Brunel University

Department of Electronics and Computer Engineering

Final Examination June 2001

EE3052B - MultiMedia Signal Processing

Time allowed 3 Hours

Answer _five_ out of eight questions

Ensure that your registration number is written clearly on the front cover.

Answers

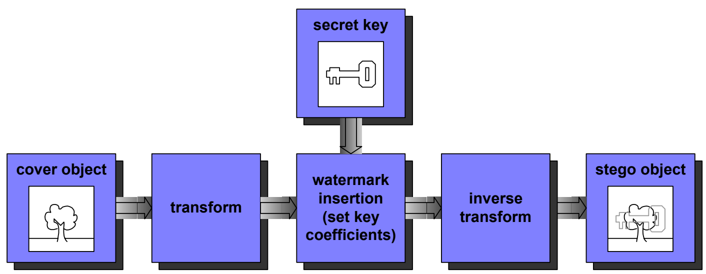
*Note more compact answers that still convey the essence and the essential substance of the materials (plus formulas and numerical answers) will be acceptable.

Q.1

a) Specifically:
   i) Image scanners have prompted copies of many original works to be stored as digital format files.
   ii) Computers allow perfect copies to be made easily.
   iii) The Internet provides a means of widely distributing copies.

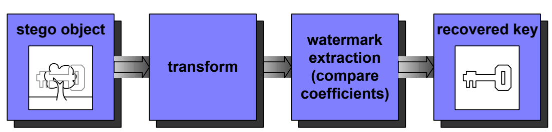
[3 marks]

b) To insert an *invisible complete* watermark:
- The *cover object* (original image) is transformed (e.g. DCT).
- The *secret key* (watermark) is embedded by modifying transform-domain coefficients, so that some relationship is created between the coefficients.
- The inverse transform (e.g. I-DCT) is applied to produce the *stego object* (watermarked image).



[2 marks]

To extract an *invisible complete* watermark:
- The stego object is transformed (e.g. DCT).
- The transform-domain coefficients are compared with each other to establish whether their relationship matches the secret key.



[2 marks]

c) Attacks on electronic watermarks:

i) Four forms of attack:
- Robustness attacks attempt to remove a watermark, but preserve the image content in a useable state.
- Presentation attacks attempt to alter the image content so a watermark cannot be detected or extracted.
- Interpretation attacks attempt to create uncertainty about how to interpret watermark evidence.
- Legal attacks attempt to exploit non-technical details to undermine a watermark.

[2 marks]

ii) Any three of the following
- **Collusion attack**
  Attack: Use multiple copies of the same image watermarked with different information, and average them to produce a composite image.
  Defence: Use highly frequency-selective scheme or always use the same watermark.

- **Inversion attack**
  Attack: Obtain knowledge of how a watermark was inserted, and then reverse the process to remove it.
  Defence: Use a secret key that describes exactly where a watermark is embedded in an image.
- **Pixel modification attack**
  Attack: Modify pixels in an image to remove a watermark directly. e.g. add random noise, modify the intensity histogram or delete whole rows of pixels.
  Defence: Increase strength of watermark (N.B. this also increases watermark detectability and decreases image quality).

- **FMLR (Frequency Mode Laplacian Removal) attack**
  Attack: Apply a Laplacian operator to a watermarked image to create the *negative Laplacian*. Then apply a Laplacian operator to the negative Laplacian to create the *positive Laplacian*. Perform some extra processing to combine the watermarked image and positive and negative Laplacians.
  Defence: Do not use block-based watermarks as the FMLR detects the ripples generated by a watermark.

[2 marks]

d) **Use #1**: Articles on the Internet could contain a content indicator watermark (e.g. an index or indication of the type of web site).
Benefit to site-owner: watermark provides useful information to end user so there is less motivation to try to remove it.
Benefit to surfers: watermarks allows effective filtering of unwanted information (e.g. pornography).

[1 mark]

Use #2: Organisations could bury URL information about their home page inside their logo to allow them to be distinguished from imitation web-sites (e.g. parody-pages).
Benefit to site-owner: It protects the organisation from having its reputation damaged.
Benefit to surfers: It protects users from inferior contents (i.e. guarantees legitimate source).

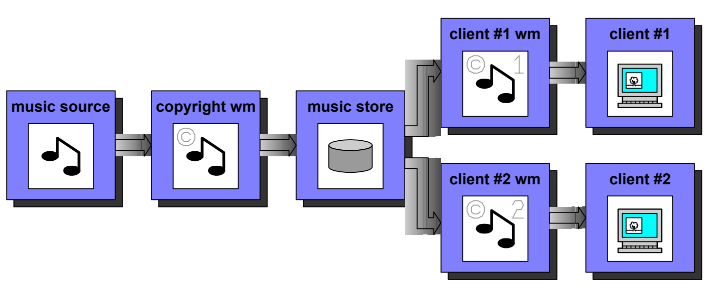
[1 mark]

e) Record companies:

i) ‘I-way robbery’ could be discouraged as follows:

- Computers will be used to read watermarks on music and track down offending web-sites.
- Copyright ownership can be identified by the 1st watermark added at source.
- Source of copyright infringement can be identified from the 2nd watermark added on distribution to clients - a unique mark for each client.

