Paper Number(s): **E4.27**
C2.3

IMPERIAL COLLEGE OF SCIENCE, TECHNOLOGY AND MEDICINE
UNIVERSITY OF LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2001

MSc and EEE PART IV: M.Eng. and ACGI

SYSTEM IDENTIFICATION

Thursday, 17 May 10:00 am

There are SIX questions on this paper.

Answer FOUR questions.

Time allowed: 3:00 hours

Corrected Copy

Examiners: Weiss,G. and Allwright,J.C.

Special information for invigilators: none

Information for candidates:

$C(\tau) = E[(u(t)-\mu)(u(t+\tau)-\mu)]$

$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{x^2}{2\sigma^2}}$ $\qquad S_{yy} = |G|^2 S_{uu}$ $\qquad Z_L = sL \qquad Z_c = \frac{1}{Cs}$

$\Phi^{\#} = (\Phi^* \Phi)^{-1}\Phi^* \qquad P = \Phi\Phi^{\#} \qquad S = \frac{1}{N-\rho}\|y - \Phi\widehat{\theta}\|^2$

$A^d = e^{Ah} \qquad B^d = \left(e^{Ah} - I\right)A^{-1}B \qquad G^d(z) \approx G\left(\frac{2}{h}\,\frac{z-1}{z+1}\right) \qquad G(s) \approx G^d\left(\frac{1+sh/2}{1-sh/2}\right)$

$C_k^{uu} g_0 + C_{k-1}^{uu} g_1 + C_{k-2}^{uu} g_2 + \ldots = C_k^{uy}$

$\mathrm{Cov}(X,Y) = E[(X-\mu_X)(Y-\mu_Y)]$

$E(X\cdot Y) = E(X)\cdot E(Y) + \mathrm{Cov}(X,Y)$

$\mathrm{Cov}(TX) = T\mathrm{Cov}(X)T^*$

$\widehat{v}(z) = \sum_{k=0}^{\infty} v_k z^{-k}$

$[(\Delta v)_k = v_{k+1}] \quad \Rightarrow \Delta v(z) = z[\widehat{v}(z) - v_0]$

$[u_k = k v_k] \quad \Rightarrow \widehat{u}(z) = -z\frac{d}{dz}\widehat{v}(z)$

$[v_k = \sin k\nu] \quad \Rightarrow \widehat{v}(z) = \frac{z\sin\nu}{(z-e^{i\nu})(z-e^{-i\nu})}$

$\left[v_k = \rho^k\right] \quad \Rightarrow \widehat{v}(z) = \frac{z}{z-\rho}$

$\left[v_k = \frac{1}{\rho}k\rho^k\right] \quad \Rightarrow \widehat{v}(z) = \frac{z}{(z-\rho)^2}$

$$y_k + a_1 y_{k-1} \cdots + a_n y_{k-n} = b_0 u_k + b_1 u_{k-1} \cdots + b_n u_{k-n} + e_k + c_1 e_{k-1} \cdots + c_n e_{k-n}$$

$$C(z) = 1 + c_1 z^{-1} \ldots + c_n z^{-n}$$

$$\hat{u}^F = C^{-1}\hat{u}, \qquad \hat{y}^F = C^{-1}\hat{y}$$

$$\overline{y_k} = (c_1 - a_1)y_{k-1}^F + (c_2 - a_2)y_{k-2}^F \cdots + (c_n - a_n)y_{k-n}^F + b_0 u_k^F + b_1 u_{k-1}^F \cdots + b_n u_{k-n}^F$$

1. Consider the system with input u and output y modeled by the ARMAX difference equation

$$y_k + a_1 y_{k-1} + a_2 y_{k-2} = b_1 u_{k-1} + e_k - 0.5 e_{k-1},$$

where a_1, a_2, b_1 are unknown real numbers. The signal e is white noise with unknown variance and such that $E(e_k) = 0$.

(a) Write the formula for the transfer function H from e to y. For which values of a_2 is it possible that H is stable?

(b) Assuming that H is stable and $u = 0$, give explicit formulas for $E(y_k)$ (the expectation of y) and for $E(y_k^2)$ (the power of y) in terms of the impulse response of H, denoted $h = (h_0, h_1, h_2, \ldots)$, and in terms of the noise variance $\sigma^2 = E(e_k^2)$.

