Prepared for:

December 1982

U.S. Department of Energy

Assistant Secretary, Conservation
and Renewable Energy
Office of Energy Systems Research
Washington, D.C. 20585

DOE/ET/29038--1

DE84 013140

**RESEARCH PROGRESS IN DYNAMIC SECURITY
ASSESSMENT**

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Prepared by:

The Analytic Sciences Corporation
One Jacob Way
Reading, MA 01867
Under Contract No. DE-AC01-79ET29038

NOTICE
PORTIONS OF THIS REPORT ARE ILLEGIBLE
It has been reproduced from the best
available copy to permit the broadest
possible availability.

tes United States United States
Energy Department of Energy
s United States United States
artment of Energy Departmen
nited States United States U
Department of Energy Depart
States **United States** United
Energy **Department of Energy**

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

ABSTRACT

This report documents the progress achieved during a three-year research effort in power system security assessment. The areas discussed are: power system modeling, state estimation, structure decomposition, state forecasting, clustering and security measure development. A detailed dynamic model of a multi-machine power system has been developed. The model includes the effects of stochastic load variations and possible transmission line and/or generator outages. A process state estimator was developed to estimate the long-term dynamic behavior of the power system. The algorithm is in a form identical to the extended Kalman filter but has a modified process noise driving term. A two-stage structure estimation technique was proposed for identifying the power system network configuration from available measurements. Two approaches to structure decomposition were investigated. A time-scale decomposition of the system equations, based on a singular perturbation approach, was evaluated using a detailed model of a generating system. This approach to system order reduction was extended to the process state estimation equations. Spatial decomposition was examined by applying an optimal network decomposition technique to a 39-bus test system. Stochastic approximation based approaches to estimator simplification were examined. Explicit expressions were obtained for the evolution of the first and second moments of the system state which take into account the possibility of changes in system network configuration. Clustering criteria were developed for use in reducing the number of system disturbances that must be considered. Research into security measures proceeded in three directions. The first area involves viewing the security assessment problem as a hyperplane crossing problem for a stochastic process. An approximate expression for the hyperplane crossing probability in terms of the moments of the expected number of crossings of the hyperplane was developed. An alternate method of analyzing this problem is based on a partial differential equation of Dynkin which yields the moments of the first crossing time. Upper and lower bounds on the first and second moments were developed which are easier to compute than the solution to Dynkin's equation. The second approach examined the stability of an unforced linear system where the system coefficients are subject to future jumps. This led to a sufficient condition for almost sure asymptotic stability. The third area of research has led to the formulation of a security measure suitable for on-line assessment of transient stability. This stability measure involves computation of the probability that a Lyapunov function will exceed a critical value in a specified future time interval.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

	Page No.		Page No.
ABSTRACT	ii	7. LONG-TERM SECURITY MEASURES	41
List of Figures	iv	7.1 Introduction	41
List of Tables	iv	7.2 Surface Crossing Problems	41
1. INTRODUCTION	1	7.2.1 Formulation of the Level Crossing Problem	42
1.1 Background	1	7.2.2 Formulation of the Surface Crossing Problem	43
1.2 Overview of the Research Effort	1	7.2.3 Higher Moments and the Probability of Level Crossing	44
1.3 Organization of the Report	2	7.2.4 Dynkin's Equation Approach to the Surface Crossing Problem	46
2. STATE ESTIMATION	3	7.3 Stability Results for Stochastic Differential Equations with Discrete Jump Coefficients	52
2.1 Introduction and Problem Definition	3	7.3.1 Background	52
2.2 General Equations for the Minimum Mean Square Error State Estimate	4	7.3.2 Equations with Jump Coefficients	53
2.3 State Estimation for a Scalar Process	5	7.4 Summary	55
2.3.1 State Estimation with Wiener Process Noise	5	8. A SECURITY MEASURE FOR TRANSIENT STABILITY	56
2.3.2 State Estimation with Generalized Poisson Process Noise	7	8.1 A Stochastic Level Crossing Problem	56
2.3.3 State Estimation with Combined Process Noise	10	8.2 A Simplified Level Crossing Problem	57
2.4 State Estimation for a Vector Valued Process	10	8.3 Phase I Computations	59
2.5 Summary	13	8.4 Phase II Computations	60
3. STRUCTURE ESTIMATION	14	8.5 Summary	60
3.1 Introduction	14	9. CLUSTERING	61
3.2 Optimal Nonlinear Filter for Continuous Measurements	14	9.1 Jump Disturbance Models	61
3.2.1 Problem Statement	14	9.2 Basic Clustering Techniques and Approach	61
3.2.2 Optimal Nonlinear Filter	16	9.3 Transient Stability Clustering	62
3.2.3 Application of the Filtering Equation to the Random Telegraph Wave	16	9.4 Dynamic Stability Clustering	63
3.3 Optimal Nonlinear Filter for Discrete Measurements	17	9.5 Steady-State Stability Clustering	63
3.4 Optimal Linear Filter for Continuous Measurements	19	10. DEMONSTRATION OF SIMPLIFIED TRANSIENT SECURITY MEASURE	65
3.4.1 Problem Formulation	19	10.1 Introduction	65
3.4.2 Optimal Linear Filter	20	10.2 Reduced-Order Network	65
3.5 Hypothesis Testing for Structure Estimation	20	10.3 Prediction of the Critical Clearing Time	66
3.5.1 Introduction	20	10.4 Simulation Studies	67
3.5.2 Multiple Model Estimation Equations	21	10.4.1 Approximations Based on Use of Machine Accelerations	67
3.6 Summary	23	10.4.2 Plots of Selected Cases	68
4. TIME-SCALE DECOMPOSITION	24	10.5 Summary	68
4.1 Introduction	24	11. SUMMARY AND RECOMMENDATIONS	71
4.2 Singularly Perturbed Estimation	24	11.1 Summary	71
4.3 Singularly Perturbed Model of a Generating System	26	11.2 The Application of Stochastic Control Concepts to Integrate Security Assessment and Security Enhancement	72
4.4 Reduced Order Estimator Design	28	11.3 Recommendations	73
5. ESTIMATOR SIMPLIFICATION FOR TRANSIENT DYNAMICS	31	APPENDIX A POWER SYSTEM MODELING	74
5.1 Introduction	31	APPENDIX B AN INTEGRATED STATE SPACE FORMULATION OF THE POWER SYSTEM EQUATIONS	92
5.2 Review of Previous Work on Transient State Estimation	31	APPENDIX C POISSON COUNTING PROCESSES	97
5.3 Nonlinear Estimation Using Observers	31	APPENDIX D DECOMPOSITION OF THE RANDOM TELEGRAPH WAVE	99
5.4 Stochastic Approximation Approach	32	APPENDIX E SOLUTION OF SINGULARLY PERTURBED KALMAN FILTER EQUATIONS	103
5.5 Nonlinear Estimation by Stochastic Approximation in Absence of Process Noise	33	APPENDIX F A BOILER-TURBINE-EXCITER- GENERATOR MODEL	108
6. STATE FORECASTING	35	APPENDIX G PROCESSES HAVING DIFFERENTIABLE SAMPLE PATHS	112
6.1 Introduction	35	REFERENCES	118
6.2 Evaluation of the First Moment	35	EXECUTIVE SUMMARY	
6.3 Evaluation of the Second Moment	38		
6.4 Recursive Updating of the State Forecast	40		

LIST OF FIGURES

<u>Figure No.</u>		<u>Page No.</u>
1.2-1	Functional Structure of TASC's Approach to Dynamic Security Assessment	2
3.5-1	Hypothesis Testing Approach to System Structure Determination	21
4.3-1	Plot of BTEG Model Eigenvalues	27
6.2-1	Block Diagram Solution for the First Moment	38
7.1-1	Hypothetical Evolution of a State Variable	41
7.2-1	First Crossing Times	42
7.2-2	Geometry of the Surface Crossing Problem	43
7.2-3	Hyperplane Geometry	43
10.1-1	WSCC 9-Bus Test System	65
10.2-1	Y-Matrix for Pre-fault WSCC Network Before Reduction	66
10.4-1	Simulation Results for Case 1	69
10.4-2	Simulation Results for Case 4	69
10.4-3	Simulation Results for Case 5	70
11.1-1	Functional Structure of TASC's Approach to Dynamic Security Assessment	72

LIST OF TABLES

<u>Table No.</u>		<u>Page No.</u>
2.3-1	Continuous-Discrete Extended Kalman Filter Equations for a Scalar Process	10
2.4-1	Continuous-Discrete Extended Kalman Filter Equations for a Vector Process	13
3.4-1	Continuous Kalman Filter Equations for a Scalar Process	20
4.3-1	Eigenvalues of the BTEG Model	26
4.3-2	Relation Between States and Eigenvalues	27
4.3-3	Eigenvalues of the (7:23) Partition	28
4.3-4	Eigenvalues of the (19:11) Partition	28
10.2-1	Reduced Admittance Matrices	66
10.2-2	Network Parameters for Lyapunov Function	67
10.4-1	Summary of Test Cases	67
10.4-2	Summary of Machine Acceleration During Fault-On Period	68

1.1 BACKGROUND

Large scale electric power systems are continually increasing in complexity and interdependence. In the United States for example, there are approximately 6000 generating units and 800,000 km of bulk power transmission lines which, under the control of more than 100 independent control centers, supply electric energy in response to the demands of over 80 million customers. Under normal operating conditions, this complex system provides reliable electric service. However, there are times when an emergency develops which disrupts the balance between demand and supply (or load and generation) and requires intervention (either human or automatic) to restore a portion of the system to normal operation. It is unlikely that a single event can trigger such an emergency. Rather, it is due to a sequence of relatively small events which tend to "weaken" part of the system, making it susceptible to an otherwise moderate disturbance. In other words, the power system is operating at low level of security and is thus more vulnerable to disturbances than it would be if it were in its normal robust state.

Digital computers have made possible on-line monitoring of power system security to identify circumstances where preventive control action must be taken to avoid hazardous operating conditions. This is formerly regarded as the security assessment function in power systems. In general terms, security assessment can be thought of as an algorithm which utilizes knowledge of the current operating state of the system, forecasts the future evolution of the system, assesses the likelihood of a security violation within a specified future time-period, and determines whether preventive control action should be taken. These functions imply that the security assessment algorithm must depend upon

- Knowledge of system dynamics
- Measurement of system variables
- Models for measurement uncertainty and system disturbances
- Well-defined criteria for security.

Because of power system complexity, present day security assessment algorithms have limitations dictated by available computer capability. In particular, they almost all use static models of the system, and evaluate security with respect to steady state operating limits on system variables. Furthermore, little attempt has been made to deal systematically with system disturbances, particularly sudden random events -- line outages, load changes, faults, etc.

The rapid pace of technological development in computer hardware, both with respect to hardware miniaturization and to various forms of parallel processing, implies that the affordable computational power available to the utilities will increase by orders of magnitude in the near future. At the same time, the narrowing gap between demand and supply of electric energy will require that operating safety margins be known more accurately in real time to avoid frequent load disruptions. This report describes the results of a three-year research program, sponsored by the Department of Energy, to develop an improved methodology for assessing power system security, which can take advantage of future computer technology, and

which can be made available in time to meet future utility needs.

Dynamic security assessment for electric power systems is a very complex and difficult problem. The magnitude of the difficulty arises from a compounding of the following issues:

- The system of interest is very large scale, dynamic, and highly nonlinear
- The inputs of interest are multidimensional statistical distributions of random events (jump processes) and perturbations
- Security assessment involves evaluating the robustness of the system over a future time-interval, which requires potentially complex and time-consuming forecasts of system behavior for all possible occurrences of discrete and random events
- Security assessment must be performed on-line.

Even assuming a substantial increase in computational capabilities over the next decade, a useful solution to the stated problem must, of necessity, incorporate efficient and effective techniques for handling the various components of the problem; i.e., dynamic state estimation, statistical modeling of events and perturbations, stochastic forecasting, and stochastic contingency evaluation. A key element of this solution is the formulation of an overall approach which is sufficiently general to incorporate all relevant components of the problem and, at the same time, is sufficiently structured to develop specific practical mathematical solution techniques.

1.2 OVERVIEW OF THE RESEARCH EFFORT

The functional structure of TASC's approach to dynamic security assessment is illustrated in Fig. 1.2-1. It incorporates the following major elements:

- Stochastic filtering of measurements to obtain an estimate of the present state of the system
- Forecasting of the system state to provide an indication of future security needs
- Establishment of security measures that are functionals of the state estimate and models of jump processes (outages, etc.), and process noise (small random fluctuation of loads, etc.)
- Off-line classification of possible jump disturbances into a smaller number of aggregates or clusters.

Throughout this effort, primary emphasis was placed on the development of power system security measures that are simultaneously useful and suitable for on-line evaluation. Subordinate to the development of security measures were the specification of computationally efficient methods for estimating the current state of the power system and forecasting its possible future evolution. These latter functions require detailed models of power system dynamic response to stochastic disturbances. In the course of this research effort, progress was made in the areas

Figure 1.2-1 Functional Structure of TASC's Approach to Dynamic Security Assessment

of power system modeling, state estimation, structure decomposition, state forecasting, clustering and security measure development. A capsule summary of this work is contained below.

A detailed dynamic model of multimachine power system was developed for use throughout this effort. The model includes a state variable description of the boiler-turbine, generator and excitation system. Stochastic components of the model consist of unpredictable load disturbances and transmission line and/or generator outages. A measurement model characterizing power system observations was also completed.

A state estimator for estimating the long-term dynamic behavior of the power system was specified. Since measurements of both the system state and network configuration are available, the long-term state estimator was partitioned into two separate (but not independent) algorithms -- the process state estimator and the structure (or network configuration) estimator. Various approaches to estimator simplification were examined including nonlinear observers and stochastic approximation based techniques.

In the area of structure decomposition, two different approaches were investigated -- spatial decomposition and time-scale decomposition. Spatial network decomposition consists of partitioning the power system into a number of subsystems, each with a small number of generators and loads. A computer program was developed to implement an optimal network decomposition technique. Time-scale decomposition, or the separation of the system into two subsystems, one containing states with slow dynamics and the other containing states with fast dynamics was applied to the equations characterizing the process state estimator.

Research in the area of state forecasting has resulted in the development of recursive expressions for calculating the first and second moments of the system state at future times.

Clustering techniques for reducing the large number of possible power system disturbances to a more manageable number were reviewed and a set of clustering criteria for the security measures of interest were proposed.

Research into security measures proceeded in three directions. The first area involves viewing the security assessment problem as a surface crossing problem for a stochastic process. The surface in this formulation separates the normal operating region from the emergency region. Approximating the surface by a hyperplane allowed the application of scalar level crossing theory to obtain moments of the expected number of crossings. An approximate expression for the

hyperplane crossing probability in terms of these moments was also developed. An alternate method of analyzing this problem utilizes a partial differential equation of Dynkin which yields the moments of the first crossing time. Upper and lower bounds on the first and second moments were developed. These bounds provide estimates of the moments which are much simpler to compute than the solution to Dynkin's equation. The second approach examined the long-term stability properties of the power system from the viewpoint of stability theory. The third area of research has led to the formulation of a security measure suitable for on-line assessment of transient stability. To alleviate the computational burden, the security measure is divided into a real-time and a non-real-time component. Simplifying approximations to estimate the kinetic energy increase during fault-on periods were evaluated along with other simplifications made to further reduce complexity.

1.3 ORGANIZATION OF THE REPORT

Chapters 2 through 5 document various aspects of the generalized state estimation effort. Dynamic state estimation is presented in two parts -- process state estimation and structure estimation. Chapter 2 contains the development of estimation equations for the power system state assuming that the network configuration is known. The problem of power system structure determination is addressed in Chapter 3. The large size of typical systems leads to algorithms with a prohibitively large number of state variables. One technique for reducing filter dimensionality involves decomposing the system into its "slow" and "fast" modes. This time-scale decomposition is discussed in Chapter 4. Another method for reducing the computational burden is to deliberately simplify the estimation algorithm. Various approaches to estimator simplification using stochastic approximation techniques are examined in Chapter 5.

State forecasting is covered in Chapter 6. Recursive expressions are developed for forecasting the state of the system subject to changes in transmission network topology.

Security measures are discussed in Chapters 7 through 10. Long-term security measures are the subject of Chapter 7. The major emphasis in this chapter is on obtaining approximate solutions to a surface crossing formulation of the security assessment problem. On a much shorter time-scale, power system transient stability becomes important. Chapter 8 contains the development of a Lyapunov-based transient stability measure suitability for on-line implementation. In evaluating the transient stability measure it is necessary to take into account network faults. The utility of this measure is enhanced if the number of faults that must be considered can be kept to a minimum. The cluster analysis techniques described in Chapter 9 provide a basis for reducing the large number of possible disturbances to a more manageable set. Finally, Chapter 10 contains a demonstration of the concepts underlying the transient stability measure.

The work documented in this report is summarized in Chapter 11. In addition, this chapter contains a short discussion of security enhancement and some recommendations for further research. The main body of the text is supplemented by seven appendices covering various technical aspects of the effort including detailed power system models referred to throughout the report.

2.1 INTRODUCTION AND PROBLEM DEFINITION

Inherent in the security assessment problem is the need for an accurate estimate of the current state of the power system. Over the last ten years, considerable attention has been devoted to the development of computer algorithms for on-line state estimation (see for example, Refs. 1 through 5). Two major approaches have been pursued -- static state estimation and tracking state estimation. The static state estimation problem is typically formulated as a weighted least squares optimization (and is usually implemented with a recursive algorithm). In the tracking state formulation it is assumed that the system state vector ($\underline{x}$) satisfies the discrete time equation

$$\underline{x}_{k+1} = \underline{x}_k + \underline{w}_k \quad (2.1-1)$$

with

$$E \{ \underline{w}_1 \underline{w}_1^T \} = Q \delta_{1j} \quad (2.1-2)$$

where δ_{ij} is the Kronecker delta. The tracking state estimator is based on the standard Kalman filter algorithm (Ref. 6) for the linear dynamical system model of Eq. 2.1-1 and a linearized measurement equation, viz ,

$$\underline{y}_k = H \underline{x}_k + \underline{v}_k \quad (2.1-3)$$

In both of these formulations, the system state vector consists of the complex voltages (typically expressed as magnitude and angle) at the system nodes (buses).

One drawback to these techniques is the use of a simplified dynamic model for the evolution of the system state. (Implicit in the least squares formulation is the assumption that $\underline{x}_{k+1} = \underline{x}_k$, while Eq. 2.1-1 is a very simplified representative of the time-variation.) If a realistic (and computationally tractable) dynamic model is available to describe the evolution of the power system state then it is reasonable to expect better estimation accuracy (i.e., lower mean square error) than that currently achievable with the simplified dynamics embodied in Eq. 2.1-1. Appendices A and B of this report contain a detailed description of such a model. This stochastic model of power system dynamics may be written in the form

$$\begin{aligned} d\underline{x}(t) = & \underline{f}[t, \underline{x}(t), \underline{z}(t), \underline{u}(t)] dt \\ & + G[t, \underline{x}(t), \underline{z}(t), \underline{u}(t)] d\underline{n} \end{aligned} \quad (2.1-4)$$

where

- $\underline{f}$ = vector function
- G = matrix function
- $\underline{x}$ = the continuous state representing voltages, generator variables, etc.
- $\underline{z}$ = a jump process representing faults, outages, etc
- $\underline{u}$ = applied controls
- $d\underline{n}$ = process disturbance

For purposes of state estimation, measurement data

$$\underline{y}_k = \underline{h}_k[\underline{x}(t_k)] + \underline{v}_k$$

$$\underline{g}_k = \underline{c}_k[\underline{z}(t_k)] + \underline{d}_k \quad (2.1-5)$$

are available at discrete instants in time. The variable $\underline{y}_k$ represents meter or instrument readings and $\underline{g}_k$ represents indications of breaker positions; $\underline{v}_k$, $\underline{d}_k$ the set of instrumentation errors.

The discussion of state estimation contained in this chapter is directed toward the development of estimation equations suitable for use in a load tracking or quasi-steady-state regime of power system operation. In this context, the long term power system equations (or a reduced order set thereof) described in Appendix A are most appropriate.

One of the fundamental issues associated with the state estimator is the manner in which the system "structure" is determined. The jump process $\underline{z}(t)$ is included in the system model to characterize sudden, unpredictable changes in the configuration (or "structure") of the power system. In the absence of the measurements $\underline{g}_k$ (see Eq. 2.1-5), the state of the process $\underline{z}(t)$ would have to be inferred from the performance of an estimator which processed only the data $\underline{y}_k$. Commonly used techniques to accomplish this employ hypothesis testing ideas to select the "best" $\underline{z}^1(t)$ from a finite collection of possible values, $\underline{z}^1(t), \dots, \underline{z}^k(t)$; this approach falls under the general heading of multiple model estimation. Under even the simplest assumptions for the jump process $\underline{z}(t)$, the number of hypothesis which must be compared rapidly becomes intractable for a power system.

It is assumed here that sufficient, noisy measurements $\underline{g}_k$ are available to decouple the problem of state estimation from the estimation of system structure. The advantages of this approach include:

- Direct estimation of system structure*
- Reduced complexity in the state estimator.

The discussion of state estimation contained in the remainder of this chapter refers only to estimation of the process $\underline{x}(t)$, assuming that the system structure is known.

Up to this point, nothing has been said about what constitutes a good estimate of the system state. It will be assumed in the following discussions of state estimation that an estimate ($\hat{\underline{x}}$) is optimal in the sense of minimum mean square error (MMSE). Let $\underline{Y}_k$ denote the set of all measurements $\underline{y}_1, 1=1, 2, \dots, k$ (i.e., from t_1 to t_k). It is well known (for example Refs. 7 and 8) that the MMSE estimate of the process $\underline{x}(t)$ at the time t_k is

$$\hat{\underline{x}}(t_k) = E \{ \underline{x}(t_k) | \underline{Y}_k \} \quad (2.1-6)$$

where $E \{ \cdot | \cdot \}$ denotes conditional expectation.

*Although measurements of structure are available, the problem of structure estimation is nontrivial. Structure identification is discussed in the next chapter.

The remaining sections in this chapter contain details on the evaluation of Eq. 2.1-6. Section 2.2 is devoted to the general equations for obtaining the MMSE estimate. The estimation of a scalar process, which illustrates the properties of the optimal estimator without being notationally cumbersome, is treated in Section 2.3. This case is used to motivate the choice of a suboptimal estimation algorithm. The suboptimal estimation equations for the general (vector) case are then developed in Section 2.4.

2.2 GENERAL EQUATIONS FOR THE MINIMUM MEAN SQUARE ERROR STATE ESTIMATE

In the previous section it was noted that the conditional expectation leads to the minimum mean square error estimate. This expectation can be evaluated in general if the conditional density function $p_{\underline{x}}(\underline{x}, t | Y_k)$ is known.* The problem formulation considered herein involves a continuous-time system model (stochastic differential equation) and discrete-time observations. Consequently, the equations expressing the evolution of the conditional density function can be broken into two distinct parts: propagation between measurements and an update at the observation time.

In the following discussion it is assumed that the $n \times 1$ dimensional state vector, $\underline{x}(t)$, is a Markov process governed by the stochastic differential equation

$$d\underline{x} = \underline{f}(\underline{x}, t) dt + G(\underline{x}, t) d\underline{w} + \Gamma(\underline{x}, t) d\underline{n}_p \quad (2.2-1)$$

where $d\underline{w}$ is an n -dimensional Wiener process with covariance matrix $Q(t)$ defined by

$$E \{d\underline{w}(t) d\underline{w}^T(t)\} = Q(t) dt \quad (2.2-2)$$

and $d\underline{n}_p$ is an n -dimensional generalized (or compound) Poisson process (see Appendix C). These two driving terms represent load perturbations to the system which may be decomposed into an almost surely continuous part ($d\underline{w}(t)$) and an almost surely discontinuous part ($d\underline{n}_p(t)$). Additional details on disturbance modeling may be found in Appendices A and B.

Let $p_{\underline{x}}(\underline{x}, t | \underline{x}(s))$ be the transition probability density function for the process $\underline{x}(t)$ defined by Eq. 2.2-1. Then, $p_{\underline{x}}(\cdot)$ satisfies the following partial differential equation (Ref. 8)

$$\begin{aligned} \frac{\partial p_{\underline{x}}}{\partial t} = & - \sum_{i=1}^n \frac{\partial (f_i p_{\underline{x}})}{\partial x_i} + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 (GQGT)_{ij} p_{\underline{x}}}{\partial x_i \partial x_j} \\ & + \sum_{i=1}^n \lambda_i(t) \left[-p_{\underline{x}} + r_{\underline{x}, a_i} \right] \end{aligned} \quad (2.2-3)$$

where $r_{\underline{x}, a_i}$ is defined as

*For notational convenience, the argument of $p(\cdot)$ is used to define the type of density function and the subscript denotes the random variable to which it applies.

$$\begin{aligned} & \left(\frac{1}{2\pi} \right)^n \int_{\underline{\xi}} \int_{\underline{v}} e^{-j\underline{\xi}^T \underline{u}} e^{j\underline{\xi}^T \underline{v}} M_{a_i}(\underline{\xi}_1 \Gamma_{1i}(\underline{v}, t) + \dots + \underline{\xi}_n \Gamma_{ni}(\underline{v}, t)) \\ & \times p_{\underline{x}}(\underline{v}, t | \underline{x}(s)) d\underline{v} d\underline{\xi} \end{aligned} \quad (2.2-4)$$

and where

$$M_{a_i}(\underline{\xi}_i) = E \{ e^{j\underline{\xi}_i^T \underline{a}_i} \} \quad (2.2-5)$$

is the characteristic function of the marks associated with the generalized Poisson process.

Solution of Eq. 2.2-3 gives the time evolution of the transition density function $p_{\underline{x}}(\underline{x}, t | \underline{x}(s))$. It should be noted that Eq. 2.2-3 contains two well known special cases. If $d\underline{n}_p$ is identically zero, Eq. 2.2-3 reduces to the well known Fokker-Planck equation

$$\frac{\partial p_{\underline{x}}}{\partial t} = \mathcal{L}^+ p_{\underline{x}} \quad (2.2-6)$$

where $\mathcal{L}^+$ is the Fokker-Planck forward operator defined by

$$\mathcal{L}^+ = - \sum_{i=1}^n \frac{\partial}{\partial x_i} (f_i) + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 (GQGT)_{ij}}{\partial x_i \partial x_j} \quad (2.2-7)$$

If $d\underline{w}$ is absent, Eq. 2.2-3 becomes the Feller-Kolmogorov equation

$$\frac{\partial p_{\underline{x}}}{\partial t} = - \sum_{i=1}^n \frac{\partial}{\partial x_i} (f_i p_{\underline{x}}) + \sum_{i=1}^n \lambda_i(t) [r_{\underline{x}, a_i} - p_{\underline{x}}] \quad (2.2-8)$$

It remains to determine the behavior of the conditional density function at the observation. Since the observations occur only at discrete times it is sufficient to characterize the conditional density after the observation (i.e., at time t_k^+) in terms of its value just prior to the observation (at time t_k^-) and the information contained in the measurement. The subsequent discussion follows Jazwinski (Ref. 7). Note that

$$p_{\underline{x}}(\underline{x}, t_k^+ | Y_k^-) = p_{\underline{x}}(\underline{x}, t_k | Y_{k-1}) \quad (2.2-9)$$

since no information has been received between time t_{k-1} and t_k^- . The key to the following derivation is Bayes' rule

$$p_{\underline{x}}(\underline{x}, t_k | Y_k) = \frac{p_Y(Y_k | \underline{x}, t_k) p_{\underline{x}}(\underline{x}, t_k | Y_{k-1})}{p_Y(Y_k)} \quad (2.2-10)$$

Since $p_{\underline{x}}(\underline{x}, t_k | Y_k) = p_{\underline{x}}(\underline{x}, t_k | Y_k, Y_{k-1})$, Eq 2 2-10 can be rewritten as

$$p_{\underline{x}}(\underline{x}, t_k | Y_k) = \frac{p_Y(Y_k, Y_{k-1} | \underline{x}, t_k) p_{\underline{x}}(\underline{x}, t_k | Y_{k-1})}{p_Y(Y_k, Y_{k-1})}$$

$$= \frac{p_Y(Y_k | \underline{x}, t_k, Y_{k-1}) p_{\underline{x}}(\underline{x}, t_k | Y_{k-1})}{p_Y(Y_k | Y_{k-1})}$$

(2.2-11)

If it is assumed that the observation error v_k (from Eq 2 1-5) is uncorrelated with previous values, then

$$p_Y(Y_k | \underline{x}_k, Y_{k-1}) = p_Y(Y_k | \underline{x}_k) \quad (2.2-12)$$

and

$$p_Y(Y_k | Y_{k-1}) = \int p_Y(Y_k | \underline{x}) p_{\underline{x}}(\underline{x}, t_k | Y_{k-1}) d\underline{x} \quad (2.2-13)$$

Inserting Eqs. 2 2-12 and 2 2-13 into Eq 2 2-11 gives the desired result, viz ,

$$p_{\underline{x}}(\underline{x}, t_k | Y_k) = \frac{p_Y(Y_k | \underline{x}) p_{\underline{x}}(\underline{x}, t_k | Y_{k-1})}{\int p_Y(Y_k | \underline{u}) p_{\underline{x}}(\underline{u}, t_k | Y_{k-1}) d\underline{u}} \quad (2.2-14)$$

Equations 2.2-3 and 2.2-14 completely specify the conditional density function $p_{\underline{x}}(\underline{x}, t | Y_k)$ from which, the optimal minimum mean square error estimate is obtained as

$$\hat{\underline{x}}(t_k) = \int \underline{u} p_{\underline{x}}(\underline{u}, t_k | Y_k) d\underline{u} \quad (2.2-15)$$

These equations formally solve the state estimation problem. However, it is an extremely difficult numerical problem to solve Eq. 2.2-3 when $n = 1$ or 2 let alone for any realistic value of n associated with power system state estimation. (A discussion of numerical solutions to this problem and their inherent difficulties may be found in Refs 9 and 10 among others.) Consequently, some form of approximation must be employed to yield a tractable solution. In the next section, the case of a scalar process is examined to identify the characteristics of the state estimator and provide a rationale for making simplifying assumptions.

2.3 STATE ESTIMATION FOR A SCALAR PROCESS

The emphasis in this section is on the characterization of solutions to the state estimation problem for the scalar process

$$dx = f(x, t) dt + g(x, t) dw + \gamma(x, t) dn_p \quad (2 3-1)$$

$$y_k = h[x(t_k)] + v_k \quad (2 3-2)$$

with

w = a Wiener process with covariance
 $E\{dw^2\} = q_w(t) dt$ (Note $q_w(t)$
 is used only in this section and
 should not be confused with q_k of
 Eq 2 1-5.)

n_p = a compound Poisson process with
 rate $\lambda(t)$ and mark space charac-
 terized by the density function
 $p_a(a)$

v_k = a discrete, uncorrelated sequence
 of measurement errors with zero
 mean and covariance, r_k

If $dn_p \equiv 0$ in Eq 2 3-1, it would reduce to the form of a (by now) standard Itô stochastic differential equation which has been extensively studied, cf, Refs. 7, 8 and 11. In contrast, the purely Poisson driven stochastic differential equation (i.e., Eq. 2.3-1 with $dw \equiv 0$) has received less attention in the engineering literature. The next section contains a review of state estimation for the dynamical system of Eq 2.3-1 driven only by a Wiener process. The structure of an approximate nonlinear filter is developed and discussed. In Section 2.3 2, the case of a Poisson driven Markov process is examined. Finally, a solution for the complete dynamical system of Eqs 2 3-1 and 2 3-2 is stated in Section 2 3.3

2 3.1 State Estimation with Wiener Process Noise

Assume the state of a dynamical system is governed by the Itô stochastic differential equation

$$dx = f(x, t) dt + g(x, t) dw \quad (2 3-3)$$

with observations of the state occurring at discrete times $t_1, 1=1, 2, \dots$, according to Eq. 2.3-2. As discussed previously, the optimal MMSE estimate of the state, $\hat{x}(t)$, is given by the conditional expectation, viz.,

$$\hat{x}(t) = E\{x(t) | Y_k\} = \int u(t) p_x(u, t | Y_k) du \quad (2 3-4)$$

The conditional density function is obtained by solving the Fokker-Planck equation (cf. Eqs 2.2-6 and 2.2-7) between measurements and updating it at an observation with Eq. 2.2-14.

Rather than attempt to evaluate the entire conditional density function, a more practical approach is to assume that the density function is adequately characterized by its low order moments and to derive equations which describe the temporal evolution of these moments. The following arguments are taken from Jazwinski (Ref. 7). Let $\phi(x)$ be a scalar function of the state x and assume that the second derivative of ϕ with respect to x is continuous. Define

$$\hat{\phi}_k(x(t)) = E\{\phi(x(t)) | Y_k\} = \int \phi(u) p_x(u, t | Y_k) du \quad (2.3-5)$$

to be the conditional expectation of $\phi(x)$. By the Itô differentiation rule,

$$d\phi(x) = \frac{\partial \phi}{\partial x} dx + \frac{1}{2} \frac{\partial^2 \phi}{\partial x^2} q_w(t) g^2(x, t) dt \quad (2.3-6)$$

Substituting Eq. 2.3-3 into 2.3-6 gives

$$d\phi(x) = \left[\frac{\partial \phi(x)}{\partial x} f(x, t) + \frac{1}{2} \frac{\partial^2 \phi(x)}{\partial x^2} q_w(t) g^2(x, t) \right] dt + \frac{\partial \phi(x)}{\partial x} g(x, t) dw \quad (2.3-7)$$

Taking the expectation of both sides of Eq. 2.3-7 yields,

$$d E\{\phi(x)\} = E\left\{ \frac{\partial \phi(x)}{\partial x} f(x, t) \right\} + \frac{1}{2} q_w(t) E\{g^2(x, t) \frac{\partial^2 \phi(x)}{\partial x^2}\} dt \quad (2.3-8)$$

since

$$E\left\{ \frac{\partial \phi(x)}{\partial x} g(x, t) dw \right\} = 0$$

Of primary interest in the sequel are the first two moments; consequently let $\phi(x) = x(t)$ in Eq. 2.3-8. Then, using the notation $\hat{(\cdot)}$ to denote $E\{\cdot\}$,

$$d\hat{x} = \hat{f}(x, t) dt \quad (2.3-9)$$

describes the time evolution (between observations) of the MMSE estimate of x . Similarly, for $\phi(x) = x^2(t)$,

$$d E\{x^2(t)\} = 2 \hat{x} \hat{f}(x, t) dt + q_w(t) \hat{g}^2(x, t) dt \quad (2.3-10)$$

Furthermore, note that

$$d(\hat{x}^2) = 2 \hat{x} \hat{f}(x, t) dt \quad (2.3-11)$$

which follows from Eq. 2.3-9. Let $p(t)$ denote the second central moment, i.e.,

$$\begin{aligned} p(t) &= E\{[x(t) - \hat{x}(t)]^2\} \\ &= E\{x^2(t)\} - \hat{x}^2(t) \end{aligned} \quad (2.3-12)$$

Then, using Eqs. 2.3-10 and 2.3-11,

$$\begin{aligned} dp(t) &= d E\{x^2(t)\} - d(\hat{x}^2) \\ &= 2[\hat{x}f(x, t) - \hat{x} \hat{f}(x, t)] dt \\ &\quad + q_w(t) \hat{g}^2(x, t) dt \end{aligned} \quad (2.3-13)$$

Equations 2.3-9 and 2.3-13 are the desired expressions for the temporal evolution (between observations) of the first two moments of the state x . It remains to determine the behavior of $\hat{x}$ and p at an observation

To this end, define

$$\hat{\phi}_k^+(x(t_k)) = E_k^+\{\phi(x(t_k))\} = \int \phi(u) p_x(u, t_k | Y_k^+) du \quad (2.3-14)$$

$$\hat{\phi}_k^-(x(t_k)) = E_k^-\{\phi(x(t_k))\} = \int \phi(u) p_x(u, t_k | Y_k^-) du \quad (2.3-15)$$

where $-$ and $+$ denote the times just before and after the observation, respectively. Now, multiplying Eq. 2.2-14 by $\phi(x)$ and integrating gives

$$\hat{\phi}_k^+(x(t_k)) = \frac{E_k^-\{\phi(x(t_k)) p(y_k | x(t_k))\}}{E_k^-\{p(y_k | x(t_k))\}} \quad (2.3-16)$$

Letting $\phi(x(t)) = x(t)$ and $x^2(t)$ results in expressions for the conditional mean and variance (respectively) at an observation

This completes the development of the general equations for the MMSE estimate. Although Eqs. 2.3-9, 2.3-13 and 2.3-16 appear to solve the estimation problem, in reality they do not. All of the expectations involved in these equations require the conditional density function for their evaluation. Thus, all lower order moments are functions of higher order moments. To obtain a computationally feasible estimator, some approximations must be made.

Two distinct assumptions will be made:

- The third and higher order moments of the conditional density function are negligible
- The nonlinear system equations can be adequately represented by a Taylor series expansion truncated after the first order term. The functions are expanded about the current state estimate $\hat{x}(t)$ *

The specific expansions are

$$f(x, t) = f(\hat{x}, t) + \frac{\partial f(\hat{x}, t)}{\partial x} (x - \hat{x}) + \dots \quad (2.3-17)$$

$$h(x) = h(\hat{x}) + \frac{\partial h(\hat{x})}{\partial x} (x - \hat{x}) + \dots \quad (2.3-18)$$

$$g(x, t) = g(\hat{x}, t) + \frac{\partial g(\hat{x}, t)}{\partial x} (x - \hat{x}) + \dots \quad (2.3-19)$$

*In what follows, $\hat{x}$ will not denote the true conditional expectation but rather the approximation that is obtained with the above assumptions.

Substituting Eq 2.3-17 into Eq 2.3-9 and taking the conditional expectation gives

$$\begin{aligned} d\hat{x} &= E\{f(\hat{x}, t) + \frac{\partial f(\hat{x}, t)}{\partial x} (x - \hat{x})\} dt \\ &= f(\hat{x}, t) dt \end{aligned} \quad (2.3-20)$$

Similarly, substituting Eqs 2.3-17 and 2.3-19 into Eq 2.3-13 yields to first order

$$\begin{aligned} dp(t) &= 2 [E\{xf(\hat{x}, t)\} \\ &+ E\{x \frac{\partial f(\hat{x}, t)}{\partial x} (x - \hat{x})\} - \hat{x}f(\hat{x}, t)] dt \\ &+ q_w(t) E\{[g(\hat{x}, t) + \frac{\partial g(\hat{x}, t)}{\partial x} (x - \hat{x})]^2\} dt \\ &= 2 \frac{\partial f(\hat{x}, t)}{\partial x} [E\{x^2\} - \hat{x}^2] dt + q_w(t) g^2(\hat{x}, t) dt \end{aligned} \quad (2.3-21)$$

Since $E\{x^2\} - \hat{x}^2 = p(t)$, Eq. 2.3-21 may be written

$$dp/dt = 2 F(\hat{x}, t) p(t) + q_w(t) g^2(\hat{x}, t) \quad (2.3-22)$$

where

$$F(\hat{x}, t) = \left. \frac{\partial f(x, t)}{\partial x} \right|_{x=\hat{x}}$$

Expressions describing how moments of the conditional density change at the observation are more complicated to obtain. Note that since the third and higher order moments are assumed to be negligible, it is necessary to treat only the conditional mean and variance. Jazwinski (Ref. 7, Section 9.4) starts with Eq. 2.3-16 and develops general equations for the behavior of the first and second moments at an observation. To limit the computational burden of the associated filter, he assumes that the mean and variance can be expressed as an affine function of the observations y_k , viz.,

$$\hat{x}_k(+) = a + B(y_k - \hat{y}_k(-)) \quad (2.3-23)$$

$$p_k(+) = C + D(y_k - \hat{y}_k(-)) \quad (2.3-24)$$

where the subscript k denotes time t_k and the $-$ and $+$ refer to times just before and after the observation, respectively. In Eqs. 2.3-23 and 2.3-24,

$$\hat{y}_k(-) \equiv E\{y_k | Y_{k-1}\} \quad (2.3-25)$$

The coefficients are

$$a = \hat{x}_k(-) \quad (2.3-26)$$

$$B = (\hat{x}h - \hat{x}\hat{h}) \{(\hat{h}-\hat{h})^2 + r\}^{-1} \quad (2.3-27)$$

$$C = p_k(-) - (\hat{x}h - \hat{x}\hat{h}) \{(\hat{h}-\hat{h})^2 + r\}^{-1} (\hat{h}x - \hat{h}\hat{x}) \quad (2.3-28)$$

$$\begin{aligned} D &= \{[(x-\hat{x})^2 h] - 2B[(\hat{h}-\hat{h})(x-\hat{x})h] \\ &+ B^2\{[(\hat{h}-\hat{h})^2 h] + r\hat{h} - \hat{C}h\} \{(\hat{h}-\hat{h})^2 + r\}^{-1} \end{aligned} \quad (2.3-29)$$

with $(\hat{\cdot})$ denoting conditional expectation using $p_x(x_k | Y_{k-1})$ and in Eqs 2.3-27 through 2.3-29, $x = x_k(-)$ and h is evaluated at $x_k(-)$. Equations 2.3-27 through 2.3-29 can be simplified using Eq. 2.3-18 and explicitly evaluating the expectations.

Let $H_k(\hat{x}(-))$ denote $\frac{\partial h(x)}{\partial x} \Big|_{x=\hat{x}_k(-)}$ and $F_k(\hat{x}(-))$ denote

$$\frac{\partial f(x, t)}{\partial x} \Big|_{x=\hat{x}_k(-), t=t_k} \quad \text{Then after some manipula-}$$

tion, the state estimate and variance update equations may be written as

$$\hat{x}_k(+) = \hat{x}_k(-) + K_k [y_k - h_k(\hat{x}_k(-))] \quad (2.3-30)$$

$$p_k(+) = [1 - K_k H_k(\hat{x}_k(-))] p_k(-) \quad (2.3-31)$$

where

$$K_k = p_k(-) H_k(\hat{x}_k(-)) [H_k^2(\hat{x}_k(-)) p_k(-) + r_k]^{-1} \quad (2.3-32)$$

Equations 2.3-20, 2.3-22 and 2.3-30 through 2.3-32 are the first order linearized or extended Kalman filter equations. The structure of these equations are the same as for the standard Kalman filter associated with a linear dynamical system. The major difference is that the filter gain (K_k) is now dependent upon the state estimate and cannot be precomputed. This results in an increased on-line computational burden. It is possible to retain higher order terms in the Taylor series expansion of the system nonlinearities and higher order moments of the density function (see Refs. 6 through 8 for examples). Such an approach would lead to computationally more complex filter structures which, when coupled with the dimensionality of the power system state estimation problem, would become impractical. The extended Kalman filter may not exhibit estimation errors as small as those obtained with higher order approximations, but relative to computational requirements it is a logical choice.

2.3.2 State Estimation with Generalized Poisson Process Noise

In this section it will be assumed that the system state is governed by the stochastic differential equation

$$dx = f(x, t) dt + \gamma(x, t) dn_p \quad (2.3-33)$$

with n_p a generalized (or compound) Poisson process. The goal of this section is to develop expressions for the statistics of the process $x(t)$. Once the statistical behavior of $x(t)$ is characterized, the general

filtering theory discussed above can be applied. As in the previous section, the scalar process of Eq. 2.3-33 will be discussed first with the natural generalization to the multidimensional problem summarized in Section 4.4. The major reference for this section is Snyder (Ref. 12).

Before discussing solutions to Eq. 2.3-33, it is beneficial to examine the characteristics of the driving term dn_p . Let $\zeta(t)$ denote a (scalar) stochastic process. The characteristic function associated with this process is defined as

$$M_\zeta(v, t) = E\{\exp[j v \zeta(t)]\} \quad (2.3-34)$$

The usefulness of the characteristic function lies in the fact that

$$E\{\zeta^n(t)\} = \frac{1}{j^n} \frac{\partial^n M_\zeta(v, t)}{\partial v^n} \Big|_{v=0} \quad (2.3-35)$$

which affords a method for determining the statistics of the process $\zeta(t)$. In the sequel, the probability density function $p_a(a)$ which characterizes the mark space of the generalized Poisson process will be assumed to be Gaussian with zero mean and variance σ_a^2 , i.e.,

$$p_a(a) = \frac{1}{\sqrt{2\pi} \sigma_a} e^{-a^2/2\sigma_a^2} \quad (2.3-36)$$

It is straightforward to show (see for example, Refs. 8 or 12) that the characteristic function for a generalized Poisson process, with intensity (or rate) $\lambda(t)$ and amplitude density given by Eq. (2.3-36), is

$$M_{n_p}(v, t) = \exp\left\{\int_0^t \lambda(\sigma) d\sigma \left[\exp\left(-\frac{1}{2} v^2 \sigma_a^2\right) - 1\right]\right\} \quad (2.3-37)$$

Calculation of the first four moments of the process $n_p(t)$ yields:

$$E\{n_p(t)\} = 0 \quad (2.3-38)$$

$$E\{n_p^2(t)\} = \sigma_a^2 \int_0^t \lambda(\sigma) d\sigma \quad (2.3-39)$$

$$E\{n_p^3(t)\} = 0 \quad (2.3-40)$$

$$E\{n_p^4(t)\} = 3 \sigma_a^4 \left\{ \left[\int_0^t \lambda(\sigma) d\sigma \right]^2 + \int_0^t \lambda(\sigma) d\sigma \right\} \quad (2.3-41)$$

Note that since the process has zero mean, the second and fourth moments are also the central moments, viz.,

$$\mu_2 = E\{[n_p - E\{n_p\}]^2\} = E\{n_p^2\} \quad (2.3-42)$$

$$\mu_4 = E\{[n_p - E\{n_p\}]^4\} = E\{n_p^4\} \quad (2.3-43)$$

For comparison, consider the Wiener process $w(t)$ of Section 2.3.1. The characteristic function is (for σ_w^2 independent of t),

$$M_w(v, t) = \exp\left[-\frac{1}{2} v^2 \sigma_w^2 t\right] \quad (2.3-44)$$

and the first four moments are

$$E\{w(t)\} = 0 \quad (2.3-45)$$

$$E\{w^2(t)\} = \sigma_w^2 t = q_w t \quad (2.3-46)$$

$$E\{w^3(t)\} = 0 \quad (2.3-47)$$

$$E\{w^4(t)\} = 3 \sigma_w^4 t^2 = 3 q_w^2 t^2 \quad (2.3-48)$$

The similarity between the moments for these two processes is evident. Assume that $\lambda(t) = \lambda_0$ then, Eqs. 2.3-39 and 2.3-41 become

$$E\{n_p^2(t)\} = \sigma_a^2 \lambda_0 t \quad (2.3-49)$$

$$E\{n_p^4(t)\} = 3 \sigma_a^4 [(\lambda_0 t)^2 + (\lambda_0 t)] \quad (2.3-50)$$

Consider the special case of $(\lambda_0 t)$ large, or equivalently, $(\lambda_0 t)^2 \gg (\lambda_0 t)$. Under this condition, the fourth moment (Eq. 2.3-39) is approximately

$$E\{n_p^4(t)\} = 3 \sigma_a^4 (\lambda_0 t)^2 \quad (2.3-51)$$

Denoting $\sigma_a^2 \lambda_0$ by σ_p^2 allows Eqs. 2.3-48 and 2.3-50 to take the same form as Eq. 2.3-45 and 2.3-47. Consequently, although the process $n_p(t)$ is not Gaussian, its low order moments approach those of a Gaussian distribution as $(\lambda_0 t)$ becomes large. This result agrees with intuition, since for λ_0 large, the marks, which are assumed to be independent Gaussian random variables, occur at a high rate (i.e., closely spaced in time). The near Gaussian density function associated with the assumed generalized Poisson process model will allow the use of the same approximations as in the previous section.

The generalized Poisson process considered here has a convenient representation; specifically, for the conditions stated above, $n_p(t)$ can be represented as

$$n_p(t) = \int_0^t \int_{-\infty}^{\infty} \alpha M(d\sigma, d\alpha) \quad (2.3-52)$$

where the integral on α is over the mark space. Under the assumption that the marks are Gaussian random variables, the mark space becomes all of R^1 . The measure, $M(d\sigma, d\alpha)$, has a convenient interpretation as the number of jumps of the Poisson process occurring in the time interval $(\sigma, \sigma+d\sigma)$ which have marks in the interval $(\alpha, \alpha+d\alpha)$. Note that for $T \subset [0, \infty)$ and $B \subset \mathcal{R} = (-\infty, \infty)$,

$$E\{M(T, B)\} = \int_T \int_B \lambda(\sigma) d\sigma p_a(\alpha) d\alpha \quad (2.3-53)$$

Using this notation, Eq. 2.3-33 may be written as

$$dx(t) = f(x, t) dt + \int_{-\infty}^{\infty} \gamma(x, t) \alpha M(dt, d\alpha) \quad (2.3-54)$$

Equation 2.3-54 defines the change in the state [i.e., $dx(t) = x(t+dt) - x(t)$] from time t to time $t+dt$. If no points of the generalized Poisson process occur in this interval, $M(dt, d\alpha) = 0$ and $dx = f(x, t) dt$. If a point with mark α does occur, then $dx = f(x, t) dt + \gamma(x, t)\alpha$. Depending upon the specific value of the mark, the process $x(t)$ will exhibit discontinuities of varying magnitude at the jump occurrence times.

Paralleling the approach taken in the discussion of the Wiener process driving term in Section 2.3.1, a stochastic differential equation for an arbitrary function of the process $x(t)$ in Eq. 2.3-33 will be used to derive differential equations for the moments of $x(t)$. Let $\psi(x)$ be a bounded, scalar function of x with continuous first derivative. Then,

$$d\psi(x(t)) = \frac{\partial \psi}{\partial x} f(x, t) dt + \int_{-\infty}^{\infty} \{\psi[x(t) + \gamma(x, t)\alpha] - \psi[x(t)]\} M(dt, d\alpha) \quad (2.3-55)$$

The derivation of this differentiation rule is in Snyder (Ref. 12, pp. 199-202). Equation 2.3-55 is analogous to the Itô differentiation rule (Eq. 2.3-6) of Section 2.3.1.

As in the case of Wiener process noise, only the first two moments are of interest. Let $\psi(x(t)) = x(t)$ in Eq. 2.3-55 and take the expected value of both sides to get

$$\begin{aligned} dE\{x(t)\} &= E\{f(x, t)\} dt + E\left\{\int_{-\infty}^{\infty} \gamma(x, t) \alpha M(dt, d\alpha)\right\} \\ &= E\{f(x, t)\} dt + E\{\gamma(x, t)\} E\{\alpha\} \lambda(t) dt \end{aligned} \quad (2.3-56)$$

since

$$E\left\{\int_{-\infty}^{\infty} \gamma(x, t) \alpha M(dt, d\alpha)\right\} = E\{\gamma(x, t)\alpha\} \lambda(t) dt$$

and $\gamma(x, t)$ is independent of α . Similarly, for $\psi[x(t)] = x^2(t)$,

$$\begin{aligned} dE\{x^2(t)\} &= 2 E\{xf(x, t)\} + E\left\{\int_{-\infty}^{\infty} 2x\gamma(x, t)\alpha \right. \\ &\quad \left. + \gamma^2(x, t)\alpha^2 M(dt, d\alpha)\right\} \\ &= 2 E\{xf(x, t)\} + 2 E\{x\gamma(x, t)\alpha\} \lambda(t) dt \\ &\quad + E\{\gamma^2(x, t)\alpha^2\} \lambda(t) dt \end{aligned} \quad (2.3-57)$$

Expanding the nonlinear functions $f(x, t)$ and $\gamma(x, t)$ in a Taylor series (cf. Eq. 2.3-17) and substituting into Eq. 2.3-56 gives

$$d\hat{x}(t) = f(\hat{x}, t) dt + \gamma(\hat{x}, t) E\{\alpha\} \lambda(t) dt = f(\hat{x}, t) dt \quad (2.3-58)$$

with $(\hat{\cdot})$ denoting $E\{\cdot\}$. As in the previous section, let $p(t) = E\{x^2\} - \hat{x}^2$, then

$$\begin{aligned} dp(t) &= [2 F(\hat{x}, t) p(t) + 2 E\{x\alpha\} \gamma(\hat{x}, t) \lambda(t) \\ &\quad + \gamma^2(\hat{x}, t) E\{\alpha^2\} \lambda(t)] dt \end{aligned} \quad (2.3-59)$$

where

$$F(\hat{x}, t) = \left. \frac{\partial f(x, t)}{\partial x} \right|_{x=\hat{x}}$$

Using the properties of the assumed mark space and the fact that $E\{x\alpha\} = E\{x\} E\{\alpha\} = 0$, Eq. 2.3-58 becomes

$$dp/dt = 2 F(\hat{x}, t) p(t) + \gamma^2(\hat{x}, t) \sigma_a^2 \lambda(t) \quad (2.3-60)$$

Equations 2.3-58 and 2.3-60 characterize the approximate time evolution of the first two moments of the Poisson driven Markov process of Eq. 2.3-33. The remaining part of the filtering algorithm specifies the effect of the observations on these moments. The relevant equations are identical to Eqs. 2.3-30 through 2.3-32 and will not be repeated here.

The generalized Poisson process model assumed in this section is sufficient for the power system problem. The relatively "nice" structure of the moments of $n_p(t)$ is a direct consequence of the Gaussian assumption for $p_a(a)$. The derivation leading up to Eqs. 2.3-56 and 2.3-57 is completely general and no restrictions are placed on the probability density function defined on the mark space. If $p_a(a)$ were chosen arbitrarily, the validity of dropping the third and higher order moments would be questionable.

2.3.3 State Estimation with Combined Process Noise

The scalar state and measurement equations are

$$dx = f(x, t) dt + g(x, t) dw + \gamma(x, t) dn_p \quad (2.3-61)$$

$$y_k = h(x(t_k)) + v_k \quad (2.3-62)$$

where the individual terms have been defined previously. Assuming that the two noise processes, w and n_p , are independent allows the previously developed solutions to the linearized state estimation problem to be superimposed. The resulting equations will be referred to as the continuous-discrete extended Kalman filter, although this terminology strictly applies to a system model with $n_p=0$.

The state estimator for the dynamic system model of Eqs. 2.3-61 and 2.3-62 is characterized by a set of differential equations which describe the time-evolution of approximations to the first two moments of the conditional density function; and a set of difference equations which reflect the information obtained from the observations. These equations are summarized in Table 2.3-1 for the scalar process. An extension to vector-valued processes is contained in Section 2.4.

TABLE 2.3-1
CONTINUOUS-DISCRETE EXTENDED KALMAN
FILTER EQUATIONS FOR A SCALAR PROCESS

SYSTEM MODEL	$dx = f(x, t) dt + g(x, t) dw + \gamma(x, t) dn_p$
MEASUREMENT MODEL	$y_k = h(x(t_k)) + v_k$
STATE ESTIMATE PROPAGATION	$d\hat{x} = f(\hat{x}, t) dt$
ERROR COVARIANCE PROPAGATION	$dp/dt = 2 F(\hat{x}, t) p(t) + g^2(\hat{x}, t) q_w(t) + \gamma^2(\hat{x}, t) \sigma_{n_p}^2 \lambda(t)$
STATE ESTIMATE UPDATE	$\hat{x}_k(+) = \hat{x}_k(-) + K_k [y_k - h_k(\hat{x}_k(-))]$
ERROR COVARIANCE UPDATE	$p_k(+) = [1 - K_k H_k(\hat{x}_k(-))] p_k(-)$
GAIN	$K_k = p_k(-) H_k(\hat{x}_k(-)) / [H_k^2(\hat{x}_k(-)) p_k(-) + r_k]$ $F(\hat{x}, t) = \frac{\partial f(\hat{x}, t)}{\partial \hat{x}} \Big _{\hat{x}=\hat{x}}$
DEFINITIONS	$H_k(\hat{x}) = \frac{\partial h(\hat{x})}{\partial \hat{x}} \Big _{\hat{x}=\hat{x}, t=t_k}$

which describes the state of a power system. Observations of the system are made at times t_i , $i=1, 2, \dots$ and are of the form

$$y_k = h(x(t_k)) + v_k \quad (2.4-2)$$

The following conditions are assumed to hold for the elements of Eqs. 2.4-1 and 2.4-2:

- $\underline{x}(t)$ is the n -dimensional system state vector
- $f(\underline{x}, t)$ is continuous in t and globally Lipschitz continuous in $\underline{x}$; furthermore, $\frac{\partial f}{\partial x_j}$ is also assumed to be continuous in t and globally Lipschitz continuous in $\underline{x}$
- $\underline{w}(t)$ is an m -dimensional Wiener process with zero-mean and covariance $E\{d\underline{w}(t)d\underline{w}^T(t)\} = Q(t) dt$
- $G(\underline{x}, t)$ is an $n \times m$ dimensional matrix, Hölder continuous in t and Lipschitz continuous in $\underline{x}$ with $G(\underline{x}, t) G^T(\underline{x}, t)$ strictly positive definite
- $\underline{n}_p(t)$ is an n -dimensional generalized Poisson process with $n_{pi}(t)$ and $n_{pj}(t)$ independent for all i and j . For each $i=1, 2, \dots, n$, the intensity of the underlying Poisson point process is $\lambda_i(t)$ and the mark space is characterized by the zero-mean Gaussian probability density function
$$p_{a_i}(a) = \frac{1}{\sqrt{2\pi} \sigma_{a_i}} \exp \{-a^2 / 2\sigma_{a_i}^2\}$$
- $\Gamma(\underline{x}, t)$ is a $n \times m$ dimensional matrix, holder continuous in t and Lipschitz continuous in $\underline{x}$
- $\underline{y}_i$ is the ℓ -dimensional measurement vector at time t_k
- $h(\underline{x}, t)$ is continuous, as is $\frac{\partial h_i}{\partial x_j}$
- $\underline{v}_k$ is an uncorrelated measurement noise sequence with zero mean and covariance matrix R_k .

2.4 STATE ESTIMATION FOR A VECTOR VALUED PROCESS

Consider the vector stochastic differential equation

$$d\underline{x} = \underline{f}(\underline{x}, t) dt + G(\underline{x}, t) d\underline{w} + \Gamma(\underline{x}, t) d\underline{n}_p \quad (2.4-1)$$

The basic approach of Section 2.3 will be followed here -- a differentiation rule will be presented and used to obtain equations describing the time-evolution of the first and second moments of the conditional density function of $\underline{x}(t)$, where $\underline{x}(t)$ satisfies Eq. 2.4-1. The nonlinear terms will be expanded in a Taylor series and truncated after the linear term to obtain approximate propagation equations. The effect of the observations on these moments are determined from an application of Bayes' rule.

Let $\phi(\underline{x})$ be a twice continuously differentiable, scalar valued function of $\underline{x}(t)$. Then, by the generalized Itô differentiation rule (cf. Ref. 13),

$$d\phi(\underline{x}(t)) = d\eta_c(t) + d\eta_d(t) \quad (2.4-3)$$

where

$$\begin{aligned} d\eta_c(t) &= \left(\frac{\partial \phi}{\partial \underline{x}} \right)^T [d\underline{x}(t) - d\underline{n}_p(t)] \\ &+ \frac{1}{2} \text{tr} [G(\underline{x}, t) Q(t) G^T(\underline{x}, t) \frac{\partial^2 \phi}{\partial \underline{x}^2}] dt \\ &= \left(\frac{\partial \phi}{\partial \underline{x}} \right)^T \underline{f}(\underline{x}, t) dt \\ &+ \frac{1}{2} \text{tr} [G(\underline{x}, t) Q(t) G^T(\underline{x}, t) \frac{\partial^2 \phi}{\partial \underline{x}^2}] dt \\ &+ \left(\frac{\partial \phi}{\partial \underline{x}} \right)^T G(\underline{x}, t) d\underline{w}(t) \end{aligned} \quad (2.4-4)$$

$$\begin{aligned} d\eta_d(t) &= \sum_{k=1}^m \int_{\mathcal{A}} [\phi(\underline{x}(t) + \underline{\gamma}^k(\underline{x}, t) \alpha_k) \\ &- \phi(\underline{x}(t))] M_k(dt, d\alpha_k) \end{aligned} \quad (2.4-5)$$

and

$$\frac{\partial \phi}{\partial \underline{x}} \text{ is the gradient of } \phi = \left(\frac{\partial \phi}{\partial x_1}, \dots, \frac{\partial \phi}{\partial x_n} \right)^T$$

$$\frac{\partial^2 \phi}{\partial \underline{x}^2} \text{ is the matrix of second partial derivatives of } \phi \text{ with } (ij)^{\text{th}} \text{ entry } \frac{\partial^2 \phi}{\partial x_i \partial x_j}$$

$\text{tr}(\cdot)$ denotes the trace operator

$\underline{\gamma}^k(\underline{x}, t)$ is the k^{th} column of $\Gamma(\underline{x}, t)$.

Evaluation of the first two moments of the vector process of Eq. 2.4-1 will be done in terms of components. Thus, let $\phi(\underline{x}(t)) = x_i(t)$ then,

$$\begin{aligned} dx_i(t) &= f_i(\underline{x}, t) dt + [G(\underline{x}, t) d\underline{w}(t)]_i \\ &+ \sum_{k=1}^m \int_{-\infty}^{\infty} \gamma_{i1}^k(\underline{x}, t) \alpha_k M(dt, d\alpha_k) \end{aligned} \quad (2.4-6)$$

and taking the expected value gives

$$d\hat{x}_i(t) = [\hat{f}_i(\underline{x}, t) + \sum_{k=1}^m \hat{\gamma}_{i1}^k(\underline{x}, t) \lambda_k(t) E\{\alpha_k\}] dt \quad (2.4-7)$$

Since $E\{\alpha_k\} = 0$ for all k , the first moment of $\underline{x}(t)$ satisfies

$$d\hat{\underline{x}}(t) = \hat{\underline{f}}(\underline{x}, t) dt \quad (2.4-8)$$

To calculate the second moment define $\phi_{ij}(\underline{x}(t)) = \underline{e}_i^T [\underline{x}(t) \underline{x}^T(t)] \underline{e}_j$ where $\underline{e}_i$ is a n -dimensional vector with the i^{th} element equal to one and all other elements zero; specifically, $\phi_{ij}(\underline{x}(t)) = x_i(t) x_j(t)$. From Eq. 2.4-3,

$$\begin{aligned} d\phi_{ij}(\underline{x}(t)) &= [(\underline{e}_i \underline{e}_j^T + \underline{e}_j \underline{e}_i^T) \underline{x}(t)]^T [\underline{f}(\underline{x}, t) dt \\ &+ G(\underline{x}, t) d\underline{w}(t)] \\ &+ \frac{1}{2} \text{tr} [\underline{e}_i^T G(\underline{x}, t) Q(t) G^T(\underline{x}, t) (2) \underline{e}_j] dt \\ &+ \sum_{k=1}^m \underline{e}_i^T \int_{-\infty}^{\infty} [\alpha_k \underline{\gamma}^k(\underline{x}, t) \underline{x}^T(t) \\ &+ \underline{x}(t) (\underline{\gamma}^k(\underline{x}, t))^T \alpha_k \\ &+ \alpha_k^2 \underline{\gamma}^k(\underline{x}, t) (\underline{\gamma}^k(\underline{x}, t))^T] M_k(dt, d\alpha_k) \underline{e}_j \end{aligned} \quad (2.4-9)$$

Taking the expected value of both sides of Eq. 2.4-9 and using the approximations

$$E\{G(\underline{x}, t)\} \approx G(\hat{\underline{x}}, t) \quad (2.4-10)$$

$$E\{\underline{\gamma}^k(\underline{x}, t)\} \approx \underline{\gamma}^k(\hat{\underline{x}}, t)$$

yields

$$\begin{aligned} d E\{x_i(t) x_j(t)\} &= E\{[(\underline{e}_i \underline{e}_j^T + \underline{e}_j \underline{e}_i^T) \underline{x}(t)]^T \underline{f}(\underline{x}, t)\} dt \\ &+ \text{tr} [\underline{e}_i^T G(\underline{x}, t) Q(t) G^T(\underline{x}, t) \underline{e}_j] dt \\ &+ \sum_{k=1}^m \underline{e}_i^T [E\{\alpha_k\} E\{\underline{\gamma}^k(\underline{x}, t) \underline{x}^T\} \\ &+ E\{\alpha_k\} E\{\underline{x} (\underline{\gamma}^k(\underline{x}, t))^T\} \\ &+ E\{\alpha_k^2\} \underline{\gamma}^k(\hat{\underline{x}}, t) (\underline{\gamma}^k(\hat{\underline{x}}, t))^T] \lambda_k(t) dt \underline{e}_j \end{aligned} \quad (2.4-11)$$

Under the assumption that $E\{\alpha_k\} = 0$ and since $\underline{e}_i^T \theta \underline{e}_j = \theta_{ij}$ for any $n \times n$ matrix θ , Eq. 2.4-11 may be simplified to

$$\begin{aligned} d E\{x_i(t)x_j(t)\} &= E\{[(\underline{e}_i \underline{e}_j^T + \underline{e}_j \underline{e}_i^T) \underline{x}(t)]^T \underline{f}(\underline{x}, t)\} dt \\ &+ \{(G(\underline{x}, t)Q(t)G^T(\underline{x}, t))_{ij} \\ &+ \sum_{k=1}^m [Y^k(\underline{x}, t)(Y^k(\underline{x}, t))^T]_{ij} \\ &\sigma_{a_k}^2 \lambda_k(t)\} dt \end{aligned} \quad (2.4-12)$$

where $\sigma_{a_k}^2 = E\{\alpha_k^2\}$. To simplify the notation let $\tilde{Q}_{ij}(t)$ denote the last term of Eq. 2.4-12. At this point it is convenient to introduce the linearization of $\underline{f}(\underline{x}, t)$, viz,

$$\underline{f}(\underline{x}, t) \sim \underline{f}(\underline{\hat{x}}, t) + F(\underline{\hat{x}}, t) (\underline{x} - \underline{\hat{x}}) + \dots \quad (2.4-13)$$

where

$F(\underline{\hat{x}}, t)$ is the $n \times n$ matrix

$$F(\underline{\hat{x}}, t) = \left[\frac{\partial f_i}{\partial x_j} \right]_{\underline{x}=\underline{\hat{x}}} \quad (2.4-14)$$

Substituting Eq. 2.4-13 into Eqs. 2.4-8 and 2.4-12 yields

$$d\underline{\hat{x}} = \underline{f}(\underline{\hat{x}}, t) dt \quad (2.4-15)$$

$$\begin{aligned} d E\{x_i(t)x_j(t)\} &= E\{[(\underline{e}_i \underline{e}_j^T + \underline{e}_j \underline{e}_i^T) \underline{x}(t)]^T [\underline{f}(\underline{\hat{x}}, t) \\ &+ F(\underline{\hat{x}}, t) (\underline{x} - \underline{\hat{x}})]\} dt + \tilde{Q}_{ij}(t) dt \end{aligned} \quad (2.4-16)$$

The covariance matrix, $P(t)$, is defined as

$$P(t) = E\{[\underline{x}(t) - \underline{\hat{x}}(t)][\underline{x}(t) - \underline{\hat{x}}(t)]^T\} \quad (2.4-17)$$

from which follows the differential form

$$dP(t) = E\{d[\underline{x}(t) \underline{x}^T(t)]\} - d[\underline{\hat{x}}(t) \underline{\hat{x}}^T(t)] \quad (2.4-18)$$

with $(ij)^{th}$ component

$$dP_{ij}(t) = E\{d[x_i(t) x_j(t)]\} - d[\hat{x}_i(t) \hat{x}_j(t)] \quad (2.4-19)$$

Using Eqs. 2.4-15 and 2.4-16 allows Eq. 2.4-19 to be written as

$$\begin{aligned} dP_{ij}(t) &= E\{[(\underline{e}_i \underline{e}_j^T + \underline{e}_j \underline{e}_i^T) \underline{x}(t)]^T [\underline{f}(\underline{\hat{x}}, t) \\ &+ F(\underline{\hat{x}}, t) (\underline{x} - \underline{\hat{x}})] \\ &- [f_i(\underline{\hat{x}}, t) \hat{x}_j(t) + \hat{x}_i(t) f_j(\underline{\hat{x}}, t)]\} dt \\ &+ \tilde{Q}_{ij}(t) dt \end{aligned} \quad (2.4-20)$$

Since $[(\underline{e}_i \underline{e}_j^T) \underline{x}(t)]^T \underline{\theta} = x_j(t) \theta_i$ for any n -vector $\underline{\theta}$, Eq. 2.4-20 can be simplified to

$$\begin{aligned} dP_{ij}(t) &= E\{x_j(t)f_i(\underline{\hat{x}}, t) + x_i(t)f_j(\underline{\hat{x}}, t) + [x_j(t)\underline{e}_i^T \\ &+ x_i(t)\underline{e}_j^T] F(\underline{\hat{x}}, t) (\underline{x} - \underline{\hat{x}}) - [f_i(\underline{\hat{x}}, t) \hat{x}_j(t) \\ &+ \hat{x}_i(t) f_j(\underline{\hat{x}}, t)]\} dt + \tilde{Q}_{ij}(t) \\ &= E\{[x_j(t) \underline{e}_i^T + x_i(t) \underline{e}_j^T] F(\underline{\hat{x}}, t) (\underline{x} - \underline{\hat{x}})\} dt \\ &+ \tilde{Q}_{ij}(t) dt \\ &= E\{x_j(t) \sum_{k=1}^n F_{ik}(\underline{\hat{x}}, t) (x_k - \hat{x}_k) \\ &+ x_i(t) \sum_{k=1}^n F_{jk}(\underline{\hat{x}}, t) (x_k - \hat{x}_k)\} dt \\ &+ \tilde{Q}_{ij}(t) dt \\ &= \left\{ \sum_{k=1}^n F_{ik}(\underline{\hat{x}}, t) [E\{x_k(t)x_j(t)\} - \hat{x}_k(t)\hat{x}_j(t)] \right. \\ &+ \sum_{k=1}^n F_{jk}(\underline{\hat{x}}, t) [E\{x_i(t)x_k(t)\} \\ &- \hat{x}_i(t)\hat{x}_k(t)] \left. \right\} dt + \tilde{Q}_{ij}(t) dt \\ &= \sum_{k=1}^n [F_{ik}(\underline{\hat{x}}, t) P_{kj} + F_{jk}(\underline{\hat{x}}, t) P_{ik}(t)] dt \\ &+ \tilde{Q}_{ij}(t) dt \end{aligned} \quad (2.4-21)$$

which in matrix notation is

$$\begin{aligned} d\mathbf{P}(t)/dt &= \mathbf{F}(\hat{\mathbf{x}}, t) \mathbf{P}(t) + \mathbf{P}(t) \mathbf{F}^T(\hat{\mathbf{x}}, t) \\ &+ \mathbf{G}(\hat{\mathbf{x}}, t) \mathbf{Q}(t) \mathbf{G}^T(\hat{\mathbf{x}}, t) + \Sigma_p(t) \end{aligned} \quad (2.4-22)$$

where

$$\Sigma_p = \sum_{k=1}^m \Sigma_p^k = \sum_{k=1}^m \mathbf{y}^k(\hat{\mathbf{x}}, t) (\mathbf{y}^k(\hat{\mathbf{x}}, t))^T \sigma_{a_k}^2 \lambda_k(t) \quad (2.4-23)$$

Equations 2.4-15 and 2.4-22 are the analogs of the scalar results in Eqs. 2.3-58 and 2.3-60.

