IMPERIAL COLLEGE LONDON

AO2
SC1

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2010

Corrected Copy

**DIGITAL SIGNAL PROCESSING AND DIGITAL FILTERS**

Monday, 26 April 2:30 pm

Time allowed: 3:00 hours

**There are FOUR questions on this paper.**

**Answer THREE questions.**

*All questions carry equal marks*

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

Examiners responsible First Marker(s) : D.P. Mandic
Second Marker(s) : M.K. Gurcan

1) Consider the problem of digital filter design.

a) Explain the role of all-pass filters. [3]

i) Write down the transfer function of a first order all-pass filter and explain the positions of its pole and zero. [2]

ii) The pole-zero diagram of a second-order all-pass filter is given below. Write down the transfer function of this filter and show that its magnitude is constant. [5]

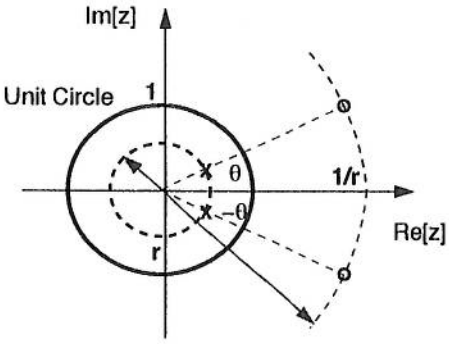



Figure 1: Pole-zero pattern of a second order all-pass digital filter.

b) Explain the operation of the frequency sampling method for the design of finite impulse response (FIR) digital filters. [2]

i) Derive the expression for the transfer function of an FIR digital filter designed using the frequency sampling method.
(Hint: start from the DFT representation of the desired frequency response) [4]

ii) The desired frequency response of an FIR digital filter is sampled over 16 points in the frequency range $[0, 2\pi)$, and is given by

$\widetilde{H}(0) = \widetilde{H}(1) = \widetilde{H}(15) = 1$
$\widetilde{H}(2) = \widetilde{H}(14) = 0.5$
$\widetilde{H}(k) = 0 \;\text{ for }\; k = 3, \ldots, 13$

Design the corresponding FIR digital filter using the frequency sampling method. [4]

2) State the aim of ARMA modelling and write down the equation for a general AR(p) model. [1]

a) Consider the first order autoregressive (Markov) process.

i) Derive the expression for the autocorrelation function for this process. [3]

ii) Write down and plot the autocorrelation function for an AR(1) process for the cases when the filter coefficient $a = 0.9$ and $a = -0.9$. [3]

iii) Derive the expressions for the variance and spectrum of such a process. Explain the shape of the spectrum for a negative value of the filter coefficient $a$. [3]

iv) Define the partial autocorrelation function and explain how the partial autocorrelation coefficients are calculated. Explain how the values of partial autocorrelation coefficents suggest the order of the AR model, and provide an example. [4]

b) Consider a general moving average (MA) process, MA(q), given by

$$x[n] = w[n] + b_1 w[n-1] + \cdots + b_q w[n-q], \qquad w[n] \sim \mathcal{N}(0, \sigma_w^2)$$

i) Write down the expression for the variance of this process. [2]

ii) Is the autocorrelation function finite or infinite in duration? [1]

iii) Consider the MA(1) process given by

$$z[n] = 0.8w[n-1] + w[n]$$

where $w$ denotes the driving white noise sequence. Write down the expression for the spectrum of this process. Explain whether the MA spectrum is suitable for the modelling of sinusoids in white noise. [3]

3) Consider the problem of sampling rate conversion.

a) Briefly define and discuss the principle of multirate sampling conversion. Give one example of a practical application. Explain the method of direct conversion within the digital domain and compare with auxiliary conversion to intermediate analogue signals. State advantages and disadvantages of both methods. [5]

b) Analyse and discuss in detail the down-sampling (decimation) method. Derive expressions for the output of the decimator, $y$, in the $z$-domain and frequency domain as a function of the input signal, $x$. Give diagrams of the spectra of the input, intermediate and output sequences. [6]

c) Consider an example of a converter shown in the Figure below. The sampled input signal is bandlimited with cut-off frequency 2 kHz. Sampling rates $F_X$ and $F_Y$ apply to the input and output signals, respectively.

