IB Paper 6, 2012: Solutions

SECTION A

1. *System response and steady-state errors*

   (a) First

   $$\bar{y}(s) = \frac{1}{s(s+3)}\left(\bar{d}(s) + K(\bar{r}(s) - \bar{y}(s))\right)$$

   $$\bar{y}(s) = \frac{\bar{d}(s) + K\bar{r}(s)}{s(s+3) + K}$$

   Set either $\bar{d}(s) = 0$ or $\bar{r}(s) = 0$ to get transfer functions

   $$\frac{\bar{y}(s)}{\bar{r}(s)} = \frac{K}{s^2 + 3s + K}$$

   $$\frac{\bar{y}(s)}{\bar{d}(s)} = \frac{1}{s^2 + 3s + K}$$

   [4]

   (b) Compare denominator of the terms in the closed loop transfer functions to

   $$s^2 + 2c\omega_n s + \omega_n{}^2$$

   Hence $\omega_n{}^2 = K, 2c\omega_n = 3$

   $$c = \frac{1.5}{\sqrt{K}}$$

   [3]

   (c) In each of these use the final value theorem to determine the final value of the error.

   $$\bar{y}(s) - \bar{r}(s) = \frac{1}{s^2 + 3s + K}\bar{d}(s) - \frac{s^2 + 3s}{s^2 + 3s + K}\bar{r}(s)$$

   Now we want $\lim_{s\to 0} s\left(\bar{y}(s) - \bar{r}(s)\right)$ for each case.

   (c)(i) $\bar{r}(s) = 1/s, \bar{d}(s) = 0$. Final value of error $= 0$. [2]

   (c)(ii) $\bar{r}(s) = 0, \bar{d}(s) = 1/s$. Final value of error $= \frac{1}{K}$. [2]

   (c)(iii) $\bar{r}(s) = 0, \bar{d}(s) = \frac{1}{s+2}$. Final value of error $= 0$. [2]

(c)(iv) $\bar{r}(s) = \frac{1}{s^2}, \bar{d}(s) = 0$.

$$\lim_{s \to 0} s \left(\bar{y}(s) - \bar{r}(s) \right) = \lim_{s \to 0} s \left(-\frac{s^2 + 3s}{s^2 + 3s + K} \frac{1}{s^2} \right)$$
$$= \frac{-3}{K}$$

[3]

(d) If the proportional controller is replaced by a PI controller with transfer function

$$K_P + \frac{K_I}{s} = \frac{K_P s + K_I}{s}$$

In this case the transfer function becomes

$$\frac{\bar{y}(s)}{\bar{d}(s)} = \frac{s}{s^3 + 3s^2 + K_P s + K_I}$$

This means with a steady state disturbance input as in (c)(ii), the steady-state error is zero.

Also note that for the case in (c)(iv) now becomes

$$\lim_{s \to 0} s\bar{y}(s) - \bar{r}(s) = \lim_{s \to 0} s \left(-\frac{s^3 + 3s^2}{s^3 + 3s^2 + K_P s + K_I} \frac{1}{s^2} \right)$$
$$= 0$$

Hence there would be no errors in tracking a ramp also. [4]

2. *Nyquist Diagrams*

(a) The closed-loop system is stable if and only if the complete Nyquist diagram of the open loop system $K(s)G(s)$ has <u>no</u> encirclements of the '-1' point . The Nyquist diagram is a plot of the open loop steady state frequency response of the system in the complex plane with the frequency varying along the plot. [3]

