Chapter 9

A Game Theoretic Fault Detection Filter

THE FAULT DETECTION FILTER was introduced by Beard (Beard 1971) in his doctoral thesis and later refined by Jones (Jones 1973) who gave it a geometric interpretation. Since then, the fault detection filter has undergone many refinements. White (White and Speyer 1987) derived an eigenstructure assignment design algorithm. Massoumnia (Massoumnia 1986) used advances in geometric theory to derive a complete and elegant geometric version of a fault detection filter and derived a reduced-order fault detector (Massoumnia et al. 1989). Most recently, Douglas robustified the filter to parameter variations (Douglas 1993) and (Douglas and Speyer 1996) and also derived a version of the filter which bounds disturbance transmission (Douglas and Speyer 1995a). The backround of Appendix A, design methods of Appendices B and C and the application to vehicle fault detection of Sections 2 through 5 all follow from these sources.

Common to all of these sources is an underlying structure of independent, invariant subspaces. Most design algorithms, an exception being (Douglas and Speyer 1995a), rely on spectral methods, that is, specifying eigenvalues and eigenvectors, since these methods lead directly to the needed filter structure. Spectral methods, however, also limit the applicability of fault detection filters to linear, time-invariant systems and filters designed by these methods can have poor robustness to parameter variations (Lee 1994).

For these reasons, we take a different approach to detection filter design. We look at the fault detection process as a disturbance attenuation problem and convert the process into a differential game which leads to the final design. The game is one in which the player is a state estimate and the adversaries are all of the exogenous signals, save the fault to be detected. The player attempts to exclude the adversaries from a specified portion of the state-space much in the same way that the invariant subspace structure of the fault detection filter restricts state trajectories when driven by faults. The end result is an $\mathcal{H}_\infty$-type filter which bounds disturbance transmission.

Since fault detection filters block transmission, it would seem reasonable to expect that in the limiting case when the $\mathcal{H}_\infty$ transmission bound is brought to zero, the game filter no longer approximates, but actually becomes a fault detection filter. We will prove that this is indeed the case. For linear time-invariant (LTI) systems, we will show, in fact, that the game filter becomes a Beard-Jones fault detector in the sense of (Douglas 1993): faults other than the one to be detected are restricted to a subspace which is invariant and unobservable.

The method developed here has wider applicability than current techniques since time-invariance is never assumed in the game solution. Thus, for a class of time-varying systems, results analogous to the LTI case exist in the limit as disturbance bounds are taken to zero. It is also possible with this method to deal with model uncertainty by treating it as another element in the differential game (Chichka and Speyer 1995, Mangoubi et al. 1994). In this manner, sensitivity to parameter variations can be reduced. Finally, by using a game theoretic approach, the designer has the freedom to choose the extent to which the game filter behaves as an $\mathcal{H}_\infty$ filter and the extent to which it behaves like a detection filter. This flexibility is unique to this method of fault detection filter design.

The development of game theoretic estimation closely followed the development of game theoretic control theory. The most notable and the most cited (and most unreadable) work in the latter was the paper by Doyle *et al.* (Doyle et al. 1989). The ascendant of the work presented here is the paper by Rhee and Speyer (Rhee and Speyer 1991) which derived the two Riccati solution of (Doyle et al. 1989) via the calculus of variations. It is hard to credit the first derivation of the game theoretic estimator, though (Banavar and Speyer 1991) or (Yaesh and Shaked 1993) are probable candidates.

In Sections 9.1 and 9.2, we pose a disturbance attenuation problem which models the fault detection process for a large class of systems which includes some time-varying systems. The solution to this problem leads to the game theoretic fault detection filter. In Section 9.3, we analyze sufficient conditions for our game cost to be non-positive. This will enable us to show the existence of the filter in the limit and analyze its structure. In Section 9.4, we return to the LTI case and prove that the limiting detection filter is equivalent to the Beard-Jones fault detection filter. In Section 9.5, we use the limiting form of the game theoretic filter to derive a reduced-order estimator for fault detection. Finally, in Section 9.6 we go through an example which shows that the filter is an effective fault detector for finite values of the disturbance attenuation bound and in the limit.

9.1 A Disturbance Attenuation Approach to Fault Detection

Consider a linear system in which q possible faults have been modeled:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t) + F_1(t)\mu_1(t) + \sum_{i=2}^{q} F_i(t)\mu_i(t) \tag{9.1}$$

$$y(t) = C(t)x(t) + v(t). \tag{9.2}$$

It is desired to detect the appearance of μ_1, the ***target fault***, in the presence of sensor noise, v, and the possible presence of other faults $\mu_i, i \neq 1$, the ***nuisance faults***. Following the standard assumptions of Appendix A, we will assume that each of the F_i's are monic and that (C,A) is an observable pair. Also, since u is a known function of $t \in [t_0, t_1]$, we will drop the Bu term for convenience. We will also neglect to explicitly show the possible time dependence of the system matrices, though the reader should keep this possibility in mind. For convenience, we define:

$$\hat{\mu}_2 = \left\{ \begin{array}{c} \mu_2 \\ \vdots \\ \mu_q \end{array} \right\},$$

and use the definition of $\hat{F}_i$ (a.10) so that the state equation becomes:

$$\dot{x} = Ax + F_1\mu_1 + \hat{F}_1\hat{\mu}_2$$

The definition that we propose is based upon disturbance attenuation. We use (a.11) and define the corresponding residual signal z_1 associated with μ_1 as the output signal. A disturbance attenuation problem would be to limit the transmission of the nuisance faults and the sensor noise to this output. For a fault detection filter problem we want to block this transmission entirely.

Definition 9.1 (Fault Detection Filter Problem). Find an estimator such that:

$$\frac{\|z_1\|^2}{\|\hat{\mu}_2\|^2} = 0 \quad \text{and} \quad \frac{\|z_1\|^2}{\|\mu_1\|^2} \neq 0.$$

Clearly, in the time-invariant case, the solution to the fault detection filter problem as defined by Definition A.1 solves the general fault detection filter problem that we have defined above. Later on, we will show that these definitions are equivalent in the time-invariant case by showing that the solution to Definition 9.1 solves the problem defined by Definition A.1. We need this alternative definition to account for time-varying systems. In such cases, we cannot talk about invariant subspaces and also observability becomes a trickier concept. Thus instead of defining the filter structure, we must content ourselves with merely describing its action.

9.2 A Game Theoretic Filter for Fault Detection in a General Class of Systems

We arrive at a solution to the fault detection filter problem as defined by Definition 9.1 by first solving the disturbance attenuation problem. The solution to the fault detection filter problem then comes when we take the limit of the disturbance attenuation solution. The results that we find here, however, are valuable in their own right. As we will see, the game filter that we get from the disturbance attenuation problem is itself a useful filter for fault detection.

We begin by quantifying the problem objective with a disturbance attenuation function, the ratio of the norm of the output to the norms of the inputs. For this problem, the function is:

$$D_{af} = \frac{\int_{t_0}^{t_1} \|\hat{H}_1 C(x-\hat{x})\|_{Q_1}^2 dt}{\int_{t_1}^{t_2} [\|\hat{\mu}_2\|_{M_2^{-1}}^2 + \|v\|_{V^{-1}}^2 + \|N_1 C(x-\hat{x})\|_{R_1}^2] dt + \|x(t_0) - \hat{x}_0\|_{P_0^{-1}}^2}$$

where $N_1 \triangleq I - \hat{H}_1$ and M_2, V, R_1, P_0 are weighting matrices. The disturbance attenuation problem is to find an estimator so that for all adversaries $\hat{\mu}_2$, $v \in L_2[t_1, t_2]$, $x(0) \in \mathcal{R}^n$:

$$D_{af} \leq \gamma. \tag{9.3}$$

We will refer to γ as the disturbance attenuation bound. Once again, the assumptions that we will make are: 1) (C, A) is a an observable pair 2) F_i, $i = 1 \ldots q$ is monic 3) i, the number of iterations of (a.9) needed to make CB_i full rank is constant over the whole time interval.

To solve (9.3), convert it into a differential game with cost function:

$$\begin{aligned} J &= -\|x(t_0) - \hat{x}_0\|_{\Pi_0}^2 \\ &+ \int_{t_0}^{t_1} \left[\|\hat{H}_1 C(x-\hat{x})\|_{Q_1}^2 - \gamma\left(\|\hat{\mu}_2\|_{M_2^{-1}}^2 + \|v\|_{V^{-1}}^2 + \|N_1 C(x-\hat{x})\|_{R_1}^2\right)\right] dt \end{aligned} \tag{9.4}$$

Note that $\Pi_0 \triangleq \gamma P_0^{-1}$. We want to find:

$$\min_{\hat{x}} \max_{v} \max_{\hat{\mu}_2} \max_{x(t_0)} J \leq 0 \tag{9.5}$$

subject to:

$$\dot{x} = Ax + \hat{F}_1\hat{\mu}_2. \tag{9.6}$$

In anticipation of the steps which will be required for the game solution, we will rewrite the sensor noise term $\|v\|^2_{V^{-1}}$ to the equivalent $\|y - Cx\|^2_{V^{-1}}$:

$$\begin{aligned} J &= -\|x(t_0) - \hat{x}_0\|^2_{\Pi_0} \\ &\quad + \int_{t_0}^{t_1} \left[\|\hat{H}_1 C(x - \hat{x})\|^2_{Q_1} - \gamma\left(\|\hat{\mu}_2\|^2_{M_2^{-1}} + \|y - Cx\|^2_{V^{-1}} + \|N_1 C(x - \hat{x})\|^2_{R_1}\right)\right] dt \end{aligned}$$

This is a common step in the solution of quadratic minimization problems. The game problem then becomes:

$$\min_{\hat{x}} \max_{y} \max_{\hat{\mu}_2} \max_{x(t_0)} J \leq 0.$$

An interpretation of the maximization of the cost with respect to y is elusive given the measurement equation (9.2), the presence of v in (9.2), and the interplay of the different players in determining the state, x. Our view is taken from (Banavar and Speyer 1991) which looks at this extremization as incorporating a "worst-case measurement" into the game. There are other interpretations (see for instance (Yaesh and Shaked 1993)), but ultimately the question of proper interpretation becomes an exercise in tail-chasing since the mechanics of the solution remains the same as does the solution itself.

An element that is missing in our problem statement (9.4), (9.5), (9.6) is the target fault, μ_1. This is not an oversight. It would seem logical to include enhancing the transmission of μ_1 as part of the game, but there is no obvious way to include such an objective in the game cost. Moreover, extremizing the cost with respect to μ_1 leads to assumptions upon the temporal behavior of the target fault. This can be quite detrimental to filter performance if these assumptions are wrong (which is why fault detection filters are designed without any such assumptions). Thus, since μ_1 is not part of the differential game, we set it to zero for convenience when we work through the solution. This places the burden on the designer to make sure the set of faults that he chooses for the filter design leads to a well-posed problem. Well-posedness is discussed in Section 9.1 and for LTI systems is easily checked by Equation a.7.

9.2.1 Maximization with Respect to $x(t_0)$ and $\hat{\mu}_2$

We will solve our problem in two steps beginning with the subproblem:

$$\max_{\hat{\mu}_2} \max_{x(t_0)} J \leq 0.$$

The reasoning for this order of the extremizations is given in (Banavar and Speyer 1991).

We begin by appending the dynamics of the system to the cost with a Lagrange multiplier, λ^T:

$$\begin{aligned} J &= -\|x(t_0) - \hat{x}_0\|^2_{\Pi_0} + \int_{t_0}^{t_1} \Big[\|\hat{H}_1 C(x-\hat{x})\|^2_{Q_1} \\ &\quad -\gamma\left(\|\hat{\mu}_2\|^2_{M_2^{-1}} + \|y - Cx\|^2_{V^{-1}} + \|N_1 C(x-\hat{x})\|^2_{R_1}\right) + \lambda^T(Ax + \hat{F}_1\hat{\mu}_2 - \dot{x})\Big]\, dt \end{aligned} \tag{9.7}$$

Integrate $\lambda\dot{x}$ by parts:

$$\begin{aligned} J = &-\|x(t_0) - \hat{x}_0\|^2_{\Pi_0} + \lambda(t_0)^T x(t_0) - \lambda(t_1)^T x(t_1) + \int_{t_0}^{t_1} \Big[\|\hat{H}_1 C(x-\hat{x})\|^2_{Q_1} \\ &-\gamma\left(\|\hat{\mu}_2\|^2_{M_2^{-1}} + \|y - Cx\|^2_{V^{-1}} + \|N_1 C(x-\hat{x})\|^2_{R_1}\right) + \lambda^T(Ax + \hat{F}_1\hat{\mu}_2) + \dot{\lambda}^T x\Big] dt \end{aligned} \tag{9.8}$$

and then take the variation of (9.8) with respect to $\hat{\mu}_2$ and $x(t_0)$:

$$\begin{aligned} \delta J &= -\left[(x(t_0) - \hat{x}_0)^T \Pi_0 + \lambda(t_0)^T\right] \delta x(t_0) - \lambda(t_1)^T \delta x(t_1) \\ &\quad + \int_{t_0}^{t_1} \Big\{\Big[(x-\hat{x})^T C^T \hat{H}_1^T Q_1 \hat{H}_1 C + \gamma(y - Cx)^T V^{-1} C - \gamma(x-\hat{x})^T C^T N_1^T R_1 N_1 C \\ &\qquad\qquad + \dot{\lambda}^T + \lambda^T A\Big] \delta x + \left[-\gamma \hat{\mu}_2^T M_2^{-1} + \lambda^T \hat{F}_1\right] \delta\hat{\mu}_2\Big\}\, dt \end{aligned} \tag{9.9}$$

The above implies that first-order necessary conditions for J to be maximized are:

$$\hat{\mu}_2 = \frac{1}{\gamma} M_2 \hat{F}_1^T \lambda \tag{9.10a}$$

$$-\dot{\lambda} = A^T\lambda + C^T(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1) C(x-\hat{x}) + \gamma C^T V^{-1}(y - Cx) \tag{9.10b}$$

$$\lambda(t_1) = 0 \tag{9.10c}$$

$$\lambda(t_0) = \Pi_0[x(t_0) - \hat{x}_0] \tag{9.10d}$$

Substituting (9.10a) into our dynamics (9.6) and using (9.10b), we obtain a two point boundary value problem:

$$\begin{Bmatrix} \dot{x} \\ \dot{\lambda} \end{Bmatrix} = \begin{bmatrix} A & \frac{1}{\gamma}\hat{F}_1 M_2 \hat{F}_1^T \\ -C^T(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma V^{-1})C & -A^T \end{bmatrix} \begin{Bmatrix} x \\ \lambda \end{Bmatrix} + \begin{Bmatrix} 0 \\ C^T(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1)C\hat{x} - \gamma C^T V^{-1} y \end{Bmatrix} \tag{9.11}$$

If we assume solutions x^* and λ^* to (9.11) and a quadratic form of the optimal return function, then:

$$\lambda^* = \Pi(x^* - x_p) \tag{9.12}$$

where x_p is a measurement dependent variable which will be shown to reduce to the estimate of the optimal state. Using (9.12) and the first equation of (9.11), the second equation of (9.11) becomes:

$$\begin{aligned} 0 \;=\; & \left[\dot{\Pi} + A^T\Pi + \Pi A + \frac{1}{\gamma}\Pi \hat{F}_1 M_2 \hat{F}_1^T \Pi + C^T\left(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma V^{-1}\right)C\right]x^* \\ & -\dot{\Pi}x_p - \Pi\dot{x}_p - A^T\Pi x_p - C^T(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1)C\hat{x} + \gamma C^T V^{-1} y \end{aligned} \tag{9.13}$$

Now, add and subtract

$$\gamma C^T V^{-1} C\hat{x}$$

and

$$\left[\Pi A + C^T\left(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma V^{-1}\right)C\right]x_p$$

to (9.13) to get:

$$\begin{aligned} 0 \;=\; & \left[\dot{\Pi} + A^T\Pi + \Pi A + \frac{1}{\gamma}\Pi \hat{F}_1 M_2 \hat{F}_1^T \Pi + C^T\left(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma V^{-1}\right)C\right](x^* - x_p) \\ & -\Pi\dot{x}_p + \Pi A x_p - \left[C^T\left(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma C^T V^{-1}\right)C\right](\hat{x} - x_p) \\ & +\gamma C^T V^{-1}(y - C\hat{x}) \end{aligned} \tag{9.14}$$

Thus, if we set:

$$-\dot{\Pi} = A^T\Pi + \Pi A + \frac{1}{\gamma}\Pi \hat{F}_1 M_2 \hat{F}_1^T \Pi + C^T\left(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma V^{-1}\right)C \tag{9.15}$$

$$\Pi\dot{x}_p = \Pi A x_p - C^T\left(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma V^{-1}\right)C(\hat{x} - x_p) + \gamma C^T V^{-1}(y - C\hat{x}) \tag{9.16}$$

(9.14) is satisfied identically. (9.15) is an estimator Riccati equation. If we set:

$$\Pi = \gamma P^{-1},$$

we can convert (9.15) into a Riccati equation:

$$\dot{P} = PA^T + PA - PC^T(V^{-1} + N_1^T R_1 N_1 - \frac{1}{\gamma}\hat{H}_1^T Q_1 \hat{H}_1^T)CP + \hat{F}_1 M_2 \hat{F}_1^T \tag{9.17}$$

as seen in (Banavar and Speyer 1991), (Rhee and Speyer 1991) and (Doyle et al. 1989). (9.16) looks like an estimator, but its final form will not become apparent until we solve the second half of the game problem.