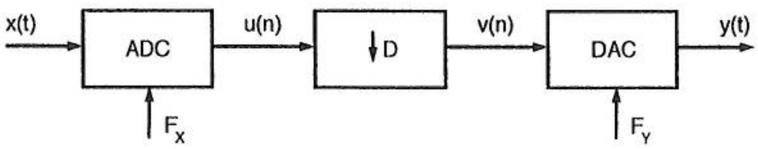



Figure 2: Block diagram of a sampling conversion scheme.

i) Plot the spectra for $u(n)$, $v(n)$ and $y(t)$ for a triangular spectrum of $x(t)$. What constraints are needed on the values of $D$, $T_X = 1/F_X$ and $T_Y = 1/F_Y$ to ensure that $y(t)$ is identical to $x(t)$? [3]

ii) For $F_X = F_Y = 10$kHz and $D = 2$, assess whether or not a distortionless reconstruction of the signal by $y(t)$ is possible in this case and explain why. Find the relationship between $y(t)$ and $x(t)$ for this converter. [2]

d) Explain the purpose of oversampling. Show how oversampling affects the total in-band noise power. Derive an expression for the effective resolution increase $\beta - b$ as a function of the oversampling ratio $M$. State the level of improvement over standard oversampling by sigma-delta oversampling and explain the reason for this. [4]

4) Consider the problem of linear estimation.

a) Define and solve the general (nonlinear) estimation problem, assuming a known forward conditional probability density function. Clearly state and interpret the best estimate. Derive the functional form of the estimator and the estimation error. Prove that the output of the estimator is orthogonal to the estimation error. [8]

b) Determine the coefficients of the optimum Wiener filter and state its relationship to the general estimation problem. Derive the squared error for the Wiener filter and derive its upper and lower bounds. [5]

c) For a Gaussian input signal to a LTI channel, derive the functional form of the estimator and discuss its relationship to the Wiener filter. [4]

d) Design a multirate sample converter for transforming a digital audio broadcast (DAB) signal with sampling frequency $F_X = 32$ kHz into a compact disc (CD) signal with sampling frequency $F_Y = 44.1$ kHz. Give a block diagram of the design. Determine the gain and cut-off frequency of the transfer function $H(\omega)$ to realize this converter. [3]

SC1
A02

1/12

**Solutions:** DSP and DIGITAL FILTERS 2010

1) **Bookwork and worked example.**

a) An all-pass filter is an IIR filter with a constant magnitude function for all digital frequency values. For a transfer function $H(z)$ to represent an all-pass filter is that for every pole $p_k = r_k e^{j\theta}$, there is a corresponding zero $z_k = \frac{1}{r_k} e^{j\theta}$. The poles and zeros will occur in conjugate pairs if $\theta_k \neq 0$ or $\pi$.

A digital filter $H(z)$ obtained by cascade connection of multiple all-pass filters $H_1(z), H_2(z) \cdots H_N(z)$ sections is itself an all-pass filter, and can be represented by

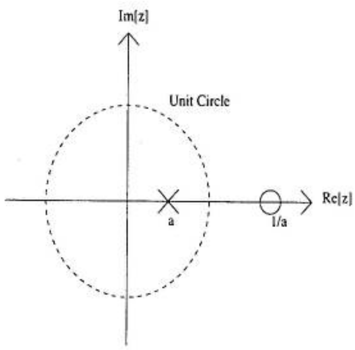
$$H(z) = H_1(z)H_2(z) \cdots H_N(z)$$

They are phase-selective (as opposed to frequency selective) and are extremely useful in the design of DSP systems.

i) A typical first-order section of an all-pass digital filter has a transfer function

$$H_1(z) = \frac{z^{-1} - a}{1 - az^{-1}}$$

where $a$ is real and to be stable, we must have $|a| < 1$. The pole-zero diagram in the $z$ plane is thus The magnitude function is unity for all frequencies, as given




by

$$|H_1(e^{j\theta})|^2 = |\frac{e^{-j\theta} - a}{1 - ae^{-j\theta}}|^2 = |\frac{\cos\theta - a - j\sin\theta}{1 - a\cos\theta + aj\sin\theta}|^2 = \frac{1 - 2a\cos\theta + a^2}{1 - 2a\cos\theta + a^2} = 1$$

ii) The transfer function of a typical second-order all-pass section is given by

$$H_2(z) = \frac{1 - (\frac{2}{r_k})\cos\theta_k z^{-1} + (\frac{1}{r_k^2})z^{-2}}{1 - 2r_k\cos\theta z^{-1} + r_k^2 z^{-2}} = \frac{[1 - (\frac{1}{r_k})z^{-1}e^{j\theta}][1 - (\frac{1}{r_k})z^{-1}e^{-j\theta}]}{[1 - r_k z^{-1}e^{j\theta}][1 - r_k z^{-1}e^{-j\theta}]}$$