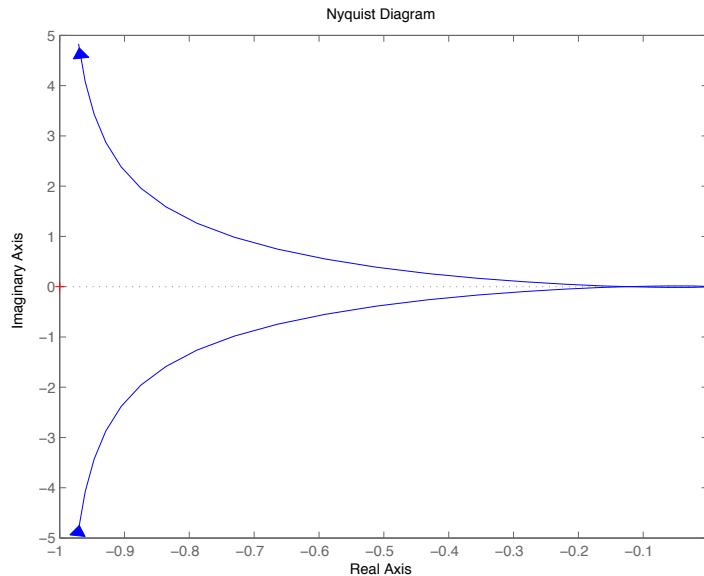(b) To construct the Nyquist diagram set $s = j\omega$ and measure $G(j\omega)$.

$$G((j\omega) = \frac{1}{j\omega(2 - \omega^2 + 4j\omega)}$$

As $\omega \to \infty$, the $|\cdot| \to 0$ and phase to $\frac{-3\pi}{2}$.
As $\omega \to 0$, $|\cdot| \to \infty$ and phase to $\frac{-\pi}{2}$.
In fact as can be seen from a Taylor expansion, as $G(j\omega \to 0) \to -1$ (which is the required asymptote). The plot crosses the negative real axis when the imaginary part is zero i.e. at $\omega = \sqrt{2}$.

Hence the Nyquist plot can be drawn:

The magnitude of the closed-loop frequency response is

$$\frac{|K.G(j\omega)|}{|1 + K.G(j\omega)|}$$

From the Nyquist plot, for $K = 1$ the numerator is the distance from the origin to the point on $G(j\omega)$ curve at a particular frequency. The denominator is the distance from (-1,0) to the point on the $G(j\omega)$ curve. [7]

3

(c) Gain margin is simply the reciprocal of the intersection with the negative real axis. Since $\omega = \sqrt{2}, G(j\omega) = 1/8$, the gain margin is equal to 8. [2]

(d) The $j\omega$ term on the denominator gives a phase of $-90°$. If the phase margin is $45°$ then at the phase margin $\omega = \omega_c$ and

$$
\begin{aligned}
-\tan^{-1} \frac{4\omega_c}{2 - \omega_c{}^2} &= -45° \\
\frac{4\omega_c}{2 - \omega_c{}^2} &= 1 \\
\omega_c &= -2 \pm \sqrt{6} \\
\omega_c &= \sqrt{6} - 2 \\
&= 0.449
\end{aligned}
$$

The magnitude is

$$
|G(j\omega)| = \frac{1}{\sqrt{16\omega^4 + (2\omega - \omega^3)^2}}
$$

However at the given phase angle we know that the real and imaginary parts of the denominator are equal hence:

$$
\begin{aligned}
|G(j\omega_c)| &= \frac{1}{\sqrt{2 \times 16\omega_c{}^4}} \\
&= \frac{1}{4\sqrt{2}\omega_c{}^2} \\
&= 0.875
\end{aligned}
$$

Hence the value of K needed to scale $|G(j\omega_c)|$ to unit magnitude is $1/0.875 = 1.143$. [4]

The magnitude of the closed loop frequency response at the phase margin doesn't depend on K, but by simple trigonometry

$$
\begin{aligned}
&= \frac{1}{2\sin 22.5°} \\
&= 1.306
\end{aligned}
$$

[2]

(e) If the peak magnitude of the closed loop frequency response is much greater than unity, the system will have a low phase margin. The step response will be oscillatory. Both of these will be undesirable. [2]

3. *Bode Plots*

(a) For an asymptotically stable open-loop system: the phase margin (if it exists) is the increase in (negative) phase before the system becomes unstable at constant gain; the gain margin (if it exists) is the factor that the gain must increase at constant phase before the system becomes unstable (measured when phase angle $= -180°$). [4]

