1964-01

Control of nonlinear systems

Gilchrist, Richard B.

Cambridge, Massachusetts: Massachusetts Institute of Technology

https://hdl.handle.net/10945/11856

CONTROL OF NONLINEAR SYSTEMS

RICHARD B. GILCHRIST

CONTROL OF NONLINEAR SYSTEMS

by

RICHARD B. GILCHRIST

Lieutenant, United States Navy

B. S., United States Naval Academy
(1955)

SUBMITTED IN PARTIAL FULFILLMENT OF THE

REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

January, 1964

CONTROL OF NONLINEAR SYSTEMS

by

RICHARD B. GILCHRIST

Lieutenant, United States Navy

Submitted to the Department of Electrical Engineering on December 10, 1963, in partial fulfillment of the requirements for the degree of Doctor of Science.

ABSTRACT

Dynamic programming is employed to obtain a solution to the problem of controlling a nonlinear system in an optimal fashion, subject to a quadratic performance index. The technique used is similar to that given by Merriam and Kalman for linear systems.

For some special nonlinear systems, the solution can be computed by direct application of this technique. As an example, the optimal control system for a freely spinning body is determined.

For more general nonlinear systems, the solution cannot be obtained directly. However, it is possible to obtain a solution indirectly. This is done by first linearizing the vector-state equations representing the nonlinear system. Next, dynamic programming is used to obtain an approximate solution based on the linearized state equations. Then an iterative procedure for improving the solution is presented. It can be shown that if the iterative procedure converges, it converges to the exact solution of the optimal nonlinear control problem.

Computer example problems are given to illustrate the method, and to indicate the convergence that is usually achieved. In addition, the performance of the optimal control system is compared with the performance of a simple sub-optimal control system for some of the example problems given.

Thesis Supervisor: Leonard A. Gould

Title: Associate Professor of Electrical Engineering

ACKNOWLEDGMENTS

The author acknowledges the supervision and assistance throughout the course of this research of Professor Leonard A. Gould, who served both as Faculty Counselor and as Thesis Supervisor. The author is also indebted to Professor George C. Newton and Dr. Michael Athans, who served as readers and who also offered advice and guidance.

The entire doctoral program of the author at M. I. T. was under the sponsorship of the Bureau of Ships, United States Navy. Captain Edward S. Arentzen, U. S. N. (Ret.), Captain William M. Nicholson, U. S. N., Commander John R. Baylis, U. S. N., and Commander William R. Porter, U. S. N., who served as Professors of Naval Architecture at M. I. T., supervised the author's program of studies and gave much advice and encouragement.

An important aspect of the author's doctoral studies was his close association with his colleague, Lieutenant Henry Cox, U. S. N. His many suggestions and discussions concerning this research are deeply appreciated.

For the arduous and tedious task of typing the thesis, the author owes a special thanks to Mrs. Gena M. Seipel and Mrs. Beverly A. Mayer of the U. S. Navy Underwater Sound Laboratory.

TABLE OF CONTENTS

TABLE OF CONTENTS (CONT'D)

CHAPTER I

SUMMARY

1.1 Introduction

During the past decade, a new approach to automatic control has been developed principally as the result of work by Bellman[1-3] and Kalman[4-7] in this country, and Pontryagin[8] in the U. S. S. R. This approach, which is now commonly called the "theory of optimal control systems," differs from the now classical approach to automatic control of Newton, Gould, and Kaiser,[9] for instance, in that it uses a vector differential equation description of the system instead of a transfer function description, and it concentrates on the time domain methods of analysis and synthesis, instead of frequency domain methods. The theory of optimal control has made use of the calculus of variation,[7,10] and the new but related "dynamic programming" of Bellman,[11,12] as well as the "maximum principle" of Pontryagin.

Useful results of the application of these methods to optimal control problems have been obtained primarily for linear systems. Useful results have been obtained for nonlinear systems in only a few very special cases.[13,14] It is the objective of this work to extend to nonlinear systems some techniques that have been successful in the design of controls for linear systems.

1.2 Notation and Terminology

An attempt has been made to keep the notation and terminology consistent with current literature. In particular, the notation used by Kalman has been used whenever practicable.

Vectors are designated by underlined lower case letters. All vectors are understood to be column vectors. For example, the vector $\underline{x}$ denotes

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_n \end{bmatrix} \tag{1.1}$$

Similarly, matrices are designated by underlined upper case letters. For example, the matrix $\underline{A}$ denotes

$$\underline{A} = \begin{bmatrix} a_{11} & a_{12} \cdots a_{1n} \\ a_{21} & a_{22} \cdots a_{2n} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{n1} & a_{n2} \cdots a_{nn} \end{bmatrix} \tag{1.2}$$

The transpose of a vector or a matrix is designated by a prime. Thus

$$\underline{x}' = [x_1 \; x_2 \cdots x_n] \tag{1.3}$$

and

$$\underline{A}' = \begin{bmatrix} a_{11} & a_{21} \cdots a_{n1} \\ a_{12} & a_{22} \cdots a_{n2} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{1n} & a_{2n} \cdots a_{nn} \end{bmatrix} \tag{1.4}$$

The inner product of two vectors is denoted by $\underline{x}'\underline{y}$, and is given by

$$\underline{x}'\underline{y} = \sum_{i=1}^{n} x_i y_i \tag{1.5}$$

Consistent with this, the square of the Euclidian norm, denoted by $||\underline{x}||^2$, is given by

$$||\underline{x}||^2 = \underline{x}'\underline{x} \tag{1.6}$$

The quadratic form of a vector with respect to a symmetric matrix $\underline{A}$, is given by $\underline{x}'\underline{Ax}$. For convenience, it is often indicated by

$$||\underline{x}||^2_{\underline{A}} = \underline{x}'\underline{Ax} \tag{1.7}$$

The derivative of a vector or a matrix with respect to the scaler variable, time, is indicated by the notation,

$$\dot{\underline{x}} = d\underline{x}/dt = \begin{bmatrix} dx_1/dt \\ dx_2/dt \\ \cdot \\ \cdot \\ \cdot \\ dx_n/dt \end{bmatrix} \tag{1.8}$$

and

$$\dot{\underline{A}} = d\underline{A}/dt = \begin{bmatrix} da_{11}/dt & da_{12}/dt \ldots da_{1n}/dt \\ da_{21}/dt & da_{22}/dt \ldots da_{2n}/dt \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ da_{n1}/dt & da_{n2}/dt \ldots da_{nn}/dt \end{bmatrix} \tag{1.9}$$

The gradient of a scalar function of $\underline{x}$ is denoted by

$$V_{\underline{x}}(\underline{x}) = \text{grad } V(\underline{x}) = \begin{bmatrix} \partial V(\underline{x})/\partial x_1 \\ \partial V(\underline{x})/\partial x_2 \\ \cdot \\ \cdot \\ \cdot \\ \partial V(\underline{x})/\partial x_n \end{bmatrix} \tag{1.10}$$

Similarly, the Jacobian matrix of a vector function of $\underline{x}$ is denoted by

$$f_{\underline{x}}(\underline{x}) = \begin{bmatrix} \partial f_1(\underline{x})/\partial x_1 & \partial f_1(\underline{x})/\partial x_2 \ldots \partial f_1(\underline{x})/\partial x_n \\ \partial f_2(\underline{x})/\partial x_1 & \partial f_2(\underline{x})/\partial x_2 \ldots \partial f_2(\underline{x})/\partial x_n \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \partial f_n(\underline{x})/\partial x_1 & \partial f_n(\underline{x})/\partial x_2 \ldots \partial f_n(\underline{x})/\partial x_n \end{bmatrix} \tag{1.10a}$$

1.3 Problem Statement

Consider the system described by the vector differential equations

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}(t), \underline{u}(t), t); \quad \underline{x}(0) = \underline{c} \tag{1.11}$$

$$\underline{y}(t) = \underline{h}(\underline{x}(t), t) \tag{1.12}$$

where $\underline{x}(t)$ is the system state vector and $\underline{y}(t)$ is the system output vector. For this system, it is desired to find the control vector, $\underline{u}(t)$, such that a performance index, $J(t)$, is a minimum. In particular, we will assume that $J(t)$ has the quadratic form,

$$J(t) = \int_t^T \left\{ \frac{1}{2} \left\| \underline{z}(\tau) - \underline{y}(\tau) \right\|^2_{\underline{Q}(\tau)} + \frac{1}{2} \left\| \underline{u}(\tau) \right\|^2_{\underline{R}(\tau)} \right\} d\tau \tag{1.13}$$

where $z(\tau)$ is the system desired output, and $\underline{Q}(\tau)$ and $\underline{R}(\tau)$ are positive definite matrices weighting the system error and control effort, respectively. We will require that the control, $\underline{u}(t)$, be expressed as $\underline{u}(\underline{x}(t), \underline{z}(t))$ so that it can be realized in a feedback configuration.

It is mathematically convenient to consider first the discrete time version of the same problem for the theoretical development. Actually, the discrete time version is a meaningful problem in its own right. It is this version that applies when a digital computer is used to synthesize the controller.

For the discrete time problem, the equations

$$\underline{x}(k+1) = \underline{f}(\underline{x}(k), \underline{u}(k), k); \quad \underline{x}(0) = \underline{c} \tag{1.14}$$

$$\underline{y}(k) = \underline{h}(\underline{x}(k), k) \tag{1.15}$$

replace equations (1.11) and (1.12), and

$$J(k) = \sum_{j=k}^{N} \frac{1}{2} \left\| \underline{z}(j) - \underline{y}(j) \right\|^2_{\underline{Q}(j)} + \sum_{j=k}^{N-1} \frac{1}{2} \left\| \underline{y}(j) \right\|^2_{\underline{R}(j)} \tag{1.16}$$

replaces equation (1.13).

1.4 Solution of the Discrete Time Problem

The solution of the discrete time nonlinear optimal control problem is sketched here. For a detailed solution, see Chapter II.

In order to proceed by dynamic programming, we define the value function

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k), \ldots, \underline{u}(N-1)}{\text{Min}} \{J(k)\} \tag{1.17}$$

Then by the "principle of optimality," it follows that

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k)}{\text{Min}} \left\{ \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{u}(k) \|^2_{\underline{R}(k)} + V_{N-k-1}(\underline{x}(k+1)) \right\} \tag{1.18}$$

An approximate solution to this equation can be obtained by assuming

$$\underline{x}(k+1) \simeq \underline{f}(\underline{x}^*(k), \underline{u}^*(k), k) + f_{\underline{x}}(\underline{x}^*(k), \underline{u}^*(k), k)[\underline{x}(k) - \underline{x}^*(k)] + f_{\underline{u}}(\underline{x}^*(k), \underline{u}^*(k), k)[\underline{u}(k) - \underline{u}^*(k)] \tag{1.19}$$

$$\underline{y}(k) \simeq \underline{h}(\underline{x}^*(k), k) + h_{\underline{x}}(\underline{x}^*(k), k)[\underline{x}(k) - \underline{x}^*(k)] \tag{1.20}$$

and

$$V_{N-k}(\underline{x}(k)) = \frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k)\, \hat{\underline{x}}(k) + a(k) \tag{1.21}$$

where $\underline{P}(k)$, $\hat{\underline{x}}(k)$, and $a(k)$ are a parametric matrix, vector, and scalar to be determined, and where $\underline{x}^*(k)$ and $\underline{u}^*(k)$ are as yet unspecified points about which we linearize.

The approximate solution obtained by combining equations (1.18), (1.19), (1.20), and (1.21) is given by the equations

$$\underline{u}(k) = -[\underline{R}(k) + f_{\underline{u}}' \underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}' [\underline{P}(k+1) f_{\underline{x}} \underline{x}(k) + \underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1)] \tag{1.22}$$

$$\underline{P}(k) = h_{\underline{x}}' \underline{Q}(k) h_{\underline{x}} + f_{\underline{x}}' \underline{M}(k) \underline{P}(k+1) f_{\underline{x}} \tag{1.23}$$

$$\hat{\underline{x}}(k) = f_{\underline{x}}' \underline{M}(k) [\underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1)] - h_{\underline{x}}' \underline{Q}(k) [\underline{z}(k) - \underline{c}(k)] \tag{1.24}$$

$$a(k) = a(k+1) + \frac{1}{2} \| \underline{z}(k) - \underline{c}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{b}(k) \|^2_{\underline{P}(k+1)} + \underline{b}'(k)\, \hat{\underline{x}}(k+1)$$
$$- \frac{1}{2} \| \underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1) \|^2_{f_{\underline{u}}[\underline{R}(k) + f_{\underline{u}}' \underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}'} \tag{1.25}$$

where

$$\underline{M}(k) = \underline{I} - \underline{P}(k+1)\, f_{\underline{u}} \left[\underline{R}(k) + f_{\underline{u}}' \, \underline{P}(k+1)\, f_{\underline{u}} \right]^{-1} f_{\underline{u}}' \tag{1.26}$$

$$\underline{b}(k) = \underline{f} - f_{\underline{x}}\, \underline{x}^{*}(k) - f_{\underline{u}}\, \underline{u}^{*}(k) \tag{1.27}$$

and

$$\underline{c}(k) = \underline{h} - h_{\underline{x}}\, \underline{x}^{*}(k) \tag{1.28}$$

In the above equations the arguments for $\underline{f}$, $f_{\underline{x}}$, $f_{\underline{u}}$, $\underline{h}$, and $h_{\underline{x}}$ have been omitted for simplicity. They are understood to be evaluated at $\underline{x}^{*}(k)$, $\underline{u}^{*}(k)$ and k, as appropriate.

The boundary conditions for equations (1.23), (1.24), and (1.25) can be obtained from equations (1.16) and (1.21). They are

$$\underline{P}(N+1) = \underline{0} \tag{1.29}$$

$$\hat{\underline{x}}(N+1) = \underline{0} \tag{1.30}$$

$$a\,(N+1) = 0 \tag{1.31}$$

Notice that equations (1.23), (1.24), and (1.25) must be solved backward in time, starting at time, $N+1$, where the boundary conditions are known, and working backward to the present time k. This implies that the desired output, $\underline{z}(k)$, must be known in advance so that the parameters $\underline{P}$ and $\underline{x}$ can be pre-computed. Once these parameters are known, the control system can be synthesized. Figure 1.1 shows a block diagram of the control system for the discrete time nonlinear optimal control problem.

From figure 1.1 it can be seen that the controller for the system consists of a time varying linear feedback portion, and a director portion. The feedback portion of the controller will insure that the system will be relatively insensitive to state or parameter perturbations occurring in the system being controlled.

The question of stability, which is of paramount importance in any control system, can be answered by the use of the second method of Lyapunov. By using this method, it can be shown that the control systems

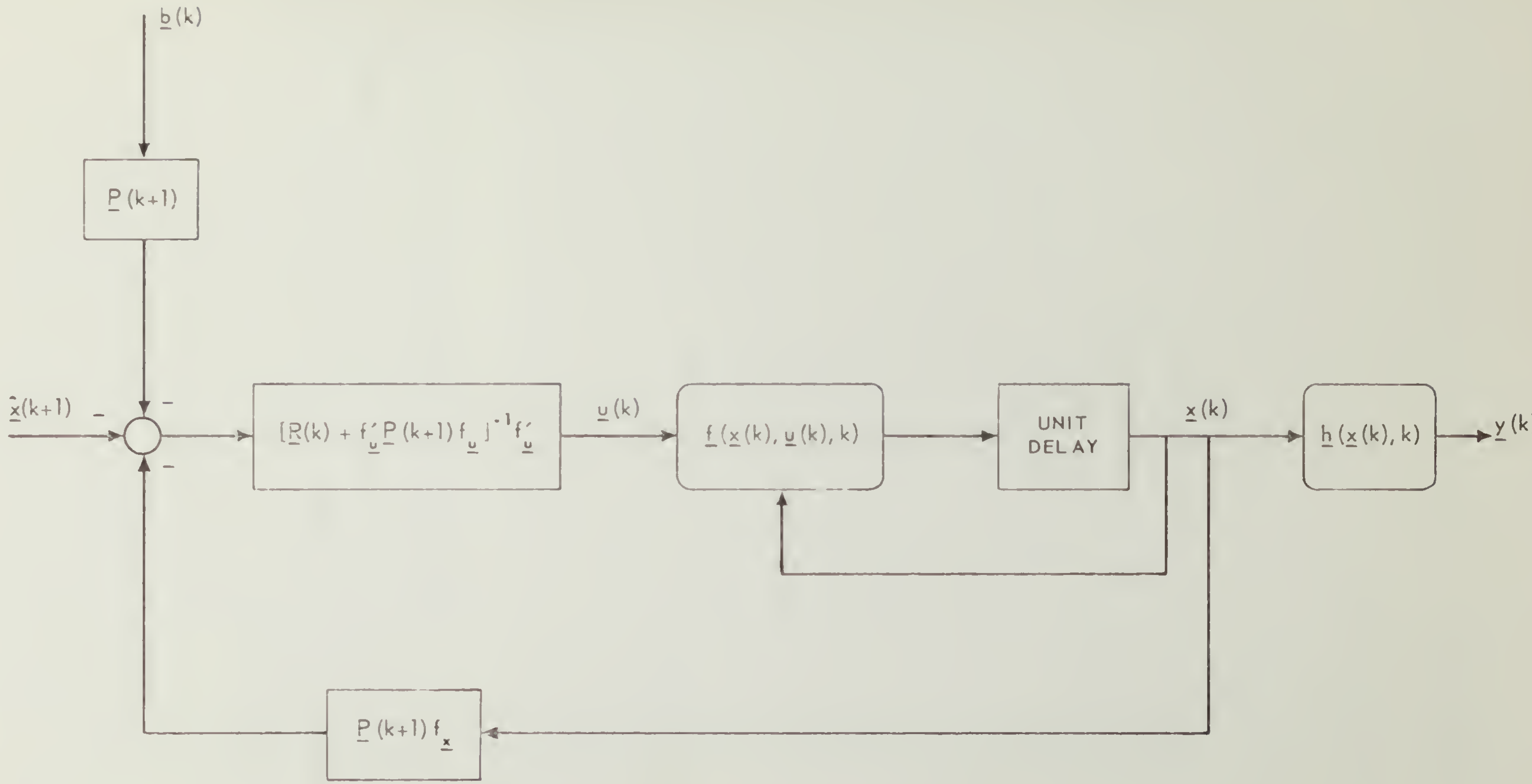


Figure 1.1 - Nonlinear Control System

designed from the theory presented here are always stable. A more detailed discussion of stability is contained in Appendix B.

The theory outlined above provides an approximately optimal solution, only. How near optimal the solution is depends on how near the vectors $\underline{x}^*(k)$ and $\underline{u}^*(k)$, which must be given beforehand, are to the actual state and control vectors, $\underline{x}(k)$ and $\underline{u}(k)$. An exact solution to the nonlinear control problem can be obtained by solving the equations for the approximately optimal solution in an iterative fashion.

At each iteration, the $\underline{x}(k)$ and $\underline{u}(k)$ determined on the previous iteration are used for the $\underline{x}^*(k)$ and $\underline{u}^*(k)$. If convergence is achieved by this procedure, the solution obtained is the exact solution to the nonlinear control problem. The question of under what conditions the iterative procedure converges is still unanswered, but experience using this algorithm on a digital computer indicates that convergence occurs for a broad range of problems, and that convergence is usually achieved in three or four iterations.

The theory presented in this section can be extended to systems with stochastic disturbances by minor modifications. However, the iterative algorithm does not produce an exact solution in this case. Details for the problem when stochastic disturbances are present are given in section 2. 7.

1. 5 Solution of the Continuous Time Problem

The equations specifying the solution to the continuous time problem may be obtained by dynamic programming in a manner analogous to that used for the discrete time problem. These equations are

$$\underline{u}(t) = -\underline{R}^{-1}(t)\, f_{\underline{u}}' [\underline{P}(t)\, \underline{x}(t) + \hat{\underline{x}}(t)] \tag{1.32}$$

$$\dot{\underline{P}}(t) = \underline{P}(t)\, f_{\underline{u}}\, \underline{R}^{-1}(t)\, f_{\underline{u}}'\, \underline{P}(t) - h_{\underline{x}}'\, \underline{Q}(t)\, h_{\underline{x}} - \underline{P}(t)\, f_{\underline{x}} - f_{\underline{x}}'\, \underline{P}(t) \tag{1.33}$$

$$\dot{\hat{\underline{x}}}(t) = h_{\underline{x}}\, \underline{Q}(t)\, [\underline{z}(t) - \underline{c}(t)] + \underline{P}(t) f_{\underline{u}} \underline{R}^{-1}(t) f_{\underline{u}}'\, \hat{\underline{x}}(t) - \underline{P}(t)\underline{b}(t) - f_{\underline{x}}\, \hat{\underline{x}}(t) \tag{1.34}$$

$$\dot{a}(t) = -\frac{1}{2}\,||\underline{z}(t) - \underline{c}(t)||^2_{\underline{Q}(t)} + \frac{1}{2}\,||\hat{\underline{x}}(t)||^2_{f_{\underline{u}} \underline{R}^{-1}(t) f'_{\underline{u}}} + \hat{\underline{x}}(t)\,\underline{b}(t) \tag{1.35}$$

with the boundary conditions

$$\underline{P}(T) = \underline{0} \tag{1.36}$$

$$\hat{\underline{x}}(T) = \underline{0} \tag{1.37}$$

$$a(T) = 0 \tag{1.38}$$

Chapter III contains a full development of the theory for the continuous time problem. The question of system stability is discussed with reference to the continuous time problem in Appendix B.

1.6 An Analytic Example

Consider the equations of motion of a freely spinning body about three mutually perpendicular axes,

$$\dot{x}_1 = a_1 x_2 x_3 + u_1; \quad x_1(0) = c_1 \tag{1.39}$$

$$\dot{x}_2 = a_2 x_1 x_3 + u_2; \quad x_2(0) = c_2 \tag{1.40}$$

$$\dot{x}_3 = a_3 x_1 x_2 + u_3; \quad x_3(0) = c_3 \tag{1.41}$$

where x_1, x_2, and x_3 are the angular velocities, where u_1, u_2, and u_3 are controls proportional to torques, and where

$$a_1 + a_2 + a_3 = 0 \tag{1.42}$$

These equations are nonlinear and coupled.

We wish to determine x_1, x_2, and x_3 such that the performance index

$$J = \int_0^T \left\{ \frac{1}{2}\, q(t)\,[x_1^2 + x_2^2 + x_3^2] + \frac{1}{2}\, r(t)\,[u_1^2 + u_2^2 + u_3^2] \right\} dt \tag{1.43}$$

is a minimum.

The solution to this problem can be obtained exactly and analytically, it turns out, if we proceed in the same manner as that indicated in the previous section. The solution is

$$u_1(t) = -k(t)x_1(t) \tag{1.44}$$

$$u_2(t) = -k(t)x_2(t) \tag{1.45}$$

$$u_3(t) = -k(t)x_3(t) \tag{1.46}$$

where

$$k(t) = p(t)/r(t) \tag{1.47}$$

and where

$$\dot{p}(t) = p^2(t)/r(t) - q(t); \qquad p(T) = 0 \tag{1.48}$$

for $r(t)$ and $q(t)$ constant, that is

$$r(t) = r \tag{1.49}$$

$$q(t) = q \tag{1.50}$$

the solution of equation (1.48) is

$$p(\tau) = rk(\tau) = r\alpha\left[\frac{1 - e^{-2\alpha\tau}}{1 + e^{-2\alpha\tau}}\right] \tag{1.51}$$

where

$$\alpha = \sqrt{q/r} \tag{1.52}$$

and

$$\tau = T - t \tag{1.53}$$

A block diagram of this control system is shown in figure 1.2.

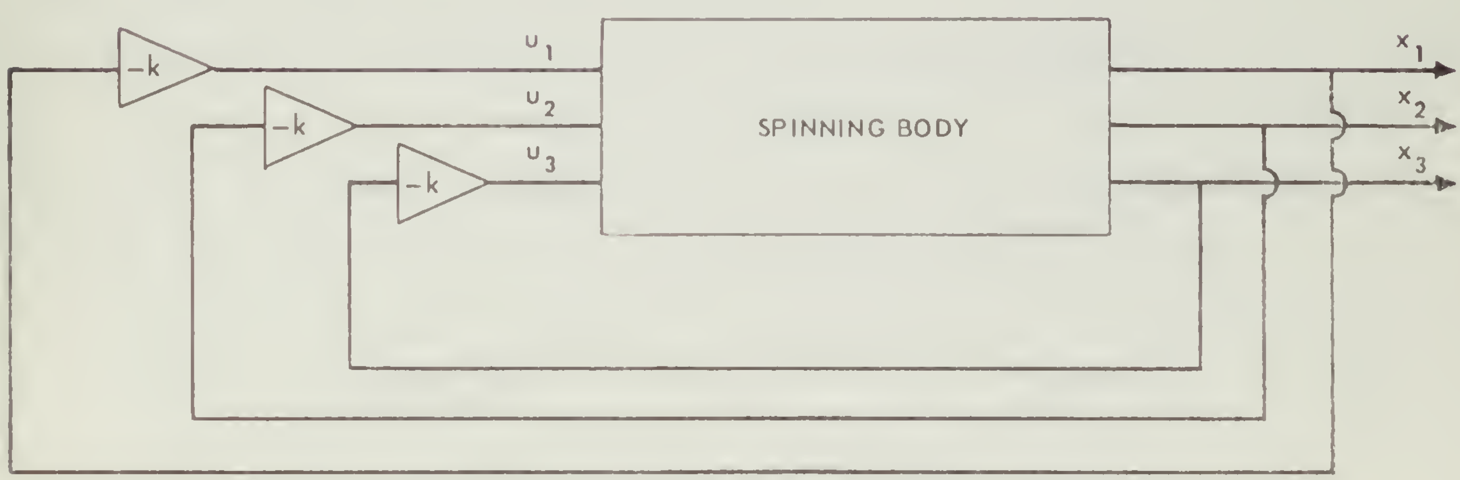


Figure 1.2

A detailed derivation of the control equations for this system, as well as a comparison of this control system with a sub-optimal one that uses constant gain linear feedback, is contained in Chapter IV.

1.7 Computer Examples

Consider the system described by the nonlinear equations

$$x(k+1) = x(k) - 0.05\,x^3(k) + 0.05\,u(k); \qquad x(1) = 1.0 \tag{1.54}$$

$$y(k) = x(k) \tag{1.55}$$

We wish to determine $u(1), \ldots, u(99)$ such that the performance index

$$J = \sum_{k=1}^{100} \frac{1}{2} Q\,[z(k) - x(k)]^2 + \sum_{k=1}^{99} \frac{1}{2} R\,u^2(k) \tag{1.56}$$

is a minimum.

The equations that form the basis for the iterative solution to this problem are given by equations (1.19), (1.22), (1.23), and (1.24). In this problem, all the variables appearing in these equations should be interpreted as scalars. Figure 1.3 shows the results of the computer solution of this problem for the case when $R = 0.01$, $Q = 10.0$, and $z(k) = 0$ for $k < 50$, but $z(k) = 1.0$ for $k \geq 50$. The iteration procedure converged (based on a convergence criterion of a 1 percent change in the performance index) in three iterations. The performance index on the third iteration was 12.272.

A sub-optimal controller, with the control determined by

$$u(k) = G\,[z(k) - x(k)] \tag{1.57}$$

where G was equal to a constant gain of 15.0, when operated with the same nonlinear system gave a performance index of 13.845.

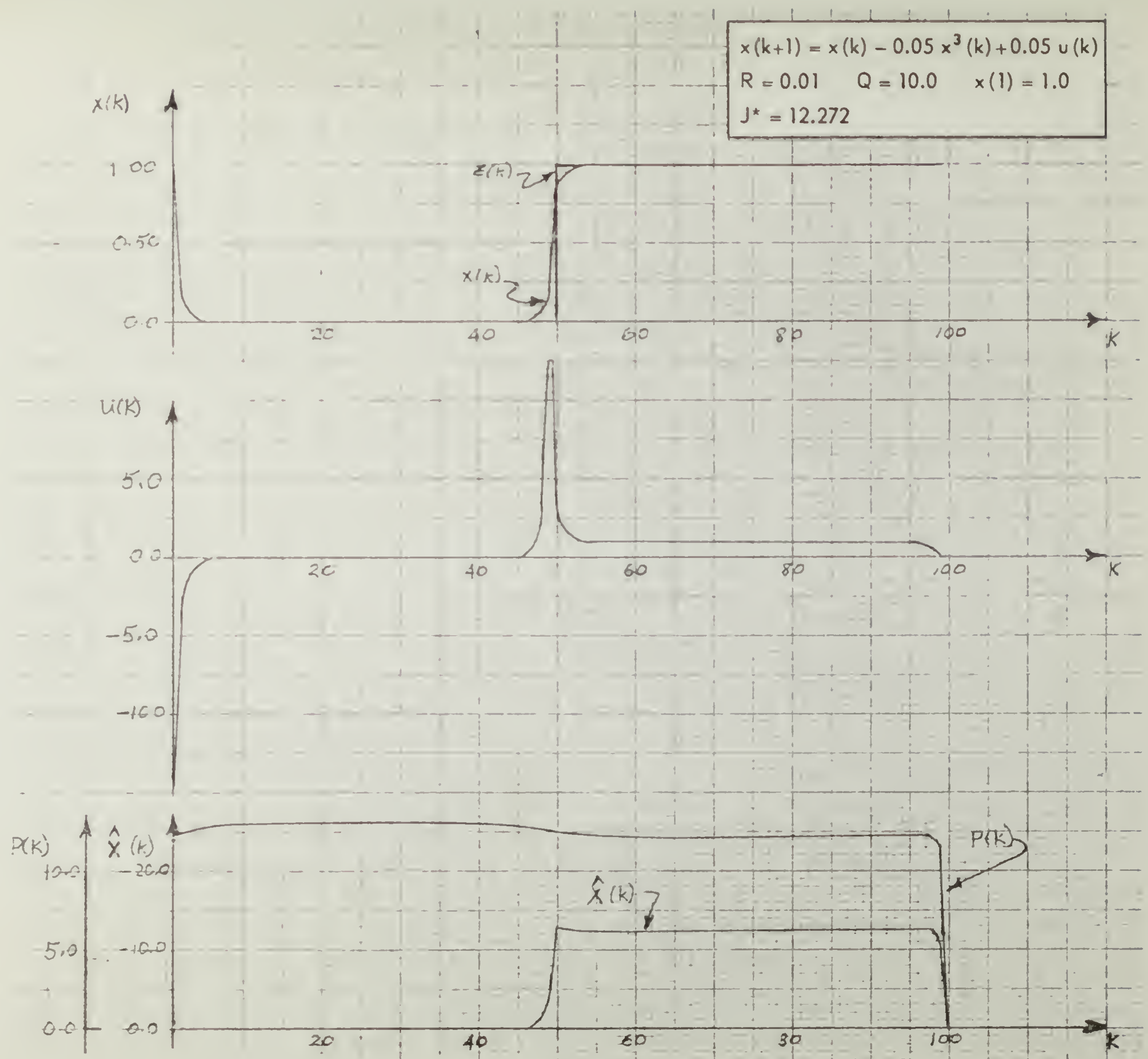
As a second example consider the system described by the equations

$$x_1(k+1) = x_1(k) + 0.01\,x_2(k); \qquad x_1(1) = 0.0 \tag{1.58}$$

$$x_2(k+1) = x_2(k) - 0.02\,x_1(k) - 0.03\,|x_2(k)|\,x_2(k) + 0.05\,u(k); \qquad x_2(1) = 3.0 \tag{1.59}$$

$$y_1(k) = x_1(k) \tag{1.60}$$

$$y_2(k) = x_2(k) \tag{1.61}$$



x(k+1) = x(k) − 0.05 x³(k) + 0.05 u(k)
R = 0.01 Q = 10.0 x(1) = 1.0
J* = 12.272
x(k)
1.00
0.50
0.0
z(k)
x(k)
20
40
60
80
100
K
u(K)
5.0
0.0
−5.0
−10.0
P(K)
X̂(k)
10.0
−20.0
5.0
−10.0
0.0

Figure 1.3

Again, we wish to control this system such that the performance index

$$J = \sum_{k=1}^{100} \left\{ \frac{1}{2} Q_1 [z_1(k) - x_1(k)]^2 + \frac{1}{2} Q_2 [z_2(k) - x_2(k)]^2 \right\} + \sum_{k=1}^{99} \frac{1}{2} R u^2(k) \quad (1.62)$$

is a minimum.

The two-dimensional version of equations (1. 19), (1. 22), (1. 23), and (1. 24) form the basis for the iterative solution procedure. Figure 1.4 shows the results of the computer solution of this problem for the case when $R = 0.01$, $Q_1 = 1.0$, $Q_2 = 1.0$, $z_1(k) = 0$ and $z_2(k) = 0$. Convergence was achieved in four iterations, and the performance index on the fourth iteration was 29. 29.

The sub-optimal controlled with $u(k)$ determined by

$$u(k) = G_1 [z_1(k) - x_1(k)] + G_2 [z_2(k) - x_2(k)] \quad (1.63)$$

with $G_1 = 8.50$ and $G_2 = 4.75$, when operated with the same nonlinear system gave a performance index of 31. 32. Chapter V contains the results of several additional computer examples.

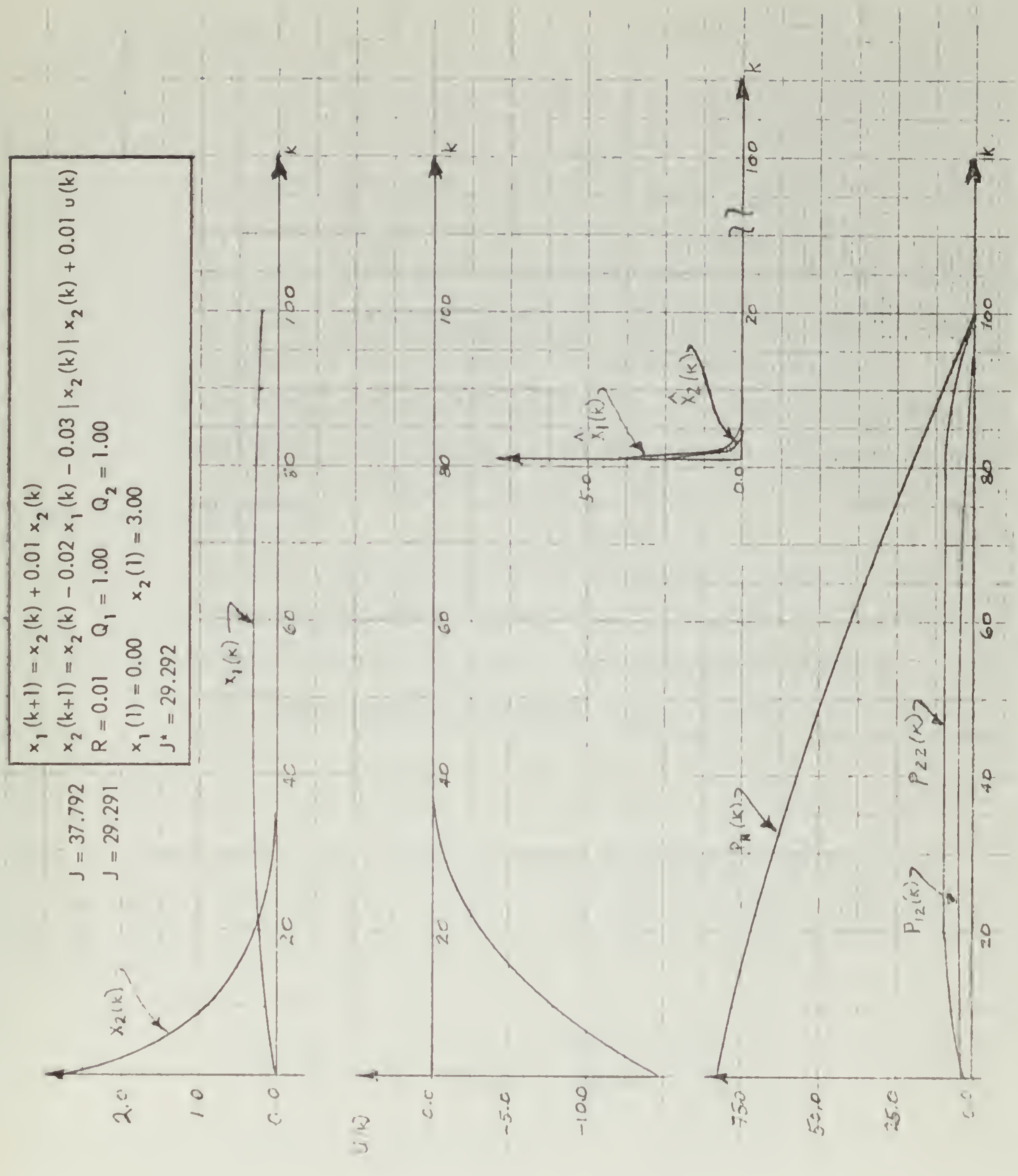

J = 37.792
J = 29.291
x1 (k+1) = x2 (k) + 0.01 x2 (k)
x2 (k+1) = x2 (k) − 0.02 x1 (k) − 0.03 | x2 (k) | x2 (k) + 0.01 u (k)
R = 0.01 Q1 = 1.00 Q2 = 1.00
x1 (1) = 0.00 x2 (1) = 3.00
J* = 29.292
x2(k)
x1(k)
u(k)
P12(k)
P22(k)
k

Figure 1.4

CHAPTER II

DISCRETE TIME SYSTEMS

2.1 Introduction

The theory for the control of discrete time systems can be developed more simply than that of continuous time systems. In particular, the discrete time theory avoids some questions about the existence of limits, etc. For this reason, the discrete time theory is presented first.

This chapter first considers linear discrete time systems thoroughly. Then, using the linear results as a guide, the theory is extended to include nonlinear systems. The exact solution of the nonlinear problem is presented in the form of an iterative algorithm. The final section of the chapter considers the problem when stochastic disturbances are present.

2.2 Linear Systems

The theory for the optimal control of deterministic linear systems has been worked out by Kalman,[4-7] Merriam,[15,16] and others.[17-19] For this case, the system considered can be described by the equations

$$\underline{x}(k+1) = \underline{F}(k)\underline{x}(k) + \underline{G}(k)\underline{u}(k); \qquad \underline{x}(0) = \underline{c} \tag{2.1}$$

$$\underline{y}(k) = \underline{H}(k)\underline{x}(k) \tag{2.2}$$

where $\underline{x}(k)$ is the n-dimensional system state vector, $\underline{u}(k)$ is the r-dimensional system control vector, and $\underline{y}(k)$ is the m-dimensional system output vector.

The performance index is

$$J(k) = \sum_{j=k}^{N} \frac{1}{2} \|\underline{z}(j) - \underline{y}(j)\|^2_{\underline{Q}(j)} + \sum_{j=k}^{N-1} \frac{1}{2} \|\underline{u}(j)\|^2_{\underline{R}(j)} \tag{2.3}$$

where $\underline{z}(j)$ is the desired output vector.