9.2.2 Minimization with Respect to $\hat{x}$ and Maximization with Respect to y

The first part of our game solution led to optimal values for μ and $x(t_0)$:

$$\begin{aligned} \mu^* &= \frac{1}{\gamma}\hat{F}_1 M_2 \hat{F}_1^T \lambda \\ x(t_0)^* &= \Pi_0^{-1}\lambda(t_0) + \hat{x}_0 \end{aligned}$$

If we substitute these optimal values into the cost function (9.4) we obtain a new cost, $\bar{J}$, which is written as:

$$\begin{aligned} \bar{J} &= -\|\lambda(t_0)\|^2_{\Pi_0^{-1}} + \\ & \int_{t_0}^{t_1} \left[\|x - \hat{x}\|^2_{C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1)C} - \|\lambda\|^2_{\frac{1}{\gamma}\hat{F}_1 M_2 \hat{F}_1^T} - \gamma \|y - Cx\|^2_{V^{-1}} \right] dt \end{aligned} \tag{9.18}$$

The game is then:

$$\min_{\hat{x}} \max_{y} \bar{J} \leq 0$$

subject to the dynamic equation (9.16). We begin towards the solution to this game by adding the identically zero term:

$$\|\lambda(t_0)\|^2_{\Pi(t_0)^{-1}} - \|\lambda(t_1)\|^2_{\Pi(t_1)^{-1}} + \int_{t_0}^{t_1} \frac{d}{dt}\|\lambda(t)\|^2_{\Pi^{-1}} dt = 0$$

to (9.18). After applying the boundary condition for λ at t_1 (9.10c) and carrying out the differentiation of the $\|\lambda\|^2_{\Pi^{-1}}$ term, we get:

$$\begin{aligned} \bar{J} = \int_{t_0}^{t_1} \Big[& \|(x - \hat{x})\|^2_{C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1)C} - \|\lambda\|^2_{\frac{1}{\gamma}\hat{F}_1 M_2 \hat{F}_1^T} - \gamma\|y - Cx\|^2_{V^{-1}} \\ & + \dot{\lambda}^T \Pi^{-1} \lambda^T + \lambda^T \dot{\Pi}^{-1} \lambda + \lambda^T \Pi^{-1} \dot{\lambda} \Big] dt + \|\lambda(t_0)\|^2_{\Pi^{-1}(t_0) - \Pi_0^{-1}} \end{aligned} \tag{9.19}$$

Note that (9.19) provides a boundary condition for (9.15):

$$\Pi(t_0) = \Pi_0$$

Applying this boundary condition and substituting the differential equation for λ, (9.10b), into (9.19) leads to:

$$\begin{aligned}\bar{J} &= \int_{t_0}^{t_1} \Big[\lambda^T\left(-A\Pi^{-1} - \Pi^{-1}A^T - \hat{F}_1 M_2 \hat{F}_1^T + \dot{\Pi}^{-1}\right)\lambda \\ &\quad + (x-\hat{x})^T C^T \left(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1\right) C(x-\hat{x}) \\ &\quad - (x-\hat{x})^T C^T \left(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1\right) C\Pi^{-1}\lambda \\ &\quad - \lambda^T \Pi^{-1} C^T \left(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1\right) C(x-\hat{x}) \\ &\quad - \gamma(y-Cx)^T V^{-1}(y-Cx) \\ &\quad + (y-Cx)^T V^{-1} C\Pi^{-1}\lambda + \lambda^T \Pi^{-1} C^T V^{-1}(y-Cx)\Big]\, dt \end{aligned} \tag{9.20}$$

From (9.15) the differential equation for Π^{-1} is:

$$\dot{\Pi}^{-1} = -\Pi^{-1}\dot{\Pi}\Pi^{-1} \tag{9.21a}$$

$$= \Pi^{-1}A^T + A\Pi^{-1} + \frac{1}{\gamma}\hat{F}_1 M_2 \hat{F}_1^T + \Pi^{-1}C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1 - \gamma V^{-1})C\Pi^{-1} \tag{9.21b}$$

After we insert (9.21) into (9.20) and cancel terms, we are left with what turns out to be a pair of quadratic terms:

$$\begin{aligned}\bar{J} &= \int_{t_0}^{t_1} \Big\{\left[\Pi^{-1}\lambda - (x-\hat{x})\right]^T C^T \left(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1\right) C \left[\Pi^{-1}\lambda - (x-\hat{x})\right] \\ &\quad - \gamma\left[C\Pi^{-1}\lambda + (y-Cx)\right]^T V^{-1}\left[C\Pi^{-1}\lambda + (y-Cx)\right]\Big\}\, dt \end{aligned} \tag{9.22}$$

Now, use the solution for the optimal value of λ (9.12) and substitute into (9.22) to get:

$$\begin{aligned}\bar{J} &= \int_{t_0}^{t_1} \Big[(\hat{x}-x_p)^T C^T \left(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1\right) C(\hat{x}-x_p) \\ &\quad - \gamma(y - Cx_p)^T V^{-1}(y - Cx_p)\Big]\, dt \end{aligned} \tag{9.23}$$

Given the cost (9.23) and the dynamics (9.16), the solutions to this game are:

$$\hat{x}^* = x_p \tag{9.24a}$$

$$y^* = Cx_p \tag{9.24b}$$

From (9.24) we can rewrite (9.16) as:

$$\Pi\dot{\hat{x}}^* = \Pi A\hat{x}^* + \gamma C^T V^{-1}(y - C\hat{x}^*) \tag{9.25}$$

Since Π is positive-definite for $\gamma > 0$, we can rewrite (9.25):

$$\dot{\hat{x}}^* = A\hat{x}^* + \gamma\Pi^{-1} C^T V^{-1}(y - C\hat{x}^*) \tag{9.26}$$

Alternatively, the analyst could use (9.17) and:

$$\dot{\hat{x}}^* = A\hat{x}^* + PC^T V^{-1}(y - C\hat{x}^*)$$

This form of the filter is equivalent to (9.26); however, experience has shown that numerical problems are more likely to be seen when trying to find a solution to (9.17) than (9.15) when γ is brought to extremely small values. For convenience, we will write $\hat{x}$ instead of $\hat{x}^*$ when referring to the optimal state estimate with the understanding that it is the estimate that comes from the game solution which is being used.

9.2.3 Steady-State Results

In many cases, it is desired to extend the finite-time solutions of game theoretic problems to the steady-state (or infinite horizon) condition. For linear-quadratic problems, the detectability and stabilizablity of (C, A, B) ensures the existence of a unique, positive semi-definite, stabilizing solution of the Riccati equation in steady-state. Unfortunately, no such conditions exist for game-theoretic problems, except in special case where the A matrix is asymptotically stable (Green and Limebeer 1995, Lemma 3.7.3).

On the other hand, when it has possible to find a steady-state solution to the disturbance attenuation problem, it has been shown (Green and Limebeer 1995) that this solution will be in the form of the estimator given by (9.26) with Π found via the solution of the algebraic Riccati equation:

$$0 = A^T\Pi + \Pi A + \frac{1}{\gamma}\Pi\hat{F}_1 M_2 \hat{F}_1^T \Pi + C^T(\hat{H}_1^T Q_1 \hat{H}_1^T - \gamma N_1^T R_1 N_1 - \gamma V^{-1})C$$

9.2.4 Finding the Limiting Solution

The solution of the fault detection filter problem exists at the limit of the game solution when γ is taken to zero. Finding the solution or even showing that it exists in the limit, however, is not a straightforward matter. In both versions of the game Riccati equation, (9.15) and (9.17), there are terms which go to infinity as γ goes to zeroA similar limit has been studied in the linear quadratic regulator problem (Kwakernaak and Sivan 1972) where the cost function is always non-negative. These results are not directly applicable here since the game cost can be either positive or negative. Furthermore, it is well known (Doyle et al. 1989) that for game Riccati equations, γ has a greatest upper bound *gam̄ma* at or below which the equation has no positive-definite solution. When $\gamma \leq$ *gam̄ma* any number of different phenomena can occur, for example, eigenvalues on the imaginary axis, which make positive-definite solutions impossible.

By decreasing the noise weighting V to zero along with γ, that is, $V \to 0$ as $\gamma \to 0$, we can find solutions to (9.15) and (9.17) for smaller and smaller γ. While solutions are obtainable for a range of $\gamma \in (0, \infty]$ where $\gamma = \infty$ corresponds to the Kalman filter, what is needed is a solution for when $\gamma = 0$. The solution follows from a pair of techniques from singular optimal control theory which are discussed in the next section.

9.3 The Limiting Case Solution via Singular Optimal Control Techniques

9.3.1 Conditions for Game Cost Non-Positivity: A Game LMI

In this section, we will find sufficient conditions for the non-positivity of the game cost. These conditions fall out after we manipulate the cost function and then set $\hat{x}$ to its optimal strategy found in Section 9.1. The game cost then becomes a single quadratic form:

$$J(\hat{x}, x(t_0), \hat{\mu}_2, v) = \int_{t_0}^{t_1} \xi^T \overline{W} \xi dt$$

where ξ is some linear vector combination of the game players. The non-negativity of the cost hinges on the sign definiteness of $\overline{W}$, giving rise to a linear matrix inequality. This technique was first seen in the singular optimal control theory (Bell and Jacobsen 1973) and (Clements and Anderson 1978) and the derivation seen here follows in that vein.

We begin with the cost function as given by (9.7). Note that the $(x - \hat{x})$ terms have been combined:

$$\begin{aligned} J &= -\|x(t_0) - \hat{x}_0\|^2_{\Pi_0} + \\ &\int_{t_0}^{t_1} \left[\|(x-\hat{x})\|^2_{C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1)C} - \gamma\|\hat{\mu}_2\|^2_{M_2^{-1}} - \gamma\|y - Cx\|^2_{V^{-1}}\right] dt \end{aligned} \quad (9.27)$$

We now append the dynamics of the system to (9.27) through the Lagrange Multiplier $(x - \hat{x})^T\Pi$:

$$\begin{aligned} J &= -\|x(t_0) - \hat{x}_0\|^2_{\Pi_0} + \int_{t_0}^{t_1} \Big[\|(x-\hat{x})\|^2_{C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1)C} \\ &\quad - \gamma\|\hat{\mu}_2\|^2_{M_2^{-1}} - \gamma\|y - Cx\|^2_{V^{-1}} + (x-\hat{x})^T\Pi(Ax + \hat{F}_1\hat{\mu}_2 - \dot{x})\Big] dt \end{aligned}$$

Add and subtract to (9.8) the terms $(x - \hat{x})^T\Pi A\hat{x}$ and $(x - \hat{x})^T\Pi\dot{\hat{x}}$. Collect terms to get:

$$\begin{aligned} J &= -\|x(t_0) - \hat{x}_0\|^2_{\Pi_0} + \int_{t_0}^{t_1} \Big\{\|(x-\hat{x})\|^2_{\Pi A + C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1)C} \\ &\quad -\gamma\|\hat{\mu}_2\|^2_{M_2^{-1}} - \gamma\|y - Cx\|^2_{V^{-1}} \\ &\quad + (x-\hat{x})^T\Pi\hat{F}_1\hat{\mu}_2 - (x-\hat{x})^T\Pi(\dot{x} - \dot{\hat{x}}) + (x-\hat{x})^T\left[\Pi A\hat{x} - \Pi\dot{\hat{x}}\right]\Big\} dt \end{aligned} \quad (9.28)$$

Note, we have moved ΠA into the weighting of $\|(x - \hat{x})\|^2$. More terms will appear in the weighting of $\|(x-\hat{x})\|^2$ as we manipulate the cost function. Now, integrate $(x-\hat{x})^T\Pi(\dot{x}-\dot{\hat{x}})$ by parts:

$$\begin{aligned} J &= -\|x(t_0) - \hat{x}_0\|^2_{\Pi_0 - \Pi(t_0)} - \|x(t_1) - \hat{x}(t_1)\|^2_{\Pi(t_1)} + \\ &\quad + \int_{t_0}^{t_1} \Big\{\|(x-\hat{x})\|^2_{\dot{\Pi} + \Pi A + C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma N_1^T R_1 N_1)C} \\ &\quad - \gamma\|\hat{\mu}_2\|^2_{M_2^{-1}} - \gamma\|y - Cx\|^2_{V^{-1}} + (x-\hat{x})^T\Pi\hat{F}_1\hat{\mu}_2 \\ &\quad + (x-\hat{x})^T\left[\Pi A\hat{x} - \Pi\dot{\hat{x}}\right] + (\dot{x} - \dot{\hat{x}})^T\Pi(x - \hat{x})\Big\} dt \end{aligned} \quad (9.29)$$

Substitute the state equation for $\dot{x}$ (9.6) and add and subtract $\hat{x}^T A^T \Pi(x - \hat{x})$:

$$
\begin{aligned}
J \;=\; & -\|x(t_0)-\hat{x}_0\|^2_{\Pi_0-\Pi(t_0)} - \|x(t_1)-x(\hat{t}_1)\|^2_{\Pi(t_1)} \\
& + \int_{t_0}^{t_1} \Big\{ \|(x-\hat{x})\|^2_{\dot{\Pi}+\Pi A+A^T\Pi+C^T(\hat{H}_1^T Q_1 \hat{H}_1-\gamma N_1^T R_1 N_1)C} \\
& \quad - \gamma\|\hat{\mu}_2\|^2_{M_2^{-1}} - \gamma\|y-Cx\|^2_{V^{-1}} \\
& \quad + (x-\hat{x})^T \Pi \hat{F}_1 \hat{\mu}_2 + \hat{\mu}_2^T \hat{F}_1^T \Pi (x-\hat{x})^T \\
& \quad + (x-\hat{x})^T \left[-\Pi\dot{\hat{x}} + \Pi A\hat{x}\right] + \left[-\Pi\dot{\hat{x}} + \Pi A\hat{x}\right]^T (x-\hat{x}) \Big\}\, dt
\end{aligned} \tag{9.30}
$$

We are now going to rewrite the $\|y-Cx\|^2_{V^{-1}}$ term by adding and subtracting $C\hat{x}$ inside of the term so that it reads $\|(y-C\hat{x})-C(x-\hat{x})\|^2_{V^{-1}}$. Expand this quadratic term out and collect terms so that we end up with:

$$
\begin{aligned}
J \;=\; & -\|x(t_0)-\hat{x}_0\|^2_{\Pi_0-\Pi(t_0)} - \|x(t_1)-\hat{x}(t_1)\|^2_{\Pi(t_1)} \\
& + \int_{t_0}^{t_1} \Big\{ \|(x-\hat{x})\|^2_{\dot{\Pi}+\Pi A+A^T\Pi+C^T(\hat{H}_1^T Q_1 \hat{H}_1-\gamma N_1^T R_1 N_1-\gamma V^{-1})C} \\
& \quad - \gamma\|\hat{\mu}_2\|^2_{M_2^{-1}} - \gamma\|y-C\hat{x}\|^2_{V^{-1}} + (x-\hat{x})^T \Pi \hat{F}_1 \hat{\mu}_2 + \hat{\mu}_2^T \hat{F}_1^T \Pi (x-\hat{x})^T \\
& \quad + (x-\hat{x})^T \left[-\Pi\dot{\hat{x}} + \Pi A\hat{x} + \gamma C^T V^{-1}(y-C\hat{x})\right] \\
& \quad - \left[\Pi\dot{\hat{x}} + \Pi A\hat{x} + \gamma C^T V^{-1}(y-C\hat{x})\right]^T (x-\hat{x}) \Big\}\, dt
\end{aligned} \tag{9.31}
$$

Using (9.25) we can eliminate a pair of terms in (9.31). We are then left with a quadratric in the form:

$$
J = \int_{t_0}^{t_1} \xi^T \overline{W} \xi dt - \|x(t_0)-\hat{x}_0\|^2_{\Pi_0-\Pi(t_0)} - \|x(t_1)-x(\hat{t}_1)\|^2_{\Pi(t_1)},
$$

where

$$
\xi = \left\{ \begin{array}{c} (x-\hat{x}) \\ \hat{\mu}_2 \\ (y-C\hat{x}) \end{array} \right\}
$$

and

$$
\overline{W} \triangleq \begin{bmatrix} W(\Pi) & 0 \\ 0 & -\gamma V^{-1} \end{bmatrix} \tag{9.32}
$$

and where $W(\Pi)$ is given by

$$
W(\Pi) \triangleq \begin{bmatrix} C^T\left(\hat{H}_1^T Q_1 \hat{H}_1 - \gamma V^{-1} - \gamma N_1^T R_1 N_1\right)C + A^T\Pi + \Pi A + \dot{\Pi} & \Pi \hat{F}_1 \\ \hat{F}_1^T \Pi & -\gamma M_2^{-1} \end{bmatrix} \tag{9.33}
$$

Clearly $\overline{W}$ is negative semi-definite for $\Pi \geq 0$ such that:

$$W(\Pi) \leq 0 \tag{9.34a}$$

$$\Pi_0 - \Pi(t_0) \geq 0 \tag{9.34b}$$

$$\Pi(t_1) \geq 0 \tag{9.34c}$$

Hence, we need only pay attention to the smaller LMI, $W(\Pi)$.

