0.00000	1.11000	1.11100
	(6)	0.00010	0.00001	0	1	0.00000	1.11100	1.11110
4	(3)	0.00010	0.00001	0	1	1.10000	1.11100	1.11110
	(6)	0.00001	0.00000	1	0	1.10000	1.01110	1.10111
5	(3)	0.00001	0.00000	1	0	0.11000	1.01110	1.10111
	(6)	0.00000	1.00000	0	0	0.11000	0.01111	0.00111
6	(3)	0.00000	1.00000	0	0	0.00000	0.01111	0.00111
	(6)	0.00000	0.10000	0	0	0.00000	0.00111	0.00011

$$\downarrow \\ y(3) = 0.21875$$

在  $b+1=6$  次迭代完后,  $y(3)$  是在寄存器 RB 中,即  $y(3) = (RB) = 0.00111 = 0.21875$ ,这与  $y(3)$  的精确值(0.234375)相比较可看出,由于滤波器实现中的有限字长,使得计算出的  $y(3)$  值有误差。

#### 数字滤波器硬件实现中的溢出误差及其消除

除了上面观察到的有限字长造成的误差外,还可能存在因运算溢出所造成的溢出误差。例如对(7-160)式的线性差分方程,如果输入信号为

$$x(n) = \begin{cases} 0, & n \leq -1 \\ 0.75, 0.5, 0.25, 0, \dots, & n = 0, 1, 2, \dots \end{cases}$$

则可算出

$$y(0) = 0.75 \times 0.75 = 0.5625$$

硬件算出  $y(0)$  的迭代进程归纳在表 7-8 中。从此表可看出，在第 5 次迭代的第(6)步中，出现运算溢出。实际上是在这次迭代的第(4)步，求  $RB = RC + RA$  所造成的，由于运算溢出，最后的答案  $y(0) = -0.875$ ，这显然是错误的结果。

表 7-8 说明溢出出现的寄存器内容的归纳

迭代号	执行步骤	移位寄存器1的内容	移位寄存器2的内容	ROM 地址		寄存器 A 内容	寄存器 B 内容	寄存器 C 内容
				A <sub>1</sub>	A <sub>2</sub>			
起始	(1)	0.00000	0.00000	0	0	0.00000	0.00000	0.00000
	(2)	0.11000 $x(0)$	0.00000 $x(-1)$	0	0	0.00000	0.00000	0.00000
1	(3)	0.11000	0.00000	0	0	0.00000	0.00000	0.00000
	(6)	0.01100	0.00000	0	0	0.00000	0.00000	0.00000
2	(3)	0.01100	0.00000	0	0	0.00000	0.00000	0.00000
	(6)	0.00110	0.00000	0	0	0.00000	0.00000	0.00000
3	(3)	0.00110	0.00000	0	0	0.00000	0.00000	0.00000
	(6)	0.00011	0.00000	1	0	0.00000	0.00000	0.00000
4	(3)	0.00011	0.00000	1	0	0.11000	0.00000	0.00000
	(6)	0.00001	1.00000	1	0	0.11000	0.11000	0.01100
5	(3)	0.00001	1.00000	1	0	0.11000	0.11000	0.01100
	(6)	0.00000	1.10000	0	0	0.11000	1.00100*	1.10010
6	(3)	0.00000	1.10000	0	0	0.00000	1.00100	1.10010
	(6)					0.00000	1.10010	1.11001

\* 溢出

$$\downarrow \\ y(0) = -0.875$$

为了减小由运算溢出所造成的误差，可将图 7-26 的系统加以改进，即，当溢出出现时，使系统有饱和输出的功能。下面举两个例子来加以说明。

(1) 第一个例子。令

$$RC = -0.75 = 1.010$$

$$RA = -0.5 = 1.100$$

求二者之和，得到

$$RB = RA + RC = 10.110$$

显然这时已出现了溢出，因而，为了减小计算误差，必须改变寄存器 RB 的内容，使之饱和成  $-1(1.000)$ ，这可用改变 RB 的符号位  $c_0$ ，并将寄存器 RB 中所有其他位全部清成逻辑零来做到。

(2) 第二个例子。令

$$RC = 0.75 = 0.110$$

$$RA = 0.5 = 0.100$$

二者之和为

$$RB = RA + RC = 1.010$$

显然这个例子也产生了溢出误差。为了减小计算误差，须寄存器 RB 的内容改变成  $1 - 2^{-6}$ ，此例中就是要将 RB 置为 0.111。这可用改变 RB 的符号位  $c_0$ ，并将寄存器 RB 中

所有其他位全部置为逻辑1来做到。

观察以上这两个例子,我们可以导出一个逻辑表达式,它描述了溢出或误差出现时的结果,如果用逻辑 1 表示溢出或误差的出现,用 O/E 来表示它,则可写成

$$O/E = [c_{o(RA)} + c_{o(RC)} + c_{o(RB)}] + [c_{o(8A)} + c_{o(8C)} + c_{o(8B)}] \quad (7-161)$$

表 7-9 是与(7-161)式相对应的真值表。

按照以上讨论,我们可将图 7-26 修改成图 7-27,它可检测溢出或误差,并且能修正寄存器 RB 的内容,即,当检测出一个误差或溢出时,逻辑上使寄存器 RB 饱和成表 7-9 中合适的对应值。

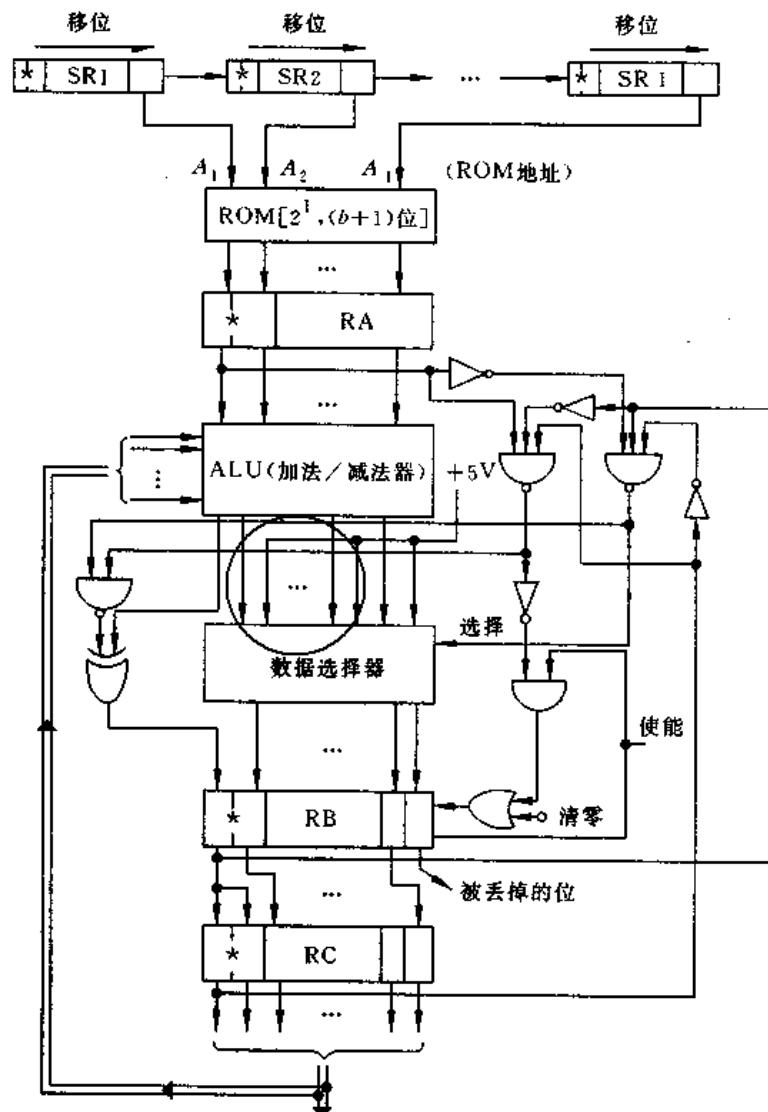


图 7-27 在 ALU 中引入饱和输出的硬件修正方案的方框图

为了看出在滤波器实现中,使用饱和加法器的效果,我们把它归纳在表 7-10 中,当出现溢出时(第 5 次迭代的第(4)步),把寄存器 RB 的内容加以修改,使得它饱和到合适的值。这个作用在表上已概括出来,可以看出,最后结果等于 0.46875,显然它比表 7-8 中

溢出后的结果 -0.875 更加接近理想值 0.5625。

表 7-9 与公式(7-161)相对应的真值表

$c_0(RA)$	$c_0(RC)$	$c_0(RB)$	O/E
0	0	0	0
0	0	1	1(饱和到 0.111...)
0	1	0	0
0	1	1	0
1	0	0	0
1	0	1	0
1	1	0	1(饱和到 1.000...)
1	1	1	0

### 减少 ROM 数量的硬件实现法

从(6-284)式  $A = \sum_{k=1}^I d_k B_k$  看出, 阵列乘法原来需要  $I$  次乘法、( $I-1$ ) 次加法, 而采用(6-286)式计算时(见表 7-7)只需  $b$  次加法, 但是其代价是需要大量的 ROM(ROM 必须有  $2^I$  个可寻址存储单元), 使得价格上升, 且存取时间要增加, 当  $I$  太大时, 就难以实现。为此, 我们可以用增加加法器的办法来减少对 ROM 的需求量。例如, 令  $I = UV$ , 则(6-284)式可重写为

$$A = \sum_{i=1}^V w_i \quad (7-162)$$

其中  $w_i$  是第  $i$  个部分和, 即

$$w_i = \sum_{k=(i-1)U+1}^U d_k B_k, \quad i = 1, 2, \dots, V \quad (7-163)$$

(7-163)式中  $w_i$  的表达式和(6-284)式的形式是一样的, 只不过相加项数是  $U$ , 可以用类似于图 7-26 的硬件结构来实现此  $w_i$ , 然后求  $V$  个  $w_i$  的总和, 即可求得  $A$ 。下面的例子表明, 在滤波器实现中, 采用增加加法次数的办法, 可大大减少对 ROM 存储量的要求。

[例 7-4] 设  $I=18$ , 试计算

(1) 用图 7-24 的方法实现时, 对 ROM 的存储要求;

表 7-10 具有饱和输出效果的寄存器内容的归纳

迭代号	执行步骤	移位寄存器1的内容	移位寄存器2的内容	ROM 地址		寄存器 A 内容	寄存器 B 内容	寄存器 C 内容
				$A_1$	$A_2$			
5	(3) 见表 7-8	0.00001	1.00000	1	0	0.11000	0.11000	0.01100
	(4)	0.00001	1.00000	1	0	0.11000	1.00100 <sup>+</sup>	0.01100
	(6)	0.00000	1.10000	0	0	0.11000	1.00100 <sup>+</sup> 0.11111	0.01111
6	(3)	0.00000	1.10000	0	0	0.00000	0.11111	0.01111
	(6)	0.00000	0.11000			0.00000	0.01111	0.00111

<sup>+</sup>溢出

$$A = y(0) = 0.46875$$

(2) 用(7-162)式及(7-163)式规定的办法计算  $A$  值时, 对 ROM 存储器的要求以及所需的额外加法次数。讨论所得到的结果。

解 (1) 图 7-26 的实现方法对 ROM 存储量的要求是

$$2^I = 2^{18} = 262114 \text{ 个存储单元(字长 } b+1 \text{ 位)}$$

加法次数是  $b$  次。

(2) 用(7-162)式及(7-163)式实现时, 可令  $I=UV=18$ , 即  $V=3, U=6$ , 此时, 对 ROM 的存储量的要求是

$$V2^U = 3 \times 2^6 = 192 \text{ 个存储单元(字长 } b+1 \text{ 位)}$$

加法次数是  $b+V-1=b+2$  次, 即只比图 7-26 的实现方法多  $V-1=2$  次加法。因而采用这种分解法时,  $I=UV$ , 可大大减少对 ROM 存储量的要求, 但其代价是

- ① 视其  $V$  的大小, 要多执行  $(V-1)$  次加法, 速度会稍有降低;
- ② 必须增加硬件, 以实现各个部分和  $w_i$ 。

采用这种分解法实现的简图如图 7-28 所示, 它是假定  $V=2$  的简单情况。

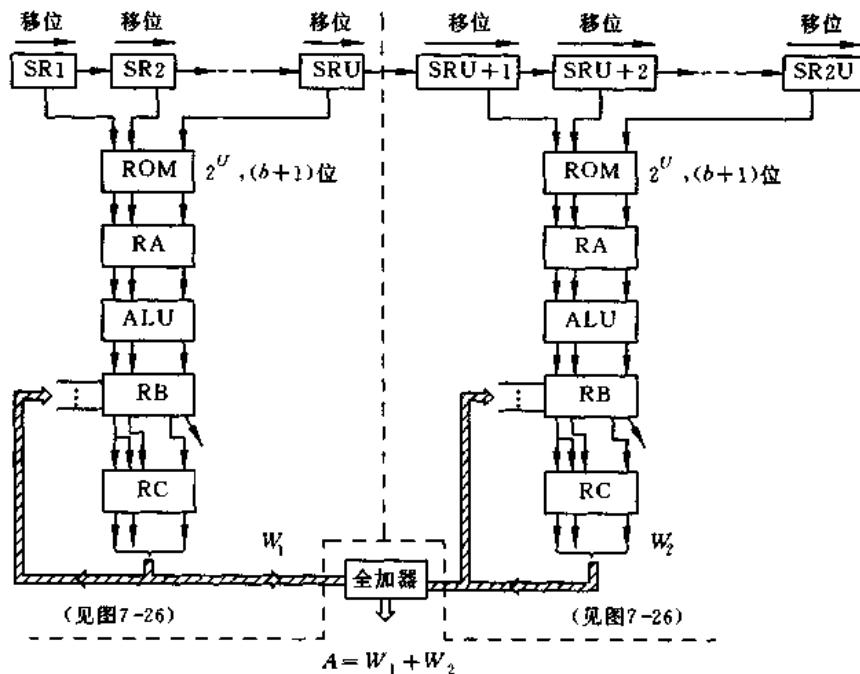


图 7-28 减少 ROM 存储量的硬件实现方框图

## 7.7 IIR 与 FIR 数字滤波器的比较

1. 在相同的技术指标下, IIR 滤波器由于存在着输出对输入的反馈, 所以可用比 FIR 滤波器较少的阶数来满足指标的要求, 所用的存储单元少, 运算次数少, 较为经济。例如, 用频率抽样法设计阻带衰减为  $-20\text{dB}$  的 FIR 滤波器, 其阶数要 33 阶才能达到, 而用双线性变换法设计只需 4~5 阶的切贝雪夫 IIR 滤波器即可达到指标要求, 所以 FIR 滤波器的

阶数要高5~10倍左右。

2. FIR滤波器可得到严格的线性相位，而IIR滤波器做不到这一点，IIR滤波器的选择性愈好，其相位的非线性愈严重。因而，如果IIR滤波器要得到线性相位，又要满足幅度滤波的技术要求，必须加全通网络进行相位校正，这同样会大大增加滤波器的阶数。从这一点上看，FIR滤波器又优于IIR滤波器。

3. FIR滤波器主要采用非递归结构，因而无论是从理论上还是从实际的有限精度的运算中它都是稳定的，有限精度运算的误差也较小。IIR滤波器必须采用递归结构，极点必须在z平面单位圆内才能稳定，对于这种结构，运算中的四舍五入处理有时会引起寄生振荡。

4. 对于FIR滤波器，由于冲激响应是有限长的，因而可以用快速傅里叶变换算法，这样运算速度可以快得多。IIR滤波器则不能这样运算。

5. 从设计上看，IIR滤波器可以利用模拟滤波器设计的现成的闭合公式、数据和表格，因而计算工作量较小，对计算工具要求不高。FIR滤波器则一般没有现成的设计公式，窗函数法只给出窗函数的计算公式，但计算通带、阻带衰减仍无显示表达式。一般FIR滤波器设计仅有计算机程序可资利用，因而要借助于计算机。

6. IIR滤波器主要是设计规格化的、频率特性为分段常数的标准低通、高通、带通、带阻、全通滤波器。FIR滤波器则要灵活得多，例如频率抽样设计法，可适应各种幅度特性及相位特性的要求，因而FIR滤波器可设计出理想正交变换器、理想微分器、线性调频器等各种网络，适应性较广。而且，目前已有许多FIR滤波器的计算机程序可供使用。

从以上比较看出，IIR滤波器与FIR滤波器各有特点，所以可以由实际应用时的要求，从多方面考虑来加以选择。

## 习题

1. 用矩形窗设计一个FIR线性相位低通数字滤波器。已知： $\omega_c = 0.5\pi$ ,  $N = 21$ 。求出  $h(n)$  并画出  $20 \lg |H(e^{j\omega})|$  曲线。

2. 用三角形窗设计一个FIR线性相位低通数字滤波器。已知： $\omega_c = 0.5\pi$ ,  $N = 51$ 。求出  $h(n)$  并画出  $20 \lg |H(e^{j\omega})|$  的曲线。

3. 用汉宁窗设计一个线性相位高通滤波器

$$H_d(e^{j\omega}) = \begin{cases} e^{-j(\omega-\pi)\alpha}, & \pi - \omega_c \leq \omega \leq \pi \\ 0, & 0 \leq \omega < \pi - \omega_c \end{cases}$$

求出  $h(n)$  的表达式，确定  $\alpha$  与  $N$  的关系，并画出  $20 \lg |H(e^{j\omega})|$  曲线（设  $\omega_c = 0.5\pi$ ,  $N = 51$ ）。

4. 用海明窗设计一个线性相位带通滤波器

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega\alpha}, & \omega_0 - \omega_c \leq \omega \leq \omega_0 + \omega_c \\ 0, & 0 \leq \omega < \omega_0 - \omega_c, \quad \omega_0 + \omega_c < \omega \leq \pi \end{cases}$$

求出  $h(n)$  的表达式并画出  $20 \lg |H(e^{j\omega})|$  曲线（设  $\omega_c = 0.2\pi$ ,  $\omega_0 = 0.5\pi$ ,  $N = 51$ ）。

5. 用布拉克曼窗设计一个理想线性相位  $90^\circ$  移相带通滤波器

$$H_d(e^{j\omega}) = \begin{cases} j e^{-j\omega\alpha}, & \omega_0 - \omega_c \leq \omega \leq \omega_0 + \omega_c \\ 0, & 0 \leq \omega < \omega_0 - \omega_c, \quad \omega_0 + \omega_c < \omega \leq \pi \end{cases}$$

求出  $h(n)$  序列，并画出  $20 \lg |H(e^{j\omega})|$  曲线（设  $\omega_c = 0.2\pi$ ,  $\omega_0 = 0.6\pi$ ,  $N = 51$ ）。

6. 用凯泽窗设计一个线性相位理想低通滤波器,若输入参数为低通截止频率  $\omega_c$ , 冲激响应长度点数  $N$  以及凯泽窗系数  $\beta$ , 求出  $h(n)$ , 并画出  $20 \lg |H(e^{j\omega})|$  曲线。

7. 试用第一种和第二种频率抽样法设计一个 FIR 线性相位数字低通滤波器(已知  $\omega_c = 0.5\pi$ ,  $N=51$ )。画出两种情况下的  $20 \lg |H(e^{j\omega})|$  曲线。

8. 如果一个线性相位带通滤波器的频率响应为

$$H_{BP}(e^{j\omega}) = H_{BP}(\omega) e^{j\theta(\omega)}$$

(1) 试证明一个线性相位带阻滤波器可以表示成

$$H_{BR}(e^{j\omega}) = [1 - H_{BP}(\omega)] \cdot e^{j\theta(\omega)}, \quad 0 \leq \omega \leq \pi$$

(2) 试用带通滤波器的单位冲激响应  $h_{BP}(n)$  来表达带阻滤波器的单位冲激响应  $h_{BR}(n)$ 。

9. 已知图 P7-9(a) 中的  $h_1(n)$  是偶对称序列  $N=8$ , 图 P7-9(b) 中的  $h_2(n)$  是  $h_1(n)$  圆周移位( $移 \frac{N}{2}=4$  位)后的序列。设

$$H_1(k) = DFT[h_1(n)], \quad H_2(k) = DFT[h_2(n)]$$

(1) 问  $|H_1(k)| = |H_2(k)|$  成立否?  $\theta_1(k)$  与  $\theta_2(k)$  有什么关系?

(2)  $h_1(n), h_2(n)$  各构成一个低通滤波器, 试问它们是否是线性相位的? 延时是多少?

(3) 这两个滤波器性能是否相同? 为什么? 若不同, 谁优谁劣?

