E4.27
C2.3
ISE4.41

IMPERIAL COLLEGE LONDON

DEPARTMENT OF ELECTRICAL AND ELECTRONIC ENGINEERING
EXAMINATIONS 2004

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

SYSTEM IDENTIFICATION

Friday, 14 May 10:00 am

Time allowed: 3:00 hours

There are SIX questions on this paper.

Answer FOUR questions.

All questions carry equal marks

Corrected Copy

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

Examiners responsible First Marker(s) : G. Weiss

Second Marker(s) : J.C. Allwright

Special information for invigilators: none

Information for candidates:

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

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

$$Z_L = sL \qquad Z_c = \frac{1}{Cs}$$

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

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

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

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

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

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

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

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

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

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

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

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

$$P_n = \frac{1}{\lambda}\left[P_{n-1} - \frac{P_{n-1}\varphi_n^*\varphi_n P_{n-1}}{\lambda + \varphi_n P_{n-1}\varphi_n^*}\right]$$

$$\varepsilon_n = y_n - \varphi_n\hat{\theta}_{n-1}$$

$$\hat{\theta}_n = \hat{\theta}_{n-1} + P_n\varphi_n^*\varepsilon_n$$

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

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

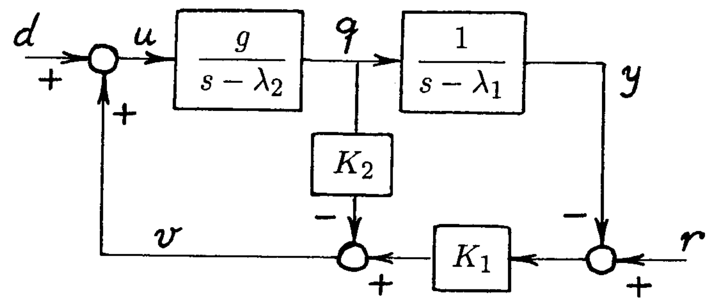
$$\hat{u}^F = C^{-1}\hat{u}, \qquad \hat{y}^F = C^{-1}\hat{y}$$

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

1. A continuous-time plant with input u and output y is known to be composed of two first order systems connected in cascade, with the transfer functions $g/(s-\lambda_2)$ and $1/(s-\lambda_1)$, see the block diagram below. The real numbers λ_1, λ_2 and g are not known, but we know that

$$|\lambda_1| < 1, \qquad |\lambda_2| < 1, \qquad \frac{1}{2} < g < 2,$$

so that for $|s| >> 1$, the transfer function from u to y is approximately g/s^2, as for a double integrator. The signal q between the two blocks can be measured. An external disturbance d acts on the plant, in addition to the control input v, and we would like the output y to track a reference signal r. We connect the system to a controller described by $v = K_1(r-y) - K_2 q$, as shown in the block diagram, where $K_1 > 0$, $K_2 > 0$.



(a) Write the equations of the closed-loop system in the state space form $\dot{x} = Ax + B_1 r + B_2 d$. It is advisable to take y as the first component of the state x. [5]

(b) Determine the set of those controller gains K_1 and K_2 for which the closed-loop system is stable, regardless of the values of λ_1, λ_2 and g in the specified range. [5]

(c) Take $K_1 = 50$ and $K_2 = 20$, so that the closed-loop system is stable. The disturbance d is an unknown constant. We have access to measurements of q, y and r taken at sampling times $t = hk$, where $k = 1, 2, \ldots 30,000$ and $h = 10^{-5}$. We denote $y_k = y(hk)$ and similarly for q_k and r_k. The reference r may be assumed to be constant on each sampling interval (since it changes slowly). Describe a least squares based method to estimate λ_1, λ_2, g and d from the given data. Hint: discretize the two blocks of the plant separately. By considering the left block only, explain how to estimate λ_2, g and d. Afterwards, consider the right block and explain how to estimate λ_1. Recall that for very small ε we have $e^{\varepsilon} \approx 1 + \varepsilon$. [10]

2. Assume that Σ is a stable discrete-time LTI system with input signal u and output signal v. The measured output y is corrupted by the noise signal w, so that $y_k = v_k + w_k$. The statistical properties of w are known: it can be modeled as filtered white noise,

$$\hat{w} = \Xi \hat{\varepsilon}, \qquad \Xi(z) = c_0 + c_1 z^{-1} \ldots + c_5 z^{-5},$$

where, $c_0, c_1, \ldots c_5$ are known and ε is Gaussian white noise with $E(\varepsilon_k) = 0$ and $\mathrm{Var}(\varepsilon_k) = 1$. ($\hat{w}, \hat{\varepsilon}$ are the $\mathcal{Z}$-transforms of w, ε.) The zeros of the polynomial $c(z) = c_0 z^5 + c_1 z^4 \ldots + c_5$ are in the open unit disk.

We have to identify Σ, based on the measurements of u_k and y_k. We would like to model Σ by a FIR filter of order 20:

$$v_k = b_0 u_k + b_1 u_{k-1} \cdots + b_{20} u_{k-20}.$$

(a) Describe the system by a standard MAX model (recall that MAX stands for "moving average with exogenous input"). By introducing new signals, reduce this to an X model in which the unknown coefficients $b_0, b_1, \ldots b_{20}$ appear. [4]

(b) Suppose that in the X model from part (a) the equation error due to model mismatch, denoted by e_k, is white noise with $E(e_k) = 0$ and e_k is independent of ε_j (for all $k, j \in \mathbb{Z}$). Assuming that the measurements u_k and y_k are available for $k = 1, 2, \ldots 6{,}000$, describe a least squares based method for estimating $b_0, b_1, \ldots b_{20}$ using the X model obtained in part (a). State clearly which cost function you are minimizing. [5]

