Solutions: 4F8 2014

ENGINEERING TRIPOS PART IIB

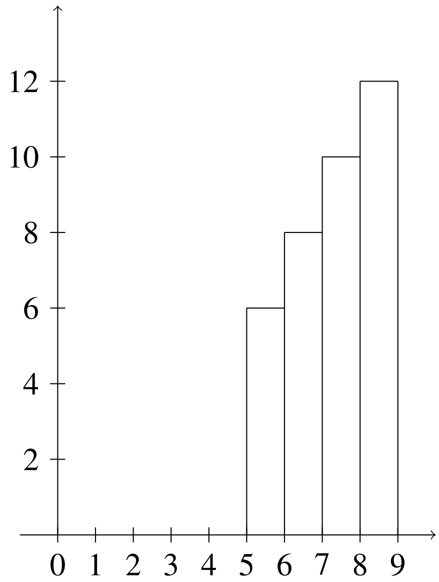
Tuesday 22 April 2014 9.30 to 11

Module 4F8

IMAGE PROCESSING AND IMAGE CODING

Version: JL/2

1 (a) (i) The histogram of the image in Figure 1 (of the paper) is shown below. We can see that the grey levels used are concentrated around the top of the range 1-18 and low and middle levels are unused. [10%]



[10%]

(ii) It helps to draw up a table when performing histogram equalisation: below let $H(i)$ be the frequency values and $C(i)$ be the cumulative frequency values

i	1	2	3	4	5	6	7	8	9
$H(i)$	0	0	0	0	0	6	8	10	12
$C(i)$	0	0	0	0	0	6	14	24	36

The transformed levels are given by

$$y_k = \sum_{i=1}^{k} L \frac{N_i}{NM}, \quad k = 1...9$$

where $N \times M$ are the dimensions of the image, N_i is the number of pixels in grey level i (equivalent to $H(i)$ above) and L is the range in grey level space. Therefore, $L = 9$, $NM = 36$ and

$$y_k = \frac{L}{NM} \sum_{i=1}^{k} N_i = \frac{1}{4} \sum_{i=1}^{k} N_i = \frac{1}{4} C(k), \quad k = 1...9$$

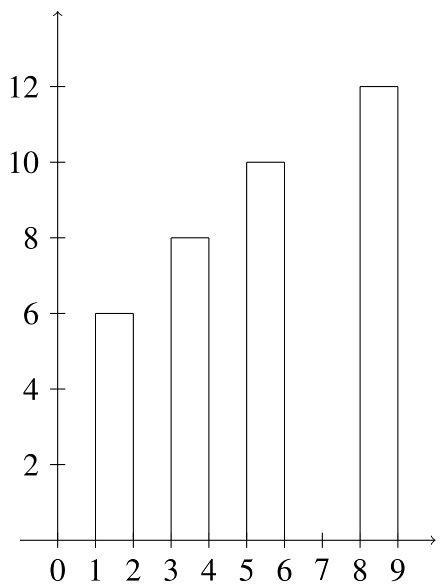
We can now add an extra line to our table to show the transformed values:

i	1	2	3	4	5	6	7	8	9
$H(i)$	0	0	0	0	0	6	8	10	12
$C(i)$	0	0	0	0	0	6	14	24	36
$y(i)$	0	0	0	0	0	1.5	3.5	6	9

If we decide to round up (so that $1.5 \to 2$ and $3.5 \to 4$ – but rounding down is acceptable if stated that this is what is being done) it is now easy to draw the new image and sketch the new histogram

9	6	4	4	6	9
4	9	6	6	9	4
2	6	9	9	6	2
2	9	6	6	9	2
9	6	4	4	6	9
9	4	2	2	4	9

Fig. 1



We can see from the new histogram that the process has succeeded in spreading out the grey levels more evenly across the scale but that the distribution is far from being uniform. The discreteness of the problem means that the equalisation process tries to do the best job it can according to the rules prescribed. [25%]