(b) Note that we have a gain ($K(s) = 2$ or 6dB. Hence the response given in Fig 4 for $G(s)$ alone is increased by 6dB. Hence we need to find the difference between -6dB and the plot when the phase plot is at $-180°$. This is about 5dB occurring at a frequency of about 65 rad/s. For the PM, find when the gain plot including the $K = 2$ is at 0dB, and find the difference between the plot and $-180°$. Hence look at where the plot of $G(s)$ crosses the -6dB line, and look for the phase difference between the actual phase and $-180°$. This occurs at about 50 rad/s and the phase is about $-165°$, so the phase margin is about $15°$ (accurately $17°$). [5]

(c)(i) First draw the plot for the compensator. It might be helpful to rearrange the response for the compensator as

$$K(j\omega) = \frac{20(30 + j\omega)}{300 + j\omega}$$
$$= \frac{2(1 + j\omega/30)}{1 + j\omega/300}$$

Hence the compensator has the same DC gain as in part (b). However we now have a breakpoint frequency of 30 rad/s on the numerator increasing at 20dB/decade. The denominator breakpoint is at 300 rad/s and cancels the numerator after that frequency to have an overall increase in gain at high frequency of 20dB. The effect on phase will be over a wide range of frequencies (approx 3 rad/s to 3000 rad/s) where it will give a positive phase. Need to compute some sample points of the phase advance due to the compensator in order to yield an accurate plot.

For instance at $\omega = 30$ phase advance is $tan^{-1}1 - tan^{-1}0.1 = 39°$. At $\omega = 100$ phase advance is $tan^{-1}10/3 - tan^{-1}1/3 = 54°$. etc. Combining these yields a Bode plot for the original (red line, lower line in phase plot), compensator alone (yellow line, upper line in phase plot) and compensated loop (green line, middle line in phase plot): see Fig. 1. [6]

(c)(ii) Measuring from the plot of the compensated loop on the revised yields a phase margin of about $45°$ (at 80rad/s) (accurately $43°$ at 81rad/s) and a gain margin of 10-15dB at about 180 rad/s (accurately 12.5dB (at 184 rad/s). [2]

(c)(iii) Note that the DC gain are the same in both cases (& hence the steady state error). However the compensated look is oscillatory (the uncompensated loop has a small GM and PM and is fairly close to instability), and also the compensated loop has a greater bandwidth. [3]
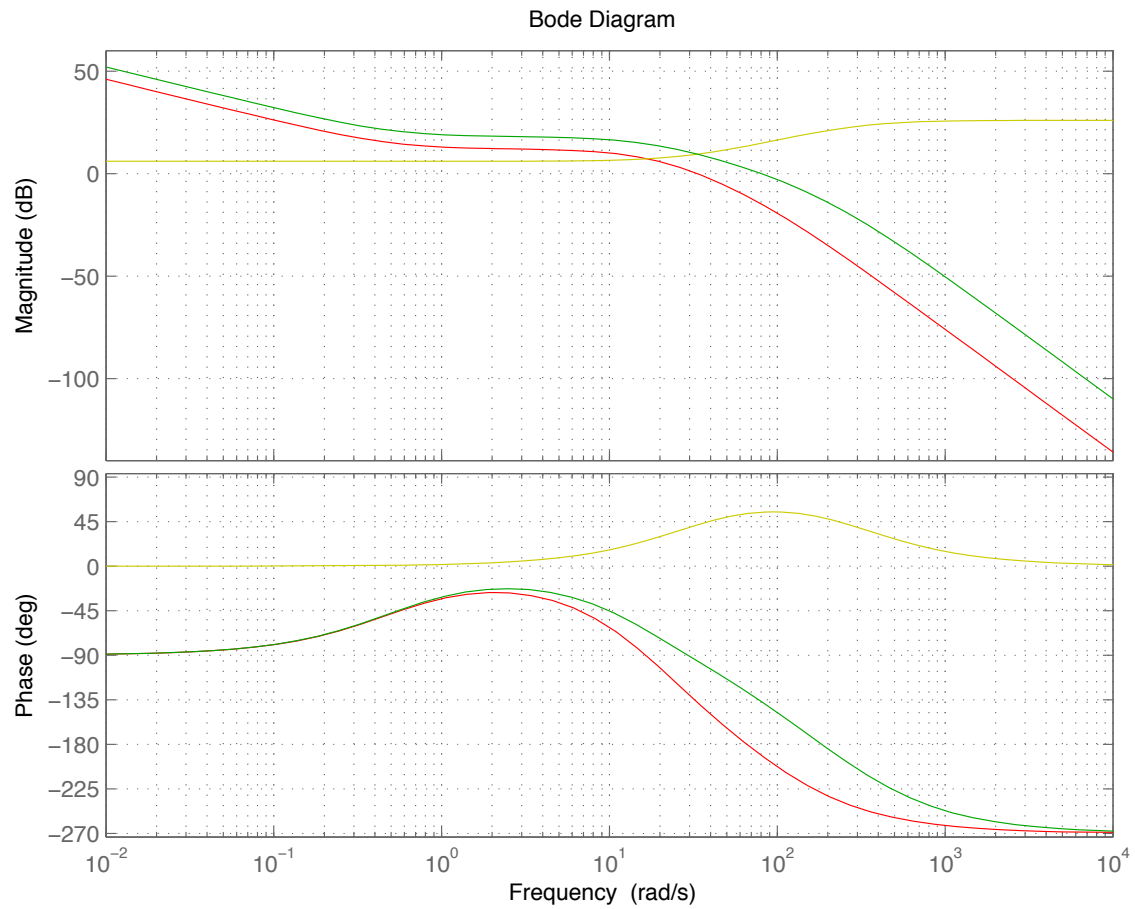
Figure 1: Bode plots for the original loop, compensator and compensated loop, for use with phase margin and gain margin calculations

