PIBEE 69-001

Final Report
on
Minimum Sensitivity Design of Attitude
Control Systems

for

National Aeronautics and Space Administration

under

NASA GRANT NGR 33-006-042

PIBEE 69-001

Prepared
by

J. J. Bongiorno, Jr. and P. Dorato
Polytechnic Institute of Brooklyn

Table of Contents

1. INTRODUCTION

The studies under NASA Grant NGR-33-006-42 were concerned with the development of design techniques for systems with uncertain parameters. Two classes of problems were treated. In the first class, statistical information concerning the system uncertainties is known. Only bounds on the uncertain parameters are known in the second class of problems investigated.

With regard to the first class of problems, considerable progress was made for linear systems. A minimum sensitivity design procedure for multivariable systems has been developed. For a given plant, plant parameter covariance matrix, input power spectral density matrix, and required overall system transfer function matrix, formulas are available which give the physically realizable compensation and feedback network transfer function matrices that minimize system sensitivity to plant parameter variations. The derivation of these formulas is contained in a paper which has been submitted for publication. This paper is included in this report as Appendix C.

Some interesting results concerning systems subjected to additive noise disturbances have also been derived. In particular, equations are available which permit one to investigate the change in system performance due to changes in the standard deviation of the additive noise. A paper treating the problem of noise intensity sensitivity was presented at the Second IFAC Symposium on System Sensitivity and Adaptivity, and is included in this report as Appendix D.

Unfortunately, most of the results obtained for the first class of problems studied have limited usefulness in spacecraft attitude control system design at the present time. The main difficulty is that the statistical information needed is not available.

The second class of problems treated is more in line with the information presently available on the uncertain parameters in spacecraft attitude control systems. The approach taken assumes that for each choice of control scheme the uncertain parameters take on the values which will cause the worst performance. One then attempts to select that control whose worst performance is less than the worst performance for any other choice of control. This design approach is referred to as minimax design.

The objective under the present grant was the delineation of spacecraft attitude control problems to which the minimax design procedure can be applied, and the development of analytical results wherever feasible. Although some notable success with regard to acquisition and equilibrium phase attitude control has been achieved, much work remains before practical minimax designs are at hand.

In the sequel, A', $\bar{A}$, A^*, A^{-1}, and $|A|$ denote the transpose, the complex conjugate, the complex conjugate transpose, the inverse, and the determinant, respectively, of the arbitrary matrix A. A diagonal matrix Λ with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_n$ is written as diag$[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Column vectors are represented by $\underset{\sim}{x}$, $\underset{\sim}{y}$, etc., or in the alternative fashion $\underset{\sim}{x} = (x_1 x_2, \ldots x_n)'$ whenever it is desirable to exhibit the components explicitly. The n x n identity matrix, the n-dimensional zero vector, and the n x m zero matrix are denoted by 1_n, $\underset{\sim}{o}_n$, and O_{mn}, respectively.

II. SENSITIVITY DESIGN FOR ACQUISITION MODE (ONE AXIS ACQUISITION)

The attitude control problem considered is that of reducing the angle between a reference axis and a body-fixed axis to zero given a large initial misalignment. Because of the large initial errors, one cannot linearize the dynamical and kinematic equations of motion. It is assumed, however, that any motion of the reference axis can be neglected. This implies that acquisition is achieved in a time interval small compared to the time variation of the reference axis; thus, for example, for an earth centered reference axis acquisition is achieved in a small fraction of the period of rotation around the earth. Let a_1, a_2, and a_3 be the direction cosines of the body fixed axes with respect to the reference axis and let ω_1, ω_2, and ω_3 be the body angular rates. The equations of motion for this problem are then [1]

$$\begin{cases} I_1 \dot{\omega}_1 = (I_2 - I_3)\,\omega_2 \omega_3 + N_1 \\ I_2 \dot{\omega}_2 = (I_3 - I_1)\,\omega_3 \omega_1 + N_2 \\ I_3 \dot{\omega}_3 = (I_1 - I_2)\,\omega_1 \omega_2 + N_3 \end{cases} \qquad (1)$$

and

$$\begin{cases} \dot{a}_1 = \omega_3 a_2 - \omega_2 a_3 \\ \dot{a}_2 = \omega_1 a_3 - \omega_3 a_1 \\ \dot{a}_3 = \omega_2 a_1 - \omega_1 a_2 \end{cases} \quad . \tag{2}$$

It is assumed for this discussion that the direction cosines can be measured exactly. The sensitivity parameters are the inertias I_1, I_2, and I_3 and the torque levels α_1, α_2, and α_3, which are defined as multiplying factors in a feedback control law, i.e.,

$$\begin{cases} N_1 = \alpha_1 \Phi_1(\underset{\sim}{x}) \\ N_2 = \alpha_2 \Phi_2(\underset{\sim}{x}) \\ N_3 = \alpha_3 \Phi_3(\underset{\sim}{x}) \; , \end{cases} \tag{3}$$

where $\Phi_i, i = 1, 2, 3$ are in general nonlinear functions of the state $\underset{\sim}{x} = (\omega_1, \omega_2, \omega_3, a_1, a_2, a_3)'$. It is assumed that the parameter vector $\underset{\sim}{\alpha} = (\alpha_1, \alpha_2, \alpha_3, I_1, I_2, I_3)'$ lies in some closed bounded set Ω. The sensitivity problem is then to design control laws $\Phi_i, i = 1, 2, 3$, such that for any $\underset{\sim}{\alpha}$ which lies in Ω the system performance is acceptable. Acceptable performance may be defined in several ways. The performance may be acceptable if (in order of increasingly stringent requirements)

1. the system is asymptotically stable,
2. if 1 and the settling time t_c never exceeds a given value,
3. if 1 and 2 and a scalar performance measure, such as [2]

$$C = \int_0^{t_c} \sum_i |N_i| \, dt \; , \tag{4}$$

never exceeds a given level.

To solve the sensitivity problems defined by 1 and 2 it appears convenient to use Lyapunov functions V, since the dependence of V and $\dot{V}$ on system parameters is rather explicit. Thus to satisfy stability criterion one must determine Φ_i such that $V(\underset{\sim}{x}) > 0$ and $\dot{V}(\underset{\sim}{x}) < 0$ for all $\underset{\sim}{\alpha}$ which belong to Ω.

To satisfy the settling time criterion one must determine Φ_i such that

$$\max_{\underset{\sim}{\alpha}, \underset{\sim}{x}} \frac{\dot{V}}{V} \tag{5}$$

is minimized*.

Finally to assure acceptable performance for 3, one chooses Φ_i to minimize

$$\max_{\underset{\sim}{\alpha}, \underset{\sim}{x}} C \quad . \tag{6}$$

In both (5) and (6) the initial states are assumed to belong to a closed bounded set X. Unfortunately for the nonlinear dynamics (1) and (2), there is no simple analytical way of computing C for a given control input. One must in this case resort to algorithms which perform the necessary evaluation of C and minimization of max C.

An example has been worked out for the stability criterion, using a Lyapunov function derived elsewhere (Sabroff et. al., op. cit.). Some difficulty arises in the case of settling time problems, since most Lyapunov functions derived for (1) and (2) do not yield a $\dot{V}$ which is negative definite. In these cases $\beta = 0$.

<u>Example</u>

Consider the following problem: $I_1=200, I_2=180,\ 72 \leq I_3 \leq 108$ with

$$\left.\begin{aligned} N_1 &= -k_1\omega_1 - c_2 a_2 \\ N_2 &= -k_2\omega_2 + c_1 a_1 \\ N_3 &= -k_3\omega_3 \end{aligned}\right\} \quad . \tag{7}$$

---

* This value of Φ_i guarantees that

$$\frac{\dot{V}}{V} \leq \min_{\Phi_i} \max_{\underset{\sim}{\alpha}, \underset{\sim}{x}} \frac{\dot{V}}{V} = \beta$$

for all $\underset{\sim}{\alpha}$ which belong to Ω. If this value β does not meet the given bound on settling time, a solution cannot be guaranteed.

Here the form of $\Phi_i(\underset{\sim}{x})$ is assumed and the design parameter vector is the vector $\underset{\sim}{a}' = (k_1, k_2, k_3, c_1, c_2)'$.

The Lyapunov function

$$V = \tfrac{1}{2}\left[I_1\omega_1^2 + I_2\omega_2^2 + I_3\omega_3^2 + 2\,I_1\omega_1 a_2 - 2\,I_2\omega_2 a_1 + 211(1-a_3)\right], \tag{8}$$

suggested in Sabroff et. al. (op. cit., page 107) is considered here. With $c_1 + k_2 = 211$, and $c_2 + k_1 = 211$, the conditions for $\dot{V}$ to be nonpositive are (Sabroff et. al.),

$$k_1 > I_1 \tag{9}$$

$$k_2 > I_2 \tag{10}$$

$$k_3 \geq \frac{I_3^2}{4(k_1 - I_1)} + \frac{I_3^2}{4(k_2 - I_2)} \tag{11}$$

$$\left.\begin{array}{l} c_1 > 0 \\ c_2 > 0 \end{array}\right\} . \tag{12}$$

For the parameter values above and $c_2 = 1$, one vector $\underset{\sim}{a}$ which satisfies the given conditions for all values of I_3 is

$$\underset{\sim}{a}' = (210, 190, 583.2, 21, 1) . \tag{13}$$

The linear control law is then

$$\left.\begin{array}{l} N_1 = -210\,\omega_1 - a_2 \\ N_2 = -190\,\omega_2 + 21\,a_1 \\ N_3 = -583.2\,\omega_3 \end{array}\right\} . \tag{14}$$

This control law is not unique. In general a trade-off is possible between the values of k_1, k_2, and k_3. With saturation effects the above control law becomes

$$N_1 = a_1 \text{ sat} \left[(-210\, \omega_1 - a_2)/a_1 \right]$$

$$N_2 = a_2 \text{ sat} \left[(-190\, \omega_2 + 21\, a_1)/a_2 \right] \qquad (15)$$

$$N_3 = a_3 \text{ sat} \left[(-583.2\, \omega_2)/a_3 \right] ,$$

where a_1, a_2, and a_3 represent in this problem, known saturation levels. The linear mode of operation near the origin is enough to guarantee asymptotic stability for sufficiently small initial disturbances. In general to allow greater initial disturbances the gains vector $\underset{\sim}{a}$ should be as small as possible (to increase initial states in linear mode).

No notable success was achieved with problems 2 and 3 . The major difficulty with problem 2 is that the Lyapunov functions developed for the dynamical equations (1) and (2) have derivatives which are only negative semi definite.* This makes it impossible to obtain an estimate of the decay time from the ratio $\dot{V}/V$. Unfortunately a solution to this has not been developed.

Problem 3 is complicated by the fact that it is very difficult to evaluate the cost function C even for a linear control law, because of the nonlinearity of the system dynamics. However, if the value of C can be evaluated for a particular control law $\underset{\sim}{\Phi}(\underset{\sim}{x})$, the following theorem may be used to obtain an improved control law.

<u>Theorem 1.</u>**

If $V(\underset{\sim}{x}, \underset{\sim}{a})$ is the value of

$$C_u = \int_0^{T_D} L(\underset{\sim}{x}(t), \underset{\sim}{u}(t))dt, \qquad \underset{\sim}{x}(o) = \underset{\sim}{x}$$

---

*See Appendix A.

**This theorem, and its proof, is quite closely related to a theorem developed by Rissanen [9] on performance deterioration of optimum systems. A proof of Theorem 1 appears in Appendix E.

for the control law $\underset{\sim}{u} = \underset{\sim}{\phi}(\underset{\sim}{x})$ and the system

$$\dot{\underset{\sim}{x}} = \underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{\alpha}, \underset{\sim}{u}),$$

where τ_D represents the first passage time to the terminal set D, and $\underset{\sim}{\alpha}$ is a parameter vector, then any control law $\underset{\sim}{v} = \underset{\sim}{\psi}(\underset{\sim}{x})$ for which

$$\frac{\partial V}{\partial \underset{\sim}{x}} \underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{\alpha}, \underset{\sim}{v}) + L(\underset{\sim}{x}, \underset{\sim}{v}) < 0 \tag{16}$$

for all $\underset{\sim}{x}$ and $\underset{\sim}{\alpha} \in \Omega$, yields a cost value which is strictly less than C_u, i.e.,

$$C_v < C_u \qquad \text{all } \underset{\sim}{x}, \underset{\sim}{\alpha} \in \Omega, \tag{17}$$

where $\frac{\partial V}{\partial \underset{\sim}{x}}$ denotes the gradient row vector $\left(\frac{\partial V}{\partial x_1} \cdots \frac{\partial V}{\partial x_n} \right)$.

The advantage of Theorem 1 is that a study of the functional inequality (17) is reduced to the algebraic inequality (16).

III. SENSITIVITY DESIGN FOR EQUILIBRIUM PHASE

The control of the attitude of a spacecraft when the spacecraft orientation is close to its desired orientation is referred to as the equilibrium phase control problem. Because only small deviations are considered in the equilibrium phase, the equations of motion can be linearized. For this purpose, it is convenient to introduce two right-handed ($\underset{\sim}{x}_3 = \underset{\sim}{x}_1 \times \underset{\sim}{x}_2$ and $\underset{\sim}{y}_3 = \underset{\sim}{y}_1 \times \underset{\sim}{y}_2$),* orthonormal sets of three-dimensional vectors $X = \{\underset{\sim}{x}_1, \underset{\sim}{x}_2, \underset{\sim}{x}_3\}$ and $Y = \{\underset{\sim}{y}_1, \underset{\sim}{y}_2, \underset{\sim}{y}_3\}$. The set Y is a basis for a reference frame rotating with respect to inertial space at the constant angular velocity $-\Omega_o \underset{\sim}{y}_2$. The set X is aligned with the principal axes of the spacecraft (assumed to be a rigid body) and represents a basis for the body-fixed frame. When the angular displacement of the X and Y bases is small so that the right-handed rotation required to bring Y into coincidence with X ($\underset{\sim}{x}_i$ aligned with $\underset{\sim}{y}_i$ for $i = 1, 2, 3$) can be approximated by a single rotation of Y about the vector

---

* $\underset{\sim}{a} \times \underset{\sim}{b}$ denotes the three-dimensional vector cross product.

$$\underset{\sim}{\theta} = \sum_{i=1}^{3} \theta_i \underset{\sim}{x}_i , \qquad (18)$$

where θ_i is the right-handed rotation of Y about $\underset{\sim}{x}_i$, $i = 1, 2, 3$ required to bring Y into coincidence with X; and when

$$\underset{\sim}{\omega} = \dot{\underset{\sim}{\theta}} = \sum_{i=1}^{3} \dot{\theta}_i \underset{\sim}{x}_i = \sum_{i=1}^{3} \omega_i \underset{\sim}{x}_i ; \qquad (19)$$

then the linearized equations of motion are

$$\dot{\omega}_1 = \frac{\Omega_o(I_1 - I_2 + I_3)\,\omega_3}{I_i} + \frac{(I_3 - I_2)\,\Omega_o^2\,\theta_1}{I_1} + \frac{\Omega_o h_3}{I_1} - \frac{\dot{h}_1}{I_1} + \frac{M_1}{I_1} \qquad (20)$$

$$\dot{\omega}_2 = \frac{M_2}{I_2} - \frac{\dot{h}_2}{I_2} \qquad (21)$$

$$\dot{\omega}_3 = \frac{\Omega_o(I_2 - I_1 - I_3)\,\omega_1}{I_3} + \frac{(I_1 - I_2)\,\Omega_o^2\,\theta_3}{I_3} - \frac{\Omega_o h_1}{I_3} - \frac{\dot{h}_3}{I_3} + \frac{M_3}{I_3} , \qquad (22)$$

where I_1, I_2, and I_3 are the principal moments of inertia about the respective axes $\underset{\sim}{x}_1, \underset{\sim}{x}_2$, and $\underset{\sim}{x}_3$; h_1, h_2, h_3 are the momenta of reaction flywheels; and M_1, M_2, M_3 account for any disturbance torques and the torques developed by gas jets. The k_i and M_i are the components of the flywheel momentum vector and the body torque vector, respectively, in the basis X.