The update equations which modify the mean and covariance to include the information obtained through the observations $\mathbf{y}_k$ are developed using the same ideas as in the scalar case. A detailed presentation is in Jazwinski (Ref. 7, pp. 363-365) and will not be repeated here. The continuous-discrete extended Kalman filter equations are summarized in Table 2.4-1.

2.5 SUMMARY

This chapter contains the development of a set of equations for estimating the state of a power system. The major advantages of this formulation over existing algorithms are the use of a more realistic dynamical description of the power system state and the perturbing effects of load uncertainties. The overall estimation problem has been broken into two parts -- estimation of system state with an assumed system model and estimation of system structure. This decomposition is a direct result of the availability of system structure measurements. The design of an estimator for structure identification is the subject of the next chapter.

The estimation equations developed in this chapter are suitable for implementation in a load tracking configuration where the expected rate of system change is relatively slow. Concurrently, it is expected that the observations will also be taken at a low rate (on the order of one every few seconds). Thus, the computational burden associated with the filtering equations (including calculation of the gain matrix) is not expected to be excessive if a judicious choice is made of system state variables. The decomposition techniques discussed in Chapter 4 are useful in this regard.

TABLE 2.4-1
CONTINUOUS-DISCRETE EXTENDED KALMAN
FILTER EQUATIONS FOR A VECTOR PROCESS

SYSTEM MODEL	$d\mathbf{x} = \mathbf{f}(\mathbf{x}, t) dt + \mathbf{G}(\mathbf{x}, t) d\mathbf{w} + \mathbf{r}(\mathbf{x}, t) d\mathbf{n}_p$
MEASUREMENT MODEL	$\mathbf{y}_k = \mathbf{h}(\mathbf{x}(t_k)) + \mathbf{v}_k$
STATE ESTIMATE PROPAGATION	$d\hat{\mathbf{x}} = \mathbf{f}(\hat{\mathbf{x}}, t) dt$
ERROR COVARIANCE PROPAGATION	$\frac{d\mathbf{P}}{dt} = \mathbf{F}(\hat{\mathbf{x}}, t) \mathbf{P}(t) + \mathbf{P}(t) \mathbf{F}^T(\hat{\mathbf{x}}, t) + \mathbf{G}(\hat{\mathbf{x}}, t) \mathbf{Q}(t) \mathbf{G}^T(\hat{\mathbf{x}}, t) + \Sigma_p(t)$
STATE ESTIMATE UPDATE	$\hat{\mathbf{x}}_k(+) = \hat{\mathbf{x}}_k(-) + \mathbf{K}_k [\mathbf{y}_k - \mathbf{h}_k(\hat{\mathbf{x}}_k(-))]$
ERROR COVARIANCE UPDATE	$\mathbf{P}_k(+) = [I - \mathbf{K}_k \mathbf{H}_k(\hat{\mathbf{x}}_k(-))] \mathbf{P}_k(-)$
GAIN MATRIX	$\mathbf{K}_k = \mathbf{P}_k(-) \mathbf{H}_k^T(\hat{\mathbf{x}}_k(-)) [\mathbf{H}_k(\hat{\mathbf{x}}_k(-)) \mathbf{P}_k(-) \mathbf{H}_k^T(\hat{\mathbf{x}}_k(-)) + \mathbf{R}_k]^{-1}$
DEFINITIONS	$\mathbf{F}(\hat{\mathbf{x}}, t) = \left. \frac{\partial \mathbf{f}(\mathbf{x}, t)}{\partial \mathbf{x}} \right _{\mathbf{x}=\hat{\mathbf{x}}}$ $\mathbf{H}_k(\hat{\mathbf{x}}_k(-)) = \left. \frac{\partial \mathbf{h}(\mathbf{x}(t_k))}{\partial \mathbf{x}} \right _{\mathbf{x}=\hat{\mathbf{x}}_k(-)}$

3.1 INTRODUCTION

This chapter addresses the problem of power system structure determination. In the context of this chapter, system structure is synonymous with the (p -dimensional) state vector z defined in Section A.3.1, since knowledge of the nonzero elements of z uniquely identifies the available generators and transmission lines. Thus, the structure determination problem is equivalent to an estimation problem for the vector z . Viewed in this manner, modern estimation techniques can be profitably employed.

As formulated in Section A.3.1, each element of z is modeled as a random telegraph wave. Under the reasonable assumption that the likelihood of a transition from one state to the other is at most a function of the current state, the process is Markovian. In addition, it is also assumed throughout this chapter that the stochastic processes $z_i(t)$ and $z_j(t)$ are independent for all t when $i \neq j$. Although this assumption ignores the very real possibility of cascading faults, it does allow the p -dimensional estimation problem to be reduced to p -independent one-dimensional problems*.

Explicit equations for computing optimal estimates of the process $z(t)$ are presented in the following sections. The case of optimal nonlinear estimation with continuous measurements is treated first in Section 3.2. Since the assumption of continuous measurements may not be realized in actual applications, Section 3.3 contains recent results on optimal nonlinear filtering with discrete measurements. Section 3.4 examines optimal linear filtering when measurements are available continuously in time. Restriction of the estimator to an equation linear in z_i results in a simplified computational algorithm; however, since the $z_i(t)$ are not Gaussian processes, the nonlinear estimator should provide better estimation accuracy (e.g., a lower mean square error) at the expense of a more complex algorithm. Equations for the optimal linear filter with discrete measurements are not treated in this chapter, since this problem can be embedded in the general formulation used for process state estimation (see Chapter 2). Section 3.5 presents an approach to power system structure estimation that is based on hypothesis testing techniques. This approach is intended to supplement the explicit estimation algorithms developed in previous sections. A brief summary of the proposed two-stage structure estimation algorithm is given in Section 3.6.

3.2 OPTIMAL NONLINEAR FILTER FOR CONTINUOUS MEASUREMENTS

3.2.1 Problem Statement

Before discussing the optimal filtering equations it is necessary to formulate the problem to be solved in as precise a manner as possible. Throughout this chapter attention is restricted to scalar-valued processes only. Extension of the results to vector-valued processes is possible but is not required for the problem at hand.

*Were the model formulated to include the possibility of cascading faults, there would still remain the extremely difficult problem of characterizing (in a probabilistic sense) the state and fault dependent process.

In Section A.3.1, the random telegraph wave was written in the form*

$$z(t) = \frac{1}{2} (-1)^{N(t)} + \frac{1}{2} \quad 0 \leq t \leq T \quad (3.2-1)$$

with $N(t)$ an inhomogeneous Poisson counting process with intensity $\lambda(t)$. An alternate characterization that is more suited to the estimation problem is to view $z(t)$ as a finite-state continuous-time (FSCT) Markov process which takes values in the countable set $E = \{\alpha, \beta, \gamma, \dots\}$. (In the specific case to be considered herein $E = \{0, 1\}$; however, to maintain generality, results will be stated for the above countable set.) From Eq. 3.2-1 and the properties of $N(t)$ it is readily apparent that the sample paths of the process $z(t)$ are continuous from the right.

Throughout this chapter it is assumed that all of the stochastic processes under consideration are defined relative to an underlying (complete) probability space $(\Omega, \mathcal{F}, P)$ where

- Ω is a set consisting of the elementary events w
- $\mathcal{F}$ is a distinguished system of subsets of Ω which forms a σ -algebra
- P is a probability measure defined on sets from the σ -algebra $\mathcal{F}$.

In the following let

$$p_\beta(t) = \Pr\{z(t) = \beta\} \quad (3.2-2)$$

and denote the transition probability by

$$p_{\alpha\beta}(s, t) = \Pr\{z(t) = \beta | z(s) = \alpha\} \quad 0 \leq s < t < T \quad (3.2-3)$$

Now, $p_{\alpha\beta}(s, t)$ satisfies the forward Kolmogorov equation (Ref. 14)

$$p_{\alpha\beta}(s, t) = \delta_{\alpha\beta} + \int_s^t \mathcal{L}^* p_{\alpha\beta}(s, u) du \quad (3.2-4)$$

where

$$\mathcal{L}^* p_{\alpha\beta}(s, u) = \sum_{\gamma \in E} \lambda_{\gamma\beta}(u) p_{\alpha\gamma}(s, u) \quad (3.2-5)$$

and $\delta_{\alpha\beta}$ is the Kronecker delta. Similarly, the probabilities $p_\beta(t)$ satisfy

$$p_\beta(t) = p_\beta(0) + \int_0^t \mathcal{L}^* p_\beta(u) du \quad (3.2-6)$$

*In the remainder of this chapter the subscript on z is dropped for notational convenience.

The function $\lambda_{\alpha\beta}(u) = \lim_{r \rightarrow 0} [p_{\alpha\beta}(u, u+r) - \delta_{\alpha\beta}]/r$ is the derivative of the transition probability from α to β at time u

To illustrate the structure of these equations consider a homogenous random telegraph wave ($E = \{0, 1\}$) with time-invariant transition probabilities. In this case,

$$\begin{aligned}\lambda_{00} &= -\mu_0 \\ \lambda_{01} &= \mu_0 \\ \lambda_{10} &= \mu_1 \\ \lambda_{11} &= -\mu_1\end{aligned}\quad (3.2-7)$$

For this homogeneous process $p_{\alpha\beta}(s, t) = p_{\alpha\beta}(\tau = t - s)$ and Eq. 3.2-4 may be written (for $\alpha=0, \beta=1$) as

$$\begin{aligned}p_{01}(\tau) &= \int_0^\tau [\lambda_{01}p_{00}(u) + \lambda_{11}p_{01}(u)] du \\ &= \int_0^\tau [\mu_0 p_{00}(u) - \mu_1 p_{01}(u)] du\end{aligned}\quad (3.2-8)$$

Differentiating Eq. 3.2-8 with respect to τ yields a differential equation for $p_{01}(\tau)$, viz ,

$$\frac{d}{d\tau} p_{01}(\tau) = \mu_0 p_{00}(\tau) - \mu_1 p_{01}(\tau) \quad (3.2-9)$$

Equations for the remaining $p_{\alpha\beta}(\tau)$ are obtained in a similar manner. This set of coupled equations may be solved for the probabilities $p_{\alpha\beta}(\tau)$ in terms of the parameters μ_0 and μ_1 . In the limit as $\tau \rightarrow \infty$ the probabilities approach the constant values

$$\begin{aligned}\lim_{\tau \rightarrow \infty} p_{00}(\tau) &= \lim_{\tau \rightarrow \infty} p_{10}(\tau) = p_0 = \frac{\mu_1}{\mu_0 + \mu_1} \\ \lim_{\tau \rightarrow \infty} p_{01}(\tau) &= \lim_{\tau \rightarrow \infty} p_{11}(\tau) = p_1 = \frac{\mu_0}{\mu_0 + \mu_1}\end{aligned}\quad (3.2-10)$$

where p_α is the probability that the (steady state) process is in state α . The mean of the stationary process z is simply

$$E\{z\} = 0 \cdot p_0 + 1 \cdot p_1 = \frac{\mu_1}{\mu_0 + \mu_1} \quad (3.2-11)$$

In the general case (i.e., $\lambda_{\alpha\beta}(t)$ time-varying), closed form solutions are more difficult to obtain.

Equation 3.2-4 characterizes the probabilities $p_{\alpha\beta}(s, t)$. Of greater interest in the sequel is an expression involving the process $z(t)$ itself. [To be more precise, the quantity of interest is a measurable function of the process, i.e., $h(z(t), t)$] For each $\beta \in E$, define

$$h(z(t), t) \triangleq \delta_{\beta z(t)} = \delta_{\beta z(0)} + \int_0^t \lambda_{z(u)\beta}(u) du + x_\beta(t) \quad (3.2-12)$$

where $x_\beta(t)$ is a square integrable martingale* with right continuous trajectories. The proof of this assertion is in Liptser and Shirayev (Ref. 16, p. 333). The usefulness of the representation given by Eq. 3.2-12 lies in the fact that $\delta_{\beta z(t)}$ is of an appropriate form for application of the optimal non-linear filtering equation discussed below.

In this section it is assumed that observations of the process $z(t)$ are of the form

$$dq(t) = c(z(t), t) dt + \sigma_w(t) dw(t) \quad (3.2-13)$$

where $w(t)$ is a standard (or normalized) Wiener process independent of $z(t)$. The filtering problem may be stated as follows: given the realization $q^t = \{q(s), s \leq t\}$ of the observation process, construct an estimate of the process $z(t)$ based upon q^t . To this end, denote the a posteriori probability by

$$\pi_\beta(t) = \Pr\{z(t) = \beta | q(s), s \leq t\} \quad (3.2-14)$$

One commonly used estimate is the conditional expectation which yields the minimum mean square error estimate of the process. This is readily evaluated using Eq. 3.2-14, viz.,

$$E\{z(t) | q(s), s \leq t\} = \sum_{\beta \in E} \beta \pi_\beta(t) \quad (3.2-15)$$

The a posteriori probability can be related to $h(z(t), t) = \delta_{\beta z(t)}$ in the following manner:

$$\begin{aligned}E\{\delta_{\beta z(t)} | q(s), s \leq t\} \\ = 0 \cdot \Pr\{z(t) \neq \beta | q(s), s \leq t\} \\ + 1 \cdot \Pr\{z(t) = \beta | q(s), s \leq t\} \\ = \pi_\beta(t)\end{aligned}\quad (3.2-16)$$

An explicit expression for computing the $\pi_\beta(t)$ is given in the next section.

*A detailed discussion of martingales may be found in Refs. 14 through 17.

3.2.2 Optimal Nonlinear Filter

For the state and observation processes defined by Eqs. 3.2-12 and 3.2-13, the optimal estimates of the a posteriori probabilities $\pi_\beta(t)$, $\beta \in E$, satisfy the following set of nonlinear stochastic equations

$$\begin{aligned} \pi_\beta(t) = & p_\beta(0) + \int_0^t \mathcal{L}^* \pi_\beta(u) du \\ & + \int_0^t \pi_\beta(u) \frac{c(\beta, u) - \bar{c}(u)}{\sigma_w(u)} d\bar{w}(u) \end{aligned} \quad (3.2-17)$$

where

$$\mathcal{L}^* \pi_\beta(u) = \sum_{\gamma \in E} \lambda_{\gamma\beta}(u) \pi_\gamma(u) \quad (3.2-18)$$

$$\bar{c}(t) = \sum_{\gamma \in E} c(\gamma, t) \pi_\gamma(u) \quad (3.2-19)$$

and $\bar{w}(t)$ is a Wiener process given by

$$\bar{w}(t) = \int_0^t \frac{dq(u) - \bar{c}(u)du}{\sigma_w(u)} \quad (3.2-20)$$

An alternate notation is to write Eq. 3.2-17 as a stochastic differential equation

$$\begin{aligned} d\pi_\beta(t) = & \mathcal{L}^* \pi_\beta(t) dt + \pi_\beta(t) \left[\frac{\bar{c}(\beta, t) - \bar{c}(t)}{\sigma_w^2(t)} \right] \\ & \times [dq(t) - \bar{c}(t) dt] \end{aligned} \quad (3.2-21)$$

This equation is to be interpreted as shorthand notation for Eq. 3.2-17. Equation 3.2-17 is derived in Ref. 16 for a wider class of processes than considered herein. [Wonham (Ref. 18) had solved the particular case of a stationary FSCT Markov process with linear observations in 1965.] Certain conditions on $\lambda_{\alpha\beta}(t)$, $x_\beta(t)$, $c(z(t), t)$, etc that must be satisfied are stated in Ref. 16 and will not be repeated here. They deal primarily with continuity and boundedness which is inherent in this physically based problem.

Since Eq. 3.2-21 is a nonlinear, stochastic differential equation, the question arises as to whether its solutions are unique. This was answered in the affirmative by Pozov's'ii and Shirayev (Ref. 19). A proof is also contained in Ref. 16, pp. 335-338.

3.2.3 Application of the Filtering Equation to the Random Telegraph Wave

An explicit representation of the optimal nonlinear filter for the random telegraph wave is given in this section. As stated above, the process takes values from the set $E = \{0, 1\}$. Associated with state 0 are the time-varying derivatives of the transition probabilities

$$\lambda_{00}(t) = -\mu_0(t) \quad (3.2-22)$$

$$\lambda_{01}(t) = \mu_0(t)$$

and with state 1,

$$\lambda_{10}(t) = \mu_1(t) \quad (3.2-23)$$

$$\lambda_{11}(t) = -\mu_1(t)$$

The parameters $\mu_0(t)$ and $\mu_1(t)$ can be interpreted as the instantaneous average rates at which the process jumps from the 0 to the 1 state and from the 1 to the 0 state, respectively. The time-dependence allows for nonconstant failure rates, e.g., increased likelihood of a transmission line outage during storms.

Substituting Eqs. 3.2-22 and 3.2-23 into Eq. 3.2-21 and using the fact that $\pi_0(t) + \pi_1(t) = 1$ yields the pair of equations

$$\begin{aligned} d\pi_1(t) = & \mu_0(t)dt - [\mu_0(t) + \mu_1(t)] \pi_1(t) dt \\ & + \left[\frac{c(1, t) - c(0, t)}{\sigma_w^2(t)} \right] \pi_1(t) [1 - \pi_1(t)] \\ & \times [dq(t) - \{c(0, t) + [c(1, t) - c(0, t)] \pi_1(t)\} dt] \end{aligned} \quad (3.2-24)$$

$$\begin{aligned} d\pi_0(t) = & \mu_1(t)dt - [\mu_0(t) + \mu_1(t)] \pi_0(t) dt \\ & + \left[\frac{c(0, t) - c(1, t)}{\sigma_w^2(t)} \right] \pi_0(t) [1 - \pi_0(t)] \\ & \times [dq(t) - \{c(1, t) + [c(0, t) - c(1, t)] \pi_0(t)\} dt] \end{aligned} \quad (3.2-25)$$

These two equations provide explicit formulae for obtaining the a posteriori probabilities $\pi_0(t)$ and $\pi_1(t)$. Since the system is in either state 0 or state 1 it is unnecessary to solve both equations. Solving for $\pi_1(t)$ via Eq. 3.2-24 then gives $\pi_0(t)$ as $1 - \pi_1(t)$. It is straightforward to verify that Eqs. 3.2-24 and 3.2-25 are self-consistent in the sense that

$$d\pi_1(t) + d\pi_0(t) = d[\pi_1(t) + \pi_0(t)] = d(1) = 0 \quad (3.2-26)$$

To complete the presentation of the optimal nonlinear filter equations, Eqs. 3.2-24 and 3.2-25 will be specialized to the case of linear measurements, viz.,

$$c(z(t), t) = z(t) \quad (3.2-27)$$

The resulting filter equations are

$$\begin{aligned} d\pi_1(t) = & \mu_0(t)dt - [\mu_0(t) + \mu_1(t)] \pi_1(t) dt \\ & + \frac{1}{\sigma_w^2(t)} \pi_1(t) [1 - \pi_1(t)] [dq(t) - \pi_1(t)dt] \end{aligned} \quad (3.2-28)$$

$$\begin{aligned} d\pi_0(t) = & \mu_1(t)dt - [\mu_0(t) + \mu_1(t)] \pi_0(t) dt \\ & - \frac{1}{\sigma_w^2(t)} \pi_0(t) [1 - \pi_0(t)] [dq(t) - (1 - \pi_0(t))dt] \end{aligned} \quad (3.2-29)$$

3.3 OPTIMAL NONLINEAR FILTER FOR DISCRETE MEASUREMENTS

The optimal nonlinear filtering equations of the previous section are applicable when observations are available continuously in time. In this section, the assumption of continuous measurements is relaxed. The scalar observations of the process $z(t)$ are taken to be of the form

$$q_k = c_k(z(t_k)) + \sigma_{v_k} v_k \quad (3.3-1)$$

where

$$c_k(t) = c_k(z(t)) \text{ is a function of } z(t), \quad t \leq t_k$$

$$v_k \text{ is a zero-mean, uncorrelated Gaussian noise sequence with } E\{v_j v_k\} = \delta_{jk}.$$

In Section 2.3.1, Bayes' rule was used to obtain an expression for the behavior of the state estimate at the discrete observation times. This result is applicable only when the system state is a Markov process and the measurement is function of the state at the current observation time, i.e., $c_k = c_k(z(t_k))$. The disadvantage of the general expression given in Eq. 2.3-16 is that the conditional density function must be known. Computationally tractable implementations of the update equations are obtained by assuming a specific functional form for the estimate, e.g., an affine function of the observation.

A more general solution to the continuous-discrete nonlinear filtering problem is given in a recent paper by Takeuchi and Akashi (Ref. 20). This formulation is in the same vein as the discussion of Section 3.2. It is included here as an alternate approach to the discrete time observation case.*

*Since $z(t)$ is a Markov process, the Bayes' formula approach could be used in this section if c_k were restricted to be a function only of $z(t_k)$. However, the update equation obtained in Chapter 2 was based on the assumption that the state estimate was linear in the measurement. This need not hold for the general case

As in the case of process state estimation, the filtering problem can be broken into two parts: propagation of the estimate between observations and a correction at the observation time. In this section, the a posteriori probability $\pi_\beta(t)$ is defined with respect to the countable sequence of observations, specifically,

$$\pi_\beta(t) = \begin{cases} E\{h(t)|q_1, q_2, \dots, q_{k-1}\} & \text{for } t \in (t_{k-1}, t_k), 1 \leq k \leq N \\ E\{h(t)|q_1, q_2, \dots, q_k\} & \text{for } t = t_k, 1 \leq k \leq N \end{cases} \quad (3.3-2)$$

where $h(t)$ denotes $\delta_{\beta z(t)}$ given by Eq. 3.2-12. The equations characterizing the optimal, nonlinear continuous-discrete filter are (Ref. 20, Theorem 1)

$$\begin{aligned} \pi_\beta(t) = & \pi_\beta(t_{k-1}^+) + \int_{t_{k-1}}^t \mathcal{L}^* \pi_\beta(s) ds \\ & t \in (t_{k-1}, t_k), 1 \leq k \leq N \end{aligned} \quad (3.3-3)$$

$$\pi_\beta(t_k^+) = \hat{\pi}_{\beta k}^t \quad t = t_k, 1 \leq k \leq N \quad (3.3-4)$$

The function $\hat{\pi}_{\beta k}^t$ appearing in Eq. 3.3-4 is a correction term which accounts for the information contained in the measurement q_k . It is shown in Ref. 20 that $\hat{\pi}_{\beta k}^t$ satisfies the stochastic differential equation

$$\begin{aligned} \frac{d\hat{\pi}_{\beta k}^t}{dt} = & \{h(t_k) - \hat{\pi}_{\beta k}^t\} u(t) \left(\frac{1}{\sigma_{v_k}^2} \right) \{q_k - \tilde{c}_k(t)\}^t \\ & t \in (t_{k-1}, t_k], 1 \leq k \leq N \end{aligned} \quad (3.3-5)$$

with initial condition

$$\hat{\pi}_{\beta k}^{t_{k-1}} = E\{h(t_k)|q_1, q_2, \dots, q_{k-1}\} \quad (3.3-6)$$

(Note that Eq. 3.3-6 is the one-step prediction of $h(t)$.)

In Eq. 3.3-5, $(\hat{\cdot})^t$ denotes conditional expectation with respect to a family of probability measures $\hat{P}_t$, $t \in [0, T]$ obtained by transformations of the underlying probability measure P . These measures are defined in terms of the process $u(t)$, $t \in [0, T]$ which generates an interpolation of the sequence $c_k(t_k)$, $1 \leq k \leq N$. Specifically, $u(t)$ is chosen such that

$$c_k(t_k) = \int_{t_{k-1}}^{t_k} u(s) ds \quad 1 \leq k \leq N \quad (3.3-7)$$

One $u(t)$ which satisfies Eq. 3.3-7 is the piecewise constant function

$$u(s) = \frac{c_k(t_k)}{(t_k - t_{k-1})} = c_k^* \quad t_{k-1} \leq t < t_k \quad (3.3-8)$$

Let t be any point in the interval $[0, T]$ and set

$$\tilde{u}^t(s) = \begin{cases} u(s) & 0 \leq s < t \\ 0 & t \leq s < T \end{cases} \quad (3.3-9)$$

Then, $\tilde{c}_k(t)$ is defined as

$$\tilde{c}_k(t) = c_k[\tilde{u}^t(s)] \quad (3.3-10)$$

For $u(s)$ given by Eq. 3.3-8, $\tilde{c}_k(t)$ becomes

$$\tilde{c}_k(t) = (t - t_{k-1}) c_k^* \quad (3.3-11)$$

which is an interpolation of the measurement $c_k(t_k)$ over the interval $[t_{k-1}, t_k]$.

Equation 3.3-5 can be rewritten in a more convenient form if $u(t)$ and $\tilde{c}_k(t)$ are defined as above. Substitution of Eqs. 3.3-8 and 3.3-11 into Eq. 3.3-5 yields

$$\begin{aligned} \frac{d\hat{\pi}_{\beta k}^t}{dt} &= \overbrace{\{h(t_k) - \hat{\pi}_{\beta k}^t\} \frac{c_k(t_k)}{(t_k - t_{k-1})} \frac{1}{\sigma_{v_k}^2} \{q_k - \frac{(t - t_{k-1})}{(t_k - t_{k-1})} c_k(t_k)\}}^t \\ &= \frac{1}{(t_k - t_{k-1}) \sigma_{v_k}^2} \overbrace{\{h(t_k) - \hat{\pi}_{\beta k}^t\} c_k(t_k) \{q_k - \frac{(t - t_{k-1})}{(t_k - t_{k-1})} c_k(t_k)\}}^t \end{aligned} \quad (3.3-12)$$

As shown in Ref. 20, $\{q_k - \frac{t - t_{k-1}}{t_k - t_{k-1}} c_k(t_k)\}$, with respect to the measure $\hat{P}_t$, is Gaussian and independent of the σ -algebras generated by the process $z(s)$, $s \leq t_k$. Thus, Eq. 3.3-12 can be simplified to

$$\begin{aligned} \frac{d\hat{\pi}_{\beta k}^t}{dt} &= \frac{1}{(t_k - t_{k-1}) \sigma_{v_k}^2} \overbrace{\{h(t_k) - \hat{\pi}_{\beta k}^t\} c(t_k)}^t \\ &\quad \{q_k - \frac{(t - t_{k-1})}{(t_k - t_{k-1})} c_k(t_k)\} \end{aligned} \quad (3.3-13)$$

In order to examine the character of Eq. 3.3-13, set $\beta=1$ and consider the case of linear measurements

$$c_k(z(t_k)) = z(t_k) \quad 1 \leq k \leq N \quad (3.3-14)$$

From the definitions of $h(t)$ and $z(t)$ it is obvious that the following relationships hold for $\beta=1$:

$$h(t) = \delta_{1z}(t) = z(t) = z^2(t) \quad (3.3-15)$$

Using Eqs. 3.3-14 and 3.3-15, Eq. 3.3-13 can be written

$$\begin{aligned} \frac{d\hat{\pi}_{1k}^t}{dt} &= \frac{1}{(t_k - t_{k-1}) \sigma_{v_k}^2} \overbrace{\{z(t_k) - \hat{\pi}_{1k}^t\} z(t_k)}^t \\ &\quad \times \{q_k - \frac{(t - t_{k-1})}{(t_k - t_{k-1})} z(t_k)\} \\ &= \frac{1}{(t_k - t_{k-1}) \sigma_{v_k}^2} \overbrace{\{z(t_k)\}^t}^t - \hat{\pi}_{1k}^t \overbrace{\{z(t_k)\}^t}^t \\ &\quad \times \{q_k - \frac{(t - t_{k-1})}{(t_k - t_{k-1})} z(t_k)\} \end{aligned} \quad (3.3-16)$$

Finally, by noting that $\overbrace{\{z(t_k)\}^t}^t = \hat{\pi}_{1k}^t$, Eq. 3.3-16 becomes

$$\frac{d\hat{\pi}_{1k}^t}{dt} = \frac{1}{(t_k - t_{k-1}) \sigma_{v_k}^2} \hat{\pi}_{1k}^t \{1 - \hat{\pi}_{1k}^t\} \{q_k - \frac{(t - t_{k-1})}{(t_k - t_{k-1})} \hat{\pi}_{1k}^t\} \quad (3.3-17)$$

A parallel development would yield a similar equation for $\hat{\pi}_{0k}^t$. The structure of Eq. 3.3-17 arises only in the case of linear measurements. In general, the first term in Eq. 3.3-13 is

$$\overbrace{\{h(t_k) - \hat{\pi}_{\beta k}^t\} c(t_k)}^t = \overbrace{h(t_k) c(t_k)}^t - \hat{\pi}_{\beta k}^t \overbrace{c(t_k)}^t \quad (3.3-18)$$

and $\overbrace{h(t_k) c(t_k)}^t$ cannot be expressed as a function of $\hat{\pi}_{\beta k}^t$.

In order to solve Eq. 3.3-17 it is necessary to have available q_k , the measurement at time t_k , and the initial condition

$$\hat{\pi}_{1k}^{t_{k-1}} = \pi_1(t_k^-) = \pi_1(t_k^+) + \int_{t_{k-1}}^{t_k} \mathcal{L}^* \pi_1(s) ds \quad (3.3-19)$$

As in the continuous measurement case, the filter equations are nonlinear differential equations that do not have simple closed-form solutions. Consequently, they must be solved on-line as the measurement data become available.

3.4 OPTIMAL LINEAR FILTER FOR CONTINUOUS MEASUREMENTS

The nonlinear filter described in the previous sections provides optimal estimates of the a posteriori probabilities $\pi_\beta(t)$. In certain applications it may prove difficult to solve the nonlinear filter equations (e.g., Eqs. 3.2-24 and 3.2-25). This section provides a computationally simpler approach to the estimation problem by restricting attention to a linear estimator.

3.4.1 Problem Formulation

The goal of this section is to obtain a linear stochastic differential equation representation for $z(t)$. (The basic approach employed herein is treated briefly by Davis, Ref. 21.) To this end it is necessary to examine the Poisson counting process, $N(t)$, in more detail. Using the notation of Section 3.2.3, the intensity of the underlying Poisson point process is

$$\lambda(t) = \begin{cases} \mu_1(t) & w_{2j} \leq t < w_{2j+1} \\ \mu_0(t) & w_{2j+1} \leq t < w_{2j+2} \end{cases} \quad j=0,1,2,\dots \quad (3.4-1)$$

(Implicit in the definition of Eq. 3.4-1 is the assumption that the process $z(t)$ is in state 1 at $t=w_0=0$.)

The point process $N(t)$ may be written (to within stochastic equivalence) as

$$N(t) = m(t) + A(t) \quad (3.4-2)$$

with $m(t)$ a local martingale and $A(t)$ an increasing process*. Equation 3.4-2 is the Doob-Meyer decomposition of the process (see Refs. 15 and 16). The increasing process $A(t)$ is called the compensator of the point process $N(t)$. Let

$$F_1(t) = \Pr\{w_1 \leq t\} \quad (3.4-3)$$

and

$$F_i(t) = \Pr\{w_i \leq t | w_{i-1}, \dots, w_1\} \quad i \geq 2 \quad (3.4-4)$$

be the conditional distribution functions associated with the point process. The compensator $A(t)$ is defined as

$$A(t) = \sum_{i \geq 1} A_i(t) \quad (3.4-5)$$

where

$$A_i(t) = \int_0^{t \wedge w_i} \frac{dF_i(u)}{1-F_i(u_-)} \quad i \geq 1 \quad (3.4-6)$$

*All of the stochastic processes discussed in this section are defined relative to an underlying probability space $(\Omega, \mathcal{F}, P)$. The right continuous stochastic process $A(t)$ is increasing if $A(t)$ is $\mathcal{F}_t$ measurable, $A(0)=0$ and $A(s) < A(t)$ almost surely (a.s.) for $s \leq t$, i.e., if $\Pr\{A(s) > A(t); s \leq t\} = 0$.

and $t \wedge w_i$ denotes the minimum of t and w_i . For a Poisson process with intensity $\lambda(t; w_{i-1}, \dots, w_1)$

$$F_i(u) = 1 - \Pr\{w_i > u | w_{i-1}, \dots, w_1\} = \begin{cases} 1 - e^{-\Lambda(u; w_{i-1})} & u \geq w_{i-1} \\ 0 & u < w_{i-1} \end{cases} \quad (3.4-7)$$

where

$$\Lambda(u; w_{i-1}, \dots, w_1) = \int_{w_{i-1}}^u \lambda(\sigma; w_{i-1}, \dots, w_1) d\sigma \quad (3.4-8)$$

Since

$$dF_i(u) = e^{-\Lambda(u; w_{i-1}, \dots, w_1)} \lambda(u; w_{i-1}, \dots, w_1) du \quad (3.4-9)$$

it follows that

$$A_i(t) = \int_0^{t \wedge w_i} \frac{dF_i(u)}{1-F_i(u_-)} = \int_{w_{i-1}}^{t \wedge w_i} \frac{e^{-\Lambda(u; w_{i-1}, \dots, w_1)} \lambda(u; w_{i-1}, \dots, w_1) du}{e^{-\Lambda(u; w_{i-1}, \dots, w_1)}} = \int_{w_{i-1}}^{t \wedge w_i} \lambda(u; w_{i-1}, \dots, w_1) du \quad (3.4-10)$$

(From Eq. 3.4-10 it is obvious that the $A_i(t)$ are random since the limits on the integral are the a priori unknown point process occurrence times w_i .)

The random telegraph wave can also be decomposed in a similar manner. Observe that $z(t)$ and $N(t)$ are related by

$$z(t) = \begin{cases} 1 & N(t) \text{ even} \\ 0 & N(t) \text{ odd} \end{cases} \quad (3.4-11)$$

Thus, each jump of $z(t)$ corresponds to a jump of $N(t)$. By definition, the jumps in $N(t)$ are positive and of unit magnitude. The jumps of $z(t)$ are also of unit magnitude but alternate in sign. For $N(t)$ even, the compensator of $z(t)$ is simply the $A_i(t)$ given by Eq. 3.4-10. For $N(t)$ odd, the process is the negative of the standard counting process and the compensator is $-A_i(t)$. From Eqs. 3.4-5 and 3.4-10 it is easy to see that $A(t)$ is continuous if $\lambda(u; w_{i-1}, \dots, w_i)$ is. However, the derivative of $A(t)$ at each of the occurrence times w_i is discontinuous. Specifically,

$$\frac{d}{dt} A(t) = \frac{d}{dt} A_i(t) = \begin{cases} \lambda(t; w_{i-1}, \dots, w_1) & i \text{ odd} \\ -\lambda(t; w_{i-1}, \dots, w_1) & i \text{ even} \end{cases} \quad (3.4-12)$$

since all the $A_j(t)$, $j \neq i$, are constant in the interval $w_{i-1} \leq t < w_i$. For the intensity of the random telegraph wave defined in Eq. 3.4-1,

$$\frac{d}{dt} A(t) = \begin{cases} \mu_1(t) & z(t)=1 \\ -\mu_0(t) & z(t)=0 \end{cases} \quad (3.4-13)$$

At this point it is useful to decompose $z(t)$ in a manner analogous to the Doob-Meyer decomposition. It is possible to show (see Appendix D) that, for μ_1 and μ_0 constant,

$$z(t) = A(t) + \tilde{\zeta}(t) + \xi \quad (3.4-14)$$

where $A(t)$ is as defined above, $\tilde{\zeta}(t)$ is an orthogonal increment process and ξ is a constant. This result is less general than the Doob-Meyer decomposition but is sufficient for the problem at hand. Using Eqs. 3.4-13 and 3.4-14,

$$\begin{aligned} dz(t) &= dA(t) + d\tilde{\zeta}(t) \\ &= \mu_1 z(t) dt - \mu_0 [1-z(t)] dt + d\tilde{\zeta}(t) \\ &= -\mu_0 dt + [\mu_0 + \mu_1] z(t) dt + d\tilde{\zeta}(t) \end{aligned} \quad (3.4-15)$$

Equation 3.4-15 is the desired linear stochastic differential equation for the process $z(t)$.

3.4.2 Optimal Linear Filter

The linear stochastic differential equation for the process $z(t)$ derived in the previous section is of the form

$$dz(t) = a_0 dt + a_1 z(t)dt + d\tilde{\zeta}(t) \quad (3.4-16)$$

with $\tilde{\zeta}(t)$ a stationary orthogonal-increment process (e.g., a Wiener process). The corresponding linear measurement equation is

$$dq(t) = b_1 z(t) dt + \sigma_v dv(t) \quad (3.4-17)$$

with $v(t)$ a Wiener process and σ_v constant.

For the linear system representation of Eqs. 3.4-16 and 3.4-17, it is well known that the Kalman filter yields the minimum mean square error estimate of the system state (Refs. 6, 16, and 21). The continuous-time filter equations for a scalar process are summarized in Table 3.4-1. The advantage of this formulation lies in the fact that the state estimate

TABLE 3.4-1
CONTINUOUS KALMAN FILTER EQUATIONS
FOR A SCALAR PROCESS

System Model	$dz(t) = a_0 dt + a_1 z(t)dt + a_2 dw(t)$
Measurement Model	$dq(t) = b_1 z(t) dt + b_2 dv(t)$
State Estimate	$d\hat{z}(t) = a_0 dt + a_1 \hat{z}(t) dt + \frac{p(t)b_1}{b_2^2} [dq(t) - b_1 \hat{z}(t)dt]$
Error Covariance	$\dot{p}(t) = 2a_1 p(t) - \frac{b_1^2 p^2(t)}{b_2^2} + a_2^2$
Kalman Gain	$k(t) = p(t) b_1 / b_2^2$
Assumptions	$E\{dwdv\} = 0,$ $E\{(dw)^2\} = E\{(dv)^2\} = dt$

is given by a linear stochastic differential equation rather than the nonlinear estimation equations of Section 3.2.3*.

This simplified structure for the estimation equations is obtained at the expense of an approximate system model. In order to obtain an explicit representation for $\tilde{\zeta}(t)$ in Eq. 3.4-15 it was necessary to assume that the transition probabilities (μ_0 and μ_1) were time-invariant and that the process $z(t)$ had reached steady-state. The validity of these assumptions is obviously dependent upon the characteristics of the problem at hand. Thus, any decision as to which estimator formulation is preferable must be made on a case by case basis.

3.5 HYPOTHESIS TESTING FOR STRUCTURE ESTIMATION

3.5.1 Introduction

In the preceding sections, algorithms have been developed for estimating power system topological structure. If a complete set of breaker status measurements were available, these algorithms would yield the best estimate of system structure. However, for a partial set of measurements, the structure estimates will be incomplete in the sense that the estimates for any unobserved breakers are based on assumed transition rates and represent predicted rather than filtered estimates.

If the number of missing measurements is small, an hypothesis testing approach can be used to identify the true system structure from a set of candidate configurations. To illustrate the concept, consider the

*The error covariance Ricatti equation is independent of both the state estimate $\hat{z}(t)$ and the measurement $q(t)$ and can be precomputed leading to a further reduction in on-line computation.

situation shown in Fig. 3.5-1 where $A_1, \dots, A_N$ represent estimation algorithms associated with the N possible configurations of the power system. Each algorithm, which is designed to be an optimal estimator with respect to the assumed system model, provides a state estimate $\hat{x}_1^*$. If the true power system configuration was identical to one of the candidate configurations (for example, the j^{th}), then $\hat{x}_j$ would be the optimal estimate of the system state and the remaining estimates $\hat{x}_1, i \neq j$, would be suboptimal. In order to determine which estimate (and hence which configuration) "best" represents the power system, consider the following hypothesis test

Figure 3.5-1 Hypothesis Testing Approach to System Structure Determination

Let $\tilde{H}_1$ denote the hypothesis that configuration 1 is the correct one and associate with each hypothesis an a priori probability $P_1(0)$. Further, let Y_k denote the set of observations $\{y_1, y_2, \dots, y_k\}$ and define

$$P_1(k) = \Pr\{\tilde{H} = \tilde{H}_1 | Y_k\} \quad (3.5-1)$$

A reasonable rule for deciding upon a candidate configuration is to choose the hypothesis (i.e., the algorithm A_1) for which the a posteriori probability $P_1(k)$ is maximized. In this situation, the estimate $\hat{x}_1$ is used as the optimal state estimate and the configuration corresponding to $\tilde{H}_1$ is taken to be the true system structure

A similar approach to the simultaneous estimation and model identification (termed multiple model estimation) was first proposed for discrete-time linear systems by Magill (Ref. 22) and later extended by Lainiotis (Ref. 23). For a linear dynamical system, the state estimates are obtained from a bank of Kalman filters each of which is based on a system model corresponding to a different hypothesis. The most useful feature of the multiple model estimation approach is the existence of a recursive formula which expresses the $P_1(k+1)$ in terms of $P_1(k)$ and the filter residuals. The structure estimation problem can be formulated as shown in Fig. 3.5-1 with each of the algorithms A_1

taken to be the process state estimator developed in Chapter 2. Equations characterizing the multiple model algorithm for this case are developed in the next section.

3.5.2 Multiple Model Estimation Equations

This section contains a brief derivation of the multiple model estimation equations relevant to the structure estimation problem. The derivation below follows one presented in Ref. 24 which may be consulted for additional details. Use of variables associated with the extended Kalman filter equations (see Table 2.4-1) will be made without additional comment.

Assume that there are N filters, each of which processes measurement data $y_k, k = 1, 2, \dots$. Let $\tilde{H}_1$ denote the hypothesis that the model for the 1^{th} filter best represents the true power system configuration which generates the y_k . Let $P_1(k)$ denote the a posteriori probability of the 1^{th} filter after the k^{th} measurement. For $P_1(k)$ to be a true probability it is necessary that

$$P_1(k) \geq 0 \text{ for all } i \text{ and } k \quad (3.5-2)$$

$$\sum_{i=1}^N P_i(k) = 1 \text{ for all } k \quad (3.5-3)$$

The goal of the following development is to obtain a recursive expression for the $P_1(k)$.

To proceed, define a discrete random variable $\tilde{H}$ that takes the "values" $\tilde{H}_1, \tilde{H}_2, \dots, \tilde{H}_N$ and let Y_k denote the measurements up to time $k, i.e., Y_k = \{y_1, y_2, \dots, y_k\}$. With these definitions, the a posteriori probability is

$$P_1(k) = \Pr\{\tilde{H} = \tilde{H}_1 | Y_k\} \quad (3.5-4)$$

Furthermore, the conditional probability density function of the random variable $\tilde{H}$ may be written as.

$$f(\tilde{H} | Y_k) = \sum_{i=1}^N P_i(k) \delta(\tilde{H} - \tilde{H}_i) \quad (3.5-5)$$

where

$$\delta(\tilde{H} - \tilde{H}_1) = \begin{cases} 1 & \tilde{H} = \tilde{H}_1 \\ 0 & \text{otherwise} \end{cases}$$

Using Bayes' rule,

*For a linear dynamical system the optimal estimator is the Kalman filter.

$$f(\tilde{H}|Y_{k+1}) = f(\tilde{H}|Y_{k+1}, Y_k)$$

$$\begin{aligned} &= \frac{f(\tilde{H}, Y_{k+1}|Y_k)}{f(Y_{k+1}|Y_k)} \\ &= \frac{f(Y_{k+1}|\tilde{H}, Y_k)f(\tilde{H}|Y_k)}{f(Y_{k+1}|Y_k)} \end{aligned} \quad (3.5-6)$$

Combining Eqs. 3.5-5 and 3.5-6 yields

$$\begin{aligned} &\sum_{i=1}^N P_i(k+1) \delta(\tilde{H}-\tilde{H}_i) \\ &= \frac{f(Y_{k+1}|\tilde{H}, Y_k)}{f(Y_{k+1}|Y_k)} \sum_{i=1}^N P_i(k) \delta(\tilde{H}-\tilde{H}_i) \end{aligned} \quad (3.5-7)$$

from which

$$P_i(k+1) = \frac{f(Y_{k+1}|\tilde{H}_i, Y_k)}{f(Y_{k+1}|Y_k)} P_i(k) \quad (3.5-8)$$

The conditional density in the denominator of Eq. 3.5-8 can be expressed in terms of $f(Y_{k+1}|\tilde{H}_i, Y_k)$ as follows. First, note that

$$\begin{aligned} f(Y_{k+1}|Y_k) &= \int f(Y_{k+1}, \tilde{H}|Y_k) d\tilde{H} \\ &= \int f(Y_{k+1}|\tilde{H}, Y_k) f(\tilde{H}|Y_k) d\tilde{H} \end{aligned} \quad (3.5-9)$$

Substituting Eq. 3.5-5 into Eq. 3.5-9 gives

$$\begin{aligned} f(Y_{k+1}|Y_k) &= \int f(Y_{k+1}|\tilde{H}, Y_k) \sum_{j=1}^N P_j(k) \delta(\tilde{H}-\tilde{H}_j) d\tilde{H} \\ &= \sum_{j=1}^N P_j(k) f(Y_{k+1}|\tilde{H}_j, Y_k) \end{aligned} \quad (3.5-10)$$

Combining Eqs. 3.5-8 and 3.5-10 results in

$$P_i(k+1) = \frac{f(Y_{k+1}|\tilde{H}_i, Y_k)}{\sum_{j=1}^N P_j(k) f(Y_{k+1}|\tilde{H}_j, Y_k)} P_i(k) \quad (3.5-11)$$

which is the desired recursive relationship. It remains to express $f(Y_{k+1}|\tilde{H}_i, Y_k)$ in terms of known (or calculable) quantities.

Consider first the case of a linear dynamical system. If the measurements y are a linear function of the state, viz.,

$$Y_{k+1} = H_{k+1} x_{k+1} + v_{k+1} \quad (3.5-12)$$

it is possible to show that the conditional density associated with the i^{th} Kalman filter is Gaussian with mean value $H_{k+1}^i \hat{x}_{k+1}^i(-)$ and covariance

$$S_{k+1}^i = H_{k+1}^i P_{k+1}^i(-) (H_{k+1}^i)^T + R_{k+1} \quad (3.5-13)$$

Thus, the estimator supplies sufficient information at each measurement time to compute the conditional density and via Eq. 3.5-11 to update the a posteriori probabilities. An explicit expression for the $P_i(k+1)$ in terms of the filter estimates can be obtained as follows. The residual for the i^{th} filter is defined as

$$r_{k+1}^i = Y_{k+1} - H_{k+1}^i \hat{x}_{k+1}^i(-) \quad (3.5-14)$$

Setting

$$\beta_{k+1}^i = (2\pi)^{-\ell/2} [\det S_{k+1}^i]^{-1/2} \quad (3.5-15)$$

and

$$w_{k+1}^i = (r_{k+1}^i)^T (S_{k+1}^i)^{-1} r_{k+1}^i \quad (3.5-16)$$

allows the conditional density function to be written in the compact form

$$f(Y_{k+1} | \tilde{H}_i, Y_k) = \beta_{k+1}^i \exp(-\frac{1}{2} w_{k+1}^i) \quad (3.5-17)$$

where ℓ in Eq. 3.5-15 is the dimension of the measurement vector y . Finally, substitution of Eq. 3.5-17 into Eq. 3.5-11 gives the desired recursive expression

$$P_i(k+1) = \frac{\beta_{k+1}^i \exp(-\frac{1}{2} w_{k+1}^i)}{\sum_{j=1}^N \beta_{k+1}^j \exp(-\frac{1}{2} w_{k+1}^j) P_j(k)} P_i(k) \quad (3.5-18)$$

In the power system problem, both the system dynamics and measurements are nonlinear functions of the power system state. As discussed in Chapter 2, an optimal filter for such a system is computationally intractable. One approximate filter for such a system is the extended Kalman filter which is based on a linearization of the dynamic and measurement equations about the current state estimate. The resulting estimation equations have the same form as the Kalman filter.

For a general nonlinear system, the conditional density $f(Y_{k+1}|\tilde{H}_i, Y_k)$ need not be Gaussian. Thus,

although the expression for $P_1(k+1)$ given by Eq. 3.5-11 is still valid, it cannot be evaluated since the conditional density is not explicitly available (or computable). If however, the nonlinear system is one for which the extended Kalman filter provides acceptable estimation accuracy*, then the approximation of the conditional density by Eq. 3.5-17 should be reasonable and the recursion formula for the a posteriori probabilities is given by Eq. 3.5-18. Although these formulae are the same for both the Kalman filter and the extended Kalman filter it must be emphasized that the symbols have different meanings. For the i^{th} Kalman filter, $\hat{x}_{k+1}^1$ is the true minimum variance estimate (i.e., conditional mean) and P_{k+1}^1 is the true error covariance matrix. In the extended Kalman filter equations, $\hat{x}_{k+1}^1$ and P_{k+1}^1 are the approximate estimate and estimation error covariance computed from linearized propagation and update equations. Since these equations are linearized about the current state estimate, P_{k+1}^1 is a function of $\hat{x}_{k+1}^1$ and cannot be precomputed.

Once the posterior probabilities have been computed at time k , it is possible to determine the "true" configuration and best estimate of the power system state. The most likely configuration at the k^{th} time step is the one with the largest $P_j(k)$. Under the assumption that the system does not undergo another configuration change, it is reasonable to expect that one of the probabilities, say $P_1(k)$, will tend to unity while the remaining $P_j(k)$ tend to zero†. If this occurs, the i^{th} configuration would be chosen as representative of the true system structure. Consider now, the best state estimate defined as the conditional expectation

$$\begin{aligned}\hat{x}_k(+) &= E\{x_k | Y_k\} \\ &= \int x f(x_k | Y_k) dx\end{aligned}\quad (3.5-19)$$

Using the same argument as in Eq. 3.5-10, $\hat{x}_k(t)$ can be expressed as

$$\begin{aligned}\hat{x}_k(t) &= \sum_{i=1}^N P_i(k) \int x f(x_k | H_i, Y_k) dx \\ &= \sum_{i=1}^N P_i(k) \hat{x}_k^i(+)\end{aligned}\quad (3.5-20)$$

where

$$\hat{x}_k^i(+) = \int x f(x_k | H_i, Y_k) dx\quad (3.5-21)$$

*The performance of such a filter can be determined only through actual implementation or detailed simulation

†This will occur if for some i and all $k > k_0$, $w_k^1 \ll w_k^j$, $j \neq 1$, i.e., if the residual process associated with the i^{th} filter is small and the other residuals are large.

3.6 SUMMARY

The problem of power system structure estimation is treated in this chapter. The proposed algorithm for determining the network configuration is based on a two-stage approach. First, a maximum a posteriori estimate of the structure (as characterized by the topological state z) is obtained from available breaker status measurements. Then, a hypothesis testing approach is used to select the true network structure from the small number of candidate configurations which result from the first stage. This approach limits the number of hypotheses that must be evaluated, resulting in an implementable technique.

Three different algorithms are developed for use in the first stage: optimal nonlinear filters for both continuous- and discrete-time measurements and linear filter based on a steady-state process model. Since the topological state is non-Gaussian, the linear filter is clearly suboptimal. In the second stage, multiple model estimation techniques are used to select the correct hypothesis (i.e., network configuration) from among the candidate configurations.

4. TIME-SCALE DECOMPOSITION

4.1 INTRODUCTION

Large scale systems invariably possess some eigenvalues representative of high bandwidth. Such systems are governed by a set of stiff differential equations whose numerical solution is generally difficult. A large disparity between the smallest and the largest eigenvalue of a high-order system suggests that the system may be amenable to order reduction through time-scaling. In principle, the system may then be partitioned into two subsystems, one containing states with slow dynamics, and the other, states with relatively fast dynamics; this partition may be viewed as a time-scale decomposition.

Systems that may be partitioned into slow and fast subsystems in a manner such that the fast subsystem is stable, can be analyzed using the formalisms of singular perturbations (SP). SP techniques attempt to obtain an approximate solution in the form of a truncated asymptotic series -- usually a convergent-divergent power series. SP theory is comprehensively discussed by Nayfeh (Ref. 25), and an excellent review of papers dealing with applications to control problems has been published by Kokotovic, O'Malley and Sannuti (Ref. 26).

Transient phenomena are a recognized feature of the behavior of electric power systems. The study of transient phenomena is especially important in the analysis of power system stability because they have a strong influence on long-term behavior of the system (Ref. 27). Consequently, from concerns of power system stability and security, it is necessary to monitor both short-term and long-term behavior, which involves fast and slow dynamics (Ref. 28).

The application of SP theory to power system analysis is embodied in papers by Allemong, Chow, Kokotovic and Winkelman (Refs. 29 through 32). These papers deal with the simulation of power systems to study high-frequency oscillations and the effect of faults. The power systems studied are comprised of generators and loads and the emphasis is on the use of block diagonalizing transformations for separating slow and fast dynamics. In the context of this report the work of Kokotovic et al, needs to be extended to the development of estimation algorithms for power systems comprised of the boiler-turbine-exciter-generator-load complex.

This chapter examines the SP methodology for state estimation and its potential application to a boiler-turbine-exciter-generator complex. Section 4.2 contains a brief review of reduced-order state estimation using the method of multiple scales. Section 4.3 presents a preliminary analysis of a typical generating system model examined in the context of singular perturbations. Section 4.4 presents details of the equations to be implemented in a reduced-order state estimator for a non-linear system. The derivations of these equations is provided in Appendix E. Appendix F contains details of the model discussed in Section 4.3.

4.2 SINGULARLY PERTURBED ESTIMATION

A large body of literature has accumulated over the past decade on the analysis of singularly perturbed systems. Most of the research in this area has been devoted to simulation and control analyses of linear systems with the hope that the methods could

easily be adapted to the problem of linear state estimation. This section presents a brief outline of the SP estimation problem and its solution.

If the estimation problem is for the linear system

$$\begin{aligned}\dot{\underline{x}} &= D \underline{x} + E \underline{w} \\ \underline{y} &= H \underline{x} + \underline{v}\end{aligned}\quad (4.2-1)$$

where $\underline{w}$ and $\underline{v}$ are white noise processes, and the state vector $\underline{x}$ can be partitioned into "slow" and "fast" vectors, $\underline{x}_1$ and $\underline{x}_2$, then the problem can be reformulated in SP form. With n_1 slow states and n_2 fast states, define the following partitions

$$\underline{x}^T = [\underline{x}_1^T \quad \underline{x}_2^T] \quad (4.2-2)$$

$$D = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \quad (4.2-3)$$

$$E = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} \quad (4.2-4)$$

$$H = [H_1 \quad H_2] \quad (4.2-5)$$

Implicit in the partition of Eqs. 4.2-2 and 4.2-3 is the assumption that the state vector can be permuted so that the states are ordered by increasing "rapidity." More precisely, if

$$\bar{D}_{11} = D_{11} - D_{12} D_{22}^{-1} D_{21} \quad (4.2-6)$$

the system of Eq. 4.2-1 can be formulated in SP form if the "smallest" eigenvalue of D_{22} is distinctly larger than the "largest" eigenvalue of $\bar{D}_{11}$.

Once the system is partitioned as described above, the fast dynamics may be scaled by a multiplying factor $\mu \ll 1$. One way of estimating this scale factor is to take the ratio of the geometric means of the eigenvalues of the slow and fast systems as follows

$$\mu \approx \frac{|\det \bar{D}_{11}|^{1/n_1}}{|\det D_{22}|^{1/n_2}} \quad (4.2-7)$$

Upon scaling by μ , the fast dynamics may be written as

$$\mu \dot{\underline{x}}_2 = F_{21} \underline{x}_1 + F_{22} \underline{x}_2 + \mu E_2 \underline{w} \quad (4.2-8)$$

where

$$F_{21} = \mu D_{21}, \quad F_{22} = \mu D_{22} \quad (4.2-9)$$

Next, a number v and two integers m and n must be found which satisfy the following constraints

$$v^m = \mu \quad (4.2-10)$$

$$0 < m \leq 2n \quad (4.2-11)$$

$$G_2 = v^{m-n} E_2 = 0(1) \quad (4.2-12)$$

Finally, the new scale factor, v , is parameterized by the variable ε and the problem of Eq. 4.2-1 is embedded in the formulation

$$\begin{aligned} \dot{\underline{x}}_1 &= F_{11}\underline{x}_1 + \varepsilon F_{12}\underline{x}_2 + G_1 \underline{w} \\ \varepsilon^m \dot{\underline{x}}_2 &= F_{21}\underline{x}_1 + F_{22}\underline{x}_2 + \varepsilon^n G_2 \underline{w} \end{aligned} \quad (4.2-13)$$

$$\underline{z} = H_1 \underline{x}_1 + H_2 \underline{x}_2$$

where

$$0 \leq \varepsilon < 1$$

$$F_{11} = D_{11}, \quad F_{12} = D_{12}/v, \quad G_1 = E_1 \quad (4.2-14)$$

The system of Eq. 4.2-13 results in a family of estimators parameterized by ε , the estimator of interest is the one with $\varepsilon=v$. While v is a number (the time-scale factor) ε is a parameter of the problem. This distinction is made so that F_{21} , F_{22} , G_2 , which are dependent on v , can be considered in the SP analysis as being independent of ε .

The constraint of Eq. 4.2-11 arises from consideration of the SP Riccati equation of the Kalman filter. If $m > 2n$, the SP form of the estimation problem is ill-posed and will result in a non-uniform asymptotic expansion for the desired solution (Ref. 33). A typical realization of this constraint has $m = 1$ and $n = 1$. In the rest of the analysis presented here, it is assumed that the fast dynamics are reformulated as

$$\varepsilon \dot{\underline{x}}_2 = F_{21} \underline{x}_1 + F_{22} \underline{x}_2 + \varepsilon G_2 \underline{w} \quad (4.2-15)$$

For the estimation problem involving continuous dynamics and discrete measurements, the differential equations for extrapolation of the state estimate, $\hat{\underline{x}}$, and its covariance, P , are

$$\begin{aligned} \dot{\hat{\underline{x}}} &= D \hat{\underline{x}} \\ \dot{P} &= DP + PD^T + EQE^T \end{aligned} \quad (4.2-16)$$

where Q is the spectral density matrix for the vector white noise process $\underline{w}$. The corresponding equations for the SP system of Eq. 4.2-13 are

$$\dot{A} = F_{11}A + AF_{11}^T + \varepsilon(F_{12}B^T + BF_{12}^T) + G_{11}$$

$$\varepsilon \dot{B} = AF_{12}^T + BF_{22}^T + \varepsilon F_{11}B + \varepsilon^2 F_{12}C + \varepsilon G_{12}$$

$$\varepsilon \dot{C} = F_{21}B + B^T F_{21}^T + F_{22}C + CF_{22}^T + \varepsilon G_{22} \quad (4.2-17)$$

$$\dot{\hat{\underline{x}}}_1 = F_{11}\hat{\underline{x}}_1 + \varepsilon F_{12}\hat{\underline{x}}_2$$

$$\varepsilon \dot{\hat{\underline{x}}}_2 = F_{21}\hat{\underline{x}}_1 + F_{22}\hat{\underline{x}}_2$$

where

$$P = \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \quad (4.2-18)$$

Equation 4.2-17 is to be solved subject to the initial conditions

$$\hat{\underline{x}}^T(0) = [\underline{a}^{*T} \quad \underline{b}^{*T}] \quad (4.2-19)$$

$$P(0) = \begin{bmatrix} A^* & B^* \\ B^{*T} & C \end{bmatrix} \quad (4.2-20)$$

The general technique for solving equations involving "fast" and "slow" modes is to examine the dynamics in two time-scales. The method of multiple scales analyzes the problem in two time-scales that are considered to be independent variables. The two scales used in solving Eq. 4.2-17 are

$$t_0 = t \quad \text{and} \quad t_1 = \varepsilon^{-1}t \quad (4.2-21)$$

These time-scales imply the equivalence

$$\varepsilon \frac{d}{dt} = \varepsilon \frac{\partial}{\partial t_0} + \frac{\partial}{\partial t_1} \quad (4.2-22)$$

so that Eq. 4.2-17 is transformed to a set of partial differential equations in the independent variables t_0 and t_1 . The next step in the analysis is to assume the form of the solution as

$$\begin{aligned} \hat{\underline{x}}_1 &= \sum \underline{a}_k(t_0, t_1) \varepsilon^k \\ \hat{\underline{x}}_2 &= \sum \underline{b}_k(t_0, t_1) \varepsilon^k \\ A &= \sum A_k(t_0, t_1) \varepsilon^k \quad (k=0, 1, \dots) \\ B &= \sum B_k(t_0, t_1) \varepsilon^k \\ C &= \sum C_k(t_0, t_1) \varepsilon^k \end{aligned} \quad (4.2-23)$$

The solution is completed by evaluating the functions $\underline{a}_k$, $\underline{b}_k$, A_k , B_k , C_k . Equations governing these functions are derived in Appendix E.

An important feature of the SP methodology is that the solution is obtained in the form of an asymptotic series (Ref 34) which may or may not be convergent. However, if the formalisms of asymptotics are realized in the analysis, the series will be a convergent-divergent series at worst. This means that typical series solutions converge to provide excellent approximations when only the first few terms of the series are considered, when more terms are evaluated, the approximation may degrade in an oscillatory or divergent fashion. In Appendix E, equations are presented for the evaluation of the first two terms of the solution series. Thus, for example, the resulting solutions have the form

$$\hat{x}_1 \approx \underline{a}_0(t_0, t_1) + \varepsilon \underline{a}_1(t_0, t_1) \quad (4.2-24)$$

$$C \approx C_0(t_0, t_1) + \varepsilon C_1(t_0, t_1)$$

etc.

The method of multiple scales is especially suited to the asymptotic approach to the problem, since it is general enough to permit the evaluation of as many terms of the series as needed. In the present formulation (Eq. 4.2-17) the solution may be extended to include the first two terms of the series, after which point further terms serve to degrade the solution and analysis becomes cumbersome.

Clearly, the SP methodology presents a trade-off between accuracy and computational savings. The exact (high-order) solution of Eq. 4.2-16 involves N_h differential equations, where

$$N_h = (n_1 + n_2)(n_1 + n_2 + 3)/2 \quad (4.2-25)$$

A k -term reduced-order solution, similar to the two-term solution of Appendix E, involves N_r differential equations, where

$$N_r = k n_1(n_1 + 3)/2 \quad (4.2-26)$$

Additionally, the reduced-order solution requires the inversion of two matrices, one of size n_2 and one other of size $n_2(n_2 + 1)/2$. For the time-invariant estimation problem these inversions may be done offline, and the computational savings may be estimated by comparing N_h and N_r . In the time-varying case, the inversions must be performed periodically depending on how rapidly F_{22} varies. In either case, the reduced-order solution permits much larger integration time steps than does the high-order solution. The savings due to this effect can be considerable.

4.3 SINGULARLY PERTURBED MODEL OF A GENERATING SYSTEM

A model of a generating system is examined in this section in the context of the SP formulation. The system is comprised of a boiler, a turbine, an

excitation system and a generator. The boiler and turbine models are derived from GE and EPRI reports (Refs 35 and 36) and do not include limit and dead-band non-linearities. They differ from the models of Appendix A in that they encompass both long-term and transient effects. Models for the excitation system and the generator are taken from Ref 37 and are identical to those presented in Appendix A.

The boiler-turbine-exciter-generator (BTEG) model has 30 state variables. The boiler and its controls are represented by states x_1 through x_{17} . Of these, states x_8 through x_{17} are artificial variables introduced to approximate the fuel-air supply transport delay inherent in the boiler model (Ref. 36). States x_{18} through x_{23} are related to the turbine and its governor. The excitation system is represented by states x_{24} through x_{26} and the generator by states x_{27} through x_{30} . Details of the dynamic relationships between the 30 state variables are given in Appendix F. A set of parameter values for the BTEG model was chosen from tables provided in Refs. 35 through 37.

An initial eigenanalysis of the BTEG model of Appendix F gives an idea of the range and separation of the eigenvalues of the system. The eigenvalues are shown in Table 4.3-1 and numbered in decreasing order of their real parts. Fig. 4.3-1 shows these eigenvalue locations in the second quadrant of the complex plane. Examination of the eigenvalue distribution in Fig. 4.3-1 suggests the four partitions (25, 5), (23, 7), (19, 11), (7, 23). The first number in each of these pairs is the number of "slow" states while the second is the number of "fast" states.

TABLE 4.3-1
EIGENVALUES OF THE BTEG MODEL

NO.	EIGENVALUE
1	-0 0
2	-0 005861
3,4	-0 017670 ±j 0.149407
5,6	-0.018628 ±j 0.009109
7	-0.033333
8,9	-0 159117 ±j 1.31710
10,11	-0.162215 ±j 0 001439
12,13	-0.163898 ±j 0.003782
14,15	-0 166647 ±j 0.004700
16,17	-0.169424 ±j 0.003819
18,19	-0.171149 ±j 0.001462
20,21	-0 260936 ±j 1 00610
22,23	-0.290482 ±j 0 032617
24,25	-0.589053 ±j 0.739933
26	-1.0
27	-1.26645
28,29	-1 92024 ±j 0 368708
30	-3.21399

The partitions indicated above were deduced from visual examination of Fig. 4.3-1. They are shown in the figure and in Table 4.3-1 by broken lines. The criterion used in choosing these partitions is that the smallest real part among the eigenvalues of the "slow" states be larger than the largest real part among the eigenvalues of the "fast" states. Moreover, there must be a distinct separation of the eigenvalues across the partition.

Figure 4.3-1 Plot of BTEG Model Eigenvalues

While it is very straightforward to identify possible partitions based on eigenvalue distributions, it is quite a difficult task to realize these partitions when the dynamics matrix has significant off-diagonal terms. The procedure for reformulating the dynamics in the SP form of Eq. 4.2-13 involves two phases -- permutation and partition. Permutation consists of reordering the state vector so that the states are arranged in order of increasing "rapidity". Partitioning consists of finding a separation between "slow" and "fast" states such that the eigenvalues of the resulting matrix F_{22} all lie to the left of the eigenvalues of the matrix $\bar{F}_{11}$ when plotted on the complex plane. Both of these tasks are difficult to do systematically -- there being no algorithm in existence to correlate the states of a linear system with the eigenvalues of the system.