SECTION B

4. *Fourier Transforms: shift theorem & convolution*

(a) The IFT is written as:

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)e^{j\omega t} d\omega$$

Thus, replacing t by $t - t_0$ in this equation gives;

$$
\begin{aligned}
f(t - t_0) &= \frac{1}{2\pi} \int_{-\infty}^{+\infty} F(\omega)e^{j\omega(t - t_0)} d\omega \\
&= \frac{1}{2\pi} \int_{-\infty}^{+\infty} \{F(\omega)e^{-j\omega t_0}\}e^{j\omega t} d\omega
\end{aligned}
$$

Which tells us that $f(t - t_o)$ and $e^{-j\omega t_0} F(\omega)$ are FT pairs. [3]

(b) Write $\sin(a\omega)$ as $\frac{1}{2j}\{e^{ja\omega} - e^{-ja\omega}\}$, so that we want to find the IFT of $\frac{1}{2j}\{F(\omega)e^{ja\omega} - F(\omega)e^{-ja\omega}\}$. Using the result of part(a), this straightforwardly gives [4]

$$\frac{1}{2j}\{f(t + a) - f(t - a)\}$$

(c) The derivative of the Gaussian $g(t)$ is given by

$$g'(t) = -2a^2 t e^{-a^2 t^2} \equiv -2a^2 t g(t)$$

and the derivative of this is given by

$$g''(t) = -2a^2\{1 - 2a^2 t^2\}g(t)$$

Thus, $g'(0) = 0$, $g'(t) \to 0$ as $t \to \pm\infty$, and $g'(t)$ is positive for $t < 0$ and negative for $t > 0$. From $g''(t)$ we see that there are max and min at $t = \pm 1/(\sqrt{2}a)$.