The poles are at $p_{1,2} = r_k e^{\pm j\theta_k}$ and the zeros at $z_{1,2} = \frac{1}{r_k} e^{\pm j\theta_k}$. For filter to be stable, $|r_k| < 1$.
The magnitude function is given by

$$|H_2(e^{j\theta})|^2 = |\frac{e^{j\theta} - (\frac{1}{r_k})e^{j\theta_k}}{e^{j\theta} - r_k e^{j\theta_k}}|^2 |\frac{e^{j\theta} - (\frac{1}{r_k})e^{-j\theta_k}}{e^{j\theta} - r_k e^{-j\theta_k}}|^2$$

where $|\frac{e^{j\theta}-(\frac{1}{r_k})e^{j\theta_k}}{e^{j\theta}-r_k e^{j\theta_k}}|^2 = |\frac{e^{j\theta}-(\frac{1}{r_k})e^{-j\theta_k}}{e^{j\theta}-r_k e^{-j\theta_k}}|^2 = r_k^{-2}$.
Hence

$$|H_2(e^{j\theta})|^2 = r_k^{-4} = c$$

where $c$ is a constant, implying that it represents an all-pass filter.

b) **Bookwork and new example.**
This method ensures that the resulting frequency response coincides with the desired characteristics at the sampled points $\theta = \frac{2\pi k}{N}$, for $k = 0, 1, \ldots, N-1$. However, the resulting frequency response may not behave well in-between the sampling frequencies. This behaviour is related to the Gibbs phenomenon, which describes the overshoot of a step function represented by a truncated Fourier series.
i) An FIR filter has equivalent DFT representation, given by

$$\widetilde{H}(k) = \sum_{n=0}^{N-1} h(n) e^{-\frac{j2\pi nk}{N}}$$

where $\widetilde{H}(k)$ is the uniformly spaced N-point sample sequence of the frequency response of the digital filter. As a consequence, the impulse response sequence $h(n)$ and transfer function $H(z)$ are given by

$$h(n) = \frac{1}{N} \sum_{k=0}^{N-1} \widetilde{H}(k) e^{\frac{j2\pi nk}{N}}$$

and

$$H(z) = \frac{1}{N} \sum_{k=0}^{N-1} \widetilde{H}(k) \frac{1 - z^{-N}}{1 - z^{-1} e^{\frac{j2\pi k}{N}}}$$

ii) In this case, the DFT sequence is given by

$$\begin{aligned} \widetilde{H}(0) = \widetilde{H}(1) = \widetilde{H}(15) &= 1 \\ \widetilde{H}(k) = 0 \text{ for } k &= 3, 4, \ldots, 14 \\ \widetilde{H}(2) = \widetilde{H}(14) &= 0.5 \end{aligned}$$

Using the frequency sampling method, the desired transfer function can be expressed as

$$
\begin{aligned}
H(z) &= \frac{1}{16}\Big[\sum_{k=0}^{15}\frac{(1-z^{-16})\widetilde{H}(k)}{1-z^{-1}e^{\frac{jk\pi}{8}}}\Big] \\
&= \frac{1-z^{-16}}{16}\Big[\frac{1}{1-z^{-1}e^{\frac{j0\pi}{4}}}+\frac{1}{1-z^{-1}e^{\frac{j\pi}{8}}}+\frac{1}{1-z^{-1}e^{\frac{j15\pi}{8}}}+ \\
&+ \frac{0.5}{1-z^{-1}e^{\frac{j2pi}{8}}}+\frac{0.5}{1-z^{-1}e^{\frac{j14\pi}{8}}}\Big] \\
&= \frac{1-z^{-16}}{16}\Big[\frac{1}{1-z^{-1}}+\frac{2(1-z^{-1}\cos(\pi/8))}{1-2z^{-1}\cos(\pi/8)+z^{-2}}+\frac{1-z^{-1}\cos(\pi/4)}{1-2z^{-1}\cos(\pi/4)+z^{-2}}\Big]
\end{aligned}
$$

It can be be shown that the frequency response will be equal to the specifications at the sampling frequencies $\theta = \frac{k\pi}{8}$ for $k = 0, 1, 2, \ldots, 15$.