[1 marks]



[2 marks]

ii) Technological problems include:

- Existing archive: There is already an enormous amount of unwatermarked digital music in the world.
- Storage: Watermarking generates data (e.g. secret keys) that need to be securely stored: who stores it, and where, and under what circumstances is it released?

[1 mark]

Legal issues include:

- Jurisdiction: The Internet is international, so who is responsible for enforcing copyright?
- Uncertainty: Only an original image, not a watermarked image, proves ownership, so what happens if a watermark is removed, and replaced with a bogus watermark?
- Defining a match: A watermark may be damaged by attack, so how much must it be affected before a court no longer considers it to match the original?
- No register of marks: No system exists to register marks with a third party, so how does an organisation prove the mark is its own?

[1 mark]

Social concerns and problems include:

- Acceptance: watermarks must not merely work, but be seen to work.
- Abuse: Watermark trackers could identify who is listening to what and misuse the information.
- Ease of use: People will demand a scheme that is easy to use, does not affect the quality of content, and permits the use of content under reasonable licence terms.
- Cheapness beats quality: Most people will accept (perhaps not even notice) the damage caused to music by an attack on a watermark if the pirate-copy is cheaper than the original.

[2 mark]

Q 2.

(a) A stationary process is one whose statistics (mean, variance, etc.) are time-invariant. A nonstationary process is one with at least one time-varying statistic.
Example of a stationary process: a sine wave with constant parameters. Example of a non-stationary process: a sine wave with time-varying parameters, and speech and most types of noise.

[2 marks]

(b) Autocorrelation and power spectrum are Fourier transform pairs

$$P_{xx}(f) = \sum_{m=-\infty}^{\infty} r_{xx}(m)\, e^{-j\omega m}$$ [1 mark]

The relationship between autocorrelation, power and variance

acf[0] = power = variance [1 mark]

(i) Correlation and power spectrum of a unit amplitude discrete-time impulse

$r_{\text{impulse}}(m) = \delta(m)$ $\qquad P_{impulse}(f) = \sum_{m=-\infty}^{\infty} \delta(m)\, e^{-j\omega m} = 1$ [2 marks]

(ii) Correlation and power spectrum of a white noise of unit variance.

$r_{\text{white\_noise}}(m) = \delta(m)$ $\qquad P_{white\_noise}(f) = \sum_{m=-\infty}^{\infty} \delta(m)\, e^{-j\omega m} = 1$ [2 marks]

(c) Information and random processes are inseparable and probability is used to model random processes. For a discrete-time process probability provides a numerical measure, between 0 and 1, of likelihood of an event. Similarly probability can be used to provide a measure of information content of an event in the range 0 (for a certain event that has probability of 1) and some maximum positive value. It turns out that log probability is a function that satisfies the requirements for an information model. The base 2 of logarithm reflects the binary nature of information.

[3 marks]

(d) Entropy is a measure of randomness and gives the lower bound on the number of bits required to encode an information source

$$H(X) = -\sum_{i=1}^{N} P(x_i)\log_2 P(x_i)$$

[3 marks]

(e) Speech is based on the use of about 40 basic acoustic symbols, known as phonemes (or phonetic units), these are used to construct words, sentences etc. Assuming that all phonetic units are equi-probable, and that the average speaking rate is 120 words per minute and that the average phone has 4 phonemes per word, calculate the minimum number of bits per second required to encode speech at the average speaking rate.

$$P(x_{\mathrm{i}})=1/40$$

$$H(X) = -\sum_{i=1}^{N} P(x_i)\log_2 P(x_i) = -\sum_{i=1}^{40} \frac{1}{40}\log_2 \frac{1}{40} = 5.32\,bits$$

Average number of phonemes per second=120*4/60=8 phonemes per second. And 8×5.32=42.4 bits per second. [6 marks]

Q. 3

(a) The source of echo in landline telephony is the impedance mismatch at hybrid of a 2 wire to 4 wire telephone line connection at the exchange.

[1 mark]

The source of acoustic echo is the acoustic leakage (feedback) coupling between the speaker and the microphone.

[1 mark]

(i)

The effect of short-delay echo is to 'enliven' the voice, where as long echo is can be disruptive. Limit of tolerable echo is about 150 ms.

[1 mark]

(ii)

The speed of sound is about 330 ms. Time delay in travelling 10 ms is negligible. The main delay is due to segment length delay in the speech coder which is about 20 ms and the network delay. If the speech goes through a satellite then a round trip delay of up to 600 ms may be added to the echo which will be disruptive.