For long life, use of gas jets should be kept to a minimum. That is, flywheels should be the principal means of controlling the spacecraft in the equilibrium phase. The only source of difficulty in using the flywheels exclusively is velocity saturation. In order to minimize the need for "unloading" the flywheels with the gas jets, the control system should attempt to achieve its objective while keeping the angular velocity of the flywheels small. This being the case, the back emf in the motors driving the flywheels is small, and a good approximation is that the motor armature current is proportional to the armature voltage. With lightly damped flywheels, it is then true that

$$\dot{h}_i = J_i \dot{w}_i = k_i v_i , \quad i = 1, 2, 3 \qquad (23)$$

where, with respect to the $\underset{\sim}{x}_i$ axis, J_i is the flywheel moment of inertia, w_i the flywheel angular velocity, k_i is the voltage-to-torque conversion constant of the flywheel motor, and v_i the armature voltage of the flywheel motor. Substituting (23) into (20), (21), and (22) and making use of (19) yields

$$\dot{\underset{\sim}{x}} = F\underset{\sim}{x} + \hat{G}\,\hat{\underset{\sim}{u}} \,, \tag{24}$$

where the state vector $\underset{\sim}{x}$ (not to be confused with the $\underset{\sim}{x}_i$) is

$$\underset{\sim}{x} = (\theta_1\ \theta_2\ \theta_3\ \omega_1\ \omega_2\ \omega_3\ w_1\ w_2\ w_3)' \,, \tag{25}$$

the input vector $\hat{\underset{\sim}{u}}$ is

$$\hat{\underset{\sim}{u}} = (M_1\ M_2\ M_3\ v_1\ v_2\ v_3) \,, \tag{26}$$

the matrix F is

$$F = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ \frac{\Omega_o^2(I_3-I_2)}{I_1} & 0 & 0 & 0 & 0 & \frac{\Omega_o(I_1-I_2+I_3)}{I_1} & 0 & 0 & \frac{\Omega_o J_3}{I_1} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{\Omega_o^2(I_1-I_2)}{I_3} & \frac{\Omega_o(I_2-I_1-I_3)}{I_3} & 0 & 0 & -\frac{\Omega_o J_1}{I_3} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} , \tag{27}$$

and the matrix $\hat{G}$ is

$$\hat{G} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{I_1} & 0 & 0 & -\frac{k_1}{I_1} & 0 & 0 \\ 0 & \frac{1}{I_2} & 0 & 0 & -\frac{k_2}{I_2} & 0 \\ 0 & 0 & \frac{1}{I_3} & 0 & 0 & -\frac{k_3}{I_3} \\ 0 & 0 & 0 & \frac{k_1}{J_1} & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{k_2}{J_2} & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{k_3}{J_3} \end{bmatrix} . \tag{28}$$

A satellite in a perfect counterclockwise (as viewed from below the plane containing $\underset{\sim}{y}_1$ and $\underset{\sim}{y}_2$) circular orbit with orbital rate Ω_o and center at the earth's center of gravity, a Y basis with $\underset{\sim}{y}_2$ perpendicular to the plane of the orbit and $\underset{\sim}{y}_3$ initially pointed toward the center of the orbit, and an earth whose center of gravity is moving with constant velocity with respect to inertial space is now considered. For this case, it follows that when $\underset{\sim}{x}_3$ and $\underset{\sim}{y}_3$ are perfectly aligned, the $\underset{\sim}{x}_3$ axis of the spacecraft always points toward the center of the orbit. Attention is restricted here to the case where it is also required that $\underset{\sim}{x}_1$ and $\underset{\sim}{y}_1$ be aligned. (With only minor modifications, the following developments can be applied to the cases in which this requirement is not imposed). That is, in addition to maintaining the flywheel speeds close to zero it is desired that X and Y be perfectly aligned. This control objective is met when $\underset{\sim}{x} = \underset{\sim}{o}_9$, and is embodied in the performance index

$$C = \frac{1}{2}\left[\underset{\sim}{x}'(t_1)L\underset{\sim}{x}(t_1) + \int_{t_o}^{t_1} (\underset{\sim}{x}'Q\underset{\sim}{x} + \hat{\underset{\sim}{u}}'R\hat{\underset{\sim}{u}})\,dt\right] \tag{29}$$

where R, L, and Q are real, symmetric, positive definite matrices. The term involving the matrix L causes the performance index to be large if $\underset{\sim}{x}(t_1)$ is significantly different from $\underset{\sim}{o}_9$. The integral term causes the performance index to be large if excessive control $\hat{\underset{\sim}{u}}$ is employed or the state $\underset{\sim}{x}$ is significantly different from $\underset{\sim}{o}_9$ in the interval $t_o \leq t \leq t_1$. The elements in R are chosen so as to weight more heavily use of the gas jets than use of the reaction flywheels. Since satisfactory performance of some spacecraft attitude control systems is achieved even though certain elements of the state vector are not near zero, and since it is desirable that results obtained be applicable to a large class of systems, in the sequel it is assumed that L and Q are only nonnegative definite.

The instants t_o and t_1 depend upon the overall control policy for the spacecraft. One possible mode of operation after the acquisition phase has been completed is to reactivate the equilibrium phase control for a time interval $t_2 = t_1 - t_o$ whenever the norm of $\underset{\sim}{x}$, written $\|\underset{\sim}{x}\|$ and defined by

$$\|\underset{\sim}{x}\| = \sqrt{\underset{\sim}{x}^* \underset{\sim}{x}} \quad , \tag{30}$$

exceeds a prescribed level, say δ. When the disturbances which cause $\|\underset{\sim}{x}\|$ to exceed δ have a frequency of occurrence which is small in relation to the time interval t_c, and when in the available time t_c the equilibrium phase control is effective in reducing $\|\underset{\sim}{x}\|$ to values significantly below δ, then no control need be exerted in intervening periods. The control period t_c should of course be small with respect to the earth's orbital period around the sun in order to justify the assumption of an earth's center of gravity moving with constant velocity in inertial space. During the periods in which control is exerted, the design objective is to realize the minimum value for the "cost" C in (29) in the face of uncertainties.

The parameters likely to have uncertain values in (24) are the body inertias I_1, I_2, and I_3 and certain constants associated with the gas jets. In this regard, it is assumed during periods in which there are no external disturbance torq--es acting, that the control of each gas jet is effected through circuitry with the characteristics illustrated in Fig. 1. When the deadzone d is small and the saturation effect is ignored, the control torques can be approximated by

$$M_i = (M_i)_{dc} = \alpha_i m_i, \quad i = 1, 2, 3 \qquad (31)$$

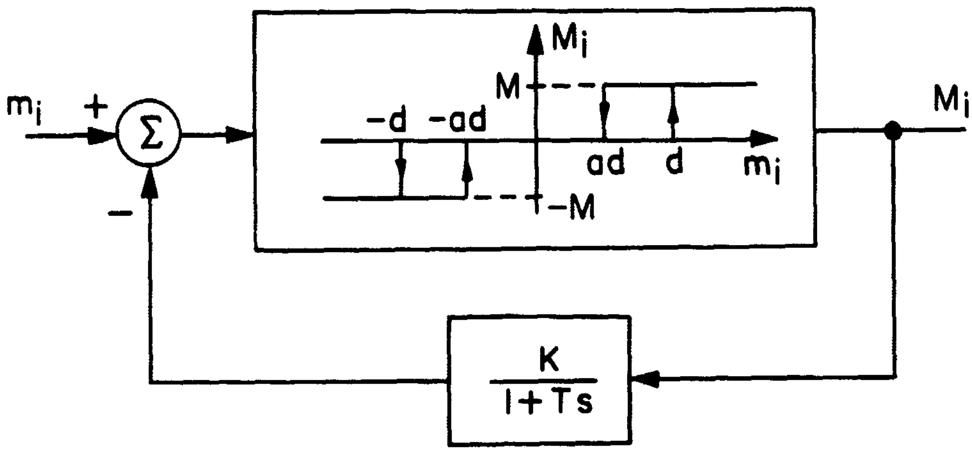
where the α_i are constants which depend on the particular linearization chosen and the assumed values for the parameters in Fig. 1. The saturation is ignored on the basis of the hypothesis that the matrices in (29) are so chosen as to preclude this possibility for optimal controls. Clearly, the α_i are uncertain parameters and any design should be as insensitive to their values as possible. It is assumed that

$$\left.\begin{aligned} (I_i)_{min} \leq I_i \leq (I_i)_{max} \\ (\alpha_i)_{min} \leq \alpha_i \leq (\alpha_i)_{max} \end{aligned}\right\} i = 1, 2, 3 \qquad (32)$$

and that the minimum and maximum value for each parameter is known. The set of values for the parameter vector

$$\underset{\sim}{\alpha} = (I_1 \, I_2 \, I_3 \, \alpha_1 \, \alpha_2 \, \alpha_3)' \qquad (33)$$

which (32) defines is denoted by A.



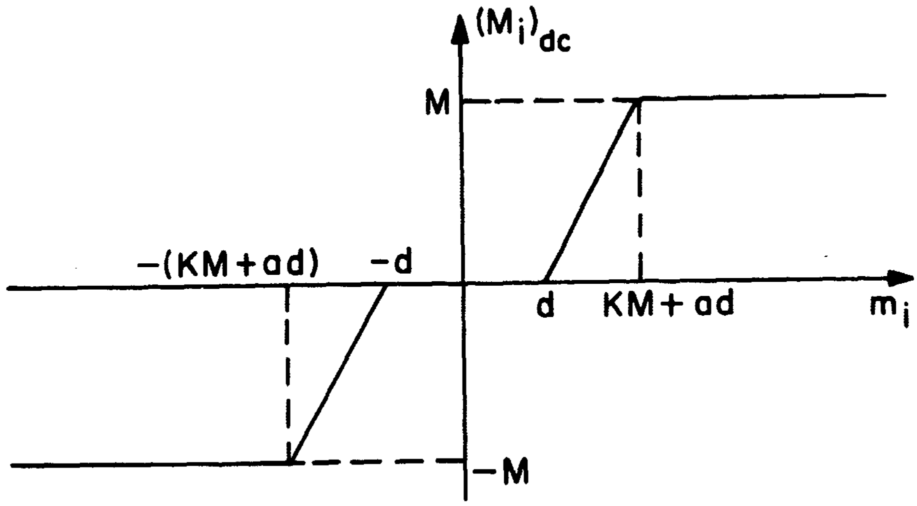


Fig. 1 - Pulse-Frequency, Pulse-Width Modulation Control
(Taken from Ref. 1, pg. 109)

Once the linearization (31) has been made, it is apparent that the control vector is

$$\underset{\sim}{u} = (m_1\ m_2\ m_3\ v_1\ v_2\ v_3)' \ . \tag{34}$$

Clearly,

$$\hat{\underset{\sim}{u}} = \text{diag}\,[\alpha_1, \alpha_2, \alpha_3, 1, 1, 1]\,\underset{\sim}{u} \equiv D\underset{\sim}{u} \ . \tag{35}$$

Hence, in terms of

$$G = \hat{G} D \tag{36}$$

the dynamics (24) become

$$\dot{\underset{\sim}{x}} = F\underset{\sim}{x} + G\underset{\sim}{u} \ . \tag{37}$$

An appropriate cost function in this case is

$$C = \tfrac{1}{2}\left[\underset{\sim}{x}'(t_1)L\underset{\sim}{x}(t_1) + \int_{t_o}^{t_1} (\underset{\sim}{x}'Q\underset{\sim}{x} + \underset{\sim}{u}'R\underset{\sim}{u})dt\right], \tag{38}$$

where some change in the value of R can be introduced to account for the change from $\hat{\underset{\sim}{u}}$ to $\underset{\sim}{u}$.

A precise mathematical statement of the design objective can now be made: Determine the control $\underset{\sim}{u}(t)$, $t_o \leq t \leq t_1$, which minimizes the maximum value of C over all $\underset{\sim}{\alpha}$. That is, the $\underset{\sim}{u}^o$ and $\underset{\sim}{\alpha}^o$ are sought for which

$$\min_{\underset{\sim}{u} \in U}\ \max_{\underset{\sim}{\alpha} \in A}\ C(\underset{\sim}{\alpha}, \underset{\sim}{u}) = C(\underset{\sim}{\alpha}^o, \underset{\sim}{u}^o) \equiv C^o \ . \tag{39}$$

The set U is the set of admissible controls $\underset{\sim}{u}(t)$. When the control $\underset{\sim}{u}^o$ is employed, one is then assured for any $\underset{\sim}{\alpha} \in A$ that $C \leq C^o$. Hence, C^o represents the minimum bound on the performance index C which can be obtained. If this value of C is acceptable, i.e., less than the value of C required for satisfactory performance, then a design is feasible.

A completely general solution to the problem posed by (39) which includes the possibility that $\underset{\sim}{u}$ depend on $\underset{\sim}{\alpha}$ through implementation of a feedback law seems prohibitive. For the case of an open-loop control law ($\underset{\sim}{u}$ is independent of $\underset{\sim}{\alpha}$), some analytical results have been achieved, and they are discussed below. A limited insight into the solution for the case of a closed-loop control law has also been established in the special case

$$\underset{\sim}{u}(t) = -K(t)\,\underset{\sim}{x}(t), \tag{40}$$

where the elements of the matrix K(t) are independent of the parameter vector $\underset{\sim}{\alpha}$ and are chosen to minimize the maximum cost. When K(t) is taken to be a constant matrix, a nonanalytical solution to the problem may possibly be achieved through use of a minimax algorithm developed by Salmon [3]. This algorithmic approach is now discussed.

Consider a cost function C which depends on a design vector $\underset{\sim}{k}$ and a parameter vector $\underset{\sim}{\alpha}$, i.e., $C = C(\underset{\sim}{\alpha}, \underset{\sim}{k})$, $\underset{\sim}{k} \in W$ and $\underset{\sim}{\alpha} \in A$, where W and A are given sets. Such a cost function arises, for example, if the constant control law

$$\underset{\sim}{u}(t) = -K\,\underset{\sim}{x}(t) \tag{41}$$

is applied to the linear system (37) and the cost (38) is evaluated for this control law. In this case the vector $\underset{\sim}{k}$ has as its components all the entries in the matrix K, or those entries of K which are not fixed. Salmon [3] has derived an algorithm which generates two sequences $\{S_i^M\}$ and $\{S_i^m\}$ such that the minimax value of C, denoted C_o, satisfies the inequalities

$$S_i^m \leq C_o \leq S_i^M \tag{42}$$

and

$$\lim_{i \to \infty} \left(S_i^M - S_i^m \right) = 0 . \tag{43}$$

The only conditions for convergence are that C be continuous in $\underset{\sim}{k}$ and $\underset{\sim}{a}$ and that A and W be closed and bounded sets. The algorithm is applied as outlined in the following steps:

1. Choose an arbitrary value of $\underset{\sim}{a} \in A$, say $\underset{\sim}{a}^o$, then minimize $C(\underset{\sim}{a}^o, \underset{\sim}{k})$ with respect to $\underset{\sim}{k} \in W$. Let $\underset{\sim}{k}^o$ be a global minimum of $C(\underset{\sim}{a}^o, \underset{\sim}{k})$, and define

$$C(\underset{\sim}{a}^o, k^o) = S_o^m. \tag{44}$$

2. Maximize $C(\underset{\sim}{a}, \underset{\sim}{k}^o)$, with respect to $\underset{\sim}{a} \in A$. Let the set of $\underset{\sim}{a}$ values which maximize $C(\underset{\sim}{a}, \underset{\sim}{k}^o)$ be denoted $\{\underset{\sim}{a}_i^1\}$. Let A_1 be any subset of $\{\underset{\sim}{a}_i^1\}$ such that

$$\max_{\underset{\sim}{a} \in A_1} C(\underset{\sim}{a}, \underset{\sim}{k}) = \max_{\underset{\sim}{a} \in \{\underset{\sim}{a}_i^1\}} C(\underset{\sim}{a}, \underset{\sim}{k}), \text{ all } \underset{\sim}{k} \in W \tag{45}$$

There may be no proper subset of $\{\underset{\sim}{a}_i^1\}$ which satisfies (45), in which case $A_1 = \{\underset{\sim}{a}_i^1\}$.

Let

$$\max_{\underset{\sim}{a} \in A} C(\underset{\sim}{a}, \underset{\sim}{k}^o) = S_o^M \tag{46}$$

3. Minimize

$$\max_{\underset{\sim}{a} \in A_1} C(\underset{\sim}{a}, \underset{\sim}{k}) \tag{47}$$

with respect to $\underset{\sim}{k} \in W$.* Denote the minimizing value of $\underset{\sim}{k}$ by $\underset{\sim}{k}^1$ and let

$$\max_{\underset{\sim}{a} \in A_1} C(\underset{\sim}{a}, \underset{\sim}{k}^1) = S_1^m \tag{48}$$

*If the set A_1 does not contain a finite number of elements, the algorithm breaks down at this point since then one is confronted with the same type of minimax problem given to start with. Unfortunately, for the stated hypothesis, there is no assurance that such a breakdown will not occur.

4. Maximize $C(\underset{\sim}{a}, \underset{\sim}{k}^1)$ with respect to $\underset{\sim}{a} \epsilon A$. Let $\{a_i^2\}$ denote the set of $\underset{\sim}{a}$ which maximize $C(\underset{\sim}{a}, \underset{\sim}{k}^1)$. Let A_2 be any subset of $A_1 \cup \{a_i^2\}$ such that

$$\max_{\underset{\sim}{a} \epsilon A_2} C(\underset{\sim}{a}, \underset{\sim}{k}) = \max_{\underset{\sim}{a} \epsilon A_1 \cup \{\underset{\sim}{a}_i^2\}} C(\underset{\sim}{a}, \underset{\sim}{k}), \text{ all } \underset{\sim}{k} \epsilon W \tag{49}$$

Let

$$\max_{\underset{\sim}{a} \epsilon A} C(\underset{\sim}{a}, \underset{\sim}{k}^1) = S_1^M \tag{50}$$

5. Repeat the above steps to form the sequences $(A_o \equiv \underset{\sim}{a}_o)$

$$\{S_i^M\} = (S_o^M, S_1^M, S_2^M, \ldots) \tag{51}$$

$$\{S_i^m\} = (S_o^m, S_1^m, S_2^m, \ldots) \tag{52}$$

$$\{\underset{\sim}{k}^i\} = (\underset{\sim}{k}^o, \underset{\sim}{k}^1, \underset{\sim}{k}^2, \ldots) \tag{53}$$

$$\{A_i\} = (A_o, A_1, A_2, \ldots) \tag{54}$$

The nth minimization step may be written

$$\min_{\underset{\sim}{k} \epsilon W} \left[\max_{\underset{\sim}{a} \epsilon A_n} C(\underset{\sim}{a}, \underset{\sim}{k}) \right] = S_n^m, \tag{55}$$

while the nth maximization-step may be written

$$\max_{\underset{\sim}{a} \epsilon A} \left[C(\underset{\sim}{a}, \underset{\sim}{k}^n) \right] = S_n^M \tag{56}$$

where $\underset{\sim}{k}^n$ is the minimizing value in (55), and A_n is such that $A_n = A_{n-1} \cup \{\underset{\sim}{a}_i^n\}$ and

$$\max_{\underset{\sim}{a} \in A_n} C(\underset{\sim}{a}, \underset{\sim}{k}) = \max_{\underset{\sim}{a} \in \left[A_{n-1} \cup \{\underset{\sim}{a}_i^n\} \right]} C(\underset{\sim}{a}, \underset{\sim}{k}), \text{ all } \underset{\sim}{k} \in W , \quad (57)$$

where $\{ a_i^n \}$ denotes the set of $\underset{\sim}{a}$ which are obtained in the maximization step (56).