For $\gamma > 0$, it is easy to see that the Riccati equation (9.15) of the previous section is embedded in (9.33). In fact, the solution of (9.15) is the solution of $W(\Pi)$ which minimizes its rank (Schumacher 1983). Thus with (9.33) and (9.25), we retain the results of the previous section, but in a form which can be easily analyzed in the limit $\gamma \to 0$. If we define $\overline{V} = \lim_{\gamma \to 0} \gamma V$, sufficient conditions for $J \leq 0$ in the limit as $\gamma \to 0$ are:

$$0 = \Pi \hat{F}_1 \tag{9.35a}$$

$$0 \geq \dot{\Pi} + A^T \Pi + \Pi A + C^T \left(\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1} \right) C \tag{9.35b}$$

along with the boundary conditions (9.34b) and (9.34c).

Condition (9.35a) shows that in the limit, the Riccati matrix Π has a non-trivial null space which contains the image of the nuisance failure map, $\hat{F}_1$. Moreover, those familiar with singular optimal control theory will recognize (9.35) as conditions seen previously for the singular LQ regulator. See, for example, (Bell and Jacobsen 1973)). This tells us, first of all, that the limiting form of this game filter is a singular filter. It is likely that similar results hold for all game theoretic ($\mathcal{H}_\infty$) filters or controllers. Secondly, singular optimal control provides a wealth of results and insights which we can apply to the analysis of this filter. This is, in fact, what we will do next.

9.3.2 A Riccati Equation for the Limiting Form of the Game Theoretic Filter

In Appendix A many components for the general fault detection filtering problem are derived using the Goh transformation. In this section, we will again use the Goh transformation on the nuisance fault input space to obtain a Riccati equation for the limiting case game filter. The existence of the solution to this equation gives the condition for the existence of the game solution in the limit. Because this Riccati Matrix must also have a non-trivial null space, we will not be able to use the solution to this Riccati equation directly in a game filter, but this matrix will prove to be useful when we look at reduced-order detection filters.

We start with the game cost for the limiting case:

$$J^* = \lim_{\gamma \to 0} J = \int_{t_0}^{t_1} \left(\|x - \hat{x}\|^2_{C^T \hat{H}_1^T Q_1 \hat{H}_1 C} - \|y - Cx\|^2_{\overline{V}^{-1}} \right) dt$$

where $\overline{V}^{-1} \triangleq \lim_{\gamma \to 0} (\gamma V)^{-1}$. Now, define a new nuisance fault vector, ρ_1 and a new state vector, α_1:

$$\rho_1 \triangleq \int_{t_0}^{t} \hat{\mu}_2 \, dt \tag{9.36}$$

$$\alpha_1 \triangleq x - \hat{F}_1 \rho_1 \equiv x - B_1 \rho_1 \tag{9.37}$$

Note that we have defined a matrix $B_1 \triangleq \hat{F}_1$. The reason for the numbered subscripts will become apparent later. Differentiating (9.37) produces a new state equation

$$\dot{\alpha}_1 = A\alpha_1 + (AB_1 - \dot{B}_1)\rho_1 \tag{9.38}$$

and a new game cost

$$\begin{aligned} J^* = \int_{t_0}^{t_1} \Big[& \|\alpha_1 - \hat{x}\|^2_{C^T \hat{H}_1^T Q_1 H_1 C} + (\alpha_1 - \hat{x})^T C^T \hat{H}_1^T Q_1 \hat{H}_1 C B_1 \rho_1 \\ & + \rho_1^T B_1^T C^T \hat{H}_1^T Q_1 \hat{H}_1 C B_1 \rho_1 - \|y - C\alpha_1\|^2_{\overline{V}^{-1}} - (y - C\alpha_1)^T \overline{V}^{-1} C B_1 \rho_1 \\ & - \rho_1^T B_1^T C^T \overline{V}^{-1} (y - C\alpha_1) - \|\rho_1\|^2_{B_1^T C^T \overline{V}^{-1} C B_1} \Big] \, dt \end{aligned} \tag{9.39}$$

Because $\hat{H}_1$ is a projector constructed so that $\hat{H}_1 C \hat{F}_1 = 0$, the cost (9.39) is simplified as:

$$\begin{aligned} J^* = \int_{t_0}^{t_1} \Big[& \|\alpha_1 - \hat{x}\|^2_{C^T \hat{H}_1^T Q_1 \hat{H}_1 C} - \|y - C\alpha_1\|^2_{\overline{V}^{-1}} - (y - C\alpha_1)^T \overline{V}^{-1} C B_1 \rho_1 \\ & - \rho_1^T B_1^T C^T \overline{V}^{-1} (y - C\alpha_1) - \|\rho_1\|^2_{B_1^T C^T \overline{V}^{-1} C B_1} \Big] \, dt. \end{aligned}$$

Now, if $B_1^T C^T \overline{V}^{-1} C B_1 > 0$, we can solve the following differential game:

$$\min_{\hat{x}} \max_{\rho_1} J^* \leq 0$$

subject to (9.38). Because of its similarity to the derivation given in Section 9.2, we do not provide the solution here. A starting point is to convert $y - C\alpha$ into $(y - C\hat{x}) + C(\alpha - \hat{x})$. The solution leads to the Riccati equation:

$$-\dot{S} = SA + A^T S + C^T \left(\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1}\right) C$$
$$+\left[S(AB_1 - \dot{B}_1) - C^T\overline{V}^{-1}CB_1\right](B_1^T C^T \overline{V}^{-1} C B_1)^{-1}\left[(AB_1 - \dot{B}_1)^T S - B_1^T C^T \overline{V}^{-1} C\right] \quad (9.40)$$

with the boundary condition:

$$S(t_0) = 0. \quad (9.41)$$

It may happen, however, that $CB_1 = 0$, which would make $B_1^T C^T \overline{V}^{-1} CB_1 = 0$ and which would invalidate our Riccati equation (9.40). The remedy to this situation is to perform the same transformation as before but on the ρ_1 input space via the recursion equations:

$$\begin{aligned}
\rho_i &= \int_{t_0}^{t} \rho_{i-1} dt \\
B_i &= AB_{i-1} - \dot{B}_{i-1} \\
\alpha_i &= x - B_i \rho_i.
\end{aligned}$$

The process stops once a B_i is found such that $CB_i \neq 0$. The game is then:

$$\begin{aligned}
\min_{\hat{x}} \max_{\rho_i} J^* = \int_{t_0}^{t_1} \Big[& \|\alpha_i - \hat{x}\|^2_{C^T \hat{H}_1^T Q_1 H_1 C} - \|y - C\alpha_i\|^2_{\overline{V}^{-1}} - (y - C\alpha_i)^T \overline{V}^{-1} C B_i \rho_i \\
& - \rho_i^T B_i^T C^T \overline{V}^{-1} (y - C\alpha_i) - \|\rho_i\|^2_{B_i^T C^T \overline{V}^{-1} C B_i} \Big] dt \quad (9.42)
\end{aligned}$$

subject to:

$$\dot{\alpha}_i = A\alpha_i + (AB_i - \dot{B}_i)\rho_i. \quad (9.43)$$

The general form of the Goh Riccati equation is then:

$$-\dot{S} = SA + A^T S + C^T (\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1}) C$$
$$+\left[S(AB_i - \dot{B}_i) - C^T \overline{V}^{-1} C B_i\right](B_i^T C^T \overline{V}^{-1} C B_i)^{-1}\left[(AB_i - \dot{B}_i)^T S - B_i^T C^T \overline{V}^{-1} C\right] \quad (9.44)$$

The following theorem shows that (9.44) is a Riccati equation for the limiting form of the game theoretic filter.

Theorem 9.1. The solution S to (9.44) satisfies the sufficient conditions for non-positivity of the game cost, that is, (9.35a) and (9.35b).

Proof. (The proof follows Bell and Jacobson (Bell and Jacobsen 1973, pg. 121). Due to its importance, we list it here.) Clearly, (9.44) implies that:

$$\dot{S} + SA + A^T S + C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1})C \leq 0, \qquad \forall t \in [t_0, t_1]. \tag{9.45}$$

which is (9.35a). Now, pre-multiply (9.44) by B_i^T and add $-\dot{B}_i^T S$ to both sides of the resulting equation to get:

$$\begin{aligned} -B_i^T \dot{S} - \dot{B}_i^T S &= B_i^T SA - \dot{B}_i^T S + B_i^T A^T S - B_i^T C^T \overline{V}^{-1} C \\ &+ B_i^T \left[S(AB_i - \dot{B}_i) - C^T \overline{V}^{-1} CB_i\right](B_i^T C^T \overline{V}^{-1} CB_i)^{-1} \left[(AB_i - \dot{B}_i)^T S - B_i^T C^T \overline{V}^{-1} C\right] \end{aligned} \tag{9.46}$$

Rearranging terms leads to a differential equation in $B_i^T S$ with (9.41) as the boundary condition:

$$\begin{aligned} -\frac{d}{dt}[B_i^T S] &= B_i^T SA \\ &+ B_i^T S(AB_i - \dot{B}_i)(B_i^T C^T \overline{V}^{-1} CB_i)^{-1} \left[(AB_i - \dot{B}_i)^T S - C^T \overline{V}^{-1} CB_i\right]. \end{aligned} \tag{9.47}$$

The solution to (9.47) given (9.41) is:

$$B_i^T(t) S(t) = 0, \qquad \forall t \in [t_0, t_1] \tag{9.48}$$

The necessary condition (9.35a) actually requires that $\hat{F}_1^T S(t) = 0$. However, $B_1 = \hat{F}_1$ and the following proposition tell us that (9.48) implies (9.35a). ◆

Proposition 9.2. Let $i \in \mathcal{N}$ be the smallest number such that $CB_i \neq 0$. Then, the solution, S, to (9.81) is such that

$$SB_j = 0, \qquad \forall j \leq i, \ \forall t \in [t_0, t_1]$$

Proof. See (Moylan and Moore 1971). The proof given there is identical to the one just used to show that $SB_i = 0$. Induction is then used to show that $SB_j = 0$ is also true for all $j < i$. ◆

9.4 An Unobservability Subspace Structure in the Limit

In this section, we return to time-invariant case and show that for these systems the solution to the fault detection filter problem as stated in Definition 9.1 also solves the problem as stated by Definition A.1. Thus, we can conclude that the limiting form of the game theoretic filter is a Beard-Jones fault detection filter.

Beard-Jones filters are constructed from invariant subspaces and so we will need to find an invariant subspace that is constructed by the game filter in order to prove our claim. This will require that we not only restrict ourselves to the time-invariant case, but also that we restrict our attention to the infinite-horizon problem. Hence, $\dot{\Pi} = 0$ and (9.35b) becomes:

$$A^T\Pi + \Pi A + C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1})C \leq 0 \tag{9.49}$$

When we specialize our analysis in this manner, we find that the required invariant subspace is the kernal of Π.

Theorem 9.3. $\operatorname{Ker}\Pi$ is a subspace which solves the fault detection filter problem

Proof. The three conditions listed by Definition A.1 are subspace inclusion, output separability and (C, A)-invariance. Condition (9.35a) clearly implies subspace inclusion. Since we are trying to detect only one fault, output separability is satisfied trivially. Thus, all that remains is to show (C, A)-invariance.

From Wonham (Wonham 1985), a necessary and sufficient condition for $\operatorname{Ker}\Pi$ to be (C, A)-invariant is that:

$$A(\operatorname{Ker}\Pi \cap \operatorname{Ker}C) \subset \operatorname{Ker}\Pi$$

Therefore, let $x \in A(\operatorname{Ker}\Pi \cap \operatorname{Ker}C)$. That is, there exists a vector ς such that:

$$x = A\varsigma \quad \text{and} \quad \Pi\varsigma = C\varsigma = 0.$$

Now consider (9.35b). If we post-multiply (9.35b) by ς we get:

$$\Pi A\varsigma \leq 0 \quad \Rightarrow \quad \varsigma^T A^T \Pi A\varsigma \leq 0.$$

Since $\Pi \geq 0$, this means that:

$$\varsigma^T A^T \Pi A \varsigma = 0.$$

which implies that:

$$\Pi A \varsigma = \Pi x = 0 \Rightarrow x \in \operatorname{Ker} \Pi$$

Therefore, $A(\operatorname{Ker} \Pi \cap \operatorname{Ker} C) \subset \operatorname{Ker} \Pi$ and so $\operatorname{Ker} \Pi$ is (C, A)-invariant. ◆

Remark 1. In practice, it is not necessary to use the limiting form of the filter. In many H_∞ designs, γ is not taken to its smallest possible value, but left at one which results in an acceptable compromise between all of the (usually competing) design objectives. The virtue of a game theoretic approach to fault detection filter design is that it provides a *knob* with which to make the filter more like a Beard-Jones filter (small γ and small V) or more like a sensor noise attenuating $\mathcal{H}_\infty$ filter (large γ and V). ◆

Remark 2. It should be noted that a Beard-Jones fault detection filter can detect all of the μ_j's. The filter that we propose here can detect only one fault. ◆

Remark 3. Lee and Gibson derive a filter for fault detection via a minimax solution in (Lee 1994). Their results are similar to ours except that they do not investigate the relationship between their filter and fault detection filters and they do not look at limiting solutions. ◆

In Section 9.1 we noted that unobservability subspaces are used in current fault detection filter design methods because they allow the designer to specify (within complex conjugate symmetry) all of the eigenvalues of the filter. Such design freedom exists with these subspaces because they include any invariant zero directions which arise out of the triple $(C, A, \hat{F}_1)$. It remains to be seen where the game theoretic filter places invariant zeros. If all of the zeros are placed in $\operatorname{Ker} \Pi$, then $\operatorname{Ker} \Pi$ would be a detection space since it would be a (C, A)-invariant subspace containing the invariant zeros. It turns out, however, that only the right-half plane and purely imaginary zeros are contained in $\operatorname{Ker} \Pi$.