Lacking a systematic methodology for permutation and partitioning according to the criteria required by SP formalisms, one is forced to resort to a trial and error procedure. This consists of first reordering the state variables intuitively, randomly or systematically, then determining the eigenvalues of F_{22} and $\bar{F}_{11}$, and finally comparing these eigenvalues to those computed for the high order system, shown in Table 4.3-1 and Fig. 4.3-1. The total number of permutations and partitions that need to be examined in this fashion is so large as to prohibit an exhaustive computer analysis. However, in the course of such trial and error analyses, it is possible to identify groups of states that influence groups of eigenvalues. This

knowledge can be used effectively in subsequent trials to eliminate improper permutations and partitions.

For the BTEG model under consideration, trial and error analyses provided some insight into the relation between state variables and eigenvalues. Table 4.3-2 displays the deduced relationships. An examination of the groups shown in this table suggests that the (7:23) partition, indicated by the broken line, should separate the "slow" and "fast" states with no overlap.

TABLE 4.3-2
RELATION BETWEEN STATES AND EIGENVALUES

STATE VARIABLES	EIGENVALUE NUMBERS
x_{24}, x_{25}, x_{26}	30, 9, 8
x_{28}, x_{30}	29, 28
x_4, x_{18}, x_{22}	27, 25, 24
x_2, x_{19}, x_{20}	26, 23, 22
x_{27}, x_{29}	21, 20
$x_8, x_9, \quad, x_{16}, x_{17}$	19, 18, \quad, 11, 10
x_7	7
x_3, x_{23}	6, 5
x_5, x_6	4, 3
x_1	2
x_{21}	1

Table 4.3-3 provides a comparison of the exact system eigenvalues with those obtained from the SP approximation with the (7:23) partition suggested by Table 4.3-2. The most significant feature of this partition is that the approximations for the "fast" eigenvalues are better than those for the "slow" eigenvalues.

Since the SP methodology of Section 4.2 and Appendix E examines a reduced order "slow" system, it is important to ensure that this reduced order system retains the "slow" eigenvalues to some degree of accuracy. The partition of Table 4.3-3 does not meet this requirement and might therefore prove unsuitable. An alternative partition, better suited to the SP methodology, was found after further intuitive trial and error analyses. The eigenvalues of this (19:11) partition are shown in Table 4.3-4. Although the (19:11) partition appears less effective than the (7:23) partition, it provides significant savings in computation from both of the considerations discussed in Section 4.2.

Considering that state x_{21} contributes a zero eigenvalue, the partition of Table 4.3-4 is effectively an (18:11) partition. From Eqs. 4.2-25 and 4.2-26, the number of filter differential equations required to be solved for the high-order system, N_h , and for the SP system, N_r , are

$$N_h = 464 \quad N_r = 378$$

In addition to reducing the number of equations to be solved, the (19:11) partition can be used to eliminate 11 "fast" stable states and consequently allow for a larger integration step size.

TABLE 4.3-3
EIGENVALUES OF THE (7.23) PARTITION

STATES	APPROXIMATE EIGENVALUES	EXACT EIGENVALUES
x_{21}	0.0	0.0
$\begin{bmatrix} x_1 \\ x_{23} \\ x_3 \end{bmatrix}$	-0.000932 -0 019941 -0 120325	-0.005861 -0.018628 $\pm j$ 0.009109
x_5, x_6	0 019252 $\pm j$ 0.163803	-0.017670 $\pm j$ 0.149407
x_7	-0.033333	-0.033333

$\begin{bmatrix} x_8 \\ x_9 \\ \cdot \\ \cdot \\ \cdot \\ x_{16} \\ x_{17} \end{bmatrix}$	-0.164308 $\pm j$ 0.000846 -0.165393 $\pm j$ 0.002146 -0.166968 -0.167060 $\pm j$ 0.002446 -0.168532 $\pm j$ 0.001599 -0.169114	-0.162215 $\pm j$ 0.001439 -0.163898 $\pm j$ 0.003782 -0.166647 $\pm j$ 0.004700 -0.169424 $\pm j$ 0.003819 -0.171149 $\pm j$ 0.001462
x_{25}, x_{26}	-0.159117 $\pm j$ 1.31710	-0.159117 $\pm j$ 1.31710
x_{27}, x_{29}	-0.261161 $\pm j$ 1.00429	-0.260936 $\pm j$ 1.00610
x_2	-0.277778	-0.290482 $\pm j$ 0.032617
x_{19}	-0.282220	
x_{18}, x_{22}	-0.590513 $\pm j$ 0 750603	-0.589053 $\pm j$ 0.739933
x_{20}	-1 0	-1 0
x_4	-1.26606	-1.26645
x_{28}, x_{30}	-1.92456 $\pm j$ 0.377219	-1.92024 $\pm j$ 0.368708
x_{24}	-3.21399	-3.21399

TABLE 4.3-4
EIGENVALUES OF THE (19:11) PARTITION

STATES	APPROXIMATE EIGENVALUES	EXACT EIGENVALUES
x_{21}	0.0	0.0
x_1	-0.005843	-0.005861
x_3, x_{23}	-0.018372 $\pm j$ 0.009044	-0.018628 $\pm j$ 0.009109
x_7	-0.033333	-0.033333
x_5, x_6	-0.020111 $\pm j$ 0.150128	-0.017670 $\pm j$ 0.149407
$\begin{bmatrix} x_8 \\ x_9 \\ \cdot \\ \cdot \\ \cdot \\ x_{16} \\ x_{17} \end{bmatrix}$	-0.162658 -0.163454 $\pm j$ 0.002380 -0.165503 $\pm j$ 0.003794 -0.167949 $\pm j$ 0.003709 -0.169834 $\pm j$ 0.002399 -0.170527	-0.162215 $\pm j$ 0.001439 -0.163898 $\pm j$ 0.003782 -0.166647 $\pm j$ 0.004700 -0.169424 $\pm j$ 0.003819 -0.171149 $\pm j$ 0.001462
$\begin{bmatrix} x_{19} \\ x_2 \end{bmatrix}$	-0.235458 -0.276647	-0.290482 $\pm j$ 0.032617

x_{25}, x_{26}	-0.159117 $\pm j$ 1.31710	-0.159117 $\pm j$ 1.31710
x_{27}, x_{29}	-0.236655 $\pm j$ 0.989079	-0.260936 $\pm j$ 1.00610
x_{18}, x_{22}	-0.664300 $\pm j$ 0.871552	-0.589053 $\pm j$ 0.739933
x_{20}	-1.0	-1.0
x_4	-1.27413	-1.26645
x_{28}, x_{30}	-1.94986 $\pm j$ 0.364753	-1.92024 $\pm j$ 0.368708
x_{24}	-3.21399	-3.21399

The conclusions to be drawn from the example studied in this section are twofold. First, there is a need for developing systematic computer algorithms which can permute and partition an arbitrary dynamics matrix in a manner compatible with the SP methodology for linear estimation problems. Secondly, the relative merits of different permutations and partitions cannot be predicated upon eigenanalysis alone and require a full-scale sub-optimal filter analysis to establish their efficacy in the context of the SP methodology.

A recently published idea for the construction of reduced order models is called selective model analysis (SMA - Ref. 38). Given a certain model behavior of interest, SMA provides a means of identifying the relevant states of the mode and suitably collapsing the remaining states to provide a low order model. This methodology is ideally suited to the permutation and partition problem discussed in this section. Applications of SMA to power system models are demonstrated in Ref. 38.

4.4 REDUCED ORDER ESTIMATOR DESIGN

The discussion of time-scale decomposition in Section 4.2 detailed the methodology of order reduction for large-scale systems and the application of singular perturbation (SP) theory to the sub-optimal estimation of a high-order linear system. A similar approach is presented here for the estimation of a high-order nonlinear system represented by the equations

$$\dot{\underline{x}} = \underline{f}(\underline{x}) + \underline{G}\underline{w} \quad (4.4-1)$$

$$\underline{y} = \underline{h}(\underline{x}) + \underline{v}$$

A standard approach to filtering for nonlinear systems is the extended Kalman filter which relies on a partial linearization of the system nonlinearities. For the case where the state has continuous dynamics and measurements are available at discrete times, the extended Kalman filter propagates the estimate and covariance between measurement times according to the following differential equations

$$\dot{\underline{\hat{x}}} = \underline{f}(\underline{\hat{x}}) \quad : \quad \underline{\hat{x}}(0) = \underline{x}^* \quad (4.4-2)$$

$$\dot{\underline{P}} = \underline{F}\underline{P} + \underline{P}\underline{F}^T + \underline{G}\underline{Q}\underline{G}^T \quad : \quad \underline{P}(0) = \underline{P}^* \quad (4.4-3)$$

where

$$\underline{F} = \frac{\partial \underline{f}(\underline{\hat{x}})}{\partial \underline{x}^T} \quad (4.4-4)$$

If the interval of propagation is small and the $\underline{F}$ -matrix elements vary slowly over this interval, Eq. 4.4-4 can be approximated by

$$\underline{F} = \frac{\partial \underline{f}(\underline{x}^*)}{\partial \underline{x}^T} \quad (4.4-5)$$

If upon examination of the linearized dynamics matrix, F , it is found that its eigenvalues have a wide spread and that the large eigenvalues have negative real parts, then the propagation equations may be amenable to order reduction by singular perturbation. A linearization of Eq. 4.4-2 is

$$\dot{\underline{x}} = \underline{f}(\underline{x}^*) - F\underline{x}^* + F\underline{\hat{x}} \quad (4.4-6)$$

Following the trial and error methodology described in Section 4.3 it is possible to reorder $\underline{x}$ and hence reorder the rows and columns of F such that the states $x_1, x_2, x_3, \dots, x_n$ are ordered according to increasing fastness of response. Then, the F matrix may be partitioned as

$$F = \begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix} \quad (4.4-7)$$

and the state vector partitioned as

$$\underline{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \quad \underline{\hat{x}}(0) = \begin{bmatrix} \underline{a}^* \\ \underline{b}^* \end{bmatrix} \quad (4.4-8)$$

so that Eq. 4.4-6 can be rewritten as

$$\begin{aligned} \dot{\underline{x}}_1 &= \underline{f}_1(\underline{a}^*, \underline{b}^*) - D_{11}\underline{a}^* - D_{12}\underline{b}^* + D_{11}\underline{\hat{x}}_1 + D_{12}\underline{\hat{x}}_2 \\ \dot{\underline{x}}_2 &= \underline{f}_2(\underline{a}^*, \underline{b}^*) - D_{21}\underline{a}^* - D_{22}\underline{b}^* + D_{21}\underline{\hat{x}}_1 + D_{22}\underline{\hat{x}}_2 \end{aligned} \quad (4.4-9)$$

Next, a scale factor $\mu \in (0,1)$ is found such that the largest eigenvalue of μD_{22} is approximately equal to the largest eigenvalue of $D_{11} - D_{12}D_{22}^{-1}D_{21}$. Setting

$$F_{11} = D_{11}, \quad F_{12} = \frac{D_{12}}{\mu} \quad (4.4-10)$$

$$F_{21} = \mu D_{21}, \quad F_{22} = \mu D_{22}$$

allows Eq. 4.4-9 to be rewritten as

$$\begin{aligned} \dot{\underline{x}}_1 &= \underline{f}_1(\underline{a}^*, \underline{b}^*) - F_{11}\underline{a}^* - \mu F_{12}\underline{b}^* + F_{11}\underline{\hat{x}}_1 + \mu F_{12}\underline{\hat{x}}_2 \\ \mu \dot{\underline{x}}_2 &= \mu \underline{f}_2(\underline{a}^*, \underline{b}^*) - F_{21}\underline{a}^* - F_{22}\underline{b}^* + F_{21}\underline{\hat{x}}_1 + F_{22}\underline{\hat{x}}_2 \end{aligned} \quad (4.4-11)$$

Instead of analyzing Eq. 4.4-11, SP methodology parameterizes the scale factor μ by a variable ε where $0 < \varepsilon \ll 1$. The distinction made here is that μ is a number used to scale the matrices D_{12}, D_{21}, D_{22} , while ε is a parameter. The filter equations are analyzed in terms of the parameter ε to obtain solution equations governing terms of a truncated power series in ε . The perturbation parameter ε is set equal to μ in the filter implementation. This device permits the use, in

analysis, of the notion $0 < \varepsilon \ll 1$, while retaining the significance of the scale factor μ , which is merely a positive number less than unity. The resulting SP equations for state estimate and covariance propagation are

$$\begin{aligned} \dot{\underline{x}}_1 &= \underline{f}_1(\underline{a}^*, \underline{b}^*) + F_{11}(\underline{\hat{x}}_1 - \underline{a}^*) + \varepsilon F_{12}(\underline{\hat{x}}_2 - \underline{b}^*) \\ \varepsilon \dot{\underline{x}}_2 &= \varepsilon \underline{f}_2(\underline{a}^*, \underline{b}^*) + F_{21}(\underline{\hat{x}}_1 - \underline{a}^*) + F_{22}(\underline{\hat{x}}_2 - \underline{b}^*) \end{aligned} \quad (4.4-12)$$

$$\begin{aligned} \dot{A} &= F_{11}A + AF_{11}^T + \varepsilon(F_{12}B^T + BF_{12}^T) + G_1QG_1^T \\ \varepsilon \dot{B} &= AF_{12}^T + BF_{22}^T + \varepsilon F_{11}B + \varepsilon^2 F_{12}C + \varepsilon G_1QG_2^T \\ \varepsilon \dot{C} &= F_{21}B + F_{22}C + B^T F_{12}^T + CF_{22}^T + \varepsilon G_2QG_2^T \end{aligned} \quad (4.4-13)$$

where

$$A = P_{11}, \quad B = P_{12}, \quad C = P_{22} \quad (4.4-14)$$

The SP equations for estimate and covariance propagation between updates can be analyzed using the method of multiple scales. Equations 4.4-12 and 4.4-13 are examined in this context in Appendix E of this report to provide a set of reduced-order equations. A summary of the equations to be implemented is presented here.

The coarsest approximations to the propagation equations are the following

$$\begin{aligned} \dot{\underline{x}}_1 &= P_0 \\ \dot{\underline{x}}_2 &= -D_0(P_0 - \underline{a}^*) + \underline{b}^* \end{aligned} \quad (4.4-15)$$

$$\begin{aligned} A &= K_0 \\ B &= -K_0 D_0^T \\ C &= D_0 K_0 D_0^T \end{aligned} \quad (4.4-16)$$

where

$$D_0 = F_{22}^{-1}F_{21} = D_{22}^{-1}D_{21} \quad (4.4-17)$$

and P_0 and K_0 are governed by

$$\begin{aligned} \dot{P}_0 &= F_{11}(P_0 - \underline{a}^*) + \underline{f}_1(\underline{a}^*, \underline{b}^*) : P_0(0) = \underline{a}^* \\ \dot{K}_0 &= F_{11}K_0 + K_0 F_{11}^T + G_1QG_1^T : K_0(0) = A^* \end{aligned} \quad (4.4-18)$$

The next level of approximation involves the inclusion of the boundary layer correction terms to the estimate and covariance expressions. Thus

$$\begin{aligned}\dot{\underline{x}}_1 &= \underline{p}_0 \\ \dot{\underline{x}}_2 &= -D_0(\underline{p}_0 - \underline{a}^*) + \underline{b}^*\end{aligned}\quad (4.4-19)$$

as before, and

$$\begin{aligned}A &= K_0 \\ B &= -K_0 D_0^T + V_0 + F_{11} V_0 t \\ C &= D_0 K_0 D_0^T - D_0 V_0 - V_0^T D_0^T + \bar{W}_0\end{aligned}\quad (4.4-20)$$

where V_0 and $\bar{W}_0$ are given by

$$\begin{aligned}V_0 &= (B^* + A^* D_0^T) \Phi^T \\ \bar{W}_0 &= \Phi(C^* + D_0 B^* + B^{*T} D_0^T + D_0 A^* D_0^T) \Phi^T\end{aligned}\quad (4.4-21)$$

and

$$\Phi = \exp\left(\frac{F_{11}}{\mu}\right)t = \exp(D_{22}t) \quad (4.4-22)$$

Another approximation that neglects the boundary layer correction terms but takes into account the scale factor μ is

$$\begin{aligned}\hat{\underline{x}}_1 &= \underline{p}_0 + \mu \underline{p}_1 \\ \hat{\underline{x}}_2 &= -D_0(\underline{p}_0 - \underline{a}^*) + \underline{b}^* - D_{22}^{-1}[\underline{f}_2(\underline{a}^*, \underline{b}^*) + D_0 \underline{f}_1(\underline{a}^*, \underline{b}^*)] \\ &\quad - \mu[D_1(\underline{p}_0 - \underline{a}^*) + D_0 \underline{p}_1]\end{aligned}\quad (4.4-23)$$

$$\begin{aligned}A &= K_0 + \mu K_1 \\ B &= -K_0 D_0^T - \mu(K_1 D_0^T + K_0 D_1^T + E_1) \\ C &= D_0 K_0 D_0^T + \mu(D_0 K_1 D_0^T + D_1 K_0 D_0^T + D_0 K_0 D_1^T + N_1)\end{aligned}\quad (4.4-24)$$

where

$$\begin{aligned}\dot{\underline{p}}_1 &= F_{11} \underline{p}_1 - F_{12} D_0(\underline{p}_0 - \underline{a}^*) : \underline{p}_1(0) = 0 \\ \dot{\underline{K}}_1 &= F_{11} \underline{K}_1 + \underline{K}_1^T F_{11}^T - F_{12} D_0 \underline{K}_0 - \underline{K}_0 D_0^T F_{12}^T : \underline{K}_1(0) = 0\end{aligned}\quad (4.4-25)$$

$$D_1 = F_{22}^{-1} D_0 F_{11} \quad (4.4-26)$$

$$E_1 = G_1 Q(G_2^T + G_1^T D_0^T) F_{22}^{-T}$$

and N_1 is the symmetric solution of

$$F_{22} N_1 + N_1 F_{22}^T = F_{21} E_1 + E_1^T F_{21}^T + D_0 G_1 Q G_1^T D_0^T - G_2 Q G_2^T \quad (4.4-27)$$

Finally, the two-term asymptotic series approximation containing all the terms derived in Appendix E is

$$\begin{aligned}\hat{\underline{x}}_1 &= \underline{p}_0 + \mu \underline{p}_1 \\ \hat{\underline{x}}_2 &= -D_0(\underline{p}_0 - \underline{a}^*) + \underline{b}^* - D_{22}^{-1}[\underline{f}_2(\underline{a}^*, \underline{b}^*) + D_0 \underline{f}_1(\underline{a}^*, \underline{b}^*)] \\ &\quad + \mu[\underline{s}_1 - D_1(\underline{p}_0 - \underline{a}^*) - D_0 \underline{p}_1]\end{aligned}\quad (4.4-28)$$

$$\begin{aligned}A &= K_0 + \mu K_1 \\ B &= -K_0 D_0^T + V_0 + F_{11} V_0 t + \mu(\bar{V}_1 - K_1 D_0^T - K_0 D_1^T - E_1) \\ C &= D_0 K_0 D_0^T - D_0 V_0 - V_0^T D_0^T + \bar{W}_0 - D_0 F_{11} V_0 t - V_0^T F_{11}^T D_0^T t \\ &\quad + \mu(D_0 K_1 D_0^T + D_1 K_0 D_0^T + D_0 K_0 D_1^T + N_1 \\ &\quad + \bar{W}_1 - D_0 \bar{V}_1 - \bar{V}_1^T D_0^T - D_1 V_0 - V_0^T D_1^T)\end{aligned}\quad (4.4-29)$$

where

$$\begin{aligned}\underline{s}_1 &= \Phi F_{22}^{-1}[\underline{f}_2(\underline{a}^*, \underline{b}^*) + D_0 \underline{f}_1(\underline{a}^*, \underline{b}^*)] \\ \bar{V}_1 &= (E_1 + A^* D_1^T) \Phi^T \\ \bar{W}_1 &= \Phi(D_0 E_1 + E_1^T D_0 + D_1 B^* + B^{*T} D_1^T \\ &\quad + D_0 A^* D_1^T + D_1 A^* D_0^T - N_1) \Phi^T\end{aligned}\quad (4.4-30)$$

The boundary layer correction terms involve the functions V_0 , W_0 , $\underline{s}_1$, $\bar{V}_1$, $\bar{W}_1$. If the time between measurements is sufficiently large and if the eigenvalues of D_{22} have real parts that are sufficiently large and negative, then the transition matrix Φ between updates will approach zero. This being the case it is reasonable to ignore the boundary layer correction terms in the solution of the propagation equations. Therefore in the implementation of the reduced order filter, Eqs. 4.4-15 through 4.4-18 may be employed as a first approximation. An improved solution may result from implementation of Eqs. 4.4-23, 4.4-24, 4.4-17, 4.4-18 and 4.4-25 through 4.4-27. Obviously, if the scale factor μ is very small the first approximation may be sufficiently precise. If not, the second approximation may have to be used. It must be noted, however, that the second approximation may result in only moderate savings in computations. This obvious trade-off between accuracy and computational burden can only be investigated by actually testing the proposed sub-optimal filtering algorithms.

5 ESTIMATOR SIMPLIFICATION FOR TRANSIENT DYNAMICS

5.1 INTRODUCTION

Power system dynamics may be divided into two categories: transient and long-term. Transient dynamics consist of fast modes with time constants up to several seconds, and are governed by the interconnection network relations, generator electrical parameters and inertia, and the excitation and turbine speed controls. Due to the short time scales involved, it is desirable to develop a state estimator for the power system transient dynamic behavior with the following properties:

- Computational simplicity (e.g., no matrix inversions on-line)
- Known convergence properties (e.g., convergence within a bound).

Various approaches to this problem are examined in this chapter.

5.2 REVIEW OF PREVIOUS WORK ON TRANSIENT STATE ESTIMATION

Two principal approaches have been used in previous studies of transient state estimation for power systems. The first, employed by Ueda, Takata, et. al. (Refs. 39 through 41), assumes deterministic, nonlinear models for both the state and measurement processes:

$$\dot{\underline{x}}(t) = \underline{f}(\underline{x}(t)) \quad (5.2-1)$$

$$\underline{y}_k = \underline{h}(\underline{x}_k) \quad (5.2-2)$$

These equations are linearized by retaining only the first order terms in a Taylor series expansion. The state is estimated by an observer-like algorithm, resulting in equations very similar to those for the extended Kalman filter. Some additional refinement of the estimator is accomplished by local iterated relinearization (Ref. 41). Example problems for testing these estimators dealt with very simple models; typically a one machine infinite bus system (e.g., Ref. 40). Besides the obvious shortcoming of ignoring noise, this method of estimation entails all the computational complexity of the extended Kalman filter.

The second approach, used recently by Modir and Schlueter (Ref. 42), assumes a simple, linearized stochastic model for the power system and uses a continuous-discrete Kalman filter for state estimation. Simulation results are presented for a six machine system. The advantage of this approach is clear: the linear, low-dimensional system model and Kalman filter afford straightforward implementation with precomputation of filter gains. In addition, Modir and Schlueter's work indicates that a data collection system, based on presently available instruments* can measure the frequency at generator buses accurately enough and at sufficiently high sampling rates (20 per second) so that generator angle variations can be tracked even if contingencies or changes in system operating conditions result in an incorrect dynamic model. However, there has not been extensive

testing of this estimation technique to determine its applicability to large power systems. In addition, since Kalman filter performance depends significantly on the accuracy of the process and measurement models, the simplified model may introduce significant errors for a large system with many machines.

5.3 NONLINEAR ESTIMATION USING OBSERVERS

Both of the approaches to transient state estimation discussed above suffer from the shortcoming of inappropriate system models. Appendices A and B of this report contain more detailed and realistic system models -- one for transient dynamics and one for long-term dynamics. These models are nonlinear, so that the problem of transient state estimation is one of nonlinear estimation for the following system of equations:

$$\begin{aligned} d\underline{x}(t) &= \underline{f}[t, \underline{x}(t), \underline{z}(t), \underline{u}(t)] dt \\ &+ G[t, \underline{x}(t), \underline{z}(t), \underline{u}(t)] d\underline{n} \end{aligned} \quad (5.3-1)$$

$$\underline{y}_k = \underline{h}_k(\underline{x}(t_k)) + \underline{v}_k \quad (5.3-2)$$

The most common approach to the estimation problem posed above has been the linearization of equations and the use of the extended Kalman filter (cf. Chapter 2). Implementation of the extended Kalman filter in real-time presents serious computational problems because filter gains cannot be precomputed and thus must be calculated on-line, with the attendant matrix inversions. Since a principal requirement for a transient state estimator is computational simplicity, the following approach was investigated.

Tarn and Rasis (Ref. 43) have presented a solution to the nonlinear estimation problem by constructing nonlinear observers using a Lyapunov-like approach. They provide sufficient conditions on the structure of a nonlinear stochastic system for the existence of an exponentially bounded observer. While they present results for both continuous- and discrete-time representations, only the continuous-time results are stated here for brevity. They consider a dynamical system modeled by the Itô equation

$$d\underline{x}(t) = \underline{f}(\underline{x}(t)) dt + \sigma_1(\underline{x}(t)) d\underline{w}(t) \quad (5.3-3)$$

with observations

$$d\underline{y}(t) = \underline{h}(\underline{x}(t)) dt + \sigma_2(\underline{x}(t)) d\underline{v}(t) \quad (5.3-4)$$

where $\underline{x}(t)$ and $\underline{f}(\underline{x}(t))$ are n -dimensional vectors, $\underline{y}(t)$ and $\underline{h}(\underline{x}(t))$ are m -dimensional vectors, $\underline{f}(\cdot)$ and $\underline{h}(\cdot)$ are continuously differentiable, $\underline{w}(t)$ and $\underline{v}(t)$ are q - and r -dimensional uncorrelated standard Wiener processes, and $\sigma_1(\underline{x}(t))$ and $\sigma_2(\underline{x}(t))$ are matrix functions of appropriate dimensions. Both Eqs. 5.3-3 and 5.3-4 are assumed to satisfy a global Lipschitz condition and a linear growth condition. Tarn and Rasis postulate a nonlinear observer for $\underline{x}(t)$ which is linear in the observed data:

$$d\underline{z}(t) = \underline{f}(\underline{z}(t)) dt + B(d\underline{y}(t) - \underline{h}(\underline{z}(t)) dt) \quad (5.3-5)$$

where B is a constant $n \times m$ matrix.

*Electric Power Research Institute, Research Project RP849

If the covariances $\sigma_1(\underline{x}(t))$, $\sigma_2(\underline{x}(t))$ are bounded and if there exist an $n \times n$ positive definite symmetric matrix Q and a constant $n \times m$ matrix B such that $Q[\nabla f - B^T h + (\alpha/2) I]$ is uniformly negative semi-definite, then the error $\underline{\varepsilon}(t) \triangleq \underline{x}(t) - \underline{z}(t)$ is exponentially bounded in mean square with given exponent $\alpha > 0$. Specifically,

$$\lim_{t \rightarrow \infty} E \|\underline{\varepsilon}(t)\|^2 \leq \frac{k}{c\alpha}$$

where

$$k = \sup_{\underline{x}(t) \in R^n} \text{tr}[\sigma_1(\underline{x}(t)) \sigma_1^T(\underline{x}(t)) + B \sigma_2(\underline{x}(t)) \sigma_2^T(\underline{x}(t)) B^T] Q < \infty$$

and c is the minimum eigenvalue of Q .

Clearly, this scheme offers the advantages of simple implementation (with reduced computations), as well as guaranteed convergence within an error bound. However, there are several drawbacks:

(i) As with all Lyapunov-like schemes, it is necessary to choose a B and a Q to satisfy the conditions of the theorem. While the theorem guarantees convergence if such a B and a Q exist, it says nothing about how to choose such matrices.

(ii) Uniform negative semi-definiteness is a difficult property to check, which further compounds the problem of choosing the appropriate B and Q .

(iii) While an error bound is provided for the observer, it may have little practical value because it may be too large, since no criterion of optimality is used to derive it. Also, the observer structure, determined by the matrices B and Q , makes no use of the noise covariances $\sigma_1^2(\underline{x}(t))$ and $\sigma_2^2(\underline{x}(t))$.

In view of these disadvantages, this approach is not useful for the transient state estimation problem.

5.4 STOCHASTIC APPROXIMATION APPROACH

Consider the following problem: estimate a constant p -dimensional vector $\underline{x}$, using the set of scalar observations y_n , $n=1,2,\dots$ defined as

$$y_n = H_n(\underline{x}) + v_n \quad (5.4-1)$$

where $\{H_n(\cdot)\}$ $n=1,2,\dots$, is known and v_n is the measurement error. A stochastic approximation scheme for estimating $\underline{x}$ may be written as

$$\underline{t}_{n+1} = \underline{t}_n + a_n(y_n - H_n(\underline{t}_n)) \quad n=1,2,\dots \quad (5.4-2)$$

In Eq. 5.4-2, $\underline{t}_1$ is arbitrary and $\{a_n\}$ is a sequence of p -dimensional gain vectors which must be suitably chosen to ensure convergence of $\{\underline{t}_n\}$ to $\underline{x}$. When $p=1$, the condition on $\{a_n\}$ are simple to derive; in higher dimensions the problem is considerably more complex. However, for any stochastic approximation, the gains $\{a_n\}$ must possess two properties:

(i) They should be signed properly, that is to say, at each successive observation, the correction to be applied to $\underline{t}_n$ must be in the correct direction (towards $\underline{x}$).

(ii) They should be "sufficiently" large to ensure convergence to $\underline{x}$.

As may be evident from an examination of Eq. 5.4-2, stochastic approximation techniques have an inherent appeal because of their simple structure and independence, in general, of the noise statistics. While the former property is desirable for filtering (and other) applications, the latter is a shortcoming for any application in which some sort of optimality or error minimization is desired.

The immediate question is whether or not stochastic approximation is applicable to the transient state estimation problem defined by nonlinear, stochastic equations for the state and measurements. For the sake of simplicity, assume the following scalar, discrete-time representation for the system.

$$\underline{x}_{n+1} = F_n(\underline{x}_n) + w_n \quad (5.4-3)$$

$$y_n = H_n(\underline{x}_n) + v_n \quad (5.4-4)$$

where

$\underline{x}_n$ = state of the system

y_n = output measurement

w_n = noise in the state dynamics

v_n = measurement noise.

The question then becomes the following: is it possible to estimate $\underline{x}_n$ using a stochastic approximation scheme such as

$$\underline{t}_{n+1} = F_n(\underline{t}_n) + a_n(y_n - H_n(\underline{t}_n)) \quad (5.4-5)$$

where $\{a_n\}$ are scalars such that $a_n \rightarrow 0$ as $n \rightarrow \infty$.

Unfortunately, it turns out that stochastic approximation is not applicable when the parameter to be estimated is subject to random disturbances. In this case, the "parameter" to be estimated is the state $\underline{x}_n$, whose dynamics are governed, in part, by the noise w_n ; hence, stochastic approximation cannot be applied to this problem. The reason stochastic approximation techniques cannot be applied to the estimation of parameters subject to random disturbances is described by Sprevak and Newmann (Ref. 44). The general condition for the stochastic approximation gains (a_n of Eq. 5.4-5) have been presented by Dvoretzky (Ref. 45), who derived these conditions for the estimation of a stationary parameter. These conditions may possibly guarantee convergence of the stochastic approximation algorithm for a changing parameter in the case where such changes are completely deterministic and can be predicted. However, if $a_n \rightarrow 0$, as required of all stochastic approximation gains, then it is impossible for the algorithm of Eq. 5.4-5 to track the variations in $\underline{x}_n$ as n

increases. For systems described by Eq. 5.4-4, it is well known that estimation algorithms of the form given in Eq. 5.4-4, must have gains a_n which do not approach zero, the Kalman filter is the most obvious illustration of this principle

Consequently, stochastic approximation techniques are not applicable to the transient state estimation problem (for the system of Eqs. 5.3-1 and 5.3-2). However, making the restrictive assumption that the state equation is free of noise allows the use of some applicable results from earlier research. This is discussed in the next section.

5.5 NONLINEAR ESTIMATION BY STOCHASTIC APPROXIMATION IN ABSENCE OF PROCESS NOISE

Several researchers have applied stochastic approximation techniques to the nonlinear state estimation problem when the state dynamics are assumed to be noise free. Both Pearson (Ref. 46) and Wolverton (Ref. 47) consider the following scalar, nonlinear system with linear measurements:

$$x_{n+1} = g(x_n) \quad n = 0, 1, 2, \dots \quad (5.5-1)$$

$$y_n = h x_n + v_n \quad (5.5-2)$$

where

x_n = state of the system

$g(\cdot)$ = nonlinear function

y_n = output measurements

h = constant > 0

(v_n) = sequence of independent, identically distributed random variables with variance σ^2 representing measurement noise.

Both authors use the following stochastic approximation scheme to estimate x_n :

$$\hat{x}_{n+1} = g(\hat{x}_n) + a_{n+1} \frac{h}{\sigma^2} [y_{n+1} - h g(\hat{x}_n)] \quad (5.5-3)$$

$$n = 0, 1, 2, \dots$$

where $\hat{x}_n$ is the estimate of x_n at time n ; $\hat{x}_0$ is chosen arbitrarily.

The $\{a_1\}$ are a sequence of positive numbers satisfying

$$\sum_{i=1}^{\infty} a_i = \infty \quad \text{and} \quad \sum_{i=1}^{\infty} a_i^2 < \infty$$

Imposing the additional restriction

$$|g(\hat{x}_n) - g(x_n)| \leq |\hat{x}_n - x_n| \quad (5.5-4)$$

allows both authors to prove convergence of $\hat{x}_n$ to x_n in the mean square.

Albert and Gardner (Ref. 48) consider the following scalar system with nonlinear state and measurement equations.

$$x_{n+1} = F_n(x_n) + w_n \quad (5.5-5)$$

$$y_n = H_n(x_n) + v_n \quad (5.5-6)$$

where F_n , H_n are known, nonlinear functions of their argument, and $\{w_n\}$ and $\{v_n\}$ are mutually and temporally independent zero-mean stochastic processes

Assume the following conditions:

A1: The noise processes w_n and v_n have finite, non-zero variances σ_w^2 and σ_v^2

A2: The derivatives of $\dot{F}_n(x)$ and $\dot{H}_n(x)$ are one-signed for each n

A3: For all x , $\beta'_n \leq |\dot{F}_n(x)| \leq \beta_n$, and $b'_n \leq |\dot{H}_n(x)| \leq b_n$, where $\beta_n \leq c_1 \beta'_n$ and $b'_n \leq c_2 b_n$ for some $1 \leq c_1, c_2 < \infty$

A4: $\beta = \limsup_n \beta_n < \frac{c}{c-1}$, where $c = c_1 c_2$

A5: $b_n \rightarrow \infty$ as $n \rightarrow \infty$

Then, t_n estimates x_n with mean square convergence where t_n is given by the stochastic approximation (with t_1 arbitrary)

$$t_{n+1} = F_n(t_n) + a_n [y_n - H_n(t_n)] \quad (5.5-7)$$

$$n = 1, 2, 3, \dots$$

$$a_n = \text{sgn}(\dot{F}_n \dot{H}_n) \frac{\beta_n}{c b_n} \quad (5.5-8)$$

The mean square prediction error $\varepsilon_n^2 = E\{[t_n - x_n]^2\}$ satisfies

$$\limsup_n \varepsilon_n^2 \leq \frac{\sigma_w^2}{1 - \beta^2(1 - 1/c)^2} \quad (5.5-9)$$

with equality in the linear case ($c=1$).

Assumption A4 does not require stability of the state equation $x_{n+1} = F_n(x_n) + w_n$, except in the limit of arbitrarily large c_1 or c_2 . At the other extreme, when $\beta'_n = \beta_n = \beta$ and $b'_n = b_n$ (the linear case), assumption A4 admits systems of the type $x_{n+1} = \pm \beta x_n + w_n$ with $\beta < \infty$ arbitrary. In such cases, the result is exact, i.e.,

$$\lim_n ||\varepsilon_n||^2 = \sigma_w^2 \quad (5.5-10)$$

This is the same error obtained at each time n if the state were directly observable.

If, instead of Assumptions A4 and A5, one assumes that

$$A4': \beta < 1$$

$$A5': b_n \rightarrow 0 \text{ as } n \rightarrow \infty$$

Then, choosing $a_n = \text{sgn}(\dot{F}_n \dot{H}_n) a \frac{\beta_n}{1 - \beta_n^2} b_n$, ($a > 0$) leads to

$$\lim_n \sup \varepsilon_n^2 \leq \frac{\sigma_v^2}{1 - \beta^2} \quad (5.5-11)$$

with equality (in Eq. 5.5-11) for $\beta'_n = \beta_n = \beta$. This is just the mean-square error that results from not using the observations at all; in other words, $t_{n+1} = \pm \beta t_n$ performs as well as the t_{n+1} given by Eq. 5.5-7

References 46 through 48 are solely concerned with the scalar state estimation problem. Fujita and Fukao (Ref. 49) have presented results for the vector case, with both linear and nonlinear measurements. Their results for the case of linear measurements are discussed below.

Consider the following state and measurement equations:

$$x_{n+1} = F_n(x_n) \quad (5.5-12)$$

$$y_n = H_n x_n + v_n$$

where

$x_n = n \times 1$ state vector at time n

$F_n =$ known $n \times n$ matrix

$y_n = m \times 1$ measurement vector at time n

$H_n =$ known $m \times n$ matrix

$\{v_n\} =$ measurement noise sequence with

$$E\{v_n\} = 0; \quad E\{v_j v_k^T\} = R_j \delta_{jk},$$

$$\sup_j ||R_j|| < \infty \text{ with each } R_j$$

positive definite

They use the following stochastic approximation for x_n :

$$t_{n+1} = F(t_n) + a_{n+1} H_{n+1}^T R_{n+1}^{-1} [y_{n+1} - H_{n+1} F(t_n)] \quad (5.5-13)$$

$$n = 0, 1, 2, \dots$$

where t_0 is chosen arbitrarily, and $\{a_n\}$ satisfies conditions as stated for the scalar case in Section 5.4.

The t_n defined by Eq. 5.5-13 converges to x_n in the mean square under the following two assumptions

$$A1: ||F_n(x_1) - F_n(x_2)|| \leq \gamma_n ||x_1 - x_2||,$$

$$0 < \gamma_n \leq 1$$

$$A2: 0 < \inf_n \lambda_{\min}(H_n^T R_n^{-1} H_n) < \sup_n \lambda_{\max}(H_n^T R_n^{-1} H_n) < \infty$$

where $\lambda_{\min}$ and $\lambda_{\max}$ denote the minimum and maximum eigenvalues, respectively

If the measurement noise is stationary and the measurement matrix constant, assumption A2 reduces to

$$A2'. \quad H^T R^{-1} H \text{ is positive definite}$$

Finally, it should be noted that the vector non-linear estimation problem with noise in the state dynamics has been solved by de Figueiredo and Dyer (Ref. 50) using modified stochastic approximation arguments. However, their algorithm is not a stochastic approximation technique in the sense that the gain sequence does not converge to zero. Rather, the authors term the resulting algorithm a "mini-max Kalman filter." The filter is specified in terms of a sequence gains which maximize the speed of convergence of the estimation error. The resulting expressions for the gain sequence involve computations very similar to those required for the extended Kalman filter (cf. Section 2.4) and consequently do not afford a significant reduction in computational burden.

6.1 INTRODUCTION

The role of system state forecasting with regard to the security assessment problem is to anticipate evolution of system security. The basic idea is to predict the system state estimate (and/or probability distribution) from the present time t_0 to a future time t_f . The predicted evolution of the state and statistics to t_f will provide the operator with an indication of future security needs.

This chapter documents research into the area of power system state forecasting. Emphasis has been placed on obtaining explicit expressions for the first and second moments of the forecasted system state. Attention in this chapter is focussed on the finite time properties of solutions to equations of the form

$$\frac{d\mathbf{x}(t)}{dt} = \mathbf{F}(t)\mathbf{x}(t), \quad \mathbf{x}(t_0) = \mathbf{x}_0 \quad (6.1-1)$$

where $\mathbf{F}(t)$ is a piecewise constant matrix which represents possible configurations of the power system. In order to introduce the possibility of system topological change, it is assumed that there exist random times $t_0 < t_1 < t_2 < \dots$, at which $\mathbf{F}(t)$ jumps from one value to another. To make this more precise, set

$$\mathbf{F}(t) = \mathbf{F}_k, \quad t_k \leq t < t_{k+1} \quad (6.1-2)$$

where $\mathbf{F}_k$ is a (constant) random matrix which can take values from a finite set of matrices $\{\mathbf{F}^{(1)}, \mathbf{F}^{(2)}, \dots, \mathbf{F}^{(N)}\}$. The time interval $[t_k, t_{k+1})$ is the sojourn time of the coefficient $\mathbf{F}_k$. It is assumed that the statistical properties of the sojourn time are known.

6.2 EVALUATION OF THE FIRST MOMENT

An expression for the first moment of the solution process of Eq. 6.1-1, denoted by $E\{\mathbf{x}(t)\}$, is developed in this section. Clearly, Eq. 6.1-1 cannot be integrated directly as an arbitrary time-varying linear system. However, it is possible to take advantage of the fact that $\mathbf{F}(t)$ is piecewise constant.

For each t , there exists a random integer $K(t)$ such that (for $t_0 = 0$)

$$0 = t_0 < t_1 < t_2 < \dots < t_{K(t)} < t$$

[Although it is possible for $t_{K(t)} = t$, the probability of having this occur is zero and hence it may be neglected.] A sample solution to Eq. 6.1-1 may be written, for $K(t) = k$, as

$$\mathbf{x}(t) = e^{\mathbf{F}_k(t-t_k)} e^{\mathbf{F}_{k-1}(t_k-t_{k-1})} \dots e^{\mathbf{F}_1(t_2-t_1)} e^{\mathbf{F}_0(t_1-t_0)} \mathbf{x}_0 \quad (6.2-1)$$

In order to calculate explicitly the expected value of Eq. 6.2-1 it is necessary to assume some statistical properties for the random quantities appearing in this equation.

Let

$$p_j = \Pr\{\mathbf{F}(t) = \mathbf{F}^{(j)}, t_{k-1} \leq t < t_k\} \quad j=1, \dots, N \quad (6.2-2)$$

be independent of the switching times $\{t_k\}$. Furthermore, assume that the switching times occur as an inhomogeneous Poisson counting process with probability law

$$\Pr\{K(t) = k\} = \frac{\left[\int_0^t \lambda(\tau) d\tau\right]^k}{k!} e^{-\int_0^t \lambda(\tau) d\tau} \quad k=0,1,2,\dots \quad (6.2-3)$$

These assumptions will allow explicit calculation of $E\{\mathbf{x}(t)\}$. From Eqs. 6.2-1 and 6.2-3,

$$\begin{aligned} E\{\mathbf{x}(t)\} &= E\{E\{\mathbf{x}(t)|K(t) = k\}\} \\ &= \sum_{k=0}^{\infty} E\{\mathbf{x}(t)|K(t) = k\} \Pr\{K(t) = k\} \\ &= \sum_{k=0}^{\infty} E\{e^{\mathbf{F}_k(t-t_k)} \dots e^{\mathbf{F}_0(t_1-t_0)} \mathbf{x}_0 | \\ &\quad K(t) = k, t_1 < t_2 < \dots < t_k < t\} \end{aligned}$$

$$\times \frac{\left[\int_0^t \lambda(\tau) d\tau\right]^k}{k!} e^{-\int_0^t \lambda(\tau) d\tau} \quad (6.2-4)$$

Consider now, the conditional expectation appearing within the summation of Eq. 6.2-4. Evaluation of this expectation requires knowledge of the statistics of the random occurrence times of the Poisson process. Let $\{\tau_k\}_1$ denote the sequence of random times. This sequence is a Markov process as shown by the following. For $0 < \tau_1 < \tau_2 < \dots < \tau_k$,

$$\begin{aligned} &\Pr\{\tau_{k+1} > t_{k+1} | \tau_1 = t_1, \tau_2 = t_2, \dots, \tau_k = t_k\} \\ &= \Pr\{K(t_{k+1}) - K(t_k) = 0 | \tau_1 = t_1, \dots, \tau_k = t_k\} \\ &= \Pr\{K(t_{k+1}) - K(t_k) = 0 | K(t), 0 \leq t \leq t_k\} \\ &= \Pr\{K(t_{k+1}) - K(t_k) = 0\} \end{aligned} \quad (6.2-5)$$

In Eq. 6.2-5 t_k is a random variable, but since the Poisson process is a strong Markov process, the

Markov property holds even for random times. The independence of the Poisson increment makes it possible to write (for a general inhomogeneous process with intensity $\lambda(t)$)

$$\begin{aligned} & \Pr\{\tau_{k+1} > t_{k+1} \mid \tau_1 = t_1, \dots, \tau_k = t_k\} \\ &= \Pr\{K(t_{k+1}) - K(t_k) = 0\} = e^{-\int_{t_k}^{t_{k+1}} \lambda(t) dt} \end{aligned} \quad (6.2-6)$$

Equivalently,

$$\begin{aligned} & \Pr\{\tau_{k+1} \leq t_{k+1} \mid \tau_1 = t_1, \dots, \tau_k = t_k\} \\ &= 1 - e^{-\int_{t_k}^{t_{k+1}} \lambda(t) dt} \\ & 0 = t_0 < t_1 < \dots < t_k < t_{k+1} \end{aligned} \quad (6.2-7)$$

which is independent of $t_1, \dots, t_k$, thus proving the Markov property for the sequence $\{\tau_k\}$. The transition probability density for the sequence is obtained by differentiating Eq. 6.2-7,

$$p(t \mid \tau_k = t_k) = \lambda(t) e^{-\int_{t_k}^t \lambda(\tau) d\tau} \quad (6.2-8)$$

Using Eq. 6.2-8, the joint density of the random time sequence is

$$\begin{aligned} p(t_1, \dots, t_k) &= p(t_k \mid t_{k-1}) p(t_{k-1} \mid t_{k-2}) \dots p(t_2 \mid t_1) p(t_1) \\ &= \lambda(t_1) \lambda(t_2) \dots \lambda(t_k) e^{-\left[\int_0^{t_1} + \int_{t_1}^{t_2} \dots + \int_{t_{k-1}}^{t_k} \lambda(\tau) d\tau\right]} \\ &= \lambda(t_1) \dots \lambda(t_k) e^{-\int_0^{t_k} \lambda(\tau) d\tau} \end{aligned} \quad (6.2-9)$$

It is now possible to solve for $p(t_1, \dots, t_k \mid K(t) = k)$, the conditional density of the occurrence times given that there are exactly k of them in the time interval $[0, t]$. From Bayes' formula,

$$\begin{aligned} & p(t_1, \dots, t_k \mid K(t) = k) \\ &= \frac{\Pr\{K(t)=k \mid \tau_1=t_1, \dots, \tau_k=t_k\} p(t_1, \dots, t_k)}{\Pr\{K(t) = k\}} \end{aligned} \quad (6.2-10)$$

Equation 6.2-10 may be written as

$$p(t_1, \dots, t_k \mid K(t) = k) = \frac{\Pr\{K(t) - K(t_k) = 0\} p(t_1, \dots, t_k)}{\Pr\{K(t) = k\}} \quad (6.2-11)$$

since the Poisson process is a strong Markov process with independent increments. Substituting Eqs. 6.2-3 and 6.2-9 into Eq. 6.2-11 gives

$$\begin{aligned} p(t_1, \dots, t_k \mid K(t) = k) &= \frac{\left[e^{-\int_{t_k}^t \lambda(\tau) d\tau} \right] \left[\prod_{i=1}^k \lambda(t_i) e^{-\int_0^{t_i} \lambda(\tau) d\tau} \right]}{\left[\frac{\left[\int_0^t \lambda(\tau) d\tau \right]^k}{k!} \right] e^{-\int_0^t \lambda(\tau) d\tau}} \\ &= k! \frac{\lambda(t_1) \dots \lambda(t_k)}{\left[\int_0^t \lambda(\tau) d\tau \right]^k} \end{aligned} \quad (6.2-12)$$

Equation 6.2-4 can now be evaluated, keeping in mind the fact that the matrices $\{F_i\}$ and the random times $\{t_i\}$ are independent random variables. First, consider

$$\begin{aligned} & E\{e^{F_k(t-t_k)} \dots e^{F_0(t_1)} \mid K(t) = k\} \\ &= E_{\{t_i\}} \{E_{\{F_i\}}\} \\ &= \{e^{F_k(t-t_k)} \dots e^{F_0(t_1)} \mid K(t) = k\} |K(t)=k\} \end{aligned} \quad (6.2-13)$$

where the expectation is evaluated first with respect to the $\{F_i\}$ and then with respect to the $\{t_i\}$. The inner term of Eq. 6.2-13 is

$$\begin{aligned} & E_{\{F_i\}} \{e^{F_k(t-t_k)} \dots e^{F_0(t_1)} \mid K(t) = k\} \\ &= E_{\{F_i\}} \{e^{F_k(t-t_k)}\} \prod_{i=1}^k E_{\{F_i\}} \{e^{F_{i-1}(t_i-t_{i-1})}\} \end{aligned} \quad (6.2-14)$$

since the transitions between the F_i are independent of the random occurrence times. Now, let

$$\Phi(\tau) = E_{\{F_i\}} \{e^{F\tau}\} = \sum_{j=1}^N e^{F^{(j)}\tau} p_j \quad (6.2-15)$$

and substitute Eqs. 6.2-12, 6.2-14 and 6.2-15 into Eq. 6.2-13 to obtain

$$\begin{aligned}
 E\{e^{F_k(t-t_k)} \dots e^{F_0(t_1)} | K(t) = k\} \\
 = E_{\{t_1\}} \{\phi(t-t_k)\phi(t_k-t_{k-1}) \dots \phi(t_2-t_1)\phi(t_1) | K(t) = k\} \\
 = \int_0^t d\tau_k \int_0^{\tau_k} d\tau_{k-1} \dots \int_0^{\tau_2} d\tau_1 \phi(t-t_k)\phi(\tau_k-\tau_{k-1}) \dots \\
 \phi(\tau_2-\tau_1)\phi(\tau_1) \left[k! \frac{\lambda(\tau_1) \dots \lambda(\tau_k)}{\left[\int_0^t \lambda(\tau) d\tau \right]^k} \right]
 \end{aligned} \quad (6.2-16)$$

For a general inhomogeneous Poisson process with arbitrary $\lambda(t)$, it is difficult to proceed beyond this point. However, in the homogeneous case with $\lambda(t) = \lambda$, a constant, the integral in Eq. 6.2-16 becomes

$$\begin{aligned}
 E\{e^{F_k(t-t_k)} \dots e^{F_0(t_1)} | K(t) = k\} \\
 = \frac{k! \lambda^k}{(\lambda t)^k} \int_0^t d\tau_k \dots \int_0^{\tau_2} d\tau_1 \\
 \times \phi(t-\tau_k)\phi(\tau_k-\tau_{k-1}) \dots \phi(\tau_2-\tau_1)\phi(\tau_1)
 \end{aligned} \quad (6.2-17)$$

Cancelling the λ terms in Eq. 6.2-17 and substituting the result into Eq. 6.2-4 yields

$$\begin{aligned}
 E\{x(t)\} &= \sum_{k=0}^{\infty} \left[\frac{k!}{t^k} \int_0^t d\tau_k \int_0^{\tau_k} d\tau_{k-1} \dots \int_0^{\tau_2} d\tau_1 \right. \\
 &\quad \times \phi(t-\tau_k) \dots \phi(\tau_2-\tau_1)\phi(\tau_1) \left. \frac{(\lambda t)^k}{k!} e^{-\lambda t} \right] E\{x_0\} \\
 &= e^{-\lambda t} \left\{ \sum_{k=0}^{\infty} \lambda^k \int_0^t d\tau_k \dots \int_0^{\tau_2} d\tau_1 \right. \\
 &\quad \times \left. [\phi(t-\tau_k) \dots \phi(\tau_2-\tau_1)\phi(\tau_1)] \right\} E\{x_0\}
 \end{aligned} \quad (6.2-18)$$

The integral in Eq. 6.2-18 is a k^{th} -order convolution which arises in the theory of integral equations. Consider the function $y(t)$ which satisfies a Volterra integral equation of the second kind,

$$y(t) = \lambda \int_a^t K(t,s)y(s)ds + g(t) \quad 0 \leq a < t < T \quad (6.2-19)$$

where λ is complex and the kernel is assumed to satisfy the following three conditions:

$$\text{i)} \quad \int_a^T \int_a^T |K(t,s)|^2 ds dt < \infty \quad (6.2-20)$$

$$\text{ii)} \quad \int_a^T |K(t,s)|^2 ds < \infty \quad (6.2-21)$$

$$\text{iii)} \quad \int_a^T |K(t,s)|^2 dt < \infty \quad (6.2-22)$$

(If i) through iii) are satisfied, $K(t,s)$ is called an L^2 kernel.) If λ is a regular value of the kernel $K(t,s)$, Eq. 6.2-19 has the unique solution (see Ref. 51)

$$y(t) = g(t) + \lambda \int_a^t H_\lambda(t,s)g(s)ds \quad (6.2-23)$$

where $H_\lambda(t,s)$ is the resolvent of $K(t,s)$ defined as the sum

$$H_\lambda(t,s) = K(t,s) + \lambda K^2(t,s) + \lambda^2 K^3(t,s) + \dots \quad (6.2-24)$$

In Eq. 6.2-24, the iterated kernels are given by

$$K^2(t,s) = \int_a^t K(t,r)K(r,s)dr \quad (6.2-25)$$

$$K^3(t,s) = \int_a^t K(t,r)K^2(r,s)dr \quad (6.2-26)$$

etc.

Consider the (vector) case with

$$K(t,s) = \begin{cases} \phi(t-s) & s \leq t \\ 0 & s > t \end{cases} \quad (6.2-27)$$

$$g(t) = \phi(t)y_0 \quad (6.2-28)$$

where y_0 is a constant. Substituting these expressions into Eq. 6.2-19 and setting $a=0$ gives

$$\underline{y}(t) = \lambda \int_0^t \phi(t-s) \underline{y}(s) ds + \phi(t) \underline{y}_0 \quad (6.2-29)$$

which has a solution

$$\begin{aligned} \underline{y}(t) &= \phi(t) \underline{y}_0 + \lambda \int_0^t \sum_{k=1}^{\infty} \lambda^{k-1} \phi^k(t-s) \phi(s) \underline{y}_0 ds \\ &= \phi(t) \underline{y}_0 + \sum_{k=1}^{\infty} \int_0^t \lambda^k \phi^k(t-\tau_1) \phi(\tau_1) d\tau_1 \underline{y}_0 \end{aligned} \quad (6.2-30)$$

where

$$\begin{aligned} \phi^k(t-\tau_1) &= \int_0^t d\tau_k \int_0^{\tau_k} d\tau_{k-1} \dots \int_0^{\tau_3} d\tau_2 \\ &\times \phi(t-\tau_k) \phi(\tau_k-\tau_{k-1}) \dots \phi(\tau_2-\tau_1) \quad k \geq 2 \end{aligned} \quad (6.2-31)$$

Inserting Eq. 6.2-31 into Eq. 6.2-30 and making use of the property that $\phi(\tau_2-\tau_1) = 0$ for $\tau_1 > \tau_2$ results in

$$\begin{aligned} \underline{y}(t) &= \phi(t) \underline{y}_0 \\ &+ \sum_{k=1}^{\infty} \lambda^k \int_0^t d\tau_k \int_0^{\tau_k} d\tau_{k-1} \dots \int_0^{\tau_2} d\tau_1 \\ &\times [\phi(t-\tau_k) \phi(\tau_k-\tau_{k-1}) \dots \phi(\tau_2-\tau_1) \phi(\tau_1)] \underline{y}_0 \end{aligned} \quad (6.2-32)$$

This is identical to the term in braces in Eq. 6.2-18 since the $k=0$ term is

$$\phi(t) = \phi(t-\tau_0) \quad (6.2-33)$$

where $\tau_0 = 0$. Thus, the first moment can be written as

$$E\{\underline{x}(t)\} = e^{-\lambda t} \underline{y}(t) \quad (6.2-34)$$

with $\underline{y}_0 = E\{\underline{x}_0\}$ and $\phi(\tau)$ given by Eq. 6.2-15.

One method for solving Eq. 6.2-29 is via Laplace transforms. Let $Y(s)$ and $\bar{\phi}(s)$ denote the Laplace transforms of $\underline{y}(t)$ and $\phi(t)$, respectively. Then

$$Y(s) = \lambda \bar{\phi}(s) Y(s) + \bar{\phi}(s) \underline{y}_0 \quad (6.2-35)$$

or

$$[I - \lambda \bar{\phi}(s)] Y(s) = \bar{\phi}(s) \underline{y}_0 \quad (6.2-36)$$

which has the formal solution

$$Y(s) = [I - \lambda \bar{\phi}(s)]^{-1} \bar{\phi}(s) \underline{y}_0 \quad (6.2-37)$$

This result admits the block diagram representation for $E\{\underline{x}(t)\}$ depicted in Fig. 6.2-1.

Figure 6.2-1 Block Diagram Solution for the First Moment

6.3 EVALUATION OF THE SECOND MOMENT

The second moment, $E\{\underline{x}(t)\underline{x}^T(t)\}$, is discussed briefly in this section. The approach described below for evaluating the second moment involves Kronecker products (see, for example Ref. 52). This is a well known device that has been applied by Bharucha (Ref. 53) in a study of the stability of stochastic differential equations and by Brockett (Ref. 54) and others to obtain moments of solutions to linear Itô equations.

The basic idea is illustrated in the following simple example. Consider the second-order system

$$\begin{aligned} \dot{x}_1(t) &= f_{11}(t) x_1(t) + f_{12}(t) x_2(t) \\ \dot{x}_2(t) &= f_{21}(t) x_1(t) + f_{22}(t) x_2(t) \end{aligned} \quad (6.3-1)$$

which may be written compactly as

$$\dot{\underline{x}}(t) = F_2(t) \underline{x}(t) \quad (6.3-2)$$

with

$$F_2(t) = \begin{bmatrix} f_{11}(t) & f_{12}(t) \\ f_{21}(t) & f_{22}(t) \end{bmatrix} \quad (6.3-3)$$

Now the second moment involves terms of the form x_1^2 , $x_1 x_2$ and x_2^2 . A computational simplification will result if it is possible to find a set of linear differential equations for the time evolution of $x_1^2(t)$, $x_1(t)x_2(t)$ and $x_2^2(t)$. Straightforward differentiation gives the following coupled equations

$$\frac{d}{dt} (x_1^2) = 2f_{11}(t) x_1^2 + 2f_{12}(t) x_1 x_2$$

$$\begin{aligned} \frac{d}{dt} (x_1 x_2) &= f_{21}(t) x_1^2 + (f_{11}(t) + f_{22}(t)) x_1 x_2 \\ &\quad + f_{12}(t) x_2^2 \end{aligned}$$

$$\frac{d}{dt} (x_2^2) = 2f_{21}(t) x_1 x_2 + 2f_{22}(t) x_2^2 \quad (6.3-4)$$

which may also be written as

$$\frac{d}{dt} \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix} = \begin{bmatrix} 2f_{11}(t) & 2f_{12}(t) & 0 \\ f_{21}(t) & f_{11}(t)+f_{22}(t) & f_{12}(t) \\ 0 & 2f_{21}(t) & 2f_{22}(t) \end{bmatrix} \begin{bmatrix} x_1^2 \\ x_1 x_2 \\ x_2^2 \end{bmatrix} \quad (6.3-5)$$

An alternate formulation is to define the vector

$$\underline{x}^{(2)} = (x_1^2, x_1 x_2, x_2 x_1, x_2^2)^T \quad (6.3-6)$$

which satisfies

$$\dot{\underline{x}}^{(2)} = \begin{bmatrix} 2f_{11}(t) & f_{12}(t) & f_{12}(t) & 0 \\ f_{21}(t) & f_{11}(t)+f_{22}(t) & 0 & f_{12}(t) \\ f_{21}(t) & 0 & f_{11}(t)+f_{22}(t) & f_{12}(t) \\ 0 & f_{21}(t) & f_{21}(t) & 2f_{22}(t) \end{bmatrix} \underline{x}^{(2)} \quad (6.3-7)$$

Equation 6.3-7 may be written compactly as

$$\dot{\underline{x}}^{(2)} = \{ [F_2(t) \otimes I] + [I \otimes F_2(t)] \} \underline{x}^{(2)} \quad (6.3-8)$$

where the Kronecker product $\otimes$ is defined (for the $m \times n$ matrix A and $l \times p$ matrix B) to be the $(ml) \times (np)$ matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \dots & a_{1n}B \\ a_{21}B & \dots & a_{2n}B \\ \vdots & & \vdots \\ a_{m1}B & \dots & a_{mn}B \end{bmatrix} \quad (6.3-9)$$

Note that the matrix in Eq. 6.3-5 is simply a condensation of the matrix in Eq. 6.3-7 obtained by replacing the second column by the sum of the second and third columns and then deleting the third row and column. In each case, it was possible to obtain linear equations (cf Eqs. 6.3-5 and 6.3-7) for the second-order products of the original variables x_1 and x_2 . Furthermore, the resulting dynamics matrix has elements that are linear combinations of the elements of the original matrix F_2 .

These ideas can be extended to the general case where the system of interest has the form

$$\frac{d\underline{x}^{(2)}(t)}{dt} = F^{(2)}(t) \underline{x}^{(2)}(t), \quad \underline{x}^{(2)}(t_0) = \underline{x}_0^{(2)} \quad (6.3-10)$$

where $\underline{x}^{(2)}(t)$ is the vector with elements $x_1^2(t)$, $x_1(t)x_2(t)$, ..., $x_2^2(t)$, $x_2(t)x_3(t)$, ..., $x_n^2(t)$. The matrix $F^{(2)}(t)$ is assumed to take values from the finite set of matrices $\{F^{(1)}(t), F^{(2)}(t), \dots, F^{(N)}(t)\}$, where each of the $F^{(1)}(t)$ is obtained from the original $F^{(1)}(t)$ via Kronecker products, as in the preceding example. Since Eq. 6.3-10 has exactly the same form as Eq. 6.1-1, the analysis of Section 6.2 can be applied to yield

$$\begin{aligned} E\{\underline{x}^{(2)}(t)\} &= e^{-\lambda t} \left\{ \sum_{k=0}^{\infty} \lambda^k \int_0^t d\tau_k \dots \int_0^{\tau_2} d\tau_1 \right. \\ &\quad \times [\Phi^{(2)}(t-\tau_k) \dots \\ &\quad \left. \Phi^{(2)}(\tau_2-\tau_1) \Phi^{(2)}(\tau_1)] \right\} E\{\underline{x}_0^{(2)}\} \end{aligned} \quad (6.3-11)$$

where

$$\Phi^{(2)}(\tau) \equiv E\{e^{F^{(2)}\tau}\} = \sum_{j=1}^N e^{F^{(j)}\tau} p_j \quad (6.3-12)$$

Thus, the solution procedure for the second moment is the same as for the first moment except that $\Phi(\tau)$ is replaced by the extended matrix $\Phi^{(2)}(\tau)$. In fact, the p^{th} -order moment can be obtained from the same expressions by computing the matrix $\Phi^{(p)}(\tau)$. Although Eq. 6.3-11 "solves" the second moment problem, a realization of this solution involves computation with $n^{(2)} \times n^{(2)}$ matrices where $n^{(2)} = \frac{n(n+1)}{2}$ is the dimension of the vector $\underline{x}^{(2)}(t)$. This computational problem becomes increasingly severe as the order of the desired moment increases.

6.4 RECURSIVE UPDATING OF THE STATE FORECAST

Expressions for the first and second moments of the forecasted system state were developed in Sections 6.2 and 6.3. From an implementation point of view, it is desirable to have recursive relationships for the forecasted variables so that, when new information about the current system state becomes available, the forecasts can be updated rather than completely regenerated. The results embodied in Eqs. 6.2-18 and 6.3-11 are of the desired form, but in an awkward notation. The goal of this section is to rewrite these equations explicitly in terms of the forecasting interval Δt_f .

The general form of the forecast equation is

$$E\{\underline{x}^{(1)}(t)\} = e^{-\lambda t} \left\{ \sum_{k=0}^{\infty} \lambda^k \int_{t_0}^t d\tau_k \cdots \int_{t_0}^{\tau_2} d\tau_1 \right. \\ \left. \times [\phi^{(1)}(t-\tau_k) \cdots \phi^{(1)}(\tau_1)] \right\} E\{\underline{x}^{(1)}(t_0)\} \quad (6.4-1)$$

where $\underline{x}^{(1)}$, $i=1,2,\dots$ contains the appropriate terms for the i^{th} moment. Under the assumption that λ is constant (which is necessary to derive Eq. 6.4-1), it is easy to see that Eq. 6.4-1 depends only on the difference $t-t_0$. (In Sections 6.2 and 6.3, t_0 was taken as zero for convenience.) Rewriting Eq. 6.4-1 in terms of the current time t and the future time for which the forecast is desired, $t+\Delta t_f$, yields

$$E\{\underline{x}^{(1)}(t+\Delta t_f)\} = e^{-\lambda \Delta t_f} \left\{ \sum_{k=0}^{\infty} \lambda^k \int_0^{\Delta t_f} d\sigma_k \cdots \int_0^{\sigma_2} d\sigma_1 \right. \\ \left. \times [\phi^{(1)}(\Delta t_f - \sigma_k) \cdots \phi^{(1)}(\sigma_1)] \right\} E\{\underline{x}^{(1)}(t)\} \quad (6.4-2)$$

where $\sigma_j = \tau_j - t$, $j=1,2,\dots,k$. The integrals in Eq. 6.4-2 are a function only of the forecast interval Δt_f and the system parameters λ , $F^{(j)}$ and p_j . Consequently, the forecasted moments can be written as

$$E\{\underline{x}^{(1)}(t+\Delta t_f)\} = \Delta\phi^{(1)}(\Delta t_f) E\{\underline{x}^{(1)}(t)\} \quad (6.4-3)$$

where

$$\Delta\phi^{(1)}(\Delta t_f) = e^{-\lambda \Delta t_f} \left\{ \sum_{k=0}^{\infty} \lambda^k \int_0^{\Delta t_f} d\sigma_k \cdots \int_0^{\sigma_2} d\sigma_1 \right. \\ \left. \times [\phi^{(1)}(\Delta t_f - \sigma_k) \cdots \phi^{(1)}(\sigma_1)] \right\} \quad (6.4-4)$$

Equation 6.4-4 can be precomputed off-line for any number of forecast intervals (Δt_f) and then used in Eq. 6.4-3 to extrapolate the known (or estimated) moments from time t to time $t+\Delta t_f$.

7. LONG-TERM SECURITY MEASURES

7.1 INTRODUCTION

Power system security is typically described by five categories of operating conditions (Ref. 55). Three of these are most relevant to the subject of security assessment:

- **Normal Operating State** - All constraints are satisfied and no constraints are in danger of being violated
- **Alert State** - All constraints are satisfied, but there is danger that some constraint will be violated unless action is taken to improve security -- i.e., to revert to the normal operating state
- **Emergency State** - System inequality constraints are violated, and there is danger of equipment damage or inability to supply the load demand.

The additional two conditions are in-extremis, in which a portion of the load is lost, and the restorative state which is a transition condition from in-extremis to alert or normal.

The primary emphasis of this research effort has been to define security measures which describe the likelihood of passing from the normal operating state to the alert or emergency states. In this sense, a security measure is a quantity which expresses the probability of a security violation. For instance, consider the inequality constraint of the form

$$x(t) \leq x_{\max} ; t_0 \leq t \leq t_f \quad (7.1-1)$$

where t_0 is the present time and t_f is a future time, perhaps based on the period required to take corrective action if it is predicted that the constraint will be violated. Let us assume that any single violation of the constraint constitutes a security violation. Therefore, the security measure, S , which describes the probability that a security violation will occur, is the probability that x will not exceed $x_{\max}$ anytime within the interval, i.e.,

$$S = P\{x(t) \leq x_{\max} ; \text{for all } t \leq [t_0, t_f]\} \quad (7.1-2)$$

The calculation of S falls in the class of level crossing problems (Ref. 56) which define whether $x(t_0)$ is in the alert state. A surface crossing formulation of the long-term security assessment problem is examined in Section 7.2. Two different approaches are described; the first makes use of scalar level crossing theory, while the second is based on a partial differential equation of Dynkin.

