FAILURE DETECTION IN LINEAR SYSTEMS

by

HAROLD LEE JONES
S. B., Massachusetts Institute of Technology, (1967)
S. M., Massachusetts Institute of Technology, (1969)

SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
at the
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
August 1973

Signature of Author _____
Department of Aeronautics and Astronautics

Certified by _____
Thesis Supervisor

_____
Thesis Supervisor

_____
Thesis Supervisor

Accepted by _____
Chairman, Departmental Graduate Committee

MASS. INST. TECH.
OCT 1 1976
LIBRARIES

FAILURE DETECTION IN LINEAR SYSTEMS

by

HAROLD LEE JONES

Submitted to the Department of Aeronautics and Astronautics on August 15, 1973, in partial fulfillment of the requirements for the degree Doctor of Philosophy.

ABSTRACT

The reference model approach developed in this thesis provides a general capability for real-time failure detection and identification in continuous, stochastic, time invariant systems with continuous measurements. A simple approximation extends the approach to stochastic systems with discrete measurements. The approach requires that an accurate model of the unfailed system dynamics be known. For a system with m sensors, a failure detection system is defined which can detect and identify failures in at least m different system components. The failure detection and identification are independent of the manner in which a component fails as long as the failure alters the system output. In the absence of a failure, the failure detection system functions as a suboptimal state estimator.

Thesis Supervisor: Wallace E. Vander Velde
Title: Professor of Aeronautics and Astronautics

Thesis Supervisor: John J. Deyst, Jr.
Title: Assistant Professor of Aeronautics and Astronautics

Thesis Supervisor: Donald C. Fraser
Title: Division Director, C. S. Draper Laboratory

ACKNOWLEDGMENT

The author wishes to thank his thesis committee: Prof. John J. Deyst, Jr, chairman, who suggested the topic and provided direction at numerous critical points during the conduct of this research, Prof. Wallace E. Vander Velde, who was instrumental in ensuring the analytical precisiness of the theoretical work, and to Dr. Donald C. Fraser, who provided the necessary motivation for the development of the filter design techniques. Their unfailing help and inspiration are deeply appreciated.

I am especially indebted to The Charles Stark Draper Laboratory, Inc., in particular the Control and Flight Dynamics Division, for supplying financial support and computing facilities, and to my associates at the Laboratory for numerous helpful discussions. In particular, I would like to thank Mr. Fred J. Marcus for his insight on analytical issues and Messers. Richard D. Goss, James C. Deckert, and Greg L. Zacharias for their many suggestions during the development of the control system examples.

A particular note of thanks is due to Ms. Margaret E. Erickson whose skill and perseverance in typing the final manuscript was greatly appreciated.

Finally, the author wishes to thank his wife Cheryl, whose moral support was instrumental in the completion of this thesis.

This work was prepared under DSR Project 51-61350, sponsored by the Johnson Space Center of the National Aeronautics and Space Administration through Contract NAS 9-13809 with The Charles Stark Draper Laboratory, Inc,. in Cambridge, Massachusetts.

The publication of this work does not constitute approval by the Charles Stark Draper Laboratory or the National Aeronautics and Space Administration of the findings or the conclusions contained herein. It is published only for the exchange and stimulation of ideas.

TABLE OF CONTENTS

CHAPTER		PAGE
1	INTRODUCTION	
	1.1 Problem Statement	9
	1.2 Background	10
	1.3 Summary of Thesis	13
2	PRELIMINARY ANALYSIS	
	2.1 Motivation - Deterministic	16
	2.2 Motivation - Stochastic	28
	2.3 Vector Space Concepts	33
	2.3.1 General Properties of a Vector Space	34
	2.3.2 Structure of a Linear Operator	36
	2.3.3 Matrix Equivalent of Linear Operator Properties	41
	2.4 Observability and Equivalent Realizations	44
3	FAILURE DETECTION THEORY	
	3.1 Detection Spaces	50
	3.2 Single Failure Detection	61
	3.2.1 Detection Theorem	62
	3.2.2 Unobservable Systems	76
	3.2.3 Algorithms	81
	3.3 Multiple Failure Detection	89
	3.3.1 Sufficient Condition for a Detectable Set	91
	3.3.2 An Example	100
	3.3.3 Detectable Subset of an Undetectable Set	104
	3.3.4 State Space Augmentation	125
	3.4 Output Stationarity	140
	3.4.1 Addition of a Stationary Space to a Detectable Set	142
	3.4.2 An Example	157
	3.4.3 Addition of a Stationary Class to a Detectable Set	162
	3.5 Necessary and Sufficient Condition for a Detectable Set.	174
	3.6 Failure Detection Theory and Linear Observers	177

CHAPTER PAGE

4 DETECTION FILTER DESIGN - DETERMINISTIC

4.1 Modeling 184
4.1.1 General Failure Model 185
4.1.2 An Example 196
4.1.3 General Modeling Considerations 204
4.2 Filter Design Procedure 208
4.2.1 Detection Spaces 208
4.2.2 Multiple Event Detection 212
4.2.3 Output Stationarity 220
4.2.4 Eigenvalue Assignment 225
4.3 An Example 230
4.3.1 Satellite Orbit Perturbation Model 230
4.3.2 Detection Filter Design 235
4.4 Base Normal Form 242
4.5 Fully Measurable Systems 255
4.6 Extensions of Failure Detection Theory 260

5 DETECTION FILTER DESIGN - STOCHASTIC

5.1 Suboptimal State Estimation - Steady State Statistics 266
5.1.1 Stochastic Error Model 268
5.1.2 Time Domain Optimization 275
5.1.3 Frequency Domain Optimization 282
5.2 Suboptimal State Estimation - Timevarying Statistics ... 287
5.3 Detection Law for Multiple Failure Detection Systems .. 295
5.4 Error Signal Analysis 302

6 FAILURE DETECTION FOR LATERAL MODE AUTOPILOT

6.1 Detection Filter Design - Deterministic 310
6.1.1 Second-order Roll Response Model 311
6.1.2 Failure Detection System Design - Deterministic. 324
6.2 Detection Filter Design - Stochastic 338
6.2.1 Stochastic Model of Autopilot - Aircraft System .. 339
6.2.2 Failure Detection System Design - Stochastic 349
6.3 Detection Law 371
6.4 Simulation Results 378
6.4.1 Failure Detection 380
6.4.2 State Estimation 405
6.4.3 Sensitivity Analysis 417

CHAPTER PAGE

7 SUMMARY AND RECOMMENDATIONS

7.1 Summary 432
7.2 Recommendations for Future Study 435

TABLE OF DEFINITIONS

		PAGE
2.1.1	Detectable failure	27
2.4.1	Observable subspace	45
2.4.1a	Observable subspace	46
2.4.2	Observable system modulo an invariant subspace	46
3.1.1	Detection equivalence	53
3.1.2	Detection space for an event vector	53
3.1.2a	Detection space for an event vector	58
3.2.1	Quasi-detectable failure	78
3.3.1	Mutually detectable failures	91
3.3.2	Output separable failures	92
3.3.3	Detection space for a set of event vectors	94
3.3.4	Excess subspace of a detection space	105
3.4.1	Output stationary failures	141
3.4.2	Normal subspace for output stationary failures	163
3.4.3	Output stationary equivalent failures	166
3.4.2	Normal subspace for output stationary failures	167
3.4.4	Extended output stationary equivalent failures	173

TABLE OF THEOREMS

		PAGE
2.4.1	Observable system	47
2.4.2	Observable system	47
3.2.1	Detection theorem for a single failure	71
3.2.1	Detection theorem for unobservable systems	79
3.3.1	Detection theorem for multiple failures	99
3.3.2	Detection theorem for a subset of mutually undetectable failures	114
3.3.3	Detection theorem for a subset of mutually undetectable failures (matrix formulation)	120
3.3.4	Minimal representation for which a set of failures are mutually detectable	133
3.4.1	Output stationarity to expand a set of detectable failures	144
3.4.2	Eigenvalues associated with output stationarity	153
3.4.3	Mutual output stationarity to expand a set of detectable failures	166
3.4.4	Mutual output stationarity to expand a set of detectable failures -- general case	172
3.5.1	Detection theorem for multiple failures -- general case	176
3.6.1	Detection theorem in conjunction with linear observers.	180
4.1	Detection theorem for measurable systems	255
4.2	Mutual output stationarity for measurable systems	256

TABLE OF TABLES

4.1	General catalog of system failure types	192
6.1	Summary of Failure Detection Properties for Failures of Lateral Mode Aircraft Autopilot.	328
6.2	Detection channels for the T_{mg}, T_{ms1}, and T_{ms2} detection systems in which each of the possible system failures generates error signals; and the system failures identified by each detection system	379
6.3	Autopilot-aircraft parameters used in simulations	381

CHAPTER 1

INTRODUCTION

1.1 Problem Statement

Failure detection and accommodation has become a critical concern as the complexity of modern systems has increased. In the general sense a failure is any change which reduces a system's effectiveness. This change may take the form of a component failure, a variation in a system parameter, or a variation in the operating environment. In many instances it is possible to determine a priori a set of changes which have a reasonable probability of occurring and to design the system to insure that the changes cannot cause a severe degradation of system performance. Frequently this can be done without determining whether or not a system change does in fact occur. Desensitization of the system in this manner is an integral part of effective system design.

There remains an important class of applications for which satisfactory system behavior requires explicit detection of a failure as an antecedent to corrective action. Failure detection problems of this nature are best described as event detection problems. The event may be either a sudden or a gradual failure, but it is characterized by the fact that it results in the degradation of the system's performance below some threshold value. The detection of such events and identification of the associated failures in an on-line basis is the central issue of this thesis. Attention is restricted to systems which can adequately be described as time invariant or piece-wise time invariant linear stochastic systems.

For a particular application, a failure detection algorithm tailored to that application may afford significant advantages over more generally applicable approaches. As systems increase in complexity and subtlety, however, the effort involved in designing

and implementing such schemes may become prohibitive. The intention of this thesis is to develop a general capability for detecting failures in complex dynamic systems which effects a reasonable compromise between ease of implementation, simplicity of operation, and efficiency of performance. Particular attention is given to maximizing the number of events which can be monitored by a single detection system and minimizing the probability of an incorrect or ambiguous decision.

1.2 Background

The greatest effort in failure detection has been related to the problem of detecting a failure in one of a set of m identical components. If the components are driven by the same input, then in the deterministic case failure detection requires only that the system outputs be measured and compared with a reference output. If a measurement deviates sufficiently from the reference, it is assumed that either the component or its accompanying sensor has failed. The reference output is generally taken to be a weighted average of the output of the m sensors.

This comparison procedure has become known as voting or the majority rule law for failure detection. It is characterized by three restrictions:

(1) under nominal conditions the sensor outputs must be algebraically related,

(2) the detection law cannot discriminate between the failure of a system component and the failure of the sensor which measures the component's output, and

(3) the detection law is memoryless and therefore may experience difficulty in detecting "soft" failures.

Because of the lack of memory, voting cannot detect a failure until the appropriate output differs from the reference by an amount greater than a preset threshold. When a failure has been detected, it is still necessary to determine which element of the component-sensor pair has failed. In the event of m-1 failures, voting will not specify which components have failed. Systems which utilize voting for failure

detection are commonly at least triply-redundant in order to minimize the occurrence of such ambiguous decisions.

Voting has been employed with considerable success in the stochastic case for inertial navigation systems by Gilmore and McKern (1970), Pejsa (1971), and Ephgrave (1969). In these instances the gyro measurements are of sufficient accuracy that the memoryless feature of voting presents no great difficulties in detecting "hard" gyro failures. For "soft" failures, more detailed analysis is required. A useful comparison of several procedures for implementing a voting algorithm in the stochastic case is presented by Potter and Deckert (1972).

If a failure mode is sufficiently soft to make memoryless voting ineffective, an alternative is to consider a weighted time average of the measurement errors as a means of reducing the error variances and hence the threshold. Chien (1972) used a sequential probability ratio test (SPRT) to develop a minimum time detection algorithm for the gyro bias problem. His work represents one of the few examples found in the literature which utilize a voting algorithm with memory for failure detection.

More recently, attempts have been made to develop failure detection algorithms which utilize a model of the system dynamics in order to detect and identify component failures. The simplest of these algorithms use a Kalman filter to monitor the system performance and then make a detection decision based upon the whiteness of the Kalman filter residuals. The WSSR algorithm proposed by Willsky et al (1974) is an example of this approach. Although simple residual-based algorithms may yield effective failure detection, they generally possess a limited capability for failure identification.

The first algorithms to offer effective failure detection and identification for dynamic systems were based upon likelihood ratio concepts discussed by Van Trees (1968) among others. These algorithms have the advantage that they are optimal in a statistical sense for stochastic systems, but they frequently are computationally unweildy. Laniotis (1971) is an example of some of the early development efforts. The Generalized Likelihood Ratio test (GLR) first investigated by

McCauley and Denlinger (1973) and later generalized by Willsky and Jones (1974), solved the computational problem without significantly degrading the failure detection and identification optimality of the likelihood ratio approach. The primary disadvantage of GLR, as well as of all likelihood ratio-based algorithms, is that it requires specific models of all possible failure modes to be formulated for each system component, and a detection test must be performed for each failure mode. For many applications, a list of precise failure modes cannot be formulated.

A variety of different approaches to the failure detection problem have appeared in the recent literature. The Bayesian test of Sanyal and Shen (1973) and the ellipsoid overlap test of Kerr (1974) are two examples. In general, these tests have performed well for their applications.

The major analytic contribution to failure detection in linear systems is due to Beard (1971). The approach used by Beard is similar to that used in state estimation theory and linear observer theory in that it utilizes a model of the nominal system dynamics to predict the system output. In the absence of failures, modeling errors, and system disturbances, the state of this reference model is identical to the state of the system. Assuming the system measurements are perfect, the reference model also predicts the system output exactly. If a failure has occurred, however, the reference model output differs from the system output. Beard showed that this output error can be used to modify the state of the reference model in such a manner that the output error generated by certain system failures possesses spatial characteristics which make the failures readily identifiable.

A very important aspect of Beard's approach is that the ability to identify the failure of a system component is independent of the manner in which the component fails as long as the failure alters the system to a sufficient degree. This is in contrast to the performance of likelihood ratio tests. Also, in contrast to the voting law performance, the approach provides a capability for discriminating between sensor

failures and changes in the system dynamics; and it does not require parallel system redundancy. A failure detection system can usually be designed to detect and identify failures in a number of system components. Beard was able to prove a sufficient condition which a set of failures must satisfy to be detectable by a single failure detection system. He did not prove a necessary condition and he did not consider the stochastic detection problem. The set of failures detectable by a single filter may be significantly larger than the set specified by Beard's sufficiency condition.

The approach used by Beard bears a marked resemblance to output decoupling. Problems of this type have received considerable attention. Among the literature are Falb and Wolovich (1967), Gilbert (1969), and Wonham and Morse (1970). The latter paper demonstrated the utility of geometric concepts as an aid in resolving difficult systems problems. Through consideration of the structure of a linear operator in an n-dimensional vector space, Worham and Morse were able to obtain simple proofs in an intuitive manner for the properties of the general decoupling problem. The derivations of detection properties in this thesis are in a similar vein.

1.3 Summary of Thesis

The subject of this thesis is a restatement and subsequent completion of the failure detection problem as posed by Beard. Beard applied matrix algebra principles to prove the essential properties of his failure detection theory. Although this constructive approach resulted in a set of useful design algorithms, it required the introduction of computational techniques which de-emphasized the basic principles of observability and controllability upon which state space theory is founded. The approach taken in this thesis is to express the failure detection problem in a geometric context in which it is possible to distinguish between properties of the system which are computationally expedient and those which define the problem. From this vantage point the derivation of failure detection theory can be simplified considerably and the relationship of the results to other

concepts in systems theory can be more clearly understood. Special properties can be proved and used where it is clear that they afford computational advantages.

The fundamental geometric property underlying detection theory is that the state space of the linear, time invariant system (A, B, C) can be partitioned into subspaces through the use of appropriate output feedback. Each control element through the appropriate column of B can control a particular subspace of the state space. This is the controllability subspace associated with that control. Wonham and Morse (1970) were able to simplify the output decoupling problem by defining a controllability subspace for each control element. Similarly, each sensor through the appropriate row of C may observe unambiguously a particular subspace of the state space. This is the observability subspace associated with that sensor. The rederivation of failure detection theory undertaken here is based upon a general definition of the observability subspace for a particular sensor.

The structure of this thesis is a consequence of two goals. The first is to present the elements of a useful theory in a clear and complete manner in the hope of spurring further research; the second is to make the theory accessible to readers not well versed in the nuances of linear maps in an n-dimensional vector space. The approach to the failure detection problem is motivated in Section 2.1. The remainder of Chapter 2 is devoted to a survey of mathematical techniques. The structure of the detection problem is investigated axiomatically in Section 3.1. The bulk of the theoretical considerations are contained in the remainder of Chapter 3. The results are summarized informally in Chapter 4 and a natural canonical form for the detection problem is presented. It is intended that a reader without a strong grasp of vector space properties should be able to develop a reasonable comprehension of the content of this thesis from Section 2.1, Section 3.1, Section 3.2.1, and Chapter 4 along with the subsequent results on stochastic systems. Some familiarity with the mathematical techniques of Chapter 2 would be useful.

An extension of the deterministic theory developed in Chapter 3 to stochastic systems is undertaken in Chapter 5. The major difficulty presented by stochastic systems is that it is necessary to develop a procedure for discriminating between output errors due to stochastic sources and those due to a combination of stochastic sources and a system failure. Failure detection theory differs from estimation theory in that the primary concern is not to minimize the difference between the system and reference model outputs, but to enhance this difference in such a manner that failures are easy to identify. Nevertheless, a failure detection system is a state estimator. Algorithms for maximizing the state estimation capability of failure detection systems are developed in Chapter 5.

The theory developed in this thesis allows a considerable amount of freedom in selecting the set of failures which a failure detection system is designed to detect. The manner in which this freedom can be used to design a failure detection system for a stochastic system with discrete measurements is illustrated by a detailed design problem in Chapter 6. The validity of the design procedure is verified by simulation results analyzed in that chapter.

A brief summary of the important results in the thesis is presented in Chapter 7. Recommendations for future work are also discussed.

CHAPTER 2

PRELIMINARY ANALYSIS

The proposed solution of the failure detection problem is based upon the reference model approach used in observer theory and state estimation theory. The dynamic relationship between the reference model and the unfailed system is reviewed in Section 2.1. The change in this relationship due to a system failure is used to motivate a definition of failure detectability and the constraints which this definition imposes on the selection of the feedback matrix in the reference model. Failure detection theory is developed in Chapter 3 from these deterministic considerations, but the theory can also be applied to stochastic systems to provide a combined failure detection and state estimation capability. Section 2.2 reviews the relevant concepts from state estimation theory.

The geometric tools used in this research are not new, but their use has been limited. The classic work by Gantmacher (1959) has been widely referenced in the literature and is gradually being introduced on the classroom level. The relevant concepts of Chapter 7 of Gantmacher are summarized in Section 2.3. The geometric approach permits some useful variations on observability to be introduced. The standard test of observability is compared in Section 2.4 to a more geometric interpretation. Some useful results on eigenvalue assignment using output feedback and on equivalent system representations are also presented.

2.1 Motivation - Deterministic

The motivation for the reference model approach to failure detection is begun in this section. Attention is restricted to deterministic systems which are linear and time invariant in their unfailed states. Such systems can be completely characterized by an input/output relationship between a control vector u(t) and a measurement vector y(t); however, the

state space representation of the systems requires the additional definition of a state vector x(t):

$$\begin{aligned} \dot{x}(t) &= A\,x(t) + B\,u(t) \qquad : x(0) \text{ given} \\ y(t) &= C\,x(t) \end{aligned} \tag{2.1.1}$$

The vectors u(t), x(t), and y(t) are p-dimensional, n-dimensional, and m-dimensional respectively and A, B, and C are constant coefficient matrices of dimension n x n, n x p, and m x n. Neither the state vector x(t) nor the triad of matrices (A, B, C) is uniquely defined for a given system. The relationship between the various equivalent choices of (A, B, C) is discussed in Section 2.4.

Unless C is invertible, the output vector y(t) at any given time does not contain sufficient information to determine x(t). Unfortunately, feedback control laws often require explicit knowledge of x(t). The standard procedure used in both observer theory and stochastic estimation theory is to approximate x(t) by the state vector $x_m(t)$ of a reference model (Figure 2.1). The two approaches define different criteria for choosing a feedback matrix D.[1] The principal concern of failure detection theory is to combine this state estimation capability with a failure detection capability. The constraint placed on the choice of D is that the output error $y(t) - y_m(t)$ must contain system failure information in a format which can be readily interpreted. Subject to this constraint, the freedom to choose D for state estimation purposes is maximized.

The reference model approach is based upon the assumption that the triad (A_m, B_m, C_m) accurately describes the input/output relationship of the system being modeled as long as the system has not failed. Since this requires that (A, B, C) and (A_m, B_m, C_m) be equivalent representations of the system (Section 2.4), no generality is lost if (A_m, B_m, C_m) is replaced in the reference model by (A, B, C). The system equations for the reference model are then similar to (2.1.1) except for the presence of the output error feedback term:

---

[1] Observer theory is somewhat more general in that it does not require an n-dimensional reference model. See Section 3.6.

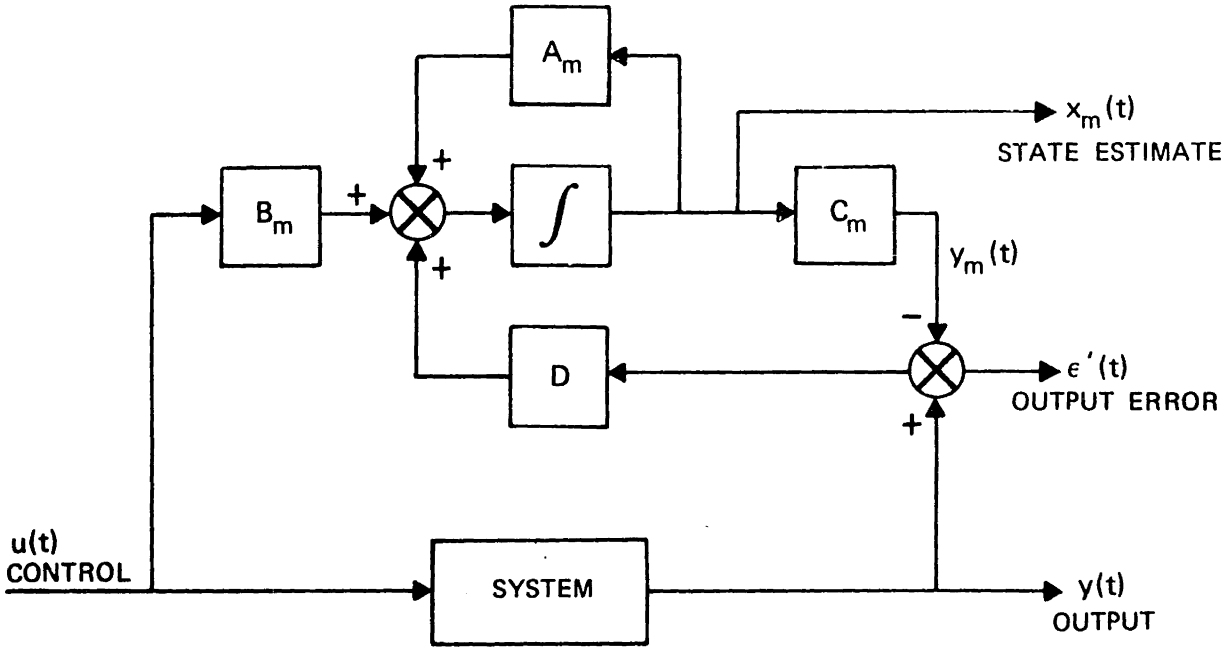


Figure 2.1 General reference model implementation for state estimation and/or failure detection.

$$\dot{x}_m(t) = Ax_m(t) + Bu(t) + D(y(t)-y_m(t)) \quad : x_m(0) \text{ given}$$
$$y_m(t) = Cx_m(t) \tag{2.1.2}$$

An alternate representation is to replace $y_m(t)$ in the differential equation by $Cx_m(t)$:

$$\dot{x}_m(t) = (A-DC)x_m(t) + Bu(t) + Dy(t) \quad : x_m(0) \text{ given}$$
$$y_m(t) = Cx_m(t) \tag{2.1.3}$$

Both (2.1.2) and (2.1.3) become identical to (2.1.1) if $y_m(t)$ equals $y(t)$.

The accuracy with which the reference model approach estimates $x(t)$ can be determined in terms of the error equations:

$$\epsilon(t) = x(t) - x_m(t)$$
$$\epsilon'(t) = y(t) - y_m(t) \tag{2.1.4}$$

In the absence of a failure, $\epsilon(t)$ satisfies a differential equation similar to (2.1.3):

$$\dot{\epsilon}(t) = (A-DC)\epsilon(t) \quad : \epsilon(0) \text{ given}$$
$$\epsilon'(t) = C\epsilon(t) \tag{2.1.5}$$

For the reference model to be an asymptotically stable state estimator, it is necessary that D be chosen such that all the eigenvalues of A-DC have negative real parts. This is the principal criterion by which observer theory defines an acceptable feedback matrix D. Failure detection theory guarantees the ability to satisfy this criterion by providing almost complete freedom to choose the eigenvalues of A-DC. The rationale for applying this freedom comes primarily from the stochastic considerations introduced in Section 2.2.

In the event of a failure or the change of a system parameter, the triad (A, B, C) upon which the reference model is based no longer describes the failed system dynamics and $x_m(t)$ is not necessarily a good approximation of $x(t)$. A general failure model for a linear, time invariant system is developed in Section 4.1. The model represents $x(t)$ and

y(t) in the failed system as a function of (A, B, C) and a series of additive vector components:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + f_1 n_1(t) + f_2 n_2(t) + \ldots \quad : x(0) \text{ given} \\ y(t) &= Cx(t) + e_{m1} n_{c1}(t) + e_{m2} n_{c2}(t) + \ldots \end{aligned} \tag{2.1.6}$$

The scalar coefficients $n_1(t)$, $n_2(t)$ ... and $n_{c1}(t)$, $n_{c2}(t)$ ... are generally functions of x(t) and u(t), but they may be arbitrary time functions. e_{mi} is the i^{th} unit vector in the m-dimensional output space, but f_1, f_2 may be any n-dimensional vectors. The usefulness of the theory developed in this thesis is predicated on the fact that system failures can frequently be modeled using (2.1.6) with a single additive term. The two failure models of most interest are referred to in what follows as the controller failure model:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + f_1 n_1(t) \quad : x(0) \text{ given} \\ y(t) &= Cx(t) \end{aligned} \tag{2.1.7}$$

and the sensor failure model:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) + e_{mi} n_{ci}(t) \end{aligned} \tag{2.1.8}$$

The labels given (2.1.7) and (2.1.8) can be partially justified by an example. The satellite orbital manuevering system analyzed in Section 4.3 is characterized by:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix}$$

$$B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin\gamma & \sin\gamma \\ 0 & 0 & 0 \\ 1 & \cos\gamma & \cos\gamma \end{bmatrix} \tag{2.1.9}$$

where y(t), x(t), and u(t) are 2-, 4-, and 3-dimensional vectors respectively:

$$y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \end{bmatrix} \qquad x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \\ x_3(t) \\ x_4(t) \end{bmatrix} \qquad u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_2(t) \end{bmatrix} \tag{2.1.10}$$

An off-failure of the first control element must be modeled such that $u_1(t)$ does not affect x(t). One possibility is to define a new control matrix B + ΔB:

$$\Delta B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix}$$

$$B + \Delta B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin\gamma & \sin\gamma \\ 0 & 0 & 0 \\ 0 & \cos\gamma & \cos\gamma \end{bmatrix} \tag{2.1.11}$$

An equivalent failure model is (2.1.7) with:

$$f_1 n_1(t) = -\begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u_1(t) \tag{2.1.12}$$

Similarly, an off-failure of the first sensor requires that $y_1(t)$ be zero. Rather than define ΔC such that the first row of C + ΔC is zero, (2.1.8) can be used with:

$$e_{mi} n_{ci}(t) = -e_{21} x_1(t) = -\begin{bmatrix} 1 \\ 0 \end{bmatrix} x_1(t) \tag{2.1.13}$$

A bias of magnitude s_1 in that sensor can also be modeled using (2.1.8) with:

$$e_{mi} n_{ci}(t) = e_{21} s_1 = s_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{2.1.14}$$

These examples are typical of the manner in which control element failures and sensor failures can be modeled using (2.1.7) and (2.1.8) respectively; however, (2.1.7) can also frequently be used to model system failures involving a change in A. Section 4.1 provides an algorithm for modeling failures using (2.1.6) or, where possible, (2.1.7) or (2.1.8).

The behavior of the errors $\epsilon(t)$ and $\epsilon'(t)$ after a system failure can be determined in terms of $\epsilon(t)$ at the time of the failure, (2.1.2), and the appropriate failure model. Because of their importance, attention is restricted to the controller and the sensor failure models. With the failure occurring at t_0, the errors corresponding to (2.1.7) and (2.1.8) are respectively:

$$\begin{aligned} \dot{\epsilon}(t) &= (A-DC)\epsilon(t) + f_1 n_1(t) \quad : \epsilon(t_0) \text{ given} \\ \epsilon'(t) &= C\epsilon(t) \qquad\qquad\qquad\quad t \geq t_0 \end{aligned} \tag{2.1.15}$$

and:

$$\begin{aligned} \dot{\epsilon}(t) &= (A-DC)\epsilon(t) - d_i n_{ci}(t) \quad : \epsilon(t_0) \text{ given} \\ \epsilon'(t) &= C\epsilon(t) + e_{mi} n_{ci}(t) \qquad t \geq t_0 \end{aligned} \tag{2.1.16}$$

where d_i is the i^{th} column of the feedback matrix D. Because of their importance in defining the properties of a failure detection system, vectors which appear in the differential equation for $\epsilon(t)$ similarly to f_1 and d_i above are referred to as event vectors. Even with a zero initial condition, $\epsilon(t)$ as defined in (2.1.15) or (2.1.16) is driven by a forcing function and is in general non-zero for $t \geq t_0$.

The central issue of failure detection theory is how to choose the feedback matrix D such that the component of $\epsilon'(t)$ generated by a failure can readily be associated with that failure. This association is not a function of the manner in which the failure occurs, but of the dynamic relationship between the failed element and the remainder of the system. With respect to the controller and sensor failure models, the association is a function of the event vectors f_1 and d_i and of $e_{mi}(t)$; but it is independent of the time functions $n_1(t)$ and $n_{ci}(t)$. To see the importance of the independence from $n_1(t)$ and $n_{ci}(t)$, consider the control element

failure modeled by (2.1.12). If the control element fails in some manner other than an off-failure, $n_1(t)$ is some time function other than $-u(t)$. The single property which characterizes all possible failure modes for the first control element is that, except for a possible scale factor, the controller failure model must use the event vector f_1 defined in (2.1.12). Similarly, $n_{ci}(t)$ in the sensor failure model for the first sensor may be $x_1(t)$, s_1, or some other time function; but the vector e_{mi} must be e_{21}. As a consequence, the first column of D is the event vector associated with any failure of the first sensor.

The basis for the theory developed in this thesis is best illustrated in terms of the controller failure model (2.1.7) and the corresponding error equations (2.1.15). The term $f_1 n_1(t)$ appears in (2.1.15) in the same manner as the control $Bu(t)$ appears in (2.1.1) and is the only driving term for $\epsilon(t)$. The time functions $\epsilon(t)$ and $\epsilon'(t)$ satisfy:

$$\epsilon(t) = e^{(A-DC)(t-t_0)}\,\epsilon(t_0) + \int_{t_0}^{t} e^{(A-DC)(t-\tau)}\, f_1 n_1(\tau)\, d\tau \qquad (2.1.17)$$

$$\epsilon'(t) = C\epsilon(t)$$

For a given D, $\epsilon(t)$ describes a trajectory in the n-dimensional state space which is a function of $\epsilon(t_0)$, the event vector f_1, and $n_1(t)$. The output error $\epsilon'(t)$ describes a corresponding trajectory in the output space. As a conceptualization aid, a pair of hypothetical trajectories is shown in Figure 2.2a for the case where:

$$Cf_1 \neq 0$$
$$\epsilon(t_0) = 0 \qquad (2.1.18)$$

The principal feature of these trajectories which is independent of $n_1(t)$ is that $\epsilon(t)$ must be contained in the controllable subspace of f_1, i.e. in the subspace spanned by the columns of W (Ogata (1967)):

$$W = [f_1 \; (A-DC)f_1 \; (A-DC)^2 f_1 \ldots (A-DC)^{n-1} f_1] \qquad (2.1.19)$$

The output error $\epsilon'(t)$ is similarly constrained to the subspace spanned by the columns of CW.

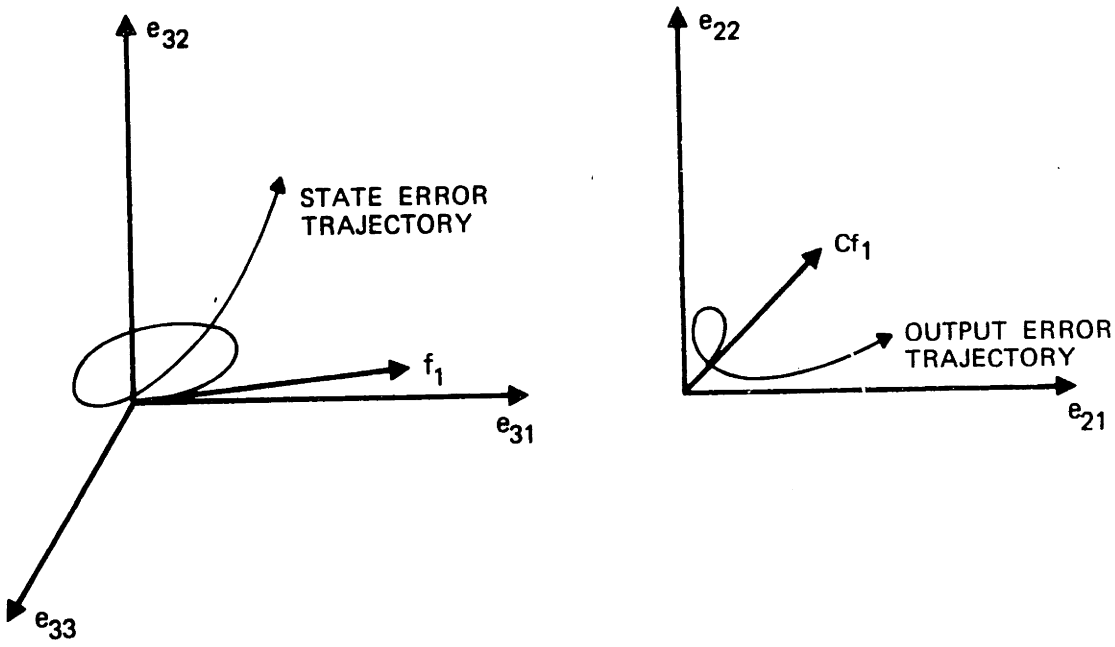


Figure 2.2a Hypothetical state estimation error (left) and output estimation error (right) for a controller failure when the feedback matrix D is chosen arbitrarily.

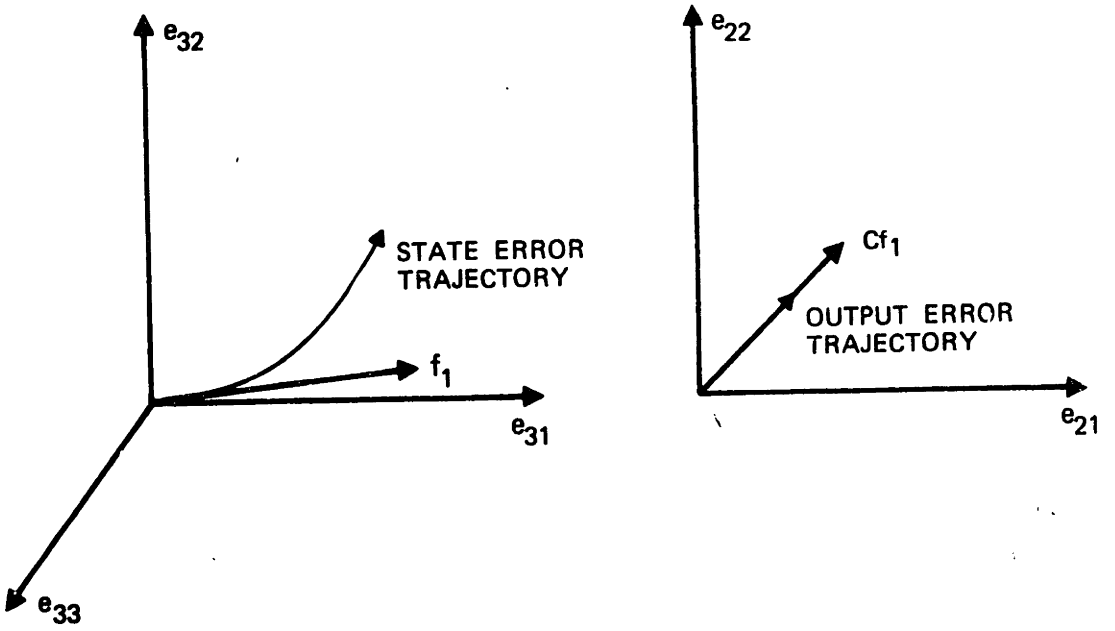


Figure 2.2b Hypothetical state estimation error (left) and output estimation error (right) for a controller failure when D is a detection filter.

The freedom to choose D provides an ability to alter W and CW to make the output error generated by $f_1 n_1(t)$ easily recognizable. The Detection Theorem in Section 3.2 proves that D can be required to satisfy:

$$\text{rk}\, CW = 1 \qquad (2.1.20)$$

while maintaining the freedom to assign the eigenvalues of A-DC. In order to discriminate between the general feedback matrix and one required to satisfy (2.1.20) for some f_1, the latter is referred to as a detection filter for that f_1. Figure 2.2b illustrates hypothetical error trajectories corresponding to those in Figure 2.2a except that D is a detection filter. Subject to certain restrictions discussed in Section 3.1, the fact that $\epsilon'(t)$ is constrained to the output direction Cf_1 allows the error to be uniquely associated with the event vector f_1 and hence with the failure modeled by $f_1 n_1(t)$. If the association is not unique, Section 3.2 provides a mechanism for determining the set of failures which can produce the same output error trajectory. It also proves that choosing D to be a detection filter for f_1 minimizes the number of failures in this set. The case where Cf_1 is zero is readily accommodated in Chapter 3.

The existence of detection filters can be illustrated using the control element failure modeled by (2.1.12). As an example of the controllable subspace of f_1 when D is not a detection filter for f_1, assume D is zero:

$$W = [f_1 \quad (A-DC)f_1 \quad (A-DC)^2 f_1 \quad (A-DC)^3 f_1] = \begin{bmatrix} 0 & 0 & 2\omega & 0 \\ 0 & 2\omega & 0 & -2\omega^3 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & -4\omega^2 & 0 \end{bmatrix} \qquad (2.1.21)$$

As long as ω is not zero, it follows that:

$$\text{rk}\,W = 4$$

$$\text{rk}\,CW = \text{rk}\begin{bmatrix} 0 & 0 & 2\omega & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = 2 \qquad (2.1.22)$$

The rank of CW means that the output estimation error due to a failure in the first control element describes some unspecified trajectory in the two-dimensional output space. By contrast, if D is a detection filter for f_1 such as:

$$D = \begin{bmatrix} 0 & 2\omega \\ 0 & 0 \\ 0 & 0 \\ 0 & -4\omega^2 \end{bmatrix} \qquad (2.1.23)$$

then W satisfies:

$$W = [f_1 \; (A-DC)f_1 \; (A-DC)^2 f_1 \; (A-DC)^3 f_1] = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 2\omega & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \qquad (2.1.24)$$

The lower dimension of W results in a lower dimension for CW:

$$\text{rk}\,W = 2$$

$$\text{rk}\,CW = \text{rk}\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = 1 \qquad (2.1.25)$$

The reference model based on the detection filter for f_1 constrains the output error generated by (2.1.12) to the single direction:

$$C(A-DC)f_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \qquad (2.1.26)$$

As is shown in Section 4.3, only a failure of the first control element can generate a unidirectional output error along $C(A-DC)f_1$ for this system. The appearance of such an output error is sufficient evidence to conclude that the first control element has failed.

The existence of a D which constrains the failure-generated output error to a single direction is guaranteed only if the failure can be modeled by the controller failure model. If the sensor failure model is used, the $d_i n_{ci}(t)$ term in (2.1.16) appears to make (2.1.20) a circuitous argument, but Section 4.1 develops an algorithm which resolves the difficulty. Even with (2.1.20) satisfied, however, the presence of the $e_{mi} n_{ci}(t)$ term in the $\epsilon'(t)$ equation means that in general the output error generated by a sensor failure can only be constrained to a two-dimensional plane. Although this implies that the output error generated by a sensor failure is more difficult for a failure detection system to analyze than that generated by a controller failure, it also implies that failure detection theory provides a mechanism for distinguishing between the failure of a control element and the failure of a sensor which measures that control element's output. This capability is in marked contrast to the performance of the majority-rule voting law discussed in Chapter 1. Detection of failures modeled by the sensor failure model or the general failure model is considered in detail in Section 4.1.

The discussion in this subsection can be summarized in terms of a formal definition of the failure detection approach used in this thesis. Because of its simplicity, the definition is stated in terms of the controller failure model. Section 4.1 extends the definition, as well as the results of Chapter 3, to the general failure model:

Definition 2.1.1 Given the inhomogeneous error equations:

$$\dot{\epsilon}(t) = (A-DC)\epsilon(t) + f_1 n_1(t) \qquad : \epsilon(0) = 0$$

$$\epsilon'(t) = C\epsilon(t)$$

the failure associated with the event vector f_1 is detectable if and only if:

i) the output generated by $f_1 n_1(t)$ maintains a fixed direction in the output space, and

ii) the eigenvalues of A-DC can be specified almost arbitrarily by the proper choice of the detection filter D.

Almost arbitrarily implies only that the n eigenvalues of A-DC need not be assignable as a single set of n eigenvalues. The algorithms developed in Chapter 3 partition the reference model into a number of subsystems. The eigenvalues of each subsystem can be assigned arbitrarily. The only constraint is that if a complex eigenvalue is associated with a particular subsystem, its complex conjugate must be also. The ability to assign the eigenvalues of A-DC almost arbitrarily is the greatest degree of freedom in choosing D which can be guaranteed if D is required to satisfy i) for a number of different event vectors, i.e. if the failure detection system is required to detect and identify several different failures.

2.2 Motivation - Stochastic

The failure detection theory developed in Chapter 3 is based upon a deterministic system model. The theory is equally valid for stochastic systems. The principal difference is that the state estimation error equations corresponding to (2.1.5) are not homogeneous for a stochastic system. It is shown in this section that in the unfailed system the error equations are driven by the stochastic sources. From a failure detection vantage point, the main concern is the problem of differentiating between the output error produced by a failure and the output error produced by the stochastic sources. If the failure detection system (reference model) is to be used for state estimation also, it is necessary to insure that a detection filter is selected which provides the necessary estimation accuracy. Both of these problems are discussed in detail in Chapter 5. This section states the optimal state estimation problem in a format which is applicable to the design of a failure detection system.

Early research in optimal state estimation was within the framework of classical control theory. Wiener solved the minimum mean-squared error estimation problem in steady state for stationary disturbances and a time invariant system (see Truxal (1955)). The Weiner filter is a free configuration design. Newton, Gould, and Kaiser (1957) solved the simpler problem of optimizing a set

of free parameters in an otherwise fixed configuration. These results were derived in the frequency domain using integral transform techniques. The general time varying problem with nonstationary statistics was solved by Kalman (1960) in the time domain. Kalman's solution is equivalent to Wiener's for the restricted problem solved by Wiener. In failure detection theory both the frequency domain approach and the time domain approach can be used to advantage. The basic mathematics for the two optimization procedures is summarized in this section.

The standard linear disturbance model is assumed for the stochastic system. The system is driven by an n_1-dimensional vector random process $\eta_1(t)$ and the measurements are corrupted by an m_1-dimensional vector random process $\eta_2(t)$:

$$\dot{x}(t) = Ax(t) + Bu(t) + G_1 \eta_1(t)$$

$$y(t) = Cx(t) + G_2 \eta_2(t) \qquad (2.2.1)$$

G_1 and G_2 are $n \times n_1$ and $m \times m_1$ matrices respectively which may be time dependent. The random processes are white and have zero means, but they are not necessarily independent:

$$E[\eta_i(t)] = 0$$

$$E[\eta_1(t)\, \eta_1(\tau)^T] = E_{n_1}\, \delta(t-\tau)$$

$$E[\eta_2(t)\, \eta_2(\tau)^T] = E_{m_1}\, \delta(t-\tau)$$

$$E[\eta_1(t)\, \eta_2(\tau)^T] = G_{12}\, \delta(t-\tau) \qquad (2.2.2)$$

E_{n_1} and E_{m_1} are $n_1 \times n_1$ and $m_1 \times m_1$ identity matrices respectively and $\delta(t-\tau)$ is the unit impulse.

In the absence of a system failure the error equations for the reference model are no longer homogeneous. Equation (2.1.5) becomes:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + G_1\dot{\eta}_1(t) - DG_2\,\eta_2(t) \;:\; \epsilon(0)\text{given}$$

$$\epsilon'(t) = C\epsilon(t) + G_2\eta_2(t) \qquad (2.2.3)$$

The state estimation error is a zero mean stochastic process with a covariance matrix:

$$P(t) = E[\,\epsilon(t)\,\epsilon(t)^T]$$

which satisfies the matrix Riccati equation (Jazwinski (1970)):

$$\dot{P}(t) = (A-DC)\,P(t) + P(t)(A-DC)^T + G_1\,G_1{}^T + DG_2G_2^TD^T - G_1G_{12}G_2^TD^T -$$

$$DG_2G_{12}^T\,G_1^T \;:\; P(0)\text{ given} \qquad (2.2.4)$$

Equation (2.2.4) is valid for an arbitrary choice of the feedback matrix D. If the eigenvalues of A-DC have negative real parts and if G_1, G_{12}, and G_2 are constant, the steady state solution P_∞ exists and satisfies:

$$(A-DC)\,P_\infty + P_\infty(A-DC)^T + G_1G_1{}^T + DG_2G_2{}^TD - G_1G_{12}G_2^TD^T - DG_2G_{12}^TG_1^T = 0$$

$$(2.2.5)$$

As long as G_1 is not zero, P_∞ is real, symmetric, and positive semi-definite. The solution of (2.2.5) must generally be found through numerical integration of (2.2.4).

The Kalman approach involves a free choice of D to minimize the variance of the state estimate generated by the reference model. The optimal choice of D is denoted here by $K_D(t)$. For the continuous system, $K_D(t)$ satisfies:

$$K_D(t) = (P(t)C^T + G_1\,G_{12}\,G_2^T)\,(G_2G_2^T)^{-1} \qquad (2.2.6)$$

A reference model based on $K_D(t)$ is referred to as a <u>minimum variance state estimator</u>. The covariance matrix for the minimum

variance state estimator satisfies (2.2.4) with $K_D(t)$ substituted for D, or alternatively:

$$\dot{P}(t) = AP(t)+P(t)A^T+G_1G_1^T-(P(t)C^T+G_1G_{12}G_2^T)(G_2G_2^T)^{-1}(CP(t)+G_2G_{12}^TG_1^T): \quad P(0) \text{ given} \qquad (2.2.7)$$

The steady state Kalman filter $K_{D\infty}$ is found by substituting the steady state solution of (2.2.7) into (2.2.6).

If D is constrained to be a detection filter, it is not in general the steady state solution of (2.2.6). As a consequence, a failure detection system does not generate state estimates which are as accurate as those generated by a minimum variance state estimator. It is possible, however, to perform a constrained optimization of a detection filter using the Kalman approach. It is shown in Section 4.4 that the detection filter for a set of m failures (m is the rank of C) can be expressed as a fixed configuration with a maximum set of n free parameters--these parameters uniquely determine the n eigenvalues of A-DC. With the aid of a coordinate transformation which places the reference model in a convenient canonical form, Section 5.1.2 uses the Kalman approach to develop an algorithm for optimizing these n parameters.

The parameter optimization procedure of Newton, Gould, and Kaiser does not perform a global minimization of the variance of the state estimates generated by the reference model. Instead, it minimizes a scalar function of the steady state variance of the output estimate $Cx_m(t)$. The error associated with this estimate is $C\epsilon(t)$ in (2.2.3) rather than $\epsilon'(t)$. The steady state covariance of $Cx_m(t)$ is $CP_\infty C^T$ where P_∞ satisfies (2.2.5), but the parameter optimization technique requires that a closed-form solution of $CP_\infty C^T$ be found using frequency domain techniques. With zero initial conditions, the Laplace transform for $C\epsilon(t)$ can be found from (2.2.3):

$$C\epsilon(s) = C(E_n s - A+DC)^{-1}(G_1\eta_1(s) - DG_2\eta_2(s)) \quad (2.2.8)$$

Since $\eta_1^{(t)}$ and $\eta_2(t)$ are zero mean, white gaussian random processes, their matrix power spectral densities are identity matrices:

$$\Phi_{\eta_1}(s) = E_{n_1}$$

$$\Phi_{\eta_2}(s) = E_{m_1}$$

$$\Phi_{\eta_1\eta_2}(s) = G_{12} \quad (2.2.9)$$

If the system has stationary statistics, then G_1, G_{12}, and G_2 are constant and the power spectral density for $C\epsilon(s)$ can be determined from (2.2.8) and (2.2.9) to be the matrix:

$$C\Phi_\epsilon(s)C^T = C(E_n s - A+DC)^{-1}(G_1G_1^T+DG_2G_2^TD^T - G_1G_{12}G_2^TD^T - DG_2G_{12}^T G_1^T)(-E_n s - A^T+C^TD^T)^{-1}C^T \quad (2.2.10)$$

The steady state output error covariance $CP_\infty C^T$ can then be expressed in terms of the inverse transform of $C\Phi_\epsilon(s)C^T$:

$$CP_\infty C^T = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} C\Phi_\epsilon(s)\, C^T ds \quad (2.2.11)$$

This expression for $CP_\infty C^T$ is of limited usefulness if $C(E_n s - A+DC)^{-1}$ and the inverse transform (2.2.11) cannot be readily computed. The transformation to the canonical form introduced in Section 4.4 permits $C(E_n s - A+DC)^{-1}$ to be determined by inspection. The elements of the inverse transform can be determined from tables as polynomials in the n degrees of freedom remaining in the choice of the detection filter. To determine the optimal detection filter-state estimator, it is necessary to use these n free parameters to minimize a set of m scalar polynomials defined from $CP_\infty C^T$ by the coordinate transformation. In a large number of cases the polynomials are of sufficiently low order to allow a direct minimization

without recursive techniques. If the failure detection system is designed for m failures, the detection filter-state estimator produced by the parameter optimization technique is identical to that yielded by the Kalman approach for systems with stationary statistics and continuous measurements. If fewer than m failures are to be monitored, however, the approaches yield different detection filter-state estimators.

2.3 Vector Space Concepts

The traditional approach for studying the properties of the system (A, B, C) has been to use matrix algebra techniques. Unfortunately, derivations obtained in this manner are frequently tedious. Often the desired result is arrived at by the construction of a coordinate transformation which places the matrix in a particular canonical form. The advantage of such an approach is that it is constructive; the disadvantage is that intuitive insight into the nature of the solution is often minimized by the indirectness of the approach.

The difficulty with matrix analysis is that it is always necessary to discriminate between properties which are indigenous to the system and those which are a function of a particular basis. The former properties are referred to as invariant properties and are identical for all possible n-dimensional realizations of the system (A, B, C). A useful method for dealing directly with those invariant properties is in terms of a linear operator in an n-dimensional vector space. The axiomatic definition of a vector space immediately separates the invariant properties from those which depend upon the choice of a particular basis.

The general nature of a linear vector space and of a linear operator defined in that space are reviewed in Section 2.3.1. Only those properties which are applicable to this thesis are mentioned. A more complete and formal development is found in Gantmacher (1959) or Halmos (1958). In Section 2.3.2 the structure of a linear operator in an n-dimensional vector space is considered. A brief interpretation of Section 2.3.2 in a matrix context is given in Section 2.3.3.

2.3.1 General Properties of a Vector Space

The axiomatic definition of a vector space is stated in terms of a set of field axioms. In essence these formalize the intuitive notion that vector addition is associative and commutative and that scalar multiplication is associative and distributive. The sum of two elements of the vector space is required to be an element of the vector space and the product of a scalar with an element of the space must be contained in the space. In what follows the scalar is restricted to the field of real numbers. The axioms also define the zero vector ϕ and the identity scalar 1.

An n-dimensional vector space is denoted by $\mathcal{S}^n$. A subspace $\mathcal{R}$ of $\mathcal{S}^n$ is denoted by $\mathcal{R} \subset \mathcal{S}^n$. The sum of any two vectors in $\mathcal{R}$ must be contained in $\mathcal{R}$ and the product of any scalar with a vector in $\mathcal{R}$ must be contained in $\mathcal{R}$. The dimension of $\mathcal{R}$ is denoted by $d(\mathcal{R})$ and is defined to be the largest number of linearly independent vectors in $\mathcal{R}$. A basis for $\mathcal{R}$ is any set of k linearly independent vectors $\xi_1 \ldots \xi_k \in \mathcal{R}$ where $k = d(\mathcal{R})$. A set of vectors $\xi_1' \ldots \xi_\ell' \in \mathcal{R}$ with $\ell \geq k$ is said to span $\mathcal{R}$ if a subset of k vectors from $\xi_1' \ldots \xi_\ell'$ can be found which is a basis for $\mathcal{R}$.

Let $\mathcal{S}$ and $\mathcal{R}$ be subspaces with bases $\rho_1 \ldots \rho_\ell$ and $\xi_1 \ldots \xi_k$ respectively. The union of $\mathcal{S}$ and $\mathcal{R}$ is denoted by $\mathcal{S} \cup \mathcal{R}$ and is the subspace spanned by $\rho_1 \ldots \rho_\ell$, $\xi_1 \ldots \xi_k$. Even though $\rho_1 \ldots \rho_\ell$ and $\xi_1 \ldots \xi_k$ are sets of linearly independent vectors, the vectors in the composite set $\rho_1 \ldots \rho_\ell$, $\xi_1 \ldots \xi_k$ are not necessarily linearly independent; therefore:

$$d(\mathcal{S} \cup \mathcal{R}) \leq d(\mathcal{S}) + d(\mathcal{R}) \qquad (2.3.1)$$

The direct sum of $\mathcal{S}$ and $\mathcal{R}$ is denoted by $\mathcal{S} \oplus \mathcal{R}$ and is reserved to designate the union of $\mathcal{S}$ and $\mathcal{R}$ when the vectors in the composite set $\rho_1 \ldots \rho_\ell$, $\xi_1 \ldots \xi_k$ are known to be linearly independent. The set $\rho_1 \ldots \rho_\ell$, $\xi_1 \ldots \xi_k$ is a basis for $\mathcal{S} \oplus \mathcal{R}$ and:

$$d(\mathcal{S} \oplus \mathcal{R}) = d(\mathcal{S}) + d(\mathcal{R}) \tag{2.3.2}$$

Any subspace $\mathcal{R}^c$ which satisfies:

$$\mathcal{E}^n = \mathcal{R} \oplus \mathcal{R}^c \tag{2.3.3}$$

is a complement of $\mathcal{R}$. The definition of a direct sum requires:

$$d(\mathcal{R}^c) = n-k \tag{2.3.4}$$

A complement of a subspace is not unique, but a complement with unique properties is defined below.

A normed vector space has an inner product defined over it. It is sufficient to restrict attention to the Euclidean norm (inner product):

$$\langle \xi, \rho \rangle = (\xi_1 \rho_1 + \xi_2 \rho_2 + \ldots \xi_n \rho_n)^{1/2} \tag{2.3.5}$$

where $\xi, \rho \in \mathcal{E}^n$ and the components ξ_i, ρ_i, $i = 1 \ldots n$, are defined for some arbitrary coordinate system. The value of the inner product is independent of the coordinate system. Two vectors ξ, ρ are orthogonal if their inner product is zero. The orthogonal complement of a subspace $\mathcal{R} \subset \mathcal{E}^n$ is a subspace $\mathcal{R}^\perp$ composed of all vectors orthogonal to every vector in $\mathcal{R}$:

$$\mathcal{R}^\perp = \{ \rho \in \mathcal{E}^n : \langle \rho, \xi \rangle = \phi, \forall\, \xi \in \mathcal{R} \} \tag{2.3.6}$$

The principal interest in the orthogonal complement in this thesis is that it provides a mechanism for determining the intersection of two subspaces. The intersection of $\mathcal{S}$ and $\mathcal{R}$ is denoted by $\mathcal{S} \cap \mathcal{R}$ and is composed of every vector in $\mathcal{S}$ which is also in $\mathcal{R}$. A basis for $\mathcal{S} \cap \mathcal{R}$ generally cannot be defined as a subset of $\rho_1 \ldots \rho_\ell$ or $\xi_1 \ldots \xi_k$. Computational algorithms for a basis of $\mathcal{S} \cap \mathcal{R}$ are usually based on the orthogonal complements of $\mathcal{R}$ and $\mathcal{S}$ and the relationship:

$$\mathcal{S} \cap \mathcal{R} = (\mathcal{S}^\perp \cup \mathcal{R}^\perp)^\perp \tag{2.3.7}$$

For every n x n matrix A, there is a unique linear operator defined in $\mathcal{E}^n$. The linear operator is denoted by the boldface symbol $\mathbf{A}$. The linear operator $\mathbf{A}$ is a map of $\mathcal{E}^n$ into $\mathcal{E}^n$. This map is generally denoted by $\mathbf{A}: \mathcal{E}^n \to \mathcal{E}^n$. For every vector $\xi \in \mathcal{E}^n$, $\mathbf{A}$ associates a vector $\rho \in \mathcal{E}^n$:

$$\rho = \mathbf{A}\xi \tag{2.3.8}$$

For every operator $\mathbf{A}$ there corresponds a class of similar matrices--the class of matrices similar to A, i.e. related to A by a coordinate transformation. $\mathbf{A}$ is the unique operator associated with that class. For a given basis (coordinate system), $\mathbf{A}$ corresponds to a unique n x n matrix which is an element of the similarity class. $\mathbf{A}$ and the elements of the similarity class are assumed to be defined over the field of real numbers.

If $\mathcal{R} \subset \mathcal{E}^n$, the image of $\mathcal{R}$ under $\mathbf{A}$ is the smallest subspace $\mathcal{R}' \subset \mathcal{E}^n$ such that for every $\xi \in \mathcal{R}$ there exists a $\xi' \in \mathcal{R}'$ which satisfies:

$$\xi' = \mathbf{A}\xi \tag{2.3.9}$$

The image of $\mathcal{R}$ under $\mathbf{A}$ is frequently denoted by $\mathbf{A}\mathcal{R}$ and the relationship of $\mathcal{R}'$ to $\mathcal{R}$ is denoted by:

$$\mathcal{R}' = \mathbf{A}\mathcal{R} \tag{2.3.10}$$

If $\mathbf{C}: \mathcal{E}^n \to \mathcal{E}^m$ and $\mathbf{D}: \mathcal{E}^m \to \mathcal{E}^n$, the composition map $\mathbf{DC}: \mathcal{E}^n \to \mathcal{E}^n$ is well defined as is the difference map $\mathbf{A}-\mathbf{DC}: \mathcal{E}^n \to \mathcal{E}^n$. The null space of $\mathbf{A}$ is denoted by $\eta(\mathbf{A})$ and is defined as the largest subspace of $\mathcal{E}^n$ whose image under $\mathbf{A}$ is the zero space $\emptyset$.

2.3.2 Structure of a Linear Operator

The structure of a linear operator $\mathbf{A}$ in $\mathcal{E}^n$ is evidenced by the manner in which it maps subspaces of $\mathcal{E}^n$ into each other and by the manner in which it maps subspaces into the zero space. Two

important concepts in this regard are those of annihilating polynomials and invariant subspaces.

Let $\xi \in \mathcal{S}^n$ be chosen arbitrarily and form the sequence $\{\xi, A\xi, A^2\xi \ldots\}$. Since $\mathcal{S}^n$ is finite dimensional, there exists some smallest integer $\nu \leq n$ such that the vectors ξ, $A\xi \ldots A^{\nu-1}\xi$ are linearly independent, but that $A^\nu\xi$ is a linear combination of the first ν vectors in the sequence:

$$A^\nu \xi = \sum_{i=0}^{\nu-1} \alpha_i A^i \xi \tag{2.3.11}$$

If the polynomial $\psi(\cdot)$ is defined:

$$\psi(A) = A^\nu - \sum_{i=0}^{\nu-1} \alpha_i A^i \tag{2.3.12}$$

it follows from (2.3.11) that:

$$\psi(A)\xi = \phi \tag{2.3.13}$$

The polynomial $\psi(\cdot)$ is defined to be the minimal annihilating polynomial of ξ with respect to A. Any polynomial $\tilde{\psi}(\cdot)$ such that

$$\tilde{\psi}(A)\xi = \phi \tag{2.3.14}$$

is an annihilating polynomial for ξ. The justification for calling $\psi(\cdot)$ the minimal annihilating polynomial[1] (or just minimal polynomial) of ξ is obvious: $\psi(\cdot)$ is clearly the polynomial of lowest order which annihilates ξ. It can also be shown that $\psi(\cdot)$ is a factor of every annihilating polynomial of ξ; hence the minimal polynomial is unique.

A minimal polynomial can be defined for a subspace. $\Psi(\cdot)$ is a minimal polynomial for $\mathcal{R}$ if it is the lowest order polynomial

[1] The qualifier "with respect to" is often omitted when the appropriate operator is well understood.

which annihilates every vector in $\mathcal{R}$. A constructive solution for $\Psi(\cdot)$ is to choose an arbitrary basis $\xi_1 \ldots \xi_k$ for $\mathcal{R}$ and determine the respective minimal polynomials $\psi_1(\cdot) \ldots \psi_k(\cdot)$ of the basis vectors. $\Psi(\cdot)$ is then the least common multiple of $\psi_1(\cdot) \ldots \psi_k(\cdot)$. It follows that the minimal polynomial of a subspace is divisible by the minimal polynomial of every vector in that subspace.

A subspace $\mathcal{R} \subset \mathcal{E}^n$ is defined to be <u>invariant</u> with respect to A if its image satisfies $A\mathcal{R} \subset \mathcal{R}$. In other words, a subspace is invariant with respect to A if it contains its own image under A. Clearly $\mathcal{E}^n$ is invariant with respect to A. For an example of another invariant subspace, consider the ν-dimensional subspace $\mathcal{L}_\xi$ for which the set of linearly independent vectors $\xi, A\xi, \ldots A^{\nu-1}\xi$ from (2.3.11) is a basis. If this basis is denoted by $e_1 \ldots e_p$, then it is apparent that $Ae_i \in \mathcal{L}_\xi$ for all $i = 1 \ldots p$. Therefore, the image under A of every vector $\rho \in \mathcal{L}_\xi$ is contained in $\mathcal{L}_\xi$ and $\mathcal{L}_\xi$ is invariant. A subspace for which a basis can be generated by the sequence $\xi, A\xi, \ldots A^{\nu-1}\xi$ is a <u>cyclic-invariant</u> subspace and the vector ξ is referred to as the <u>generator</u> of that subspace. The generator for a particular subspace is not unique; $\xi + A\xi$ may also be a generator for $\mathcal{L}_\xi$. It is apparent that every vector $\rho \in \mathcal{E}^n$ generates some invariant subspace of $\mathcal{E}^n$. The opposite is not true; not every invariant subspace of $\mathcal{E}^m$ has a generator.

Every invariant subspace has associated with it a minimal polynomial $\Psi(\cdot)$. The roots of the polynomial are the eigenvalues of A associated with that subspace. The solution in Chapter 3 is based upon the construction of an operator A-DC which has a specified set of invariant subspaces and for which the eigenvalues of A-DC associated with these subspaces assume preassigned values.

A concept which plays an important role in the latter part of Chapter 3 is that of congruent vectors. The vectors $\xi, \rho \in \mathcal{E}^n$ are defined to be <u>congruent</u> (modulo $\mathcal{R}$) if their difference is an element of $\mathcal{R}$, i.e. $\xi - \rho \in \mathcal{R}$. Notationally, congruence is specified by:

$$\xi \equiv \rho \pmod{\mathcal{R}} \qquad (2.3.15)$$

It follows trivially that congruence is reflexive and symmetric. It is also transitive since (2.3.15) and:

$$\rho \cong \chi \pmod{\mathcal{R}} \tag{2.3.16}$$

together imply:

$$\xi \equiv \chi \pmod{\mathcal{R}} \tag{2.3.17}$$

Similarly, congruent vectors multiplied by a constant are congruent:

$$a\xi \equiv a\rho \pmod{\mathcal{R}} \tag{2.3.18}$$

Congruence relationships are also additive, i.e. given (2.3.15) and:

$$\xi' \equiv \rho' \pmod{\mathcal{R}} \tag{2.3.19}$$

then it follows that:

$$\xi + \xi' \equiv \rho + \rho' \pmod{\mathcal{R}} \tag{2.3.20}$$

The practical applications of congruence require $\mathcal{R}$ to be invariant. In this case (2.3.15) implies:

$$A^j \xi \equiv A^j \rho \pmod{\mathcal{R}} \tag{2.3.21}$$

for all j.

Congruence is a generalization of equality. The (mod $\mathcal{R}$) argument can be interpreted as an expansion of the zero space ϕ tp include $\mathcal{R}$. All the vector space concepts introduced in this section can be defined (mod $\mathcal{R}$). The relative minimal polynomial of a subspace $\mathcal{S} \subset \mathcal{E}^n$ is denoted by $\Psi_{\mathcal{R}}(\cdot)$ and satisfies:

$$\Psi_{\mathcal{R}}(A)\,\mathcal{S} \equiv \phi \pmod{\mathcal{R}} \tag{2.3.22}$$

The vectors $\xi_1 \ldots \xi_k \in \mathcal{S}$ are linearly independent (mod $\mathcal{R}$) if there exist no $\alpha_1 \ldots \alpha_k$ such that:

$$\alpha_1 \xi_1 + \alpha_2 \xi_2 + \ldots \alpha_k \xi_k \equiv \phi \pmod{\mathcal{R}} \tag{2.3.23}$$

The dimension of $\mathcal{S}$ (mod $\mathcal{R}$) is written $d_{\mathcal{R}}(\mathcal{S})$ and satisfies:

$$d(\mathcal{S}) \geq d_{\mathcal{R}}(\mathcal{S}) \tag{2.3.24}$$

Similarly, the subspace $\mathcal{S}$ is invariant (mod $\mathcal{R}$) if:

$$A\mathcal{S} \subset \mathcal{S} \pmod{\mathcal{R}} \tag{2.3.25}$$

This section is concluded with a brief discussion of factor spaces. Factor spaces are of minor importance to this thesis. They are used in several instances to suggest a proper generalization of a concept. A thorough understanding of the following is not a necessary prerequisite for Chapter 3.

Let ξ be an arbitrary vector in the subspace $\mathcal{S}$ and let the class ξ^* be defined as the collection of all vectors in $\mathcal{S}$ which are congruent (mod $\mathcal{R}$) to ξ. For every vector $\rho \in \mathcal{S}$ there is a corresponding class ρ^*. Given ξ^* and ρ^*, (2.3.20) suggests that arbitrary elements of the two classes can be added and the set of all such combinations is a class of vectors (denoted by $\xi^* + \rho^*$) whose elements are congruent (mod $\mathcal{R}$) to $\xi + \rho$. Similarly, multiplication by a scalar can be defined and denoted by $\alpha \xi^*$. There is a zero class ϕ^* composed of $\mathcal{R}$ and it follows that:

$$\xi^* - \xi^* = \phi^* \tag{2.3.26}$$

From the above it is clear that a vector space $\mathcal{S}^*$ can be defined whose elements are the classes ξ^*, ρ^*, ... The vector space $\mathcal{S}^*$ is referred to as the <u>factor space</u> of $\mathcal{S}$ with respect to $\mathcal{R}$. If $d(\mathcal{R}) = \nu_{\mathcal{R}}$ and $d(\mathcal{S}) = \nu_{\mathcal{S}}$, then:

$$d(\mathcal{S}^*) = \nu_{\mathcal{S}}^* = \nu_{\mathcal{S}} - \nu_{\mathcal{R}} \tag{2.3.27}$$

The operator A can be defined on $\mathcal{L}^*$. If $\mathcal{R}$ is invariant, the image $A\xi^*$ is a class in $\mathcal{L}^*$ for all $\xi^* \in \mathcal{L}^*$. The factor space is a vector space in every sense and any property defined in this section can be defined on $\mathcal{L}^*$ through consideration of (2.3.15) - (2.3.27).

2.3.3 Matrix Equivalent on Linear Operator Properties

The essential difference between a matrix and a linear operator is that a linear operator is defined without appending a basis to the vector space. For every property of linear operators there is an equivalent property of matrices. This equivalence insures that the results of Chapter 3 can be converted into matrix algorithms. It also allows for the possibility of following the derivations in Chapter 3 on an informal basis. The matrix properties summarized below are sufficient for both purposes. Gantmacher (1959), Hildebrand (1965), and Wonham and Morse (1971) are references for this subsection.

The matrix theory equivalent to linear operator properties can be illustrated in terms of $A: \mathcal{E}^n \to \mathcal{E}^n$, the subspace $\mathcal{R} \subset \mathcal{E}^n$, and the image of $\mathcal{R}$ under A:

$$\mathcal{R}' = A\mathcal{R} \tag{2.3.28}$$

If $\mathcal{R}$ has a basis $\xi_1 \ldots \xi_k$, let R be the matrix whose columns are the vectors $\xi_1 \ldots \xi_k$ expressed in the appropriate coordinate system. By convention, R is denoted by:

$$R = [\xi_1 \ \xi_2 \ldots \xi_k] \tag{2.3.29}$$

The range space of R is the matrix equivalent of the subspace $\mathcal{R}$. Similarly, if the vectors $\xi_1' \ldots \xi_k' \in \mathcal{R}'$ satisfy:

$$\xi_i' = A\xi_i \qquad i = 1 \ldots k \tag{2.3.30}$$

then R' can be defined as the matrix whose columns are the vectors $\xi_1' \ldots \xi_k'$ expressed in the same coordinate system:

$$R' = [\xi_1' \; \xi_2' \ldots \xi_k'] \tag{2.3.31}$$

For this coordinate system, (2.3.28) and (2.3.30) imply that the linear operator A corresponds to a n x n matrix A which satisfies:

$$R' = AR \tag{2.3.32}$$

Just as $\mathcal{R}'$ is the image of $\mathcal{R}$ under A, the range space of R' can be defined as the range space of R under A.

Matrices such as R and R' are useful for expressing the relationship between two subspaces in matrix form. Let $\mathcal{S} \subset \mathcal{E}^n$ be an arbitrary subspace with a basis $\rho_1 \ldots \rho_\ell$ and let the matrix S be defined in terms of $\rho_1 \ldots \rho_\ell$. If $\mathcal{S}$ satisfies $\mathcal{S} \subset \mathcal{R}$, the range space of S is contained in the range space of R. If $\mathcal{S} \not\subset \mathcal{R}$ and if the union of $\mathcal{S}$ and $\mathcal{R}$ is the direct sum as defined by (2.3.2), the subspace $\mathcal{S} \oplus \mathcal{R}$ is the range space of the matrix [S: R]. The subspaces $\mathcal{S} \cup \mathcal{R}$ and $\mathcal{S} \cap \mathcal{R}$ are considerably more difficult than $\mathcal{S} \oplus \mathcal{R}$ to define in matrix terms, but they are of relatively minor importance in Chapter 3 and their matrix equivalents are not developed here.

If the state space $\mathcal{E}^n$ is partitioned into two disjoint subspaces $\mathcal{S}$ and $\mathcal{R}$:

$$\mathcal{E}^n = \mathcal{R} \oplus \mathcal{S} \tag{2.3.33}$$

the operator $A: \mathcal{E}^n \to \mathcal{E}^n$ can be uniquely defined by four operators on the partitioned state space:

$$\begin{aligned} A_{RR}&: \mathcal{R} \to \mathcal{R} \\ A_{SS}&: \mathcal{S} \to \mathcal{S} \\ A_{SR}&: \mathcal{R} \to \mathcal{S} \\ A_{RS}&: \mathcal{S} \to \mathcal{R} \end{aligned} \tag{2.3.34}$$

If a matrix A is given in some coordinate system and the matrices R and S in that coordinate system are defined similarly to (2.3.29),

then the direct sum in (2.3.33) guarantees that R and S can be used to define a similarity (coordinate) transformation on A:

$$A' = [R: S]^{-1} A [R: S] = \left[\begin{array}{c|c} A_{RR} & A_{RS} \\ \hline A_{SR} & A_{SS} \end{array}\right] \qquad (2.3.35)$$

The submatrix A_{RR} is the matrix equivalent of the operator A_{RR} etc. in precisely the same manner that A (or A) is the matrix equivalent of A.

The minimal polynomial of a subspace with respect to A can be related to the characteristic equation $\Delta(\cdot)$ of the matrix $\lambda E_n - A$. Specifically, if the roots of $\Delta(\cdot)$ are distinct, then the minimal polynomial of $\mathcal{S}^n$ with respect to A is $\Delta(\cdot)$. This follows from the fact that the minimal polynomial of $\mathcal{S}^n$ is defined to be the lowest order polynomial which satisfies:

$$\Psi(\mathsf{A})\, \mathcal{S}^n = \phi \qquad (2.3.36)$$

In matrix terms, $\mathcal{S}^n$ corresponds to the identity matrix and (2.3.36) becomes:

$$\Psi(A)E_n = \Psi(A) = A^n - \alpha_{n-1} A^{n-1} \cdots - \alpha_1 A - \alpha_0 E_n = 0 \qquad (2.3.37)$$

The Cayley-Hamilton Theorem guarantees that $\Delta(\cdot)$ is the lowest order polynomial which satisfies (2.3.37) if the roots of $\Delta(\cdot)$ are distinct; hence $\Psi(\cdot)$ is equal to $\Delta(\cdot)$. If the roots of $\Delta(\cdot)$ are not distinct, the minimal polynomial of $\mathcal{S}^n$ is a divisor of $\Delta(\cdot)$. If $\mathcal{R}$ is a subspace of $\mathcal{S}^n$, the minimal polynomial of $\mathcal{R}$ is a divisor of both the minimal polynomial of $\mathcal{S}^n$ and of $\Delta(\cdot)$.

The most important concept introduced in Section 2.3.2 is that of an invariant subspace. The subspace $\mathcal{R} \subset \mathcal{S}^n$ is invariant with respect to A if the image of $\mathcal{R}$ under A is a subspace of $\mathcal{R}$:

$$\mathsf{A}\mathcal{R} \subset \mathcal{R} \qquad (2.3.38)$$

The matrix equivalent is that there must exist a matrix Π such that:

$$AR = R\Pi \tag{2.3.39}$$

Since the range space of $R\Pi$ is a subspace of the range space of R for any matrix Π, the range space of AR is a subspace of the range space of R and (2.3.39) is equivalent to (2.3.38). It can be shown that the eigenvalues of Π are a subset of the eigenvalues of A. If $\mathcal{R}$ is not invariant with respect to A, the matrix equivalent of:

$$A\mathcal{R} \subset \mathcal{R} \oplus \mathcal{S} \tag{2.3.40}$$

is the matrix equation:

$$A R = R\Pi + \circledS \tag{2.3.41}$$

In this case the eigenvalues of Π are not a subset of the eigenvalues of A. If:

$$\mathcal{S}^n = \mathcal{R} \oplus \mathcal{S}$$

it is simple to show that Π in (2.3.39) is A_{RR} in (2.3.35) and that $\circledS$ in (2.3.41) is related to A_{SR}.

2.4 Observability and Equivalent Realizations

The discussion in Section 2.1 was stated in terms of the controllable subspace W of an event vector f_1. For the moment, assume that Cf_1 is not zero. Section 3.2 demonstrates that as long as CW is not zero, it is a simple matter to choose a detection filter D for f_1. A more difficult problem is to determine if the failure detection system can uniquely associate a unidirectional output error along the output direction Cf_1 with the event vector f_1 and, hence, with the failure associated with f_1. If it cannot, it is important to know which additional event vectors can generate unidirectional output errors along Cf_1 and whether or not any physically meaningful failures are associated with these event vectors. This problem, as well as the problem of eigenvalue assignability for A-DC, requires

a careful definition of several observability concepts. The concepts of this section are stated using matrix notation. The analogous results for linear operators are used in Chapter 3 without further comment.

A system is completely observable (Ogata 1967) on the time interval $t_0 \leq t \leq t_1$ if for every t_0 and some t_1, every initial state $x(t_0)$ can be determined from knowledge of the system output for that time interval. In essence a system is completely observable if every transition of the system's state eventually affects the output. A number of criteria have been developed for testing observability. The most generally used for time invariant systems is similar to the controllability criterion used in Section 2.1. With M defined as:

$$M = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \qquad (2.4.1)$$

the system (A, B, C) is completely observable if and only if M is of rank n. If the rank is less, there is a theorem (Brockett (1970)) which says that any initial state $x(t_0)$ can be determined from the system output to within an additive constant vector which lies in $\eta(M)$.

The null space of M is invariant with respect to A--from an inspection of M, it follows that if $\xi \in \eta(M)$, then $A\xi \in \eta(M)$. It is a simple matter to show that no initial condition $x(t_0) \in \eta(M)$ can ever affect the output of the system (A, B, C). For this reason, $\eta(M)$ is said to be the unobservable space of (A, B, C).

Various definitions of the observable space of (A, B, C) are found in the literature. A simple choice would be the row space of M. For the failure detection problem a more general definition is desirable:

Definition 2.4.1 A subspace $\mathcal{P}$ is observable for (A, B, C) if:

$$\mathcal{R} \cap \eta(M) = \phi$$

An equivalent statement of the definition is:

Definition 2.4.1a Let $\mathcal{J}_\xi$ be the cyclic invariant subspace generated by some $\xi \in \mathcal{R}$. The subspace $\mathcal{R}$ is observable if for every $\xi \in \mathcal{R}$:

$$\mathcal{J}_\xi \not\subset \eta(C)$$

The first definition clearly implies the second since $\xi \notin \eta(M)$ is sufficient to insure that $\mathcal{J}_\xi \notin \eta(C)$. The opposite is also true. If $\xi \in \eta(M)$, then $\mathcal{J}_\xi \in \eta(C)$ because $\eta(M)$ is invariant. Therefore if $\mathcal{R}$ is not observable by the first definition, it is not observable by the second.

The intention of these definitions is to permit partial observability to be considered in some context other than the row space of M. The row space of M is orthogonal to $\eta(M)$; the subspace $\mathcal{R}$ need not be to be observable by Definition 2.4.1. As an example of the use of Definition 2.4.1a, consider the controllable subspace for the event vector f_1. The controllable subspace is spanned by the vectors f_1, $(A-DC)\,f_1$, $(A-DC)^2\,f_1$ ... and therefore is a cyclic-invariant subspace. With W given by (2.1.19), Definition 2.4.1a states that the controllable subspace of f_1 is observable as long as the rank of W is not zero, i.e., as long as the failure associated with f_1 affects the system output.

The proof of the Detection Theorem is based on a definition of system observability which is not normally found in the literature. This definition is useful because of the concern of failure detection theory with certain output directions:

Definition 2.4.2 Let $\mathcal{R}$ be an invariant subspace and $\mathcal{J}_\xi$ by the cyclic-invariant subspace generated by some vector $\xi \in \mathcal{S}^n$ which satisfies $\xi \notin \mathcal{R}$. The system (A, B, C) is observable (mod $\mathcal{R}$) if for every such ξ:

$$C\mathcal{J}_\xi \not\subset C\mathcal{R}$$

An equivalent condition is:

$$\mathcal{J}_\xi \not\subset \eta(C) \cup \mathcal{R}$$

(A, B, C) is observable (mod $\mathcal{R}$) if no event vector which is not in $\mathcal{R}$ can generate an output error which is constrained to the output subspace $C\mathcal{R}$. If D is a detection filter for f_1 and Cf_1 is not zero, Section 3.2 uses Definition 2.4.2 to determine if any event vectors other than f_1 can generate unidirectional output errors along Cf_1.

The material on observability is concluded with two theorems dealing with output feedback. Since they are not commonly found in textbooks, they are presented formally. The first due to Wonham (1967) is stated in his paper as a theorem on controllability and state feedback. The theorem presented here is the dual of Wonham's.

Theorem 2.4.1 The system (A, B, C) is observable if and only if for every choice of a set Λ of n eigenvalues, there is a matrix D such that A-DC has Λ for its set of eigenvalues.

The theorem states necessary, but not sufficient, conditions for a failure to be detectable according to Definition 3.1.1.

The final theorem concludes that observability is independent of the choice of output feedback. Since no proof is readily available in the literature, a rigorous proof is included.

Theorem 2.4.2 Let the feedback matrix D be defined arbitrarily. The system (A-DC, B, C) is observable if and only if (A, B, C) is observable.

Proof Define the observability matrix for (A-DC, B, C) similarly to (2.4.1):

$$M_D = \begin{bmatrix} C \\ C(A-DC) \\ C(A-DC)^2 \\ \vdots \\ C(A-DC)^{n-1} \end{bmatrix}$$

It is sufficient to show that the unobservable spaces for the two systems are equal, i.e.:

$$\eta(M) = \eta(M_D)$$

a) To prove that $\eta(M) \subset \eta(M_D)$, let ξ be an arbitrary vector in $\eta(M)$. Then $C\xi$ is zero and it follows that:

$$C(A-DC)\xi = CA\xi$$

From (2.4.1) and the fact that $\xi \in \eta(M)$, $CA\xi$ must also be zero and:

i) $$C(A-DC)\xi = 0$$

Because $\xi \in \eta(M)$ implies that $C\xi$, $CA\xi$, $CA^2\xi$ ... are all zero, an expansion of $C(A-DC)^j\xi$ for arbitrary j yields the more general result:

ii) $$C(A-DC)^j\xi = CA^j\xi = 0 \qquad j = 0, 1, 2...$$

Therefore, $\xi \in \eta(M_D)$. Because $\xi \in \eta(M)$ was chosen arbitrarily, it follows that $\eta(M) \subset \eta(M_D)$.

b) Similarly, $\eta(M_D) \subset \eta(M)$ since if $\xi \in \eta(M_D)$, then $C\xi = \phi$ and the equality of ii) holds for all $\xi \in \eta(M_D)$.

By i) and ii),

$$\eta(M) = \eta(M_D)$$

and the theorem is proved.

Δ

A system is completely characterized by its input/output history. If a model of the system dynamics is assumed, that model is satisfactory only if it can reproduce that input/output history. If a model is not known or for some reason is not acceptable, the question becomes one of how to choose a model which does adequately describe the system behavior. A general theory has been developed

which defines the equivalent realizations of (A, B, C) for a given input/output history (Brockett (1970) and Arbib (1969)). Only a few special results for linear, time-invariant systems are needed here.

The system models (A, B, C) and $(\tilde{A}, \tilde{B}, \tilde{C})$ are equivalent realizations of a system if:

$$C\, e^{A(t-t_0)}\, B = \tilde{C}\, e^{\tilde{A}(t-t_0)}\, \tilde{B} \quad \forall\, t \geq t_0$$

This can be expressed in simple matrix form:

$$CA^j B = \tilde{C}\, \tilde{A}^j \tilde{B} \qquad j = 1, 2 \ldots \tag{2.4.2}$$

Because of the cyclic nature of A and $\tilde{A}$ it is not necessary to consider any $j \geq \max(n, \tilde{n})$. Two particular examples of equivalent realizations are of continuing interest. If T is a coordinate transformation, (A, B, C) and $(T^{-1}AT, T^{-1}B, CT)$ are equivalent. Also, (A, B, C) and $(\tilde{A}, \tilde{B}, \tilde{C})$ are equivalent where:

$$\tilde{C} = [C : 0] \qquad A = \left[\begin{array}{c:c} A & \Gamma \\ \hdashline 0 & A_1 \end{array}\right] \qquad B = \left[\begin{array}{c} B \\ \hdashline 0 \end{array}\right] \tag{2.4.3}$$

This second example illustrates that system models of different orders can be equivalent. A good deal of attention has been devoted to determining the smallest possible dimension for the state space of a model. The minimal realization is an observable and controllable system. A non-minimal realization cannot be both observable and controllable.

The minimal realization is of little value in failure detection theory. The smallest permissible order of the model is prescribed by the set of failures to be detected and is generally greater than the minimal order. Once the non-minimal model (A, B, C) is chosen, the results of Chapter 3 may indicate that complete eigenvalue assignment is possible only if the state space of the model is increased from n to $\tilde{n}$.

CHAPTER 3

FAILURE DETECTION THEORY

The body of failure detection theory is derived in this chapter. The first step is to complete the motivation for the definition of detectability and of the vector space techniques. The basic concept of a detection space associated with each event vector is introduced in Section 3.1. The detection space for an event vector f is comprised of the event vectors which cannot be discerned from f on the basis of output error direction alone. It is the sufficient characteristic by which a failure may be related to the system state space. The axiomatic definition of the detection space is the basis for the theoretical results on detection theory proved in this chapter. The premise is that of a continuous, linear, time invariant, deterministic system. The admission of stochastic properties does not alter the results of this chapter. The possibility of modification of the theory for discrete and stochastic systems is discussed in Chapters 4 and 5 respectively.

The basic theorem of the detectability of a single failure is proved in Section 3.2. The properties of the detection space are derived and algorithms for determining the detection space and assigning the eigenvalues are presented.

Detection of multiple failures with a single filter is treated in Section 3.3. A sufficient condition for a set of failures to be detectable by a single failure detection system is proved in Section 3.3.1. If the set of failures cannot be detected by a single failure detection system, Section 3.3.3 provides an algorithm for finding a subset of mutually detectable vectors from the undetectable set. Unfortunately, the algorithm does not guarantee that an acceptable number of failures can be included in this subset. Section 3.3.4 provides a stronger result by allowing for the possibility of adopting a reference model with

a dimension greater than n. As long as the failures satisfy a simple relationship, it is shown that a failure detection system can be defined for a set of m failures (m is the rank of C). It is proved in Sections 3.5 and 3.6 that this is the largest number of failures which can be detected unambiguously by a single failure detection system.

A modified definition of detectability is introduced in Section 3.4. This definition differs from Definition 2.1.1 in that a weaker control over the eigenvalue assignment for A-DC is accepted. The new definition does not require that the eigenvalues of A-DC be individually assignable, but only that they be assignable with a multiplicity of one or greater. The consequence of the change may be a several-fold increase in the number of failures detectable by a single failure detection system. Constructive algorithms are developed for determining the increased capability of a filter.

The principal theorems of Section 3.2 and Section 3.3.3 were proved by Beard (1971) using matrix algebra techniques--the remaining results are new. The approach and the derivations presented here are radically different from those of Beard. Although Beard's derivations are considerably more lengthy, a reader comfortable with matrix analysis might find them simpler to follow. Unfortunately, the transition from the matrix algebra approach back to the linear algebra approach upon which the new results are based would not be trivial.

It is intended that the applications-oriented reader should be able to use failure detection theory without reading past Section 3.2.1. Except for developing the computational algorithms, the detection problem is abstracted in this chapter to a study of the properties of an operator A-DC in an n-dimensional space. The important results are summarized and interpreted in Chapter 4. A careful reading of the previous chapters and of Sections 3.1 and 3.2.1 should suffice as preparation for following the developments of this chapter on an intuitive basis without reading the proofs of the theorems.

3.1 Detection Spaces

Section 2.1 motivated failure detection theory in terms of the controllable subspace of the event vector f_1 associated with a failure. In terms of the vector space notation introduced in Section 2.3, the controllable subspace of f_1 is the vector space $\mathcal{V}_1$ spanned by the vectors in the sequence f_1, $(A-DC)f_1$, $(A-DC)^2 f_1 \ldots$. The vector space equivalent of the failure detection requirement in (2.1.20) is for the output subspace $C\mathcal{V}_1$ to be one-dimensional:

$$d(C\mathcal{V}_1) = 1 \qquad (3.1.1)$$

Although (3.1.1) states a necessary and sufficient condition for D to be a detection filter for f_1, it does not yield an acceptable algorithm for assigning the eigenvalues of $A-DC$. In order to develop such an algorithm, it is desirable to state the failure detection problem in a more general context. The problem is most conveniently stated in terms of the detection space of f_1 as defined in this section.

The detection space of f_1 has a second application which further justifies its introduction. It was suggested in Section 2.1 that certain failures might be indistinguishable on the basis of output error direction alone. It is obvious that this difficulty is related to the difference between the dimension of the state space and the number of independent sensors in the system. If only a single sensor exists, no discrimination between sensor failures can be based upon the output error direction. If a full set of n linearly independent measurements is available, it is reasonable to suspect that little (in fact, no) ambiguity exists. The possible ambiguity is formally defined in Definition 2.1.1 in terms of the event vectors f_1 and f_2 associated with two different failures. The definition is applicable to failures modeled by either the controller failure model or the sensor failure model since Section 2.1 defines event vectors for both models.

Definition 3.1.1 Let f_1 and f_2 be event vectors associated with failures in a system which is modeled in the unfailed state by the triad (A, B, C). The event vector f_2 is detection equivalent to f_1 if:

i) every detection filter for f_1 is a detection filter for f_2, and

ii) the unidirectional output error generated by the failure associated with f_2 is in the same output direction as that associated with f_1.

If f_2 is detection equivalent to f_1, a failure detection system based on output error direction alone cannot discriminate between failures associated with the two event vectors. It can be shown that the detection space of f_1 contains all the event vectors which are detection equivalent to f_1.

The detection space of f_1 is denoted by $\bar{\mathcal{R}}_1$ and is defined axiomatically in Definition 3.1.2 in terms of a set of four properties which uniquely characterize it. To simplify the notation, the definition assumes that Cf_1 is not zero. The modification of the definition which is necessary if Cf_1 is zero is discussed later in this section.

Definition 3.1.2 Let f_1 be an event vector associated with a failure in a system which is modeled in the unfailed state by (A, B, C). Assume that Cf_1 is not zero. The detection space for f_1 is denoted by $\bar{\mathcal{R}}_1$ and is the direct sum:

i) $\bar{\mathcal{R}}_1 = f_1 \oplus \mathcal{R}_1$

where $\mathcal{R}_1 \subset \mathcal{S}^n$ is the largest subspace which satisfies the three conditions:[1]

ii) $\eta(M) \cap \mathcal{R}_1 = \phi$

iii) $\mathcal{R}_1 \subset \eta(C)$

iv) $A\,\mathcal{R}_1 \subset \bar{\mathcal{R}}_1$

[1] The observability operator $\eta(M)$ is defined in Section 2.4

The remainder of this section is concerned with a motivation of Definition 3.1.2 in terms of the controllable subspace of f_1 and the vectors which are detection equivalent to f_1. Algorithms for finding $\bar{\mathcal{R}}_1$ are based upon Definition 3.1.2 and are developed in Section 3.2.

Condition i) of Definition 3.1.2 insures that f_1 is a vector in its own detection space. The use of the direct sum operator "$\oplus$" in i) rather than the more general union operator "$\cup$" is possible because of iii). To see this, assume that "$\cup$" is used in i) and note that the partition of $\bar{\mathcal{R}}_1$ specified by i) allows $C\bar{\mathcal{R}}_1$ to be expressed:

$$C\bar{\mathcal{R}}_1 = C(f_1 \cup \mathcal{R}_1) = Cf_1 \cup C\mathcal{R}_1 \qquad (3.1.2)$$

Because the statement of the definition requires that Cf_1 is not zero and iii) defines $\mathcal{R}_1$ such that $C\mathcal{R}_1$ is zero, (3.1.2) becomes:

$$C\bar{\mathcal{R}}_1 = Cf_1 \cup \phi = Cf_1$$

$$d(C\bar{\mathcal{R}}_1) = 1 \qquad (3.1.3)$$

This distinction between Cf_1 and $C\mathcal{R}_1$ is sufficient to guarantee that:

$$d(\bar{\mathcal{R}}_1) = d(f_1) + d(\mathcal{R}_1) = 1 + d(\mathcal{R}_1) \qquad (3.1.4)$$

Comparison of (3.1.4) with (2.3.2) verifies that $\bar{\mathcal{R}}_1$ is the direct sum of f_1 and $\mathcal{R}_1$.

Condition ii) is in the form of Definition 2.4.1 and defines $\mathcal{R}_1$ to be an observable subspace for (A, B, C). The event vector f_1 is also an observable subspace for (A, B, C) by virtue of the fact that Cf_1 is not zero. The direct sum of two observable subspaces does not necessarily satisfy Definition 2.4.1, but it is simple to show that $\bar{\mathcal{R}}_1$ is also an observable subspace for (A, B, C). Because $\bar{\mathcal{R}}_1$ is the direct sum of f_1 and $\mathcal{R}_1$, every

vector $\bar{\xi}$ in $\bar{\mathcal{R}}_1$ can be uniquely represented as a linear combination of f_1 and some vector ξ in $\mathcal{R}_1$:

$$\bar{\xi} = \alpha_f f_1 + \alpha_\xi \xi \tag{3.1.5}$$

If α_f is zero, then $\bar{\xi}$ is contained in $\mathcal{R}_1$ and the fact that $\mathcal{R}_1$ is observable for (A, B, C) guarantees that $\bar{\xi} \notin \eta(M)$. On the other hand, if α_f is not zero, ii) and the fact that $C f_1$ is not zero imply that:

$$C\bar{\xi} = \alpha_f C f_1 + \alpha_\xi \phi = \alpha_f C f_1 + \phi \tag{3.1.6}$$

and again $\bar{\xi} \notin \eta(M)$. Since these two cases include every vector in $\bar{\mathcal{R}}_1$, they imply that:

$$\bar{\mathcal{R}}_1 \cap \eta(M) = \phi \tag{3.1.7}$$

and consequently that $\bar{\mathcal{R}}_1$ is an observable subspace for (A, B, C).

With these points established, it is possible to relate Definition 3.1.2 to the failure detection problem. Because iii) states that $C\mathcal{R}_1$ is zero, iv) is equivalent to requiring that $\mathcal{R}_1$ satisfy:

$$(A-DC)\mathcal{R}_1 \subset \bar{\mathcal{R}}_1 \tag{3.1.8}$$

Assume for the moment that D is chosen such that:

$$(A-DC)f_1 = \bar{\xi}_f \in \bar{\mathcal{R}}_1 \tag{3.1.9}$$

for some arbitrary $\bar{\xi}_f$ in $\bar{\mathcal{R}}_1$. From (3.1.8), (3.1.9), and i) it follows that $\bar{\mathcal{R}}_1$ is an invariant subspace with respect to $A-DC$ for this choice of D since:

$$(A-DC)(f_1 \oplus \mathcal{R}_1) = (A-DC)\bar{\mathcal{R}}_1 \subset \bar{\mathcal{R}}_1 \tag{3.1.10}$$

This fact in turn insures that all of the vectors f_1, $(A-DC)f_1$, $(A-DC)^2f_1$... are contained in $\bar{\mathcal{R}}_1$. Since these vectors span the controllable subspace of f_1, it follows that $\mathcal{V}_1$ is a subspace of $\bar{\mathcal{R}}_1$:

$$\mathcal{V}_1 \subset \bar{\mathcal{R}}_1 \tag{3.1.11}$$

Coupled with (3.1.3), this implies that $\mathcal{V}_1$ satisfies (3.1.1) and, hence, that D is a detection filter for f_1.

It is shown in Section 3.2 that (3.1.9) is equivalent to (3.1.1) as a means of defining a detection filter for f_1. For every $\bar{\xi}_f$ in $\bar{\mathcal{R}}_1$ there is at least one detection filter for f_1 which satisfies (3.1.9) with that $\bar{\xi}_f$. Furthermore, there are no detection filters for f_1 which do not satisfy (3.1.9) for some $\bar{\xi}_f$ in $\bar{\mathcal{R}}_1$. It follows that if D is a detection filter for f_1, then $\bar{\mathcal{R}}_1$ is invariant with respect to $A-DC$ and $\mathcal{V}_1$ is a subspace of $\bar{\mathcal{R}}_1$. The advantage of (3.1.9) over (3.1.1) is the relative ease with which the desired eigenvalues of $A-DC$ can be associated with a particular $\bar{\xi}_f$ in $\bar{\mathcal{R}}_1$. The algorithm for selecting D developed in this chapter is based upon (3.1.9).

The use of $\bar{\mathcal{R}}_1$ to characterize the failure detection problem also simplifies the task of determining the event vectors which are detection equivalent to f_1. Let f_2 be an event vector associated with some failure and let $\mathcal{V}_2$ denote the controllable subspace of f_2. If D is a detection filter for f_1 and if f_2 is contained in $\bar{\mathcal{R}}_1$, an argument identical to that leading to (3.1.11) shows that $\mathcal{V}_2$ is a subspace of $\bar{\mathcal{R}}_1$:

$$\mathcal{V}_2 \subset \bar{\mathcal{R}}_1 \tag{3.1.12}$$

Furthermore, because (3.1.7) guarantees that $\bar{\mathcal{R}}_1$ is an observable subspace, the failure associated with f_2 generates a non-zero output error, i.e. f_2 is not in the unobservable space of C. It follows from (3.1.3) and (3.1.12) that D is a detection filter for f_2 and that the output error direction associated with f_2 is Cf_1:

$$C\mathcal{V}_2 = C\bar{\mathcal{R}}_1 = Cf_1 \qquad (3.1.13)$$

This is the same output error direction associated with f_1 and by Definition 3.1.1, f_2 is detection equivalent to f_1.

It is shown in Section 3.2 that an event vector cannot be detection equivalent to f_1 unless it is contained in $\bar{\mathcal{R}}_1$. Since (3.1.12) and (3.1.13) are valid for an arbitrary event vector in $\bar{\mathcal{R}}_1$, it follows that the detection space of f_1 is comprised of all the event vectors which are detection equivalent to f_1. In many instances $\mathcal{R}_1$ is zero and f_1 is its own detection space. This corresponds to the case in which no event vectors are detection equivalent to f_1. This case is important because it guarantees that a failure detection system can uniquely associate a unidirectional error in the direction Cf_1 with the failure associated with f_1. Although no general rule can be stated, the fact that $\mathcal{R}_1$ is a subspace of $\eta(C)$ implies that $\mathcal{R}_1$ is most likely to be zero if m is almost as large as n, i.e. if there are almost as many independent measurements as there are state variables. If m equals n, $\eta(C)$ is zero and therefore $\mathcal{R}_1$ must be zero. The case where m equals n is a trivial application of failure detection theory and is treated in Section 4.5.

The development of the properties of $\bar{\mathcal{R}}_1$ has been predicated upon the assumption that Cf_1 is not zero. For notational convenience, the results derived in this chapter are based upon this assumption, but it is simple to extend them to the more general case. If Cf_1 is zero, then:

$$(A-DC)f_1 = Af_1 \qquad (3.1.14)$$

and $\mathcal{V}_1$ is spanned by the sequence of vectors f_1, Af_1, $(A-DC)Af_1$, $(A-DC)^2Af_1$.... Similarly, if Cf_1, CAf_1,.....$CA^{\mu-2}f_1$ are all zero, it is possible to show that:

$$(A-DC)^k f_1 = A^k f_1 \quad k \leq \mu \qquad (3.1.15)$$

and that $\mathcal{H}_1$ is spanned by the sequence of vectors f_1, Af_1,... $A^{\mu-2}f_1$, $A^{\mu-1}f_1$,$(A-DC)A^{\mu-1}f_1$,$(A-DC)^2A^{\mu-1}f_1$... The output space, $C\mathcal{H}_1$ associated with $\mathcal{H}_1$ in this case is spanned by the corresponding sequence ϕ, ϕ...ϕ, $CA^{\mu-1}f_1$,$C(A-DC)A^{\mu-1}f_1$, $C(A-DC)^2A^{\mu-1}f_1$.... The primary importance to failure detection theory of the zero vectors in the sequence for $C\mathcal{H}_1$ is that it focusses attention on the output direction $CA^{\mu-1}f_1$ rather than Cf_1. This shift of attention can be emphasized by modifying Definition 3.1.2:

Definition 3.1.2a Let f_1 be an event vector associated with a failure in a system which is modeled in the unfailed state by (A, B, C) and assume that μ is the smallest integer such that $CA^{\mu-1}f_1$ is not zero. The detection space for f_1 is denoted by $\bar{\mathcal{R}}_1$ and is the direct sum:

$$\text{i)}\quad \bar{\mathcal{R}}_1 = A^{\mu-1}f_1 \oplus \mathcal{R}_1$$

where $\mathcal{R}_1 \subset \mathcal{S}^n$ is the largest subspace which satisfies the three conditions:

$$\text{ii)}\quad \eta(M)\cap\mathcal{R}_1 = \phi$$

$$\text{iii)}\quad \mathcal{R}_1 \subset \eta(C)$$

$$\text{iv)}\quad A\mathcal{R}_1 \subset \bar{\mathcal{R}}_1$$

It is simple to show that the vectors f_1, Af_1...$A^{\mu-2}f_1$ satisfy ii) - iv) and are contained in $\mathcal{R}_1$. In general, however, they span only a subspace of $\mathcal{R}_1$ and the algorithm developed in Section 3.2 must still be used to determine $\mathcal{R}_1$ and $\bar{\mathcal{R}}_1$.

The definitions introduced in this section can be illustrated with the example used in Section 2.1. When specific applications of failure detection theory are considered, the system is usually expressed in matrix form. Similarly, for computational purposes, it is more convenient to deal with detection spaces and detection filters in a matrix context than in terms of vector spaces and linear operators. The theory developed in this chapter is based upon the vector space techniques summarized in Section 2.3, but the final product is a set of matrix algorithms based upon

the matrix properties summarized in Section 2.3.3. The application of these algorithms to the example in Section 2.1 is analyzed in detail in Section 4.3.

The system in Section 2.1 is characterized by the matrices:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin\gamma & \sin\gamma \\ 0 & 0 & 0 \\ 1 & \cos\gamma & \cos\gamma \end{bmatrix} \tag{2.1.9}$$

The specific example used here is an off-failure of the first control element as modeled by the controller failure model and (2.1.12):

$$\dot{x}(t) = Ax(t) + Bu(t) + f_1 n_1(t) \tag{2.1.7}$$

$$y(t) = Cx(t)$$

$$f_1 n_1(t) = - \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} u_1(t) \tag{2.1.12}$$

Since Cf_1 is zero, it is necessary to use Af_1 and Definition 3.1.2a in order to define the detection space of f_1:

$$Af_1 = \begin{bmatrix} 0 \\ 2\omega \\ 1 \\ 0 \end{bmatrix} \tag{3.1.16}$$

It is simple to verify that CAf_1 is not zero. It can be shown that $\mathcal{R}_1$ for this example is a one-dimensional subspace which is spanned by the vector f_1. In the coordinate system on which

(A, B, C) is based, the detection space $\bar{R}_1$ of f_1 corresponds to the range space of the matrix $\bar{R}_1$:

$$\bar{R}_1 = [Af_1 \quad R_1] = [Af_1 \quad f_1] = \begin{bmatrix} 0 & 0 \\ 2\omega & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \tag{3.1.17}$$

Although a detection space cannot always be so conveniently related to its event vector, (3.1.17) illustrates the dynamic relationship between f_1 and the other event vectors in $\bar{R}_1$. Assume that $u_1(t)$ is a unit impulse at time t_0:

$$u_1(t) = \delta(t-t_0) \tag{3.1.18}$$

With zero initial conditions, the estimation errors generated by the controller failure can be found from (2.1.17):

$$\epsilon(t) = -\int_{t_0}^{t} e^{(A-DC)(t-\tau)} f_1 \, \delta(t_0) \, d\tau$$

$$\epsilon'(t) = C\epsilon(t) \tag{3.1.19}$$

Equation (3.1.19) is valid whether or not D is a detection filter for f_1. The presence of the impulse makes the integral simple to evaluate:

$$\epsilon(t) = -e^{(A-DC)(t-t_0)} f_1$$

$$\epsilon'(t) = C\epsilon(t) \tag{3.1.20}$$

Because Cf_1 is zero, (3.1.20) is precisely the same estimation error which would result from a failure modeled by:

$$\dot{x}(t) = Ax(t) + Bu(t) - Af_1\, n_f(t)$$

$$y(t) = Cx(t) \qquad (3.1.21)$$

if the time function $n_f(t)$ were the unit step:

$$n_f(t) = \begin{cases} 0 & t < t_0 \\ 1 & t \geq t_0 \end{cases} \qquad (3.1.22)$$

In essence, this is the property defined as detection equivalence in Definition 3.1.1. The event vector Af_1 is detection equivalent to f_1 because the output error given by (3.1.20) can be attributed to a failure modeled by (2.1.7) and (2.1.12), or by (3.1.21) and (3.1.22) -- or by some other failure model based upon a linear combination of f_1 and Af_1. If (2.1.7) and (3.1.21) are failure models for two different failures, no failure detection system can distinguish between the output errors generated by the two failures without specific knowledge of the possible values of $n_1(t)$ and $n_f(t)$ -- and this knowledge may not be sufficient to allow a decision to be made. The fact that $\bar{R}_1$ contains every event vector which is detection equivalent to f_1 insures that the output error direction alone is sufficient information to distinguish the controller failure modeled by (2.1.12) from every failure which is not modeled by either (3.1.21) or a linear combination of (2.1.7) and (3.1.21).

3.2 Single Failure Detection

In this section the basic result in failure detection theory is derived: A failure is detectable if and only if the system (A, C) is observable. This simple result guarantees the ability to choose D to assign the eigenvalues of A-DC almost arbitrarily while insuring that the output estimation error due to the failure is unidirectional. The Detection Theorem is proved in Section 3.2.1 and the properties of the result are discussed. A generalization

of the notion of detectability (Definition 2.1.1) is made in Section 3.2.2 to extend the Detection Theorem to unobservable reference models. The modification amounts to an acceptance of the fact that not all of the eigenvalues of an unobservable reference model can be assigned by output feedback (Theorem 2.4.1). Algorithms for determining $\bar{\mathcal{R}}_f$ and assigning the eigenvalues of $A-DC$ are developed in 3.2.3. The derivations in this section are in terms of an arbitrary event vector f and its detection space $\bar{\mathcal{R}}_f$. For convenience it is assumed that $C\,f \neq \phi$.

3.2.1 Detection Theorem

The proof of the Detection Theorem is based upon a set of three lemmas. The first is a general result for linear operators in a vector space and has a well-known matrix theory counterpart. Given two sets of vectors $y_1 \ldots y_q \in \mathcal{S}^m$ and $x_1 \ldots x_q \in \mathcal{S}^n$, the lemma states a necessary and sufficient condition for there to exist an operator D: $\mathcal{S}^m \rightarrow \mathcal{S}^n$ such that:

$$x_i = D\,y_i \qquad i = 1 \ldots q \tag{3.2.1}$$

The lemma does not require $y_1 \ldots y_q$ to be linearly independent--if they are, a solution always exists. The proof involves the construction of two operators which map the q-dimensional coefficient space $\mathcal{S}^Q$ into the vector spaces $\mathcal{S}^n$ and $\mathcal{S}^m$ respectively.[1] It adds little to the understanding of detection theory and may be bypassed.

Lemma 3.2.1 Let $x_i \in \mathcal{S}^n$ and $y_i \in \mathcal{S}^m$, where $i = 1 \ldots q$, be arbitrarily chosen sets of vectors and define the operators X: $\mathcal{S}^Q \rightarrow \mathcal{S}^n$ and Y: $\mathcal{S}^Q \rightarrow \mathcal{S}^m$:

$$X = \begin{bmatrix} x_1 & x_2 & \cdots & x_q \end{bmatrix}$$

$$Y = \begin{bmatrix} y_1 & y_2 & \cdots & y_q \end{bmatrix}$$

[1] The elements of a q-dimensional coefficient space $\mathcal{S}^Q$ are q-tuplets of coefficients $\{\alpha_{11}, \alpha_{12} \cdots \alpha_{1q}\}$.

There exists an operator D such that:

$$D y_i = x_i \qquad i = 1 \ldots q$$

if and only if $\eta(Y) \subset \eta(X)$. In particular, a solution always exists if the y_i are independent since then $\eta(Y) = \phi$.

Proof

Necessity The operators X and Y are operators from the coefficient space $\mathcal{Q}^Q$ into $\mathcal{S}^m$ and $\mathcal{S}^n$ respectively. If a D which satisfies the hypothesis exists, then by linearity:

$$DY\,\alpha = \alpha$$

for all $\alpha \in \mathcal{S}^Q$.

Define a partition of $\eta(Y)$:

$$\eta(Y) = \eta'(Y) \oplus \eta''(Y)$$

such that $\eta'(Y)$ is a subspace of $\eta(X)$ and that no vector in $\eta''(Y)$ is in $\eta(X)$:

$$\eta'(Y) \subset \eta(X)$$

$$\eta''(Y) \cap \eta(X) = \phi$$

If $\eta''(Y)$ is not the null space ϕ, then let $\alpha_f \in \mathcal{S}^Q$ be an arbitrary q-tuple in $\eta''(Y)$. Since by definition:

$$Y\alpha_f = \phi$$

it follows that:

$$DY\,\alpha_f = D\phi = \phi$$

But by construction $\alpha_f \notin \eta(X)$ and therefore:

$$X\ \alpha_f \neq \phi$$

and no satisfactory D exists. Therefore a necessary condition for a satisfactory D to exist is that $\eta''(Y)$ be the null space ϕ, i.e.:

$$\eta(Y) = \eta'(Y) \subset \eta(X)$$

Sufficiency Let $\mathcal{Y}$ be the r-dimensional subspace spanned by $y_1 \ldots y_q$ and $\mathcal{X}$ the s-dimensional subspace spanned by $x_1 \ldots x_q$. Clearly $\eta(Y) \subset \eta(X)$ implies $r \geq s$. If $y'_1, y'_2 \ldots y'_r$ denote the first r independent vectors in the set $y_1 \ldots y_q$; choose $\alpha_1 \ldots \alpha_r$ such that:

$$Y\alpha_i = y'_i \quad i = 1, \ldots r$$

A set of coefficients $\alpha_{r+1}, \ldots \alpha_q$ which span $\eta(Y)$ can also be chosen:

$$Y\alpha_j = \phi \quad j = r+1, \ldots q$$

The q q-tuples $\alpha_1, \ldots \alpha_r, \alpha_{r+1}, \ldots \alpha_q$ form a basic for $\mathcal{S}^Q$. Since $\eta(Y) \subset \eta(X)$, then:

$$X\alpha_j = \phi \qquad j = r+1, \ldots q$$

and if the vectors $x'_1 \ldots x'_r$ are defined by:

$$X\alpha_i = x'_i \quad i = 1 \ldots r$$

some subset of $x'_1 \ldots x'_r$ must form a basis for $\mathcal{X}$. The operator $D: \mathcal{S}^m \rightarrow \mathcal{S}^n$ is then any solution to the r equations:

$$D\, y'_i = x'_i \qquad i = 1 \ldots r$$

If $r = m$, D is unique.

Δ

The following lemma formalizes the statements made in Section 3.1 concerning the nature of the detection space $\bar{\mathcal{R}}_f$ of an event vector f. Definition 3.1.2 and a trivial application of Lemma 3.2.1 are used to prove that there exists an operator D such that $\bar{\mathcal{R}}_f$ is invariant with respect to A-DC:

$$(A\text{-}DC)\,\bar{\mathcal{R}}_f \subset \bar{\mathcal{R}}_f \tag{3.1.8}$$

The controllable subspace $\mathcal{U}_f$ of f is spanned by f, $(A\text{-}DC)f$, $(A\text{-}DC)^2 f \ldots$ and by virtue of (3.1.8) must be a subspace of $\bar{\mathcal{R}}_f$. Equation (3.1.3) then guarantees:

$$d(C\mathcal{U}_f) = 1 \tag{3.1.1}$$

and the output error generated by the failure associated with f is unidirectional. This satisfies the unidirectionality requirement of Definition 2.1.1.

Lemma 3.2.2 There exists a D such that $\bar{\mathcal{R}}_f$ is invariant with respect to A-DC and observable for the system (A-DC, C). The failure associated with f generates unidirectional output along Cf for (A-DC, C).

Proof

Since $Cf \neq \phi$, Lemma 3.2.1 guarantees that for any ξ_R there exists a corresponding D such that:

$$DCf = Af + \xi_R$$

In particular, there exists a satisfactory D for any $\xi_R \in \bar{\mathcal{R}}_f$. For that D:

i)' $\qquad (A\text{-}DC)f = -\xi_R \in \bar{\mathcal{R}}_f$

Condition iii) of Definition 3.1.2 states that $\mathcal{R}_f$ is a subspace of $\eta(C)$. Consequently:

$$(A\text{-}DC)\mathcal{R}_f = A\mathcal{R}_f$$

and $A\mathcal{R}_f$ in condition iv) can be replaced by $(A\text{-}DC)\mathcal{R}_f$:

ii)' $$(A\text{-}DC)\mathcal{R}_f \subset \bar{\mathcal{R}}_f$$

Equations i)' and ii)' coupled with condition i) of Definition 3.1.2 imply that $\bar{\mathcal{R}}_f$ is invariant with respect to $A\text{-}DC$:

$$(A\text{-}DC)\bar{\mathcal{R}}_f \subset \bar{\mathcal{R}}_f$$

This in turn implies that:

$$(A\text{-}DC)^j f \in \bar{\mathcal{R}}_f$$

for all $j > 0$ and the controllable subspace $\mathcal{V}_f$ of f is a subspace of $\bar{\mathcal{R}}_f$. Since:

$$d(C\bar{\mathcal{R}}_f) = d(Cf) = 1$$

the failure associated with f generates unidirectional output errors along Cf.

Observability follows from Theorem 2.4.2 and Definition 2.4.1. Suppose that $\bar{\mathcal{R}}_f$ is not observable for (A-DC, C) and let $\xi_D \in \bar{\mathcal{R}}_f$ satisfy $\xi_D \in \mathcal{N}(M_D)$. Then Theorem 2.4.2 requires $\xi_D \in \mathcal{N}(M)$ -- which contradicts condition ii) of Definition 3.1.2. Therefore, $\bar{\mathcal{R}}_f$ is observable for (A-DC, C). Δ

Lemma 3.2.2 satisfies the first condition of Definition 2.1.1 -- that a feedback matrix be chosen such that the output generated by a particular failure is unidirectional. It remains to show that the eigenvalues of $A\text{-}DC$ can be assigned arbitrarily. Since $\bar{\mathcal{R}}_f$ is invariant for $A\text{-}DC$, there are a set of:

$$\nu_f = d(\bar{\mathcal{R}}_f) \qquad (3.2.2)$$

eigenvalues of $A\text{-}DC$ associated with $\bar{\mathcal{R}}_f$; the remaining $n\text{-}\nu_f$ are associated with some arbitrary completion of $\mathcal{B}^n$. If $\mathcal{B}^n$ is partitioned:

$$\mathcal{B}^n = \bar{\mathcal{R}}_f \oplus \bar{\mathcal{R}}_f^c \tag{3.2.3}$$

the additional $n\text{-}\nu_f$ eigenvalues are associated with $\bar{\mathcal{R}}_f^c$. It is a simple matter to show that the eigenvalues associated with $\bar{\mathcal{R}}_f$ can be assigned arbitrarily by proper choice of D; the question concerns the remaining $n\text{-}\nu_f$ eigenvalues. The following lemma provides the key to the answer.

Assume for the moment that (A, C) is observable. The essential implication of Lemma 3.2.3 is that if D is chosen to satisfy Lemma 3.2.2 for f, and f′ is some arbitrary event vector not contained in $\bar{\mathcal{R}}_f$, then the failure associated with f′ cannot generate a unidirectional output error along Cf for (A-DC, C) -- the output error must have a component along some direction other than Cf. In terms of Definition 3.1.1, f′ is not detection equivalent to f unless f′ is an element of $\bar{\mathcal{R}}_f$. The proof of the Detection Theorem demonstrates that this fact is sufficient to insure eigenvalue assignability. The first corollary to the Detection Theorem states the above interpretation of Lemma 3.2.3 formally.

The statement of Lemma 3.2.3 employs a more restrictive definition of observability than the standard definition. The requirement is a formalization of the discussion above: A system (A-DC, C) is observable (mod $\bar{\mathcal{R}}_f$) if no event vector f′ which is not contained in $\bar{\mathcal{R}}_f$ generates output errors which are constrained to the output subspace $C\bar{\mathcal{R}}_f$. If $\mathcal{L}_f'$ is the controllable subspace for f′, i.e. the space spanned by f′, $(A\text{-}DC)\,f'$, $(A\text{-}DC)^2 f'\ldots$, Definition 2.4.2 states the condition for observability (mod $\bar{\mathcal{R}}_f$) as:

$$C\mathcal{L}_f' \not\subset C\bar{\mathcal{R}}_f \tag{3.2.4}$$

for all $f' \notin \bar{\mathcal{R}}_f$. Lemma 3.2.3 proves that observability and observability (mod $\bar{\mathcal{R}}_f$) are equivalent definitions as long as $\bar{\mathcal{R}}_f$ is a detection space.

Lemma 3.2.3 Let D satisfy Lemma 3.2.2. The system (A-DC, C) is observable (mod $\bar{\mathcal{R}}_f$) if and only if it is observable.

Proof

Necessity If (A-DC, C) is not observable, then $\eta(M_D) \neq \phi$. $\eta(M_D)$ is an invariant subspace (Section 2.4) and it is disjoint from $\bar{\mathcal{R}}_f$ (condition ii) of Definition 3.1.2). Definition 2.4.2 clearly implies that (A-DC, C) is not observable (mod $\bar{\mathcal{R}}_f$)since:

$$C(\eta(M_D)) = \phi \subset C\bar{\mathcal{R}}_f$$

Sufficiency Assume (A-DC, C) is observable. If (A-DC, C) is not observable (mod $\bar{\mathcal{R}}_f$), the equivalent condition in Definition 2.4.2 states that there must exist some vector $f' \notin \bar{\mathcal{R}}_f$ such that $\mathcal{V}_f'$, the controllable (invariant) subspace of f' for (A-DC, C), satisfies:

$$\mathcal{V}_f' \subset \eta(C) \cup \bar{\mathcal{R}}_f$$

The fact that $\bar{\mathcal{R}}_f$ is a detection space means that conditions i) and iii) of Definition 3.1.2 can be used to reduce the above to a disjoint sum:

i)' $$\mathcal{V}_f' \subset \eta(C) \oplus f$$

Assume that i)' above is satisfied. Now consider the union of $\mathcal{V}_f'$ and $\bar{\mathcal{R}}_f$:

$$\bar{\mathcal{R}}_f' = \mathcal{V}_f' \cup \bar{\mathcal{R}}_f$$

By construction, $\mathcal{V}_f'$ and $\bar{\mathcal{R}}_f$ are invariant with respect to A-DC ; therefore, $\bar{\mathcal{R}}_f'$ must be also. Because of i)' above and condition i) of Definition 3.1.2, $\bar{\mathcal{R}}_f'$ can be partitioned similarly to $\bar{\mathcal{R}}_f$:

i) $\bar{\mathcal{R}}_f' = \mathcal{R}_f' \oplus f$

where:

iii) $\mathcal{R}_f' \subset \eta(C)$

Since (A-DC, C) is observable, Therem 2.4.2 requires (A, C) to be observable as well and:

ii) $\mathcal{R}_f' \cap \eta(M) = \phi$

Finally, since iii) above implies:

$$(A-DC)\mathcal{R}_f' = A\mathcal{R}_f'$$

the invariance of $\bar{\mathcal{R}}_f'$ with respect to $A-DC$ yields

iv) $A\ \mathcal{R}_f' \subset \bar{\mathcal{R}}_f'$

Conditions i) - iv) above are identical to conditions i) - iv) of Definition 3.1.2. Since $\bar{\mathcal{R}}_f$ is the largest subspace to satisfy these conditions, it is necessary that $\mathcal{L}_f' \subset \bar{\mathcal{R}}_f$ and therefore $f' \in \bar{\mathcal{R}}_f$ for all $\mathcal{L}_f$ which satisfy condition i)'. Consequently, (A-DC, C) is observable (mod $\bar{\mathcal{R}}_f$). Δ

Let $\mathcal{S}^n$ be partitioned as in (3.2.3), but with the additional constraint on the choice of $\bar{\mathcal{R}}_f^c$:

$$d(C\bar{\mathcal{R}}_f^c) = m-1 \qquad (3.2.5)$$

For an observable (A, C), Lemmas 3.2.2 and 3.2.3 insure that there exists a D such that $A-DC$ can be decomposed into two subsystems--one associated with $\bar{\mathcal{R}}_f$ and the other with $\bar{\mathcal{R}}_f^c$-- such that both systems are observable and the subsystem associated with $\bar{\mathcal{R}}_f$ is <u>decoupled</u>, i.e. an input command to that subsystem does not affect the state of the subsystem associated with $\bar{\mathcal{R}}_f^c$.