[2 marks]

(b) An echo suppresser is primarily a detector switch that lets the speech through during the speech-active periods and attenuates the echo during speech-inactive periods. An echo suppresser is controlled by a speech/echo detection device. The echo detector monitors the signal levels on the incoming and outgoing lines, and decides if the signal on a line from, say, speaker B to speaker A is the speech from the speaker B to the speaker A, or the echo of speaker A. If the echo detector decides that the signal is an echo then the signal is heavily attenuated. The performance of an echo suppresser depends on the accuracy of the echo/speech classification subsystem. Echo of speech often has a smaller amplitude level than the speech signal, but otherwise it has mainly the same spectral characteristics and statistics as those of the speech. Therefore the only basis for discrimination of speech from echo is the signal level. As a result, the speech/echo classifier may wrongly classify and let through high-level echoes as speech, or attenuate low-level speech as echo. The performance of an echo suppresser depends on the time delay of the echo. In general, echo suppressers perform well when the round-trip delay of the echo is less than 100 ms. For a conversation routed via a geostationary satellite the round-trip delay may be as much as 600 ms. Such long delays can

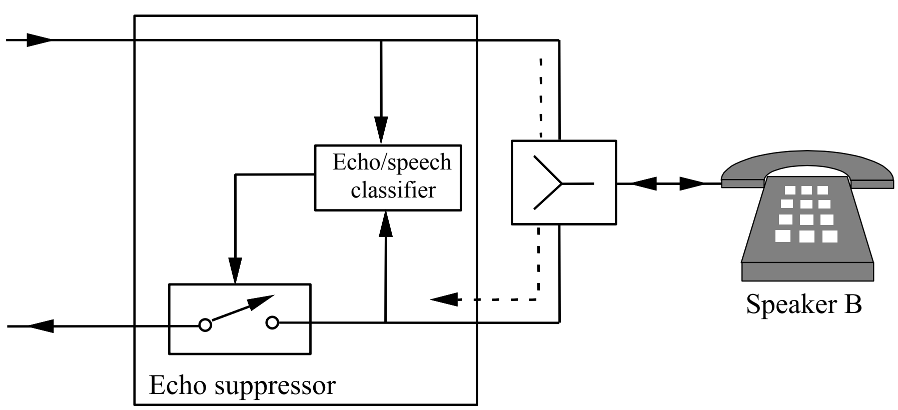


Figure 1 Block diagram illustration of an echo suppression system.

[2 marks]

change the pattern of conversation and result in a significant increase in speech/echo classification errors.

[3 marks]

(c) The figure shows the block diagram of an echo canceller. The adaptive filter models the echo path and attempts to synthesis a replica of the echo which is then subtracted from the echo signal.

(i) The echo canceller can be an infinite impulse response (IIR) or a finite impulse response (FIR) filter. The main advantage of an IIR filter is that a long-delay echo can be synthesised by a relatively small number of filter coefficients. In practice, echo cancellers are based on FIR filters. This is mainly due to the practical difficulties associated with the adaptation and stable operation of adaptive IIR filters.

(ii) The adaptation algorithm can be RLS or LMS. LMS is computationally less expensive than RLS, but RLS converges much faster and has less error at convergence.

LMS is usually chosen for its computationally inexpensive advantage.

[3 marks]

(d) A subband echo cancellation (with R bands) has three advantages

(i) *Reduction in filter length*. The length of the FIR filter for each down-sampled sub-band is $1/R$ of the full band filter.

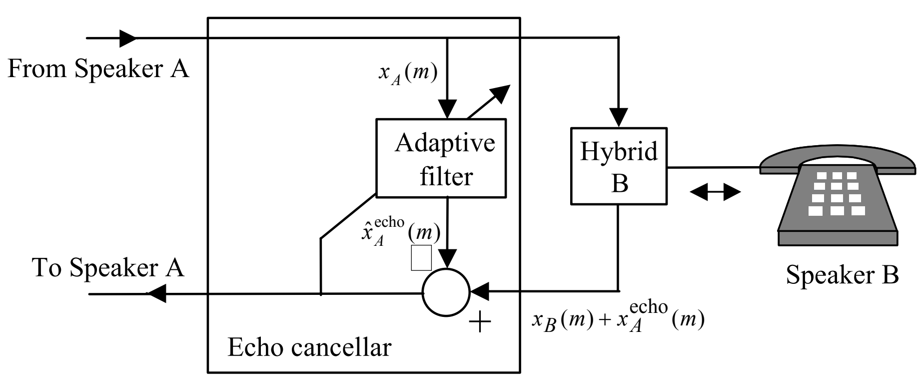
[1 mark]

(ii) *Reduction in computational complexity*. The computational complexity of an LMS-type adaptive filter depends on the product of the filter length and the sampling rate. As for each sub-band, the number of samples per second and the filter length decrease with $1/R$, it follows that the computational complexity of each sub-band filter is $1/R^2$ of that of the full band filter. Hence the overall gain in computational complexity of a sub-band system is R^2/N of the full band system.

