Design Method for Minimizing Sensitivity to Plant Parameter Variations

Zeev Hadass*
Ministry of Defense, Haifa, Israel

and

J. David Powell†
Stanford University, Stanford, Calif.

A method is described for minimizing the sensitivity of multivariable systems to parameter variations. The variable parameters are considered as random variables and their effect is included in a quadratic performance index. The performance index is a weighted sum of the state and control covariances that stem from both the random system disturbances and the parameter uncertainties. The numerical solution of the problem is described and application of the method to several initially sensitive tracking systems is discussed. The sensitivity factor of reduction was typically 2 or 3 over a system based on random system noise only, and yet resulted in state rms increases of only about a factor of two.

I. Introduction

THE design of controllers for linear time invariant systems has become well advanced in recent years. Both time domain and frequency domain methods are well developed and a considerable body of experience has accumulated. Unfortunately, in most realistic design problems, the system is not linear nor is it time invariant. However, in many uses, if the nonlinearities are mild and the time variation of the parameters is bounded, a linear time invariant approximation of the system can be used for controller design. Both the design techniques and the resulting controllers are then much simpler.

Using such an approximation, however, gives rise to the problem of parameter sensitivity. A controller that performs adequately when the plant parameters have their assumed values, may become unsatisfactory when these parameters vary. The desirability of low parameter sensitivity is therefore obvious. This sensitivity may be defined, depending on the way in which the performance criteria are given, as trajectory sensitivity,[1] eigenvalue sensitivity,[2] or transfer function sensitivity[3]

In the present paper, the sensitivity of controllers designed by time domain techniques is examined, and a method for its reduction is developed. Special emphasis is placed on the investigation of state feedback controllers in which the states that are used for feedback are estimated rather than measured. This is the case for most such controllers since, in general, the number of measurements is much less than the order of the system.

The problem of sensitivity is treated extensively in the literature. A collection of the most representative papers dealing with this subject was recently published by Cruz.[4] Most of these papers, and others that are not included in the collection, treat various theoretical aspects of the problem. Design papers, in which methods of sensitivity reduction are described, are scarce.

For controllers that are designed by time domain techniques, the most common approach given in the literature for sensitivity minimization (for example[5-7] is the augmentation of the state by the trajectory sensitivity vector σ defined as $\sigma = dx/d\mu$ where x is the state vector and μ is the variable parameter.

A performance index for this augmented system is defined and the control structure is specified. It may be full-state feedback (for the augmented system)[5,6] or feedback of some or all of the plant states,[8,9] The feedback gains that minimize the performance index are determined by computational or experimental methods.

The application results for these methods are mostly given for low-order systems. The results that are reported for higher order systems seem to be inconclusive.[10] Also, these methods are computationally difficult to apply to systems which have more than one variable parameter. No published results have been found concerning the effect of using state estimate feedback rather than state feedback in conjunction with this method.

A different approach is used by Palsson and Whittaker[11] for the sensitivity minimization of single-input single-output systems subjected to step inputs and which may have several variable parameters. In this approach, the variable parameters are considered as random variables with known means and variances. A performance index of the form

$$J = \int_0^\infty x^T A x \mathrm{d}t$$

is minimized where the mean is taken over the values of the variable parameters. The minimization is accomplished by adjusting a set of parameters that are at the designers disposition such as feedback gains. The method developed in the present paper is based on the Palsson-Whittaker approach but is applicable to multivariable systems subject to arbitrary inputs.

The organization of the paper is as follows: In Sec. II, the sensitivity of state estimate feedback controllers is examined and illustrated by means of an example. In Sec. III, the sensitivity reduction method is described. Section IV contains application results of this method. Conclusions are given in Sec. V.

Presented as Paper 74-927 at the AIAA Mechanics and Control of Flight Conference, Anaheim, California, August 5-9, 1974; submitted September 3, 1974; revision received March 12, 1975. This work was partially supported by NASA under Contract NGR-05-020-019.

Index categories: Navigation, Control, and Guidance Theory; Spacecraft Attitude Dynamics and Control.

*Research Engineer.

†Assistant Professor of Aeronautics and Astronautics. Associate Fellow AIAA.

II. Sensitivity of State Estimate Feedback Controllers

Consider the system

$$\dot{x} = Fx + Gu + \Gamma w \tag{1a}$$

$$y = Hx + v \tag{1b}$$

where w = process noise and Γ = disturbance distribution matrix. The control is defined as

$$u = -C\hat{x} + u_0 \tag{2}$$

where u is the matrix of controls, C is the state feedback matrix, $\hat{x}$ is the state vector of an estimator, and u_0 is an external control. The estimator dynamic equation is

$$\dot{\hat{x}} = F_n\hat{x} + G_n u + K(y - H_n\hat{x}) \tag{3}$$

where K is the estimator gain matrix, and the subscript n signifies the nominal value of the plant parameters.

The dynamic equations of the augmented system consisting of the system and the estimator are obtained by combining Eqs. (1-3) and by using

$$\tilde{x} = x - \hat{x}$$

When the plant parameters have their nominal values these equations are

$$\begin{bmatrix} \dot{x} \\ \dot{\tilde{x}} \end{bmatrix} = \begin{bmatrix} F_n - G_nC & G_nC \\ 0 & F_n - KH_n \end{bmatrix} \begin{bmatrix} x \\ \tilde{x} \end{bmatrix} + \begin{bmatrix} \Gamma \\ \Gamma \end{bmatrix} w + \begin{bmatrix} 0 \\ K \end{bmatrix} v + \begin{bmatrix} G_n \\ 0 \end{bmatrix} u_0 \tag{4}$$

The eigenvalues of the controller $(F-GC)$ and of the estimate error $(F-KH)$ appear separately in the augmented system.