(c) With the assumptions from part (b), how could we estimate $\mathrm{Var}(e_k)$ based on the available data described in part (b)? Hint: first estimate $\mathrm{Var}(\varepsilon_k + e_k)$. [5]

(d) Suppose that we want the estimation of the coefficients $b_0, b_1, \ldots b_{20}$ to be performed on-line, in order to track these coefficients if they are slowly changing. How can you modify the least squares minimization problem that you have solved in part (b), to ensure that the algorithm gradually "forgets" old data? Describe a recursive algorithm which minimizes the modified minimization problem. [6]

3. Suppose that $u = (u_k)$ is a stationary and ergodic Gaussian random signal in discrete time ($k \in \mathbb{Z}$). An LTI system with transfer function $\mathbf{G}$ has u as its input signal and $y = (y_k)$ as its output signal. We model $\mathbf{G}$ by

$$\mathbf{G}(z) = \frac{b}{z+a},$$

where the parameters $a \in (0,1)$ and b are unknown.

A company (your employer) requires you to produce formulas which give good predictions of u_n and y_n based on the measurements of u_k and y_k for $k = n-1, n-2, \ldots n-50$. The company supplies you with the measurements of u_k and y_k for $k = 1, 2, \ldots 5,000$, and you must design your prediction formulas using these data.

(a) Denote by C^{uu} the auto-correlation function of u. Explain how to estimate $\mu_u = E(u_k)$ and C^{uu}_j for $j = 0, 1, \ldots 50$. [3]

(b) Explain how to estimate the coefficients of a stable auto-regressive filter of order m, with transfer function denoted by Ξ, such that u can be regarded as the output function of this filter, when the input function of the filter is white noise $e = (e_k)$. Explain how to estimate $\mu_e = E(e_k)$ and $\sigma_e^2 = \mathrm{Var}(e_k)$ after Ξ has been identified. [5]

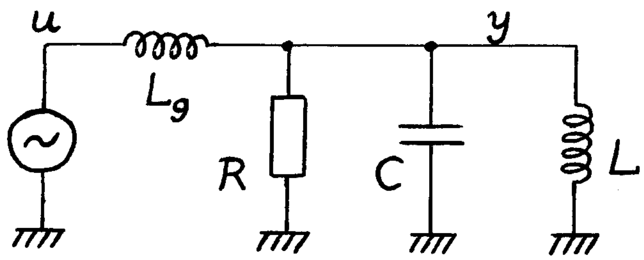
(c) Explain how to obtain the formula for the prediction of u_n, denoted by $\overline{u_n}$, using the results you obtained in part (b). Assuming that you have identified Ξ, μ_e and σ_e^2 with very high accuracy, give an estimate for the variance of the prediction error $u_n - \overline{u_n}$. [5]

(d) Outline a least squares based method to estimate a and b. Do not give any proofs. [3]

(e) Explain how to obtain the formula for the prediction of y_n, denoted by $\overline{y_n}$, using the results you obtained in the earlier parts. Give an estimate for the variance of the prediction error $y_n - \overline{y_n}$.

Hint: do not rush your answer, think carefully which data are available for predicting y_n. [4]

4. A power generator is represented as a voltage source of voltage u in series with a known inductor L_g. The load circuit is not known, but we want to model it as the parallel connection of a resistor R, a capacitor C and an inductor L, as shown in the circuit below. The values R, C and L are unknown but positive. For the purposes of the load identification, we may vary the mechanical speed of the generator in the relevant range, which results in sinusoidal voltages u at various frequencies (in the relevant range) and various amplitudes. We can measure u and also the voltage y on the load. We cannot expect a perfect match between the true circuit and our model, but we would like to get a close match in the relevant frequency range.



(a) Compute the transfer function $\mathbf{G}$ of the model circuit (from u to y), in terms of L_g, R, C and L. Is $\mathbf{G}$ stable? [5]

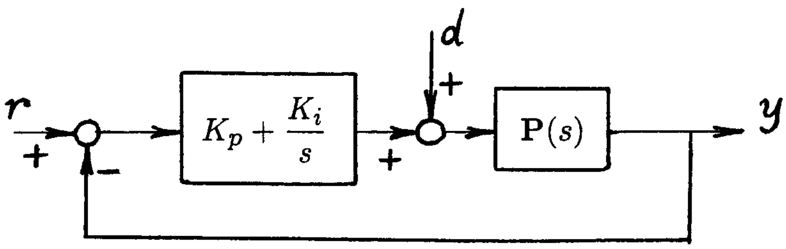
(b) Suppose that by measurements (using different frequencies for u) we have obtained estimates for $\mathbf{G}$ at 40 angular frequencies $\omega_1, \ldots \omega_{40}$, in the frequency range of interest. Using these data, how could we estimate R, C and L using a least squares based algorithm? Write down the formulas which give the estimated R, C and L, taking care to define all the symbols that you use. Hints: To avoid writing complicated formulas, introduce suitable intermediate variables. Avoid getting complex numbers as estimates for R, C and L. [6]

(c) Construct a realization of the transfer function $\mathbf{G}$, of the form $\dot{x} = Ax + Bu$, $y = C_o x + Du$, where A, B, C_o and D are matrices. What are the eigenvalues of A, expressed in terms of quantities that are known or have been estimated earlier, such as L_g, R, C and L? [4]

(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 h. 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]

5. We want to connect a stable linear SISO plant with an unknown transfer function $\mathbf{P}$ to a PI controller, in order to ensure tracking of a constant reference r, as shown in the block diagram below. The DC gain of the plant is known to be positive. The relevant frequency range on which the closed-loop system will operate is from 0 to 1000 Hz. At higher frequencies we expect $\mathbf{P}(i\omega)$ to be practically zero.