To find the optimal control sequence, $\underline{u}(k)$, $\underline{u}(k+1)$, . . . , $\underline{u}(N-1)$, the method of dynamic programming is used. For this purpose, we define the value function,

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k), \ldots, \underline{u}(N-1)}{\text{Min}} \{J(k)\} \tag{2.4}$$

We then invoke the "principle of optimality," which states:

> "an optimal policy has the property that, whatever the initial state and the initial decision are, the remaining decisions must constitute an optimal policy with regard to the state resulting from the first decision."[3]

Thus, it follows that

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k)}{\text{Min}} \left\{ \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{u}(k) \|^2_{\underline{R}(k)} + V_{N-k-1}(\underline{x}(k+1)) \right\} \tag{2.5}$$

A solution for V_{N-k} $(\underline{x}(k))$ and $\underline{u}(k)$, $(k = 0, 1, \ldots, N-1)$, can be obtained by assuming

$$V_{N-k}(\underline{x}(k)) = \frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k)\hat{\underline{x}}(k) + a(k) \tag{2.6}$$

where $\underline{P}(k)$, $\hat{\underline{x}}(k)$, and $a(k)$ are a parameter matrix, vector, and scalar, respectively, to be determined. By combining equations (2.5) and (2.6), we get

$$\begin{aligned} \frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k)\hat{\underline{x}}(k) + a(k) = \underset{\underline{u}(k)}{\text{Min}} & \left\{ \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{u}(k) \|^2_{\underline{R}(k)} \right. \\ & \left. + \frac{1}{2} \| \underline{x}(k+1) \|^2_{\underline{P}(k+1)} + \underline{x}'(k+1)\hat{\underline{x}}(k+1) + a(k+1) \right\} \end{aligned} \tag{2.7}$$

The vector variable $\underline{x}(k+1)$ can be eliminated from this equation by using equation (2.1). This gives

$$\begin{aligned} \frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k)\underline{x}(k) + a(k) = \underset{\underline{u}(k)}{\text{Min}} & \left\{ \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} + a(k+1) \right. \\ & \left. + \frac{1}{2} \| \underline{F}(k)\underline{x}(k) + \underline{G}(k)\underline{u}(k) \|^2_{\underline{P}(k+1)} + [\underline{F}(k)\underline{x}(k) + \underline{G}(k)\underline{u}(k)]' \, \hat{\underline{x}}(k+1) \right\} \end{aligned} \tag{2.8}$$

The minimizing value of $\underline{u}(k)$ for the expression on the right-hand side of equation (2.8) can be determined by ordinary methods of calculus.

This value is

$$\underline{u}_{Min}(k) = -[\underline{R}(k) + \underline{G}'(k)\underline{P}(k+1)\underline{G}(k)]^{-1}\,\underline{G}'(k)\,[\underline{P}(k+1)\,\underline{F}(k)\underline{x}(k) + \hat{\underline{x}}(k+1)] \qquad (2.9)$$

By substituting the expression for the minimizing value of $\underline{u}(k)$ into equation (2.8), we get

$$\frac{1}{2}\,||\underline{x}(k)||^2_{\underline{P}(k)} + \underline{x}'(k)\hat{\underline{x}}(k) + a(k) = \frac{1}{2}\,||\underline{z}(k) - \underline{H}(k)\underline{x}(k)||^2_{\underline{Q}(k)}$$

$$-\frac{1}{2}\,||\underline{P}(k+1)\,\underline{F}(k)\underline{x}(k) + \hat{\underline{x}}(k+1)||^2_{\underline{G}(k)[\underline{R}(k)+\underline{G}'(k)\underline{P}(k+1)\underline{G}(k)]^{-1}\underline{G}'(k)} \qquad (2.10)$$

$$+\frac{1}{2}\,||\underline{F}(k)\underline{x}(k)||^2_{\underline{P}(k+1)} + \underline{x}'(k)\underline{F}'(k)\hat{\underline{x}}(k+1) + a(k+1)$$

This equation will be satisfied for all $\underline{x}(k)$ if and only if the following recursion equations are satisfied.

$$\underline{P}(k) = \underline{H}'(k)\underline{Q}(k)\underline{H}(k) + \underline{F}'(k)\underline{M}(k)\underline{P}(k+1)\underline{F}(k) \qquad (2.11)$$

$$\hat{\underline{x}}(k) = \underline{F}'(k)\underline{M}(k)\hat{\underline{x}}(k+1) - \underline{H}'(k)\underline{Q}(k)\underline{z}(k) \qquad (2.12)$$

$$a(k) = a(k+1) - \frac{1}{2}\,||\underline{z}(k)||^2_{\underline{Q}(k)} - \frac{1}{2}\,||\hat{\underline{x}}(k+1)||^2_{\underline{G}(k)[\underline{R}(k)+\underline{G}'(k)\underline{P}(k+1)\underline{G}(k)]^{-1}\underline{G}'(k)} \qquad (2.13)$$

where

$$\underline{M}(k) = \underline{I} - \underline{P}(k+1)\underline{G}(k)[\underline{R}(k) + \underline{G}'(k)\underline{P}(k+1)\underline{G}(k)]^{-1}\,\underline{G}'(k) \qquad (2.14)$$

The boundary conditions for this set of equations can be determined from equations (2.3), (2.4), and (2.6) evaluated at $k = N$. Thus

$$\underline{P}(N+1) = \underline{0} \qquad (2.15)$$

$$\hat{\underline{x}}(N+1) = \underline{0} \qquad (2.16)$$

$$a(N+1) = 0 \qquad (2.17)$$

form the appropriate boundary conditions.

Notice that equations (2.11), (2.12), and (2.13) must be solved backwards in time. For this reason, the system must be "deterministic" in the sense that $\underline{z}(j)$ must be known on the entire interval, $j = k, k+1, \ldots, N$, in order to compute the optimal control vector at time $j = k$. Also notice that $a(k)$ is required to determine $V_{N-k}(\underline{x}(k))$, but is not required to determine $\underline{u}(k)$. Thus, if we want to synthesize

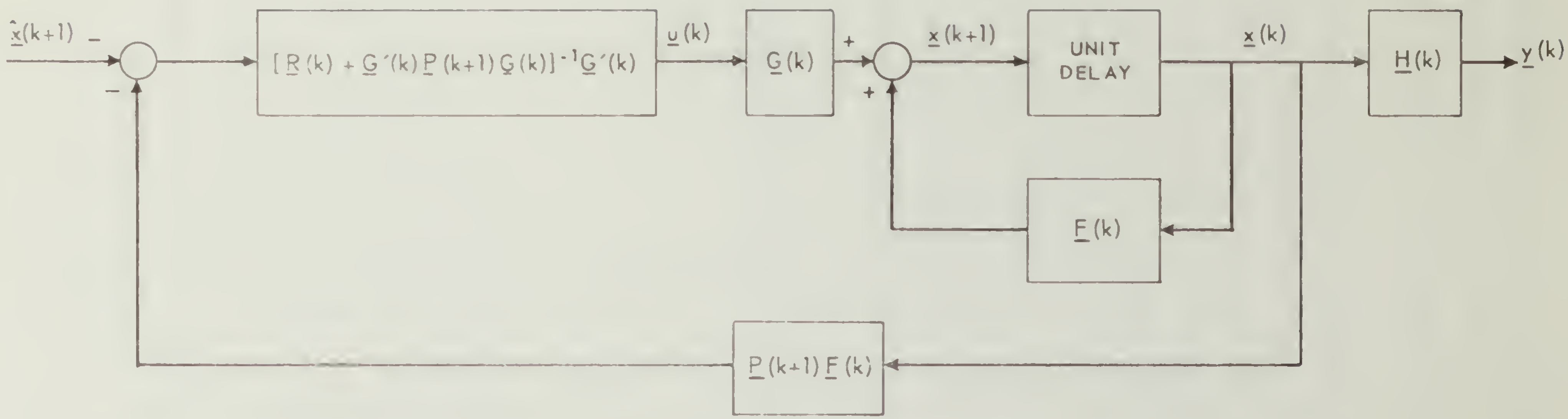


Figure 2.1 - Optimal Linear Control System

the optimal control system and are not interested in computing the minimum value of the performance index, we need not compute the $a(k)$ sequence. A block diagram of the optimal linear control system is shown in figure 2. 1.

As can be seen from the block diagram, the controller consists of a time varying linear feedback portion and a feed-forward or director portion. The feedback signal is simply the system state vector amplified by the time varying gain matrix $\underline{P}(k+1)\underline{F}(k)$. The feed-forward signal, $\hat{\underline{x}}(k)$, may be interpreted as a modified desired output. In other words, the closed loop portion of the system tries to follow $\hat{\underline{x}}(k)$ instead of $\underline{z}(k)$ because it is more economical.

From equation (2. 12), it can be seen that $\hat{\underline{x}}(k)$ is derived from $\underline{z}(k)$ by the feedback system shown in figure 2. 2. As has been stated previously, this system operates backward in time.

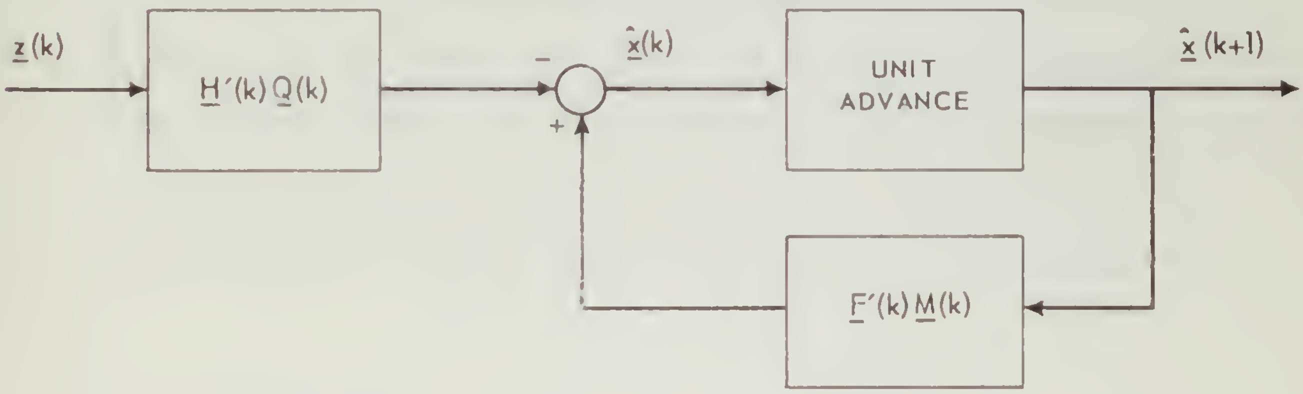


Figure 2. 2 - System for $\hat{\underline{x}}(k)$

If the output of the system shown above follows the input reasonably well, using $-\underline{H}'(k)\,\underline{Q}(k)\,\underline{z}(k)$ in place of $\hat{\underline{x}}(k)$ for the feed-forward input to the control system of figure 2. 1 should give nearly optimal performance. This would eliminate the objectionable requirement of having to know $\underline{z}(j)$ over the entire interval in advance.

The computational procedure for determining the optimal control is evident from the nature of the equations. The matrices $\underline{P}(N)$, $\underline{P}(N-1)$, . . . , $\underline{P}(0)$, and the vectors $\hat{\underline{x}}(N)$, $\hat{\underline{x}}(N-1)$, . . . , $\hat{\underline{x}}(0)$ must be pre-computed by backwards recursion of equations (2. 11) and (2. 12). These quantities

would then be used along with $\underline{x}(k)$ to determine $\underline{u}(k)$ as the actual control system evolves forward in time.

Another consideration concerning the optimal control system is that of the measurement of state variables. For the preceeding development, we have tacitly assumed that the state variables are exactly measurable. This frequently is not a reasonable assumption. For the linear problem, Gunckel[20,21] has shown that the optimal control system for the case when the state variables are not exactly measurable consists of the control system derived above with an optimum filter inserted in the control loop to estimate the state variables. When the state variables are not exactly measurable in the case of nonlinear control systems, we have no assurance that an optimal filter to estimate the state variables inserted in the control system will result in optimal performance. In this case, however, as Cox[22] has pointed out, if the state variables are not exactly measurable, we have no alternative to determining the optimal control system by assuming the state variables are exactly measurable and then inserting an optimal filter in the control loop. In all that follows, we will assume that the state variables are exactly measurable. Cox[22] has treated the problem of estimating state variables in noisy nonlinear systems.

2.3 Nonlinear Systems

The theory for the optimal control of deterministic linear systems is extended to a fairly general class of nonlinear systems in this section. Actually the solution derived in this section is only approximately optimal. Section 2.4 presents an iterative procedure based on this approximate solution that leads to the exact solution.

For the nonlinear case, the system considered can be described by the state equations

$$\underline{x}(k+1) = \underline{f}(\underline{x}(k), \underline{u}(k), k); \qquad \underline{x}(0) = \underline{c} \tag{2.18}$$

$$\underline{y}(k+1) = \underline{h}(\underline{x}(k), k) \tag{2.19}$$

The performance index is

$$J(k) = \sum_{j=k}^{N} \frac{1}{2} \| \underline{z}(j) - \underline{y}(j) \|^2_{\underline{Q}(j)} + \sum_{j=k}^{N-1} \frac{1}{2} \| \underline{u}(j) \|^2_{\underline{R}(j)} \qquad (2.20)$$

We follow the procedure of the previous section and define

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k), \ldots, \underline{u}(N-1)}{\text{Min}} \{ J(k) \} \qquad (2.21)$$

By the principle of optimality, it follows that

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k)}{\text{Min}} \left\{ \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{u}(k) \|^2_{\underline{R}(k)} + V_{N-k-1}(\underline{x}(k+1)) \right\} \qquad (2.22)$$

We cannot solve this equation by direct methods; so we resort to linearization.

The approximations are

$$\underline{x}(k+1) \simeq \underline{f}(\underline{x}^*(k), \underline{u}^*(k), k) + f_{\underline{x}}(\underline{x}^*(k), \underline{u}^*(k), k)\, [\underline{x}(k) - \underline{x}^*(k)] + f_{\underline{u}}(\underline{x}^*(k), \underline{u}^*(k), k) [\underline{u}(k) - \underline{u}^*(k)] \qquad (2.23)$$

and

$$\underline{y}(k+1) \simeq \underline{h}(\underline{x}^*(k), k) + h_{\underline{x}}(\underline{x}^*(k), k)\, [\underline{x}(k) - \underline{x}^*(k)] \qquad (2.24)$$

As before, we assume

$$V_{N-k}(\underline{x}(k)) = \frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k) \hat{\underline{x}}(k) + a(k) \qquad (2.25)$$

By combining equations (2. 22), (2. 23), (2. 24), and (2. 25) we obtain the single equation

$$\frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k) \hat{\underline{x}}(k) + a(k) = \underset{\underline{u}(k)}{\text{Min}} \left\{ \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} \right.$$

$$+ \frac{1}{2} \| \underline{u}(k) \|^2_{\underline{R}(k)} + \frac{1}{2} \| \underline{f} + f_{\underline{x}} [\underline{x}(k) - \underline{x}^*(k)] + f_{\underline{u}} [\underline{u}(k) - \underline{u}^*(k)] \|^2_{\underline{P}(k+1)} \qquad (2.26)$$

$$\left. + [\underline{f} + f_{\underline{x}}(\underline{x}(k) - \underline{x}^*(k)) + f_{\underline{u}}(\underline{u}(k) - \underline{u}^*(k))]' \hat{\underline{x}}(k+1) + a(k+1) \right\}$$

(When the arguments of $\underline{f}$, $f_{\underline{x}}$, and $f_{\underline{u}}$ are omitted, they are understood to be evaluated at $\underline{x}^*(k)$, $\underline{u}^*(k)$, and k. Similarly, when the arguments of $\underline{h}$ and $h_{\underline{x}}$ are omitted, they are understood to be evaluated at $\underline{x}^*(k)$ and k.)

The minimizing value of $\underline{u}(k)$ can be computed by the ordinary methods of calculus, and is given by

$$\underline{u}_{Min}(k) = -[\underline{R}(k) + f_{\underline{u}}' \underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}' [\underline{P}(k+1) f_{\underline{x}} \underline{x}(k) + \underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1)] \tag{2.27}$$

where

$$\underline{b}(k) = \underline{f} - f_{\underline{x}} \underline{x}^*(k) - f_{\underline{u}} \underline{u}^*(k) \tag{2.28}$$

When the minimum value of $\underline{u}(k)$ from equation (2.27) is substituted into equation (2.26), it becomes

$$\begin{aligned}\frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k)\hat{\underline{x}}(k) + a(k) &= \frac{1}{2} \| \underline{z}(k) - h_{\underline{x}}(k) - \underline{c}(k) \|^2_{\underline{Q}(k)} \\ &- \frac{1}{2} \| \underline{P}(k+1) f_{\underline{x}} \underline{x}(k) + \underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1) \|^2_{f_{\underline{u}}[\underline{R}(k) + f_{\underline{u}}'\underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}'} \\ &+ \frac{1}{2} \| f_{\underline{x}} \underline{x}(k) + \underline{b}(k) \|^2_{\underline{P}(k+1)} + [f_{\underline{x}} \underline{x}(k) + \underline{b}(k)]' \hat{\underline{x}}(k+1) + a(k+1)\end{aligned} \tag{2.29}$$

where

$$\underline{c}(k) = \underline{h} - h_{\underline{x}} \underline{x}^*(k) \tag{2.30}$$

This equation will be satisfied for all $\underline{x}(k)$ if and only if the following set of equations are satisfied:

$$\underline{P}(k) = h_{\underline{x}}' \underline{Q}(k) h_{\underline{x}} + f_{\underline{x}}' \underline{M}(k) \underline{P}(k+1) f_{\underline{x}} \tag{2.31}$$

$$\hat{\underline{x}}(k) = f_{\underline{x}}' \underline{M}(k) [\underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1)] - h_{\underline{x}}' \underline{Q}(k) [\underline{z}(k) - \underline{c}(k)] \tag{2.32}$$

$$\begin{aligned}a(k) = a(k+1) &+ \frac{1}{2} \| \underline{z}(k) - \underline{c}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{b}(k) \|^2_{\underline{P}(k+1)} + \underline{b}'(k)\hat{\underline{x}}(k+1) \\ &- \frac{1}{2} \| \underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1) \|^2_{f_{\underline{u}}[\underline{R}(k) + f_{\underline{u}}'\underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}'}\end{aligned} \tag{2.33}$$

where

$$\underline{M}(k) = \underline{I} - \underline{P}(k+1) f_{\underline{u}} [\underline{R}(k) + f_{\underline{u}}' \underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}' \tag{2.34}$$

The boundary values for this set of equations can be obtained in the same manner as for the linear problem. The boundary values are

$$\underline{P}(N+1) = \underline{0} \tag{2.35}$$

$$\hat{\underline{x}}(N+1) = \underline{0} \tag{2.36}$$

$$a(N+1) = 0 \tag{2.37}$$

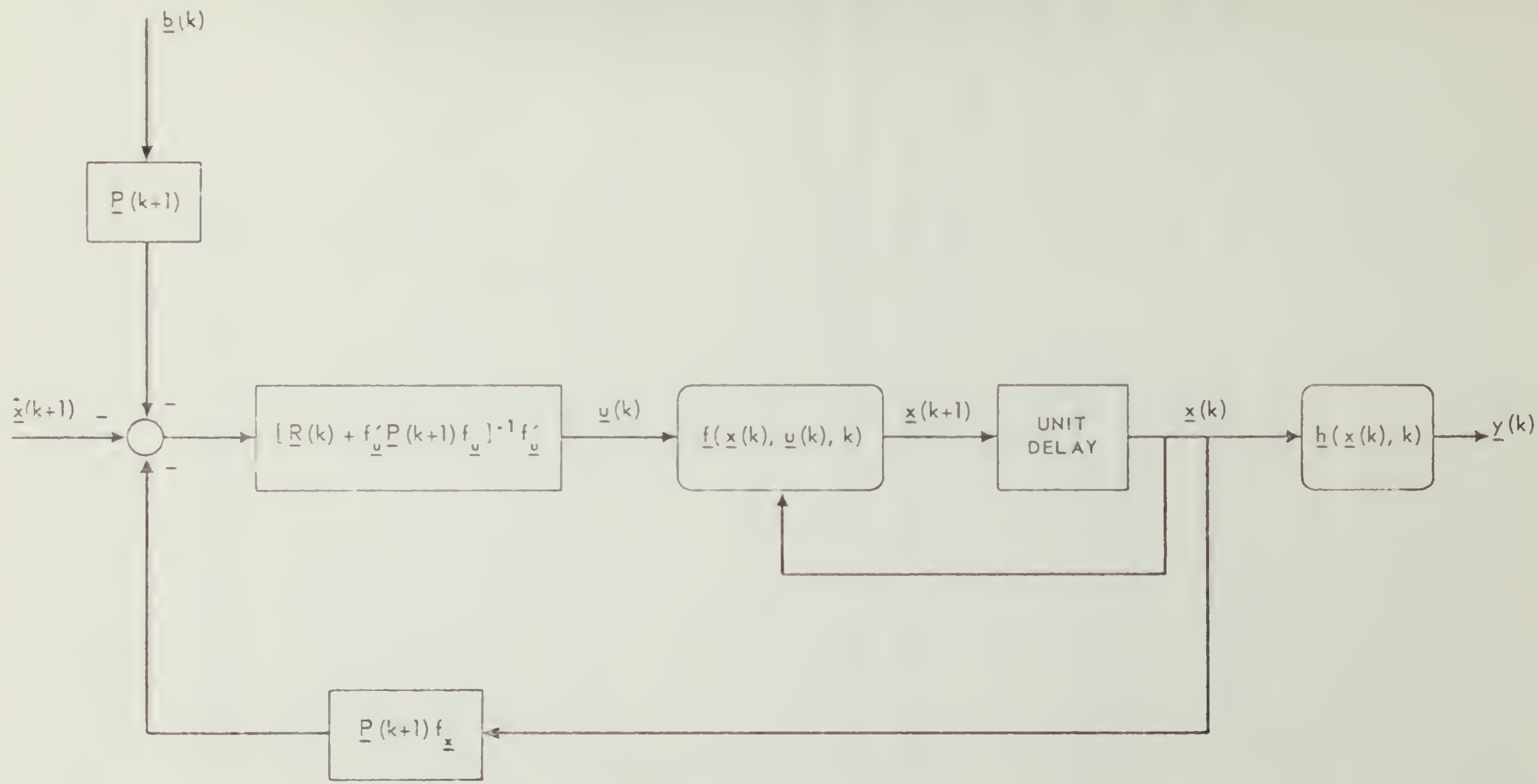


Figure 2.3 - Nonlinear Control System

Once the sequence of points, $\underline{x}^*(k)$ and $\underline{u}^*(k)$, are given, the sequences $\underline{P}(k)$, $\hat{\underline{x}}(k)$, and $a(k)$ may be computed by backward recursion of equations (2.31), (2.32), and (2.33). After these quantities have been pre-computed, the system may be operated forward in time under the approximately optimal control given by equation (2.27). The problem is, of course, to determine a sequence, $\underline{x}^*(k)$ and $\underline{u}^*(k)$, about which to linearize such that the approximation is a good one. This is the subject of the next section.

Figure 2.3 shows a block diagram of the nonlinear control system. Notice that although the system being controlled is nonlinear, the controller is time varying linear.

2.4 Solution by Iteration

The development of the theory in this section requires us to attack the optimal nonlinear control problem from a different point of view. Consider again the system

$$\underline{x}(k+1) = \underline{f}(\underline{x}(k), \underline{u}(k), k); \qquad \underline{x}(0) = \underline{c} \tag{2.38}$$

$$\underline{y}(k) = \underline{h}(\underline{x}(k), k) \tag{2.39}$$

subject to the performance criterion

$$J = \sum_{k=0}^{N} \frac{1}{2} \|\underline{z}(k) - \underline{y}(k)\|^2_{\underline{Q}(k)} + \sum_{k=0}^{N-1} \frac{1}{2} \|\underline{u}(k)\|^2_{\underline{R}(k)} \tag{2.40}$$

We wish to choose $\underline{u}(k)$ such that the performance criterion is a minimum.

The minimization can be performed by calculus techniques using Lagrange multipliers.† For this purpose we define the function

$$\begin{aligned} I = {} & \sum_{k=0}^{N} \frac{1}{2} \|\underline{z}(k) - \underline{h}(\underline{x}(k), k)\|^2_{\underline{Q}(k)} + \sum_{k=0}^{N-1} \frac{1}{2} \|\underline{u}(k)\|^2_{\underline{R}(k)} \\ & + \sum_{k=0}^{N-1} \underline{\lambda}'(k) [\underline{x}(k+1) - \underline{f}(\underline{x}(k), \underline{u}(k), k)] + \underline{\lambda}(-1)[\underline{x}(0) - \underline{c}] \end{aligned} \tag{2.41}$$

† This approach is similar to that used by Kipiniak[23] and the entire development of this section, including the iterative procedure, is closely related to the nonlinear smoothing problem treated by Cox.[22]

By equating partial derivatives of I with respect to $\underline{u}(k)$, $\underline{\lambda}(k)$, and $\underline{x}(k)$ to zero, we obtain the following set of equations which define the optimum control system.

$$\underline{u}(k) = \underline{R}^{-1}(k) f_{\underline{u}}'(\underline{x}(k), \underline{u}(k), k)\, \underline{\lambda}(k) \tag{2.42}$$

$$\underline{x}(k+1) = \underline{f}(\underline{x}(k), \underline{u}(k), k) \tag{2.43}$$

$$\underline{\lambda}(k-1) = f_{\underline{x}}'(\underline{x}(k), \underline{u}(k), k)\, \underline{\lambda}(k) + h_{\underline{x}}' \underline{Q}(k) [\underline{z}(k) - \underline{h}(\underline{x}(k), k)] \tag{2.44}$$

The boundary conditions are

$$\underline{x}(0) = \underline{c} \tag{2.45}$$

and

$$\underline{\lambda}(N) = \underline{0} \tag{2.46}$$

This set of equations is nonlinear, and an analytic solution is not known. However, we can obtain an approximate solution by using the linearizations

$$\begin{aligned} \underline{x}(k+1) \simeq\; & \underline{f}(\underline{x}^*(k), \underline{u}^*(k), k) + f_{\underline{x}}(\underline{x}^*(k), \underline{u}^*(k), k) [\underline{x}(k) - \underline{x}^*(k)] \\ & + f_{\underline{u}}(\underline{x}^*(k), \underline{u}^*(k), k) [\underline{u}(k) - \underline{u}^*(k)] \end{aligned} \tag{2.47}$$

and

$$\underline{y}(k) \simeq \underline{h}(\underline{x}^*(k), k) + h_{\underline{x}}(\underline{x}^*(k), k) [\underline{x}(k) - \underline{x}^*(k)] \tag{2.48}$$

When we use these approximations instead of equations (2.38) and (2.39), the equations for $\underline{u}(k)$, $\underline{x}(k+1)$, and $\underline{\lambda}(k-1)$ become

$$\underline{u}(k) = \underline{R}^{-1}(k) f_{\underline{u}}' \underline{\lambda}(k) \tag{2.49}$$

$$\underline{x}(k+1) = \underline{f} + f_{\underline{x}} [\underline{x}(k) - \underline{x}^*(k)] + f_{\underline{u}} [\underline{u}(k) - \underline{u}^*(k)] \tag{2.50}$$

$$\underline{\lambda}(k-1) = f_{\underline{x}}' \underline{\lambda}(k) + h_{\underline{x}}' \underline{Q}(k) \{\underline{z}(k) - \underline{h} - h_{\underline{x}} [\underline{x}(k) - \underline{x}^*(k)]\} \tag{2.51}$$

where $\underline{f}$, $f_{\underline{x}}$, and $f_{\underline{u}}$ are understood to be evaluated at $\underline{x}^*(k)$, $\underline{u}^*(k)$, and k, and $\underline{h}$ and $h_{\underline{x}}$ are understood to be evaluated at $\underline{x}^*(k)$ and k.

We can solve the above set of equations by assuming

$$\underline{\lambda}(k-1) = -\underline{P}(k)\underline{x}(k) - \hat{\underline{x}}(k) \tag{2.52}$$

The solution proceeds by combining equations (2.49), (2.50), (2.51), and (2.52) to eliminate $\underline{u}(k)$, $\underline{\lambda}(k)$, and $\underline{\lambda}(k-1)$, obtaining

$$\underline{x}(k+1) = \underline{f} + f_{\underline{x}}[\underline{x}(k) - \underline{x}^*(k)] - f_{\underline{u}}\underline{u}^*(k) + f_{\underline{u}}\underline{R}^{-1}(k)f_{\underline{u}}'[\underline{P}(k+1)\underline{x}(k+1) + \hat{\underline{x}}(k+1)] \tag{2.53}$$

and

$$\underline{P}(k)\underline{x}(k) + \hat{\underline{x}}(k) = f_{\underline{x}}'[\underline{P}(k+1)\underline{x}(k+1) + \hat{\underline{x}}(k+1)] - h_{\underline{x}}'\underline{Q}(k)\{\underline{z}(k) - \underline{h} - h_{\underline{x}}[\underline{x}(k) - \underline{x}^*(k)]\} \tag{2.54}$$

These two equations can be combined to eliminate $\underline{x}(k+1)$, giving

$$\begin{aligned}\underline{P}(k)\underline{x}(k) + \hat{\underline{x}}(k) &= f_{\underline{x}}'\hat{\underline{x}}(k+1) - h_{\underline{x}}'\underline{Q}(k)\{\underline{z}(k) - \underline{h} - h_{\underline{x}}[\underline{x}(k) - \underline{x}^*(k)]\} \\ &+ f_{\underline{x}}'\underline{P}(k+1)[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k)f_{\underline{u}}'\underline{P}(k+1)]^{-1}\{\underline{f} + f_{\underline{x}}[\underline{x}(k) - \underline{x}^*(k)] - f_{\underline{u}}\underline{u}^*(k) - f_{\underline{u}}\underline{R}^{-1}(k)f_{\underline{u}}'\hat{\underline{x}}(k+1)\}\end{aligned} \tag{2.55}$$

provided the inverse indicated exists. (Section 2.5 contains a proof that the inverse required above does indeed exist.)

Equation (2.55) will be satisfied for all $\underline{x}(k)$ if and only if the following set of equations are satisfied.

$$\underline{P}(k) = f_{\underline{x}}'\underline{P}(k+1)[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k)f_{\underline{u}}'\underline{P}(k+1)]^{-1}f_{\underline{x}} + h_{\underline{x}}'\underline{Q}(k)h_{\underline{x}} \tag{2.56}$$

$$\hat{\underline{x}}(k) = f_{\underline{x}}'\hat{\underline{x}}(k+1) - h_{\underline{x}}'\underline{Q}(k)[\underline{z}(k) - \underline{c}(k)] - f_{\underline{x}}'\underline{P}(k+1)[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k)f_{\underline{u}}'\underline{P}(k+1)]^{-1}[f_{\underline{u}}\underline{R}^{-1}(k)f_{\underline{u}}'\hat{\underline{x}}(k+1) - \underline{b}(k)] \tag{2.57}$$

We are now in a position to obtain an exact solution to equations (2.42), (2.43), and (2.44), and hence an exact solution to the nonlinear control problem. The exact solution is obtained by solving equations (2.49), (2.50), (2.52), (2.56), and (2.57) iteratively.

First, we denote the state sequence and the control sequence obtained on the *ith* iteration as $\underline{x}_i(0), \ldots, \underline{x}_i(N)$ and $\underline{u}_i(0), \ldots, \underline{u}_i(N-1)$, respectively. Then, for the $i+1$*st* iteration we linearize about the points $\underline{x}_i(k)$ and $\underline{u}_i(k)$. The procedure for the $i+1$*st* iteration is as follows:

Step 1. Solve equations (2.56) and (2.57) backward in time using

$$\underline{x}^*(k) = \underline{x}_i(k) \tag{2.58}$$

and

$$\underline{u}^*(k) = \underline{u}_i(k) \tag{2.59}$$

to compute $\underline{P}(N), \ldots, \underline{P}(0)$ and $\hat{\underline{x}}(N), \ldots, \hat{\underline{x}}(0)$.

Step 2. Solve equations (2.49), (2.50), and (2.52) forward in time using

$$\underline{x}^*(k) = \underline{x}_i(k) \tag{2.60}$$

and

$$\underline{u}^*(k) = \underline{u}_i(k) \tag{2.61}$$

as before, to compute $\underline{u}_{i+1}(0), \ldots, \underline{u}_{i+1}(N-1)$ and $\underline{x}_{i+1}(0), \ldots, \underline{x}_{i+1}(N)$.

Steps 1 and 2 are repeated until convergence is achieved, i.e., until the norms of the quantities $[\underline{x}_{i+1}(k) - \underline{x}_i(k)]$ and $[\underline{u}_{i+1}(k) - \underline{u}_i(k)]$ are less than some previously specified convergence criteria. It can be seen by comparing equations (2.49), (2.50), and (2.51) with equations (2.42), (2.43), and (2.44) that if convergence is achieved using the iterative procedure, that is if

$$\underline{x}_{i+1}(k) = \underline{x}_i(k) \tag{2.62}$$

and

$$\underline{u}_{i+1}(k) = \underline{u}_i(k) \tag{2.63}$$

then the solution obtained is the exact solution for equations (2.42), (2.43), and (2.44) as well. In other words, the solution obtained by convergence of the iterative procedure is the exact solution to the optimal nonlinear control problem. The question of under what conditions convergence can be assured is a difficult one, and as yet has not been answered by the author. This remains a challenging area for possible future research. However, computer studies using this iteration procedure indicate that convergence usually occurs in a few iterations. Chapter V contains some of these results.

Because the inverse in equations (2.56) and (2.57) is generally more difficult to compute than the inverse occurring in the solution of the last section, we would prefer to use equations (2.31) and (2.32) as the basis for the iterative algorithm in lieu of equations (2.56) and (2.57). However, nothing we have shown thus far would permit us to do this and still guarantee that a convergent solution for the iterative algorithm is the exact solution to the nonlinear control problem.

We can show that the iteration scheme based on equations (2.31) and (2.32) does lead to the exact solution, and in fact is identical to the scheme based on equations (2.56) and (2.57) by using the following matrix identities.

$$[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k)f'_{\underline{u}}\underline{P}(k+1)]^{-1} \equiv \underline{I} - f_{\underline{u}}[\underline{R}(k) + f'_{\underline{u}}\underline{P}(k+1)f_{\underline{u}}]^{-1} f'_{\underline{u}}\underline{P}(k+1) \qquad (2.64)$$

and

$$[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k)f'_{\underline{u}}\underline{P}(k+1)]^{-1} f_{\underline{u}}\underline{R}^{-1}(k)f'_{\underline{u}} \equiv f_{\underline{u}}[\underline{R}(k) + f'_{\underline{u}}\underline{P}(k+1)f_{\underline{u}}]^{-1} f'_{\underline{u}} \qquad (2.65)$$

(Appendix A contains a proof of these identities.) The application of identities (2.64) and (2.65) to equations (2.56) and (2.57) immediately transforms them into equations (2.31) and (2.32). In addition, since by equation (2.49)

$$\underline{u}(k) = \underline{R}^{-1}(k)f'_{\underline{u}}\underline{\lambda}(k) \qquad (2.66)$$

or, using (2.52),

$$\underline{u}(k) = -\underline{R}^{-1}(k)f'_{\underline{u}}[\underline{P}(k+1)\underline{x}(k+1) + \hat{\underline{x}}(k+1)] \qquad (2.67)$$

and by (2.50)

$$\underline{u}(k) = -\underline{R}^{-1}(k)f'_{\underline{u}}\{\underline{P}(k+1)[\underline{f} + f_{\underline{x}}(\underline{x}(k) - \underline{x}^*(k)) + f_{\underline{u}}(\underline{u}(k) - \underline{u}^*(k))] + \hat{\underline{x}}(k+1)\} \qquad (2.68)$$

Solving this equation for $\underline{u}(k)$ explicitly yields

$$\underline{u}(k) = -[\underline{R}(k) + f'_{\underline{u}}\underline{P}(k+1)f_{\underline{u}}]^{-1} f'_{\underline{u}}\{\underline{P}(k+1)[\underline{f} + f_{\underline{x}}(\underline{x}(k) - \underline{x}^*(k)) - f_{\underline{u}}\underline{u}^*(k)] + \hat{\underline{x}}(k+1)\} \qquad (2.69)$$

which is identical to equation (2.27). Thus we have shown the solution based on the equations derived in this section is identical to the solution based on the equations of the previous section.

2.5 On $\underline{P}(k)$ and $[I + f_{\underline{u}}\underline{R}^{-1}(k)f'_{\underline{u}}\underline{P}(k+1)]^{-1}$

This section contains two theorems of importance to the material in this chapter. The first theorem concerns the existence of $[I + f_{\underline{u}}\underline{R}^{-1}(k)f'_{\underline{u}}\underline{P}(k+1)]^{-1}$, and the second theorem concerns the non-negative definiteness of $\underline{P}(k)$. The proof of these theorems will require some elementary results from matrix theory. These are

a. If the $n \times n$ matrix $\underline{P}$ is non-negative definite, then the matrix $\underline{G}'\underline{PG}$ is non-negative definite, where $\underline{G}$ is any $n \times r$ matrix.

b. The inverse of a positive definite matrix exists and is positive definite.

c. The sum of a positive definite matrix and a non-negative definite matrix is positive definite.

d. The sum of two non-negative definite matrices is non-negative definite.

Theorem 1: If $\underline{R}(k)$ is positive definite, and $\underline{P}(k+1)$ is non-negative definite, then the inverse $[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}'\underline{P}(k+1)]^{-1}$ exists.

Proof: Consider the matrix expression

$$\underline{I} - f_{\underline{u}} [\underline{R}(k) + f_{\underline{u}}' \underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}'\underline{P}(k+1) \qquad (2.70)$$

If $\underline{P}(k+1)$ is non-negative definite, then by a., $f_{\underline{u}}'\underline{P}(k+1) f_{\underline{u}}$ is non-negative definite. If $\underline{R}(k)$ is positive definite, then by c., $\underline{R}(k) + f_{\underline{u}}'\underline{P}(k+1) f_{\underline{u}}$ is positive definite, and hence by b., $[\underline{R}(k) + f_{\underline{u}}'\underline{P}(k+1) f_{\underline{u}}]^{-1}$ exists. Thus the whole expression exists. But, by the first identify of section 2.4, $[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}'\underline{P}(k+1)]^{-1}$ is identical to the expression above and hence must exist.

Theorem 2: If $\underline{R}(k)$ is positive definite, and if $\underline{Q}(k)$ and $\underline{P}(k+1)$ are non-negative definite, then $\underline{P}(k)$ is non-negative definite.

Proof: Consider equation (2.56), rewritten here.

$$\underline{P}(k) = f_{\underline{x}}'\underline{P}(k+1) [\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}'\underline{P}(k+1)]^{-1} f_{\underline{x}} + h_{\underline{x}}'\underline{Q}(k) h_{\underline{x}} \qquad (2.71)$$

If $\underline{Q}(k)$ is non-negative definite, then by a., $h_{\underline{x}}'\underline{Q}(k) h_{\underline{x}}$ is non-negative definite. As for the first term on the right of (2.56), it must be non-negative definite also if $\underline{P}(k+1)$ is non-negative definite. To show that this is so, let

$$[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}'\underline{P}(k+1)]^{-1} f_{\underline{x}} = \underline{A} \qquad (2.72)$$

then

$$f_{\underline{x}} = [\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}' \underline{P}(k+1)] \underline{A} \qquad (2.73)$$

Thus we see that

$$\underline{f}_{\underline{x}}' \underline{P}(k+1)[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}' \underline{P}(k+1)]^{-1} f_{\underline{x}} = \underline{A}'[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}' \underline{P}(k+1)]' \underline{P}(k+1)\underline{A} \qquad (2.74)$$

or

$$f_{\underline{x}}' \underline{P}(k+1)[\underline{I} + f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}' \underline{P}(k+1)]^{-1} f_{\underline{x}} = \underline{A}'\underline{P}(k+1)\underline{A} + \underline{A}'\underline{P}(k+1) f_{\underline{u}}\underline{R}^{-1}(k) f_{\underline{u}}' \underline{P}(k+1)\underline{A} \qquad (2.75)$$

But by a., and d., the right-hand side of equation (2. 75) is non-negative definite. Hence for the same reason, the right-hand side of equation (2. 56) is non-negative definite, completing the proof.