If the plant parameters differ from their assumed values so that

$$F = F_n + \delta F \tag{5a}$$

$$G = G_n + \delta G \tag{5b}$$

$$H = H_n + \delta H \tag{5c}$$

the dynamic equation of the augmented systems becomes

$$\begin{bmatrix} \dot{x} \\ \dot{\tilde{x}} \end{bmatrix} = \begin{bmatrix} F - GC & GC \\ \delta F - \delta GC - K\delta H & F_n \delta KH_n + \delta GC \end{bmatrix} \begin{bmatrix} x \\ \tilde{x} \end{bmatrix} + \begin{bmatrix} \Gamma \\ \Gamma \end{bmatrix} w - \begin{bmatrix} 0 \\ K \end{bmatrix} v + \begin{bmatrix} G \\ \delta G \end{bmatrix} u_0 \tag{6}$$

The salient fact that can be observed from this equation is that the parameter perturbations couple the state into the estimate error and therefore destroy the separation of the eigenvalues. This coupling in many instances causes the system to be much more sensitive to parameter variations than it would be if true state feedback were used. This is illustrated in the following example.

The plant used in this example is derived from the Stanford Relativity Satellite for which a controller was designed by J. Bull.[2] This satellite is shown in Fig. 1. It consists of an outer body in which a helium filled dewar is mounted elastically. The dewar contains a telescope which is also connected to it elastically. The attitude of the telescope is controlled by means of two controllers: 1) an actuator between the telescope and the dewar; 2) a thrustor mounted on the outer body.

The attitude of the telescope is measured. The actuator provides the high bandwidth precision control for the telescope, and the thrustor controls the outer body so that the relative attitude between the three bodies remains small.

The plant used in the example is a low-frequency approximation of the full system. This approximation is ob-

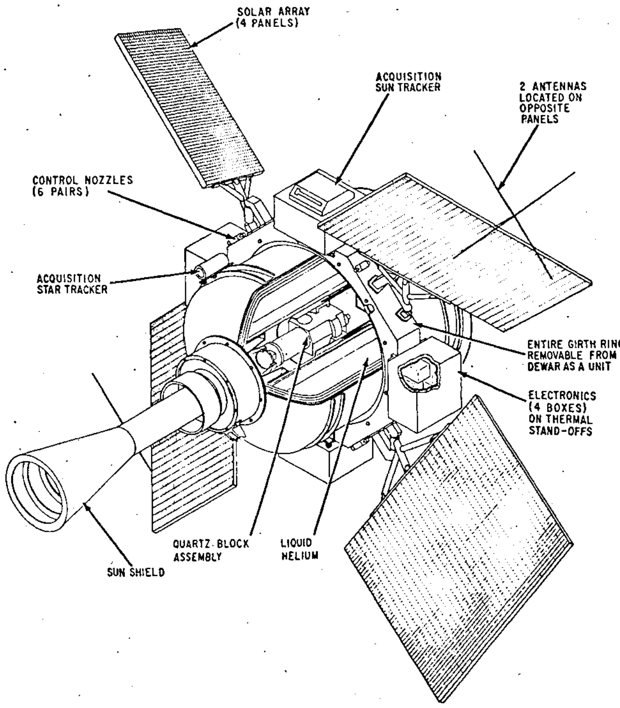

Fig. 1 Configuration of the Stanford relativity satellite.[12]

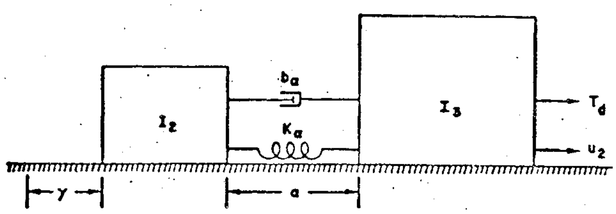

Fig. 2 Dynamic model of low frequency approximation of the Stanford relativity satellite.[12]

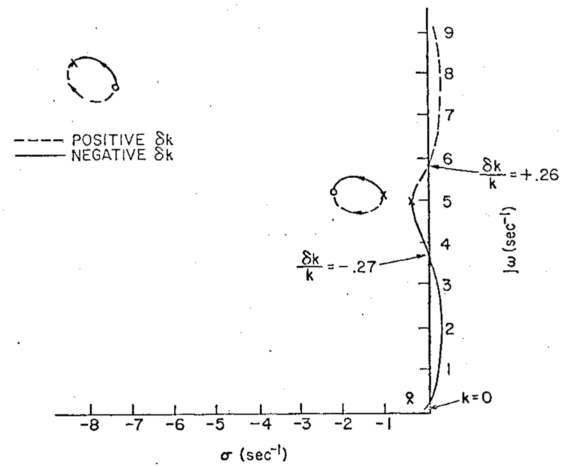

Fig. 3 Root locus vs δk for state estimate feedback controller.

tained by assuming that the spring force in the stiff inner spring is at all times cancelled by the actuator control force so that no net force is applied by the telescope on the middle body. The dynamic model of this approximation is shown in Fig. 2. The damping constant is low and may be neglected in the analysis.