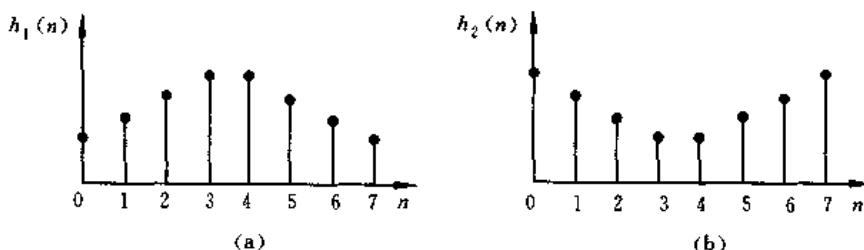


图 P7-9

10. 请选择合适的窗函数及  $N$  来设计一个线性相位低通滤波器

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega}, & 0 \leq \omega \leq \omega_c \\ 0, & \omega_c \leq \omega \leq \pi \end{cases}$$

要求其最小阻带衰减为  $-45\text{dB}$ , 过渡带宽为  $\frac{8}{51}\pi$ 。

(1) 求出  $h(n)$  并画出  $20 \lg |H(e^{j\omega})|$  曲线(设  $\omega_c = 0.5\pi$ )。

(2) 保留原有轨迹, 画出用满足所给条件的其他几种窗函数设计出的  $20 \lg |H(e^{j\omega})|$  曲线。

11. 已知一限带理想微分器, 它的延时为  $\tau$ , 对应的频率特性为

$$H_d(j\Omega) = \begin{cases} j\Omega e^{-j\Omega\tau}, & |\Omega| < \Omega_c \\ 0, & \text{其他 } \Omega \end{cases}$$

(1) 利用冲激响应不变法, 求有延时的理想数字微分器的单位抽样响应

$$h_d(n) = Th_d(nT) (\text{设 } \Omega_c = \pi/T)。$$

(2) 由  $h_d(n)$  求相应的频率特性  $H_d(e^{j\omega})$ , 并画出其示意图; 这个系统的延时是多少个抽样间隔。

(3) 如果用点数为  $N$  的矩形窗对  $h_d(n)$  截断, 得到因果性 FIR 逼近的数字微分器  $h(n)$ , 则在

(a)  $N$  为偶数

(b)  $N$  为奇数

这两种情况下应如何选择  $\tau$ , 画出每种情况下  $h(n)$  的示意图。

12. 若最小相位的 FIR 系统的单位抽样响应为  $h_{min}(n), n=0, 1, \dots, N-1$ , 另一个 FIR 系统的单位

抽样响应  $h(n)$  为

$$h(n) = h_{\min}(N-1-n), \quad n=0, 1, \dots, N-1$$

试证明：

(1) 系统  $h(n)$  和系统  $h_{\min}(n)$  有着相同的幅频响应。

(2) 系统  $h(n)$  是最大相位延时系统。

13. 若一个离散时间系统的系统函数是

$$H(z) = (1 - 0.85e^{j0.2\pi}z^{-1})(1 - 0.85e^{-j0.2\pi}z^{-1})(1 - 1.5e^{j0.3\pi}z^{-1})(1 - 1.5e^{-j0.3\pi}z^{-1}) \\ \times (1 - 0.7z^{-1})(1 - 1.2z^{-1})$$

移动其零点可得到新的系统，但要在下述条件下：

(1) 新系统和  $H(z)$  具有相同的幅频响应。

(2) 新系统的单位抽样响应仍为实数，且其长度和原系统的一样。

试问：

(1) 可得几个不同的系统？

(2) 哪一个是最小相位的？哪一个是最小相位的？

(3) 对所有系统求  $h(n)$ ，并计算  $E(m) = \sum_{n=0}^m h^2(n)$ ,  $m \leq 6$ ，试比较各系统的能量累积情况。

14. 要设计一个理想线性相位带通滤波器

$$H_d(e^{j\omega}) = \begin{cases} e^{-j\omega}, & \omega_0 - \omega_c \leq \omega \leq \omega_0 + \omega_c \\ 0, & 0 \leq \omega \leq \omega_0 - \omega_c, \quad \omega_0 + \omega_c < \omega \leq \pi \end{cases}$$

若需阻带衰减大于(1) 50dB；(2) 60dB。

试用窗函数法设计这两个滤波器(取  $\omega_0 = 0.5\pi$ ,  $\omega_c = 0.1\pi$ )。

15. 如图 P7-15 有 8 个有限长序列  $N=4$ ，各序列的傅里叶变换都相同，试问哪些序列的  $z$  变换部零点都位于单位圆内部。

16. 已知  $h_1(n), h_2(n), \dots, h_M(n)$  为  $M$  个点数皆为  $N$  点 ( $0 \leq n \leq N-1$ ) 的有限长序列 ( $h_i(n) = 0, n > N, n < 0, i=1, \dots, M$ )，各序列的傅里叶变换的幅度函数都相同，且各序列对实常数或复常数而言不成正比。

试讨论：(1) 若各序列可以是实序列，也可以是复序列，则  $M$  的最大值是多少？

(2) 若各序列必须是实序列，则  $M$  的最大值是多少？

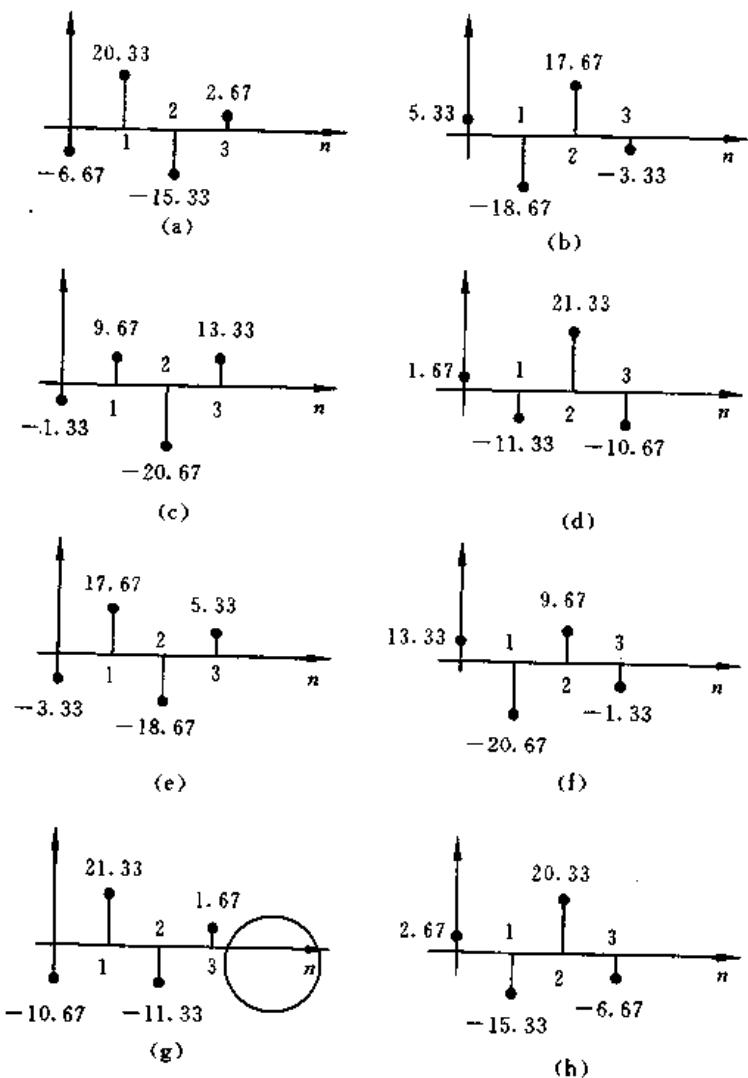


图 P7-15

## 第八章 数字信号处理中的有限字长效应

### 8.1 引言

上面各章所讨论的数字信号与系统都是无限精度的。但是，实际上无论是用专用硬件还是用计算机软件来实现，其数字信号处理系统的有关参数以及运算过程中的结果都是存储在有限字长的存储单元中的。如果处理的是模拟信号，即抽样信号处理系统，输入的模拟量经过抽样及模数(A/D)变换后，也变成有限字长的数字信号。这样两种结果都是有限精度的，因而相对于原设计系统就产生了误差。在数字系统中有三种因有限字长的影响而引起误差的因素：

- (1) A/D 变换器将模拟输入信号变为一组离散电平时产生的量化效应；
- (2) 把系数用有限位二进制数表示时产生的量化效应；
- (3) 在数字运算过程中，为限制位数而进行尾数处理以及为防止溢出而压缩信号电平的有限字长效应，包括低电平极限环振荡效应以及溢出振荡效应。

上述三种误差与系统结构形式、数的表示方法、所采用的运算方式、字的长短以及尾数的处理方式有关。但是，将上面三种误差因素综合起来分析是困难的，只能分别对三种效应单独加以分析，以计算出它们的影响。

研究有限字长效应的目的是：

- (1) 如果数字信号处理是在通用计算机上实现时，字长已经固定，做误差分析是为了知道结果的可信度，否则要采取改进措施。但是一般计算机字长较长，可不考虑字长的影响。
- (2) 用专用硬件实现数字信号处理时，一般是采用定点实现，涉及到硬件采用的字长问题，因而必须了解为达到所需精度所必须选用的最小字长，以便在设备价格和达到精度之间作合适的折衷。

下面我们首先从二进制数表示法开始讨论。

### 8.2 二进制数的表示及其对量化的影响

在数字系统中所采用的二进制算法最基本的有定点制、浮点制及成组浮点制。

#### 一、三种算术运算法

##### 1. 定点二进制数

在整个运算中，二进制小数点在数码中的位置是固定不变的称为定点制。在定点制中，小数点右边各位表示数的分数部分，左边各位表示数的整数部分。例如，六位字长的二进制数 10.1001 等于十进制数

$$1 \times 2^1 + 0 \times 2^0 + 1 \times 2^{-1} + 0 \times 2^{-2} + 0 \times 2^{-3} + 1 \times 2^{-4} = 2.50625$$

其中小数点固定在二进制数码的第二位后是不变的。

原则上说,定点制的小数点可固定在任意位上,但是为了运算方便,通常定点制都把数  $M$  限制在  $\pm 1$  之间,即  $-1 < M < 1$ 。这时将小数点固定在第一位二进制码之前,而整数位则作为“符号位”,代表数的正负号(“0”表示正数,“1”表示负数),数的本身只有小数部分,称为“尾数”。例如,0.375 表示成二进制数为 0.011。定点制在整个运算过程中,所有运算结果的绝对值都不能超过 1。为此,当数很大时,就乘一个比例因子,使整个运算中数的最大绝对值不超过 1;运算完以后,再除以同一比例因子,还原成真值输出。如运算过程中出现绝对值超过 1 时,数就进位到整数部分的符号位,这就出现错误,称为溢出,这时应该修正比例因子。但是,在 IIR 滤波器中,分母的系数决定着极点的位置,所以不适合用比例因子。

定点制的加法运算不会增加字长,但若没选择合适的比例因子,则加法运算会出现溢出的可能性,例如  $0.1001 + 0.1101 = ①.0110$ ,产生溢出。

定点制的乘法运算不会产生溢出,因为绝对值小于 1 的两个数相乘后,其绝对值仍小于 1,但是相乘后字长却要增加一倍,一般说  $(b+1)$  位的定点数(其中  $b$  位为字长,1 位为符号位)相乘后字长为  $2b$  位,因此在定点制每次相乘运算后需要进行尾数处理,使结果仍然保持  $b$  位字长。对超过字长的尾数有两种处理办法:一是简单地去掉超过字长  $b$  的各尾数位,称为“截尾”,另一是在舍去超过字长的各尾数位时,若舍掉部分的值大于或等于保留部分最低位的权值的一半,则给留下部分的最低位处加 1,结果字长仍为  $b$  位。这相当于十进制中的四舍五入近似法,故称为“舍入”。尾数处理结果会带来截尾误差或舍入误差。

## 2. 浮点二进制数

定点制的缺点是动态范围小,有溢出现象。浮点制则可避免这一缺点,它的动态范围大,可以避免溢出,不需要比例因子。浮点制是将一个数表示成尾数和指数两部分,即

$$x = \pm 2^c M$$

$M$  是数  $x$  的尾数部分,  $2^c$  是数  $x$  的指数部分,  $c$  是阶数,称为阶码。例如  $x = 0.11 \times 2^{010}$  表示的十进制数是

$$x = 0.75 \times 2^2 = 3$$

浮点制中把尾数用带符号位的定点数来表示,因而尾数的第一位就表示浮点数的符号,一般为了充分利用尾数的有效位数,总是使尾数字长的最高位(符号位除外)为 1,称为规格化形式,这时尾数  $M$  是小数,且满足  $1/2 \leq M < 1$ ,例如  $x = 0.0101 \times 2^{011}$  就是非规格化形式,应调整尾数的小数点并相应地调整阶码,使它成为  $x = 0.101 \times 2^{010}$ ,这就是规格化的浮点数。

阶码  $c$  也是带符号位的定点数,这是因为要用负的阶码表示数值小于 0.5 的数。

浮点表示数的小数点是浮动的,一般通用计算机中往往同时使用定点、浮点两种运算方式。

在浮点制中,位数必须分成两部分。尾数为  $b_m + 1$  位,其中的 1 是符号位;阶码为  $b_c + 1$  位,其中的 1 位也是符号位。浮点数的尾数字长决定了浮点制的运算精度,而阶码

字长决定了浮点制的动态范围。例如  $b_m + b_c + 2 = 16$  位浮点数,  $b_m = 8$ ,  $b_c = 6$ (采用规格化尾数), 它所表示的动态范围(若阶码和尾数都用补码表示)为(如图 8-1 所示)

$$2^{-2^{b_c}} \cdot 2^{-1} \leq |x|_{10} \leq 2^{(2^{b_c}-1)} \cdot (1 - 2^{-b_m}) < 2^{(2^{b_c}-1)}$$

其中  $|x|_{10}$  表示十进制数的绝对值。代入  $b_c, b_m$ , 值, 则为

$$2^{-65} \leq |N_{10}| \leq 0.99609375 \cdot 2^{63} < 2^{63}$$

尾数精度为

$$2^{-8} = \frac{1}{256} = 3.9\%$$

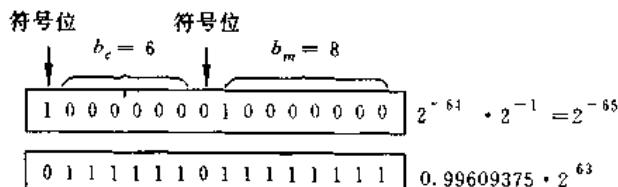


图 8-1 16 位浮点数的动态范围

浮点制的乘法是尾数相乘阶码相加。尾数相乘的过程与定点制相同, 因此也要作截尾或舍入处理, 由于尾数的乘积是  $\frac{1}{4}$  到 1 之间的数, 故需要加以规格化, 并同时调整阶码。

浮点制的加法, 如果两个数的阶码相同, 则只要两个尾数相加就得到和数的尾数, 和数的阶码则为两数原来的阶码。阶码不等的两浮点数相加则要分三步进行: 第一, 先进行对阶, 这时需将小阶码向大阶码看齐, 同时, 阶小的数的尾数的小数点要左移, 左移一位, 阶码加 1, 直至两数阶码相等; 第二, 将所得两尾数用定点运算相加; 第三, 使所得结果的尾数归一化, 并相应调整阶码。例如两个二进制浮点数

$$x_1 = 2^{010} \times 0.1100, \quad x_2 = 2^{000} \times 0.1001$$

求它们的浮点相加结果。首先对阶, 将阶码较小的  $x_2$  的阶码变成与  $x_1$  的阶码一样, 这只需将  $x_2$  的尾数小数点左移两位而阶码加 2, 即得

$$x_2 = 2^{010} \times 0.001001$$

然后将两数相加, 得

$$x = x_1 + x_2 = 2^{010} \times 0.111001$$

最后进行规格化。在这个例子中  $x$  的尾数正好在  $\frac{1}{2}$  到 1 之间, 已是规格化的数, 否则还要规格化。如果尾数是  $b+1=5$  位, 则所得结果还要舍入(或截尾)处理成 5 位, 例如截尾处理后, 得到  $x = 2^{010} \times 0.1110$ 。

由以上讨论看出, 浮点制运算中, 不论是相乘还是相加, 尾数的位数都可能超过寄存器长度, 都要做尾数的量化处理, 因而都有量化误差。

### 3. 分组浮点二进制数

定点制优点是快速简单, 只有乘法才出现舍入或截尾误差, 其缺点是动态范围小, 可能出现溢出, 为防止溢出, 就要压缩输入信号电平, 这样就减小了输出信号与量化噪声的比值。而浮点制的优点是数的动态范围大, 缺点是运算速度慢, 并且其加法和乘法运算都

会产生舍入或截尾误差。

分组浮点制则兼有定点制与浮点制的某些优点，是将这两种表示法结合起来。这种制式，一组数具有一个共同的阶码，这个阶码是这一组数中最大的那个数的阶码。由于一组数中各数的大小不同，因而只是这组中最大的数具有规格化的尾数，其他数则不可能刚好都是规格化的。由于分组浮点法只用一个单一的阶码表示一组数的阶码，因而节约了存储器，简化了系统。这种制式，对于要运算的数比较多，数值相近的情况特别适用。它最适宜实现快速傅里叶变换算法，当然也可用来实现数字滤波器。

## 二、负数的表示法——原码、补码、反码

不论是定点制还是浮点制的尾数都是将整数位用作符号位，小数位代表尾数值。 $(b+1)$ 位码的形式为

$$a_0 \cdot a_1 a_2 \cdots a_b \quad (8-1)$$

其中整数位  $a_0$  表示符号位，小数位  $a_1, a_2, \dots, a_b$  表示  $b$  位字长的尾数值， $a_i$  表示第  $i$  位二进制码 ( $i=0, 1, \dots, b$ )，取值可为 0 或 1。

对于正数，上面的表示是很清楚的。对于负数，根据需要的不同，有原码、反码、补码三种。

### 1. 原码

原码也称“符号-幅度码”，它的尾数部分代表数的绝对值（即幅度大小），符号位代表数的正负号， $a_0=0$  时代表正数， $a_0=1$  时，代表负数，例如  $x=0.011$  表示的是  $+0.375$ ， $x=1.011$  表示的是  $-0.375$ 。如果用(8-1)式表示原码，则可定义为

$$[x]_{\text{原}} = \begin{cases} |x|, & 0 \leq x < 1 \\ 1 + |x| (\text{或 } 1 - x), & -1 < x \leq 0 \end{cases} \quad (8-2)$$

它所代表的十进制数值为

$$x = (-1)^{a_0} \sum_{i=1}^b a_i 2^{-i} \quad (8-3)$$

原码的优点是乘除运算方便，以两数符号位的逻辑加就可简单决定结果的正负号，而数值则是两数数值部分的乘除结果。但原码的加减运算则不方便，因为两数相加，先要判断两数符号是否相同，相同则做加法，不同则做减法，做减法时还要判断两数绝对值大小，以便用大者作为被减数，这样增加了运算时间。用以下的补码则可使加法和减法运算方便。

在原码表示中，“零”有两种表示方法，例如  $b=3$  时， $0.000$  及  $1.000$  都表示零，故  $b+1$  位字长，只能表示  $(2^{b+1}-1)$  个数，即表达  $-(1-2^{-b})$  到  $(1-2^{-b})$  间的数。

### 2. 补码

补码又称“2 的补码”。补码中正数与原码正数表示一样。补码中负数是采用 2 的补数来表示的，即把负数先加上 2，以便将正数与负数的相加转化为正数与正数相加，从而克服原码表示法做加减法的困难。因此，给补码定义如下：

$$[x]_{\text{补}} = \begin{cases} |x|, & 0 \leq x < 1 \\ 2 - |x|, & -1 < x \leq 0 \end{cases} \quad (8-4)$$

例如,  $x = -0.625$  在原码中表示为 1.101, 在补码中为  $2 - 0.625 = 1.375$ , 因此补码的表示为 1.011, 其整数 1 正好表示负数。对于(8-1)式的表示形式, 补码所代表的二进制数值可表示为

$$x = -a_0 + \sum_{i=1}^b a_i 2^{-i} \quad (8-5)$$

例如补码 1.011, 按照上式就知道其所表示的数为

$$x = -1 + 0.375 = -0.625$$

补码表示法可把减法与加法统一起来。都采用补码加法。例如做减法时, 若减数是正数, 则将其变为负数的补码与被减数的补码相加, 如减数是负数, 则将其变成正数的补码与被减数的补码相加。采用补码做加法, 符号位也同样参加运算, 如果符号位发生进位, 则把进位的 1 去掉就行了。

在补码表示中, “零”的表示是惟一的, 为 0.000 ( $b=3$ ), 故  $(b+1)$  位字长可表示  $2^{b+1}$  个不同的数, 即表示从  $-1$  到  $(1 - 2^{-b})$  之间的数。

由于负数的补码是  $2 - |x|$ , 故求负数的补码时, 实际上要做一次减法, 这是不希望的。可以发现, 只要将原码正数的每位取反码 ( $1 \rightarrow 0, 0 \rightarrow 1$ ), 再在所得数的末位加 1, 则正好得到负数的补码, 这简称为对尾数的“取反加 1”。

### 3. 反码

又称“1 的补码”。和补码一样, 反码的正数与原码的正数表示相同。反码的负数, 则是将该数的正数表示形式中的所有 0 改为 1, 所有 1 改为 0, 即“求反”。例如  $x = -0.375$ , 其正数表达式为 0.011, 将它 0,1 全部颠倒则得 1.100, 这就是  $x = -0.375$  的反码表示。

$x$  的反码就是  $x$  按位对 1 的补码, 即

$$[x]_{\text{反}} = \underbrace{1.11\dots11}_{(b+1) \text{ 位}} - |x| = (2 - 2^{-b}) - |x|$$

因而可给反码定义为

$$[x]_{\text{反}} = \begin{cases} |x|, & 0 \leq x < 1 \\ (2 - 2^{-b}) - |x|, & -1 < x \leq 0 \end{cases} \quad (8-6)$$

由此得出

$$\begin{aligned} [x]_{\text{反}} &= 2 - |x| - 2^{-b} = [x]_{\text{补}} - 2^{-b} \\ [x]_{\text{补}} &= [x]_{\text{反}} + 2^{-b} \end{aligned} \quad (8-7)$$

由此看出, 补码负数与反码负数之间的简单关系。

由(8-7)式, 利用求补码的十进制数值表达式(8-5), 求得反码的十进制数值为

$$x = -a_0(1 - 2^{-b}) + \sum_{i=1}^b a_i 2^{-i} \quad (8-8)$$

“零”在反码中有两种表示, 0.000 与 1.111, 因而  $(b+1)$  位字长可表示  $(2^{b+1} - 1)$  个不同的数, 即表示从  $-(1 - 2^{-b})$  到  $(1 - 2^{-b})$  之间的数。

习惯上, 加法器硬件多采用补码制, 而串行乘法器通常用原码表示。

表 8-1 以  $b=3$  为例, 列表表示了这三种码各自所表达的数值。

表 8-1 三种码的表示法 ( $b=3$ )

二进制数	原码值	补码值	反码值
0.111	7/8	7/8	7/8
0.110	6/8	6/8	6/8
0.101	5/8	5/8	5/8
0.100	4/8	4/8	4/8
0.011	3/8	3/8	3/8
0.010	2/8	2/8	2/8
0.001	1/8	1/8	1/8
0.000	0	0	0
1.000	-0	-1	-7/8
1.001	-1/8	-7/8	-6/8
1.010	-2/8	-6/8	-5/8
1.011	-3/8	-5/8	-4/8
1.100	-4/8	-4/8	-3/8
1.101	-5/8	-3/8	-2/8
1.110	-6/8	-2/8	-1/8
1.111	-7/8	-1/8	0

### 三、量化方式——舍入与截尾

上面已说到, 定点制的乘法以及浮点制的加法和乘法在运算结束后都会使字长增加, 因而都需要对尾数进行截尾或舍入处理, 由此引入的误差取决于所用二进制数的位数  $b$ 、数的运算方式(定点制或浮点制)、负数的表示法以及对尾数的处理方法(舍入或截尾)。下面分别加以分析。

#### 1. 定点制的截尾与舍入误差

设原来是  $b$  位字长, 运算后增加到  $b_1$  位字长, 需要对尾数进行量化处理, 使  $b_1$  位字长减小为  $b$  位字长。

(1) 定点制截尾: 截尾是保留  $b$  位, 把余下的尾数丢掉。

对于正数  $x$ , 三种码的表示法是相同的, 因而量化影响也是相同的。一个  $b_1$  位正数  $x$  的十进制数值为

$$x = \sum_{i=1}^{b_1} a_i 2^{-i}$$

截尾处理后为  $b$  位字长, 显然  $b < b_1$ , 我们用  $Q[\cdot]$  表示量化处理, 加下标 T 后,  $Q_T[\cdot]$  表示截尾量化处理, 则有

$$Q_T[x] = \sum_{i=1}^b a_i 2^{-i}$$

若以  $E_T$  表示截尾误差, 则有

$$E_T = Q_T[x] - x = - \sum_{i=b+1}^{b_1} a_i 2^{-i} \quad (8-9)$$

看出  $E_T$  为负值或零, 当被弃位  $a_i (i=b+1 \text{ 到 } i=b_1)$  均为 1 时, 为最大截尾误差

$$E_{T\max} = - \sum_{i=b+1}^{b_1} 2^{-i} = -(2^{-b} - 2^{-b_1}) \quad (8-10)$$

因而有

$$-(2^{-b} - 2^{-b_1}) \leq E_T \leq 0, \quad x > 0 \quad (8-11)$$

一般  $2^{-b} \gg 2^{-b_1}$ , 并令

$$\Delta = 2^{-b} \quad (8-12)$$

$\Delta$  表示最小码位所表示的数值, 称为“量化宽度”或“量化步阶”。因而定点正数的截尾误差是负数, 满足

$$-\Delta < E_T \leq 0, \quad x > 0 \quad (8-13)$$

对于负数  $x$ , 截尾误差与负数的表示法有关, 分别讨论如下:

① 原码负数 ( $a_0=1$ ), 则有

$$\begin{aligned} x &= - \sum_{i=1}^{b_1} a_i 2^{-i} \\ Q_T[x] &= - \sum_{i=1}^b a_i 2^{-i} \\ E_T &= Q_T[x] - x = \sum_{i=b+1}^{b_1} a_i 2^{-i} \end{aligned}$$

这时截尾误差是正数, 满足

$$0 \leq E_T \leq (2^{-b} - 2^{-b_1}), \quad x < 0 \quad (8-14)$$

因而定点制原码负数的截尾误差是正数, 满足

$$0 \leq E_T < \Delta, \quad x < 0 \quad (8-15)$$

例如:  $b_1=4, b=2$ , 原码负数  $x=1.1001$  表示  $-9/16$ ,  $Q_T[x]=1.10$  表示  $-1/2$ ,

$$E_T = Q_T[x] - x = -\frac{1}{2} - \left(-\frac{9}{16}\right) = \frac{1}{16} > 0$$

② 补码负数 ( $a_0=1$ ), 则有

$$\begin{aligned} x &= -1 + \sum_{i=1}^{b_1} a_i 2^{-i} \\ Q_T[x] &= -1 + \sum_{i=1}^b a_i 2^{-i} \\ E_T &= Q_T[x] - x = - \sum_{i=b+1}^{b_1} a_i 2^{-i} \end{aligned}$$

