E4.12
AO2
SC1
ISE4.7

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2003

MSc and EEE/ISE PART IV: M.Eng. and ACGI

DIGITAL SIGNAL PROCESSING AND DIGITAL FILTERS

Thursday, 1 May 10:00 am

Time allowed: 3:00 hours

There are FIVE questions on this paper.

Answer THREE questions.

Corrected Copy

Any special instructions for invigilators and information for candidates are on page 1.

Examiners responsible First Marker(s) : A.G. Constantinides
Second Marker(s) : T. Stathaki

1. The order N of a Finite Impulse Response filter may be estimated through the empirical formula

$$N = \frac{A}{20}\frac{f_S}{\Delta f}$$

where:

A = passband-to-stopband discrimination in dB

Δf = transition bandwidth in Hz

f_S = sampling frequency in Hz

Justify the above expression. Explain particularly why it is that the order increases as the transition width is reduced. [4]

A single-stage decimator is required to reduce the sampling rate from 32 kHz to 800 Hz. The decimation filter is to be designed as a Finite Impulse Response filter with a cutoff frequency at 300 Hz, a transition frequency at 350 Hz and a passband-to-stopband discrimination of 40 dB. Estimate the order of the required FIR transfer function. [3]

By using the total number of multiplications per second as a measure of computational complexity, determine the computational complexity of this single stage decimation process. The decimator is to be designed as a two-stage structure having a 10:1 decimator as the first stage. Determine the computational complexity of this arrangement and compare it to the single stage realisation. [8]

In a cascade connection of more than two decimators for the above problem how would you order the decimators on the basis of their transition bandwidths? [5]

2. Explain what is meant by terms *computational complexity* and *twiddle factors* in the context of evaluating the Discrete Fourier Transform (DFT). Derive the computational complexity of a N-point DFT. [3]

It is given that $N = N_1 N_2$ with N_1 and N_2 co-prime. It is proposed to carry out on the data array $\{x(n)\}$, $0 \le n \le N-1$, the following 1-D to 2-D mapping

$$n = \langle An_1 + Bn_2 \rangle_N \quad \begin{Bmatrix} 0 \le n_1 \le N_1 - 1 \\ 0 \le n_2 \le N_2 - 1 \end{Bmatrix}$$

$$k = \langle Ck_1 + Dk_2 \rangle_N \quad \begin{Bmatrix} 0 \le k_1 \le N_1 - 1 \\ 0 \le k_2 \le N_2 - 1 \end{Bmatrix}$$

where $\langle M \rangle_N$ means a reduction of the number M modulo N.

Derive the conditions that must prevail on the products AC, BD, AD, and BC in order that all possible twiddle factors in the 2-D DFT computation are eliminated. [10]

Show that the following set of parameters satisfies these conditions

$$A = N_2,\ B = N_1,\ C = N_2 \langle N_2^{-1} \rangle_{N_1},\ D = N_1 \langle N_1^{-1} \rangle_{N_2}$$

where $\langle L^{-1} \rangle_P$ denotes the multiplicative inverse of L evaluated modulo P. [3]

Hence outline the algorithm for the computation of the N-point DFT. [4]

3. In an audio application the structure in Figure 3.1 is used as a “reverberator” to reproduce attenuated versions of an impulse applied as input. Determine its transfer function and show that it is allpass. [3]

Propose an alternative canonic or non-canonic signal flow graph with one multiplier that realises the transfer function. Explain every step in your answer [4]

Determine the impulse response of the structure, and hence the period τ_{rev} during which the impulse response has an absolute value more than 1% of its value at the instant $n=1$. [5]

The reverberator is to be used in an application in which signals are sampled at 10 kHz. It is required to have τ_{rev} =800 msecs. Determine the value for the parameter α and comment on the problems likely to be encountered in an implementation with such a value. [4]

Suggest enhancements to the structure below to produce a more complex reverberation behaviour and justify your answer. [4]