(iii) The spread of greylevels can be improved by interpolation after the histogram equalisation process. One simple interpolation rule would be simply to do a linear interpolation of the histogram above, replacing any zero values with the average of the values either side, so that our new H values have the following values (take $H(0) = 0$):

i	1	2	3	4	5	6	7	8	9
$H(i)$	3	6	7	8	9	10	10.67	11.33	12
$\frac{36}{77} \times H(i)$	1.4026	2.8052	3.2727	3.7403	4.2078	4.6753	4.9886	5.2971	5.6104
[int]	1	3	3	4	4	5	5	5	6

where the bottom two rows have been renormalised (so that they sum to 36). Other interpolations (if they are reasonable) are OK. [20%]

(b) (i) Take the Fourier transform of each side of the convolution equation to give:

$$Y(\omega_1, \omega_2) = H(\omega_1, \omega_2)X(\omega_1, \omega_2)$$

$$\therefore X(\omega_1, \omega_2) = \frac{Y(\omega_1, \omega_2)}{H(\omega_1, \omega_2)}$$

Now, if $H(\omega_1, \omega_2)$ has zeros, then the inverse filter, $1/H$, will have infinite gain. i.e. if $1/H$ is very large (or indeed infinite), small noise in the regions of the frequency plane where these large values of $1/H$ occur can be hugely amplified. To counter this we can threshold the frequency response, leading to the so-called, pseudo-inverse or generalised inverse filter $H_g(\omega_1, \omega_2)$ given by

$$H_g(\omega_1, \omega_2) = \begin{cases} \frac{1}{H(\omega_1,\omega_2)} & \frac{1}{|H(\omega_1,\omega_2|} < \gamma \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

or

$$H_g(\omega_1, \omega_2) = \begin{cases} \frac{1}{H(\omega_1,\omega_2)} & \frac{1}{|H(\omega_1,\omega_2|} < \gamma \\ \gamma\frac{|H(\omega_1,\omega_2|}{H(\omega_1,\omega_2)} & \text{otherwise} \end{cases} \tag{2}$$

Clearly for $\frac{1}{|H(\omega_1,\omega_2|} \geq \gamma$ in equation 2, the modulus of the filter is set as γ, whereas in previous equation it is set as 0.

Although the *generalised inverse filter* may perform reasonably well on *noiseless images*, the performance is unsatisfactory with even mildly noisy images due to the still significant noise gain at frequencies where $H(\omega_1, \omega_2)$ is relatively small. [20%]

(ii) An expression in terms of the relevant quantities is given below:

$$G(\omega) = \frac{H^*(\omega) P_{xx}(\omega)}{|H(\omega)|^2 P_{xx}(\omega) + P_{dd}(\omega)} \tag{3}$$

[10%]

(iii) The matrix form of the Wiener filter is given by:

$$W = (C^{-1} + L^T N^{-1} L)^{-1} L^T N^{-1}$$

A comparison with the form of the Wiener filter in (b)ii tells us that, $C \equiv P_{xx}$, $N \equiv P_{dd}$ and $L \equiv H$ (therefore $L^T \equiv H^*$). [15%]

Parts a)(i) and a)(ii) of this question were done well by almost all candidates (though many lost marks needlessly as they ignored the fact that the question asked for the new equalised image). Part a)(iii) was less well done – many gave one line answers and many, who gave a perfectly good interpolation, lost marks by interpolating and forgetting to normalise so that the total number of values was still 36.

Part b)(i) and (ii) were generally well done. Part (iii) was done well by most, but there were a good number who just substituted C, L and N straight into the Weiner filter and came up with the wrong answer.