In matrix terms, if a basis for $\mathcal{S}^n$ is specified as ν_f linearly independent vectors from $\bar{\mathcal{R}}_f$ and n-ν_f from $\bar{\mathcal{R}}_f^c$, the matrix A-DC for that basis is upper block diagonal:

$$\text{A-DC} = \left[\begin{array}{c|c} A'_{11} & A'_{12} \\ \hline 0 & A'_{22} \end{array}\right]$$

$$C = [C_1 \mid C_2] \tag{3.2.6}$$

where A'_{11} is a $\nu_f \times \nu_f$ matrix, C_1 is $m \times \nu_f$ etc.[1] The subsystems (A'_{11}, C_1) and (A'_{22}, C_2) are observable with:

$$\text{rk } C_1 = 1$$

$$\text{rk } C_2 = m - 1 \tag{3.2.7}$$

A'_{11}, is associated with $\bar{\mathcal{R}}_f$ and A'_{22} is associated with $\bar{\mathcal{R}}_f^c$--f is a linear combination of the first ν_f unit vectors. The zero off-diagonal block reflects the fact that $\bar{\mathcal{R}}_f$ is invariant with respect to A-DC. Along with the fact that the rank of C_1 is unity, it also implies that the failure associated with f generates unidirectional output errors along Cf. These results are verified in the proof of Theorem 3.2.1.

The partitioning of (A,C) into subsystems is useful for proving the eigenvalue assignability required by Definition 2.1.1. In matrix notation the zero submatrix in (3.2.6) implies that the eigenvalues of A-DC are the sum of those of A'_{11} and A'_{22}. Theorem 2.4.1 and the observability of (A'_{11}, C_1) and (A'_{22}, C_2) then guarantee that matrices D_{21} and D_{22}

[1]The relationship between (3.2.6) and the definition of A-DC in terms of operators on $\bar{\mathcal{R}}_f$ and $\bar{\mathcal{R}}_f^c$ is outlined in Section 2.3.3.

(n x 1 and n x (m-1) respectively) exists such that the eigenvalues of $A'_{11} - D_{21} C_1$ and $A'_{22} - D_{22} C_2$ are any desired set. The proof of the Detection Theorem is completed by showing that Lemma 3.2.1 implies that a matrix D_2 exist such that:

$$A - (D+D_2) C = \left[\begin{array}{c|c} A'_{11} - D_{21}C_1 & A'_{12} \\ \hline 0 & A'_{22} - D_{22} C_2 \end{array}\right] \quad (3.2.8)$$

for any choice of D_{21} and D_{22}.

Theorem 3.2.1 (Detection Theorem) The failure associated with the event vector f is detectable if and only if (A, C) is observable.

Proof

Let $\mathbb{D}$ be any operator which satisfies Lemma 3.2.2 for f. A satisfactory $\mathbb{D}$ always exists. Define a partition of $\mathcal{S}^n$ in terms of $\bar{\mathcal{R}}_f$ and some subspace $\bar{\mathcal{R}}_f^c$, where $\bar{\mathcal{R}}_f$ and $\bar{\mathcal{R}}_f^c$ are disjoint:

$$\mathcal{S}^n = \bar{\mathcal{R}}_f \oplus \bar{\mathcal{R}}_f^c$$

For simplicity choose $\bar{\mathcal{R}}_f^c$ such that its image $C\bar{\mathcal{R}}_f^c$ in $\mathcal{S}^m$ does not include Cf. Since conditions i) and iii) of Definition 3.1.2 imply that:

$$C\,\bar{\mathcal{R}}_f = Cf$$

such a $\bar{\mathcal{R}}_f^c$ exists. $C\,\bar{\mathcal{R}}_f^c$ is then an (m-1) - dimensional subspace of $\mathcal{S}^m$:

$$C\bar{\mathcal{R}}_f^c = \mathcal{N}_f^c$$

With $\mathcal{N}_f^c$ defined, the partition of $\mathcal{S}^m$ is consistent with that of $\mathcal{S}^n$:

$$\mathcal{S}^m = Cf \oplus \mathcal{N}_f^c$$

Since $\bar{\mathcal{R}}_f$ is invariant with respect to A-DC, the operator A-DC is uniquely defined by four operators on the partitions of $\mathcal{S}^n$, one of which is the null operator:

i) $A'_{11}: \bar{\mathcal{R}}_f \to \bar{\mathcal{R}}_f$

ii) $A'_{22}: \bar{\mathcal{R}}_f^c \to \bar{\mathcal{R}}_f^c$

iii) $A'_{12}: \bar{\mathcal{R}}_f^c \to \bar{\mathcal{R}}_f$

iv) $\phi : \bar{\mathcal{R}}_f \to \bar{\mathcal{R}}_f^c$ (ϕ is the null operator)

C is uniquely defined by operators from $\bar{\mathcal{R}}_f$ and $\bar{\mathcal{R}}_f^c$ onto the appropriate output subspaces:

v) $C_1: \bar{\mathcal{R}}_f \to Cf$

vi) $C_2: \bar{\mathcal{R}}_f^c \to \mathcal{N}_f^c$

The subsystem of A-DC associated with $\bar{\mathcal{R}}_f$ can be denoted by (A'_{11}, C_1) and the one associated with $\bar{\mathcal{R}}_f^c$ by (A'_{22}, C_2). Because of the null operator in iv) the eigenvalues of A-DC are the eigenvalues of A'_{11} and A'_{22}.

Let D_2 be an arbitrary operator $D_2: \mathcal{S}^m \to \mathcal{S}^n$ defined on the same partitions as C in v) and vi):

vii) $D_{21}: Cf \to \bar{\mathcal{R}}_f$

vii) $D_{22}: \mathcal{N}_f^c \to \bar{\mathcal{R}}_f^c$

Clearly D_{21} and D_{22} can be chosen freely. The composite operator $D_2C: \mathcal{S}^n \to \mathcal{S}^n$ can then be defined as a pair of operators on $\bar{\mathcal{R}}_f$ and $\bar{\mathcal{R}}_f^c$:

ix) $D_{21}\, C_1: \bar{\mathcal{R}}_f \to \bar{\mathcal{R}}_f$

x) $D_{22}\, C_2 : \bar{\mathcal{R}}_f^c \to \bar{\mathcal{R}}_f^c$

Combining ix) and x) with i) - iv) above, $A-(D+D_2)C$ is uniquely defined by:

i)' $A'_{11} - D_{21}C_1 : \bar{\mathcal{R}}_f \rightarrow \bar{\mathcal{R}}_f$

ii)' $A'_{22} - D_{22}C_2 : \bar{\mathcal{R}}_f^c \rightarrow \bar{\mathcal{R}}_f^c$

iii)' $A'_{12} : \bar{\mathcal{R}}_f^c \rightarrow \bar{\mathcal{R}}_f$

iv)' $\phi : \bar{\mathcal{R}}_f \rightarrow \bar{\mathcal{R}}_f^c$

From iv)' it is apparent that $\bar{\mathcal{R}}_f$ is invariant with respect to $A-(D+D_2)C$ and that the eigenvalues of $A-(D+D_2)C$ are those of $A'_{11} - D_{21}C_1$ and $A'_{22}-D_{22}C_2$.

Let Λ_f and Λ_f^c be sets of ν_f and $n-\nu_f$ desired eigenvalues of $A-(D+D_2)C$ associated with $\bar{\mathcal{R}}_f$ and $\bar{\mathcal{R}}_f^c$ respectively. Theorem 2.4.1 states that there exists operators D_{21} and D_{22} such that the $A'_{11}-D_{21}C_1$ and $A'_{22}-D_{22}C_2$ have the desired eigenvalues for all sets Λ_f and Λ_f^c if and only if (A'_{11}, C_1) and (A'_{22}, C_2) are both observable. By Lemma 3.2.2, (A'_{11}, C_1) is always observable. From the output space associated with C_2 by v) and vi), Lemma 3.2.3 states that (A'_{22}, C_2) is observable if and only if $(A-DC, C)$ is observable. If follows that the eigenvalues of $A-(D+D_2)C$ are assignable if and only if $(A-DC, C)$ is observable. Theorem 2.4.2 then allows the observability of (A, C) to be substituted for that of $(A-DC, C)$.

Δ

The following corollary to the Detection Theorem insures that if (A, C) is observable no event vector $\rho \notin \bar{\mathcal{R}}_f$ can produce an error signal which is unidirectional along C_f. Thus, the detection space $\bar{\mathcal{R}}_f$ is the space of all event vectors which are detection equivalent to f--which is formally stated in Corollary 2. Corollary 2 provides the justification for referring to the detection space as the sufficient characteristic by which a failure detection problem can be formulated.

Corollary 1 Assume D satisfies Lemma 3.2.2. No vector $\rho \notin \bar{\mathcal{R}}_f$ can produce an output for (A-DC, C) which is unidirectional along Cf if (A, C) is observable.

Proof

The proof is the sufficiency proof for Lemma 3.2.3.

Δ

Corollary 2 The detection space $\bar{\mathcal{R}}_f$ is isomorphic to the set $\{\xi_f\}$ of vectors which are detection equivalent to f.

Proof

By Corollary 1 and the definition of detection equivalence (Definition 3.1.1), $\bar{\mathcal{R}}_f$ clearly contains every vector of $\{\xi_f\}$. Since Lemma 3.2.2 guarantees that $\bar{\mathcal{R}}_f$ is observable and invariant with respect to A-DC, every $\xi \in \bar{\mathcal{R}}_f$ produces a unidirectional output along Cf for every acceptable D. Thus every vector of $\bar{\mathcal{R}}_f$ is in $\{\xi_f\}$.

Δ

Only one structural property of detection spaces has been proved thus far: For every detection space there exists an operator D such that the space is invariant with respect to A-DC. It is possible to show that in fact the detection space becomes cyclic invariant and that the generator for $\bar{\mathcal{R}}_f$ with respect to A-DC is not unique[1]. It is also possible to show that there exists a unique vector g in $\bar{\mathcal{R}}_f$ such that the vectors g, Ag, ... A^{ν_f-1} g are a basis for $\bar{\mathcal{R}}_f$. The vector g was defined by Beard (1971) to be the detection generator for the failure associated with the event vector f. The detection generator is particularly useful in defining computational algorithms for choosing D. Its existence is proved in Lemma 3.2.4. An algorithm for computing g based upon the proof of Lemma 3.2.4 is given in Appendix A.2.

Lemma 3.2.4 Let $d(\bar{\mathcal{R}}_f) = \nu_f$. There exists a unique vector $g \in \bar{\mathcal{R}}_f$, called the detection generator of $\bar{\mathcal{R}}_f$, such that:

i) $A^k g \in \mathcal{R}_f \qquad k < \nu_f - 1$

[1]Cyclic invariant vector spaces are discussed in Section 2.3.2

ii) $CA^{\nu_f - 1} g = Cf$

The statements i) and ii) imply that the vectors $g, Ag \ldots A^{\nu_f - 1} g$ are a basis for $\bar{\mathcal{R}}_f$.

Proof

a) Existence

By conditions ii) and iv) of Definition 3.1.2, for every vector ξ in $\bar{\mathcal{R}}_f$, there is an integer $\nu \leq \nu_f$ such that ξ satisfies i) and ii) of the hypothesis with ν_f replaced by ν. Define g to be the vector associated with the largest integer $\nu_g \leq \nu_f$:

iii) $A^k g \in \mathcal{R}_f \qquad k < \nu_g - 1$

iv) $CA^{\nu_g - 1} g = Cf$

and let $\mathcal{R}_g$ be the vector space spanned by $g, Ag \ldots A^{\nu_g - 1} g$. It is simple to prove that ν_g must equal ν_f. Otherwise, there must exist some largest integer $\nu' > 0$ and a vector $g' \notin \mathcal{R}_g$ such that:

v) $A^k g' \in \mathcal{R}_f \qquad k < \nu' - 1$

vi) $CA^{\nu' - 1} g' = Cf$

However, the vector:

vii) $g'' = g' - A^{\nu_g - \nu'} g$

also must satisfy $g'' \notin \mathcal{R}_g$. Conditions iii) and v) clearly imply:

viii) $A^k g'' \in \mathcal{R}_f \qquad k < \nu' - 1$

But by the construction of g'':

ix) $CA^{\nu' - 1} g'' = CA^{\nu' - 1} g' - CA^{\nu_g - 1} g$

$= Cf - Cf = 0$

Since $\mathcal{R}_f$ is observable (Lemma 3.2.2) and $\alpha\, g'' \notin \mathcal{R}_g$ for arbitrary α, this contradicts the maximality of ν'.

b) Uniqueness

The detection generator must be unique. If there existed a second detection generator, g_α, then $g_\alpha - g$ would satisfy:

$$A^k(g_\alpha - g) \in \mathcal{R}_f \qquad k \leq \nu_f - 1$$

Since $d(\mathcal{R}_f) = \nu_f - 1$, this would imply that $\mathcal{R}_f$ is invariant and hence unobservable. $\mathcal{R}_f$ is observable and therefore $g_\alpha = g$.

Δ

3.2.2 Unobservable System

Unfortunately from a mathematical vantage point, the detection problem loses a number of desirable properties if the reference model (A, C) upon which the failure detection system is to be based is not observable. It is not necessary that the system be observable; only that the a suitable system model be observable. In practice an observable model can be obtained simply by eliminating the unobservable dynamics from the model. This does not alter the effectiveness of the failure detection system because failures in the unobservable part of a system cannot be detected. If for some reason it is necessary to design a detection filter for an unobservable system model, the difficulties mentioned below should be acknowledged.

Although it is not done here, it is possible to show that if (A, C) is not observable, the detection space as defined in Definition 3.1.2 need not be unique. This shortcoming can be eliminated in several ways. One is to modify the definition to insure a unique solution. As an example, it would be necessary to change condition i) to:

$$\text{i)}'\ \bar{\mathcal{R}}_f' = f' \oplus \mathcal{R}_f' \qquad (3.2.9)$$

where f′ is in the orthogonal complement of $\eta(M)$:

$$f' \in \eta(M)^{\perp}$$

$$f' \equiv f \pmod{\eta(M)}$$

The required change in condition ii) would be similar:

$$ii)' \quad \mathcal{R}'_f \subset \eta(M)^{\perp} \tag{3.2.10}$$

The end result would be to restrict attention to the projection of $\bar{\mathcal{R}}_f$ onto $\eta(M)^{\perp}$. $\eta(M)^{\perp}$ is the row space of M and is classically referred to as the observable space of (A, C). Such a specialization of the definition of a detection space would not add to the conceptual understanding of the failure detection problem and it would not allow any additional results to be obtained. The conclusions of this subsection demonstrate that when an algorithm is implemented, it is simple to choose $\bar{\mathcal{R}}_f \in \eta(M)^{\perp}$ as the detection space of f′ if that seems desirable.

The most algebraically consistent procedure for eliminating the ambiguity is to replace Definition 3.1.2 with a similar definition in terms of factor spaces. (Factor spaces are discussed in Section 2.3.2). Although there is no practical necessity for such a generalization, it is the logical extension of Section 3.2.1. The generalization to factor spaces is presented in the remainder of this section without rigorous proof. This material is not necessary to the remainder of the chapter.

If (A, C) is not observable, then $\bar{\mathcal{R}}_f^c$ can be chosen such that:

$$\mathcal{S}^n = \bar{\mathcal{R}}_f \oplus \bar{\mathcal{R}}_f^c \oplus \eta(M) \tag{3.2.11}$$

The unobservable space $\eta(M)$ of the system is invariant. This space and the eigenvalues associated with it constitute a subsystem which never affects the system output. It is clear that these eigenvalues can never be altered by output feedback. If:

$$\nu_o = d(\eta(M)) \qquad (3.2.12)$$

only $n-\nu_o$ eigenvalues of $A-DC$ can be assigned by choosing D. The implication of such free trajectory subsystems is studied in detail by Davison and Wang (1973).

In the proof of the Detection Theorem it was pointed out that the choice of $\bar{R}_f^c$ is not unique even if (A, C) is observable. From (3.2.9) it is clear that for an unobservable system the decomposition does not require $\bar{R}_f$ to be unique. Any $\bar{R}'$ which satisfies:

$$\bar{R}' \oplus \eta(M) = \bar{R}_f \oplus \eta(M) \qquad (3.2.13)$$

is invariant (mod $\eta(M)$) with respect to $A-DC$ if $\bar{R}_f$ is invariant with respect to $A-DC$. As an example, it is possible to choose $\bar{R}_f'$as defined by (3.2.9) and (3.2.10). The importance of (3.2.13) is that it implies that it is not necessary to determine $\bar{R}_f$ in order to solve the failure detection problem. The problem is completely characterized by any solution of the equation. There are ν_f assignable eigenvalues associated with that $\bar{R}'$.

If (A, C) is observable, $\bar{R}_f$ is isomorphic to the set $\{\xi_f\}$ of vectors which are detection equivalent to f. If (A, C) is not observable, a more general means for characterizing $\{\xi_f\}$ is needed. Clearly every $\xi_f \in \bar{R}'$ for all $\bar{R}'$ which satisfy (3.2.13) is detection equivalent to f. This suggests that $\{\xi_f\}$ can be specified in terms of the factor space of $\bar{R}_f$ with respect to $\eta(M)$. The factor space is uniquely specified by any solution of (3.2.13).

In order to state these arguments rigorously it is necessary to adopt a modified definition of detectability.

Definition 3.2.1 The failure associated with the event vector f is quasi-detectable if and only if there exists a D such that:

i) The output generated by f n(t) maintains a fixed direction in the output space, and

ii) the $n-\nu_o$ eigenvalues associated with the observable dynamics of (A-DC, C) can be specified almost arbitrarily by the proper choice of the detection filter.

The Detection Theorem can be interpreted in terms of this definition. For ease of comparison with the Detection Theorem, the proof which follows is based upon some $\bar{\mathcal{R}}_f$ which satisfies Definition 3.1.2. The proof is very similar to that of Theorem 3.2.1 and is only sketched here.

Theorem 3.2.1′ (Detection Theorem for Unobservable Systems) The failure associated with the event vector f is quasi-detectable if and only if $\bar{\mathcal{R}}_f$ is observable.

Proof

Necessity If $\bar{\mathcal{R}}_f$ is not observable, $\bar{\mathcal{R}}_f \subset \eta(M)$ and condition i) of Definition 3.1.2 is violated.

Sufficiency Let $\mathcal{S}^n$ be partitioned:

$$\mathcal{S}^n = \bar{\mathcal{R}}_f \oplus \bar{\mathcal{R}}_f^c \oplus \eta(M)$$

As was done in the proof of Theorem 3.2.1, choose $\bar{\mathcal{R}}_f^c$ such that

$$C\,\bar{\mathcal{R}}_f^c = \mathcal{V}_f^c$$

and $d(\mathcal{V}_f^c) = m - 1$. The proof is identical to that of Theorem 3.2.1 except that to define A-DC uniquely the additional operators must be given:

v) $A_o: \eta(M) \to \eta(M)$

vi) $A_{o2}: \bar{\mathcal{R}}_f^c \to \eta(M)$

vii) $\phi: \bar{\mathcal{R}}_f \to \eta(M)$

viii) $\phi: \eta(M) \to (\bar{\mathcal{R}}_f \oplus \bar{\mathcal{R}}_f^c)$

Of these operators, only A_{o2} can be altered by choice of D_2 and this does not alter the eigenvalues associated with A_o.

It remains only to specify D_2 in terms of the desired operators i) - iv) as in Theorem 3.2.1.

Δ

Implicit in the proof of Theorem 3.2.1′ is a generalization of Lemma 3.2.3 to the effect that (A-DC, C) is observable (mod $\bar{\mathcal{R}}_f \oplus \eta(M)$) Corollary 1 reflects this fact.

Corollary 1 Let D be chosen such that $\bar{\mathcal{R}}_f$ is invariant. No vector $\rho \notin (\bar{\mathcal{R}}_f \oplus \eta(M))$ can produce an output for (A-DC, C) which is unidirectional along Cf.

Proof

The proof is a generalization of that of Corollary 1 of Theorem 3.2.1 to unobservable systems.

Δ

Since every vector $\xi \notin \eta(M)$ affects the system output, it is apparent that every vector $\xi_f \in (\bar{\mathcal{R}}_f \oplus \eta(M))$ either produces output which is unidirectional along Cf or it produces no output at all. From the previous comment and Corollary 1, it follows that the set of event vectors detection equivalent to f is related to the factor space of $\bar{\mathcal{R}}_f$ with respect to $\eta(M)$.

Corollary 2 Let $\bar{\mathcal{R}}_f^*$ be the factor space of $\bar{\mathcal{R}}_f$ with respect to $\eta(M)$. The set $\{\xi_f\}$ of vectors which are detection equivalent to f is the set of vectors which are elements of some equivalence class in $\bar{\mathcal{R}}_f^*$ other than ϕ^*.

Proof

The statement is a formal characterization of all vectors $\xi_f \in (\eta(M) \oplus \bar{\mathcal{R}}_f)$ which satisfy $\xi_f \notin \eta(M)$. By Corollary 1 and the invariance of $\eta(M)$, these are the vectors which are detection equivalent to f.

Δ

In summary, if (A, C) is not observable, the failure detection problem is altered in three ways:

i) ν_o of the eigenvalues of A-DC become unassignable,

ii) the detection space is not unique and its function may be satisfied by any solution of (3.2.13), and

iii) the set of events detection equivalent to f is specified by the factor space $\bar{\mathcal{R}}_f^*$ of $\bar{\mathcal{R}}_f$ with respect to $\eta(M)$.

Just as the detection problem for which (A, C) is observable is completely characterized by the detection space $\bar{\mathcal{R}}_f$; when (A, C) is not observable, the problem is completely specified by the ν_o unassignable eigenvalues and the factor space $\bar{\mathcal{R}}_f^*$. The ambiguity of condition ii) above is removed if it is possible to determine the factor space directly. Alternatively, the relevant property of (3.2.13) is that any solution of the equation uniquely specifies $\bar{\mathcal{R}}_f^*$. The choice of whether to accept condition ii) or redefine the detection spaces of this chapter as factor spaces is one of personal taste. Either approach makes the subsequent results of this chapter readily generalizable to the unobservable case.

If a system is unobservable, the system model can be made observable by eliminating the unobservable dynamics from the reference model upon which the failure detection system is based. The detection space $\tilde{\bar{\mathcal{R}}}$ in the reduced $\tilde{n}$-dimensional model can then be determined simply. This is tantamount to dealing directly with the factor space $\bar{\mathcal{R}}_f^*$ in the n-dimensional model. All of the algorithms of this chapter can be applied to the $\tilde{n}$-dimensional model and then extrapolated to the n-dimensional model. It is assumed in Sections 3.2 - 3.6 that such a procedure is followed.

3.2.3 Algorithms

In order to utilize detection spaces as a practical procedure for determining the operator D, it is necessary to define a generating algorithm for the detection spaces. Two such algorithms are introduced here. The first demonstrates the

fundamental relationship of detection spaces to the failure detection problem by applying Definition 3.1.2 to determine $\bar{\mathcal{R}}_f$ recursively. It is included primarily to emphasize this relationship. The second algorithm is more circuitous but in many respects simpler to implement. It necessitates the arbitrary choice of a D which makes $\bar{\mathcal{R}}_f$ invariant and suggests one such choice. An example of the application of the algorithm is given in Chapter 4. The algorithms assume (A, C) is observable. If it is not, both converge to $\eta(M) \oplus \bar{\mathcal{R}}_f$ (although a constraint on the eigenvalues associated with $\eta(M)$ must be made for the first algorithm to converge). The discussion in the preceding subsection guarantees that this is sufficient to fully characterize the detection problem.

The first algorithm determines $\bar{\mathcal{R}}_f$ through a recursive solution for $\mathcal{R}_f$ and then condition i) of Definition 3.1.2:

$$\bar{\mathcal{R}}_f = \mathcal{R}_f \oplus f \tag{3.2.14}$$

Although the algorithm is expressed in terms of A^{-1}, it can be interpreted in terms of A. The starting point for the recursion is the null space of C:

$$\mathcal{R}^{(0)} = \eta(C) \tag{3.2.15}$$

The first iteration determines $\mathcal{R}^{(1)}$ by eliminating all vectors from $\mathcal{R}^{(0)}$ whose image under A is not contained in $\mathcal{R}^{(0)} \oplus f$. The recursion terminates when $\mathcal{R}^{(j)}$ is found for which condition iv) of Definition 3.1.2 is satisfied:

$$A\mathcal{R}^{(j)} \subset \mathcal{R}^{(j)} \oplus f \tag{3.2.16}$$

where it is necessary that $j \leq n-m$. The structure of the algorithm guarantees that $\mathcal{R}^{(j)}$ is the largest subspace of $\eta(C)$ which satisfies (3.2.16); hence, $\mathcal{R}^{(j)}$ is $\mathcal{R}_f$. A variation upon the algorithm can be implemented even if A^{-1} does not exist,

but the simplicity of the recursion is masked by the necessity of having to define inverse operators on restricted subspaces of $\mathcal{S}^n$.

Lemma 3.2.5 Assume (A, C) is observable and A^{-1} exists. The detection space $\bar{\mathcal{R}}_f$ is given by:

$$\bar{\mathcal{R}}_f = \mathcal{R}_f \oplus f$$

where:

$$\mathcal{R}_f = \mathcal{R}^{(n-m)}$$

and $\mathcal{R}^{(n-m)}$ is defined recursively from:

$$\mathcal{R}^{(0)} = \eta(C)$$

$$\mathcal{R}^{(i)} = \mathcal{R}^{(i-1)} \cap A^{-1}(\mathcal{R}^{(i-1)} \oplus f) \quad i = 1 \ldots n-m$$

Proof

The sequence $\mathcal{R}^{(0)}, \mathcal{R}^{(1)} \ldots \mathcal{R}^{(n-m)}$ is a decreasing nested sequence since $\mathcal{R}^{(i)} \subset \mathcal{R}^{(i-1)}$. This implies $d(\mathcal{R}^{(i)}) \leq d(\mathcal{R}^{(i-1)})$. If $d(\mathcal{R}^{(i)})$ equals $d(\mathcal{R}^{(i-1)})$, then it follows that $\mathcal{R}^{(i)}$ and $\mathcal{R}^{(i-1)}$ are the same subspace and the sequence becomes repetitive. Because $\eta(C)$ is (n-m)-dimensional, the recursion therefore converges in n-m steps or less and:

$$\mathcal{R}^{(n-m)} = \mathcal{R}^{(n-m)} \cap A^{-1}(\mathcal{R}^{(n-m)} \oplus f)$$

$$\text{i)} \qquad \subset A^{-1}(\mathcal{R}^{(n-m)} \oplus f)$$

It can be shown inductively that $\mathcal{R}_f \subset \mathcal{R}^{(n-m)}$. Clearly $\mathcal{R}_f \subset \mathcal{R}^{(0)}$. Assume $\mathcal{R}_f \subset \mathcal{R}^{(i-1)}$. Conditions iii) and iv) of Definition 3.1.2 states that $\mathcal{R}_f$ is the largest subspace of $\eta(C)$ for which:

$$A\,\mathcal{R}_f \subset \mathcal{R}_f \oplus f$$

Since A^{-1} is unique, this can be stated as:

ii) $$\mathcal{R}_f \subset A^{-1}(\mathcal{R}_f \oplus f)$$

and then by:

$$\mathcal{R}_f \subset \mathcal{R}^{(i-1)} \cap A^{-1}(\mathcal{R}^{(i-1)} \oplus f)$$

The expression on the right is $\mathcal{R}^{(i)}$ which yields $\mathcal{R}_f \subset \mathcal{R}^{(i)}$. Therefore by induction, $\mathcal{R}_f \subset \mathcal{R}^{(n-m)}$. Conditions i) and ii) above are identical relationships. Since $\mathcal{R}_f$ is the largest subspace of $\eta(C)$ to satisfy ii) and $\mathcal{R}_f \subset \mathcal{R}^{(n-m)}$, it follows that:

$$\mathcal{R}_f = \mathcal{R}^{(n-m)}$$

Δ

Lemma 3.2.5 is useful in that it demonstrates a straightforward and analytically appealing procedure for determining $\bar{\mathcal{R}}_f$ directly from Definition 3.1.2. Unfortunately, implementation of the algorithm can be computationally prohibitive. No direct algorithms exist for finding the intersection of two vector spaces. The standard procedure involves finding a series of orthogonal complements. The intersection for the lemma becomes:

$$\mathcal{R}^{(i-1)} \cap A^{-1}(\mathcal{R}^{(i-1)} \oplus f) = [\mathcal{R}^{(i-1)^\perp} \cup [A^{-1}(\mathcal{R}^{(i-1)} \oplus f)]^\perp]^\perp \tag{3.2.17}$$

A procedure for determining orthogonal complements is presented in Appendix A.1. The union of two subspaces can be found easily from their bases.

A computationally simpler procedure for determining $\bar{\mathcal{R}}_f$ can be based upon Lemma 3.2.3. The algorithm presented below is based upon the arbitrary choice of an operator D which makes $\bar{\mathcal{R}}_f$ invariant for the system (A-DC, C). Once this is accomplished it is a simple matter to determine $\bar{\mathcal{R}}_f$

from the properties of the new system. Lemma 3.2.6 delineates the fundamental requirements for the algorithm and Lemma 3.2.6′ provides a simple construction. The algorithms determine $\bar{\mathcal{R}}_f$ as the null space of a matrix. A straight-forward algorithm for determining this null space is described in Appendix A.2.

Lemma 3.2.6 Let D_f satisfy Lemma 3.2.2 and let C' be any operator such that:

i) $C'f = \phi$

ii) $C'\xi = C\xi \pmod{Cf}$

for all $\xi \in \mathcal{B}^n$. If (A, C) is observable, $\bar{\mathcal{R}}_f$ is the unobservable space of the system $(A-D_fC, C')$.

Proof

Let the unobservable space of $(A-D_fC, C')$ be denoted by $\eta(M'_D)$. Since $\bar{\mathcal{R}}_f$ is invariant with respect to $A-D_fC$ and:

$$C'\bar{\mathcal{R}}_f = \phi$$

it follows that $\bar{\mathcal{R}}_f \subset \eta(M'_D)$. By Corollary 1 of Theorem 3.2.1 and condition ii) above, it is also obvious that $\eta(M'_D) \subset \bar{\mathcal{R}}_f$. Δ

In matrix notation, M'_D is given by:

$$M'_D = \begin{bmatrix} C' \\ C'(A-D_fC) \\ \vdots \\ C'(A-D_fC)^{n-1} \end{bmatrix} \qquad (3.2.18)$$

It remains only to determine matrices D_f and C' which satisfy the lemma. A very reasonable choice is made below. These are the matrices used by Beard in his constructive derivation of Theorem 3.2.1.

Lemma 3.2.6′ For a given basis, let:

$$D_f = Af\left[(Cf)^T Cf\right]^{-1}(Cf)^T$$

and:

$$C' = \left[E_m - Cf\left[(Cf)^T Cf\right]^{-1}(Cf)^T\right]C$$

where E_m is the identity matrix, Then:

$$\bar{\mathcal{R}}_f = \eta(M_D')$$

Proof

Trivial application of Lemma 3.2.6.

Δ

Lemma 3.2.6′ is gratifying in that it provides a straightforward algorithm for finding detection spaces for event vectors: The detection space is $\eta(M_D')$. There is nothing special about the choice of D_f and C' in Lemma 3.2.6′. For a particular application a more convenient choice may suggest itself. Any D_f and C' which satisfy Lemma 3.2.6 determine both g and $\bar{\mathcal{R}}_f$.[1] Lemma 3.2.6′ is included here only to demonstrate that even in the most complex of systems, there is a simple, constructive algorithm for determining detection spaces.

After $\bar{\mathcal{R}}_f$ has been determined, the remaining task is to choose a D which satisfies Definition 2.1.1 for f:

i) the failure associated with f must generate unidirectional output errors for (A-DC, C), and

ii) the eigenvalues of A-DC must equal an almost arbitrarily chosen set.

If the reference model is transformed to the base normal form developed in Section 4.4, eigenvalue assignment is trivial. The algorithm presented in the remainder of this section is of less practical significance, but it emphasizes the importance of invariant subspaces in the study of linear operators.

It is simplest to consider assignment of the eigenvalues associated with $\bar{\mathcal{R}}_f$ first. The detection generator of $\bar{\mathcal{R}}_f$ is

[1] If (A, C) is not observable a g (mod $\eta(M)$) and some solution of (3.2.13) are determined.

useful in this context. The most convenient manner by which to assign the eigenvalues of $A-DC$ associated with $\bar{\mathcal{R}}_f$ is in terms of the desired minimal polynomial (characteristic equation) $\Psi_d(\cdot)$ for $\bar{\mathcal{R}}_f$ (Section 2.3.2):

$$\Psi_d(\lambda) = \lambda^{\nu_f} + p_{\nu_f}\lambda^{\nu_f-1} + p_{\nu_{f-1}}\lambda^{\nu_f-2} + \ldots p_2\lambda + p_1 \tag{3.2.19}$$

Lemma 3.2.7 Let g be the detection generator for $\bar{\mathcal{R}}_f$ and (A, C) be observable. If $\Psi_d(\cdot)$ is the ν_f-order desired minimal polynomial for $\bar{\mathcal{R}}_f$ with respect to $A-DC$, D must satisfy:

$$DCA^{\nu_f-1}g = \Psi_d(A)g$$

$\bar{\mathcal{R}}_f$ is invariant with respect to $A-DC$ for any such D.

Proof

If $\Psi_d(\cdot)$ is to be the minimal polynomial for $\bar{\mathcal{R}}_f$, then since $g \in \bar{\mathcal{R}}_f$, $\Psi_d(\cdot)$ must satisfy:

i) $\Psi_d(A-DC)\,g = (A-DC)^{\nu_f}g + p_{\nu_f}(A-DC)^{\nu_f-1}g + \ldots p_2(A-DC)g + p_1 g = \phi$

By Lemma 3.2.4, the vectors $g, Ag, \ldots A^{\nu_f-1}g$ form a basis for $\bar{\mathcal{R}}_f$. The lemma also states that:

ii) $$CA^k g = \phi \qquad k < \nu_f - 1$$

which implies that $g, (A-DC)g \ldots (A-DC)^{\nu_f-1}g$ are the same basis for $\bar{\mathcal{R}}_f$ for any choice of D. Equation i) is then equivalent to requiring $\bar{\mathcal{R}}_f$ to be a cyclic invariant subspace with respect to $A-DC$ (Section 2.3.2).

Because of ii) above, i) has only one non-zero term involving DC:

$$A^{\nu_f}g - DCA^{\nu_f-1}g + p_{\nu_f}A^{\nu_f-1}g + \ldots p_2 Ag + p_1 g = \phi$$

This can be reorganized in terms of $\Psi_d(A)g$:

$$\Psi_d(A)g - DCA^{\nu_f - 1} g = \phi$$

which is the desired result. Δ

In matrix terms, Lemma 3.2.7 does not specify a unique detection filter unless $\nu_f = n$.[1] Theorem 3.2.1 asserts that the remaining freedom is sufficient to assign the remaining $n - \nu_f$ eigenvalues. One approach for utilizing this freedom is to define additional event vectors $f_1 .. f_k$ to detect along with f. If each event vector has its own detection space and if its eigenvalues can be assigned independently of those associated with the other detection spaces, the net effect is to increase the capability of the detection system from monitoring one failure to monitoring k + 1 failures. This possibility is investigated in the next section. It is safe to assume that in general such an increase in capability is desirable and useful.

If f is the only event vector to be detected by A-DC, the partial observability of the system (A-DC, C') can be used to assign the remaining eigenvalues. (Any D which satisfies Lemma 3.2.7 also satisfies Lemma 3.2.6). Since $\bar{\mathcal{R}}_f \subset \eta(M'_D)$, the eigenvalues associated with $\bar{\mathcal{R}}_f$ cannot be altered by output feedback for (A-DC, C'). A number of algorithms for assigning the eigenvalues of the observable subsystem can be found in the literature (see Wonham and Morse (1970), Howze and Pearson (1970), Brasch and Pearson (1970), and Davison (1970)). The algorithm presented here is based upon the ideas presented in this section.

Lemma 3.2.8 Assume D has been chosen such that $\Psi_d(\cdot)$ is the minimal polynomial of $\bar{\mathcal{R}}_f$ for A-DC. Let $\bar{\mathcal{R}}'$ be a ν'-order

[1] The non-uniqueness of D is discussed in Sections 4.2.4 and 4.4.

detection space for some arbitrarily chosen event vector f' for $(A\text{-}DC, C')$, g' be the detection generator for $\bar{R}'$, and $\Psi'_d(\cdot)$ be its desired minimal polynomial (mod $\bar{R}_f$) for $A\text{-}DC\text{-}D'C'$. If D' satisfies:

$$D'C'(A\text{-}DC)^{\nu'-1}g' = \Psi'_d(A\text{-}DC)g'$$

then $\Psi_d(\cdot)$ is the minimal polynomial of $\bar{R}_f$ for $A\text{-}DC\text{-}D'C'$ and $\Psi'_d(\cdot)$ is the minimal polynomial of $\bar{R}'$ (mod $\bar{R}_f$).

Proof

Since Lemma 3.2.6 provides:

$$\eta(M'_D) = \bar{R}_f$$

it follows that:

$$(A\text{-}DC\text{-}D'C')\ \bar{R}_f = (A\text{-}DC)\bar{R}_f$$

and the minimal polynomial of $\bar{R}_f$ for $A\text{-}DC\text{-}D'C'$ is the same as its minimal polynomial for $A\text{-}DC$. A generalization of the proof of Lemma 3.2.7 shows that the minimal polynomial (mod $\bar{R}_f$) of $\bar{R}'$ is $\Psi'_d(\cdot)$.

Δ

Because $(A\text{-}DC, C')$ is unobservable, Lemma 3.2.8 makes $\bar{R}'$ invariant (mod $\bar{R}_f$) with respect to $A\text{-}DC\text{-}D'C'$. It does not necessarily insure that $\bar{R}'$ produces unidirectional output. Although the algorithm requires $d(C'\bar{R}') = 1$, it does not imply that $d(CR') = 1$.

3.3 Multiple Failure Detection

In a system with a large number of possible failures to monitor, it is useful to know the conditions under which more than one event can be detected by a single detection filter. The emphasis in this section is upon an abstraction of the

Detection Theorem to the case of r failures where $r \leq m$. Attention is restricted to the generation of sufficient conditions for a set of failures to be detected by a single failure detection system. It can be shown that these conditions are also necessary (Section 3.5), but the proof is based upon concepts not introduced until Section 3.4. Algorithms for testing mutual detectability are developed in this section, but their application is not discussed until Section 4.2.2.

The basic theorem on mutual detectability is Theorem 3.3.1. The theorem states two conditions which together are sufficient for a set of r failures to be mutually detectable. The first condition is that the output directions $Cf_1 \ldots Cf_r$ be linearly independent, where $f_1 \ldots f_r$ are the event vectors associated with the failures. The second condition involves the respective detection spaces $\bar{R}_1 \ldots \bar{R}_r$. If $Cf_1 \ldots Cf_r$ are not linearly independent and none of the event vectors are detection equivalent, the failures cannot all be detected by a single failure detection system. (This statement is not proved until Section 3.5). Unfortunately, even when $Cf_1 \ldots Cf_r$ are linearly independent, the second condition of Theorem 3.3.1 need not be satisfied. The physical meaning behind the constraint on $\bar{R}_1 \ldots \bar{R}_r$ is motivated by an example in Section 3.3.2.

If a set of r failures are mutually detectable, the next step is to implement design algorithms for the failure detection system. If they are not mutually detectable but $Cf_1 \ldots Cf_r$ are linearly independent, several options are available. One of these is to violate Defintion 2.1.1 and allow some eigenvalues of A-DC to be unassignable. Another is to find a subset of the r failures which can be detected by a single failure detection system. A third option is to increase the dimension of the reference model in such a manner that all of the failures can be detected by a single failure detection system based upon the higher dimensional model. The first possibility is discussed in Section 4.2.2. The last two are studied in detail in Sections 3.3.3 and 3.3.4 respectively. Complete algorithms

for constructing mutually detectable sets of failures are developed in these sections.

Throughout this and subsequent sections, it is assumed that the reference model (A, C) is observable (the real system need not be observable). If (A, C) is not observable, the results are equally valid as a study of mutually quasi-detectable events (Section 3.2.2). It is also assumed that $Cf_i \neq \phi$ for $i = 1, \ldots r$. If this is not true for some f_i, that f_i is simply replaced by the first element of the sequence f_i, Af_i, $A^2 f_i \ldots$ which is not contained in $\mathcal{N}(C)$. The validity of this substitution was discussed in Section 3.1.

3.3.1 Sufficient Conditions for a Mutually Detectable Set

The culmination of this subsection is Theorem 3.3.1 on mutual detectability. The definition of mutual detectability is a generalization of Definition 2.1.1 to multiple failures:

Definition 3.3.1 Given the inhomogeneous error equations:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + f_i\, n_i(t) \quad i = 1 \ldots r$$

$$\epsilon'(t) = C\,\epsilon(t)$$

the failures associated with the event vectors $f_1 \ldots f_r$ are mutually detectable by a single failure detection system if:

i) The output generated by each of $f_1\, n_1(t) \ldots f_r\, n_r(t)$ maintains a fixed direction in the output space, and

ii) the eigenvalues of $A-DC$ can be specified almost arbitrarily by the proper choice of D.

Definition 3.3.1 does not require any particular relationship between the output directions $Cf_1 \ldots Cf_r$. Certain special cases can be examined immediately, however. Assume that the event vectors f_i and f_j have the same output direction:

$$Cf_i = Cf_j$$

and f_j is a vector in the detection space $\bar{\mathcal{R}}_i$ for f_i. By Corollary 2 of the Detection Theorem, the failures associated with f_i and f_j are detection equivalent. If one of the failures is constrained to generate unidirectional output estimation errors along Cf_i, the second failure is also. A single failure detection system can always be designed to detect both failures, but the decision as to which of the failures occurs cannot be based upon output error direction alone. On the other hand, if Cf_j and Cf_i are equal but f_j is not a vector in $\bar{\mathcal{R}}_i$, Corollary 1 of the Detection Theorem asserts that a failure detection system which detects the failure associated with f_i cannot simultaneously detect the one associated with f_j. The output error for the second failure cannot be constrained to a single direction.

The results of the previous section are directly generalizable to a discussion on mutual detectability only in these two instances. Additional results are developed in this section for the case where the output directions $Cf_1 \ldots Cf_r$ are linearly independent. Except for the possible inclusion of detection equivalent failures, this limits attention to a maximum of m possible failures. A consequence of Theorem 3.5.1 is that m is an upper limit unless a modified control over the eigenvalues of $A\text{-}DC$ is accepted. Because of the importance of linearly independent output directions, the following definition is made:

Definition 3.3.2 Let $\mathcal{I} \subset \mathcal{E}^n$ be the subspace spanned by the event vectors $f_1 \ldots f_r$. The vectors $f_1 \ldots f_r$ are output separable if:

$$d(C\mathcal{I}) = r$$

It is a simple matter to show that output separability is sufficient to guarantee that a D can be found for which the failures associated with $f_1 \ldots f_r$ produce unidirectional output errors. If the dimension of $\bar{\mathcal{R}}_i$ is denoted by ν_i:

$$d(\bar{\mathcal{R}}_i) = \nu_i \qquad i = 1 \ldots r \qquad (3.3.1)$$

a generalization of Lemma 3.2.7 using Lemma 3.2.1 shows that in addition a total of ν'_J eigenvalues of A-DC can be almost arbitrarily assigned by the choice of D where:

$$\nu'_J = \sum_{i=1,r} \nu_i \qquad (3.3.2)$$

These facts are proved in Lemma 3.3.1. The proof could be based upon a partition of A-DC as in the Detection Theorem, however, an approach based upon Lemma 3.2.7 is used because of its simplicity.

Lemma 3.3.1 If $f_1 \ldots f_r$ are output separable, there exists a D such that the failures associated with $f_1 \ldots f_r$ generate unidirectional output errors for the failure detection system (A-DC, C) and the ν'_J eigenvalues of A-DC associated with $\bar{R}_1 \ldots \bar{R}_r$ are the elements of a desired set Λ.

Proof

By virtue of Lemma 3.2.2, unidirectionality is assured if D satisfies:

i) $(A\text{-}DC)\, f_i \in \bar{R}_i \quad i = 1 \ldots r$

A particular D which satisfies these r equations and which assigns the eigenvalues of A-DC can be found using Lemma 3.2.7. Assume $\Psi_1(\cdot) \ldots \Psi_r(\cdot)$ are the desired minimal polynomials of A-DC for the respective detection spaces $\bar{R}_1 \ldots \bar{R}_r$. The roots of the polynomials must be the elements of Λ. If $g_1 \ldots g_r$ are the detection generators for $\bar{R}_1 \ldots \bar{R}_r$, Lemma 3.2.7 requires that the r equations be satisfied:

ii) $DCA^{\nu_i - 1} g_i = \Psi_i(A)\, g_i \quad i = 1 \ldots r$

Define the operators Y and X :

$$Y = [\, CA^{\nu_1 - 1} g_1 \ldots CA^{\nu_r - 1} g_r \,]$$

$$X = [\, \psi_1(A) g_1 \ldots \psi_r(A) g_r \,]$$

Since Lemma 3.2.4 states:

$$CA^{\nu_i - 1} g_i = C f_i \qquad i = 1 \ldots r$$

and $f_1 \ldots f_r$ are output separable, it follows that the vectors $CA^{\nu_1 - 1} g_1 \ldots CA^{\nu_r - 1} g_r$ are linearly independent and therefore $\eta(Y) = \phi$. Lemma 3.2.1 then guarantees that a D exists which satisfies ii). By Lemma 3.2.7, $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant for this D and i) is satisfied also.

Δ

If $\nu'_{\mathcal{I}}$ equals n, Lemma 3.3.1 implies that the r failures associated with $f_1 \ldots f_r$ are mutually detectable. The failures can also be mutually detectable if $\nu'_{\mathcal{I}} < n$ as long as sufficient freedom in the choice of D exists to assign the remaining $n - \nu'_{\mathcal{I}}$ eigenvalues of $A - DC$. In order to characterize the eigenvalue assignability, a generalization of the concept of a detection space is necessary.

Definition 3.3.3 Let $\mathcal{I} \subset \mathcal{S}^n$ be spanned by $f_1 \ldots f_r$ and let $f_1 \ldots f_r$ be output separable. The detection space for $\mathcal{I}$ is the subspace $\bar{\mathcal{R}}_{\mathcal{I}} \subset \mathcal{S}^n$ such that:

i) $\bar{\mathcal{R}}_{\mathcal{I}} = \mathcal{R}_{\mathcal{I}} \oplus \mathcal{I}$

where $\mathcal{R}_{\mathcal{I}}$ is the largest subspace which satisfies:

ii) $\eta(M) \cap \mathcal{R}_{\mathcal{I}} = \phi$

iii) $\mathcal{R}_{\mathcal{I}} \subset \eta(C)$

iv) $A\,\mathcal{R}_{\mathcal{I}} \subset \bar{\mathcal{R}}_{\mathcal{I}}$

The conditions of Definition 3.3.3 are precisely those of Definition 3.1.2 with the subspace $\mathcal{J}$ substituted for the vector f. $\mathcal{R}_{\mathcal{J}}$ must be in the null space of C and it must be observable. The image of $\mathcal{R}_{\mathcal{J}}$ under A must be contained in $\bar{\mathcal{R}}_{\mathcal{J}}$. Recall that if D is chosen such that $\bar{\mathcal{R}}_f$ is invariant with respect to $A-DC$, every event vector in $\bar{\mathcal{R}}_f$ generates output errors which are unidirectional along Cf (Corollary 2 of the Detection Theorem). Similarly, if $D_{\mathcal{J}}$ is chosen such that $\bar{\mathcal{R}}_{\mathcal{J}}$ is invariant with respect to $A-D_{\mathcal{J}}C$, every event vector in $\bar{\mathcal{R}}_{\mathcal{J}}$ generates output errors which are restricted to the hyperplane $C\mathcal{J}$. The output errors may be unidirectional, but it is not necessary.

The fact that a $D_{\mathcal{J}}$ which makes $\bar{\mathcal{R}}_{\mathcal{J}}$ invariant exists can be proved in a manner similar to that used in Lemma 3.3.1. A different approach can be developed from a straightforward generalization of Lemma 3.2.6 by use of Lemma 3.2.1. The result is analogous to the construction in Lemma 3.2.6′ and is presented in Lemma 3.3.2. The proof is a direct consequence of the generalization of Lemma 3.2.6 and is omitted.

Lemma 3.3.2 For a given basis let F be the matrix:

$$F = [f_1 \quad f_2 \quad \dots \quad f_r]$$

and assume rkF = rkCF = r. Let:

$$D_{\mathcal{J}} = AF[(CF)^T CF]^{-1}(CF)^T$$

and:

$$C'_{\mathcal{J}} = [E_m - CF[(CF)^T CF]^{-1}(CF)^T]\, C$$

where E_m is the identity matrix. If (A, C) is observable, $\bar{\mathcal{R}}_{\mathcal{J}}$ is the unobservable space of $(A-D_{\mathcal{J}}C, C'_{\mathcal{J}})$

$$\bar{\mathcal{R}}_{\mathcal{J}} = \eta(M'_D)$$

Also, $\bar{\mathcal{R}}_{\mathcal{J}}$ is invariant with respect to $A-D_{\mathcal{J}}C$.

The detection space $\bar{R}_{\mathcal{J}}$ is more that an arbitrary generalization of $\bar{R}_f$. It is a detection space which must be considered along with $\bar{R}_1 \ldots \bar{R}_r$ when mutual detectability is considered. The relationship of $\bar{R}_{\mathcal{J}}$ to $\bar{R}_1 \ldots \bar{R}_r$ is the prime concern of this section. Before considering this relationship, however, it is necessary to determine the relationship of $\bar{R}_1 \ldots \bar{R}_r$ to each other. The following lemma demonstrates that the union $\bar{R}_1 \cup \bar{R}_2 \cup \ldots \bar{R}_r$ can be properly expressed as the direct sum $\bar{R}_1 \oplus \bar{R}_2 \oplus \ldots \bar{R}_r$. This property is true only if $f_1 \ldots f_r$ are output separable.

Lemma 3.3.3 If $f_1 \ldots f_r$ are output separable event vectors, then $\bar{R}_1 \ldots \bar{R}_r$ are mutually distinct, i.e.:

$$d(\bar{R}_1 \cup \ldots \bar{R}_r) = \sum_{i=1,r} d(\bar{R}_i)$$

Proof

Assume the contrary. This requires that for some $1 \leq k \leq r$ there exist $\xi_k \in \bar{R}_k$ such that:

$$\xi_k = \sum_{i=1,r} \alpha_i \rho_i$$

where $\rho_i \in \bar{R}_i$ and the α_i can be chosen such that some $\alpha_i \neq 0$ for $i \neq k$. For this set of α_i, define:

$$\xi_k' = \xi_k - \alpha_k \rho_k = \sum_{i \neq k} \alpha_i \rho_i$$

Clearly since $\rho_k \in \bar{R}_k$, it is still true that $\xi_k' \in \bar{R}_k$. However, ξ_k' can also be expressed as an element of the union of the remaining subspaces:

i) $\xi_k' \in \bar{R}_1 \cup \ldots \bar{R}_{k-1} \cup \bar{R}_{k+1} \cup \ldots \bar{R}_r$

Let μ_k be the smallest integer such that $CA^{\mu_k - 1} \xi_k' \neq \emptyset$. Since $\xi_k' \in \bar{R}_k$, condition iv) of Definition 3.1.2 implies:

$$CA^{\mu_k - 1} \xi_k' = \beta\, Cf_k$$

where β is a constant. But from i) above it also follows that there exists some subset I_J of the integers 1, 2 ... k-1, k+1... r such that:

$$CA^{\mu_k - 1} \xi_k' = \sum_{j \in I_J} \beta_j\, Cf_j$$

These two conditions are consistent only if Cf_k is a linear combination of the vectors Cf_j or if $\alpha_i = 0$ for all $i \neq k$. Since $f_1 \ldots f_r$ are output separable, $Cf_1 \ldots Cf_r$ are linearly independent and therefore $\alpha_i = 0$ for all $i \neq k$. Δ

The simplest relationship to prove between $\bar{\mathcal{R}}_J$ and $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ is that each of $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ is a subspace of $\bar{\mathcal{R}}_J$. This result requires little more than a manipulation of Definitions 3.1.2 and 3.3.3 to obtain. It is stated as Lemma 3.3.4 because of its importance and because of the contrast to Theorem 3.3.1.

Lemma 3.3.4 If $f_1 \ldots f_r$ are output separable, the individual detection spaces are subspaces of $\bar{\mathcal{R}}_J$:

$$\bar{\mathcal{R}}_1 \oplus \bar{\mathcal{R}}_2 \oplus \ldots \bar{\mathcal{R}}_r \subset \bar{\mathcal{R}}_J$$

Proof

The direct sum is a consequence of Lemma 3.3.3. From Definition 3.1.2, $\mathcal{R}_1 \ldots \mathcal{R}_r$ are observable and in the null space of C:

ii) $\eta(M) \cap \mathcal{R}_i = \phi \qquad i = 1 \ldots r$

iii) $\mathcal{R}_i \subset \eta(C)$

Their images under A are contained in the respective detection spaces:

iv) $A\mathcal{R}_i \subset \mathcal{R}_i \oplus f \qquad i = 1 \ldots r$

But since $f_1 \ldots f_r$ are contained in $\mathcal{J}$, the last statement can be replaced by:

iv)' $A\mathcal{R}_i \subset \mathcal{R}_i \oplus \mathcal{J} \qquad i = 1 \ldots r$

By Definition 3.3.3, $\mathcal{R}_{\mathcal{J}}$ is the largest subspace to satisfy ii), iii), and iv)' above and it necessarily follows that $\mathcal{R}_1 \subset \mathcal{R}_{\mathcal{J}}$, $i = 1 \ldots r$. Since the definitions require:

$$\bar{\mathcal{R}}_i = \mathcal{R}_i \oplus f_i \qquad i = 1 \ldots r$$

$$\bar{\mathcal{R}}_{\mathcal{J}} = \mathcal{R}_{\mathcal{J}} \oplus \mathcal{J}$$

it follows that $\bar{\mathcal{R}}_i \subset \bar{\mathcal{R}}_{\mathcal{J}}$ and the lemma is proved.

Δ

The four lemmas of this section allow the problem of mutual detectability of output separable event vectors to be stated in a solvable format. If the dimension of $\bar{\mathcal{R}}_{\mathcal{J}}$ is denoted by $\nu_{\mathcal{J}}$:

$$d(\bar{\mathcal{R}}_{\mathcal{J}}) = \nu_{\mathcal{J}} \tag{3.3.3}$$

a consequence of Lemma 3.3.2 is that if $\bar{\mathcal{R}}_{\mathcal{J}}$ is invariant with respect to A-DC, there are $\nu_{\mathcal{J}}$ eigenvalues of A-DC associated with $\bar{\mathcal{R}}_{\mathcal{J}}$. It was proved in conjunction with Lemma 3.3.1 that the $\nu'_{\mathcal{J}}$ eigenvalues of A-DC associated with $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ can be assigned almost arbitrarily. Lemma 3.3.4 clearly requires that:

$$\nu_{\mathcal{J}} \geq \nu'_{\mathcal{J}} \tag{3.3.4}$$

The proof of Theorem 3.3.1 shows that any D which makes $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_f$ invariant makes $\bar{\mathcal{R}}_{\mathcal{J}}$ invariant. These statements suggest that it may be difficult to make $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ invariant and assign all eigenvalues of A-DC almost arbitrarily unless $\nu_{\mathcal{J}}$ equals $\nu'_{\mathcal{J}}$, i.e., unless:

$$C\bar{R}_{\mathcal{J}}^{c} = \mathcal{J}_{\mathcal{J}}^{c}$$

where $d(\mathcal{J}_{\mathcal{J}}^{c}) = m-r$. A construction similar to that employed in the proof of Theorem 3.2.1 shows that the $\nu_{\mathcal{J}}$ eigenvalues associated with $\bar{R}_{\mathcal{J}}$ and the $n-\nu_{\mathcal{J}}$ associated with $\bar{R}_{\mathcal{J}}^{c}$ can be assigned as two independent groups. If D is chosen such that $\bar{R}_1 \ldots \bar{R}_r$ are invariant as well as $\bar{R}_{\mathcal{J}}$, Lemma 3.3.1 demonstrates that only $\nu'_{\mathcal{J}}$ of the $\nu_{\mathcal{J}}$ eigenvalues associated with $\bar{R}_{\mathcal{J}}$ can be assigned. Therefore $\bar{R}_1 \ldots \bar{R}_r$ are mutually detectable if and only if:

$$\nu_{\mathcal{J}} = \nu'_{\mathcal{J}}$$

Δ

3.3.2 An Example

The analysis in Section 3.2.2 indicated that not all the eigenvalues of $A-DC$ can be assigned almost arbitrarily if (A, C) is unobservable. Theorem 3.3.1 indicates a second set of circumstances in which eigenvalue assignment cannot be complete. Before investigating this problem analytically, it is useful to indicate the type of application in which it occurs. The example presented in this subsection provides a degree of intuitive motivation for the problem as well as a summary of the previous subsections. The physical system described below is utilized repeatedly in later sections.

Consider the simple problem of a spinning cylinder in a viscous fluid. Since there is no restoring force, the differential equation describing the motion is:

$$\ddot{\varphi}(t) + \beta_r \dot{\varphi}(t) = \alpha_r T(t) \quad : \varphi(0), \dot{\varphi}(0) \text{ given} \qquad (3.3.6)$$

where $T(t)$ is the applied torque (Figure 3.1). Assume the rotational angle and the rotational rate of the cylinder are measured and that these measurements are incorporated into a first-order feedback law:

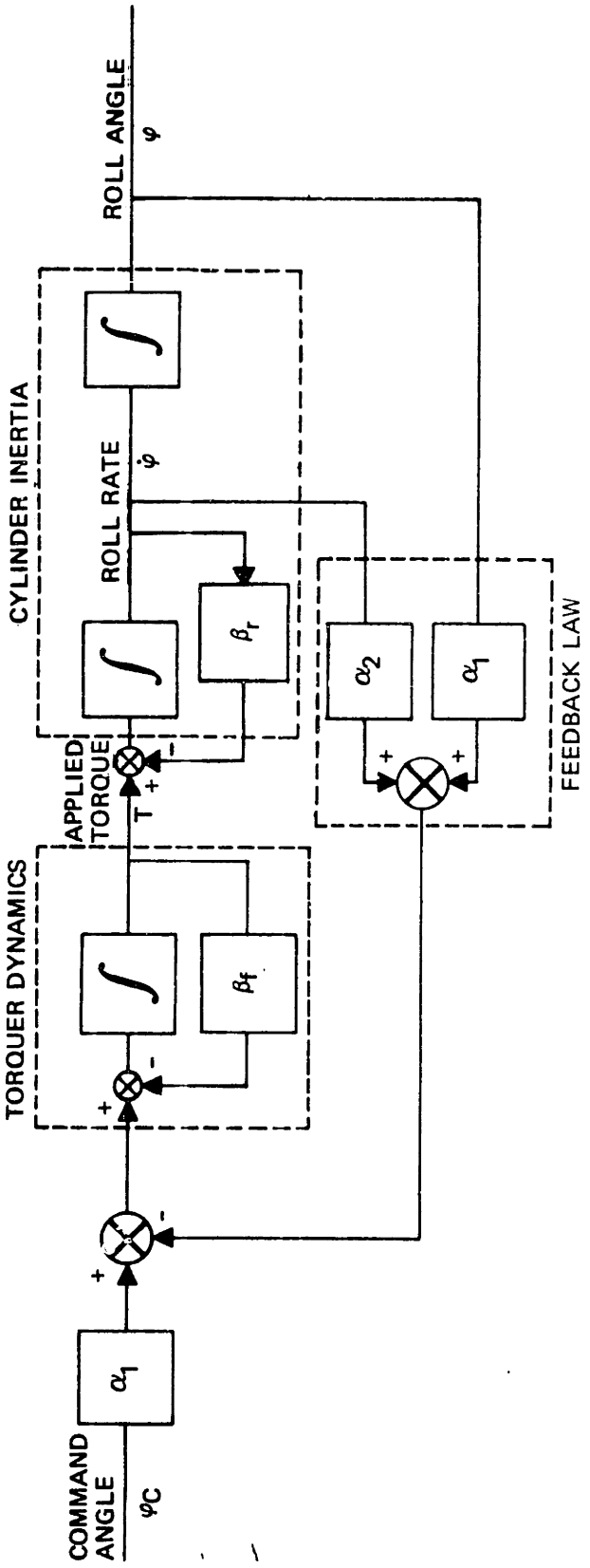


Figure 3.1 Block diagram for spinning cylinder example

$$\dot{T}(t) + \beta_f T(t) = -\alpha_1 \varphi_m(t) - \alpha_2 \dot{\varphi}_m(t) + \alpha_1 \varphi_c(t) \quad (3.3.7)$$

where $\varphi_c(t)$ is the commanded roll angle. The total applied torque T(t) is the sum of the feedback and the command torques. To simplify the analysis, let the total applied torque also be measured.

One possible state space model for the closed loop system is:

$$\dot{x}(t) = A_4 x(t) + B_4 u(t)$$

$$y(t) = C_4 x(t) \quad (3.3.8)$$

where:

$$C_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_4 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\beta_r & 0 & \alpha_r \\ 0 & -\beta_r & 0 & \alpha_r \\ -\alpha_1 & 0 & -\alpha_2 & -\beta_f \end{bmatrix} \quad B_4 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \alpha_1 \end{bmatrix}$$

(3.3.9)

The state vector is:

$$x(t) = \begin{bmatrix} \varphi_m(t) \\ \dot{\varphi}(t) \\ \dot{\varphi}_m(t) \\ T(t) \end{bmatrix} \quad (3.3.10)$$

The model (3.3.9) is chosen to be nonminimal because the true rotation rate, $\dot{\varphi}(t)$, and the measured rotation rate, $\dot{\varphi}_m(t)$, affect the closed loop system response in different manners. If the rotation rate sensor fails, the system model must reflect the fact that only the rotation rate feedback to the control law is affected; the angle feedback remains correct. Considerations

involved in choosing a reference model for which failures can be properly modeled are discussed at length in Section 4.1.

With (3.3.9), both the angle sensor failure and the rotation rate sensor failure can be modeled using the controller failure model (2.1.7):

$$\dot{x}(t) = A_4 x(t) + B_4 u(t) + f_1 n_1(t) \quad : x(0) \text{ given} \qquad (2.1.7)$$

$$y(t) = C_4 x(t)$$

where f_1 for the respective failures is:

$$f_1 = f_\varphi = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = e_{41}$$

$$f_1 = f_{\dot{\varphi}} = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = e_{43} \qquad (3.3.11)$$

The fact that the sensor failures can be modeled using the controller failure model rather than the more cumbersome sensor failure model (2.1.8) is a consequence of using the closed loop reference model rather than an open loop reference model. This property is also discussed in Section 4.1.

In this example, the detection spaces $\bar{\mathcal{R}}_\varphi$ and $\bar{\mathcal{R}}_{\dot{\varphi}}$ are one-dimensional and contain only their respective event vectors, as can be verified by application of Lemma 3.3.6'. If $\mathcal{J}$ is spanned by f_φ, $f_{\dot{\varphi}}$, it can be shown by application of Lemma 3.3.2 that $\bar{\mathcal{R}}_\mathcal{J}$ is spanned by the basis $\{e_{41}, e_{42}, e_{43}\}$. Therefore $\nu_\mathcal{J} > \nu'_\mathcal{J}$ and f_φ, $f_{\dot{\varphi}}$ are not mutually detectable. If D_4 is chosen such that f_φ and $f_{\dot{\varphi}}$ both generate unidirectional output, one eigenvalue of $A_4 - D_4 C_4$ is unassignable. This eigenvalue is associated with e_{42} and is equal to $-\beta_r$ for all such D_4.

The vectors f_φ, $f_{\dot\varphi}$ are not mutually detectable because of the presence of $\dot\varphi(t)$ in the state model. Physically, $\dot\varphi(t)$ is not in either detection space because it appears in the derivative of both $\varphi_m(t)$ and $\dot\varphi_m(t)$. In terms of event vectors and operator notation, since:

$$e_{42} \in \eta(C_4) \tag{3.3.12}$$

for e_{42} to be in either $\bar{R}_\varphi$ or $\bar{R}_{\dot\varphi}$ requires one of:

$$\begin{aligned} A_4\, e_{42} &\in \bar{R}_\varphi \\ A_4\, e_{42} &\in \bar{R}_{\dot\varphi} \end{aligned} \tag{3.3.13}$$

to be true (condition iv) of Definition 3.1.2). Computation reveals that $A_4\, e_{42}$ does not satisfy (3.3.13); however, it is contained in the union of $\bar{R}_\varphi$ and $\bar{R}_{\dot\varphi}$:

$$A_4\, e_{42} \in \bar{R}_\varphi \oplus \bar{R}_{\dot\varphi} \tag{3.3.14}$$

Lemma 3.3.4 and Definition 3.3.3 then imply that e_{42} is contained in $\bar{R}_\mathcal{J}$. Hence Theorem 3.3.1 states that $f_\varphi, f_{\dot\varphi}$ are not mutually detectable. The problem could be circumvented if $\dot\varphi(t)$ were removed from the reference model; but then it would not be possible to model the failures properly.

The event vector e_{42} is referred to as the excess subspace of $\bar{R}_\mathcal{J}$ -- it is mapped by A into the union of the component detection spaces $\bar{R}_\varphi$, $R_{\dot\varphi}$ but is not contained in either of them. In the following two subsections the properties of the excess subspace are studied further.

3.3.3 Mutually Detectable Subset of an Undetectable Set

Theorem 3.3.1 states that the output separable event vectors $f_1 \ldots f_r$ are not mutually detectable by a single failure

detection system if the union of their detection spaces $\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$ is not equal to $\bar{\mathcal{R}}_{\mathcal{J}}$. Since Lemma 3.3.4 requires that the union $\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$ be a subspace of $\bar{\mathcal{R}}_{\mathcal{J}}$, the theorem recognizes that $\bar{\mathcal{R}}_{\mathcal{J}}$ may for some reason be larger than $\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$. The "excess space" of $\bar{\mathcal{R}}_{\mathcal{J}}$ results in a set of unassignable eigenvalues of A-DC if D is chosen such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant with respect to A-DC. The initial effort in this subsection is to develop the structure of $\bar{\mathcal{R}}_{\mathcal{J}}$ more completely than Lemma 3.3.4 in order to establish the reason for the unassignable eigenvalues. With this accomplished, algorithms are devised for choosing a subset of $f_1 \ldots f_r$ which are mutually detectable.

When $f_1 \ldots f_r$ are not mutually detectable, the results of Section 3.3.1 do not provide a representation of $\bar{\mathcal{R}}_{\mathcal{J}}$ in terms of $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$. Definition 3.3.4 supplies such a representation by formally defining the excess subspace $\mathcal{R}_0$ of $\bar{\mathcal{R}}_{\mathcal{J}}$.

Definition 3.3.4 Assume $f_1 \ldots f_r$ are not mutually detectable. The excess subspace of $\bar{\mathcal{R}}_{\mathcal{J}}$ is any subspace $\mathcal{R}_0 \subset \eta(C)$ which satisfies:

$$\bar{\mathcal{R}}_{\mathcal{J}} = \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r \oplus \mathcal{R}_0$$

The stipulation that $\mathcal{R}_0$ be in the null space of C is not necessary to the definition of an excess subspace. It is added to simplify the proofs in this subsection. The fact that some $\mathcal{R}_0$ exists can be readily demonstrated. Definition 3.3.3 requires that:

$$d(C\bar{\mathcal{R}}_{\mathcal{J}}) = d(C^{\mathcal{J}}) = r \qquad (3.3.15)$$

Similarly, the fact that $f_1 \ldots f_r$ are output separable implies:

$$d(C(\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r)) = d(C\bar{\mathcal{R}}_1 \oplus \ldots C\bar{\mathcal{R}}_r) = d(Cf_1 \oplus \ldots Cf_r) = r \qquad (3.3.16)$$

Equations (3.3.15) and (3.3.16) together with Lemma 3.3.4 require that there must exist some $\mathcal{R}_0 \subset \eta(C)$ such that:

$$\bar{\mathcal{R}}_{\mathcal{J}} = \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r \oplus \mathcal{R}_0 \qquad (3.3.17)$$

The excess subspace of $\bar{\mathcal{R}}_{\mathcal{J}}$ is not unique. Any subspace which satisfies (3.3.17) and is in $\eta(C)$ is an excess subspace for $\bar{\mathcal{R}}_{\mathcal{J}}$. The majority of the properties of an excess subspace can be determined without choosing a specific $\mathcal{R}_0$. For instance, the dimension of $\mathcal{R}_0$ is defined by $\nu_{\mathcal{J}}$ and $\nu'_{\mathcal{J}}$:

$$d(\mathcal{R}_0) = \nu_0 = \nu_{\mathcal{J}} - \nu'_{\mathcal{J}} \qquad (3.3.18)$$

It is shown in Lemma 3.3.5 that if D is chosen such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant with respect to $A\text{-}DC$, the ν_0 eigenvalues of $A\text{-}DC$ associated with $\mathcal{R}_0$ are the same for any choice of D or $\mathcal{R}_0$. These are the eigenvalues which prevent $f_1 \ldots f_r$ from being mutually detectable. The lemma does not imply that the unassignable eigenvalues are a subset of the original eigenvalues of A and in general they are not. They are determined by the system and the detection spaces.

Lemma 3.3.5 Let $\mathcal{R}_0$ be an excess subspace for $\bar{\mathcal{R}}_{\mathcal{J}}$ and $f_1 \ldots f_r$ be output separable. If D is chosen such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant with respect to $A\text{-}DC$, the ν_0 eigenvalues of $A\text{-}DC$ associated with $\mathcal{R}_0$ are independent of the choice of D.

Proof

For the operator $A\text{-}DC$, let the relative minimal polynomial (mod $\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$) of $\mathcal{R}_0$ be denoted by $\Psi_0(\cdot)$. From Section 2.3, the roots of $\Psi_0(\cdot)$ are eigenvalues of $A\text{-}DC$ and by definition:

$$\Psi_0(A\text{-}DC)\,\mathcal{R}_0 \equiv \phi \qquad (\text{mod } \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r)$$

But since $\mathcal{R}_0 \subset \eta(C)$, it follows that this is equivalent to:

$$\Psi_0(A)\,\mathcal{R}_0 \equiv \phi \qquad (\text{mod } \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r)$$

and the eigenvalues of $A\text{-}DC$ associated with $\mathcal{R}_0$ are independent of D.

Δ

Lemma 3.3.5 permits a study of the mutual detectability of $f_1 \ldots f_r$ to focus on $\mathcal{R}_0$. From condition iv) in Definition 3.3.3 and from Definition 3.3.4, the image of $\mathcal{R}_0$ under A can be expressed:

$$A\mathcal{R}_0 \subset \mathcal{R}_0 \oplus \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r \tag{3.3.18}$$

Since $\mathcal{R}_0$ is in $\eta(C)$, but also observable, there can be no subspace $\mathcal{R}_0'$ of $\mathcal{R}_0$ which is invariant with respect to A (or with respect to A-DC):

$$A\mathcal{R}_0' \not\subset \mathcal{R}_0' \qquad \forall \mathcal{R}_0' \subset \mathcal{R}_0 \tag{3.3.19}$$

Similarly, there can be no subspace $\mathcal{R}_0'$ of $\mathcal{R}_0$ whose image under A is contained in $\mathcal{R}_0'$ and some single detection space $\bar{\mathcal{R}}_k$:

$$A\mathcal{R}_0' \not\subset \mathcal{R}_0' \oplus \bar{\mathcal{R}}_k \quad \forall \mathcal{R}_0' \subset \mathcal{R}_0 \quad 1 \geq k \geq r \tag{3.3.20}$$

Otherwise, Definition 3.1.2 could be used to demonstrate that $\mathcal{R}_0'$ must be a subspace of $\bar{\mathcal{R}}_k$. Equations (3.3.18) - (3.3.20) support the intuitive argument made in connection with the example of Section 3.3.2 that $\mathcal{R}_0$ is related by the operator A to $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ in such a manner that it cannot be a part of any one of the detection spaces.

If $\mathcal{R}_0$ is chosen arbitrarily to satisfy Definition 3.3.4, (3.3.18) indicates the relationship of $\mathcal{R}_0$ to $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$. There is one particular choice of $\mathcal{R}_0$ for which (3.3.18) can be simplified. This excess subspace is denoted by $\mathcal{R}_{0g}$ and its existence is demonstrated constructively in Lemma 3.3.6. The proof of Lemma 3.3.6 proceeds by first considering only the image of $\mathcal{R}_0$ under A into $\bar{\mathcal{R}}_1$, and then into $\bar{\mathcal{R}}_2$, etc. This is accomplished by defining the subspace:

$$\bar{\mathcal{R}}_j^D = \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_{j-1} \oplus \bar{\mathcal{R}}_{j+1} \oplus \ldots \bar{\mathcal{R}}_r \quad 1 \geq j \geq r \tag{3.3.21}$$

and considering the image of $\mathcal{R}_0$ under $A \pmod{\bar{\mathcal{R}}_j^D}$. Obviously, $\bar{\mathcal{R}}_J$ in (3.3.17) can be expressed in terms of $\bar{\mathcal{R}}_j^D$:

$$\bar{\mathcal{R}}_J = \bar{\mathcal{R}}_j \oplus \bar{\mathcal{R}}_j^D \oplus \bar{\mathcal{R}}_0 \qquad 1 \geq j \geq r \tag{3.3.22}$$

Lemma 3.3.6 Let $f_1 \ldots f_r$ be output separable and $g_1 \ldots g_r$ be their respective detection generators. There exists a unique excess subspace $\mathcal{R}_{0g}$ of $\bar{\mathcal{R}}_J$ such that:

$$A\mathcal{R}_{0g} \subset \mathcal{R}_{0g} \oplus g_1 \oplus \ldots g_r$$

Proof

Let $\xi_1 \ldots \xi_{\nu_0}$ be an arbitrary basis for $\mathcal{R}_0$. Choose some D such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant. Clearly since $\mathcal{R}_0 \subset \eta(C)$:

$$(A-DC)\mathcal{R}_0 = A\mathcal{R}_0 \subset \bar{\mathcal{R}}_J$$

Consider first the image of $\mathcal{R}_0$ under A into $\bar{\mathcal{R}}_1 \pmod{\bar{\mathcal{R}}_1^D}$. Let μ_i, $i = 1 \ldots \nu_0$ be the smallest integers such that the vectors $(A-DC)^{\mu_i}\xi_i$ have components along Cf_1:

$$\text{i)}\quad C(A-DC)^{\mu_i}\xi_i = CA^{\mu_i}\xi_i \equiv \alpha_i Cf_1 \pmod{C\bar{\mathcal{R}}_1^D} \quad i = 1 \ldots \nu_0$$

Because $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ (and hence $\bar{\mathcal{R}}_J$) are invariant, the image of ξ_i under $A \pmod{\bar{\mathcal{R}}_1^D}$ is contained in $\mathcal{R}_1 \oplus \mathcal{R}_0$ for $k < \mu_i$:

$$A^k\xi_i \in \mathcal{R}_1 \oplus \mathcal{R}_0 \pmod{\bar{\mathcal{R}}_1^D} \qquad \begin{matrix} i = 1 \ldots \nu_0 \\ k < \mu_i \end{matrix}$$

If $\mu_i < \nu_1$, define the new basis:

$$\xi_i' = \xi_i - \alpha_i A^{\nu_1 - \mu_i - 1} g_1 \qquad i = 1 \ldots \nu_0$$

Lemma 3.2.4 guarantees that $g_1, Ag_1 \ldots A^{\nu_1 - 1}g_1$ are in $\bar{\mathcal{R}}_1$. It follows that $\xi_1' \ldots \xi_{\nu_0}'$ are in $\bar{\mathcal{R}}_J$ and define a new excess

subspace $\mathcal{R}_0'$ for $\bar{\mathcal{R}}_{\mathcal{J}}$. If condition i) is applied to the basis of $\mathcal{R}_0'$, it follows that $\mu_i' \geq \mu_i$ since:

$$CA^{\mu_i} \xi_i' = CA^{\mu_i} \xi_i - \alpha_i CA^{\nu_1 - 1} g_1 = \emptyset \pmod{C\bar{\mathcal{R}}_1^D} \quad i = 1 \ldots \nu_0 \quad \forall \mu_i < \nu_1$$

Continue in this manner until an excess subspace $\mathcal{R}_0^{(j)}$ is generated whose basis vectors $\xi_1^{(j)} \ldots \xi_{\nu_0}^{(j)}$ satisfy:

ii) $$A^k \xi_i^{(j)} \in \mathcal{R}_1 \oplus \mathcal{R}_0^{(i)} \pmod{\bar{\mathcal{R}}_1^D} \quad k < \nu_1 \quad i = 1 \ldots \nu_0$$

But g_1 is the unique vector $g_1 \in \bar{\mathcal{R}}_1$ such that for $k < \nu_1 - 1$:

$$A^k g_1 \in \mathcal{R}_1 \qquad k < \nu_1 - 1$$

Therefore, the image of $\xi_i^{(j)}$ under $A \pmod{\bar{\mathcal{R}}_1^D}$ must be restricted to $\mathcal{R}_0^{(i)} \oplus g_1$:

$$A \xi_i^{(j)} \in \mathcal{R}_0^{(j)} \oplus g_1 \pmod{\bar{\mathcal{R}}_1^D} \quad i = 1 \ldots \nu_0$$

The construction of $\mathcal{R}_{0g}$ is completed by continuing the iteration for $\bar{\mathcal{R}}_2$, $\bar{\mathcal{R}}_3$ etc.

To show uniqueness, let $\mathcal{R}_0$ be any excess subspace for $\bar{\mathcal{R}}_{\mathcal{J}}$. By Definition 3.3.4, $\mathcal{R}_{0g}$ and $\mathcal{R}_0$ must be equivalent (mod $\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$):

$$\mathcal{R}_{0g} \equiv \mathcal{R}_0 \pmod{\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r}$$

If the two subspaces are not equal, equivalence (mod $\bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$) implies that there must exist same vector $\xi_0 \in \mathcal{R}_0$ such that there exists a $\xi_{0g} \in \mathcal{R}_{0g}$ and $\xi_{\mathcal{J}} \in \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$ which are both non-zero and for which:

$$\xi_0 = \xi_{0g} + \alpha \xi_{\mathcal{J}}$$

For convenience, assume $\xi_{\mathcal{J}} \in \bar{\mathcal{R}}_1$. This implies that there exist some $k < \nu_1$ for which

$$CA^k \xi_{\mathcal{J}} = \alpha_{\mathcal{J}} Cf_1 \quad k < \nu_1$$

It then follows that for the same k:

$$CA^k \xi_0 = CA^k \xi_{0g} + \alpha\, CA^k \xi_{\mathcal{J}} = \alpha\, CA^k \xi_{\mathcal{J}} \neq \emptyset \quad k < \nu_1$$

Therefore, ξ_0 violates ii) and $\mathcal{R}_0$ cannot satisfy the hypothesis.

Δ

A variation on the constructive proof of Lemma 3.3.6 can be used to define an algorithm for $\mathcal{R}_{0g}$. One such algorithm is suggested in Appendix A.3. The utility of $\mathcal{R}_{0g}$ is that it allows the excess subspace to be viewed as an entity in itself. The study of the $\nu_{\mathcal{J}}$-dimensional detection space $\bar{\mathcal{R}}_{\mathcal{J}}$ is reduced to the study of the $(\nu_0 + r)$-dimensional space $\mathcal{R}_{0g} \oplus g_1 \oplus \ldots\ g_r$.

Assume D is chosen such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant with respect to $A\text{-}DC$. Assume that a partition of $\mathcal{R}_{0g}$ exists:

$$\mathcal{R}_{0g} = \mathcal{R}_{0gk} \oplus \mathcal{R}^c_{0gk} \tag{3.3.23}$$

such that the image of $\mathcal{R}_{0gk}$ under A does not have a component along g_k:

$$A\,\mathcal{R}_{0gk} \subset \mathcal{R}_{0gk} \oplus g_1 \oplus \ldots\ g_{k-1} \oplus g_{k+1} \oplus \ldots\ g_r \tag{3.3.24}$$

If $\mathcal{R}_{0gk}$ is the largest subspace which satisfies (3.3.23) and (3.3.24), then the observability of $\mathcal{R}_{0g}$ implies that $\mathcal{R}^c_{0gk}$ satisfies:

$$\mathcal{R}^c_{0gk} \subset \mathcal{R}_{0g}$$

$$A\mathcal{R}^{c}_{0gk} \subset \mathcal{R}^{c}_{0gk} \oplus g_k \pmod{\bar{\mathcal{R}}^{D}_{k} \oplus \mathcal{R}_{0gk}} \quad k = 1 \ldots r \qquad (3.3.25)$$

and that $\mathcal{R}^{c}_{0gk}$ is observable (mod $\bar{\mathcal{R}}^{D}_{k}$), i.e. every event vector in $\mathcal{R}^{c}_{0gk}$ generates an output error signal which has a component along Cf_k. The dimension of $\mathcal{R}^{c}_{0gk}$ is:

$$\nu^{c}_{0k} = d(\mathcal{R}^{c}_{0gk}) \qquad (3.3.26)$$

$\mathcal{R}^{c}_{0gk}$ would be a subspace of $\bar{\mathcal{R}}_k$ except that the output error signal for every event vector in $\mathcal{R}^{c}_{0gk}$ also has a component along some direction other than Cf_k--hence the need for the (mod $\bar{\mathcal{R}}^{D}_{k} \oplus \mathcal{R}_{0gk}$) argument in (3.3.25).

The importance of $\mathcal{R}^{c}_{0gk}$ is that it represents the contribution of f_k to the excess subspace of $\bar{\mathcal{R}}_{\mathcal{I}}$. If f_k is removed from $f_1 \ldots f_r$, (3.3.25) implies that no vector in $\mathcal{R}^{c}_{0gk}$ can be in the new excess subspace because the output error generated by event vectors in $\mathcal{R}^{c}_{0gk}$ cannot be constrained to the output space spanned by $Cf_1 \ldots C f_{k-1}, Cf_{k+1} \ldots Cf_r$. On the other hand, (3.3.24) suggests that the event vectors in $\mathcal{R}_{0gk}$ remain in the new excess subspace since they generate no output component along $C f_k$. It is proved in Lemma 3.3.7 that the new excess subspace consists only of $\mathcal{R}_{0gk}$. The result of removing f_k is that the dimension of the excess subspace is decreased to:

$$\nu_{0k} = d(\mathcal{R}_{0gk}) = \nu_0 - \nu^{c}_{0k} \qquad (3.3.27)$$

Given that $f_1 \ldots f_r$ are not mutually detectable, the issue is to determine a procedure for finding a subset of $f_1 \ldots f_r$ which are mutually detectable. Two algorithms for accomplishing this are developed in the remainder of this section. The first adhers to the vector space concepts upon which the earlier results have been based. The resulting algorithm is simple to obtain, but computationally inefficient. The second algorithm is formulated in a matrix algebra context and is the recommended

algorithm for implementation. Both are similar in that they restrict attention to the excess subspace $\mathcal{R}_{0g}$ and determine which subset of the event vectors $f_1 \ldots f_r$ can be removed such that the remaining event vectors are mutually detectable.

The algorithms begin by considering the set of event vectors $f_1 \ldots f_r$ with a single event vector f_k removed. If $\mathcal{I}_k$, $\bar{\mathcal{R}}_{\mathcal{I}k}$, and $\mathcal{R}_{0gk}$ for the remaining r-1 event vectors are defined similarly to $\mathcal{I}$, $\bar{\mathcal{R}}_{\mathcal{I}}$, and $\mathcal{R}_{0g}$; it is shown in Lemma 3.3.7 that:

$$\mathcal{I}_k \subset \mathcal{I} \tag{3.3.28}$$

$$\bar{\mathcal{R}}_{\mathcal{I}k} \subset \bar{\mathcal{R}}_{\mathcal{I}}$$

$$\mathcal{R}_{0gk} \subset \bar{\mathcal{R}}_{\mathcal{I}}$$

$$\mathcal{R}_{0gk} \subset \mathcal{R}_{0g}$$

$\mathcal{R}_{0gk}$ and $\bar{\mathcal{R}}_{\mathcal{I}k}$ could be computed using the algorithms of Appendix A.3, but there are computational advantages to developing algorithms in terms of $\mathcal{R}_{0g}$ and $\bar{\mathcal{R}}_{\mathcal{I}}$. The algorithm developed in Lemma 3.3.7 is iterative and is patterned after Lemma 3.2.5. The iteration begins with an initial excess subspace:

$$\mathcal{R}_0^{(0)} = \mathcal{R}_{0g} \tag{3.3.29}$$

and removes all vectors from $\mathcal{R}_0^{(0)}$ whose image under A is not contained in $\mathcal{R}_0^{(0)} \oplus g_1 \oplus \ldots g_{k-1} \oplus g_{k+1} \oplus \ldots g_r$. The deleted vectors are those in $\mathcal{R}_{0gk}^c$. The remaining space is denoted by $\mathcal{R}_0^{(1)}$ where:

$$d(\mathcal{R}_0^{(1)}) \leq d(\mathcal{R}_0^{(0)}) \tag{3.3.30}$$

The test is then repeated on $\mathcal{R}_0^{(1)}$, etc. until no further vectors are removed--a maximum of ν_0 cycles. The terminating subspace in the sequence is $\mathcal{R}_{0gk}$.

As with Lemma 3.2.5, the algorithm of Lemma 3.3.7 is based upon A^{-1}. If A^{-1} does not exist but (A, C) is observable, it is sufficient to have the inverse operator $A_\eta^{-1} : \mathcal{S}^n \to \eta(C)$ which satisfies:

$$A_\eta^{-1} \; A(\eta(C)) \; = \; \eta(C) \tag{3.3.31}$$

This operator always exists. Implementation of the algorithm must be along the lines mentioned in Section 3.2.3.

Lemma 3.3.7 Assume (A, C) is observable. If f_k is removed from the set of output separable vectors $f_1 \ldots f_r$, $\mathcal{R}_{0gk}$ satisfies:

$$\mathcal{R}_{0gk} \; = \; \mathcal{R}_0^{(\nu_0)}$$

where $\mathcal{R}_0^{(0)}$ is defined by the recursive relationship:

$$\mathcal{R}_0^{(0)} = \mathcal{R}_{0g}$$

and:

$$\mathcal{R}_0^{(j)} = \mathcal{R}_0^{(j-1)} \cap \left[A_\eta^{-1} \left(\mathcal{R}_0^{(j-1)} \oplus g_1 \oplus \ldots \; g_{k-1} \oplus g_{k+1} \oplus \ldots \; g_r\right)\right] \qquad j = 1 \ldots \nu_0$$

Under suitable invertibility conditions, A_η^{-1} can be replaced by A^{-1}.

Proof

Clearly $\mathcal{J}_k \subset \mathcal{J}$ which, along with Definition 3.3.3, implies that $\bar{\mathcal{R}}_{\mathcal{J}k} \subset \bar{\mathcal{R}}_{\mathcal{J}}$ By Lemma 3.3.6, the image of $\mathcal{R}_{0gk}$ under A satisfies:

$$A\mathcal{R}_{0gk} \subset \bar{\mathcal{R}}_{0gk} \oplus g_1 \oplus \ldots \; g_{k-1} \oplus g_{k+1} \oplus \ldots \; g_r$$

The inequality is unchanged if g_k is added on the right:

$$A\mathcal{R}_{0gk} \subset \mathcal{R}_{0gk} \oplus g_1 \oplus \ldots \; g_r$$

$\mathcal{R}_{0g}$ is the largest subspace of η (C) which satisfies this relationship. By the uniqueness of $\mathcal{R}_{0g}$ and $\mathcal{R}_{0gk}$, and the fact that $\bar{\mathcal{R}}_{\mathcal{J}k} \subset \bar{\mathcal{R}}_{\mathcal{J}}$ it follows that $\mathcal{R}_{0gk} \subset \mathcal{R}_{0g}$. With respect to the iteration, $\mathcal{R}_{0gk} \subset \mathcal{R}_0^{(0)}$. With this assertion the proof of Lemma 3.2.5 is directly applicable.

If: Δ

$$\mathcal{R}_{0gk} = \phi \qquad 1 \leq j \leq r \tag{3.3.32}$$

the event vectors $f_1 \ldots f_{k-1}, f_{k+1}, \ldots f_r$ are mutually detectable. If (3.3.32) is not satisfied for any $1 \leq k \leq r$, two or more event vectors from $f_1 \ldots f_r$ must be removed in order to yield a mutually detectable set of event vectors. It would be possible to generalize Lemma 3.3.7 to the removal of ℓ event vectors, but a simpler procedure is to use $\mathcal{R}_{0gk}$ and $\mathcal{R}_{0gj}$ in a direct manner to determine the new excess subspace if both f_k and f_j are removed from $f_1 \ldots f_r$. If I is the index set composed of the two indices k and j, and $\mathcal{R}_{0gI}$ denotes the new excess subspace with f_k and f_j removed, Theorem 3.3.2 states that $\mathcal{R}_{0gI}$ is the intersection of $\mathcal{R}_{0gk}$ and $\mathcal{R}_{0gj}$:

$$\mathcal{R}_{0gI} = \mathcal{R}_{0gk} \cap \mathcal{R}_{0gj} \qquad \begin{matrix} 1 \leq j \leq r \\ 1 \leq k \leq r \end{matrix} \tag{3.3.33}$$

Equation (3.3.33) is applicable to a general set of ℓ vectors. The only vectors in $\mathcal{R}_{0gI}$ are those which are contained in both $\mathcal{R}_{0gk}$ and $\mathcal{R}_{0gj}$, i.e. they are the vectors in $\mathcal{R}_{0g}$ whose presence is not affected by the removal of either f_k or f_j. The theorem suggests that a reasonable algorithm for finding a detectable subset of $f_1 \ldots f_r$ is to compute $\mathcal{R}_{0g1} \ldots \mathcal{R}_{0gr}$ and then remove the ℓ event vectors for which the generalization of (3.3.33) yields the null space ϕ.

Theorem 3.3.2 Given the output separable event vectors $f_1 \ldots f_r$, let I be an arbitrary subset of the index set

$\{1, 2, \ldots r\}$. If the vectors f_k with $k \in I$ are removed from $f_1 \ldots f_r$, the excess subspace associated with the remaining event vectors is:

$$\mathcal{R}_{0gI} = \bigcap_{k \in I} \mathcal{R}_{0gk}$$

Proof

Lemma 3.3.6 states that $\mathcal{R}_{0gk}$ satisfies:

$$A\mathcal{R}_{0gk} \subset \mathcal{R}_{0gk} \oplus g_1 \oplus \ldots g_{k-1} \oplus g_{k+1} \oplus \ldots g_r \quad 1 \leq k \leq r$$

From the above and the equation defining $\mathcal{R}_{0gI}$, it follows that $\mathcal{R}_{0gI}$ satisfies a similar relationship if the vectors g_k, $k \in I$, are removed from $g_1 \oplus \ldots g_r$:

i) $$A\mathcal{R}_{0gI} \subset \mathcal{R}_{0gI} \bigcup_{j \in I} g_j$$

Lemma 3.3.7 proves that $\mathcal{R}_{0gk} \subset \mathcal{R}_{0g}$ for $1 \leq k \leq r$. It follows that $\mathcal{R}_{0gI} \subset \mathcal{R}_{0g}$. From the discussion preceding Lemma 3.3.7 (also, the omitted portion of the proof of Lemma 3.3.7), it is also true that $\mathcal{R}_{0gI}$ is the largest subspace of $\mathcal{R}_{0g}$ to satisfy i). By Definition 3.3.4, the excess subspace of $f_1 \ldots f_r$ with all f_k, $k \in I$, removed must satisfy i). A generalization of the proof of Lemma 3.3.7 proves that it must also be a subspace of $\mathcal{R}_{0g}$. Hence, $\mathcal{R}_{0gI}$ is the new excess subspace.

Δ

Except for a few minor additions, Theorem 3.3.2 completes the development of the structure of $\bar{\mathcal{R}}_{\mathcal{J}}$ and of algorithms for finding a mutually detectable subset of $f_1 \ldots f_r$. If $\mathcal{R}_{0gI}$ is the zero space ϕ, the corresponding set of event vectors is mutually detectable. The remainder of this subsection is devoted to matrix-based algorithms which are computationally simpler than Lemma 3.3.5, Lemma 3.3.7, and Theorem 3.3.2. The algorithms are essentially those developed by Beard. The following algorithms could be stated as corollaries of the previous results and their proofs omitted; however, proofs are included in an effort to bridge the gap between the constructive approach to failure detection used by Beard and the abstract approach which has dominated this chapter. The proofs are similar to those given by Beard, but emphasize the notation and nomenclature of this thesis.

Generalization of matrix algorithms requires that the nomenclature developed in Section 2.3.3 for the matrix representation of a vector space be utilized. If a coordinate system is adopted for the state space $\mathcal{S}^n$, matrices $\bar{R}_1 \ldots \bar{R}_r$, $\bar{R}_{\mathcal{J}}$ and R_{0g} can be chosen such that their columns form a basis for the respective vector spaces $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$, $\bar{\mathcal{R}}_{\mathcal{J}}$, $\mathcal{R}_{0g}$. Similarly, the matrix G can be defined in terms of the vectors $g_1 \ldots g_r$. In the remainder of this chapter, capital letters are used to define matrices in this manner. When no coordinate system is explicitly defined, an arbitrarily chosen coordinate system is assumed.

The invariance of the detection space $\bar{\mathcal{R}}_1$ with respect to A-DC is denoted by:

$$(A\text{-}DC)\bar{\mathcal{R}}_i \subset \bar{\mathcal{R}}_i \qquad (3.3.34)$$

The excess subspace $\mathcal{R}_{0g}$ of $\bar{\mathcal{R}}_J$ is not invariant with respect to A-DC, but Lemma 3.3.6 indicates that it satisfies a similar relationship:

$$A\mathcal{R}_{0g} \subset \mathcal{R}_{0g} \oplus g_1 \oplus \ldots g_r \tag{3.3.35}$$

The matrix equivalent to invariance is:

$$(A-DC)\bar{R}_i = \bar{R}_i A_{ii} \quad i = 1 \ldots r \tag{3.3.36}$$

Every column of $\bar{R}_i$ when multiplied by A-DC yields a vector in the range space of $\bar{R}_i$. The equivalent to (3.3.35) is:

$$AR_{0g} = R_{0g}\Pi + G\Theta \tag{3.3.37}$$

Matrices $A_{11} \ldots A_{rr}$, Π which satisfy these equations are not difficult to find. If the matrix T has a rank of n:

$$T = [\bar{R}_1 : \bar{R}_2 : \ldots : R_{0g}] \tag{3.3.38}$$

$A_{11} \ldots A_{rr}$, Π are the block diagonal submatrices of the transformed matrix $\hat{A}-\hat{D}\hat{C}$:

$$\hat{A}-\hat{D}\hat{C} = T^{-1}(A-DC)\,T \tag{3.3.39}$$

Appendix A.3 defines algorithms for Π and Θ for the general case.

Equations (3.3.38) and (3.3.39) are a special case of the transformation to base normal form and are discussed further in Section 4.4. The advantage of that coordinate system is that it is a simple matter to define invariant subsystems of ($\hat{A}-\hat{D}\hat{C}$, $\hat{C}$) in terms of

$A_{11} \ldots A_{rr}$ and the appropriate detection space. A subsystem associated with R_{0g} can be defined in terms of Π, and Θ determines the manner in which that subsystem interacts with the invariant subsystems. Because the eigenvalues of a matrix are not altered by a similarity transformation, these facts suggest that the eigenvalues of $A_{11} \ldots A_{rr}$ must be the roots of the minimal polynominals of A-DC for $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ and that the eigenvalues of Π must be the roots of the relative minimal polynomial (mod $\bar{\mathcal{R}}_1 \oplus, \ldots \bar{\mathcal{R}}_r$) of A-DC for $\mathcal{R}_{0g}$. This last fact is proved in Lemma 3.3.8. The lemma is the matrix equivalent to Lemma 3.3.5.

Lemma 3.3.8 Let D be chosen such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant with respect to A-DC. The relative minimal polynomial $\Psi_0(\cdot)$ of $\mathcal{R}_{0g}$ for A-DC is the characteristic equation of Π where Π is defined by:

$$A R_{0g} = R_{0g} \Pi + G\Theta$$

Proof

Equation (3.3.35) is a consequence of Lemma 3.3.6. Since the range space of $\mathcal{R}_{0g}$ is a subspace of $\eta(C)$, the matrix equivalent of (3.3.35) can be replaced by

$$(A\text{-}DC) R_{0g} = R_{0g} \Pi + G\Theta$$

To emphasize the properties of interest this can be expressed (mod $[\bar{\mathcal{R}}_1 : \ldots \bar{\mathcal{R}}_r]$)[1]:

[1] In matrix notation, "(mod R)" is to be interpreted as "modulo the range space of R".

$$(\text{A-DC})\, R_{0g} \equiv R_{0g}\, \Pi \quad (\text{mod } [\bar{R}_1 : \ldots : \bar{R}_r])$$

Because $\bar{\mathcal{R}}_1, \ldots \bar{\mathcal{R}}_r$ are invariant, this relationship is valid for arbitrary powers of A-DC:

i) $$(\text{A-DC})^k\, R_{0g} \equiv R_{0g}\, \Pi^k \quad (\text{mod } [\bar{R}_1 : \ldots : \bar{R}_r])$$

where $k \geq 0$. By definition, the relative minimal polynomial of $\mathcal{R}_{0g}$ satisfies:

$$\Psi_0 (\text{A-DC})\, R_{0g} \equiv 0 \quad (\text{mod } [\bar{R}_1 : \ldots : \bar{R}_r])$$

and it follows from i) that:

$$R_{0g}\, \Psi_0 (\Pi) = 0$$

Since $\Psi_0 (\cdot)$ is a ν_0-order polynomial and Π is $\nu_0 \times \nu_0$, $\Psi_0(\cdot)$ is the characteristic equation of Π.

Δ

Lemma 3.3.8 supports the statement that Π and Θ along with R_{0g} define a subsystem of A-DC. This fact allows the matrix version of Theorem 3.3.2 to be reduced from consideration of an n-dimensional system to a ν_0-dimensional system. One method for accomplishing this reduction is to make the transformation suggested by (3.3.38) and (3.3.39) and then restate the theorem in terms of Π^{-1} instead of A^{-1}. The approach taken in Theorem 3.3.3 is more straightforward in that it procedes directly from (3.3.37) to a solution for the new excess subspace $\mathcal{R}_{0gL}$ when ℓ event vectors are removed from $f_1 \ldots f_r$.

The product $G\Theta$ in (3.3.37) can be expressed:

$$G\Theta = \sum_{k=1,r}' g_k \theta_k \qquad (3.3.40)$$

where θ_k is the k^{th} row of Θ. Equation (3.3.40) clearly implies that each row of Θ is uniquely associated with the corresponding column of G and hence the corresponding detection space. The k^{th} row of Θ determines the manner in which the product AR_{0g} is mapped into the range space of $\bar{R}_k$. If β_i is some ν_0-dimensional coefficient vector, then $R_{0g}\,\beta_i$ is a vector in the range space of R_{0g}. If, in addition, $R_{0g}\beta_i$ is a vector in R_{0gk}, (3.3.37) and (3.3.40) imply that $\theta_k\,\beta_i$ must be zero. Theorem 3.3.3 extends this line of reasoning to determine R_{0gL}. For notational convenience, it is assumed that the first ℓ event vectors in $f_1 \ldots f_r$ are removed.

Theorem 3.3.3 Given the output separable vectors $f_1 \ldots f_r$, assume the first ℓ vectors are removed from the set and let Θ_L^c be the first ℓ rows of Θ. The excess subspace $\mathcal{R}_{0gL}$ for the remaining $r - \ell$ vectors is spanned by the ν_{0L} columns of R_{0gL}:

$$R_{0gL} = R_{0g}\,\beta$$

where the columns of the $\nu_0 \times \nu_{0L}$ matrix β span the null space of:

$$M_{0L} = \begin{bmatrix} \Theta_L^c \\ \Theta_L^c\,\Pi \\ \vdots \\ \Theta_L^c\,\Pi^{\nu_0 - 1} \end{bmatrix}$$

Proof

Let Θ_L be Θ with the first ℓ rows deleted and G_L be G with the first ℓ columns deleted:

$$\Theta = \begin{bmatrix} \Theta_L^c \\ \hline \Theta_L \end{bmatrix}$$

$$G = \left[G_L^c \vdots G_L \right]$$

Because $\mathcal{R}_{0gL}$ is a subspace of $\mathcal{R}_{0g}$, there is a unique $\nu_0 \times \nu_{0L}$ matrix β of rank ν_{0L} such that R_{0gL} satisfies:

i) $$R_{0gL} = R_{0g}\,\beta$$

Lemma 3.3.6 as interpreted for matrices in (3.3.37) requires $\mathcal{R}_{0gL}$ to satisfy:

$$(A\text{-}DC)\,R_{0gL} = R_{0gL}\,\Pi_L + G_L\,\Theta_L$$

for some $\nu_{0L} \times \nu_{0L}$ matrix Π_L. To emphasize the following arguments, this result can be written (mod $[\bar{\mathcal{R}}_{\ell+1} \vdots \ldots \vdots \bar{R}_r]$):

ii) $$(A\text{-}DC)R_{0gL} \equiv R_{0gL}\,\Pi_L \pmod{[\bar{R}_{\ell+1} \vdots \ldots \vdots \bar{R}_r]}$$

The matrix equivalent of Lemma 3.3.6 for R_{0g} (mod $[\bar{R}_{\ell+1} \vdots \ldots \vdots \bar{R}_r]$) differs from ii) in that the range space of (A-DC) R_{0g} may contain vectors which have a component in the range space of $\bar{R}_1 \ldots \bar{R}_\ell$:

$$(A\text{-}DC)\quad R_{0g} \equiv R_{0g}\,\Pi + G_L^c\,\Theta_L^c \pmod{[\bar{R}_{\ell+1} \vdots \ldots \vdots \bar{R}_r]}$$

When post-multiplied by β, this equation becomes:

iii) $$(A\text{-}DC)\,R_{0g}\,\beta \equiv R_{0g}\,\Pi\,\beta + G_L^c\,\Theta_L^c\,\beta \pmod{[\bar{R}_{\ell+1} \vdots \ldots \vdots \bar{R}_r]}$$

In order for substitution of i) into iii) to have the general form of ii), it is necessary that, for all acceptable β, a $\nu_{0L} \times \nu_{0L}$ matrix Π_L exist for which:

$$\beta \Pi_L = \Pi \beta$$

If Π_L exists, iii) becomes:

iii)' (A-DC) $R_{0g} \beta \equiv R_{0g} \beta \Pi_L + G_L^c \Theta_L^c \beta \pmod{[\bar{R}_{\ell+1} : \ldots : \bar{R}_r]}$

Because β is $\nu_0 \times \nu_{0L}$ and of rank ν_{0L}, it always has a left inverse and such a Π_L exists. Equation iii)' then places the first constraint of β since it is necessary for the last term to be zero. Because $g_1 \ldots g_\ell$ are linearly independent (Lemma 3.3.3), $\mathrm{rk} G_L^c = \ell$ and β must satisfy:

iv) $$\Theta_L^c \beta = 0$$

If D is chosen such that $\bar{R}_{\ell+1} \ldots \bar{R}_r$ are invariant, $(A\text{-}DC)^2 R_{0gL}$ is also contained in $R_{0gL} \pmod{[\bar{R}_{\ell+1} : \ldots : \bar{R}_r]}$. The equivalent equation for $R_{0g} \beta$ is (assuming β satisfies iv)):

$$(A\text{-}DC)^2 R_{0g} \beta \equiv (A\text{-}DC) R_{0g} \Pi \beta \pmod{[\bar{R}_{\ell+1} : \ldots : \bar{R}_r]}$$

$$= R_{0g} \Pi^2 \beta + G_L^c \Theta_L^c \Pi \beta$$

$$= R_{0g} \beta \Pi_L^2 + G_L^c \Theta_L^c \Pi \beta$$

and comparison with ii) requires that β meet an additional constraint:

$$\Theta_L^c \Pi \beta = 0$$

A total of ν_0 such constraints on β can be formulated. They are equivalent to requiring that β be in $\eta(M_{0L})$. Since R_{0gL} must be the largest subspace of R_{0g} to satisfy ii), β must span $\eta(M_{0L})$.

Δ

Theorem 3.3.3 is not as useful as Theorem 3.3.2 in that it does not allow R_{0gL} to be related to the R_{0gk}, $k = 1 \ldots r$, which result from removing the corresponding f_k from $f_1 \ldots f_r$. This shortcoming is eliminated by the following two corollaries. Corollary 1 provides a simple mechanism for finding R_{0gL} and R^c_{0gL} along with Π_L and Θ_L. The corollary is particularly useful when state space augmentation is considered in Section 3.3.4.

Corollary 1 Given the conditions of Theorem 3.3.3, define the transformations:

$$\Pi^0_L = [\Delta : \beta]^{-1}_L \Pi [\Delta : \beta]_L$$

$$(G\Theta)^0_L = G\Theta [\Delta : \beta]_L$$

$$R^0_{0g} = R_{0g} [\Delta : \beta]_L$$

where the columns of β and Δ span the null space and row space of M_{0L} respectively. If $(G\Theta)^0_L$ and Π^0_L are partitioned:

$$\Pi^0_L = \begin{bmatrix} \Pi^c_L & 0 \\ \Pi^c_{L0} & \Pi_L \end{bmatrix}$$

$$(G\Theta)^0_L = [G^c_L \Theta^c_L \Delta + G_L \Theta_L \Delta : G_L \Theta_L \beta]$$

with $G_L \Theta_L \beta$ a $n \times \nu_{0L}$ submatrix, Π_L a $\nu_{0L} \times \nu_{0L}$ submatrix etc., then R_{0gL} and R^c_{0gL} satisfy:

$$AR^c_{0gL} = R^c_{0g} \Pi^c_L + R_{0gL} \Pi^c_{L0} + G_L \Theta_L \Delta + G^c_L \Theta^c_L \Delta$$

$$AR_{0gL} = R_{0gL} \Pi_L + G_L \Theta_L \beta$$

where R_{0gL} and R^c_{0gL} can be determined from a partion of R^0_{0g}:

$$R^0_{0g} = [R^c_{0gL} : R_{0gL}]$$

Proof

R_{0g} is not a unique matrix representation for $\mathcal{R}_{0g}$ -- R^0_{0g} is an equivalent representation. By Theorem 3.3.3, $R_{0g}\,\beta$ is R_{0gL}. Since the columns of Δ are not in $\eta(M_{0L})$, no vector in the range space of $R_{0g}\,\Delta$ is in R_{0gL}. It follows that R^0_{0g} can be expressed:

$$R^0_{0g} = [R^c_{0gL} : R_{0gL}]$$

The argument used to show the existence of Π_L in Theorem 3.3.3 can be generalized to show that with Π^0_L and $(G\Theta)^0_L$ as defined, (3.3.37) can be written for R^0_{0g}:

i) $$A\,R^0_{0g} = R^0_{0g}\,\Pi^0_L + (G\,\Theta)^0_L$$

The corollary follows from a partition of i). The zero submatrix in Π^0_L appears because AR_{0gL} must satisfy ii) from the proof of Theorem 3.3.3. The $G_L\,\Theta_L\,\beta$ term appears in $(G\,\Theta)^0_L$ because $G\Theta$ can be written:

$$G\Theta = G^c_L\,\Theta^c_L + G_L\,\Theta_L$$

and β is in the null space of $G^c_L\,\Theta^c_L$. Δ

Theorem 3.3.2 provides an algorithm for $\mathcal{R}_{0gI}$ in terms of $\mathcal{R}_{0gk}$, $k \in I$. Corollary 2 provides a similar test based upon the eigenvalues of the matrices Π_k, $k \in I$, obtained by applying Corollary 1 for each f_k, $k \in I$. The proof is similar to that of Theorem 3.3.2 and is omitted.

Corollary 2 Let $\Pi_1 \ldots \Pi_r$ be solutions of Corollary 1 if the corresponding f_k is removed from $f_1 \ldots f_r$ and let the elements of the sets $\Lambda_1 \ldots \Lambda_r$ be the eigenvalues of $\Pi_1 \ldots \Pi_r$ respectively. If the index set I is an arbitrary subset of $\{1, 2, \ldots r\}$ and all f_k for which $k \in I$ are removed from $f_1 \ldots f_r$, the eigenvalues associated with the excess subspace of the remaining f_j with $j \notin I$ are the elements of:

$$\Lambda_I = \bigcap_{k \in I} \Lambda_k$$

3.3.4 State Space Augmentation

There are a number of circumstances in which a change in the dimension of the state space of the reference model may make the failure detection problem simpler. It was mentioned in Section 3.2.2 that if (A, C) is not observable the computational requirements in the design of a detection system can be decreased by accepting a lower order, observable system model. It is assumed here that (A, C) is observable. The results of this subsection show that if $f_1 \ldots f_r$ are output separable, there is an observable reference model $(\tilde{A}, \tilde{C})$ of order $\tilde{n} \geq n$ for which they are mutually detectable. Following the convention of this chapter, the result is derived first as a general theorem on linear operators and then restated as a matrix-based algorithm. The state space augmentation techniques complement the techniques of Section 3.3.3. If $f_1 \ldots f_r$ are not mutually detectable, state space augmentation is an alternative to removing vectors from the set. Criteria for choosing between the two approaches are discussed in Section 4.2.

Wonham and Morse (1970b) first indicated that certain systems could be decoupled only if the dimension of the state space was increased. Their solution to the extended decoupling problem was based upon geometric arguments similar to those

used here. Howze and Pearson (1970) among others introduced a matrix augmentation scheme known as dynamic compensation. They proved that for the decoupling problem it is sufficient to add a series of integrators to the system model which are controllable but not observable. For the failure detection problem the reverse is true. It is necessary to duplicate the dynamics associated with a portion of the excess subspace in such a manner that they are observable but uncontrollable.

Section 3.3.3 revealed that for any $\bar{\mathcal{R}}_k$, $\mathcal{R}_{0g}$ can be partitioned into a subspace $\mathcal{R}^c_{0gk}$ whose event vectors generate output errors with a component along $\mathcal{C}f_k$ and a subspace $\mathcal{R}_{0gk}$ whose event vectors do not:

$$\mathcal{R}_{0g} = \mathcal{R}_{0gk} \oplus \mathcal{R}^c_{0gk} \qquad (3.3.23)$$

The images under A of these partitions are:

$$A\mathcal{R}^c_{0gk} \subset \mathcal{R}^c_{0gk} \oplus g_k \pmod{\bar{\mathcal{R}}^D_k \oplus \mathcal{R}_{0gk}} \qquad (3.3.25)$$

$$A\mathcal{R}_{0gk} \subset \mathcal{R}_{0gk} \oplus g_1 \oplus \cdots g_{k-1} \oplus g_{k+1} \oplus \cdots g_r \qquad (3.3.24)$$

$\mathcal{R}_{0gk}$ is completely independent of $\bar{\mathcal{R}}_k$ and the presence of f_k in $f_1 \cdots f_r$. Conversely, the presence of $\mathcal{R}^c_{0gk}$ in $\mathcal{R}_{0g}$ is dependent upon f_k being in $f_1 \cdots f_r$. If f_k is removed from $f_1 \cdots f_r$, $\mathcal{R}_{0gk}$ is the new excess subspace (Lemma 3.3.7). Section 3.3.3 concentrated on the properties of $\mathcal{R}_{0gk}$ and an algorithm for finding a mutually detectable subset of $f_1 \cdots f_r$. In this subsection attention is restricted to $\mathcal{R}^c_{0gk}$ because the emphasis is upon finding a higher dimensional reference model $(\tilde{A}, \tilde{C})$ for which the presence or absence of f_k in the set $f_1 \cdots f_r$ no longer affects the dimension of $\mathcal{R}_{0g}$. A prerequisite for this effort is a more detailed definition of the structure of A than was required in Section 3.3.3.

Equations (3.3.23) - (3.3.25) imply that the image of $\mathcal{R}_{0gk}$ under A has no component in $\bar{\mathcal{R}}_k$, but that the image of $\mathcal{R}^c_{0gk}$ under A does. Adopting the linear operator equivalent of the notation introduced in Corollary 1 of Theorem 3.3.3, if a partition of $\mathcal{S}^n$ is established in terms of (3.3.23) for some particular $1 \le k \le r$:

$$\begin{aligned}\mathcal{S}^n &= \bar{\mathcal{R}}_{\mathcal{J}} \oplus \bar{\mathcal{R}}^c_{\mathcal{J}} \\ &= \bar{\mathcal{R}}_k \oplus (\bar{\mathcal{R}}^D_k \oplus \mathcal{R}_{0gk}) \oplus \mathcal{R}^c_{0gk} \oplus \bar{\mathcal{R}}^c_{\mathcal{J}}\end{aligned} \qquad (3.3.41)$$

then A can be specified in terms of sixteen operators in the four subspaces of $\mathcal{S}^n$. Six of these reflect the mappings indicated in (3.3.23) - (3.3.25):

$$\begin{aligned}\Pi^c_k &: \mathcal{R}^c_{0gk} \to \mathcal{R}^c_{0gk} \\ \Pi^c_{k0} &: \mathcal{R}^c_{0gk} \to \mathcal{R}_{0gk} \\ \Theta^c_k &: \mathcal{R}^c_{0gk} \to \bar{\mathcal{R}}_k\end{aligned} \qquad (3.3.42)$$

$$\begin{aligned}\Pi_k &: \mathcal{R}_{0gk} \to \mathcal{R}_{0gk} \\ \phi &: \mathcal{R}_{0gk} \to \mathcal{R}^c_{0gk} \oplus \bar{\mathcal{R}}_k \\ \Theta_k &: \mathcal{R}_{0gk} \to \bar{\mathcal{R}}^D_k\end{aligned}$$

The approach used for the state space augmentation is to enlarge the state space to an $(n+\nu^c_{0k})$ - dimensional state space:

$$\mathcal{S}^{\tilde{n}} = \mathcal{S}^n \oplus \mathcal{S}^{\nu^c_{0k}} \qquad (3.3.43)$$

and to define an operator $\tilde{A}$ on $\mathcal{S}^{\tilde{n}}$ such that the image of $\mathcal{S}^{n}$ under $\tilde{A}$ is the same as that of $\mathcal{S}^{n}$ under A, and the image of $\mathcal{S}^{\nu^c_{0k}}$ under $\tilde{A}$ is similar to that of $\mathcal{R}^c_{0gk}$ under A (i.e. under $\tilde{A}$). In this augmented state space, the detection spaces of $f_1 \ldots f_r$ and $\mathcal{J}$ may also be augmented and are denoted by $\tilde{\bar{R}}_1 \ldots \tilde{\bar{R}}_r$ and $\tilde{\bar{R}}_{\mathcal{J}}$. The construction of $\tilde{A}$ insures that:

$$\tilde{\bar{R}}_k = \bar{R}_k \oplus \mathcal{S}^{\nu^c_{0k}}$$

$$\tilde{\bar{R}}_{\mathcal{J}} = \bar{R}_{\mathcal{J}} \oplus \mathcal{S}^{\nu^c_{0k}} \qquad (3.3.44)$$

More importantly, the excess subspace $\tilde{R}_{0g}$ for $\tilde{\bar{R}}_{\mathcal{J}}$ has the desired independence from the presence of f_k in $f_1 \ldots f_r$ i.e. $\tilde{R}^c_{0gk}$ is zero. $\tilde{R}_{0g}$ then satisfies a relationship similar to that of R_{0gk} in (3.3.25):

$$\tilde{A}\,\tilde{R}_{0g} \subset \tilde{R}_{0g} \oplus \tilde{g}_1 \oplus \ldots \oplus \tilde{g}_{k-1} \oplus \tilde{g}_{k+1} \oplus \ldots \tilde{g}_r \qquad (3.3.45)$$

where $\tilde{g}_1 \ldots \tilde{g}_{k-1}$, $\tilde{g}_{k+1} \ldots \tilde{g}_r$ may be different from the detection generators for the original reference model. The independence of $\tilde{R}_{0g}$ from f_k implies that if a subset of $f_1 \ldots f_r$ which is mutually detectable for $(\tilde{A}, \tilde{C})$ is found, f_k can be added to that set and the enlarged subset is also mutually detectable. It is not necessarily true that the dimension of $\tilde{R}_{0g}$ is less than that of R_{0g}, but it cannot be greater. The properties of $\tilde{R}_{0g}$ and $\tilde{\bar{R}}_1 \ldots, \tilde{\bar{R}}_{k-1}, \tilde{\bar{R}}_{k+1} \ldots \tilde{\bar{R}}_r$ are studied in detail following the lemma.

Lemma 3.3.9. Assume (A, C) is observable and let ν^c_{0k} be the dimension of R^c_{0gk} for some $1 \le k \le r$. There is an observable extension $(\tilde{A}, \tilde{C})$ of (A, C) onto:

$$\mathcal{S}^{\tilde{n}} = \mathcal{S}^{n} \oplus \mathcal{S}^{\nu^c_{0k}}$$

such that the excess subspace of $\tilde{\mathcal{R}}_{\mathcal{J}}$ for $(\tilde{A}, \tilde{C})$ satisfies:

$$\text{i)} \quad \tilde{A}\,\tilde{\mathcal{R}}_{0g} \subset \tilde{\mathcal{R}}_{0g} \oplus \tilde{g}_1 \oplus \ldots \tilde{g}_{k-1} \oplus \tilde{g}_{k+1} \oplus \ldots \tilde{g}_r$$

i.e. $\tilde{\mathcal{R}}_{0g}$ is not altered by the removal of f_k from $f_1 \ldots f_r$. In addition:

$$\text{ii)} \quad d(\tilde{\mathcal{R}}_{0g}) \leq d(\mathcal{R}_{0g})$$

$$\text{iii)} \quad d(\tilde{\bar{\mathcal{R}}}_k) = \nu_k + \nu_{0k}^c$$

$$\text{iv)} \quad d(\tilde{\bar{\mathcal{R}}}_j) \geq d(\bar{\mathcal{R}}_j) \qquad j \neq k$$

Proof

Let $\mathcal{S}^n$ be partitioned into four subspaces:

$$\mathcal{S}^n = \bar{\mathcal{R}}_k \oplus \mathcal{R}_{0gk}^c \oplus (\bar{\mathcal{R}}_k^D \oplus \mathcal{R}_{0gk}) \oplus \bar{\mathcal{R}}_{\mathcal{J}}^c$$

where $\bar{\mathcal{R}}_{\mathcal{J}}^c$ is as defined in the proof of Theorem 3.3.1. A is uniquely defined by the sixteen operators which can be defined on the partitions. Of particular interest are:

$$\text{i)}' \quad \Theta_k^c : \mathcal{R}_{0gk}^c \to \bar{\mathcal{R}}_k$$

$$\text{ii)}' \quad \Pi_k^c : \mathcal{R}_{0gk}^c \to \mathcal{R}_{0gk}^c$$

If the partion of $\mathcal{S}^n$ is included in the definition of $\tilde{A}$, twenty-five operators are necessary to uniquely specify A. This can be reduced to five by defining operators on larger subspaces when convenient:

$$\text{iii)}' \quad A : \mathcal{S}^n \to \mathcal{S}^n$$

$$\text{iv)}' \quad \phi : \mathcal{S}^n \to \mathcal{S}^{\nu_{0k}^c}$$

v)' ϕ : $\mathcal{B}^{\nu^c_{0k}} \rightarrow (\bar{\mathcal{R}}_k^D \oplus \mathcal{R}_{0gk}) \oplus \mathcal{R}^c_{0gk} \oplus \bar{\mathcal{R}}^c_{\mathcal{J}}$

vi)' $-\Theta^c_k$: $\mathcal{B}^{\nu^c_{0k}} \rightarrow \bar{\mathcal{R}}_k$

vii)' Π^c_k : $\mathcal{B}^{\nu^c_{0k}} \rightarrow \mathcal{B}^{\nu^c_{0k}}$

$\tilde{C}$ can readily be defined in terms of C by the two operators:

viii)' C : $\mathcal{B}^n \rightarrow \mathcal{B}^m$

ix)' ϕ : $\mathcal{B}^{\nu^c_{0k}} \rightarrow \mathcal{B}^m$

It remains to show that $\tilde{A}$ and $\tilde{C}$ as defined in iii)' - ix)' satisfy the hypothesis.

From v)' - vii)', the image of $\mathcal{B}^{\nu^c_{0k}}$ under $\tilde{A}$ satisfies:

x)' $\tilde{A}\, \mathcal{B}^{\nu^c_{0k}} \subset \mathcal{B}^{\nu^c_{0k}} \oplus \bar{\mathcal{R}}_k$

By construction $\mathcal{R}^c_{0gk}$ is observable (mod $\bar{\mathcal{R}}^D_k$) for (A, C). Comparison of i)' and ii)' with vi)' and vii)' then indicates $\mathcal{B}^{\nu^c_{0k}}$ is observable (mod $\bar{\mathcal{R}}^D_k$) also. Hence $\mathcal{B}^{\nu^c_{0k}}$ is observable. A similar argument proves that $\bar{\mathcal{R}}_k \oplus \mathcal{B}^{\nu^c_{0k}}$ is an observable subspace. Corollary 1 of Theorem 3.2.1 also implies that no event vector in $\bar{\mathcal{R}}_{\mathcal{J}} \oplus \mathcal{B}^{\nu^c_{0k}}$ which is not also in $\bar{\mathcal{R}}_k \oplus \mathcal{B}^{\nu^c_{0k}}$ can be constrained to generate unidirectional output along $\tilde{C}$ f. It then follows that $(\tilde{A}, \tilde{C})$ is an observable reference model and $\tilde{\bar{\mathcal{R}}}_{\mathcal{J}}$ can be defined by:

xi)' $\tilde{\bar{\mathcal{R}}}_{\mathcal{J}} = \bar{\mathcal{R}}_{\mathcal{J}} \oplus \mathcal{B}^{\nu^c_{0k}} = \tilde{\bar{\mathcal{R}}}_1 \oplus \ldots \tilde{\bar{\mathcal{R}}}_r \oplus \tilde{\mathcal{R}}_{0g}$

and $\tilde{\bar{\mathcal{R}}}_k$ satisfies:

xii)' $\tilde{\bar{\mathcal{R}}}_k = \bar{\mathcal{R}}_k \oplus \mathcal{B}^{\nu^c_{0k}}$

Therefore iii) is true. Similarly, since it is obvious that for all $j \neq k$:

$$\text{xiii)}' \qquad \bar{\mathcal{R}}_j \subset \tilde{\bar{\mathcal{R}}}_j$$

it follows that iv) is true and xi)' - xiii)' imply ii).