2) **Bookwork and worked examples:**

Autoregressive (AR) models are linear models which model the unknown data as a liner combination of fixed filter coefficients and the regressor vector (data in filter memory), and are driven by white noise, that is

$$z[n] = a_1 z[n-1] + a_2 z[n-2] + \cdots + a_p z[n-p] + w[n]$$

where $a_1, \ldots, a_p$ are the model parameters and $\{w[n]\}$ is the driving white noise.

a) The first order Markov process is given by

$$z[n] = az[n-1] + w[n]$$

By applying the expectation operator $E\{\cdot\}$ to

$$z[n-k]z[n]$$

we have

$$\rho(k) = a\rho(k-1) \quad or \quad \rho(k) = a^k, \quad k \geq 0$$

where $\rho(0) = 1$ and $\rho(1) = a$.

ii) **Bookwork and worked example:**
The ACF for $a = \pm 0.9$ is $\rho(k) = (\pm 0.9)^k, k \geq 0$. The plots are a decaying function with or without alternating the sign (for a negative $a$).

iii) For $k = 0$ the variance becomes

$$\sigma_z^2 = \frac{\sigma_w^2}{1 - a\rho(1)} = \frac{\sigma_w^2}{1 - a^2}$$

The spectrum of an AR(1) process is given by

$$S(f) = \frac{2\sigma_w^2}{|1 - ae^{-\jmath 2\pi f}|^2} = \frac{2\sigma_w^2}{1 + a^2 - 2a\cos(2\pi f)}$$

For a negative $a$ this represents a high–pass filter.

iv) Initially we may not know which order of autoregressive process to fit to an observed time series. This problem is analogous to deciding on the number of independent variables to be included in a multiple regression.
The partial autocorrelation function is a device which exploits the fact that whereas an AR(p) process has an autocorrelation function which is infinite in extent, it can by its very nature be described in terms of $p$ nonzero functions of autocorrelations. Denote by $a_{kj}$ the $j$th coefficient in an autoregressive representation of order $k$, so that $a_{kk}$ is the last coefficient. The $a_{kj}$ satisfy the set of equations

$$\rho(j) = a_{k1}\rho(j-1) + \cdots + a_{kk}\rho(j-k) \quad j = 1, 2, \ldots, k$$

leading to the Yule–Walker equations. The quantity $a_{kk}$, regarded as a function of lag $k$ is called the partial autocorrelation function. The large values of the partial autocorrelation function may therefore indicate undermodelling, whereas small values indicate over modelling.

b) **Bookwork and new example:**

i) For the MA(q) process

$$z[n] = b_1 w[n-1] + \cdots + b_q w[n-q] + w[n]$$

the variance is given by $E\{z^2\}$, that is

$$var(MA(q)) = \left(1 + b_1^2 + \cdots + b_q^2\right)\sigma_w^2$$

ii) The ACF is finite in duration and has a length $q$.

iii) The spectrum of an MA(q) process is given by

$$S(f) = 2\sigma_w^2 \left|1 - b_1 e^{-\jmath 2\pi f} - \cdots - b_q e^{-\jmath 2\pi q f}\right|^2$$

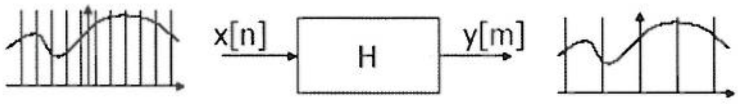
Theferore, for the given MA(1) we have

$$S(f) = 2\sigma_w^2[1 + 0.8^2 - 2 * 0.8\cos(2\pi f)], \quad 0 \le f \le 0.5$$

Since this spectrum is an all-zero system, it is appropriate for spectra with pronounced minima, and is not suitable for the modelling of peaky power spectra. Hence is not the best choice for the modelling of narrowband signals, such as the sinewave.

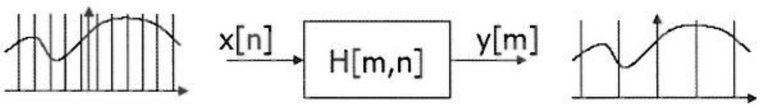
3 **[Bookwork and worked example]**

(a) multirate conversion = increasing or decreasing the sampling frequency of a signal to another sampling frequency




examples: CD (44.1 kHz) to/from DAB (32 kHz) to/from DAT (48 kHz); composite video signals NTSC (14.818 MHz) to/from PAL (17.734 MHz); also DVD-R, DVD+R, Blueray