2 (a) (i) Consider the ideal filter given in the fig.2 – one way to construct this is to say that the ideal frequency response of this filter, $H(\omega_1,\omega_2)$, can be written as

$$H(\omega_1,\omega_2) = H_1(\omega_1,\omega_2) + H_2(\omega_1,\omega_2)$$

where H_1 is a rectangular lowpass filter given by

$$H_1(\omega_1,\omega_2) = \begin{cases} 1 & \text{if } |\omega_1| < \Omega_{L1} \text{ and } |\omega_2| < \Omega_{L2} \\ 0 & \text{otherwise} \end{cases}$$

and H_2 is the following alternative ideal bandpass filter:

$$H_2(\omega_1,\omega_2) = H_a(\omega_1)\,H_b(\omega_2)$$

where $H_a(\omega_1)$ and $H_b(\omega_2)$ are ideal 1-D bandpass filters:

$$H_a(\omega_1) = \begin{cases} 1 & \text{if } \Omega_{L1} < |\omega_1| < \Omega_{U1} \\ 0 & \text{otherwise} \end{cases}$$

$$H_b(\omega_2) = \begin{cases} 1 & \text{if } \Omega_{L2} < |\omega_2| < \Omega_{U2} \\ 0 & \text{otherwise} \end{cases}$$

Thus our impulse response of H is written as the sum of the impulse responses of H_1 and H_2

$$h(n_1\Delta_1, n_2\Delta_2) = \frac{\Delta_1\Delta_2}{\pi^2}[\Omega_{L2}\,\Omega_{L1}\,\mathrm{sinc}(\Omega_{L2}n_2\Delta_2)\,\mathrm{sinc}(\Omega_{L1}n_1\Delta_1)]$$
$$+\frac{\Delta_1\Delta_2}{\pi^2}\,[\Omega_{U1}\mathrm{sinc}(\Omega_{U1}n_1\Delta_1)) - \Omega_{L1}\mathrm{sinc}(\Omega_{L1}n_1\Delta_1))]\times$$
$$[\Omega_{U2}\mathrm{sinc}(\Omega_{U2}n_2\Delta_2) - \Omega_{L2}\mathrm{sinc}(\Omega_{L2}n_2\Delta_2)]$$

which simplifies to

$$h(n_1\Delta_1, n_2\Delta_2) = \frac{\Delta_1\Delta_2}{\pi^2}[2\Omega_{L2}\,\Omega_{L1}\,\mathrm{sinc}(\Omega_{L2}n_2\Delta_2)\,\mathrm{sinc}(\Omega_{L1}n_1\Delta_1)+$$
$$\Omega_{U2}\,\Omega_{U1}\,\mathrm{sinc}(\Omega_{U2}n_2\Delta_2)\,\mathrm{sinc}(\Omega_{U1}n_1\Delta_1) - \Omega_{L2}\,\Omega_{U1}\,\mathrm{sinc}(\Omega_{L2}n_2\Delta_2)\,\mathrm{sinc}(\Omega_{U1}n_1\Delta_1) -$$
$$\Omega_{U2}\,\Omega_{L1}\,\mathrm{sinc}(\Omega_{U2}n_2\Delta_2)\,\mathrm{sinc}(\Omega_{L1}n_1\Delta_1)$$

The figure shows that the maximum frequency ranges are within the values of π/Δ_1 (for ω_1) and π/Δ_2 (for ω_2) – thus, the Nyquist criteria are satisfied and there will be no aliasing. [40%]

(ii) If $\Omega_{L1} \to 0$ and $\Omega_{L2} \to 0$ the equation above reduces to

$$h(n_1\Delta_1, n_2\Delta_2) = \frac{\Delta_1\Delta_2}{\pi^2}[\Omega_{U2}\,\Omega_{U1}\,\mathrm{sinc}(\Omega_{U2}n_2\Delta_2)\,\mathrm{sinc}(\Omega_{U1}n_1\Delta_1)$$

Which is indeed the impulse response of the following lowpass filter:

$$H_1(\omega_1,\omega_2) = \begin{cases} 1 & \text{if } \omega_1| < \Omega_{U1} \text{ and } \omega_2| < \Omega_{U2} \\ 0 & \text{otherwise} \end{cases}$$

[10%]