这个误差与正数时的误差是一样的, 也是负的, 即

$$-(2^{-b} - 2^{-b_1}) \leq E_T \leq 0 \quad (8-16)$$

因而定点制补码负数的截尾误差是负数, 满足

$$-\Delta < E_T \leq 0, \quad x < 0 \quad (8-17)$$

例如  $x=1.1010 (-0.375)$ ,  $Q_T[x]=1.10 (-0.5)$

$$E_T = Q_T[x] - x = -0.5 - (-0.375) = -0.125 < 0$$

③ 反码负数 ( $a_0=1$ ), 则有

$$x = -1 + \sum_{i=1}^{b_1} a_i 2^{-i} + 2^{-b_1}$$

$$Q_T[x] = -1 + \sum_{i=1}^b a_i 2^{-i} + 2^{-b}$$

$$E_T = Q_T[x] - x = -\sum_{i=b+1}^{b_1} a_i 2^{-i} + 2^{-b} - 2^{-b_1}$$

可以看出,当被弃位  $a_i$  ( $i=b+1$  到  $i=b_1$ ) 全为零时误差最大,全为 1 时误差最小,可得出

$$0 \leq E_T \leq (2^{-b} - 2^{-b_1}) \quad (8-18)$$

则定点制反码负数的截尾误差是正数,满足

$$0 \leq E_T < \Delta, \quad x < 0$$

它与原码负数时相同,  $E_T$  为正值或为零。

例如:  $x = 1.1100 (-0.1875), Q_T[x] = 1.11(0)$

$$E_T = Q_T[x] - x = 0.1875 > 0$$

总括起来, 补码的截尾误差皆为负数, 其截尾量化的非线性特性如图 8-2(a) 所示。原码与反码的截尾误差与数的正负有关, 正数时误差为负, 负数时误差为正, 其截尾量化的非线性特性如图 8-2(b) 所示。用式子表示为:

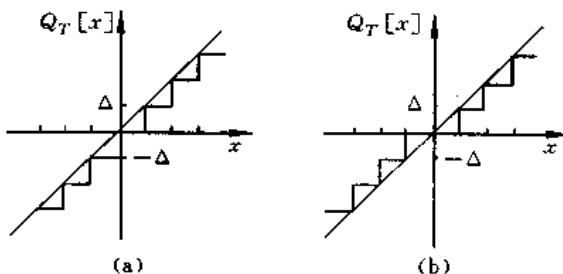


图 8-2 定点制截尾处理的量化特性 ( $\Delta = 2^{-b}$ )

(a) 补码; (b) 原码、反码

正数及补码负数  $-\Delta < E_T \leq 0$

原码负数及反码负数  $0 \leq E_T < \Delta$

(2) 定点制舍入: 舍入是按最接近的值取  $b$  位码, 因而舍入后各数值按  $\Delta = 2^{-b}$  的间距被量化, 即两个数间最小非零差是  $2^{-b}$ , 舍入是选择靠得最近的量化层标准值为舍入后的值, 因此不论是正数、负数, 也不论是原码、补码、反码, 其误差总是在  $\pm \frac{\Delta}{2} = \pm \frac{2^{-b}}{2}$  之间, 我们用  $Q_R[\cdot]$  表示舍入处理,  $E_R$  表示舍入误差, 则

$$E_R = Q_R[x] - x$$

$$-\frac{1}{2} \cdot 2^{-b} < E_R \leq \frac{1}{2} \cdot 2^{-b} \quad (8-20)$$

即定点制舍入误差为

$$-\frac{\Delta}{2} < E_R \leq \frac{\Delta}{2} \quad (8-21)$$

有时被舍入的数恰好在两个量化层标准值的正中间,这时可规定恒取上入,或恒取下舍,或是采用随机舍入。例如取  $b=2$ ,则

$$x=0.1001, Q_R[x]=0.10, \text{舍去 } 0.0001, \text{误差 } E_R=-2^{-4};$$

$$x=0.1011, Q_R[x]=0.11, \text{将 } 0.0011 \text{ 上入为 } 0.01, \text{误差 } E_R=+2^{-4};$$

$x=0.1010$ ,则  $x$  与 0.10 及 0.11 的距离相等,因此  $Q_R[x]$  既可下舍为 0.10,也可上入为 0.11,一般可按十进制中四舍五入的规则,因此取  $Q_R[x]=0.11$ ,则  $E_R=+2^{-3}$ 。

例如:对于补码数可表示为

$$x = -a_0 + \sum_{i=1}^{b_1} a_i 2^{-i}$$

舍入处理可表示为

$$Q_R[x] = -a_0 + \sum_{i=1}^b a_i 2^{-i} + a_{b+1} 2^{-b}$$

最后一项表示的是逢 5 进 1。其他码可类似地表示。

图 8-3 画出了定点舍入的量化特性。

比较图 8-2 和图 8-3 可看出,舍入误差是对称分布的,而补码截尾误差是单极性分布的,因而它们的统计特性是不同的。一般来说舍入误差的影响要小,所以应用得多一些。

## 2. 浮点制的截尾与舍入误差

浮点制中截尾或舍入处理只涉及尾数的字长,因而可运用定点误差分析的结果。但是所产生的误差的大小却与阶码有关,两个尾数相同而阶码不同的数,在同样尾数量化处理的情况下,误差的大小是不一样的。例如, $x_1$  与  $x_2$  是尾数相同而阶码不同的两个数,即

$$x_1 = 0.0101 \times 2^{000} (0.3125), \quad x_2 = 0.0101 \times 2^{011} (2.5)$$

则

$$Q_R[x_1] = 0.01 \times 2^{000} (0.25), \quad Q_R[x_2] = 0.01 \times 2^{011} (2.0)$$

因而

$$E_{R_1} = Q_R[x_1] - x_1 = -0.0625, \quad E_{R_2} = Q_R[x_2] - x_2 = -0.5$$

从这个例子可以看出,在同样的尾数舍去情况下,由于  $x_2$  是  $x_1$  的 8 倍,相应的量化误差  $|E_{R_2}|$  也是  $|E_{R_1}|$  的 8 倍。因而,在浮点制中,用相对误差比用绝对误差更能反映它的特点,我们以  $\epsilon$  来表示这个相对误差,即

$$\epsilon = \frac{Q[x] - x}{x} \quad (8-22)$$

根据此式,绝对误差可表示成

$$E = Q[x] - x = \epsilon x \quad (8-23)$$

这是相乘性误差,而不是像定点制那样是相加性误差。

下面分别就浮点舍入和浮点截尾分析  $\epsilon$  的误差范围。

(1) 浮点舍入:尾数误差在  $\pm \frac{\Delta}{2}$  之间,若  $x$  的阶码为  $c$ ,则

$$-2^c \frac{\Delta}{2} < Q_R[x] - x \leq 2^c \frac{\Delta}{2}$$

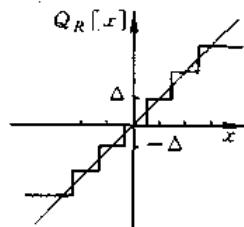


图 8-3 定点制舍入处理的量化特性

即

$$-2^e \frac{\Delta}{2} < \epsilon_R x \leq 2^e \frac{\Delta}{2} \quad (8-24)$$

由于  $x = \pm M 2^e$ , 其尾数部分是规格化的, 即  $\frac{1}{2} \leq M < 1$ , 故有

$$2^{e-1} \leq |x| < 2^e \quad (8-25)$$

分以下两种情况讨论:

① 由(8-25)式知, 当  $x$  为正时, 满足

$$2^{e-1} \leq x < 2^e \quad (8-26)$$

$x$  为正, 若  $\epsilon_R$  为正, 则有

$$\epsilon_R 2^{e-1} \leq \epsilon_R x$$

再利用(8-24)右边不等式与此式联立, 可得

$$\epsilon_R 2^{e-1} \leq 2^e \frac{\Delta}{2}$$

所以

$$\epsilon_R \leq \Delta \quad (8-27)$$

$x$  为正, 若  $\epsilon_R$  为负, 利用(8-26)式, 有

$$\epsilon_R 2^{e-1} \geq \epsilon_R x$$

再利用(8-24)左边的不等式与此式联立, 可得

$$\begin{array}{c} \epsilon_R 2^{e-1} > -2^e \frac{\Delta}{2} \\ \epsilon_R > -\Delta \end{array}$$

所以

$$(8-28)$$

这样当  $x$  为正时, 有

$$-\Delta < \epsilon_R \leq \Delta, \quad x > 0 \quad (8-29)$$

② 同样, 若  $x$  为负, 则有

$$-2^{e-1} \geq x > -2^e, \quad (8-30)$$

当  $x$  为负, 讨论  $\epsilon_R$  为正,  $\epsilon_R$  为负两种情况, 则可得

$$-\Delta \leq \epsilon_R < \Delta, \quad x < 0 \quad (8-31)$$

由(8-29)式与(8-31)式可知浮点制舍入的相对误差范围应为

$$-\Delta \leq \epsilon_R \leq \Delta \quad (8-32)$$

(2) 浮点截尾: 要分正数与负数两种情况, 同样要利用定点制的结果。

$x$  为正数, 其尾数误差为

$$-\Delta < E_T \leq 0$$

因此

$$-2^e \Delta < \epsilon_T x \leq 0, \quad x > 0$$

考虑到

$$2^{c-1} \leq x < 2^c$$

由于  $x > 0$ , 故必有  $\epsilon_T \leq 0$ , 因而有  $\epsilon_T x \leq 2^{c-1} \epsilon_T$ , 从而可得出, 浮点制正数的截尾相对误差  $\epsilon_T$  满足

$$-2\Delta < \epsilon_T \leq 0, \quad x > 0 \quad (8-33)$$

$x$  为原码、反码的负数, 尾数截尾误差为

$$0 \leq E_T < \Delta$$

因此

$$0 \leq \epsilon_T x < 2^c \Delta, \quad x < 0$$

由于  $x < 0$ , 故有

$$-2^c < x \leq -2^{c-1}$$

$x < 0$ , 必有  $\epsilon_T \geq 0$ , 因而有  $\epsilon_T x \geq -\epsilon_T 2^{c-1}$ , 由此得出浮点制原码、反码的负数的截尾相对误差满足

$$-2\Delta < \epsilon_T \leq 0, \quad x < 0 \quad (8-34)$$

$x$  为补码负数, 尾数截尾误差为

$$-\Delta < E_T \leq 0,$$

因而

$$-\Delta 2^c < \epsilon_T x \leq 0, \quad x < 0$$

由于  $x < 0$ , 故有

$$-2^c < x \leq -2^{c-1}$$

$x < 0$ , 必有  $\epsilon_T \geq 0$ , 因而有  $\epsilon_T x \leq -2^{c-1} \epsilon_T$ , 因此得出浮点制补码负数截尾相对误差  $\epsilon_T$  满足

$$0 \leq \epsilon_T < 2\Delta, \quad x < 0 \quad (8-35)$$

概括起来, 浮点制截尾的相对误差为:

$$\text{原码、反码} \quad -2\Delta < \epsilon_T \leq 0 \quad (8-36)$$

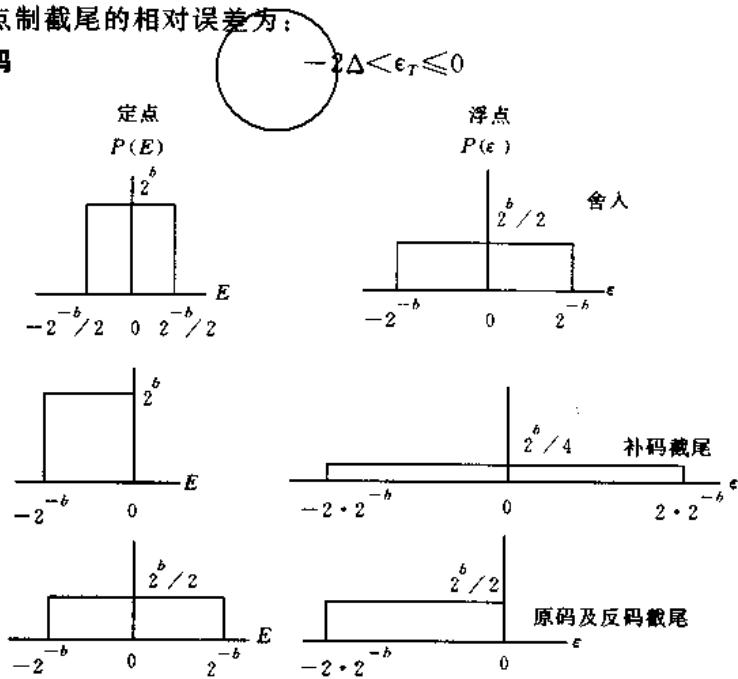


图 8-4 等概假设下, 量化误差的概率密度函数

补码

$$\begin{cases} -2\Delta < \epsilon_T \leq 0, & x > 0 \\ 0 \leq \epsilon_T < 2\Delta, & x < 0 \end{cases} \quad (8-37)$$

由以上分析看出,舍入和截尾都产生非线性关系。为了研究量化误差对数字信号处理系统精度的影响,必须了解舍入和截尾误差的特性,一般最方便的方法是把这些量化误差看成随机变量,对每种误差求出概率密度函数,现在虽然知道了量化误差的范围,但是并不知道在此范围内误差的概率。一个较为合理的假设,是设量化误差在整个可能出现的范围内是等概率的,也就是均匀分布的,在此假设下,图 8-4 示出了定点制和浮点制舍入误差及截尾误差的概率密度函数:对于定点制,变量为绝对误差  $E = Q[x] - x$ ,对于浮点制,变量为相对误差  $\epsilon = \frac{Q[x] - x}{x}$ 。

### 8.3 A/D 变换的量化效应

A/D(模数)变换器是将输入模拟信号  $x_a(t)$  转换为  $b$  位二进制数字信号的器件。 $b$  的数值可以是 8, 12 或高至 20。一个 A/D 变换器从功能上讲可以分为两部分:抽样器与量化器。A/D 变换器的非线性模型如图 8-5 所示。抽样器产生序列  $x(n) = x_a(t)|_{t=nT} = x_a(nT)$ ,  $x(n)$  具有无限精度,量化器对每个抽样序列  $x(n)$  进行截尾或舍入的量化处理,从而给出  $\hat{x}(n) = Q[x(n)]$ 。实际上这两部分是同时完成的。

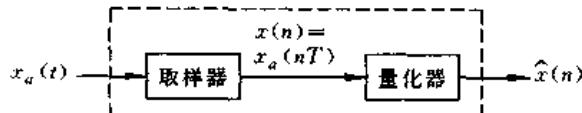


图 8-5 A/D 变换器的非线性模型

分析 A/D 变换器量化效应的目的在于选择合适的字长,以满足信噪比指标。

为了使抽样后不产生混叠失真,模拟信号必须是限带的,因而 A/D 变换器前一般都加一个前置模拟低通滤波器,它对于高于折叠频率(抽样频率之半)的频率应有足够的衰减(例如至少大于 40dB),且要求信号基带内的波纹要足够小。此外,由于 A/D 变换器总是定点制的,必须使信号不超过 A/D 变换的动态范围,为此模拟输入信号必须乘一个比例因子,使得它满足 A/D 变换器动态范围的要求,即

$$x(n) = Ax_a(t)|_{t=nT} = Ax_a(nT) \quad (8-38)$$

设量化器输出抽样值表示成  $(b+1)$  位的补码定点小数,二进制小数点后为  $b$  位。输入到量化器的精确抽样值  $x(n)$  要舍入到最靠近的量化层标准值,以得到量化抽样值  $\hat{x}(n)$ ,量化器对补码定点制输入信号的动态范围为

$$-\left(1 + \frac{2^{-b}}{2}\right) < x(n) < \left(1 - \frac{2^{-b}}{2}\right) \quad (8-39)$$

量化误差为

$$e(n) = Q[x(n)] - x(n) = \hat{x}(n) - x(n) \quad (8-40)$$

A/D 变换器的量化特性主要取决于所采用的数的表示方式和量化方式,对于补码舍入处理,由(8-21)式知

$$-\frac{\Delta}{2} < \epsilon_R(n) \leq \frac{\Delta}{2}, \quad \Delta = 2^{-b} \quad (8-41)$$

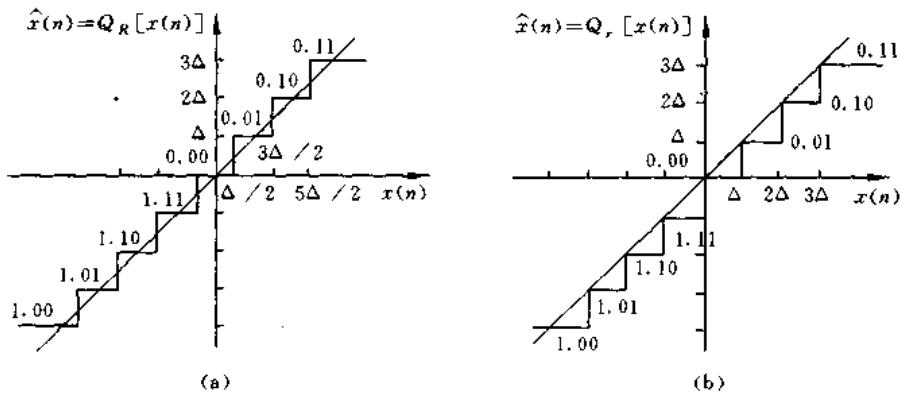


图 8-6 补码表示时,A/D 变换的量化特性( $b=2, \Delta=2^{-6}$ )

(a) 舍入; (b) 截尾

这种 A/D 变换器的量化特性如图 8-6(a)所示。

如果输入抽样的精确值落在(8-39)式规定的范围之外,将会产生错误,如图 8-6(a)所示,超过  $0.11(3\Delta)$  将变成  $1.00(-4\Delta)$ 。此时,可采用限幅办法,即当抽样值超过  $1 - \frac{2^{-b}}{2}$  时,皆取为量化值  $1 - 2^{-b}$ ;抽样值小于  $-1 - \frac{2^{-b}}{2}$  时,皆取成量化值  $-1$ 。一般我们不希望出现这种限幅作用,因而必须将输入幅度加权以满足(8-39)式。

对于补码截尾处理,由(8-13)式及(8-17)式知,A/D 变换器的量化误差为

$$-\Delta < \epsilon_T(n) \leq 0 \quad (8-42)$$

这种 A/D 变换器的量化特性如图 8-6(b)所示。

### 一、量化误差的统计分析

上面虽然分析了 A/D 变换的量化误差范围,但要精确知道所有  $n$  下的量化误差  $\epsilon(n)$  几乎是不可能的,也无此必要。一般只要知道量化误差的一些平均效应就够了,就可用来做为确定 A/D 变换器所需字长的依据,所以,对量化误差适合于采用统计分析方法。

在统计分析中,对误差  $\epsilon(n)$  的统计特性作了如下的一些假定:

- (1)  $\epsilon(n)$  是平稳随机序列;
- (2)  $\epsilon(n)$  与抽样信号  $x(n)$  是不相关的;
- (3)  $\epsilon(n)$  序列本身的任意两个值之间是不相关的,即  $\epsilon(n)$  是白噪声序列;
- (4)  $\epsilon(n)$  在其误差范围内为均匀等概分布的。

根据这些假定,量化误差  $\epsilon(n)$  就是一个与信号序列完全不相关的白色噪声序列,也称为量化噪声。它与信号的关系是相加性的。在这些假定下,实际的 A/D 变换的非线性

模型(如图 8-5 所示)就变成等效线性过程的统计模型,如图 8-7 所示,即在理想抽样器的输出端加入了一个量化白噪声序列  $e(n)$ 。

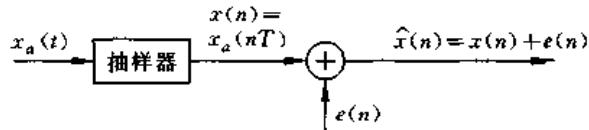


图 8-7 A/D 变换的统计模型

以上假定是为了简化量化误差的分析,实际工作中,有时并不满足这些要求,例如如果输入  $x_a(t)$  是阶跃、方波、正弦等简单信号,量化后显然不能认为误差是统计独立的和白色的。但是信号越不规则,例如语言信号及音乐信号,这种假定就越接近于实际,也就是说只要信号足够复杂且量化台阶足够小,则此统计模型就更加有效。注意对于反码截尾和原码截尾来说,不能假定误差与信号互相独立,因为误差的符号与信号的符号总是相反的。

下面我们来研究量化误差  $e(n)$  的统计性能,也就是研究它的平均值  $m_e$ 、方差  $\sigma_e^2$ 。

对于定点舍入情况,误差序列  $e(n)$  的概率密度函数为(如图 8-4 所示)

$$p[e(n)] = \begin{cases} \frac{1}{\Delta}, & -\frac{\Delta}{2} < e(n) \leq \frac{\Delta}{2} \\ 0, & \text{其他} \end{cases} \quad (8-43)$$

可求出其均值  $m_e$  与方差  $\sigma_e^2$  分别为

$$m_e = E[e(n)] = \int_{-\Delta/2}^{\Delta/2} ep(e) de = 0 \quad (8-44)$$

$$\begin{aligned} \sigma_e^2 &= E((e(n) - m_e)^2) = \int_{-\Delta/2}^{\Delta/2} [e - m_e]^2 p(e) de \\ &= \int_{-\Delta/2}^{\Delta/2} \frac{1}{\Delta} e^2 de = \frac{\Delta^2}{12} = \frac{2^{-2b}}{12} \end{aligned} \quad (8-45)$$

其中  $E[\cdot]$  表示求数学期望,即统计平均。由于  $e(n)$  是平稳的,在求数学期望时与  $n$  无关,所以可以不标  $n$ 。

对于定点补码截尾情况,误差序列  $e(n)$  的概率密度函数为(如图 8-4 所示)

$$p[e(n)] = \begin{cases} \frac{1}{\Delta}, & -\Delta < e(n) \leq 0 \\ 0, & \text{其他} \end{cases} \quad (8-46)$$

可求出其均值  $m_e$  与方差  $\sigma_e^2$  分别为

$$m_e = \int_{-\Delta}^0 \frac{1}{\Delta} ede = -\frac{\Delta}{2} = -\frac{2^{-b}}{2} \quad (8-47)$$

$$\sigma_e^2 = \int_{-\Delta}^0 \left[ e + \frac{\Delta}{2} \right]^2 \frac{1}{\Delta} de = \frac{\Delta^2}{12} = \frac{2^{-2b}}{12} \quad (8-48)$$

由于假定  $e(n)$  为白噪声序列,故舍入误差的自协方差序列为

$$r_{ee}(n) = E\{[e(m) - m_e][e(m+n) - m_e]\} = E[e(m)e(m+n)] = \sigma_e^2 \delta(n) \quad (8-49)$$

其中  $E[\cdot]$  表示求统计平均值。

从上面的关系式看出,量化噪声的方差和 A/D 变换的字长  $b$  有关,字长越长,则量化间距  $\Delta = 2^{-b}$  越小,因而量化噪声的方差  $\sigma_e^2$  就越小。同时我们也看出,采用定点制补码截尾时,A/D 变换器的量化噪声有直流分量(均值不为零),这将影响信号的频谱结构,所以很少采用它。

由于在抽样模拟信号的数字处理中,把量化噪声看成相加性噪声序列,量化过程看成是无限精度的信号与量化噪声的叠加,因而信噪比是一个衡量量化效应的重要指标。

对于舍入处理,设信号  $x(n)$  的功率为  $\sigma_x^2$ ,则信号功率与噪声功率之比为

$$\frac{\sigma_x^2}{\sigma_e^2} = \frac{\sigma_x^2}{2^{-2b}/12} = 12 \cdot 2^{2b} \cdot \sigma_x^2$$

表示成分贝数,则为

$$\frac{S}{N} = 10 \log_{10} \frac{\sigma_x^2}{\sigma_e^2} = 6.02b + 10.79 + 10 \log_{10} \sigma_x^2 \quad (dB) \quad (8-50)$$