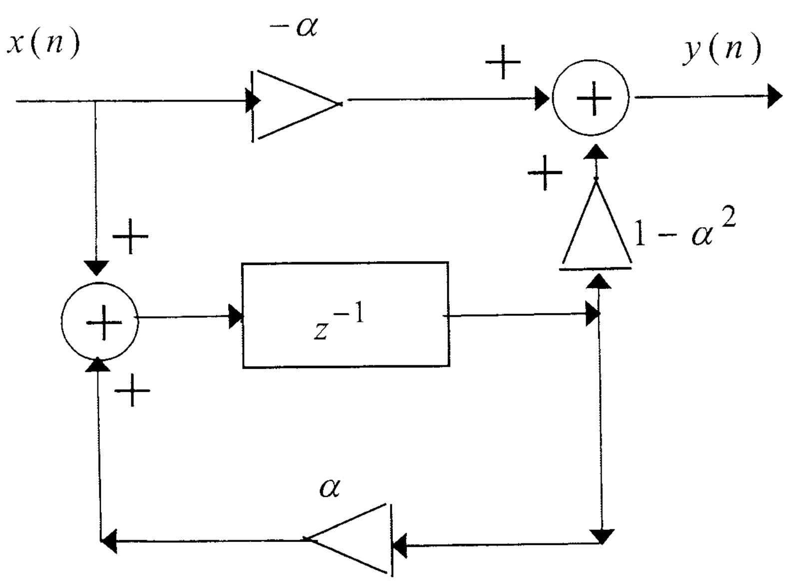


Figure 3.1

4. Two different signals $X_1(z)$ and $X_2(z)$ are transmitted through independent channels in close proximity so that they are received linearly mixed as $Y_1(z) = AX_1(z) + BX_2(z)$ and $Y_2(z) = CX_1(z) + DX_2(z)$ modelled as shown in the Figure 4.1

The mixing matrix is $A = 1 \quad B = H_{12}(z) \quad C = H_{21}(z) \quad D = 1$.

The signals $Y_1(z)$ and $Y_2(z)$ are further processed by a linear system which has a mixing matrix $A = 1 \quad B = -G_{12}(z) \quad C = -G_{21}(z) \quad D = 1$ to produce two new signals $U_1(z)$ and $U_2(z)$ each of which is a function of only one of the original signals $X_1(z)$ and $X_2(z)$.

Determine the two alternative solutions possible and the associated proportionality transfer functions in terms of the mixing parameters above. [12]

It is not known a priori whether the mixing channel transfer functions are minimum phase or not. Comment on possible limitations of one of the alternative solutions [8]

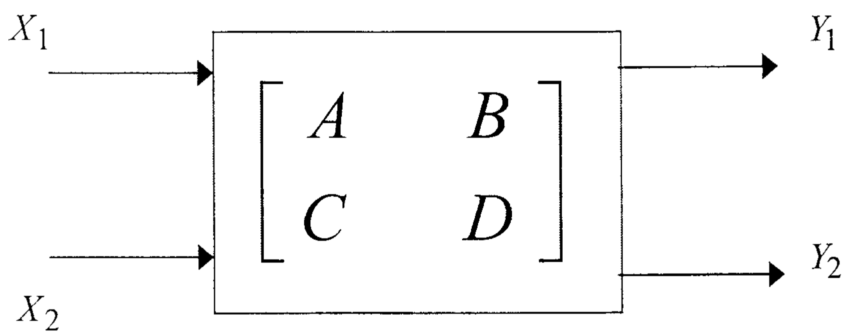


Figure 4.1

5. The transfer function of an ideal real coefficient highpass filter $H_{HP}(z)$ has cutoff frequency θ_p and impulse response $h_{HP}(n)$. Show that $H_{LP}(z) = H_{HP}(-z)$ is a lowpass filter and determine its cutoff frequency. Indicate on the unit circle the correspondence between the passbands of the two filters. Find an expression for the impulse response $h_{LP}(n)$ of the ~~highpass~~ lowpass filter in terms of the impulse response $h_{HP}(n)$ of the original ~~lowpass~~ highpass filter. [5]

In a specific application it is required to split the entire baseband into two equal bands. Indicate how the above relationships can be used to produce a minimal coefficient realisation. [5]

Let be $\theta_p < \dfrac{\pi}{2}$ and form $G(z) = H_{LP}(ze^{j\theta_0}) + H_{LP}(ze^{-j\theta_0})$. Show that $G(z)$ is a real coefficient bandpass filter with a passband centred at θ_0. Determine the impulse response and bandwidth of the bandpass filter in terms of the impulse response of the lowpass filter and the centre frequency θ_0. [5]

Explain why in practice with realisable transfer functions such an approach may not produce a bandpass response [5]

Q1. The fractional transition BW increases the rate of fall off of the frequency response, and for linear phase FIR filters this is clearly related to the "frequency" of the maximum cosine term in the amplitude response. Thus the higher this frequency, the narrower the transition width and hence the relationship. 4

