第5章　数字滤波器的结构

数字滤波器是指完成信号滤波处理功能的、用有限精度算法实现的离散时间线性非时变系统，其输入是一组由模拟信号采样和量化的数字量，其输出是经过变换或处理的另一组数字量。数字滤波器既可以是用数字硬件装配成的一台完成给定运算的专用数字计算机，也可以是将所需的运算编成程序，由通用计算机来执行。数字滤波器的结构是指滤波器的物理或逻辑布局，这决定了它们的工作方式和特性。

数字滤波器可以分成两大类。一类称为经典滤波器，即一般滤波器，特点是输入信号中有用的频率成分和希望滤除的频率成分各占有不同的频带，通过一个合适的选频滤波器达到滤波的目的。例如，当输入信号中含有干扰时，如果信号和干扰的频带互不重叠，即可滤除干扰得到想要的信号。但对于一般滤波器，如果信号和干扰的频带互相重叠，则不能完成对干扰的有效滤除，这时需要采用另一类所谓的现代滤波器，如维纳(Wiener)滤波器、卡尔曼(Kalman)滤波器、自适应滤波器等，这些滤波器可按照随机信号内部的一些统计规律，从干扰中最佳地提取信号。

本书仅介绍经典滤波器的设计，本章介绍数字滤波器的2种基本结构，包括IIR(无限冲激响应)滤波器和FIR(有限冲激响应)滤波器。

视频讲解

5.1　数字滤波器概述

理想滤波器是不可能实现的，因为它们的单位冲激响应均是非因果且无限长的，设计者只能按照某些准则设计实际滤波器，使之尽可能逼近理想滤波器。数字滤波器的传输函数$H(e^{j\omega})$都是以2π为周期的，滤波器的低通频带处于2π的整数倍处，而高通频带处于π的奇数倍附近，如图5.1.1所示。数字滤波器按实现的网络结构或单位冲激响应分类，可以分为无限冲激响应(IIR)滤波器和有限冲激响应(FIR)滤波器。它们的系统函数分别表示为

$$H(z)=\frac{Y(z)}{X(z)}=\frac{\sum_{r=0}^{M}b_r z^{-r}}{1-\sum_{k=1}^{N}a_k z^{-k}} \tag{5.1.1}$$

$$H(z)=\sum_{n=0}^{N-1}h(n)z^{-n} \tag{5.1.2}$$

式(5.1.1)中，当满足$M<N$时，这类系统称为N阶IIR系统；当$M\geqslant N$时，可视为一

个 N 阶 IIR 子系统与一个 $M-N$ 阶 FIR 子系统(多项式)的级联。式(5.1.2)所示系统称为 $N-1$ 阶 FIR 系统。这两种类型的滤波器的设计方法有很大区别,下面将分别进行介绍。

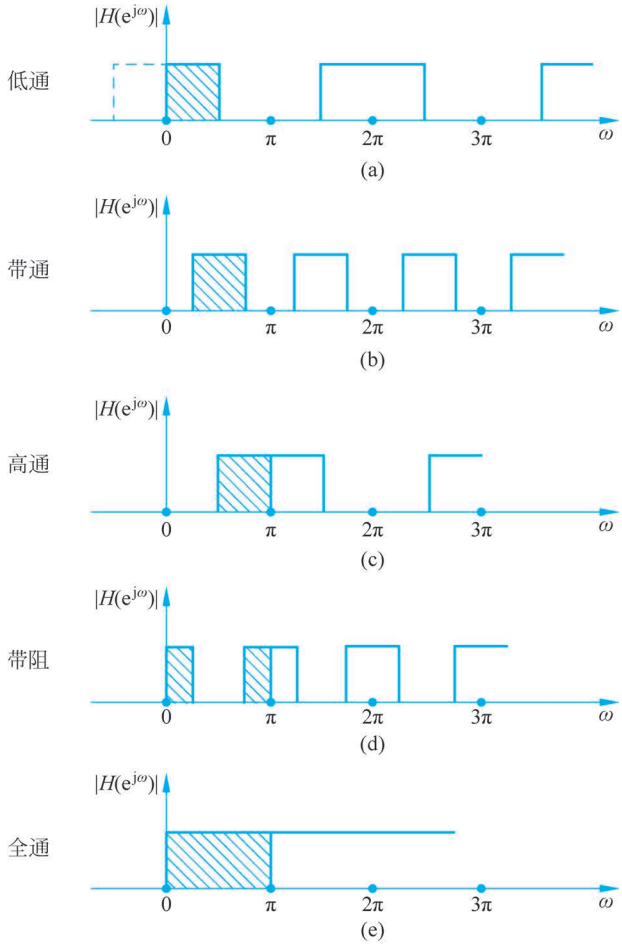


图 5.1.1　各种数字滤波器的理想幅频响应

5.1.1　滤波器的技术指标

数字滤波器的传输函数 $H(e^{j\omega})$ 可以表示为

$$H(e^{j\omega})=|H(e^{j\omega})|e^{j\theta(\omega)}$$

$|H(e^{j\omega})|$ 称为幅频特性,$\theta(\omega)$ 称为相频特性。幅频特性表示信号通过该滤波器后各频率的衰减情况,而相频特性反映各频率成分通过滤波器后在时间上的延时情况。因此,即使两个滤波器幅频特性相同,而相频特性不一样,对相同的输入,滤波器输出的信号波形也是不一样的。一般选滤波器的技术要求由幅频特性给出,而对相频特性一般不做要求,但如果对输出波形有要求,则需要考虑相频特性的技术指标,例如语音合成、波形传输、图像信号处理等,需要设计成线性相位数字滤波器。

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,即通带和阻带都允

许有一定的误差容限。按照实际需要,确定滤波器的性能要求。通常是在频域中给定数字滤波器的性能要求。图 5.1.2 所示为低通滤波器的性能要求。虚线表示满足预定性能要求的系统的频率响应,ω_p 为通带截止频率,在通带内幅度响应以$\pm\delta_1$ 的误差接近于 1,即

$$1-\delta_1 \leqslant |H(e^{j\omega})| \leqslant 1+\delta_1 \quad |\omega| \leqslant \omega_p$$

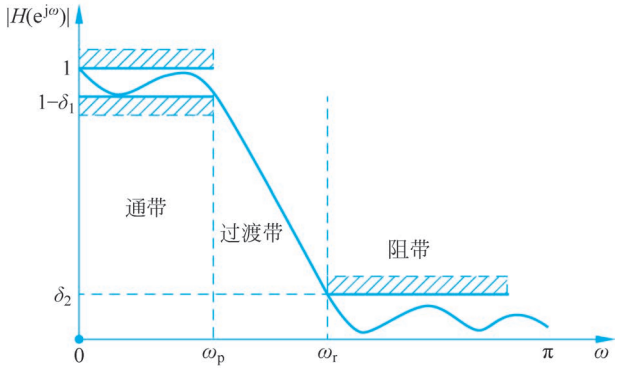


图 5.1.2 逼近理想低通滤波器的误差容限

ω_s 为阻带起始频率,在阻带内幅度响应以小于 δ_2 的误差接近于零,即

$$|H(e^{j\omega})| \leqslant \delta_2 \quad \omega_s \leqslant |\omega| \leqslant \pi$$

为了使逼近理想低通滤波器的方法成为可能,还必须提供宽度为($\omega_s-\omega_p$)的不为零的过滤频带。在这个频带内幅度响应从通带平滑地下落到阻带。这里 ω,包括 ω_p 和 ω_s,指的是数字域频率,而不是真实频率,或者说 ω 是沿着单位圆周的相角变化。相位特性除了受稳定性和因果性要求的限制,即要求系统函数的极点必须位于单位圆内部,没有其他任何限制。通带内和阻带内允许的衰减一般用分贝数表示,通带内允许的最大衰减用 α_p 表示,阻带内允许的最小衰减用 α_s 表示,α_p 和 α_s 分别定义为

$$\alpha_p = 20\log\frac{|H(e^{j0})|}{|H(e^{j\omega_p})|} \tag{5.1.3}$$

$$\alpha_s = 20\log\frac{|H(e^{j0})|}{|H(e^{j\omega_s})|} \tag{5.1.4}$$

对于低通滤波器,如将 $H(e^{j0})$归一化为 1,式(5.1.3)和式(5.1.4)可分别表示为

$$\alpha_p = -20\log|H(e^{j\omega_p})| \tag{5.1.5}$$

$$\alpha_s = -20\log|H(e^{j\omega_s})| \tag{5.1.6}$$

当幅度下降到$\sqrt{2}/2\approx0.707$ 时,$\omega=\omega_c$,此时 $\alpha_p=3\text{dB}$,称 ω_c 为 3dB 通带截止频率。

5.1.2 数字滤波器的设计过程

实际中的数字滤波器设计都是用有限精度算法实现,一般包括以下设计步骤。

(1) 根据实际需要确定数字滤波器的技术指标。例如滤波器的频率响应的幅度特性和截止频率等。

(2) 用一个因果稳定的离散线性非时变系统的系统函数去比逼近这些性能指标。具体而言,就是用这些指标来计算系统函数。IIR 滤波器的系统函数是 z^{-1} 的有理函数,FIR 滤

波器的系统函数是 z^{-1} 的多项式。这样，滤波器的设计问题，变成了一个数学逼近问题，即用一个因果稳定的系统函数去逼近给定的性能要求，以确定滤波器系数。

(3) 用有限精度的运算实现所设计的系统。包括选择运算结构，及对滤波器的系数、输入变量、中间变量和输出变量量化到固定字长。

(4) 通过模拟，验证所设计的系统是否符合给定性能要求。根据验证的结果决定是否对第(2)步和第(3)步作修改，以满足技术要求。

IIR 数字滤波器和 FIR 数字滤波器的设计方法是很不相同的。IIR 数字滤波器设计方法经常借助于模拟滤波器的设计方法来进行。其设计步骤是：先设计模拟原型滤波器，得到其传输函数，然后按某种方法转换成数字滤波器的系统函数。模拟滤波器设计方法已经很成熟，有完整的设计公式，还有完善的图表可供查阅，并且还有一些典型的滤波器类型可供设计者使用，得到的是闭合形式的公式。FIR 数字滤波器不能采用先设计模拟滤波器然后再转换成数字滤波器的方法，经常使用的设计方法是窗函数法和频率采样法，需要通过计算机辅助设计来完成。