(c) Introduce new input and output variables u^F and y^F such that (i) u^F and y^F can be computed from u and y, (ii) u and y can be computed from u^F and y^F, (iii) the relation between u^F, y^F and e is described by an ARX difference equation with the same unknown parameters a_1, a_2 and b_1. Write down this ARX equation.

(d) Describe a least squares based method to estimate a_1, a_2 and b_1 from measurements of u_k and y_k for $k \leq 200$. Give a formula for an unbiased estimate of $Var(e_k)$. Hint: use the variables u^F and y^F introduced in part (c).

2. The proposed mathematical model of a static system with two inputs u and v and with one output w is

$$w = \sqrt{1 - \left(\frac{u}{\alpha}\right)^2 - \left(\frac{v}{\beta}\right)^2}. \tag{1}$$

The variables u, v and w can be measured and α, β are unknown parameters. We have 50 measurements available from experiments, $u_1, u_2, \ldots u_{50}$ and similarly for v and w, and we have $0 < w_k < 1$ (so that they do not contradict the model (1)). Because of measurement and modeling errors, the measurements do not fit any model of the form (1) exactly.

(a) By defining new variables if necessary, rewrite the model of the system in the form $y_k = \varphi_k\theta + e_k$, where y_k and φ_k are known, θ is the vector of unknown parameters and e_k are the equation errors.

(b) Write the formula for the vector of estimated parameters $\hat{\theta}$ which minimizes $e_1^2 + e_2^2 \ldots + e_{50}^2$.

(c) Assume that $e_1, e_2, \ldots e_{50}$ are independent and identically distributed. Assuming also that $E(e_k) = 0$, give a formula for an unbiased estimate of $Var(e_k)$.

(d) Still assuming independent and identically distributed equation errors, give a formula for an unbiased estimate of $Cov(\hat{\theta})$, where $\hat{\theta}$ is the estimate from part (b). Note that $Var(e_k)$ is not known, but it can be estimated, as was required in part (c).

(e) If, instead of 50 measurements, we would have 500 measurements available, would you expect $\|Cov(\hat{\theta})\|$ to be larger or smaller? Give a brief reason for your answer.

(f) We return to the problem in part (e) in a more precise framework. Suppose that each of the sequences u_k, v_k and e_k consists of independent and identically distributed random variables, and the three sequences are also independent of each other. Let $\hat{\theta}$ be the estimate from part (b). If, instead of 50 measurements, we have 500 measurements, approximately how many times do you expect $Cov(\hat{\theta})$ to decrease? Give, briefly, a reason for your answer.

3. We have a stable linear system with an unknown stable transfer function $\mathbf{G}$. The relevant frequency range on which this system operates is from 0 to 300 Hz. At higher frequencies we expect the transfer function of the system to be practically zero.

(a) Suppose that the input signal is

$$u(t) = 2(1 - e^{-3t})[1 + \cos 7t] \tag{2}$$

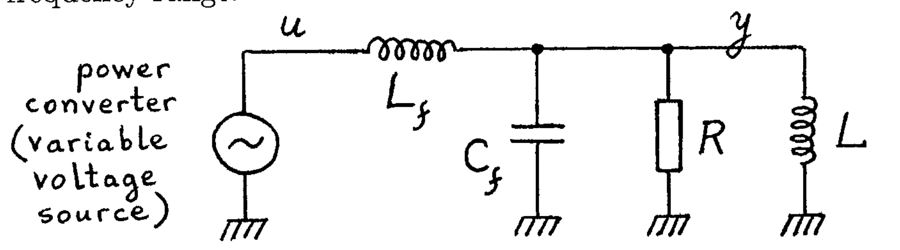
and the corresponding output signal is denoted by $y(t)$. Suppose that the limit $\lim_{T\to\infty} \frac{1}{T}\int_0^T y(t)\,dt = 10$ has been measured. What conclusion can we draw concerning $\mathbf{G}$ (e.g., its value at some point)?

(b) If we give the input signal u of (2) to our system, what sort of output function do we expect to see a long time after the start of the experiment (i.e., in steady state)? Express this output function in the time domain in terms of the transfer function $\mathbf{G}$.

(c) Assume that we want to incorporate our system into a feedback loop, and for stability studies we would like to plot an approximate Nyquist plot of $\mathbf{G}$. What sort of identification experiments could provide us with the necessary data for the Nyquist plot? Describe these experiments very briefly (assuming that looking at an oscilloscope is not accurate enough), and also describe briefly the computations necessary to process the data from these experiments.

