Negative microfilm
received from university

66–5044

CHAN, Shu–Yun, 1930–
ANALYSIS OF VARIATIONS IN OPTIMAL CONTROL SYSTEMS SUBJECT TO DE–TERMINISTIC AND STOCHASTIC DIS–TURBANCES.

The University of Michigan, Ph.D., 1965
Engineering, electrical

University Microfilms, Inc., Ann Arbor, Michigan

ANALYSIS OF VARIATIONS IN OPTIMAL CONTROL SYSTEMS SUBJECT TO DETERMINISTIC AND STOCHASTIC DISTURBANCES

by
Shu-Yun Chan

A dissertation submitted in partial fullfillment
of the requirements for the degree of
Doctor of Philosophy in the
University of Michigan
1965

Committee in charge:

Associate Professor Kuei Chuang, Co-Chairman
Professor Louis F. Kazda, Co-Chairman
Professor Chiao-Min Chu
Professor Louis J. Cutrona
Professor Gunnar Hok
Professor Wilfred Kaplan

ACKNOWLEDGEMENTS

The author wishes to express his deepest gratitude to Professor Kuei Chuang for his guidance, invaluable suggestions constructive criticism and encouragement throughout the preparation of this thesis. Acknowledgement is given to Professor Louis F. Kazda for his helpful and stimulating discussions on the material in this work; to Professors Hok and Wilfred Kaplan for their invaluable comments and advice; to Professors Chiao-Min Chu and Louis J. Cutrona for their encouragement and help in various ways.

The author is indebted to the National Science Foundation for its generous support under the grant GP-1381.

TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS	ii
LIST OF ILLUSTRATIONS	vi
LIST OF SYMBOLS	ix
LIST OF TABLES	viii
CHAPTER I - INTRODUCTION	1
1.1 Historical Review of Prior Work	1
1.2 Research Objectives	5
CHAPTER II ANALYSIS OF VARIATIONS IN TRAJECTORY AND COST OF OPEN-LOOP OPTIMAL CONTROL SYSTEMS DUE TO DETERMINISTIC CHANGES OF PLANT PARAMETERS	8
2.1 Specific Problem and Discussion	8
2.2 Evaluation of Change in Cost and of Variation in Trajectory Due to Changes in Plant Parameters	10
2.3 Error Introduced by Linearization	13
2.4 Inverse Sensitivity Problem of Determining Actual Variation of Parameters from known Trajectory	23
2.5 Estimation of Variations of Time Dependent Plant Parameters I	30
2.6 Estimation of Variations of Time Dependent Parameters II	47
CHAPTER III ANALYSIS OF VARIATION IN TRAJECTORY AND COST OF CLOSED LOOP OPTIMAL CONTROL SYSTEMS DUE TO DETERMINISTIC CHANGES OF PLANT PARAMETERS	59
3.1 Specific System	59
3.2 Evaluation of Change in Cost and of	

TABLE OF CONTENTS (Con't)

		Page
	Variation in Trajectory Due to Changes in Plant Parameters	61
	3.3 Estimation of Plant Parameters with Known Trajectory	64
CHAPTER IV	THE FOKKER-PLANCK EQUATION	66
	4.1 Probabilistic Characterization of a Markov Process	66
	4.2 Definition of a Continuous Process	68
	4.3 The Fokker-Planck Equation	69
	4.4 The Fokker-Planck Equation Corresponding to a Linear Time-Varying System Excited by White Noise	70
	4.5 Reduction of the Fokker-Planck Equation to Solving a Set of Ordinary Differential Equations with Variable Coefficients	74
	4.6 Linear Time-Varying System with Disturbances Generated by Differential Equations Excited by White Noise	89
CHAPTER V	DETERMINATION OF AVERAGE CHANGES IN TRAJECTORY AND IN COST OF OPTIMAL CONTROL SYSTEMS DUE TO RANDOM DISTURBANCES	94
	5.1 Systems Excited by White Noise Disturbances in Plant Parameters	94
	5.2 Systems Subject to Disturbances Generated by Differential Equations Excited by White Noise in Plant Parameters	96
	5.3 Estimation of Changes in Trajectory and Cost due to Noisy Observation of State Variables	97

TABLE OF CONTENTS (Con't)

		Page
	5.4 Application of the above Method for Two Special Classes of Optimal Control Systems	99
	5.5 Stability Considerations	106
CHAPTER VI	EXPERIMENTAL VERIFICATIONS	109
	6.1 Measurement of the Power Spectral Density of the Random Noise Generator	109
	6.2 An Optimal Control System	116
	6.3 Minimum Energy Control System Disturbed by Random Noise	123
CHAPTER VII	CONCLUSIONS AND FUTURE RESEARCH	147
	7.1 Conclusions	147
	7.2 Suggestions for Future Research	150
APPENDIX A	DERIVATION OF (2.22)	151
APPENDIX B	VERIFICATION OF (2.36)	153
APPENDIX C	DERIVATION OF THE FOKKER-PLANCK EQUATION	155
APPENDIX D	DERIVATION OF (5.17)	159
APPENDIX E	DERIVATION OF (5.31)	163
APPENDIX F	DERIVATION OF (5.38)	165
APPENDIX G	DERIVATION OF (6.19)	167
BIBLIOGRAPHY		170

LIST OF ILLUSTRATIONS

Figure	Title	Page
1.1	Basic Open-Loop System	5
1.2	Basic Idealized Closed-Loop System	6
2.1	Basic Open-Loop System	9
2.2	Rotational Body in Space	54
3.1	Basic Closed-Loop System	60
5.1	Closed-Loop System with Random Disturbances	97
6.1	Block Diagram for Measurement of Standard Deviation	111
6.2	Computer Setup for the Measurement of Power Spectral Density of the Output of Random Noise Generator at 10 and 20 Radians per Second	112
6.3	Computer Setup for the Measurement of Power Spectral Density of the Output of Random Noise Generator at 40, 60, 80 and 100 Radians per Second	114
6.4	Power Spectrum of Noise Source	115
6.5	Vehicle and Reaction Wheel with a D.C. Servomotor	116
6.6	Computer Setup of the System (6.22)	121
6.7	Minimum Energy Control for D.C. Motor Reaction Wheel System	122
6.8	Computer Setup for the Simulation of the System (6.29) for the Generation of the Matrix Ψ	126
6.9	Experimental Determination of the Elements of the Matrix Ψ	127
6.10	Computer Setup for the Simulation of the System (6.24)	129
6.11	Experimental Determination of the Elements of the Matrix $\bar{\Phi}$	130

LIST OF ILLUSTRATIONS (Con't) Page

6.12 The Means of the System (6.24) as a Function of Dimensionless Time 133

6.13 APPlot of $\frac{f_{22}}{D_1} \psi_{21}^2$ versus Dimensionless Time τ 135

6.14 A Plot of $\frac{f_{22}}{D_1} \psi_{21} \psi_{22}$ versus Dimensionless Time τ 136

6.15 A Plot of $\frac{f_{22}}{D_1} \psi_{22}^2$ versus Dimensionless Time τ 137

6.16 The Variances of the System (6.16) as a Function of Dimensionless Time 139

6.17 A Typical Sample of the Output of the System (6.24) with Noise Disturbance n_1 144

LIST OF SYMBOLS

Symbol	Page Defined	Symbol	Page Defined
T	2	$\|x\|$	13
S_p^T	2	$[x,y]$	13
P	2	E_n	14
$a(t)$	3	S	10
L	4	t_o	9
$\alpha_k(t)$	4	t_1	14
$\delta(\tau)$	4	$\lceil$	14
E	4	π	14
$\ddot{x}$	9	K	14
α	9	K_1	14
α_o	10	K_2	14
u	9	K_3	14
f	8	K_4	15
f_i	8	K_o'	15
f_o	9	K_o	16
J	9	C	16
φ	10	ρ	17
g_i	4	F^i	17
g	11	G^i	17
F	11	H^i	17
G	11	K^i	17
Φ	12	E^i	17
I	12	ε_1	18
$\|x\|$	13	$\|M(t)\|$	19

Symbol	Page Defined	Symbol	Page Defined
M_1	20	$k(t,\tau)$	33
M_2	20	γ_o	33
F_o	21	L^*	33
ε_2	21	λ_i	34
M_3	22	ϕ_i	35
M_4	22	e_i	35
ϕ_{ij}	24	$R(L^*)$	35
g_{ik}	24	$\oplus$	35
$\lambda_{ik}(t)$	25	$N(L)$	35
$\Delta x^*(\alpha_o,t)$	25	h	36
$\Delta\alpha_k^a$	25	h_N	37
$\Delta\alpha_k^b$	25	Δx_1	36
$\Delta\alpha_k^c$	26	$Q(\gamma)$	38
$\{\beta_k(t)\}$	26	$Q_k(\gamma)$	40
$\{y_i(t)\}$	27	Q_o	41
a_{ji}	29	K_o	41
ξ	31	z_n	46
H^n	31	ϵ_n	46
$\gamma(t)$	32	γ_N	46
$x'(t)$	31	$F_i(\alpha,t)$	47
L	32	$h(\alpha)$	48
H^p	32	R	48
C	31	$K(\alpha)$	48
$x_a(t)$	32	u	48
$\Delta x_{ao}(t)$	33	M	48
δ	33	$\Delta k(\alpha)$	48

Symbol	Page Defined	Symbol	Page Defined
$\phi(\lambda)$	49	$\delta(t_1 - t_2)$	71
τ_i	54	$f_{ij}(t)$	74
I_i	54	$g(\xi,t)$	75
I	54	$\psi(t_1 t_o)$	77
$\theta(t)$	55	$\Phi(t_1 t_o)$	78
$f(t)$	55	Γ	79
$f_1(t)$	55	$F(\xi,t)$	79
$f_2(t)$	55	$K_{rs}(t)$	79
$f_3(t)$	55	$m_j(t)$	86
C	54	$\lambda_{rs}(t)$	87
F	56	λ	87
$x^*(t)$	57	$\beta_{ij}(t)$	90
$u^*(t)$	57	w_i	90
$A(t)$	62	$E(t)$	92
$B(t)$	62	$F(t)$	93
$P(t,x/\tau,y)$	66	$N(f)$	109
$F(t,x/S_g Z^i)$	68	$Y(s)$	110
Z^i	67	σ	109
$\overline{Z}$	68	f_o	110
$a_i(t,x)$	69	ω_o	110
$b_{ij}(t,x)$	69	L	116
		θ	116
$n(t)$	70	ω	116
$C(t)$	70	I	117
$D(t)$	70	J_w	116
D_i	70	J	116

Symbol	Page Defined
K_L	117
K_m	117
R_a	117
τ_m	117
$V(t)$	117
$u(t)$	117
τ	117
e_1	118
e_2	118
$\bar{m}_1(t)$	146

LIST OF TABLES

Table	Title	Page
6.1	Measurement of Variance of Filter Output and Calculation of Power Spectrum of Noise Source	113
6.2	Tabulation of x_2, ϕ_{11}, ϕ_{21}, ϕ_{12}, ϕ_{22}, Ψ_{21} and Ψ_{22} at Different Values of Dimensionless Time	131
6.3	Theoretical and Experimental Values of Means of the System (6.24)	132
6.4	Tabulation of $\frac{f_{22}}{D_1}\Psi_{21}^2 \times 10^3$, $\frac{f_{22}}{D_1}\Psi_{21}\Psi_{22} \times 10^3$, $\frac{f_{22}}{D_1}\Psi_{22}^2 \times 10^3$ and their Integrals at Different Values of Dimensionless Time	134
6.5	Theoretical and Experimental Values of Variances of the System (6.24)	138
6.6	Tabulation of Values of Z_1, and Z_2 of the System (6.24) taken from Experimental Curves	143

Chapter I

INTRODUCTION

1.1 Historical Review of Prior Work

Little attention was paid to the subject of optimal control systems prior to 1950. During the latter part of the decade the analysis and synthesis of optimal control systems has been extensively studied both in this country and in Russia. Since 1950 a considerable body of theory has been developed. Pontryagin's maximum principle[1] and Bellman's dynamic programming[2] are well known in the field of control systems. Other variational techniques[3,4,5] were also modified so that they could be applied to optimal control systems. Direct effective computational techniques for solving these systems were not developed until recently. Kelley[6,7] and Bryson[8] have independently revived the method of steepest descent in solving open-loop control systems. Neustadt[9] and Ho[10] also developed computational techniques for finding the optimal control law by using the known properties of state transition matrix in the case of linear dynamic systems with inequality constraints on the control function. Ho[11] further extended the method to nonlinear systems with state variable constraint by introducing slack variables into the system. Recently functional analysis has been used to attack optimal control problems. Krassovskii[12] first observed that certain properties of functionals in normed linear spaces are applicable to time optimal control problems. Kulikowski[13] used results from functional anaysis in problems dealing with nonlinear plants and his work was further extended by Kranc and Sarachik[14] to linear systems with multiple inputs.

In order to analyze the behavior of a given physical system it is necessary to select a mathematical form which will represent the physical system adequately. This mathematical form is in general described by

nonlinear differential equations. To simplify the analysis of the behavior of the given system, assumptions are usually made such as replacing the nonlinear terms by their linear approximations at the point of interest so that linear methods can be applied. Distributed parameters are further replaced by lumped parameters and time-varying parameters are replaced by constant ones so that transfer function technique in complex frequency domain can be used. Therefore, the behavior of the physical system, in general, cannot be assumed to be identical to the solution of its mathematical model. Furthermore, manufacturing tolerances of components, changes in component values over periods of time caused by environmental factors such as temperature, shock and radiation, and noise in the system are some of the factors causing the actual physical system to behave differently from that of the idealized one.

For linear time invariant systems some attention has been paid to the evaluation of the deterioration of the system performance due to changes in components of the plant. The definition ofasensitivity function was first introduced in feedback systems by Bode[15]. The commonly used definition of sensitivity function S_p^T, denoting the sensitivity of the system transfer function T to the variations in the plant parameter P, is defined as[16,17]

$$S_p^T = \frac{dT/T}{dP/P} \tag{1.1}$$

However, for linear time-varying and nonlinear systems the above definition cannot be applied to the evaluation of system deterioration. Hence a new sensitivity function has been defined which denotes the change in the system response due to the component inaccuracies, and is given by[18]

$$\left[\frac{\partial x(t,\alpha)}{\partial \alpha}\right]_{\alpha = \alpha_o} , \tag{1.2}$$

where $x(t,\alpha_o)$ represents the response of the nominal system and α_o is the nominal value of the component. This new definition of sensitivity function can be applied to all systems.

The stability and sensitivity of dynamic systems are closely related. For example with the definition of sensitivity function given in (1.2) it is clear that for any n^{th} order linear differential equation with constant coefficients the differential equation of the sensitivity has the same characteristic equation as the original differential equation. Since the characteristic equation is used to study the system's stability, this implies that there exists a connection between a dynamic system's stability and its sensitivity equation.

Some sensitivity analyses has been performed in the case of variation of system parameters. Huang[19] studied the changes in the pole positions of linear closed-loop systems due to small variations of a parameter. Miller and Murray[20] developed a method for analyzing the errors in the output of an analog computer due to errors in simulation.

Linear randomly varying systems play an important role in problems involved in the study of optimal control systems. Very few attempts were made at studying the behavior of randomly varying systems prior to 1950[21]. The analysis of a first order system

$$\dot{y} + a(t)y = u(t) \tag{1.3}$$

in which $a(t)$ is a randomly varying coefficient and $u(t)$ a nonrandom input was studied by Rosenbloom[22]. He obtained explicit results in the case where the random variation of $\{a(t)\}$ is a Gaussian process. For higher order systems some results were obtained by Samuels and Eringer[23,24]. For the system of the form

$$Ly + \alpha_k(t) \frac{d^k y}{dt^k} = u(t) \tag{1.4}$$

in which L is a linear differential operator with constant coefficients, $\alpha_k(t)$ is a randomly varying parameter with correlation function $\delta(\tau)$, and u(t) is the random input, they obtained an explicit expression for the expectation $E\left\{y^2(t)\right\}$ as a function of time, which enabled them to determine whether the system of the form (1.4) is stable in mean square. That is, if

$$\lim_{t\to\infty} E\left[y^2(t)\right] < M \tag{1.5}$$

with M being a finite constant. The concept of stability in the case of a system with randomly varying parameters described by the vector differential equation

$$\dot{x} = f(x,\alpha(t),t)$$

in which $\alpha(t)$ denotes a set of randomly varying parameters was given by Bertram and Sarachik[25]. They showed that sufficient conditions for stability in the mean square can be obtained by extending in stochastic sense Lyapunov's second method. Similiar work was also carried out by Kats and Krasovskii[26] in U. S. S. R.

In the investigation of technical literature on control theory the sensitivity analysis of optimal control systems in their general forms is almost nonexistent to the author's knowledge. In his correspondence in IEEE Transactions on Automatic Control Dorato[27] emphasized the importance of sensitivity analysis of optimal control systems and outlined a procedure to study the changes in system behavior due to small parameter variations. Gavrilovic and others[28] have devised a computational scheme for the determination of small perturbations along the optimal trajectory in the state space when the system is subject to disturbances. As pointed out by Wonham[29] there is no workable theory

for the changes in system behavior when the observations of the state of the system are mixed with noise.

1.2 Research Objectives

It is intended in this research to investigate in detail the variations of system behavior of both open- and closed-loop optimal control systems caused by changes in parameters of the plant in a deterministic or a stochastic sense or caused by the presence of noise in the observations of the state vector. The noise vector may be a white noise vector or it may be generated by a set of differential equations excited by white noise. In both cases only small disturbances are considered.

This work may be classified into two parts (1) open-loop optimal control systems, (2) closed-loop optimal control systems.

(1) Open-Loop Optimal Control Systems

A typical problem treated in this part may be described schematically by Fig. 1.1.

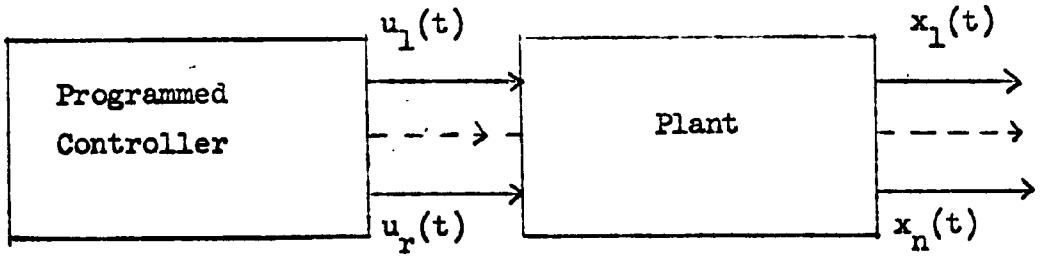


Fig. 1.1 Basic Open-Loop System

The optimal control $u(t) = (u_1(t),\ldots,u_r(t))$ is assumed to be a known vector function of t. The variation of the trajectory of the system, that is, the variation in the output of the plant, $x(t) = (x_1(t),\ldots,x_n(t))$, and hence the change in system performance are determined once the changes in plant parameters are known.

The inverse sensitivity problem of determining the actual variations of parameters of the plant are also investigated if the actual trajectory of the system can be observed. In the case of randomly varying parameters the evaluation of variations of the system behavior is carried out in stochastic sense.

(2) Closed-Loop Optimal Control Systems

A closed-loop system can be represented schematically by Fig. 1.2

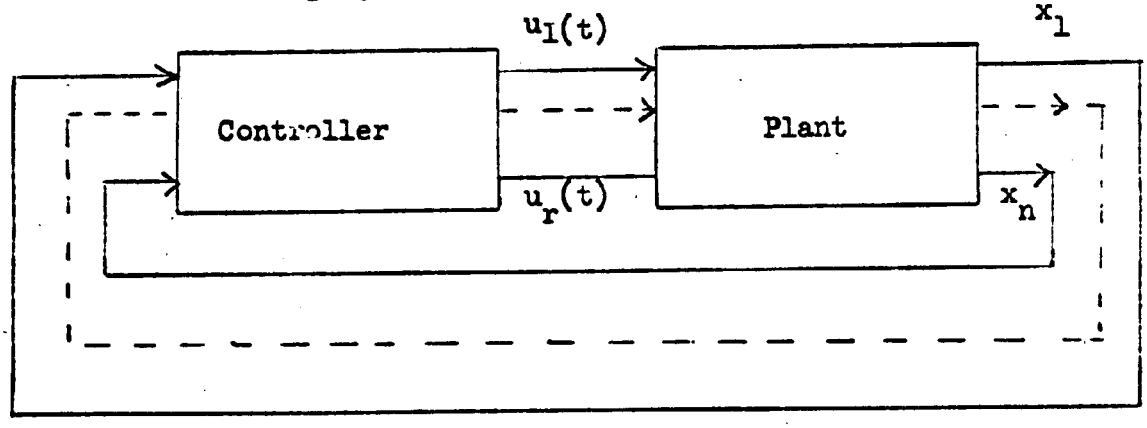


Fig. 1.2 Basic Idealized Closed-Loop System

in which the optimal control $u(t) = (u_1(t),\ldots,u_r(t))$ is assumed to be a known vector function of state vector $x(t) = (x_1(t),\ldots,x_n(t))$. The investigation of variations of the system performance due to parameter variations can be carried out in the same manner as in the open-loop case. However, when the observations of the state vector are mixed with noise due to inaccuracies of sensing the components of the state or due to estimation error in state variables, because certain state variables cannot be obtained directly, the variation of the system behavior under the assumptions given in Chapter 4 can be described as a multi-dimensional Markov process. The method of characteristics is used in solving the Fokker-Planck equation. It is shown in Chapter 4 that the probability density function is Gaussian, hence the average

change along the optimal trajectory and the average change in cost can be determined accordingly.

Chapter II

ANALYSIS OF VARIATIONS OF TRAJECTORY AND COST OF OPEN-LOOP OPTIMAL CONTROL SYSTEMS DUE TO DETERMINISTIC CHANGES OF PLANT PARAMETERS

In this chapter only open loop optimal control systems are considered. The first part of the chapter is devoted to calculating the variations in the trajectory and cost of the system due to the changes in plant parameters about their nominal values. This is done by solving the linearized equation of the system along its nominal trajectory. The error between the actual trajectory and its linear approximation, which is defined as the nominal trajectory plus its first variation, is obtained. The remaining part of the chapter is devoted to the estimation of the variations of plant parameters under the assumption that the actual trajectory can be observed. For the constant plant parameter case the Schmidt orthogonalization process is used. In the case of time-varying plant parameters the method of steepest descent is used.

2.1 Specific Problem and Discussion

An open-loop optimal control system is described schematically in Fig. 2.1 in which the behavior of the plant is assumed to be governed by a set of differential equations

$$\frac{dx_i}{dt} = f_i(x_1,\ldots,x_n,\alpha_1,\ldots,\alpha_p,u_1,\ldots,u_r,t)$$

and (2.1)

$$x_i(t_o) = x_{io} \quad , \qquad i = 1,\ldots,n$$

Equations 2.1 can be expressed in vector form as

$$\frac{dx}{dt} = f(x,\alpha,u,t) \tag{2.2}$$

and

$$x(t_o) = x_o \quad ,$$

where

x = n-dimensional state vector describing the instantaneous behavior of the plant (object to be controlled)

α = p-dimensional vector representing the set of plant parameters

u = r-dimensional control vector

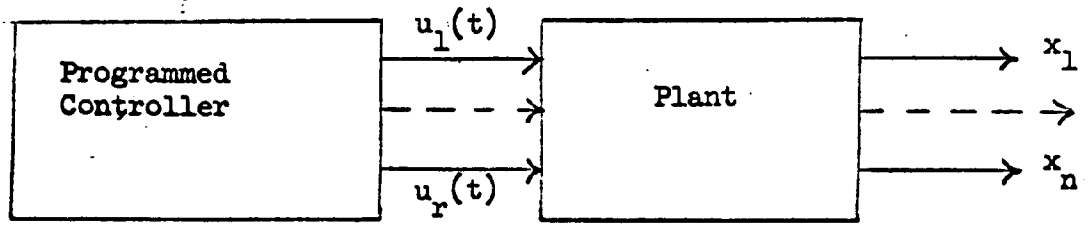


Fig.2.1 Basic Open-Loop System

The cost to be minimized is given by the functional

$$J = \int_{t_o}^{t_1} f_o(x,u)dt \quad , \tag{2.3}$$

where t_1 is the terminal time. It is assumed that the following conditions are satisfied:

a) the functions f_o and $f_i (i = 1,\ldots,n)$ as well as their first partial derivatives with respect to x_j and $\alpha_k (j = 1,\ldots, n, k = 1,\ldots,p)$ are defined and continuous in some domain Γ of the space R of variables $x_1,\ldots,x_n,\alpha_1,\ldots,\alpha_p,u_1,\ldots,u_r,t$. In other words, the space R can be taken to be the Cartesian product $R = X \times A \times U \times T$.

b) the optimal control $u(t)$, which transfers the state of the system (2.2) from a given initial position x_o to the final

position x_1 that is restricted to lie on a given hypersurface $S(x_1) = 0$, and which at the same time minimizes the functional (2.3), exists and is a known piecewise continuous vector function of time t.

Then for $\alpha_o \in A$ there exists a unique solution of the system (2.1) such that

$$x(\alpha_o, t) \triangleq \varphi(\alpha_o, t) = (\varphi_1(\alpha_o, t), \ldots, \varphi_n(\alpha_o, t)) \qquad (2.4)$$

is defined and continuous on some interval containing the point t_o and satisfies the initial condition

$$\varphi(\alpha_o, t_o) = x_o \ ,$$

where α_o is the nominal parameter vector[1]. We shall assume that on the interval $[t_o, t_1]$ the solution is defined.

2.2 Evaluation of Change in Cost and of Variation in Trajectory Due to Changes in Plant Parameters

As discussed in Chapter I the actual plant parameter vector $\alpha = (\alpha_1, \ldots, \alpha_p)$ of the physical system in general cannot be expected to be identical to the nominal vector $\alpha_o = (\alpha_{o1}, \ldots, \alpha_{op})$. As a result the actual trajectory $\varphi(\alpha, t)$ is actually not the optimal trajectory. This causes an error in terminal state and a change in value of the cost expressed by the integral (2.3).

In this research we shall assume that the variation of the parameter vector α from its nominal vector is very small so that the solution of the first variation of (2.2) can be considered as a sufficiently accurate estimate of the changes in optimal trajectory.

Since the optimal control $u(t) = (u_1(t), \ldots, u_r(t))$ is a known piecewise continuous function, upon substitution of u(t)

into (2.1) we find

$$\frac{dx_i}{dt} = g_i(x_1,..,x_n,\alpha_1,...,\alpha_p,t) \qquad i = 1,...,n$$

$$x_i(t_o) = x_{io} \tag{2.6}$$

or in vector form

$$\frac{dx}{dt} = g(x,\alpha,t)$$

$$x(t_o) = x_o \ . \tag{2.7}$$

Expanding the right hand side of (2.6) in a Taylor series about the point $(x(\alpha_o,t),\alpha_o,t)$ and then neglecting the second and higher order terms, the variation along the optimal trajectory satisfies the linear differential equation

$$\frac{d(\Delta x_i)}{dt} = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j} \Delta x_j + \sum_{k=1}^{p} \frac{\partial g_i}{\partial \alpha_k} \Delta \alpha_k \tag{2.8}$$

$$\Delta x_i(t_o) = 0, \qquad i = 1,...,n$$

where the partial derivatives of g_i are evaluated at $(x(\alpha_o t),\alpha_o,t)$.

Writing equation (2.7) in vector form, we find

$$\Delta \dot{x} = F \Delta x + G \Delta \alpha \tag{2.9}$$

$$\Delta x(t_o) = 0,$$

where

$$F = \begin{bmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{bmatrix}, \qquad G = \begin{bmatrix} \frac{\partial g_1}{\partial \alpha_1} & \cdots & \frac{\partial g_1}{\partial \alpha_p} \\ \cdots & \cdots & \cdots \\ \frac{\partial g_n}{\partial \alpha_1} & \cdots & \frac{\partial g_n}{\partial \alpha_p} \end{bmatrix}. \tag{2.10}$$

The differential equation (2.8) is linear and time-varying, its solution may be written explicitly in terms of the transition matrix $\Phi(t,\tau)$, that is, $\Phi(t,\tau)$ is the solution of the matrix equation

$$\dot{\Phi} = F\Phi, \tag{2.11}$$

which satisfies the initial condition

$$\Phi(t_o,t_o) = I \ , \tag{2.12}$$

where I is the identity matrix. Thus the variation along the optimal trajectory is given by

$$\Delta x(\alpha_o,t) = \int_{t_o}^{t} \Phi(t,\tau)\, G(\tau)\Delta\alpha\, d\tau \ , t_1 \geq t \geq t_o \tag{2.13}$$

Similarly, the change in cost can be evaluated approximately by expanding the function f_o into a Taylor series about the point $(x(\alpha_o,t),\alpha_o,t)$ with the second and higher order terms neglected. Denoting the change in cost by

$$\Delta J = \int_{t_o}^{t_1} \sum_{j=1}^{n} \frac{\partial f_o}{\partial x_j} \Delta x_i(\alpha_o,t)\, dt \ . \tag{2.14}$$

It can be seen that once the variation along the optimal trajectory is known, the change in cost as well as the deviation in terminal state can be calculated accordingly.

2.3 Error Introduced by Linearization

In the evaluation of the variation along the optimal trajectory and of the change in cost due to changes in plant parameters we have used the equation of first variation (2.8). In this section we shall find the errors introduced by approximating the actual trajectory by the equation of first variation, namely, the difference between the actual trajectory and $x(\alpha_o,t) + \Delta x(\alpha_o,t)$ and the difference between the actual value of the cost and $J(\alpha_o) + \Delta J$. In order to do so, we need, in addition to assumptions a) and b) which are stated in section 2.1, one additional assumption, namely that c) the partial derivatives of f_o and $f_i (i = 1,\ldots,n)$ with respect to x_j and $\alpha_k (j = 1,\ldots,n, k = 1,\ldots,p)$ of the second order are also continuous. In order to use vector notations freely we shall establish some natural definitions and simple inequalities for vectors and vector functions

Definitions and Inequalities

1) The length of the vector $x(t) = (x_1(t),\ldots,x_n(t))$ is defined by the formula

$$|x(t)| = \left\{ \sum_{i=1}^{n} x_i^2(t) \right\}^{1/2} \qquad t_o \le t \le t_1 \tag{2.15}$$

2) The norm of the vector function x is defined as

$$\|x\| = \underset{t_o \le t \le t_1}{\text{Max}} |x(t)| , \tag{2.16}$$

which is commonly referred to as the Chebyshev metric.

3) The inner product of two vectors x and y is defined as

$$[x,y] = \sum_{i=1}^{n} x_i y_i \tag{2.17}$$

It can be shown that for any two vectors $x(t)$ and $y(t)$ in E_n, the n-dimensional Euclidean space, the following inequalities are satisfied.

3) $|x+y| \leq |x| + |y|$ (2.18)

4) $\left| \int_{t_o}^{t} x(\tau)\, d\tau \right| \leq \int_{t_o}^{t} |x(\tau)|\, d\tau, \quad t > t_o$ (2.19)

The initial values t_o and x_o of the solution of (2.7) corresponding to the nominal parameter vector $\alpha_o = (\alpha_{o1}, \ldots, \alpha_{op})$ are coordinates of the point (x_o, t_o) in Γ. Because the solution $\varphi(\alpha_{o1} t)$ is continuous and passes through Γ we wish to show in the following paragraphs that for any positive numbers a and b such that the set π defined by the relations

$$|x - \varphi(\alpha_{o1} t)| = \left\{ \sum_{j=1}^{n} (x_i - \varphi_i(\alpha_{o1} t))^2 \right\}^{1/2} \leq a \qquad (2.20)$$

$$|\alpha - \alpha_o| = \left\{ \sum_{k=1}^{p} (\alpha_i - \alpha_{oi})^2 \right\}^{1/2} \leq b$$

can be constructed and it is contained in Γ. It is clear that the set π so defined is closed and bounded.. Since the functions f_o and g_i as well as their partial derivatives are piecewise continuous, therefore positive numbers can be found such that for (x,α,t) in π, the following inequalities hold

$$\left| \frac{\partial g_i}{\partial x_j} \right| \leq K \qquad \left| \frac{\partial g_i}{\partial \alpha_s} \right| \leq K_1$$

$$\left| \frac{\partial^2 g_i}{\partial x_i \partial x_j} \right| \leq K_2 \qquad \left| \frac{\partial^2 g_i}{\partial \alpha_s \partial \alpha_j} \right| \leq K_3 \qquad \left| \frac{\partial^2 g_i}{\partial \alpha_k \partial \alpha_j} \right| \leq K_4 \qquad (2.21)$$

$$\left|\frac{\partial f_o}{\partial x_1}\right| \leq K_o' \qquad \left|\frac{\partial^2 f_o}{\partial x_i \partial x_j}\right| \leq K_o$$

$i, j = 1, \ldots, n$

$k, s = 1, \ldots, p$

Then for any two points (x_1, α, t) and (x_2, α, t) of π we can show that the relation

$$|g(x_1, \alpha, t) - g_1(x_2, \alpha, t)| \leq n^2 K |x_1 - x_2| \tag{2.22}$$

is valid (see Appendix A). Similarly we have

$$|g(x, \alpha, t) - g(x, \alpha_o, t)| \leq p^2 K_1 |\alpha - \alpha_o| \tag{2.23}$$

and

$$\begin{aligned} &|g(\varphi(\alpha,\tau),\alpha,\tau) - g(\varphi(\alpha_o,\tau),\alpha_o,\tau)| \\ &\leq |g(\varphi(\alpha,\tau),\alpha,\tau) - g(\varphi(\alpha_o,\tau),\alpha,\tau)| + |g(\varphi(\alpha_o,\tau),\alpha,\tau) - g(\varphi(\alpha_o,\tau),\alpha_o,\tau)| \\ &\leq n^2 K |\varphi(\alpha,\tau) - \varphi(\alpha_o,\tau)| + p^2 K_1 |\alpha - \alpha_o| \end{aligned} \tag{2.24}$$

Using relation (2.24) we find for any $t_o \leq t \leq t_1$

$$|\varphi(\alpha,t) - \varphi(\alpha_o,t)| \leq \int_{t_o}^{t} |g(\varphi(\alpha,\tau),\alpha,\tau) - g(\varphi(\alpha_o,\tau),\alpha_o,\tau)| \, d\tau$$

$$\leq \int_{t_o}^{t} [n^2 K |\varphi(\alpha,\tau) - \varphi(\alpha_o,\tau)| + p^2 K_1 |\alpha - \alpha_o|] d\tau$$

$$\leq \frac{p^2 K_1 |\alpha - \alpha_o|}{n^2 K} \left[e^{n^2 K (t - t_o)} - 1\right]^* \tag{2.25}$$

---

* If $u(t) \leq \int_{t_o}^{t} (\alpha u(t) + \beta) d\tau$, $\tau > 0$, $\beta > 0$ then the relation

$u(t) \leq \frac{\beta}{\alpha} (e^{\alpha(t-t_o)} - 1)$ is valid.

$$\leq \frac{p^2 K_1}{n^2 K} \left[e^{n^2 K(t_1 - t_o)} - 1 \right] \left| \alpha - \alpha_o \right| \qquad (2.25)$$

$$= C \left| \alpha - \alpha_o \right| ,$$

where

$$C = \frac{p^2 K_1}{n^2 K} \left[e^{n^2 K(t_1 - t_o)} - 1 \right] . \qquad (2.26)$$

Let a positive number ρ be chosen such that

$$\left| \alpha - \alpha_o \right| \leq \rho \leq b \text{ and } C\rho \leq a \qquad (2.27)$$

From (2.27) and (2.25) it is obvious that relations (2.20) and hence (2.21) are satisfied.

To find the difference between the actual trajectory and its approximate solution we expand the right hand side of equation (2.6) into a Taylor series about $(x(\alpha_o t), \alpha_o, t)$.