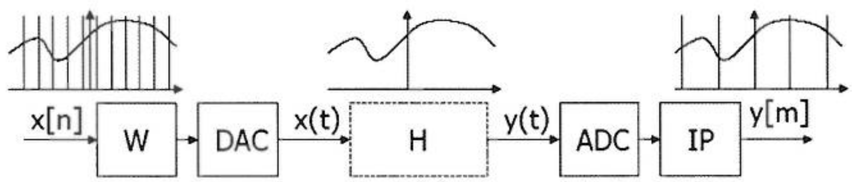
*Direct conversion:*




advantage: no additional distortion through intermediate ADC/DAC (only single ADC/DAC at front-ends)

disadvantages:
- fractional (rational) rates of conversion only
- large non-integer ratios require multi-stage approach

*Intermediate conversion to analogue signals:*




advantages:
- simple/obvious concept with low computation, hence fast
- high ratio of pre- and post-processing rates is possible in 1 stage

disadvantages:
- extra distortion by anti-aliasing filter (W), intermediate DAC and interpolating filter (IP)
- extra quantization error due to extra ADC

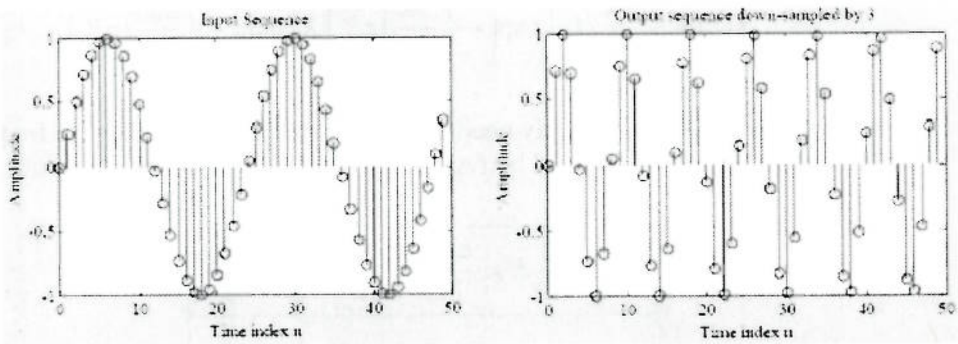
(b) Decimation: $T_y = DT_x$ with $D>1$, i.e., subsampling if $D$ is integer

$$y(mT_y) = \sum_{n=-\infty}^{+\infty} x(nT_x)h((mD-n)T_x)$$

Example for $D=2$:

$x_d[n] = x[2n] \quad n = 0, \pm 2, \pm 4, \ldots$

$$X_d(z)=\sum_{n=-\infty}^{+\infty}x_d[n]z^{-n}=\sum_{n=-\infty}^{+\infty}x[2n]z^{-n}=\sum_{m=-\infty}^{+\infty}x[m]z^{-m/2}=\frac{1}{2}\sum_{k=0}^{1}H\left(e^{-\frac{j2\pi k}{2}}z^{\frac{1}{2}}\right)X\left(e^{-\frac{j2\pi k}{2}}z^{\frac{1}{2}}\right)$$




decimation reduces data size and, consequently, processing time (less/no overhead)

result: time axis is effectively being "compressed"

decimation requires anti-aliasing low-pass filter, because of the lower output sampling rate: missing samples may cause ambiguity. Mathematically, owing to the sampling theorem, we require that

$$-F_X/(2D)\le F\le +F_X/(2D)\Leftrightarrow -\pi/D\le\omega\le +\pi/D$$

This pre-LPF may potentially cause distortion in its output signal

$$x(n)\longrightarrow\boxed{h_D(n)}\xrightarrow[(F_v=F_x)]{v(n)}\boxed{\downarrow D}\xrightarrow[(F_y=F_x/D)]{y(m)}$$

$$v(n)=\sum_{k=0}^{+\infty}h_D(k)x(n-k),$$

hence

$$y(m)=v(mD)=\sum_{k=0}^{+\infty}h_D(k)x(mD-k)\equiv\sum_{k=0}^{+\infty}h_D(k)\delta(mD-k)x(mD-k),$$