The hypotheses of theorem 2 are satisfied by the original assumptions of the problem statement. The hypotheses of theorem 1 are satisfied by the original assumptions in the problem statement, and by the results of theorem 2. Thus theorem 1 applies to equation (2. 55) in section 2. 4.

2. 6 An Alternative Linearization Procedure

There are other possible linearization procedures that can be applied to the nonlinear control problem. One procedure suggested by Pearson[24] has the advantage of being computationally simpler than the methods of sections 2. 3 and 2. 4, but it is theoretically less attractive.

To present the theory for this method, we follow the approach used in section 2. 3. However, instead of the linearization used there, we use the following linearizations.

$$\underline{x}(k+1) \simeq \underline{F}(\underline{x}^*(k), \underline{u}^*(k), k)\, \underline{x}(k) + \underline{G}(\underline{x}^*(k), \underline{u}^*(k), k)\, \underline{u}(k) \qquad (2.76)$$

$$\underline{y}(k) \simeq \underline{H}(\underline{x}^*(k), k)\, \underline{x}(k) \qquad (2.77)$$

where $\underline{F}$ and $\underline{G}$ are determined such that

$$\underline{f}(\underline{x}(k), \underline{u}(k), k) = \underline{F}(\underline{x}(k), \underline{u}(k), k)\, \underline{x}(k) + \underline{G}(\underline{x}(k), \underline{u}(k), k)\, \underline{u}(k) \qquad (2.78)$$

$$\underline{h}(\underline{x}(k), k) = \underline{H}(\underline{x}(k), k)\, \underline{x}(k) \qquad (2.79)$$

This type of linearization is not unique, and it is an open question as to which linearization of this type is best. However, in many instances there is an obvious intuitively appealing way to proceed.

As an example of such a linearization, consider the scalar nonlinear function

$$f(x(k), u(k), k) = -x^3(k) + \sqrt[3]{u(k)} \tag{2.80}$$

One possible linearization is

$$f(x(k), u(k), k) \simeq -x^{*2}(k)\, x(k) + u^{*-2/3}(k)\, u(k) \tag{2.81}$$

Another one, arbitrarily chosen, is

$$f(x(k), u(k), k) \simeq [-x^{*2}(k) - u^*(k)]\, x(k) + [u^{*-2/3}(k) + x^*(k)]\, u(k) \tag{2.82}$$

The first, of course, is intuitively more appealing.

By using the linearizations outlined above instead of equations (2.23) and (2.24), equation (2.26) becomes

$$\frac{1}{2}\|\underline{x}(k)\|^2_{\underline{P}(k)} + \underline{x}'(k)\hat{\underline{x}}(k) + a(k) = \underset{\underline{u}(k)}{\text{Min}} \left\{ \frac{1}{2}\|\underline{z}(k) - \underline{y}(k)\|^2_{\underline{Q}(k)} + \frac{1}{2}\|\underline{u}(k)\|^2_{\underline{R}(k)} + \frac{1}{2}\|\underline{F}\,\underline{x}(k) + \underline{G}\,\underline{u}(k)\|^2_{\underline{P}(k+1)} + [\underline{F}\,\underline{x}(k) + \underline{G}\,\underline{u}(k)]'\hat{\underline{x}}(k+1) + a(k+1) \right\} \tag{2.83}$$

(When the arguments of $\underline{F}$, $\underline{G}$, and $\underline{H}$ are omitted, they are understood to be evaluated at the points $\underline{x}^*(k)$, $\underline{u}^*(k)$, and k).

The minimizing value of $\underline{u}(k)$ is

$$\underline{u}(k) = -[\underline{R}(k) + \underline{G}'\underline{P}(k+1)\underline{G}]^{-1}\,\underline{G}'[\underline{P}(k+1)\underline{F}\,\underline{x}(k) + \hat{\underline{x}}(k+1)] \tag{2.84}$$

When this value of $\underline{u}(k)$ is substituted into equation (2.83), we get

$$\begin{aligned} \frac{1}{2}\|\underline{x}(k)\|^2_{\underline{P}(k)} + \underline{x}'(k)\hat{\underline{x}}(k) + a(k) = {} & \frac{1}{2}\|\underline{z}(k) - \underline{H}\,\underline{x}(k)\|^2_{\underline{Q}(k)} \\ & - \frac{1}{2}\|\underline{P}(k+1)\underline{F}\,\underline{x}(k) + \hat{\underline{x}}(k+1)\|^2_{\underline{G}[\underline{R}(k)+\underline{G}'\underline{P}(k+1)\underline{G}]^{-1}\underline{G}'} \\ & + \frac{1}{2}\|\underline{F}\,\underline{x}(k)\|^2_{\underline{P}(k+1)} + \underline{x}'(k)\underline{F}'\hat{\underline{x}}(k+1) + a(k+1) \end{aligned} \tag{2.85}$$

This equation will be satisfied for all $\underline{x}(k)$ if and only if the following equations are satisfied.

$$\underline{P}(k) = \underline{H}'\underline{Q}(k)\underline{H} + \underline{F}'\underline{M}(k)\underline{P}(k+1)\underline{F} \tag{2.86}$$

$$\hat{\underline{x}}(k) = \underline{F}'\underline{M}(k)\hat{\underline{x}}(k+1) - \underline{H}'\underline{Q}(k)\underline{z}(k) \tag{2.87}$$

$$a(k) = \frac{1}{2}\|\underline{z}(k)\|^2_{\underline{Q}(k)} - \frac{1}{2}\|\hat{\underline{x}}(k+1)\|^2_{\underline{G}[\underline{R}(k)+\underline{G}'\underline{P}(k+1)\underline{G}]^{-1}\underline{G}'} \tag{2.88}$$

where

$$\underline{M}(k) = \underline{I} - \underline{P}(k+1)\underline{G}[\underline{R}(k) + \underline{G}'\underline{P}(k+1)\underline{G}]^{-1}\underline{G}' \tag{2.89}$$

The boundary conditions are again

$$\underline{P}(N+1) = \underline{0} \tag{2.90}$$

$$\hat{\underline{x}}(N+1) = \underline{0} \tag{2.91}$$

$$a(N+1) = 0 \tag{2.92}$$

As can be seen, these equations are identical in form to the solution equations for the linear system. The only difference is that the matrices $\underline{F}$, $\underline{G}$, and $\underline{H}$ in this section are functions of $\underline{x}^*(k)$ and $\underline{u}^*(k)$ as well as of k.

An iterative type solution, similar to that introduced in section 2.4 is possible here also. However, we cannot show that this iterative solution converges to the exact optimal nonlinear solutions. The reason for this can be seen by comparing equation (2.44) of the exact optimal nonlinear solution, rewritten here,

$$\underline{\lambda}(k-1) = \underline{f}_{\underline{x}}'(\underline{x}(k), \underline{u}(k), k)\,\underline{\lambda}(k) + \underline{h}_{\underline{x}}'\underline{Q}(k)[\underline{z}(k) - \underline{h}(\underline{x}(k), k)] \tag{2.44}$$

with the equation corresponding to equation (2.51) when the approximations of this section are used. This equation would be

$$\underline{\lambda}(k-1) = \underline{F}'\underline{\lambda}(k) + \underline{H}'\underline{Q}(k)[\underline{z}(k) - \underline{H}\underline{x}(k)] \tag{2.93}$$

It is obvious that equation (2.93) will not approach equation (2.44) as $\underline{x}(k)$ approaches $\underline{x}^*(k)$. Thus the convergent solution of the iteration procedure based on the equations of this section will not in general be the exact optimal solution. We could only hope that this solution would be very near the true optimum.

2.7 Nonlinear Systems with Stochastic Disturbances

This section presents a technique for controlling a nonlinear system that is subject to stochastic disturbances. Such a system can be described by the equations

$$\underline{x}(k+1) = \underline{f}(\underline{x}(k), \underline{u}(k), k) + \underline{r}(k); \qquad \underline{x}(0) = \underline{c} \tag{2.94}$$

$$\underline{y}(k) = \underline{h}(\underline{x}(k), k) \tag{2.95}$$

where $\underline{r}(k)$ is an n-dimensional random vector such that $\underline{r}(j)$ is independent of $\underline{r}(k)$ for $j \neq k$. Thus $\underline{r}(k)$ is essentially the discrete time equivalent of white noise.

If the nonlinear system we are interested in controlling is disturbed by a random input that is not independent as described above, but instead is disturbed by a random vector that can be described by the difference equation

$$\underline{r}(k) = \underline{\phi}(\underline{r}(k), k) + \underline{w}(k); \qquad \underline{r}(0) = \underline{w} \tag{2.96}$$

where $\underline{w}(k)$ is an independent random sequence, then the system equations can be transformed into the form of (2. 94) and (2. 95) by augmenting the state variables. This can best be illustrated by a simple example.

Suppose the system is described by the equations

$$x(k+1) = x(k)\, u(k)\, r(k) \tag{2.97}$$

and

$$r(k+1) = ar(k) + w(k) \tag{2.98}$$

where $w(k)$ is an independent scalar random variable. We can define an augmented state vector

$$\underline{x}(k) = \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} = \begin{bmatrix} x(k) \\ r(k) \end{bmatrix} \tag{2.99}$$

and write the system equations as

$$\underline{x}(k+1) = \underline{f}(\underline{x}(k), u(k), k) + \underline{r}(k) \tag{2.100}$$

where

$$\underline{f}(\underline{x}(k), u(k), k) = \begin{bmatrix} x_1(k)\, u(k)\, x_2(k) \\ ax_2(k) \end{bmatrix} \tag{2.101}$$

and

$$\underline{r}(k) = \begin{bmatrix} 0 \\ w(k) \end{bmatrix} \tag{2.102}$$

which is in the form of equation (2. 94).

Because the variables involved in equations (2.94) and (2.95) are stochastic, a reasonable performance index will involve an expectation. Thus we assume the performance index is

$$J(k) = \underset{\underline{r}(k), \ldots, \underline{r}(N-1)}{\text{Exp}} \left\{ \sum_{j=k}^{N} \frac{1}{2} \| \underline{z}(j) - \underline{y}(j) \|^2_{\underline{Q}(j)} + \sum_{j=k}^{N-1} \frac{1}{2} \| \underline{u}(j) \|^2_{\underline{R}(j)} \right\} \tag{2.103}$$

In order to proceed by dynamic programming, we define the value function

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k), \ldots, \underline{u}(N-1)}{\text{Min}} \{ J(k) \} \tag{2.104}$$

Bellman[3] shows that when the $\underline{r}(k)$ sequence is independent, the principle of optimality implies

$$V_{N-k}(\underline{x}(k)) = \underset{\underline{u}(k)}{\text{Min}} \ \underset{\underline{r}(k)}{\text{Exp}} \left\{ \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{u}(k) \|^2_{\underline{R}(k)} + V_{N-k-1}(\underline{x}(k+1)) \right\} \tag{2.105}$$

As before, if we assume

$$\underline{x}(k+1) \simeq \underline{f} + f_{\underline{x}} [\underline{x}(k) - \underline{x}^*(k)] + f_{\underline{u}} [\underline{u}(k) - \underline{u}^*(k)] + \underline{r}(k) \tag{2.106}$$

$$\underline{y}(k) \simeq \underline{h} + h_{\underline{x}} [\underline{x}(k) - \underline{x}^*(k)] \tag{2.107}$$

and

$$V_{N-k}(\underline{x}(k)) = \frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k) \hat{\underline{x}}(k) + a(k) \tag{2.108}$$

then we obtain

$$\begin{aligned} \frac{1}{2} \| \underline{x}(k) \|^2_{\underline{P}(k)} + \underline{x}'(k) \hat{\underline{x}}(k) + a(k) = \underset{\underline{u}(k)}{\text{Min}} \ \underset{\underline{r}(k)}{\text{Exp}} \Big\{ & \frac{1}{2} \| \underline{z}(k) - \underline{y}(k) \|^2_{\underline{Q}(k)} + \frac{1}{2} \| \underline{u}(k) \|^2_{\underline{R}(k)} \\ & + \frac{1}{2} \| f_{\underline{x}} \underline{x}(k) + \underline{b}(k) + \underline{r}(k) \|^2_{\underline{P}(k+1)} + [f_{\underline{x}} \underline{x}(k) + f_{\underline{u}} \underline{u}(k) + \underline{b}(k) + \underline{r}(k)]' \hat{\underline{x}}(k+1) + a(k+1) \Big\} \end{aligned} \tag{2.109}$$

where

$$\underline{b}(k) = \underline{f} - f_{\underline{x}} \underline{x}^*(k) - f_{\underline{u}} \underline{u}^*(k) \tag{2.110}$$

Performing the expectation operation and then the minimization operation yields, assuming $\underset{\underline{r}(k)}{\text{Exp}} \{ \underline{r}(k) \} = \underline{0}$,

$$\underline{u}(k) = -[\underline{R}(k) + f_{\underline{u}}' \underline{P}(k+1) f_{\underline{u}}]^{-1} f_{\underline{u}}' [\underline{P}(k+1) f_{\underline{x}} \underline{x}(k) + \underline{P}(k+1) \underline{b}(k) + \hat{\underline{x}}(k+1)] \tag{2.111}$$

and

$$\frac{1}{2}\,||\underline{x}(k)||^2_{\underline{P}(k)} + \underline{x}'(k)\hat{\underline{x}}(k) + a(k) = \frac{1}{2}\,||\underline{z}(k) - h_{\underline{x}}\underline{x}(k) - \underline{c}(k)||^2_{\underline{Q}(k)} + \underset{\underline{r}(k)}{\text{Exp}}\,\{\underline{r}'(k)\underline{P}(k+1)\underline{r}(k)\} + a(k+1)$$

$$- \frac{1}{2}\,||\underline{P}(k+1)f_{\underline{x}}\underline{x}(k) + \underline{P}(k+1)\,\underline{b}(k) + \hat{\underline{x}}(k+1)||^2_{f_{\underline{u}}[\underline{R}(k) + f'_{\underline{u}}\underline{P}(k+1)f_{\underline{u}}]^{-1}f'_{\underline{u}}} \qquad (2.112)$$

$$+ \frac{1}{2}\,||f_{\underline{x}}\,\underline{x}(k) + \underline{b}(k)||^2_{\underline{P}(k+1)} + [f_{\underline{x}}\,\underline{x}(k) + \underline{b}(k)]'\,\hat{\underline{x}}(k+1)$$

where

$$\underline{c}(k) = \underline{h} - h_{\underline{x}}\,\underline{x}^*(k) \qquad (2.113)$$

This equation will be satisfied for all $\underline{x}(k)$ if and only if the following equations are satisfied:

$$\underline{P}(k) = h'_{\underline{x}}\,\underline{Q}(k)\,h_{\underline{x}} + f'_{\underline{x}}\,\underline{M}(k)\,\underline{P}(k+1)\,f_{\underline{x}} \qquad (2.114)$$

$$\hat{\underline{x}}(k) = f'_{\underline{x}}\,\underline{M}(k)\,[\underline{P}(k+1) + \underline{b}(k) + \hat{\underline{x}}(k+1)] - h'_{\underline{x}}\,\underline{Q}(k)\,[\underline{z}(k) - \underline{c}(k)] \qquad (2.115)$$

$$a(k) = a(k+1) + \frac{1}{2}\,||\underline{z}(k) - \underline{c}(k)||^2_{\underline{Q}(k)} + \frac{1}{2}\,||\underline{b}(k)||^2_{\underline{P}(k+1)} + \underline{b}'(k)\hat{\underline{x}}(k+1) + \underset{\underline{r}(k)}{\text{Exp}}\,\{\underline{r}'(k)\underline{P}(k+1)\underline{r}(k)\}$$

$$- \frac{1}{2}\,||\underline{P}(k+1)\underline{b}(k) + \hat{\underline{x}}(k+1)||^2_{f_{\underline{u}}[\underline{R}(k) + f'_{\underline{u}}\underline{P}(k+1)f_{\underline{u}}]^{-1}f'_{\underline{u}}} \qquad (2.116)$$

where

$$\underline{M}(k) = \underline{I} - \underline{P}(k+1)\,f_{\underline{u}}\,[\underline{R}(k) + f'_{\underline{u}}\,\underline{P}(k+1)\,f_{\underline{u}}]^{-1}\,f'_{\underline{u}} \qquad (2.117)$$

The boundary values are

$$\underline{P}(N+1) = \underline{0} \qquad (2.118)$$

$$\hat{\underline{x}}(N+1) = \underline{0} \qquad (2.119)$$

$$a(N+1) = 0 \qquad (2.120)$$

These equations are identical to equations (2.31) through (2.37) except for the additional expectation term in equation (2.116).

In essence, these equations are the solution to the optimal control problem for the linearized system. This solution differs from the exact optimal nonlinear solution because the linearized system only approximates the nonlinear system. In section 2.4 we were able to improve this approximation by an iterative technique so that eventually the exact solution was obtained. What are the prospects of a similar procedure in this case?

An examination of the iterative procedure of section 2.4 reveals that the technique was dependent on being able to predict exactly the state at time k+1 which results from the application of a known control signal to the system in a known state at time k. Unfortunately, because of the random disturbance, $\underline{r}(k)$, this is impossible for the system considered in this section.

We can, however, use the following iterative algorithm to obtain an approximate solution:

Step 1. Solve equations (2.114) and (2.115) backward in time using

$$\underline{x}^*(j) = \underline{x}_i(j) \tag{2.121}$$

$$\underline{u}^*(j) = \underline{u}_i(j) \tag{2.122}$$

to compute $\underline{P}(N), \ldots, \underline{P}(k)$ and $\hat{\underline{x}}(N), \ldots, \hat{\underline{x}}(k)$.

Step 2. Solve equations (2.111) and (2.106) forward in time with

$$\underline{r}(j) = \underline{0}; \qquad j \geq k \tag{2.123}$$

and again using

$$\underline{x}^*(j) = \underline{x}_i(j) \tag{2.124}$$

$$\underline{u}^*(j) = \underline{u}_i(j) \tag{2.125}$$

to compute $\underline{u}_{i+1}(k), \ldots, \underline{u}_{i+1}(N-1)$ and $\underline{x}_{i+1}(k), \ldots, \underline{x}_{i+1}(N)$.

The i+1st iteration would then proceed using the extrapolated control vectors, $\underline{u}_{i+1}(j)$, and the extrapolated state vectors, $\underline{x}_{i+1}(j)$, just computed in place of $\underline{u}^*(j)$ and $\underline{x}^*(j)$. The procedure would be repeated until satisfactory convergence had been achieved.

This algorithm should provide nearly optimal performance when the $\underline{P}(j)$ and $\hat{\underline{x}}(j)$ obtained in this fashion are used to generate the control for the real system. As time goes on, and the true state deviates more and more from the extrapolated state, the performance will slowly be degraded.

One way to overcome partially this degradation of performance is to update the solution periodically by measuring the current state of the system, and then using this state as the starting point for a recomputation

of $\underline{P}(j)$ and $\hat{\underline{x}}(j)$, using the same iterative procedure as before. Of course, this would require that the iterative algorithm be executed in much faster time than the real system evolves.

Some computer results using this approach are presented in Chapter V.

CHAPTER III

CONTINUOUS TIME SYSTEMS

3.1 Introduction

Even a cursory examination of the results of Chapter II shows that the control systems required by the theory are of such complexity that a high speed digital computer will generally be required to investigate or to synthesize the control system. However, for the few analog control system applications that may be possible, and for a few special nonlinear control problems that can be solved analytically, a continuous time theory is required.

The purpose of this chapter is to develop the theory for the control of continuous time nonlinear systems. This theory is developed in a manner analogous to that used in Chapter II for the discrete time systems. It should be mentioned here that Kalman[7] and Merriam[15,16] have developed the theory for linear continuous time systems.

3.2 Linear Systems

Consider the linear control system described by the equations

$$\dot{\underline{x}}(t) = \underline{F}(t)\underline{x}(t) + \underline{G}(t)\underline{u}(t); \qquad \underline{x}(0) = \underline{c} \tag{3.1}$$

$$\underline{y}(t) = \underline{H}(t)\underline{x}(t) \tag{3.2}$$

where $\underline{x}(t)$ is the n-dimensional system state vector, $\underline{u}(t)$ is the r-dimensional control or input vector, and $\underline{y}(t)$ is the m-dimensional system output vector. As indicated by the notation, the transformation matrices $\underline{F}(t)$, $\underline{G}(t)$, and $\underline{H}(t)$ as well as the vectors $\underline{x}(t)$, $\underline{u}(t)$, and $\underline{y}(t)$ can vary continuously with time. For this system we wish to find the control $\underline{u}(\tau)$ on the interval $t \leq \tau \leq T$ such that the performance index

$$J(t) = \int_t^T \left[\frac{1}{2}\|\underline{z}(\tau) - \underline{y}(\tau)\|^2_{\underline{Q}(\tau)} + \frac{1}{2}\|\underline{u}(\tau)\|^2_{\underline{R}(\tau)}\right] d\tau \tag{3.3}$$

is a minimum. Here, $\underline{z}(\tau)$ is the desired output of the system.

We define the value function

$$V(\underline{x}(t), t) = \underset{t \leq \tau \leq T}{\underset{\underline{u}(\tau)}{\text{Min}}} \{J(t)\} \tag{3.4}$$

By the principle of optimality, we have

$$V(\underline{x}(t), t) = \underset{t \leq \tau \leq t+\Delta t}{\underset{\underline{u}(\tau)}{\text{Min}}} \left\{ \int_t^{t+\Delta t} \left[\frac{1}{2} || \underline{z}(\tau) - \underline{y}(\tau) ||^2_{\underline{Q}(\tau)} + \frac{1}{2} || \underline{u}(\tau) ||^2_{\underline{R}(\tau)} \right] d\tau + V(\underline{x}(t+\Delta t), t+\Delta t) \right\} \tag{3.5}$$

If we expand $V(\underline{x}(t+\Delta t), t+\Delta t)$ in a Taylor series about the point $[\underline{x}(t), t]$, we get

$$V(\underline{x}(t), t) = \underset{t \leq \tau \leq t+\Delta t}{\underset{\underline{u}(\tau)}{\text{Min}}} \left\{ \int_t^{t+\Delta t} \left[\frac{1}{2} || \underline{z}(\tau) - \underline{y}(\tau) ||^2_{\underline{Q}(\tau)} + \frac{1}{2} || \underline{u}(\tau) ||^2_{\underline{R}(\tau)} \right] d\tau \right.$$
$$\left. + V(\underline{x}(t), t) + V_t(\underline{x}(t), t)\Delta t + V_{\underline{x}}'(\underline{x}(t), t)[\underline{x}(t+\Delta t) - \underline{x}(t)] + 0(\Delta t) \right\} \tag{3.6}$$

When we take the limit as Δt approaches zero (provided it exists, etc.), equation (3.6) becomes

$$V_t + \underset{\underline{u}(t)}{\text{Min}} \left[\frac{1}{2} || \underline{z}(t) - \underline{y}(t) ||^2_{\underline{Q}(t)} + \frac{1}{2} || \underline{u}(t) ||^2_{\underline{R}(t)} + V_{\underline{x}}' \dot{\underline{x}}(t) \right] = 0 \tag{3.7}$$

or

$$V_t + \underset{\underline{u}(t)}{\text{Min}} \left[\frac{1}{2} || \underline{z}(t) - \underline{y}(t) ||^2_{\underline{Q}(t)} + \frac{1}{2} || \underline{u}(t) ||^2_{\underline{R}(t)} + V_{\underline{x}}' \underline{F}(t)\underline{x}(t) + V_{\underline{x}}' \underline{G}(t)\underline{u}(t) \right] = 0 \tag{3.8}$$

The minimization can be performed by ordinary methods of calculus yielding

$$\underline{u}_{Min}(t) = -\underline{R}^{-1}(t)\underline{G}'(t)V_{\underline{x}} \tag{3.9}$$

Substituting this value of $\underline{u}(t)$ into equation (3.8) yields the Hamilton-Jacobi equation,

$$V_t + \frac{1}{2} || \underline{z}(t) - \underline{H}(t)\underline{x}(t) ||^2_{\underline{Q}(t)} - \frac{1}{2} || V_{\underline{x}} ||^2_{\underline{G}(t)\underline{R}^{-1}(t)\underline{G}'(t)} + V_{\underline{x}}' \underline{F}(t)\underline{x}(t) = 0 \tag{3.10}$$

The solution for this equation can be obtained by assuming

$$V(\underline{x}(t), t) = \frac{1}{2} || \underline{x}(t) ||^2_{\underline{P}(t)} + \underline{x}'(t)\hat{\underline{x}}(t) + a(t) \tag{3.11}$$

Hence,

$$V_t = \frac{1}{2} || \underline{x}(t) ||^2_{\dot{\underline{P}}(t)} + \underline{x}'(t)\dot{\hat{\underline{x}}}(t) + \dot{a}(t) \tag{3.12}$$

and

$$V_{\underline{x}} = \underline{P}(t)\underline{x}(t) + \hat{\underline{x}}(t) \tag{3.13}$$

After substituting these expressions into equation (3.10), we obtain

$$\frac{1}{2}\|\underline{x}(t)\|^2_{\dot{\underline{P}}(t)} + \underline{x}'(t)\dot{\hat{\underline{x}}}(t) + \dot{a}(t) + \frac{1}{2}\|\underline{z}(t) - \underline{H}(t)\underline{x}(t)\|^2_{\underline{Q}(t)}$$
$$-\frac{1}{2}\|\underline{P}(t)\underline{x}(t) + \hat{\underline{x}}(t)\|^2_{\underline{G}(t)\underline{R}^{-1}(t)\underline{G}'(t)} + [\underline{P}(t)\underline{x}(t) + \hat{\underline{x}}(t)]'\underline{F}(t)\underline{x}(t) = 0 \qquad (3.14)$$

This equation can be satisfied for all $\underline{x}(t)$ if and only if the following equations are satisfied:

$$\dot{\underline{P}}(t) = \underline{P}(t)\underline{G}(t)\underline{R}^{-1}(t)\underline{G}'(t)\underline{P}(t) - \underline{P}(t)\underline{F}(t) - \underline{F}'(t)\underline{P}(t) - \underline{H}'(t)\underline{Q}(t)\underline{H}(t) \qquad (3.15)$$

$$\dot{\hat{\underline{x}}}(t) = [\underline{P}(t)\underline{G}(t)\underline{R}^{-1}(t)\underline{G}'(t) - \underline{F}'(t)]\hat{\underline{x}}(t) + \underline{H}'(t)\underline{Q}(t)\underline{z}(t) \qquad (3.16)$$

$$a(t) = \frac{1}{2}\|\hat{\underline{x}}(t)\|^2_{\underline{G}(t)\underline{R}^{-1}(t)\underline{G}'(t)} - \frac{1}{2}\|\underline{z}(t)\|^2_{\underline{Q}(t)} \qquad (3.17)$$

The boundary conditions for these equations can be obtained from equations (3.3) and (3.11). They are

$$\underline{P}(T) = \underline{0} \qquad (3.18)$$

$$\hat{\underline{x}}(T) = \underline{0} \qquad (3.19)$$

$$a(T) = 0 \qquad (3.20)$$

Here again, these equations must be solved backwards in time, but they do not depend on the state of the system. Therefore, they can be pre-computed, as in the discrete time case if the desired output, $\underline{z}(\tau)$, is known on the interval $t \leq \tau \leq T$. The control can be realized in the form of the block diagram shown in figure 3.1.

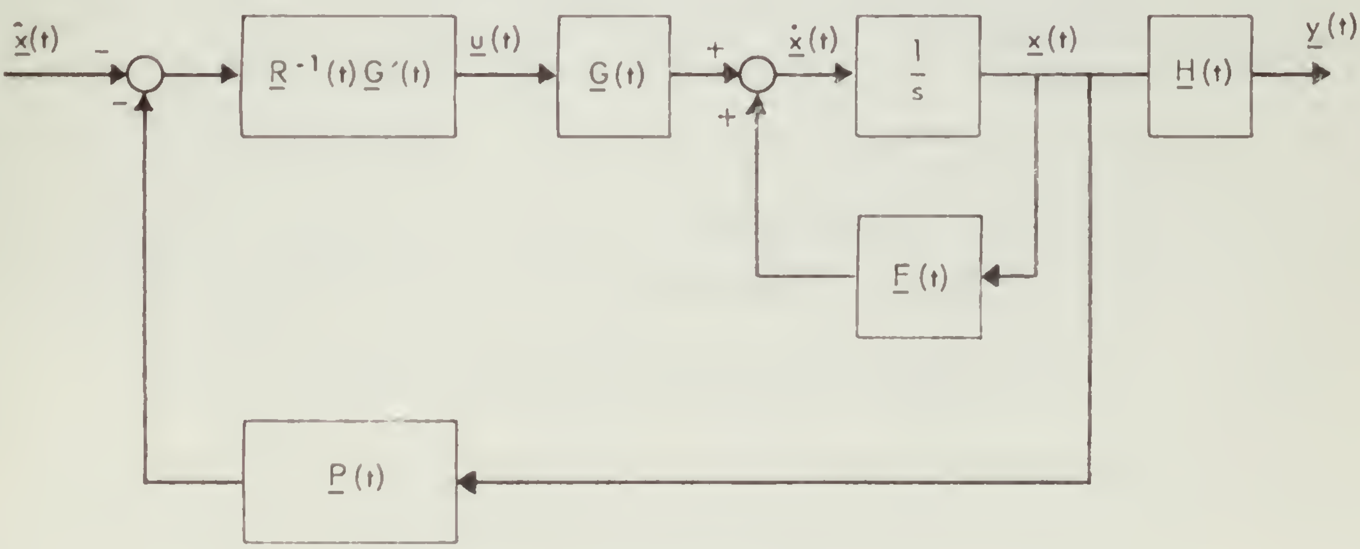


Figure 3.1 - Continuous Time Optimal Linear Control System

3.3 Nonlinear Systems

The nonlinear systems we consider here can be described by the equations

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}(t), \underline{u}(t), t); \qquad \underline{x}(0) = \underline{c} \tag{3.21}$$

$$\underline{y}(t) = \underline{h}(\underline{x}(t), t) \tag{3.22}$$

where $\underline{x}(t)$, $\underline{u}(t)$, and $\underline{y}(t)$ are state, control, and output vectors, as before, and where $\underline{f}(\underline{x}(t), \underline{u}(t), t)$ and $\underline{h}(\underline{x}(t), t)$ are continuous time vector valued functions. It is necessary to assume that $\underline{f}$ and $\underline{h}$ satisfy certain differentiability conditions in what follows. Whenever derivatives of these functions appear, we will tacitly assume that they exist.

For the system just described, we wish to find the control, $\underline{u}(\tau)$, on the interval, $t \leq \tau \leq T$, such that the performance index

$$J(t) = \int_t^T \left[\frac{1}{2} \| \underline{z}(\tau) - \underline{y}(\tau) \|^2_{\underline{Q}(\tau)} + \frac{1}{2} \| \underline{u}(\tau) \|^2_{\underline{R}(\tau)} \right] d\tau \tag{3.23}$$

is a minimum.

We define the value function

$$V(\underline{x}(t), t) = \underset{t \leq \tau \leq T}{\underset{\underline{u}(\tau)}{\text{Min}}} \{ J(t) \} \tag{3.24}$$

Then by the principle of optimality,

$$V(\underline{x}(t), t) = \underset{t \leq \tau \leq t+\Delta t}{\underset{\underline{u}(\tau)}{\text{Min}}} \left\{ \int_t^{t+\Delta t} \left[\frac{1}{2} \| \underline{z}(\tau) - \underline{y}(\tau) \|^2_{\underline{Q}(\tau)} + \frac{1}{2} \| \underline{u}(\tau) \|^2_{\underline{R}(\tau)} \right] d\tau + V(\underline{x}(t+\Delta t), t+\Delta t) \right\} \tag{3.25}$$

By expanding $V(\underline{x}(t+\Delta t), t+\Delta t)$ in a Taylor series about $\underline{x}(t)$ and t, and then taking the limit as Δt approaches 0, we get

$$V_t + \underset{\underline{u}(t)}{\text{Min}} \left[\frac{1}{2} \| \underline{z}(t) - \underline{y}(t) \|^2_{\underline{Q}(t)} + \frac{1}{2} \| \underline{u}(t) \|^2_{\underline{R}(t)} + V_{\underline{x}}' \dot{\underline{x}}(t) \right] = 0 \tag{3.26}$$

Since the system is nonlinear, we cannot solve this Hamilton-Jacobi equation directly in general. So, as in Chapter II, we resort to linearization. We use the approximations

$$\dot{\underline{x}}(t) \simeq \underline{f}(\underline{x}^*(t), \underline{u}^*(t), t) + f_{\underline{x}}(\underline{x}^*(t), \underline{u}^*(t), t)\,[\underline{x}(t) - \underline{x}^*(t)] + f_{\underline{u}}(\underline{x}^*(t), \underline{u}^*(t), t)\,[\underline{u}(t) - \underline{u}^*(t)] \quad (3.27)$$

and

$$\underline{y}(t) \simeq \underline{h}(\underline{x}^*(t), t) + h_{\underline{x}}(\underline{x}^*(t), t)\,[\underline{x}(t) - \underline{x}^*(t)] \quad (3.28)$$

With these approximations, equation (3.26) becomes

$$V_t + \underset{\underline{u}(t)}{\text{Min}} \left\{ \frac{1}{2} \| \underline{z}(t) - \underline{y}(t) \|^2_{\underline{Q}(t)} + \frac{1}{2} \| \underline{u}(t) \|^2_{\underline{R}(t)} + V_{\underline{x}}' \underline{f} + V_{\underline{x}}' f_{\underline{x}} [\underline{x}(t) - \underline{x}^*(t)] + V_{\underline{x}}' f_{\underline{u}} [\underline{u}(t) - \underline{u}^*(t)] \right\} = 0 \quad (3.29)$$

The minimization operation yields

$$\underline{u}_{Min}(t) = -\underline{R}^{-1}(t) f_{\underline{u}}' V_{\underline{x}} \quad (3.30)$$

and

$$V_t + \frac{1}{2} \| \underline{z}(t) - h_{\underline{x}} \underline{x}(t) - \underline{c}(t) \|^2_{\underline{Q}(t)} - \frac{1}{2} \| V_{\underline{x}} \|^2_{f_{\underline{u}} \underline{R}^{-1}(t) f_{\underline{u}}'} + V_{\underline{x}}' [f_{\underline{x}} \underline{x}(t) + \underline{b}(t)] = 0 \quad (3.31)$$

where

$$\underline{b}(t) = \underline{f} - f_{\underline{x}} \underline{x}^*(t) - f_{\underline{u}} \underline{u}^*(t) \quad (3.32)$$

and

$$\underline{c}(t) = \underline{h} - h_{\underline{x}} \underline{x}^*(t) \quad (3.33)$$

A solution for equation (3.31) can be obtained by assuming

$$V(\underline{x}(t), t) = \frac{1}{2} \| \underline{x}(t) \|^2_{\underline{P}(t)} + \underline{x}'(t) \hat{\underline{x}}(t) + a(t) \quad (3.34)$$

which implies

$$V_t(\underline{x}(t), t) = \frac{1}{2} \| \underline{x}(t) \|^2_{\dot{\underline{P}}(t)} + \underline{x}'(t) \dot{\hat{\underline{x}}}(t) + \dot{a}(t) \quad (3.35)$$

and

$$V_{\underline{x}}(\underline{x}(t), t) = \underline{P}(t) \underline{x}(t) + \hat{\underline{x}}(t) \quad (3.36)$$

Combining equations (3.31), (3.35), and (3.36) yields

$$\frac{1}{2} \| \underline{x}(t) \|^2_{\dot{\underline{P}}(t)} + \underline{x}'(t) \dot{\hat{\underline{x}}}(t) + \dot{a}(t) + \frac{1}{2} \| \underline{z}(t) - h_{\underline{x}} \underline{x}(t) - \underline{c}(t) \|^2_{\underline{Q}(t)} - \frac{1}{2} \| \underline{P}(t) \underline{x}(t) + \hat{\underline{x}}(t) \|^2_{f_{\underline{u}} \underline{R}^{-1}(t) f_{\underline{u}}'} + [\underline{P}(t) \underline{x}(t) + \hat{\underline{x}}(t)]' [f_{\underline{x}} \underline{x}(t) + \underline{b}(t)] = 0 \quad (3.37)$$

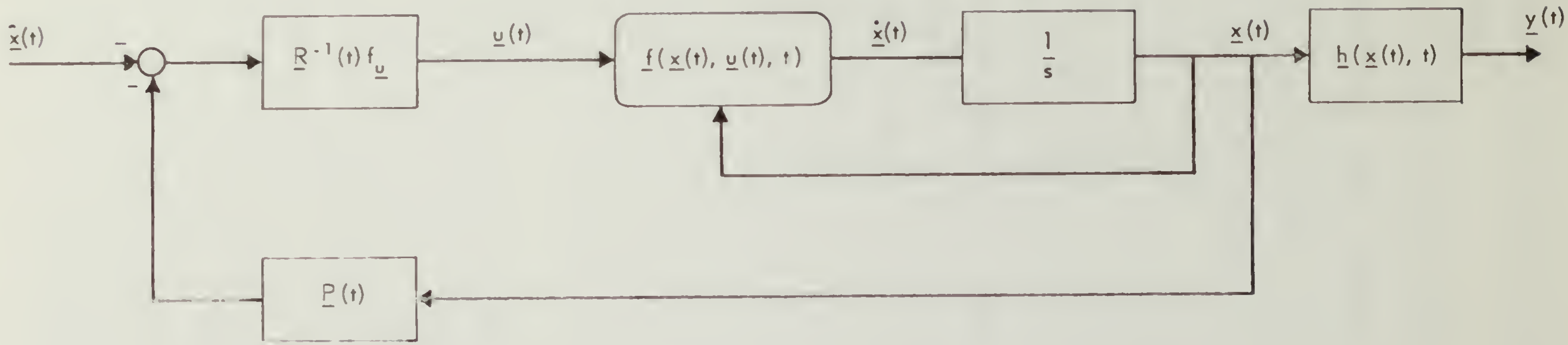


Figure 3.2 - Continuous Time Nonlinear Control System

This equation will be satisfied for all $\underline{x}(t)$ if and only if the following equations are satisfied:

$$\dot{\underline{P}}(t) = \underline{P}(t) f_{\underline{u}} \underline{R}^{-1}(t) f_{\underline{u}}' \underline{P}(t) - \underline{P}(t) f_{\underline{x}} - f_{\underline{x}}' \underline{P}(t) - h_{\underline{x}}' \underline{Q}(t) h_{\underline{x}} \tag{3.38}$$

$$\dot{\hat{\underline{x}}}(t) = [\underline{P}(t) f_{\underline{u}} \underline{R}^{-1}(t) f_{\underline{u}}' - f_{\underline{x}}'] \hat{\underline{x}}(t) - \underline{P}(t) \underline{b}(t) + h_{\underline{x}}' \underline{Q}(t) [\underline{z}(t) - \underline{c}(t)] \tag{3.39}$$

$$\dot{a}(t) = \frac{1}{2} \| \hat{\underline{x}}(t) \|^2_{f_{\underline{u}} \underline{R}^{-1}(t) f_{\underline{u}}'} - \frac{1}{2} \| \underline{z}(t) - \underline{c}(t) \|^2_{\underline{Q}(t)} + \underline{b}'(t) \hat{\underline{x}}(t) \tag{3.40}$$

The boundary conditions are the same as for the linear case.