5.2　无限冲激响应数字滤波器的结构

视频讲解

数字滤波器可以用差分方程、单位采样响应，以及系统函数等表示。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。这些基本的运算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两种表示方法，如图 5.2.1 所示。用方框图表示比较直观，用信号流图表示则更加简单方便。

一个给定的输入输出关系，可以用多种不同的数字网络来实现。在不考虑量化影响时，这些不同的实现方法是等效的；但在考虑量化影响时，这些不同的实现方法性能上就有差异。因此，运算结构是很重要的，同一系统函数 $H(z)$，运算结构不同，将会影响系统的精度、误差、稳定性、经济性以及运算速度等许多重要性能。

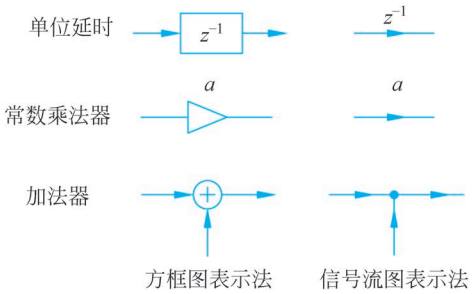


图 5.2.1　基本运算单元的方框图表示法和信号流图表示法

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，后者则影响运算速度。无限冲激响应滤波器与有限冲激响应滤波器在结构上有各自不同的特点，本书将分别对它们加以讨论。

无限冲激响应数字滤波器的基本结构包括直接型(直接Ⅰ型、直接Ⅱ型、转置型)、级联型、并联型。

5.2.1　直接型

1. 直接Ⅰ型

IIR 数字滤波器的系统函数

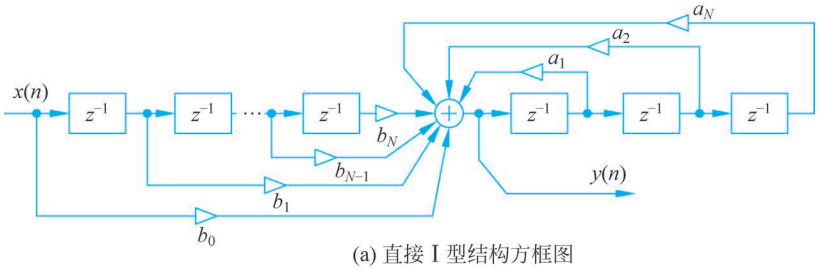
$$H(z)=\frac{\sum_{r=0}^{M} b_r z^{-r}}{1-\sum_{k=1}^{N} a_k z^{-k}} \tag{5.2.1}$$

对应的差分方程为

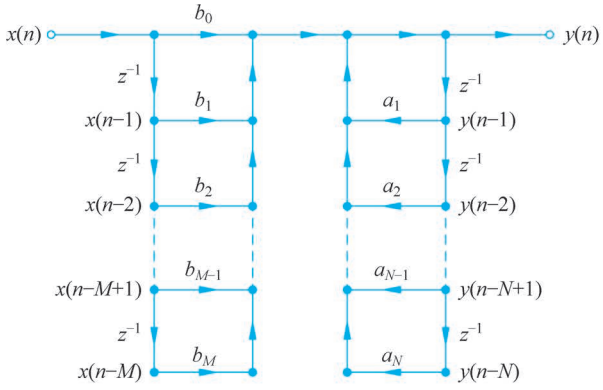
$$y(n)=\sum_{r=0}^{M} b_r x(n-r)+\sum_{k=1}^{N} a_k y(n-k) \tag{5.2.2}$$

$y(n)$由两部分相加构成：第一部分 $\sum_{r=0}^{M} b_r x(n-r)$ 是一个对输入 $x(n)$ 的 M 节延时链结构，每节延时抽头后加权相加；第二部分 $\sum_{k=1}^{N} a_k y(n-k)$ 是一个对 $y(n)$ 的延时链结构，每级延时抽头后加权相加，因此是一个反馈网络。

这种结构形式称为直接I型，如图 5.2.2 所示。其中图 5.2.2(a)为方框图，图 5.2.2(b)为信号流图。为了方便，图中画的是 $M=N$ 的情况，如果 $M\neq N$，N 阶滤波器需要 $N+M$ 级延时单元。



(a) 直接 Ⅰ 型结构方框图



(b) 直接 Ⅰ 型结构信号流图

图 5.2.2　直接 Ⅰ 型结构图

【例 5.1】　采用直接 Ⅰ 型实现系统函数如下的 IIR 数字滤波器，并求单位脉冲响应和阶跃响应。

$$H(z)=\frac{\sum_{r=0}^{M} b_r z^{-r}}{1-\sum_{k=1}^{N} a_k z^{-k}}=\frac{1-3z^{-1}+11z^{-2}+27z^{-3}+18z^{-4}}{16+12z^{-1}+2z^{-2}-4z^{-3}-2z^{-4}}$$

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import signal

# 差分方程的参数
b = np.array([1, -3, 11, 27, 18])                    # 分子
a = np.array([16, 12, 2, -4, -2])                    # 分母

# 输入信号
N = 30
delta = signal.unit_impulse(N)                       # 单位样本信号
y = np.ones(N)                                       # 单位阶跃信号

# IIR 数字滤波器
zi = signal.lfilter_zi(b, a) * 0                     # 零初始条件
z1, _ = signal.lfilter(b, a, delta, zi=zi)
z2, _ = signal.lfilter(b, a, y, zi=zi)