(b) (i) If we do not sample sufficiently frequently, we will have aliasing which will occur via frequencies from the repeated spectrum (the spectrum repeats at intervals of the sampling frequency – therefore we need the sampling frequency to be at least twice the highest frequency in the image, to avoid aliasing) falling into the 'main' spectrum. In many cases we will get aliased frequencies which will then lie very close to each other in the frequency domain. With two close frequencies, the effect will be to produce artefacts at the sum and difference of the two frequencies – it is generally the difference frequency [which is the lower frequency] which will then manifest itself as ringing/beating (ie moire fringe effects) artefacts in the aliased image.

[15%]

(ii) We can write s as a Fourier series:

$$s(u_1,u_2) = \sum_{p_1=-\infty}^{\infty}\sum_{p_2=-\infty}^{\infty} c(p_1,p_2)\,e^{j(p_1\Omega_1u_1+p_2\Omega_2u_2)}$$

where $\Omega_1 = \frac{2\pi}{\Delta_1}$ and $\Omega_2 = \frac{2\pi}{\Delta_2}$.

We can then find the Fourier coefficients c in the usual way:

$$c(p_1,p_2) = \frac{1}{\Delta_1\,\Delta_2}\int_{-\frac{\Delta_2}{2}}^{\frac{\Delta_2}{2}}\int_{-\frac{\Delta_1}{2}}^{\frac{\Delta_1}{2}} s(u_1,u_2)\,e^{-j(p_1\Omega_1u_1+p_2\Omega_2u_2)}\,du_1\,du_2$$

$$= \frac{1}{\Delta_1\,\Delta_2}\int_{-\frac{\Delta_2}{2}}^{\frac{\Delta_2}{2}}\int_{-\frac{\Delta_1}{2}}^{\frac{\Delta_1}{2}}\left[\sum_{n_1=-\infty}^{\infty}\sum_{n_2=-\infty}^{\infty}\delta(u_1-n_1\Delta_1,u_2-n_2\Delta_2)\right]$$

$$\times e^{-j(p_1\Omega_1 u_1 + p_2\Omega_2 u_2)}\, du_1\, du_2$$

$$\Longrightarrow c(p_1, p_2) = \frac{1}{\Delta_1\,\Delta_2} \quad \text{for all} \quad p_1, p_2$$

The sampled image may then be expressed as:

$$g_s(u_1, u_2) = g(u_1, u_2)\,\frac{1}{\Delta_1\,\Delta_2}\sum_{p_1=-\infty}^{\infty}\sum_{p_2=-\infty}^{\infty} \mathrm{e}^{j(p_1\Omega_1 u_1 + p_2\Omega_2 u_2)}$$

Using the frequency shift or spatial modulation theorem to take the Fourier transform

$$g(u_1, u_2)e^{j(p_1\Omega_1 u_1 + p_2\Omega_2 u_2)} \Leftrightarrow G(\omega_1 - \Omega_1 p_1, \omega_2 - \Omega_2 p_2)$$

gives:

$$G_s(\omega_1, \omega_2) = \frac{1}{\Delta_1\,\Delta_2}\sum_{p_1=-\infty}^{\infty}\sum_{p_2=-\infty}^{\infty} G(\omega_1 - p_1\Omega_1, \omega_2 - p_2\Omega_2)$$

It can therefore be seen that the Fourier transform or spectrum of the sampled 2d signal is the periodic repetition of the spectrum of the unsampled 2d signal – precisely analogous to the 1d case. It is therefore clear that for a bandlimited 2d signal, we must sample at more than twice the largest frequencies in the signal to keep these copies of the FT separate. Hence

$$\frac{2\pi}{\Delta_1} > 2\Omega_{B1} \quad \frac{2\pi}{\Delta_2} > 2\Omega_{B2}$$

These are the Nyquist frequencies, and if we sample below these we observe artefacts which we call aliasing.