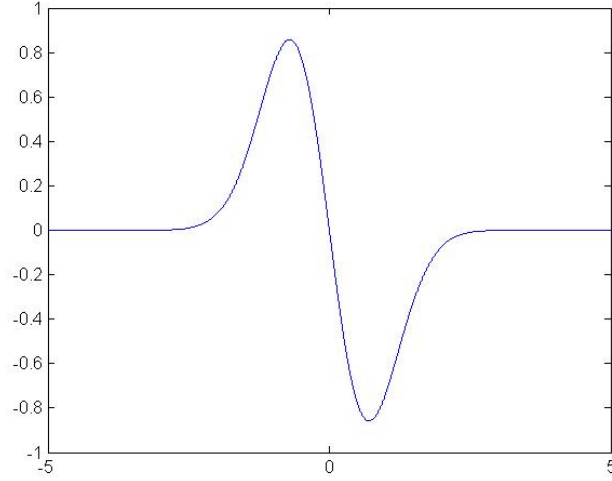This is shown in figure 2 for $a = 1$: [5]

Figure 2: Derivative of a Gaussian

(d) We now convolve this derivative of a Gaussian with the pulse $p(t)$. Since we know that convolution in the spatial domain is equivalent to multiplication in the Fourier domain, we have that

$$H(\omega) = FT(g'(t))FT(p(t))$$

(i) But we also know that if $F(\omega)$ is the FT of $f(t)$, then $j\omega F(\omega)$ is the FT of $f'(t)$. Therefore the FT of $g'(t)$ is

$$FT(g'(t)) = j\omega G(\omega) = \frac{j\omega\sqrt{\pi}}{a}e^{-\omega^2/(4a^2)}$$

and from the Databook we know that $P(\omega)$ is given by

$$P(\omega) = cb\,\text{sinc}\frac{\omega b}{2}$$

Thus, $H(\omega)$ can be written as $\{cb\,\text{sinc}\frac{\omega b}{2}\}\{j\omega G(\omega)\}$, which is

$$H(\omega) = cb\frac{\sin(\omega b/2)}{\omega b/2}j\omega G(\omega) = 2jc\sin\frac{\omega b}{2}G(\omega)$$

[4]

(ii) Using the result of part b), we know that the IFT of $H(\omega)$ can be written as

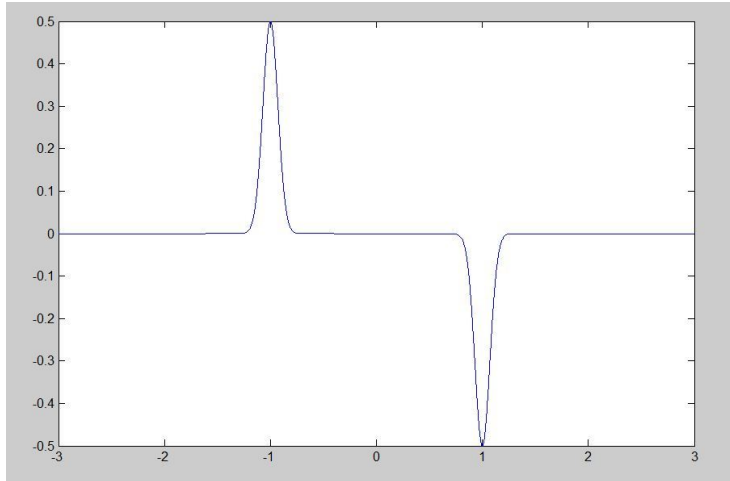$$IFT(H(\omega)) = h(t) = \frac{2jc}{2j}\{g(t+b/2) - g(t-b/2)\} = c\{g(t+b/2) - g(t-b/2)\}$$

8

Figure 3: Value of convolution if $b \gg 1/a$ $[c = 0.5, b = 2, a = 10]$

ie the difference of two gaussians centred on $\pm b/2$. If $b \gg 1/a$ then, since $1/a$ determines the width of the gaussian (ie $\sigma \sim 1/a$) then we can see that the convolution will approx be two gaussians which interfere with each other minimally, see figure 3. If $b \ll 1/a$ then the width of our pulse is very small and we approximately have a delta function, which means that the convolution will simply give the derivative of a gaussian as shown earlier in figure 2. We therefore see that convolving with a gaussian derivative (of appropriate width) can be used as an edge detector (not necessary in the answer). [4]