# 绘图
fig, axs = plt.subplots(2, 1, constrained_layout=True)
axs[0].stem(z1, basefmt="")
axs[1].stem(z2, basefmt="")
plt.rcParams['font.sans-serif'] = ['SimHei']         # 用来正常显示中文标签
axs[0].set_title('IIR 直接型 h(n)')
axs[1].set_title('IIR 直接型 y(n)')
plt.show()
fig.savefig('./iir_dir_I_sequence.png', dpi=500)
```

运行程序，结果如图 5.2.3 所示。

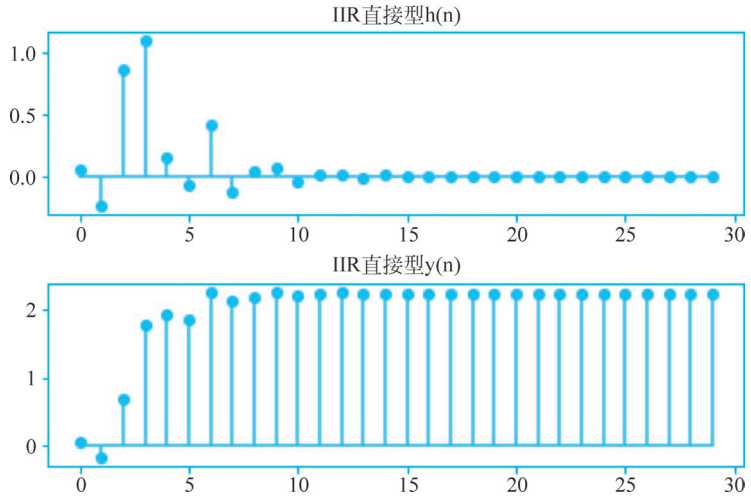


图 5.2.3　例 5.1 单位脉冲响应和阶跃响应

2. 直接Ⅱ型

将式(5.2.1)改写成下式(当 $M=N$ 的情况)

$$H(z)=\frac{Y(z)}{X(z)}=\frac{Y(z)}{W(z)}\cdot\frac{W(z)}{X(z)}=\left(\sum_{r=0}^{N}b_r z^{-r}\right)\left(\frac{1}{1-\sum_{k=1}^{N}a_k z^{-k}}\right) \tag{5.2.3}$$

因此 $H(z)$ 可视为分子多项式 $\sum_{r=0}^{N} b_r z^{-r}$ 与分母多项式 $1-\sum_{k=1}^{N} a_k z^{-k}$ 的倒数所构成的两个子系统函数的乘积，这相当于两子系统的级联。其中第一个子系统实现零点为

$$H_1(z)=\frac{Y(z)}{W(z)}=\sum_{r=0}^{N} b_r z^{-r}$$

故得

$$Y(z)=\sum_{r=0}^{N} b_r z^{-r} W(z)$$

其时域表示为

$$y(n)=\sum_{r=0}^{N} b_r w(n-r)$$

第二个子系统实现极点为

$$H_2(z)=\frac{W(z)}{X(z)}=\frac{1}{1-\sum_{k=1}^{N} a_k z^{-k}}$$

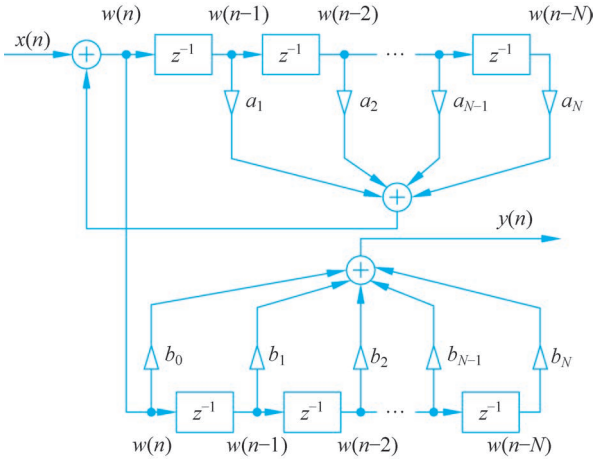
经整理后得

$$W(z)=X(z)+\sum_{k=1}^{N} a_k z^{-k} W(z)$$

其时域表示为

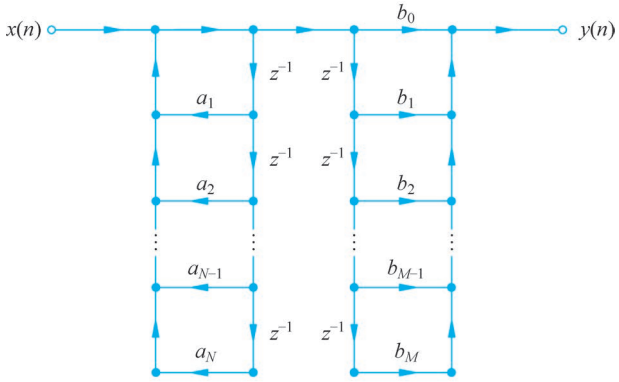
$$w(n)=x(n)+\sum_{k=1}^{N} a_k w(n-k) \tag{5.2.4}$$

综上所述，可得图 5.2.4 所示的实现结构。其中图 5.2.4(a)为方框图，图 5.2.4(b)为信号流图。



(a) 直接Ⅱ型的变型结构方框图

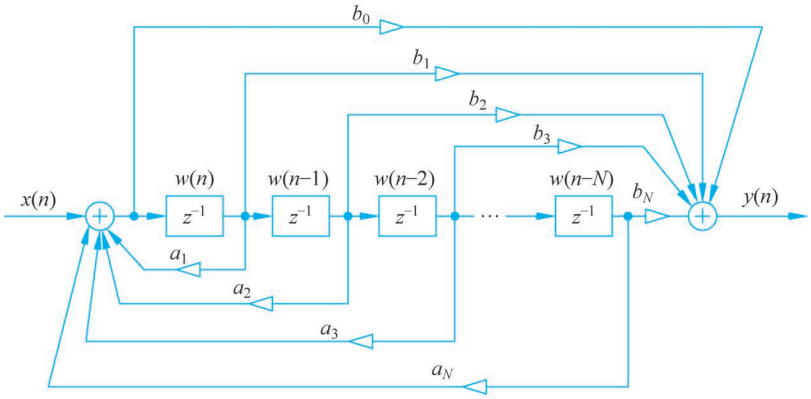
图 5.2.4 直接Ⅱ型的变型结构图



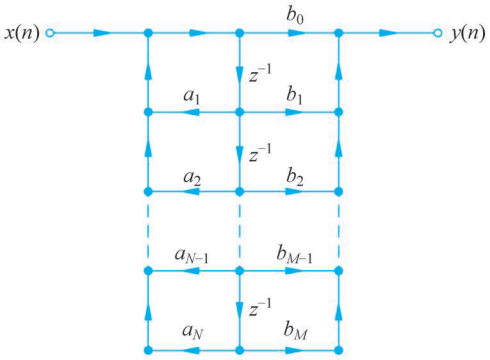
(b) 直接Ⅱ型的变型结构的信号流图

图 5.2.4 （续）

如果将图 5.2.4 中相同输出的延迟单元合成一个，则得图 5.2.5 所示的结构图。它的延迟单元少一倍，N 阶滤波器只需 N 级延迟单元，这是实现 N 阶滤波器所必需的最少数量的延迟单元。这种结构称为直接Ⅱ型。有时将直接Ⅰ型简称为直接型，而将直接Ⅱ型称为典范型。



(a) 直接Ⅱ型结构方框图



(b) 直接Ⅱ型结构信号流图

图 5.2.5　直接Ⅱ型结构图

【例 5.2】 用直接Ⅱ型实现系统函数如下的 IIR 数字滤波器，并求单位脉冲响应和单位阶跃响应。

$$H(z)=\frac{\sum_{r=0}^{M} b_r z^{-r}}{1-\sum_{k=1}^{N} a_k z^{-k}}=\frac{1-3z^{-1}+11z^{-2}+27z^{-3}+18z^{-4}}{16+12z^{-1}+2z^{-2}-4z^{-3}-2z^{-4}}$$

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import signal

# 差分方程的参数
b = np.array([[1, 0, 0, 0, 0], [1, -3, 11, 27, 18]])     # 分子
a = np.array([[16, 12, 2, -4, -2], [1, 0, 0, 0, 0]])     # 分母

# 输入信号
M = a.shape[0]
N = 30
delta = signal.unit_impulse(N)                            # 单位样本信号
y = np.ones(N)                                            # 单位阶跃信号

# IIR 直接Ⅱ型滤波器
z1 = np.zeros((M + 1, N))
z2 = np.zeros((M + 1, N))
z1[0, :] = delta
z2[0, :] = y
for i in range(M):                                        # 循环滤波,计算最终结果
    zi = signal.lfilter_zi(b[i, :], a[i, :]) * 0          # 零初始条件
    z1[i + 1, :], _ = signal.lfilter(b[i, :], a[i, :], z1[i, :], zi=zi)
    z2[i + 1, :], _ = signal.lfilter(b[i, :], a[i, :], z2[i, :], zi=zi)

# 绘图
fig, axs = plt.subplots(2, 1, constrained_layout=True)
axs[0].stem(z1[M, :], basefmt="")
axs[1].stem(z2[M, :], basefmt="")
plt.rcParams['font.sans-serif'] = ['SimHei']              # 用来正常显示中文标签
axs[0].set_title('IIR 直接Ⅱ型 h(n)')
axs[1].set_title('IIR 直接Ⅱ型 y(n)')
plt.show()
fig.savefig('./iir_dir_Ⅱ_sequence.png', dpi=500)
```

程序运行结果如图 5.2.6 所示。

3. 转置型

线性信号流图理论上有多种运算处理方法，可在保持输入和输出之间的传输关系不变的情况下，将信号流图变换成各种不同的形式。其中流图转置的方法可导出一种转置滤波器结构。具体地讲，就是把网络中所有支路方向都颠倒成反向，且输入、输出的位置互相调换一下。对于单输入、单输出系统来说，倒转后的结构和原结构的系统函数相同，但对有限字长的影响而言，转置结构与原结构性质不同。

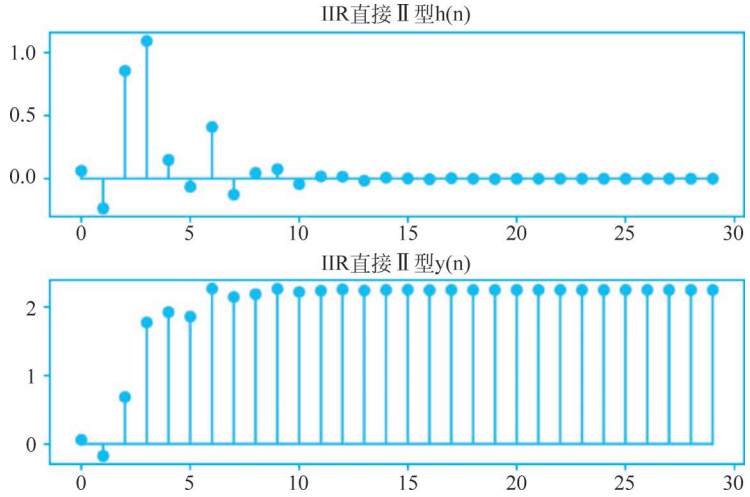


图 5.2.6　例 5.2 单位脉冲响应和单位阶跃响应

将转置原理应用于图 5.2.5 所示的直接Ⅱ型网络，并按照习惯，使输入在左边输出在右边，则得到图 5.2.7 所示的直接Ⅱ型的转置。

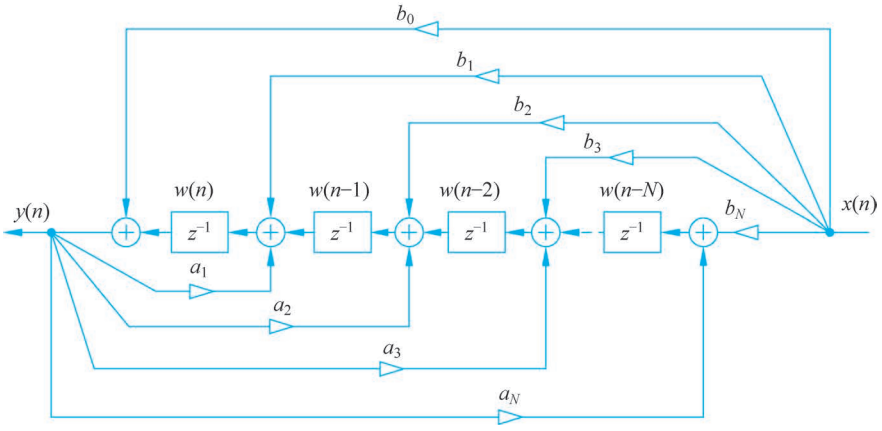


图 5.2.7　直接Ⅱ型的转置结构方框图

直接Ⅰ型和直接Ⅱ型结构优点是简单直观。它们的共同缺点是系数 a_k 对滤波器性能的控制关系不直接，因此调整不便。更严重的是这种结构的极点位置灵敏度太大，对字长效应太敏感，易于出现不稳定现象，产生较大误差。

5.2.2　级联型

将式(5.2.1)按零点和极点进行因式分解，则可表示成

$$H(z)=\frac{\sum_{r=0}^{M} b_r z^{-r}}{1-\sum_{k=1}^{N} a_k z^{-k}}=A\frac{\prod_{r=1}^{M}(1-c_r z^{-1})}{\prod_{k=1}^{N}(1-d_k z^{-1})}$$

式中 A 为归一化常数。由于系统函数 $H(z)$ 的系数都是实系数，故零点 c_r 和极点 d_k 只有两种情况：或者是实根，或者是共轭复根，即

$$H(z)=A\frac{\prod_{i=1}^{M_1}(1-g_iz^{-1})\prod_{i=1}^{M_2}(1-h_iz^{-1})(1-h_i^*z^{-1})}{\prod_{i=1}^{N_1}(1-p_iz^{-1})\prod_{i=1}^{N_2}(1-q_iz^{-1})(1-q_i^*z^{-1})} \tag{5.2.5}$$