For a single stage decimation we have:–

$$32\text{kHz} \longrightarrow \boxed{D:1\downarrow} \longrightarrow 800\ \text{kHz}$$

the FIR filter has an order $N \cong \frac{40}{20} \cdot \frac{32{,}000}{50}$

or $N = 1{,}280$

The computational complexity is $1{,}280 \times 32 \times 10^3$ or $\approx 41 \times 10^6$ mults/second. 3

As a two-stage operation we have:–

The first stage can have wider transition BW by $10\times$ then $N_1 = \frac{40}{20} \cdot \frac{32{,}000}{500} = 128$

and the associated computational complexity is 4×10^6

The second stage has the required transition BW but at a reduced sampling rate ie 3.2 kHz

Hence $N_2 = \frac{40}{20} \cdot \frac{3{,}200}{50} = 128$ 8

Thus the total computational complexity is 8×10^6 – a gain of $5\times$

The computational complexity can be reduced further by a multistage decimation with the early stages arranged to have higher decimation rates. 5

20

AS Alle

Q2. <u>Computational complexity</u> in DFT is taken to be the total number of complex multiplications required for its computation. (Sometimes implicit symmetries can be used for its reduction).

<u>Twiddle factors</u> are phasing factors in the form $\exp(-j \frac{2\pi}{N_i} k_i n_j)$ between stages that modify partial computations in a multi-stage DFT evaluation. 3

For N-point DFT there are N complex multiplications per point producing a total of $O(N^2)$ multiplications.

For
$$\left.\begin{aligned} n &= \langle An_1 + Bn_2 \rangle_N \\ k &= \langle Ck_1 + Dk_2 \rangle_N \end{aligned}\right\} \quad N = N_1 N_2$$
we have,

$$X(k) = X(\langle Ck_1 + Dk_2 \rangle_N) = \sum_{n_1=0}^{N_1-1} \sum_{n_2=0}^{N_2-1} x(\langle An_1 + Bn_2 \rangle_N) W_N^P$$

where $P = (An_1 + Bn_2)(Ck_1 + Dk_2)$ & $W_N = e^{-j\frac{2\pi}{N}}$

Thus $W_N^P = W_N^{ACn_1k_1} \cdot W_N^{ADn_1k_2} \cdot W_N^{BDn_2k_2} \cdot W_N^{BCn_2k_1}$

For the complete removal of twiddle factors we need

$$\langle AD \rangle_N = 0 \quad \langle BC \rangle_N = 0 \quad \langle AC \rangle_N = N_2$$
$$\text{and} \quad \langle BD \rangle_N = N_1$$ 10

Let $\langle N_1^{-1} \rangle_{N_2} = \alpha$ or $\langle \alpha N_1 \rangle_{N_2} = 1$ or $\alpha N_1 = \beta N_2 + 1$

$\langle N_2^{-1} \rangle_{N_1} = \gamma$ $\quad \gamma N_2 = \delta N_1 + 1$

Then
$$\langle AC \rangle_N = \langle N_2(\delta N_1 + 1) \rangle_N = N_2$$
$$\langle BD \rangle_N = \langle N_1(\beta N_2 + 1) \rangle_N = N_1$$

$\langle AD \rangle_N = \langle N_1 N_2 \langle N_1^{-1} \rangle_{N_2} \rangle_N = 0$ i.e. multiple of $N_1 N_2$ and similarly with $\langle BC \rangle_N$ 3

The algorithm maps a 1-D array $\{x(n)\}$ to a 2-D array. Then a line-by-line followed by a column-by-column N_1 & N_2 point (1D) DFTs (or v.v.) are carried out to return frequency samples to $\{X(\langle Ck_1 + Dk_2 \rangle_N)\}$ 4

All 20 PTS

Q3. Let the output at the intermediate adder be U.
Then in the z-transform domain we can write

$$U(z) = X(z) + \alpha z^{-1} U(z) \quad \text{or} \quad U(z) = X(z) / (1-\alpha z^{-1})$$

At the output adder we have

$$Y(z) = -\alpha X(z) + (1-\alpha^2) . z^{-1} U(z)$$

$$= -\alpha X(z) + (1-\alpha^2) . z^{-1} \frac{X(z)}{1-\alpha z^{-1}}$$

or
$$\frac{Y(z)}{X(z)} = -\alpha + \frac{(1-\alpha^2) z^{-1}}{1-\alpha z^{-1}} = \frac{-\alpha + \alpha^2 z^{-1} + z^{-1} - \alpha^2 z^{-1}}{1-\alpha z^{-1}}$$

Thus

$$H(z) = \frac{Y(z)}{X(z)} = \frac{z^{-1}-\alpha}{1-\alpha z^{-1}} \qquad \text{——— (1)}$$