The reasons for the use of this approximation rather than the full system are: a) It exhibits the same sensitivity properties as the full system to variations of the spring constant K_α, and it is simpler to analyze. Better understanding can therefore be obtained about the underlying reasons for its sensitivity, b) It demonstrates the sensitivity problem for the class of systems which have an elastic element between the controller and the controlled element.

The disturbances acting on this system consist of the thrustor noise and a noise torque between the outer and middle body, both assumed white. The intensity matrix of these disturbances is obtained from Ref. 12 as

$$Q=\begin{bmatrix} 1.1\times 10^{-14} & 0 \\ 0 & 2.3\times 10^{-14} \end{bmatrix}[\mathrm{rad}^2\ \mathrm{sec}^{-3}]$$

The measurement noise intensity is $R=5.5\times 10^{-17}$ rad^2 sec.

Using these noises and state and control weighting matrices obtained from the same reference an optimal controller and filter are designed. The resulting augmented system is given in the form of Eq. (6) as

$$\begin{bmatrix} \dot{x} \\ \hline \dot{\tilde{x}} \end{bmatrix} = \left[\begin{array}{cccc|cccc} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.78k_\alpha & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ -0.44 & -1.08 & -(1.13+k_\alpha) & -1.5 & 0.44 & 1.08 & 1.13 & 1.5 \\ \hline 0 & 0 & 0 & 0 & -18.7 & 1.0 & 0 & 0 \\ 0 & 0 & 0.78k_\alpha-19.3 & 0 & -175 & 0 & 19.3 & 0 \\ 0 & 0 & 0 & 0 & -13.9 & 0 & 0 & 1.0 \\ 0 & 0 & 25-k_\alpha & 0 & 30.5 & 0 & -25.0 & 0 \end{array}\right] \begin{bmatrix} x \\ \\ \tilde{x} \end{bmatrix}$$

$$+\left[\begin{array}{ccc} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & -18.7 \\ 1 & 0 & -175 \\ 0 & 0 & -13.9 \\ 0 & 1 & 30.5 \end{array}\right] \begin{bmatrix} w_1 \\ w_2 \\ v \end{bmatrix} \tag{7}$$

where

$$x^T \triangleq [\gamma,\ \dot{\gamma},\ \alpha,\ \dot{\alpha}]$$

and

$$k_\alpha = K_\alpha \times (I_2+I_3)/I_2 I_3$$

(see Fig. 2). When $k_\alpha=25$, this equation takes the form of Eq. (4).

The eigenvalue change as a function of $\delta k_\alpha = k_\alpha - 25$ is conveniently represented in the form of a root locus. This root locus is obtained by inserting different values of k_α into Eq. (7) and computing the system eigenvalues. The root locus is shown in Fig. 3.

The range of variation of k_α for which the system is stable is

$$0.73 < (\delta k_\alpha / k_\alpha) < 1.26 \tag{8}$$

This range is not sufficient since k_α is ill defined and may vary considerably from its nominal value. Also, the range of permitted variation is smaller than the stability range since at least some damping is required for adequate response.

For comparison, the root locus vs δk is shown in Fig. 4 for the same system but with state feedback instead of state estimate feedback. No parameter sensitivity problem exists for this case, and it is therefore obvious that the sensitivity problem stems from the use of state estimates, instead of states, for feedback. In the root locus of Fig. 3, the sensitivity problem is apparent. More insight into its underlying reasons can, however, be gained by considering the open loop transfer function, given in Ref. 13.

$$G_0(s) = \frac{\gamma(s)}{\gamma_{\mathrm{ref}}(s)} = H(sI-F)^{-1}$$

$$\times\ G\ \frac{C\,\mathrm{adj}(sI-F+KH)K}{\det(sI-F+KH+GC)} \tag{9}$$

where adj is the adjoint matrix; det is the determinant. $G_0(j\omega)$ is shown in Fig. 5. Because of the undamped plant natural frequency, the frequency response has three zero crossings: points A, B, and C. The stability criteria for points B and C are found from the polar plot of the system. They are $\varphi_B > 180°$ and $\varphi_C < 180°$ where φ_B and φ_C are the negative phase angles at the points B and C, respectively. The sensitivity to variations in $\omega_n = \sqrt{k_\alpha}$ can be found from this criterion.

Figure 6 is an amplitude-frequency plot of the region of $\omega = \omega_n$ with part of the phase frequency plot overlaid. Since

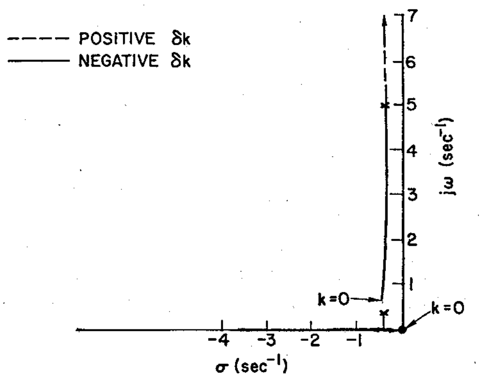

Fig. 4 Root locus vs δk for state feedback controller.