式(5.2.5)中 $M=M_1+2M_2$，$N=N_1+2N_2$，g_i 表示实零点，p_i 表示实极点，h_i 和 h_i^* 表示复共轭零点，q_i 和 q_i^* 表示复共轭极点。每一对共轭因子合并起来，就可以构成一个实系数的二阶因子。因此

$$H(z)=A\frac{\prod_{i=1}^{M_1}(1-g_iz^{-1})\prod_{i=1}^{M_2}(1+\beta_{1i}z^{-1}+\beta_{2i}z^{-2})}{\prod_{i=1}^{N_1}(1-p_iz^{-1})\prod_{i=1}^{N_2}(1-\alpha_{1i}z^{-1}-\alpha_{2i}z^{-2})} \tag{5.2.6}$$

式(5.2.6)给出了任意系统均可由一阶和二阶子系统级联构成的表达式。如果假设实数极点和实数零点已成对合并，并把单实根因子看作二阶因子的特例，即二次项系数 α_{2i}，β_{2i} 等于零的二阶因子，同时假设 $M\leqslant N$，则整个函数 $H(z)$ 可分解为实系数二项因子的形式。

$$H(z)=A\prod_{i=1}^{L}\frac{1+\beta_{1i}z^{-1}+\beta_{2i}z^{-2}}{1-\alpha_{1i}z^{-1}-\alpha_{2i}z^{-2}}=A\prod_{i=1}^{L}H_i(z) \tag{5.2.7}$$

其中 $H_i(z)=\dfrac{1+\beta_{1i}z^{-1}+\beta_{2i}z^{-2}}{1-\alpha_{1i}z^{-1}-\alpha_{2i}z^{-2}}$，$L$ 表示 $\dfrac{(N+1)}{2}$ 中最大整数。$H_i(z)$ 称为滤波器的二阶基本节，可以采用直接Ⅱ型结构实现，如图 5.2.8 所示。其中图 5.2.8(a)为方框图，图 5.2.8(b)为信号流图。

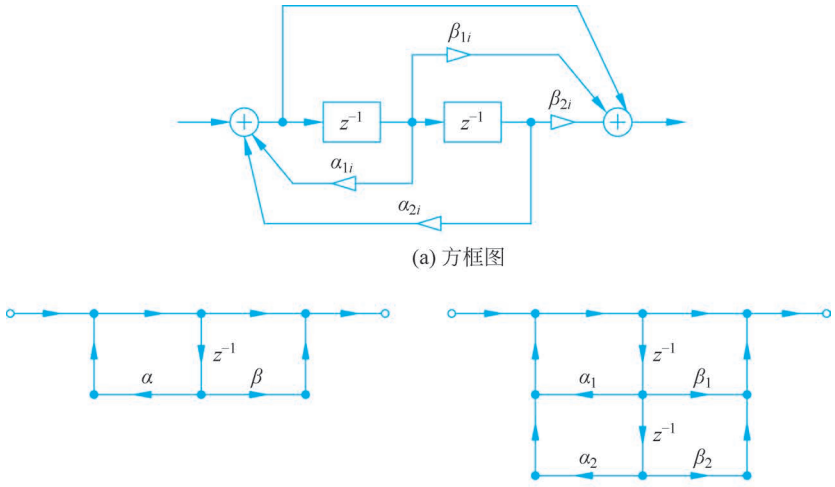


图 5.2.8　直接Ⅱ型级联结构的二阶基本节

整个滤波器则是 $H_i(z)$ 的级联，如图 5.2.9 所示。

级联型结构的一个重要优点是需要较少的存储单元，硬件实现时，可以用一个二阶基本节进行时分复用。

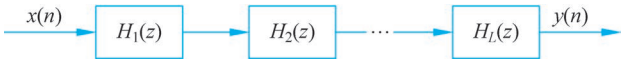


图 5.2.9　级联型结构

级联型结构的另一特点是，它的每个基本节都是关系到滤波器的一对极点和一对零点。调整系数 β_{1i}、β_{2i}，就单独地调整了滤波器的第 i 对零点而不影响其他任何零点、极点。同样，调整系数 α_{1i}、α_{2i} 也就单独调整了第 i 对极点。因此，这种结构便于准确地实现滤波器的极点、零点，也便于调整滤波器的频率响应性能。

由式(5.2.7)可见，$H(z)$中的分子分母各有 L 个二阶因子，它们可以任意两两搭配形成 $L!$ 个基本节，这些基本节中选出的任意 L 个的级联次序又可有 $L!$ 种排法，但完成的却是同一个系统函数 $H(z)$。实际工作时，由于二进制数的字长有一定限度，因此不同的排列，运算误差会各不相同。如何才能得到最好的排列，以使运算误差最小，这是最优化问题，此处不作讨论。另外级联的各基本节间要有电平的放大或缩小，以使级间输出变量不要太大或太小。级间输出变量太大，易使数字滤波器在运算过程中产生溢出；级间输出变量太小，则输出端的信号噪声比会太小。

【例 5.3】 用级联结构实现如下系统函数的 IIR 数字滤波器，并求单位脉冲响应和单位阶跃响应。

$$H(z)=\frac{\sum_{r=0}^{M} b_r z^{-r}}{1-\sum_{k=1}^{N} a_k z^{-k}}=\frac{3(1+z^{-1})(1-3.14z^{-1}+z^{-2})}{(1-0.6z^{-1})(1+0.7z^{-1}+0.72z^{-2})}$$

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import signal

# 差分方程的参数
A = 3
b = np.array([[1, 1, 0], [1, -3.14, 1]])                # 分子
a = np.array([[1, -0.6, 0], [1, 0.7, 0.72]])            # 分母

# 输入信号
M = a.shape[0]
N = 30
delta = signal.unit_impulse(N)                          # 单位样本信号
y = np.ones(N)                                          # 单位阶跃信号

# IIR 级联型滤波器
z1 = np.zeros((M + 1, N))
z2 = np.zeros((M + 1, N))
z1[0, :] = delta
z2[0, :] = y
for i in range(M):                                      # 循环滤波,计算最终结果
    zi = signal.lfilter_zi(b[i, :], a[i, :]) * 0        # 零初始条件
    z1[i + 1, :], _ = signal.lfilter(b[i, :], a[i, :], z1[i, :], zi=zi)
    z2[i + 1, :], _ = signal.lfilter(b[i, :], a[i, :], z2[i, :], zi=zi)