5. *DFT and signal quantisation*

(a) If we represent the sampled version of $x(t)$ as $x_s(t)$ where

$$x_s(t) = \sum_{n=-\infty}^{\infty} x(nT)\delta(t - nT)$$

$$= x(t) \sum_{n=-\infty}^{\infty} \delta(t - nT) = x(t)\,\delta_p(t)$$

with $f_s = 1/T$, we can write the train of delta functions $\delta_p(t)$ as a Fourier series (since it is obviously periodic) as follows:

$$\delta_p(t) = \sum_{n=-\infty}^{\infty} \frac{1}{T} e^{jn\omega_0 t}$$

with $\omega_0 = 2\pi/T$.

Therefore $X_s(\omega)$ is given by the FT of $\frac{1}{T}\sum_{n=-\infty}^{\infty} x(t)e^{jn\omega_0 t}$. Hence, by the shift theorem, we can write this as

$$X_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} X(\omega - n\omega_0)$$

ie the FT of the sampled signal is the FT of the original signal repeated at intervals of the sampling frequency.

Thus, provided the width of the FT $X(\omega)$ is less than half of the sampling frequency, we do not get overlaps in the Fourier domain and the original FT can be filtered and inverted to obtain a perfect reconstruction of the original signal. If we do get overlap of the repeated FTs, we observe artefacts known as aliasing [note, sometimes the whole process of repeating the spectrum is referred to as *aliasing*]. Thus, the condition for exact recovery is that $f_s > 2f_{max}$. [6]

(b) The formula for the inverse DFT is

$$x_n = \frac{1}{N} \sum_{m=0}^{N-1} X_m e^{jnm2\pi/N} \quad n = 0, .., N-1$$

(i) In the case of the DFT sequence $[0, 1, 0, 1]$, $N = 4$ and so our original signal samples $\{x_n\}$ are given by:

$$x_0 = \frac{1}{4}\{0 + 1 + 0 + 1\} = \frac{1}{2}$$

$$x_1 = \frac{1}{4}\{0 + 1e^{j2\pi/4} + 0 + 1e^{j3\times2\pi/4}\} = \frac{1}{4}\{e^{j\pi/2} + e^{j3\pi/2}\} = \frac{1}{4}\{j - j\} = 0$$

$$x_2 = \frac{1}{4}\{0 + 1e^{j2\pi/2} + 0 + 1e^{j3\times2\pi/2}\} = \frac{1}{4}\{e^{j\pi} + e^{j3\pi}\} = \frac{1}{4}\{-1 - 1\} = -\frac{1}{2}$$

$$x_3 = \frac{1}{4}\{0 + 1e^{j3\pi/2} + 0 + 1e^{j3\times3\pi/2}\} = \frac{1}{4}\{e^{j3\pi/2} + e^{j9\pi/2}\} = \frac{1}{4}\{-j + j\} = 0$$

Therefore $\{x_n\} = [\, 1/2, \ 0, \ -1/2, \ 0\,]$. [4]

(ii) If we sample at 4Hz the kth frequency component in the DFT is given by $k/(NT)$ where $k = 0, 1, 2$ and $N = 4$, $T = 1/4$, ie kHz – note that we can only pick out frequencies at half of the sampling frequency (recall that $X_{N-m} = X_m^*$). Thus the non-zero components in the DFT represent a frequency of 1Hz. Given this information about the spectrum of the signal, we can infer that the simplest form of the signal (ie no aliasing) is

$$x(t) = a_1 \sin 2\pi t \quad \text{or} \quad a_1 \cos 2\pi t$$

Taking cos, $x_0 = x(0) = a_1$, $x_1 = x(1/4) = 0$, $x_2 = x(2/4) = -a_1$, $x_3 = x(3/4) = 0$, which agrees with inverse DFT if $a_1 = 1/2$, ie $x(t) = 0.5 \cos 2\pi t$. [4]

(c) If we sample a signal with 3kHz bandwidth, our sampling frequency must be greater than or equal to 6kHz to prevent aliasing, call this sampling frequency f_s. If we use 2^N quantisation levels, the number of bits per second that we need to transmit is $f_s \times N$. Thus, if the maximum transmission rate is 64kbits per sec, we have that

$$f_s N \leq 64000$$

So that $N \leq 64000/f_s$. Thus, N is maximised when f_s is minimised and we know that the minimum value of f_s is 6kHz, therefore $N \leq 64000/6000$ and the maximum number of levels we can use is $2^{10} = 1024$.