To prove i), assume first that:

$$\text{xiv)}' \qquad d(\tilde{\mathcal{R}}_{0g}) = d(\mathcal{R}_{0g})$$

R_{0gk} satisfies (3.3.24) with A replaced by $\tilde{A}$ and is a subspace of $\tilde{\mathcal{R}}_{0g}$. Because ii)' and vii)' are identical and vi)' is the negative of i)', x)' and (3.3.25) imply that there must also exist a ν_{0k}^{c} - dimensional subspace:

$$\mathcal{R}_{0gk}^{D} \subset \mathcal{R}_{0gk}^{c} \oplus \mathcal{S}^{\nu_{0k}^{c}}$$

whose image under $\tilde{A}$ does not have a component in $\tilde{\bar{\mathcal{R}}}_k$ i.e. an expression similar to (3.3.25) is:

$$\tilde{A}\mathcal{R}_{0gk}^{D} \subset \mathcal{R}_{0gk}^{D} \pmod{\tilde{\bar{\mathcal{R}}}_k^{D} \oplus \mathcal{R}_{0gk}}$$

If $\tilde{\mathcal{R}}_D$ is defined as the direct sum of $\mathcal{R}_{0gk}$ and $\mathcal{R}_{0gk}^{D}$:

$$\tilde{\mathcal{R}}_D = \mathcal{R}_{0gk}^{D} \oplus \mathcal{R}_{0gk}$$

then the above equations require $\tilde{\mathcal{R}}_D$ to be a ν_0 -dimensional subspace which satisfies Lemma 3.3.6 with g_k removed:

$$\text{xv)}' \qquad \tilde{A}\,\tilde{\mathcal{R}}_D \subset \tilde{\mathcal{R}}_D \oplus g_1 \oplus \cdots g_{k-1} \oplus g_{k+1} \oplus \cdots g_r$$

Equations xi)', xii)', and xiv)' require xiii)' to be an equality and therefore $g_1 \cdots g_{k-1}$, $g_{k+1} \cdots g_r$ are the detection generators

for $\tilde{\bar{\mathcal{R}}}_1 \ldots \tilde{\bar{\mathcal{R}}}_{k-1}$, $\tilde{\bar{\mathcal{R}}}_{k+1} \ldots \tilde{\bar{\mathcal{R}}}_r$. Lemma 3.3.6 then implies that $\tilde{\mathcal{R}}_{0g}$ is the unique ν_0-dimensional solution to xv)' which requires:

$$\tilde{\mathcal{R}}_{0g} = \tilde{\mathcal{R}}_D$$

To complete the proof of i) assume that x(v)' is not satisfied, i.e. that $\tilde{\mathcal{R}}_{0g}$ is smaller than $\mathcal{R}_{0g}$. For convenience, let:

$$\tilde{\nu}_0 = d(\tilde{\mathcal{R}}_{0g}) = d(\mathcal{R}_{0g}) - 1 = \nu_0 - 1$$

Equations xi)' - xiii)' then imply that for some $j \neq k$:

$$\tilde{\nu}_j = d(\tilde{\bar{\mathcal{R}}}_j) = d(\bar{\mathcal{R}}_j) + 1 = \nu_j + 1$$

and a new generator $\tilde{g}_j$ for that $\tilde{\bar{\mathcal{R}}}_j$ must be found. Because $\tilde{\mathcal{R}}_D$ is a ν_0-dimensional subspace which satisfies both xv)' and:

$$\tilde{\bar{\mathcal{R}}}_{\mathcal{J}} = \tilde{\bar{\mathcal{R}}}_k \oplus \bar{\mathcal{R}}_k^D \oplus \tilde{\mathcal{R}}_D$$

it follows from Lemma 3.2.4 that $\tilde{\mathcal{R}}_D$ must contain $\tilde{g}_j$. $\tilde{\mathcal{R}}_D$ can then be partitioned:

xvi)' $$\tilde{\mathcal{R}}_D = \tilde{\mathcal{R}}_0 \oplus \tilde{g}_j$$

where $\tilde{\mathcal{R}}_0$ is some arbitrary excess subspace of $\tilde{\bar{\mathcal{R}}}_{\mathcal{J}}$. $\tilde{\mathcal{R}}_{0g}$ which satisfies:

$$\tilde{A}\,\tilde{\mathcal{R}}_{0g} \subset \tilde{\mathcal{R}}_{0g} \oplus g_1 \oplus \ldots g_{j-1} \oplus \tilde{g}_j \oplus g_{j+1} \oplus \ldots g_{k-1} \oplus g_{k+1} \oplus \ldots g_r$$

can be found by applying the construction of Lemma 3.3.6 to $\tilde{\mathcal{R}}_0$. This argument is generalizable for all $d(\tilde{\mathcal{R}}_{0g}) \leq d(\mathcal{R}_{0g})$ and i) is proved.

Δ

The independence of $\tilde{\mathcal{R}}_{0g}$ from the presence of f_k in $f_1 \ldots f_r$ is in marked contrast to Lemma 3.3.6. The fundamental importance of this change can be seen if the procedure is repeated sequentially for $f_1 \ldots f_r$. If the reference model after the r^{th} application of Lemma 3.3.9 is denoted by $(\tilde{\tilde{A}}, \tilde{\tilde{C}})$, the corresponding $\tilde{\tilde{\mathcal{R}}}_{0g}$ must be completely independent of every vector in $f_1 \ldots f_r$:

$$\tilde{\tilde{A}}\, \tilde{\tilde{\mathcal{R}}}_{0g} \subset \tilde{\tilde{\mathcal{R}}}_{0g} \tag{3.3.46}$$

Lemma 3.3.9 constrains $(\tilde{A}, \tilde{C})$, and hence $(\tilde{\tilde{A}}, \tilde{\tilde{C}})$, to be observable if the original reference model (A, C) is observable. Because Definition 3.3.4 requires $\tilde{\tilde{\mathcal{R}}}_{0g}$ to be a subspace of $\eta(\tilde{\tilde{C}})$ and (3.3.46) states that $\tilde{\tilde{\mathcal{R}}}_{0g}$ is invariant, observability then requires that:

$$\tilde{\tilde{\mathcal{R}}}_{0g} = \phi \tag{3.3.47}$$

It follows that $f_1 \ldots f_r$ are mutually detectable for the augmented reference model $(\tilde{\tilde{A}}, \tilde{\tilde{C}})$. Theorem 3.3.4 provides a rigorous proof of this result:

Theorem 3.3.4 Assume $f_1 \ldots f_r$ are output separable, but not mutually detectable for (A, C). The minimal equivalent representation $(\tilde{\tilde{A}}, \tilde{\tilde{C}})$ for which $f_1 \ldots f_r$ are mutually detectable is of order:

$$\tilde{\tilde{n}} = n + \sum_{k=1,r} \nu^c_{0k} - \nu_0$$

Proof

Let Lemma 3.3.9 be applied for f_k. The similarity of xv)′ in the proof of Lemma 3.3.9 to Lemma 3.3.6 suggests that, as in (3.3.23), there is a partition of $\tilde{\mathcal{R}}_D$:

i) $\tilde{\mathcal{R}}_D = \tilde{\mathcal{R}}^c_{Dj} \oplus \mathcal{R}_{Dj} \qquad j \neq k$

for which $\tilde{\mathcal{R}}^c_{Dj}$ is a ν^c_{0j}-dimensional subspace of $\tilde{\mathcal{R}}_D$ and is the largest such subspace which is observable (mod $\bar{\mathcal{R}}^D_j$):

$$A\tilde{\mathcal{R}}^c_{Dj} \subset \tilde{\mathcal{R}}^c_{Dj} \oplus g_j \quad (\text{mod } \bar{\mathcal{R}}^D_j \oplus \mathcal{R}_{Dj}) \quad j \neq k$$

If some $\tilde{\bar{\mathcal{R}}}_j$, $j \neq k$, for $(\tilde{A}, \tilde{C})$ has a greater dimension than $\bar{\mathcal{R}}_j$ for (A, C), i.e. if $\tilde{\nu}_j > \nu_j$, the proof of i) in Lemma 3.3.9 indicates that there are $\tilde{\nu}_j - \nu_j$ independent vectors in $\tilde{\mathcal{R}}_D$ (hence in $\tilde{\mathcal{R}}^c_{Dj}$) which are vectors of $\tilde{\bar{\mathcal{R}}}_j$. If $\tilde{\mathcal{R}}^c_{0g1} \ldots \tilde{\mathcal{R}}^c_{0gr}$ are defined for $\tilde{\mathcal{R}}_{0g}$ similarly to $\mathcal{R}^c_{0g1} \ldots \mathcal{R}^c_{0gr}$ for $\mathcal{R}_{0g}$; a generalization of the argument leading to xvi)' implies that $\tilde{\nu}^c_{0j}$, the dimension of $\mathcal{R}^c_{0gj}$, must be smaller than ν^c_{0j}. In fact, the difference must be equal to $\tilde{\nu}_j - \nu_j$:

ii) $$\tilde{\nu}^c_{0j} = \nu^c_{0j} - [\tilde{\nu}_j - \nu_j] \qquad j \neq k$$

$$\tilde{\nu}^c_{0k} = 0$$

Equations xi)' - xiii)' of Lemma 3.3.9 also imply that if $d(\tilde{\bar{\mathcal{R}}}_j) > d(\bar{\mathcal{R}}_j)$, then $d(\tilde{\mathcal{R}}_{0g})$ must be correspondingly smaller than $d(\mathcal{R}_{0g})$:

$$\nu_0 - \tilde{\nu}_0 = \sum_{j \neq k} [\tilde{\nu}_j - \nu_j]$$

It follows from ii) that this difference can be related to the dimensions of $\tilde{\mathcal{R}}^c_{0g1} \ldots \tilde{\mathcal{R}}^c_{0gr}$:

iii) $$\sum_{j \neq k} \tilde{\nu}^c_{0j} = \sum_{j \neq k} \nu^c_{0j} - [\nu_0 - \tilde{\nu}_0]$$

If Lemma 3.3.9 is applied sequentially to $\tilde{\bar{\mathcal{R}}}_j$ and so on, the final reference model $(\tilde{\tilde{A}}, \tilde{\tilde{C}},)$ after r applications of the lemma has a dimension:

$$\tilde{\tilde{n}} = n + \nu^c_{0k} + \tilde{\nu}^c_{0j} + \ldots$$

which from iii) can be written as:

$$\text{iv)} \qquad \tilde{\tilde{n}} = \sum_{i=1,r} \nu_{0i}^{c} - [\nu_0 - \tilde{\tilde{\nu}}_0]$$

But $\tilde{\tilde{\mathcal{R}}}_0'$ is invariant since the image under $\tilde{\tilde{A}}$ of $\tilde{\tilde{\mathcal{R}}}_0$ must be contained in $\tilde{\tilde{\mathcal{R}}}_0$:

$$\tilde{\tilde{A}}\,\tilde{\tilde{\mathcal{R}}}_0 \subset \tilde{\tilde{\mathcal{R}}}_0$$

Because $(\tilde{\tilde{A}}, \tilde{\tilde{C}})$ is observable, this implies that $\tilde{\tilde{\mathcal{R}}}_0 = \phi$ and iv) becomes the desired result.

Δ

If $f_1 \ldots f_r$ are not mutually detectable for (A, C), Theorem 3.3.4 proves that sequential application of Lemma 3.3.9 for $f_1 \ldots f_r$ yields a $\tilde{\tilde{n}}$-dimensional reference model for which $f_1 \ldots f_r$ are mutually detectable. The remaining task is to formulate algorithms for implementing the results. This requires Lemma 3.3.9 to be restated in a matrix format. The following two lemmas establish the necessary results. Lemma 3.3.10 defines the reference model $(\tilde{A}, \tilde{C})$ and $\tilde{\tilde{R}}_k$. The matrix representation of the operators Π_k^c: $\mathcal{S}^{\nu_{0k}^c} \to_c \mathcal{S}^{\nu_{0k}^c}$ and $-\Theta_k^c : \mathcal{S}^{\nu_{0k}^c} \to \bar{\mathcal{R}}_k$ depends upon the basis chosen for $\mathcal{S}^{\nu_{0k}^c}$. (Note that the matrix equivalent of Θ_k^c is the k^{th} row θ_k of Θ). The construction of $(\tilde{A}, \tilde{C})$ in Lemma 3.3.10 is made trivial by choice of an appropriate basis.

Lemma 3.3.10 Assume (A, C) is observable and $f_1 \ldots f_r$ are not mutually detectable. If $[\Delta : \beta]_k$ is the transformation matrix resulting from the application of Corollary 1 of Theorem 3.3.3 for f_k, a reference model $(\tilde{A}, \tilde{B}, \tilde{C})$ which satisfies Lemma 3.3.9 for f_k and is input-output equivalent to (A, B, C) is the $(n+\nu_{0k}^c)$-dimensional model:

$$\tilde{C} = [C : 0] \qquad \tilde{A} = \left[\begin{array}{c|c} A & -g_k \theta_k \Delta \\ \hline 0 & \Pi_k^c \end{array}\right] \qquad \tilde{B} = \left[\begin{array}{c} B \\ \hline 0 \end{array}\right]$$

The range space of the matrix $\tilde{\bar{R}}_k$ is the detection space $\tilde{\bar{\mathcal{R}}}_k$:

$$\tilde{\bar{R}}_k = \left[\begin{array}{c|c} \bar{R}_k & 0 \\ \hline 0 & E_{\nu_{0k}^c} \end{array}\right]$$

where $E_{\nu_{0k}^c}$ is the ν_{0k}^c-dimensional identity matrix.

Proof

Input-output equivalence follows from (2.4.2)-(2.4.3). The most convenient choice of bases is to let $\mathcal{S}^n$ be spanned by the unit vectors in the columns of $\left[\begin{array}{c} E_n \\ \hline 0 \end{array}\right]$ and $\mathcal{S}^{\nu_{0k}^c}$ by those of $\left[\begin{array}{c} 0 \\ \hline E_{\nu_{0k}^c} \end{array}\right]$. The choice of $\tilde{\bar{R}}_k$ follows immediately from xii)' of Lemma 3.3.9.

From Corollary 1 of Theorem 3.3.3, R_{0gk}^c satisfies:

$$\text{i)} \qquad A R_{0gk}^c = R_{0gk}^c \Pi_k^c + R_{0gk} \Pi_{k0}^c + G_k \Theta_k \Delta + g_k \theta_k \Delta$$

Lemma 3.3.9 is concerned only with the operators $\Pi_k^c : \mathcal{R}_{0gk}^c \to \mathcal{R}_{0gk}^c$ and $\Theta_k^c : \mathcal{R}_{0gk}^c \to \bar{\mathcal{R}}_k$ and, therefore, only the first and last terms of the above are important. Let the $(n + \nu_{0k}^c) \times (n + \nu_{0k}^c)$ matrix $\tilde{A}$ be partitioned consistently with the bases for $\mathcal{S}^n$ and $\mathcal{S}^{\nu_{0k}^c}$:

$$\tilde{A} = \left[\begin{array}{c|c} \tilde{A}_{11} & \tilde{A}_{12} \\ \hline \tilde{A}_{21} & \tilde{A}_{22} \end{array}\right]$$

It is immediately obvious from iii)' and iv)' of Lemma 3.3.9 that $\tilde{A}_{11}$ and $\tilde{A}_{21}$ must satisfy:

$$\tilde{A}_{11} = A$$

$$\tilde{A}_{21} = 0$$

The matrix version of the image of $\mathcal{S}^{\nu^c_{0k}}$ under $\tilde{A}$ is:

$$\tilde{A}\begin{bmatrix} 0 \\ \hline E_{\nu^c_{0k}} \end{bmatrix} = \begin{bmatrix} \tilde{A}_{12}\, E_{\nu^c_{0k}} \\ \hline \tilde{A}_{22}\, E_{\nu^c_{0k}} \end{bmatrix}$$

Because $E_{\nu^c_{0k}}$ is the identity matrix, this can be rewritten in a form similar to i):

$$\tilde{A}\begin{bmatrix} 0 \\ \hline E_{\nu^c_{0k}} \end{bmatrix} = \begin{bmatrix} \tilde{A}_{12} \\ \hline E_{\nu^c_{0k}}\tilde{A}_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ \hline E_{\nu^c_{0k}} \end{bmatrix}\tilde{A}_{22} + \begin{bmatrix} \tilde{A}_{12} \\ \hline 0 \end{bmatrix}$$

from which it is apparent that Lemma 3.3.9 requires:

$$\tilde{A}_{22} = \Pi^c_k$$

$$\tilde{A}_{12} = -g_k\, \theta_k\, \Delta$$

Δ

As mentioned in Theorem 3.3.4, it may be necessary to apply Lemma 3.3.9 repeatedly to remove the excess subspace of $\bar{\mathcal{R}}_j$. Before continuing to the second iteration it is necessary to determine $\tilde{\bar{\mathcal{R}}}_j$, $j \neq k$, and $\tilde{\mathcal{R}}_{0g}$. These could be found using Lemmas 3.2.6 and 3.3.7 respectively, but a computationally simpler approach can be based upon $(\tilde{A}, \tilde{C})$ and the known spaces $\bar{\mathcal{R}}_j$, $j \neq k$, and $\mathcal{R}_{0g}$. An algorithm for $\tilde{R}_{0g}$ is given in Appendix A.3; Lemma 3.3.11 supplies an algorithm for the detection

spaces which utilizes techniques discussed in Appendix A.2.
Both algorithms require the matrix $\tilde{R}_D$ from the proof of Lemma 3.3.9.

Lemma 3.3.11 If Lemma 3.3.10 is applied for f_k, a basis for $\bar{\tilde{\mathcal{R}}}_j$, $j \neq k$, is the range space of:

$$\bar{\tilde{R}}_j = \left[\begin{array}{c|c} \bar{R}_j & \\ \hline 0 & \tilde{R}_D Z_j \end{array}\right]$$

where $\tilde{R}_D$ is:

$$\tilde{R}_D = \left[\begin{array}{c|c} R^c_{0gk} & R_{0gk} \\ \hline E^c_{\nu_{0k}} & 0 \end{array}\right] = \left[\begin{array}{c|c} R_{0g}\,\Delta & R_{0g}\,\beta \\ \hline E^c_{\nu_{0k}} & 0 \end{array}\right]$$

and the columns of Z_j span the null space of:

$$M_{0Lj} = \left[\begin{array}{l} (\Theta_k)_j\,[\Delta:\ \beta]_k \\ (\Theta_k)_j\,[\Delta:\ \beta]_k \Pi^0_k \\ \vdots \\ (\Theta_k)_j\,[\Delta:\ \beta]_k (\Pi^0_k)^{\nu_0 - 1} \end{array}\right]$$

$(\Theta_k)_j$ is Θ with the k^{th} and j^{th} rows removed and $[\Delta:\beta]_k$ and Π^0_k are defined in Corollary 1 of Theorem 3.3.3.

Proof

It is simple to show that $\tilde{R}_D$ corresponds to $\tilde{\mathcal{R}}_D$ in the proof of Lemma 3.3.9. Using $\tilde{A}$ from Lemma 3.3.10, $\tilde{A}\,\tilde{R}_D$ is:

$$\tilde{A}\,\tilde{R}_D = \left[\begin{array}{c|c} A\,R^c_{0gk} - g_k\,\theta_k\,\Delta\,E^c_{\nu_{0k}} & A\,R_{0gk} \\ \hline \Pi^c_k\ E^c_{\nu_{0k}} & 0 \end{array}\right]$$

Corollary 1 of Theorem 3.3.3 can be used to evaluate this matrix. Cancellation of the terms in $g_k\ \theta_k\ \Delta$ yields:

$$\tilde{A}\tilde{R}_D = \left[\begin{array}{c|c} R^c_{0gk}\ \Pi^c_k + R_{0gk}\ \Pi^c_{k0} & R_{0gk}\ \Pi_k \\ \hline E_{\nu_{0k}}{}^c\ \Pi^c_k & 0 \end{array}\right] + \left[\begin{array}{c} G_k \Theta_k \\ \hline 0 \end{array}\right] [\Delta : \beta]_k$$

The definition of Π^0_k allows this to be stated as the matrix equivalent of xv)' from Lemma 3.3.9:

$$\tilde{A}\ \tilde{R}_D = \tilde{R}_D \Pi^0_k + \left[\begin{array}{c} G_k \Theta_k \\ \hline 0 \end{array}\right] [\Delta : \beta]_k$$

The definition of $\tilde{\tilde{R}}_j$ in terms of $\tilde{R}_D\ Z_j$ is similar to the definition of R_{0gk} in terms of β. Recall that $(\tilde{A}, \tilde{C})$ is observable. The image of $\tilde{R}_D\ Z_j$ under $\tilde{A}$ satisfies

$$\text{i)}\quad \tilde{A}\tilde{R}_D Z_j = \tilde{R}_D \Pi^0_k Z_j + \left[\begin{array}{c} g_j \beta_j + (G_k)_j (\Theta_k)_j \\ \hline 0 \end{array}\right] [\Delta : \beta]_k Z_j$$

If the range space of $\tilde{R}_D\ Z_j$ is to be in the range space of $\tilde{\tilde{R}}$, an argument similar to that used in the proof of Theorem 3.3.3 requires i) to reduce to:

$$\tilde{A}\ \tilde{R}_D Z_j = \tilde{R}_D \Pi_k{}^0 Z_j + g_j \theta_j [\Delta : \beta]_k Z_j$$

Since $rk(G_k)_j = r-2$, this requires Z_j to satisfy:

$$(\Theta_k)_j [\Delta : \beta]_k Z_j = 0$$

Continuation of this argument for $\tilde{A}^2 \tilde{R}_D Z_j \ldots \tilde{A}^{\nu_0 - 1} \tilde{R}_D Z_j$ yields the desired result.

Δ

3.4 Output Stationarity

Section 3.3 guarantees that if the event vectors $f_1 \ldots f_m$ are output separable, there exists a lowest order state space augmentation such that $f_1 \ldots f_m$ are mutually detectable. One of the consequences of Theorem 3.4.1 is that no larger collection of event vectors can be detected by a single detection filter. Proof of this limitation is deferred until Section 3.5.

Adherence to the principles underlying Definition 2.1.1 implies that only m directions in the output space of the system model can be associated uniquely with potential failures in the system. The only reasonable manner in which Definition 2.1.1 could be relaxed would be to relinquish at least partial control over the assignability of the eigenvalues of A-DC. Section 3.3 supplies no ready indication that this maneuver would increase the number of failures which could be monitored by a single detection system. The possibilities are studied in detail in this section. As in Section 3.3, the theory is developed first in a linear operator context and then in a matrix context.

For a particular D selected for $f_1 \ldots f_m$, there may be a set of event vectors $h_1 \ldots h_q$ in addition to $f_1 \ldots f_m$ which produce unidirectional output. The event vectors $f_1 \ldots f_m$, $h_1 \ldots h_q$ are not mutually detectable since their output is unidirectional only if D satisfies a set of constraints in addition to those required to make $f_1 \ldots f_m$ mutually detectable. It is shown in this section that the minimal cost associated with expanding the set of vectors to $f_1 \ldots f_m$, $h_1 \ldots h_q$ is a loss of freedom to assign the eigenvalues of each of $\bar{R}_1 \ldots \bar{R}_m$ independently of the remaining detection spaces. If a set of eigenvalues Λ_1 is to be associated with $\bar{R}_1$, the same set must be associated with each of $\bar{R}_2 \ldots \bar{R}_m$, i.e. the assignable eigenvalues of A-DC must be given a multiplicity of m. If the dimensions of $\bar{R}_1 \ldots \bar{R}_m$ differ or if $h_1 \ldots h_q$ are poorly chosen, some eigenvalues of A-DC may become unassignable.

Definition 3.4.1 Assume $f_1 \ldots f_m$ are output separable and let $\mathbf{D}$ be the class of operators such that for every $\mathbf{D} \in \mathbf{D}^*$ the output generated by each of $f_1 \ldots f_m$ with respect to $(\mathbf{A}-\mathbf{DC}, \mathbf{C})$ is unidirectional. The subspace $\bar{\mathcal{S}}$ can be made output stationary with $f_1 \ldots f_m$ if there exists a $\mathbf{D}' \in \mathbf{D}^*$ such that the output from every element $\xi_i \in \bar{\mathcal{S}}$ for $(\mathbf{A}-\mathbf{D}'\mathbf{C}, \mathbf{C})$ is unidirectional along $\mathbf{C}\xi_i$ (or $\mathbf{C}\mathbf{A}\xi_i$ if $\mathbf{C}\xi_i = \emptyset$ etc.).

Output stationarity is an attempt to use the last degree of freedom left in the choice of $\mathbf{D}$--the freedom to choose the eigenvalues of $\mathbf{A}-\mathbf{DC}$--to obtain failure information about events other than $f_1 \ldots f_m$ from the failure detection system. The method may be quite successful. It is possible to determine a class $\bar{\mathcal{S}}^*$ (actually, an m-dimensional subspace) of detection spaces which are invariant with respect to $\mathbf{A}-\mathbf{DC}$ and such that with each output direction in $\mathcal{E}^m$ there is a unique detection space in $\bar{\mathcal{S}}^*$ which produces unidirectional output in that direction. If a unidirectional error signal $\epsilon'(t)$ appears in the detection system output, it can uniquely be associated with the appropriate detection space.

The potential value of output stationarity is dependent upon the freedom available to specify $\bar{\mathcal{S}}^*$. Unfortunately, this freedom varies with the individual system. A consequence of Theorem 3.4.1 is that the detection spaces corresponding to certain event vectors h' cannot be elements of $\bar{\mathcal{S}}^*$. Such events are incompatible with $f_1 \ldots f_m$. Theorem 3.4.1 offers a simple test for compatibility. If the event vectors $h_1 \ldots h_q$ are individually compatible with $f_1 \ldots f_m$, Lemma 3.4.4 provides a test for mutual compatibility. There is no maximum number of event vectors $h_1 \ldots h_q$ which can be chosen to have their detection spaces in $\bar{\mathcal{S}}^*$. The number depends strictly upon the nature of the system and of the potential failures. As a rule-of-thumb, the more elements $h_1 \ldots h_q$ specified, the smaller the range of acceptable $\mathbf{D} \in \mathbf{D}^*$ and the less freedom to assign the eigenvalues of $\mathbf{A}-\mathbf{DC}$. Once the $h_1 \ldots h_q$ are chosen, the remaining elements of $\bar{\mathcal{S}}^*$ are automatically fixed. In general, the remaining detection spaces in $\bar{\mathcal{S}}^*$ correspond to physically meaningless failures and may be disregarded.

The systems with m = 1 and m = n are degenerate with respect to output stationarity. In the first case $\bar{R}_1 = \mathcal{S}^n$ and output stationarity is meaningless. In the latter $\bar{R}_i = f_i$ for $i = 1 \ldots m$ and it follows trivially that $\mathcal{S}^n$ can be made output stationary. This example was used by Beard (1971) to introduce the Detection Theorem and is summarized in Section 4.5. In the following it is assumed that $\bar{R}_1 \ldots \bar{R}_m$ may be of arbitrary dimension. Conditions under which a single detection space $\bar{\mathcal{J}}$ can be made output stationary with $f_1 \ldots f_m$ are determined in Section 3.4.1 and generalized to $\bar{\mathcal{J}}^*$ in Section 3.4.3. A simple example of an application of output stationarity is given in Section 3.4.2.

The conclusions of this section are stated under the assumption that (A, C) is observable and that $f_1 \ldots f_m$ are not mutually detectable. The results are equally valid if $f_1 \ldots f_m$ are not mutually detectable and such occurrences are treated as corollaries to the principal theorems. It is also possible to consider a basic set of events $f_1 \ldots f_r$ with $r < m$; however, a full complement of m event vectors is the most general case. Treatment of the set $f_1 \ldots f_r$ follows trivially from the stated theorems.

3.4.1 Addition of a Stationary Space to a Detectable Set

Based upon the importance which detection spaces play in determining the mutual detectability of a set of event vectors, it should not come as a surprise that they also assume a dominant role in the solution of the output stationarity problem. The basic problem can still be posed in terms of event vectors. In design studies the question arises as to whether or not an event vector h_i can be made output stationary with $f_1 \ldots f_m$. This question can be answered without consideration of detection spaces; however, the determination of which $\mathbb{D} \in \mathbb{D}^*$ make the failure associated with h_i generate unidirectional output for A-DC requires knowledge of $\bar{\mathcal{J}}_i$, the detection space of h_i. Detection spaces are also necessary to determine the elements of $\bar{\mathcal{J}}^*$ (and their corresponding event vectors) which are output stationary with $f_1 \ldots f_m$ when h_i is output stationary.

The over-riding importance of the detection space is proved in Lemma 3.4.1. The lemma is no more than a restatement of Corollary 2 of Theorem 3.2.1 in terms of output stationarity: If for some $D \in D^*$, the failure associated with h_i produces unidirectional output along Ch_i, then so does every $\rho_i \in \bar{\mathcal{P}}_i$ and vice versa. In this section detection spaces are denoted by both $\bar{R}_i$ and $\bar{\mathcal{P}}_i$. $\bar{R}_1 \ldots \bar{R}_m$ are used exclusively to denote the detection space for $f_1 \ldots f_m$; $\bar{\mathcal{P}}_i$ is the detection space for an arbitrary event vector h_i which is not an element of one of $\bar{R}_1 \ldots \bar{R}_m$. $\mathcal{P}_i$ and $\bar{\mathcal{P}}_i$ are defined for h_i by Definition 3.1.2 just as R_i and $\bar{R}_i$ are defined for f_i. The minimal polynomials of $\bar{R}_i$ and $\bar{\mathcal{P}}_i$ are denoted by $\Psi(\cdot)$ and $\Upsilon(\cdot)$ respectively.

It is important to note that the set $\{\rho_i\}$ where $\rho_i \in \bar{\mathcal{P}}_i$ is not defined to be output stationary equivalent with h_i. The set $\{\rho_i\}$ is detection equivalent to h_i as was proved in Section 3.2. Output stationary equivalence is defined in Section 3.4.3 as an operator on detection spaces and is a much broader concept than detection equivalence.

Lemma 3.4.1 Let $\bar{\mathcal{P}}_i$ be the detection space for an event vector h_i and assume both $\bar{\mathcal{P}}_i$ and h_i can be made output stationary with $f_1 \ldots f_m$. If $\bar{D}_i^*$ and D_i^* are the subclasses of D^* for which $\bar{\mathcal{P}}_i$ and h_i respectively are output stationary with $f_1 \ldots f_m$, then:

$$\bar{D}_i^* = D_i^*$$

Proof

Since $h_i \in \bar{\mathcal{P}}_i$, it necessarily follows that $\bar{D}_i^* \subset D_i^*$. By Corollary 2 of Theorem 3.2.1 every $\rho_i \in \bar{\mathcal{P}}_i$ is detection equivalent to h_i. Therefore, every $D \in D_i^*$ satisfies $D \in \bar{D}_i^*$ and $D_i^* \subset \bar{D}_i^*$. Along with the prior inequality, this result completes the proof. Δ

Assume that $f_1 \ldots f_m$ are mutually detectable and it is desirable to determine whether or not the vectors $h_1 \ldots h_q$ can be made output stationary with $f_1 \ldots f_m$. Lemma 3.4.1 states

that an equivalent test is to determine whether or not the respective detection spaces $\bar{\mathcal{S}}_1 \ldots \bar{\mathcal{S}}_q$ can be made output stationary with $f_1 \ldots f_m$; however, the lemma does not suggest a procedure for implementing the test. Theorem 3.4.1 provides necessary and sufficient conditions which $\bar{\mathcal{S}}_i$ must satisfy in order for $\bar{D}_i^*$ to be nonempty. The prescribed test is simple to implement, but from the statement of the theorem it is clear that in general there are many vectors and subspaces which cannot be made output stationary with $f_1 \ldots f_m$. The theorem should properly be stated in terms of index sets on $\{1, 2, \ldots m\}$. As in previous sections, if any of $Cf_1 \ldots Cf_m$, Ch_i are zero, they can be replaced by the first nonzero element of the sequence CAf_1, $CA^2f_1 \ldots$ etc.

Theorem 3.4.1 Assume $f_1 \ldots f_m$ are mutually detectable, $\bar{\mathcal{S}}_i$ is the detection space for the event vector h_i, and $f_1 \ldots f_\ell$ is the smallest subset of $f_1 \ldots f_m$ such that:

$$\bar{\mathcal{S}}_i \subset \bar{R}_1 \oplus \ldots \bar{R}_\ell$$

Then $\bar{\mathcal{S}}_i$ (and h_i) can be made output stationary with $f_1 \ldots f_m$ if and only if

$$C\bar{\mathcal{S}}_i = \alpha_1 C\bar{R}_1 + \ldots \alpha_\ell C\bar{R}_\ell$$

and all $\alpha_1 \ldots \alpha_\ell$ are nonzero. Assuming $Cf_1 \ldots Cf_\ell$, Ch_i are all

$$\beta_i Ch_i = \alpha_1 Cf_1 + \ldots \alpha_\ell Cf_\ell$$

for the same $\alpha_1 \ldots \alpha_r$, where β_i is a scaling factor for h_i.

Proof

Necessity Assume $\alpha_j = 0$ for some $1 \leq j \leq \ell$. The hypothesis then implies that the output direction:

$$C\bar{\mathcal{S}}_i = \sum_{k \neq j} \alpha_k C \bar{R}_k$$

is associated with $\bar{\mathcal{S}}_i$ (and h_i). From Lemma 3.2.2, for $\bar{R}_1 \ldots \bar{R}_\ell$ to be invariant requires that $D \in \mathbb{D}^*$ satisfy:

$$(A - DC)\,\bar{R}_k \subset \bar{R}_k \qquad k = 1 \ldots \ell$$

Choose $\rho_i \in \bar{\mathcal{S}}_i$ such that ρ_i may be expressed:

$$\rho_i = \sum_{k=1,\ell} \delta_k \xi_k$$

where $\xi_k \in \bar{R}_k$, $k = 1 \ldots \ell$, and $\delta_k \neq 0$. Then since $\bar{R}_j$ is observable and $\bar{R}_1 \ldots \bar{R}_\ell$ are invariant, there is some smallest integer μ_j such that:

$$(A-DC)^{\mu_j} \rho_i = (A-DC)^{\mu_j} \xi_j + \ldots = \delta_j C f_j + \ldots$$

for some $\delta_j \neq 0$ and for all $D \in \mathbb{D}^*$. But since $\alpha_j = 0$, this is a second output direction associated with $\bar{\mathcal{S}}_i$ and therefore $\bar{\mathcal{S}}_i$ cannot be made output stationary with $f_1 \ldots f_m$. By Lemma 3.4.1, then h_i cannot be made output stationary with $f_1 \ldots f_m$.

Sufficiency Assume $D \in \mathbb{D}^*$ is to be chosen such that $\bar{\mathcal{S}}_i$ is invariant with respect to $A-DC$. From the hypothesis, h_i can be expressed uniquely by:

$$\beta_i h_i = \alpha_1 f_1 + \ldots \alpha_\ell f_\ell + \xi_{hi}$$

where $\xi_{hi} \in \mathcal{R}_1 \oplus \ldots \mathcal{R}_\ell \subset \eta(C)$. If $D \in \mathbb{D}^*$ can be chosen such that:

i) $$(A-DC)\,\beta_i h_i = \xi_{\mathcal{S}i}$$

where $\xi_{\mathcal{J}i} \in \bar{\mathcal{J}}_i$, then Lemma 3.2.2 states that $\bar{\mathcal{J}}_i$ is invariant with respect to A-DC, and therefore is output stationary with $f_1 \ldots f_m$. To specify this D define:

$$h_{i\mathcal{R}} \equiv A\,\xi_{hi} \pmod{\mathcal{J}_i}$$

Clearly $h_{i\mathcal{R}} \in \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_\ell$. Theorem 3.3.1 insures that a D exists such that:

$$(A\text{-}DC) f_k = \xi_k \quad k = 1 \ldots \ell$$

for all $\xi_k \in \bar{\mathcal{R}}_k$. To satisfy i), these $\xi_1 \ldots \xi_\ell$ must be solutions of:

$$\text{ii)} \quad (A\text{-}DC)\,\beta_i h_i = \alpha_1 \xi_1 + \ldots \alpha_\ell \xi_\ell + h_{i\mathcal{R}} = \xi_{\mathcal{J}i}$$

But since $h_{i\mathcal{R}}, \xi_{\mathcal{J}i} \in \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_\ell$ and the $\alpha_1 \ldots \alpha_\ell$ are nonzero, a solution is guaranteed and $\bar{\mathcal{J}}_i$ can be made output stationary with $f_1 \ldots f_\ell$. Δ

The proof of Theorem 3.4.1 stipulates that if $\bar{\mathcal{J}}_i$ is to be made output stationary with $f_1 \ldots f_m$, h_i must be expressible in the form:

$$\beta_i h_i = \alpha_1 f_1 + \ldots \alpha_\ell f_\ell + \xi_{hi} \tag{3.4.1}$$

where $\xi_{hi} \in \mathcal{R}_1 \oplus \ldots \mathcal{R}_\ell \subset \eta(C)$ and the $\alpha_1 \ldots \alpha_\ell$ are all nonzero. The vector ξ_{hi} was not present when the Detection Theorem was derived in Section 3.2. It is present in (3.4.1) because Definition 3.4.1 along with Lemma 3.4.1, requires that $\bar{\mathcal{J}}_i$ be made invariant for A-DC only under the additional constraint that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_m$ also be invariant. Thus $\bar{\mathcal{J}}_i$ must be an invariant subspace which is contained in the direct sum of the invariant subspaces $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_\ell$. The vector $\xi_{hi} \pmod{\mathcal{J}_i}$ in essence defines

the relationship of $\bar{\mathscr{S}}_i$ to the component subspaces $\bar{R}_1 \ldots \bar{R}_m$. It is shown in Section 3.4.3 that ξ_{hi} and:

$$h_{i\mathscr{R}} \equiv A\xi_{hi} \pmod{\mathscr{S}_i} \tag{3.4.2}$$

are the critical factors in determining the mutual output stationarity of an arbitrary set of subspaces $\bar{\mathscr{S}}_1 \ldots \bar{\mathscr{S}}_q$ with $f_1 \ldots f_m$. Before studying the implications of Theorem 3.4.1, it is instructive to define the test in matrix terms. The test defined in Corollary 1 reverses the test in Theorem 3.4.1 in that it first finds the nonzero $\alpha_1 \ldots \alpha_\ell$ in the expression for Ch_i, and then determines whether or not $\bar{\mathscr{S}}_i$ is contained in the direct sum of the corresponding detection spaces. The discussions in Section 2.3.3 verify that Corollary 1 is the matrix equivalent of Theorem 3.4.1.

Corollary 1 Assume $f_1 \ldots f_\ell$ are the smallest subset of $f_1 \ldots f_m$ such that:

$$CF_L^c \alpha_L^c = \beta_i \, Ch_i$$

has a solution where:

$$\alpha_L^c = [\alpha_1 \ldots \alpha_\ell]^T$$

$$F_L^c = [f_1 \ldots f_\ell]$$

The event vector h_i can be made output stationary with $f_1 \ldots f_m$ if and only if there is a solution to:

$$\bar{R}_L^c \, \Xi = \bar{S}_i$$

where:

$$\bar{R}_L^c = [\bar{R}_i : \ldots : \bar{R}_\ell]$$

and $\bar{R}_1 \ldots \bar{R}_\ell$, $\bar{S}_i$ are matrix equivalents of the appropriate detection space S. If a Ξ exists, it satisfies:

$$\Xi = (\bar{R}_L^{c\,T} \bar{R}_L^c)^{-1} \bar{R}_L^{c\,T} \bar{S}_i$$

If $f_1 \ldots f_m$ are output separable, but not mutually detectable, Theorem 3.4.1 must be modified. The modifications include a change in the test of $\bar{\mathcal{J}}_i$ plus a test on $h_{i\mathcal{R}}$ which was not required before. Corollary 2 summarizes the general case.

Corollary 2 Assume $f_1 \ldots f_m$ are output separable. Let $\bar{\mathcal{R}}^c_{\mathcal{J}L}$ be the detection space of $\mathcal{J}^c_L$ where $\mathcal{J}^c_L$ is spanned by $f_1 \ldots f_\ell$, and let $f_1 \ldots f_\ell$ be the smallest subset of $f_1 \ldots f_m$ such that:

$$\bar{\mathcal{J}}_i \subset \bar{\mathcal{R}}^c_{\mathcal{J}L}$$

Then $\bar{\mathcal{J}}_i$ can be made output stationary with $f_1 \ldots f_m$ if and only if:

$$\beta_i h_i = \alpha_1 f_1 + \ldots \alpha_\ell f_\ell + \xi_{hi}$$

with $\alpha_1 \ldots \alpha_\ell$ nonzero and $h_{i\mathcal{R}} \in \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_\ell \pmod{\bar{\mathcal{J}}_i}$ where:

$$h_{i\mathcal{R}} \equiv A\, \xi_{hi} \pmod{\mathcal{J}_i}$$

Proof

The additional constraint on $h_{i\mathcal{R}}$ is necessary to insure that ii) in the proof of Theorem 3.4.1 has a solution for which $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_\ell$ are invariant.

Δ

Lemma 3.4.1 and Theorem 3.4.1 permit some simple observations on the possibility of making more than one event vector output stationary with $f_1 \ldots f_m$. If $\bar{\mathcal{J}}_i$ is the detection

space for two different event vectors h_i and h_j, i.e. if h_i and h_j are detection equivalent, Lemma 3.4.1 states that making h_i output stationary with $f_1 \ldots f_m$ results in both h_i and h_j being output stationary with $f_1 \ldots f_m$. The following corollary introduces mutual output stationarity for the special case where detection spaces $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ are contained in the appropriate union $\bigcup_{k \in I_i} \bar{R}_k$, $k = 1 \ldots q$, and the index sets $I_1 \ldots I_q$ are non-intersecting. For this example $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ are clearly noninteractive and can be considered separately. The general result is derived in Section 3.4.3.

Corollary 3 Assume $f_1 \ldots f_m$ are mutually detectable and define the index sets I_i to be the smallest sets such that:

$$\bar{\mathcal{J}}_i \subset \bigcup_{k \in I_i} \bar{R}_k \qquad i = 1 \ldots q$$

If each of $\bar{\mathcal{J}}_i$ can be made output stationary with $f_1 \ldots f_m$, then the set of subspaces $\bar{\mathcal{J}}_i \ldots \bar{\mathcal{J}}_q$ can be made output stationary with $f_1 \ldots f_m$ if:

$$I_i \cap I_j = \emptyset$$

for all $i \neq j$.

Proof

Application of Theorem 3.4.1 for individual $\bar{\mathcal{J}}_i$.

Δ

If $\bar{\mathcal{J}}_i$ can be made output stationary with $f_1 \ldots f_m$, the effect on the assignability of the n eigenvalues associated with A-DC must be determined. As a prerequisite, it is necessary to develop the relationship of $\bar{\mathcal{J}}_i$ to $\bar{R}_1 \ldots \bar{R}_m$ more fully. Lemma 3.4.2 provides the necessary result. If $\nu_{\mathcal{J}}$ is the dimension of $\bar{\mathcal{J}}_i$, then:

$$\nu_{\mathcal{J}} \leq \nu_k \qquad k = 1 \ldots \ell \tag{3.4.3}$$

In fact, $\bar{\mathcal{J}}_i$ has $\nu_{\mathcal{J}}$-dimensional projections $\bar{P}_1 \ldots \bar{P}_\ell$ onto the respective $\bar{R}_1 \ldots \bar{R}_\ell$. The interpretation of $\bar{P}_1 \ldots \bar{P}_\ell$ is that for every vector $\rho_j \in \bar{\mathcal{J}}_i$, there are unique vectors $\xi_k \in \bar{P}_k \subset \bar{R}_k$, $k = 1 \ldots \ell$, such that ρ_j is a linear combination of $\xi_1 \ldots \xi_k$:

$$\rho_j = \xi_1 + \ldots \xi_k \tag{3.4.4}$$

As discussed in Sections 2.3.2 and 2.3.3, projections can be defined in terms of equivalent subspaces. Thus $\bar{P}_1 \ldots \bar{P}_\ell$ are defined by:

$$\begin{aligned} &P_k \subset \bar{R}_k \\ &\bar{P}_k \equiv \mathcal{J}_i \pmod{\bar{R}_k^D} \quad k = 1 \ldots \ell \end{aligned} \tag{3.4.5}$$

where $\bar{R}_k^D$ is as defined in Section 3.3.3:

$$\bar{R}_k^D = \bar{R}_1 \oplus \ldots \bar{R}_{k-1} \oplus \bar{R}_{k+1} \oplus \ldots \bar{R}_m \tag{3.3.21}$$

Lemma 3.4.2 Assume $f_1 \ldots f_\ell$ satisfy Theorem 3.4.1 for $\bar{\mathcal{J}}_i$. Then every $\rho_j \in \bar{\mathcal{J}}_i$ has a unique decomposition:

$$\rho_j = \xi_{ij} + \ldots \xi_{\ell j}$$

where $\xi_{kj} \in \bar{R}_k$, $k = 1 \ldots \ell$. If $\rho_1 \ldots \rho_{\nu_{\mathcal{J}}}$ form a basis for $\bar{\mathcal{J}}_i$, then the corresponding vectors $\xi_{k1}, \ldots \xi_{k\nu_{\mathcal{J}}}$ form a basis for $\bar{P}_k$, where $\bar{P}_k$ is the $\nu_{\mathcal{J}}$-dimensional subspace defined by:

$$\begin{aligned} &\bar{P}_k \subset \bar{R}_k \\ &\bar{P}_k \equiv \bar{\mathcal{J}}_i \pmod{\bar{R}_k^D} \end{aligned} \qquad k = 1 \ldots \ell$$

Proof

By implication, every vector $\rho_j \in \bar{\mathcal{P}}_i$ must satisfy a relationship similar to that of h_i to $f_1 \ldots f_\ell$ in Theorem 3.4.1. If $C\rho_j \neq \phi$, then:

$$\beta_j C\rho_j = \alpha_1 Cf_1 + \ldots \alpha_\ell Cf_\ell$$

where $\alpha_1 \ldots \alpha_\ell$ are the same as for h_i, but the scale factor β_j may be different. This relationship implies that their exist vectors $\xi_{kj} \in \bar{\mathcal{R}}_k$ such that:

i) $$\rho_j = \xi_{1j} + \ldots \xi_{\ell j}$$

and that all of $\xi_{1j} \ldots \xi_{\ell j}$ are nonzero. The uniqueness of $\xi_{1j} \ldots \xi_{\ell j}$ is a consequence of Lemma 3.3.3.

If $C\rho_j = \phi$, let μ be the smallest integer such that $CA^\mu \rho_j \neq \phi$. Theorem 3.4.1 then requires:

$$\beta_j CA^\mu \rho_j = \alpha_1 Cf_1 + \ldots \alpha_\ell Cf_\ell$$

Since D can be chosen such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_m$ are invariant with respect to $A+DC$ the output separability of $f_1 \ldots f_m$ plus the fact that $\alpha_1 \ldots \alpha_\ell$ are nonzero requires ρ_j to have a representation similar to i) with $\xi_{kj} \in \mathcal{R}_k$, i.e. $C\xi_{kj} = \phi$. Uniqueness is again guaranteed by Lemma 3.3.3.

Assume the above procedure has been applied for a basis $\rho_1 \ldots \rho_{\nu_\ell}$ of $\bar{\mathcal{P}}_i$. By definition the vectors sets $\xi_{k1} \ldots \xi_{k\nu_\ell}$ span the respective $\bar{\mathcal{P}}_k$ for $k = 1 \ldots \ell$. It remains only to show that the vectors $\xi_{k1} \ldots \xi_{k\nu_\ell}$ in each set are linearly independent and hence that $\bar{\mathcal{P}}_k$ is ν_ℓ-dimensional. Assume that $\xi_{k1} \ldots \xi_{k\nu_\ell}$ are not linearly independent for some $k \leq \ell$ and that ξ_{k1} is a linear combination at $\xi_{k2} \ldots \xi_{k\nu_\ell}$:

ii) $\xi_{k1} = \sigma_2 \xi_{k2} + \ldots \sigma_{\nu_{\mathcal{J}}} \xi_{k\nu_{\mathcal{J}}} \qquad k \leq \ell$

Now define $\rho' \in \bar{\mathcal{J}}_i$ as a similar linear combination of $\rho_1 \ldots \rho_{\nu_{\mathcal{J}}} \in \bar{\mathcal{J}}_i$:

iii) $\rho' = \rho_1 - \sigma_2 \rho_2 - \ldots \sigma_{\nu_{\mathcal{J}}} \rho_{\nu_{\mathcal{J}}}$

If ρ' is expressed as a sum of vectors $\xi'_1 \ldots \xi'_\ell$ in $\bar{R}_1 \ldots \bar{R}_\ell$, then ii) and iii) imply that ξ'_1 is zero, i.e. ρ' does not have a component in $\bar{R}_1$. This contradicts the argument attendant with i) which in turn implies that ξ_{k1} cannot satisfy ii). An extension of this reasoning completes the proof that $\bar{P}_1 \ldots \bar{P}_\ell$ are $\nu_{\mathcal{J}}$-dimensional. Δ

With Lemma 3.4.2, the effect of output stationarity on eigenvalue assignment can be determined. If:

$$\nu^c_{\mathcal{J}L} = \sum_{k=1,\ell} \nu_k \qquad (3.4.6)$$

and $f_1 \ldots f_m$ are mutually detectable, it is obvious that the $n\text{-}\nu^c_{\mathcal{J}L}$ eigenvalues not associated with $\bar{R}_1 \ldots \bar{R}_\ell$ can still be assigned almost arbitrarily. The assignability of the remaining $\nu^c_{\mathcal{J}L}$ eigenvalues is summarized in Theorem 3.4.2. The effect of making $\bar{\mathcal{J}}_i$ output stationary with $f_1 \ldots f_m$ is that $\ell \cdot \nu_{\mathcal{J}}$ eigenvalues associated with $\bar{R}_1 \oplus \ldots \bar{R}_\ell$ can be assigned. The assignment allows only $\nu_{\mathcal{J}}$ distinct eigenvalues, each of which must have multiplicity ℓ.

A consequence of Lemma 3.4.2 is that $\nu_{\mathcal{J}} \leq \nu_k$ for $k = 1 \ldots \ell$ and thus $\nu_{\mathcal{J}} \leq \min(\nu_k,\ k = 1 \ldots \ell)$. If $\bar{\mathcal{J}}_i$ is to be made output stationary with $f_1 \ldots f_m$, this suggests two design procedures. The first is that $f_1 \ldots f_m$ be chosen such that their detection spaces are of approximately equal dimension and the second is that h_i be chosen such that its detection space is as large as possible. To the extent that this can be accomplished, the number of unassignable eigenvalues can be minimized.

Theorem 3.4.2 Assume $\bar{\mathcal{J}}_i$ satisfies Theorem 3.4.1 for $\bar{R}_1 \ldots \bar{R}_\ell$. If $\bar{\mathcal{J}}_i$ is made output stationary with $f_1 \ldots f_m$, the $\nu_\mathcal{J}$ eigenvalues of A-DC associated with $\bar{\mathcal{J}}_i$ can be assigned almost arbitrarily. The eigenvalues of A-DC associated with each of $\bar{R}_1 \ldots \bar{R}_\ell$ consist of the $\nu_\mathcal{J}$ associated with $\bar{\mathcal{J}}_i$ plus $\nu_k - \nu_\mathcal{J}$ eigenvalues which are independent of the choice of $D \in D^*$.

Proof

Let $\Lambda_\mathcal{J}$ be the desired set of $\nu_\mathcal{J}$ eigenvalues associated with $\bar{\mathcal{J}}_i$. By the Detection Theorem, D can be chosen using Lemma 3.2.7 such that $\bar{\mathcal{J}}_i$ is invariant and the eigenvalues of A-DC associated with $\bar{\mathcal{J}}_i$ are the elements of $\Lambda_\mathcal{J}$. The choice of D is not unique. If $g_\mathcal{J}$ is the detection generator for $\bar{\mathcal{J}}_i$, the net effect of Lemma 3.2.7 is to define some vector $\xi_{\mathcal{J}i} \in \bar{\mathcal{J}}_i$ and require D to satisfy:

$$\text{i)} \qquad (\text{A-DC})\, \beta_i h_i = -\beta_i \Upsilon(A) g_\mathcal{J} + \beta_i A h_i = \xi_{\mathcal{J}i}$$

where the roots of the minimal polynomial $\Upsilon(\cdot)$ are the elements of $\Lambda_\mathcal{J}$. The first requirement is to show that i) leaves sufficient freedom in the choice of D to insure that $D \in D^*$, i.e. that $\bar{\mathcal{J}}_i$ can be made output stationary with $f_1 \ldots f_m$. But i) above is identical to i) in Theorem 3.4.1 and the same argument used in Theorem 3.4.1 verifies that a satisfactory D exists.

With D specified by i), it remains to determine the eigenvalues of A-DC associated with $\bar{R}_1 \ldots \bar{R}_\ell$. Denote the relative minimal polynomial (mod $\bar{R}_k^D$) of $\bar{\mathcal{J}}_i$ by $\Upsilon_{Dk}(\cdot)$:

$$\Upsilon_{Dk}(\text{A-DC})\, \bar{\mathcal{J}}_i \equiv \emptyset \pmod{\bar{R}_k^D} \quad k = 1 \ldots \ell$$

Since the projections $\bar{P}_1 \ldots \bar{P}_\ell$ are defined by:

$$\text{ii)} \qquad \bar{P}_k \subset \bar{R}_k \qquad k = 1 \ldots \ell$$

$$\bar{P}_k \equiv \bar{\mathcal{J}}_i \pmod{\bar{R}_k^D}$$

then $\gamma_{D1}(\cdot) \ldots \gamma_{D\ell}(\cdot)$ are the minimal polynomials of $\bar{P}_1 \ldots \bar{P}_\ell$:

iii) $$\gamma_{Dk}(A-DC)\bar{P}_k = \phi \qquad k = 1 \ldots \ell$$

$\gamma_{D1}(\cdot) \ldots \gamma_{D\ell}(\cdot)$ are also divisors of $\gamma(\cdot)$ (Section 2.3.2). In fact, because $\bar{P}_1 \ldots \bar{P}_\ell$ are ν_μ-dimensional subspaces, $\gamma_{D1}(\cdot) \ldots \gamma_{D\ell}(\cdot)$ must equal $\gamma(\cdot)$ and therefore:

$$\gamma(A-DC)\bar{P}_k = \phi \qquad k = 1 \ldots \ell$$

It follows that the ν_μ eigenvalues of $A-DC$ associated with each of $\bar{P}_1 \ldots \bar{P}_\ell$ are those in Λ_μ. Because $\bar{P}_k \subset \bar{R}_k$, they are also ν_μ of the eigenvalues associated with each of $\bar{R}_1 \ldots \bar{R}_\ell$.

To determine the remaining $\nu_k - \nu_\mu$ eigenvalues of $A-DC$ associated with $\bar{R}_k$, $k = 1 \ldots \ell$, define the partitions:

$$\bar{R}_k = \bar{P}_k \oplus P_k^c$$

where $P_k^c \subset \eta(C)$. Equation iii) implies that each of $\bar{R}_1 \ldots \bar{R}_\ell$ is an invariant subspace. The operator $A-DC$ on $\bar{R}_k$ is therefore specified by the four operators:

i)' $A_{\bar{P}_k} : \bar{P}_k \to \bar{P}_k$

ii)' $A_{P_k^c} : P_k^c \to P_k^c$

iii)' $A_{\bar{P}_k P_k^c} : P_k^c \to \bar{P}_k$

iv)' $\phi : \bar{P}_k \to P_k^c$

The eigenvalues of $A-DC$ associated with $\bar{R}_k$ are those of $A_{\bar{P}_k}$ and $A_{P_k^c}$. Those of $A_{\bar{P}_k}$ are the solution of $\gamma_{Dk}(\cdot)$ and by the previous reasoning are the elements of Λ_μ. Since $P_k^c \subset \eta(C)$, those of $A_{P_k^c}$ are independent of the choice of $D \in D^*$.

Δ

There are two sets of algorithms to be supplied for Theorem 3.4.2: Algorithms for assigning the ν_ℓ eigenvalues in Λ_ℓ and algorithms for determining the unassignable eigenvalues of A-DC if $\bar{\mathcal{S}}_i$ is made output stationary with $f_1 \ldots f_m$. Lemma 3.2.7 and the construction used in the proof of Theorem 3.4.1 satisfy the first requirement. This algorithm is presented in matrix form in Corollary 1. The basis for the corollary is the matrix counterpart of equivalent subspaces.

Corollary 1 Assume h_i satisfies Theorem 3.4.1. Then h_i is output stationary with $f_1 \ldots f_m$ and each element of Λ_ℓ is an eigenvalue of A-DC of multiplicity ℓ if D is a solution to the ℓ linear equations:

$$\text{(A-DC) } f_k = \frac{1}{\alpha_k} (PR)_k \, (\bar{R}_L^{c\,T} \bar{R}_L^c)^{-1} \left[-\beta_i Y(A) g_\ell + \sum_{j=1,\ell} \alpha_j A f_j \right] \quad k \neq 1 \ldots \ell$$

$\bar{R}_L^c$ is specified by the detection spaces of $f_1 \ldots f_\ell$:

$$\bar{R}_L^c = [\bar{R}_1 : \ldots : \bar{R}_\ell]$$

and $(PR)_k$ is $\bar{R}_L^c$ with all but the k^{th} submatrix set equal to zero.

Proof

Lemma 3.2.7 and the matrix equivalent of i) in Theorem 3.4.1 require:

$$\text{i) } \text{(A-DC) } \beta_i h_i = -\beta_i Y(A) g_\ell + \beta_i A h_i = \xi_{\ell i}$$

where $\beta_i h_i$ can be expressed as:

$$\beta_i h_i = \alpha_1 f_1 + \ldots \alpha_\ell f_\ell + \xi_{hi}$$

and ξ_{hi} satisfies $C \xi_{hi} = 0$. Substituting into i) above:

$$\text{(A-DC) } (\alpha_1 f_1 + \ldots \alpha_\ell f_\ell + \xi_{hi}) = -\beta_i Y(A) g_\ell + A(\alpha_1 f_1 + \ldots \alpha_\ell f_\ell + \xi_{hi})$$

which becomes:

$$\text{ii)}\quad (\text{A-DC})\,(\alpha_1 f_1 + \ldots \alpha_\ell f_\ell) = -\beta_i\, Y(A)\, g_\mathscr{I} + \sum_{j=1,\ell} \alpha_j\, Af_j$$

Invariance of $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_\ell$ adds the requirement that (A-DC) f_k must be in the range space of $\bar{\mathcal{R}}_k$ for $k = 1 \ldots \ell$. This decouples ii) into a set of ℓ linear equations:

$$(\text{A-DC})\, f_k = \xi_k \qquad k = 1 \ldots \ell$$

where ξ_k is $1/\alpha_k$ times the projection of the right hand side of ii) into the range space of $\bar{\mathcal{R}}_k$. ξ_k is therefore:

$$\xi_k = \frac{1}{\alpha_k} (PR)_k \left(\bar{R}_L^{c\,T} \bar{R}_L^c\right)^{-1} \left[-\beta_i Y(A) g_\mathscr{I} + \sum_{j=1,\ell} \alpha_j Af_j\right] \quad k = 1 \ldots \ell$$

Theorem 3.4.1 guarantees that a solution for D exists.

Δ

The second algorithm required for Theorem 3.4.2 is for the unassignable eigenvalues of A-DC resulting from application of Corollary 1. Let $\Psi_k(\cdot)$ be the minimal polynomial of $\bar{\mathcal{R}}_k$ with respect to A-DC:

$$\Psi_k(\text{A-DC})\, \bar{\mathcal{R}}_k = 0 \qquad k = 1 \ldots \ell \qquad (3.4.7)$$

The unassignable eigenvalues of A-DC associated with $\bar{\mathcal{R}}_k$ are the roots of $\Psi_k(\cdot)$ which are not also roots of $Y(\cdot)$, i.e. they are the roots of the quotient $\Psi(\cdot)/Y(\cdot)$. If $\Psi(\cdot)$ is expressed:

$$\Psi_k(\lambda) = \lambda^{\nu_k} + p_{k\nu_k} \lambda^{\nu_{k-1}} + \ldots p_{k2} \lambda + p_{k1} \qquad (3.4.8)$$

then the coefficients $p_{k1} \ldots p_{k\nu_k}$ are the components of the coefficient vector p_k:

$$p_k = \begin{bmatrix} p_{k1} \\ \vdots \\ p_{k\nu_k} \end{bmatrix}$$

where p_k is the solution to:

$$Wp_k = -(A-DC)^{\nu_k} g_k \qquad (3.4.4)$$

and W is the controllability matrix of g_k:

$$W = [g_k \ (A-DC)g_k \ldots \ (A-DC)^{\nu_k - 1} g_k] \qquad (3.4.10)$$

3.4.2 An Example

The simplest examples of the usefulness of output stationarity are to be found in systems with parallel redundancy. For discussion purposes it is assumed that the redundant components are dynamically equivalent and that they have inputs which are measured. The performance level for the components is assumed to be a function of both system parameters (e.g. temperature or input commands) and parameters unique to the individual components (e.g. time constants or jet thrust levels). The theory developed in Section 3.3.3 permits a failure detection system for parameters of the latter type to be designed. The results of Section 3.4.1 permit a single failure detection system to be designed for both types of failures.

Consider a modification of the spinning cylinder example developed in Section 3.3.2. The cylinder dynamics remain the same:

$$\ddot{\phi}(t) + \beta_r \dot{\phi}(t) = \alpha_r T(t) \ : \ \phi(0), \dot{\phi}(0) \text{ given} \qquad (3.4.11)$$

but the applied torque is the sum of the torque from two independent control surfaces:

$$T(t) = T_1(t) + T_2(t) \tag{3.4.12}$$

Each of the control surfaces is driven by the same control law as before:

$$\dot{T}_1(t) + \beta_1 T_1(t) = -\alpha_1 \varphi_m(t) - \alpha_2 \dot{\varphi}_m(t) + \alpha_1 \varphi_c(t) \tag{3.4.13}$$

$$\dot{T}_2(t) + \beta_2 T_2(t) = -\alpha_1 \varphi_m(t) - \alpha_2 \dot{\varphi}_m(t) + \alpha_1 \varphi_c(t)$$

Besides $\varphi(t)$ and $\dot{\varphi}(t)$, the feedback torques $T_1(t)$ and $T_2(t)$ are measured, but they are not used in the control law. The measurements of $T_1(t)$ and $T_2(t)$ are used only for the failure detection algorithm. A block diagram for this system is given in Figure 3.2.

The system model differs from that developed in Section 3.3.2 only in the inclusion of an additional measurement and state component:

$$C_5 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \qquad A_5 = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & -\beta_r & 0 & \alpha_r & \alpha_r \\ 0 & -\beta_r & 0 & \alpha_r & \alpha_r \\ -\alpha_1 & 0 & -\alpha_2 & -\beta_1 & 0 \\ -\alpha_1 & 0 & -\alpha_2 & 0 & -\beta_2 \end{bmatrix}$$

$$B_5 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \alpha_1 \\ \alpha_1 \end{bmatrix} \tag{3.4.14}$$

The state vector is:

$$x(t) = \begin{bmatrix} \varphi_m(t) \\ \dot{\varphi}(t) \\ \dot{\varphi}_m(t) \\ T_1(t) \\ T_2(t) \end{bmatrix} \tag{3.4.15}$$

In the previous example, failures in the roll and roll rate sensors were discussed. In this example, failures of the control law and the control

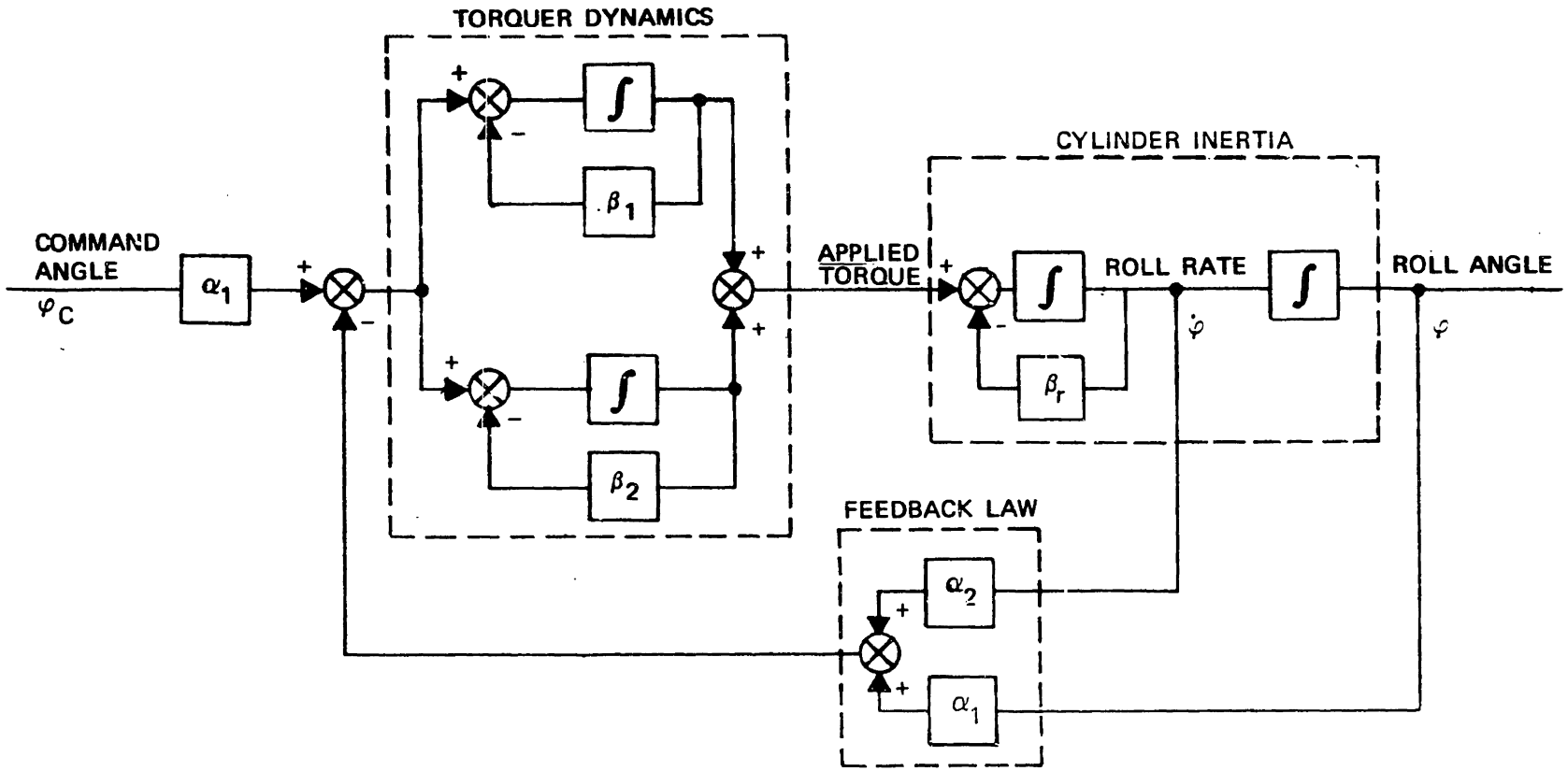


Figure 3.2 Block diagram for spinning cylinder example with parallel redundant torquers. Measurements are roll angle, roll rate and outputs of both torquers.

surfaces are analyzed. It is assumed that the torque sensors cannot fail.

The control surface failures can clearly be modeled using the controller failure model:

$$\dot{x}(t) = A_5 x(t) + B_5 u(t) + f_1 n(t) : x(0) \text{ given} \qquad (2.1.7)$$

$$y(t) = C_5 x(t)$$

where f_1 is:

$$f_1 = f_{T1} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} = e_{54}$$

$$f_1 = f_{T2} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = e_{55} \qquad (3.4.16)$$

for failures in the respective control surfaces. A failure of the control law can be modeled as a change in either of the gains α_1 and α_2. These changes can also be modeled using the controller failure model with:

$$f_1 = f_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} = e_{45} + e_{55} \qquad (3.4.17)$$

If either α_1 or α_2 changes, both $T_1(t)$ and $T_2(t)$ differ from their nominal values by the same amount. Thus a control law failure has the same effect on the closed loop system and its measured outputs as simultaneous, identical failures in the control surfaces.

The event vectors f_{T1}, f_{T2}, f_c are not output separable since:

$$\text{rk}\, C_5[f_{T1} \; f_{T2} \; f_c] = 2 \qquad (3.4.18)$$

As a consequence, a failure detection system for the three failures cannot be designed using Section 3.3.4 -- in fact, it is proved in Section 3.5 that f_{T1}, f_{T2}, f_c are not mutually detectable. However, the two event vectors f_{T1}, f_{T2}, are output separable since:

$$\text{rk}\, C_5 [f_{T1} \quad f_{T2}] = 2 \tag{3.4.19}$$

Each event vector is its own detection space, $\bar{\mathcal{R}}_{T1}$ and $\bar{\mathcal{R}}_{T2}$ respectively, and it follows from Theorem 3.3.1 that f_{T1}, f_{T2} are mutually detectable.

A single failure detection system can be designed for f_{T1}, f_{T2}, f_c only if output stationarity is used. The event vector f_c is its own detection space, which is denoted by $\bar{\mathcal{J}}_c$. From (3.4.16) and (3.4.17) it follows that:

$$\bar{\mathcal{J}}_c \subset \bar{\mathcal{R}}_{T1} \oplus \bar{\mathcal{R}}_{T2} \tag{3.4.20}$$

since f_c is a linear combination of f_{T1} and f_{T2}. It also follows that:

$$C_5 f_c = C_5 f_{T1} + C_5 f_{T2} \tag{3.4.21}$$

and Theorem 3.4.1 verifies that f_c can be made output stationary with f_{T1}, f_{T2}. If this is done, a failure in the first control surface causes a unidirectional error along $C_5 f_{T1}$ in the reference model output. Similarly, a failure in the second control surface generates an error along $C_5 f_{T2}$ and a change in the control law generates an error along $C_5 f_c$. The only restriction which Theorem 3.4.2 places on the eigenvalue assignment for $A_5 - DC_5$ is that the eigenvalue associated with $\bar{\mathcal{R}}_1$ must be identical to the eigenvalue associated with $\bar{\mathcal{R}}_{T2}$. Since $\bar{\mathcal{J}}_c$, $\bar{\mathcal{R}}_{T1}$, and $\bar{\mathcal{R}}_{T2}$ all have the same dimension, no unassignable eigenvalues are introduced.

This simple example illustrates the basic utility of output stationarity. If f_c is not made output stationary with f_{T1}, f_{T2} and a control law change occurs, the failure results in output errors along both $C_5 f_{T1}$ and $C_5 f_{T2}$ which do not have any apparent correlation. With the two eigenvalues of $A_5 - DC_5$ constrained to be equal, however, the errors along $C_5 f_{T1}$ and $C_5 f_{T2}$ are correlated to the extent that the resultant

output error is unidirectional along $C_5 f_c$. If the probability of simultaneous, identical control surface failures is small, a unidirectional output error along $C_5 f_{T1}$ is sufficient information to conclude that a control law change has occurred rather than two control surface failures.

3.4.3 Addition of a Stationary Class to a Detectable Set

If each of $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ can be made output stationary with $f_1 \ldots f_m$, it does not necessarily follow that $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ can be made mutually output stationary with $f_1 \ldots f_m$. The study of mutual output stationarity can be undertaken at three levels. The simplest level was summarized in Corollary 3 of Theorem 3.4.1. If $I_1 \ldots I_q \in \{1, 2 \ldots m\}$ are the smallest index sets such that:

$$\bar{\mathcal{J}}_i \subset \bigcup_{k \in I_i} \bar{R}_k \qquad i = 1 \ldots q \qquad (3.4.22)$$

the corollary asserts that $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ can be made mutually output stationary if the index sets are all non-intersecting:

$$I_i \cap I_j = \emptyset \qquad \forall\, i \neq j \qquad (3.4.23)$$

The second level of difficulty occurs if the index sets are identical:

$$I_i = I_j \qquad \forall\, i \neq j \qquad (3.4.24)$$

In this case, Theorem 3.4.3 proves that the test for mutual output stationarity with $f_1 \ldots f_m$ need be applied for only one subspace from the set $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$. If there are ℓ indicial elements in I_i, $\bar{\mathcal{J}}_i$ defines a class $\bar{\mathcal{J}}_i^*$ of detection spaces (actually, an ℓ-dimensional subspace of detection spaces) which are always mutually output stationary with $f_1 \ldots f_m$ when $\bar{\mathcal{J}}_i$ is output stationary with $f_1 \ldots f_m$. If $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ can be made output stationary with $f_1 \ldots f_m$, it is shown that all of $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ must be elements of $\bar{\mathcal{J}}_i^*$, for any $i \leq q$. The output associated with each element of $\bar{\mathcal{J}}_i^*$ has a unique direction in the ℓ-dimensional subspace of $\mathcal{S}^m$ spanned by $\bigcup_{k \in I_i} Cf_k$.

An extension of the above is if some $\mu \leq q$ exists such that

$$I_i \subset I_\mu \qquad i = 1 \ldots q \tag{3.4.35}$$

Technically, it is necessary to use Theorem 3.4.4 in this instance, but the difficulty is one of definitions rather than theory. Theorem 3.4.3 can again be applied, but only for $\bar{\mathscr{I}}_\mu$ and the corresponding $\bar{\mathscr{I}}_\mu^*$. Whereas $\bar{\mathscr{I}}_1 \ldots \bar{\mathscr{I}}_q$ may be elements of $\bar{\mathscr{I}}_\mu^*$, it is not true that $\bar{\mathscr{I}}_\mu$ is an element of the various $\bar{\mathscr{I}}_i^*$ for $i \neq \mu$.

As the size of the index set I_i increases, the dimension of the class $\bar{\mathscr{I}}_i^*$ increases. In the general case where $I_1 \ldots I_q$ are intersecting but independent, the class $\bar{\mathscr{I}}_\mathscr{K}^*$ must be defined on the index set:

$$I_\mathscr{K} = \bigcup_{i=1,q} I_i \tag{3.4.26}$$

If none of the individual index sets is equal to $I_\mathscr{K}$, it is necessary to define an event vector $h_\mathscr{K}$ and its detection space $\bar{\mathscr{I}}_\mathscr{K}$ on $\bigcup_{k \in I_\mathscr{K}} \bar{R}_k$. If $\bar{\mathscr{I}}_1 \ldots \bar{\mathscr{I}}_q$ can be made mutually output stationary -- that is, if $\bar{\mathscr{I}}_\mathscr{K}^*$ exists-- Lemma 3.4.4 provides an algorithm for finding $\bar{\mathscr{I}}_\mathscr{K}$ and the corollary to Theorem 3.4.4 then yields $\bar{\mathscr{I}}_\mathscr{K}^*$.

The first two levels of mutual output stationarity are simple extensions of the arguments of Section 3.4.1. The algebra preceding Theorem 3.4.4 is considerably more involved and execution of Lemma 3.4.4 is nontrivial.

Before considering the general output stationarity problem, it is necessary to find a convenient means for characterizing $\bar{\mathscr{I}}_i^*$. The elements of $\bar{\mathscr{I}}_i^*$ are related by a coordinate transformation. In order to exploit this property, it is useful to construct a basis for $\bar{\mathscr{I}}_i$. The basis can be arbitrarily chosen for $\bar{\mathscr{I}}_i$ and once it is used to define the normal subspace $\bar{\mathscr{I}}_i^\circ$ for $\bar{\mathscr{I}}_i^*$, it can be dropped. The different elements of $\bar{\mathscr{I}}_i^*$ need not have related bases.

Definition 3.4.2 Assume $\bar{\mathscr{I}}_i$ satisfies Theorem 3.4.1 and h_i satisfies:

$$\beta_i h_i = \sum_{j \in I_i} \alpha_j f_j + \xi_{hi}$$

Choose a basis $\rho_1 \ldots \rho_{\nu_{\mathcal{J}}}$ for $\bar{\mathcal{J}}_i$. The basis vectors have a unique representation:

$$\rho_j = \sum_{k \in I_i} \xi_{kj} \qquad j = 1 \ldots \nu_{\mathcal{J}}$$

where $\xi_{kj} \in \bar{\mathcal{R}}_k$, $k \in I_i$. Define the vectors:

$$\rho_j^{\infty} = \sum_{k \in I_i} \frac{1}{\alpha_k} \xi_{kj}$$

The subspace $\bar{\mathcal{J}}_i^{\circ}$ spanned by $\rho_1^{\circ} \ldots \rho_{\nu_{\mathcal{J}}}^{\circ}$ is the normal subspace associated with $\bar{\mathcal{J}}_i$.

The normal subspace $\bar{\mathcal{J}}_i^{\circ}$ clearly has the same dimension as $\bar{\mathcal{J}}_i$; for every vector in $\bar{\mathcal{J}}_i$, Definition 3.4.2 uniquely defines a corresponding vector in $\bar{\mathcal{J}}_i^{\circ}$. The effect of the normalization is that the subspace $\bar{\mathcal{J}}_i$ is "rotated" in $\mathcal{E}^n$ until a subspace $\bar{\mathcal{J}}_i^{\circ}$ is found for which there exists a vector h_i° which satisfies:

$$\beta_i h_i^{\circ} = \sum_{k \in I_i} f_k + \xi_{hi}^{\circ} \tag{3.4.27}$$

and:

$$h_{i\mathcal{R}}^{\circ} = A \xi_{hi}^{\circ} \tag{3.4.28}$$

The important distinction between h_i° and h_i is that the α_k, $k \in I_i$, do not appear in the representation of h_i° given in (3.4.27). The construction used in Definition 3.4.2 suggest that the simplest procedure for determining h_i° is to find coefficients $\mu_1 \ldots \mu_{\nu_{\mathcal{J}}}$ such that:

$$h_i = \sum_{j=1,\, \nu_{\mathcal{J}}} \mu_j \rho_j \tag{3.4.29}$$

h_i° is then specified by:

$$h_i^\circ = \sum_{j=1,\nu_\ell} \mu_j \rho_j^\circ \qquad (3.4.30)$$

and ξ_{hi}° and $h_{i\ell}^\circ$ follow from expressing (3.4.30) in the form of (3.4.27). The output direction associated with $\bar{\mathcal{I}}_i^\circ$ is:

$$\beta_i \, Ch_i = \sum_{k \in I_i} Cf_k \qquad (3.4.31)$$

It is a simple matter to show that $\bar{\mathcal{I}}_i^\circ$ is the detection space for h_i° -- Lemma 3.4.3 provides the proof.

Lemma 3.4.3 Let $\bar{\mathcal{I}}_i^\circ$ be the normal subspace of $\bar{\mathcal{I}}_i$ and $h_i^\circ \in \bar{\mathcal{I}}_i^\circ$ correspond to $h_i \in \bar{\mathcal{I}}_i$. Then $\bar{\mathcal{I}}_i^\circ$ is the detection space for h_i°.

Proof

Choose the basis of $\bar{\mathcal{I}}_i$ in Definition 3.4.2 as $h_i, \rho_2 \ldots \rho_{\nu_\ell}$ where $\rho_2 \ldots \rho_{\nu_\ell} \in \eta(C)$. Since this requires that the corresponding vectors $\xi_{k2} \cdots \xi_{k\nu_\ell}$, $k \in I_i$, are elements of $\eta(C)$, it follows that $\rho_2^\circ \cdots \rho_{\nu_\ell}^\circ \in \eta(C)$ and conditions i) and iii) of Definition 3.1.2 apply for $\bar{\mathcal{I}}_i^\circ$. Conditions ii) and iv) are guaranteed by the one-to-one relationship between $\bar{\mathcal{I}}_i$ and $\bar{\mathcal{I}}_i^\circ$. Δ

$\bar{\mathcal{I}}_i^\circ$ has no intrinsically useful properties except that it provides a useful benchmark for comparing two detection spaces. If $\bar{\mathcal{I}}_i$ is output stationary with $f_1 \ldots f_m$, it is proved in Theorem 3.4.3 that $\bar{\mathcal{I}}_i^\circ$ is also output stationary with $f_1 \ldots f_m$. Therefore, it is possible to determine whether or not $\bar{\mathcal{I}}_1 \ldots \bar{\mathcal{I}}_q$ can be made mutually output stationary with $f_1 \ldots f_m$ by investigating the properties of the respective normal subspaces $\bar{\mathcal{I}}_1^\circ \ldots \bar{\mathcal{I}}_q^\circ$. Definition 3.4.3 is the first step in this effort. The definition restricts attention to the case where (3.2.24) is satisfied, i.e. the index sets $I_1 \ldots I_q$ for $\bar{\mathcal{I}}_1 \ldots \bar{\mathcal{I}}_q$ are equal. A generalization of Definition 3.4.3 is considered later.

Definition 3.4.3 The detection spaces $\bar{\mathcal{I}}_i$ and $\bar{\mathcal{I}}_j$ are output stationary equivalent if their normal subspaces $\bar{\mathcal{I}}_i^\circ$ and $\bar{\mathcal{I}}_j^\circ$ are equal.

It has not yet been shown that output stationary equivalence is tantamount to mutual output stationarity -- Theorem 3.4.3 provides the necessary result. For the special case where $I_1 \ldots I_q$ are identical index sets, Theorem 3.4.3 defines a necessary and sufficient condition for $\bar{\mathcal{I}}_1 \ldots \bar{\mathcal{I}}_q$ to be mutually output stationary with $f_1 \ldots f_m$.

Theorem 3.4.3 Let $\bar{\mathcal{I}}_i^*$ be the class of detection spaces which are output stationary equivalent to $\bar{\mathcal{I}}_i$. The elements of $\bar{\mathcal{I}}_i^*$ are mutually output stationary with $f_1 \ldots f_m$ if and only if $\bar{\mathcal{I}}_i$ is output stationary with $f_1 \ldots f_m$.

Proof

A consequence of the proof of Theorem 3.4.1 is that $\bar{\mathcal{I}}_i$ cannot be output stationary with $f_1 \ldots f_m$ unless there exists vectors $\xi_k \in \bar{R}_k$, $k \in I_i$, such that h_i satisfies:

$$\text{i)} \quad (\text{A-DC})\beta_i h_i = \sum_{k \in I_i} \alpha_k \xi_k + h_{i\mathcal{R}} = \xi_{\mathcal{I}_i}$$

where $\xi_{\mathcal{I}_i} \in \bar{\mathcal{I}}_i$. But the construction used in Definition 3.4.2 guarantees that a constraint which is identical to i) is to require $\xi_k \in \bar{R}_k$ to satisfy:

$$\text{ii)} \quad (\text{A-DC})\beta_i h_i^\circ = \sum_{k \in I_i} \xi_k + h_{i\mathcal{R}}^\circ = \xi_{\mathcal{I}_i}^\circ$$

where $\xi_{\mathcal{I}_i}^\circ \in \bar{\mathcal{I}}_i^\circ$. It follows that $\bar{\mathcal{I}}_i^\circ$ can be output stationary with $f_1 \ldots f_m$ if and only if $\bar{\mathcal{I}}_i$ is output stationary with $f_1 \ldots f_m$. Hence detection spaces in $\bar{\mathcal{I}}_i^*$ can be output stationary with $f_1 \ldots f_m$ if and only if $\bar{\mathcal{I}}_i$ is output stationary with $f_1 \ldots f_m$. Δ

The class $\bar{\mathcal{I}}_i^*$ is the bonus gained if $\bar{\mathcal{I}}_i$ is made output stationary with $f_1 \ldots f_m$. For every ℓ-tuple $\{\alpha'_k\}$, $k \in I_i$,[1] there is an event vector h_j and detection space $\bar{\mathcal{I}}_j$ such that:

[1] Strictly speaking, it is necessary to require $\alpha'_k \neq 0$. This restriction is removed later when Definition 3.4.3 is generalized.

$$C\beta_i h_j = \sum_{k \in I_i}' \alpha'_k C f_k \qquad (3\ 4.32)$$

and such that the corresponding normal subspace $\bar{\mathcal{J}}_j^\circ$ satisfies Definition 3.4.3:

$$\bar{\mathcal{J}}_i^\circ = \bar{\mathcal{J}}_j^\circ \qquad (3.4.33)$$

Indeed, for every $\{\alpha'_k\}$, $k \in I_i$, the reverse of the construction utilized in Definition 3.4.2 can be used to determine $\bar{\mathcal{J}}_j$ from $\bar{\mathcal{J}}_i^\circ$. Theorem 3.4.3 guarantees that all such detection spaces are output stationary with $f_1 \ldots f_m$ if $\bar{\mathcal{J}}_i$ is output stationary with $f_1 \ldots f_m$. Equation (3.4.32) and Definition 3.4.3 verify that each element of $\bar{\mathcal{J}}_i^*$ is associated with a unique output direction in $\mathcal{S}^m$. If a number of event vectors have detection spaces which are elements of $\bar{\mathcal{J}}_i^*$, output stationarity can increase the number of failures monitored by $\mathbb{D}$ considerably from the original set $f_1 \ldots f_m$.

Theorem 3.4.3 is the only result on mutual output stationarity which is necessary if the index sets $I_1 \ldots I_q$ corresponding to $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ are equal. Before continuing to the general case, it is convenient to solidify the result by providing a matrix equivalent to the construction prescribed in Definition 3.4.2. For notational convenience, Definition 3.4.2′ restricts attention to the case where I_i is $\{1, 2 \ldots \ell\}$.

Definition 3.4.2′ Assume $\bar{\mathcal{J}}_i$ can be made output stationary with $f_1 \ldots f_m$ and that Ξ is a solution to Corollary 1 of Theorem 3.4.1. The matrix $\bar{S}_i^\circ$ corresponding to the normal subspace of $\bar{\mathcal{J}}_i$ satisfies:

$$\bar{S}_i^\circ = T_i \bar{R}_L^c \Xi$$

where T_i is:

$$T_i = \begin{bmatrix} \frac{1}{\alpha_1} E_{\nu_1} & 0 \cdots\cdots 0 \\ 0 \ddots \frac{1}{\alpha_2} E_{\nu_2} & 0 \\ 0 \cdots\cdots 0 & \frac{1}{\alpha_\ell} E_{\nu_\ell} \end{bmatrix}$$

and E_{ν_k} are ν_k-dimensional identity matrices.

The remainder of this section is concerned with mutual output stationarity in the general case where $I_1 \ldots I_q$ do not satisfy (3.4.23) or (3.4.24). If I_χ is then defined by:

$$I_\chi = \bigcup_{i=1,q} I_i \tag{3.4.26}$$

a generalization of Theorem 3.4.3 requires that a normal subspace $\bar{\mathcal{J}}_\chi{}^\circ$ be defined on the corresponding $\bar{\mathcal{R}}_k$, $k \in I_\chi$ and that $\bar{\mathcal{J}}_\chi{}^\circ$ then be compared with $\bar{\mathcal{J}}_1{}^\circ \ldots \bar{\mathcal{J}}_q{}^\circ$ in order to determine mutual output stationarity. Unfortunately, the computation of $\bar{\mathcal{J}}_\chi{}^\circ$ is not easy unless I_χ is equal to some I_μ, $1 \leq \mu \leq q$.

It is assumed in what follows that there exist no subsets I_u and I_m of $\{1, 2 \ldots q\}$ and $\{1, 2, \ldots m\}$ respectively such that:

$$\bigcup_{i \in I_u} \bar{\mathcal{J}}_i{}^\circ \subset \bigcup_{k \in I_m} \bar{\mathcal{R}}_k \tag{3.4.34}$$

and:

$$\bigcup_{i \notin I_u} \bar{\mathcal{J}}_i{}^\circ \subset \bigcup_{k \notin I_m} \bar{\mathcal{R}}_k \tag{3.4.35}$$

If such a partition exists, Corollary 3 of Theorem 3.4.1 asserts that the detection filter design problem can be dissected into two independent output stationarity problems. The following results can then be applied to each problem separately. Consideration of two independent subsets of $\bar{\mathcal{J}}_1{}^\circ \ldots \bar{\mathcal{J}}_q{}^\circ$ increases the flexibility in eigenvalue assignment and may render some otherwise unassignable eigenvalues assignable.

For the normal subspaces $\bar{\mathcal{J}}_1{}^\circ \ldots \bar{\mathcal{J}}_q{}^\circ$, Definition 3.4.2 defines event vectors $h_1{}^\circ \ldots h_q{}^\circ$ and (3.4.27) in turn defines vectors $\xi_{h1}{}^\circ \ldots \xi_{hq}{}^\circ$. Lemma 3.4.4 states that $\bar{\mathcal{J}}_1{}^\circ \ldots \bar{\mathcal{J}}_q{}^\circ$ can be mutually output stationary with $f_1 \ldots f_m$ if and only if $\xi_{h1}{}^\circ \ldots \xi_{hq}{}^\circ$ satisfy a particular equivalence relationship. If this equivalence relationship is satisfied, it is then a straightforward matter to define $\bar{\mathcal{J}}_\chi{}^\circ$.

Lemma 3.4.4 Let I_i be the smallest index sets such that:

$$\bar{\mathcal{J}}_i{}^\circ \subset \bigcup_{k \in I_i} \bar{\mathcal{R}}_k \qquad i = 1 \ldots q$$

and assume that each of $\bar{\mathcal{P}}_1^\circ \ldots \bar{\mathcal{P}}_q^\circ$ can be made output stationary with $f_1 \ldots f_m$. Then $\bar{\mathcal{P}}_1^\circ \ldots \bar{\mathcal{P}}_q^\circ$ can be made mutually output stationary with $f_1 \ldots f_m$ if and only if there exists some vector $\xi_{\mathcal{N}}^\circ \in \mathcal{R}_1 \oplus \ldots \mathcal{R}_m$ such that:

$$\xi_{hi}^\circ \equiv \xi_{\mathcal{N}}^\circ \qquad (\text{mod } \mathcal{P}_i^\circ \cup_{k \notin I_i} \bar{\mathcal{R}}_k) \qquad i = 1 \ldots q$$

Proof

Sufficiency Let D^* be the class of operators such that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_m$ are invariant with respect to A-DC . Given that $\xi_{\mathcal{N}}^\circ$ exists, choose $D \in D^*$ such that:

$$\text{i) } (\text{A-DC}) f_k = \xi_k \qquad k = 1 \ldots m$$

where $\xi_k \in \bar{\mathcal{R}}_k$ and:

$$\text{ii) } \xi_1 + \ldots \xi_m = -A \xi_{\mathcal{N}}^\circ = -h_{\mathcal{N}\mathcal{R}}^\circ$$

Clearly $h_{\mathcal{N}\mathcal{R}}^\circ \in \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r$. For this choice of D, the images of $h_1^\circ \ldots h_q^\circ$ under A-DC are:

$$\text{iii) } (\text{A-DC}) \beta_i h_i^\circ = (\text{A-DC}) \sum_{k \in I_i} f_k + (\text{A-DC}) \xi_{hi}^\circ \qquad i = 1 \ldots q$$

$$= \sum_{k \in I_i} \xi_k + A \xi_{hi}^\circ$$

From ii) and the fact that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_m$ are disjoint, it follows that; $\sum_{k \in I_i} \xi_k$ and $A \xi_{\mathcal{N}}^\circ$ are equivalent vectors:

$$\sum_{k \in I_i} \xi_k \equiv -A \xi_{\mathcal{N}}^\circ \qquad (\text{mod } \bigcup_{k \notin I_i} \bar{\mathcal{R}}_k) \qquad i = 1 \ldots q$$

The equivalence argument clearly can be expanded to include $\bar{\mathcal{P}}_i^\circ$:

$$\text{iv) } \sum_{k \in I_i} \xi_k \equiv -A \xi_{\mathcal{N}}^\circ \qquad (\text{mod } \bar{\mathcal{P}}_i^\circ \bigcup_{k \notin I_i} \bar{\mathcal{R}}_k) \qquad i = 1 \ldots q$$

But by the hypothesis, $\xi_{\mathcal{N}}{}^{\circ}$ and $\xi_{hi}{}^{\circ}$ are equivalent (mod $\mathcal{I}_i{}^{\circ} \bigcup_{k \notin I_i} \bar{R}_k$). Since $\bar{R}_1 \ldots \bar{R}_m$ are invariant, the images $A\xi_{\mathcal{N}}{}^{\circ}$ and $A\xi_{hi}{}^{\circ}$ are also equivalent if $\mathcal{I}_i{}^{\circ}$ in the modulo argument is replaced by $\bar{\mathcal{I}}_i{}^{\circ}$:

$$v) \quad A\xi_{\mathcal{N}}{}^{\circ} \equiv A\xi_{hi}{}^{\circ} \qquad (\text{mod } \bar{\mathcal{I}}_i{}^{\circ} \bigcup_{k \notin I_i} \bar{R}_k) \qquad i = 1 \ldots q$$

Congruence is transitive so that iv) and v) together imply:

$$\sum_{k \in I_i} \xi_k \equiv -A\,\xi_{hi}{}^{\circ} \quad (\text{mod } \bar{\mathcal{I}}_i{}^{\circ} \bigcup_{k \notin I_i} \bar{R}_k) \qquad i = 1 \ldots q$$

However, by construction neither of the above terms has a projection onto $\bigcup_{k \notin I_i} \bar{R}_k$ and the congruence argument can be modified:

$$vi) \sum_{k \in I_i} \xi_k \equiv -A\xi_{hi}{}^{\circ} \qquad (\text{mod } \bar{\mathcal{I}}_i{}^{\circ}) \qquad i = 1 \ldots q$$

Substitution into iii) then yields:

$$(A\text{-}DC)\beta_i h_i{}^{\circ} \equiv \phi \qquad (\text{mod } \bar{\mathcal{I}}_i{}^{\circ}) \qquad i = 1 \ldots q$$

By Lemma 3.2.2, this is sufficient to guarantee that $\bar{\mathcal{I}}_i{}^{\circ} \ldots \bar{\mathcal{I}}_q{}^{\circ}$ are invariant and therefore are output stationary with $f_1 \ldots f_m$.

Necessity For every $D \in D^*$, i) and ii) above define a unique set of vectors $\xi_1 \ldots \xi_m$ and a unique $\xi_{\mathcal{N}}{}^{\circ}$. Assume that for every $D \in D^*$ and the associated $\xi_{\mathcal{N}}{}^{\circ}$ there is some integer $\mu \leq q$ such that:

$$iv) \quad A\xi_{h\mu}{}^{\circ} \not\equiv A\xi_{\mathcal{N}}{}^{\circ} \qquad (\text{mod } \bar{\mathcal{I}}_{\mu}{}^{\circ} \bigcup_{k \notin I_\mu} \bar{R}_k)$$

By the same argument leading to vi):

$$A\xi_{h\mu}{}^{\circ} \not\equiv -\sum_{k \in I_\mu} \xi_k \qquad (\text{mod } \bar{\mathcal{I}}_{\mu}{}^{\circ})$$

Substitution of the above into iii) then yields:

$$(A\text{-}DC)^{\beta_\mu} h_\mu \not\equiv \phi \qquad (\mathrm{mod}\ \bar{\mathcal{J}}_\mu^\circ)$$

and therefore $\bar{\mathcal{J}}_\mu^\circ$ is not output stationary with $f_1 \ldots f_m$. Since the procedure exhausts D^β, $\bar{\mathcal{J}}_1^\circ \ldots \bar{\mathcal{J}}_q^\circ$ cannot be made mutually output stationary with $f_1 \ldots f_m$. Δ

Lemma 3.4.3 states a necessary and sufficient condition which $\bar{\mathcal{J}}_1^\circ \ldots \bar{\mathcal{J}}_q^\circ$ (and therefore $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$) must satisfy in order for mutual output stationarity with $f_1 \ldots f_m$ to be possible. If $\xi_{\mathcal{K}}^\circ$ exists and it is know, the sufficiency proof of the lemma furnishes an algorithm for determining a satisfactory D. The lemma does not give any indication of the number of eigenvalues of $A\text{-}DC$ which can be almost arbitrarily assigned, however. The answer to this problem requires that a new subspace $\bar{\mathcal{J}}_{\mathcal{K}}^\circ$ be determined. With $h_{\mathcal{K}}^\circ$ defined by:

$$h_{\mathcal{K}}^\circ = \sum_{k \in I_{\mathcal{K}}} f_k + \xi_{\mathcal{K}}^\circ \qquad (3.4.36)$$

$\bar{\mathcal{J}}_{\mathcal{K}}^\circ$ is defined to be the detection space for $h_{\mathcal{K}}^\circ$.

Since $\xi_{\mathcal{K}}^\circ \in \bigcup_{k \in I_{\mathcal{K}}} \mathcal{R}_k$, comparison with (3.4.27) verifies that the definition of $h_{\mathcal{K}}^\circ$ is identical to that of $h_1^\circ \ldots h_q^\circ$. Lemma 3.4.3 also requires that $h_{\mathcal{K}}^\circ$ and h_i° be equivalent:

$$\beta_i \ h_i^\circ \equiv h_{\mathcal{K}}^\circ \qquad (\mathrm{mod}\ \mathcal{J}_i \bigcup_{k \notin I_i} \bar{\mathcal{R}}_k) \qquad i = 1 \ldots q$$

Theorem 3.4.4 establishes the fundamental nature of $\bar{\mathcal{J}}_{\mathcal{K}}^\circ$; $\bar{\mathcal{J}}_1^\circ \ldots \bar{\mathcal{J}}_q^\circ$ are mutually output stationary with $f_1 \ldots f_m$ if and only if $\bar{\mathcal{J}}_{\mathcal{K}}^\circ$ is output stationary with $f_1 \ldots f_m$. Once $\mathcal{J}_{\mathcal{K}}^\circ$ is found, this reduces the mutual output stationarity problem to consideration of output stationarity of a single subspace -- and this problem was treated fully in Section 3.4.1. The proof of Theorem 3.4.4 utilizes the fact that partitions of the form outlined in (3.4.34) and (3.4.35) have already been used to define smaller sets than $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ where possible.

Theorem 3.4.4 The subspaces $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q \subset \bigcup_{k \in I_{\mathcal{K}}} \bar{\mathcal{R}}_k$ are mutually output stationary with $f_1 \ldots f_m$ if and only if $\bar{\mathcal{J}}_{\mathcal{K}}^{\circ}$ is output stationary with $f_1 \ldots f_m$.

Proof

Sufficiency If $\bar{\mathcal{J}}_{\mathcal{K}}^{\circ}$ can be made output stationary with $f_1 \ldots f_m$, Theorem 3.4.2 states that the eigenvalues of A-DC associated with $\bar{\mathcal{J}}_{\mathcal{K}}^{\circ}$ can be assigned almost arbitrarily. In particular, the proof of Theorem 3.4.2 verifies that D can be chosen such that:

$$\text{i)} \quad (\text{A-DC}) \sum_{k \in I_{\mathcal{K}}} f_k = -\text{A}\, \xi_{\mathcal{K}}^{\circ}$$

since this implies that:

$$(\text{A-DC})\, h_{\mathcal{K}}^{\circ} = \phi$$

But i) above is identical to i) and ii) in the proof of Lemma 3.4.4. Therefore, $\bar{\mathcal{J}}_1^{\circ} \ldots \bar{\mathcal{J}}_q^{\circ}$ are invariant for this D and $\bar{\mathcal{J}}_1^{\circ} \ldots \bar{\mathcal{J}}_q^{\circ}$ (and therefore $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$) are output stationary with $f_1 \ldots f_m$.

Necessity Given that $\bar{\mathcal{J}}_1^{\circ} \ldots \bar{\mathcal{J}}_q^{\circ}$ are invariant, Lemma 3.4.4 and the definition of $h_{\mathcal{K}}^{\circ}$ require:

$$\text{i)} \quad h_{\mathcal{K}}^{\circ} \equiv \beta_i\, h_i^{\circ} \qquad (\text{mod } \mathcal{J}_i \bigcup_{k \notin I_i} \bar{\mathcal{R}}_k) \qquad i = 1 \ldots q$$

To prove that $\bar{\mathcal{J}}_{\mathcal{K}}^{\circ}$ is invariant it is sufficient to show:

$$(\text{A-DC})\, h_{\mathcal{K}}^{\circ} \in \bar{\mathcal{J}}_{\mathcal{K}}^{\circ}$$

Assume $\bar{\mathcal{J}}_{\mathcal{K}}^{\circ}$ is not invariant, Then Corollary 1 of Theorem 3.2.1 insures that there must exist some smallest integer μ such that:

$$\text{ii)} \quad \text{C}\, (\text{A-DC})^{\mu} h_{\mathcal{K}}^{\circ} = \sum_{k \in I_u} \text{C}\, f_k$$

where the non-empty set I_u satisfies $I_u \neq I_{\mathcal{N}}$. Since $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_m$ are invariant and $h_{\mathcal{N}}^\circ \in \bigcup_{k \in I_{\mathcal{N}}} \bar{\mathcal{R}}_k$, it follows that $I_u \subset I_{\mathcal{N}}$. Equations i) and ii) together imply that:

$$\text{iii)} \quad C\left(A\text{-}DC\right)^{\mu} \beta_i h_i^{\circ} = \sum_{k \in I_u \cap I_i} C f_k$$

Since $\bar{\mathcal{J}}_1^{\circ} \ldots \bar{\mathcal{J}}_q^{\circ}$ are invariant, iii) therefore requires that each of $I_1 \ldots I_q$ satisfy either $I_u \cap I_i = \phi$ or $I_i \subset I_u$. But by assumption (see (3.4.34) and (3.4.35)) this is possible only if $I_u = I_{\mathcal{N}}$ or $I_u = \phi$. In either case, this contradicts the assumptions associated with ii). It follows that $\bar{\mathcal{J}}_{\mathcal{N}}^{\circ}$ is output stationary with $f_1 \ldots f_m$. Δ

Theorem 3.4.4 is the immediate goal which had been sought. The mutual output stationarity problem for $f_1 \ldots f_m$ has been reduced to a simple output stationarity problem for $\bar{\mathcal{J}}_{\mathcal{N}}^{\circ}$ where $\bar{\mathcal{J}}_{\mathcal{N}}^{\circ} \in \bigcup_{k \in I_{\mathcal{N}}} \bar{\mathcal{R}}_k$. A complete set of algorithms for the single subspace problem were developed in Section 3.4.1. To fully appreciate the importance of $\bar{\mathcal{J}}_{\mathcal{N}}^{\circ}$, it is convenient to introduce an extended definition of output stationary equivalence in contrast to Definition 3.4.3. The new definition provides the basis for the promised generalization of Theorem 3.4.3.

Definition 3.4.4 Let $\bar{\mathcal{J}}_{\mathcal{N}}^{\circ}$ be the detection space for $h_{\mathcal{N}}^{\circ}$, where $h_{\mathcal{N}}^{\circ}$ is defined by Lemma 3.4.4 and (3.4.36). A subspace $\bar{\mathcal{J}}_w \in \bigcup_{k \in I_w} \bar{\mathcal{R}}_k$ is extended output stationary equivalent to $\bar{\mathcal{J}}_{\mathcal{N}}^{\circ}$ if its normal subspace $\bar{\mathcal{J}}_w^{\circ}$ satisfies:

$$\bar{\mathcal{J}}_{\mathcal{N}}^{\circ} \subset \bar{\mathcal{J}}_w^{\circ} \qquad (\text{mod} \bigcup_{k \notin I_w} \bar{\mathcal{R}}_k)$$

Clearly Definition 3.4.4 is a generalization of Definition 3.4.3. The essential difference is that it is not necessary for $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ to be contained in the same I_i in order for them to be extended output stationary equivalent. All that is required is that it must be possible to define $\bar{\mathcal{J}}_{\mathcal{N}}^{\circ}$ for them since that is equivalent to the congruence constraint of the definition. If $\bar{\mathcal{J}}_w$ satisfies Definition 3.4.4, $\bar{\mathcal{J}}_w^{\circ}$ could be added to the set $\bar{\mathcal{J}}_1^{\circ} \ldots \bar{\mathcal{J}}_q^{\circ}$ in Lemma 3.4.4 and the same $\xi_{\mathcal{N}}^{\circ}$ could be found.

If $\bar{\mathcal{J}}_w$ is not extended output stationary equivalent to $\bar{\mathcal{J}}_\chi{}^\circ$, then if $\bar{\mathcal{J}}_w{}^\circ$ were added to the set $\bar{\mathcal{J}}_1{}^\circ \ldots \bar{\mathcal{J}}_q{}^\circ$, no solution $\xi_\chi{}^\circ$ would exist. The following corollary completes the theory of this section.

Corollary 1 Let $\bar{\mathcal{J}}_\chi{}^*$ be the class of detection spaces which are extended output stationary equivalent to $\bar{\mathcal{J}}_\chi{}^\circ$. The elements of $\bar{\mathcal{J}}_\chi{}^*$ are mutually output stationary with $f_1 \ldots f_m$ if and only if $\bar{\mathcal{J}}_\chi{}^\circ$ is output stationary with $f_1 \ldots f_m$.

Proof

Sufficiency is given by the sufficiency proof of Theorem 3.4.4; necessity follows since $\bar{\mathcal{J}}_\chi{}^\circ \in \bar{\mathcal{J}}_\chi{}^*$.

Δ

3.5 Necessary and Sufficient Condition for a Detectable Set

The following theorem belongs properly in Section 3.3. It is presented here because its proof requires consideration of output stationarity as defined in Section 3.4. One of the weaknesses of detection theory to this point is that it has dealt in sufficient conditions for mutual detectability, but has not established a necessary and sufficient condition. Beard derived sufficient conditions for a set of event vectors to be mutually detectable for an n-dimension system model if they are output separable (Theorem 3.3.1). Theorem 3.3.4 carries the argument a step further by demonstrating that all output separable sets of event vectors can be made mutually detectable by means of state space augmentation. The final step in the progression is provided by Theorem 3.5.1: A set of event vectors can be made mutually detectable with state space augmentation if and only if they are output separable. Although the proof is straightforward, it could not be accomplished without the introduction of output stationarity and the additional insight into the structure of linear operators that it provides.

Theorem 3.5.1 and the results of Section 3.6 on linear observers provide the first definitive statement that the capability of a single detection filter for monitoring failures in a system (A, C) with $m \leq n$ is limited. If a system has m independent sensors, the output errors

generated by each of m output separable event vectors $f_1 \ldots f_m$ can be made unidirectional along the respective $Cf_1 \ldots Cf_m$ for the (possibly augmented) system (A-DC, C) with complete freedom in assigning the eigenvalues of A-DC , but this is the maximum number of such events. The detection spaces $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_m$ are fixed by the choice of $f_1 \ldots f_m$. If by chance there is an additional event vector ξ_f contained in some $\bar{\mathcal{R}}_k$, $1 \leq k \leq m$, it also produces a unidirectional output error along f_k, but it is not possible to discriminate between the failures corresponding to ξ_f and f_k on the basis of output error direction information alone.

The only possibility for increasing the number of failures which can be monitored by a single detection filter is to accept a less stringent definition of detectability than Definition 2.1.1. Definition 3.4.1 for output stationarity provides a mechanism for constraining the output produced by $f_1 \ldots f_m$ to be unidirectional along $Cf_1 \ldots Cf_m$ respectively, and for associating with every other direction Ch_i in the output space an event vector $h_i \in \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_m$ such that h_i generates unidirectional output along Ch_i for A-DC . The event vectors h_i cannot be chosen arbitrarily, but subject to certain restrictions stated in Section 3.4, a set of event vectors $h_1 \ldots h_q$ corresponding to meaningful failures in (A, C) may be chosen. The cost of defining a failure detection system for $f_1 \ldots f_m$, $h_1 \ldots h_q$ is a partial sacrifice of the ability to assign the eigenvalues of A-DC.

The proof of Theorem 3.5.1 requires a result that was alluded to in Section 3.3, but never stated. Lemma 3.5.1 is a generalization of Corollary 1 of Theorem 3.2.1 for the detection space $\bar{\mathcal{R}}_{\mathcal{J}}$ as defined in Definition 3.3.3. For generality, the lemma is stated for $f_1 \ldots f_r$ where $r \leq m$.

Lemma 3.5.1 Let $\mathcal{J}$ be spanned by $f_1 \ldots f_r$ and assume $f_1 \ldots f_r$ are output separable. If (A, C) is observable and D is chosen such that $\bar{\mathcal{R}}_{\mathcal{J}}$ is invariant, no vector $\xi \notin \bar{\mathcal{R}}_{\mathcal{J}}$ can produce an output for the system (A-DC, C) which is restricted to $C\mathcal{J}$.

Proof

Lemma 3.2.1 and an extended version of Lemma 3.2.2 guarantee that D can be chosen such that $\bar{\mathcal{R}}_{\mathcal{J}}$ is invariant. A proof identical to that of Lemma 3.2.3 shows that (A-DC, C) is observable (mod $\bar{\mathcal{R}}_{\mathcal{J}}$) if and only if (A, C) is observable. The proof of Lemma 3.5.1 is then identical to that of Corollary 1 of Theorem 3.2.1 with the above fact replacing Lemma 3.2.3.

Theorem 3.5.1 is restricted to the case where no two of $f_1 \ldots f_r$ are detection equivalent. This is the general case since if f_i and f_j are detection equivalent, then $\bar{\mathcal{R}}_i = \bar{\mathcal{R}}_j$ and only one of f_i, f_j need be included in the set of event vectors.

Theorem 3.5.1 Assume no two event vectors in $f_1 \ldots f_r$ are detection equivalent. The vectors $f_1 \ldots f_r$ can be made mutually detectable with state space augmentation if and only if they are output separable.

Proof

Sufficiency Direct consequence of Theorem 3.3.4.

Necessity Let $\mathcal{J}$ be spanned by $f_1 \ldots f_r$ and assume $d(\mathcal{J}) = p < r$. Choose p vectors from the set which are output separable. For convenience let them be $f_1 \ldots f_p$. Also, let $\mathcal{J}_p^c$ be the space spanned by $f_1 \ldots f_p$ and let $\bar{\mathcal{R}}_{\mathcal{J}_p^c}$ be the detection space of $\mathcal{J}_p^c$. Augment the state space if necessary so that $f_1 \ldots f_p$ are mutually detectable. Two cases need be considered.

i) Assume $f_k \notin \bar{\mathcal{R}}_{\mathcal{J}_p^c}$ for some $p < k \leq r$. For every D such that $f_1 \ldots f_p$ generate unidirectional output errors along the respective $Cf_1 \ldots Cf_p$, the output error generated by f_k can only be constrained to the output space spanned by the vectors in the sequence Cf_k, $C(A-DC)f_k \ldots$ By construction, $Cf_k \in C\mathcal{J}_p^c$. But Lemma 3.5.1 requires that there exist some integer μ such that $C(A-DC)^{\mu} f_k \notin C\mathcal{J}_p^c$ and therefore f_k cannot be constrained to generate unidirectional output errors. Therefore $f_1 \ldots f_p$, f_k are not mutually detectable.

ii) If all f_k for $p < k \leq r$ satisfy $f_k \in \bar{\mathcal{R}}_{\mathcal{J}_p^c}$, then since no two of $f_1 \ldots f_r$ are detection equivalent, for each f_k there must be some minimal index set $I_k \in \{1 \ldots p\}$ with two or more elements such that $f_k \in \bigcup_{j \in I_k} \bar{\mathcal{R}}_j$. By Theorem

3.4.2, the output errors produced by $f_1 \ldots f_p$, f_k can be constrained to the respective directions $Cf_1 \ldots Cf_p$, Cf_k only if some of the eigenvalues associated with $\bar{R}_1 \ldots \bar{R}_p$ have a multiplicity equal to the number of elements in I_k. Therefore Definition 2.1.1 cannot be satisfied for $f_1 \ldots f_p$, f_k for any $p < k \leq r$.

Since the dimension of the augmented state space may be chosen arbitrarily subject to the minimal dimension of Theorem 3.3.4, i) and ii) imply sufficiency.

Δ

3.6 Failure Detection Theory and Linear Observers

The purpose of a linear observer is to reconstruct the state of a system from a partial measurement of the state by utilizing prior knowledge of the dynamics of the system. In instances in which the measurements are corrupted by disturbances, the observer may assume the role of a state estimator which is optimal in some statistical sense. A general survey of observer theory is provided by Luenberger (1964 and 1971) and specific interpretations are discussed in Wonham (1970), Leondes and Novak (1972), and Tse (1973) among others.

Any state estimator is an observer; which implies that the failure detection system presented in this thesis is an observer. The interest in observers with regard to failure detection theory is due to the difficulties involved with detection spaces. If m = n, every event vector is its own detection space and it is shown in Section 4.5 that failure detection theory is simplified tremendously. There is an intuitive argument suggesting that if an observer were used to reconstruct the system state, the reconstructed state vector could be used as an n-dimensional pseudo- measurement and the simplified theory of Section 4.5 could be applied. The argument is specious, as a careful examination of Theorem 3.5.1 would show. Nonetheless, it is instructive to consider the possibility in detail.

Only a cursory review of linear observer theory is necessary as preparation for Theorem 3.6.1. The basic premise of observer theory is that there exists a system (F, G, H):

$$\dot{z}(t) = F\, z(t) + G\, y(t) + TB\, u(t) \qquad (3.6.1)$$

for which z(t) is a constant linear transformation of x(t):

$$z(t) = Tx(t) \qquad : \text{given } z(0) = Tx(0) \tag{3.6.2}$$

The choice of T is not unique; however, it must satisfy:

$$\text{rk} \begin{bmatrix} C \\ \hline T \end{bmatrix} = n \tag{3.6.3}$$

With T specified as a solution of (3.6.3), F and G are any solutions of:

$$TA - FT = GC \tag{3.6.4}$$

It is then possible to choose H and K from:

$$\left[K \,\vdots\, H\right] \begin{bmatrix} C \\ \hline T \end{bmatrix} = E_n \tag{3.6.5}$$

where E_n is the n x n identity matrix. With these constraints on (F, G, H) and K, the state estimate supplied by the observer is:

$$\hat{x}(t) = Hz(t) + Ky(t) \tag{3.6.6}$$

There is a certain amount of design freedom in the choice of an observer. The minimal order solution of (3.6.3) requires T to be a (n-m) x n matrix. There are corresponding solutions of (3.6.4) for both F (n-m)x(n-m) and G (n-m) x m. These solutions are not unique in that sufficient freedom remains to allow the eigenvalues of F to be freely chosen. If T is minimal order, the corresponding system (F, G, H) coupled with K is a minimal-order observer for (A, B, C).

The proposed detection system configuration is indicated in Figure 3.3. The failure detection system does not see the output from the system directly, but rather the reconstructed state vector $\hat{x}(t)$. It is assumed that the observer has no failure modes so that any aberration in $\hat{x}(t)$ is due to a change in y(t) plus the generated error from passing y(t) through (F, G, H). For convenience, it is assumed that the observer is minimal order. The system-observer model is then represented by the triple (A_0, B_0, C_0):

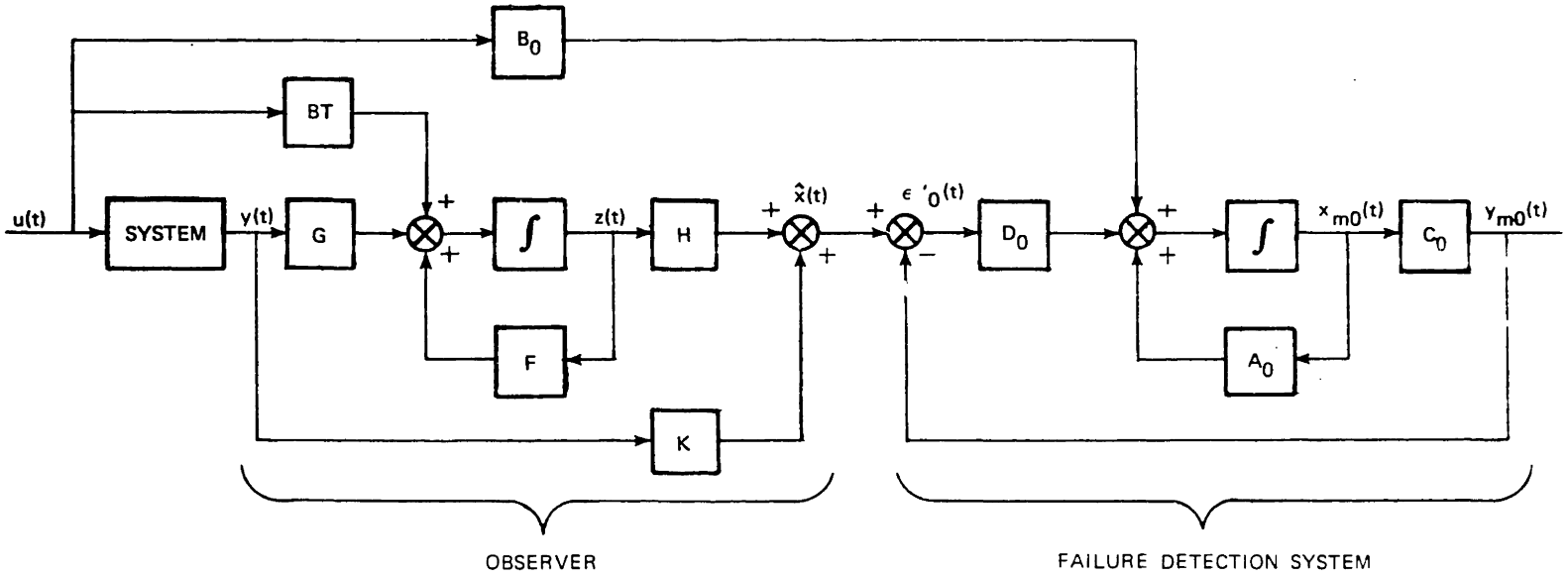


Figure 3.3 Observer prefilter to supply failure detection system with an estimate of the full state vector of the system.

$$C_0 = [KC \mid H] \qquad A_0 = \left[\begin{array}{c|c} A & 0 \\ \hline GC & F \end{array}\right] \qquad B_0 = \left[\begin{array}{c} B \\ \hline TB \end{array}\right] \qquad (3.6.7)$$

This is the system model upon which the failure detection system must operate.

Just as detection spaces were defined for (A, C) by Definition 3.1.2, they can be defined in the augmented state space for (A_0, C_0). The following lemma demonstrates a trivial but useful relationship between the two.

Lemma 3.6.1 Let the system model (A_0, C_0) be defined on $\mathcal{S}^{2n-m} = \mathcal{S}^n \oplus \mathcal{S}^{n-m}$ with A_0 specified by:

i) $A : \mathcal{S}^n \to \mathcal{S}^n$
ii) $GC : \mathcal{S}^n \to \mathcal{S}^{n-m}$
iii) $F : \mathcal{S}^{n-m} \to \mathcal{S}^{n-m}$
iv) $\phi : \mathcal{S}^{n-m} \to \mathcal{S}^n$

and C_0 specified by:

v) $KC : \mathcal{S}^n \to \mathcal{S}^m$
vi) $H : \mathcal{S}^{n-m} \to \mathcal{S}^{n-m}$

If $f_0 \in \mathcal{S}^n$ is an event vector for (A_0, C_0), then $\bar{\mathcal{R}}_0 \subset \mathcal{S}^n$.

Proof

Since the observer is minimal order, $d(H\mathcal{S}^{n-m}) = n-m$ and $\eta(C_0) \subset \mathcal{S}^n$. The desired result follows from Definition 3.1.2 since $\bar{\mathcal{R}}_0 = \mathcal{R}_0 \oplus f_0$ and $\mathcal{R}_0 \subset \eta(C_0)$.

Δ

If the system model is as defined in Lemma 3.6.1, all event vectors associated with the system are constrained to $\mathcal{S}^n$. If no failures can occur in the observer, all event vectors for the system-observer combination are contained in $\mathcal{S}^n$. Theorem 3.6.1 verifies that the observer does not increase the number of failures associated with (A, C) which can be monitored by a single failure detection system.

Theorem 3.6.1 Let $f_1 \ldots f_r$ be a set of event vectors for (A, C) and $f_{10} \ldots f_{r0}$ be the corresponding set for (A_0, C_0). The properties of:

i) output separability

ii) mutual detectability (with state space augmentation)

iii) mutual stationarity

are true of $f_{10} \ldots f_{r0}$ with respect to (A_0, C_0) only if the corresponding properties are true of $f_1 \ldots f_r$ with respect to (A, C).

Proof

i) Clearly if $\mathcal{J}$ and $\mathcal{J}_0$ are spanned respectively by $f_1 \ldots f_r$ and $f_{10} \ldots f_{r0}$, then $d(C\mathcal{J}) \geq d(C_0\mathcal{J}_0)$ since $d(C_0\mathcal{J}_0) = d(KC\mathcal{J})$. Therefore, output separability of $f_{10} \ldots f_{r0}$ implies output separability of $f_1 \ldots f_r$.

ii) Assume state space augmentation has been used if necessary. By Theorem 3.5.1, mutual detectability of $f_{10} \ldots f_{r0}$ implies output separability of $f_{10} \ldots f_{r0}$. From i) and Theorem 3.5.1 again, $f_1 \ldots f_r$ are mutually detectable.

iii) Suppose $f_{10} \ldots f_{r0}$ and $f_1 \ldots f_r$ are mutually detectable with respect to (A_0, C_0) and (A, C) respectively and let h_1 and h_{10} be additional event vectors similar to $f_1 \ldots f_r$ and $f_{10} \ldots f_{r0}$. If $f_{10} \ldots f_{\ell 0}$ is the smallest subset of $f_{10} \ldots f_{r0}$ such that $\mathcal{J}_{10} \subset \bar{\mathcal{R}}_{10} \oplus \ldots \bar{\mathcal{R}}_{\ell 0}$, Theorem 3.4.1 requires that:

$$\beta_{10}\, C_0 h_{10} = \alpha_{10} C_0 f_{10} + \ldots \alpha_{\ell 0}\, C_0 f_{\ell 0}$$

with $\alpha_{10} \ldots \alpha_{\ell 0}$ nonzero if f_{10} is to be output stationary with $f_{10} \ldots f_{r0}$. By the argument used in i), this implies:

$$\beta_1\, Ch_1 = \alpha_1\, Cf_1 + \ldots \alpha_\ell\, Cf_\ell$$

with $\alpha_1 \ldots \alpha_\ell$ nonzero. From Lemma 3.5.1 it follows that $\mathcal{J}_1 \subset \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_\ell$ and h_1 can be made output stationary with $f_1 \ldots f_r$.