Equation 7.1-1 is a precise expression for a (hypothetical) condition under which security is violated. However, in some cases, practical considerations may dictate an alternative definition. To illustrate, consider Fig. 7.1-1 which portrays the hypothetical evolution of a power system state variable. Times t_1 and t_2 are occurrence times of jump events that change the system topology. Thus, $x(t)$

Figure 7.1-1 Hypothetical Evolution of a State Variable

can be viewed as being composed of transient effects produced by the events, and long-term dynamic behavior caused by the load. The transients are governed by high-order differential equations and have relatively fast "time-constants." Although level crossings may occur during the transients (e.g., at time t_2), prediction of such events in real time, using level-crossing theory is extremely difficult. An alternate approach to the assessment of transient stability based upon the use of Lyapunov stability theory, is described in Chapter 8.

The final section in this chapter documents research on the problem of long-term stability from a viewpoint which differs from the level crossing based approach. The goal of this effort was to determine the long term stability of the system, accounting for potential future contingencies. The basic formulation consists of analyzing the stability properties of the linearized system

$$\dot{\underline{x}} = F(t)\underline{x} \quad (7.1-3)$$

where $F(t)$ includes system coefficients that are subject to future jumps. The main result is a sufficient condition for almost sure (a.s.) asymptotic stability.

7.2 SURFACE CROSSING PROBLEMS

One approach to the security assessment problem is to view it as a "surface crossing problem" for a stochastic process. The idea here is to consider a set of stochastic differential equations for the electrical power system in the form derived in Appendices A and B. Assume that these equations are sufficiently detailed to be able to characterize those features of the system behavior that are of interest. Then a transition of the current system state to an "emergency" state in a period τ is described by a trajectory of the system running from the current state into the emergency region in time τ .

For a perfectly observed, deterministic system whether or not this event occurs can be answered with certainty. However, as discussed in Appendix A, the overall power system is a partially observed stochastic system. This leads us to examine the security of an electrical power system via the probability of transition from the current state -- or estimate thereof -- into the emergency region during a given period τ . In its full generality this constitutes a "surface crossing" problem for a partially observed stochastic process generated by nonlinear stochastic differential equations.

In the sequel two special cases of the surface crossing problem of security assessment are considered. The first is the level crossing problem for Gaussian non-Markov processes. The second is the surface crossing problem for Markov processes generated by linear Itô equations, the approach in this case is an extension of the classical analysis using Dynkin's equation

7.2.1 Formulation of the Level Crossing Problem

By way of introduction, consider the surface crossing problem for scalar Markov processes (i.e., the level crossing problem for such processes)

Given a scalar stochastic process $x(t, \omega) \in \mathbb{R}^1$ [defined on some probability space $(\Omega, \mathcal{F}, P)$] and given a number $\alpha \in \mathbb{R}^1$ and an initial condition $x_0 < \alpha$, determine the probability of exceeding the level α in an interval $[0, T]$.

Defining

$$P_\alpha(T|x_0) \triangleq P\{x(t) < \alpha, \forall t \in [0, T], \text{ given } x(0) = x_0\} \quad (7.2-1)$$

results in $1 - P_\alpha(T|x_0)$ being the probability that at least one crossing occurs in $[0, T]$. Furthermore,

$$\begin{aligned} P_\alpha(T|x_0) &= P\{\text{no crossing in } [0, T] \text{ and} \\ &\quad \text{a crossing in } [T, T+dT]\} \\ &\quad + P\{\text{no crossing in } [0, T] \\ &\quad \text{and no crossing in } [T, T+dT]\} \\ &= P\{x(t) \text{ first crosses } \alpha \text{ in } [T, T+dT]\} \\ &\quad + P_\alpha(T+dT|x_0) \end{aligned} \quad (7.2-2)$$

from which

$$-\frac{\partial P_\alpha(T|x_0)}{\partial T} = \lim_{dT \rightarrow 0} \frac{1}{dT} [P_\alpha(T|x_0) - P_\alpha(T+dT|x_0)] = q_\alpha(T) \quad (7.2-3)$$

where $q_\alpha(T)dT$ is defined to be the probability that the first crossing occurs in $[T, T+dT]$.

It may be verified that if $x > \alpha > x_0$, and if x is a Markov process with almost surely continuous sample paths, then

$$\begin{aligned} P(x(t_0 + \tau) = x | x(t_0) = x_0) \\ = \int_{t_0}^{t_0 + \tau} q_\alpha(\theta) [P(x(t_0 + \tau) = x | x(\theta) = \alpha)] d\theta \end{aligned} \quad (7.2-4)$$

which is depicted in Fig. 7.2-1. Solving Eq. 7.2-4 for $q_\alpha(T)$ yields $P_\alpha(T|x_0)$ via Eq. 7.2-3

Figure 7.2-1 First Crossing Times

A substantial part of the level crossing literature is concerned with attempts to solve this equation for different types of Markov processes. In general this turns out to be a very difficult task. A more tractable measure of the chance of crossing some level α is given by the moments of $C_\alpha[0, T]$, where $C_\alpha[0, T]$ denotes the integer random variable given by the number of upcrossings, downcrossings and tangencies in the period $[0, T]$. (For many processes the measure of the latter set of events is zero.) In particular consider a stationary, scalar Gaussian process (which is not necessarily Markovian), which has zero mean and for which the $(2i)^{\text{th}}$ spectral moments of the spectral distribution $F(\lambda)$ are denoted by $\rho_0, \rho_2, \dots$. Then, the expected number of crossings is given by the Rice-Kač formula

$$E\{C_\alpha[0, T]\} = \frac{T}{\pi} \left(\frac{\rho_2}{\rho_0}\right)^{1/2} e^{-\alpha^2/2\rho_0} \quad (7.2-5)$$

a clear derivation of which is presented in (Ref. 57). It should be noted that for a stationary process, $E\{C_\alpha[0, T]\} = TE\{C_\alpha[0, 1]\}$ and it is thus sufficient to consider the interval $[0, 1]$.

A generalized form of the Rice-Kač formula due to Cramér and Leadbetter is applied in the sequel to a class of non-stationary vector process level crossing problems.

Equation 7.2-5 imposes an important restriction on the random process; namely, it must have a finite second spectral moment,

$$\rho_2 = \int_{-\infty}^{\infty} \lambda^2 \Phi(\lambda) d\lambda < \infty \quad (7.2-6)$$

where $\Phi(\lambda)$ is the spectral density function of the process. In general this constitutes a restriction on Markov processes defined by equations having the form of Eq. 2.1-4

Appendix G develops a linear model of the form

$$d\mathbf{x}(t) = \mathbf{F}(t) \mathbf{x}(t)dt + \mathbf{G}(t) \mathbf{u}(t)dt + \mathbf{M}(t) d\mathbf{m}(t) \quad (7.2-7)$$

$$d\mathbf{y}(t) = \mathbf{H}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)dt + \mathbf{N}(t) d\mathbf{n}(t) \quad (7.2-8)$$

where $\mathbf{x}$ and $\mathbf{y}$ are the state and measurement vectors, respectively, and $\mathbf{m}$ and $\mathbf{n}$ are Gaussian stationary increment processes with almost surely (a.s.) differentiable sample paths. This model has the property that $\mathbf{x}$ and $\mathbf{y}$ have finite second spectral moments and are differentiable. This latter restriction is necessary for much of the subsequent discussion.

7.2.2 Formulation of the Surface Crossing Problem

The general surface crossing problem can be described as follows: Let S be a compact, connected surface in R^n , i.e., a compact, connected $(n-1)$ -dimensional manifold in R^n . Assume S has no boundary and observe that one of the two connected pieces into which S divides R^n has finite volume (Ref. 58). Call this open, connected n -dimensional manifold $\text{int}(S)$, the interior of S . Then by analogy with Eq. 7.2-1, define

$$P_S(T|x_0) \triangleq P[\mathbf{x}(t) \in \text{int}(S), t \in [0, T], \text{ given } \mathbf{x}(0) = \mathbf{x}_0 \in \text{int}(S)] \quad (7.2-9)$$

and write $q_S(T)$ for the first crossing probability density whenever this exists (see Fig. 7.2-2). As in the level crossing problem, the goal in the surface crossing problem is to compute the quantities $P_S(T|x_0)$ and $q_S(T)$.

Figure 7.2-2 Geometry of the Surface Crossing Problem

The first approach to the surface crossing problem may be motivated in either of the following ways: First, imagine that one of the constraint surfaces is nearly linear and assume that it is significantly closer to the initial position $\mathbf{x}_0$ of the process $\mathbf{x}$ than the other constraints. Approximate this

constraint surface by a hyperplane (denoted by H) and treat the general surface crossing problem as a simpler hyperplane crossing problem. Or second, suppose that the emergency states form a convex set V at some distance from the origin. By the separating hyperplane theorem (Ref. 59) there exists a hyperplane which isolates the region V from the origin.

Either of these formulations may be used to justify the following approximations to the surface crossing problem. Describe either of the hyperplanes mentioned above by

$$H = \{\mathbf{x}: \underline{\lambda}^T(\mathbf{x} - \mathbf{x}_0) = 0, \mathbf{x} \in R^n, \text{ for some } \mathbf{x}_0 \in R^n\} \quad (7.2-10)$$

Of course, for such $\mathbf{x}_0$ and $\underline{\lambda} \in R^n$, the hyperplane H is also described by

$$||\mathbf{x}|| \cos \theta(\mathbf{x}) = \text{constant} \quad (7.2-11)$$

with $\theta(\mathbf{x})$ as shown in Fig. 7.2-3. Now, consider the variation of the process $\mathbf{x}(t) \in R^n$ along the direction given by the fixed vector $\underline{\lambda} \in R^n$. The frequency with which $\mathbf{x}(t)$ crosses H may be treated as an approximation to the frequency with which it passes through the surface S . The important consequence of this approximation is that the hyperplane crossing problem for the process $\mathbf{x}$ is equivalent to the level crossing problem for the scalar process $\underline{\lambda}^T \mathbf{x}$.

Figure 7.2-3 Hyperplane Geometry

Let z be any zero-mean, scalar Gaussian process with covariance $r(t, s)$ and which possesses a quadratic mean derivative (see Appendix G). The following paragraphs contain the development of a general expression for $E\{C[0, T]\}$ for the process z . Assume $r(t, s)$ has second mixed partial derivatives on $t=s$. Then, with z' denoting the quadratic mean derivative of z , define

$$\gamma^2(t) \triangleq \text{Var}(z'_t) = r_{11}(t, t) \triangleq \frac{\partial^2 r(t, s)}{\partial t \partial s} \bigg|_{t=s} \quad (7.2-12)$$

and

$$\mu(t) \triangleq \frac{\text{cov}(z_t, z'_t)}{\gamma(t)\sigma(t)} = \frac{r_{01}(t,t)}{\gamma(t)\sigma(t)} = \left(\frac{1}{\gamma(t)\sigma(t)} \right) \frac{\partial r(t,s)}{\partial s} \Big|_{t=s} \quad (7.2-13)$$

where $\sigma^2(t) = r(t,t)$.

The crossings of the level α by the process z are equivalent to the crossings of the level 0 by the process $z - \alpha$. Consequently it is possible to apply the result of Cramér and Leadbetter (Ref. 57, pp. 286-288), which is valid for a process satisfying the stated hypotheses upon z , to get $E\{C[0,T]\}$. First, set

$$\eta_t = \frac{-\mu(t)\alpha}{\sigma(t)(1-\mu^2(t))^{1/2}} \quad (7.2-14)$$

then the formula of Cramér and Leadbetter reads

$$E\{C[0,T]\} = 2 \int_0^T \gamma(t)\sigma^{-1}(t) (1-\mu^2(t))^{1/2} \{2\phi(\eta(t)) + \eta(t)(2\phi(\eta(t))-1)\} dt \quad (7.2-15)$$

where $\phi(\cdot)$ denotes the Gaussian density $N(0,1)$ and

$$\phi(s) \triangleq \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^s e^{-u^2/2} du = \int_{-\infty}^s \phi(u) du$$

It should be stressed here that Eq. 7.2-15 holds for any mean square differentiable, scalar, Gaussian, non-stationary process and that such a process is not in general Markovian.

In order to apply Eq. 7.2-15 to the hyperplane crossing problem for the process $\underline{x}$ generated by Eq. 7.2-7 for any given $\underline{\lambda} \in R^n$ and any $\alpha \in R^1$, set $z(t) = \underline{\lambda}^T \underline{x}(t)$. Then the appropriate formulae for $\gamma^2(t)$, $\mu(t)$ and $\eta(t)$ are

$$\gamma^2(t) = \frac{\partial^2}{\partial t \partial s} \underline{\lambda}^T R(t,s) \underline{\lambda} \Big|_{t=s} \quad (7.2-16)$$

$$\mu(t) = \frac{1}{\gamma(t)\sigma(t)} \frac{\partial}{\partial s} \underline{\lambda}^T R(t,s) \underline{\lambda} \Big|_{t=s} \quad (7.2-17)$$

where $\sigma^2(t) = \underline{\lambda}^T R(t,t) \underline{\lambda}$,

$$\eta_t = \frac{-\mu(t)\alpha}{\sigma(t)(1-\mu^2(t))^{1/2}} \quad (7.2-18)$$

and where $R(t,s)$ and the partial differentials of $R(t,s)$ appearing in Eqs. 7.2-16 and 7.2-17, are given by Eqs. G.3-8, G.3-12 and G.3-13 in Appendix G. Thus, it is now possible to compute $E\{C[0,T]\}$ for a process generated by Eqs. 7.2-8 and 7.2-9. To summarize, the procedure is as follows:

- Use Eqs. G.3-8, G.3-12 and G.3-13 to compute $R(t,s)$, $\frac{\partial R(t,s)}{\partial t}$ and $\frac{\partial^2 R(t,s)}{\partial t \partial s}$, and hence to compute $\gamma^2(t)$ and $\mu(t)$ via Eqs. 7.2-16 and 7.2-17
- Evaluate Eq. 7.2-18 for the specified $\underline{\lambda}$ and α
- Substitute the result into Eq. 7.2-15 and integrate over the interval $[0,T]$.

If the hyperplane crossing problem is posed for the system output process $\underline{y}$ given in Eq. 7.2-8, the analysis proceeds in a similar manner. For simplicity, assume that the processes $\underline{m}$ and $\underline{n}$ are independent. Furthermore, let ψ_m and ψ_n denote the spectral densities associated with these stationary, independent-increment processes. Then the covariance of the $\underline{y}$ process is given by

$$R_y(t,s) = H \phi(t,0) V_0 \phi^T(s,0) H^T + H \left[\int_0^t \int_0^s \phi(t,u) K(u,v) \phi^T(s,v) du dv \right] H^T + \int_0^t \int_0^s L(u,v) du dv \quad (7.2-19)$$

where

$$K(u,v) = \int_{-\infty}^{\infty} e^{2\pi i(u-v)\lambda} \psi_m(\lambda) d\lambda \quad (7.2-20)$$

and

$$L(u,v) = \int_{-\infty}^{\infty} e^{2\pi i(u-v)\lambda} \psi_n(\lambda) d\lambda \quad (7.2-21)$$

The approach described above can now be applied using $R_y(t,s)$. It should be noted that the higher order moments of $C[0,T]$ may be computed by an extension of this technique but as one might expect the formulae become proportionately more complicated (Ref. 57).

7.2.3 Higher Moments and the Probability of Level Crossing

In this section, one method is illustrated for approximating

$$P[0,T] = \Pr\{\underline{x}(t) \text{ crosses } H \text{ in the interval } [0,T]\} \quad (7.2-22)$$

in terms of the moments of the integer random variable $C[0,T]$. Define

$$\Sigma^{(k)} = E\{C^k[0,T]\} \quad k \geq 1 \quad (7.2-23)$$

and let $P^{(k)}[0,T]$ denote the probability that $x(t)$ crosses H exactly k times during the interval $[0,T]$. Now, the $\Sigma^{(k)}$ may be written in terms of the $P^{(k)}[0,T]$ as

$$\begin{aligned} \Sigma^{(k)} &= \sum_{n=1}^{\infty} n^k P^{(n)}[0,T] \\ &= \sum_{n=1}^{\infty} n^k P^{(n)}[0,T] + P[0,T] - \sum_{n=1}^{\infty} P^{(n)}[0,T] \\ &= P[0,T] + \sum_{n=2}^{\infty} (n^k - 1) P^{(n)}[0,T] \quad k \geq 1 \end{aligned} \quad (7.2-24)$$

Under the assumption that the summations defining the $\Sigma^{(k)}$ can be truncated at $n=J$ with acceptable error, Eq. 7.2-24 can be expressed in the form

$$\begin{bmatrix} \Sigma^{(1)} \\ \Sigma^{(2)} \\ \vdots \\ \Sigma^{(J)} \end{bmatrix} = \begin{bmatrix} 1 & (2-1) & \dots & (J-1) \\ 1 & (2^2-1) & \dots & (J^2-1) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & (2^J-1) & \dots & (J^J-1) \end{bmatrix} \begin{bmatrix} P[0,T] \\ P^{(2)}[0,T] \\ \vdots \\ P^{(J)}[0,T] \end{bmatrix} \quad (7.2-25)$$

This equation may be solved for the unknowns which include the desired probability $P[0,T]$.

In order to solve Eq. 7.2-25, it is necessary to know the moments $\Sigma^{(k)}$, $1 \leq k \leq J$. Rather than compute the higher order moments themselves, it is more convenient to evaluate the factorial moments defined as

$$M^{(k)}[0,T] = E\{[C-1] \dots [C-k+1]\} \quad k \geq 1 \quad (7.2-26)$$

where the time interval $[0,T]$ has been suppressed on the right hand side of Eq. 7.2-26 for simplicity. It is clear that, for each N , the moments $E\{C\}$, ..., $E\{C^N\}$ can be obtained by a linear transformation of the $M^{(1)}[0,T]$, ..., $M^{(N)}[0,T]$. Cramér and Leadbetter (Ref. 57) give the following formula for the higher factorial moments:

$$\begin{aligned} M^{(k)}[0,T] &= 2 \int_0^T \dots \int_0^T dt_1 \dots dt_k \int_0^\infty \dots \\ &\int_0^\infty z_1 \dots z_k p_t(u; z') dz_1 \dots dz_k \end{aligned} \quad (7.2-27)$$

where

$$\begin{aligned} p_t(u; z') &\triangleq p_t(u_1, \dots, u_k; z'_1, \dots, z'_k), \\ t &= (t_1, \dots, t_k), \end{aligned}$$

and where this expression denotes the joint density of the random variables

$$(z(t_1), \dots, z(t_k); \frac{dz}{dt} \Big|_{t=t_1}, \dots, \frac{dz}{dt} \Big|_{t=t_k})$$

Since these random variables are jointly Gaussian, the integrand in Eq. 7.2-27 is completely specified once the covariance matrix is known. To compute this it is sufficient to find

$$\Sigma_{t_1, t_j}^{0,0} \triangleq E\{z_{t_1} z_{t_j}\} \quad 1 \leq j \leq k \quad (7.2-28)$$

$$\left(\Sigma_{t_j, t_1}^{1,0} \right)^T = \Sigma_{t_1, t_j}^{0,1} \triangleq E\left\{ z_{t_1} \frac{dz}{dt} \Big|_{t=t_j} \right\} \quad 1 \leq j \leq k \quad (7.2-29)$$

$$\Sigma_{t_1, t_j}^{1,1} = E\left\{ \frac{dz}{dt} \Big|_{t=t_1} \frac{dz}{dt} \Big|_{t=t_j} \right\} \quad 1 \leq j \leq k \quad (7.2-30)$$

The covariance of $(z_{t_1}, \dots, z_{t_k}, \frac{dz}{dt} \Big|_{t=t_1}, \dots, \frac{dz}{dt} \Big|_{t=t_k})$ is then given by

$$\Sigma_t = \begin{bmatrix} \begin{matrix} 0,0 & 0,0 \\ \Sigma_{t_1, t_1}^{0,0} \dots \Sigma_{t_1, t_k}^{0,0} \\ \vdots & \vdots \end{matrix} & \begin{matrix} 0,1 & 0,1 \\ \Sigma_{t_1, t_1}^{0,1} \dots \Sigma_{t_1, t_k}^{0,1} \\ \vdots & \vdots \end{matrix} \\ \begin{matrix} 0,0 & 0,0 \\ \Sigma_{t_k, t_1}^{0,0} \dots \Sigma_{t_k, t_k}^{0,0} \\ \vdots & \vdots \end{matrix} & \begin{matrix} 0,1 & 0,1 \\ \Sigma_{t_k, t_1}^{0,1} \dots \Sigma_{t_k, t_k}^{0,1} \\ \vdots & \vdots \end{matrix} \\ \hline \begin{matrix} 1,0 & 1,0 \\ \Sigma_{t_1, t_1}^{1,0} \dots \Sigma_{t_1, t_k}^{1,0} \\ \vdots & \vdots \end{matrix} & \begin{matrix} 1,1 & 1,1 \\ \Sigma_{t_1, t_1}^{1,1} \dots \Sigma_{t_1, t_k}^{1,1} \\ \vdots & \vdots \end{matrix} \\ \begin{matrix} 1,0 & 1,0 \\ \Sigma_{t_k, t_1}^{1,0} \dots \Sigma_{t_k, t_k}^{1,0} \\ \vdots & \vdots \end{matrix} & \begin{matrix} 1,1 & 1,1 \\ \Sigma_{t_k, t_1}^{1,1} \dots \Sigma_{t_k, t_k}^{1,1} \\ \vdots & \vdots \end{matrix} \end{bmatrix} \quad (7.2-31)$$

Expressions for $\Sigma_{t_1, t_j}^{0,0}$, $\Sigma_{t_1, t_j}^{0,1}$, $\Sigma_{t_1, t_j}^{1,1}$ may be obtained directly from the formulae derived above. To be specific,

$$\Sigma_{t_1, t_j}^{0,0} = \lambda^T R(t_1, t_j) \lambda \quad (7.2-32)$$

where, from Eq G.3-8,

$$R(t_1, t_j) = \Phi(t_1, 0) V_0 \Phi^T(t_j, 0) + \int_0^{t_1} \int_0^{t_j} \Phi(t_1, u) \int_{-\infty}^{\infty} e^{2\pi i(u-v)\lambda} \psi_m(\lambda) d\lambda \times \Phi^T(t_j, v) du dv \quad (7.2-33)$$

Let $z'(t_j)$ denote $\left. \frac{dz(t)}{dt} \right|_{t=t_j}$, then

$$\begin{aligned} \Sigma_{t_1, t_j}^{0,1} &= \lambda^T E\{z(t_1) z'(t_j)\} \lambda \\ &= \lambda^T \frac{\partial}{\partial t} \left[E\{z(t_1) z(t)\} \right]_{t=t_j} \lambda \\ &= \lambda^T \frac{\partial}{\partial t} R(t_1, t) \Big|_{t=t_j} \lambda \end{aligned} \quad (7.2-34)$$

The interchange of expectation and differentiation is justified by Fubini's theorem since

$$\int_0^t E\{ |(\lambda^T z(t_1))(\lambda^T z'(s))| \} ds < \infty, \quad 0 < t < \infty \quad (7.2-35)$$

implies

$$\begin{aligned} \int_0^t E\{(\lambda^T z(t_1))(\lambda^T z'(s))\} ds \\ = E\left\{ \int_0^t (\lambda^T z(t_1))(\lambda^T z'(s)) ds \right\} \end{aligned} \quad (7.2-36)$$

and differentiating Eq. 7.2-36 gives

$$E\{(\lambda^T z(t_1))(\lambda^T z'(t))\} = \frac{\partial}{\partial t} E\{(\lambda^T z(t_1))(\lambda^T z(t))\}, \quad 0 < t < \infty \quad (7.2-37)$$

Evaluating Eq. 7.2-34 by use of Eq. G.3-13 results in

$$\begin{aligned} \Sigma_{t_1, t_j}^{0,1} &= \lambda^T R(t_1, t_j) F^T(t_j) \lambda \\ &+ \lambda^T \left[\int_0^{t_j} \Phi(t_j, u) R_m(t_1 - u) du \right] \lambda \end{aligned} \quad (7.2-38)$$

Similar reasoning yields

$$\begin{aligned} \Sigma_{t_1, t_j}^{1,0} &= \lambda^T F(t_1) R(t_1, t_j) \lambda \\ &+ \lambda^T \left[\int_0^{t_1} R_m(t_1 - v) \Phi^T(t_1 - v) dv \right] \lambda \end{aligned} \quad (7.2-39)$$

and finally using Eq G.3-14,

$$\begin{aligned} \Sigma_{t_1, t_j}^{1,1} &= \lambda^T F(t_1) R(t_1, t_j) F^T(t_j) \lambda \\ &+ \lambda^T \left[F(t_1) \int_0^{t_1} \Phi(t_1, u) R_m^T(t_j - u) du \right. \\ &+ \left. \int_0^{t_1} R_m(t_1 - v) \Phi^T(t_j, v) dv F^T(t_j) \right] \lambda \\ &+ \lambda^T R_m(t_1 - t_j) \lambda \end{aligned} \quad (7.2-40)$$

Using these formulae it may be possible, in certain special cases, to evaluate the factorial moments $M^{(1)}[0, T], \dots, M^{(k)}[0, T]$ for reasonably high values of k . It would then be straightforward to calculate the moments $\Sigma^{(k)}$ and, via Eq. 7.2-25, obtain an approximate value for the level crossing probability $P[0, T]$.

7.2.4 Dynkin's Equation Approach to the Surface Crossing Problem

In the previous section, an approximate solution to the surface crossing problem was obtained using scalar level crossing theory. An alternate approach to the general surface crossing problem is discussed in this section. This approach is based on a partial differential equation of Dynkin (Ref 13). The solution to this equation yields the expected value of the first exit time of the process $\underline{x}(t)$ and closely related "Dynkin equations" give higher moments of this random variable. This is shown below.

Let the Markov process $\underline{x}$ satisfy

$$d\underline{x} = \underline{f}(\underline{x})dt + G(\underline{x})d\underline{w} \quad \underline{x} \in R^n \quad (7.2-41)$$

where $\underline{w}$ is a Brownian motion with incremental covariance

$$E\{d\underline{w} d\underline{w}^T\} = Q dt \quad \text{with } Q > 0 \quad (7.2-42)$$

Then, the infinitesimal generator of the process $\underline{x}$ is given by

$$A_{\underline{x}} = \sum_1 f_1(\underline{x}) \frac{\partial}{\partial x_1} + \frac{1}{2} \sum_{1j} (GQG^T)_{1j} \frac{\partial^2}{\partial x_1^2} \quad (7.2-43)$$

i.e., for any twice continuously differentiable function $\alpha(\underline{x})$

$$\lim_{t \rightarrow 0} \frac{E\{\alpha(\underline{x}(t+t_0)) - \alpha(\underline{x}(t_0))\}}{t} = A_{\underline{x}} \alpha(\underline{x}) \Big|_{\underline{x}=\underline{x}(t_0)} \quad (7.2-44)$$

Let D be a compact set in R^n , containing the origin O , whose boundary $S = \partial D$ is a connected $n-1$ dimensional C^∞ manifold. Let the twice continuously differentiable function v satisfy Dynkin's equation on D , i.e.,

$$\begin{aligned} -1 &= A_{\underline{x}} v(\underline{x}) \\ &= \underline{f}^T(\underline{x}) \frac{\partial v(\underline{x})}{\partial \underline{x}} + \frac{1}{2} \text{Tr} \left[G(\underline{x}) Q G^T(\underline{x}) \frac{\partial^2 v(\underline{x})}{\partial \underline{x}^2} \right] \quad \underline{x} \in D \\ 0 &= v(\underline{x}) \quad \underline{x} \in S = \partial D \end{aligned} \quad (7.2-45)$$

The basic theorem for the random first exit time from D for a process governed by Eq. 7.2-41 is that

$$E\{\tau - t_0\} = v(\underline{x}_0) \quad (7.2-46)$$

when v satisfies Eq. 7.2-45 and where τ is the random first exit time from D for the process starting at $\underline{x}_0 \in D$ at $t=t_0$.

Briefly, the proof of Eq. 7.2-46 proceeds as follows:

Let $t = T \wedge \tau = \min(T, \tau)$, to ensure the integrals below are finite, and let $\underline{x}_0$ be the initial value of $\underline{x}(t)$ at $t=t_0$. Then, applying the Itô differentiation rule (see e.g., Ref. 16) to $v(\underline{x})$ and using Eqs. 7.2-41 and 7.2-42 gives

$$\begin{aligned} v(\underline{x}(t)) - v(\underline{x}_0) &= \int_{t_0}^t \left[\frac{\partial v(\underline{x}(s))}{\partial \underline{x}} \right]^T G(\underline{x}(s)) d\underline{w}(s) \\ &+ \int_{t_0}^t \left\{ \underline{f}^T(\underline{x}(s)) \frac{\partial v(\underline{x}(s))}{\partial \underline{x}} \right. \\ &+ \left. \frac{1}{2} \text{Tr} \left[\frac{\partial^2 v(\underline{x}(s))}{\partial \underline{x}^2} G(\underline{x}(s)) Q G^T(\underline{x}(s)) \right] \right\} ds \\ &= \int_{t_0}^t \left[\frac{\partial v(\underline{x}(s))}{\partial \underline{x}} \right]^T G(\underline{x}(s)) d\underline{w}(s) - (t - t_0) \end{aligned} \quad (7.2-47)$$

where the last step follows from Dynkin's equation. Hence,

$$E_{t_0} \{v(\underline{x}(t)) - v(\underline{x}_0)\} = -E_{t_0} \{t - t_0\} \quad (7.2-48)$$

where E_{t_0} denotes expectation conditional on $\underline{x}(t_0)$

Let $T \rightarrow \infty$, then $t \rightarrow \tau$ and since $\underline{x}(\tau) \in \partial D$, $v(\underline{x}(\tau)) = 0$. Further, since v is bounded on D , it follows by dominated convergence that $E_{t_0} \{v(\underline{x}(t))\} \rightarrow 0$ and thus Eq. 7.2-48 yields

$$E_{t_0} \{\tau - t_0\} = v(\underline{x}_0) \quad (7.2-49)$$

Consider now the function $v_1(\underline{x})$ which satisfies

$$\underline{f}^T(\underline{x}) \frac{\partial v_1}{\partial \underline{x}} + \frac{1}{2} \text{Tr} \left[\frac{\partial^2 v_1}{\partial \underline{x}^2} G(\underline{x}) Q G^T(\underline{x}) \right] = -2v(\underline{x}) \quad \underline{x} \in D \quad (7.2-50)$$

with $v_1(\underline{x}) = 0$, $\underline{x} \in \partial D$ For $t = T \wedge \tau$,

$$\begin{aligned} v_1(\underline{x}(t)) - v_1(\underline{x}_0) &= -2 \int_{t_0}^t v(\underline{x}(s)) ds \\ &+ \int_{t_0}^t \left[\frac{\partial v_1(\underline{x}(s))}{\partial \underline{x}} \right]^T G(\underline{x}(s)) d\underline{w}(s) \end{aligned} \quad (7.2-51)$$

Taking the expected value of Eq. 7.2-51 gives

$$\begin{aligned} E_{t_0} \{v_1(\underline{x}(t)) - v_1(\underline{x}_0)\} &= -2E_{t_0} \left\{ \int_{t_0}^{t=T \wedge \tau} v(\underline{x}(s)) ds \right\} + 0 \\ &= -2E_{t_0} \left\{ \int_{t_0}^{t=T \wedge \tau} [E_s \{T \wedge \tau - s\}] ds \right\} \\ &= -E_{t_0} \{[t - t_0]^2\} \end{aligned} \quad (7.2-52)$$

Let $T \rightarrow \infty$ to obtain

$$v_1(\underline{x}_0) = E_{t_0} \{(\tau - t_0)^2\} \quad (7.2-53)$$

as required. Higher moments may be obtained recursively, for example the third moment is given by

$$A_{\underline{x}} v_2(\underline{x}) = -3 v_1(\underline{x}) \quad \underline{x} \in D \quad (7.2-54)$$

$$v_2(\underline{x}) = 0 \quad \underline{x} \in \partial D, \text{ etc.}$$

For the remainder of this development, it will be assumed that $G(\underline{x}) = I$ and hence the covariance of the driving Brownian motion entering Eq. 7.2-41 will be independent of $\underline{x}$. In addition, it is also assumed that $\underline{f}(\underline{x})$ is continuous in $\underline{x}$.

This section describes a method for estimating the first, second and, in principle, higher moments of the random variable $\tau - t_0$. As with the moment estimates (of the number of hyperplane crossings by a smooth random process) given in Section 7.2.2, these moments are to be viewed as a proxy for the complete information contained in the level crossing probability function $P(t < T, \underline{x})$ for $\underline{x} = \underline{x}_0$ at t_0 . However, this density function satisfies the partial differential equation

$$-\frac{\partial P}{\partial t}(t, \underline{x}) = A_{\underline{x}} P(t, \underline{x})$$

with

$$\begin{aligned} P(\underline{x}, t) &= 1, & t_0 \leq t < T, & \quad \underline{x}_0 \in \partial D \\ P(\underline{x}_0, t) &= 0, & \underline{x}_0 \in D \end{aligned} \quad (7.2-55)$$

and even the numerical solution of this equation in all but trivial cases in more than three dimensions is impossible (see Refs. 60 through 62).

The basic idea of our estimation method is as follows: choose a domain D specified via the level set of a suitable function $v(\underline{x})$, apply Dynkin's formula:

$$E \left\{ v(\underline{x}(\tau)) - v(\underline{x}_0) \mid \underline{x}_0 \right\} = E \left\{ \int_{t_0}^{\tau} A_{\underline{x}} v(\underline{x}) ds \mid \underline{x}_0 \right\}$$

and then by, estimating terms in the integral, extract a bound on $E\{(\tau - t_0) \mid \underline{x}_0\}$.

To be specific, consider the function $v(\underline{x}) = \underline{x}^T M \underline{x}$ with $M = M^T > 0$ and let $D = \{\underline{x}: \underline{x}^T M \underline{x} \leq c\}$ and $S = \partial D = \{\underline{x}: \underline{x}^T M \underline{x} = c\}$, $c > 0$.

Observe that $v(\underline{x})$ solves the equation

$$\begin{aligned} \Delta_Q v(\underline{x}) &\equiv \sum_{i,j=1}^n Q_{ij} \frac{\partial^2}{\partial x_i \partial x_j} v(\underline{x}) = 2 \sum_{i,j=1}^n Q_{ij} M_{ij} \\ &= 2 \operatorname{Tr}[QM] \end{aligned} \quad (7.2-56)$$

We shall denote $\operatorname{Tr}[QM]$ by ρ^{QM} in the calculations below, emphasizing the dependence upon Q and the matrix M characterizing the surface ∂D .

As in the derivation that led to Eq. 7.2-49, we have

$$E_{t_0} \{v(\underline{x}(\tau)) - v(\underline{x}_0)\} = E_{t_0} \left\{ \int_{t_0}^{\tau} A_{\underline{x}} v(\underline{x}(s)) ds \right\} \quad (7.2-57)$$

and hence

$$\begin{aligned} c - v(\underline{x}_{t_0}) &= E_{t_0} \left\{ \int_{t_0}^{\tau} \left[\frac{1}{2} \Delta_Q v(\underline{x}(s)) \right. \right. \\ &\quad \left. \left. + \sum_{i=1}^n f_i(\underline{x}(s)) \frac{\partial v(\underline{x}(s))}{\partial x_i} \right] ds \right\} \end{aligned} \quad (7.2-58)$$

(Since $\Delta_Q v = \rho^{QM}$, Eq. 7.2-58 yields

$$\begin{aligned} E_{t_0} \left\{ \int_{t_0}^{\tau} \left[\sum_{i=1}^n f_i(\underline{x}(s)) \frac{\partial v(\underline{x}(s))}{\partial x_i} \right] ds \right\} \\ = c - v(\underline{x}_0) - \frac{1}{2} \rho^{QM} E_{t_0} \{\tau - t_0\} \end{aligned} \quad (7.2-59)$$

However, this explicit expression involves $E_{t_0} \{\tau - t_0\}$, and our object is to estimate just this quantity.)

Equation 7.2-58 yields the inequality

$$\begin{aligned} c - v(\underline{x}_{t_0}) &\leq E_{t_0} \left\{ \int_{t_0}^{\tau} \frac{1}{2} \Delta_Q v(\underline{x}(s)) ds \right\} + \alpha E_{t_0} \left\{ \int_{t_0}^{\tau} 1 ds \right\} \\ &= E_{t_0} \{\tau - t_0\} \left[\frac{1}{2} \rho^{QM} + \alpha \right] \end{aligned} \quad (7.2-60)$$

yielding the lower bound

$$\left[\frac{c - v(\underline{x}_0)}{\frac{1}{2} \rho^{QM} + \alpha} \right] \leq E_{t_0} \{\tau - t_0\} \quad (7.2-61)$$

$$\text{when } \frac{1}{2} \rho^{QM} + \alpha > 0 \quad (7.2-62)$$

where

$$\begin{aligned} \alpha &= \max_{\underline{x} \in \bar{D}} \sum_{i=1}^n f_i(\underline{x}) \frac{\partial v(\underline{x})}{\partial x_i} \\ &\equiv \max_{\underline{x} \in \bar{D}} \underline{f}^T(\underline{x}) \operatorname{grad} v(\underline{x}), \quad \bar{D} \equiv \text{closure of } D, \\ &= \max_{\underline{x} \in \bar{D}} \left(2 \sum_{i,j=1}^n f_i(\underline{x}) Q_{ij} x_j \right), \\ &\quad (\text{by definition of } v(\underline{x})) \end{aligned} \quad (7.2-63)$$

In order to obtain an upper bound on $E\{\tau - t_0\}$, let

$$\min_{\underline{x} \in \bar{D}} \underline{f}^T(\underline{x}) \operatorname{grad} v(\underline{x}) \geq \gamma$$

Then Eq. 7.2-58 gives

$$c - v(\underline{x}_0) \geq E_{t_0} \{\tau - t_0\} \left[\frac{1}{2} \rho^{QM} + \gamma \right], \quad (7.2-64)$$

yielding the upper bound

$$\frac{c - v(\underline{x}_0)}{\frac{1}{2} \rho^{QM} + \gamma} \geq E_{t_0} \{ \tau - t_0 \}. \quad (7.2-65)$$

$$\text{when } \frac{1}{2} \rho^{QM} + \gamma > 0 \quad (7.2-66)$$

We make the following remarks:

(i) There is no assumption in the derivations above that $f(\underline{x})$ is linear in $\underline{x}$.

(ii) To estimate these bounds as tightly as possible one seeks $\hat{\alpha} \geq \alpha$ with $\hat{\alpha}$ as small as possible, and $\hat{\gamma} \leq \gamma$, $\hat{\gamma}$ as large as possible.

(iii) Equation 7.2-61 need not hold; for example, if $f(\underline{x}) = F\underline{x}$, then $\underline{f}^T(\underline{x}) \text{ grad } v(\underline{x}) = \underline{x}^T F^T M \underline{x}$, and for Eq. 7.2-61 to fail we only require $F^T M \not\leq 0$ with

$$\min_{\underline{x} \in \{ \underline{x}^T M \underline{x} = c \}} \underline{x}^T [F^T M + M F] \underline{x} + \text{Tr } QM \leq 0 \quad (7.2-67)$$

(iv) An extension of this method holds when $v(\underline{x})$ is not quadratic in $\underline{x}$, and hence $\Delta_Q v(\underline{x})$ is possibly a function of $\underline{x}$; see Example 4 below.

Higher Moments

We shall give an explicit calculation for the case of $E\{\tau - t_0\}$; with a prohibitive increase in complexity, the procedure may be continued for yet higher moments.

Following the original derivation of the formula for $E\{(\tau - t_0)^2\}$ via Dynkin's method, set

$$v_1(\underline{x}) = v^2(\underline{x}) \quad (7.2-68)$$

and notice that the level surfaces of $v_1(\underline{x})$ are identical to those of $v(\underline{x})$.

Proceeding as before

$$\begin{aligned} E_{t_0} \{ v_1(\underline{x}(\tau)) - v_1(\underline{x}_0) \} &= \\ E_{t_0} \left\{ \int_{t_0}^{\tau} \left[\frac{1}{2} \Delta_Q v_1(\underline{x}(s)) + \underline{f}^T(\underline{x}(s)) \text{ grad } v_1(\underline{x}(s)) \right] ds \right\} \\ &= E_{t_0} \left\{ \int_{t_0}^{\tau} \left[v(\underline{x}(s)) \Delta_Q v(\underline{x}(s)) + || \text{ grad } v(\underline{x}(s)) ||^2 \right. \right. \\ &\quad \left. \left. + 2 v(\underline{x}(s)) \underline{f}^T(\underline{x}(s)) \text{ grad } v(\underline{x}(s)) \right] ds \right\} \\ &\geq E_{t_0} \left\{ \int_{t_0}^{\tau} \left[\rho^{QM} v(\underline{x}(s)) + 2 v(\underline{x}(s)) \underline{f}^T(\underline{x}(s)) \right. \right. \\ &\quad \left. \left. \cdot \text{ grad } v(\underline{x}(s)) \right] ds \right\} \\ &\geq E_{t_0} \left\{ \int_{t_0}^{\tau} \rho^{QM} \left[c - \left(\frac{1}{2} \rho^{QM} + \alpha \right) E_s(\tau - s) \right] ds \right\} \end{aligned}$$

$$\begin{aligned} &+ 2 \mu_1 \int_{t_0}^{\tau} ds \Big\} \\ &= (\rho^{QM} c + 2 \mu_1) E_{t_0} \{ \tau - t_0 \} \\ &+ \frac{1}{2} \rho^{QM} \left(\frac{1}{2} \rho^{QM} + \alpha \right) E\{(\tau - t_0)^2\}, \quad (7.2-69) \end{aligned}$$

$$\text{when } \mu_1 \equiv \min_{\underline{x} \in \bar{D}} v(\underline{x}) \underline{f}^T(\underline{x}) \text{ grad } v(\underline{x}).$$

Hence we obtain

$$\begin{aligned} E_{t_0} \{ (\tau - t_0)^2 \} &\leq \frac{2}{\rho^{QM}} \left(\frac{1}{2} \rho^{QM} + \alpha \right)^{-1} (c^2 - v^2(\underline{x}_0)) \\ &- (\rho^{QM} c + 2 \mu_1) E_{t_0} \{ \tau - t_0 \} \quad (7.2-70) \end{aligned}$$

when Eq. 7.2-61 holds, and, since the lower bound of Eq. 7.2-61 is available for $E_{t_0} \{ \tau - t_0 \}$, an upper bound on $E_{t_0} \{ (\tau - t_0)^2 \}$ in terms of the constants ρ^{QM} , α , c , μ_1 is immediately available. As to a lower bound for $E_{t_0} \{ (\tau - t_0)^2 \}$, from Eq. 7.2-65 we have

$$\begin{aligned} &E_{t_0} \{ v_1(\underline{x}(\tau)) - v_1(\underline{x}_0) \} \\ &\leq E_{t_0} \left\{ \rho^{QM} \int_{t_0}^{\tau} v(\underline{x}(s)) ds \right\} + \lambda_1 E_{t_0} \left\{ \int_{t_0}^{\tau} ds \right\} \\ &+ \mu_2 E_{t_0} \int_{t_0}^{\tau} ds \\ &= \rho^{QM} c E_{t_0} \{ \tau - t_0 \} + \frac{1}{2} \rho^{QM} \left(\frac{1}{2} \rho^{QM} + \gamma \right) E_{t_0} \{ (\tau - t_0)^2 \} \\ &+ (\lambda_1 + \mu_2) E_{t_0} \{ \tau - t_0 \} \quad (7.2-71) \end{aligned}$$

$$\text{when } \mu_2 \equiv \max_{\underline{x} \in \bar{D}} v(\underline{x}) \underline{f}^T(\underline{x}) \text{ grad } v(\underline{x})$$

and

$$\lambda_1 = \max_{\underline{x} \in \bar{D}} || \text{ grad } v(\underline{x}(s)) ||^2 = \max_{\underline{x} \in \bar{D}} || \underline{x}^T M^2 \underline{x} || \quad (7.2-72)$$

Hence, when Eq. 7.2-65 holds,

$$\begin{aligned} E_{t_0} \{ (\tau - t_0)^2 \} &\geq \frac{2}{\rho^{QM}} \left(\frac{1}{2} \rho^{QM} + \gamma \right)^{-1} \left[(c^2 - v^2(\underline{x}_0)) \right. \\ &\quad \left. - (\rho^{QM} c + \lambda_1 + \mu_2) E_{t_0} \{ \tau - t_0 \} \right] \quad (7.2-73) \end{aligned}$$

with the upper bound of Eq 7.2-65 available to give a lower bound on $E_{t_0} \{(\tau-t_0)^2\}$ in terms of the constants ρ^{QM} , γ , c , λ_1 , and μ_2 .

Examples

In order to illustrate the application of the bounds in Eqs. 7.2-61 and 7.2-65, we need to consider $\max[\underline{f}^T(\underline{x}) \text{ grad } v(\underline{x})]$ and $\min[\underline{f}^T(\underline{x}) \text{ grad } v(\underline{x})]$ subject to $\underline{x}^T M \underline{x} \leq c$ $M > 0$

To begin with, we take the linear case $\underline{f}(\underline{x}) = F\underline{x}$, both for tractability and for relevance to the linear power systems models described earlier in this report. Let $L \triangleq F^T M + MF$. Then, by use of Lagrange multipliers, it is readily seen that if $L \not\leq 0$ or $L \not\geq 0$, respectively, the maximum and minimum values of $\underline{f}^T(\underline{x}) \text{ grad } v(\underline{x}) = 2 \underline{x}^T F^T M \underline{x} = 2 \underline{x}^T L \underline{x}$, subject to $\underline{x}^T M \underline{x} \leq c$, are given by the procedure

- (i) Find all eigenvectors $\tilde{x}_1$ of $M^{-1}L$
- (ii) Normalize $\tilde{x}_1$ so that $\tilde{x}_1^T M \tilde{x}_1 = c$, $1 \leq i \leq n$
- (iii) Pick $\tilde{x}_1$ such that $\tilde{x}_1^T L \tilde{x}_1 = \max$ or $\min$.

(Notice that $L \geq 0$ makes the corresponding minimum zero, and $L \leq 0$ the corresponding maximum zero.)

So the computational task of finding the bounds α and γ is equivalent to that of finding the eigenvectors of $M^{-1}L$. In case $M = I$ the bounds are

$$\frac{c-v(\underline{x}_0)}{\frac{1}{2} \rho^{QM} + c \max_{1 \leq i \leq n} \lambda_1 (F+F^T)} \leq E\{\tau-t_0\} \leq \frac{c-v(\underline{x}_0)}{\frac{1}{2} \rho^{QM} + c \min_{1 \leq i \leq n} \lambda_1 (F+F^T)} \quad (7.2-74)$$

where λ_1 is the i^{th} eigenvector.

EXAMPLE 0: Brownian Motion

We observe that when $F = 0$ we are in the special case of a multidimensional Brownian motion (crossing $\|\underline{x}\|^2 = c$) and in that case $\lambda_1(F+F^T) = 0$ for all i and we have the exact formula

$$E\{\tau-t_0\} = \frac{c - \|\underline{x}_0\|^2}{\frac{1}{2} \text{Trace } [Q]} \quad (7.2-75)$$

with, of course, $c - \|\underline{x}_0\|^2 \geq 0$.

We remark that in the asymptotically stable case we have $\lambda_1(F+F^T) < 0$ for all i . Hence, as the maximum real part of the eigenvalues of F goes toward $-\frac{1}{2} \rho^{QM}$, from 0, we get $E\{\tau-t_0\} \rightarrow \infty$, which fits intuition (and

other information) about asymptotically stable systems.

For unstable F , we see that as $\text{Re } \lambda_1(F) \rightarrow \infty$, the lower bound goes to 0, which is uninformative, but the upper bound also goes to zero, which, again, fits intuition (and other information) on the unstable case

EXAMPLE 1: Exact Bounds with $L = MF^T + FM$ Indefinite

Consider the system $\dot{\underline{x}} + \underline{x} = \dot{\underline{B}}$ ($\dot{\underline{B}}$ = "derivative" of Brownian motion). More precisely,

$$dx_1 = x_2 dt$$

$$dx_2 = -x_1 dt + dw$$

$$\text{or } d\underline{x} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \underline{x} dt + \begin{bmatrix} 0 \\ 1 \end{bmatrix} dw \quad (7.2-76)$$

Then

$$\begin{aligned} A_x &= \frac{1}{2} \Delta_Q - x_1 \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_1} \\ &= \frac{1}{2} \frac{\partial^2}{\partial x_2^2} - x_1 \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_1}, \end{aligned} \quad (7.2-77)$$

$$\text{with } Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$$

Let $v(x_1, x_2) = x_1^2 + x_2^2$ i.e. $M=I$. Then $\rho^{QM}=2$ $\text{Tr } [QM] = 2 = \Delta_Q v(\underline{x})$. Take $v(x_1, x_2) = r^2 = c$ as the level surface of interest and consider the case where the process starts at the origin at $t=0$. Then we can solve the associated Dynkin equation (the solution is $(r^2 - (x_1^2 + x_2^2))$) to get $r^2 = E\{\tau_{r^2}\}$, where τ_{r^2} denotes the first crossing time of a sphere of radius r . In particular, $E\{\tau_1\}=1$. In this case, the upper and lower bounds in Eq 7.2-74 are identical, since $\lambda_1(F+F^T) = \lambda_1(0) = 0$, and we have

$$E\{\tau-t_0\} = \frac{c-v(\underline{x}_0)}{\frac{1}{2} \rho^{QM}} = r^2, \quad (7.2-78)$$

in agreement with the solution just derived

EXAMPLE 2: Exact Bounds With Elliptic Regions

Consider $\dot{\underline{x}} + \omega^2 \underline{x} = \dot{\underline{B}}$, i.e.,

$$d\underline{x} = \begin{bmatrix} 0 & 1 \\ -\omega^2 & 0 \end{bmatrix} \underline{x} dt + \begin{bmatrix} 0 \\ 1 \end{bmatrix} dw \quad (7.2-79)$$

but now take the level surfaces of the function

$$v(\underline{x}) = x_2^2 + \omega^2 x_1^2 \text{ -- so here } M = \begin{bmatrix} \omega^2 & 0 \\ 0 & 1 \end{bmatrix}; \text{ then}$$

$$\begin{aligned}
A_x &= \frac{1}{2} \Delta_Q - (w^2 x_1) \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_1} \\
&= \frac{1}{2} \frac{\partial}{\partial x_2^2} - (w^2 x_1) \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_1},
\end{aligned}
\tag{7.2-80}$$

with $Q = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}$ Further,

$$\rho^{QM} = 2 \operatorname{Tr}[QM] = 2 = \Delta_Q v(\underline{x}),$$

and

$$L = MF + P^T M = \begin{bmatrix} w^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -w^2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -w^2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} w^2 & 0 \\ 0 & 1 \end{bmatrix} = 0.
\tag{7.2-81}$$

Hence the eigenvalues of $M^{-1}L$ are all 0.

Now in this case we can again solve Dynkin's equation to obtain the solution $r^2 - (w^2 x_1^2 + x_2^2)$ and so $E\{\tau_{r^2}\} = r^2$. Hence, we see that we again obtain the exact "estimates"

$$\frac{c-v(\underline{x}_0)}{\frac{1}{2}\rho^{QM} + \alpha} = \frac{c-v(\underline{x}_0)}{\frac{1}{2}\rho^{QM} + \gamma} = \frac{r^2 - 0}{\frac{2}{2} + 0} = r^2 = E\{\tau_{r^2}\}
\tag{7.2-82}$$

EXAMPLE 3: Damped System

Consider the damped system

$$dx_1 = x_2 dt \tag{7.2-83}$$

$$dx_2 = -x_1 dt - 2\eta x_2 dt + dw, \quad \eta \geq 0$$

Hence

$$A_x = \frac{1}{2} \frac{\partial^2}{\partial x_2^2} - (x_1 + 2\eta x_2) \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_1}
\tag{7.2-84}$$

The simple addition of the damping term makes the corresponding Dynkin equation analytically intractable. However, our estimation procedure goes as follows:

$$v(\underline{x}) = x_1^2 + x_2^2 \text{ yields } M = I; \text{ clearly}$$

$$F = \begin{bmatrix} 0 & 1 \\ -1 & -2\eta \end{bmatrix}
\tag{7.2-85}$$

and so

$$\lambda_i (M^{-1}L) \equiv \lambda_i (M^{-1}[MF + F^T M])$$

$$= \lambda_i \begin{bmatrix} 0 & 1 \\ -1 & -2\eta \end{bmatrix} + \begin{bmatrix} 0 & -1 \\ 1 & -2\eta \end{bmatrix} = \lambda_i \begin{bmatrix} 0 & 0 \\ 0 & -4\eta \end{bmatrix}
\tag{7.2-86}$$

and so we have the eigenvector, eigenvalue pairs

$$\lambda_1 = 0, \quad \underline{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}; \quad \lambda_2 = -4\eta, \quad \underline{x}_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \tag{7.2-87}$$

Normalizing so that $\underline{x}_i^T M \underline{x}_i = r^2$, yields

$$\lambda_1 = 0, \quad \tilde{\underline{x}}_1 = \begin{bmatrix} r^2 \\ 0 \end{bmatrix}; \quad \lambda_2 = -4\eta, \quad \tilde{\underline{x}}_2 = \begin{bmatrix} 0 \\ r^2 \end{bmatrix} \tag{7.2-88}$$

and hence

$$\begin{aligned}
\max_{1 \leq i \leq 2} \tilde{\underline{x}}_i^T L \tilde{\underline{x}}_i &= [r^2 \ 0] \begin{bmatrix} 0 & 0 \\ 0 & -4\eta \end{bmatrix} \begin{bmatrix} r^2 \\ 0 \end{bmatrix} = 0, \quad (\text{at } i = 1) \\
&\tag{7.2-89}
\end{aligned}$$

and

$$\begin{aligned}
\min_{1 \leq i \leq 2} \tilde{\underline{x}}_i^T L \tilde{\underline{x}}_i &= [0 \ r^2] \begin{bmatrix} 0 & 0 \\ 0 & -4\eta \end{bmatrix} \begin{bmatrix} 0 \\ r^2 \end{bmatrix} = -4\eta r^2, \\
&\tag{at } i=2 \tag{7.2-90}
\end{aligned}$$

Hence, from Eqs. 7.2-61 and 7.2-65, for $\eta < \frac{1}{4r^2}$

$$r^2 \leq E\{\tau_{r^2}\} \leq \frac{r^2}{1 - (4\eta)r^2}, \tag{7.2-91}$$

with the degenerate case $\eta=0$ agreeing with Example 1.

EXAMPLE 4: Nonlinear, Damped System

Consider

$$dx_1 = x_2 dt$$

$$dx_2 = -x_1^3 dt - 2\eta x_2 dt + dw \tag{7.2-92}$$

here

$$A_x = \frac{1}{2} \frac{\partial^2}{\partial x_2^2} - (x_1^3 + 2\eta x_2) \frac{\partial}{\partial x_2} + x_2 \frac{\partial}{\partial x_1},
\tag{7.2-93}$$

and systems of this general type also yield intractable Dynkin's equations.

Now take the function $v(\underline{x})$ to be $\frac{1}{2} x_1^4 + x_2^2$ and the level surface of interest to be $v(\underline{x}) = r^2$. Then the maximum and the minimum of

$$\begin{aligned} \underline{f}^T(\underline{x}) \text{ grad } v(\underline{x}) &= [x_2, -x_1^3 - 2\eta x_2] \begin{bmatrix} 2x_1^3 \\ 2x_2 \end{bmatrix} \\ &= 2x_1^3 x_2 - 2x_1^3 x_2 - 4\eta x_2^2 \end{aligned} \quad (7.2-94)$$

over $\frac{1}{2} x_1^4 + x_2^2 \leq r$ are 0 and $-4\eta r^2$ respectively. Hence we obtain the estimate

$$\frac{r^2}{1} \leq E\{\tau_2\} \leq \frac{r^2}{1 - (4\eta)r^2} \quad (7.2-95)$$

EXAMPLE 5: Interlinked Undamped Systems

Suppose the system

$$d\underline{z} = \underline{F} \underline{z} dt + \underline{G} d\underline{w} \quad (7.2-96)$$

is such that, for $\underline{x} = \underline{T}_J \underline{z}$,

$$d\underline{x} = \underline{F}_J \underline{x} dt + \underline{G}_J d\underline{w} \quad (7.2-97)$$

where $\underline{F}_J$ has $2n$ 1×1 Jordan blocks $\pm i\omega_k$. We regard such a system as a general set of interlinked undamped oscillators.

Let a second state space transformation $\underline{T}$ be applied such that $\Delta \equiv \underline{T} \underline{F}_J \underline{T}^{-1}$ has n 2×2 blocks of the form $\begin{bmatrix} 0 & 1 \\ -\omega_k^2 & 0 \end{bmatrix}$. Take $\underline{M} \equiv \underline{T}^T \underline{N} \underline{T} \equiv \underline{T}^T [\text{diag} \begin{bmatrix} \omega_k^2 & 0 \\ 0 & 1 \end{bmatrix}] \underline{T}$. So then, $\partial D = \{\underline{x} : \underline{x}^T \underline{T}^T \underline{N} \underline{T} \underline{x} = r^2\}$

and

$$\underline{L} = (\underline{M} \underline{F}_J + \underline{F}_J^T \underline{M}) = \underline{T}^T [\underline{N} \Delta + \Delta^T \underline{N}] \underline{T} = 0 \quad (7.2-98)$$

since

$$\begin{bmatrix} \omega_k^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ -\omega_k^2 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -\omega_k^2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_k^2 & 0 \\ 0 & 1 \end{bmatrix} = 0 \quad (7.2-99)$$

for each k . Hence, for all matrices $\underline{G}$ in the original system, we may exactly estimate $E\{\tau_2\}$ by $\frac{2r^2}{\rho_{QM}}$ with

$$\underline{Q} = \underline{T} \underline{G}_J \underline{Q}_w \underline{G}_J^T \underline{T}^T$$

and the given $\underline{M}$.

7.3 STABILITY RESULTS FOR STOCHASTIC DIFFERENTIAL EQUATIONS WITH DISCRETE JUMP COEFFICIENTS

7.3.1 Background

The discussion in this section is concerned with the general linear equation

$$\dot{\underline{x}}(t) = \underline{F}(t) \underline{x}(t), \quad \underline{x}(0) = \underline{x}_0, \quad t \geq 0 \quad (7.3-1)$$

where $\underline{x}(t)$ is an n vector and $\underline{F}(t)$ is a real $n \times n$ matrix. Two basic results from ordinary differential equation theory (see e.g., Ref. 63) that are used in the sequel are stated below.

Theorem 1 - In the system defined by Eq. 7.3-1, assume that $\underline{F}$ is a constant matrix, let $[\lambda_i]$ denote the set of eigenvalues of $\underline{F}$, and set

$$\eta = \max_i \text{Re}(\lambda_i) \quad (7.3-2)$$

Then there exists a constant $b > 0$ such that

$$||\phi(t)|| \leq b e^{\eta t} \quad (7.3-3)$$

where $\phi(t)$ is the fundamental (or transition) matrix for Eq. 7.3-1 and

$$||C(t)|| \equiv (\sum_{ij} C_{ij}^2)^{1/2} \quad (7.3-4)$$

Theorem 2 (Wintner, Ref. 64) - For time-varying $\underline{F}(t)$, set

$$\underline{H}(t) = \underline{F}(t) + \underline{F}^T(t) \quad (7.3-5)$$

which has characteristic values $(\lambda_i(t))$. Let

$$\lambda_{\max}(t) \equiv \max_i [\lambda_i(t)] \quad (7.3-6)$$

then the solution $\underline{x}(t)$, with initial value $\underline{x}_0$ satisfies

$$||\underline{x}(t)||^2 \leq ||\underline{x}_0||^2 e^{\int_0^t \lambda_{\max}(\tau) d\tau} \quad (7.3-7)$$

Proof of Theorem 2:

Since $||\underline{x}(t)||^2 = \underline{x}^T(t) \underline{x}(t)$, differentiation gives

$$\begin{aligned} \frac{d||\underline{x}(t)||^2}{dt} &= \dot{\underline{x}}^T(t) \underline{x}(t) + \underline{x}^T(t) \dot{\underline{x}}(t) \\ &= \underline{x}^T(t) \underline{F}^T(t) \underline{x}(t) + \underline{x}^T(t) \underline{F}(t) \underline{x}(t) \\ &= \underline{x}^T(t) [\underline{F}^T(t) + \underline{F}(t)] \underline{x}(t) \end{aligned} \quad (7.3-8)$$

or

$$\frac{1}{\|\underline{x}(t)\|^2} \frac{d\|\underline{x}(t)\|^2}{dt} = \frac{\underline{x}^T(t) [F^T(t) + F(t)] \underline{x}(t)}{\|\underline{x}(t)\|^2} \quad (7.3-9)$$

From the properties of quadratic forms,

$$\lambda_{\min}(t) \leq \frac{\underline{x}^T(t) [F^T(t) + F(t)] \underline{x}(t)}{\|\underline{x}(t)\|^2} \leq \lambda_{\max}(t) \quad (7.3-10)$$

and, the result follows from

$$\frac{1}{\|\underline{x}(t)\|^2} \frac{d\|\underline{x}(t)\|^2}{dt} \leq \lambda_{\max}(t) \quad (7.3-11)$$

7.3.2 Equations with Jump Coefficients

To include jump coefficients, assume that the system of Eq 7.3-1 is random in the sense that the matrix $F(t)$ contains elements $\{f_{ij}(t)\}$ that are discrete state Markov processes. Thus, each f_{ij} can take any one of a possible set of values, $\{\alpha_1, \alpha_2, \dots, \alpha_N\}$. Assume further that the time spent in each state is a random variable whose distribution function may depend upon the state. Specifically, assume that the times (usually called the "sojourn time") spent in a state are independent, identically distributed random variables, such that with probability one only a finite number of switchings can occur in any finite time. Such processes are referred to as semi-Markov processes (see Chung, Ref. 65) with sample functions that are step functions.

For a random matrix process as defined above, the system of Eq. 7.3-1 is a constant coefficient system between the random transitions or switchings. Clearly, the sample solutions exist and are unique and continuous in each region. Since the boundary conditions are matched at each switching time, the solution process is an almost surely continuous stochastic process.

Before proceeding, a further assumption will be made about the semi-Markov process coefficients of the system -- that the coefficient matrix $F(t)$ is a stationary positive recurrent Markov chain with statistically independent sojourn times which may depend upon the state. (That is, assume that the chain is an ergodic chain, in which all states can recur).

Notice, that a Markov chain is determined by its initial probability as well as its transition matrix,

$$P = (p_{ij}), \text{ where } p_{ij} = \Pr\{x_k = j | x_{k-1} = i\}$$

and $\sum_{j=1}^N p_{ij} = 1$ for each $i=1, \dots, N$. If the chain is ergodic, then there is a stationary probability of being in any state called $\pi = (\pi_1, \dots, \pi_N)$, independent of the initial probability, $(p_0)_j = \Pr\{F(0)=F_j\}$. This condition implies that

$$\lim_{P_0^T} P^m = \pi^T; \text{ for all } P_0^T \quad (7.3-12)$$

or equivalently,

$$\pi P = \pi \text{ or } P^T \pi^T = \pi^T \quad (7.3-13)$$

This P (or P^T) must have an eigenvalue $\lambda = 1$ with π the corresponding eigenvector. A necessary and sufficient condition is

$$\begin{bmatrix} P_{11}^{-1} & P_{12} & \dots & P_{1N} \\ P_{21} & P_{22}^{-1} & \dots & P_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ P_{N1} & P_{N2} & \dots & P_{NN}^{-1} \end{bmatrix} = 0 \quad (7.3-14)$$

By adding the sum of the second through n^{th} columns to the first column and noting that the row sums equal unity, viz.,

$$\begin{bmatrix} (\sum_j P_{1j})^{-1} & P_{12} & \dots \\ (\sum_j P_{2j})^{-1} & P_{22}^{-1} & \dots \\ \vdots & \vdots & \ddots \\ (\sum_j P_{Nj})^{-1} & P_{N2} & \dots \end{bmatrix} = \begin{bmatrix} 0 & P_{12} & \dots \\ 0 & P_{22}^{-1} & \dots \\ 0 & \vdots & \ddots \\ 0 & P_{N2} & \dots \end{bmatrix} = 0 \quad (7.3-15)$$

which satisfies the condition.

In the remainder of this section assume that for the j^{th} state of the matrix $F(t)$ (denoted by the matrix F_j) the corresponding terms in Theorem 1 are (b_j, η_j) and $E\{\tau_j\}$ is the expected sojourn time for the state F_j . The following theorem states the major result obtained for the system defined by Eq. 7.3-1.

Theorem 3 - Let $(b_j, \eta_j, E\{\tau_j\}, \pi_j)$ be defined as above, then a sufficient condition for almost sure asymptotic stability is that

$$\sum_{j=0}^N \pi_j \eta_j E\{\tau_j\} + \sum_{j=0}^N \pi_j \log(b_j) < 0 \quad (7.3-16)$$

Proof:

Denote the (random) switching times by, $t_1 < t_2 < t_3 < \dots$, and let the state of the matrix in the interval $t_k \leq t < t_{k+1}$ be F_{j_k} , where F_{j_k} is an element of $\{F_1, F_2, \dots, F_N\}$. Then, if the initial value at time t_k is $\underline{x}(t_k) = \underline{x}_k$,

$$\underline{x}(t) = e^{F_{j_k}(t-t_k)} \underline{x}_k \quad (7.3-17)$$

in the time interval $t_k \leq t < t_{k+1}$. Therefore,

$$\begin{aligned} \underline{x}(t_{k+1}) &= \underline{x}_{k+1} = e^{F_{J_k}(t_{k+1}-t_k)} \underline{x}_k \\ &= e^{F_{J_k} \tau_{J_k}} \underline{x}_k = \Phi_k \underline{x}_k \end{aligned} \quad (7.3-18)$$

where τ_{J_k} is the sojourn time for the state F_{J_k} . Thus, successively applying Eq. 7.3-18 at each switching time yields,

$$\underline{x}_{k+1} = \Phi_k \Phi_{k-1} \dots \Phi_1 \Phi_0 \underline{x}_0 \quad (7.3-19)$$

Upon taking norms,

$$\|\underline{x}_{k+1}\| \leq \prod_{i=0}^k \|\Phi_i\| \|\underline{x}_0\| \quad (7.3-20)$$

However, from Theorem 1 on constant coefficient systems,

$$\|\Phi_i\| \leq b_{J_i} e^{\eta_{J_i} \tau_{J_i}} \quad (7.3-21)$$

therefore

$$\|\underline{x}_{k+1}\| \leq e^{\sum_{i=0}^k \eta_{J_i} \tau_{J_i} + \sum_{i=0}^k \log b_{J_i}} \|\underline{x}_0\| \quad (7.3-22)$$

Now since irreducible, positive, recurrent Markov chains are ergodic, they satisfy a law of large numbers. Thus,

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k \log b_{J_i} = E\{\log b\} = \sum_{j=1}^N \pi_j \log b_j \quad \text{a.s.} \quad (7.3-23)$$

and

$$\begin{aligned} \lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k \eta_{J_i} \tau_{J_i} &= E\{\eta \tau\} = \sum_{j=1}^N \pi_j \eta_j \tau_j \\ &= E\{\eta E[\tau|\eta]\} \\ &= \sum_{j=1}^N \pi_j \eta_j E\{\tau|\eta=\eta_j\} \\ &= \sum_{j=1}^N \pi_j \eta_j E\{\tau_j\} \quad \text{a.s.} \end{aligned} \quad (7.3-24)$$

(These results can be obtained in a more direct way by counting the number of times a given state is visited and taking the limits.) Now, taking the limit in Eq. 7.3-22, viz.,

$$\lim_{k \rightarrow \infty} \|\underline{x}_{k+1}\| \leq$$

$$\lim_{k \rightarrow \infty} e^{k \left[\frac{1}{k} \sum_{i=0}^k \eta_{J_i} \tau_{J_i} + \frac{1}{k} \sum_{i=0}^k \log b_{J_i} \right]} \|\underline{x}_0\| \quad (7.3-25)$$

and using Eqs. 7.3-4 and 7.3-5 gives

$$\sum_{j=0}^N \pi_j \eta_j E\{\tau_j\} + \sum_{j=0}^N \pi_j \log b_j < 0 \quad (7.3-26)$$

which is the sufficient condition.