We find

$$\dot{x}_i = g_i(x(\alpha_o t), \alpha_o, t) + \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j} (x_j(\alpha, t) - x_j(\alpha_o, t))$$

$$+ \sum_{k=1}^{+p} \frac{\partial g_i}{\partial \alpha_k} (\alpha_k - \alpha_{ok}) + \frac{1}{2} \sum_{l,j=1}^{n} \frac{\partial^2 g_i}{\partial x_l \, \partial x_j} (x_j(\alpha_l t) - x_j(\alpha_o, t))(x_l(\alpha, t) - x_l(\alpha_o, t))$$

$$+ \sum_{j=1}^{n} \sum_{k=1}^{p} \frac{\partial^2 g_i}{\partial \alpha_k \, \partial x_j} (\alpha_k - \alpha_{ok})(x_j(\alpha, t) - x_j(\alpha_o, t))$$

$$+ \frac{1}{2} \sum_{k,s=1}^{p} \frac{\partial^2 g_i}{\partial \alpha_k \, \partial \alpha_s} (\alpha_k - \alpha_{ok})(\alpha_s - \alpha_{os}) , \qquad i = 1, \ldots, n \qquad (2.28)$$

where the second partial derivatives of g_i are evaluated at the point

$$\beta(x(\alpha_o,t),\alpha_o,t) + (1-\beta)(x(\alpha,t),\alpha,t),\ 0 \le \beta \le 1.$$

As a matter of convenience let us define for $i = 1,\ldots, n$

$$F^i = \left(\frac{\partial g_i}{\partial x_1},\ldots,\frac{\partial g_i}{\partial x_n}\right),\quad G^i = \left(\frac{\partial g_i}{\partial \alpha_1},\ldots,\frac{\partial g_i}{\partial \alpha_p}\right) \tag{2.29}$$

$$H^i = \left[\frac{\partial^2 g_i}{\partial x_j\,\partial x_\ell}\right],\quad K^i = \left[\frac{\partial^2 g_i}{\partial \alpha_k\,\partial x_j}\right],\quad E^i = \left[\frac{\partial^2 g_i}{\partial \alpha_k\,\partial \alpha_s}\right] \qquad \begin{matrix} i,j = 1,\ldots,n \\ k,s = 1,\ldots,p \end{matrix}$$

With the notations introduced in (2.29), equation (2.28) can be written in terms of inner products.

$$\dot{x}_i = g_i(x(\alpha_o,t),\alpha_o,t) + [F^i, x(\alpha,t) - x(\alpha_o,t)]$$

$$+ [G^i, \alpha - \alpha_o] + \tfrac{1}{2}\,[H^i(x(\alpha,t) - x(\alpha_o,t)),(x(\alpha,t) - x(\alpha_o,t))]$$

$$+ [K^i(x(\alpha,t) - x(\alpha_o,t)),\ (\alpha - \alpha_o)] + \tfrac{1}{2}\,[E^i(\alpha - \alpha_o),(\alpha - \alpha_o)]\ .$$

$$i = 1,\ldots, n \tag{2.30}$$

Writing equation (2.30) in the form of an integral, we obtain

$$\varphi_i(\alpha,t) = x_{io} + \int_{t_o}^{t} g_i(\varphi(\alpha_o,t),\alpha_o,t)dt + \int_{t_o}^{t} [F^i,\ \varphi(\alpha,t) - \varphi(\alpha_o,t)]dt$$

$$+ \int_{t_o}^{t} [G^i, \alpha - \alpha_o]dt + \frac{1}{2}\int_{t_o}^{t} [H^i(\varphi(\alpha,t) - \varphi(\alpha_o,t)),(\varphi(\alpha,t) - \varphi(\alpha_o,t))]dt$$

$$+ \int_{t_o}^{t} [K^i(\varphi(\alpha,t) - \varphi(\alpha_o,t)), \alpha - \alpha_o]dt + \frac{1}{2} \int_{t_o}^{t} [E^i(\alpha - \alpha_o), (\alpha - \alpha_o)]dt$$

$$i = 1, \ldots, n \qquad (2.31)$$

In the following analysis we wish to determine the error, ξ_1, in the evaluation of variation along the optimal trajectory by using the approximate solution $\varphi^{\Delta}(\alpha,t) \triangleq \varphi(\alpha_o,t) + \Delta x$. ξ_1 is defined by the equation

$$\xi_1^2 = \| \varphi(\alpha,t) - \varphi^{\Delta}(\alpha,t) \|^2 \qquad (2.32)$$

$$= \max_{t_o \le t \le t_1} \sum_{i=1}^{n} \left| \varphi_i - \varphi_i \right|^2$$

From equation (2.8), Δx can be obtained and consequently the approximate solution $\varphi^{\Delta}(\alpha,t)$ of equation (2.6) is given by

$$\varphi_i^{\Delta}(\alpha,t) = x_{io} + \int_{t_o}^{t} g_i(\varphi(\alpha_o,t), \alpha_o,t)dt$$

$$+ \int_{t_o}^{t} [F^i, \varphi^{\Delta}(\alpha,t) - \varphi(\alpha_o,t)]dt \qquad (2.33)$$

$$+ \int_{t_o}^{t} [G^i, \alpha - \alpha_o]dt \, .$$

$$i = 1, \ldots, n$$

From equations (2.31) and (2.33) we find

$$\left|\varphi_i - \varphi_i^{\Delta}\right| = \left| \int_{t_o}^{t} [F^i,\ \varphi - \varphi_{\alpha o}]dt + \frac{1}{2} \int_{t_o}^{t} [H^i(\varphi - \varphi_{\alpha o}),\ \varphi - \varphi_{\alpha o}]dt \right.$$

$$+ \int_{t_o}^{t} [K^i(\varphi - \varphi_{\alpha o}),\ \alpha - \alpha_o]dt + \frac{1}{2} \int_{t_o}^{t} [E^i(\alpha - \alpha_o),\ (\alpha - \alpha_o)]dt$$

$$\left. - \int_{t_o}^{t} [F^i,\ \varphi^{\Delta} - \varphi_{\alpha o}]dt \right| \qquad (2.34)$$

If we define the norm of a matrix

$$\|M(t)\| = \max_{t_o \le t \le t_1} [\text{trace } M^* M]^{1/2}, \qquad (2.35)$$

then it can be shown that the following inequality is satisfied (Appendix B)

$$\|y\| \le \|M(t)\| \ \|x\| , \qquad (2.36)$$

where $y(t) = M(t)\, x(t)$.

Using Schwarz inequality $|[x,y]| \le |x||y|$ together with relations (2.21), (2.34) and (2.35), we find

$$\xi_1^2 \le n \max_{t,i} \left|\varphi_i - \varphi_i^{\Delta}\right|^2$$

$$\le n \max_{t} \left\{ \sqrt{n}\, K \int_{t_o}^{t} |\varphi - \varphi_o| dt + \frac{n}{2} K_2 \int_{t_o}^{t} |\varphi - \varphi_o|^2 dt \right.$$

$$+ \sqrt{np}\, K_4 \int_{t_o}^{t} |\varphi - \varphi_{\alpha o}|\,|\alpha - \alpha_o|\, dt + \frac{pK_3}{2} \int_{t_o}^{t} |\alpha - \alpha_o|^2\, dt$$

$$+ \sqrt{n}\, K \int_{t_o}^{t} |\varphi - \varphi_{\alpha_o}|\, dt \Big\}^2$$

Hence

$$\xi_1 \le n^{1/2} \Big\{ 2\sqrt{n}\, K\, (t_1 - t_o)\,|\alpha - \alpha_o| + \frac{n}{2} K_2 c^2 + \sqrt{np}\, K_4 c + \frac{pK_3}{2} (t_1 - t_o)\,|\alpha - \alpha_o| \Big\}^2$$

Relation (2.25) has been used in the last inequality. Let us define

$$M_1 = 2nK(t_1 - t_o) \qquad M_2 = n^{1/2}\Big(\frac{n}{2} K_2 c^2 + (np)^{1/2} K_4 c + \frac{p}{2} K_3 \Big)(t_1 - t_o) \tag{2.37}$$

then

$$\xi_1 \le M_1\,|\alpha - \alpha_o| + M_2\,|\alpha - \alpha_o|^2 \tag{2.38}$$

To find the error introduced in the evaluation of the change in cost by using the approximate solution (2.32) we expand the function $f_o(x,u)$ in (2.3) in a Taylor series about the point $(\varphi(\alpha_o,t),u)$

$$f_o(\varphi(\alpha,t),u) = f_o(\varphi(\alpha_o,t),u) + \sum_{i=1}^{n} \frac{\partial f_o}{\partial x_i} (\varphi_i(,t) - \varphi_i(\alpha_o,t))$$

$$+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f_o}{\partial x_i\, \partial x_j} (\varphi_i(\alpha,t) - \varphi_i(\alpha_o,t))(\varphi_j(\alpha,t) - \varphi_j(\alpha_o,t))$$

$$= f_o(\varphi_{\alpha o}, u) + \sum_{i=1}^{n} \frac{\partial f_o}{\partial x_i} (\varphi_i^{\Delta} - \varphi_{i\alpha o}) \qquad (2.39)$$

$$+ \sum_{i=1}^{n} \frac{\partial f_o}{\partial x_i} (\varphi_i - \varphi_i^{\Delta}) + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f_o}{\partial x_i \partial x_j} (\varphi_i - \varphi_{i\alpha o})(\varphi_j - \varphi_{j\alpha o})$$

The second partial derivatives in (2.39) are evaluated at $(\beta \varphi(\alpha,t) + (1 - \beta)\varphi(\alpha_o,t),u)$ with $0 \leq \beta \leq 1$. Upon comparison between (2.14) and (2.39), the error ξ_2 introduced in the evaluation of the change in cost by the approximate solution $\varphi^{\Delta}(\alpha,t)$ is given by

$$\xi_2 = \left| \int_{t_o}^{t_1} \sum_{i=1}^{n} \frac{\partial f_o}{\partial x_i} (\varphi_i - \varphi_i^{\Delta}) dt + \frac{1}{2} \int_{t_o}^{t_1} \sum_{i,j=1}^{n} \frac{\partial^2 f_o}{\partial x_i \partial x_j} (\varphi_i - \varphi_{i\alpha_o}) \right.$$

$$\left. (\varphi_j - \varphi_{j\alpha_o}) dt \right|$$

$$= \left| \int_{t_o}^{t_1} [F_o, \varphi - \varphi^{\Delta}] dt + \frac{1}{2} \int_{t_o}^{t_1} [G_o(\varphi - \varphi_{\alpha o}), (\varphi - \varphi_{\alpha o})] dt \right| \qquad (2.40)$$

where

$$F_o = \left(\frac{\partial f_o}{\partial x_i}, \ldots, \frac{\partial f_o}{\partial x_n}\right) \quad \text{and} \quad G_o = \left[\frac{\partial^2 f_o}{\partial x_i \partial x_j}\right]$$

Using relations (2.21), (2.25) and (2.38) we obtain

$$\xi_2 \leq \int_{t_o}^{t_1} \|F_o\| \, \|\varphi - \varphi^{\Delta}\| dt + \frac{1}{2} \int_{t_o}^{t_1} \|G_o\| \, \|\varphi - \varphi_o\|^2 dt$$

$$\leq (\sqrt{n}\, K_o'\, M_1|\alpha - \alpha_o| + \sqrt{n}\, K_o'\, M_2|\alpha - \alpha_o|^2 + \frac{nK_o}{2} c^2|\alpha - \alpha_o|^2)(t_1 - t_o)$$

Define

$$M_3 = \sqrt{n}\, K_o'\, M_1(t_1 - t_o)\,, \quad M_4 = (\sqrt{n}\, K_o'\, M_2 + \frac{nK_o}{2} c^2)(t_1 - t_o) \tag{2.41}$$

then

$$\xi_2 \leq M_3 |\alpha - \alpha_o| + M_4 |\alpha - \alpha_o|^2 \tag{2.42}$$

Let us summarize the above results into the following theorem.

Theorem 2.1 Given the open-loop control system of Fig. 2.1 the input-output relationship is given by (2.1). The cost to be minimized is given by (2.3). If the assumptions a,b (section 2.1) and c (section 2.3) are satisfied, the variation along the optimal trajectory and the change in cost because of changes in plant parameters can be evaluated by using the approximate solution (2.33) with the error introduced in the trajectory given by (2.38) and the error introduced in the cost by (2.42) provided that the relation

$$|\alpha - \alpha_o| \leq \rho \leq \min\left[\frac{a}{c}\,,\ b\right] \tag{2.43}$$

is satisfied. The constants a and b are given in (2.26) whereas c and ρ are described in (2.20) and (2.27) respectively.

2.4 Inverse Sensitivity Problem of Determining Actual Variation of Parameters from Known Trajectory

It was shown that for small changes in plant parameters the variation along the optimal trajectory can be accurately determined by using the approximate solution $y'(\alpha,t)$. In performing a specific task the allowable deviation from the optimal trajectory must be very small. Otherwise, the terminal state will not be reached by the plant. Hence, it is of interest to determine how much the plant parameters can vary from their nominal values once the allowable change in trajectory is specified. The physical significance of this lies in the fact that a physical system can never be made identical to the ideal one. Components with high degree of precision are extremely costly while poor quality components will result in unsatisfactory performance. Therefore manufacturing tolerances must be set up on one hand to minimize the cost and on the other hand to meet the requirements for performing a specific task.

For a physical system with one dimensional output Tomovic[30] showed that the changes in parameter values may be determined. However, he failed to specify what conditions the system must satisfy in order that the changes in parameters be determined uniquely.

In the following development we shall proceed to determine the variation of parameter values of the system (2.1) if the actual trajectory can be observed. It is assumed that the following conditions are satisfied:

1) The difference between the actual and optimal trajectories are so small , as judged by the results of section 2.3, that the difference $\Delta x(\alpha_o,t)$ can be considered as the solution of the equation of first variation

$$\frac{d\Delta x_i}{dt} = \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j} \Delta x_j + \sum_{k=1}^{p} \frac{\partial g_i}{\partial \alpha_k} \Delta \alpha_k \qquad (2.8)$$

$$\Delta x_i(t_o) = 0 \qquad i = 1,\ldots,n$$

2) The vectors $\frac{\partial g}{\partial \alpha_i} = \left(\frac{\partial g_1}{\partial \alpha_i}, \ldots, \frac{\partial g_n}{\partial \alpha_i} \right)$ $\quad i = 1,\ldots,p.$

formed from the columns of the matrix G in (2.10) are linearly independent.

Then from equation (2.13) the $i^{\underline{th}}$ component of $\Delta x(\alpha_o,t)$ can be written as

$$\Delta x_i(\alpha_o,t) = \int_{t_o}^{t} \sum_{k=1}^{p} \left(\sum_{j=1}^{n} \phi_{ij}(t,\tau) g_{jk}(\tau) \right) \Delta \alpha_k \, d\tau \,, \qquad (2.44)$$

where

$\phi_{ij}(t,\tau) \triangleq$ the element located at the $i^{\underline{th}}$ row and the $j^{\underline{th}}$ column of the transition matrix $\underline{\Phi}(t,\tau)$, and

$g_{jk}(\tau) \triangleq \frac{\partial g_j}{\partial \alpha_k}$ = the element located at the $j^{\underline{th}}$ row and the $k^{\underline{th}}$ column of $G(\tau)$.

Let us define the function $\lambda_{ik}(t)$ as

$$\lambda_{ik}(t) = \int_{t_o}^{t} \sum_{j=1}^{n} \phi_{ij}(t,\tau)\, g_{jk}(\tau)\, d\tau \,. \qquad (2.45)$$

$$i = 1,\ldots,n \;\; ; \; k = 1,\ldots,p$$

Then equation (2.44) reduces to

$$\Delta x_i(\alpha_o,t) = \Delta\alpha_1 \lambda_{i1}(t) + \cdots + \Delta\alpha_p \lambda_{ip}(t) \qquad (2.46)$$

$$i = 1,\ldots,n$$

Writing (2.46) in vector form we find

$$\Delta x(\alpha_o,t) = \Delta\alpha_1 \lambda_1(t) + \cdots + \Delta\alpha_p \lambda_p(t) \,, \qquad (2.47)$$

where

$$\lambda_k(t) \triangleq (\lambda_{1k}(t), \cdots, \lambda_{nk}(t)).$$

$$k = 1,\ldots,p.$$

Every vector $\Delta x(\alpha_o,t)$ has a unique representation of the form

$$\Delta x(\alpha_o,t) = \sum_{k=1}^{p} \Delta\alpha_k \lambda_k(t) \,.$$

To show that this is true let us suppose on the contrary that there exist two different sets of $\{\Delta\alpha_k\}$ that yield the same $\Delta x^*(\alpha_o,t)$. We write

$$\Delta x^*(\alpha_o,t) = \sum_{k=1}^{p} \Delta\alpha_k^a \lambda_k(t) \,, \qquad (2.48)$$

and

$$\Delta x^*(\alpha_o,t) = \sum_{k=1}^{p} \Delta\alpha_k^b \lambda_k(t) \,. \qquad (2.49)$$

Then upon subtraction of (2.49) from (2.48) we find

$$0 = \sum_{k=1}^{p} \Delta\alpha_k^c \ \lambda_k(t) \ , \tag{2.50}$$

where

$$\Delta\alpha_k^c \triangleq \Delta\alpha_k^a - \Delta\alpha_k^b \ .$$

Equation (2.50) means that the set $\left\{\Delta\alpha_k^c\right\}$ will yield a zero solution for all time in the interval $[t_o,t_1]$. From equation (2.13) it is possible only when

$$G\ \Delta\alpha^c = 0 \ . \qquad \text{for } t_o \le t \le t_1 \tag{2.51}$$

Let us expand (2.51) into the following form

$$\frac{\partial g}{\partial\alpha_1}\Delta\alpha_1^c + \frac{\partial g}{\partial\alpha_2}\Delta\alpha_2^c + \cdots + \frac{\partial g}{\partial\alpha_p}\Delta\alpha_p^c = 0. \tag{2.52}$$

This contradicts the assumption that the set of vectors $\frac{\partial g}{\partial\alpha_i}$ are linearly independent. Therefore the set $\left\{\lambda_k(t)\right\}$ form a basis for $\Delta x(\alpha_o,t)$. In terms of reciprocal basis[31] equation (2.47) can be written

$$\Delta x(\alpha_o,t) = \sum_{k=1}^{p} \Delta\alpha_k \ \lambda_k(t) = \sum_{k=1}^{p} \left(\Delta x(\alpha_o,t),\beta_k(t)\right) \lambda_k(t), \tag{2.53}$$

where

$\left\{\beta_k(t)\right\}$ is the reciprocal basis of $\left\{\lambda_k(t)\right\}$,

and

$$\left(\Delta x(\alpha_o,t),\beta_k(t)\right) \triangleq \int_{t_o}^{t_1} \sum_{i=1}^{n} \Delta x_i(\alpha_o,t)\ \beta_{ik}(t)\ dt \ . \tag{2.54}$$

From equation (2.53) it is obvious that the determination of $\{\Delta\alpha_k\}$ amounts to finding the reciprocal basis of $\{\lambda_k(t)\}$. It is well known[31] that an arbitrary basis can always be transformed into an orthonormal basis and that the reciprocal basis of an othonormal set is the orthonormal set itself. Therefore we shall proceed to construct an orthogonal set from the given set $\{\lambda_k(t)\}$. By Schmidt orthogonalization process[31] let

$$y_1(t) = \lambda_1(t)$$

$$y_2(t) = \lambda_2(t) - \frac{(y_1(t), \lambda_2(t))}{(y_1(t), y_1(t))}\, y_1(t) \tag{2.55}$$

. . .

$$y_i(t) = \lambda_i(t) - \sum_{j=1}^{i-1} \frac{(y_j(t), \lambda_i(t))}{(y_j(t), y_j(t))}\, y_j(t)$$

$$i = 2,\ldots p$$

or solving for $\lambda_i(t)$ in terms of the orthogonal set $\{y_i(t)\}$ we obtain

$$\lambda_1(t) = y_1(t)$$

$$\lambda_2(t) = y_2(t) + \frac{(y_1(t), \lambda_2(t))}{(y_1(t), y_1(t))}\, y_1(t) \tag{2.56}$$

. . .

$$\lambda_i(t) = y_i(t) + \sum_{j=1}^{i-1} \frac{(y_j(t), \lambda_i(t))}{(y_j(t), y_j(t))}\, y_j(t)$$

$$i = 2,\ldots p \quad .$$

By inspection of (2.56) if we define

$$a_{ji} = \begin{cases} -\dfrac{(y_j(t),\ \lambda_i(t))}{(y_j(t),\ y_j(t))} & j < i \\ 1 & j = i \\ 0 & j > i \end{cases} \tag{2.57}$$

then equation (2.56) reduces to the more convenient form,

$$\lambda_i(t) = \sum_{j=1}^{i} a_{ji}\, y_j(t) \quad . \qquad i = 1,\ldots,p \tag{2.58}$$

Upon substitution of (2.58) into (2.53) we obtain

$$\Delta x(\alpha_o,t) = \sum_{i=1}^{p} \Delta\alpha_i \sum_{j=1}^{i} a_{ji}\, y_j(t)$$

$$= \sum_{i=1}^{p} y_i(t) \left(\sum_{j=1}^{p} a_{ji}\, \Delta\alpha_j \right) \tag{2.59}$$

and in terms of the orthogonal basis $\left\{y_i(t)\right\}$, $\Delta x(\alpha_o,t)$ can be expressed as

$$\Delta x(\alpha_o,t) = \sum_{i=1}^{p} y_i(t)\, \frac{(\Delta x(\alpha_o,t),\ y_i(t))}{(y_i(t)\ ,\ y_i(t))} \quad . \tag{2.60}$$

Since the set $\{y_i(t)\}$ is linearly independent, we obtain by comparing the coefficients of equations (2.59) and (2.60)

$$\sum_{j=1}^{p} a_{ji} \Delta\alpha_j = \frac{(\Delta x(\alpha_o,t),\ y_i(t))}{(y_i(t),\ y_i(t))} \qquad i = 1,\ldots,p \tag{2.61}$$

If we define

$$\xi_i = \frac{(\Delta x(\alpha_o,t),\ y_i(t))}{(y_i(t),\ y_i(t))} \ , \qquad i = 1,\ldots,p \tag{2.62}$$

equation (2.61) can be expressed in the vector form as

$$A\ \Delta\alpha = \xi \ , \tag{2.63}$$

where

$$A = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1p} \\ & 1 & a_{22} & \cdots & a_{2p} \\ & & \ddots & & \\ & & & \ddots & \\ 0 & & & & 1 \end{bmatrix} \qquad \xi = (\xi_1, \ldots, \xi_p) \tag{2.64}$$

Since the matrix A is nonsingular and triangular, it can be easily inverted. Hence

$$\Delta\alpha = A^{-1}\xi \tag{2.65}$$

or

$$\alpha = \alpha_o + A^{-1}\xi$$

Let us summarize the above results into the following theorem.

Theorem 2.2 If the change along the optimal trajectory of the system (2.1) due to variation of the plant parameters can be observed, and satisfies equation (2.8), and if the vectors formed from the columns of the matrix G in (2.10) are linearly independent, then the change in plant parameters can be uniquely determined.

2.5 Estimation of Variations in Time-Dependent Parameters I

In section 2.4 the determination of changes in parametric values was based on the assumption that the actual plant parameters of the system were constants, and in this case we were able to show that there is only one set of parameters corresponding to a given trajectory. However, in the case of time-varying parameters there may be more than one set of parameters which give the same trajectory. The question arises as to which set of parameters one should determine if there is more than one set of parameters corresponding to the same trajectory. In this section only small variations along the optimal trajectory are considered. Due to the fact that the trajectory of the system under consideration is assumed to be continuous function of the plant parameters, if the variation along the optimal trajectory is small, then it is reasonable to state that the variation in the plant parameters of the system is also small. Hence, the relation between the variation along the optimal trajectory and the variation in plant parameters is governed approximately by a set of linear differential equations, i.e., the equations of first variation of the given system. In inverse sensitivity analysis we are interested in determining the parameter variations about their nominal values corresponding to a given variation along the optimal trajectory. Using the method described below one is able to estimate the particular set $\alpha(t)$ corresponding to the actual trajectory such that the norm of the variation of parameters about their nominal values is minimal. We shall use the method of steepest descent in function space developed by Kantarovich[32] and later extended by Balakrishnan[33].

The advantage of using the last mentioned method is as follows: As stated above, there is, in general, more than one set of parametric variations corresponding to the same variation in trajectory. This method enables one to determine the unique element of minimal norm in a closed convex set C. This is essentially a generalized inverse problem. By making the radius, δ , of the convex set sufficiently small, the solution of the linearized system will be an accurate estimate of the actual solution.

In the following discussion we shall follow closely the work of Balakrishnan. Let H^n be the real Hilbert space formed of vector functions $x(t) = (x_1(t),\ldots,x_n(t))$ on $[t_o,t_1]$ for which each component of $x(t)$ is in $L_2[t_o,t_1]$. Let us define $x'(t)$ as the transpose of the vector function $x(t)$, then the inner product is defined by .

$$[y,x] = \int_{t_o}^{t_1} x'(t)\, y(t)dt = \int_{t_o}^{t_1} \sum_{i=1}^{n} x_i(t)\, y_i(t)dt \qquad (2.66)$$

for $y,x \in H^n$.

The norm of this space will be

$$\|x\| = \left\{ \int_{t_o}^{t_1} |x(t)|^2 dt \right\}^{1/2} = \left\{ \int_{t_o}^{t_1} x'(t)\, x(t)\, dt \right\}^{1/2} \qquad (2.67)$$

For a given set of plant parameters $\alpha(t)$, the corresponding approximate solution $x^{\Delta}(\alpha,t)$ of the system (2.1) is defined as the sum of nominal solution $x(\alpha_o,t)$ and the variational term along the nominal solution. Mathematically, $x^{\Delta}(\alpha,t)$ is given by

$$x^{\Delta}(\alpha,t) = x(\alpha_o,t) + \int_{t_o}^{t} \underline{\Phi}(t,\tau)\, G(\tau)\gamma(\tau)d\tau \tag{2.68}$$

$$= x(\alpha_o,t) + \int_{t_o}^{t_1} \underline{\Phi}(t,\tau)\, G(\tau)\gamma(\tau)d\tau$$

where $\gamma(t) \triangleq \alpha(t) - \alpha_o(t)$, (2.69)

the variations in plant parameters, and

$$\underline{\Phi}(t,\tau) \triangleq 0 \qquad \text{for } \tau > t \tag{2.70}$$

The condition (2.70) is always met in a physical system since no response exists for a system before it is excited.

Let us define a linear operator L on H^p by

$$L\gamma(t) = \int_{t_o}^{t_1} \underline{\Phi}(t,\tau)\, G(\tau)\gamma(\tau)d\tau \tag{2.71}$$

with $\gamma(t) \in H^p$, a p-dimensional Hilbert vector function space over $[t_o,t_1]$. L is a bounded linear transformation of H^p into H^n.

Let us assume that the actual trajectory can be observed and is given by

$$x_a(t) = (x_{a_1}(t),\ldots,x_{a_n}(t)) \tag{2.72}$$

To estimate the set of parameters with minimal norm about the nominal plant parameter vector we consider the problem of minimizing

$$\| x_a(t) - x^{\Delta}(\alpha,t)\| = \| x_a(t) - x(\alpha_o,t) - \int_{t_o}^{t_1} \underline{\Phi}(t,\tau)\, G(\tau)\gamma(\tau)\, d\tau \|$$

$$= \| L\gamma - \Delta x_{ao}(t) \| \tag{2.73}$$

where

$$\Delta x_{ao}(t) = x_a(t) - x(\alpha_o, t) \tag{2.74}$$

In order that the approximate solution $x^A(\alpha,t)$ be an accurate estimate of the actual solution, the norm of $\gamma(t)$ must be small so that $\gamma(t)$ is assumed to lie in a closed convex set C, for example a sphere of radius, 8 . Since the kernel $k(t,\tau) \triangleq \Phi(t,\tau)G(\tau)$ satisfies the condition

$$\|k(t,\tau)\| = \left[\int_{t_o}^{t_1}\int_{t_o}^{t_1} |k(t,\tau)|^2 \, dt d\tau\right]^{1/2} = \left[\int_{t_o}^{t_1}\int_{t_o}^{t_1}\sum_{j=1}^{n}\sum_{i=1}^{p} k_{ji}^2(t,\tau) dt d\tau\right]^{1/2}$$

$$< \infty$$

it is called a Hilbert-Schmidt kernal. It is known in functional anaysis[34] that every integral operator with a Hilbert-Schmidt kernel is compact or completely continuous in Hilbert space. Therefore, L maps every bounded set in H^p into a compact set in H^n .

It has been shown by Balakrishnan[33] that there is a unique element γ_o of minimal norm in convex set C such that

$$\inf_{\gamma} \|L\gamma - \Delta x_{ao}(t)\| = \|L\gamma_o - \Delta x_{ao}(t)\| \tag{2.75}$$

Our problem is to find this γ_o . Let us first consider the case when no restriction is placed on the norm of γ_o . If we denote the adjoint of the operator L by L^* , then it can be shown that L^* is given by

$$L^*y = \int_{t_o}^{t_1} G'(t)\,\Phi'(\tau,t)\,y(\tau)\,d\tau, \quad y(\tau) \in H^n \tag{2.76}$$

Since

$$[L\gamma, y] = \int_{t_o}^{t_1} y'(t) \int_{t_o}^{t_1} \Phi(t,\tau) G(\tau)\gamma(\tau)\, d\tau\, dt$$

$$= \int_{t_o}^{t_1} \int_{t_o}^{t_1} y'(\tau) \Phi(\tau,t) G(t)\gamma(t)\, dt\, d\tau \qquad (2.77)$$

and

$$[L\gamma, y] = [\gamma, L^* y] = \int_{t_o}^{t_1} (L^* y)' \gamma(t)\, dt \qquad (2.78)$$

Upon comparison of (2.77) and (2.78) we find

$$(L^* y)' = \int_{t_o}^{t_1} y'(\tau) \Phi(\tau,t) G(t)\, d\tau \qquad (2.79)$$

Equation (2.79) reduces to (2.76) by a transpose operation. L^* is a bounded linear operator. It is known in functional analysis[35] that the product of a bounded operator and a completely continuous operator is completely continuous. Thus, both LL^* and L^*L are completely continuous. Furthermore, the operators LL^* and L^*L are self-adjoint and nonnegative definite, since

$$(LL^*)^* = LL^*$$

and

$$[LL^* x, x] = [L^* x, L^* x] \geq 0 \qquad (2.80)$$

Similarly for L^*L. Since a self-adjoint completely continuous operator cannot have a continuous spectrum[35], LL^* and L^*L can have at most a countable number of positive eigenvalues. Let $\{\lambda_i\}$, $(i=1,2,\ldots)$ be the nonzero eigenvalues of LL^* with the magnitudes of λ_i arranged in

non-increasing order and the corresponding orthonormal eigenvectors be designated by $\{e_i\}$. It can be shown that

$$\phi_i = \frac{L^* e_i}{\lambda_i^{1/2}} \tag{2.81}$$

are orthonormal eigenvectors of L^*L, for

$$L^*L\,\phi_i = L^*\,\frac{LL^* e_i}{\lambda_i^{1/2}} = \lambda_i^{1/2}\,L^* e_i = \lambda_i\,\phi_i \tag{2.82}$$

From equation (2.82) it is obvious that LL^* and L^*L have the same nonzero eigenvalues. Let $R(L^*)$ be the subspace of H^p spanned by $\{\phi_i\}$, then H^p can be decomposed into the direct sum of $R(L^*)$ and its orthogonal compliment $R(L^*)^\perp$ since H^p is a complete linear inner product space[36]

$$\begin{aligned} H^p &= R(L^*) \oplus R(L^*)^\perp \\ &= R(L^*) \oplus N(L) \end{aligned} \tag{2.83}$$

where $N(L)$ is the null space of L. Any element γ in H^p can be written uniquely as

$$\gamma = \sum_{i=1}^{\infty} [\gamma, \phi_i]\,\phi_i + \gamma_N \quad , \qquad \gamma_N \in N(L) \tag{2.84}$$

and

$$L\gamma = \sum_{i=1}^{\infty} [\gamma, \phi_i]\,L\,\phi_i \tag{2.85}$$

Similarly, for any element $x \in H^n$ we have

$$L^* x = \sum_{i=1}^{\infty} [x, e_i] L^* e_i \tag{2.86}$$

Let us expand

$$\| L\gamma - \Delta x_{ao} \|^2 = [L\gamma - \Delta x_{ao}, L\gamma - \Delta x_{ao}]$$

$$= [L^* L\gamma, \gamma] - 2[\gamma, L^* \Delta x_{ao}] + [\Delta x_{ao}, \Delta x_{ao}] \tag{2.87}$$

Since $\Delta x_{ao}(t)$ is known and is in H^n, it can be decomposed into

$$\Delta x_{ao} = Lh + \Delta x_1 \tag{2.88}$$

where $h \in H^p$ and $\Delta x_1 \in R(L)^{\perp} = N(L^*)$. Thus

$$L^* \Delta x_{ao} = L^* Lh \tag{2.89}$$

Substituting (2.89) into (2.87) gives

$$\| L\gamma - \Delta x_{ao} \|^2 = [L^* L(\gamma - h), (\gamma - h)] + [\Delta x_{ao}, \Delta x_{ao}] - [Lh, Lh], \tag{2.90}$$

and hence

$$\| L\gamma - \Delta x_{ao} \|^2 \geq [\Delta x_{ao}, \Delta x_{ao}] - [Lh, Lh], \tag{2.91}$$

since the first term on the right-hand side of (2.90) is non negative. The equality sign in (2.91) holds if and only if

$$\gamma = h$$

Using (2.82), (2.85) and (2.89) we find

$$
\begin{aligned}
L^* I h &= \sum_{i=1}^{\infty} [h, \phi_i]\, L^* L\, \phi_i \\
&= \sum_{i=1}^{\infty} \lambda_i [h, \phi_i]\, \phi_i \\
&= \sum_{i=1}^{\infty} [\Delta x_{ao}, e_i]\, L^* e_i
\end{aligned}
\tag{2.92}
$$

and hence the element h is given by

$$
h = \sum_{i=1}^{\infty} [h, \phi_i]\, \phi_i + h_N \qquad h_N \in N(L) \tag{2.93}
$$

Due to the fact that h_N is in the null space of L it only increases the norm of h without having any effect on the minimization of (2.91). Therefore the element h must be of the form

$$
\begin{aligned}
h &= \sum_{i=1}^{\infty} [h, \phi_i]\, \phi_i \\
&= \sum_{i=1}^{\infty} \frac{[L^* \Delta x_{ao}\, ,\, \phi_i]}{\lambda_i}\, \phi_i
\end{aligned}
\tag{2.94}
$$

It should be noted that h is in the range of L^* and that it would have been the minimizing element sought if there were no restriction on its norm. Since γ is confined to lie in C, its norm must satisfy

the condition

$$\|\gamma\|^2 = \sum_{i=1}^{\infty} \alpha_i^2 \leq \delta^2 \tag{2.95}$$

where

$$\gamma = \sum_{i=1}^{\infty} \alpha_i \phi_i \tag{2.96}$$

Our immediate problem is to find the element γ_o which will minimize $\|L\gamma - \Delta x_{a\theta}\|$ and at the same time will satisfy condition (2.95). From equation (2.87) to minimize $\|L\gamma - \Delta x_{a\theta}\|^2$ is equivalent to minimizing

$$Q(\gamma) \triangleq [L^*L\gamma, \gamma] - 2[\gamma, L^*\Delta x_{a\theta}] \tag{2.97}$$

since $\Delta x_{a\theta}$ is known.

Since L^*L is non-negative, then for any positive number K, the operator $(L^*L + KI)$ satisfies the relation

$$K\|\gamma\| \leq \|(L^*L + KI)\gamma\| \quad , \tag{2.98}$$

and the operator $(L^*L + KI)$ has a bounded inverse.[36]

Let

$$\gamma_K \triangleq [L^*L + KI]^{-1} L^*\Delta x_{a\theta} \tag{2.99}$$

then it can be shown that $\|\gamma_K\|^2$ is a monotonic decreasing function of K, since

$$\| \gamma_{K_2} \|^2 - \| \gamma_{K_1} \|^2$$

$$= [(L^*L + K_2 I)^{-1} L^* \Delta x_{ao} \, , \, (L^*L + K_2 I)^{-1} L^* \Delta x_{ao}] -$$

$$[(L^*L + K_1 I)^{-1} L^* \Delta x_{ao}, \, (L^*L + K_1 I)^{-1} L^* \Delta x_{ao}]$$

$$= [(L^*L + K_2 I)(L^*L + K_2 I)^{-2} L^* \Delta x_{ao} \, , \, (L^*L + K_2 I)^{-1} L^* \Delta x_{ao}] -$$

$$[(L^*L + K_1 I)(L^*L + K_1 I)^{-2} L^* \Delta x_{ao} \, , \, (L^*L + K_1 I)^{-1} L^* \Delta x_{ao}]$$

$$= [(L^*L + K_2 I)^{-2} L^* \Delta x_{ao}, \, (L^*L + K_2 I)^* (L^*L + K_2 I)^{-1} L^* \Delta x_{ao}] -$$

$$[(L^*L + K_1 I)^{-2} L^* \Delta x_{ao}, \, (L^*L + K_1 I)^* (L^*L + K_1 I)^{-1} L^* \Delta x_{ao}]$$

$$= [(L^*L + K_2 I)^{-2} L^* \Delta x_{ao} - (L^*L + K_1 I)^{-2} L^* \, x_{ao}, \, L^* \Delta x_{ao}]$$

$$= [(L^*L + K_1 I)^{-2} \, (L^*L + K_1 I)^2 \, (L^*L + K_2 I)^{-2} \, L^* \Delta x_{ao} -$$

$$(L^*L + K_1 I)^{-2} \, (L^*L + K_2 I)^2 (L^*L + K_2 I)^{-2} \, L^* \Delta x_{ao}, \, L^* \Delta x_{ao}]$$

$$= [(L^*L + K_1 I)^{-2} \left\{ (L^*L + K_1 I)^2 - (L^*L + K_2 I)^2 \right\} (L^*L + K_2 I)^{-2} L^* \Delta x_{ao}, L^* \Delta x_{ao}]$$

$$= [(L^*L + K_1 I)^{-2} \left\{ 2L^*L(K_1 - K_2) + (K_1 - K_2)(K_1 + K_2) \right\} (L^*L + K_2 I)^{-2} L^* \Delta x_{ao}, L^* \Delta x_{ao}]$$

$$= (K_1 - K_2)[(L^*L + K_1 I)^{-2} \left\{ 2L^*L + (K_1 + K_2) I \right\} (L^*L + K_2 I)^{-2} \, L^* \Delta x_{ao}, \, L^* \Delta x_{ao}] \tag{2.100}$$

The second factor is positive. Therefore

$$\| \gamma_{K_2} \|^2 - \| \gamma_{K_1} \|^2 < 0 \qquad \text{for } K_1 < K_2 \tag{2.101}$$

Furthermore, as $K \longrightarrow \infty$, $\| \gamma_K \| \longrightarrow 0$.

Let us consider the problem of minimizing, for fixed K,

$$Q_K(\gamma) \triangleq [(L^*L + KI)\gamma, \gamma] - 2[\gamma, L^*\Delta x_{ao}] \tag{2.102}$$

$$= [(L^*L + KI)\gamma, \gamma] - 2[\gamma, (L^*L + KI)\gamma_K]$$

$$+ [(L^*L + KI)\gamma_K, \gamma_K)] - [(L^*L + KI)\gamma_K, \gamma_K]$$

$$= [(L^*L + KI)(\gamma - \gamma_K), (\gamma - \gamma_K)] - [(L^*L + KI)\gamma_K, \gamma_K] \tag{2.103}$$

It is obvious that minimum occurs when $\gamma = \gamma_K$:

$$\inf_{\gamma \in H^p} Q_K(\gamma) = -[(L^*L + KI)\gamma_K, \gamma_K] \tag{2.104}$$

We need to consider two cases.