Taking the z-transform, using the expansion $\delta(n)=\frac{1}{D}\sum_{k=0}^{D-1}\exp\left(\frac{j2\pi n}{D}k\right)$ for the sampling function, and making the change of variable $z'=\exp\left(-\frac{j2\pi k}{D}\right)z^{\frac{1}{D}}$ yields

$$Y(z)=\sum_{n=-\infty}^{+\infty}v(n)\left[\frac{1}{D}\sum_{k=0}^{D-1}\exp\left(\frac{j2\pi n}{D}k\right)\right]z^{-\frac{n}{D}}=\frac{1}{D}\sum_{k=0}^{D-1}V\left(\exp\left(-\frac{j2\pi k}{D}\right)z^{\frac{1}{D}}\right),$$

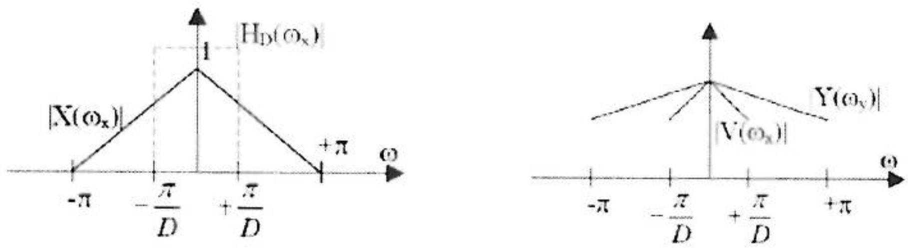
Using the relationship between the z-transforms of the input and output of a linear system $H_D$,

$$Y(z) = \frac{1}{D}\sum_{k=0}^{D-1} H_D\left(\exp\left(-\frac{j2\pi k}{D}\right)z^{\frac{1}{D}}\right)X\left(\exp\left(-\frac{j2\pi k}{D}\right)z^{\frac{1}{D}}\right)$$

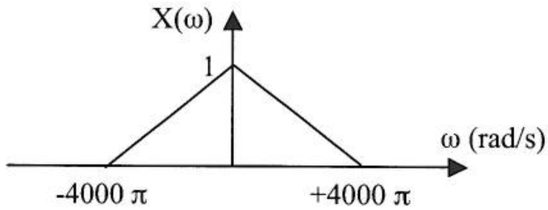
For the anti-aliasing filter, we only need to retain the term $k$=0, so that the frequency spectrum of $Y$ (generally obtained by replacing $z$ with exp($j\omega$)) is (final result):

$$\boxed{Y(\omega_y) = \frac{1}{D}H_D\left(\frac{\omega_y}{D}\right)X\left(\frac{\omega_y}{D}\right) = \frac{1}{D}X\left(\frac{\omega_y}{D}\right)}$$

Spectra:



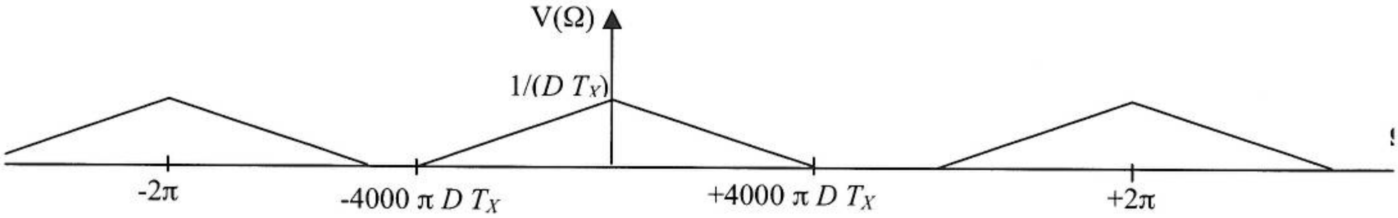

(c) [**New example**]
1. spectrum of $x(t)$ is



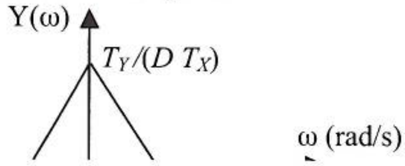

Since $v(n) = u(nM) = x(nD\,T_X)$, aliasing of $u(t)$ is avoided whenever

$$DT_X < \frac{1}{4000}s = 0.25\,\text{ms}$$

If this condition is satisfied, then the DTFT of the output of the decimator is




The Fourier transform of the output is then



Thus, in order for $y(t) = x(t)$, it is necessary that
(A) $D\,T_X < 0.25$ ms to prevent frequency aliasing, and
(B) $T_Y = D\,T_X$. to prevent frequency scaling

2. For $T_X = T_Y = 1/10000 = 0.1$ ms and $D = 2$, the condition

$$DT_X = \frac{1}{5000} < \frac{1}{4000}$$

is satisfied. Therefore,

$$Y(\omega) = \frac{1}{2} X\left(\frac{\omega}{2}\right), \qquad y(t) = x(2t)$$

(d) [**Bookwork and critical reasoning**]