If we are given $\underline{x}^*(t)$ and $\underline{u}^*(t)$, we can compute $\underline{P}(t)$ and $\hat{\underline{x}}(t)$ in advance. Then these parameters can be used to determine a near optimum control for the system. Of course, how near optimal the control system is depends on how good the approximations (3. 27) and (3. 28) are. Computationally, we can proceed in a manner analogous to the discrete time iterative procedure. To do this, we can use $\underline{x}(t)$ and $\underline{u}(t)$ determined by the *ith* iteration as $\underline{x}^*(t)$ and $\underline{u}^*(t)$ for the i+1*st* iteration. Similar to the iterative algorithm of section 2. 4, this algorithm can be shown to yield the exact solution to the continuous time nonlinear control problem.

The control system can be synthesized in the form of the block diagram of figure 3. 2. As can be seen from figure 3. 2, the continuous time control system is almost identical in form to the discrete time nonlinear control system.

Some additional insight into the problem of optimal control can be gained by examining the nature of the equations for $\underline{P}(t)$ and $\hat{\underline{x}}(t)$. As the quantity, T-t, approaches zero, $\underline{P}(t)$ and $\hat{\underline{x}}(t)$ approach zero. Hence, the optimum control signal approaches zero as the terminal time nears. On the other hand, when T-t is very large, and the system being controlled is linear time invariant, $\dot{\underline{P}}(t)$ is very small. We would expect that when T-t is very large, and when the time variations and nonlinearities of the system being controlled are not severe, $\dot{\underline{P}}(t)$ should be small, also. The director part of the input, $\hat{\underline{x}}(t)$, is derived from the desired output, $\underline{z}(t)$, by the feedback system shown in figure 3. 3.

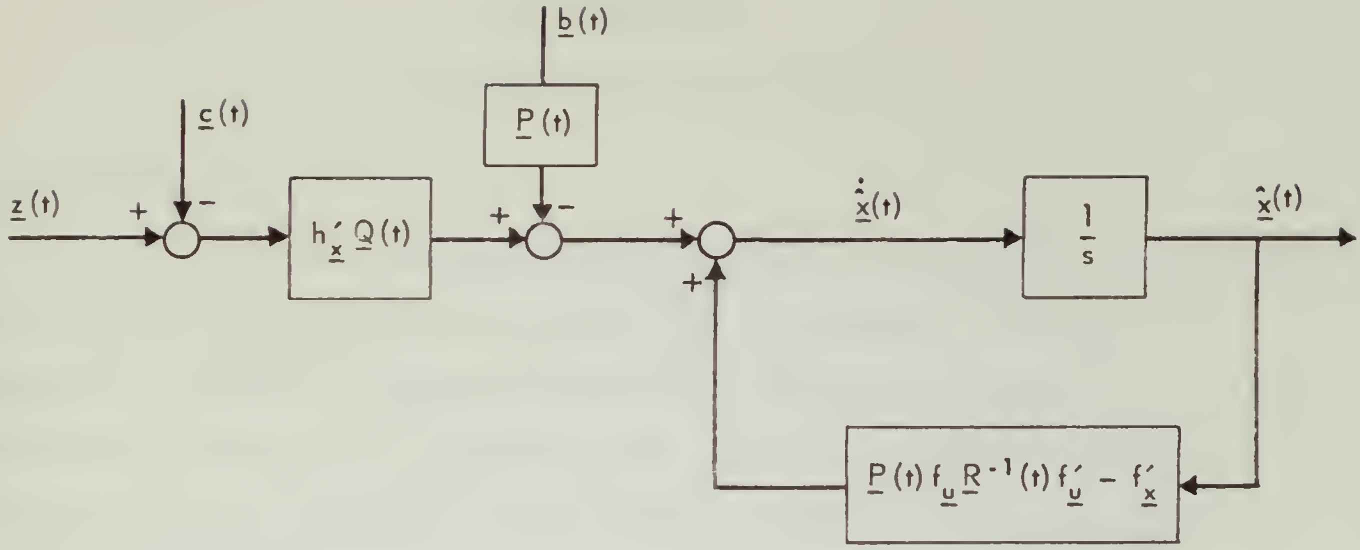


Figure 3.3 - Block Diagram of System for $\hat{\underline{x}}(t)$

If the output of this system follows the input reasonably well, the system synthesized using $h_{\underline{x}}'\underline{Q}(t)\,\underline{z}(t)$ in place of $\hat{\underline{x}}(t)$ might perform near optimally, provided $\underline{b}(t)$ and $\underline{c}(t)$ are reasonably small in magnitude.

The comments above have been imprecise, and were meant only to convey some insight into the problem beyond the bare mathematical statements.

CHAPTER IV

CONSERVATIVE SYSTEMS

4.1 Introduction

A special class of nonlinear systems which we shall call "conservative," can be treated analytically and exactly by the methods introduced in Chapters II and III. The purpose of this chapter is to study this class of nonlinear problems by means of two examples. Often, as much can be learned from the study of one analytic example as from a hundred numerical examples.

4.2 General

Consider the nonlinear system

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{u}; \qquad \underline{x}(0) = \underline{c} \tag{4.1}$$

If the performance criterion is

$$J = \int_{t}^{T} L(\underline{x}, \underline{u})\, dt \tag{4.2}$$

then the loss equation, equation (3.26), is

$$V_t + \underset{\underline{u}(t)}{\text{Min}} \left[L(\underline{x}, \underline{u}) + V_{\underline{x}}' \underline{f}(\underline{x}) + V_{\underline{x}}' \underline{u} \right] = 0 \tag{4.3}$$

If the term, $V_{\underline{x}}' \underline{f}(\underline{x})$, in equation (4.3) vanishes identically for all $\underline{x}$, it is possible for a great simplification to result. Of course V, and hence $V_{\underline{x}}$, depend strongly on the form of $L(\underline{x}, \underline{u})$. Thus $V_{\underline{x}}' \underline{f}(\underline{x})$ will vanish only if $L(\underline{x}, \underline{u})$ has a special form. Fortunately, this is sometimes the case in practical problems. The example problems which follow will serve to illustrate the nature of the special form $L(\underline{x}, \underline{u})$ must have to permit this simplification. In addition, the example problems will permit us to study the analytic solutions of some optimal nonlinear control problems, and compare them with some sub-optimal solutions.

4.3 Spinning Body Problem

The equations of motion for the angular velocities of a freely spinning body about three mutually perpendicular axes can be written as

$$\left.\begin{aligned} \dot{x}_1 &= a_1 x_2 x_3; & x_1(0) &= c_1 \\ \dot{x}_2 &= a_2 x_1 x_3; & x_2(0) &= c_2 \\ \dot{x}_3 &= a_3 x_1 x_2; & x_3(0) &= c_3 \end{aligned}\right\} \qquad (4.4)$$

where x_1, x_2, and x_3 are the angular velocities, and where

$$a_1 + a_2 + a_3 = 0 \qquad (4.4A)$$

These equations of motions are coupled and nonlinear. If we wish to control the spin of this system by exerting torques about each of the three axes, the equations of motion become

$$\left.\begin{aligned} \dot{x}_1 &= a_1 x_2 x_3 + u_1; & x_1(0) &= c_1 \\ \dot{x}_2 &= a_2 x_1 x_3 + u_2; & x_2(0) &= c_2 \\ \dot{x}_3 &= a_3 x_1 x_2 + u_3; & x_3(0) &= c_3 \end{aligned}\right\} \qquad (4.5)$$

where u_1, u_2, and u_3 are the control variables proportional to the torques.

If we wish to reduce the angular velocities to a minimum, subject to a constraint of the control effort expended, an appropriate performance criterion might be

$$J = \int_0^T \left\{ \frac{1}{2} q(t) \left[x_1^2 + x_2^2 + x_3^2 \right] + \frac{1}{2} r(t) \left[u_1^2 + u_2^2 + u_3^2 \right] \right\} dt^{\dagger} \qquad (4.6)$$

Optimal Control

The control which minimizes J can be found by the method of Chapter III. The loss equation is

$$\begin{aligned} V_t + \min_{u_1, u_2, u_3} \Big\{ & \frac{1}{2} q(t) \left[x_1^2 + x_2^2 + x_3^2 \right] + \frac{1}{2} r(t) \left[u_1^2 + u_2^2 + u_3^2 \right] \\ & + V_{x_1} a_1 x_2 x_3 + V_{x_2} a_2 x_1 x_3 + V_{x_3} a_3 x_1 x_2 + V_{x_1} u_1 + V_{x_2} u_2 + V_{x_3} u_3 \Big\} = 0 \end{aligned} \qquad (4.7)$$

---

† The spinning body control problem has been treated by Athans[14] and Windeknecht,[25] but their methods differ from that used here in significant respects.

If we assume

$$V = \frac{1}{2} p(t) \left[x_1^2 + x_2^2 + x_3^2 \right] \tag{4.8}$$

then the optimal control is

$$\left.\begin{aligned} u_1 &= -k(t) x_1 \\ u_2 &= -k(t) x_2 \\ u_3 &= -k(t) x_3 \end{aligned}\right\} \tag{4.9}$$

where

$$k(t) = p(t) / r(t), \tag{4.10}$$

and equation (4. 7) becomes

$$\frac{1}{2} [\dot{p}(t) - p^2(t)/r(t) + q(t)] \left[x_1^2 + x_2^2 + x_3^2 \right] = 0 \tag{4.11}$$

But since this equation must be true for all x_1, x_2, and x_3, we must have

$$\dot{p}(t) - p^2(t)/r(t) + q(t) = 0 \tag{4.12}$$

From the definition of V, the boundary condition is

$$p(T) = 0 \tag{4.13}$$

If q and r are constant, the solution for equation (4. 12) is

$$p(\tau) = r k(\tau) = r\alpha \left[\frac{1 - e^{-2\alpha\tau}}{1 + e^{-2\alpha\tau}} \right] \tag{4.14}$$

where

$$\alpha = \sqrt{q/r} \tag{4.15}$$

and

$$\tau = T - t \tag{4.16}$$

A plot of $p(\tau)$ is shown in figure 4. 1.

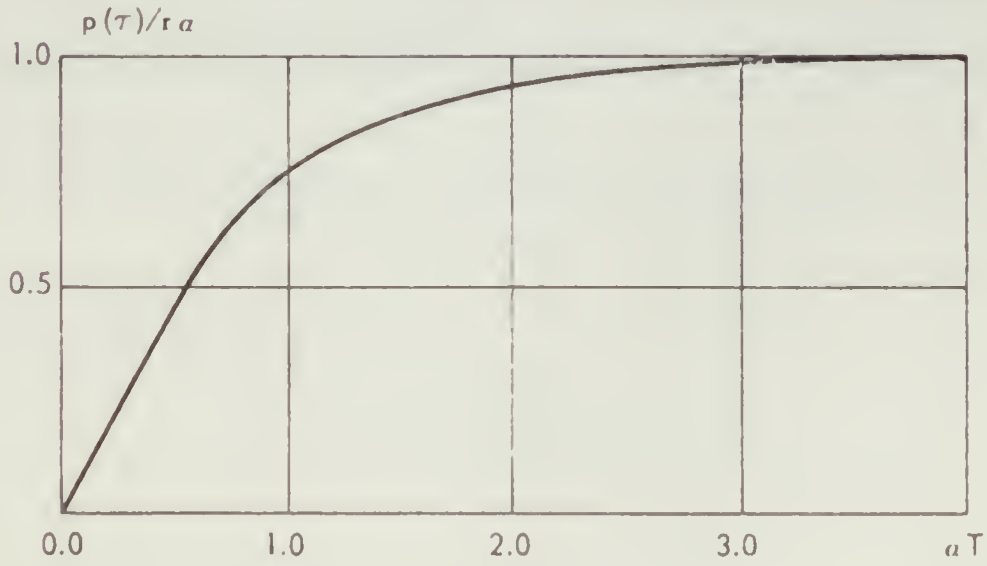


Figure 4. 1 - Plot of $p(\tau)/r\alpha$ versus $\alpha\tau$

Notice that the optimal controller is linear with time varying gains even though the system controlled is nonlinear. Also notice that the time varying gains reach 76 percent of their steady-state value in $\tau = 1/a$ seconds, 95 percent in $\tau = 2/a$ seconds, and 99.5 percent in $\tau = 3/a$ seconds. As is evident the quantity, $1/a$, plays the role of a time constant.

The controller may be realized in the form of the block diagram of figure 4.2.

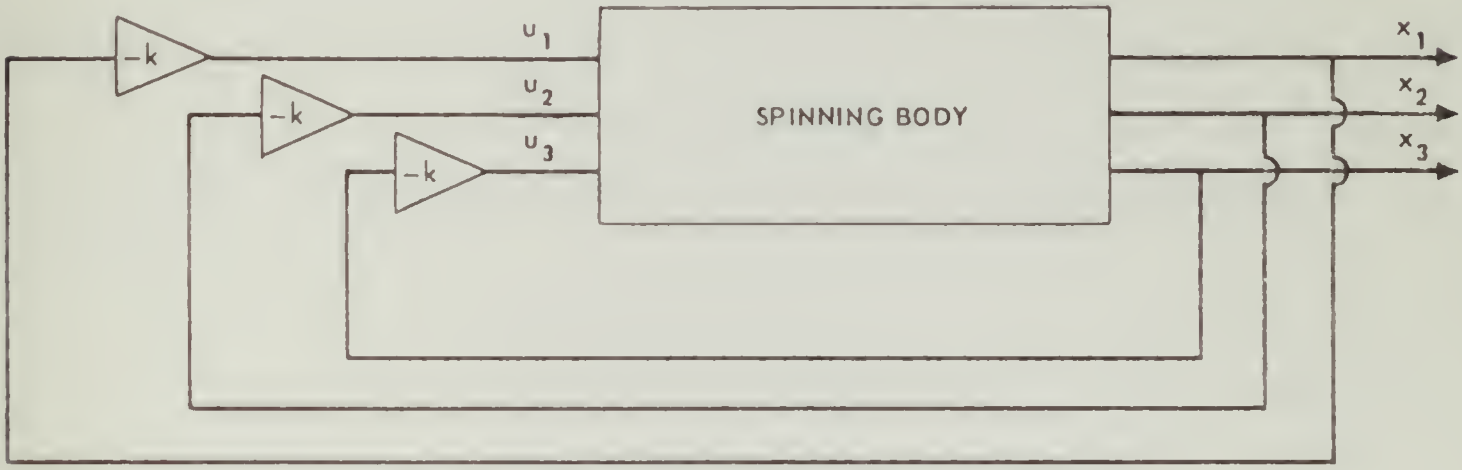


Figure 4.2 - Spinning Body Control System Block Diagram

Sub-optimal Control

It is instructive to compare the optimal control system of the last section with the sub-optimal control system which simply uses constant gains. In order to make this comparison, the performance criterion must be computed for the optimal and sub-optimal controls on the time interval $[0, T]$.

For the optimal control the performance criterion is

$$J^* = V(\underline{c}, 0) \tag{4.17}$$

or, for this problem,

$$J^* = \frac{1}{2} ra \left[\frac{1 - e^{-2aT}}{1 + e^{-2aT}}\right]\left[c_1^2 + c_2^2 + c_3^2\right] \tag{4.18}$$

For the sub-optimal control with

$$\left.\begin{aligned} u_1 &= -kx_1 \\ u_2 &= -kx_2 \\ u_3 &= -kx_3 \end{aligned}\right\} \tag{4.19}$$

the performance criterion is

$$J = \frac{1}{2}\left[q + rk^2\right]\int_0^T \left[x_1^2 + x_2^2 + x_3^2\right] dt \tag{4.20}$$

or

$$J = \frac{1}{2}\left[q + rk^2\right]\int_0^T W(t)\, dt \tag{4.21}$$

where

$$W(t) = x_1^2(t) + x_2^2(t) + x_3^2(t) \tag{4.22}$$

It is possible to compute W(t) from equation (4. 5) in the following manner:

$$\left.\begin{aligned} x_1\dot{x}_1 &= a_1 x_2 x_3 x_1 - kx_1^2 \\ x_2\dot{x}_2 &= a_2 x_1 x_3 x_2 - kx_2^2 \\ x_3\dot{x}_3 &= a_3 x_1 x_2 x_3 - kx_3^2 \end{aligned}\right\} \tag{4.23}$$

Adding, we get

$$\frac{1}{2}\, d/dt\left[x_1^2 + x_2^2 + x_3^2\right] = -k\left[x_1^2 + x_2^2 + x_3^2\right] \tag{4.24}$$

or

$$\dot{W}(t) + 2kW(t) = 0 \qquad W(0) = c_1^2 + c_2^2 + c_3^2 \;. \tag{4.25}$$

The solution of equation (4. 25) is

$$W(t) = e^{-2kt}\, W(0) \tag{4.26}$$

From this the sub-optimal performance criterion may be computed. This gives

$$J = \frac{[q + rk^2]}{4k}\,[1 - e^{-2kT}]\left[c_1^2 + c_2^2 + c_3^2\right] \tag{4.27}$$

For $k = a = \sqrt{q/r}$, J becomes

$$J = \frac{1}{2}\, ra\,[1 - e^{-2aT}]\left[c_1^2 + c_2^2 + c_3^2\right] \tag{4.28}$$

The ratio, J/J^*, is then simply

$$J/J^* = 1 + e^{-2aT} \tag{4.29}$$

A plot of J/J^* is shown in figure 4.3.

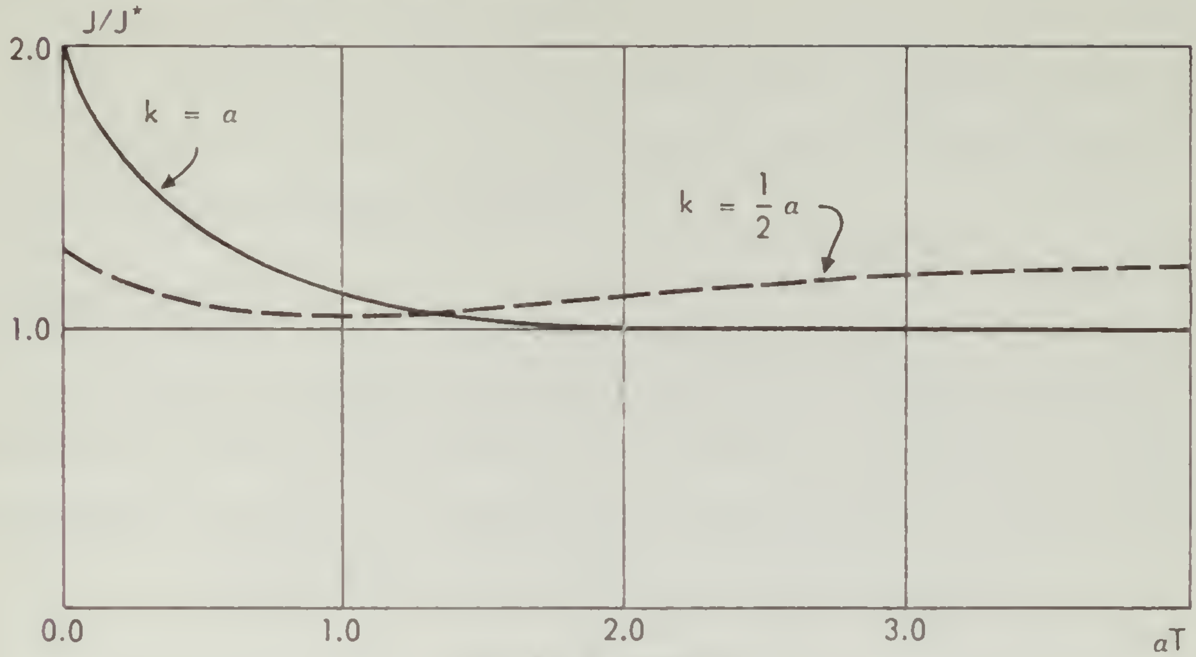


Figure 4.3 - J/J^* versus aT

The maximum value of J/J^* is 2 when T, the control interval, is infinitesimal, and the ratio tends to unity as T increases. In fact, when T is just $1/a$ seconds, the ratio is only 1.13.

To get an indication of the sensitivity of the performance index, we can compute the sub-optimal control with $k = \frac{1}{2}a$ and compare the results with those for $k = a$.

The value of the performance index for $k = \frac{1}{2}a$ is given by

$$J = \frac{\left[q + \frac{1}{4}ra^2\right]}{2a}\left[1 - e^{-aT}\right]\left[c_1^2 + c_2^2 + c_3^2\right] \tag{4.30}$$

or, since $a = \sqrt{q/r}$,

$$J = \frac{5}{8}ra\left[1 - e^{-aT}\right]\left[c_1^2 + c_2^2 + c_3^2\right] \tag{4.31}$$

For this case, the ratio, J/J^*, is

$$J/J^* = \frac{5}{4}\,\frac{\left[1 - e^{-aT}\right]\left[1 + e^{-2aT}\right]}{\left[1 - e^{-2aT}\right]} \tag{4.32}$$

A plot of this ratio is also shown in figure 4.3.

We can see from figure 4.3 that the constant gain sub-optimal control provides a nearly optimal system. The gain setting with $k = a$ would be better if the control interval is much greater than $1/a$ and $k = \frac{1}{2} a$ would be better if the control interval is much less than $1/a$. In any case the system is relatively insensitive to variations in the gain setting, and this is the reason that the optimal control system is little better than the constant gain sub-optimal systems.

Terminal Control

If we desire to reduce the angular velocities of the spinning body to a minimum at the terminal time only, subject to a constraint on the control effort expended, an appropriate performance criterion might be

$$J = \frac{1}{2} q \left[x_1^2(T) + x_2^2(T) + x_3^2(T) \right] + \int_0^T \frac{1}{2} r \left[u_1^2 + u_2^2 + u_3^2 \right] dt \qquad (4.33)$$

The results of the sub-section on optimal control apply directly to this problem if we let

$$q(t) = q u_o(t-T) \qquad (4.34)$$

where $u_o(t)$ is the unit impulse function.

The equation for $p(t)$ then is

$$\dot{p}(t) - p^2(t)/r = 0; \qquad p(T) = q \qquad (4.35)$$

The solution of equation (4.35) is

$$p(t) = \frac{q}{1 + a\tau} \qquad (4.36)$$

where again

$$\tau = T - t \qquad (4.37)$$

and

$$a = q/r \qquad (4.38)$$

Thus the optimal value of the performance index is

$$J^* = \frac{1}{2} \frac{qW(0)}{1 + aT} \qquad (4.39)$$

A plot of p(t) is shown in figure 4.4.

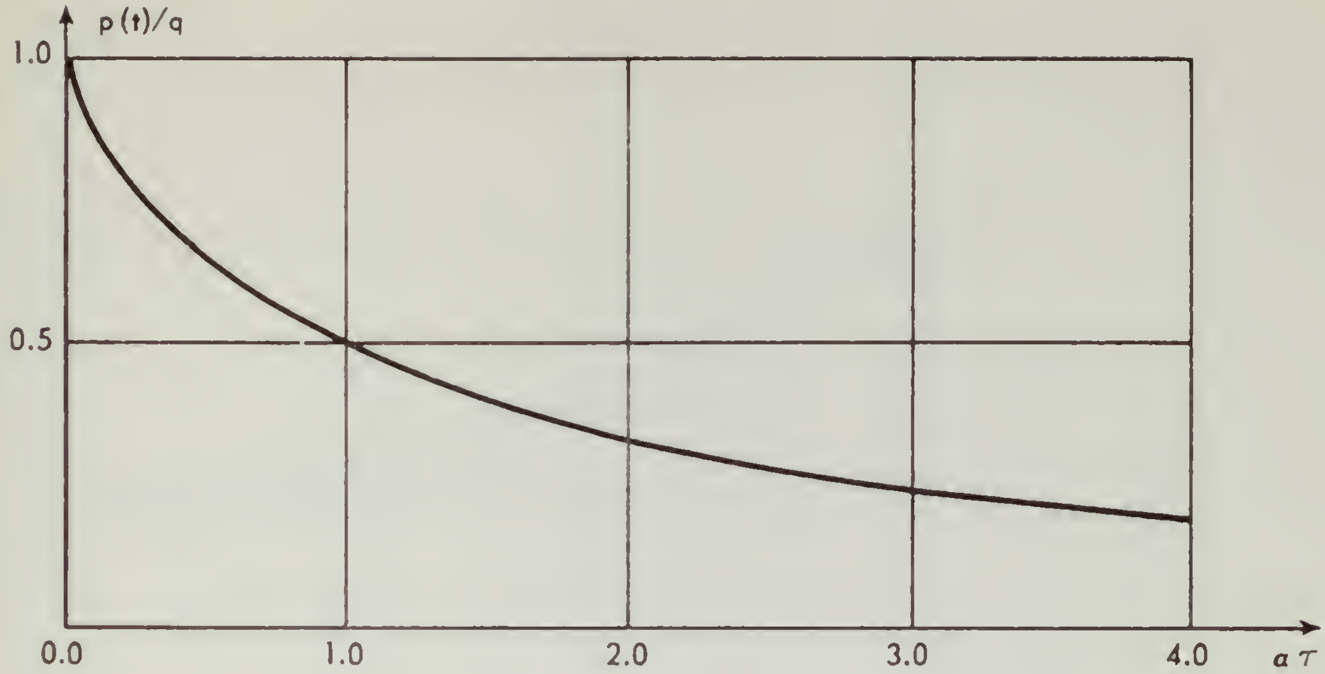


Figure 4.4 - Plot of p(t)/q versus $a\tau$

Again, it is interesting to compare the optimal controller with a sub-optimal constant gain linear controller. In terms of the constant gain, k, the performance criterion for the sub-optimal controller is

$$J = \frac{1}{2} q W(0) \left[e^{-2kT} - \frac{1}{2} rke^{-2kT} / q + \frac{1}{2} rk/q \right] \tag{4.40}$$

If the gain, k, is set equal to a, the sub-optimal performance criterion, J, approaches the optimal performance criterion, J^*, for very short control intervals. In this case, the ratio, J/J^*, is

$$J/J^* = \frac{1}{2}(1 + aT)(e^{-2aT} + 1) \tag{4.41}$$

A plot of J/J^* is shown in figure 4.5.

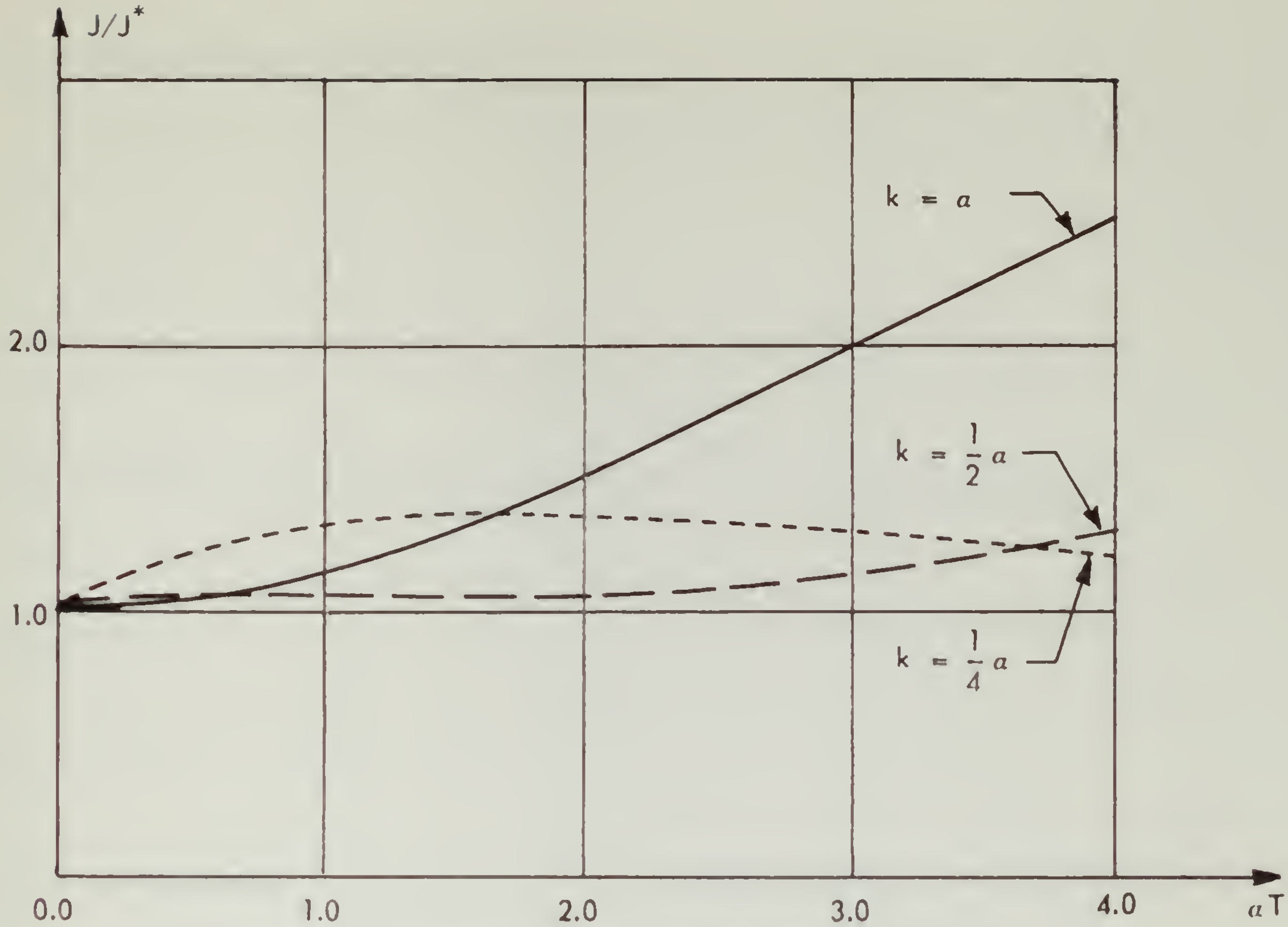


Figure 4.5 - Performance Ratio, J/J^*, for Terminal Control

It should be noted that the value for k in this example was chosen to give near optimal performance over relatively short control intervals. Better performance could be achieved over longer control intervals with a lower gain setting. For instance if $k = \frac{1}{2}\,a$, the value of the performance criterion is

$$J = \frac{1}{2} qW(0) \left[\frac{1}{4} + \frac{3}{4} e^{-aT} \right] \tag{4.42}$$

The ratio, J/J^*, then is

$$J/J^* = \frac{1}{4}(1 + aT)(1 + 3e^{-aT}) \tag{4.43}$$

A plot of this is shown in figure 4.5 also. A plot is also shown for $k = \frac{1}{4}\,a$.

As can be seen from the plot, there exists a constant gain for any particular value of control interval which will give very nearly optimal performance. For instance, with a control interval of $1/a$, $k = \frac{1}{2}\,a$ will give a performance index of about 1.05 times the true optimal performance index.

4.4 Nonlinear Spring Problem

The equation of motion for a mass attached to a cubic spring can be written as

$$\ddot{x} + x^3 = 0 \tag{4.44}$$

or if control is exerted, i.e., the system is forced, the equation is

$$\ddot{x} + x^3 = u \tag{4.45}$$

This equation can be written as the system of first order equations

$$\left.\begin{aligned} \dot{x}_1 &= x_2 + u_1 \\ \dot{x}_2 &= -x_1^3 + u_2 \end{aligned}\right\} \tag{4.46}$$

where

$$x_1 = x \tag{4.47}$$

and

$$u = \dot{u}_1 + u_2 \tag{4.48}$$

The state variable, x_2, is not as easily identified with the original system variables, but this is of little consequence.

Suppose that we wish to control the system (4.46) such that

$$J = \int_0^T \left[q\left(\frac{1}{4} x_1^4 + \frac{1}{2} x_2^2\right) + r\left(x_1^2 u_1^2 + \frac{1}{2} u_2^2\right)\right] dt \tag{4.49}$$

is a minimum. The loss equation for this system is

$$V_t + \underset{u_1, u_2}{\text{Min}} \left[q\left(\frac{1}{4} x_1^4 + \frac{1}{2} x_2^2\right) + r\left(x_1^2 u_1^2 + \frac{1}{2} u_2^2\right) + V_{x_1}(x_2 + u_1) + V_{x_2}\left(-x_1^3 + u_2\right)\right] = 0 \tag{4.50}$$

If we assume

$$V = p(t)\left(\frac{1}{4} x_1^4 + \frac{1}{2} x_2^2\right) \tag{4.51}$$

then the optimal control is

$$\left.\begin{aligned} u_1 &= -\frac{1}{2} p(t) x_1 / r(t) \\ u_2 &= -p(t) x_2 / r(t) \end{aligned}\right\} \tag{4.52}$$

and equation (4.50) becomes

$$\left[\dot{p}(t) - p^2(t)/r(t) + q(t)\right] \left[\frac{1}{4} x_1^4 + \frac{1}{2} x_2^2\right] = 0 \tag{4.53}$$

Since this equation must be satisfied for all values of x_1 and x_2, we must have

$$\dot{p}(t) - p^2(t)/r(t) + q(t) = 0 \tag{4.54}$$

The boundary condition is

$$p(T) = 0 \tag{4.55}$$

This equation is identical to equation (4. 12), and the results of section 4. 3 of this chapter, including the sub-optimal control results, are equally applicable to this problem.

Since

$$u = \dot{u}_1 + u_2 \tag{4.56}$$

the control, u, may be expressed as

$$u = -\frac{1}{2} p(t)\dot{x}_1/r(t) - p(t)x_2/r(t) \tag{4.57}$$

For the actual synthesis of the controller, however, this expression for u is unsatisfactory because the state variable, x_2, has not been identified with the original system variables. We can get around this by expressing x_2 in terms of x_1 and u_1, thus

$$x_2 = \dot{x}_1 - u_1 \tag{4.58}$$

or

$$x_2 = \dot{x}_1 + \frac{1}{2} p(t) x_1/r(t) \tag{4.59}$$

The control, u, then is

$$u = 3p(t)\dot{x}_1/2r(t) + \frac{1}{2} p^2(t) x_1/r^2(t) \tag{4.60}$$

The block diagram for this control system is shown in figure 4. 6.

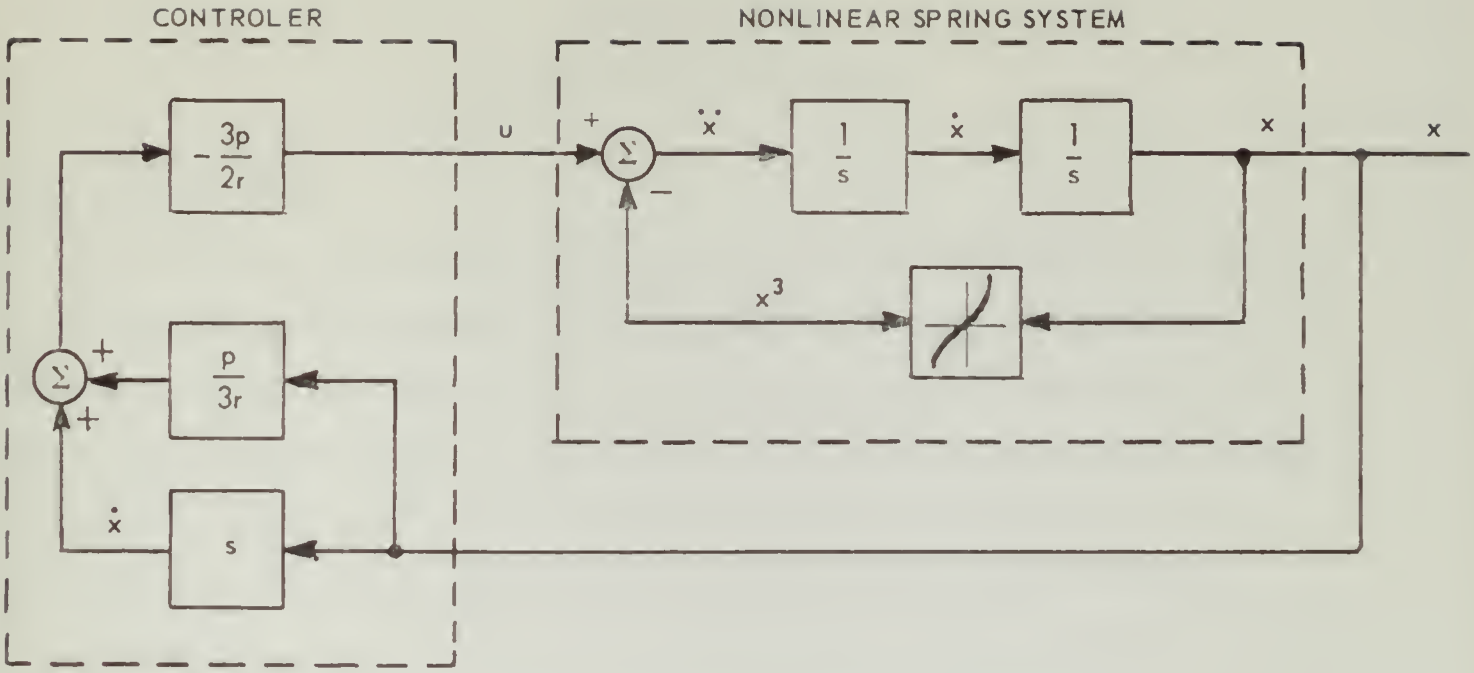


Figure 4.6 - Nonlinear Spring Control System Block Diagram

Admittedly, the nonlinear systems and the performance criteria used in this example problem and the previous one are very special. However, because we are able to obtain analytical solutions, a great deal of insight can be gained from them about the nature and behavior of optimal controllers in nonlinear systems. In particular, we have found that simple constant gain linear controllers can provide very near optimal performance over a wide range of conditions.

CHAPTER V

COMPUTER RESULTS

5.1 Introduction

The results of several computer problems illustrating the methods of Chapter II are presented in this chapter. Several variations of each problem are presented in order to show the effect of changes in the initial state and changes in the performance index. It should be borne in mind that since the system being controlled is nonlinear, the controller parameters depend on the initial state of the system.

In addition, the results of controlling some of the nonlinear systems with simple sub-optimal linear controllers are presented and compared with the optimal results.

The results of this section were obtained on the IBM 7090 computer at the MIT computation center. The Fortran programs used to obtain the solutions for the two state-variable deterministic problems are given in Appendix D. In all cases, the change in the performance index from one iteration to the next was used as a convergence criterion. When the magnitude of this change was less than one per cent of the value of the performance index, the iteration procedure was terminated.