Case 1. If

$$\sup_{K>0} \left\| \gamma_K \right\|^2 \leq \delta^2 \tag{2.105}$$

Since $L^*\Delta x_{ao} \in R(L^*)$, it can be represented as

$$L^*\Delta x_{ao} = \sum_{i=1}^{\infty} [L^*\Delta x_{ao}, \phi_i]\, \phi_i \tag{2.106}$$

From (2.96), (2.99) and (2.106) we find

$$\gamma_K = \sum_{i=1}^{\infty} \frac{[L^*\Delta x_{ao}, \phi_i]}{\lambda_i + K} \phi_i \tag{2.107}$$

so that

$$\| \gamma_K \|^2 = \sum_{i=1}^{\infty} \frac{[L^* \Delta x_{ao}, \phi_i]^2}{(\lambda_i + K)^2} \leq \delta^2 \tag{2.108}$$

Hence

$$\sum_{i=1}^{\infty} \frac{[L^* \Delta x_{ao}, \phi_i]^2}{\lambda_i^2} \leq \delta^2 \tag{2.109}$$

and we can define

$$\gamma_o = \sum_{i=1}^{\infty} \frac{[L^* \Delta x_{ao}, \phi_i]}{\lambda_i} \phi_i \tag{2.110}$$

Then the sequence $\{\gamma_K\}$ converges to γ_o as $K \longrightarrow 0$. This element γ_o is the element of minimal norm which minimizes $Q(\gamma)$ in (2.97). It can also be seen that this element γ_o is the same as the one expressed in (2.94).

Case 2. If

$$\sup_{K>0} \| \gamma_K \|^2 > \delta^2 , \tag{2.111}$$

Then there is a positive number K_o such that

$$\| \gamma_{K_o} \|^2 = \delta^2 , \tag{2.112}$$

since $\| \gamma_K \|^2$ is a monotone decreasing function of K.

Let

$$Q_o = \inf_{\gamma \in C} Q(\gamma) \tag{2.113}$$

Also let us designate the element γ in C corresponding to Q_o as γ_o so that

$$Q_o = Q(\gamma_o) \quad . \tag{2.114}$$

Then

$$\|\gamma_o\|^2 = m^2 \leq \delta^2 \quad . \tag{2.115}$$

We wish to show that $m^2 = \delta^2$. Since, for any $K > 0$

$$Q(\gamma_o) = \inf_{\|\gamma\|^2 = m^2} Q(\gamma)$$

$$= \inf_{\|\gamma\|^2 = m^2} \left\{ [(L^*L + KI)\gamma, \gamma] - 2[\gamma, L^*\Delta x_{ao}] - Km^2 \right\}$$

$$\geq \inf_{\gamma \in H^p} \left\{ [(L^*L + KI)\gamma, \gamma] - 2[\gamma, L^*\Delta x_{ao}] - Km^2 \right\}$$

$$= \inf_{\gamma \in H^p} Q_K(\gamma) - Km^2$$

$$= Q_K(\gamma_K) - Km^2$$

Since $Q_K(\gamma_K) = Q(\gamma_K) + K\|\gamma_K\|^2$, then

$$Q(\gamma_o) \geq Q(\gamma_K) + K(\|\gamma_K\|^2 - m^2) \tag{2.116}$$

Let $K = K_o$, then

$$Q(\gamma_o) \geq Q(\gamma_{K_o}) + K_o(\|\gamma_{K_o}\|^2 - m^2) \tag{2.117}$$

Since γ_o is by definition a minimizing element in C and $\gamma_{K_o} \in C$ then (2.117) implies that

$$\left\| \gamma_{K_o} \right\|^2 - m^2 = 0 \tag{2.118}$$

or

$$m^2 = \delta^2 \text{ and } \left\| \gamma_{K_o} \right\| = \left\| \gamma_o \right\| \tag{2.119}$$

otherwise $Q(\gamma_{K_o}) < Q(\gamma_o)$, contradicting the assumption that $Q(\gamma_o) = \inf_{\gamma \in C} Q(\gamma)$.

Knowing the fact that every closed convex set in a Hilbert space has a unique element of minimal norm, we write

$$L\gamma_o = L\gamma_{K_o} \tag{2.120}$$

or

$$L(\gamma_o - \gamma_{K_o}) = 0$$

$$\therefore \quad \gamma_o = \gamma_{K_o} + \gamma_N \qquad \text{where } \gamma_N \in N(L) \tag{2.121}$$

since

$$\begin{aligned} [\gamma_N, \gamma_{K_o}] &= [\gamma_N, (L^*L + K_oI)^{-1} L^* \Delta x_{ao}] \\ &= [(L^*L + K_oI)^{-1} \gamma_N, L^* \Delta x_{ao}] \\ &= [\frac{\gamma_N}{K_o}, L^* \Delta x_{ao}] \\ &= [\frac{L\gamma_N}{K_o}, \Delta x_{ao}] = 0 \end{aligned} \tag{2.122}$$

and hence

$$\left\|\gamma_o\right\|^2 = \left\|\gamma_{K_o}\right\|^2 + \left\|\gamma_N\right\|^2 \quad , \tag{2.123}$$

From (2.119), γ_N must be zero and the minimizing element is

$$\gamma_o = \gamma_{K_o} = (L^*L + K_oI)^{-1}L^*\Delta x_{ao} \tag{2.124}$$

with K_o adjusted so that $\left\|\gamma_o\right\| = \delta$.

The above development can be summarized as follows:
In order to find the element γ_o of minimal norm which lies in C and minimizes equation (2.72) we construct a sequence $\left\{\gamma_K\right\}$ of the form

$$\gamma_K = [L^*L + KI]^{-1} L^*\Delta x_{ao} \tag{2.125}$$

which will converge to the minimizing element γ_o .
If

$$\sup_{K>0} \left\|(L^*L + KI)^{-1} L^*\Delta x_{ao}\right\| \leq \delta \quad , \tag{2.126}$$

then

$$\gamma_o = (L^*L)^{-1} L^*\Delta x_{ao} \quad , \tag{2.127}$$

and

$$\lim_{K\to 0} \left\|L\gamma_K - \Delta x_{ao}\right\|^2 = \inf_{\gamma\in C} \left\|L\gamma - \Delta x_{ao}\right\|$$
$$= \left\|L\gamma_o - \Delta x_{ao}\right\|^2 \tag{2.128}$$

However, if

$$\sup_{K \geq 0} \left\| (L^*L + KI)^{-1} L^* \Delta x_{ao} \right\| > \delta \,, \tag{2.129}$$

then

$$\gamma_o = (L^*L + K_o I)^{-1} L^* \Delta x_{ao} \,, \tag{2.130}$$

with K_o adjusted so that $\left\| \gamma_o \right\| = \delta$.

In order to construct the sequence $\left\{ \gamma_K \right\}$ the method of steepest descent is used. As indicated before the minimization problem is equivalent to minimizing

$$Q(\gamma) = [L^*L\,\gamma, \gamma] - 2[\gamma, L^* \Delta x_{ao}] \tag{2.131}$$

with γ confined to lie in C.

Let us denote the nth iteration by γ_n and consider the minimization of

$$Q_K(\gamma_{n+1}) \triangleq [(L^*L + KI)\,\gamma_{n+1}, \gamma_{n+1}] - 2[\gamma_{n+1}, L^* \Delta x_{ao}] \tag{2.132}$$

Let $\gamma_{n+1} = \gamma_n - \epsilon_n z_n$ with $z_n \epsilon H^p$ and ϵ_n , a scalar then

$$Q_K(\gamma_{n+1}) = [(L^*L + KI)\,\gamma_n, \gamma_n] - 2\epsilon_n [(L^*L + KI) z_n, \gamma_n] + \epsilon_n^2 [(L^*L + KI) z_n, z_n]$$

and the directional derivative at γ_n is given by

$$\frac{d}{d\epsilon_n} Q_K(\gamma_n - \epsilon_n z_n) \Big|_{\epsilon_n = 0} = 2[z_n, L^* \Delta x_{ao} - (L^*L + KI)\,\gamma_n] \tag{2.133}$$

By Schwarz inequality, (2.133) will be minimum if z_n is chosen to be

$$z_n = (L^*L + KI)\gamma_n - L^*\Delta x_{ao} \tag{2.134}$$

This z_n determines the direction of steepest descent. The step size is obtained by setting $\frac{\partial Q_K}{\partial \epsilon_n}(\gamma_{n+1}) = -2[(L^*L+KI)z_n, \gamma_n] + 2\epsilon_n[(L^*L+KI)z_n, z_n]$ to zero, and ϵ_n is given by

$$\epsilon_n = \frac{[z_n, z_n]}{[(L^*L+KI)z_n, z_n]} \tag{2.135}$$

in which (2.134) has been used.

We define as a continuous function of K,

$$\gamma_{n+1}(K) = \gamma_n - \epsilon_n z_n \tag{2.136}$$

If $K = 0$ and $\| \gamma_{n+1}(0) \| \leq \delta$, then we define the nth iteration as

$$\gamma_{n+1} = \gamma_{n+1}(0) \tag{2.137}$$

otherwise, $K = K_n$ is chosen such that $\| \gamma_{n+1}(K_n) \| = \delta$ and define

$$\gamma_{n+1} = \gamma_{n+1}(K_n) \tag{2.138}$$

Remark: The minimizing element γ_o found by this method is in L_2 space. There is no guarantee that γ_o is a continuous function. However, it can be approximated as closely as desired by a continuous function due to the fact that the set of all continuous functions is dense in L_2 space.[37] The approximating sequence will always provide such continuous functions.

2.6 Estimation of Variations in Time-Dependent Parameters II

In this section we assume that the actual trajectory differs considerably from that of the optimal one and we are intersted in how the plant parameters vary about their nominal values. In other words we wish to obtain some information about the structure of the actual physical system. Again the method of steepest descent can be used to estimate a set of plant parameters corresponding to the actual trajectory. We assume as before that the actual trajectory can be observed and is designated by $x_a(t)$. Then from equation (2.6) we write

$$x_{a_i}(t) = F_i(\alpha,t) = \int_{t_o}^{t} g_i[x_a(t), \alpha(t), t]\, dt \qquad (2.139)$$

The initial value x_{io} is taken to be zero for convenience without loss of generality. Our problem is to find a set of plant parameters satisfying (2.139). In general equation (2.139) does not guarantee a one to one correspondence between $x_a(t)$ and $\alpha(t)$. We must define our solution of (2.139) in the following manner. Let us consider the problem of minimizing the functional defined by

$$h(\alpha) = \int_{t_o}^{t_1} k[\alpha(t), t]dt = \int_{t_o}^{t_1} \sum_{i=1}^{n} [x_{a_i}(t) - F_i(\alpha,t)]^2 dt \qquad (2.140)$$

If we can find a $\alpha^*(t)$ such that $h(\alpha^*) = 0$ then our problem is solved. We shall consider the following theorem.

Theorem. Let $H(\alpha) = \int_{t_o}^{t_1} k[\alpha(t), t]dt$ be a non-negative continuous functional defined on the Hilbert space H^p of square integrable functions

of dimension p, $\alpha(t) = (\alpha_1(t), \ldots, \alpha_p(t))$.

Let $\alpha^{(0)}(t)$ be some point in H^p such that the set R defined by

$$R = \left\{ \alpha(t) \;\middle|\; h(\alpha) \leq h(\alpha^{(0)}) \right\} \tag{2.141}$$

is bounded in norm. It is also assumed that $h(\alpha)$ is twice continuously weakly differentiable on R and that the hessian matrix satisfies the condition

$$\left| [K(\alpha)u, u] \right| \leq M[u, u] \tag{2.142}$$

for every $\alpha \in R$ and $u \in H^p$, where

$$[K(\alpha)u, u] = \int_{t_o}^{t_1} \sum_{i,j=1}^{p} \frac{\partial^2 k(\alpha, t)}{\partial \alpha_i \partial \alpha_j} u_i(t)\, u_j(t)\, dt \tag{2.143}$$

and M is a positive real constant.

Define

$$\alpha^{(K+1)}(t) = \alpha^{(K)}(t) - \lambda_K \nabla k(\alpha^{(K)}(t)) \tag{2.144}$$

such that

$$h[\alpha^{(K)} - \lambda_K \nabla k(\alpha^{(K)})] \leq h[\alpha^{(K)} - \lambda \nabla k(\alpha^{(K)})] \tag{2.145}$$

for all $\lambda \geq 0$, where

$$\nabla k(\alpha) = \left(\frac{\partial k(\alpha, t)}{\partial \alpha_1}, \ldots, \frac{\partial k(\alpha, t)}{\partial \alpha_p} \right) . \tag{2.146}$$

Then

(1) A sequence $\left\{\alpha^{(K)}\right\}$ can be found such that

$$\lim_{k \to \infty} \alpha^{(K)} = \alpha^* \text{ in R kfor which } \nabla k(\alpha^*) = 0$$

(2) $h(\alpha^{(K)})$ decreases monotonically to $h(\alpha^*)$.

Proof. We wish to show that if the function $\alpha^{(K)}(t)$ belongs to R, $\alpha^{(K+1)}(t)$ also belongs to R; let $\alpha^{(K)}(t) \in R$ and $\nabla k(\alpha^{(K)}) \neq 0$, then the set R_K defined by

$$R_K = \left\{\alpha(t) \;\middle|\; h(\alpha) \leq h(\alpha^{(K)})\right\} \tag{2.147}$$

is a subset of R, since being in R, $h(\alpha^{(K)}) \leq h(\alpha^{(0)})$.

Let R be the set of $\alpha(t)$ of the form

$$\alpha(t) = \alpha^{(K)}(t) - \lambda u^{(K)}(t) \tag{2.148}$$

for $\lambda \geq 0$ and the function $\phi(\lambda)$ be defined as

$$\phi(\lambda) = h(\alpha^{(K)} - \lambda\, u^{(K)}) - h(\alpha^{(K)}) \qquad \text{for } \lambda \geq 0 \tag{2.149}$$

Expanding $h(\alpha^{(K)} - \lambda\, u^{(K)})$ into a Taylor series in λ about $\lambda = 0$, we find

$$\phi(\lambda) = -\lambda[\nabla k(\alpha^{(K)}),\, u^{(K)}] + \frac{1}{2}\lambda^2[K(\xi)u^{(K)},\, u^{(K)}] \;, \tag{2.150}$$

where

$$\xi = \beta\alpha^{(K)} + (1 - \beta)(\alpha^{(K)} - \lambda\, u^{(K)}) \;. \qquad 0 \leq \beta \leq 1 \tag{2.151}$$

From (2.150) we find the directional derivative of ϕ at $\alpha^{(K)}$

$$\left.\frac{d\phi}{d\lambda}\right|_{\lambda=0} = -[\nabla k(\alpha^{(K)}),\ u^{(K)}] \quad . \tag{2.152}$$

By Schwarz inequality, $\left.\frac{d\phi}{d\lambda}\right|_{\lambda=0}$ will be minimum if

$$u^{(K)} = \nabla k(\alpha^{(K)}) \quad . \tag{2.153}$$

Equation (2.153) provides the direction of steepest descent. Upon substituion of equation (2.153) into (2.150) we find

$$\phi(\lambda) = -\lambda[u^{(K)},\ u^{(K)}] + \frac{1}{2}\lambda^2[K(\xi)u^{(K)},\ u^{(K)}] \tag{2.154}$$

Knowing the fact that $\phi(0) = 0$ and $d\phi(0)/d\lambda < 0$ together with the boundedness of the set R and the continuity of $\phi(\lambda)$ there exists a finite interval $[0,\bar{\lambda}]$ such that

$$\phi(\bar{\lambda}) = 0 \tag{2.155}$$

otherwise all $\alpha(t)$ in (2.148) for $\lambda \geq 0$ will belong to R. In this case the norm of $\alpha(t)$ cannot be bounded. This contradicts the assumption in the theorem that the set $\alpha(t)$ in R is bounded in norm. Also, we have

$$\phi(\lambda) < 0 \quad \text{for} \quad \lambda \in (o,\bar{\lambda}) \tag{2.156}$$

Since $\phi(\lambda)$ is continuous and $\phi(\lambda) \leq 0$ for $\lambda \in [0,\bar{\lambda}]$, there exists a $\lambda \in [o,\bar{\lambda}]$, say λ_K, at which $\phi(\lambda_K)$ is minimum and $\phi(\lambda_K) < 0$. Thus the element $\alpha^{(K+1)}(t)$ defined by

$$\alpha^{(K+1)}(t) = \alpha^{(K)}(t) - \lambda_K u^{(K)}(t) \tag{2.157}$$

belongs to R_K and hence to R.

From (2.154) and (2.155), we find

$$\bar{\lambda} = \frac{2[u^{(K)},u^{(K)}]}{[K(\bar{\xi})u^{(K)},u^{(K)}]} , \tag{2.158}$$

where $\bar{\xi} = \beta\alpha^{(K)} + (1 - \beta)\,(\alpha^{(K)} - \bar{\lambda}\, u^{(K)})$.

Since $\bar{\lambda} > 0$, from equation (2.142), the following inequality holds

$$0 < [K(\bar{\xi})u^{(K)},u^{(K)}] \leq M[u^{(K)},u^{(K)}] , \tag{2.159}$$

and hence

$$\bar{\lambda} \geq \frac{2}{M} . \tag{2.160}$$

Next, we wish to show that the sequence $\{\alpha^{(K)}\}$ converges to α^* which also belongs to R. From (2.142) and (2.154) we find

$$\begin{aligned} \phi(\lambda) &\leq -\lambda(u^{(K)},u^{(K)}) + \tfrac{1}{2}\lambda^2 M(u^{(K)},u^{(K)}) \\ &= -(u^{(K)},u^{(K)})\,\lambda\,(1 - \tfrac{1}{2}\lambda M) \end{aligned} \tag{2.161}$$

The last expression assumes its maximum value at either end point of the interval $[\delta,\ (2/M) - \delta]$, a subinterval of $(0,\bar{\lambda})$, with $\delta > 0$ and $\bar{\lambda} \geq 2/M$. Thus for every K,

$$\phi(\lambda_K) \leq \phi(\lambda) \leq -\delta(1 - \tfrac{M\delta}{2})(u^{(K)},u^{(K)}) \tag{2.162}$$

since

$$\begin{aligned} \sum_{k=1}^{\infty} -\phi(\lambda_K) &= \sum_{k=1}^{\infty} [h(\alpha^{(K)}) - h(\alpha^{(K+1)})] \\ &\leq 2h(\alpha^{(1)}) < \infty \end{aligned} \tag{2.163}$$

It follows from (2.162) that

$$\sum_{K=1}^{\infty} (u^{(K)}, u^{(K)}) < \infty \quad . \tag{2.164}$$

Consequently

$$(u^{(K)}, u^{(K)}) \to 0 \quad \text{as} \quad K \to \infty \tag{2.165}$$

or from (2.153)

$$\nabla k(\alpha^{(K)}) \to 0 \quad \text{as} \quad K \to \infty \tag{2.166}$$

It can be shown that the sequence $\{\alpha^{(K)}(t)\}$ is a Cauchy sequence; for $n > m$ we find from (2.157) and (2.165)

$$\left|\left| \alpha^{(n)}(t) - \alpha^{(m)}(t) \right|\right| \leq \left|\left| \alpha^{(n)}(t) - \alpha^{(n-1)}(t) \right|\right| + \ldots +$$

$$\left|\left| \alpha^{(m+1)}(t) - \alpha^{(m)}(t) \right|\right|$$

$$= \lambda_{n-1} \left|\left| u^{(n-1)}(t) \right|\right| + \ldots + \lambda_m \left|\left| u^{(m)}(t) \right|\right|$$

$$\leq \bar{K} \sum_{K=m}^{n-1} \left|\left| u^{(K)}(t) \right|\right|$$

$$= \epsilon \text{ for } n, m > N \quad , \tag{2.167}$$

where $\bar{K} = \max [\lambda_{n-1}, \ldots, \lambda_m]$ and N is chosen to be a sufficiently large number.

Since $\alpha^{(K)}(t) \in H^p$ and H^p is complete, there exists a $\alpha^*(t)$ such that $\{\alpha^{(K)}(t)\}$ converges to $\alpha^*(t)$, or

$$\lim_{K \to \infty} \left|\left| \alpha^{(K)}(t) - \alpha^*(t) \right|\right| = 0 \tag{2.168}$$

and hence from (2.163)

$$\nabla k\,(\alpha^{(K)}) \to \nabla k(\alpha^{*}) = 0 \qquad (2.169)$$

Note that the element $\alpha^{*}(t)$ also belongs to R since $h(\alpha^{*}) < h(\alpha^{(0)})$. Also that $h(\alpha^{(K)})$ decreases monotonically to $h(\alpha^{*})$ follows from the fact that $\phi(\lambda_K) < 0$ for every K.

Remark 1. In order to use the above theorem it is necessary to solve for λ_K at which the function $\phi(\lambda)$ is minimum for each iteration. This may be difficult if $h(\alpha)$ is not a quadratic function. A slight modification in the above procedure will enable one to avoid finding the exact value of λ_K. From equation (2.159) it can be seen that if λ_K is taken to be any value in the interval $\delta \leq \lambda_K \leq (2/M) - \delta$ the condition $\phi(\lambda_K) < 0$ is still valid and hence $h(\alpha^{(K+1)}) < h(\alpha^{(K)})$ for each K.

Remark 2. The above theorem can be considered as a modification of Kantarovich's steepest descent method in function space[32].

Example The rotational motion of a body in space is described by the equations[38,39]

$$\begin{aligned} I_1\dot{X}_1(t) &= (I_2 - I_3)\,X_2(t)\,X_3(t) + \tau_1(t) \\ I_2\dot{X}_2(t) &= (I_3 - I_1)\,X_3(t)\,X_1(t) + \tau_2(t) \\ I_3\dot{X}_3(t) &= (I_1 - I_2)\,X_1(t)\,X_2(t) + \tau_3(t) \end{aligned} \qquad (2.169)$$

where 1, 2, 3 are the axes fixed on the body passing through its center of gravity as indicated in Fig. 2.2. I_1, I_2, I_3 are the moments of

inertia and $X_1(t)$, $X_2(t)$, $X_3(t)$ are angular velocities of the body about the axes 1, 2, 3, respectively. τ_1, τ_2, τ_3 are the torques applied about the 1, 2, 3 axes, respectively.

If $I_1 = I_2 = I$, then the body has a single axis of symmetry, the 3 axis. Equations (2.169) reduce to

$$\begin{aligned} \dot{X}_1(t) &= \frac{I-I_3}{I} X_2(t) X_3(t) + \frac{\tau_1(t)}{I} \\ \dot{X}_2(t) &= -\frac{I-I_3}{I} X_3(t) X_1(t) + \frac{\tau_2(t)}{I} \\ \dot{X}_3(t) &= \frac{\tau_3(t)}{I_3} \end{aligned} \tag{2.170}$$

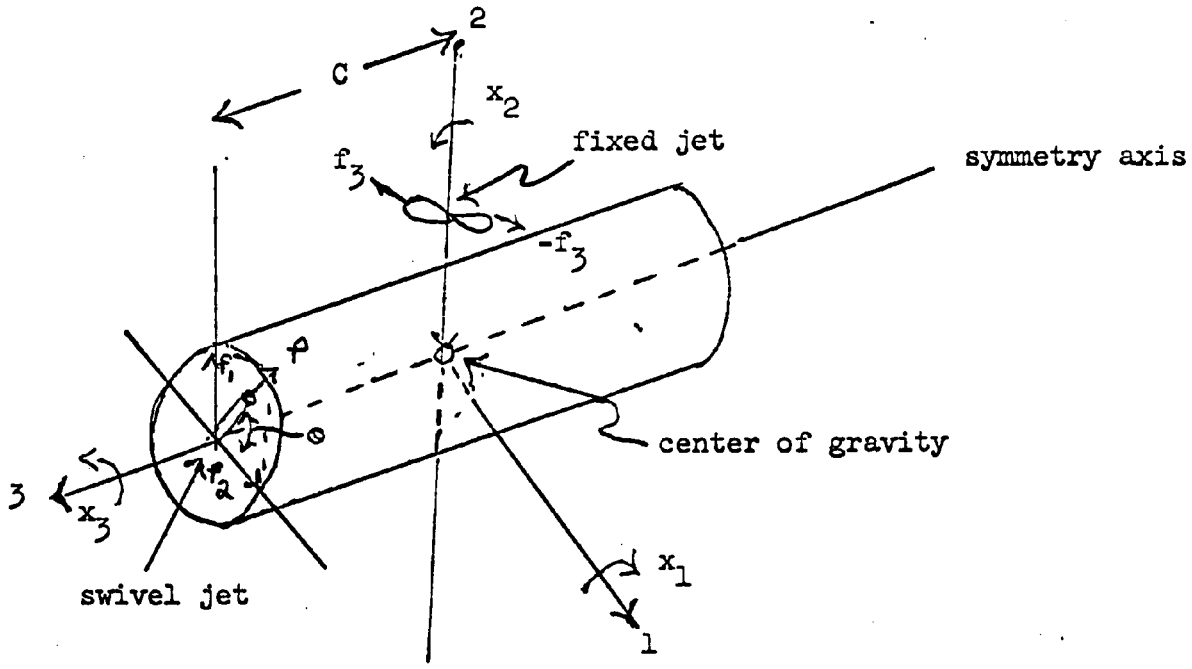


Fig. 2.2 Rotational Body in Space.

The torque is obtained by a single reaction jet, with a thrust $f_3(t)$ such that $\tau_3(t) = C_3 f_3(t)$. The torques $\tau_1(t)$ and $\tau_2(t)$ are provided by a swivelled reaction jet with a thrust $f(t)$. If $\theta(t)$ is the angle between $f(t)$ and the 1 axis, then the components of f along the 1, 2 axes become

$$\begin{aligned} f_1(t) &= | f(t) | \sin\theta(t) \\ f_2(t) &= | f(t) | \cos\theta(t) \end{aligned} \tag{2.171}$$

If the torque $\tau_3(t)$ is known, then $X_3(t)$ is a known function of time. Equations (2.170) reduce to

$$\begin{bmatrix} \dot{X}_1(t) \\ \dot{X}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & \alpha(t) \\ -\alpha(t) & 0 \end{bmatrix} \begin{bmatrix} X_1(t) \\ X_2(t) \end{bmatrix} + \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \tag{2.172}$$

where

$$\alpha(t) \triangleq \frac{I-I_3}{I} X_3(t) \quad , \qquad u_1(t) \triangleq \frac{\tau_1(t)}{I} = \frac{Cf_1(t)}{I}$$

$$u_2(t) \triangleq \frac{\tau_2(t)}{I} = \frac{Cf_2(t)}{I} \quad .$$

If the thrust delivered by the swivelled jet is limited in magnitude

$$\left|\left| f(t) \right|\right| = \left[f_1^2(t) + f_2^2(t) \right]^{1/2} \leq m \quad , \tag{2.173}$$

then $u(t)$ is also limited in magnitude

$$\left|\left| u(t) \right|\right| \leq m_1 = \frac{C}{I} m \quad . \tag{2.174}$$

It can be shown that the system (2.172) is norm-invariant, that is, the unforced part

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & \alpha(t) \\ -\alpha(t) & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_1 \\ x_2 \end{bmatrix} \tag{2.175}$$

has the property

$$\left|\left|X(t)\right|\right| = \left|\left|X(0)\right|\right| \quad \text{for all } t \geq 0 \tag{2.176}$$

since from (2.175)

$$\frac{d}{dt}\left|\left|X(t)\right|\right| = \frac{[\dot{X}(t), X(t)]}{X(t)} =$$

$$= \frac{\alpha(t)X_2(t)X_1(t) - \alpha(t)X_2(t)X_1(t)}{\left|\left|X(t)\right|\right|} = 0$$

It is known[39] that the optimal control $u^*(t)$ can be found such that it drives any initial state $X(0) = \xi$ to 0 in minimum time with minimum fuel

$$F = \int_0^t K \left|\left|u(t)\right|\right| dt \tag{2.177}$$

The optimal trajectory $X^*(t)$ is the solution of

$$\begin{bmatrix} \dot{X}_1^*(t) \\ \dot{X}_2^*(t) \end{bmatrix} = \begin{bmatrix} 0 & \alpha(t) \\ -\alpha(t) & 0 \end{bmatrix} \begin{bmatrix} X_1^*(t) \\ X_2^*(t) \end{bmatrix} + \begin{bmatrix} u_1^*(t) \\ u_2^*(t) \end{bmatrix} \tag{2.178}$$

Any deviation in $\alpha(t)$ will cause the actual trajectory to be different from that of the optimal one. In order to determine the actual $\alpha(t)$, we denote the actual trajectory by $X_a(t)$. Since $X_a(t)$ is assumed to be known, its derivative $\dot{X}_a(t)$ can be determined. We form the functional

$$h(\alpha) = \int_0^T [\dot{X}_{a_i}(t) - F_i(\alpha\, t)]^2 \, dt \quad , \tag{2.179}$$

where

$$F_1(\alpha, t) = \alpha(t)\, X_{a_2}(t) + u_1^*(t)$$

and

$$F_2(\alpha, t) = -\alpha(t)\, X_{a_1}(t) + u_2^*(t) \tag{2.180}$$

Using the notations introduced above, we find

$$k(\alpha, t), t) = \sum_{i=1}^{2} [\dot{X}_{a_i}(t) - F_i(\alpha, t)]^2 \tag{2.181}$$

To minimize $h(\alpha)$ we need the first and second partial derivatives of $k(\alpha t), t)$ with respect to $\alpha(t)$.

$$\frac{\partial k}{\partial \alpha}(\alpha(t),t) = -2\left[\dot{x}_{a_1}(t) - F_1(\alpha(t), \;]\, x_{a_2}(t) + 2\left[\dot{x}_{a_2}(t) - F_2(\alpha,t)\right] x_{a_1}(t)$$

$$\frac{\partial^2 k}{\partial \alpha^2}(\alpha(t),t) = 2\left[x_{a_1}{}^2(t) + x_{a_2}{}^2(t)\right] \tag{2.182}$$

therefore

$$M = 2 \max_{t\epsilon[o,T]} \left|\left|X_a(t)\right|\right|^2 \tag{2.183}$$

Let $\alpha^{(K)}(t)$ be the Kth approximation, then

$$\alpha^{(K+1)}(t) = \alpha^{(K)}(t) - \lambda_K \nabla k(\alpha^{(K)}(t)) \tag{2.184}$$

with λ_K restricted to lie in the interval

$$0 \leq \delta \leq \lambda_K \leq \frac{2}{M} - \delta \tag{2.185}$$

In this chapter the change in cost and the variation in trajectory due to changes in plant parameters are evaluated by means of the solution of the equations of first variation of the original system. The error involved in using the approximate solution of the system instead of the exact solution is discussed. The inverse problem of estimating the variations in plant parameters corresponding to a given variation along the optimal trajectory is explored. This covers the case of constant variations in plant parameters as well as the cases of time-varying parameters.

CHAPTER III

Analysis of Variation in Trajectory and Cost of Closed-Loop Optimal Control Systems Due to Deterministic Changes of Plant Parameter

The analysis of variation in trajectory and cost as well as the estimation of actual parameter variations under different situations are repeated in this chapter for closed loop optimal control systems. It is found that the methods developed in the previous chapter can be applied directly to the closed-loop control systems as well.

3.1 Specific Problem

A closed-loop optimal control system can be described schematically in Fig. 3.1 in which the behavior of the plant is assumed to be governed by a set of differential equations

$$\frac{dx_i}{dt} = f_i(x_1,\ldots,x_n, \alpha_{1o},\ldots,\alpha_{p_o}, u_1,\ldots,u_r, t)$$

$$x_i(0) = x_{io} \qquad i = 1,\ldots,n \tag{3.1}$$

or in vector form

$$\frac{dx}{dt} = f(x,\alpha_o,u,t)$$

$$x(0) = x_o \tag{3.2}$$

where x = n-dimensional state vector describing the instantaneous behavior of the plant (object to be controlled)

α_o = p-dimensional vector representing the set of nominal plant parameters

u = r-dimensional control vector

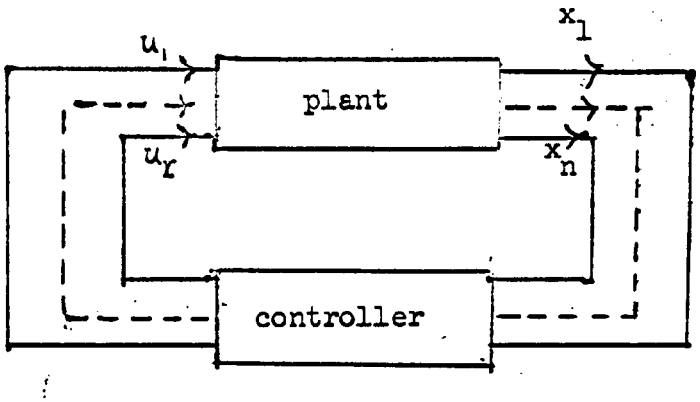


Fig. 3.1 Basic Closed-Loop System.

The cost to be minimized is given by the functional

$$J = \int_{t_o}^{t_1} f_o(x,u)\, dt \qquad (3.3)$$

It is assumed that the following conditions are satisfied:

a) the functions f_o and f_i $(i = 1,\ldots,n)$ as well as their first partial derivatives with respect to x_j and α_K $(j = 1,\ldots,n;\ K = 1,\ldots,p)$ are defined and continuous in some domain Γ of the space R of variables $x_1,\ldots,x_n,\ \alpha_1,\ldots,\alpha_p,\ u_1,\ldots,u_r$. In other words, the space R can be taken to be the Cartesian product $R = X \times A \times U \times T$

b) the optimal control u(t), which transfers the state of the system (3.2) from a given initial position x_o to the final position x_1 that is restricted to lie on a given hypersurface S and which at the same time minimizes the functional (3.3), is a continuous function of the state variables $u(t) = u(\bar{x}(t))$. Its partial derivatives

$$\frac{\partial u_k}{\partial x_j} \quad (j = 1,\ldots,n,\ k = 1,\ldots,r)$$

are also assumed to be continuous.

3.2 Evaluation of Change in Cost and of Variation in Trajectory Due to Changes in Plant Parameters.

Just as in Chapter 2, the actual plant parameter vector $\alpha = (\alpha_1,\ldots,\alpha_p)$ of the physical system cannot be expected to be identical to the nominal vector $\alpha_o = (\alpha_{o_1},\ldots,\alpha_{o_p})$. Instead of (3.2) the actual behavior of the plant is given by

$$\frac{dy}{dt} = f(y,\alpha,\ u(y),t)$$

$$y(o) = y_o \ . \tag{3.4}$$

By expanding the right hand side of (3.4) about the point $(x,\ \alpha_o,\ u(x),t)$ with the second and higher order terms neglected we find

$$\frac{dy_i}{dt} = f_i(x,\ \alpha_o,\ u(x),t) + \sum_{j=1}^{n}\left(\frac{\partial f_i}{\partial x_j} + \sum_{k=1}^{r}\frac{\partial f_i}{\partial u_k}\ \frac{\partial u_k}{\partial x_j}\right)(y_j - x_j)$$

$$+\sum_{s=1}^{p}\frac{\partial f_i}{\partial \alpha_s}\ (\alpha_s - \alpha_{os}) \qquad n=1,\ldots,n \tag{3.5}$$

Subtracting equation (3.1) from (3.5) we obtain

$$\frac{dZ_i}{dt} = \sum_{j=1}^{n} \left(\frac{\partial f_i}{\partial x_j} + \sum_{k=1}^{r} \frac{\partial f_i}{\partial u_k} \frac{\partial u_k}{\partial x_j} \right) Z_j + \sum_{s=1}^{p} \frac{\partial f_i}{\partial \alpha_s} \gamma_s \qquad (3.6)$$

$$Z_i(0) = Z_{io} \ , \qquad i = 1, \cdots, n$$

where $Z_i = y_i - x_i$, the ith component of the variation along the optimal trajectory,

and

$\gamma_s = \alpha_s - \alpha_{os}$, the variation of the parameter α_s about its nominal value.

Equation (3.6) can be written in vector form as

$$\frac{dZ}{dt} = A(t)\, Z + B(t)\, \gamma(t) \ , \qquad (3.7)$$

$$Z(0) = Z_o \ ,$$

where

$$A(t) = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial u} \frac{\partial u}{\partial x}$$

$$B(t) = \frac{\partial f}{\partial \alpha}$$

and

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1}, \ldots, \frac{\partial f_1}{\partial x_n} \\ \cdot \quad \cdot \quad \cdot \\ \frac{\partial f_n}{\partial x_1}, \ldots, \frac{\partial f_n}{\partial x_n} \end{bmatrix} \quad \frac{\partial f}{\partial u} = \begin{bmatrix} \frac{\partial f_1}{\partial u_1}, \ldots, \frac{\partial f_1}{\partial u_r} \\ \cdot \quad \cdot \quad \cdot \\ \frac{\partial f_n}{\partial u_1}, \ldots, \frac{\partial f_n}{\partial u_r} \end{bmatrix} \tag{3.8}$$

$$\frac{\partial u}{\partial x} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1}, \ldots, \frac{\partial u_1}{\partial x_n} \\ \cdot \quad \cdot \quad \cdot \\ \frac{\partial u_r}{\partial x_1}, \ldots, \frac{\partial u_r}{\partial x_n} \end{bmatrix}$$

The partial derivatives in (3.8) are evaluated along the optimal trajectory.

Just as in Section 2.2, the variation along the optimal trajectory is is given by

$$Z(\alpha_o, t) = \Phi(t)Z_o + \int_0^t \Phi(t,\tau)\, B(\tau)\, \gamma(\tau)\, d\tau \ , \tag{3.9}$$

where $\Phi(t)$ is the transition matrix of (3.7) satisfying the differential equation

$$\dot{\Phi} = A(t)\, \Phi \ , \qquad \Phi(0) = I \tag{3.10}$$

The corresponding change in cost is given by

$$\Delta J = \int_0^T \left(\frac{\partial f_o}{\partial x} + \frac{\partial f_o}{\partial u} \frac{\partial u}{\partial x} \right) Z \, dt \quad , \tag{3.11}$$

with $\partial f_o/\partial x$ and $\partial f_o/\partial u$ given by

$$\frac{\partial f_o}{\partial x} = \left(\frac{\partial f_o}{\partial x_1}, \ldots, \frac{\partial f_o}{\partial x_n} \right) \quad , \quad \frac{\partial f_o}{\partial u} = \left(\frac{\partial f_o}{\partial u_1}, \ldots, \frac{\partial f_o}{\partial u_r} \right) \tag{3.12}$$

The partial derivatives are again evaluated at the optimal path.

3.3 Estimation of Plant Parameters with Known Trajectory.

Sections 2.3 through 2.5 inclusive can be applied to the estimation of plant parameters of closed loop optimal control systems since in both cases the variation along the optimal trajectory was described by a linear time varying vector differential equation.