Oversampling:

Optimum sampling (=at Nyquist rate $2F_B$) allows in principle for perfect reconstruction, but faces problem of needing an steep anti-aliasing filter that is infinitely steep (because spectral images of sampled signal are adjacent in frequency with zero spacing). If done with an analogue filter, this becomes expensive and difficult, imprecise, with large phase distortion, etc.
By using oversampling, i.e., sampling at rates much higher than Nyquist rate, spectral images (periods) in sampled signal spectrum acquire larger separation, which can be filtered more easily with a low roll-off filter. Can be done adequately with analogue filter (cheaper than digital filter and reduces computational requirements)

Noise reduction through oversampling:

For a $b$-bit ADC with range $R$, the quantization step is $Q = \dfrac{R}{2^b}$

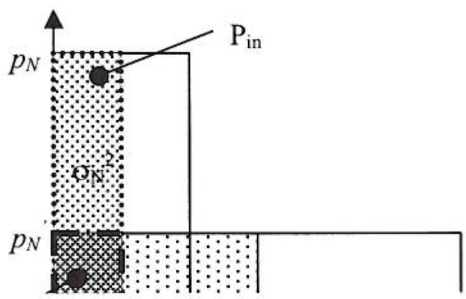
The noise power density per unit sampling bandwidth is then

$$p_N = \frac{\sigma_N^2}{F_s/2} = \frac{Q^2/12}{F_s/2} = \frac{Q^2}{6F_s} \quad \text{W/Hz}$$

The total in-band noise power is then

$$P_{in} = \int_0^{F_B} p_N(f)\,\mathrm{d}f = \frac{Q^2 F_B}{6F_s} = \frac{(R/2^b)^2}{12} \times \frac{F_B}{F_s/2}$$

Thus, a high $F_s$ produces a low $P_{in}$: $F_s' >> 2F_B \Rightarrow P_{in}' << \sigma_N^2$ with the following diagram:



A $\beta$-bit ADC operating at Nyquist rate will be equivalent to a $b$-bit ADC operating at oversampling rate over the same range $R$, where "equivalent" means in the sense of producing the same noise power, if

$$\frac{(R/2^b)^2}{12} \times \frac{F_B}{F_s/2} = \frac{(R/2^\beta)^2}{12} \times \frac{F_B}{2F_s/2}$$

i.e., when

$$\beta - b = \frac{\log_2\left(F_s/(2F_B)\right)}{2} = \frac{\log_2(M)}{2}$$

where $M$ represents the oversampling ratio and $\beta$-$b$ represents the effective increase in resolution

As a rule of thumb, each doubling of $M$ allows for a reduction of resolution of the ADC by half a bit.

Oversampling using a sigma-delta converter produces

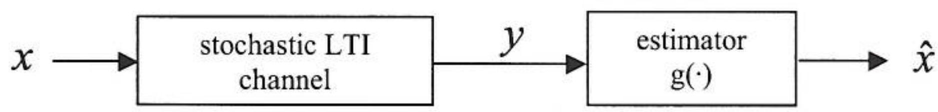
$$10\log_{10}\left(\frac{P_{in}}{P_y}\right) = 10\log_{10}\left(\frac{3M^2}{\pi^2}\right) = [-5.17 + 20\log_{10}(M)] \text{ dB}$$

now yielding 1.5 bit reduction per doubling of $M$.

Reason for improvement: noise shaping by the transfer function of the sigma-delta converter (due to noise reduction for frequencies below bandwidth $F_B$).

4 [**Bookwork and worked example**]

(a)




Given:

- a source generating <u>nonobservable</u> stochastic signals $X$, e.g., signal+noise, bits, characters, etc.