(d) Assume that the input to the system is a stationary random signal u with power spectral density $S_{uu}(i\omega) = (1 + 4\omega^2)^{-1}$ and that the estimated power spectral density of the output signal y is $S_{yy}(i\omega) = (1 + 2\omega^2 + \omega^4)^{-1}$. Determine $|\mathbf{G}(i\omega)|$ (as a function of ω) and try to find one possible stable transfer function $\mathbf{G}$ whose absolute values match with those that you have determined.

4. We want to model the output circuit of a power converter by the simplified circuit shown below, where the filter inductor L_f and the filter capacitor C_f are known, and the load resistor R and load inductance L are unknown. We can choose the waveform of u (which is the output voltage of the converter) and we can measure the load voltage y. R and L are to be determined (they should be positive). We cannot expect a perfect match between our true circuit and this model, but we would like to get a close match in a certain frequency range.



(a) Compute the transfer function $\mathbf{G}$ of the model circuit (from u to y), in terms of L_f, C_f, R and L. Is $\mathbf{G}$ stable?

(b) Suppose that by measurements that use sinusoidal u, we have obtained estimates for $\mathbf{G}$ at 20 angular frequencies $\omega_1, \ldots \omega_{20}$, in the frequency range of interest. Using these data, how could we estimate R and L using a least squares based algorithm? Write down the formulas which give the estimated R and L, taking care to define all the symbols that you use. Hints: do not rush to write formulas, first think carefully what is known and what has to be estimated. To avoid writing complicated formulas, introduce suitable intermediate variables.

(c) Construct a realization of the transfer function $\mathbf{G}$, of the form $\dot{x} = Ax + Bu$, $y = Cx + Du$, where A, B, C and D are matrices. What are the eigenvalues of your matrix A, expressed in terms of quantities that have been determined earlier, such as L_f, C_f, R and L?

(d) We connect a hold device (D/A converter) at the input of our system (i.e., we use a digitally controlled converter) and we connect a sampler (A/D converter) at its output (e.g., a digital voltmeter), both converters working with the sampling period T. How can we compute the transfer function of the resulting discrete-time LTI system? There is no need to perform any computations to answer this part.

5. Consider the discrete-time LTI system with transfer function

$$\mathbf{G}(z) = \frac{3z^{-1}}{1 - 1.7z^{-1} + 0.72z^{-2}}.$$

(a) Write the difference equation corresponding to $\mathbf{G}$ (this should be an ARMA model).

(b) Determine if $\mathbf{G}$ is a stable transfer function, and compute its DC-gain. Is $\mathbf{G}$ proper? Is it strictly proper? Is it a FIR transfer function?

(c) Consider the signal u given by $u_0 = 0$ and

$$u_k = \frac{2}{5^k} \quad (k = 1, 2, 3, \ldots). \tag{3}$$

Compute the $\mathcal{Z}$-transform $\hat{u}(z)$.

(d) Let y be the response of the system to the input signal u given in (3). Assume that the initial state of the system is zero. Compute the $\mathcal{Z}$-transform $\hat{y}(z)$.

(e) Explain why the signal y from (d) is of the form

$$y_k = c_1(0.8)^k + c_2(0.9)^k + c_3(0.2)^k.$$

Explain briefly how the constants c_1, c_2, c_3 can be computed, but do not compute them numerically.

(f) Assume that the discrete-time transfer function $\mathbf{G}$ given above has been obtained by a discrete-time identification procedure applied to a continuous-time LTI system, via sample and hold blocks (i.e., D/A and A/D converters) with a sampling frequency of 5 KHz. Give a formula for an estimate of the transfer function $\mathbf{P}$ of the continuous-time system, which should be valid for frequencies that are significantly lower than the sampling frequency.

(g) Suppose that the output measurements of the above system are subject to measurement errors, in that

$$\hat{y}(z) = \mathbf{G}(z)\hat{u}(z) + \hat{e}(z),$$

where $\hat{e}$ is the $\mathcal{Z}$-transform of the sequence e_k which is normalized white noise (so that $E(e_k) = 0$ and $E(e_k^2) = 1$). Given measurements $y_1, y_2, \ldots y_{200}$, how can you compute an unbiased prediction of y_{201}? How large is the variance of the prediction error?