In the case that the actual trajectory differs substantially from that of the idealized system we can rewrite equation (3.1) into the following form

$$x_{a_i}(t) = F_i(\alpha, t) = \int_{t_o}^{t} f_i(x, \alpha, u(x), t) \, dt \tag{3.12}$$

$$n=1, \ldots, n$$

and form the functional

$$f(\alpha) \triangleq \int_{t_o}^{t_1} \sum_{i=1}^{n} [x_{a_i}(t) - F_i(\alpha, t)]^2 \, dt$$

Using the successive approximation method developed in Section (2.6) one can readily find a set of parameters corresponding to the actual trajectory of the system.

CHAPTER IV

The Fokker-Planck Equation

In this chapter a continuous multidimensional Markov process is defined. The probabilistic characterization of a Markov process is given. The Fokker-Planck equation corresponding to a linear time-varying system excited by white noise is derived. The reduction of the Fokker-Planck equation associated with a given system of linear time-varying differential equations excited by white noise to the solution of a set of ordinary linear differential equations with variable coefficients is carried out by Fourier transformation and the method of characteristics. By means of the Fokker-Planck equation the expected values as well as the moments of the solution of the given system can be determined even when the initial position of the system is not exactly known, but only given by its probability distribution. Since a white noise random process can be considered only as a limiting case which will never occur in practice, we shall in the later part of the chapter, analyze somewhat more general problems, that is, linear time-varying systems with disturbances generated by differential equations excited by white noise. The analysis is based on the method of expanding the state space of the original system into a higher-dimensional state space[40]. By doing so we will be able to determine the behavior of the system when the disturbances are non-white as well as non-stationary.

4.1 Probabilistic Characterization of a Markov Process.

Let $P(t,x/\tau,y)dy$ be the conditional probability that at time τ the random vector $\xi(\tau) = (\xi_1(\tau),\ldots,\xi_n(\tau))\epsilon[y,y+dy]$ under the condition that at t, $\xi(t) = x$. Then a Markov process can be defined by stating that for

such a process the conditional probability $P(t,x/\tau,y)dy$ that the random vector $\xi(t) = (\xi_1(t),\ldots,\xi_n(t))$ assumes the values in $[y_i, y_i + dy_i]$ $(i=1,\ldots,n)$ at time τ depends besides on y_i and τ only on the vector ξ at the previous time t. Any additional information concerning the states of the system at instants of time earlier than t for the process has no effect on the function $P(t,x/\tau,y)$. In other words, a Markov process is completely described by the conditional probability density function $p(t,x/\tau,y)$.

Like any conditional probability density function, $P(t,x/\tau,y)$ must satisfy the following conditions

$$P(t,x/\tau,y) = P(t,x_1,\ldots,x_n \,/\, \tau,\, y_1,\ldots,y_n) \;\geq\; 0$$

and

$$\int P(t,x/\tau,y)dy$$

$$= \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} P(t,x_1,\ldots,x_n \,/\, \tau,y_1,\ldots,y_n)dy_1,\ldots,dy_n = 1. \qquad (4.1)$$

It is assumed in what follows that any integration is to be carried out over the entire n-dimensional state space R_n unless the limits of integration are specified otherwise.

Let us divide the domain of definition of the random vector $\xi(t)$ into the subregions $\xi \in [z^i, z^{i+1})$ by means of the points $z^i(i=0,\pm1,\pm2,\ldots)$ and consider the instants of time, t, s, τ such that $t < s < \tau$; the conditional probability density function $P(t,x/\tau,y)$ can be expressed as an infinite sum

$$P(t,x\,\tau,y) = \sum_{i=-\infty}^{\infty} [F(t,x/s,z^i + \Delta z^i) - F(t,x/s,z^i)]\, P(s,\bar{z}/\tau,y) \quad , \qquad (4.2)$$

where $F(t,x/s,z^i)$ is the probability distribution function such that

$$F(t,x/s,z^i) = P(\xi_1 < z_1^i, \ldots, \xi_n < z_n^i \text{ at time } s \text{ given that}$$

$$\xi_1 = x_1, \ldots, \xi_n = x_n \text{ at time } t) \quad , \qquad (4.3)$$

$\Delta z^i = z^{i+1} - z^i$, and $\bar{z}^i \in [z^i, z^i + \Delta z^i]$. As $\max |\Delta z^i| \to 0$, equation (4.2) reduces to

$$P(t,x/\tau,y) = \int dF_z(t,x/s,z)\, P(s,z/\tau,y) = \int P(s,z/\tau,y)\, P(t,x/s,z)\, dz \quad . \qquad (4.4)$$

Equation (4.4) is called the generalized Markov equation, the Markov identity or the Smoluchowski equation.

4.2 Definition of a Continuous Process

We say that a stochastic process $\xi(t) = (\xi_1(t), \ldots, \xi_n(t)$ is continuous if there is only a small probability that the coordinates of the random point take on appreciable increments in a short interval of time. However, in order to use the Fokker-Planck equation we shall need a stronger continuity for a Markov process; namely, we require that the relation

$$\lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_{|y-x| \geq \delta} P(t-\Delta t, x/t, y)\, dy = 0 \qquad (4.5)$$

holds for every positive number δ. The quantity $|y-x|$ can be considered as the Euclidean length in R_n.

4.3 The Fokker-Planck Equation

The derivation of Fokker-Planck equation for a one-dimensional Markov process has been treated very elegantly by Gnedenko[41]. The same method can be used to derive the Fokker-Planck equation for a Markov process when the random point is a vector $\xi(t) = (\xi_1(t),\ldots,\xi_n(t))$.

In addition to condition (4.5) we further assume that:

a) for any $\delta > 0$, the limits

$$\lim_{\Delta t\to 0} \frac{1}{\Delta t} \int_{|y-x|<\delta} (y_i-x_i)\, P(t-\Delta t,x/t,y)dy = a_i(t,x) \qquad (4.6)$$

$$i = 1,\ldots,n$$

$$\lim_{\Delta t\to 0} \frac{1}{\Delta t} \int_{|y-x|<\delta} (y_i- x_i)(y_j- x_j)\, P(t-\Delta t,x/t,y)dy = b_{ij}(t,x) \qquad (4.7)$$

$$i,j, = 1,\ldots,n$$

exist and the convergence of (4.6) and (4.7) is uniform in x.

b) the partial derivatives

$$\frac{\partial P(t,x/\tau,v)}{\partial \tau} \quad , \quad \frac{\partial}{\partial y_i}\,[a_i(\tau,y)\, P(t,x/\tau,y)] \quad ,$$

$$\frac{\partial^2}{\partial y_i \partial y_j}\,[b_{ij}(\tau,y)\, P(t,x/\tau,y)] \qquad (4.8)$$

exist and are continuous for all t, x, $\tau > t$ and y.

Under the assumptions (4.5) through (4.8) it can be shown (see Appendix C) that function $P(t,x/\tau,y)$ satisfies the following second order parabolic differential equation (Fokker-Planck equation)

$$\frac{\partial P(t,x/\tau,y)}{\partial \tau} = -\sum_{i=1}^{n} \frac{\partial}{\partial y_i} [a_i(\tau,y)\ P(t,x/\tau,y)]$$

$$+\frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial y_i \partial y_j} [b_{ij}(\tau,y)\ P(t,x/\tau,y)] \quad (4.9)$$

4.4 The Fokker-Planck Equation Corresponding to a Linear Time-Varying System Excited by White Noise

Let us consider a linear time-varying differential equation of the following form

$$\frac{dZ_i}{dt} = \sum_{k=1}^{n} c_{ik}(t)\ Z_k + \sum_{k=1}^{n} d_{ik}(t)\ n_k(t) \quad , \qquad (4.10)$$

$$Z_i(0) = Z_{oi} \quad , \quad i = 1,\ldots,n$$

where $c_{ik}(t)$ and $d_{ik}(t)$ are assumed to be continuous functions of $t \geq 0$ and $n_k(t)$ are the components of a random noise vector $n(t) = (n_1(t),\ldots, n_n(t))$. Equations (4.10) can be put into vector form

$$\frac{dZ}{dt} = C(t)\ Z + D(t)\ n(t)$$
$$Z(0) = Z_o \quad . \qquad (4.11)$$

In order to derive the Fokker-Planck equation corresponding to the system (4.11) we shall need the following assumptions on the random noise vector $n(t)$.

a) Each component $n_i(i = 1,\ldots,n)$ is a white noise with a constant spectral density of $4D_i$, or $E[n_i(t)] = 0 \quad i = 1,\ldots,n$ (4.12)

and

$$E[n_i(t_1)\ n_i(t_2)] = 2D_i\ \delta(t_1 - t_2) \quad i = 1,\ldots,n \qquad (4.13)$$

where $\delta(t_1 - t_2)$ is the Dirac singular function. Relation (4.13) is valid because the spectral density and auto correlation function form a Fourier Transform pair.[42]

b) The components $n_i(t)$ and $n_j(t)$, $i \neq j$, are not correlated so that

$$E[n_i(t_1)\, n_j(t_2)] = 0 \qquad i,j = 1,\ldots,n \tag{4.14}$$

To determine a_i and b_{ij} of the Fokker-Planck equation we integrate (4.10) over a very short time interval Δt

$$\Delta Z_i \cong \sum_{k=1}^{n} c_{ik}(t)\, Z_k\, \Delta t + \sum_{k=1}^{n} \int_t^{t+\Delta t} d_{ik}(t)\, n_k(t)\, dt \tag{4.15}$$

and

$$\begin{aligned}
\Delta Z_i\, \Delta Z_j \cong & \left(\sum_{k=1}^{n} c_{ik}(t) Z_k\, \Delta t \right)\left(\sum_{l=1}^{n} c_{jl}(t)\, Z_l \Delta t \right) \\
& + \left(\sum_{K=1}^{n} c_{iK}(t)\, Z_K \Delta t \right) \int_t^{t+\Delta t} \sum_{l=1}^{n} d_{jl}(t)\, n_l(t)\, dt \\
& + \left(\sum_{l=1}^{n} c_{jl}(t)\, Z_l \Delta t \right) \int_t^{t+\Delta t} \sum_{K=1}^{n} d_{iK}(t)\, n_K(t)\, dt \\
& + \left(\int_t^{t+\Delta t} \sum_{K=1}^{n} d_{iK}(t)\, n_K(t)\, dt \right) \int_t^{t+\Delta t} \sum_{l=1}^{n} d_{jl}(t) n_l(t)\, dt
\end{aligned} \tag{4.16}$$

Let $P^*\, dz\, dn$ be the probability of the random vector $(\xi, \xi^*) = (\xi_1(t), \xi_2(t), \ldots, \xi_n^*(t), \ldots, \xi_n^*(t))$ such that at time $t+\Delta t$

$$\xi(t+\Delta t) \in [Z,\ Z+dZ]$$

and

$$\xi^*(t+\Delta t) \in [n,\ n+dn]$$

under the condition that at time t, $\xi(t) = y$. Then the mean of equation (4.15) can be written in the form

$$\int_R Z_i\, P^*\, dZ\, dn = \int_R \sum_{k=1}^{n} c_{ik}(t)\, Z_k \Delta t\, P^*\, dZ\, dn + \int_R \sum_{k=1}^{n} \int_t^{t+\Delta t} d_{ik}(t)\, n_k(t)\, P^*\, dZ\, dn \quad , \qquad (4.17)$$

where R is the space of the 2n variables $(\xi,\ \xi^*)$. The right hand side of (4.17) reduces to

$$\sum_{k=1}^{n} c_{ik}(t)\, Z_k\, \Delta t + \sum_{k=1}^{n} \int_t^{t+\Delta t} d_{ik}(t) \int n_k(t)\, P(n)\, dn \quad ,$$

where P(n) is the probability density function of the random noise vector, and the left hand side of (4.17) can be written as

$$\int \Delta Z_i\, P(t, y/t+\Delta t,\ Z)\, dZ = E\,[\,\Delta Z_i\,] \quad .$$

Consequently, (4.17) reduces to the following form

$$E\,[\,\Delta Z_i\,] = \sum_{k=1}^{n} c_{ik}(t)\, Z_k\, \Delta t + \sum_{k=1}^{n} \int_t^{t+\Delta t} d_{ik}(t)\, E[n_k(t)]\, dt \quad .$$

It can be seen from (4.6) and (4.12) that

$$a_i(t,Z) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} E[\Delta Z_i] = \sum_{K=1}^{n} c_{ik}(t) Z_k , \qquad (4.18)$$

$$i = 1,\ldots,n$$

and, similiarly, from (4.7)

$$b_{ij}(t,Z) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} E[\Delta Z_i \Delta Z_j]$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \int_t^{t+\Delta t} \sum_{k,\ell=1}^{n} d_{ik}(\xi) d_{j\ell}(\alpha) E(n_k(\xi) n_\ell(\alpha)) \, d\xi \, d\alpha$$

$$i,j = 1,\ldots,n \qquad (4.19)$$

Making use of the conditions (4.13) and (4.14) we find that

$$b_{ij}(t,Z) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \int_t^{t+\Delta t} \sum_{K=1}^{n} d_{iK}(\xi) d_{jK}(\alpha) \, 2D_K \, \delta(\xi-\alpha) \, d\xi \, d\alpha$$

$$= \lim_{\Delta t \to 0} \frac{1}{\Delta t} \int_t^{t+\Delta t} \sum_{K=1}^{n} d_{iK}(\xi) \, d_{jK}(\xi) \, 2D_K \, d\xi .$$

Hence

$$b_{ij}(t,Z) = \sum_{K=1}^{n} 2D_K \, d_{iK}(t) \, d_{jK}(t) \triangleq 2f_{ij}(t) , \qquad (4.20)$$

$$i,j = 1,\ldots,n$$

where

$$f_{ij}(t) = \sum_{k=1}^{n} D_k \, d_{ik}(t) \, d_{jk}(t) \quad . \tag{4.21}$$

The Fokker-Planck equation (4.9) corresponding to the system (4.10) reduces to

$$\frac{\partial P}{\partial t} = - \sum_{i,k=1}^{n} c_{ik}(t) \frac{\partial}{\partial Z_i} (Z_k P) + \sum_{i,j=1}^{n} f_{ij}(t) \frac{\partial^2 P}{\partial Z_i \partial Z_j} \quad , \tag{4.22}$$

which for $t = 0$ becomes

$$P\Big|_{t=0} = \delta(Z-Z_o) = \prod_{n=1}^{n} \delta(Z_i - Z_{oi}) \quad . \tag{4.23}$$

4.5 Reduction of the Fokker-Planck Equation to Solving a Set of Ordinary Differential Equations with Variable Coefficients

To find the fundamental solution of (4.22) with the initial condition (4.23) we shall generalize the method of Wang and Uhlenbeck[43] which was first used in the theory of Brownian motion with Brownian particles governed by a set of linear differential equations with constant coefficients. We assume that the conditional probability density function satisfies the following condition.

$$\lim_{\text{any } Z_i \to \pm \infty} P = 0 \tag{4.24}$$

Let us define the Fourier transform of the conditional probability density function P by

$$g(\xi,t) = \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dZ_1 \cdots dZ_n \, P \exp\left[-i \sum_{k=1}^{n} \xi_k Z_k\right] . \qquad (4.25)$$

Taking the Fourier transform of both sides of (4.22) we find

$$\frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial t} = -\sum_{i,K=1}^{n} c_{iK}(t) \int \cdots \int_{-\infty}^{\infty} \frac{\partial}{\partial Z_i} [Z_k P] \exp[-i \sum_{k=1}^{n} \xi_k Z_k] dZ_1 \cdots dZ_n$$

$$+ \sum_{i,j=1}^{n} f_{ij}(t) \int_{-\infty}^{\infty} \cdots \int \frac{\partial^2 P}{\partial Z_i \partial Z_j} \exp[-i \sum_{k=1}^{n} \xi_k Z_k] dZ_1 \cdots dZ_n \, . \qquad (4.26)$$

Integrating the right hand side of (4.26) by parts with respect to Z_i and Z_j and making use of the condition (4.24) we find

$$\frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial t} = \sum_{i,K=1}^{n} c_{iK}(t) \int_{-\infty}^{\infty} \cdots \int (-i\xi_i) Z_k P \exp[-i \sum_{k=1}^{n} \xi_k Z_k] dZ_1 \cdots dZ_n$$

$$- \sum_{i,j=1}^{n} f_{ij}(t) \int_{-\infty}^{\infty} \cdots \int \xi_i \xi_j P \exp\left[-i \sum_{k=1}^{n} \xi_k Z_k\right] dZ_1 \cdots dZ_n \, .$$

Making use of (4.25), the above equation reduces to

$$\frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial t} = \sum_{i,j=1}^{n} c_{ij}(t) \, \xi_i \frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial \xi_j}$$

$$- g(\xi_1,\ldots,\xi_n,t) \sum_{i,j=1}^{n} f_{ij}(t) \, \xi_i \xi_j \, . \qquad (4.27)$$

Equation (4.27) is a linear first order partial differential equation and its solution can be obtained by the method of characteristics. The auxiliary system of equations[44] (Lagrange subsidiary equations) is of the following form:

$$\frac{dt}{1} = -\frac{d\xi_j}{\sum_{i=1}^{n} c_{ij}(t)\xi_i} = \frac{-dg}{g\sum_{i,j=1}^{n} f_{ij}(t)\xi_i\xi_j} \quad . \tag{4.28}$$

Equation (4.28) can be written out explicitly as the following set of differential equations

$$\begin{aligned}
\frac{d\xi_1}{dt} &= -c_{11}(t)\xi_1 - \cdots - c_{n1}(t)\xi_n \\
&\cdots \\
\frac{d\xi_n}{dt} &= -c_{1n}(t)\xi_1 - \cdots - c_{nn}(t)\xi_n \\
\frac{dg}{dt} &= -g\sum_{i,j=1}^{n} f_{ij}(t)\xi_i\xi_j
\end{aligned} \tag{4.29}$$

Writing (4.29) in matrix form, we have,

$$\begin{aligned}
\dot{\xi} &= -C'(t)\xi \quad , \\
\dot{g} &= -g\sum_{i,j=1}^{n} f_{ij}(t)\xi_i\xi_j \quad ,
\end{aligned} \tag{4.30}$$

where $C'(t)$ is the transpose of the matrix $C(t)$ expressed in (4.11). It can be seen that the first equation of (4.30) is the adjoint equation of

of the linear system (4.11).

Let $\psi(t,t_o)$ be the fundamental matrix solution of the first equation of (4.30), that is, $\psi(t,t_o)$ is the solution of

$$\dot{\psi}(t,t_o) = -C'(t)\,\psi(t,t_o) \ ,$$

$$\psi(t_o,t_o) = I \ , \tag{4.31}$$

where I is the identity matrix. Then the solution of the first equation of (4.30) is given by

$$\xi(t) = \psi(t,t_o)\,\xi_o \tag{4.32}$$

or in terms of components

$$\xi_i(t) = \sum_{j=1}^{n} \psi_{ij}(t,t_o)\xi_{oj} \ , \qquad i = 1,\ldots,n \tag{4.33}$$

where $\psi_{ij}(t,t_o)$ is the ij^{th} element of the matrix ψ, and ξ_{oj} is the jth component of an arbitrary initial vector ξ_o. By the property of nonsingularity of fundamental matrix[45] we find from (4.32)

$$\xi_o = \psi^{-1}(t,t_o)\,\xi(t) = \Phi(t,t_o)\,\xi(t) \tag{4.34}$$

or

$$\xi_{oi} = \sum_{j=1}^{n} \phi_{ij}(t,t_o)\,\xi_j(t) \ , \qquad i = 1,\ldots,n \tag{4.35}$$

where $\phi_{ij}(t,t_o)$ is the ij^{th} element of $\Phi(t,t_o)$, which is the transpose of the fundamental matrix of the system (4.11).

The solution of the second equation of (4.30) can be found as follows:

$$\frac{dg}{g} = - \sum_{i,j=1}^{n} f_{ij}(t)\, \xi_i \xi_j dt$$

$$g = g_o \exp\left\{- \sum_{i,j=1}^{n} \int_{t_o}^{t} f_{ij}(t)\, \xi_i(t)\, \xi_j(t)\, dt\right\} . \qquad (4.36)$$

Upon substitution of (4.33) into (4.36) we find

$$g = g_o \exp\left\{- \sum_{i,j=1}^{n} \int_{t_o}^{t} f_{ij}(t) \sum_{k,\ell=1}^{n} \psi_{ik}(t,t_o)\, \psi_{j\ell}(t,t_o)\, \xi_{ok} \xi_{o\ell}\, dt\right\}$$

Using relation (4.35) with $t_o = 0$ we have

$$g = g_o \exp\left\{- \sum_{\substack{i,j=1 \\ k,\ell}}^{n} \sum_{r,s=1}^{n} \phi_{kr}(t)\phi_{\ell s}(t)\, \xi_r \xi_s \int_0^t f_{ij}(t)\psi_{ik}(t)\psi_{j\ell}(t)\, dt\right\}$$

$$= g_o \exp\left\{- \sum_{\substack{i,j,k,\ell=1 \\ r,s}}^{n} \phi_{kr}(t)\phi_{\ell s}(t)\xi_r \xi_s \int_0^t f_{ij}(t)\psi_{ik}(t)\psi_{j\ell}(t)\, dt\right\} \qquad (4.37)$$

The general solution of (4.27) is of the form

$$\Gamma(\xi_{o1}, \xi_{o2}, \ldots, \xi_{on}, g_o) = 0 \tag{4.38}$$

where Γ is an arbitrary function. Making use of (4.35), (4.37) and (4.38) one can solve for g

$$g(\xi_1, \ldots \xi_n, t) = F\left[\sum_{j=1}^{n} \phi_{1j}(t)\xi_j, \ldots, \sum_{j=1}^{n} \phi_{nj}\xi_j, t)\right]$$

$$\cdot \exp\left\{-\sum_{\substack{i,j,k \\ \ell,r,s=1}}^{n} \phi_{kr}(t)\phi_{\ell s}(t)\xi_r\xi_s \int_0^t f_{ij}(t)\psi_{ik}(t)\psi_{j\ell}(t)dt\right\} \tag{4.39}$$

where F is an arbitrary function. Let us define

$$K_{rs}(t) \triangleq \sum_{\substack{i,j \\ k,\ell=1}}^{n} \phi_{kr}(t)\phi_{\ell s}(t) \int_0^t f_{ij}(t)\psi_{ik}(t)\psi_{j\ell}(t)\, dt \tag{4.40}$$

then equation (4.39) reduces to

$$g(\xi_1, \ldots, \xi_n, t) = F\left[\sum_{j=1}^{n} \phi_{1j}(t)\xi_j, \ldots, \sum_{j=1}^{n} \phi_{nj}(t)\xi_j, t\right] \exp -\left[\sum_{r,s=1}^{n} K_{rs}(t)\xi_r\xi_s\right] . \tag{4.41}$$

To find the explicit expression of F one uses the initial condition (4.23). From (4.25) we find at t = 0

$$g(\xi_1,\ldots,\xi_n,0) = \exp\left[-i\sum_{k=1}^{n}\xi_k Z_{ok}\right] . \qquad (4.42)$$

Also at $t = 0$, (4.41) reduces to

$$g(\xi_1,\ldots,\xi_n,0) = F(\xi_1,\ldots,\xi_n,0) . \qquad (4.43)$$

Upon comparison between (4.42) and (4.43) it is evident that the arbitrary function F must be of the following form:

$$F(\xi_1,\ldots,\xi_n,0) = \exp\left[-i\sum_{k=1}^{n}\xi_K Z_{oK}\right]$$

and at any time $t > 0$ one finds

$$F\left[\sum_{j=1}^{n}\phi_{ij}(t)\xi_j,\ldots,\sum_{j=1}^{n}\phi_{nj}(t)\xi_j,t\right] = \exp\left\{-i\sum_{k=1}^{n}\left[\sum_{j=1}^{n}\phi_{kj}(t)\xi_j\right]Z_{oK}\right\} \qquad (4.44)$$

From (4.41) the fundamental solution of (4.27) becomes

$$g(\xi_1,\ldots,\xi_n,t) = \exp\left\{-i\sum_{K,j=1}^{n}\phi_{kj}(t)\xi_j Z_{ok} - \sum_{r,s=1}^{n}K_{rs}(t)\xi_r\xi_s\right\}. \qquad (4.45)$$

Our next step is to verify that (4.45) is actually the solution of (4.27). At $t = 0$

$$g(\xi_1,..,\xi_n,0) = \exp\left\{-i \sum_{k,j=1}^{n} \phi_{kj}(0)\ \xi_j\ Z_{ok}\right\} .$$

From the property of the fundamental matrix $\Phi(0) = I$, we find that its element $\phi_{ij}(0)$ must be

$$\phi_{ij}(0) = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases} . \qquad (4.46)$$

Hence

$$g(\xi_1,\ldots,\xi_n,0) = \exp\left\{-i \sum_{k=1}^{n} \xi_k\ Z_{ok}\right\} \qquad (4.47)$$

which is simply the initial condition expressed in (4.42)

Taking the partial derivative of g with respect to t we find'

$$\frac{\partial g}{\partial t}(\xi_1,\ldots,\xi_n,t) = g(\xi_1,\ldots,\xi_n,t)\left\{-i \sum_{k,j=1}^{n} \dot{\phi}_{kj}\xi_j Z_{ok}\right.$$

$$\left. - \sum_{r,s=1}^{n} \dot{K}_{rs}(t)\xi_r\xi_s \right\} , \qquad (4.48)$$

where

$$\dot{K}_{rs}(t) = \sum_{\substack{i,j\\K,\ell=1}}^{n} \dot{\phi}_{Kr}\phi_{\ell s}\int_0^t f_{ij}\psi_{iK}\psi_{j\ell}\,dt + \sum_{\substack{i,j\\K,\ell=1}}^{n} \phi_{Kr}\dot{\phi}_{\ell s}\int_0^t f_{ij}\psi_{iK}\psi_{j\ell}\,dt$$

$$+ \sum_{\substack{i,j\\K,\ell=1}}^{n} \phi_{Kr}\,\phi_{\ell s} f_{ij}\psi_{iK}\psi_{j\ell} \quad . \qquad (4.49)$$

Also from((4.35) we have the relation

$$\sum_{j=1}^{n} \dot{\phi}_{ij}(t)\,\xi_j + \sum_{j=1}^{n} \phi_{ij}(t)\,\dot{\xi}_j = 0 \qquad (4.50)$$

$$i = 1,\ldots,n$$

Upon substitution of (4.49) and (4.50) into (4.48) we find

$$\frac{\partial g}{\partial t}(\xi_1,\ldots,\xi_n,t) = g(\xi_1,\ldots,\xi_n,t)\Big\{ i \sum_{k,j=1}^{n} \phi_{kj}\dot{\xi}_j Z_{ok}$$

$$- \sum_{\substack{i,j,k=1\\ \ell,r,s}}^{n} \dot{\phi}_{kr}\,\phi_{\ell s}\,\xi_r\xi_s \int_0^t f_{ij}\psi_{ik}\psi_{j\ell}\,dt$$

$$- \sum_{\substack{i,j,k=1\\ \ell,r,s}}^{n} \phi_{kr}\dot{\phi}_{\ell s}\xi_r\xi_s \int_0^t f_{ij}\psi_{ik}\psi_{j\ell}\,dt - \sum_{\substack{i,j,k=1\\ \ell,r,s}}^{n} \phi_{kr}\phi_{\ell s} f_{ij}\psi_{ik}\psi_{j\ell}\dot{\xi}_r\xi_s \Big\}$$

$$= g(\xi_1,\ldots,\xi_n,t)\Big\{ i \sum_{k,j=1}^{n} \phi_{kj}\,\dot{\xi}_j\,Z_{ok}$$

$$+ \sum_{\substack{i,j,k=1\\ \ell,r,s}}^{n} \phi_{kr}\,\phi_{\ell s}\,\dot{\xi}_r\,\xi_s \int_0^t f_{ij}\,\psi_{ik}\,\psi_{j\ell}\,dt$$

$$+ \sum_{\substack{i,j,k=1\\ \ell,r,s}}^{n} \phi_{kr}\,\phi_{\ell s}\,\xi_r\,\dot{\xi}_s \int_0^t f_{ij}\,\psi_{ik}\,\psi_{j\ell}\,dt$$

$$- \sum_{\substack{i,j,k=1\\ \ell,r,s}}^{n} \phi_{kr}\,\phi_{\ell s}\,f_{ij}\,\psi_{ik}\,\psi_{j\ell}\,\xi_r\,\xi_s \Big\} \quad . \qquad (4.51)$$

Making use of the property $\psi\Phi = I$ and the definition of $K_{rs}(t)$, equation (4.51) reduces to

$$\frac{\partial g}{\partial t}(\xi_1,\ldots,\xi_n,t) = g(\xi_1,\ldots,\xi_n,t)\Big\{ i \sum_{k,j=1}^{n} \phi_{kj}\,\dot{\xi}_j\, Z_{ok} + \sum_{r,s=1}^{n} K_{rs}(t)\,\dot{\xi}_r\,\xi_s + \sum_{r,s=1}^{n} K_{rs}(t)\,\xi_r\,\dot{\xi}_s - \sum_{i,j=1}^{n} f_{ij}\,\xi_i\,\xi_j \Big\} . \quad (4.52)$$

Taking the partial derivative of g in (4.45) with respect to ξ_j we obtain

$$\frac{\partial g}{\partial \xi_j}(\xi_1,\ldots,\xi_n,t) = g(\xi_1,\ldots,\xi_n,t)\Big\{ - i \sum_{k=1}^{n} \phi_{kj}\, Z_{ok} - \sum_{s=1}^{n} K_{js}(t)\,\xi_s - \sum_{r=1}^{n} K_{rj}(t)\,\xi_r \Big\} . \quad (4.53)$$

Let us use the expression of (4.53) to form the right hand side of (4.27)

$$\sum_{i,j=1}^{n} c_{ij}\xi_i \frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial \xi_j} - g(\xi_1,\ldots,\xi_n,t) \sum_{i,j=1}^{n} f_{ij}\xi_i\xi_j$$

$$= \sum_{j=1}^{n} \left(\sum_{i=1}^{n} c_{ij}\xi_i \right) \frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial \xi_j} - g(\xi_1,\ldots,\xi_n,t) \sum_{i,j=1}^{n} f_{ij}\xi_i\xi_j$$

$$= - \sum_{j=1}^{n} \dot{\xi}_j \frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial \xi_j} - g(\xi_1,\ldots,\xi_n,t) \sum_{i,j=1}^{n} f_{ij}\xi_i\xi_j$$

$$= g(\xi_1,\ldots,\xi_n,t) \left\{ i \sum_{k,j=1}^{n} \dot{\xi}_j \phi_{kj} Z_{ok} + \sum_{s,j=1}^{n} K_{js}(t)\xi_s\dot{\xi}_j + \sum_{r,j=1}^{n} K_{rj}(t)\xi_r\dot{\xi}_j \right.$$

$$\left. - \sum_{i,j=1}^{n} f_{ij}\xi_i\xi_j \right\}$$

$$= g(\xi_1,\ldots,\xi_n,t) \left\{ i \sum_{k,j=1}^{n} \varphi_{kj}\dot{\xi}_j Z_{ok} + \sum_{r,s=1}^{n} K_{rs}(t)\dot{\xi}_r\xi_s + \sum_{r,s=1}^{n} K_{rs}(t)\xi_r\dot{\xi}_s \right.$$

$$\left. - \sum_{i,j=1}^{n} f_{ij}\xi_i\xi_j \right\} . \quad (4.54)$$

Upon comparison with (4.52) we find that the right hand side of (4.54) is simply equal to $\frac{\partial g}{\partial t}(\xi_1,\ldots,\xi_n,t)$, therefore we finally obtain:

$$\frac{\partial g}{\partial t}(\xi_1,\ldots,\xi_n,t) = \sum_{i,j=1}^{n} c_{ij}(t)\xi_i \frac{\partial g(\xi_1,\ldots,\xi_n,t)}{\partial \xi_j} - g(\xi_1,\ldots,\xi_n,t) \sum_{i,j=1}^{n} f_{ij}\xi_i\xi_j \quad (4.55)$$

which is the original equation expressed in (4.27). Therefore, (4.45) is actually the fundamental solution of (4.27). Rearranging the terms and changing the notations of indexes (4.45) reduces to:

$$g(\xi_1,\ldots,\xi_n,t) = \exp\left\{-i \sum_{k,j=1}^{n} \phi_{kj}(t)\xi_j Z_{ok} - \sum_{r,s=1}^{n} K_{rs}(t)\ \xi_r\xi_s \right\} \quad (4.56)$$

or

$$g(\xi_1,\ldots,\xi_n,t) = \exp\left\{-i \sum_{j=1}^{n} \xi_j m_j(t) - \frac{1}{2} \sum_{r,s=1}^{n} \lambda_{rs}(t)\xi_r\xi_s \right\}, \quad (4.57)$$

where

$$m_j(t) \triangleq \sum_{k=1}^{n} \phi_{kj}(t)\ Z_{ok} \qquad j = 1,\ldots,n \quad (4.58)$$

and

$$\lambda_{rs}(t) \triangleq 2\, K_{rs}(t) \quad . \tag{4.59}$$

Equation (4.57) is simply the Fourier transform of an n-dimensional distribution[41] with means and variances expressed respectively by (4.58) and (4.59). The corresponding conditional probability density function is given by

$$P(0,Zo/t,z) = \frac{1}{(2\pi)^{n/2}|\lambda|^{1/2}} \exp\left[-\frac{1}{2|\lambda|} \sum_{i,j=1}^{n} |\lambda|_{ij}(Z_i-m_i)(Z_j-m_j)\right] \tag{4.60}$$

where $|\lambda|_{ij}$ is the cofactor of the element λ_{ij} in the determinant $|\lambda|$ of the covariance matrix

$$\lambda = \begin{bmatrix} \lambda_{11} & \lambda_{12} \cdots \lambda_{1n} \\ \lambda_{21} & \lambda_{22} \cdots \lambda_{2n} \\ \cdots & \cdots \\ \lambda_{n1} & \lambda_{n2} \cdots \lambda_{nn} \end{bmatrix} , \tag{4.61}$$

in which the element λ_{ij} is given by (4.59). Equation (4.6) can be written in the following form:

$$P(0\ Z_o/t,Z) = \frac{1}{(2\pi)^{n/2}|\lambda|^{1/2}} \exp\left\{-\frac{1}{2|\lambda|}\left[|\lambda|_{11}(Z_1-m_1)^2\right.\right.$$

$$+ \sum_{j=2}^{n} |\lambda|_{ij}(Z_j-m_j)(Z_i-m_i) + \sum_{j=2}^{n} |\lambda|_{j1}(Z_1-m_1)(Z_j-m_j)$$

$$\left.\left. + \sum_{i,j=2}^{n} |\lambda|_{ij}(Z_i-m_i)(Z_j-m_j)\right]\right\} . \quad (4.62)$$

It is readily seen that as $Z_1 \to \pm \infty$, the probability distribution $P(0,Z_o/t,Z)$ also tends to zero provided that the covariance matrix λ is positive definite for all $t > 0$. Same argument holds true for all Z_j, $j=1,\ldots,n$.

If the initial position of the system (4.10) or (4.11) is known exactly then the probability density function $P(0,Z_o)$ can be written as

$$P(0,Z_o) = \delta(Z_o-a_o) = \prod_{i=1}^{n} \delta(Z_{oi}-a_{oi}) , \quad (4.63)$$

and the probability density function $P(t,Z)$ will be the same as equation (4.69) with means and variances given by (4.58) and (4.59) with Z_{ok} in (4.58) replaced by a_{ok}.

However, if only the inital distribution $P(0,Z_o)$ is known, then the probability density function $P(t,Z)$ is given in the form of an integral

$$P(t,Z) = \int P(0,Z_o;t,Z)\, dZ_o = \int P(0,Z_o)\, P(0,Z_o/t,Z)\, dZ_o \quad , \quad (4.64)$$

where $P(0,Z_o/t,Z)$ is given by (4.60) and the integration is understood to be carried over the whole space R_n. The average value of Z_i of the system (4.10) is simply

$$E[Z_i(t)] = \int Z_i(t)\, P(t,Z)\, dZ$$

$$= \int_{-\infty}^{+\infty} \cdots \int Z_i(t)\, P(t,Z)\, dZ_1, \ldots, dZ_n \quad . \qquad (4.65)$$

4.6 Linear Time-Varying System with Disturbances Generated by Differential Equations Excited by White Noise

In previous sections we assumed that the disturbances $n(t)$ of the linear time-varying system were caused by white noise; namely, random processes with constant spectral densities. A white-noise random process can be considered only as a limiting case which will never occur in practice. The most common examples of power spectral density, especially in noise applications to electric circuits, are those which can be considered as the spectral density of the output from a lumped-parameter linear circuit excited by white noise. If the linear circuit is also time-invariant then the spectral density is a rational function.

For the above reason the natural generalization is to assume that the random noise vector $n(t) = (n_1(t), \ldots, n_n(t))$ in the optimal control system is governed by the following set of differential equations.

$$\frac{d^{\alpha_i} n_i}{dt^{\alpha_i}} + \beta_{i(\alpha_i-1)}(t) \frac{d^{\alpha_i-1} n_i}{dt^{\alpha_i-1}} + \ldots + \beta_{il}(t) \frac{dn_i}{dt} + \beta_{io}(t)\, n_i = w_i \qquad i = 1,\ldots,n \qquad (4.66)$$

where w_i is the ith component of the white noise vector $w(t) = (w_1(t),\ldots,w_n(t))$ and α_i indicates the order of the differential equation for each n_i.

For each i, equation (4.66) can be transformed into a set of first order linear differential equations.