5.2 One State-Variable Example

The system considered for this example can be described by the equations

$$x(k+1) = x(k) - 0.05\,x^3(k) + 0.05\,u(k); \qquad x(1) = c \tag{5.1}$$

$$y(k) = x(k) \tag{5.2}$$

The system may be thought of as the discrete time approximation of the continuous time system

$$\dot{x}(t) = -x^3(t) + u(t); \qquad x(0) = c \tag{5.3}$$

$$y(t) = x(t) \tag{5.4}$$

The performance index used was

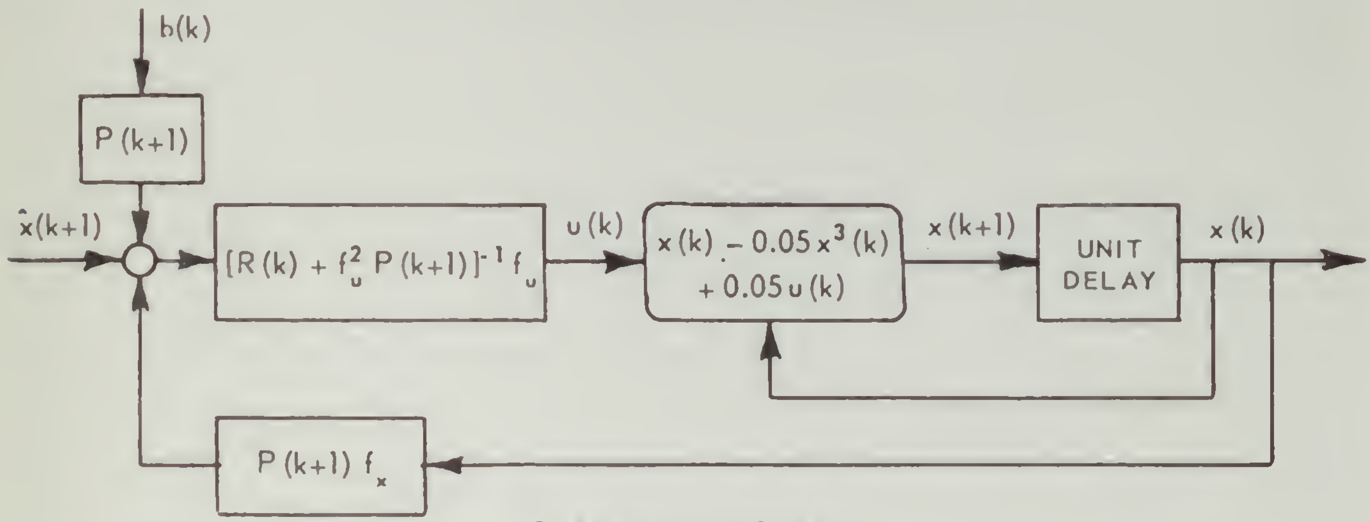
$$J = \sum_{k=1}^{100} \frac{1}{2} Q(k) [z(k) - x(k)]^2 + \sum_{k=1}^{99} \frac{1}{2} R(k) u^2(k) \qquad (5.5)$$

The equations used as the basis of the iterative procedure for this problem may be determined from equations (2. 23), (2. 27), (2. 31), and (2. 32).

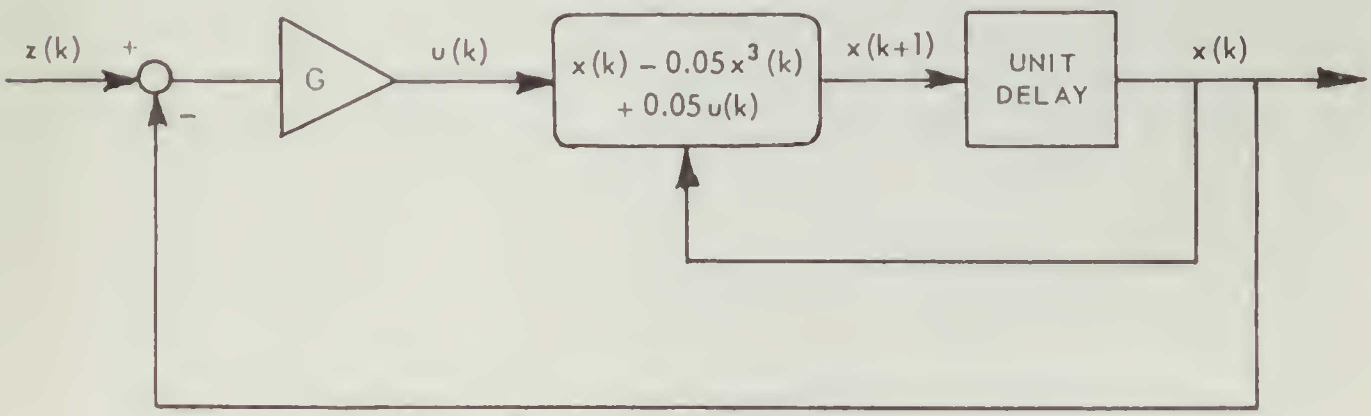
The sub-optimal system used is given by the same equations except that u (k) is given by

$$u(k) = G(k) [z(k) - x(k)] \qquad (5.6)$$

where G is a constant gain factor. Block diagrams of the optimal and the sub-optimal control systems are given in figure 5. 1.



Optimal Control System



Sub-optimal Control System

Figure 5. 1 - One State-Variable Control Systems

Figures 5. 2 through 5. 10 give the plotted results from several data sets for this system. Comments on each of the figures are given below.

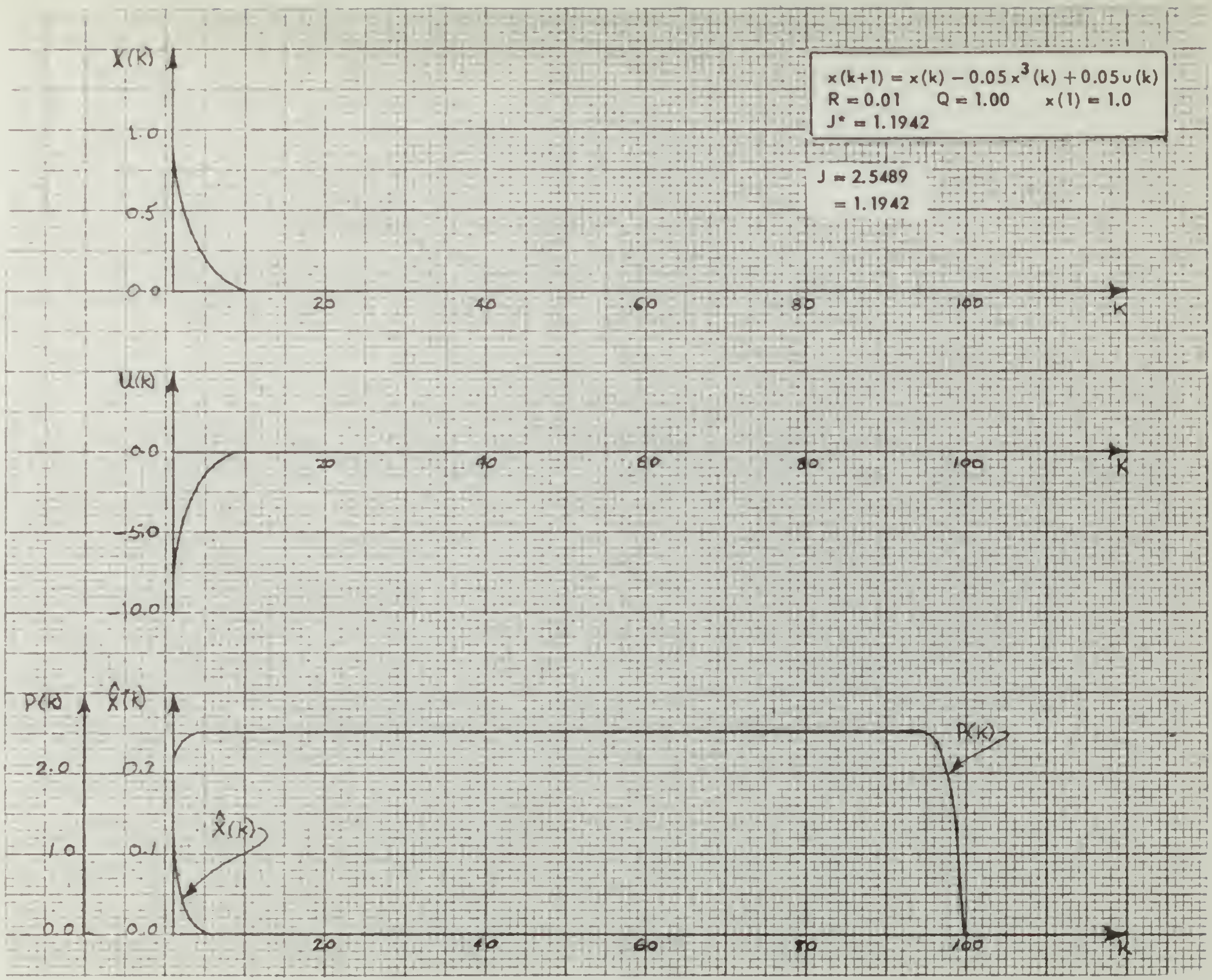

x(k+1) = x(k) − 0.05 x³(k) + 0.05 u(k)
R = 0.01 Q = 1.00 x(1) = 1.0
J* = 1.1942
J = 2.5489
= 1.1942
X(k)
1.0
0.5
0.0
20
40
60
80
100
K
U(k)
-5.0
-10.0
P(k)
X̂(k)
2.0
1.0
0.2
0.1

Figure 5.2

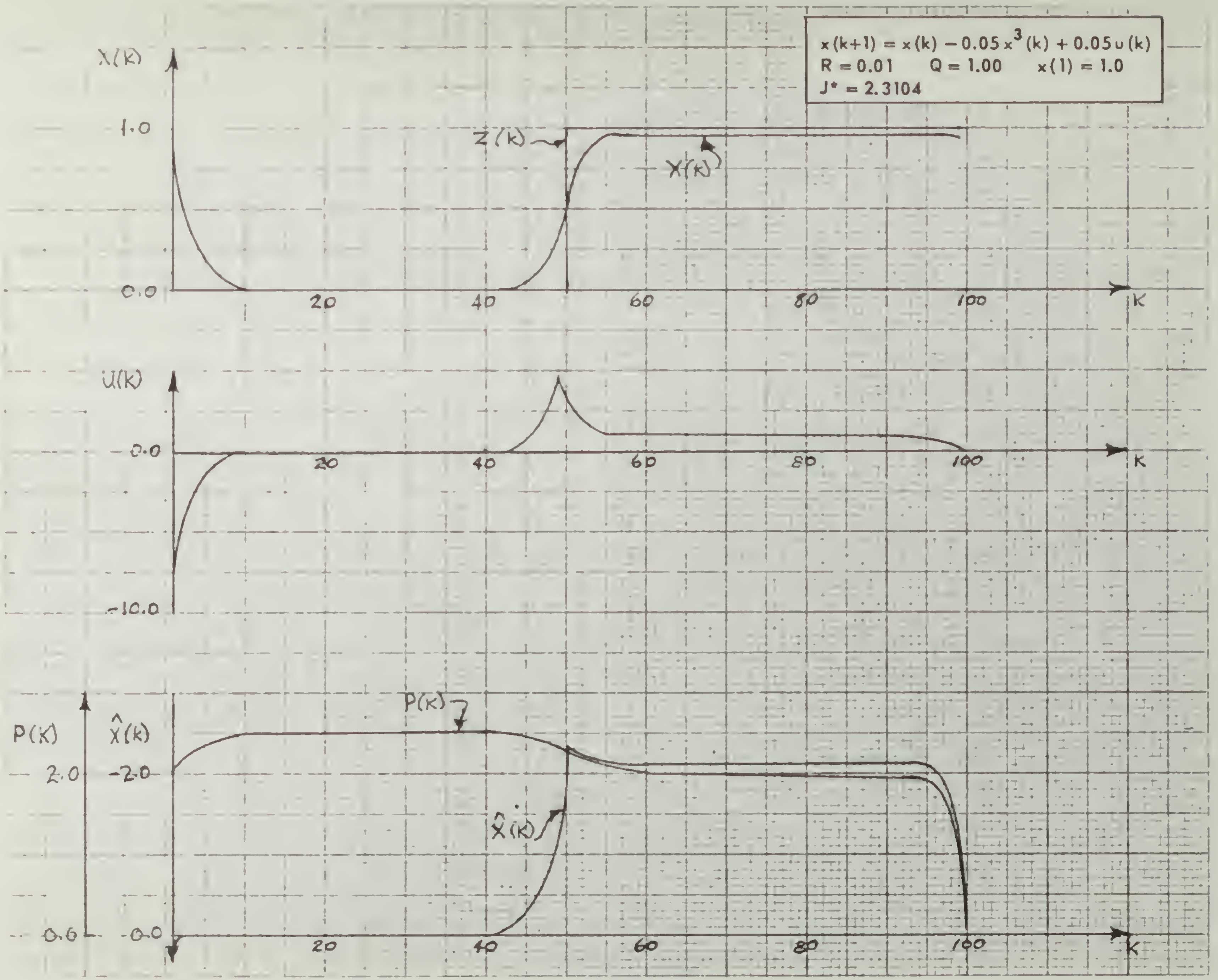
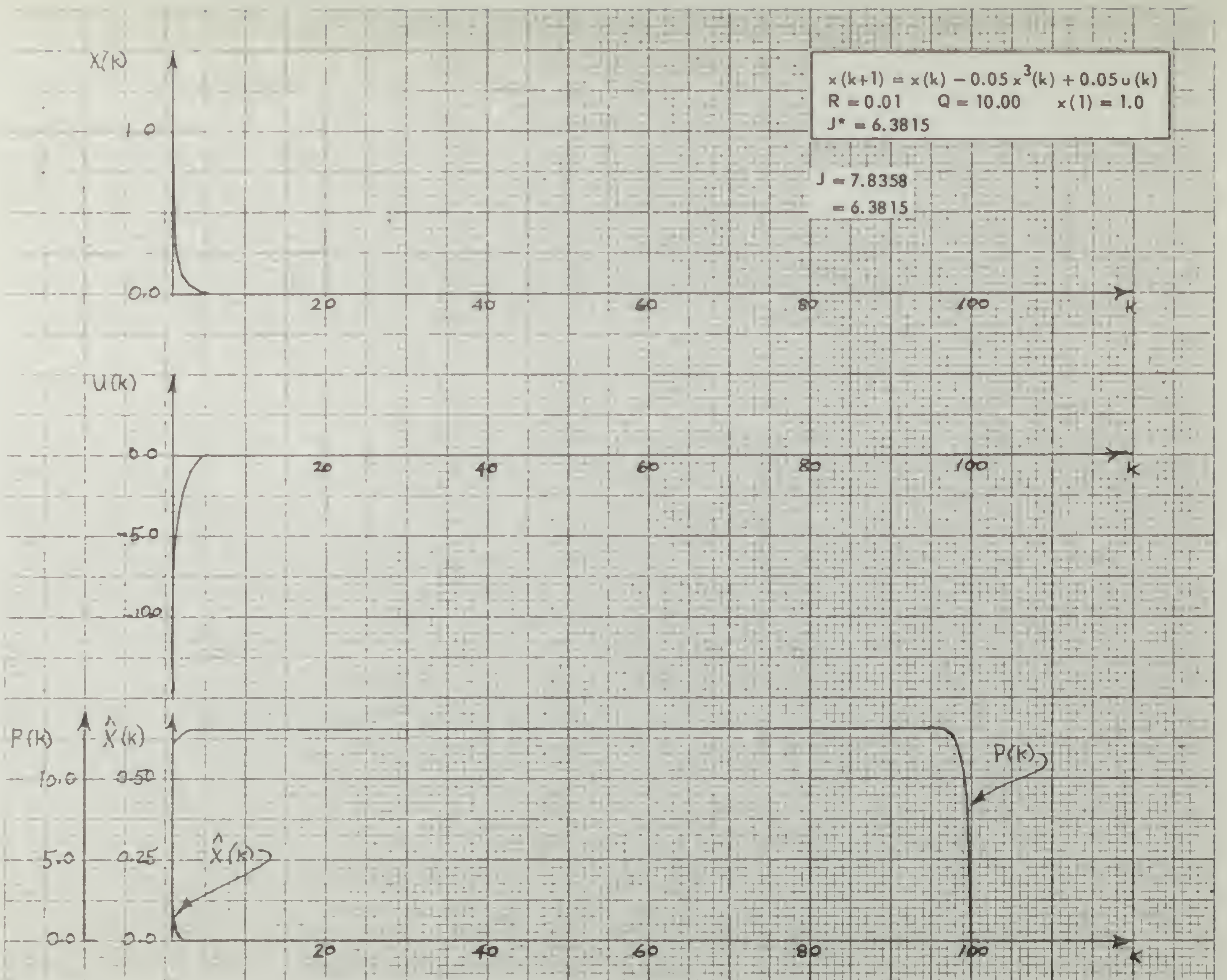
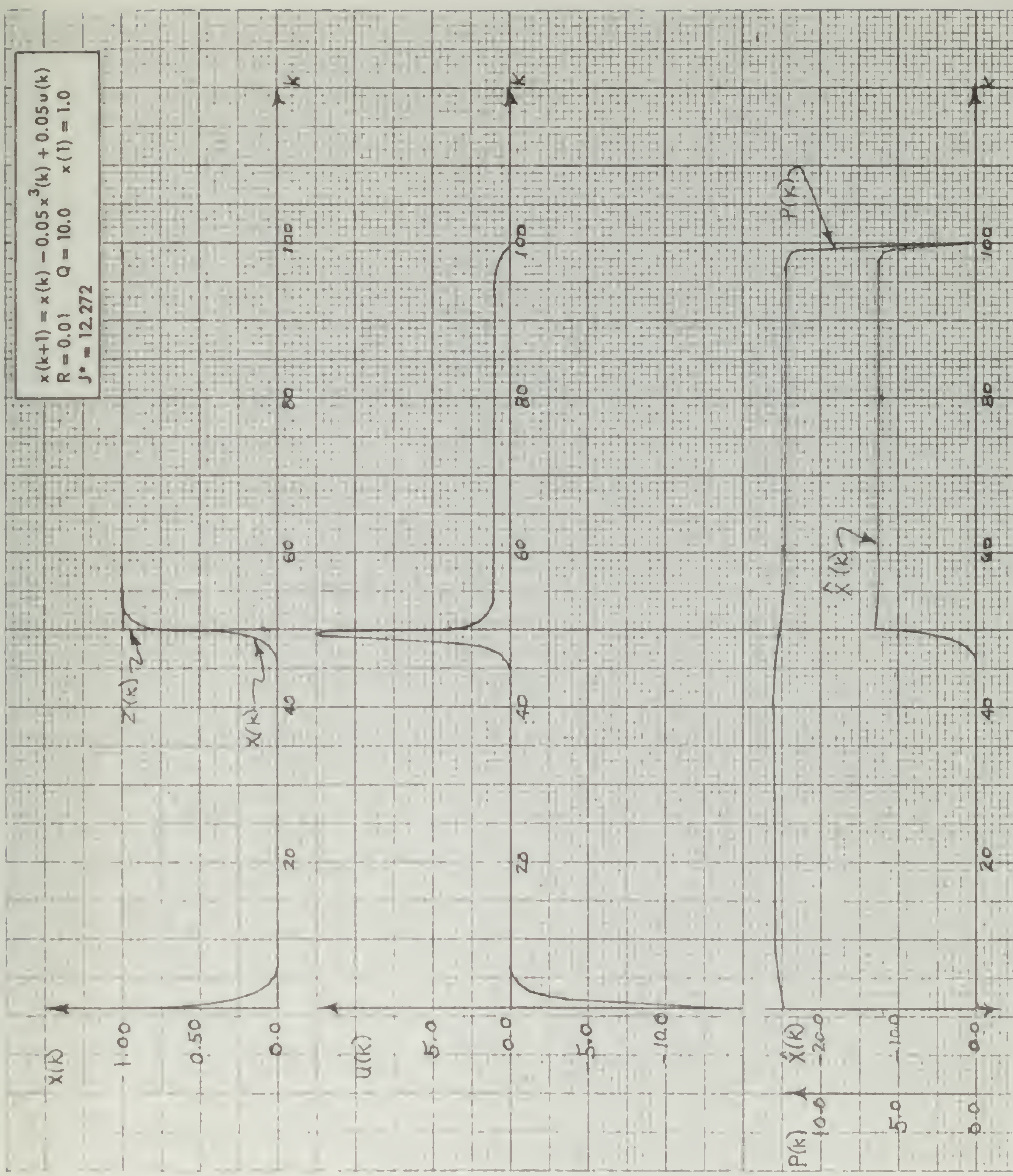

x(k+1) = x(k) − 0.05x3(k) + 0.05u(k)
R = 0.01 Q = 1.00 x(1) = 1.0
J* = 2.3104
X(K)
1.0
0.0
Z(K)
X(K)
U(K)
0.0
-10.0
P(K)
X̂(K)
2.0
-2.0
0.0
X̂̇(K)
20
40
60
80
100
K

Figure 5.3



x(k+1) = x(k) − 0.05 x³(k) + 0.05 u(k)
R = 0.01 Q = 10.00 x(1) = 1.0
J* = 6.3815
J = 7.8358
= 6.3815
X(k)
1.0
0.0
U(k)
-5.0
-10.0
P(k)
X̂(k)
10.0
0.50
5.0
0.25
20
40
60
80
100
k

Figure 5.4



x(k+1) = x(k) − 0.05 x³(k) + 0.05 u(k)
R = 0.01 Q = 10.0 x(1) = 1.0
J* = 12.272
X(k)
1.00
0.50
0.0
Z(k)
X(k)
U(k)
5.0
0.0
-5.0
-10.0
P(k)
X̂(k)
10.0
5.0
0.0
-20.0
-10.0
0.0
20
40
60
80
100
k

Figure 5.5

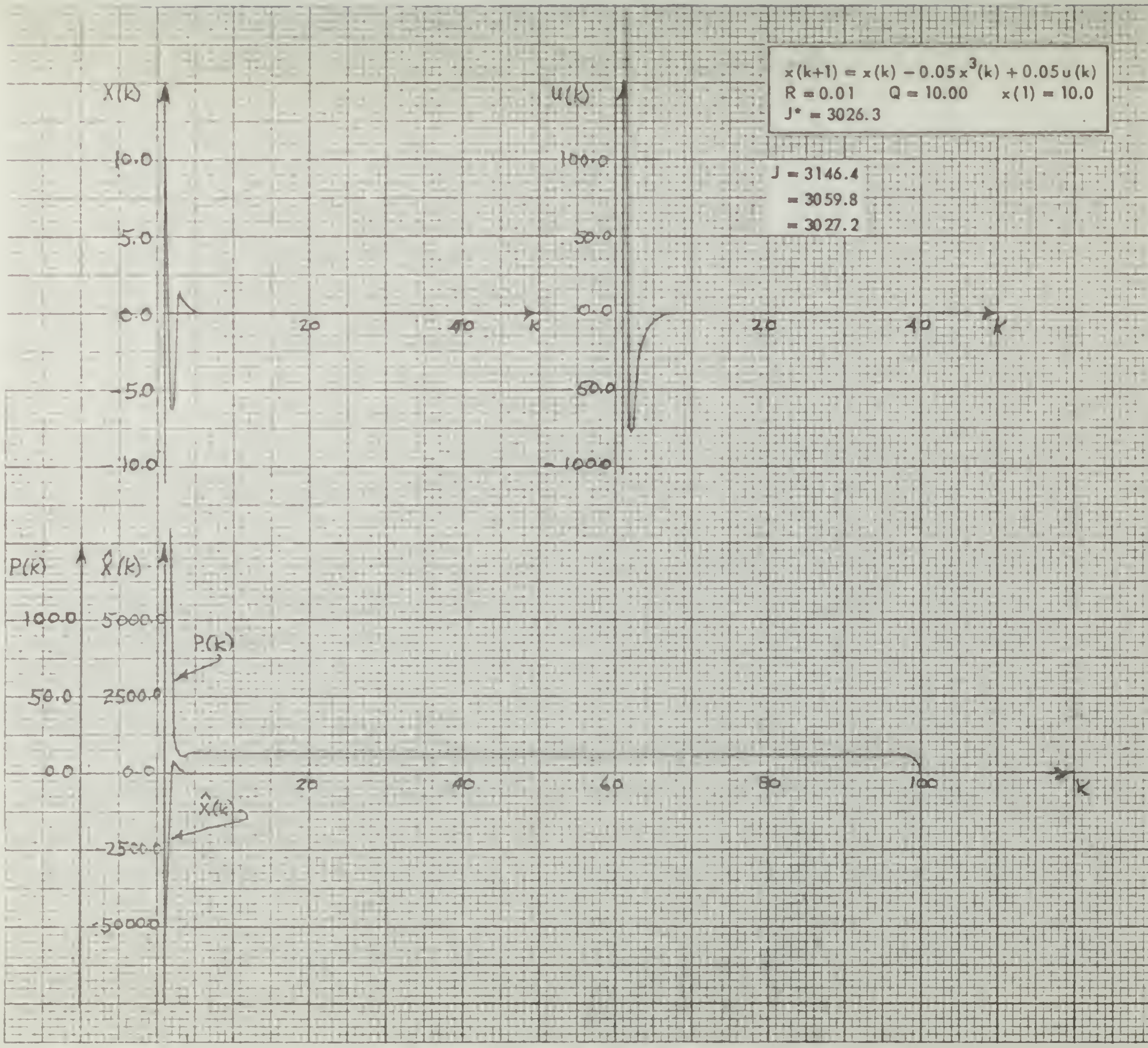

x(k+1) = x(k) − 0.05 x³(k) + 0.05 u(k)
R = 0.01 Q = 10.00 x(1) = 10.0
J* = 3026.3
J = 3146.4
= 3059.8
= 3027.2
X(k)
10.0
5.0
0.0
-5.0
-10.0
20
40
k
u(k)
100.0
50.0
0.0
-50.0
-100.0
20
40
k
P(k)
X̂(k)
100.0
50.0
0.0
5000.0
2500.0
0.0
-2500.0
-5000.0
P(k)
X̂(k)
20
40
60
80
100
k

Figure 5.6

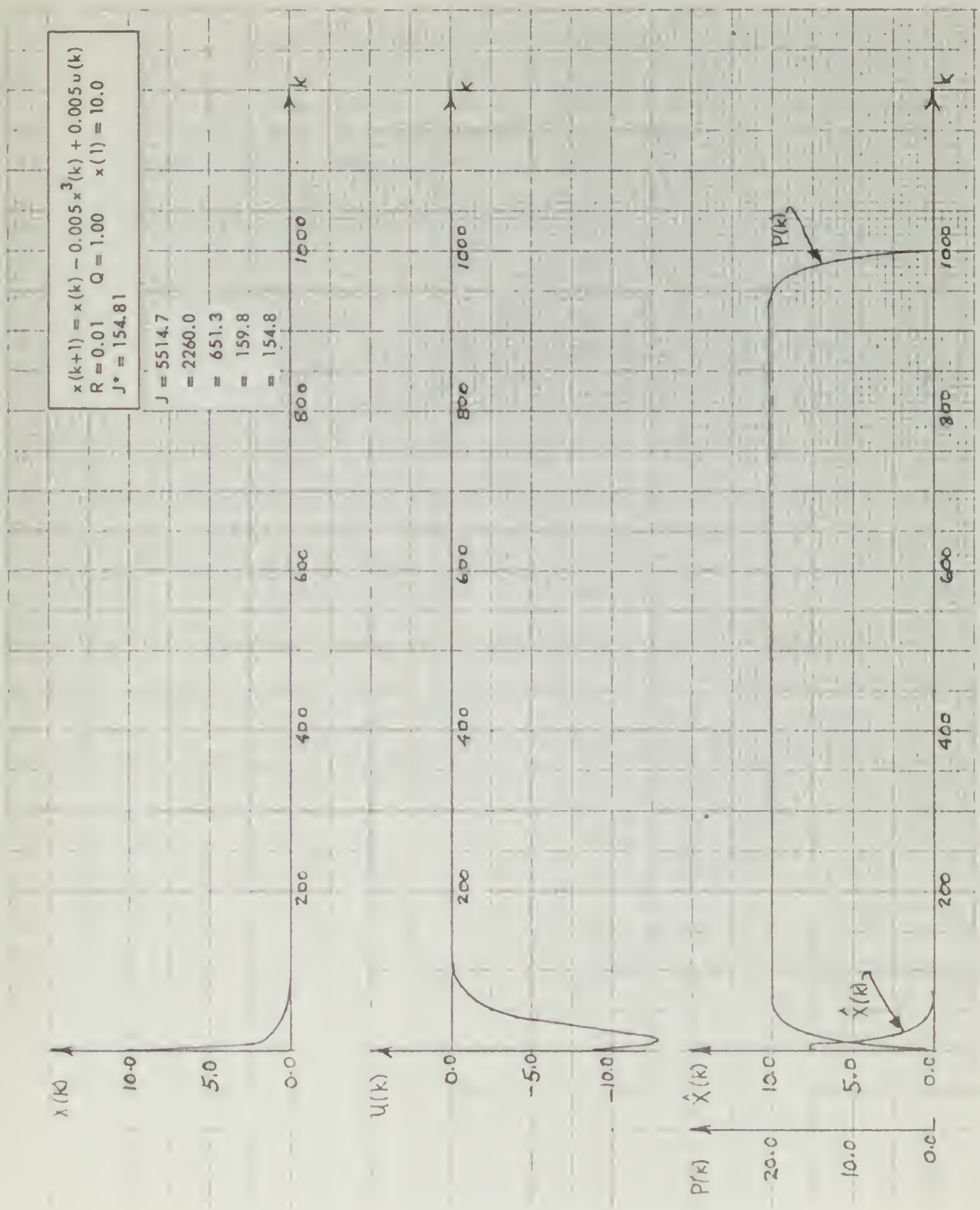

x(k+1) = x(k) − 0.005 x³(k) + 0.005 u(k)
R = 0.01   Q = 1.00   x(1) = 10.0
J* = 154.81
J = 5514.7
= 2260.0
= 651.3
= 159.8
= 154.8
λ(k)
10.0
5.0
0.0
U(k)
0.0
-5.0
-10.0
P(k)
X̂(k)
20.0
10.0
5.0
0.0
P(k)
X̂(k)
200
400
600
800
1000
k

Figure 5.7

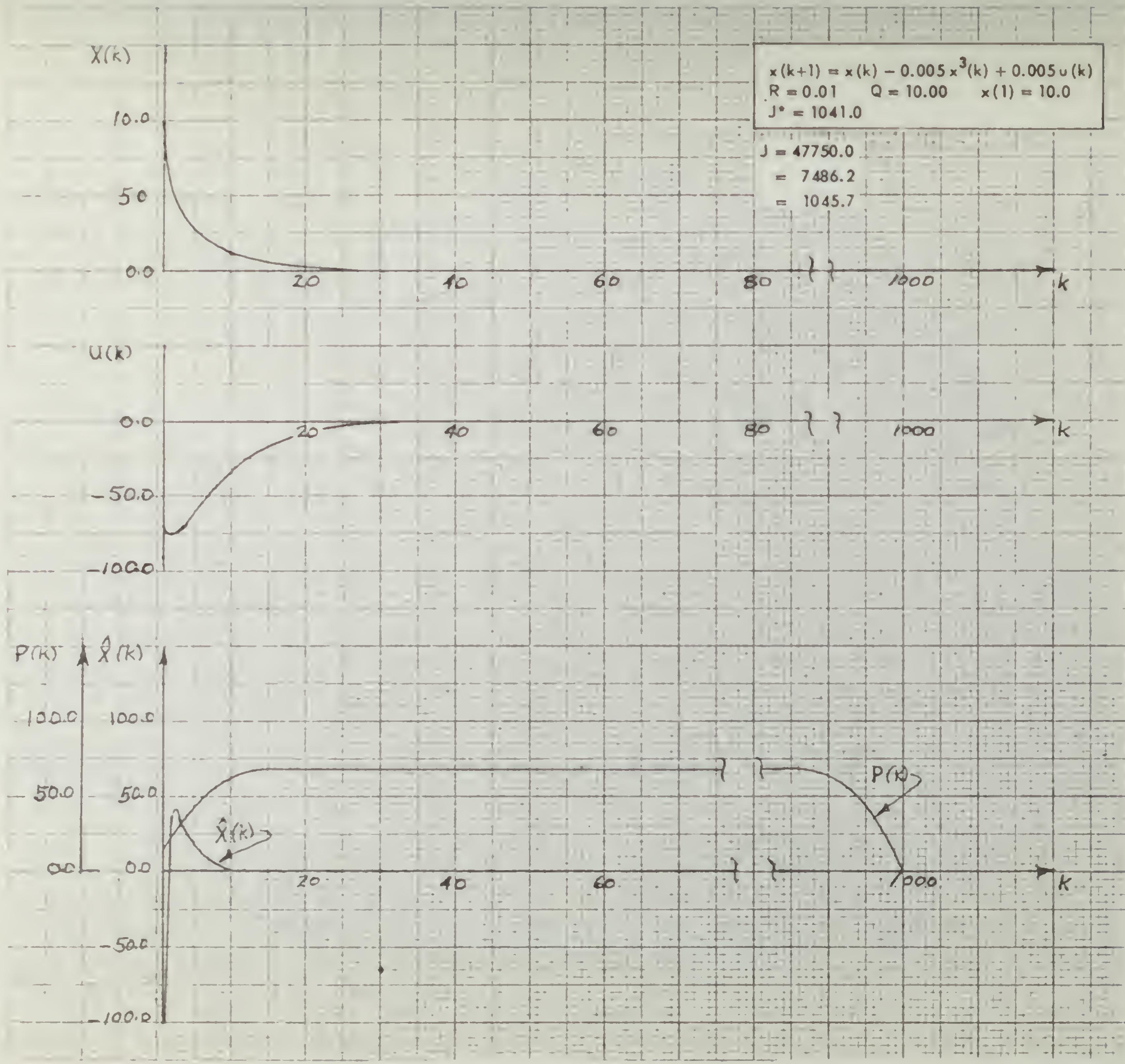
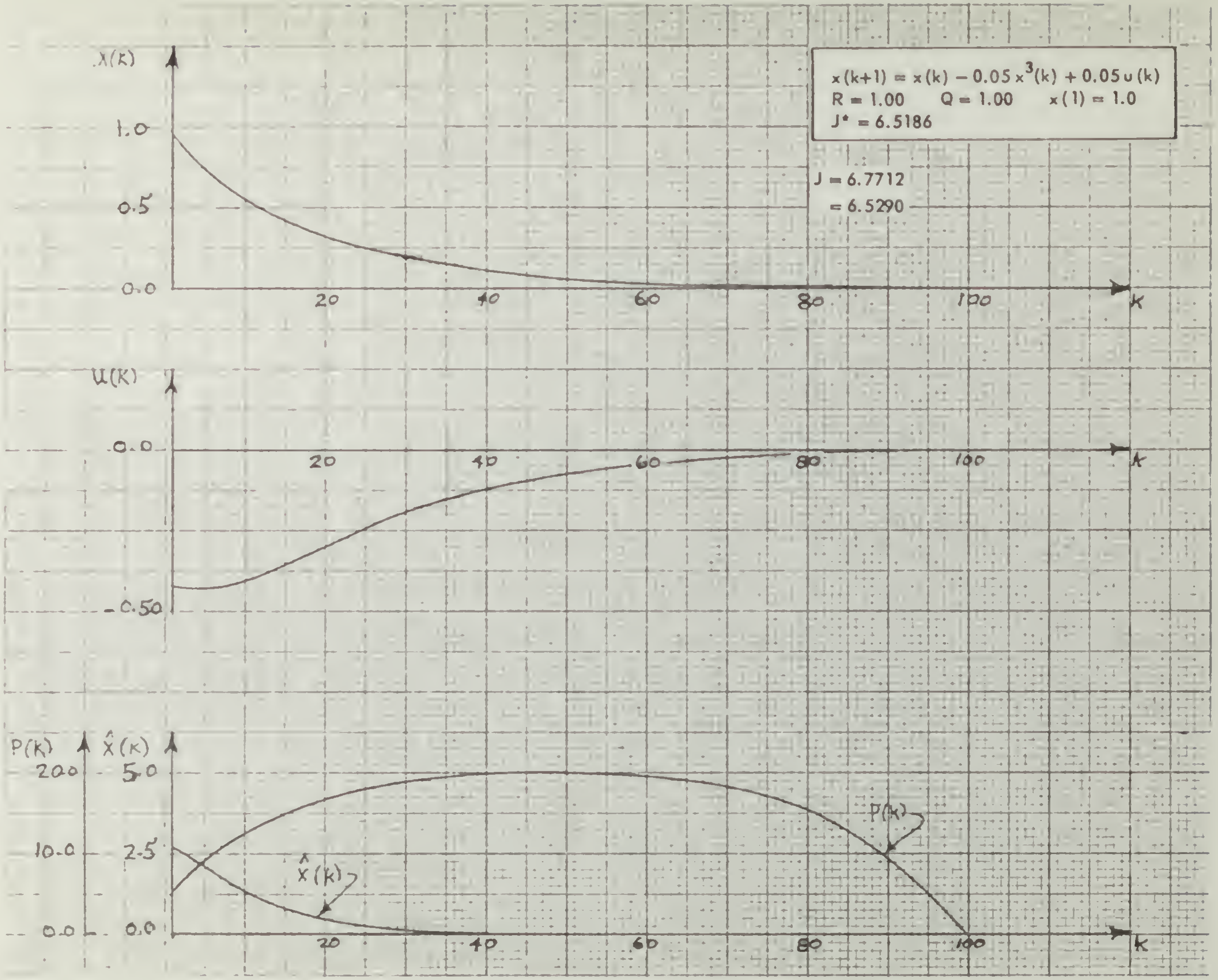

x(k+1) = x(k) − 0.005 x³(k) + 0.005 u(k)
R = 0.01 Q = 10.00 x(1) = 10.0
J* = 1041.0
J = 47750.0
= 7486.2
= 1045.7
X(k)
U(k)
P(k)
X̂(k)
k

Figure 5.8



x(k+1) = x(k) − 0.05 x³(k) + 0.05 u(k)
R = 1.00 Q = 1.00 x(1) = 1.0
J* = 6.5186
J = 6.7712
= 6.5290
X(K)
1.0
0.5
0.0
U(K)
0.0
-0.50
P(K)
X̂(K)
200
5.0
10.0
2.5
0.0
0.0
20
40
60
80
100
K
P(k)
X̂(k)

Figure 5. 9

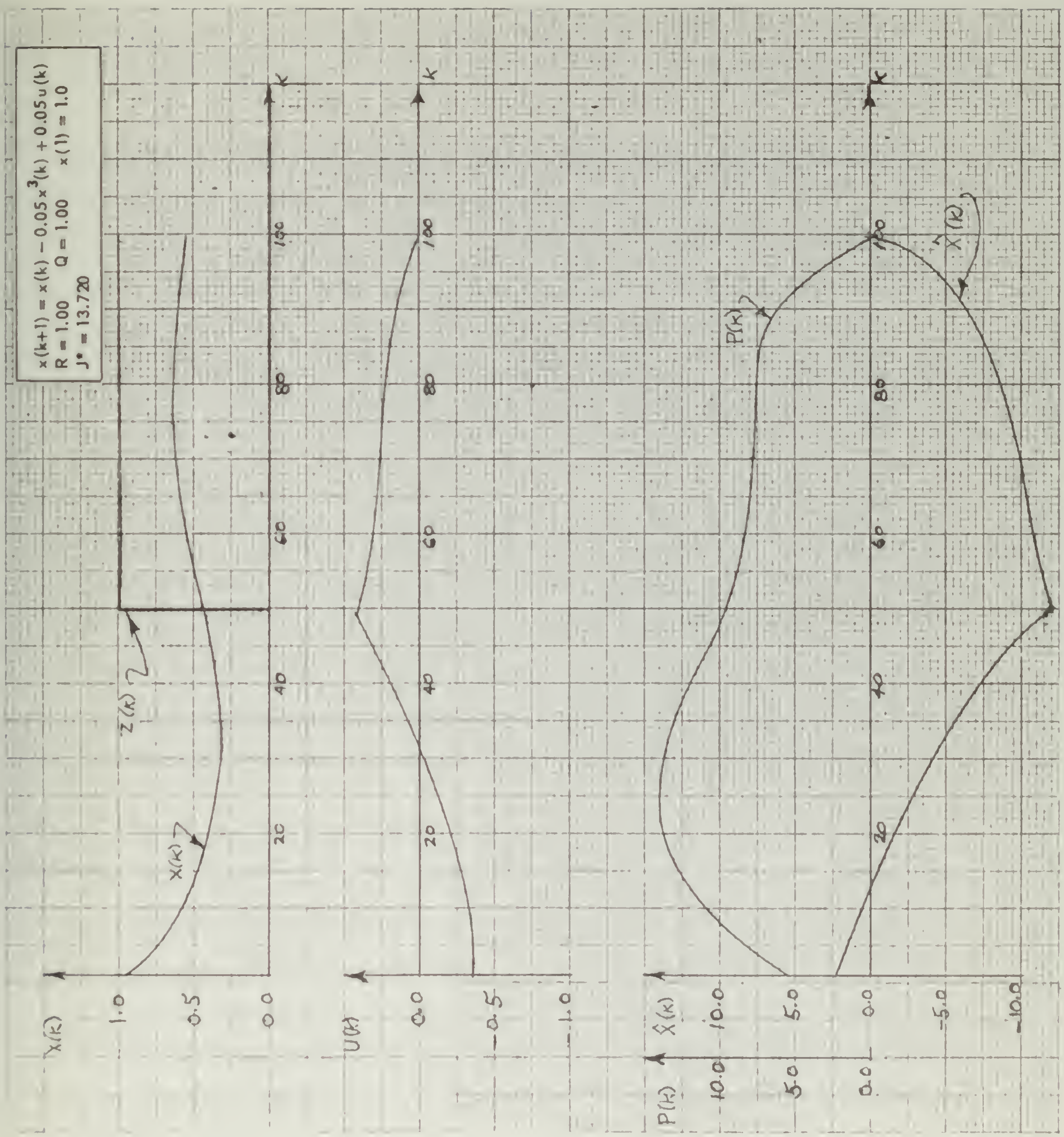


Figure 5.10

Figure 5.2: For this data set, $R(k) = 0.01$, $Q(k) = 1.00$, $x(1) = 1.00$, and $z(k) = 0.0$. Convergence was achieved in three iterations. The linear sub-optimal control system with a gain equal to 7.5 gave a performance index of 1.1951, just 0.1 per cent higher than the optimal.