由此式看出,信号功率  $\sigma_x^2$  越大,信噪比当然越高;另一方面,随着字长  $b$  的增加,信噪比也增大。字长  $b$  每增加一位,则信噪比增加约 6dB。

前面已说到,当输入信号超过 A/D 变换器的动态范围时,必须压缩输入信号幅度,以避免采用限幅,也就是要将输入信号缩小为  $Ax(n)$ , $0 < A < 1$ ,然后对其量化。由于  $Ax(n)$  的方差是  $A^2 \sigma_x^2$ ,故此时的信噪比为

$$\frac{S}{N} = 10 \log_{10} \left( \frac{A^2 \sigma_x^2}{\sigma_e^2} \right) = 6.02b + 10.79 + 10 \log_{10} \sigma_x^2 + 20 \log_{10} A \quad (dB) \quad (8-51)$$

上式中, $0 < A < 1$ ,则  $\log_{10} A$  为负数,因此压缩输入信号幅度,将使信噪比减小。

许多模拟信号,例如语言和音乐,可以视为一个随机过程。因此可用概率分布来表示这些信号,它们在幅值为零附近,概率分布曲线出现峰值,且随幅度增加分布曲线值急剧下降,抽样幅度超过信号均方根值 3 倍或 4 倍的概率极小。因此,如令  $A = 1/4\sigma_x$ ,则不出现限幅失真的概率是极高的,此时信噪比为

$$\begin{aligned} \frac{S}{N} &= 10 \log_{10} \left( \frac{A^2 \sigma_x^2}{\sigma_e^2} \right) = 10 \log_{10} \left( \frac{1}{16\sigma_e^2} \right) = 10 \log_{10} \left( \frac{3}{4} \cdot 2^{2b} \right) \\ &\approx 6b - 1.25 \quad (dB) \end{aligned} \quad (8-52)$$

这个关系式很重要,由它可以看出,若需得到信噪比大于 70dB,至少需要满足  $b = 12$  bit。当然,字长越长,A/D 变换的信噪比越高。但是,字长过长也无必要,因为输入信号  $x_a(t)$  本身也有一定的信噪比,字长长到 A/D 变换器的量化噪声比  $x_a(t)$  的噪声电平更低就没有意义了。

这样,我们看出,提高信噪比的办法,一是增大输入信号,但这受到 A/D 变换器动态范围的限制;二是增加字长  $b$ ,但它受到输入信号  $x_a(t)$  的信噪比的限制。

## 二、白噪声(量化噪声)通过线性系统

下面讨论量化的序列  $\hat{x}(n) = x(n) + e(n)$  通过线性移不变系统,而且假定系统是完全理想的,即是无限精度的,也就是说系统实现时带来的误差以及运算带来的误差暂都不考虑,把它们看成是独立于量化噪声而引起的误差,可单独计算,然后将结果叠加。

因为我们已认为  $x(n)$  和  $e(n)$  不相关, 且系统是线性的, 根据叠加原理, 系统的输出为

$$\begin{aligned}\hat{y}(n) &= \hat{x}(n) * h(n) = x(n) * h(n) + e(n) * h(n) \\ &= \sum_{m=0}^{\infty} h(m)x(n-m) + \sum_{m=0}^{\infty} h(m)e(n-m) = y(n) + f(n)\end{aligned}$$

其中  $y(n)$  是  $x(n)$  的响应, 即

$$y(n) = x(n) * h(n) = \sum_{m=0}^{\infty} h(m)x(n-m)$$

$f(n)$  是量化噪声  $e(n)$  的响应, 即

$$f(n) = \hat{y}(n) - y(n) = e(n) * h(n) = \sum_{m=0}^{\infty} h(m)e(n-m)$$

图 8-8 表示了量化噪声通过线性系统的框图。

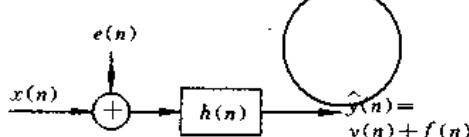


图 8-8 量化噪声通过线性系统

由于  $e(n)$  与  $x(n)$  互不相关, 故在计算输出噪声功率时, 可以不管  $x(n)$  的影响。因此, 舍入噪声  $e(n)$  造成的输出噪声  $f(n)$  的方差为

$$\begin{aligned}\sigma_f^2 &= E[f^2(n)] = E\left[\sum_{m=0}^{\infty} h(m)e(n-m) \sum_{l=0}^{\infty} h(l)e(n-l)\right] \\ &= \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} h(m)h(l)E[e(n-m)e(n-l)]\end{aligned}$$

由于  $e(n)$  是白色的, 故  $e(n)$  的各序列值之间互不相关, 因而

$$E[e(n-m)e(n-l)] = \delta(m-l)\sigma_e^2$$

将它代入  $\sigma_f^2$  式中, 可得

$$\sigma_f^2 = \sum_{m=0}^{\infty} \sum_{l=0}^{\infty} h(m)h(l)\delta(m-l)\sigma_e^2 = \sigma_e^2 \sum_{m=0}^{\infty} h^2(m) \quad (8-53)$$

按照帕塞瓦定理, 考虑到  $h(n)$  是实序列, 则有

$$\sum_{m=0}^{\infty} h^2(m) = \frac{1}{2\pi j} \oint_c H(z)H(z^{-1}) \frac{dz}{z}$$

这样, (8-53)式可改写成

$$\sigma_f^2 = \frac{\sigma_e^2}{2\pi j} \oint_c H(z)H(z^{-1}) \frac{dz}{z} \quad (8-54)$$

或者在单位圆上计算, 可得

$$\sigma_f^2 = \frac{\sigma_e^2}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 d\omega = \frac{\sigma_e^2}{2\pi} \int_{-\pi}^{\pi} |H(e^{j\omega})|^2 d\omega \quad (8-55)$$

如果  $e(n)$  是补码截尾白噪声, 经过分析可知, 输出噪声的方差仍为(8-54)式或(8-55)式, 但还有一个直流分量

$$m_f = E[f(n)] = E\left[\sum_{m=0}^{\infty} h(m)e(n-m)\right] = \sum_{m=0}^{\infty} h(m)E[e(n-m)]$$

$$= m_e \sum_{m=0}^{\infty} h(m) = m_e H(e^{j\omega}) \quad (8-56)$$

以上这些分析对于白噪声通过线性系统都是合适的,因此这些结果在下面还将用到。

## 8.4 数字滤波器的系数量化效应

理想数字滤波器的系统函数为

$$H(z) = \frac{\sum_{k=0}^M b_k z^{-k}}{1 - \sum_{k=1}^N a_k z^{-k}} = \frac{B(z)}{A(z)} \quad (8-57)$$

由理论设计出的理想数字滤波器系统函数的各系数  $b_k, a_k$  都是无限精度的,但实际实现时,滤波器的所有系数都必须以有限长的二进制码形式存放的存储器中,因而必须对理想的系数值加以量化,就会与原系数值有偏差,这就造成滤波器的零点、极点位置发生偏移,这样实际系统函数将与原设计的有所不同,也就是系统的实际频率响应与按要求设计出的频率响应有偏离,甚至严重时,如果  $z$  平面单位圆内极点偏移到单位圆外,系统就不稳定,滤波器就不能使用了。

系数量化对滤波器性能的影响当然和字长有关,但是也和滤波器的结构形式密切相关。因而选择合适的结构,对减小系数量化的影响是非常重要的。分析数字滤波器系数量化误差的目的在于选择合适的字长,以满足频率响应指标的要求。

### 一、系数量化对滤波器零、极点位置的影响

#### 1. 系统极点(零点)位置对系数量化的灵敏度

系数量化后使滤波器的特性与所要求的频率响应不同,或说表现在极点、零点离开了它们应有的位置,所以一个网络结构对系数量化的灵敏度是用系数量化引起的极点、零点的位置误差来衡量的。不同形式的系统结构,在相同的系数“量化步距”情况下,其量化灵敏度是不同的,这是比较各种结构形式的重要标准。

极点位置灵敏度是指每个极点位置对各系数偏差的敏感程度。其分析方法,同样适用于零点,但是极点对系统的影响更大,直接影响到系统的稳定性,更为人们所注意。

(8-57)式表示了一个无限精度的  $N$  阶直接型结构的 IIR 滤波器的系统函数,其中  $a_k, b_k$  是系统直接型结构所求出的无限精度的系数。若系数在实际实现时已被量化为  $\hat{a}_k, \hat{b}_k$ ,即

$$\begin{aligned} \hat{a}_k &= a_k - \Delta a_k \\ \hat{b}_k &= b_k - \Delta b_k \end{aligned} \quad (8-58)$$

其中  $\Delta a_k, \Delta b_k$  是由量化造成的系数误差,则实际实现的系数函数为

$$\hat{H}(z) = \frac{\sum_{k=0}^M \hat{b}_k z^{-k}}{1 - \sum_{k=1}^N \hat{a}_k z^{-k}} \quad (8-59)$$

下面我们来讨论系数量化误差对极点的影响。由(8-57)式,原系统函数  $H(z)$  的分母多项式为

$$A(z) = 1 - \sum_{k=1}^N a_k z^{-k} \quad (8-60)$$

也可表示成因式形式

$$A(z) = \prod_{i=1}^N (1 - z_i z^{-1}) \quad (8-61)$$

令  $A(z)=0$ ,就得到  $H(z)$  的极点

$$z = z_i, \quad i = 1, 2, \dots, N$$

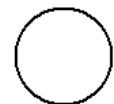
设系数量化后  $\hat{H}(z)$  的极点为

$$z_i + \Delta z_i, \quad i = 1, 2, \dots, N$$

$\Delta z_i$  为极点位置的偏差量,它是由各个系数偏差  $\Delta a_k$  引起的,因此

$$\Delta z_i = \sum_{k=1}^N \frac{\partial z_i}{\partial a_k} \Delta a_k, \quad i = 1, 2, \dots, N \quad (8-62)$$

由此式看出, $\partial z_i / \partial a_k$  值的大小决定着系数  $a_k$  的偏差  $\Delta a_k$  对极点位置偏差  $\Delta z_i$  的影响程度, $\partial z_i / \partial a_k$  越大, $\Delta a_k$  对  $\Delta z_i$  的影响也越大, $\partial z_i / \partial a_k$  越小, $\Delta a_k$  对  $\Delta z_i$  的影响就越小。所以  $\partial z_i / \partial a_k$  就是极点  $z_i$  对系数  $a_k$  变化的灵敏度。下面根据  $A(z)$  来求这个极点位置灵敏度  $\partial z_i / \partial a_k$  的表达式,根据复合函数的微分法则可得



$$\left( \frac{\partial A(z)}{\partial z_i} \right)_{z=z_i} \left( \frac{\partial z_i}{\partial a_k} \right) = \left( \frac{\partial A(z)}{\partial a_k} \right)_{z=z_i}$$

由此得出

$$\frac{\partial z_i}{\partial a_k} = \frac{\partial A(z)/\partial a_k}{\partial A(z)/\partial z_i} \Big|_{z=z_i} \quad (8-63)$$

由(8-60)式,可以得出

$$\frac{\partial A(z)}{\partial a_k} = -z^{-k} \quad (8-64)$$

根据(8-61)式的因式表达式,可以求出(假定各  $z_i$  全是单根)

$$\frac{\partial A(z)}{\partial z_i} = -z^{-1} \prod_{\substack{l=1 \\ l \neq i}}^N (1 - z_l z^{-1}) = -z^{-N} \prod_{\substack{l=1 \\ l \neq i}}^N (z - z_l) \quad (8-65)$$

将(8-64)式及(8-65)式代入到(8-63)式,就得到极点位置灵敏度为

$$\frac{\partial z_i}{\partial a_k} = \frac{z_i^{N-k}}{\prod_{\substack{l=1 \\ l \neq i}}^N (z_i - z_l)} \quad (8-66)$$

这个公式非常重要。它表示由  $H(z)$  的分母的第  $k$  个系数  $a_k$  的偏差造成第  $i$  个极点  $z_i$  的偏差的灵敏度。此式只对单阶极点有效,多阶极点可进行类似的推导。对于直接型结构,由于它的零点只取决于分子多项式的系数  $b_k$ ,因而对于零点可得到完全相似的结果。即由  $b_k$  的偏差造成零点位置偏差的灵敏度表达式。

将(8-66)式代入到(8-62)式,可以得到各  $a_k$  的偏差  $\Delta a_k$  引起的第  $i$  个极点位置的变化量

$$\Delta z_i = \sum_{k=1}^N \frac{z_i^{N-k}}{\prod_{l=1, l \neq i}^N (z_i - z_l)} \Delta a_k, \quad i = 1, 2, \dots, N \quad (8-67)$$

(8-66)式的分母中的每一个因子( $z_i - z_l$ )是一个由极点 $z_l$ 指向 $z_i$ 的矢量,而整个分母正是所有其他极点 $z_l$ ( $l \neq i$ )指向该极点 $z_i$ 的矢量积。这些矢量越长,即极点彼此间越远时,极点位置灵敏度就越低;这些矢量越短,即极点彼此越密集时,极点位置灵敏度就越高。例如图8-9(a)表示共轭极点在 $z$ 平面虚轴附近的带通滤波器,图8-9(b)表示共轭极点在实轴附近的低通滤波器,前者极点间距离比后者长,因此前者极点位置灵敏度比后者小,也就是说,在相同程度的系数量化下所造成的极点位置误差前者比后者要小。

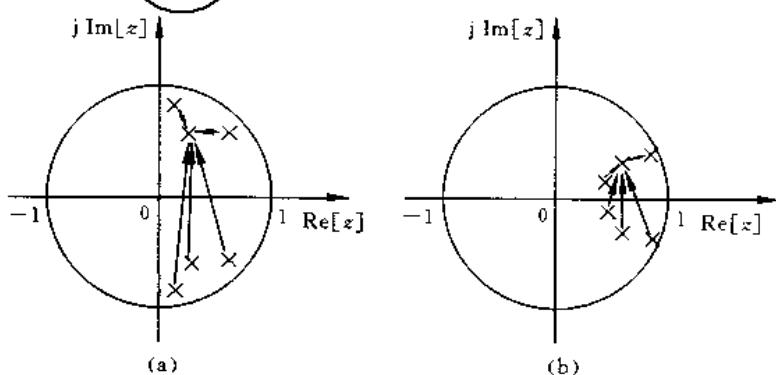


图 8-9 极点位置灵敏度与极点间距离成反比

(a) 极点间距离长;(b) 极点间距离短

高阶直接型结构滤波器的极点数目多而密集,而低阶直接型结构滤波器的极点数目少而稀疏,因而前者对系数量化误差要敏感得多,同理,并联型结构及级联型结构将比直接型结构要好得多。这是因为,在级联型和并联型结构中,每一对共轭极点是单独用一个二阶子系统实现的,其他二阶子系统的系数变化对本节子系统的极点位置不产生任何影响,由于每对极点只受与之有关的两个系数的影响,而且级联或并联后,每个子系统的极点密集度就比直接型高阶网络的要稀疏得多,因而极点位置受系数量化的影响比直接型结构要小得多,如图8-10所示。

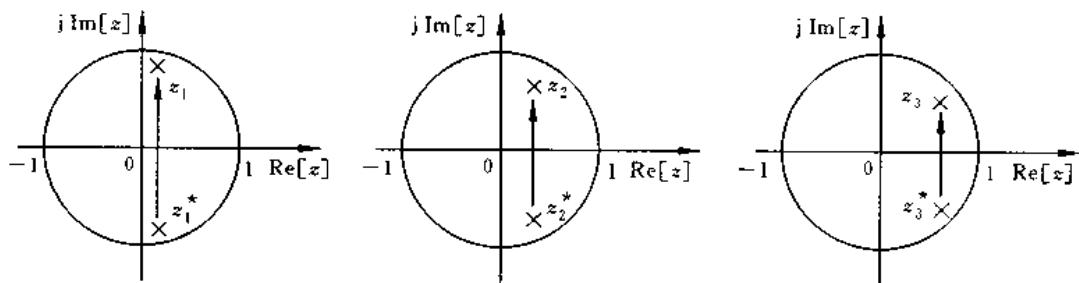


图 8-10 级联型、并联型极点密度(与图 8-9(a)相对应)

下面用一个例子,通过极点位置灵敏度来说明极点位置变化与系数字长的关系。

[例 8-1] 设数字滤波器的系统函数为

$$H(z) = \frac{0.0373}{1 + 1.7z^{-1} + 0.745z^{-2}} = \frac{0.373}{1 - a_1 z^{-1} - a_2 z^{-2}}$$

利用  $a_2$  变化造成的极点位置灵敏度, 为保持极点在其正常值的 0.5% 内变化, 试确定所需要的最小字长。

解 令  $H(z)$  的分母为零, 即

$$1 + 1.7z^{-1} + 0.745z^{-2} = 0$$

由此得出  $H(z)$  的两个极点为

$$z_1 = -0.85 + j0.15, \quad z_2 = -0.85 - j0.15$$

则

$$|z_1| = |z_2| = 0.863$$

可直接利用(8-72)式, 看  $a_2$  变化的影响:

$$\frac{\partial z_1}{\partial a_2} = \frac{1}{z_1 - z_2} = \frac{1}{j0.3} = 3.3333e^{-j90^\circ}$$

$$\frac{\partial z_2}{\partial a_2} = \frac{1}{z_2 - z_1} = \frac{1}{-j0.3} = 3.3333e^{j90^\circ}$$

可以看出,  $a_2$  的变化对  $z_1, z_2$  影响的大小是相同的。由于题意要求研究  $a_2$  对  $z_2$  (或  $z_1$ ) 的影响, 故只需考虑绝对值即可, 因而有

$$|\Delta z_2| = \left| \frac{\partial z_2}{\partial a_2} \right| |\Delta a_2|$$

所以

$$|\Delta a_2| = \left| \frac{\Delta z_2}{\partial z_2 / \partial a_2} \right| = \frac{0.5\% \times |z_2|}{3.3333} = 1.295 \times 10^{-3}$$

这样, 所需的系数“量化步距”应为  $2|\Delta a_2| = 2.590 \times 10^{-3}$ 。如果采用定点二进制小数表示, 设小数点后为  $b$  位, 则分辨率为  $2^{-b}$ , 因为应满足  $2^{-b} < 2.590 \times 10^{-3}$ ,  $b$  取整数, 可得  $b=9$ 。由此可知, 系数字长  $b=9$  才能满足性能要求。

总括说来, 系数量化对极点位置的影响与极点本身状态以及滤波器的结构都有密切关系。对于高阶滤波器来说, 应该避免采用直接型结构, 而应采用分解为基本二阶节或一阶节的级联型结构或并联型结构。这样, 在给定字长的情况下, 可以使系数量化的影响最小。对于极点灵敏度很高的场合, 可以采用双精度的系数, 以便有效地达到精度的要求。

## 2. 系数量化对二阶子系统极点位置的影响

上面说到, 就极点位置灵敏度来看, 级联型和并联型优于直接型。但是级联型和并联型的基本子系统是二阶节, 如何实现这个基本二阶节也有着不同的情况。

设二阶 IIR 系统的差分方程为

$$y(n) = x(n) - a_1 y(n-1) - a_2 y(n-2)$$

其系统函数为

$$H(z) = \frac{1}{1 + a_1 z^{-1} + a_2 z^{-2}}$$

设  $H(z)$  有一对共轭复极点

$$z_{1,2} = re^{\pm j\theta} \quad (8-68)$$

则有

$$1 + a_1 z^{-1} + a_2 z^{-2} = (1 - r e^{j\theta} z^{-1})(1 - r e^{-j\theta} z^{-1}) = 1 - 2r \cos\theta \cdot z^{-1} + r^2 z^{-2}$$

$H(z)$ 的直接型实现表示在图 8-11 上。由此看出

$$r^2 = a_2, \quad r \cos\theta = -\frac{a_1}{2} \quad (8-69)$$

若系数量化,也就是将  $a_1, a_2$  量化,由于  $a_2 = r^2$  决定了极点的半径,而  $a_1 = 2r \cos\theta$  则决定了极点在实轴上的坐标。如果  $a_1, a_2$  用三位字长  $b=3$  表示(不包括符号位),表 8-2 表示了三位字长只能有 8 种不同值,因而只能表示 8 种半径  $r$  值和  $\pm 7/8$  之间的 15 种实轴坐标  $r \cos\theta$ ,这样,三位字长的系数所能表达的极点位置就是在同心圆(对应于  $a_2 = r^2$  的量化)及垂直线(对应于  $a_1 = -2r \cos\theta$  的量化)的网格交

点上,如图 8-12 所示。可以看出极点在  $z$  平面的网格点子很不均匀,实轴附近分布较稀,半径大的地方分布很密,这就使得实轴附近的极点(对应于高通、低通滤波器)量化误差大,而虚轴附近的极点(对应于带通滤波器)量化误差小。当然这种分布只是二阶直接型结构的情况,不同结构的滤波器,系数量化对零、极点位置的影响是不同的,这在前面已经讨论过了。

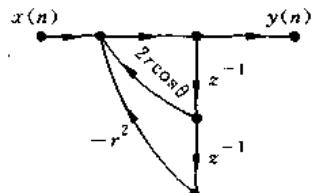


图 8-11 复共轭极点对组成的二阶基本节的直接型实现

表 8-2 三位字长( $b=3$ )系数所能表达的共轭极点参数

$ a_1 $ 三位二进制 码 $a_1, a_2, a_3$	所表达的 $ a_1 $ 值	极点横坐标 $ r \cos\theta  = \left  \frac{a_1}{2} \right $	$a_2$ 三位二进制 码 $a_1, a_2, a_3$	所表达的 $a_2$ 值	极点半径 $r = \sqrt{a_2}$
0.00	0.00	0.000	0.000	0.000	0.000
0.01	0.25	0.125	0.001	0.125	0.354
0.10	0.50	0.250	0.010	0.250	0.500
0.11	0.75	0.375	0.011	0.375	0.612
1.00	1.00	0.500	0.100	0.500	0.707
1.01	1.25	0.625	0.101	0.625	0.791
1.10	1.50	0.750	0.110	0.750	0.866
1.11	1.75	0.875	0.111	0.875	0.935

如果所需要的理想极点不在这些网格节点上时,就只能以最靠近的一个节点来代替这一极点位置,这样就会引入极点位置误差,有时,可能使共轭极点变成实极点(实极点位置不在图 8-12 所示的网格点上),甚至使滤波器不稳定。

由雷德(Rader)和戈尔德(Gold)提出的二阶节对偶式结构如图 8-13 所示。其差分方程可表示为。

$$\begin{cases} y_1(n) = r(\cos\theta)y_1(n-1) - r(\sin\theta)y(n-1) + x(n) \\ y(n) = r(\sin\theta)y_1(n-1) + r(\cos\theta)y(n-1) \end{cases} \quad (8-70)$$

对上两式取  $z$  变换,且利用  $z$  变换的移位定理,即可得出

$$H(z) = \frac{Y(z)}{X(z)} = \frac{rz^{-1}(\sin\theta)}{1 - 2r(\cos\theta)z^{-1} + r^2z^{-2}} \quad (8-71)$$