Δ

A failure detection system based upon (A_0, C_0) as a system model does precisely what it is designed to do. If a system failure occurs which would affect z(t), the detection system senses the failure and predicts the effect it would have on z(t). The predicted change is then added to $y_{m0}(t)$ (which is now n-dimensional) to insure that the error in the estimate:

$$\epsilon_0'(t) = \hat{x}(t) - y_{m0}(t) \qquad (3.6.8)$$

remains unidirectional. The cascaded observer does not hinder failure detection operations, but it does not help them.

CHAPTER 4

DETECTION FILTER DESIGN -- DETERMINISTIC SYSTEMS

The presentation of failure detection theory has been within the context of a general theoretical development. The motivation for the proposed solution was given in Chapter 2 in terms of an abstraction of the properties of a general reference model (A, B, C). In Chapter 3 a general solution to the failure detection problem was derived from a study of linear operators. In the present chapter the results of Chapter 3 are translated into an orderly design procedure. The discussions are limited to deterministic systems; consideration of stochastic systems is deferred until Chapter 5. An understanding of Chapter 3 is not necessary to use the design procedure outlined in this chapter, but a comprehension of the ideas introduced in Sections 2.1 and 3.1 is mandatory.

The theory of Chapter 3 states the failure detection properties of a system explicitly as a function of the reference model (A, B, C) and the set of event vectors associated with the possible failures in the system. The choice of the reference model has a profound effect upon the flexibility of the failure detection system. To a certain extent this sensitivity is negated by the results of Section 3.3.4 which provide a constructive means for improving the reference model; however, the modeling of failures still requires an acceptable initial choice of a reference model. With an acceptable reference model, the general failure model developed in Section 4.1 permits the event vector (or vectors) associated with any system failure to be determined. A number of important factors in choosing the reference model and in modeling failures are discussed and illustrated in Section 4.1.

The formal presentation of the design procedure is made in Section 4.2 for the general case of a partially measurable state vector. The procedure is applied in Section 4.3 to the design of a failure detection system.

An important canonical form for the failure detection problem is introduced in Section 4.4. For a system stated in this form many of the computations required in the design procedure are trivial. Also, once a detection filter is selected, the closed loop system model is decoupled into a set of subsystems with one potential failure associated with each subsystem. This canonical form is the basis of the analysis of stochastic systems in Chapter 5. Examples of the canonical form are given in Section 4.5 for the degenerate case of a fully measurable state vector and in Chapter 6 for the general case.

The analytical results of Chapter 3 are for a continuous, time invariant system with continuous measurements. The theory is equally valid for discrete systems. It can also be used to design a detection filter for continuous system with discrete measurements, but the failure detection system performance is a function of the sampling interval for the measurements. These results, along with limited observations on time-varying systems, are discussed in Section 4.6.

The intent of this chapter is to present a reasonable procedure for designing a failure detection system. The discussion of design considerations is more informal than that of Chapter 3. Even when referring to particular theorems of Chapter 3, results are stated in the more familiar matrix notation. Attention is restricted to the design of a single failure detection system. If the number of failures to be monitored is large, two or more failure detection systems may be required. The material of this section can be applied to the design of each of the detection filters. The detection law required to interpret the output of multiple failure detection systems is developed in Section 5.3.

4.1 Modeling

Modeling is considered in two contexts in this section -- the choice of the reference model for the system in the nominal case and a characterization of the system in the event of a failure. The

modeling problems are discussed in the reverse of the order in which they would normally be considered. The general failure model is developed in Section 4.1.1 and illustrated by the spinning cylinder example in Section 4.1.2. The concepts introduced in these sections permit some specific comments on choosing a reference model to be made in Section 4.1.3.

4.1.1 General Failure Model

The intention of this subsection is to generalize the intuitively motivated failure models presented in Section 2.1. It is a simple matter to model a failure in terms of event vectors, but the emphasis is on techniques for reducing a failure model with multiple event vectors to the form of either a controller failure model (2.1.7) or a sensor failure model (2.1.8). As mentioned in Section 2.1, the usefulness of failure detection theory rests upon the fact that failures other than controller failures and sensor failures can frequently be modeled by one of these two special failure models. This subsection presents a general algorithm for modeling system failures in the desired form.

The simplest failures to design failure detection systems for are control element failures. The system in the unfailed case satisfies:

$$\dot{x}(t) = Ax(t) + Bu(t) \; : \; x(0) \text{ given} \qquad (4.1.1)$$

$$y(t) = Cx(t)$$

where u(t), x(t) and y(t) are p-dimensional, n-dimensional and m-dimensional vectors respectively. The relationship of u(t) to B is clearer if (4.1.1) is re-written in summation form:

$$\dot{x}(t) = Ax(t) + \sum_{i=1,p} b_i u_i(t) \; : \; x(0) \text{ given} \qquad (4.1.2)$$

where b_i is the i^{th} column of B and $u_i(t)$ is the i^{th} component of u(t). Each term $b_i u_i(t)$ constitutes a control channel for the system. The p components of u(t) are commands issued to the

p control elements of the system; the columns of B describe the manner in which an excitation of the control elements affects the system state.

If a failure occurs in the i^{th} control channel, that failure must result in a change in the $b_i u_i(t)$ term in (4.1.2). The general failure model for the failure is:

$$\dot{x}(t) = Ax(t) + \sum_{i=1,p}[b_i u_i(t) + \Delta b_i u(t) + b_i \Delta u_i(t)] : x(0) \text{ given}$$

$$y(t) = Cx(t) \qquad (4.1.3)$$

The $\Delta u_i(t)$ term reflects the fact that the control command received by the control element may differ from the intended command. As an example, if the failure results in no input command being received by the i^{th} control element, the failure can be modeled by:

$$\Delta u_i(t) = -u_i(t) \qquad (4.1.4)$$

$$\Delta b_i = 0$$

When $\Delta u_i(t)$ is a function of $u_i(t)$ the failure is defined to be control-dependent; otherwise, it is control-independent. As examples of the latter, if a control element receives a noisy command, $\Delta u_i(t)$ may be a random process; if the control element receives a biased command, $\Delta u_i(t)$ may be a constant, etc. The event vector associated with any failure which can be modeled by $\Delta u_i(t) b_i$ is simply b_i.

The $\Delta b_i(t)$ term in (4.1.3) is necessary for modeling failures in which the control element no longer responds properly to a received command. In this case, $\Delta b_i(t)$ is the event vector. It is important to note that $\Delta b_i(t)$ can be time-varying and hence that B(t) in the failed system can be time-varying. Failure detection theory requires the system to be time invariant in the unfailed state, but requires neither time invariance nor linearity in the failed system. An important special case, however, is when $\Delta b_i(t)$ is used to indicate a change in the effectiveness of the control element:

$$\Delta b_i(t) = k(t)b_i \qquad (4.1.5)$$

If k(t) is -1, (4.1.5) is used to model a control channel which is completely inoperative, i.e. the i^{th} control is not applied. Other values of k(t) indicate a change in the gain in the i^{th} control channel. Most control channel failures can be modeled using (4.1.5) and some $\Delta u_i(t)$:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + b_i(k(t)u_i(t) + \Delta u_i(t)): x(0) \text{ given} \\ y(t) &= Cx(t) \end{aligned} \qquad (4.1.6)$$

b_i is the event vector for all such failures.

There are certain component failures which are classically thought of as control element failures, but which cannot be modeled by (4.1.3). An example is a change in the time constant of a hydraulic actuator. Such failures must technically be modeled as system dynamics changes since they involve a change in at least one coefficient of the state matrix A. The example in Section 4.1.2 suggests that in most cases such failures modes can be modeled in terms of a single event vector, and that the event vector is frequently the same as the appropriate column of B.

The next failure model to be considered is the model for a system dynamics change. There are two types of dynamics changes to be considered -- those that involve changes in the state matrix A and those which do not. Restricting attention for the moment to changes in A, the general failure model is:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + \Delta Ax(t) : x(0) \text{ given} \\ y(t) &= Cx(t) \end{aligned} \qquad (4.1.7)$$

If a failure can be modeled as a change in only the a_{ij}^{th} element of A it is simple to show that (4.1.7) reduces to:

$$\dot{x}(t) = Ax(t) + Bu(t) + e_{ni}\Delta a_{ij}(t)\, x_j(t) : x(0) \text{ given} \qquad (4.1.8)$$

$$y(t) = Cx(t)$$

where e_{ni} is the i^{th} unit vector in the state space. Failures modeled by (4.1.8) require only the single event vector e_{ni} and therefore are modeled in the same manner as the control element failures in (4.1.6). The change in the time constant of a hydraulic actuator mentioned earlier is an example of a failure which can be modeled by (4.1.8).

There are two additional instances in which a change in A can be modeled with a single event vector. The first is if only the elements of the i^{th} row of A change. Physically, this corresponds to the case where the derivative of the i^{th} state variable is altered by the failure. Equation (4.1.7) then reduces to:

$$\dot{x}(t) = Ax(t) + Bu(t) + e_{ni}\Big(\sum_{j=1,n} \Delta a_{ij}(t)\, x_j(t)\Big) : x(0) \text{given} \qquad (4.1.9)$$

$$y(t) = Cx(t)$$

As was the case with (4.1.8) the event vector associated with this failure is e_{ni}. If only a single column of A changes, and if the elements of the non-zero column of ΔA are constant (or at least have a constant ratio), the event vector can be defined:

$$f_j = \sum_{i=1,n} e_{ni}\, \Delta a_{ij} \qquad (4.1.10)$$

and the failure model is:

$$\dot{x}(t) = Ax(t) + f_j\, x_j(t) \quad : x(0) \text{ given} \qquad (4.1.11)$$

$$y(t) = Cx(t)$$

This corresponds to the case where the derivatives of several state variables change in similar manners. Failures for which

(4.1.9) and (4.1.11) are suitable models are illustrated in Section 4.1.2.

For an arbitrary ΔA, (4.1.7) can be expressed:

$$\dot{x}(t) = Ax(t) + Bu(t) + \sum_{i=1,n} \left(e_{ni} \sum_{j=1,n} \Delta a_{ij}(t)\, x_j(t)\right) \;:\; x(0) \text{ given} \qquad (4.1.12)$$

$$y(t) = Cx(t)$$

From the preceding paragraph, it is clear that (4.1.12) does not necessarily imply that the failure model requires n event vectors. If ΔA is constant, and:

$$\text{rk}\, \Delta A = \ell \qquad (4.1.13)$$

the manipulations illustrated by (4.1.9) - (4.1.11) can be used to express the failure model in terms of ℓ event vectors. If ΔA is time-varying, (4.1.13) defines a lower limit on the number of event vectors necessary to model the failure. The upper limit is defined by the number of non-zero rows in ΔA.

In the development of failure detection theory in Chapter 3, it became clear that it is difficult to detect a large number of event vectors with a single detection filter. Added to this is the problem that for arbitrary ΔA, the output error signal cannot be constrained to a single direction, but to a hyperplane with a dimension possibly equal to the number of event vectors in the failure model. Understandably, such error signals are more difficult to associate with their sources. These considerations suggest that a failure detection system may not provide a satisfactory failure detection capability for such failures.

The discussion to this point has concentrated upon dynamics changes which are described by changes in A. There is an important class of failures which are best described as dynamics changes, but which are independent of both A and the value of x(t). As an example, a change in the i^{th} component of x(t) can be modeled by:

$$\dot{x}(t) = Ax(t) + Bu(t) + e_{ni}\,\nu(t) \quad : x(0) \text{ given}$$

$$y(t) = Cx(t) \tag{4.1.14}$$

where $\nu(t)$ is an arbitrary time function of the jump and e_{ni} is the event vector. Such failures are referred to as state-independent dynamics changes. The presence of a single event vector implies that the state-independent dynamics change in (4.1.14) is modeled by the controller failure model (2.1.7).

The final failure model to be considered is that for a sensor failure. As was the case with dynamics changes, sensor failures can be divided into two categories -- those which are a function of the state vector and those which are not. The former include failures in which the functional relationship between the state and the sensor output changes. A failure of this type can be modeled by:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad : x(0) \text{ given} \tag{4.1.15}$$

$$y(t) = (C+\Delta C(t))\,x(t)$$

If the failure is in the i^{th} sensor, only the i^{th} row of $\Delta C(t)$ is non-zero and (4.1.15) reduces to:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad : x(0) \text{ given} \tag{4.1.16}$$

$$y(t) = Cx(t) + e_{mi}\left(\sum_{j=1,n} \Delta\, c_{ij}(t)\, x_j(t)\right)$$

Scale factor changes and hard sensor failures are the most common examples of sensor failures which can be modeled by (4.1.16).

The failure model for a state-independent sensor failure requires a single additive term in the output equation. A failure of this type in the i^{th} sensor can be modeled by:

$$\dot{x}(t) = Ax(t) + Bu(t) \; : \; x(0) \text{ given} \qquad (4.1.17)$$

$$y(t) = Cx(t) + e_{mi}\, n_{ci}(t)$$

where $n_{ci}(t)$ is some time function. Sensor biases and random drifts are typical failures which can be modeled by (4.1.17). An important difference between sensor failure models and previous failure models is that there are no event vectors in (4.1.16) and (4.1.17) -- the vector e_{mi} appears in the output equation rather than the state equation. This difference requires sensor failures to be handled differently from other failures.

The failure models developed in this section are all of the form:

$$\dot{x}(t) = Ax(t) + Bu(t) + f_i\, n_i(t) \; : x(0) \text{ given} \qquad (4.1.18)$$

$$y(t) = Cx(t) + e_{mi}\, n_{ci}(t)$$

where for notational convenience only a single event vector is included in the state equation. The additive terms for the various failure types are summarized in Table 4.1. The event vectors associated with sensor failures in Table 4.1 have not been discussed yet. They are columns of the detection filter D and are a consequence of the additive terms in the output equations in (4.1.15) - (4.1.18).

The failure detection system is based upon a reference model which has the same general structure as a linear observer. The reference model dynamics are defined by:

$$\dot{x}_m(t) = Ax_m(t) + Bu(t) + D(y(t) - y_m(t)) \; : x(0) \text{ given} \qquad (2.1.2)$$

$$y_m(t) = Cx_m(t)$$

With the state estimation error and output estimation error defined:

$$\epsilon(t) = x(t) - x_m(t)$$

$$\epsilon'(t) = y(t) - y_m(t) \qquad (2.1.4)$$

	Additive term in state equation	Additive term in output equation	Event vector	Description
General failure model	$f_i\, n_i(t)$	$e_{mi}\, n_{ci}(t)$	$f_i, -d_i$	(4.1.18)
Control-dependent	$\Delta b_i\, u(t)$	0	Δb_i	(4.1.3), (4.1.6)
Controller failure	$b_i\, \Delta u_i(t)$	0	b_i	(4.1.3), (4.1.6)
Control-independent Controller failure	$b_i\, n_i(t)$	0	b_i	(4.1.3)
State-dependent dynamics change	$e_{ni}\, \Delta a_{ij}(t)\, x_j(t)$	0	e_{ni}	(4.1.7) - (4.1.12)
State-independent dynamics change	$\nu(t)\, e_{ni}$	0	e_{ni}	(4.1.14)
State-dependent sensor failure	0	$e_{mi}\, \Delta c_{ij}(t)\, x_j(t)$	$-d_i$	(4.1.15), (4.1.16)
State-independent sensor failure	0	$e_{mi}\, n(t)$	$-d_i$	(4.1.17)

Table 4.1 General catalog of system failure types

the error equations for the failure detection system estimates after a control element failure or dynamics change have a similar form to that of the appropriate failure model:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + f_i\, n_i(t) \;:\; \epsilon(0) \text{ given} \qquad (2.1.15)$$

$$\epsilon'(t) = C\epsilon(t)$$

The error equations corresponding to failure of the i^{th} sensor are different from the sensor failure model in that they are driven by an event vector:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) - d_i\, n_{ci}(t) \;:\; \epsilon(0) \text{ given} \qquad (2.1.16)$$

$$\epsilon'(t) = C\epsilon(t) + e_{mi}\, n_{ci}(t)$$

The algorithms developed in Chapter 3 use the event vectors in the error equations to chose D. Since the event vector in (2.1.16) is a column of D, this results in a circular argument. In order to eliminate this difficulty, it is necessary to replace (2.1.16) with error equations in which d_i does not appear explicitly.

Let f be any vector which satisfies:

$$e_{mi} = Cf \tag{4.1.19}$$

and consider the consequences of requiring D in (2.1.16) to be a detection filter for f. The discussion in Section 3.1 (and Lemma 3.2.2) motivated the fact that D must then satisfy an equation of the form:

$$(A-DC)\, f = \xi_f \tag{4.1.20}$$

where ξ_f is some vector in the detection space of f. With (4.1.19), (4.1.20) can be solved explicitly for d_i:

$$d_i = Af - \xi_f \tag{4.1.21}$$

It is apparent from (4.1.21) that requiring the failure detection system to detect the "failure" associated with f results in the i^{th} column of D being fixed and (2.1.16) no longer implies a circular argument. It is then possible to require the remaining columns of D to detect a second "failure" associated with $d_i\, n_{ci}(t)$, i.e. with $(Af - \xi_f)\, n_{ci}(t)$. To emphasize this requirement, (2.1.16) can be replaced by an error equation with two event vectors:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + (Af - \xi_f)\, n_{ci}(t) + f \tag{4.1.22}$$

$$\epsilon'(t) = C\epsilon(t)$$

Equation (4.1.22) associates the event vector pair $\{f,\ Af-\xi_f\}$ with a failure of the i^{th} sensor (see Section 4.1.2 for an example).

Two points with regard to (4.1.22) deserve further comment. The first is that utilization of two event vectors to characterize a sensor failure in general implies that the output estimation error generated by a sensor failure cannot be constrained to a single output direction. If Cf and $C(Af - \xi_f)$ are not colinear and if $C(Af - \xi_f)$ is not zero, the output error can only be constrained to the plane spanned by Cf and $C(Af - \xi_f)$. This non-unidirectionality of output errors generated by sensor failures is a general property of linear observers and was not brought about by the definition of f in (4.1.19).

The second point is that the only constraint on f is that it satisfy (4.1.19) and the only constraint on ξ_f is that it be in the detection space of f. Frequently, an f can be chosen which is an event vector for a control element failure or a dynamics change. When such an f exists, a failure detection system designed for $\{f, Af - \xi_f\}$ can detect and isolate both the particular component failure associated with f and a failure of the sensor. The component failure generates an output error which is unidirectional along Cf; the sensor failure generates an output error which is constrained to the plane spanned by Cf and $C(Af - \xi_f)$. Failure detection algorithms such as the majority rule voting law cannot discriminate between such failures.

The last difficulty with sensor failures concerns sensors which are linearly dependent. If $rkC < m$, the output of some sensor is a linear combination of the output of the remaining sensors. This fact introduces a certain ambiguity in determining sensor failures. As an example, if two sensors are identical, it is impossible to determine from the failure detection system output which of them has failed. The remedy involves an increase in the dimension of the state model.

Assume $rkC = m-2$ and the last two sensors are linearly dependent upon the first m-2 sensors. Denote by C_{m-2} the matrix formed from the first m-2 rows of C and let the last

two rows of C be c_{m-1} and c_m respectively. If the failure detection system is designed for the reference model $(\bar{A}, \bar{B}, \bar{C})$ where:

$$\bar{C} = \left[\begin{array}{c|cc} & 0 & 0 \\ C_{m-2} & \vdots & \vdots \\ & 0 & 0 \\ \hline c_{m-1} & 1 & 0 \\ c_m & 0 & 1 \end{array}\right] \quad \bar{A} = \left[\begin{array}{c|cc} & 0 & 0 \\ A & \vdots & \vdots \\ & 0 & 0 \\ \hline 0 \cdots\cdots 0 & 0 & 0 \\ 0 \cdots\cdots 0 & 0 & 0 \end{array}\right] \quad \bar{B} = \left[\begin{array}{c} B \\ \hline 0 \cdots 0 \\ 0 \cdots\cdots 0 \end{array}\right] \tag{4.1.23}$$

then, assuming the initial conditions on the last two elements of the augmented state vector $\bar{x}(t)$ are zero, $(\bar{A}, \bar{B}, \bar{C})$ is input-output equivalent to (A, B, C). Since the rows of $\bar{C}$ are linearly independent, no ambiguity in sensor failure detection for $(\bar{A}, \bar{B}, \bar{C})$ exists. A detection filter $\bar{D}$ can be chosen to assign all n+2 eigenvalues of $\bar{A}-\bar{D}\bar{C}$ and all event vectors can be defined on the enlarged state space, etc. The intuitive meaning of (4.1.23) is that, even though the added state vector components are always zero, their presence allows the failure detection system to use the last two elements in the state estimate $\bar{x}_m(t)$ to determine which sensor has failed. If one of the last two sensors fails, the corresponding added component to $\bar{x}_m(t)$ becomes non-zero.

Equation (4.1.23) is obviously generalizable to as many dimensions as necessary. The trick of adding observable but uncontrollable dynamics to the reference model suggests that the problem with detecting failures in linearly dependent sensors is more algebriac than physical. A similar state space augmentation was used in Section 3.3.4 to make unassignable eigenvalues of A-DC assignable.

4.1.2 An Example

The spinning cylinder example of Section 3.3.2 can be used to illustrate the comments on modeling in the previous subsection. For convenience, the basic mathematics of the problem are restated.

The problem is one of controlling the spin of a cylinder in a viscous medium. The dynamic response of the cylinder to an externally applied torque is given by:

$$\ddot{\varphi}(t) + \beta_r \dot{\varphi}(t) = \alpha_r T(t): \quad \varphi(0), \ \dot{\varphi}(0) \text{ given} \tag{4.1.24}$$

where T(t) is the applied torque. The torquing device is first-order and described by:

$$\dot{T}(t) + \beta_f T(t) = \alpha_1 \varphi_c(t) + T_f(t) \ : \quad T(0) \text{ given} \tag{4.1.25}$$

where $\varphi_c(t)$ is the commanded roll angle and $T_f(t)$ is a feedback torque to improve the system response. It is assumed that the rotational angle and the rotational rate of the cylinder are measured as $\varphi_m(t)$ and $\dot{\varphi}_m(t)$ respectively. The control law is given by:

$$T_f(t) = -\alpha_1 \varphi_m(t) - \alpha_2 \dot{\varphi}_m(t) \tag{4.1.26}$$

As in Section 3.3.2, it is assumed that the applied torque T(t) is measured. A block diagram for this system is given in Figure 3.1.

The open loop dynamics of the system do not include the feedback law (4.1.26) and can be described by the state vector:

$$x_o(t) = \begin{bmatrix} \varphi(t) \\ \dot{\varphi}(t) \\ T(t) \end{bmatrix}$$

and the triad (A_o, B_o, C_o):

$$C_o = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad A_o = \begin{bmatrix} 0 & 1 & 0 \\ 0 & -\beta_r & \alpha_r \\ 0 & 0 & -\beta_f \end{bmatrix} \quad B_o = \begin{bmatrix} 0 \\ 0 \\ \alpha_1 \end{bmatrix} \qquad (4.1.27)$$

Equation (4.1.27) is an acceptable model upon which to design a failure detection system. The failures of possible interest are failures in:

i) the attitude sensor

ii) the rate sensor

iii) the torquing device

iv) the torque measuring device

Since the control law is not included in the model, failures in the controller cannot be detected. Only the first two failures are discussed here. The complete failure detection system is considered in Section 4.5.

Assume for the moment that the sensor failures may be hard failures only. The relevant failure model is the state-dependent failure model of (4.1.15):

$$\begin{aligned} \dot{x}_o(t) &= A_o x_o(t) + B_o u_o(t) \\ y_o(t) &= C_o x_o(t) + \Delta C_o x_o(t) \end{aligned} \qquad (4.1.15)$$

The attitude and rate sensor failures can be described by choosing ΔC_o such that the appropriate sensor output becomes zero, i.e.:

$$\Delta C_o x_o(t) = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} x_o(t) = -e_{31}\varphi(t) \qquad (4.1.28)$$

for an attitude sensor failure and:

$$\Delta C_o x_o(t) = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix} x_o(t) = -e_{32}\dot{\varphi}(t) \quad (4.1.29)$$

for a rate sensor failure. In accordance with (2.1.16) (see Section 4.1.1), if $\epsilon_a(t)$ and $\epsilon'_a(t)$ denote the state estimation error and output estimation error generated by the attitude sensor failure, the error equations are:

$$\dot{\epsilon}_a(t) = (A_o - DC_o)\,\epsilon_a(t) + d_1\,\varphi(t) \;:\; \epsilon_a(0) = 0$$

$$\epsilon'_a(t) = C_o\,\epsilon_a(t) - e_{31}\,\varphi(t) \quad (4.1.30)$$

Similarly, if $\epsilon_r(t)$ and $\epsilon'_r(t)$ are the corresponding estimation errors generated by the rate sensor failure, then the error equations are:

$$\dot{\epsilon}_r(t) = (A_o - DC_o)\,\epsilon_r(t) + d_2\,\dot{\varphi}(t) \;:\; \epsilon_r(0) = 0$$

$$\epsilon'_r(t) = C_o\,\epsilon_r(t) - e_{32}\,\dot{\varphi}(t) \quad (4.1.31)$$

where d_1 and d_2 above are the first and second columns of D respectively.

The design algorithms developed in Chapter 3 do not permit columns of D to appear in the error state equations in the same manner as event vectors. An algorithm for replacing (4.1.30) and (4.1.31) with equivalent error equations without d_1 and d_2 was detailed in Section 4.1.1. Consider (4.1.30) as an example. The first step is to define an event vector f_a such that:

$$e_{31} = C_o f_a \quad (4.1.32)$$

i.e. the output direction associated with f_a must be e_{31}. Since C_o is the identity matrix, the unique solution is:

$$f_a = e_{31} \quad (4.1.33)$$

If D is required to be a detection filter for f_a, (4.1.21) states that d_1 is:

$$d_1 = A_o f_a + \alpha_a f_a \qquad (4.1.34)$$

$$= \alpha_a f_a$$

where α_a can be chosen arbitrarily (f_a is its own detection space).

With (4.1.34), the explicit appearance of d_1 in (4.1.30) is eliminated and, since $A_o f_a$ is zero, the output error generated by the attitude sensor failure is:

$$\dot{\epsilon}(t) = (A_o - D C_o)\, \epsilon_a(t) + \alpha_a \varphi(t) f_a : \epsilon_a(0) = 0 \qquad (4.1.35)$$

$$\epsilon'_a(t) = C_o \epsilon(t) - e_{31} \varphi(t)$$

But D was required to be a detection filter for f_a, and e_{31} is the output direction associated with f_a; therefore, $e'_a(t)$ is unidirectional. For failure detection purposes the additive term in the output equation is redundant -- an equivalent set of equations is:

$$\dot{\epsilon}_a(t) = (A_o - DC_o)\, \epsilon_a(t) + \alpha_a \varphi(t) f_a : \epsilon_a(0) = 0 \qquad (4.1.36)$$

$$\epsilon'_a(t) = C_o \epsilon(t)$$

f_a is the event vector associated with an attitude sensor failure and (4.1.36) is the equation which must be used with the algorithms in Chapter 3.

The attitude sensor failure is a special case in that D can be chosen such that the output error in (4.1.36) (hence in (4.1.35) and (4.1.30)) is unidirectional. In general, the output error

for a sensor failure can only be constrained to be planar. Consider (4.1.31) and the rate sensor failure. The algorithm from Section 4.1.1 again requires that an event vector f_r be found such that:

$$e_{32} = C_o f_r \qquad (4.1.37)$$

and again there is a unique solution. If D is required to be a detection filter for f_r, however, d_2 is the sum of two linearly independent vectors:

$$\begin{aligned} d_2 &= A_o f_r + \alpha_r f_r \\ &= (e_{31} - \beta_r e_{32}) + \alpha e_{32} \end{aligned} \qquad (4.1.38)$$

where α is arbitrary. With d_2 satisfying (4.1.38), a set of equations equivalent to (4.1.31) for failure detection purposes is:

$$\dot{\epsilon}_r(t) = (A_o - D C_o)\, \epsilon_r(t) + \dot{\varphi}(t)\, A_o f_r + \dot{\varphi}(t)\, \alpha\, f_r: \quad \epsilon_r(0) = 0$$

$$\epsilon_r'(t) = C_o\, \epsilon_r(t) \qquad (4.1.39)$$

In contrast to (4.1.36), (4.1.39) has two event vectors. It is always true that the choice of d_2 leaves sufficient freedom in the choice of D to insure that the output error from (4.1.39) is planar, but in this case it is not possible to choose D such that the output error is unidirectional. Hence, the output error generated by a rate sensor failure can only be constrained to a two-dimensional plane.

The notation adopted in this chapter is that sensor failures are associated with a pair of event vectors. Thus the rate sensor failure is associated with the pair $\{f_r, A_o f_r\}$ and the attitude sensor failure is associated with the pair $\{f_a, 0\}$.* These event

*For convenience, α_a and α are chosen to be zero; however, any value is acceptable. The choice of α_a and α affects the direction of the event vectors, but not the output directions associated with the failure.

vectors are independent of the manner in which the sensor fails. If a failure mode other than the hard failure modeled by (4.1.28) and (4.1.29) were assumed, the only difference in (4.1.30) - (4.1.39) would be in the time functions which multiply the different vectors.

The preceeding discussion on sensor failures is pertinent only if an open loop reference model is the basis for the failure detection system. If a sensor output is used in the control loop, then the sensor's function in the closed loop is that of a dynamic element. If the failure detection system is then based upon the closed loop reference model, a single event vector is associated with the sensor failure and D can be chosen such that the sensor failure generates a unidirectional output error. This feature can be demonstrated using the spinning cylinder example.

The closed loop reference model is the model analyzed in Section 3.3.2, which includes the control law (4.1.26). The state vector is:

$$x_c(t) = \begin{bmatrix} \varphi_m(t) \\ \dot{\varphi}(t) \\ \dot{\varphi}_m(t) \\ T(t) \end{bmatrix}$$

and the system model is:

$$C_c = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad A_c = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & -\beta_r & 0 & \alpha_r \\ 0 & -\beta_r & 0 & \alpha_r \\ -\alpha_1 & 0 & -\alpha_2 & -\beta_f \end{bmatrix} \quad B_c = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \alpha_1 \end{bmatrix} \tag{4.1.40}$$

The control law is algebraic, but it is still necessary to increase the dimension of the system model to four. Since the measured attitude $\varphi_m(t)$ and rate $\dot{\varphi}_m(t)$ are used in the control law, these measurements are included in the state vector. (The first two rows of C_c no longer correspond to failures in the sensors.) It is also necessary to include the true rotation rate $\dot{\varphi}(t)$ in the state vector in order to properly model $\varphi_m(t)$ after a rate sensor failure. The true attitude $\varphi(t)$ can be omitted since it has no dynamic effect on the other state variables.

A failure detection system for (A_c, B_c, C_c) can be designed for the failures i) - iv) mentioned earlier plus:

v) a failure in the controller

As discussed in Section 3.3.2, the attitude and rate sensor failures are not mutually detectable for this model since there is an unassignable eigenvalue of $A_c - D_c C_c$ associated with $\varphi(t)$. Means for compensating for this fact by a further increase in the dimension of the model are discussed in Section 4.2. For present purposes, it is convenient to assume that the unassignable eigenvalue is acceptable and to ignore it.

Although the closed loop model can detect failures in both the torquing device and the control law, these failures are detection equivalent, i.e. it is not possible to discern between them on the basis of output error direction alone (Definition 3.1.1). The same difficulty is encountered in attempting to determine whether a controller failure is due to a change in the attitude feedback gain α_1 or the rate feedback gain α_2. A general control law failure can be modeled as a change in the last row of A:

$$\Delta A = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \Delta\alpha_1 & 0 & \Delta\alpha_2 & 0 \end{bmatrix} \qquad (4.1.41)$$

By comparison with the state-dependent dynamics change model in (4.1.9), the failure can be modeled by a single event vector e_{44} and the error equations become:

$$\dot{\epsilon}_c(t) = (A_c - DC_c)\,\epsilon_c(t) + e_{44}(\Delta\alpha_1 \varphi_m(t) + \Delta\alpha_2\,\dot{\varphi}_m(t)):$$

$$\epsilon_c(0) = 0$$

$$\epsilon'_c(t) = C_c\,\epsilon_c(t) \qquad (4.1.42)$$

The output direction associated with the failure is:

$$e_{33} = C_c e_{44} \qquad (4.1.43)$$

Since $\Delta\alpha_1$ and $\Delta\alpha_2$ have the same event vector in (4.1.42), output error direction alone is not sufficient to determine whether $\Delta\alpha_1$, $\Delta\alpha_2$, or both cause the failure. If a torquing device failure is modeled by $\Delta\beta_f$ and $\Delta\alpha_c$, the event vector is also e_{44} and the output error for this failure is also along e_{33}. The correlation techniques of Section 5.4 suggest one procedure for discriminating among $\Delta\alpha_1$, $\Delta\alpha_2$, and torquing device failures.

An attitude sensor failure can be modeled in the closed loop case as a change in any or all of the elements of the first row of A_c. Similarly the third row corresponds to a rate sensor failure. Suitable state-dependent failure models can be based on:

$$\Delta A_a = \begin{bmatrix} \Delta a_{11} & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad \Delta A_r = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & +\beta_r & \Delta a_{33} & -\alpha_r \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

(4.1.44)

The event vectors are e_{41} and e_{43} respectively and the error equations are:

$$\dot{\epsilon}_a(t) = (A_c - DC_c)\epsilon_a(t) + e_{41}(\Delta a_{11}\varphi_m(t) - \dot{\varphi}(t)): \epsilon_a(0) = 0$$

$$\epsilon_a'(t) = C_c \epsilon_a(t) \qquad (4.1.45)$$

and:

$$\dot{\epsilon}_r(t) = (A_c - DC_c)\epsilon_r(t) + e_{43}(\Delta a_{33}\dot{\varphi}_m(t) + \beta_r\dot{\varphi}(t) - \alpha_r T(t)): \epsilon_r(0) = 0$$

$$\epsilon_r'(t) = C_c \epsilon_r(t) \qquad (4.1.46)$$

In both cases ΔA has been chosen such that the driving terms for $\varphi_m(t)$ and $\dot{\varphi}_m(t)$ are canceled and, if Δa_{11} and Δa_{33} are negative, that $\varphi_m(t)$ and $\dot{\varphi}_m(t)$ after the failure decay to zero. Each failure model required only a single event vector and it follows that a detection filter can be chosen such that the sensor failures produce unidirectional output--along e_{31} and e_{32} respectively. The same event vectors and output directions would have been obtained if (4.1.14) had been used to model the sensor failures as biases.

The torque sensor is not used in the control law and an output error signal due to its failure can only be constrained to a plane. This distinction between measurements incorporated in the control law and measurements not incorporated in the control law is a general property of closed loop reference models. The implication of this to the design of a failure detection system is discussed further in Section 4.1.3.

4.1.3 General Modeling Considerations

The choice of an acceptable reference model (A, B, C) for the failure detection system involves an element of judgement on the part of the designer. It was pointed out in Section 3.2.2

that no advantage is to be gained from accepting an unobservable model. A detection filter can be designed for an unobservable model and the resulting filter will perform precisely as if it had been designed for a lower dimensional, observable reference model. The cost of an unobservable model is an increase in both the computation required in the design of the detection filter and in the computation necessary to use the model in real-time. Even if the state estimate produced by the failure detection system is to be used in the control law for the system, no useful information can be gathered from the unobservable dynamics of the model. If the reference model becomes unobservable because of a sensor failure in the system, the model may still be used. Otherwise, the reference model should be observable.

The designer has considerable freedom in choosing the manner in which a failure is to be modeled. Every mathematical model is an approximation of reality. The choice of the reference model dictates to a certain extent the manner in which a failure can be modeled. If the model is too precise, more event vectors may be required to model the failure of a component than if a less precise model were used. It is generally true that a reference model in which only the principal effects of the failure are noticed provides adequate failure detection without requiring excessive computation.

The most basic modeling decision the designer has to make--the choice between modeling the open loop or closed loop dynamics--has a major effect upon the structure of the failure detection problem. Implementations in both the open loop and closed loop cases are illustrated in Figure 4.1. One advantage of using an open loop model is that the state estimate $x_m(t)$ produced by the failure detection system can be used by the controller in the feedback control law. This possibility is indicated in Figure 4.1a. Application of a failure detection system in this context is discussed in Chapter 5.

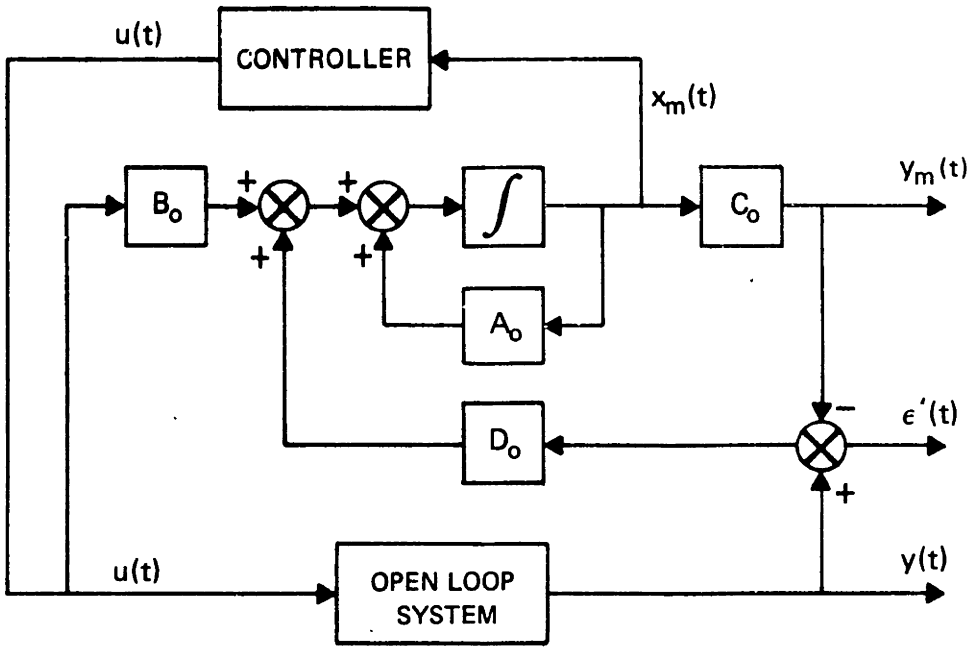


Figure 4. 1a Implementation of a failure detection system to monitor failures in the open loop system $x_m(t)$ can be used in the control law.

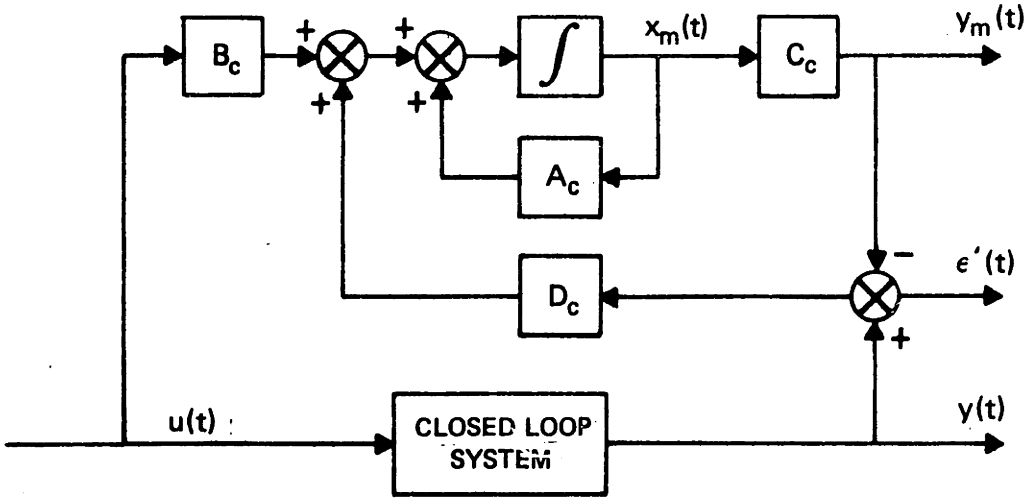


Figure 4. 1b Implementation of a failure detection system to monitor failures in the closed loop system. Failures in both the plant and the controller can be monitored.

Aside from the state estimator possibilities for an open loop reference model, there are a number of failure detection advantages and disadvantages to be considered in choosing the reference model. In general, three advantages can be gained if the closed loop reference model is substituted for the open loop model:

i) failures in the control law can be detected,

ii) for sensors whose output is used in the control law, the output error signals generated by a failure can be made unidirectional, and

iii) for a set of sensors which are not linearly independent but whose output is used in the control law, a failure can be associated unambiguously with the appropriate sensor.

The last two items are in marked contrast to the open loop case, although with an increase in the order of the open loop model failures in linearly dependent sensors can be detected unambigously for that case also.

Along with these advantages of a closed loop reference model, there are also two disadvantages:

i) the closed loop model may be of higher dimension than the open loop model, and

ii) failures which are mutually detectable for an open loop model may not be mutually detectable for a closed loop model.

The increase in dimension may occur even if the control law in the controller is algebraic rather than dynamic. The loss of mutual detectability is caused by the possible introduction of unassignable eigenvalues for $A_c - D_c C_c$. All of these factors were illustrated in Section 4.1.2.

4.2 Filter Design Procedure

The general design problem for a detection filter is outlined in this section. This section and Section 4.4 on base normal form present a reasonable procedure for designing a failure detection system based upon Chapter 3. Many of the design algorithms presented in this section are trivial if the system is expressed in base normal canonical form. In particular, a detection filter for a set of $r < m$ event vectors can be chosen by inspection. The design algorithms are illustrated for a satellite orbit perturbation problem in Section 4.3. The base normal canonical form is used exclusively in Chapters 5 and 6.

In the algorithms to follow, it is assumed that the reference model (A, B, C) is observable. As discussed in Sections 3.2.2 and 4.1, there is no advantage in having an unobservable model. If an unobservable model is used, Sections 3.2.2 and 3.2.3 should be used to reinterpret the following algorithms.

Many of the statements of this section concern the event vectors $f_1 \ldots f_r$, $h_1 \ldots h_q$. It is implicitly assumed that if some Cf_i is zero, f_i is replaced by the first element of Af_i, $A^2 f_i \ldots$ for which the corresponding element of CAf_i, $CA^2 f_i \ldots$ is non-zero. A similar procedure is assumed for h_i. Whenever event vectors are used, it is understood that this substitution is made where necessary.

4.2.1 Detection Spaces

The first step in the design procedure is to determine acceptable event vectors for the failures of interest and then find their detection spaces. The remarks of Section 4.1.1 concerning the replacement of columns of D by equivalent sets of event vectors are applicable. Before the design procedure is completed, pairs of event vectors $\{f, Af - \xi_f\}$, where ξ_f is in the detection space of f, should be associated with all sensor failures, or means for uniquely determining the

appropriate columns of D before D is completely specified should be developed. In general it is desirable to determine the detection spaces for failures with well-defined event vectors before considering an improvement in the sensor failure model. This permits a judicious choice of ξ_f in the set $\{f, Af + \xi_f\}$.

Assume $f_1 \ldots f_s$ are a set of event vectors for a set of s failures of interest. Assume also that a failure in the i^{th} sensor is to be monitored. Placing the sensor failure in the form of (4.1.22) requires an f which satisfies (4.1.19). In order of preference, the procedures for accomplishing this are:

i) If any of $f_1 \ldots f_s$ satisfy:

$$e_{mi} = Cf_k \qquad k \in \{1, 2 \ldots s\}$$

let:

$$f = f_k$$

and define the event vector set $\{f, Af\}$. This set is the set of event vectors necessary for (4.1.22).

ii) If $f_1 \ldots f_\ell$ are a subset of $f_1 \ldots f_s$ and a linear combination of f_1 and f_ℓ satisfies:

$$e_{mi} = \sum_{k=1,\ell} \alpha_k C f_k$$

then define the set $\{f_1 \ldots f_\ell, d_i\}$. This option is discussed further below.

iii) Find any f which satisfies:

$$e_{mi} = Cf$$

and define the set $\{f, Af\}$.

Option ii) above differs from the algorithm suggested in Section 4.1.1 in that d_i need not be eliminated. If a detection filter is chosen which detects $f_1 \ldots f_\ell$ and the eigenvalues of A-DC associated with $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_\ell$ are assigned using Section 4.2.4, d_i is uniquely defined. If sufficient freedom remains in the choice of D to detect d_i, then the sensor failure can be detected along with the failures for $f_1 \ldots f_\ell$ and it produces output in the output space defined by e_{mi} and Cd_i.

In i) and iii) above Af can be replaced by $Af - \xi_f$ where ξ_f is an arbitrary element of $\mathcal{R}_f$. If Af is contained in $\bar{\mathcal{R}}_f$, the sensor error can be made unidirectional and only f need be considered. The three steps are presented in order of the efficiency with which they utilize the design freedom in the choice of D. Since iii) requires two new event vectors to be defined, it is obviously the least efficient. Option ii) is the most difficult to implement, iii) is the easiest. Option i) should always be examined first.

Starting from an initial set of event vectors $f_1 \ldots f_s$, application of i) - iii) produces a larger set $f_1 \ldots f_p$. As indicated, some of the event vectors in the set must be considered in pairs for the design algorithm. If a sensor failure is characterized by the pair of event vectors {f, Af}, it is of no value to design a detection filter for only one event vector of the pair. In the following subsections, this paired dependency is not explicitly denoted, but it must be recognized in the design process.

Conceptually, detection spaces are simple to determine. If an event vector f_i satisfies $Cf_i \neq 0$ and if D_f and C' are chosen as suggested in Lemma 3.2.6′:

$$D_f = Af_i \left[(Cf_i)^T Cf_i \right]^{-1} (Cf_i)^T$$

$$C' = \left[E_m - Cf_i \left[(Cf_i)^T C f_i \right]^{-1} (Cf_i)^T \right] C \qquad (4.2.1)$$

Then $\bar{\mathcal{R}}_i$ is the ν_i-dimensional unobservable space of the system $(A-D_fC, C')$. The principal difference between C' and C is that the output direction Cf_i has been excised from the output space

of C'; a similar relationship holds for D_f and the image of f_i under $A\text{-}D_f C$:

$$C'f_i = 0$$

$$(A\text{-}D_f C)\, f_i = 0 \qquad (4.2.2)$$

If Cf_i is zero, it should be replaced by the first element of Af_i, A^2f_i.... for which the corresponding CAf_i, CA^2f_i... is non-zero.

A computational algorithm for $\bar{\mathcal{R}}_i$ is given in Appendix A.2. The algorithm determines a $n \times \nu_i$ matrix $\bar{R}_i$ whose range space is $\bar{\mathcal{R}}_i$ where ν_i is the dimension of $\bar{\mathcal{R}}_i$. The matrix is partitioned:

$$\bar{R}_i = [f_i : R_i] \qquad (4.2.3)$$

and the range space of R_i is $\mathcal{R}_i$ as defined in Definition 3.1.2, i.e., R_i is in the null space of C:

$$C R_i = 0 \qquad (4.2.4)$$

The algorithm in Appendix A.2 is less direct than it could be because it determines several other quantities needed for designing a detection filter.

In assigning the eigenvalues of A-DC, or in defining the transformation to base normal canonical form, the detection generator g_i for $\bar{\mathcal{R}}_i$ is useful. The detection generator is defined in Lemma 3.2.4. It possesses two important properties:

i) If $\bar{\mathcal{R}}_i$ is ν_i-dimensional, g_i, $Ag_i \ldots A^{\nu_i - 1} g_i$ form a basis for $\bar{\mathcal{R}}_i$, and

ii) The vectors g_i, $Ag_i \ldots A^{\nu_i - 2} g_i$ are in the null space of C.

Since f_i is in $\bar{\mathcal{R}}_i$, i) implies that constants $p_{i1} \ldots p_{i\nu_i}$ exist such that:

$$f_i = p_{i\nu_i} A^{\nu_i - 1} g_i + \ldots p_{i2} Ag_i + p_{i1} g_i \qquad (4.2.5)$$

Appendix A.3 provides an algorithm for g_i similar to that used for $\bar{R}_i$.

4.2.2 Multiple Event Detection

Once a set of event vectors $f_1 \ldots f_p$ has been chosen, the next step in the design of a failure detection system is to determine the subsets of $f_1 \ldots f_p$ which can be detected with a single detection filter. The Detection Theorem (Theorem 3.2.1) guarantees that a detection filter for a single event vector can always be found; the matrix D_f is an acceptable choice of D. Unfortunately the analysis involved in finding subsets of $f_1 \ldots f_p$ for which a detection filter exists is more difficult. Section 3.3 provides a complete set of algorithms for this purpose.

Consider a subset $f_1 \ldots f_r$ of $f_1 \ldots f_p$ and define the matrix:

$$F = [f_1 \ldots f_r] \tag{4.2.6}$$

The subspace $\mathcal{J}$ is spanned by $f_1 \ldots f_r$. The vectors $f_1 \ldots f_r$ are defined to be output separable (Definition 3.3.2) if:

$$\text{rk}F = \text{rk}CF \tag{4.2.7}$$

Lemma 3.3.1 proves output separability is sufficient to guarantee that D can be chosen such that the failures associated with $f_1 \ldots f_r$ generate unidirectional output errors -- an acceptable choice of D is $D_{\mathcal{J}}$ as given below. The difficulty is that the definition of mutual detectability (Definition 3.3.1) requires that a D exist such that:

i) the output errors generated by the failures associated with $f_1 \ldots f_r$ are unidirectional, and

ii) the eigenvalues of A-DC can be specified almost arbitrarily by the proper choice of D.

Output separability is not sufficient to guarantee that both conditions i) and ii) above can be satisfied.

The determination of a sufficient condition for mutual detectability requires a generalization of the concept of detection spaces. If $\mathcal{J}$ is the range space of F, Definition 3.3.3 defines a detection space $\bar{R}_{\mathcal{J}}$ for $\mathcal{J}$. $\bar{R}_{\mathcal{J}}$ is similar to the ordinary detection spaces discussed in Section 4.2.1 except that the single vector f is replaced by the space $\mathcal{J}$. In essence $\bar{R}_{\mathcal{J}}$ defines the group detection properties of $f_1 \ldots f_r$ -- this interpretation is motivated in the example in Section 3.3.2. It can be found in the same manner as $\bar{R}_1 \ldots \bar{R}_r$. If $D_{\mathcal{J}}$ and C' are defined:

$$D_{\mathcal{J}} = AF[(CF)^T CF]^{-1}(CF)^T$$

$$C'_{\mathcal{J}} = [E_m - CF[(CF)^T CF]^{-1}(CF)^T]C \qquad (4.2.8)$$

then Lemma 3.3.2 states that $\bar{R}_{\mathcal{J}}$ is the unobservable space of the system $(A-D_{\mathcal{J}}C, C')$. The algorithm in Appendix A.2 determines $\bar{R}_{\mathcal{J}}$ as the range space of a matrix $\bar{R}_F$. The matrices R_F and $\bar{R}_F$ are defined similarly to R_i and $\bar{R}_i$:

$$\bar{R}_F = [F : R_F] \qquad (4.2.9)$$

The algorithm suggests that $\bar{R}_{\mathcal{J}}$ be computed first and then $\bar{R}_1 \ldots \bar{R}_r$ be determined from $\bar{R}_{\mathcal{J}}$ instead of $\eta(C) \oplus F$.

The important feature of $\bar{R}_{\mathcal{J}}$ is that it contains each of $\bar{R}_1 \ldots \bar{R}_r$ as well as their direct sum (Lemma 3.3.4):

$$\bar{R}_1 \oplus \bar{R}_2 \oplus \ldots \bar{R}_r \subset \bar{R}_{\mathcal{J}} \qquad (4.2.10)$$

In matrix terms, this implies that:

$$\begin{aligned} \text{rk } \bar{R}_F &\geq \text{rk } [\bar{R}_1 : \ldots : \bar{R}_r] \\ &= \text{rk } [F : R_1 : \ldots : R_r] \end{aligned} \qquad (4.2.11)$$

and that the range space of $\bar{R}_F$ contains the range spaces of all of $\bar{R}_1 \ldots \bar{R}_r$. The difficulties involved with detecting $f_1 \ldots f_r$ with a single detection filter concern the inequalities in (4.2.10) and (4.2.11). Theorem 3.3.1 states that a necessary and sufficient condition for $f_1 \ldots f_r$ to be mutually detectable is for (4.2.10) (and therefore (4.2.11)) to be an equality.

If (4.2.10) is an equality, the remainder of this subsection is not needed. Assuming (A,C) is observable, a detection filter can be chosen for $f_1 \ldots f_r$ and all n eigenvalues of A-DC can be assigned using the techniques of Section 4.2.4 or Section 4.4. If (4.2.10) is not an equality, Lemma 3.3.1 guarantees that a D can still be chosen for which each failure associated with one of $f_1 \ldots f_r$ produces a unidirectional output error, but some of the eigenvalues of A-DC cannot be assigned. If ν_0 is defined:

$$\nu_0 = \text{rk } \bar{R}_F - \sum_{i=1,r} \text{rk } \bar{R}_i$$

$$= \nu_{\mathcal{J}} - \sum_{i=1,r} \nu_i \qquad (4.2.12)$$

a total of ν_0 of the eigenvalues of A-DC have a fixed value which cannot be altered (Lemmas 3.3.5 and 3.3.8).

Lemma 3.3.6 constructs a ν_0-dimensional subspace $\mathcal{R}_{0g}$ -- called the excess subspace of $\bar{\mathcal{R}}_{\mathcal{J}}$ -- such that:

$$\bar{\mathcal{R}}_{\mathcal{J}} = \bar{\mathcal{R}}_1 \oplus \ldots \bar{\mathcal{R}}_r \oplus \mathcal{R}_{0g} \qquad (4.2.13)$$

If D is chosen such that each of $f_1 \ldots f_r$ generate a unidirectional output error, the unassignable eigenvalues of A-DC are the eigenvalues associated with $\mathcal{R}_{0g}$. Appendix A.3 provides an algorithm for determining these eigenvalues. A matrix R_{0g} is constructed whose range space is $\mathcal{R}_{0g}$. With G defined in terms of the detection generators for $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$:

$$G = [g_1 \dots g_r] \tag{4.2.14}$$

a $\nu_0 \times \nu_0$ matrix Π and a $r \times \nu_0$ matrix Θ are constructed in Appendix A.3 such that:

$$A R_{0g} = R_{0g} \Pi + G \Theta \tag{4.2.15}$$

The unassignable eigenvalues of A-DC are the eigenvalues of Π (Lemma 3.3.8).

Once the excess subspace of $\bar{\mathcal{R}}_{\mathcal{J}}$ and the unassignable eigenvalues of A-DC have been determined, three options are available to the designer:

i) choose a D for which each of $f_1 \dots f_r$ generate unidirectional output errors and accept the ν_0 unassignable eigenvalues of A-DC,

ii) find a subset of $f_1 \dots f_r$ for which their are no unassignable eigenvalues of A-DC, or

iii) increase the dimension of the reference model to $\tilde{n}$ in such a manner that all the eigenvalues of $\tilde{A} - \tilde{D}\tilde{C}$ are assignable.

The choice between these options obviously depends upon the acceptability of the eigenvalues of Π as eigenvalues of A-DC. If either of the last two options is chosen, Sections 3.3.3 and 3.3.4 respectively provide a complete analytical basis for the two algorithms discussed in the remainder of this subsection.

Consider the problem of removing event vectors from $f_1 \dots f_r$ first. The Detection Theorem (Theorem 3.2.1) guarantees that any one vector of $f_1 \dots f_r$ is detectable. Therefore, as vectors are removed from the set $f_1 \dots f_r$, the number of unassignable eigenvalues of A-DC should decrease. It is possible to remove one or more vectors from $f_1 \dots f_r$, find a new $\bar{\mathcal{R}}_{\mathcal{J}}$ for the remaining vectors, check for mutual detectability, etc. The problem is that if r is large, the computation required to ultimately obtain a mutually detectable subset

of $f_1 \ldots f_r$ may be prohibitive. Theorem 3.3.3 and its corollaries present a more efficient algorithm.

If the first ℓ vectors of $f_1 \ldots f_r$ are removed, Theorem 3.3.3 defines a matrix M_{0L}:

$$M_{0L} = \begin{bmatrix} \Theta_L^c \\ \Theta_L^c \Pi \\ \vdots \\ \Theta_L^c \Pi^{\nu_0 - 1} \end{bmatrix} \qquad (4.2.16)$$

where Θ_L^c is the first L rows of Θ in (4.2.15). Corollary 1 of Theorem 3.3.3 then defines a $\nu_0 \times \nu_\theta$ transformation matrix $[\Delta : \beta]_L$ where the ν_{0L} columns of β span the null space of M_{0L} and the $\nu_0 - \nu_{0L}$ columns of Δ span the row space of M_{0L} -- Appendix A.1 provides an algorithm which yields $[\Delta : \beta]_L$. Both Π and R_{0g} for the original r event vectors can then be transformed:

$$\Pi_L^0 = [\Delta : \beta]_L^{-1} \Pi [\Delta : \beta]_L = \left[\begin{array}{c|c} \Pi_L^c & 0 \\ \hline \Pi_{L0}^c & \Pi_L \end{array}\right] \qquad (4.2.17)$$

$$R_{0g}^0 = R_{0g}[\Delta : \beta]_L = [R_{0gL}^c : R_{0gL}]$$

where Π_L is $\nu_{0L} \times \nu_{0L}$, R_{0gL} is $n \times \nu_{0L}$ etc. Corollary 1 of Theorem 3.3.3 states that the excess subspace for $f_{\ell+1} \ldots f_r$ is the range space of R_{0gL}. If D is chosen such that each of $f_{\ell+1} \ldots f_r$ generates a unidirectional output error, the unassignable eigenvalues of A-DC are the ν_{0L} eigenvalues of Π_L. Since the eigenvalues of a matrix are not changed by a similarity transformation, the ν_{0L} eigenvalues of Π_L must be a subset of the ν_0 eigenvalues of Π. Clearly the theorem is applicable to the removal of ℓ arbitrarily chosen event vectors from $f_1 \ldots f_r$ if Θ_L^c is chosen to be the appropriate rows of Θ.

Although (4.2.16) and (4.2.17) are computationally more efficient than determining $\bar{R}_{\mathcal{J}}$ for the event vectors $f_{\ell+1}\ldots f_r$ and then finding R_{0gL} and Π_L directly, Corollary 2 of Theorem 3.3.3 permits a much greater savings. The procedure is to apply Theorem 3.3.3 and Corollary 1 by first eliminating only f_1 from $f_1\ldots f_r$, and then only f_2 from $f_1\ldots f_r$ and so on for r iterations. This yields a set of r equations similar to (4.2.15):

$$A R_{0gk} = R_{0gk} \Pi_k + G_k \Theta_k \qquad k = 1\ldots r \tag{4.2.18}$$

where G_k is G with the k^{th} column deleted, Θ_k is Θ with the k^{th} row deleted, the range space of R_{0gk} is the excess subspace after f_k is removed from $f_1\ldots f_r$, and the eigenvalues of Π_k are the corresponding unassignable eigenvalues of A-DC. If any R_{0gk} is zero, the r-1 event vectors $f_1\ldots f_{k-1}$, $f_{k+1}\ldots f_r$ are mutually detectable.

With the r solutions from (4.2.18), let $\Lambda_1\ldots\Lambda_r$ be the sets of eigenvalues for the corresponding $\Pi_1\ldots\Pi_r$. If any arbitrary subset of $f_1\ldots f_r$ is removed from $f_1\ldots f_r$, Corollary 2 states that the unassignable eigenvalues of A-DC associated with the remaining event vectors can be determined from the intersections of the appropriate set of $\Lambda_1\ldots\Lambda_r$. As an example, if both f_1 and f_2 are removed from $f_1\ldots f_r$, the unassignable eigenvalues of A-DC associated with $f_3\ldots f_r$ are the intersection of Λ_1 and Λ_2:

$$\Lambda_I = \Lambda_1 \cap \Lambda_2 \tag{4.2.19}$$

If Λ_I is the empty set, then $f_3\ldots f_r$ are mutually detectable; if not, other combinations can be tried. Corollary 2 of Theorem 3.3.3 generalizes this simple test to all possible subsets of $f_1\ldots f_r$.

The intersections prescribed by Corollary 2 of Theorem 3.3.3 are trivial to evaluate once $\Pi_1\ldots\Pi_r$ and then $\Lambda_1\ldots\Lambda_r$ have been determined from (4.2.18) using Corollary 1 of Theorem 3.3.3 and Appendix A.3. It then becomes a simple matter to determine all possible mutually detectable subsets of $f_1\ldots f_r$. As an alternative,

it is possible to determine larger subsets of $f_1 \ldots f_r$ which are not mutually detectable, but for which the corresponding unassignable eigenvalues of A-DC are acceptable. This amounts to a compromise between options i) and ii) mentioned earlier. In some applications, either of these alternatives yields a subset of $f_1 \ldots f_r$ which is smaller than desired, or which contains only one of a pair of event vectors associated with a sensor failure. In such cases option iii) can be utilized to yield a larger set of mutually detectable event vectors than option ii).

An unfortunate feature of failure detection theory is that the mutual detectability of certain failures is a function of the choice of the reference model upon which the failure detection system is to be based. This model dependence is shared by other aspects of systems theory -- most notably, by output decoupling theory (Wonham and Morse (1970b) and Howze and Pearson (1970)). Just as dynamic compensation was developed to eliminate model-dependency in the output decoupling problem, the state space augmentation techniques developed in Section 3.3.4 eliminate model-dependency in the mutual detectability problem. The power of state space augmentation is summarized in Theorem 3.5.1: The vectors $f_1 \ldots f_r$ can be made mutually detectable with state space augmentation if and only if they are output separable. Since the rank of C is m, this guarantees that m appropriately chosen failures can be detected with a single failure detection system. The cost is that the reference model upon which the failure detection system is based may have a larger dimension than the original reference model.

Consider (4.2.18) again. If:

$$\mathrm{rk}\Pi_k = \mathrm{rk}\Pi \qquad (4.2.20)$$

then removing f_k from $f_1 \ldots f_r$ does not alter the mutual detectability of $f_1 \ldots f_r$ -- there are still ν_0 unassignable eigenvalues of A-DC associated with the remaining r-1 event vectors. Corollary 2 of Theorem 3.3.3 implies that (4.2.20) is sufficient to guarantee

that f_k can be in every mutually detectable subset of $f_1 \ldots f_r$. Lemma 3.3.10 constructs a reference model from (A, B, C) such that (4.2.20) is satisfied for f_k and the new reference model:

$$\tilde{C} = [C \vdots 0] \qquad \tilde{A} = \left[\begin{array}{c|c} A & -g_k \theta_k \Delta \\ \hline 0 & \Pi_k^c \end{array}\right] \quad \tilde{B} = \left[\begin{array}{c} B \\ \hline 0 \end{array}\right] \tag{4.2.21}$$

where g_k is the k^{th} column of G, θ_k is the k^{th} row of Θ, and Δ and Π_k^c are the appropriate terms in (4.2.17) if Theorem 3.3.3 is evaluated for f_k and the original reference model (A, B, C). The new reference model ($\tilde{A}$, $\tilde{B}$, $\tilde{C}$) is input-output equivalent to the original reference model. If Theorem 3.3.3 were then evaluated for f_k and the new reference model, $\tilde{\Pi}_k^c$ would be zero and

$$rk\,\tilde{\Pi}_k = rk\,\tilde{\Pi} \tag{4.2.22}$$

as desired.

Lemma 3.3.10 can be applied sequentially for the event vectors in the set $f_1 \ldots f_r$ until a reference model ($\tilde{\tilde{A}}$, $\tilde{\tilde{B}}$, $\tilde{\tilde{C}}$) is finally obtained for which $f_1 \ldots f_r$ are mutually detectable. Regardless of the order in which the state space augmentations are undertaken, Theorem 3.3.4 states that the final reference model will have a dimension of:

$$\tilde{\tilde{n}} = n + \sum_{k=1,r} (\nu_0 - rk\,\Pi_k) - \nu_0 \tag{4.2.23}$$

$$= n + \sum_{k=1,r} rk\,\Pi_k^c - \nu_0$$

where ν_0 is the rank of Π. After each application of Lemma 3.3.10, it is necessary to determine the excess subspace $\tilde{\mathcal{R}}_{0g}$ and the detection spaces for $f_1 \ldots f_r$ for the new reference model. Lemmas 3.3.10 and 3.3.11, and Appendix A.3, provide algorithms which are computationally more efficient than returning to Appendix A.2.

In certain applications, state space augmentation may result in a large $\tilde{n}$ for the final reference model. This shortcoming is partially compensated for by the fact that the reference model can be decoupled into a set of r smaller reference models, the sum of whose dimensions is $\tilde{n}$ (Section 4.4). If this still yields an unacceptable computational burden, some trade-off between options i) - iii) must be accepted.

4.2.3 Output Stationarity

The analysis of the preceding subsection begins with an output separable subset $f_1 \ldots f_r$ of event vectors for which a detection filter may be chosen. The algorithm in Section 4.2.2 may have increased the dimension of the reference model and it may have subsequently removed vectors from $f_1 \ldots f_r$. For notational convenience, however, the final reference model and set of event vectors are still denoted by (A, B, C) and $f_1 \ldots f_r$ respectively. It is not necessary that $f_1 \ldots f_r$ be mutually detectable. The vectors of $f_{r+1} \ldots f_p$ which are not output separable with $f_1 \ldots f_r$ are denoted by $h_1 \ldots h_q$ in this subsection and their detection spaces are denoted by $\bar{\mathcal{S}}_1 \ldots \bar{\mathcal{S}}_q$.

Even if unassignable eigenvalues of A-DC are acceptable, the results on mutual detectability provide for a maximum of rk C failures to be detected by a single failure detection system. Output stationarity (Definition 3.4.1) seeks to determine whether or not particular choices of D can be found which increase this number. In the context of the previous subsection, it was determined that $f_1 \ldots f_r$ is the largest subset of $f_1 \ldots f_p$ for which a failure detection system based on mutual detectability can be found. Output stationarity considers the problem of trying to find a D such that each of $f_1 \ldots f_r, h_i$ generates a unidirectional output error, where h_i is one of $h_1 \ldots h_q$ defined above. The cost of using output stationarity to increase the number of failures which can be detected by a single failure detection system is that certain eigenvalues of A-DC may have to be assigned a multiplicity greater than one. The advantage of this approach, however, is that it may significantly increase the number of failures which can be detected by a single failure detection system.

By definition, $f_1 \ldots f_r$, h_i are not output separable. Since $f_1 \ldots f_r$ are output separable, this implies that Ch_i is a linear combination $Cf_1 \ldots Cf_r$. Assume that $f_1 \ldots f_\ell$ are the smallest subset of $f_1 \ldots f_r$ such that:

$$C F_L^c \alpha_L^c = \beta_i Ch_i \tag{4.2.24}$$

has a solution where:

$$\alpha_L^c = [\alpha_1 \ldots \alpha_\ell]^T$$

$$F_L^c = [f_1 \ldots f_\ell] \tag{4.2.25}$$

and $\alpha_1 \ldots \alpha_\ell$ are a set of non-zero coefficients. If $f_1 \ldots f_\ell$ are mutually detectable, Corollary 1 of Theorem 3.4.1 guarantees that h_i can be made output stationary with $f_1 \ldots f_r$ if there exists a solution to:

$$\bar{R}_L^c \Xi = \bar{S}_i \tag{4.2.26}$$

where:

$$\bar{R}_L^c = [\bar{R}_1 : \ldots . : \bar{R}_\ell] \tag{4.2.27}$$

and where $\bar{S}_i$ corresponds to $\bar{\mathcal{S}}_i$ in the same manner that $\bar{R}_i$ corresponds to $\bar{\mathcal{R}}_i$:

$$\bar{S}_i = [h_i : S_i] \tag{4.2.28}$$

If any solution to (4.2.26) exists, one solution is:

$$\Xi = (\bar{R}_L^{c\,T} \bar{R}_L^c)^{-1} \bar{R}_L^{c\,T} \bar{S}_i \tag{4.2.29}$$

Corollary 2 of Theorem 3.4.1 presents a similar test for the case where $f_1 \ldots f_\ell$ are not mutually detectable.

If h_i can be made output stationary with $f_1 \ldots f_r$, the next question concerns an algorithm for selecting D and whether or not the eigenvalues of A-DC can still be chosen almost arbitrarily. If:

$$\nu_{\mathscr{l}} = rk\, \bar{S}_i \qquad (4.2.30)$$

Theorem 3.4.2 asserts that a set of $\nu_{\mathscr{l}}$ eigenvalues $\Lambda_{\mathscr{l}}$ can be chosen arbitrarily and that a D can be found for which:

i) h_i is output stationary with $f_1 \ldots f_r$ for A-DC, and

ii) $\nu_{\mathscr{l}}$ of the eigenvalues of A-DC associated with each of $\bar{R}_1 \ldots \bar{R}_{\mathscr{l}}$ are the elements of $\Lambda_{\mathscr{l}}$.

The constraint on the choices of D is that $\mathscr{l}$ eigenvalues of A-DC must be equal to every element in $\Lambda_{\mathscr{l}}$. The eigenvalues of A-DC associated with the remaining detection spaces $\bar{R}_{\mathscr{l}+1} \ldots \bar{R}_r$ remain freely assignable. If $\nu_k > \nu_{\mathscr{l}}$ for some $k \leq \mathscr{l}$, however, the $\nu_k - \nu_{\mathscr{l}}$ additional eigenvalues associated with the corresponding $\bar{R}_k$ are unassignable. An algorithm for choosing D is presented in Section 4.2.4.

The major difficulty with output stationarity is the possible introduction of unassignable eigenvalues of A-DC. For a large number of applications such as the satellite example in Section 4.3 or the spinning cylinder example in Section 4.1.2, the $\nu_1 \ldots \nu_{\mathscr{l}}$ are equal and no unassignable eigenvalues are produced. If $\nu_1 \ldots \nu_{\mathscr{l}}$ are not equal, however, unassignable eigenvalues of A-DC are a necessary by-product of output stationarity.[1] An algorithm for determining the unassignable eigenvalues is presented in Section 4.2.4. If the eigenvalues are unacceptable, nothing can be done except remove the output stationarity constraint on h_i.

Equations (4.2.24) and (4.2.26) can be applied to each of $h_1 \ldots h_q$ to determine which can be made output stationary with $f_1 \ldots f_r$. If some h_j cannot, it should be eliminated from the set. It is then possible to consider the mutual output stationarity of the remaining

[1] It is possible to define circumstances under which output stationarity introduces unassignable eigenvalues of A-DC even if the ν_k, $k \in I$, are equal.

elements of the set. For notational convenience, it is assumed here that each of $h_1 \ldots h_q$ can be made output stationary with $f_1 \ldots f_r$.

Unlike the mutual detectability problem, there is no limit to the number event vectors which can be made mutually output stationary with $f_1 \ldots f_r$ -- but the fact that each of two event vectors can be made individually output stationary with $f_1 \ldots f_r$ does not necessarily imply that they can be made mutually output stationary with $f_1 \ldots f_r$. Section 3.4 provides a series of tests for determining mutual output stationarity. Assume h_i and h_j are two elements of $h_1 \ldots h_q$. The simplest test concerns the subsets of $Cf_1 \ldots Cf_r$ which are necessary to construct Ch_i and Ch_j similarly to (4.2.24). If I_i is the smallest index set such that Ch_i is a linear combination of Cf_k, $k \in I_i$, and I_j is a corresponding index set for h_j, then Corollary 3 of Theorem 3.4.1 states that h_i, h_j can be made mutually output stationary with $f_1 \ldots f_r$ if:

$$I_i \cap I_j = 0 \tag{4.2.31}$$

This is a sufficient, but not a necessary condition for mutual output stationarity.

The remaining tests are based upon the construction of normal subspaces. If Ξ is the solution of (4.2.29), the normal subspace $\bar{\mathcal{J}}_i^\circ$ of $\bar{\mathcal{J}}$ is defined in Definition 3.4.2′ to be the range space of the $n \times \nu_{\mathcal{J}}$ matrix:

$$\bar{S}_i^\circ = T_i \bar{R}_L^c \Xi \tag{4.2.32}$$

where T_i is a normalizing matrix defined in terms of the coefficients $\alpha_1 \ldots \alpha_\ell$ in (4.2.25):

$$T_i = \begin{bmatrix} \frac{1}{\alpha_1} E_{\nu_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\alpha_\ell} E_{\nu_\ell} \end{bmatrix} \qquad (4.2.33)$$

$E_{\nu_1} \ldots E_{\nu_\ell}$ are $\nu_1 \ldots \nu_\ell$-dimensional identity matrices where $\nu_1 \ldots \nu_\ell$ are the ranks of $\bar{R}_1 \ldots \bar{R}_\ell$ respectively. The normal event vector h_i° is the first column of $\bar{S}_i^\circ$:

$$\bar{S}_i^\circ = [h_i^\circ : S_i^\circ] \qquad (4.2.34)$$

Assume normal subspaces $\bar{\mathcal{S}}_i^\circ$ and $\bar{\mathcal{S}}_j^\circ$ have been determined for $\bar{\mathcal{S}}_i$ and $\bar{\mathcal{S}}_j$ respectively. If the index sets I_i and I_j are equal:

$$I_k = I_j \qquad (4.2.35)$$

a necessary and sufficient condition for h_i and h_j to be made mutually output stationary with $f_1 \ldots f_r$ is that $\bar{\mathcal{S}}_i^\circ$ and $\bar{\mathcal{S}}_j^\circ$ be identical (Theorem 3.4.3), i.e. the range spaces of $\bar{S}_i^\circ$ and $\bar{S}^\circ$ must be equal. A necessary and sufficient condition for this is that $\bar{S}_i^\circ$ and $\bar{S}_j^\circ$ have the same dimensions and that there be a solution K to:

$$\bar{S}_i^\circ K = \bar{S}_j^\circ \qquad (4.2.36)$$

If (4.2.35) is not satisfied, Lemma 3.4.3 and Theorem 3.4.4 define a general, but considerably more difficult, test for mutual output stationarity.

The difficulty with mutual output stationarity is that the algebra necessary to find a mutually output stationary subset of $h_1 \ldots h_q$ may be rewarded in the design algorithm for D by a set of unassignable and unacceptable eigenvalues of A-DC. There is no fail-safe rule for avoiding this problem, but several rules-of-thumb are helpful. In the order in which they should be applied:

i) discard any h_i for which the dimensions ν_k of the detection spaces $\bar{\mathcal{R}}_k$, $k \in I_i$, are not nearly equal,

ii) discard any h_i for which $\nu_{\mathcal{J}}$ is much less than the minimum ν_k, $k \in I_i$, and

iii) do not try to make h_i, h_j mutually output stationary with $f_1 \ldots f_r$ unless the dimensions of $\bar{\mathcal{J}}_i$ and $\bar{\mathcal{J}}_j$ are approximately equal or unless (4.2.31) applies.

These rules do not eliminate the possibility of unacceptable eigenvalues of A-DC occurring, but they do minimize the probability by minimizing the number of unassignable eigenvalues resulting from application of Theorem 3.4.2.

4.2.4 Eigenvalue Assignment

If the reference model is expressed in base normal form (Section 4.4), the eigenvalues of A-DC can be assigned trivially in a manner similar to that used in Section 4.5. Otherwise it is necessary to use the algorithms outlined in this subsection. The results of Section 3.2.2, Section 3.4.1, and Appendix A.4 are applied here to provide algorithms to:

i) assign the eigenvalues of A-DC associated with $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$,

ii) assign the remaining $n-\nu_{\mathcal{J}}$ eigenvalues of A-DC,

iii) choose D such that h_i is output stationary with $f_1 \ldots f_r$, and

iv) determine the unassignable eigenvalues of A-DC if h_i is output stationary with $f_1 \ldots f_r$.

As in the previous subsection, it is assumed that $f_1 \ldots f_r$ are output separable and $h_1 \ldots h_q$ are candidates for output stationarity.

The solution to the eigenvalue assignment problem depends on the set of event vectors $f_1 \ldots f_r$, but it depends even more fundamentally upon A and C. Assume for a moment that it is desired to choose D such that A-DC is equal to some chosen matrix A_Λ. If C is invertible, the unique choice of D is:

$$D = (A - A_\Lambda) C^{-1} \qquad (4.2.37)$$

If C is not invertible, however, a solution to:

$$DC = A - A_\Lambda \qquad (4.2.38)$$

may not exist -- Lemma 3.2.1 states necessary and sufficient conditions for a D which satisfies (4.2.38) to exist. If Λ is the set of n desired eigenvalues of A-DC, these eigenvalues do not uniquely specify a desired A_Λ. Given Λ and a set of conditions to be discussed later, the algorithms developed in Chapter 3 choose A_Λ such that a solution to (4.2.38) exists.

The basic result on eigenvalue assignment using output feedback is Theorem 2.4.1. This theorem states that for any set of eigenvalues Λ, a matrix A_Λ with the desired eigenvalues can be chosen such that (4.2.38) has a solution D if and only if (A, C) is an observable reference model. This is a well-known result in system theory and does not consider the constraints which $f_1 \ldots f_r$ may place upon the permissible choices of D. In the remainder of this subsection, it is assumed that (A, C) is observable. The generalization of detection theory to unobservable reference models is discussed in Section 3.2.2.

The detection spaces $\bar{R}_1 \ldots \bar{R}_r$ are $\nu_1 - \ldots \nu_r$-dimensional respectively. Because of the requirement that the failures associated with each of $f_1 \ldots f_r$ must generate unidirectional output errors, ν_1 of the eigenvalues of A-DC become uniquely associated with $\bar{R}_1$, ν_2 with $\bar{R}_2$, etc. (Theorem 3.3.1). Let $\Lambda_1 \ldots \Lambda_r$ be the sets of desired eigenvalues associated with $\bar{R}_1 \ldots \bar{R}_r$. The elements of $\Lambda_1 \ldots \Lambda_r$ are the roots of an associated set of polynomials $\Psi_1(.) \ldots \Psi_r(.)$:

$$\Psi_k(\lambda) = (\lambda - \lambda_{k1})(\lambda - \lambda_{k2}) \ldots (\lambda - \lambda_{k\nu_k}) \quad k = 1 \ldots r \qquad (4.2.39)$$
$$= \lambda^{\nu_k} + p_{k\mu_k} \lambda^{\mu_k - 1} + \ldots p_{k2}\lambda + p_{k1}$$

where $\lambda_{kj} \in \Lambda_k$ etc. These are the desired relative minimal polynomials of $\bar{R}_1 \ldots \bar{R}_r$ with respect to A-DC (Section 2.3). If $\nu'_{\mathcal{J}}$ is defined as:

$$\sum_{k=1,r} \nu_k = \nu'_{\mathcal{J}} \leq n \tag{4.2.40}$$

then the union of $\Lambda_1 \ldots \Lambda_r$:

$$\Lambda'_{\mathcal{J}} = \bigcup_{k=1,r} \Lambda_k \tag{4.2.41}$$

is a $\nu'_{\mathcal{J}}$- element subset of the desired Λ.

Since $f_1 \ldots f_r$ are cutput separable (Definition 3.3.2), Theorem 3.3.1 states that a D can be chosen such that:

i) the output errors generated by the failures associated with $f_1 \ldots f_r$ are unidirectional, and

ii) $\Lambda'_{\mathcal{J}}$is a $\nu'_{\mathcal{J}}$ - element subset of the eigenvalues of A-DC.

A straightforward generalization of Lemma 3.2.7 provides an algorithm for finding this D. With the detection generators $g_1 \ldots g_r$ (Section 4.2.1), and with $\Psi_1(.) \ldots \Psi_r(.)$ written as polynomials in A, an n x r matrix is formed:

$$\Psi_F(A) = [\Psi_1(A)\, g_1 \ldots \Psi_r(A)\, g_r] \tag{4.2.42}$$

D is then any solution to:

$$DCF = \Psi_F(A) \tag{4.2.43}$$

A solution to (4.2.43) is guaranteed and can be found using conventional matrix analysis techniques (Gantmacher(1959) or Helmos(1958)). The difference between (4.2.38) and (4.2.43) is that (4.2.43) implicitly contains the constraint that any solution D satisfies both conditions i) and ii) above.

If r equals rkC, (4.2.43) has a unique solution. If not, the algorithm developed in Appendix A.4 choses a particular solution. In either case, if $f_1 \ldots f_r$ are mutually detectable (Definition 3.3.1),

a solution of (4.2.43) is chosen such that the eigenvalues of A-DC are the elements of Λ. If $f_1 \ldots f_r$ are not mutually detectable, however, for every solution of (4.2.43) certain eigenvalues of A-DC are fixed and unassignable. These are the ν_0 eigenvalues of Π as discussed in Section 4.2.2.

The remaining task is to define an eigenvalue assignment algorithm when output stationarity is being utilized. As in Section 4.2.3, assume h_i is to be made output stationary with h_i and that $f_1 \ldots f_\ell$ is the smallest subset of $f_1 \ldots f_r$ such that:

$$C F_L^c \alpha_L^c = \beta_i C h_i \tag{4.2.24}$$

has a solution where:

$$\begin{aligned} \alpha_L^c &= [\alpha_1 \ldots \alpha_\ell]^T \\ F_L^c &= [f_1 \ldots f_\ell]^T \end{aligned} \tag{4.2.25}$$

The eigenvalues of A-DC associated with $\bar{R}_{\ell+1} \ldots \bar{R}_r$ are unaffected by the output stationarity requirement. They can be assigned as before by requiring D to satisfy the $r-\ell$ equations:

$$DC F_L = \Psi_{FL}(A) \tag{4.2.44}$$

where F_L is the last $r-\ell$ columns of F and $\Psi_{FL}(A)$ is the last $r-\ell$ columns of $\Psi_F(A)$.

The effect of output stationarity is to restrict the freedom with which the eigenvalues of A-DC associated with $\bar{R}_1 \ldots \bar{R}_\ell$ can be assigned. If ν_j is the dimension of $\bar{\mathcal{J}}_i$, a single ν_j-element set Λ_j and an associated polynomial $Y_j(\cdot)$ can be freely chosen. Theorem 3.4.2 then requires that the elements of Λ_j be eigenvalues of A-DC associated with each of $\bar{R}_1 \ldots \bar{R}_\ell$. The net effect is that only ν_j unique eigenvalues can be chosen and each must be assigned a multiplicity of ℓ. Corollary 1 of Theorem 3.4.2 provides an algorithm for assuring that an acceptable D is chosen. The ℓ linear equations from the corollary plus the $r-\ell$ linear equations implied by (4.2.44) together replace (4.2.43). A

solution of these equations is guaranteed and it is unique if r is equal to rkC. Otherwise, the algorithm is Appendix A.4 can be used to find a unique solution.

If $\nu_1 \ldots \nu_\ell$ are the dimensions of $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_\ell$, Theorem 3.4.2 states that ν_ρ satisfies:

$$\nu_\rho \leq \min (\nu_1 \ldots \nu_\ell) \tag{4.2.45}$$

Whereas (4.2.43) permits all $\nu'_\jmath$ eigenvalues of A-DC associated with $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_\ell$ to be assigned, the above discussion implies that output stationarity of h_i permits only $\ell\nu_\rho$ of the eigenvalues to be assigned. The remaining $\nu'_\jmath - \ell\nu_\rho$ eigenvalues are fixed for every acceptable choice of D. These unassignable eigenvalues should not be confused with those of Π discussed earlier; these are a consequence of the desire to make h_i output stationary with $f_1 \ldots f_r$. Whether or not it is feasible to design a failure detection system to detect the failures associated with each of $f_1 \ldots f_r$ and also the failure associated h_i depends upon the acceptability of these unassignable eigenvalues. An algorithm for determining the unassignable eigenvalues is presented in the discussion following Corollary 1 of Theorem 3.4.2.

Generalization of the eigenvalue assignment algorithms to $h_1 \ldots h_q$ follows the same line developed in Section 4.2.3. If h_i and h_j are two event vectors which satisfy (4.2.31), then Corollary 1 of Theorem 3.4.2 can be applied independently for each of them. If they satisfy (4.2.35) and their normal subspaces $\bar{\mathcal{N}}^\circ_i$ and $\bar{\mathcal{N}}^\circ_j$ are equal, then the corollary need only be applied for one of h_i and h_j. The general case is discussed in conjunction with Theorem 3.4.4. When used with the appropriate variation of (4.2.44), each of these algorithms defines a set of r consistent equations which D must satisfy.

4.3 An Example

This section provides an application of the detection filter design algorithm of Section 4.2. To the extent that it is possible, an attempt is made to motivate the design procedure physically. The satellite example was chosen for its simplicity and because the subsystems defined by the failure detection requirements are the same subsystems which would be defined for a particular decoupling problem. Although the decoupling analogy is not generally useful, it is helpful in this instance for generating a physical understanding of the solution. Unfortunately, the satellite example is simple enough that some results of Chapter 3 are not required. The problem of unassignable eigenvalues is treated in Chapter 6 for the lateral autopilot example.

4.3.1 Satellite Orbit Perturbation Model

The failure detection system is to monitor the performance of an orbital maneuvering system on a satellite in a near circular orbit. Only the open loop system without the control law is to be modeled. The satellite is assumed to have an autonomous navigation system and to have onboard thrusters for orbit perturbation. To keep the analysis simple, only linearized equations of motion about a nominal circular orbit are considered and all orbit perturbations are assumed to be in the orbital plane. The geometry of the problem is indicated in Figure 4.2.

The output of the navigation system is assumed to be an altitude measurement r(t) and a true anomaly measurement $\varphi(t)$. These measurements uniquely determine the satellite's position in the orbital plane. The measurements may be provided either by an altimeter, a star tracker, and horizon indicators; or by an inertial navigation system. The dynamics of the measurement system are assumed to be negligible and the measurements to be noise free. It is assumed, however, that the sensors providing r(t) and $\varphi(t)$ may fail.

The thrusters are grouped as three sets of paired (opposite polarity) thrusters, each set having its own fuel supply. If one of a pair of thrusters fails, that pair is considered to have failed.

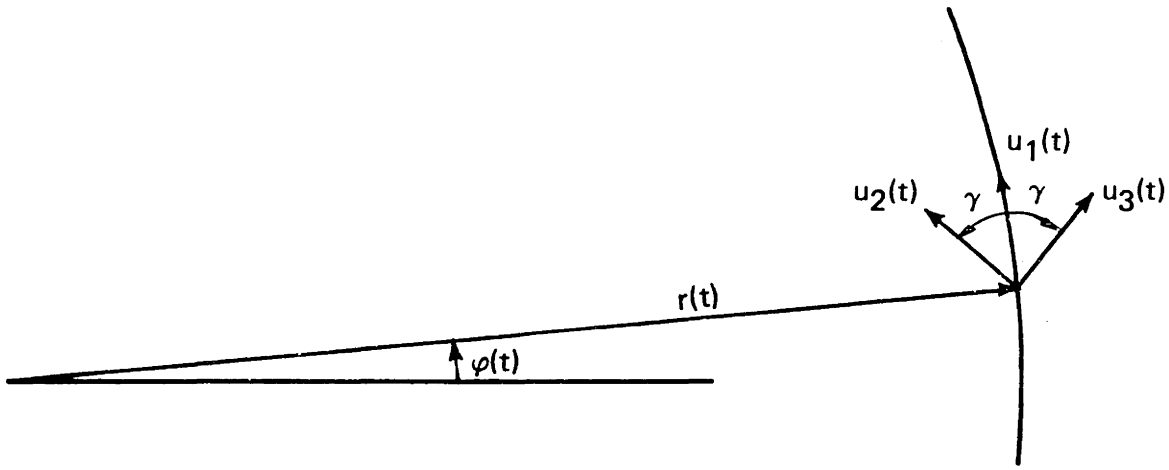


Figure 4.2 Geometry for satellite orbital maneuvering system. Thruster pairs are aligned along the tangent to the orbit, $u_1(t)$, and along $u_2(t)$ and $u_3(t)$.

If a thruster fails in the "on" mode and the failure is detected, the failure is compensated for by the controller which switches on the remaining thruster of that set until the fuel supply is depleted. The thruster pairs are aligned with one set along the velocity vector and the remaining two at angles of $\pm \gamma$ respectively to the velocity vector. The attitude control system for the satellite is assumed to maintain this alignment perfectly. As long as $\gamma \neq 90^{\circ}$ only one operational thruster pair is necessary to effect an orbit change. This provides an example of a triply redundant system as discussed in Section 3.4.2. Under nominal operating conditions the controller may use any or all of the thruster pairs to effect a maneuver.

The system failures to be monitored for the open loop system are:

i) a failure of any thruster pair, and
ii) a failure of either sensor.

The output space is only two dimensional and sensor failures in general produce planar output for failure detection systems. Because of this, if an error signal occurs which is planar, it is not possible to determine whether a sensor failure or some unknown and unmodeled failure produced the error signal. Unambiguous detection of sensor failures requires more than two sensors unless one can be constrained to generate a unidirectional output error signal. With this ambiguity accepted, a single detection system can be designed which monitors all of the above failures.

The general planar equations of motion for a satellite can be specified in polar coordinates:

$$\ddot{r}(t) = r(t)\,\dot{\varphi}(t)^2 - K/r(t)^2 - u_2(t)\sin\gamma + u_3(t)\sin\gamma \quad (4.3.1)$$

$$r(t)\,\ddot{\varphi}(t) = -2\dot{\varphi}(t)\,\dot{r}(t) + u_1(t) + u_2(t)\cos\gamma + u_3(t)\cos\gamma \quad (4.3.2)$$

where K is the gravitational constant and $u_1(t)$, $u_2(t)$, and $u_3(t)$ are the accelerations due to the repective thruster pairs. The equations are to be linearized about the circular orbit defined by the radius σ and angular rate ω:

$$\begin{aligned} r(t) &= \sigma \\ \varphi(t) &= \omega t \\ \sigma^3 \omega^2 &= K \end{aligned} \tag{4.3.3}$$

If a system is described by the nonlinear equation:

$$\dot{x}(t) = f[x, u, t] \tag{4.3.4}$$

then the linearized model is:

$$\delta \dot{x}(t) = A\delta x(t) + B\delta u(t) \tag{4.3.5}$$

The elements of A and B satisfy:

$$\begin{aligned} a_{ij} &= \frac{\partial f_i}{\partial x_j}[x, u, t] \Big|_{x_o, u_o} \\ b_{ij} &= \frac{\partial f_i}{\partial u_j}[x, u, t] \Big|_{x_o, u_o} \end{aligned} \tag{4.3.6}$$

where $x_o(t)$, $u_o(t)$ are the nominal state and control vectors respectively. If A and B are to be time invariant, the solutions of (4.3.6) must not be either explicit or implicit functions of time.

Equations (4.3.1) and (4.3.2) are linear in $u_1(t)$, $u_2(t)$, and $u_3(t)$ and the variational control vector can be defined by:

$$\delta u(t) = \begin{bmatrix} u_1(t) \\ u_2(t) \\ u_3(t) \end{bmatrix} \qquad (4.3.7)$$

The choice of $\delta x(t)$ is slightly more difficult. Define $\delta x(t)$ by:

$$\delta x(t) = \begin{bmatrix} \delta x_1(t) \\ \delta x_2(t) \\ \delta x_3(t) \\ \delta x_4(t) \end{bmatrix} = \begin{bmatrix} r(t) - \sigma \\ \dot{r}(t) \\ \sigma(\varphi(t) - \omega t) \\ \sigma(\dot{\varphi}(t) - \omega) \end{bmatrix} \qquad (4.3.8)$$

The first two components are the satellite displacement in a radial direction from the nominal trajectory and the rate of that displacement respectively. The last two components are the displacement and displacement rate from the nominal trajectory in a tangential direction. Equation (4.3.8) allows the perturbation to be expressed in an orthogonal coordinate system centered at the nominal state. Then the triad (A, B, C) for the orbital maneuvering system is:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 3\omega^2 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\sin\gamma & \sin\gamma \\ 0 & 0 & 0 \\ 0 & \cos\gamma & \cos\gamma \end{bmatrix}$$

(4.3.9)

The choice of $\delta x(t)$ yields the desired time invariant system model.

It is simple to show that the variational system (A, B, C) is observable and controllable. In fact, the orbital state is completely controllable by any of the thruster pairs and observable by the true anomaly sensor. It is not observable by the altimeter.

4.3.2 Detection Filter Design

The failure models for the satellite can be determined in accordance with Section 4.1. As suggested in Section 4.2.1, it is best to determine the detection spaces and failure models for failures with known event vectors before considering the sensor failures.

Since the attitude control system of the satellite is assumed to be perfect, thruster pair failures can be modeled as errors $\Delta u_1(t)$, $\Delta u_2(t)$, $\Delta u_3(t)$ respectively in the achieved thrust. A failure in the first thruster set can then be modeled as an additional $B\Delta u(t)$ term in (4.3.5). This is the control-dependent controller failure model (Table 4.1) with:

$$B\Delta u(t) = e_{44}\, \Delta u_1(t) \tag{4.3.10}$$

and the event vector is:

$$f_1 = e_{44} \tag{4.3.11}$$

The failure could also be modeled as a control-independent failure if $\Delta u_1(t)$ were a bias etc. The precise choice of the model is not important since the event vector is the same for all models.

Failures in the second and third thruster pairs involve changes in the second and fourth rows of $B\delta u(t)$. For the second thruster pair:

$$B\Delta u(t) = \begin{bmatrix} 0 \\ -\Delta u_2(t) \sin\gamma \\ 0 \\ \Delta u_2(t) \cos\gamma \end{bmatrix} \tag{4.3.12}$$

Since the rows are linearly dependent for all t, the discussion in conjunction with (4.1.10) implies that a single event vector is necessary for the failure model and the control-dependent controller failure model is:

$$B\Delta u(t) = (-e_{42} \sin\gamma + e_{44} \cos\gamma)\, \Delta u_2(t) \qquad (4.3.13)$$

The event vector is the constant vector in parentheses:

$$f_2 = (-e_{42} \sin\gamma + e_{44} \cos\gamma) \qquad (4.3.14)$$

Similarly for a failure in the third thruster set the event vector is:

$$f_3 = (e_{42} \sin\gamma + e_{44} \cos\gamma) \qquad (4.3.15)$$

All three of the event vectors f_1, f_2, f_3 satisfy:

$$Cf_i = 0 \qquad i = 1,2,3$$

Therefore, to determine the detection spaces using D_f and C' from (4.2.1), it is necessary to use Af_1, Af_2, Af_3. With D_f and C', the orthogonal reduction procedure of Appendix A can be used to find first R_1, R_2, and R_3 and then $\bar{R}_1$, $\bar{R}_2$, $\bar{R}_3$. It can also be used to find the respective detection generators. In this case, the detection generators are f_1, f_2, and f_3 respectively:

$$\bar{R}_1 = [Af_1 \;\; f_1] = \begin{bmatrix} 0 & 0 \\ 2\omega & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad g_1 = f_1$$

$$\bar{R}_2 = [Af_2 \;\; f_2] = \begin{bmatrix} -\sin\gamma & 0 \\ 2\omega\cos\gamma & -\sin\gamma \\ \cos\gamma & 0 \\ 2\omega\sin\gamma & \cos\gamma \end{bmatrix} \qquad g_2 = f_2 \qquad (4.3.16)$$

$$\bar{R}_3 = [Af_3 \;\; f_3] = \begin{bmatrix} \sin\gamma & 0 \\ 2\omega\cos\gamma & \sin\gamma \\ \cos\gamma & 0 \\ -2\omega\sin\gamma & \cos\gamma \end{bmatrix} \qquad g_3 = f_3$$

The detection spaces $\bar{R}_1$, $\bar{R}_2$, $\bar{R}_3$ are spanned by the columns of the appropriate matrices in (4.3.16).

The physical importance of a detection space can be motivated by considering $\bar{R}_1$ for the first thruster pair. This thruster pair is tangential to the orbital path. Physical reasoning suggests that a firing of the tangential thrusters first causes a change in the true anomaly perturbation $\delta x_3(t)$ and that the ensuing perturbation $\delta x_1(t)$ in the radius of the orbit requires a portion of an orbital period to develop. The columns of $\bar{R}_1$ are Af_1 and f_1 respectively. They satisfy the output relations:

$$\begin{aligned} Cf_1 &= 0 \\ CAf_1 &= \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{aligned} \tag{4.3.17}$$

Equation (4.3.17) and the discussion in Section 2.1 verify the intuitive notion that a failure in the tangential thrusters is first noticed as an error in the true anomaly perturbation estimate $\delta x_{m3}(t)$. A detection filter chooses A-DC such that a failure in the tangential thrusters affects only $\delta x_{m3}(t)$. If the satellite system had only a pair of tangential thrusters and a pair of radial thrusters, the failure detection system also would constrain the output estimation error due to a radial thruster failure to $\delta x_{m1}(t)$. For this particular two thruster pair system, the detection filter would be the solution to the output decoupling problem with output feedback. The decoupling analogy is not directly applicable to arbitrary systems.

After $\bar{R}_1$, $\bar{R}_2$, $\bar{R}_3$ have been found, i) - iii) outlined in Section 4.2.1, should be evaluated for the sensor failure models. The results are not important since the output space is only two dimensional but they do illustrate the algorithm.

The failure model for the true anomaly sensor is a good example of option i). From (4.3.17) it follows that Af_1 satisfies

the requirement and the true anomaly sensor failure can be characterized by the event vectors $\{Af_1, A^2f_1\}$. The corresponding detection spaces are defined by the set of matrices $\{\bar{R}_1, \bar{R}_4\}$. $\bar{R}_4$ is the detection space for A^2f_1:

$$\bar{R}_4 = [A^2f_1 \;\; g_4] = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ -2\omega & 0 \end{bmatrix} \qquad g_4 = e_{42}$$

A failure of the altitude sensor cannot be satisfied by option i) for any vector in $\bar{R}_1$, $\bar{R}_2$, or $\bar{R}_3$. The alternative requires use of option ii). Af_2 and Af_3 from $\bar{R}_2$ and $\bar{R}_3$ respectively satisfy:

$$C\left(-\frac{\csc\gamma}{2} Af_2 + \frac{\csc\gamma}{2} Af_3\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad (4.3.19)$$

If f_2 and f_3 are detected by D, then d_1 is completely specified and an altitude sensor failure can be characterized by the three event vectors $\{f_2, f_3, d_1\}$. An obvious second alternative is to use option iii) to find some arbitrary set $\{f, Af\}$ rather than the three event vectors.

For the remainder of this section, only the thruster failures are considered. Mutual detectability of the three failures can be investigated using the techniques of Section 4.2.2, although the results are simple. If F is defined for the thruster pair event vectors:

$$F = [f_1 \; f_2 \; f_3] \qquad (4.3.20)$$

it is simple to show that rk CF < rk F and f_1, f_2, f_3 are not output separable (Definition 3.3.1). By Theorem 3.5.1, they are not mutually detectable by a single detection filter.

The next step is to attempt to find subsets of f_1, f_2, f_3 which are mutually detectable. f_2, f_3 are considered as an example. If F_1 is defined by removing f_1 from F, $\bar{R}_{F1}$ can be found using (4.2.8), (4.2.9), and Appendix A.2. In this case, however, since the state space is four-dimensional and:

$$\text{rk}\,[\bar{R}_2 : \bar{R}_4] = 4 \tag{4.3.21}$$

inspection of (4.2.11) verifies that:

$$\text{rk}\,\bar{R}_{F1} = \text{rk}\,\bar{R}_2 + \text{rk}\,\bar{R}_3 = 4 \tag{4.3.22}$$

By Theorem 3.3.1, f_2 and f_3 are mutually detectable. It is possible to show through similar reasoning that any two of f_1, f_2, f_3 are mutually detectable.

Since all three thruster pair failures are to be monitored, one solution is to design two failure detection systems-- one for f_2 and f_3 and the other for f_1. A more acceptable alternative is to use only a single failure detection system and make f_1 output stationary with f_2 and f_3. In this context, f_1 and $\bar{\mathcal{R}}_1$ correspond to h_i and $\bar{\mathcal{I}}_i$ of Section 4.2.3. Since Af_1 satisfies:

$$CAf_1 = \frac{\sec\gamma}{2} CAf_2 + \frac{\sec\gamma}{2} CAf_3 \tag{4.3.23}$$

the index set I_1 can be defined:

$$I_1 = \{2,3\} \tag{4,3.24}$$

Theorem 3.4.1 requires that $\bar{\mathcal{R}}_1$ satisfy (4.2.26) if f_1 is to be a candidate for mutual output stationarity. In this case, since:

$$\text{rk}\,\bar{R}_{FI_1} = \text{rk}[\bar{R}_2 : \bar{R}_3] = 4 \tag{4.3.25}$$

$\bar{R}_{FI_1}$ is non-singular and a unique solution to (4.2.26) is guaranteed.

It follows immediately that f_1 can be made output stationary with f_2 and f_3.

In general, it is necessary to apply Theorem 3.4.1 to determine whether or not a vector h_i can be made output stationary with $f_1 \ldots f_r$. For the satellite application, there is an intuitive reason to suspect that f_1 can be made output stationary with f_2 and f_3. The three thruster pairs are an example of parallel redundancy. Although each thruster pair has a different effect on the satellite and its sensors, the effects are dynamically similar. An intuitive agrument suggests that output stationarity is often applicable in designing failure detection systems for reference models with parallel redundancy.

The remaining problem concerns eigenvalue assignment. Since each of $\bar{R}_1$, $\bar{R}_2$ and $\bar{R}_3$ is two-dimensional, Theorem 3.4.2 guarantees that all four eigenvalues of A-DC can be assigned. Because of the output stationarity constraint, however, the two eigenvalues associated with $\bar{R}_2$ must be the same as those associated with $\bar{R}_3$. This requirement is a necessary, but not sufficient condition for output stationarity.

The eigenvalue assignment algorithm is given by Corollary 1 of Theorem 3.4.2. The first step is to choose a set of desired eigenvalues in terms of the polynomial $\Upsilon_1(\cdot)$:

$$\Upsilon_1(\lambda) = \lambda^2 + p_2\lambda + p_1 \tag{4.3.26}$$

and then to determine:

$$\Upsilon_1(A)g_1 = \begin{bmatrix} 2\omega \\ 0 \\ 0 \\ -4\omega^2 \end{bmatrix} + p_2 \begin{bmatrix} 0 \\ 2\omega \\ 1 \\ 0 \end{bmatrix} + p_1 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 2\omega \\ 2\omega p_2 \\ p_2 \\ p_1 - 4\omega^2 \end{bmatrix} \tag{4.3.27}$$

The next step is to determine the bracketed quantity in the corollary. Since Cf_2 and Cf_3 are both zero, Af_2 and Af_3 must be replaced by A^2f_2 and A^2f_3 in the algorithm:

$$[Y(A)g_1 - \sum_{j=2,3} \alpha_j A^2 f_j] = p_2 \begin{bmatrix} 0 \\ 2\omega \\ 1 \\ 0 \end{bmatrix} + p_1 \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \tag{4.3.28}$$

An evaluation of the projections of this vector on the range spaces of R_2 and R_3 respectively then yields the two equations for D:

$$DCAf_2 = p_2 \begin{bmatrix} -\sin\gamma \\ 2\omega\cos\gamma \\ \cos\gamma \\ 2\omega\sin\gamma \end{bmatrix} + p_1 \begin{bmatrix} 0 \\ -\sin\gamma \\ 0 \\ \cos\gamma \end{bmatrix} + A^2 f_2$$

$$DCAf_3 = p_2 \begin{bmatrix} \sin\gamma \\ 2\omega\cos\gamma \\ \cos\gamma \\ -2\omega\sin\gamma \end{bmatrix} + p_1 \begin{bmatrix} 0 \\ \sin\gamma \\ 0 \\ \cos\gamma \end{bmatrix} + A^2 f_3 \tag{4.3.29}$$

Equations (4.3.29) uniquely define the detection filter D:

$$D = \begin{bmatrix} p_2 & 2\omega \\ -\omega^2 + p_1 & 2\omega p_2 \\ -2\omega & p_2 \\ -2\omega\, p_2 & -4\omega^2 + p_1 \end{bmatrix} \tag{4.3.30}$$

and the design algorithm is complete.

The simplest check of the design procedure is to determine A-DC:

$$\text{A-DC} = \begin{bmatrix} -p_2 & 1 & -2\omega & 0 \\ 4\omega^2 - p_1 & 0 & -2\omega p_2 & 2\omega \\ 2\omega & 0 & -p_2 & 1 \\ 2\omega p_2 & -2\omega & 4\omega^2 - p_1 & 0 \end{bmatrix} \qquad (4.3.31)$$

It is possible to verify that the characteristic equation of (4.3.31) is:

$$\chi_{\text{A-DC}}(\lambda) = (\lambda^2 + p_2\lambda + p_1)^2 \qquad (4.3.32)$$

A simpler and equivalent test is to verify that each of f_1, f_2, f_3 satisfies the equation:

$$Y_1(\text{A-DC})\, g_i = 0 \qquad i = 1,2,3 \qquad (4.3.33)$$

Equation (4.3.33) verifies that $\bar{\mathcal{R}}_1$, $\bar{\mathcal{R}}_2$ and $\bar{\mathcal{R}}_3$ are all invariant for A-DC and that the eigenvalues associated with each are the roots of $Y_1(\cdot)$.

This completes the design algorithm for the satellite example. If there were more than two sensors, (4.3.29) would not assign all n eigenvalues of A-DC. The remaining eigenvalues could be assigned using Appendix A.4.

4.4 Base Normal Form

The design algorithms in Section 4.2 were based upon the derivation of failure detection theory presented in Chapter 3. The development in Chapter 3 emphasized linear algebra techniques which are analytically straightforward rather than computationally efficient. In particular, the development did not utilize a particular coordinate system which can make the selection of a detection filter trivial. A reference model expressed in this coordinate system is defined to be in base normal canonical form. The base normal

canonical form has an additional advantage in that it permits the reference model to be decoupled into a set of lower-dimensioned subsystems. The properties of the base normal form for deterministic systems and the transformation matrix to base normal form are developed in this section. The importance of the base normal form for stochastic systems is developed in Chapter 5.

Assume that the event vectors $f_1 \ldots f_r$ are output separable. If D is a detection filter for $f_1 \ldots f_r$, then the failures associated with $f_1 \ldots f_r$ generate unidirectional output errors for (A-DC, C), i.e. the detection spaces $\bar{R}_1 \ldots \bar{R}_r$ are invariant with respect to A-DC (Lemma 3.3.1). If ξ is an arbitrary vector in $\bar{R}_j$ for some $j \leq r$, then invariance requires that (A-DC)ξ be a linear combination of the vectors which span $\bar{R}_j$. This suggests that if a coordinate system for $\mathcal{E}^n$ is chosen such that the basis vectors for each of $\bar{R}_1 \ldots \bar{R}_r$ are unit vectors, then A-DC is at least partially block diagonal. This block diagonality when the reference model is expressed in the proper coordinate system is a general property of invariant subspaces.

The crucial fact in the above discussion is that a coordinate system in which A-DC is partially block diagonal is a function only of $\bar{R}_1 \ldots \bar{R}_r$. The transformation to this coordinate system can be constructed from the basis vectors for $\bar{R}_1 \ldots \bar{R}_r$ independent of any choice of the detection filter D. The basis defined in Lemma 3.2.4 for each of $\bar{R}_1 \ldots \bar{R}_r$ is particularly useful. If g_1 is the detection generator for $\bar{R}_1$, the indicated basis for $\bar{R}_1$ is $g_1, Ag_1 \ldots A^{\nu_1 - 1} g_1$; similar bases exist for $\bar{R}_2 \ldots \bar{R}_r$. The base normal canonical form is defined in terms of these basis vectors.

The transformation matrix from the given coordinate system to base normal canonical form is defined to be T^{-1}. If $r = m$ and $f_1 \ldots f_r$ are mutually detectable, T^{-1} is uniquely determined by the bases for $\bar{R}_1 \ldots \bar{R}_r$ given in Lemma 3.2.4; otherwise an additional set of vectors must be specified. The general structure for T^{-1} is:

$$T^{-1} = [\, g_1 \ldots A^{\nu_1 - 1} g_1 \quad g_2 \ldots A^{\nu_2 - 1} g_2 \quad \ldots g_r \ldots A^{\nu_r - 1} g_r : T_0 \,]^{-1} \tag{4.4.1}$$

where the principal constraint on the submatrix T_0 is that it be chosen such that T is nonsingular:

$$T_0 = [\, z_1 \ldots z_{\nu_0} \quad w_{r+1} \ldots (A - D_{\mathcal{J}} C)^{q_{r+1} - 1} w_{r+1} \ldots w_m \ldots (A - D_{\mathcal{J}} C)^{q_m - 1} w_m \,]$$

This particular T_0 is chosen such that $z_1 \ldots z_{\nu_0}$ are a basis for the excess subspace $\mathcal{R}_{0g}$ associated with $\bar{R}_1 \ldots \bar{R}_r$, and the remaining vectors are defined in Appendix A.4.[1] All but m of the vectors on the right-hand side of (4.4.1) are in the null space of C. These m columns can be used to define a transformation of the output space compatible with T^{-1}:

$$T_m^{-1} = [\, CA^{\nu_1 - 1} g_1 \ \ldots CA^{\nu_r - 1} g_r \quad C'_{\mathcal{J}} (A - D_{\mathcal{J}} C)^{q_{r+1} - 1} w_{r+1} \ \ldots C'_{\mathcal{J}} (A - D_{\mathcal{J}} C)^{q_m - 1} w_m \,]^{-1} \tag{4.4.2}$$

The transformation to base normal form is a similarity transformation--T^{-1} transforms the state vector to base normal form and T transforms it back to the original coordinate system. The reference model and state vector in base normal form are denoted by:

$$\begin{aligned}
\hat{A} &= T^{-1} A T \\
\hat{B} &= T^{-1} B \\
\hat{C} &= T_m^{-1} C T \\
\hat{D} &= T^{-1} D T_m \\
\hat{\bar{R}}_i &= T^{-1} \bar{R}_i \\
\hat{x}_m(t) &= T^{-1} x_m(t) \\
\hat{y}_m(t) &= T_m^{-1} y_m(t)
\end{aligned} \tag{4.4.3}$$

---

[1] Recall from Section 4.2.2 that

$$\bar{R}_{\mathcal{J}} = \bar{R}_1 \oplus \ldots \bar{R}_r \oplus \mathcal{R}_{0g} \tag{4.2.13}$$

and that $\mathcal{R}_{0g}$ is zero if $f_1 \ldots f_r$ are mutually detectable. The dimension of $\bar{R}_{\mathcal{J}}$ is denoted by $\nu_{\mathcal{J}}$ and the dimension of $\mathcal{R}_{0g}$ is denoted by ν_0.

An implementation of the reference model in base normal form is shown in Figure 4.3. The reference models $(\hat{A}-\hat{D}\hat{C},\ \hat{C})$ and $(A-DC, C)$ would be equivalent realizations of the same system if $\hat{y}_m(t)$ were in the original coordinate system, i.e., if the output transformation matrix T_m^{-1} did not appear in (4.4.3). The advantage of the transformed output space is demonstrated later. The disadvantage is that the system output y(t) must be transformed before it is incorporated into the reference model estimates.

With T^{-1} and T_m^{-1} defined, it is possible to begin a systematic development of the structure of $\hat{A}, \hat{C}$, and $\hat{D}$. The system matrix $\hat{A}$ is of the form:

$$\hat{A} = \begin{bmatrix} \hat{A}'_{11} & \hat{A}_{12} & \cdots & \hat{A}_{1r} & \hat{\Theta}_1 & \hat{\Gamma}'_1 \\ \hat{A}_{21} & \hat{A}'_{22} & & & \hat{\Theta}_2 & \hat{\Gamma}'_2 \\ \vdots & & & & \vdots & \vdots \\ \hat{A}_{r1} & & & \hat{A}'_{rr} & \hat{\Theta}_r & \hat{\Gamma}'_r \\ \hat{A}_{01} & & \cdots & \hat{A}_{0r} & \Pi & \hat{\Gamma}'_0 \\ \hat{A}_{\Gamma 1} & & \cdots & \hat{A}_{\Gamma\Gamma} & 0 & \hat{A}'_{\Gamma\Gamma} \end{bmatrix} \qquad (4.4.4)$$

The submatrices $\hat{A}'_{\Gamma\Gamma}$ and $\hat{\Gamma}'_1 \ldots \hat{\Gamma}'_r$, Γ'_0 describe the dynamics of A which are not associated with $\bar{\mathcal{R}}_{\mathcal{J}}$ (see previous footnote). $\hat{A}_{\Gamma\Gamma}$ is $(n-\nu_{\mathcal{J}}) \times (n-\nu_{\mathcal{J}})$ and $\hat{\Gamma}'_1 \ldots \hat{\Gamma}'_r$ are dimensioned accordingly. The matrix Π is the $\nu_0 \times \nu_0$ matrix associated with $\mathcal{R}_{0g}$ in (4.2.15) -- this can be verified by transforming (4.2.15) to base normal form. Similarly, $\hat{\Theta}_1 \ldots \hat{\Theta}_r$ are matrices which have only one non-zero row corresponding to a row of Θ as defined in (4.2.15):

$$\hat{\Theta}_i = \begin{bmatrix} \theta_i \\ \hline 0 \end{bmatrix} \qquad (4.4.5)$$

The remaining submatrices of $\hat{A}$ are related to the detection spaces $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$. The purpose of the transformation is to place these

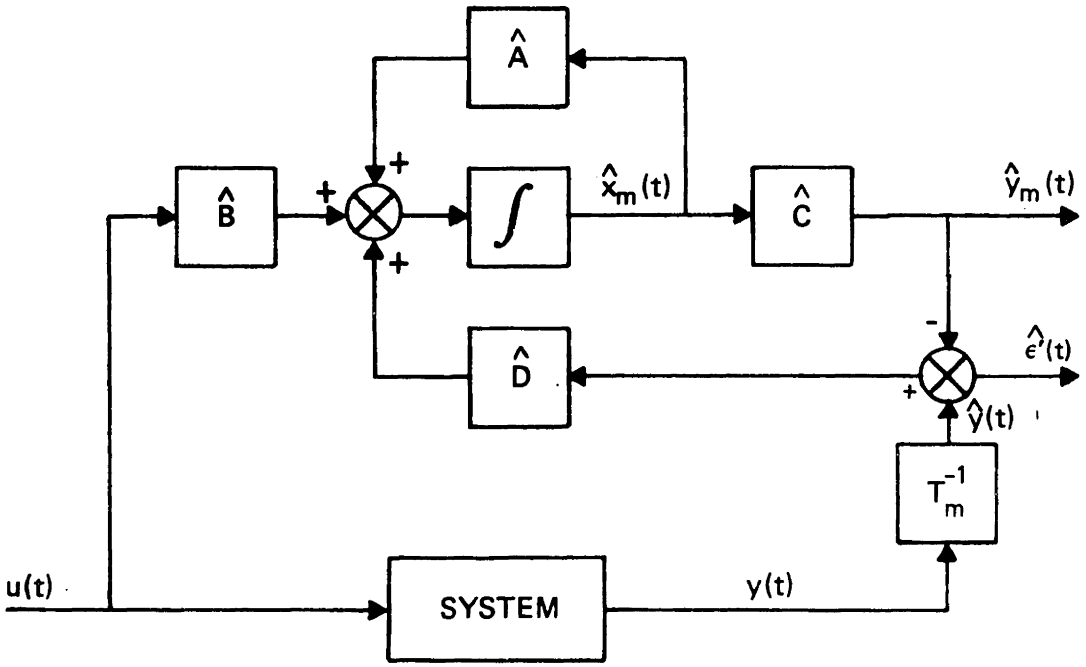


Figure 4.3 Implementation of failure detection system in base normal form.

submatrices in a standard form. The off-diagonal matrices $\hat{A}_{ij}$ are $\nu_i \times \nu_j$ matrices with only a single non-zero column:

$$\hat{A}'_{ij} = \begin{bmatrix} 0 \cdots\cdots 0 & \hat{a}_{ij1} \\ \vdots \quad\quad \vdots & \vdots \\ 0 \cdots\cdots 0 & \hat{a}_{ij\nu_i} \end{bmatrix} = \left[\begin{array}{c|c} 0 & \hat{a}_{ij} \end{array}\right] \begin{array}{l} i = 1,..r,0,\Gamma \\ j = 1...r \\ i \neq j \end{array} \qquad (4.4.6)$$

The diagonal submatrices are $\nu_i \times \nu_i$ matrices of the form:

$$\hat{A}'_{ii} = \begin{bmatrix} 0 & \cdots & 0 & -p'_{i1} \\ 1 & \ddots & \vdots & \vdots \\ 0 & \ddots & 0 & \\ 0 & \cdots & 1 & -p'_{i\nu_i} \end{bmatrix} \qquad (4.4.7)$$

The structure of the submatricies in (4.4.6) and (4.4.7) can be motivated by considering $\hat{A}'_{11}$. As suggested earlier, a basis for $\bar{\mathcal{R}}_1$ in the base normal coordinate system is the first ν_1 unit vectors $\hat{e}_{n1} \cdots \hat{e}_{n\nu_1}$. The detection generator g_1 in base normal form is $\hat{e}_{n1}$ since g_1 is the first column of T:

$$g_1 = T\,\hat{e}_{n1} \qquad (4.4.8)$$

It follows similarly that:

$$A^j g_1 = T\,\hat{e}_{n(j+1)} \qquad j \leq \nu_1 - 1 \qquad (4.4.9)$$

However, $A^j g_1$ can be expressed recursively in terms of $A^{j-1} g_1$:

$$A^j g_1 = A(A^{j-1}) g_1 \qquad (4.4.10)$$

and it follows that (4.4.9) and (4.4.10) then imply:

$$T\,\hat{e}_{n(j+1)} = A\,T\,\hat{e}_{nj} \qquad j \leq \nu_1 - 1 \qquad (4.4.11)$$

Premultiplication by T^{-1} and substitution of $\hat{A}$ then yields an expression for $\hat{A}\,\hat{e}_{nj}$:

$$\hat{e}_{n(j+1)} = \hat{A}\,\hat{e}_{nj} \qquad j \leq \nu_1 - 1 \qquad (4.4.12)$$

This requires that the first $\nu_1 - 1$ columns of $\hat{A}$ be zero, except that the $(j+1)^{th}$ element of the j^{th} column is 1 for $j \leq \nu_1 - 1$. Because $A^{\nu_1} g_1$ need not be contained in $\bar{R}_1$ (Lemma 3.2.4), the ν_1^{th} column of A is arbitrary. Hence, $\hat{A}_{11}$ and $\hat{A}_{21} \ldots \hat{A}_{\Gamma 1}$ are given by (4.4.6) and (4.4.7) respectively. Equations (4.4.8) - (4.4.12) can be generalized to $\bar{R}_2 \ldots \bar{R}_r$ and subsequent columns of $\hat{A}$.

The output matrix $\hat{C}$ has a simpler structure than $\hat{A}$:

$$\hat{C} = \begin{bmatrix} \hat{c}_1 & 0 & \cdots & 0 \\ 0 & \hat{c}_2 & 0 & 0 \\ & & \ddots & \\ 0 & \cdots & 0 & \hat{c}_r \\ 0 & \cdots & 0 & \hat{c}_{\Gamma m} \end{bmatrix} \qquad (4.4.13)$$

where $\hat{c}_1 \ldots \hat{c}_r$ are unit vectors in $\nu_1 - \ldots \nu_r -$ dimensional subspaces:

$$\hat{c}_i = \hat{e}^T_{\nu_i \nu_i} = [0 \ldots 0 \quad 1] \qquad i = 1 \ldots r$$

To verify (4.4.13), consider the basis $\hat{e}_{n1} \ldots \hat{e}_{n\nu_1}$ for $\bar{R}_1$ again. From the definitions of $\hat{C}$ and T:

$$\hat{C}\, \hat{e}_{nj} = T_m^{-1} CA^{j-1} g_1 \qquad j = 1 \ldots \nu_1 \qquad (4.4.14)$$

Lemma 3.2.4 states that Cg_1, $CAg_1 \ldots CA^{\nu_1 - 2} g_1$ are zero; therefore, (4.4.14) requires that the first $\nu_1 - 1$ columns of $\hat{C}$ be zero. Since $CA^{\nu_1 - 1} g_1$ is the first column of T_m, however, (4.4.14) yields:

$$\hat{C}\, \hat{e}_{n\nu_1} = \hat{e}_{m1} \qquad (4.4.15)$$

and the ν_1^{th} column of $\hat{C}$ is as defined in (4.4.13). Equations (4.4.14) and (4.4.15) can be generalized to $\bar{R}_2 \ldots \bar{R}_r$ to complete the definition of $\hat{C}$.