The process may be terminated when the difference $S_i^M - S_i^m$ becomes less than some specified value ϵ.

There are several practical difficulties with applying this algorithm. One problem is that at each minimization and maximization step one must determine global minima and maximum. This is computationally very difficult to do if one does not assume some structure, e.g., convexity, for the cost function. Another problem, which has already been noted, is that the set A_n may grow too rapidly for practical computation. In particular the set A_n may contain an infinite number of points, in which case the algorithm is no longer useful.

The algorithm described above was not applied to the attitude control problem considered in this study. A description of the algorithm is included here to indicate the type of computational procedures which are available for the sensitivity design of systems using a minimax criterion.

The analytical results obtained for the open-loop problem are for the case in which the only variable parameters are the $a_i, i = 1, 2, 3$. That is, the body inertias are assumed to be known, and the parameter vector is simply

$$\underset{\sim}{a} = (a_1 \; a_2 \; a_3)' . \quad (58)$$

For this case, it can be shown (see Appendix B) that the cost C is a quadratic form in $\underset{\sim}{a}$. It then follows that only those $\underset{\sim}{a} \in A$ which cannot be written in the form

$$\underset{\sim}{a} = (1-s)\, \underset{\sim}{a}_a + s\, \underset{\sim}{a}_b , \quad \underset{\sim}{a}_a, \underset{\sim}{a}_b \in A , \quad 0 < s < 1 \quad (59)$$

need be considered when maximizing the cost C (again, see Appendix B). Since the boundaries of the set A are hyperplanes, this means that only those $\underset{\sim}{a}$ which correspond to corner points need be considered. For convenience, this set of vertices is denoted by

$$V = \left\{ \underset{\sim}{a}_1, \underset{\sim}{a}_2, \ldots, \underset{\sim}{a}_8 \right\} . \quad (60)$$

One can then write

$$C^o = \min_{\underset{\sim}{u} \in U} \max_{\underset{\sim}{\alpha} \in A} C = \min_{\underset{\sim}{u} \in U} \max_{\underset{\sim}{\alpha} \in V} C \qquad (61)$$

and

$$C_o \equiv \max_{\underset{\sim}{\alpha} \in A} \min_{\underset{\sim}{u} \in U} C \leq C^o \leq \min_{\underset{\sim}{u} \in U^o} \max_{\underset{\sim}{\alpha} \in V} C \equiv C_1 , \qquad (62)$$

where the left-hand side inequality is a well known result [4] and

$$U^o = \{\underset{\sim}{u}_o, \underset{\sim}{u}_1, \underset{\sim}{u}_2, \ldots, \underset{\sim}{u}_8\} \qquad (63)$$

is a subset of U. Specifically, $\underset{\sim}{u}_i$, is the optimal control with respect to C for $\underset{\sim}{\alpha} = \underset{\sim}{\alpha}_i$, $i = 1, 2, \ldots, 8$ and $\underset{\sim}{u}_o$ satisfies

$$\max_{\underset{\sim}{\alpha} \in A} \min_{\underset{\sim}{u} \in U} C = C(\underset{\sim}{\alpha}_o, \underset{\sim}{u}_o) = C_o. \qquad (64)$$

Since an analytical solution for the minimax control $\underset{\sim}{u}^o$ is prohibitive in general, the inequality (62) suggests the following approach: choose for the control $\underset{\sim}{u}$ the $\underset{\sim}{u}_i \in U^o$ which minimizes $\max_{\underset{\sim}{\alpha} \in V} C$. When $\underset{\sim}{u}_o$ is the minimizing control and $\underset{\sim}{\alpha}_o$ turns out to be the associated maximizing $\underset{\sim}{\alpha}$, then

$$C_1 = \min_{\underset{\sim}{u} \in U} \max_{\underset{\sim}{\alpha} \in V} C = C(\underset{\sim}{\alpha}_o, \underset{\sim}{u}_o) = C_o \qquad (65)$$

and it follows from (62) that

$$C_o = C^o = C_1 \qquad (66)$$

and

$$\underset{\sim}{u}^o = \underset{\sim}{u}_o \qquad (67)$$

In this case, the minimax solution is realized. Of course, this can only happen when $\underset{\sim}{\alpha}_o \in V$. One can expect more often that $C_1 > C_o$. Since the set

U^o is a set of optimal controls, however, it is reasonable to expect that $C_1 - C_o$ is small in most cases of practical interest. It then follows from

$$C_1 - C^o \leq C_1 - C_o \tag{68}$$

that C_1 is close to C^o and a satisfactory design is realized.

The analytical procedures necessary to carry out the above steps for the linear plant (37) with quadratic cost function (38) are now discussed. This type of system has been treated extensively [5], [6] and the results available in the literature are freely used in the following discussion. First, the vertices in the set V must be identified. This is not difficult to do when the parameter vector $\underset{\sim}{\alpha}$ contains only three elements. The vertices are simply the eight corners of the rectangular parallelepiped bounded by the planes

$$\alpha_i = (\alpha_i)_{min}, \ \alpha_i = (\alpha_i)_{max}; \ i = 1, 2, 3 \tag{69}$$

and are easily enumerated.

The next step is the determination of the set U^o. First the computation of $\underset{\sim}{u}_o$ is discussed. It is well known for the system (37) that

$$\min_{\underset{\sim}{u} \in U} C = \tfrac{1}{2} \underset{\sim}{x}_o' M(t_1, t_o) \underset{\sim}{x}_o, \quad \underset{\sim}{x}_o \equiv \underset{\sim}{x}(t_o) \tag{70}$$

where

$$M(t_1, t) = \left[\Omega_{22}(t_1, t) - L\Omega_{12}(t_1, t)\right]^{-1} \left[L\Omega_{11}(t_1, t) - \Omega_{21}(t_1, t)\right] \tag{71}$$

and the 9 x 9 matrices $\Omega_{11}(t_1, t)$, $\Omega_{12}(t_1, t)$, $\Omega_{21}(t_1, t)$, and $\Omega_{22}(t_1, t)$ are the four partitions of

$$\Omega(t_1, t) = \left[\begin{array}{c|c} \Omega_{11}(t_1, t) & \Omega_{12}(t_1, t) \\ \hline \Omega_{21}(t_1, t) & \Omega_{22}(t_1, t) \end{array}\right] \tag{72}$$

the 18 x 18 state transition matrix for the homogeneous system of equations

$$\dot{z} = \left[\begin{array}{c|c} F & -GR^{-1}G' \\ \hline -Q & -F' \end{array}\right] z . \tag{73}$$

The control is given by

$$u = -R^{-1}G'M(t_1, t)\,\Psi(t, t_o)\,x_o \tag{74}$$

where

$$\dot{\Psi}(t, t_o) = [F - GR^{-1}G'M(t_1, t)]\,\Psi(t, t_o) , \tag{75}$$

$$\Psi(t_o, t_o) = 1_9 . \tag{76}$$

Clearly, $M(t_1, t)$ depends on G and, therefore, the parameter vector α. If one now maximizes (70) with respect to α, the value of α_o can be established. This procedure, however, leads to an α_o which depends on the initial state x_o. Moreover, the evaluation of α_o is not uncomplicated. For these reasons, use is made of the relationship

$$x_o' M(t_1, t_o)\, x_o \geq \lambda_m \|x_o\|^2 \tag{77}$$

where λ_m is the smallest eigenvalue of the real symmetric positive-definite matrix $M(t_1, t_o)$ and the equality sign holds when x_o is an eigenvector associated with the eigenvalue λ_m. It immediately follows that

$$\max_{\alpha \epsilon A} \min_{u \epsilon U} C = C_o \geq \widetilde{C}_o = \tfrac{1}{2}\|x_o\|^2 \lambda_{m_o} \tag{78}$$

where

$$\lambda_{m_o} = \max_{\alpha \epsilon A} \lambda_m \equiv \lambda_m \Big|_{\alpha = \hat{\alpha}_o} . \tag{79}$$

In line with the above developments, one takes

$$u_o = \hat{u}_o = -R^{-1}G_o' M_o(t_1, t)\,\Psi_o(t, t_o)\,x_o \equiv P_o(t_1, t, t_o)\,x_o \tag{80}$$

where

$$G_o = G\Big|_{\underset{\sim}{\alpha} = \hat{\underset{\sim}{\alpha}}_o} , \tag{81}$$

$$M_o(t_1, t) = M(t_1, t)\Big|_{\underset{\sim}{\alpha} = \hat{\underset{\sim}{\alpha}}_o} , \tag{82}$$

and

$$\Psi_o(t, t_o) = \Psi(t, t_o)\Big|_{\underset{\sim}{\alpha} = \hat{\underset{\sim}{\alpha}}_o} . \tag{83}$$

The computation of $\hat{\underset{\sim}{u}}_o$ (although simpler than the computation of $\underset{\sim}{u}_o$) is complicated because (75) is a time-varying matrix differential equation and the calculation of $\hat{\underset{\sim}{\alpha}}_o$ from (79) is not easily accomplished. Computer solutions are required, therefore, in almost all cases of interest.

It is not difficult to show that the remaining elements in the set U^o are given by

$$\underset{\sim}{u}_i = -R^{-1}G_i' M_i(t_1, t)\, \Psi_i(t, t_o)\, \underset{\sim}{x}_o \equiv P_i(t_1, t, t_o)\underset{\sim}{x}_o, \tag{84}$$

where

$$G_i = G\Big|_{\underset{\sim}{\alpha} = \underset{\sim}{\alpha}_i} , \tag{85}$$

$$M_i(t_1, t) = M(t_1, t)\Big|_{\underset{\sim}{\alpha} = \underset{\sim}{\alpha}_i} , \tag{86}$$

and

$$\Psi_i(t, t_o) = \Psi(t, t_o)\Big|_{\underset{\sim}{\alpha} = \underset{\sim}{\alpha}_i} . \tag{87}$$

After the set U^o is established, the next step is the computation of

$$c_i = \max_{\underset{\sim}{\alpha} \in V} C(\underset{\sim}{\alpha}, \underset{\sim}{u}_i) \tag{88}$$

for each $i = 0, 1, 2, \ldots, 8$. One can easily establish after a number of straightforward steps, that

$$c_i = \max_{\underset{\sim}{\alpha} \in V} \tfrac{1}{2}\, \underset{\sim}{x}_o' J_i(t_1, t_o)\underset{\sim}{x}_o \le \tfrac{1}{2} \|\underset{\sim}{x}_o\|^2 \max_{\underset{\sim}{\alpha} \in V} \lambda_{M_i} \equiv \tilde{c}_i , \tag{89}$$

where the real, symmetric, non-negative definite matrix $J_i(t_1, t_o)$ is a function of $\underset{\sim}{a}$, and λ_{M_i} is the largest eigenvalue of J_i. The matrix J_i is given by

$$J_i\,(t_1, t_o) = \frac{1}{2}\,\Big\{\tilde{P}_i'\,(t_1, t_1, t_o)\, L\, \tilde{P}_i(t_1, t_1, t_o)$$

$$+ \int_{t_o}^{t_1} [\,\tilde{P}_i'(t_1, t, t_o)\, Q\, \tilde{P}_i\,(t_1, t, t_o)$$

$$+ P_i'\,(t_1, t, t_o)\, R\, P_i\,(t_1, t, t_o)\,]\, dt \Big\}\ , \qquad (90)$$

where

$$\tilde{P}_i(t_1, t, t_o) = \epsilon^{F(t-t_o)} + \int_{t_o}^{t} \epsilon^{F(t-\tau)}\, G\, P_i(t_1, \tau, t_o)\, d\tau\ , \qquad (91)$$

Since

$$C_1 = \min\{c_o, c_1, \ldots, c_8\} \leq \min\{\tilde{c}_o, \tilde{c}_1, \ldots, \tilde{c}_8\} = \tilde{C}_1 \qquad (92)$$

one chooses for reasons identical to those given for choosing $\underset{\sim}{u}_o = \hat{\underset{\sim}{u}}_o$ the control $\underset{\sim}{u}_{i_o}$ where $i_o \epsilon \{0, 1, 2, \ldots, 8\}$ and

$$\tilde{C}_1 = \tilde{c}_{i_o} \leq \tilde{c}_i\ ,\quad i = 0, 1, \ldots, 8\ . \qquad (93)$$

This choice insures that

$$C \leq \tilde{C}_1 = \tfrac{1}{2}\, \|\underset{\sim}{x}_o\|^2\, \lambda_{M_{io}} \qquad (94)$$

for all $\underset{\sim}{a} \epsilon A$. Clearly,

$$\tilde{C}_o = \tfrac{1}{2}\, \|\underset{\sim}{x}_o\|^2\, \lambda_{m_o} \leq C^o \leq \tfrac{1}{2}\, \|\underset{\sim}{x}_o\|^2\, \lambda_{M_{io}} = \tilde{C}_1\ , \qquad (95)$$

and the effectiveness of the approach presented is measured by the smallness of

$$\eta = \frac{\tilde{C}_1 - \tilde{C}_o}{\tilde{C}_o} = \left(\frac{\lambda_{M_{i_o}}}{\lambda_{m_o}} - 1\right)\ . \qquad (96)$$

It is obvious from the above that the steps required to calculate $\underset{\sim}{u}_{i_o}$ are clear, but not uncomplicated.

The above discussions are directed toward the case in which the dimension of $\underset{\sim}{x}$ is 9 and the dimension of $\underset{\sim}{\alpha}$ is three. All essential aspects of the developments remain unchanged when $\underset{\sim}{x}$ is an n-dimensional vector and $\underset{\sim}{\alpha}$ is a k-dimensional vector. Of course, the larger n and k, the more complicated the computations become.

IV. SOME ADDITIONAL RESULTS AND OBSERVATIONS

When the earth-orbiting spacecraft described in the preceding section is not in a perfect circular orbit and/or when the motion of the earth about the sun is taken into account, then $\underset{\sim}{x} = \underset{\sim}{o}_9$ is no longer the desired result. One now requires that

$$\underset{\sim}{y} = [\theta_1 \; \theta_2 \; \theta_3]' = \underset{\sim}{y}_d \neq \underset{\sim}{o}_3 \qquad (97)$$

in order that the $\underset{\sim}{x}_3$ body axis be pointed toward the center of the earth and there be no rotation of the spacecraft about this axis. The problem then is one of controlling the spacecraft so that it is properly aligned in the face of parameter uncertainties. Since this type of control is long term control, and since the errors build up only slowly without any control, it is reasonable to assume that the control is effected with the use of inertia flywheels only. This is desirable on two counts: first, the gas jets are conserved and second, the control is insensitive to variations in the uncertain α_i, $i = 1, 2, 3$. Hence, for the problem under study the spacecraft input and output vectors have only three components. The same is true for the parameter vector $\underset{\sim}{\alpha}$ whose components are the body inertias I_1, I_2, and I_3.

The block diagram of a possible control system for this purpose is shown in Fig. 2. $G_c(s)$ and $H(s)$ are 3×3 transfer function matrices to be determined and $G_p(s, \underset{\sim}{\alpha})$ denotes the 3×3 spacecraft transfer function matrix after feedback stabilization is introduced. (The spacecraft is normally unstable and must first be stabilized before applying the technique presented in the sequel.) The symbol $\underset{\sim}{\alpha}$ is here used to denote the nominal value of the

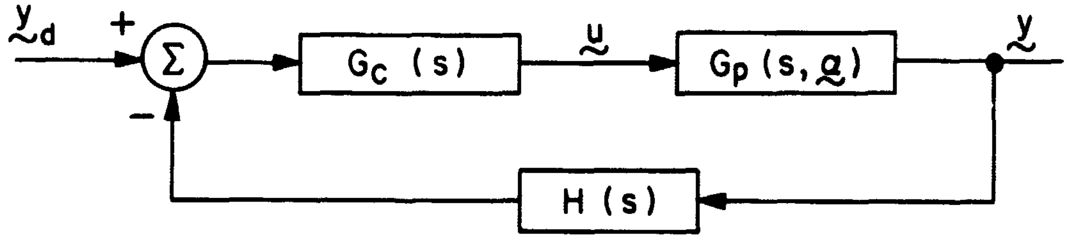


Fig. 2- Spacecraft Control System

plant parameter vector and any deviations from this nominal value are denoted by $\delta\alpha$. Only small deviations are assumed. It is also understood that a physically realizable non-anticipatory and stable nominal transfer function matrix

$$W(s,\alpha) = G_p(s,\alpha)G_c(s)\left[1_3 + H(s)G_p(s,\alpha)G_c(s)\right]^{-1} \tag{98}$$

has been specified.