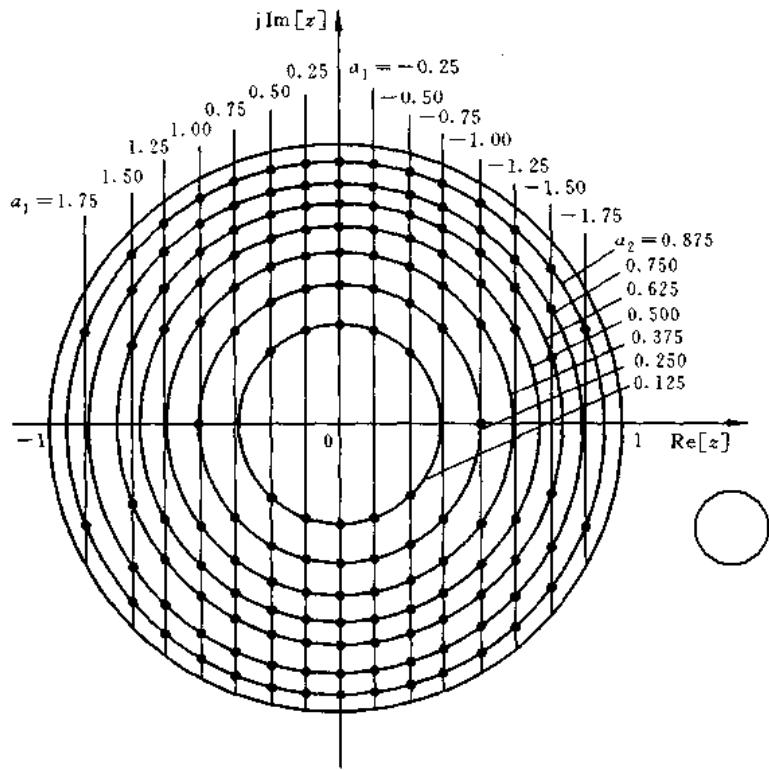


图 8-12 三位字长( $b=3$ )系数所能表达的共轭极点位置

因而对于无限精度的系数,图 8-13 与图 8-11 两网络结构的系统函数的极点是相同的。但是图 8-13 中,当系数量化时,是对  $r\cos\theta$  及  $r\sin\theta$  进行量化,因而所得到的网格点子在  $z$  平面上是均匀分布的,如图 8-14 所示,这和图 8-12 是不同的。因而这里系数量化后对  $z$  平面上的所有区域,所产生的误差是相同的。

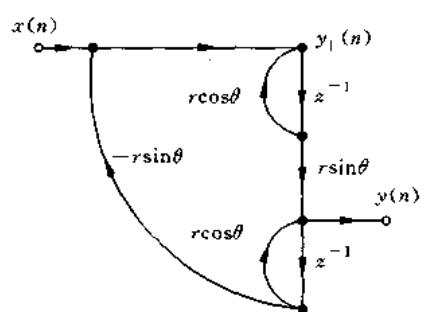


图 8-13 实现复共轭极点的对偶式结构

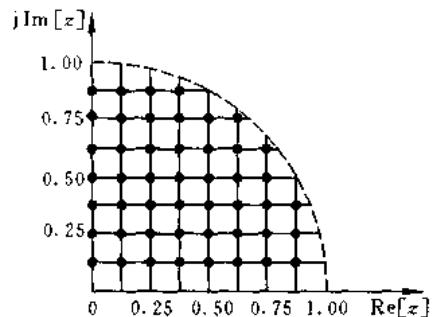


图 8-14 当系数  $r\cos\theta$  和  $r\sin\theta$  量化成三位码时,  
图 8-13 网络极点可能位置的网格点

## 二、系数量化效应的统计分析

当然,我们最感兴趣的是系数量化所造成的滤波器频率特性的偏差。由上面的讨论

已知,因极点位置灵敏度形成的极点偏差并不能直接求得频率响应的偏差,尤其是在高阶情况下,系数多,它们的量化误差就更具有随机性,因而可采用统计方法,把系数量化误差“视为”一个随机变量,这样来估计滤波器频率响应的偏差。

### 1. IIR 数字滤波器系数量化的统计分析

$N$  阶 IIR 直接型结构当系数无限精度时的系统函数为

$$\hat{H}(z) = \frac{\sum_{i=0}^M b_i z^{-i}}{1 - \sum_{i=1}^N a_i z^{-i}} = \frac{B(z)}{A(z)} \quad (8-72)$$

设量化后的系数为

$$\hat{a}_i = Q[a_i] = a_i + \alpha_i \quad (8-73)$$

$$\hat{b}_i = Q[b_i] = b_i + \beta_i \quad (8-74)$$

而  $\alpha_i, \beta_i$  是系数的量化误差。如果系数  $\hat{a}_i, \hat{b}_i$  采用小数点后  $b$  位字长,并假定采用舍入方式进行量化,则误差  $\alpha_i, \beta_i$  的变化范围皆为  $(-\frac{\Delta}{2}, +\frac{\Delta}{2})$ ,一般都假定在此范围内误差是均匀分布的因而均值为零,方差为  $\frac{\Delta^2}{12}$ 。

故系数量化后,实际的系统函数为

$$\hat{H}(z) = \frac{\sum_{i=0}^N \hat{b}_i z^{-i}}{1 - \sum_{i=1}^M \hat{a}_i z^{-i}} \quad (8-75)$$

为了研究系数量化造成的频率响应的偏差,我们先来研究系统函数的偏差,即  $\hat{H}(z)$  与无限精度的  $H(z)$  之偏差。令输入为  $x(n)$ ,系数为无限精度的滤波器输出为  $y(n)$ ,系数量化后滤波器输出为  $\hat{y}(n)$ (当然不考虑运算的有限字长影响),则可定义输出误差序列为

$$\begin{aligned} e(n) &= \hat{y}(n) - y(n) \\ &= \left[ \sum_{i=0}^N \hat{b}_i x(n-i) + \sum_{i=1}^N \hat{a}_i \hat{y}(n-i) \right] - \left[ \sum_{i=0}^N b_i x(n-i) + \sum_{i=1}^N a_i y(n-i) \right] \\ &= \sum_{i=0}^N \beta_i x(n-i) + \sum_{i=1}^N a_i e(n-i) + \sum_{i=1}^N \alpha_i y(n-i) + \sum_{i=1}^N \alpha_i e(n-i) \end{aligned} \quad (8-76)$$

将等式右边最后一项(二阶无穷小)忽略,可得

$$e(n) = \sum_{i=0}^N \beta_i x(n-i) + \sum_{i=1}^N a_i e(n-i) + \sum_{i=1}^N \alpha_i y(n-i)$$

对此式取  $z$  变换,得

$$E(z) = \beta(z)X(z) + \sum_{i=1}^N \alpha_i z^{-i} E(z) + \alpha(z)Y(z)$$

即

$$\beta(z)X(z) + \alpha(z)Y(z) - E(z)A(z) = 0 \quad (8-77)$$

其中

$$\begin{aligned}\alpha(z) &= \sum_{i=1}^N \alpha_i z^{-i}, & \beta(z) &= \sum_{i=0}^N \beta_i z^{-i} \\ A(z) &= 1 - \sum_{i=1}^N a_i z^{-i}, & E(z) &= \sum_{i=0}^N e(i) z^{-i}\end{aligned}\quad (8-78)$$

将  $Y(z) = H(z)X(z)$  代入(8-77)式, 可得

$$E(z) = \left[ \frac{\beta(z) + \alpha(z)H(z)}{A(z)} \right] X(z) \quad (8-79)$$

这就是系数量化造成的滤波器输出误差的  $z$  变换, 由于  $\hat{y}(n) = y(n) + e(n)$ , 故  $\hat{Y}(z) = Y(z) + E(z)$ , 因而有

$$\hat{Y}(z) = \left[ H(z) + \frac{\beta(z) + \alpha(z)H(z)}{A(z)} \right] X(z)$$

则可得出

$$\hat{H}(z) = \frac{\hat{Y}(z)}{X(z)} = H(z) + \frac{\beta(z) + \alpha(z)H(z)}{A(z)} \quad (8-80)$$

从而得到由系数量化造成的系统函数的偏差

$$H_E(z) = \hat{H}(z) - H(z) = \frac{\beta(z) + \alpha(z)H(z)}{A(z)} \quad (8-81)$$

所以, 系数量化后实际滤波器的系统函数  $\hat{H}(z)$  可表示成无限精度滤波器的系统函数  $H(z)$  和偏差滤波器系统函数  $H_E(z)$  的并联, 如图 8-15 所示。在(8-81)式中, 代入  $z = e^{j\omega}$ , 即可得到系数量化造成的频响的偏差

$$H_E(e^{j\omega}) = \hat{H}(e^{j\omega}) - H(e^{j\omega}) \quad (8-82)$$

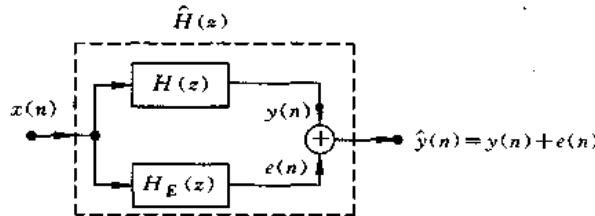


图 8-15 系数量化后滤波器的等效框图结构

一般用频响的均方偏差  $\epsilon^2$  作为频响的偏差的度量, 即

$$\epsilon^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |H_E(e^{j\omega})|^2 d\omega = \frac{1}{2\pi j} \oint_c H_E(z) H_E(z^{-1}) \frac{dz}{z} \quad (8-83)$$

将  $H_E(z)$  的表达式(8-81)式代入(8-83)式, 即可算出均方偏差  $\epsilon^2$ 。但是  $\alpha_i, \beta_i$  值一般并不能精确知道, 因而我们用  $\epsilon^2$  的统计平均, 即频响偏差的均方值  $\sigma^2$  来估计  $\epsilon^2$  的大小。首先我们假定  $\alpha_i, \beta_i$  都是独立的均匀分布的随机变量, 则在舍入情况下, 它们的平均值及方差分别为

$$E[\alpha_i] = E[\beta_i] = 0, \quad \sigma^2 = E[\alpha_i^2] = E[\beta_i^2] = \frac{\Delta^2}{12}$$

对均方偏差求统计平均,得到频响偏差的均方值为

$$\begin{aligned}\sigma_{\epsilon}^2 &= E[\epsilon^2] = E\left[\frac{1}{2\pi} \int_{-\pi}^{\pi} |H_E(e^{j\omega})|^2 d\omega\right] \\ &= E\left[\frac{1}{2\pi j} \oint_c H_E(z) H_E(z^{-1}) \frac{dz}{z}\right]\end{aligned}\quad (8-84)$$

将(8-81)式代入(8-84)式,并考虑到(8-78)式,则有

$$\sigma_{\epsilon}^2 = E\left\{\left[\frac{1}{2\pi j} \oint_c \frac{\sum_{i=0}^N \beta_i z^{-i} + H(z) \sum_{i=1}^N \alpha_i z^{-i}}{A(z)}\right] \cdot \left[\frac{\sum_{j=0}^N \beta_j z^j + H(z^{-1}) \sum_{j=1}^N \alpha_j z^j}{A(z^{-1})}\right] \frac{dz}{z}\right\}$$

考虑到  $\alpha_i, \beta_i$  是互不相关、各自统计独立的,  $E[\alpha_i \beta_i] = 0$ , 且当  $i \neq j$  时,  $E[\alpha_i \alpha_j] = E[\beta_i \beta_j] = 0$ , 因而

$$\begin{aligned}\sigma_{\epsilon}^2 &= \left(\sum_{i=0}^N E[\beta_i^2]\right) \frac{1}{2\pi j} \oint_c \frac{1}{A(z) A(z^{-1})} \frac{dz}{z} \\ &\quad + \left(\sum_{i=1}^N E[\alpha_i^2]\right) \frac{1}{2\pi j} \oint_c \frac{H(z) H(z^{-1})}{A(z) A(z^{-1})} \frac{dz}{z}\end{aligned}\quad (8-85)$$

如果  $\mu, \gamma$  分别是分子系数  $b_i$  和分母系数  $a_i$  的非零非 1 的数目, 则有

$$\sum_{i=1}^N E[\alpha_i^2] = \gamma \frac{\Delta^2}{12}, \quad \sum_{i=0}^N E[\beta_i^2] = \mu \frac{\Delta^2}{12} \quad (8-86)$$

因此

$$\sigma_{\epsilon}^2 = \frac{\Delta^2}{12} \left\{ \frac{\mu}{2\pi j} \oint_c \frac{1}{A(z) A(z^{-1})} \frac{dz}{z} + \frac{\gamma}{2\pi j} \oint_c \frac{H(z) H(z^{-1})}{A(z) A(z^{-1})} \frac{dz}{z} \right\} \quad (8-87)$$

滤波器设计好后, 则可利用此式估算在一定系数字长  $b$  之下 ( $\Delta = 2^{-b}$ ), 频响偏差的均方值  $\sigma_{\epsilon}^2$  或者估算在一定的频响偏差的均方值  $\sigma_{\epsilon}^2$  下, 系数所需字长。

应该注意到, 对一个具体滤波器来说, 系数量化误差是固定值而不是随机变量, 其频响的均方偏差  $\epsilon^2$  也是固定值。上面把它们都看成随机变量是为了对  $\epsilon^2$  的大小作一个概率估计, 即  $\sigma_{\epsilon}^2$  是  $\epsilon^2$  最有可能出现的估值, 当滤波器阶数越高, 系数越多时, 这种估计的收敛性就越好。

## 2. FIR 数字滤波器系数量化的统计分析

由于有  $h(n)$  为奇对称、偶对称, 加之  $N$  为奇、偶之分, 故共有 4 种线性相位 FIR 数字滤波器情况, 但它们的频率响应表达式都是相似的。我们仅以  $N$  为奇数、 $h(n) = h(N-1-n)$  偶对称的 FIR 线性相位数字滤波器的频响为例, 来进行讨论。

$$\begin{aligned}H(e^{j\omega}) &= \left\{ \sum_{n=0}^{\frac{N-1}{2}} 2h(n) \cos\left[\left(\frac{N-1}{2} - n\right)\omega\right] + h\left(\frac{N-1}{2}\right) \right\} e^{-j\left(\frac{N-1}{2}\right)\omega} \\ &= H(\omega) e^{-j\left(\frac{N-1}{2}\right)\omega}\end{aligned}$$

其中

$$H(\omega) = \sum_{n=0}^{\frac{N-1}{2}} 2h(n) \cos\left[\left(\frac{N-1}{2} - n\right)\omega\right] + h\left(\frac{N-1}{2}\right) \quad (8-88)$$

是我们要研究的, 而由于  $e^{-j\omega(N-1)/2}$  相角部分不受系数量化的影响, 故不必去考虑它。

设  $h(n)$  序列以步阶为  $\Delta = 2^{-b}$  舍入成  $\hat{h}(n)$  序列, 则有

$$\hat{h}(n) = h(n) + e(n), \quad 0 \leq n \leq \frac{N-1}{2} \quad (8-89)$$

且舍入后的  $\hat{h}(n)$  也一定满足偶对称关系

$$\hat{h}(n) = \hat{h}(N-1-n), \quad 0 \leq n \leq \frac{N-1}{2} \quad (8-90)$$

而  $e(n)$  是随机变量, 假定它在  $(-\frac{\Delta}{2}, +\frac{\Delta}{2})$  间隔内是均匀分布的, 令

$$\hat{H}(z) = \mathcal{Z}[\hat{h}(n)], \quad \hat{H}(e^{j\omega}) = \hat{H}(\omega)e^{-j\omega(N-1)/2}$$

则有

$$\hat{H}(\omega) = \sum_{n=0}^{\frac{N-3}{2}} 2[h(n) + e(n)] \cos\left[\left(\frac{N-1}{2} - n\right)\omega\right] + \left[h\left(\frac{N-1}{2}\right) + e\left(\frac{N-1}{2}\right)\right] \quad (8-91)$$

定义频响误差函数为

$$E(\omega) = \hat{H}(\omega) - H(\omega) = \sum_{n=0}^{\frac{N-3}{2}} 2e(n) \cos\left[\left(\frac{N-1}{2} - n\right)\omega\right] + e\left(\frac{N-1}{2}\right) \quad (8-92)$$

很明显  $E(\omega)$  是一个线性相位滤波器的幅度响应(有正负), 此线性相位滤波器的单位冲激响应  $e(n)$  是偶对称的, 即



$$e(n) = e(N-1-n), \quad 0 \leq n \leq \frac{N-1}{2}$$

所以, 系数量化后的滤波器可以表示成无限精度的滤波器与一个频率响应为  $E(\omega)e^{-j(N-1)\omega/2}$  的滤波器的并联(可见(8-92)式)。

由于  $|e(n)| \leq \frac{\Delta}{2}$ , 故从(8-92)式可导出  $E(\omega)$  的上限为

$$\begin{aligned} |E(\omega)| &\leq \sum_{n=0}^{\frac{N-3}{2}} 2|e(n)| \cdot \left|\cos\left[\left(\frac{N-1}{2} - n\right)\omega\right]\right| + \left|e\left(\frac{N-1}{2}\right)\right| \\ &\leq \frac{\Delta}{2} \left[1 + 2 \sum_{n=1}^{\frac{N-1}{2}} |\cos(\omega n)|\right] \end{aligned}$$

即

$$|E(\omega)| \leq \frac{N\Delta}{2} \quad (8-93)$$

这个误差上限没有考虑到各系数量化误差的统计特性, 因而过于偏大, 故用处不大。

### 3. FIR 数字滤波器系数量化的统计界限

利用前面的假定, 即假定系数量化产生的舍入误差是统计独立的, 且在  $(-\frac{\Delta}{2}, +\frac{\Delta}{2})$  区间内是均匀分布的, 其均值为零, 方差  $E[e^2(n)] = \frac{\Delta^2}{12}$ , 则可得频响误差  $E(\omega)$ (见(8-92)式)的均方差为

$$\begin{aligned}\sigma^2(\omega) &= E[E^2(\omega)] = \sum_{n=0}^{\frac{N-3}{2}} 4E[e^2(n)]\cos^2\left[\left(\frac{N-1}{2}-n\right)\omega\right] + E\left[e^2\left(\frac{N-1}{2}\right)\right] \\ &= \frac{\Delta^2}{12}\left[1 + 4\sum_{n=1}^{\frac{N-1}{2}}\cos^2(\omega n)\right]\end{aligned}$$

若设

$$W_N(\omega) = \left\{\frac{1}{2N-1}\left[1 + 4\sum_{n=1}^{\frac{N-1}{2}}\cos^2(\omega n)\right]\right\}^{\frac{1}{2}}$$

则可得频响误差的标准偏差为

$$\sigma(\omega) = \sqrt{E[E^2(\omega)]} = \sqrt{\frac{2N-1}{3}} \cdot \frac{\Delta}{2} W_N(\omega) \quad (8-94)$$

很明显,对任意  $N$ ,皆有

$$0 < W_N(\omega) \leqslant 1, \quad W_N(0) = W_N(\pi) = 1$$

因此可得

$$\sigma(\omega) \leqslant \frac{\Delta}{2} \sqrt{\frac{2N-1}{3}} \quad (8-95)$$

又因

$$\lim_{N \rightarrow \infty} W_N(\omega) = \frac{1}{\sqrt{2}}, \quad 0 < \omega < \pi \quad (8-96)$$

故(8-95)式是频响误差的标准偏差的估计的上限,当  $N$  较大时,还可降低  $1/\sqrt{2}$ 。(8-95)式还表明,当预先不知道 FIR 滤波器系数数值而只给定所需频响误差的标准偏差  $\sigma(\omega)$  及冲激响应长度  $N$  时,可给设计者提供一个预计所需系数字长  $b$  的办法( $\Delta = 2^{-b}$ ),也就是使设计者知道系数需要的精度。

由于从(8-92)式看出误差函数  $E(\omega)$  是独立的随机变量之和,且这些随机变量的概率密度函数只在有限区间  $(-\frac{\Delta}{2}, +\frac{\Delta}{2})$  之内不为零,故只要相加的项数(即  $N$ )充分大,则  $E(\omega)$  趋近于正态分布,因而  $E(\omega)$  的平均值和方差就构成对  $E(\omega)$  的极好的统计描述。

## 8.5 数字滤波器运算中的有限字长效应

实现数字滤波器所包含的基本运算有延时、乘系数和相加三种。因为延时并不造成字长的变化,所以只需讨论乘系数和相加运算造成的影响。在定点制运算中,相乘的结果尾数位数会增加。例如两个  $b$  位尾数的数相乘后尾数是  $2b$  位,必须被舍入或截尾成  $b$  位尾数;相加的结果,尾数字长不变,不必舍入或截尾,但相加的结果可能超出有限寄存器长度,产生溢出,故有动态范围问题。浮点制运算中,相加及相乘都可能使尾数位数增加,故都会有舍入或截尾,但动态范围则不成问题。

分析数字滤波器运算误差的目的,是为了选择滤波器运算位数(即寄存器长度),以便满足信号噪声比值的技术要求。

前面已分析到,舍入或截尾的处理是非线性过程,分析起来非常麻烦,精确计算不仅不大可能,也没有必要,因而可以采用前面提出的统计方法,得到舍入或截尾的平均效果即可。下面以 IIR 滤波器为例,讨论运算中的有限字长效应。

### 1. IIR 滤波器定点运算舍入误差的统计分析

在定点制中,每次相乘运算  $y(n) = ax(n)$ (如图 8-16(a)所示)之后都要作一次舍入或截尾处理,因此会引入非线性,前面说过,一般多采用舍入处理,如图 8-16(b)所示。采用统计分析方法,可以将舍入误差

$$e(n) = Q_R[ax(n)] - ax(n) = Q_R[y(n)] - y(n) \quad (8-97)$$

作为独立噪声叠加到信号上(和前面一样,此处  $Q_R[\cdot]$  表示舍入处理)。这样仍可用线性流图来表示,如图 8-16(c)所示。

采用图 8-16(c)的统计模型,在分析数字滤波器由于乘法舍入的影响时,需对实现滤波器所出现的各种噪声源作以下假定:

- (1) 所有误差  $e(n)$  是平稳的白噪声序列(均值为零);
- (2) 每个误差在它的量化范围内都是均匀分布的;
- (3) 任何两个不同乘法器形成的噪声源互不相关;

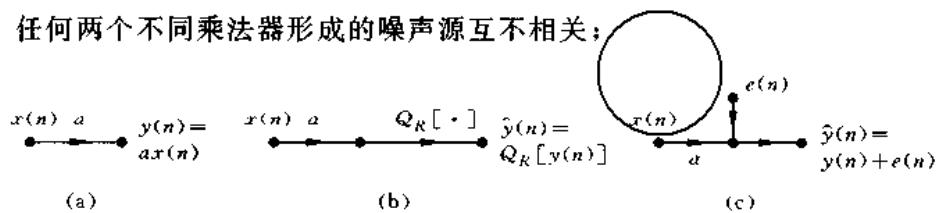


图 8-16 定点制相乘运算的模型

(a) 理想相乘;(b) 实际相乘舍入的非线性流图;(c) 实际相乘的定点舍入噪声统计模型的线性流图

- (4) 误差  $e(n)$  与输入  $x(n)$  及中间计算结果不相关,从而和输出序列  $y(n)$  也不相关。

当信号波形越复杂,量化步距越小时,这些假定越接近实际。根据这些假定,可认为舍入噪声是在  $\left(-\frac{2^{-b}}{2}, +\frac{2^{-b}}{2}\right]$  范围内均匀分布的,因而均值为  $m_e = E[e(n)] = 0$ ,方差为  $\sigma_e^2 = E[e^2(n)] = \frac{\Delta^2}{12}$ ,  $\Delta = 2^{-b}$ 。

这样,我们可以按照统计模型,即按线性系统的原则来求各噪声  $e(n)$  所产生的总输出噪声  $e_f(n)$ 。设  $y(n)$  是理想的没作尾数处理的输出,则经定点舍入处理后的实际输出为

$$\hat{y}(n) = y(n) + e_f(n) \quad (8-98)$$

而每一个噪声源  $e(n)$  所造成的输出噪声,可以利用白噪声通过线性系统的(8-54)式及(8-56)式分别求  $e(n)$  所造成的输出噪声的方差及均值,重写如下:

$$\sigma_f^2 = \frac{\sigma_e^2}{2\pi j} \oint_c H_e(z) H_e(z^{-1}) \frac{dz}{z} = \sigma_e^2 \sum_{n=-\infty}^{\infty} h_e^2(n) \quad (8-99)$$

$$m_f = m_e \sum_{n=-\infty}^{\infty} h_e(n) \quad (8-100)$$

其中  $h_e(n)$  是从  $e(n)$  加入的节点到输出节点间的系统的单位抽样响应,  $H_e(z)$  是  $h_e(n)$  的  $z$  变换。由于可以作线性系统处理,因此最后将所有的输出噪声线性叠加就得到总的输