The result can be made sharper in the following fashion by applying Theorem 2 which is due to Wintner (Ref. 64). Suppose the maximum characteristic value of $F + F^T$ is denoted λ_j , which must be real since $F_j + F_j^T$ is symmetric. In the k^{th} time-interval $t_k \leq t < t_{k+1}$, $\|\underline{x}\|^2$ can be written as the solution to the differential equation

$$\frac{d\|\underline{x}(t)\|^2}{dt} = \underline{x}^T(t) (F_{J_k} + F_{J_k}^T) \underline{x}(t) \quad (7.3-27)$$

(This follows from Theorem 2.) Dividing Eq. 7.3-27 by $\|\underline{x}(t)\|^2$ yields

$$\frac{1}{\|\underline{x}(t)\|^2} \frac{d\|\underline{x}(t)\|^2}{dt} = \frac{\underline{x}^T(t) (F_{J_k} + F_{J_k}^T) \underline{x}(t)}{\|\underline{x}(t)\|^2} \quad (7.3-28)$$

Therefore, in the interval $t_k \leq t < t_{k+1}$,

$$\|\underline{x}(t)\|^2 \leq e^{\lambda_{J_k}(t-t_k)} \|\underline{x}_k\|^2 \quad (7.3-29)$$

Using a similar argument at each successive switching time yields

$$\|\underline{x}_{k+1}\|^2 \leq e^{\sum_{i=0}^k \lambda_{J_i} \tau_{J_i}} \|\underline{x}_0\|^2 \quad (7.3-30)$$

As in Theorem 3 above, simply take the limit of the exponent in Eq. 7.3-30 to obtain

$$\lim_{k \rightarrow \infty} \frac{1}{k} \sum_{i=0}^k \lambda_{J_i} \tau_{J_i} = E\{\lambda \tau\} = \sum_{j=1}^N \pi_j \lambda_j E\{\tau_j\} \quad \text{a.s.} \quad (7.3-31)$$

by the assumed ergodic properties (law of large numbers) of the chain. This, from Eq. 7.3-30,

$$\lim_{k \rightarrow \infty} ||x_{k+1}||^2 \leq \lim_{k \rightarrow \infty} e^{\left[\frac{1}{k} \sum_{i=0}^k \lambda_{j_1} \tau_{j_1} \right]} ||x_0||^2 \quad (7.3-32)$$

From the above inequality it is immediate that the sufficient condition is

$$\sum_{j=1}^N \pi_j \lambda_j E\{\tau_j\} < 0 \quad (7.3-33)$$

Note that it is possible using a scheme first proposed by Infante (Ref 66) to modify, somewhat, the approach discussed above. Consider a quadratic Lyapunov function

$$V(\underline{x}) = \underline{x}^T B \underline{x}, \quad (7.3-34)$$

where B is positive definite. Then, in the interval $t_k \leq t < t_{k+1}$,

$$\begin{aligned} & \frac{1}{V(\underline{x}(t))} \frac{dV(\underline{x}(t))}{dt} \\ &= \frac{\underline{x}(t)(F_{jk} B + B F_{jk}) \underline{x}(t)}{\underline{x}^T(t) B \underline{x}(t)} \leq \xi_{jk} \end{aligned} \quad (7.3-35)$$

where ξ_{jk} is the maximum eigenvalue of $[F_{jk} + B F_{jk} B^{-1}]$.

Equation 7.3-11 corresponds to Eq. 7.3-7. This leads to a formula equivalent to Eq. 7.3-9 which is

$$\sum_{j=1}^N \pi_j \xi_j E\{\tau_j\} < 0. \quad (7.3-36)$$

7.4 SUMMARY

The two security measures described in this chapter are tailored to different aspects of the long-term security assessment function. The transition from the normal operating state to the alert or emergency state can naturally be viewed in terms of a general surface crossing problem for a nonlinear stochastic dynamical system. The drawback to formulating the problem with this level of generality is the impossibility of obtaining a (numerical) solution in other than trivial cases. By imposing simplifying assumptions of the process and surface, it is possible to use scalar level crossing theory to obtain expressions for the moments of the number of surface crossings in any time interval of interest. An approximation to the probability that the process crosses the surface in this interval can be obtained from a finite number of these moments. While this approach is computationally feasible, its efficacy is dependent upon the degree to which the requisite smooth dynamical system model and separating hyperplane correspond to the actual processes and security regions of interest. These issues need to be addressed prior to the use of this technique.

The alternate method of analyzing this problem via a partial differential equation of Dynkin also yields the moments of the first crossing time without requiring any simplifying assumptions. However, it is

in a practical sense impossible to compute solutions to the moment equations for systems of realistic dimension. This difficulty can be circumvented by computing upper and lower bounds on the moments via much simpler equations. The utility of this approach is dependent upon the tightness of the bounds. This issue should also be examined in the context of a realistic power system model.

The sufficient condition for almost sure asymptotic stability developed in the previous section is readily computed but may be of limited value because of the necessary assumption that the system dynamics are linear. Consequently, this security measure might best be employed in conjunction with other, more detailed, measures.

8. A SECURITY MEASURE FOR TRANSIENT STABILITY

The transient behavior of power systems can be studied using Lyapunov stability theory (e.g., Refs. 67 and 68). This approach has the advantage that qualitative behavior can be assessed without explicitly solving the differential equations of motion. The research documented in this chapter has concentrated on the development of a security measure based on Lyapunov methods which is suitable for on-line evaluation. While there is a considerable amount of literature on the application of Lyapunov methods to off-line power system transient stability analysis, results for on-line evaluation are almost nonexistent. The thrust of the two approaches is conceptually different. Off-line procedures are an attempt to provide highly accurate estimates of critical clearing times without having to utilize step-by-step integration techniques. Current research by Athay (Ref. 68) is directed along these lines. Conversely, the goal for an on-line procedure is to provide an indication to the dispatcher of system susceptibility to loss of synchronism in the presence of random network faults. If needed, suitable corrective action can then be taken. Approximate (conservative) measures may prove to be sufficient to indicate a potential weakness in the grid.

Typical off-line analysis begins by assuming a specific fault and a specific post-fault network topology. These specifics are not known a priori for on-line analysis. Since transient stability must be evaluated with respect to the post-fault network, as well as the specific fault locations, on-line stability evaluation using pre-fault information must consider all possible network faults and post-fault network configurations. Transient stability margins must consider the inherent stability of the post-fault network. The pre-fault margin is irrelevant except in those cases where the pre-fault and post-fault configurations are identical. Also essential is the kinetic energy acquired during the fault-on period which is a function of the time it takes to clear the fault as well as the specific fault location. These quantities need to be estimated prior to the occurrence of the fault. Thus, emphasis should be placed on devising simplifying assumptions which reduce computational complexity to the point where the security measure can be computed in real-time with typical computers in use by dispatchers. It is intended, by design, that any simplifications result in a more conservative security measure.

8.1 A STOCHASTIC LEVEL CROSSING PROBLEM

In this section, a measure transient stability is formulated as a stochastic level crossing problem. This formulation is obtained in terms of a Lyapunov function. Lyapunov stability concerns the behavior of solutions to a system of differential equations of the form

$$\dot{\underline{x}} = \underline{f}(\underline{x}) \quad (8.1-1)$$

in the vicinity of an equilibrium point $\underline{x}_0$ which satisfies

$$\underline{f}(\underline{x}_0) = 0 \quad (8.1-2)$$

The stability analysis depends on finding a function $V(\underline{x}-\underline{x}_0)$ which has the following properties

$$V(\underline{x}-\underline{x}_0) > 0; \quad \underline{x} \neq \underline{x}_0 \quad (8.1-3)$$

$$V(0) = 0 \quad (8.1-4)$$

$$\dot{V}(\underline{x}-\underline{x}_0) \leq 0 \quad (8.1-5)$$

If the above conditions hold in a closed region about the equilibrium point, then $|\underline{x}-\underline{x}_0|$ remains bounded and the system is said to be stable. If strict inequality holds in Eq. 8.1-5, then $\underline{x}$ approaches $\underline{x}_0$ asymptotically and $\underline{x}$ is asymptotically stable. One approach (Ref. 67) which has been used to establish this region of stability is based on the set of unstable equilibrium solutions $\{\underline{x}_i^u\}$ to Eq. 8.1-2. The conditions established in Eqs. 8.1-3 through 8.1-5 are satisfied if

$$V(\underline{x}-\underline{x}_0) \leq \bar{V} = \min_i V(\underline{x}_i^u - \underline{x}_0) \quad (8.1-6)$$

It has been pointed out (Ref. 68) that this region is overly conservative since the lowest unstable equilibrium may be far from the actual post-fault trajectory. If the post-fault trajectory were known, then the nearest unstable equilibrium is usually the governing critical point for first-swing instability.

Regardless of how $\bar{V}$ is chosen, Eq. 8.1-6 establishes the criterion for stability in terms of a level crossing problem. In the remainder of this section, a stochastic level crossing problem is formulated which indicates how the probabilistic occurrence of disturbances can be incorporated in defining a transient stability security measure.

Consider the Lyapunov function V to be dependent upon the known system parameters (reactances, inertias, etc.) $\underline{a}$ which define the transient dynamics, the unknown network disturbance parameters $\underline{b}_m$, and the electrical state (generator rotor speed and angle deviations from the system reference) $\underline{x}$ of the power system:

$$V = V(\underline{a}, \underline{b}_m, \underline{x})$$

The index m refers to the topological description of the disturbance (e.g., a single-phase-to-earth fault in a particular line), and $\underline{b}_m$ is the parameteric description of the disturbance (e.g., the position of the fault on the line relative to the protective circuit breakers). Both $\underline{b}_m$ and $\underline{x}$ are continuous random variables, and m is a discrete random variable. Assume that

- The system is in steady state at time t_0 with known probability density function $p_{t_0}(\underline{x}_0)$, obtained with the aid of a state estimator.
- The m^{th} disturbance is characterized by a homogeneous Poisson process with rate λ_m

- The probability density p_m for b_m is known, and is independent of $x(t)$

This analysis addresses the possibility of a single disturbance occurring at some time $t_1 > t_0$. The probability of a particular disturbance occurring is denoted P_m , which is simply the probability that that disturbance occurs first. In terms of the Poisson process models,

$$P_m = \frac{\lambda_m}{\sum_{m=1}^N \lambda_m} \quad (8.1-7)$$

where N is the number of possible different disturbances.

If a disturbance occurs after time t_0 the system will remain stable if

$$\Delta V_m(t) \equiv \bar{V}(a, b_m) - V(a, b_m, x(t)) > 0 \quad (8.1-8)$$

where $\bar{V}$ is the minimum saddle point of V . Both $\bar{V}$ and V are specified in terms of the post-disturbance network configuration. Thus, a measure of security over the interval $t_0 \leq t \leq t_1$, with respect to the m^{th} disturbance, is defined by

$$S_m(t_0, t_1) \equiv \Pr\{\Delta V_m(t) > 0, t_0 \leq t \leq t_1\} \quad (8.1-9)$$

which is a level crossing condition.

With the above stated Poisson process assumption for the m^{th} disturbance, the probability that the disturbance will occur in the time-interval (t_0, t_1) is

$$\lambda_m e^{-\lambda_m(t_1-t_0)} (t_1 - t_0)$$

given that the m^{th} disturbance is the first one to occur. Furthermore, the probability that it will occur in (t_0, t_1) is independent of the probability that it will occur in (t_1, t_2) .

Now, divide time into successive intervals of length Δt . It follows from the above that

$$S_m(t_0, t_n) \equiv \lambda_m \Delta t \sum_{j=1}^n e^{-\lambda_m(t_j-t_{j-1})} \int_{\bar{V}-V>0} p_m(b_m) \times p_t(x(t_j)) db_m dx_j \quad (8.1-10)$$

where

$$t_j = j\Delta t$$

and the integration is performed over the set of values for b_m and $x(t_j)$ for which Eq. 8.1-8 is satisfied. Finally, averaging S_m over all disturbances, yields the measure of system security,

$$S(t_0, t_n) = \sum_m P_m \lambda_m \Delta t \sum_{j=1}^n e^{-\lambda_m(t_j-t_{j-1})} \times \int_{\bar{V}-V>0} p_m(b_m) p_t(x(t_j)) db_m dx_j \quad (8.1-11)$$

Equation 8.1-11 requires a great deal of computation, too much to be done on-line. However, the quantities needed in Eq. 8.1-11 are presumed known or can be computed in terms of known quantities. The analysis presented in the next section is an attempt to reduce significantly the amount of computation needed to evaluate an approximation to this transient security measure in order to make on-line evaluation feasible

8.2 A SIMPLIFIED LEVEL CROSSING PROBLEM

The proposed method for on-line assessment of transient stability involves computation of the probability that the Lyapunov function will not exceed a critical value in a specified interval of time. To simplify the computations, the time-interval will be chosen to be sufficiently short that the following assumptions hold:

- No more than one fault occurs in the interval $t_0 \leq t \leq t_1$
- The specific fault does not occur prior to time t_0
- The system is in steady state at time t_0 and will remain so during the interval $t_0 \leq t \leq t_1$ if the fault does not occur
- A state estimator provides accurate estimates of all parameters needed for evaluating the Lyapunov function.

These assumptions simplify the computations since a stochastic forecast of the state variables is not needed. The validity of this approach will be addressed in the technology demonstration phase of this effort (see Chapter 10). Further, the thrust of this research effort is to develop a simplified on-line procedure, not to develop a better (i.e., less conservative) Lyapunov function. Therefore, a "classical" choice -- the transient energy function -- will be used. Transfer conductances will not be considered since methods to include them are still a current research topic. A measure of transient security $S_T(t_0, t_1)$ is given by:

$$S_T(t_0, t_1) = \Pr\{\Delta V > 0, t_0 \leq t \leq t_1\} \quad (8.2-1)$$

where

$$\Delta V = \bar{V}^{PF} - V^{PF} \quad (8.2-2)$$

which must be greater than zero for first-swing stability. The superscript, PF, is used to denote post-fault values and is used to emphasize that only post-fault margins are important. The pre-fault stability margin is meaningless except for those few cases where the post-fault network configuration is identical to the pre-fault network (i.e., the fault is cleared without line switching).

The Lyapunov function used herein is given by:

$$V(\Omega, \theta) = \Delta V_K(\Omega) + \sum_{i=1}^n \left\{ P_{i1} \theta_{i1}^e - P_{i1}^{PF} \theta_{i1} \right\} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\{ K_{ij} \cos \theta_{ij}^e - K_{ij}^{PF} \cos \theta_{ij} \right\} \quad (8.2-3)$$

where

$$\Delta V_K = \frac{1}{2} \sum_{i=1}^n M_{i1} \Omega_{i1}^2 \quad (8.2-4)$$

$$\Omega_{i1} = \dot{\delta}_{i1} - \bar{\omega} \quad (8.2-5)$$

$$\theta_{i1} = \delta_{i1} - \bar{\delta} \quad (8.2-6)$$

$$\theta_{ij} = \theta_{i1} - \theta_{j1} \quad (8.2-7)$$

$$P_{i1} = P_{mi} - E_{i1}^2 G_{i1} \quad (8.2-8)$$

$$K_{ij} = E_{i1} E_{j1} B_{ij} \quad (8.2-9)$$

$$\bar{\omega} = \frac{\sum_{i=1}^n M_{i1} \dot{\delta}_{i1}}{\sum_{i=1}^n M_{i1}} \quad (8.2-10)$$

and, for the i^{th} generator,

δ_{i1} = generator rotor angle

P_{mi} = mechanical power input

E_{i1} = voltage behind direct axis transient reactance

B_{ij} = transfer susceptance (imaginary term of the reduced admittance matrix)

M_{i1} = moment of inertia

ΔV_K = kinetic energy increase during the fault-on period

Ω_{i1} = generator speed with respect to center-of-angle

n = number of generators in the network.

The superscript, e , denotes equilibrium values. The critical value, $\bar{V}$, is given by.

$$\bar{V} = V(0, \theta^{UPF}) \quad (8.2-11)$$

where the superscript, UPF, denotes the unstable saddle point corresponding to the pattern of generators which lose synchronism (post-fault). All rotor angles are measured with respect to center-of-angle (COA) coordinates. This coordinate system allows the motion of the individual generators to be differentiated from motion of the system as a whole. If all generator rotors accelerate together the system will be in synchronous equilibrium but not in frequency equilibrium. Only relative motion will contribute to a loss of synchronism.

Let the discrete random variable m denote a specific network fault with specific post-fault network topology. The system will be first-swing-stable with respect to this fault if

$$\Delta V^m = \bar{V}^m - V^m > 0 \quad (8.2-12)$$

Straightforward substitution of Eq. 8.2-3 into Eq. 8.2-12 yields

$$\Delta V^m = \sum_{i=1}^n P_{i1}^{PF} \left\{ \theta_{i1} - \theta_{i1}^{UPF} \right\} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\{ K_{ij}^{PF} (\cos \theta_{ij} - \cos \theta_{ij}^{UPF}) \right\} - \Delta V_K^m \quad (8.2-13)$$

Define V_R^m to be the first two terms on the right-hand-side of Eq. 8.2-13, viz.,

$$V_R^m = \sum_{i=1}^n P_{i1}^{PF} \left\{ \theta_{i1} - \theta_{i1}^{UPF} \right\} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\{ K_{ij}^{PF} (\cos \theta_{ij} - \cos \theta_{ij}^{UPF}) \right\} \quad (8.2-14)$$

The condition for stability with respect to the m^{th} fault is thus

$$\Delta V_K^m < V_R^m \quad (8.2-15)$$

- Equation 8.2-15 establishes a limit on the allowable increase in kinetic energy during the fault-on period. The quantity V_R^m therefore represents a threshold on
- the amount of energy that can be absorbed by the post-fault network without having the generators lose synchronism on the first-swing.

With the assumptions stated earlier, only two terms will be considered random variables. These are ΔV_K^m , which represents the kinetic energy increase, and the distribution of the network faults indexed by the discrete random variable m . One approach for simplifying the computations is to consider only a restricted set of faults. The resulting table of faults may be established during system planning efforts where a significant number of simulations are performed to establish key network parameters, critical clearing times, etc. It is expected that this table will consist of those faults which are most significant from a transient stability point of view. In this research only 3-phase zero-impedance bus faults which are cleared both with and without line switching are considered. These faults are assumed to be Poisson distributed with rate, λ_m , and having a probability of occurrence, P_m , where

$$P_m = \frac{\lambda_m}{N \sum_{i=1}^N \lambda_i} \quad (8.2-16)$$

where N = number of faults in the table.

The fault table must contain, at a minimum, the following data:

- Description of the fault (type and location in the network).
- The method by which the fault is to be cleared (e.g., opening the line without reclosure). This information is needed to establish the post-fault topology.
- Clearing time statistics.
- The fault occurrence rate, λ_m .

The network faults are assumed to be distributed according to a Poisson process. If the m^{th} failure, as enumerated in the set of faults, has not occurred prior to t_0 , the probability that it will occur in the short time interval $[t_0, t_1]$ is

$$\lambda_m e^{-\lambda_m(t_1-t_0)} \quad (8.2-17)$$

With this assumption the transient security measure defined in Eq. 8.2-1 becomes

$$S_T(t_0, t_1) = \sum_{m=1}^N P_m \lambda_m e^{-\lambda_m(t_1-t_0)} (t_1-t_0) \Pr\{\Delta V_K^m < V_R^m\} \quad (8.2-18)$$

To alleviate the real-time computational burden, evaluation of this transient security measure is divided into two parts. The first part (Phase I) operates in a non-real-time mode, periodically updating the large collection of parameters for each fault configuration necessary for the real-time computations. The second part (Phase II) provides a real-time evaluation of transient security utilizing inputs from the state estimator as well as parameters from the Phase I computations.

8.3 PHASE I COMPUTATIONS

The Phase I computations are sufficiently complex to preclude their evaluation in "real time". These computations are needed to determine the post-fault characterization for each fault listed in the fault table and to estimate the increase in kinetic energy obtained in the fault-on period. These computations consist of a power-flow problem for each post-fault network configuration. The boundary of separation which defines the region of stability must also be estimated. This is equivalent to determining those generators most likely to lose synchronism during the fault. The kinetic energy obtained during the fault is a function of the clearing time which is assumed to be a random variable with known statistics. The computations are described below:

Step I-1

For each post-fault network topology, a power flow problem is solved to obtain θ_1^{ePF} which are the equilibrium post-fault rotor angles with respect to center-of-angle coordinates. The power flow problem requires P_1 and K_{1j} which are assumed to be available from the state estimator (or by suitable computation).

Step I-2

For each fault, the initial acceleration of the rotor angles is computed from the swing equation. The machine (or machines) with the largest initial acceleration with respect to center-of-angle coordinates defines (approximately) the boundary of separation. Identification of the exact boundary of separation is a very complex problem and a subject of current research. The justification for using acceleration to estimate the boundary of separation is based on recent research by Ribbens-Pavella (Ref. 69) and Krishnaparandhama (Ref. 70). The latter research describes a method for identifying coherent generators based on the property of equal acceleration for coherent machines. This approach is attractive for use with on-line security measures since only the admittance matrix of the system and moment of inertia of the generators are needed.

The rotor angles of the unstable equilibrium are approximated by

$$\theta^{\text{UPF}} = \begin{cases} \pi - \theta^{\text{ePF}} + 2\theta_r^{\text{ePF}} & , \theta^{\text{ePF}} \geq 0 \\ -(\pi + \theta^{\text{ePF}} - \theta_r^{\text{ePF}}) & , \theta^{\text{ePF}} < 0 \end{cases} \quad (8.3-1)$$

for the generator(s) with maximum initial acceleration. For all others

$$\theta^{\text{UPF}} = \theta^{\text{ePF}} \quad (8.3-2)$$

The subscript, r , in Eq. 8.3-1 denotes the machine with largest inertia (reference machine). This approximation is based on the work of Prabhakara and El-Abiad (Ref. 71) transformed to center-of-angle coordinates, and is justified on the basis that the Lyapunov function is fairly constant in the vicinity of an unstable equilibrium point.

Step I-3

For each fault it is necessary to compute the kinetic energy increase during the fault-on period. A conservative assumption is that the machine accelerations are constant during the fault-on period at their initial values. Thus, the kinetic energy increase for the m^{th} fault is given by:

$$\Delta V_K^m = \frac{1}{2} t_{cl}^2 \sum_{i=1}^n M_i \alpha_i^2 \quad (8.3-3)$$

where α_i is the initial machine acceleration and t_{cl} is the clearing time. The probability density function for ΔV_K^m can be computed assuming M_i and α_i are deterministic and that the statistics of t_{cl} are known.

8.4 PHASE II COMPUTATIONS

The Phase II computations utilize the output of the state estimator and the result of the Phase I computations to generate the transient security measure in real-time. The Phase I computations provide θ^{UPF} and ΔV_K^m for each fault in the fault table. The state estimator provides current estimates of θ_i , P_i and K_{ij} .

Step II-1

For each fault under consideration, the deterministic threshold V_R^m can be computed from Eq. 8.2-14 since all quantities are either known as outputs from the state estimator or are computed as part of the Phase I computations. The probability that $\Delta V_K^m < V_R^m$ can be established for each fault since the distribution for ΔV_K^m was evaluated in Step I-3. This probability is denoted by

$$\phi_T^m = \Pr \{ \Delta V_K^m < V_R^m \} \quad (8.4-1)$$

Step II-2

The measure of security $S_T(t_0, t_1)$ is now evaluated by summing over all faults included in the fault table, thus

$$S_T(t_0, t_1) = \sum_{m=1}^N P_m \lambda_m e^{-\lambda_m (t_1 - t_0)} (t_1 - t_0) \phi_T^m \quad (8.4-2)$$

The transient security measure defined in Eq. 8.4-2 is significantly easier to compute than the

measure defined in Eq. 8.1-11. The computational savings result from the following assumptions:

- A list of enumerated faults and the clearing time statistics are the only random variables
- The system is in quasi-steady-state prior to the fault.
- Rotor acceleration is sufficient to establish the coherent generators defining the boundary of separation
- Unstable post-fault equilibrium angles are approximated from their pre-fault equilibrium values
- Known clearing time statistics and constant rotor acceleration are sufficient to estimate kinetic energy increase during the fault-on period
- The energy threshold is computed on the basis of the post-fault topology associated with each fault.

8.5 SUMMARY

A measure of power system transient stability suitable for on-line evaluation was developed in this chapter. In order to reduce the computational burden to a reasonable level, certain simplifying assumptions were made. A small-scale demonstration of this security measure was performed and is documented in Chapter 10. The results of this demonstration afford an assessment of the validity of the simplifying assumptions made in the course of developing this security measure.

The number of possible jump disturbances in a power system which could reduce security levels is substantial. The number of system configurations arising from line and/or generator outages for a reasonably-sized power system is intractable in terms of algorithmic implementation in real time. For example, the New England power system model with 45 major transmission lines would have a total of $\sim 10^{13}$ possible configurations due to line outages. Obviously, methods are needed to reduce the number of jump disturbances that must be considered to a manageable yet meaningful number. Some techniques for accomplishing this are discussed in this chapter.

9.1 JUMP DISTURBANCE MODELS

As discussed in Section A.3, the discrete nature of a line or generator outage is conveniently modeled as a network structural change. The network topology (i.e., structure) may be characterized by a vector $\underline{z}$ of binary random variables with elements

$$z_1(t) = \frac{1}{2}(-1)^{N_1(t)} + \frac{1}{2} \quad (9.1-1)$$

where $N_1(t)$ is a Poisson counting process. This model is consistent with an analysis of available data (reported in Ref. 72) which showed that the time between outages for electrical transmission equipment is approximately exponentially distributed, indicating that the failure occurrences follow a Poisson model.

For a linear system model with parameters that are functions of a finite-state continuous-time Markov process (e.g., $\underline{z}(t)$), the results of Brockett and Blankenship (Ref. 73) can be applied to transform the system in a manner such that the matrix parameters are independent of the finite-state random variables. However, the number of state variables for the transformed system must be increased by a factor equal to the number of finite states of the Markov process. In the line outage context this number could be as high as $n \times 2^q$, where n is the dimension of the original state vector and q is the number of possible network configurations. For the New England power system model, n may be 100 or more so that the number of state variables would have to be increased by $\sim 10^{15}$ -- hardly a manageable number! Clearly a reduction in the number of state variables is needed.

One method for relaxing the dimension problem would be to reduce the size of the network being modeled to a small number of generators, transmission lines, and load buses. An approach to this type of order reduction via a spatial decomposition technique is described in Section A.6. Another approach to reducing system dimension would be to substantially reduce the number of random coefficients in the system equations.

If a few random coefficients in the system model can effectively represent the system response that a model with the larger number of random coefficients would give (or better yet the actual system response), then the reduced order model would be much easier to analyze while remaining an adequate system representation. But the question is: What does representing the system accurately mean? What are the criteria

upon which to judge the model's closeness to reality? Based on these criteria, what method should be chosen to reduce the number of random coefficients? These questions are addressed in detail in the following sections.

9.2 BASIC CLUSTERING TECHNIQUES AND APPROACH

The goal of cluster analysis is to separate a set of data into groups of clusters that are in some way related. An application from the area of power system stability would be to group those jump disturbances (e.g., line faults) which yield the same splitting of the network into islands. Cluster analysis techniques may roughly be assigned to one of the following types:

- Hierarchical techniques - in which the classes themselves are classified into groups, the process being repeated at different levels to form a tree
- Optimization techniques - in which the clusters are formed by the optimization of a clustering criterion. The classes are mutually exclusive, thus forming a partition of the set of entities
- Density or mode-seeking techniques - in which clusters are formed by searching for regions containing a relatively dense concentration of entities
- Clumping techniques - in which the classes or clumps can overlap
- Others - methods which do not fall clearly into any of the four previous groups.

In a hierarchical classification technique, the data are not partitioned into classes in a single step, but rather, in a sequence of steps. Hierarchical techniques fall into two major groups: agglomerative methods and divisive methods. In the former method, the individual entities are fused into successively larger groups, while the latter method proceeds by generating finer and finer partitions of the data. Unless stopped at some intermediate stage, a hierarchical technique will either split the entire data set into groups each containing a single entity or merge all the data into one large group.

Optimization techniques differ from the preceding technique in that the solutions do not necessarily exhibit an hierarchical relationship among the entries. Optimization techniques seek a partition of the data which optimizes a predefined numerical measure indicating the relative "closeness" of a pair of data points. Most, although not all, methods require the a priori selection of the number of groups into which the data is to be separated.

If the data are depicted as points in a metric space, a natural concept of clustering suggests that there should be parts of the space in which the points are very dense, separated by parts of low density. This concept forms the basis of density search techniques which seek regions of high density or modes in the data. In general, each mode is taken to signify a different group.

Most classification techniques lead to distinct or disjoint clusters, in other cases it is useful to permit an overlap between the classes. In general, classification techniques which allow overlapping

clusters are known as clumping techniques. Clumping techniques begin with the computation of a similarity index from the original data to give an estimate of the similarity between each pair of entities on the basis of the properties they exhibit. These methods seek a partition of the entities into two groups, the smaller of which is generally considered to be the class sought.

The use of cluster analysis in practice is not the simple application of one particular technique to the data being studied, but instead involves a series of steps each of which may be dependent on the results of the preceding one. It is generally impossible a priori to anticipate what combinations of variables, similarity measures and clustering techniques are likely to lead to interesting and informative classifications. Consequently the analysis may proceed through several stages including the alteration of variables, choosing a different similarity measure, concentration of a particular subset of entities, etc.

Many clustering techniques are sensitive to the presence of outliers and so before performing a cluster analysis some attempt should be made to detect such observations. (See Ref. 74 for a discussion of outlier identification techniques). Although outliers are probably best removed before application of any clustering technique, they should not necessarily be regarded as mere errors of observation. They should be examined carefully in a bid to find a rational explanation of how they might have arisen. For example, one possible explanation could be that they are members of a cluster poorly represented in the available data set. After any outliers have been removed, it is generally helpful to use some ordination technique to obtain a two-dimensional mapping of the data. Such a mapping is an invaluable aid in the interpretation of clusters produced by some method at a later stage of the analysis.

The preceding paragraphs are only a brief overview of clustering techniques. A more detailed discussion, including algorithms, may be found in Ref. 75.

In order to reduce the number of jumps variables to a manageable, yet meaningful number by some method of clustering, criteria must be established by which to judge equivalence of jump disturbance effects. Obviously, the clustering criteria will vary according to which measure of power system security is of interest. The clustering criteria discussed herein correspond to three stability categories: transient, dynamic, and steady-state. A brief description of each of these stability category follows.

Transient instability results in loss of synchronism among generators. Loss of synchronism usually occurs within a very short time (less than a second) after a large disturbance. Transient stability is discussed in detail in Chapter 8.

Dynamic instability refers to the lack of damping in a power system leading to growth of subsynchronous oscillations or power surges. Meisel (Ref. 76) gives many historical accounts of negative damping in power systems that have led to power oscillations which ultimately reach equipment limits. These resonant-like oscillations are caused by interconnections of generator-regulator systems and are set off by relatively small perturbations (Ref. 76).

Steady-state instability is characterized by variables such as voltages, currents, and frequencies reaching values beyond equipment limits causing load shedding, generator outages, etc. The time frame

of this form of instability is on the order of several minutes following disturbances.

The general approach to clustering for all three stability categories is

- Choose the appropriate deterministic model (see Appendices A and B)
- Compute an appropriate measure of equivalence for given jump disturbances and given initial conditions
- Treat each measure (computed above) as a sample of a random variable
- Cluster the samples according to one of the techniques described above
- Work with a reduced number of jumps in estimation and forecasting of security measures.

Specific information on the clustering criteria for these three stability categories are discussed in the following sections.

9.3 TRANSIENT STABILITY CLUSTERING

The security measure proposed for transient stability evaluation is discussed in Section 8.2. In the analysis of that section, it was assumed that there existed a set of m network faults to be evaluated. For computational tractability, this set must be much smaller than the total number of possible faults to which the system is susceptible. Cluster analysis can be used to identify those network faults that are, in some sense, equivalent. In this context, a measure of "closeness" between faults can be defined in terms of the generator rotor angle response. Two different faults are equivalent if the individual generator rotor oscillations arising from the faults exhibit a similar (i.e., coherent) time response relative to center-of-angle (COA) coordinates. An approach for clustering faults on the basis of rotor angle coherency is discussed in this section.

Several approaches to coherency analysis have been used in the past. For example, Lee and Schweppe (Ref. 77) used an electrical distance measure. Podmore (Ref. 78) developed a method utilizing linear system models, closeness of swing curves as a criterion, and clustering in the classical sense. Schlueter, et al. (Refs. 79 and 80) have utilized an rms coherency measure for probabilistic disturbance analysis.

The coherency measure proposed here is essentially a measure of covariance among faults for each of the generator angles. Let $C_{ij}^k(t)$ be the covariance matrix associated with the k^{th} angle with i, j^{th} element

$$C_{ij}^k(t) = E \left\{ \int_0^t \delta_{ik}(\tau) \delta_{jk}(\tau) d\tau \right\} \quad (9.3-1)$$

where δ_{ik}^k is the response of the k^{th} rotor angle (relative to COA) following the occurrence of the i^{th} fault at $t = 0$. The motivation for choosing a covariance measure is that well-established multivariate statistical methods (i.e., principal components or factor analysis, Ref. 81) are available to divide the

faults into groups based on the covariances or inter-correlations among generator rotor angle deviations.

One approach for estimating $C^k(t)$ that appears to be promising requires direct simulation of the generator swing equations (as embodied in the EPRI midterm simulation, Ref.36). For each generator (k) and pair of faults (i and j), estimate the appropriate element of the covariance matrix based on the sampled values of the δ_{ik} via

$$\hat{C}_{ij}^k(t_k) = \sum_{m=1}^L \delta_{ik}(t_m) \delta_{jk}(t_m) \quad (9.3-2)$$

For each generator, use a principal components or factor analysis technique to project the covariance vectors onto two or more orthogonal factors which account for the highest proportion of variance. The projections of the covariance or correlation vectors on the factors, called factor loadings, can be used as a basis for deciding which faults may be grouped together. The above procedure is then repeated for each generator. Those disturbances which fall into the same group for a number of generators are then considered to belong to a cluster.

Once the faults are grouped according to their effect on system response*, it would be valuable to assess whether the various faults cluster in the same way with respect to the parameter V_R (which represents the amount of energy that can be absorbed in the post-fault network). Faults which lie in the same clusters based on both $C^k(t)$ and V_R are readily grouped. However, in those cases for which the fault is grouped differently using the two criteria, it would be necessary for an analyst to judge into which cluster the fault should be placed, or alternately, whether it should be placed into a unique group.

9.4 DYNAMIC STABILITY CLUSTERING

Dynamic instability results from positively damped oscillations of a power system at one or more characteristic frequencies. This form of instability is analogous to structural vibrations at resonant frequencies. Dynamic instability appears to arise when locally stable generators and their controls are coupled in a network which gives rise to resonant underdamped oscillations at a frequency typically on the order of several Hertz (Ref. 76). Fairly minor perturbations can set off these oscillations so that a linearized system model seems appropriate for analyzing dynamic stability (Ref. 76). One basic test of dynamic stability then involves examining a linearized dynamic model for positive eigenvalues.

On approach to clustering for dynamic stability analysis would be to form a linearized model (for each jump disturbance and for each set of initial conditions about which the system is linearized) and then to compute the eigenvalues (cf. Refs. 82 and 83). Disturbances (which would conceptually include initial conditions) could be classified on the basis of eigenvalue sensitivity. This technique is outlined in the remainder of this section.

The eigenvalue sensitivity approach to jump clustering involves the computation of the partial derivative of each eigenvalue of a linearized dynamic system

*This is in essence a grouping based on the kinetic energy increase ΔV_k of Section 8.2.

model with respect to the parameters (and initial states or conditions) subject to jump disturbances. Those jumps giving rise to similar eigenvalue changes (in magnitude and direction) would be clustered and represented by a single jump variable in the estimation of security assessment. This approach can be easily linked to security enhancement, since information on which parameters or control variables will make the eigenvalues more negative is generated in the process of clustering analysis and security assessment.

Since the proposed dynamic stability clustering criterion involves eigenvalues, an (unforced) system model of the form

$$\dot{\underline{x}} = F_o(\xi)\underline{x} \quad (9.4-1)$$

will be assumed, where the linearized dynamics matrix F_o is a function of the parameter ξ which is influenced by the jump disturbances. The sensitivity of the eigenvalues of F_o to the parameter ξ is given by

$$\frac{\partial \lambda_i}{\partial \xi} = \frac{(\underline{v}_i^o)^T \frac{\partial F_o}{\partial \xi} \underline{u}_i^o}{(\underline{v}_i^o)^T \underline{u}_i^o} \quad i=1,2,\dots,n \quad (9.4-2)$$

where

λ_i is an eigenvalue of F_o
 $\underline{u}_i^o, \underline{v}_i^o$ are the eigenvectors of F_o and F_o^T respectively, corresponding to λ_i
 n is the dimension of F_o

The jumps are then clustered according to the closeness in magnitude and sign (positive or negative) of the eigenvalue sensitivities. Classical clustering analysis methods (Ref. 75) or factor analysis (Ref. 81) can be used.

If the eigenvalues are sensitive to the initial states (or nominal operating point of the power system around which the model is linearized), another set of likely initial states should be chosen. The procedure should then be repeated for the new model. If the jumps cluster into similar groups for the key operating points of the power system, the same reduced number of jump variables could be used for assessing dynamic stability.

9.5 STEADY-STATE STABILITY CLUSTERING

Steady-state instability is assumed, for the purposes of this analysis, to refer to the condition of a power system wherein variables have reached or surpassed equipment limits and as a result, significant loads are lost. The basic hypothesis test for security assessment would be:

$$H_0: \underline{x} \geq \underline{x}_m$$

$$H_1: \underline{x} < \underline{x}_m$$

where

$\underline{x}$ is the system state vector (see Appendices A and B)

and

$\underline{x}_m$ represents the equipment limit values corresponding to $\underline{x}$

Two approaches could be used to cluster jumps for steady-state stability analysis. The first is a straight-forward method wherein the steady-state values of the power system nonlinear model are computed for each jump disturbance. In the second method, a linearized model is used and clustering is done according to an eigenvalue criterion. The two approaches are described briefly below.

The first approach uses a more accurate nonlinear model (e.g., the long-term model of Appendix A) but requires more computation. Steady-state values are computed for each model corresponding to a jump disturbance. The jumps are clustering according to the closeness of steady-state values in state space.

The second approach involves a system model linearized about key operating points as in the previous dynamic stability algorithm. The steady-state values can be related to eigenvalues in the following way. Let

$$\dot{\underline{x}} = \underline{F}\underline{x} + \underline{L}\underline{u} \quad (9.5-1)$$

where

$\underline{u}$ is the input vector

$\underline{L}$ is the input distribution matrix with constant elements which may be functions of operating points (initial conditions).

Using the coordinate transformation

$$\underline{y} = \underline{T}\underline{x} \quad (9.5-2)$$

where $\underline{T}$ is a matrix consisting of the eigenvectors of $\underline{F}$ gives the alternate representation

$$\dot{\underline{y}} = \underline{\Lambda}\underline{y} + \underline{T}\underline{L}\underline{u} \quad (9.5-3)$$

with

$$\underline{\Lambda} = \underline{T}\underline{F}\underline{T}^{-1} \quad (9.5-4)$$

a diagonal matrix whose entries are the eigenvalues of $\underline{F}$. The steady-state solution to Eq. 9.5-3 is

$$\underline{y}_{ss} = -\underline{\Lambda}^{-1}\underline{T}\underline{L}\underline{u} \quad (9.5-5)$$

From which

$$\underline{x}_{ss} = -\underline{T}^{-1}\underline{\Lambda}^{-1}\underline{T}\underline{L}\underline{u} \quad (9.5-6)$$

Thus, the steady-state solution can be written in terms of eigenvectors and eigenvalues which have already been computed for use in the dynamic stability clustering algorithm. In fact, it would not appear unreasonable to use the eigenvalues alone as the clustering criteria so that the clustering results from the dynamic stability analysis could be used directly for the steady-state analysis.

10. DEMONSTRATION OF SIMPLIFIED TRANSIENT SECURITY MEASURE

10.1 INTRODUCTION

A model of the 9-bus WSCC network described below was used to evaluate the transient security measure developed in Section 8.2 and to test its various assumptions. The model was simulated using the EPRI Transient-Midterm Simulation Program (Ref. 36). For this study, each of the three generators was represented by the classical model consisting of a constant voltage source behind the direct axis transient reactance.

The transient security measure defined by Eq. 8.2-18 requires a stochastic representation of the distribution of the network faults and associated clearing time statistics. Rather than select representative statistical models and perform a Monte Carlo analysis, it is more instructive to use the simplified procedure to predict the critical clearing times for a known set of network fault/post-fault topology combinations and compare these with the exact values obtained with the EPRI simulation. Five such combinations were selected for analysis and the results are discussed below. The following sections document the procedure by which the results were obtained.

The WSCC 9-bus test system is depicted schematically in Fig. 10.1-1. The buses associated with the three generators are labeled GEN BUS-1, -2 and -3. These buses are connected to buses labeled STA-1, -2, and -3 via step-up transformers. Transmission lines connect these distribution buses with three load centers shown as buses labeled STA-A, -B and -C.

Figure 10.1-1 WSCC 9-Bus Test System

10.2 REDUCED-ORDER NETWORK

When using the classical generator model, it is standard procedure to reduce the network to generator internal nodes. In order to accomplish this, the network nodes, x , are partitioned into three groups:

$$x = \begin{bmatrix} x_a \\ x_b \\ x_c \end{bmatrix} \quad (10.2-1)$$

where

x_a = generator internal nodes (dimension n)

x_b = generator terminal nodes (dimension n)

x_c = remaining network nodes (dimension m)

n = number of generators

Using this formulation, Ohm's law can be written in partitioned form as

$$\begin{bmatrix} I_a \\ I_b \\ I_c \end{bmatrix} = \begin{bmatrix} Y_a & Y_b \\ Y_c & Y_d \end{bmatrix} \begin{bmatrix} V_a \\ V_b \\ V_c \end{bmatrix} \quad (10.2-2)$$

with

$$Y_d = \begin{bmatrix} Y_{d1} & Y_{d2} \\ Y_{d3} & Y_{d4} \end{bmatrix} \quad (10.2-3)$$

where

I = complex currents injected into each node

V = complex node voltages

Y = complex admittances

The admittance matrices Y_a , Y_b , Y_c contain the generator direct axis transient reactances. For the WSCC network this also includes the reactance of the transformers connected to each generator terminal bus. Matrix Y_d contains the admittances of the transmission network. The generator plus transformer reactances are included in the diagonal elements of Y_{d1} . The

loads at the three load buses are modeled as constant impedances with the appropriate values added to the diagonal elements of Y_{d4} . For the WSCC network pre-

fault configuration, the Y matrix is shown in Fig. 10.2-1. With reference to Fig. 10.1-1 the ordering of the nodes is

$$x = \begin{bmatrix} \text{GEN 1 (Internal)} \\ \text{GEN 2 (Internal)} \\ \text{GEN 3 (Internal)} \\ \text{STA 1} \\ \text{STA 2} \\ \text{STA 3} \\ \text{STA A} \\ \text{STA B} \\ \text{STA C} \end{bmatrix}$$

A similar Y matrix is generated for each post-fault network configuration.

Since current injection occurs only at the generator internal nodes

$$I_b = I_c = 0 \quad (10.2-4)$$

TABLE 10.2-1

LINE REMOVED	Y_r (REAL)			Y_r (IMAGINARY)		
Pre-Fault Configuration	1 073 0 266 0 196	0 266 0 374 0 191	0 196 0 191 0 266	-3 299 1 662 1 395	1 662 -2 759 0 993	1 395 0 993 -2 432
STA 1 - STA A	0.676 0 135 0 095	0 135 0 680 0 293	0 095 0 293 0 293	-2 024 0 735 1 205	0 735 -2 252 1.070	1 205 -2 476 0 627
STA 1 - STA B	0 744 0 188 0 142	0 188 0 614 0 289	0 142 0 289 0 467	-2 289 1 413 0 627	1 413 -2 729 1 121	0 627 1 121 -1 971
STA 2 - STA A	1 355 0 109 0 152	0 109 0 362 0 188	0 152 0 188 0 264	-2 443 0 757 1 175	0 757 -2 010 1 160	1 175 0 160 -2 397
STA 2 - STA C	1 068 0 176 0 262	0 176 0 301 0 045	0 262 0 045 0 545	-3 322 1 638 1 370	1 638 -1 934 0 254	1 370 0 254 -1 867
STA 3 - STA B	1 241 0 229 0 110	0 229 0 368 0 179	0 110 0 179 0 239	-2 519 1 391 0 636	1 391 -2 706 1 169	0 636 1 169 -1 883
STA 3 - STA C	1 048 0 317 0 110	0 317 0 621 0 041	0 110 0 041 0 270	-3 319 1 549 1 459	1 549 -2 130 0 257	1 459 0 257 -1 691

Figure 10.2-1 Y-Matrix for Pre-fault WSCC
Network Before Reduction.
(All admittances expressed
per unit.)

Thus, Eq. 10.2-2 can be written as

$$I_a = Y_a V_a + Y_b \begin{bmatrix} V_b \\ V_c \end{bmatrix} \quad (10.2-5a)$$

$$0 = Y_c V_a + Y_d \begin{bmatrix} V_b \\ V_c \end{bmatrix} \quad (10.2-5b)$$

Equation 10.2-5 can be rearranged to yield

$$I_a = Y_r V_a \quad (10.2-6)$$

$$Y_r = Y_a - Y_b Y_d^{-1} Y_c \quad (10.2-7)$$

Admittance matrix Y_r is the reduced admittance matrix and has dimension equal to the number of generators. The elements of the reduced admittance matrix for the pre-fault network and six different post-fault configurations are given in Table 10.2-1.

Given the Y_r matrix for the reduced network, the remaining parameters needed for evaluation of the Lyapunov function, viz., P_i and K_{ij} , can be computed from a knowledge of the post-fault stable equilibrium angles. These parameters for each post-fault configuration are obtained using the power flow equation for the reduced network

$$P_{ei} = E_i^2 G_{ii} + \sum_{\substack{j=1 \\ j \neq i}}^n E_i E_j \left\{ B_{ij} \sin (\delta_i - \delta_j) + G_{ij} \cos (\delta_i - \delta_j) \right\} \quad (10.2-8)$$

where

P_{e_i} = real electrical power input to the node

E_1 = magnitude of the voltage-behind-transient reactance in the classical generator model

δ_i = voltage angle (generator rotor angle in classical model)

$$G_{1,j} = \operatorname{Re} \{Y_r\}$$

$$B_{1j} = \text{Im} \{Y_r\}$$

Under steady-state conditions,

$$P_{m1} = P_{ei} \quad (10.2-9)$$

where P_{mi} is the mechanical power input to each node (from the turbine). Assuming steady-state conditions and a specific post-fault network configuration, Eq. 10.2-8 can be solved for E_i . The parameters P_i and K_{ij} , are then computed as

$$P_i = P_{mi} - E_{i,ii}^2 \quad (10.2-10)$$

$$K_{ij} = E_i E_j B_{ij} \quad (10.2-11)$$

Table 10.2-2 lists the values of these parameters for the five cases examined. Also listed are the stable equilibrium angles measured with respect to center-of-angle.

10.3 PREDICTION OF THE CRITICAL CLEARING TIME

The condition for stability for a known fault was derived in Section 8.2 (cf. Eqs. 8.2-14 and 8.2-15). The condition is

$$\Delta V_K < V_P \quad (10.3-1)$$

where the stability threshold is

TABLE 10.2-2
NETWORK PARAMETERS FOR LYAPUNOV FUNCTION

CASE	P ₁	P ₂	P ₃	K ₁₂	K ₁₃	K ₂₃	θ ₁ ^e	θ ₂ ^e	θ ₃ ^e
1	-0.874	1.222	0.567	0.894	1.354	1.276	-10.17	29.95	14.09
2	-0.438	1.292	0.199	1.861	1.607	0.295	-5.66	24.17	-5.90
3	-0.438	1.215	0.577	1.816	1.464	1.058	-4.97	14.47	7.19
4	0.024	0.707	0.525	0.898	1.331	1.314	-2.05	5.84	3.19
5	-0.447	0.891	0.576	1.786	1.554	0.282	-4.68	11.31	10.96

$$V_R = \sum_{i=1}^n P_i^{PF} \left\{ \theta_i - \theta_i^{UPF} \right\} + \sum_{i=1}^{n-1} \sum_{j=i+1}^n \left\{ K_{ij}^{PF} (\cos \theta_{ij} - \cos \theta_{ij}^{UPF}) \right\} \quad (10.3-2)$$

The kinetic energy increase during the fault-on period, ΔV_K , is approximated as (cf. Eq. 8.3-3)

$$\Delta V_K = \frac{1}{2} t_{cl}^2 \sum_{i=1}^n M_i \alpha_i^2 \quad (10.3-3)$$

At the critical clearing time, t_{cr} , the kinetic energy increase is exactly equal to the threshold

$$\Delta V_K = V_R \text{ at } t_{cl} = t_{cr} \quad (10.3-4)$$

Equations 10.3-2 and 10.3-3 can be equated and solved for t_{cr} to yield

$$t_{cr} = \left(\frac{2V_R}{\sum_{i=1}^n M_i \alpha_i^2} \right)^{1/2} \quad (10.3-5)$$

In the evaluation of critical clearing time several assumptions are made. The machine accelerations, α_i , are assumed to remain constant at the values established at the beginning of the fault-on period. The post-fault unstable equilibrium angles, which appear in V_R , are computed on the basis of the approximation given by Eq. 8.3-1. In this approach, the separation boundary is determined by the machine(s) which has the largest magnitude acceleration with respect to the center-of-angle. Consequently, there is no need to integrate the equations of motion during the fault-on period. Finally, it is assumed that V_R is evaluated at the start of the fault-on period, $V_R(0)$. These assumptions are discussed further in the next section.

10.4 SIMULATION STUDIES

A total of five cases were selected for evaluation including at least one case representing a 3-phase fault at each generator terminal bus. A summary of these test case is given in Table 10.4-1. Also shown in the table is a comparison of the critical clearing times predicted by the simplified procedure of Eq. 10.3-5 with actual results obtained from detailed simulation of the unreduced network using the EPRI Transient-Midterm simulation program. Since the EPRI simulation provides an integration of the non-linear differential equations of motion, several runs of each case are needed to bracket the actual critical clearing times.

TABLE 10.4-1
SUMMARY OF TEST CASES

CASE	FAULTED BUS	FINAL TOPOLOGY	CLEARING TIMES FROM SIMULATION		CRITICAL CLEARING TIME (SIMPLIFIED PROCEDURE)
			STABLE	UNSTABLE	
1	STA 2	STA 2 - STA A Open	0.16	0.18	0.16
2	STA 2	STA 2 - STA C Open	0.18	0.19	0.18
3	STA 2	No Line Switching	0.23	0.24	0.27
4	STA 1	STA 1 - STA A Open	0.33	0.34	0.30
5	STA 3	STA 3 - STA C Open	0.26	0.27	0.30

As can be seen from Table 13.4-1, there is reasonable agreement between the actual and approximate clearing times; the largest deviation being approximately 15 percent. A detailed analysis of these cases along with recommendations for further research follow.

10.4.1 Approximations Based on Use of Machine Accelerations

The use of machine acceleration measured with respect to the center-of-angle is fundamental to the simplified approach. First, it is used to approximate the kinetic energy increase during the fault-on period and second, it is used to establish coherency groups of machines and therefore identify those machine(s) most likely to lose synchronism. Table 10.4-2 provides a summary of results based on the EPRI simulation. With the exception of Case 5, the largest error in approximating the kinetic energy increase using machine acceleration at the start of the fault-on period is 8 percent. This would translate into an error in calculation of the critical clearing time by about 4 percent as a consequence of the square-law relationship. The error in Case 5 can be traced to the fact that, for this case, the fault is applied to the generator with the smallest inertia in the system. While the corresponding acceleration of Generator 3 remain constant, the accelerations of the machines with higher inertia do not.

In all cases except one, the actual machine(s) which loses synchronism is the one with the largest initial acceleration. The lone exception is Case 1 where in addition to Generator 2 (with the largest acceleration), Generator 3 (with a much smaller initial acceleration) also loses synchronism as measured by angle difference with respect to center-of-angle. Generator 3 has the smallest inertia of the generators and its influence on the critical clearing time was found to be minimal for this case.

TABLE 10.4-2
SUMMARY OF MACHINE ACCELERATION
DURING FAULT-ON PERIOD

CASE	ACCELERATION (deg/sec ²) (START OF FAULT-ON PERIOD/ END OF FAULT-ON PERIOD)			ΔV_K (PER UNIT)		ACTUAL MACHINE WHICH LOSES SYNCHRONISM
	GEN 1	GEN 2	GEN 3	APPROX- IMATE	ACTUAL	
1	-640/-567	1994/1990	690/200	0 71	0 68	GEN 2 AND GEN 3
2	-640/-552	1994/1990	690/102	0 88	0 84	GEN 2
3	-640/507	1994/1986	690/-201	1 48	1 37	GEN 2
4	-390/-390	939/931	927/941	1 05	1 05	GEN 2 AND GEN 3
5	-541/-338	934/204	1955/1921	1 38	1 12	GEN 3

10.4 2 Plots of Selected Cases

The following series of plots, is representative of the cases studied. The results for Case 1, shown in Fig 10 4-1, represent a case where the calculated critical clearing time is quite close to that obtained by the detailed simulation. Results of Case 4 (Fig. 10 4-2) and Case 5 (Fig 10.4-3) represent extremes in terms of under and over-estimating the critical clearing times.

For each case, a 3-phase bus fault (short) is introduced at $t = 0.0833$ sec and subsequently cleared after the clearing time specified in Table 10.4-1. Each figure includes a pair of plots for both the stable and unstable clearing times which bracket the actual critical clearing time. For the cases shown, the faults are cleared by line switching. The top graph of each pair is a plot of the rotor angle of each generator in degrees from center-of-angle; typically referred to as a swing curve. The lower graph is a plot of the evolution of individual terms in the security measure for the critical "first-swing" of the transient response. The curves plotted are: ΔV_K , the actual increase in kinetic energy of the rotors from center-of-angle; V_R , the energy threshold from Eq. 10.3-2. Also shown is $V_R(0)$, which is the constant value of V_R used to calculate the critical clearing time in Eq. 10 3-5

Figure 10.4-1 shows the simulation results for Case 1. Inspection of Fig. 11.4-1 (b) reveals that the kinetic energy increase during the fault-on period almost equals the calculated threshold $V_R(0)$ while in Fig. 10.4-1 (d) the threshold is exceeded with a corresponding loss of stability. Consequently, the predicted critical clearing time agrees closely with the true value (Table 10.4-1). This figure also shows that V_R does not remain constant during the fault which is another contributor to error. The change in V_R is associated with the fact that, during the fault, the rotor angles change due to machine acceleration. This moves the operating point in angle-space to regions which may be more or less stable with respect to the post-fault topology. An undesirable integration of the differential equations would be needed to include this effect.

Figures 10.4-2 and 10.4-3 show the extremes of results obtained with classical models producing the largest errors in computed critical clearing time. Only the latter is undesirable as it would indicate to the operator more margin than the system really had.

10.5 SUMMARY

The objective of this research was to develop a measure of transient security which was sufficiently simple that it could be computed online for use by the system dispatcher. Since transient stability margins are a function of both specific fault as well as eventual post-fault topology, which are not known *a priori*, simplifying assumptions were made to be able to encompass all possible combinations with a minimum of computation. More accurate approaches are known and are used for off-line planning studies. Nevertheless, the simplified security measure developed herein produced results within 15 percent of the true values when used to predict critical clearing times. This accuracy may prove adequate for real-time operations when qualitative information, such as trend analysis, can alert the system operator of the need to take action. Areas in which additional research is needed to improve the transient security measure are discussed in the following paragraphs.

It is apparent from inspection of elements of the reduced bus network in Table 10 2-1, that for the WSCC system the magnitude of the transfer conductance terms (off-diagonal terms in real Y) is significant. A Lyapunov function such as suggested in Ref 08 may yield more accurate results. The best way to include these effects (which precludes the existence of a true Lyapunov function) is still an open question.

Alternative approximations to the unstable equilibrium angles should be explored. Equation 8 3-1 could still be used as an initial approximation if a more accurate solution is to be found. However any technique which requires integration of the faulted network equations will result in a significant increase in computational complexity when summed over all possible network faults and post-fault topology combinations.

Consideration should be given to more complex (non-classical) generator models. Further, Lyapunov functions which include the characteristics of exciters and turbines are needed since the application of dynamic breaking and fast valving can have a significant impact on transient stability.

Finally, simulation studies should be conducted for significantly larger networks in terms of the number of generators and buses. In particular, a Monte Carlo study to evaluate the complete security measure of Eq. 8.4-2 should be undertaken to assess its suitability for on-line monitoring and control of a system subject to random network faults. The results obtained thus far in predicting critical clearing times with a simple computational procedure provide encouragement in this regard.

Figure 10.4-1 Simulation Results for Case 1

Figure 10.4-2 Simulation Results for Case 4

Figure 10.4-3 Simulation Results for Case 5

11. SUMMARY AND RECOMMENDATIONS

11.1 SUMMARY

This report documents the results obtained during a three-year research effort into the development of a dynamic security assessment methodology for power systems. Progress was made in the areas of:

- Power system modeling
- State estimation
- Structure decomposition
- State forecasting
- Clustering
- Security measure development.

Each of these topics is discussed briefly in the following paragraphs.

A detailed dynamic model of a multimachine power system was developed for use throughout this effort. The model includes a state variable description of the boiler-turbine, generator and excitation system. Two sets of equations were formulated -- one describing the fast response (on the order of seconds) associated with transient dynamics and the other modeling the slower long-term dynamics (on the order of minutes). The interconnection network is characterized by a system of simultaneous algebraic equations. Stochastic components of the model consist of unpredictable load disturbances and transmission line and/or generator outages. An integrated state space formulation of the power system model was also developed. This model combines the component dynamic models, the algebraic network equations and the load disturbances into a unified state space representation suitable for state estimation. A measurement model characterizing power system observations was also completed.

A state estimator for estimating the long-term dynamic behavior of the power system was specified. Since measurements of both the system state and network configuration are available, the long-term state estimator was partitioned into two separate (but not independent) algorithms -- the process state estimator and the structure (or network configuration) estimator. The process state estimation effort led to the development of an estimation algorithm identical in form to the extended Kalman filter, but containing a modified process noise driving term to account for the Poisson component of the process noise. The structure estimation algorithm is based on a two-stage approach: first, a maximum a posteriori estimate of the structure (as characterized by the topological state z) is obtained from available breaker status measurements. Then, a hypothesis testing approach is used to select the true network structure from the small number of candidate configurations which result from the first stage.

Various approaches to estimator simplification were examined including nonlinear observers and stochastic approximation based techniques. The most promising techniques involve algorithms of the same form as the extended Kalman filter, but with gain sequences chosen in a simplified manner.

In the area of structure decomposition, two different approaches were investigated -- spatial decomposition and time-scale decomposition. Spatial network decomposition consists of partitioning the power

system into a number of subsystems, each with a small number of generators and loads. A computer program was developed to implement an optimal network decomposition technique. This technique minimizes the interaction between the subnetworks. The program was used to decompose the IEEE 39-Bus New England Test System into six subsystems, each containing one or two generators. Time-scale decomposition, or the separation of the system into two subsystems, one containing states with slow dynamics and the other containing states with fast dynamics, was applied to the extended Kalman filter equations of the process state estimator. This approach to system order reduction was based on an application of singular perturbation theory. This methodology provides a tradeoff between accuracy and computational savings.

State forecasting provides information on the anticipated evolution of the power system state. Research in this area has resulted in the development of recursive expressions for the first and second moments of the system state under the assumption that there are a finite number of possible network configurations and that the transition from one configuration to another is governed by a homogeneous Poisson process.

Clustering techniques for reducing the large number of possible power system disturbances to a more manageable number were reviewed and a set of clustering criteria for the security measures of interest were proposed. These criteria were chosen so as to reflect a natural measure of closeness for each security measure.

Research into security measures has proceeded in three directions. The first area involves viewing the security assessment problem as a surface crossing problem for a stochastic process. The transition of the power system from the normal operating state to an emergency state in a period τ is described by the trajectory of the system in a multidimensional space. The surface in this formulation separates the normal operating region from the emergency region. In order to obtain computationally tractable results, it was assumed that the surface could be approximated by a hyperplane. This allowed the application of scalar level crossing theory to obtain moments of the expected number of crossings of the hyperplane. An approximate expression for the hyperplane crossing probability in terms of these moments was also developed. An alternate method of analyzing this problem utilizes a partial differential equation of Dynkin which yields the moments of the first crossing time. Upper and lower bounds on the first and second (and in principal higher) moments were developed. These bounds provide estimates of the moments which are much simpler to compute than the solution to Dynkin's equation.

The second approach examined the long-term stability properties of the power system from the viewpoint of stability theory. The basic method was to analyze the stability properties of an unforced linear system where the system coefficients are subject to future jumps. Under the assumption that the elements of the dynamics matrix are discrete-state Markov processes, a sufficient condition for almost sure asymptotic stability was developed.

The third area of research has led to the formulation of a security measure suitable for on-line assessment of transient stability. This stability measure involves computation of the probability that a Lyapunov function will exceed a critical value in a specified future time interval. Specific distributions are assumed for the faults and expected clearing

times. To alleviate the computational burden, the security measure is divided into a real-time and a non-real-time component. The non-real-time computations are based on a set of network faults and resulting topological network changes. Simplifying approximations to estimate the kinetic energy increase during fault-on periods were evaluated along with other simplifications made to further reduce complexity

The interrelationship among the elements of the security assessment function is depicted schematically in Fig. 11 1-1. State estimation provides information on the current operating state and configuration of the power system. State forecasting provides a statistical indication of the future behavior of the system given its current state. The various security measures developed during this effort use this information to compute a statistical measure that may be related to the likelihood that the power system will pass from the current operating region to an emergency region. Translating these security measures (e.g., the expected first crossing time) into operator action to forestall loss of system security is beyond the scope of this effort. However, some general comments on the application of stochastic control concepts to security enhancement are given in the next section.

Figure 11 1-1 Functional Structure of TASC's Approach to Dynamic Security Assessment

To summarize, this research has led to the development of new long-term and transient dynamic security measures which are capable of providing the system operator with a real-time indication of power system security in terms of the likelihood that, given the present system operating state and a probabilistic description of future events, the system will undergo a transition from the normal operating state to the alert or emergency state. While similar in concept to other security indices, the security measures developed in the course of this effort are unique in their compatibility with the computational limitations imposed by real-time implementation. In addition to the security measures, new results were obtained in the area of time-scale decomposition of the long-term estimator equations and with the development of a closed form solution for the moments of the forecasted system state

11.2 THE APPLICATION OF STOCHASTIC CONTROL CONCEPTS TO INTEGRATE SECURITY ASSESSMENT AND SECURITY ENHANCEMENT

The application of the techniques of optimal stochastic control theory to an electrical generating, distribution and load network possesses a myriad of theoretical and practical problems. Basically these are the problems of nonlinearity, discontinuity and complexity (or scale). This report has discussed at

length the difficult problems of structure estimation and, in association with this, estimation of the system state (for given system structure) for those parts of the system with slow and fast dynamics. The first estimation procedure is carried out in a form of real-time hypothesis testing while the second uses a mixture of linear filtering techniques and singular perturbation theory

In the face of these complexities the derivation of a genuinely optimal stochastic control law for the system is, quite simply, impossible. Recall that in the current state of the art for the optimal control of stochastic systems with partial observations (see e.g., Ref. 84) a form of "separation theorem" can be established for such problems. That is to say, for a reasonably large class of nonlinear systems with partial observations, it can be shown that (subject to certain technical conditions) the optimal value of the loss function is attained by feedback controls depending only upon past inputs to and outputs from the system. The "state" of the "separated version" of the control problem is the unnormalized conditional distribution of the state of the original system if (again subject to certain technical restrictions) a conditional density exists for the state given past observations. Now, all of this has been derived for systems without jump disturbances, but it is reasonable to conjecture that the work may be extended to such cases.

At the cost of great computational complexity, the variable topology problem can be combinatorially embedded in a single large-scale state space problem. One could, for instance, conceive of a super-state and its conditional distribution evolving on the set of all possible configurations. The true state, of course, would evolve on only one of the configurations during the time interval between jump changes of the topology. In principle, the stochastic control problem could be formulated for such a set-up. Even if a conditional distribution could be recursively computed for the power system set-up described above (with or without random topology changes), one is faced with the problem of the computation of the optimal control law.

By use of the separation principle (which we have seen is justified) we have "reduced" the problem to the optimal control of a Markovian stochastic system with state dependent optimal controls. This, in principle, is solved (as is the more general partial observations case) by the Dynamic Programming Equation (DPE) of stochastic control. However, our Markovian system is the (linear) Zakai equation of nonlinear filtering which has a function space for a state space -- even in the case where conditional densities exist for the state of the original system.

Hence, an attempt to apply optimal stochastic control theory to the power systems problem results, at best, in a separated problem including a DPE for a infinite dimensional filtered system represented by the Zakai equation. Furthermore, techniques to account for the jump changes in system topology increase combinatorially the complexity of the underlying nonlinear stochastic system whose state space may need to be on the order of hundreds in order to give a reasonable system description.

We conclude that the application of optimal stochastic control theory to the power system problem is effectively beyond current theory, in so far as the theory gives no tractable solutions conceptually, let alone analytically. Further, by virtue of its combinatoric complexity and the function space nature of its state space, the problem is beyond the current

state of computer science and numerical analysis. None of this rules out, however the possibility that intelligent approximations and sufficient ingenuity could devise suboptimal (say, stabilizing) feedback stochastic controls. Such stabilizing stochastic controls would depend upon the deterministic controllability of the overall system and its stochastic stability. It should be emphasized at this point that the security assessment criteria, described elsewhere in this report, constitute a form of stochastic stability measure since they estimate, in one form or another, the probability of escape of the system from some specified region of the state space.

11.3 RECOMMENDATIONS

The research documented in this report has led to the development of two security measures: a long-term measure formulated in terms of a surface crossing problem and a Lypaunov function based transient stability measure suitable for on-line implementation. Neither of these measures were brought to a final stage of development. Additional effort in the following areas is recommended to assess the full utility of the proposed security measures:

- Obtain a numerical solution to Dynkin's equation for several example cases and compare these solutions with the upper and lower bounds derived during this effort.
- Investigate the adequacy of approximating the level crossing probability in terms of a computable number of factorial moments.
- Assess which of the two system models considered for the long-term security measure (Brownian motion-driven Markov process or differentiable sample path process) is a more suitable description of actual power system behavior
- Extend the scope of the simulation studies of the transient stability measure and assess the efficacy of alternative approximations or models, e.g., a different expression for the unstable equilibrium angles or more complex generator models.

Although the problem of interpreting the security measures and relating them to necessary control actions for security enhancement has not been examined in this effort, it is a crucial aspect of developing a dynamic security assessment/enhancement methodology. Further research in this area should focus on translating a warning of possible transition into an unsecure operating region (as embodied in one of the security measures discussed above) into a realistic control strategy (out of computational necessity suboptimal) for steering the power system into a more secure operating state. A computer-based operator decision aid using Artificial Intelligence to identify the problem and select an appropriate solution may prove to be useful in this regard.

APPENDIX A

POWER SYSTEM MODELING

A.1 INTRODUCTION

An electric power system is a complex dynamical system largely consisting of generating stations and load centers, connected by transmission lines. A generating station typically includes boiler-turbines, generators, and excitation systems. A load center may be a substation where a large sum of electric power is distributed among customers, or it may be a large factory whose power consumption is significant. In such a power system, the system dynamics cover a wide spectrum of electrical, electro-mechanical, and thermo-mechanical phenomena subject to continual random disturbances, such as load fluctuations and component failures. In addition, today's power systems are interconnected with each other throughout the nation, and therefore involve a large number of interacting systems. Modern power systems are therefore large-scale, time-varying, and stochastic in nature.

In order to provide a mathematical basis for studying security assessment issues, TASC has developed a model that encompasses all the above aspects of a power system. The model contains the state and input variables, nonlinearities, and disturbances, which are important for dynamic security assessment. The development of the power system model has relied upon presently available power system component models for the boiler, turbine, boiler and turbine controls, generator, exciter, and interconnection network models. The level of the selected modeling detail is appropriate for characterizing the effect of these components on power system security. The emphasis of this work is in combining the various component models, and on developing models for significant disturbance effects.