6. Assume that we have to identify the input/output behaviour of a stable discrete-time LTI plant which is part of a larger system. We do not have the possibility to apply input signals of our choice to the plant, we can only observe the existing input and output signals. These appear to be stationary random signals.

(a) Describe a method for estimating the terms $g_0, g_1, \ldots$ in the impulse response of the plant, based on estimating certain auto-correlation and cross-correlation functions first.

(b) What is the meaning of a random signal being "persistent of order N"? What is the significance of this concept in the context of part (a) above?

(c) After having estimated the first N terms of the impulse response, $g_0, g_1, \ldots g_{N-1}$, how can we build a FIR filter whose transfer function is a good approximation to the true transfer function? Write the corresponding difference equation.

(d) Suppose that a discrete-time stationary random signal u_k has power spectral density S^{uu} such that $S^{uu}(\zeta) \geq \varepsilon$ for all ζ on the unit circle, where $\varepsilon > 0$. Explain why this implies that u is persistent of order N for any $N = 1, 2, 3, \ldots$.

Hint for (d): consider the infinite Toeplitz matrix T with entries $T_{jk} = C^{uu}_{j-k}$. We regard this matrix as acting on square summable sequences x_k (with $k \geq 0$), which are regarded as infinite column vectors. The scalar product of two such sequences is defined by $\langle x, y\rangle = \sum_{k=0}^{\infty} x_k \overline{y_k}$ and the norm is defined by $\|x\|^2 = \langle x, x\rangle$. We have (using the Parseval equality)

$$\langle Tx, x\rangle = \langle C^{uu} * x, x\rangle = \frac{1}{2\pi}\int_{|\zeta|=1} S^{uu}(\zeta)|\hat{x}(\zeta)|^2 d\zeta \geq \varepsilon\|x\|^2.$$

Thus, we see that the infinite Toeplitz matrix is strictly positive, in fact, $T \geq \varepsilon I$. This inequality can be used as the starting point for the explanation requested in part (d).

[END]

SYSTEMS IDENTIFICATION, May 2001
Solutions

master- 21.6.01

(E4 27 C2.3)

Question 1. (a) ④ $H(z) = \dfrac{1-0.5z^{-1}}{1+a_1 z^{-1} + a_2 z^{-2}} =$

$$= \frac{z^2 - 0.5z}{z^2 + a_1 z + a_2} = \frac{z(z-0.5)}{(z-\lambda_1)\cdot(z-\lambda_2)}, \quad a_2 = \lambda_1 \lambda_2 .$$

If H is stable, then $|\lambda_1| < 1$, $|\lambda_2| < 1$, hence $|a_2| < 1$.

(b) ⑦ $E(y_k) = H(1) \cdot E(e_k) = 0$. $E(y_k^2) =$

$$= E\big((h_0 e_k + h_1 e_{k-1} + h_2 e_{k-2} + \dots)(h_0 e_k + h_1 e_{k-1} + h_2 e_{k-2} + \dots)\big)$$

$$= E\Big(h_0^2 e_k^2 + h_1^2 e_{k-1}^2 + h_2^2 e_{k-2}^2 + \dots + \sum_{j \neq \ell} h_j h_\ell e_{k-j} e_{k-\ell}\Big)$$

$$= h_0^2 \underbrace{E(e_k^2)}_{\sigma^2} + h_1^2 E(e_{k-1}^2) + h_2^2 \underbrace{E(e_{k-2}^2)}_{\sigma^2} + \dots + \sum_{j \neq \ell} h_j h_\ell \underbrace{E(e_{k-j} e_{k-\ell})}_{0}$$

Thus,

$$E(y_k^2) = \sigma^2 (h_0^2 + h_1^2 + h_2^2 + \dots).$$

The same result could be obtained from $C_{yy} = \check{h} * h * C_{ee}$, using $C_{ee} = \sigma^2 \delta_0$ and computing $E(y_k^2) = C_{yy}(0)$. (Here, $\check{h}_k = h_{-k}$.)