Let $n_i = Z_{n+\alpha_1+\ldots+\alpha_{i-1}+1}$, equation (4.66) can be written into the following form:

$$\frac{dZ_{n+\alpha_1+\ldots+\alpha_{i-1}+1}}{dt} = Z_{n+\alpha_1+\ldots+\alpha_{i-1}+2}$$

$$\frac{dZ_{n+\alpha_1+\ldots+\alpha_{i-1}+2}}{dt} = Z_{n+\alpha_1+\ldots+\alpha_{i-1}+3}$$

$$\frac{dZ_{n+\alpha_1+\ldots+\alpha_i}}{dt} = -\beta_{i(\alpha_i-1)}(t) Z_{n+\alpha_1+\ldots+\alpha_i} - \ldots - \beta_{io}(t) Z_{n+\alpha_1+\ldots+\alpha_{i-1}+1} + w_i \qquad i = 1,\ldots,n \qquad (4.67)$$

Adding equations (4.67) to the linear system (4.10)

$$\frac{dZ_i}{dt} = \sum_{k=1}^{n} c_{ik}(t)\, Z_k + \sum_{k=1}^{n} d_{ik}(t)\, n_k(t) \qquad i = 1,\ldots,n \qquad (4.10)$$

one finds a set of linear first order differential equations in the expanded state space of dimension $N \triangleq n+\alpha_1+\ldots+\alpha_n$ excited by white noise. In terms of components we have

$$\frac{dZ_1}{dt}=c_{11}(t)Z_1+\ldots+c_{1n}(t)Z_n+d_{11}(t)Z_{n+1}+d_{12}(t)Z_{n+\alpha_1+1}+\ldots+d_{1n}(t)Z_{n+\alpha_1+\ldots+\alpha_{n-1}+1}$$

. . .

$$\frac{dZ_n}{dt}=c_{n1}(t)Z_1+\ldots+c_{nn}(t)Z_n+d_{n1}(t)Z_{n+1}+d_{n2}(t)Z_{n+\alpha_1+1}+\ldots+d_{nn}(t)Z_{n+\alpha_1+\ldots+\alpha_{n-1}+1}$$

$$\frac{dZ_{n+1}}{dt}=Z_{n+2}$$

. . .

$$\frac{dZ_{n+\alpha_1}}{dt}=-\beta_{1(\alpha_1-1)}(t)Z_{n+\alpha_1}-\ldots-\beta_{10}(t)Z_{n+1}+F_{n+\alpha_1}(t)$$

$$\frac{dZ_{n+\alpha_1+1}}{dt}=Z_{n+\alpha_1+2}$$

. . .

$$\frac{dZ_{n+\alpha_1+\alpha_2}}{dt}=-\beta_{2(\alpha_2-1)}Z_{n+\alpha_1+\alpha_2}-\ldots-\beta_{20}(t)Z_{n+\alpha_1+1}+F_{n+\alpha_1+\alpha_2}(t)$$

. . .

$$\frac{dZ_{n+\alpha_1+\ldots+\alpha_{n-1}+1}}{dt}=Z_{n+\alpha_1+\ldots+\alpha_{n-1}+2}$$

. . .

$$\frac{dZ_{n+\alpha_1+\ldots+\alpha_n}}{dt}=-\beta_{n(\alpha_{n-1})}(t)Z_{n+\alpha_1+\ldots+\alpha_n}-\ldots-\beta_{no}(t)Z_{n+\alpha_1+\ldots+\alpha_{n-1}+1}$$

$$+F(t)_{n+\alpha_1+\ldots+\alpha_n} \qquad (4.68)$$

Writing equation (4.68) in matrix form we obtain

$$\frac{dZ}{dt} = E(t)\, Z + F(t) \quad , \tag{4.69}$$

where

$$E(t) = \begin{bmatrix} c_{11}(t)\cdots c_{1n}(t) & d_{11}(t)0\cdots0 & d_{12}(t)0\cdots0 & \cdots & \cdots & d_{1n}(t)0\cdots0 \\ \vdots & \vdots & \vdots & & & \vdots \\ c_{n1}(t)\cdots c_{nn}(t) & d_{n1}(t)0\cdots0 & d_{n2}(t)0\cdots0 & \cdots & \cdots & d_{nn}(t)0\cdots0 \\ \hline 0 & \begin{matrix} 0 & 1 & 0\cdots0 \\ 0 & 0 & 1\cdots0 \\ \vdots & & \\ 0 & 0\cdots0 & 1 \\ -\beta_{10}(t) & -\beta_{11}(t) & \cdots -\beta_{1(\alpha_1-1)}(t) \end{matrix} & 0 & \cdots & \cdots & 0 \\ \hline 0 & 0 & \begin{matrix} 0 & 1 & 0\cdots0 \\ 0 & 0 & 1\cdots0 \\ \vdots & & \\ 0 & 0\cdots0 & 1 \\ -\beta_{20}(t) & \cdots & -\beta_{2(\alpha_2-1)}(t) \end{matrix} & \cdots & \cdots & 0 \\ \hline \vdots & \vdots & \vdots & \cdots & \cdots & \vdots \\ \hline 0 & 0 & \cdots & \cdots & 0 & \begin{matrix} 0 & 1 & 0\cdots0 \\ 0 & 0 & 1\cdots0 \\ \vdots & & \\ 0 & 0\cdots0 & 1 \\ -\beta_{n0}(t) & \cdots & -\beta_{n(\alpha_n-1)}(t) \end{matrix} \end{bmatrix} \tag{4.70}$$

and F(t) is a column matrix in which the non zero elements are given by

$$F_{n+\sum_{j=1}^{i} \alpha_j} = w_i(t) \qquad i = 1,\ldots,n \tag{4.71}$$

It can be seen that equation (4.69) is a linear system excited by white noise and is in essence the same as equation (4.11). The development in the previous section therefore applies to this section as well.

Chapter V

DETERMINATION OF AVERAGE CHANGES IN TRAJECTORY AND COST OF OPTIMAL CONTROL SYSTEMS DUE TO RANDOM DISTURBANCES

The results developed in Chapter 4 are applied for the determination of average changes in trajectory and cost of optimal control systems due to random disturbances in plant parameters or noisy observation of state variables.

5.1 Systems Subject to White Noise Disturbances in Plant Parameters

In chapters 2 and 3 the analysis of variations of trajectory and cost of optimal control systems due to deterministic changes of plant parameters was discussed in detail. In this chapter we shall assume that the plant parameters about their nominal values may vary in a random manner. However, the disturbances in the system are again assumed to be small, so that the solution of the equation of first variation of the system can be considered as an accurate determination of the deviations from the optimal trajectory of the system. The deviations from the optimal trajectory in such a situation, which is random in nature and is not amenable to exact a priori specification, can only be evaluated by a method of statistical reasoning. In the following we shall discuss the open and closed-loop systems separately.

a) Open-loop control systems

The behavior of the plant of an open-loop optimal control system is again assumed to be governed by equation (2.2) which is repeated here for convenience

$$\frac{dx}{dt} = f(x,\alpha,u,t)$$
$$x(0) = x_0 \qquad (5.1)$$

The cost to be minimized is again given by the functional

$$J = \int_0^T f_0(x,u)dt \tag{5.2}$$

If the variations about the nominal plant parameter values are small we can represent the changes along the optimal trajectory by the equation of first variation

$$\dot{Z} = F Z + G\gamma ,$$

$$Z(0) = Z_0 , \tag{5.3}$$

where Z is the variation along the optimal trajectory and F, G are matrices expressed in (2.10). The vector γ represents the variation about the nominal parameter vector and is considered to be a white noise vector satisfying the assumptions stated in Chapter 4.

Using the results developed in Chapter 4, the average change along the optimal trajectory is given by

$$E[Z_i(t)] = \int Z_i(t)\, P(t,Z)dZ \quad i = 1,\ldots,n \tag{5.4}$$

where P(t,Z) is expressed in (4.74) and the integral is carried out over R_n.

From equation (2.14) the average change in cost is expressed by

$$E(\Delta J) = \int_0^T \sum_{j=1}^{n} \frac{\partial f_0}{\partial x_j} \; E[Z_j(t)]dt \tag{5.5}$$

b) <u>Closed-loop control systems</u>

Just as in the open-loop case, for small parameters variation

about the nominal vector, the equation of first variation is given by (3.7) which is repeated here for convenience

$$\dot{Z} = A(t)\, Z + B(t)\, \gamma(t) \qquad (5.6)$$

$$Z(0) = Z_0$$

with A(t) and B(t) expressed in (3.8). If γ (t) is again a white noise vector, we find

$$E[Z_i(t)] = \int Z_i(t)\, P(t,Z)\, dZ \qquad i = 1,\ldots,n \qquad (5.7)$$

In the same manner from equation (3.11) we obtain the average change in cost

$$E(\Delta J) = \int_0^T \sum_{j=1}^{n} \left(\frac{\partial f_o}{\partial x_j} + \sum_{k=1}^{r} \frac{\partial f_o}{\partial u_k} \frac{\partial u_k}{\partial x_j} \right) E[Z_i(t)]\, dt$$

5.2 Systems Subject to Disturbances Generated by Differential Equations Excited by White Noise in Plant Parameters

As discussed in section 4.5, if the parameter variations $\gamma(t)$ in equations (5.3) and (5.6) are generated by differential equations excited by white noise, we can expand the state space of the original system into another state space of higher dimension. By doing so we will arrive at another linear time-varying system excited by white noise. The average change along the optimal trajectory is again given by equations (5.4) and (5.7) and the average change in cost is given by (5.5) and (5.8), respectively, for the open- and closed-loop control systems.

5.3 Estimation of Changes in Trajectory and Cost due to Noisy Observation of State Variables

An ideal closed-loop control system was represented schematically in Fig. 3.1 with the behavior of the plant described by equations (3.1) or (3.2).

However, in reality, all the state variables are not available for direct observation, and therefore an estimation of those state variables must be made. As a result of the estimation errors, that is, errors due to estimation, a real physical system can be described more closely by the schematic diagram of Fig. 5.1 in which $n(t) = (n_1(t),\ldots,n_n(t))$ is an estimation error which can be

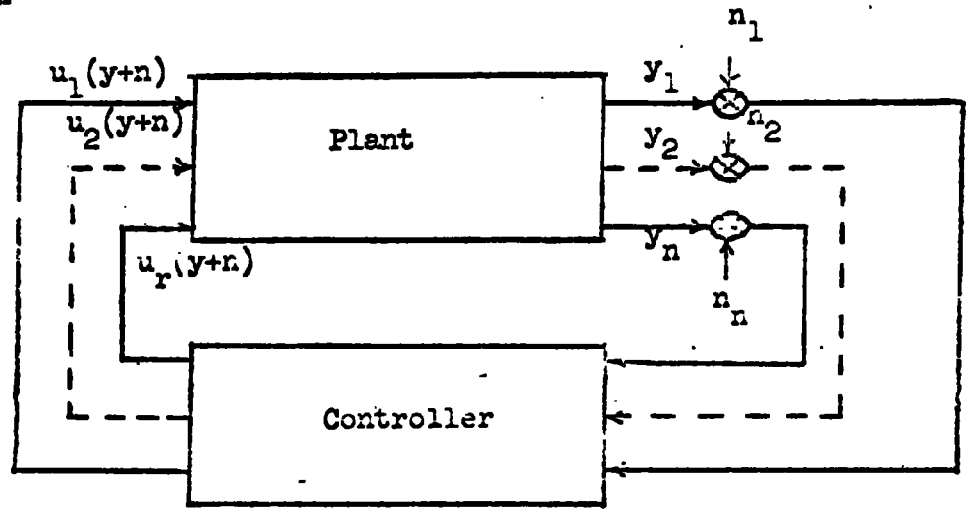


Fig. 5.1 Closed-Loop System with Random Disturbances

considered as a random noise vector[58]. The corresponding equations governing the motion of the system of Fig. 5.1 become

$$\frac{dy}{dt} = f(y,u) \qquad (5.9)$$

$$y(0) = y_o$$

and

$$u(t) = u(y + n) \qquad (5.10)$$

In order to estimate the deviations along the optimal trajectory due to estimation error, we expand the right hand side of equation (5.9) about the point (x,u(x)) with the second and higher order terms neglected. Taking the difference between this equation and equation (3.2) gives the equation of first variation of the system.

$$\frac{d(y-x)}{dt} = \left(\frac{\partial f(x,u(x))}{\partial x} + \frac{\partial f(x,u(x))}{\partial u}\ \frac{\partial u}{\partial x}\right)(y-x) + \frac{\partial f(x,u(x))}{\partial u}\ \frac{\partial u(x)}{\partial x}\ n. \tag{5.11}$$

where

$$\frac{\partial f}{\partial x} = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial x_1} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix} \qquad \frac{\partial f}{\partial u} = \begin{bmatrix} \frac{\partial f_1}{\partial u_1} & \cdots & \frac{\partial f_1}{\partial u_r} \\ \cdots & \cdots & \cdots \\ \frac{\partial f_n}{\partial u_1} & \cdots & \frac{\partial f_n}{\partial u_r} \end{bmatrix} \tag{5.12}$$

$$\frac{\partial u}{\partial x} = \begin{bmatrix} \frac{\partial u_1}{\partial x_1} & \cdots & \frac{\partial u_1}{\partial x_n} \\ \cdots & \cdots & \cdots \\ \frac{\partial u_r}{\partial x_1} & \cdots & \frac{\partial u_r}{\partial x_n} \end{bmatrix}$$

The partial derivatives are evaluated along the optimal trajectory. If we write $Z = y - x$, the variation along the optimal path, then equation (5.11) reduces to

$$\frac{dZ}{dt} = C(t)\,Z + D(t)n$$

$$Z(0) = Z_0 \qquad (5.13)$$

where

$$C(t) = -\frac{\partial f}{\partial x} + \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} \qquad (5.14)$$

and

$$D(t) = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x}$$

If n, the estimation error, has the property of a white noise vector, the method developed in section (4.5) can be used to obtain the average change along the optimal trajectory and the average change in cost. If n is generated by a set of linear time-varying differential equations excited by white noise, the results of section (4.6) can be applied.

5.4 Applications of the Above Method for Two Special Classes of Optimal Control Systems

In this section we shall apply the method developed above to two special classes of control systems, in which the closed-loop optimal control laws are functions of state variables.

a) All linear systems of the form

$$\frac{dx}{dt} = Bx + Mu$$

$$x(0) = x_0\,,\ x(\infty) = 0 \qquad (5.15)$$

$$u(0) = u_0\,,\ u(\infty) = 0$$

where

B = n x n constant matrix

M = n x 1 constant column vector

The cost to be minimized is given by the functional

$$J = \int_0^\infty V dt = \int_0^\infty \left(\sum_{k=1}^{n} a_k x_k^2 + Cu^2 \right) dt \tag{5.16}$$

where $a_k (k=1,\ldots,n)$ and C are constants.

Letov[47] has shown that the control law minimizing (5.16) is given by

$$u(t) = \sum_{i=1}^{n} p_i x_i \quad , \quad p_i = \text{constant} \tag{5.17}$$

provided that certain conditions are satisfied (see Appendix D). Substituting (5.17) into (5.15) we find that the optimal trajectory will satisfy the following equation:

$$\dot{x} = Bx + Mp^T x = (B + Mp^T)x \tag{5.18}$$

where p^T is the transpose of the constant vector $p = (p_1,\ldots,p_n)$. However, if there is estimation error in the system the behavior of the plant is governed by

$$\begin{aligned} \dot{y} &= By + Mp^T(y+n) \\ y(0) &= y_0 \end{aligned} \tag{5.19}$$

where n is the equivalent noise vector of the estimation error.

Since the system is linear, the exact change along the optimal trajectory x(t) can be obtained by subtracting (5.18) from (5.19). Let Z = y-x then Z satisfies the equation:

$$\dot{Z} = (B + Mp^T)Z + Mp^T n$$
$$Z(0) = Z_0 \qquad (5.20)$$

Equation (5.20) is in the same form as equation (4.22), hence the methods developed in sections (4.4) and (4.5) can be used to determine the average change along the optimal path once the probability probability distribution of the noise is specified.

To estimate the average change in cost we first write the cost functional corresponding to the actual trajectory.

$$J = \int_0^\infty \sum_{k=1}^{n} a_k(Z_k+x_k)^2 + C\left[\sum_{k=1}^{n} P_k(Z_k+x_k+n_k)\right]^2 \qquad (5.21)$$

The change in cost is given by the following linear approximation

$$\Delta J = 2 \left[\left(\sum_{k=1}^{n} a_k x_k Z_k + C \sum_{k=1}^{n} P_k x_k (Z_k+n_k)\right] dt \qquad (5.22)$$

The expected value of ΔJ becomes

$$E\left\{\Delta J\right\} = 2 \int \sum_{k=1}^{n} (a_k+Cp_k)x_k E\left\{Z_k\right\} dt$$

where $E\left\{n_k\right\} = 0$ by assumption.

b) Norm-invariant systems

Consider a nonlinear system of the form

$$\begin{aligned} \dot{x}_i(t) &= f_i[x_1(t),\ldots,x_n(t),t] + u_i(t) \\ x_i(0) &= \xi_i \end{aligned} \tag{5.23}$$

When written in vector form (5.23) becomes

$$\begin{aligned} \dot{x}(t) &= f(x(t),t) + u(t) \\ x(0) &= \xi \end{aligned} \tag{5.24}$$

The vector x(t) is the state of the system at time t and u(t) is the control vector at time t.

It is assumed that the following conditions are satisfied.

a) The control vector u(t) is restricted to lie within a sphere of radius M, or

$$\| u(t) \| \triangleq \left\{ \sum_{i=1}^{n} u_i^2 \right\}^{1/2} \leq M \text{ for all } t \geq 0 \tag{5.25}$$

b) The unforced part of the system (5.24)

$$\begin{aligned} \dot{x}(t) &= f(x(t),t) \\ x(0) &\neq 0 \end{aligned} \tag{5.26}$$

has the property

$$\| x(t) \| = \| x(0) \| \qquad \text{for all } t \geq 0 \tag{5.27}$$

that is $\frac{d}{dt} \|x(t)\| = \frac{[\dot{x}(t),x(t)]}{x(t)} = 0$. The quantity $[\dot{x},x]$ is the inner product of $x(t)$ and $\dot{x}(t)$. Equation (5.27) is equivalent to

$$[f(x(t),t),\ x(t)] = 0 \tag{5.28}$$

The cost functional to be minimized is given by

$$J = \int_0^T V\, dt \quad . \tag{5.29}$$

Depending on the form of the function V we have different types of minimization problems. Let us consider two specific cases.

Case A The Minimum Fuel Problem

In this case we wish to find the control $u(t)$ subject to the constraint (5.25), as a function of state $x(t)$, which drives the solution of (5.24) from any initial state ξ to 0 and which minimizes

$$J = \int_0^T k \|u(t)\|\, dt \ , \tag{5.30}$$

where k is a positive constant and T is not fixed in advance. The expression (5.30) represents the fuel consumed in some physical problems. It has been shown by Athans[39] (appendix E) that the optimal control which will drive the system (5.24) from any initial state ξ to 0 and which will minimize (5.30) is given by

$$u(t) = -M \frac{x(t)}{\|x(t)\|} \ . \tag{5.31}$$

x(t) is the optimal trajectory and is the solution of

$$\dot{x}(t) = f(x(t),t) - \frac{Mx(t)}{\|x(t)\|} \tag{5.32}$$

In the presence of noise the equation of motion becomes

$$\dot{y}(t) = f(y(t),t) - M\frac{y(t)+n(t)}{\|y(t)+n(t)\|} \tag{5.33}$$

From equations (5.32) and (5.33) the equation of first variation is given by

$$\dot{Z} = A(t)\,Z + B(t)n\ , \tag{5.34}$$

where the element of the matrices A(t) and B(t) are given by

$$a_{ij}(t) = \frac{\partial f_i}{\partial x_j}(x,t) + M\left(\frac{x_i x_j}{\|x\|^3} - \frac{x_i}{\|x\|}\right) \tag{5.35}$$

and

$$b_{ij}(t) = M\left(\frac{x_i x_j}{\|x\|^3} - \frac{x_i}{\|x\|}\right)\ , \tag{5.36}$$

Once the probability distribution of the estimation error vector is specified the average deviation along the optimal trajectory can be determined. In this case there is no change in fuel due to estimation error.

Case B The Minimum Energy Problem

This problem requires to find the control u(t) as a function of the state x(t), which will drive the solution of (5.24) from any initial state ξ to 0 and will minimize

$$J = \frac{1}{2}\int_0^T \|u(t)\|^2 \, dt = \frac{1}{2}\int_0^T [u(t),u(t)]dt \ , \quad (5.37)$$

where the terminal time is given in advance. (5.37) represents dissipated energy for some physical problems.

It can be shown (appendix F) that the control

$$u(t) = -\frac{\|\xi\|}{T}\frac{x(t)}{\|x(t)\|} \quad (5.38))$$

is the optimal control for the system and x(t) is the optimal trajector satisfying

$$\dot{x}(t) = f(x(t),t) - \frac{\|\xi\|}{T}\frac{x(t)}{\|x(t)\|} \quad (5.39)$$

If there is estimation error in the system then the equation of first variation is again given by (5.34) with the constant M replaced by $\frac{\|\xi\|}{T}$.

5.5 Stability Considerations

As mentioned in Chapter I, Bertram and Sarachik in the USA as well as Kats and Krasovskii in USSR have extended the second method of Lyapunov to investigate the stability, inastochastic sense, of circuits or systems described by ordinary differential equations with randomly time-varying parameters $\alpha(t)$ of the following form

$$\dot{x} = f[x,\alpha(t),t] \tag{5.40}$$

where $x(t)$ is an n-dimensional state vector and $\alpha(t)$ is a p-dimensional randomly time-varying parameter vector.

It is assumed that $x(t) = 0$ is an equilibrium solution, i.e.,

$$f(0,\alpha(t),t) = 0, \tag{5.41}$$

whose stability is to be examined. However, the concept of Lyapunov's stability cannot be applied directly to the investigation of stability of optimal control systems for the following reasons: Firstly, as pointed out by Lebedev and several other authors,[48] the stability of a system in the sense of Lyapunov, that is, motions (solutions) which are once near together remain near together for all future time as functions of time, can never be verified in practice, since a physical system can only be observed for a finite interval of time. This is especially true for optimal control systems since the terminal time is in general finite. Secondly, the variations along the optimal trajectory of a control system investigated in this dissertation are caused not only by initial disturbances but also by variations of the control function in the form of persistent random disturbances which cannot be treated by the method of Bertram and Sarachik men-

mentioned above for the following reason: As discussed in section 5.1 the system under consideration is given by

$$\dot{x} = f[x,\alpha(t), u(t),t] \tag{5.42}$$
$$x(0) = x_o \quad ,$$

where $\alpha(t)$ is a deterministic plant parameter vector. What we are concerned with is the variation along the optimal trajectory if the plant parameter $\alpha(t)$ is subject to random disturbances $n(t)$. The disturbed system can be represented by the following equation

$$\dot{y} = f[y, \alpha(t) + n(t), u(t),t] \tag{5.43}$$
$$y(\theta) = y_o \quad .$$

If the random disturbance $n(t)$ is small, then the variation along the optimal trajectory is approximately given by the equation of first variation

$$\dot{Z} = FZ + Gn \ , \tag{5.44}$$

where F and G are matrices expressed in (2.10). Perhaps it is worthwhile to emphasize again that the equation of first variation under consideration is of a completely different nature as the equation considered by Bertram and others. Therefore, when one speaks of stability of such a system in a finite time interval, the concept of bounded input variation and bounded output variation in stochastic sense should be used.

Some initial attempts to study the stability of motion for a finite time interval without stochastic disturbances have been made by various authors such as Lebedev[49] and Chzhan[50]. They discussed different types of stability over a finite interval of time such as stability

under finite initial disturbances. The ideas of the second method of Lyapunov are modified so that they can be applied to this type of problems.

As suggested by Landa[51] there is a possibility of using the Fokker-Planck equation to investigate the stability of automatic control systems in the presence of random disturbances. In order to investigate the stability of motion of optimal control systems in a finite time interval one can modify the definition of stability in (1.5) as follows: Let n be either the random parameter variation vector or estimation error vector in the estimation of state variables and Z be the variation vector along optimal trajectory of a given system. Then the system is said to be stable in the mean square sense if, for some $\delta > 0$,

$$\underset{t_o \leq t \leq T}{E} [\| n(t) \|^2] \leq \delta ,$$

the following condition is satisfied, that is,

$$\underset{t_o \leq t \leq T}{E} [\| Z \|^2] < M ,$$

where M may depend on the initial condition Z_o, on the finite time interval, on the value δ and on the characteristics of the system.* By means of the Fokker-Planck equation the analysis given in Chapter 4 enables one to determine the value of each compnnent of Z at every instant of time in a stochastic sense and therefore reveals the stability properties of the system.

---

*This concept of stability is a generalization of the concept of practical stability, as defined by LaSalle and Lefschetz[52], to stochastic systems in a finite time interval.

Chapter VI

EXPERIMENTAL VERIFICATIONS

In this chapter an optimal control system subject to stochastic disturbances is simulated on an Applied Dynamics analog computer, model AD-1-64PB. Experimental verifications of the theoretical analyses in the previous chapters are presented. The results of the experiment are in close agreement with the theoretical calculations.

6.1 Measurement of the power spectral density of the random noise generator

The spectral density N(f) of the Pace random noise generator output, model 201A, as a function of frequency can be obtained by measuring the variance of the noise output. If the noise with zero mean is fed into a narrow band-pass linear filter with transfer function Y(s), then according to Wiener-Kinchine theorem[53] the standard deviation σ of the output of the filter is given by

$$\sigma^2 = \int_0^\infty N(f)\,|Y(f)|^2 df \quad . \tag{6.1}$$

It is assumed that the power spectral density N(f) is flat over the narrow bandwidth of the filter. Taking N(f) to be a constant, then the value of the power spectral density at the resonant frequency f_o of the filter is given by

$$N(f_o) = \frac{\sigma^2}{\int_0^\infty |Y(f)|^2 df} \tag{6.2}$$

In the following measurement a filter of the transfer function

$$Y(s) = \frac{s/\omega_o}{(1 + s/\omega_o)^3} \tag{6.3}$$

is used. ω_o is the resonant frequency of the filter in radians per second.

Combining equations (6.2) and (6.3) one obtains

$$N(\omega_o) = \frac{\sigma^2}{f_o \int_0^\infty \frac{\left(\frac{\omega}{\omega_o}\right)^2}{\left[1 + \left(\frac{\omega}{\omega_o}\right)^2\right]^3} d\left(\frac{\omega}{\omega_o}\right)}$$

$$= \frac{\sigma^2}{\frac{f_o}{4}\left[\int_0^\infty \frac{d\left(\frac{\omega}{\omega_o}\right)}{\left[1 + \left(\frac{\omega}{\omega_o}\right)^2\right]^2}\right]}$$

$$= \frac{\sigma}{\frac{f_o}{8}\left[\int_0^\infty \frac{d\left(\frac{\omega}{\omega_o}\right)}{\left[1 + \left(\frac{\omega}{\omega_o}\right)^2\right]}\right]}$$

$$= \frac{\sigma^2}{\frac{f_o}{8} \times \frac{\pi}{2}} = \frac{\sigma^2}{\frac{\omega_o}{32}} = \frac{32\,\sigma^2}{\omega_o} \tag{6.4}$$

It can be seen from equation (6.4) that if the standard deviation of the output of the filter is known, the power spectral density of the output of the random noise generator can be calculated accordingly.

The technique of measuring the standard deviation of the output of the filter excited by random noise can be explained with the help of the following block diagram, Fig. 6.1.



Fig. 6.1 Block Diagram for Measurement of Standard Deviation

The output of the random noise generator $n_o(t)$ is fed into the narrow-band linear filter whose output, $x(t)$, is squared by an electronic multiplier. The output of the electronic multiplier is then integrated for a time interval T by an integrator and the output of the integrator is $T\,\sigma^2$ with σ given by

$$\sigma \approx \left[\frac{1}{T}\int_o^T x^2(t)\,dt\right]^{1/2} \tag{6.5}$$

The block diagram of Fig. 6.1 is simulated by an analog computer. Six measurements of σ^2 are made with resonant frequencies ω_o of the filter set at 10, 20, 40, 60, 80, and 100 radians per second. For the resonant frequencies of 10 and 20 radians per second the circuit of Fig. 6.2 is used.

In Fig. 6.2 the values of all resistors are in megohms and the values of the capacitors are in microfarads. For ω_o = 10 radians per second the resistor R is set at 0.1 megohms while for ω_o = 20 radians per second R is set at 0.05 megohms.

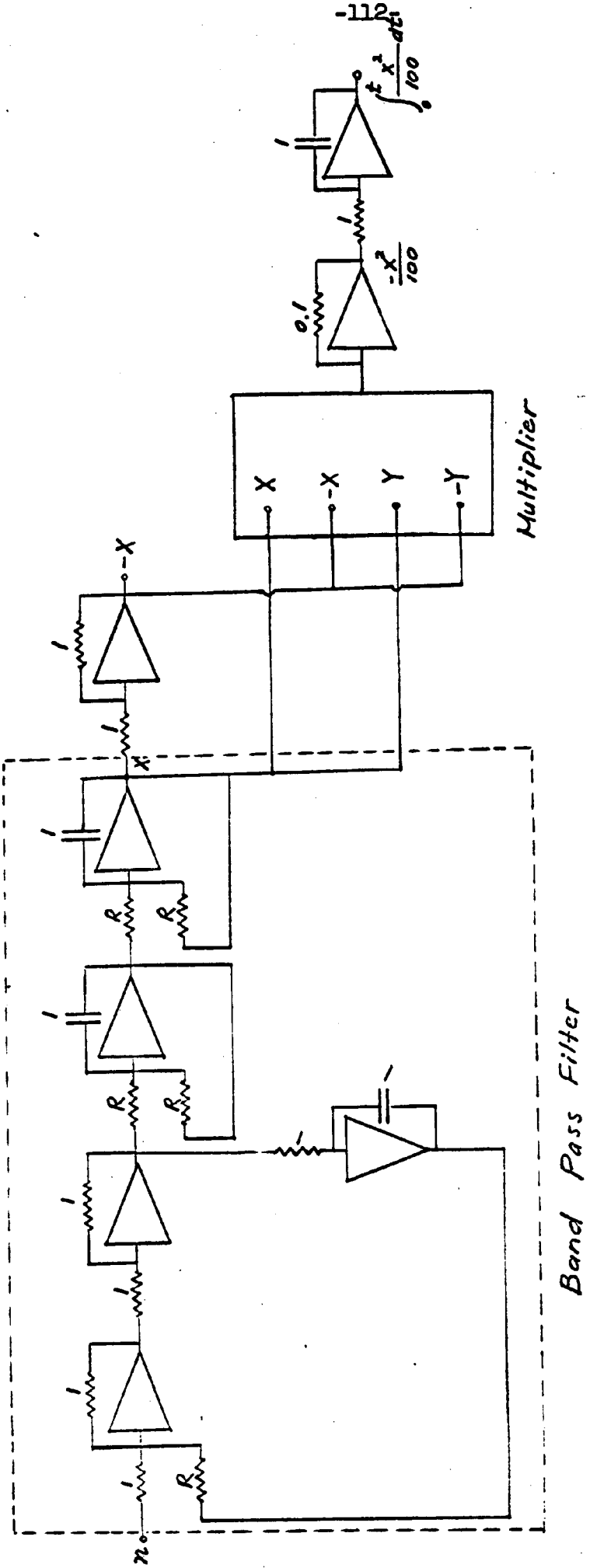


Fig. 6.2 Computer Setup for the Measurement of Power Spectral Density of the Output of Random Noise Generator at 10 and 20 Radians per Second.

For the higher resonant frequencies the circuit of Fig. 6.2 is modified as indicated in Fig. 6.3 due to the fact that the proper values of the resistors R are not available in the computer.

It can be seen that the transfer function of the filter in Fig. 6.3 is given by

$$Y(s) = \frac{\left(\frac{s}{100\alpha}\right)}{\left(1 + \frac{s}{100\alpha}\right)^3} \qquad (6.6)$$

If all the potentiometers are set at $\alpha = 0.4$, 0.6, 0.8 and 1.0 then the resonant frequencies of the filter become respectively 40, 60, 80 and 100 radians per second.

Using the circuits of Fig. 6.2 and Fig. 6.3 with the time interval for each measurement taken to be 30 minutes the approximate values of the variances of the filter and hence the power spectral density of the random noise source at various frequencies are tabulated in Table 6.1. The power spectrum of the noise source is given in Fig. 6.4

ω_0 rad/sec	10	20	40	60	80	100
σ^2(volts)^2	1.72	3.44	5.11	10.22	12.90	19.00
N(volts)^2/cps	5.52	5.33	4.18	5.45	5.15	6.07

Table 6.1 Measurement of Variance of Filter Output and Calculation of Power Spectrum of Noise Source

It can be seen from Fig. 6.4 that the power spectral density of the output of the random noise generator is reasonably flat as a

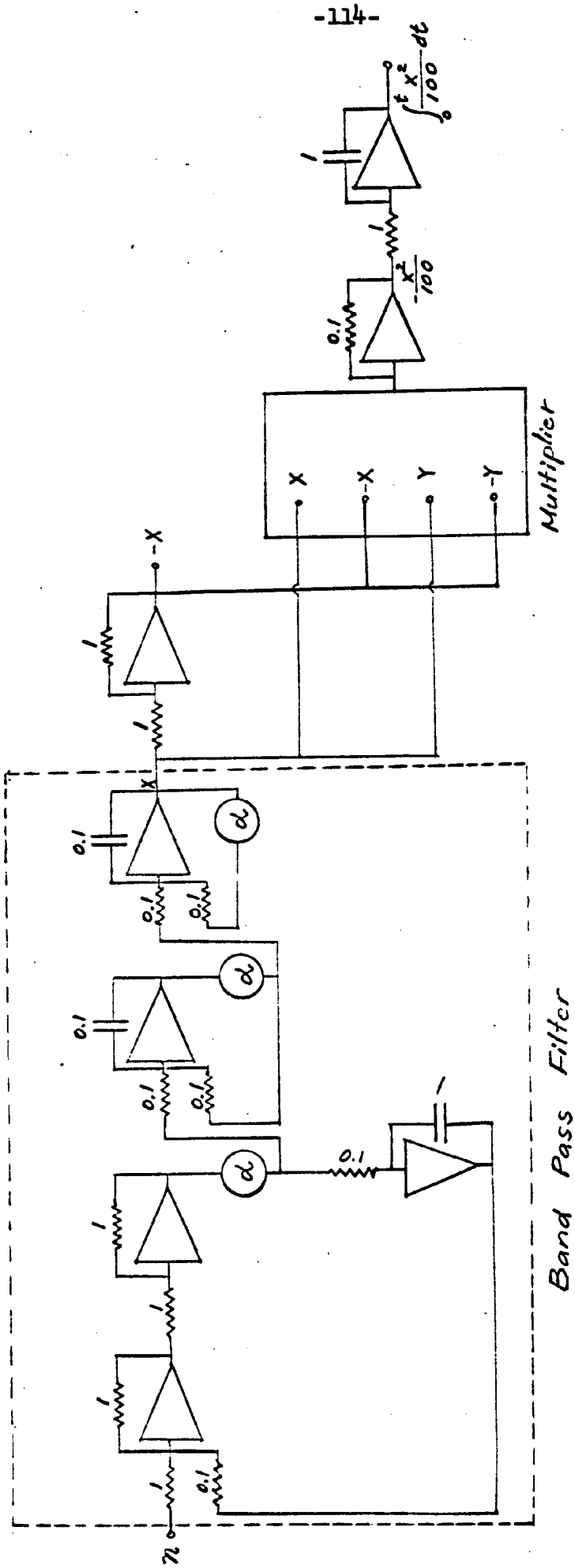


Fig. 6.3 Computer Setup for the Measurement of Power Spectral Density of the Output of Random Noise Generator at 40, 60, 80 and 100 Radians per Second.

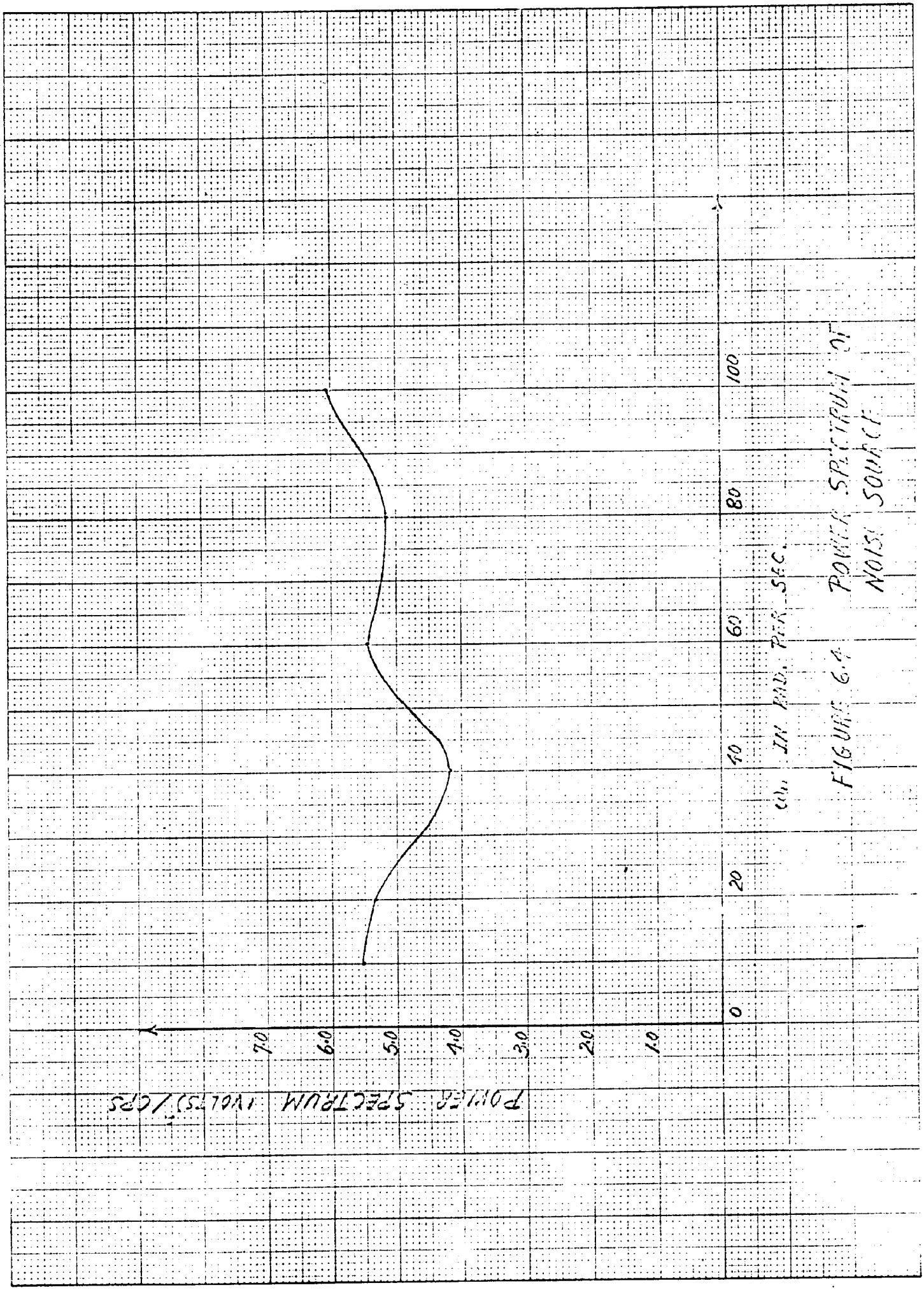
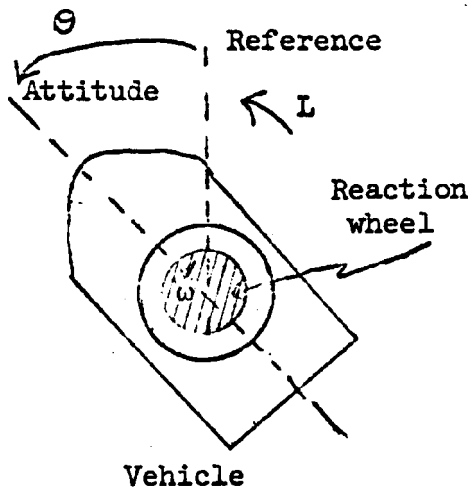


FIGURE 6.4 POWER SPECTRUM OF NOISE SOURCE

function of frequency with the average value found to be 5.28 $(\text{volts})^2/\text{cps}$.