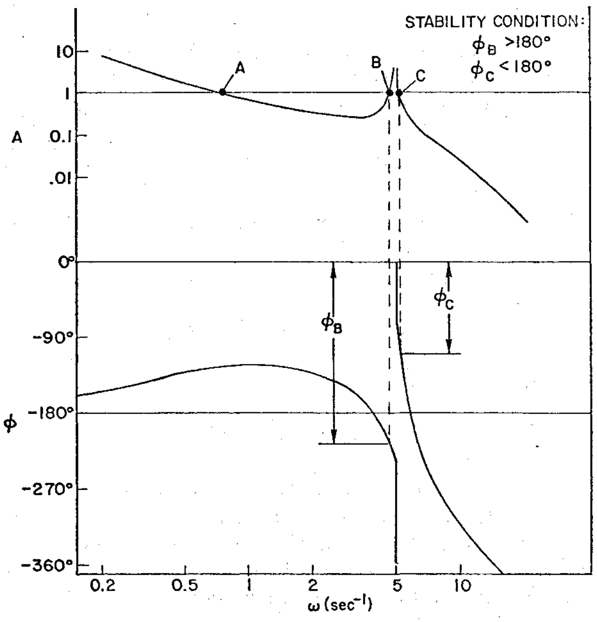

Fig. 5 Open loop frequency response of the controller.

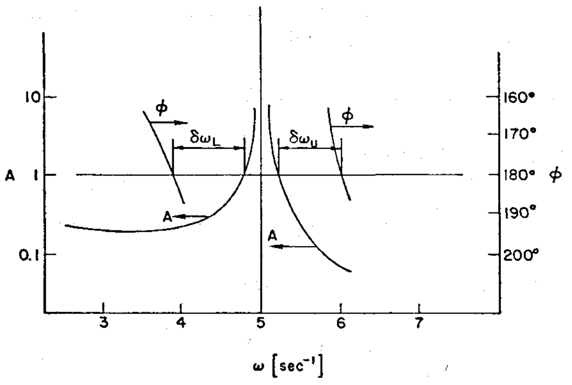

Fig. 6 Frequency response in the region of resonance.

the plant root at $\omega = \omega_n$ has no damping, its influence on the phase frequency plot consists of the addition of a phase lag of 180° at this frequency without modifying the plot at other frequencies. Also, the shape of the amplitude-frequency plot in the vicinity of this frequency is determined mostly by this root. Changes in the natural frequency of the plant will therefore cause the amplitude plot to move relative to the phase plot in Fig. 6 without changing the shape of the phase plot and with little change in the shape of the amplitude plot. It can be seen from the figure that changes in ω_n will cause instability as follows: 1) decrease by 1 rad/sec→instability by $\varphi_B < 180°$; 2) increase by 0.8 rad/sec→ instability by $\varphi_C > 180°$. The region of stability found by this approximate method is somewhat larger than that found by the root locus method.

To decrease the sensitivity, the frequency margins $\delta\omega_L$ and $\delta\omega_u$ have to be increased. This can be done by modifying the compensator roots so that the phase slopes in the region of the zero crossings at B and C are decreased. However, decreasing the slope in the vicinity of B will also decrease the phase margin at the first zero crossing, A. An acceptable compromise may be difficult to find and no systematic way exists to achieve it even for this simple system. The example given in Sec. II demonstrates the sensitivity problem of state estimate feedback controllers and motivates the requirement for a method to reduce this sensitivity.

III. Description of the Sensitivity Reduction Method

The method described in this section is applicable to the sensitivity reduction of linear systems that can be described by Eq. (1). The systems may be multivariable and subject to arbitrary inputs, both deterministic and stochastic. The basic equations of this method are similar to those given by Palsson and Whittaker[11] but the method of their solution is totally different.

A. Problem Statement

Consider the system of Eq. (1) where the matrices F, G, Γ, and H contain parameters the values of which are uncertain. If these parameters are considered as Gaussian random variables, a Gaussian random vector z may be formed of which they are the components. This vector is specified by

$$E(z) = z_0 \tag{10}$$

where the components of z_0 are the nominal values of the parameters, and by

$$E(zz^T) = V \tag{11}$$

a covariance matrix which is assumed known. Equation (1) may then be written as

$$\begin{aligned} \dot{x} &= F(z)x + G(z)u + \Gamma(z)w \\ y &= H(z)x + u \\ x(0) &= 0 \end{aligned} \tag{12}$$

A quadratic performance index (PI) for this sytem is

$$J = \lim \frac{1}{t_f} \int_0^{t_f} E(x^T A x + u^T B u)\,dt \tag{13}$$

where A and B are selected by the designer. In this expression the expected value is taken over the probability distributions of x and u that are derived from both the distributions of the random process w and of the random vector z. Since w and z are independent, the expected value of a function of x and u is

$$E[(x,u)] = \int_w E_z p(w)\,dw = \int_z E_w p(z)\,dz \tag{14}$$

where

$$E_z = \int_z g(w,z) p(z)\,dz$$

is the expected value over the distribution of z and

$$E_w = \int_w g(w,z) p(w)\,dw$$

is the expected value over the distribution of w. The control is defined as

$$u = -Cx \tag{15}$$

where the structure of the matrix C is given and the values of its coefficients may be left free or defined by functional relationships to other parameters of the system. For this definition of u to be valid, Eq. (1) in general is required to describe an augmented system that includes the plant and the compensations. The matrices F, H, and Γ then have to be defined accordingly. For state estimate feedback controllers, the augmented system is given by Eq. (4). For this case, u is defined as

$$u = [0, -C_E] \; [x, \hat{x}]^T = [-C_E, C_E] \; [x, \tilde{x}]^T$$

A free parameter vector q is defined by the designer. This vector consists of n_q parameters of the system matrices that can be varied by the designer.