(c) ⑦ $\hat{u}^F(z) = \dfrac{1}{1-0.5z^{-1}} \hat{u}(z)$, $\quad \hat{y}^F(z) = \dfrac{1}{1-0.5z^{-1}} \hat{u}(z)$, i.e., u^F and y^F are the solutions of

$$u_k^F - 0.5\, u_{k-1}^F = u_k \,, \qquad y_k^F - 0.5\, y_{k-1}^F = y_k$$

with initial conditions zero : $u_0^F = y_0^F = 0$ (the influence of the initial conditions is anyway negligible for large k). Then the ARX equation is

$$y_k^F + a_1 y_{k-1}^F + a_2 y_{k-2}^F = b_1 u_{k-1}^F + e_k \,.$$

(d)[7]

$$y_k^F = \underbrace{\begin{bmatrix} -y_{k-1}^F & -y_{k-2}^F & u_{k-1}^F \end{bmatrix}}_{\varphi_k} \underbrace{\begin{bmatrix} a_1 \\ a_2 \\ b_1 \end{bmatrix}}_{\theta} + e_k$$

$$\Phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{200} \end{bmatrix}, \quad \Phi^\# = (\Phi^*\Phi)^{-1}\Phi^*,$$

unbiased estimate of θ

$$\hat{\theta} = \Phi^\# y^F, \qquad y^F = \begin{bmatrix} y_1^F \\ y_2^F \\ \vdots \\ y_{200}^F \end{bmatrix}.$$

$$\widehat{\mathrm{Var}(e_k)} = \frac{1}{197} \| y^F - \Phi\hat{\theta} \|^2 \qquad \left(\text{because } 200-3=197\right)$$

unbiased estimate of $\mathrm{Var}(e_k)$

$$= \frac{1}{197}\, y^{F*}(I - \Phi\Phi^\#)\, y^F .$$

Question 2.

(a)[4] $$w_k^2 = 1 - \frac{u_k^2}{\alpha^2} - \frac{v_k^2}{\beta^2} - e_k$$

it is more convenient to start with a minus sign here

$$\underbrace{1 - w_k^2}_{y_k} = \underbrace{\begin{bmatrix} u_k^2 & v_k^2 \end{bmatrix}}_{\varphi_k} \underbrace{\begin{bmatrix} \frac{1}{\alpha^2} \\ \frac{1}{\beta^2} \end{bmatrix}}_{\theta} + e_k$$

θ = vector of unknown parameters

(b)[4] $\hat{\theta} = \Phi^{\#} y$, where $\Phi^{\#} = (\Phi^* \Phi)^{-1} \Phi^*$,

$$\Phi = \begin{bmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_{50} \end{bmatrix}, \quad y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{50} \end{bmatrix} \qquad (\Phi \in \mathbb{R}^{50 \times 2})$$

(c)[4] $$\widehat{\mathrm{Var}(e_k)} = \frac{1}{48} \| y - \Phi \hat{\theta} \|^2 = \frac{1}{48} y^* (I - \Phi \Phi^{\#}) y$$

$48 = 50 - 2$

(d)[4] $$\widehat{\mathrm{Cov}(\hat{\theta})} = \widehat{\mathrm{Var}(e_k)} \, (\Phi^* \Phi)^{-1},$$ where

$\widehat{\mathrm{Var}(e_k)}$ has been computed in (c).

(e)[4] With more measurements, $\| \mathrm{Cov}(\hat{\theta}) \|$ would decrease, because $\Phi^* \Phi$ would increase (as a positive matrix) and $\mathrm{Cov}\, \hat{\theta} = \sigma^2 (\Phi^* \Phi)^{-1}$, where $\sigma^2 = \mathrm{Var}(e_k)$ (σ^2 is not known).

(f)[5] For N measurements we have
$$\Phi^* \Phi = \begin{bmatrix} \sum_{k=1}^{N} u_k^4 & \sum_{k=1}^{N} u_k^2 v_k^2 \\ \sum_{k=1}^{N} u_k^2 v_k^2 & \sum_{k=1}^{N} v_k^4 \end{bmatrix}$$
so that by ergodicity (signals obtained from white noise are ergodic)
$$\Phi^* \Phi \approx N \begin{bmatrix} E(u_k^4) & E(u_k^2 v_k^2) \\ E(u_k^2 v_k^2) & E(v_k^4) \end{bmatrix}.$$
Hence, for $N = 500$ we expect $\mathrm{Cov}\, \hat{\theta}$ to be about 10 times smaller than for $N = 50$.