The objective is the determination of physically realizable transfer function matrices $G_c(s)$ and $H(s)$ for which

$$S = E\left\{ (y_a - y)'(y_a - y) \right\} \tag{99}$$

is a minimum. In (99), y_a denotes the response when instead of α the parameters take on the value $\alpha + \delta\alpha$. Since long term effects are being considered, an infinite range of integration is chosen. The symbol $E\{\cdot\}$ denotes the expected value. The problem posed is a special case of the one whose solution appears in Appendix C. The solution requires that the covariance matrix

$$E\{\delta\alpha\,\delta\alpha'\} = [\sigma_{ij}],\ \sigma_{ji} = \sigma_{ij} \tag{100}$$

be known. Since this knowledge is not likely to be available in the design of spacecraft attitude control system, the results are of limited utility for this application. They are presented, however, since they extend earlier results [7] to the multi-variable case as promised in the proposal for the present grant.

For long term control it is also of interest to determine the effect of additive noise-like torque disturbances. Consider the linearized equations

$$\dot{\underset{\sim}{x}} = F \underset{\sim}{x} + G \underset{\sim}{u} + H \dot{\underset{\sim}{\xi}}, \tag{101}$$

where $\dot{\underset{\sim}{\xi}}$ is a white noise (formally the derivative of a Wiener process $\underset{\sim}{\xi}(t)$ torque disturbance). Consider a performance index

$$C = \frac{1}{2} E \left\{ \underset{\sim}{x}'(t_1) L \underset{\sim}{x}(t_1) + \int_{t_o}^{t_1} (\underset{\sim}{x}' Q \underset{\sim}{x} + \underset{\sim}{u}' R \underset{\sim}{u}) \, dt \right\} . \tag{102}$$

If the variance parameter σ^2 of the white noise, defined by

$$E \left\{ \dot{\underset{\sim}{\xi}}(t) \dot{\underset{\sim}{\xi}}'(\tau) \right\} = \sigma^2 I \, \delta(t-\tau), \tag{103}$$

where I represents the unit matrix, is small, then an estimate of the change in C, due to changes in $\alpha = \sigma^2$, can be obtained from

$$\Delta C = \frac{\partial C}{\partial \alpha} \Delta \alpha \quad . \tag{104}$$

When $\frac{\partial C}{\partial \alpha}$ is evaluated at $\alpha = 0$, formula (104) can be used to study the effect of low intensity noise on a deterministic system. Note that $\alpha = 0$ corresponds to a deterministic system since for a Wiener process $\alpha = \sigma^2 = 0$ implies $\underset{\sim}{\xi}(t) \equiv 0$. It can be shown* that $v = \frac{\partial C}{\partial \alpha}$ satisfies

* This equation is developed in appendix D for the case where ξ is a scalar Wiener process and the system is optimal for the nominal value of α. Appendix D duplicates [10].

the partial differential equation

$$0 = \{ (F\underset{\sim}{x} + G\underset{\sim}{u})\, v_{\underset{\sim}{x}} + \frac{1}{2}\ \mathrm{tr}\ [H H' V_{\underset{\sim}{x}\underset{\sim}{x}}] + v_t + \frac{\alpha}{2}\ \mathrm{tr}\ [H H' v_{\underset{\sim}{x}\underset{\sim}{x}}] \}, \tag{105}$$

where $\underset{\sim}{u} = -K\underset{\sim}{x}$ and

$$v_{\underset{\sim}{x}} = \left[\frac{\partial v}{\partial x_1}, \ldots, \frac{\partial v}{\partial x_n} \right] \tag{106}$$

$$v_{\underset{\sim}{x}\underset{\sim}{x}} = \begin{bmatrix} \frac{\partial^2 v}{\partial x_1^2} & \frac{\partial^2 v}{\partial x_1 \partial x_2} & \cdots \\ \frac{\partial^2 v}{\partial x_2 \partial x_1} & & \\ \vdots & & \end{bmatrix} \tag{107}$$

$$v_t = \frac{\partial v}{\partial t} \tag{108}$$

and where tr A denotes the trace of A, i.e.,

$$\mathrm{tr} \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & \\ \vdots & \end{bmatrix} = a_{11} + a_{22} + \ldots a_{nn}, \tag{109}$$

The function V denotes the value of the cost (102) for the nominal value of α. The boundary condition for (105) is

$$v(t, \underset{\sim}{x}) \Big|_{t=T} = 0, \text{ for all } \underset{\sim}{x}. \tag{110}$$

For the linear problem with additive noise the computation of v is particularly simple, v is given by $v(t, \underset{\sim}{x}) = r(t)$, where $r(t)$ satisfies the equation

$$\dot{r} + \mathrm{tr}\,[H H' P] = 0 , \quad r(T) = 0 ,$$

and P satisfies the equation

$$\begin{aligned} \dot{P} + (F - GK)' P + P (F - GK) &= - (Q + K' RK) \\ P(T) &= O_{nn}. \end{aligned} \tag{112}$$

Since the long term effect of parameter and additive torque disturbances were not considered major problems in this study, no numerical computations are included here for the above results.

V. CONCLUSIONS

A main objective of the initial studies under the present grant was directed toward isolating spacecraft attitude control problems to which the ideas of modern control theory can be successfully applied. Substantial progress has been made in this regard for design in the face of uncertain parameter values, but much yet remains to be done. Difficult computational hurdles must be crossed before results useful in practice are at hand.

ACKNOWLEDGEMENT

The authors wish to thank Professor D. C. Youla for his interest and help in the studies reported herein.

Appendix A

Asymptotic Stability Criterion for $\dot{V} \leq 0$.

Most of the Lyapunov functions, V, which have been found for attitude control systems are such that the time derivative of V, $\dot{V}$, is negative semi-definite. For asymptotic stability the usual stability theorems require that $\dot{V}$ be negative definite (see reference [1], Theorem II, page 37). There does exist a theorem, however, which can be used to guarantee asymptotic stability for a semi-definite $\dot{V}$. Such a theorem appears in references [2] and [3] (Theorem VI, page 58 of reference [1], and Theorem 26.2, page 108 of reference [3]), with a proof outline. Since this theorem is especially relevent to attitude control systems it is presented here with a detailed proof. The proof given here follows that of Youla (reference [2]).

Theorem

Given a nonlinear system

$$\dot{\underset{\sim}{x}} = \underset{\sim}{f}(\underset{\sim}{x}) \tag{1}$$

where $\underset{\sim}{f}(\underset{\sim}{x})$ satisfies the conditions

(a) $\underset{\sim}{f}(\underset{\sim}{o}) = \underset{\sim}{o}$, i.e., $\underset{\sim}{x} = \underset{\sim}{o}$ is an equilibrium point,

(b) $\underset{\sim}{f}(\underset{\sim}{x})$ is continuous and locally Lipshitz* in the region $\|\underset{\sim}{x}\| < h$,

and if there exists a positive definite decrescent function** V whose

---

* A function $\underset{\sim}{f}(\underset{\sim}{x})$ is said to be locally Lipshitz in a Region R if there exists positive numbers b and K such that all $\underset{\sim}{x}_1$ and $\underset{\sim}{x}_2$ in the closed sphere $\|\underset{\sim}{x}\| \leq b$, which lies in R, satisfy

$$\|\underset{\sim}{f}(\underset{\sim}{x}_2) - \underset{\sim}{f}(\underset{\sim}{x}_1)\| \leq K \|\underset{\sim}{x}_2 - \underset{\sim}{x}_1\|$$

** A function $V(\underset{\sim}{x})$ is said to be positive definite if

1. $V(\underset{\sim}{x}) = 0$, if $\underset{\sim}{x} = \underset{\sim}{0}$,
2. $V(\underset{\sim}{x}) \geq \psi(\|\underset{\sim}{x}\|)$, where $\psi(r)$ is real continuous scalar function,

defined for $0 \leq r < h$, which is monotonically increasing in r, i.e., $\psi(r_2) > \psi(r_1)$ if $r_2 > r_1$, and which vanishes for r = 0. A function $V(\underset{\sim}{x})$ is said to be decrescent if a function $\Phi(r)$ with the same properties of $\psi(r)$ above, exists such that

$$|V(\underset{\sim}{x})| \leq \Phi(\|\underset{\sim}{x}\|)$$

total derivative V is not positive, i.e., $\dot{V} \leq 0$, then the null solution $\underset{\sim}{x}(t) = \underset{\sim}{o}$ is asymptotically stable provided the set M of points $\underset{\sim}{x}$ for which

$$\dot{V} = \frac{\partial V}{\partial \underset{\sim}{x}} \underset{\sim}{f}(\underset{\sim}{x}) \equiv 0 , \tag{2}$$

where $\frac{\partial V}{\partial \underset{\sim}{x}}$ denotes the row (gradient) vector $\left(\frac{\partial V}{\partial x_1}, \frac{\partial V}{\partial x_2}, \ldots \frac{\partial V}{\partial x_n} \right)$,

contain no nontrivial solution of (1). It is assumed here that the above gradient exists and is continuous in $\underset{\sim}{x}$.

Proof.

Since $V(\underset{\sim}{x})$ is positive definite there exists a function $\psi(r)$, with previously delineated properties, such that $V(\underset{\sim}{x}) \geq \psi(\|\underset{\sim}{x}\|)$. Given an ϵ such that $0 < \epsilon < h$, from the decrescent property of V it follows that there exists a vector $\underset{\sim}{x}_o$ such that $\|\underset{\sim}{x}_o\| < \epsilon$ and $V(\underset{\sim}{x}_o) < \psi(\epsilon)$. From $\dot{V} \leq 0$ it follows that the system is Lyapunov stable and, hence, with the above choice of $\underset{\sim}{x}_o$, that

$$\|\underset{\sim}{x}(t, \underset{\sim}{x}_o)\| < \epsilon \quad \text{for} \quad t \geq 0 , \tag{3}$$

where $\underset{\sim}{x}(t, \underset{\sim}{x}_o)$ denotes the solution of (1) with $\underset{\sim}{x}(0) = \underset{\sim}{x}_o$. It also follows from $\dot{V} \leq 0$ that V does not increase with time along a trajectory and, hence,

$$V(\underset{\sim}{x}(t, \underset{\sim}{x}_o)) \geq \lim_{t \to \infty} V(\underset{\sim}{x}(t, \underset{\sim}{x}_o)) = V_\infty \geq 0 . \tag{4}$$

It will now be shown that $V_\infty > 0$ leads to a contraction. The proof is then complete since from the properties of V, if $V_\infty = 0$, then

$$\lim_{t \to \infty} \|\underset{\sim}{x}(t, \underset{\sim}{x}_o)\| = 0$$

To continue with the proof, let $\underset{\sim}{x}_*$ be a limiting value of $\underset{\sim}{x}(t, x_o)$; i.e.,

$$\lim_{t \to \infty} \underset{\sim}{x}(t, \underset{\sim}{x}_o) = \underset{\sim}{x}_*$$

It follows from (4) and the continuity of V that $V(\underset{\sim}{x}_*) = V_\infty$. It is evident from (3) that the norm of the limiting value of $\underset{\sim}{x}(t, \underset{\sim}{x}_o)$ cannot exceed ϵ, so that $||\underset{\sim}{x}_*|| \leq \epsilon$. It will now be shown that the limiting norm is actually less than ϵ. If $||\underset{\sim}{x}_*|| = \epsilon$, then since V is positive definite

$$V(\underset{\sim}{x}_*) \geq \psi(||\underset{\sim}{x}_*||) = \psi(\epsilon) \tag{5}$$

But $\dot{V} \leq 0$ implies

$$\lim_{t \to \infty} V(\underset{\sim}{x}(t, \underset{\sim}{x}_o)) = V(\underset{\sim}{x}_*) \leq V(\underset{\sim}{x}_o) . \tag{6}$$

Recall that $\underset{\sim}{x}_o$ was chosen so that $V(\underset{\sim}{x}_o) < \psi(\epsilon)$. Thus (6) yields,

$$V(\underset{\sim}{x}_*) < \psi(\epsilon) , \tag{7}$$

which contradicts (5), hence, $||\underset{\sim}{x}_*|| < \epsilon$. Since $||\underset{\sim}{x}_*|| < \epsilon$, the solution $\underset{\sim}{x}(t, \underset{\sim}{x}_*)$ is well defined and nontrivial. The solution starting at $\underset{\sim}{x}_*$ is nontrivial since $||\underset{\sim}{x}_*|| > 0$ (recall that $V(\underset{\sim}{x}_*) > 0$ by assumption, and $V(\underset{\sim}{x})$ is decrescent so that $0 < V(\underset{\sim}{x}_*) \leq \phi(||\underset{\sim}{x}_*||)$; hence, $||\underset{\sim}{x}_*|| > 0$) and from the uniqueness of solutions $\underset{\sim}{x}(t, \underset{\sim}{x}_*)$ cannot be the null solution if $\underset{\sim}{x}_* \neq 0$. By assumption all nontrivial solutions leave the set M where $\dot{V} = 0$; hence, there exists some finite time t_1 such that

$$V(\underset{\sim}{x}(t_1, \underset{\sim}{x}_*)) < V(\underset{\sim}{x}(0, \underset{\sim}{x}_*)) = V(\underset{\sim}{x}_*) = V_\infty \tag{8}$$

Since $\lim_{t \to \infty} \underset{\sim}{x}(t, \underset{\sim}{x}_o) = \underset{\sim}{x}_*$ and solutions of $\underset{\sim}{x}(t, \underset{\sim}{x}_o)$, are continuous in the initial state $\underset{\sim}{x}_o$ it follows that

$$\lim_{t \to \infty} \underset{\sim}{x}(t_1, \underset{\sim}{x}(t)) = \underset{\sim}{x}(t_1, \underset{\sim}{x}_*) . \tag{9}$$

From the continuity in $V(\underset{\sim}{x})$,

$$\lim_{t \to \infty} V(\underset{\sim}{x}(t_1, \underset{\sim}{x}(t)\ )\ = V(\underset{\sim}{x}(t_1, \underset{\sim}{x}_*)\ ) \tag{10}$$

From the semi-group properties of solutions* $\underset{\sim}{x}(t, \underset{\sim}{x}_o)$ of (1),it follows that

$$\underset{\sim}{x}(t_1 + t, \underset{\sim}{x}_o) = \underset{\sim}{x}(t_1, \underset{\sim}{x}(t) \tag{11}$$

Therefore,from (11) , (10) and (8) it follows that

$$\lim_{t \to \infty} V(\underset{\sim}{x}(t_1, \underset{\sim}{x}(t)\ ) = \lim_{t \to \infty} V(\underset{\sim}{x}(t_1+t, \underset{\sim}{x}_o) = V(\underset{\sim}{x}(t_1, \underset{\sim}{x}_*) < V_\infty \tag{12}$$

But ,

$$\lim_{t \to \infty} V(\underset{\sim}{x}(t_1+t, \underset{\sim}{x}_o)\ ) = V(\underset{\sim}{x}_*) < V_\infty \tag{13}$$

contradicts

$$\lim_{t \to \infty} V(\underset{\sim}{x}(t, \underset{\sim}{x}_o)\ ) = V(\underset{\sim}{x}_*) = V_\infty \tag{14}$$

obtained previously. Thus the assumption $V_\infty > 0$ must be false and the proof is complete.

---

*Solutions of (1) which has unique solutions and is stationary have the so-called semi-group property,

$$\underset{\sim}{x}(t_1 + t_2, \underset{\sim}{x}_o) = \underset{\sim}{x}(t_2, \underset{\sim}{x}(t_1, \underset{\sim}{x}_o)\ )$$

for all positive t_1 and t_2.