Theorem 9.4. Let $\overline{\mathcal{V}}^+$ be the subspace spanned by the invariant zero directions that correspond to the invariant zeros lying in the right-half plane. Let $\overline{\mathcal{V}}^0$ be the corresponding subspace for purely imaginary zeros. The (C, A)-invariant subspace, $\operatorname{Ker}\Pi$, created by the game-theoretic fault detection filter is such that

$$\overline{\mathcal{V}}^+ \subset \operatorname{Ker}\Pi.$$

If $(A, \hat{F}_1)$ is stabilizable, then

$$\overline{\mathcal{V}}^+ + \overline{\mathcal{V}}^0 \subset \operatorname{Ker}\Pi$$

Proof. Our proof is essentially the same as the one given in (Francis 1979), though modified to fit the particulars of our problem. The arguments that we present here rely on geometric control theory, which means that we will have to spend a fair amount of time defining subspaces and mappings between these subspaces. Once this is done, however, the actual proof comes together quickly.

We begin by defining a new subspace, $\mathcal{V}^*$, the maximal $(A, \hat{F}_1)$-invariant subspace contained in $\operatorname{Ker} C$. $\mathcal{V}^*$ is the dual of the minimal (C, A)-invariant subspace $\mathcal{W}_*$ defined by Theorem A.1 and in a similar manner it can be found as the limit of an iteration (Wonham 1985):

$$\begin{aligned}\mathcal{V}_0 &= \operatorname{Ker} C \\ \mathcal{V}_{i+1} &= \operatorname{Ker} C \cap A^{-1}(\operatorname{Im}\hat{F}_1 + \mathcal{V}_i)\end{aligned}$$

The notation A^{-1} should be understood as an inverse mapping and not an inverse of the matrix A. That is:

$$A^{-1}(\operatorname{Im}\hat{F}_1 + \mathcal{V}_i) \triangleq \left\{x \in \mathcal{X} : Ax \in \operatorname{Im}\hat{F}_1 + \mathcal{V}_i\right\}$$

To be $(A, \hat{F}_1)$-invariant means that if μ were a control input then for any $x(t_0) \in \mathcal{V}^*$ there exists a matrix K such that $\mu = Kx$ and:

$$x(t) = e^{A+\hat{F}_1 K}x(t_0) \;\in\; \mathcal{V}^* \qquad \forall t \in [t_0, t_1].$$

This is not to say that we are specifying the time history of $\mu(t)$ to be a linear feedback of the states. It is just a way of illustrating the meaning of $\mathcal{V}^*$. In fact we do not need all of the space $\mathcal{V}^*$, but a portion of it. This portion, it turns out, corresponds to the invariant zeros. We define the following factor spaces:

$$\begin{aligned}\overline{\mathcal{X}} &= \mathcal{X}/(\mathcal{V}^* \cap \mathcal{W}_*)\\ \overline{\mathcal{V}} &= \mathcal{V}^*/(\mathcal{V}^* \cap \mathcal{W}_*)\end{aligned}$$

The significance of these factor spaces is through the relationship between $\overline{\mathcal{V}}$ and the $(C, A, \hat{F}_1)$ invariant zeros. If $\mathcal{M}$ is the failure input space and $K : \mathcal{M} \rightarrow \mathcal{X}$ is a feedback matrix which makes $\mathcal{V}^*$ an $(A, \hat{F}_1)$-invariant subspace, the spectrum of $A + \hat{F}_1 K$ induced on $\overline{\mathcal{V}}$ is precisely the set of invariant zeros of the triple $(C, A, \hat{F}_1)$. The invariant zero directions span $\overline{\mathcal{V}}$. Given that we are trying to prove a result about the invariant zeros, the space $\overline{\mathcal{V}}$ will clearly play a key role in our proof.

The equivalence of $\overline{\mathcal{V}}$ and the space spanned by the invariant zero directions follows from a pair of results from geometric control theory. The first, which can be found in (Morse 1973), is that the space $\mathcal{V}^* \cap \mathcal{W}_*$ is equal to the maximal controllability subspace, which we will label $\mathcal{R}^*$. $\mathcal{R}^*$ is the largest $(A, \hat{F}_1)$-invariant subspace on which the spectrum of $A + \hat{F}_1 K$ can be arbitrarily specified, hence $\mathcal{R}^* \subseteq \mathcal{V}^*$. Moreover, $\mathcal{R}^*$ is the dual to the unobservability spaces, or detection spaces, which we described earlier. The second result is that the factor space $\mathcal{V}^*/\mathcal{R}^*$, which is our space $\overline{\mathcal{V}}$, is the space spanned by the invariant zero directions. This result can be found in many places, in particular (Wonham 1985).

Define $\mathcal{V}^+$ to be the subspace of $\mathcal{V}$ on which the restriction of $A + \hat{F}_1 K$ yields eigenvalues with positive real parts. $\mathcal{V}^0$ is the corresponding space for purely imaginary eigenvalues and $\mathcal{V}^-$ the space for eigenvalues with negative real parts. Let $M : \mathcal{X} \rightarrow \overline{\mathcal{X}}$ be the canonical projection. Therefore:

$$\overline{\mathcal{V}}^+ = M\mathcal{V}^+, \qquad \overline{\mathcal{V}}^0 = M\mathcal{V}^0, \qquad \overline{\mathcal{V}}^- = M\mathcal{V}^-$$

and

$$\overline{\mathcal{V}} = \overline{\mathcal{V}}^+ + \overline{\mathcal{V}}^0 + \overline{\mathcal{V}}^-.$$

Finally, let $L : \mathcal{V} \rightarrow \mathcal{X}$ and $\overline{L} : \overline{\mathcal{V}} \rightarrow \overline{\mathcal{X}}$ be natural insertions.

To aid our understanding, we will make use of a commutative diagram. Commutative diagrams are a common tool in abstract algebra and show, pictorily, the relationships between the different subspaces and the maps which take vectors from one space to another. For this proof the corresponding commutative diagram is given by Figure 9.1.

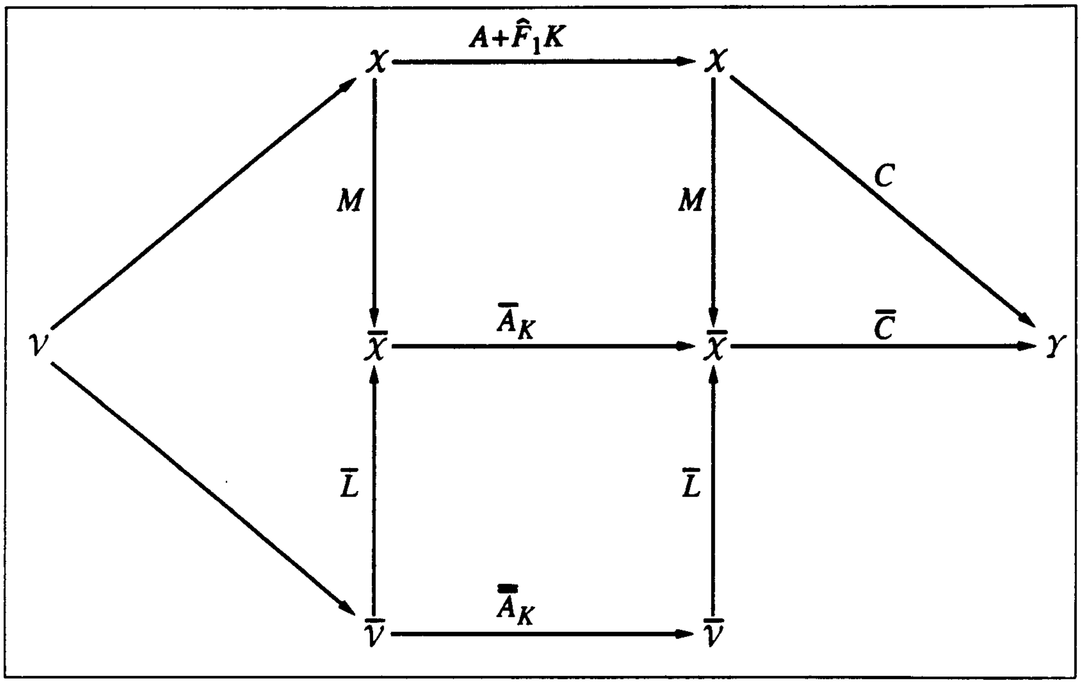

Figure 9.1: Commutative diagram for fault detection filter structure.

Through the actions of M and $\overline{L}$ on the invariant subspaces $\overline{\mathcal{X}}$ and $\overline{\mathcal{V}}$ we can infer the existence of a number of induced mappings. $\overline{A}_K : \overline{X} \rightarrow \overline{X}$ is the map induced by $A + \hat{F}_1 K$ on $\overline{\mathcal{X}}$. From Figure 9.1, $\overline{A}_K$ is related to $A + \hat{F}_1 K$ via:

$$(A + \hat{F}_1 K)M = M\overline{A}_K \tag{9.50}$$

$\overline{\overline{A}}_K$ is the restriction of $\overline{A}_K$ to $\overline{\mathcal{V}}$. Its existence is guaranteed by the $\overline{A}_K$-invariance of $\overline{V}$ and it is related to $\overline{A}_K$ by:

$$\overline{L}\,\overline{A}_K = \overline{\overline{A}}_K \overline{L} \tag{9.51}$$

Finally, the map $\overline{C}$ is the unique solution to:

$$\overline{C}M = C \tag{9.52}$$

Its existence and uniqueness is guaranteed by the fact that $(\mathcal{V}^* \cap \mathcal{W}_*) \subset \operatorname{Ker} C$.

We can now begin with the actual proof. We begin by asserting that:

$$(\mathcal{V}^* \cap \mathcal{W}_*) \subset \operatorname{Ker} \Pi. \tag{9.53}$$

We know that this is true because $\operatorname{Ker}\Pi$ is a (C, A)-invariant subspace containing the range of $\hat{F}_1$ and $\mathcal{W}_*$ is the smallest of all such subspaces. Hence, $\mathcal{W}_* \subset \operatorname{Ker}\Pi$, which implies (9.53). From (9.53) we can assert that there exists a unique symmetric matrix $\tilde{\Pi}$ such that:

$$\Pi = M^T \tilde{\Pi} M. \tag{9.54}$$

Using (9.54),(9.50), and (9.52), we can rewrite (9.49) as:

$$M^T \left[\overline{A}_K^T \tilde{\Pi} + \tilde{\Pi}\overline{A}_K + \overline{C}^T (\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1})\overline{C}\right] M \leq 0$$

Because M is a canonical projector, it has a right inverse which means that we can rework the above inquality into:

$$\overline{A}_K^T \tilde{\Pi} + \tilde{\Pi}\overline{A}_K + \overline{C}^T (\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1})\overline{C} \leq 0 \tag{9.55}$$

We need now need to go one step further and consider the system restricted to the subspace $\overline{\mathcal{V}}$. Pre-multiply (9.55) by $\overline{L}^T$ and post-multiply by $\overline{L}$. Since $\overline{L}$ is insertion map of a space which lies in $\operatorname{Ker} C$, it follows that $\overline{C}\,\overline{L} = 0$. Thus, from (9.51) we can rewrite (9.55) as:

$$\overline{\overline{A}}_K^T \overline{L}^T \tilde{\Pi} \overline{L} + \overline{L}^T \tilde{\Pi} \overline{L}\, \overline{\overline{A}}_K \leq 0 \tag{9.56}$$

Now, let λ_j be the jth eigenvalue of $\overline{\overline{A}}_K$ such that $\operatorname{Re}\lambda > 0$ and let $z_{j_{i_0}}$, $j_i = 1 \dots j_{i_0} \dots \alpha_j$ be one of the corresponding generalized eigenvectors. Here α_j is the algebraic multiplicity of λ_j. Pre-multiply (9.56) by $z^*_{j_{i_0}}$, the conjugate transpose of $z_{j_{i_0}}$, and post-multiply by $z_{j_{i_0}}$ to get:

$$z^*_{j_{i_0}} \left(\overline{\overline{A}}_K^T \overline{L}^T \tilde{\Pi} \overline{L} + \overline{L}^T \tilde{\Pi} \overline{L}\, \overline{\overline{A}}_K\right) z_{j_{i_0}} = (2\operatorname{Re}\lambda)\, z^*_{j_{i_0}} \overline{L}^T \tilde{\Pi} \overline{L} z_{j_{i_0}} \leq 0.$$

The latter inequality implies that $\tilde{\Pi}\overline{L}z_{j_{i_0}} = 0$ since $\tilde{\Pi}$ is positive semi-definite. As stated earlier, the eigenvalues of $\overline{\overline{A}}_K$ are the invariant zeros of the triple $(C, A, \hat{F}_1)$, meaning that $\overline{L}z_{j_{i_0}}$ is the invariant zero direction. We have just shown that this direction lies in the kernal of $\tilde{\Pi}$ which is sufficient to claim that it lies in the kernal of Π itself. Since $z_{j_{i_0}}$ was chosen arbitrarily out of the set of generalized eigenvalues, this holds for all z_{j_i} in the set. Since λ_j was chosen arbitrarily out of the set of unstable eigenvalues of $\overline{\overline{A}}_K$, this holds for all such eigenvalues. This proves the first half of our theorem.

To prove the second half of our theorem we need to make the additional assumption that $(A, \hat{F}_1)$ is stabilizable. This new assumption is fairly benign and was also made by (Banavar and Speyer 1991). $(A, \hat{F}_1)$ stabilizable implies that $(A + \hat{F}_1 K, \hat{F}_1)$ and $(\overline{A}_K, M\hat{F}_1)$ are stabilizable. The latter is proven in (Wonham 1985)). Now let $\lambda_k = j\omega$ be an eigenvalue of $\overline{\overline{A}}_K$ and let $z_{j_{k_0}}$, $j_k = 1 \dots j_{k_0} \dots \alpha_k$ be one of the corresponding generalized eigenvectors. Pre-multiplying (9.55) by $z^*_{j_{k_0}}\overline{L}^T$ and post-multiplying by $\overline{L}z_{j_{k_0}}$ leads to:

$$z^*_{j_{k_0}}\overline{L}^T\left(\overline{A}_K\tilde{\Pi} + \tilde{\Pi}\overline{A}_K\right)\overline{L}z_{j_{k_0}} = (2\,\mathrm{Re}\,\lambda)\, z^*_{j_{k_0}}\overline{L}^T\tilde{\Pi}\overline{L}z_{j_{k_0}} = 0$$

which implies

$$z^*_{j_{k_0}}\overline{L}^T\left(\overline{A}_K\tilde{\Pi} + \tilde{\Pi}\overline{A}_K\right) = z^*_{j_{k_0}}\overline{L}^T\tilde{\Pi}(-\lambda I + \overline{A}_K) = 0$$

We also know that $\tilde{\Pi}\hat{F}_1 M = 0$ since $\Pi\hat{F}_1 = 0$. Hence we can augment the above equation to read:

$$z^*_{j_{k_0}}\overline{L}^T\tilde{\Pi}\left[\overline{A}_K - \lambda I,\ \hat{F}_1 M\right] = 0$$

This implies that $z^*_{j_{k_0}}\overline{L}^T\tilde{\Pi} = 0$, since the stabilizability assumption implies $[\overline{A}_K - \lambda I,\ \hat{F}_1 M]$ is full rank. From this we can conclude that $\overline{L}z_{j_{k_0}} \in \mathrm{Ker}\,\tilde{\Pi}$ which, by using the same arguments as before, leads to the conclusion that the invariant zero directions corresponding to the purely imaginary zeros lie in the kernal of Π. ◆

Even though invariant zeros will not destabilize the game-theoretic filter as was just shown, it is still possible that a left-half plane zero could be in a location which is undesirable. This potential shortcoming is mitigated somewhat by the fact that zeros are rare for non-square systems.

9.5 Fault Detection with the Limiting Form of the Game Theoretic Filter

In this section, we will show that a reduced-order fault detector can be derived from the limiting form of the game theoretic filter. The results from this section are more easily applied to time-invariant systems, but we will give an overview of how to apply these results to time-varying systems.