Figure 5.3: For this data set, $R(k)$, $Q(k)$, $x(1)$, and $z(k)$ are the same as for the previous data set except that $z(k) = 1.0$ for $k \geq 50$. Convergence occurred in three iterations. The plot clearly shows that $u(k)$ anticipates the step in $z(k)$ indicating the sense in which this control system is "unrealizable." The sub-optimal control system, which is non-anticipative, with a gain of 7.5 had a performance index of 3.238, about 30 per cent higher than the anticipative optimal system.

Figure 5.4: For this data set, $R(k) = 0.01$, $Q(k) = 10.0$, $x(1) = 1.0$, and $z(k) = 0.0$. Convergence was achieved in three iterations. Notice that since output error is relatively more important in this case, the control effort used is higher and the system response is faster. The sub-optimal control for this set had a gain of 15.0 and gave a performance index of 6.386, about 0.1 per cent higher than the optimal.

Figure 5.5: For this data set, $R(k) = 0.01$, $Q(k) = 10.0$, $x(1) = 1.0$, and $z(k) = 0.0$ for $k < 50$, but $z(k) = 1.0$ for $k \geq 50$. Convergence was achieved in two iterations. With a gain of 15.0, the sub-optimal system gave a performance index of 13.845, which is 13 per cent higher than the performance index for the optimal system. When the control system response is relatively fast, as in this case, anticipation of the optimal system does not improve the system performance as much.

Figure 5.6: For this data set, $R(k) = 0.01$, $Q(k) = 10.0$, $x(1) = 10.0$, and $z(k) = 0.0$. When the initial state is as large as it is in this case, the system is open loop unstable. (The continuous time system, $\dot{x} + x^3 = 0$, is always stable, but the sampling introduced to make the discrete time approximation causes the system to be unstable for $|x(1)|$ greater than about 6.0. The closed loop control system is, nevertheless, stable, at the expense of a very large performance index. Because the discrete time system is unstable, it is not a good representation of the continuous time system for this case. For this reason figures 5.7 and 5.8 have been included.

Figure 5.7: For this data set, the sampling interval has been decreased by a factor of 10 and the number of steps has been increased by a factor of 10. This makes the system open loop stable, and once again a reasonable discrete time approximation to the continuous time system. Here $R(k) = 0.01$, $Q(k) = 1.0$, $x(1) = 10.0$, and $z(k) = 0.0$. Convergence occurred in six iterations.

Figure 5.8: For this data set, the comments of the previous set apply except that $Q(k) = 10.0$. Convergence occurred in four iterations.

Figure 5.9: For this data set, $R(k) = 1.0$, $Q(k) = 1.0$, $x(1) = 1.0$, and $z(k) = 0.0$. Convergence occurred in three iterations. Because the cost of control is so high relative to the cost of output error, the control effort expended is small and the system response is slow. As a matter of fact, it can be shown that for the one state-variable system the speed of response is proportional to the ratio, $Q(k)/R(k)$. In general, we can expect the speed of response to depend on the ratio of the norm of the $\underline{Q}(k)$ matrix to the $\underline{R}(k)$ matrix.

Figure 5.10: The data for this set is the same as for the last set except that z(k) = 0.0 for k < 50 and z(k) = 1.0 for k $\geq$ 50. Convergence occurred in four iterations.

5.3 Two State-Variable Examples

The system considered for the first two state-variable example can be described by the equations

$$x_1(k+1) = x_1(k) + 0.01\, x_2(k); \qquad x_1(1) = c_1 \tag{5.7}$$

$$x_2(k+1) = x_2(k) - 0.02\, x_1(k) - 0.03\, |x_2(k)|\, x_2(k) + 0.01\, u(k); \qquad x_2(1) = c_2 \tag{5.8}$$

$$y_1(k) = x_1(k) \tag{5.9}$$

$$y_2(k) = x_2(k) \tag{5.10}$$

A block diagram of this system is shown in figure 5.11. The system

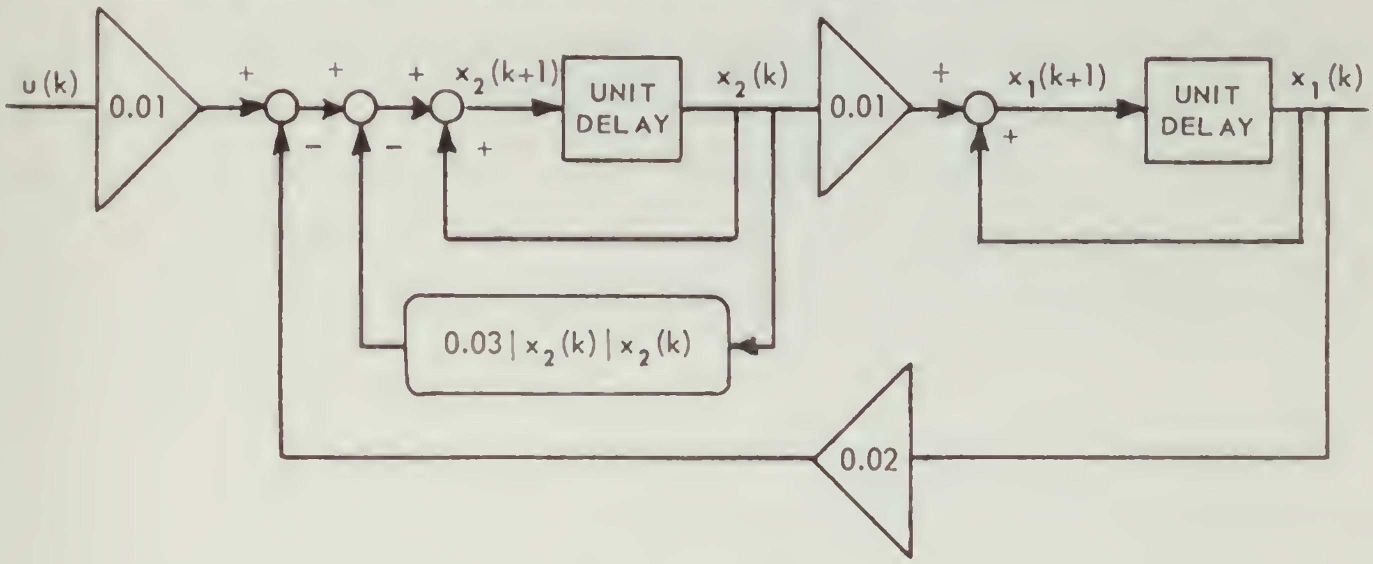


Figure 5.11 - Two State-Variable Nonlinear System

described above may be thought of as the discrete approximation for the continuous time system

$$\ddot{x}(t) + 3\, |\dot{x}(t)|\, \dot{x}(t) + 2x(t) = u(t) \tag{5.11}$$

$$y_1(t) = x(t) \tag{5.12}$$

$$y_2(t) = \dot{x}(t) \tag{5.13}$$

The performance index used was

$$J = \sum_{k=1}^{100} \frac{1}{2} \{Q_1(k)\, [z_1(k) - x_1(k)]^2 + Q_2(k)\, [z_2(k) - x_2(k)]^2\} + \sum_{k=1}^{99} \frac{1}{2} R(k) u^2(k) \tag{5.14}$$

The equations that form the basis of the iterative routine follow from equations (2.23), (2.27), (2.31), and (2.32). The equations for the sub-optimal systems are the same except that

$$u(k) = G_1 [z_1(k) - x_1(k)] + G_2 [z_2(k) - x_2(k)] \tag{5.15}$$

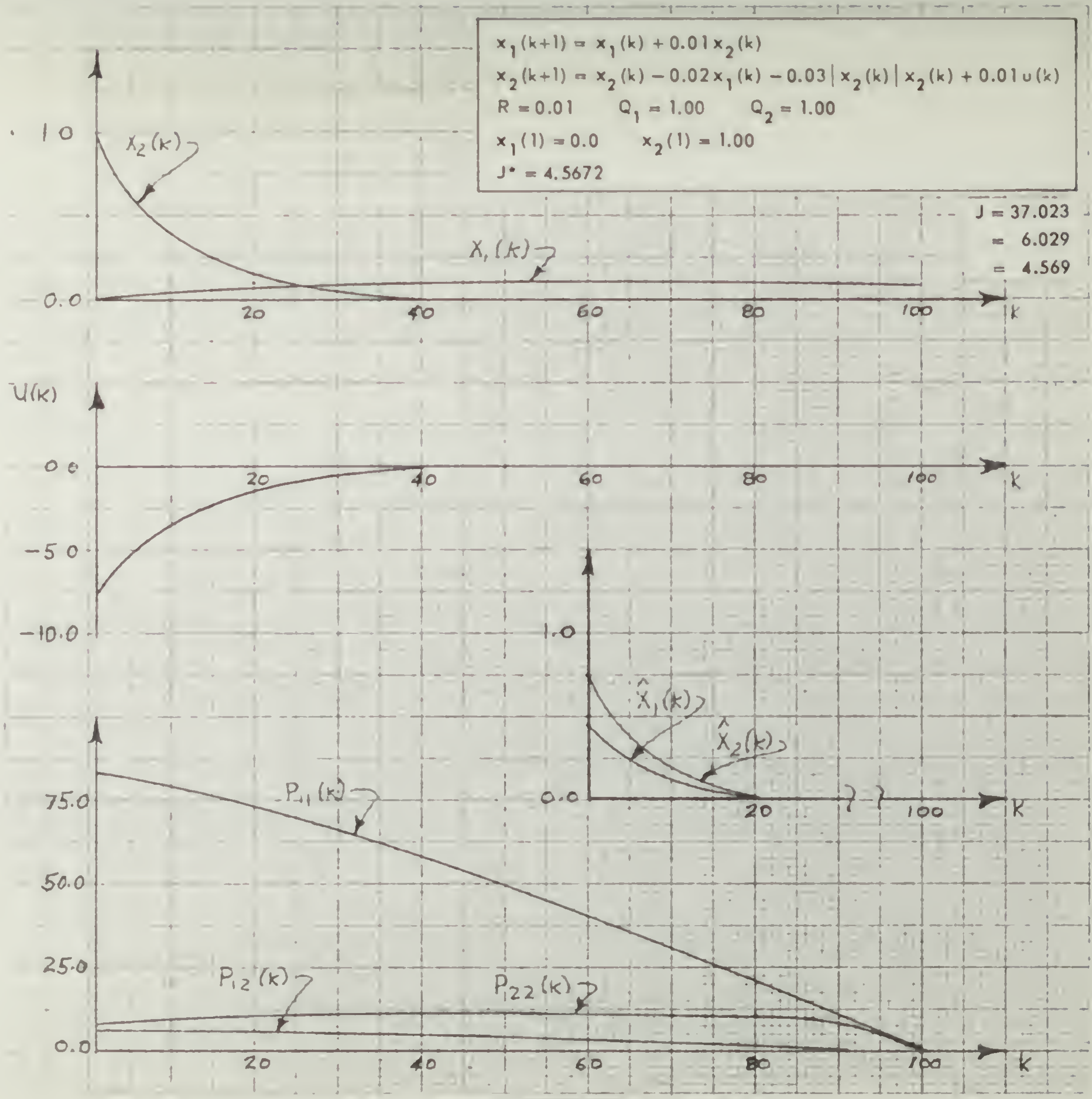
where G_1 and G_2 are constant gain factors.

Figures 5.12 through 5.21 give the plotted results from ten data sets for this example. Comments on these figures follow.

Figure 5.12 - 5.14: For these data sets, $R(k) = 0.01$, $Q_1(k) = 1.00$, $Q_2(k) = 1.00$, $x_1(1) = 0.0$, $z_1(k) = 0.0$, and $z_2(k) = 0.0$. In figure 5.12, $x_2(1) = 1.0$, in figure 5.13, $x_2(1) = 3.0$, and in figure 5.14, $x_2(1) = 10.0$. For each of these convergence occurred in three or four iterations. The sub-optimal control with $G_1 = 8.5$ and $G_2 = 4.75$ gave performance indices of 5.360, 31.32, and 158.62 for $x_2(1) = 1.0$, $= 3.0$, and $= 10.0$, respectively. The sub-optimal control system performance indices were 17.5, 7.0, and 7.5 percent higher than the optimal performance indices.

Figures 5.15 - 5.17: For these figures, the data were the same as for figures 5.12 - 5.14 except that $z_1(k) = 0.0$ for $k < 50$ and $z_1(k) = 1.0$ for $k \geq 50$. In each case convergence occurred in three iterations.

Figure 5.18: For this data set, $R(k) = 0.01$, $Q_1(k) = 1.0$, $Q_2(k) = 0.0$, $x_1(1) = 0.0$, $x_2(1) = 1.0$, $z_1(k) = 0.0$, and $z_2(k) = 0.0$. Convergence occurred in five iterations. The sub-optimal control system with $G_1 = 11.0$ and $G_2 = 2.0$ gave a performance index of 1.305, about 5 per cent higher than the optimal system performance index.



x1(k+1) = x1(k) + 0.01 x2(k)
x2(k+1) = x2(k) − 0.02 x1(k) − 0.03 | x2(k) | x2(k) + 0.01 u(k)
R = 0.01 Q1 = 1.00 Q2 = 1.00
x1(1) = 0.0 x2(1) = 1.00
J* = 4.5672
J = 37.023
= 6.029
= 4.569
1.0
0.0
X2(k)
X1(k)
20
40
60
80
100
k
U(k)
0.0
−5.0
−10.0
75.0
50.0
25.0
P11(k)
P12(k)
P22(k)
X̂1(k)
X̂2(k)

Figure 5. 12

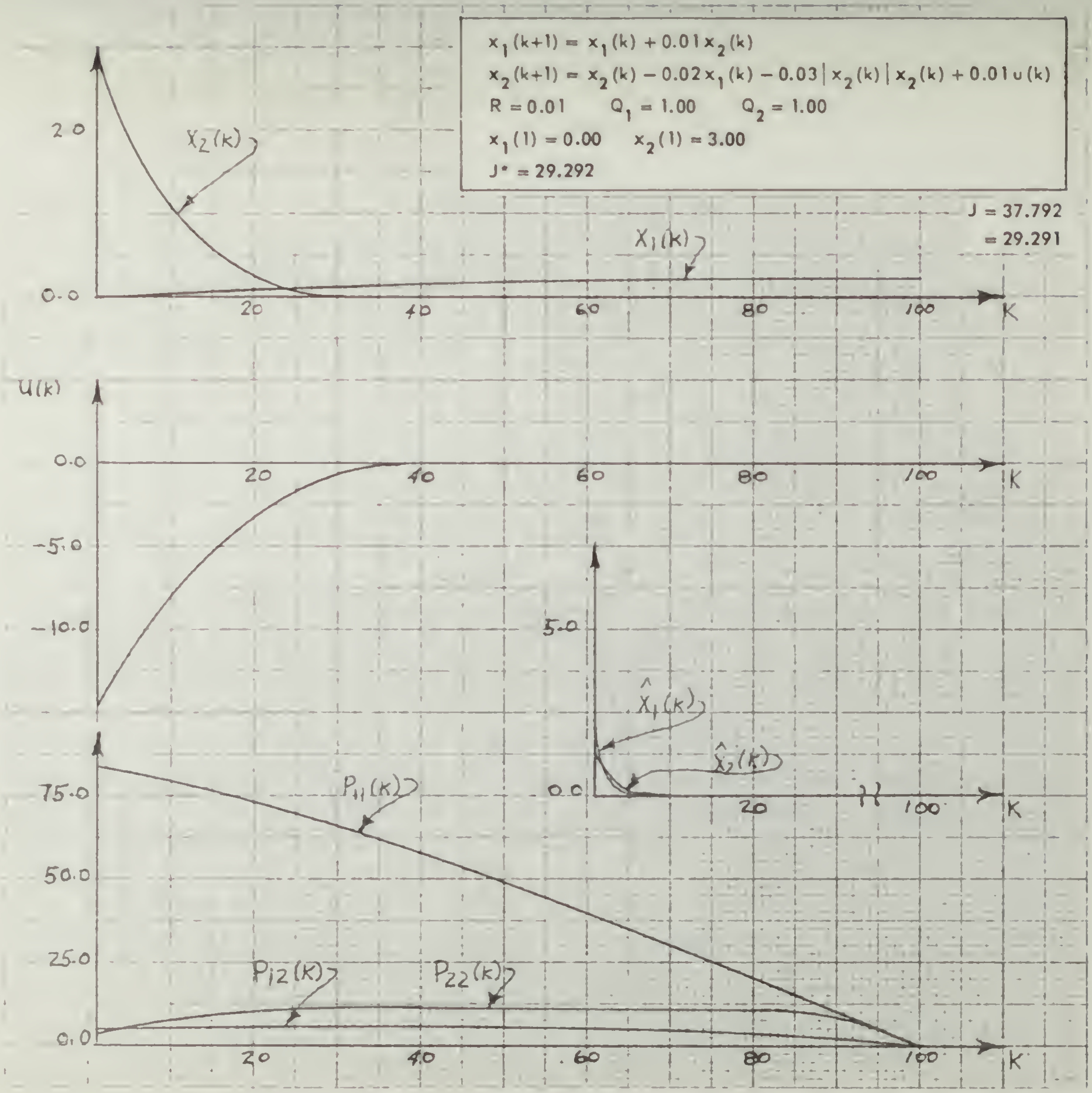
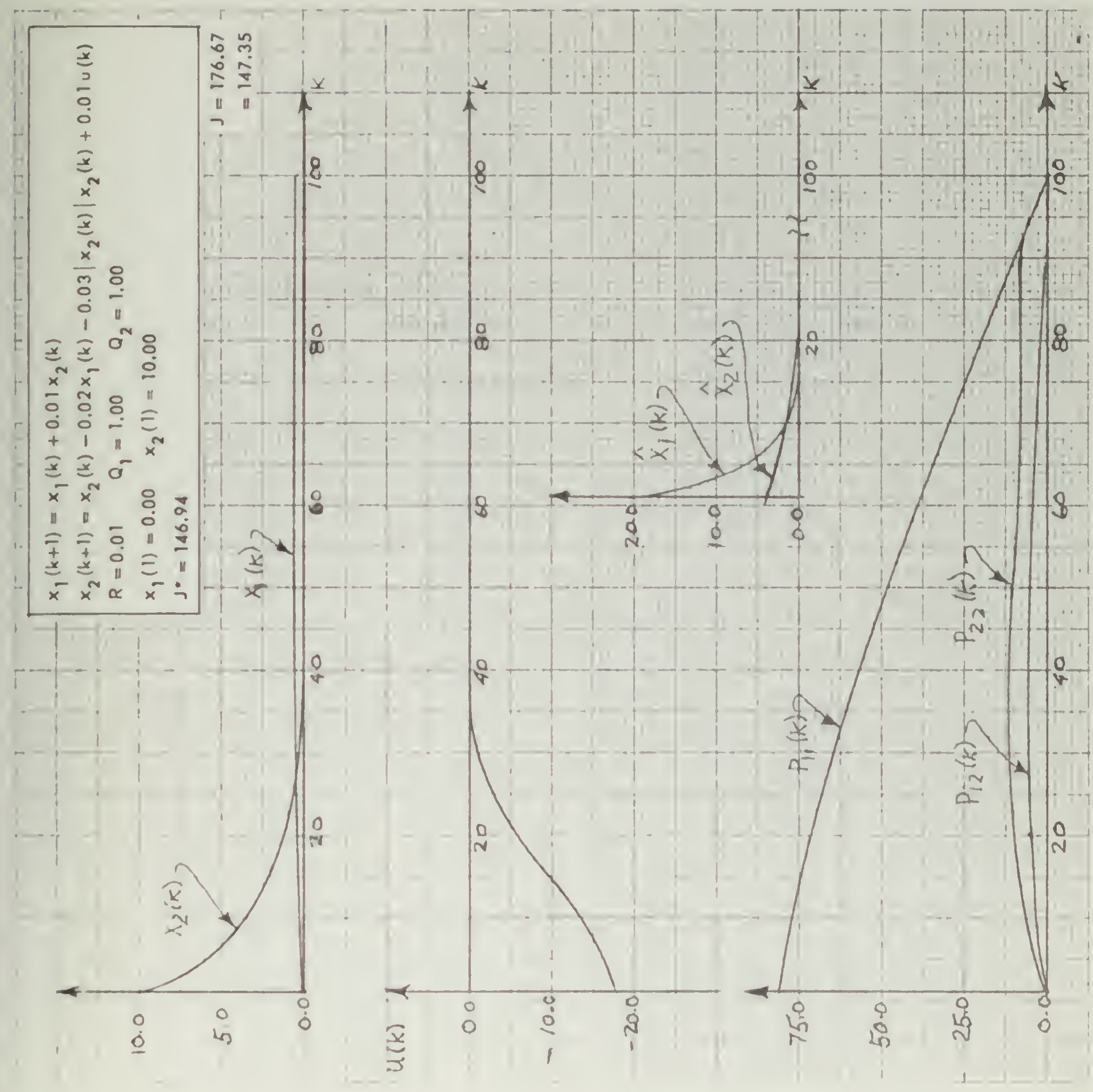

$x_1(k+1) = x_1(k) + 0.01 x_2(k)$
$x_2(k+1) = x_2(k) - 0.02 x_1(k) - 0.03 |x_2(k)| x_2(k) + 0.01 u(k)$
$R = 0.01$ $Q_1 = 1.00$ $Q_2 = 1.00$
$x_1(1) = 0.00$ $x_2(1) = 3.00$
$J^* = 29.292$
J = 37.792
= 29.291
$X_2(k)$
$X_1(k)$
$U(k)$
$\hat{X}_1(k)$
$\hat{X}_2(k)$
$P_{11}(k)$
$P_{12}(k)$
$P_{22}(k)$
K

Figure 5. 13



x1(k+1) = x1(k) + 0.01 x2(k)
x2(k+1) = x2(k) − 0.02 x1(k) − 0.03 |x2(k)| x2(k) + 0.01 u(k)
R = 0.01 Q1 = 1.00 Q2 = 1.00
x1(1) = 0.00 x2(1) = 10.00
J* = 146.94
J = 176.67
= 147.35
X2(K)
X1(k)
U(k)
P11(k)
P12(k)
P22(k)

Figure 5.14

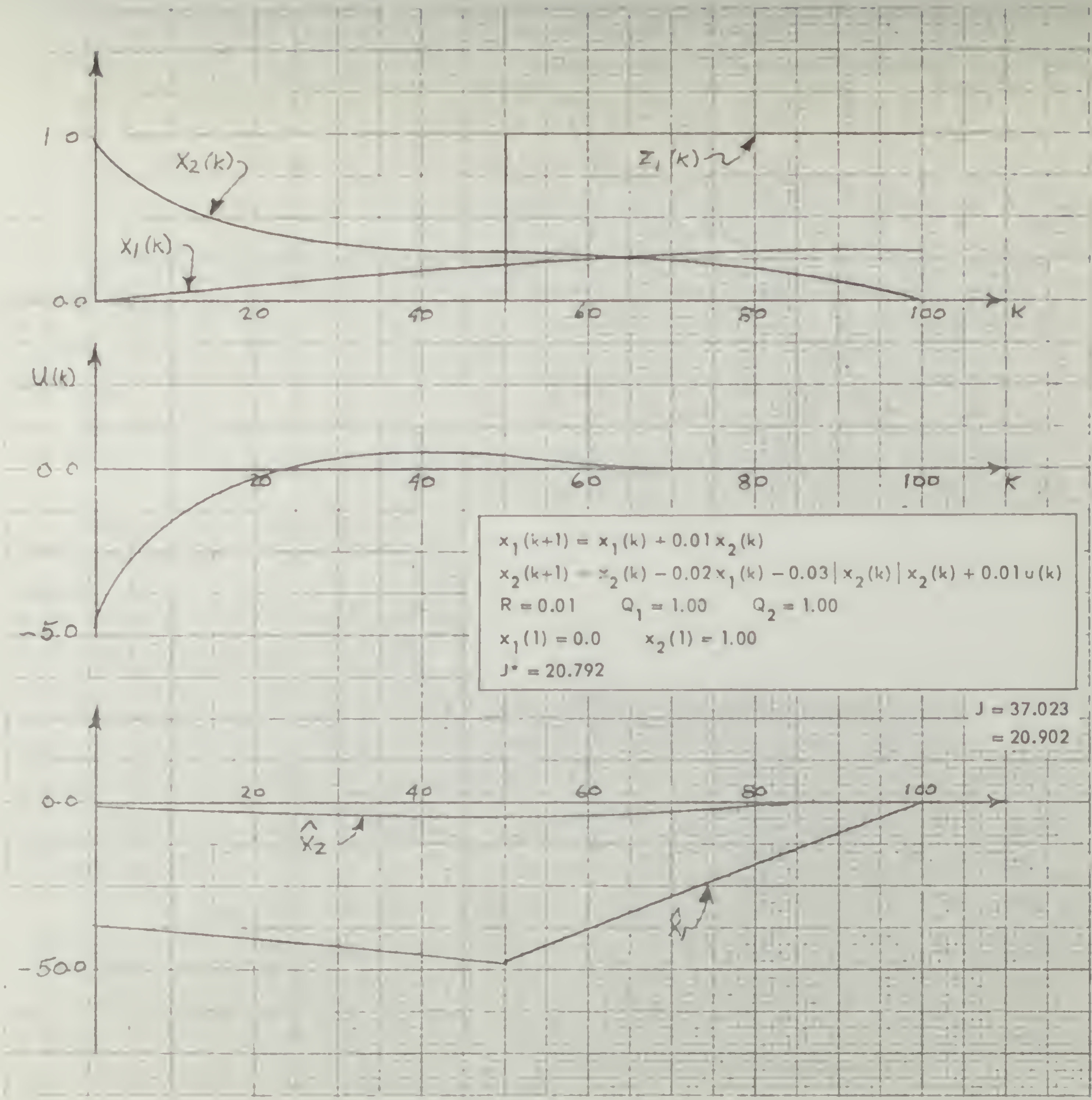

1.0
0.0
$X_2(k)$
$X_1(k)$
$Z_1(k)$
U(k)
-5.0
-500
20
40
60
80
100
K
$x_1(k+1) = x_1(k) + 0.01\,x_2(k)$
$x_2(k+1) = x_2(k) - 0.02\,x_1(k) - 0.03\,|x_2(k)|\,x_2(k) + 0.01\,u(k)$
$R = 0.01 \quad Q_1 = 1.00 \quad Q_2 = 1.00$
$x_1(1) = 0.0 \quad x_2(1) = 1.00$
$J^* = 20.792$
J = 37.023
= 20.902
$\hat{X}_2$

Figure 5. 15

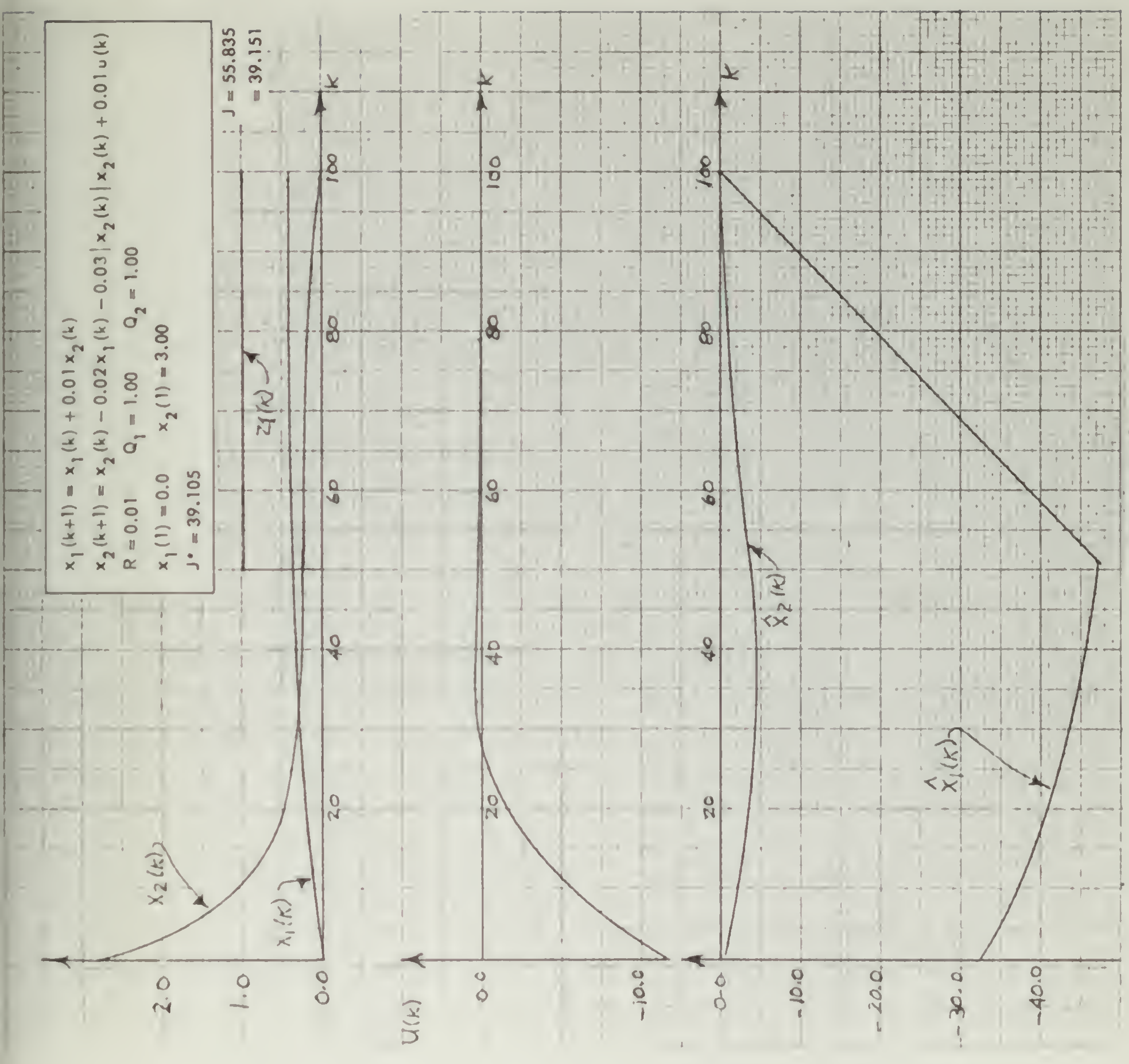
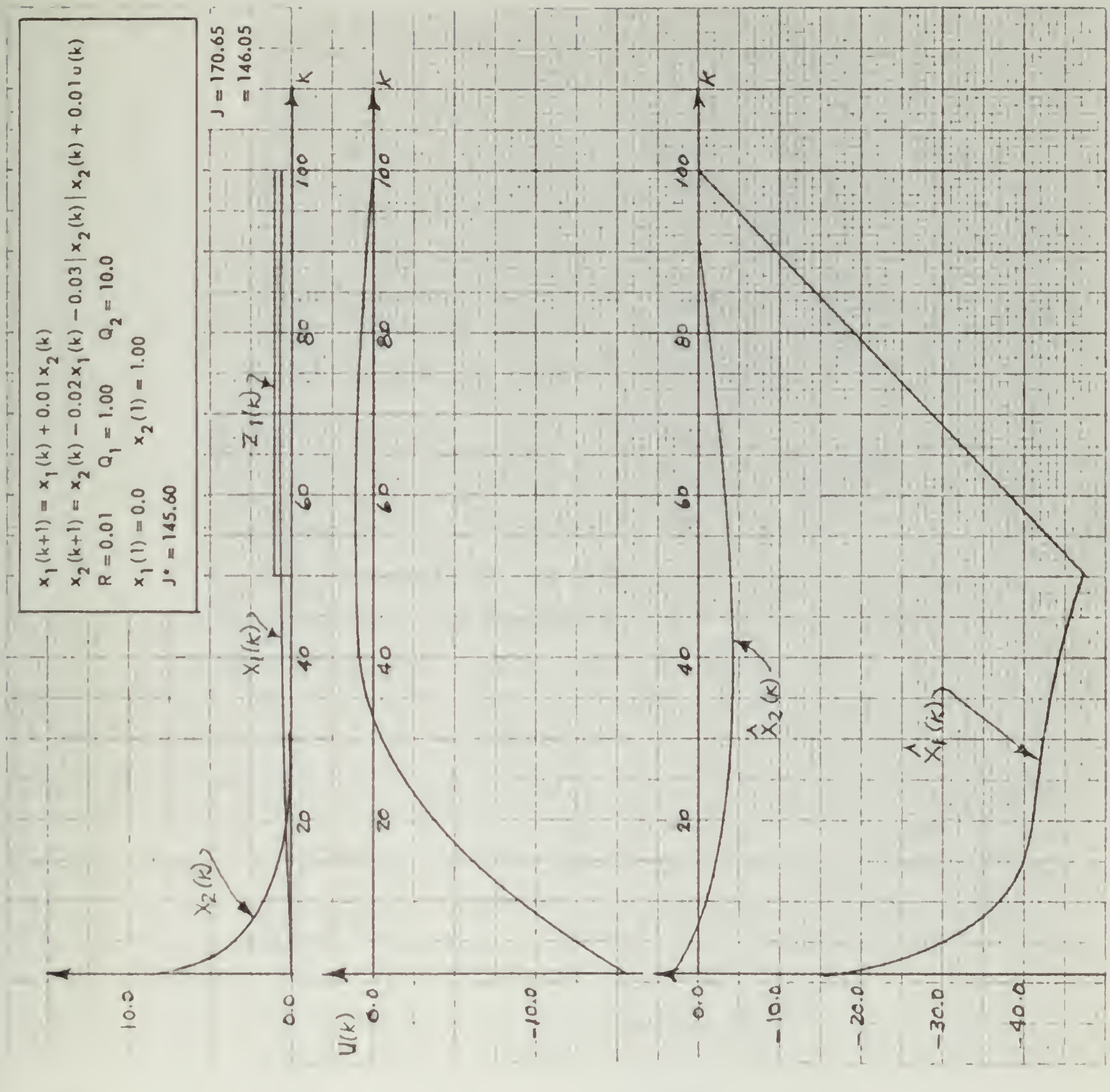
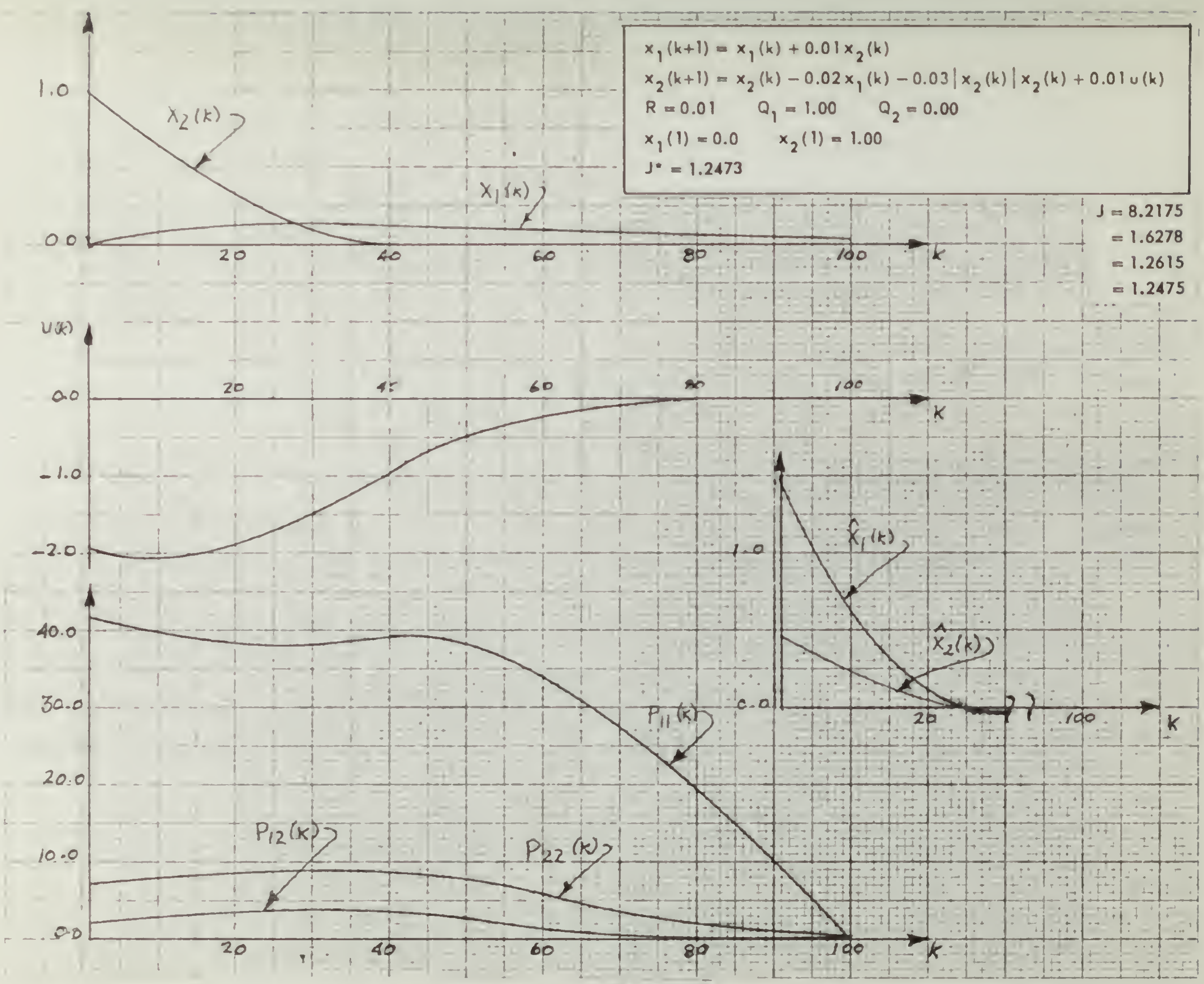

x1(k+1) = x1(k) + 0.01 x2(k)
x2(k+1) = x2(k) − 0.02 x1(k) − 0.03 |x2(k)| x2(k) + 0.01 u(k)
R = 0.01   Q1 = 1.00   Q2 = 1.00
x1(1) = 0.0   x2(1) = 3.00
J* = 39.105
J = 55.835
= 39.151
X2(k)
X1(k)
Z1(k)
U(k)
X̂2(k)
X̂1(k)
k
2.0
1.0
0.0
-10.0
-20.0
-30.0
-40.0
20
40
60
80
100