[2 marks]

(iii) *Speed of convergence*. The speed of convergence depends on both the filter length and the eigenvalue spread of the signal. The speed of convergence increases with the decrease in the length of the FIR filter for each sub-band and the decrease in the eigen value spread.

[2 marks]



Block diagram illustration of an adaptive echo cancellation system.

[2 marks]

Q4.

(a)

(i) Given a noisy signal modelled as

$$y(m) = x(m) + n(m)$$

the noisy signal Equation in frequency domain is

$$Y(f) = X(f) + N(f)$$

[1 mark]

(ii) The Wiener filter is based on the minimum mean squared optimisation criterion. The filter is derived via minimising the difference between the filter output and a desired signal that is correlated with the input.

[1 mark]

Prove that in frequency domain the Wiener filter is given by

$$W(f) = \frac{P_{xx}(f)}{P_{xx}(f) + P_{nn}(f)}$$

where $P_{xx}(f)$ and $P_{nn}(f)$ are the signal and the noise power spectra.

In the frequency domain, the Wiener filter output $\hat{X}(f)$ is the product of the input signal $Y(f)$ and the filter frequency response $W(f)$:

$$\hat{X}(f) = W(f)Y(f)$$

[1 mark]

The estimation error signal $E(f)$ is defined as the difference between the desired signal $X(f)$ and the filter output $\hat{X}(f)$,

$$\begin{aligned} E(f) &= X(f) - \hat{X}(f) \\ &= X(f) - W(f)Y(f) \end{aligned}$$

[1 marks]

and the mean square error at a frequency f is given by

$$\mathcal{E}\left[\left|E(f)\right|^2\right] = \mathcal{E}\left[\left(X(f) - W(f)Y(f)\right)^* \left(X(f) - W(f)Y(f)\right)\right]$$

[2 marks]

where $\mathcal{E}[\cdot]$ is the expectation function, and the symbol * denotes the complex conjugate.

To obtain the least mean square error filter we set the complex derivative of thw mean squared error Equation with respect to filter $W(f)$ to zero

$$\frac{\partial \mathcal{E}[|E(f)|^2]}{\partial W(f)} = 2W(f)P_{YY}(f) - 2P_{XY}(f) = 0$$

[2 marks]

where $P_{YY}(f) = \mathcal{E}[Y(f)Y^*(f)]$ and $P_{XY}(f) = \mathcal{E}[X(f)Y^*(f)]$ are the power spectrum of $Y(f)$, and the cross-power spectrum of $Y(f)$ and $X(f)$ respectively. The least mean square error Wiener filter in the frequency domain is given as

$$W(f)=\frac{P_{XY}(f)}{P_{YY}(f)}$$

Since the signal and noise are uncorrelated $P_{XY}(f)=P_{XX}(f)$ and $P_{YY}(f)=P_{XX}(f)+P_{NN}(f)$

$$W(f) = \frac{P_{xx}(f)}{P_{xx}(f) + P_{nn}(f)}$$

[2 marks]

(b)

Communication channel distortions may be modelled by a combination of a linear filter and an additive random noise source as shown above. The input/output signals of a linear time invariant channel can be modelled as

$$y(m)=\sum_{k=0}^{P-1} h_k\, x(m-k)+n(m)$$ [1 mark]

where $x(m)$ and $y(m)$ are the transmitted and received signals, $[h_k]$ is the impulse response of a linear filter model of the channel, and $n(m)$ models the channel noise. In the frequency domain the channel output is

$$Y(f)=X(f)H(f)+N(f)$$ [1 mark]

where $X(f)$, $Y(f)$, $H(f)$ and $N(f)$ are the signal, noisy signal, channel and noise spectra respectively. To remove the channel distortions, the receiver is followed by an equaliser. The equaliser input is the distorted channel output, and the desired signal is the channel input. Using the frequency domain Wiener filter, the Wiener equaliser in the frequency domain is given by

$$W(f)=\frac{P_{XX}(f)H^*(f)}{P_{XX}(f)|H(f)|^2+P_{NN}(f)}$$

where it is assumed that the channel noise and the signal are uncorrelated. In the absence of channel noise, $P_{NN}(f)=0$, and the Wiener filter is simply the inverse of the channel filter model $W(f)=H^{-1}(f)$.

[6 marks]

A practical method of training a channel equaliser is to send a known binary sequence at periodic preset intervals to the receiver. The known sequence and the received sequence can be used to derive the Wiener filter equation.

[2 marks]

Q5.

(a)

Adaptation is affected by the choice of the filter type, the filter length, the use of fullband or subband filters and the choice of adaptation method.