We want to find controller gains K_p and K_i such that the closed-loop system is stable, and K_p, K_i should not be too small (to avoid a very slow response of the closed-loop system).



(a) In order to choose suitable parameters for the controller, we would like to plot an approximate Nyquist plot of $\mathbf{P}$. What sort of identification experiments could provide us with the necessary data for the Nyquist plot? Describe these experiments very briefly, and also describe briefly the computations necessary to process the data from these experiments. [6]

(b) Suppose that K_p, K_i have been chosen such that the closed-loop system is stable. Suppose that the reference signal is

$$r(t) = 30(1 - e^{-7t}) + 60te^{-6t}\cos 300t$$

and the corresponding output signal is denoted by y, as shown in the block diagram. Assume that $d = 3$ (constant in time). Describe the structure of $y(t)$ for large t (i.e., in steady state), computing all the relevant constants. [5]

(c) Assume that the closed-loop system is stable, $r = 0$ and d is a stationary ergodic random signal with expectation $E(d) = 3$ and a certain known power spectral density S^{dd}. Is y a stationary random signal? Is y ergodic? Compute $E(y)$ and write a formula for computing $Var(y)$ (the power of y), in terms of $\mathbf{P}$, K_p, K_i and S^{dd}. [6]

(d) If K_p, K_i and r are as in part (b) and d is as in part (c), is y a stationary random signal? Give a very brief reasoning. [3]

6. Assume that v is a stationary ergodic random signal with the expectation $E(v(t)) = 3$ and the auto-correlation function

$$C^{vv}(\tau) = e^{-100|\tau|}.$$

The signal v is sampled with a sampling period $h = 10^{-4}$, and the resulting discrete-time signal $u_k = v(hk)$ is applied to an unknown stable discrete-time LTI system Σ with the (unknown) transfer function $\mathbf{G}$. The output signal of Σ, denoted by y, is corrupted by measurement noise e, such that in terms of $\mathcal{Z}$-transforms

$$\hat{y} = \mathbf{G}\hat{u} + \hat{e}.$$

The measurement noise e is white noise with $E(e_k) = 0$, independent of v. The measurements u_k and y_k are available for $k = 1, 2, \ldots 5,000$.

(a) Are u and y jointly stationary? Explain very briefly what this concept means. Are u and y jointly ergodic? Again, explain very briefly what this means. [3]

(b) Compute the power and the power spectral density of the discrete time signal u. [3]

(c) Describe a method for estimating $E(y_k)$, $\mathrm{Var}(y_k)$ and C^{uy}_τ (the cross-correlation function of u and y), using the available measurements u_k and y_k. For what values of τ can we obtain reasonable estimates of C^{uy}_τ? [3]

(d) Assume that the system Σ is sufficiently stable so that its impulse response (g_k) is negligible for $k \geq 50$. Describe a method for estimating the first 50 terms $g_0, g_1, \ldots g_{49}$ from the results of part (c). [4]

(e) What is the meaning of a random signal being "persistent of order m"? What is the significance of this concept in the context of part (d) above? Is u persistent of order 50? Give a brief reasoning. Hint: use your result from part (b). [3]

(f) Compute $\mathrm{Var}(y_k)$ in terms of the impulse response (g_k), the cross-correlation function C^{uy} and the noise power $\mathrm{Var}(e_k)$ (the latter is not known). Hence, give a formula for estimating $\mathrm{Var}(e_k)$ in terms of quantities estimated earlier. [4]

[END]

E4.27
C2.3
ISE 4.41

SYSTEM IDENTIFICATION, Exam of May 2004, Solutions

Question 1. [5] (a) Take the state $x = \begin{bmatrix} y \\ q \end{bmatrix}$.

Then $\dot{y} = \lambda_1 y + q$, $\dot{q} = \lambda_2 q + gd + gv$,

$v = K_1(r-y) - K_2 q$, hence

$$\underbrace{\frac{d}{dt}\begin{bmatrix} y \\ q \end{bmatrix}}_{\dot{x}} = \underbrace{\begin{bmatrix} \lambda_1 & 1 \\ -gK_1 & \lambda_2 - gK_2 \end{bmatrix}}_{A} \cdot \underbrace{\begin{bmatrix} y \\ q \end{bmatrix}}_{x} + \underbrace{\begin{bmatrix} 0 \\ gK_1 \end{bmatrix}}_{B_1} r + \underbrace{\begin{bmatrix} 0 \\ g \end{bmatrix}}_{B_2} d.$$

[5]
(b) A is stable iff trace $A < 0$ and det $A > 0$. We have trace $A = \lambda_1 + \lambda_2 - gK_2$, so we must have $gK_2 > \lambda_1 + \lambda_2$. Since $\lambda_1 + \lambda_2 < 2$ (and any value slightly below 2 is possible) we must ensure $gK_2 \geq 2$. Since $g > \frac{1}{2}$ (and any value slightly above $\frac{1}{2}$ is possible) we must ensure $\boxed{K_2 \geq 4}$.

We have det $A = \lambda_1\lambda_2 - \lambda_1 g K_2 + gK_1$, so K_1 has to be such that $gK_1 > \lambda_1 g K_2 - \lambda_1\lambda_2$.

Dividing by g: $\quad K_1 > \lambda_1 K_2 - \frac{\lambda_1\lambda_2}{g}$.

We have $-\frac{\lambda_1\lambda_2}{g} < \frac{1}{g} < 2$ (and any value slightly below 2 is possible), so we must ensure

$$K_1 \geq \lambda_1 K_2 + 2.$$

Since $\lambda_1 < 1$ (and any value slightly below 1 is possible), we must ensure

$\boxed{K_1 \geq K_2 + 2.}$ (Remember that the first condition was $K_2 \geq 4$.)