In addition to the power system components and disturbances, the sensor measurement data needed for performing the various security assessment functions -- state estimation, state forecasting, and computation of security measures -- are characterized. This includes both the types of system variables which are measured and the sensor error characteristics.

The remainder of this appendix is divided into five sections. A deterministic, dynamic model is discussed first in Section A.2. Disturbances necessary for a stochastic model of the power system are the subject of Section A.3, while measurement data characteristics are treated in Section A.4. Methods for dealing with the high dimensionality of the power system equations are covered in the last two sections. Section A.5 contains a review of dynamic equivalencing techniques and the selection of a suitable method for representing the external portions of the power network to which the system under study is connected. Finally, in Section A.6, network decomposition ideas are employed to decompose the large scale network into a number of subnetworks of manageable size.

A.2 DYNAMIC MODEL

A power system consists of the following principal components or subsystems:

- Interconnection Network
- Boiler-Turbine
- Generator

- Excitation System
- Load

These subsystems are interrelated as shown in Fig. A.2-1. The input variables are the throttle pressure set point, megawatt demand, reference frequency and reference voltage. The throttle pressure set point is generally fixed at a certain value for a boiler operating condition. The megawatt demand, reference frequency, and reference voltage are provided by a system dispatcher who oversees the power supply and demand balance in a system. An economic dispatch program or an automatic generation control program is used to compute megawatt demands; and a load frequency control program calculates a required reference frequency. Reference voltage is determined from the load demand, sometimes an excessive load demand is met by a voltage drop.

Figure A.2-1 Power System Functional Diagram

The megawatt demand and the difference between the reference frequency and the turbine speed drive the boiler-turbine subsystem to provide a mechanical power to the generator, which then converts this mechanical power into an electrical power. The generator maintains a certain terminal voltage level; this is done through an excitation system which varies the generator field voltage in response to an error signal consisting of the difference between the instantaneous terminal voltage and the reference voltage. Electrical power is transmitted to the loads through an interconnection network. In general, the transmission network receives power from many generators and supplies many loads. Power consumed by a load is affected by voltage and system frequency which are determined by the power balance in the interconnection network.

Power system dynamics describing the above process may be divided into two categories: transient and long-term. Transient dynamics consist of fast modes with time constants up to several seconds, and are governed by the interconnection network relations, generator electrical parameters and inertia, and the excitation and turbine speed controls. In this model, system frequency, which is an average speed of all generators in a system, is assumed to be constant. In contrast, the long-term dynamic model consists of slow modes with time constants up to several minutes, where slowly acting boiler and turbine dynamics, and interconnection network behavior are of primary interest. In this model, the generator representation is simplified to the extent that the turbine speed of each individual machine is not identified. Only an average speed, system frequency, is modeled.

In the remainder of this section, dynamic equations characterizing the individual subsystems are presented under the respective headings. References 35, 37 and 85 through 88 contain details on the various models, for each subsystem, a model reflecting dynamic properties important in security assessment has been selected. The fast and slow mode dynamics of each subsystem, termed transient and long-term dynamics, respectively, are also of concern. For some subsystems, different sets of equations are formulated for the transient and long-term dynamic models.

A 2.1 Interconnection Network

The interconnection network in a power system refers to the transmission system which connects generators and load centers (loads) as shown in Fig A 2-2. In the figure, n generators are connected to buses 1, 2, ..., n and r loads are connected to buses $n+1$, $n+2$, ..., $n+r$.

Figure A.2-2 Transmission System

For most power system analyses, the network is modeled by a set of simultaneous algebraic equations, as it is adequate to regard the network as though it were passing directly from one steady state to another. This is usually justified for two reasons:

- The electrical transients associated with the network are very much faster than the mechanical transients of interest.
- It is necessary to handle a large network by algebraic rather than differential equations from a computational standpoint

The algebraic equations describing the network make use of the nodal variables at each bus -- admittance, voltage, current and power -- and the Kirchhoff laws relating them, and are called the bus equations. The bus equations can be written in various ways depending on what variables are specified and what variables are to be computed. Commonly, three different types of buses are considered in forming the equations, namely P-Q bus, P-V bus, and swing bus.

Before introducing the bus equations for each bus type, some related variables are first explained. At each bus, the equations are written for either the nodal injected power, $\bar{S}_1$, or the nodal injected current, $\bar{I}_1$, where the subscript 1 indicates 1th bus. The nodal injected power is generally divided into real and reactive components, viz.,

$$\bar{S}_1 = P_1 + jQ_1 \quad (A\ 2-1)$$

The complex bus voltage at the 1th bus is denoted by $\bar{V}_1$, and may be written either in polar form

$$\bar{V}_1 = V_1 e^{j\delta_1} \quad (A\ 2-2)$$

or in rectangular form

$$\bar{V}_1 = V_{Q1} + jV_{D1}$$

where the subscripts Q and D indicate that the voltage is expressed in the Q-D frame of Park's transformation, which is a commonly used reference frame for the network. It is rotating at the synchronous velocity ω_{sys} . An element of the nodal admittance matrix denoted by $\bar{Y}_{1k}$ may be written as

$$\bar{Y}_{1k} = G_{1k} + jB_{1k}$$

or

$$\bar{Y}_{1k} = Y_{1k} e^{j\theta_{1k}}$$

(A 2-3)

In the case $1=k$, $\bar{Y}_{11}$ represents a self-admittance, otherwise, $\bar{Y}_{1k}$ represents a mutual admittance

P-Q Bus - At a P-Q bus, the nodal injected real and reactive power (P_1 and Q_1) are known, and the bus voltage is calculated from,

$$P_1 - jQ_1 = \bar{V}_1^* \left(\sum_{k=1}^N \bar{Y}_{1k} \bar{V}_k \right) \quad (A\ 2-4)$$

where N is the total number of buses in a system. For the system shown in Fig. A.2-2, $N = n+r$. Equation A.2-4 can be rewritten in the following way

$$\begin{aligned} (P_1 - jQ_1) - V_1 e^{-j\delta_1} (\bar{Y}_{11} V_1 e^{j\delta_1} + \sum_{\substack{k=1 \\ k \neq 1}}^N \bar{Y}_{1k} V_k e^{j\delta_k}) \\ = 0 \end{aligned} \quad (A\ 2-5)$$

In this study, load buses are considered P-Q buses exclusively. The nodal injected real and reactive powers are assumed to be inputs over which there is only coarse control. Section A.2.5 contains additional detail on load characteristics

P-V Bus - At a P-V bus, the nodal injected real power (P_1) and bus voltage magnitude (V_1) are known, and the bus voltage angle (rotor angle) is calculated from,

$$P_i - V_i^2 G_{ii} - \sum_{\substack{k=1 \\ k \neq i}}^N V_i V_k [B_{ik} \sin(\delta_i - \delta_k) + G_{ik} \cos(\delta_i - \delta_k)] = 0 \quad (A.2-6)$$

The above equation is simply the real part of Eq. A.2-5.

In the long-term dynamic model, all the generator buses are characterized P-V buses, since it is assumed that frequency is uniform throughout the system, and thus rotor angle swings of individual machines are not considered. The nodal injected real power is supplied from the boiler and the bus voltage magnitude is supplied from the exciter-generator.

Swing Bus - A bus connected to a swing machine where the rotor accelerates or decelerates according to mechanical and electrical torques supplied to the machine is termed a swing bus. At a swing bus, the bus current equation is written as,

$$\bar{I}_i - \sum_{k=1}^N \bar{Y}_{ik} \bar{V}_k = 0 \quad (A.2-7)$$

or

$$\bar{I}_i - (\bar{Y}_{ii} \bar{V}_i + \sum_{\substack{k=1 \\ k \neq i}}^N \bar{Y}_{ik} \bar{V}_k) = 0$$

Bus current $\bar{I}_i$ and voltage $\bar{V}_i$ are equivalent to terminal current and voltage of the swing machine since the latter is directly connected to the bus. The relationship between the network and machine variables is treated in Section A.2.3 after the swing machine equations are developed. Generators in the transient dynamic model are characterized as swing machines and are associated with swing buses.

A.2.2 Boiler-Turbine

Basic elements of a boiler-turbine system are:

- Boiler controls
- Boiler dynamics
- Turbine controls
- Turbine dynamics.

Interrelationships between the components are shown in Fig. A.2-3. Each component is discussed below.

Boiler Control Model - A typical fossil-fired, drum-type, reheat boiler unit with boiler-following type of control is chosen for this study, since it represents the majority of presently operating units. This typical unit employs the following control loops: a combustion control loop; reheat and main steam temperature control loops; and a drum level control loop. Among them, the combustion control loop is the only boiler control loop of importance in power system dynamic studies since it controls the drum pressure, which is a rough indication of stored thermal energy. The reheat and main steam temperature control loops

Figure A.2-3 Elements of a Boiler-Turbine System

may be ignored since steam temperature response is very slow and small temperature changes have little effect on stored energy and it is reasonable to assume that steam temperatures are always at the desired value. The drum level control loop may be ignored since the response of drum level and feedwater flow are very fast; consequently it is reasonable to assume that drum level is held constant.

The combustion control loop includes a pressure control and a fuel/air flow rate control, as shown in Fig. A.2-4 (Ref. 35). The pressure controller consists of a proportional-integral controller (with K_{BP} and K_{BI} as the respective gains) followed by a rate element. The output of the pressure controller is summed with the throttle steam flow w_{HP} . This signal then represents a demand. The fuel/air dynamics are represented by a simple first-order lag with time constant T_{BF} .

Figure A.2-4 Boiler Control System

Dynamic equations representing the boiler controls can be written from inspection of the block diagram

$$\begin{aligned} \dot{x}_{B1} &= K_{BI} \psi_e \\ \dot{x}_{B2} &= \frac{10}{T_{BR}} (-x_{B2} + 9 x_{B1} + 9 K_{BP} \psi_e) \\ \dot{L}_{CH} &= \frac{1}{T_{BF}} [-L_{CH} + K_{BF} (-x_{B2} + 10 x_{B1} + 10 K_{BP} \psi_e + w_{GV})] \end{aligned} \quad (A.2-8)$$

where

$$\psi_e = \psi_{SP} - \psi_{TH} = \text{pressure error}$$

ψ_{SP} = throttle pressure set point
which is an input variable

ψ_{TH} = throttle pressure fed from
the boiler

w_{GV} = governor value steam flow

L_{CH} = steam charging signal which is
an input to the boiler

x_{B1}, x_{B2} = intermediate state variables of
the boiler control system.

Since the above model involves slowly varying variables, it is employed only in the long-term dynamic model

Boiler Dynamic Model - A boiler, as shown in Fig. A.2-5, may be represented by a storage element or capacitance C_{BD} and a superheater pressure drop (friction coefficient K_{BS}) proportional to the square of the steam flow (Ref. 35). The input is a steam charging signal from the boiler control model and the output is throttle pressure ψ_{TH} .

Figure A.2-5 Boiler Model

The boiler dynamic equation can be written as,

$$\dot{\psi}_{TH} = \frac{1}{C_{BD}} (L_{CH} - w_{GV}) - K_{BS} w_{GV}^2 \quad (A.2-9)$$

As in the case of the boiler control model, this model is employed only in the long-term dynamic model.

Turbine Control Model - A turbine control model developed and published by the IEEE Task Force on Overall Plant Response (Ref. 86) has been adopted herein. The model is shown in Fig. A.2-6, and represents an electro-hydraulic system which applies to conventional fossil-fired units.

Figure A.2-6 Turbine Control Model

In this model, the reference load P_o is treated as an external input provided by a dispatching center that oversees the operation of all the generators in the power system and dispatches a required power generation for each generator. This initial value is combined with an increment due to speed deviation to obtain the valve outlet variable P_{GV} , subject to a first-order lag introduced by a servomotor mechanism. The output P_{GV} is either a desired governor valve position or a desired mechanical power produced by the turbine, depending on the turbine model. If the turbine system is modeled in detail to include a governor valve mechanism, P_{GV} represents the valve position, P_{GV} represents desired power output when the turbine system is modeled by a first-order lag. The input P_o has the same units as P_{GV} .

In power system dynamic studies, nonlinearities in the turbine speed control mechanism are normally neglected except for rate limits and the limits on valve position, as shown in Fig. A.2-6. Governor dead-band is not considered in the model. In studies where the performance of the speed-governing system itself is the primary concern, it may be necessary to represent dead-band, but it is normally not represented in large system studies.

Dynamic equations representing the turbine control are,

$$P_o' = P_o - K_{TC} \Delta\omega$$

$$P_o'' = \begin{cases} P_{MIN} & \text{if } P_o' < P_{MIN} \\ P_o' & \text{if } P_{MIN} \leq P_o' \leq P_{MAX} \\ P_{MAX} & \text{if } P_o' > P_{MAX} \end{cases}$$

$$\dot{P}_{GV}' = \frac{1}{T_{TC}} (P_o'' - P_{GV})$$

$$\dot{P}_{GV} = \begin{cases} \dot{P}_{DOWN} & \text{if } \dot{P}_{GV}' < \dot{P}_{DOWN} \\ \dot{P}_{GV}' & \text{if } \dot{P}_{DOWN} \leq \dot{P}_{GV}' \leq \dot{P}_{UP} \\ \dot{P}_{UP} & \text{if } \dot{P}_{GV}' > \dot{P}_{UP} \end{cases}$$

(A.2-10)

where $\Delta\omega$ is a frequency deviation given by,

$$\Delta\omega = \omega - \omega_o$$

in the transient dynamic model, and

$$\Delta\omega = \omega_{SYS} - \omega_o$$

in the long-term dynamic model. The turbine control model is used in both transient and long-term dynamic studies.

Turbine Dynamic Model - A tandem compound, single reheat steam turbine may be modeled as shown in Fig A.2-7 (Ref.86). In this model, the turbine control output P_{GV} represents a valve position. The latter is multiplied by a throttle pressure ψ_{TH} and a valve flow coefficient K_{GV} , describing governor valve action. The block following the governor valve action represents a steam bowl or chest. This vessel introduces a time-delay between changes in valve steam flow (w_{GV}) and steam flow in the high pressure turbine (w_{HP}); it is modeled by a first-order lag with time constant T_{CH} . Inlet piping to the first turbine cylinder, reheaters, and crossover downstream, all introduce delays in steam flow. The principal objective in modeling the steam system for dynamic security assessment is to account for these delays. Flows into and out of any steam vessel are related by a simple time constant.

Figure A.2-7 Tandem Compound, Single Reheat Steam Turbine Model

Dynamic equations characterizing the tandem compound, single reheat steam turbine as obtained from the block diagram are,

$$\begin{aligned}\dot{w}_{HP} &= \frac{K_{GV}}{T_{CH}} \psi_{TH} P_{GV} - \frac{1}{T_{CH}} w_{HP} \\ \dot{w}_{RH} &= \frac{1}{T_{RH}} w_{HP} - \frac{1}{T_{RH}} w_{RH} \\ \dot{w}_{CO} &= \frac{1}{T_{CO}} w_{RH} - \frac{1}{T_{CO}} w_{CO} \\ P_m &= F_{HP} w_{HP} + F_{IP} w_{RH} + F_{LP} w_{CO}\end{aligned}\quad (A.2-11)$$

where

- w_{HP} = high pressure turbine steam flow rate
- w_{RH} = reheater steam flow rate
- w_{CO} = crossover pipe steam flow rate
- T_{CH} = steam chest time constant
- T_{RH} = reheat time constant
- T_{CO} = crossover time constant
- F_{HP} = high pressure turbine power fraction
- F_{IP} = intermediate pressure turbine power fraction

F_{LP} = low pressure turbine power fraction

P_m = mechanical power

For the transient dynamic model, the turbine may be modeled as the simple first-order lag shown in Fig. A.2-8 (Ref 86). In this model, the output of turbine control model P_{GV} represents power at valve outlet. The corresponding dynamic equation is

$$\dot{P}_m = \frac{1}{T_{TT}} P_{GV} - \frac{1}{T_{TT}} P_m \quad (A.2-12)$$

where

T_{TT} = turbine system time constant used for transient dynamics.

Figure A.2-8 Turbine Model for Transient Dynamic Model

A.2.3 Generator

Large generators used in a power system typically consist of synchronous machines. For the transient dynamic model, these machines are treated in detail to reveal their interconnection relationships with the rest of the power system. In the long-term dynamic model, however, generators may be modeled as devices delivering fixed real power.

Synchronous Machine - A synchronous machine consists of two basic parts: a stator and rotor. The stator is a hollow laminated ferromagnetic cylindrical structure with longitudinal slots located on the inside surface. Placed in these slots are coils that are connected to form three separate windings. The stator therefore serves two basic functions: it is a mounting structure for the stator windings, and it provides a low reluctance return path for the rotor magnetic field.

The rotor is a solid ferromagnetic cylinder in turbine generators. Located on the rotor are a field winding excited from a dc source (exciter) and two amortisseur or damper windings. The purpose of the field winding is to produce an intense magnetic field. Since the rotor is mounted so that it can rotate within the stator, this field sweeps across the stator windings, inducing a voltage by Faraday's law.

The three windings on the stator and the three windings on the rotor are magnetically coupled. The magnetic coupling between the windings is a function of rotor position. Thus, the flux linking of each winding is also a function of rotor position. An instantaneous terminal voltage v of any winding is of the form,

$$\begin{aligned} v &= \Sigma r i \pm \Sigma \dot{\lambda} \\ &= \Sigma r / \ell \lambda \pm \Sigma \dot{\lambda} \end{aligned} \quad (\text{A.2-13})$$

where λ is the flux linkage, r is the winding resistance, ℓ is the winding inductance, and i is the current, with positive stator current flowing out of the generator terminals. The notation Σ indicates the summation of all appropriate terms with due regard to signs.

The well-known swing equation governs motion of a machine rotor relating the inertia torque to the resultant of the mechanical and electrical powers on the rotor, i.e.,

$$M \ddot{\delta}_m + D \dot{\delta}_m = P_a \quad (\text{A.2-14})$$

where M is the inertia constant of all rotating masses attached to the shaft, D is the damping coefficient, δ_m is the mechanical torque angle of the shaft with respect to a fixed reference, and P_a is the accelerating power acting on the shaft. Since the machine is a generator, the driving power P_m is mechanical and the retarding or load power P_e is electrical. Thus P_a can be written,

$$P_a = P_m - P_e \quad (\text{A.2-15})$$

which establishes a useful sign convention, namely, that in which a positive P_m accelerates the shaft, whereas a positive P_e is a decelerating power.

In general, a synchronous machine may be described by a set of coupled nonlinear differential equations of the form,

$$\dot{\underline{x}} = \underline{f}(\underline{x}, \underline{u}) \quad (\text{A.2-16})$$

Before the equations are stated, some simplifications will be made to reduce the order of the equations. First, it is assumed that voltages induced by the three stator windings are balanced, so that when these voltages are expressed in the o-d-q frame of Park's transformation, the voltage along the o-axis is reduced to zero. This assumption permits writing the equations in single-phase format.

Another assumption is that amortisseur effects are neglected, and the machine is considered to have a solid round rotor. The solid round rotor acts like a q-axis damper winding. The effect of amortisseur windings may be included in rotor damping power by increasing the damping coefficient in the power equation. In addition, the effect of air-gap flux saturation is ignored in developing the machine equations. The saturation effect slightly reduces induced terminal voltage in a nonlinear fashion, but it is not significant.

With the above assumptions, a set of simplified equations for stator and rotor windings of a synchronous machine may be obtained. A derivation is not provided here, but may be found in Ref. 37. The resulting equations are:

$$\dot{\lambda}_d = -\frac{r}{L_d'} \lambda_d + \frac{r}{L_d'} E_q' - \omega \lambda_q - V_d$$

$$\dot{\lambda}_q = \omega \lambda_d - \frac{r}{L_q'} \lambda_q - V_q$$

$$\dot{E}_q' = -\frac{L_d}{L_d' \tau_{do}'} E_q' + \frac{L_d - L_d'}{\tau_{do}'} \lambda_d + \frac{1}{\tau_{do}'} E_{FD}$$

$$\dot{\omega} = \frac{1}{\tau_H} P_m - \frac{1}{\tau_H L_d'} E_q' \lambda_q + \left(\frac{1}{\tau_H L_d'} - \frac{1}{\tau_H L_q'} \right) \lambda_d \lambda_q - \frac{D}{\tau_H} \omega$$

$$\dot{\delta} = \omega - \omega_R \quad (\text{A.2-17})$$

where

λ_d = d-axis stator equivalent flux linkage

λ_q = q-axis stator equivalent flux linkage

E_q' = q-axis voltage proportional to the field flux linkage resulting from the combined effect of the field and armature currents (voltage behind transient reactance)

ω = angular velocity of the revolving magnetic field

δ = torque angle corresponding to the electrical angle between the generated EMF and the resultant stator voltage phasors

L_d' = d-axis transient inductance

L_d = d-axis synchronous inductance

L_q = q-axis synchronous inductance

r = stator winding resistance

τ_{do}' = d-axis transient open circuit time constant

τ_H = rotating inertia time constant

D = machine load damping coefficient

V_d = d-axis stator equivalent voltage

V_q = q-axis stator equivalent voltage

E_{FD} = rotor field voltage by excitation

P_m = mechanical power output from the turbine

ω_R = generator rated speed.

In the above equations, Park's transformation is utilized so that the stator and rotor winding equations (the first three equations) are expressed in a q-d axis frame. Therefore, these equations represent all three-phase quantities. To be consistent, swing equations (the last two equations) are also given for three-phase base. All the variables and parameters

(including time) in Eq. A.2-17 are expressed as per-unit quantities. Thus, the time-constants must be in radians; the relationship between per-unit time and standard time is given by,

$$\tau_{pu} = \tau_{sec} \omega_R \text{ (rad)}$$

Equation A.2-17 includes four external variables which relate the machine equations to the turbine, exciter, and network equations. The mechanical power, P_m , is associated with the turbine, and the excitation voltage, E_{FD} , is associated with the exciter. The stator equivalent voltages, V_d and V_q , are associated with the network equations. This is explained below.

Relation Between Machine and Network Equations - Machine and network equations are related through the terminal (or bus) voltage. The terminal voltage in the machine equations is expressed in the individual machine frame (q-d frame); while it is expressed in a common reference frame (Q-D frame) in the network equations. The two frames are related by the torque angle δ , as shown in Fig. A.2-9. The terminal voltage of machine 1 can be expressed in either the common frame $V_1^{(r)}$ or the machine frame $V_1^{(m)}$; by inspection of Fig. A 2-9, they are related by,

$$\begin{aligned} V_{Q1} + j V_{D1} &= (V_{q1} \cos \delta_1 - V_{d1} \sin \delta_1) \\ &+ j (V_{q1} \sin \delta_1 + V_{d1} \cos \delta_1) \end{aligned}$$

or

$$V_1^{(r)} = V_1^{(m)} e^{j\delta_1} \quad (\text{A.2-18})$$

Figure A.2-9 Two Frames of Reference for Phasor Quantities for a Voltage V_1

The above relationship also holds for current, i.e.,

$$I_1^{(r)} = I_1^{(m)} e^{j\delta_1} \quad (\text{A.2-19})$$

A swing bus equation was written in Eq. A.2-7 without specifying its coordinate frame. In the common reference frame, it is expressed as,

$$I_1^{(r)} = \sum_{k=1}^N \bar{Y}_{1k} V_k^{(r)} \quad (\text{A.2-20})$$

Combining Eqs. A.2-18 through A.2-20, and separating generator and load buses gives

$$I_1^{(m)} = \sum_{k=1}^n \bar{Y}_{1k} V_k^{(m)} e^{j\delta_k} e^{-j\delta_1} + \sum_{k=n+1}^{n+r} \bar{Y}_{1k} V_k^{(r)} e^{-j\delta_1} \quad (\text{A 2-21})$$

Dividing $I_1^{(m)}$ into real and imaginary parts yields

$$\begin{aligned} I_{q1} &= \sum_{k=1}^n Y_{1k} V_k \cos[\theta_{1k} + \theta'_1 - (\delta_1 - \delta_k)] \\ &+ \sum_{k=n+1}^{n+r} Y_{1k} V_k \cos[\theta_{1k} - (\delta_1 - \delta_k)] \\ I_{d1} &= \sum_{k=1}^n Y_{1k} V_k \sin[\theta_{1k} + \theta'_1 - (\delta_1 - \delta_k)] \\ &+ \sum_{k=n+1}^{n+r} Y_{1k} V_k \sin[\theta_{1k} - (\delta_1 - \delta_k)] \end{aligned} \quad (\text{A 2-22})$$

where

$$\begin{aligned} Y_{1k} &= \sqrt{G_{1k}^2 + B_{1k}^2} \\ V_k &= \sqrt{V_{Qk}^2 + V_{Dk}^2} = \sqrt{V_{qk}^2 + V_{dk}^2} \\ \theta'_1 &= \tan^{-1} \frac{V_{d1}}{V_{q1}} \end{aligned}$$

The current can also be calculated using flux linkages and voltage behind transient reactance, i.e.,

$$\begin{aligned} I_{q1} &= \frac{1}{\bar{L}_{q1}} \Lambda_{q1} \\ I_{d1} &= \frac{1}{\bar{L}_{d1}} \Lambda_{d1} - \frac{1}{\bar{L}_{d1}} E'_{q1} \end{aligned} \quad (\text{A.2-23})$$

Combining Eqs. A.2-22 and A.2-23 gives, with some rearrangement,

$$\begin{aligned} &\frac{1}{\bar{L}_{q1}} \Lambda_{q1} - \sum_{k=1}^n Y_{1k} \{V_{qk} \cos[\theta_{1k} - (\delta_1 - \delta_k)] \\ &- V_{dk} \sin[\theta_{1k} - (\delta_1 - \delta_k)]\} \\ &- \sum_{k=n+1}^{n+r} Y_{1k} V_{\ell k} \cos[\theta_{1k} - (\delta_1 - \delta_k)] = 0 \end{aligned}$$

$$\begin{aligned}
\frac{1}{L'_{di}} \Lambda_{di} - \frac{1}{L'_{di}} E'_{qi} + \sum_{k=1}^n Y_{ik} \{V_{qk} \sin[\theta_{ik} - (\delta_i - \delta_k)] \\
- V_{dk} \cos[\theta_{ik} - (\delta_i - \delta_k)]\} \\
+ \sum_{k=n+1}^{n+r} Y_{ik} V_{dk} \sin[\theta_{ik} - (\delta_i - \delta_k)] = 0
\end{aligned}
\quad (A.2-24)$$

Equations A.2-18 and A.2-24 relate the machine and network equations.

Simplified Generator Representation for the Long-Term Dynamic Model - In long-term dynamic studies, a simplified generator model is employed to describe a fixed generation power supplied to a system. The generator model shown in Fig. A.2-10 (Ref. 85) is a first-order lag with a time constant T_G and a gain K_G . The power delivered by the generator is equivalent to the mechanical power output from the turbine. The differential equation for the terminal voltage is

$$\dot{V}_t = \frac{K_G}{T_G} E_{FD} - \frac{1}{T_G} V_t \quad (A.2-25)$$

Figure A.2-10 A Simplified Generator Model

A basic assumption in the long-term dynamic model is that frequency is uniform throughout the system. Thus, rotor angle swings of the individual generators are not considered. The rate of change of the uniform system frequency, ω_{SYS} , is proportional to the accelerating power P_{acc} , viz.,

$$\dot{\omega}_{SYS} = \frac{P_{acc}}{2H_{SYS} \omega_{SYS}} \quad (A.2-26)$$

where H_{SYS} is the system inertia given by,

$$H_{SYS} = \sum_{i=1}^n H_i \quad (A.2-27)$$

where H_i is the inertia of generator i , and n is the number of generators in the system. Accelerating power is the difference between the power out of the turbines and the sum of the system losses and loads, viz.,

$$P_{acc} = \sum_{i=1}^n P_{mi} - P_L - \sum_{j=1}^r P_{Lj} \quad (A.2-28)$$

The system losses, P_L , are due mainly to power losses in transmission lines. Power loss in a transmission line may be computed from bus voltages as illustrated below.

The current at bus p in a line connecting bus p to q is given by,

$$\bar{i}_{pq} = (\bar{V}_p - \bar{V}_q) \bar{y}_{pq} + \bar{V}_p \frac{\bar{y}'_{pq}}{2} \quad (A.2-29)$$

where

$\bar{y}_{pq}$ = line admittance

$\bar{y}'_{pq}$ = total line charging admittance.

The real power flow is computed as,

$$\begin{aligned}
P_{pq} &= \text{Re}\{\bar{V}_p^* \bar{i}_{pq}\} \\
&= \text{Re}\{\bar{V}_p^* (\bar{V}_p - \bar{V}_q) \bar{y}_{pq} + \bar{V}_p^* \bar{V}_p \frac{\bar{y}'_{pq}}{2}\}
\end{aligned}
\quad (A.2-30)$$

where at bus p the real power flow from bus p to q is P_{pq} . Similarly, at bus q the real power flow from q to p is given by,

$$P_{qp} = \text{Re}\{\bar{V}_q^* (\bar{V}_q - \bar{V}_p) \bar{y}_{pq} + \bar{V}_q^* \bar{V}_q \frac{\bar{y}'_{pq}}{2}\} \quad (A.2-31)$$

The power loss in line p - q is the sum of the real power flows of Eqs. A.2-30 and A.2-31, and is given by,

$$\begin{aligned}
P_{Lpq} &= \{V_p^2 + V_q^2 - V_p V_q \cos(\delta_p - \delta_q)\} y_{pq} \cos \theta_{pq} \\
&\quad + \frac{1}{2} (V_p^2 + V_q^2) y'_{pq} \cos \theta'_{pq}
\end{aligned}
\quad (A.2-32)$$

Total power loss P_L in a system is obtained by summing all the line losses given above, viz.,

$$P_L = \sum_{p=1}^{n+r} \sum_{q=1}^{n+r} (P_{Lpq}) \quad (A.2-33)$$

A.2.4 Excitation System

Since excitation controllers affect almost all aspects of power system behavior, they should be modeled whenever detailed generator models are used. There are presently a wide variety of excitation systems in use. However, the majority in service can be represented adequately by the IEEE Type 1 model presented by the Working Group of the Excitation Systems Subcommittee of the IEEE Power Generation Committee (Ref. 88). Basic components of the model are the regulator, amplifier, and exciter each of which is represented by a first-order lag. For most systems, the time constant of the regulator is very small, and it is ignored here. Figure A.2-11 shows a simplified excitation system.

Figure A.2-11 Excitation System Model

Most excitation systems use an input signal provided by a stabilizer, in addition to the generator terminal voltage. The stabilizer is, however, a secondary system in a generator unit, and it is omitted here. The model shown in Fig. A.2-11 contains two nonlinearities. The regulated voltage is limited by the exciter mechanical limit. The other nonlinearity compensates for exciter saturation.

Dynamic equations representing the excitation system are (from Fig. A.2-11),

$$\dot{V}_F = -\frac{1}{T_{EF}} V_F + \frac{K_{EF}}{T_{EE}} [(V_R - V_S) - K_{EE} E_{FD}]$$

$$\dot{V}_R = \frac{K_{EA}}{T_{EA}} (V_{REF} - V_t - V_F) - \frac{1}{T_{EA}} V_R$$

$$\dot{E}_{FD} = \frac{1}{T_{EE}} (V_R - V_S) - \frac{K_{EE}}{T_{EE}} E_{FD}$$

$$V_R = \begin{cases} V_{RMAX} & \text{if } V'_R > V_{RMAX} \\ V'_R & \text{if } V_{RMIN} \leq V'_R \leq V_{RMAX} \\ V_{RMIN} & \text{if } V'_R < V_{RMIN} \end{cases}$$

$$V_S = A \cdot e^{B \cdot E_{FD}}$$

(A.2-34)

where

- V_{REF} = input reference voltage
- V_t = terminal voltage
- A, B = exciter saturation characteristic constants
- T_{EA} = amplifier time constant
- K_{EA} = amplifier gain
- T_{EE} = exciter time constant
- K_{EE} = exciter gain
- T_{EF} = stabilizing loop time constant
- K_{EF} = stabilizing loop gain.

This excitation system model is used in both the transient and long-term dynamic models.

A.2.5 Load Model

The effect of the load plays a significant role in dynamic studies of a power system. Loads are spatially distributed and are represented by a load model at each P-Q bus in the system. Loads vary due to weather patterns, living schedules, etc. A load model should exhibit the following major effects

- Voltage sensitivity
- Frequency sensitivity
- Load tap changer action to maintain a constant voltage or constant volt/frequency ratio
- Load tripping due to underfrequency and undervoltage relays
- Load management by scheduled or remote load control
- Weather sensitivity
- Periodic energy consumption patterns
- Breaking resistors, series capacitors, and shunt reactors used for stability augmentation.

The load at bus 1 can be modeled as follows:

$$P_1 = P_{01} R_1(f)^{a_1} (C_1 V_1)^{b_1} \quad (A.2-35)$$

$$Q_1 = Q_{01} R_1(f)^{c_1} (C_1 V_1)^{d_1} \quad (A.2-36)$$

where

P_1, Q_1 = actual real and reactive powers at bus 1

P_{01}, Q_{01} = actual real and reactive powers consumed at bus 1 under nominal voltage, frequency, and system conditions

R_1 = ratio of load reduction due to load control and/or load management action

C_1 = voltage ratio of load tap changer

V_1 = voltage at bus i (on system side of load tap changer)

a_1, b_1, c_1, d_1 = load model parameters

The dynamic behavior of the load is given by its voltage and frequency sensitivity, tap changer action, and load control management. These characteristics of

the load can be considered as being deterministic. In the transient dynamic model, the frequency is held constant and the frequency dependent terms in Eqs. A.2-35 and A.2-36 may be dropped.

A.2.6 Summary of the Dynamic Equations

The detailed component models described above are collected together in Eqs. A.2-37 and A.2-38 which represent the transient and long-term dynamic models, respectively, for an n -generator and r -load system. For convenience generator buses are numbered 1 through n and load buses are numbered $n+1$ through $n+r$. All the variables (including time) and parameters are as per-unit quantities with a three-phase base.

TRANSIENT DYNAMIC DIFFERENTIAL EQUATIONS

$$\begin{bmatrix} \dot{p}_{CV1} \\ \dot{p}_{m1} \\ \dot{\delta}_1 \\ \dot{\omega}_1 \\ \dot{v}_{F1} \\ \dot{v}_{R1} \\ \dot{p}_{FD1} \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{T_{T1}} & -\frac{1}{T_{T1}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{r_1}{L_{d1}} & 0 & \frac{r_1}{L_{d1}} & 0 & 0 \\ 0 & 0 & 0 & -\frac{r_1}{L_{q1}} & 0 & 0 & 0 \\ 0 & 0 & \frac{L_{d1}'}{L_{d1}'} - \frac{L_{d1}'}{L_{d1}'} & 0 & -\frac{L_{d1}'}{L_{d1}'} & 0 & \frac{1}{T_{dof}} \\ 0 & \frac{1}{H_1} & 0 & 0 & 0 & -\frac{D_1}{H_1} & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{1}{T_{EF1}} \\ 0 & 0 & 0 & 0 & 0 & \frac{E_{A1}}{T_{EA1}} - \frac{1}{T_{EA1}} & \frac{E_{A1} E_{F1}}{T_{EA1} T_{EF1}} \\ 0 & 0 & 0 & 0 & 0 & 0 & -\frac{E_{F1}}{T_{EF1}} \end{bmatrix} \begin{bmatrix} p_{CV1} \\ p_{m1} \\ \delta_1 \\ \omega_1 \\ v_{F1} \\ v_{R1} \\ p_{FD1} \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{E_{A1}}{T_{EA1}} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} p_{o1} \\ \omega_o \\ v_{REF1} \\ p_{soj} \\ Q_{soj} \\ R_j \\ C_{zj} \end{bmatrix}$$

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_{d1} \\ v_{q1} \\ v_{zj} \\ \delta_j \end{bmatrix} + \begin{bmatrix} E_{LIN} \left\{ \frac{1}{T_{C1}} \left(E_{LIN}(p_{o1} - E_{TL1}(\omega_1 - \omega_o)) - p_{CV1} \right) \right\} \\ 0 \\ -\omega_1 \omega_{q1} \\ \omega_1 \omega_{d1} \\ 0 \\ -\frac{1}{T_{H1}} \frac{E_{q1}}{L_{d1}} \omega_{q1} + \frac{1}{T_{H1}} \left(\frac{1}{L_{d1}} - \frac{1}{L_{q1}} \right) \omega_{d1} \omega_{q1} \\ -\omega_{R1} \\ 0 \\ -\frac{E_{A1}}{T_{EA1}} \sqrt{v_{q1}^2 + v_{d1}^2} \\ \frac{1}{T_{EE1}} \left\{ E_{LIN}(v_{R1}) - A_1 + \theta_1 p_{FD1} \right\} \end{bmatrix}$$

ALGEBRAIC EQUATIONS

$$\begin{aligned} \frac{1}{L_{q1}} \omega_{q1} - \sum_{k=1}^n V_{ik} \{ V_{qk} \cos[\theta_{ik} - (\delta_1 - \delta_k)] - V_{dk} \sin[\theta_{ik} - (\delta_1 - \delta_k)] \} \\ - \sum_{k=n+1}^{n+r} V_{ik} V_{zk} \cos[\theta_{ik} - (\delta_1 - \delta_k)] = 0 \\ \frac{1}{L_{d1}} \omega_{d1} - \frac{1}{L_{d1}} E_{q1} + \sum_{k=1}^n V_{ik} \{ V_{qk} \sin[\theta_{ik} - (\delta_1 - \delta_k)] + V_{dk} \cos[\theta_{ik} - (\delta_1 - \delta_k)] \} \\ + \sum_{k=n+1}^{n+r} V_{ik} V_{zk} \sin[\theta_{ik} - (\delta_1 - \delta_k)] = 0 \\ p_{soj} R_j (C_{zj} V_{zj})^{b_j} - \sum_{k=1}^n V_{jk} V_{zj} \{ V_{qk} \cos[\theta_{jk} - (\delta_1 - \delta_k)] \\ - V_{dk} \sin[\theta_{jk} - (\delta_1 - \delta_k)] \} - \sum_{k=n+1}^{n+r} V_{jk} V_{zk} V_{zj} \cos[\theta_{jk} - (\delta_j - \delta_k)] = 0 \\ Q_{soj} R_j (C_{zj} V_{zj})^{d_j} + \sum_{k=1}^n V_{jk} V_{zj} \{ V_{qk} \sin[\theta_{jk} - (\delta_1 - \delta_k)] \\ + V_{dk} \cos[\theta_{jk} - (\delta_1 - \delta_k)] \} + \sum_{k=n+1}^{n+r} V_{jk} V_{zk} V_{zj} \sin[\theta_{jk} - (\delta_j - \delta_k)] = 0 \\ i = 1, 2, \dots, n, \quad j = n+1, n+2, \dots, n+r \end{aligned}$$

(A.2-37)

[illegible]
$$P_{acc} = \sum_{k=1}^n P_{mk} - P_L - \sum_{k=n+1}^{n+r} P_{zk}$$

$$P_{acc} = \sum_{k=1}^n P_{mk} - P_L - \sum_{k=n+1}^{n+r} P_{zk}$$

$$P_{m1} = F_{HP1} W_{HP1} + F_{IP1} W_{RH1} + F_{LP1} W_{CO1}$$

$$P_L = \sum_{p=1}^{n+r} \sum_{q=1}^{n+r} \left[\{v_p^2 + v_q^2 - v_p v_q \cos(\delta_p - \delta_q)\} y_{pq} \cos \theta_{pq} + \frac{1}{2}(v_p^2 + v_q^2) y'_{pq} \cos \theta_{pq} \right]$$

$$P_{LJ} = P_{L0J} R_J (\omega_{SYS})^{a_J} (C_{LJ} V_J)^{b_J}$$

$$Q_{Lj} = Q_{Loj} R_j (\omega_{SYS})^{c_j} (C_{Lj} V_j)^{d_j}$$

$$P_{m1} = \sum_{k=1}^{n+r} \gamma_{1k} v_1 v_k \cos \{\theta_{1k} - (\delta_1 - \delta_k)\} = 0$$

$$P_{Lj} = \sum_{k=1}^{n+r} Y_{jk} V_j V_k \cos [\theta_{jk} - (\delta_j - \delta_k)] = 0$$

$$Q_{2j} + \sum_{k=1}^{n+r} Y_{jk} V_j V_k \sin [\theta_{jk} - (\delta_j - \delta_k)] = 0$$

$$N_{\text{SYS}} = \sum_{k=1}^n N_k; \quad i = 1, 2, \dots, n, \quad j = n+1, n+2, \dots, n+r.$$

Equations A.2-37 and A.2-38 are of the form:

$$\dot{\underline{x}} = \underline{f}(\underline{x}, \underline{u}, \underline{p})$$

(A.2-40)

$$\dot{\underline{x}} = \underline{F}\underline{x} + \underline{G}\underline{u} + \underline{C}\underline{p} + \underline{f}_n(\underline{x}, \underline{u}, \underline{p})$$

$$\underline{0} = \underline{g}(\underline{x}, \underline{u}, \underline{p})$$

$$\underline{0} = \underline{g}(\underline{x}, \underline{u}, \underline{p})$$

The above equations model the power system by a coupled set of differential equations and algebraic equations. The presence of the algebraic equations in the general model is somewhat awkward. A more tractable form for the model is developed in Appendix B where auxiliary equations involving the network variables are used to obtain an integrated dynamical descriptor of the power system.

Power systems may be subject to three distinct disturbance classes -- line outages, generator outages and load disturbances. Line outages arise from faults on transmission lines or cascading events due to protective relay action following severe overloads or a violation of operating limits. Generator outages may be due to loss of synchronism or equipment malfunctions. Both line and generator outages result in circuit breaker trips which disconnect the troubled

where

x = state vector

u = input vector

p = network variable vector

$$\underline{f}_n = \text{nonlinear vector function.}$$

The above form is adopted to more clearly indicate the structure of the equations. It should be noted that the C_p term does not appear in Eq. A.2-38. Equation A.2-39 may be written more compactly as

lines or generators from the rest of the system. Load disturbances may consist of smooth, random fluctuations resulting from small, slowly varying load changes or possibly by sudden, large load jumps.

Due to their inherent similarity, line and generator outages are discussed together in Section A 3.1. A load model which includes stochastic disturbance terms is treated in Section A 3 2

A.3.1 Models for Line and Generator Outage

The discrete nature of a line or generator outage is conveniently modeled as a network structural change. In other words, if a line is out that line is removed from the network configuration, if a generator is out, then the generator bus together with its connected lines are removed from the network configuration.

Consider first, the case of a transmission line outage. Each line is modeled in terms of the complex mutual admittance $\bar{Y}_{ij}$ between its endpoints*. Loss of the line is equivalent to the condition $\bar{Y}_{ij} = 0$, thus it is straightforward to introduce the binary random variable

$$\xi_{ij} = \begin{cases} 1 & \text{line between nodes } i \text{ and } j \\ & \text{intact} \\ 0 & \text{line is open} \end{cases}$$

and use the product $\xi_{ij}\bar{Y}_{ij}$ instead of $\bar{Y}_{ij}$ in the network equations (Eq. A.2-5). The random nature of line outages are thus embodied in the ξ_{ij} .

For a generator outage, the situation is slightly more complicated. If generator k is lost, all interaction between the k^{th} generator and the remaining $n-1$ generators and r loads drops to zero. The interaction between the k^{th} generator and the remaining generators can be eliminated by introducing another set of binary random variables ξ_k defined by

$$\xi_k = \begin{cases} 1 & \text{Generator } k \text{ operative} \\ 0 & \text{Generator } k \text{ inoperative} \end{cases}$$

and then multiplying all the state variables associated with generator k in the dynamic equations by ξ_k . The effect of setting $\xi_k = 0$ is to isolate the k^{th} generator from the network.

For notational convenience, introduce the vector $\underline{z}$ which has components ξ_{ij} and ξ_j . A fixed ordering of the elements within $\underline{z}$ can be obtained with a standard enumeration scheme. Since the number of generators is fixed set $\underline{z}_\ell = \xi_\ell$, $\ell = 1, 2, \dots, n$. To include the ξ_{ij} , $i, j = 1, 2, \dots, n+r$, construct the array shown in Fig. 2.3-1 and order the ξ_{ij} in the sequence shown, i.e., $\xi_{11}, \xi_{21}, \xi_{12}, \xi_{31}, \xi_{22}, \dots$

Figure A.3-1 Ordering of the ξ_{ij}

Pass through the sequence deleting those ξ_{ij} for which $\bar{Y}_{ij} = 0$ (i.e., for which there is no line connecting Bus i to Bus j). Finally, assign to z_{n+1} the first nonzero element, to z_{n+2} the second, continuing until the sequence is exhausted. This uniquely defines the vector $\underline{z}$ for a specific network configuration.

A reasonable general model to describe the behavior of z_1 is that of the random telegraph wave (see Fig. A 3-2), viz,

$$z_1(t) = \frac{1}{2} (-1)^{N_1(t)} + \frac{1}{2} \quad (\text{A.3-1})$$

with $N_1(t)$ a Poisson counting process* with intensity $\lambda_1(t)$. To illustrate the properties of this model, consider the case of one generator which is in either an operating state ($z(t) = 1$) or a failed state ($z(t) = 0$). Assume that the failure and repair times are exponentially distributed with mean time to failure τ_F and mean repair time τ_R . If the generator is initially in the operative state, the intensity of the underlying Poisson process is given by

$$\lambda(t) = \begin{cases} 1/\tau_F & w_{2j} \leq t < w_{2j+1} \\ 1/\tau_R & w_{2j+1} \leq t < w_{2j+2} \end{cases} \quad j = 0, 1, 2, \dots \quad (\text{A.3-2})$$

Figure A.3-2 Sample Function of a Random Telegraph Wave

*Parallel transmission lines require a slightly more complicated approach which is not discussed herein.

*See Appendix C for a brief overview of Poisson counting processes.

A.3.2 Stochastic Load Disturbance Model

An ideal situation from the point of view of a power system control center would be to have complete knowledge of customer load demand. This situation is obviously impossible. However, at the level of a distribution substation a pattern of load demand tends to emerge. At the subtransmission and transmission levels the load is very repeatable such that it becomes possible to predict (or forecast) average load demand quite accurately. The problems associated with electric load forecasting and the identification of load model parameters from physical data are described in Ref. 89 (and other work referenced therein). It will be assumed here that adequate models exist which allow the real and reactive load demand to be forecast over the periods of interest.

Any forecast of user demand cannot be perfect; thus there are discrepancies between the actual load demand and the forecasted value. A general model for such differences is

$$\epsilon L(t) = v(t) + n_p(t) \quad (A.3-3)$$

where

$\epsilon L(t)$ = load power misprediction

$v(t)$ = random process modeling the continuous portion of the load misprediction

$n_p(t)$ = random (jump) process modeling the discontinuous portion of the load misprediction

The almost surely (a.s.) continuous sample path process $v(t)$ can be adequately represented as the output of a linear dynamical system driven by a Wiener process w . To illustrate this technique, let $\underline{v} = (v_1, \dots, v_p)^T$ satisfy the stochastic differential equation

$$d\underline{v} = \underline{A}\underline{v} dt + \underline{B}dw \quad (A.3-4)$$

and set

$$v(t) = \underline{c}^T \underline{v} \quad (A.3-5)$$

If the process $v(t)$ has a known, rational spectrum, then stochastic realization theory (Ref. 90) can be used to determine the elements of $\underline{A}$, $\underline{B}$ and $\underline{c}$.

The jump process $n_p(t)$ (which is a.s. discontinuous) models discrete changes in bus load power. Since these changes may tend to increase or decrease the load in varying amounts, the jumps in $n_p(t)$ must be both positive and negative and in addition have a distribution of amplitudes. One random process that does exhibit these characteristics is the generalized or compound Poisson process. A brief description of this process may be found in Appendix C.

Determination of the specific parameters associated with these processes is beyond the scope of this work. If sufficient physical data is available, it is possible to apply the techniques discussed in Ref. 89

as well as the conventional tools of time series analysis (Refs. 91 and 92) to obtain a load forecast model and the parameters necessary to characterize the processes $v(t)$ and $n_p(t)$. In the remainder of this report it will be assumed that the actual real and reactive loads take the form

$$P_j(t) = P_j^f(t) + v_j^P(t) + n_{p_j}^P(t) \quad (A.3-6)$$

$$Q_j(t) = Q_j^f(t) + v_j^Q(t) + n_{p_j}^Q(t) \quad (A.3-7)$$

$$j = n+1, \dots, n+r$$

where

P_j^f, Q_j^f = forecasted (i.e., modeled) real and reactive load powers at the j^{th} bus

v_j^P, v_j^Q = a.s. continuous sample path components of the difference between the actual and forecasted real and reactive powers at the j^{th} bus

$n_{p_j}^P, n_{p_j}^Q$ = a.s. discontinuous sample path components of the difference between the actual and forecasted real and reactive powers at the j^{th} bus

A.4 MEASUREMENT MODEL

Presently, measurements of various mechanical and electrical variables in a power system are made primarily for the purposes of steady-state estimation, dispatching, billing, and control. A typical example of the measurement requirements from a dispatching center is the New England Power Exchange (NEPEX) Metering and Telemetering Criteria (Ref. 93); it is used here as a guideline to specify a measurement model.

With some accuracy improvements and an increase in metering frequency, presently available measurements can be utilized for dynamic state estimation and forecasting. Typical measurements made for steady-state estimation include line power flows, bus real and reactive powers, and bus voltage magnitudes. For billing purposes, a utility is required to measure and record tie-line power flows and MW generation in all power plants. For dispatching purposes, generated real power (MW) and reactive power (MVAR) from generators, bus voltages, MW and MVAR from transmission lines and transformers, and breaker status from transmission line and generator breakers are measured. For boiler-turbine control purposes, turbine or generator rotor speed, turbine throttle and first stage steam pressures, valve position, and boiler feed water flow rate are measured. System frequency is measured for load frequency control. Generator terminal voltage and rotor field voltage are measured for excitation control.

The measurements discussed above are formulated in terms of the variables in the dynamic equations

specified by Eqs. A.2-37 and A.2-38. First, the measurement equations associated with the transient dynamic equations (Eq. A.2-37) are given

$$y_1 = P_{GV1}$$

$$y_2 = P_{m1}$$

$$y_3 = \sqrt{\left(\frac{\Delta_{q1}}{L_{q1}}\right)^2 + \left(\frac{\Delta_{d1} - E'_{q1}}{L_{d1}}\right)^2}$$

$$y_4 = \omega_1$$

$$y_5 = E_{FD1}$$

$$y_6 = V_{t1}$$

$$y_7 = V_{\ell j}$$

$$y_8 = V_p V_q y_{pq} \sin(\delta_p - \delta_q)$$

$$y_9 = -V_p^2 y_{pq} - V_p V_q y_{pq} \cos(\delta_p - \delta_q) + \frac{1}{2} V_p^2 y'_{pq}$$

$$i = 1, 2, \dots, n$$

$$j = n+1, n+2, \dots, n+r$$

$$p, q = 1, 2, \dots, n+r$$

(A.4-1)

where

n = number of generator buses

r = number of load buses.

Measurement y_1 indicates power through the governor valve which can be inferred from governor valve position. The second equation indicates mechanical power supplied to a generator. The third equation represents a generator current measurement which can be measured easily with an ammeter. Measurement ω_1 is rotor speed deviation measured by tachometers. The rotor field voltage E_{FD1} is measured from a voltmeter; V_{t1} and $V_{\ell j}$ are bus terminal voltages of a generator and load, respectively. The last two measurements in Eq. A.4-1 represent real and reactive line power flows; since these expressions are not obvious, the derivation is presented below.

Power flow from bus p to bus q can be expressed using bus voltages, line admittance, and shunt admittance, as follows;

$$\bar{S}_{pq} = \bar{V}_p^* (\bar{V}_p - \bar{V}_q) \bar{y}_{pq} + \bar{V}_p^* \bar{V}_p \frac{\bar{y}'_{pq}}{2} \quad (A.4-2)$$

where

$\bar{S}_{pq}$ = complex power flow from bus p to bus q

$\bar{V}_p, \bar{V}_q$ = bus voltages

$\bar{y}_{pq}$ = p - q line admittance

$\bar{y}'_{pq}$ = shunt admittance

Using polar expressions for voltages and admittances, $\bar{S}_{pq}$ can be rewritten,

$$\begin{aligned} \bar{S}_{pq} = & [V_p^2 y_{pq} \cos \theta_{pq} - V_p V_q y_{pq} \cos(-\delta_p + \delta_q + \theta_{pq}) \\ & + \frac{1}{2} V_p^2 y'_{pq} \cos \theta'_{pq}] + j[V_p^2 y_{pq} \sin \theta_{pq} \\ & - V_p V_q y_{pq} \sin(-\delta_p + \delta_q + \theta_{pq}) + \frac{1}{2} V_p^2 y'_{pq} \sin \theta'_{pq}] \end{aligned} \quad (A.4-3)$$

Transmission lines have much larger reactance and capacitance than resistance, and therefore the above equation can be simplified by assuming

$$\begin{aligned} \theta_{pq} & \approx -90^\circ \\ \theta'_{pq} & \approx 90^\circ \end{aligned} \quad (A.4-4)$$

and dividing $\bar{S}_{pq}$ into its real and reactive parts,

$$\begin{aligned} \bar{S}_{pq} & = P_{pq} + j Q_{pq} \\ P_{pq} & = V_p V_q y_{pq} \sin(\delta_p - \delta_q) \\ Q_{pq} & = -V_p^2 y_{pq} - V_p V_q y_{pq} \cos(\delta_p - \delta_q) + \frac{1}{2} V_p^2 y'_{pq} \end{aligned} \quad (A.4-5)$$

The quantities P_{pq} and Q_{pq} represent real and reactive power flows, respectively, measured from the bus p -side, and appear in Eq. A.4-1.

Measurement equations associated with the long-term dynamics model (Eq. A.2-38) are formulated as follows.

$$y_1 = I_{CH1}$$

$$y_2 = \psi_{TH1}$$

$$y_3 = P_{GV1}$$

$$y_4 = \psi_{HP1} = \frac{w_{HP1}}{K_{HP1}}$$

$$y_5 = P_{m1} = F_{HP1} w_{HP1} + F_{IP1} w_{RH1} + F_{LP1} w_{CO1}$$

$$y_6 = V_1$$

$$y_7 = w_{SYS} \quad (A.4-6)$$

$$y_8 = E_{FD1}$$

$$y_9 = V_{Lj}$$

$$y_{10} = V_p V_q y_{pq} \sin(\delta_p - \delta_q)$$

$$y_{11} = -V_p^2 y_{pq} - V_p V_q y_{pq} \cos(\delta_p - \delta_q) + \frac{1}{2} V_p^2 y'_{pq}$$

$$i = 1, 2, \dots, n$$

$$j = n+1, n+2, \dots, n+r$$

$$p, q = 1, 2, \dots, n+r$$

where

I_{CH1} = steam charging signal

ψ_{TH1} = throttle pressure

P_{GV1} = governor valve position

ψ_{HP1} = first stage steam pressure

w_{HP1} = turbine first-stage steam flow rate

K_{HP1} = equivalent reaction turbine flow coefficient.

The above measurement equations are either duplicates of Eq. A.4-1 or are obvious.

Measurement equations for transient and long-term dynamic studies have been given in Eqs. A.4-1 and A.4-6, each of which has the general form

$$y = h(x, p) \quad (A.4-7)$$

where

x = state vector

p = network variable vector

Considering instrumentation errors, the above equation can be further generalized,

$$y = h(x, p) + \tilde{v} \quad (A.4-8)$$

where $\tilde{v}$ represents an additive noise vector. This may be considered as a Gaussian white noise vector; additional instrumentation errors may be present -- biases, scale factor errors, etc -- which can be handled by augmenting the state vector x to include the error parameters.

A.5 DYNAMIC EQUIVALENCING

A.5.1 Motivation and Overview of Techniques

A dynamic equivalent is a reduced order differential equation model of the dynamic characteristics of a power system. The usual application of a dynamic equivalent is to model those portions of an interconnected power system which are beyond the boundaries of the system under investigation. For the security assessment problem it is possible (and desirable) to divide the large scale power system into two parts: the study system and the external system as shown in Fig. A.5-1. The study system (e.g., a single utility or power pool) is the one for which disturbances are to be examined and dynamic security assessment performed. The external system is of interest only because of its effect on the dynamic security of the study system. It is therefore desirable to reduce the representation of the external system to as low an order as possible while maintaining the essential effects upon the study system.

Figure A.5-1 Study and External Systems

At present there are two commonly used techniques for constructing dynamic equivalents: modal analysis and the coherent machine method. The modal analysis technique identifies a dynamic equivalent by determining the natural modes of the external system and eliminating the state variables corresponding to the modes which are deemed to have a negligible effect on the study system. It generally involves the following steps (Ref. 94):

- Construction of a comprehensive linear differential equation describing the external system

- Separation of natural modes of response by transforming the equations into a simplified canonical form
- Reduction of the order of the canonical form equations by the deletion of selected natural modes
- Integration of the reduced linear differential equations, in conjunction with the nonlinear differential equations of the study system, to provide a simulation of the study system as it is affected by the external system.

The coherent machine technique reduces the order by lumping together groups of machines in the external system which tend to swing together when perturbed by disturbances in the study system. Use of the technique involves the following steps (Ref. 95):

- Modeling an equivalent generating unit, with the individual mechanical powers replaced by the total mechanical power, and the electrical powers by the total electrical power output for each group of coherent generators
- Specification of the characteristics of the equivalent mode, given the model of each individual unit, by considering separately the rotor dynamics, the governor and turbine model, the synchronous machine model, the excitation system model, and the power system stabilizer model
- Adjustment of the linear parameters of each equivalent model in order to obtain a minimal error between its transfer function and the sum of the transfer functions of the individual units. The error to be minimized is the sum of the squares of the magnitudes of the relative differences, for specified discrete frequencies.

The above techniques assume that sufficient information is available to construct a complete model of both the study and external systems. The external system is then reduced to an equivalent. Therefore, these approaches are not adequate for use in an on-line, operational mode in which dynamic security assessment is performed. In dynamic security assessment, an up-to-the-minute equivalent of neighboring systems is desirable for use in the analysis of possible disturbances or contingencies in the study system, under the circumstances that enough information about the current network configuration and state of neighboring systems is not available. The system identification technique is appropriate for that problem since it determines the essential characteristics of a dynamic system by observing the response of system variables to random system inputs, either natural or intentional.

Price, et al., (Ref 96) employed the maximum likelihood method, and Ibrahim, et al., (Ref 97), used the recursive least-mean-square method to study the system identification technique. The general procedure of both methods is:

- Formulate a hypothetical structure to represent the external system

- Construct a filter (Kalman filter for the maximum likelihood method and least-mean-square filter for the recursive least-mean-square method) to process the observed data
- Estimate parameters of the hypothetical structure.

A.5.2 Selected Technique

Based upon a review of previous work on equivalencing techniques, the most applicable approach for use in dynamic security assessment is the modal analysis technique coupled with maximum likelihood parameter identification. A low order dynamic model for the external system is constructed using the modal analysis technique and a nominal representation of the external system structure. One method for permuting and partitioning a state vector according to its natural modes response is discussed in Section 4.3. The maximum likelihood technique is used to update estimates of the model parameters continually, based upon measurements taken within the study system and at the external system/study system interface. This process is depicted schematically in Fig. A.5-2 which was adapted from Ref 96. The remainder of this section contains a summary of the maximum likelihood identification technique proposed in Ref 96

Figure A.5-2 On-Line Maximum Likelihood Identification Technique (see Ref. 96)

Let Y_N denote the set of all measurements made on the power system up to time N . The likelihood function is the conditional probability of obtaining the observed measurement sequence from the assumed system model with parameters α . The parameter vector (α) contains the coefficients of the external system model (dynamic equivalent to be estimated from on-line measurements). In essence, the maximum likelihood technique generates estimates ($\hat{\alpha}$) which maximize the conditional probability density $p(Y_N|\alpha)$. For computational reasons, it is more convenient to maximize the log likelihood function $\xi(N|\alpha) = \ln p(Y_N|\alpha)$. Since the logarithm is a monotonic function,

$$\max_{\alpha} \xi(N|\alpha) = \max_{\alpha} p(Y_N|\alpha) \quad (A.5-1)$$

In the following development (which follows Refs 96 and 98), it is assumed that the system model is linear (or linearized), time-invariant and driven by a zero-mean gaussian process. These assumptions result in an analytically tractable form for the likelihood function; specifically, $p(Y_N|\alpha)$ is Gaussian. Although the detailed power system model is nonlinear, nominal

system operation (i.e., in the absence of severe faults) can be adequately characterized by a linearized model.

Since the on-line parameter identification technique is implemented in conjunction with a state estimator (see Chapter 2), it is desirable to have a recursive expression for the log likelihood function. From Bayes' rule,

$$p(Y_N|\underline{\alpha}) = p(Y_{N-1}, y_N|\underline{\alpha}) = p(y_N|Y_{N-1}, \underline{\alpha}) p(Y_{N-1}|\underline{\alpha}) \quad (A.5-2)$$

and

$$\xi(N|\underline{\alpha}) = \xi(N-1|\underline{\alpha}) + \ln[p(y_N|Y_{N-1}, \underline{\alpha})] \quad (A.5-3)$$

which is the desired relation. This expression is true in general, but it takes on a particularly useful form if $y(n)$ is assumed to be Gaussian. Define

$\hat{y}_n(-|\underline{\alpha})$ = the conditional expectation of $y(n)$ given $y(1), \dots, y(n-1)$ and $\underline{\alpha}$.

$$\delta_y(n|\underline{\alpha}) = y(n) - \hat{y}_n(-|\underline{\alpha})$$

$$P_{y_n}(-|\underline{\alpha}) = E\{\delta_y(n|\underline{\alpha}) \delta_y^T(n|\underline{\alpha})\}$$

Then, the log likelihood function has the form

$$2\xi(N|\underline{\alpha}) = \xi_{BIAS}(N|\underline{\alpha}) + \xi_{OBS}(N|\underline{\alpha}) \quad (A.5-4)$$

with

$$\xi_{BIAS}(N|\underline{\alpha}) = -m N \ln(2\pi) - \sum_{n=1}^N \ln |P_{y_n}(-|\underline{\alpha})| \quad (A.5-5)$$

$$\xi_{OBS}(N|\underline{\alpha}) = - \sum_{n=1}^N \delta_y^T(n|\underline{\alpha}) P_{y_n}^{-1}(-|\underline{\alpha}) \delta_y(n|\underline{\alpha}) \quad (A.5-6)$$

and m = dimension of y . Note that the term subscripted "BIAS" is independent of the observations if the system is linear.

The basic procedure for implementing the identification algorithm is as follows:

- The on-line measurements are processed in the state estimator with the residuals (δ_y) and measurement covariance matrix (P_{y_n}) computed and used to calculate the log likelihood function via Eqs. A.5-5 and A.5-6

- After N measurements have been processed, the likelihood function is optimized. Price et al suggest using an approximate Newton method
- The values of the parameter vector $\underline{\alpha}$ which maximize the log likelihood function are chosen as the new estimates and the procedure is repeated as desired. To track slowly varying changes in the external system, this procedure would be repeated at selected intervals

A.6 SPATIAL DECOMPOSITION OF POWER SYSTEM MODELS

One of the difficulties in power system security assessment is the dimensionality of models. The dimension of a power system model increases rapidly with the number of buses, particularly with the number of generators. The power system model discussed in Section A.2, for example, has 14 equations governing the dynamics associated with each generator and two equations governing the dynamics associated with each load bus. A typical system considered for security assessment may be a power pool consisting of several regional power companies that are electrically interconnected. Since each company may have many generating plants and distribution centers it is apparent that the model for the entire power pool may be very large and computationally cumbersome.

It is computationally cheaper to analyze two systems of orders m and n than to analyze the composite system of order $m+n$. Conversely, computational economics can be achieved by decomposing a large power system into two or more smaller systems with few or weak interconnections. One way to do this is via spatial decomposition -- partitioning a system by taking advantage of the topological properties of its structure. Spatial decomposition makes use of information about the location of generating plants and load centers, the interconnections between physically distinct components and the distances between them.

A power system can be represented topologically by a network of nodes and branches in which the nodes represent buses and the branches represent transmission lines. The spatial decomposition discussed in this section is restricted to partitioning the network by cutting through nodes. The number of possible cut-node decompositions for an n -node network is of the order of $n!$, which is a large number even for moderate n . In order to limit the number of decompositions that may be of value, the problem of generating partitions by node-cutting is posed as a constrained optimization problem.

The problem of optimal network decomposition has the objective of minimizing the coupling between the resulting partitions. The constraints of the problem are the maximum number of subsystems desired and the maximum numbers of generator and load buses that each partition may have. The problem formulated in Ref. 99 seeks a solution that minimizes the number of cut-nodes for a prescribed number of partitions while constraining the size of each sub-network. The solution technique uses methods of dynamic programming and has been incorporated in a computer program.

An illustration of an optimal network decomposition is shown in Figs. A.6-1 and A.6-2. The first figure shows the IEEE 39-bus New England Test System which has 39 buses, 46 lines and transformers, 9 generators and 20 loads. The computer program was used

Figure A.6-1 One-Line Diagram of the IEEE 39 Bus Test System

Figure A.6-2 Decomposed Network Structure of the IEEE 39 Bus Test System

to partition this network into six subsystems each having at least one generator and at most ten buses. The resulting optimal decomposition, shown in Fig. A.6-2 has six cut-nodes.

APPENDIX B

AN INTEGRATED STATE SPACE FORMULATION OF THE POWER SYSTEM EQUATIONS

B.1 INTRODUCTION

In Appendix A, detailed dynamical models are developed to describe the behavior of both synchronous machines and their associated prime movers. These models are written in the useful state space form

$$d\tilde{x} = \tilde{f}(\tilde{x}, u, p) dt \quad (B.1-1)$$

where the state vector $\tilde{x}$ contains all the dynamical variables of interest. In addition to these equations, a multimachine system is governed by a set of nonlinear algebraic equations -- the network equations -- which account for the interconnections in the power transmission system. These algebraic equations relate the generator and prime mover variables ($\tilde{x}$) to variables (p) characterizing the electrical network (specifically, complex bus voltages). They are of the form

$$g(\tilde{x}, u, p) = 0 \quad (B.1-2)$$

In dynamic security assessment studies, it is more convenient to deal with a set of integrated state space equations

$$dx = f(x, u) dt \quad (B.1-3)$$

having an augmented state vector

$$x = \begin{bmatrix} p \\ \tilde{x} \end{bmatrix} \quad (B.1-4)$$

rather than the coupled set of differential and algebraic equations. This appendix contains the development of one such set of combined equations.

Of primary interest for state estimation are the complex voltage (typically expressed as magnitude and angle) and frequency offset at the power system buses. The approach taken below starts with the desired network variable (e.g., voltage angle at a generator bus) and, where possible, relates it to one (or more) of the dynamical variables ($\tilde{x}$) through an algebraic or differential equation.

Throughout this development, the algebraic constraint equations are used repeatedly. For convenience they are repeated below:

$$P_1 - \sum_{k=1}^N Y_{1k} V_1 V_k \cos[\theta_{1k} - (\delta_1 - \delta_k)] = 0 \quad (B.1-5a)$$

$$Q_1 + \sum_{k=1}^N Y_{1k} V_1 V_k \sin[\theta_{1k} - (\delta_1 - \delta_k)] = 0$$

$i = 1, 2, \dots, N = n+r$

(B.1-5b)

B.2 FORMULATION OF THE DYNAMIC EQUATIONS

The first variable to be considered is the voltage angle at a generator bus, δ_1 . (In this appendix, a variable with subscript 1 is associated with a generator bus. The subscript j denotes a load bus.) A common assumption in power system analyses is that the voltage angle (δ_1) is approximately equal to the synchronous machine rotor angle (δ_r) when it is expressed in electrical radians relative to a synchronously rotating reference. The validity of this assumption is dependent upon the magnitude of disturbances to the electrical network and the time-periods of interest. Under nominal operating conditions (i.e., small rather than large disturbances) and for time-periods commensurate with the long-term dynamics model (i.e., a few seconds or longer), this assumption is reasonable.