[2 marks]

All these methods are based on least mean squared optimisation. The Wiener and RLS are similar the main difference being that RLS a recursive version of Wiener filter. The Steepest descent and LMS use gradient search optimisation methods. The LMS uses the gradient of the instantaneous squared error.

[2 marks]

(b) The steepest-descent adaptation method can be expressed as

$$\boldsymbol{w}(m+1) = \boldsymbol{w}(m) + \mu \left[-\frac{\partial \mathcal{E}[e^2(m)]}{\partial \boldsymbol{w}(m)} \right] \qquad 5.1$$

The gradient of the mean square error function is given by

$$\frac{\partial \mathcal{E}[e^2(m)]}{\partial \boldsymbol{w}(m)} = -2\boldsymbol{r}_{yx} + 2\boldsymbol{R}_{yy}\boldsymbol{w}(m) \qquad 5.2$$

Hence

$$\boldsymbol{w}(m+1) = \boldsymbol{w}(m) + \mu \left[\boldsymbol{r}_{yx} - \boldsymbol{R}_{yy}\boldsymbol{w}(m) \right] \qquad 5.3$$

Let $\boldsymbol{w}_O$ denote the optimal LSE filter coefficient vector, we define a filter coefficients error vector $\tilde{\boldsymbol{w}}(m)$ as

$$\tilde{\boldsymbol{w}}(m) = \boldsymbol{w}(m) - \boldsymbol{w}_o \qquad 5.4$$

For a stationary process, the optimal LSE filter $\boldsymbol{w}_O$ is obtained from the Wiener filter as

$$\boldsymbol{w}_o = \boldsymbol{R}_{yy}^{-1}\boldsymbol{r}_{yx} \qquad 5.5$$

Subtracting $\boldsymbol{w}_O$ from both sides of Equation (5.3), and then substituting $\boldsymbol{R}_{yy}\boldsymbol{w}_o$ for $\boldsymbol{r}_{yx}$ yields

$$\tilde{\boldsymbol{w}}(m+1) = \left[\mathbf{I} - \mu\boldsymbol{R}_{yy} \right] \tilde{\boldsymbol{w}}(m) \qquad 5.6$$

[2 marks]

It is desirable that the filter error vector $\tilde{\boldsymbol{w}}(m)$ vanishes as rapidly as possible. The correlation matrix can be expressed in terms of the matrices of eigenvectors and eigenvalues as

$$\boldsymbol{R}_{yy} = \boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}} \qquad 5.7$$

Substituting $\boldsymbol{R}_{yy}$ in Equation 5.6 yields

$$\tilde{\boldsymbol{w}}(m+1) = \left[\mathbf{I} - \mu\boldsymbol{Q}\boldsymbol{\Lambda}\boldsymbol{Q}^{\mathrm{T}} \right] \tilde{\boldsymbol{w}}(m) \qquad 5.8$$

[2 marks]

Multiplying both sides of Equation 5.8 by $\boldsymbol{Q}^{\mathrm{T}}$ and using the relation $\boldsymbol{Q}^{\mathrm{T}}\boldsymbol{Q} = \boldsymbol{Q}\boldsymbol{Q}^{\mathrm{T}} = \mathbf{I}$ yields

$$\boldsymbol{Q}^{\mathrm{T}}\tilde{\boldsymbol{w}}(m+1)=[\mathbf{I}-\mu\boldsymbol{\Lambda}]\boldsymbol{Q}^{\mathrm{T}}\tilde{\boldsymbol{w}}(m) \qquad 5.9$$

Let $$\boldsymbol{v}(m)=\boldsymbol{Q}^{\mathrm{T}}\tilde{\boldsymbol{w}}(m) \qquad 5.10$$

Then $$\boldsymbol{v}(m+1)=[\mathbf{I}-\mu\boldsymbol{\Lambda}]\,\boldsymbol{v}(m) \qquad 5.11$$

[2 marks]

(c)

Calculate the eigenvalue spread. Eigenvalue spread $= \dfrac{\lambda_{\max}}{\lambda_{\min}} = 10$

(i) The bounds on adaptation step size.

$$0<\mu<\frac{2}{\lambda_{\max}}=1$$

[2 marks]

(ii) The decay factor of the error equations for the fastest and the slowest converging coefficients of the filter given that the adaptation stepsize is 0.4.

$$v_{\max}(m+1)=(1-\mu\lambda_{\max})v_{\max}(m)$$
$$v_{\min}(m+1)=(1-\mu\lambda_{\min})v_{\min}(m)$$

Max decay factor= 1-0.4x2=0.2, Min decay factor =1-0.4x0.2=0.92

[3 marks]