To show that $H(z)$ is allpass we can write the above as

$$H(z)\Big|_{C} = z^{-1} \frac{(1-\alpha z)}{1-\alpha z^{-1}} = z^{-1} \frac{(1-\alpha z^{-1})^*}{(1-\alpha z^{-1})}\Bigg|_{C:|z|=1}$$

Thus $\quad |H(z)|_{C:|z|=1} = 1.$ 3

From equ(1) we can write

$$Y(z)(1-\alpha z^{-1}) = X(z)(z^{-1}-\alpha)$$

and hence

$$Y(z) = \alpha z^{-1} Y(z) + X(z)(z^{-1}-\alpha) \qquad \text{——— (2)}$$

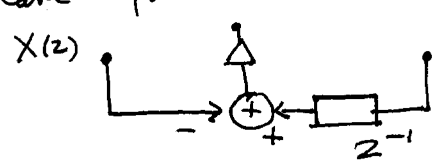
For single multiplier realisation write equ(2) as

$$Y(z) = z^{-1} X(z) + \alpha [z^{-1} Y(z) - X(z)] \qquad \text{—(3)}$$

At this stage we can have many SFGs depending on any additional constraints to be taken into consideration.

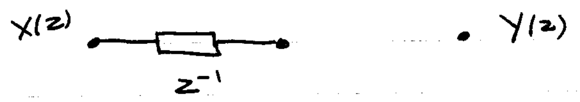
A direct non-canonic realisation is to generate separately the components on the RHS of equ(3).

With the input and output nodes defined we have for the second term

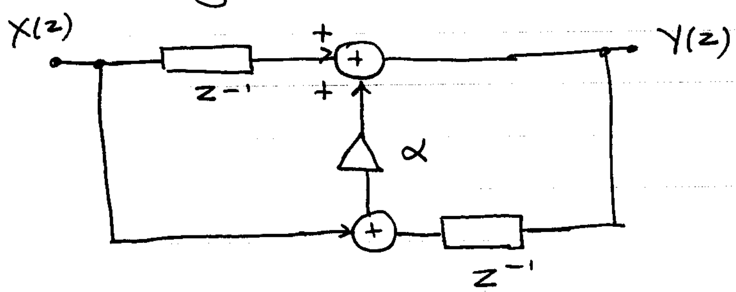


AWP75

while the first term of equ(3) is directly implementable as



On combining the two SFGs we obtain 4



Impulse response $h(n)$

$$h(n) = \mathcal{Z}^{-1}\left[\frac{z^{-1}-\alpha}{1-\alpha z^{-1}}\right]$$

$$= \mathcal{Z}^{-1}\left[\frac{-\alpha}{1-\alpha z^{-1}} + \frac{z^{-1}}{1-\alpha z^{-1}}\right]$$

$$= \mathcal{Z}^{-1}\left[\frac{-\alpha}{1-\alpha z^{-1}}\right] + {}^{-1}\left[\frac{z^{-1}}{1-\alpha z^{-1}}\right]$$

$$= -\alpha\,(\alpha)^{n} + \alpha^{n-1} \qquad n \geq 1$$

$$= -\alpha \qquad n = 0$$

$$= 0 \qquad n < 0$$

When $T_{rev} = 800$ msecs we have $10 \times 800 = 8{,}000$ samples i.e. 5

$$\alpha^{8{,}000}(1-\alpha^2) = \frac{1}{100}(1-\alpha^2)$$

or $\alpha = 0.9994$

Problems of precision are likely to be encountered with such a value of α in a fixed point realisation. 4

A richer impulse response can be obtained when the allpass is of higher order. 4

20

PTS

Q4 For the given system we have

$$\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} 1 & H_{12}(z) \\ H_{21}(z) & 1 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

The new signals U_1 and U_2 are given by

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} 1 & -G_{12}(z) \\ -G_{21}(z) & 1 \end{bmatrix}\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix}$$

By combining the above we have

$$\begin{bmatrix} U_1 \\ U_2 \end{bmatrix} = \begin{bmatrix} 1 & H_{12}(z) \\ H_{21}(z) & 1 \end{bmatrix}\begin{bmatrix} 1 & -G_{12}(z) \\ -G_{21}(z) & 1 \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$= \begin{bmatrix} 1 - H_{12}(z)\, G_{21}(z) & H_{12}(z) - G_{12}(z) \\ H_{21}(z) - G_{21}(z) & 1 - H_{21}(z)\, G_{12}(z) \end{bmatrix}\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

Thus there are two possible outcomes

(A) the main diagonal is zero

ie. $H_{12}(z) - G_{12}(z) = 0$

$H_{21}(z) - G_{21}(z) = 0$

giving

$$U_1 = (1 - H_{12}(z) \cdot G_{21}(z))\, X_1$$

$$U_2 = (1 - H_{21}(z) \cdot G_{12}(z))\, X_2$$

(B) the minor diagonal is zero

ie.