We can compute the quantisation noise by finding the expected value of the squared error over an quantisation interval – if the step size is Δ, it is not hard to show (see notes) that the RMS noise is given by $\Delta/\sqrt{12}$, if we assume that the real signal is equally likely to lie anywhere in the interval. Thus, for an N-bit quantiser with 2^N levels, assuming our range is from $-V$ to $+V$, the step is given by

$$\Delta = \frac{2V}{2^N} = 2^{1-N}V$$

Therefore the SNR is given by the square of the RMS signal over the square of the RMS noise, ie

$$\text{SNR} = (V^2/2)/(\Delta^2/12) = 6V^2/\Delta^2 = 6V^2/(2^{2-2N}V^2) = 3 \times 2^{2N-1}$$

In dB, this is $10\log_{10}(3 \times 2^{2N-1})$ which is

$$= 10\log_{10}(3/2) + 20N\log_{10}2 = 1.76 + 6.02N$$

[See Databook!].

So, the SNR in this case is $1.76 + 60.2 = 61.96$dB. [6]

6. *Modulation (analogue & digital) and PSK*

(a) This first part is basically bookwork. Some summarised version of the notes is expected, which includes the following key points:

Modulation is the process of *shaping* one or more parameters of a carrier wave with the information signal $x(t)$. The carrier wave will be tuned (in terms of frequency, bandwidth etc.) to the channel and the information signal will be extracted at the receiver via *demodulation*. The parameters of the carrier wave that are *shaped* are *amplitude, phase* and *frequency*. This is necessary in general as the medium may not be able to transport the information in its raw form – reasons for this could be; size of antennae required, sharing of medium with other users, propagation properties as a function of signal frequency.

In *amplitude modulation* (AM) of analogue signals, the amplitude of the carrier wave, where the carrier is given by $a_0 \cos(2\pi f_c t)$, is shaped such that the AM modulated signal, $s_{AM}(t)$, is given by

$$s_{AM}(t) = [a_0 + x(t)] \cos(2\pi f_c t)$$

$m_A = \max_t x(t)/a_0$ is known as the *modulation index* and is normally chosen to be less than 1 to enable simple demodulation.

In *phase modulation* (PM) of analogue signals, the *instantaneous phase* of the carrier wave is shaped such that the PM modulated signal, $s_{PM}(t)$, is given by

$$s_{PM}(t) = a_0 \cos(2\pi f_c t + \phi_\Delta x(t))$$

where ϕ_Δ is known as the phase deviation. Analogue PM is not used widely in practice, mainly because it is harder to demodulate. However, PM can be effectively encompassed in a discussion of FM – see below.

In *frequency modulation* (FM) of analogue signals, the *instantaneous frequency* of the carrier wave is shaped such that the FM modulated signal, $s_{FM}(t)$, is given by $s_{FM}(t) = a_o \cos(2\pi f_i(t))$ where

$$f_i(t) = f_c + k_f x(t)$$

This can be integrated to give an instantaneous phase, so that we can write $s_{FM}(t)$ as

$$s_{FM}(t) = a_0 \cos\left(2\pi f_c t + 2\pi k_f \int_0^t x(\tau)d\tau\right)$$

FM is a non-linear modulation technique that distorts the signal and spreads it over a wider frequency bandwith than that of the original signal. This leads to FM being more robust to noise than AM. [5]

(b) When our information is digital we employ digital modulation techniques – again this is bookwork, and the following main points should be covered (noting that we are only covering the binary case):