(d) The LMS adaptation method is defined as

$$\boldsymbol{w}(m+1)=\boldsymbol{w}(m)+\mu\left(-\frac{\partial e^2(m)}{\partial \boldsymbol{w}(m)}\right)$$

where the error signal $e(m)$ is given by

$$e(m)=x(m)-\boldsymbol{w}^{\mathrm{T}}(m)\boldsymbol{x}(m)$$

The instantaneous gradient of the squared error can be re-expressed as

$$\frac{\partial e^2(m)}{\partial \boldsymbol{w}(m)}=\frac{\partial}{\partial \boldsymbol{w}(m)}[x(m)-\boldsymbol{w}^{\mathrm{T}}(m)\boldsymbol{y}(m)]^2$$
$$=-2\boldsymbol{y}(m)[x(m)-\boldsymbol{w}^{\mathrm{T}}(m)\boldsymbol{y}(m)]^2$$
$$=-2\boldsymbol{y}(m)e(m)$$

[3 marks]

Substituting the above equation into the recursion update equation of the filter parameters, yields the LMS adaptation equation:

$$\boldsymbol{w}(m+1)=\boldsymbol{w}(m)+\mu\,[\boldsymbol{y}(m)e(m)]$$

At the point of convergence the steepest descent converges to a constant value, whereas the LMS fluctuates depending on the values of the step size.

[2 marks]

Q. 6

(a)

(i) The equation of a Pth order linear predictor

$$x(m) = a_1 x(m-1) + a_2 x(m-2) + \cdots + a_P x(m-P) + e(m)$$

[1 mark]

(ii) The "best" predictor coefficients are obtained by minimising a mean square error criterion defined as

$$\mathcal{E}[e^2(m)] = \mathcal{E}\left[\left(x(m) - \sum_{k=1}^{P} a_k x(m-k)\right)^2\right]$$

$$= \mathcal{E}[x^2(m)] - 2\sum_{k=1}^{P} a_k \mathcal{E}[x(m)x(m-k)] + \sum_{k=1}^{P} a_k \sum_{j=1}^{P} a_j \mathcal{E}[x(m-k)x(m-j)]$$

$$= r_{xx}(0) - 2\boldsymbol{r}_{xx}^{\mathrm{T}}\boldsymbol{a} + \boldsymbol{a}^{\mathrm{T}}\boldsymbol{R}_{xx}\boldsymbol{a}$$

[4 marks]

The mean square prediction error with respect to the predictor coefficient vector $\boldsymbol{a}$ is given by

$$\frac{\partial}{\partial \boldsymbol{a}} \mathcal{E}[e^2(m)] = -2\boldsymbol{r}_{xx}^{\mathrm{T}} + 2\boldsymbol{a}^{\mathrm{T}}\boldsymbol{R}_{xx}$$ [1 marks]

The least mean square error solution is

$$\boldsymbol{R}_{xx}\boldsymbol{a} = \boldsymbol{r}_{xx}$$

From Equation the predictor coefficient vector is given by

$$\boldsymbol{a} = \boldsymbol{R}_{xx}^{-1}\boldsymbol{r}_{xx}$$ [2 marks]

Equation may also be written in an expanded form as

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_P \end{pmatrix} = \begin{pmatrix} r_{xx}(0) & r_{xx}(1) & r_{xx}(2) & \cdots & r_{xx}(P-1) \\ r_{xx}(1) & r_{xx}(0) & r_{xx}(1) & \cdots & r_{xx}(P-2) \\ r_{xx}(2) & r_{xx}(1) & r_{xx}(0) & \cdots & r_{xx}(P-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_{xx}(P-1) & r_{xx}(P-2) & r_{xx}(P-3) & \cdots & r_{xx}(0) \end{pmatrix}^{-1} \begin{pmatrix} r_{xx}(1) \\ r_{xx}(2) \\ r_{xx}(3) \\ \vdots \\ r_{xx}(P) \end{pmatrix}$$

(iii)

Taking the z-transform of the linear prediction model results in an all-pole digital filter with z-transfer function

$$H(z) = \frac{X(z)}{E(z)} = \frac{1}{1 - \sum_{k=1}^{P} a_k z^{-k}}$$

[2 marks]

(b)

(i) A second order linear prediction model for the process,

$$x(m) = a_1 x(m-1) + a_2 x(m-2) + e(m)$$

To obtain the coefficients multiply both sides by $x(m-1)$ and take expectation to obtain

$$r(1) = a_1 r(0) + a_2 r(1)$$

Multiply both sides by $x(m-2)$ and take expectation to obtain

$$r(2) = a_1 r(1) + a_2 r(0)$$

From these two correlation equations we can obtain the predictor coefficients as

$$a_2 = -0.81 \quad a_1 = 0.9$$

[4 marks]

the polar form coefficients can be obtained as

pole radius $a_2 = -r^2 = -0.81$, $r = 0.9$,

pole angle $2r\cos(\theta) = a_1 = 0.9$ $\theta = 60$

[2 marks]