Appendix B

Properties of the Open-Loop Cost Function

The dynamical system considered here is described by

$$\dot{\underset{\sim}{x}} = F \underset{\sim}{x} + G \underset{\sim}{u} , \tag{1}$$

where $\underset{\sim}{x}$ and $\underset{\sim}{u}$ are, respectively, n-dimensional and r-dimensional column vectors, and

$$G = \left[\alpha_1 g_1 \mid \alpha_2 g_2 \mid \dots \mid \alpha_N g_N \mid \tilde{G} \right] . \tag{2}$$

The g_i are n-dimensional column vectors and $\tilde{G}$ is an n x (r-N) matrix independent of the α_i . The N-dimensional parameter vector

$$\underset{\sim}{\alpha} = (\alpha_1 \, \alpha_2 \dots \alpha_N)' \tag{3}$$

belongs to a closed, bounded, convex set A and the boundaries of this set are hyperplanes. The system performance is characterized by the cost index

$$C = \frac{1}{2} \left\{ \underset{\sim}{x}'(t_1) L \underset{\sim}{x}(t_1) + \int_{t_o}^{t_1} \left[\underset{\sim}{x}'(t) Q \underset{\sim}{x}(t) + \underset{\sim}{u}'(t) R \underset{\sim}{u}(t) \right] dt \right\} . \tag{4}$$

In (4), L and Q are real, symmetric, nonnegative definite matrices and R is a real, symmetric, positive definite matrix. It is shown in this appendix that

$$\max_{\underset{\sim}{\alpha} \in A} C = \max_{\underset{\sim}{\alpha} \in V} C , \tag{5}$$

where V is the subset of A consisting only of the corner points of A. The corner points of A are all those points which cannot be written as

$$(1-s) \underset{\sim}{\alpha}_a + s \underset{\sim}{\alpha}_b , \quad \underset{\sim}{\alpha}_a, \underset{\sim}{\alpha}_b \in A , \quad 0 < s < 1 . \tag{6}$$

It is first noted that

$$\underset{\sim}{x}(t) = \Phi(t,t_o)\,\underset{\sim}{x}_o + \int_{t_o}^{t} \Phi(t,\tau)\,G\,\underset{\sim}{u}(\tau)\,d\tau\,, \tag{7}$$

where $\Phi(t,\tau)$ is the state transition matrix for the homogeneous system $\dot{\underset{\sim}{x}} = F\,\underset{\sim}{x}$. When $\underset{\sim}{u}$ is partitioned according to

$$\underset{\sim}{u} = (\underset{\sim}{u}_1 \mid \underset{\sim}{u}_2)'\,, \tag{8}$$

where $\underset{\sim}{u}_1$ is an N-dimensional column vector, then

$$G\,\underset{\sim}{u} = [a_1\,g_1 \mid a_2\,g_2 \mid \ldots \mid a_N\,g_N]\,\underset{\sim}{u}_1 + \tilde{G}\,\underset{\sim}{u}_2\,. \tag{9}$$

Substituting (9) into (7) yields

$$\underset{\sim}{x}(t) = \underset{\sim}{v}(t,t_o) + \Psi(t,t_o)\,\underset{\sim}{a}\,, \tag{10}$$

where

$$\underset{\sim}{v}(t,t_o) = \Phi(t,t_o)\,\underset{\sim}{x}_o + \int_{t_o}^{t} \Phi(t,\tau)\,\tilde{G}\,\underset{\sim}{u}_2(\tau)\,d\tau \tag{11}$$

and

$$\Psi(t,t_o) = \int_{t_o}^{t} \Phi(t,\tau)\,[u_1(\tau)\,g_1 \mid u_2(\tau)\,\underset{\sim}{g}_2 \mid \ldots \mid u_N(\tau)\,\underset{\sim}{g}_N]\,d\tau. \tag{12}$$

It follows, therefore, that

$$C = \frac{1}{2}\,(\underset{\sim}{a}'\Gamma\underset{\sim}{a} + 2\,\underset{\sim}{b}'\underset{\sim}{a} + c)\,, \tag{13}$$

where

$$\Gamma = \Gamma(t_1, t_o) = \Psi'(t_1, t_o)\, L\, \Psi(t_1, t_o) + \int_{t_o}^{t_1} \Psi'(t, t_o)\, Q\, \Psi(t, t_o)\, dt, \tag{14}$$

$$\underset{\sim}{b} = \underset{\sim}{b}(t_1, t_o) = \underset{\sim}{v}'(t_1, t_o)\, L\, \Psi(t_1, t_o) + \int_{t_o}^{t_1} \underset{\sim}{v}'(t, t_o)\, Q\, \Psi(t, t_o)\, dt, \tag{15}$$

and

$$c = c(t_1, t_o) = \underset{\sim}{v}'(t_1, t_o)\, L\, \underset{\sim}{v}(t_1, t_o) + \int_{t_o}^{t_1} [\underset{\sim}{v}'(t, t_o)\, Q\, \underset{\sim}{v}(t, t_o) + \underset{\sim}{u}'(t)\, R\, \underset{\sim}{u}(t)]\, dt. \tag{16}$$

It is clear from (13) that the cost C is a quadratic form in $\underset{\sim}{a}$. Moreover, since Γ is a nonnegative definite matrix, the cost is a convex function of $\underset{\sim}{a}$. That is, for $0 \leq s \leq 1$, and for $\underset{\sim}{a}_a$ and $\underset{\sim}{a}_b$ any two vectors in A, then

$$C\,[(1-s)\,\underset{\sim}{a}_a + s\,\underset{\sim}{a}_b] \leq (1-s)\,C(\underset{\sim}{a}_a) + s\,C(\underset{\sim}{a}_b). \tag{17}$$

The result (17) follows immediately from

$$C[(1-s)\,\underset{\sim}{a}_a + s\,\underset{\sim}{a}_b] = (1-s)\,C(\underset{\sim}{a}_a) + s\,C(\underset{\sim}{a}_b) - s(1-s)(\underset{\sim}{a}_a' - \underset{\sim}{a}_b')\,A(\underset{\sim}{a}_a - \underset{\sim}{a}_b). \tag{18}$$

The two possibilities which exist when C takes on its maximum value for $\underset{\sim}{a}$ not a corner point are now considered: First, that $C(\underset{\sim}{a}) \leq C(\underset{\sim}{a}_M)$ where $\underset{\sim}{a}_M$ is an interior point of the set A. Second, that $C(\underset{\sim}{a}) \leq C(\underset{\sim}{a}_M)$ where $\underset{\sim}{a}_M$ is on the boundary of A, but not at a corner point. In the first case, one can immediately conclude that $C(\underset{\sim}{a})$ is constant with respect to $\underset{\sim}{a}$ (see Theorem 4, pg. 71, of Ref. [1]). Hence, the cost C is maximized in this case for any $\underset{\sim}{a}$ in A and, hence, one need only consider the corner points.

In the second case, it is now shown that a similar result holds. When $\underset{\sim}{a}_M$ is on a bounding hyperplane, but not at a corner point, one can always choose an $\epsilon > 0$ and any other point $\underset{\sim}{a} \in A$ which is on the particular hyperplane under consideration such that

$$\underset{\sim}{\beta} = \underset{\sim}{a}_M + \epsilon(\underset{\sim}{a}_M - \underset{\sim}{a}) \tag{19}$$

is in A and also on this hyperplane. With

$$0 < s = \frac{\epsilon}{1+\epsilon} < 1 , \tag{20}$$

one gets

$$\underset{\sim}{a}_M = (1-s)\,\underset{\sim}{\beta} + s\,\underset{\sim}{a} . \tag{21}$$

Using the convexity of the cost function yields

$$C(\underset{\sim}{a}_M) \leq (1-s)\,C(\underset{\sim}{\beta}) + s\,C(\underset{\sim}{a}) . \tag{22}$$

One also has

$$C(\underset{\sim}{a}) \leq C(\underset{\sim}{a}_M) \tag{23}$$

and

$$C(\underset{\sim}{\beta}) \leq C(\underset{\sim}{a}_M). \tag{24}$$

If $C(\underset{\sim}{a}) < C(\underset{\sim}{a}_M)$, it follows from (22) and (24) that

$$C(\underset{\sim}{a}_M) < (1-s)\,C(\underset{\sim}{a}_M) + s\,C(\underset{\sim}{a}_M) = C(\underset{\sim}{a}_M), \tag{25}$$

a contradiction. Hence, $C(\underset{\sim}{a}) = C(\underset{\sim}{a}_M)$ for all points $\underset{\sim}{a}$ including the corner points on the bounding hyperplane containing $\underset{\sim}{a}_M$, and again the maximum C can be realized by considering only the corner points of A.

Appendix C

Minimum Sensitivity Design of Linear Multivariable Feedback Control Systems by Matrix Spectral Factorization*

Introduction

The results of an earlier effort [1] are extended to linear lumped stationary multivariable control systems in this paper. The system considered is shown in Fig. 1. The plant is represented by the rational transfer function matrix $G_p(s, \underset{\sim}{\alpha})$. It is assumed that the plant is asymptotically stable. (When the plant is not asymptotically stable, but is completely controllable it can always be made asymptotically stable with state variable feedback [2] or with output feedback through a compatible observer [3].) The N-dimensional column vector $\underset{\sim}{\alpha}$ represents the mean or expected value of the plant parameters, and any deviation from the mean is denoted by $\delta\underset{\sim}{\alpha}$. Thus,

$$E\{\delta\alpha_i\} = 0, \; i = 1, 2, \ldots, N \tag{1}$$

where $E\{.\}$ denotes the expected value and $\delta\alpha_i$ is the element in the i-row of $\delta\underset{\sim}{\alpha}$. It is assumed that the covariance matrix (the prime denotes the transpose)

$$\Sigma = E\{\delta\underset{\sim}{\alpha}\,\delta\underset{\sim}{\alpha}'\} = [\sigma_{ij}], \sigma_{ji} = \sigma_{ij} = E\{\delta\alpha_i\,\delta\alpha_j\} \tag{2}$$

is known, and that the variations $\delta\alpha_i$ are small and independent of the signals in the system. The input $\underset{\sim}{R}$ is generated by a stationary stochastic process with known power spectral density matrix.

The rational transfer function matrices $G_c(s)$ and $H(s)$ represent, respectively, the tandem compensation network and the feedback network.

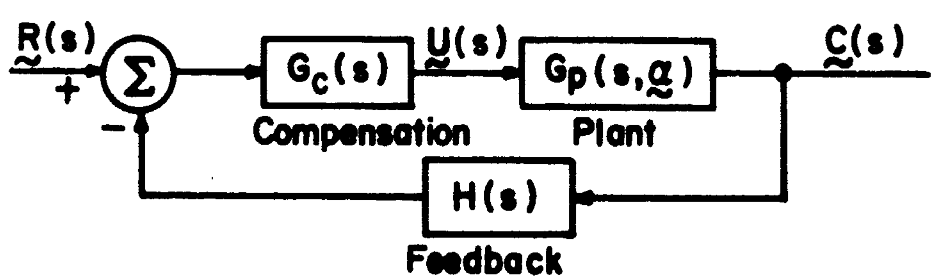


Fig. 1 The System

*This appendix has been submitted to the IEEE Tra tions on Automatic Control by J. J. Bongiorno, Jr.

The bilateral Laplace transform is used exclusively and attention is restricted to only those cases in which the strip of convergence for all transforms includes the imaginary axis of the complex s-plane. In this setting, a transfer function matrix is physically realizable (i. e., the impulse response matrix is causal) if, and only if, all of its elements are analytic in Re s $\geq$ 0.

The objective is the determination of physically realizable transfer function matrices $G_c(s)$ and H(s) for which

$$W(s, \underset{\sim}{a}) = G_p(s, \underset{\sim}{a})\, G_c(s)\, [1_n + H(s)\, G_p(s, \underset{\sim}{a}) G_c(s)]^{-1} \tag{3}$$

(1_n denotes the n x n identity matrix and n is the dimension of $\underset{\sim}{R}$) satisfies the dynamic performance requirements placed on the system and for which the scalar sensitivity measure

$$S = E\,\{(\underset{\sim}{c}_a - \underset{\sim}{c})'\; Q\, (\underset{\sim}{c}_a - \underset{\sim}{c})\} \tag{4}$$

is a minimum. The square matrix Q is real, symmetric, constant, and non-negative definite. The response $\underset{\sim}{c}_a(t)$ is the output response $\underset{\sim}{c}(t)$ when instead of $\underset{\sim}{a}$ the parameter vector takes on the value $\underset{\sim}{a} + \delta\underset{\sim}{a}$.

For the case of single-input-output systems, the sensitivity index (4) reduces to one similar to that employed by Mazer [13]. Here, however, the expectation is taken over the random plant parameters as well as the stochastic inputs. The sensitivity measure (4), except for taking the expected value, is also identical to the one considered by Perkins and Cruz [4], [5]. These same authors in collaboration with Gonzales [6] recently treated the design of the system shown in Fig. 1 from a minimax parameter optimization point of view. Using a computational algorithm they obtain the values of parameters which determine $G_c(s)$ and H(s). When statistical information on the variable plant parameters is available, the approach taken here leads to an analytical solution. Also, no constraint is imposed on the structure of $G_c(s)$ and H(s). It is only required that these matrices be physically realizable. This freedom in the choice of $G_c(s)$ and H(s), however, leads to many designs in which differentiators are required. These differentiators, then, must be approximated with practical circuits.

The solution of the problem posed here is accomplished by first solving the multivariable semi-free-configuration Wiener problem. The solution of this Wiener problem is in itself of some theoretical interest. It was first treated by Hsieh and Leondes [7]. They reduced the problem to the solution of a system of algebraic equations, but never proved that this system of equations has a solution. Indeed, for the free configuration problem Davis [8] states that their method fails in the case of a predictor. The solution for the semi-free-configuration problem is achieved here using the idea of matrix spectral factorization. The conditions under which a matrix can be spectrally factored were first derived by Youla [9]. A computer program for factoring those square rational matrices which can be factored has recently been developed by Tuel [10].

The notation used in this paper is now summarized for easy reference. For an arbitrary matrix A the transpose, the complex conjugate, the adjoint (the complex conjugate transpose), the inverse, the trace, and the determinant of A are denoted by A', $\bar{A}$, A^*, A^{-1}, Tr [A], and $|A|$, respectively. A diagonal matrix Λ with diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_n$ is written as $\Lambda = \text{diag}\,[\lambda_1, \lambda_2, \ldots, \lambda_n]$. Column vectors are represented by $\underset{\sim}{x}$, $\underset{\sim}{y}$, etc., or in the alternative fashion $\underset{\sim}{x} = (x_1, x_2 \ldots x_n)'$ whenever it is desirable to indicate the components explicitly. The n x n identity matrix, the n-dimensional zero vector, and the n x m zero matrix are denoted by 1_n, $\underset{\sim}{o}_n$, and O_{nm}, respectively. The n-dimensional column vector with unity in the i-row and all other elements equal to zero is denoted by $\underset{\sim}{e}_i$. The right inverse of a p x q matrix A is the q x p matrix A^{-1} which has the property $A\,A^{-1} = 1_p$.

A matrix A(s) is rational when each of its elements are rational. The matrix A(s) is analytic in a region when each of its elements are analytic in the region. A(s) is said to be real if $\bar{A}(s) = A(\bar{s})$. When for the matrix A(s) there exists one minor of order ν which does not vanish identically, and when all minors of order greater than ν vanish identically, then A(s) is said to be a matrix with normal rank ν. A point s_o is a pole of A(s) if some element of A(s) has a pole at $s = s_o$. It is also convenient to introduce the notation

$$A_*(s) = A^*(-\bar{s}) \tag{5}$$

which for real matrices - the only kind of interest here - reduces to

$$A_*(s) = A'(-s). \tag{6}$$

Preliminary Analysis

When the number of plant outputs exceeds the normal rank of the plant transfer function matrix, one can always restrict attention to the n independent outputs. Since the number of independent outputs is always less than or equal to the number of inputs, it is always possible to choose n inputs to control the plant. One can choose for the n inputs those associated with the n columns of any nonzero minor of $G_p(s, \underset{\sim}{a})$ of order n. Once attention is restricted to n x n plant transfer function matrices, it immediately follows for input vectors $\underset{\sim}{r}(t)$ of dimension n that both $G_c(s)$ and H(s) are n x n matrices. In the sequel, therefore, all transfer function matrices are square and of order n, and the normal rank of the plant transfer function matrix is n.

The sensitivity index (4) is equivalent to

$$S = E\left\{\int_0^{\infty}\int_0^{\infty} \underset{\sim}{r}'(t-\tau_1)\,\delta W'(\tau_1)\,Q\,\delta W(\tau_2)\,\underset{\sim}{r}(t-\tau_2)\,d\tau_1\,d\tau_2\right\}, \qquad (7)$$

where in terms of impulse-response-matrices

$$\delta W(t) = W(t, \underset{\sim}{a} + \delta\underset{\sim}{a}) - W(t, \underset{\sim}{a}) . \qquad (8)$$

Using the fact that the δa_i and $r_i(t)$ are independent, denoting the expectation with respect to the δa_i by $E_{\delta a}$ and the expectation with respect to the $r_i(t)$ by E_r, and recognizing

$$\mathrm{Tr}\,[A\,B] = \mathrm{Tr}\,[B\,A] \qquad (9)$$

whenever the indicated matrix products are defined, one easily obtains (provided the double integral exists)

$$S = E_{\delta a}\left\{\mathrm{Tr}\left[\int_0^{\infty}\int_0^{\infty} \Phi_r(\tau_2 - \tau_1)\,\Psi(\tau_1, \tau_2)\,d\tau_1\,d\tau_2\right]\right\}, \qquad (10)$$

where

$$\Phi_r(\tau) = E_r\left\{\underset{\sim}{r}(t)\,\underset{\sim}{r}'(t+\tau)\right\} \qquad (11)$$

and

$$\Psi(\tau_1, \tau_2) = \delta W'(\tau_1)\,Q\,\delta W(\tau_2) . \qquad (12)$$

Attention is restricted here to those cases in which the elements of both $\Phi_r(\tau)$ and $\delta W(\tau)$ are bounded by a decaying exponential function of τ. The existence of the double integral in (10) is guaranteed, then, and Parseval's relationship leads to

$$S = E_{\delta a}\left\{ \mathrm{Tr}\left[\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \Phi_r(s)\, \delta W'(s)\, Q\, \delta W(-s)\, ds \right] \right\}, \tag{13}$$

where $\Phi_r(s)$ and $\delta W(s)$ are the bilateral Laplace transforms of $\Phi_r(\tau)$ and $\delta W(\tau)$, respectively. The matrix $\Phi_r(s)$ is the power-spectral-density matrix for the process generating the inputs $\underset{\sim}{r}(t)$.