The reduced-order filter falls out from the fact that positive semi-definite, symmetric matrices such as Π always have non-singular, transformations - say Γ - that are orthonormal ($\Gamma^T\Gamma = I$) and that convert the matrix into the form:

$$\Gamma\Pi\Gamma^T = \begin{bmatrix} \overline{\Pi} & 0 \\ 0 & 0 \end{bmatrix}, \tag{9.57}$$

where $\overline{\Pi}$ is positive definite. From (9.57), we can derive transformations on system matrices which will allow us to factor out the portion of the state-space which corresponds to $\operatorname{Ker}\Pi$. First define:

$$C\Gamma^T = \begin{bmatrix} C_1 & C_2 \end{bmatrix}, \qquad \Gamma A\Gamma^T = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \qquad \Gamma\hat{F}_1 = \begin{bmatrix} F_{11} \\ F_{12} \end{bmatrix}$$

Because $\Pi\hat{F}_1 = 0$ implies $\Gamma\Pi\hat{F}_1 = 0$, we can immediately conclude that:

$$\Gamma\Pi\Gamma^T\Gamma\hat{F}_1 = \begin{bmatrix} \overline{\Pi} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} F_{11} \\ F_{22} \end{bmatrix} = \overline{\Pi}F_{11} = 0.$$

Which, since $\overline{\Pi}$ is positive-definite, implies:

$$F_{11} = 0.$$

Now, using Γ we can partition the state-space as:

$$\hat{\eta} = \begin{Bmatrix} \hat{\eta}_1 \\ \hat{\eta}_2 \end{Bmatrix} = \Gamma\hat{x}.$$

Pre-multiply (9.25) by Γ and make use of the identity $\Gamma^T\Gamma = I$ to get:

$$(\Gamma\Pi\Gamma^T)\dot{\hat{\eta}} = (\Gamma\Pi\Gamma^T)(\Gamma A\Gamma^T)\hat{\eta} + \Gamma C^T\overline{V}^{-1}(y - C\Gamma^T\hat{\eta}) \tag{9.58}$$

The transformed filter equation (9.58) is seen to be:

$$\begin{bmatrix} \overline{\Pi} & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} \dot{\hat{\eta}}_1 \\ \dot{\hat{\eta}}_2 \end{Bmatrix} = \begin{bmatrix} \overline{\Pi} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{Bmatrix} \hat{\eta}_1 \\ \hat{\eta}_2 \end{Bmatrix} + \begin{Bmatrix} C_1^T \\ C_2^T \end{Bmatrix} \overline{V}^{-1} \left(y - \begin{bmatrix} C_1 & C_2 \end{bmatrix} \begin{Bmatrix} \hat{\eta}_1 \\ \hat{\eta}_2 \end{Bmatrix} \right) \tag{9.59}$$

From (9.59) we get a dynamic equation for $\hat{\eta}_1$:

$$\overline{\Pi}\dot{\hat{\eta}}_1 = \overline{\Pi}A_{11}\hat{\eta}_1 + \overline{\Pi}A_{12}\hat{\eta}_2 + C_1^T\overline{V}^{-1}(y - C_1\hat{\eta}_1 - C_2\hat{\eta}_2) \tag{9.60}$$

and a static equation for $\hat{\eta}_2$:

$$\hat{\eta}_2 = (C_2^T\overline{V}^{-1}C_2)^{-1}C_2^T\overline{V}^{-1}(y - C_1\hat{\eta}_1). \tag{9.61}$$

Define

$$K \triangleq (C_2^T\overline{V}^{-1}C_2)^{-1}C_2^T\overline{V}^{-1} \tag{9.62}$$

so that the substitution of (9.62) and (9.61) into (9.60) gives us an estimator for $\hat{\eta}_1$:

$$\dot{\hat{\eta}}_1 = A_{11}\hat{\eta}_1 + \left[\overline{\Pi}^{-1}C_1^T\overline{V}^{-1}(I - C_2K) + A_{12}K\right](y - C_1\hat{\eta}_1). \tag{9.63}$$

To see that the reduced-order estimator (9.63) is unaffected by the nuisance fault $\hat{\mu}_2$, we will derive the error equation for the reduced-order filter. Define:

$$\eta = \begin{Bmatrix} \eta_1 \\ \eta_2 \end{Bmatrix} \triangleq \Gamma x, \qquad e_1 \triangleq \hat{\eta}_1 - \eta_1, \qquad e_2 \triangleq \hat{\eta}_2 - \eta_2$$

We begin by premultiplying the dynamic equation (9.6) by the Riccati matrix Π. Since $\Pi\hat{F}_1 = 0$, we get:

$$\Pi\dot{x} = \Pi Ax.$$

This can be pre-multiplied by Γ and manipulated into:

$$\begin{bmatrix} \overline{\Pi} & 0 \\ 0 & 0 \end{bmatrix} \begin{Bmatrix} \dot{\eta}_1 \\ \dot{\eta}_2 \end{Bmatrix} = \begin{bmatrix} \overline{\Pi} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{Bmatrix} \eta_1 \\ \eta_2 \end{Bmatrix}. \tag{9.64}$$

As with the estimator equation, (9.64) shows that only a portion of the state-space possesses dynamics:

$$\overline{\Pi}\dot{\eta}_1 = \overline{\Pi}A_{11}\eta_1 + \overline{\Pi}A_{12}\eta_2 \tag{9.65}$$

Using (9.65) to get an error equation would leave terms in η_2 or e_2. In anticipation of this, we transform the measurement equation:

$$y = Cx + v = C\Gamma^T\Gamma x + v = C_1\eta_1 + C_2\eta_2 + v \tag{9.66}$$

and use (9.61) to solve for e_2:

$$e_2 = (C_2^T\overline{V}^{-1}C_2)^{-1}(C_2^T\overline{V}^{-1}C_1e_1 + C_2^T\overline{V}^{-1}v) = K(C_1e_1 - v) \tag{9.67}$$

Subtract (9.65) from (9.60) and substitute (9.66) for y:

$$\overline{\Pi}\dot{e}_1 = \overline{\Pi}A_{11}e_1 + \overline{\Pi}A_{12}e_2 + C_1^T\overline{V}^{-1}C_1e_1 + C_1^T\overline{V}^{-1}C_2e_2 + C^T\overline{V}^{-1}v$$

Using (9.67) and collecting terms, we can turn the previous equation into:

$$\begin{aligned} \dot{e}_1 &= \left[A_{11} - \overline{\Pi}^{-1}C_1^T\overline{V}^{-1}(I - C_2K)C_1 - A_{12}KC_1\right]e_1 \\ &\quad + \left[\overline{\Pi}^{-1}C_1^T\overline{V}^{-1}(I - C_2K) + A_{12}K\right]v. \end{aligned} \tag{9.68}$$

Note that the nuisance fault, $\hat{\mu}_2$, appears nowhere in the estimator (9.63) nor in the error equation (9.68). Thus, in the limit, we get a reduced-order estimator completely uninfluenced by the nuisance faults. The term $(C_2^T\overline{V}^{-1}C_2)^{-1}$ appears in various places in the reduced-order estimator. This inverse will always exist since $\overline{V}$ is positive definite and since the assumption of (C, A) observability guarantees that C_2 will have full column rank.

Remark 4. The reduced-order filter derived here is similar to the residual generator derived by Massoumnia, *et al.* in (Massoumnia et al. 1989). An important difference, however, is that Massoumnia begins his design process by factoring out the reachable space of the nuisance faults. As a result, he has the freedom to use any kind of filter design technique for the lower dimensional state-space. The trade-off, however, is that the system reduction in Massoumnia's filter is sensitive to the inexactness of the plant model. Variations in the plant will change the reachable subspace and may, as a result, degrade the performance of the reduced-order detector. In the game filter, the order reduction comes at the end of the design process. Thus, there is no design freedom left to tune the reduced-order filter, but the game formulation used to obtain the filter makes it possible to account for model uncertainties.

The Goh transformation and corresponding Riccati equation greatly extend our ability to analyze the reduced-order estimator. In fact with the Goh Riccati equation we can show that there always exists a stabilizing solution for the reduced order estimator. Applying the transformation Γ to (9.44), we get:

$$\begin{aligned} -\Gamma\dot{S}\Gamma^T &= \Gamma S\Gamma^T\Gamma A\Gamma^T + \Gamma A^T\Gamma^T\Gamma S\Gamma^T \\ &\quad + \Gamma C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1})C\Gamma^T + \bar{\Gamma}\left(B_i^T C^T \overline{V}^{-1} C B_i\right)^{-1}\bar{\Gamma}^T \end{aligned}$$

where, for notational convenience, $\bar{\Gamma}$ is defined as

$$\bar{\Gamma} = \left[\Gamma S\Gamma^T\left(\Gamma A\Gamma^T\Gamma B_i - \Gamma\dot{B}_i\right) - \Gamma C^T\overline{V}^{-1}C\Gamma^T\Gamma B_i\right]$$

Define:

$$\Gamma B_i = \begin{bmatrix} B_{11} \\ B_{12} \end{bmatrix}.$$

As in section 9.4, the necessary condition $SB_i = 0$ will lead to $B_{11} = 0$ since $\Gamma S\Gamma^T\Gamma B_i = 0 \quad \Rightarrow \quad \overline{S}B_{11} = 0$ and $\overline{S}$ is positive-definite. Also, if we carry the transformation through, a number of terms fall out because the projector $\hat{H}_1$ has been constructed so that:

$$\begin{aligned} \hat{H}_1 C B_i = 0 \quad &\Rightarrow \quad \hat{H}_1 C\Gamma^T\Gamma B_i = 0 \\ &\Rightarrow \quad \begin{bmatrix} \hat{H}_1 C_1 & \hat{H}_1 C_2 \end{bmatrix}\begin{bmatrix} 0 \\ B_{12} \end{bmatrix} = 0 \\ &\Rightarrow \quad \hat{H}_1 C_2 B_{12} = 0 \end{aligned} \tag{9.69}$$

We we show later that B_i can always be augmented so that B_{12} is an invertible square matrix. Hence (9.69) implies:

$$\hat{H}_1 C_2 = 0. \tag{9.70}$$

Using (9.70) and working through all of the transformations leads to:

$$\begin{bmatrix} -\dot{\overline{S}} & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} \overline{S}A_{11} & \overline{S}A_{12} \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} A_{11}^T\overline{S} & 0 \\ A_{12}^T\overline{S} & 0 \end{bmatrix}$$

$$+\left(\begin{bmatrix} \overline{S}A_{12}B_{12}-\overline{S}\dot{B}_{11} \\ 0 \end{bmatrix}-\begin{bmatrix} C_1^T\overline{V}^{-1}C_2B_{12} \\ C_2^T\overline{V}^{-1}C_2B_{12} \end{bmatrix}\right)(B_{12}^TC_2^T\overline{V}^{-1}C_2B_{12})^{-1}$$
$$\times\left(\begin{bmatrix} B_{12}^TA_{12}\overline{S}-\dot{B}_{11}\overline{S} & 0 \end{bmatrix}-\begin{bmatrix} B_{12}^TC_2^T\overline{V}^{-1}C_1^T & B_{12}^TC_2^T\overline{V}^{-1}C_2^T \end{bmatrix}\right)$$
$$+\begin{bmatrix} C_1^T(\hat{H}_1^TQ_1\hat{H}_1-\overline{V}^{-1})C_1 & -C_1^T\overline{V}^{-1}C_2 \\ -C_2^T\overline{V}^{-1}C_1 & -C_2^T\overline{V}^{-1}C_2 \end{bmatrix} \tag{9.71}$$

From (9.71) we get three equations:

$$-\dot{\overline{S}} = C_1^T(\hat{H}_1^TQ_1\hat{H}_1-\overline{V}^{-1})C_1+\overline{S}A_{11}+A_{11}^T\overline{S}$$
$$+\left(\overline{S}A_{12}B_{12}-\overline{S}\dot{B}_{11}-C_1^T\overline{V}^{-1}C_2B_{12}\right)\left(B_{12}^TC_2^T\overline{V}^{-1}C_2B_{12}\right)^{-1}$$
$$\times\left(\overline{S}A_{12}B_{12}-\overline{S}\dot{B}_{11}-C_1^T\overline{V}^{-1}C_2B_{12}\right)^T \tag{9.72}$$
$$0 = -C_1^T\overline{V}^{-1}C_2+\overline{S}A_{12}-\left(\overline{S}A_{12}B_{12}-\overline{S}\dot{B}_{11}-C_1^T\overline{V}^{-1}C_2^TB_{12}\right)$$
$$\times\left(B_{12}^TC_2^T\overline{V}^{-1}C_2B_{12}\right)^{-1}B_{12}^TC_2^T\overline{V}^{-1}C_2 \tag{9.73}$$
$$0 = -C_2^T\overline{V}^{-1}C_2+C_2^T\overline{V}^{-1}C_2B_{12}\left(B_{12}^TC_2^T\overline{V}^{-1}C_2B_{12}\right)^{-1}B_{12}^TC_2^T\overline{V}^{-1}C_2. \tag{9.74}$$

However, if we post-multiply (9.74) by B_{12} and cancel terms we obtain the identity $0 = 0$. If we post-multiply (9.73) by B_{12} we obtain:

$$0=\overline{S}\dot{B}_{11} \quad\Rightarrow\quad \dot{B}_{11}=0. \tag{9.75}$$

Thus, we need only (9.72), which thanks to (9.75) can be simplified to:

$$-\dot{\overline{S}} = C_1^T(\hat{H}_1^TQ_1\hat{H}_1-\overline{V}^{-1})C_1+\overline{S}A_{11}+A_{11}^T\overline{S}+\left(\overline{S}A_{12}B_{12}-C_1^T\overline{V}^{-1}C_2B_{12}\right)$$
$$\times\left(B_{12}^TC_2^T\overline{V}^{-1}C_2B_{12}\right)^{-1}\left(\overline{S}A_{12}B_{12}-C_1^T\overline{V}^{-1}C_2B_{12}\right)^T. \tag{9.76}$$

Now if i=1, then $B_i = \hat{F}_1$ and the rank of $\hat{F}_1$ equals the dimension of the kernal of S. $B_{12} = F_{12}$ will then be square and, moreover, it will be invertible since $\hat{F}_1$ was assumed monic. Given this, we can simplify (9.76) to:

$$-\dot{\overline{S}} = C_1^T\left(\hat{H}_1^TQ_1\hat{H}_1-\overline{V}^{-1}\right)C_1+\overline{S}A_{11}+A_{11}^T\overline{S}$$
$$+\left(\overline{S}A_{12}-C_1^T\overline{V}^{-1}C_2\right)\left(C_2^T\overline{V}^{-1}C_2\right)^{-1}\left(\overline{S}A_{12}-C_1^T\overline{V}^{-1}C_2\right)^T \tag{9.77}$$
$$\overline{S}(t_0) = 0 \tag{9.78}$$

where the boundary condition comes from (9.41). This leads us to the key result of this section.

Theorem 9.5. The solution $\overline{S}$ to (9.77) gives a stabilizing solution for the reduced-order estimator (9.63).

Proof. Using the same transformation to derive both (9.77) and (9.63) will ensure that $\overline{S}$ is of proper dimension for (9.63). Substitute $\overline{S}$ into (9.63) directly for $\overline{\Pi}$. The resulting estimator is:

$$\dot{\hat{\eta}}_1 = \left(A_{11} - \left[\overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K) + A_{12}K\right]C_1\right)\hat{\eta}_1 + \left[\overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K) + A_{12}K\right]y.$$

where $K \triangleq (C_2^T\overline{V}^{-1}C_2)^{-1}C_2^T\overline{V}^{-1}$. Clearly, the stability of the estimator depends upon the closed-loop state matrix, $(A_{11} - [\overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K) + A_{12}K]C_1)$. Now, if we go back to (9.77), multiply out the quadratic, and use the definition for K, we get:

$$\begin{aligned} -\dot{\overline{S}} &= \overline{S}(A_{11} - A_{12}KC_1) + (A_{11} - A_{12}KC_1)^T\overline{S} \\ &\quad + C_1^T\left[\hat{H}_1^TQ_1\hat{H}_1 - \overline{V}^{-1}(I - C_2K)\right]C_1 + \overline{S}A_{12}(C_2^T\overline{V}^{-1}C_2)^{-1}A_{12}^T\overline{S}. \end{aligned} \tag{9.79}$$

If we add and subtract $C_1^T\overline{V}^{-1}(I - C_2K)C_1$ to (9.79) and rearrange terms we get:

$$\begin{aligned} -\dot{\overline{S}} &= \overline{S}\left[A_{11} - A_{12}KC_1 - \overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K)C_1\right] \\ &\quad + \left[A_{11} - A_{12}KC_1 - \overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K)C_1\right]^T\overline{S} \\ &\quad + C_1^T\left[\hat{H}_1^TQ_1\hat{H}_1 + \overline{V}^{-1}(I - C_2K)\right]C_1 + \overline{S}A_{12}(C_2^T\overline{V}^{-1}C_2)^{-1}A_{12}^T\overline{S}. \end{aligned} \tag{9.80}$$

Note that $C_1^T\overline{V}^{-1}(I - C_2K)C_1$ is symmetric. (9.80) implies:

$$\begin{aligned} &\dot{\overline{S}} + \overline{S}\left[A_{11} - A_{12}KC_1 - \overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K)C_1\right] \\ &\quad + \left[A_{11} - A_{12}KC_1 - \overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K)C_1\right)^T\overline{S} \le 0, \end{aligned}$$

which by Lyapunov's direct method (Brogan 1991) implies that

$$A_{11} - A_{12}KC_1 - \overline{S}^{-1}C_1^T\overline{V}^{-1}(I - C_2K)C_1$$

is stable. For time-invariant systems, this implies that the closed-loop eigenvalues lie in the open left-half plane.