Now, the rotor angle at the 1th bus (and by assumption the voltage angle) satisfies the so-called swing equation (Ref. 37)

$$M_1 \ddot{\delta}_1 + D_1 \dot{\delta}_1 = P_{m_1} - P_{e_1} \quad (B.2-1)$$

where

δ_1 = voltage (equivalently rotor) angle

$M_1 = H_1 / (\pi f^0)$, where H_1 is the per-unit inertia constant and f^0 is the nominal system frequency

D_1 = damping coefficient

P_{m_1} = mechanical input power from the turbine (cf eqs. A.2-4 and A.2-5)

$$P_{e_1} = \sum_{k=1}^N Y_{1k} V_1 V_k \cos[\theta_{1k} - (\delta_1 - \delta_k)]$$

= electrical power transported through the network.

Note that if the system is in equilibrium, $P_{m_1} = P_{e_1}$ and $\dot{\delta}_1 = 0$; thus, the generator voltage angle is constant (with respect to the synchronous reference). If P_{e_1} should change (e.g., due to an increase or decrease in load), the difference ($P_{m_1} - P_{e_1}$) is nonzero and the rotor (and thus voltage angle) will increase or decrease according to Eq. B.2-1. In addition, changes in system frequency (i.e., δ_1) are sensed and used to drive P_{m_1} toward P_{e_1} ; equations describing this response are developed in Section A.3.1.

The excitation of Eq. B.2-1 arises from unpredictable variations in the (real) load P . Assume that the load at the j th bus changes by the amount ΔP_j . The resulting imbalance between total generation and load excites transient behavior throughout the network. These transients are the response of individual generators to the power imbalance at each of the generator buses. Thus, it is necessary to determine the

effects of the load changes ΔP_j on each of the P_{e_i} . The first part of the following treatment is similar to one presented in Anderson and Fouad (Ref. 37).

For a small change in load power, the nonlinear constraint equations may be linearized about the current operating point $\underline{V}^0, \underline{\delta}^0$. If Eq. B.1-5a is expressed in vector form as

$$\underline{P} = \underline{g}_P(\underline{V}, \underline{\delta}) \quad (\text{B.2-2})$$

with $\underline{P} = (P_1, \dots, P_N)$, $\underline{V} = (V_1, \dots, V_N)^T$, $\underline{\delta} = (\delta_1, \dots, \delta_N)^T$ and

$$g_{P_i}(\underline{V}, \underline{\delta}) = \sum_{k=1}^N Y_{ik} V_i V_k \cos[\theta_{ik} - (\delta_i - \delta_k)] \quad i=1, \dots, N \quad (\text{B.2-3})$$

Then, by defining the matrices G_V^P and G_δ^P as

$$(G_V^P)_{ij} = \left. \frac{\partial g_{P_i}}{\partial V_j} \right|_{\underline{V}^0, \underline{\delta}^0} \quad (\text{B.2-4})$$

$$(G_\delta^P)_{ij} = \left. \frac{\partial g_{P_i}}{\partial \delta_j} \right|_{\underline{V}^0, \underline{\delta}^0} \quad (\text{B.2-5})$$

$$\underline{P}^0 + \Delta \underline{P} = \underline{g}_P(\underline{V}^0, \underline{\delta}^0) + G_V^P \Delta \underline{V} + G_\delta^P \Delta \underline{\delta} + \dots \quad (\text{B.2-6})$$

from which

$$\Delta \underline{P} = G_V^P \Delta \underline{V} + G_\delta^P \Delta \underline{\delta} \quad (\text{B.2-7})$$

to first order. At this point the following assumptions will be made:

- The voltage magnitudes at all of the system buses remain approximately constant* (i.e., $\Delta V=0$), hence

$$\Delta \underline{P} = \begin{bmatrix} \Delta P_G \\ \Delta P_L \end{bmatrix} = G_\delta^P \Delta \underline{\delta} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} \Delta \underline{\delta}_G \\ \Delta \underline{\delta}_L \end{bmatrix} \quad (\text{B.2-8})$$

with the partitions corresponding to generator and load buses.

- The voltage angles at the generator buses $(\delta_1, \dots, \delta_n)$ do not change instantaneously because of the nonzero rotor inertia.

*This is in effect a P- δ , Q-V decoupling of the power system dynamic response.

Under these assumptions, it follows that the instantaneous change in load bus voltage angle is related to a change in load power via

$$\Delta \underline{P}_L = D \Delta \underline{\delta}_L \quad (\text{B.2-9})$$

Equations B.2-8 and B.2-9 may be combined to eliminate $\Delta \underline{\delta}_L$ if D^{-1} exists, viz.,

$$\Delta \underline{P}_G = A \Delta \underline{\delta}_G + BD^{-1} \Delta \underline{P}_L \quad (\text{B.2-10})$$

Since $P_{e_i} = (P_G)_i$, it follows that

$$P_{e_i}(t) = (A \underline{\delta}_G + BD^{-1} \underline{P}_L(t))_i \quad (\text{B.2-11})$$

for values of $\underline{P}_L(t)$ near $\underline{P}_L^0(t)$, where $\underline{P}_L^0(t)$ is the load power which satisfies the equality

$$\begin{aligned} P_{e_i}^0 &= \sum_{k=1}^N Y_{ik} V_i^0 V_k^0 \cos[\theta_{ik} - (\delta_i^0 - \delta_k^0)] \\ &= (A \underline{\delta}_G^0 + BD^{-1} \underline{P}_L^0(t))_i \end{aligned} \quad (\text{B.2-12})$$

It is important to realize that the matrices, A, B, and D are evaluated at the point $\underline{V}^0, \underline{\delta}^0$ and thus Eq. B.2-11 holds for values of $\underline{V}$ and $\underline{\delta}$ (and consequently $\underline{P}_L(t)$) near $\underline{V}^0, \underline{\delta}^0$. Substituting Eq. B.2-11 into Eq. B.2-1 yields

$$M_i \ddot{\delta}_i + D_i \dot{\delta}_i = P_{m_i} - (A \underline{\delta}_G + BD^{-1} \underline{P}_L(t))_i \quad (\text{B.2-13})$$

$$i=1, \dots, n$$

This equation may be interpreted in the following manner. If the load does not change, viz.,

$$\Delta \underline{P}_L = \underline{P}_L(t) - \underline{P}_L^0 = 0 \quad (\text{B.2-14})$$

and if the system is in equilibrium (i.e., $P_{m_i} = P_{e_i}$), the rotor angle δ_i does not change. However, if the load changes by the amount $\Delta \underline{P}_L$, each of the generator rotor angles is accelerated in proportion to the apparent load change $(BD^{-1} \Delta \underline{P}_L)_i$.

Before proceeding a few comments on Eq. B.2-9 are necessary. In the above development it is assumed that the voltage angle at a load would change much more rapidly than the angle at a generator bus. For a purely resistive load this indeed is the case. However, most loads are not purely resistive but contain numerous inductive devices (e.g., motors). Consequently, a jump change in load power $\Delta \underline{P}_L$ does not result in an instantaneous change in $\underline{\delta}_L$ of the magnitude predicted by Eq. B.2-9.

The actual situation is closer to the one depicted in Fig. B.2-1. (This figure illustrates the response of the voltage angle at the j^{th} load bus due to the effective load change at that bus.) At the instant t_0 when the load changes by $\Delta P_j = (D^{-1} \Delta P_L)_j$ there is a step change in the voltage angle δ_j in the amount $\delta_j^- - \delta_j^+$. This is followed by a more gradual decay (possibly oscillatory) from δ_j^+ to δ_j^f where

$$\delta_j^f = \delta_j^- - (D^{-1} \Delta P_L)_j \quad (\text{B.2-15})$$

The decay time constant τ is a function of the actual impedance of the load. If the decay time constants at the load buses are sufficiently small, the voltage angle decays to δ_j^f before the generator rotor angle (which would normally have much higher inertias) could change. In this case the assumptions leading to Eq. B.2-8 are still valid and Eq. B.2-10 is a good approximation

Figure B.2-1 Approximate Response of Load Bus Voltage Angle to a Load Change

If the decay in Fig. B.2-1 can be approximated by an exponential then δ_j^d may be written as the sum of a discontinuous part δ_j^d and a continuous part δ_j^c , with

$$\delta_j^d(t) = \begin{cases} \delta_j^-(t) & t = t_0^- \\ \delta_j^+(t) & t = t_0^+ \end{cases} \quad (\text{B.2-16})$$

with $\delta_j^c(t)$ governed by the differential equation

$$\dot{\delta}_j^c(t) = -\frac{1}{\tau} [\delta_j^c(t) - \delta_j^f], \quad \delta_j^c(t_0^+) = \delta_j^+ \quad (\text{B.2-17})$$

It should be noted that the intermediate quantity δ_j^+ must be determined empirically since it is dependent upon the characteristics of the load.

Equations B.2-13, B.2-16 and B.2-17 are the desired equations relating the network variables (δ) to the dynamical system model (via the state variable P_m). To specify the complex voltage completely at the system buses, it is necessary to develop analogous expressions for voltage magnitude. The following paragraphs outline the development

In Section A.4, equations were presented which describe the behavior of the excitation system (or

voltage regulator). At each generator bus, the voltage regulator attempts to maintain V_1 at its setpoint V_{REF} via a feedback control loop. The long-term dynamic model incorporates the reasonable assumption that $V_1 \approx V_{\text{REF}}$ for the time periods of interest. However, in this development it is necessary to include the dynamical behavior of the voltage regulator explicitly in the model. On the average $V_1 \approx V_{\text{REF}}$ due to the influence of the voltage regulator. However, at any instant, V_1 will differ from V_{REF} as the effects of load changes are reflected through the network.

At this point it would be desirable to express the voltage change at the i^{th} generator bus (ΔV_i) in terms of a load change at the j^{th} load bus. Unfortunately, a load change at the j^{th} bus results in a complicated readjustment of the voltage magnitudes at all of the system buses such that a reactive power balance is maintained. To proceed, it is necessary to solve Eq. B.1-5b for the bus voltage changes ΔV_i , $i = 1, 2, \dots, N$ that result from a small change in load reactive power. Using the same techniques as before, Eq. B.1-5b may be written in vector form as

$$Q = g_Q(V, \delta) \quad (\text{B.2-18})$$

with

$$g_{Q_1}(V, \delta) = - \sum_{k=1}^N Y_{1k} V_1 V_k \sin[\theta_{1k} - (\delta_1 - \delta_k)] \quad i = 1, \dots, N \quad (\text{B.2-19})$$

For a small change in reactive power ΔQ , Eq. B.2-18 may be written as

$$Q^o + \Delta Q = g_Q^o(V^o, \delta^o) + G_V^Q \Delta V + G_\delta^Q \Delta \delta + \dots \quad (\text{B.2-20})$$

where the superscript o denotes values which satisfy Eq. B.2-18 and the matrices G_V^Q and G_δ^Q are defined as

$$(G_V^Q)_{1j} = \left. \frac{\partial g_{Q_1}}{\partial V_j} \right|_{V^o, \delta^o} \quad (\text{B.2-21})$$

$$(G_\delta^Q)_{1j} = \left. \frac{\partial g_{Q_1}}{\partial \delta_j} \right|_{V^o, \delta^o} \quad (\text{B.2-22})$$

Under the assumptions that the reactive power at a generator bus cannot change instantaneously* (i.e., $\Delta Q_1 = \dots = \Delta Q_n = 0$) and that the δ_1 are constant

*Reactive power balance is maintained through use of the exciter which is modeled by Eq. B.2-24.

(since changes in reactive power have little effect on system frequency), Eq. B.2-20 can be solved to yield

$$\Delta \underline{V} = G_V^{-1} \tilde{\Delta Q} \quad (B.2-23)$$

where

$$\tilde{\Delta Q} = (0, \dots, 0, \Delta Q_{n+1}, \dots, \Delta Q_{n+r})^T = (0, \dots, \Delta Q_L^T)^T$$

It is now possible to express the generator bus voltage magnitude as a superposition of two terms -- the perturbation described by Eq. B.2-23 which tends to drive V_1 away from V_{REF} and the restorative effects of the voltage regulator. The voltage regulator is governed by an ordinary differential equation of the form

$$\frac{dV_1}{dt} = f_{VR}^1(V_1, V_{REF}) \quad (B.2-24)$$

which is described in Section A.4. The load changes act as a driving term, thus by letting the increment $\tilde{\Delta Q}$ become a differential $d\tilde{Q}$, Eq. B.2-23 may be added to Eq. B.2-22 to yield the desired equation for the i^{th} generator bus, viz ,

$$dV_1 = f_{VR}^1(V_1, V_{REF})dt + (G_V^{-1} d\tilde{Q})_1 \quad (B.2-25)$$

$i = 1, 2, \dots, n$

Although the preceding development was focussed on describing the behavior of the voltage magnitude at a generator bus, it is necessary to solve for the voltage change at all of the system buses. Thus, Eq. B.2-24 also provides the relationship between a load change and load bus voltage variations. Since there is no direct mechanism for maintaining the load bus voltage magnitude at a predetermined value, the magnitude of the load bus voltage tends to vary proportionally to changes in reactive power. Specifically, if $d\tilde{Q}$ were "noiselike," then $dV_{n+1}, \dots, dV_{n+r}$ would behave in a similar manner. This has obvious implications for the state estimation problem.

This completes the development of an integrated set of dynamical equations for a multimachine power system. The combined system state vector contains voltage magnitude and angle for each system bus and all of the relevant dynamical variables described in Appendix A. In the next section the dynamical model is completed with the addition of process noise.

B.3 INCORPORATION OF PROCESS NOISE

The equations developed in the previous section are driven by changes in the real and reactive power at the load buses. In Section A.3, a model for load disturbances of the form

$$\varepsilon_L(t) = v(t) + n_p(t) \quad (B.3-1)$$

is postulated. The remainder of this appendix illustrates how this model for load changes can be integrated with the dynamic system model.

From Section A.3, the actual real and reactive loads take the form

$$P_j(t) = P_j^f(t) + v_j^P(t) + n_{P_j}^P(t) \quad (B.3-2)$$

$$Q_j(t) = Q_j^f(t) + v_j^Q(t) + n_{P_j}^Q(t) \quad (B.3-3)$$

$$j = n+1, \dots, n+r$$

where

P_j^f, Q_j^f are the forecasted (i.e., modeled) real and reactive load powers at the j^{th} bus

v_j^P, v_j^Q are the a.s. continuous sample path components of the difference between the actual and forecasted real and reactive powers at the j^{th} bus

$n_{P_j}^P, n_{P_j}^Q$ are the a.s. discontinuous sample path components of the difference between the actual and forecasted real and reactive powers at the j^{th} bus.

In the following paragraphs, Eqs. B.3-2 and B.3-3 are combined with Eqs. B.2-13, B.2-16, B.2-17 and B.2-25 which are repeated below (in a slightly modified form) for convenience:

$$d(\delta_1) + \frac{D_1}{M_1} \delta_1 dt = \frac{1}{M_1} [P_{m_1} - (A\delta_G + BD^{-1} P_L(t))_1] dt \quad (B.3-4)$$

$i=1, \dots, n$

$$\delta_j^d(t) = \begin{cases} \delta_j^-(t) & t = t_0^- \\ \delta_j^+(t) & t = t_0^+ \end{cases} \quad (B.3-5)$$

$$d\delta_j^c = -\frac{1}{T} [\delta_j^c(t) - \delta_j^f] dt, \quad \delta_j^c(t_0^+) = \delta_j^+ \quad (B.3-6)$$

$j=n+1, \dots, n+r$

$$dV_1 = f_{VR}^1(V_1, V_{REF})dt + (G_V^{-1} d\tilde{Q}_L)_1 \quad (B.3-7)$$

$i=1, \dots, n$

$$dV_j = (G_V^{-1} d\tilde{Q}_L)_j \quad j=n+1, \dots, n+r \quad (B.3-8)$$

Consider first Eq. B.3-4; substituting Eq. B.3-2 yields

$$\begin{aligned} d(\delta_1) + \frac{D_1}{M_1} \delta_1 + \frac{A_{1i}}{M_1} \delta_1 dt \\ = \frac{1}{M_1} [P_{m_1} - (\tilde{A}\delta_G + BD^{-1}(P_L^f + v_L^P + n_P^P))_1] dt \end{aligned} \quad (B.3-9)$$

where $\tilde{A}$ is equal to A with A_{11} set to zero. Since $d\delta_1 = \omega_1 dt$ (equivalently, $\dot{\delta}_1 = \omega_1$), it is possible to associate a state variable representation with Eq. B.3-9. Specifically,

$$d \begin{bmatrix} \delta_1 \\ \omega_1 \\ \eta_1 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{A_{11}}{M_1} & -\frac{D_1}{M_1} & -\frac{1}{M_1} \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_1 \\ \omega_1 \\ \eta_1 \end{bmatrix} dt + \frac{1}{M_1} \begin{bmatrix} 0 \\ P_{m_1} - (\tilde{A} \delta_G + BD^{-1} P_L^f)_1 \\ 0 \end{bmatrix} dt + d \begin{bmatrix} 0 \\ 0 \\ (BD^{-1}(\underline{v}^P + \underline{n}_P))_1 \end{bmatrix} \quad (B.3-10)$$

which is of the form

$$d\underline{x} = \underline{f}(\underline{x}, \underline{u})dt + \underline{\Gamma} d\underline{n} \quad (B.3-11)$$

The control vector $\underline{u}$ which is implicit in Eq. B.3-10 enters through the turbine mechanical power P_{m_1} .

Equations B.3-5 and B.3-6 are slightly more complicated. The response of the load bus voltage angle consists of two parts. The first part is the immediate change in angle which reflects the discontinuous behavior of the load, viz.,

$$d\delta_j^d = \delta_j^+ - \delta_j^- = \alpha d[(D^{-1}(\underline{P}_L^f + \underline{v}^P + \underline{n}_P))_j] \\ = \alpha [D^{-1}(d\underline{P}_L^f + d\underline{v}^P + d\underline{n}_P)]_j \quad (B.3-12)$$

where $0 \leq \alpha \leq 1$ is an empirically determined coefficient. The continuous portion of the response is modeled by Eq. B.3-6. The term δ_j^f in Eq. B.3-6 may be expressed as a function of the load power if the change

$$\Delta \underline{P}_L(t) = \underline{P}_L(t) - \underline{P}_L^0$$

is not so large as to invalidate the first-order approximation of Eq. B.2-6; specifically,

$$\delta_j^f(t) = \delta_j^0 + (D^{-1} \Delta \underline{P}_L(t))_j \\ = (D^{-1} \underline{P}_L(t))_j \quad (B.3-13)$$

Combining the discontinuous and continuous portions of the response yields

$$d\delta_j(t) = \alpha [D^{-1} d\underline{P}_L^d(t)]_j - \frac{1}{\tau} [\delta_j(t) - (D^{-1} \underline{P}_L^c(t))_j] dt \\ j=n+1, \dots, n+r \quad (B.3-14)$$

where $\underline{P}_L^d(t)$ and $\underline{P}_L^c(t)$ are discontinuous and continuous components of the load. Since $\underline{v}^P(t)$ is a continuous and $\underline{n}_P(t)$ is a discontinuous,

$$\underline{P}_L^d(t) = \underline{P}_L^{fd}(t) + \underline{n}_P^P(t) \quad (B.3-15)$$

$$\underline{P}_L^c(t) = \underline{P}_L^{fc}(t) + \underline{v}^P(t) \quad (B.3-16)$$

Equation B.3-14 can be written in state space form as

$$d \begin{bmatrix} \delta_j \\ \eta_j^c \end{bmatrix} = \begin{bmatrix} -1/\tau & 1/\tau \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \delta_j \\ \eta_j^c \end{bmatrix} dt + d \begin{bmatrix} \alpha (D^{-1}(\underline{P}_L^{fd} + \underline{n}_P))_j \\ (D^{-1}(\underline{P}_L^{fc} + \underline{v}^P))_j \end{bmatrix} \\ j=n+1, \dots, n+r \quad (B.3-17)$$

The final variables to be discussed are the voltage magnitudes. Defining

$$f_{VR}^k(V_k, V_{REF}) = \begin{cases} f_{VR}^k(V_k, V_{REF}) & k=1, \dots, n \\ 0 & k=n+1, \dots, n+r \end{cases} \quad (B.3-18)$$

allows Eqs. B.3-7 and B.3-8 to be written compactly as

$$dV_k = f_{VR}^k(V_k, V_{REF})dt + [G_V^{-1}(dQ_L^f + d\underline{v}^Q + d\underline{n}_P^Q)]_k \quad (B.3-19) \\ k=1, \dots, n+r$$

which is already of the desired form.

Equations B.3-10, B.3-17 and B.3-19 constitute the integrated power system equations. The general form of these equations includes the effects of a control input $\underline{u}$. Throughout this report, the control input is dropped from the system equations as is explicit consideration of the forecasted loads $\underline{P}_L^f$ and $\underline{Q}_L^f$. The equations developed herein and in Appendix A can be brought into this alternate form by assuming that the control inputs are chosen such that the mechanical power P_m and generator bus voltages match the forecasted real and reactive loads. The closed-loop response of the power system to the disturbance inputs $\underline{v} + \underline{n}_P$ is still governed by equations of the form discussed above.

APPENDIX C
POISSON COUNTING PROCESSES

This appendix lists the relevant properties of the Poisson counting process necessary for this report. This material is presented without proof and the reader is directed to Snyder (Ref 12) for additional details

A point process is a mathematical model for highly localized events which may occur at random points in a continuum. In this report, it is sufficient to treat the continuum as time, i.e., the real line parameterized by the variable $t \geq 0$. For example, lightning discharges observed at times $w_1, w_2, \dots$ (with $0 \leq w_1 \leq w_2 \leq \dots < T$) are a point process

A counting process $N(t)$ is a nonnegative, integer valued random process which indicates the number of points which occur in an interval of time. Introducing the indicator function

$$I_{\{t \geq T\}} = \begin{cases} 0 & t < T \\ 1 & t \geq T \end{cases} \quad (C-1)$$

allows $N(t)$ to be written as

$$N(t) = \sum_1 I_{\{t \geq w_1\}} \quad (C-2)$$

with $0 \leq w_1 \leq w_2 \leq \dots$ an increasing sequence of random variables

A Poisson counting process is a counting process which satisfies the following properties:

- $\Pr\{N(0) = 0\} = 1$,
- Let $\Lambda(t) = \int_0^t \lambda(\sigma) d\sigma$ where $\lambda(\sigma) \geq 0$ is termed the intensity of the process. For $0 \leq s < t$, the increment $N(s, t) = N(t) - N(s)$ is Poisson distributed with parameter $\Lambda(t) - \Lambda(s)$, i.e.,

$$\Pr\{N(s, t) = n\} = \frac{[\Lambda(t) - \Lambda(s)]^n}{n!} e^{-[\Lambda(t) - \Lambda(s)]} \quad (C-3)$$

for $n=0, 1, 2, \dots$,

- $N(t)$, $t \geq 0$ has independent increments, i.e., if $[s_1, t_1], [s_2, t_2], \dots, [s_k, t_k]$ are k disjoint intervals on $[0, \infty)$ then

$$\Pr\{N(s_1, t_1) = n_1, N(s_2, t_2) = n_2, \dots, N(s_k, t_k) = n_k\} = \prod_{i=1}^k \Pr\{N(s_i, t_i) = n_i\}. \quad (C-4)$$

For a Poisson counting process the expected number of points in the interval $[s, t)$ may be evaluated as

$$\begin{aligned} E\{N(s, t)\} &= \sum_{n=0}^{\infty} n \cdot \Pr\{N(s, t) = n\} \\ &= \sum_{n=0}^{\infty} n \frac{[\Lambda(t) - \Lambda(s)]^n}{n!} e^{-[\Lambda(t) - \Lambda(s)]} \\ &= \sum_{n=1}^{\infty} n \frac{[\Lambda(t) - \Lambda(s)]^n}{n!} e^{-[\Lambda(t) - \Lambda(s)]} \\ &= [\Lambda(t) - \Lambda(s)] e^{-[\Lambda(t) - \Lambda(s)]} \sum_{n=1}^{\infty} \frac{[\Lambda(t) - \Lambda(s)]^{n-1}}{(n-1)!} \\ &= [\Lambda(t) - \Lambda(s)] e^{-[\Lambda(t) - \Lambda(s)]} e^{[\Lambda(t) - \Lambda(s)]} \\ &= \Lambda(t) - \Lambda(s) = \int_s^t \lambda(\sigma) d\sigma \quad (C-5) \end{aligned}$$

From Eq. C-5, the intensity, $\lambda(\sigma)$, may be interpreted as the instantaneous average rate that points occur at time σ . If $\lambda(\sigma)$ is a constant, the Poisson counting process is homogeneous, otherwise it is inhomogeneous. Furthermore, for $\lambda(\sigma)$ constant,

$$\Lambda(t) - \Lambda(s) = \lambda(t-s) \quad (C-6)$$

which illustrates that the expected number of occurrences is dependent only on the length of the interval $[s, t)$; thus for constant λ the counting process is stationary.

Associated with the Poisson counting process is the characteristic function

$$\begin{aligned} M(v; s, t) &= E\{e^{jvN(s, t)}\} \\ &= \exp\{(e^{jv} - 1)(\Lambda(t) - \Lambda(s))\} \quad (C-7) \end{aligned}$$

The usefulness of the characteristic function lies in its moment generating property, viz.,

$$E\{N^k(s, t)\} = \frac{1}{j^k} \frac{\partial^k M(v; s, t)}{\partial v^k} \bigg|_{v=0} \quad (C-8)$$

From Eq. C-8,

$$E\{N(s, t)\} = \frac{1}{j} \frac{\partial M(v; s, t)}{\partial v} \bigg|_{v=0} = \Lambda(t) - \Lambda(s) \quad (C-9)$$

and

$$E\{N^2(s,t)\} = \frac{1}{J^2} \frac{\partial^2 M(v,s,t)}{\partial v^2} \Big|_{v=0}$$

$$= \Lambda(t) - \Lambda(s) + [\Lambda(t) - \Lambda(s)]^2 \quad (C-10)$$

Equations C-9 and C-10 can be combined to yield the variance of $N(s,t)$,

$$\text{Var}\{N(s,t)\} = E\{N^2(s,t)\} - [E\{N(s,t)\}]^2$$

$$= \Lambda(t) - \Lambda(s) \quad (C-11)$$

A marked point process is a point process with an auxiliary variable associated with each jump of the point process. Let a_k be the mark associated with the k^{th} jump. It will be assumed that $\{a_k\}$ is a sequence of mutually independent, identically distributed random variables (with density function $p_a(a)$) that is independent of the underlying point process. (Note that the assumed mark space is continuous.) Furthermore, let $N(t)$, $t \geq 0$ be an inhomogeneous Poisson counting process with rate parameter $\lambda(t)$. A generalized Poisson process is of the form

$$n_p(t) = \sum_{k=0}^{N(t)} a_k \quad (C-12)$$

The processes $N(t)$ and $n_p(t)$ are illustrated in Fig C-1. A much more complete discussion of this process may be found in Refs. 8 and 12.

Figure C-1 Generalized Poisson Process
a) Marked Point Process,
b) Poisson Counting Process,
c) Generalized Poisson Process

APPENDIX D

DECOMPOSITION OF THE RANDOM TELEGRAPH WAVE

This appendix contains details on decomposing the random telegraph wave $z(t)$ into the sum of the compensator $A(t)$ and an orthogonal increment process $\zeta(t)$ * In order to obtain a useful decomposition, certain restrictions will be placed on the parameters associated with the process $z(t)$. Their necessity is illustrated by the following observation.

Consider the compensator $A(t)$ defined via

$$A(t) = \sum_{i \geq 1} \int_{w_{i-1}}^{t \wedge w_i} \lambda(u; w_{i-1}, \dots, w_i) du \quad (D-1)$$

Since the w_i are the random occurrence times of the underlying Poisson process, evaluation of the first and second moments of $A(t)$ would require far more knowledge of the intensity process $\lambda(u; w_{i-1}, \dots, w_i)$ than is feasible. In the sequel it is assumed that the intensity is of the form

$$\lambda(u) = \begin{cases} \mu_1 & z(u)=1 \\ \mu_0 & z(u)=0 \end{cases} \quad (D-2)$$

This eliminates the complex dependence of $\lambda(u)$ on the time-history of the Poisson process. Taking μ_1 and μ_0 to be constant is not unduly restrictive since any time-variation would probably take place over periods which are large relative to the occurrence rate of the underlying Poisson process.†

At this point it is convenient to examine the process $z(t)$ in more detail. Define

$$P(t) = \Pr\{z(t) = 1 \mid z(0) = 1\} \quad (D-3)$$

$$Q(t) = \Pr\{z(t) = 1 \mid z(0) = 0\} \quad (D-4)$$

From the definition of $z(t)$ it is straightforward to show that:

$$\Pr\{z(t+\Delta t) = 0 \mid z(t) = 1\} = \mu_1 \Delta t + o(\Delta t) \quad (D-5)$$

$$\Pr\{z(t+\Delta t) = 1 \mid z(t) = 0\} = \mu_0 \Delta t + o(\Delta t) \quad (D-6)$$

*For the linear estimator of interest, it is sufficient that the process noise have orthogonal rather than independent increments. Consequently, the discussion which follows is limited to the second-order statistical properties of the relevant processes.

†Slow variations in the μ_i can be taken into account by constructing a piecewise constant approximation to $\mu_i(t)$ and computing the estimator parameters for each segment.

Now,

$$P(t+\Delta t) = P(t) (1-\mu_1 \Delta t) + (1-P(t)) \mu_0 \Delta t + o(\Delta t) \quad (D-7)$$

which yields the differential equation

$$\frac{dP(t)}{dt} + (\mu_0 + \mu_1) P(t) = \mu_0 \quad (D-8)$$

The solution to Eq D-8 with initial condition $P(0) = 1$ is

$$P(t) = \frac{1}{\mu_0 + \mu_1} [\mu_0 + \mu_1 e^{-(\mu_0 + \mu_1)t}] \quad (D-9)$$

A similar argument gives

$$Q(t) = \frac{\mu_0}{\mu_0 + \mu_1} [1 - e^{-(\mu_0 + \mu_1)t}] \quad (D-10)$$

Of primary interest is the probability that $z(t) = 1$. Let

$$P_1(t) = \Pr\{z(t) = 1\} \quad (D-11)$$

and denote the probability that $z(0) = 1$ by P_0 . Then,

$$\begin{aligned} P_1(t) &= P(t) P_0 + Q(t) (1-P_0) \\ &= \frac{\mu_0}{\mu_0 + \mu_1} (1 - e^{-(\mu_0 + \mu_1)t}) + P_0 e^{-(\mu_0 + \mu_1)t} \end{aligned} \quad (D-12)$$

[In the limit as $t \rightarrow \infty$,

$$P_1(t) \rightarrow \frac{\mu_0}{\mu_0 + \mu_1} \quad (D-13)$$

which agrees with Eq. 3.2-10.] Equation D-12 may be specialized to the case $P_0 = 1$, viz.,

$$P_1(t) = \frac{\mu_0}{\mu_0 + \mu_1} + \frac{\mu_1}{\mu_0 + \mu_1} e^{-(\mu_0 + \mu_1)t} \quad (D-14)$$

The first and second moments of $z(t)$ are readily obtained using Eq D-14 and the fact that $z^2(t) = z(t)$:

$$E\{z(t)\} = E\{z^2(t)\} = 1 \cdot \Pr\{z(t)=1\} + 0 \cdot \Pr\{z(t)=0\}$$

$$= \frac{\mu_0}{\mu_0 + \mu_1} + \frac{\mu_1}{\mu_0 + \mu_1} e^{-(\mu_0 + \mu_1)t} \quad (D-15)$$

For $t \gg \frac{1}{\mu_0 + \mu_1}$, these moments approach the constant value $\frac{\mu_0}{\mu_0 + \mu_1}$.

Heuristically, this may be interpreted as the fraction of time the process $z(t)$ spends in state 1. Let $\xi_1, \eta_1, \xi_2, \eta_2, \xi_3, \dots$ denote the times spent in states 1 and 0, alternately. From the definition of $z(t)$ it follows that $\{\xi_i\}$ and $\{\eta_i\}$, $i=1,2,\dots$, are non-negative, independent random variables with

$$\Pr\{\xi_1 < \tau\} = F_1(\tau) = 1 - e^{-\mu_1 \tau} \quad (D-16)$$

$$\Pr\{\eta_1 < \tau\} = F_0(\tau) = 1 - e^{-\mu_0 \tau} \quad (D-17)$$

The average time spent in state 1 is

$$\bar{\tau}_1 = E\{\xi_1\} = \int_0^\infty \tau dF_1(\tau) = \frac{1}{\mu_1} \quad (D-18)$$

Similarly, for state 0

$$\bar{\tau}_0 = E\{\eta_1\} = \int_0^\infty \tau dF_0(\tau) = \frac{1}{\mu_0} \quad (D-19)$$

Define the quantity $E\{\xi_1 + \eta_1\}$, $i=1,2,\dots$ to be the period T of one cycle of the alternating process 1, 0, 1, 0, Assume that the process has gone through N cycles at time t . Then,

$$t = NT = N \left[\frac{1}{\mu_1} + \frac{1}{\mu_0} \right]$$

and the fraction of time spent in state 1 is

$$E\{z(t)\} = \frac{N E\{\xi_1\}}{t} = \frac{1/\mu_1}{1/\mu_1 + 1/\mu_0} = \frac{\mu_0}{\mu_0 + \mu_1} \quad (D-20)$$

Consider now the compensator $A(t)$. Figure D-1 depicts $A(t)$ for one specific realization of the process $z(t)$. Substituting Eq. D-2 into Eq. 3.4-12 shows that $A(t)$ is piecewise linear with slope μ_1 when $z(t) = 1$ and slope $-\mu_0$ when $z(t) = 0$. Hence, $A(t)$ may be written as

$$A(t) = \mu_1 \xi_1 - \mu_0 \eta_1 + \mu_1 \xi_2 - \dots + r(t) \quad (D-21)$$

where

$$r(t) = \begin{cases} \mu_1(t - w_{2n}) & w_{2n} \leq t < w_{2n+1} \\ \mu_1 \xi_{n+1} - \mu_0(t - w_{2n+1}) & w_{2n+1} \leq t < w_{2n+2} \end{cases} \quad (D-22)$$

and n is the number of complete cycles of the alternating process prior to time t .

R 79704

Figure D-1 Realization of the Compensator $A(t)$

Equation D-21 may be rewritten as

$$A(t) = \mu_1 \chi_1(t) - \mu_0 \chi_0(t) \quad (D-23)$$

with $\chi_1(t) = (\xi_1 + \xi_2 + \dots)$ and $\chi_0(t) = (\eta_1 + \eta_2 + \dots)$ the total time spent in states 1 and 0, respectively, during the interval $(0, t)$. Since $\chi_0(t) + \chi_1(t) = t$,

$$\begin{aligned} A(t) &= \mu_1 [t - \chi_0(t)] - \mu_0 \chi_0(t) \\ &= \mu_1 t - (\mu_0 + \mu_1) \chi_0(t) \end{aligned} \quad (D-24)$$

Let

$$\Omega(t, \tau) = \Pr\{\chi_0(t) \leq \tau\} \quad (D-25)$$

denote the distribution function of the random variable $\chi_0(t)$. For $F_0(\tau)$ and $F_1(\tau)$ defined by Eqs D-16 and D-17, $\Omega(t, \tau)$ has the form (Ref. 100)

$$\Omega(t, \tau) = e^{-\mu_1(t-\tau)} [1 +$$

$$\sqrt{\mu_0 \mu_1 (t-\tau)} \int_0^\tau e^{-\mu_0 \rho} \rho^{-1/2} I_1(2\sqrt{\mu_0 \mu_1 (t-\tau) \rho}) d\rho] \quad (D-26)$$

where $I_1(\alpha)$ is the Bessel function of order 1 defined by

$$I_1(\alpha) = \sum_{j=0}^{\infty} \frac{(\alpha/2)^{2j+1}}{j'(j+1)'} \quad (D-27)$$

Using Eq D-26, the r^{th} moment of $\chi_0(t)$ is

$$\begin{aligned} E\{\chi_0^r(t)\} &= \int_0^t \tau^r d\Omega(t, \tau) \\ &= r \int_0^t \tau^{r-1} [1 - \Omega(t, \tau)] d\tau \end{aligned} \quad (D-28)$$

It is obvious from Eqs. D-26 and D-28 that even the low-order moments of $\chi_0(t)$ are quite complicated. However, explicit results are available in the limiting case $t \rightarrow \infty$.

In Ref. 100 it is shown that if $\bar{\tau}_0 + \bar{\tau}_1 < \infty$, then

$$\lim_{t \rightarrow \infty} \frac{E\{\chi_0(t)\}}{t} = \frac{\bar{\tau}_0}{\bar{\tau}_0 + \bar{\tau}_1} = \frac{\mu_1}{\mu_0 + \mu_1} \quad (D-29)$$

Taking the expected value of Eq D-24 and substituting Eq D-25 yields

$$\begin{aligned} \lim_{t \rightarrow \infty} E\{A(t)\} &= \lim_{t \rightarrow \infty} [\mu_1 t - (\mu_0 + \mu_1) E\{\chi_0(t)\}] \\ &= \lim_{t \rightarrow \infty} [\mu_1 t - \mu_1 t] = 0 \end{aligned} \quad (D-30)$$

Thus, in steady state, the compensator $A(t)$ is a zero-mean process.

To calculate the second moment take the expected value of the square of Eq D-24, viz.,

$$\begin{aligned} E\{A^2(t)\} &= E\{[\mu_1 t - (\mu_0 + \mu_1) \chi_0(t)]^2\} \\ &= \mu_1^2 t^2 - 2\mu_1(\mu_0 + \mu_1)t E\{\chi_0(t)\} \\ &\quad + (\mu_0 + \mu_1)^2 E\{\chi_0^2(t)\} \end{aligned} \quad (D-31)$$

From Ref. 100,

$$\lim_{t \rightarrow \infty} E\{\chi_0^2(t)\} = \frac{2\mu_0\mu_1}{(\mu_0 + \mu_1)^3} t + \frac{\mu_1^2}{(\mu_0 + \mu_1)^2} t^2 \quad (D-32)$$

Substituting Eq D-32 into Eq. D-31 and cancelling terms gives

$$\lim_{t \rightarrow \infty} E\{A^2(t)\} = \frac{2\mu_0\mu_1}{\mu_0 + \mu_1} t \quad (D-33)$$

Having examined the steady state statistical properties of $z(t)$ and $A(t)$, it is now possible to give an explicit representation for the process

$$\zeta(t) = z(t) - A(t) \quad (D-34)$$

Observe first that

$$\lim_{t \rightarrow \infty} E\{\zeta(t)\} = \frac{\mu_0}{\mu_0 + \mu_1} \quad (D-35)$$

In the following it is more convenient to deal with the zero-mean process

$$\tilde{\zeta}(t) = [z(t) - E\{z(t)\}] - A(t) = \zeta(t) - \frac{\mu_0}{\mu_0 + \mu_1} \quad (D-36)$$

It was stated earlier that $\zeta(t)$ is an orthogonal increment process. Actually, $\tilde{\zeta}(t)$, the zero-mean modification of $\zeta(t)$, has orthogonal increments. A general proof of this for any renewal process $z(t)$ is given in Davis (Ref 21) and is not repeated here.

The second moment of $\tilde{\zeta}(t)$ is obtained from Eqs. D-34 and D-36 as

$$\begin{aligned} E\{\tilde{\zeta}^2(t)\} &= E\{[z(t) - A(t) - \frac{\mu_0}{\mu_0 + \mu_1}]^2\} \\ &= E\{z^2(t)\} + E\{A^2(t)\} \\ &\quad - 2 E\{z(t)A(t)\} - \left(\frac{\mu_0}{\mu_0 + \mu_1}\right)^2 \end{aligned} \quad (D-37)$$

The only unknown in this expression is $E\{z(t)A(t)\}$. From Fig. D-1 it is evident that $A(t)$ is composed of alternating segments with slopes of μ_1 and $-\mu_0$. The product $z(t)A(t)$ has the form shown in Fig. D-2 for the realization of Fig D-1. This function consists of the original portions of $A(t)$ with positive slope, and is zero elsewhere.

R 79703

Figure D-2 Realization of the Product $z(t)A(t)$

Let

$$A^+(t) = \begin{cases} A(t) & \text{when the slope is positive} \\ 0 & \text{when the slope is negative} \end{cases} \quad (D-38)$$

and

$$A^-(t) = \begin{cases} A(t) & \text{when the slope is negative} \\ 0 & \text{when the slope is positive} \end{cases} \quad (D-39)$$

Since $A(t) = A^+(t) + A^-(t)$ it follows that for large t ,

$$E\{A(t)\} = E\{A^+(t)\} + E\{A^-(t)\} = 0 \quad (D-40)$$

From the steady-state structure of $A(t)$ it is readily evident that the segments with positive slope have the same statistical character as the segments with negative slope, hence

$$E\{A^+(t)\} = E\{A^-(t)\} = 0 \quad (D-41)$$

Finally, since $A^+(t)$ is just $z(t)A(t)$, the expectation of the product is zero. Thus, Eq. D-37 reduces to:

$$E\{\xi^2(t)\} = E\{z^2(t)\} + E\{A^2(t)\} - \left(\frac{\mu_0}{\mu_0 + \mu_1}\right)^2 \quad (D-42)$$

Of primary interest is the behavior of $\xi(t)$ for t large. Substituting Eqs. D-15 and D-33 into Eq. D-42 yields

$$\begin{aligned} \lim_{t \rightarrow \infty} E\{\xi^2(t)\} &= \lim_{t \rightarrow \infty} \left[\frac{\mu_0}{\mu_0 + \mu_1} + \frac{\mu_1}{\mu_0 + \mu_1} e^{-(\mu_0 + \mu_1)t} \right. \\ &\quad \left. + \frac{2\mu_0\mu_1}{\mu_0 + \mu_1} t - \left(\frac{\mu_0}{\mu_0 + \mu_1}\right)^2 \right] = \frac{2\mu_0\mu_1}{\mu_0 + \mu_1} t \end{aligned} \quad (D-43)$$

To summarize, the random telegraph wave can be decomposed as

$$z(t) = A(t) + \xi(t) + \xi \quad (D-44)$$

where

$A(t)$ is the compensator defined in Eq. D-1

$\xi(t)$ is a zero-mean, orthogonal-increment process with

$$E\{\xi^2(t)\} = \frac{2\mu_0\mu_1}{\mu_0 + \mu_1} t \text{ for } t \text{ large}$$

ξ is a constant dependent upon the transition rate of the point process defining $z(t)$.

APPENDIX E
SOLUTION OF SINGULARLY PERTURBED
KALMAN FILTER EQUATIONS

E.1 INTRODUCTION

In this appendix solution equations are derived for the SP Kalman filter equations for the system

$$\begin{aligned}\dot{\underline{x}}_1 &= F_{11}\underline{x}_1 + \varepsilon F_{12}\underline{x}_2 + G_1 \underline{w} \\ \varepsilon \dot{\underline{x}}_2 &= F_{21}\underline{x}_1 + F_{22}\underline{x}_2 + \varepsilon G_2 \underline{w} \\ \underline{y} &= H_1 \underline{x}_1 + H_2 \underline{x}_2 + \underline{v}\end{aligned}\quad (E.1-1)$$

This formulation admits a computationally satisfying solution using the method of multiple scales.

By viewing Eq. E.1-1 as

$$\begin{aligned}\dot{\underline{x}} &= \underline{F}\underline{x} + \underline{G}\underline{w} \\ \underline{y} &= \underline{H}\underline{x} + \underline{v}\end{aligned}\quad (E.1-2)$$

where

$$\underline{F} = \begin{bmatrix} F_{11} & \varepsilon F_{12} \\ F_{21} & F_{22} \\ \varepsilon & \varepsilon \end{bmatrix}\quad (E.1-3)$$

$$\underline{G} = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix}\quad (E.1-4)$$

$$\underline{H} = \begin{bmatrix} H_1 & H_2 \end{bmatrix}\quad (E.1-5)$$

the SP Kalman filter propagation equations can be derived as

$$\begin{aligned}\dot{\underline{A}} &= F_{11}\underline{A} + \underline{A}F_{11}^T + \varepsilon(F_{12}\underline{B}^T + \underline{B}F_{12}^T) + G_{11} \\ \varepsilon \dot{\underline{B}} &= \underline{A}F_{12}^T + \underline{B}F_{22}^T + \varepsilon F_{11}\underline{B} + \varepsilon^2 F_{12}\underline{C} + \varepsilon G_{12} \\ \varepsilon \dot{\underline{C}} &= F_{21}\underline{B} + \underline{B}^T F_{21}^T + F_{22}\underline{C} + \underline{C}F_{22}^T + \varepsilon G_{22} \\ \dot{\underline{\hat{x}}}_1 &= F_{11}\underline{\hat{x}}_1 + \varepsilon F_{12}\underline{\hat{x}}_2 \\ \varepsilon \dot{\underline{\hat{x}}}_2 &= F_{21}\underline{\hat{x}}_1 + F_{22}\underline{\hat{x}}_2\end{aligned}\quad (E.1-6)$$

$$\text{where cov } \begin{bmatrix} \underline{x}_1 \\ \underline{x}_2 \end{bmatrix} = \begin{bmatrix} \underline{A} & \underline{B} \\ \underline{B}^T & \underline{C} \end{bmatrix}\quad (E.1-8)$$

and

$$G_{ij} = G_i Q G_j^T \quad (E.1-9)$$

Using the two time-scales

$$t_0 = t, \quad t_1 = \varepsilon^{-1}t \quad (E.1-10)$$

implies the equivalence

$$\varepsilon \frac{d}{dt} = \varepsilon \frac{\partial}{\partial t_0} + \frac{\partial}{\partial t_1} \quad (E.1-11)$$

so that Eq. E.1-6 can be rewritten as

$$\begin{aligned}\varepsilon \frac{\partial \underline{A}}{\partial t_0} + \frac{\partial \underline{A}}{\partial t_1} &= \varepsilon(F_{11}\underline{A} + \underline{A}F_{11}^T) \\ &\quad + \varepsilon^2(F_{12}\underline{B}^T + \underline{B}F_{12}^T) + \varepsilon G_{11} \\ \varepsilon \frac{\partial \underline{B}}{\partial t_0} + \frac{\partial \underline{B}}{\partial t_1} &= \underline{A}F_{21}^T + \underline{B}F_{22}^T + \varepsilon F_{11}\underline{B} + \varepsilon^2 F_{12}\underline{C} + \varepsilon G_{12} \\ \varepsilon \frac{\partial \underline{C}}{\partial t_0} + \frac{\partial \underline{C}}{\partial t_1} &= F_{21}\underline{B} + \underline{B}^T F_{21}^T + F_{22}\underline{C} + \underline{C}F_{22}^T + \varepsilon G_{22}\end{aligned}\quad (E.1-12)$$

and Eq. E.1-7 as

$$\begin{aligned}\varepsilon \frac{\partial \underline{\hat{x}}_1}{\partial t_0} + \frac{\partial \underline{\hat{x}}_1}{\partial t_1} &= \varepsilon F_{11}\underline{\hat{x}}_1 + \varepsilon^2 F_{12}\underline{\hat{x}}_2 \\ \varepsilon \frac{\partial \underline{\hat{x}}_2}{\partial t_0} + \frac{\partial \underline{\hat{x}}_2}{\partial t_1} &= F_{21}\underline{\hat{x}}_1 + F_{22}\underline{\hat{x}}_2\end{aligned}\quad (E.1-13)$$

The solution equations corresponding to Eqs. E.1-12 and E.1-13 are derived in the following sections.

E.2 THE COVARIANCE EQUATIONS

In analyzing Eq. E.1-12 using the multiple scales approach the solution forms assumed are

$$\begin{aligned}\underline{A} &\sim \underline{A}_0 + \varepsilon \underline{A}_1 + \varepsilon^2 \underline{A}_2 \dots \\ \underline{B} &\sim \underline{B}_0 + \varepsilon \underline{B}_1 + \varepsilon^2 \underline{B}_2 \dots \\ \underline{C} &\sim \underline{C}_0 + \varepsilon \underline{C}_1 + \varepsilon^2 \underline{C}_2 \dots\end{aligned}\quad (E.2-1)$$

where the $\underline{A}_i$, $\underline{B}_i$ and $\underline{C}_i$ are functions of t_0 and t_1 . Use of the forms of Eq. E.2-1 in Eq. E.1-12 allows the separation of terms according to order of magnitude in powers of ε . The following sets of partial differential equations result:

Terms of $O(1)$:

$$\frac{\partial A_0}{\partial t_1} = 0$$

$$\frac{\partial B_0}{\partial t_1} = A_0 F_{21}^T + B_0 F_{22}^T \quad (E.2-2)$$

$$\frac{\partial C_0}{\partial t_1} = F_{21} B_0 + B_0^T F_{21}^T + F_{22} C_0 + C_0 F_{22}^T$$

Terms of $O(\varepsilon)$:

$$\frac{\partial A_0}{\partial t_0} + \frac{\partial A_1}{\partial t_1} = F_{11} A_0 + A_0 F_{11}^T + G_{11}$$

$$\frac{\partial B_0}{\partial t_0} + \frac{\partial B_1}{\partial t_1} = A_1 F_{21}^T + B_1 F_{22}^T + F_{11} B_0 + G_{12} \quad (E.2-3)$$

$$\frac{\partial C_0}{\partial t_0} + \frac{\partial C_1}{\partial t_1} = F_{21} B_1 + B_1^T F_{21}^T + F_{22} C_1 + C_1 F_{22}^T + G_{22}$$

Terms of $O(\varepsilon^2)$

$$\frac{\partial A_1}{\partial t_0} + \frac{\partial A_2}{\partial t_1} = F_{11} A_1 + A_1 F_{11}^T + F_{12} B_0^T + B_0 F_{12}^T$$

$$\frac{\partial B_1}{\partial t_0} + \frac{\partial B_2}{\partial t_1} = A_2 F_{21}^T + B_2 F_{22}^T + F_{11} B_1 + F_{12} C_0$$

$$\frac{\partial C_1}{\partial t_0} + \frac{\partial C_2}{\partial t_1} = F_{21} B_2 + B_2^T F_{21}^T + F_{22} C_2 + C_2 F_{22}^T$$

etc.

(E.2-4)

Since the partial differential equations listed above are linear, they may be resolved into ordinary differential equations by separation of variables. Accordingly, it is assumed that

$$A_i = K_i(t_0) + U_i(t_1)$$

$$B_i = L_i(t_0) + V_i(t_1) \quad (E.2-5)$$

$$C_i = M_i(t_0) + W_i(t_1)$$

with the added stipulation that the functions U_i , V_i , W_i be exponentially bounded. The ordinary differential equations governing the functions K_i , L_i , M_i , U_i , V_i , W_i are presented below:

$$\frac{dU_0}{dt_1} = 0$$

$$\frac{dV_0}{dt_1} = V_0 F_{21}^T + V_0 F_{22}^T$$

$$0 = K_0 F_{21}^T + L_0 F_{22}^T$$

$$\frac{dW_0}{dt_1} = F_{21} V_0 + V_0^T F_{21}^T + F_{22} W_0 + W_0 F_{22}^T$$

$$0 = F_{21} L_0 + L_0^T F_{21}^T + F_{22} M_0 + M_0 F_{22}^T$$

$$\frac{dK_0}{dt_0} = F_{11} K_0 + K_0 F_{11}^T + G_{11}$$

$$\frac{dU_1}{dt_1} = F_{11} U_0 + U_0 F_{11}^T$$

$$\frac{dL_0}{dt_0} = K_1 F_{21}^T + L_1 F_{22}^T + F_{11} L_0 + G_{12}$$

$$\frac{dV_1}{dt_1} = U_1 F_{21}^T + V_1 F_{22}^T + F_{11} V_0$$

$$\frac{dM_0}{dt_0} = F_{21} L_1 + L_1^T F_{21}^T + F_{22} M_1 + M_1 F_{22}^T + G_{22}$$

$$\frac{dW_1}{dt_1} = F_{21} V_1 + V_1^T F_{21}^T + F_{22} W_1 + W_1 F_{22}^T$$

$$\frac{dK_1}{dt_0} = F_{11} K_1 + K_1 F_{11}^T + F_{12} L_0^T + L_0 F_{12}^T \quad (E.2-6)$$

etc.

Recalling that the functions of t_1 are all required to be exponentially bounded, further simplifications result. The equations to be "solved" can be derived, after some algebraic manipulation, to be [(with $(\cdot)'$ denoting $\frac{d(\cdot)}{dt_1}$ and $(\dot{\cdot})$ denoting $\frac{d(\cdot)}{dt_0}$]

$$U_0 = 0$$

$$V_0' = V_0 F_{22}^T \quad (E.2-7)$$

$$W_0' = F_{22} W_0 + W_0 F_{22}^T + F_{21} V_0 + V_0^T F_{21}^T$$

$$\dot{K}_0 = F_{11}K_0 + K_0F_{11}^T + G_{11}$$

$$L_0 = -K_0D_0^T \quad (E.2-8)$$

$$M_0 = D_0K_0D_0^T$$

$$U_1 = 0$$

$$V_1' = V_1F_{22}^T + F_{11}V_0 \quad (E.2-9)$$

$$W_1' = F_{21}V_1 + V_1F_{21}^T + F_{22}W_1 + W_1F_{22}^T$$

$$\dot{K}_1 = F_{11}K_1 + K_1F_{11}^T - F_{12}D_0K_0 - K_0D_0^TF_{12}^T$$

$$L_1 = -K_1D_0^T - K_0D_1^T - E_1 \quad (E.2-10)$$

$$M_1 = D_0K_1D_0^T + D_0K_0D_1^T + D_1K_0D_0^T + N_1$$

where

$$D_0 = F_{22}^{-1}F_{21}$$

$$D_1 = F_{22}^{-1}D_0F_{11} \quad (E.2-11)$$

$$E_1 = (G_{12} + G_{11}D_0^T)F_{22}^{-T}$$

and N_1 is the symmetric solution of

$$F_{22}N_1 + N_1F_{22}^T = F_{21}E_1 + E_1^TF_{21}^T + D_0G_{11}D_0^T - G_{22} \quad (E.2-12)$$

Inspection of the foregoing equations suggests that the exponential boundedness of the t_1 -functions requires that the matrix F_{22} be stable. Furthermore, the equations governing W_0 , V_1 , W_1 can be simplified to

$$W_0 = \bar{W}_0 - D_0V_0 - V_0^TD_0^T$$

$$V_1 = \bar{V}_1 + F_{11}V_0t_1 \quad (E.2-13)$$

$$W_1 = \bar{W}_1 - D_0V_1 - V_1^TD_0^T - D_1V_0 - V_0^TD_1^T$$

where $\bar{W}_0$, $\bar{V}_1$, $\bar{W}_1$ are governed by

$$\bar{W}_0' = F_{22}\bar{W}_0 + \bar{W}_0F_{22}^T$$

$$\bar{V}_1' = \bar{V}_1F_{22}^T \quad (E.2-14)$$

$$\bar{W}_1' = F_{22}\bar{W}_1 + \bar{W}_1F_{22}^T$$

The analysis of the covariance equations is completed by specifying the initial conditions for the differential equations governing K_0 , K_1 , V_0 , $\bar{V}_1$, $\bar{W}_0$, $\bar{W}_1$. It is assumed here that the initial covariance matrix is composed of

$$A(0) = A^*, B(0) = B^*, C(0) = C^* \quad (E.2-15)$$

This suggests that

$$K_0(0) = A^*$$

$$L_0(0) + V_0(0) = B^*$$

$$M_0(0) + W_0(0) = C^* \quad (E.2-16)$$

$$K_1(0) = 0$$

$$L_1(0) + V_1(0) = 0$$

$$M_1(0) + W_1(0) = 0$$

After some algebraic manipulation the complete set of the desired initial conditions are found to be

$$K_0(0) = A^*$$

$$K_1(0) = 0$$

$$V_0(0) = B^* + A^*D_0^T \quad (E.2-17)$$

$$\bar{V}_1(0) = E_1 + A^*D_1^T$$

$$\bar{W}_0(0) = C^* + D_0B^* + B^{*T}D_0^T + D_0A^*D_0^T$$

$$\begin{aligned} \bar{W}_1(0) = & D_0E_1 + E_1^TD_0 + D_1B^* + B^{*T}D_1^T + D_0A^*D_1^T \\ & + D_1A^*D_0^T - N_1 \end{aligned}$$

E.3 THE ESTIMATE EQUATIONS

The analysis of the previous section applies equally to the covariance propagation of the nonlinear system

$$\begin{aligned}\dot{\hat{x}}_1 &= f_1(\hat{x}_1, \hat{x}_2) + G_1 w & : & \quad \hat{x}_1(0) = \underline{a}^* \\ \dot{\hat{x}}_2 &= f_2(\hat{x}_1, \hat{x}_2) + G_2 w & : & \quad \hat{x}_2(0) = \underline{b}^*\end{aligned}\quad (E.3-1)$$

where

$$F_{11} = \frac{\partial f_1}{\partial \hat{x}_1}(\underline{a}^*, \underline{b}^*), \quad F_{12} = \frac{1}{\varepsilon} \frac{\partial f_1}{\partial \hat{x}_2}(\underline{a}^*, \underline{b}^*) \quad (E.3-2)$$

$$F_{21} = \varepsilon \frac{\partial f_2}{\partial \hat{x}_1}(\underline{a}^*, \underline{b}^*), \quad F_{22} = \varepsilon \frac{\partial f_2}{\partial \hat{x}_2}(\underline{a}^*, \underline{b}^*)$$

However, the extended Kalman filter approach to the propagation of the state estimate for the nonlinear system of Eq. E.3-1 requires the solution of

$$\begin{aligned}\dot{\hat{x}}_1 &= f_1(\hat{x}_1, \hat{x}_2) & : & \quad \hat{x}_1(0) = \underline{a}^* \\ \dot{\hat{x}}_2 &= f_2(\hat{x}_1, \hat{x}_2) & : & \quad \hat{x}_2(0) = \underline{b}^*\end{aligned}\quad (E.3-3)$$

To order $O(\varepsilon)$, these propagation equations may be approximated as

$$\begin{aligned}\dot{\hat{x}}_1 &= F_{11}\hat{x}_1 + \varepsilon F_{12}\hat{x}_2 + f_1(\underline{a}^*, \underline{b}^*) - F_{11}\underline{a}^* - \varepsilon F_{12}\underline{b}^* \\ \dot{\hat{x}}_2 &= F_{21}\hat{x}_1 + F_{22}\hat{x}_2 + \varepsilon f_2(\underline{a}^*, \underline{b}^*) - F_{21}\underline{a}^* - F_{22}\underline{b}^*\end{aligned}\quad (E.3-4)$$

As in the preceding section a series form is assumed for the state estimates

$$\begin{aligned}\hat{x}_1 &\sim \underline{a}_0 + \varepsilon \underline{a}_1 + \varepsilon^2 \underline{a}_2 \dots \\ \hat{x}_2 &\sim \underline{b}_0 + \varepsilon \underline{b}_1 + \varepsilon^2 \underline{b}_2 \dots\end{aligned}\quad (E.3-5)$$

which yields the following pairs of partial differential equations:

Terms of $O(1)$:

$$\frac{\partial \underline{a}_0}{\partial t_1} = 0 \quad (E.3-6)$$

$$\frac{\partial \underline{b}_0}{\partial t_1} = F_{21}\underline{a}_0 + F_{22}\underline{b}_0 - F_{21}\underline{a}^* - F_{22}\underline{b}^*$$

Terms of $O(\varepsilon)$:

$$\frac{\partial \underline{a}_1}{\partial t_0} + \frac{\partial \underline{a}_1}{\partial t_1} = F_{11}\underline{a}_0 + f_1(\underline{a}^*, \underline{b}^*) - F_{11}\underline{a}^* \quad (E.3-7)$$

$$\frac{\partial \underline{b}_0}{\partial t_0} + \frac{\partial \underline{b}_1}{\partial t_1} = F_{21}\underline{a}_1 + F_{22}\underline{b}_1 + f_2(\underline{a}^*, \underline{b}^*)$$

Terms of $O(\varepsilon^2)$:

$$\frac{\partial \underline{a}_1}{\partial t_0} + \frac{\partial \underline{a}_2}{\partial t_1} = F_{11}\underline{a}_1 + F_{12}\underline{b}_0 - F_{12}\underline{b}^* \quad (E.3-8)$$

$$\frac{\partial \underline{b}_1}{\partial t_0} + \frac{\partial \underline{b}_2}{\partial t_1} = F_{21}\underline{a}_2 + F_{22}\underline{b}_2$$

etc.

To effect a separation of variables it is assumed that

$$\underline{a}_1 = \underline{p}_1(t_0) + \underline{r}_1(t_1) \quad (E.3-9)$$

$$\underline{b}_1 = \underline{q}_1(t_0) + \underline{s}_1(t_1)$$

where the functions $\underline{r}_1$ and $\underline{s}_1$ are required to be exponentially bounded. The ordinary differential equations that result from the separation of variables are

$$\begin{aligned}\dot{\underline{r}}_0 &= 0 \\ \dot{\underline{s}}_0 &= F_{21}\underline{r}_0 + F_{22}\underline{s}_0 \\ \dot{\underline{q}}_0 &= F_{21}\underline{p}_0 + F_{22}\underline{q}_0 - F_{21}\underline{a}^* - F_{22}\underline{b}^* \\ \dot{\underline{r}}_1 &= F_{11}\underline{r}_0 \\ \dot{\underline{s}}_1 &= F_{21}\underline{r}_1 + F_{22}\underline{s}_1 \\ \dot{\underline{p}}_0 &= F_{11}\underline{p}_0 + f_1(\underline{a}^*, \underline{b}^*) - F_{11}\underline{a}^* \\ \dot{\underline{q}}_0 &= F_{21}\underline{p}_1 + F_{22}\underline{q}_1 + f_2(\underline{a}^*, \underline{b}^*) \\ \dot{\underline{p}}_1 &= F_{11}\underline{p}_1 + F_{12}\underline{q}_0 - F_{12}\underline{b}^*\end{aligned}\quad (E.3-10)$$

Then, since $\underline{r}_0$ and $\underline{r}_1$ are to be exponentially bounded, the following sets of equations result

$$\begin{aligned}\underline{r}_0 &= 0 \\ \dot{\underline{s}}_0 &= F_{22}\underline{s}_0\end{aligned}\quad (E.3-11)$$

$$\begin{aligned}\dot{\underline{p}}_0 &= F_{11}(\underline{p}_0 - \underline{a}^*) + \underline{f}_1(\underline{a}^*, \underline{b}^*) \\ \underline{q}_0 &= -D_0(\underline{p}_0 - \underline{a}^*) + \underline{b}^*\end{aligned}\quad (E.3-12)$$

$$\begin{aligned}\underline{r}_1 &= 0 \\ \dot{\underline{s}}_1 &= F_{22}\underline{s}_1\end{aligned}\quad (E.3-13)$$

$$\begin{aligned}\dot{\underline{p}}_1 &= F_{11}\underline{p}_1 - F_{12}D_0(\underline{p}_0 - \underline{a}^*) \\ \underline{q}_1 &= -D_0\underline{p}_1 - D_1(\underline{p}_0 - \underline{a}^*) - F_{22}^{-1}\{\underline{f}_2(\underline{a}^*, \underline{b}^*) \\ &\quad + D_0\underline{f}_1(\underline{a}^*, \underline{b}^*)\}\end{aligned}\quad (E.3-14)$$

The appropriate initial conditions for the differential equations governing $\underline{p}_0$, $\underline{p}_1$, $\underline{s}_0$, $\underline{s}_1$ are

$$\begin{aligned}\underline{p}_0(0) &= \underline{a}^* \\ \underline{p}_1(0) &= \underline{0} \\ \underline{s}_0(0) &= \underline{0} \\ \underline{s}_1(0) &= F_{22}^{-1}\{\underline{f}_2(\underline{a}^*, \underline{b}^*) + D_0\underline{f}_1(\underline{a}^*, \underline{b}^*)\}\end{aligned}\quad (E.3-15)$$

The fact that $\underline{s}_0$ has zero initial conditions, together with F_{22} being a stable matrix, implies that

$$\underline{s}_0 = \underline{0} \quad (E.3-16)$$

This completes the analysis of the estimate equations for the nonlinear system of Eq. E.3-1.

E.4 THE UPDATE EQUATIONS

The SP' approach does not modify the update equations for the Kalman filter or the extended Kalman filter. If the measurements are

$$\underline{y} = \underline{h}(\underline{x}_1, \underline{x}_2) + \underline{v} \quad (E.4-1)$$

$$\begin{aligned}\text{and} \quad H_1 &= \frac{\partial \underline{h}}{\partial \underline{x}_1^T} (\hat{\underline{x}}_1(-), \hat{\underline{x}}_2(-)) \\ H_2 &= \frac{\partial \underline{h}}{\partial \underline{x}_2^T} (\hat{\underline{x}}_1(-), \hat{\underline{x}}_2(-))\end{aligned}\quad (E.4-2)$$

Then the estimate update equation can be written and

$$\hat{\underline{x}}(+) = \hat{\underline{x}}(-) + K[\underline{y} - \underline{h}(\hat{\underline{x}}_1(-), \hat{\underline{x}}_2(-))] \quad (E.4-3)$$

where the Kalman gain is given by

$$K = P(-)H^T[HP(-)H^T + R]^{-1} \quad : H = [H_1 \ H_2] \quad (E.4-4)$$

and the covariance update equation can be written as

$$P(+) = P(-) - KHP(-) \quad (E.4-5)$$

The updated estimate of Eq. E.4-3 and the updated covariance of Eq. E.4-5 can be partitioned to obtain the initial conditions for the next propagation cycle. Thus

$$\hat{\underline{x}} = \begin{bmatrix} \underline{a}^* \\ \underline{b}^* \end{bmatrix} \quad (E.4-6)$$

and

$$P = \begin{bmatrix} A^* & B^* \\ B^{*T} & C^* \end{bmatrix} \quad (E.4-7)$$

APPENDIX F

A BOILER-TURBINE-EXCITER-GENERATOR MODEL

The boiler-turbine-exciter-generator (BTEG) model presented in this appendix is an aggregation of models presented in Appendix A. The present model was developed with a view toward studying singularly perturbed (SP) formulations of a generating system. This requires a model in which the coupling between fast and slow modes is retained. The models of Appendix A provide equations that govern transient behavior and long-term behavior separately. As such, additional details were extracted from GE and EPRI reports (Refs. 35 and 36) to provide a composite model representative of both transient and long-term behavior. The only significant differences between the present model and the models of Appendix A are in the boiler control, boiler and turbine control representations.

The objective of this appendix is to present the mathematical relationships between the state variables of the BTEG system. These relationships provide the linearized dynamics matrix that would be used in a simulation or other numerical analysis of the system. Specifically, the dynamics matrix derived here is used to study the feasibility of the SP methodology for sub-optimal estimation of the BTEG system. As such, the details addressed in this appendix are limited to those that influence the dynamics matrix of the BTEG model.