The problem is now stated as follows: given the system of Eqs. (1) and (15) in which the system matrices are functions of a variable parameter vector defined by Eqs. (10) and (11), determine the value of the free parameter vector q so that the PI of Eq. (13) is minimized.

B. Method of Solution

Using Eq. (15) in Eq. (13), the PI becomes

$$J = \lim_{t_f \to \infty} \frac{1}{t_f} \int_0^{t_f} E[x^T(A + C^T BC)x] \, dt \tag{16a}$$

$$J = \mathrm{tr}\left[(A + C^T BC) \lim_{t_f \to \infty} \frac{1}{t_f} \int_0^{t_f} E(xx^T) \, dt \right] \tag{16b}$$

where tr is the trace operator. The state vector x may be written as

$$x(w, z) = x_n(w, z_n) + \delta x(w, \delta z) \tag{17}$$

where x_n is the state vector obtained when the parameters have their nominal values, and δx is the perturbation in the state vector due to a perturbation δz in the variable parameter vector.

Assuming small perturbations in z so that a first-order expansion is satisfactory, substituting Eq. (15) into Eq. (12) and defining $F_c = F - GC$, the governing equations for x_n and δ_x become

$$\dot{x}_n = F_c(z_n)x_n + \Gamma(z_n)w \tag{18a}$$

$$\dot{\delta}_x = F_c(z_n)\delta x + \delta F_c x_n + \delta\Gamma w \tag{18b}$$

$$x_n(0) = 0; \quad \delta x(0) = 0 \tag{18c}$$

where

$$\delta F_c = \sum_{i=1}^{n_z} \frac{\partial F_c}{\partial z_i} \delta z_i \quad \text{and} \quad \delta\Gamma = \sum_{i=1}^{n_z} \frac{\partial \Gamma}{\partial z_i} \delta z_i$$

The expected value of δx can be shown to be

$$E[\delta x(t)] = 0 \tag{19}$$

The expression $E(xx^T)$ in the PI of Eq. (16) can now be evaluated in terms of x_n and δx,

$$E(xx^T) = E[(x_n + \delta x) \; (x_n + \delta x)^T] \tag{20a}$$

$$E(xx^T) = E(x_n x_n^T) + E(x_n \delta x^T) + E(\delta x x_n^T) + E(\delta x \delta x^T) \tag{20b}$$

In this equation,

$$E[\delta x x_n^T] = E[x_n \delta x^T] = 0$$

and therefore

$$E(xx^T) = E(x_n x_n^T) + E(\delta x \delta x^T) = X_n + \delta X \tag{21}$$

where X_n is the covariance matrix of the nominal state. δX can thus be interpreted as the addition to the covariance due to the parameter uncertainties. Substituting Eq. (21) into Eq. (16) yields

$$J = \mathrm{tr}\left\{ (A + C^T BC) \lim_{t_f \to \infty} \frac{1}{t_f} \int_0^{t_f} [X_n(t) + \delta X(t)] \, dt \right\} \tag{22a}$$

$$J = \mathrm{tr}[(A + C^T BC) \; (\bar{X}_n + \overline{\delta X})] \tag{22b}$$

where $\bar{X}_n$ and $\overline{\delta X}$ are the time averages over time of X_n and δX, respectively. For a stable system, X_n and δX tend to constant values $X_n(\infty)$ and $\delta X(\infty)$. Since the averaging in Eq. (22) is performed over a large time interval, it can be assumed that $\bar{X}_n \to X_n(\infty)$ and $\overline{\delta X} \to \delta X(\infty)$. The PI therefore becomes

$$J \cong \mathrm{tr}\{ [A + C^T BC)X_n(\infty) + \delta X(\infty)] \} = J_0 + J_a \tag{23}$$

where

$$J_0 = \mathrm{tr}[(A + C^T BC)X_n(\infty)]$$

is the nominal PI and

$$J_A = \mathrm{tr}[(A + C^T BC)\delta X(\infty)]$$

is the additional PI due to the parameter variations.

The PI of Eq. (23) can be minimized computationally by a two-step sequence: 1) the matrices $X_n(\infty)$ and $\delta X(\infty)$ are found for a given value of the free parameter vector q. The PI that corresponds to this value is then determined. 2) The vector is modified in a direction that decreases J.

This sequence is repeated until the decrease in J in one cycle is less than a predetermined value. The two parts of this sequence are independent and computer programs for each one can be developed separately.

C. Governing Equations for X_n and δX

For brevity, the subscripts of x_n and X_n are now dropped. To find X, use $(d/dt) \; (xx^T)$, viz.,

$$(d/dt) \; (xx^T) = \dot{x}x^T + x\dot{x}^T \tag{24}$$

Substituting Eq. (18) into Eq. (24) yields

$$(d/dt) \; (xx^T) = (F_c x + \Gamma w)x^T + X(F_c x + \Gamma w)^T$$

$$= F_c xx^T + \Gamma wx^T + xx^T F_c^T + xw^T \Gamma^T \tag{25}$$

Taking the expected value of both sides of this equation and using the defintion

$$E(xx^T) = X$$

Equation (25) becomes

$$\dot{X} = F_c X + \Gamma E(wx^T) + XF_c^T + E(xw)\Gamma^T \tag{26}$$

But

$$E(xw^T) = E\{ [e^{F_c t}x(0) + \int_0^t e^{F_c(t-\tau)} \Gamma w(\tau) d\tau] w^T(t) \}$$