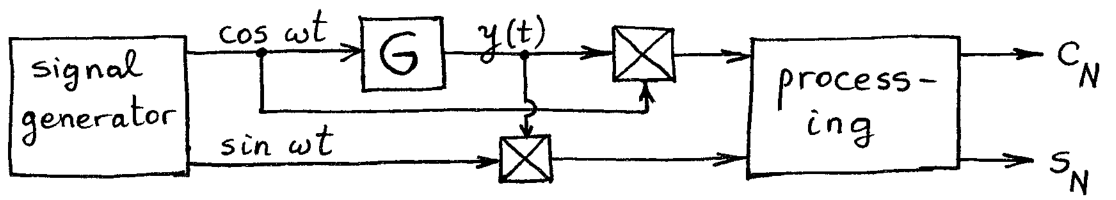
Question 3. (a) ⑥ $u(t) = 2 + 2\cos 7t + e(t)$

where $e(t) \to 0$ (fast). Thus, the output will have a constant (DC) component corresponding to $u_1(t) = 2$, a sinusoidal (AC) component corresponding to $u_2(t) = 2\cos 7t$ and a transient response, which $\to 0$. The DC component is (according to the measurement) 10, so the DC gain is 5, i.e., $G(0) = 5$.

(b) ⑤ $$y(t) = G(0)\cdot 2 + |G(7i)|\cdot \cos(7t+\varphi) + \varepsilon(t),$$

where $\varphi = \arg G(7i)$ and $\varepsilon(t) \to 0$.

(c) ⑦



We would try several frequencies up to $\omega \approx 300\cdot 2\pi$.

From

$$\left(T = \frac{2\pi}{\omega}\right) \begin{cases} C_N = \int_0^{NT} y(t)\cos\omega t\, dt = A_\omega \cos\varphi_\omega \cdot \frac{NT}{2} \\ S_N = \int_0^{NT} y(t)\sin\omega t\, dt = -A_\omega \sin\varphi_\omega \cdot \frac{NT}{2} \end{cases}$$

(these are computed when the system has reached steady state) it is easy to determine A_ω and φ_ω, hence $G(i\omega) = A_\omega e^{i\varphi_\omega}$. This gives several points on the Nyquist plot, from which we can guess its shape.

(d) ⑦ $$|G(i\omega)|^2 = S_{yy}(i\omega)/S_{uu}(i\omega) = \frac{1+4\omega^2}{1+2\omega^2+\omega^4}.$$

We write this as a function of $i\omega$ and we factor it:

$$|G(i\omega)|^2 = \frac{(1-2i\omega)(1+2i\omega)}{[1-(i\omega)^2]^2} = \frac{1-2i\omega}{(1-i\omega)^2}\cdot\frac{1+2i\omega}{(1+i\omega)^2}.$$

The two factors above are complex conjugate of each other, and the second is stable $\Rightarrow$ the second could be G.

Question 4. (a) ⑥ The impedance Z of the three circuit elements connected in parallel is given by

$$\frac{1}{Z(s)} = C_f s + \frac{1}{R} + \frac{1}{Ls} = \frac{C_f R L s^2 + Ls + R}{RLs}$$

so that $Z(s) = \dfrac{RLs}{C_f R L s^2 + Ls + R}$. The transfer function is

$$G(s) = \frac{Z(s)}{Z(s) + L_f s} = \frac{RLs}{RLs + L_f s(C_f R L s^2 + Ls + R)}$$

$$= \frac{\dfrac{1}{L_f C_f}}{s^2 + \dfrac{1}{RC_f}\cdot s + \dfrac{L_f + L}{L_f C_f L}} = \frac{b_0}{s^2 + a_1 s + a_0}$$

Note that b_0 is known, while a_1, a_0 are unknown. G is stable, because a_1, a_0 are positive.

(b) ⑧ Recall that G is only a model. We denote by $G^e(i\omega_k)$ the experimentally determined values of the true transfer function at frequencies ω_k (these are subject to measurement errors, of course). We have

$$b_0 = \left[(i\omega_k)^2 + a_1 (i\omega_k) + a_0 \right] G^e(i\omega_k) - e_k ,$$

with $k = 1, \dots 20$, where e_k represents the combined effect of measurement and modeling errors. We rewrite:

$$\underbrace{(i\omega_k)^2 G^e(i\omega_k) - b_0}_{y_k} = \underbrace{[-i\omega_k \ \ -1]\, G^e(i\omega_k)}_{\varphi_k} \begin{bmatrix} a_1 \\ a_0 \end{bmatrix} + e_k$$

Note that y_k and φ_k are known.