- a stochastic linear time invariant channel (extendable to time-varying), e.g., 1/f noise, noise from resistor, ionospheric propagation, etc.
- *observable* outputs $Y$ of the channel
- assumption: the forward probability $f_{Y|X}(y \mid X = x)$ is known, but backward probability $f_{X|Y}(x \mid Y = y)$ is not

Problem statement: find best estimate $\hat{X}$ of $X$, based on observations $Y$

Solution

We can compute the joint probability density function (pdf) of $X$ and $Y$ as

$$f_{X,Y}(x, y) = f_{Y|X}(y \mid X = x) \cdot f_X(x)$$

We define and design an optimal estimator for $X$, in the sense that out of all possible transfer functions g(.) for this estimator, we pick the one $g_{opt}(.)$ for which

$$\hat{x} = g_{\mathbf{opt}}(y)$$

is closest to $x$, in the sense of minimizing the mean squared deviation from $x$, i.e.,

$$\begin{aligned}
\mathrm{E}[(X - \hat{X})^2] &= \iint [x - g(y)]^2 f_{X,Y}(x, y)\, dx\, dy \\
&= \int dy\, f_Y(y) \int [x - g(y)]^2 f_X(x \mid Y = y)\, dx \\
&\doteq \int f_Y(y) K(y) dy
\end{aligned}$$

Since $K(y) \geq 0$ , we should minimize

$$\begin{aligned}
K(y) &= \int [x - g(y)]^2 f_{X|Y}(x \mid Y = y)\, dx \\
&= \mathrm{E}(X^2 \mid Y = y) - 2g(y)\mathrm{E}(X \mid Y = y) + g^2(y)
\end{aligned}$$

Thus, from $\frac{dK(y)}{dg} = 0$ for the minimum deviation, this leads to the solution

$$g(y) = \mathrm{E}(X \mid Y = y)$$

which yields the best estimate in a MMSE sense. The associated mean squared deviation (error) itself with this choice for g(.) is

$$\varepsilon_{\mathrm{opt}}^2 = E(X^2) - \int E^2(X \mid Y = y) f_Y(y)\, dy$$

In general, the optimal estimator is nonlinear.

Proof of orthogonality:

$$\iint\left[x-\mathrm{E}(X\mid y)\right]g(y)\,f_{X,Y}(x,y)\,dx\,dy=\iint g(y)\,f_{X,Y}(x,y)\,dx\,dy$$
$$-\iint \mathrm{E}(X\mid y)\,g(y)\,f_{X,Y}(x,y)\,dx\,dy$$
$$=\iint x\,g(y)\,f_{X,Y}(x,y)\,dx\,dy-\int \mathrm{E}(X\mid y)g(y)\,f_Y(y)\,dy$$
$$=\iint x\,g(y)\,f_{X,Y}(x,y)\,dx\,dy-\int x\,f_{X|Y}(x\mid y)dx\int g(y)\,f_Y(y)\,dy$$
$$=\iint x\,g(y)\,f_{X,Y}(x,y)\,dx\,dy-\iint x\,g(y)\,f_{X,Y}(x,y)\,dx\,dy$$
$$=0$$

(b) Wiener filter = best <u>linear</u> estimator, i.e., of the form $\hat{x}=\alpha y+\beta$ Since the estimator minimizes the mean square error under this constraint, its coefficients are the solutions of

$$\begin{cases}\dfrac{\partial\,\mathrm{E}[(X-\alpha Y-\beta)^2]}{\partial\alpha}=0\\[2ex]\dfrac{\partial\,\mathrm{E}[(X-\alpha Y-\beta)^2]}{\partial\beta}=0\end{cases}$$

This yields (intermediate calculations to be demonstrated)

$$\begin{cases}\alpha=r\dfrac{\sigma_X}{\sigma_Y}\\[1ex]\beta=\mathrm{E}(X)-\alpha\mathrm{E}(Y)\end{cases}$$

where

$r=\dfrac{\sigma_{XY}}{\sigma_X\sigma_Y}$ is the correlation coefficient between $X$ and $Y$

The associated mean squared error for the Wiener filter is

$$\varepsilon^2_{\text{opt}}=\sigma_X^2(1-r^2)$$

(c) For Gaussian $X$ and $Y$

$$\mathrm{E}(X\mid Y=y)=r\frac{\sigma_X}{\sigma_Y}\left[y-\mathrm{E}(Y)\right]+\mathrm{E}(X)$$

Interpretation: <u>linear</u> estimate for Gaussian random variables is also its <u>optimal</u> estimate

12/12

(d) We require the ratio I/D for rational (fractional) sample rate conversion, which in this case is

$$\frac{I}{D} = \frac{44.1}{32} = \frac{441}{320}$$

Thus, we upsample (interpolate) by a factor $I$=441then downconvert (decimate) by factor $D$=430.

The low-pass filter required has a cut-off frequency given by

$$\omega_c = \min\left(\frac{\pi}{I}, \frac{\pi}{D}\right) = \frac{\pi}{441}$$

and the gain of the filter should be $I$=441.