The digital counterpart of analogue AM is *amplitude shift keying* (ASK, in this case BASK). Here the carrier signal is altered according to the following rule:

$$s_{ASK}(t) = \begin{cases} a\cos(2\pi f_c t) & \text{for information bit=1} \\ 0 & \text{for information bit=0} \end{cases}$$

The digital counterpart of analogue PM is *phase shift keying* (in this case BPSK). Here the carrier signal is altered according to the following rule:

$$s_{PSK}(t) = \begin{cases} a\cos(2\pi f_c t) & \text{for information bit=1} \\ a\cos(2\pi f_c t + \pi) & \text{for information bit=0} \end{cases}$$

The digital counterpart of analogue FM is *frequency shift keying* (in this case BFSK). Here the carrier signal is altered according to the following rule:

$$s_{FSK}(t) = \begin{cases} a\cos(2\pi f_c^1 t) & \text{for information bit=1} \\ a\cos(2\pi f_c^2 t) & \text{for information bit=0} \end{cases}$$

So that a change of bit value causes a flip between carrier frequencies. [4]

(c)(i) In a pulse amplitude modulation (PAM) scheme the pulse $p(t)$ must satisfy the following conditions:

$$p(t = 0) = 1 \quad \text{and} \quad p(t = \pm nT) = 0, \quad n = 1, 2, 3...$$

where T is the symbol period.

If $p(t)$ is a unit amplitude rectangular pulse of width T the result of applying PAM to the signal $[-A, A, A, A, -A, A, -A, -A, A, A]$ is shown in figure 4 [3]
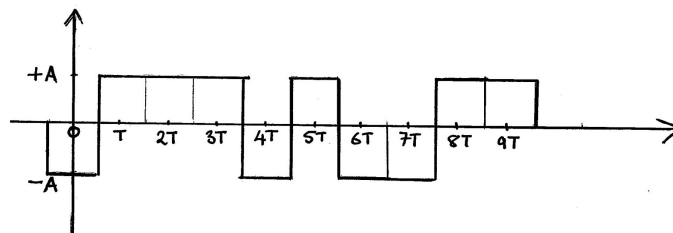

Figure 4: PAM signal

14

(c)(ii) The binary signal above is now transmitted using PSK with $f_c = 1/T$ – this means that the phase of the carrier will be shifted at every symbol, with the transmitted signal shown in figure 5 (note, OK to also have the negative of plot shown in figure 5) [3]
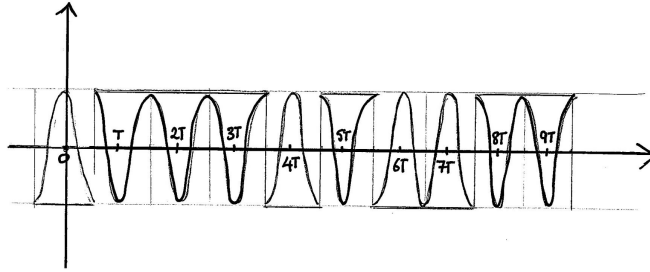

Figure 5: Transmitted PSK signal

(c)(iii) The probability of an error occurring is, P_e, where

$$P_e = p(\hat{X} = +A|X = -A)p(X = -A) + p(\hat{X} = -A|X = +A)p(X = +A)$$

where $\hat{X}$ is our estimate of X.

Therefore

$$P_e = \frac{1}{2}\left\{p(\hat{X} = +A|X = -A) + p(\hat{X} = -A|X = +A)\right\}$$

as both $+A$ and $-A$ are equally likely. If Y is our observation the above can be written as

$$P_e = \frac{1}{2}\left\{p(Y > 0|X = -A) + p(Y < 0|X = +A)\right\}$$

If $Y = X + Z$ where Z is the noise and is a zero mean Gaussian random variable with variance σ^2, the above condition reduces to

$$P_e = \frac{1}{2}\left\{p(Z > A) + p(Z < -A)\right\}$$

Since $Z/\sigma \sim N(0, 1)$ we can write this as

$$P_e = p(Z > A) = 1 - \Phi(A/\sigma) \equiv Q(a/\sigma)$$

[5]

15