出噪声  $e_f(n)$ 。按照上面 4 项假定，则总的输出噪声的方差也等于每个输出噪声方差之和。

下面以 IIR 滤波器为例来讨论滤波器运算中的有限字长效应。

【例 8-2】有一个 IIR 滤波器的系统函数为

$$H(z) = \frac{0.2}{(1 - 0.7z^{-1})(1 - 0.6z^{-1})}$$

用定点制算法，尾数舍入，分别计算直接型、级联型、并联型三种结构的舍入误差。

(1) 直接型结构

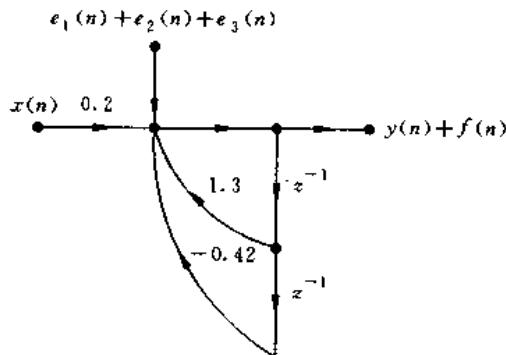


图 8-17 例 8-2 的直接型结构舍入误差统计模型

$$H(z) = \frac{0.2}{(1 - 0.7z^{-1})(1 - 0.6z^{-1})} = \frac{0.2}{1 - 1.3z^{-1} + 0.42z^{-2}} = \frac{0.2}{A(z)}$$

其中  $A(z) = 1 - 1.3z^{-1} + 0.42z^{-2} = (1 - 0.7z^{-1})(1 - 0.6z^{-1})$

图 8-17 画出了直接型结构定点相乘舍入后的统计模型，三个系数相乘，有三个舍入噪声  $e_0(n) + e_1(n) + e_2(n)$ ，它只通过  $H_0(z) = \frac{1}{A(z)}$  网络（而不是  $H(z) = \frac{0.2}{A(z)}$  网络）。

所以

$$f(n) = [e_0(n) + e_1(n) + e_2(n)] \times h_0(n)$$

而

$$h_0(n) = \mathcal{Z}^{-1}[H_0(z)]$$

输出噪声方差为

$$\sigma_f^2 = 3\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_C \frac{1}{A(z)A(z^{-1})} \frac{dz}{z}$$

先计算

$$\frac{1}{2\pi j} \oint_C \frac{1}{A(z)A(z^{-1})} \frac{dz}{z} = \frac{1}{2\pi j} \oint_C \frac{1}{(1 - 0.7z^{-1})(1 - 0.6z^{-1})(1 - 0.7z)(1 - 0.6z)} \frac{dz}{z}$$

围线  $C$  为单位圆，围线内只有两个极点  $z=0.6, z=0.7$ ，求此两个极点的留数，即

$$\begin{aligned} \frac{1}{2\pi j} \oint_C \frac{1}{A(z)A(z^{-1})} \frac{dz}{z} &= \frac{1}{2\pi j} \oint_C \frac{z}{(z-0.7)(z-0.6)(1-0.7z)(1-0.6z)} dz \\ &= \sum_k (\text{被积函数在单位圆内极点 } z_k \text{ 上的留数}) \\ &= \frac{0.7}{0.1 \times (1-0.49)(1-0.42)} + \frac{0.6}{(-0.1) \times (1-0.42)(1-0.36)} \end{aligned}$$

$$= 23.6646 - 16.1638 = 7.5008$$

所以

$$\sigma_f^2 = 3\sigma_e^2 \times 7.5008 = 3 \times \frac{\Delta^2}{12} \times 7.5008 = 1.8752\Delta^2 \quad (\Delta = 2^{-b})$$

### (2) 级联型结构

$$\begin{aligned} H(z) &= 0.2 \times \frac{1}{(1-0.7z^{-1})(1-0.6z^{-1})} = 0.2 \times \frac{1}{A_1(z)} \cdot \frac{1}{A_2(z)} = 0.2 \times \frac{1}{A(z)} \\ &= 0.2 \times H_0(z) \end{aligned}$$

$$\text{其中 } A_1(z) = 1 - 0.7z^{-1}, \quad A_2(z) = 1 - 0.6z^{-1}$$

图 8-18 画出了级联型结构定点相乘舍入后的统计模型, 每一次相乘相应节点上引入一个舍入噪声。要注意, 噪声经过的网络是不同的。

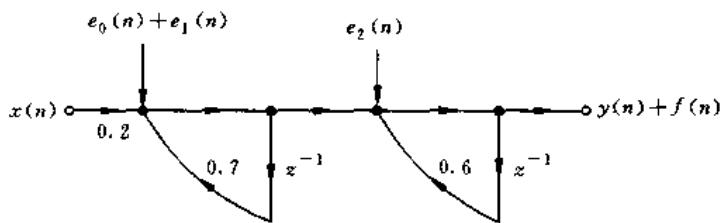


图 8-18 例 8-2 的级联结构舍入误差统计模型

$$f(n) = [e_0(n) + e_1(n)] \times h_0(n) + e_2(n) \times h_2(n)$$

$$h_0(n) = \mathcal{Z}^{-1}\left[\frac{1}{A(z)}\right], \quad h_2(n) = \mathcal{Z}^{-1}\left[\frac{1}{A_2(z)}\right]$$

所以

$$\begin{aligned} \sigma_f^2 &= 2\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{1}{A_1(z)A_2(z)A_1(z^{-1})A_2(z^{-1})} \frac{dz}{z} \\ &\quad + \sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{1}{A_2(z)A_2(z^{-1})} \frac{dz}{z} \end{aligned}$$

所以

$$\begin{aligned} \sigma_f^2 &= 2\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{dz}{(1-0.7z^{-1})(1-0.6z^{-1})(1-0.7z)(1-0.6z)z} \\ &\quad + \sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{dz}{(1-0.6z^{-1})(1-0.6z)z} \\ &= 2\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{z dz}{(z-0.7)(z-0.6)(1-0.7z)(1-0.6z)} \\ &\quad + \sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{dz}{(z-0.6)(1-0.6z)} \\ &= 2\sigma_e^2 \left[ \frac{0.7}{0.1 \times (1-0.49)(1-0.42)} + \frac{0.6}{(-0.1)(1-0.42)(1-0.36)} \right] + \sigma_e^2 \frac{1}{1-0.36} \\ &= 16.5641\sigma_e^2 = 16.5641 \times \frac{\Delta^2}{12} = 1.3803\Delta^2 \end{aligned}$$

### (3) 并联型结构

$$H(z) = \frac{1.4}{1 - 0.7z^{-1}} + \frac{-1.2}{1 - 0.6z^{-1}}$$

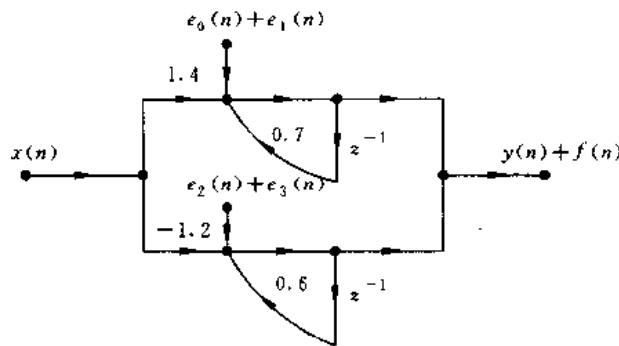


图 8-19 例 8-2 的并联结构的舍入误差统计模型

图 8-19 画出了并联型结构定点相乘舍入后的统计模型，有 4 个相乘系数，故有 4 个相乘舍入噪声。由图 8-19 看出

$$e_0(n) + e_1(n) \text{ 只通过 } \frac{1}{A_1(z)} = \frac{1}{1 - 0.7z^{-1}}$$

$$e_2(n) + e_3(n) \text{ 只通过 } \frac{1}{A_2(z)} = \frac{1}{1 - 0.6z^{-1}}$$

因此输出  $f(n)$  的方差为

$$\begin{aligned} \sigma_f^2 &= 2\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{1}{A_1(z)A_1(z^{-1})} \frac{dz}{z} + 2\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{1}{A_2(z)A_2(z^{-1})} \frac{dz}{z} \\ &= 2\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{dz}{(1 - 0.7z^{-1})(1 - 0.7z)} + 2\sigma_e^2 \cdot \frac{1}{2\pi j} \oint_c \frac{dz}{(1 - 0.6z^{-1})(1 - 0.6z)} \\ &= 2\sigma_e^2 \cdot \left[ \frac{1}{2\pi j} \oint_c \frac{dz}{(z - 0.7)(1 - 0.7z)} + \frac{1}{2\pi j} \oint_c \frac{dz}{(z - 0.6)(1 - 0.6z)} \right] \\ &= 2\sigma_e^2 \left[ \frac{1}{1 - 0.49} + \frac{1}{1 - 0.36} \right] \\ &= 7.0466\sigma_e^2 = 7.0466 \times \frac{\Delta^2}{12} = 0.5872\Delta^2 \end{aligned}$$

由此看出

$$\sigma_{f \text{ 直接}}^2 > \sigma_{f \text{ 级联}}^2 > \sigma_{f \text{ 并联}}^2$$

直接型结构的所有舍入误差都要经过全部网络的反馈环节，误差积累起来了，所以误差最大。级联型结构的每个舍入误差只通过其后面的反馈环节（不通过前面的），故舍入误差比直接型的小（在某些排序情况下，其误差性能可接近甚至超过并联型结构）。并联型结构的每个并联回路的舍入误差只通过本回路，与其他回路无关，误差累积作用更小，故在一般情况下，其输出误差最小。

## 2. IIR 滤波器定点运算的动态范围

定点制的加法运算不会出现尾数字长增加，故不会有舍入误差问题，但它却可能出现溢出。上面讨论的舍入噪声只有在系统不溢出时才是输出误差的主要来源，当溢出时，溢出会产生很大的误差。为了防止溢出，需要在网络内加入适当的压缩比例因子。要保证每一个加法器的输出都不溢出，惟而广之就是要滤波器每一个节点上都不产生溢出，也就

是使每一节点信号的值都小于 1。

令  $y_k(n)$  表示第  $k$  个节点上的输出, 对定点制要满足

$$|y_k(n)| < 1 \quad (8-101)$$

用  $h_k(n)$  表示从滤波器输入到第  $k$  个节点的单位抽样响应, 则有

$$y_k(n) = \sum_{m=-\infty}^{\infty} h_k(m)x(n-m) = h_k(n) * x(n)$$

$y_k(n)$  的动态范围取决于输入信号类型及  $h_k(n)$ , 对于不同类信号,  $y_k(n)$  的动态范围可见表 8-3。

下面来分析使任一节点  $k$  不发生溢出的条件。由于

$$|y_k(n)| = \sum_{m=-\infty}^{\infty} |h_k(m)x(n-m)| \leq x_{\max} \sum_{m=-\infty}^{\infty} |h_k(m)|$$

其中  $x_{\max}$  表示输入序列最大绝对值。要求

$$|y_k(n)| < 1 \quad k = 1, 2, \dots$$

即

$$x_{\max} \sum_{m=-\infty}^{\infty} |h_k(m)| < 1 \quad (8-102)$$

也就是要求

$$x_{\max} < \frac{1}{\sum_{m=-\infty}^{\infty} |h_k(m)|} \quad (8-103)$$

这就是保证第  $k$  个节点上不出现溢出时, 最大输入值上限。

表 8-3 输入  $x(n)$  为不同类信号时  $y_k(n) = x(n) * h_k(n)$  的动态范围

输入	$y_k(n)$ 的动态范围
1. 正弦 $x(n) = \cos(\omega n)$	$ y_k(n)  \leq \max  H_k(e^{j\omega}) $
2. 有界信号 $ x(n)  \leq 1$	$ y_k(n)  \leq \sum_{n=0}^{\infty}  h_k(n) $
3. 有限能量信号 $\sum_{m \leq n} x^2(m) \leq 1$	$ y_k(n)  \leq \left[ \sum_{n=0}^{\infty} h_k^2(n) \right]^{1/2}$
4. 广义平稳信号	$\{E[y_k^2(n)]\}^{1/2} = \left[ \frac{1}{2\pi} \int_{-\pi}^{\pi} s_{xx}(\omega)  H_k(e^{j\omega}) ^2 d\omega \right]^{1/2}$
5. 白色信号 $s_{xx}(\omega) = 1$	$\{E[y_k^2(n)]\}^{1/2} = \left[ \sum_{n=0}^{\infty}  h_k(n) ^2 \right]^{1/2}$

注:  $S_{xx}(\omega)$  为广义平稳随机信号的功率谱密度,  $H_k(e^{j\omega}) = DTFT[h_k(n)]$  在表中前三种情况下给出  $y_k(n)$  的真正上限, 后两种情况则只给出随机变量的标准偏差, 而不是真正的上限, 因而仍可能存在溢出。

但是  $X_{\max}$  不一定满足(8-103)式的要求, 为此可将输入信号乘上一个适当的压缩比例因子  $A$  ( $A < 1$ ), 来衰减输入信号的幅度, 以使得任意第  $k$  个节点不发生溢出, 即

$$y_k(n) = Ax_{\max} \sum_{m=-\infty}^{\infty} |h_k(m)| < 1 \quad (8-104)$$

则

$$A < \frac{1}{x_{\max} \sum_{m=-\infty}^{\infty} |h_k(m)|} \quad (8-105)$$

### 3. IIR 滤波器的定点运算的零输入极限环振荡和溢出振荡

数字滤波器中由于运算过程中的尾数处理产生量化的非线性作用,因而使系统中引入了非线性环节,而 IIR 滤波器又存在着反馈,因而在一定的条件下,也可以引起振荡。一个稳定的数字滤波器,在某时刻令输入为零,则从此时刻开始输出应逐渐衰减到零。但是同一滤波器,若用有限寄存器长度来实现,则在输入为零的时刻开始,输出并不逐渐衰减为零,而有可能衰减到某一非零的幅度范围,产生振荡特性(包括  $\omega=0$  的等幅序列),这种效应称为零输入极限环振荡。

数字滤波器的极限环特性很复杂,很难分析,下面仅就一阶 IIR 数字滤波器的简单情况加以讨论。

#### (1) 定点运算有限寄存器长度引起的零输入极限环振荡

设一阶 IIR 网络的系统函数为

$$H(z) = \frac{1}{1 - az^{-1}} \quad (8-106)$$

在无限精度运算下,其差分方程为

$$y(n) = ay(n-1) + x(n) \quad (8-107)$$

在定点制中,每次乘法运算后都要对尾数进行舍入处理,舍入处理是非线性的,这时的非线性方程可表示为

$$\hat{y}(n) = Q_R[\hat{a}\hat{y}(n-1)] + x(n) \quad (8-108)$$

$Q_R[\cdot]$  表示舍入量化处理。此式可用图 8-20 的非线性流图来表示。

下面我们来看一个实际例子。设此一阶系统中尾数字长为  $b=3$  位(不包括符号位),系数  $a=0.5$  用二进制表示为  $a=0.100$ ,系统的极点  $z=a=0.5<1$ ,在单位圆内,故系统稳定,系统的单位冲激响应为  $h(n)=a^n u(n)$ 。如果输出序列由 0.75 降到零,即(以二进制数表示)

$$x(n) = \begin{cases} 0.110, & n=0 \\ 0.000, & n>0 \end{cases} \quad \text{即 } x(n) = 0.75\delta(n)$$

在无限精度运算下,输出  $y(n)$  也将逐渐衰减到零,即

$$y(n) = x(n) * h(n) = 0.75a^n u(n) = 0.75 \times (0.5)^n u(n) \xrightarrow{n \rightarrow \infty} 0$$

如果对运算结果进行尾数处理,则当  $n \rightarrow \infty$  时,输出就不趋于零了。将以上的(8-108)式的非线性方程的一步一步运算结果列表表示在表 8-4 中。

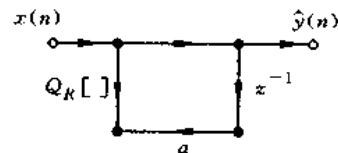


图 8-20 一阶 IIR 网络的  
非线性流图

表 8-4  $a=0.100$ (二进制)的一阶网络运算过程

n	x(n)	$\hat{y}(n-1)$	$a\hat{y}(n-1)$	$Q_R[a\hat{y}(n-1)]$	$\hat{y}(n)$	
					二进制	十进制
0	0.110	0.000	0.000	0.000	0.110	0.75
1	0.000	0.110	0.0110	0.011	0.011	0.375
2	0.000	0.011	0.0011	0.010	0.010	0.25
3	0.000	0.010	0.0010	0.001	0.001	0.125
4	0.000	0.001	0.0001	0.001	0.001	0.125
⋮	⋮	⋮	⋮	⋮	⋮	⋮

可以看出,最后输出停留在  $\hat{y}(n)=0.001$ (二进制)上,不会衰减到零。进入  $\hat{y}(n)=0.001$ (二进制)后被称为“死带”区域,其结果如图 8-21(a)所示。如果  $a$  为负数,则每乘一次  $a$ ,输出改变一次符号,则输出是正负相间的不衰减振荡。例如,当  $a=-0.5$ (十进制),则每乘一次  $a$  就改变一次符号,就得到如图 8-21(b)所示的  $\hat{y}(n)$ 。这两种现象就是“零输入极限环振荡”。

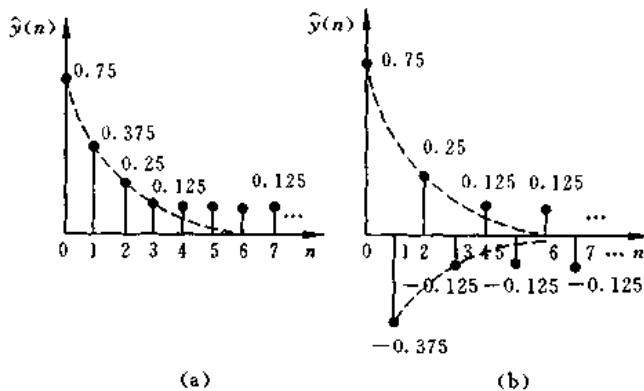


图 8-21 一阶 IIR 网络的零输入极限环振荡

(a)  $a=0.5$ ; (b)  $a=-0.5$

这种现象产生的原因是什么呢? 观察表 8-4 的最后一行可以看出,当  $\hat{y}(n-1)=0.001$  时,  $a\hat{y}(n-1)=0.00010$ , 数值被衰减了,但经过舍入处理后,  $Q_R[a\hat{y}(n-1)]=0.001$ , 又变成原来  $\hat{y}(n-1)$  的数值,因而输出就保持不变,也就是说,只要满足

$$|Q_R[a\hat{y}(n-1)]|=|\hat{y}(n-1)| \quad (8-109)$$

时,舍入处理就使系数  $a$  失效,也就是等效于使  $a$  换成绝对值为 1 的系数  $a'$ ,  $a'=\frac{a}{|a|}$ ,  $|a'|=1$ ,这时一阶滤波器的极点变成  $a'=\pm 1$ ,代入(8-106)式,可得等效系统函数为

$$H'(z)=\frac{1}{1-a'z^{-1}}=\frac{1}{1\pm z^{-1}} \quad (8-110)$$

如图 8-22 所示,当  $a=a_1=0.5$  时,等效  $H'(z)$  的极点在  $z=1$  处,当  $a=a_2=-0.5$  时,等效  $H'(z)$  的极点在  $z=-1$  处,也就是极点都迁移到单位圆上,因而系统就是临界稳定的,出现等幅振荡。 $a=0.5$  时的极点迁移到零频位置  $z=a'=1$  处,故所产生的死带称为零频极限环振荡。

下面讨论极限环振荡幅度与字长  $b$  的关系。

由于舍入误差的范围为  $\Delta/2$  以内,故有

$$|Q_R[a\hat{y}(n-1)]-a\hat{y}(n-1)|\leq \frac{\Delta}{2} \quad (8-111)$$

或

$$|Q_R[a\hat{y}(n-1)]|-|a||\hat{y}(n-1)|\leq \frac{\Delta}{2}$$

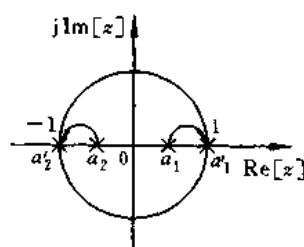


图 8-22 舍入后极点的等效迁移

将极限环振荡时的(8-109)式代入此式,可得

$$|\hat{y}(n-1)| - |a| |\hat{y}(n-1)| \leq \frac{\Delta}{2}$$

则可得

$$|\hat{y}(n-1)| \leq \frac{\Delta/2}{1-|a|} \quad (8-112)$$

此式定义了一阶 IIR 网络的死带范围,表明极限环幅度与量化宽度成正比,因此增加字长(减小量化宽度)将使极限环振荡减弱。例如,设  $b=3, \Delta=2^{-b}=1/8, |a|=0.5$  时,有

$$|\hat{y}(n)| = |\hat{y}(n-1)| = \frac{1/16}{1-0.5} = \frac{1}{8} = 0.125$$

这与表 8-4 的结果一致。

用同样的方法可分析二阶系统的零输入极限环振荡现象。

当采用高阶 IIR 滤波器时,极限环振荡的分析更为复杂。如果采用并联形式来实现高阶系统,由于每个并联节的输出是独立的,故可分别直接应用上面的分析。用级联形式实现高阶系统,只有第一节输入为零,后续节可呈现出自己的极限环振荡且过滤前节的极限环振荡输出。

在实际问题中,要尽量克服极限环振荡现象,例如在通信中,空载线路就不允许有振荡存在,但是利用极限环振荡现象,可以设计周期性信号发生器。

## (2) 定点运算的溢出振荡

在上面的讨论中,都假定已经选择合适的比例因子,使相加结果不出现溢出。如果在 IIR 滤波器的定点加法运算中存在溢出,则在一定条件下会引起振荡,称为溢出振荡。我们以定点补码运算的二阶 IIR 滤波器为例进行讨论,滤波器结构如图 8-23 所示。

此二阶滤波器的差分方程为

$$y(n) = a_1 y(n-1) + a_2 y(n-2) + x(n) \quad (8-113)$$

其系统函数为

$$\begin{aligned} H(z) &= \frac{1}{1-a_1 z^{-1}-a_2 z^{-2}} = \frac{z^2}{z^2-a_1 z-a_2} = \frac{z^2}{(z-\alpha_1)(z-\alpha_2)} \\ &= \frac{z^2}{z^2-(\alpha_1+\alpha_2)z+\alpha_1\alpha_2} \end{aligned} \quad (8-114)$$

$$\alpha_{1,2} = \frac{a_1 \pm \sqrt{a_1^2 + 4a_2}}{2} \quad (8-115)$$

首先讨论稳定问题。稳定条件是  $H(z)$  的极点  $\alpha_1, \alpha_2$  在单位圆内,即  $|\alpha_1| < 1, |\alpha_2| < 1$ , 又由(8-114)式知  $\alpha_1 + \alpha_2 = a_1, \alpha_1 \alpha_2 = -a_2$ , 由于  $|\alpha_1 \alpha_2| < 1$ , 故  $|a_2| < 1$ 。下面分两种情况讨论:

① 设  $\alpha_1, \alpha_2$  为实根,则  $(1-\alpha_1^2)(1-\alpha_2^2) > 0$

即  $1 + \alpha_1^2 \alpha_2^2 > \alpha_1^2 + \alpha_2^2$

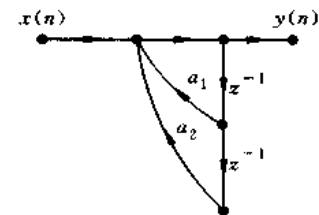


图 8-23 二阶 IIR 滤波器

因而

$$(1 + \alpha_1 \alpha_2)^2 > (\alpha_1 + \alpha_2)^2$$

由于  $\alpha_1, \alpha_2$  为实数, 将  $\alpha_1, \alpha_2$  与  $a_1, a_2$  的关系代入, 则

$$(1 - a_2)^2 > a_1^2, \text{ 即 } 1 - a_2 > |a_1|, \text{ 即 } |a_1| + a_2 < 1 \quad (8-116)$$