What happens, however, when $i > 1$ and $\dim(\operatorname{Ker} S) > \operatorname{Rank} B_i$? The matrix B_{12} will no longer be square and the reduced-order Riccati equation will be stuck in the form of (9.76) which is not the same as what is needed in the proof for stability (9.77). It would seem that we cannot guarantee stability in the general case.

It turns out, however, that by augmenting the failure map in the original problem statement, we can always convert the reduced-order Riccati equation into the desired form (9.77). The necessary augmentation turns out to be:

$$\overline{F}_1 = \begin{bmatrix} B_i & B_{i-1} & \dots & B_1 \end{bmatrix}$$

The new game problem for the limiting case is:

$$\begin{aligned}
\min_{\hat{x}} \max_{\overline{\mu}_2} J^* \;=\; & \int_{t_0}^{t_1} \Big[\|x-\hat{x}\|^2_{C^T \hat{H}_1^T Q_1 \hat{H}_1 C} + (x-\hat{x})^T C^T \hat{H}_1^T Q_1 \hat{H}_1 C \overline{F}_1 \overline{\mu}_2 \\
& + \|\overline{\mu}_2\|^2_{\overline{F}_1^T C^T \hat{H}_1^T Q_1 \hat{H}_1 C \overline{F}_1} - \|y - Cx\|^2_{\overline{V}^{-1}} - (y-Cx)^T \overline{V}^{-1} C \overline{F}_1 \overline{\mu}_2 \\
& - \overline{\mu}_2^T \overline{F}_1^T C^T \overline{V}^{-1} (y - Cx) - \|\overline{\mu}_2\|^2_{\overline{F}_1^T C^T \overline{V}^{-1} C \overline{F}_1} \Big] dt
\end{aligned}$$

subject to:

$$\dot{x} = Ax + \overline{F}_1 \overline{\mu}_2$$

where $\overline{\mu}_2$ is the augmented failure signal which has as many inputs as there are columns in $\overline{F}_1$. Note, that here we have gone back to the pre-transformed problem where the state is x, not α_i. We will show that this new problem leads to a Riccati equation which is equivalent to (9.44). In this equation, however, the reduced-order version is easily seen to reduce to the desired form (9.77). The equivalence of the two equations then implies that the same reduced form holds for both.

The augmented failure map, $\overline{F}_1$ is such that $C\overline{F}_1 \neq 0$, so the transformation process converges after one iteration. The solution to this game leads to a Goh Riccati equation:

$$\begin{aligned}
-\dot{S} \;=\; & SA + A^T S + C^T(\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1})C \\
& + \left[S(A\overline{F}_1 - \dot{\overline{F}}_1) - C^T \overline{V}^{-1} C \overline{F}_1 \right] (\overline{F}_1^T C^T \overline{V}^{-1} C \overline{F}_1)^{-1} \\
& \times \left[(A\overline{F}_1 - \dot{\overline{F}}_1)^T S - \overline{F}_1^T C^T \overline{V}^{-1} C \right] \qquad (9.81)
\end{aligned}$$

with a boundary condition given by (9.41). The solution, S, to (9.81) is such that

$$\dim(\operatorname{Ker}\overline{S}) = \operatorname{Rank}\overline{F}_1.$$

Hence, after the transformation and defining:

$$\begin{bmatrix} \overline{F}_{11} \\ \overline{F}_{12} \end{bmatrix} = \Gamma\overline{F}_1,$$

the reduced-order Riccati equation:

$$\begin{aligned} -\dot{\overline{S}} &= C_1^T\left(\hat{H}_1^T Q_1 \hat{H}_1 - \overline{V}^{-1}\right)C_1 + \overline{S}A_{11} + A_{11}^T\overline{S} \\ &\quad + \left(\overline{S}A_{12}\overline{F}_{12} - C_1^T\overline{V}^{-1}C_2\overline{F}_{12}\right)\left(\overline{F}_{12}^T C_2^T \overline{V}^{-1} C_2\overline{F}_{12}\right)^{-1}\left(\overline{S}A_{12}\overline{F}_{12} - C_1^T\overline{V}^{-1}C_2\overline{F}_{12}\right)^T. \end{aligned}$$

can be simplified to (9.77) because $\overline{F}_{12}$ is square and invertible. We know that $\overline{F}_{12}$ is square and invertible because the construction of $\overline{F}_1$ ensures that $\overline{F}_1$ has full column rank and that the size of Ker S, which determines the order reduction, is equal to this column rank.

Proposition 9.6. The Goh Riccati equation of the augmented system (9.81) is equivalent to the Goh Riccati equation of the original system (9.44).

Proof. It is immediate that

$$C\overline{F}_1 = C\begin{bmatrix} B_i & B_{i-1} & \ldots & B_1 \end{bmatrix} = CB_i \tag{9.82}$$

If we examine the term $SA\overline{F}_i - \dot{\overline{F}}_1$ in (9.81):

$$\begin{aligned} S(A\overline{F}_1 - \dot{\overline{F}}_1) &= SA\left[B_i,\ B_{i-1},\ \ldots,\ B_1\right] + S\left[\dot{B}_i,\ \dot{B}_{i-1},\ \ldots,\ \dot{B}_1\right] \\ &= \left[SAB_i - S\dot{B}_i,\ SAB_{i-1} - S\dot{B}_i,\ \ldots,\ SAB_1 - S\dot{B}_1\right] \\ &= \left[SAB_i - S\dot{B}_i,\ SB_i,\ SB_{i-1},\ \ldots,\ SB_2\right]. \end{aligned}$$

Because of Proposition 9.2, this simplifies to

$$S(A\overline{F}_1 - \dot{\overline{F}}_1) = S(AB_i - \dot{B}_i). \tag{9.83}$$

Given, (9.82) and (9.83), the Goh Riccati equation for the augmented system (9.81) reduces to (9.44). ◆

Remark 5. The proposed "augmentation" is simply a restatement of the problem.

Reduced-order filters for the time-varying case are much harder to come by since the transformation matrix, Γ, will now be a function of time. In this case, the only likely option left to the analyst is to use the results of (Oshman and Bar-Itzhack 1985) which give differential equations for the eigenvectors and eigenvalues of the solution to a time-varying Riccati equation. From here the reduced-order Riccati matrix, the transformed system equation and finally the reduced-order filter can be formed through a transformation matrix based upon the eigenvectors. Needless to say, the computation required here will be quite intensive. The state and measurement matrices will also have to be transformed at each time step and only then can the filter be formed and propagated. The point here is that it is possible to find a reduced filter for the time-varying case, though the effort may outweigh the benefits. Since the full-order filter is always available, this is not a serious problem.

The analyst has many options when designing a game theoretic filter. In the case of the full-order filter he has the freedom to choose the different weighting matrices and γ. For reduced-order filters, he can use either the solution to the Goh Riccati equation (9.44) or the solution of linear matrix inequality (9.33) with $\gamma = 0$ to find the needed transformation matrix and reduced-order filter gain. He also has the reduced-order Riccati equation (9.77). Moreover, he can mix the two approaches, for example, by using the LMI to find the transformation matrix and using the reduced-order Goh Riccati equation to find the gain. This flexibility is important, because the solution to the Goh equations may be ill-conditioned when several iterations of the Goh transformation are needed to generate the Riccati equation. The appearance of powers of A in the resulting equation may cause problems with the numerical solution.

9.6 Application to AVCS: An Engine Air Mass Sensor Fault Detection Filter

To demonstrate the effectiveness of the game theoretic filter, we will apply our results to an example derived from (Douglas et al. 1995). In that report, a fault detection and identification system consisting of a bank of Beard-Jones fault detection filters was designed a for a single automobile using the methodology of (Douglas and Speyer 1996). Since we are only trying to provide a design example, we will not attempt to repeat the entire FDI system construction of (Douglas et al. 1995), but will merely design a game theoretic filter for one of the subproblems given in (Douglas et al. 1995): the monitoring of the engine air mass sensor.

In (Douglas et al. 1995), the nonlinear dynamics of a single vehicle was linearized about at straight line path at the constant speed of $25\frac{\text{m}}{\text{sec}}$. The resulting linear dynamics were then further reduced via spectral separation and balanced realizations until a 2-input, 7-output, 7^{th}-order state-space model representing the longitudinal dynamics was found:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du + v. \end{aligned}$$

The measurements are:

$$y = \left\{ \begin{array}{c} y_m \\ y_\omega \\ y_{\ddot{x}} \\ y_{\ddot{z}} \\ y_q \\ y_{y_{fs}} \\ y_{y_{rs}} \end{array} \right\} \begin{array}{l} \text{Engine Manifold Air Mass (kg)} \\ \text{Engine Speed } (\frac{\text{rad}}{\text{sec}}) \\ \text{longitudinal acceleration } (\frac{\text{m}}{\text{sec}^2}) \\ \text{heave acceleration } (\frac{\text{m}}{\text{sec}^2}). \\ \text{Pitch Rate } (\frac{\text{rad}}{\text{sec}}). \\ \text{Forward Symmetric Wheel Speed } (\frac{\text{rad}}{\text{sec}}). \\ \text{Rear Symmetric Wheel Speed } (\frac{\text{rad}}{\text{sec}}). \end{array} \tag{9.84}$$

The inputs are:

$$u = \left\{ \begin{array}{c} \alpha \\ \beta \end{array} \right\} \begin{array}{l} \text{Throttle Angle (deg)} \\ \text{Brake Torque (N-m)} \end{array} \tag{9.85}$$

Because of the balanced realization, the states have no physical meaning.

In all, there are 9 possible actuator/sensor faults. As we discussed earlier, the sensor faults will require detection spaces which are at least 2^{nd}-order. Actuator faults typically need no more than a 1^{st}-order detection space, but because of the direct feedthrough matrix D, the actuator faults in this example will require 3^{rd}-order detection spaces. See (Douglas et al. 1995) for details. Given that we have only 7 states, we will not be able to monitor all of the sensor and actuator faults with a single filter. In (Douglas et al. 1995), the 9 failures were divied up among 4 fault detection filters with some of the failures included in more than one filter for dynamical reasons. To keep our example simple, we will apply the game theoretic filter to only one of the failure sets, which is designated "Filter 1" in (Douglas et al. 1995). In that filter, the following three failures were grouped together:

$$
\begin{aligned}
F_{y_m} &: \text{Air Mass Sensor Failure} \\
F_{y_\omega} &: \text{Engine Speed Sensor Failure} \\
F_{y_{\ddot{x}}} &: \text{Forward Acceleration Sensor Failure}
\end{aligned}
$$

In this example we will attempt to detect the air mass sensor failure, μ_{y_m}, given the possible presence of an engine speed sensor failure, μ_{y_ω}, and forward acceleration sensor failure, $\mu_{y_{\ddot{x}}}$. For comparision, the filter designed in (Douglas et al. 1995) was able to detect and identify each of the three faults. As we noted before, a limitation of the game theoretic filter is that, in its present form, it can only look for one fault per filter and in this example we see this limitation brought to the forefront. Finally, we should also note that the filter we design here will detect μ_{y_m} in the presence of any other failure that enters the system in the same way as μ_{y_ω} and $\mu_{y_{\ddot{x}}}$ or in the presence of any failure whose reachable subspace lies in the sum of the reachable subspaces of F_{y_ω} and $F_{y_{\ddot{x}}}$.

The failure model for this example is:

$$\dot{x} = Ax + F_{y_\omega}\mu_{y_\omega} + F_{y_{\ddot{x}}}\mu_{y_{\ddot{x}}} = Ax + \hat{F}_{y_m}\hat{m}_{y_m} \tag{9.86}$$

$$y = Cx + v, \tag{9.87}$$

where the system matrices are:

$$
A = \begin{bmatrix}
-0.0521 & -0.2213 & 0.2681 & -0.0121 & 0.0136 & 0.0084 & -0.0078 \\
-0.3007 & -8.0277 & -19.0734 & -1.1013 & 0.0795 & 0.2471 & 0.0378 \\
-0.3263 & -19.7571 & -51.0638 & -3.2675 & -4.8766 & -2.4258 & 0.0040 \\
0.0454 & 2.4036 & 15.7922 & -2.1857 & 6.4655 & -0.2062 & 0.0495 \\
0.0219 & 1.1136 & 8.6428 & -7.1817 & -0.6526 & -0.2171 & 0.9316 \\
0.0116 & 0.5928 & 3.8335 & -1.0926 & -0.6513 & -0.9851 & 5.9628 \\
0.0154 & 0.7868 & 4.8494 & -1.4900 & -1.0329 & -6.5688 & -2.5996
\end{bmatrix}
$$

$$C = \begin{bmatrix} 0.0075 & 0.4605 & 0.3710 & 0.1023 & 0.0513 & 0.0340 & -0.0137 \\ 0.7318 & 2.7938 & -2.8640 & 0.1680 & -0.0415 & -0.0491 & -0.0029 \\ 0.0028 & 0.1711 & -0.2654 & 0.0765 & -0.0161 & 0.0093 & -0.0008 \\ 0.0000 & -0.0007 & -0.0005 & -0.0216 & -0.0496 & -0.0438 & 0.0697 \\ -0.0000 & -0.0024 & 0.0050 & 0.0111 & 0.0205 & -0.0027 & 0.0009 \\ 0.4214 & -0.1440 & 0.0371 & 0.2203 & -0.1764 & -0.0129 & 0.1051 \\ 0.4211 & 0.1318 & -0.4410 & -0.2741 & -0.0304 & -0.0734 & 0.0585 \end{bmatrix}$$

For simplicity, the inputs u will be disregarded.