Figure 5.18



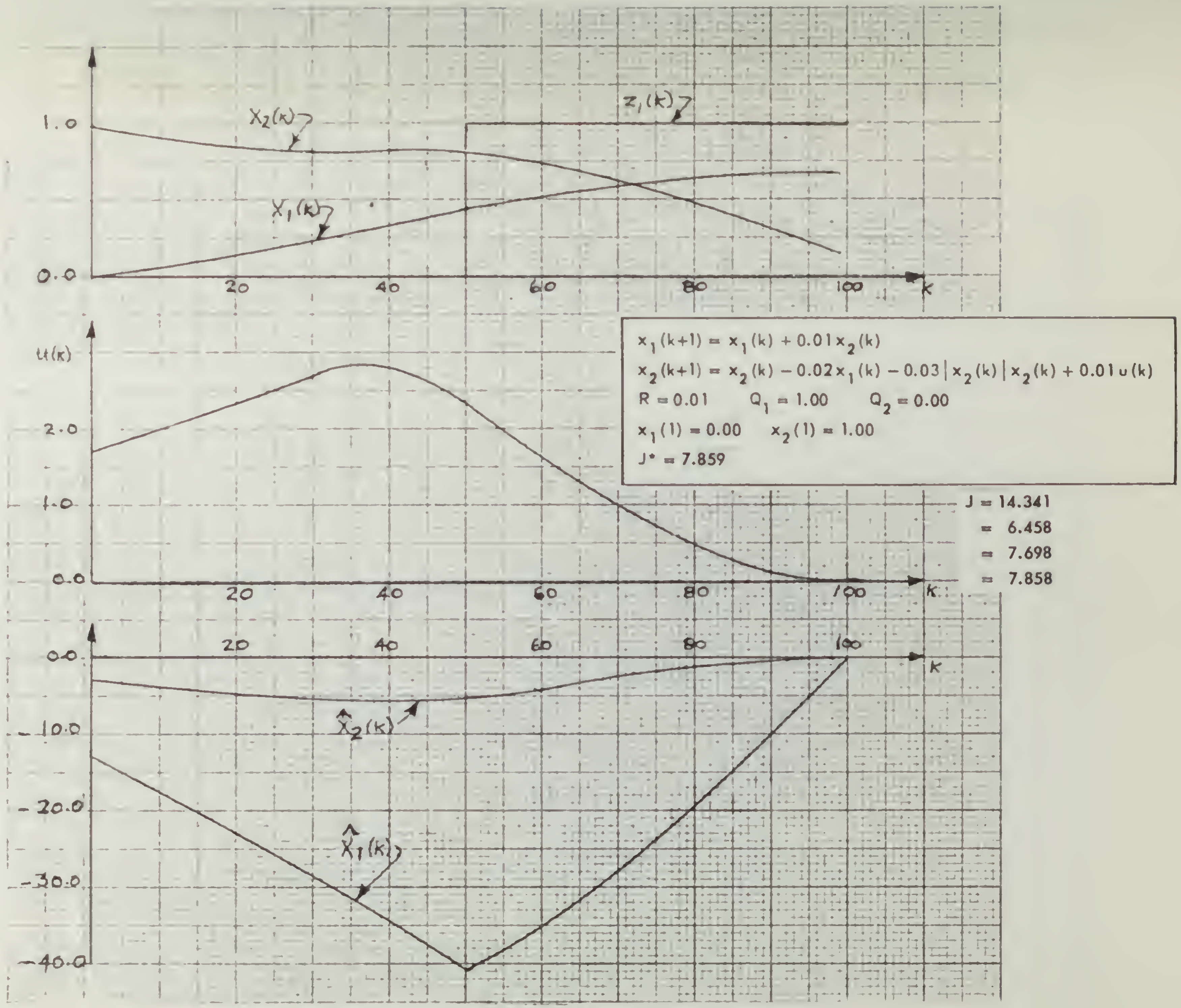
x1(k+1) = x1(k) + 0.01 x2(k)
x2(k+1) = x2(k) − 0.02 x1(k) − 0.03 | x2(k) | x2(k) + 0.01 u(k)
R = 0.01 Q1 = 1.00 Q2 = 10.0
x1(1) = 0.0 x2(1) = 1.00
J* = 145.60
J = 170.65
= 146.05
10.0
0.0
X2(k)
X1(k)
Z1(k)
k
20
40
60
80
100
U(k)
0.0
-10.0
0.0
-10.0
-20.0
-30.0
-40.0
X2(k)
X1(k)

Figure 5.17



$x_1(k+1) = x_1(k) + 0.01x_2(k)$
$x_2(k+1) = x_2(k) - 0.02x_1(k) - 0.03|x_2(k)|x_2(k) + 0.01u(k)$
R = 0.01 Q_1 = 1.00 Q_2 = 0.00
$x_1(1)$ = 0.0 $x_2(1)$ = 1.00
J* = 1.2473
J = 8.2175
= 1.6278
= 1.2615
= 1.2475
$X_2(k)$
$X_1(k)$
u(k)
$\hat{X}_1(k)$
$\hat{X}_2(k)$
$P_{11}(k)$
$P_{12}(k)$
$P_{22}(k)$
k

Figure 5.18



$x_1(k+1) = x_1(k) + 0.01x_2(k)$
$x_2(k+1) = x_2(k) - 0.02x_1(k) - 0.03|x_2(k)|x_2(k) + 0.01u(k)$
$R = 0.01 \quad Q_1 = 1.00 \quad Q_2 = 0.00$
$x_1(1) = 0.00 \quad x_2(1) = 1.00$
$J^* = 7.859$
J = 14.341
= 6.458
= 7.698
= 7.858
$X_2(k)$
$X_1(k)$
$z_1(k)$
u(k)
$\hat{X}_2(k)$
$\hat{X}_1(k)$
k

Figure 5.19

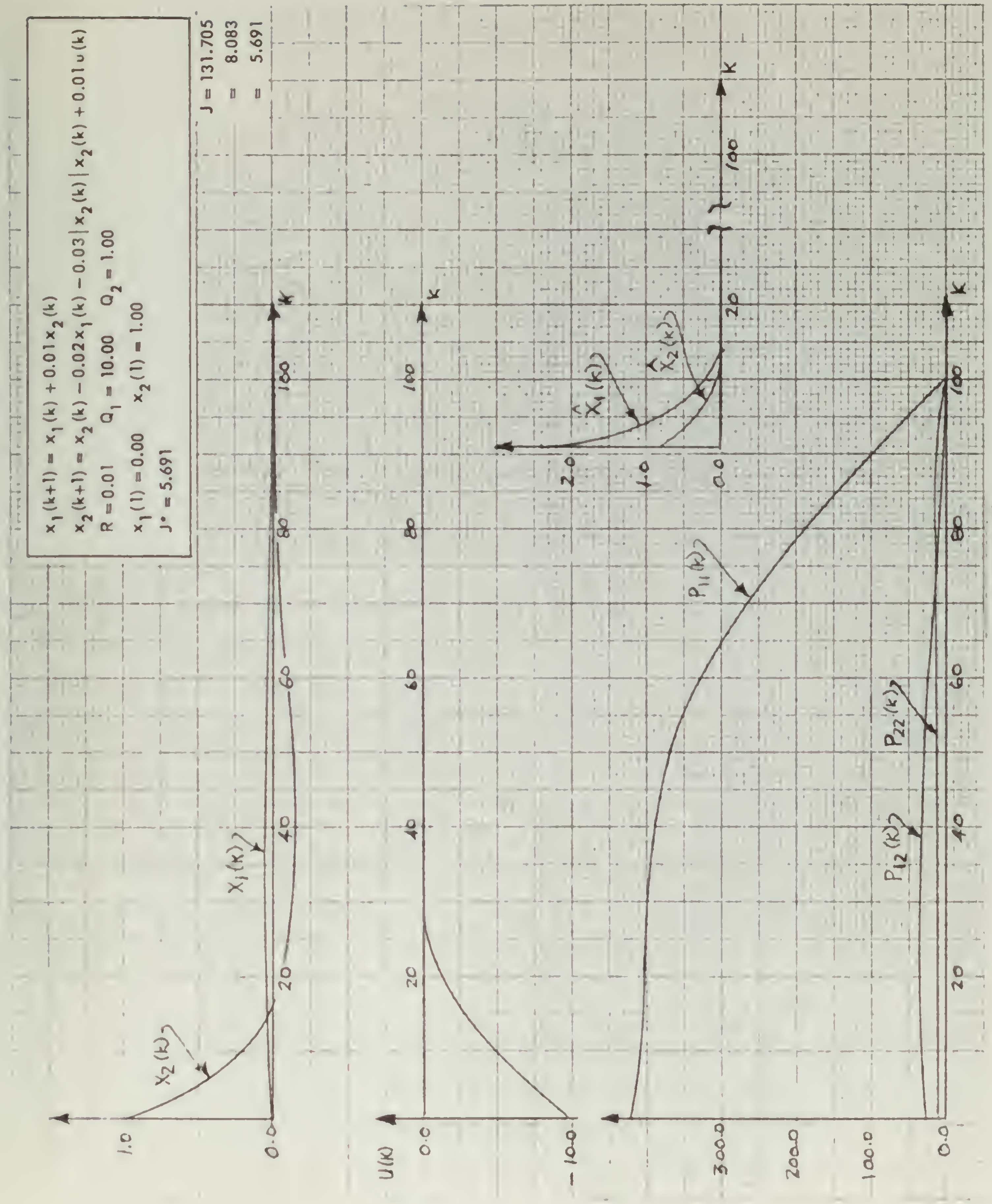
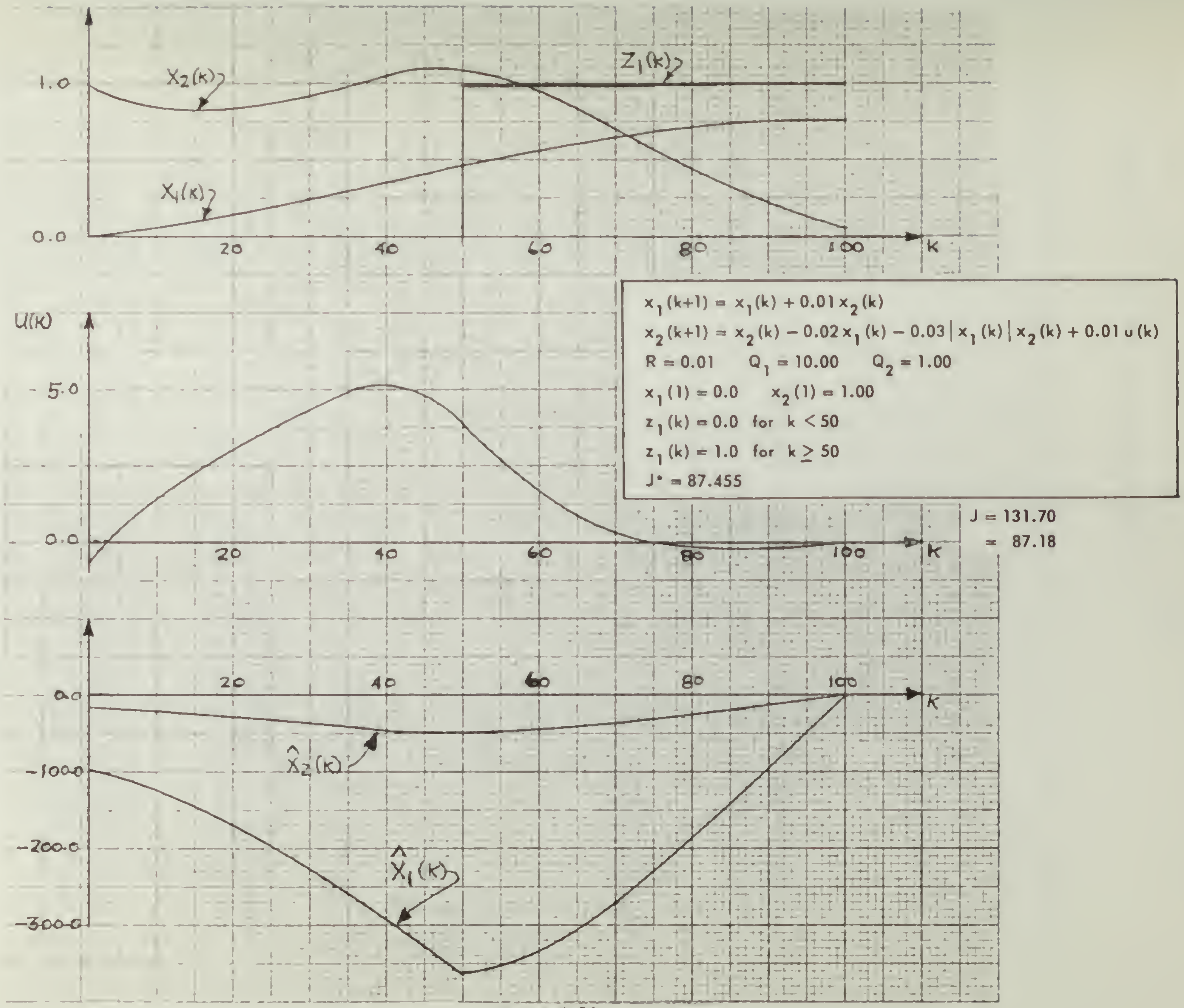

$x_1(k+1) = x_1(k) + 0.01 x_2(k)$
$x_2(k+1) = x_2(k) - 0.02 x_1(k) - 0.03 |x_2(k)| x_2(k) + 0.01 u(k)$
$R = 0.01 \quad Q_1 = 10.00 \quad Q_2 = 1.00$
$x_1(1) = 0.00 \quad x_2(1) = 1.00$
$J^* = 5.691$
$J = 131.705$
$= 8.083$
$= 5.691$
1.0
0.0
$X_2(k)$
$X_1(k)$
U(k)
-10.0
300.0
200.0
100.0
$P_{11}(k)$
$P_{12}(k)$
$P_{22}(k)$
2.0
1.0
$\hat{X}_1(k)$
$\hat{X}_2(k)$
20
40
60
80
100
K

Figure 5.20



1.0
0.0
$X_2(k)$
$X_1(k)$
$Z_1(k)$
20
40
60
80
100
k
U(K)
5.0
$x_1(k+1) = x_1(k) + 0.01\,x_2(k)$
$x_2(k+1) = x_2(k) - 0.02\,x_1(k) - 0.03\,|x_1(k)|\,x_2(k) + 0.01\,u(k)$
$R = 0.01 \quad Q_1 = 10.00 \quad Q_2 = 1.00$
$x_1(1) = 0.0 \quad x_2(1) = 1.00$
$z_1(k) = 0.0$ for $k < 50$
$z_1(k) = 1.0$ for $k \geq 50$
$J^* = 87.455$
$J = 131.70$
$= 87.18$
0.0
-100.0
-200.0
-300.0
$\hat{X}_2(k)$
$\hat{X}_1(k)$

Figure 5. 21

Figure 5.19: The same data applies here as in the previous figure except that $z_1(k) = 0.0$ for $k < 50$ and $z_1(k) = 1.0$ for $k \geq 50$. Convergence occurred in five iterations.

Figure 5.20: For this data set, $R(k) = 0.01$, $Q_1(k) = 10.0$, $Q_2(k) = 1.0$, $x_1(1) = 0.0$, $x_2(1) = 1.0$, $z_1(k) = 0.0$, and $z_2(k) = 0.0$. Convergence occurred in four iterations. The sub-optimal system with $G_1 = 28.0$ and $G_2 = 10.7$ gave a performance index of 5.71, or less than one per cent higher than that for the optimal system.

Figure 5.21: The same data applies here as in the previous figure except that $z_1(k) = 0.0$ for $k < 50$ and $z_1(k) = 1.0$ for $k \geq 50$. Convergence was achieved in three iterations.

The system considered for the second two state-variable example can be described by the equations

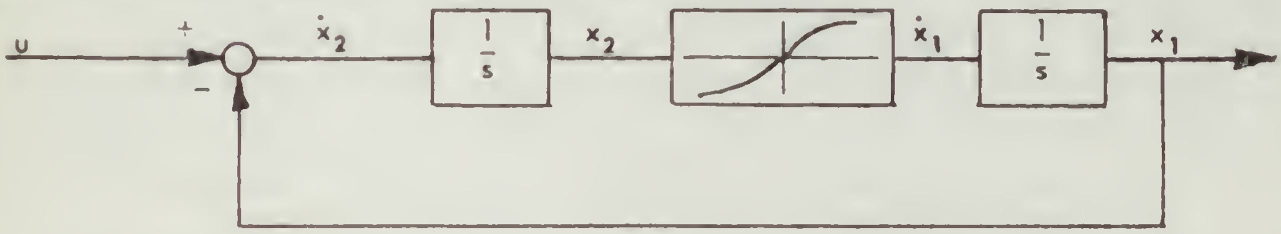
$$x_1(k+1) = x_1(k) + 0.01\,x_2(k)/(1 + |x_2(k)|); \qquad x_1(1) = c_1 \tag{5.16}$$

$$x_2(k+1) = x_2(k) - 0.01\,x_1(k) + 0.01\,u(k); \qquad x_2(1) = c_2 \tag{5.17}$$

$$y_1(k) = x_1(k) \tag{5.18}$$

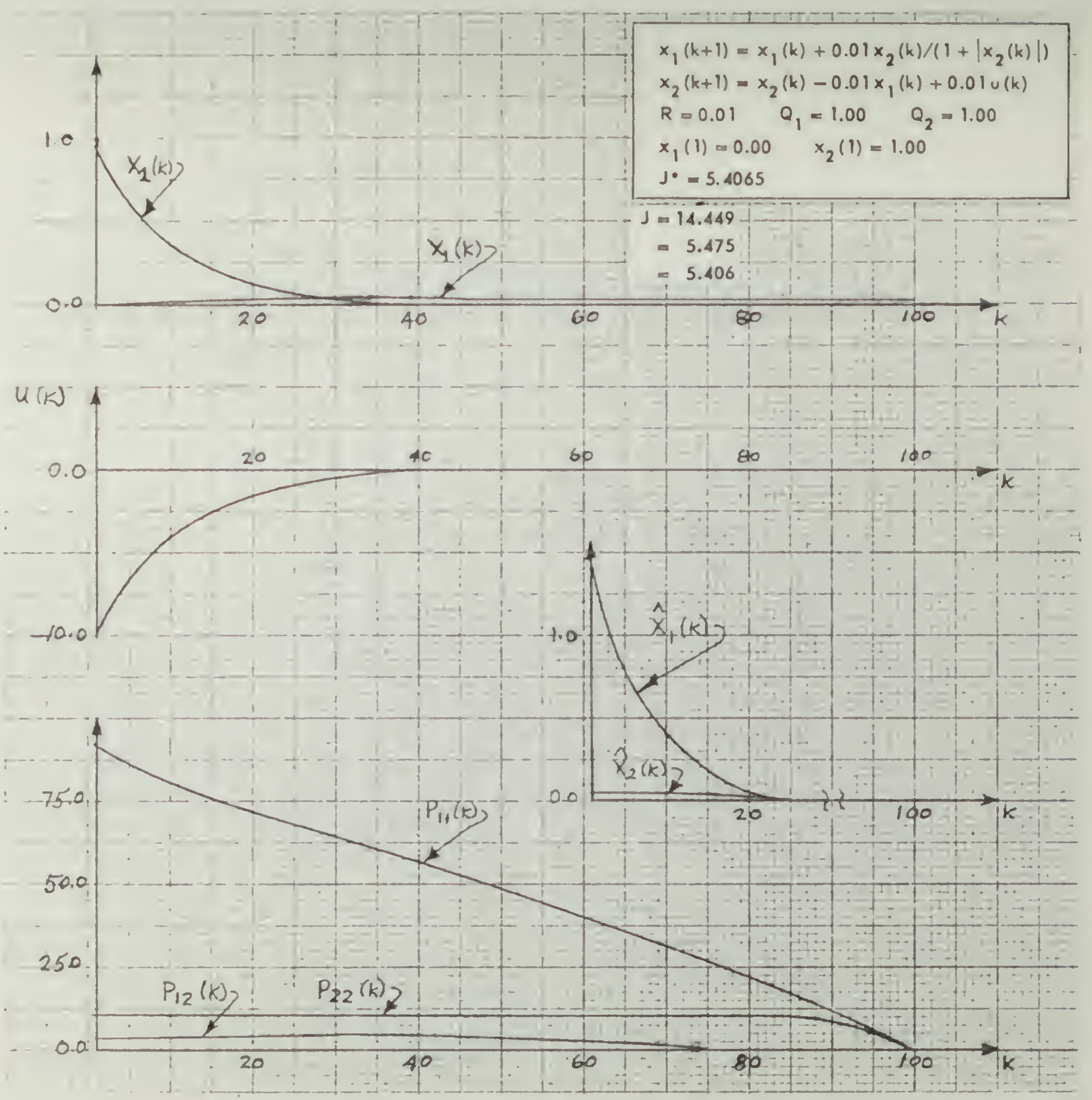
$$y_2(k) = x_2(k) \tag{5.19}$$

This system can be thought of as the discrete approximation to the system described by the block diagram below



The performance index for this example is the same as that for the previous example, and the equations used in the iterative procedure are the same except for the system equations.

Figures 5.22 through 5.25 give the plotted results from four data sets for this system.



x1(k+1) = x1(k) + 0.01x2(k)/(1 + |x2(k)|)
x2(k+1) = x2(k) − 0.01x1(k) + 0.01u(k)
R = 0.01 Q1 = 1.00 Q2 = 1.00
x1(1) = 0.00 x2(1) = 1.00
J* = 5.4065
J = 14.449
= 5.475
= 5.406
X2(k)
X1(k)
U(K)
P11(k)
P12(k)
P22(k)
k

Figure 5.22

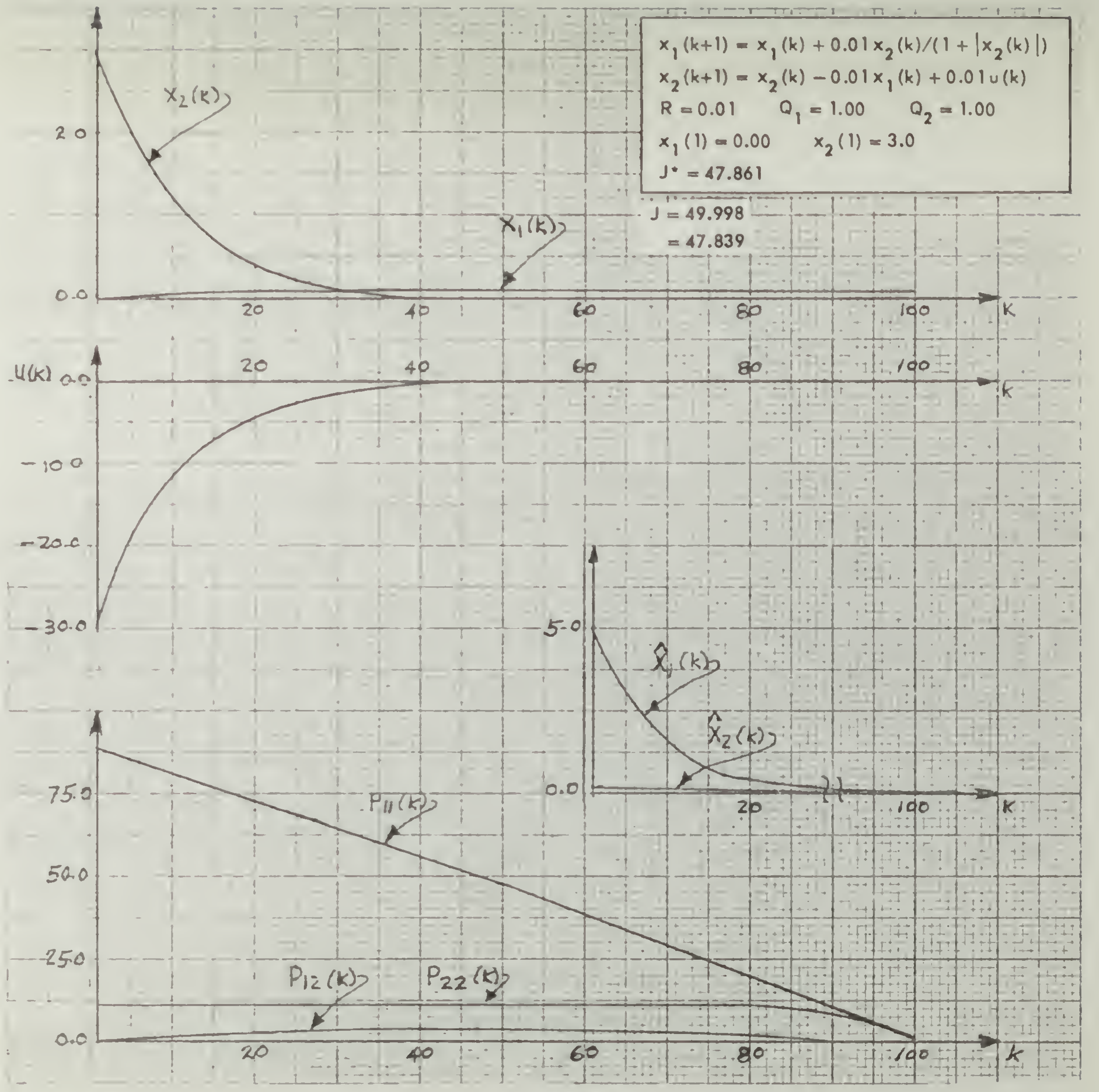
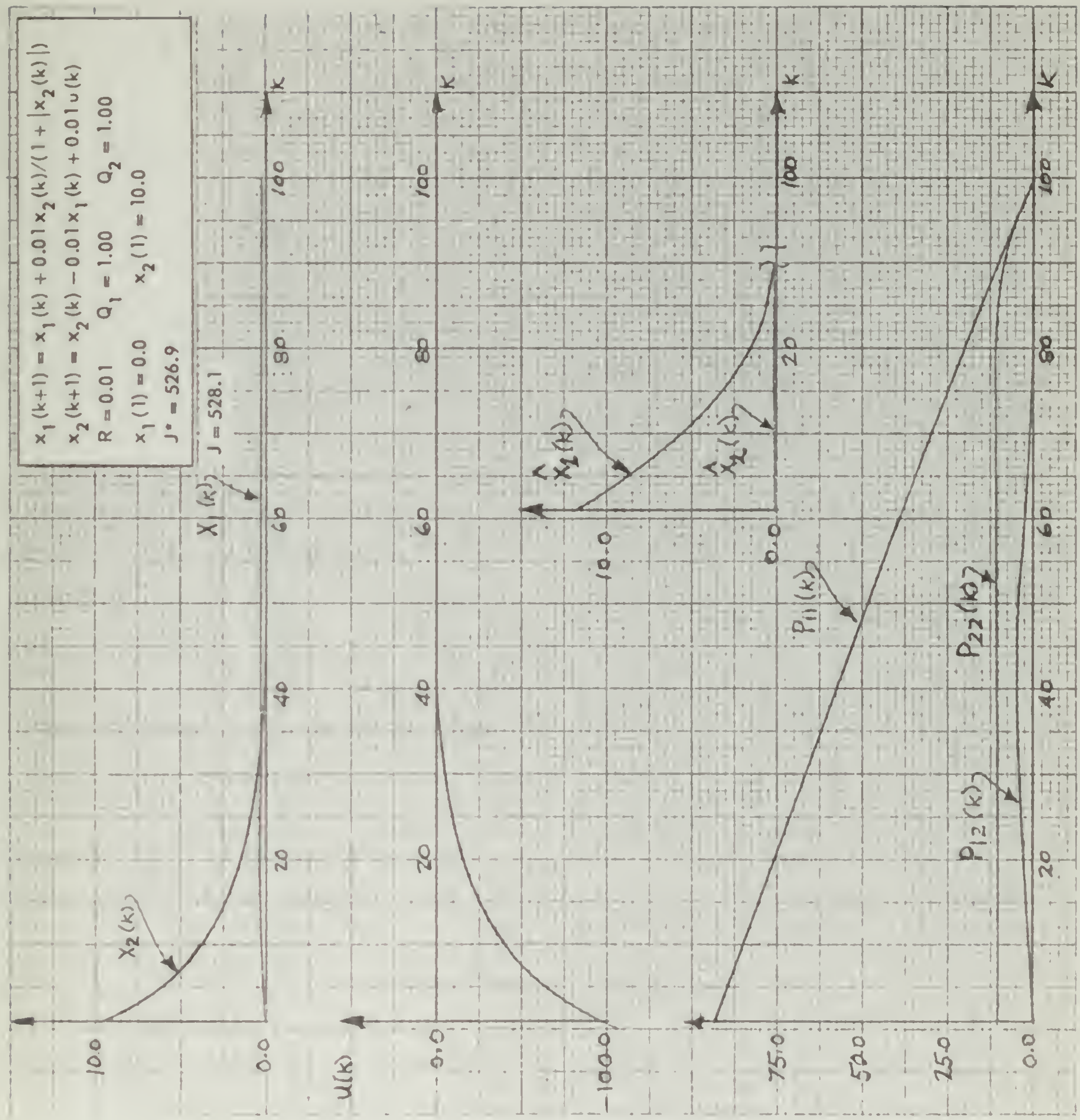

x1(k+1) = x1(k) + 0.01 x2(k)/(1 + |x2(k)|)
x2(k+1) = x2(k) − 0.01 x1(k) + 0.01 u(k)
R = 0.01 Q1 = 1.00 Q2 = 1.00
x1(1) = 0.00 x2(1) = 3.0
J* = 47.861
J = 49.998
= 47.839
X2(k)
X1(k)
U(k)
P11(k)
P12(k)
P22(k)
K

Figure 5. 23



x1(k+1) = x1(k) + 0.01 x2(k)/(1 + |x2(k)|)
x2(k+1) = x2(k) − 0.01 x1(k) + 0.01 u(k)
R = 0.01   Q1 = 1.00   Q2 = 1.00
x1(1) = 0.0   x2(1) = 10.0
J* = 526.9
J = 528.1

Figure 5.24

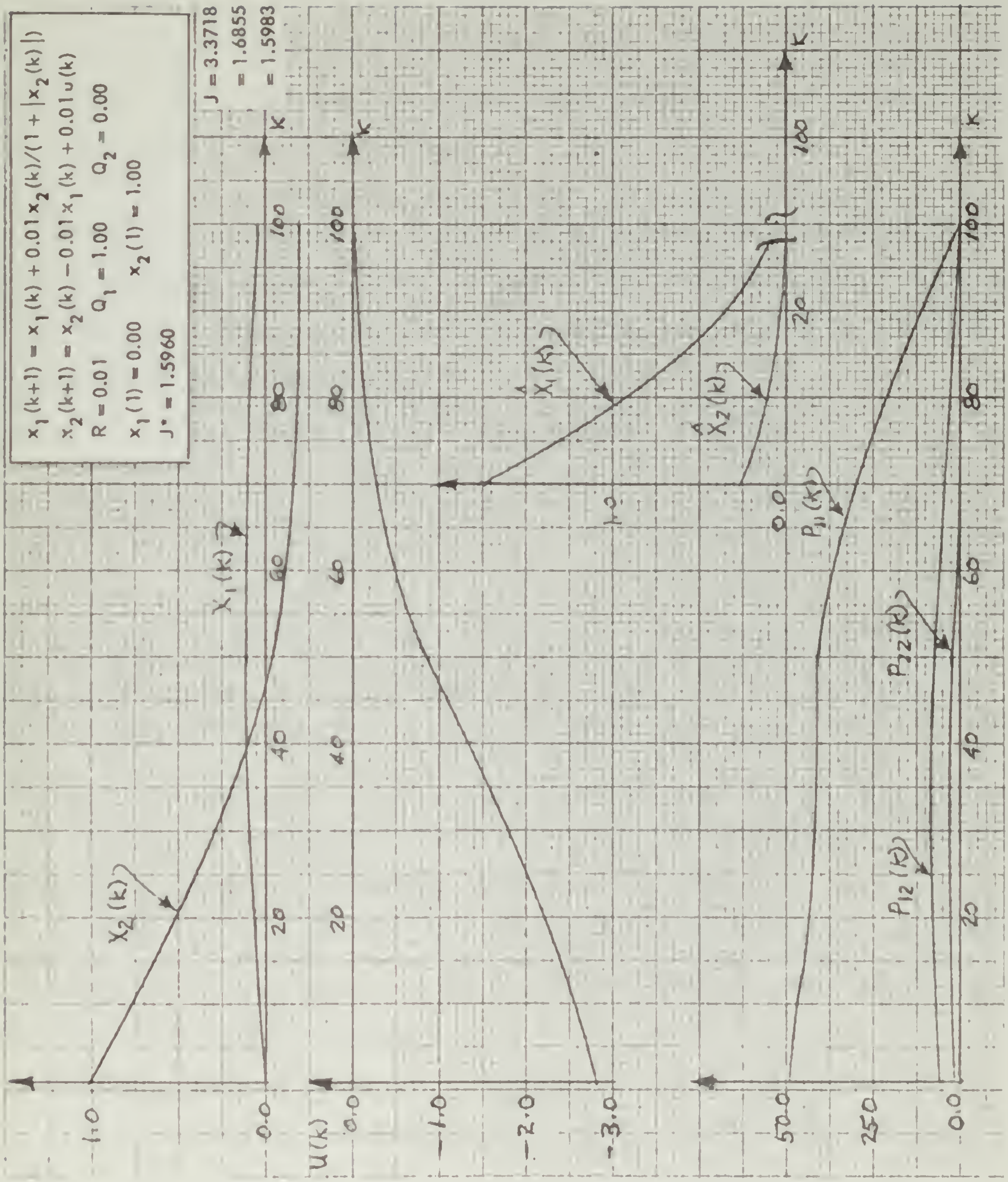

x1(k+1) = x1(k) + 0.01 x2(k)/(1 + |x2(k)|)
x2(k+1) = x2(k) − 0.01 x1(k) + 0.01 u(k)
R = 0.01 Q1 = 1.00 Q2 = 0.00
x1(1) = 0.00 x2(1) = 1.00
J* = 1.5960
J = 3.3718
= 1.6855
= 1.5983
X2(k)
X1(k)
U(k)
P11(k)
P12(k)
P22(k)
K

Figure 5. 25

Figures 5.22 - 5.24: For these data sets, $R(k) = 0.01$, $Q_1(k) = 1.0$, $Q_2(k) = 1.0$, $x_1(1) = 0.0$, $z_1(k) = 0.0$, and $z_2(k) = 0.0$. In figures 5.22, 5.23, and 5.24, $x_2(1) = 1.0$, 3.0, and 10.0, respectively. The number of iterations required for convergence was three, two, and two.

Figure 5.25: For this data set, $R(k) = 0.01$, $Q_1(k) = 1.0$, $Q_2(k) = 0.0$, $x_1(1) = 0.0$, $x_2(1) = 1.0$, $z_1(k) = 0.0$, and $z_2(k) = 0.0$. Four iterations were required for convergence.

One additional variation of this problem was run in an effort to get some indication of under what conditions the iterative routine might not converge. For this purpose, the nonlinearity was made more violent by changing the system equations to

$$x_1(k+1) = x_1(k) + x_2(k)/(1.0 + 10.0\,|x_2(k)|); \qquad x_1(1) = c_1 \tag{5.20}$$

$$x_2(k+1) = x_2(k) - 0.01\,x_1(k) + 0.01\,u(k); \qquad x_2(1) = c_2 \tag{5.21}$$

$$y_1(k) = x_1(k) \tag{5.22}$$

$$y_2(k) = x_2(k) \tag{5.23}$$

For each of these data sets, $R = 0.01$, $Q_1 = 1.00$, $Q_2 = 1.00$, and $x_1(1) = 0.00$. For the data set with $x_2(1) = 10.0$, convergence occurred in five iterations. For the data set with $x_2(1) = 3.00$, convergence occurred in four iterations. For the data set with $x_2(1) = 1.00$, convergence occurred after some rather severe oscillations in the convergence criterion, and then only after 19 iterations.

The convergence was slower when a small initial condition was used probably because in this case the system spent more time operating in the highly nonlinear regions.

We can conclude from this variation of the example problem, that when the nonlinearity is severe, the iterative routine may converge slowly or not at all.

For purposes of comparing convergence rates, the value of the performance index, J, computed on each iteration has been included

in most of the preceding figures. The value of the performance index computed on the convergent iteration is denoted by J^*.

5.4 Stochastic Examples

The results of three stochastic examples are presented in this section. In each of these examples, the nonlinear system being controlled is disturbed by a random input.

The computer algorithm that was used is outlined below.

Step 1. Using $\underline{P}(k) = \underline{0}$, $\hat{\underline{x}}(k) = \underline{0}$, and $\underline{r}(k) \equiv \underline{0}$, the control and the state variables are extrapolated ahead to determine $\underline{u}(1), \ldots,$ $\underline{u}(99)$ and $\underline{x}(2), \ldots, \underline{x}(100)$.

Step 2. Using the $\underline{u}(k)$ and the $\underline{x}(k)$ just determined, $\underline{P}(99), \ldots,$ $\underline{P}(11)$ and $\hat{\underline{x}}(99), \ldots, \hat{\underline{x}}(11)$ are computed by backward recursion.

Step 3. The control, $\underline{u}(1), \ldots, \underline{u}(10)$, and the state, $\underline{x}(2), \ldots,$ $\underline{x}(11)$, are computed with $\underline{r}(1), \ldots, \underline{r}(10)$ taking on random values, simulating the actual evolution of the nonlinear system.

Step 4. Using $\underline{P}(k)$ and $\hat{\underline{x}}(k)$ previously determined, and $\underline{r}(k) \equiv \underline{0}$, the control and the state variables are extrapolated ahead to determine $\underline{u}(11), \ldots, \underline{u}(99)$ and $\underline{x}(12), \ldots, \underline{x}(100)$.

Step 5. Using the $\underline{x}(k)$ and the $\underline{u}(k)$ just determined, $\underline{P}(99), \ldots,$ $\underline{P}(21)$ and $\hat{\underline{x}}(99), \ldots, \hat{\underline{x}}(21)$ are computed.

Steps 3, 4, and 5 are then repeated, starting at $k = 11$, $k = 21$, etc., until the actual simulation has evolved to $k = 100$. The system should be visualized with steps 4 and 5 simulating the controller in fast time, and step 3 simulating the actual evolution of the nonlinear system in real time.

Figure 5.26: The results given in this figure are for the example using the system of section 5.2, but with an independent random disturbance, $r(k)$, added. For this data set, $R(k) = 0.01$, $Q(k) = 1.0$, $x(1) = 1.0$, and $z(k) = 0.0$. Notice that by the time $k = 21$, the P and $\hat{x}$ variables

are well determined with no more jumps, indicating that despite the random disturbance, the control system is operating near optimally. In this example and the others of this section, $r(k)$ is a zero mean, unit variance, independent random sequence.