Now we can estimate a_1 and a_0 using the usual least squares algorithm:

$$\hat{\theta} = \Phi^{\#} y, \quad y = \begin{bmatrix} y_1 \\ \vdots \\ y_{20} \end{bmatrix}, \quad \Phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_{20} \end{bmatrix}, \quad \Phi^{\#} = (\Phi^* \Phi)^{-1} \Phi^*.$$

From a_1 and a_0 we can compute R and L:

$$R = \frac{1}{a_1 C_f}, \qquad L = \frac{L_f}{a_0 L_f C_1 - 1}.$$

(c) 5

$$A = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

$$C = [\, b_0 \quad 0 \,], \qquad D = [0],$$

$$\sigma(A) = \left\{ \frac{-a_1 \pm \sqrt{a_1^2 - 4a_0}}{2} \right\}.$$

(d) 6

$$A^d = e^{AT}, \quad B^d = (e^{AT} - I) A^{-1} B,$$

$$G^d(z) = C (zI - A^d)^{-1} B^d + D.$$

Alternatively, we get a good approximation of G^d by Tustin's formula:

$$G^d(z) \approx G\left(\frac{2}{T} \cdot \frac{z-1}{z+1} \right).$$

Question 5.

(a) (2) The difference equation is

$$y_k - 1.7\, y_{k-1} + 0.72\, y_{k-2} = 3\, u_{k-1} \qquad (k=1,2,\dots).$$

(b) (3) The poles of G are 0.8 and 0.9, so that it is stable. Its DC gain is $G(1) = \dfrac{3}{1.72-1.7} = 150.$

G is strictly proper (i.e., $G(\infty)=0$) and not **FIR**.

(c) (3) If the formula for u_k were true also for $k=0$, then it Z-transform would be $\dfrac{2z}{z-0.2}$. From this we have to subtract $2\hat{\delta}_0 = 2$, so that

$$\hat{u}(z) = \frac{2z}{z-0.2} - 2 = \frac{0.4}{z-0.2}.$$

(d) (3) $$\hat{y}(z) = G(z)\,\hat{u}(z) = \frac{1.2\,z}{(z-0.8)(z-0.9)(z-0.2)}$$

(e) (5) We use the partial fractions decomposition

$$\underbrace{\frac{1.2}{(z-0.8)(z-0.9)(z-0.2)}}_{\text{this is } z^{-1}\hat{y}(z)} = \frac{c_1}{z-0.8} + \frac{c_2}{z-0.9} + \frac{c_3}{z-0.2}.$$

Now multiply both sides with z and take inverse Z-transforms.:

$$y_k = c_1\,(0.8)^k + c_2\,(0.9)^k + c_3\,(0.2)^k, \quad k=0,1,2,\dots$$

$$c_1 = \lim_{z\to 0.8} (z-0.8)\cdot\frac{\hat{y}(z)}{z}$$ and similarly for c_2, c_3.

(f) (4) $$\underbrace{P(s) \approx}_{\text{by the inverse of Tustin's formula}} G\left(\frac{1+\frac{hs}{2}}{1-\frac{hs}{2}}\right), \text{ where } h = \frac{1}{5000}.$$

(g) ⑤ The general formula for one step ahead prediction is

$$\bar{y}_k = (c_1 - a_1) y^F_{k-1} + (c_2 - a_2) y^F_{k-2} \dots + (c_n - a_n) y^F_{k-n}$$
$$+ b_0 u^F_k + b_1 u^F_{k-1} \dots + b_n u^F_{k-n}$$

where
$$y^F_k + c_1 y^F_{k-1} \dots + c_n y^F_{k-n} = y_k$$
$$u^F_k + c_1 u^F_{k-1} \dots + c_n u^F_{k-n} = u_k,$$

$$y_k + a_1 y_{k-1} \dots + a_n y_{k-n} = b_0 u_k + b_1 u_{k-1} \dots + b_n u_{k-n}$$
$$+ e_k + c_1 e_{k-1} \dots + c_n e_{k-n}.$$

the original ARMAX model

(There is no need to write all this down in the exam, this is only a reminder of the theory.)

$$y_k = \bar{y}_k + e_k.$$

In our case (output error) we have $c_\ell = a_\ell$ ($\ell = 1, \dots n$), which implies that

$$\hat{\bar{y}}(z) = B(z)\hat{u}^F(z) = \frac{B(z)}{C(z)}\hat{u}(z) = G(z)\hat{u}(z).$$

Thus, $\bar{y}_k$ is obtained by filtering u through G, and <u>ignoring all the measurements</u> $y_1, \dots y_{k-1}$. In particular, this is true for $k = 201$. From $y_k = \bar{y}_k + e_k$ we see that the variance of the prediction error is $\mathrm{Var}(e_k) = 1$.