and

$$E[w(\tau)w^T(t)]=Q\delta(t-\tau)$$

$$E[x(0)w(t)] = 0 \text{ for } t\geq 0$$

Therefore

$$E(xw^T)=\tfrac{1}{2}\Gamma Q \tag{27}$$

Similarly,

$$E(wx^T)=\tfrac{1}{2}Q\Gamma^T \tag{28}$$

Using Eqs. (27) and (28) in Eq. (26), the covariance is

$$\dot{X}=F_cX+XF_c^T+\Gamma Q\Gamma^T \tag{29}$$

In the steady-state, $\dot{X}=0$ and the final equation for X_∞ is

$$F_cX_\infty+X_\infty F_c^T+\Gamma Q\Gamma^T=0 \tag{30}$$

The governing equation for δX can be formed in a similar way, using

$$(\mathrm{d}/\mathrm{d}t)\ (\delta x\delta x^T)=\delta x\delta\dot{x}^T+\delta\dot{x}\delta x^T \tag{31}$$

and Eq. (18b). The equation for δX_∞ is obtained as

$$\delta X_\infty F_c^T+F_c\delta X_\infty=-E_z[\delta\Gamma Q\delta\Gamma^T+Y_\infty^T\delta F_c^T+\delta F_cY_\infty] \tag{32}$$

where $Y=E_w(x\delta x^T)$. To get the governing equation for Y_∞, $(\mathrm{d}/\mathrm{d}t)\ (x\delta x^T)$ is used, viz.,

$$(\mathrm{d}/\mathrm{d}t)\ (x\delta x^T)=\dot{x}\delta x^T+x\delta\dot{x}^T$$

The steady-state value of Y as $t\rightarrow\infty$ is then given by

$$F_cY_\infty+Y_\infty F_c^T+\Gamma Q\delta\Gamma^T+X_\infty\delta F_c^T=0 \tag{33}$$

The governing equations for X_∞, Y_∞ and δX_∞ all have the same form: $AX+XA^T=B$, where A and B are given. They must be solved in the order in which they are written because the right-hand sides of the second and third equations contain expressions found in the solution of the previous equations. For their actual solution, these equations must be written in a slightly different form. Equation (33) is rewritten as

$$F_cY_\infty+Y_\infty F_c^T=-\sum_{i=1}^{n_z}\left[\Gamma Q\frac{\partial\Gamma}{\partial z_i}-x_\infty\frac{\partial F_c}{\partial z_i}\right]\delta z_i \tag{34}$$

Y is therefore obtained in the form of a sum

$$Y_\infty=\sum_{i=1}^{n_z}Y_i\delta z$$

where Y_i is the solution of

$$F_cY_i+Y_iF_c=-\Gamma Q\frac{\partial\Gamma}{\partial z_i}-X_\infty\frac{\partial F_c}{\partial z_i} \tag{35}$$

This equation is solved n_z times with different right-hand sides.

The right-hand side of Eq. (32) can be written as

$$E_z[\delta\Gamma Q\delta\Gamma^T+\delta F_cY_\infty+Y_\infty^T\delta F_c^T]$$

$$=E_z\left[\sum_{j=1}^{n_z}\sum_{i=1}^{n_z}\frac{\partial\Gamma}{\partial z_i}Q(\frac{\partial\Gamma}{\partial z_j})^T+\frac{\partial F_c}{\partial z_i}Y_j+(\frac{\partial F_c}{\partial z_i}Y_j)^T\right]\delta z_i\delta z_j$$

$$=\sum_{j=1}^{n_z}\sum_{i=1}^{n_z}\left[\frac{\partial\Gamma}{\partial z_i}Q(\frac{\partial\Gamma}{\partial z_j})^T+\frac{\partial F_c}{\partial z_i}Y_j+(\frac{\partial F_c}{\partial z_i}Y_j)^T\right]\mathrm{V}_{ij} \tag{36}$$

where V is the parameter covariance matrix [see Eq. (11)].

The final equations for X_∞ and δX_∞ are therefore

$$F_cX_\infty+X_\infty F_c^T=-\Gamma Q\Gamma^T \tag{37a}$$

$$F_cY_i+Y_iF_c^T=-\Gamma Q\frac{\partial\Gamma}{\partial z_i}-X_\infty\frac{\partial F_c}{\partial z_i}\qquad i=1\ldots n_z \tag{37b}$$

$$F_c\delta X_\infty+\delta X_\infty F_c^T=\sum_{j=1}^{n_z}\sum_{j=1}^{n_z}\left[\frac{\partial\Gamma}{\partial z_i}Q(\frac{\partial\Gamma}{\partial z_j})^T+\frac{\partial F_c}{\partial z_i}Y_j+(\frac{\partial F_c}{\partial z_i}Y_j)^T\right]V_{ij} \tag{37c}$$

In the problem statement, the parameter covariance matrix V is assumed known [see Eq. (11)]. Its elements are a measure of the uncertainties in the parameters. In reality, however, these uncertainties are ill defined and the matrix V is considered mainly as a design tool. The magnitude of its elements is selected according to the importance of the sensitivity reduction for the respective parameters.

Equations (37) remain valid if the system is forced by inputs other than white noise. If the disturbance is colored, it can be modeled by means of a shaping filter forced by white noise. An augmented system can then be formed consisting of the states of the system and of the shaping filter. This augmented system is excited by white noise and Eqs. (37) are therefore valid for it.