Figure 5.27: The results given in this figure are for the example using the same system as the first example in section 5.3, but with an independent random disturbance, $r(k)$, added to the x_2 component. For this data set, $R(k) = 0.01$, $Q_1(k) = 1.0$, $Q_2(k) = 1.0$, $x_1(1) = 0.0$, $x_2(1) = 1.0$, $z_1(k) = 0.0$, and $z_2(k) = 0.0$.

Figure 5.28: The results given in this figure are for a nonlinear system disturbed by dependent noise. In this example, $x_1(k)$ represents the dependent noise which is obtained from independent noise by the system

$$x_1(k+1) = 0.95\,x_1(k) + 0.05\,r(k); \qquad x_1(1) = 0.0 \tag{5.24}$$

where $r(k)$ is an independent random variable. The state of the nonlinear system being controlled is represented by $x_2(k)$, and is determined by the equation

$$x_2(k+1) = x_2(k) - 0.05\,x_2^3(k) + 0.05\,u(k) + 0.05\,x_1(k); \qquad x_2(1) = 1.0 \tag{5.25}$$

Together $x_1(k)$ and $x_2(k)$ make up an augmented two-dimensional state vector. For this data set, $R(k) = 0.01$, $Q_1(k) = 0.0$ (as we have no control over the noise), $Q_2(k) = 1.0$, $z_1(k) = 0.0$, and $z_2(k) = 0.0$.

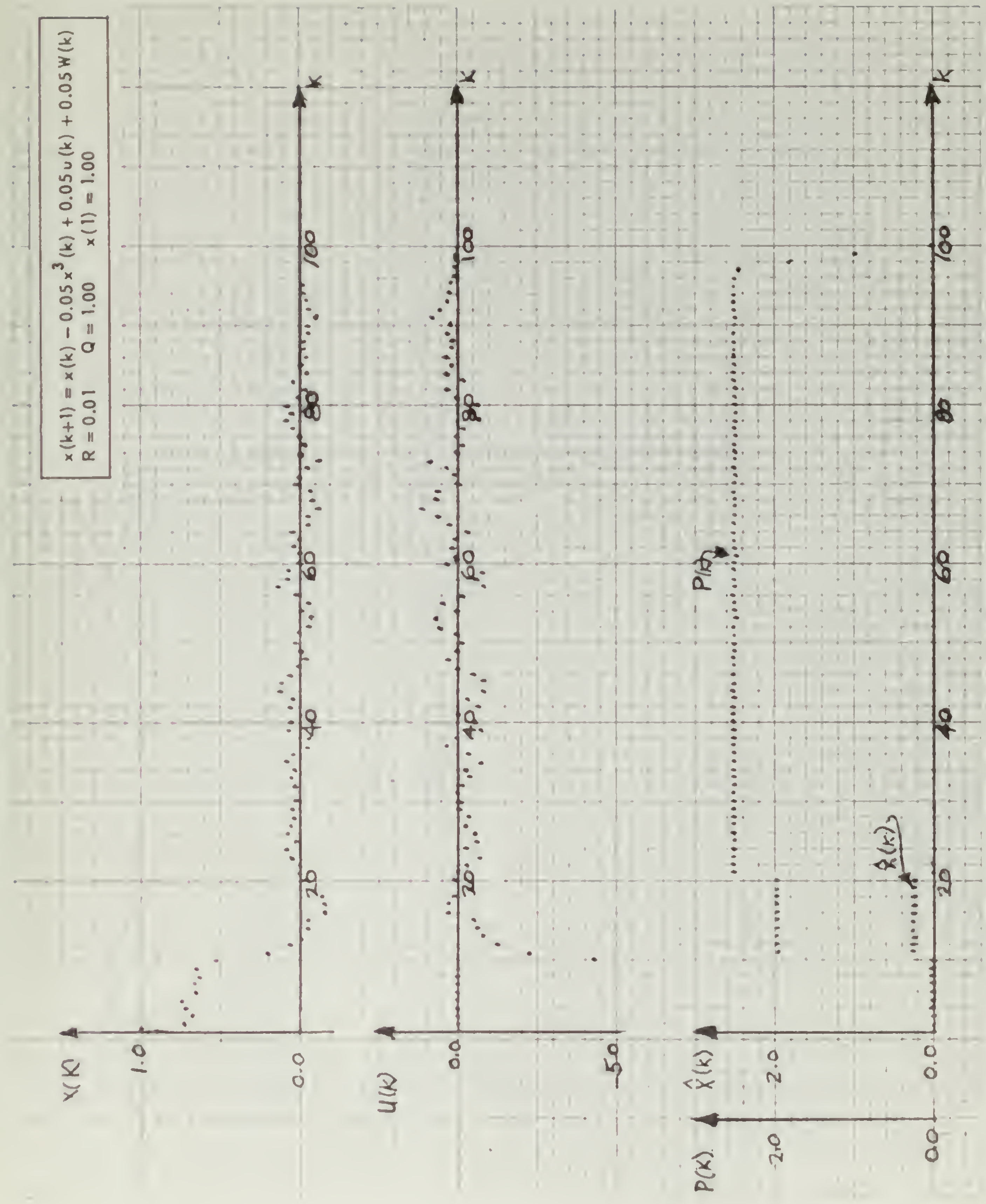

x(k+1) = x(k) − 0.05 x³(k) + 0.05 u(k) + 0.05 W(k)
R = 0.01 Q = 1.00 x(1) = 1.00
Y(K)
1.0
0.0
U(k)
0.0
-5.0
X̂(k)
2.0
0.0
P(K)
2.0
0.0
20
40
60
80
100
k
P(k)
x̂(k)

Figure 5. 26

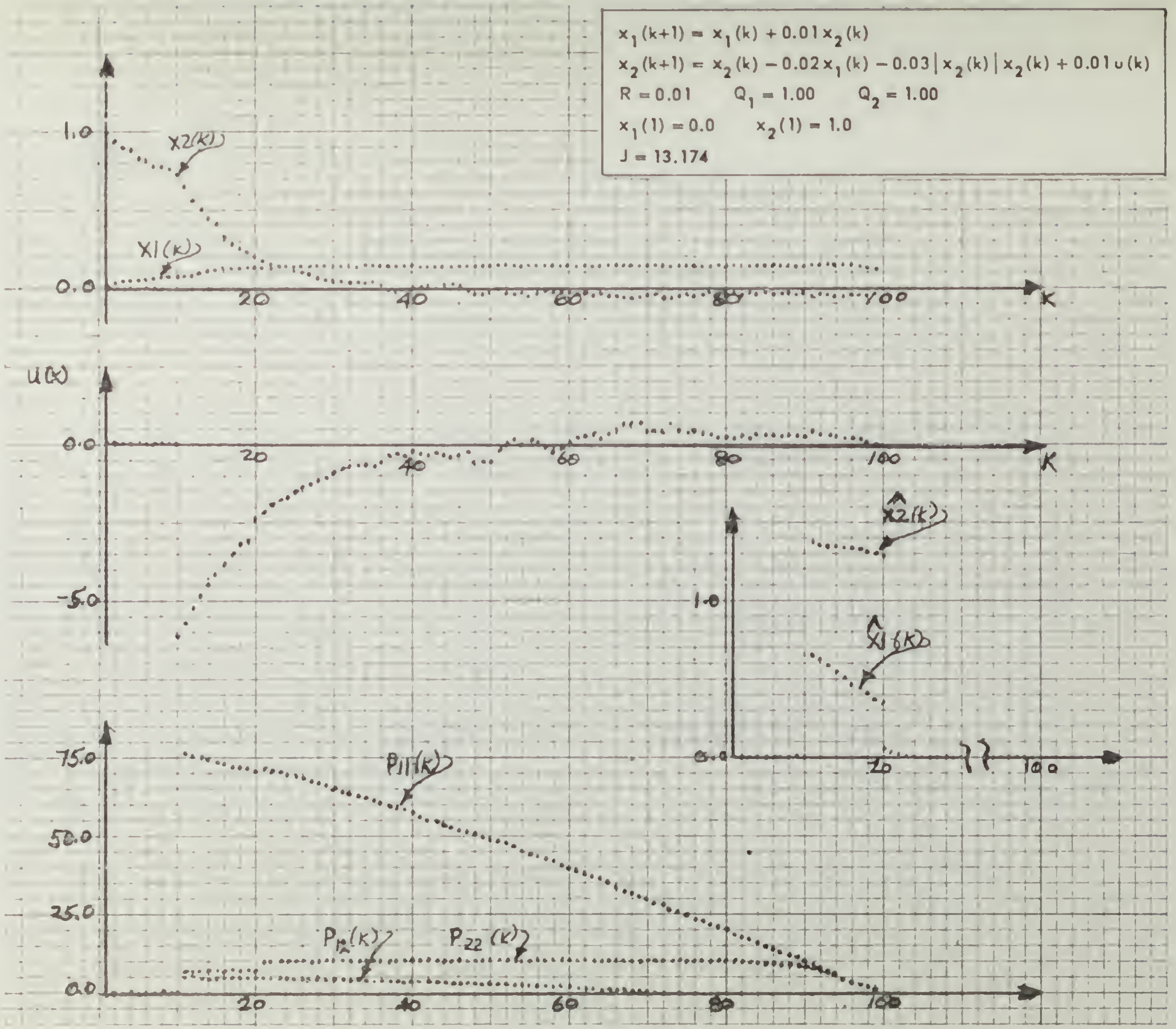
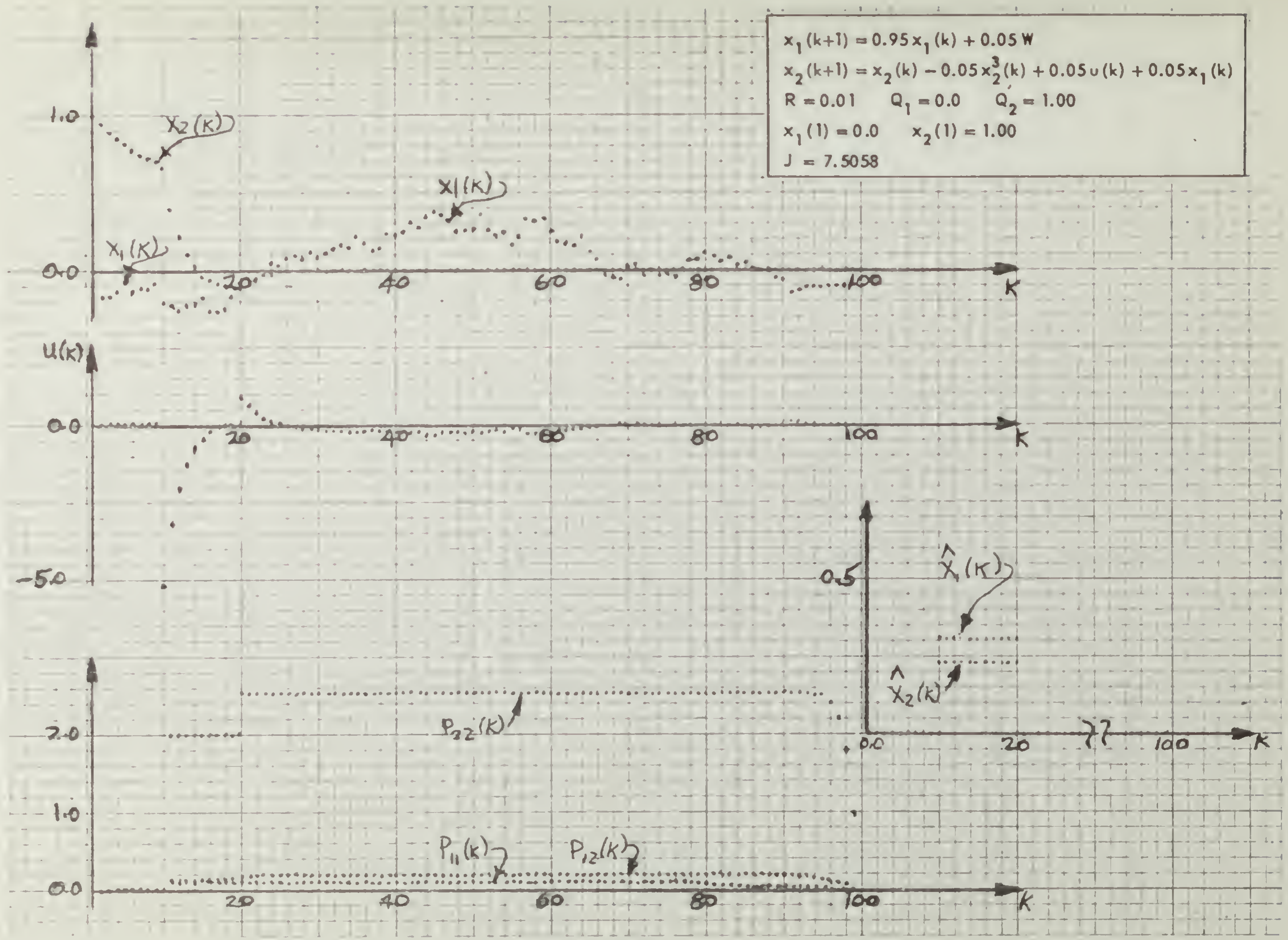

x1(k+1) = x1(k) + 0.01 x2(k)
x2(k+1) = x2(k) − 0.02 x1(k) − 0.03 | x2(k) | x2(k) + 0.01 u(k)
R = 0.01 Q1 = 1.00 Q2 = 1.00
x1(1) = 0.0 x2(1) = 1.0
J = 13.174
1.0
0.0
X2(K)
X1(K)
20
40
60
80
100
K
U(K)
-5.0
75.0
50.0
25.0
P11(K)
P12(K)
P22(K)
1.0
20
100

Figure 5. 27



x1(k+1) = 0.95x1(k) + 0.05W
x2(k+1) = x2(k) − 0.05x2^3(k) + 0.05u(k) + 0.05x1(k)
R = 0.01 Q1 = 0.0 Q2 = 1.00
x1(1) = 0.0 x2(1) = 1.00
J = 7.5058
1.0
0.0
X2(K)
X1(K)
x1(K)
u(k)
-5.0
2.0
P22(K)
P11(k)
P12(k)
0.5
X̂1(K)
X̂2(K)
20
40
60
80
100
K

Figure 5. 28

CHAPTER VI

CONCLUSIONS

The major contribution of this work has been the presentation of a theory along with an iterative algorithm for the solution of optimal nonlinear control problems subject to quadratic performance criteria. In addition, the results of the computer examples presented in Chapter V have demonstrated the feasibility of the method.

A by-product of the theory has been the analytic solution of the problems of Chapter IV. In Chapters IV and V, comparisons of sub-optimal systems with the optimal ones determined by the theory have shown that often near-optimal performance is possible with simple linear controllers, a possibility that has been suspected but not demonstrated previously.

All is not rosy, however. Appendix C shows that the method is essentially limited to problems of no more than five state variables and control intervals of no more than 1000 steps by the size and speed of presently available digital computers.

Many questions have been raised, but not answered. Of prime importance among these is the question of under what conditions can the convergence of the iterative algorithm be guaranteed. Further research on the problem with stochastic disturbances is required in order to determine under what conditions the control procedure presented in section 2.7 is reasonable.

It would be highly desirable to be able to rephrase the problem in such a way that the optimal control system determined by the theory would be restricted to be non-anticipative. This problem has been worked on briefly by the author, but without results.

Finally, although it is conceivable that actual control systems may be synthesized by this method, it is far more likely that the main use for the theory will be to establish ultimate performance figures for comparison purposes in design studies. Further research in this direction seems warranted.

APPENDIX A

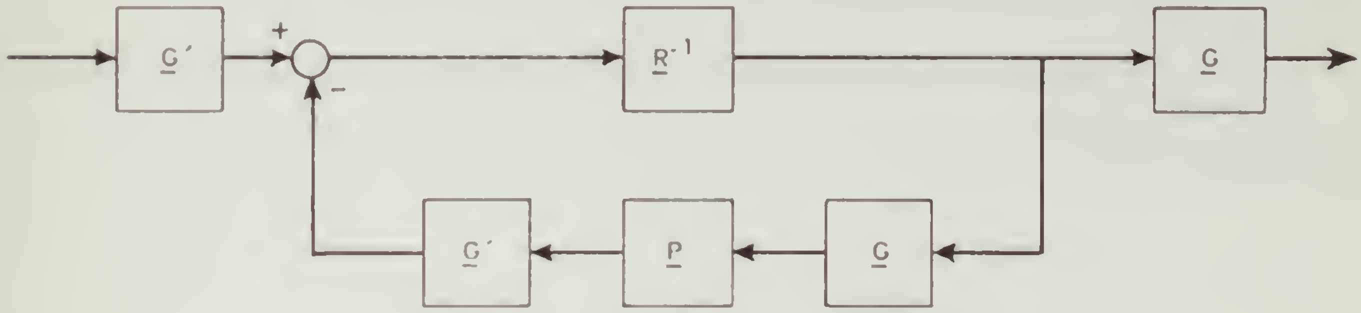
TWO MATRIX IDENTITIES

Theorem A1: If $\underline{R}^{-1}$ and $[\underline{R} + \underline{G}'\underline{PG}]^{-1}$ exist, then

$$[\underline{I} + \underline{GR}^{-1}\underline{G}'\underline{P}]^{-1}\,\underline{GR}^{-1}\underline{G}' \equiv \underline{G}\,[\underline{R} + \underline{G}'\underline{PG}]^{-1}\,\underline{G}' \qquad \text{(A. 1)}$$

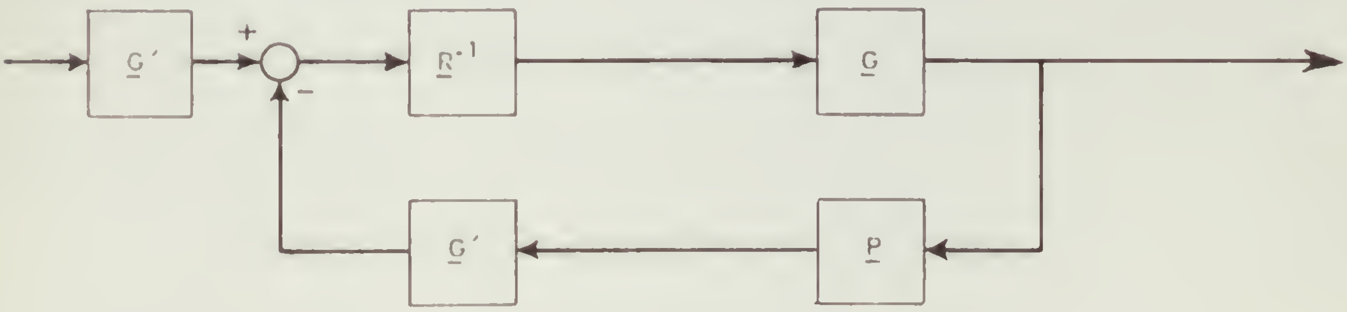
Proof: The proof uses a method of matrix manipulations given by Cox.[22] This method views a matrix as a linear transformation and shows that such transformations obey all the rules for block diagram manipulation provided order of blocks is preserved. In other words, block diagram manipulations may be used to prove matrix identities.

For the proof of this theorem, it is easy to show that the expression on the right-hand side of equation (A. 1) can be represented by the block diagram

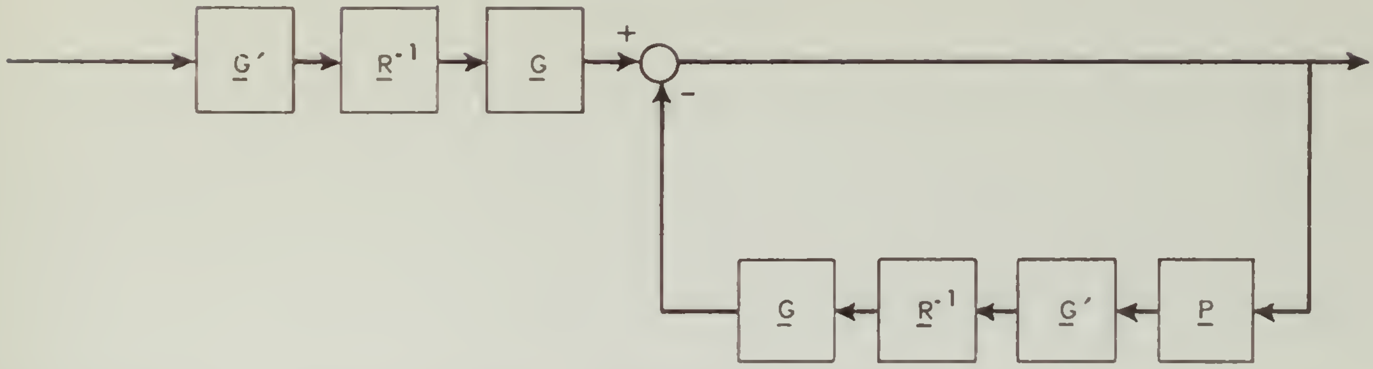


provided $\underline{R}^{-1}$ and $[\underline{R} + \underline{G}'\underline{PG}]^{-1}$ exist.

By moving $\underline{G}$ into the loop we get



Then moving $\underline{R}^{-1}$ and $\underline{G}$ back out the other side of the loop gives



But this block diagram is equivalent to the expression

$$[\underline{I} + \underline{GR}^{-1}\underline{G}'\underline{P}]^{-1}\ \underline{GR}^{-1}\underline{G}' \qquad \text{(A. 2)}$$

which proves the theorem.

Theorem A2: If $\underline{R}^{-1}$ and $[\underline{R} + \underline{G}'\underline{PG}]^{-1}$ exist, then

$$[\underline{I} + \underline{GR}^{-1}\underline{G}'\underline{P}]^{-1} = \underline{I} - \underline{G}[\underline{R} + \underline{G}'\underline{PG}]^{-1}\underline{G}'\underline{P} \qquad \text{(A. 3)}$$

Proof: The proof proceeds by using the definition of an inverse. Thus if the right-hand side of (A. 3) is truly the inverse of $\underline{I} + \underline{GR}^{-1}\underline{G}'\underline{P}$ then we must have

$$[\underline{I} + \underline{GR}^{-1}\underline{G}'\underline{P}]\{\underline{I} - \underline{G}[\underline{R} + \underline{G}'\underline{PG}]^{-1}\underline{G}'\underline{P}\} = \underline{I} \qquad \text{(A. 4)}$$

or

$$\underline{I} + \underline{GR}^{-1}\underline{G}'\underline{P} - \underline{G}[\underline{R} + \underline{G}'\underline{PG}]^{-1}\underline{G}'\underline{P} - \underline{GR}^{-1}\underline{G}'\underline{PG}[\underline{R} + \underline{G}'\underline{PG}]^{-1}\underline{G}'\underline{P} = \underline{I} \qquad \text{(A. 5)}$$

By regrouping terms we get

$$\underline{I} + \underline{G}\{\underline{R}^{-1} - [\underline{R} + \underline{G}'\underline{PG}]^{-1} - \underline{R}'\underline{G}'\underline{PG}[\underline{R} + \underline{G}'\underline{PG}]^{-1}\}\underline{G}'\underline{P} = \underline{I} \qquad \text{(A. 6)}$$

But since $[\underline{R} + \underline{G}'\underline{PG}]^{-1}$ exists, we can write

$$\underline{I} + \underline{G}\{\underline{R}^{-1}[\underline{R} + \underline{G}'\underline{PG}] - \underline{I} - \underline{R}^{-1}\underline{G}'\underline{PG}\}[\underline{R} + \underline{G}'\underline{PG}]^{-1}\underline{G}'\underline{P} = \underline{I} \qquad \text{(A. 7)}$$

The bracketed term is the zero matrix, hence

$$\underline{I} = \underline{I} \qquad \text{(A. 8)}$$

proving that $\underline{I} - \underline{G}[\underline{R} + \underline{G}'\underline{PG}]^{-1}\underline{G}'\underline{P}$ is indeed the inverse given in (A. 3).

APPENDIX B

STABILITY

In the design of any control system, the question of stability is of paramount importance. For this reason, the stability of control systems synthesized using the theory of Chapters II and III is considered here briefly. For simplicity we shall consider first the continuous time system and use the second method of Lyapunov.

For the unperturbed control system (i. e., $\underline{z}(t) \equiv \underline{0}$), the value function (3. 24) is positive definite, provided $\underline{f} = \underline{0}$ and $\underline{h} = \underline{0}$ when $\underline{x}(t) = \underline{0}$ and $\underline{u}(t) = \underline{0}$. In addition $V(\underline{x}(t), t)$ approaches infinity as $\underline{x}(t)$ approaches infinity.

The derivative of V with respect to time along an optimal trajectory is given by

$$\dot{V}(\underline{x}(t), t) = V_t + V'_{\underline{x}} \dot{\underline{x}}(t) \tag{B. 1}$$

or, by (3. 19),

$$\dot{V}(\underline{x}(t), t) = -\frac{1}{2} \| \underline{h}(\underline{x}(t), t) \|^2_{\underline{Q}(t)} - \frac{1}{2} \| \underline{u}_{min}(t) \|^2_{\underline{R}(t)} \tag{B. 2}$$

The right-hand side of equation (B. 2) is non-positive definite. A function which posses these properties is called a Lyapunov function, and the second method of Lyapunov states that when a Lyapunov function exists for a system, the system is stable.

As a matter of fact, $\dot{V}(\underline{x}(t), t)$ is usually negative definite, although it is difficult to give general conditions under which this is true. In this case the second method of Lyapunov guarantees that the system will be asymptotically stable.

For the discrete time control system, analogous results can be drawn using a discrete version of the second method of Lyapunov.

APPENDIX C

COMPUTATIONAL CONSIDERATIONS

C. 1 Computer Storage Requirements

The discrete time problem is analyzed in this section to determine computer storage requirements, and in the next section to determine computer time requirements. Because we are attempting to get approximate answers, many simplifying assumptions will be made.

The first assumption we will make is that we are interested in computing the optimum control only. For instance we are not interested in computing a(k). By considering equations (2. 23), (2. 27), (2. 31), and (2. 32), we can determine the computer storage requirements for the iterative algorithm of section 2. 4. These requirements are given in Table I.

Table I

Variables	Number of Registers Required
$\underline{P}(k)$	$\frac{1}{2}n(n+1)N$
$\hat{\underline{x}}(k)$	nN
$\underline{x}(k)$	nN
$\underline{u}(k)$	rN
Total	$\left[\frac{1}{2}n(n+5)+r\right]N$

Assuming a single input system, that is, r = 1, and for a computer with 50, 000 registers, the dimension of n must be less than 5 and N = 1000 in order to fit the problem on the computer. For the same computer with N = 100, the dimension of n must be less than 20. Even from this quick look into the storage requirements aspect of the problem, we can immediately see that the method is going to be severely restricted by the size of present day computers.

C. 2 Time Requirements

For computer time requirements, we will determine the total number of mathematical operations involved in one iteration of the algorithm. We will assume that all operations require the same amount of time. The total time required can then be determined by multiplying the total number of operations by the average time required per operation. In addition, to simplify matters more, we will assume that the input $\underline{u}(k)$ is a scalar (i. e., $r = 1$), and enters in only one component of $\underline{f}$.

Table II was determined by examination of the same equations as were used in determining Table I.

Table II

Variables	Number of Operations
$\underline{P}(k)$	$\frac{1}{2}n(n+1)(7n^2+3m^2)N$
$\hat{\underline{x}}(k)$	$n(7n^2+3m^2)N$
$\underline{x}(k)$	$2n(n+m)N$ (estimated)
$\underline{u}(k)$	$(3n^2+2n)N$
Total	$\left[5n^2+2n(m+1)+\frac{1}{2}n(n+3)(7n^2+3m^2)\right]N$

As an example, suppose $N = 1000$ and $n = m = 10$. The total number of operations would be on the order of 6×10^7. If the computer could process, on the average, one operation every ten microseconds, the total time required for one iteration would be about ten minutes. Again the limitations of this algorithm using present day computers becomes plainly evident.

As a second example suppose $N = 1000$ but $n = m = 5$. Then the total number of operations required for one iteration would be on the order of 4.5×10^6. At a computer speed of one operation every ten microseconds, this would require about 45 seconds per iteration.

These figures are somewhat conservative because they neglect the time saving possible when repeated factors are encountered. Nevertheless, the figures agree in order of magnitude with the times observed on actual computer problems. (The actual computer times are about one-half to two-thirds of that predicted.)

From these example problems, we can conclude that a problem with 5 state variables and $N = 1000$ steps, represents about the largest size problem that can be handled by this algorithm with presently available computers.

APPENDIX D

FORTRAN PROGRAMS FOR TWO STATE-VARIABLE EXAMPLES

```
   NONLINEAR CONTROLLER
   DIMENSION X1(1000), X2(1000), Z1(1000), Z2(1000), U(1000),
  1XE1(1000), XE2(1000), P11(1000), P12(1000), P22(1000)
   COMMON X1, X2, Z1, Z2, U, XE1, XE2, P11, P12, P22, F1, F2, FX11,
  1FX12, FX21, FX22, FU, R, Q1, Q2, K
20 DO 1 K=1,1000
 1 X1(K) = 0.0
 2 READ 3, X1(1), X2(1), Z1(1), Z2(1), R, Q1, Q2, KF, ITYPE
 3 FORMAT (7F5.2, 2I5)
   TEST = 0.0
   DO 7 I=1,10
   V = Q1*(Z1(1) - X1(1))*(Z1(1) - X1(1)) + Q2*(Z2(1) - X2(1))*
  1(Z2(1) - X2(1))
   X1TEM = X1(1)
   X2TEM = X2(1)
   DO 4 K=1,KF
   K = K
   CALL NONLIN
   R1 = F1 - FX11*X1(K) - FX12*X2(K)
   R2 = F2 - FX21*X1(K) - FX22*X2(K) - FU*U(K)
   RINV = 1.0/(R + FU*P22(K+1)*FU)
   PHX2 = P12(K+1)*R1 + P22(K+1)*R2 + XE2(K+1)
   X1(K) = X1TEM
   X2(K) = X2TEM
   U(K) = -RINV*FU*((P12(K+1)*FX11 + P22(K+1)*FX21)*X1(K) +
  1(P12(K+1)*FX12 + P22(K+1)*FX22)*X2(K) + PHX2)
   X1TEM = FX11*X1(K) + FX12*X2(K) + R1
```

```
      X2TEM = FX21*X1(K) + FX22*X2(K) + FU*U(K) + B2
    4 V = V + Q1*(Z1(K+1)-X1TEM)*(Z1(K+1)-X1TEM) + Q2*(Z2(K+1)-X2TEM)*
     1(Z2(K+1)-X2TEM) + R*U(K)*U(K)
      PCT = (V - TEST)/V
      IF (ABSF(PCT) - 0.01) 8,8,5
    5 TEST = V
      PRINT 6, V
    6 FORMAT (10X, 3HV =1PE15.4)
      DO 7 J =1,KF
      K = KF + 1 - J
      CALL NONLIN
      B1 = F1 - FX11*X1(K) - FX12*X2(K)
      B2 = F2 - FX21*X1(K) - FX22*X2(K) - FU*U(K)
      RINV = 1.0/(R + FU*P22(K+1)*FU)
      R1FU = FU*RINV*FU
      FA11 = FX11
      FA12 = FX21 - R1FU*(FX11*P12(K+1) + FX21*P22(K+1))
      FA21 = FX12
      FA22 = FX22 - R1FU*(FX12*P12(K+1) + FX22*P22(K+1))
      PBX1 = P11(K+1)*B1 + P12(K+1)*B2 + XE1(K+1)
      PBX2 = P12(K+1)*B1 + P22(K+1)*B2 + XE2(K+1)
      XE1(K) = FA11*PBX1 + FA12*PBX2 - Q1*Z1(K)
      XE2(K) = FA21*PBX1 + FA22*PBX2 - Q2*Z2(K)
      P11(K) = Q1 + FA11*(P11(K+1)*FX11 + P12(K+1)*FX21) +
     1 FA12*(P12(K+1)*FX11 + P22(K+1)*FX21)
      P12(K) = FA11*(P11(K+1)*FX12 + P12(K+1)*FX22) + FA12*(P12(K+1)*
     1FX12 + P22(K+1)*FX22)
    7 P22(K) = Q2 + FA21*(P11(K+1)*FX12 + P12(K+1)*FX22) +
     1FA22*(P12(K+1)*FX12 + P22(K+1)*FX22)
```

```
8 PRINT 9, ITYPE, R, Q1, Q2, V, (X1(K), X2(K), Z1(K), Z2(K), U(K),
 1XE1(K), XE2(K), P11(K), P12(K), P22(K), K=1,N)
9 FORMAT (1H1, 14X, 25HTWO STATE-VARIABLE SYSTEM//15X, 25HTHIS IS NO
 1NLINEARITY TYPEI5/15X,2HR=F5.2,4H Q1=F5.2,4H Q2=F5.2///15X, 9HFINA
 2L V =1PE15.4///10X, 2HX1, 8X, 2HX2, 8X, 2HZ1, 8X, 2HZ2, 8X, 1HU,
 39X, 3HXE1, 7X, 3HXE2, 7X, 3HP11, 7X, 3HP12, 7X, 3HP22//
 4(1PE10.3, 1P9E10.3))
  GO TO 20
  END
```

```
      SUBROUTINE NONLIN
      DIMENSION X1(1000), X2(1000), Z1(1000), Z2(1000), U(1000),
     1XE1(1000), XE2(1000), P11(1000), P12(1000), P22(1000)
      COMMON X1, X2, Z1, Z2, U, XE1, XE2, P11, P12, P22, F1, F2, FX11,
     1FX12, FX21, FX22, FU, R, Q1, Q2, K
      F1 = X1(K) + 0.01*X2(K)
      F2 = X2(K) - 0.02*X1(K) - 0.03*ABSF(X2(K))*X2(K) + 0.01*U(K)
      FX11 = 1.0
      FX12 = 0.01
      FX21 = -0.02
      FX22 = 1.0 - 0.06*ABSF(X2(K))
      FU = 0.01
      IF (K - 50) 1,2,2
    1 Z1(K) = 0.0
      RETURN
    2 Z1(K) = 1.0
      RETURN
      END
```

```
      SUBROUTINE NONLIN
      DIMENSION X1(1000), X2(1000), Z1(1000), Z2(1000), U(1000),
     1XE1(1000), XE2(1000), P11(1000), P12(1000), P22(1000)
      COMMON X1, X2, Z1, Z2, U, XE1, XE2, P11, P12, P22, F1, F2, FX11,
     1FX12, FX21, FX22, FU, R, Q1, Q2, K
      DEN = 1.0 + ABS(X2(K))
      F1 = X1(K) + 0.01*X2(K)/DEN
      F2 = X2(K) - 0.01*X1(K) + 0.01*U(K)
      FX11 = 1.0
      FX12 = 0.01/(DEN*DEN)
      FX21 = -0.01
      FX22 = 1.0
      FU = 0.01
      RETURN
      END
```

REFERENCES

1. R. E. Bellman, Stability Theory of Differential Equations, McGraw-Hill Book Company, New York, 1953.

2. R. E. Bellman, Introduction to Matrix Analysis, McGraw-Hill Book Company, New York, 1960.

3. R. E. Bellman, Adaptive Control Processes: A Guided Tour, Princeton University Press, Princeton, New Jersey, 1961.

4. R. E. Kalman and R. W. Koepcke, "Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indices," ASME Transactions, Vol. 80, November 1958, pp. 1820-1826.

5. R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems," J. Basic Engr. (ASME Trans.), Vol. 82-D, 1960, pp. 35-45.

6. R. E. Kalman, "New Methods and Results in Linear Prediction and Filtering Theory," RIAS TR 61-1, Baltimore, Maryland, 1961.

7. R. E. Kalman, "The Theory of Optimal Control and the Calculus of Variations," RIAS TR 61-3, Baltimore, Maryland, 1961.

8. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mischenko, The Mathematical Theory of Optimal Processes, John Wiley and Sons, Inc., New York, 1962.

9. G. C. Newton, L. A. Gould, and J. F. Kaiser, Analytical Design of Linear Feedback Controls, John Wiley and Sons, Inc., New York, 1957.

10. S. E. Dreyfus, "Dynamic Programming and the Calculus of Variations," Journal of Mathematical Analysis and Applications, Vol. 1, 1960, pp. 228-239.

11. R. E. Bellman, Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

12. R. E. Bellman and S. E. Dreyfus, Applied Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1962.

13. A. T. Fuller, "Bibliography of Optimum Nonlinear Control of Deterministic and Stochastic-definite Systems," Journal of Electronics and Control, Vol. XIII, No. 6, December 1962.

14. M. Athans, P. Falb, and R. T. Lacoss, "Time-Optimal Velocity Control of a Spinning Space Body," IEEE Transactions on Applications and Industry, No. 67, July 1963, pp. 206-214.

15. C. W. Merriam, III, "Synthesis of Adaptive Control," MIT Servomechanisms Laboratory Report 7793-R-3, 1959.

16. C. W. Merriam, III, "An Optimization Theory for Feedback Control System Design," Information and Control, Vol. 8, 1960, pp. 32-59.

17. J. E. Bertram, "On Optimal Computer Control," Automatic and Remote Control, Proceedings of the First International Congress, IFAC, Moscow 1960, Butterworth and Co., London, 1961.

18. D. S. Adorno, "Optimum Control of Certain Linear Systems with Quadratic Loss," Information and Control, Vol. 5, No. 1, March 1962, pp. 1-12.

19. M. Athanassiades, "Optimal Control for Linear Time-Invariant Plants with Time, Fuel, and Energy Constraints," IEEE Transactions on Applications and Industry, No. 64, January 1963, pp. 321-325.

20. T. L. Gunckel, II, "Optimum Design of Sample Data Systems with Random Parameters," Stanford Electronics Laboratory Report No. 2102-2, April 1961.

21. T. L. Gunckel, II, and G. F. Franklin, "A General Solution for Linear, Sampled-Data Control," Preprints for the 1962 JACC.

22. H. Cox, "Estimation of State Variables for Noisy Dynamic Systems," ScD. Thesis, Department of Electrical Engineering, Massachusetts Institute of Technology, May 1963.

23. W. Kipiniak, Dynamic Optimization and Control; A Variational Approach, John Wiley and Sons, Inc., New York, 1961.

24. J. D. Pearson, "Approximation Methods in Optimal Control," Journal of Electronics and Control, Vol. XIII, No. 5, November 1962, pp. 453-470.

25. T. G. Windeknecht, "Optimal Stabilization of Rigid Body Attitude," Journal of Mathematical Analysis and Applications, Vol. 6, No. 2, May 1963, pp. 325-335.

BIOGRAPHICAL NOTE

Richard B. Gilchrist was born in Lehi, Utah, on October 10, 1932. In December, 1955, he married Janice Chipman of American Fork, Utah. They have four children, Karen 7, Katy 5, Richard 3, and Geaniel 1-1/2.

Lieutenant Gilchrist graduated from Lawton High School, Lawton, Oklahoma. He attended the University of Utah, Salt Lake City, Utah, for one year before entering the United States Naval Academy, Annapolis, Maryland, where he graduated in 1955 with a Bachelor of Science degree.

After graduation, he served for one year aboard the U. S. S. SKAGIT (AKA-105). He then attended submarine school, and served aboard the U. S. S. TUNNY (SSG-282) until June, 1959, when he entered the Massachusetts Institute of Technology.

Lieutenant Gilchrist has recently been designated an Engineering Duty Officer. He is a member of the Institute of Electrical and Electronics Engineers.