[10]

(c) $K_1 = 50$, $K_2 = 20$, d is constant, $h = 10^{-5}$, $y_k = y(hk)$, $r_k = r(hk)$, $q_k = q(hk)$ are known for $k = 1, 2, \ldots 30{,}000$. We can compute from here $v_k = K_1(r_k - y_k) - K_2 q_k$. Now $u_k = v_k + d$, but d is not known. The discretization of the left block of the plant is $G^d(z) = \frac{b}{z-a}$, where $a = e^{\lambda_2 h}$, $b = \frac{g}{\lambda_2}(e^{\lambda_2 h} - 1)$.

Note that $\lambda_2 h < 10^{-5}$, so that $b \approx gh$. The difference equation is $q_k - a q_{k-1} = b(d + v_{k-1})$ or, equivalently,

$$q_k = \underbrace{\begin{bmatrix} q_{k-1} & v_{k-1} & 1 \end{bmatrix}}_{\varphi_k} \underbrace{\begin{bmatrix} a \\ b \\ bd \end{bmatrix}}_{\theta} + e_k$$

θ = unknown parameters

where e_k is the modeling error.

From here, we can obtain a least squares estimate for θ in the usual way:

$$\Phi = \begin{bmatrix} \varphi_2 \\ \vdots \\ \varphi_N \end{bmatrix} \quad (N = 30{,}000), \quad \Phi^{\#} = (\Phi^* \Phi)^{-1} \Phi^*,$$

$\hat{\theta} = \Phi^{\#} \tilde{y}$, where $\tilde{y} = \begin{bmatrix} q_2 \\ \vdots \\ q_N \end{bmatrix}$. Once we have the estimate $\hat{\theta}$, we immediately get estimates for a, b and d, denoted $\hat{a}, \hat{b}, \hat{d}$. We have $\hat{g} = \hat{b}/h$, $\hat{\lambda}_2 = \frac{1}{h} \log \hat{a}$.

Now we discretize the right block of the plant, getting $G^{d'}(z) = \frac{b'}{z - a'}$, where $a' = e^{\lambda_1 h}$, $b' = \frac{1}{\lambda_1}(e^{\lambda_1 h} - 1) \approx h.$ Thus, b' is practically known. The difference equation is $y_k - a' y_{k-1} = b' q_{k-1}$, or, equivalently,

$$y_k - b' q_{k-1} = y_{k-1} a' + e'_k ,$$

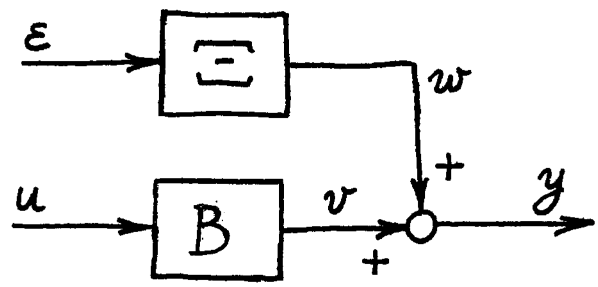
where e'_k is the modeling error. Introduce

$$\Phi' = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{N-1} \end{bmatrix}, \quad \Phi'^* \Phi' = \sum_{k=1}^{N-1} y_k^2, \quad \Phi^{\#} = \frac{[y \ \dots \ y_{N-1}]}{\sum_{k=1}^{N-1} y_k^2},$$

and we get the estimates

$$\hat{a'} = \sum_{k=1}^{N-1} (y_k - b' q_{k-1})\, y_k \Big/ \sum_{k=1}^{N-1} y_k^2, \quad \hat{\lambda}_1 = \frac{1}{h} \log \hat{a'}.$$

Question 2. [4] (a) Here is a block diagram corresponding to the problem statement. We have denoted by B the transfer function of Σ, so that



$$B(z) = b_0 + b_1 z^{-1} + b_2 z^{-2} \dots + b_{20} z^{-20}.$$

Recall that $\Xi(z) = c_0 + c_1 z^{-1} \dots + c_5 z^{-5}$. The MAX model is $\hat{y}(z) = B(z)\hat{u}(z) + \Xi(z)\hat{\varepsilon}(z)$, or, in time domain

$$y_k = b_0 u_k + b_1 u_{k-1} \dots + b_{20} u_{k-20} + c_0 \varepsilon_k \dots + c_5 \varepsilon_{k-5}.$$

Since Ξ^{-1} is (by assumption) stable, we can introduce the filtered signals

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

(these signals can be obtained by solving the auto-regressive equations $\Xi(z)\widehat{y^F}(z) = \hat{y}(z)$ and $\Xi(z)\widehat{u^F}(z) = \hat{u}(z)$ in the time domain) and then we get the X model

$$\widehat{y^F}(z) = B(z)\widehat{u^F}(z) + \hat{\varepsilon}(z).$$

[5]
(b) Taking the modeling error e into account, the X model derived earlier becomes

$$y_k^F = b_0 u_k^F + b_1 u_{k-1}^F \dots + b_{20} u_{k-20}^F + \varepsilon_k + e_k.$$

Since (ε_k) and (e_j) are independent white noise signals, the same is true for their sum.

Denote $\eta_k = \varepsilon_k + e_k$, then η is white noise.