For deterministic systems which are required to recover from nonzero initial conditions the white noise vector w is replaced by an impulsive input vector w_i at $t=0$: $w_i=w_0\delta(t)$. The initial conditions are then represented by the equivalent initial impulses. Equations (37) can be used, unchanged, if the following terms are redefined:

$$Q=w_0w_0^T \tag{38a}$$

$$X_\infty=\int_0^\infty xx^T\mathrm{d}t \tag{38b}$$

$$Y=\int_0^\infty x\delta x^T\mathrm{d}t \tag{38c}$$

$$\delta X_\infty=\int_0^\infty\delta x\delta x^T\mathrm{d}t \tag{38d}$$

With these redefined terms, the deterministic PI is given by Eq. (23). If the system is to be optimized for some deterministic input other than an initial impulse, this can be handled by state augmentations.

D. Description of the Computer Program

The computer program PAROPT for the minimization of the PI of Eq. (13) is described in detail in Ref. 13. Only its main features are described in this section. It consists of two main parts: 1) a search subprogram and 2) a subprogram for the solution of Eqs. (37) and determination of the PI from Eq. (23)

1) Search Subprogram

This subprogram is part of the program library of the Computer Science Department at Stanford University. Developed by Gill, Murray, and Pitfield,[4] it iteratively seeks the minimum value of a scalar function $F(q)$ where q is a vector of dimension n_q by modifying the components of q. The value of $F(q)$ for a given q is on input to the subprogram. For the current problem, it is the performance index of Eq. (23). The operation of this program consists of sequences of numerical gradient evaluations and linear searches along the direction of the conjugate gradient. These sequences are repeated until termination criteria are satisfied.

2) Subprogram for the Evaluation of J

It is obvious from the description of the search subprograms that the values of J must be computed a large number of times. In one iteration, n_q evaluations are required for the gradient determination and 3-4 for the linear search. In an average program, 8-10 iterations may be expected and therefore it may be required to compute J 100-200 times. For the program to be of any practical usefulness, a very efficient method for this computation must therefore be developed. The principal part of this evaluation is the solution of Eqs. (37).

The same equation with different right-hand sides is to be solved $2+n_z$ times. It is to be noted, however, the Eq. (37a) and (37c) have symmetric right-hand sides and therefore symmetric solutions (Lyapunov equations), whereas the right-hand side of Eq. (37b) is not symmetric. Several methods are available for the solution of this equation. These methods are compared by Hagander[15] and Pace and Barnett.[16]

A direct method was elected in this case. It consists of transforming an equation of the form

$$AZ+ZA^T=B \tag{39}$$

into the form $\alpha z=\beta$ and solving this linear equation by Gaussian elimination. $z(\ell\times 1)$ and β are vectors of the elements of Z and B, respectively. α is obtained from A by algorithms which depend on the form of A (symmetric, antisymmetric, or general).[17,18]

Gaussian elimination consists of two steps: 1) forward elimination [$\ell^3/3$ operations] and 2) back substitution [$\ell^2/2$ operations]. Forward elimination is required only if α is changed. If only β is changed, just the much less costly backsubstitution is required.

In our case, for one evaluation of the PI, only one forward elimination of order $\frac{1}{2}(n+1)n$ is required (for X), and one of order $\frac{1}{2}(n-1)n$ (for the antisymmetric part of Y_l). Moreover, since the points at which the cost is computed for the gradient determination correspond to only slightly perturbed q vectors, the changes in X and δX that stem from these perturbations can be obtained by expanding Eq. (37) about the nominal point and neglecting second-order terms.

The evaluation of the PI and its gradient, therefore, can be done with only one forward elimination of order $\frac{1}{2}(n-1)n$. This method was implemented in the computer program PAROPT, which has been applied to the sensitivity reduction of several systems. Results of its application are described in Sec. IV. The largest system to which it was applied is of 12th order, with 2 variable and 20 free parameters. One iteration for this system (gradient evaluation + linear search) required about 40 sec on an IBM 360/67 computer. This method therefore seems practical for even fairly large systems.

IV. Applications

In this section, an application of the sensitivity reduction program is described. The system to which it is applied is the low-frequency approximation of the Stanford Relativity Satellite as described in Sec. II. The criteria used for the comparison of the nominal and desensitized designs (for the same system) are: 1) sensitivity criterion: the range of variation of the variable parameter for which the system is still stable (high range ≡ low sensitivity) and 2) nominal performance criteria: output and control rms values, and the square root of the nominal PI, which is a weighted rms value.

Some points have to be kept in mind when using these criteria: 1) The stability range of the variable parameters should not be construed as defining the actual permitted range of variation of these parameters. In general, the performance will become unacceptable for variations that are considerably less than those that cause instability. 2) The rms values of the outputs and the controls depend on the assumed covariance matrices of the process and measurement noises. These covariance matrices are in general not well known and in some cases, they are artificially determined in order to get acceptably damped roots of the estimator.[12] Small differences (less than a factor of two) in the rms values of different designs cannot, therefore, be considered as significant.

More precise criteria are difficult to define in general, although for specific cases, they may exist. In some cases, the time response envelope to a specific input may be restricted, or limits may be posed on the phase and gain margins. It is important to verify whenever such criteria are used that they reflect actual system requirements and do not pose artificial restraints on the design.

The parameters that are controlled by the designer (some or all of which can be used as free variables) are: C, the feedback gains; K, the estimator gains; and F, the representations of the variable parameters in the estimator. The last item may require some clarification. If there are no variable parameters, the estimator parameters will obviously be selected to be the same as the plant parameters. If, however, some plant parameters are variable, the sensitivity may be decreased if their representations in the estimator differ from their nominal values. It is therefore desirable to include these representations among the free parameters. The actual selection of these program parameters will be discussed for each example separately. The nominal optimal point (found without considering sensitivity) was selected as the initial point.