The approximation

$$\delta W(s) = \sum_{\ell=1}^{N} \frac{\partial W(s, \underset{\sim}{a})}{\partial a_\ell} \delta a_\ell \tag{14}$$

is now made since the δa_i are small. Substituting (14) into (13), recalling (2), and interchanging the expectation, trace, and integral operations yields

$$S = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \mathrm{Tr}\left[\sum_{\ell=1}^{N} \sum_{m=1}^{N} \sigma_{\ell m} \Phi_r(s) \frac{\partial W'(s, \underset{\sim}{a})}{\partial a_\ell} Q \frac{\partial W(-s, \underset{\sim}{a})}{\partial a_m} \right] ds . \tag{15}$$

Since Q is a symmetric non-negative definite matrix, it can always be written as

$$Q = \tilde{Q}' \tilde{Q} , \tag{16}$$

where the rank of Q is equal to the number of rows of $\tilde{Q}$. This being the case, one can take $Q = 1_n$ without any loss in generality: formulas for the case $Q \neq 1_n$ are simply obtained from the formulas derived in the sequel by replacing W with $\tilde{Q}W$. With $Q = 1_n$ in (15) one obtains after substituting

$$\frac{\partial W}{\partial a_k} = (1_n - W H) \frac{\partial G_p}{\partial a_k} G_p^{-1} W \tag{17}$$

and defining

$$\Phi = \sum_{\ell=1}^{N} \sum_{m=1}^{N} \sigma_{\ell m} A_{m*} W_*' \Phi_r W' A_\ell , \tag{18}$$

where

$$A_k = \left[\frac{\partial G_p}{\partial a_k} G_p^{-1} \right]' , \tag{19}$$

the relationship

$$S = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(1_n - WH)'\,(1_n - WH)_*'\;\Phi]\,ds\,. \tag{20}$$

Since the normal rank of G_p is n, the existence of G_p^{-1} is guaranteed almost everywhere in the complex s-plane. In (17) thru (20) and in the sequel, the dependence of Φ, G_p, W, and A_k on s and $\underset{\sim}{\alpha}$, and the dependence of Φ_r, G_c, and H on s is not shown explicitly unless necessary for clarity.

<u>The Multivariable Semi-Free-Configuration Wiener Problem</u>

It is not difficult to show and interesting to note that the mean-square-error for the system shown in Fig. 2,

$$S = E\left\{\underset{\sim}{y}_e'(t)\,\underset{\sim}{y}_e(t)\right\}\,, \tag{21}$$

is given by the value of s in (20) when the power-spectral-density matrix for the process generating the input $\underset{\sim}{y}(t)$ is Φ.

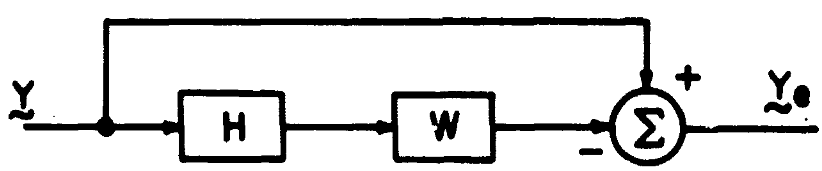


Fig. 2 System Defining Optimum H(s)

Finding the physically realizable H which minimizes S given W is, therefore, equivalent to solving the multivariable semi-free-configuration Wiener problem for the system shown in Fig. 2. This is done here using in part the techniques in Section 4-2 and 4-3 of Reference [11].

Replacing H by $H + \epsilon H_1$ in (20), one obtains

$$S = S_0 - 2\epsilon S_1 + \epsilon^2 S_2\,, \tag{22}$$

where

$$S_0 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(1_n - WH)'\,(1_n - WH)_*'\;\Phi]\,ds\,, \tag{23}$$

$$S_1 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(1_n - WH)'\ W'_*\ H'_{1*}\ \Phi]\, ds, \tag{24}$$

and

$$S_2 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(WH_1)'\ (WH_1)'_*\ \Phi]\, ds\,. \tag{25}$$

When $S_1 = 0$ and $S_2 \geq 0$ for all physically realizable H_1, then there is no physically realizable choice for the feedback network transfer function matrix other than H which gives a smaller value for the sensitivity index S.

The condition

$$S_1 = \frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \mathrm{Tr}\,[\Phi\,(1_n - WH)'\ W'_*\ H'_{1*}]\, ds = 0 \tag{26}$$

is necessary for H to be the optimum transfer function matrix. It must be satisfied for all physically realizable H_1. Hence, it must be satisfied for

$$H_1(s) = h(s)\,[\,\underset{\sim}{e}_i\ \underset{\sim}{e}'_j\,], \tag{27}$$

where the scalar function h(s) is analytic in Re $s \geq 0$ and satisfies

$$\lim_{s \to \infty} s\, h(s)\, \mathrm{Tr}\,[\Phi\,(1_n - WH)'\ W'_*\ (\underset{\sim}{e}_i\ \underset{\sim}{e}'_j)] = 0 \tag{28}$$

Substituting (27) into (26) and using (28) gives

$$\lim_{\rho \to \infty} \int_{C_\rho} [x_{ji}(s)]\ h(-s)\, ds = 0, \tag{29}$$

where $x_{ji}(s)$ is the element in the j-row, i-column of

$$X \equiv \Phi\,(1_n - WH)'\ W'_* \tag{30}$$

and C_ρ is the contour in the complex s-plane consisting of the imaginary axis for $|\omega| \leq \rho$ and the semi-circle $s = \rho\, e^{j\theta}$, $\frac{\pi}{2} \leq \theta \leq \frac{3\pi}{2}$. It follows from (29) and the fact that h(-s) is analytic in Re $s \leq 0$, that $x_{ji}(s)$, $j, i = 1, 2, \ldots, n$ must be analytic everywhere in the half-plane Re $s \leq 0$. That is, H must satisfy (30) where the matrix X is analytic in Re $s \leq 0$, but is otherwise arbitrary.

Conditions are discussed in the next section under which it is possible to write

$$\Phi = \Delta_*\ \Delta \tag{31}$$

and

$$\Omega \equiv W_* \; W \; = \; \Gamma_* \, \Gamma, \tag{32}$$

where the n x n square matrices Δ and Γ together with their inverses are analytic in Re s $\geq$ 0. Assuming these conditions are met, one obtains from (30) after transposing and making the substitutions (31) and (32)

$$\Gamma_*^{-1} \, (W_* - \Gamma_* \, \Gamma H) \, \Delta' \; = \; \Gamma_*^{-1} \, X' \, (\Delta_*')^{-1} \equiv \tilde{X} \tag{33}$$

where $\tilde{X}$ is analytic in Re s $\leq$ 0. Now one can write

$$\Gamma_*^{-1} W_* \, \Delta' \; = \{ \Gamma_*^{-1} \, W_* \, \Delta' \}_+ \; + \{\Gamma_*^{-1} W_* \, \Delta' \}_- \; , \tag{34}$$

where

$$\{\Gamma_*^{-1} \; W_* \; \Delta' \}_+ = \int_0^\infty \epsilon^{-st} \left[\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \Gamma_*^{-1} W_* \Delta' \epsilon^{st} \, ds \right] dt \tag{35}$$

is analytic in Re s $\geq$ 0 and

$$\{\Gamma_*^{-1} W_* \, \Delta' \}_- \; = \int_{-\infty}^{0} \epsilon^{-st} \left[\frac{1}{2\pi j} \int_{-j\infty}^{j\infty} \Gamma_*^{-1} W_* \; \Delta' \; \epsilon^{st} \, ds \right] dt \tag{36}$$

is analytic in Re s $\leq$ 0. It therefore follows from (33) that

$$\{\Gamma_*^{-1} W_* \Delta' \}_+ \; - \Gamma H \; \Delta' \; = \; \tilde{X} - \{\Gamma_*^{-1} \; W_* \; \Delta' \}_- \tag{37}$$

The left-hand side of (37) is analytic everywhere in the half-plane Re s $\geq$ 0 and the right-hand side of (37) is analytic everywhere in the half-plane Re s $\leq$ 0. Thus, the left-hand side of (37) must be analytic everywhere. This is the case if, and only if,

$$\{\Gamma_*^{-1} \, W_* \; \Delta' \}_+ - \Gamma H \; \Delta' \; = K \; , \tag{38}$$

where K is an arbitrary polynomial matrix. Solving for H and recognizing that (32) implies $\Gamma_*^{-1} \; W_* = \Gamma W^{-1}$ yields($|W| \not\equiv 0$ is assumed)

$$H \; = H_o - \Gamma^{-1} \, K \, (\Delta')^{-1} \; , \tag{39}$$

where

$$H_o = \Gamma^{-1} \{ \, \Gamma W^{-1} \, \Delta' \, \}_+ \, (\Delta')^{-1} \; . \tag{40}$$

Equation (39) is arrived at after consideration of a special class of variations for H_1. It can not be stated, therefore, that $S_1 = 0$ for all physically realizable H_1 when H is given by (39). It can be stated, however, that if an optimum H exists it must be included among the family of functions defined by (39). It is now shown that H_o is the optimum H. Substituting (39) into (20) one obtains

$$S = \tilde{S}_0 + 2\,\tilde{S}_1 + \tilde{S}_2 , \tag{41}$$

where

$$\tilde{S}_0 = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(1_n - WH_o)'\,(1_n - WH_o)'_*\,\Phi]\,ds , \tag{42}$$

$$\tilde{S}_1 = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(1_n - WH_o)'\,W'_*(\Gamma'_*)^{-1}K'_*\Delta_*^{-1}\,\Phi]\,ds , \tag{43}$$

and

$$\tilde{S}_2 = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \mathrm{Tr}\,[\Delta^{-1}K'\,(\Gamma')^{-1}\,W'W'_*\,(\Gamma'_*)^{-1}\,K'_*\Delta_*^{-1}\,\Phi]\,ds. \tag{44}$$

[Equations (41) thru (44) can also be arrived at by substituting $\epsilon = 1$, $H = H_o$, and $H_1 = -\Gamma^{-1}K(\Delta')^{-1}$ in (22) thru (25).] From (40) it follows that

$$(1_n - WH_o) = W\,\Gamma^{-1}\,(\Gamma W^{-1}\Delta' - \{\Gamma W^{-1}\Delta'\}_+)\,(\Delta')^{-1} \tag{45}$$

or

$$(1_n - WH_o) = W\,\Gamma^{-1}\,\{\Gamma W^{-1}\Delta'\}_-\,(\Delta')^{-1} . \tag{46}$$

Using (31), (32), and (46) in (42) thru (44) gives

$$\tilde{S}_0 = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(\{\Gamma W^{-1}\Delta'\}_-)'_*\,(\{\Gamma W^{-1}\Delta'\}_-)]\,ds , \tag{47}$$

$$\tilde{S}_1 = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \mathrm{Tr}\,[(\{\Gamma W^{-1}\Delta'\}_-)\,K_*]\,ds , \tag{48}$$

and

$$\tilde{S}_2 = \frac{1}{2\pi j}\int_{-j\infty}^{j\infty} \mathrm{Tr}\,[K\,K_*]\,ds . \tag{49}$$

The above results are arrived at with the aid of (9) and the additional fact that a matrix and its transpose have the same trace.

It follows from (41) and (47) thru (49) that if either

$$K = O_{nn} \tag{50}$$

or

$$K \rightarrow - \{\Gamma W^{-1} \Delta\}_- , \quad |s| \rightarrow \infty \tag{51}$$

then, and only then, can the sensitivity index be finite. The condition (51) is never satisfied unless W^{-1} has no poles in Re $s \geq 0$: in this case $\{\Gamma W^{-1} \Delta'\}_- = O_{nn}$ and the best result possible, S = 0, is obtained. When (50) and (51) are not satisfied it follows that

$$Tr\,[K K_*] = q\,(s^2) \not\equiv 0, \tag{52}$$

where $q(s^2)$ is an arbitrary polynomial in s^2. This is the case because

$$Tr\,[KK_*] = \sum_{i=1}^{n} k_i' \, k_{i*}' = 0 \tag{53}$$

if, and only if, $k_i = \varrho_n$, i = 1, 2, . . . , n where k_i' is the i-row of K. Substituting (52) into (49) immediately leads to $\tilde{S}_2$ and, therefore, S being infinite. The only possible choice for H in (39) is, then, $H = H_o$.

It is now verified that $S_2 \geq 0$ for all physically realizable H_1 and, therefore, that $H = H_o$ is the optimum choice for the feedback network transfer function matrix. From (25), it suffices to show that

$$I(j\omega) = [\,(WH_1)'_* \, \Phi \, (WH_1)'\,)\,]\Big|_{s = j\omega} = [\,(WH_1)'_* \, \Delta_* \, \Delta\,(WH_1)'\,]\Big|_{s = j\omega} \tag{54}$$

is a non-negative definite hermitian matrix. Applying the definition (5) with $s = j\omega$ to (54) gives

$$I(j\omega) = Z^*(j\omega)\, Z\,(J\omega), \tag{55}$$

where

$$Z(j\omega) = [\Delta\,(WH_1)'\,]\Big|_{s = j\omega} . \tag{56}$$

Clearly, $I(j\omega)$ is a non-negative definite hermitian matrix. Thus, $H = H_o$ is optimum, and (47) is a compact formula for the minimum value of S:

$$\min S = S\Big|_{H = H_o} = \hat{S}_0 \tag{57}$$

On the Spectral Factorizations

The fundamental theorem regarding the spectral factorization of rational matrices is contained in [9]. An abbreviated statement of the theorem suitable for the problem being treated here is

Theorem 1: When the n x n real rational matrix $A(s)$ satisfies

a_1) $A(s) = A'(-s) = A_*(s)$,

a_2) $A(s)$ is analytic on the finite $s = j\omega$ axis, and

a_3) $A(j\omega)$ is positive definite for all finite ω, then there exists an n x n real rational matrix $B(s)$ such that

b_1) $A(s) = B'(-s)\,B(s) = B_*(s)\,B(s)$ and

b_2) $B(s)$ and $B^{-1}(s)$ are both analytic in Re $s \geq 0$.

An immediate consequence of Theorem 1 is

Theorem 2: Sufficient conditions for the n x n real rational matrix $\Omega = W_* W$ to have the spectral factorization $\Omega = \Gamma_* \Gamma$ where Γ and Γ^{-1} are both n x n real rational matrices analytic in Re $s \geq 0$ are :

c_1) $W(s)$ be physically realizable and

c_2) $|W(j\omega)| \neq 0$ for all finite ω, or

c_3) $W^{-1}(s)$ be analytic on the finite $s = j\omega$ axis.

Proof: Clearly, $\Omega(s) = W'(-s)\,W(s) = \Omega'(-s)$ and condition a_1) is satisfied. Since $W(s)$ is physically realizable, $W(s)$ is analytic in Re $s \geq 0$ and $W'(-s)$ is analytic in Re $s \leq 0$. Hence, $\Omega(s)$ is analytic for $s = j\omega$, and condition a_2) is met. Finally, $|W(j\omega)| \neq 0$ guarantees that the hermitian matrix $\Omega(j\omega)$ is positive definite, and a_3) is satisfied. The equivalence of c_2) and c_3) follows immediately from $W(s)\,W^{-1}(s) = 1_n$. Wherever $W^{-1}(s)$ is analytic on the $s = j\omega$ axis, $|W^{-1}(j\omega)|$ is finite and $|W(j\omega)|\,|W^{-1}(j\omega)| = 1$ leads to $|W(j\omega)| \neq 0$. On the other hand, wherever $|W(j\omega)| \neq 0$, $|W^{-1}(j\omega)|$ must be finite. This is the case for finite ω only if $W^{-1}(s)$ is analytic on the finite $s = j\omega$ axis.

The spectral factorization of the matrix Φ defined by (18) is now considered. The conditions under which it is possible to accomplish the required spectral factorization are embodied in

Theorem 3: The n x n real rational matrix Φ has the spectral factorization

$\Phi = \Delta_* \Delta$ where Δ and Δ^{-1} are both analytic in Re s $\geq$ 0 whenever all of the following conditions are satisfied:

d_1) Φ_r satisfies a_1) thru a_3),

d_2) W satisfied c_1) and c_2),

d_3) G_p^{-1} is analytic on the finite s = jω axis,

d_4) The N x N covariance matrix $\Sigma = [\sigma_{ij}]$ is positive definite,

and

d_5) the rank of the n x nN matrix

$$\nabla = \left[\frac{\partial G_p}{\partial \alpha_1} \middle| \frac{\partial G_p}{\partial \alpha_2} \middle| \cdots \middle| \frac{\partial G_p}{\partial \alpha_N} \right]$$

is n everywhere on the finite s = jω axis.

Proof: It is not difficult to establish from d_1) and d_2) that the representation

$$W'_* \Phi_r W' = V_* V \tag{58}$$

is possible where the real rational n x n matrices V and V^{-1} are analytic in Re s $\geq$ 0. Equation (18) is, therefore, of the form

$$\Phi = \sum_{\ell=1}^{N} \sum_{m=1}^{N} \sigma_{\ell m} M_{m*} M_{\ell} , \tag{59}$$

where

$$M_k = V A_k = V \left[\frac{\partial G_p}{\alpha_k} G_p^{-1} \right]' . \tag{60}$$

Since G_p is real and rational it follows that the M_k are also. Hence, Φ is an n x n real rational matrix. Moreover,

$$\Phi_* = \sum_{\ell=1}^{N} \sum_{m=1}^{N} \sigma_{\ell m} M_{\ell *} M_m = \sum_{m=1}^{N} \sum_{\ell=1}^{N} \sigma_{m\ell} M_{\ell *} M_m = \Phi , \tag{61}$$

and condition a_1) is satisfied.

