SYSTEM IDENTIFICATION, Exam of May 2003, Solutions

Question 1. (a) A state space realization of $\mathbf{P}$ is $\dot{x} = -\frac{1}{T}x + \frac{1}{T}u$, $w = cx$. This, together with the A/D and D/A blocks, gives the exact discretization of the plant, described by

$$x_{k+1} = e^{-h/T} x_k + (e^{-h/T} - 1)(-T) \cdot \frac{1}{T}(p_k + d)$$
$$= e^{-h/T} x_k + (1 - e^{-h/T})(p_k + d), \quad y_k = cx_k,$$

which has the transfer function

$$\mathbf{P}^d(z) = c\left(z - e^{-h/T}\right)^{-1}\left(1 - e^{-h/T}\right).$$

Denoting $a = e^{-h/T}$ and $b = c(1 - e^{-h/T})$, we obtain $\mathbf{P}^d(z) = b/(z-a)$. This is stable for $|a| < 1$, i.e., for all $T > 0$.

(b) Since $\mathbf{C}$ is known and stable, we can compute

$p_k = c_0(r_k - y_k) + c_1(r_{k-1} - y_{k-1}) \ldots + c_9(r_{k-9} - y_{k-9})$.

The first nine values $p_1, \ldots p_9$ may be affected by the (unknown) initial state of the controller, afterwards the initial state has no influence. Since we have 2000 measurements, the effect of the initial state of the controller will be small. From $y_k - a y_{k-1} = b(p_{k-1} + d) + e_k$ we get, with $\theta^T = [a \;\; b \;\; bd]$,

$$y_k = [y_{k-1} \;\; p_{k-1} \;\; 1]\,\theta + e_k.$$

(c) Denoting $\varphi_k = [y_{k-1} \ p_{k-1} \ 1]$ and

$$\Phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_{2000} \end{bmatrix}, \quad \Phi^{\#} = (\Phi^* \Phi)^{-1} \Phi^*,$$

the vector θ which minimizes $\|e\|^2 = |e_1|^2 + |e_2|^2 \ldots + |e_{2000}|^2$ is given by $\hat{\theta} = \Phi^{\#} y$, where $y = [y_1 \ y_2 \ \ldots \ y_{2000}]^T$. $\hat{\theta}$ is an unbiased estimate of θ, if $E(e_k) = 0$.

(d) The white noise r will lead to the smallest $\mathrm{Cov}\,\hat{\theta}$, because it subjects the plant to a wide range of frequencies. The possibilities (i) and (ii) test the system at one frequency only. (We remark that (ii) will work better than (i), i.e., will lead to smaller $\mathrm{Cov}\,\hat{\theta}$, because a sinusoidal signal has both amplitude and phase information, while a constant has only an amplitude. Moreover, for constant r, we cannot distinguish between the effect of r and the effect of d). For $\widehat{\mathrm{Var}}\, e_k$, the picture is very different: we expect it to be smallest for constant r_k. Indeed, for constant r_k, since the whole feedback system is stable, all signals will converge to constants. Then the equation $y_k = \varphi_k \theta + e_k$ can be satisfied with a suitable choice of θ such that practically $e_k = 0$.

(e)
$$P^d(z) = \frac{b z^{-1}}{1 - a z^{-1}} = b z^{-1} (1 + a z^{-1} + a^2 z^{-2} + \ldots),$$

$$P^d(z) \approx b z^{-1} + b a z^{-2} + b a^2 z^{-3} \ldots + b a^9 z^{-10},$$

where $a = e^{-h/T}$, $b = c(1 - e^{-h/T})$.

Question 2. (a) The factor z^{-m} represents a delay by m sampling periods, so that the impulse response of the model G^d is zero for $k = 0, 1, \dots m-1$. Thus, mh should be approximately the time when the second ball starts to move ($m \approx 100$).

(b) The structure of G^d implies

$$y_k + a_1 y_{k-1} + a_2 y_{k-2} = b_0 u_{k-m} + b_1 u_{k-m-1} + b_2 u_{k-m-2} + e_k$$

where e_k are the equation errors due to the imperfection of the model. For $k = 1, 2, \dots m-1$ the equation is irrelevant (both sides should be zero). Denoting $\theta = [a_1 \ a_2 \ b_0 \ b_1 \ b_2]^T$, we rewrite the equations as

$$y_k = \underbrace{\left[-y_{k-1} \ -y_{k-2} \ u_{k-m} \ u_{k-m-1} \ u_{k-m-2}\right]}_{\varphi_k} \theta + e_k$$

for $k = m, m+1, \dots 300$.