6.2 An Optimal Control System

In this section we shall investigate an optimal control system which can be considered as a single axis model of vehicle with a reaction wheel as indicated in Fig. 6.5[54]. It is assumed that the control torque L is provided by a d.c., armature-controlled, permanent magnet servomotor and that the motor is driven by an ideal voltage source. The control problem is to transfer the vehicle from some fixed initial attitude to some fixed final attitude and the control V(t) will be chosen such that the energy delivered to the servomotor is minimum.



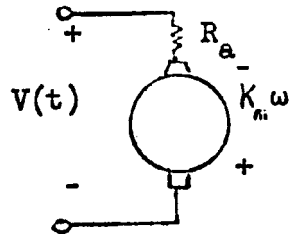


Fig. 6.5 Vehicle and Reaction Wheel with a D.C. Servomotor.

Let the vehicle attitude and the wheel velocity with respect to the vehicle be designated as θ and ω respectively, then the equations of motion can be written as

$$J\ddot{\theta} = L$$
$$I_w(\dot{\omega} + \ddot{\theta}) = -L, \qquad (6.7)$$

where J and I_w are moments of inertia of the vehicle and the wheel, respectively. Assuming that the vehicle and the wheel are initially nonrotating, equations (6.7) can be reduced to

$$I_w(\dot{\omega} + \ddot{\theta}) = - J\ddot{\theta}$$

$$\dot{\omega} = - \frac{I_w + J}{I_w} \ddot{\theta}$$

$$\omega = - \frac{I_w + J}{I_w} \dot{\theta}$$

$$\omega = - \frac{J}{I} \dot{\theta} \ , \tag{6.8}$$

where

$$I = \frac{J_w}{\frac{J_w}{J} + 1}$$

The torque L is given by

$$L = K_L i = K_L \left(\frac{V + K_m \omega}{R_a} \right) , \tag{6.9}$$

where K_L and K_m are respectively the torque and speed constants of the motor. Substituting (6.7) and (6.8) into (6.9) we find

$$J\ddot{\theta} = \frac{K_L V}{R_a} - \frac{K_L K_m J}{R_a I} \dot{\theta} \ . \tag{6.10}$$

Rearranging the terms in (6.10) gives

$$\tau_m^2 \ddot{\theta} + \tau_m \dot{\theta} = \frac{\tau_m}{K_m} \left(\frac{I}{J} \right) V \tag{6.11}$$

where $\tau_m = \frac{IR_a}{K_L K_m}$ is the motor time constant. For convenience, let us define

$$u(t) = V(t)$$

$$\tau = \frac{t}{\tau_m} \ , \text{ the dimensionless time} \tag{6.12}$$

and

$$x(t) = \frac{K_m \theta}{\tau_m \left(\frac{I}{J} \right)} \ .$$

By means of these definitions equation (6.11) reduces to the following form

$$\frac{d^2x}{d\tau^2} + \frac{dx}{d\tau} = u(\tau). \tag{6.13}$$

Let $x_1 = x$ and $x_2 = \frac{dx_1}{d\tau}$, then equation (6.13) can be written in the vector form

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_2 + u(\tau) \end{aligned} \tag{6.14}$$

with initial conditions, $x_1(0) = x_{10}$, $x_2(0) = 0$.

Equations (6.14) can be reduced further to decuupled form by a linear transformation[55]

$$\begin{aligned} e_1 &= x_1 + x_2 \\ e_2 &= - x_2 \end{aligned} \tag{6.15}$$

Then equations (6.14) can be expressed in terms of e_1 and e_2

$$\begin{aligned} \dot{e}_1 &= u(\tau) \\ \dot{e}_2 &= - e_2 - u(\tau) \end{aligned} \tag{6.16}$$

with the initial conditions $e_1(0) = x_{10}$, $e_2(0) = 0$.

The power delivered to the servomotor is given by

$$P = Vi = V \frac{(V + K_m\omega)}{R_a} \tag{6.17}$$

$$= \frac{u(\tau)}{R_a}(u(\tau) + K_m\omega) = \frac{u(\tau)}{R_a}(u(\tau) + \dot{e}_2(\tau)).$$

Then the control problem is to find $u(\tau)$ such that it transfers the initial state $(x_{10},0)$ to the final state $(0\ ,\ 0)$ and at the same time it minimizes the functional

$$J = \int_0^T u(u + e_2)d\tau \tag{6.18}$$

for a given transition time T.

It can be shown (Appendix G) that for a given time T the energy in (6.18) is minimized if the control function $u(\tau)$ assumes the following form

$$u(x_1,x_2) = x_2 + \frac{36}{T^3} x_{10} \left[\frac{2x_1 - x_{10}}{2x_2 - \frac{6}{T} x_{10}} \right] \tag{6.19}$$

Substituting equation (6.19) into (6.14) gives

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \frac{36}{T^3} x_{10} \left[\frac{2x_1 - x_{10}}{2x_2 - \frac{6}{T} x_{10}} \right] \end{aligned} \tag{6.20}$$

If we let

$$\begin{aligned} x_1(0) &= x_{10} = 20v \\ x_2(0) &= 0 \\ x_1(T) &= 0 \\ x_2(T) &= 0 \\ T &= 10 \ , \end{aligned}$$

then equations (6.20) are reduced to

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= 0.72 \left[\frac{x_1 - 10}{x_2 - 6} \right] \ , \end{aligned} \tag{6.22}$$

which can be simulated on an analog computer. The computer set up is indicated in Fig. 6.6, with resistors in megohms and capacitors in microfarads. The results are shown in Fig. 6.7.

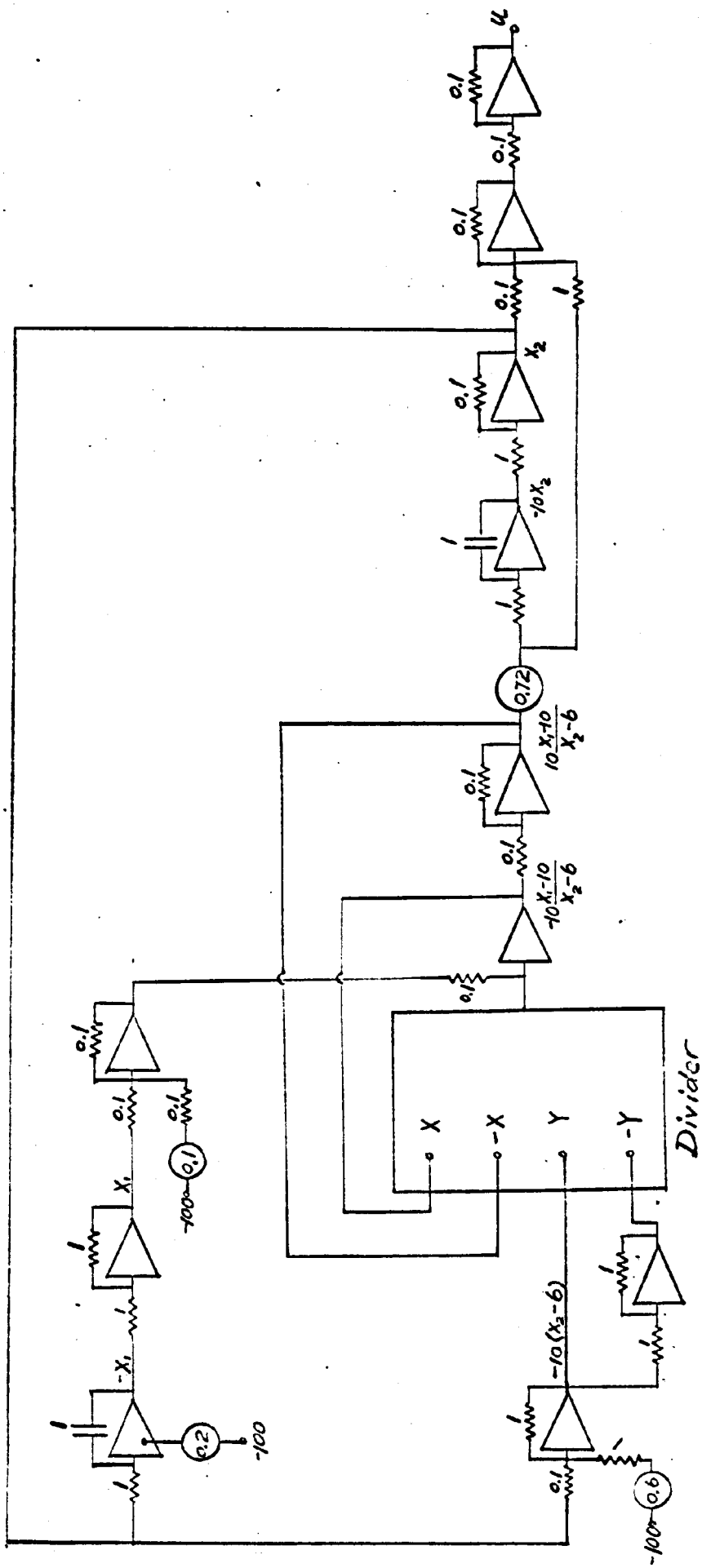


Fig. 6.6 Computer Setup of the System (6.22).

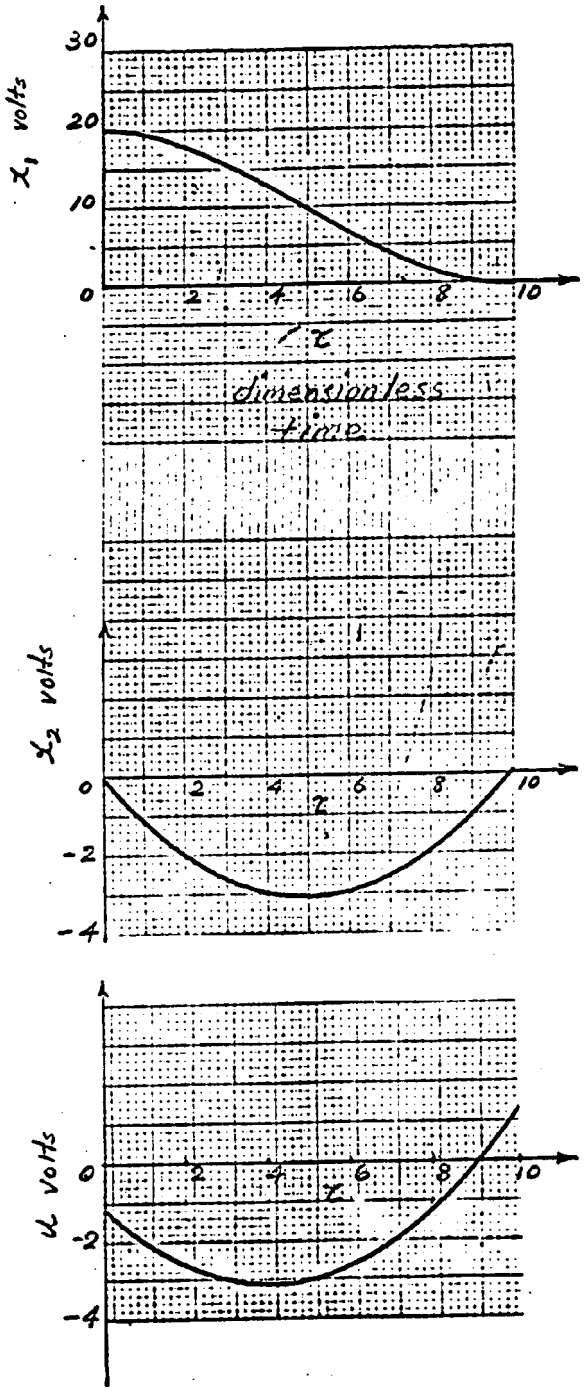


Fig. 6.7 Minimum Energy Control for D.C. Motor Reaction Wheel System.

6.3 Minimum Energy Control System Disturbed by Random Noise.

If there is random noise n_1 in the optimal control problem discussed in the previous section, the equations of motion will be

$$\dot{y}_1 = y_2$$

$$\dot{y}_2 = 0.72 \left[\frac{y_1 + n_1 - 10}{y_2 - 6} \right] , \tag{6.23}$$

where y's are the components of the actual trajectory of the system. Let us assume that the initial values of the system (6.23) are given by

$$y_1(0) = 18$$

$$y_2(0) = 0 ,$$

then by combining equations (6.20) and (6.23) the equations of first variation become

$$\dot{Z}_1 = Z_2$$

$$\dot{Z}_2 = 0.72 \left[\frac{1}{x_2 - 6} (Z_1 + n_1) - \frac{x_1 - 10}{(x_2 - 6)^2} Z_2 \right] \tag{6.24}$$

and

$$Z_1(0) = -2$$

$$Z_2(0) = 0 ,$$

where $Z_1 = y_1 - x_1$ and $Z_2 = y_2 - x_2$. Equations (6.24) can be written in the following vector form

$$\dot{Z} = C(t)\, Z + D(t)\, n , \tag{6.25}$$

where

$$C(t) = \begin{bmatrix} 0 & 1 \\ \dfrac{0.72}{x_2 - 6} & -0.72\ \dfrac{x_1 - 10}{(x_2 - 6)^2} \end{bmatrix}$$

$$D(t) = \begin{bmatrix} 0 & 0 \\ \dfrac{0.72}{x_2 - 6} & 0 \end{bmatrix}, \qquad (6.26)$$

and

$$n(t) = \begin{bmatrix} n_1(t) \\ 0 \end{bmatrix} .$$

The expected values of Z_1 and Z_2 as well as their variances of the linearized system (6.24) are given by equations (4.40), (4.58) and (4.59) which can be written as

$$m_j(t) = E[Z_j(t) = \phi_{1j}(t)\ Z_j(0) \qquad (6.27)$$

and

$$\lambda_{rs}(t) = 2 \left[\sum_{k,\ell=1}^{2} \phi_{kr}(t)\ \phi_{\ell s}(t) \int_0^t f_{22}(t)\ \psi_{2k}(t)\ \psi_{2\ell}(t)\ dt \right]$$

$$r,s = 1,2.$$

From equation (4.31), the elements $\psi_{ij}(t)$ satisfy the equation

$$\dot{\psi}(t) = - C'(t)\, \psi(t)$$
$$\psi(0) = I \ , \qquad (6.28)$$

where $C'(t)$ is the transpose of $C(t)$ expressed in (6.26). To generate the elements of ψ, the computer setup corresponding to the following equation

$$\dot{\beta} = - C'(t)\, \beta \qquad (6.29)$$

or

$$\begin{bmatrix} \dot{\beta}_1 \\ \dot{\beta}_2 \end{bmatrix} = \begin{bmatrix} 0 & - \dfrac{0.72}{x_2 - 6} \\ -1 & + 0.72 \dfrac{x_1 - 10}{(x_2 - 6)} \end{bmatrix} \begin{bmatrix} \beta_1 \\ \beta_2 \end{bmatrix}$$

is shown in Fig. 6.8. x_1 and x_2 are again the components of the optimal trajectory. The elements ψ_{11} and ψ_{21} are obtained by setting the initial conditions of β to be

$$\beta_1(0) = 1$$

and

$$\beta_2(0) = 0 \ , \qquad (6.30)$$

then

$$\beta_1(t) = \psi_{11}(t)$$

and

$$\beta_2(t) = \psi_{21}(t) \ . \qquad (6.31)$$

The elements ψ_{12} and ψ_{22} are obtained by setting the initial conditions of (6.29) equal to

$$\beta_1(0) = 0$$

and

$$\beta_2(0) = 1 \ ,$$

then

$$\beta_1(t) = \psi_{12}(t)$$

and

$$\beta_2(t) = \psi_{22}(t)$$

The functions ψ_{ik} obtained from the experiment are shown in Fig. 6.9.

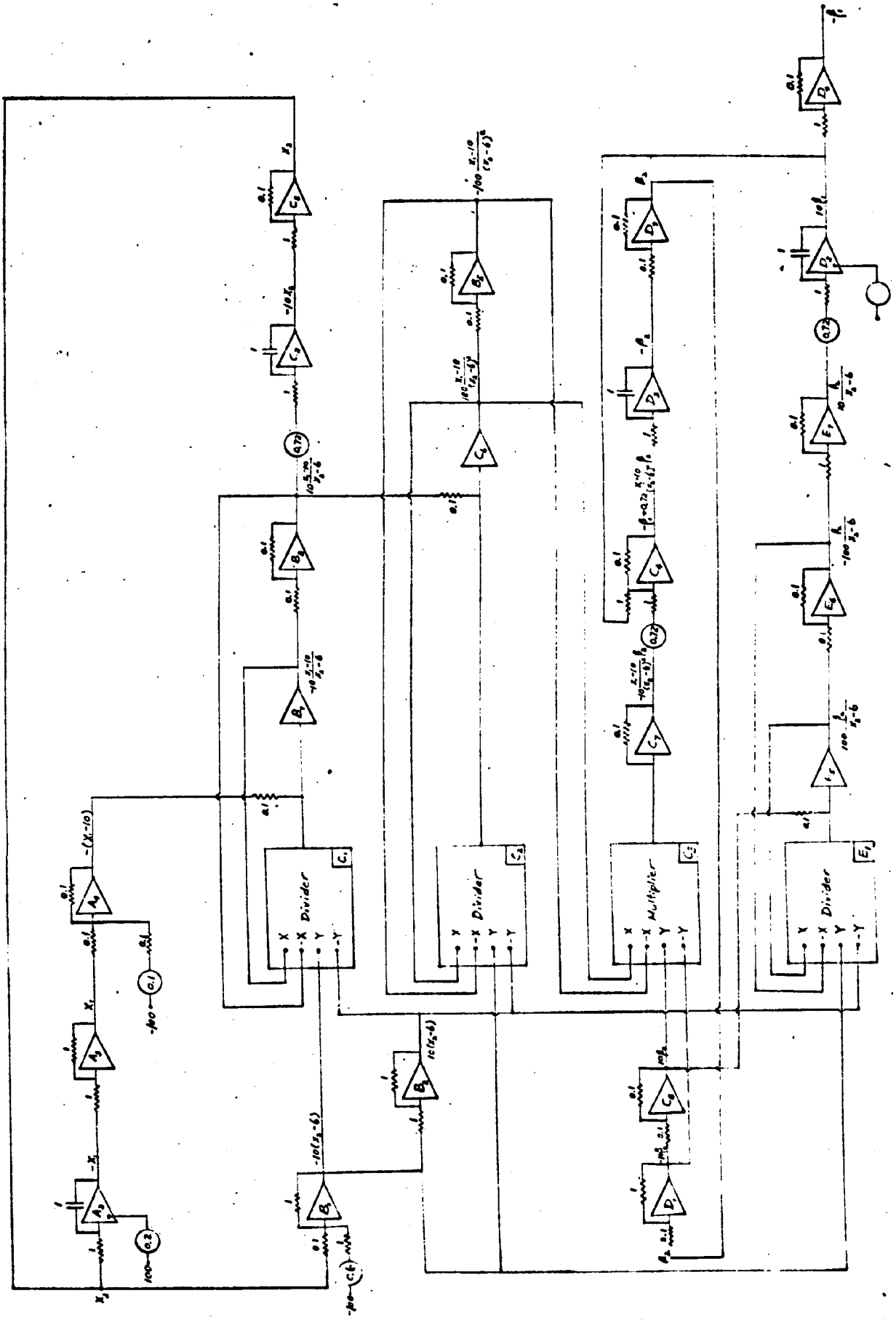


Fig. 6.8 Computer Setup for the Simulation of the System (6.29) for the Matrix

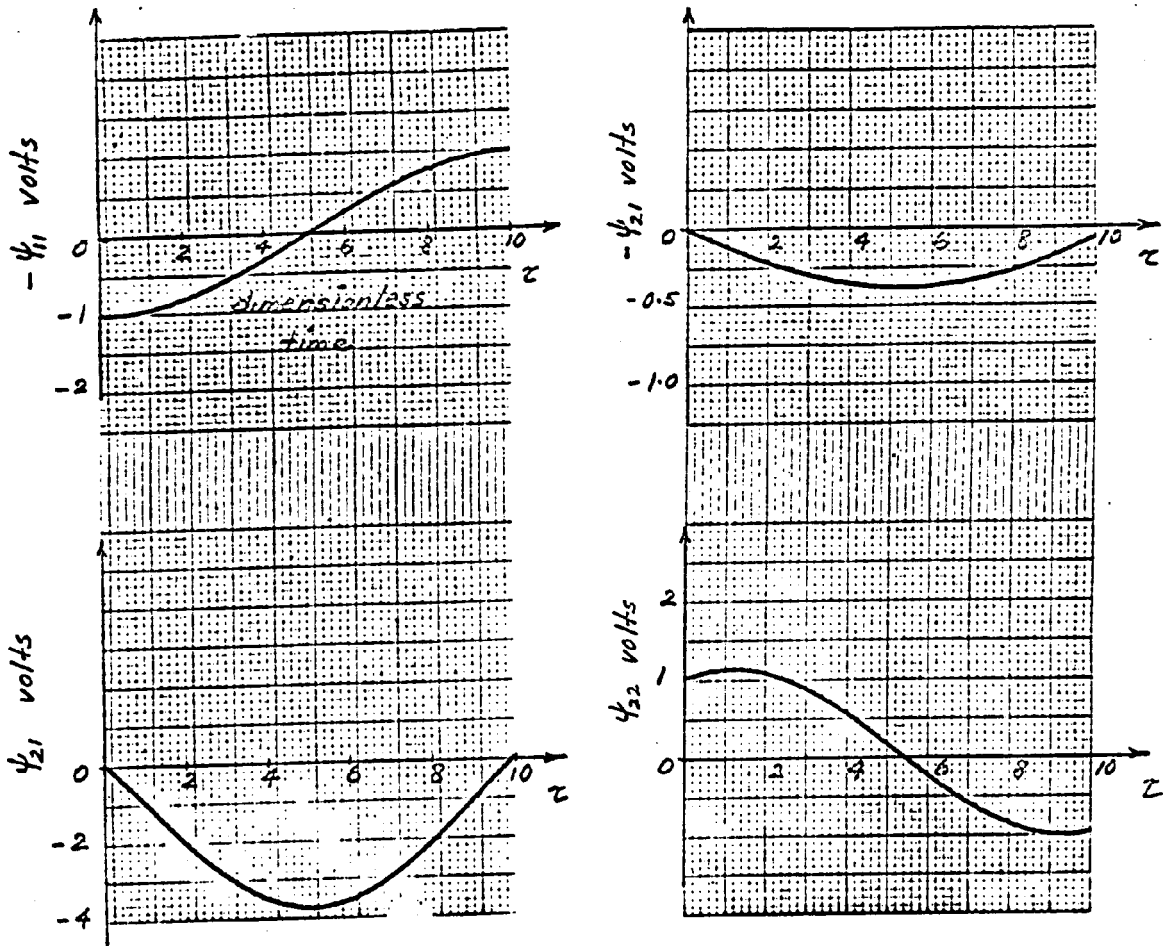


Fig. 6.9 Experimental Determination of the Elements of the Matrix Ψ

To generate the state transition matrix Φ, which is the transpose of the fundamental matrix of the system (6.24), we set up the computer diagram corresponding the system (6.24) which is shown in Fig. 6.10 the fundamental matrix $\Phi'(t)$ which satisfies the equation

$$\dot{\Phi}'(t) = C(t)\,\Phi'(t) \tag{6.32}$$

and

$$\Phi'(0) = I$$

can be found in the same manner as Ψ. $\Phi(t)$ can be obtained by a transpose operation of $\Phi'(t)$. The functions ϕ_{ik} of the matrix Φ are shown in Fig. 6.11.

As indicated in Fig. 6.10 the output of the random noise generator, n_o, is connected to the system of (6.24) through an amplifier of voltage gain $\frac{1}{10}$. This reduction in amplitude of the noise is necessary because we are considering only small disturbances in the system so that linearization of the system is valid. As indicated in section 6.2, the average power spectral density was found to be 5.28 volts^2/cps. Because of the amplifier gain of $\frac{1}{10}$, the actual random noise n_1 fed into the system (6.24) will have an average spectral density of 5.28×10^{-2} (volts^2/cps. As discussed in section 4.4 the random noise is assumed to have a constant spectral density of $4D_1$, therefore D_1 is given by

$$D_1 = \frac{5.28 \times 10^{-2}}{4} = 1.32 \times 10^{-2} \text{ volts}^2/\text{cps}. \tag{6.33}$$

Using the results obtained in the experiment, equations (6.27) can be used to calculate the theoretical values of the means and variances of the system (6.24) as functions of time.

From figures 6.7 and 6.11 one can read off the values of x_2, and of the elements of the matrices Φ and Ψ. These values are tabulated in Table 6.2

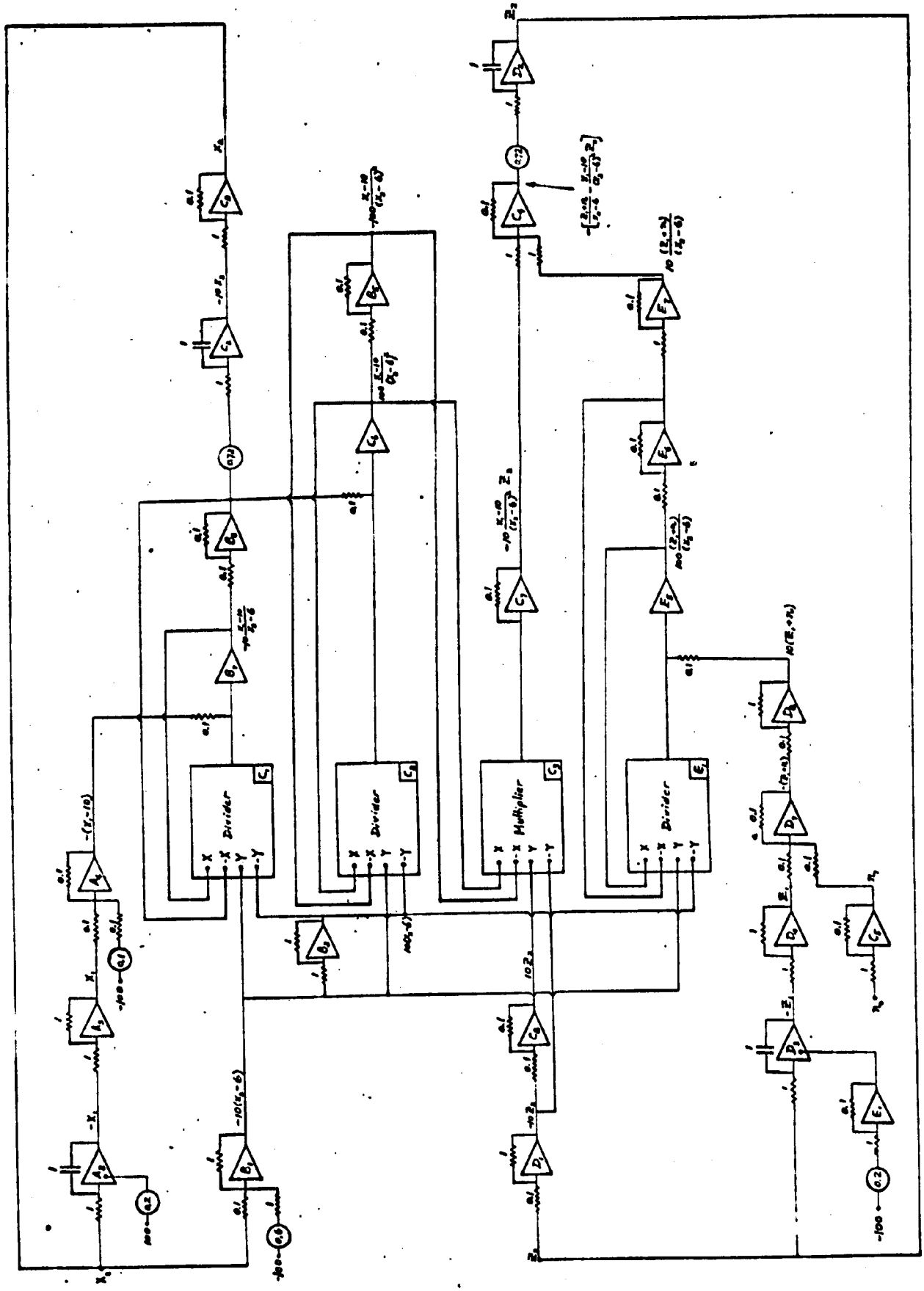


Fig. 6.10 Computer Setup for the Simulation of the System (6.24)

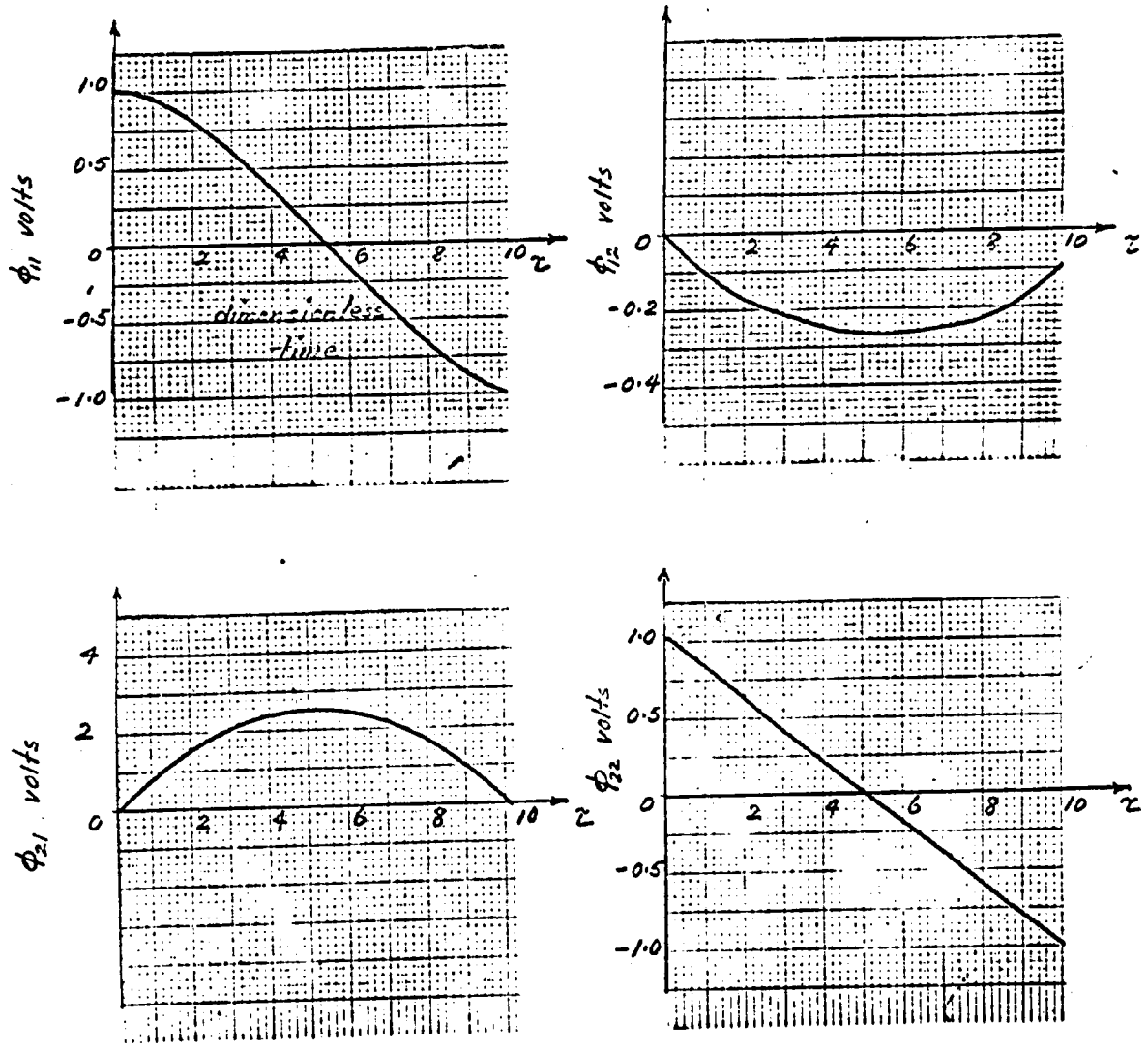


Fig. 6.11 Experimental Determination of the Elements of the Matrix $\underline{\Phi}$.

τ	x_2	ϕ_{11}	ϕ_{21}	ϕ_{12}	ϕ_{22}	ψ_{21}	ψ_{22}
0	0	1	0	0	1	0	1
0.5	-0.60	0.99	0.50	-0.05	0.92	-0.55	1.10
1.0	-1.20	0.95	0.90	-0.10	0.82	-1.10	1.13
1.5	-1.60	0.88	1.30	-0.12	0.72	-1.60	1.13
2.0	-2.00	0.80	1.65	-0.17	0.60	-2.20	1.08
2.5	-2.40	0.70	1.90	-0.19	0.50	-2.60	1.00
3.0	-2.60	0.60	2.15	-0.21	0.40	-3.00	0.87
3.5	-2.80	0.48	2.30	-0.23	0.30	-3.30	0.75
4.0	-3.00	0.37	2.40	-0.25	0.20	-3.50	0.56
4.5	-3.05	0.24	2.50	-0.26	0.10	-3.60	0.37
5.0	-3.10	0.10	2.50	-0.26	0	-3.70	0.18
5.5	-3.05	-0.03	2.50	-0.27	-0.10	-3.60	-0.02
6.0	-2.95	-0.15	2.40	-0.26	-0.20	-3.50	-0.22
6.5	-2.80	-0.30	2.30	-0.26	-0.30	-3.25	-0.40
7.0	-2.60	-0.43	2.10	-0.25	-0.40	-2.90	-0.60
7.5	-2.30	-0.55	1.85	-0.24	-0.50	-2.45	-0.73
8.0	-1.90	-0.67	1.55	-0.23	-0.60	-2.00	-0.85
8.5	-1.50	-0.78	1.20	-0.21	-0.70	-1.45	-0.93
9.0	-1.00	-0.87	0.80	-0.18	-0.80	-0.90	-0.98
9.5	-0.50	-0.95	0.40	-0.14	-0.90	-0.40	-1.00
10.0	0	-1.00	-0.10	-0.09	-1.00	+0.10	-0.95

Table 6.2 Tabulation of x_2, ϕ_{11}, ϕ_{21}, ϕ_{12}, ϕ_{22}, ψ_{21} and ψ_{22} at Different Values of Dimensionless Time.

From the first equation of (6.27) and Table 6.2, the theoretical expected values of Z_1 and Z_2 can be calculated with the results shown in Table 6.3 and plotted in Fig. 6.12.

τ	Calculated Theoretical $m_1(t)$	Experimental $m_1(t)$	Calculated Theoretical $m_2(t)$	Experimental $m_2(t)$
1	-1.90		0.20	
2	-1.60	-1.61	0.34	0.37
3	-1.20		0.42	
4	-0.74	-0.70	0.50	0.54
5	-0.20		0.53	
6	0.30	0.40	0.53	0.59
7	0.86		0.50	
8	1.34	1.51	0.46	0.51
9	1.74		0.36	
10	2.00	2.28	0.18	0.21

Table 6.3

Theoretical and Experimental Values of Means of the System (6.24)

From the second equation of (6.27), the variances λ_{rs} can be calculated as follows: From equation (4.21) the function f_{22} corresponding to the system (6.25) is given by

$$\frac{f_{22}(t)}{D_1} = d_{21}^2(t) = \frac{(0.72)^2}{(x_1 - 6)^2} . \tag{6.34}$$

From Table 6.2 and equation (6.34), Table 6.3 can be established as an intermediate step to calculate λ_{rs}'s.