We rewrite the X model:

$$y_k^F = \underbrace{\begin{bmatrix} u_k^F & u_{k-1}^F & \dots & u_{k-20}^F \end{bmatrix}}_{\varphi_k} \cdot \underbrace{\begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_{20} \end{bmatrix}}_{\theta} + \eta_k .$$

Here, θ is the vector of unknown parameters. The least squares approach is to minimize $\eta_1^2 + \eta_2^2 \dots \eta_N^2$, where $N = 6{,}000$. (It is also reasonable to start not from $k=1$ but from a higher value, in order to diminish the effect of initial conditions. For example, we could take the cost function $\eta_{100}^2 + \eta_{101}^2 + \dots + \eta_{6{,}000}^2$.)

Introduce $\Phi = \begin{bmatrix} \varphi_1 \\ \vdots \\ \varphi_N \end{bmatrix}$ (or, if you minimize only from $k=100$ to $k=6{,}000$ then in Φ start with φ_{100})

and denote $\Phi^{\#} = (\Phi^*\Phi)^{-1}\Phi^*$, $y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$.

Then the optimal estimate for θ is

$$\hat{\theta} = \Phi^{\#} y .$$

[5]

(c) According to the least squares theory, an estimate for $\operatorname{Var}(\eta)$ is $\widehat{\operatorname{Var}}(\eta) = \frac{1}{6{,}000-20} \| y - \Phi\hat{\theta} \|^2$, which can be computed from the given data. Then use $\operatorname{Var}(\eta) = \operatorname{Var}(\varepsilon) + \operatorname{Var}(e)$

where $Var(\varepsilon) = 1$ by assumption. Hence,

$$\widehat{Var(e)} = \frac{1}{5{,}980} \| y - \Phi \hat{\theta} \|^2 - 1 .$$

[6]

(d) To make the least squares estimate gradually "forget" old data, we introduce the cost function

$$V_\lambda = \eta_N^2 + \lambda \eta_{N-1}^2 + \lambda^2 \eta_{N-2}^2 \dots + \lambda^{N-1} \eta_1^2 ,$$

where $0 < \lambda < 1$. A recursive algorithm to minimize V_λ is

$$\begin{cases} P_n = \frac{1}{\lambda}\left[P_{n-1} - \frac{P_{n-1}\varphi_n^* \varphi_n P_{n-1}}{\lambda + \varphi_n P_{n-1} \varphi_n^*}\right], \\ \tilde{\varepsilon}_n = y_n - \varphi_n \hat{\theta}_{n-1} , \\ \hat{\theta}_n = \hat{\theta}_{n-1} + P_n \varphi_n^* \tilde{\varepsilon}_n , \end{cases}$$

where we can start from $\hat{\theta}_0 = 0$ (or any other guess, if we have a better one) and $P_0 = \alpha I$, with $\alpha > 0$. For large n, $\hat{\theta}_n$ computed recursively will converge to the optimal estimate based on the cost function V_λ. Practically, only the last $(1-\lambda)^{-1}$ data points matter. (The recursive formulas are listed in the section "Information for candidates", p. 1 of the exam paper.)

Question 3.

The block diagram corresponding to the problem statement and point (b):

$$e \longrightarrow \boxed{\Xi} \xrightarrow{\;u\;} \boxed{G} \longrightarrow y$$

[3]

(a) Since u is ergodic, we can estimate expectations by time averages:

$$\widehat{\mu_u} = \frac{1}{N}\sum_{k=1}^{N} u_k\,, \qquad \widehat{C_j^{uu}} = \frac{1}{N-j}\sum_{k=1}^{N-j} \tilde{u}_k \cdot \tilde{u}_{k+j}\,,$$

where $\tilde{u}_k = u_k - \widehat{\mu}_u$. This will work reasonably well since $j \le 50$, which is much less than $N = 5{,}000$.

[5]

(b) We represent u as $\hat{u} = \Xi \hat{e}$, where $e = (e_k)$ is white noise and $\Xi^{-1}(z) = 1 + g_1 z^{-1} + g_2 z^{-2} + \dots$, Ξ^{-1} is stable (like Ξ), so that $g_k \to 0$. We have

$$\begin{bmatrix} C_0^{uu} & C_1^{uu} & C_2^{uu} & \dots \\ C_1^{uu} & C_0^{uu} & C_1^{uu} & \dots \\ \vdots & \vdots & \vdots & \end{bmatrix} \cdot \begin{bmatrix} g_1 \\ g_2 \\ \vdots \end{bmatrix} = - \begin{bmatrix} C_1^{uu} \\ C_2^{uu} \\ \vdots \end{bmatrix},$$

which is an infinite sequence of equations in infinitely many unknowns $g_1, g_2, \dots$. However, if we assume that g_k is negligible for $k > m$ then we can truncate this to a linear system of m equations with m unknowns $g_1, \dots g_m$. The coefficients C_k^{uu} can be estimated, see part (a). If u is persistent of order m, then we can solve for $g_1, \dots g_m$. This gives us an auto-regressive realization of the estimated Ξ.

We have $E(u) = \Xi(1) \cdot E(e)$. Since $\Xi(1) = (1 + g_1 + g_2 \dots + g_m)^{-1}$ has been estimated and also $E(u) = \mu_u$ (see part (a)), we can now estimate $E(e)$. To estimate $Var(e)$, we use the Wiener-Lee formula for $\nu = 0$ $(e^{i\nu} = 1)$:

$$\underbrace{S^{uu}(1)}_{\sum_{j=-\infty}^{\infty} C_j^{uu}} = \underbrace{|\Xi(1)|^2}_{(1+g_1+g_2\dots)^{-2}} \underbrace{S^{ee}(1)}_{Var(e)}$$