Denoting

$$y = \begin{bmatrix} y_m \\ y_{m+1} \\ \vdots \\ y_{300} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \varphi_m \\ \varphi_{m+1} \\ \vdots \\ \varphi_{300} \end{bmatrix}, \quad e = \begin{bmatrix} e_m \\ e_{m+1} \\ \vdots \\ e_{300} \end{bmatrix},$$

and assuming that $\Phi^* \Phi$ is invertible, the vector $\hat{\theta}$ which minimizes $\|e\|^2 = |e_m|^2 + \dots |e_{300}|^2$ is given by $\hat{\theta} = (\Phi^*\Phi)^{-1}\Phi^* y$. The corresponding optimal e is given by $\hat{e} = y - \Phi\hat{\theta} = [I - \Phi(\Phi^*\Phi)^{-1}\Phi^*]y$.

(c) The best m is the one which minimizes the corresponding $\|\hat{e}\|^2$. Even better, we should take into account also e_k for k slightly less than m, since for such k, y_k may already be nonzero.

Thus, we may define $\hat{e}_k = y_k - \varphi_k \hat{\theta}$ for all $k = 1, 2, \dots 300$ and then choose m for which $|\hat{e}_1|^2 + |\hat{e}_2|^2 \dots + |\hat{e}_{300}|^2$ is minimal (recall that $\hat{\theta}$ depends on m). We will have $\hat{e}_1 = \hat{e}_2 = \dots \hat{e}_{90} = 0$ in practice, because of the finite propagation speed of the waves.

(d) We should take $\tau = mh$, because the factor $e^{-\tau s}$ represents a delay by τ. The factor G_0 could be estimated by the Tustin transformation, substituting $z^{-1} = \dfrac{1 - hs/2}{1 + hs/2}$ into

$$G_0^d(z) = \frac{b_0 + b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} .$$

(e) One way would be to use averaging, i.e.,

$$y_k^{ave} = (y_k + \tilde{y}_k)/2$$

and then use y_k^{ave} instead of y_k in the whole identification process. If the measurements are affected by noise, then averaging would reduce the standard deviation of the measurement noise $1/\sqrt{2}$ times.

Another possibility is to consider both sets of equations $y_k = \varphi_k \theta + e_k$, $\tilde{y}_k = \tilde{\varphi}_k \theta + \tilde{e}_k$, $k = m, m+1, \dots 300$, where

$$\tilde{\varphi}_k = [-\tilde{y}_{k-1} \ -\tilde{y}_{k-2} \ \ u_{k-m} \ u_{k-m-1} \ \ u_{k-m-2}]$$

so that we have $2(300 - m + 1)$ equations.

Question 3. (a) $S(s) = \left[1 + P(s)C(s)\right]^{-1}$

$$= \frac{s^2 + (100\pi)^2}{\left[s^2+(100\pi)^2\right]\cdot\left[1 + k_1 P(s)\right] + k_2 P(s) s}.$$

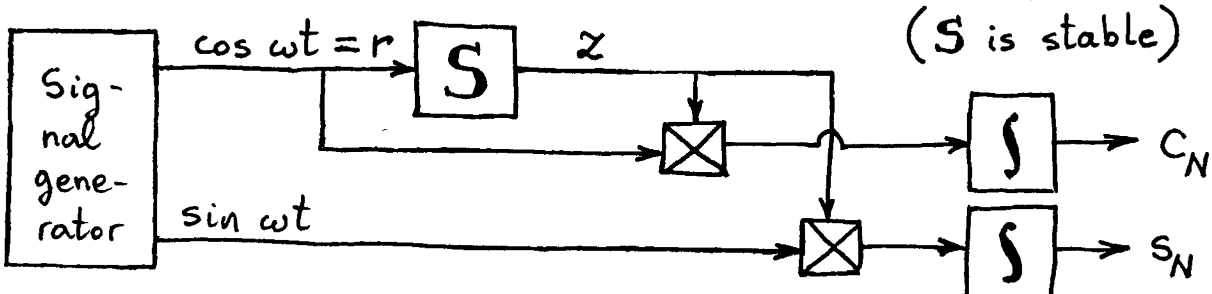
If $d=0$ and $r(t) = R\cos\omega t$, then for large t we have $z(t) \approx Z\cos(\omega t + \varphi_\omega)$, where