With $\hat{A}$ and $\hat{C}$ completely specified, attention can now be directed toward an algorithm for selecting $\hat{D}$. The matrix $\hat{D}$ is to be selected as the sum of three terms:

$$\hat{D} = \hat{D}_{FR} + \hat{D}_{\Psi} + \hat{D}_{\Gamma} \tag{4.4.16}$$

The first term is chosen such that $\bar{\mathcal{R}}_1$ and $\bar{\mathcal{R}}_r$ are invariant with respect to $\hat{A}$-$\hat{D}\hat{C}$, i.e. it is chosen to insure that $\hat{D}$ is a detection filter for $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$. The matrices $\hat{D}_{\Psi}$ and $\hat{D}_{\Gamma}$ are then used to assign the eigenvalues of $\hat{A}$-$\hat{D}\hat{C}$ in such a way that this invariant property is not violated. (If the detection filter is being designed to detect m output separable failures, i.e., if r=m, then $\hat{D}_{\Gamma}$ is zero.) With a given set of desired eigenvalues $\hat{D}$ is generally unique, however the component matrices $\hat{D}_{FR}$, $\hat{D}_{\Psi}$ and $\hat{D}_{\Gamma}$ are generally not unique.

An acceptable choice for $\hat{D}_{FR}$ would be $\hat{D}_{\mathcal{J}}$ from (4.2.8), but a preferable choice is:

$$\hat{D}_{FR} = \begin{bmatrix} 0 & \hat{a}_{12} & \cdots & \hat{a}_{1r} & 0 & 0 \\ \hat{a}_{21} & 0 & \cdots & \hat{a}_{2r} & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots & \vdots \\ \hat{a}_{r1} & \hat{a}_{r2} & \cdots & 0 & 0 & 0 \\ \hat{a}_{01} & \hat{a}_{02} & \cdots & \hat{a}_{0r} & 0 & 0 \\ \hat{a}_{\Gamma 1} & \hat{a}_{\Gamma 2} & \cdots & \hat{a}_{\Gamma r} & 0 & 0 \end{bmatrix} \tag{4.4.17}$$

where the vectors $\hat{a}_{12}$ etc. are defined in (4.4.6). The effect of assigning $\hat{D}_{FR}$ using (4.4.17) is that $\hat{A} - \hat{D}_{FR}\hat{C}$ becomes nearly block diagonal:

$$\hat{A} - \hat{D}_{FR}\,\hat{C} = \begin{bmatrix} \hat{A}'_{11} & 0 & \cdots & 0 & \hat{\Theta}_1 & \hat{\Gamma}'_1 \\ 0 & \hat{A}'_{22} & & 0 & \vdots & \vdots \\ \vdots & & \ddots & & \hat{\Theta}_r & \hat{\Gamma}'_r \\ & & & \hat{A}'_{rr} & \Pi & \hat{\Gamma}'_0 \\ 0 & & \cdots & 0 & & \hat{A}'_{\Gamma\Gamma} \end{bmatrix} \qquad (4.4.18)$$

The block diagonality of $\hat{A} - \hat{D}_{FR}\,\hat{C}$ is sufficient to guarantee that $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are invariant with respect to $\hat{A} - \hat{D}_{FR}\,\hat{C}$. As an example, the unit vectors $\hat{e}_{n1} \ldots \hat{e}_{n\nu_1 - 1}$ satisfy an equation similar to (4.4.12) and $(\hat{A} - \hat{D}_{FR}\hat{C})\,e_{n\nu_1}$ is a linear combination of $\hat{e}_{n1} \ldots \hat{e}_{n\nu_1}$:

$$\begin{aligned} \hat{e}_{n(j+1)} &= (\hat{A} - \hat{D}_{FR}\,\hat{C})\,\hat{e}_{nj} \quad j \le \nu_1 - 1 \\ (\hat{A} - \hat{D}_{FR}\,\hat{C})\,\hat{e}_{n\nu_1} &= -\sum_{j=1}^{\nu_1} p'_{1j}\,\hat{e}_{nj} \end{aligned} \qquad (4.4.19)$$

where $p'_{11} \ldots p'_{1\nu_1}$ are the nonzero components of $\hat{A}_{11}$ in (4.4.7). Equation (4.4.19) satisfies the definition given in Section 2.2 for the invariance of $\bar{\mathcal{R}}_1$.

A variation of the above argument can be used to show that the necessary and sufficient requirement for $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ to be invariant with respect to $\hat{A} - \hat{D}\hat{C}$ is that $\hat{A} - \hat{D}\hat{C}$ exhibit the partial block diagonality of (4.4.18). The matrix $\hat{D}_\Psi$ can be chosen to assign the $\nu_J - \nu_0$ eigenvalues (see previous footnote) of $\hat{A} - \hat{D}\,\hat{C}$ associated with $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ without violating this constraint. The general solution for $\hat{D}_\Psi$ is:

$$\hat{D}_\Psi = \begin{bmatrix} d_{\Psi 1} & 0 & \cdots & 0 & 0 \\ 0 & d_{\Psi 2} & \cdots & 0 & 0 \\ \vdots & & & & \\ 0 & \cdots & 0 & d_{\Psi r} & 0 \\ 0 & \cdots & & 0 & 0 \\ 0 & \cdots & & 0 & 0 \end{bmatrix} \qquad (4.4.20)$$

where the ν_1-...ν_r-dimensional vectors $d_{\Psi 1}...d_{\Psi r}$ satisfy:

$$d_{\Psi i} = \begin{bmatrix} -p'_{i1} + p_{i1} \\ -p'_{i2} + p_{i2} \\ \vdots \\ -p'_{i\nu_i} + p_{i\nu_i} \end{bmatrix} \quad i = 1...r \qquad (4.4.21)$$

The coefficients $p'_{11}...p'_{1\nu}$ etc. are the elements of $\hat{A}'_{11}...\hat{A}'_{rr}$ from (4.4.7); the coefficients $p_{11}...p_{1\nu_1}$ are the coefficients of the desired relative minimal polynomials $\Psi_1(\cdot)...\Psi_r(\cdot)$:

$$\begin{aligned} \Psi_i(\lambda) &= \lambda^{\nu_i} + p_{i\nu_i}\lambda^{\nu_i - 1} + ...p_{i2}\lambda + p_{i1} \qquad i = 1...r \\ &= (\lambda - \lambda_{i1})(\lambda - \lambda_{i2})...(\lambda - \lambda_{i\nu_i}) \end{aligned} \qquad (4.4.22)$$

where $\lambda_{i1}...\lambda_{i\nu_1}$ etc. are the desired eigenvalues of $\hat{A} - \hat{D}\hat{C}$. With the aid of (4.4.13) it is simple to verify that:

$$A - (\hat{D}_{FR} + \hat{D}_{\Psi})\hat{C} = \begin{bmatrix} \hat{A}_{11} & 0 & \cdots & 0 & \hat{\Theta}_1 & \hat{\Gamma}'_1 \\ 0 & \hat{A}_{22} & & 0 & \hat{\Theta}_2 & \hat{\Gamma}'_2 \\ \vdots & & \ddots & \vdots & \vdots & \vdots \\ & & & \hat{A}_{rr} & \hat{\Theta}_r & \hat{\Gamma}'_r \\ & & & 0 & \Pi & \hat{\Gamma}'_0 \\ 0 & & \cdots & 0 & 0 & \hat{A}'_{\Gamma\Gamma} \end{bmatrix} \qquad (4.4.23)$$

where the diagonal submatrices are:

$$\hat{A}_{ii} = \begin{bmatrix} 0 & \cdots & 0 & -p_{i1} \\ 1 & \ddots & \vdots & \vdots \\ & \ddots & 0 & -p_{i(\nu_i-1)} \\ 0 & & 1 & -p_{i\nu_i} \end{bmatrix} \qquad (4.4.24)$$

From (4.4.24), it follows that the product $\Psi_1(\cdot)\,\Psi_2(\cdot)\ldots\Psi_r(\cdot)$ is a divisor of the characteristic equation of $\hat{A}-(\hat{D}_{FR}+\hat{D}_{\Psi})\hat{C}$ and, therefore, that each of $\lambda_{i1}\ldots\lambda_{i\nu_1}$ etc. is an eigenvalue of $\hat{A}-(\hat{D}_{FR}+\hat{D}_{\Psi})\hat{C}$ associated with the appropriate detection space.

If $r=m$, the specification of $\hat{D}$ is complete. Otherwise, the final step in the choice of $\hat{D}$ is to choose $\hat{D}_{\Gamma}$ to assign the remaining $n-\nu_{\mathcal{J}}$ assignable eigenvalues of $\hat{A}-\hat{D}\hat{C}$-- the ν_0 eigenvalues of $\hat{A}-\hat{D}\hat{C}$ associated with $\mathcal{R}_{0g}$ are not assignable. The general solution for $\hat{D}_{\Gamma}$ is:

$$\hat{D}_{\Gamma} = \begin{bmatrix} 0 & \hat{D}_{\Gamma 1} \\ 0 & \hat{D}_{\Gamma 2} \\ \vdots & \\ 0 & \hat{D}_{\Gamma r} \\ 0 & \hat{D}_{\Gamma 0} \\ 0 & \hat{D}_{\Gamma m} \end{bmatrix} \qquad (4.4.25)$$

where $\hat{D}_{\Gamma m}$ is a $(n-\nu_{\mathcal{J}})$ x $(m-r)$ submatrix. The final expression for $\hat{A}-\hat{D}\hat{C}$ is then:

$$\hat{A}-\hat{D}\hat{C} = \begin{bmatrix} \hat{A}_{11} & 0 & \hat{\Theta}_1 & \hat{\Gamma}_1 \\ 0 & \hat{A}_{22} & \hat{\Theta}_2 & \hat{\Gamma}_2 \\ & \ddots & \vdots & \vdots \\ & & \hat{\Theta}_r & \hat{\Gamma}_r \\ & & \Pi & \hat{\Gamma}_0 \\ 0 & & 0 & \hat{A}_{\Gamma\Gamma} \end{bmatrix} \qquad (4.4.26)$$

where:

$$\hat{\Gamma}_i = \hat{\Gamma}'_i - \hat{D}_{\Gamma i}\hat{C}_{\Gamma m} \qquad i = 1 \ldots r, 0 \tag{4.4.27}$$

$$\hat{A}_{\Gamma\Gamma} = \hat{A}'_{\Gamma\Gamma} - \hat{D}_{\Gamma m}\hat{C}_{\Gamma m}$$

Because of the structure of $\hat{C}$, the only effect of $\hat{D}_\Gamma$ is to alter the last $n-\nu_{\mathcal{J}}$ columns of $\hat{A} - \hat{D}\hat{C}$. Therefore, its value affects only the $n-\nu_{\mathcal{J}}$ eigenvalues of $\hat{A} - \hat{D}\hat{C}$ not associated with $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_1, \mathcal{R}_0$.

The computational usefulness of base normal form is evidenced by the ease with which $\hat{D}$ in (4.4.16) can be specified. $\hat{D}_{FR}$ can be chosen by inspection using (4.4.17). Once a set of desirable eigenvalues $\lambda_{i1} \ldots \lambda_{i\nu_i}$ etc. has been chosen, all that is required to determine $\hat{D}_\Psi$ in (4.4.20) is to evaluate the polynomial coefficients in (4.4.22). If the detection filter is being designed to detect m output separable failures, i.e. $r = m$, then $\hat{D}_\Gamma$ is zero. Otherwise, $\hat{D}_\Gamma$ is specified by the algorithm in Appendix A.4. In essence, base normal form permits $\hat{D}$ to be chosen by inspection.

There is a second important advantage of base normal form. Because of the partial block diagonality of $\hat{A} - \hat{D}\hat{C}$ and the corresponding partitions of $\hat{C}$, the reference model $(\hat{A} - \hat{D}\hat{C},\ \hat{C})$ can be decoupled into a set of $r + 2$ subsystems of dimension $\nu_1 \ldots \nu_r,\ \nu_0,\ n-\nu_{\mathcal{J}}$ respectively. With $\hat{x}_m(t)$, $\hat{y}_m(t)$, and $\hat{B}$ partitioned similarly to $(\hat{A} - \hat{D}\hat{C})$ and $\hat{C}$ respectively:

$$x_m(t) = \begin{bmatrix} x_{m1}(t) \\ \hline \\ x_{mr}(t) \\ \hline x_{m0}(t) \\ \hline x_{m\Gamma}(t) \end{bmatrix} \tag{4.4.28}$$

$$\hat{y}_m(t) = \begin{bmatrix} y_{m1}(t) \\ \hline \vdots \\ \hline y_{mr}(t) \\ \hline y_{m\Gamma}(t) \end{bmatrix} \qquad (4.4.24)$$

$$\hat{B} = \begin{bmatrix} \hat{B}_1 \\ \hline \vdots \\ \hline \hat{B}_r \\ \hline B_0 \\ \hline B_\Gamma \end{bmatrix} \qquad (4.4.30)$$

then the decoupled reference model satisfies the r equations:

$$\dot{\hat{x}}_{mi}(t) = \hat{A}_{ii}\,\hat{x}_{mi}(t) + \hat{B}_i u(t) + \hat{\Theta}_i\,\hat{x}_{m0}(t) + \hat{\Gamma}_i\,\hat{x}_{m\Gamma}(t)$$

$$y_{mi}(t) = \hat{c}_i\,\hat{x}_{mi}(t) \qquad i = 1 \ldots r \qquad (4.4.31)$$

where $\hat{x}_{m0}(t)$ and $\hat{x}_{m\Gamma}(t)$ are solutions to the equations:

$$\dot{\hat{x}}_{m0}(t) = \Pi\,\hat{x}_{m0}(t) + \hat{B}_0 u(t) + \hat{\Gamma}_0\,\hat{x}_{m\Gamma}(t) \qquad (4.4.32)$$

$$\dot{\hat{x}}_{m\Gamma}(t) = \hat{A}_{\Gamma\Gamma}\,\hat{x}_{m\Gamma}(t) + \hat{B}_\Gamma u(t) \qquad (4.4.33)$$

$$y_{m\Gamma}(t) = \hat{C}_{\Gamma m}\,\hat{x}_{m\Gamma}(t)$$

If the failures associated with $f_1 \ldots f_r$ are mutually detectable, then (4.4.32) disappears. If in addition, r=m, then (4.4.33) disappears and only (4.4.31) remains. The decoupling can result in a significant decrease in the computational requirements for the reference model. In addition, Chapter 5 uses the decoupling property to optimize the state estimation, properties of the reference model.

4.5 Fully Measurable Systems

The simplest failure detection problem is for systems in which the complete state vector is measured. The open loop model for the spinning cylinder given in Section 4.1.2 is an example. Beard referred to such systems as fully measurable systems. They are characterized by the fact that rkC = n. The measurement matrix can therefore be inverted to determine the system state directly. Fully measurable systems are a degenerate example of the general failure detection problem. The majority of the numerical algorithms in the design process can be executed trivially. The considerations pertinent to designing a failure detection system for a fully measurable system are summarized in this section. The algorithm developed in this section utilizes the base normal canonical form developed in Section 4.4.

Fully measurable systems were not analyzed per se in Chapter 3. In Section 3.1 it was suggested that every event vector is its own detection space for fully measurable systems. Because of this, the major results of Chapter 3 can be reduced to two rather simple statements. To give credence to the discussion to follow, these results are state formally, but without proof.

Recall from Section 3.3.1 that the event vectors $f_1 \ldots f_r$ are defined to be output separable if $Cf_1 \ldots Cf_r$ are mutually independent. For fully measurable systems, independence of $f_1 \ldots f_r$ implies output separability. The results of Sections 3.2, 3.3, and 3.5 on mutually detectable failures can be summarized for fully measurable systems by a single theorem.

Theorem 4.1 If the event vectors $f_1 \ldots f_r$ are linearly independent, they are mutually detectable for any fully measurable system (A, B, C).

The theorem states that the failures represented by the event vectors $f_1 \ldots f_r$ can be detected by a single failure detection system if the event vectors are independent. Detectability is interpreted as in Definition 2.1.1. Independence of $f_1 \ldots f_r$ is sufficient to insure

that a D can be chosen such that the error signal associated with each of the event vectors is unidirectional along the corresponding $Cf_1 \ldots Cf_r$ and that the eigenvalues of A-DC can be freely chosen. The difficulty of unassignable eigenvalues considered in Section 3.3 does not exist for fully measurable systems.

To complete the results, it is necessary to consider sets of event vectors which are not mutually independent. This problem corresponds to output stationarity as developed in Section 3.4. Event vectors $h_1 \ldots h_q$ which are linear combinations of $f_1 \ldots f_r$ can be added to $f_1 \ldots f_r$, but it is then necessary to assign some of the eigenvalues of A-DC to have identical values. Output stationarity for fully measurable systems is summarized in Theorem 4.2.

Theorem 4.2 Let $f_1 \ldots f_r$ be mutually detectable for the fully measurable system (A, B, C) and let $\lambda_1 \ldots \lambda_r$ be the eigenvalues of A-DC associated with $f_1 \ldots f_r$ respectively. Assume D is chosen such that the error signal associated with each of $f_1 \ldots f_r$ is unidirectional for (A-DC, C) and a subset $\lambda_1 \ldots \lambda_k$ of the r eigenvalues are equal. Any failure with an event vector h_1 which is a linear combination of $f_1 \ldots f_k$ produces unidirectional output along Ch_1.

This is a restatement of a familiar matrix algebra property. D is chosen such that $f_1 \ldots f_r$ are eigenvectors of A-DC. Therefore:

$$(A\text{-}DC)\, f_i = \lambda_i f_i \qquad i = 1 \ldots r \qquad (4.5.1)$$

and the arguments of Section 2.1 show that the failures associated with $f_1 \ldots f_r$ produce unidirectional output. If $\lambda_1 \ldots \lambda_k$ are equal, any linear combination h_1 of $f_1 \ldots f_k$ is also an eigenvector of A-DC and the failure associated with it must generate unidirectional output along Ch_1. If all n eigenvalues are chosen to be equal, every vector is an eigenvector of A-DC and every failure which can be described by a single event vector produces a unidirectional output error. Subject to the comments in Section 4.1, however, the output error produced by sensor failures can only be constrained to a plane.

With these two theorems, all that is needed is a design algorithm for choosing D. A straightforward algorithm is presented below. Each step is illustrated by the open loop model for the spinning cylinder as given in (4.1.27).

Step 1 Choose a largest set of linearly independent event vectors corresponding to failures of interest.

Obviously it is advantageous to choose as large a set of failures as possible Assume the failures of interest in the example are failures in:

i) the attitude sensor
ii) the rate sensor
iii) the torquing device
iv) the torque measuring device

It was shown in Section 4.1.2 that the first two failures can be modeled adequately by the sets of event vectors $\{e_{31}, 0\}$ and $\{e_{32}, e_{31}\}$ respectively. Since a failure in the torquing device can be modeled as changes in the third row of A_o and B_o, that failure is associated with the event vector e_{33}. In a manner similar to the analysis of the rate sensor failure, the torque measuring device can be described by the set of event vectors $\{e_{33}, e_{32}\}$.

After this analysis, the event vectors for the four failures can be summarized by the sets:

i) $\{e_{31}, 0\}$
ii) $\{e_{32}, e_{31}\}$
iii) e_{33}
iv) $\{e_{33}, e_{32}\}$

From these entries, it is clear that if the set of event vectors $\{e_{31}, e_{32}, e_{33}\}$ is chosen for Step 1, then all four failures can be detected by a single detection system. Each failure has a unique set of directions in which the output error signal generated by that failure appears and no ambiguity is involved.

Step 2 If the set of event vectors is $f_1 \ldots f_r$, define the coordinate transformations to base normal canonical form:

$$T = [f_1 \ldots f_r \; w_{r+1} \ldots w_n] \tag{4.5.2}$$

and:

$$T_m = [Cf_1 \ldots Cf_r \; Cw_{r+1} \ldots Cw_n]$$

where $w_{r+1} \ldots w_n$ are chosen to make T invertible and define the transformed system:

$$\begin{aligned} \hat{A} &= T^{-1}AT \\ \hat{B} &= T^{-1}B \\ \hat{C} &= T_m{}^{-1}CT \\ \hat{D} &= T^{-1}DT_m \end{aligned} \tag{4.5.3}$$

The system $(\hat{A}, \hat{B}, \hat{C})$ is input-output equivalent to $(A, B, T_m{}^{-1}C)$. The salient properties of the base normal canonical form are that the event vectors $\hat{f}_1 \ldots \hat{f}_r$ are the new unit vectors $\hat{e}_{n1} \ldots \hat{e}_{nr}$ and that, since $\hat{C}$ is diagonal, $\hat{C}\hat{f}_1 \ldots \hat{C}f_r$ equal $\hat{e}_{n1} \ldots \hat{e}_{nr}$ respectively. In this canonical form the choice of $\hat{D}$ is trivial. D can then be found from the inverse transformation.

For the general application, the transformation places the system model in a coordinate system in which the elements of the state vector are linear combinations of the original elements and may not have apparent physical meaning. For the cylinder example, T and T_m are the identity matrix and $(\hat{A}, \hat{B}, \hat{C})$ is the same as (A, B, C).

Step 3 $\hat{D}$ is to be chosen as $\hat{D}_{FR} + \hat{D}_\psi$ where $\hat{D}_{FR}$ makes the output generated by $\hat{f}_1 \ldots \hat{f}_r$ unidirectional for $\hat{A}\text{-}\hat{D}_{FR}\hat{C}$ and $\hat{D}_\psi$ assigns the eigenvalues of $\hat{A}\text{-}\hat{D}\hat{C}$ without interferring with the unidirectionality. $\hat{D}_{FR}$ is chosen such that $\hat{f}_1 \ldots \hat{f}_r$ are eigenvectors of $\hat{A}\text{-}\hat{D}_{FR}\hat{C}$. Since $\hat{f}_1 \ldots \hat{f}_r$ are the unit vectors $\hat{e}_{n1} \ldots \hat{e}_{nr}$, this requires:

$$(\hat{A}-\hat{D}_{FR}\hat{C})\,\hat{e}_{ni} = \mu_i\,\hat{e}_{ni} \qquad i = 1\ldots r \qquad (4.5.4)$$

where $\hat{C}$ is the identity matrix. A satisfactory choice for $\hat{D}_{FR}$ is the off-diagonal elements of $\hat{A}$. This choice allows all n eigenvalues of $\hat{A}-\hat{D}\hat{C}$ to be assigned in one step.

For the example, $\hat{D}_{FR}$ can be written from (4.1.27):

$$\hat{D}_{FR} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & \alpha_r \\ 0 & 0 & 0 \end{bmatrix} \qquad (4.5.5)$$

and $\hat{A}-\hat{D}_{FR}\hat{C}$ is diagonal as it should be:

$$\hat{A}-\hat{D}_{FR}\hat{C} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -\beta_r & 0 \\ 0 & 0 & -\beta_f \end{bmatrix} \qquad (4.5.6)$$

Step 4 Let the diagonal elements of $\hat{A}-\hat{D}_{FR}\hat{C}$ be $\mu_1\ldots\mu_n$. If the desired eigenvalues for $\hat{A}-\hat{D}\hat{C}$ are $\lambda_1\ldots\lambda_n$, where the first r are to be associated with $\hat{f}_1\ldots\hat{f}_r$ respectively, then $\hat{D}_\psi$ must be:

$$\hat{D}_\psi = \begin{bmatrix} \mu_1-\lambda_1 & & & 0 \\ & 0 & & \\ & & & 0 \\ 0 & & & \mu_n-\lambda_n \end{bmatrix} \qquad (4.5.7)$$

Since $\hat{A}-\hat{D}_{FR}\hat{C}$ is diagonal, this choice of $\hat{D}_\psi$ yields:

$$\hat{A}-(\hat{D}_{FR}+\hat{D}_\psi)\,\hat{C} = \hat{A}-\hat{D}\hat{C} = \begin{bmatrix} \lambda_1 & & & 0 \\ & 0 & & \\ & & & 0 \\ 0 & & & \lambda_n \end{bmatrix} \qquad (4.5.8)$$

and the eigenvalues of $\hat{A}-\hat{D}\hat{C}$ have been assigned arbitrarily. The detection filter D in the original system coordinates is:

$$D = T(\hat{D}_{FR} + \hat{D}_{\psi})\ T_m^{-1} \tag{4.5.9}$$

For the example, $\hat{D}_{\psi}$ is:

$$\hat{D}_{\psi} = \begin{bmatrix} -\lambda_1 & 0 & 0 \\ 0 & -\beta_r-\lambda_2 & 0 \\ 0 & 0 & -\beta_f-\lambda_3 \end{bmatrix} \tag{4.5.10}$$

Step 5 Assume there are additional event vectors $h_1 \ldots h_q$ which are linear combinations of $f_1 \ldots f_r$. If I is the smallest index set such that $h_1 \ldots h_q$ are a linear combination of f_i, $i \in I$, then each of $\hat{f}_1 \ldots \hat{f}_r$, $\hat{h}_1 \ldots \hat{h}_q$ produces unidirectional output along $\hat{C}\hat{f}_1 \ldots \hat{C}\hat{f}_r$, $\hat{C}\hat{h}_1 \ldots \hat{C}\hat{h}_q$ respectively if $\hat{D}_{\psi}$ of Step 4 is chosen such that the λ_i, $i \in I$, are equal.

This step constrains the choice of $\hat{D}_{\psi}$ to insure that $\hat{h}_1 \ldots \hat{h}_q$ are output stationary with $\hat{f}_1 \ldots \hat{f}_r$, i.e., that $\hat{h}_1 \ldots \hat{h}_q$ are also eigenvectors of $\hat{A}-\hat{D}\hat{C}$. The procedure is a direct application of Theorem 4.2. This completes the design algorithm.

4.6 Extensions of Failure Detection Theory

The formulation of failure detection theory in Chapter 3 was in an algebraic context. The motivation was a desire to detect failures in continuous, linear, time invariant systems; but the results are applicable to a broader range of problems. In particular, they are directly applicable to discrete systems. They can also be used effectively to design a failure detection system for sampled data systems. Attention is restricted in this section to deterministic systems. Stochastic systems and sampled data systems are considered in Chapter 5.

The dynamics of a discrete system can be expressed by the sequence:

$$x(k+1) = A_\Delta x(k) + B_\Delta u(k) \quad : x(0) \text{ given}$$
$$y(k+1) = C\, x(k+1) \tag{4.6.1}$$

The important feature of discrete systems is that the state is defined only at the discrete times k, k+1... Similarly to continuous systems, failures in discrete systems can be modeled as either a change in dynamics ($A_\Delta + \Delta A_\Delta$, $B_\Delta + \Delta B_\Delta$, $C + \Delta C$) or an arbitrary additive sequence $f_i n_i(k)$ in (4.6.1). If the reference model is based upon perfect knowledge of the state transition dynamics, then:

$$x_m(k+1) = A_\Delta x_m(k) + B_\Delta u(k) + D_\Delta (y(k) - y_m(k))$$
$$y_m(k+1) = C\, x_m(k+1) \qquad : x_m(0) \text{ given} \tag{4.6.2}$$

With the errors defined in the normal manner, a state equation for $\epsilon(k)$ can be found for the nominal system:

$$\epsilon(k+1) = (A_\Delta - D_\Delta C)\, \epsilon(k) \quad : \epsilon(0) \text{ given}$$
$$\epsilon'(k+1) = C\, \epsilon(k+1) \tag{4.6.3}$$

It is simple to show that the estimation errors produced by failures can be modeled similarly to (2.1.6):

$$\epsilon(k+1) = (A_\Delta - D_\Delta C)\, \epsilon(k) + f_i\, n_i(k)$$
$$\epsilon'(k+1) = C\epsilon(k+1) + e_{mi}\, n_{ci}(k+1) \tag{4.6.4}$$

The important similarity between continuous systems and discrete systems is that failures in either can be characterized by event vectors. The event vectors are determined by the associated failure, but they are unique and independent of the particular failure mode. Equations (4.6.1) - (4.6.4) are of the same form as (2.1.1) - (2.1.6). Because of this similarity, the results of Chapter 3 can be used directly for discrete systems.

The distinction between discrete systems and continuous systems with discrete measurements (sampled data systems) is that in the

latter case the state is defined as a continuous function of time and a failure must be also. Because of the discrete measurements, however, only the estimation errors at the measurement times can be defined meaningfully. For failure detection purposes, this requires that sampled data systems be modeled in discrete form similar to (4.6.1) - (4.6.4). Not surprisingly, the deficiency of measurements means that the analytic results of Chapter 3 are only approximate for sampled data systems.

The sampled data equivalent of (4.6.1) involves the state transition matrix $\Phi(t, \tau)$ for A. As an example of the difficulties introduced by discrete measurements, the state equation is written in the case of a failure in the i^{th} control element:

$$x(t+\Delta t) = \Phi(t+\Delta t, t)\, x(t) + \int_t^{t+\Delta t} \Phi(t+\Delta t, \tau)\, Bu(\tau)\, d\tau + \int_t^{t+\Delta t} \Phi(t+\Delta t, \tau)\, \Delta b_i u_i(\tau)\, d\tau$$

$$y(t+\Delta t) = Cx(t+\Delta t) \qquad (4.6.5)$$

The reference model is again defined for the nominal system:

$$x_m(t+\Delta t) = \Phi(t+\Delta t, t)\, x_m(t) + \int_t^{t+\Delta t} \Phi(t+\Delta t, \tau)\, Bu(\tau)\, d\tau + D_\Delta(y(t) - y_m(t))$$

$$y_m(t+\Delta t) = Cx_m(t+\Delta t) \qquad (4.6.6)$$

The error equations in this instance are driven by the failure:

$$\epsilon(t+\Delta t) = (\Phi(t+\Delta t, t) - D_\Delta C)\epsilon(t) + \int_t^{t+\Delta t} \Phi(t+\Delta t, \tau)\, \Delta b_i\, u_i(\tau)\, d\tau$$

$$\epsilon'(t+\Delta t) = C\epsilon(t+\Delta t) \qquad (4.6.7)$$

The reference model in (4.6.6) is not unique; $D_\Delta(y(t) - y_m(t))$ could be replaced by $D_\Delta(y(t+\Delta t) - y_m(t))$. Both representations converge to (2.1.2) as Δt approaches zero. Equation (4.6.6) is used here because of the parallel with (4.6.2). The second representation is used in Section 5.2 for stochastic systems.

The difference between discrete systems and sampled data systems is evident in (4.6.7). It is possible to define an event vector $f_i'(t+\Delta t)$ and a function $n_i(t+\Delta t)$ such that:

$$\int_t^{t+\Delta t} \Phi(t+\Delta t, \tau)\, \Delta b_i\, u_i(\tau)\, d\tau = f_i'(t)\, n_i(t) \qquad (4.6.8)$$

The difficulty is that the event vector is a function of $u_i(\tau)$ in the interval $t < \tau < t+\Delta t$. If $u_i(\tau)$ is not constant within every interval, the event vector $f_i'(t)$ changes from one interval to the next. As indicated in (2.1.19), $f_i'(t)$ is some linear combination of Δb_i, $A\Delta b_i$, $A^2\Delta b_i \ldots$. The sequence $f_i'(t)$, $f_i'(t+\Delta t)$, $f_i'(t+2\Delta t)\ldots$ are in general different linear combinations of those vectors. Since Δb_i, $A\Delta b_i$, $A^2\Delta b_i \ldots$ are not necessarily in the same detection space, it is not likely that the event vectors are either. If they are not, it is not possible to define a detection filter which constrains the output error generated by a failure in the i^{th} controller to a single direction.

One possible solution for the sampled data problem is to constrain the control vector to be piece-wise constant within each measurement interval. (If the same measurements are used for failure detection and for choosing the control, this is always true). If the magnitude of Δb_i is not a function of time, $f_i'(t)$ as defined in (4.6.8) is then a constant. If the controller failure is control-independent (Table 4.1), the time function $n_i(t)$ must also be piece-wise constant. With these constraints, Chapter 3 can be used to define a detection filter for controller and sensor failures. For changes in the state matrix A, however, the equivalent of (4.6.8) involves state vector components and unique event vectors cannot be found. It follows that Chapter 3 cannot be used directly for such failures.

For small time increments, $f_i'(t)$, $f_i'(t+\Delta t)\ldots$ lie approximately along Δb_i. This suggests that a first order approximation of (4.6.5) - (4.6.7) by a discrete system would place the failure detection problem for sampled data systems in tractable form. With the approximations:

$$\Phi(t+\Delta t,\ t) \approx E_n + \Delta t\, A$$

$$D_\Delta \approx \Delta t\, D \tag{4.6.9}$$

(4.6.5) - (4.6.7) can be written as the first order equations:

$$x(t+\Delta t) = (E_n + \Delta t\, A)\, x(t) + \Delta t\, Bu(t) + \Delta t\, \Delta b_i\, u_i(t)$$

$$y(t+\Delta t) = Cx(t+\Delta t) \tag{4.6.10}$$

$$x_m(t+\Delta t) = (E_n + \Delta t\, A)\, x_m(t) + \Delta t\, Bu(t) + \Delta t\, D(y(t) - y_m(t))$$

$$y_m(t+\Delta t) = C\, x_m(t+\Delta t) \tag{4.6.11}$$

$$\epsilon(t+\Delta t) = (E_n + \Delta t\, A - \Delta t\, DC)\, \epsilon(t) + \Delta t\, \Delta b_i\, u_i(t)$$

$$\epsilon'(t+\Delta t) = C\, \epsilon(t+\Delta t) \tag{4.6.12}$$

As desired, the event vectors for (4.6.12) are constant and independent of the failure mode. The results of Chapter 3 can be applied directly to the discrete system defined by (4.6.10) - (4.6.12), but with A, B, C, and D replaced by $E_n + \Delta t\, A$, $\Delta t B$, C, and $\Delta t D$ respectively. That detection filter can then be used with the sampled-data reference model (4.6.6). If (4.6.10) - (4.6.12) are used to determine a detection filter, the degree to which the filter satisfies the detection constraints is a function of the validity of the first order approximations. If it is sufficiently accurate, the detected output error signals will be approximately unidirectional. Otherwise, a failure detection filter based upon Chapter 3 cannot be found. The possible effectiveness of the linearization is illustrated in Chapter 6.

CHAPTER 5

DETECTION FILTER DESIGN -- STOCHASTIC

Two important issues relative to failure detection systems were omitted in Chapter 4. The first concerns criteria for the selection of eigenvalues for the error state matrix A-DC. The second is the problem of differentiating between the output error produced by a failure and that produced by measurement and dynamic uncertainties. Unlike the analytical results of Chapter 3, there is not a single best answer to either problem. The intent of this chapter is to present reasonable solutions to both problems and discuss their advantages and disadvantages viz-a-viz other approaches. The emphasis is on solutions which are applicable to a broad range of design applications. All of the discussions in this chapter rely upon the base normal form of Section 4.4.

The reference model approach to failure detection defines a state estimator. If a particular application requires both state estimation and failure detection, a useful question is how good a state estimator a failure detection system can be. If its accuracy is acceptable, it can be used to perform the combined functions of failure detection and state estimation. It is shown in Sections 5.1 and 5.2 that, with certain restrictions, the freedom to assign the eigenvalues of A-DC can be used to optimize the detection filter as a state estimator. Because of the detection constraints, the optimal detection filter-state estimator does not provide the minimum variance estimate produced by a Kalman filter, but it does enjoy certain advantages over a Kalman filter. The relative merits of the two estimators are discussed qualitatively in this chapter and illustrated in Chapter 6.

If a state estimation capability is not needed, it is still reasonable to suppose that an optimal detection filter-state estimator would be a good detection filter for many applications. If the output estimation errors (residuals) are as small as possible in the absence of a failure, the output errors due to a failure should be easy to detect. It is

possible to define a simple threshold crossing detection law to discriminate between output estimation errors due to stochastic sources and errors due to a combination of stochastic sources and a failure. The implementation of such a law is discussed in Section 5.3. The possibility of using other criteria is discussed in Section 5.4.

The optimal detection filter-state estimator assumes no apriori knowledge of failure mode. If specific failure mode information is available, there is a trade-off between minimizing the output errors due to stochastic sources and enhancing those due to failures. The detection problem is placed in a general framework in Section 5.4.

5.1 Suboptimal State Estimation -- Steady State Statistics

If D is a detection filter for the failures associated with $f_1 \ldots f_r$, the base normal canonical form of Section 4.4 allows (A-DC, C) to be expressed as a set of r + 1 subsystems, r of which are interrelated only in that they may be driven by the same or similar inputs. It is shown in this section that these r subsystems are decoupled for state estimation purposes. If a large state estimation error arises in one subsystem, that error does not degrade the state estimates associated with any other subsystems. If a system failure occurs, the accuracy of the state estimator is decreased only for the subsystem associated with that failure. The other subsystems of the state estimator remain completely functional.

The isolation of the effect of system failures on the state estimates is the principal difference between a failure detection system and a minimum variance state estimator (Kalman filter). A Kalman filter optimizes the state estimates for a system by making maximum use of the dynamic interdependence of the state components of the system. If a large error occurs in the estimate of one state component, that error ultimately degrades the accuracy of the remaining state component estimates. If a system failure occurs, the propagation of the ensuing state estimation error may make the Kalman filter completely useless.

For state estimation purposes, the choice between a Kalman filter and an optimal detection filter-state estimator is a trade-off between three considerations:

i) the greater accuracy of a Kalman filter under nominal operating conditions,

ii) the ability of an optimal detection filter-state estimator to provide partial state estimation for a failed system, and

iii) the computational efficiency of the optimal detection filter-state estimator.

This last property is due to the decoupling of the reference model into $r + 1$ lower dimensional subsystems. The autopilot example of Chapter 6 is used to illustrate the difference in the behavior of the two types of state estimators.

The failure detection constraints of Chapter 3 allow the n eigenvalues of A-DC to be freely assigned by the choice of D. If $r = m$, base normal form translates this freedom into a set of n freely assignable components of the transformed matrix $\hat{D}$. If $f_1 \ldots f_m$ are mutually detectable, Section 5.1.2 uses the Kalman approach to optimize the choice of the n components. If $f_1 \ldots f_m$ are not mutually detectable, it is necessary to apply the state space augmentation technique discussed in Sections 3.3.4 and 4.2.3 to obtain an $\tilde{n}$-dimensional reference model then optimize the resulting $\tilde{n}$ assignable components of the new $\tilde{\hat{D}}$. The Kalman approach does not yield an optimal detection filter-state estimator if $r < m$, but a reasonable approximation can be made in order to obtain a suboptimal $\hat{D}$. In general this requires further state space augmentation.

When output stationarity is considered, i.e. when $r > m$, fewer than n components of $\hat{D}$ can be assigned freely and there is not sufficient freedom remaining to optimize $\hat{D}$ as a state estimator. The remaining freedom can be used to minimize a scalar function of the residuals, however. The utility of the parameter optimization techniques of Section 2.2 in this regard is illustrated in Section 5.1.3. The parameter optimization technique can also be used in lieu of the

Kalman approach when the latter requires an unacceptable increase in the dimension of the reference model. If $r = m$ and $f_1 \ldots f_m$ are mutually detectable, both approaches yield the same $\hat{D}$.

Only stochastic systems with continuous measurements and stationary statistics are considered in this section. Nonstationary statistics and sampled data systems are considered in Section 5.2. The results are all developed from the base normal canonical form.

5.1.1 Stochastic Error Model

The standard linear disturbance model is introduced in Section 2.2. The system is driven by a white, Gaussian system noise $\eta_1(t)$ and the measurements are corrupted by a white Gaussian measurement noise $\eta_2(t)$. The random processes $\eta_1(t)$ and $\eta_2(t)$ are continuous with zero means:

$$
\begin{aligned}
&E[\eta_k(t)] = 0 \qquad k = 1,2 \\
&E[\eta_1(t)\,\eta_1(\tau)^T] = E_{n_1}\,\delta(t-\tau) \\
&E[\eta_2(t)\,\eta_2(\tau)^T] = E_{m_1}\,\delta(t-\tau) \\
&E[\eta_1(t)\,\eta_2(\tau)^T] = G_{12}\,\delta(t-\tau)
\end{aligned}
\qquad (2.2.2)
$$

In many applications $\eta_1(t)$ and $\eta_2(t)$ are independent, but the general treatment presented here assumes that they are not.

The nominal dynamics for the stochastic system are still defined by (A, C), but the differential equations must include the disturbances:

$$
\begin{aligned}
\dot{x}(t) &= A\,x(t) + B\,u(t) + G_1\,\eta_1(t) \;:\; x(0) \text{ given} \\
y(t) &= C\,x(t) + G_2\,\eta_2(t)
\end{aligned}
\qquad (2.2.1)
$$

The failure detection system is affected by the disturbances only through the output feedback of y(t):

$$
\begin{aligned}
\dot{x}_m(t) &= (A-DC)\,x_m(t) + Bu(t) + Dy(t) \\
y_m(t) &= C\,x_m(t)
\end{aligned}
\qquad (2.1.3)
$$

Since the error sources are not included in the reference model, they appear as driving terms for the error equations:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + G_1\,\eta_1(t) - D\,G_2\,\eta_2(t):\ \epsilon(0) \text{ given}$$

$$\epsilon'(t) = C\,\epsilon(t) + G_2\,\eta_2(t) \qquad (2.2.3)$$

In the event that there is a failure which can be modeled using the controller failure model (2.1.7), an additional driving term appears in the error equations:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + G_1\,\eta_1(t) - D\,G_2\,\eta_2(t) + f_i\,n(t):\ \epsilon(0) \text{ given}$$

$$\epsilon'(t) = C\,\epsilon(t) + G_2\,\eta_2(t) \qquad (5.1.1)$$

If the sensor failure model must be used, there is a failure - related term in each of the error equations:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + G_1\,\eta_1(t) - DG_2\,\eta_2(t) - d_i\,n_c(t)$$

$$\epsilon'(t) = C\,\epsilon(t) + G_2\,\eta_2(t) + e_{mi}\,n_c(t) \qquad (5.1.2)$$

It is assumed in this section that the techniques of Section 4.2.1 have been used to determine equivalent event vectors for sensor failures where necessary and that (5.1.1) is the general stochastic error model.

If the event vectors $f_1 \ldots f_r$ are output separable, the state estimator equations and the error equations can be expressed in base normal form. This requires the transformations T and T_m from (4.4.1) and (4.4.2) respectively. If D is a detection filter for $f_1 \ldots f_r$, then $\hat{A} - \hat{D}\hat{C}$ is nearly block diagonal:

$$\hat{A} - \hat{D}\hat{C} = \begin{bmatrix} \hat{A}_{11} & 0 & \hat{\Theta}_1 & \hat{\Gamma}_1 \\ 0 & \hat{A}_{22} & \hat{\Theta}_2 & \hat{\Gamma}_2 \\ & & \hat{\Theta}_r & \hat{\Gamma}_r \\ & & \Pi & \hat{\Gamma}_0 \\ 0 & & 0 & \hat{A}_{\Gamma\Gamma} \end{bmatrix} \qquad (4.4.26)$$

With $\hat{\epsilon}(t)$, $\hat{\epsilon}'(t)$, $\hat{G}_1$, and $\hat{G}_2$ defined:

$$\hat{\epsilon}(t) = T^{-1}\epsilon(t) = \begin{bmatrix} \hat{\epsilon}_1(t) \\ \vdots \\ \hat{\epsilon}_r(t) \\ \hat{\epsilon}_0(t) \\ \hat{\epsilon}_\Gamma(t) \end{bmatrix} \qquad (5.1.3)$$

$$\hat{\epsilon}'(t) = T^{-1}\epsilon'(t) = \begin{bmatrix} \hat{\epsilon}_1'(t) \\ \vdots \\ \hat{\epsilon}_r'(t) \\ \hat{\epsilon}_\Gamma'(t) \end{bmatrix}$$

$$\hat{G}_1 = T_m^{-1} G_1$$

$$\hat{G}_2 = T_m^{-1} G_2$$

the error equations in the absence of a failure become:

$$\dot{\hat{\epsilon}}(t) = (\hat{A} - \hat{D}\hat{C})\,\hat{\epsilon}(t) + \hat{G}_1\,\eta_1(t) - \hat{D}\,\hat{G}_2\,\eta_2(t) \quad : \epsilon(0) \text{ given}$$

$$\hat{\epsilon}'(t) = \hat{C}\,\hat{\epsilon}(t) + \hat{G}_2\,\eta_2(t) \qquad (5.1.4)$$

The base normal canonical form defines $\hat{D}$ as the sum of three matrices:

$$\hat{D} = \hat{D}_{FR} + \hat{D}_\Psi + \hat{D}_\Gamma \qquad (4.4.16)$$

The non-zero elements of $\hat{D}_\Psi$ and $\hat{D}_\Gamma$ can be freely chosen to optimize $\hat{D}$ as a detection filter-state estimator; $\hat{D}_{FR}$ is completely specified. In order to emphasize the fixed nature of $\hat{D}_{FR}$, it is convenient to define a new system disturbance model for (5.1.4):

$$\eta_1'(t) = \hat{G}_1\eta_1(t) - \hat{D}_{FR}\hat{G}_2\eta_2(t) = \begin{bmatrix} \hat{\eta}_{1_1}'(t) \\ \hat{\eta}_{1_2}'(t) \\ \vdots \\ \hat{\eta}_{1_r}'(t) \\ \hat{\eta}_{1_0}'(t) \\ \hat{\eta}_{1_\Gamma}'(t) \end{bmatrix} \qquad (5.1.5)$$

The random process $\eta_1'(t)$ is not actually a system disturbance, but if (4.4.16) is substituted into (5.1.4), $\eta_1'(t)$ does appear explicitly in the $\hat{\epsilon}(t)$ equation in precisely the manner that a system disturbance would. The covariance of $\eta_1'(t)$ is:

$$E[\eta_1'(t)\,\eta_1'(\tau)^T] = (\hat{G}_1 - \hat{D}_{FR}\,\hat{G}_2)\,(\hat{G}_1 - \hat{D}_{FR}\,\hat{G}_2)^T\,\delta(t-\tau)$$

$$= \hat{G}_1'\,\delta(t-\tau) = \delta(t-\tau)\begin{bmatrix} \hat{G}'_{11} & \hat{G}'_{112} & \cdots & \hat{G}'_{11r} & \hat{G}'_{110} & \hat{G}'_{11\Gamma} \\ \hat{G}'_{121} & \hat{G}'_{22} & & & & \\ \vdots & & & \vdots & \vdots & \vdots \\ \hat{G}'_{1r1} & & & \hat{G}'_{rr} & \hat{G}'_{1r0} & \hat{G}'_{1r\Gamma} \\ \hat{G}'_{101} & & & \hat{G}'_{10r} & \hat{G}'_{00} & \hat{G}'_{10\Gamma} \\ \hat{G}'_{1\Gamma1} & & & \hat{G}'_{1\Gamma r} & \hat{G}'_{1\Gamma0} & \hat{G}'_{\Gamma\Gamma} \end{bmatrix} \qquad (5.1.6)$$

The matrix $\hat{G}_1'$ is n x m and the dimensions of the partitions are identical to those of $\hat{A} - \hat{D}\hat{C}$ in (4.4.26). The off-diagonal partitions of $\hat{G}_1'$ are in general non-zero, but they are of little importance to failure detection theory. Even if $\eta_1(t)$ and $\eta_2(t)$ are not correlated, it is clear that $\eta_1'(t)$ and $\eta_2(t)$ are correlated:

$$E[\eta_1'(t)\eta_2(\tau)^T\hat{C}_2^T] = (\hat{G}_1 G_{12}\hat{G}_2^T - \hat{D}_{FR}\hat{G}_2\hat{G}_2^T)\delta(t-\tau) = \delta(t-\tau)\begin{bmatrix} \hat{G}'_{12:1} & \hat{G}'_{1212} & \cdots & \hat{G}'_{121r} & \hat{G}'_{121\Gamma} \\ \hat{G}'_{122} & \hat{G}'_{1222} & & & \\ \vdots & & & \vdots & \vdots \\ G'_{12r1} & & & \hat{G}'_{12rr} & \hat{G}'_{12r\Gamma} \\ G'_{1201} & & & \hat{G}'_{120r} & \hat{G}'_{120\Gamma} \\ G'_{12\Gamma1} & & & \hat{G}'_{12\Gamma r} & \hat{G}'_{12\Gamma\Gamma} \end{bmatrix}$$

(5.1.7)

The importance of defining $\eta'(t)$ is that it allows (5.1.4) to be expressed in the standard form of the estimation error for a set of lower dimensional state estimators. The first step is to substitute (4.4.16) and (5.1.5) into (5.1.4):

$$\dot{\hat{\epsilon}}(t) = (A - \hat{D}_{FR}\hat{C} - \hat{D}_\Psi\hat{C} - \hat{D}_\Gamma\hat{C})\,\hat{\epsilon}(t) + \eta_1'(t) - (\hat{D}_\Psi + \hat{D}_\Gamma)\,\hat{G}_2\eta_2(t): \ \epsilon(0) \text{ given}$$

$$\hat{\epsilon}'(t) = \hat{C}\,\hat{\epsilon}(t) + \hat{G}_2\,\eta_2(t) \qquad (5.1.8)$$

The matrix $\hat{A} - \hat{D}_{FR}\,\hat{C}$ has the same form as $\hat{A} - \hat{D}\,\hat{C}$:

$$\hat{A} - \hat{D}_{FR}\,\hat{C} = \begin{bmatrix} \hat{A}'_{11} & 0 & \cdots & 0 & \hat{\Theta}_1 & \hat{\Gamma}'_1 \\ 0 & \hat{A}'_{22} & & 0 & & \\ & & \ddots & & & \\ & & & \hat{A}'_{rr} & \hat{\Theta}_r & \hat{\Gamma}' \\ & & & & \Pi & \hat{\Gamma}'_0 \\ 0 & & \cdots & & 0 & \hat{A}'_{\Gamma\Gamma} \end{bmatrix} \qquad (4.4.18)$$

With the partitions of $\hat{D}_\Psi$, $\hat{C}$, and $\hat{D}_\Gamma$ as defined in Section 4.4:

$$\hat{D}_\Psi = \begin{bmatrix} d_{\Psi 1} & 0 & \cdots & 0 & 0 \\ 0 & d_{\Psi 2} & 0 & \cdots & 0 & 0 \\ & & & & & \\ 0 & \cdots & 0 & d_{\Psi r} & 0 \\ & & & & & \\ 0 & \cdots & & 0 & 0 \\ & & & & & \\ 0 & \cdots & & 0 & 0 \end{bmatrix} \qquad (4.4.20)$$

$$\hat{D}_\Gamma = \begin{bmatrix} 0 & \cdots & 0 & \hat{D}_{\Gamma 1} \\ 0 & \cdots & 0 & \hat{D}_{\Gamma 2} \\ & & & \\ 0 & \cdots & 0 & \hat{D}_{\Gamma r} \\ 0 & \cdots & 0 & \hat{D}_{\Gamma 0} \\ 0 & \cdots & 0 & \hat{D}_{\Gamma m} \end{bmatrix} \qquad (4.4.25)$$

$$\hat{C} = \left[\begin{array}{ccccc} \hat{c}_1 & 0 & \cdots & 0 & \\ 0 & \hat{c}_2 & 0 & \cdots & 0 \\ & & \ddots & & \\ 0 & \cdots & 0 & \hat{c}_r & 0 \\ 0 & \cdots & 0 & & \hat{C}_{\Gamma m} \end{array}\right] \qquad (4.4.13)$$

the term $\hat{D}_\Psi \hat{C} + \hat{D}_\Gamma \hat{C}$ in (5.1.8) is:

$$\hat{D}_\Psi \hat{C} + \hat{D}_\Gamma \hat{C} = \left[\begin{array}{ccccc} d_{\Psi 1}\hat{c}_1 & \cdots & 0 & & \hat{D}_{\Gamma 1}\hat{C}_{\Gamma m} \\ 0 & d_{\Psi 2}\hat{c}_2 & 0 & & \hat{D}_{\Gamma 2}\hat{C}_{\Gamma m} \\ & & \ddots & & \\ 0 & \cdots & d_{\Psi r}\hat{c}_r & & \hat{D}_{\Gamma r}\hat{C}_{\Gamma m} \\ 0 & \cdots & 0 & & \hat{D}_{\Gamma 0}\hat{C}_{\Gamma m} \\ 0 & \cdots & 0 & & \hat{D}_{\Gamma m}\hat{C}_{\Gamma m} \end{array}\right] \qquad (5.1.9)$$

The partitions in (5.1.9) have the same dimension as those of $\hat{A}-\hat{D}\hat{C}$.

With the definitions in the previous paragraph, and with the partitions of $\hat{\epsilon}(t)$ and $\hat{\epsilon}'(t)$ given in (5.1.3), (5.1.8) can be stated as a set of r + 2 partially decoupled subsystems. The first r subsystems are those associated with $f_1 \ldots f_r$:

$$\dot{\hat{\epsilon}}_i(t) = (\hat{A}'_{ii} - d_{\Psi i}\,\hat{c}_i)\,\hat{\epsilon}_i(t) + \eta'_{1i}(t) - d_{\Psi i}\hat{G}_{2i}\eta_2(t) + \hat{\Theta}_i\hat{\epsilon}_0(t) + \hat{\Gamma}_i\,\hat{\epsilon}_\Gamma(t) - \hat{D}_{\Gamma i}\,\hat{G}_{2\Gamma}\,\eta_2(t)$$

$$\hat{\epsilon}'_i(t) = c_i\,\epsilon_i(t) + \hat{G}_{2i}\,\eta_2(t) \qquad i = 1 \ldots r \;:\; \hat{\epsilon}_i(0) \text{ given} \qquad (5.1.10)$$

where $\hat{G}_{2i}$ is the i^{th} row of $\hat{G}_2$ and $\hat{G}_{2\Gamma}$ is the last m-r rows of $\hat{G}_2$. Equation (5.1.10) is the standard error equation for a

state estimator with a scalar measurement. In addition to $\eta'_{1i}(t)$, the estimator is driven by the colored random processes $\hat{\epsilon}_0(t)$ and $\hat{\epsilon}_\Gamma(t)$. The process $\hat{\epsilon}_0(t)$ is defined by the dynamics of the excess subspace:

$$\dot{\hat{\epsilon}}_0(t) = \Pi\, \hat{\epsilon}_0(t) + \eta'_{10}(t) + \hat{\Gamma}_0\, \hat{\epsilon}_\Gamma(t) - \hat{D}_{\Gamma 0}\, \hat{G}_{2\Gamma}\, \eta_2(t) \;:\; \epsilon_0(0) \text{ given} \tag{5.1.11}$$

The second colored random process is defined by the $(A_{\Gamma\Gamma}, C_\Gamma)$ subsystem, i.e. the subsystem which is not directly associated with the r failures:

$$\dot{\hat{\epsilon}}_\Gamma(t) = (\hat{A}'_{\Gamma\Gamma} - \hat{D}_{\Gamma m}\, \hat{C}_{\Gamma m})\, \hat{\epsilon}_\Gamma(t) + \eta'_{1\Gamma}(t) - \hat{D}_{\Gamma m}\, \hat{G}_{2\Gamma}\, \eta_2(t) : \hat{\epsilon}_\Gamma(0) \text{ given}$$

$$\hat{\epsilon}'_\Gamma(t) = \hat{C}_{\Gamma m}\, \hat{\epsilon}_\Gamma(t) + \hat{G}_{2\Gamma}\, \eta_2(t) \tag{5.1.12}$$

From a failure detection vantage point, (5.1.11) and (5.1.12) are important only to the extent that they are helpful in choosing $\hat{D}_\Psi$ and $\hat{D}_\Gamma$ which minimize the variance of $\hat{\epsilon}_1(t) \ldots \hat{\epsilon}_r(t)$ in (5.1.10). From a state estimation vantage point, it is desirable that the variance of $\hat{\epsilon}_0(t)$ and $\hat{\epsilon}_\Gamma(t)$ be minimized as well as that of $\hat{\epsilon}_1(t) \ldots \hat{\epsilon}_r(t)$. Since $d_{\Psi 1} \ldots d_{\Psi r}$ do not affect $\hat{\epsilon}_0(t)$ or $\hat{\epsilon}_\Gamma(t)$, no conflict arises in the choice of $\hat{D}_\Psi$. The choice of $\hat{D}_\Gamma$ affects the variance of all of $\hat{\epsilon}_1(t) \ldots \hat{\epsilon}_r(t)$, $\hat{\epsilon}_0(t)$, $\hat{\epsilon}_\Gamma(t)$; however, and an optimal choice of $\hat{D}_\Gamma$ is not well defined. A reasonable choice of $\hat{D}_\Gamma$ is suggested in the following subsection.

If a failure associated with one of $f_1 \ldots f_r$ occurs, the corresponding driving term $\hat{f}_{ii} n_i(t)$ must be added to the equations describing the subsystem for that $\hat{\epsilon}_i(t)$ in (5.1.10). None of the additional terms appear in (5.1.11) or (5.1.12).

5.1.2 Time Domain Optimization

This subsection is concerned with the application of the time domain approach of estimation theory to the selection of an optimal detection filter-state estimator. For the important case where r equals m and $f_1 \ldots f_m$ are mutually detectable, the Kalman filter equations can be used to specify an optimal $\hat{D}$. Otherwise, a state space augmentation is generally necessary before the Kalman algorithm can be applied. The development presented here is restricted to the steady state case; the extension to the nonstationary case is deferred until Section 5.2 in order to emphasize equivalence to the frequency domain techniques in Section 5.1.3.

If m failures are to be detected by the failure detection system and $f_1 \ldots f_m$ are mutually detectable, $(\hat{A}-\hat{D}\hat{C}, \hat{C})$ is block diagonal. The m subsystems of (5.1.4) are then special cases of (5.1.10):

$$\dot{\hat{\epsilon}}_i(t) = (\hat{A}'_{ii} - d_{\psi i}\,\hat{c}_i)\,\hat{\epsilon}_i(t) - d_{\psi i}\,\hat{G}_{2i}\,\eta_2(t) + \eta'_{1i}(t) \quad i = 1 \ldots m$$

$$\hat{\epsilon}'_i(t) = \hat{c}_i\,\hat{\epsilon}_i(t) + \hat{G}_{2i}\,\eta_2(t) \qquad (5.1.13)$$

The matrix $\hat{D}_\Gamma$ is zero for this case. $\hat{D}$ is uniquely specified if the $d_{\psi 1} \ldots d_{\psi m}$ in the m subsystems described by (5.1.13) are known. The subsystems are decoupled and each can be used to optimize the particular $d_{\psi i}$ associated with it independently of the remaining subsystems.

The effect of the transformation to base normal form is to replace a single n^{th}-order state estimator with a set of m state estimators of order $\nu_1 \ldots \nu_m$ respectively. With the transformed state estimates defined similarly to $\hat{\epsilon}(t)$ and $\hat{\epsilon}'(t)$ in (5.1.3):

$$\hat{x}_m(t) = T^{-1} x_m(t) = \begin{bmatrix} \hat{x}_{m1}(t) \\ \vdots \\ \hat{x}_{mm}(t) \end{bmatrix}$$

(5.1.14)

$$\hat{y}_m(t) = T_m^{-1} y_m(t) = \begin{bmatrix} \hat{y}_{m1}(t) \\ \vdots \\ \hat{y}_{mm}(t) \end{bmatrix}$$

the decoupled state estimators can be expressed:

$$\dot{\hat{x}}_{mi}(t) = (\hat{A}'_{ii} - d_{\psi i}\,\hat{c}_i)\,\hat{x}_{mi}(t) + [\hat{D}\,T_m^{-1} y(t)]_i \;;\; \hat{x}_{mi}(0) \text{ given}$$

$$\hat{y}_{mi}(t) = \hat{c}_i\,\hat{x}_{mi}(t) \qquad i = 1 \ldots m \tag{5.1.15}$$

where $[\hat{D}T_m^{-1}y(t)]_i$ is the appropriate rows of $\hat{D}T_m^{-1}y(t)$. The total computational requirements for these m state estimators is considerably less than that for the single n^{th}-order state estimator. The decoupling is a consequence of the desire to constrain the output error produced by a system failure associated with one of $f_1 \ldots f_m$ to a single direction. If one of the failures occurs, only the accuracy of the state estimator assigned to the appropriate subsystem of $(\hat{A}\text{-}\hat{D}\hat{C},\ \hat{C})$ is degraded -- the state estimates for the remaining state estimators are unaffected.

Since each of $(\hat{A}'_{11},\ \hat{c}_1) \ldots (\hat{A}'_{mm},\ \hat{c}_m)$ has a scalar measurement, the ν_i^{th}-order vectors $d_{\psi 1} \ldots d_{\psi m}$ are the general

feedback filters for the state estimators. The choices of $d_{\psi 1} \cdots d_{\psi m}$ which minimize the steady-state variances of $\hat{\epsilon}_1(t) \cdots \hat{\epsilon}_m(t)$ in (5.1.13) are the steady-state Kalman filters for the appropriate subsystem. These filters are denoted by $d_{\psi K1} \cdots d_{\psi Km}$ and satisfy (2.2.6):

$$d_{\psi Ki} = [\hat{P}_{\infty i} \hat{c}_i^T + \hat{G}'_{12ii}] [\hat{G}_{2i} \hat{G}_{2i}^T]^{-1} \quad i = 1 \cdots m \tag{5.1.16}$$

where $\hat{P}_{\infty i}$ is the positive definite solution of:

$$\hat{A}'_{ii} \hat{P}_{\infty i} + \hat{P}_{\infty i} \hat{A}'^{T}_{ii} + \hat{G}'_{ii} - [\hat{P}_{\infty i} \hat{c}_i^T + \hat{G}'_{12ii}] [\hat{G}_{2i} \hat{G}_{2i}^T]^{-1} [\hat{c}_i \hat{P}_{\infty i} + \hat{G}'_{12ii}] = 0 \tag{5.1.17}$$

With $\hat{D}_{\psi K}$ defined by (4.4.20) for $d_{\psi K1} \cdots d_{\psi Km}$, the optimal detection filter-state estimator is:

$$\hat{D}_K = \hat{D}_{FR} + \hat{D}_{\psi K} \tag{5.1.18}$$

$\hat{D}_K$ minimizes the variance of the state estimation error (5.1.4) subject to the failure detection constraints, but a failure detection system based on $\hat{D}_K$ is not a minimum variance state estimator. That would require all $n \times m$ elements of $\hat{D}_K$ to be optimized rather than the n associated with $d_{\psi K1} \cdots d_{\psi Km}$. No general comparison of $\hat{D}_K$ with a Kalman filter is possible, but several comparisons for particular systems are made in Chapter 6.

The closed form solution for the optimal detection filter-state estimator given by (5.1.16), (4.4.20), and (5.1.18) is valid only if $r = m$ and $f_1 \cdots f_m$ are mutually detectable. If $f_1 \cdots f_m$ are not mutually detectable, the driving terms $\hat{\Theta}_1 \hat{\epsilon}_0(t) \cdots \hat{\Theta}_m \hat{\epsilon}_0(t)$ must be included in (5.1.13) where:

$$\dot{\hat{\epsilon}}_0(t) = \Pi \hat{\epsilon}_0(t) + \eta'_{10}(t) \tag{5.1.19}$$

The error $\hat{\epsilon}_0(t)$ appears in (5.1.13) as colored system noise. The standard estimation theory procedure for dealing with colored system noise is to "whiten" the random process by modeling it as the output of a shaping filter driven by white noise. If the detection filter is to be optimized as a detection filter-state estimator, the reference model may be augmented to include the dynamics and state components of the shaping filter in each of the r subsystems:

$$\bar{\hat{\epsilon}}_i(t) = \begin{bmatrix} \hat{\epsilon}_i(t) \\ \text{-----} \\ \hat{\epsilon}_0(t) \end{bmatrix} \tag{5.1.20}$$

The new state matrices corresponding to (5.1.20) are:

$$\bar{\hat{A}}'_{ii} = \begin{bmatrix} \hat{A}'_{ii} & \hat{\Theta}_i \\ 0 & \Pi \end{bmatrix} \tag{5.1.21}$$

and (5.1.10) must be modified accordingly. The augmented Kalman filters can then be found for each of the resulting $(\nu_i + \nu_0)$ - dimensional subsystems.

In a failure detection system, the shaping filter approach may lead to an unobservable reference model. The procedure for avoiding this difficulty is to use the techniques developed in Section 4.2.3 to remove the excess subspace through state space augmentation. The net effect is the same as in the shaping filter approach, but the new subsystems are $(\nu_i + \nu_0)$ - dimensional or less.

A different approach is to write the steady-state covariance matrices $\hat{P}_{\infty 1} \ldots \hat{P}_{\infty m}$ for (5.1.10) and determine some suboptimal technique for choosing $d_{\psi 1} \ldots d_{\psi m}$. Equation (5.1.17) cannot be used to find the covariances in this instance. A set of coupled differential equations for $\hat{P}_i(t)$, $\hat{P}_{i0}(t)$, and $\hat{P}_0(t)$ must be integrated for a specific choice of $d_{\psi i}$ in order to determine $\hat{P}_{\infty i}$ (Jazwinski (1970)):

$$\dot{\hat{P}}_i(t) = (\hat{A}'_{ii} - d_{\psi i}\hat{c}_i)\,\hat{P}_i(t) + \hat{P}_i(t)(\hat{A}'_{ii} - d_{\psi i}\hat{c}_i)^T + \Theta_i\hat{P}_{i0}(t) + \hat{P}_{i0}(t)\,\hat{\Theta}_i^T + \hat{G}'_{ii} + d_{\psi i}\hat{G}_{2i}\,\hat{G}_{2i}^T\,d_{\psi i}^T - \hat{G}'_{12ii}\,d_{\psi i}^T - d_{\psi i}\,\hat{G}_{12ii}^{\prime T} \quad : \hat{P}_i(0) \text{ given} \tag{5.1.22}$$

$$\dot{\hat{P}}_{i0}(t) = (\hat{A}'_{ii} - d_{\psi i}\,\hat{c}_i)\,\hat{P}_{i0}(t) + \hat{P}_{i0}(t)\,\Pi^T + \hat{G}'_{1i0} - d_{\psi i}\,\hat{G}'^{\,T}_{120i} + \hat{\Theta}_i\hat{P}_0(t)\Pi^T \quad : \hat{P}_{i0}(0) \text{ given} \tag{5.1.23}$$

$$\dot{\hat{P}}_0(t) = \Pi\,\hat{P}_0(t) + \hat{P}_0(t)\,\Pi^T + \hat{G}'_{00} \quad : \hat{P}_0(0) \text{ given} \tag{5.1.24}$$

In general, $d_{\psi 1} \ldots d_{\psi m}$ do not exist which minimize the variance of $\hat{P}_{\infty i} \ldots \hat{P}_{\infty m}$; but it is possible to minimize some scalar function of the covariance. In particular, Section 5.1,3 develops a technique for minimizing the steady state output covariance:

$$\hat{P}'_{\infty i} = \lim_{t\to\infty} \hat{P}_i'(t) = \lim_{t\to\infty} \hat{c}_i\hat{P}_i(t)\,\hat{c}_i^T \qquad i = 1 \ldots m \tag{5.1.25}$$

Once $d_{\psi 1} \ldots d_{\psi m}$ have been chosen according to some criterion, the state estimation error covariances can be ascertained from (5.1.22) - (5.1.24).

A more difficult problem arises if $r < m$. In this case $\hat{D}_\Gamma$ is not zero and (5.1.12) must be considered in the optimization along with (5.1.10). Because of the coupling between $\hat{\epsilon}_1(t) \ldots \hat{\epsilon}_r(t)$ and $\hat{\epsilon}_\Gamma(t)$, a proper Kalman optimization cannot be formulated for both $\hat{D}_\psi$ and $\hat{D}_\Gamma$. A straightforward application of (2.2.5) and (2.2.6) can be used to find the Kalman filter $\hat{D}_{\Gamma Km}$ for (5.1.12), but minimizing the variance of $\hat{\epsilon}_\Gamma(t)$ does not necessarily minimize the contribution of the $\Gamma_1\hat{\epsilon}_\Gamma(t) \ldots \Gamma_r\hat{\epsilon}_\Gamma(t)$ terms to the variances of $\hat{\epsilon}_1(t) \ldots \hat{\epsilon}_r(t)$. Nevertheless, it can be argued that:

$$\hat{D}_{\Gamma K} = \begin{bmatrix} 0 & \cdots & 0 \\ \hline & \vdots & \\ \hline 0 & \cdots & 0 \\ \hline 0 & \cdots & 0 \quad \hat{D}_{\Gamma K m} \end{bmatrix} \qquad (5.1.26)$$

is a reasonable selection. With $\hat{D}_{\Gamma K}$ given, a $\hat{D}_{\Psi K}$ which minimizes the variance of $\hat{\epsilon}_1(t) \ldots \hat{\epsilon}_r(t)$ can be found using a technique similar to that used when $\hat{\epsilon}_0(t)$ was considered. Before pursuing this possibility, however, it is worthwhile to consider some simpler approaches to implement.

The treatment of the $r < m$ case which leads to the simplest design procedure is to determine an additional set of m-r event vectors such that the total of m event vectors are output separable. If possible, it is desirable to make the selection such that the m event vectors are mutually detectable; in which case the optimization problem reduces immediately to (5.1.13). Otherwise, state space augmentation must be used to eliminate the excess subspace. Because they serve primarily as a computational artifice, the additional event vectors need not corresond to physically meaningful failures. The general approach for finding satisfactory event vectors is illustrated in Section 6.2.2. If the design procedure does not lead to satisfactory state estimation properties, a possible remedy is to keep the optimal $\hat{D}_{\Psi K}$ based upon the m event vectors, but replace $\hat{D}_{FR}$ with the $\hat{D}_{FR}$ corresponding tp the original r event vectors, i.e. remove the added event vectors from consideration. Because $\hat{D}_{\Psi K}$ then has terms which should more correctly appear in $\hat{D}_{\Gamma K}$, it is possible to set $\hat{D}_{\Gamma K}$ equal to zero. The net effect of these manipulations is a suboptimization similar to that based on (5.1.26).

A second alternative which requires no state space augmentation is to use the parameter optimization techniques of Section 5.1.3 directly with the n-dimensional model. This approach does not have a single best solution and it may require some trial and error. The best implementation is to choose $\hat{D}_{\Gamma K}$ as (5.1.26) and then use the parameter optimization techniques to select $d_{\Psi K1} \ldots d_{\Psi Kr}$. If this

does not lead to acceptable state estimation properties for some of the r subsystems, a new optimization which includes the components of $\hat{D}_\Gamma$ can be formulated for those subsystems.

The two alternatives mentioned above have the advantage that they define simple suboptimal algorithms for selecting both $\hat{D}_\Gamma$ and $\hat{D}_\Psi$. If, however, (5.1.26) is used, there is a somewhat lengthy procedure for optimizing $\hat{D}_\Psi$ conditioned on this choice of $\hat{D}_\Gamma$. This procedure is outlined in the remainder of this subsection.

Equation (5.1.12) differs from (5.1.11) in that the $(\hat{A}'_{\Gamma\Gamma}, C_\Gamma)$ subsystem has measurements associated with it. This difference is important from a failure detection vantage point, but the difficulties it poses can be circumvented. With $\hat{D}_{\Gamma K}$ specified by (5.1.26), it is necessary to define an augmented reference model which duplicates the dynamics associated with $\hat{\epsilon}_\Gamma(t)$, but not the measurements:

$$\tilde{\hat{A}} - (\tilde{\hat{D}}_{FR} + \tilde{\hat{D}}_{\Gamma K})\tilde{\hat{C}} = \begin{bmatrix} \hat{A}'_{11} & 0 \cdots 0 & \hat{\Gamma}'_1 & \hat{\Gamma}'_1 \\ 0 & 0 & \vdots & \vdots \\ \vdots & \hat{A}'_{rr} & \hat{\Gamma}'_r & \hat{\Gamma}'_r \\ & & \hat{A}'_{\Gamma\Gamma} - \hat{D}_{\Gamma Km}\hat{C}_{\Gamma m} & 0 \\ 0 & \cdots\cdots 0 & & \hat{A}'_{\Gamma\Gamma} - \hat{D}_{\Gamma Km}\hat{C}_{\Gamma m} \end{bmatrix} \tag{5.1.27}$$

where:

$$\begin{aligned} \tilde{\hat{C}} &= [\hat{C} \,\vdots\, 0] \\ \tilde{\hat{D}}_{FR} &= \left[\begin{array}{c:c} \hat{D}_{FR} & 0 \\ \hdashline 0 & 0 \end{array}\right] \\ \tilde{\hat{D}}_{\Gamma K} &= \left[\begin{array}{c:c} \hat{D}_{\Gamma K} & 0 \\ \hdashline 0 & 0 \end{array}\right] \end{aligned} \tag{5.1.28}$$

If the reference model $(\tilde{\hat{A}} - (\hat{D}_{FR} + \tilde{\hat{D}}_{\Gamma K})\, \tilde{\hat{C}}, \tilde{\hat{C}})$ defined by (5.1.27) and (5.1.28) is not observable, the unobservable portion of the augmented subsystem can be removed.

The effect of the state space augmentation is to define a new excess subspace. If the original excess subspace had already been removed using state space augmentation, the state estimation error vector corresponding to (5.1.27) is:

$$\tilde{\hat{\epsilon}}(t) = \begin{bmatrix} \hat{\epsilon}_1(t) \\ \vdots \\ \hat{\epsilon}_r(t) \\ \hat{\epsilon}_\Gamma(t) \\ \tilde{\epsilon}_0(t) \end{bmatrix} \tag{5.1.29}$$

The dynamics associated with $\tilde{\epsilon}_0(t)$ by (5.1.27) are deterministic; however, because $\tilde{\epsilon}_0(t)$ and $\tilde{\epsilon}_\Gamma(t)$ have the same dynamics, it is possible to assume that $\hat{\epsilon}_\Gamma(t)$ is deterministic and that $\tilde{\epsilon}_0(t)$ is driven by the random processes which appear in (5.1.12):

$$\dot{\tilde{\epsilon}}_0(t) = (\hat{A}'_{\Gamma\Gamma} - \hat{D}_{\Gamma K m}\hat{C}_{\Gamma m})\tilde{\epsilon}_0(t) + \eta'_{1\Gamma}(t) - \hat{D}_{\Gamma K m}\hat{G}_{2\Gamma}\,\eta_2(t) \tag{5.1.30}$$

The error $\tilde{\epsilon}_0(t)$ can then be substituted for $\hat{\epsilon}_\Gamma(t)$ in (5.1.10) and the algorithms discussed in conjunction with (5.1.11) can be applied to eliminate the new excess subspace through further state space augmentations. This procedure is circuitous and may lead to a reference model with an unacceptably large dimension, but conditioned on (5.1.26), it yields an optimal detection filter state estimator.

5.1.3 Frequency Domain Optimization

In order to specify $\hat{D}_K$, it is necessary to choose a set of filters $d_{\psi 1} \ldots d_{\psi r}$, $\hat{D}_\Gamma$. If $\hat{D}_K$ is to be an optimal detection filter-state estimator, the set of r+1 filters must be chosen optimally. This subsection presents a frequency domain alternative to the time domain solution of Section 5.1.2. In a design problem, different approaches

may be used to select different elements of the set $d_{\Psi 1} \ldots d_{\Psi r}$, $\hat{D}_\Gamma$. All that is required is that each of $d_{\Psi 1} \ldots d_{\Psi r}$, $\hat{D}_\Gamma$ be chosen optimally. For convenience, the optimal detection filter-state estimator is still denoted by $\hat{D}_K$; however, the solutions defined in this subsection for $d_{\Psi 1} \ldots d_{\Psi r}$ are denoted by $d_{\Psi N1} \ldots d_{\Psi Nr}$.

If the Kalman approach of Section 5.1.2 cannot be used to find a $d_{\Psi Ki}$ which minimizes the variance of $\hat{\epsilon}_i(t)$ for some subsystem of $(\hat{A}-\hat{D}\hat{C}, \hat{C})$ without an unacceptable state space augmentation, or if output stationarity has been used, an alternative is to find a function of $\hat{\epsilon}_i(t)$ which can be minimized. The most obvious choice from a failure detection vantage point is to minimize the variance of the output error $\hat{c}_i\hat{\epsilon}_i(t)$. This is a scalar minimization problem with stationary statistics and the parameter optimization technique of Newton, Gould, and Kaiser (1957) discussed in Section 2.2 can be applied. The parameter optimization yields a unique solution $d_{\Psi Ni}$. If the noise sources for a subsystem are white ($\hat{\Theta}_i$ and $\hat{\Gamma}_i'$ are zero), $d_{\Psi Ni}$ is equal to the Kalman filter $d_{\Psi Ki}$ for that subsystem and the parameter optimization minimizes the variance of the error state vector $\hat{\epsilon}_i(t)$ as well as the output. It follows that all of $d_{\Psi Ni} \ldots d_{\Psi Nr}$ can be determined in lieu of using the Kalman approach if desired.

The steady-state output error variances $\hat{c}_1\hat{P}_{\infty 1}\hat{c}_1^T \ldots \hat{c}_r\hat{P}_{\infty r}\hat{c}_r^T$ can be expressed as the integral of the corresponding power spectral densities $\hat{c}_1\hat{\Phi}_1(s)\hat{c}_1^T \ldots \hat{c}_r\hat{\Phi}_r(s)\hat{c}_r^T$. These in turn can be developed from (2.2.10) in terms of the transfer functions for $\hat{c}_1\hat{\epsilon}_1(s) \ldots \hat{c}_r\hat{\epsilon}_r(s)$ and the power spectral densities of the noise sources. The Laplace transform for (5.1.10) with zero initial conditions is:

$$(sE_{\nu_i} - \hat{A}_{ii}' + d_{\Psi i}\hat{c}_i)\hat{\epsilon}_i(s) = \eta_{1i}'(s) - d_{\Psi i}\hat{G}_{2i}\eta_2(s) + \hat{\Theta}_i\hat{\epsilon}_0(s) + \hat{\Gamma}_i\hat{\epsilon}_\Gamma(s)$$

$$i = 1 \ldots r \qquad (5.1.31)$$

The matrix on the left-hand side of (5.1.31) is invertible. By taking the inverse, the equation can be solved for the transfer functions of $\hat{\epsilon}_1(s) \ldots \hat{\epsilon}_r(s)$ and then $\hat{c}_1\hat{\epsilon}_1(s) \ldots \hat{c}_r\hat{\epsilon}_r(s)$. First, however, it is desirable to simplify the expression using $\hat{A}_{ii}$ from (4.4.23):

$$sE_{\nu i} - \hat{A}_{ii} = sE_{\nu i} - \hat{A}'_{ii} + d_{\psi i}\hat{c}_i = \begin{bmatrix} s & 0 & \cdots & 0 & p_{i1} \\ -1 & \ddots & \ddots & 0 & \vdots \\ 0 & \ddots & \ddots & s & p_{i(\nu_i-1)} \\ \vdots & \ddots & \ddots & -1 & s+p_{i\nu_i} \\ 0 & \cdots & 0 & & \end{bmatrix}$$

$$i = 1 \ldots r \qquad (5.1.32)$$

The coefficients $p_{i1} \ldots p_{i\nu_i}$ for $i = r$ are those of the desired minimal polynomials $\psi_1(\cdot) \ldots \psi_r(\cdot)$ for the r subsystems in (5.1.10) and are the parameters to be optimized. The filters $d_{\psi 1} \ldots d_{\psi r}$ can then be determined from the coefficients using (4.4.21). Equation (5.1.32) is a standard form and its inverse can be evaluated analytically:

$$(sE_{\nu_i} - \hat{A}_{ii})^{-1} = \frac{1}{\psi_i(s)} \begin{bmatrix} x & \cdots & x \\ \vdots & & \vdots \\ x & \cdots & x \\ 1 & s & \ldots & s^{\nu_i - 1} \end{bmatrix} i = 1 \ldots r \qquad (5.1.33)$$

The x's are not zero, but they are not important here. It is a simple matter to verify that:

$$\hat{c}_i(sE_{\nu_i} - \hat{A}_{ii})^{-1} = \frac{1}{\psi_i(s)} \begin{bmatrix} 1 & s & \ldots & s^{\nu_i - 1} \end{bmatrix} = \frac{1}{\psi_i(s)} z_i(s)^T \qquad (5.1.34)$$

The transfer functions for $\hat{c}_1\hat{\epsilon}_1(s) \ldots \hat{c}_r\hat{\epsilon}_r(s)$ can be expressed in terms of the vectors $z_1(s) \ldots z_r(s)$:

$$\hat{c}_i\hat{\epsilon}_i(s) = \frac{1}{\psi_i(s)} z_i(s)^T \left\{ \eta'_{1i}(s) - d_{\psi i}\hat{G}_{2i}\,\eta_2(s) + \hat{\Theta}_i\hat{\epsilon}_0(s) + \hat{\Gamma}_i\,\hat{\epsilon}_\Gamma(s) \right\} \quad i = 1 \ldots r \qquad (5.1.35)$$

The power spectral densities of each of the noise sources in (5.1.35) must be determined along with the cross-power spectral densities. The computation is straightforward but tedious (see Section 2.2). Ignoring the cross-power spectral densities, the resulting power spectral densities for the output errors are:

$$\hat{c}_i\hat{\Phi}_i(s)\hat{c}_i^T = \frac{1}{\psi_i(s)\psi_i(-s)} z_i(s)^T\Big\{\Phi'_{\eta 1i}(s)+d_{\psi i}\hat{e}_{mi}^T\hat{G}_2\hat{G}_2^T\hat{e}_{mi}d_{\psi i}^T$$

$$+ \hat{\Theta}_i\hat{\Phi}'_0(s)\hat{\Theta}_i^T + \hat{\Gamma}_i\Phi_\Gamma(s)\hat{\Gamma}_i^T\Big\} z_i(-s) \qquad (5.1.36)$$

where $\Phi'_{\eta 1}$ is the power spectral density of $\eta_1'(t)$ etc. The integrals of $\hat{c}_1\hat{\Phi}_1(s)\hat{c}_1^T \ldots \hat{c}_r\hat{\Phi}_r(s)\hat{c}_r^T$ are the output error variances $\hat{c}_1\hat{P}_{\infty 1}\hat{c}_1^T \ldots \hat{c}_r\hat{P}_{\infty r}\hat{c}_r^T$. If each pair of cross-power spectral densities is combined in one term, every term of $\hat{c}_1\hat{\Phi}_1(s)\hat{c}_1^T \ldots \hat{c}_r\hat{\Phi}_r(s)\hat{c}_r^T$ can be factored into a rational polynomial of s and the same polynomial of -s:

$$\hat{c}_i\hat{\Phi}_i(s)\hat{c}_i^T = v_{1i}(s)v_{1i}(-s)+v_{2i}(s)v_{2i}(-s) + \ldots \quad i=1 \ldots r, \Gamma \qquad (5.1.37)$$

The variances can then be expressed:

$$\hat{c}_i\hat{P}_{\infty i}\hat{c}_i^T = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\hat{c}_i\hat{\Phi}_i(s)\hat{c}_i^T ds = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty}\Big(v_{1i}(s)v_{1i}(-s)$$

$$+ v_{2i}(s)v_{2i}(-s) + \ldots \Big) ds \qquad (5.1.38)$$

The above is a standard integral form and can be evaluated using the tables in Appendix B.

The variances $\hat{c}_1\hat{P}_{\infty 1}\hat{c}_1^T \ldots \hat{c}_r\hat{P}_{\infty r}\hat{c}_r^T$ are algebraic functions of the appropriate coefficients $p_{i1} \ldots p_{i\nu_1}$, $i = 1 \ldots r$, of $d_{\psi 1} \ldots d_{\psi r}$. If the optimal coefficients are denoted by $p_{iN1} \ldots p_{iN\nu_i}$, $i = 1 \ldots r$, they must satisfy:

$$\hat{c}_i\hat{P}_{\infty i}\hat{c}_i^T(p_{i1} \ldots p_{i\nu_i}) \geq \hat{c}_i\hat{P}_{\infty i}\hat{c}_i^T(p_{iN1} \ldots p_{iN\nu_i}) \qquad (5.1.39)$$

for all real coefficients $p_{i1} \ldots p_{i\nu_i}$, $i = 1 \ldots r$. If the optimal coefficients are finite, they define a stationary point with positive second partial derivatives:

$$\frac{\partial(\hat{c}_i \hat{P}_{\infty i} \hat{c}_i^T)}{\partial p_{ij}} = 0 \qquad \begin{matrix} j = 1 \ldots \nu_i \\ i = 1 \ldots r \end{matrix} \tag{5.1.40}$$

$$\frac{\partial^2(\hat{c}_i \hat{P}_{\infty i} \hat{c}_i^T)}{\partial p_{ij}} > 0 \tag{5.1.41}$$

For each subsystem, (5.1.40) provides a set of ν_i equations which specify the optimal coefficients and hence $d_{\Psi Ni}$. If $\nu_i > 2$, the equations are generally nonlinear and digital function minimization techniques must be used to find $p_{iN1} \ldots p_{iN\nu_i}$.

If the event vectors $h_1 \ldots h_q$ are to be output stationary with $f_1 \ldots f_r$, Section 4.2.4 specifies $p_{i1} \ldots p_{i\nu_i}$, $i = 1 \ldots r$, as functions of the coefficients of the desired polynomials $Y_1(\cdot) \ldots Y_q(\cdot)$ associated with $\bar{\mathscr{P}}_1 \ldots \bar{\mathscr{P}}_q$. For convenience, assume that only h_j is to be made output stationary with $f_i \ldots f_r$ and the index set I_j exists for which:

$$\beta_j \, Ch_j = \sum_{k \in I_j} \alpha_{kj} Cf_k \tag{5.1.42}$$

where α_{kj}, $k \in I_j$, are nonzero. The parameters $p_{i1} \ldots p_{i\nu_i}$, $i \notin I_j$, are not affected by the output stationarity and can still be optimized using either (5.1.40) or the Kalman approach. The only remaining freedom in the choice of $\hat{D}_\Psi$ is in the coefficients of $Y_1(\cdot)$. If a set of weighting functions β'_k, $k \in I_j$, are selected, the coefficients of $Y_1(\cdot)$ can be constrained to minimize the cost function:

$$\hat{P}_{\infty I_j} = \sum_{k \in I_j} \beta'_k \hat{c}_k \hat{P}_{\infty k} \hat{c}_k^T \tag{5.1.43}$$

in a manner identical to (5.1.40). If the β'_k's, $k \in I_j$, are equal to the squares of the respective α_{ki}'s, $\hat{P}_{\infty I_j}$ is the covariance of the output error associated with Ch_j. Otherwise, they reflect the relative cost associated with output errors along each of $Cf_1 \ldots Cf_r$. Equation (5.1.43), together with (5.1.40) for $i \notin I_j$, uniquely specify an optimal $\hat{D}_K$ as a function of β'_k, $k \in I_j$.

5.2 Suboptimal State Estimation--Time-varying Statistics

In certain applications, the steady state covariance matrices $\hat{P}_{\infty 1} \ldots \hat{P}_{\infty r}$ upon which Section 5.1 was predicted do not exist. This may be due to the fact that the random processes in the stochastic system model are nonstationary, or it may be due to the fact that the response of the system before it reaches steady state is of interest. The Kalman filter approach of Section 5.1.2 for the optimal detection filter-state estimator is generalizable to nonstationary problems. The parameter optimization technique of Section 5.1.3 cannot be applied rigorously, but acceptable approximations can be made in many instances. The extension of both these subsections to the nonstationary statistics case is developed in this subsection. Because of their extensive utilization in practice, attention is also directed toward sampled-data systems.

The Kalman filter for continuous systems with nonstationary statistics is time-varying. If the noise sources remain white and zero mean, the stochastic model of Section 5.1.1 can be modified by making P_i, G_{12}, G_1, and G_2 functions of time. Instead of the steady state solution (5.1.16) and (5.1.17), the Kalman filter is a function of $\hat{P}_i(t)$:

$$d_{\psi K i}(t) = [\hat{P}_i(t)\, \hat{c}_i^{\,T} + \hat{G}'_{12ii}(t)]\, [\hat{G}_{2i}(t)\, \hat{G}_{2i}(t)^T]^{-1} \quad i = 1 \ldots r \qquad (5.2.1)$$

where $\hat{P}_i(t)$ satisfies the differential equation:

$$\dot{\hat{P}}_i(t) = \hat{A}'_{ii}\, \hat{P}_i(t) + \hat{P}_i(t)\, \hat{A}'^{\,T}_{ii} + \hat{G}'_{ii}(t)$$

$$-[\hat{P}_i(t)\, \hat{c}_i^{\,T} + \hat{G}'_{12ii}(t)]\, [\hat{G}_{2i}(t)\, \hat{G}_{2i}(t)^T]^{-1} [\hat{c}_i\, \hat{P}_i(t) + \hat{G}'_{12ii}(t)^T]$$

$$i = 1 \ldots r : P_i(0) \text{ given} \qquad (5.2.2)$$

If the nonstationary Kalman filter is used, (5.2.2) and (5.2.1) must either be precomputed and stored or solved in real-time along with the reference model equations in order to determine $d_{\psi K1}(t) \ldots d_{\psi Kr}(t)$

as a function of time. Utilization of (5.2.1) and (5.2.2) leads to a time-varying detection filter-state estimator, but this does not affect the failure detection system properties.

An alternative to (5.2.1) is a piece-wise stationary approximation to the system which would allow the parameter optimization procedure of Section 5.1.3 to be applied. The technique is useful if $P(t)$, $G_{12}(t)$, $G_1(t)$, and $G_2(t)$ can be approximated as piece-wise constant over time intervals appreciably longer than the response time of the system. Each time a step change in the variances is made, the algorithm of Section 5.1.3 must be repeated to choose new vectors $d_{\psi N1}(t) \ldots d_{\psi Nr}(t)$ as a function of the time. The performance of the resulting detection filter-state estimator depends on the rate of change of the noise variances relative to the system response time and the magnitude of the steps.

The algorithms of the preceding paragraphs select different vectors $d_{\psi 1} \ldots d_{\psi r}$ when the noise processes for the system are nonstationary. The Kalman algorithm has the advantage that it is optimal, but it may require a large state space augmentation. The selection between the two alternatives must be made subjectively on the basis of performance and computational requirements.

Section 4.6 motivated the fact that failure detection theory is not analytically valid for continuous deterministic systems with discrete measurements (sampled-data systems). If a first order expansion of the state transition matrix $\Phi(t, t+\Delta t)$ is sufficiently accurate, however, a detection filter can be specified for a discrete system approximation to the sampled-data system. The detection filter chosen for the discrete system can be used with the sampled-data system to constrain the output errors produced by failures to be approximately unidirectional. A similar result is true for stochastic sampled-data systems, but for state estimation purposes it is desirable to modify the representation of the reference model given in (4.6.6) to be consistent with that commonly used for state estimators. No analytical difficulties are introduced because the representation given below and that in (4.6.6) both converge to the continuous reference model as Δt approaches zero.

The system equation for a continuous stochastic system with discrete measurements and no failure is given by:

$$x(t+\Delta t) = \Phi(t+\Delta t, t)\, x(t) + \int_{t}^{t+\Delta t} \Phi(t+\Delta t, \tau) B u(\tau)\, d\tau + G_{1\Delta}\, \eta_{1\Delta}(t+\Delta t)$$

$$y(t+\Delta t) = C\, x(t+\Delta t) + G_{2\Delta}\, \eta_{2\Delta}(t+\Delta t) \qquad (5.2.3)$$

where $\eta_{1\Delta}(t)$ and $\eta_{2\Delta}(t)$ are discrete, zero mean, white vector processes with finite variances:

$$E[\eta_{1\Delta}(t)\, \eta_{1\Delta}(t)^T] = \frac{1}{\Delta t} E_{n_1}$$

$$E[\eta_{2\Delta}(t)\, \eta_{2\Delta}(t)^T] = \frac{1}{\Delta t} E_{m_1} \qquad (5.2.4)$$

A state estimation system for a linear sampled-data system can be defined as a two stage processor. The first stage involves the propagation of the state estimate $x_m(t)$ to time $t+\Delta t$ and the measurement of the propagated estimate $\bar{x}_m(t+\Delta t)$:

$$\bar{x}_m(t+\Delta t) = \Phi(t+\Delta t, t)\, x_m(t) + \int_{t}^{t+\Delta t} \Phi(t+\Delta t, \tau)\, B u(\tau)\, d\tau$$

$$\bar{y}_m(t+\Delta t) = C\, \bar{x}_m(t+\Delta t) \qquad (5.2.5)$$

Because of the stochastic nature of the problem, $\bar{x}_m(t+\Delta t)$ contains an error even if the system has not failed. The standard state estimation technique for minimizing the state estimation error variance is to use the new system measurement $y(t+\Delta t)$ to update the state estimate:

$$x_m(t+\Delta t) = \bar{x}_m(t+\Delta t) + D_{\Delta}(y(t+\Delta t) - \bar{y}_m(t+\Delta t)) \qquad (5.2.6)$$

The two stages of the state estimator make it necessary to define errors for the propagated state estimate as well as the updated state estimate:

$$\bar{\epsilon}(t+\Delta t) = x(t+\Delta t) - \bar{x}_m(t+\Delta t) \tag{5.2.7}$$

$$\epsilon(t+\Delta t) = x(t+\Delta t) - x_m(t+\Delta t)$$

The measured error is a third quantity which is frequently referred to as the measurement residual:

$$\bar{\epsilon}'(t+\Delta t) = y(t+\Delta t) - C\bar{x}_m(t+\Delta t) \tag{5.2.8}$$

The error equations follow from (5.2.4) - (5.2.7):

$$\bar{\epsilon}(t+\Delta t) = \Phi(t+\Delta t, t)\,\epsilon(t) + G_{1\Delta}\,\eta_{1\Delta}(t+\Delta t) \tag{5.2.9}$$

$$\bar{\epsilon}'(t+\Delta t) = C\,\bar{\epsilon}(t+\Delta t) + G_{2\Delta}\,\eta_{2\Delta}(t+\Delta t)$$

$$\epsilon(t+\Delta t) = \bar{\epsilon}(t+\Delta t) - D_\Delta\,\bar{\epsilon}'(t+\Delta t) \tag{5.2.10}$$

The covariances for $\bar{\epsilon}(t+\Delta t)$ and $\epsilon(t+\Delta t)$ respectively are:

$$\bar{P}(t+\Delta t) = \Phi(t+\Delta t, t)\,P(t)\,\Phi(t+\Delta t, t)^T + \frac{1}{\Delta t}\,G_{1\Delta}\,G_{1\Delta}^T$$

$$P(t+\Delta t) = (E_n - D_\Delta C)\,\bar{P}(t+\Delta t)\,(E_n - D_\Delta C)^T + \frac{1}{\Delta t}\,D_\Delta G_{2\Delta}\,G_{2\Delta}^T\,D_\Delta^T \tag{5.2.11}$$

The variance of $\bar{\epsilon}'(t+\Delta t)$ differs from $C\bar{P}(t+\Delta t)\,C^T$ in that it includes the effect of the measurement uncertainty:

$$\bar{P}'(t+\Delta t) = C\bar{P}(t+\Delta t)\,C^T + \frac{1}{\Delta t}\,G_{2\Delta}\,G_{2\Delta}^T \tag{5.2.12}$$

Equations (5.2.5) and (5.2.6) have the standard form of a state estimator for sampled-data systems, but it is not immediately obvious that they have the form of a failure detection system. It is also not apparent how D_Δ should be chosen if the equations do define a failure detection system. These questions can be answered by considering a continuous system approximation of (5.2.3) - (5.2.10). The necessary procedure is to first approximate (5.2.3) by a

discrete system and then consider the discrete to continuous limit. Linearizing the state transition matrix:

$$\Phi(t+\Delta t, t) \approx E_n + \Delta t A \tag{5.2.13}$$

and keeping only first order terms, (5.2.3) becomes:

$$\begin{aligned} x(t+\Delta t) &= (E_n + \Delta t A)\, x(t) + \Delta t\, B\, u(t) + G_{1\Delta}\eta_{1\Delta}(t+\Delta t) \\ y(t+\Delta t) &= C\, x(t+\Delta t) + G_2\, \eta_{2\Delta}(t+\Delta t) \end{aligned} \tag{5.2.14}$$

Equation (5.2.14) is the stochastic version of (4.6.11) with no failure. The state equations for the failure detection system can be expanded similarly:

$$\begin{aligned} \bar{x}_m(t+\Delta t) &= (E_n + \Delta t A)\, x_m(t) + \Delta t\, B\, u(t) \\ x_m(t+\Delta t) &= \bar{x}_m(t+\Delta t) + D_\Delta(y(t+\Delta t) - C\bar{x}_m(t+\Delta t)) \end{aligned} \tag{5.2.15}$$

The error equations can be developed in the same manner.

From (5.2.13) - (5.2.15) it is possible to consider the discrete-to-continuous limit and reduce (5.2.15) to a single equation. With the substitutions:

$$\begin{aligned} G_{1\Delta} &= \Delta t\, G_1 \\ G_{2\Delta} &= G_2 \\ D_\Delta &= \Delta t\, D \end{aligned} \tag{5.2.16}$$

(5.2.14) becomes the standard continuous stochastic system of Section 5.1:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + G_1\, \eta_1(t) \\ y(t) &= C\, x(t) + G_2\, \eta_2(t) \end{aligned} \tag{5.2.17}$$

Equation (5.2.15) can be placed in incremental form:

$$\frac{x_m(t+\Delta t) - x_m(t)}{\Delta t} = A\, x_m(t) + B\, u(t) + D(y(t+\Delta t) - C\bar{x}_m(t+\Delta t)) \tag{5.2.18}$$

In the limit as Δt approaches zero, $\bar{x}_m(t+\Delta t)$ approaches $x_m(t)$ and (5.2.18) becomes the standard continuous reference model of Chapter 3:

$$\dot{x}_m(t) = A\, x_m(t) + Bu(t) + D(y(t) - C\, x_m(t))$$

$$y_m(t) = C\, x_m(t) \tag{5.2.19}$$

The error equations become the stochastic error equations of Section 5.1.

Equations (5.2.16) and (5.2.19) say that a detection filter can be determined for the continuous approximation of (5.2.3) and then used to specify D_Δ. With the base normal form for (5.2.19) as defined in (4.4.3), this suggests that Section 5.1 can be used to determine a continuous optimal detection filter-state estimator $\hat{D}_K$ and that a reasonable choice for a suboptimal sampled-data detection filter-state estimator is:

$$\hat{D}_{\Delta K} = \Delta t\, \hat{D}_K = \Delta t\, (\hat{D}_{FR} + \hat{D}_{\Psi K} + \hat{D}_{\Gamma K}) \tag{5.2.20}$$

For many applications (5.2.20) is satisfactory. If the sampled-data system has precise measurements and a large sampling period, however, (5.2.20) has a tendency to overweight the measurements and $x_m(t+\Delta t)$ can become a poorer state estimate than $\bar{x}_m(t+\Delta t)$.

Improvement over (5.2.20) can be made by considering a new approximation of (5.2.3), (5.2.5), and (5.2.6). The suboptimal detection filter-state extimator can be decomposed into the sum of three matrices be defining $\hat{D}_{\Delta FR}$, $\hat{D}_{\Delta \Psi K}$, and $\hat{D}_{\Delta \Gamma K}$ analogously to $\hat{D}_{FR}$, $\hat{D}_{\Psi K}$, and $\hat{D}_{\Gamma K}$:

$$\hat{D}_{\Delta K} = \hat{D}_{\Delta FR} + \hat{D}_{\Delta \Psi K} + \hat{D}_{\Delta \Gamma K} \tag{5.2.21}$$

$\hat{D}_{\Delta FR}$ is uniquely specified as:

$$\hat{D}_{\Delta FR} = \Delta t\, \hat{D}_{FR} \tag{5.2.22}$$

by the desire to constrain the output error produced by failures to be nearly unidirectional. $\hat{D}_{\Delta \Psi K}$ and $\hat{D}_{\Delta \Gamma K}$ can be chosen to improve the state estimation properties of (5.2.5) and (5.2.6).

Consider the transformation of the sampled-data system (5.2.3) along with (5.2.5) and (5.2.6) into the coordinate system of the base normal form defined for (5.2.19):

$$\begin{aligned} \hat{\Phi}(t,\tau) &= T^{-1}\, \Phi(t,\tau) T \\ \hat{C} &= T_m{}^{-1}\, C\, T \\ \hat{D}_{\Delta K} &= T^{-1}\, D_{\Delta K}\, T_m \end{aligned} \tag{5.2.23}$$

Although $\hat{A}$-$\hat{D}\hat{C}$ may be upper block diagonal for some choice of $\hat{D}$, in general $\hat{\Phi}(t,\tau)$ is not upper block diagonal for any choice of $\hat{D}_{\Delta}$:

$$\hat{\Phi}(t,\tau) = \begin{bmatrix} \hat{\Phi}_{11} & \hat{\Phi}_{12} \cdots \hat{\Phi}_{1r} & \hat{\Phi}_{10} & \hat{\Phi}_{1\Gamma} \\ \hat{\Phi}_{21} & \hat{\Phi}_{22} \searrow & \vdots & \vdots \\ \hat{\Phi}_{r1} \cdots\cdots & \hat{\Phi}_{rr} & \hat{\Phi}_{r0} & \hat{\Phi}_{r\Gamma} \\ \hat{\Phi}_{01} \cdots\cdots & \hat{\Phi}_{0r} & \hat{\Phi}_{00} & \hat{\Phi}_{0\Gamma} \\ \hat{\Phi}_{\Gamma 1} \cdots\cdots & \hat{\Phi}_{\Gamma r} & \hat{\Phi}_{\Gamma 0} & \hat{\Phi}_{\Gamma\Gamma} \end{bmatrix} \tag{5.2.24}$$

where $\hat{\Phi}_{11}$ is $\nu_1 \times \nu_1$ etc. As long as $\hat{D}_{\Delta K}$ is chosen to satisfy (5.2.21) and (5.2.22) and the continuous approximation (5.2.19) is reasonable, a failure related to the first subsystem of the continuous approximation (5.2.19) primarily affects the accuracy of only the first ν_1 components of the state estimate $\hat{x}_m(t)$ in the transformed sampled data system. The presence of $\hat{\Phi}_{21} \cdots \hat{\Phi}_{r1}$

in (5.2.24) means that estimation errors are generated in the remaining $n-\nu_1$ components of $\bar{x}_m(t+\Delta t)$, but the output feedback due to the $\hat{D}_{\Delta FR}$ component of $\hat{D}_\Delta$ minimizes the growth of these errors components. This error isolation property can be extrapolated to the first r subsystems of $\hat{A}-\hat{D}\hat{C}$.

It follows from the preceding arguments that a reasonable procedure for choosing $\hat{D}_{\Delta\Psi}$ and $\hat{D}_{\Delta\Gamma}$ in (5.2.22) is to replace $\Phi(t,\tau)$ by:

$$\hat{\Phi}'(t,\tau) = \begin{bmatrix} \hat{\Phi}_{11} & 0 & \hat{\Phi}_{10} & \hat{\Phi}_{1\Gamma} \\ 0 & \hat{\Phi}_{22} & \vdots & \vdots \\ & \hat{\Phi}_{rr} & \hat{\Phi}_{r0} & \hat{\Phi}_{r\Gamma} \\ 0 & 0 & \hat{\Phi}_{00} & \hat{\Phi}_{0\Gamma} \\ 0 & 0 & \hat{\Phi}_{\Gamma 0} & \hat{\Phi}_{\Gamma\Gamma} \end{bmatrix} \qquad (5.2.25)$$

The failure detection system can then be decomposed into a set of r+2 subsystems precisely as was done in Section 5.1.1 for $\hat{A}-\hat{D}\hat{C}$ -- the base normal form for (5.2.25) is that of (5.2.19). The matrix $\hat{D}_{\Delta\Gamma K}$ can be chosen optimally using the discrete Kalman filter for sampled data systems (Bryson and Ho (1969) or Jazwinski (1970)) rather than the continuous filter. Similarly, if $\hat{D}_{\Delta\Psi K}$ is defined by (4.4.20):

$$\hat{D}_{\Delta\Psi K} = \begin{bmatrix} d_{\Delta\Psi K1} & 0 & 0 & 0 \\ 0 & d_{\Delta\Psi K2} & \vdots & \vdots \\ & & d_{\Delta\Psi Kr} & \\ 0 & 0 & 0 & 0 \end{bmatrix} \qquad (5.2.26)$$

the discrete Kalman filters $d_{\Delta\Psi K1} \cdots d_{\Delta\Psi Kr}$ which minimize the variance of the block diagonalized system can be found using the

procedure in Section 5.1.2. If the Kalman approach is not used for the i^{th} subsystem of $(\hat{A}-\hat{D}\hat{C},\hat{C})$, the parameter optimization technique can be used to define the optimal $d_{\Psi Ni}$ for the corresponding subsystem in (5.2.19) and $d_{\Delta\Psi Ki}$ in (5.2.26) can be defined by:

$$d_{\Delta\Psi Ki} = \Delta t\, d_{\Psi Ni} \tag{5.2.27}$$

An alternative to (5.2.25) is to replace $\hat{\Phi}(t,\tau)$ by:

$$\hat{\Phi}''(t,\tau) = e^{(\hat{A}-\hat{D}_{FR}\hat{C})\,(t-\tau)} \tag{5.2.28}$$

Although (5.2.25) and (5.2.28) are not identical, they are nearly equal for small Δt. The state transition matrix $\hat{\Phi}''(t,\tau)$ is upper block diagonal and can be used to determine $\hat{D}_{\Delta K}$ and $\hat{D}_{\Delta\Gamma}$ using either (5.2.26) or (5.2.27).

In general (5.2.21) can be expected to yield a better state estimator than (5.2.20). The validity of $\hat{D}_{\Delta K}$ as a detection filter is a function of the degree to which it constrains failure-induced output errors to a single direction. The amount of acceptable "leakage" of the error signal is a function of the strength of the stochastic noise sources. The potential value as well as the limitations of a failure detection system for a stochastic sampled data system are illustrated in Chapter 6.

5.3 Detection Law for Multiple Failure Detection Systems

If the number of failures to be monitored is large, two or more failure detection systems may be required. The failure detection systems can be implemented independently (Figure 5.1). One advantage which accrues from the use of multiple failure detection systems is that each generates a state estimate. Although the state estimates are not independent, they can be combined to yield a better state estimate than any of the individual estimates. A disadvantage is that failure isolation is more difficult when multiple failure detection systems are used. If a failure occurs, it affects the state estimates in all of the failure detection systems and each interprets the resulting output estimation errors in terms

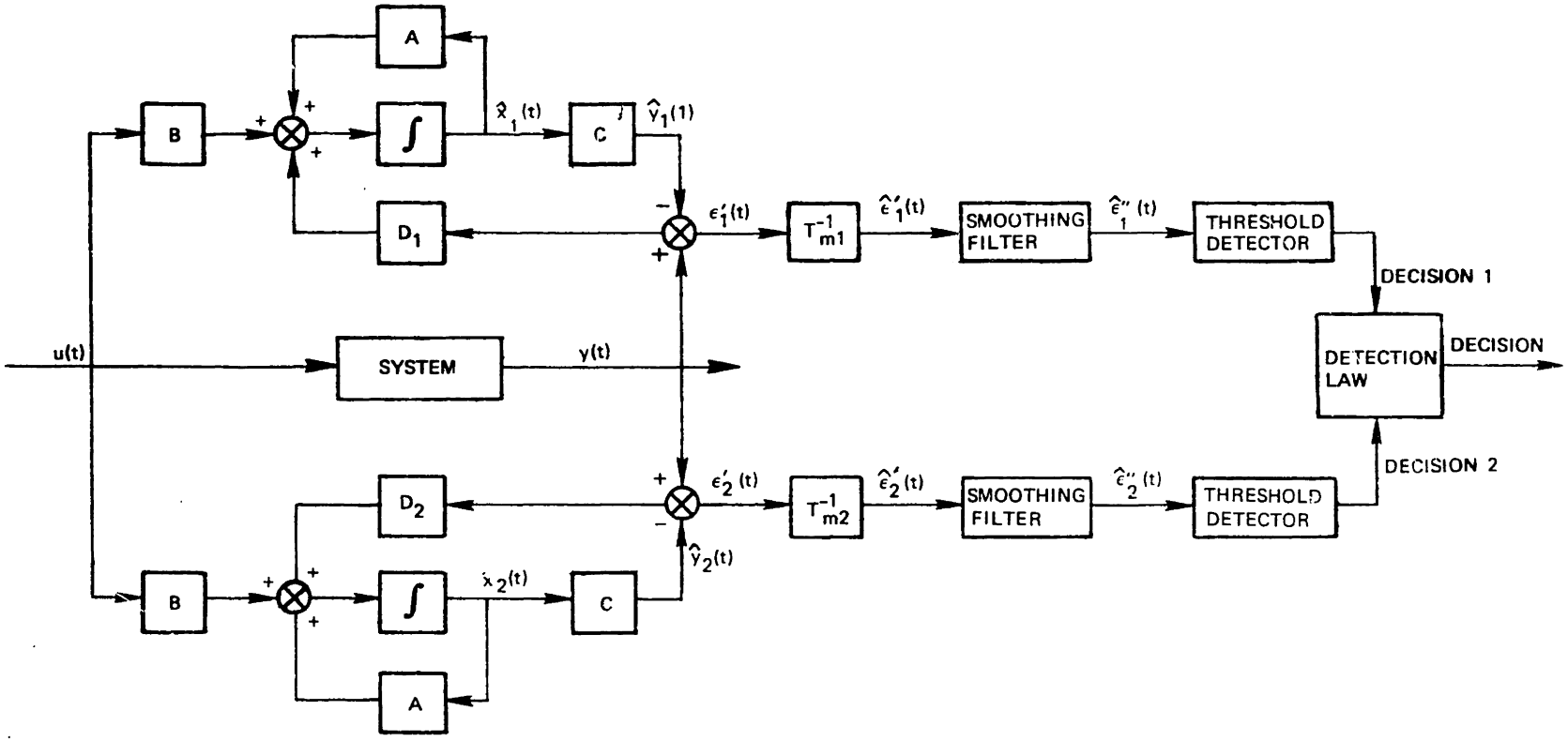


Figure 5.1 Implementation of parallel detection systems with threshold detection logic

of the failures it was designed to detect. It is necessary to have a detection law which resolves conflicts between the detection decisions made by the individual failure detection systems.

Assume the two failure detection systems of Figure 5.1 are implemented and that $(A\text{-}D_1C, C)$ is designed to detect the failure associated with an event vector f_i. If that failure occurs, the output estimation error for $(A\text{-}D_1C, C)$ is unidirectional along Cf_i and $(A\text{-}D_1C, C)$ detects the failure correctly. An output error signal is also generated by the second failure detection system and a second failure detection decision is made. If $(A\text{-}D_2C, C)$ is also designed to detect the failure associated with f_i, it produces a unidirectional output error along Cf_i and identifies the failure correctly. If $(A\text{-}D_2, C, C)$ is not designed for the failure associated with f_i, however, Corollary 1 of Theorem 3.2.1 (Detection Theorem) guarantees that f_i does not produce a unidirectional output for $(A\text{-}D_2C, C)$. The $(A\text{-}DC_2, C)$ failure detection system therefore determines that either a sensor failure or some multiple failure has occurred. In this instance a detection law makes the correct decision if it accepts the decision of the failure detection system in which the output error was restricted to the lowest dimensional subspace of the output space $\mathcal{Y}^m$.

The detection law stated in the preceding paragraph is generalizable to sensor failures. If a sensor failure for which $(A\text{-}D_2C, C)$ was designed occurs, the output error signal for $(A\text{-}D_1C, C)$ cannot be constrained to a two-dimensional subspace unless $(A\text{-}D_1C, C)$ was also designed to detect the sensor failure. In the event of a single failure, the detection law always identifies the failure correctly. If several failures occur simultaneously, the detection law may yield an ambiguous decision. If a failure occurs which none of the failure detection systems is designed to detect, the detection law can be expected to yield an ill-defined decision. When designing a failure detection system, it is to the designer's advantage to know how the failure detection system and the detection law will respond in all possible situations. An orderly procedure for developing and implementing this detection law is illustrated in Section 6.3.

The detection law requires that each of the individual failure detection systems make detection decisions before a final decision is made. In the case of the failure associated with f_i, this requires that $(A\text{-}D_1C, C)$ indicate the correct failure and $(A\text{-}D_2C, C)$ indicate either a sensor failure or a multiple failure. A certain waiting period must be allowed for each failure detection system to observe the behavior of its output error signal. The length of the waiting period is a function of the system, the failure, and the failure detection system. It can only be determined through a trial and error process.

The implementation of a detection law is complicated by the presence of stochastic disturbances. It is necessary for each failure detection system to differentiate between output errors generated by stochastic sources and those generated by failures. The simplest procedure is to establish a threshold criterion -- other criteria are discussed in Section 5.4. A detection law based upon the threshold criterion is developed in this section for systems with continuous measurements. A similar law for sampled-date systems is developed in Section 6.3. To simplify the analysis, it is assumed here that the estimation errors in the unfailed system satisfy (5.1.13).

The stochastic model upon which Section 5.1 is based assumes that the measurement random process $\eta_2(t)$ is white noise. In many respects this model is a reasonable approximation of the uncertainty in continuous measurements; but, unfortunately, the threshold criterion makes its decision on the basis of the magnitude of the residuals. Because of the additive white noise in the measurement:

$$\epsilon'(t) = C\epsilon(t) + G_2\,\eta_2(t) \qquad (2.2.3)$$

the variance of $\epsilon'(t)$ is infinite. In order to define a threshold test, it is necessary to first smooth $\epsilon'(t)$. The smoothing interval can be made arbitrarily small so that it attenuates the high frequency portion of $\epsilon'(t)$ without altering $C\epsilon(t)$ appreciably. Although some smoothing algorithms may yield better signal-to-noise ratios than others, their implementation is largely a function of how much is

known about the expected nature of $C\epsilon(t)$. For present purposes it is sufficient to restrict attention to a simple first-order smoother of the form:

$$\dot{\epsilon}''_i(t) = -\frac{1}{\tau_s}\epsilon'_i(t) \quad i = 1...m \tag{5.3.1}$$

where τ_s is approximately the length of the smoothing interval.

The variance of $\epsilon''(t)$ can be used to define a 99% (3σ) likelihood error ellipse centered at the origin in $\mathcal{E}^m$. (If the variance of $\epsilon''(t)$ is not steady state, the ellipse is a function of time.) In the absence of any failures, $\epsilon''(t)$ is contained in the ellipse 99% of the time. As long as $\epsilon''(t)$ is contained in the ellipse, its magnitude can reasonable be attributed to the effect of the noise sources. If $\epsilon''(t)$ is outside the ellipse it is reasonable to assume that the system has failed. The remaining problem is to associate the direction $\epsilon''(t)$ with the failures which the failure detection system is designed to detect.

Procedures for determining the direction of a vector which is the sum of a deterministic component and a random component can become complicated. The simplest for $\epsilon'_i(t)$ is to transform it into the coordinate system defined by $Cf_1...Cf_r$ (i.e. the base normal canonical form) and define a separate threshold criterion for each of the error components in the new coordinate system (see Figure 5.1). Although it is immaterial whether the transformation is made before or after the smoothing operation, it is notationally convenient to assume that it occurs before:

$$\hat{\epsilon}'(t) = T_m^{-1}\epsilon'(t)$$

$$\dot{\hat{\epsilon}}''_i(t) = -\frac{1}{\tau_s}\hat{\epsilon}'_i(t) \quad i = 1...m \tag{5.3.2}$$

A threshold crossing for $\hat{\epsilon}''_1(t)$ then indicates that the failure associated with the output direction Cf_1 has occurred, etc.

Equation (5.2.2) defines an additional state for each of the m subsystems and consequently state space augmentations are necessary in order to determine the variances of $\hat{\epsilon}''_1(t)...\hat{\epsilon}''_m(t)$. Since

(5.1.13) and (5.3.2) together imply that the residual components satisfy:

$$\dot{\hat{\epsilon}}_i''(t) = -\frac{1}{\tau_s}\,\hat{c}_i\,\hat{\epsilon}_i(t) + \hat{G}_{2i}\,\eta_2(t) \quad i=1\ldots m \qquad (5.3.3)$$

the necessary augmentations are to add each of $\hat{\epsilon}_1''(t) \ldots \hat{\epsilon}''_m(t)$ to the corresponding state estimation error vector:

$$\hat{\hat{\epsilon}}_i(t) = \begin{bmatrix} \hat{\epsilon}_i(t) \\ \hline \hat{\epsilon}_i''(t) \end{bmatrix} \qquad i = 1\ldots m \qquad (5.3.4)$$

In order to write differential equations for $\hat{\hat{\epsilon}}_1(t) \ldots \hat{\hat{\epsilon}}_m(t)$ and their variances, it is also necessary to define:

$$\hat{\hat{A}}_{ii} = \left[\begin{array}{c|c} \hat{A}'_{ii} - d_{\psi i}\hat{c}_i & 0 \\ \hline -\frac{1}{\tau_s}\,\hat{c}_i & 0 \end{array}\right] \quad i = 1\ldots m \qquad (5.3.5)$$

$$\hat{\hat{d}}_{\psi i} = \begin{bmatrix} d_{\psi i} \\ \hline 0 \end{bmatrix}$$

$$\eta''_{1i}(t) = \begin{bmatrix} \eta'_{1i}(t) \\ \hline \hat{G}_{2i}\,\eta_2(t) \end{bmatrix}$$

The manner in which $\eta'_1(t)$ was defined insures that the covariances of $\eta''_{11}(t) \ldots \eta''_{1m}(t)$ are block diagonal:

$$E[\,\eta''_{1i}(t)\,\eta''_{1i}(\tau)^T\,] = \delta(t-\tau)\left[\begin{array}{c|c} \hat{G}'_{ii} & 0 \\ \hline 0 & \hat{G}_{2i}\,\hat{G}_{2i}^T \end{array}\right] = \hat{\hat{G}}'_{ii}\,\delta(t-\tau) \qquad (5.3.6)$$

With (5.3.5) and (5.3.6), the differential equations for the smoothed residuals satisfy:

$$\dot{\hat{\hat{\epsilon}}}_i(t) = \hat{\hat{A}}_{ii}\,\hat{\hat{\epsilon}}_i(t) + \eta''_{1i}(t) - \hat{\hat{d}}_{\psi i}\,\hat{G}_{2i}\,\eta_2(t)$$
$$\hat{\epsilon}''_i(t) = e^T_{\nu_i+1,\,\nu_i+1}\,\hat{\hat{\epsilon}}_i(t) \qquad (5.3.7)$$

The covariances of $\hat{\hat{\epsilon}}_1(t) \ldots \hat{\hat{\epsilon}}_m(t)$ satisfy:

$$\dot{\hat{\hat{P}}}_i(t) = \hat{\hat{A}}_i\,\hat{\hat{P}}_i(t) + \hat{\hat{P}}_i\,\hat{\hat{A}}_i^T + \hat{\hat{G}}_{ii}' + \hat{\hat{d}}_{\psi i}\,\hat{G}_{2i}\,\hat{G}_{2i}^T\,\hat{\hat{d}}_{\psi i}^T \qquad (5.3.8)$$

Finally the variances for the residual components are the last diagonal elements of the corresponding covariance matrices:

$$E[\hat{\epsilon}''_i(t)^2] = e^T_{\nu_i+1,\,\nu_i+1}\,\hat{\hat{P}}_i(t)\,e_{\nu_i+1,\nu_i+1} = [\hat{\sigma}''_i(t)]^2 \qquad (5.3.9)$$

If any of $\hat{\epsilon}''_1(t) \ldots \hat{\epsilon}''_m(t)$ satisfies:

$$\hat{\epsilon}''_i > 3\,\hat{\sigma}''_i(t) \qquad i = 1 \ldots m \qquad (5.3.10)$$

the threshold crossing infers that the failure associated with that residual component has failed.

Identification of the planar output error generated by a sensor failure is more difficult. The output error plane associated with a failure in the i^{th} sensor is spanned by Cd_i and e_{mi} (Section 4.1). The threshold crossing law (5.3.10) determines the components of $\hat{\epsilon}''(t)$ along Cd_i and e_{mi} and then checks to determine whether or not either component exceeds the 3σ threshold in that direction. It is possible for $\hat{\epsilon}''(t)$ to be contained in the plane defined by Cd_i and e_{mi} and to be outside the 3σ ellipse, but for its components along Cd_i and e_{mi} to be inside the ellipse (if the magnitude of $\epsilon''(t)$ is greater than $4.25\,\sigma$ one of its components must be outside the ellipse). If that sensor has failed, however, it is reasonable to expect that the rotation of $\hat{\epsilon}''(t)$ in the plane would cause both of the thresholds to be exceeded at various times. Nevertheless, when applying the threshold crossing law for sensor failures, a conservative approach is to assume that at least a 4σ failure-induced error is required in order to insure a crossing of the 3σ thresholds in (5.3.10)

5.4 Output Estimation Error Analysis

The threshold detection law developed in Section 5.3 is predicated on the assumption that a failure detection system has a large error in its state estimate if the system which it is monitoring has failed. There are a number of applications for which this assumption is not valid. The purpose of this section is to formulate the general problem of detecting a signal imbedded in noise and to review techniques commonly used in its solution. For convenience, attention is restricted to the case of a failure detection system for m mutually detectable failures; however, the algorithms of Section 5.1.2 allow the techniques discussed in this section to be applied for any failure detection system for which $r \leq m$.

One of the important contributions of failure detection theory is that it permits the decoupling of the error equations for a state estimator (2.3.2) to a set of m scalar output subsystems. When the residual smoothing introduced in Section 5.3 is employed, these subsystems are described by (5.3.7). If only one failure detection system is used, each subsystem contains all the information relevant to the occurrence of a particular failure. The problem of interpreting that information can be formulated as a set of m independent hypothesis tests. The hypothesis tests which require the least amount of a priori knowledge of failure modes are the unary tests on the null hypotheses $H_{01} \ldots H_{0m}$. The null hypotheses are that the smoothed residual for the subsystems are generated solely by the stochastic sources modeled in those subsystems (5.3.7). For zero initial conditions, they are denoted by:

$$H_{0i}: \hat{\epsilon}_i''(t) = \hat{e}^T_{\nu_i+1,\,\nu_i+1} \int_0^t e^{\hat{\hat{A}}_{ii}(t-\tau)} (\eta_{1i}''(\tau) - \hat{\hat{d}}_{\psi i} \hat{G}_{2i} \eta_2(\tau)) d\tau \quad i = 1 \ldots m \tag{5.4.1}$$

The null hypotheses state that the residual components are zero mean Gaussian random processes with variances $[\hat{\sigma}_1'']^2 \ldots [\hat{\sigma}_m'']^2$ respectively. The hypothesis tests can be conditioned on any sufficient statistic of such processes, e.g. the threshold test in

Section 5.3 utilizes the fact that at any given time there is only a 1% probability that the residual for a subsystem exceeds its 3σ threshold. Null hypothesis tests conditioned on sample means and sample variances are outlined in the following paragraphs. To the extent that failure mode knowledge of a general nature is available, a general indication of the types of output errors the various hypothesis tests are most sensitive to is given.