# 绘图
fig, axs = plt.subplots(2, 1, constrained_layout=True)
```

```
axs[0].stem(A * z1[M, :], basefmt = "")
axs[1].stem(A * z2[M, :], basefmt = "")
plt.rcParams['font.sans-serif'] = ['SimHei']          # 用来正常显示中文标签
plt.rcParams['axes.unicode_minus'] = False            # 用来显示负号
axs[0].set_title('IIR级联型h(n)')
axs[1].set_title('IIR级联型y(n)')
plt.show()
fig.savefig('./iir_cas_sequence.png', dpi = 500)
```

运行程序,结果如图 5.2.10 所示。

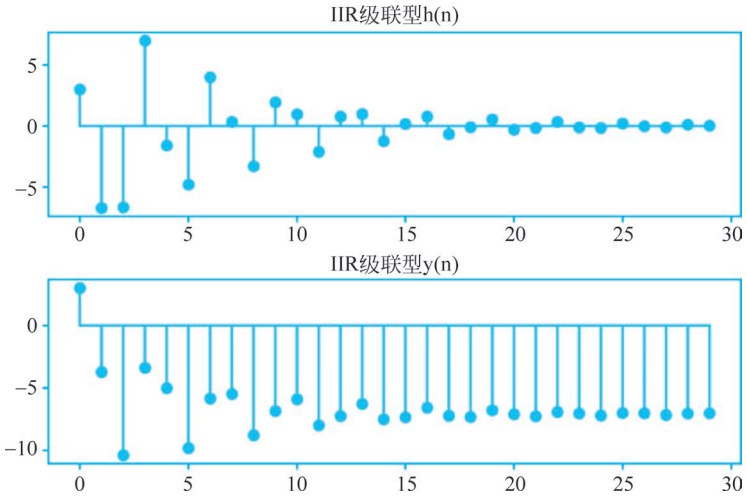


图 5.2.10　例 5.3 单位脉冲响应和单位阶跃响应

5.2.3　并联型

将式(5.2.1)的系统函数 $H(z)$展成部分分式之和,即

$$H(z)=\sum_{k=1}^{N_1}\frac{A_k}{1-g_kz^{-1}}+\sum_{k=1}^{N_2}\frac{B_k(1-e_kz^{-1})}{(1-d_kz^{-1})(1-d_k^*z^{-1})}+\sum_{k=0}^{M-N}G_kz^{-k} \tag{5.2.8}$$

式中,$N_1+2N_2=N$,d_k^* 是 d_k 的共轭复数。由于式(5.2.1)中系数 a_k 和 b_r 是实数,故 A_k、B_k、G_k、g_k 及 e_k 全是实数。如果 $M<N$,则式(5.2.8)中不包含 $\sum_{k=0}^{M-N}G_kz^{-k}$ 项,如果 $M=N$,则 $\sum_{k=0}^{M-N}G_kz^{-k}$ 项变为G_0。一般 IIR 系统皆满足 $M\leqslant N$ 的条件。当 $M=N$ 时,式(5.2.8)成为

$$H(z)=\sum_{k=1}^{N_1}\frac{A_k}{1-g_kz^{-1}}+\sum_{k=1}^{N_2}\frac{r_{0k}+r_{1k}z^{-1}}{1-\alpha_{1k}z^{-1}-\alpha_{2k}z^{-2}}+G_0 \tag{5.2.9}$$

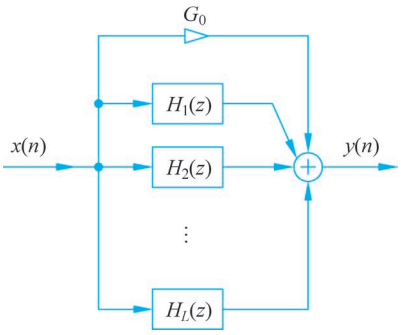
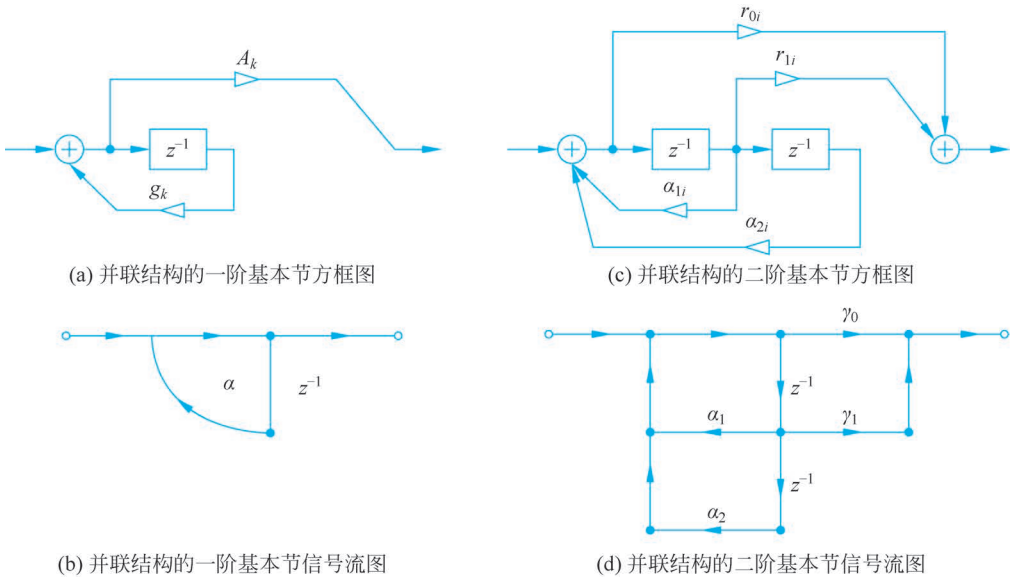


图 5.2.11　并联型结构方框图

总系统函数为各部分系统函数之和时,表示总系统为各相应子系统的并联。因此,式(5.2.9)可以解释为一阶和二阶系统的并联组合,其结构实现如图 5.2.11 所示。每个一阶或二阶子系统可用直接Ⅱ型实现,如

图 5.2.12 所示。也可将式(5.2.9) 中实根部分两两合并，形成二阶分式，以便系统全部采用二阶系统结构。



(a) 并联结构的一阶基本节方框图　　(c) 并联结构的二阶基本节方框图

(b) 并联结构的一阶基本节信号流图　　(d) 并联结构的二阶基本节信号流图

图 5.2.12　直接Ⅱ型实现的并联结构

显然，并联结构运算速度快，也可以单独调整极点位置，但不能像级联型那样直接调整零点，因为并联型各二阶节网络的零点，并非整个系统函数的零点。因此，当要求准确传输零点时，以采用级联型为宜。另外，并联型各基本节的误差互不影响。

基本节构成的并联型信号流图如图 5.2.13 所示。

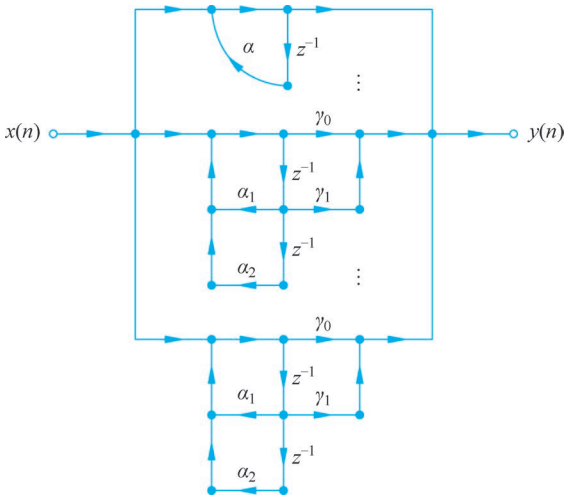


图 5.2.13　基本节构成的并联型信号流图

【例 5.4】　用并联结构实现如下系统函数的 IIR 数字滤波器，并求单位脉冲响应和单位阶跃响应。

$$H(z)=\frac{-13.65-14.81z^{-1}}{1-2.95z^{-1}+3.14z^{-2}}+\frac{32.6-16.37z^{-1}}{1-z^{-1}+0.5z^{-2}}$$

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import signal

# 差分方程的参数
b = np.array([[-13.65, -14.81, 0], [32.6, -16.37, 0]])  # 分子
a = np.array([[1, -2.95, 3.14], [1, -1, 0.5]])          # 分母

# 输入信号
M = a.shape[0]
N = 30
delta = signal.unit_impulse(N)                          # 单位样本信号
y = np.ones(N)                                          # 单位阶跃信号

# IIR 并联型滤波器
z1 = np.zeros((M, N))
z2 = np.zeros((M, N))
for i in range(M):                                      # 分别通过 M 个滤波器,计算最终结果
    zi = signal.lfilter_zi(b[i, :], a[i, :]) * 0        # 零初始条件
    z1[i, :], _ = signal.lfilter(b[i, :], a[i, :], delta, zi=zi)
    z2[i, :], _ = signal.lfilter(b[i, :], a[i, :], y, zi=zi)

# 绘图
fig, axs = plt.subplots(2, 1, constrained_layout=True)
axs[0].stem(np.sum(z1, axis=0), basefmt="")
axs[1].stem(np.sum(z2, axis=0), basefmt="")
plt.rcParams['font.sans-serif'] = ['SimHei']            # 用来正常显示中文标签
axs[0].set_title('IIR 并联型 h(n)')
axs[1].set_title('IIR 并联型 y(n)')
plt.show()
fig.savefig('./iir_par_sequence.png', dpi=500)
```

程序运行结果如图 5.2.14 所示。

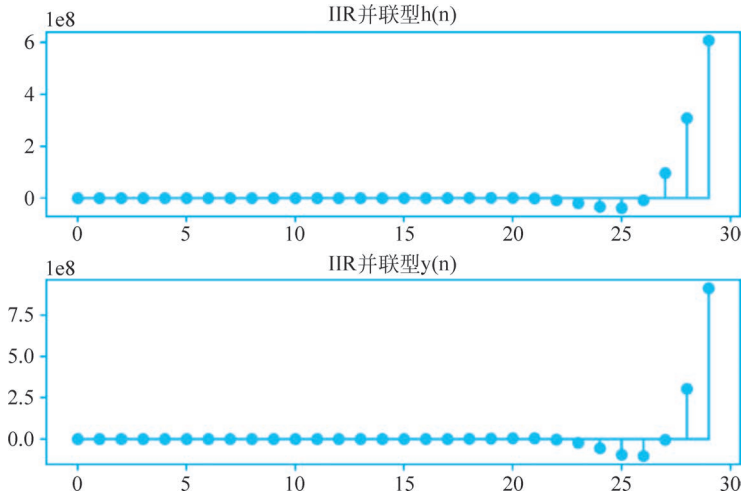


图 5.2.14　例 5.4 单位脉冲响应和单位阶跃响应

除以上 3 种最基本的递归型结构外，还有级并联混合结构、最小乘法结构、把 $H(z)$ 展成连分式实现的梯形结构等。

5.3　有限冲激响应数字滤波器的结构

有限冲激响应数字滤波器的基本结构包括直接型（横截型、卷积型）、级联型、频率采样型、快速卷积型、线性相位的 FIR 数字滤波器结构。

有限冲激响应滤波器的系统函数为

$$H(z)=\sum_{n=0}^{N-1}h(n)z^{-n} \tag{5.3.1}$$

其差分方程为

$$y(n)=\sum_{k=0}^{N-1}h(k)x(n-k) \tag{5.3.2}$$

其基本结构型式有下述几种。

5.3.1　直接型

由式(5.3.2)可得出图 5.3.1 所示的直接型结构，由于式(5.3.2)就是信号的卷积型式，故称为卷积型结构。图 5.3.1 也可看成是图 5.2.4 在各 $a_k=0$ 和 $b_k=h(k)$ 时的特例。

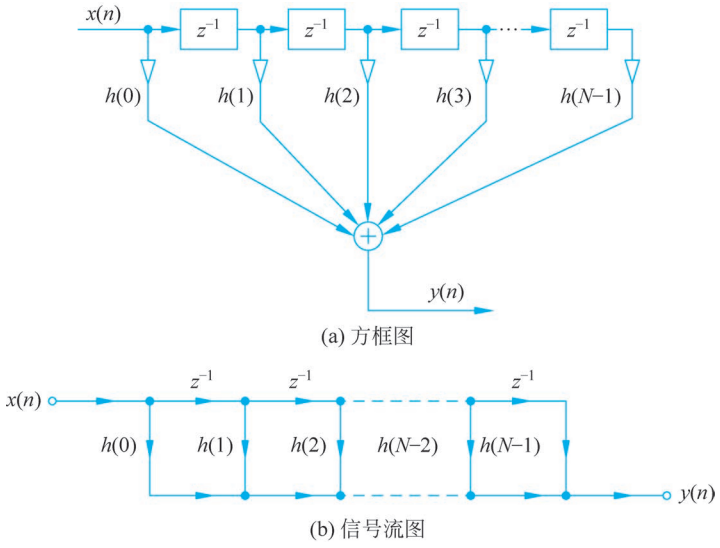
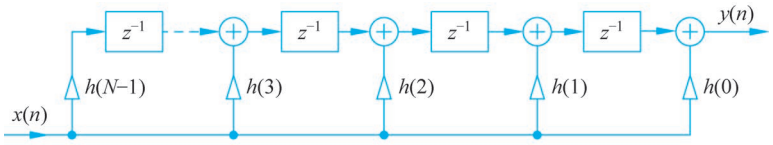


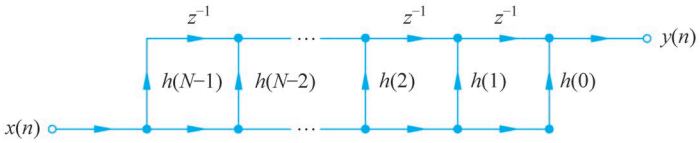
图 5.3.1　FIR 数字滤波器直接型结构

将转置理论应用于图 5.3.1，可得到图 5.3.2 的转置直接型结构。

【例 5.5】 用直接型结构实现如下系统函数的 FIR 数字滤波器，并求单位脉冲响应和单位阶跃响应。



(a) 转置型方框图



(b) 转置型信号流图

图 5.3.2　FIR 数字滤波器转置直接型结构

$$H(z)=\sum_{n=0}^{N-1}h(n)z^{-n}=\sum_{n=0}^{10}0.9^{n}z^{-n}$$

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import signal

# 差分方程的参数
M0 = 11
M = np.arange(0, 11, 1)
a = 1                                                    # 分母
b = np.power(0.9, M)                                     # 分子