Because G_p is physically realizable and rational it is true that the $\frac{\partial G_p}{\partial \alpha_k}$ are analytic in Re s $\geq$ 0. Condition d_3) insures, therefore, that the M_k are analytic on the finite s = jω axis. It immediately follows that Φ is analytic on the finite s = jω axis and condition a_2) is satisfied. It only remains to show that Φ satisfies condition a_3).

Equation (59) is equivalent to

$$\Phi = M_* \quad \tilde{\Sigma} \; M \tag{62}$$

where

$$M = [M_1' \,|M_2'|\ldots|M_N' \,]' \tag{63}$$

and

$$\tilde{\Sigma} = \Sigma \times 1_n = [\sigma_{ij} \; 1_n \,] \, . \tag{64}$$

The matrix $\tilde{\Sigma}$ is a Kronecker product (see page 227 of [12]). Since both Σ and 1_n are positive definite or, equivalently, have only positive eigenvalues, it follows that all eigenvalues of $\tilde{\Sigma}$ are positive. Thus, $\Phi(j\omega)$ is positive definite for every finite ω if, and only if, there exists no n-dimensional non-zero column vector a and no finite ω for which $M(j\omega) \; a = 0_{(nN)}$. This is the case if, and only if, the rank of $M'(j\omega)$ is n for all finite ω. Now

$$M' = \nabla \; \text{diag}\,[G_p^{-1}, \; G_p^{-1}, \ldots, G_p^{-1} \,] \; \text{diag}\,[V', V', \ldots, V' \,], \tag{65}$$

and it immediately follows from d_5) that rank $M'(j\omega)$ is indeed n for all finite ω provided V and G_p are nonsingular in the finite $s = j\omega$ axis. That $V(j\omega)$ is nonsingular follows from the fact that the right-hand side of (58) is positive definite on the finite $s = j\omega$ axis. Arguments identical with those used to establish the equivalence of c_2) and c_3) can be used with d_3) to establish that $G_p(j\omega)$ is nonsingular for all finite ω.

It is not difficult to verify that the plant transfer functic . matrix has the form

$$G_p(s, a) = G(s) \; \text{diag}\,[a_1, \; a_2, \ldots, a_n\,] \tag{66}$$

when the plant is described by the vector differential equation

$$\left.\begin{aligned} \dot{x} &= A\,x + B \; \text{diag}\,[a_1, a_2, \ldots, a_n\,]\,u \\ c &= D\,x \; . \end{aligned}\right\} \tag{67}$$

The δa_i then represent variations in control effort gains. It is now shown for the case in which G^{-1} is analytic on the finite $s = j\omega$ axis and $a_i \neq 0$

for each $i = 1, 2, \ldots, n$, that condition d_5) is satisfied. The α_i are nonzero in keeping with the fact that only plant transfer function matrices with normal rank n need be considered. The analyticity of G^{-1} on the finite $s = j\omega$ axis assures the satisfaction of d_3).

When the j-column of G is denoted by the column vector $\underset{\sim}{g}_j$, then one can write

$$\frac{\partial G_p}{\partial \alpha_j} = \underset{\sim}{g}_j \underset{\sim}{e}_j' \quad . \tag{68}$$

Substituting (68) into the matrix ∇ defined in d_5) establishes that the only nonzero columns of ∇ are the columns $\underset{\sim}{g}_j$, $j = 1, 2, \ldots, n$. Since G^{-1} is analytic on the finite $s = j\omega$ axis, it follows that the minor of order n formed from these n columns of ∇ is nonzero on the finite $s = j\omega$ axis. The rank of ∇ on the finite $s = j\omega$ axis is n, therefore, and condition d_5) is satisfied.

Physical Realizability of $G_c(s)$

The above developments are concerned with the determination of H_o once W is specified. Attention is now turned to the computation of G_c, and the determination of conditions which guarantee that G_c is physically realizable. Solving (3) with $H = H_o$ for G_c and using (46) in the result yields

$$G_c = G_p^{-1}(1_n - WH_o)^{-1} W = G_p^{-1} \Delta' L^{-1} \Gamma, \tag{69}$$

where

$$L = \{ \Gamma W^{-1} \Delta' \}_- \quad . \tag{70}$$

It is clear from (69) that G_c is not, generally, physically realizable. The matrices G_p^{-1} and L^{-1} can have poles in Re $s \geq 0$. The problem facing the designer and the one discussed here is the specification of W so that the G_c given by (69) is physically realizable.

Any W satisfying conditions c_1) and c_2) or c_3) can be written in the form

$$W = \frac{q}{q_*} \tilde{W} , \tag{71}$$

where $q = q(s)$ is a monic polynomial with zeros in Re $s > 0$ only. The zeros of q include any zeros in Re $s > 0$ common to every element of W and, also, any other possible poles of W^{-1} in Re $s > 0$. From c_1), c_3), and the definition of q, it follows that $\tilde{W}^{-1}$ is analytic in Re $s \geq 0$. Substituting (71) into (70), recalling that Γ and Δ are analytic in Re $s \geq 0$, and collecting terms in the partial fraction expansion of $\Gamma W^{-1} \Delta'$ associated with the zeros of q one obtains

$$L = \frac{\tilde{L}}{q} , \tag{72}$$

where $\tilde{L}$ is a polynomial matrix. When for example the zeros of q are all simple and denoted by s_i, $i = 1, 2, \ldots, \ell$, then

$$\tilde{L} = \sum_{i=1}^{\ell} \left[\left. \frac{(s - s_i)\, q_* \, \Gamma \tilde{W}^{-1} \Delta'}{q} \right|_{s = s_i} \right] \frac{q}{(s - s_i)} . \tag{73}$$

From (69) and (72) it follows that

$$G_c = q\, G_p^{-1} \; \Delta' \; \tilde{L}^{-1} \, \Gamma = q \, \tilde{G}_c \tag{74}$$

is physically realizable if, and only if the poles of $\tilde{G}_c$ in Re $s \geq 0$ are cancelled by zeros of q.

It is of interest to examine the case in which G_p^{-1} has only one pole in Re $s \geq 0$ at $s = \sigma_o > 0$, σ_o real. When it is possible to meet the dynamic performance requirements placed on the system with

$$W = - \left(\frac{s - \sigma_o}{s + \sigma_o} \right) \tilde{W} , \tag{75}$$

where $\tilde{W}$ and $\tilde{W}^{-1}$ are analytic in Re $s \geq 0$, then one can choose $\Gamma = -\tilde{W}$ and obtain from (73)

$$\tilde{L} = 2 \, \sigma_o \, \Delta' (\sigma_o). \tag{76}$$

It now follows from (74) that

$$G_c = - \frac{1}{2 \, \sigma_o} (s - \sigma_o) \, G_p^{-1} \, \Delta' \, [\Delta' \; (\sigma_o)]^{-1} \, \tilde{W} . \tag{77}$$

Since Δ^{-1} is analytic in Re $s \geq 0$, the matrix $[\Delta' \, (\sigma_o)]^{-1}$ is finite. Moreover, $(s - \sigma_o) \, G_p^{-1}$, Δ, and $\tilde{W}$ are analytic in Re $s \geq 0$, and G_c is, therefore, physically realizable.

Another interesting case for which general conclusions can be drawn occurs when G_p^{-1} has no poles in Re s $\geq$ 0, and the dynamic performance requirements of the system are met by a physically realizable W whose inverse is analytic in Re s $\geq$ 0. Under these conditions

$$\{\Gamma W^{-1} \Delta'\}_+ = \Gamma W^{-1} \Delta' \tag{78}$$

and

$$\{\Gamma W^{-1} \Delta'\}_- = O_{nn} \, . \tag{79}$$

Hence, (40) reduces to $H_o = W^{-1}$, and (57) leads to min $S = 0$. Substituting $H_o = W^{-1}$ into (69) indicates that

$$G_c = G_p^{-1} \; K W \tag{80}$$

which aside from the fact that K is a constant matrix each of whose elements are infinitely large is physically realizable. That is, the sensitivity can be made arbitrarily small at the expense of high gains in the tandem compensation network. This is simply a generalization to the multivariable case of the well known result for single-input-output systems that overall system sensitivity to plant parameter variations can be made arbitrarily small when the specified plant and overall system transfer functions are minimum phase.

Conclusions

The approach taken in this paper is significant in that it leads to the analytical design of minimum sensitivity feedback systems. The method is applicable when the plant is a linear, time-invariant, lumped, finite-dimensional dynamical system and the uncertain plant parameters can be viewed as random variables. A part of the development important in its own right is the solution of the semi-free-configuration Wiener problem for the multivariable case.

Acknowledgement

The author wishes to thank Professor D. C. Youla for his interest in and help with the study reported herein.

Appendix D

NOISE-INTENSITY SENSITIVITY IN OPTIMAL STOCHASTIC SYSTEMS*

1. Introduction

In the design of optimal control systems for plants with random attributes (noise signals or randomly varying parameters), a question of some interest is the effect of the noise intensity on the performance index. For plants where zero noise-intensity reduces the system to a deterministic one, a noise-intensity sensitivity analysis yields some insight into the first order effects of noise on a deterministic design.

Consider a plant characterized by the stochastic differential equation

$$d\underset{\sim}{x} = \underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{u})\,dt + \underset{\sim}{\sigma}(\underset{\sim}{x})\,d\xi, \tag{1}$$

where

$\underset{\sim}{x}$ - n dimensional state vector,

$\underset{\sim}{u}$ - m dimensional control input vector,

ξ_t - scalar Wiener Process with variance parameter α, i.e.,

$$E\left\{(\xi_{t+\Delta t} - \xi_t)^2\right\} = \alpha\,\Delta t,$$

$\underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{u})$, $\underset{\sim}{\sigma}(\underset{\sim}{x})$ - n dimensional vector functions of $\underset{\sim}{x}$ and $\underset{\sim}{u}$.

For more details on this type of representation of stochastic systems, including mathematical conditions for the existence and uniqueness of solutions of equation (1), see Kushner [1]. Only a single noise source is considered here to simplify the subsequent computation, however, the results can easily be extended to multiple noise sources.

Assume that a performance index of the form

$$C = E_x\left\{\int_t^T k(\underset{\sim}{x}_s, \underset{\sim}{u}_s)\,ds + b(\underset{\sim}{x}_T)\right\}, \tag{2}$$

*Paper presented by P. Dorato at the Second IFAC Symposium on System Sensitivity and Adaptivity, Dubrovnik, Yugoslavia, Aug. 26-31, 1968.

is given, where

$k(\underset{\sim}{x}_s, \underset{\sim}{u}_s)$, $b(\underset{\sim}{x}_T)$ - non-negative loss functions,

$E_x\{\cdot\}$ - expectation operator conditioned on $\underset{\sim}{x} = \underset{\sim}{x}_t$

T - fixed end time.

The optimal stochastic control problem is to determine a control law $\underset{\sim}{u} = \underset{\sim}{\Phi}(x, t)$, within a class of admissible control inputs, which minimizes the performance index (2). As is well known [1] the optimization equation appropriate to the above problem is given by the stochastic Hamilton-Jacobi equation

$$0 = \min_{u} \left[k(\underset{\sim}{x}, \underset{\sim}{u}) + \underset{\sim}{f}'(\underset{\sim}{x}, \underset{\sim}{u}) V_x + \frac{\alpha}{2} \underset{\sim}{\sigma}'(\underset{\sim}{x}) V_{xx} \underset{\sim}{\sigma}(\underset{\sim}{x}) + V_t \right], \tag{3}$$

with the boundary condition

$$V(\underset{\sim}{x}, T) = 0, \tag{4}$$

for all $\underset{\sim}{x}$. In (3), V_x is a column vector with entries $\left[\frac{\partial V}{\partial x_i}\right]$ and V_{xx} is a matrix with entries $\left[\frac{\partial^2 V}{\partial x_i \partial x_j}\right]$, $\underset{\sim}{f}'$ denotes the transpose of $\underset{\sim}{f}$, and V_t denotes the partial derivative with respect to t.

Let $\underset{\sim}{\Phi}^o(\underset{\sim}{x}, t)$ denote the value of $\underset{\sim}{u}$ which minimizes the bracketed term in (3). Then the minimal value of C is given by the solution $V(\underset{\sim}{x}, t)$ of the equation

$$0 = k(\underset{\sim}{x}, \underset{\sim}{\Phi}^o) + \underset{\sim}{f}'(\underset{\sim}{x}, \underset{\sim}{\Phi}^o) V_x + \frac{\alpha}{2} \underset{\sim}{\sigma}'(\underset{\sim}{x}) V_{xx} \underset{\sim}{\sigma}(\underset{\sim}{x}), + V_t, \tag{5}$$

subject to the boundary condition (4).

A solution of (5) yields performance value V which depends on $\underset{\sim}{x}$, t, and the noise parameter α, i.e., $V = V(\underset{\sim}{x}, t, \alpha)$. This then implies that the control law $\underset{\sim}{\Phi}^o(\underset{\sim}{x}, t)$ also depends on $\underset{\sim}{x}$, t, and α. The sensitivity problem considered here relates to the study of the variations in the performance index (2) with variations in the variance parameter α in the dynamics (1), for the fixed control law $\underset{\sim}{\phi}^{o}(\underset{\sim}{x}, t)$

2. Sensitivity Equation

Let the nominal value of noise parameter be α and the perturbed value be $\alpha + \Delta\alpha$. It is assumed that the only parameter variation of significance is in the plant equations, so that the control law can be considered fixed once the nominal value α, is fixed. In this case the equation which the performance sensitivity function [2]

$$v = \frac{\partial V}{\partial \alpha} \quad ,$$

satisfies may be formally determined by taking the partial derivative of (5) with respect to α , and interchanging derivatives with respect to α and $\underset{\sim}{x}$. This yields the sensitivity equation

$$0 = \left[\underset{\sim}{f}' v_x + \frac{1}{2} \underset{\sim}{\sigma}' v_{xx} \underset{\sim}{\sigma} + v_t + \frac{\alpha}{2} \underset{\sim}{\sigma}' v_{xx} \underset{\sim}{\sigma} \right], \tag{6}$$

where in (6) V_{xx} is evaluated at the nominal value of the variance parameter.

Equation (6) is a linear second order partial differential equation for the sensitivity function $v = v(\underset{\sim}{x}, t, \alpha)$. The boundary condition for (6) is

$$v(\underset{\sim}{x}, T, \alpha) = 0 , \tag{7}$$

for all $\underset{\sim}{x}$. The boundary condition (7) results from the fact that at the termination point the value of v is identically zero.

3. Some Solutions to the Sensitivity Equation

Even though the sensitivity equation (6) is linear, solutions to this equation are difficult to obtain in general*. In this section certain linear problems will be considered where at least a partial solution is possible.

Problem 1. (Additive Noise)

Plant dynamics:

$$d\,\underset{\sim}{x} = A\,\underset{\sim}{x}\,dt + B\,\underset{\sim}{u}\,dt + \underset{\sim}{g}\,d\,\xi\,,$$

where A and B are matrices and $\underset{\sim}{g}$ is a column vector.

Performance index:

$$C = E_x \left\{ \int_t^T (\underset{\sim}{x}_s' C\,\underset{\sim}{x}_s + \underset{\sim}{u}_s' D\,\underset{\sim}{u}_s)\,ds \right\} .$$

As is well known [1], [3] the optimal control law for this stochastic problem is

$$\underset{\sim}{\Phi}^o = -\frac{D^{-1}B}{2}\,V_x\,,$$

where $V = V(\underset{\sim}{x}, t)$ satisfies the equation

$$0 = \left[\underset{\sim}{x}' C\,\underset{\sim}{x} + (\underset{\sim}{\Phi}^o)' D\,\underset{\sim}{\Phi}^o + V_t + (A\underset{\sim}{x} + B\,\underset{\sim}{\Phi}^o)' V_x + \frac{a}{2}\,g' V_{xx}\underset{\sim}{g} \right], \qquad (9)$$

with the boundary condition

$$V(\underset{\sim}{x}, T) = 0\,.$$

The solution of (9) is given by [1], [3]

$$V(\underset{\sim}{x}, t) = \underset{\sim}{x}' P(t)\,\underset{\sim}{x} + q(t),$$

where P(t) satisfies the equation

---

* It should be noted that since the sensitivity function v satisfies a diffusion type equation [1], it is possible to determine $v(\underset{\sim}{x}, t)$ from Monte Carlo methods.

$$\dot{P} + A'P + PA + C - P(BD^{-1}B')P = 0, \quad P(T) = 0,$$

and q(t) satisfies the equation

$$\dot{q} + \alpha \underset{\sim}{g}' P(t) \underset{\sim}{g} = 0, \; q(t) = 0.$$

Since P(t) is independent of α and

$$\underset{\sim}{\Phi}^{o}(\underset{\sim}{x}) = -D^{-1}B'P(t)\underset{\sim}{x},$$

it follows that for this particular problem, the control law is independent of the noise intensity. This result is a consequence of the additive nature of the noise signal ξ_t in plant dynamics (8). The results are quite different for the multiplicative-noise case considered in the next problem. The sensitivity equation (6) for this problem becomes, after some manipulation,

$$0 = \left[\underset{\sim}{x}'(A - BD^{-1}B'P)'v_x + \underset{\sim}{g}'P\underset{\sim}{g} + v_t + \frac{\alpha}{2}\underset{\sim}{g}'v_{xx}\underset{\sim}{g} \right]. \tag{10}$$

The solution of (10) is given by

$$v(\underset{\sim}{x}, t) = \underset{\sim}{x}' S(t) \underset{\sim}{x} + r(t),$$

where S(t) and r(t) satisfy the equations

$$\begin{cases} \dot{S} + (A - BD^{-1}B'P)'S + S(A - BD^{-1}B'P) = 0 \\ \dot{r} + \underset{\sim}{g}'P\underset{\sim}{g} + a\,\underset{\sim}{g}'S\underset{\sim}{g} = 0 \end{cases} \tag{11}$$

with the boundary conditions

$$S(\underset{\sim}{x}, T) = 0, \quad r(t) = 0.$$

Since the equation for S in (11) is linear and homogeneous and $S(\underset{\sim}{x}, T) = 0$, it follows that $S(\underset{\sim}{x}, t) = 0$. Thus $v(\underset{\sim}{x}, t)$ becomes simply $v(\underset{\sim}{x}, t) = r(t)$ where r(t) satisfies the equation, $\dot{r} + \underset{\sim}{g}'P\underset{\sim}{g} = 0$.