We estimate $\widehat{S^{uu}(1)} = \sum_{j=-m+1}^{m-1} \widehat{C_j^{uu}}$ (recall that $C_{-j}^{uu} = C_j^{uu}$) and from here we obtain an estimate for $Var(e)$:

$$\widehat{Var(e)} = (1 + g_1 + g_2 \dots + g_m)^2 \sum_{j=-m+1}^{m-1} \widehat{C_j^{uu}} .$$

[5]

(c) We have (from $\Xi \hat{u} = \hat{e}$)

$$u_n = -g_1 u_{n-1} - g_2 u_{n-2} \dots - g_m u_{n-m} + e_n . \quad (*)$$

$\bar{u}_n$ is the conditional expectation of u_n, given the measurements $u_{n-1}, u_{n-2}, \dots$. If we take conditional expectations of both sides of $(*)$, we obtain

$$\bar{u}_n = -g_1 u_{n-1} - g_2 u_{n-2} \dots - g_m u_{n-m}.$$

Comparing this with $(*)$, we see that the prediction error is e_n. Hence, the variance of the prediction error is $Var(e)$ (which has been estimated in part (b)).

[3]

(d) We have $G(z) = \dfrac{bz^{-1}}{1+az^{-1}}$, hence

$y_k + a y_{k-1} = b u_{k-1}$, which we write as

$$y_k = \underbrace{\begin{bmatrix} -y_{k-1} & u_{k-1} \end{bmatrix}}_{\varphi_k} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\theta} + \varepsilon_k$$

where ε_k is the modeling error and θ is the vector of unknown parameters. Introducing

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

the estimate which minimizes $\varepsilon_1^2 + \varepsilon_2^2 \ldots + \varepsilon_N^2$ is $\hat{\theta} = \Phi^{\#} y$, where $y = [y_1 \; y_2 \ldots y_N]^T$.

[4]

(e) From $y_n = -a y_{n-1} + b u_{n-1} + \varepsilon_n$,

taking conditional expectations, we get

$$\overline{y_n} = -a y_{n-1} + b u_{n-1}.$$

Note that this has nothing to do with e_n and with predicting u_n, because we only need u_{n-1}. The prediction error is ε_n, and an estimate for $\mathrm{Var}(\varepsilon)$ is

$$\widehat{\mathrm{Var}(\varepsilon)} = \frac{1}{N-2} \| y - \Phi \hat{\theta} \|^2,$$

as we know from the least squares theory.

[5]

Question 4. (a) The impedance of the load model is Z given by

$$\frac{1}{Z(s)} = \frac{1}{R} + \frac{1}{Ls} + Cs = \frac{RLCs^2 + Ls + R}{RLs},$$

so that
$$Z(s) = \frac{RLs}{RLCs^2 + Ls + R}.$$

Hence
$$G(s) = \frac{Z(s)}{Z(s) + L_g s} = \frac{RLs}{RLs + L_g s(RLCs^2 + Ls + R)}$$

$$= \frac{\frac{1}{L_g C}}{s^2 + \underbrace{\frac{1}{RC}}_{a_1} s + \underbrace{\frac{L_g + L}{L_g L C}}_{a_0}} = \frac{b_0}{s^2 + a_1 s + a_0}.$$

The coefficients b_0, a_1, a_0 are unknown, but it is clear that they are positive, hence G is stable (because $a_0 > 0$ and $a_1 > 0$).

[6]
(b) The transfer function G discussed above is only a model. We denote by $G^e(i\omega_k)$ the experimentally determined values of the true transfer function at the frequencies ω_k (these values are subject to measurement errors). We have for $k = 1, 2, \ldots 40$

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

where e_k represents the combined effect of modeling and measurement errors (e_k is complex). We rewrite this:

$$\underbrace{(i\omega_k)^2 G^e(i\omega_k)}_{\eta_k} = \underbrace{\begin{bmatrix} -i\omega_k G^e(i\omega_k) & -G^e(i\omega_k) & 1 \end{bmatrix}}_{\varphi_k} \cdot \underbrace{\begin{bmatrix} a_1 \\ a_0 \\ b_0 \end{bmatrix}}_{\theta} + e_k$$

which looks like a standard least squares identification problem with θ the vector of unknown parameters. Since η_k and φ_k are complex, but we want θ to be real, we are searching for the optimal real vector θ which minimizes $|e_1|^2 + |e_2|^2 \ldots + |e_{40}|^2$. For this, we introduce

$$\tilde{\varphi}_k = \begin{cases} \mathrm{Re}\, \varphi_k & \text{for } k = 1, 2, \ldots 40, \\ \mathrm{Im}\, \varphi_{k-40} & \text{for } k = 41, 42, \ldots 80, \end{cases}$$

and similarly we introduce $\tilde{\eta}_k$, $\tilde{e}_k$ for $k = 1, \ldots 80$. Then we get 80 real equations $\tilde{\eta}_k = \tilde{\varphi}_k \theta + \tilde{e}_k$. Here, $|\tilde{e}_1|^2 + |\tilde{e}_2|^2 \ldots + |\tilde{e}_{80}|^2 = |e_1|^2 + |e_2|^2 \ldots + |e_{40}|^2$, so that we are still minimizing the same cost, but now θ is forced to be real. The optimal estimate $\hat{\theta}$ is found by introducing

$$\Phi = \begin{bmatrix} \tilde{\varphi}_1 \\ \vdots \\ \tilde{\varphi}_{80} \end{bmatrix}, \quad \Phi^{\#} = (\Phi^* \Phi)^{-1} \Phi^*, \quad \tilde{\eta} = \begin{bmatrix} \tilde{\eta}_1 \\ \vdots \\ \tilde{\eta}_{80} \end{bmatrix},$$

as in the standard least squares theory, and then

$$\hat{\theta} = \Phi^{\#} \tilde{\eta}.$$

After having estimated a_1, a_0 and b_0, we compute the estimated C (from b_0), then the estimated R (from a_1) and finally L from a_0:

$$C = \frac{1}{b_0 L_g}, \quad R = \frac{1}{a_1 C}, \quad L = \frac{L_g}{a_0 L_g C - 1}$$

(I have omitted the hats on top of each symbol in the last three formulas).