If $\hat{D}_K$ is an optimal detection filter-state estimator, the unsmoothed residuals $\hat{\epsilon}'_1(t) \ldots \hat{\epsilon}'_m(t)$ corresponding to the null hypotheses are white noise. The only time correlation in the smoothed subsystem residuals is that imposed by the smoothing operation. If the sample means for the subsystem residuals are computed from a set of ℓ data points:

$$\bar{\hat{m}}''_i(t, \ell, \Delta t) = \frac{1}{\ell} \sum_{k=0}^{\ell-1} \hat{\epsilon}''_i(t - k\Delta t) \qquad i = 1 \ldots m \qquad (5.4.2)$$

then the sample means are random variables. If $\Delta t >> \tau_s$ (recall that τ_s can be chosen arbitrarily small), the null hypotheses state that the data points in (5.4.2) are independent and that the sample means are zero mean Gaussian random variables with variances:

$$E\left[\bar{\hat{m}}''_i(t, \ell, \Delta t)^2\right] = \left[\frac{\hat{\sigma}''_i}{\ell}\right]^2 \qquad (5.4.3)$$

Threshold-type hypothesis tests can be conditioned on (5.4.2) and (5.4.3). Such tests are most sensitive to failures which impart a bias to the residuals or which given them a correlation time longer than $\ell\,\Delta t$. The sample mean does not provide a good test for a particular subsystem if the failures related to that subsystem generate residuals with a correlation time significantly shorter than $\ell\,\Delta t$, even if the residuals have large magnitudes.

The sample variance can be defined in a manner similar to (5.4.2):

$$\left[\bar{\hat{\sigma}}''_i(t, \ell, \Delta t)\right]^2 = \frac{1}{\ell} \sum_{k=0}^{\ell-1} \left[\hat{\epsilon}''_i(t - k\Delta t)\right]^2 \quad i = 1 \ldots m \qquad (5.4.4)$$

The sample variances are random variables. If $\Delta t >> \tau_s$, the null hypotheses state that the sample variances have χ^2-distributions with $\ell - 1$ degrees of freedom. The probability distribution for χ^2-distributed random variables can be used to define threshold-type hypothesis tests for the sample variances. For a particular subsystem, a sample variance test is most likely to be superior to the tests mentioned previously if the failure mode generates residuals with small magnitudes and short correlation times.

There are a variety of other statistics on which null hypothesis tests can be conditioned -- some of these are discussed by Mehra and Peschon (1971). The usefulness of the various criteria is a function of the particular application and the possible failure modes. If there is reason to believe that for a particular subsystem no single criterion can satisfactorily detect all possible failure modes, it may be desirable to use several of the criteria concurrently for the null hypothesis test on that subsystem.

The null hypothesis test is characterized by the fact that it requires virtually no information about possible failure modes. If the failure mode for a particular subsystem is known, this information can be incorporated into the detection law by defining a second hypothesis which describes the system in the failed state. The decision as to which hypothesis to accept is then based upon the evaluation of a likelihood function. The general approach for implementing this binary hypothesis test is detailed by Van Trees (1968). The literature on the subject is extensive and cannot be reviewed here, but it is worthwhile to relate the applicability of some of the more important results to failure detection theory.

In the failed state, the i^{th} subsystem of (5.1.13) is described by:

$$\dot{\hat{\epsilon}}(t) = (\hat{A}' - d_{\psi i}\,\hat{c}_i)\,\hat{\epsilon}_i(t) + \eta'_{1i}(t) - d_{\psi i}\,\hat{G}_{2i}\,\eta_2(t) + \hat{f}_{ii}\,n_i(t)$$

$$\hat{\epsilon}'(t) = \hat{c}_i\hat{\epsilon}(t) + \hat{G}_{2i}\,\eta_2(t) \quad : \quad \epsilon(0) \text{ given} \qquad (5.4.5)$$

Because the evaluation of the likelihood function provides the necessary smoothing of $\epsilon'(t)$, the two hypotheses can be formulated in terms of the solutions of (5.1.13) and (5.4.5):

$$H_0: \ \hat{\epsilon}_i'(t) = \eta_\epsilon(t)$$

$$H_1: \ \hat{\epsilon}_i'(t) = \hat{c}_i \int_0^t e^{\hat{A}_{ii}(t-\tau)} \hat{f}_{ii} \, n_i(\tau) \, d\tau + \eta_\epsilon(t) \qquad (5.4.6)$$

where:

$$\hat{A}_{ii} = \hat{A}'_{ii} - d_{\psi i} \hat{c}_i$$

and $\eta_\epsilon(t)$ is the solution to (5.1.13) with zero initial conditions:

$$\eta_\epsilon(t) = \hat{c}_i \int_0^t e^{\hat{A}_{ii}(t-\tau)} [\eta'_{1i}(\tau) - d_{\psi i} \hat{G}_{2i} \eta_2(\tau)] \, d\tau + \hat{G}_{2i} \eta_2(t) \qquad (5.4.7)$$

If the likelihood function for the two hypothesis is denoted by $\lambda(t)$, the hypothesis test is based upon the threshold η:

$$\lambda(t) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \qquad (5.4.8)$$

If $\lambda(t)$ is greater than η, the hypothesis that the failure described by H_1 has occurred is accepted. Otherwise the hypothesis H_0 that no failure has occurred is accepted.

The amount of processing required to determine $\lambda(t)$ is dependent upon the characteristics of $n_i(t)$ and $\eta_\epsilon(t)$. The simplest algorithms result if $\eta_\epsilon(t)$ is a white random process. Since $\eta_\epsilon(t)$ is white if $\hat{D}$ is required to be an optimal detection filter-state estimator, it is convenient to restrict attention to this case. Procedures for determing $\lambda(t)$ when $\eta_\epsilon(t)$ is white and $n_i(t)$ has one of two special forms are reviewed in the remainder of this section. More general algorithms are presented in the literature.

The classic signal detection problem (Van Trees (1968)) assumes that $n_i(t)$ is a known waveform and that the integral in (5.4.6) is non-zero only in the interval $t_0 \le t \le T$. In this case the hypothesis test is made at time T and $\lambda(T)$ is the output of a matched filter. With $w_i(t)$ defined by:

$$w_i(t) = \hat{c}_i \int_{t_0}^{t} e^{\hat{A}_{ii}(t-\tau)} \hat{f}_{ii}\, n_i(\tau)\, d\tau \qquad (5.4.9)$$

the likelihood function is:

$$\lambda(T) = \int_{t_0}^{T} w_i(\tau)\, \hat{\epsilon}_i'(\tau)\, d\tau \qquad (5.4.10)$$

Computation of $\lambda(T)$ can be done in several different fashions. The algorithm illustrated in Figure 5.2a is consistent with (5.4.10) and is frequently referred to as a correlation receiver. Since $\epsilon_i'(t)$ is the output of a ν_i-dimensional subsystem, in general an additional ν_i-dimensional subsystem is required to compute $w_i(t)$.

An important generalization of the signal detection problem is the case where $n_i(t)$, and hence $w_i(t)$, is a Gaussian random process. In the context of failure detection theory, $n_i(t)$ could be white noise, it could be an impulse of magnitude ξ at t_0:

$$n_i(t) = \xi\, \delta(t-t_0) \qquad (5.4.11)$$

where ξ is a Gaussian random variable, etc. The algorithm for $\lambda(t)$ was first derived by Schweppe (1965) and is summarized in Van Trees (1970). It requires that $\hat{\epsilon}_i'(t)$ be passed through a Kalman filter to estimate $w_i(t)$. The likelihood function then satisfies:

$$\lambda(T) = \int_{t_0}^{T} \left[2\, \hat{\epsilon}_i'^{T}(\tau)\, \hat{G}_\epsilon^{-1}\, \hat{w}_i(\tau) - \hat{w}_i(\tau)^T\, \hat{G}_\epsilon^{-1}\, \hat{w}_i(\tau) \right] d\tau \qquad (5.4.12)$$

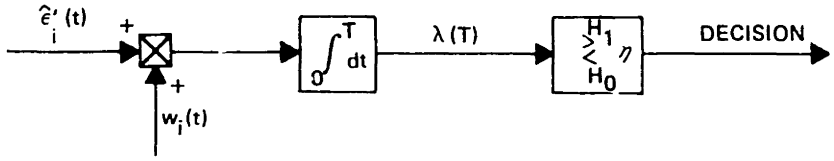


Figure 5.2a Correlation receiver to detect presence of known signal in detection system residuals

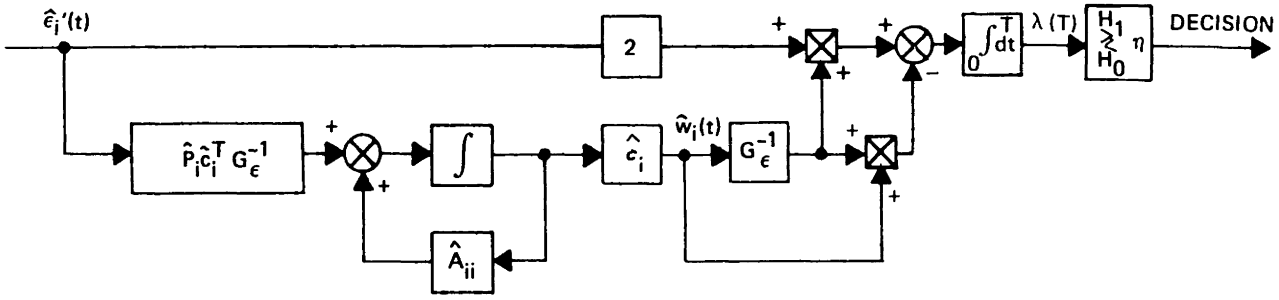


Figure 5.2b Correlation receiver to detect presence of gaussian signal in detection system residuals (assumes optimal detection filter-state estimator)

where $\hat{G}_\epsilon$ is the variance of $\eta_\epsilon(t)$ and $\hat{w}_i(t)$ is the Kalman filter estimate of $w_i(t)$. This implementation is indicated in Figure 5.2b. Even though the Kalman filter is ν_i-dimensional, the computational requirements are greater than in the deterministic case because it is also necessary to compute the Kalman gains.

In addition to the computational requirements, the two binary test algorithms presented in this section have three major disadvantages:

i) Their performance is uncertain if a failure mode other than that described by H_1 occurs,

ii) the statistics of the hypothesis test are well-defined only $T-t_0$ seconds after the failure occurs, and

iii) the statistics for the hypothesis test are often difficult to compute.

The first difficulty is important only if the output error trajectories corresponding to the various failure modes are significantly different. A number of approximate techniques have been developed to overcome the last two difficulties. The generalized likelihood ratio test treated by McAulay and Denlinger (1973) and also by Willsky and Jones (1974) eliminates these difficulties for a particular class of problems.

CHAPTER 6

FAILURE DETECTION FOR LATERAL MODE AUTOPILOT

Chapter 4 and Chapter 5 summarize the theoretical aspects of failure detection theory. The algorithms presented in those chapters permit a failure detection system to be designed for any given linear stochastic system. A question which has not been addressed concerns the effectiveness of the resultant failure detection system. No general answer to this question exists. For a wide range of applications the theory provides a capability for real-time failure detection where none had previously existed. For other applications, use of the theory without supplemental error signal analysis cannot be expected to completely satisfy the failure detection requirements. The purpose of this chapter is to use a particular example to demonstrate the validity of the theory and investigate the circumstances under which it can be expected to yield useful failure detection information.

The example used in this chapter is a failure detection system for the roll mode autopilot in an aircraft. The simplest mathematical model for the aircraft-autopilot combination is the closed loop spinning cylinder example of Sections 3.3.2 and 4.1.2. A stochastic version of this model is sufficiently complicated to exercise the algorithms of Chapters 4 and 5 and is used in Section 6.2 to design a pair of optimal detection filter-state estimators for a prescribed set of failures. The detection law necessary to interpret the output of the two failure detection systems is developed in Section 6.3. The performance of the resulting failure detection systems when the nominal aircraft response is also described by the simplified dynamics is analyzed in Section 6.4.1. When the aircraft response satisfies the full lateral equations of motion, the modeling error in the reference model provides a source of state estimation errors in addition to those associated with failures and stochastic sources. The sensitivity of

the failure detection systems to the modeling errors and procedures for reducing that sensitivity are discussed in Section 6.4.3.

The emphasis in this chapter is upon the detection filter as an optimal detection filter-state estimator. The failure detection systems designed in this chapter are compared with a Kalman filter in Section 6.4.2. The aircraft-autopilot combination provides several examples of instances in which the objectives of good failure detection and good state estimation can conflict. These examples are discussed in Section 6.4.3 and nonoptimal detection filters which increase the detectability of the failures are presented.

6.1 Detection Filter Design - Deterministic

The lateral equations of motion for an aircraft consist of three coupled equations for roll, yaw, and sideslip angle. The effective control surfaces are the ailerons and the rudder. The dynamics of the lateral mode are fourth-order and can be roughly categorized as:

1) The dutch roll mode which is a second-order oscillation coupling yaw, roll, and sideslip. Characteristically, the dutch roll mode is stable and lightly damped with a period of several seconds.

2) The spiral divergence mode which is a coupling of yaw angle and roll angle. The mode may be stable or unstable with a slow divergence. If it is unstable, a roll angle results in a broad spirally dive.

3) The roll subsidence mode which is first-order and is the roll response of the aircraft to an aileron input.

A standard design procedure for lateral mode autopilots (see Blakelock (1968)) is to utilize the rudder to:

i) damp the dutch roll mode,

ii) make the spiral divergence mode stable and negligible, and

iii) minimize sideslip in turns.

The rudder control loop is the inner loop of the autopilot and performs all of the functions required of a lateral mode autopilot except heading

control and roll angle control (Figure 6.1). The ailerons are then used in a second (middle) loop of the autopilot to control roll angle. Heading angle is controlled by defining a third (outer) loop to command the roll sequence necessary to maintain heading.

The dynamics of the rudder control system-aircraft combination appear much simpler than those of the aircraft alone. With the rudder automatically damping sideslip and dutch roll, the dominant response to an aileron input is the same second-order roll response exhibited by the spinning cylinder of Sections 3.3.2 and 4.1.2. If the rudder and the inner loop autopilot are operating nominally, the roll response of the aircraft can reasonably be approximated by this second-order model. The failure detection systems designed in this section assume that no failures occur in the inner loop and that the second-order model is valid.

6.1.1 Second-order Roll Response Model

The basic vehicle configuration for the simplified roll dynamics is indicated in Figure 6.2. The autopilot consists of roll rate and roll angle feedback with the appropriate gains. As an illustration of parallel redundancy, the ailerons are assumed to be segmented into inboard and outboard sections. The inboard ailerons respond independently of the outboard ailerons and vice-versa. Either set of control surfaces is capable of controlling the roll mode dynamics. The measured quantities for the system are:

i) rate gyro output
ii) attitude gyro output
iii) inboard aileron deflection
iv) outboard aileron deflection

The aileron deflections are open loop in that the servos do not use the deflection measurements to position the ailerons. The measurements are used strictly for failure analysis. This open loop feature is not typical for autopilots, but it is useful for contrasting the difference between detecting failures in sensors whose output is incorporated into the control law and those whose output is not.

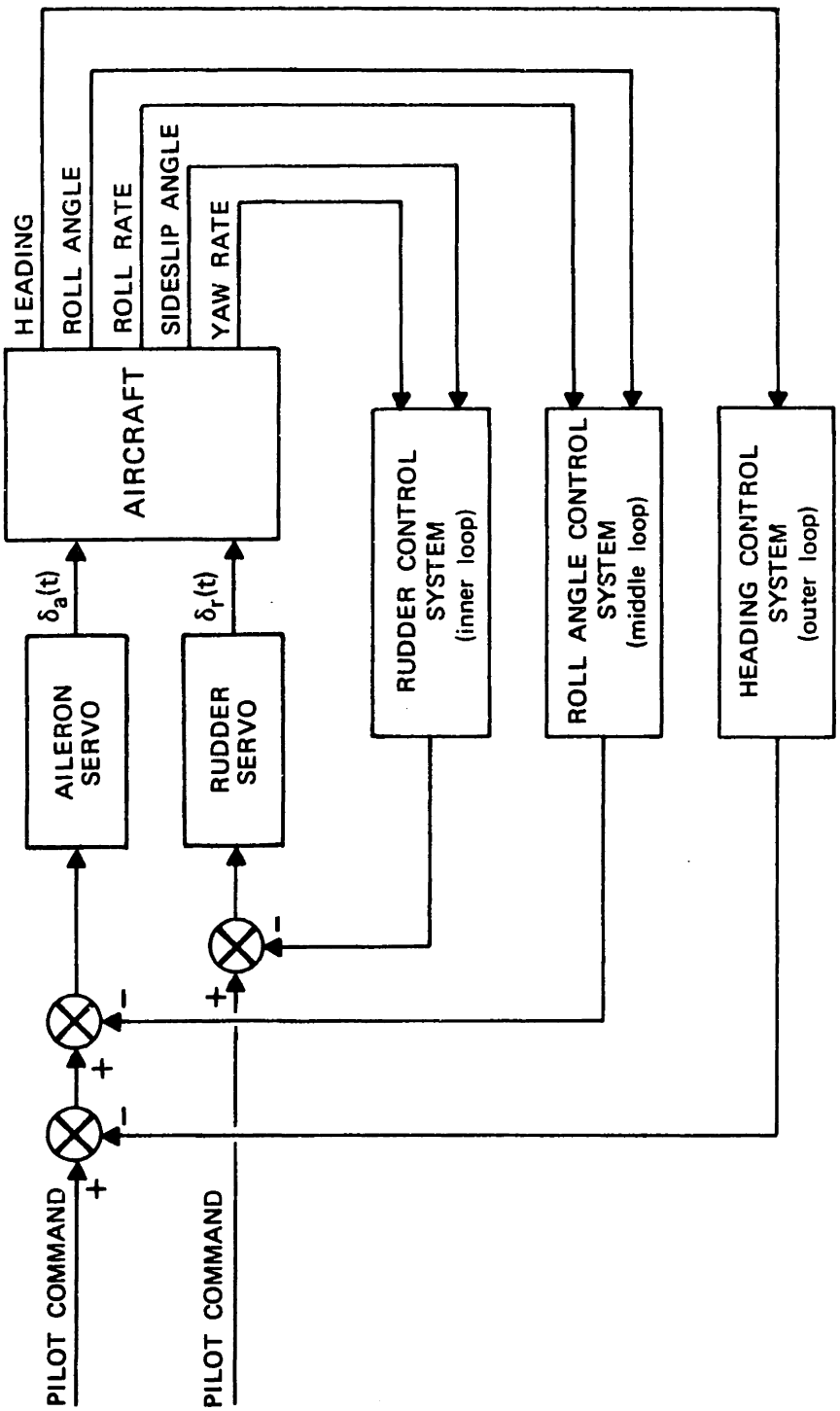
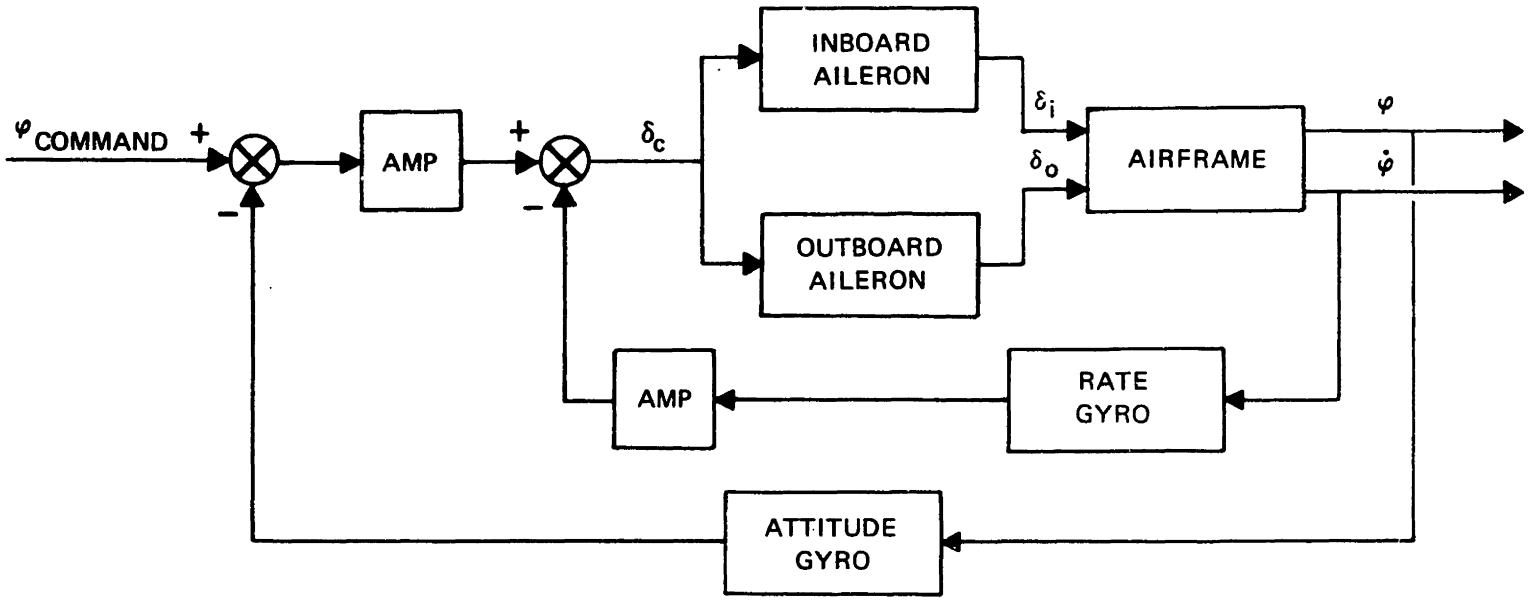


Figure 6.1 Feedback structure of a typical lateral mode autopilot



MEASURED QUANTITIES

1) RATE GYRO OUTPUT
2) ATTITUDE GYRO OUTPUT
3) INBOARD AILERON DEFLECTION
4) OUTBOARD AILERON DEFLECTION

Figure 6.2 Simplified roll response model of autopilot and aircraft

The vehicle described in Figure 6.2 can suffer a failure in any of the system components and survive if that failure can be detected and if the failed component can be removed from service. If one set of ailerons fails, the remaining set can control the vehicle unless the failed ailerons remain fully deflected. If either gyro fails, it can be replaced by a standby or the pilot can assume control. The important feature of the system is that failure detection and isolation is necessary in order to maintain system performance and that there is not sufficient redundancy for a majority rule voting procedure to be used. Voting could be used only if there were three sets of ailerons with identical dynamics--but even then it could not discriminate between an aileron failure and a sensor failure. Failure detection theory eliminates these restrictions as well as the ambiguity.

The reference model dynamics are those of Section 4.1.2. The second-order differential equation approximating the aircraft response to aileron deflections $\delta_{ai}(t)$ and $\delta_{ao}(t)$ is:

$$\ddot{\varphi}(t) + \beta_{\varphi} \dot{\varphi}(t) = \alpha_i \delta_{ai}(t) + \alpha_o \delta_{ao}(t) \qquad (6.1.1)$$

where $\varphi(t)$ is the roll angle, $\frac{1}{\beta_{\varphi}}$ is the roll time constant, and α_i and α_o are the aileron effectivenesses. The aileron servomechanisms are assumed to be first-order:

$$\dot{\delta}_{ai}(t) + \tau_i \delta_{ai}(t) = k_i \delta_c(t)$$

$$\delta_{ao}(t) + \tau_o \delta_{ao}(t) = k_o \delta_c(t) \qquad (6.1.2)$$

where $\delta_c(t)$ is the deflection command supplied from the autopilot. The time constants $\frac{1}{\tau_i}$, $\frac{1}{\tau_o}$ and the proportionality constants k_i, k_o may be the same for the two ailerons or they may be different. If the time constants are different, the aileron deflection measurements cannot be compared directly to determine if their outputs are consistent--but failure detection theory does not require that they be compared. The control law is an algebraic function of the measure roll angle and roll rate:

$$\delta_c(t) = k_a(\varphi_c(t) - \varphi_a(t)) - k_r \dot{\varphi}_r(t) \qquad (6.1.3)$$

where $\varphi_c(t)$ is the commanded roll angle from the outer loop of the autopilot and $\varphi_a(t)$, $\dot{\varphi}_r(t)$ are the roll angle and roll rate measurements.

In the unfailed system, $\varphi_a(t)$ and $\varphi(t)$ are identical and so are $\dot{\varphi}_r(t)$ and $\dot{\varphi}(t)$. This raises the possibility of modeling the autopilot-aircraft system as a fourth-order system with $\varphi_a(t)$ (or $\varphi(t)$), $\dot{\varphi}_r(t)$ (or $\dot{\varphi}(t)$),$\delta_i(t)$ and $\delta_o(t)$ as the state vector components. The resulting system model would by fully measurable in the sense of Section 4.5 and detection filter design would appear to be a simple matter. Unfortunately, a fourth-order model does not allow rate gyro failures to be modeled properly. If the rate gyro fails, the model should reflect the fact that the output of the attitude gyro continues to be the true aircraft roll angle. On the other hand, the model should reflect the fact that a rate gyro failure alters the autopilot-aircraft dynamics significantly. These two requirements necessitate that both the measured roll rate and the true roll rate be components of the state vector:

$$x_m(t) = \begin{bmatrix} \varphi_a(t) \\ \dot{\varphi}(t) \\ \dot{\varphi}_r(t) \\ \delta_{ai}(t) \\ \delta_{ao}(t) \end{bmatrix} \qquad (6.1.4)$$

A reference model based upon (6.1.4) is observable but not controllable. The lack of controllability exists only because the control law cannot control $\dot{\varphi}(t)$ and $\dot{\varphi}_r(t)$ independently and does not imply that the aircraft is uncontrollable. The reference model based on (6.1.4) is:

$$C = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & -\beta_{\varphi} & 0 & \alpha_i & \alpha_o \\ 0 & -\beta_{\varphi} & 0 & \alpha_i & \alpha_o \\ -k_i k_a & 0 & -k_i k_r & -\tau_i & 0 \\ -k_o k_a & 0 & -k_o k_r & 0 & -\tau_o \end{bmatrix} \quad B = \begin{bmatrix} 0 \\ 0 \\ 0 \\ k_i k_a \\ k_o k_a \end{bmatrix}$$

(6.1.5)

Once the reference model is chosen, the next step is to define the failures to be detected and their event vectors. As would have been the case if a fourth order model had been chosen instead of (6.1.5) this may involve changing the model. The failures to be monitored by the failure detection systems designed in this section are:

i) attitude gyro,
ii) rate gyro,
iii) inboard aileron servo,
iv) outboard aileron servo,
v) inboard aileron deflection sensor,
vi) outboard aileron deflection sensor,
vii) a change in the roll dynamics of the aircraft, and
viii) either amplifier (k_a or k_r).

If a system has m sensors, failure detection theory does not guarantee that more than m failures can be detected by a single failure detection system. Since the autopilot-aircraft system has only four sensors, it is to be expected that more than one failure detection system is required to monitor all eight failures. Because failure detection theory is independent of the failure mode for a component, the event vector associated with each failure is independent of any particular model for the failed system. Nevertheless, in designing a failure detection system it is useful to work with a particular failure model to insure that the reference model adequately reflects the manner in which the failure affects the system. The failure models and the error equations developed in

the remainder of this subsection are applications of the general error models of Section 4.1.1.

Since the gyro outputs are used in the control law, a gyro failure model must indicate a change in the value of the appropriate state variable, $\varphi_a(t)$ or $\dot{\varphi}_r(t)$. If the failure mode for an attitude gyro failure consists of a bias of magnitude s_a at time t_o, the failure model is:

$$\dot{x}(t) = Ax(t) + Bu(t) + s_a f_a \delta(t-t_o) : x(0) \text{ given} \qquad (6.1.6)$$

$$y(t) = Cx(t)$$

where $\delta(t-t_o)$ is the unit impulse and f_a is the event vector:

$$f_a = e_{51} \qquad (6.1.7)$$

$$Cf_a = e_{41}$$

Equation (6.1.6) is in the form of the controller failure model developed in Section 4.1.1. Cf_a is the output direction associated with the attitude gyro bias. The effect of the bias on the state estimates generated by the failure detection system is described by the corresponding error equations:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + e_{51} s_a \delta(t-t_o) : \epsilon(0) \text{ given} \qquad (6.1.8)$$

$$\epsilon'(t) = C\epsilon(t)$$

The error generated by the failure first appears in the state component associated with e_{51}, i.e., in the first component of the state vector. This is the estimate of $\varphi_a(t)$. If D is a detection filter for an attitude gyro failure, the estimates of the remaining components of the state vector are not affected by the failure.

The attitude gyro may possess failure modes other than a bias, but the corresponding failure model must have the same general form as (6.1.6) and the event vector must be f_a as defined in (6.1.7). For example, a total failure of the attitude gyro can be modeled by the triad $(A + \Delta A_a, B, C)$ where:

$$\Delta A_a = \begin{bmatrix} \Delta a_{11} & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (6.1.9)$$

The -1 element in ΔA_a means that $\varphi_a(t)$ for the failed system $(A + \Delta A_a, B, C)$ is no longer related to the true aircraft attitude; $\varphi_a(t)$ is dependent solely upon its value at the time of the failure, the new time constant $\frac{1}{\Delta a_{11}}$, and whatever noise sources continue to be present. Because the only non-zero elements in ΔA_a are in the first row, application of the algorithms developed in Section 4.1.1 yields a failure model similar to (6.1.6):

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) + e_{51}\,(\Delta a_{11}\,\varphi_a(t) - \dot{\varphi}(t)): \quad x(0) \text{ given} \\ y(t) &= Cx(t) \end{aligned} \qquad (6.1.10)$$

Equation (6.1.10) supports the assertion made in Section 4.1.1 that frequently failure models which involve a change in A can be reduced to the more manageable controller failure model.

A line of reasoning similar to that which lead to (6.1.9) allows a complete failure of the rate gyro to be modeled by the triad $(A + \Delta A_r, B, C)$ where:

$$\Delta A_r = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & \beta_\varphi & \Delta a_{33} & -\alpha_i & -\alpha_o \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (6.1.11)$$

The event vector corresponding to this failure model is the third unit vector:

$$f_r = e_{53}$$

$$Cf_r = e_{42} \qquad (6.1.12)$$

The failure model is similar to (6.1.10) in its dependence upon the vehicle state. The estimation error equations follow in a straightforward manner from the failure model:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + e_{53}\,(\Delta a_{33}\dot{\varphi}_r(t) + \beta_{\varphi}\dot{\varphi}(t) - \alpha_i\delta_{ai}(t) - \alpha_o\,\delta_{ao}(t)):$$

$$\epsilon\,(0) \text{ given} \qquad (6.1.13)$$

$$\epsilon'(t) = C\,\epsilon(t)$$

As was the case with the attitude gyro, failure models (such as a gyro bias) exist for which the estimation error may not be a function of the state variables -- but f_r is the event vector for all possible failure models.

The gyro failure models can be contrasted to failures in sensors whose output is not used in the control law by considering failures in the aileron deflection sensors. These failures do not affect the state of the aircraft -- only the failure detection system's knowledge of that state. A reasonable failure model then includes changes only in the measurement matrix (A, B, C + ΔC). If the sensors are modeled as having failed completely, ΔC for the inboard and outboard aileron deflection sensors respectively is:

$$\Delta C_{si} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \quad \Delta C_{so} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix} \qquad (6.1.14)$$

Although event vectors do not exist per se for sensor failures, ΔC_{si} and ΔC_{so} satisfy:

$$\Delta C_{si}\, x(t) = -e_{43}\, \delta_{ai}(t) \qquad (6.1.15)$$

$$\Delta C_{so}\, x(t) = -e_{44}\, \delta_{ao}(t)$$

and the output directions e_{43} and e_{44} are sufficient to characterize the failures. The general estimation error equations for a state-dependent sensor failure can again be determined from the general failure model:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) - D\Delta C\, x(t) \;:\; \epsilon(0) \text{ given} \qquad (6.1.16)$$

$$\epsilon'(t) = C\,\epsilon(t) + \Delta C\, x(t)$$

Equation (6.1.16) is in the form of the sensor failure model developed in Section 4.1.1. The $D\Delta C\, x(t)$ term appears in (6.1.16) because the detection filter-state estimator D uses the output feedback for failure isolation and state estimation. If a sensor fails, the erroneous measurement induces erroneous state estimates. In general this means that the output estimation error generated by a sensor failure cannot be constrained to a single direction. For an inboard aileron deflection sensor failure, (6.1.16) becomes:

$$\dot{\epsilon}(t) = (A-DC)\,\epsilon(t) + d_3\, \delta_{ai}(t) \qquad (6.1.17)$$

$$\epsilon'(t) = C\,\epsilon(t) - e_{43}\, \delta_{ai}(t)$$

where d_3 is the third column of D. For failure detection system design it is desirable to replace (6.1.17) with a set of error equations which do not involve columns of D, but this is best done when the actual design is considered in Section 6.1.2.

The estimation error equations for the aileron deflection sensor failures can be compared to the error equations if the ailerons themselves fail. If a failed aileron is assumed to return to the trim

position and remain there, a simple failure model can be developed. Because the control law (6.1.3) requires the aileron to respond to both the measured roll angle and roll rate as well as an externally supplied roll angle command, the failure model must include a change in dynamics and a change in the control matrix. A suitable model for the inboard aileron is:

$$\Delta A_i = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ k_i k_a & 0 & k_i k_r & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Delta B_i = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -k_i k_a \\ 0 \end{bmatrix} \qquad (6.1.18)$$

The inboard aileron for the system $(A + \Delta A_i, B + \Delta B_i, C)$ is undriven and any initial condition at the time of the failure decays to the trim position at a rate determined by the time constant of the aileron servo. The general error models of Section 4.1.1 indicate that the event vector associated with both ΔA_i and ΔB_i is e_{54}:

$$f_i = e_{54} \qquad (6.1.19)$$

$$Cf_i = e_{43}$$

The estimation error equations can then be written:

$$\dot{\epsilon}(t) = (A\text{-}DC)\,\epsilon(t) + e_{54}\,(k_i k_a\,(\varphi_a(t) - \varphi_c(t)) + k_i k_r\,\dot{\varphi}_r(t)):\ \epsilon(0) \text{ given}$$

$$\epsilon'(t) = C\,\epsilon(t) \qquad (6.1.20)$$

The output error direction with an inboard aileron failure is the same direction as the additive output error term $e_{43}\,\delta_{ai}(t)$ if the corresponding sensor fails; however, the presence of a single event vector in (6.1.20) is sufficient for a failure detection system to distinguish between the two failures. Similar equations to (6.1.17) - (6.1.20) can be developed for the outboard aileron and its sensor.

The amplifiers are characterized by their gains k_r and k_a. If either amplifier fails, the control law no longer functions properly. Because the control law command $\delta_c(t)$ is not directly measured,[1] the failure is first noted as an unexpected increase in the residuals of the inboard and the outboard aileron deflection estimates. The fact that both aileron deflection estimates are affected in dynamically related fashions is sufficient to allow the failure detection system to determine that the control law has failed. A failure in the attitude gyro amplifier can be modeled as a change in the gain k_a. If k_a becomes zero, the failure model requires changes in A and B:

$$\Delta A_{ka} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ k_i k_a & 0 & 0 & 0 & 0 \\ k_o k_a & 0 & 0 & 0 & 0 \end{bmatrix} \qquad \Delta B_{ka} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -k_i k_a \\ -k_o k_a \end{bmatrix} \qquad (6.1.21)$$

[1] In practical applications, $\delta_c(t)$ would generally be measured. A change in $\delta_c(t)$ which was not accompanied by corresponding changes in $\varphi_a(t)$ and $\dot{\varphi}_r(t)$ would imply that either an amplifier had failed or the control law had changed. In order to illustrate one aspect of failure detection theory, however, it is assumed that $\delta_c(t)$ is not available.

Because ΔA_{ka} has two non-zero rows, it would in general require two event vectors. Since the rows are linear combinations of each other, however, and:

$$\Delta A_{ka}\, x(t) = \begin{bmatrix} 0 \\ 0 \\ 0 \\ k_i \\ k_o \end{bmatrix} k_a \varphi_a(t) \qquad (6.1.22)$$

a single event vector is adequate:

$$f_{ka} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ k_i \\ k_o \end{bmatrix} \qquad Cf_{ka} = \begin{bmatrix} 0 \\ 0 \\ k_i \\ k_o \end{bmatrix} \qquad (6.1.23)$$

The vector f_{ka} is also the event vector associated with ΔB_{ka}. The resulting error equations are:

$$\dot{\epsilon}(t) = (A\text{-}DC)\,\epsilon(t) + k_a f_{ka}\,(\varphi_a(t) - \varphi_c(t)) \;:\; \epsilon(0) \text{ given} \qquad (6.1.24)$$

$$\epsilon'(t) = C\epsilon(t)$$

The error model if k_r is zero is similar to (6.1.21) except that no change in B is required. The event vector is also f_{ka}. As a consequence, a failure detection system cannot determine which amplifier failed on the basis of output error direction alone.

The last failure to be monitored is a change in aircraft roll response dynamics. The change may be due to a variation in the roll time constant or the fact that the dutch roll mode is not adequately surpressed. It is not so much a component failure as it is a failure

of the reference model to adequately predict the aerodynamic response of the aircraft to an aileron deflection. As such, it allows a check on the validity of the reference model as well as on the performance of the inner loop of the autopilot. For failure detection purposes, it is sufficient to assume that the failure is a change in the roll time constant β_φ:

$$\Delta A_\beta = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & \Delta\beta_\varphi & 0 & 0 & 0 \\ 0 & \Delta\beta_\varphi & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (6.1.25)$$

The event vector and output direction can be determined:

$$f_\beta = \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{bmatrix} = e_{52} + e_{53} \qquad Cf_\beta = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} = e_{42} \qquad (6.1.26)$$

The state estimation error is a function of $\Delta\beta_\varphi$ and the true vehicle roll rate:

$$\dot{\epsilon}(t) = (A\text{-}DC)\,\epsilon(t) + f_\beta\,\Delta\beta_\varphi\,\varphi(t) \;:\; \epsilon(0) \text{ given} \qquad (6.1.27)$$

$$\epsilon'(t) = C\,\epsilon(t)$$

A change in the aircraft dynamics causes an initial increase in the error of the estimate of $\dot{\varphi}_r(t)$, and a related error in the estimate of $\dot{\varphi}(t)$. The output error first appears as an increase in the residual associated with $\dot{\varphi}_r(t)$.

6.1.2 Failure Detection System Design--Deterministic

Design of the failure detection system can be accomplished in two stages. The first is a deterministic design stage and involves the selection of a set of failures to be monitored. The selection

is accomplished through consideration of the inter-relationships of the detection spaces associated with the failures. The second stage is concerned with the actual selection of the detection filter, i.e. the choice of D which assigns A-DC the desired eigenvalues. The criteria involved in selecting the desired eigenvalues are very much a function of the context in which the failure detection system is used, and in particular, the noise environment in which the system is to operate. For this reason, the second stage is referred to as the stochastic design stage. This subsection and Section 6.2.2 apply the design algorithms for the aircraft-autopilot example. The design procedure is intended to be complimentary to that of Section 4.2. Algorithms on which Section 4.2 focused are treated lightly in favor of concentration on subsequent algorithms.

The deterministic design procedure begins with the selection of a set of failures whose event vectors are output separable. The first step is then to determine the associated detection spaces (Section 4.2.1.) and investigate the mutual detectability of the failures (Section 4.2.2.). If they are mutually detectable, the next step in the procedure can be by-passed. If they are not mutually detectable, four options are available to the designer:

i) design a failure detection system for the set of failures and accept the resulting unassignable eigenvalues of A-DC,

ii) find a subset of the failures which are mutually detectable,

iii) select a higher dimensional reference model for which the set of failures is mutually detectable, or

iv) choose a different set of failures and re-initiate the design procedure.

Typically, trial and error is required to choose the best option (or options) for a particular application. Once a satisfactory set of output separable event vectors is chosen, the final step is to determine whether or not output stationarity can be used to add additional failures to the set.

The result of Section 6.1.1 and several additional results developed below are summarized in Table 6.1 for the eight failures of interest. A preliminary step of the design procedure is to replace the column of D in (6.1.17) (and the corresponding outboard aileron sensor equation) with a suitable event vector. The output direction associated with the inboard aileron deflection sensor is e_{43}. The algorithm of Section 4.2.1 suggests that the event vectors associated with the six non-sensor failures be inspected to determine if any satisfy:

$$e_{43} = Cf \tag{6.1.28}$$

f_i for an inboard aileron failure clearly does. The arguments of Section 4.1 then indicate that d_3 and e_{43} can be removed from (6.1.17) in favor of the two event vectors f_i and Af_i or by f_i and $Af_i + \alpha_\xi \xi_i$ where ξ_i is any vector in the detection space of f_i and α_ξ can be chosen arbitrarily. The purpose in selecting ξ_i and α_ξ is to minimize the total number of event vectors to be considered. Because Af_i satisfies:

$$Af_i = \begin{bmatrix} 0 \\ \alpha_i \\ \alpha_i \\ -\tau_i \\ 0 \end{bmatrix} = \alpha_i f_\beta - \tau_i f_i \tag{6.1.29}$$

it is convenient to let α_ξ and ξ_i equal τ_i and f_i respectively. The inboard aileron deflection sensor failure is then completely characterized by the set of event vectors $\{f_i, Af_i + \tau_i f_i\}$. This set is the same as $\{f_i, f_\beta\}$. Similar manipulations yield the set $\{f_o, f_\beta\}$ for an outboard aileron deflection sensor.

The sets of event vectors $\{f_i, f_\beta\}$ and $\{f_o, f_\beta\}$ establish a connection between the sensor failures and the failures for which the event vectors were originally defined. If a failure detection system is designed to detect failures in:

i) the inboard aileron,

ii) the outboard aileron, and

iii) a change in system dynamics

it also detects failures in:

iv) the inboard aileron deflection sensor, and

v) the outboard aileron deflection sensor.

The output estimation errors for i) - iii) are restricted to the single directions Cf_i, Cf_o, and Cf_β respectively. The sensor failures are differentiated from these failures by virtue of the fact that they produce output errors restricted to the planes defined by $\{Cf_i, Cf_\beta\}$ and $\{Cf_o, Cf_\beta\}$. A certain ambiguity is introduced in that it is not possible to differentiate between a failure of the inboard sensor and the simultaneous occurrence of i) and iii) above, etc. It is reasonable to assume that this ambiguity is acceptable if the failures have a low probability of occurrence and if they are reasonably independent.

With $\{f_i, f_\beta\}$ and $\{f_o, f_\beta\}$ associated with the sensor failures, the failure detection problem is well formulated. It is then necessary to determine the detection spaces for the failures using the algorithms of Appendix A.2. The results are summarized in Table 6.1. All of the detection spaces are one-dimensional and are spanned by their respective event vectors. The event vectors are also their own detection generators. In order to detect all eight failures with only two failure detection systems, however, it will be necessary to augment the dimension of the reference model (6.1.5) in such a manner that some of the detection spaces become higher dimensional.

The possibility of being able to detect all of i) -v) in Table 6.1 with a single failure detection system makes it attractive to begin the search for a set of mutually detectable failures with the event vectors f_i, f_o, f_β. This is possible only because the three event vectors are output separable (Definition 3.3.1). Of the remaining failures listed in Table 6.1 the only event vector which is output separable with f_i, f_o, f_β is f_a since:

Failure	Event Vector	Output Direction	Detection Space	Detection Generators
vi) attitude gyro	$f_a = e_{51}$	$Cf_a = e_{41}$	$\bar{R}_a = f_a$	$g_a = f_a$
vii) rate gyro	$f_r = e_{53}$	$Cf_r = e_{42}$	$\bar{R}_r = f_r$	$g_r = f_r$
i) inboard aileron	$f_i = e_{54}$	$Cf_i = e_{43}$	$\bar{R}_i = f_i$	$g_i = f_i$
ii) outboard aileron	$f_o = e_{55}$	$Cf_o = e_{44}$	$\bar{R}_o = f_o$	$g_o = f_o$
viii) amplifier (k_a only)	$f_{ka} = k_i e_{54} + k_o e_{55}$	$Cf_{ka} = k_i e_{43} + k_o e_{44}$	$\bar{R}_{ka} = f_{ka}$	$g_{ka} = f_{ka}$
iii) dynamics change	$f_\beta = e_{52} + e_{53}$	$Cf_\beta = e_{42}$	$\bar{R}_\beta = f_\beta$	$g_\beta = f_\beta$
iv) inboard aileron deflection sensor	$\{f_i, f_\beta\}$	$\{Cf_i, Cf_\beta\}$	$\{\bar{R}_i, \bar{R}_\beta\}$	
v) outboard aileron deflection sensor	$\{f_o, f_\beta\}$	$\{Cf_o, Cf_\beta\}$	$\{\bar{R}_o, \bar{R}_\beta\}$	

Table 6.1 Summary of Failure Detection Properties for Failures of Lateral Mode Aircraft Autopilot.

$$F = [f_i \; f_o \; f_\beta \; f_a] \tag{6.1.30}$$

$$rkCF = rkF = 4$$

This implies that f_i, f_o, f_β, f_a is the largest set of output separable event vectors which contains f_i, f_o, f_β. If a failure detection system can be found for the event vectors in F, it detects i) - vi).

The detection space $\overline{R}_F$ for F can be determined from Appendix A.2. Because CF spans the output space and (A, C) is observable, the desired result is that $\overline{R}_F$ spans the state space:

$$rk\overline{R}_F = 5 \tag{6.1.31}$$

The individual detection spaces are one-dimensional and:

$$rk\overline{R}_i + rk\overline{R}_o + rk\overline{R}_\beta + rk\overline{R}_a = 4 \tag{6.1.32}$$

By Theorem 3.3.1, the difference in rank between (6.1.31) and (6.1.32) means that i) - vi) are <u>not</u> mutually detectable with a single failure detection system. A failure detection system can be found for which each of the non-sensor failures produces a undirectional output error signal, but one eigenvalue of the closed loop state matrix A-DC is unassignable. Assuming this eigenvalue is stable, its acceptability can only be determined from the stochastic design procedure of the next subsection. The principal disadvantage of the unassignable eigenvalue is that is restricts the freedom to tailor the failure detection system to improve state estimation properties or to enhance the error signals.

The unassignable eigenvalue of A-DC is associated with the <u>excess subspace</u> R_{0g} of $\overline{R}_F$. $\overline{R}_F$ can be expressed in terms of the detection spaces in (6.1.32) and R_{0g}:

$$\overline{R}_F = [\overline{R}_i : \overline{R}_o : \overline{R}_\beta : \overline{R}_a : R_{0g}] \tag{6.1.33}$$

Appendix A.3 develops an algorithm for R_{0g} based upon the orthogonal reduction procedure. In this case R_{0g} corresponds to the true roll rate $\dot{\varphi}(t)$ of the aircraft:

$$R_{0g} = e_{52} \tag{6.1.34}$$

$$\nu_0 = 1$$

Because the event vectors in F are their own detection generators:

$$G = [g_i \; g_o \; g_\beta \; g_a] = F \tag{6.1.35}$$

It is shown in Section 3.3.3 that R_{0g} must satisfy the equation:

$$AR_{0g} = \begin{bmatrix} 1 \\ -\beta_\varphi \\ -\beta_\varphi \\ 0 \\ 0 \end{bmatrix} = R_{0g}\Pi + G\Theta \tag{6.1.36}$$

Π and Θ can be determined from (6.1.36) or Appendix A.3:

$$\Pi = 0$$

$$\Theta = \begin{bmatrix} 0 \\ 0 \\ -\beta_\varphi \\ 1 \end{bmatrix} \tag{6.1.37}$$

The unassignable eigenvalue of A-DC is the eigenvalue of Π, which in this case is zero. If a failure detection system is designed for i) - vi), (A-DC, C) must have at least one pure integration and marginal stability.

The excess subspace of $\overline{R}_F$ is brought about by the presence of the three coupled state variables $\varphi_a(t)$, $\dot{\varphi}(t)$, and $\dot{\varphi}_r(t)$ in the state vector, only two of which are measured. The unmeasured state variable $\dot{\omega}(t)$ is the derivative of $\varphi_a(t)$ and it is closely related to the combination of state variables $\dot{\varphi}(t) + \dot{\varphi}_r(t)$ which is associated with f_β and a dynamics change. Because of this, it cannot be uniquely associated with either $\overline{R}_a$ or $\overline{R}_\beta$. The excess subspace and the unassignable eigenvalue can be eliminated, but only by either removing f_a or f_β from the set f_i, f_o, f_β, f_a or through a change in the dimension of (A, C). In general, both alternatives should be tried to determine their impact on the failure detection problem. In this instance, it is convenient to remove f_a and require the second failure detection system to detect both gyro failures.

The fact that removal of f_a or f_β from the set of event vectors eliminates the excess subspace can be verified by a simple application of Theorem 3.3.3. Removing f_a from F requires that the fourth column of G and fourth row of Θ be removed to define:

$$G_a = [g_a] \tag{6.1.38}$$

$$\Theta_a^c = [1]$$

If F_a is defined similarly:

$$F_a = [f_i \; f_o \; f_\beta] \tag{6.1.39}$$

the theorem says that the excess subspace for $\overline{R}_{Fa}$ is then:

$$R_{0ga} = R_{0g}\beta$$

Since the rank of R_{0g} is one, β is a column vector (in this case, a scalar) which spans the null space of M_{0a} where:

$$M_{0a} = \Theta_a^c \tag{6.1.40}$$

M_{0a} is a scalar and has no null space. Therefore, both β and R_{0ga} are zero and:

$$rk\bar{R}_{Fa} = 3 = rk\bar{R}_i + rk\bar{R}_o + rk\bar{R}_\beta \qquad (6.1.41)$$

Equation (6.1.41) in turn implies that:

$$\bar{R}_{Fa} = [\bar{R}_i : \bar{R}_o : \bar{R}_\beta] \qquad (6.1.42)$$

This can be verified by applying the algorithm of Appendix A.2 for $\bar{R}_{Fa}$. Theorem 3.3.1 states that (6.1.41) is sufficient for the failures i) - v) associated with f_i, f_o, f_β to be mutually detectable.

If a failure detection system is designed for i) - v), the last step of the deterministic design procedure is to determine if any of the remaining failures listed in Table 6.1 can be made output stationary with i) - v) (Definition 4.1.1). The requirements for output stationarity are proved in Theorem 3.4.1. and stated in matrix form in Section 4.2.3 as a set of two necessary and sufficient conditions. The first is that the event vector associated with the failure under consideration cannot be output separable with f_i, f_o, f_β. This immediately rules out f_a and the attitude gyro failure. The second condition is tested below for f_r and f_{ka} and their corresponding failures. Because of the observation in Section 6.1.1 that an amplifier failure produces an output error similar to that generated by simultaneous aileron failures, there is prior reason to suspect that f_{ka} satisfies both conditions.

For an amplifier failure to be output stationary with i) - v), the second condition of Theroem 3.4.1 requires that the non-zero coefficients in the expression for Cf_{ka} be determined:

$$Cf_{ka} = \alpha_{ia} Cf_i + \alpha_{oa} Cf_o + \alpha_\beta Cf_\beta \qquad (6.1.43)$$

Because f_i, f_o, f_β are output separable and f_i, f_o, f_β, f_{ka} are not (the first condition), the coefficients in (6.1.43) exist and they are unique:

$$\alpha_{ia} = k_i \qquad (6.1.44)$$
$$\alpha_{oa} = k_o$$
$$\alpha_{\beta a} = 0$$

If the matrix $\overline{R}_{io}^{c}$ is formed from the subset of $\overline{R}_i$, $\overline{R}_o$, $\overline{R}_\beta$ for which the coefficients are non-zero:

$$\overline{R}_{io}^{c} = [\overline{R}_i : \overline{R}_o] \qquad (6.1.45)$$

then an amplifier failure can be made output stationary with i) - v) if some ξ_{ka} exists which satisfies (4.2.26):

$$\overline{R}_{ka} = \overline{R}_{io}^{c}\, \xi_{ka} \qquad (6.1.46)$$

In this case a unique solution exists:

$$\xi_{ka} = \begin{bmatrix} k_i \\ k_o \end{bmatrix} \qquad (6.1.47)$$

and an amplifier failure can be made output stationary with i) - v). When the same test is repeated for a rate gyro failure, (4.2.26) requires ξ_r to satisfy:

$$\overline{R}_r = \overline{R}_\beta\, \xi_r \qquad (6.1.48)$$

and clearly no such ξ_r exists.

Even though f_{ka} can be made output stationary with f_i and f_o, it will not be unless the eigenvalues of A-DC are assigned properly. Because $\overline{R}_{ka}$ satisfies (6.1.46), Theorem 3.4.2 requires that the eigenvalues of A-DC associated with $\overline{R}_i$ and $\overline{R}_o$ be equal in order to insure output stationarity. An algorithm for satisfying this constraint is illustrated in Section 6.2.2. The necessary procedure for discriminating between an amplifier failure and aileron failures is developed in Section 6.3.

The conclusion of the analysis is that, with output stationarity, a failure detection system can be designed to monitor i) - v) and viii) from Table 6.1. If it is assumed that no simultaneous failures occur, the failure detection system can unambiguously detect the occurrence of any of the six failures. The one exception is that it cannot determine on the basis of output estimation error alone which amplifier has failed. With state space augmentation, the failure detection system would also be able to detect an attitude gyro failure. If this were done the rate gyro failure would still violate the output stationarity requirements and a second failure detection system would be required. For state estimation reasons discussed in Section 6.4, the second failure detection system is required to detect both gyro failures.

The difficulties with detecting an attitude gyro failure and a dynamics change with a single detection system persist if the dynamics change is replaced by a rate gyro failure requirement. With F defined:

$$F = [f_a \; f_r] \tag{6.1.49}$$

$\overline{R}_F$ can be computed from the algorithm of Appendix A.2 and:

$$rk\,\overline{R}_F = 3 > rk\,\overline{R}_a + rk\,\overline{R}_r \tag{6.1.50}$$

It follows that the gyro failures are not mutually detectable. The excess subspace is again associated with $\dot{\phi}(t)$:

$$R_{0g} = e_{52} \tag{6.1.51}$$

$$\nu_0 = 1$$

but Π and Θ are different. Since $\overline{R}_a$ and $\overline{R}_r$ are one-dimensional, G is equal to F as before:

$$G = [g_a \; g_r] = F \tag{6.1.52}$$

Both G and F are different from the previous case. Π and Θ now must satisfy:

$$A R_{0g} = \begin{bmatrix} 1 \\ -\beta_{\varphi} \\ -\beta_{\varphi} \\ 0 \\ 0 \end{bmatrix} = R_{0g} \Pi + G \Theta \qquad (6.1.53)$$

Because G has changed, Π and Θ are:

$$\Pi = -\beta_{\varphi}$$

$$\Theta = \begin{bmatrix} 1 \\ -\beta_{\varphi} \end{bmatrix} = \begin{bmatrix} \theta_a \\ \theta_r \end{bmatrix} \qquad (6.1.54)$$

This is an improvement over the earlier result in that the unassignable eigenvalue of A-DC (i.e., the eigenvalue of Π) is stable. For this particular application it is smaller than desired for state estimation purposes, however, and it is desirable to make the gyro failures mutually detectable through state space augmentation.

The procedure for implementing the state space augmentation is summarized in Section 4.2.3 and stated in matrix form in Lemma 3.3.10. The construction in the lemma requires quantities defined in Theorem 3.3.3 and its Corollary. Assume f_a were to be removed from the set f_a, f_r. The proof of Theorem 3.3.3 defines a transformation $[\Delta : \beta]_a$ on the excess subspace R_{0g} such that:

$$[R^c_{0ga} : R_{0ga}] = R_{0g}[\Delta : \beta]_a \qquad (6.1.55)$$

R_{0ga} is the new excess subspace if f_a is removed and R^c_{0ga} identifies the portion of the excess subspace which is removed. But the only remaining event vector in the set is f_r and the Detection Theorem (Theorem 3.2.1) guarantees that f_r is detectable. Therefore the new excess subspace is zero and the transformation in (6.1.55) is more correctly defined in this case by:

$$R^{c}_{0ga} = R_{0g}[\Delta]_a = R_{0g} = e_{52} \tag{6.1.56}$$

The partitions of Π in the Corollary of Theorem 3.3.3 can then be restated along the lines of (6.1.56):

$$\Pi^{c}_{a} = [\Delta]_a^{-1} \Pi [\Delta]_a = -\beta_{\varphi} \tag{6.1.57}$$

Lastly, the decrease in the dimension of R_{0g} if f_a is removed from f_a, f_r is defined:

$$\nu^{c}_{0a} = \nu_0 = 1 \tag{6.1.58}$$

The preceding arguments could have been made for the removal of f_r from f_a, f_r. In particular, f_a is detectable and the excess subspace associated with f_a is zero; therefore:

$$\nu^{c}_{0r} = \nu_0 = 1 \tag{6.1.59}$$

Theorem 3.3.4 then guarantees that f_a, f_r can be made mutually detectable for a reference model of dimension $\tilde{n}$:

$$\tilde{n} = n + \nu^{c}_{0a} + \nu^{c}_{0r} - \nu_0 = 6 \tag{6.1.60}$$

Because $\tilde{n}$ is greater than n, it is necessary to add a state component to the state vector with a single application of Lemma 3.3.10. With g_a and θ_a given by Table 6.1 and (6.1.54) respectively, Lemma 3.3.10 yields the six-dimensional reference model $(\tilde{A}, \tilde{C})$ when applied for f_a:

$$\tilde{A} = \left[\begin{array}{c|c} A & -g_a \\ \hline 0 & -\beta_{\varphi} \end{array}\right] = \left[\begin{array}{c|c} A & -e_{51} \\ \hline 0 & -\beta_{\varphi} \end{array}\right] \tag{6.1.61}$$

Furthermore, the detection space for $\tilde{f}_a$ in the augmented state space and its detection generator are:

$$\tilde{\bar{R}}_a = \left[\begin{array}{c|c} f_a & \\ \hline 0 & e_{66} \end{array}\right] = [\tilde{f}_a : e_{66}] \qquad (6.1.62)$$

$$\tilde{g}_a = -e_{66}$$

In the augmented state space, the detection space for $\tilde{f}_a$ is two-dimensional.

When a failure detection system is being designed for more than two failures, the state space augmentation may increase the dimensions of the detection spaces of the remaining event vectors or it may leave them unchanged. When only two failures are concerned, however, the dimensions of both detection spaces must be increased. The detection space $\tilde{\bar{R}}_r$ for $\tilde{f}_r$ can be determined from Appendix A.4 or Lemma 3.3.11. Lemma 3.3.11 defines $\tilde{\bar{R}}_r$ by:

$$\tilde{\bar{R}}_r = \left[\begin{array}{c|cc} \bar{R}_r & & \\ \hline 0 & \tilde{R}_D & Z_r \end{array}\right] \qquad (6.1.63)$$

where the substitution of $[\Delta]_a$ for $[\Delta : \beta]_a$ allows $\tilde{R}_D$ to be defined:

$$\tilde{R}_D = \left[\begin{array}{c} R^c_{0ga} \\ \hline 1 \end{array}\right] = \left[\begin{array}{c} e_{52} \\ \hline 1 \end{array}\right] = e_{62} + e_{66} \qquad (6.1.64)$$

Z_r spans the null space of M_{0ar}. Since $(\Theta_a)_r$ is Θ with both rows removed, it follows that M_{0ar} is zero and :

$$\tilde{\bar{R}}_r = \left[\begin{array}{c|c} f_r & \\ \hline 0 & e_{62} + e_{66} \end{array}\right] = [\tilde{f}_r : e_{62} + e_{66}] \qquad (6.1.65)$$

$$\tilde{g}_r = -\frac{1}{\beta_\varphi}(e_{62} + e_{66})$$

As a check on the state space augmentation procedure, $\tilde{\bar{R}}_F$ can be computed from Appendix A.2. For the six-dimensional reference model, F must be replaced by:

$$\tilde{\tilde{F}} = \begin{bmatrix} F \\ --- \\ 0 \end{bmatrix} \tag{6.1.66}$$

The result is that:

$$rk\tilde{\tilde{R}}_F = 4 = rk\tilde{\tilde{R}}_a + rk\tilde{\tilde{R}}_r \tag{6.1.67}$$

and $\tilde{\tilde{f}}_a$, $\tilde{\tilde{f}}_r$ are mutually detectable for $(\tilde{\tilde{A}}, \tilde{\tilde{C}})$.

Introduction of a six-dimensional reference model for detecting gyro failures requires that a six-component state vector be defined. Because A is a diagonal submatrix of $\tilde{\tilde{A}}$, the first five components of $\tilde{\tilde{x}}(t)$ are the same as those of x(t) as defined in (6.1.4). The new component is denoted by $\dot{\phi}_0(t)$:

$$\tilde{\tilde{x}}(t) = \begin{bmatrix} \varphi_a(t) \\ \dot{\varphi}(t) \\ \dot{\varphi}_r(t) \\ \delta_{ai}(t) \\ \delta_{ao}(t) \\ \dot{\varphi}_0(t) \end{bmatrix} \tag{6.1.68}$$

The notation suggests that $\dot{\varphi}_0(t)$ is in some way related to the true roll rate $\dot{\varphi}(t)$. The time constant associated with $\dot{\varphi}_0(t)$ is the roll dynamics time constant $\frac{1}{\beta_\varphi}$. The difference between $\dot{\varphi}_0(t)$ and $\dot{\varphi}(t)$ is that $\dot{\varphi}_0(t)$ is not driven by the ailerons and that it interacts with $\varphi_a(t)$ and $\dot{\varphi}_r(t)$ in a different manner than $\dot{\varphi}(t)$ does. For practical purposes, however, $\dot{\varphi}_0(t)$ can be thought of as a duplication of $\dot{\varphi}(t)$ which permits the reference model to decouple the relationship between $\varphi_a(t)$ and $\dot{\varphi}(t)$ from that between $\dot{\varphi}_r(t)$ and $\dot{\varphi}(t)$.

6.2 Detection Filter Design - Stochastic

The deterministic design procedure requires a continuous, linear, deterministic reference model. The reference model in

Section 6.1.1 was chosen to satisfy this constraint. The stochastic design procedure chooses the failure detection systems to perform state estimation as well as failure detection. Optimization of the state estimation properties of the failure detection systems requires knowledge of the stochastic environment in which the aircraft is to operate and of the sensor accuracies. A stochastic model for the autopilot-aircraft system is developed in Section 6.2.1. The model is used in Section 6.2.2 to determine optimal detection filter-state estimators for the two failure detection systems of Section 6.1.2.

6.2.1 Stochastic Model of Autopilot - Aircraft System

The failure detection theory developed in Chapter 3 is analytically valid for continuous systems only if the systems have continuous measurements. Approximation procedures for applying the theory to sampled-data systems were developed in Section 5.2. In order to demonstrate the range of applicability of these procedures as completely as possible, the autopilot for the stochastic reference model is chosen to be a digital autopilot with a sampling period of .01 seconds. This is a shorter sampling period than commonly used in digital autopilots, but it is useful for evaluating the ability of a failure detection system to detect failures in sampled data systems. The measurements available to the failure detection systems are discrete with a sampling period which can be varied from .01 seconds to .2 seconds. With measurements every .01 seconds, the failure detection system has complete knowledge of the autopilot inputs and outputs and it is reasonable to expect the failure detection system to provide good failure isolation. With a measurement sampling period of .2 seconds, the .01 seconds autopilot sampling period allows a system failure to propagate through the control loop between measurements and makes it difficult for the failure detection system to determine the source of the failure.

The sampled-data formulation of the stochastic reference model is developed in this subsection. An approximation of the autopilot-aircraft system as a continuous stochastic system with continuous

measurements is then made utilizing standard stochastic approximation procedures. For a short measurement sampling period, the sampled-data and continuous measurement reference models have nearly equivalent statistics. The use of both reference models to define optimal detection filter-state estimators is illustrated in Section 6.2.2. In the following, the measurement sampling period is denoted by Δt. It is assumed to be some integer multiple L of the autopilot sampling period T:

$$\Delta t = LT \tag{6.2.1}$$

For state estimation purposes, the dynamic response of a sampled-data system can be characterized by the propagation of the state from one measurement to the next:

$$x(k\Delta t) = \Phi(k\Delta t, (k-1)\Delta t)\, x((k-1)\Delta t) + \int_{(k-1)\Delta t}^{k\Delta t} \Phi(k\Delta t, \tau)\, Bu(\tau)\, d\tau + G_{1\Delta}\, \eta_{1\Delta}(k\Delta t)$$

$$y(k\Delta t) = Cx(k\Delta t) + G_{2\Delta}\, \eta_{2\Delta}(k\Delta t) \tag{5.2.3}$$

$G_{1\Delta}\, \eta_{1\Delta}(k\Delta t)$ models the effect of system disturbances in the interval $(k-1)\Delta t < t \leq k\Delta t$ on the state at $k\Delta t$. The system disturbances may be either continuous or discrete. $G_{1\Delta}$ and $G_{2\Delta}$ are $n \times n_1$ and $m \times m_1$ matrices respectively. $\eta_{1\Delta}(k\Delta t)$ and $\eta_{2\Delta}(k\Delta t)$ are discrete, zero mean, white Gaussian random processes with variances:

$$E\left[\eta_{1\Delta}(k\Delta t)\, \eta_{1\Delta}(k\Delta t)^T\right] = \frac{1}{\Delta t} E_{n_1}$$

$$E\left[\eta_{2\Delta}(k\Delta t)\, \eta_{2\Delta}(k\Delta t)^T\right] = \frac{1}{\Delta t} E_{m_1} \tag{5.2.4}$$

Figure 6.3 is a mathematical block diagram for the autopilot-aircraft system except that the measurements are not indicated. The system disturbances to be modeled by $\eta_{1\Delta}(k\Delta t)$ are:

i) a random aerodynamic moment,
ii) an uncertainty in the attitude gyro output, and
iii) an uncertainty in the rate gyro output.

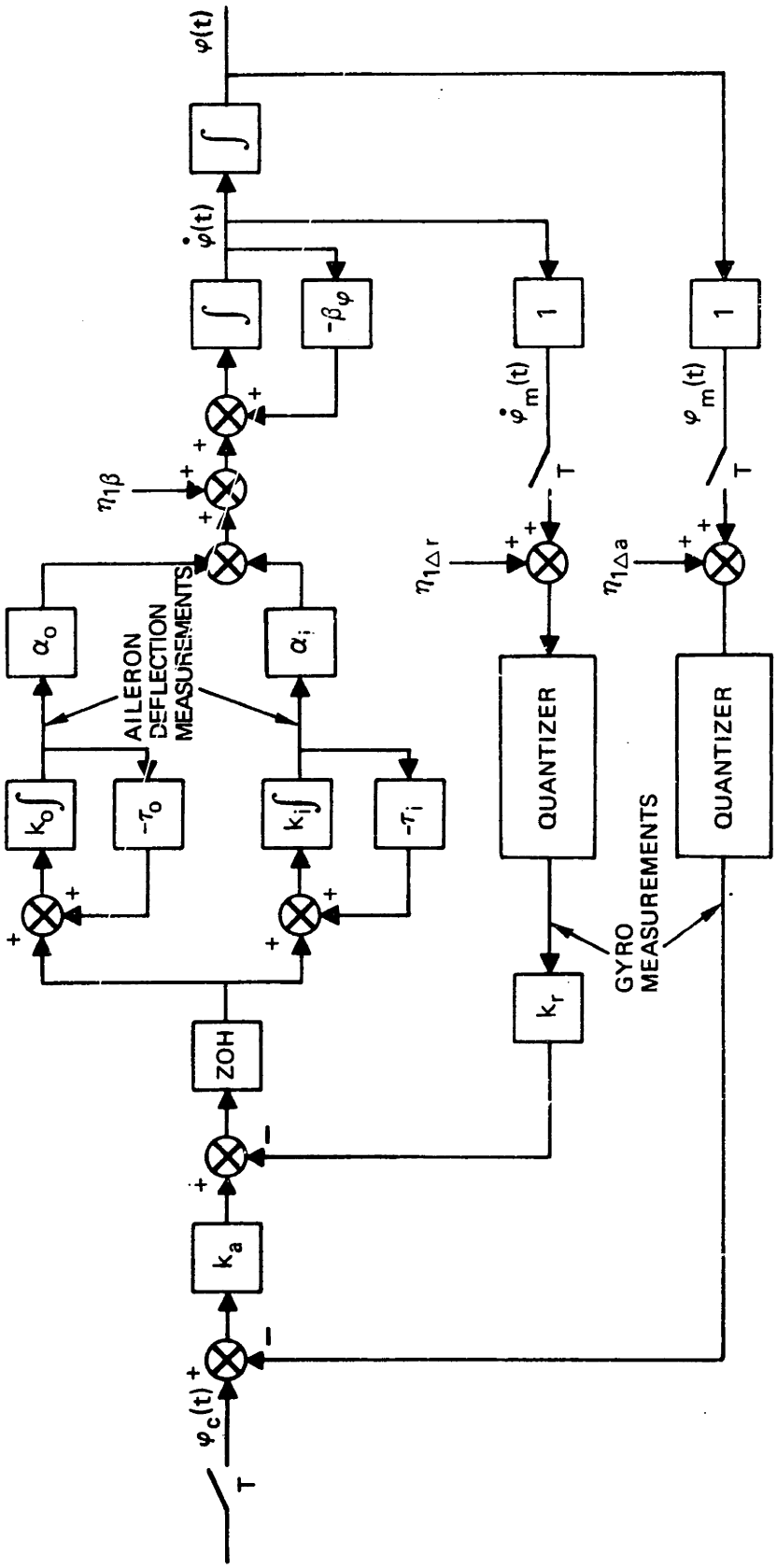


Figure 6.3 Mathematical block diagram for lateral mode autopilot-aircraft combination

The gyro output uncertainties are system disturbances because they appear as inputs to the respective amplifiers. Figure 6.3 indicates that they also appear as additive components in the measurements. Therefore, the measurement uncertainties to be modeled by $\eta_{2\Delta}(k\Delta t)$ include:

iv) a random uncertainty in the output of both aileron deflection sensors, and

v) both gyro uncertainties.

It follows that $\eta_{1\Delta}(k\Delta t)$ and $\eta_{2\Delta}(k\Delta t)$ can be defined as three-dimensional and four-dimensional vector processes respectively. Because of the gyro uncertainties, they are not independent.

The aerodynamic moment is a consequence of wind gusts and the modeling errors attendant with the selection of simplified aircraft dynamics. It can be represented as a continuous white Gaussian random process with a power spectral density W_β. Because of the fact that $\dot{\varphi}(t)$ and $\dot{\varphi}_r(t)$ must experience the same aircraft dynamics, $G_{1\Delta}$ must be chosen such that $\dot{\varphi}(k\Delta t)$ and $\dot{\varphi}_r(k\Delta t)$ in (5.2.3) are driven by the random aerodynamic moment in identical manners. If the vector $g_{1\beta}$ is defined similarly to f_β in (6.1.26):

$$g_{1\beta} = \begin{bmatrix} 0 \\ \sqrt{W_\beta} \\ \sqrt{W_\beta} \\ 0 \\ 0 \end{bmatrix} \tag{6.2.2}$$

and if $\eta_{1\beta}(t)$ is a continuous white Gaussian random process with zero mean:

$$E\left[\eta_{1\beta}(t)\,\eta_{1\beta}(\tau)\right] = \delta(t-\tau) \tag{6.2.3}$$

then $g_{1\beta}\eta_{1\beta}(t)$ is a model of the continuous random aerodynamic moment. The discrete process $g_{1\beta\Delta}\,\eta_{1\beta\Delta}(k\Delta t)$ necessary for (5.2.3) is the effect

of the random aerodynamic moment in the interval $(k-1)\Delta t < t \leq k\Delta t$ on $x(k\Delta t)$:

$$g_{1\beta\Delta}\,\eta_{1\beta\Delta}(k\Delta t) = \int_{(k-1)\Delta t}^{k\Delta t} \Phi(k\Delta t, \tau)\, g_{1\beta}\,\eta_{1\beta}(\tau)\, d\tau \qquad (6.2.4)$$

The variance of this random process is:

$$E[g_{1\beta\Delta}\,\eta_{1\beta\Delta}(k\Delta t)^2\, g_{1\beta\Delta}^T] = E[\iint_{(k-1)\Delta t}^{k\Delta t} \Phi(k\Delta t, \tau)\, g_{1\beta}\,\eta_{1\beta}(\tau)\,\eta_{1\beta}(\tau')$$

$$g_{1\beta}^T\,\Phi(k\Delta t, \tau')^T\, d\tau\, d\tau'] \qquad (6.2.5)$$

If $\eta_{1\beta\Delta}(k\Delta t)$ is defined to have a variance of $\frac{1}{\Delta t}$:

$$E[\eta_{1\beta\Delta}(k\Delta t)\,\eta_{1\beta\Delta}(k\Delta t)] = \frac{1}{\Delta t} \qquad (6.2.6)$$

then (6.2.3) - (6.2.5) imply that $g_{1\beta\Delta}$ must satisfy:

$$\frac{1}{\Delta t}\, g_{1\beta\Delta}\, g_{1\beta\Delta}^T = \int_{(k-1)\Delta t}^{k\Delta t} \Phi(k\Delta t, \tau)\, g_{1\beta}\, g_{1\beta}^T\, \Phi(k\Delta t, \tau)^T\, d\tau \qquad (6.2.7)$$

The attitude gyro and the rate gyro are interrogated by the autopilot every T seconds. The gyro output uncertainties are assumed to result from the addition of discrete white Gaussian random processes to $\varphi_a(\ell T)$ and $\dot{\varphi}_r(\ell T)$ and the quantization of the resulting signals. If the variances of the additive processes are $\sigma_{1a\Delta}^2$ for the attitude gyro and $\sigma_{1r\Delta}^2$ for the rate gyro, and the quantization intervals are q_a and q_r, the combined output uncertainty for each gyro can be approximated for error analysis purposes by discrete Gaussian random processes with variances:

$$\sigma'^{2}_{1a\Delta} = \sigma^{2}_{1a\Delta} + \frac{q_a^2}{12}$$

$$\sigma'^{2}_{1r\Delta} = \sigma^{2}_{1r\Delta} + \frac{q_r^2}{12} \qquad (6.2.8)$$

Under the assumption that the aircraft roll mode is sufficiently active that the gyro outputs can be expected to change by at least one quantization interval over the autopilot sampling period, these random processes are approximately white. Because the uncertainties are additive at the gyro outputs, they first appear in (5.2.3) as system disturbances driving the ailerons. If $g'_{1a\Delta}$ and $g'_{1r\Delta}$ are defined:

$$g'_{1a\Delta} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -k_i\, k_a\, \sigma'_{1a\Delta} \\ -k_o\, k_a\, \sigma'_{1a\Delta} \end{bmatrix}$$

$$g'_{1r\Delta} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -k_i\, k_r\, \sigma'_{1r\Delta} \\ -k_o\, k_r\, \sigma'_{1r\Delta} \end{bmatrix} \qquad (6.2.9)$$

along with the independent discrete white Gaussian noises $\eta'_{1a\Delta}(\ell T)$ and $\eta'_{1r\Delta}(\ell T)$:

$$E[\eta'_{1a\Delta}(\ell T)\ \eta'_{1a\Delta}(\ell T)] = 1$$

$$E[\eta'_{1r\Delta}(\ell T)\ \eta'_{1r\Delta}(\ell T)] = 1 \qquad (6.2.10)$$

then $g'_{1a\Delta}\eta'_{1a\Delta}(\ell T)$ and $g'_{1r\Delta}\eta'_{1r\Delta}(\ell T)$ are acceptable models of the gyro output uncertainties.

The contribution of the gyro output uncertainties in the interval $(k-1)\,\Delta t < t \le k\Delta t$ to $\mathbf{x}(k\Delta t)$ can be determined in the same manner used for the random aerodynamic moment. Let $\eta_{1a_\Delta}(k\Delta t)$ and $\eta_{1r_\Delta}(k\Delta t)$ be discrete white Gaussian random processes derived from $\eta'_{1a_\Delta}(\ell T)$ and $\eta'_{1r_\Delta}(\ell T)$:

$$E[\eta_{1a_\Delta}(k\Delta t)\,\eta_{1a_\Delta}(k\Delta t)] = \frac{1}{\Delta t}$$

$$E[\eta_{1r_\Delta}(k\Delta t)\;\eta_{1r_\Delta}(k\Delta t)] = \frac{1}{\Delta t} \qquad (6.2.11)$$

and let g_{1a_Δ} and g_{1r_Δ} satisfy:

$$g_{1a_\Delta}\,\eta_{1a_\Delta}(k\Delta t) = \sum_{\ell=0,\,L-1} \int_{(k-1)\Delta t+\ell T}^{(k-1)\,\Delta t+(\ell+1)T} \Phi(k\Delta t,\tau)\, g'_{1a_\Delta}\,\eta'_{1a_\Delta}((k-1)\Delta t+\ell T)\,d\tau$$

$$g_{1r_\Delta}\,\eta_{1r_\Delta}(k\Delta t) = \sum_{\ell=0,\,L-1} \int_{(k-1)\Delta t+\ell T}^{(k-1)\Delta t+(\ell+1)T} \Phi(k\Delta t,\tau)\, g'_{1r_\Delta}\,\eta'_{1r_\Delta}((k-1)\Delta t+\ell T)\,d\tau \qquad (6.2.12)$$

Because $\eta'_{1a_\Delta}(\ell T)$ and $\eta'_{1r_\Delta}(\ell T)$ are discrete random processes, the covariance equations which define g_{1a_Δ} and g_{1r_Δ} differ from (6.2.7):

$$\frac{1}{\Delta t}\, g_{1a_\Delta}\, g_{1a_\Delta}^{T} = \sum_{\ell=0,L-1} \int_{(k-1)\Delta t+\ell T}^{(k-1)\Delta t+(\ell+1)T} \int \Phi(k\Delta t,\tau)\, g'_{1a_\Delta}\, g'^{\,T}_{1a_\Delta}\, \Phi(k\Delta t,\tau')^{T} d\tau\, d\tau'$$

$$\frac{1}{\Delta t}\, g_{1r_\Delta}\, g_{1r_\Delta}^{T} = \sum_{\ell=0,\,L-1} \int_{(k-1)\Delta t+\ell T}^{(k-1)\Delta t+(\ell+1)T} \int \Phi(k\Delta t,\tau)\, g'_{1r_\Delta}\, g'^{\,T}_{1r_\Delta}\, \Phi(k\Delta t,\tau')^{T} d\tau\, d\tau' \qquad (6.2.13)$$

$g_{1a\Delta}\,\eta_{1a\Delta}(k\Delta t)$ and $g_{1r\Delta}\,\eta_{1r\Delta}(k\Delta t)$ are the desired gyro output uncertainty models for the attitude gyro and the rate gyro respectively. With (6.2.4) and (6.2.7), they allow $G_{1\Delta}$ and $\eta_{1\Delta}(k\Delta t)$ to be defined:

$$G_{1\Delta} = [g_{1\beta\Delta}\ g_{1a\Delta}\ g_{1r\Delta}]$$

$$\eta_{1\Delta}(k\Delta t) = \begin{bmatrix} \eta_{1\beta\Delta}(k\Delta t) \\ \eta_{1a\Delta}(k\Delta t) \\ \eta_{1r\Delta}(k\Delta t) \end{bmatrix} \qquad (6.2.14)$$

The model for the measurement uncertainties in (5.2.3) is considerably simpler to arrive at. Gyro output measurements are made every L^{th} autopilot sampling period. Their variances are given in (6.2.8). The gyro output uncertainties are the same uncertainties which appear as new system disturbances at the measurement times. Let the random processes $\eta_{2a\Delta}(k\Delta t)$ and $\eta_{2r\Delta}(k\Delta t)$ be defined in terms of every L^{th} sample of $\eta'_{1a\Delta}(\ell T)$ and $\eta'_{1r\Delta}(\ell T)$:

$$\eta_{2a\Delta}(k\Delta t) = \frac{1}{\sqrt{\Delta t}}\ \eta'_{1a\Delta}(kLT)$$

$$\eta_{2r\Delta}(k\Delta t) = \frac{1}{\sqrt{\Delta t}}\ \eta'_{1r\Delta}(kLT) \qquad (6.2.15)$$

From (6.2.10) it follows that $\eta_{2a\Delta}(k\Delta t)$ and $\eta_{2r\Delta}(k\Delta t)$ are independent white Gaussian random processes. The measurement errors in the aileron deflection sensors are modeled as white random processes with variances of $\sigma^2_{2i\Delta}$ for the inboard sensor and $\sigma^2_{2o\Delta}$ for the outboard sensor. These errors can be modeled in terms of two additional zero mean white Gaussian random processes $\eta_{2i\Delta}(k\Delta t)$ and $\eta_{2o\Delta}(k\Delta t)$:

$$E\left[\eta_{2i_\Delta}(k\Delta t)\,\eta_{2i_\Delta}(k\Delta t)\right] = \frac{1}{\Delta t}$$
$$E\left[\eta_{2o_\Delta}(k\Delta t)\,\eta_{2o_\Delta}(k\Delta t)\right] = \frac{1}{\Delta t} \qquad (6.2.16)$$

The vector measurement uncertainty can be defined in terms of (6.2.15) and (6.2.16):

$$\eta_{2_\Delta}(k\Delta t) = \begin{bmatrix} \eta_{2a_\Delta}(k\Delta t) \\ \eta_{2r_\Delta}(k\Delta t) \\ \eta_{2i_\Delta}(k\Delta t) \\ \eta_{2o_\Delta}(k\Delta t) \end{bmatrix}$$

$$G_{2_\Delta} = \left[g_{2a_\Delta} g_{2r_\Delta}\, g_{2i_\Delta} g_{2o_\Delta}\right] = \sqrt{\Delta t}\begin{bmatrix} \sigma'_{1a_\Delta} & 0 & 0 & 0 \\ 0 & \sigma'_{1r_\Delta} & 0 & 0 \\ 0 & 0 & \sigma_{2i_\Delta} & 0 \\ 0 & 0 & 0 & \sigma_{2o_\Delta} \end{bmatrix} \qquad (6.2.17)$$

As mentioned earlier, the random processes $\eta_{1_\Delta}(k\Delta t)$ and $\eta_{2_\Delta}(k\Delta t)$ as defined in (6.2.14) and (6.2.17) are not independent. The measurement noise $G_{2_\Delta}\eta_{2_\Delta}(k\Delta t)$ contains information regarding $G_{1_\Delta}\eta_{1_\Delta}((k+1)\Delta t)$. The algebra is somewhat involved but the correlation can be developed from (6.2.12) - (6.2.17):

$$E\left[G_{1_\Delta}\eta_{1_\Delta}((k+1)\Delta t)\,\eta_{2_\Delta}^{T}(k\Delta t)\,g_{2_\Delta}^{T}\right] = \frac{1}{\sqrt{\Delta t}}\int_{k\Delta t}^{k\Delta t+T} \Phi((k+1)\Delta t,\tau)(g'_{1a_\Delta} g_{2a_\Delta}^{T} + g'_{1r_\Delta} g_{2r_\Delta}^{T})$$

(6.2.18)

If a reference model has continuous measurements, the standard stochastic system model is:

$$\dot{x}(t) = Ax(t) + Bu(t) + G_1 \eta_1(t)$$

$$y(t) = Cx(t) + G_2 \eta_2(t) \qquad (2.2.1)$$

where G_1 and G_2 are $n \times n_1$ and $m \times m_1$ matrices respectively and $\eta_1(t)$ and $\eta_2(t)$ are continuous, zero mean, white Gaussian random processes:

$$E[\eta_1(t)\ \eta_1(\tau)^T] = \delta(t-\tau)\ E_{n_1}$$

$$E[\eta_2(t)\ \eta_2(\tau)^T] = \delta(t-\tau)\ E_{m_1} \qquad (2.2.2)$$

An approximation of the sample data system in (5.2.3) by a reference model with continuous measurements can be made by defining:

$$G_1 = \frac{G_{1\Delta}}{\Delta t}$$

$$G_2 = G_{2\Delta} \qquad (6.2.19)$$

$G_{2\Delta}$ is a conveniently defined diagonal matrix. $G_{1\Delta}$ is defined in integral form by (6.2.7), (6.2.13), and (6.2.14). For small Δt, $\Phi(k\Delta t, \tau)$ can be approximated by E_n and (6.2.13) becomes:

$$\frac{1}{\Delta t}\ g_{1a\Delta}\, g_{1a\Delta}^T \approx \sum_{\ell=0,\,L-1} \frac{T^2}{2}\ g'_{1a\Delta}\, g'^{T}_{1a\Delta} = \Delta t^2\ \frac{1}{2L}\ g'_{1a\Delta}\, g'^{T}_{1a\Delta}$$

$$\frac{1}{\Delta t}\ g_{1r\Delta}\, g_{1r\Delta}^T \approx \sum_{\ell=0,\,L-1} \frac{T^2}{2}\ g'_{1r\Delta}\, g'^{T}_{1r\Delta} = \Delta t^2\ \frac{1}{2L}\ g'_{1r\Delta}\, g'^{T}_{1r\Delta} \qquad (6.2.20)$$

G_1 in (6.2.19) then becomes:

$$G_1 \approx \begin{bmatrix} 0 & 0 & 0 \\ \sqrt{W_\beta} & 0 & 0 \\ \sqrt{W_\beta} & 0 & 0 \\ 0 & \sqrt{\Delta t}\,\alpha_L\, k_i k_a\, \sigma'_{1a\Delta} & \sqrt{\Delta t}\,\alpha_L\, k_i k_r\, \sigma'_{1r\Delta} \\ 0 & \sqrt{\Delta t}\,\alpha_L\, k_o k_a\, \sigma'_{1a\Delta} & \sqrt{\Delta t}\,\alpha_L\, k_o k_r\, \sigma'_{1r\Delta} \end{bmatrix} \qquad (6.2.21)$$

where:

$$\alpha_L = \left(\frac{1}{2L}\right)^{\frac{1}{2}}$$

A similar approximation can be made in (6.2.18) to determine the correlation between $\eta_1(t)$ and $\eta_2(t)$:

$$G_{12} = E[\eta_1(t)\,\eta_2(\tau)^T] \approx \delta(t-\tau) \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2\alpha_L & 0 & 0 & 0 \\ 0 & 2\alpha_L & 0 & 0 \end{bmatrix} \qquad (6.2.22)$$

The acceptability of (2.2.1), (2.2.2), and (6.2.19) - (6.2.22) as an approximation of the autopilot-airplane system is a function of the relationship of Δt to the response time of the system. The continuous approximation represents in some sense the "average" stochastic behavior of the system. If T and Δt are large enough that a single gyro interrogation significantly alters either the system state or the estimate of its state, the continuous measurement approximation does not accurately portray the stochastic properties of the system.

6.2.2 Failure Detection System Design--Stochastic

Section 6.1.2 outlined the basic function of two failure detection systems for the lateral mode autopilot-aircraft combination. One system is designed to monitor failures in:

i) the inboard aileron,

ii) the outboard aileron,

iii) roll dynamics changes,

iv) the inboard aileron deflection sensor,

v) the outboard aileron deflection sensor,

viii) either amplifier, and

ix) (certain simultaneous, identical gyro failures).

The second is to monitor failures in:

vi) the attitude gyro,

vii) the rate gyro,

i) (the inboard aileron), and

ii) (the outboard aileron).

The failures in parentheses are added as a consequence of the analysis in this subsection. With the stochastic reference model of the previous subsection, it is possible to use the design freedom available in the selection of the detection filters to improve the state estimates generated by the failure detection systems. Because the stochastic reference model has discrete measurements, it is necessary to use one of the approximate algorithms of Section 5.2 to define detection filter-state estimators for the failure detection systems. If the set of event vectors $f_1 \ldots f_r$ associated with the failure to be detected satisfies:

$$\text{rk } C = m = \sum_{k=1,r} \text{rk } C f_k \qquad (6.2.23)$$

each of the algorithms of Section 5.2 converges to a unique detection filter-state estimator. Neither of the failure detection systems to be designed satisfies (6.2.23), however, and it is necessary to further constrain the optimization problem before applying the algorithms.

A failure detection system for a sampled-data system is a two-stage processor. The first stage involves a propagation of the state estimate $x_m((k-1)\Delta t)$ to $k\Delta t$ and the measurement of the propagated estimate $\bar{x}_m(k\Delta t)$:

$$\bar{x}_m(k\Delta t) = \Phi(k\Delta t, (k-1)\Delta t)\, x_m((k-1)\Delta t) + \int_{(k-1)\Delta t}^{k\Delta t} \Phi(t, \tau)\, Bu(\tau)\, d\tau$$

$$\bar{y}_m(k\Delta t) = C\, \bar{x}_m(k\Delta t) \qquad (5.2.5)$$

The second stage uses the new measurement to update the state estimate:

$$x_m(k\Delta t) = \bar{x}_m(k\Delta t) + D_{\Delta K}(y(k\Delta t) - \bar{y}_m(k\Delta t))$$

$$= \bar{x}_m(k\Delta t) + D_{\Delta K}\bar{\epsilon}'(k\Delta t) \qquad (5.2.6)$$

The algorithms of Section 5.2 choose $D_{\Delta K}$ to provide as good a state estimate as possible while satisfying the failure detection constraints. Interpretation of these constraints requires that the deterministic reference model (A, C) from Section 6.1.1 be employed. Using the appropriate base normal canonical form for each failure detection system (Section 4.4), a matrix $\hat{D}_{FR}$ is determined such that each of the failures being monitored generates a unidirectional output error for the continuous measurement system $(\hat{A}-\hat{D}_{FR}\hat{C}, \hat{C})$. $\hat{D}_{\Delta K}$ is then defined as the sum of three matrices:

$$\hat{D}_{\Delta K} = \hat{D}_{\Delta FR} + \hat{D}_{\Delta \Psi K} + \hat{D}_{\Delta \Gamma K} \qquad (5.2.21)$$

where:

$$\hat{D}_{\Delta FR} = \Delta t\, \hat{D}_{FR} \qquad (5.2.22)$$

Each element of $\hat{D}_{\Delta \Psi K}$ and $\hat{D}_{\Delta \Gamma K}$ is either constrained to be zero, or is a freely assignable parameter. The algorithms of Section 5.2 differ

only in the manner in which they assign these parameters. With $\hat{D}_{\Delta K}$ completely specified, $D_{\Delta K}$ is determined by a similarity transformation.

Techniques for formulating a solvable optimization for $\hat{D}_{\Delta\Psi K}$ and $\hat{D}_{\Delta\Gamma K}$ when (6.2.23) is not satisfied are discussed at the conclusion of Section 5.1.2. No rigid criteria exist for choosing between the options. The technique adopted in this subsection is to designate additional failures for each failure detection system to monitor so that (6.2.23) is satisfied. In the case of the failure detection system for the gyro failures, the added failures are physically meaningful and provide useful information; for the other failure detection system they are only a computational artifice. The advantage of adding the failures is that it eliminates $\hat{D}_{\Delta\Gamma K}$, it does not require an increase in the dimension of the state space, and it simplifies the detection law for isolating the failure. The disadvantage of this technique relative to the others is illustrated in Section 6.4.2. Section 6.4.3 discusses particular failure modes for which optimal or near-optimal detection filter-state estimators may yield unsatisfactory failure detection and demonstrates modifications in the choice of $\hat{D}_{\Delta K}$ which can be expected to improve the failure detection properties.

The event vectors associated with failures i) - v) and viii) in Table 6.1 are f_i, f_o, f_β. It was determined in Section 6.1.2 that the only event vector in Table 6.1 which is output separable with f_i, f_o, f_β, is f_a; but that the failures associated with f_i, f_o, f_β, f_a are not mutually detectable. The detection spaces for these event vectors satisfy:

$$\text{rk}\,\bar{R}_i + \text{rk}\,\bar{R}_o + \text{rk}\,\bar{R}_\beta + \text{rk}\,\bar{R}_a = 4 \qquad (6.1.32)$$

Theorem 3.3.1 required the summation in (6.1.32) to equal the dimension of the state space in order for the failures to be mutually detectable. To satisfy (6.2.23) for the first failure detection system and maintain mutual detectability, it is necessary to find a new event vector f_{ar} such that f_i, f_o, f_β, f_{ar} are output separable and the failures associated with f_i, f_o, f_β, f_{ar} satisfy Theorem 3.3.1--even though f_{ar} may not correspond to a physically meaningful failure. The search is essentially trial and

error. Equation (6.1.32) and the requirement made by Theorem 3.3.1 suggest that if $\overline{R}_{ar}$ were two-dimensional; f_i, f_o, f_β, f_{ar} might be mutually detectable. By starting with the null space of C and defining:

$$\xi_{ar} = e_{52} = \eta(C)$$

$$f_{ar} = A\xi_{ar} = \begin{bmatrix} 1 \\ -\beta_\varphi \\ -\beta_\varphi \\ 0 \\ 0 \end{bmatrix} \qquad (6.2.24)$$

it is possible to show that f_{ar} as defined in (6.2.24) has the desired two-dimensional detection space:

$$\overline{R}_{ar} = [f_{ar} \quad g_{ar}] = [f_{ar} \quad \xi_{ar}] \qquad (6.2.25)$$

Furthermore; f_i, f_o, f_β, f_{ar} are output separable--which satisfies (6.2.23)--and mutually detectable. It can be shown from the failure modeling considerations in Section 4.1.1 that f_{ar} is the event vector for certain simultaneous, identical failures in the rate gyro and attitude gyro. These failures are indicated as failure ix) in the introduction to this subsection, but the fact that such failures exist is not of great importance. The first failure detection system is required to detect ix) along with i) - v), viii) because the algorithms for selecting $D_{\Delta K}$ are simplified; not because ix) is a likely failure to occur.

The process of adding failures for the second failure detection system is much simpler. Equation (6.2.23) is satisfied by requiring the failure detection system to detect aileron failures in addition to the gyro failures. The event vectors $\tilde{f}_i$ and $\tilde{f}_o$ corresponding to inboard and outboard aileron failures can be defined for the augmented reference model $(\tilde{A}, \tilde{C})$ in the same manner in which $\tilde{\overline{R}}_r$ was determined:

$$\tilde{\tilde{f}}_i = \tilde{\tilde{R}}_i = \tilde{\tilde{g}}_i = e_{64}$$

$$\tilde{\tilde{f}}_o = \tilde{\tilde{R}}_o = \tilde{\tilde{g}}_o = e_{65} \qquad (6.2.26)$$

The event vectors $\tilde{\tilde{f}}_a$, $\tilde{\tilde{f}}_r$, $\tilde{\tilde{f}}_i$, $\tilde{\tilde{f}}_o$ are output separable and therefore satisfy (6.2.23). From the fact that the detection spaces satisfy:

$$\text{rk}\,\tilde{\tilde{R}}_a + \text{rk}\,\tilde{\tilde{R}}_r + \text{rk}\,\tilde{\tilde{R}}_i + \text{rk}\,\tilde{\tilde{R}}_o = 6 \qquad (6.2.27)$$

it can be shown (Theorem 3.3.1) that the gyro failures and aileron failures are detectable by a single failure detection system. Since $\tilde{\tilde{f}}_\beta$ is not included in $\tilde{\tilde{f}}_a$, $\tilde{\tilde{f}}_r$, $\tilde{\tilde{f}}_i$, $\tilde{\tilde{f}}_o$; the failure detection system cannot detect dynamics changes or aileron sensor failures.

The stochastic design procedure is detailed only for the failure detection system for $\tilde{\tilde{f}}_a$, $\tilde{\tilde{f}}_r$, $\tilde{\tilde{f}}_i$, $\tilde{\tilde{f}}_o$ in the remainder of this subsection. The considerations involved for the remaining failure detection system are similar to those discussed below.

The first step in the algorithms of Section 5.2 is to find matrices T and T_m which transform $(\tilde{\tilde{A}}, \tilde{\tilde{C}})$ into the base normal form for the failures associated with $\tilde{\tilde{f}}_a$, $\tilde{\tilde{f}}_r$, $\tilde{\tilde{f}}_i$, $\tilde{\tilde{f}}_o$. Equation (6.2.25) guarantees that T and T_m can be defined without using Appendix A.4. From (6.1.62), (6.1.65), (6.2.26) and (4.4.1):

$$T = [\tilde{\tilde{g}}_a \; \tilde{\tilde{A}}\tilde{\tilde{g}}_a \; \tilde{\tilde{g}}_r \; \tilde{\tilde{A}}\tilde{\tilde{g}}_r \; \tilde{\tilde{g}}_i \; \tilde{\tilde{g}}_o] = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{\beta_\varphi} & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ -1 & \beta_\varphi & -\frac{1}{\beta_\varphi} & 1 & 0 & 0 \end{bmatrix}$$

$$T^{-1} = \begin{bmatrix} \beta_\varphi & 1 & 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\beta_\varphi & \beta_\varphi & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

$$T_m = T_m^{-1} = [\tilde{C}\tilde{A}\tilde{g}_a \;\; \tilde{C}\tilde{A}\tilde{g}_r \;\; \tilde{C}\tilde{g}_i \;\; \tilde{C}\tilde{g}_o] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

(6.2.28)

If the aileron failures had not been added to the set of failures to monitor, Appendix A.4 would have been needed to determine vectors $\tilde{w}_3$ and $\tilde{w}_4$ to replace $\tilde{g}_i$ and $\tilde{g}_o$ in (6.2.28). In this particular example, $\tilde{g}_i$ and $\tilde{g}_o$ are the vectors which would have been selected by Appendix A.4 and the transformation matrices would not have changed. In general, however, Appendix A.4 does not yield additional detection spaces.

Given T and T_m, the base normal canonical form for $(\tilde{A}, \tilde{C})$ follows directly from (6.2.28):

$$\hat{A} = T^{-1}\tilde{A}T = \left[\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & \alpha_i & \alpha_o \\ 1 & -\beta_\varphi & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\beta_\varphi & \alpha_i & \alpha_o \\ \hline 0 & -k_i k_a & 0 & -k_i k_r & -\tau_i & 0 \\ 0 & -k_o k_a & 0 & -k_o k_r & 0 & -\tau_o \end{array}\right] = \left[\begin{array}{c|c|c|c} \hat{A}'_{aa} & 0 & x & x \\ \hline 0 & \hat{A}'_{rr} & x & x \\ \hline x & x & \hat{A}'_{ii} & 0 \\ \hline x & x & 0 & \hat{A}'_{oo} \end{array}\right]$$

$$\hat{C} = T_m^{-1} \tilde{\tilde{C}} T = \left[\begin{array}{cc|cc|cc} 0 & 1 & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 1 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right] = \left[\begin{array}{cc|cc|cc} \hat{c}_a & & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & \hat{c}_r & & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & \hat{c}_i & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & \hat{c}_o \end{array}\right]$$

$$\hat{D}_{\Delta K} = T^{-1} \tilde{\tilde{D}}_{\Delta K} T_m = \hat{D}_{\Delta FR} + \hat{D}_{\Delta \Psi K} + \hat{D}_{\Delta \Gamma K}$$

$$\hat{x}_m(t) = T^{-1} \tilde{\tilde{x}}_m(t) = \left[\begin{array}{c} \hat{x}_{ma}(t) \\ \hline \hat{x}_{mr}(t) \\ \hline \hat{x}_{mi}(t) \\ \hline \hat{x}_{mo}(t) \end{array}\right] = \left[\begin{array}{c} \beta_\varphi \varphi_{ma}(t) + \dot{\varphi}_m(t) - \dot{\varphi}_{m0}(t) \\ \varphi_{ma}(t) \\ \hline \beta_\varphi (\dot{\varphi}_{mr}(t) - \dot{\varphi}_m(t)) \\ \dot{\varphi}_{mr}(t) \\ \hline \delta_{mi}(t) \\ \hline \delta_{mo}(t) \end{array}\right] \qquad (6.2.29)$$

The x's in $\hat{A}$ denote non-zero off-diagonal submatrices. Because the set of failures was increased to satisfy (6.2.23), the $\hat{A}'_{\Gamma\Gamma}$ subsystem in (4.4.4) is missing from (6.2.29). Each of the partitions of $\hat{x}_m(t)$ is associated with the detection space with the same subscript.-- $\varphi_{ma}(t)$ is the estimate of $\varphi_a(t)$ etc. The transformation of $\tilde{\tilde{R}}_a$, $\tilde{\tilde{R}}_r$, $\tilde{\tilde{R}}_i$, and $\tilde{\tilde{R}}_o$ into base normal form results in:

$$\hat{\tilde{R}}_a = T^{-1} \tilde{\tilde{R}}_a = [\hat{f}_a \; \hat{g}_a] = [\hat{e}_{62} + \beta_\varphi \hat{e}_{61} | \hat{e}_{61}] = \left[\begin{array}{cc} \beta_\varphi & 1 \\ 1 & 0 \\ \hline 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array}\right] = \left[\begin{array}{cc} \hat{f}_{aa} & \hat{g}_{aa} \\ \hline 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{array}\right]$$

$$\hat{\tilde{R}}_r = T^{-1} \tilde{\tilde{R}}_r = [\hat{f}_r \; \hat{g}_r] = [\hat{e}_{64} + \beta_\varphi \hat{e}_{63} | \hat{e}_{63}]$$

$$\tilde{\tilde{R}}_i = T^{-1} \tilde{\tilde{R}}_i = \hat{f}_i = \hat{g}_i = \hat{e}_{65}$$

$$\tilde{\tilde{R}}_o = T^{-1} \tilde{\tilde{R}}_o = \hat{f}_o = \hat{g}_o = \hat{e}_{66}$$

(6.2.30)

The transformed detection generators $\hat{g}_a$, $\hat{g}_r$, $\hat{g}_i$, and $\hat{g}_o$ are unit vectors. The definition of T_m insures that the output directions associated with the event vectors $\hat{f}_a$, $\hat{f}_r$, $\hat{f}_i$, and $\hat{f}_o$ are also unit vectors:

$$\begin{aligned} \hat{C}\,\hat{f}_a &= \hat{e}_{41} \\ \hat{C}\,\hat{f}_r &= \hat{e}_{42} \\ \hat{C}\,\hat{f}_i &= \hat{e}_{43} \\ \hat{C}\,\hat{f}_o &= \hat{e}_{44} \end{aligned} \tag{6.2.31}$$

With the base normal canonical form established, the next step in choosing $\hat{D}_{\Delta K}$ is to determine $\hat{D}_{FR}$ in (5.2.22). A sufficient condition for aileron and gyro failures to generate unidirectional output errors for $(\hat{A} - \hat{D}_{FR}\,\hat{C},\ \hat{C})$ is for $\hat{A} - \hat{D}_{FR}\,\hat{C}$ to be block diagonal. A satisfactorily $\hat{D}_{FR}$ always exists and (4.4.17) expresses it in terms of the elements in the last column of each of the off-diagonal submatrices of $\hat{A}$:

$$\hat{D}_{FR} = \left[\begin{array}{cccc} 0 & 0 & \alpha_i & \alpha_o \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & \alpha_i & \alpha_o \\ \hline -k_i k_a & -k_i k_r & 0 & 0 \\ \hline -k_o k_a & -k_o k_r & 0 & 0 \end{array}\right] = \left[\begin{array}{c} \hat{D}_{Fa} \\ \hline \hat{D}_{Fr} \\ \hline \hat{D}_{Fi} \\ \hline \hat{D}_{Fo} \end{array}\right] \tag{6.2.32}$$

With (6.2.32), the only non-zero elements of $\hat{A} - \hat{D}_{FR}\,\hat{C}$ are those of the block diagonal submatrices $\hat{A}'_{aa}$, $\hat{A}'_{rr}$, $\hat{A}'_{ii}$, and $\hat{A}'_{oo}$; and these submatrices are unchanged:

$$\hat{A} - \hat{D}_{FR}\hat{C} = \left[\begin{array}{cc|cc|cc} 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -\beta_\varphi & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\beta_\varphi & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -\tau_i & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -\tau_o \end{array}\right] = \left[\begin{array}{c|c|c|c} \hat{A}'_{aa} & 0 & 0 & 0 \\ \hline 0 & \hat{A}'_{rr} & 0 & 0 \\ \hline 0 & 0 & \hat{A}'_{ii} & 0 \\ \hline 0 & 0 & 0 & \hat{A}'_{oo} \end{array}\right] \quad (6.2.33)$$

To see that $\hat{D}_{FR}$ satisfies the failure detection constraint, consider an attitude gyro failure as modeled by (6.1.9). The state estimation error due to the failure is:

$$\dot{\hat{\epsilon}}(t) = (\hat{A} - \hat{D}_{FR}\hat{C})\,\hat{\epsilon}(t) + \hat{f}_a(\Delta a_{11}\,\varphi_a(t) - \dot{\varphi}(t)) \;:\; \hat{\epsilon}(0) \text{ given} \quad (6.2.34)$$

The error generated by the driving term in (6.2.34) remains in the controllable subspace of $\hat{f}_a$. This subspace is spanned by the columns of:

$$[\hat{f}_a \,|\, (\hat{A} - \hat{D}_{FR}\hat{C})\hat{f}_a \,|\, (\hat{A} - \hat{D}_{FR}\hat{C})^2\hat{f}_a \;..] = [\hat{e}_{62} + \beta_\varphi\,\hat{e}_{61} \,|\, 0 \quad 0 \ldots] \quad (6.2.35)$$

The controllable subspace is one-dimensional--an attitude gyro failure affects only the estimate of the state component associated with the direction $\hat{e}_{62} + \beta_\varphi\,\hat{e}_{61}$. A transformation back to the original coordinate system using (6.2.29) verifies that this is the estimate of $\varphi_m(t)$. The output estimation error is restricted to the output subspace defined by $\hat{C}$ and (6.2.35):

$$\hat{C}[\hat{e}_{62} + \beta_\varphi\,\hat{e}_{61} \,|\, 0 \quad 0 \ldots] = [\hat{e}_{41} \quad 0 \quad 0 \quad \ldots] \quad (6.2.36)$$

and is clearly unidirectional. Similar results follow for the remaining failures.

The last step in the choice of $\hat{D}_{\Delta K}$ is to choose $\hat{D}_{\Delta \Psi K}$ and $\hat{D}_{\Delta \Gamma}$. The absence of a $\hat{x}_\Gamma(t)$ partition of $\hat{x}(t)$ in (6.2.29) means that $\hat{D}_{\Delta \Gamma K}$

is zero (Section 4.4). The only constraint on the choice of $\hat{D}_{\Delta\Psi K}$ is that it must not alter the output error unidirectionality provided by $\hat{D}_{FR}$. It is not necessary that $\hat{D}_{\Delta\Psi K}$ satisfy the same continuous-discrete approximation defined for $\hat{D}_{\Delta FR}$ in (5.2.22), but this approximation is useful for interpreting the failure detection constraint on $\hat{D}_{\Delta\Psi K}$:

$$\hat{D}_{\Delta\Psi K} = \Delta t \hat{D}_{\Psi K} \tag{6.2.37}$$

Equation (4.4.20) defines $\hat{D}_{\Psi K}$ to complement $\hat{D}_{FR}$. Where an element of $\hat{D}_{FR}$ is required to equal an element in one of the off-diagonal submatrices of $\hat{A}$, the corresponding element of $\hat{D}_{\Psi K}$ must be zero. Where an element of $\hat{D}_{FR}$ was set arbitrarily equal to zero, that element in $\hat{D}_{\Psi K}$ is freely assignable:

$$\hat{D}_{\Psi K} = \left[\begin{array}{cccc} & 0 & 0 & 0 \\ d_{\Psi Ka} & 0 & 0 & 0 \\ \hline 0 & & 0 & 0 \\ & d_{\Psi Kr} & & \\ 0 & & 0 & 0 \\ \hline 0 & 0 & d_{\Psi Ki} & 0 \\ \hline 0 & 0 & 0 & d_{\Psi Ko} \end{array}\right] = \left[\begin{array}{c} D_{\Psi Ka} \\ \hline D_{\Psi Kr} \\ \hline D_{\Psi Ki} \\ \hline D_{\Psi Ko} \end{array}\right] \tag{6.2.38}$$

If the non-zero elements of $\hat{D}_{\Psi K}$ are expressed as functions of the coefficients in the last columns of $\hat{A}'_{aa}$, $\hat{A}'_{rr}$, $\hat{A}'_{ii}$, and $\hat{A}'_{oo}$ in (6.2.29):

$$\begin{aligned} d_{\Psi Ka} &= \begin{bmatrix} p_{a1} \\ -\beta_{\varphi} + p_{a2} \end{bmatrix} \\ d_{\Psi Kr} &= \begin{bmatrix} p_{r1} \\ -\beta_{\varphi} + p_{r2} \end{bmatrix} \\ d_{\Psi Ki} &= -\tau_i + p_{i1} \\ d_{\Psi Ko} &= -\tau_o + p_{o1} \end{aligned} \tag{6.2.39}$$

then $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}$ is block diagonal and expressible as a function of the freely assignable coefficients p_{a1}, p_{a2}, p_{r1}, p_{r2}, p_{i1}, and p_{o1}:

$$\hat{A}-(\hat{D}_{FR}+\hat{D}_{\Psi K})\hat{C} = \left[\begin{array}{cc|cc|c|c} 0 & -p_{a1} & 0 & 0 & 0 & 0 \\ 1 & -p_{a2} & 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & -p_{r1} & 0 & 0 \\ 0 & 0 & 1 & -p_{r2} & 0 & 0 \\ \hline 0 & 0 & 0 & 0 & -p_{i1} & 0 \\ \hline 0 & 0 & 0 & 0 & 0 & -p_{o1} \end{array}\right] = \left[\begin{array}{c|c|c|c} \hat{A}_{aa} & 0 & 0 & 0 \\ \hline 0 & \hat{A}_{rr} & 0 & 0 \\ \hline 0 & 0 & \hat{A}_{ii} & 0 \\ \hline 0 & 0 & 0 & \hat{A}_{oo} \end{array}\right] \qquad (6.2.40)$$

Arguments similar to (6.2.35) and (6.2.36) verify that the monitored failures generate unidirectional output errors for $(\hat{A}-(\hat{D}_{FR}+\hat{D}_{\Psi K})\hat{C},\ \hat{C})$. The characteristic equations of $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}$ is :

$$|\lambda\hat{E}_6 - \hat{A} + (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}| = (\lambda^2 + p_{a2}\lambda + p_{a1})(\lambda^2 + p_{r2}\lambda + p_{r1})(\lambda + p_{i1})(\lambda + p_{o1})$$

$$= \Psi_a(\lambda)\,\Psi_r(\lambda)\,\Psi_i(\lambda)\,\Psi_o(\lambda) \qquad (6.2.41)$$

where $\Psi_a(\lambda)$ is the minimal polynomial of $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K}\hat{C})$ associated with $\hat{\bar{R}}_a$ etc. Equations (6.2.39) and (6.2.41) together guarantee the freedom to select the eigenvalues of $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}$.

Section 5.2 provides three algorithms for selecting $\hat{D}_{\Delta\Psi K}$ to minimize the variance of the state estimates provided by (5.2.5) and (5.2.6). All these involve approximations of the stochastic system model defined in Section 6.2.1. Two of the algorithms are based on the continuous measurement stochastic model developed in Section 6.2.1. The free parameters of $\hat{D}_{\Psi K}$ in (6.2.38) are used in these algorithms to choose $\hat{D}_{FR} + \hat{D}_{\Psi K}$ as an optimal detection filter-state estimator for the continuous state estimator:

$$\dot{\hat{x}}_m(t) = \hat{A}\,\hat{x}_m(t) + (\hat{D}_{FR} + \hat{D}_{\Psi K})(\hat{y}(t) - \hat{y}_m(t))$$
$$= (\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\,\hat{C})\,\hat{x}_m(t) + (\hat{D}_{FR} + \hat{D}_{\Psi K})\,\hat{y}(t)$$
$$\hat{y}_m(t) = \hat{C}\,\hat{x}_m(t) \qquad (6.2.42)$$

$\hat{D}_{\Delta\Psi K}$ is then approximated by $\Delta t\hat{D}_{\Psi K}$ and $\hat{D}_{\Delta K}$ becomes:

$$\hat{D}_{\Delta K} = \Delta t\,(\hat{D}_{FR} + \hat{D}_{\Psi K}) \qquad (6.2.43)$$

The only distinction between the two algorithms is that one optimizes $d_{\Psi Ka}$, $d_{\Psi Kr}$, $d_{\Psi Ki}$, and $d_{\Psi Ko}$ using the Kalman filter algorithm of Section 5.1.2 and the other uses the parameter optimization technique of Section 5.1.3.

The third algorithm optimizes $\hat{D}_{\Delta\Psi K}$ by working directly with the sampled-data stochastic model of Section 6.2.1 and the sampled-data state estimators presented at the beginning of this subsection. Failure detection constraints equivalent to (6.2.38) are imposed on $\hat{D}_{\Delta\Psi K}$ by replacing $\hat{\Phi}(t,\tau)$ in both the stochastic model and state estimator by the state transition matrix for $\hat{A} - \hat{D}_{FR}\,\hat{C}$:

$$\hat{\Phi}'(t,\tau) = e^{(\hat{A} - \hat{D}_{FR}\,\hat{C})\,(t-\tau)} \qquad (6.2.44)$$

The above approximation results in a block diagonal state transition matrix:

$$\hat{\Phi}'(t,\tau) = \begin{bmatrix} \hat{\Phi}'_{aa} & 0 & 0 & 0 \\ 0 & \hat{\Phi}'_{rr} & 0 & 0 \\ 0 & 0 & \hat{\Phi}'_{ii} & 0 \\ 0 & 0 & 0 & \hat{\Phi}_{oo} \end{bmatrix} \qquad (6.2.45)$$

where $\hat{\Phi}'_{aa}(t,\tau)$ has the same dimension as $\hat{A}'_{aa}$ and is approximately the corresponding block of $\hat{\Phi}(t,\tau)$ for small time increments etc. The general form for $\hat{D}_{\Delta\Psi K}$ is then the same as that of $\hat{D}_{\Psi K}$ in (6.2.38):

$$\hat{D}_{\Delta\Psi K} = \left[\begin{array}{cccc} d_{\Delta\Psi Ka} & 0 & 0 & 0 \\ 0 & d_{\Delta\Psi Kr} & 0 & 0 \\ 0 & 0 & d_{\Delta\Psi Ki} & 0 \\ 0 & 0 & 0 & d_{\Delta\Psi Ko} \end{array}\right] = \left[\begin{array}{c} D_{\Delta\Psi Ka} \\ D_{\Delta\Psi Kr} \\ D_{\Delta\Psi Ki} \\ D_{\Delta\Psi Ko} \end{array}\right] \qquad (6.2.46)$$

The algorithm optimizes $d_{\Delta\Psi Ka}$, $d_{\Delta\Psi Kr}$; $d_{\Delta\Psi Ki}$, and $d_{\Delta\Psi Ko}$ using the discrete Kalman filter of Section 5.2.

The two algorithms based on the continuous measurement stochastic model are illustrated in the remainder of this section. With (6.2.44) - (6.2.46), the third algorithm is procedurally analogous to the continuous Kalman optimization but uses the discrete formulation. The effectiveness of the detection filter-state estimates for the continuous and discrete formulations is compared in Section 6.4.

The two algorithms for optimizing $\hat{D}_{\Psi K}$ require that the expressions for G_1, G_2, and G_{12} given in (6.2.19), (6.2.21), and (6,2.22) be restated for the augmented reference model ($\tilde{\tilde{A}}$, $\tilde{\tilde{C}}$). The added state component has no physical counterpart and is not driven by system dynamics. It follows that $\tilde{\tilde{G}}_1$ is defined similarly to $\tilde{\tilde{B}}$; $\tilde{\tilde{G}}_2$ is unchanged:

$$\begin{aligned} \tilde{\tilde{G}}_1 &= \left[\begin{array}{c} G_1 \\ \hline 0\ 0\ 0 \end{array}\right] \\ \tilde{\tilde{G}}_2 &= G_2 \qquad (6.2.47) \\ \tilde{\tilde{G}}_{12} &= G_{12} \end{aligned}$$

The matrices $\tilde{\tilde{G}}_1$ and $\tilde{\tilde{G}}_2$ can be transformed into base normal canonical form; no transformation for $\tilde{\tilde{G}}_{12}$ is necessary:

$$\hat{G}_1 = T^{-1}\,\tilde{\tilde{G}}_1 = \begin{bmatrix} \sqrt{W_\beta} & 0 & 0 \\ 0 & 0 & 0 \\ \hline 0 & 0 & 0 \\ \sqrt{W_\beta} & 0 & 0 \\ \hline 0 & \sqrt{\Delta t}\,\alpha_L k_i k_a \sigma'_{1a\Delta} & \sqrt{\Delta t}\,\alpha_L k_i k_r \sigma'_{1r\Delta} \\ \hline 0 & \sqrt{\Delta t}\,\alpha_L k_o k_a \sigma'_{1a\Delta} & \sqrt{\Delta t}\,\alpha_L k_o k_r \sigma'_{1r\Delta} \end{bmatrix} = \begin{bmatrix} \hat{G}_{1a} \\ \hline \hat{G}_{1r} \\ \hline \hat{G}_{1i} \\ \hline \hat{G}_{1o} \end{bmatrix}$$

$$\hat{G}_2 = T_m^{-1}\,\tilde{\tilde{G}}_2 = \sqrt{\Delta t} \begin{bmatrix} \sigma'_{1a\Delta} & 0 & 0 & 0 \\ \hline 0 & \sigma'_{1r\Delta} & 0 & 0 \\ \hline 0 & 0 & \sigma_{2i\Delta} & 0 \\ \hline 0 & 0 & 0 & \sigma_{2o\Delta} \end{bmatrix} = \begin{bmatrix} \hat{G}_{2a} \\ \hline \hat{G}_{2r} \\ \hline \hat{G}_{2i} \\ \hline \hat{G}_{2o} \end{bmatrix} \qquad (6.2.48)$$

With these definitions and the assumption that the continuous measurement stochastic model is precise, the error in the state estimates provided by (6.2.42) satisfies (5.1.4):

$$\dot{\hat{\epsilon}}(t) = (\hat{A} - (\hat{D}_{FR} + \hat{D}_{\psi K})\,\hat{C})\,\hat{\epsilon}(t) - \hat{D}_{\psi K}\,\hat{G}_2\,\eta_2(t) - \hat{D}_{FR}\,\hat{G}_2\,\eta_2(t) + \hat{G}_1\eta_1(t)$$

$$\hat{\epsilon}'(t) = \hat{C}\,\hat{\epsilon}(t) + \hat{G}_2\,\eta_2(t) \qquad (5.1.4)$$

Because $\hat{D}_{FR}$ is independent of the choice of $d_{\psi Ka}$, $d_{\psi Kr}$, $d_{\psi Ki}$, and $d_{\psi Ko}$; Section 5.1.1 defines a new system noise from the last two terms in (5.1.4):

$$\eta_1'(t) = \hat{G}_1\,\eta_1(t) - \hat{D}_{FR}\,\hat{G}_2\,\eta_2(t) = \begin{bmatrix} \eta'_{1a} \\ \eta'_{1r} \\ \eta'_{1i} \\ \eta'_{1o} \end{bmatrix} \qquad (6.2.49)$$

It is shown in Section 6.4.3 that the second term in (6.2.49) can become the dominant system noise for the error equations. $\eta_1'(t)$ is a zero mean, white Gaussian random process with a correlation function:

$$E[\eta_1'(t)\,\eta_1'(\tau)] = \delta(t-\tau)\begin{bmatrix} \sigma^2_{\beta io} & 0 & 0 & \sigma^2_{\beta io} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \sigma^2_{\beta io} & 0 & 0 & \sigma^2_{\beta io} & 0 & 0 \\ 0 & 0 & 0 & 0 & (1-3\alpha_L^2)\,k_i^2\,\sigma^2_{ar} & (1-3\alpha_L^2)\,k_i k_o\,\sigma^2_{ar} \\ 0 & 0 & 0 & 0 & (1-3\alpha_L^2)\,k_i k_o\,\sigma^2_{ar} & (1-3\alpha_L^2)\,k_o^2\,\sigma^2_{ar} \end{bmatrix} \cdot \delta(t-\tau)\begin{bmatrix} \hat{G}'_{aa} & x & 0 & 0 \\ x & \hat{G}'_{rr} & 0 & 0 \\ 0 & 0 & \hat{G}'_{ii} & x \\ 0 & 0 & x & \hat{G}'_{co} \end{bmatrix}$$

$$= \delta(t-\tau)\,\hat{G}'_{11} \qquad (6.2.50)$$

where:

$$\sigma^2_{\beta io} = W_\beta + \alpha_i^2\,\sigma^2_{2i\Delta} + \alpha_o^2\,\sigma^2_{2o\Delta}$$

$$\sigma^2_{ar} = \Delta t\,(k_a^2\,\sigma'^2_{1a\Delta} + k_r^2\,\sigma'^2_{1r\Delta})$$

Equation (6.2.48) implies that $\eta_1'(t)$ and $\eta_2(t)$ are in general not independent. Their covariance satisfies:

$$E[\eta_1'(t)\,\eta_2(t)^T\,\hat{G}_2^T] = \delta(t-\tau)\begin{bmatrix} 0 & 0 & -\Delta t\alpha_i\,\sigma^2_{2i\Delta} & -\Delta t\alpha_o\,\sigma^2_{2o\Delta} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -\Delta t\alpha_i\,\sigma^2_{2i\Delta} & -\Delta t\alpha_o\,\sigma^2_{2o\Delta} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \cdot \delta(t-\tau)\begin{bmatrix} \hat{G}'_{12a} & 0 & x & x \\ 0 & \hat{G}'_{12r} & x & x \\ 0 & 0 & \hat{G}'_{12i} & 0 \\ 0 & 0 & 0 & \hat{G}'_{12o} \end{bmatrix} = \delta(t-\tau)\hat{G}'_{12} \qquad (6.2.51)$$

The block diagonal nature of $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}$ in (6.2.40) and the partitions in (6.2.29), (6.2.48), and (6.2.50) allow the error equations (5.1.4) to be decoupled into four sets of equations. With $\hat{\epsilon}(t)$ partitioned in the same manner as $\hat{x}_m(t)$ in (6.2.29):

$$\hat{\epsilon}(t) = \begin{bmatrix} \hat{\epsilon}_a(t) \\ \hat{\epsilon}_r(t) \\ \hat{\epsilon}_i(t) \\ \hat{\epsilon}_o(t) \end{bmatrix} \tag{6.2.52}$$

and (6.2.33) substituted for $\hat{A}\text{-}\hat{D}_{FR}\,\hat{C}$, the four error equations become:

$$\dot{\hat{\epsilon}}_a(t) = (\hat{A}'_{aa} - d_{\Psi Ka}\,\hat{c}_a)\,\hat{\epsilon}_a(t) - d_{\Psi Ka}\,\hat{G}_{2a}\,\eta_2(t) + \eta'_{1a}(t)$$
$$\hat{\epsilon}'_a(t) = \hat{c}_a\,\hat{\epsilon}_a(t) + \hat{G}_{2a}\,\eta_2(t) \tag{6.2.53a}$$

$$\dot{\hat{\epsilon}}_r(t) = (\hat{A}'_{rr} - d_{\Psi Kr}\,\hat{c}_r)\,\hat{\epsilon}_r(t) - d_{\Psi Kr}\,\hat{G}_{2r}\,\eta_2(t) + \eta'_{1r}(t)$$
$$\hat{\epsilon}'_r(t) = \hat{c}_r\,\hat{\epsilon}_r(t) + \hat{G}_{2r}\,\eta_2(t) \tag{6.2.53r}$$

$$\dot{\hat{\epsilon}}_i(t) = (\hat{A}'_{ii} - d_{\Psi Ki}\,\hat{c}_i)\,\hat{\epsilon}_i(t) - d_{\Psi Ki}\,\hat{G}_{2i}\,\eta_2(t) + \eta'_{1i}(t)$$
$$\hat{\epsilon}'_i(t) = \hat{c}_i\,\hat{\epsilon}_i(t) + \hat{G}_{2i}\,\eta_2(t) \tag{6.2.53i}$$

$$\dot{\hat{\epsilon}}_o(t) = (\hat{A}'_{oo} - d_{\Psi Ko}\,\hat{c}_o)\,\hat{\epsilon}_o(t) - d_{\Psi Ko}\,\hat{G}_{2o}\,\eta_2(t) + \eta'_{1o}(t)$$
$$\hat{\epsilon}'_o(t) = \hat{c}_o\,\hat{\epsilon}_o(t) + \hat{G}_{2o}\,\eta_2(t) \tag{6.2.53o}$$

These equations associate each of the vectors $d_{\Psi Ka}$, $d_{\Psi Kr}$, $d_{\Psi Ki}$, and $d_{\Psi Ko}$ uniquely with the partition of $\hat{\epsilon}(t)$ with the appropriate subscript. The fact that the four equations are decoupled implies that the choice of $d_{\Psi Ka}$ affects the accuracy of $\hat{x}_{ma}(t)$, but not of $\hat{x}_{mr}(t)$ etc.