$$Z = |S(i\omega)|\cdot R, \qquad \varphi_\omega = \arg S(i\omega).$$

For 50Hz we have $\omega = 100\pi$, so that if $k_2 P(i100\pi) \neq 0$ then $S(i\omega) = 0$. In this particular case $z(t) \to 0$, which is called asymptotic tracking.

(b) The identification experiments could be done separately for a number of different frequencies ω, as follows (making $d=0$):

(S is stable)



Denoting $T = \frac{2\pi}{\omega}$, we compute for large $t_0 > 0$

$$C_N = \int_{t_0}^{t_0+NT} z(t)\cos\omega t\, dt = A_\omega \cos\varphi_\omega \frac{NT}{2},$$

$$S_N = \int_{t_0}^{t_0+NT} z(t)\sin\omega t\, dt = A_\omega \sin\varphi_\omega \frac{NT}{2}.$$

From here we can compute A_ω and φ_ω, and hence $S(i\omega) = A_\omega e^{i\varphi_\omega}$.

Once we have estimated $\mathbf{S}(i\omega)$, we can obtain

$$\mathbf{P}(i\omega) = \frac{1}{\mathbf{C}(i\omega)}\left(\mathbf{S}^{-1}(i\omega) - 1\right).$$

This formula will lead to large errors if $\mathbf{C}(i\omega)$ is very small (in which case $\mathbf{S}(i\omega) \approx 1$), i.e., a small relative error in estimating $\mathbf{S}(i\omega)$ will be amplified and become a large relative error in the estimate of $\mathbf{P}(i\omega)$. If $\mathbf{C}(i\omega)$ is very large, then the signal z will be very small, and this will increase the likelihood of errors corrupting the measurement of $\mathbf{S}(i\omega)$. To avoid these problems, we can choose $k_1 = 1$ and $k_2 = 0$ (these values are stabilizing).

(c) Denote $w = \begin{bmatrix} r \\ d \end{bmatrix}$. z is obtained by passing the stationary signal w through a stable LTI system with transfer function $G = [\mathbf{S}\ \ -\mathbf{PS}]$, hence z is stationary. Since $E(w) = 0$, we have $E(z) = 0$.

We have $\quad Var(z) = C_{zz}(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{zz}(i\omega)\, d\omega.$

S_{zz} can be obtained from the matrix version of the Wiener-Lee formula: $S_{zz} = \overline{G}\, S_{ww}\, G^T$, where $S_{ww} = \begin{bmatrix} S_{rr} & 0 \\ 0 & S_{dd} \end{bmatrix}$ (the zeros in S_{ww} appear because r, d are independ.).

(d) From $\hat{z}(i\omega) = \mathbf{S}(i\omega)\hat{r}(i\omega) - \mathbf{P}(i\omega)\mathbf{S}(i\omega)\hat{d}(i\omega)$, since $\mathbf{S}(100\pi i) = 0$, the first term is zero (since time started at $-\infty$, we are already in steady state), and hence $E(z) = -\mathbf{P}(0)\mathbf{S}(0)E(d) = 0$. Since $z = r - y$, we conclude that $E(y(t)) = r(t)$.

Question 4. (a) We denote by z the voltage on the capacitor C_2 (from top to bottom), so that $y = kz$. We denote by v the voltage on the capacitor C_1 (from left to right). Then the currents through the capacitors are

$$C_1 \dot{v} = \frac{u-v}{R} + \frac{u-z-v}{R},$$

$$C_2 \dot{z} = \frac{u-z-v}{R}.$$

The state vector should be $x = \begin{bmatrix} v \\ z \end{bmatrix}$.