$1 - H_{12}(z) \cdot G_{21}(z) = 0$

$1 - H_{21}(z) \cdot G_{12}(z) = 0$

Hcl PTS

giving

$$U_1 = [H_{12}(z) - G_{12}(z)]\, X_2$$
$$U_2 = [H_{21}(z) - G_{21}(z)]\, X_1$$ 12

The second case requires

$$G_{12}(z) = \frac{1}{H_{21}(z)}$$

$$G_{21}(z) = \frac{1}{H_{12}(z)}$$

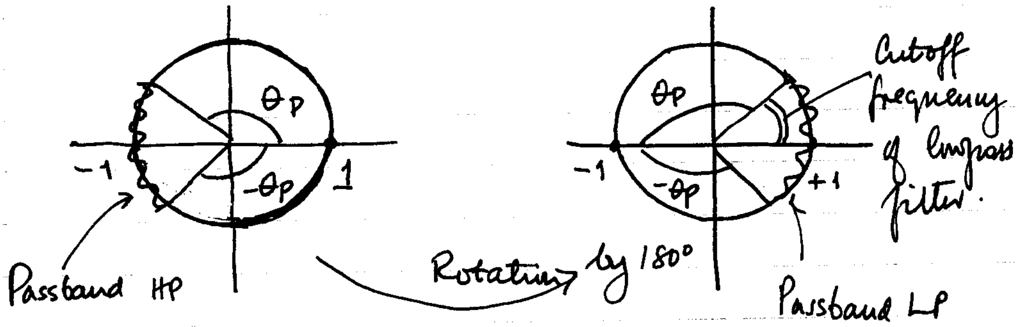
If the mixing matrix (ie channel transfer functions) have zeros on or near the circumference of the unit circle then the dynamic range requirements for $G_{12}(z)$ and $G_{21}(z)$ would be large and may not be practically attainable.

Moreover if $H_{12}(z)$ and $H_{21}(z)$ are non-minimum phase then $G_{12}(z)$ and $G_{21}(z)$ would be unstable and hence the entire scheme unrealisable. 8

20

ACC PTS

Q5 The z-plane sketch for the given $H_{HP}(z)$ is



Rotation by 180° is equivalent to replacing z by -z. From the figures it follows that the cutoff frequency of the lowpass filter θ_c will be such that

$$\theta_c + \theta_p = \pi$$

Let

$$H_{HP}(z) = \sum_{n=-\infty}^{+\infty} h_{HP}(n)\, z^{-n}$$

so that $$H_{LP}(z) = H_{HP}(-z) = \sum_{n=-\infty}^{+\infty} h_{HP}(n) . (-1)^n z^{-n}$$

i.e. $$h_{LP}(n) = (-1)^n . h_{HP}(n)$$ 5

For bandsplitting into two equal bands $\theta_p = \theta_c = \frac{\pi}{2}$ Since alternate coefficients of the impulse are of opposite sign to their corresponding counterparts in the complementary filter, and in absolute values they are all equal we can group even indexed terms together and odd indexed terms together. Their sum would produce one filter while their difference would produce the complementary filter.

$$G(z) = \sum_{n=-\infty}^{+\infty} h_{LP}(n)\, (e^{-j\theta_0} z)^{-n} + \sum_{n=-\infty}^{+\infty} h_{LP}(n)\, (e^{+j\theta_0} z)^{-n}$$ 5

$$= \sum_{n=-\infty}^{+\infty} h_{LP}(n) \left[e^{-jn\theta_0} + e^{+jn\theta_0} \right] . z^{-n}$$

ACC

PTS

ie. the impulse response is given by

$$g(n) = 2h_{LP}(n) \cdot \cos n\theta_0$$

The bandwidth is then equal to

$$BW = (\theta_0 + \theta_c) - (\theta_0 - \theta_c) = 2\theta_c$$ 5

In realisable systems the amplitude response does not necessarily fall off rapidly to zero outside the passband.

Hence the addition of shifted versions of the LP response will produce interactions between the passband and/or stopband of one with the corresponding part of the other. These interactions will combine destructively due to phase relationships. 5

20

AW

PTS