(ii)

Transfer function is given by

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

and the frequency response is

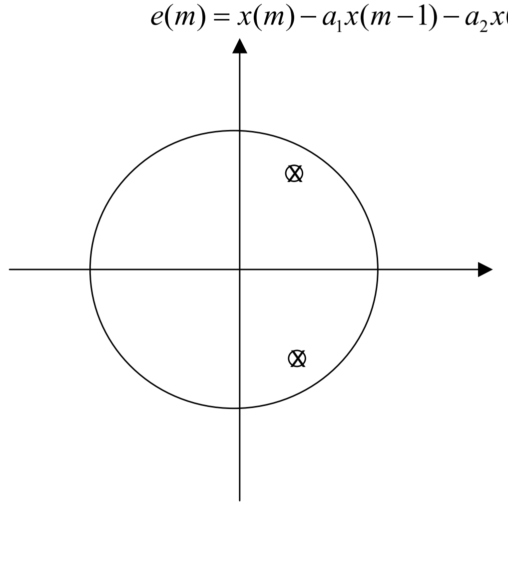
$$H(\omega) = \frac{1}{1 - 0.9e^{-j\omega} + 0.81e^{-j2\omega}}$$

[2 marks]

(iii)

The inverse filter equation is given by

$$e(m) = x(m) - a_1 x(m-1) - a_2 x(m-2)$$



[2 Marks]

Q. 7

(a) A HiFi audio signal has its energy up to a highest frequency of 20 kHz. It is required that after quantisation the signal should have a minimum quatisation to signal to noise ratio of 90 dB.

(i) The minimum required sampling rate.

LPF cutoff=20 kHz,

The minimum sampling rate is therefore 40 kHz. We can use the Hi Fi audio standard rate

Sampling rate=44100,

(ii)The minimum number of bits required.

No of bits using the 6dB per bit rule gives 16 bits,

(iii) How many minutes of a stereo audio signal can be stored on compact Disc with a capacity of 650 M.

Recording time =650,000,000/(2*44100*16*60)= 61 minutes

[3 marks]

(b)

(i)

The sampling rate can be increased by a factor of I through interpolation of I–1 samples between every two samples of $x(m)$. Consider the zero-inserted signal $x_Z(m)$ obtained by inserting I–1 zeros between every two samples of $x(m)$ and expressed as

$$x_z(m) = \begin{cases} x\left(\dfrac{m}{I}\right), & m = 0, \pm I, \pm 2I, \ldots \\ 0, & \text{otherwise} \end{cases}$$

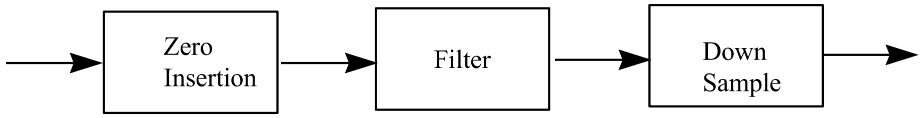
The spectrum of the zero-inserted signal is related to the spectrum of the original discrete-time signal by

$$\begin{aligned} X_z(f) &= \sum_{m=-\infty}^{\infty} x_z(m) e^{-j2\pi fm} \\ &= \sum_{m=-\infty}^{\infty} x(m) e^{-j2\pi fmI} \\ &= X(I.f) \end{aligned}$$

[4 Marks]

(ii)

Figure shows the block diagram of a re-sampling system.



The system is composed of an up-sampling subblock, this inserts 3 zeros in-between every two samples taking the sampling rate up by 5 times to 220 kHz. This is followed by a lowpass filter to replace the zeros with interpolated values. The cutoff frequency of the low pass filter should be $\pi/5$, or Fs/10 (which is 22 kHz at an Fs=220 kHz or 2 kHz at the new sampling rate of Fs=20) . The final sub block is a 1 to 11 down sampler.

[4 Marks]

(c)

Assume the sampling frequency is 8 kHz. For uniformly spaced band-pass filters the filter cutoff points for the four band-pass filters are 0 and 1 kHz, 1 and 2 kHz, 2 and 3 kHz, and 3 and 4 kHz respectively. First we design a low-pass filter with a bandwidth of 1 kHz. At a sampling rate of 8 kHz, a frequency of 1 kHz corresponds to a normalized angular frequency of $\omega_N = 2\pi \times 1/8$ radians or a normalized frequency of $f_N = 1/8$. Using the window design technique the FIR impulse response is obtained from the inverse Fourier transform as

$$h_d(m) = \int_{-0.125}^{0.125} 1.0\, e^{j2m\pi f}\, df$$

we obtain the FIR filter response as

$$h_1(m) = w(m) \times 0.25\,\mathrm{sinc}(0.25\pi(m - M/2)) \qquad 0 \le m \le M$$

Note for causality the filter impulse response is windowed and delayed.