The four error equations are decoupled but they are not statistically independent. The non-diagonal form of $\hat{G}'_{11}$ and $\hat{G}'_{12}$ imply that the residuals $\hat{\epsilon}'_a(t)$, $\hat{\epsilon}'_r(t)$, $\hat{\epsilon}'_i(t)$, and $\hat{\epsilon}'_o(t)$ are correlated. It follows that the error $\hat{\epsilon}'_a$ in the estimate of the nominal attitude gyro output contains information about the errors in the estimates of the aileron deflections etc. A true minimum variance state estimator would use this information to improve the estimates of all of the state variables. An optimal detection filter-state estimator cannot do this because of the requirement to properly isolate failures. If an attitude gyro failure modeled by (6.1.9)

occurs, the partition of $\hat{f}_a$ in (6.2.30) implies that only the error equations for $\hat{\epsilon}_a(t)$ are driven by the failure:

$$\dot{\hat{\epsilon}}_a(t) = (\hat{A}'_{aa} - d_{\psi Ka}\,\hat{c}_a)\,\hat{\epsilon}_a(t) - d_{\psi Ka}\,\hat{G}_{2a}\,\eta_2(t) + \eta'_{1a}(t) + \hat{f}_{aa}\,n(t)$$
$$\hat{\epsilon}'_a(t) = \hat{c}_a\,\hat{\epsilon}_a(t) + \hat{G}_{2a}\,\eta_2(t) \qquad (6.2.54)$$

If $\hat{\epsilon}'_a(t)$ were used to decrease the variances of $\hat{\epsilon}'_r(t)$, $\hat{\epsilon}'_i(t)$, and $\hat{\epsilon}'_o(t)$; it would appear as a driving term in (6.2.53r), (6.2.53i), and (6.2.53o). The error in $\hat{\epsilon}'_a(t)$ resulting from the attitude gyro failure would then propagate through the state estimates and the desired unidirectionality of failure-induced output estimation errors would be lost. It follows that in optimizing the state estimates produced by the failure detection system, it is necessary to consider the four error equations (6.2.53a) - (6.2.53o) to be decoupled and to disregard their cross-correlations.

Equations (6.2.53a) - (6.2.53o) are the classic error equations for state estimators monitoring a system with a single continuous measurement. The four systems can be denoted by $(\hat{A}'_{aa}, \hat{c}_a)$, $(\hat{A}'_{rr}, \hat{c}_r)$, $(\hat{A}'_{ii}, \hat{c}_i)$, and $(\hat{A}'_{oo}, \hat{c}_o)$. The vectors $d_{\psi Ka}$, $d_{\psi Kr}$, $d_{\psi Ki}$, and $d_{\psi Ko}$ are the general output feedback matrices for such systems. The state estimates associated with each system are the components of the partition of $\hat{x}_m(t)$ with the same subscript. The continuous Kalman filter and parameter optimization techniques can be applied to each of the four subscripts to determine the optimal choice of $\hat{D}_{\psi K}$. Both algorithms are illustrated in the following paragraphs for $(\hat{A}'_{aa}, \hat{c}_a)$ and $d_{\psi Ka}$. Because additional failures were selected in order to satisfy (6.2.23), the algorithms yield identical results. When (6.2.23) is not satisfied, the Kalman algorithm of Section 5.1.2 generally requires a state space augmentation and the Kalman and parameter optimization techniques yield different results.

The continuous Kalman filter for $(\hat{A}'_{aa}, \hat{c}_a)$ minimizes the variance of the state estimate $\hat{x}_{ma}(t)$. The steady state covariance matrix $\hat{P}_{\infty a}$ is the solution of (5.1.17). Because $\hat{G}'_{12a}$ in (6.2.51) is zero, $\hat{G}_{2a}\,\eta_2(t)$ and $\eta'_{1a}(t)$ in (6.2.53a) are uncorrelated and $\hat{P}_{\infty a}$ satisfies:

$$\hat{A}'_{aa}\,\hat{P}_{\infty a} + \hat{P}_{\infty a}\,\hat{A}'^{T}_{aa} + \hat{G}'_{aa} - \hat{P}_{\infty a}\hat{c}_a^T\left[\hat{G}_{2a}\,\hat{G}_{2a}^T\right]^{-1}\hat{c}_a\,\hat{P}_{\infty a} = 0 \tag{6.2.55}$$

Substitution of the appropriate partitions from (6.2.33), (6.2.48), and (6.2.50) yields:

$$\hat{P}_{\infty a} = \Delta t \begin{bmatrix} \sigma_{\beta io}\,\sigma'_{1a\Delta}\sqrt{\beta_\varphi^2 + \dfrac{2\sigma_{\beta io}}{\sigma'_{1a\Delta}}} & \sigma_{\beta io}\,\sigma'_{1a\Delta} \\ \sigma_{\beta io}\,\sigma'_{1a\Delta} & \sigma'^2_{1a\Delta}\left(-\beta_\varphi + \sqrt{\beta_\varphi^2 + \dfrac{2\sigma_{\beta io}}{\sigma'_{ia\Delta}}}\right) \end{bmatrix} \tag{6.2.56}$$

The Kalman filter then satisfies (5.1.16):

$$d_{\Psi Ka} = \hat{P}_{\infty a}\,\hat{c}_a^T\left[\hat{G}_{2a}\,\hat{G}_{2a}^T\right]^{-1} = \begin{bmatrix} \dfrac{\sigma_{\beta io}}{\sigma'_{1a\Delta}} \\ -\beta_\varphi + \sqrt{\beta_\varphi^2 + \dfrac{2\sigma_{\beta io}}{\sigma'_{1a\Delta}}} \end{bmatrix} \tag{6.2.57}$$

Comparison of $d_{\Psi Ka}$ with (6.2.39) indicates that the coefficients of the minimal polynominal $\Psi_a(\cdot)$ are:

$$\begin{aligned} p_{a1} &= \frac{\sigma_{\beta io}}{\sigma'_{1a\Delta}} \\ p_{a2} &= \sqrt{\beta_\varphi^2 + \frac{2\sigma_{\beta io}}{\sigma'_{1a\Delta}}} \end{aligned} \tag{6.2.58}$$

The parameter optimization technique minimizes the variance of the output $\hat{c}_a\,\hat{x}_{ma}(t)$. Because the noise sources driving $(\hat{A}'_{aa},\ \hat{c}_a)$ are white, it yields the continous Kalman filter and therefore minimizes the variance of $\hat{x}_{ma}(t)$. The technique requires the Laplace transform of $\hat{c}_a\,\hat{\epsilon}_a(t)$. The general equation for the transform is given by (5.1.35). Since $\hat{R}_a$ has a rank of two, z from (5.1.34) is:

$$z = \begin{bmatrix} 1 \\ s \end{bmatrix} \qquad (6.2.59)$$

Substituting the appropriate variables into (5.1.35) yields the scalar equation:

$$\hat{c}_a \hat{\epsilon}_a(s) = \frac{1}{\Psi_a(s)} z^T(\eta'_{1a}(s) - d_{\psi Ka} \hat{G}_{2a} \eta_2(s))$$

$$= \frac{1}{\Psi_a(s)}\Big[\sqrt{W_\beta}\eta_{1\beta}(s) - \sqrt{\Delta t}\,\alpha_i \sigma_{2i\Delta} \eta_{2i}(s) - \sqrt{\Delta t}\,\alpha_o \sigma_{2o\Delta} \eta_{2o}(s) -$$

$$\sqrt{\Delta t}\,\sigma'_{1a\Delta}(p_{a1} - s(p_{a2} - \beta_\varphi))\eta_{2a}(s)\Big] \qquad (6.2.60)$$

All of the random processes in (6.2.60) are statistically independent and (2.2.2) implies that their power spectral densities are unity. The power spectral density of $\hat{c}_a \hat{\epsilon}_a(s)$ can then be found from (5.1.36):

$$c_a \hat{\Phi}_{aa}(s) c_a^T = \frac{1}{\Psi_a(s)\Psi_a(-s)}\Big[W_\beta + \Delta t\,\alpha_i^2 \sigma_{2i\Delta}^2 + \Delta t\,\alpha_o^2 \sigma_{2o\Delta}^2 +$$

$$\Delta t \sigma'^2_{1a\Delta}(p_{a1} - s(p_{a2} - \beta_\varphi))(p_{a1} + s(p_{a2} - \beta_\varphi))\Big]$$

$$(6.2.61)$$

This can be factored in polynomials of s and -s as in (5.1.37):

$$c_a \Phi_{aa}(s) c_a^T = \frac{(c_o + c_1 s)(c_o - c_1 s)}{\Psi_a(s)\Psi_a(-s)} \qquad (6.2.62)$$

where:

$$c_o^2 = W_\beta + \Delta t(\alpha_i^2 \sigma_{2i\Delta}^2 + \alpha_o \sigma_{2o\Delta}^2 + \sigma'^2_{1a\Delta} p_{a1}^2)$$

$$c_1^2 = \Delta t\,\sigma'^2_{1a\Delta}(\beta_\varphi - p_{a2})^2$$

Equation (6.2.62) is in the form required for the tables in Appendix B. From I_2 in the tables, the steady state output covariance $\hat{c}_a \hat{P}_{\infty a} \hat{c}_a^T$ is:

$$\hat{c}_a \hat{P}_{\infty a} \hat{c}_a^T = \frac{\Delta t\, \sigma'^2_{1a\Delta} (\beta_\omega - p_{a2})^2 + W_\beta + \Delta t(\alpha_i^2 \sigma^2_{2i\Delta} + \alpha_o^2 \sigma^2_{2o\Delta} + \sigma'^2_{1a\Delta} p^2_{a1})}{2 p_{a1} p_{a2}} \tag{6.2.63}$$

$\hat{c}_a \hat{P}_{\infty a} \hat{c}_a^T$ is an explicit function of the coefficients p_{a1} and p_{a2}. The optimal coefficients are the solution to the coupled algebraic equations:

$$\frac{\partial(\hat{c}_a \hat{P}_{\infty a} \hat{c}_a^T)}{\partial p_{a1}} = 0$$

$$\frac{\partial(\hat{c}_a \hat{P}_{\infty a} \hat{c}_a^T)}{\partial p_{a2}} = 0 \tag{6.2.64}$$

The solutions obtained are the same as (6.2.57) for the Kalman algorithm.

In general, when output stationarity is involved, the vectors of $\hat{D}_{\Delta K}$ as defined in (6.2.38) or (6.2.46) cannot be optimized independently. A suboptimal detection filter-state estimator can still be found but a variation on the two algorithms presented above is necessary. The failure detection system designed for the failures i) - v), viii), and ix) from Table 6.1 requires output stationarity in order to detect amplifier failures. It was noted in Section 6.1.1 and again in conjunction with (6.1.44) - (6.1.48) that the output error generated by an amplifier failure is similar to that generated by simultaneous aileron failures of a particular type. Because the failure detection system for the gyro failures has been required to monitor aileron failures in order to satisfy (6.2.23), output stationarity could be used to add amplifier failure monitoring to its responsibilities. The effect of this constraint would be that $d_{\Psi Ki}$ and $d_{\Psi Ko}$ could not be optimized independently. The necessary modification of the parameter optimization procedure is illustrated below.

Equation (6.2.41) defines the desired minimal polynomials $\Psi_i(\lambda)$ and $\Psi_o(\lambda)$ of $\hat{\bar{R}}_i$ and $\hat{\bar{R}}_o$ with respect to $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}$. The detection space for an amplifier failure can be defined from Table 6.1 and the augmented system $(\tilde{\hat{A}}, \tilde{\hat{C}})$:

$$\hat{\bar{R}}_{ka} = \hat{f}_{ka} = T^{-1} \tilde{\hat{R}}_{ka} = T^{-1} \begin{bmatrix} f_{ka} \\ 0 \end{bmatrix} = k_i \hat{e}_{65} + k_o \hat{e}_{66} \tag{6.2.65}$$

The one-dimensionality of $\hat{\bar{R}}_{ka}$ means that its desired minimal polynomial with respect to $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}$ must be first-order also:

$$\Upsilon_{ka}(\lambda) = \lambda + p_{ka1} \tag{6.2.66}$$

Because $\Upsilon_{ka}(\lambda)$, $\Psi_i(\lambda)$, and $\Psi_o(\lambda)$ are all first-order, the output stationarity constraint on $d_{\Psi Ki}$ and $d_{\Psi Ko}$ is:

$$\begin{aligned} \Psi_i(\lambda) &= \Upsilon_{ka}(\lambda) \\ \Psi_o(\lambda) &= \Upsilon_{ka}(\lambda) \end{aligned} \tag{6.2.67}$$

and the characteristic equation of $\hat{A} - (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}$ is required to have a double eigenvalue:

$$|\lambda \hat{E}_6 - \hat{A} + (\hat{D}_{FR} + \hat{D}_{\Psi K})\hat{C}| = \Psi_a(\lambda)\,\Psi_r(\lambda)\,\Upsilon_{ka}(\lambda)^2 \tag{6.2.68}$$

The optimization procedure suggested in Section 5.1.3 for assigning the double eigenvalue is to replace p_{i1} and p_{o1} in $\hat{c}_i \hat{P}_{\infty i} \hat{c}_i^T$ and $\hat{c}_o \hat{P}_{\infty o} \hat{c}_o^T$ by p_{ka1} and then optimize p_{ka1} according to a variation on (6.2.64):

$$\frac{\partial(\rho_i^2 \hat{c}_i \hat{P}_{\infty i} \hat{c}_i^T + \rho_o^2 \hat{c}_o \hat{P}_{\infty o} \hat{c}_o^T)}{\partial p_{ka1}} = 0 \tag{6.2.69}$$

The weighting factors ρ_i^2 and ρ_o^2 can be chosen by the designer. The double eigenvalue in (6.2.41) is then defined by assigning:

$$\begin{aligned} p_{i1} &= p_{ka1} \\ p_{o1} &= p_{ka1} \end{aligned} \tag{6.2.70}$$

Without the output stationarity requirement, p_{i1} and p_{o1} could be optimized according to the two independent equations:

$$\begin{aligned} \frac{\partial(\hat{c}_i \hat{P}_{\infty i} \hat{c}_i^T)}{\partial p_{i1}} &= 0 \\ \frac{\partial(\hat{c}_o \hat{P}_{\infty o} \hat{c}_o^T)}{\partial p_{o1}} &= 0 \end{aligned} \tag{6.2.71}$$

Unless p_{i1} and p_{o1} in (6.2.70) coincidently satisfy (6.2.71) the addition of the output stationarity requirement decreases the output estimate accuracies of the failure detection system.

6.3 Detection Law

The use of multiple failure detection systems presents difficulties in interpreting the output of the individual failure detection systems. It is necessary to avoid a situation where one failure detection system indicates a particular failure has occurred, a second indicates a different failure has occurred, and there is no means of determining which failure detection system to believe. Section 5.3 presents a heuristically derived detection law for choosing between conflicting decisions made by the individual failure detection systems. That detection law is developed in this section for the autopilot-aircraft example.

A detection law must be implemented as a two-tiered structure. The first tier specifies the manner in which the individual failure detection systems interpret the output error signals which they monitor. The premise is that each failure detection system is to interpret its own output signals (residuals) in the context of the failures which it is designed to detect. The structure of this decision-making process consists of three segments:

i) an analysis of the residuals to determine the presence of a failure-generated output error,

ii) a waiting period to determine the direction or hyperplane in the output space to which that error signal is constrained, and

iii) an interpretation of the observed error signal in terms of the failures which the failure detection system was designed to detect.

The second tier in the detection law provides for the selection of one of the failure detection system decisions as a correct identification of the failure. If one failure detection system constrains the output error to a single direction and the other constrains it to a plane, the decision of the first failure detection system is to be accepted. A considerable amount of freedom is available in implementing the two tiers of the detection law. The algorithm used for the autopilot-aircraft is that developed in the second half of Section 5.3.

In the absence of any failure, the estimation errors for the sampled-data state estimator of (5.2.5) - (5.2.6) satisfy:

$$\bar{\epsilon}(k\,\Delta t) = \Phi(k\,\Delta t, (k-1)\,\Delta t)\,\epsilon((k-1)\Delta t) + G_{1\Delta}\eta_{1\Delta}(k\Delta t)$$

$$\bar{\epsilon}'(k\Delta t) = C\,\bar{\epsilon}(k\Delta t) + G_{2\Delta}\eta_{2\Delta}(k\Delta t) \qquad (5.2.9)$$

The residual $\bar{\epsilon}'(k\Delta t)$ is the difference between the measurement at $k\Delta t$ and the predicted measurement; and $\epsilon(k\Delta t)$ is the error in the state estimate at $k\Delta t$ after being updated with the new measurement:

$$\epsilon(k\Delta t) = \bar{\epsilon}(k\Delta t) - D_{\Delta K}\bar{\epsilon}'(k\Delta t) \qquad (5.2.10)$$

The covariance matrices are:

$$\bar{P}_\infty = E[\bar{\epsilon}(k\Delta t)\bar{\epsilon}(k\Delta t)^T] = \Phi(\Delta t,\ 0)\,P_\infty\,\Phi(\Delta t,\ 0)^T + \frac{1}{\Delta t}G_{1\Delta}G_{1\Delta}^T$$

$$P_\infty = E[\,\epsilon(k\Delta t)\,\epsilon(k\Delta t)^T] \qquad (5.2.11)$$

$$= (E_n - D_{\Delta K}C)\,\bar{P}_\infty\,(E_n - D_{\Delta K}C)^T + \frac{1}{\Delta t}D_{\Delta K}G_{2\Delta}G_{2\Delta}^T D_{\Delta K}^T$$

The variance of the residuals is $C\bar{P}_\infty C^T + \frac{1}{\Delta t}G_{2\Delta}G_{2\Delta}^T$.

The first step in the implementation of the detection law is to resolve the residuals for each failure detection system into the base normal coordinates associated with that failure detection system:

$$\hat{\bar{\epsilon}}'(k\Delta t) = T_m^{-1}\,\bar{\epsilon}'(k\Delta t) \qquad (6.3.1)$$

Each component of the transformed residual is then treated as the output of an independent detection channel. The test for the presence of a failure-generated component in the residual is a simple threshold crossing test on the magnitude of the residual. A failure is assumed to have

occurred if the magnitude of the residual in any detection channel exceeds the 3σ projection of $C\bar{P}_\infty C^T + \frac{1}{\Delta t} G_{2\Delta} G_{2\Delta}^T$ onto that channel. As an example the four detection channels for the gyro-aileron failure detection system are defined by the residuals $\hat{\bar{\epsilon}}'_a(k\Delta t)$, $\hat{\bar{\epsilon}}'_r(k\Delta t)$, $\hat{\bar{\epsilon}}'_i(k\Delta t)$, and $\hat{\bar{\epsilon}}'_o(k\Delta t)$:

$$T_{mg} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\hat{\bar{\epsilon}}'_g(k\Delta t) = T_{mg}^{-1} \bar{\epsilon}'(k\Delta t) = \begin{bmatrix} \hat{\bar{\epsilon}}'_a(k\Delta t) \\ \hat{\bar{\epsilon}}'_r(k\Delta t) \\ \hat{\bar{\epsilon}}'_i(k\Delta t) \\ \hat{\bar{\epsilon}}'_o(k\Delta t) \end{bmatrix} \tag{6.3.2}$$

where T_{mg} is the output transformation T_m defined in (6.2.28). A failure is assumed to have occurred if any of the four inequalities is satisfied:

$$\hat{\bar{\epsilon}}'_a(k\Delta t) > 3\left[\hat{e}_{41}^T T_m^{-1} \left(C\bar{P}_\infty C^T + \frac{1}{\Delta t} G_{2\Delta} G_{2\Delta}^T\right) T_m \hat{e}_{41} \right]^{1/2}$$

$$\hat{\bar{\epsilon}}'_r(k\Delta t) > 3\left[\hat{e}_{42}^T T_m^{-1} \left(C\bar{P}_\infty C^T + \frac{1}{\Delta t} G_{2\Delta} G_{2\Delta}^T\right) T_m \hat{e}_{42} \right]^{1/2}$$

$$\hat{\bar{\epsilon}}'_i(k\Delta t) > 3\left[\hat{e}_{43}^T T_m^{-1} \left(C\bar{P}_\infty C^T + \frac{1}{\Delta t} G_{2\Delta} G_{2\Delta}^T\right) T_m \hat{e}_{43} \right]^{1/2}$$

$$\hat{\bar{\epsilon}}'_o(k\Delta t) > 3\left[\hat{e}_{44}^T T_m^{-1} \left(C\bar{P}_\infty C^T + \frac{1}{\Delta t} G_{2\Delta} G_{2\Delta}^T\right) T_m \hat{e}_{44} \right]^{1/2}$$

(6.3.3)

Once a failure detection system has detected the existence of a failure, ii) and iii) require that the failure be isolated and identified. As implemented for the example, isolation requires a waiting period to determine which detection channels the failure-generated error signal appears in. The fact that the gyro-aileron failure detection system cannot contain the output error generated by a dynamics change to a single detection channel does not imply that the threshold crossings due to a dynamics change must appear in two detection channels simultaneously. If a threshold crossing appears first in the rate gyro channel, the failure detection system cannot identify the failure as a rate gyro failure until it is confident that a threshold crossing will not subsequently appear in the attitude gyro channel. With one exception discussed later, the number of channels in which the threshold is exceeded defines the dimension of the output subspace to which the failure detection system constrains the error signal. Because the second tier of the detection law requires this dimension in order to accept the correct failure identification, it is important that each failure detection system correctly isolate its failure-generated residuals. The minimum length of the waiting period is a function of the system, the failure mode, and the eigenvalues of A-DC.

Following the failure isolation by a failure detection system, the failure identification is made in terms of the failures which the failure detection system was designed to detect. If only a rate gyro channel threshold crossing occurs, the gyro-aileron failure detection system reports a rate gyro failure. If an attitude gyro threshold crossing occurs also, the failure detection system reports that both gyros have failed -- even though the true failure may be a dynamics change. If a dynamics change has in fact occurred, the other failure detection system will report it as such. The second tier of the detection law is then responsible for accepting the correct failure identification. When a failure detection system is designed to monitor sensor failures and a threshold crossing appears in two detection channels, the identified failure may be a single sensor failure rather than two separate failures. This possibility underscores the necessity of correctly isolating the residuals generated by a failure before making a failure identification.

If output stationarity has been used to increase the number of failures which a failure detection system is to monitor, the failure identification problem becomes more complex. In effect, it is necessary to adopt two or more parallel failure identification algorithms for that failure detection system and then require the second tier of the failure detection law to accept the correct identification. This procedure can be illustrated in the context of the failure detection system for the aileron and aileron sensor failures.

Ignoring the amplifier failure for a moment, the event vectors associated with the failure detection system are f_i, f_o, f_β, f_{ar}. The transformation to define the detection channels for $\bar{\epsilon}'_{ar}(k\Delta t)$, $\bar{\epsilon}'_{\beta}(k\Delta t)$, $\bar{\epsilon}'_{i}(k\Delta t)$, and $\bar{\epsilon}'_{o}(k\Delta t)$ is the inverse of T_{ms1}:

$$T_{ms1} = [\, Cf_{ar}\ Cf_{\beta}\ Cf_{i}\ Cf_{o}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\beta_{\varphi} & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\hat{\bar{\epsilon}}'_{s1}(k\Delta t) = T_{ms1}^{-1}\, \bar{\epsilon}'(k\Delta t) = \begin{bmatrix} \hat{\bar{\epsilon}}'_{ar}(k\Delta t) \\ \hat{\bar{\epsilon}}'_{\beta}(k\Delta t) \\ \hat{\bar{\epsilon}}'_{i}(k\Delta t) \\ \hat{\bar{\epsilon}}'_{o}(k\Delta t) \end{bmatrix} \qquad (6.3.4)$$

The detection channel defined by $\bar{\epsilon}'_i(k\Delta t)$ monitors an inboard aileron failure, etc. If either aileron fails, a threshold crossing appears in the appropriate detection channel. The event vector f_{ka} for an amplifier failure is a linear combination of f_i and f_o:

$$f_{ka} = +k_i f_i + k_o f_o \qquad (6.1.23)$$

Because $D_{\Delta K}$ for this failure detection system is chosen such that f_{ka} is output stationary with f_i, f_o, f_β, f_{ar}; (6.1.23) implies that an amplifier

failure creates threshold crossings in both aileron detection channels. Output stationarity guarantees that the output error generated by an amplifier is unidirectional along Cf_{ka}. The two threshold crossings occur because $T_{ms1}^{-1} Cf_{ka}$ is not in any single detection channel defined by (6.3.4). This difficulty can be overcome by defining a new output transformation with one of Cf_i, Cf_o replaced by Cf_{ka}:

$$T_{ms2} = [\, Cf_{ar} \;\; Cf_{\beta} \;\; Cf_i \;\; Cf_{ka}] = \begin{bmatrix} 1 & 0 & 0 & 0 \\ -\beta_{\varphi} & 1 & 0 & 0 \\ 0 & 0 & 1 & k_i \\ 0 & 0 & 0 & k_o \end{bmatrix}$$

$$\hat{\bar{\epsilon}}'_{s2}(k\Delta t) = T_{ms2}^{-1} \hat{\bar{\epsilon}}'(k\Delta t) = \begin{bmatrix} \hat{\bar{\epsilon}}'_{ar}(k\Delta t) \\ \hat{\bar{\epsilon}}'_{\beta}(k\Delta t) \\ \hat{\bar{\epsilon}}'_{i}(k\Delta t) \\ \hat{\bar{\epsilon}}'_{ka}(k\Delta t) \end{bmatrix} \qquad (6.3.5)$$

In the transformed output defined by (6.3.5), an outboard aileron failure is not associated with a single detection channel, but an amplifier failure is. As far as the second tier of the detection law is concerned, it is as if a third failure detection system had been defined. An outboard aileron failure causes a single threshold crossing only for (6.3.4); an amplifier failure causes a single threshold crossing only for (6.3.5); and an inboard aileron failure causes a single threshold crossing in both output spaces. With both (6.3.4) and (6.3.5) providing failure identifications, the detection law can then accept the correct identification.

The addition of a third coordinate transformation results in a set of twelve detection channels to be monitored. Figure 6.4 defines three detection systems based upon these detection channels -- the T_{mg}, T_{ms1}, and T_{ms2} detection systems. With the aid of deterministic simulations with nearly continuous measurements, the threshold crossings generated in each detection system for every potential failure can be tabulated.

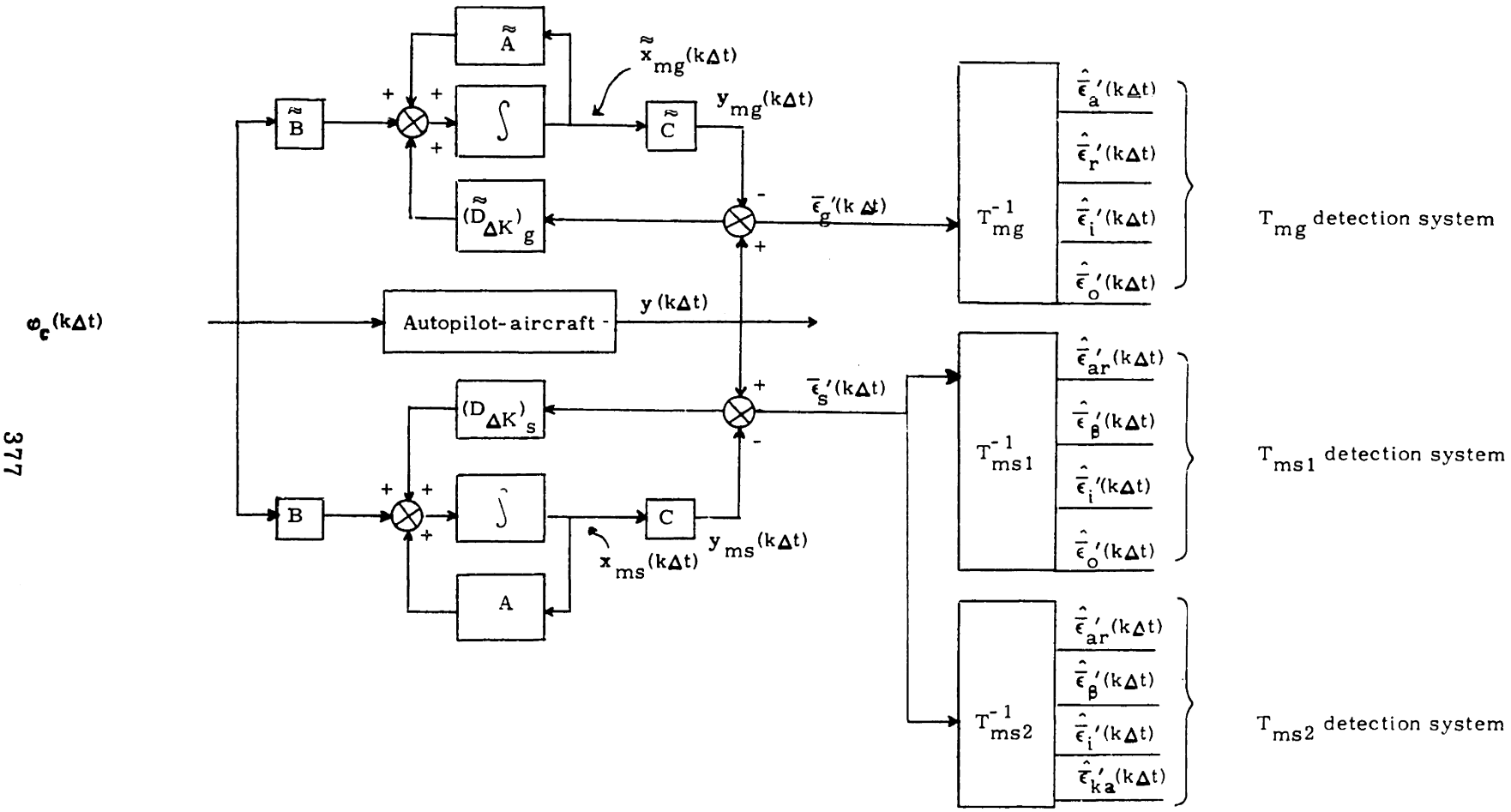


Figure 6.4: Schematic diagram of the relationship of the detection systems T_{mg}, T_{ms1}, and T_{ms2} nased upon the failure detection systems designed in Sections 6.1 and 6.2. The figure does not indicate the sampled-data nature of the measurements or the state estimates.

These results are summarized in Table 6.2 for the three detection systems. The failure identification which each detection system makes is also indicated in the table. It can be verified from the table that the correct failure identification is always made by the detection system which experiences the least number of threshold crossings. This further justifies the choice of the second tier of the detection law.

6.4 Simulation Results

The purpose of this section is to emphasize the potential value of failure detection theory for providing real-time failure detection and identification. The analysis is restricted to the autopilot-aircraft system, but the intention is to provide insight into the class of applications for which failure detection theory can be expected to yield useful information. Particular attention is given to the possibility of a failure detection system performing the dual functions of failure detection and state estimation.

Section 6.4.1 presents a series of simulation results which indicate that the failure detection systems and detection law designed in Sections 6.1 - 6.3 can provide nearly instantaneous failure detection and identification for a variety of failures and failure modes. The emphasis in that subsection is on failure modes which can significantly impair the short-term stability, controllability, or observability of the autopilot-aircraft system. It is always possible to find component degradations which do not significantly alter the dynamic response of a vehicle, but which have an important cumulative effect over a long period of time -- a bias in one of the gyros in the autopilot is a typical example. Section 6.4.3 demonstrates that failure detection systems can also be used to detect such degradations, but that the detection filter-state estimators designed in Section 6.2.2 may not provide an adequate failure detection capability. Several techniques for modifying the detection filters to improve the failure detection properties are illustrated. Some of these techniques also improve the accuracy of the state estimates; some do not.

The simulations in Section 6.4.1 are based upon a single measurement sampling period (.1 seconds) and a reasonable choice of $G_{1\Delta}$ and $G_{2\Delta}$ in the stochastic system model. The sensitivity analysis in Section

	T_{mg} Detection System		T_{ms1} Detection System		T_{ms2} Detection System	
Failure	Channels	Identification	Channels	Identification	Channels	Identification
Inboard aileron	$\hat{\bar{\epsilon}}_i'(k\Delta t)$	Inboard aileron*	$\hat{\bar{\epsilon}}_i'(k\Delta t)$	Inboard aileron*	$\hat{\bar{\epsilon}}_i'(k\Delta t)$	Inboard aileron*
Outboard aileron	$\hat{\bar{\epsilon}}_o'(k\Delta t)$	Outboard aileron*	$\hat{\bar{\epsilon}}_o'(k\Delta t)$	Outboard aileron*	$\hat{\bar{\epsilon}}_i'(k\Delta t)$; $\hat{\bar{\epsilon}}_{ka}'(k\Delta t)$	Inboard aileron; amplifier*
Aircraft dynamics	$\hat{\bar{\epsilon}}_r'(k\Delta t)$; $(k\Delta t)$	Rate gyro; Attitude gyro	$\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Aircraft dynamics*	$\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Aircraft dynamics
Inboard aileron deflection sensor	$\hat{\bar{\epsilon}}_r'(k\Delta t)$; $\hat{\bar{\epsilon}}_a'(k\Delta t)$; $\hat{\bar{\epsilon}}_i'(k\Delta t)$	Rate gyro; Attitude gyro; Inboard aileron	$\hat{\bar{\epsilon}}_l'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Inboard aileron deflection sensor*	$\hat{\bar{\epsilon}}_i'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Inboard aileron deflection sensor
Outboard aileron deflection sensor	$\hat{\bar{\epsilon}}_r'(k\Delta t)$; $\hat{\bar{\epsilon}}_a'(k\Delta t)$ $\hat{\bar{\epsilon}}_o'(k\Delta t)$	Rate gyro; Attitude gyro; Outboard aileron	$\hat{\bar{\epsilon}}_o'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Outboard aileron deflection sensor*	$\hat{\bar{\epsilon}}_i'(k\Delta t)$; $\hat{\bar{\epsilon}}_{ka}'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Inboard aileron; amplifier; dynamics change
Attitude gyro	$\hat{\bar{\epsilon}}_a'(k\Delta t)$	Attitude gyro*	$\hat{\bar{\epsilon}}_{ar}'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Attitude-rate gyros; Aircraft dynamics	$\hat{\bar{\epsilon}}_{ar}'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Attitude-rate gyros; Aircraft dynamics
Rate gyro	$\hat{\bar{\epsilon}}_r'(k\Delta t)$	Rate gyro*	$\hat{\bar{\epsilon}}_{ar}'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Attitude-rate gyros; Aircraft dynamics	$\hat{\bar{\epsilon}}_{ar}'(k\Delta t)$; $\hat{\bar{\epsilon}}_\beta'(k\Delta t)$	Attitude-rate gyros; Aircraft dynamics
Amplifier	$\hat{\bar{\epsilon}}_l'(k\Delta t)$; $\hat{\bar{\epsilon}}(k\Delta t)$	Inboard aileron; Outboard aileron	$\hat{\bar{\epsilon}}_l'(k\Delta t)$; $\hat{\bar{\epsilon}}_o'(k\Delta t)$	Inboard aileron; Outboard aileron	$\hat{\bar{\epsilon}}_{ka}'(k\Delta t)$	Amplifier*

Table 6.2 Detection channels for the T_{mg}, and T_{ms2} detection systems in which each of the possible system failures generates error signals; and the system failures identified by each detection system. Asterisks denote the identification accepted by the second tier of the detection law.

6.4.3 demonstrates the dependence of good failure detection on a short measurement sampling period. Section 6.4.2 also indicates that sub-optimal detection filter-state estimators can provide state estimates comparable to those of a Kalman filter for short measurement sampling periods, but that the accuracy of the state estimates may become unacceptable for long measurement sampling periods. The effectiveness of the various optimization procedures of Section 5.2 in finding good detection filter-state estimators is shown to be a function of $G_{1\Delta}$ and $G_{2\Delta}$ as well as the measurement sampling period.

6.4.1 Failure Detection

The simulations presented in this subsection are based upon the system parameters given in Table 6.3. The roll time constant $(1/\beta_\varphi)$ and total aileron effectiveness $(\alpha_i + \alpha_o)$ are those of the 040A Space Shuttle Vehicle at Mach 2 and 70,000 feet. The airframe has a roll response time $(3/\beta_\varphi)$ of 8.6 seconds. The amplifier gains (k_a and k_r) are chosen to provide a closed-loop response time of 1.8 seconds with no overshoot. k_a is sufficiently large that the autopilot-aircraft system is unstable without the damping provided by the rate gyro feedback. The basic autopilot sampling period is .01 seconds. As indicated in Section 6.2.1, the purpose of the short autopilot sampling period is to evaluate the ability of a failure detection system to use discrete measurements to detect and identify failures in nearly continuous systems.

The measurement sampling period used in this subsection is .1 seconds. The measurement variances indicated in Table 6.3 are typical of the measurement uncertainties in currently available instruments and the strength of the random aerodynamic moment is representative of the disturbances expected at Mach 2 for the 040A vehicle. This information permits the optimization algorithms outlined in Section 6.2.2 to be solved for the optimal detection filter-state estimator $D_{\Delta K}$ for each of the two failure detection systems. The detection filters used for the simulation were optimized using the parameter optimization procedure based upon the continuous measurement stochastic model of Section 6.2.1. As indicated in Section 6.2.2, the parameter optimization technique yields the

System parameters

$$\beta_{\varphi} = .34912 \text{ sec}^{-1}$$
$$k_a = 1. \text{ sec}^{-1}$$
$$k_r = .85$$
$$k_i = 5.$$
$$k_o = 5.$$
$$\alpha_i = 1.081 \text{ sec}^{-2}$$
$$\alpha_o = 1.622 \text{ sec}^{-2}$$

System disturbances

$$W_{\beta} = 9 \times 10^{-8} \text{ (rad/sec}^2\text{)}^2$$

$$q_a^2 = \sigma_{1a\Delta}^2 = 4 \times 10^{-8} \text{ rad}^2$$

$$q_r^2 = \sigma_{1r\Delta}^2 = 4 \times 10^{-8} \text{ (rad/sec)}^2$$

Measurement noise

$$\sigma_{2o\Delta}^2 = \sigma_{2i\Delta}^2 = 6.4 \times 10^{-7} \text{ rad}^2$$

Table 6.3: Autopilot-aircraft parameters used in simulations

same detection filter as the continuous Kalman algorithm, but not the same detection filter as the discrete Kalman algorithm.

The detection law developed in Section 6.3 defines three detection systems, each of which must make a failure detection decision before the second tier of the detection law can correctly identify a system failure. The T_{mg} detection system is based upon the residuals of the gyro-aileron failure detection system; both the T_{ms1} and T_{ms2} detection systems are based upon the residuals of the aileron-aileron sensor failure detection system. For the failures analyzed in this subsection, the T_{ms1} and T_{ms2} detection systems present redundant information. Because of this, only the outputs of the T_{mg} and T_{ms1} detection systems are presented and discussed. Each detection system is required to wait .3 seconds after a threshold crossing in one of its detection channels before making a failure identification. The .3 second delay was determined through trial and error to be the shortest waiting period within which the detection systems could correctly isolate the residuals generated by the failure modes analyzed in this subsection. Sections 6.4.2 and 6.4.3 discuss other possible criteria involved in selecting a waiting period.

The nominal performance of the autopilot-aircraft system is illustrated in Figure 6.5. This simulation is presented for future reference and to demonstrate the state estimation capabilities of a failure detection system in the absence of a failure. Figure 6.5a presents the measured aileron deflections (inboard and outboard), the rate gyro output, and the attitude gyro output for a roll command of 2° (.034 radians) at 1.6 seconds. Figures 6.5b and 6.5c present the T_{mg} and T_{ms1} detection system outputs respectively. Because T_{mg} as defined in (6.3.2) is the identity matrix, the outboard aileron detection channel residual $\hat{\epsilon}'_o(k\Delta t)$ is the difference between the measured aileron deflection at $k\Delta t$ and the estimated aileron deflection at $k\Delta t$, etc. The 3σ magnitudes for the residual components are indicated in Figures 6.5b and 6.5c by upper and lower bounds. Since no system failure occurs during the run, the residuals in all four detection channels remain within the 3σ thresholds.

Figure 6.5a Sensor outputs for a roll command of 2° (.034 radians) at 1.6 seconds and no failures.

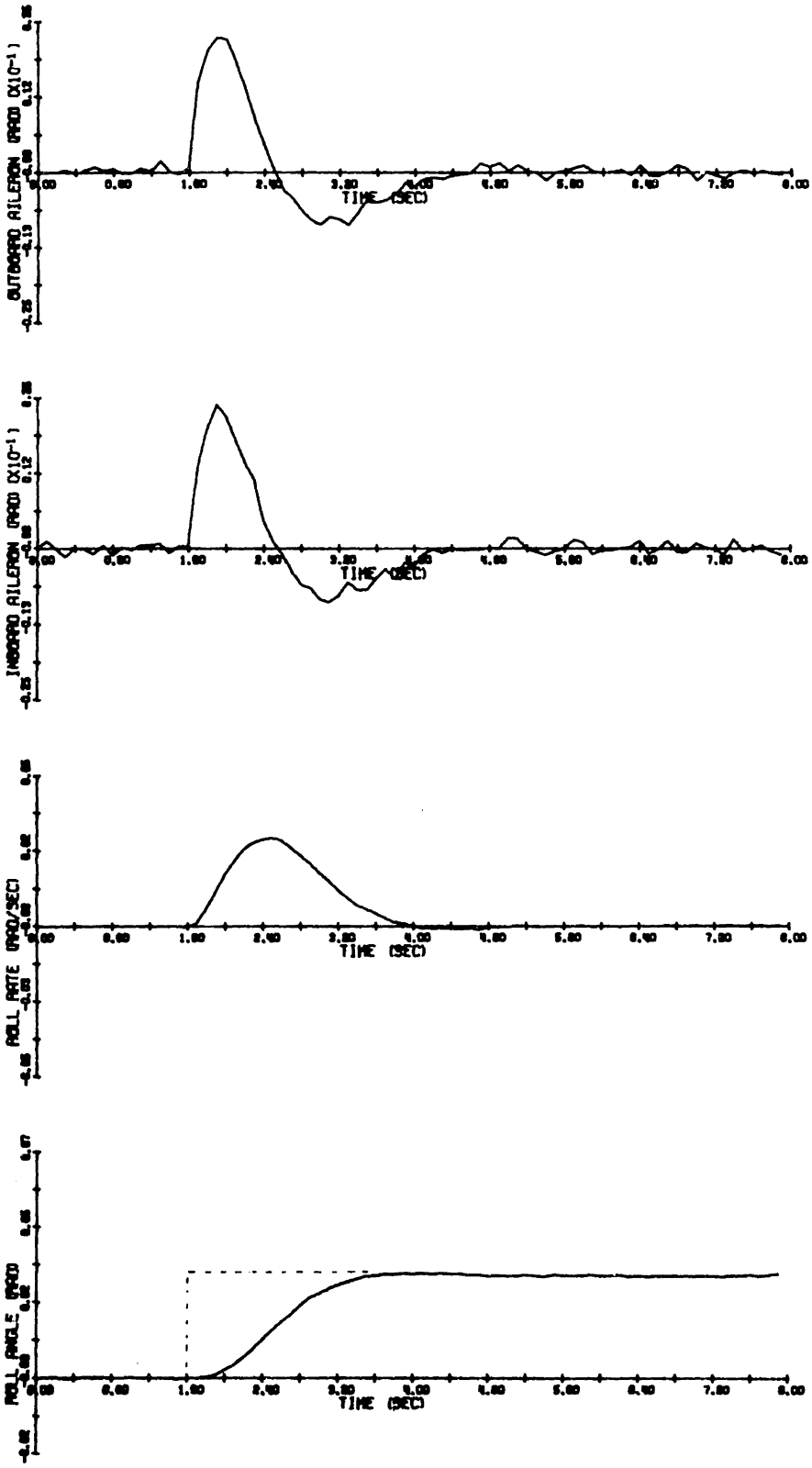


Figure 6.5b T_{mg} detection system output for simulation in Figure 6.5a.

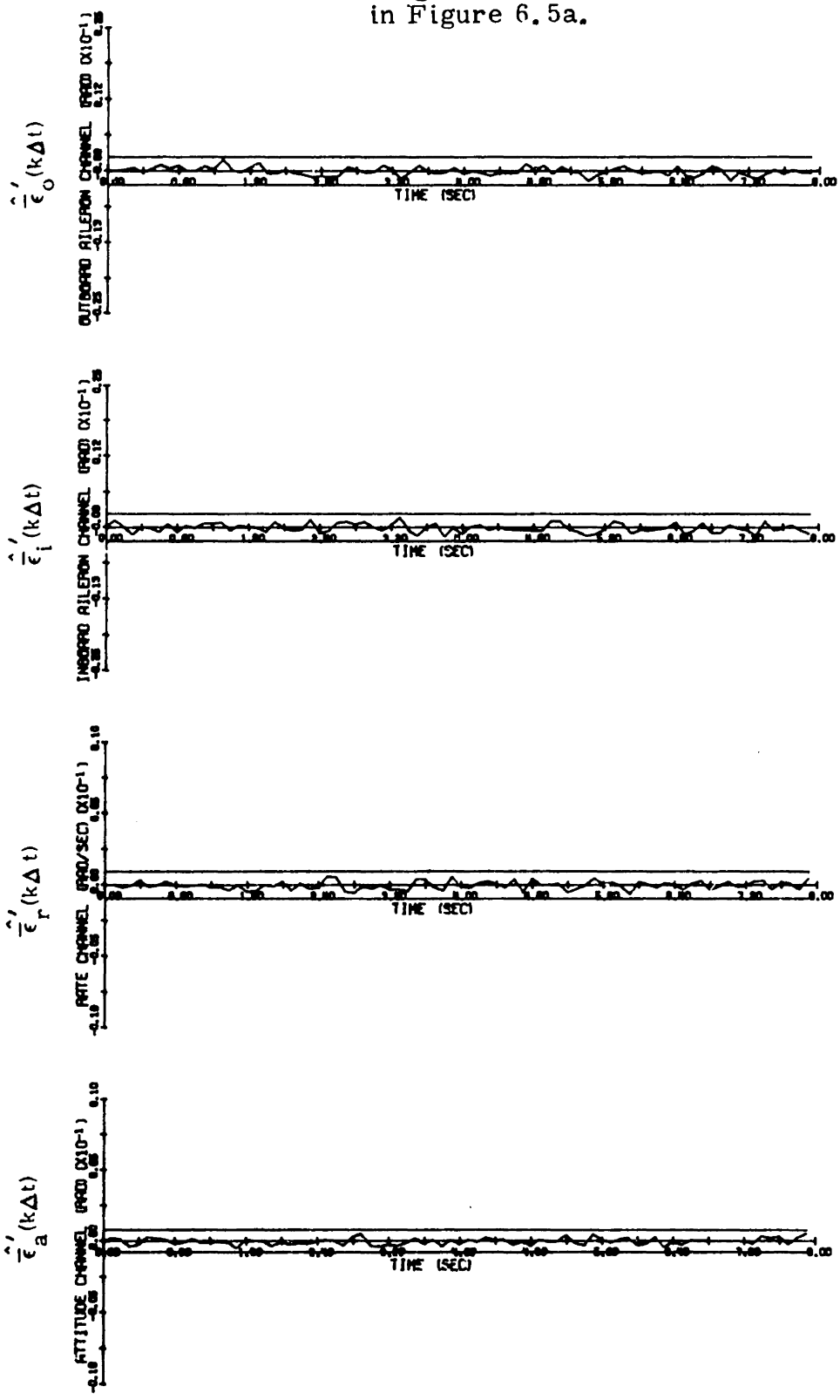
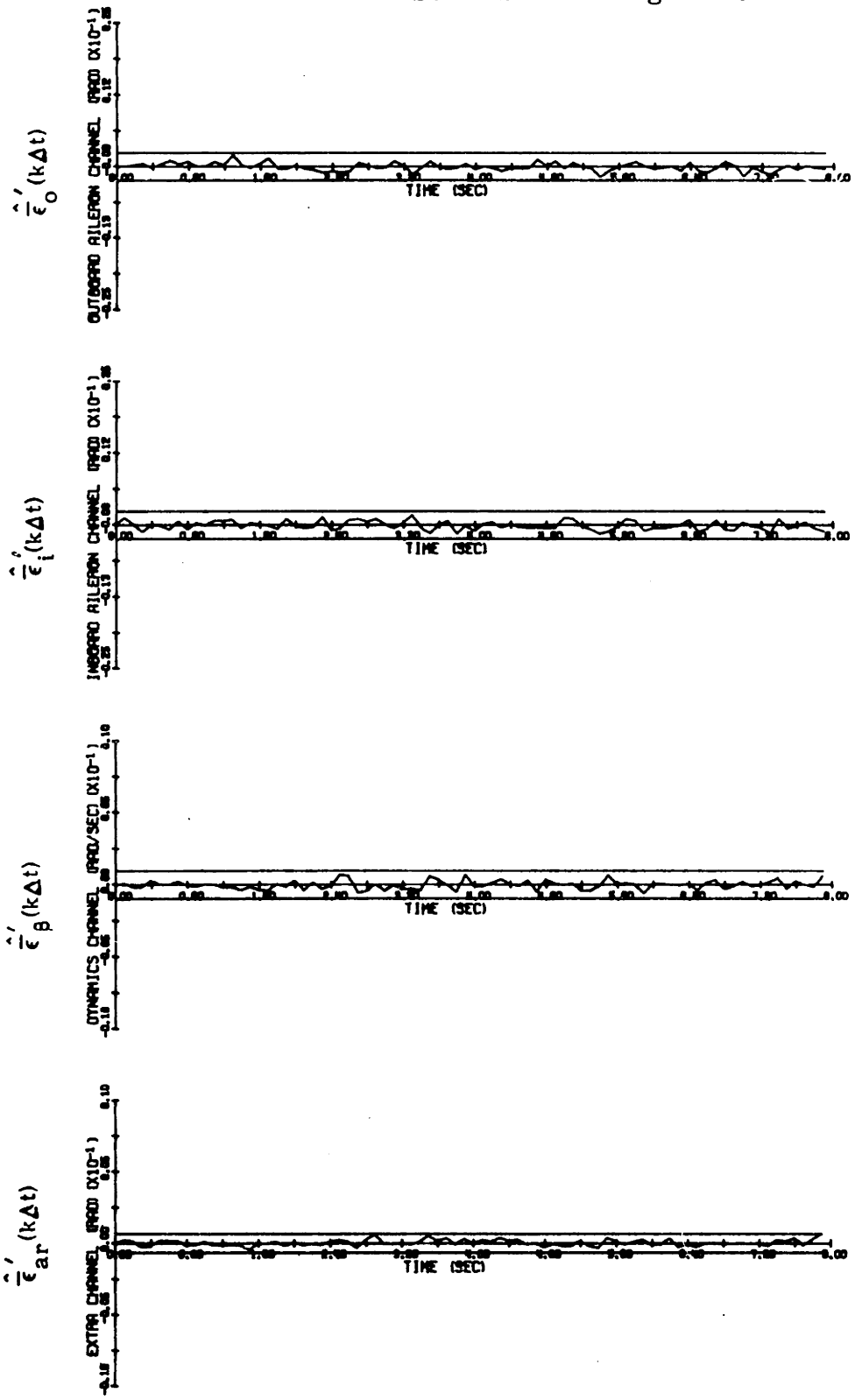


Figure 6.5c T_{ms1} detection system output for simulation in Figure 6.5a



One of the more difficult failure detection problems is to discriminate between the failure of a system component and the failure of the sensor which measures the output variable associated with that component. Unless the sensor is at least triply redundant, this discrimination cannot be accomplished solely on the basis of the sensor output. The only possibility for correctly identifying the failure is to capitalize upon the fact that the component failure and the sensor failure affect the system performance in different manners. This difficulty and the solution which failure detection theory provides are illustrated by the following two simulations.

Figure 6.6 presents the system measurement history and detection channel outputs for a simulation which differs from Figure 6.5 in that an inboard aileron deflection sensor failure occurs at the time at which the 2° command is applied (1.6 seconds). The failure mode is that modeled by ΔC_{s_i} in (6.1.14) -- the sensor output consists only of the white measurement noise. Because the sensor failure model does not involve a change in the autopilot-aircraft system dynamics, the nominal appearance of the gyro output histories and outboard aileron deflection history suggests that a sensor failure has occurred.

The intuitive argument of the previous paragraph is verified by the detection system outputs. The T_{ms1} detection system is designed to detect an inboard aileron deflection sensor failure. Table 6.2 indicates that the failure should produce threshold crossings in the inboard aileron channel ($\hat{\bar{\epsilon}}_i'(k\Delta t)$) and the roll dynamics channel ($\hat{\bar{\epsilon}}_\beta'(k\Delta t)$) for the T_{ms1} detection system. Figure 6.6c verifies that these crossings occur at 1.7 seconds and 1.8 seconds respectively -- well within the .3 second waiting period -- and the T_{ms1} detection system correctly identifies the sensor failure at 2.0 seconds. The T_{mg} detection system is not designed to detect sensor failures, but its detection decision must be made before the second tier of the detection law can make a final failure identification. Three threshold crossings appear in Figure 6.6b; the first at 1.7 seconds in the inboard aileron channel ($\hat{\bar{\epsilon}}_i'(k\Delta t)$) and the third at 2.0 seconds in the attitude channel ($\hat{\bar{\epsilon}}_a'(k\Delta t)$). Table 6.2 indicates that at the end of the .3 second waiting period (2.0 seconds) the T_{mg} detection system reports that the inboard ailerons, the rate

Figure 6.6a Sensor outputs for a roll command of 2° at 1.6 seconds and an inboard aileron sensor failure at 1.6 seconds.

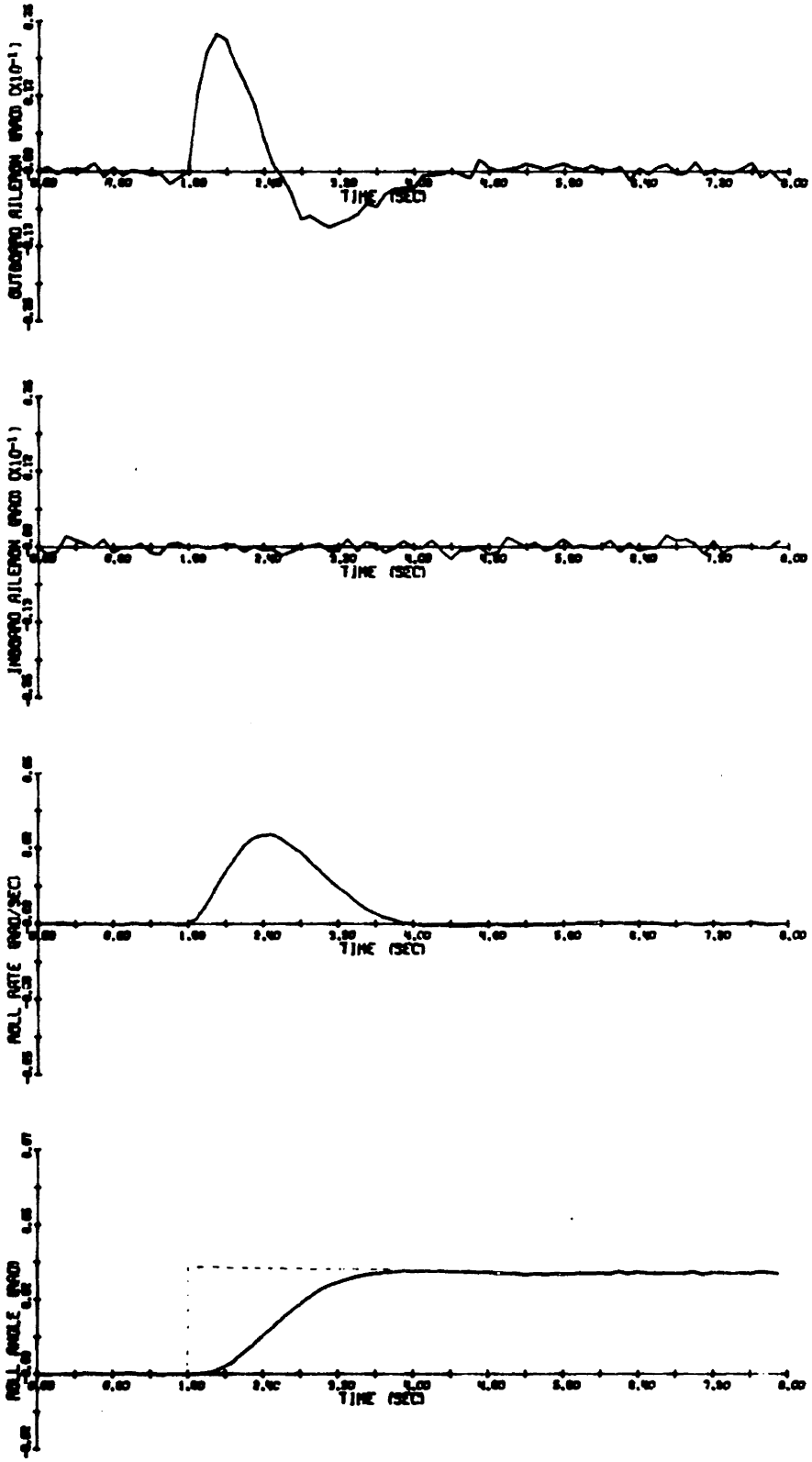


Figure 6.6b T_{mg} detection system output for simulation in Figure 6.6a. Arrows indicate time of threshold crossings.

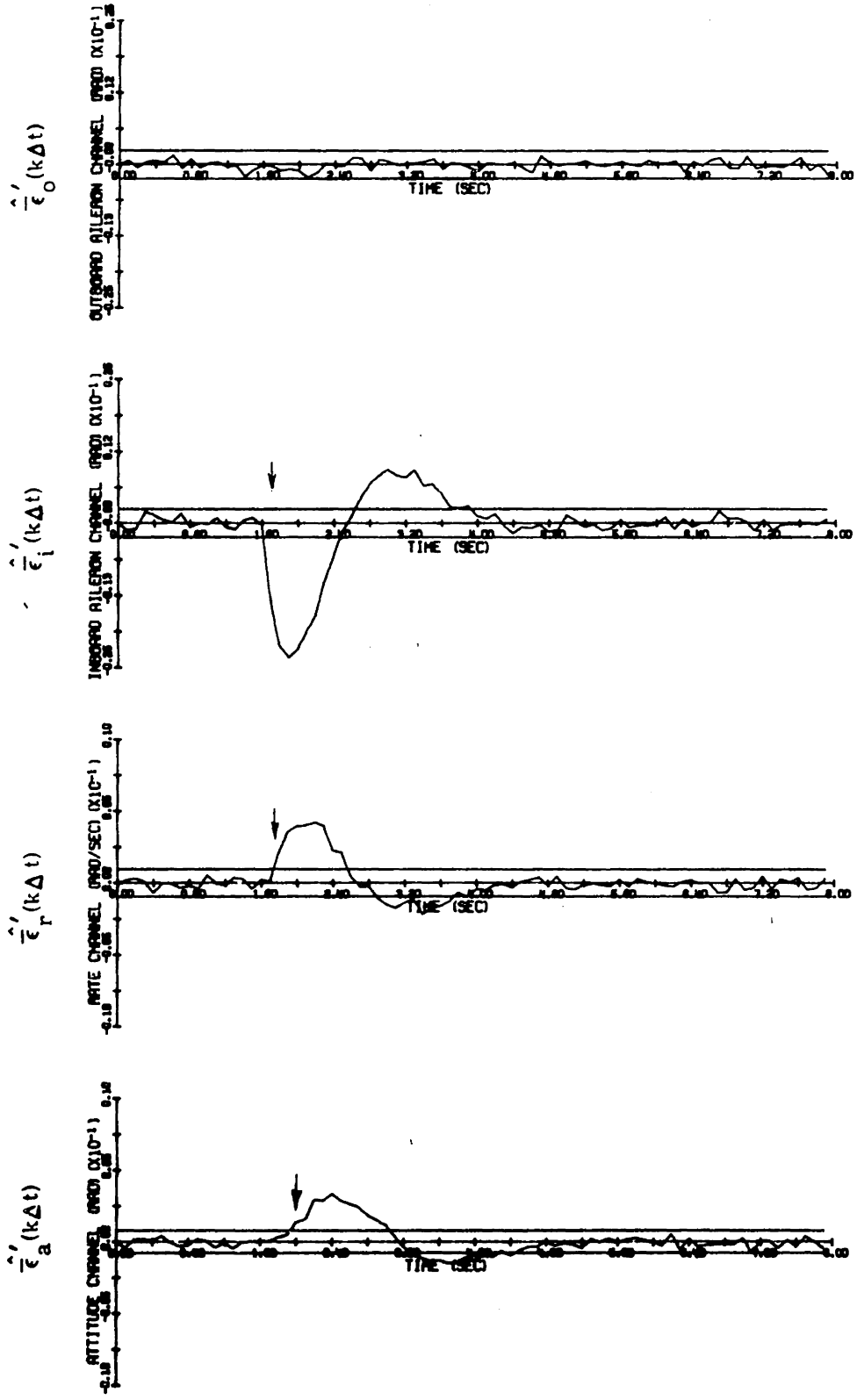
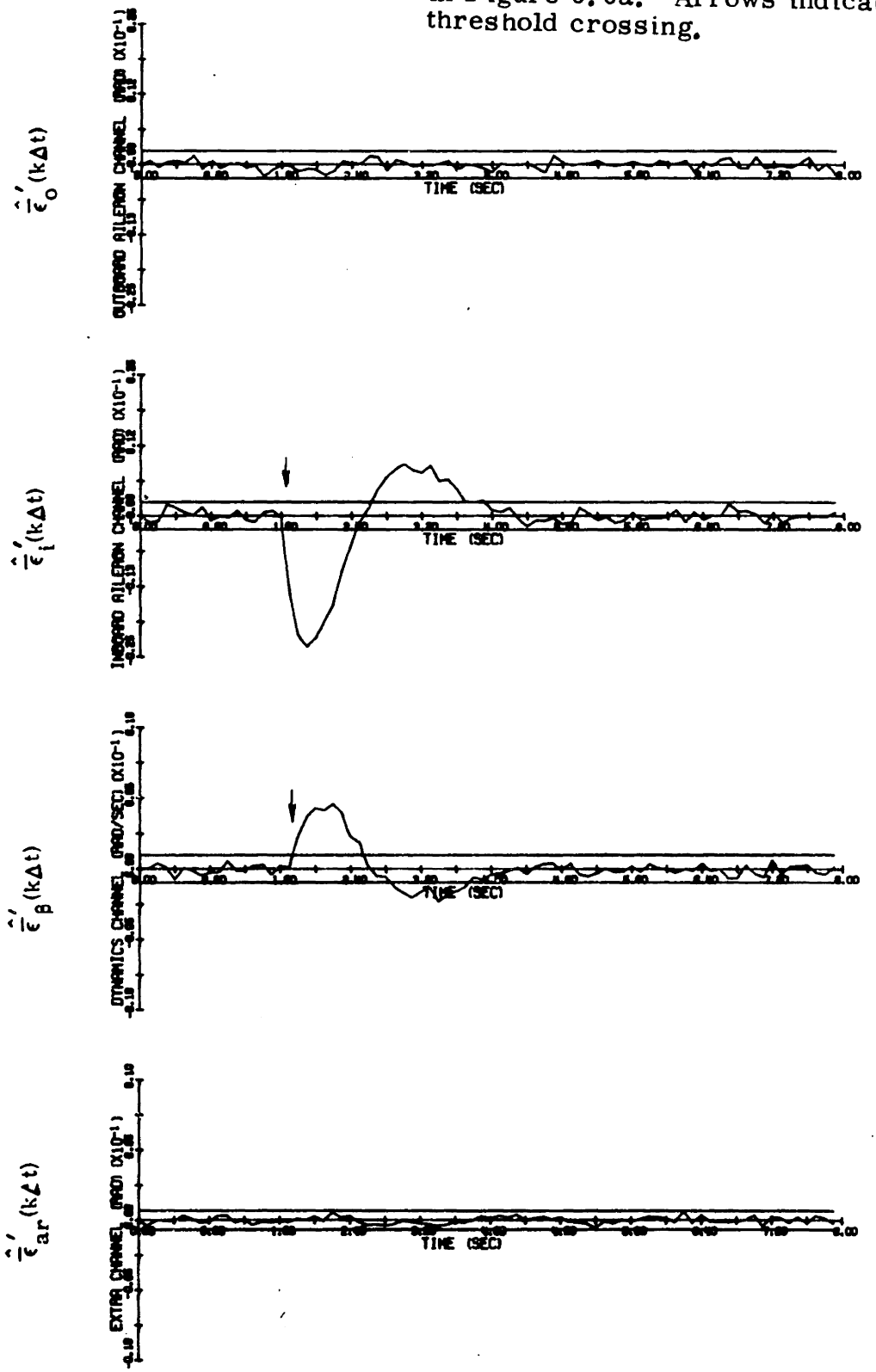


Figure 6.6c T_{ms1} detection system output for simulation in Figure 6.6a. Arrows indicate time of threshold crossing.



gyro, and the attitude gyro have all failed. By virtue of the fact that only two threshold crossings occur in the T_{ms1} detection system, the second tier of the detection law correctly accepts the T_{ms1} detection system's identification of the inboard aileron deflection sensor failure.

The difference between an aileron failure and an aileron sensor failure is illustrated by the simulation presented in Figure 6.7. The run differs from Figure 6.6 in that an inboard aileron servo failure occurs rather than the sensor failure. The failure mode is that modeled by ΔA_i and ΔB_i in (6.1.18). The inboard ailerons remain fixed in the trim position and do not respond to the deflection commands generated by the autopilot. Figure 6.7a indicates that the inboard aileron deflection sensor output consists only of the white measurement process as it did in Figure 6.6a -- the difference is that in this instance the sensor is providing accurate deflection measurements. The only means of determining that the ailerons rather than the sensor have failed is by analyzing the output histories of the remaining sensors. The ratio of α_i to α_o in Table 6.3 indicates that inoperative inboard ailerons create a 40% reduction in aileron effectiveness and a less responsive control system. Comparison of Figure 6.7a with Figure 6.5a reveals that the roll response time for the autopilot-aircraft system has increased from 1.8 seconds to approximately 5 seconds with some overshoot. This indicates that a closed-loop dynamics change has occurred and that an inboard aileron failure is a likely candidate as an explanation of the change.

The detection system outputs corroborate the subjective analysis in the preceding paragraph. Both the T_{mg} and T_{ms1} detection systems experience threshold crossings in the inboard aileron detection channels at 1.7 seconds. No further threshold crossings occur in the .3 second waiting period and both detection systems report inboard aileron failures at 2.0 seconds.[1] The presence of a single threshold crossing in Figure

[1] The small threshold crossing which appears in the $\hat{\epsilon}'_{ar}(k\Delta t)$ detection channel in Figure 6.7c is a consequence of the fact that failure detection theory is not analytically precise for sampled-data systems. Such crossings are discussed in Section 6.4.3.

Figure 6.7a Sensor outputs for a roll command of 2° at 1.6 seconds and an inboard aileron failure at 1.6 seconds.

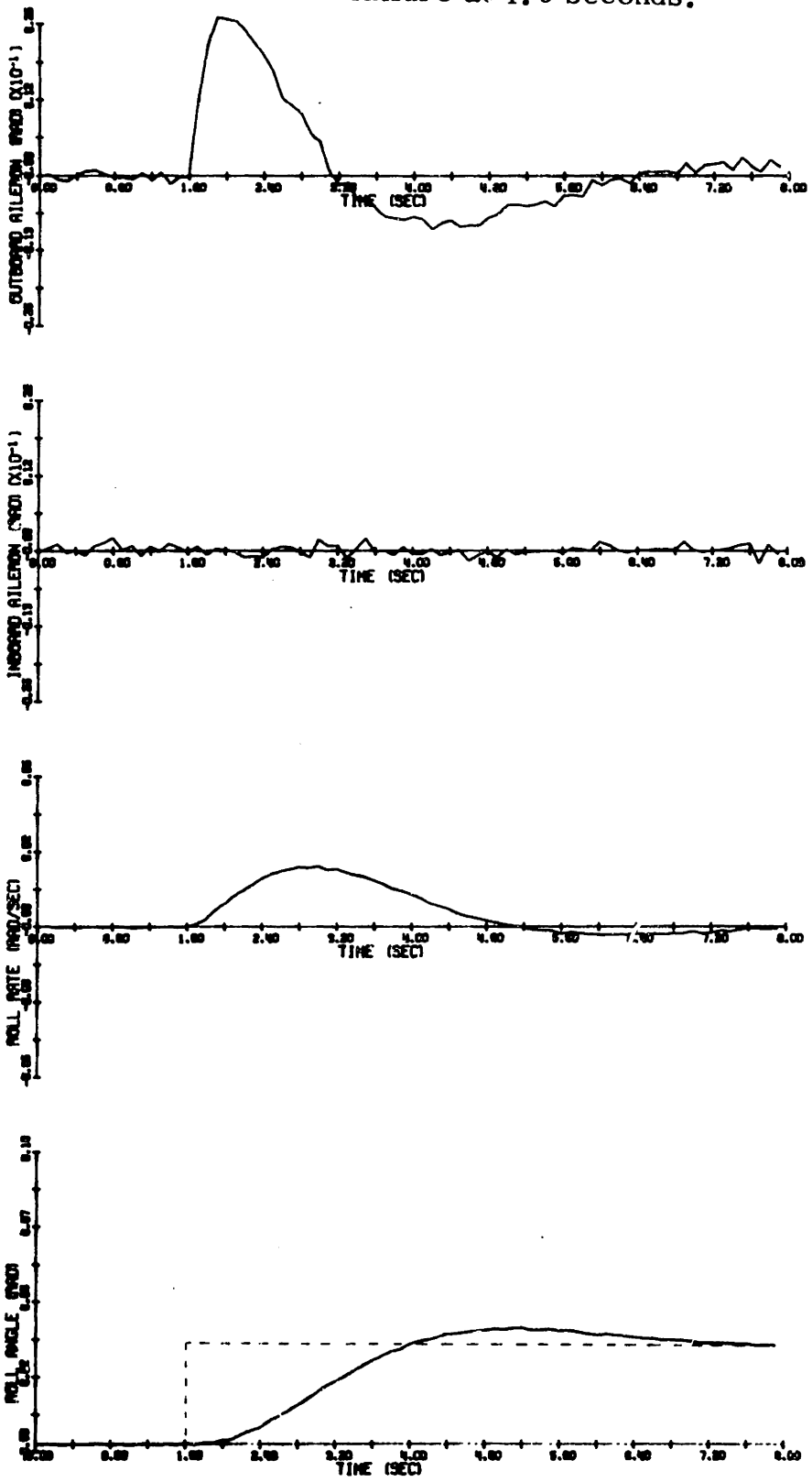


Figure 6.7b T_{mg} detection system output for simulation in Figure 6.7a. Arrow indicates time of threshold crossing.

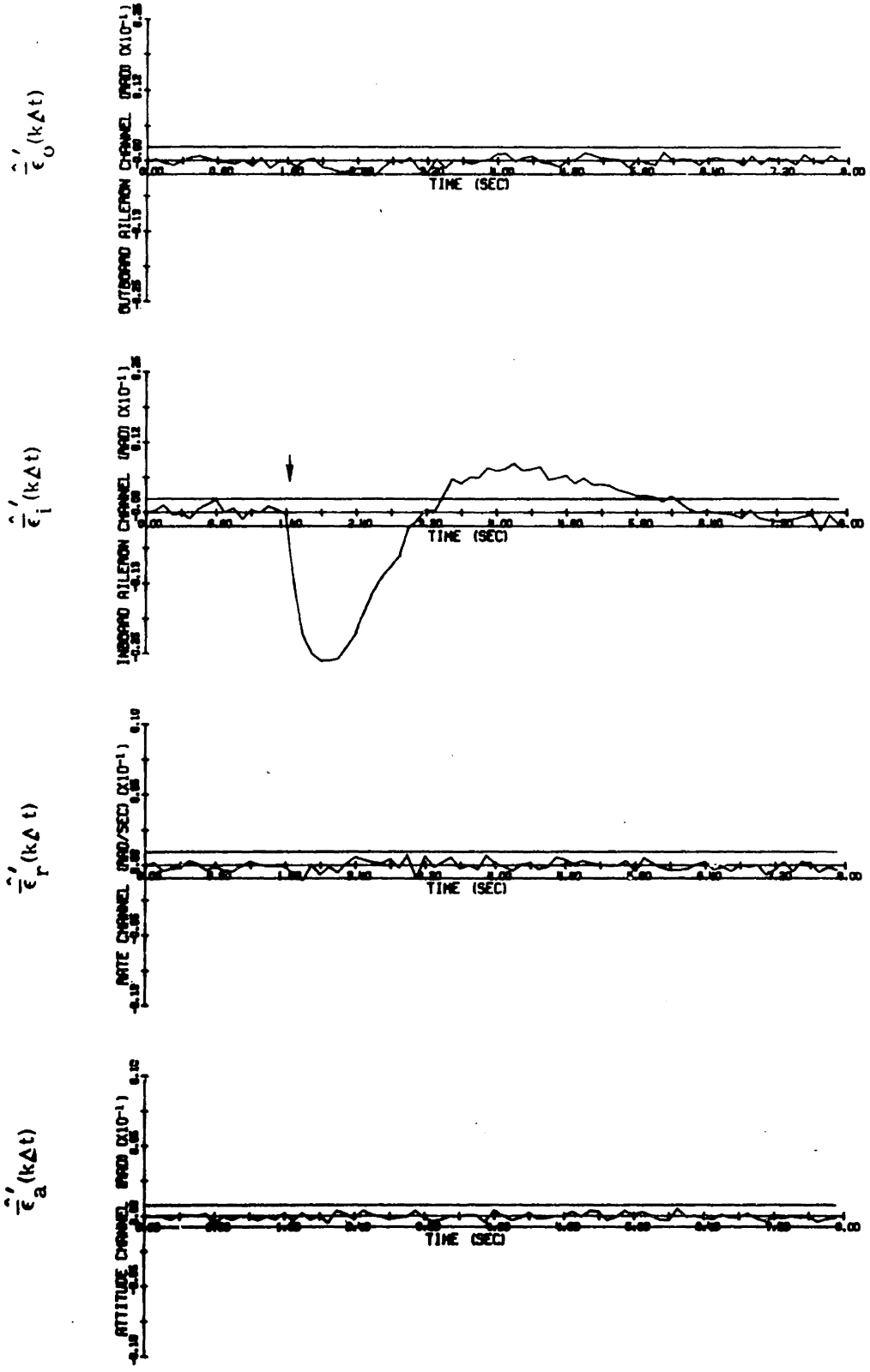
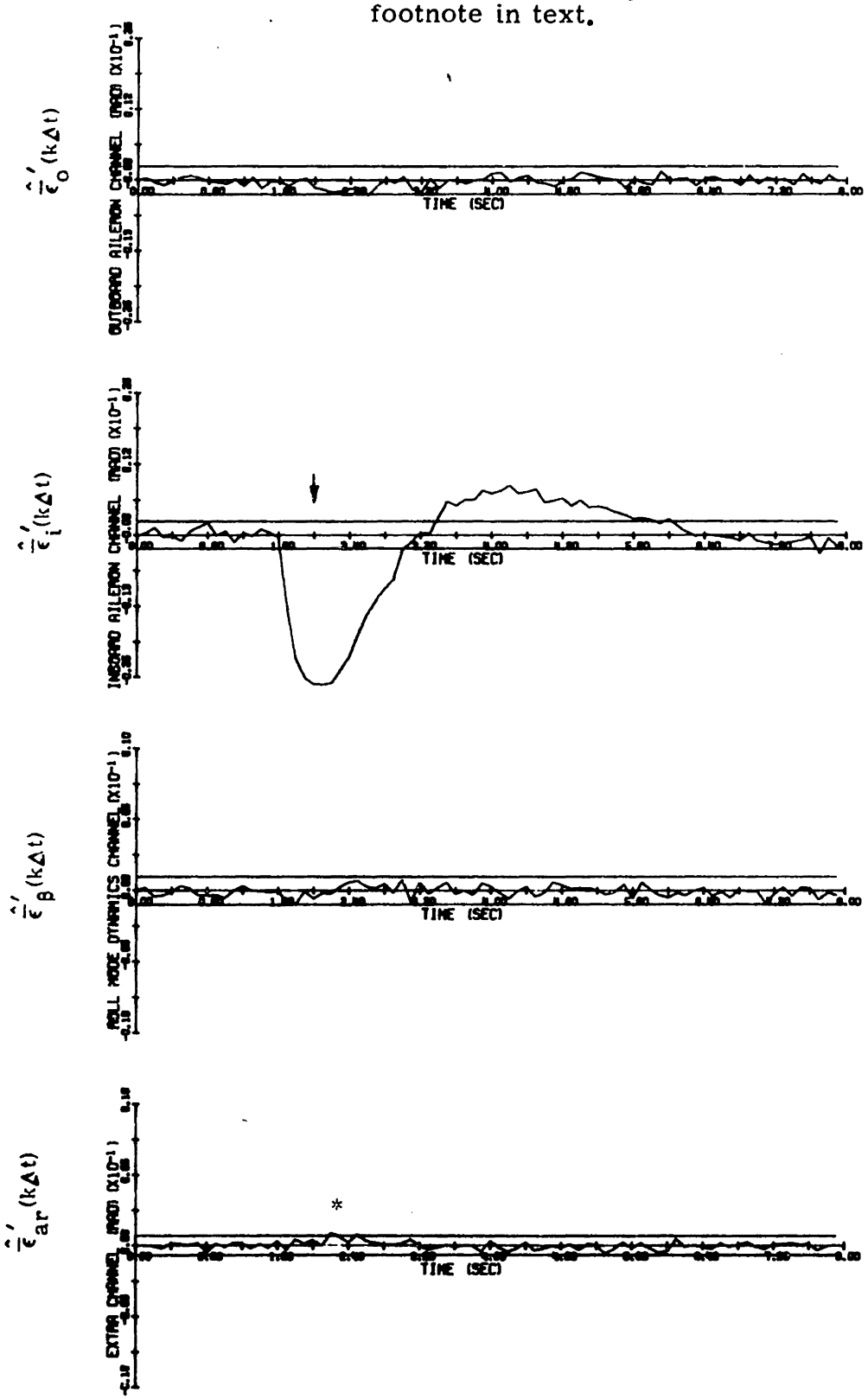


Figure 6.7c T_{ms1} detection system output for simulation in Figure 6.7a. Arrow indicates time of a threshold crossing. Asterisk is discussed in footnote in text.



6.7b indicates that an inboard aileron failure does not significantly degrade the estimates of the outboard aileron deflections, the roll rate, or the roll angle. The ability to provide accurate estimates of some state variables in a failed system is unique to failure detection systems.

It is interesting at this point to support a statement made in the previous paragraph by contrasting the simulation in Figure 6.7 with a similar run in which the detection filters $D_{\Delta K}$ in both failure detection systems are replaced by Kalman filters. As demonstrated in Section 6.4.2, the Kalman filters provide better state estimation accuracies in an unfailed system, but the simulation results in Figure 6.8 verify that a Kalman filter cannot be expected to provide any failure detection information or any acceptable state estimates in a failed system. The inboard aileron servo failure generates threshold crossings in all eight detection channels in Figures 6.8b and 6.8c. Neither detection system output can readily be deciphered to determine the source of the unexpectedly large residuals. Similarly, the fact that the residuals shown in Figure 6.8b for the outboard aileron deflection estimate, the roll rate estimate, and the roll angle estimate are significantly greater than the corresponding residuals in Figure 6.7b indicates that the failure detection system provides better estimates of these variables in the failed system than the Kalman filter does. Figures 6.8b and 6.8c are typical of the detection channel outputs generated by Kalman filters for any of the failures listed in Table 6.1.

Figures 6.6 and 6.7 demonstrate that the failure detection systems can readily detect and identify failures in which the failure mode results in a passively inoperative system component. It is worthwhile to demonstrate that the failure detection systems perform equally well for a non-passive failure. The following simulation illustrates an instance in which real-time use of failure detection information can allow a system to recover from an otherwise unstable configuration.

Figure 6.9 summarizes a simulation run in which no roll angle command was given. At 1.6 seconds the inboard aileron failed in an unstable manner and began a hard-over deflection. The failure model is similar to that in (6.1.18) except that the servo time constant becomes unstable:

Figure 6.8a Kalman filter residuals transformed into T_{mg} coordinates for simulation in Figure 6.7a

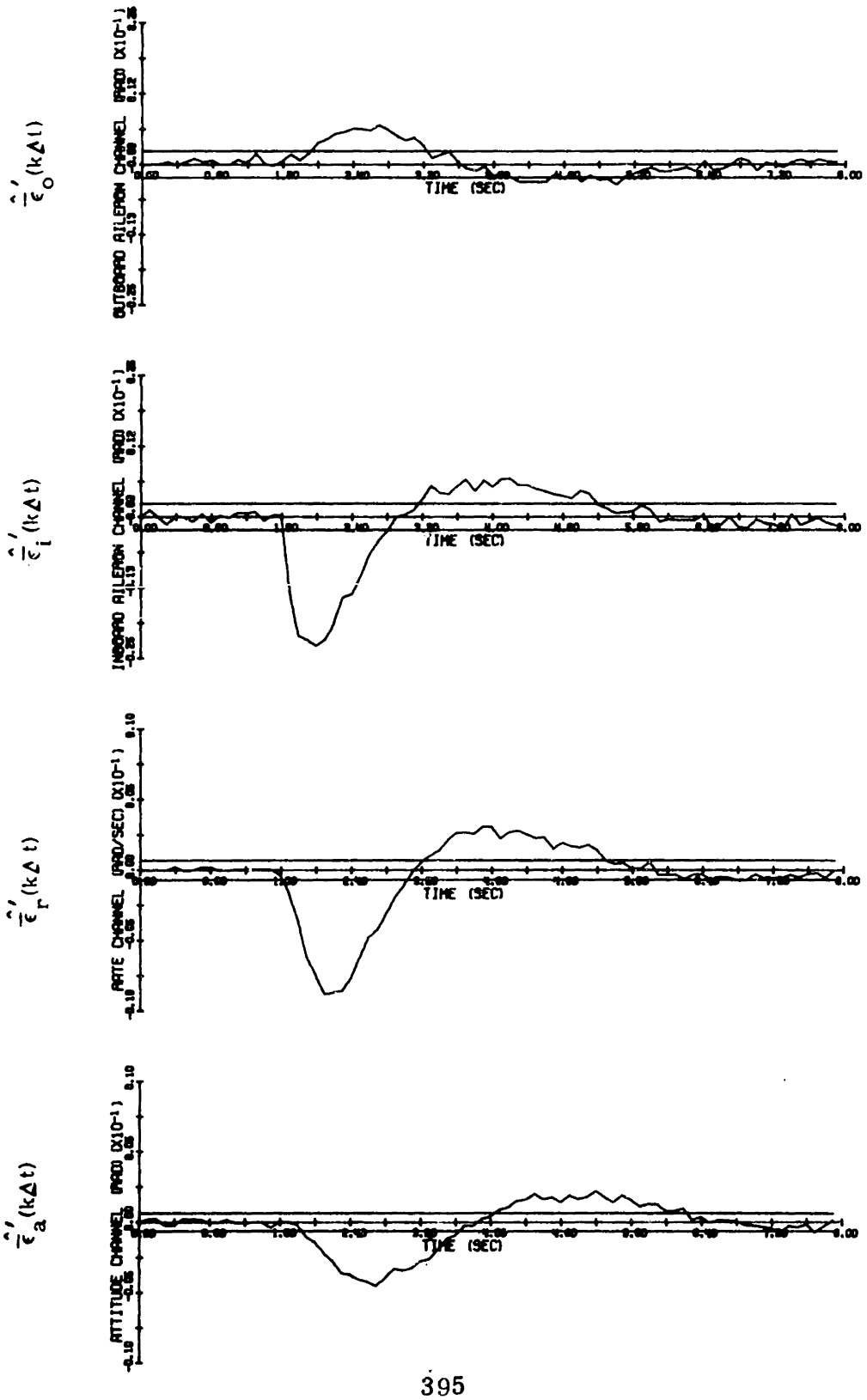


Figure 6.8b Kalman filter residuals transformed into T_{ms1} coordinates for simulation in Figure 6.7a.

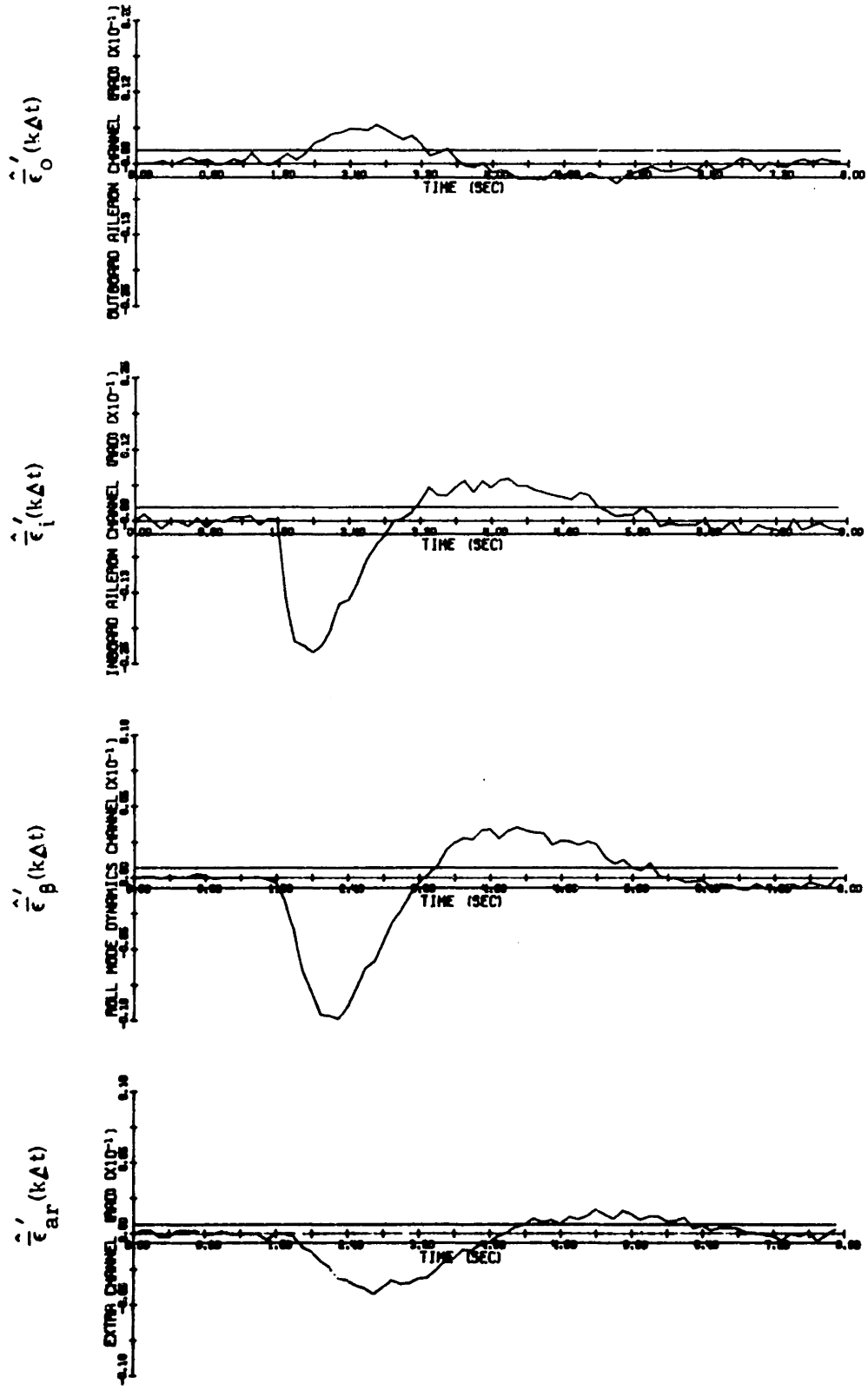


Figure 6.9a Sensor outputs for a simulation with no roll command. The inboard ailerons entered a hard-over failure mode at 1.6 seconds. At 2.0 seconds, the inboard ailerons were inactivated and at 2.4 seconds they had returned to the trim position

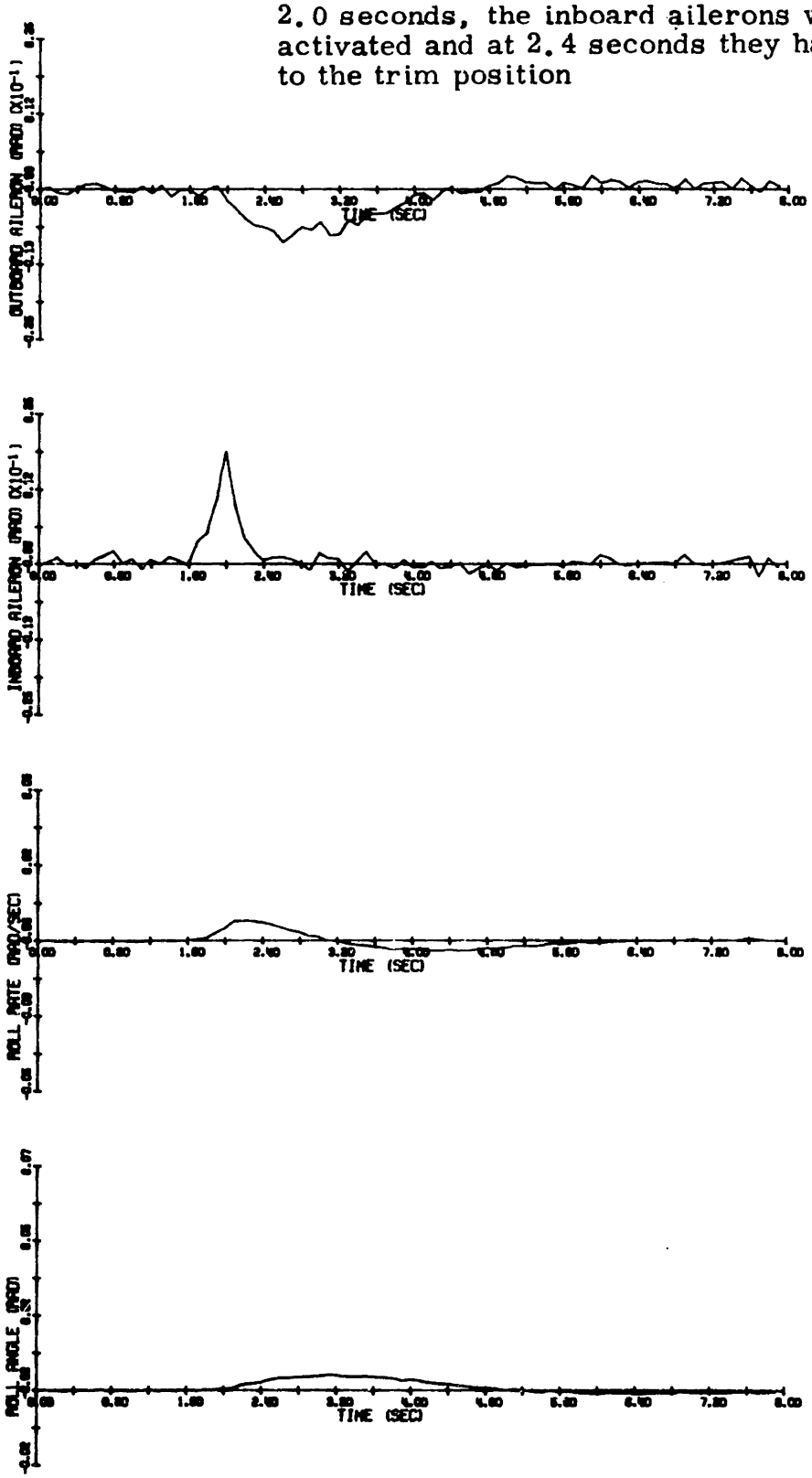


Figure 6.9b T_{mg} detection system output for simulation in Figure 6.9a. Arrow indicates time of threshold corssing.

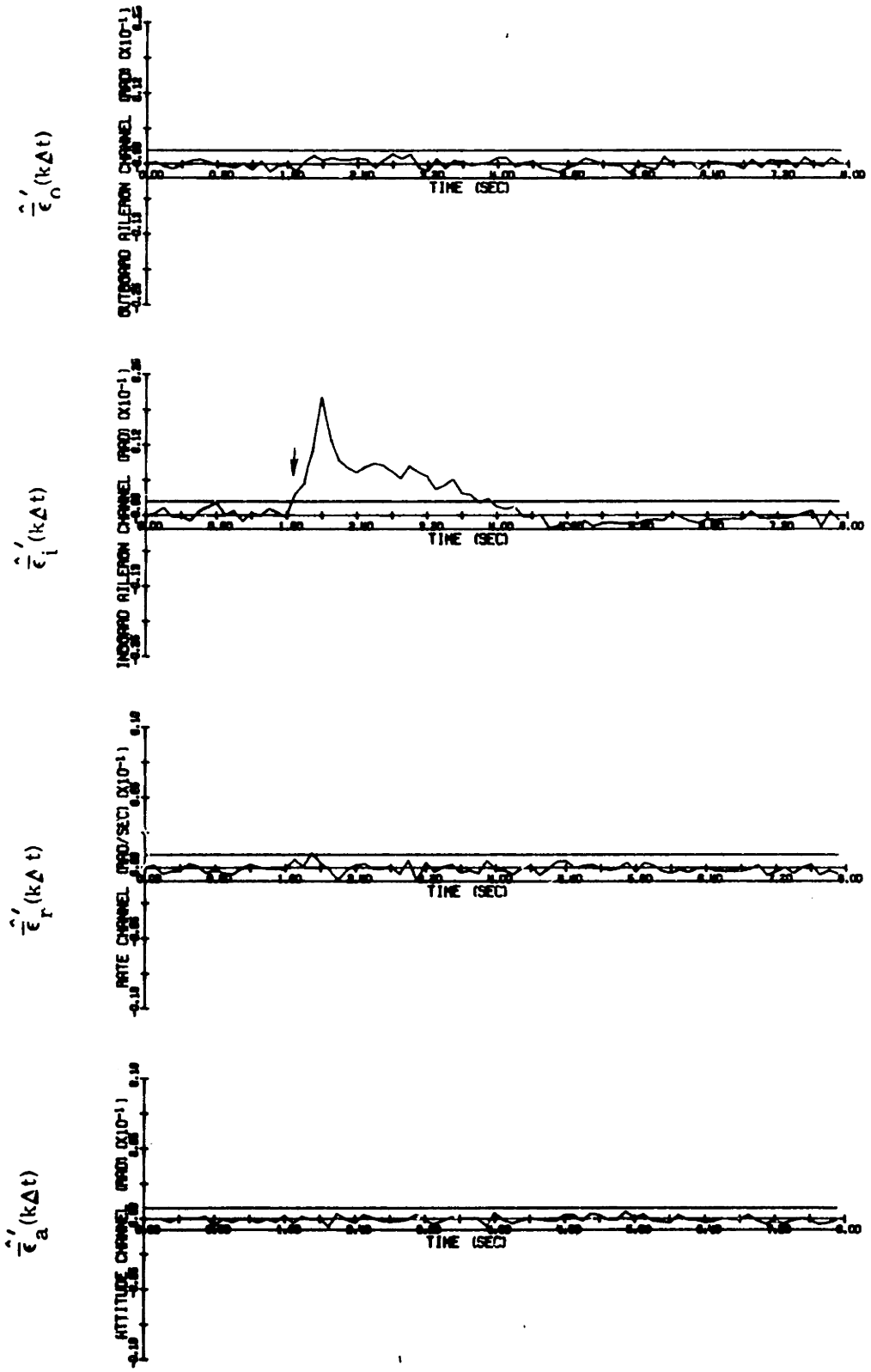
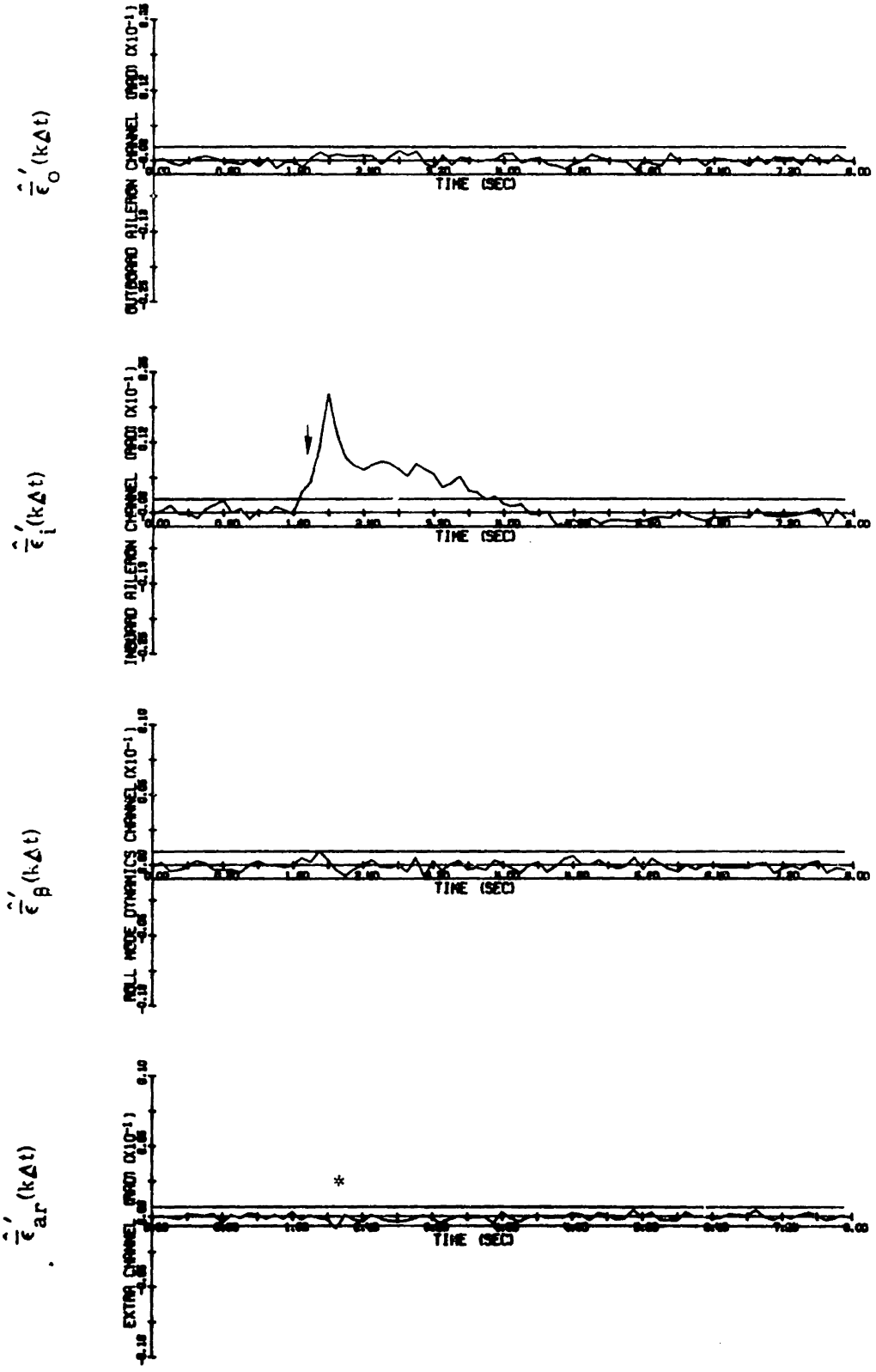


Figure 6.9c. T_{mgl} detection system output for simulation in Figure 6.9a. Arrow indicates time of threshold crossing. Asterisk is discussed in footnote following Figure 6.7.



$$\Delta A_i' = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ k_i k_a & 0 & k_i k_r & 8+\tau_i & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\Delta B_i' = \begin{bmatrix} 0 \\ 0 \\ 0 \\ -k_i k_a \\ 0 \end{bmatrix} \qquad (6.4.1)$$

Figure 6.9a depicts the inboard aileron divergence in the interval between 1.6 seconds and 2.0 seconds. At 2.0 seconds the aileron deflection is approximately 1°. By this time the autopilot has begun to command negative aileron deflections to counteract the measured roll rate, but only the unfailed outboard aileron responds to the command. Because the outboard aileron effectiveness happens to be greater than the inboard aileron effectiveness, the autopilot can ultimately regain control of the aircraft once the inboard ailerons are fully deflected.

Figures 6.9b and 6.9c present the detection system responses to the failure. At 1.7 seconds, both the T_{mg} and T_{ms1} detection systems experience threshold crossings in the inboard aileron channel ($\hat{\bar{\epsilon}}_i'(k\Delta t)$). No subsequent threshold crossings occur in the .3 second waiting period and both detection systems correctly identify the inboard aileron failure at 2.0 seconds. The detection law decision based upon these identifications is used at 2.0 seconds to inactivate the inboard ailerons. After a .4 seconds transient, they return to a trim position and the autopilot uses the outboard aileron to return the aircraft to the commanded attitude. The corrective action changes the aileron failure mode from active to passive. As indicated by the threshold crossings in the inboard aileron detection channels in the interval between 2.0 and 4.0 seconds, the inboard ailerons are no longer responsive to autopilot commands. For

any subsequent autopilot command the threshold crossings would reappear. The means for eliminating this possibility is to update k_i to zero in A, B, and D when the inboard ailerons are deactivated. The failure detection systems would then no longer expect the inboard ailerons to respond to autopilot commands and the threshold crossings would reappear only if the inboard ailerons were somehow re-activated.

The last simulation in this subsection is concerned with a rate gyro failure. One purpose is to contrast the detection of a sensor failure when the sensor output is used in the control law (rate gyro) with the case when the sensor output is not used in the control law (inboard aileron deflection sensor). In the latter instance, Figure 6.6 verified that the sensor failure must generate threshold crossings in at least two detection channels of each detection system. Because a rate gyro failure affects the dynamics of the autopilot-aircraft system, its failure can be constrained to a single detection channel. The failure mode presented in the following example is passive -- gyro drifts and gyro biases are analyzed in Section 6.4.3.

The simulation in Figure 6.10 further illustrates the manner in which real-time failure detection, identification, and correction can be used to prevent a failure from disrupting the autopilot-aircraft system performance. The aircraft is given a 2° commanded roll angle at 1.6 seconds and a 4° commanded roll angle at 5.6 seconds. The rate gyro fails in a null output mode at .8 seconds. The failure model is ΔA_r in (6.1.11) with Δa_{33} equal to zero. Because the rate gyro output does not deviate from the nominal output until after a roll command is applied, no threshold crossings appear in the detection systems until after 1.6 seconds. At 1.7 seconds the T_{mg} detection system detects a threshold crossing in the rate gyro channel ($\hat{\epsilon}_r'(k\Delta t)$) and at 2.0 seconds correctly identifies the rate gyro failure. The T_{ms1} detection system experiences threshold crossings in the roll dynamics channel ($\hat{\epsilon}_\beta'(k\Delta t)$) at 1.7 seconds and the attitude-rate channel ($\hat{\epsilon}_{ar}'(k\Delta t)$) at 1.9 seconds. At 2.0 seconds, the T_{ms1} detection system reports that a series of simultaneous failures has occurred (Table 6.2), but the second tier of the detection law correctly accepts the T_{mg} identification of a rate gyro failure because of the

Figure 6.10a Sensor outputs for a simulation with a 2° roll command at 1.6 seconds and a 4° roll command at 5.6 seconds. The rate gyro failed at .8 seconds and was repaired at 2.0 seconds.

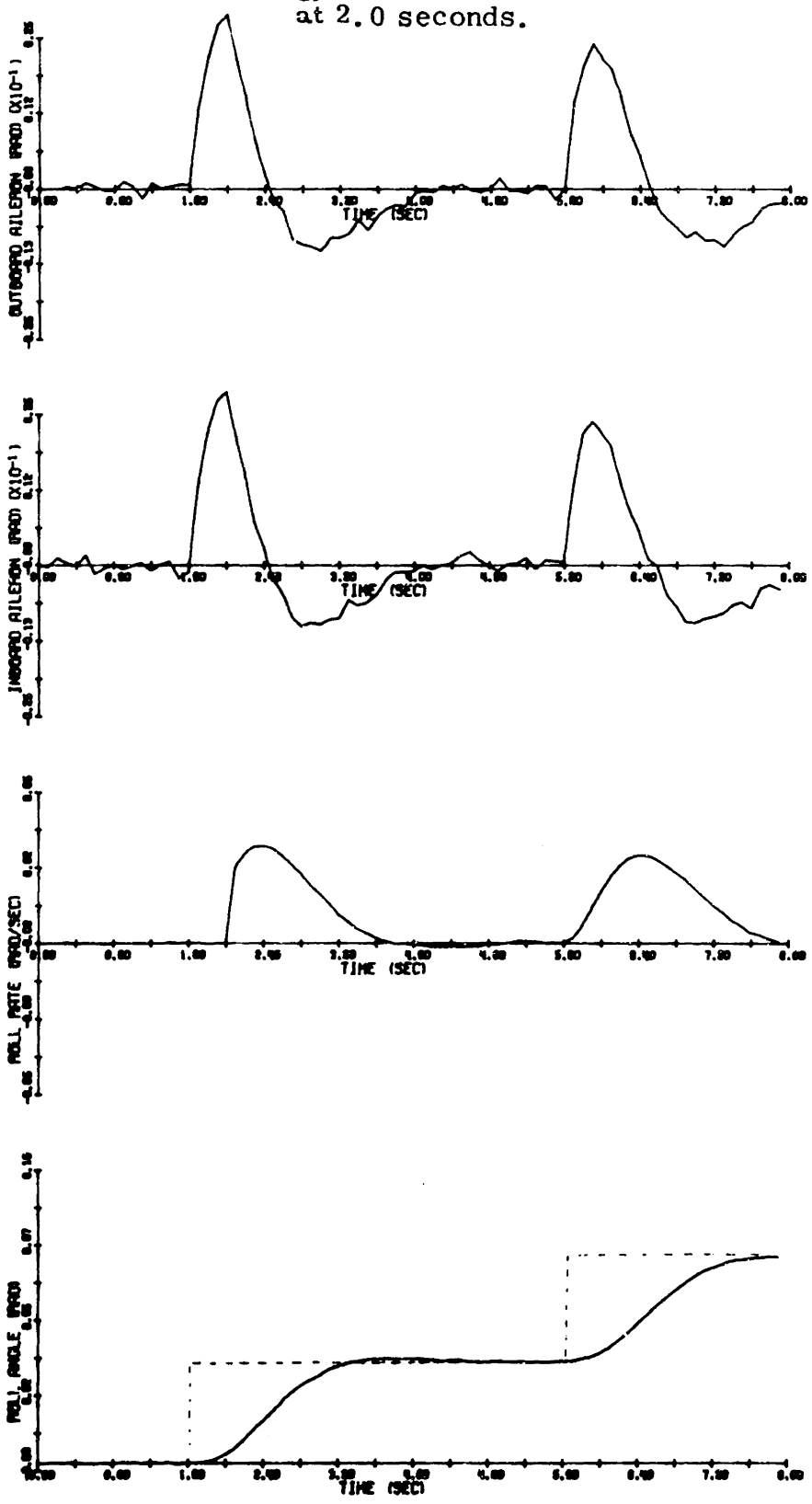


Figure 6.10b T_{mg} detection system output for simulation in Figure 6.10a. Arrow indicates time of threshold crossing. $\hat{\bar{\epsilon}}_i'(k\Delta t)$ and $\hat{\bar{\epsilon}}_o'(k\Delta t)$ threshold crossings are transients associated with rate gyro repair.

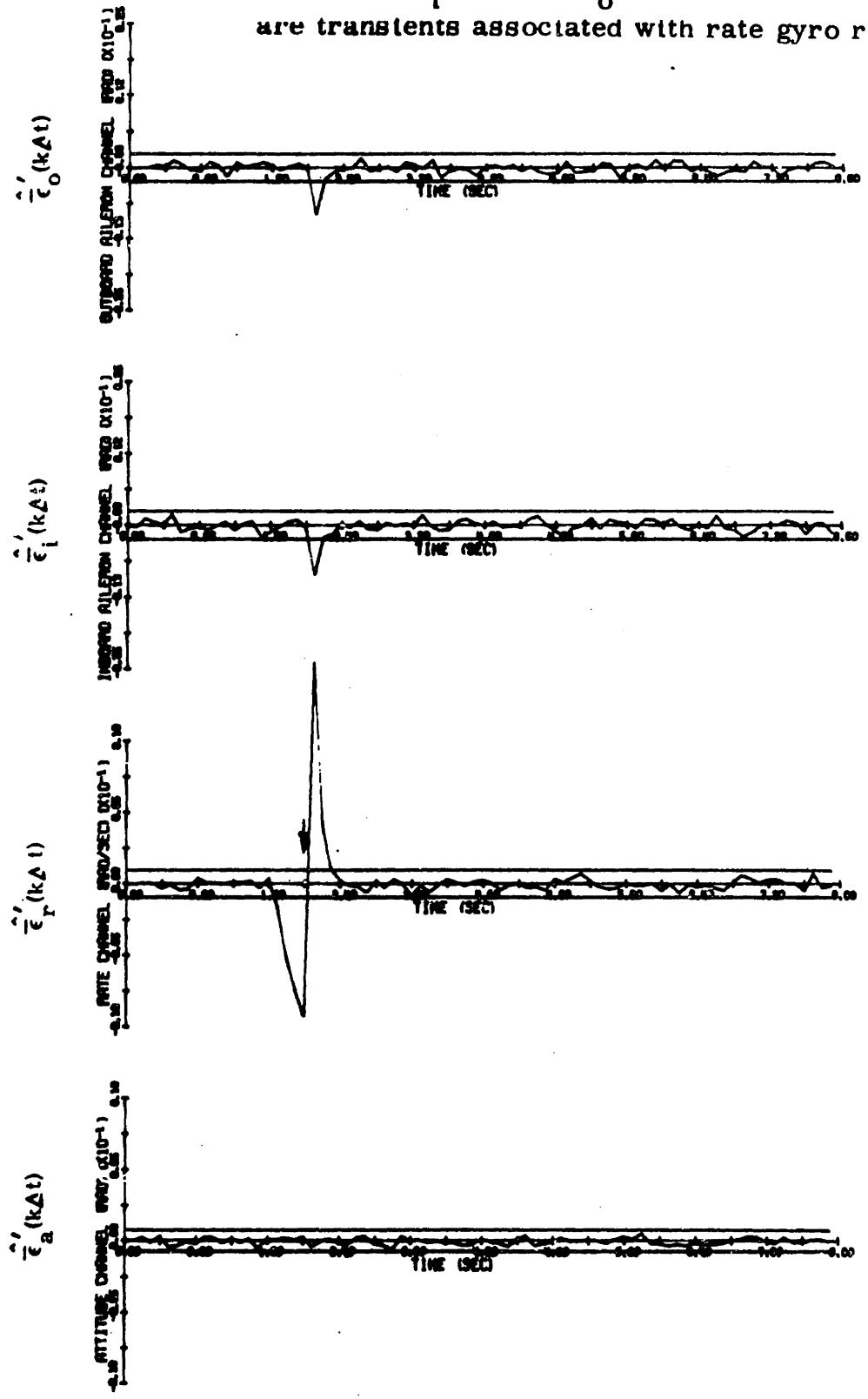
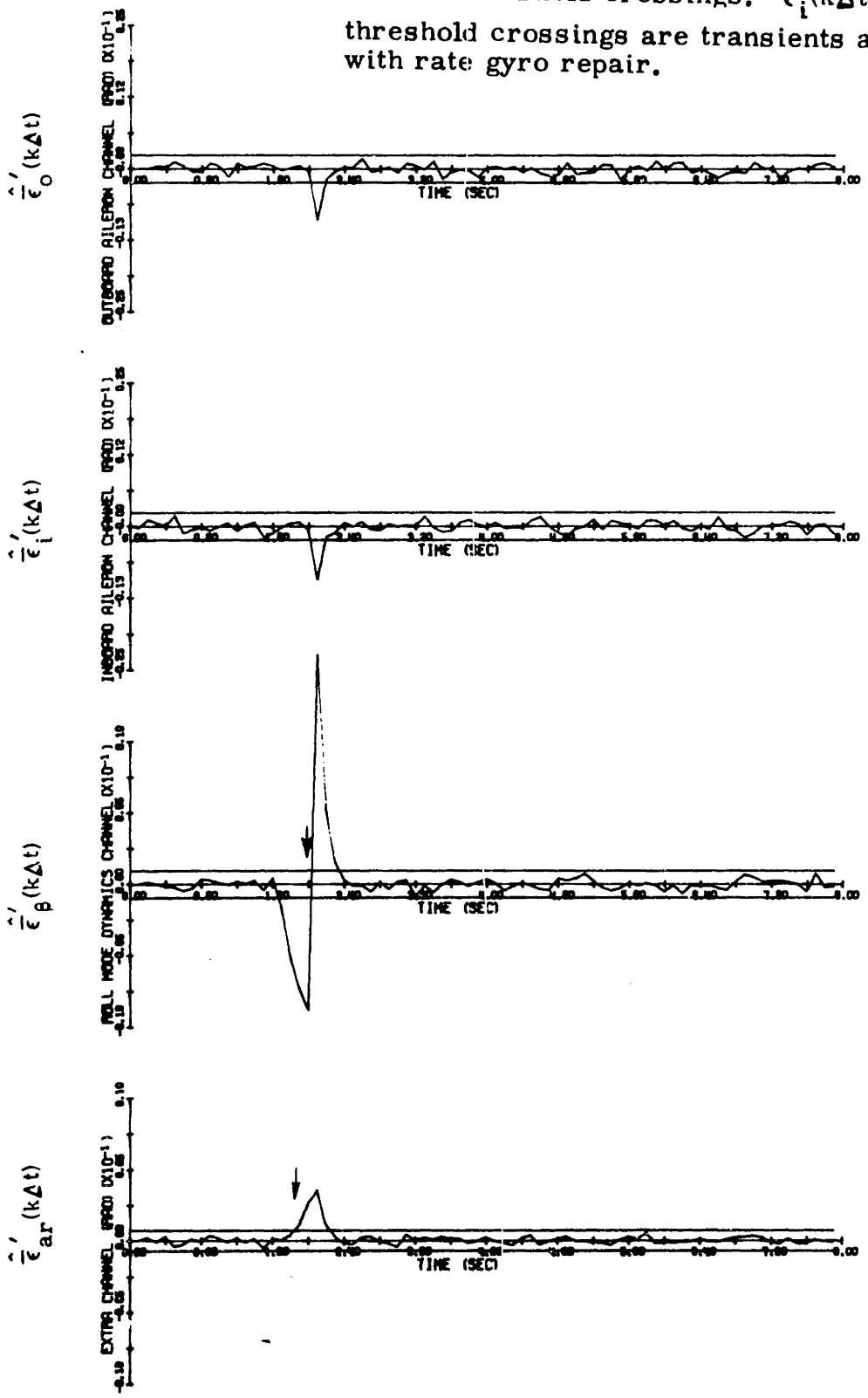


Figure 6.10c T_{ms1} detection system output for simulation in Figure 6.10a. Arrow indicates time of threshold crossings. $\hat{\bar{\epsilon}}_i'(k\Delta t)$ and $\hat{\bar{\epsilon}}_o'(k\Delta t)$ threshold crossings are transients associated with rate gyro repair.



single threshold crossing reported by that detection channel. In the simulation, a standby rate gyro is brought on-line at 2.0 seconds and the autopilot-aircraft system resumes nominal behavior. Comparison of the attitude gyro response to the 2° roll command with the response to the 4° command or with Figure 6.5a verifies that the rapid identification and correction results in a nearly nominal roll response. The replacement of the rate gyro introduces transient threshold crossings at 2.1 seconds in seven of the eight detection channels in Figures 6.10b and 6.10c, but all the transients decay by 2.4 seconds. This behavior illustrates the failure detection system inability to constrain a residual to the appropriate detection channel when the residual undergoes a large change from one measurement to the next. In this case, the rate gyro repair causes a 40σ change in the rate gyro residual $\hat{\epsilon}_r'(k\Delta t)$ between 2.0 seconds and 2.1 seconds. This phenomenon and its importance in the design of failure detection systems for sampled-data systems is discussed in Section 6.4.2.

6.4.2 State Estimation

The single run simulations in Section 6.4.1 indicate that for a measurement sampling period of .1 seconds and the system parameters of Table 6.3, the failure detection systems and detection law designed in Sections 6.1 - 6.3 can provide useful failure detection information for the autopilot-aircraft system. Although it was not verified in Section 6.4.1, in certain instances the failure detection systems can also provide state estimates comparable to those produced by a Kalman filter. This subsection is concerned with the accuracy of those state estimates as a function of:

i) the measurement sampling period,
ii) the relative variances of the measurement and system noises,
iii) the failures which a failure detection system is designed to detect, and
iv) the failures which have occurred in the system.

The analysis presented is not complete; but, along with Section 6.4.3, it serves to focus attention on the critical factors in the design of a failure

detection system. Only the roll attitude and roll rate estimates generated by the gyro-aileron failure detection system are considered in detail.