此外, 若  $\alpha_1, \alpha_2$  为实根, 则由(8-115)式必有

$$a_1^2 + 4a_2 \geq 0 \quad (8-117)$$

可得以下两种情况:

$a_2 > 0$  时, (8-117)式一定成立;

$a_2 < 0$  时, 要求

$$a_2 \geq -\frac{a_1^2}{4} \quad (8-118)$$

才能使(8-117)式成立。

将(8-116)式及(8-118)式综合, 可得  $\alpha_1, \alpha_2$  为实根时系统稳定的条件为

$$|a_1| + a_2 < 1, \quad a_2 \geq -\frac{a_1^2}{4} \quad (8-119)$$

这就是图 8-24 中横线阴影区的内部及  $a_2 = -\frac{a_1^2}{4}$  的边界上。

② 设  $\alpha_1, \alpha_2$  为共轭复根  $z_{1,2} = re^{\pm j\theta}, r < 1$ , 则有

$$|z_1 z_2| = r^2 = -a_2 < 1, \text{ 即 } a_2 > -1 \quad (8-120)$$

此外, 若  $\alpha_1, \alpha_2$  为共轭复根, 则由(8-115)式必有

$$a_1^2 + 4a_2 < 0 \quad (8-121)$$

将(8-120)式和(8-121)式综合, 可得  $\alpha_1, \alpha_2$  为共轭复根时系统稳定的条件为

$$a_2 > -1, \quad a_2 < -\frac{a_1^2}{4} \quad (8-122)$$

这就是图 8-24 中竖线的阴影区内部(不包括边界)。

这样, 将①, ②两种情况归纳可知, 系统稳定必须在  $|a_1| + a_2 = 1$  (即  $a_2 + a_1 = 1$  直线和  $a_2 - a_1 = 1$  直线)以及  $a_2 = -1$  这三条直线方程所围的三角形之内, 包括横线阴影区及竖线阴影区, 如图 8-24 所示。

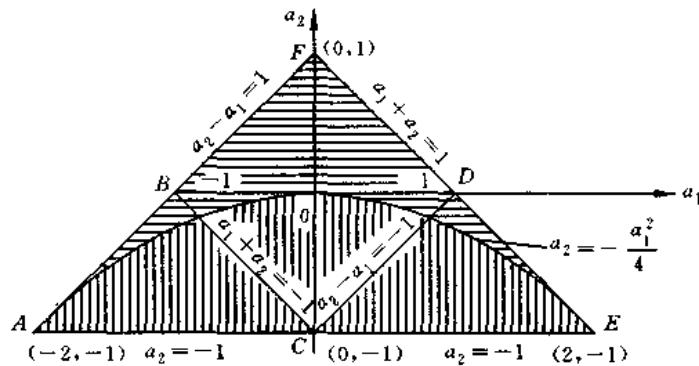


图 8-24 二阶 HR 系统的稳定性三角形

其次讨论不溢出问题。我们来研究,当输入  $x(n)=0$  后,  $a_1, a_2$  为何值时才能保证不溢出。仍利用补码加法器,并暂时忽略舍入误差。补码加法器的非线性特性见图 8-25 所示。其中  $v$  是该加法器输入之和,而  $f[v]$  是加法器的输出,也就是滤波器的输出(见图 8-23 所示),只有当各输入之和处于  $|v| < 1$  的数值范围之内时加法器才相当于真正的加法运算。当  $v$  的值在 1 与 2 之间时,相加后正的输出将被理解为负的分数输出,这时就出现严重的非线性。当输入  $x(n)=0$  时,则二阶滤波器的输出  $y(n)$  为

$$y(n) = a_1 y(n-1) + a_2 y(n-2)$$

经补码加法器后,其真实输出应满足图 8-25 的非线性变换关系

$$y(n) = f[a_1 y(n-1) + a_2 y(n-2)]$$

不出现溢出,要求  $|y(n)| < 1$  对所有  $n$  都成立,即要求  $|y(n)| < 1, |y(n-1)| < 1, |y(n-2)| < 1$ 。

由图 8-25 知,只要输入之和满足

$$|a_1 y(n-1) + a_2 y(n-2)| < 1$$

则  $y(n)$  会被正确理解,即  $|y(n)| < 1$ 。

假定  $|y(n-1)| < 1, |y(n-2)| < 1$ , 如果

$$|a_1| + |a_2| \leq 1 \quad (8-123)$$

则有

$$|a_1 y(n-1) + a_2 y(n-2)| \leq |a_1 y(n-1)| + |a_2 y(n-2)| < |a_1| + |a_2| \leq 1$$

所以  $|a_1| + |a_2| \leq 1$  是不溢出( $|y(n)| < 1$ )的充分条件。同样,用反证法可证明  $|a_1| + |a_2| \leq 1$  也是不溢出的必要条件。

(8-123)式就是图 8-24 中  $a_1 + a_2 = 1, a_2 - a_1 = 1, a_1 + a_2 = -1, a_2 - a_1 = -1$  四条直线所围的正方形。

最后讨论输出极限环振荡问题。分以下两种情况:

(1) 输入  $x(n)=0$  时,输出一个等幅序列  $y(n)=y_0 > 0$  的可能性,即

$$y_0 = f[a_1 y_0 + a_2 y_0] = 2K + a_1 y_0 + a_2 y_0, \quad K = 0, \pm 1, \pm 2, \dots$$

选为  $2K$ ,是考虑到图 8-25 的特性,输入相差偶数时,输出相同。一定能找到一个  $K$  值,使上式相等,即

$$y_0 = \frac{2K}{1 - (a_1 + a_2)}$$

注意,由于  $f[\cdot] < 1$ ,故要求  $|y_0| < 1$ 。

在没有溢出时,若  $K=0$ ,则  $y_0 = (a_1 + a_2)y_0$ ,故只要  $a_1 + a_2 = 1$  即可,但这一直线不在稳定区域内。

如果  $K=\pm 1$ ,则  $y_0 = \frac{\pm 2}{1 - (a_1 + a_2)}$ ,要求  $|y_0| = \frac{2}{1 - (a_1 + a_2)} < 1$ ,即必须

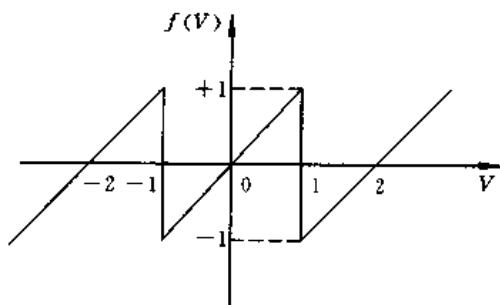


图 8-25 补码加法器的输入输出特性

$$a_1 + a_2 < -1 \quad (8-124)$$

从而得到输入为零时输出为等幅振荡序列,这正好是图 8-24 的左下角的三角形  $\triangle ABC$  内部区域。因为  $BC$  边  $a_1 + a_2 = -1$ , 故  $y_0 = 1$ ; 顶点  $A$  上为  $a_1 + a_2 = -3$ , 故  $y_0 = 1/2$ 。由此看出, 在  $\triangle ABC$  内有  $1/2 < y_0 < 1$ 。 $k \geq 2$  时,  $a_1, a_2$  不可能在稳定区域内。

(2) 输入  $x(n) = 0$  时, 输出是周期为 2 的零输入极限环振荡, 即输出为  $y(n) = (-1)^n y_0$  的情况, 其中  $0 < y_0 < 1$ , 由补码加法器的非线性特性可得

$$(-1)^n y_0 = f[a_1(-1)^{n-1} y_0 + a_2(-1)^{n-2} y_0] = f[(-1)^n (a_2 - a_1) y_0]$$

同样有

$$(-1)^n y_0 = (-1)^n (a_2 - a_1) y_0 + 2K, \quad K = 0, \pm 1, \pm 2, \dots$$

可解得

$$(-1)^n y_0 = \frac{2K}{1 - (a_2 - a_1)} = \frac{2K}{1 + (a_1 - a_2)}$$

如果  $K = 0$ , 则  $a_2 - a_1 = 1$ , 不在稳定区域内。如果  $K = \pm 1$ , 则上式变成  $(-1)^n y_0 = \frac{\pm 2}{1 + (a_1 - a_2)}$ , 要求  $|y_0| < 1$ , 则有

$$a_1 - a_2 > 1 \quad (8-125)$$

在此条件下, 才能满足  $x(n) = 0$  时, 输出是周期为 2 的等幅振荡, 这一条件限定的区域为图 8-24 的右下角的三角形  $\triangle CDE$  的内部区域。其一条边界为  $a_1 - a_2 = 1$ , 则  $y_0 = 1$ ; 而顶点  $E$  处,  $a_1 - a_2 = 3$ ,  $y_0 = 1/2$ 。于是在  $\triangle CDE$  内有  $1/2 < y_0 < 1$ 。 $k \geq 2$  时,  $a_1, a_2$  皆不在稳定区内。

综上所述, 若采用补码加法器, 则必须满足  $|a_1| + |a_2| < 1$  (即在图 8-24 中  $FBCD$  方形的内部(不包括边界)), 才能既稳定, 又不产生溢出。否则在没有输入时可能出现振荡。

为了避免溢出振荡, 必须修改加法器, 使非线性函数  $f[\cdot]$  由图 8-25 的形式变成图 8-26 的形式, 也就是变成饱和加法器, 当相加器的输入之和大于 1 或小于 -1 时, 就分别以 1 和 -1 代表相加结果, 这样就能克服溢出振荡。

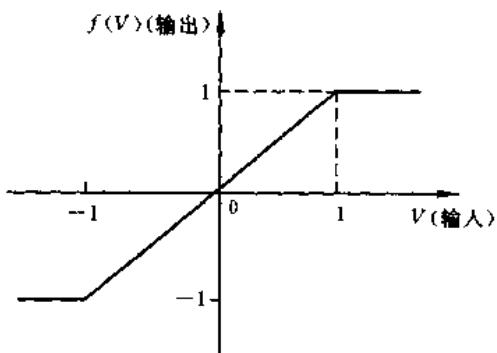


图 8-26 补码饱和加法器特性

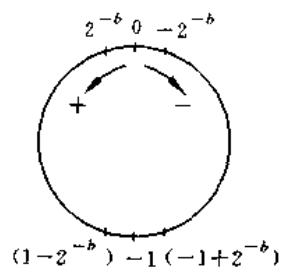


图 8-27 补码运算的表示

当然, 若用正常的不饱和加法器, 采用补码运算, 在两个或多个输入值相加时, 只要肯定最后结果的绝对值一定小于 1, 则相加的中间结果的溢出不会影响最后结果的正确性。这是因为, 例如由(8-5)式定义的补码二进制数表示法可用图 8-27 来表示, 溢出虽然产生

暂时的错误,但得到的数仍在  $b+1$  位字的动态范围之中。

[例] 补码加法  $\frac{3}{8} + \frac{3}{4} + \frac{1}{8} - \frac{1}{2} = \frac{3}{4}$

可表示为  $\frac{3}{8} = 0.011, \frac{3}{4} = 0.110, \frac{1}{8} = 0.001, -\frac{1}{2} = 1.100$

则

$$\frac{3}{8} + \frac{3}{4} = 0.011 + 0.110 = 1.001 = -\frac{7}{8} \quad (\text{溢出,产生错误})$$

继续相加

$$-\frac{7}{8} + \frac{1}{8} = 1.001 + 0.001 = 1.010 = -\frac{3}{4}$$

再继续相加

$$-\frac{3}{4} + \left(-\frac{1}{2}\right) = 1.010 + 1.100 = 10.110$$

这是最后的结果,把进位项(最高位)1丢掉,则为  $0.110 = 3/4$ ,这正是所需的正确答案。因此,暂时的溢出是允许的。

## 8.6 FFT 算法的有限字长效应

FFT 算法是计算 DFT 的快速算法。由于在数字滤波器和频谱分析中广泛采用 DFT,所以弄清楚 DFT 尤其是 FFT 中有限寄存器长度的影响是非常重要的。和数字滤波器一样,精确分析这种影响是困难的,一般为了选择寄存器长度,采用简化的分析方法就够了,即和前面一样,借助于用可加性噪声来进行分析。下面的讨论都以时间抽取(DIT)基-2 FFT 为例,分析运算有限字长效应及系数量化效应,并且是针对舍入情况进行分析的,对于其他 FFT 算法及截尾运算是不难加以推广的。我们首先讨论直接 DFT 计算中的有限字长误差,然后讨论 FFT 算法中的有限字长误差。

### 一、DFT 计算中有限字长效应的分析

DFT 的定义为

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^{nk}, \quad k = 0, 1, \dots, N-1 \quad (8-126)$$

其中  $W_N = e^{-j\frac{2\pi}{N}}$ ,当只需计算少数几个点处的  $X(k)$  值时,可直接把乘积相加即可,对指定的  $k$  值来说,可把(8-126)式看成卷积运算,与下式相当:

$$y(n) = \sum_{i=0}^{N-1} h(i)x(n-i)$$

即把  $W_N^{nk}$  看成线性移不变系统的单位抽样响应  $h(n)$ , $x(n)$  看作系统输入,而把  $X(k)$  看作系统的输出  $y(n)$ ,因而只要稍加说明就可完全利用 FIR 滤波器运算有限字长效应的分析结果,在定点实现时,每个乘积均需作尾数处理。令  $e(n, k)$  表示对乘积  $x(n)W_N^{nk}$  舍入引起的量化误差,因此等效的统计模型可见图 8-28。图中  $X(k)$  是理想的无限精度运算结

果,  $\hat{X}(k)$  表示有限字长运算的结果,  $F(k)$  表示第  $k$  个值计算结果的误差。由于各误差源直接加到输出端, 故总输出误差为

$$F(k) = \sum_{n=0}^{N-1} e(n, k) \quad (8-127)$$

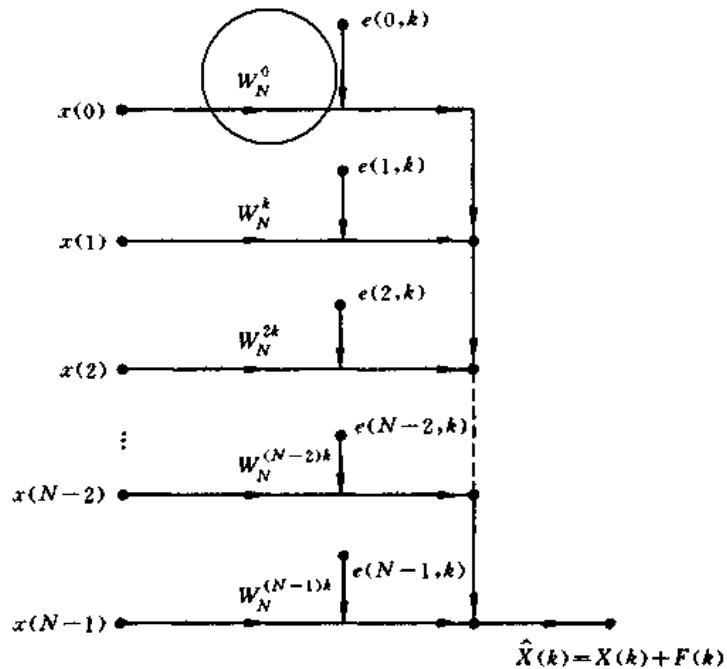


图 8-28 DFT 计算中定点舍入的统计模型

一般  $x(n)$  及  $W_N^k$  皆为复数, 因而  $W_N^k x(n)$  为复数, 它是由四个实数乘法来完成的, 而四个实乘引入四个实的舍入误差  $e_1(n, k), e_2(n, k), e_3(n, k)$  及  $e_4(n, k)$ 。如果不考虑系数  $W_N^k$  的量化误差, 则经舍入后, 可表示成

$$\begin{aligned} Q_R[x(n)W_N^k] &= \operatorname{Re}[x(n)]\cos\left(\frac{2\pi}{N}kn\right) + e_1(n, k) + \operatorname{Im}[x(n)]\sin\left(\frac{2\pi}{N}kn\right) + e_2(n, k) \\ &\quad + j\left[\operatorname{Im}[x(n)]\cos\left(\frac{2\pi}{N}kn\right) + e_3(n, k)\right] - j\left[\operatorname{Re}[x(n)]\sin\left(\frac{2\pi}{N}kn\right) + e_4(n, k)\right] \end{aligned} \quad (8-128)$$

为了计算  $\hat{X}(k)$  中误差的方差, 对舍入误差  $e_i(n, k)$  ( $i=1, 2, 3, 4$ ) 的统计特性作以下假设:

(1) 误差  $e_i(n, k)$  是白噪声, 在  $\left(-\frac{2^{-b}}{2}, \frac{2^{-b}}{2}\right]$  范围内是均匀分布的, 故其平均值为零,

方差为  $\frac{2^{-2b}}{12}$ ;

(2) 各误差  $e_i(n, k)$  彼此互不相关, 且某一复乘的四个误差源与其他复乘的误差源也互不相关;

(3) 所有误差  $e_i(n, k)$  与输入不相关, 因而与输出也不相关。

一个复乘舍入后误差均值为零, 误差的模的平方为

$$|e(n, k)|^2 = [e_1(n, k) + e_2(n, k)]^2 + [e_3(n, k) - e_4(n, k)]^2$$

由于各  $e_i(n, k)$  互不相关, 故  $|e(n, k)|^2$  的统计平均为

$$E[|e(n,k)|^2] = 4 \frac{2^{-2b}}{12} = \frac{1}{3} \cdot 2^{-2b} \quad (8-129)$$

同样,利用(8-127)式及(8-129)式可得输出误差的均方幅度(即为方差,因均值为零)

$$E[|F(k)|^2] = E\left[\left|\sum_{n=0}^{N-1} e(n,k)\right|^2\right] = \sum_{n=0}^{N-1} E[|e(n,k)|^2] = \frac{N}{3} 2^{-2b} \quad (8-130)$$

即和 FIR 滤波器的直接型实现一样,输出噪声的总方差正比于  $N$ 。但是按(8-130)式估计出的误差是偏大的,因为与某些系数相乘时,并不出现舍入误差。

DFT 的定点计算也受到动态范围的限制,因而应防止出现溢出。即要求

$$|X(k)| = \left| \sum_{n=0}^{N-1} x(n) W_N^{nk} \right| \leq \sum_{n=0}^{N-1} |x(n)| < 1 \quad (8-131)$$

满足这一充分条件,就不会出现溢出。由于  $x(n)$  在定点运算中都是小于 1 的定点数,所以只要把输入  $x(n)$  乘一个  $1/N$  因子,就可以防止溢出,这是一种在最坏的情况下也不会出现溢出的条件。一般为了提高输出信号/噪声比值,总是想法使输入尽量选得大一些。

## 二、定点 FFT 计算中有限字长效应的分析

定点 FFT 的算法不同,则运算的有限字长效应就不同。我们只讨论用得最多的基-2 FFT 算法,且以 DIT(时间抽选)法为例来加以分析。所得结果只要稍加修正即可适用于 DIF(频率抽选)基-2FFT 算法以及其他一些算法。

设序列长度为  $N=2^L$ ,需计算  $L=\log_2 N$  级,每级为  $N$  个数构成的数列,讨论原位运算的 DIT 的蝶形运算,每级有  $N/2$  个单独的蝶形结计算,由  $m$  列到  $m+1$  列的蝶形计算可表示为

$$\begin{aligned} X_{m+1}(i) &= X_m(i) + W_N^r X_m(j) \\ X_{m+1}(j) &= X_m(i) - W_N^r X_m(j) \end{aligned} \quad (8-132)$$

$i, j$  表示同一列中,一对蝶形节点在这一列中的位置(行的数值),此蝶形结构已示于图 4-7 中。用定点法实现时,只有相乘才需舍入,仍以加性误差来考虑相乘舍入的影响,则蝶形结的定点舍入统计模型如图 8-29 所示,对于此图  $m=0$  时,表示输入序列,  $m+1=L$  时表示输出序列,即为所要求的离散傅里叶变换。图中,  $e(m,j)$  表示  $X_m(j)$  与  $W_N^r$  相乘

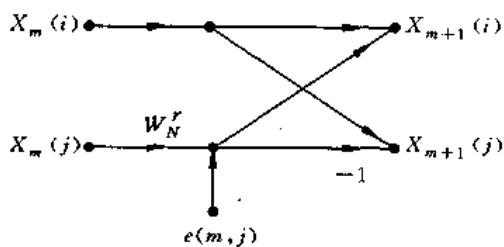


图 8-29 基-2FFT,DIT 算法蝶形结的定点舍入统计模型

所引入的舍入误差源,这一误差源是复数,每个复乘包括 4 个实乘,每个定点实乘产生一个舍入误差源。我们假定每个误差源具有和上述 DFT 误差源相同的统计特性,因而和(8-129)式一样,一个复乘运算所引入的误差的方差为

$$E[|e(m,j)|^2] = \frac{2^{-2b}}{3} = \sigma_B^2 \quad (8-133)$$

误差源  $e(m,j)$  通过后级蝶形结时, 其方差是不会变化的; 通过加、减运算是不影响方差的, 通过乘系数  $W_N^0$  后方差也无影响, 因为

$$E[|e(m,j)W_N^0|^2] = |W_N^0|^2 E[|e(m,j)|^2] = E[|e(m,j)|^2]$$

所以  $e(m,j)$  通过所有蝶形结时, 其方差均保持不变。这样, 计算 FFT 的最后输出误差, 只需知道结点共联结多少个蝶形结即可, 每个蝶形结产生误差的方差为  $\sigma_B^2$ 。若以  $F(k)$  表示  $X(k)$  上叠加的输出误差, 它和末级的一个蝶形结联接, 和末前级的两个蝶形结联接, 依此类推, 每往前一级, 引入的误差源就增加一倍, 因此, 连接到  $X(k)$  末端的误差源总数为

$$1 + 2 + 2^2 + \cdots + 2^{L-1} = 2^L - 1 = N - 1$$

图 8-30 表示  $N=8$  时, DIT 算法中, 联接到  $X(0)$  的各蝶形结情况, 因而, 在终端, 即在离散傅里叶变换  $X(k)$  上叠加的输出噪声  $F(k)$  的均方值(即为方差, 因均值为零)为

$$E[|F(k)|^2] = (N-1)\sigma_B^2$$

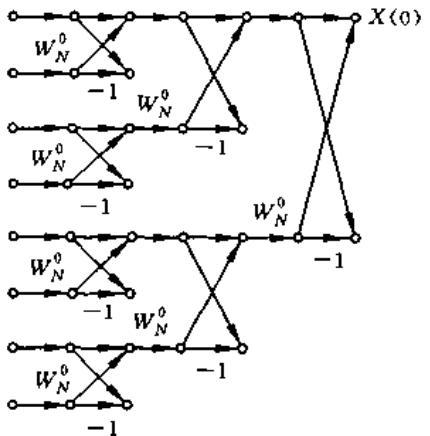


图 8-30 对  $X(0)$  起作用的蝶形( $N=8$ )

当  $N$  很大时, 可近似认为

$$E[|F(k)|^2] \approx N\sigma_B^2 = \frac{N\Delta^2}{3} \quad (8-134)$$

下面要进一步估算输出的噪声/信号比值。此时要讨论定点制的动态范围问题, 也就是要保证不发生溢出的条件下来讨论输出噪声/信号比值。

从(8-132)式看出

$$\begin{aligned} \max(|X_m(i)|, |X_m(j)|) &\leq \max(|X_{m+1}(i)|, |X_{m+1}(j)|) \\ &\leq 2\max(|X_m(i)|, |X_m(j)|) \end{aligned} \quad (8-135)$$

这表示, 从前一级到后一级, 最大模值是逐级非减的, 只要最后一级不出现溢出, 则前一级计算一定不会溢出。

有以下三种防止溢出的办法:

第一种办法: 由于蝶形结输出的最大模值小于等于输入最大模值的两倍, 总共有  $L = \log_2 N$  级蝶形结, 因此 FFT 最后输出最大值小于等于输入最大值的  $2^L = N$  倍, 即

$$\max |X(k)| \leq 2^L \max |x(n)| = N \max |x(n)|$$

要使  $X(k)$  不溢出, 需  $\max |X(k)| < 1$ , 因而要求

$$|x(n)| < \frac{1}{N}, \quad 0 \leq n \leq N-1 \quad (8-136)$$

如果假设  $x(n)$  是白色的, 且假定其实部虚部是在  $(-\frac{1}{N\sqrt{2}}, \frac{1}{N\sqrt{2}})$  区间内均匀等概分布的, 且二者互不相关, 则  $x(n)$  在  $(-\frac{1}{N}, \frac{1}{N})$  区间内是均匀等概分布的, 因而  $x(n)$  的方差为

$$\sigma_x^2 = E[|x(n)|^2] = \frac{1}{3N^2} \quad (8-137)$$

由于

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk}$$

因而

$$E[|X(k)|^2] = \sum_{n=0}^{N-1} E[|x(n)|^2 + W_N^{nk}]^2 = N\sigma_x^2 = \frac{1}{3N} \quad (8-138)$$

这样, 我们就得到在这种防止溢出的办法 ( $|x(n)| < 1/N$ ) 下, 输出的噪声/信号比值

$$\frac{E[|F(k)|^2]}{E[|X(k)|^2]} = 3N^2\sigma_B^2 = N^2\Delta^2 = N^22^{-2b} \quad (8-139)$$

从此式看出, 当输入为白噪声且满足  $|x(n)| < 1/N$  时, 也就是如果原来输入满足定点小数要求  $x(n) < 1.0$ , 那么在输入端乘上  $1/N$  的比例因子, 此时输出噪声/信号比值与  $N^2$  成正比, 同时还看出, 如果 FFT 每增加一级运算 ( $L$  增加 1), 即  $N$  加倍, 噪声/信号比值将增加到 4 倍, 或者说, 为了保持同一运算精度 (即同一噪声/信号比值), 每增加一级运算,  $\Delta^2$  必须降低 4 倍, 也就是字长 ( $b$ ) 需要增加一位。

如果输入不是白噪声, 噪声/信号比值仍然正比于  $N^2$ , 只不过比例常数有所改变而已。

这种防止溢出的办法, 使得输入幅度被限制得过小, 造成输出信号/噪声比值过小。我们要想法加以改善。

第二种办法: 由 (8-135) 式看出, 一个蝶形结的最大输出幅度不超过输入的两倍, 又知输入是满足  $|x(n)| < 1$  的, 因而, 如果对每个蝶形结的两个输入支路都乘上  $1/2$  个比例因子, 其统计模型如图 8-31 所示, 就可保证蝶形结运算不发生溢出, 对  $L = \log_2 N$  级蝶形, 就相当于设置了  $\left(\frac{1}{2}\right)^L = \frac{1}{N}$  的比例因子。和第一种办法不同之处在于这里把  $1/N$  的比例因子分散到各级运算之中。因此这里在保持输出信号方差 (8-138) 式不变的情况下, 输入幅度却增加了  $N$  倍, 达到

$$|x(n)| < 1 \quad (8-140)$$

这时, 对白色输入信号, 这里得到的最大输出信号幅度仍和前面一样, 但是输出噪声电平却比 (8-134) 式中的要小得多, 这是因为 FFT 前几级引入的噪声都被后面几级的比例因子衰减掉了。

由图 8-31 看出,由于引入了  $1/2$  的比例因子,因而每个蝶形结都有两个噪声源与之相连。和前面一样,假定噪声源的实部与虚部不相关,并与其它噪声源也不相关,并假设实部与虚部是在  $(-\frac{1}{2} \cdot 2^{-b}, \frac{1}{2} \cdot 2^{-b}]$  区间内均匀等概分布的。由于  $e(m, i)$  是由  $x_m(i)$  乘以实系数  $1/2$  所引入的,而  $e(m, j)$  是由  $X_m(j)$  乘以复系数  $W_N^r$  所引入的,因此  $e(m, i)$  的方差应不大于  $e(m, j)$  的方差  $\sigma_B^2$  ( $\sigma_B^2$  见(8-133)式的表达式,是由一个复乘所引入的误差的方差),这样,一个蝶形结乘  $1/2$  后所形成的总的误差方差为

$$\sigma_{B'}^2 = E[|e(m, i)|^2] + E[|e(m, j)|^2] \leq 2\sigma_B^2$$

由于这个误差每通过一级蝶形结,受比例因子的加权作用,其幅度要下降到  $1/2$ ,方差要下降到  $1/4$ ,因而,输出总方差中末级引入误差是不受衰减的,但每往前一级,其引入误差的方差要降低到  $1/4$  倍。参照图 8-30 蝶形结的联接,则可求出总的输出噪声方差为

$$\begin{aligned} E[|F(k)|^2] &= \sigma_k^2 = \sigma_{B'}^2 + \frac{1}{4}2\sigma_{B'}^2 + \frac{1}{4^2}4\sigma_{B'}^2 + \cdots + \frac{1}{4^{L-1}}2^{L-1}\sigma_{B'}^2 \\ &= \frac{1 - \left(\frac{1}{2}\right)^L}{1 - \frac{1}{2}}\sigma_{B'}^2 \approx 2\sigma_{B'}^2 = 4\sigma_B^2 = \frac{4}{3} \cdot 2^{-2b} \end{aligned} \quad (8-141)$$

考虑到输出信号的方差仍为  $E[|X(k)|^2] = 1/(3N)$  (见(8-138)式),则可得到输出噪声/信号比值为

$$\frac{E[|F(k)|^2]}{E[|X(k)|^2]} = 12N\sigma_B^2 = 4N2^{-2b} \quad (8-142)$$

由此式看出,此时噪声/信号比值不再是和  $N^2$  成正比,而是和  $N$  成正比,即当  $N$  每增加到 4 倍时(增加两级蝶形运算),为了保持输出噪声/信号比值不变,寄存器长度( $b$ )只需增加一位。这显然比把  $1/N$  的比例因子全放在输入端的情况要好多了。这个结论不仅对白噪声输入序列正确,对其他很多输入序列也都成立,只不过和  $N$  成正比的比例系数有所不同罢了。

因而,只要有可能,就应把  $1/N$  的衰减分散到各级蝶形中去,在每一级中插入  $1/2$  的比例因子,以提高输出的信号/噪声比值。

第三种办法:采用成组浮点运算。由于每次运算结果,只要保证  $\text{Re}[X_{m+1}(i)]$ ,  $\text{Im}[X_{m+1}(i)]$  不溢出就行了,故先将原始数据用成组浮点制表示为

$$\text{Re}[x(i)] = 2^p \text{Re}[x'(i)], \quad \text{Im}[x(i)] = 2^p \text{Im}[x'(i)]$$

这样来保证  $\text{Re}[x'(i)] < 1$ ,  $\text{Im}[x'(i)] < 1$ ,这里  $p$  为某个整数,称为共阶数。我们以  $[x'(i)]$  作为 FFT 的初始数据进行变换,计算按定点运算进行,一旦某个蝶形结计算中出现溢出,则将整个这一级的序列(运算过的,未运算的)全部右移一位(即乘  $1/2$  因子),并

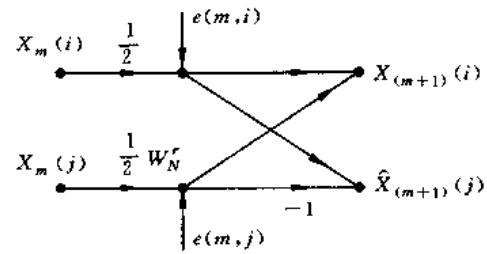
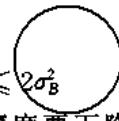


图 8-31 乘比例因子  $\frac{1}{2}$  后的蝶形结统计模型



在阶码  $p$  上加 1, 然后运算从发生溢出的蝶形继续下去。当以后某级蝶形运算又出现溢出时, 再对该级的输入乘上  $1/2$  衰减因子, 并在阶码  $p$  上加 1 继续运算, 直到  $L$  级计算完为止。这种办法要增加检验溢出的判断步骤。

为了确定整个数列最后的比例因子, 要计算移位的次数。此法的误差分析较为困难, 这是因为输出信号/噪声比值极大地依赖于溢出多少次以及溢出发生在哪一级。后者又与输入信号有关, 因而还需要知道输入信号的性质。

### 三、系数量化对 FFT 的影响

当系统确定后, 系数值就是已知的, 因而量化后系数值也是可知的, 所以对一个具体系统, 系数量化误差不是随机性的。下面我们利用统计的分析方法, 目的是为了在不知系数具体数值时, 在一定字长下, 对系数量化造成的影响作一统计估计。

所采用的办法是在系数上引入一个随机扰动, 也就是系数用它的真值加上一个白噪声序列来代替, 从而估计这个噪声引起的输出噪声/信号比值。这与量化引起的系数误差有细微的不同, 但二者大体上相似。

理想 FFT 运算为

$$X(k) = \sum_{n=0}^{N-1} x(n) W_N^{nk} \quad (8-143)$$

当系数量化后, 上式可表示成

$$\hat{X}(k) = \sum_{n=0}^{N-1} x(n) \hat{W}_N^{nk} = X(k) + F(k) \quad (8-144)$$

式中  $F(k)$  是由系数量化引起的 DFT 计算误差, 从某个  $x(n)$  计算某个  $X(k)$  要经过  $L = \log_2 N$  个蝶形结, 故  $\hat{W}_N^{nk}$  中有  $L$  个因子, 这些因子中有的是  $W_N$  的各次幂系数, 有的可能只是  $\pm 1$ , 即不需要乘任何系数, 但为了统计方便, 我们假定  $x(n)$  通过每级蝶形结时, 都乘了一个系数  $W_N^a$  ( $i=1, 2, \dots, L$ ), 故总乘积为  $\prod_{i=1}^L W_N^a = W_N^{ak}$ , 也就是  $x(n) W_N^{nk}$  是  $X(k)$  中的一项。系数量化后, 每个支路的  $W_N^a$  变成  $W_N^a + \delta_i$ , 因而有

$$\hat{X}(k) = \sum_{n=0}^{N-1} x(n) \hat{W}_N^{nk} = \sum_{n=0}^{N-1} x(n) \left[ \prod_{i=1}^L (W_N^a + \delta_i) \right] = X(k) + F(k) \quad (8-145)$$

故系数量化所引起的 DFT 计算误差可表示为

$$F(k) = \hat{X}(k) - X(k) = \sum_{n=0}^{N-1} x(n) [\hat{W}_N^{nk} - W_N^{nk}] \quad (8-146)$$

由于

$$\hat{W}_N^{nk} = \prod_{i=1}^L (W_N^a + \delta_i) = \prod_{i=1}^L W_N^a + \sum_{i=1}^L \delta_i \prod_{j=1, j \neq i}^L W_N^a + (\delta_i \text{ 的高阶项}) \quad (8-147)$$

故

$$\hat{W}_N^{nk} - W_N^{nk} = \sum_{i=1}^L \delta_i \prod_{j=1, j \neq i}^L W_N^a + (\delta_i \text{ 的高阶项}) \quad (8-148)$$

略去  $\delta_i$  的高阶项, 并假定各  $\delta_i$  是统计独立、白色等概分布的随机变量, 则有

$$E[\hat{W}_N^{nk} - W_N^{nk}] \approx E\left[\sum_{i=1}^L \delta_i \prod_{\substack{j=1 \\ j \neq i}}^L W_N^{nj}\right] = \prod_{i=1}^L E[\delta_i] \prod_{\substack{j=1 \\ j \neq i}}^L W_N^{nj} = 0$$

由于  $x(n)$  为已知, 因而从(8-146)式知

$$E[F(k)] = 0$$

因为  $\delta_i$  的方差为

$$\sigma_\delta^2 = E[|\delta_i|^2] = E[|\operatorname{Re}(\delta_i)|^2 + |\operatorname{Im}(\delta_i)|^2] = 2\sigma_\delta^2 = \frac{\Delta^2}{6} = \frac{1}{6} \cdot 2^{-2b} \quad (8-149)$$

为了方便, 令

$$\prod_{\substack{j=1 \\ j \neq i}}^L W_N^{nj} = b_i \quad (8-150)$$

则可得到  $(\hat{W}_N^{nk} - W_N^{nk})$  的方差为

$$\begin{aligned} E[|\hat{W}_N^{nk} - W_N^{nk}|^2] &\approx E\left[\left(\sum_{i=1}^L \delta_i b_i\right)\left(\sum_{j=1}^L \delta_j b_j\right)^*\right] = E\left[\sum_{i=1}^L |\delta_i b_i|^2\right] \\ &= E\left[\sum_{i=1}^L |b_i|^2 |\delta_i|^2\right] \end{aligned}$$

因为

$$|b_i| = \prod_{\substack{j=1 \\ j \neq i}}^L |W_N^{nj}| = 1$$

所以有

$$E[|\hat{W}_N^{nk} - W_N^{nk}|^2] \approx L E[|\delta_i|^2] = L \sigma_\delta^2 = \frac{L}{6} \Delta^2 = \frac{L}{6} \cdot 2^{-2b} \quad (8-151)$$

由(8-146)式可得输出误差  $F(k)$  的方差为

$$\begin{aligned} \sigma_F^2 &= E[|F(k)|^2] = \sum_{n=0}^{N-1} E[|x(n)(\hat{W}_N^{nk} - W_N^{nk})|^2] \\ &= \sum_{n=0}^{N-1} |x(n)|^2 E[|\hat{W}_N^{nk} - W_N^{nk}|^2] \approx \sum_{n=0}^{N-1} |x(n)|^2 \frac{L}{6} \cdot 2^{-2b} \quad (8-152) \end{aligned}$$

根据 DFT 的帕塞瓦定理

$$\sum_{n=0}^{N-1} |x(n)|^2 = \frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2 \quad (8-153)$$

如果把  $X(k)$  看成输出信号, 则(8-153)式右端为输出信号的均方值, 从而我们可得到输出均方误差  $\sigma_F^2$  与输出均方信号的比值为

$$\frac{\sigma_F^2}{\frac{1}{N} \sum_{k=0}^{N-1} |X(k)|^2} = \frac{L}{6} \cdot 2^{-2b} \quad (8-154)$$

由此得出结论: 输出噪声/信号比值随  $N$  增大而增加的速度极为缓慢, 只与  $L = \log_2 N$  成正比。 $N$  加倍时,  $L$  只增加 1, 输出噪声/信号比值增加极小。上面的式子通常要比实际误差稍大一些, 但仍可作粗略的估计之用。

## 习 题

1. 设数字滤波器的系统函数为

$$H(z) = \frac{0.017221333z^{-1}}{1 - 1.7235682z^{-1} + 0.74081822z^{-2}}$$

现用 8bit 字长的寄存器来存放其系数, 试求此时该滤波器的实际  $H(z)$  表示式。

2. 图 P8-2(a) 为一阶系统的流图

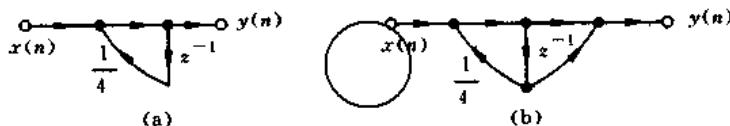


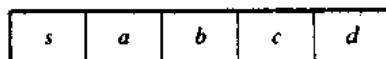
图 P8-2

(a) 求系统对如入输入的响应:

$$x(n) = \begin{cases} \frac{1}{2}, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

$n$  较大时, 系统的响应是什么?

(b) 系统用定点算法实现。网络中的系数和所有变量都用 5 位寄存器表示成原码, 即



$s$  为符号位, 寄存器值  $= a \times 2^{-1} + b \times 2^{-2} + c \times 2^{-3} + d \times 2^{-4}$ , 其中  $a, b, c, d$  是 1 或 0。乘法的结果作截尾处理, 即只保留符号位和前四位。

试计算已量化的系统对 (a) 中输入的响应, 画出未量化系统在  $0 \leq n \leq 5$  时响应。问  $n$  比较大时如何比较这两种响应?

(c) 研究 P8-2(b) 所示系统, 其输入为

$$x(n) = \begin{cases} \frac{1}{2}(-1)^n, & n \geq 0 \\ 0, & n < 0 \end{cases}$$

重作 (a), (b)。

(d) 当尾数采用舍入处理时, 重作 (b), (c)。

3. A/D 变换器的字长为  $b$ , 其输出端接一网络, 网络的单位抽样响应为

$$h(n) = [a^n + (-a)^n]u(n)$$

试求网络输出的 A/D 量化噪声方差  $\sigma_f^2$ 。

4. 一个二阶 IIR 滤波器, 其差分方程为

$$y(n) = y(n-1) - ay(n-2) + x(n)$$

现采用  $b=3$  位的定点制运算, 舍入处理。

(a) 当系数  $a=0.75$ , 零输入  $x(n)=0$ , 初始条件为  $\hat{y}(-2)=0, \hat{y}(-1)=0.5$ 。求  $0 \leq n \leq 9$  的 10 点输出  $\hat{y}(n)$  值。

(b) 证明当  $Q_R[a\hat{y}(n-2)] = \hat{y}(n-2)$  时发生零输入极限环振荡, 并用等效极点迁移来解释这个现象。

5. 一个一阶 IIR 网络, 差分方程为

$$y(n) = ay(n-1) + x(n)$$

采用定点制原码运算,尾数作截尾处理。

(a) 证明:只要系统稳定,即 $|a|<1$ ,就不会发生零输入极限环振荡。

(b) 若采用定点补码运算,尾数作截尾处理,这时以上结论仍然成立吗?

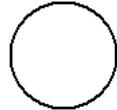
#### 6. 两个一阶 IIR 网络

$$H_1(z) = \frac{1}{1 - 0.9z^{-1}}, \quad H_2(z) = \frac{1}{1 - 0.1z^{-1}}$$

用定点制运算,舍入处理,要求输出精度 $\sigma_f^2/\sigma_v^2 = -80\text{dB}$ ,问各需几位尾数字长。

7. 在定点制运算中为了输出不发生溢出,往往必须在网络的输入端加一比例因子 $A$ ,即网络输出为

$$y(n) = A \sum_{m=0}^{\infty} h(m)x(n-m)$$



若输入 $x(n)$ 的动态范围为 $\pm x_{\max}$ ,则比例因子 $A$ 可以这样来确定

$$|y(n)| \leq A \sum_{m=0}^{\infty} |h(m)| |x(n-m)|$$

因此

$$y_{\max} \leq Ax_{\max} \sum_{m=0}^{\infty} |h(m)|$$

为了保证不发生溢出,必须使 $y_{\max} \leq 1$ ,故

$$A \leq \frac{1}{x_{\max} \sum_{m=0}^{\infty} |h(m)|}$$

现有二阶网络

$$H(z) = \frac{1}{(1 - 0.9z^{-1})(1 - 0.8z^{-1})}$$

采用定点制运算,输入动态范围为 $x_{\max} \leq 1$ 。

(a) 用直接型结构时[见图 P8-7(a)],为使运算过程中任何地方都不出现溢出,比例因子 $A_1$ ,应该选多大?

(b) 采用级联型结构[见图 P8-7(b)],比例因子 $A_2$ ,应选多大?

(c) 在级联结构中每一单元网络分别加一比例因子[见图 P8-7(c)],以使该环节不出现溢出,这时比例因子 $A_3, A_4$ ,应选多大?

(d) 在以上三种情况下,信号的最大输出 $y_{\max}$ 各为多少? 输出信号噪声比 $y_{\max}^2/\sigma_f^2$ ,谁最高谁最低?

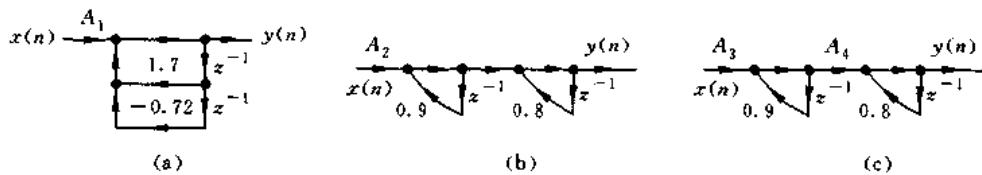


图 P8-7

#### 8. 考虑离散傅里叶变换

$$X(k) = \sum_{n=0}^{N-1} x(n)W_N^k, \quad 0 \leq k \leq N-1$$

其中 $W_N = e^{-j2\pi/N}$ .假设序列值 $x(n)$ 是一均值为零的平稳白噪声序列的 $N$ 个相邻序列值,即

$$E[x(n)x(m)] = \sigma_x^2 \delta(n-m), \quad E[x(n)] = 0$$

(a) 试确定  $|X(k)|^2$  的方差。

(b) 试确定离散傅里叶变换诸值间的互相关, 即确定  $E[X(k)X^*(r)]$ , 并把它表示为  $k$  和  $r$  的函数。

### 9. 一个二阶网络

$$H(z) = \frac{0.2}{(1 - 0.4z^{-1})(1 - 0.9z^{-1})}$$

用 6 位字长舍入方式对其系数进行量化, 试用统计方法估算在以下三种结构下, 系数量化所引起的频响偏移的方差值  $\sigma_e^2$ 。

(a) 直接型结构; (b) 级联型结构; (c) 并联型结构。

### 10. 设数字滤波器

$$H(z) = \frac{0.06}{1 - 0.6z^{-1} + 0.25z^{-2}} = \frac{0.06}{1 + a_1z^{-1} + a_2z^{-2}}$$

利用  $a_1, a_2$  变化造成的极点位置灵敏度,  $a_1, a_2$  分别造成极点在正常值的 0.2%, 0.3% 内变化, 试确定所需的最小字长。

### 11. 一个二阶 IIR 滤波器的系统函数为

$$H(z) = \frac{0.6 - 0.42z^{-1}}{(1 - 0.4z^{-1})(1 - 0.8z^{-1})}$$

现用  $b$  位字长的定点制运算实现它, 尾数作舍入处理。

(1) 试计算直接 I 型及直接 II 型结构的输出舍入噪声方差。

(2) 如果用一阶网络的级联结构来实现  $H(z)$ , 则共有六种网络流图。试画出有运算舍入噪声时的每种网络流图并计算每种流图的输出舍入噪声方差。

(3) 用并联结构实现  $H(z)$ , 计算输出舍入噪声方差, 几种结构相比较, 运算精度哪种最高、哪种最低?

(4) 考虑动态范围, 因为系统中任一节点的输出值(包括整个系统的输出节点)等于从输入到此节点的单位冲激响应与系统输入的卷积和。可以表示成

$$y_i(n) = \sum_{k=-\infty}^{\infty} h_i(k)x(n-k)$$

其中  $y_i(n)$  为第  $i$  个节点的输出,  $h_i(n)$  为从输入到第  $i$  个节点的单位抽样响应。对于输出节点来说  $y_i(n) = y(n)$ ,  $h_i(n) = h(n)$ 。由上式可得

$$|y_i(n)| \leq \sum_{k=-\infty}^{\infty} |h_i(k)| |x(n-k)|$$

也就是说, 一个网络的最大输出电平不一定在输出端, 可能在某一中间节点, 利用这一关系以及  $x_{\max}$ , 试求以上各种网络中每一个的最大  $y_{\max}$ , 要求网络的所有节点上都不发生溢出, 即要最大输出  $y_{\max} < 1$ , 这样即可求得最大的输入  $x_{\max}$ (不发生溢出时)。试求以上各个网络的  $x_{\max}$ 。

(5) 设输入信号是白噪声序列, 它的幅度在  $-x_{\max}$  到  $x_{\max}$  之间均匀分布, 按照已求出的每一滤波器结构的最大输入  $x_{\max}$ , 求每种结构在输出端的噪声信号比值(输出噪声方差与输出信号均方值之比)。问哪种结构输出噪声信号比值最低。

### 12. 一个 $N$ 阶 FIR 滤波器

$$H(z) = \sum_{i=0}^N a_i z^{-i}$$

采用直接型结构, 用 6 位字长舍入方式对其系数取量化。

(a) 试用统计方法估算由于系数量化所引起的频率响应偏移的方差  $\sigma_e^2$ 。

(b) 当  $N=1024$  时, 若要求  $\sigma_e^2 \leq 10^{-8}$ , 则系数字长  $b$  应该需要多少位?