Then

$$\underbrace{\frac{d}{dt}\begin{bmatrix} v \\ z \end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix} -\frac{2}{RC_1} & -\frac{1}{RC_1} \\ -\frac{1}{RC_2} & -\frac{1}{RC_2} \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} v \\ z \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} \frac{2}{RC_1} \\ \frac{1}{RC_2} \end{bmatrix}}_{B} u,$$

$$y = \underbrace{\begin{bmatrix} 0 & k \end{bmatrix}}_{C} \cdot \begin{bmatrix} v \\ z \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \end{bmatrix}}_{D} \cdot u$$

We have $\det A = \frac{2}{R^2 C_1 C_2} - \frac{1}{R^2 C_1 C_2}$

$= \frac{1}{R^2 C_1 C_2} > 0$. Since $\operatorname{trace} A < 0$ and $\det A > 0$, A is stable.

(b) The characteristic polynomial of A is

$$p(s) = \det(sI - A) = s^2 - (\operatorname{trace} A)s + \det A$$

$$= s^2 + \underbrace{\frac{1}{R}\left(\frac{2}{C_1} + \frac{1}{C_2}\right)}_{a_1} s + \underbrace{\frac{1}{R^2 C_1 C_2}}_{a_0} .$$

We have $G(s) = C(sI - A)^{-1} B =$

$$= \begin{bmatrix} 0 & k \end{bmatrix} \frac{1}{p(s)} \begin{bmatrix} s + 1/RC_2 & -1/RC_1 \\ -1/RC_2 & s + 2/RC_1 \end{bmatrix} \begin{bmatrix} 2/RC_1 \\ 1/RC_2 \end{bmatrix}$$

$$= \frac{k}{p(s)} \begin{bmatrix} -\frac{1}{RC_2} & s + \frac{2}{RC_1} \end{bmatrix} \begin{bmatrix} \frac{2}{RC_1} \\ \frac{1}{RC_2} \end{bmatrix}$$

$$= \frac{k}{p(s)} \begin{bmatrix} -\nu_2 & s + \nu_1 \end{bmatrix} \begin{bmatrix} \nu_1 \\ \nu_2 \end{bmatrix} = \frac{k s \nu_2}{p(s)} ,$$

where we have used the notation

$$\nu_1 = \frac{2}{RC_1} , \qquad \nu_2 = \frac{1}{RC_2} .$$

Thus,

$$G(s) = \frac{\overbrace{k \nu_2}^{b_1} s}{s^2 + \underbrace{(\nu_1 + \nu_2)}_{a_1} s + \underbrace{0.5\, \nu_1 \nu_2}_{a_0}} .$$

We clearly have

$$\lim_{\omega \to 0} G(i\omega) = \lim_{\omega \to \infty} G(i\omega) = 0 .$$

(c) We denote by $G^e(i\omega_k)$ the experimentally determined values of the true transfer function, $k=1,\dots 50$, which are subject to measurement errors. We have

$$b_1 i\omega_k = \left[(i\omega_k)^2 + a_1(i\omega_k) + a_0\right] G^e(i\omega_k) + e_k$$

where e_k is due to the combined effect of the measurement and modeling errors. We rewrite this:

$$\underbrace{\omega_k^2 G^e(i\omega_k)}_{y_k} = \underbrace{\left[i\omega_k G^e(i\omega_k) \quad G^e(i\omega_k) \quad -i\omega_k\right]}_{\varphi_k} \theta + e_k$$

where $\theta^T = \left[a_1 \;\; a_0 \;\; b_1\right]$ are the unknown parameters, while y_k and φ_k are known. Denoting

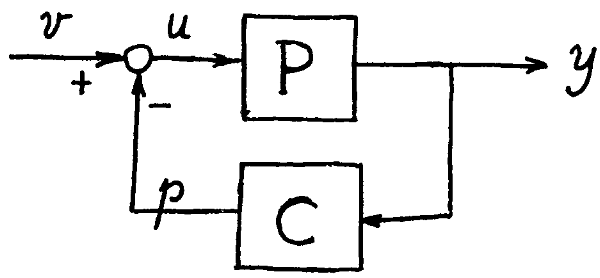
$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_{50} \end{bmatrix}, \quad \phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_{50} \end{bmatrix} \in \mathbb{C}^{50\times 3}, \quad e = \begin{bmatrix} e_1 \\ \vdots \\ e_{50} \end{bmatrix},$$

we obtain the usual equation $y = \phi\theta + e$. However, ϕ and y are complex, while we are searching for the optimal <u>real</u> θ. Thus we decompose into the 2×50 equations

$$\begin{cases} \operatorname{Re} y = (\operatorname{Re}\phi)\theta + \operatorname{Re} e \\ \operatorname{Im} y = (\operatorname{Im}\phi)\theta + \operatorname{Im} e \end{cases}$$