# 输入信号
N = 30
delta = signal.unit_impulse(N)                           # 单位样本信号
y = np.ones(N)                                           # 单位阶跃信号

# FIR 直接型滤波器
zi = signal.lfilter_zi(b, a) * 0                         # 零初始条件
z1, _ = signal.lfilter(b, a, delta, zi = zi)
z2, _ = signal.lfilter(b, a, y, zi = zi)

# 绘图
fig, axs = plt.subplots(2, 1, constrained_layout = True)
axs[0].stem(z1, basefmt = "")
axs[1].stem(z2, basefmt = "")
plt.rcParams['font.sans - serif'] = ['SimHei']           # 用来正常显示中文标签
axs[0].set_title('FIR 直接型 h(n)')
axs[1].set_title('FIR 直接型 y(n)')
plt.show()
fig.savefig('./fir_dir_sequence.png', dpi = 500)
```

运行程序，结果如图 5.3.3 所示。

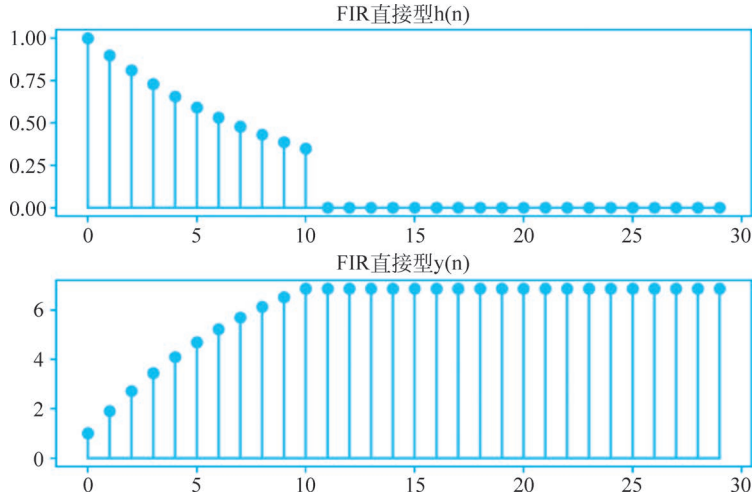


图 5.3.3　例 5.5 单位脉冲响应和单位阶跃响应

5.3.2　级联型

将式(5.3.1)所示的系统函数分解成若干个一阶和二阶多项式的连乘积

$$H(z)=\prod_{k=1}^{M_1}H_{1k}(z)\prod_{k=1}^{M_2}H_{2k}(z)$$

则可构成如图 5.3.4(a)所示的级联型结构。其中，$H_{1k}(z)=a_{0k}^{(1)}+a_{1k}^{(1)}z^{-1}$ 为一阶节；$H_{2k}(z)=a_{0k}^{(2)}+a_{1k}^{(2)}z^{-1}+a_{2k}^{(2)}z^{-2}$ 为二阶节。每个一阶节和二阶节可用图 5.3.1 所示的直接型结构实现。当 $M_1=M_2=1$ 时，即可得图 5.3.4(b)所示的具体结构。这种结构的每节都便于控制零点，在需要控制传输零点时可以采用。但是它所需要的系数 a 比直接型 $h(n)$ 的多，所需要的乘法运算也比直接型多。直接型结构和级联型结构在雷达信号处理中作为相关器和对消器等获得了广泛的应用。

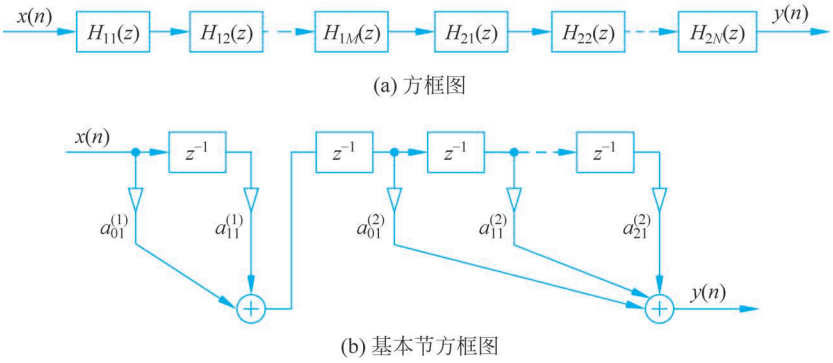


图 5.3.4　FIR 级联型结构

【例 5.6】 用级联型结构实现如下系统函数的 FIR 数字滤波器，并求单位脉冲响应和单位阶跃响应。

$$H(z)=(1+1.72z^{-1}+0.81z^{-2})(1+1.17z^{-1}+0.85z^{-2})$$

```
import numpy as np
import matplotlib.pyplot as plt
from scipy import signal

# 差分方程的参数
A = 1
a = np.array([[1, 0, 0], [1, 0, 0]])                    # 分母
b = np.array([[1, 1.72, 0.81], [1, 1.17, 0.85]])        # 分子

# 输入信号
M = a.shape[0]
N = 30
delta = signal.unit_impulse(N)                          # 单位样本信号
y = np.ones(N)                                          # 单位阶跃信号

# FIR 级联型滤波器
z1 = np.zeros((M + 1, N))
z2 = np.zeros((M + 1, N))
z1[0, :] = delta
z2[0, :] = y
for i in range(M):                                      # 循环滤波,计算最终结果
    zi = signal.lfilter_zi(b[i, :], a[i, :]) * 0        # 零初始条件
    z1[i + 1, :], _ = signal.lfilter(b[i, :], a[i, :], z1[i, :], zi = zi)
    z2[i + 1, :], _ = signal.lfilter(b[i, :], a[i, :], z2[i, :], zi = zi)