The base normal form defined for the gyro-aileron failure detection system in (6.2.29) can be used to partition the update equation for the sampled-data state estimator defined by (5.2.5) and (5.2.6):

$$\hat{x}_m(k\Delta t) = \hat{\bar{x}}_m(k\Delta t) + \hat{D}_{\Delta K}\hat{\bar{\epsilon}}'(k\Delta t)$$

$$= \hat{\bar{x}}_m(k\Delta t) + (\Delta t\hat{D}_{FR} + \hat{D}_{\Delta\Psi K} + \hat{D}_{\Delta\Gamma K})\hat{\bar{\epsilon}}'(k\Delta t) \tag{5.2.6}$$

The partitions for the state estimates are similar to those defined for the continuous estimator:

$$\hat{\bar{x}}_m(k\Delta t) = \begin{bmatrix} \hat{\bar{x}}_{ma}(k\Delta t) \\ \hline \hat{\bar{x}}_{mr}(k\Delta t) \\ \hline \hat{\bar{x}}_{mi}(k\Delta t) \\ \hline \hat{\bar{x}}_{mo}(k\Delta t) \end{bmatrix} = \begin{bmatrix} \beta_\varphi \bar{\varphi}_{ma}(k\Delta t) + \dot{\bar{\varphi}}_m(k\Delta t) - \dot{\bar{\varphi}}_{m0}(k\Delta t) \\ \bar{\varphi}_{ma}(k\Delta t) \\ \hline \beta_\omega(\dot{\bar{\varphi}}_{mr}(k\Delta t) - \dot{\bar{\varphi}}_m(k\Delta t)) \\ \dot{\bar{\varphi}}_{mr}(k\Delta t) \\ \hline \bar{\delta}_{mi}(k\Delta t) \\ \hline \bar{\delta}_{mo}(k\Delta t) \end{bmatrix}$$

$$\hat{x}_m(k\Delta t) = \begin{bmatrix} \hat{x}_{ma}(k\Delta t) \\ \hat{x}_{mr}(k\Delta t) \\ \hat{x}_{mi}(k\Delta t) \\ \hat{x}_{mo}(k\Delta t) \end{bmatrix} = \begin{bmatrix} \beta_\varphi \varphi_{ma}(k\Delta t) + \dot{\varphi}_m(k\Delta t) - \dot{\varphi}_{m0}(k\Delta t) \\ \overline{\varphi}_{ma}(k\Delta t) \\ \beta_\varphi(\dot{\varphi}_{mr}(k\Delta t) - \dot{\varphi}_m(k\Delta t)) \\ \dot{\varphi}_{mr}(k\Delta t) \\ \delta_{mi}(k\Delta t) \\ \delta_{mo}(k\Delta t) \end{bmatrix} \qquad (6.4.2)$$

$\overline{\varphi}_{ma}(k\Delta t)$ is the estimate of the roll angle at $k\Delta t$ based only on measurements made before $k\Delta t$, and $\varphi_{ma}(k\Delta t)$ is the updated estimate based on both $\overline{\varphi}_{ma}(k\Delta t)$ and the measurements at $k\Delta t$ etc. The measurement residuals are those defined in (6.3.2) for the T_{mg} detection system:

$$\hat{\overline{\epsilon}}_g'(k\Delta t) = \begin{bmatrix} \hat{\overline{\epsilon}}_a'(k\Delta t) \\ \hat{\overline{\epsilon}}_r'(k\Delta t) \\ \hat{\overline{\epsilon}}_i'(k\Delta t) \\ \hat{\overline{\epsilon}}_o'(k\Delta t) \end{bmatrix} \qquad (6.3.2)$$

where $\hat{\overline{\epsilon}}_a'(k\Delta t)$ is the difference between the attitude gyro output and its predicted output etc. With $\hat{D}_{\Delta\Gamma K}$ equal to zero and the partitions of $\hat{D}_{FR}$ and $\hat{D}_{\Delta\Psi K}$ as defined in (6.2.32) and (6.2.46) respectively; the first two partitions of the update equation can be written:

$$\hat{x}_{ma}(k\Delta t) = \hat{\overline{x}}_{ma}(k\Delta t) + (\Delta t\, \hat{D}_{Fa} + \hat{D}_{\Delta\Psi Ka})\, \hat{\overline{\epsilon}}_g'(k\Delta t)$$

$$\hat{x}_{mr}(k\Delta t) = \hat{\overline{x}}_{mr}(k\Delta t) + (\Delta t\, \hat{D}_{Fr} + \hat{D}_{\Delta\Psi Kr})\, \hat{\overline{\epsilon}}_g'(k\Delta t) \qquad (6.4.3)$$

Further substitution yields the desired relationships:

$$\hat{x}_{ma}(k\Delta t) = \bar{\hat{x}}_{ma}(k\Delta t) + d_{\Delta\Psi Ka}\bar{\hat{\epsilon}}_a'(k\Delta t) + \begin{bmatrix} \Delta t\alpha_i \\ 0 \end{bmatrix}\bar{\hat{\epsilon}}_i'(k\Delta t) + \begin{bmatrix} \Delta t\alpha_o \\ 0 \end{bmatrix}\bar{\hat{\epsilon}}_o'(k\Delta t)$$

$$\hat{x}_{mr}(k\Delta t) = \bar{\hat{x}}_{mr}(k\Delta t) + d_{\Delta\Psi Kr}\bar{\hat{\epsilon}}_r'(k\Delta t) + \begin{bmatrix} 0 \\ \Delta t\alpha_i \end{bmatrix}\bar{\hat{\epsilon}}_i'(k\Delta t) + \begin{bmatrix} 0 \\ \Delta t\alpha_o \end{bmatrix}\bar{\hat{\epsilon}}_o'(k\Delta t) \qquad (6.4.4)$$

If $D_{\Delta\Psi K}$ were defined by (6.2.37) and (6.2.38), $d_{\Delta\Psi Ka}$ and $d_{\Delta\Psi Kr}$ in (6.4.4) would be replaced by $\Delta t\, d_{\Psi Ka}$ and $\Delta t\, d_{\Psi Kr}$.

These equations show the difference between a detection filter-state estimator and a Kalman filter. A Kalman filter would use the residuals of the four measurements at $k\Delta t$ in an optimal sense to update the estimates. In contrast, only $\bar{\hat{\epsilon}}_a'(k\Delta t)$ in the first equation and $\bar{\hat{\epsilon}}_r'(k\Delta t)$ in the second equation can be used optimally because $d_{\Delta\Psi Ka}$ and $d_{\Delta\Psi Kr}$ are the only gains in (6.4.4) which can be varied to improve the state estimates. The gains associated with $\bar{\hat{\epsilon}}_i'(k\Delta t)$ and $\bar{\hat{\epsilon}}_o'(k\Delta t)$ are defined by the $\hat{D}_{Fa}$ and $\hat{D}_{Fr}$ partitions of $\hat{D}_{FR}$ and are fixed by the failure detection requirement that $\hat{A} - \hat{D}_{FR}\hat{C}$ be block diagonal. In particular, they are fixed by the requirement to detect aileron failures. Contrary to the case of a Kalman filter, the use of the aileron deflection residuals in (6.4.4) does not necessarily improve the state estimates. For the measurement accuracies presented in Table 6.3, the $\Delta t\alpha_i$ and $\Delta t\alpha_o$ gains are large enough that $\bar{\hat{\epsilon}}_i'(k\Delta t)$ and $\bar{\hat{\epsilon}}_o'(k\Delta t)$ increase the uncertainty in $\bar{\hat{x}}_{ma}(k\Delta t)$ for large measurement sampling periods. In the event that an aileron failure occurs, however, Figures 6.7 and 6.8 verified that the roll angle and roll rate estimates generated by a Kalman filter are severely degraded, but those generated by the gyro-aileron failure detection system are unaltered. The general rule-of-thumb is that if a failure detection system is designed to detect a particular failure, the only state

estimates degraded by the occurrence of the failure are those directly associated with the failed component, e.g., if the inboard aileron fails, the gyro-aileron failure detection system can estimate every state component except the inboard aileron deflection.

One procedure for improving the state estimates in the unfailed system is to remove the requirement for detecting aileron failures -- although this makes the state estimates sensitive to aileron failures. The gyro aileron failure detection system was required to detect aileron failures in order to satisfy (6.2.23) and thereby make the algorithms in Section 5.2 easier to implement. The algorithms are still applicable if the failure detection system is designed for the gyro failures only. As indicated in Section 6.2.2, definition of the base normal form for the two event vectors $\tilde{\tilde{f}}_a$ and $\tilde{\tilde{f}}_r$ necessitates the use of Appendix A.4, but for this example the same transformations T and T_m are found. From a state estimation vantage point, the important difference in the failure detection constraints on the choice of $\hat{D}_{\Delta K}$ is that (4.4.17) and the discrete equivalent of (4.4.20) require the last two columns of $\hat{D}_{FR}$ and $\hat{D}_{\Delta\Psi K}$ to be zero:

$$\hat{D}_{FR} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -k_i k_a & -k_i k_r & 0 & 0 \\ -k_o k_a & -k_o k_r & 0 & 0 \end{bmatrix} = \begin{bmatrix} \hat{D}_{Fa} \\ \hat{D}_{Fr} \\ \hat{D}_{Fi} \\ \hat{D}_{Fo} \end{bmatrix}$$

$$\hat{D}_{\Delta\Psi K} = \begin{bmatrix} d_{\Delta\Psi Ka} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & d_{\Delta\Psi Kr} & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \hat{D}_{\Delta\Psi Ka} \\ \hat{D}_{\Delta\Psi Kr} \\ \hat{D}_{\Delta\Psi Ki} \\ \hat{D}_{\Delta\Psi Ko} \end{bmatrix} \qquad (6.4.5)$$

The change in $\hat{D}_{FR}$ from (6.2.32) eliminates the $\Delta t\alpha_i$ and $\Delta t\alpha_o$ terms in the equivalent to (6.4.4). The increased freedom in the choice of $\hat{D}_{\Delta K}$ which results from the decrease in the failure detection requirements is indicated by the fact that the discrete equivalent of (4.4.25) no longer requires $\hat{D}_{\Delta\Gamma K}$ to be zero:

$$\hat{D}_{\Delta\Gamma K} = \begin{bmatrix} 0 & 0 & \hat{D}_{\Delta\Gamma a} \\ 0 & 0 & \\ 0 & 0 & \hat{D}_{\Delta\Gamma r} \\ 0 & 0 & \\ 0 & 0 & \hat{D}_{\Delta\Gamma m} \\ 0 & 0 & \end{bmatrix} \qquad (6.4.6)$$

$\hat{D}_{\Delta\Gamma a}$, $\hat{D}_{\Delta\Gamma r}$, and $\hat{D}_{\Delta\Gamma m}$ are 2 x 2 matrices and can be chosen arbitrarily.

If the failure detection system is designed only to detect gyro failures, the fact that $\hat{D}_{Fa}$ and $\hat{D}_{Fr}$ in (6.4.5) are zero implies that the equivalent to (6.4.4) is:

$$\hat{x}_{ma}(k\Delta t) = \hat{\bar{x}}_{ma}(k\Delta t) + d_{\Delta\Psi Ka}\,\hat{\bar{\epsilon}}_a'(k\Delta t) + \hat{D}_{\Delta\Gamma a}\begin{bmatrix} \hat{\bar{\epsilon}}_i'(k\Delta t) \\ \hat{\bar{\epsilon}}_o'(k\Delta t) \end{bmatrix}$$

$$\hat{x}_{mr}(k\Delta t) = \hat{\bar{x}}_{mr}(k\Delta t) + d_{\Delta\Psi Kr}\,\hat{\bar{\epsilon}}_r'(k\Delta t) + \hat{D}_{\Delta\Gamma r}\begin{bmatrix} \hat{\bar{\epsilon}}_i'(k\Delta t) \\ \hat{\bar{\epsilon}}_o'(k\Delta t) \end{bmatrix} \qquad (6.4.7)$$

The difference between (6.4.4) and (6.4.7) is that the gain matrices $\hat{D}_{\Delta\Gamma a}$ and $\hat{D}_{\Delta\Gamma r}$ are not fixed by failure detection constraints and can be selected solely for the purpose of improving the accuracy of $\hat{x}_{ma}(k\Delta t)$ and $\hat{x}_{mr}(k\Delta t)$. The new failure detection system is still not a Kalman filter because only three of the four residuals at $k\Delta t$ can be used to update each of $\hat{\bar{x}}_{ma}(k\Delta t)$ and $\hat{\bar{x}}_{mr}(k\Delta t)$, but the freedom to choose the

gains for $\hat{\bar{\epsilon}}_i'(k\Delta t)$ and $\hat{\bar{\epsilon}}_o'(k\Delta t)$ can be expected to improve the value of the failure detection system as a state estimator. The one difficulty is that the algorithms provided in Section 5.2 for choosing $\hat{D}_{\Delta\Gamma a}$ and $\hat{D}_{\Delta\Gamma r}$ cannot be stated in a closed form unless the parameter optimization procedure is used. For the measurement uncertainties given in Table 6.3, the gyro outputs are significantly more accurate than the aileron deflection sensors. As is verified in the analysis which follows, a reasonable choice is to not use the aileron residuals at all:

$$\hat{D}_{\Delta\Gamma a} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

$$\hat{D}_{\Delta\Gamma r} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \qquad (6.4.8)$$

(6.4.7) then becomes:

$$\hat{x}_{ma}(k\Delta t) = \hat{\bar{x}}_{ma}(k\Delta t) + d_{\Delta\Psi Ka}\,\hat{\bar{\epsilon}}_a'(k\Delta t)$$

$$\hat{x}_{mr}(k\Delta t) = \hat{\bar{x}}_{mr}(k\Delta t) + d_{\Delta\Psi Kr}\,\hat{\bar{\epsilon}}_r'(k\Delta t) \qquad (6.4.9)$$

Because (6.4.9) differs from (6.4.4), the optimal $d_{\Delta\Psi Ka}$ for the gyro failure detection system in general differs from the optimal choice for the gyro-aileron failure detection system, etc.

The second component of $\hat{x}_{ma}(k\Delta t)$ is $\varphi_{ma}(k\Delta t)$ and the second component of $\hat{x}_{mr}(k\Delta t)$ is $\dot{\varphi}_{mr}(k\Delta t)$. In the unfailed system these correspond to the roll angle and the roll rate estimates respectively. Figure 6.11 compares the standard deviation of $\varphi_{ma}(k\Delta t)$ and $\dot{\varphi}_{mr}(k\Delta t)$ in the unfailed system for:

i) the gyro-aileron failure detection system
ii) the gyro failure detection system defined above, and
iii) a Kalman filter

as a function of the measurement sampling period. The measurement variances and aerodynamic moment variances are those given in Table

6.3. Both the discrete optimization procedure based on (6.2.45) and the continuous optimization procedure based on (6.2.37) and (6.2.38) were used to determine filters for the gyro-aileron failure detection system; only the discrete optimization procedure was used for the gyro failure detection system.

Figure 6.11a indicates that the $\varphi_{ma}(k\Delta t)$ estimate from the gyro-aileron failure detection system has approximately twice the standard deviation of the estimate from the corresponding Kalman filter. For a measurement sampling period of .1 seconds, the gyro-aileron failure detection system estimate has a σ of 1.2×10^{-4} radians and the Kalman filter estimate has a σ of $.56 \times 10^{-4}$ radians. By comparison, the σ for a single attitude gyro measurement is 2.1×10^{-4} radians. The gyro failure detection system offers a modest improvement in estimation accuracy over the gyro-aileron failure detection system -- a σ of $.97 \times 10^{-4}$ radians at .1 second -- but it does not approach the accuracy of a Kalman filter. To some extent this is because $\hat{D}_{\Delta\Gamma a}$ and $\hat{D}_{\Delta\Gamma r}$ in (6.4.8) were not chosen optimally, but the principal reason is that the Kalman filter can use the relatively accurate rate gyro output to improve the roll angle estimate and (6.4.9) states that the gyro failure detection system cannot. The offsetting advantage of the gyro failure detection system is that it can generate accurate roll angle estimates if the rate gyro has failed, but a Kalman filter cannot.

The significant difference in the state estimation capabilities of the two failure detection systems is indicated in Figure 6.11b. With the .1 second measurement sampling period, the $\dot{\varphi}_{mr}(k\Delta t)$ estimate from the gyro-aileron failure detection system has a σ of 2.34×10^{-4} radians/second when the discrete optimization procedure is used to design the failure detection filter. This uncertainty is greater than the σ of 2.1×10^{-4} radians/second if the rate gyro outputs were used directly for the roll rate estimation without any measurement filtering at all. It implies that the gyro-aileron failure detection system cannot be expected to yield useful roll rate estimates for a measurement sampling period as long as .1 seconds. By contrast, the gyro failure detection system generates a roll rate estimate of comparable accuracy to the Kalman filter estimate. This accuracy justifies the reasonableness of the choice

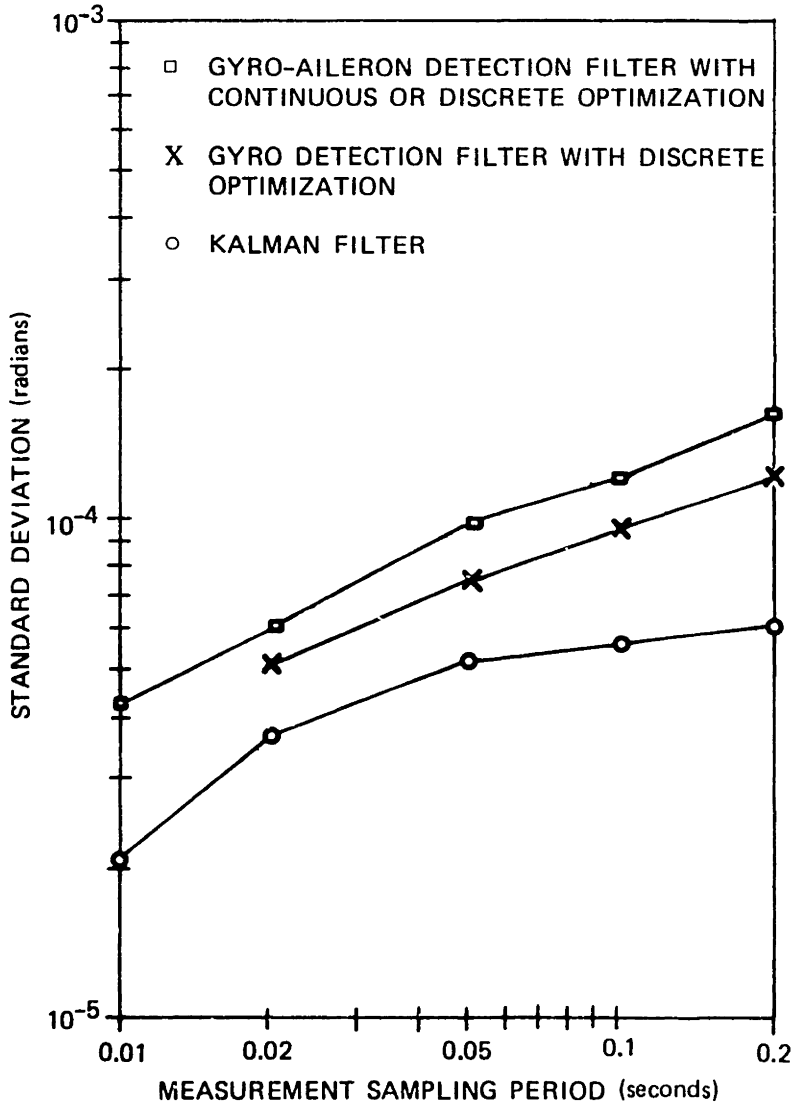


Figure 6.11a Roll estimation accuracies for gyro-aileron failure detection system, gyro failure detection system, and Kalman filter as a function of measurement sampling period. Measurement variances and aerodynamic moment variance are those of Table 6.3

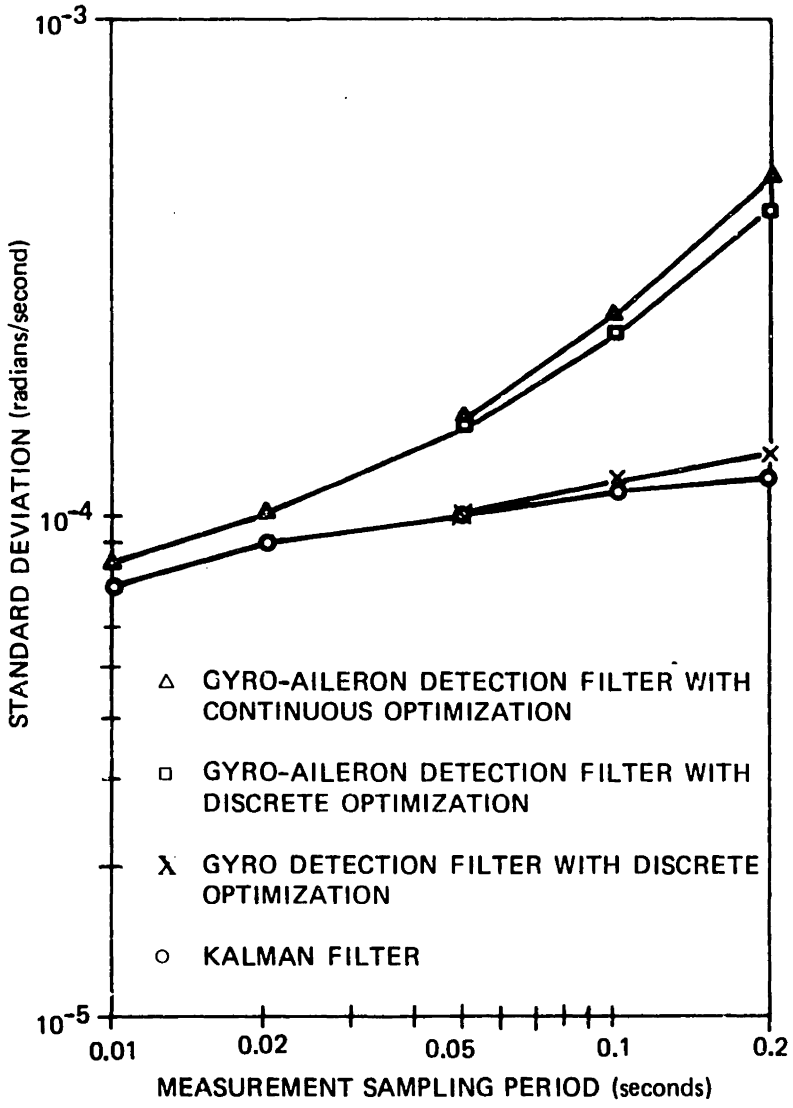


Figure 6.11b Roll rate estimation accuracies for gyro-aileron failure detection system, gyro failure detection system, and Kalman filter as a function of measurement sampling period. Measurement variance and aerodynamic moment variances are those of Table 6.3

of $\hat{D}_{\Delta\Gamma a}$ and $\hat{D}_{\Delta\Gamma r}$ in (6.4.8). The large difference in the roll rate estimates generated by the two failure detection systems is attributable to the fact that the $\Delta t\alpha_i \hat{\bar{\epsilon}}_i'(k\Delta t)$ and $\Delta t\alpha_o \hat{\bar{\epsilon}}_o'(k\Delta t)$ terms add directly into the $\dot{\varphi}_{mr}(k\Delta t)$ update portion of (6.4.4) and the magnitude of their gains is disproportionate to their state estimation usefulness for large measurement sampling periods. This difficulty is a general shortcoming of sampled-data failure detection systems. Because it is not required to detect aileron failures, the gyro failure detection system does not place such heavy emphasis on the noisy aileron deflection residuals. On the other hand, the gyro failure detection system, like the Kalman filter, cannot generate useful roll rate estimates in the event of an aileron failure.

Figure 6.12 indicates that the relative state estimation performance of the gyro-aileron failure detection system in the unfailed system is improved viz-a-viz the Kalman filter if the power of the random aerodynamic moment increases. The measurement uncertainties used in the comparison are those of Table 6.3, but the power of the random aerodynamic moment is a factor of 10 greater. The roll angle and roll rate estimates from both the Kalman filter and the gyro-aileron failure detection system are worse than in the previous case; however, with the detection filter designed using the discrete optimization procedure, the gyro-aileron failure detection system estimates are less affected than the Kalman filter estimates. An increase in the random aerodynamic moment power or a decrease in the gyro-aileron deflection sensor variances has the effect of reducing the importance of the $\Delta t\alpha_i \hat{\bar{\epsilon}}_i'(k\Delta t)$ and $\Delta t\alpha_o \hat{\bar{\epsilon}}_o'(k\Delta t)$ terms in (6.4.4) and thereby improving the relative value of the gyro-aileron failure detection system as a state estimator.

Figure 6.12 also indicates that the discrete and continuous optimization procedures can yield detection filters which produce state estimates with significantly different accuracies. This single result cannot be generalized to conclude that the discrete optimization procedure is always superior. For large measurement sampling periods, both algorithms are approximate and it is not possible to predict which will yield the better detection filter. For certain applications, both algorithms may be unac-

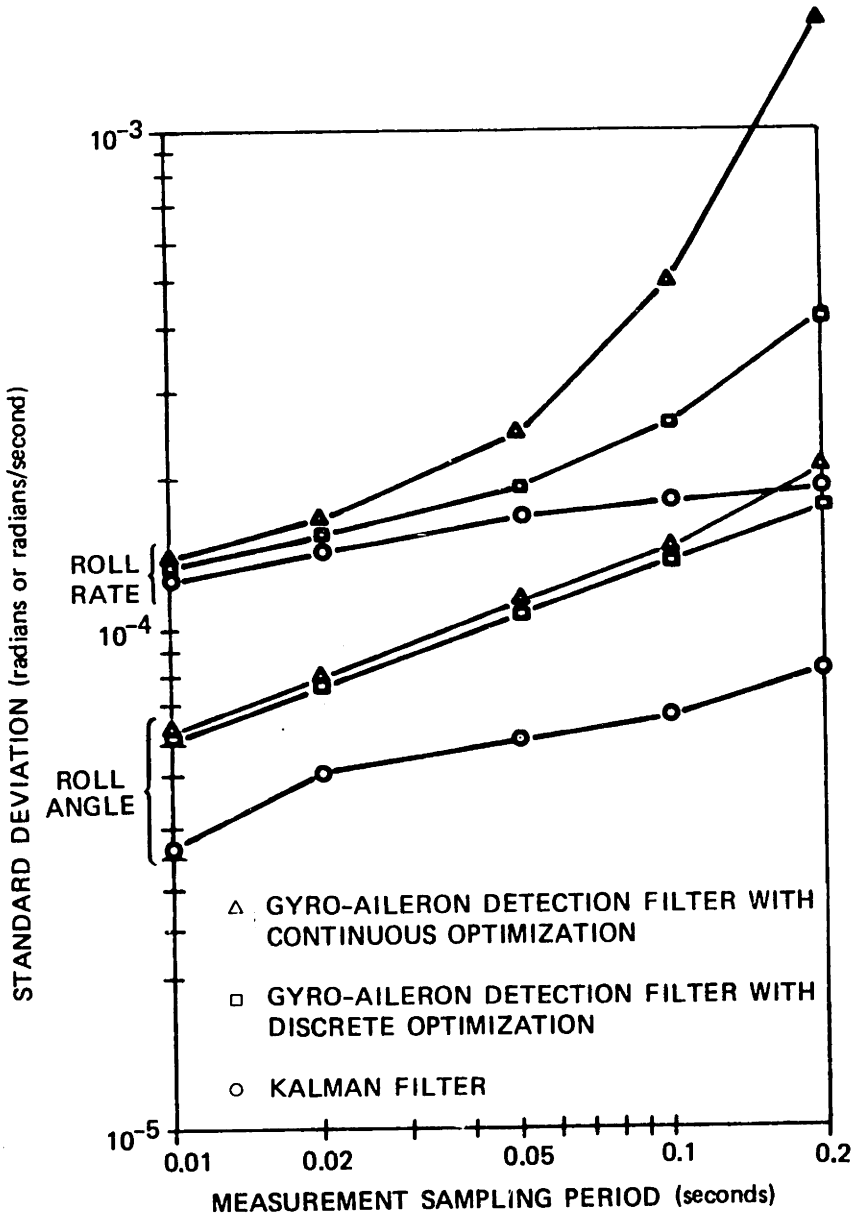


Figure 6.12 Roll angle and roll rate estimation accuracies for gyro-aileron failure detection system and Kalman filter as a function of measurement sampling period. Measurement variances are those of Table 6.3; W_{β} for aerodynamic moment is 10^{-6} $(rad/sec^2)^2$

ceptable and it may be necessary to resort to a trial and error procedure to choose $\hat{D}_{\Delta\Psi K}$.

6.4.3 Sensitivity Analysis

Section 6.4.1 demonstrated the ability of the T_{mg} and T_{ms1} detection systems to perform nearly instantaneous failure detection and identification when the failure significantly alters the short-term dynamics of the autopilot-aircraft system. In conjunction with Section 6.4.2, it also demonstrated that the failure detection systems can be used to generate state estimates which are partially insensitive to system failures, e.g. the roll angle and roll rate estimates produced by the gyro-aileron failure detection system are not affected by aileron failures. This subsection is concerned with circumstances in which a failure detection system using the threshold detection law may not yield satisfactory failure detection information. The specific areas discussed are:

i) an inability to constrain a failure-generated output error to the appropriate detection channel,
ii) an inability to discriminate between failure-generated output errors and those resulting from modeling errors, and
iii) a difficulty in detecting failure modes which do not significantly alter the short-term dynamics of the system.

In each instance modifications to the failure detection system or the detection law are suggested which may achieve the desired performance. Particular attention is given to the effect of choosing a detection filter which is not an optimal detection filter-state estimator.

It was mentioned in Section 5.2 that failure detection theory is not analytically precise for sampled-data systems. $\hat{D}_{\Delta K}$ as defined in (5.2.20) and (5.2.21) can only constrain the output error generated by a failure to be approximately unidirectional. In the context of the T_{mg} detection system, the residual generated by a rate gyro failure cannot be constrained completely to the rate gyro detection channel, but must have small projections into the attitude gyro detection channel, etc. The detection system outputs in Figure 6.10 demonstrated that when $\hat{\bar{\epsilon}}_r(k\Delta t)$ undergoes large changes from one sampling time to the next, the magnitudes

of the projections can be sufficiently great to create extraneous threshold crossings in detection channels in which they should not appear. In that simulation, the threshold crossings were predictable because they were a consequence of the rate gyro repair. The major significance of the extraneous signals in Figure 6.10 is that they indicate that the failure detection systems may not provide reliable failure detection information immediately after a system repair has been made.

Figure 6.13 reveals that the approximate nature of failure detection theory for sampled-data systems may have a more practical importance than suggested by Figure 6.10. The only command in the simulation in Figure 6.13 is a 2° roll command at 1.6 seconds. The failure is the same rate gyro failure as in Figure 6.10. The difference is that the detection law used for Figure 6.13 requires a waiting period of 2.9 seconds after a threshold crossing before identifying the failure (only the T_{mg} detection system output is presented). The measurement histories in Figure 6.13a show that the longer waiting period allows an unstable oscillation to build up in the roll response, but that an identification and correction of the rate gyro failure at 4.6 seconds (2.9 seconds after the first threshold crossing appears) permits the system to recover. The T_{mg} detection system output in Figure 6.13b indicates that the detection law developed in Section 6.3 might have difficulty correctly identifying the failure after the longer waiting period. A threshold crossing appears in the rate gyro detection channel at 1.7 seconds, but a threshold crossing also appears in the attitude gyro detection channel at 3.0 seconds. Except for the transients which occur in response to the rate gyro replacement, the residual is almost completely constrained to the rate gyro detection channel, but a comparative evaluation must be made in order to properly interpret the T_{mg} detection system output. The threshold crossing detection law as developed in Section 6.3 does not make such an evaluation.

The possibility of extraneous threshold crossings is a function of four parameters:

Figure 6.13a Sensor outputs for a simulation with a 2° roll command at 1.6 seconds. The rate gyro failed at .8 seconds and was repaired at 4.6 seconds.

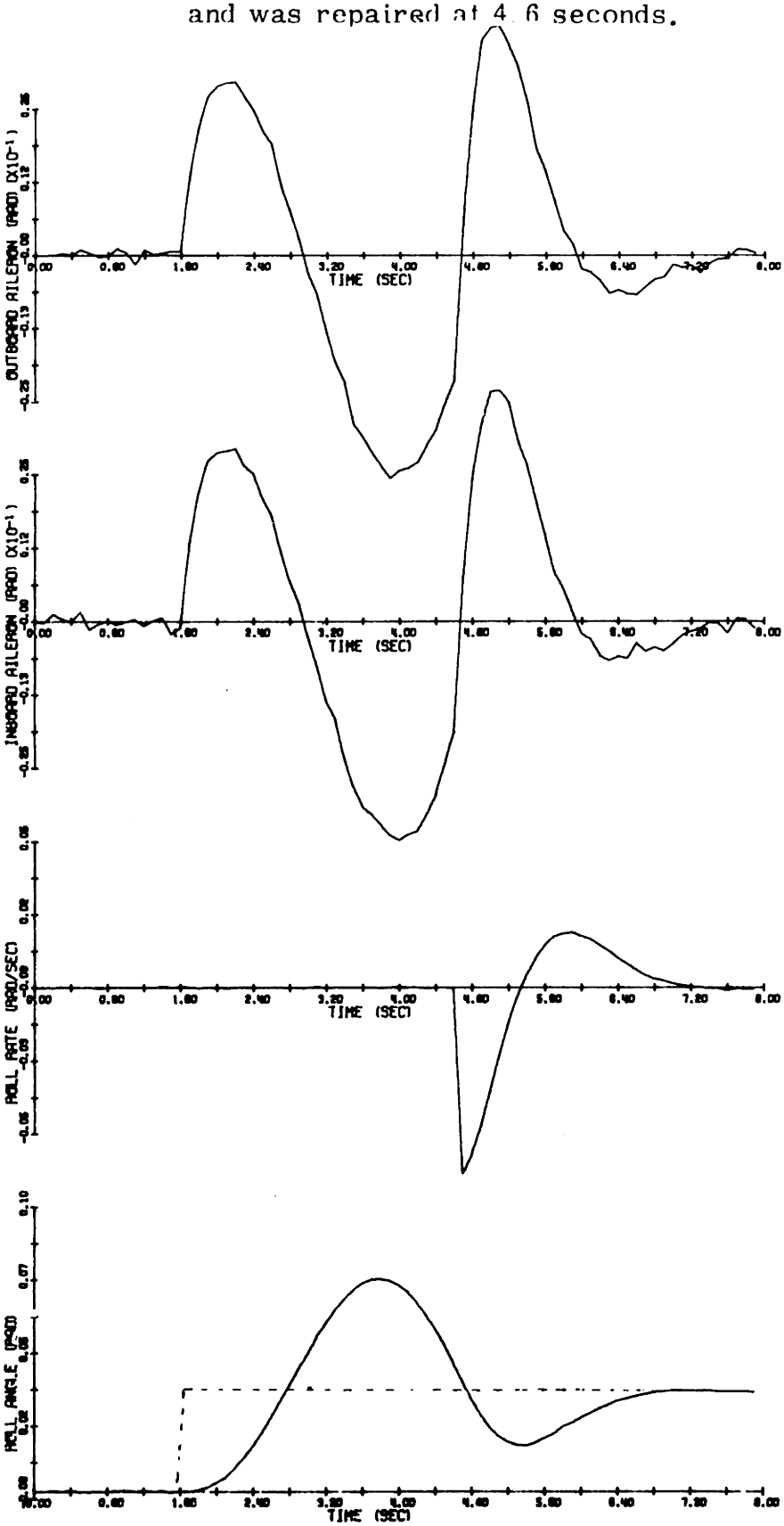
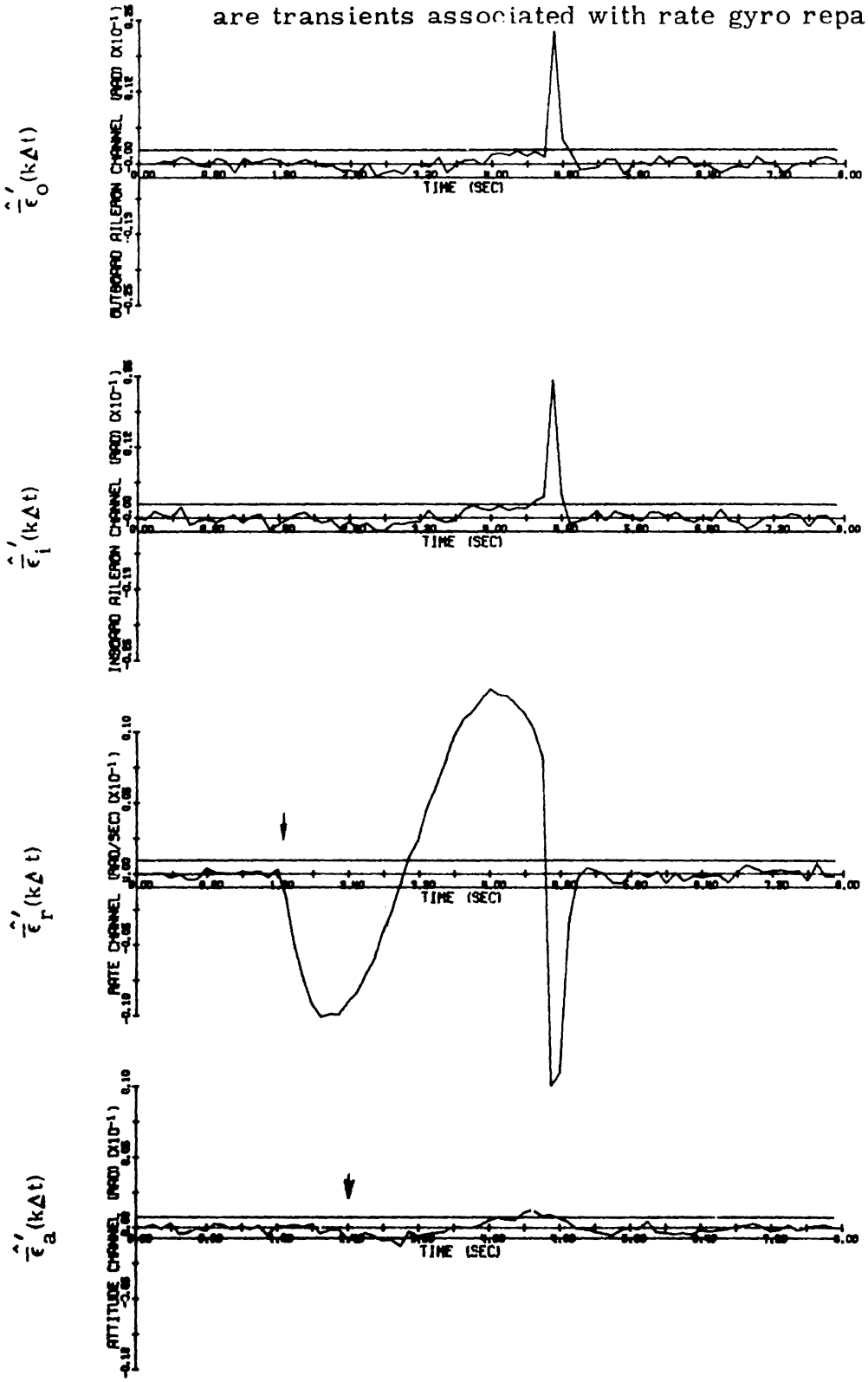


Figure 6.13b T_{mg} detection system output for simulation in Figure 6.13a. Arrows indicate time of threshold crossings. $\hat{\bar{\epsilon}}_i'(k\Delta t)$ and $\hat{\bar{\epsilon}}_o'(k\Delta t)$ threshold crossings are transients associated with rate gyro repair.



i) the size and variation of the residual magnitude,

ii) the threshold height,

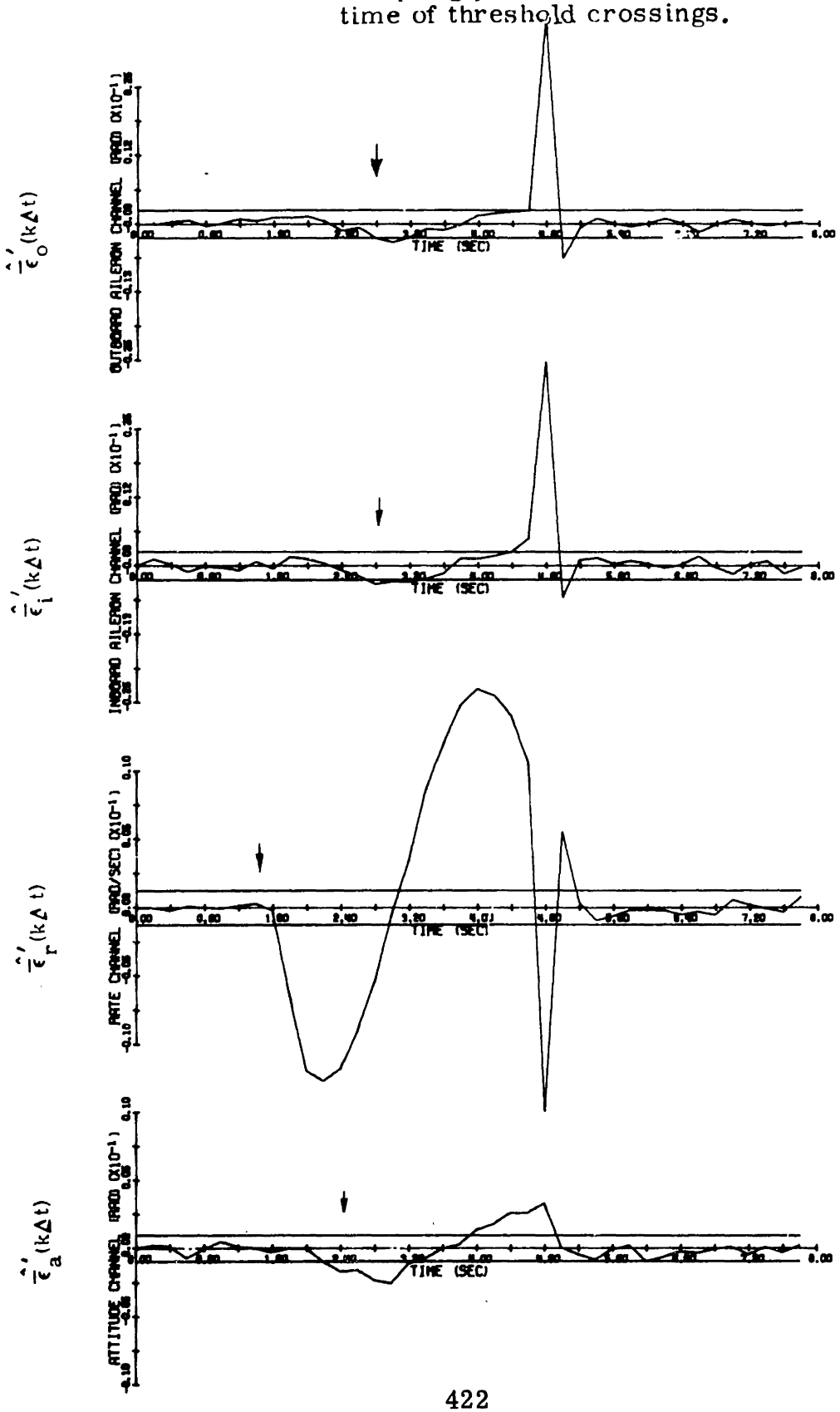
iii) the measurement sampling interval, and

iv) the noise statistics.

The noise statistics are accounted for by setting the thresholds at or above the 3σ noise level. If the thresholds were higher in the T_{mg} detection system, the extraneous threshold crossing in Figure 6.13b would not have appeared. The stochastic model and the choice of detection filters define the 3σ lower limit for the threshold, but the thresholds can be placed above the 3σ limit. On the other hand, if the roll command issued at 1.6 seconds were 20° instead of 2°, the residual in the rate gyro detection channel would be an order of magnitude greater and the extraneous threshold crossing in the attitude gyro detection channel would quite likely cross a 5σ or 6σ threshold. The extraneous threshold crossing would also be worse if the measurement sampling period were longer. The T_{mg} detection system output is shown in Figure 6.14 for the same simulation as Figure 6.13a except that the failure detection system is based upon a measurement sampling period of .2 seconds. Comparison of Figure 6.13b with Figure 6.14 verifies that the second failure detection system is considerably less effective in constraining the residual to the proper detection channel.

The most effective means of preventing extraneous threshold crossings is to have a short measurement sampling period. Beyond that, a short waiting period between the first threshold crossing and the failure detection decision is useful, but it is necessary to insure that the waiting period is sufficiently long to allow accurate isolation of the residual. After these possibilities have been tried, and if raising the thresholds is not effective, it is necessary to consider more sophisticated algorithms for interpreting the failure detection system output. One possibility is a comparison of the errors in each detection channel based upon peak magnitudes, average power, or some other reasonable criterion in order to determine the detection channel to which the output error is primarily constrained. A more general statement of the necessary conditions for isolating a failure is made in Section 5.3.

Figure 6.14 T_{mg} detection system output for simulation in Figure 6.13a except that the measurement sampling period is .2 seconds. Arrows indicate time of threshold crossings.



All the analysis of this chapter has been made under the assumption that the reference model defined in (6.1.5) is exact. The approximations made in ignoring the inner loop of the autopilot mean that (6.1.5) is not exact. Although the failure detection systems could have been designed for the full system dynamics, the fact that they were not affords an opportunity to study their performance when the system is inaccurately modeled. Figure 6.15a presents the sensor outputs for a simulation in which the full fourth-order aircraft dynamics and both the inner and middle loops of the autopilot were used. The simulation is analogous to that presented in Figure 6.5a. There are no failures and the only command is a 2° roll command at 1.6 seconds. Comparison with Figure 6.5a indicates that the response time has increased from 1.6 seconds to 3.5 seconds. The gains k_a and k_r could be modified to improve the response time, but that is not important here.

The T_{mg} and T_{ms1} detection system outputs corresponding to Figure 6.15a are presented in Figures 6.15b and 6.15c respectively. The single threshold crossing in the roll dynamics detection channel of the T_{ms1} detection system indicates that the roll mode dynamics have changed, or more precisely, that the roll response of the aircraft to an aileron deflection differs from the roll response defined in (6.1.5). Reference to Table 6.2 indicates that the threshold crossings in the T_{mg} detection channel are consistent with this observation. Unfortunately, this information is not useful since it is known a priori. Of more importance is the possibility that the threshold crossings associated with the modeling error might prevent the detection law from correctly identifying a failure when it occurs. In essence, the problem is one of the detection law's ability to detect simultaneous failures when a roll dynamics change is one of the failures. Reference to Table 6.2 indicates that a simultaneous roll dynamics change and rate gyro failure results in threshold crossings in the $\hat{\epsilon}_a'(k\Delta t)$ and $\hat{\epsilon}_r'(k\Delta t)$ detection channels of the T_{mg} detection system and the $\hat{\epsilon}_{ar}'(k\Delta t)$ and $\hat{\epsilon}_\beta'(k\Delta t)$ detection channels of the T_{ms1} detection system. The table also shows that a simultaneous

Figure 6.15a Sensor outputs for a simulation in which the autopilot aircraft system is described by the full lateral dynamics equations and the autopilot of Figure 6.2. A roll command of 2° occurs at 1.6 seconds. There are no failures.

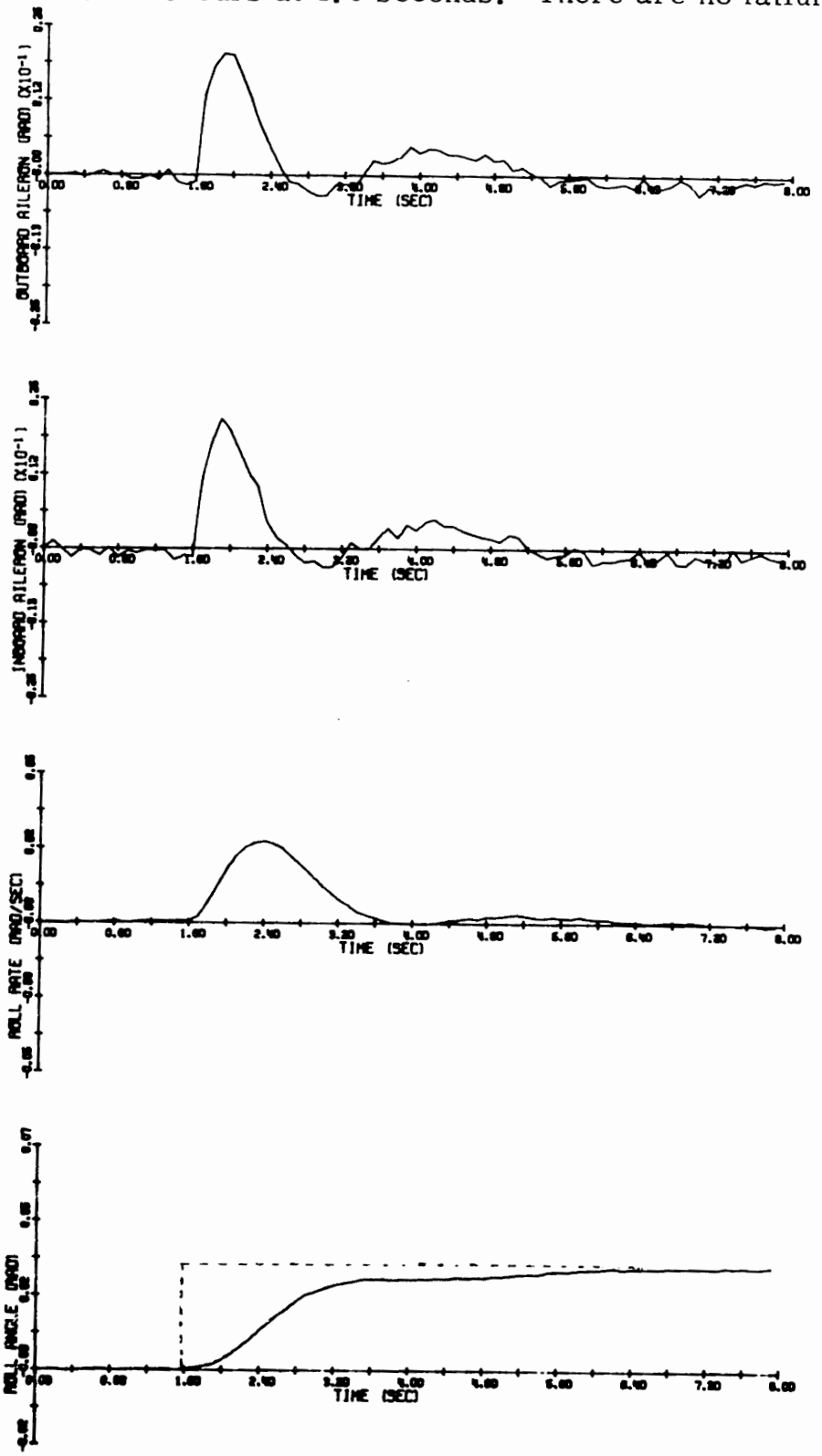


Figure 6.15b T_{mg} detection system output for simulation in Figure 6.15a. Arrows indicate time of threshold crossings.

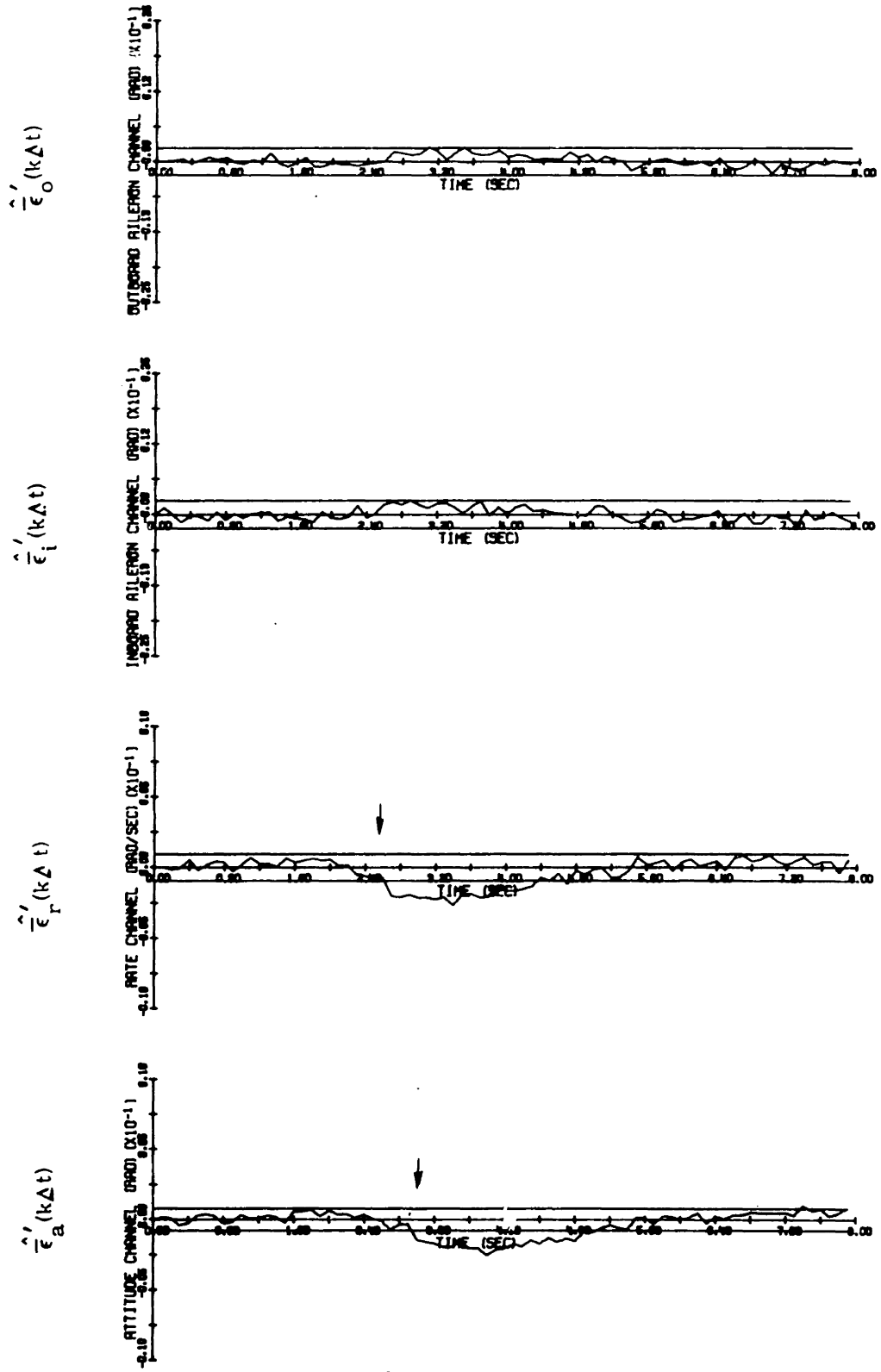
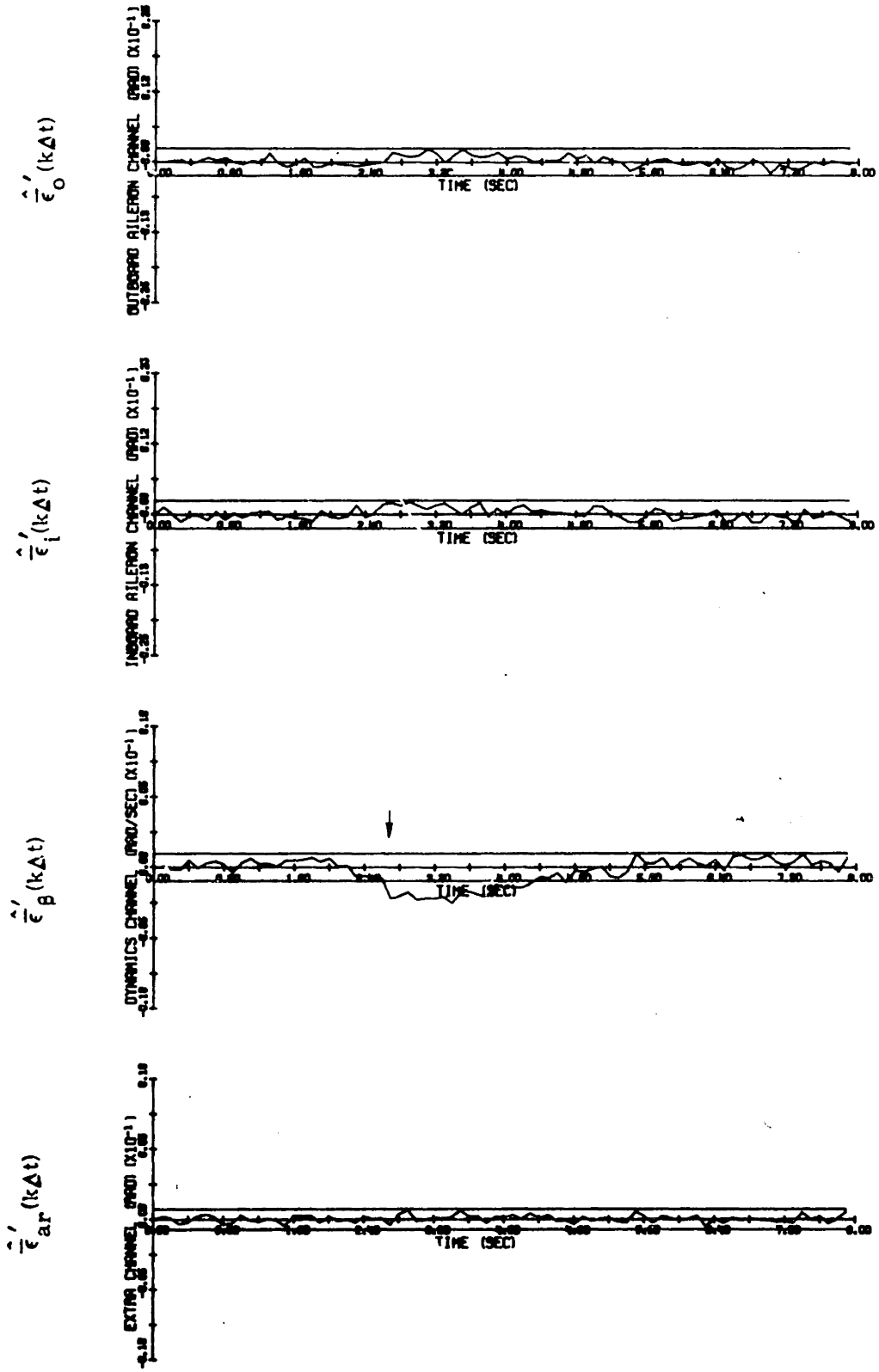


Figure 6.15c T_{mgl} detection system output for simulation in Figure 6.15a. Arrow indicates time of threshold crossing.



roll dynamics change and attitude gyro failure generate the same pattern of threshold crossings. It follows that the detection law has lost its ability to discriminate between a rate gyro failure and an attitude gyro failure unless the threshold crossings due to the modeling error can be accomodated in some way.

The most effective means of coping with the unwanted threshold crossings is to eliminate the modeling error. An alternative is to use some special property of the output error generated by the modeling error to distinguish it from the output error generated by a failure. In this instance, the output error due to the simplified dynamics has a 6 second period. In principle, the output errors $\hat{\bar{\epsilon}}_g'(k\Delta t)$ and $\hat{\bar{\epsilon}}_s'(k\Delta t)$ denoted in Figure 6.4 can be passed through a notch filter to remove this component after they are transformed into the detection channels. The delay introduced by the filter increases the length of time between a failure and its detection, but it allows the threshold crossing detection law to be used.

An alternative similar to the notch filter is to choose the $\hat{D}_{\Delta\Psi K}$ segment of $\hat{D}_{\Delta K}$ to attenuate the Dutch roll components of the output errors rather than to optimize the detection filter as a state estimator. With reference to the continuous approximations in (6.2.53a) and (6.2.53r), the minimal polynomials $\Psi_a(\cdot)$ and $\Psi_r(\cdot)$ are the denominators for the transfer functions:

$$\frac{1}{\Psi_a(s)} = \left| sE_2 - \hat{A}'_{aa} + d_{\Psi Ka}\hat{c}_a \right|^{-1}$$

$$\frac{1}{\Psi_r(s)} = \left| sE_2 - \hat{A}'_{rr} + d_{\Psi Kr}\hat{c}_r \right|^{-1} \qquad (6.4.10)$$

$\Psi_a(s)$ and $\Psi_r(s)$ are second-order polynomials which can be freely chosen. To some extent, they can be chosen to attenuate the error equation frequency responses to the Dutch roll frequency. Because they are of such low order, however, they are not as effective as the notch filter.

The final example in this subsection concerns the ability of the failure detection system to detect a degradation in a system component which does not significantly affect the short-term dynamics of the system. Figure 6.16 presents a computer simulation in which a .002 radians/second rate gyro bias appears at 1.6 seconds and a 2° roll command is issued at 4.0 seconds. When the bias appears, the autopilot responds by developing a roll angle offset. Other than the offset roll angle, the dynamic response of the system is not affected, but the long-term effect of the offset roll angle is to cause the aircraft to drift from its nominal flight path. The T_{mg} detection system output in Figure 6.16b indicates that the bias causes a threshold crossing in the rate gyro channel and can theoretically be detected, but the threshold crossing occurs only at one sampling time and its magnitude is only 5σ.

Unlike the majority of the failures studied in this section, the magnitude of the threshold crossing for the bias is neither state-dependent nor control dependent. If the failure detection system were to be designed to detect gyro biases, it is reasonable to tailor the detection filter to enhance the detectability of biases. The result is presented graphically in Figure 6.17. Figure 6.17 is the T_{mg} detection system output corresponding to the simulation presented in Figure 6.16a, but with two changes. The first change is to replace the gyro-aileron failure detection system with the gyro failure detection system. As shown in Figure 6.11b, this increases the accuracy of the roll rate estimate and therefore narrows the thresholds. The second change is that $d_{\Delta\Psi Kr}$ in (6.4.9) is chosen to weight the residual more lightly than an optimal detection filter-state estimator does. The result is to give the failure detection system a slower time response and cause the threshold crossing generated by a gyro bias to persist longer. Figure 6.17 indicates that the bias generates threshold crossings at five sampling times and that the $\hat{\bar{\epsilon}}_r'(k\Delta t)$ detection channel exhibits a definitive bias in the interval between 1.6 seconds and 5.2 seconds. A detection law which used a bias test on the $\hat{\bar{\epsilon}}_r'(k\Delta t)$ detection channel rather than the threshold crossing test could presumably detect the $\hat{\bar{\epsilon}}_r'(k\Delta t)$ bias and associate it with the rate gyro bias.

Figure 6.16a Sensor outputs for a simulation with a rate gyro bias of .002 radians/second at 1.6 seconds and a 2° roll command at 4.0 seconds.

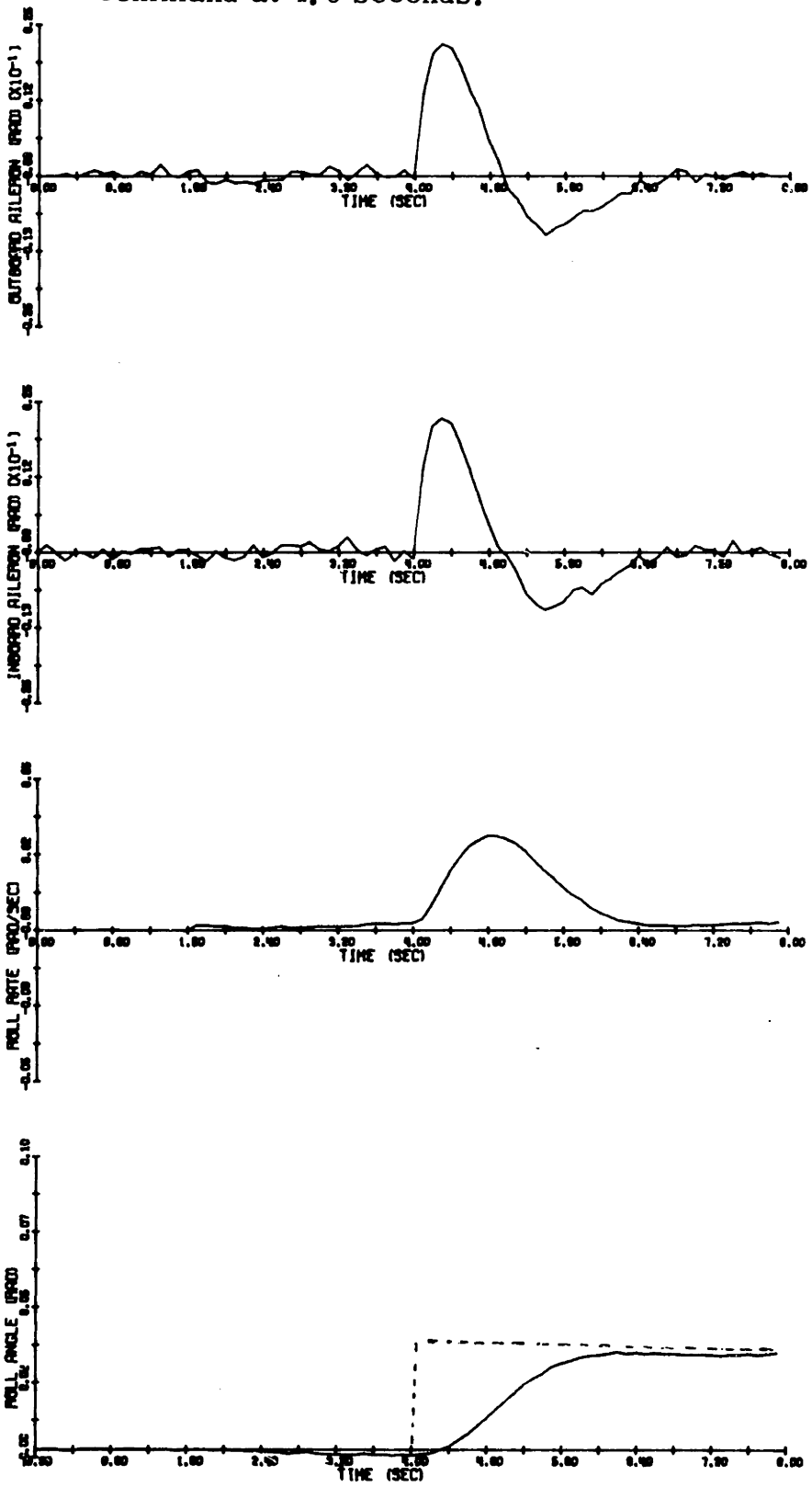


Figure 6.16b T_{mg} detection system output for simulation in Figure 6.16a. Arrow indicates time of threshold crossing.

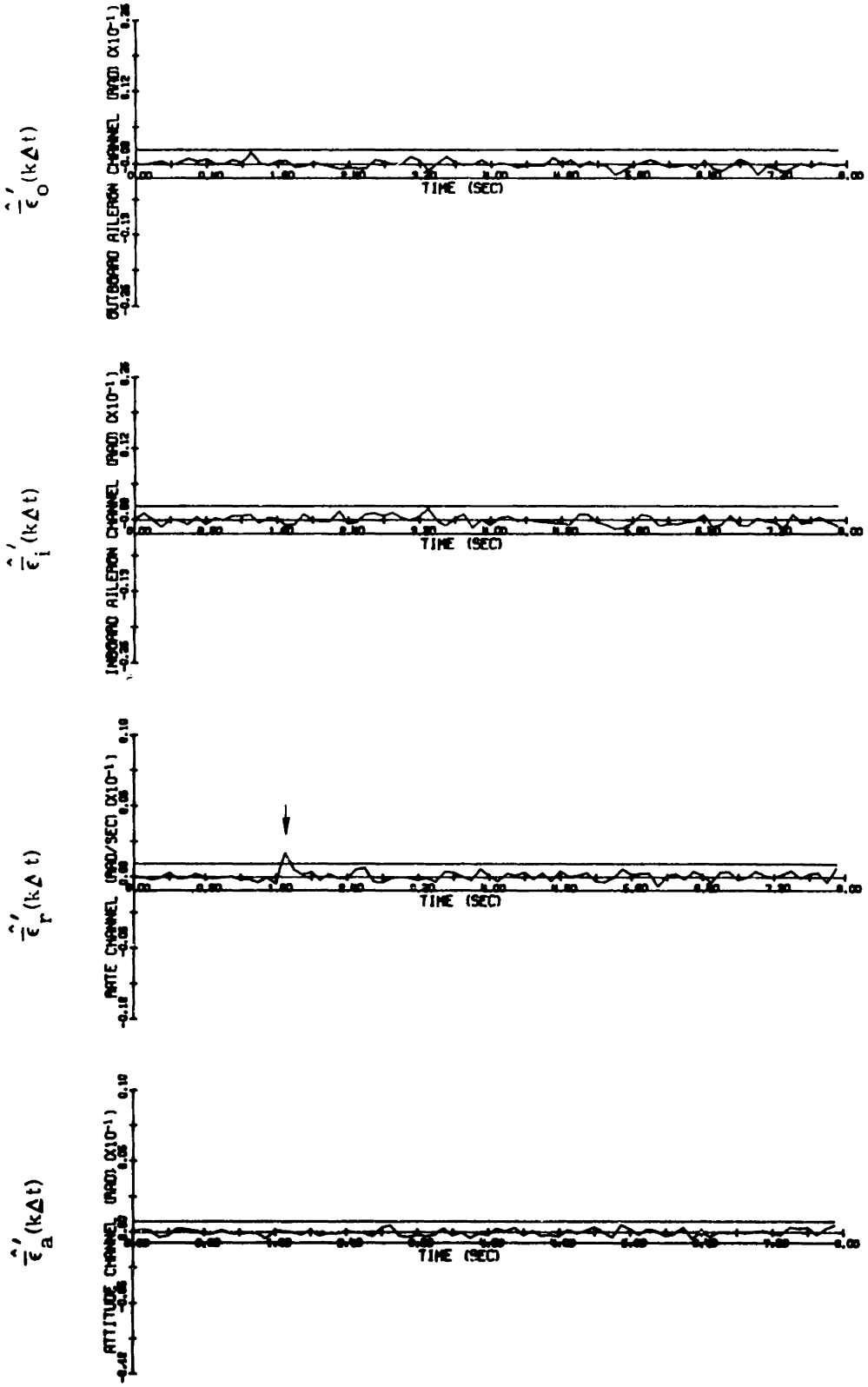
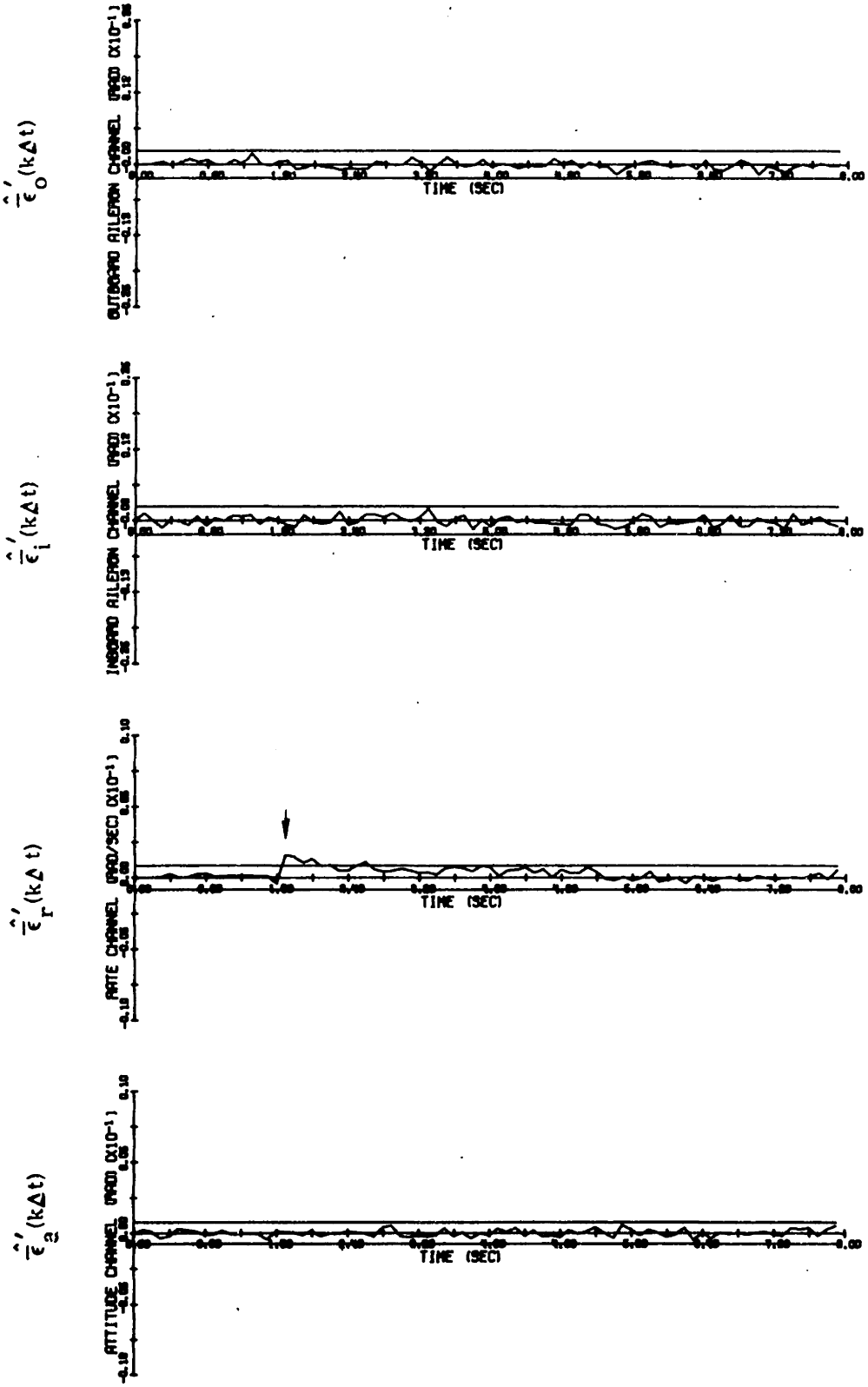


Figure 6.17 T_{mg} detection system residuals for Figure 6.16a. Arrow indicates time of threshold crossing



CHAPTER 7

SUMMARY AND RECOMMENDATIONS

7.1 Summary

The reference model approach used in this thesis provides a general capability for on-line detection of failures in complex dynamic systems. The theory is developed specifically for continuous, deterministic, time invariant linear systems; but it is equally valid for discrete systems. The addition of a simple detection law for interpreting the failure detection system output permits the theory to be extended to stochastic systems. A procedure for applying the theory to sampled-data systems has also been devised, but it is based upon a discrete-to-continuous approximation and lacks the mathematical preciseness of the previous results.

The basic analytical results of failure detection theory were developed in Chapter 3. These results were converted to a design algorithm in Chapter 4. Chapter 4 also presented a general failure model for linear systems (Section 4.1) and the transformation matrix to a canonical form for the reference model which greatly simplifies the design algorithm (Section 4.4). Extension of the theory to stochastic systems was treated in Chapter 5--of particular importance was the discussion of detection laws in Section 5.3. The simulation results presented in Chapter 6 verify that failure detection systems based upon the theory presented in this thesis can provide effective real-time detection and identification of failures in stochastic systems.

A reference model consists of a model (A, B, C) of the nominal system dynamics and a feedback matrix D which uses the output error (the difference between the system and model outputs) to update the state vector estimate of the model. The general failure model developed in Section 4.1 expresses the failed system in terms of the original model (A, B, C) and a set of additive vector terms which drive the state and the output equations. The usefulness of

the theory developed in this thesis is predicated on the fact that the model for a component failure frequently requires only one additive term. All failures of a single control element or a single sensor can be modeled with one such term. The additive term appears in the state equation for control element failures and in the output equation for sensor failures.

If the failure model for a particular system component is similar to the failure model for a control element, and if the system is observable, the algorithm presented in Section 4.2 can be used to chose D such that the output error generated by the failure is constrained to a single direction in the output space. Furthermore, that direction is the same for every possible failure mode for the component. With the single exception noted below the unidirectional output error is sufficient information to correctly identify the failure. This ability to identify failures independently of the particular failure mode is unique to failure detection theory. A D which yields a unidirectional output error for a given failure is referred to as a detection filter for that failure.

If more than one component is associated with the same output direction it is not possible to determine which of the components has failed on the basis of the output error direction alone. Knowledge of the possible failure modes for each component will generally permit a unique failure identification, but in certain cases no failure detection algorithm can remove the ambiguity. Different components are associated with the same output direction only if their individual failures have identical effects on the system output. This situation can arise for a system with redundant (or dynamically similar) components if the outputs of these components are not measured, but it does not arise if the outputs are measured.

Component failures which cannot be modeled similarly to control element failures cannot in general be constrained to generate unidirectional output errors. The principal examples of such failures are sensor failures. In most instances the output error generated by a sensor failure can only be constrained to a two-dimensional

plane. Although this implies that the output error generated by a sensor failure is more difficult to detect and identify than that generated by a control element failure, it also implies that the reference model approach offers a capability for distinguishing between the failure of a control element and the failure of the sensor which measures that control element's output. This capability is not present in many failure detection algorithms.

A single failure detection system can detect and identify failures in several different components. The precise number depends upon the number of sensors in the system, the dynamic relationships between the various components, and the degree to which the available design options are utilized. If the system has m sensors, a simple algorithm can be used to determine a set of m components which have "output separable" failures. It is possible to choose D such that a failure of any of the components generates a unidirectional output error along the appropriate direction. A number of design options are available at this point, but D can be constrained to be a detection filter for the m component failures and still retain sufficient freedom in the choice of D to arbitrarily choose the closed loop eigenvalues of the reference model. One of the design options may increase the number of component failures for which D is the detection filter. If a single failure detection system cannot detect and identify failures in all system components, additional failure detection systems can be defined.

The structure of a failure detection system differs from that of a Kalman filter only in that D is not chosen to minimize the variance of the state estimates. Nevertheless, a failure detection system is a state estimator. The algorithms developed in Sections 5.1 and 5.2 optimize the failure detection system as a state estimator subject to the constraint that D must be a detection filter for a specified set of component failures. The simulation results presented in Chapter 6 verify that a failure detection system chosen to detect a suitable set of failures can be a good state estimator.

An important feature of the failure detection theory developed in this thesis is that the algorithms employed in selecting D are algebraic. Unlike Kalman filter theory where computation of the feedback matrix requires the solution to a matrix Riccati equation, the detection filter D is the solution of a set of linear, algebraic equations. As a consequence, the total computation required to find D is considerably less than that required to find a Kalman filter gain. In particular, if the reference model is transformed to base normal canonical form (Section 4.4), D can be chosen by inspection.

The theory presented in this thesis is not entirely new. The major theorems in Sections 3.2, 3.3.1, and 3.3.3 were first derived by Beard (1971) using matrix algebra techniques. The significance of these sections is that they provide a geometric derivation of failure detection theory which emphasizes the basic principles of controllability and observability upon which state space theory is founded. The remainder of the derivations in Chapter 3, the base normal canonical form in Chapter 4, and the optimization techniques in Chapter 5 are new results.

7.2 Recommendations for Further Study

In the course of this study, a number of areas for further research have been identified. Those which appear to offer the greatest potential yield are outlined in the following paragraphs.

The theorems derived in Chapter 3 provide the major results needed to design failure detection systems for linear, time invariant, deterministic systems. Although time-varying systems were not specifically analyzed, there are classes of time-varying systems for which the majority of the analytical results are applicable. In particular, it appears that, with some minor restrictions, the theory is applicable to the class of systems in which the time variation is restricted to a small number of parameters. The possibility of extending failure detection theory to large classes of time-varying systems deserves attention. In this regard, a group-theoretic generalization of the basic concepts discussed in Chapter 3 might provide valuable insights.

The general concept of output stationarity was introduced in this thesis and analytical results were developed which indicate that output stationarity can be used to significantly increase the number of failures which can be detected by a single failure detection system. Necessary and sufficient conditions for output stationarity were developed; but efforts to incorporate the concept into an efficient computational algorithm were only partially successful. The algorithm developed in Sections 3.4 and 4.2.3 performs satisfactorily if a single event vector is to be made output stationary, but it is cumbersome in dealing with mutual output stationarity. Further study of output stationarity should yield a more computationally satisfactory algorithm.

The treatment of stochastic systems presented in Chapter 5 is valuable in that it defines algorithms by which failure detection theory can be applied to stochastic systems, but it does not directly address an important philosophical question: The theory permits complete freedom in assigning the closed loop eigenvalues of the reference model, but what values for the eigenvalues should be selected? In certain applications, the optimal detection filter-state estimator developed in Section 5.1 may be the best answer, but the question deserves a more thorough examination.

The problem of detecting the presence of a failure-generated error signal in a stochastic environment should also be investigated further. Failure detection theory has an important property in that it decouples a multidimensional output error analysis problem into a set of scalar output error analysis problems. As discussed in Section 5.4, this feature could potentially be used in conjunction with existing hypothesis testing techniques to yield reliable detection and identification of "soft" component failures with a reasonable requirement for real-time computation. Utilization of hypothesis testing techniques which yield minimum detection time, or maximum probability of detection, etc., could increase the range of stochastic problems for which the theory is applicable.

The most important single recommendation is that serious consideration should be given to using the algorithm developed in Chapters 4 and 5 to design a failure detection system for a real system. The simulation results in Chapter 6 have verified the important elements of the theory and have demonstrated that a failure detection system based on the theory can provide reliable failure detection and identification performance. For a properly chosen application, the algorithms should yield a reliable failure detection system with the ability to detect and identify component failures independent of the manner in which the components fail.

APPENDIX A

NUMERICAL ALGORITHMS

The algorithms of this appendix are those required to implement a failure detection system. As indicated in Chapter 3, they are not unique and they are not necessarily the most direct solution to the design problem or the most computationally efficient. Their primary attribute is that a single artifice -- the orthogonal reduction procedure -- can be used to analyze the basic characteristics of a system (A, C) which determine its failure detection properties. The algorithms presented here are a variation upon those proposed by Beard. Alternate algorithms are discussed in the appropriate sections of Chapter 3 and Chapter 4.

A. 1 Orthogonal Reduction Procedure

Consider an $n' \times n$ matrix:

$$V = \begin{bmatrix} v_1^T \\ \vdots \\ v_{n'}^T \end{bmatrix} \qquad \text{(A. 1)}$$

where the v_j, $j = 1 \ldots n'$, are arbitrary n-vectors. The orthogonal reduction procedure is an iterative process which generates an $n \times n$ symmetric, positive semi-definite matrix $\Omega^{(n')}$ whose range space coincides with $\eta(V)$. In the j^{th} iteration the j^{th} row of V is tested to determine if it is orthogonal to the range space of a symmetric matrix. If it is not, the range space of the matrix is reduced to insure orthogonality. The algorithm begins with an initial matrix $\Omega^{(1)}$ which is symmetric and positive-definite. The removal is done in such a manner that the $\Omega^{(j)}$, $j = 1 \ldots n'$, remain symmetric and positive semi-definite. At the completion of the procedure, the range space of $\Omega^{(n')}$ is the orthogonal complement of the row space of V, i.e., it is $\eta(V)$. The algorithm is a generalization of the Gram-Schmidt orthogonalization procedure.

The algorithm begins with any symmetric, positive-definite matrix $\Omega^{(1)}$. The iteration is a two step procedure in terms of the auxiliary vector w_j:

i) $$w_j = \Omega^{(j)} v_j \tag{A. 2}$$

w_j is the inner product of v_j with the row space of $\Omega^{(j)}$ -- which is also the range space of $\Omega^{(j)}$ since $\Omega^{(j)}$ is symmetric. If w_j is zero, v_j is orthogonal to the range space of $\Omega^{(j)}$ and no modification of $\Omega^{(j)}$ is required:

ii) $$\Omega^{(j+1)} = \Omega^{(j)} \tag{A. 3}$$

If w_j is not zero, $\eta(\Omega^{(j+1)})$ is enlarged to include v_j:

ii)′ $$\Omega^{(j+1)} = \Omega^{(j)} - \frac{w_j w_j^T}{w_j^T v_j} \tag{A. 4}$$

After a maximum of n' iterations, the row space of V is $\eta(\Omega^{(n')})$ and the range space of $\Omega^{(n')}$ is $\eta(V)$.

Each step of the reduction yields a symmetric, positive definite matrix $\Omega^{(j)}$. If $\Omega^{(1)}$ is positive definite, $\eta(\Omega^{(j)})$ is spanned by $v_1 \ldots v_{j-1}$. If $\Omega^{(1)}$ is positive semi-definite, $\eta(\Omega^{(j)})$ is spanned by the sum of $v_1 \ldots v_{j-1}$ and any basis of $\eta(\Omega^{(1)})$. This fact is useful in finding the detection generator for f from $\overline{\mathcal{R}}_f$. The number of reductions made is rkV if $\mathrm{rk}\,\Omega^{(1)} = n$, but will be less if $\eta(\Omega^{(1)})$ is chosen to contain part of the row space of V.

A. 2 A Detection Space and Its Detection Generator

If C′ and D_f are defined as in Lemma 3. 2. 6′ and (A, C) is observable, $\overline{\mathcal{R}}_f = \eta(M_D')$ where:

$$M_D' = \begin{bmatrix} C' \\ C'(A-D_fC) \\ \vdots \\ C'(A-D_fC)^{n-1} \end{bmatrix} \qquad (A.5)$$

A basis for $\bar{\mathcal{R}}_f$ can be determined as the range space of some matrix $\Omega^{(n)}$ by using the orthogonal reduction procedure on M_D'. The most straightforward manner in which to initiate the algorithm is to let $\Omega^{(1)}$ be the identity matrix. The matrix R_f can then be composed from the largest number of independent columns of $\Omega^{(n)}$. Obviously, $\text{rk}\,\Omega^{(n)} = \nu_f$.

A more efficient procedure can be developed to eliminate as much repetition in the computation as possible. If a set of event vectors $f_1 \ldots f_r$ is output separable, an investigation of their mutual detectability ultimately requires that $\bar{\mathcal{R}}_{\mathcal{J}}$, $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$, and $g_1 \ldots g_r$ all be determined. Further, it might be desirable to add a set of output stationary vectors $h_1 \ldots h_q$ in the design of a detection filter. This would require that the additional detection spaces $\bar{\mathcal{J}}_1 \ldots \bar{\mathcal{J}}_q$ and their generators be determined. The algorithm presented below is segmented into four steps and the circumstances under which the steps must be executed are discussed.

Step 1. Find $\eta(C)$

If detection spaces other than $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ are to be determined, their computation is more efficient if it is not necessary to initiate the orthogonal reduction procedure with a positive definite $\Omega^{(1)}$. A more acceptable approach is to start with a symmetric, positive semi-definite matrix Ω_C whose range space is $\eta(C)$. This can be accomplished by applying the orthogonal reduction procedure to C with $\Omega^{(1)}$ equal to the identity matrix:

$$\Omega^{(1)} = E_n \qquad (A.6)$$

Ω_C is then the terminating matrix from the algorithm:

$$\Omega_C = \Omega^{(m)} \qquad (A.7)$$

Since rkC = m, a total of m reductions must be made using ii)' of the orthogonal reduction procedure. It is only necessary to compute Ω_C once.

Step 2. Compute $\bar{\mathcal{R}}_{\mathcal{J}}$

If $\mathcal{J}$ is the subspace spanned by $f_1 \ldots f_r$, the mutual detectability of $f_1 \ldots f_r$ is governed by the detection space $\bar{\mathcal{R}}_{\mathcal{J}}$ for $\mathcal{J}$ as specified in Definition 3.3.3. $\bar{\mathcal{R}}_{\mathcal{J}}$ satisfies:

$$\bar{\mathcal{R}}_{\mathcal{J}} = \mathcal{R}_{\mathcal{J}} \oplus \mathcal{J}$$

where $\mathcal{R}_{\mathcal{J}} \subset \eta(C)$. Since $\mathcal{J}$ is known, $\bar{\mathcal{R}}_{\mathcal{J}}$ is known if $\mathcal{R}_{\mathcal{J}}$ can be found. If $D_{\mathcal{J}}$ and $C'_{\mathcal{J}}$ are defined as in Lemma 3.3.2 and M_D' is defined:

$$M_D' = \begin{bmatrix} C'_{\mathcal{J}} \\ C'_{\mathcal{J}}(A-D_{\mathcal{J}}C) \\ \vdots \\ C'_{\mathcal{J}}(A-D_{\mathcal{J}}C)^{n-1} \end{bmatrix} \qquad \text{(A.8)}$$

then $\bar{\mathcal{R}}_{\mathcal{J}}$ can be found by applying the orthogonal reduction procedure to M_D' and starting with the initial matrix:

$$\Omega^{(1)} = \Omega_C \qquad \text{(A.9)}$$

From Lemma 3.3.2 it is clear that $\eta(C) \subset \eta(C'_{\mathcal{J}})$ and therefore the row space of $C'_{\mathcal{J}}$ is contained in the row space of C. If the reduction procedure is applied to M_D' with the initial matrix (A.9), the first m rows of M_D' are contained in $\eta(\Omega^{(1)})$ and the reduction can start with the $(m+1)^{th}$ row of M_D'. A total of $n - m - \nu_{\mathcal{J}} + r$ reductions using ii)' are necessary. Let $\Omega_{\mathcal{J}}$ be the terminating matrix $\Omega^{(nm-m+1)}$:

$$\Omega_{\mathcal{J}} = \Omega^{(nm-m+1)} \qquad \text{(A.10)}$$

The matrix R_F can be composed from the maximum number of independent columns of $\Omega_{\mathcal{J}}$ and:

$$\bar{R}_F = [\, F \,\vdots\, R_F \,] \tag{A.11}$$

A basis for $\bar{\mathcal{R}}_{\mathcal{J}}$ is provided from the columns of $\bar{R}_F$. This computation need occur only once unless event vectors are added to the set $f_1 \ldots f_r$.

If an event vector f_k is removed from the set or if the state space is augmented, the algorithms of Theorem 3.3.3 (and Corollary 1) or Lemmas 3.3.10, 3.3.11, and Appendix A.3 can be used to determine $\bar{\mathcal{R}}_{\mathcal{J}k}$ or $\tilde{\bar{\mathcal{R}}}_{\mathcal{J}}$ respectively. Both algorithms require use of the orthogonal reduction procedure to find the null space of a matrix.

Step 3. Compute $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$ and $\bar{\mathscr{I}}_1 \ldots \bar{\mathscr{I}}_q$

The matrices $\mathcal{R}_1 \ldots \mathcal{R}_r$, $\mathscr{I}_1 \ldots \mathscr{I}_q \subset \eta(C)$ can be determined in a manner similar to Step 2 by an orthogonal reduction on M_D' as defined in (A.5) with C' and D_f computed for the appropriate element of $f_1 \ldots f_r$, $h_1 \ldots h_q$. In the case of $\mathcal{R}_1 \ldots \mathcal{R}_r$, the initial matrix for the reduction should be:

$$\Omega^{(1)} = \Omega_{\mathcal{J}} \tag{A.12}$$

A total of $\nu_{\mathcal{J}} - r - \nu_i + 1$ reductions using ii)' are necessary for each subspace and the terminating matrices are denoted by $\Omega_1 \ldots \Omega_r$:

$$\Omega_k = \Omega^{(nm-m+1)} \qquad k = 1 \ldots r \tag{A.13}$$

$R_1 \ldots R_r$ are then composed from the independent columns of the appropriate $\Omega_1 \ldots \Omega_r$ and $\bar{R}_1 \ldots \bar{R}_r$ are constructed by:

$$\bar{R}_k = [\, f_k \,\vdots\, R_k \,] \tag{A.14}$$

Since it may not be true that $\mathscr{I}_1 \ldots \mathscr{I}_q \subset \mathcal{R}_{\mathcal{J}}$, the reduction for them must be initiated with (A.9). A total of $n - m - \nu_{\mathscr{I}_i} + 1$ reductions are necessary, the terminating matrices are denoted by $\Omega_{\mathscr{I}_1} \ldots \Omega_{\mathscr{I}_q}$:

$$\Omega_{\nu i} = \Omega^{(nm-m+1)} \tag{A.15}$$

and $\bar{S}_1 \ldots \bar{S}_q$ are constructed similarly to (A.14).

If the state space is augmented, the new $\tilde{\bar{R}}_1 \ldots \tilde{\bar{R}}_r$ can be found by applying Lemmas 3.3.10 and 3.3.11. The latter requires a matrix $[\ \Delta\ |\ \beta\ ]$ to be determined where β spans the null space of the matrix M_{0k} as defined in Theorem 3.3.3, and Δ spans the row space of M_{0k}. If the orthogonal reduction procedure is applied to M_{0k} with the initial matrix:

$$\Omega^{(1)} = E_{\nu_k} \tag{A.16}$$

then Δ is composed of the non-zero auxiliary vectors w_j and β is composed of the independent columns of the terminating matrix. A similar operation is then required to determine Z_j as the set of vectors which span the null space of M_{0Lj} as defined in Lemma 3.3.11.

In general, the computation leading to (A.15) must be repeated for $(\tilde{A},\ \tilde{C})$ in order to determine the new $\tilde{\bar{S}}_1 \ldots \tilde{\bar{S}}_q$. For this reason, it may be desirable to postpone computing $\bar{S}_1 \ldots \bar{S}_q$ until the state space augmentation has been completed.

Step 4. Compute detection generators

Once the detection spaces are determined, a modified reduction procedure is needed to determine the detection generators $g_1 \ldots g_r$. This time the reduction is on the observability matrix M:

$$M = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{n-1} \end{bmatrix} \tag{A.17}$$

The generator is contained in the range space of the appropriate $\Omega_1 \ldots \Omega_r$ and the reduction should be initiated with that matrix.

$$\Omega^{(1)} = \Omega_k \qquad k = 1 \ldots r \tag{A.18}$$

Since $\mathcal{R}_1 \ldots \mathcal{R}_r$ are observable, each reduction procedure terminates on the zero matrix and for $\mathcal{R}_i$, $\nu_i - 1$ reductions using ii)' are needed. The last vector to be removed from the range space of $\Omega^{(j)}$ is a multiple of the detection generator. Since $\mathrm{rk}\,\Omega^{(1)} = \mathrm{rk}\,\mathcal{R}_i = \nu_i - 1$, if the last non-zero auxiliary w_j from (A. 2) is denoted by $w_{\nu_i - 1}$, then:

$$g_i = \alpha_i w_{\nu_i - 1} \tag{A. 19}$$

where α_i is an appropriate scale factor. The same algorithm is used to find the detection generators for $\mathcal{J}_1 \ldots \mathcal{J}_q$.

All the reduction procedures of this section can be simplified if special attention is paid to the manner in which M and $M_D{}'$ are constructed. Consider M as an example. A row of M can be immediately removed from the reduction procedure without applying (A. 2) if it is known to be a linear combination of the rows above it in M -- previous reductions will have modified $\Omega^{(j)}$ such that the row is contained in $\eta(\Omega^{(j)})$. If $c_k A^{\ell}$ is a linear combination of the rows above it, $c_k A^{\ell+i}$ must be also for all i and the reduction can be terminated for all subsequent rows of M in which c_k appears. Beard refers to the first zero auxiliary vector associated with c_k as the intermediate terminating point for that row of C.

If $\mathrm{rk}\,\Omega^{(1)} < n$, it is possible to show that the intermediate terminating point may occur even if $c_k A$ is not a linear combination of the rows above it. For the algorithms proposed in this appendix, that does not matter. If the auxiliary vector for $c_k A^{\ell}$ is zero, it is zero for all $c_k A^{\ell+i}$ and the reduction can be terminated for subsequent rows of M in which c_k appears. The same result holds for the reductions on $M_D{}'$. Once m intermediate terminating points have been encountered, the reduction process itself can be terminated. This occurs after at most $\mathrm{rk}\,\Omega^{(1)} + m$ rows of M (or $M_D{}'$) have been inspected.

After $\bar{R}_F$, $\bar{R}_1 \ldots \bar{R}_r$ have been determined, $f_1 \ldots f_r$ can be inspected for mutual detectability. Theorem 3. 3. 3 assures mutual detectability if and only if

$$\mathrm{rk}[\ \bar{R}_1 : \ldots : \bar{R}_r\ ] = \mathrm{rk}\,\bar{R}_F \tag{A.20}$$

Standard numerical analysis techniques can be utilized to determine whether or not (A.20) is an equality.

A.3 Excess Subspace

The excess subspace associated with $\bar{\mathcal{R}}_{\mathcal{J}}$ is not unique, but the excess subspace $\mathcal{R}_{0g}$ defined in Lemma 3.3.6 is. The algorithm defined in this section for $\mathcal{R}_{0g}$ is a variation upon the construction used in the proof of Lemma 3.3.6. It is also a variation on the algorithm proposed by Beard.

The procedure is to remove from $\mathcal{R}_{\mathcal{J}}$ of Step 2 the subspaces $\mathcal{R}_1 \ldots \mathcal{R}_r$ in such a manner that only $\mathcal{R}_{0g}$ remains. Assume the rows of the following matrix are known:

$$[\ (CF)^T CF\ ]^{-1}(CF)^T C = \begin{bmatrix} \bar{c}_1 \\ \vdots \\ \bar{c}_m \end{bmatrix} \tag{A.21}$$

Since the matrix is a portion of the expression for $C'_{\mathcal{J}}$, it is computed in finding $\mathcal{R}_{\mathcal{J}}$. Let the matrix M_0' be defined in terms of these rows:

$$M_0' = \begin{bmatrix} \bar{c}_1(A-D_{\mathcal{J}}C) \\ \vdots \\ \bar{c}_1(A-D_{\mathcal{J}}C)^{\nu_1-1} \\ \bar{c}_2(A-D_{\mathcal{J}}C) \\ \vdots \\ \bar{c}_m\,(A-D_{\mathcal{J}}C)^{\nu_m-1} \end{bmatrix} \tag{A.22}$$

If the orthogonal reduction procedure is applied to M_0' with the initial matrix $\Omega^{(1)}$ given by (A. 12), the terminating matrix can be designated by Ω_{0g}:

$$\Omega_{0g} = \Omega^{(\nu_J')} \tag{A.23}$$

The matrix R_{0g} can then be defined:

$$R_{0g} = [\, z_1 \cdots z_{\nu_0} \,] \tag{A.24}$$

where $z_1 \cdots z_{\nu_0}$ are the first ν_0 linearly independent columns of Ω_{0g}.

This result can be motivated by considering C_J' and M_0' in detail. It is easy to see that:

$$[\,(CF)^T CF\,]^{-1}(CF)^T C f_i = e_{mi} \tag{A.25}$$

The matrix is a transformation of the output space $\mathcal{S}^m$ such that, if $\bar{\mathcal{R}}_i$ is invariant, f_i produces output only along the i^{th} coordinate direction in the output space. The observability of $\bar{\mathcal{R}}_i$, or the manner in which some arbitrary excess space $\mathcal{R}_0$ maps into $\bar{\mathcal{R}}_i$, can be determined from consideration of only the i^{th} row of (A. 25). Specifically, if $\mathcal{R}_{0g}$ maps into $\bar{\mathcal{R}}_i$, $\mathcal{R}_{0g}$ must satisfy:

$$\begin{bmatrix} \bar{c}_i \\ \bar{c}_i(A-D_J C) \\ \vdots \\ \bar{c}_i(A-D_J C)^{\nu_i - 1} \end{bmatrix} R_{0g} = 0 \tag{A.26}$$

and:

$$\bar{c}_i(A-D_J C)^{\nu_i} R_{0g} \neq 0 \tag{A.27}$$

No vector in $\mathcal{R}_i$ satisfies (A. 26). Therefore, the reduction on M_0' removes $\mathcal{R}_1 \ldots \mathcal{R}_r$ from the range space of $\Omega_{\mathcal{J}}$ to yield Ω_{0g}. By construction Ω_{0g} satisfies (A. 26). Since $\mathcal{R}_{\mathcal{J}}$ is observable and $\bar{\mathcal{R}}_{\mathcal{J}}$ is invariant with respect to A-$D_{\mathcal{J}}$C, Ω_{0g} also must satisfy (A. 27) for some subset of i = 1 ... r. Because:

$$\mathrm{rk}\,\Omega_{0g} = \nu_0 = \mathrm{rk}\,\mathcal{R}_{0g} \tag{A. 28}$$

$\mathcal{R}_{0g}$ is uniquely determined by the range space of Ω_{0g}.

With $\mathcal{R}_{0g}$ determined, the excess space system (Π, Θ) as defined in Lemma 3. 3. 8:

$$AR_{0g} = R_{0g}\Pi + G\Theta \tag{A. 29}$$

can be determined trivially. The i^{th} row of Θ is given by (A. 27):

$$\theta_i = \bar{c}_i(A\text{-}D_{\mathcal{J}}C)^{\nu_i}R_{0g} \tag{A. 30}$$

since $\mathcal{R}_{0g}$ satisfies (A. 27) and g_i must satisfy:

$$e_{mi} = \bar{c}_i(A\text{-}D_{\mathcal{J}}C)^{\nu_i - 1}g_i$$

Once Θ is known, Π can be found simply by solving (A. 29):

$$\Pi = (R_{0g}^T R_{0g})^{-1}R_{0g}^T[AR_{0g} - G\Theta] \tag{A. 31}$$

If a vector f_k is removed from the set $f_1 \ldots f_r$, Θ_k is determined by removing the k^{th} row from Θ and Corollary 1 of Theorem 3. 3. 3 provides an algorithm for Π_k and R_{0gk}. The matrix $[\Delta : \beta]_k$ used in the corollary can be determined using the algorithm developed in Appendix A. 2.

A special case for the excess subspace computation exists if the state space augmentation algorithms of Section 3.3.4 are used. If the excess subspaces in $\mathcal{S}^n$ and $\mathcal{S}^{\tilde{n}}$ are denoted by $\mathcal{R}_{0g}$ and $\tilde{\mathcal{R}}_{0g}$ respectively, the range space of $\tilde{R}_{0g}$ is contained in the range space of the matrix $\tilde{R}_D$ defined in Lemma 3.3.11:

$$\tilde{R}_D = \left[\begin{array}{c|c} R_{0g}\Delta & R_{0g}\beta \\ \hline E_{\nu_{0k}}^{c} & 0 \end{array}\right] \tag{A.32}$$

$\tilde{R}_{0g}$ can be determined from $\tilde{\mathcal{R}}_{\mathcal{J}}$ and $\tilde{\mathcal{R}}_1 \ldots \tilde{\mathcal{R}}_r$ using the techniques of this section if $\tilde{\Omega}_{\mathcal{J}}$ is defined:

$$\tilde{\Omega}_{\mathcal{J}} = \left[\begin{array}{c|c} \Omega_{\mathcal{J}} & 0 \\ \hline 0 & E_{\nu_1} \end{array}\right] \tag{A.33}$$

but this involves a greater amount of computation than that used for R_{0g} originally.

A simpler procedure involves the matrix:

$$\tilde{\Omega}_D = \tilde{R}_D \tilde{R}_D^T \tag{A.34}$$

$\tilde{\Omega}_D$ is symmetric, positive semi-definite and the range space of $\tilde{\Omega}_D$ is equal to the range space of $\tilde{R}_D$. If $\tilde{M}_0$ is defined as the $(\nu_0 - \tilde{\nu}_0) \times \tilde{n}$ matrix:

$$\tilde{M}_0 = \begin{bmatrix} (\tilde{R}_D Z_1)^T \\ \vdots \\ (\tilde{R}_D Z_r)^T \end{bmatrix} \tag{A.35}$$

with Z_j, $j \neq k$, given in Lemma 3.3.11, the rows of $\tilde{M}_0$ are the rows which must be removed from the range space of $\tilde{R}_D$ to determine $\tilde{R}_{0g}$. If the orthogonal reduction procedure is applied to $\tilde{M}_0$ with:

$$\Omega^{(1)} = \tilde{\Omega}_D \tag{A.36}$$

the independent columns of the terminating matrix $\Omega^{(\nu_0 - \tilde{\nu}_0)}$ define $\tilde{R}_{0g}$. $\tilde{\Theta}$ and $\tilde{\Pi}$ can then be determined from (A.30) and (A.31) respectively.

A.4 A Completion for the State Space

The orthogonal reduction procedure applied in Steps 1 and 2 in Appendix A.2 eliminated all vectors from the range space of $\Omega^{(j)}$ which were not elements of $\bar{R}_J$, i.e. all vectors which were not in $\eta(M_D')$. These vectors are not important in determining the detection space for $f_1 \ldots f_r$, but they are important when assignment of the $n - \nu_J$ eigenvalues of A-DC not associated with $\bar{R}_J$ is considered. Also, they define a completion of the state space which is needed to determine the transformation to base normal canonical form as defined in Section 4.4. The relevant properties of these vectors are discussed below.

Consider Steps 1 and 2 of Appendix A.2. The orthogonal reduction procedure was initiated with:

$$\Omega^{(1)} = E_n \tag{A.6}$$

and terminated with

$$\Omega_J = \Omega^{(nm-m+1)} \tag{A.10}$$

A total of $n - \nu_J$ reductions using ii)' were necessary. Because of the manner in which C_J' was constructed, there were only $m - r$ rows of C_J' in (A.8) for which i) yielded a non-zero auxiliary vector[1]. Assume for

---

[1] If $\mathrm{rk}\, C < m$, m should be replaced by $\mathrm{rk}\, C$ in the discussion.

convenience that these are the last m - r rows $c'_{r+1} \ldots c'_m$ of $C'_{\mathcal{J}}$. For i = r+1 ... m, let q_i be the largest integer such that all of $c'_i \ldots c'_i(A-D_{\mathcal{J}}C)^{q_i-1}$ have non-zero auxiliary vectors and let w'_i denote the auxiliary vector associated with $c'_i(A-D_{\mathcal{J}}C)^{q_i-1}$. The implication of the discussion of intermediate terminating points in Appendix A.2 is that there are $n - \nu_{\mathcal{J}}$ linearly independent vectors in the sequence

$$w'_{r+1} \ldots (A-D_{\mathcal{J}}C)^{q_{r+1}-1} w'_{r+1} \ldots (A-D_{\mathcal{J}}C)^{q_m-1} w'_m$$

and that these vectors, in conjunction with a basis for $\bar{\mathcal{R}}_{\mathcal{J}}$, define a basis for the state space. Thus, with $z_1 \ldots z_{\nu_0}$ defined in (A.24), the matrix T_0 required in Section 4.4 is

$$T_0 = [\, z_1 \ldots z_{\nu_0} \, w'_{r+1} \ldots (A-D_{\mathcal{J}}C)^{q_{r+1}-1} w'_{r+1} \ldots (A-D_{\mathcal{J}}C)^{q_m-1} w'_m \,] \tag{A.37}$$

Section 4.2.4 states that the $\nu_{\mathcal{J}}$ eigenvalues of A-DC associated with $\bar{\mathcal{R}}_{\mathcal{J}}$ can be assigned almost arbitrarily by requiring D to be any solution to:

$$DCF = \Psi_F(A) \tag{4.2.43}$$

where:

$$\Psi_F(A) = [\, \Psi_1(A)g_1 \ldots \Psi_r(A)g_r \,] \tag{4.2.42}$$

and $\Psi_1(1) \ldots \Psi_r(1)$ are the desired minimal polynomials associated with $\bar{\mathcal{R}}_1 \ldots \bar{\mathcal{R}}_r$. If $r < rkC$, the solution of (4.2.23) is not unique. A unique solution is obtained, however, if the remaining $n - \nu_{\mathcal{J}}$ eigenvalues of A-DC are specified. The most convenient manner to accomplish this is to use the vectors $w'_{r+1} \ldots w'_m$.

Lemma 3.2.8 provides the basis for the proposed algorithm. In essence, the lemma states that $w'_{r+1} \ldots w'_m$ can be treated as detection generators (mod $\bar{\mathcal{R}}_{\mathcal{J}}$) and that (4.2.43) leaves sufficient freedom in the choice of D to make the spaces spanned by $w'_{r+1} \ldots (A-DC)^{q_{r+1}-1} w'_{r+1}$ etc. invariant (mod $\bar{\mathcal{R}}_{\mathcal{J}}$). Let D'_Ψ be any solution of (4.2.43).

If $\psi'_{r+1}(\cdot) \ldots \psi'_m(\cdot)$ are the desired relative minimal polynomials associated with these spaces:

$$\psi'_i(\lambda) = \lambda^{q_i} + p'_{iq_i}\lambda^{q_i-1} + \ldots p'_{i1} \qquad i = r+1 \ldots m \tag{A.38}$$

then a unique solution for D which assigns all n eigenvalues of A-DC to the desired values is:

$$D = D'_\psi + D_\Gamma \tag{A.39}$$

where D_Γ is the solution of:

$$D_\Gamma C'_\mathcal{J} W = \psi'_W(A-D'_\psi C) \tag{A.40}$$

and:

$$W = [\,(A-D'_\psi C)^{q_{r+1}-1} w'_{r+1} \ldots (A-D'_\psi C)^{q_m-1} w'_m\,] \tag{A.41}$$

$$\psi'_W(A-D'_\psi C) = [\,\psi'_{r+1}(A-D'_\psi C)w'_{r+1} \ldots \psi'_m(A-D'_\psi C)w'_m\,] \tag{A.42}$$

Regardless of the manner in which D'_ψ is selected, (A.39) provides a unique representation for D. Thus, if base normal canonical form is used to define D'_ψ as:

$$D'_\psi = D_{FR} + D_\psi \tag{A.43}$$

where D_{FR} and D_ψ are transforms of the solutions of (4.4.17) and (4.4.20), then (A.39) and (A.40) can be used to determine D_Γ and complete the selection of D.

APPENDIX B

A USEFUL INTEGRAL

Computation of the power spectral density for the output of a linear system involves an integral of the form

$$\Phi_0(s) = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} H(s)\, \Phi_i(s)\, H(-s)^T ds \qquad (B.1)$$

where $\Phi_i(s)$ and $H(s)$ involve polynomials in the variable s. This integral has been evaluated for polynomials up to tenth-order by Newton, Gould and Kaiser (1957). The results for polynomials up to sixth-order are reproduced in this appendix.

The integral in (B.1) can be written in the general form

$$I_n = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} ds\, \frac{c(s)c(-s)}{d(s)d(-s)} \qquad (B.2)$$

where

$$c(s) = c_{n-1}s^{n-1} + \ldots + c_0 \qquad (B.3)$$

$$d(s) = d_n s^n + \ldots + d_0 \qquad (B.4)$$

The solution for I_n for $n \leq 6$ is expressed in Table B.1 as a function of the coefficients $c_0 \ldots c_{n-1}$ and $d_0 \ldots d_n$.

Table B. 1

$$I_1 = \frac{c_0^2}{2d_0d_1}$$

$$I_2 = \frac{c_1^2d_0 + c_0^2d_2}{2d_0d_1d_2}$$

$$I_3 = \frac{c_2^2d_0d_1 + (c_1^2 - 2c_0c_2)d_0d_3 + c_0^2d_2d_3}{2d_0d_3(-d_0d_3 + d_1d_2)}$$

$$I_4 = \frac{c_3^2(-d_0^2d_3 + d_0d_1d_2) + (c_2^2 - 2c_1c_3)d_0d_1d_4 + (c_1^2 - 2c_0c_2)d_0d_3d_4 + c_0^2(-d_1d_4^2 + d_2d_3d_4)}{2d_0d_4(-d_0d_3^2 - d_1^2c_4 + d_1d_2d_3)}$$

$$I_5 = \frac{1}{2\Delta_5}\left[c_4^2m_0 + (c_3^2 - 2c_2c_4)m_1 + (c_2^2 - 2c_1c_3 + 2c_0c_4)m_2 + (c_1^2 - 2c_0c_2)m_3 + c_0^2m_4\right]$$

where

$$m_0 = \frac{1}{d_5}(d_3m_1 - d_1m_2) \qquad m_3 = \frac{1}{d_0}(d_2m_2 - d_4m_1)$$

$$m_1 = -d_0d_3 + d_1d_2 \qquad m_4 = \frac{1}{d_0}(d_2m_3 - d_4m_2)$$

$$m_2 = -d_0d_5 + d_1d_4 \qquad \Delta_5 = d_0(d_1m_4 - d_3m_3 + d_5m_2)$$

$$I_6 = \frac{1}{2\Delta_6}\left[c_5^2m_0 + (c_4^2 - 2c_3c_5)m_1 + (c_3^2 - 2c_2c_4 + 2c_1c_5)m_2 + (c_2^2 - 2c_1c_3 + 2c_0c_4)m_3 + (c_1^2 - 2c_0c_2)m_4 + c_0^2m_5\right]$$

where

Table B. 1 (cont.)

$$m_0 = \frac{1}{d_6}(d_4 m_1 - d_2 m_2 + d_0 m_3)$$

$$m_1 = -d_0 d_1 d_5 + d_0 d_3^2 + d_1^2 d_4 - d_1 d_2 d_3$$

$$m_2 = d_0 d_3 d_5 + d_1^2 d_6 - d_1 d_2 d_5$$

$$m_3 = d_0 d_5^2 + d_1 d_3 d_6 - d_1 d_4 d_5$$

$$m_4 = \frac{1}{d_0}(d_2 m_3 - d_4 m_2 + d_6 m_1)$$

$$m_5 = \frac{1}{d_0}(d_2 m_4 - d_4 m_3 + d_6 m_2)$$

$$\Delta_6 = d_0(d_1 m_5 - d_3 m_4 + d_5 m_3)$$

Bibliography

1. Arbib, M. A., Kalman, R. E. and Falb, P. L., Topics in Mathematical Systems Theory, McGraw-Hill, 1969.

2. Baggeroer, A. B., State Variables in Communication Theory, M. I. T. Press, Cambridge, Mass., 1970.

3. Beard, R. V., Failure Accommodation in Linear Systems through Self-Reorganization, Report No. MVT-71-1, Man Vehicle Laboratory, M. I. T., Cambridge, Mass., February 1971.

4. Blakelock, J. H., Automatic Control of Aircraft and Missiles, Wiley, 1965.

5. Brasch, F. M. Jr. and Pearson, J. B., "Pole Placement Using Dynamic Compensators" IEEE Transactions on Automatic Control, Vol 15 No.1, February 1970, pp 34-43.

6. Brockett, R. W., Finite Dimensional Linear Systems, Wiley, 1970.

7. Bryson, A. E. Jr. and Ho, Y.C. Applied Optimal Control, Blaisdell, Waltham, Mass., 1969.

8. Bucy, R. S. "Canonical Forms for Multivariable Systems" IEEE Transactions on Automatic Control, Vol 13 No. 5, October 1968, pp 567-569.

9. Buxbaum, P. J. and Haddad, R. A., "Recursive Optimal Estimation for a Class of Nongaussian Processes, "Proc. of Symp. on Computer Processing in Communications, Polytechnic Institue of Brooklyn, April 8-10, 1969.

10. Chien, T. T., An Adaptive Technique for a Redundant Navigation System, The Charles Stark Draper Laboratory, Inc., T-560, Cambridge, Mass., February 1972.

11. Davison, E. J. "On Pole Assignment in Linear Systems with Incomplete State Feedback" IEEE Transactions on Automatic Control, Vol 15 No. 3, June 1970, pp 348-351.

12. Davison, E. J. and McLane, P. J., "Disturbance Localization and Decoupling in Stationary Linear Multivariable Systems" IEEE Transactions on Automatic Control, Vol 15 No. 1, February 1970. pp 33-134.

13. Davison, E. J. and Wang, S. H., "Properties of Linear Time-Invariant Multivariable Systems Subject to Arbitrary Output and State Feedback" IEEE Transactions on Automatic Control, Vol 18 No. 1, February 1970, pp 24-32.

14. Ephgrave, J. T., "Redundant Adaptive Strapdown Inertial Navigation System, The Aerospace Corporation, TOR-0066(5306)-10, October 1969.

15. Falb, P. L. and Wolovich, W. A., "Decoupling in the Design and Synthesis of Multivariable Control Systems" IEEE Transactions on Automatic Control, Vol 12 No. 6, December 1967, pp 65 -659.

16. Gantmacher, F. R., The Theory of Matrices (Vols. I and II), Chelsea Publishing Co., 1959.

17. Gilbert, E. G., "The Decoupling of Multivariate Systems by State Feedback" SIAM Journal of Control , Vol 7 No. 1, February 1969, pp 50-63.

18. Gilmore, J. P. and McKern, R. A., "A Redundant Strapdown Inertial System Mechanization - SIRU" M. I. T. C. S. Draper Laboratory, E-2527, August 1970.

19. Halmos, P. R., Finite Dimensional Vector Space, D. Van Nostrand Co., Princeton, N. J., 1958.

20. Heymann, M. "A Unique Canonical Form for Multivariable Linear Systems" International Journal of Control, Vol 12 No. 6, 1970, pp 913-927.

21. Hildebrand, F. B., Methods of Applied Mathematics, Prentiss-Hall, 1965.

22. Howze, J. W. and Pearson, J. B., "Decoupling and Arbitrary Pole Placement in Linear Systems Using Output Feedback" IEEE Transactions on Automatic Control, Vol. 15 No. 6, December 1970, pp 660-663.

23. Jazwinski, A. H., Stochastic Processes and Filtering Theory, Academic Press, New York, 1970.

24. Kalman, R. E., "A New Approach to Linear Filtering and Prediction Problems", Trans. ASME, Ser. D, J. Basic Eng., Vol 82, 1960, pp 35-45.

25. Kerr, T. H., "A Two Ellipsoid Overlap Test for Real-Time Failure Detection and Isolation by Confidence Regions" presented at Pittsburgh Conference on Modeling and Simulation, April 24-26, 1974.

26. Lainiotis, D. G., "Joint Detection, Estimation and System Identification" Information and Control, Vol. 19, No. 1, August 1971, pp 75-92.

27. Lee, R. C. K., Optimal Estimation, Identification, and Control, Cambridge, Mass., M.I.T. Press, 1964.

28. Leondes, C. T., and Novak, L. M., "Optimal Minimal-Order Observers for Discrete-Time Systems -- A Unified Theory", Proceedings at the 5th IFAC Congress, June 1972.

29. Luenberger, D. G., "Observing the State of a Linear System", IEEE Transactions on Military Electronics, Vol. MIL-8, No. 2, April 1964, pp 74-80.

30. Luenberger, D. G., "Canonical Forms for Linear Multivariable Systems" IEEE Transactions on Automatic Control" Vol 12 No.3, June 1967, pp 290-293.

31. Luenberger, D. G., "An Introduction to Observers," IEEE Trans. Automatic Control, Vol 16 No. 6, December 1971, pp 596-602.

32. McAuley, R. J. and Denlinger, E., "A Decision-Directed Adaptive Tracker", IEEE Trans. on Aerospace and Electronic Systems, Vol. AES-9, March 1973, pp 229-236.

33. Mehra, R. K. and Peschon, J., "An Innovations Approach to Fault Detection and Diagnosis in Dynamic Systems", Automatica, Vol. 7, 1971, pp 637-640.

34. Newton, G. C. Jr., Gould, L. A. and Kaiser, J. F., Analytical Design of Linear Feedback Controls, Wiley, 1957.

35. Ogata, K., State Space Analysis of Control Systems, Prentice-Hall, 1967.

36. Papoulis, A. Probability, Random Variables, and Stochastic Processes, McGraw-Hill, 1965.

37. Pejsa, A. J., "Optimum Orientation and Accuracy of Redundant Sensor Arrays", Honeywell, Inc., 1971.

38. Potter, J. E. and Deckert, J. C., "Gyro and Accelerometer Failure Detection and Identification in Redundant Systems", M. I. T. C. S. Draper Laboratory, E-2686, May 1972.

39. Sanyal, P. and Shen, C. N., "Bayes Decision Rule for Rapid Detection and Adaptive Estimation Scheme with Space Applications," Preprints of the 4th JACC, Columbus, Ohio, June 1973, pp 23-26.

40. Schweppe, F. C., "Evaluation of Likelihood Functions for Gaussian Signals", IEEE Transactions on Information Theory, Vol 11, p 61-70, January 1966.

41. Tse, E., "Observer-Estimators for Discrete-time Systems", IEEE Transactions on Automatic Control, Vol 18 No. 1, February 1973, pp 10-16.

42. Truxal, J. G., Control System Synthesis, McGraw-Hill, 1955.

43. Van Tress, H. L., Detection, Estimation and Modulation Theory, Part I, Wiley, 1968.

44. Van Tress, H. L., "Application of State-Variable Techniques in Detection Theory". Proceedings of the IEEE, Vol 58 No. 5, May 1970.

45. Willsky, A. S. and Jones, H. L., "A Generalized Likelihood Ratio Approach to State Estimation in Linear Systems Subject to Abrupt Changes", Proc. of the IEEE Conf. on Decision and Control, (Phoenix, Arizona) November 1974.

46. Wonham, W. M., "On Pole Assignment in Multi-Input Controllable Linear Systems", IEEE Transactions on Automatic Control, Vol 2 No. 6, December 1967, pp 660-665.

47. Wonham, W. M., "Dynamic Observers - Geometric Theory", IEEE Transactions on Automatic Control, Vol 15 No. 2, April 1970, pp 258-259.

48. Wonham, W. M. and Morse, A. S., "Decoupling and Pole Assignment in Linear Multivariable Systems: A Geometric Approach", SIAM Journal of Control, Vol 8 No. 1, February 1970, pp 1-18.

49. Wonham, W. M. and Morse, A. S., "Decoupling and Pole Assignment by Dynamic Compensation" SIAM Journal of Control, Vol 8 No. 3, August 1970, pp 317-337.

50. Wonham, W. M., and Morse, A. S., "Status of Noninteracting Control", IEEE Trans. Automatic Control, Vol 16 No. 6, December 1971, pp 568-581.

BIOGRAPHY

Harold Lee Jones was born on May 13, 1945 in Great Bend, Kansas. He attended public schools in Great Bend and graduated from Great Bend High School in 1963. He received his B.S. and M.S. degrees from MIT in 1967 and 1969, respectively. As an undergraduate, Mr. Jones held a National Merit Scholarship; as a graduate student, he was supported in part by an AC Electronics Fellowship.

Upon entering graduate school, Mr. Jones became a research assistant with the Measurement Systems Laboratory. His subsequent research activities included participation in the design of a hydrofoil-based hydrographic survey system and of several satellite-based experiments. In June 1971, he became a full-time research assistant with The Charles Stark Draper Laboratory, Inc.

Mr. Jones lives in Reading, Mass. with his wife, the former Cheryl Bowen. He is a member of Tau Beta Pi, and Sigma Gamma Tau honorary fraternities.