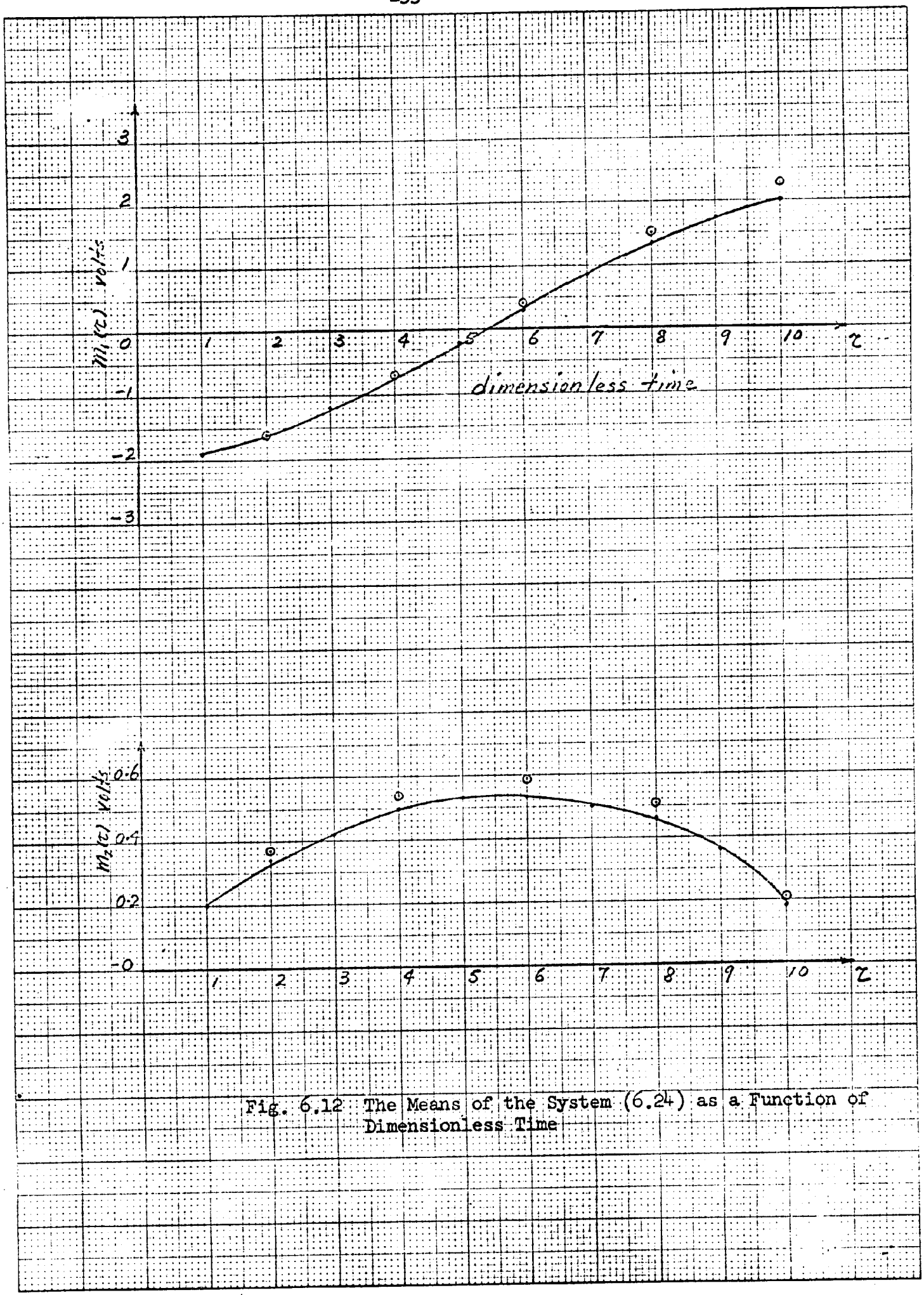


Fig. 6.12 The Means of the System (6.24) as a Function of Dimensionless Time

τ	$\frac{f_{22}}{D_1}\psi_{21}^2 \times 10^3$	$\frac{f_{22}}{D_1}\psi_{21}\psi_{22} \times 10^3$	$\frac{f_{22}}{D_1}\psi_{22}^2 \times 10^3$	$\int_0^\tau \frac{f_{22}}{D_1}\psi_{21}^2 d\tau \times 10^3$	$\int_0^\tau \frac{f_{22}}{D_1}\psi_{21}\psi_{22} d\tau \times 10^3$	$\int_0^\tau \frac{f_{22}}{D_1}\psi_{22}^2 d\tau \times 10^3$
0.5	3.57	- 7.14	14.40			
1.0	12.10	-12.40	12.77	4.89	- 7.44	14.17
1.5	23.01	-16.18	11.48			
2.0	39.15	-19.25	9.43	30.45	-23.56	26.00
2.5	50.23	-19.08	7.34			
3.0	63.09	-18.30	5.31	82.67	-43.33	33.33
3.5	72.85	-16.52	3.76			
4.0	78.40	-12.54	2.01	157.53	-59.56	37.22
4.5	84.31	- 8.54	0.87			
5.0	85.70	- 4.19	0.20	245.33	-68.33	38.17
5.5	84.31	0.44	0.02			
6.0	79.26	4.98	0.31	331.33	-67.78	38.39
6.5	70.65	8.70	1.07			
7.0	59.95	12.20	2.52	402.66	-59.89	39.28
7.5	45.12	13.46	4.01			
8.0	33.20	14.11	6.00	450.00	-44.56	43.44
8.5	19.34	12.43	7.96			
9.0	8.57	9.31	10.16	470.67	-31.11	52.11
9.5	1.96	4.90	12.27			
10.0	0.14	- 1.29	13.00	473.55	-27.00	64.83

Table 6.4

Tabulation of $\frac{f_{22}}{D_1}\psi_{21}^2 \times 10^3$, $\frac{f_{22}}{D_1}\psi_{21}\psi_{22} \times 10^3$ $\frac{f_{22}}{D_1}\psi_{22}^2 \times 10^3$ and Their Integrals at Different Values of Dimensionless Time.

The first three columns are plotted Figs. 6.13, 6.14 and 6.15 respectively and the curves in these figures are integrated by means of a planimeter with the results entered at the last three columns of Table 6.4. Using the equation of (6.27) together with (6.33) and Tables 6.2 and 6.4, the theoretical values of the variances of the system (6.24) are calculated. The results are tabulated in Table 6.5 and are plotted in Fig. 6.16.

12-282
5780-20
20 SQUARES TO THE INCH

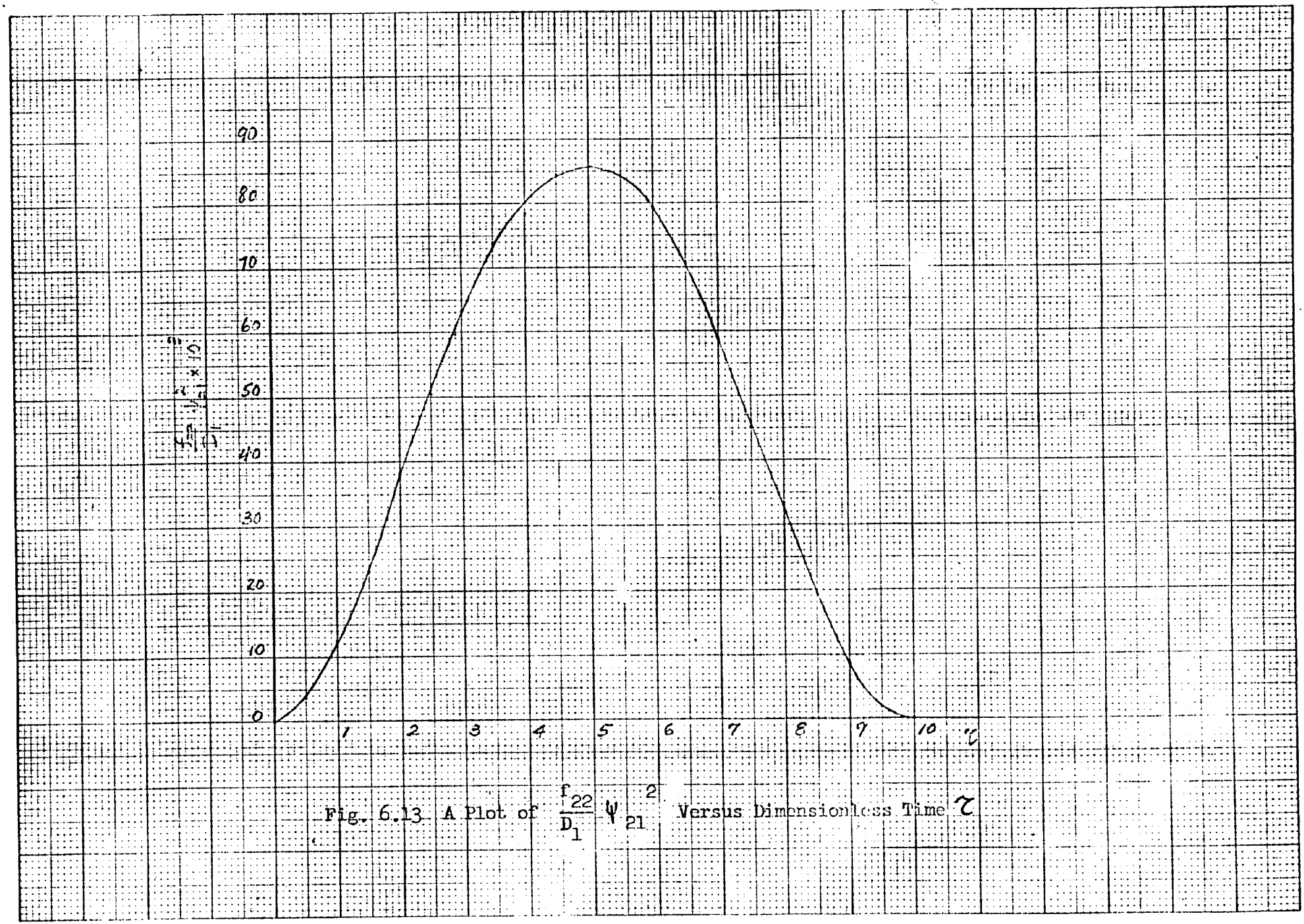


Fig. 6.13 A Plot of $\frac{f_{22}}{D_1}\psi_{21}^2$ Versus Dimensionless Time τ

NATIONAL
12-282
5780-20
20 SQUARES TO THE INCH

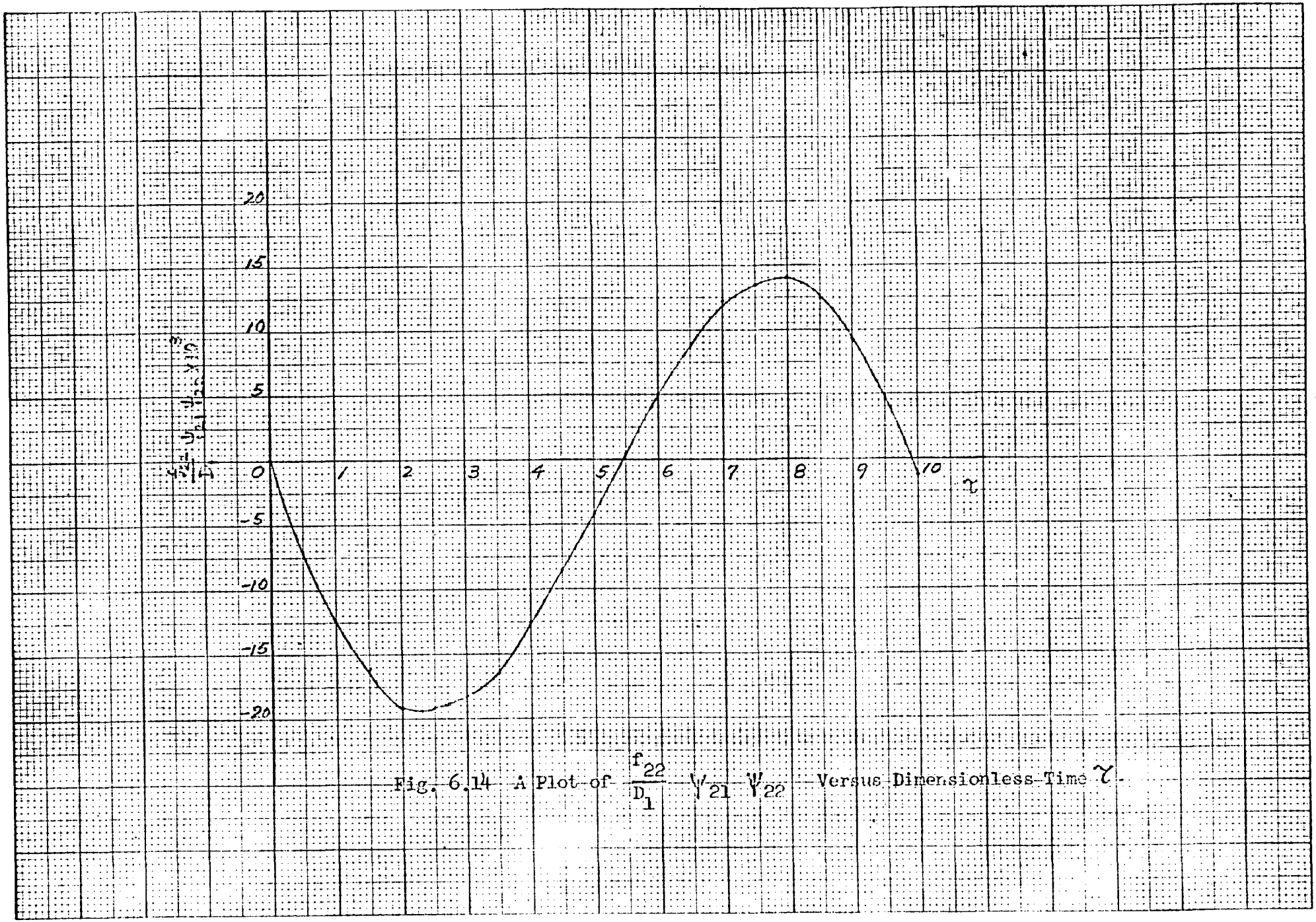


Fig. 6.14 A Plot of $\frac{f_{22}}{D_1} \psi_{21} \psi_{22}$ Versus Dimensionless Time τ.

12-282
5780-20
20 SQUARES TO THE INCH

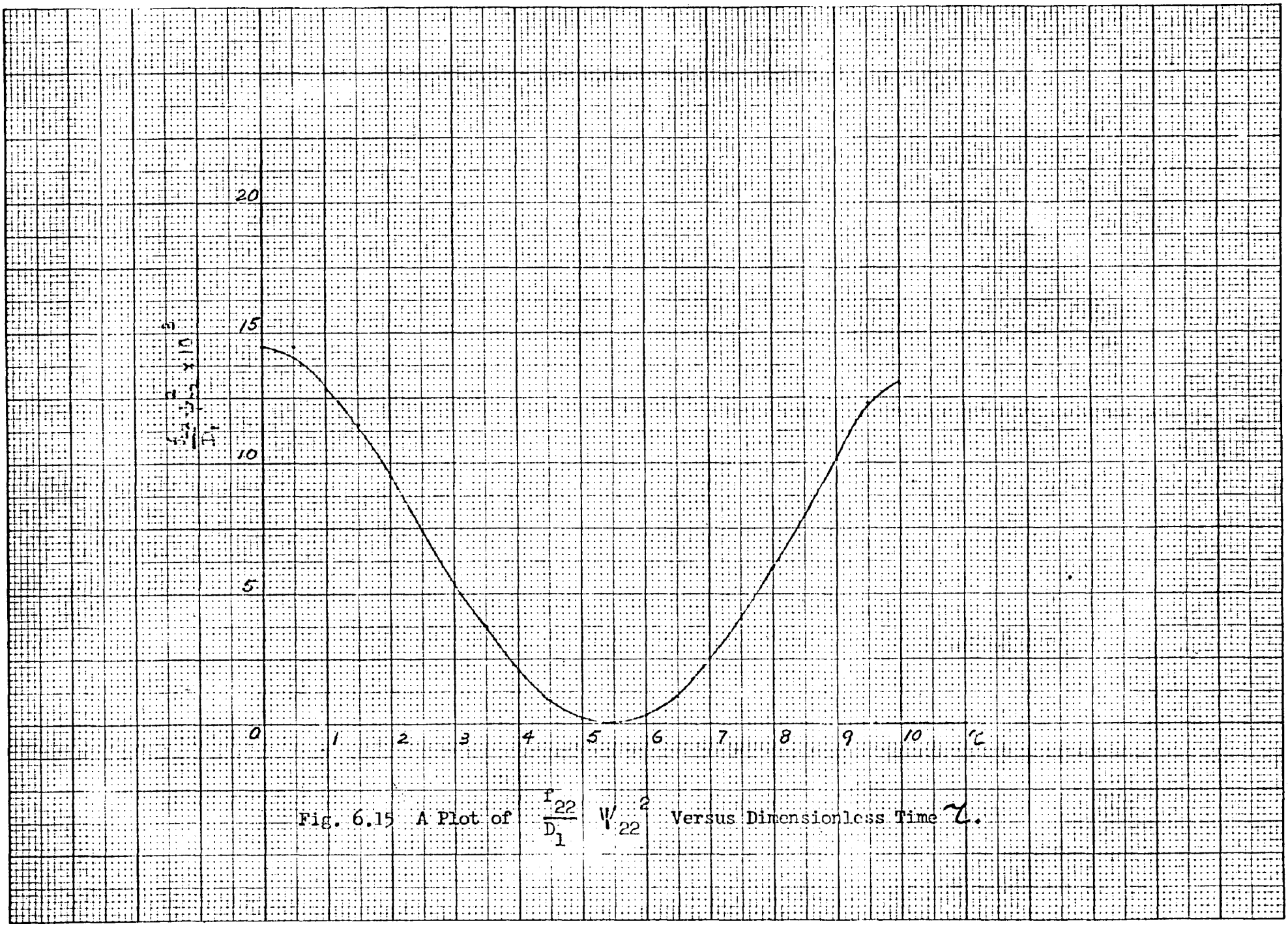


Fig. 6.15 A Plot of $\frac{f_{22}}{D_1} \Psi_{22}^2$ Versus Dimensionless Time τ.

τ	Calculated Theoretical $\lambda_{11} \times 10^4$	Experimental $\lambda_{11} \times 10^4$	Calculated Theoretical $\lambda_{12}=\lambda_{21} \times 10^4$	Experimental $\lambda_{12}=\lambda_{21} \times 10^4$	Calculated Theoretical $\lambda_{22} \times 10^4$	Experimental $\lambda_{22} \times 10^4$
1	0.84		1.28		2.85	
2	7.42	8.95	4.46	3.94	3.97	5.43
3	19.03		7.24		4.29	
4	34.40	42.94	9.14	9.68	4.57	5.26
5	54.60		10.23		4.55	
6	73.20	91.87	9.45	15.47	4.83	6.24
7	94.00		8.30		5.13	
8	105.30	123.70	7.10	8.26	7.17	11.21
9	114.20		6.17		10.47	
10	124.30	170.55	5.77	4.87	16.84	24.03

Table 6.5

Theoretical and Experimental Values of Variances of the System (6.24)

In order to check the theoretical means as well as the variances of the system (6.24), the random noice n_1 is applied to the system (6.24) as indicated in Fig. 6.9. The experiment is repeated for 100 times with a time interval of 10 seconds for each experiment. The results are recorded by a Sanborn recorder model 322 and tabulated in Table 6.6 with a typical sample shown in Fig. 6.17..

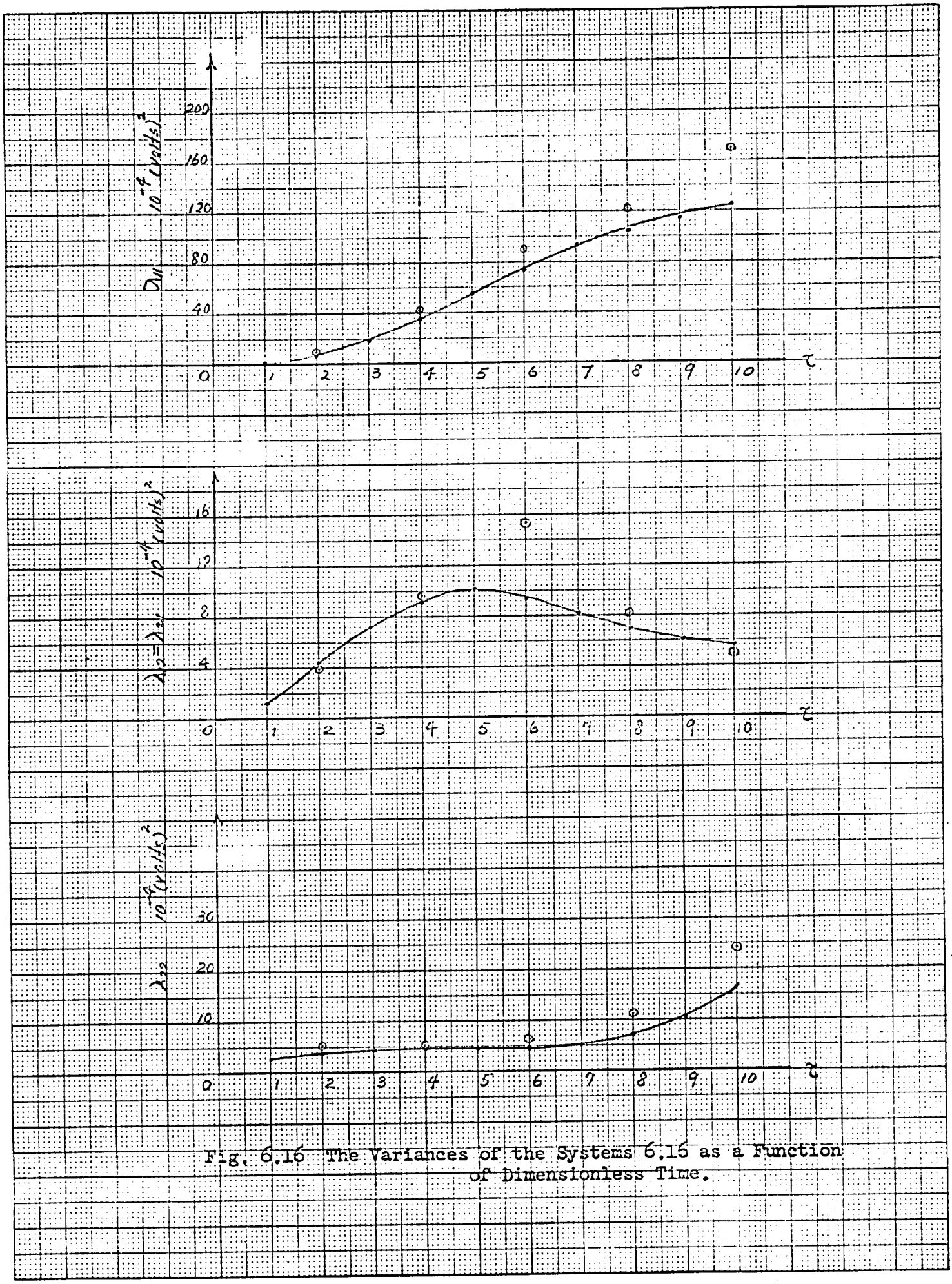


Fig. 6.16 The Variances of the Systems 6.15 as a Function of Dimensionless Time.

No. of Experiment	Z_1 Volts					Z_2 Volts				
	τ=2	τ=4	τ=6	τ=8	τ=10	τ=2	τ=4	τ=6	τ=8	τ=10
1	-1.60	-0.65	0.40	1.60	2.40	0.38	0.53	0.60	0.55	0.31
2	-1.60	-0.65	0.40	1.50	2.20	0.40	0.54	0.58	0.50	0.28
3	-1.60	-0.65	0.40	1.40	2.20	0.40	0.54	0.55	0.50	0.23
4	-1.60	-0.70	0.50	1.60	2.40	0.35	0.57	0.62	0.54	0.25
5	-1.60	-0.70	0.50	1.60	2.40	0.36	0.55	0.63	0.55	0.22
6	-1.60	-0.65	0.50	1.60	2.40	0.42	0.58	0.60	0.50	0.20
7	-1.65	-0.80	0.40	1.50	2.40	0.37	0.55	0.62	0.59	0.25
8	-1.65	-0.80	0.30	1.35	2.15	0.37	0.52	0.57	0.56	0.25
9	-1.65	-0.80	0.40	1.50	2.40	0.37	0.56	0.62	0.55	0.30
10	-1.65	-0.80	0.40	1.40	2.20	0.35	0.55	0.60	0.51	0.25
11	-1.60	-0.70	0.50	1.55	2.30	0.40	0.57	0.62	0.45	0.20
12	-1.60	-0.70	0.40	1.50	2.20	0.42	0.55	0.58	0.52	0.17
13	-1.60	-0.70	0.45	1.55	2.30	0.35	0.55	0.58	0.52	0.15
14	-1.65	-0.80	0.30	1.40	2.30	0.34	0.52	0.58	0.55	0.30
15	-1.65	-0.80	0.30	1.40	2.10	0.37	0.52	0.56	0.48	0.23
16	-1.65	-0.70	0.40	1.55	2.30	0.37	0.52	0.58	0.50	0.20
17	-1.60	-0.70	0.40	1.50	2.20	0.37	0.53	0.55	0.51	0.20
18	-1.60	-0.65	0.50	1.60	2.30	0.40	0.55	0.58	0.45	0.13
19	-1.60	-0.70	0.35	1.40	2.15	0.37	0.52	0.55	0.50	0.15
20	-1.60	-0.70	0.40	1.40	2.20	0.37	0.52	0.56	0.48	0.13
21	-1.60	-0.75	0.40	1.50	2.20	0.35	0.55	0.57	0.50	0.18
22	-1.65	-0.75	0.35	1.40	2.10	0.35	0.50	0.54	0.50	0.25
23	-1.60	-0.60	0.50	1.60	2.35	0.40	0.56	0.60	0.45	0.13
24	-1.60	-0.70	0.40	1.60	2.40	0.37	0.53	0.58	0.55	0.23
25	-1.65	-0.70	0.40	1.55	2.40	0.37	0.52	0.58	0.55	0.23
26	-1.60	-0.80	0.35	1.40	2.20	0.37	0.53	0.57	0.47	0.17
27	-1.60	-0.70	0.40	1.50	2.40	0.33	0.53	0.57	0.55	0.25
28	-1.60	-0.70	0.40	1.50	2.30	0.35	0.53	0.57	0.52	0.25

	τ=2	τ=4	τ=6	τ=8	τ=10	τ=2	τ=4	τ=6	τ=8	τ=10
29	-1.60	-0.60	0.60	1.70	2.35	0.40	0.58	0.60	0.43	0.08
30	-1.60	-0.60	0.65	1.75	2.50	0.38	0.60	0.61	0.53	0.17
31	-1.60	-0.65	0.50	1.60	2.40	0.38	0.56	0.62	0.45	0.23
32	-1.60	-0.60	0.60	1.70	2.45	0.38	0.58	0.62	0.45	0.16
33	-1.55	-0.55	0.60	1.70	2.45	0.40	0.55	0.62	0.45	0.10
34	-1.60	-0.70	0.40	1.50	2.30	0.37	0.52	0.57	0.51	0.20
35	-1.60	-0.65	0.50	1.65	2.50	0.37	0.55	0.60	0.55	0.25
36	-1.65	-0.70	0.40	1.55	2.35	0.37	0.55	0.60	0.52	0.20
37	-1.60	-0.70	0.40	1.50	2.25	0.38	0.53	0.57	0.50	0.19
38	-1.65	-0.70	0.50	1.55	2.25	0.40	0.58	0.60	0.50	0.15
39	-1.60	-0.70	0.40	1.45	2.20	0.38	0.53	0.57	0.49	0.25
40	-1.65	-0.70	0.40	1.45	2.30	0.35	0.53	0.60	0.56	0.17
41	-1.60	-0.65	0.50	1.60	2.40	0.40	0.55	0.60	0.52	0.18
42	-1.65	-0.80	0.40	1.60	2.45	0.35	0.54	0.62	0.52	0.27
43	-1.65	-0.70	0.40	1.50	2.30	0.38	0.55	0.59	0.50	0.24
44	-1.60	-0.65	0.50	1.65	2.40	0.39	0.57	0.63	0.54	0.20
45	-1.60	-0.70	0.50	1.60	2.50	0.37	0.55	0.63	0.57	0.26
46	-1.60	-0.60	0.60	1.70	2.40	0.42	0.58	0.61	0.44	0.14
47	-1.60	-0.65	0.55	1.70	2.50	0.42	0.58	0.61	0.55	0.17
48	-1.65	-0.70	0.50	1.65	2.50	0.38	0.57	0.61	0.52	0.25
49	-1.65	-0.70	0.45	1.55	2.40	0.34	0.56	0.60	0.55	0.25
50	-1.60	-0.65	0.45	1.50	2.30	0.38	0.55	0.57	0.53	0.28
51	-1.55	-0.65	0.50	1.60	2.45	0.38	0.57	0.62	0.56	0.25
52	-1.55	-0.60	0.60	1.70	2.45	0.42	0.57	0.61	0.52	0.17
53	-1.55	-0.55	0.55	1.55	2.30	0.38	0.53	0.57	0.52	0.25
54	-1.60	-0.60	0.60	1.55	2.30	0.40	0.57	0.62	0.57	0.30
55	-1.62	-0.75	0.40	1.60	2.45	0.33	0.56	0.64	0.56	0.29
56	-1.60	-0.70	0.42	1.55	2.35	0.37	0.52	0.59	0.55	0.24
57	-1.60	-0.70	0.34	1.40	2.20	0.37	0.53	0.55	0.52	0.30
58	-1.60	-0.60	0.55	1.60	2.30	0.40	0.54	0.62	0.53	0.15
59	-1.55	-0.60	0.55	1.70	2.50	0.40	0.58	0.60	0.54	0.22
60	-1.60	-0.55	0.60	1770	2.50	0.41	0.57	0.61	0.54	0.22

	τ=2	τ=4	τ=6	τ=8	τ=10	τ=2	τ=4	τ=6	τ=8	τ=10
61	-1.60	-0.65	0.45	1.60	2.45	0.38	0.56	0.63	0.56	0.25
62	-1.60	-0.65	0.60	1.60	2.30	0.39	0.57	0.61	0.51	0.13
63	-1.55	-0.60	0.55	1.60	2.35	0.40	0.58	0.60	0.44	0.13
64	-1.60	-0.70	0.50	1.60	2.35	0.35	0.55	0.61	0.54	0.22
65	-1.60	-0.70	0.50	1.55	2.35	0.35	0.52	0.58	0.52	0.27
66	-1.60	-0.70	0.40	1.50	2.20	0.36	0.55	0.58	0.48	0.18
67	-1.60	-0.70	0.40	1.50	2.30	0.37	0.55	0.62	0.53	0.24
68	-1.60	-0.70	0.40	1.45	2.15	0.40	0.54	0.58	0.50	0.18
69	-1.65	-0.80	0.30	1.30	2.10	0.36	0.50	0.55	0.50	0.28
70	-1.65	-0.75	0.30	1.35	2.00	0.37	0.54	0.57	0.48	0.20
71	-1.60	-0.70	0.35	1.40	2.10	0.35	0.52	0.55	0.47	0.15
72	-1.65	-0.80	0.30	1.35	2.15	0.32	0.50	0.56	0.48	0.22
73	-1.65	-0.70	0.45	1.55	2.25	0.35	0.55	0.59	0.50	0.18
74	-1.65	-0.82	0.25	1.30	2.10	0.32	0.50	0.58	0.48	0.29
75	-1.65	-0.70	0.38	1.40	2.20	0.35	0.52	0.56	0.50	0.28
76	-1.65	-0.80	0.30	1.35	2.05	0.35	0.50	0.55	0.47	0.17
77	-1.65	-0.75	0.30	1.35	2.05	0.35	0.52	0.55	0.47	0.18
78	-1.65	-0.80	0.25	1.35	2.10	0.35	0.51	0.55	0.49	0.22
79	-1.60	-0.70	0.35	1.40	2.10	0.36	0.52	0.55	0.47	0.17
80	-1.60	-0.75	0.32	1.40	2.15	0.36	0.52	0.57	0.49	0.28
81	-1.65	-0.75	0.35	1.40	2.10	0.35	0.53	0.56	0.48	0.20
82	-1.60	-0.70	0.35	1.40	2.15	0.37	0.53	0.57	0.48	0.21
83	-1.60	-0.75	0.30	1.40	2.20	0.34	0.52	0.59	0.52	0.29
84	-1.65	-0.80	0.30	1.40	2.10	0.35	0.52	0.55	0.48	0.19
85	-1.65	-0.80	0.20	1.30	2.00	0.33	0.50	0.58	0.47	0.21
86	-1.65	-0.70	0.38	1.40	2.20	0.38	0.52	0.56	0.50	0.20
87	-1.65	-0.80	0.25	1.30	2.10	0.35	0.50	0.56	0.50	0.29
88	-1.60	-0.65	0.48	1.55	2.25	0.38	0.56	0.59	0.48	0.18

	τ=2	τ=4	τ=6	τ=8	τ=10	τ=2	τ=4	τ=6	τ=8	τ=10
89	-1.65	-0.80	0.30	1.40	2.20	0.34	0.53	0.58	0.50	0.21
90	-1.65	-0.75	0.35	1.40	2.10	0.37	0.53	0.55	0.48	0.17
91	-1.65	-0.75	0.38	1.50	2.30	0.37	0.53	0.60	0.53	0.22
92	-1.65	-0.65	0.35	1.50	2.30	0.35	0.52	0.57	0.54	0.22
93	-1.65	-0.75	0.33	1.45	2.20	0.34	0.52	0.58	0.50	0.20
94	-1.65	-0.70	0.45	1.60	2.40	0.35	0.57	0.62	0.54	0.24
95	-1.60	-0.75	0.50	1.60	2.25	0.35	0.55	0.58	0.46	0.16
96	-1.65	-0,75	0.40	1.47	2.20	0.35	0.53	0.58	0.48	0.19
97	-1.65	-0.80	0.30	1.35	2.10	0.34	0.52	0.56	0.48	0.18
98	-1.60	-0.70	0.45	1.50	2.20	0.39	0.53	0.57	0.47	0.16
99	-1.60	-0.65	0.50	1.65	2.45	0.38	0.56	0.60	0.55	0.22
100	-1.65	-0.75	0.35	1.45	2.25	0.37	0.53	0.58	0.54	0.21

Table 6.6

Tabulation of Values of Z_1 and Z_2 of the System (6.24) taken from Experimental Curves

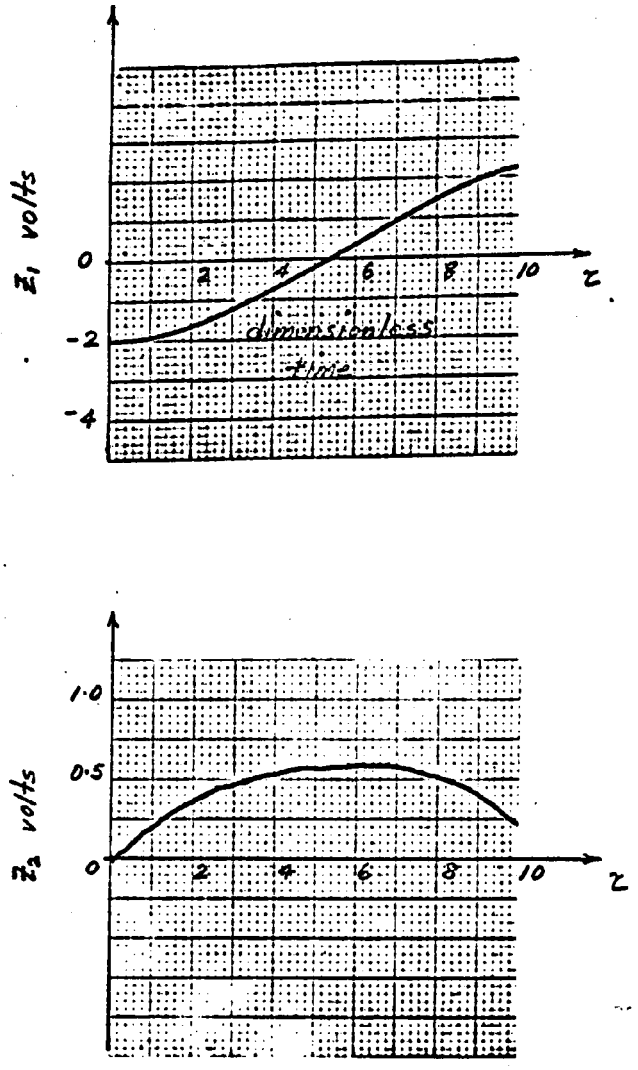


Fig. 6.17 A Typical Sample of the Output of the System (6.24) with Noise Disturbance n_1 .

From the data shown in Table 6.6 the means and variances of the system (6.24) can be calculated according to the following formulas

$$m_j(t) \cong \sum_{i=1}^{100} \frac{z_j^{\,i}(t)}{100} \qquad j = 1,2 \tag{6.35}$$

and

$$\lambda_{jk}(t) \cong \sum_{i=1}^{100} \frac{(z_j^{\,i}(t) - m_j(t))(z_k^{\,i}(t) - m_k(t))}{100} \qquad j,k = 1,2$$

with the results tabulated in Tables 6.3 and 6.5 and plotted in Figs. 6.12 and 6.16 respectively.

It is found from Figs. 6.12 and 6.16 that the experimental results are in close agreement with theory. However, it should be pointed out that confidence intervals are not used to compare the experimental and theoretical curves. This is due to the fact the confidence intervals are exceedingly small and in fact that they are well within the limits of experimental errors. For example, the 95% confidence limits for the mean $m_1(t)$ is given by [56]

$$\bar{m}_1(t) \pm 1.96 \frac{\sigma_1(t)}{\sqrt{N}} \tag{6.36}$$

where $\bar{m}_1(t)$ is the sample mean of Z_1, $\sigma_1(t)$ is the theoretical standard deviation of Z_1 and N is the number of observations, which in this case is 100. From Table 6.5 the 95% confidence limits range from $\pm 1.8 \times 10^{-3}$ volts to $\pm 2.19 \times 10^{-2}$ volts. In the measurement of Z_1, the Sanborn recorder was set at the sensitivity of 0.2 volts per millimeter. With a reading error of 0.2 of a millimeter this will amount to an error of 4×10^{-2} volts. For this reason the confidence intervals cannot be used for comparing the theoretical and experimental values.

The discrepancies between the experimental and theoretical values are believed mainly due to the following factors: (1) The measured spectral density of the random noise generator varies as a function of frequency as opposed to the assumption that it is a constant. In the

calculation of theoretical values of variances, the average value of the spectral density is used. (2) As explained in previous paragraph reading, error is a very important factor. (3) As can be seen in Fig. 6.12, the experimental and theoretical values of the mean values of Z_1 and Z_2 are numerically very close together. However, all experimental values are slightly higher than the corresponding theoretical values. It is believed that these discrepancies are likely due to (a) a slight error made in the calibration of the recorder, (b) a slight error in setting initial conditions, and (c) a small d-c component in the noise generator.