[35%]

Parts a)(i) and (ii) were well done. A good number of candidates lost marks (not many) by leaving their answer in terms of normalised frequencies. Parts b)(i) and (ii) were also well done. This was the most popular question on the paper and the easiest (if you had done your revision).

3 (a)

$$T\,T^T = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

Thus, since $T^T = T^{-1}$, T is orthonormal.

Or say that:

A square orthonormal matrix has all rows orthogonal to each other ($\mathbf{t}_i \cdot \mathbf{t}_j = 0$ for $i \neq j$ – where $\mathbf{t}_i$ is the ith row of T, and each row vector is of unit length ($\mathbf{t}_i \cdot \mathbf{t}_i = 1$ for all i). Since T satisfies this, it is orthonormal.

If $\mathbf{y} = T\mathbf{x}$, then $|\mathbf{y}|^2 = \mathbf{y}^T\mathbf{y} = \mathbf{x}^T T^T T\mathbf{x} = \mathbf{x}^T\mathbf{x} = |\mathbf{x}|^2$ since $T^T T = I$. Thus the energy in $\mathbf{y}$ is the same as the energy in $\mathbf{x}$, and T therefore preserves total energy. [15%]

(b) Assume N is even. Since $\mathbf{y}_i = T\mathbf{x}_i$, we have that $C_{yy} = E[\mathbf{y}_i\mathbf{y}_i^T]$ so that

$$C_{yy} = E[\mathbf{y}_i\mathbf{y}_i^T] = E[T\mathbf{x}_i\mathbf{x}_i^T T^T] = TE[\mathbf{x}_i\mathbf{x}_i^T]T^T$$

So that

$$C_{yy} = \frac{\sigma^2}{2}T\begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}T^T = \sigma^2\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} 1 & \alpha \\ \alpha & 1 \end{bmatrix}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= 2\sigma^2\begin{bmatrix} (1+\alpha) & 0 \\ 0 & (1-\alpha) \end{bmatrix}$$

The diagonal terms of this covariance matrix represent the variance of the two input elements $\{\mathbf{y}_i\}_1$ and $\{\mathbf{y}_i\}_2$. Thus the energy of the $N/2$ coefficients $\{\mathbf{y}_i\}_1$ is

$$\sigma^2(1+\alpha)\frac{N}{2}$$

and the energy of the $N/2$ coefficients $\{\mathbf{y}_i\}_2$ is

$$\sigma^2(1-\alpha)\frac{N}{2}$$

Note that there is no DC energy as the pixels of $\mathbf{x}$ have zero mean and hence their sums and differences, $\{\mathbf{y}_i\}_1$ and $\{\mathbf{y}_i\}_1$ also have zero mean: recall

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} x_1+x_2 \\ x_1-x_2 \end{bmatrix}$$

[25%]

(c) $\mathbf{y} = T\mathbf{x}$ transforms the column vector $\mathbf{x}$ to the column vector $\mathbf{y}$. Therefore TX, where X is a 2×2 matrix, transforms the 2 columns of X.

Similarly, $\mathbf{y}^T = \mathbf{x}^T T^T$ transforms the row vector $\mathbf{x}^T$ to the row vector $\mathbf{y}^T$. Therefore XT^T, transforms the 2 rows of X.

Thus, $Y = TXT^T$ transforms both the rows and columns of X.

To invert the transform we simply take $X = T^T YT$ since T is orthonormal. [25%]

(d) To perform 1st-level of transformation on a big image $\hat{X}$, we calculate $Y_{ij} = TX_{ij}T^T$ for each 2×2 block of pixels X_{ij} for $i = 1,..,768/2 = 384$ and for $j = 1,...,1024/2 = 512$. We then form 4 sub-images, each of size 384×512, from the top left, top right, lower left and lower right elements of all the coefficient blocks Y_{ij}. The top left subimage is the result of averaging or lowpass filtering of the input and still has strong correlations between adjacent coefficients; we therefore apply the 2nd-level transform to this subimage. The remaining 3 subimages, each of size 384×512, are the outputs from level 1.