Denoting $\tilde{y} = \begin{bmatrix} \operatorname{Re} y \\ \operatorname{Im} y \end{bmatrix}$, $\tilde{\phi} = \begin{bmatrix} \operatorname{Re}\phi \\ \operatorname{Im}\phi \end{bmatrix}$, $\tilde{e} = \begin{bmatrix} \operatorname{Re} e \\ \operatorname{Im} e \end{bmatrix}$, we arrive at the real equation $\tilde{y} = \tilde{\phi}\theta + \tilde{e}$. Now we can find the optimal (real) θ by the least squares algorithm: $\hat{\theta} = \tilde{\phi}^{\#} y$, where $\tilde{\phi}^{\#} = (\tilde{\phi}^*\tilde{\phi})^{-1}\tilde{\phi}^*$.
From a_1, a_0, b_1 we can compute ν_1, ν_2 and k, and hence also C_1, C_2.

Question 5. (a)



$$P(z) = \frac{1-0.8z^{-1}}{1-0.9z^{-1}}$$

$$C(z) = \delta(1-z^{-1})$$

$$G(z) = P(z)\left[1+P(z)C(z)\right]^{-1}$$

$$= \frac{1-0.8z^{-1}}{1-0.9z^{-1}+(1-0.8z^{-1})\,\delta\,(1-z^{-1})}$$

$$G(z) = \frac{1-0.8z^{-1}}{1+\delta-0.9(1+2\delta)z^{-1}+0.8\delta z^{-2}} .$$

(b) For $\delta = -1$ the highest order term in the denominator disappears, so that **G** is not proper. Hence, **G** is not stable (even though its pole is in the unit disk).

For $\delta = 0$ the zero of the denominator $1-0.9z^{-1}$ is at 0.9, so **G** is stable.

For $\delta = 1$ the zeros of the denominator $2-2.7z^{-1}+0.8z^{-2}$ are in the open unit disk, so that again **G** is stable.

(c) The DC-gain of **G** is

$$\mathbf{G}(1) = P(1)\left[1+P(1)C(1)\right]^{-1} = P(1)$$

(because $C(1)=0$). Since P does not depend on δ, we get $\mathbf{G}(1) = P(1) = 2$, regardless of δ.

(d) If h denotes the step response of $\mathbf{G}$, then its $\mathcal{Z}$-transform is $\hat{h}(z) = \mathbf{G}(z)\frac{z}{z-1}$. The poles of $\mathbf{G}$ are the zeros λ_1, λ_2 of $z^2 - 1.35z + 0.4$ (see the case $\delta = 1$ discussed at part (b)) which are in the open unit disk. We use the partial fractions decomposition

$$\frac{\mathbf{G}(z)}{z-1} = \frac{c_1}{z-\lambda_1} + \frac{c_2}{z-\lambda_2} + \frac{c_3}{z-1},$$

where $c_1 = \lim\limits_{z\to\lambda_1} (z-\lambda_1)\frac{\mathbf{G}(z)}{z-1}$, and c_2 is computed similarly. Then

$$\hat{h}(z) = c_1\frac{z}{z-\lambda_1} + c_2\frac{z}{z-\lambda_2} + c_3\frac{z}{\lambda-1},$$

whence

$$h_k = c_1(\lambda_1)^k + c_2(\lambda_2)^k + c_3, \quad k=0,1,\ldots$$

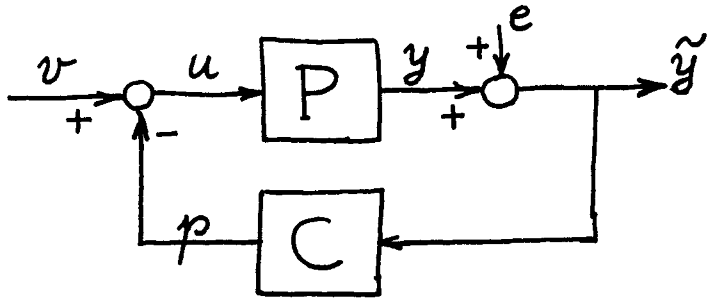
in particular $h_k \to c_3$. Hence, c_3 must be the DC-gain of $\mathbf{G}$, which is 2, see part (c).