[4]
(c)

$$\left.\begin{array}{ll} A = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix}, & B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ C_0 = \begin{bmatrix} b_0 & 0 \end{bmatrix}, & D = [0], \end{array}\right\} \begin{array}{l}\text{minimal} \\ \text{realiza-} \\ \text{tion of} \\ G\end{array}$$

the eigenvalues of A are $\dfrac{-a_1 \pm \sqrt{a_1^2 - 4a_0}}{2}$.

[5]
(d) The exact discretization is:

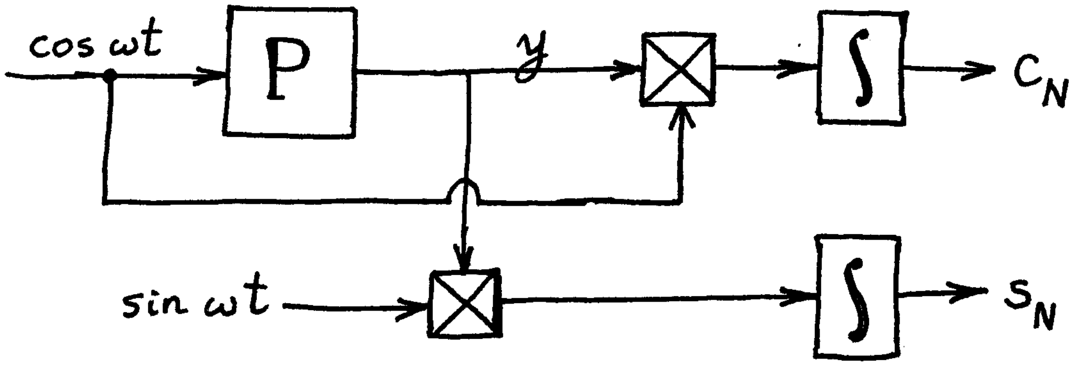
$$A^d = e^{Ah}, \quad B^d = (e^{Ah} - I) A^{-1} B,$$

$$G^d(z) = C_0 (zI - A^d)^{-1} B^d + D.$$

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

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

Question 5. [6] (a) The identification experiments can be done applying sinusoidal inputs at various frequencies in the relevant range ($\omega \leqslant 2\pi \cdot 1000$), in order to measure the corresponding gain A_ω and phase shift φ_ω :



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

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

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

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

[5] (b) Denote $C(s) = K_p + K_i/s$ (this is the transfer function of the PI controller), then the closed-loop transfer function from r to y is

$$G_1(s) = \frac{C(s)P(s)}{1 + C(s)P(s)} \quad \text{(assumed to be stable)}$$

and from d to y:

$$G_2(s) = \frac{P(s)}{1 + C(s)P(s)} \quad \text{(also stable).}$$

We have assumed that $\mathbf{P}(0) > 0$. Since $\mathbf{C}(0) = \infty$, it follows that $G_1(0) = 1$, $G_2(0) = 0$.

The reference r can be decomposed as

$$r(t) = 30 + e(t), \quad \text{where } \lim_{t \to \infty} e(t) = 0.$$

Thus, for large values of t, y will be a constant given by

$$y(t) = G_1(0) \cdot 30 + G_2(0) \cdot \underset{(d)}{3} = 30$$

(this shows that the PI controller solves the problem of tracking a constant r).

[6]

(c) $r = 0$, d = stationary, ergodic random signal, $E(d) = 3$. Then $\begin{bmatrix} d \\ y \end{bmatrix}$ is stationary and ergodic, in particular, y has these properties. We have $E(y) = G_2(0) \cdot E(d) = 0 \cdot 3 = 0$. By Wiener-Lee, $S^{yy} = |G_2|^2 \cdot S^{dd}$, hence

$$\mathrm{Var}(y) = C^{yy}(0) = \frac{1}{2\pi} \int_{-\infty}^{\infty} S^{yy}(i\omega)\, d\omega$$

$$G_2 = \frac{\mathbf{P}}{1 + \mathbf{C}\mathbf{P}}$$

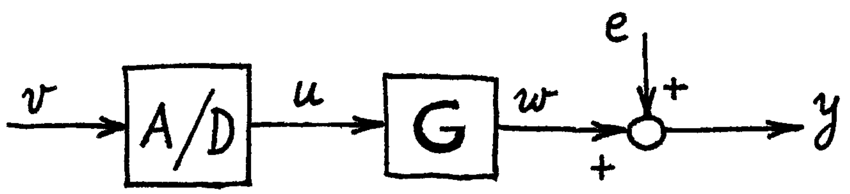
$$= \frac{1}{2\pi} \int_{-\infty}^{\infty} |G_2(i\omega)|^2 S^{dd}(i\omega)\, d\omega$$

[3]

(d) If r is a given function (of time) and d is a stationary random signal, denote by y_r and y_d the components of y due to r and d. y_r is a given (non-constant) function and y_d is a stationary random signal. Now $E(y(t)) = y_r(t) + E(y_d(t))$.

The second component $E(y_d(t))$ is constant (actually it is zero, see part (c)) so that $E(y(t))$ is not constant. Hence, y cannot be stationary.