Similarly, there are 3 subimages of size $384/2 \times 512/2 = 192 \times 256$, which are the outputs of level 2. We then apply the 3rd-level transform to the top left level-2 subimage. This generates 3 subimages of size $192/2 \times 256/2 = 96 \times 128$, and a 4th lowpass subimage also of size 96×128, containing all the low frequency components of the image.

Given this process, the total number of output coefficients is given by:

$$3 \times 384 \times 512 + 3 \times 192 \times 256 + 4 \times 96 \times 128 = 96 \times 128[3 \times 4 \times 4 + 3 \times 2 \times 2 + 4 \times 1 \times 1]$$

$$= 96 \times 128[48 + 12 + 4) = 96 \times 128 \times 64 = 768 \times 1024$$

which is equal to the number of input pixels. Thus the transform is non-redundant (ie a 1-1 mapping). [35%]

Part (a) was well done by almost all candidates who attempted it. Part (b) was less well done, with almost no fully correct answers. Candidates clearly did not understand what was being asked for in 'estimating the energy in the two subbands' and very few noted why it was necessary for the input pixels to come from a pdf with zero mean. Part (c) was well done, but a good number of candidates gave remarkably sketch answers even though they knew the material. Part (d) was well done by almost all who attempted it.

Version: JL/2

4 (a) See figure 2 below. [15%]

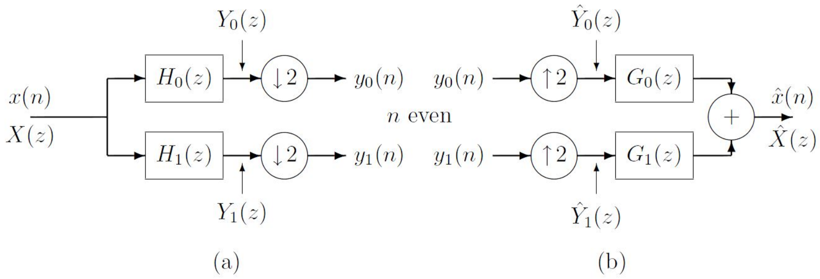


Fig. 2

(b) Normally $H_0(z)$ and $G_0(z)$ are lowpass filters which pass a lot more of the signal energy than H_1 and G_1, which are highpass filters. To get good energy compaction into as few coefficients as possible, we repeatedly split the lowpass outputs with further levels of band-splitting, at lower and lower sample rates. A 4-level transform is shown in figure 3 below.

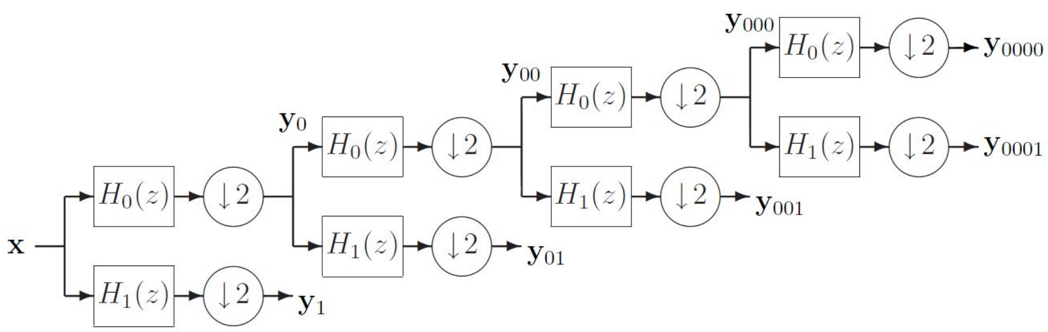


Fig. 3

The inverse transform is the reverse of this, using G_0 and G_1 instead of H_0 and H_1. [20%]