Chapter 7

CONCLUSIONS AND FUTURE RESEARCH

7.1 Conclusions

This dissertation has presented a general study of variations in optimal control systems subject to deterministic and stochastic disturbances. This covers both open-and closed-loop cases. The first part of Chapter 2 was devoted to calculating the variations in the trajectory and cost of an open-loop optimal control system due to changes in plant parameters about their nominal values. This was done by solving the linearized equation of the system (2.9) along its nominal trajectory. For a given change in the plant parameter vector α, the variation along the optimal trajectory and the change in the cost were found to be, respectively,

$$\Delta x(\alpha_o,t) = \int_{t_o}^{t} \underline{\Phi}(t,\tau)\, G\,(t)\, \Delta\alpha\, d\tau \qquad (7.1)$$

and

$$\Delta J = \int_{t_o}^{t_1} \sum_{j=1}^{n} \frac{\partial f_o}{\partial x_j}\ \Delta x_i\ (\alpha_o,t)\ dt\ , \qquad (7.2)$$

where $\underline{\Phi}$ is the transition matrix of the system (2.9) satisfying equation (2.11). The remaining part of Chapter 2 was devoted to the estimation of the variations of plant parameters under the assumption that the actual trajectory can be observed. For the constant parameter case the Schmidt orthogonalization process was used to construct an orthogonal basis for $\Delta x\ (\alpha_o,t)$ so that the variations in plant parameters can be uniquely determined. In the case of time-dependent parameters the modified methods of steepest descent were used so that variations in plant parameters can be approximated by means of iteration processes.

The analysis of variations in trajectory and cost as well as the estimation of actual parameter variations under different situations were repeated in Chapter 3 for closed-loop optimal control systems. It was found that if the optimal control u(t) is a continuous vector function of the state variables $x(t) = (x_1(t),\ldots,x_n(t)$, the methods developed in Chapter 2 can be applied directly to closed-loop optimal control systems as well.

In Chapter 4 a continuous multidimensional process was defined. The probabilistic characterization of a Markor process was given. The Fokker-Planck equation corresponding to a linear time-varying system excited by white noise was derived. The reduction of the Fokker-Planck equation associated with a given system of linear time-varying differential equations excited by white noise to the solution of a set of ordinary linear differential equations with variable coefficients was carried out by means of a Fourier transformation and the method of characteristics. As a result one is able to determine the expected values as well as the moments of the solution of a given system at every instant of time even when the initial position of the system is not exactly known, but only given by its probability distribution. A white noise random process can be considered only as a limiting case which will not occur in practice so that the remaining part of Chapter 4 was devoted to analyze somewhat more general problems, that is, linear time-varying systems with disturbances generated by linear differential equations excited by white noise. The analysis was based on the method of expanding the state space of the original system

into another state space of a higher dimension. By doing so one is able to determine the behavior of the system when the disturbances are non-white as well as non-stationary.

The results developed in Chapter 4 were applied directly in Chapter 5 for the determination of average changes in trajectory and cost of optimal control systems due to random disturbances in plant parameters or noisy observation of state variables.

Chapter 6 was devoted to experimental verification of the theoretical results obtained in Chapters 4 and 5. A simple second-order optimal control system subject to stochastic disturbance was simulated on an analog computer. It was found that the results of the experiment were in close agreement with theoretical calculations.

The following ideas and developments contained in this dissertation are believed, to the best knowledge of the writer, to be of a contributory nature, i.e., original contributions.

1) A study of variations in trajectory and cost of optimal control systems of a very general nature.

2) The application of the modified Kantorovich's steepest descent method to the estimation of plant parameter variations in the inverse sensitivity problem.

3) The generalization of the method of Wang and Uhlenbeck to linear time-varying systems excited by white noise.

4) By expanding the state space of the original system into a higher dimensional state space the behavior of a system subject to disturbances or estimation errors which are non-white as well as non-stationary can be determined by the last mentioned method.

7.2 Suggestions for Further Research

1) In the analysis of the closed-loop control systems the control law is assumed to be a continuous function of the state variables. The analysis should be extended to include other forms of control laws such as those which are discontinuous functions of of state variables.

2) The problem of compensating the sensitivity of the system to parameter variations is very important. The writer believes that this problem for nonlinear optimal control systems is very worthwhile but a difficult task.

3) As discussed in Chapter 6 the method of confidence intervals cannot be used to compare the experimental and theoretical curves due to the fact that the confidence intervals are well within the limits of experimental errors, it is suggested that simulations of physical systems be made by means of a digital computer so that experimental results can be improved.

Appendix A

DERIVATION OF (2.22)

The proof of $|g(x_1,\alpha,t) - g(x_2,\alpha,t)| \le |n^2K|\,|x_2 - x_1|$ can be found in the book by Pontryagin[57]. However, it is repeated here for convenience.

$$|g(x_1,\alpha,t) - g(x_2\alpha,t)|^2 = \sum_{i=1}^{n} |g_i(x_1,\alpha,t) - g_i(x_2,\alpha,t)|^2$$

$$= \sum_{i=1}^{n} \Big| \sum_{j=1}^{n} \frac{\partial g_i}{\partial x_j}(x_s,\alpha,t)(x_{1_j} - x_{2_j}) \Big|^2$$

$$x_s = x_1 + s(x_2 - x_1)$$

$$0 \le s \le 1$$

$$\le \sum_{i=1}^{n} \Big(\sum_{j=1}^{n} \Big| \frac{\partial g_i}{\partial x_j}(x_s,\alpha,t) \Big|\, |x_{1_j} - x_{2_j}| \Big)^2$$

$$\le \sum_{i=1}^{n} K^2 \Big(\sum_{j=1}^{n} |x_{1_j} - x_{2_j}| \Big)^2$$

$$\le \sum_{i=1}^{n} K^2 \Big(\sum_{j=1}^{n} |x_1 - x_2| \Big)^2$$

$$= \sum_{i=1}^{n} n^2K^2\,|x_1 - x_2|^2$$

$$= n^3K^2\,|x_1 - x_2|^2 .$$

Therefore,

$$|g(x_1,\alpha,t) - g(x_2,\alpha,t)| \leq n^{3/2}K|x_1 - x_2| \leq n^2K|x_1 - x_2|.$$

Appendix B

VERIFICATION OF (2.36)

Let us define the norms of a vector function $x(t)$ and of a matrix $M(t)$, respectively, by

$$\|x(t)\| = \underset{t\in[t_o,t_1]}{\text{Max}} |x(t)| = \underset{t\in[t_o,t_1]}{\text{Max}} \left[\sum_{i=1}^{n} |x_i(t)|^2\right]^{1/2}, \tag{B.1}$$

and

$$\|M(t)\| = \underset{t\in[t_o,t_1]}{\text{Max}} \; [\text{trace } M^* M]^{1/2}, \tag{B.2}$$

where M^* is the transpose of M. Let $y = Mx$, then

$$\|y\|^2 = \max_{t\in[t_o,t_1]} \sum_{j=1}^{m} \left| \sum_{i=1}^{n} a_{ji} \; x_i \right|^2$$

$$\leq \max_{t\in[t_o,t_1]} \sum_{j=1}^{m} \left\{ \sum_{i=1}^{n} a_{ji}^2 \right\} \left\{ \sum_{i=1}^{n} x_i^2 \right\} \text{ by Schwarz inequality}$$

$$= \max_{t\in[t_o,t_1]} \left\{ \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ji}^2 \right\} \left\{ \sum_{i=1}^{n} x_i^2 \right\}$$

$$\leq \max_{t\in[t_o,t_1]} \left\{ \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ji}^2 \right\} \max_{t\in[t_o,t_1]} \left\{ \sum_{i=1}^{n} x_i^2 \right\}$$

therefore

$$\|y\| \leq \max_{t\in[t_o,t_1]} \left\{ \sum_{j=1}^{m} \sum_{i=1}^{n} a_{ji}^2 \right\}^{1/2} \|x\|$$

$$= \max_{t\in[t_o,t_1]} [\text{trace } M^*M]^{1/2} \|x\|$$

$$= \|M(t)\| \|x\|$$

Appendix C

DERIVATION OF THE FOKKER-PLANCK EQUATION

In order to show that the function $P(t,x/\tau,y)$ satisfies (4.9) let us consider the integral

$$\int R(y) \frac{\partial P(t,x/\tau,y)}{\partial \tau} dy \quad , \tag{C.1}$$

where $R(y)$ is assumed to be a non-negative continuous function having continuous partial derivatives up to and including those of the second order. Furthermore, let it be required that

$$R(y) = 0 \quad \text{for } y_i < a \text{ and } y_i > b \quad (b \gtrless a)$$

$$i = 1,\ldots, n \tag{C.2}$$

It follows immediately from the continuity condition imposed on $R(y)$ that

$$R(y) = \frac{\partial R(y)}{\partial y_i} = \frac{\partial^2 R(y)}{\partial y_i \partial y_j} = 0 \quad \text{at } y_i = a \text{ and } y_i = b \tag{C.3}$$

$$i,j = 1,\ldots, n$$

We first note that

$$\int R(y) \frac{\partial P(t,x/\tau,y)dy}{\partial \tau} = \lim_{\Delta\tau \to 0} \frac{1}{\Delta\tau} \int R(y)[P(t,x/\tau + \Delta\tau,y) - P(t,x/\tau,y)]dy$$

Making use of the generalized Markov equation

Making use of the generalized Markov equation

$$P(t,x/\tau + \Delta\tau,y) = \int P(t,x/\tau,z)\, P(\tau,z/\tau + \Delta\tau,y)dz \qquad (C.4)$$

we find

$$\int R(y)\frac{\partial P(t,x/\tau,y)dy}{\partial\tau} = \lim_{\Delta\tau\to 0}\frac{1}{\Delta\tau}\left\{\int R(y)\left[\int P(t,x/\tau,z)P(\tau,z/\tau+\Delta\tau,y)dz\right]dy - \int R(y)\,P(t,x/\tau,y)dy\right\}$$

$$= \lim_{\Delta\tau\to 0}\frac{1}{\Delta\tau}\left\{\int P(t,x/\tau,z)dz\int R(y)P(\tau,z/\tau+\Delta\tau,y)dy - \int R(y)P(t,x/\tau,y)dy\right\}$$

$$= \lim_{\Delta\tau\to 0}\frac{1}{\Delta\tau}\left\{\int P(t,x/\tau,y)dy\int R(z)P(\tau,y/\tau+\Delta\tau,z)dz - \int R(y)P(t,x/\tau,y)dy\right\}$$

$$\int R(y)\,\frac{\partial P(t,x/\tau,y)dy}{\partial\tau} = \lim_{\Delta\tau\to 0}\frac{1}{\Delta\tau}\left\{P(t,x\,\tau,y)\left[\int R(z)P(\tau,y/\tau,+\Delta\tau,z)dz - R(y)\right]dy\right\} \qquad (C.5)$$

By Taylor's theorem

$$R(z)=R(y) + \sum_{i=1}^{n}\frac{\partial R(y)}{\partial y_i}(z_i-y_i) + \frac{1}{2}\sum_{i,j=1}^{n}\frac{\partial^2 R(y)}{\partial y_i\,\partial y_j}(z_i-y_i)(z_j-y_j) + O[|z-y|^2] \qquad (C.6)$$

Because of the boundedness of R(z) and the strong continuity condition (4.5) we find

$$\int_{|z-y| \geq \delta} R(z)\, P(\tau, y/\tau + \Delta\tau, z)dz = o(\Delta\tau) \tag{C.7}$$

and

$$\int_{|z-y| < \delta} P(\tau, y/\tau + \Delta\tau, z)dz = 1 + o(\Delta\tau) \tag{C.8}$$

It follows that from (C.6), (C.7), and (C.8)

$$\int R(z)\, P(\tau, y/\tau + \Delta\tau, z)dz - R(y) =$$

$$= \int_{|z-y| \geq \delta} P(\tau, y/\tau + \Delta\tau, z)R(z)dz + \int_{|z-y| < \delta} P(\tau, y/\tau + \Delta\tau, z)R(z)dz - R(y)$$

$$= \sum_{i=1}^{n} \frac{\partial R(y)}{\partial y_i} \int_{|z-y|<\delta} (z_i - y_i)P(\tau, y/\tau + \Delta\tau, z)dz$$

$$+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 R(y)}{\partial y_i\, \partial y_j} \int_{|z-y|<\delta} [(z_i - y_i)(z_j - y_j) + o[|z-y|^2]]P(\tau, y/\tau + \Delta\tau, z)dz + o(\Delta\tau)$$

Hence (C.5) becomes

$$\int R(y)\, \frac{\partial P(t, x/\tau, y)dy}{\partial \tau} =$$

$$= \lim_{\Delta\tau \to 0} \frac{1}{\Delta\tau} \left\{ \int P(t, x/\tau, y)\left[\sum_{i=1}^{n} \frac{\partial R(y)}{\partial y_i} \int_{|z-y|<\delta} (z_i - y_i)P(\tau, y/\tau + \Delta\tau, z)dz + \right.\right.$$

$$\left. + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 R(y)}{\partial y_i\, \partial y_j} \int [(z_i - y_i)(z_j - y_j) + o(|z-y|^2)]P(\tau, y/\tau + \Delta\tau, z)dz \right.$$

$$\left. + o(\Delta\tau)]dy \right\}$$

Let $\Delta\tau \to 0$, by the uniform convergence of (4.6) and (4.7) we obtain

$$\int \frac{\partial P(t,x/\tau,y)}{\partial \tau} R(y)dy = \int P(t,x/\tau,y)\left[\sum_{i=1}^{n} \frac{\partial R(y)}{\partial y_i} a_i(\tau,y) + \right.$$

$$\left. + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 R(y)}{\partial y_i \partial y_j} b_{ij}(\tau,y)\right] dy \qquad (C.9)$$

Integrate by parts the expressions on the right hand side of (C.9) with respect to the variables y_i and y_j , we find

$$\int_a^b \left\{ \frac{\partial P(t,x/\tau,y)}{\partial \tau} + \sum_{i=1}^{n} \frac{\partial}{\partial y_i} [a_i(\tau,y)P(t,x/\tau,y)] \right.$$

$$\left. - \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial y_i \partial y_j} [b_{ij}(\tau,y)P(t,x/\tau,y)] \right\} R(y)\, dy = 0 \qquad (C.10)$$

The end conditions of R(y) in (C.3) have been used to obtain the above result.

Since R(y) is arbitrary, we obtain from (C.10) the expression

$$\frac{\partial P(t,x/\tau,y)}{\partial \tau} = -\sum_{i=1}^{n} \frac{\partial}{\partial y_i} [a_i(\tau,y)P(t,x/\tau,y)]$$

$$+ \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2}{\partial y_i \partial y_i} [b_{ij}(\tau,y)P(t,x/\tau,y)] , \qquad (C.11)$$

which is the Fokker-Planck equation.

Appendix D

DERIVATION OF (5.17)

To show that the optimal control law of the system (5.15) is given by

$$u(t) = \sum_{i=1}^{n} P_i x_i \quad , \tag{D.1}$$

we use calculus of variations since we are dealing with Lagrange variational problem. We form the following function

$$H = V + \sum_{K} \lambda_K (\dot{x}_K - \sum_{\alpha} b_{K\alpha} x_\alpha - M_K u)$$

$$= \sum_{K} a_K x_K^2 + Cu^2 + \sum_{K} \lambda_K (\dot{x}_K - \sum_{\alpha} b_{K\alpha} x_\alpha - M_K u) \quad , \tag{D.2}$$

where the λ_K are arbitrary multipliers. Then

$$\frac{\partial H}{\partial \dot{x}_K} = \lambda_K \quad , \quad \frac{\partial H}{\partial x_K} = 2a_K x_K - \sum_{\alpha} \lambda_\alpha b_{\alpha K}$$

$$\tag{D.3}$$

$$\frac{\partial H}{\partial \dot{u}} = 0 \quad , \quad \frac{\partial H}{\partial u} = 2Cu - \sum_{\alpha} M_\alpha \lambda_\alpha$$

Using Euler's equation we have

$$\dot{\lambda}_K = -\sum_j b_{jK}\lambda_j + 2a_K x_K \,, \quad 0 = 2Cu - \sum_j m_j \lambda_j \tag{D.4}$$

From the second equation of (D.4)

$$u = \sum_j \frac{m_j \lambda_j}{2C} \tag{D.5}$$

In terms of components equation (5.15) can be written

$$\dot{x}_K = \sum_j b_{Kj} x_j + \sum_j \frac{m_K m_j \lambda_j}{2C} \quad K = 1,\ldots,n \tag{D.6}$$

Adjoining (D.4) into (D.6) we have

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_n \\ \dot{\lambda}_1 \\ \dot{\lambda}_2 \\ \vdots \\ \dot{\lambda}_n \end{pmatrix} = \begin{pmatrix} b_{11},\ldots,b_{1n}, & \frac{m_1^2}{2C}, & \frac{m_1 m_2}{2C},\ldots, & \frac{m_1 m_n}{2C} \\ \cdots\cdots\cdots\

The system's characteristic determinant becomes

$$\Delta(\mu)=\begin{vmatrix} b_{11}-\mu,\ldots, b_{1n}, & \frac{m_1^2}{2C}, \ldots, & \frac{m_1 m_n}{2C} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \\ b_{n1},\ldots, b_{nn}-\mu, & \frac{m_n m_1}{2C}, \ldots, & \frac{m_n^2}{2C} \\ 2a_1,\ldots, 0, & -b_{11}-\mu, \ldots, & -b_{n1} \\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots & & \\ 0,\ldots,2a_n, & -b_{1n},\ldots, & -b_{nn}-\mu \end{vmatrix} \qquad (D.8)$$

It can be shown that if $\mu_1,\ldots,\mu_n$ are simple roots of $\Delta(\mu)$ then $-\mu_1,\ldots,-\mu_n$ are also simple roots. We assume that $\mu_1,\ldots,\mu_n$ are simple roots with negative real parts. If we let $y = (x_1,\ldots,x_n,\lambda_1,\ldots,\lambda_n)$ then the solution of (D.7) is given by[55]

$$y = C_1 e^{\mu_1 t} v_1 + \ldots + C_n e^{\mu_n t} v_n + C_{n+1} e^{-\mu_1 t} v_{n+1} + \ldots \pm C_{2n} e^{-\mu_n t} v_{2n} \qquad (D.9)$$

where $C_1,\ldots,C_{2n}$ are arbitrary constants and $v_1,\ldots,v_{2n}$ are linearly independent eigenvectors. By virtue of boundary conditions we get

$$C_{n+K} = 0 \ (K = 1,\ldots, n),$$

where the constants C_K can be chosen from initial conditions.

In the remaining 2n formulas we eliminate the n functions of time $C_K e^{\mu_K t}$, as a result we obtain

$$\lambda_j = \sum_\alpha d_j x_\alpha \tag{D.10}$$

where $d_{j\alpha}$ are completely defined constants.

Combining (D.5) and (D.10) we obtain

$$u = \sum_j \frac{m_j}{2C} \sum_\alpha d_{j\alpha} x_\alpha = \sum_\alpha \left(\sum_j \frac{m_j d_j}{2C} \right) x_\alpha$$

or

$$u = \sum_\alpha p_\alpha x_\alpha$$

which is what we want to show.

Appendix E

DERIVATION OF (5.31)

To show that the optimal control which drives t'.e system (5.24) from any initial state ξ to 0 and at the same time minimizes the fuel

$$J = \int_0^T k \, \|u(t)\| \, dt \tag{E.1}$$

is given by (5.31).

For any vector x(t) we differentiate

$$\frac{d}{dt} \|x(t)\| = \frac{d}{dt} \left[\sum_{i=1}^{n} x_i^2(t) \right]^{1/2} \tag{E.2}$$

$$= \frac{1}{2} \left[\sum_{i=1}^{n} x_i^2(t) \right]^{-1/2} 2[\dot{x}_i(t) x_i(t)]$$

$$= \frac{[\dot{x}(t), x(t)]}{\|x(t)\|} = \left[\dot{x}(t), \frac{x(t)}{\|x(t)\|} \right]$$

By virtue of (5.24) and (5.28)

$$\frac{d}{dt} \|x(t)\| = [u(t), \frac{x(t)}{\|x(t)\|}] \tag{E.3}$$

If $x(0) = \xi$, then

$$\|x(t)\| - \|\xi\| = \int_0^t [u(\tau), \frac{x(\tau)}{\|x(\tau)\|}] \, d\tau \tag{E.4}$$

Let $x(T) = 0$. Then

$$-\|\xi\| = \int_0^T [u(\tau) , \frac{x(\tau)}{\|x(\tau)\|}] \, d\tau \quad . \tag{E.5}$$

By Schwarz inequality

$$\|\xi\| \leq \int_0^T \|u(\tau)\| \, d\tau \tag{E.6}$$

or

$$k\|\xi\| \leq \int_0^T k\|u(\tau)\| \, d\tau = J \tag{E.7}$$

But the right hand side of (E.7) is the fuel consumed by u. In particular, if the control (5.31) is used one finds

$$k\|\xi\| = \int_0^{t^*} kM \, dt = kMt^* = J^*$$

It has been shown by Athan that the control (5.31) is also the time optimal control. Therefore

$$J^* \leq J$$

for any u.

Appendix

DERIVATION OF (5.38)

To show that

$$u(t) = - \frac{\|\xi\|}{T} \frac{x(t)}{\|x(t)\|} \tag{F.1}$$

provides a solution to the minimum energy problem we write

$$\frac{d}{dt} \|x(t)\| = [u(t), \frac{x(t)}{\|x(t)\|}] \tag{F.2}$$

If u(t) satisfies (1), then

$$\frac{d}{dt} \|x(t)\| = - \frac{\|\xi\|}{T} \tag{F.3}$$

Hence $$\|x(t)\| = \|\xi\| - \frac{\|\xi\|}{T} t \tag{F.4}$$

which gives $\|x(T)\| = 0$. The energy corresponding to this u is given by

$$J^* = \frac{\|\xi\|^2}{2T} \tag{F.5}$$

Consider any other control $\hat{u}(t)$ which drives ξ to 0 in T with the corresponding energy $\hat{J}$. Let $\hat{x}(t)$ be the corresponding solution of (5.24). Then

$$\left| \frac{d}{dt} \|\hat{x}(t)\| \right| = \left| [\hat{u}(t), \frac{\hat{x}(t)}{\|\hat{x}(t)\|}] \right| \leq \|\hat{u}(t)\| \tag{F.6}$$

Therefore

$$\frac{d}{dt} \|\hat{x}(t)\| \geq - \|\hat{u}(t)\| \tag{F.7}$$

and

$$\|\hat{x}(t)\| - \|\xi\| \geq - \int_0^t \|\hat{u}(\tau)\| \, d\tau \ . \tag{F.8}$$

But it is required that

$$\|\hat{x}(T)\| = 0 \ . \tag{F.9}$$

Therefore,

$$\|\xi\| \leq \int_0^T \|\hat{u}(\tau)\| \, d\tau \tag{F.10}$$

The energy corresponding to $\hat{u}(t)$ is

$$\hat{J} = \frac{1}{2} \int_0^T \|\hat{u}(t)\|^2 \, dt \tag{F.11}$$

By Schwarz inequality

$$T \int_0^T \|\hat{u}(t)\|^2 \, dt \geq \left[\int_0^T \|\hat{u}(t)\| \, dt \right]^2 \tag{F.12}$$

From (F.10) and (F.12), it follows that

$$T \int_0^T \|\hat{u}(t)\|^2 dt \geq \|\xi\|^2 \tag{F.13}$$

From (F.5) and (F.13), we conclude that

$$\hat{J} \geq J^* \tag{F.14}$$

Appendix G

DERIVATION OF (6.19)

To show the optimal control $u(\tau)$ of the system (6.14) is given by

$$u(x_1,x_2) = x_2 + \frac{36\, x_{10}}{T^3}\left[\frac{2x_1 - x_{10}}{2x_2 - \frac{6}{T}x_{10}}\right] \qquad (G.1)$$

we form the following function

$$H = u(e_2 + u) + \lambda_1(\dot{e}_1 - u) + \lambda_2(\dot{e}_2 + e_2 + u) \qquad (G.2)$$

where λ_1 and λ_2 are arbitrary multipliers. Then

$$\frac{\partial H}{\partial \dot{u}} = 0 \quad , \quad \frac{\partial H}{\partial u} = e_2 + 2u - \lambda_1 + \lambda_2$$

$$\frac{\partial H}{\partial \dot{e}_1} = \lambda_1 \quad , \quad \frac{\partial H}{\partial e_1} = 0 \qquad (G.3)$$

$$\frac{\partial H}{\partial \dot{e}_2} = \lambda_2 \quad , \quad \frac{\partial H}{\partial e_2} = u + \lambda_2$$

Using Euler's equation we have

$$u = \frac{1}{2}(\lambda_1 - \lambda_2 - e_2)$$

$$\lambda_1 = 2c_1$$

$$\dot{\lambda}_2 = \frac{1}{2}\lambda_1 + \frac{1}{2}\lambda_2 - \frac{1}{2}e_2 \quad , \qquad (G.4)$$

where C_1 is a constant of integration. In terms of components together with the first two equations of (G.4), equation (6.16) can be written as

$$\dot{e}_1 = -\frac{1}{2} e_2 + \frac{1}{2} \lambda_2 - C_1$$

$$\dot{e}_2 = -\frac{1}{2} e_2 + \frac{1}{2} \lambda_2 + C_1 \qquad \text{(G.5)}$$

Combining the last two equations of (G.4) and the second equation of (G.5) one finds

$$\lambda_2 = e_2 + 2C_1\tau + C_2 \quad , \qquad \text{(G.6)}$$

where C_2 is an arbitrary constant. Upon elimination of λ_2 between the second equation of (G.5) and (G.6) we obtain

$$e_2(\tau) = C_1 \frac{\tau^2}{2} + \left(\frac{C_2}{2} - C_1 \right) \tau + C_3 - C_2 \qquad \text{(G.7)}$$

Combining (G.7) with the first equation of (G.5) the expression of e_1 is found to be

$$e_1(\tau) = -\frac{1}{2} C_1 \frac{\tau^3}{3} - \frac{1}{2} C_2 \frac{\tau^2}{2} + \left(C_1 + \frac{C_2}{2} - C_3\right)\tau + C_4 \qquad \text{(G.8)}$$

The constants of integration C_1, C_2, C_3, and C_4 can be determined by applying the boundary conditions of the system, i.e.,

$$e_1(0) = x_{10}$$

$$e_2(0) = 0$$

$$e_1(T) = 0$$

$$e_2(t) = 0 \quad . \qquad \text{(G.9)}$$

As a result the constants are found to be

$$C_1 = -\frac{12}{T^3} x_{10}$$

$$C_2 = C_3 = \frac{12(T-2)}{T^3} x_{10}$$

$$C_4 = x_{10} \quad . \tag{G.10}$$

Substituting equation (G.10) into (G.7) and (G.8) gives

$$e_1(\tau) = x_{10} \left[2\left(\frac{\tau}{T}\right)^3 + \left(\frac{6}{T} - 3\right)\tau^2 - \frac{6}{T^2}\tau + 1\right] \tag{G.11}$$

and

$$e_2(\tau) = \frac{6}{T} x_{10} \left[-\left(\frac{\tau}{T}\right)^2 + \left(\frac{\tau}{T}\right)\right] \quad .$$

From equations (G.4) , (G.6) and (G.11), $u(\tau)$ is found to be

$$u(\tau) = \frac{6x_{10}}{T^2} \left[\frac{1}{T}\tau^2 + \left(\frac{2}{T} - 1\right)\tau - 1\right] \tag{G.12}$$

From equations (6.15) and (G.11) the expressions for x_1 and x_2 are given by

$$x_1(\tau) = x_{10} \left[1 - 3\left(\frac{\tau}{T}\right)^2 + 2\left(\frac{\tau}{T}\right)^3\right] ,$$

and

$$x_2(\tau) = \frac{6}{T} x_{10} \left[\left(\frac{\tau}{T}\right)^2 - \frac{\tau}{T}\right] \quad . \tag{G.13}$$

Upon elimination of τ between (G.12) and (G.13) one finds

$$u(x_1, x_2) = x_2 + \frac{36}{T^3} x_{10} \left[\frac{2x_1 - x_{10}}{2x_2 - \frac{6}{T} x_{10}}\right]$$

which is what we want to show.

BIBLIOGRAPHY

1. Pontryagin, L.S., Boltyanskii, V.G., Gamkrelidze, R.V. and Mischenko, E.F.: "The Mathematical Theory of Optimal Processes", (book) Interscience Publishers, 1962.

2. Bellman, R.E. and Dreyfus, S.E.: "Applied Dynamic Programming", (book) Princeton Univ. Press, N.J., 1962.

3. Desver, C.A.: "The Bang Bang Serve Problem Treated by Variational Techniques", Information and Control, vol. 2, pp. 333-348; Dec. 1959.

4. Friedland, B.: "The Structure of Optimal Control Systems", Jour. of Basic Engr., (Trans. A.S.M.E.), vol. 83, 1961.

5. Lee, E.B.: "Mathematical Aspects of Synthesis of Linear Minimum Response Time Controllers", Trans. P.G.A.C., vol. AC-5, pp. 283-289; Sept. 1960.

6. Kelley, H.J.: "Gradient Theory of Optimal Flight Paths", ARS Jour. 30, pp. 947-954, 1960.

7. Kelley, H.J.: "Optimization Techniques", (book) edited by G. Leitmann, Academic Press, N.Y., 1962.

8. Bryson, A.E. and Denham, W.F.: "A Steepest Ascend Method for Solving Optimum Programming Problems", Am. Soc. Mech. Eng. Conf. on Applied Mechanics, Chicago, Ill., June, 1961.

9. Neustadt, L.: "Synthesizing Time Optimal Controls", Jour. of Math. Anal. and Appl. 1, pp. 484-493, 1960.

10. Ho, Y.C.: "A Successive Approximation Technique for Optimal Control Systems Subject to Input Saturation", Jour. Basic Eng., vol. 84, pp. 33-40, 1962.

11. Ho, Y.C.: "A Computational Technique for Optimal Control Problems with State Variable Constraint", Jour. Math. Anal. and Appl. 5, pp. 216-224, 1962.

12. Krasovskii, N.N.: "On the Theory of Optimum Regulation, Automation and Remote Control", vol. 18, pp. 1005-1016; Nov. 1957.

13. Kulikowski, R.: "Optimizing Processes and Synthesis of Optimizing Automatic Control Systems with Nonlinear Invariable Elements", Proc. I.F.A.C. Moscow Conf. (Butterworth Scientific Publications) pp. 473-477, 1961.

14. Kranc, G.M. and Sarachik, P.E.: "An Application of Functional Analysis to Optimal Control System", Joint Automatic Control Conference, AIEE, 1962.

15. Bode, H.: "Network Analysis and Feedback Amplifier Design," (book) Van Nostrand, N.Y., 1950.

16. Truxal, J.G.: "Automatic Feedback Control System Synthesis", (book) McGraw Hill Book Co., N.Y., 1955.

17. Horowitz, I.M.: "Fundamental Theory of Automatic Linear Feedback Systems", IRE Trans. on A.C., vol. AC-4, pp. 5-19; Dec. 1959.

18. Chang, S.S.L.: "Synthesis of Optimum Control Systems", (book) McGraw-Hill Book Co., Inc., N.Y., 1961.

19. Huang, R.Y.: "The Sensitivity of the Poles of Linear Closed-Loop Systems". Trans. AIEE, part II, vol. 77, pp. 182-186; 1958.

20. Miller, K.S. and Murray, F.J.: "A Mathematical Basis for Error Analysis of Differential Analyzers", Jour. Math. Phys., vol. 32, nos. 2-3, July-Oct., 1953.

21. Zadeh, L.A.: "Time-Varying Networks, I", Proc. IRE., vol. 49, pp. 1488-1503; Oct. 1961.

22. Rosenbloom, A.: "Analysis of Linear Systems with Randomly Varying Parameters", Proc. Symp. on Information Networks, Poly. Institute of Brooklyn, 1954.

23. Samuels, J.C. and Eringen: "On Stochastic Linear Systems", Jour. Math. and Phys., vol. 38, pp. 83-103; July, 1959.

24. Samuels, J.C.: "On the Mean Square Stability of Random Linear Systems", IRE Trans. on Circuit Theory (Special suppl.) vol. CT-6, pp. 248-259; May, 1959.

25. Bertram, J.E. and Sarachik, P.E.: "Stability of Circuits with Randomly Time-Varying Parameters", IRE Trans. on Circuit Theory (special suppl.) vol. CT-6, pp. 260-270; May, 1959.

26. Kats, I.I. and Krasovskii, N.N.: "On Stability of a System with Random Parameters", PMM 24, pp. 809-823, 1960.

27. Dorato, P.: "On Sensitivity in Optimal Control Systems", IEEE Trans. on AC, pp. 256-257; July, 1963.

28. Gavrilovic, M., Petrovic, R., and Siljak, D.: "Adjoint Method in the Sensitivity Analysis of Optimal Systems", Jour. of Franklin Inst., vol. 276, no. 1, pp. 26-38, July, 1963.

29. Wonham, W.M.: "Stochastic Problems in Optimal Control," RIAS Report, Baltimore, Md., 1962.

30. Tomovic, R.: "Sensitivity Analysis of Dynamic Systems," (book) McGraw-Hill Book Co., N.Y., 1963.

31. Friedman, B.: "Principles and Techniques of Applied Mathematics", (book) John Wiley and Sons, Inc., N.Y., 1956.

32. Kantarovich, L.V.: "Functional Analysis and Applied Mathematics", Usp. Mat. Nauk., 3 (1948) pp. 89-185; Engl. Trans. National Bureau of Standards, 1955.

33. Balakrishnan, A.V.: "An Operator Theoretic Formulation of a Class of Control Problems and a Steepest Descent Method of Solution", Jour. SIAM on Control, Series A, vol. 1, no. 2, pp. 109-127, 1963.

34. Akhiezer, N.I. and Glazman, I.M.: "Theory of Linear Operators in Hilbert Space," (book) vol. 2, Frederick Ungar Publishing Co., N.Y., 1961.

35. Kolmogorov, A.N. and Fomin, S.V.: "Elements of the Theory of Functions and Functional Analysis", (book) vol. 1, Graylock Press, Rochester, N.Y. , 1957.

36. Taylor, A.E.: "Introduction to Functional Analysis", (book) John Wiley and Sons, Inc., N.Y., 1958.

37. Kolmogorov, A.N. and Fomin, S.V.: "Elements of the Theory of Functions and Functional Analysis", (book) vol. 2, Graylock Press, Albany, N.Y., 1961.

38. Athans, M., Falb, P. and Lacoss, R.: "On Optimal Control of Self-Adjoint Systems", 1963 Joint Automatic Control Conference, pp. 113-120, Minneapolis, Minn., June, 1963.

39. Athans, M. Falb, P., and Lacoss, R.: "Time-Fuel-and-Energy-Optimal Control of Nonlinear Norm-Invariant Systems", IEEE on AC., pp. 196-202; July, 1963.

40. Chuang, K.: Private Communication.

41. Gnedenko, B.V.: "The Theory of Probability", (book), Chelsea Publishing Co., N.Y., N.Y., 1962.

42. Davenport, W.B., Jr., and Root, W.L.: "Random Signals and Noise", (book) McGraw-Hill Book Company, Inc., N.Y., 1958.

43. Wang, M.C., and Uhlenbeck, G.E.: "On the Theory of Brownian Motion II", Review of Modern Physics, Vol. 17, nos. 2 and 3, April-July, 1945.

44. Elgolts, L.E.: "Differential Equations", (book) Hindustan Publishing Corp., Delhi, India, 1961.

45. Zadeh, L.A. and Desoer, C.A.: "Linear System Theory", (book) McGraw-Hill Book Co., Inc., N.Y., 1963.

46. Cramer, H.: "Mathematical Methods of Statistics", (book), Princton University Press., 1946.

47. Letov, A.M.: "Analytic Controller Design 1," Automatika Telemekhanika 21, no. 4 (1960). (Engl. trans. Automation and Remote Control, vol. 21, no. 4, pp. 303-306, 1960).

48. Hahn, S.H.: "Theory and Application of Liapunov's Direct Method", (book) Prentice-Hall, Inc., Englewood Cliffs, N.J., 1963.

49. Lebedev, A.A.: "The Problem of Stability in a Finite Interval of Time", Prikladnaja Matematika i Mekhanika 18, 75-94, 1954. (English Summary may be found in Mathematical Review 16, p. 132, 1955).

50. Chzhan, S.: "On Stability of Motion for a Finite Interval of Time", Prikladnaja Matemtika i Mekhanika 23, pp. 230-238, 1959.

51. Barrett, J.F.: "Application of Kolmogorov's Equations to Randomly Disturbed Automatic Control Systems", International Federation of Automatic Control, Moscow, USSR, Vol. 2, pp. 724-733, 1960.

52. La Salle, J., and Lefschetz, S.: "Stability by Liapunov's Direct Method", (book) Academic Press, N.Y., N.Y., 1961.

53. Schwartz, N.: "Information Transmission, Modulation, and Noise", (book) McGraw-Hill Book Company, Inc., N.Y., N.Y., 1959.

54. Flugge-Lotz, I. and Marbach, H., : "The Optimal Control of Some Attitude Control Systems for Different Performance Criteria", 1960, Joint Automatic Control Conference, N.Y., N.Y.

55. Kaplan, W.: "Ordinary Differential Equations", (book) Addison-Wesley Publishing Co., Inc. Reading, Mass., 1958.

56. Hall A.: "Statistical Theory with Engineering Applications", (book) John Wiley and Sons, Inc., N.Y., N.Y., 1952.

57. Pontryagin, L.S.: "Ordinary Differntial Equations", (book) Addison-Wesley Publishing Co., Inc., Reading, Mass., 1962.

58. Kalman, R.E.: "New Methods in Wiener Filtering Theory", Proc. of the First Symposium on Engineering Applications of Random Function Theory and Probability edited by J.L. Bogdanoff and F. Kozin (book), John Wiley & Sons, Inc., 1963.