Question 6. The block diagram corresponding to the problem statement:



$\left(\begin{array}{l} e = \text{white noise,} \\ \text{independent of } v \end{array} \right)$

[3]

(a) $u_k = v(hk)$ is stationary and ergodic, hence $\begin{bmatrix} u \\ w \end{bmatrix}$ is stationary and ergodic. Since e is independent of v, it is also independent of $\begin{bmatrix} u \\ w \end{bmatrix}$ (which is generated from v). Hence, $\begin{bmatrix} u \\ y \end{bmatrix}$ is stationary and ergodic. Stationarity of $\begin{bmatrix} u \\ y \end{bmatrix}$ (also called "joint stationarity of u, y") means that the distribution functions of

$$\begin{bmatrix} u_m \\ y_n \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} u_{m+\tau} \\ y_{n+\tau} \end{bmatrix}$$

are equal, for any combination of $m, n, \tau \in \mathbb{Z}$. Ergodicity of $\begin{bmatrix} u \\ y \end{bmatrix}$ (also called "joint ergodicity") means that for any measurable function g of several variables,

$$E\left(g\left(\begin{bmatrix} u_{k_1} \\ y_{k_1} \end{bmatrix}, \dots \begin{bmatrix} u_{k_n} \\ y_{k_n} \end{bmatrix}\right)\right) = \lim_{N \to \infty} \frac{1}{2N} \sum_{j=-N}^{N} g\left(\begin{bmatrix} u_{k_1+j} \\ y_{k_1+j} \end{bmatrix}, \dots \begin{bmatrix} u_{k_n+j} \\ y_{k_n+j} \end{bmatrix}\right)$$

with probability 1 (i.e., expectations are equal to infinite time averages).

[3]

(b) We have $C_k^{uu} = C^{vv}(hk) = e^{-100h|k|}$, where $h = 10^{-4}$. In particular, the power of u is $\mathrm{Var}(u) = C_0^{uu} = 1$. The power spectral density of u is (for $|z| = 1$)

$$S^{uu}(z) = \sum_{k=-\infty}^{\infty} z^{-k} C_k^{uu}$$

denote $\rho = e^{-100h}$

$$= \sum_{k=-\infty}^{-1} z^{-k} e^{-100h|k|} + \sum_{k=0}^{\infty} z^{-k} \left(e^{-100h}\right)^k$$

$$= \underbrace{\sum_{k=1}^{\infty} z^k \rho^k}_{} + \underbrace{\sum_{k=0}^{\infty} z^{-k} \rho^k}_{}$$

$$= \frac{\rho}{z^{-1} - \rho} + \frac{\rho}{z - \rho} + 1$$

because $z^{-1} = \bar{z}$

$$= \frac{z - \rho + \bar{z} - \rho}{(\bar{z} - \rho)(z - \rho)} \cdot \rho + 1$$

$$= \frac{(z + \bar{z}) - 2\rho}{1 - (z + \bar{z})\rho + \rho^2} \cdot \rho + 1$$

If we denote $z = e^{i\nu}$ ($\nu \in (-\pi, \pi]$) then we obtain

$$\boxed{\begin{aligned} S^{uu}(e^{i\nu}) &= 1 + 2\rho \frac{\cos\nu - \rho}{1 - 2\cos\nu\,\rho + \rho^2} \\ &= \frac{1 - \rho^2}{1 - 2\cos\nu\,\rho + \rho^2}. \end{aligned}}$$

(Any of the last 5 formulas is a good answer)

[3]
(c) We can estimate $E(y_k)$ and C^{uy}_{τ}, using ergodicity, as follows:

$$\widehat{E(y_k)} = \frac{1}{N}\sum_{k=1}^{N} y_k \, , \quad \widehat{C^{uy}_{\tau}} = \frac{1}{N-\tau}\sum_{k=1}^{N-\tau} \tilde{u}_k \tilde{y}_{k+\tau}$$

where $\tilde{u}_k = u_k - 3$, $\tilde{y}_k = y_k - \widehat{E(y_k)}$, $N = 5{,}000$ and $\tau << N$. Similarly, $\widehat{Var(y_k)} = \frac{1}{N}\sum_{k=1}^{N} (\tilde{y}_k)^2$.

[4]
(d) We can estimate $g_0, g_1, \dots g_{49}$ by solving

$$\begin{bmatrix} C^{uu}_0 & C^{uu}_1 & \dots & C^{uu}_{49} \\ C^{uu}_1 & C^{uu}_0 & \dots & C^{uu}_{48} \\ \vdots & \vdots & & \vdots \\ C^{uu}_{49} & C^{uu}_{48} & \dots & C^{uu}_0 \end{bmatrix} \cdot \begin{bmatrix} g_0 \\ g_1 \\ \vdots \\ g_{49} \end{bmatrix} = \begin{bmatrix} \widehat{C^{uy}_0} \\ \widehat{C^{uy}_1} \\ \vdots \\ \widehat{C^{uy}_{49}} \end{bmatrix} .$$

This is because $C^{uw}_k = C^{uy}_k$, since $y = w + e$ and e is independent of u.

[3]
(e) u is persistent of order m if the $m \times m$ Toeplitz matrix shown above for $m = 50$ is invertible. If S^{uu} is rational and $S^{uu}(e^{i\nu}) > 0$ for all $\nu \in (-\pi, \pi]$, then u is persistent of any order — this is the case for our u, as can be seen from part (b).

[4]
(f) $C^{ww} = \check{g} * C^{uw} = \check{g} * C^{uy}$, hence $C^{ww}_0 = \sum_{k=0}^{\infty} g_k C^{uy}_k = Var(w)$. Since e is independent of u, we have $Var(y) = Var(w) + Var(e)$, hence $Var(e) = Var(y) - \sum_{k=0}^{\infty} g_k C^{uy}_k$.