(c) From figure 2 we can analyse the 2-band filterbank system:

$$\hat{X}(z) = G_0(z)\hat{Y}_0(z) + G_1(z)\hat{Y}_1(z)$$

$$= G_0(z)\frac{1}{2}(Y_0(z) + Y_0(-z)) + G_1(z)\frac{1}{2}(Y_1(z) + Y_1(-z))$$

But $Y_0(z) = H_0(z)X(z)$ and $Y_1(z) = H_1(z)X(z)$, so that

$$\hat{X}(z) = \frac{1}{2}G_0(z)\left(H_0(z)X(z) + H_0(-z)X(-z)\right) + \frac{1}{2}G_1(z)\left(H_1(z)X(z) + H_1(-z)X(-z)\right)$$

$$= \frac{1}{2}X(z)\left(G_0(z)H_0(z) + G_1(z)H_1(z)\right) + \frac{1}{2}X(-z)\left(G_0(z)H_0(-z) + G_1(z)H_1(-z)\right)$$

For P-R we require that $\hat{X}(z) = X(z)$, and therefore the $X(-z)$ term must become zero and the $X(z)$ term must be $\frac{1}{2}X(z)2$. Therefore

$$G_0(z)H_0(-z) + G_1(z)H_1(-z) = 0$$

and

$$G_0(z)H_0(z) + G_1(z)H_1(z) = 2$$

as required. [25%]

(d) Substituting the given expressions for H_1 and G_1 into the LHS of the first equation above, we have

$$G_0(z)H_0(-z) + G_1(z)H_1(-z) = G_0(z)H_0(-z) + zH_0(-z)(-z)^{-1}G_0(z)$$

$$= G_0(z)H_0(-z)\left(1 + \frac{z}{-z}\right) = 0$$

so the first condition is satisfied.

If $Z = \frac{1}{2}(z + z^{-1})$ then $-Z = \frac{1}{2}((-z) + (-z)^{-1})$.

Now

$$H_0(z)G_0(z) + H_1(z)G_1(z) = H_0(z)G_0(z) + z^{-1}G_0(-z)zH_0(-z)$$

$$= H_0(z)G_0(z) + G_0(-z)H_0(-z) = P(z) + P(-z) = 2$$

where $P(z) = H_0(z)G_0(z)$.

In terms of Z, let $P_t(Z) = P(z)$, so that $P_t(-Z) = P(-z)$, because Z contains only *odd powers* of z. Hence, for P-R, we require $P_t(Z) + P_t(-Z) = 2$. Now all the odd powers of Z will become zero when $P_t(-Z)$ is added to $P_t(Z)$, so the only constraint for P-R is that the even powers of Z in $P_t(Z)$ should be zero except for the Z^0 term which should be equal to unity.

Hence

$$\begin{aligned} P_t(Z) &= (1+Z)(1-\frac{2}{7}Z)(1+Z)(1+aZ+bZ^2) \\ &= \frac{1}{7}(1+Z)^2(7-2Z)(1+aZ+bZ^2) \\ &= \frac{1}{7}(7+12Z+3Z^2-2Z^3)(1+aZ+bZ^2) \end{aligned}$$

This is a 5th order polynomial in Z, so for P-R we require that the coefficients in Z^2 and Z^4 be zero and the coefficient in Z^0 is unity.

The coefficient in Z^0 is indeed unity. The coefficient in Z^2 is $\frac{1}{7}(7b+2a+3)$ and the coefficient in Z^4 is $\frac{1}{7}(3b-2a)$. Setting these to zero therefore gives:

$$b = \frac{2}{3}a \quad \text{and} \quad 7b + 12a = -3$$

therefore

$$a = -\frac{9}{50} \quad \text{and} \quad b = -\frac{3}{25}$$

[40%]

Parts a) and b) were essentially bookwork and were done well by most who attempted them (though marks were deducted in part b) for answers which gave a diagram only and no written explanation). Part c) was also answered well. Part d) was the part where most marks were lost – some candidates made only a token attempt, while many made algebraic errors when multiplying out products of polynomials.

END OF SOLUTIONS