[3 marks]

To design the band-pass filters we can use the amplitude modulation (AM) method to convert a low-pass filter to a band-pass filter. For a band-pass width of 1 kHz the low-pass filter should have cutoff frequencies of ±500 Hz (note that from –500 to +500 Hz we have a bandwidth of 1 kHz). Thus the required low pass FIR filter equation is given by

$$h(m) = w(m) \times 0.125\,\mathrm{sinc}(0.125\pi(m - M/2)) \qquad 0 \le m \le M$$

[3 marks]

To translate this low pass filter to the specified band pass filters we need AM sinusoidal carriers with frequencies 1.5 kHz, 2.5 kHz and 3.5 kHz. The modulated band pass filter equations are given by

$$h_2(m) = 2 \times w(m) \times 0.125\,\mathrm{sinc}(0.125\pi(m - M/2))\sin(3\pi m/8) \quad 0 \le m \le M$$
$$h_3(m) = 2 \times w(m) \times 0.125\,\mathrm{sinc}(0.125\pi(m - M/2))\sin(5\pi m/8) \quad 0 \le m \le M$$
$$h_4(m) = 2 \times w(m)\, 0.125\,\mathrm{sinc}(0.125\pi(m - M/2))\sin(7\pi m/8) \quad 0 \le m \le M$$

The process of amplitude modulation halves the amplitude of each sideband, hence we have the multiplying factor of 2.

[3 marks]

Q8.

(a) The discrete Fourier transform (DFT) is given by

$$X(k) = \sum_{m=0}^{N-1} x(m) e^{-j\frac{2\pi}{N}mk} \qquad k = 0, \ldots, N-1$$

(i) Multiply both sides of the DFT equation by $e^{-j2\pi kn/N}$ and take the summation as

$$\sum_{k=0}^{N-1} X(k)\, e^{-j\frac{2\pi}{N}kn} = \sum_{k=0}^{N-1}\sum_{m=0}^{N-1} x(m)\, e^{-j\frac{2\pi}{N}km}\, e^{-j\frac{2\pi}{N}kn}$$

Using the principle of orthogonality, the inverse DFT equation can be derived as

$$x(m) = \frac{1}{N}\sum_{k=0}^{N-1} X(k)\, e^{-j\frac{2\pi}{N}kn}$$

[5 marks]

(ii) Time-Frequency resolution. The DFT takes as input a window of N uniformly spaced time domain samples [$x(0)$, $x(1)$, …, $x(N\text{-}1)$] of duration $\Delta T = N.T_s$, and outputs N samples spaced uniformly in frequency domain between zero Hz and $f_s = 1/T_s$ Hz. The frequency resolution of the DFT spectrum Δf, that is the space between successive frequency samples, is proportional to the time window length NT_s and is given by

$$\Delta f = \frac{1}{\Delta T} = \frac{1}{NT_s} = \frac{f_s}{N}$$

Note that frequency resolution and time resolution are inversely proportional that is they can not both be small, in fact $\Delta T \Delta f = 1$. This is known as the uncertainty principle.

[3 marks]

(iii) The time resolution for audio signals is typically 20 ms giving a frequency resolution of 1/0.02=50Hz.

[2 Marks]

(a)

(i) The number of time domain samples required to achieve a frequency resolution of 100 Hz at the minimum sampling rate.

Assume a sampling rate =220 kHz

100Hz=1/(N/220000)=220000/N N=2200

[3 marks]

(ii) The spectrum of a short length signal can be interpolated to obtain a smoother spectrum. Interpolation of the frequency spectrum $X(k)$ is achieved by *zero-padding* of the time domain signal $x(m)$. Consider a signal of length N samples [$x(0), \ldots, x(N-1)$].

Increase the signal length from N *to* $2N$ samples by padding N zeros to obtain the padded sequence $[x(0), \ldots, x(N-1), 0, \ldots, 0]$. The DFT of the padded signal is given by

$$
\begin{aligned}
X(k) &= \sum_{m=0}^{2N-1} x(m) e^{-j\frac{2\pi}{2N}mk} \\
&= \sum_{m=0}^{N-1} x(m) e^{-j\frac{\pi}{N}mk}
\end{aligned}
\qquad k = 0, \ldots, 2N-1
$$

The spectrum of the zero-padded signal, is composed of $2N$ spectral samples; N of which, $[X(0), X(2), X(4), X(6), \ldots X(2N-2)]$ are the same as those that would be obtained from a DFT of the original N samples, and the other N samples $[X(1), X(3), X(5), X(6), \ldots X(2N-1)]$ are interpolated spectral lines that result from zero-padding. Note that zero padding does not increase the spectral resolution, it merely has an *interpolating or smoothing* effect in the frequency domain

[5 marks]

(iii) Actual resolution is 220000/4000=55 Hz, apparent resolution 27.5 Hz.

[2 marks]