# 绘图
fig, axs = plt.subplots(2, 1, constrained_layout = True)
axs[0].stem(A * z1[M, :], basefmt = "")
axs[1].stem(A * z2[M, :], basefmt = "")
plt.rcParams['font.sans - serif'] = ['SimHei']          # 用来正常显示中文标签
axs[0].set_title('FIR 级联型 h(n)')
axs[1].set_title('FIR 级联型 y(n)')
plt.show()
fig.savefig('./fir_cas_sequence.png', dpi = 500)
```

运行程序,结果如图 5.3.5 所示。

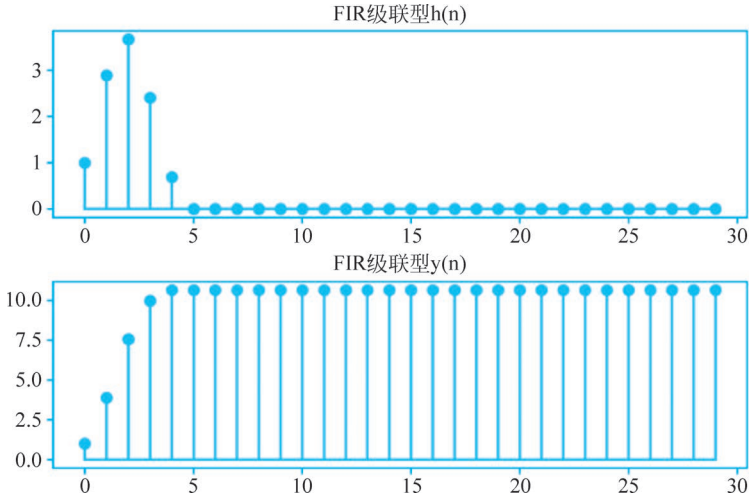


图 5.3.5　例 5.6 单位脉冲响应和单位阶跃响应

5.3.3　线性相位的 FIR 系统网络结构

FIR 系统的最主要特性之一就是它可以构成具有线性相位特性的滤波器。所谓线性相位特性是指滤波器对不同频率的正弦波所产生的相移和正弦波的频率成直线关系。因此，在滤波器通带内的信号通过滤波器后，除了由相频特性的斜率决定的延迟外，可以不失真地保留通带以内的全部信号。

线性相位的 FIR 系统的单位采样响应具有如下特性

$$h(n)=\pm h(N-1-n) \tag{5.3.3}$$

此时式(5.3.1)可分两种情况写出。

当 N 为偶数时

$$\begin{aligned}H(z)&=\sum_{n=0}^{N-1}h(n)z^{-n}\\&=\sum_{n=0}^{\frac{N}{2}-1}h(n)z^{-n}+\sum_{n=\frac{N}{2}}^{N-1}h(n)z^{-n}\\&=\sum_{n=0}^{\frac{N}{2}-1}h(n)z^{-n}+\sum_{n=0}^{\frac{N}{2}-1}h(N-1-n)z^{-(N-1-n)}\end{aligned}$$

将式(5.3.3)代入可得到

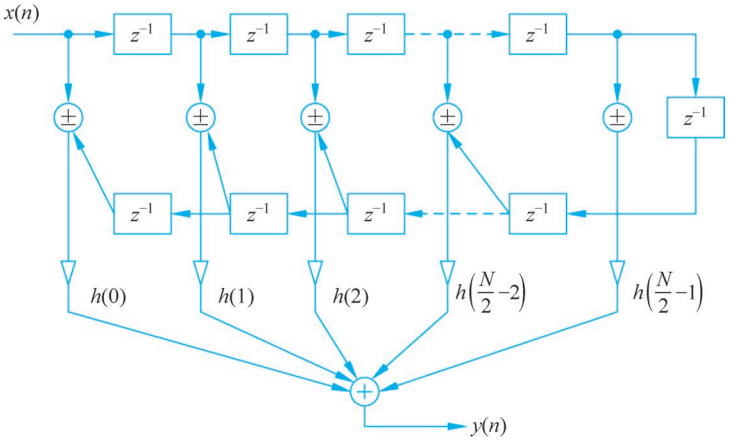
$$H(z)=\sum_{n=0}^{\frac{N}{2}-1}h(n)[z^{-n}\pm z^{-(N-1-n)}] \tag{5.3.4}$$

当 N 为奇数时，可证明

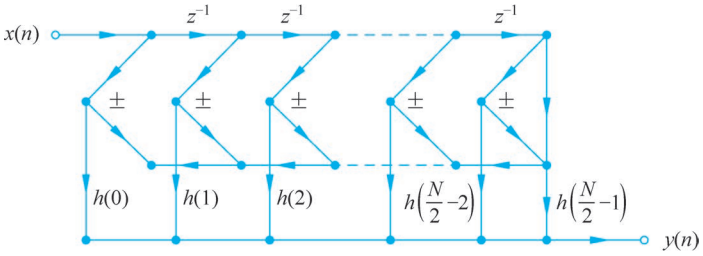
$$\begin{aligned}H(z)&=\sum_{n=0}^{N-1}h(n)z^{-n}\\&=\sum_{n=0}^{\frac{N-1}{2}-1}h(n)[z^{-n}\pm z^{-(N-1-n)}]+h\left(\frac{N-1}{2}\right)z^{-\left(\frac{N-1}{2}\right)}\end{aligned} \tag{5.3.5}$$

式(5.3.4)及式(5.3.5)的 FIR 系统的非递归型实现结构如图 5.3.6 所示。由图可见，仅需$\frac{N}{2}$(N 为偶数时)个或$\frac{N+1}{2}$(N 为奇数时)个乘法运算，而不是如图 5.3.1 所示的需 N 次乘法。在 $h(n)=-h(N-1-n)$(即 $h(n)$为奇对称)，N 为奇数的情况下，$h\left(\frac{N-1}{2}\right)=-h\left(N-1-\frac{N-1}{2}\right)=-h\left(\frac{N-1}{2}\right)$，因此 $h(n)$的中间项 $h\left(\frac{N-1}{2}\right)$必须为零，这时图 5.3.6(c)和图 5.3.6(d)中的 $h\left(\frac{N-1}{2}\right)$为零。也可从图 5.3.6 得到其相应的转置形式。

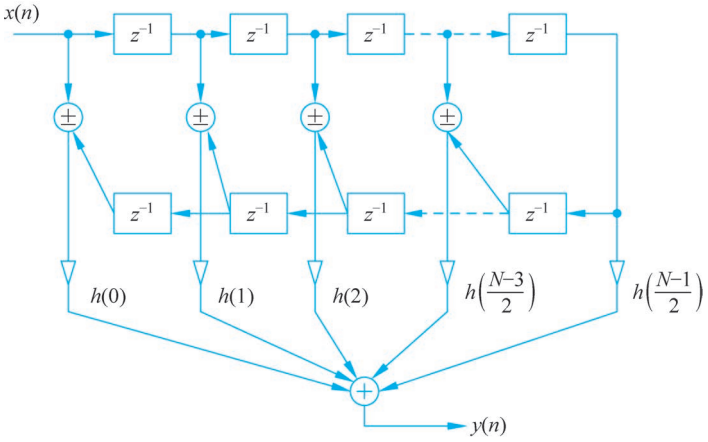
其中，$h(n)$偶对称时取$+1$，$h(n)$奇对称时取-1。



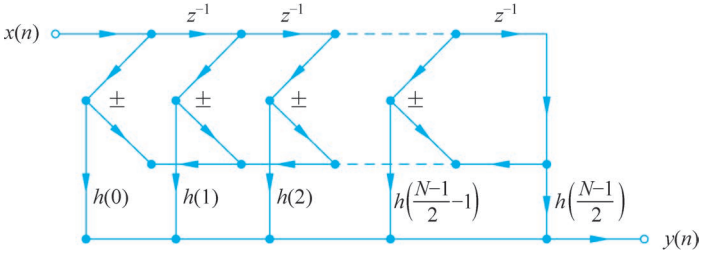
(a) N为偶数时方框图



(b) N为偶数时信号流图



(c) N为奇数时方框图



(d) N为奇数时信号流图

图 5.3.6　线性相位 FIR 数字滤波器直接型结构

5.3.4　频率采样型

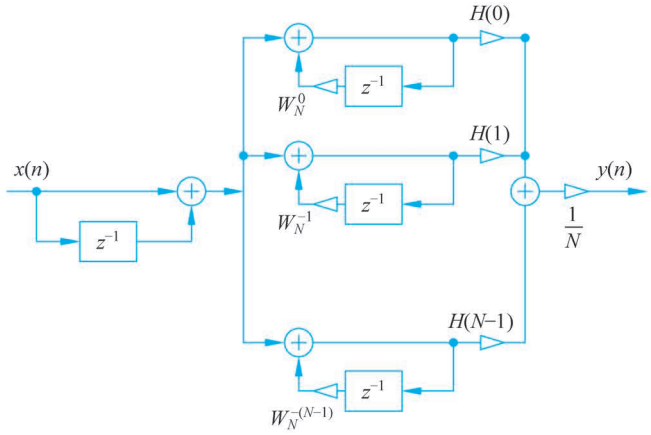
系统函数 $H(z)$ 在单位圆上作 N 等分采样的采样值就是 $h(n)$ 的离散傅里叶变换 $H(k)$

$$H(k)=H(W_N^{-k})=|H(k)|\,\mathrm{e}^{\mathrm{j}\theta(k)}=\sum_{n=0}^{N-1}h(n)\mathrm{e}^{-\mathrm{j}\frac{2\pi}{N}kn}$$

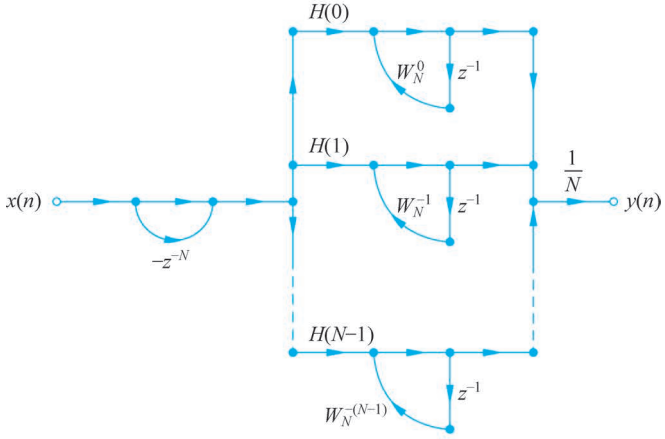
根据式(3.5.5)可得

$$H(z)=(1-z^{-N})\frac{1}{N}\sum_{k=0}^{N-1}\frac{H(k)}{1-W_N^{-k}z^{-1}} \tag{5.3.6}$$

由式(5.3.6)可见，FIR 系统可用一个子 FIR 系统 $(1-z^{-N})$ 和一个子 IIR 系统 $\sum_{k=0}^{N-1}\frac{H(k)}{1-W_N^{-k}z^{-1}}$ 级联实现，如图 5.3.7 所示。



(a) 方框图



(b) 信号流图

图 5.3.7　频率采样型结构

子 FIR 系统 $(1-z^{-N})$ 是一个由 N 节延迟单元组成的梳状滤波器，如图 5.3.8 所示。$(1-z^{-N})$ 在单位圆上有 N 个等分的零点

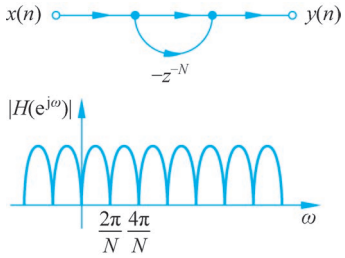


图 5.3.8 梳状滤波器的结构及频率响应幅度

$$1-z^{-N}=0$$

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

梳状滤波器的频率响应

$$H(e^{j\omega})=1-e^{-jN\omega}$$

其幅度特性为

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

子 IIR 系统是 N 个 $\frac{H(k)}{1-W_N^{-k}z^{-1}}$ 型的分式和的形式，所以其实现采用了图 5.3.7 所示的并联结构。每个一阶网络 $\frac{H(k)}{1-W_N^{-k}z^{-1}}$ 在单位圆上有一个极点 $z_k=W_N^{-k}=e^{j\frac{2\pi}{N}k}$，因此网络对频率为 $\omega=\frac{2\pi}{N}k$ 的响应是∞，是一个谐振频率为 $\frac{2\pi}{N}k$ 的无耗谐振器。并联谐振器的极点正好各自抵消一个梳状滤波器的零点，从而使在频率点 $\omega=\frac{2\pi}{N}k$ 处的响应就是 $H(k)$。因此控制滤波器的响应很直接，这正是频率采样型结构的特点。

频率采样型结构有两个问题。第一个问题是所有谐振器的极点都在单位圆上，由系数 W_N^{-k} 决定，当系数量化时，这些极点会移动，因此，系统稳定裕度为零，实际上是不能用的。实践中将所有谐振器的极点设置在半径 r 小于 1 又接近于 1 的圆周上，而子 FIR 系统的零点又需和这些极点重合以互相抵消，故梳状滤波器的零点也移到 r 圆上，如图 5.3.9 所示。

实现修正后的系统函数为

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

其中 $H_r(k)$是修正点上的采样值，因 $r\approx1$，因此

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