(e) $$\mathbf{P}^c(s) \approx \mathbf{P}\left(\frac{1+\frac{hs}{2}}{1-\frac{hs}{2}}\right), \quad \text{where } h = \frac{1}{2000}$$

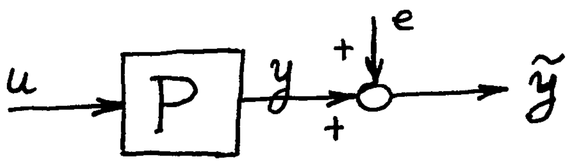
and $\mathbf{P}$ has been computed in part (a). This follows from Tustin's formula. Thus,

$$\mathbf{P}^c(s) \approx \frac{1+\frac{hs}{2} - 0.8\left(1-\frac{hs}{2}\right)}{1+\frac{hs}{2} - 0.9\left(1-\frac{hs}{2}\right)}.$$

(f) The block diagram with the measurement errors is :



Since $\tilde{y}_k$ and v_k are available and C is known, we can compute u_k, for $k = 1, \dots 100$. Now we can forget about the feedback configuration and consider only the following subsystem:



We know from prediction theory that if the error is added to the output, as it is here, then $\tilde{y}_0, \tilde{y}_1, \dots \tilde{y}_{100}$ are not relevant for the prediction of $\tilde{y}_{101}$. The unbiased prediction $\bar{\tilde{y}}_{101}$ is obtained by processing u_k $(k=0,1,\dots 100)$ through the filter P. From $\tilde{y}_k = \bar{\tilde{y}}_k + e_k$ (also from prediction theory) we see that the prediction error is e_k, so that its variance is 1. Note that $\bar{\tilde{y}}$ is in fact the uncorrupted output y of P $(\bar{\tilde{y}}_k = y_k)$, but y_k cannot be measured.

(Note that $\tilde{y}_0, \tilde{y}_1, \dots \tilde{y}_{100}$ have been used at the beginning, to compute $u_0, u_1, \dots u_{100}$.)

a model (or a copy) of

Question 6. (a) Since $\ldots u_{-1}, u_0, u_1, u_2, \ldots$ are i.i.d. (independent and identically distributed), the same is true for $\ldots u_{-1}^2, u_0^2, u_1^2, u_2^2, \ldots$ (or any other function of u). Thus, u^2 is white noise, in particular, it is stationary and ergodic. Based on ergodicity, we can estimate (a hat denotes an estimate)

$$\widehat{E}(u_k^2) = \frac{1}{N+1}\sum_{k=0}^{N} u_k^2 , \quad \widehat{Var}(u_k^2) = \frac{1}{N}\sum_{k=0}^{N}\left(\widetilde{u_k^2}\right)^2,$$

where $N = 10{,}000$ and $\widetilde{u_k^2} = u_k^2 - \widehat{E}(u_k^2)$. Since u^2 is white noise, its autocorrelation function is estimated as

$$\widehat{C}_\tau^{u^2u^2} = \widehat{Var}(u_k^2)\cdot\delta_0(\tau),$$

where δ_0 is the discrete unit pulse, $\delta_0(\tau) = \begin{cases} 1 & \text{for } \tau = 0, \\ 0 & \text{else.}\end{cases}$

(b) Since u^2 is white noise, it is ergodic. Hence, the random signal $[u^2 \; p \; q]^T$, where q is the output signal of $\mathbf{H}$, is also ergodic. Since $y = q + e$, where e is again white noise, independent of u^2, we get that also $[u^2 \; p \; y]^T$ is (stationary and) ergodic. In particular, it follows that we can estimate $C_\tau^{u^2p}$ and C_τ^{py} (for $\tau = 0, 1, \ldots 100$) by

$$\widehat{C}_\tau^{u^2p} = \frac{1}{N-\tau+1}\sum_{k=0}^{N-\tau} \widetilde{u_k^2}\,\tilde{p}_{k+\tau} , \quad \widehat{C}_\tau^{py} = \frac{1}{N-\tau+1}\sum_{k=0}^{N-\tau} \tilde{p}_k\,\tilde{y}_{k+\tau},$$

where $\widetilde{u_k^2}$ has been defined earlier, and similarly $\tilde{p}_k = p_k - \widehat{E}(p_k)$, $\tilde{y}_k = y_k - \widehat{E}(y_k)$. (Obviously, y is ergodic, being a component of $[u^2 \; p \; y]^T$.)