What remains is to calculate F_{y_ω} and $F_{y_{\ddot{x}}}$. Following the the modeling techniques described in Section 9.1, we begin by augmenting the measurment equation to reflect the presence of the engine speed and accelerometer sensor failures:

$$\dot{x} = Ax \tag{9.88}$$

$$y = Cx + E_{y_\omega} m_{y_\omega} + E_{y_{\ddot{x}}} m_{y_{\ddot{x}}} + v. \tag{9.89}$$

where

$$E_{y_\omega} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$
$$E_{y_{\ddot{x}}} = \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix}^T$$

We then calculate f_{y_ω} as the solution of $E_{y_\omega} = Cf_{y_\omega}$ and $f_{y_{\ddot{x}}}$ as the solution to $E_{y_{\ddot{x}}} = Cf_{y_{\ddot{x}}}$. The second column of the failure map is then obtained by multiplying $f_{y_{\ddot{x}}}$ and $f_{y_{\ddot{x}}}$ by the state matrix A. We then have the following failure maps:

$$F_{y_\omega} = \begin{bmatrix} f_{y_\omega} & Af_{y_\omega} \end{bmatrix} = \begin{bmatrix} 0.2107 & -0.0681 \\ 0.2986 & -1.1171 \\ 0.3791 & 14.0532 \\ 1.7301 & -9.9008 \\ -2.3516 & -13.4314 \\ -13.8538 & -43.7274 \\ -9.8358 & 118.5002 \end{bmatrix}$$

and

$$F_{y_{\ddot{x}}} = \begin{bmatrix} f_{y_{\ddot{x}}} & Af_{y_{\ddot{x}}} \end{bmatrix} = \begin{bmatrix} 0.0873 & 0.0209 \\ 0.9262 & 7.7252 \\ 0.2544 & -99.5538 \\ -3.0910 & 35.2772 \\ 4.0831 & 33.4690 \\ 24.1122 & 80.5043 \\ 17.1083 & -200.5111 \end{bmatrix}$$

For the purposes of the filter design we combine the two failure maps into a single complementary failure map:

$$\hat{F}_{y_m} = \begin{bmatrix} F_{y_\omega} & F_{y_{\ddot{x}}} \end{bmatrix}$$

Since $C\hat{F}_{y_m}$ is full rank we do not need to go into a Goh iteration sequence to form the projector $\hat{H}_1$. Thus, this projector is simply:

$$\begin{aligned} \hat{H}_1 &= I - (C\hat{F}_{y_m})[(C\hat{F}_{y_m})^T(C\hat{F}_{y_m})]^{-1}(C\hat{F}_{y_m})^T \\ &= \begin{bmatrix} 0.9986 & -0.0000 & 0.0000 & 0.0098 & -0.0008 & 0.0165 & -0.0317 \\ -0.0000 & 0.0000 & 0.0000 & -0.0000 & 0.0000 & -0.0000 & -0.0000 \\ 0.0000 & 0.0000 & 0.0000 & -0.0000 & -0.0000 & 0.0000 & -0.0000 \\ 0.0098 & -0.0000 & -0.0000 & 0.6340 & 0.0062 & -0.4785 & -0.0540 \\ -0.0008 & 0.0000 & -0.0000 & 0.0062 & 0.9995 & 0.0102 & -0.0179 \\ 0.0165 & -0.0000 & -0.0000 & -0.4785 & 0.0102 & 0.3620 & 0.0397 \\ -0.0317 & -0.0000 & -0.0000 & -0.0540 & -0.0179 & 0.0397 & 0.0058 \end{bmatrix} \end{aligned} \quad (9.90)$$

9.6.1 Full-Order Filter Design

Equation 9.15, the Riccati equation in terms of Π, was used for this example. To bring sensor noise weighting, V $(= \nu I)$, to zero with the disturbance bound, it is assumed that ν is some multiple of γ. By trial and error, it was found that:

$$\nu = 1 \times 10^{-8}, \quad \frac{\nu}{\gamma} = 0.8, \quad Q_1 = R_1 = M_2 = I$$

gave the results seen in Figure 9.2. For the parameters above, the solution of (9.15) is:

$$\Pi = \begin{bmatrix} 0.0108 & -0.0001 & 0.0009 & 0.0043 & -0.0035 & 0.0011 & 0.0003 \\ -0.0001 & 0.0044 & -0.0003 & -0.0033 & -0.0034 & 0.0005 & -0.0004 \\ 0.0009 & -0.0003 & 0.0014 & 0.0020 & 0.0011 & 0.0000 & 0.0001 \\ 0.0043 & -0.0033 & 0.0020 & 0.0059 & 0.0025 & 0.0000 & 0.0005 \\ -0.0035 & -0.0034 & 0.0011 & 0.0025 & 0.0051 & -0.0009 & 0.0003 \\ 0.0011 & 0.0005 & 0.0000 & 0.0000 & -0.0009 & 0.0002 & 0.0000 \\ 0.0003 & -0.0004 & 0.0001 & 0.0005 & 0.0003 & 0.0000 & 0.0000 \end{bmatrix} \tag{9.91}$$

resulting in a gain:

$$L = 10^6 \times \begin{bmatrix} -0.0000 & 0.0037 & -0.0344 & -0.0002 & 0.0000 & -0.0003 & 0.0007 \\ 0.0003 & 0.0218 & -0.0470 & 0.2636 & -0.0004 & 0.3208 & 0.2517 \\ -0.0003 & -0.1411 & -0.1145 & 0.3172 & -0.0007 & 0.3878 & 0.2879 \\ -0.0006 & 0.1147 & -0.1078 & -0.3230 & 0.0006 & -0.3935 & -0.3032 \\ 0.0015 & 0.1110 & 0.3183 & -0.5540 & 0.0012 & -0.6768 & -0.5083 \\ 0.0066 & 0.2818 & 1.7919 & -2.4235 & 0.0050 & -2.9591 & -2.2383 \\ 0.0035 & -1.2066 & 0.1269 & 6.9371 & -0.0120 & 8.4546 & 6.5149 \end{bmatrix} \tag{9.92}$$

When applied to the 7^{th}-order car model, the result is a stable filter with closed-loop poles at: -2, 128, 332.1, -458867.7, -11, 157.0, -856.2, -259.7, -9.1 and -0.31. As Figure 9.2 shows, the filter achieves roughly 80 db. of separation in transmission between the target fault (an engine air mass sensor failure) and the larger of the two nuisance faults. As a comparison, Figure 9.3 plots the results of the Beard-Jones filter design from (Douglas et al. 1995) for the same set of faults. The closed-loop poles for this filter were selected to be: -3, -4, -5, -6, -7, -8 and -9.

A comparison of the two filters shows that they both do an adequate job of separating the target fault and the nuisance faults. The Beard-Jones filter has less separation, but it also amplifies the target fault signal. For the residual processing stage of fault detection and identification, this might prove to be useful side effect. Moreover, the game theoretic filter achieves its impressive transmission separation at the cost of extremely high gains. This is due the agressively low value of γ chosen for this design example. Higher values of γ can be chosen which achieve less separation but also result in smaller gains. We will also show, in the next section, how to design a reduced-order filter which achieves our fault detection goals and which also possesses very reasonable gains.

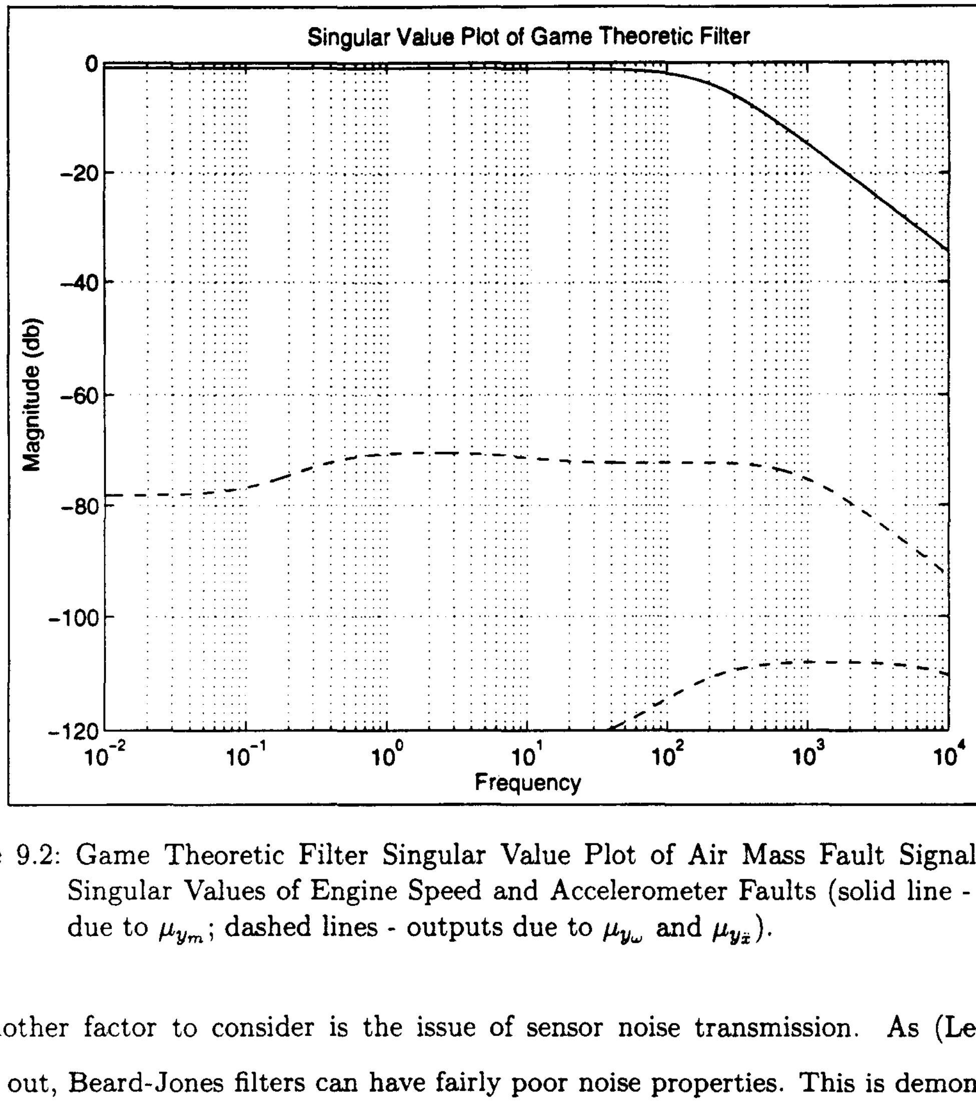

Figure 9.2: Game Theoretic Filter Singular Value Plot of Air Mass Fault Signal versus Singular Values of Engine Speed and Accelerometer Faults (solid line - output due to μ_{y_m}; dashed lines - outputs due to μ_{y_ω} and $\mu_{y_{\ddot{x}}}$).

Another factor to consider is the issue of sensor noise transmission. As (Lee 1994) points out, Beard-Jones filters can have fairly poor noise properties. This is demonstrated by Figure 9.4 which shows that the largest singular value for noise tranmission is consistently larger than the singular value for the target fault transmission. On the other hand, Figure 9.5 shows that the game theoretic filter achieves separation between sensor noise and target fault transmission at frequencies above $10\frac{\text{rad}}{\text{sec}}$ for all of the the noise channels except for the one which comes into the filter dynamics in the same way as the target fault itself. This noise signal is indistinguishable from the target fault and its singular value plot is identical to the target faults over all frequencies. Separating the fault signal from measurement noise will then have to come in the residual evaluation stage. Typically, this

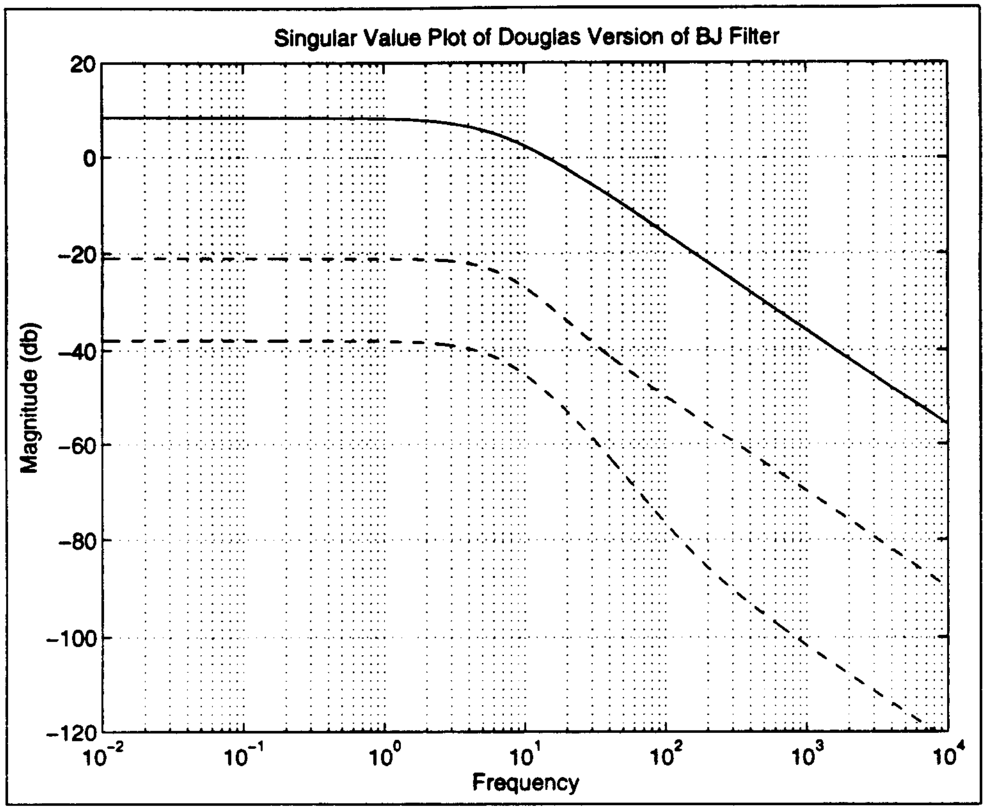

Figure 9.3: Beard-Jones Filter Singular Value Plot of Air Mass Fault Signal versus Singular Values of Engine Speed and Accelerometer Faults (solid line - output due to μ_{y_m}; dashed lines - outputs due to μ_{y_ω} and $\mu_{y_{\ddot{x}}}$).

involves making assumptions about the failure signal and about the statistics of the sensor noise. See for example (Douglas et al. 1995) and (Emami-Naeini et al. 1988).

9.6.2 Reduced-Order Filter Design via the Goh Riccati Equations

We now repeat the example, but now we will design a lower-order filter using the Goh Riccati equations. The first step is to derive the transformation matrix, Γ. Since the transformation is determined via the null space of the full-order Riccati matrix, the design process begins by finding the solution to the full-order Goh Riccati equation (9.44). Because $C\hat{F}_1$ is full-rank, we are spared the step of going through a Goh iteration to set up the correct Goh Riccati equation.

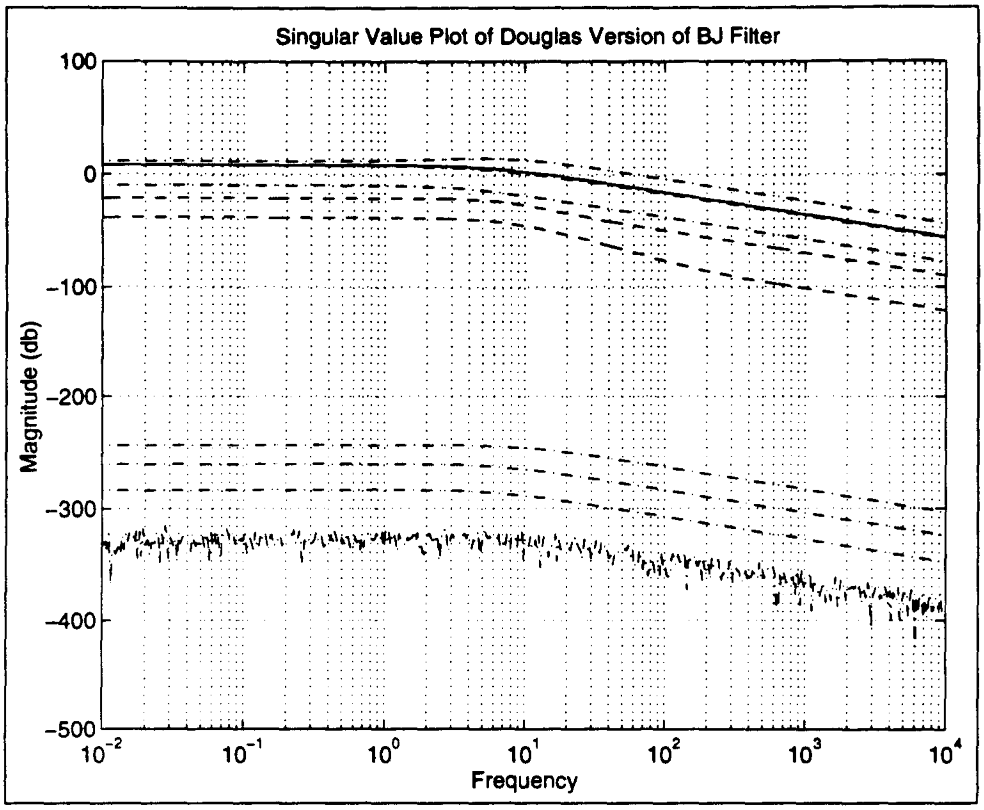

Figure 9.4: Beard-Jones Filter Singular Value Plot of Air Mass Fault Signal versus Singular Values of Engine Speed and Accelerometer Faults (solid line - output due to μ_{y_m}; dashed lines - nuisance faults, dot-dashed lines - noise).