Problem 2. (Multiplicative Noise)

Plant dynamics

$$d\underset{\sim}{x} = A\underset{\sim}{x}\,dt + B\underset{\sim}{u}\,dt + G\underset{\sim}{x}\,d\xi$$

where A, B, and G are matrices.

Performance Index: Same as problem one.

The optimal control law for this problem is given by

$$\underset{\sim}{\Phi}^{o} = -\frac{D^{-1}B'V_x}{2},$$

where V_x satisfies the equation (9) with $\underset{\sim}{g}$ replaced by $G\underset{\sim}{x}$. The solution of this equation is known [3], [4] to be given by

$$V(\underset{\sim}{x}, t) = \underset{\sim}{x}'\,P(t)\,\underset{\sim}{x},$$

where P(t) satisfies the equation

$$\dot{P} + A'P + PA + C + \alpha G'PG - P(BD^{-1}B')P = 0, \qquad (12)$$

with

$$P(T) = 0.$$

Since

$$\Phi^{o} = -D^{-1}B'P\underset{\sim}{x},$$

and P(t), from (12), depends on α the control law is dependent on the noise intensity. The sensitivity equation for this problem is,

$$0 = \left[\underset{\sim}{x}'(A - BD^{-1}B'P)\,v_x + \underset{\sim}{x}'G'PG\underset{\sim}{x} + v_t + \frac{\alpha}{2}\underset{\sim}{x}'G'v_{xx}G\underset{\sim}{x}\right]. \qquad (13)$$

The solution of (13) is given by

$$v(\underset{\sim}{x}, t) = \underset{\sim}{x}' S(t) \underset{\sim}{x} ,$$

where S(t) satisfies the equation

$$\dot{S} + (A-BD^{-1}B'P) S + S(A-BD^{-1}B'P) + G'PG + \alpha G'SG = 0 , \qquad (14)$$

with S(x, T) = 0. Here, as in problem one P(t) is evaluated at the nominal value of noise intensity α .

Example:

Consider the scalar system

$$dx = (-x + u) dt + x d\xi ,$$

with a performance index

$$C = E_x \left\{ \int_0^\infty (x^2 + u^2) ds \right\} .$$

This represents a control problem with a time constant equal to -1 perturbed by "white noise" (derivative of Wiener process). For this multiplicative noise problem equation (12) becomes*

$$p^2 - (2 + \alpha) p - 1 = 0 ,$$

with positive solution

$$p = \left(1 + \frac{\alpha}{2}\right) + \sqrt{\left(1 + \frac{\alpha}{2}\right)^2 + 1}$$

The sensitivity equation (14) becomes

$$[2(1-p) + \alpha] s + p = 0 ,$$

where p is given above. The solution of this sensitivity equation is

---

* For T = ∞ , $\dot{P}$ and $\dot{S}$ are zero and P is the positive definite solution of the resulting algebraic equation (12).

$$s = \frac{1}{2\sqrt{\left(1 + \frac{a}{2}\right)^2 + 1}}$$

Let s_N denote a normalized sensitivity parameter defined by $s_N = {}^{s}/C$, where C is the optimal cost at the nominal setting. Since $C = p$, s_N can be written

$$s_N = \frac{1}{2\sqrt{\left(1 + \frac{a}{2}\right)^2 + 1}}$$

It is interesting to note in this case that the normalized sensitivity is greatest for the no noise case ($a = 0$) and decreases monotonically with increasing noise intensity.

4. Conclusion

A sensitivity equation is derived for the analysis of optimal performance index sensitivity to variations in the variance parameter of a random disturbance signal. Since statistical parameters, such as variances, are never known exactly, this type of sensitivity analysis is essential in the practical design of optimal stochastic systems. It is difficult to arrive at general conclusions from the sensitivity equation (6), since an analytic solution of this equation is not available except for some low order problems. However, the first order example worked out in section 3, indicates that performance index sensitivity decreases with increasing values of variance. This appears to be a reasonable conclusion since one would expect the design of a system for large noise perturbation signals to be less sensitive to these perturbation signals.

Appendix E
Proof of Theorem 1

First note that $V(\underset{\sim}{x}, \underset{\sim}{\alpha})$ satisfies the partial differential equation

$$\frac{\partial V}{\partial \underset{\sim}{x}} \underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{\alpha}, \underset{\sim}{\Phi}(\underset{\sim}{x})) + L(\underset{\sim}{x}, \underset{\sim}{\Phi}(\underset{\sim}{x})) = 0 \tag{1}$$

for all $\underset{\sim}{x}$ outside of D and

$$V(\underset{\sim}{x}) = 0 , \tag{2}$$

for $\underset{\sim}{x}$ on the boundary of D. The boundary condition (2) follows directly from the definition of τ_D, that is if $\underset{\sim}{x}$ is on the boundary of D, then $\tau_D = 0$ and $V(\underset{\sim}{x}) = 0$. The partial differential equation (1) may be derived as follows. From the integral form of $V(\underset{\sim}{x}, \underset{\sim}{\alpha})$, i.e.,

$$V(\underset{\sim}{x}, \underset{\sim}{\alpha}) = \int_0^{\tau_D} L(\underset{\sim}{x}, \underset{\sim}{\Phi}(\underset{\sim}{x}))\, dt$$

it follows that one can write

$$V(\underset{\sim}{x}, \underset{\sim}{\alpha}) = \int_0^{\Delta t \wedge \tau_D} L\, dt + \int_{\Delta t \wedge \tau_D}^{\tau_D} L\, dt, \quad \underset{\sim}{x}(o) = \underset{\sim}{x} . \tag{3}$$

where $\Delta t \wedge \tau_D$ denotes $\min(\Delta t, \tau_D)$. The second integral in (3) is given by

$$V(\underset{\sim}{x}(\Delta t \wedge \tau_D), \underset{\sim}{\alpha}) . \tag{4}$$

If now $\Delta t \to 0$ and a Taylor series expansion is made of (4), about the point $\underset{\sim}{x}(t) = \underset{\sim}{x}$, there results*

$$V(\underset{\sim}{x}, \underset{\sim}{\alpha}) = L\, \Delta t + V(\underset{\sim}{x}, \underset{\sim}{\alpha}) + \frac{\partial V}{\partial \underset{\sim}{x}} \Delta \underset{\sim}{x} + \ldots \tag{5}$$

---

* It is assumed here that

$$\lim_{\Delta t \to 0} (\Delta t \wedge \tau_D) = \Delta t ,$$

This is true, for example, if $\underset{\sim}{x}(t)$ is continuous in t.

If (5) is divided by Δt and the limit taken as $\Delta t \to 0$, equation (1) results. Conversely, if $V(\underset{\sim}{x}, \underset{\sim}{\alpha})$ satisfies the equation (1) and the boundary condition (2), then $V(\underset{\sim}{x}, \underset{\sim}{\alpha})$ is given by

$$V(\underset{\sim}{x}, \underset{\sim}{\alpha}) = \int_0^{\tau_D} L(\underset{\sim}{x}(t), \underset{\sim}{\Phi}(\underset{\sim}{x}(t))) \, dt , \tag{6}$$

where $\underset{\sim}{x}(t)$ is the solution of

$$\dot{\underset{\sim}{x}} = \underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{\alpha}, \underset{\sim}{\Phi}(\underset{\sim}{x})) \tag{7}$$

To see this let $\underset{\sim}{x}_{\Phi}$ be the solution of (7), with $\underset{\sim}{x}_{\Phi}(o) = \underset{\sim}{x}$; then, since (1) is satisfied for all $\underset{\sim}{x}$, outside of D, it is satisfied for $\underset{\sim}{x}_{\Phi}$, i.e.,

$$\frac{\partial V}{\partial \underset{\sim}{x}} \underset{\sim}{f}(\underset{\sim}{x}_{\Phi}, \underset{\sim}{\alpha}, \underset{\sim}{\Phi}(\underset{\sim}{x}_{\Phi})) + L(\underset{\sim}{x}_{\Phi}, \underset{\sim}{\Phi}(\underset{\sim}{x}_{\Phi})) = 0 \tag{8}$$

Now integration (8) yields

$$V(\underset{\sim}{x}_{\Phi}(\tau_D)) - V(\underset{\sim}{x}) + \int_0^{\tau_D(\Phi)} L(\underset{\sim}{x}_{\Phi}, \underset{\sim}{\Phi}(\underset{\sim}{x}_{\Phi})) \, dt = 0 , \tag{9}$$

since the integrand in the first term of (8) is a perfect differential. From the boundary condition (2), $V(\underset{\sim}{x}_{\Phi}(\tau_D)) = 0$, and the required result follows (the value of $\tau_D(\Phi)$ corresponds to the time required for $\underset{\sim}{x}_{\Phi}$ to reach D).

Consider now the inequality

$$\frac{\partial V}{\partial \underset{\sim}{x}} \underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{\alpha}, \underset{\sim}{v}) + L(\underset{\sim}{x}, \underset{\sim}{v}) < 0 , \tag{10}$$

where $\underset{\sim}{v} = \underset{\sim}{\psi}(\underset{\sim}{x})$ and where $V = V(\underset{\sim}{x}, \underset{\sim}{\alpha})$ satisfies (1) and (2). Let

$\underset{\sim}{x}_{\psi}$ be the solution of

$$\dot{\underset{\sim}{x}} = \underset{\sim}{f}(\underset{\sim}{x}, \underset{\sim}{a}, \underset{\sim}{v})$$

with $\underset{\sim}{v} = \underset{\sim}{\psi}(\underset{\sim}{x})$ and substitute $\underset{\sim}{x}_{\psi}$ in (10). Let $\tau_D(\psi)$ be the time required for $\underset{\sim}{x}_{\psi}$ to first reach D. Then (10) becomes after integrating

$$V(\underset{\sim}{x}_{\psi}(\tau_D)) - V(\underset{\sim}{x}) + \int_0^{\tau_D} L(\underset{\sim}{x}_{\psi}(t), \underset{\sim}{\psi}(\underset{\sim}{x}_{\psi}(t))) \, dt < 0 \; . \tag{11}$$

But $V(\underset{\sim}{x}_{\psi}(\tau_D)) = 0$ since $\underset{\sim}{x}_{\psi}(\tau_D)$ is in D and V satisfies (1); therefore,

$$\int_0^{\tau_D} L(\underset{\sim}{x}_{\psi}(t), \underset{\sim}{\psi}(\underset{\sim}{x}_{\psi}(t))) \, dt < V(\underset{\sim}{x}) = \int_0^{\tau_D} L(\underset{\sim}{x}_{\Phi}(t), \underset{\sim}{\Phi}(\underset{\sim}{x}_{\Phi}(t))) \, dt \; . \tag{12}$$

REFERENCES (Body of Report)

[1] W. Redisch, "Equilibrium Analysis and its Application to Attitude Control Systems", NASA TR R - 270, August, 1967.

[2] A. Sabroff, et. al., "Investigation of the Acquisition Problem in Satellite Attitude Control", Tech. Report AFFDL-TR-65-115, December, 1965.

[3] D. A. Salmon, "Minimax Controller Design", IEEE Trans. on Automatic Control, Vol. AC-13, No. 4, pp 369-376, August, 1968.

[4] S. Karlin, "Mathematical Methods and Theory in Games, Programming, and Economics", Vol. II, Addison-Wesley Co. Inc., Reading, Mass., 1959.

[5] M. Athans and P. T. Falb, "Optimal Control", Chap. 9, McGraw-Hill, New York, 1966.

[6] D. C. Youla, "Fundamentals of Optimal Control Theory", Lecture notes, Course EE 674, Polytechnic Institute of Brooklyn.

[7] J. J. Bongiorno, Jr., "Minimum Sensitivity Designs", IEEE Convention Record, March, 1967.

[8] D. C. Youla, "On the Factorization of Rational Matrices", IRE Trans. on Information Theory, pp. 172-189, July, 1961.

[9] J. J. Rissanen, " Performance Deterioration of Optimal Systems", IEEE Trans. on Automatic Control, Vol. AC-11, No. 3, pp. 530-532, July, 1966.

[10] P. Dorato, "Noise-Intensity Sensitivity in Optimal Stochastic Systems", 2nd IFAC Symposium on System Sensitivity and Adaptivity, Dubrovnik, Yugoslavia, Auguest 26-31, 1968.

REFERENCES (Appendix A).

[1] J. LaSalle, "Stability by Liapunov's Direct Method", Academic Press, New York, 1961.

[2] D. C. Youla, "Fundamentals of Optimal Control Theory", Lecture Notes, Polytechnic Institute of Brooklyn, 1966.

[3] W. Hahn, "Stability of Motion", Springer-Verlag Inc., New York, 1967.

REFERENCE (Appendix B).

[1] R. J. Duffin, E. L. Peterson, C. Zener, "Geometric Programming - Theory and Application", Wiley, New York, 1967.

REFERENCES (Appendix C).

[1] J. J. Bongiorno, Jr., "Minimum Sensitivity Designs," IEEE International Convention Record, Vol. 15, Part 3, pp. 129-135, March, 1967.

[2] M. W. Wonham, "On Pole Assignment in Multi-Input Controllable Linear Systems," IEEE Trans. on Automatic Control, Vo. AC-12, No. 6, pp 660-665, December, 1967.

[3] J. J. Bongiorno. Jr. and D. C. Youla, "On Observers in Multi-Variable Control Systems," International Journal of Control, Vol. 8, No 3, pp. 221-243, September 1968.

[4] J. B. Cruz, Jr. and W. R. Perkins, "A New Approach to the Sensitivity Problem in Multivariable Feedback System Design," IEEE Trans. on Automatic Control, Vol. AC-9, pp. 216-223, July, 1964.

[5] J. B. Cruz, Jr. and W. R. Perkins, "The Parameter Variation Problem in State Feedback Control Systems," ASME Journal of Basic Engineering, Vol. 87, Series D, No. 1, pp. 120-124, March, 1965.

[6] W. R. Perkins, J. B. Cruz, Jr., and R. L. Gonzales, "Design of Minimum Sensitivity Systems," IEEE Trans. on Automatic Control, Vol. AC-13, No. 2, pp. 159-167, April, 1968.

[7] H. C. Hsieh and C. T. Leondes, "Techniques for the Optimum Synthesis of Multipole Control Systems with Random Processes as Inputs", IRE Trans. on Automatic Control Systems, Vol. AC-4, No. 3, pp. 212-231, December, 1959.

[8] M. C. Davis, "Factoring the Spectral Matrix," IEEE Trans. on Automatic Control, Vol. AC-8, No. 4, pp. 296-305, October, 1963.

[9] D. C. Youla, "On the Factorization of Rational Matrices," IRE Trans. on Information Theory, Vol. IT-7, No. 3, pp. 172-189, July, 1961.

[10] W. G. Tuel, Jr., "Computer Algorithm for Spectral Factorization of Rational Matrices," IBM Journal of Research and Development, Vol. 12, No. 2, pp. 163-170, March, 1968.

[11] S. S. L. Chang, Synthesis of Optimal Control Systems, McGraw-Hill, New York, 1961.

[12] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.

[13] W. M. Mazer, "Specification of the Linear Feedback System Sensitivity Function," IRE Trans. on Automatic Control, Vol. AC-5, No. 2, pp. 85-93, June, 1960.

REFERENCES (Appendix D).

[1] H. J. Kushner, "Stochastic Stability and Control," Academic Press, 1967.

[2] B. Pagurek, "Sensitivity of the Performance of Optimal Control Systems to Plant Parameter Variations", Sensitivity Methods in Control Theory, L. Radanovic (Editor), Pergamon Press, 1966.

[3] W. M. Wonham, "Optimal Stationary Control of a Linear System with State-Dependent Noise", SIAM Journal Control, Vo. 5, no. 3, 1967, pp. 486-500.

[4] W. Gersch and F. Kozin, "Optimal Control of Multiplicatively Perturbed Stochastic Systems," Proc. 1963, Allerton Conf. on Circuit and System Theory, University of Illinois.