Question 6. (a)(7) For $\tau = 0,1,2,\dots N$ we estimate the auto-correlations C^{uu}_τ and cross-correlations C^{uy}_τ by time averaging (hoping that they are ergodic):

$$\widehat{C^{uu}_\tau} = \frac{1}{n+1}\sum_{k=0}^{n} u_k u_{k+\tau}, \quad \widehat{C^{uy}_\tau} = \frac{1}{n+1}\sum_{k=0}^{n} u_k y_{k+\tau}.$$

The larger n, the better these estimates become. Then we solve

$$\begin{bmatrix} \widehat{C^{uu}_0} & \widehat{C^{uu}_1} & \cdots & \widehat{C^{uu}_{N-1}} \\ \widehat{C^{uu}_1} & \widehat{C^{uu}_0} & \cdots & \widehat{C^{uu}_{N-2}} \\ \vdots & \vdots & & \vdots \\ \widehat{C^{uu}_{N-1}} & \widehat{C^{uu}_{N-2}} & \cdots & \widehat{C^{uu}_0} \end{bmatrix} \cdot \begin{bmatrix} \hat{g}_0 \\ \hat{g}_1 \\ \vdots \\ \hat{g}_{N-1} \end{bmatrix} = \begin{bmatrix} \widehat{C^{uy}_0} \\ \widehat{C^{uy}_1} \\ \vdots \\ \widehat{C^{uy}_{N-1}} \end{bmatrix}$$

Now $\hat{g}_0, \dots \hat{g}_{N-1}$ are our estimate of the first N terms of the impulse response $(g_0, g_1, g_2, \dots)$.

(b)(6) u is called persistent of order N if the square matrix appearing above, but with the true values instead of the estimates, is invertible. If this is the case, then there is a good chance for the estimated matrix to be invertible as well, so that we can solve for $\hat{g}_0, \dots \hat{g}_{N-1}$. Since the square matrix in question is positive (≥ 0), if it is invertible then all its diagonal submatrices are also invertible, so that persistence of order N implies persistence of order K for all $K \leq N$.

(c)(5) The FIR filter with impulse response

$$\hat{g} = (\hat{g}_0, \hat{g}_1, \dots \hat{g}_{N-1}, 0, 0, 0, \dots)$$

is described by the MA equation

$$y_k = \hat{g}_0 u_k + \hat{g}_1 u_{k-1} + \hat{g}_2 u_{k-2} \cdots + \hat{g}_{N-1} u_{k-N+1}.$$

(d) 7 In the "hint" part of point (d) it was shown that $\langle Tx, x\rangle \geqslant \varepsilon \|x\|^2$ for any square summable sequence x. In particular, if we take

$$x = [x_0 \; x_1 \; x_2 \; \ldots \; x_{N-1} \; 0 \; 0 \; 0 \; \ldots]^T$$

then this can be written in matrix form:

$$|x^* T x| \geqslant \varepsilon \|x\|^2 .$$

Since only the first N components of x are nonzero, this can be rewritten as

$$|x_N^* T_N x_N| \geqslant \varepsilon \|x_N\|^2 ,$$

where $x_N = [x_0 \; x_1 \; \ldots \; x_{N-1}]^T$ and

$$T_N = \begin{bmatrix} c_0^{uu} & c_1^{uu} & \ldots & c_{N-1}^{uu} \\ c_1^{uu} & c_0^{uu} & \ldots & c_{N-2}^{uu} \\ \vdots & \vdots & & \vdots \\ c_{N-1}^{uu} & c_{N-2}^{uu} & \ldots & c_0^{uu} \end{bmatrix} .$$

It follows that $T_N x_N \neq 0$ for all $x_N \in \mathbb{C}^N$, so that T_N is invertible. This means that u is persistent of order N.