(c) Since $C^{u^2p} = g * C^{u^2u^2}$ and

$$C^{u^2u^2} = \sigma^2 \delta_0, \quad \text{where } \sigma^2 = \mathrm{Var}(u_k^2),$$

we get $\boxed{C^{u^2p} = \sigma^2 \cdot g}$. We can estimate both C^{u^2p} and σ^2 (see the answer to parts (a) and (b)), and then we can obtain an estimate of g from the framed formula above.

(d) Denote again by q the output signal of the subsystem with transfer function $\mathbf{H}$. Then

$C^{pq} = h * C^{pp}$. Since $y = q + e$, we have

$C^{py} = C^{pq} + C^{pe}$. Since p is obtained from u and e is independent of u, we have that e is independent of p, hence $C^{pe} = 0$. Thus, $C^{py} = h * C^{pp}$, or, written in matrix form,

$$\begin{bmatrix} C_0^{pp} & C_1^{pp} & C_2^{pp} & \cdots \\ C_1^{pp} & C_0^{pp} & C_1^{pp} & \cdots \\ C_2^{pp} & C_1^{pp} & C_0^{pp} & \cdots \\ \vdots & \vdots & \vdots & \end{bmatrix} \cdot \begin{bmatrix} h_0 \\ h_1 \\ h_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} C_0^{py} \\ C_1^{py} \\ C_2^{py} \\ \vdots \end{bmatrix}.$$

Since $\mathbf{H}$ is stable, we have $h_k \to 0$. We truncate the above infinite system of equations by keeping only the first 101 equations and the first 101 unknowns $h_0, h_1, \ldots h_{100}$. We replace C_τ^{pp} and C_τ^{py} (which are not known) by their estimates, see part (b), and then we solve for $h_0, h_1, \ldots h_{100}$.

(e) The approximate transfer function is

$$G^a(z) = \hat{g}_0 + \hat{g}_1 z^{-1} + \hat{g}_2 z^{-2} \dots + \hat{g}_{100}\, z^{-100}.$$

This corresponds to the MA equation

$$p_k = \hat{g}_0\, v_k + \hat{g}_1 v_{k-1} + \hat{g}_2 v_{k-2} \dots + \hat{g}_{100}\, v_{k-100}.$$

Here, v denotes the input signal of the FIR filter, p is the output signal, and $\hat{g}_k$ the estimate of g_k, $k = 0, 1, \dots 100$.

(f) We have $\widehat{y} = \mathbf{H}\mathbf{G}\,\widehat{u^2} + \widehat{e}$, where (for a change) a hat denotes the $\mathcal{Z}$ transformation. Thus,

$$\hat{y} = \underbrace{[\mathbf{H}\mathbf{G} \quad \mathbf{I}]}_{\mathbf{L}}\,\hat{w}, \quad \text{where} \quad w = \begin{bmatrix} u^2 \\ e \end{bmatrix}.$$

By the matrix version of the Wiener-Lee formula (discrete-time version), we have $S^{yy} = \bar{\mathbf{L}}\, S^{ww}\, \mathbf{L}^T$. Since u^2 and e are independent white noises, we have

$$S^{ww} = \begin{bmatrix} \sigma^2 & 0 \\ 0 & \sigma_1^2 \end{bmatrix}, \quad \text{where}$$

$$\sigma^2 = \mathrm{Var}(u_k^2) = E(u_k^4) - E(u_k^2)^2,$$

$$\sigma_1^2 = \mathrm{Var}(e_k^2) = E(e_k^2) - E(e_k)^2.$$

Multiplying out, we obtain

$$S^{yy}(e^{i\nu}) = |\mathbf{H}(e^{i\nu})|^2 \cdot |\mathbf{G}(e^{i\nu})|^2 \sigma^2 + \sigma_1^2.$$