Block diagrams of the boiler control and boiler dynamics are presented in Figs. 5-3 and 5-4 of Ref. 36. For the purposes of this model, the rate limits and deadband non-linearities have been eliminated to obtain equations relevant to normal operations. The state variables of the boiler control model are x_1 through x_4 ; their dynamics are governed by

$$\dot{x}_1 = x_2 + F_1 \frac{K_{MWB}}{R_2} x_{30} + F_1 f_1(x_{22}, x_{23}) \dots \quad (F-1)$$

$$\dot{x}_2 = -\frac{10}{T_R} x_2 - F_2 \frac{K_{MWB}}{R_2} x_{30} - F_2 f_1(x_{22}, x_{23}) \dots \quad (F-2)$$

$$\dot{x}_3 = -\frac{K_{IC}}{R_2} x_{30} + K_{IC} K_{PB} f_1(x_{22}, x_{23}) \dots \quad (F-3)$$

$$\dot{x}_4 = x_3 - C_1 x_4 - \frac{K_{PC}}{R_2} x_{30} + K_{PC} K_{PB} f_1(x_{22}, x_{23}) \dots \quad (F-4)$$

where

$$F_1 = 10 \left(\frac{9K_P}{T_R} - K_I \right) \quad (F-5)$$

$$F_2 = \frac{90}{T_R} \left(\frac{10K_P}{T_R} - K_I \right) \quad (F-6)$$

$$f_1(x_{22}, x_{23}) = \frac{1}{2K_{SH}} \left[\frac{\sqrt{1 + 4K_{SH}^2 x_{22} x_{23}} - 1}{x_{22}^2} \right] \quad (F-7)$$

The feedwater control path of the boiler model is represented by states x_5 and x_6 whose dynamics are governed by:

$$\begin{aligned} \dot{x}_5 = & F_3 x_1 + x_6 - F_3 \frac{(K_{FF} + 10K_P K_{MWB})}{R_2} x_{30} \\ & - 10F_3 K_P f_1(x_{22}, x_{23}) + F_3 (1-K_F) f_2(x_{22}, x_{23}) \dots \end{aligned} \quad (F-8)$$

$$\begin{aligned} \dot{x}_6 = & F_4 x_1 - \frac{K_{AUX2} K_{W2}}{T_{F2}} x_5 - \frac{(1 + K_{AUX2} K_{W1})}{T_{F2}} x_6 \\ & + F_4 \frac{(K_{FF} + 10K_P K_{MWB})}{R_2} x_{30} \\ & - 10F_4 K_P f_1(x_{22}, x_{23}) + F_4 (1-K_F) f_2(x_{22}, x_{23}) \end{aligned} \quad (F-9)$$

where

$$F_3 = \frac{K_{AUX2} K_{W1}}{T_{F2}} (1-K) \quad (F-10)$$

$$F_4 = F_3 \left[\frac{K_{W2}}{K_{W1}} + \frac{(1 + K_{AUX2} K_{W1})}{T_{F2}} \right] \quad (F-11)$$

$$f_2(x_{22}, x_{23}) = x_{22} f_1(x_{22}, x_{23}) \quad (F-12)$$

States x_7 through x_{17} represent the fuel-air control path of the boiler dynamics. States x_8 through x_{17} are artificially introduced to approximate the transport delay in the fuel supply system. The equations governing the dynamics of states x_7 through x_{17} are

$$\begin{aligned} \dot{x}_7 = & F_5 x_1 - \frac{1}{T_{F1}} x_7 - F_5 \frac{(K_{FF} + 10K_P K_{MWB})}{R_2} x_{30} \\ & - 10F_5 K_P f_1(x_{22}, x_{23}) + F_5 (1-K_F) f_2(x_{22}, x_{23}) \dots \end{aligned} \quad (F-13)$$

$$\dot{x}_k = \frac{10}{D_1} x_{k-1} - \frac{10}{D_1} x_k \quad k=8,9,10,\dots,16,17 \quad (F-14)$$

The remainder of the boiler model is combined with the turbine model.

The governor-turbine model is shown in Figs 5-5 and 5-11 of Ref. 36. For present purposes, the limit, backlash and least value gate blocks have been eliminated, as has been the IPL path. States x_{18} through x_{23} represent the boiler output and the governor-turbine system; their dynamics are governed by

$$\dot{x}_{18} = -\frac{1}{T_4} x_{18} + \frac{1}{T_4} f_2(x_{22}, x_{23}) \quad (F-15)$$

$$\dot{x}_{19} = \frac{1}{T_5} x_{18} - \frac{1}{T_5} x_{19} \quad (F-16)$$

$$\dot{x}_{20} = \frac{1}{T_6} x_{19} - \frac{1}{T_6} x_{20} \quad (F-17)$$

$$\dot{x}_{21} = K_1 K_3 x_4 - K_2 K_3 x_{18} - \frac{K_1 K_3}{R_1} x_{30} \dots \quad (F-18)$$

$$\begin{aligned} \dot{x}_{22} = & \frac{K_1 K_4}{T_3} x_4 - \frac{K_2 K_4}{T_3} x_{18} + \frac{1}{T_3} x_{21} - \frac{1}{T_3} x_{22} \\ & - \frac{K_1 K_4}{R_1 T_3} x_{30} \dots \end{aligned} \quad (F-19)$$

$$\dot{x}_{23} = \frac{1}{C_D} x_5 + \frac{1}{C_D} x_{17} - \frac{1}{C_D} f_2(x_{22}, x_{23}) \dots \quad (F-20)$$

The excitation system and generator models are taken from Ref. 37 and are identical to those given in Appendix A of this report. States x_{24} through x_{26} represent the excitation system and are governed by

$$\dot{x}_{24} = -\frac{1}{T_A} x_{24} - \frac{K_A}{T_A} x_{25} \dots \quad (F-21)$$

$$\dot{x}_{25} = -\frac{1}{T_F} x_{25} + \frac{K_F}{T_F} x_{26} \quad (F-22)$$

$$\dot{x}_{26} = \frac{1}{T_E} x_{24} - \frac{K_E}{T_E} x_{26} - \frac{1}{T_E} f_3(x_{26}) \quad (F-23)$$

where

$$f_3(x_{26}) = A e^{B x_{26}} \quad (F-24)$$

The generator is modeled by states x_{27} through x_{30} ; their dynamics are governed by

$$\dot{x}_{27} = -\frac{r}{L_q} x_{27} + \frac{L'_d}{L_q} x_{28} x_{30} + \frac{1}{L_q} x_{29} x_{30} \dots \quad (F-25)$$

$$\begin{aligned} \dot{x}_{28} = & -\frac{1}{\tau'_{do}} x_{26} - \frac{1}{L'_d} \left[r + \frac{(L_d - L'_d)}{\tau'_{do}} \right] x_{28} \\ & + \frac{L_1}{\tau'_{do}} x_{29} - \frac{q}{L'_d} x_{27} x_{30} \dots \end{aligned} \quad (F-26)$$

$$\dot{x}_{29} = \frac{1}{\tau'_{do}} x_{26} + \frac{(L_d - L'_d)}{\tau'_{do}} x_{28} - \frac{1}{\tau'_{do}} x_{29} \quad (F-27)$$

$$\begin{aligned} \dot{x}_{30} = & \frac{K_{HP}}{\tau_H} x_{18} + \frac{K_{IP}}{\tau_H} x_{19} + \frac{K_{LP}}{\tau_H} x_{20} - \frac{D}{\tau_H} x_{30} \\ & - \frac{1}{\tau_H} x_{27} x_{29} + \frac{(L - L'_d)}{\tau_H} x_{27} x_{28} \end{aligned} \quad (F-28)$$

Of the 30 state variables, 10 are artificially introduced to account for fuel-air transport delay and five others cannot be directly identified as distinct physical variables. Table F-1 gives the list of state variables that represent physical system variables. The state variables are all normalized as per unit quantities so that they have unit value in the steady state.

TABLE F-1
PHYSICAL STATES OF THE BTEG MODEL

STATE VARIABLE	DESCRIPTION
x_4	Reference power
x_5	Boiler charging - feedwater
x_{17}	Boiler charging - fuel and air
x_{18}	First stage turbine pressure
x_{19}	Second stage turbine pressure
x_{20}	Third stage turbine pressure
x_{22}	Throttle position
x_{23}	Boiler pressure
x_{24}	Regulator output voltage
x_{25}	Exciter derivative feedback
x_{26}	Exciter output voltage
x_{27}	q-axis generator current
x_{28}	d-axis generator current
x_{29}	RMS stator voltage
x_{30}	Electrical frequency

In the equations for the BTEG model, there are five types of non-linearities -- the throttle related non-linearities of Eqs. F-7 and F-12, the exciter

saturation characteristic of Eq. F-24 and the product non-linearities in Eqs. F-25, F-26 and F-28. These non-linearities were incorporated into the linear part of the dynamics using Taylor's series expansions about an assumed mean value of unity for the variables involved.

A single set of parameter values was chosen for the BTEG model from Refs. 35 through 37. Table F-2 lists these values. The linearized dynamics matrix resulting from this choice of parameters has the non-zero elements shown in Table F-3. The data of Table F-3 can be used for eigenanalysis of the BTEG system and for subsequent investigation of the SP formulation and methodology.

TABLE F-2
PARAMETER VALUES FOR BTEG MODEL

PARAMETER	VALUE	DESCRIPTION
C_1	1.0	LRM control feedback gain
C_2	1.0	Generator power signal gain
C_D	125.0	Boiler drum storage capacity (sec)
D_1	60.0	Fuel-air transport delay (sec)
K_{IC}	0.05	LRM control integral gain
K_{PC}	1.00	LRM control proportional gain
K_I	0.011	Pressure control reset gain
K_P	1.1	Pressure control proportional gain
K_{MWB}	1.0	Megawatt bias error gain
K_{FF}	1.0	Megawatt demand gain
K_F	1.0	Steam flow signal to boiler demand
K_{AUX1}	0.964	Fuel-air auxiliary gain
K_{AUX2}	0.950	Feedwater auxiliary gain
K_{W1}	1.5	Feedwater control proportional gain
K_{W2}	0.5	Feedwater control reset gain
T_{F1}	30.0	Fuel-air time constant (sec)
T_{F2}	30.0	Feedwater time constant (sec)
K_{SH}	0.08	Superheater frictional pressure drop coefficient
T_R	36.0	Pressure control time constant
R_2	0.05	Frequency error gain inverse
K	0.0	Drum/once-through weight
K_1	1.0	Reference power signal to governor
K_2	0.0	First stage pressure signal to governor
K_3	0.0	Governor integral gain

TABLE F-2
PARAMETER VALUES FOR BTEG MODEL (Continued)

PARAMETER	VALUE	DESCRIPTION
K_4	1.0	Governor proportional gain
R_1	0.05	Frequency error gain inverse
K_{HP}	0.28	High pressure fraction
K_{IP}	0.72	Intermediate pressure fraction
K_{LP}	0.00	Low pressure fraction
T_3	0.9	Governor time constant (sec)
T_4	0.95	Steam bowl time constant (sec)
T_5	8.0	Turbine first stage time constant (sec)
T_6	1.0	Turbine second stage time constant (sec)
A	0.115	Exciter saturation parameters
B	0.713	
K_A	30.0	Exciter amplifier gain
K_E	0.02	Exciter gain
K_F	0.05	Exciter stabilizer loop gain
T_A	0.40	Amplifier time constant (sec)
T_E	0.56	Exciter time constant (sec)
T_F	1.30	Stabilizer time constant (sec)
D	20.0	Generator damping constant
r	0.0016	Stator winding resistance
L_d	1.7	d-axis synchronous inductance
L'_d	0.256	d-axis transient inductance
L_q	1.62	q-axis synchronous inductance
τ'_{do}	4.8	d-axis transient time constant (sec)
τ_H	8.26	Rotor inertia time constant (sec)

TABLE F-3
NON-ZERO ELEMENTS OF BTEG DYNAMICS MATRIX

ROW	COLUMN	VALUE
1	2	1.0
1	22	-0.318463
1	23	2.297825
1	30	52.8
2	2	-0.277778
2	22	0.088831
2	23	-0.640944
2	30	-14.727778
3	22	-0.006032
3	23	0.043519
3	30	-1.0
4	3	1.0
4	4	-1.0

TABLE F-3

NON-ZERO ELEMENTS OF BTEG DYNAMICS MATRIX (Continued)

ROW	COLUMN	VALUE
4	22	-0.120630
4	23	0.870388
4	30	-20.00
5	1	0.0475
5	6	1.0
5	22	0.063029
5	23	-0.454778
5	30	-11.4
6	1	0.019673
6	5	-0.015833
6	6	-0.080833
6	22	0.026105
6	23	-0.188354
6	30	-4.721500
7	7	-0.033333
8	7	0.166667
8	8	-0.166667
9	8	0.166667
9	9	-0.166667
:	:	:
:	:	:
16	15	0.166667
16	16	-0.166667
17	16	0.166667
17	17	-0.166667
18	18	-1.052632
18	22	0.852709
18	23	0.916198
19	18	0.125
19	19	-0.125
20	19	1.0
20	20	-1.0
22	4	1.111111
22	21	1.111111
22	22	-1.111111
22	30	-22.222222
23	5	0.008
23	17	0.008
23	22	-0.006481
23	23	-0.006963
24	24	-2.5
24	25	-75.0
25	25	-0.769231
25	26	0.038462
26	24	1.785714
26	26	-0.262997
27	27	-0.000988
27	28	0.158025
27	29	0.617284
27	30	0.775309
28	26	-0.208333
28	27	-6.328125
28	28	-1.181380
28	29	0.208333
28	30	-6.328125
29	26	0.208333
29	28	0.300833
29	29	-0.208333
30	18	0.033898
30	19	0.087167
30	27	0.044068
30	28	0.165133
30	29	-0.121065
30	30	-2.421308

APPENDIX G

PROCESSES HAVING DIFFERENTIABLE SAMPLE PATHS

G 1 BACKGROUND

To begin, consider an Itô stochastic differential equation (Itô-SDE) of the general form

$$d\mathbf{x}(t) = \mathbf{f}(\mathbf{x}(t), t)dt + \mathbf{G}(\mathbf{x}(t), t)d\mathbf{w}(t) \quad (G.1-1)$$

where $\mathbf{w}$ is a Wiener process and where $\mathbf{f}$, $\mathbf{G}$, etc. were defined following Eq 2.4-2. It is well known that, subject to smoothness conditions on the coefficients, a unique sample path solution $\mathbf{x}$ exists almost surely (a.s.) for this equation and that the process $\mathbf{x}$ is an a.s. continuous Markov process (see, for example, Refs 11, 101 and 102).

In a vast range of applications, particularly in systems science and control engineering, partial observations are available on the underlying state process $\mathbf{x}$ generated by Eq. G.1-1. These observations are described by an equation of the form

$$d\mathbf{y}(t) = \mathbf{h}(\mathbf{x}(t), t)dt + \mathbf{D}(\mathbf{x}(t), t)d\mathbf{v}(t) \quad (G.1-2)$$

where $\mathbf{v}$ is another Wiener process possibly correlated with $\mathbf{w}$. If the coefficients in Eqs. G.1-1 and G.1-2 depend upon a control function $\mathbf{u}$ -- which depends non-anticipatively on the past of $\mathbf{y}$ -- it is possible to formulate a range of state estimation and optimal control problems concerning the behavior of $\mathbf{x}$ over some given interval of time (see e.g., Refs. 21, 103 and 104). For a large class of nonlinear systems, implicit solutions for the optimal control problem have been derived; they are in terms of the evolution of a conditional density for $\mathbf{x}(t)$ -- this being a function of observations on the past of $\mathbf{y}$, together with the Hamilton-Jacobi-Bellman equation for the cost-to-go. (See for example, Refs. 103 and 105). Specializing to the case where $\mathbf{f}$ and $\mathbf{h}$ depend linearly on $\mathbf{x}$ and $\mathbf{u}$, and making $\mathbf{G}$ and $\mathbf{D}$ independent of $\mathbf{x}$ and $\mathbf{u}$, gives the linear system

$$d\mathbf{x}(t) = \mathbf{F}(t)\mathbf{x}(t)dt + \mathbf{G}(t)\mathbf{u}(t)dt + \mathbf{M}(t)d\mathbf{w}(t) \quad (G.1-3)$$

$$d\mathbf{y}(t) = \mathbf{H}(t)\mathbf{x}(t)dt + \mathbf{D}(t)\mathbf{u}(t)dt + \mathbf{N}(t)d\mathbf{v}(t) \quad (G.1-4)$$

For this familiar case, state estimates are readily obtained by use of the well known Kalman filter (Ref. 6); moreover, when quadratic cost criteria on the states and controls are employed, the optimal feedback control law is given via a linear function of the state estimate (Refs. 21, 103 and 104). A central feature of these optimal estimation and control problems is that the filter and the state feedback equations involve the solution of a pair of dual (deterministic) Riccati equations. These generate the estimation error covariance matrix and a matrix quantity known as the cost index, respectively.

The tractability of the estimation and control problem for the system described by Eqs G 1-1 and G.1-2 depends upon the fact that the state process $\mathbf{x}$ is Markovian -- i.e., for $s < t$,

$$P(\mathbf{x}(t) \in A | \{\mathbf{x}(\tau), 0 \leq \tau \leq s\}) = P(\mathbf{x}(t) \in A | \{\mathbf{x}(s)\})$$

where $\{\cdot\}$ denotes the σ -field generated by the indicated random variable. Now, the defining property of a Wiener process is that it has an independent increment over any interval Δt with a Gaussian distribution of variance $(\Delta t)^{1/2}$. A well known consequence of these assumptions is that the sample paths are almost surely nowhere differentiable (Refs. 101 and 106). In fact the random variable

$$\frac{\mathbf{w}(t+dt) - \mathbf{w}(t)}{dt}$$

has variance proportional to $(dt)^{-1/2}$. This infinitely rough feature of $\mathbf{w}$ and $\mathbf{v}$ is inherited by the processes, $\mathbf{x}$ and $\mathbf{y}$.

G.2 CONDITIONS FOR SAMPLE PATH DIFFERENTIABILITY

Assume a model of the form given in Eqs G 1-3 and G.1-4, but let the process $\mathbf{w}$ be replaced by a more general Gaussian stationary increment process $\mathbf{m}$. That is to say, assume $\mathbf{m}$ is a zero-mean Gaussian process such that

$$\begin{aligned} r_m(t_1, s_1; t_2, s_2) &\triangleq E\{[(\mathbf{m}(t_1) - \mathbf{m}(s_1))][\mathbf{m}(t_2) - \mathbf{m}(s_2)]^T\} \\ &= E\{[\mathbf{m}(t_1+h) - \mathbf{m}(s_1+h)][\mathbf{m}(t_2+h) - \mathbf{m}(s_1+h)]^T\} \\ &= r_m(t_1+h, s_1+h; t_2+h, s_2+h) \end{aligned} \quad (G.2-1)$$

By a standard result (e.g., see Doob, Ref. 14), when $\mathbf{m}$ is scalar it possesses a spectral representation in terms of an orthogonal increment process $\zeta(\cdot)$ evolving in the frequency domain. To be more specific

$$\begin{aligned} \mathbf{m}(t) - \mathbf{m}(s) &= \int_s^t d\mathbf{m}(\tau) \\ &= \int_{-\infty}^{\infty} \frac{e^{2\pi i t \lambda} - e^{2\pi i s \lambda}}{2\pi i \lambda} (1+\lambda^2)^{1/2} d\zeta(\lambda) \end{aligned} \quad (G.2-2)$$

with

$$E\{d\zeta(\lambda) d\zeta(\mu)\} = \begin{cases} 0 & \lambda \neq \mu \\ dH(\lambda) & \lambda = \mu \end{cases} \quad (G.2-3)$$

where H is a bounded, increasing function on $\mathbb{R}^1$. Extending Eqs G.2-2 and G.2-3 to the vector case and applying the results to Eq. G.2-1 produces

$$r_m(s_1, t_1; s_2, t_2) = \int_{-\infty}^{\infty} \frac{(e^{2\pi i t_1 \lambda} - e^{2\pi i s_1 \lambda})(e^{-2\pi i t_2 \lambda} - e^{-2\pi i s_2 \lambda})}{4\pi^2 \lambda^2} (1+\lambda^2) dH(\lambda) \quad (G.2-4)$$

where we assume the matrix inequality $H(\lambda+d\lambda) - H(\lambda) > 0$ holds for all $\lambda \in R^1$ and $d\lambda \in R_+$.

This spectral representation of a stationary increment process is a fundamental device and will be invoked many times in the sequel. It should be noted that the orthogonal increment or "white noise" case, where $\underline{m} = \underline{w}$, corresponds to

$$(1+\lambda^2) dH(\lambda) = (\text{constant}) d\lambda$$

Suppose that $\underline{m}$ is a "colored" noise process with rational spectrum (i.e., one whose stationary increments are correlated); then

$$(1+\lambda^2) dH(\lambda) = \Phi(\lambda) d\lambda \quad (G.2-5)$$

with $\Phi(\lambda)$ a rational matrix function of λ . In the following assume that $\Phi(\lambda)$ is bounded at ∞ . In this case we can find a spectral factor $Z(i\lambda)$ of $\Phi(\lambda)$ (see Refs. 107 and 108), which has the property that

$$\Phi(\lambda) = Z(i\lambda) \Sigma Z^T(-i\lambda) \quad (G.2-6)$$

where $Z(s)$ is a rational $n \times n$ matrix with $Z(\infty) = I$, $Z(s)$ is analytic in $\text{Re}[s] \geq 0$ (i.e., Z is a stable linear system), and Σ denotes the innovations covariance matrix of the process $\underline{m}$. It is then possible to construct the so-called strong stochastic time-invariant realization of $\underline{m}$ (see e.g., Ref. 108). This is an almost sure sample path realization of $\underline{m}$ of the form

$$d\underline{a}(t) = \tilde{F}\underline{a}(t)dt + \tilde{G}d\tilde{w}(t) \quad (G.2-7)$$

$$d\underline{m}(t) = \tilde{H}\underline{a}(t)dt + d\tilde{w}(t) \quad (G.2-8)$$

where

$$Z(s) = \tilde{H}(sI - \tilde{F})^{-1} \tilde{G} + I$$

This yields the desired spectral representation of the stationary increment process $\underline{m}$; namely,

$$\underline{m}(t) - \underline{m}(s) = \int_{-\infty}^{\infty} \frac{e^{2\pi i t \lambda} - e^{2\pi i s \lambda}}{2\pi i \lambda} Z(i\lambda) d\xi(\lambda) \quad (G.2-9)$$

$$E\{d\xi(\lambda) d\xi^T(\mu)\} = \begin{cases} 0 & \lambda \neq \mu \\ \Sigma & \lambda = \mu \end{cases} \quad (G.2-10)$$

(See Ref. 90, or for the discrete time parameter case, Ref. 108.) By replacing $\underline{w}$ by $\underline{m}$ in Eqs. G.1-3 and G.1-4 this representation maneuver allows Eqs. G.1-3 and G.1-4 to be written

$$d \begin{bmatrix} \underline{x}(t) \\ \underline{a}(t) \end{bmatrix} = \begin{bmatrix} F(t) & M(t)\tilde{H} \\ 0 & \tilde{F} \end{bmatrix} \begin{bmatrix} \underline{x}(t) \\ \underline{a}(t) \end{bmatrix} dt + \begin{bmatrix} G(t) \\ 0 \end{bmatrix} \underline{u}(t)dt + \begin{bmatrix} M(t) \\ \tilde{G} \end{bmatrix} d\tilde{w}(t) \quad (G.2-11)$$

$$d\underline{y}(t) = [H(t) \quad 0] \begin{bmatrix} \underline{x}(t) \\ \underline{a}(t) \end{bmatrix} dt + D(t)\underline{u}(t)dt + N(t) d\underline{v}(t) \quad (G.2-12)$$

(The reader should note which of the above matrices are necessarily time-invariant and which are not.) This leads to a model for $\underline{y}$ which is driven by the Wiener processes $\tilde{w}$ and $\underline{v}$ and consequently the estimation and control theory alluded to earlier can be applied directly.

Exactly the same state extension trick works if the observation disturbance $\underline{v}(t)$ is such that its associated spectral process is given by

$$(1+\lambda^2) dH_v(\lambda) = \Psi(\lambda) d\lambda \quad (G.2-13)$$

where $\Psi(\lambda)$ satisfies the same conditions as $\Phi(\lambda)$, i.e., $\Psi(\lambda)$ possesses a factorization of the form given by Eq. G.2-6. From here on the increasing matrix function H corresponds to the $\underline{w}$ process or the joint $\begin{bmatrix} \underline{w} \\ \underline{v} \end{bmatrix}$ process depending upon the context.

The result of all these steps is to produce a stochastic state space realization for Eqs. G.1-3 and G.1-4, i.e., a model for $\underline{y}(t)$ driven by Wiener processes and possessing a Markovian state process. In general giving $\tilde{w}$ and $\underline{v}$ rational spectra, instead of constant spectra, does not result in smooth state and output processes. This is because a necessary and sufficient condition for a stationary increment process $\underline{m}$ to have a mean square differentiable sample path (and thus for Gaussian separable processes to be pathwise differentiable) is that (Ref. 14, p. 558)

$$\int_{-\infty}^{\infty} (1+\lambda^2) E\{d\xi(\lambda) d\xi^*(\lambda)\} = \int_{-\infty}^{\infty} (1+\lambda^2) dH(\lambda) < \infty \quad (G.2-14)$$

where $(\cdot)^*$ denotes complex conjugate transpose and Eq. G.2-14 will not hold in general.

If Eq. G.2-14 holds, then $d\underline{m}(t)/dt$ is meaningful and

$$\frac{d\mathbf{m}(t)}{dt} \text{ a.s. } = \int_{-\infty}^{\infty} e^{2\pi i t \lambda} (1+\lambda^2)^{1/2} d\zeta(\lambda) \quad (\text{G.2-15})$$

$$E \left\{ \frac{d\mathbf{m}(t)}{dt} \frac{d\mathbf{m}^T(s)}{dt} \right\} = \int_{-\infty}^{\infty} e^{2\pi i \lambda(t-s)} (1+\lambda^2) dH(\lambda) \quad (\text{G.2-16})$$

and, in particular,

$$E \frac{d\mathbf{m}(t)}{dt} \frac{d\mathbf{m}^T(t)}{dt} = \int_{-\infty}^{\infty} (1+\lambda^2) dH(\lambda) < \infty \quad (\text{G.2-17})$$

In the remainder of this section, it will be assumed that a density exists for the distribution H , thus, $(1+\lambda^2) dH(\lambda)$ can be replaced by $\Psi(\lambda)d\lambda$.

Observe that in general the presence of the identity term in the rational spectral density

$$\begin{aligned} \Phi(\lambda) &= Z(1\lambda) \Sigma Z^T(-1\lambda) = [H(I(1\lambda) - F)^{-1} G + I] \\ &\Sigma [H(I(-1\lambda) - F)^{-1} G + I]^T \end{aligned} \quad (\text{G.2-18})$$

prevents the integrability condition of Eq. G.2-14 from holding and, in particular, this applies to the Itô-SDE case where

$$(1+\lambda^2) dH = (\text{constant}) d\lambda$$

This may be interpreted in the sense that the Wiener process driven Itô equations generate "infinitely rough" processes that are consequently not mean square differentiable. However, the rational spectral density

$$\Phi(\lambda) = 1/(1+\lambda^2)$$

satisfies

$$\int_{-\infty}^{\infty} (1+\lambda^2) dH(\lambda) = \int_{-\infty}^{\infty} \Phi(\lambda) d\lambda = \int_{-\infty}^{\infty} \frac{d\lambda}{(1+\lambda^2)} < \infty$$

and provides an example where a rational spectral density which decays sufficiently fast at ∞ to satisfy Eq. G.2-14.

Two methods are now available to study processes with a.s. differentiable sample paths which are generated by linear time-invariant, finite-dimensional systems:

- Impose the conditions (i) $(1+\lambda^2) dH =$ (rational function of λ) $d\lambda$ and (ii) Eq. G.2-14 holds on an initial model of the form given by Eqs G.1-3 and G.1-4. Then transform to equations of the form given by Eqs G.2-11 and G.2-12 via the factorization techniques described above.

- Remain with the initial model of Eqs G.1-3 and G.1-4 and impose Eq. G.2-14 on $\underline{w}$ and $\underline{v}$, on H

The first approach has the feature that the state estimation and control theory can be applied directly, however the second has the advantage that one explicitly retains the original system equations and these may be in a form of physical or engineering significance. Furthermore, the possibility of using irrational spectra for $\underline{w}$ and $\underline{v}$ is retained, permitting one to choose spectra which decay rapidly outside some given frequency range

In the remainder of this appendix, the second alternative will be used because of the advantages just described. Before proceeding, it should be noted that there are some curious consequences for the process $\underline{x}$ arising from the assumption $\Phi(\lambda) = 0$ for $|\lambda| > \lambda_0 > 0$, for some λ_0 , when F is asymptotically stable and $\underline{u}$ takes, say, the value 0. In this case the process $\underline{x}(t+h) - \underline{x}(t)$ generated by

$$\begin{aligned} \underline{x}(t+h) - \underline{x}(t) &= \int_{-\infty}^{\infty} \frac{e^{2\pi i(t+h)\lambda} - e^{2\pi i t \lambda}}{2\pi i \lambda} \\ &\times (I(1\lambda) - F)^{-1} (1+\lambda^2)^{1/2} d\zeta(\lambda) \end{aligned} \quad (\text{G.2-19})$$

is stationary (as a function of t) with spectral density

$$\begin{aligned} \Psi(\lambda) &= \frac{|e^{2\pi i h \lambda} - 1|^2}{4\pi^2 \lambda^2} (1+\lambda^2) (I(1\lambda) - F)^{-1} \Phi(\lambda) \\ &\times (I(-1\lambda) - F)^{-T} \end{aligned} \quad (\text{G.2-20})$$

Since this vanishes outside $|\lambda| > \lambda_0$, the process $\underline{x}(t+h) - \underline{x}(t)$ can be predicted with zero mean square error given observations on the past of $\underline{x}(t+h) - \underline{x}(t)$ over the interval $(-\infty, t_0]$. This follows from the Szegő formula for the linear least squares prediction error. The reader is referred to Ref. 109 for the relevant analysis and for some examples, and to Ref. 108 for a presentation of the corresponding discrete time parameter theory.

G.3 COVARIANCE FUNCTIONS

In the preceding section, conditions on the stationary increment processes $\underline{m}(t)$ and $\underline{n}(t)$ have been developed such that solutions to

$$d\underline{x}(t) = F\underline{x}(t)dt + G\underline{u}(t)dt + d\underline{m}(t) \quad (\text{G.3-1})$$

$$d\underline{y}(t) = H\underline{x}(t)dt + D\underline{u}(t)dt + d\underline{n}(t) \quad (\text{G.3-2})$$

have differentiable sample paths. In this section the equations for the covariance functions of $\underline{x}(t)$ and $\underline{y}(t)$ are developed; these functions are needed in the discussion of the surface-crossing problem of Section 7.2.

Some remarks on the existence and uniqueness of the solutions to Eqs. G.3-1 and G.3-2 are appropriate. In these equations, $\underline{m}$ and $\underline{n}$ are taken to be separable zero-mean, Gaussian, stationary increment processes, defined, along with $\underline{x}(t_0)$, on the underlying probability space $(\Omega, \mathcal{F}, P)$. Assume that $\underline{m}$ and $\underline{n}$ are independent of $\underline{x}(t_0)$, possess a spectral density matrix $\Psi(\lambda) > 0$ with block components $\Psi_m(\lambda)$, $\Psi_n(\lambda)$, $\Psi_{mn}(\lambda)$, etc. and that $\Psi(\lambda)$ satisfies

$$\int_{-\infty}^{\infty} (1+\lambda^2) dH \left[\frac{w}{v} \right] (\lambda) = \int_{-\infty}^{\infty} \Psi(\lambda) d\lambda < \infty \quad (G.3-3)$$

The existence of unique solutions to Eqs. G.3-1 and G.3-2 is shown as follows. As stated previously, Eq. G.3-3 implies that the sample paths of $\underline{m}$ and $\underline{n}$ are almost surely differentiable. It follows from the construction of solutions to Itô-SDEs that Eqs. G.3-1 and G.3-2 have solution processes identical to the solutions of the ordinary differential equations

$$\frac{d\underline{x}(t)}{dt} = F(t)\underline{x}(t) + G(t)\underline{u}(t) + \frac{d\underline{m}(t)}{dt} \quad (G.3-4)$$

$$\frac{d\underline{y}(t)}{dt} = H(t)\underline{x}(t) + D(t)\underline{u}(t) + \frac{d\underline{n}(t)}{dt} \quad (G.3-5)$$

if these exist. However, on the subset of Ω of probability one on which $\frac{d\underline{m}(t, \omega)}{dt}$ and $\frac{d\underline{n}(t, \omega)}{dt}$ exist for all $t \in [t_0, T]$, the right hand sides of Eqs. G.3-4 and G.3-5 satisfy a Lipschitz condition in $\underline{x}$ which holds uniformly in t . As a result, the classical theorems (Ref. 110) on the existence and uniqueness of solutions to ordinary differential equations apply. This establishes the a.s. existence and uniqueness of solutions to Eqs. G.3-1 and G.3-2.

The remainder of this section deals with computing the covariance functions of the processes $\underline{x}$ and $\underline{y}$ under the assumption that $\underline{m}$ and $\underline{n}$ are mutually orthogonal processes, an assumption which is made purely for convenience and which can be relaxed subject to an increase in the complexity of the resulting formulae.

Assume that $\underline{x}(t_0)$ has zero mean and for convenience, set $t_0 = 0$. Since $\underline{m}$ and $\underline{n}$ have zero mean, so do their contributions to $\underline{x}$ and $\underline{y}$. Thus, the mean value of the $\underline{x}$ process is given by

$$E\{\underline{x}(t)\} = \int_0^t \Phi(t, s) G(s) \underline{u}(s) ds \quad (G.3-6)$$

where

$$\frac{\partial \Phi}{\partial t}(t, s) = F(t) \Phi(t, s); \quad \Phi(s, s) = I, \quad (G.3-7)$$

i.e., $\Phi(\cdot, \cdot)$ is the fundamental matrix of $F(t)$ for $t \in [0, \infty)$. If we assume $\underline{u} \equiv 0$, then $\underline{x}$ has zero mean and the formulae simplify in a convenient manner. Subject to this assumption, the covariance function is obtained as follows.

$$\begin{aligned} R(t, s) &\triangleq E\{\underline{x}(t)\underline{x}^T(s)\} \\ &= E\left\{\left[\Phi(t, 0)\underline{x}(0) + \int_0^t \Phi(t, u) \frac{d\underline{m}(u)}{du} du\right] \right. \\ &\quad \left. \left[\Phi(s, 0)\underline{x}(0) + \int_0^s \Phi(s, v) \frac{d\underline{m}(v)}{dv} dv\right]^T\right\} \\ &= \Phi(t, 0) E\{\underline{x}(0)\underline{x}^T(0)\} \Phi^T(s, 0) \\ &\quad + \int_0^t \int_0^s \Phi(t, u) E\left\{\frac{d\underline{m}(u)}{du} \frac{d\underline{m}^T(v)}{dv}\right\} \\ &\quad \Phi^T(s, v) du dv \\ &= \Phi(t, 0) V_0 \Phi^T(s, 0) \\ &\quad + \int_0^t \int_0^s \Phi(t, u) E\left\{\int_{-\infty}^{\infty} e^{2\pi i u \lambda} \right. \\ &\quad \left. (1+\lambda^2)^{1/2} d\underline{\xi}(\lambda) \right. \\ &\quad \left. \times \int_{-\infty}^{\infty} e^{-2\pi i v \mu} (1+\mu^2)^{1/2} d\underline{\xi}^*(\mu) \right\} \\ &\quad \Phi^T(s, v) du dv \\ &= \Phi(t, 0) V_0 \Phi^T(s, 0) \\ &\quad + \int_0^t \int_0^s \Phi(t, u) \left[\int_{-\infty}^{\infty} e^{2\pi i (u-v)\lambda} \Psi(\lambda) d\lambda \right] \\ &\quad \Phi^T(s, v) du dv \quad (G.3-8) \end{aligned}$$

Equation G.3-8 is invoked frequently in the sequel. One important feature that should be noted is

that the second mixed partial derivative $\frac{\partial^2 R(t, s)}{\partial t \partial s}$ ex-

ists and is finite along the diagonal in R^2 . This fact is equivalent to the mean square differentiability of the solution trajectories $\{\underline{x}(t); t > 0\}$ (Ref. 14). In fact, observe that $R(t, s)$ in Eq. G.3-8 has continuous mixed partial derivatives of all orders at all points in R^2 . (An elementary example of a non-second order differentiable covariance corresponding to a non-mean-square differentiable process is the covariance $r(t, s) = \min(t, s)$ of the Wiener process.)

It is now possible to derive an integro-differential equation for $R(t,t)$ and expressions for $\frac{\partial R(t,s)}{\partial s}$ and $\frac{\partial^2 R(t,s)}{\partial t \partial s}$. From Eq. G.3-8,

$$R(t,t) = \Phi(t,0) V_0 \Phi^T(t,0) + \int_{-\infty}^{\infty} \left[\int_0^t \Phi(t,u) e^{2\pi i u \lambda} du \right] \Psi(\lambda) \left[\int_0^t \Phi^T(t,v) e^{-2\pi i v \lambda} dv \right] d\lambda \quad (G.3-9)$$

Now,

$$\begin{aligned} \frac{dR(t,t)}{dt} &= F(t) [\Phi(t,0) V_0 \Phi^T(t,0)] \\ &+ [\Phi(t,0) V_0 \Phi^T(t,0)] F^T(t) \\ &+ \int_{-\infty}^{\infty} \left[e^{2\pi i t \lambda} + \int_0^t F(t) \Phi(t,u) e^{2\pi i u \lambda} du \right] \Psi(\lambda) \left[\int_0^t \Phi^T(t,v) e^{-2\pi i v \lambda} dv \right] d\lambda \\ &+ \int_{-\infty}^{\infty} \left[\int_0^t \Phi(t,u) e^{2\pi i u \lambda} du \right] \Psi(\lambda) \left[e^{-2\pi i t \lambda} + \int_0^t F(t) \Phi(t,v) e^{-2\pi i v \lambda} dv \right]^T d\lambda \\ &= F(t) R(t,t) + R(t,t) F^T(t) \\ &+ \int_0^t \left[\int_{-\infty}^{\infty} e^{2\pi i (t-v) \lambda} \Psi(\lambda) d\lambda \right] \Phi^T(t,v) dv \\ &+ \int_0^t \Phi(t,u) \left[\int_{-\infty}^{\infty} e^{2\pi i (u-t) \lambda} \Psi(\lambda) d\lambda \right] du \end{aligned} \quad (G.3-10)$$

But by Eq. G.2-16

$$\int_{-\infty}^{\infty} e^{2\pi i (t-v) \lambda} \Psi(\lambda) d\lambda = R_m^*(t-v) = R_m^T(v-t) \quad (G.3-11)$$

where $R_m^*(\cdot)$ denotes the covariance function of the stationary process $\frac{dm}{dt}$. This yields the following integro-differential equation for $R(t,t)$

$$\begin{aligned} \frac{dR(t,t)}{dt} &= F(t) R(t,t) + R(t,t) F^T(t) \\ &+ \int_0^t \Phi(t,u) R_m^T(t-u) du \\ &+ \int_0^t R_m^*(t-v) \Phi^T(t,v) dv \end{aligned} \quad (G.3-12)$$

Similarly, it may be seen that

$$\begin{aligned} \frac{\partial R(t,s)}{\partial t} &= F(t) \Phi(t,0) V_0 \Phi^T(s,0) \\ &+ F(t) \left\{ \int_0^t \Phi(t,u) \left[\int_{-\infty}^{\infty} (1+\lambda^2) e^{2\pi i u \lambda} dH(\lambda) du \right] \right. \\ &\quad \times \left. \int_0^s \Phi^T(s,v) e^{-2\pi i v \lambda} dv \right. \\ &\quad \left. + \int_0^s \int_{-\infty}^{\infty} e^{2\pi i t \lambda} e^{-2\pi i v \lambda} \Psi(\lambda) d\lambda \Phi^T(s,v) dv \right\} \\ &= F(t) R(t,s) + \int_0^s R_m^*(t-v) \Phi^T(s,v) dv \end{aligned} \quad (G.3-13)$$

with

$$\left. \frac{\partial R(t,s)}{\partial t} \right|_{t=s} = \left. \frac{\partial R(t,s)}{\partial s} \right|_{t=s}$$

and

$$\begin{aligned} \frac{\partial^2 R(t,s)}{\partial t \partial s} &= \frac{\partial^2 R(t,s)}{\partial s \partial t} \\ &= F(t) \left[R(t,s) F^T(s) + \int_0^t \Phi(t,u) R_m^*(u-s) du \right] \\ &+ R_m^*(t-s) \Phi^T(s,s) \\ &+ \int_0^s R_m^*(t-v) \Phi^T(s,v) F^T(s) dv \end{aligned}$$

$$\begin{aligned}
&= F(t) R(t,s) F^T(s) + F(t) \left[\int_0^t \Phi(t,u) R_m^T(s-u) du \right] \\
&+ \left[\int_0^s R_m(t-v) \Phi^T(s,v) dv \right] F^T(s) \\
&+ R_m(t-s). \qquad \qquad \qquad (G.3-14)
\end{aligned}$$

Equations G.3-9 and G.3-12 through G.3-14 constitute the full set of functions of the process covariances that are needed for the surface crossing problem in Section 7.2.

REFERENCES

1. Schweppe, F.C., and Wildes, J., "Power System Static-State Estimation, Part I: Exact Model," IEEE Transactions on Power Apparatus and Systems, Vol PAS-89, January 1970, pp 120-125
2. Schweppe, F.C., and Rom, D.B., "Power System Static-State Estimation, Part II: Approximate Model," IEEE Transactions on Power Apparatus and Systems, Vol PAS-89, January 1970, pp 125-130.
3. Schweppe, F.C., "Power System Static-State Estimation, Part III: Implementation," IEEE Transactions on Power Apparatus and Systems, Vol PAS-89, January 1970, pp 130-135.
4. Debs, A.S., and Larson, R.E., "A Dynamic Estimator for Tracking the State of a Power System," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-89, September/October 1970, pp. 1670-1678.
5. Schweppe, F.C., and Handschin, E.J., "Static State Estimation in Electric Power Systems," Proceedings of the IEEE, Vol. 62, July 1974, pp. 972-982.
6. Gelb, A., ed, Applied Optimal Estimation, M.I.T Press, Cambridge, 1974.
7. Jazwinski, A.H., Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.
8. McGarty, T.P., Stochastic Systems and State Estimation, Wiley-Interscience, New York, 1974.
9. Bucy, R.S., and Senne, K.D., "Digital Synthesis of Nonlinear Filters," Automatica, Vol. 7, May 1971, pp. 287-298
10. Alspach, D.L., "Practical Control Algorithms for Nonlinear Stochastic Systems and Investigations of Nonlinear Filters," ORICON Corporation, Report No. OC-R-79-0002AA-1, March 1979.
11. Arnold, L., Stochastic Differential Equations: Theory and Applications, Wiley-Interscience, New York, 1974.
12. Snyder, D.L., Random Point Processes, Wiley-Interscience, New York, 1975.
13. Gihman, I.I., and Skorohod, A.V., Stochastic Differential Equations, Springer-Verlag, New York, 1972.
14. Doob, J.L., Stochastic Processes, Wiley, New York, 1953.
15. Meyer, P.A., Probability and Potentials, Blaisdell, Waltham, MA, 1966.
16. Liptser, R.S., and Shirayev, A.N., Statistics of Random Processes I - General Theory, Springer-Verlag, New York, 1977.
17. Neveu, J., Discrete-Parameter Martingales, Elsevier (North-Holland), New York, 1975.
18. Wonham, W.M., "Some Applications of Stochastic Differential Equations to Optimal Nonlinear Filtering," SIAM Journal on Control, Vol. 2, 1965, pp. 347-369.
19. Rozovskii, B.L., and Shirayev, A.N., "On Infinite Order Systems of Stochastic Differential Equations Arising in the Theory of Optimal Nonlinear Filtering," Theory of Probability and its Applications, Vol XVII, No 2, 1972, pp. 218-226.
20. Takeuchi, Y., and Akashi, H., "Nonlinear Filtering Formulas for Discrete-Time Observations," SIAM Journal on Control and Optimization, Vol. 19, No. 2, 1981, pp. 244-261
21. Davis, M.H.A., Linear Estimation and Stochastic Control, Halsted Press, New York, 1977.
22. Magill, D.T., "Optimal Adaptive Estimation of Sampled Stochastic Processes," IEEE Transactions on Automatic Control, Vol. AC-10, October 1965, pp. 434-439
23. Lainiotis, D.G., "Optimal Adaptive Estimation: Structure and Parameter Adaptation," IEEE Transactions on Automatic Control, Vol. AC-15, April 1971, pp. 160-170.
24. Athans, M., and Chang, C.B., "Adaptive Estimation and Parameter Identification Using Multiple Model Estimation Algorithm," Lincoln Laboratory Technical Note No 1976-28, June 1976
25. Nayfeh, A.H., Perturbation Methods, John Wiley and Sons, New York, 1973
26. Kokotovic, P.V., O'Malley, R.E., and Sannuti, P., "Singular Perturbations and Order Reduction in Control Theory -- An Overview," Automatica, Vol. 12, 1976, pp. 123-132.
27. Blankenship, G.L., and Fink, L.H., "Statistical Characterizations of Power System Stability and Security," Proceedings Second Lawrence Symposium on System and Decision Sciences, Berkeley, 1978
28. Concordia, C., and Schulz, R.P., "Appropriate Component Representation for the Simulation of Power System Dynamics," IEEE Report No 75CH0970-4-PWR, Symposium on Adequacy and Philosophy of Modeling: Dynamic System Performance, 1975, pp. 16-23.
29. Chow, J.H., Allemong, J.J., and Kokotovic, P.V., "Singular Perturbation Analysis of Systems with Sustained High-Frequency Oscillations," Automatica, Vol. 14, 1978, pp. 271-279.
30. Winkelman, J.R., Chow, J.H., Allemong, J.J., and Kokotovic, P.V., "Singular Perturbation Methods in Multiple Time-Scale Power System Simulation, Part I: Analytical Development, Part II: Application Experience," Proceedings of the Twelfth Asilomar Conference on Circuits, Systems, & Computers, Pacific Grove, 1978.
31. Kokotovic, P.V., Allemong, J.J., Winkelman, J.R., and Chow, J.H., "Singular Perturbation and Iterative Separation of Time Scales," Automatica, Vol. 16, 1980, pp. 23-33.
32. Winkelman, J.R., Chow, J.H., Allemong, J.J., and Kokotovic, P.V., "Multi-Time-Scale Analysis of a Power System," Automatica, Vol. 16, 1980, pp. 35-43.
33. Acharya, N.A., "Singularly Perturbed Estimation Using the Method of Multiple Scales," Ph.D., Dissertation, M.E. Department, University of Massachusetts, Amherst, February 1979.

34. Wasow, W.A., "Asymptotic Expansions for Ordinary Differential Equations", John Wiley and Sons, New York, 1965.
35. Schulz, R.P., Turner, A.E., and Ewart, D.N., "Long Term Power System Dynamics," Electric Power Research Institute, Report No. EPRI 907-0, Vol. I, June 1974.
36. "Midterm Simulation of Electric Power Systems," Electric Power Research Institute, Report No. EL-596, June 1979.
37. Anderson, P.M., and Fouad, A.A., Power System Control and Stability, The Iowa State University Press, Ames, 1977.
38. Perez-Arriaga, I.J., Verghese, G.C., and Schweppe, F.C., "Selective Modal Analysis with Applications to Electric Power Systems," Proceedings of the IEEE Power Engineering Society, Winter Meeting, Paper Nos. 82WM070-1 and 82WM071-9, 1982.
39. Ueda, R., et al., "On the Estimation of Transient State of Power System by Discrete Nonlinear Observer," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-94, November/December 1975, pp. 2135-2140.
40. Ueda, R., et al., "Estimation of Transient State of Multi-Machine Power System by Extended Linear Observer," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-96, March/April 1977, pp. 681-687.
41. Takata, H., et al., "An Iterative Sequential Observer for Estimating the Transient States of a Power System," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-96, March/April 1977, pp. 673-680.
42. Modir, H., and Schlueter, R., "A Dynamic State Estimator for Dynamic Security Assessment," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-100, November 1981, pp. 4644-4652.
43. Tarn, T.J., and Rasis, Y., "Observers for Nonlinear Systems," IEEE Transactions on Automatic Control, Vol. AC-21, August 1976, pp. 441-448.
44. Sprevak, D., and Newmann, M.M., "On the Use of Stochastic Approximation in the State-Parameter Identification Problem," Electronics Letters, Vol. 14, June 1978, p. 351-352.
45. Dvoretzky, A., "On Stochastic Approximation," Proceedings of the Third Berkeley Symposium on Mathematical Statistics and Probability, J. Neyman, ed. University of California Press, 1956, p. 39-55.
46. Pearson, J.B., "A Note on Nonlinear Filtering," IEEE Transactions on Automatic Control, Vol. AC-13, February 1968, pp. 103-105.
47. Wolverton, C.T., "A Note on the Application of Dvoretzky's Theorem to Nonlinear Filtering," IEEE Transactions on Automatic Control, AC-14, February 1969, p. 117.
48. Albert, A.E., and Gardner, L.A., Jr., Stochastic Approximation and Nonlinear Regression, MIT Press, Cambridge, 1967.
49. Fujita, S., and Fukao, T., "Convergence Conditions of Dynamic Stochastic Approximation Method for Nonlinear Stochastic Discrete-Time Dynamic Systems," IEEE Transaction Automatic Control, Vol. AC-17, October 1972, pp. 715-717.
50. de Figueiredo, R.J.P., Dyer, L., and Netravali, A.N., "Estimation of the States and Parameters of Nonlinear Dynamical Systems by Stochastic Approximation Methods," Proceedings of the Symposium on Nonlinear Estimation Theory and its Applications, San Diego, California, 1970, pp. 196-206.
51. Smithies, F., Integral Equations, Cambridge University Press, Cambridge, 1958.
52. Bellman, R., Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.
53. Bharucha, B.H., "On the Stability of Randomly Varying Systems," Ph.D. Thesis, Department of Electrical Engineering, University of California (Berkeley), 1961.
54. Brockett, R.W., "Lie Theory and Control Systems Defined on Spheres," SIAM Journal on Applied Mathematics, Vol. 25, September 1973, pp. 213-225.
55. Fink, L.H., and Carlsen, K., "Operating Under Stress and Strain," IEEE Spectrum, March 1978.
56. Blake, I.F., and Lindsey, W.C., "Level-Crossing Problems for Random Processes," IEEE Transactions on Information Theory, Vol. IT-19, May 1973, pp. 295-315.
57. Cramér, H., and Leadbetter, M.R., Stationary and Related Stochastic Processes, John Wiley and Sons, New York, 1967.
58. Spanier, E.N., Algebraic Topology, McGraw-Hill, New York, 1966.
59. Rockafellar, R.T., Convex Analysis, Princeton University Press, Princeton, 1970.
60. Mitchell, A.R., "Computational Methods in Partial Differential Equations," John Wiley and Sons, New York, 1969.
61. Smith, G.D., "Numerical Solution of Partial Differential Equations," Oxford University Press, 1975.
62. Gladwell, I. and Wait, R., Editors. "A Survey of Numerical Methods for Partial Differential Equations," Clarendon Press, Oxford, 1979.
63. Cesari, L., Asymptotic Behavior and Stability Problems in Ordinary Differential Equations, Springer-Verlag, New York, 1959.
64. Wintner, A., "Asymptotic Integration Constants," American Journal of Mathematics, Vol. 68, 1946, pp. 553-559.
65. Chung, K.L., Markov Chains With Stationary Transition Probabilities, Springer-Verlag, New York, 1967.
66. Infante, E.F., "On the Stability of Some Linear Non-Autonomous Random Systems," ASME Journal of Applied Mechanics, Vol. 35, March 1968, pp. 7-12.

67. Lüders, G., "Transient Stability of Multimachine Power Systems via the Direct Method of Lypaunov," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-90, January/February 1971, pp. 23-36.
68. Athay, T., et al. "A Practical Method for the Direct Analysis of Transient Stability," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-98, March/April 1979, pp. 573-584.
69. Ribbens-Pavella, M., et al., "Transient Stability Index for On-Line Evaluation," Proceedings of the IEEE Power Engineering Society Winter Meeting, Paper No. A80013-3, 1980.
70. Krishnaparandhama, T., et al. "Method for Identifying Coherent Generators," Electrical Power and Energy Systems, Vol. 3, April 1981, pp. 85-90.
71. Prabhakara, F.S., and El-Abiad, A.H., "A Simplified Determination of Transient Stability Regions for Lyapunov Methods," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-94, March/April 1975, pp. 672-689.
72. Donalek, P.J., and Whysong, J.L., "Reliability Information for Electric Utility Transmission and Distribution Systems," Harza Engineering Co., Report for Argonne National Laboratory, Report No. ANL/EES-TM-31, December 1978.
73. Brockett, R.W., and Blankenship, G.L., "A Representation Theorem for Linear Differential Equations with Markovian Coefficients," 1977 Allerton Conference on Circuits and System Theory, Urbana, Illinois, 1977.
74. Barnett, V., and Lewis, T., Outliers in Statistical Data, John Wiley and Sons, New York, 1978.
75. Duran, B.S., and Odell, P.L., Cluster Analysis - A Survey, Springer-Verlag, New York, 1974.
76. Meisel, J., "Dynamic and Transient Stability Augmentation," Engineering Foundations Conference, Systems Engineering for Power: Status and Prospects, Henniker, New Hampshire, 1975, pp. 454-481.
77. Lee, S.T.J., and Schweppe, F.C., "Distance Measures and Coherency Recognition Transient Stability Equivalents," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-92, September/October 1973, pp. 1550-1557.
78. Podmore, R., "Identification of Coherent Generators for Dynamic Equivalents," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-97, July/August 1978, pp. 1344-1354.
79. Schlueter, R.A., Akhtar, H., and Modir, H., "An RMS Coherency Measure: A Basis for Unification of Coherency Model Analysis Model Aggregation Techniques," Proceedings of the IEEE Power Engineering Society Summer Meeting, Paper No. A78 533-2, 1978.
80. Schlueter, R.A., and Akhtar, H., "A Transient Coherency Measure and Its Use to Analyze System Transient Response," Proceedings of the IEEE Power Engineering Society Summer Meeting, Paper No. A79 480-5, 1979.
81. Overall, J.E., and Klett, C.J., Applied Multivariate Analysis, McGraw-Hill, New York, 1972.
82. Van Ness, J.E., Boyle, J.M., and Imad, F.P., "Sensitivities of Large, Multiple-Loop Control Systems," IEEE Transactions on Automatic Control, Vol. AC-10, July 1965, pp. 308-315.
83. Nolan, P.J., Sinha, N.K., and Alden, R.T.H., "Eigenvalue Sensitivities of Power Systems Including Network and Shaft Dynamics," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-95, July/August 1976, pp. 1318-1324.
84. Fleming, W.H., and Pardoux, E., "Optimal Control for Partially Observed Diffusions," SIAM Journal on Control and Optimization, Vol. 20, March 1982, pp. 261-285.
85. Stagg, G.W., and El-Abiad, A.H., Computer Methods in Power System Analysis, McGraw-Hill, New York, 1968.
86. IEEE Committee Report, "Dynamic Models for Steam and Hydro Turbines in Power System Studies," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-92, December 1973, pp. 1904-1916.
87. McDonald, J.P., "Adequacy and Philosophy of Modeling, Dynamic System Performance Representation of the Boiler-Turbine," Symposium on Adequacy and Philosophy of Modeling Dynamic System Performance, IEEE Publication No. 75CH0970-4-PWR, 1975.
88. IEEE Committee Report, "Computer Representation of Excitation Systems," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-87, June 1978, pp. 1460-1464.
89. Galiana, F.D., Handschin, E., and Fiechter, A.R., "Identification of Stochastic Electric Load Models from Physical Data," IEEE Transactions on Automatic Control, Vol. AC-19, December 1974, pp. 887-893.
90. Lindquist, A., and Picci, G., "Realization Theory for Multivariate Stationary Gaussian Processes: I-State Space Construction," Proc 4th International Symposium on Mathematical Theory of Networks and Systems, July 1979, Delft, pp. 140-148.
91. Box, G.E.P., and Jenkins, G.M., Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco, 1970.
92. Bendat, J.S., and Piersol, A.G., Measurement and Analysis of Random Data, John Wiley and Sons, New York, 1966.
93. "NEPEX Metering and Telemetering Criteria," New England Power Exchange, West Springfield, MA., April 15, 1976.
94. Undrill, J.M., Casazza, J.A., Gulachenski, E.M., and Kirchmayer, L.K., "Electro-Mechanical Equivalents for Use in Power System Stability Studies," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-90, September/October 1971, pp. 2060-2071.
95. Germond, A.J., and Podmore, R., "Dynamic Aggregation of Generating Unit Models," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-97, July/August 1978, pp. 1060-1069.

LIST OF PUBLICATIONS

96. Price, W.W., Ewart, D.N., Gulachenski, E.M., and Silva, R.F., "Dynamic Equivalents from On-Line Measurements," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-94, July/August 1975, pp. 1349-1357.
97. Ibrahim, M.A.H., Mostafa, D.M., and El-Abiad, A.H., "Dynamic Equivalents Using Operating Data and Stochastic Modeling," IEEE Transactions on Power Apparatus and Systems, Vol. PAS-95, September/October 1976, pp. 1713-1922.
98. Schweppe, F.C., Uncertain Dynamic Systems, Prentice-Hall, Englewood Cliffs, 1973.
99. Lee, J.G., Vogt, W.G., and Mickle, M.H., "Optimal Decomposition of Large-Scale Networks," IEEE Transactions on Systems, Man, and Cybernetics, Vol. SMC-9, July 1979, pp. 369-375.
100. Takacs, L., "On Certain Sojourn Time Problems in the Theory of Stochastic Processes," Acta Mathematica (Academiae Scientiarum Hungaricae), Vol. 8, 1957, pp. 169-191.
101. McKean, H.P., Stochastic Integrals, Academic Press, New York, 1969.
102. Bucy, R.S. and Joseph, P.D., Filtering for Stochastic Processes with Applications to Guidance, Wiley-Interscience, New York, 1968.
103. Fleming, W.H., and Rishel, R.W., Deterministic and Stochastic Optimal Control, Springer-Verlag, New York, 1975.
104. Åström, K.J., Introduction to Stochastic Control Theory, Academic Press, New York, 1970.
105. Davis, M.H.A., and Varaiya, P.P., "Dynamic Programming Conditions for Partially Observable Stochastic Systems," SIAM Journal on Control, Vol. 11, May 1973, pp. 226-261.
106. Billingsley, P., Probability and Measure, John Wiley and Sons, New York, 1979.
107. Youla, D.C., "On the Factorization of Rational Matrices," IRE Transactions on Information Theory, Vol. IT-7, July 1961, pp. 172-189.
108. Caines, P.E., Linear Discrete Time Parameter Stochastic Systems, to be published by John Wiley and Sons.
109. Yaglom, A.M., An Introduction to Stationary Random Functions, Prentice-Hall, Englewood Cliffs, 1962.
110. Coddington, E.A. and Levinson, N., Theory of Ordinary Differential Equations, McGraw-Hill, New York, 1955.

Mealy, G.L., Lee, J.G and Acharya, N.A., "Power System Dynamic Modeling and State Estimation," Proceedings of the IEEE International Symposium on Circuits and Systems, (Chicago), April 1981, pp. 919-922.

Acharya, N.A., "Formulation of Singularly Perturbed Estimation Problems," Proceedings of the IEEE International Symposium on Circuits and Systems, (Chicago), April 1981, pp. 923-926.

Caines, P.E., "On the Hyperplane Crossing Problem for Processes Generated by State Space Systems," Proceedings on the 20th IEEE Conference on Decision and Control, (San Diego), December 1981, pp. 1370-1376.

RESEARCH PROGRESS IN
DYNAMIC SECURITY ASSESSMENT

U.S. Department of Energy
Contract No. DE-AC01-79-ET-29038

Period of Performance

September 1979 - December 1982

EXECUTIVE SUMMARY

ES 1 INTRODUCTION

Large scale electric power systems are continually increasing in complexity and interdependence. In the United States for example, there are approximately 6000 generating units and 800,000 km of bulk power transmission lines which, under the control of more than 100 independent control centers, supply electric energy in response to the demands of over 80 million customers. Under normal operating conditions, this complex system provides reliable electric service. However, there are times when an emergency develops which disrupts the balance between demand and supply (or load and generation) and requires intervention (either human or automatic) to restore a portion of the system to normal operation. It is unlikely that a single event can trigger such an emergency. Rather, it is due to a sequence of relatively small events which tend to "weaken" part of the system, making it susceptible to an otherwise moderate disturbance. In other words the power system is operating at low level of security and is thus more vulnerable to disturbances than it would be if it were in its normal robust state.

Digital computers have made possible on-line monitoring of power system security to identify circumstances where preventive control action must be taken to avoid hazardous operating conditions. This is formerly regarded as the security assessment function in power systems. In general terms, security assessment can be thought of as an algorithm which utilizes knowledge of the current operating state of the system, forecasts the future evolution of the system, assesses the likelihood of a security violation within a specified future time-period, and determines whether preventive control action should be taken. These functions imply that the security assessment algorithm must depend upon

- Knowledge of system dynamics
- Measurement of system variables
- Models for measurement uncertainty and system disturbances
- Well-defined criteria for security.

Because of power system complexity, present day security assessment algorithms have limitations dictated by available computer capability. In particular, they almost all use static models of the system, and evaluate security with respect to steady state operating limits on system variables. Furthermore, little attempt has been made to deal systematically with system disturbances, particularly sudden random events -- line outages, load changes, faults, etc.

The rapid pace of technological development in computer hardware, both with respect to hardware miniaturization and to various forms of parallel processing, implies that the affordable computational power available to the utilities will increase by orders of magnitude in the near future. At the same time, the narrowing gap between demand and supply of electric energy will require that operating safety margins be known more accurately in real time to avoid frequent load disruptions. This report describes the results of a three-year research program, sponsored by the Department of Energy, to develop an improved methodology for assessing power system security, which can take advantage of future computer technology, and

which can be made available in time to meet future utility needs.

Dynamic security assessment for electric power systems is a very complex and difficult problem. The magnitude of the difficulty arises from a compounding of the following issues:

- The system of interest is very large scale, dynamic, and highly nonlinear
- The inputs of interest are multidimensional statistical distributions of random events (jump processes) and perturbations
- Security assessment involves evaluating the robustness of the system over a future time-interval, which requires potentially complex and time-consuming forecasts of system behavior for all possible occurrences of discrete and random events
- Security assessment must be performed on-line.

Even assuming a substantial increase in computational capabilities over the next decade, a useful solution to the stated problem must, of necessity, incorporate efficient and effective techniques for handling the various components of the problem; i.e., dynamic state estimation, statistical modeling of events and perturbations, stochastic forecasting, and stochastic contingency evaluation. A key element of this solution is the formulation of an overall approach which is sufficiently general to incorporate all relevant components of the problem and, at the same time, is sufficiently structured to develop specific practical mathematical solution techniques.

The functional structure of TASC's approach to dynamic security assessment is illustrated in Fig. ES.1-1. It incorporates the following major elements:

- Stochastic filtering of measurements to obtain an estimate of the present state of the system
- Forecasting of the system state to provide an indication of future security needs
- Establishment of security measures that are functionals of the state estimate and models of jump processes (outages, etc.), and process noise (small random fluctuation of loads, etc.)
- Off-line classification of possible jump disturbances into a smaller number of aggregates or clusters.

ES.2 SUMMARY OF WORK PERFORMED

This report documents the results obtained during a three-year research effort into the development of a dynamic security assessment methodology for power systems. Progress was made in the areas of:

Figure ES.1-1 Functional Structure of TASC's Approach to Dynamic Security Assessment

- Power system modeling
- State estimation
- Structure decomposition
- State forecasting
- Clustering
- Security measure development.

Each of these topics is discussed briefly in the following paragraphs.

A detailed dynamic model of a multimachine power system was developed for use throughout this effort. The model includes a state variable description of the boiler-turbine, generator and excitation system. Two sets of equations were formulated -- one describing the fast response (on the order of seconds) associated with transient dynamics and the other modeling the slower long-term dynamics (on the order of minutes). The interconnection network is characterized by a system of simultaneous algebraic equations. Stochastic components of the model consist of unpredictable load disturbances and transmission line and/or generator outages. An integrated state space formulation of the power system model was also developed. This model combines the component dynamic models, the algebraic network equations and the load disturbances into a unified state space representation suitable for state estimation. A measurement model characterizing power system observations was also completed.

A state estimator for estimating the long-term dynamic behavior of the power system was specified. Since measurements of both the system state and network configuration are available, the long-term state estimator was partitioned into two separate (but not independent) algorithms -- the process state estimator and the structure (or network configuration) estimator. The process state estimation effort led to the development of an estimation algorithm identical in form to the extended Kalman filter, but containing a modified process noise driving term to account for the Poisson component of the process noise. The structure estimation algorithm is based on a two-stage approach: first, a maximum a posteriori estimate of the structure (as characterized by the topological state z) is obtained from available breaker status measurements. Then, a hypothesis testing approach is used to select the true network structure from the small number of

candidate configurations which result from the first stage.

Various approaches to estimator simplification were examined including nonlinear observers and stochastic approximation based techniques. The most promising techniques involve algorithms of the same form as the extended Kalman filter, but with gain sequences chosen in a simplified manner.

In the area of structure decomposition, two different approaches were investigated -- spatial decomposition and time-scale decomposition. Spatial network decomposition consists of partitioning the power system into a number of subsystems, each with a small number of generators and loads. A computer program was developed to implement an optimal network decomposition technique. This technique minimizes the interaction between the subnetworks. The program was used to decompose the IEEE 39-Bus New England Test System into six subsystems, each containing one or two generators. Time-scale decomposition, or the separation of the system into two subsystems, one containing states with slow dynamics and the other containing states with fast dynamics, was applied to the extended Kalman filter equations of the process state estimator. This approach to system order reduction was based on an application of singular perturbation theory. This methodology provides a tradeoff between accuracy and computational savings.

State forecasting provides information on the anticipated evolution of the power system state. Research in this area has resulted in the development of recursive expressions for the first and second moments of the system state under the assumption that there are a finite number of possible network configurations and that the transition from one configuration to another is governed by a homogeneous Poisson process.

Clustering techniques for reducing the large number of possible power system disturbances to a more manageable number were reviewed and a set of clustering criteria for the security measures of interest were proposed. These criteria were chosen so as to reflect a natural measure of closeness for each security measure.

Research into security measures has proceeded in three directions. The first area involves viewing the security assessment problem as a surface crossing problem for a stochastic process. The transition of the power system from the normal operating state to an emergency state in a period t is described by the trajectory of the system in a multidimensional space. The surface in this formulation separates the normal operating region from the emergency region. In order to obtain computationally tractable results, it was assumed that the surface could be approximated by a hyperplane. This allowed the application of scalar level crossing theory to obtain moments of the expected number of crossings of the hyperplane. An approximate expression for the hyperplane crossing probability in terms of these moments was also developed. An alternate method of analyzing this problem utilizes a partial differential equation of Dynkin which yields the moments of the first crossing time. Upper and lower bounds on the first and second (and in principal higher) moments were developed. These bounds provide estimates of the moments which are much simpler to compute than the solution to Dynkin's equation.

The second approach examined the long-term stability properties of the power system from the view-

point of stability theory. The basic method was to analyze the stability properties of an unforced linear system where the system coefficients are subject to future jumps. Under the assumption that the elements of the dynamics matrix are discrete-state Markov processes, a sufficient condition for almost sure asymptotic stability was developed.

The third area of research has led to the formulation of a security measure suitable for on-line assessment of transient stability. This stability measure involves computation of the probability that a Lyapunov function will exceed a critical value in a specified future time interval. Specific distributions are assumed for the faults and expected clearing times. To alleviate the computational burden, the security measure is divided into a real-time and a non-real-time component. The non-real-time computations are based on a set of network faults and resulting topological network changes. Simplifying approximations to estimate the kinetic energy increase during fault-on periods were evaluated along with other simplifications made to further reduce complexity.

The interrelationship among the elements of the security assessment function is depicted schematically in Fig. ES.1-1. State estimation provides information on the current operating state and configuration of the power system. State forecasting provides a statistical indication of the future behavior of the system given its current state. The various security measures developed during this effort use this information to compute a statistical measure that may be related to the likelihood that the power system will pass from the current operating region to an emergency region. Translating these security measures (e.g., the expected first crossing time) into operator action to forestall loss of system security is beyond the scope of this effort.

ES.3 RECOMMENDATIONS

The research documented in this report has led to the development of two security measures: a long-term measure formulated in terms of a surface crossing problem and a Lyapunov function based transient stability measure suitable for on-line implementation. Neither of these measures were brought to a final stage of development. Additional effort in the following areas is recommended to assess the full utility of the proposed security measures:

- Obtain a numerical solution to Dynkin's equation for several example cases and compare these solutions with the upper and lower bounds derived during this effort.
- Investigate the adequacy of approximating the level crossing probability in terms of a computable number of factorial moments.
- Assess which of the two system models considered for the long-term security measure (Brownian motion-driven Markov process or differentiable sample path process) is a more suitable description of actual power system behavior.
- Extend the scope of the simulation studies of the transient stability measure and assess the efficacy of alternative approximations or models, e.g.,

a different expression for the unstable equilibrium angles or more complex generator models.

Although the problem of interpreting the security measures and relating them to necessary control actions for security enhancement has not been examined in this effort, it is a crucial aspect of developing a dynamic security assessment/enhancement methodology. Further research in this area should focus on translating a warning of possible transition into an unsecure operating region (as embodied in one of the security measures discussed above) into a realistic control strategy (out of computational necessity, suboptimal) for steering the power system into a more secure operating state. A computer-based operator decision aid using Artificial Intelligence to identify the problem and select an appropriate solution may prove to be useful in this regard.