故

$$H(z)\approx\frac{(1-r^Nz^{-N})}{N}\sum_{k=0}^{N-1}\frac{H(k)}{1-rW_N^{-k}z^{-1}}\qquad(5.3.7)$$

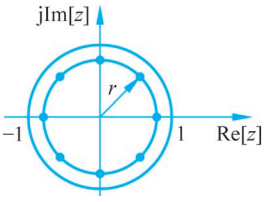


图 5.3.9 采样点改到 r 小于 1 或近似等于 1 的圆上

第二个问题：因 W_N^{-k} 及 $H(k)$都是复数，因此按图 5.3.7 结构实现 FIR 系统需要进行大量的复数运算，比实数运算复杂。但在系统的单位采样响应 $h(n)$为实序列时，可得到局部改善。

由第 3 章讨论的 DFT 奇偶对称特性可知，当时间函数是实函数时，频率函数的模是偶函数，而相角是奇函数，即

$$|H(k)|=|H(N-k)|$$

$$\theta(k)=-\theta(N-k)$$

或者表示为

$$H(k)=H^*(N-k)\quad k=1,2,\cdots,N-1$$

另外，为了使系数为实数，可将谐振器的共轭根合并。若谐振器的根为

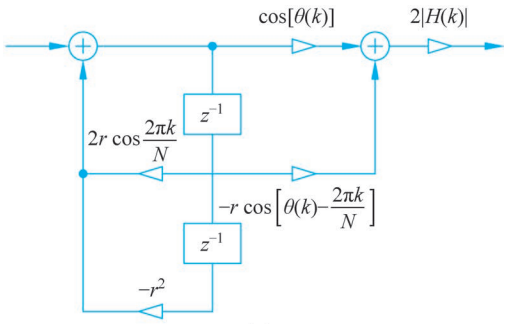
$$z_k=W_N^{-k}=e^{j\frac{2\pi}{N}k}$$

与其对称的根 $W^{-(N-k)}=W_N^k=(W_N^{-k})^*$ 是共轭的。

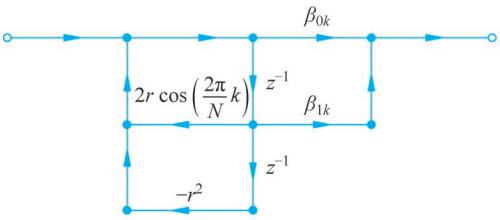
综上所述，可将第 k 个及第 $N-k$ 个谐振器合并为一个二阶网络

$$
\begin{aligned}
H_k(z) &\approx \frac{H(k)}{1-rW_N^{-k}z^{-1}}+\frac{H(N-k)}{1-rW_N^{-(N-k)}z^{-1}} \\
&= \frac{H(k)}{1-rW_N^{-k}z^{-1}}+\frac{H^*(k)}{1-(W_N^{-k})^*z^{-1}} \\
&= \frac{2\mid H(k)\mid\left\{\cos[\theta(k)]-rz^{-1}\cos\left[\theta(k)-\frac{2\pi k}{N}\right]\right\}}{1-2rz^{-1}\cos\frac{2\pi k}{N}+r^2z^{-2}}, \quad 0<k<\frac{N}{2}
\end{aligned}
$$

此系统函数所对应的结构如图 5.3.10 所示，全部都是实数运算。



(a) 方框图



(b) 信号流图

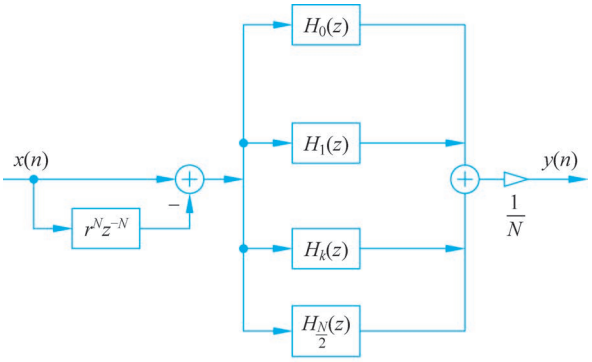
图 5.3.10　二阶网络结构

可以发现，按照上述合并法，图 5.3.7 中的极点位于 W_N^0 的一阶网络 $H_0(z)$，与 N 为偶数时极点位于 $W_N^{-\frac{N}{2}}$ 的一阶网络 $H_{\frac{N}{2}}(z)$合并不了。

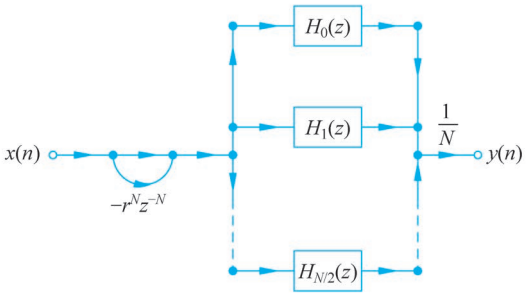
综上所述，改进后的总结构如图 5.3.11 所示。

当 N 为偶数时

$$
H(z)=\frac{(1-r^Nz^{-N})}{N}\left[\frac{H(0)}{1-rz^{-1}}+\frac{H\left(\frac{N}{2}\right)}{1+rz^{-1}}+\sum_{k=1}^{\frac{N}{2}-1}\frac{2\mid H(k)\mid\left\{\cos[\theta(k)]-rz^{-1}\cos\left[\theta(k)-\frac{2\pi k}{N}\right]\right\}}{1-2rz^{-1}\cos\frac{2\pi k}{N}+r^2z^{-2}}\right]
$$



(a) 方框图



(b) 信号流图

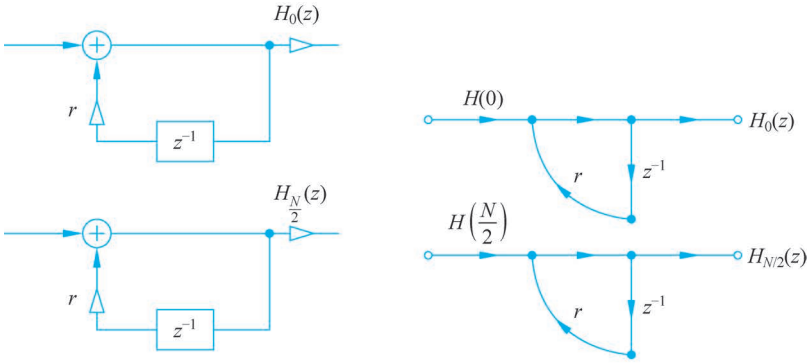
图 5.3.11　修正后的频率采样型结构

其中

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

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

是一阶网络，其结构如图 5.3.12 所示。其余网络都是二阶的。



(a) 方框图　　(b) 信号流图

图 5.3.12　实根一阶网络结构

当 N 为奇数时

$$H(z)=\frac{(1-r^{N}z^{-N})}{N}\left[\frac{H(0)}{1-rz^{-1}}+\sum_{k=1}^{\frac{N}{2}-1}\frac{2\mid H(k)\mid\left\{\cos[\theta(k)]-rz^{-1}\cos\left[\theta(k)-\frac{2\pi k}{N}\right]\right\}}{1-2rz^{-1}\cos\frac{2\pi k}{N}+r^{2}z^{-2}}\right]$$

只有一个 $H_0(z)$一阶网络。

一般来说，频率采样结构比较复杂，所需存储器及乘法器也比较多。但频率采样法也有其优点。由图 5.3.11 可见，每个二阶节均要乘以与频率采样值成比例的数值 $2|H(k)|$，如果这些值中某些是零(比如窄带低通或带通滤波器的情况)，则对应的二阶节就可省去，从而使结构大为简化。另外，它的每个部分都具有很高的规范性，二阶节很多时，设计也并不复杂。

5.4　习题

1. 按照下面所给滤波器系统函数，求出该系统直接Ⅰ型和直接Ⅱ型两种形式的实现方案。

$$H(z)=\frac{3+3.6z^{-1}+0.6z^{-2}}{1+0.1z^{-1}-0.2z^{-2}}$$

2. 给出题 1 系统函数的级联和并联实现方案。

3. 用一阶节和二阶节的级联形式实现下面所给的系统函数。

$$H(z)=\frac{2(z-1)(z^{2}+1.4142136z+1)}{(z+0.5)(z^{2}-0.9z+0.81)}$$

4. 给出题 3 系统函数的并联实现方案。

5. (1) 确定出由下面差分方程所表示的系统的频率响应(幅度与相位)。

$$y(n)=0.5y(n-1)+x(n)+x(n-1)$$

(2) 设采样频率为 1kHz，输入正弦波幅度为 10，频率为 100Hz，求稳态输出。

6. 用可编程序的计算器来对一组测量的随机数据 $x(n)$进行平均处理。当接收到一个测量数据以后，计算器算出这一测量数据与前 3 次测量数据的平均值，并计算出这一运算过程的频率响应。

7. 已知滤波器单位采样响应为

$$h(n)=\begin{cases}0.2^{n}, & 0\leqslant n\leqslant 5\\ 0, & \text{其他 } n\end{cases}$$

求 FIR 数字滤波器直接型结构。

8. 已知 FIR 数字滤波器的 6 个频率采样值为

$$H(0)=12,\qquad H(1)=-3-\mathrm{j}\sqrt{3}$$
$$H(2)=1+\mathrm{j},\qquad H(3)\sim H(13)\text{ 都为 }0$$
$$H(14)=1-\mathrm{j},\qquad H(15)=-3+\mathrm{j}\sqrt{3}$$

计算滤波器的频率采样结构，设选择修正半径 $r=1$(即不修正极点位置)。

9. 已知 FIR 线性系统的系统函数为 $H(z)=\left(1+\frac{1}{2}z^{-1}\right)(1+2z^{-1})\left(1+\frac{1}{4}z^{-1}\right)(1-4z^{-1})$，画出下列每种类型的实现流程图。

(1) 级联型

(2) 直接型

(3) 线性相位型

(4) 频率采样型

10. 试求图 5.4.1 中两个网络的系统函数，并证明它们有相同的极点。

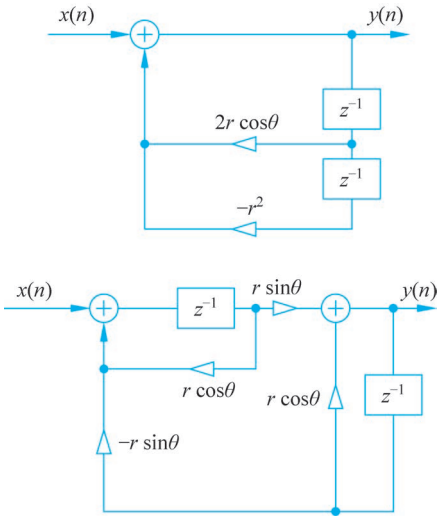


图 5.4.1 题 10 图