Using the same weightings as in the full-order design, we find that the solution to the Goh Riccati equation (9.44) is:

$$S = \begin{bmatrix} 21.8547 & -0.2217 & -0.0277 & -0.0358 & 0.0271 & 0.0141 & 0.0114 \\ 63.2776 & -0.8201 & -0.0969 & -0.0807 & 0.1369 & 0.0093 & 0.0352 \\ -21.9515 & 0.2891 & 0.0331 & 0.0270 & -0.0496 & -0.0023 & -0.0122 \\ -61.5141 & 0.8211 & 0.0953 & 0.0749 & -0.1416 & -0.0047 & -0.0345 \\ -76.2310 & 0.9668 & 0.1138 & 0.0996 & -0.1586 & -0.0148 & -0.0421 \\ 10.9799 & -0.1357 & -0.0162 & -0.0148 & 0.0216 & 0.0028 & 0.0060 \\ -6.5160 & 0.0860 & 0.0101 & 0.0081 & -0.0146 & -0.0007 & -0.0036 \end{bmatrix}$$

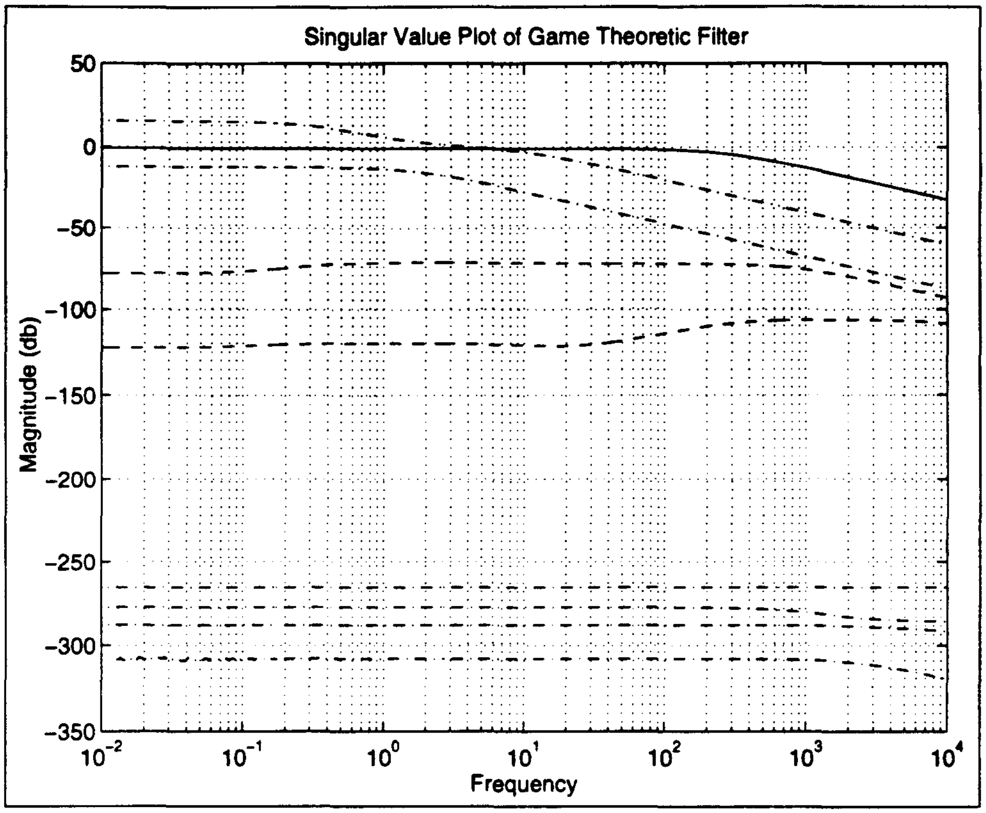

Figure 9.5: Game Theoretic Filter Singular Value Plot of Air Mass Fault Signal versus Nuisance Faults and Noise (solid line - output due to μ_{y_m}; dashed lines - nuisance faults, dot-dashed lines - noise).

Using the QR decompostion we find obtain a transformation matrix:

$$\Gamma^T = \begin{bmatrix} -0.1801 & 0.8639 & 0.0800 & -0.0329 & -0.3166 & 0.3369 & -0.0035 \\ -0.5215 & -0.0913 & -0.6879 & -0.4917 & -0.0018 & 0.0687 & -0.0056 \\ 0.1809 & 0.0982 & -0.6580 & 0.7204 & -0.0020 & 0.0693 & -0.0312 \\ 0.5070 & 0.4348 & -0.2304 & -0.3580 & 0.4084 & -0.4414 & -0.1051 \\ 0.6283 & -0.1984 & -0.1801 & -0.3258 & -0.5190 & 0.3693 & 0.1467 \\ -0.0905 & 0.0797 & -0.0416 & 0.0575 & -0.3232 & -0.5328 & 0.7695 \\ 0.0537 & 0.0312 & 0.0166 & -0.0237 & 0.5992 & 0.5118 & 0.6118 \end{bmatrix} \tag{9.93}$$

Using this transformation, we reduce our state-space to a third-order system, that is, we find the matrices A_{11}, C_1 etc. From here we employ the reduced-order system matrices in the reduced order Goh Riccati equation, (9.81). The soltution to (9.81) using (9.93) is:

$$\overline{S} = \begin{bmatrix} -0.0417 & 0.0216 & -0.3085 \\ 0.0216 & -0.0073 & 0.1923 \\ -0.3085 & 0.1923 & -2.1336 \end{bmatrix} \tag{9.94}$$

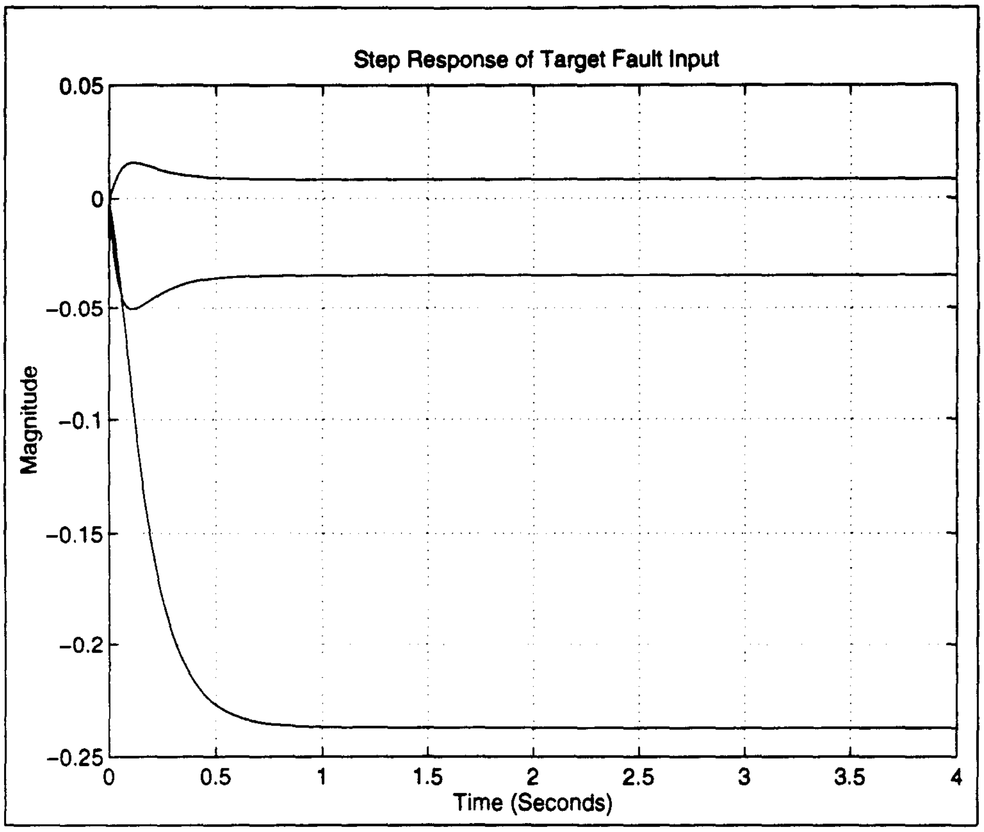

Figure 9.6: Reduced-Order Goh Filter Residual due to step in μ_{A_z} (fault to be detected).

with a corresponding gain:

$$L = \begin{bmatrix} -4.8971 & 0.0001 & 0.0000 & -213.6617 & 39.9079 & 154.1616 & 30.1926 \\ -1.4088 & 0.0000 & 0.0001 & -98.3020 & 24.3178 & 73.0444 & 11.9234 \\ 0.2842 & -0.0001 & 0.0002 & 21.6742 & -3.5535 & -16.2795 & -1.9279 \end{bmatrix} \tag{9.95}$$

The closed-loop eigenvalues are: −7.0976, −23.3114 and −35.2309. To demonstrate the effectiveness of the reduced-order filter a linear simulation of the system was run for two cases: one with a engine air mass sensor fault input (modeled as a step) the other with a engine speed sensor fault input (also a step). Figures 9.7 and 9.6 show that the reduced-order

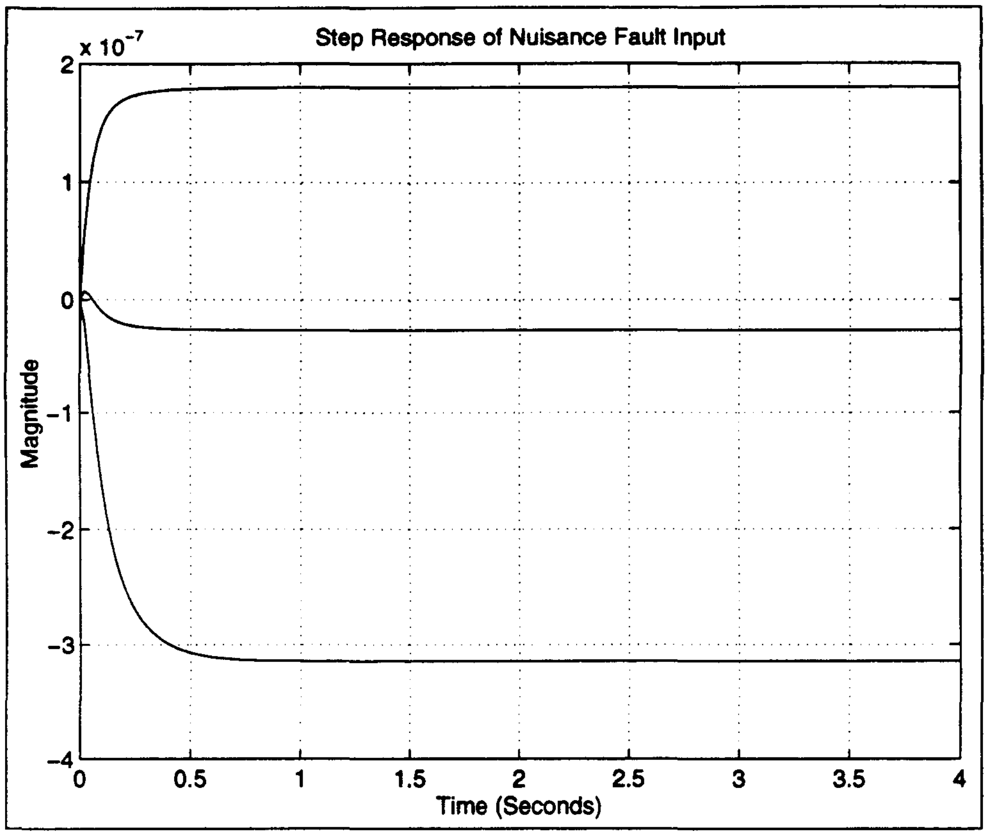

Figure 9.7: Reduced-Order Goh Filter Residual due to step in μ_{wg} (nuisance fault).

filter responds to the air mass sensor fault input and is relatively insensitive to the engine speed sensor fault.

9.7 Discussion

By solving the fault detection problem via disturbance attenuation, we obtain a game theoretic filter that bounds the transmission of disturbances and nuisance faults. By going to the limit of this solution, we get a fault detection filter which in the time-invariant case is equivalent to the Beard-Jones fault detection filter. That is, the presence of the nuisance faults is restricted to an invariant subspace that can be made unobservable through a projection. This unobservable subspace can be factored out of total space to get a lower-order system which is uninfluenced by the nuisance faults. The same factoring process can then be applied to the game filter to get a reduced-order fault detector for the newly reduced state-space. Extensions of this latter result exist for the time-varying case, though the computation involved may be intensive.

The game theoretic approach to fault detection filter design is more flexible than current design methods. The designer can choose the degree to which the game filter possesses the structure of the Beard-Jones filter. This allows him to make tradeoffs between nuisance fault blocking and sensor noise rejection. The linear quadratic game used to solve the disturbance attenuation problem admits time-varying systems and can be used to incorporate parameter uncertainty into the filter design. Recent extensions of robust control such as designs which constrain pole-placement and designs with multiple objectives, for example, the so-called mixed $\mathcal{H}_2/\mathcal{H}_\infty$ problems, suggest that the same can be done here. The latter is of particular interest since it appears to be a logical way to detect and identify multiple faults with a single game theoretic filter.

Finally, we have shown that the limiting form of the game filter is a singular filter. Since any disturbance attenuation problem can be solved in the same manner as this one, it is likely that this result applies to all such problems. That is, the limiting form of a disturbance attenuation problem is a singular optimization problem. This makes applicable a wealth of results from singular control and it provides a new way to understand $\mathcal{H}_\infty$ problems by looking at them as "almost" singular optimal control problems.

CHAPTER 10
Conclusions

ANALYTIC REDUNDANCY is a viable approach to vehicle health monitoring. The fault detection filters developed here perform well in a high-fidelity nonlinear simulation. The filter residuals quickly and clearly respond to the introduction of faults even in the presence of significant vehicle nonlinearities from both longitudinal and lateral modes. Two candidate residual processing systems both effectively automate fault announcement. A Bayesian neural network examines the fault detection filter residual for activity characteristic of a static pattern associated with a fault. A fault and an associated probability of occurance are announced by the neural network soon after the fault is introduced in the vehicle nonlinear simulation. A modified Shiryayev sequential probability ratio test extended to include multiple hypotheses examines the filter residuals and tests for a fault hypotheis change. Both systems respond well to hard and soft failures in the presence of sensor noise, dynamic disturbances and vehicle nonlinearities.

By directing development of the project components in parallel and seeing significant progress in all areas, we are able to identify several important areas for future work: model refinement, robust fault detection filter design, time-varying fault detection filter design, system integration and platoon health monitoring.

Model Refinement: This year, a refined nonlinear vehicle model and simulation was completed. This model allows for arbitrarily changing road gradients for each of the four wheels. Work will now continue by developing uncertainty models associated with process disturbances such as rough and hilly roads, winds, system parameter uncertainty and unmodeled dynamics. Through a good working relation with the Berkeley PATH researchers, model fidelity will be improved further using empirically derived data. Fidelity of the modeled nonlinearities and uncertainties is very important for a realistic assessment of any health monitoring system performance.

Robust Fault Detection Filter Design: Development of robust fault detection filters will continue with two directions of investigation. First, the system will be examined for the possibility of treating nonlinearities and disturbances as pseudo-fault directions. This approach effectively decouples the nonlinearity or disturbance from fault identifying residuals. Second, parameter uncertainty in the linearized vehicle dynamics is modeled as an input-output decomposition. This allows model uncertainty to be treated as a disturbance.

Time-Varying Detection Filter Design: Automated vehicles engaged in merge and split maneuvers may follow a trajectory that induces time-varying vehicle dynamics. The notion of a fault detection filter for time-varying systems was introduced in the game theoretic fault detection filter development described in this report. It is expected that these notions will be extended to invariant subspace filter structures.

System Integration: Having developed preliminary fault detection and isolation system designs for one longitudinal and one lateral mode, work will procede by considering several other design points and then combining all the designs into one integrated package.

Platoon Health Monitoring: Work will begin towards extending the health monitoring system for one vehicle to include the presence of multiple vehicles in a controlled platoon configuration. Sensors required for control such as distance measurements will be included in the fault set. Transmission of vehicle sensor outputs will be transmitted to all vehicles. Feasibility and performance of an expanded health monitoring system will be evaluated in an extended nonlinear simulation.