A. Low Frequency Approximation of the Stanford Relativity Satellite

This system is described in Sec. II. Only the estimator gains and the variable parameter representations in the estimator were used as free parameters since it was found in preliminary runs that the feedback gains do not vary appreciably. The application of the sensitivity reduction program in this case is therefore a redesign of the estimator. The spring constant is the only variable parameter. The covariance matrix is therefore a scalar, $V=\sigma_k{}^2$ where σ_k is the assumed rms value of the spring constant.

As explained in Sec. III, this rms value does not represent an actual expected uncertainty of the variable parameter but is used as a design tool for the relative weighting of the nominal and additional performance indices. The essential features of this example are exhibited in the following two designs.

Design 1: $\sigma_k/k_\alpha=0.28$

Design 2: $\sigma_k/k_\alpha=0.9$

The estimator in these designs is a steady-state Kalman filter using the nominal values of R and Q given in Sec. II, and the sensitivity reduction program seeks a balance between the nominal optimal estimator parameters and those required for minimum sensitivity.

B. Results

The results of this application are presented below. Designs 1 and 2 are compared to the nominal design. The eigenvalues of the different designs are shown in Table 1; the sensitivity properties are given in Table 2; the nominal performance

Table 1 Eigenvalues of nominal and desensitized systems: reduced Stanford relativity satellite

	Nominal	Design 1	Design 2
Controller	$-0.36 \pm 5.01j$	$-0.38 \pm 4.98j$	$-0.38 \pm 5.05j$
	$-0.41 \pm 0.41j$	$-0.41 \pm 0.41j$	$-0.41 \pm 0.41j$
Estimator	$-1.02 \pm 5.11j$	$-1.41 \pm 3.17j$	$-1.16 \pm 9.19j$
	$-8.33 \pm 8.35j$	$-0.8 \pm 13.2j$	-27.6
			-1.14

Table 2 Stability range-reduced SRS

		Nominal	Design 1	Design 2
Range of	$+\delta k_\alpha / k_\alpha$	+0.26	+1	+1.27
Stability	$-\delta k_\alpha / k_\alpha$	−0.52	−0.52	−0.8

Table 3 Nominal performance criteria (scaled to initial values)

	Nominal	Design 1	Design 2
Output rms Noise ($\sigma_\theta / \sigma_{\theta i}$)	1	1.14	1.72
Control rms Noise (σ_u / σ_{ui})	1	1.12	1.12
Weighted rms Noise $(J_0/J_{0i})^{1/2}$	1	1.09	1.12

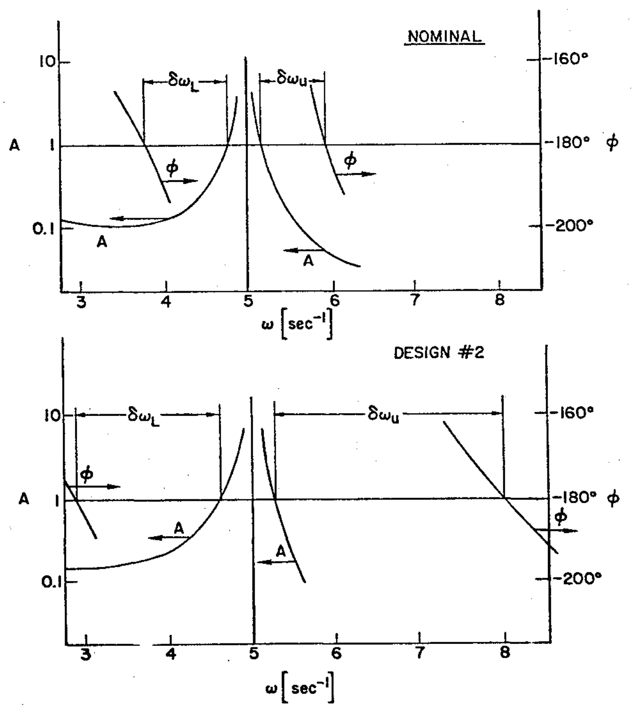

Fig. 7 Frequency margin comparison.

criteria are compared in Table 3; and the frequency margins of the nominal design and design 2 are compared in Fig. 7.

C. Evaluation of the Results

1) As expected, increasing the assumed uncertainty of the variable parameter decreases the sensitivity. Since the initial point is an optimal one, the decrease in sensitivity is linked to an increase in output and control noises.

2) The lower sensitivity of design 2 can also be observed from its increased frequency margin (Fig. 7).

3) From additional designs that were performed, it was found that if noise considerations are ignored the resulting designs are similar to design 2. That means that the uncertainty in k_α that was assumed in design 2 causes the estimator gains in this design to be determined primarily by sensitivity considerations. Such considerations can therefore serve as an important design tool for estimators.

4) Even for this simple system, it would have been difficult to determine the minimum sensitivity eigenvalue locations without some general method such as was used.

5) The estimator root at $-1.02 \pm 5.11j$, that was close to a controller root (see Fig. 3 and Table 1), was significantly shifted by the desensitizing procedure, thus reducing the tendency of the locus to go into the right half-plane.

The method was also applied to the full relativity satellite as described in Sec. II.[13] The stability range for this case was increased by a factor of about 2.5 with an output noise increase by a factor of 2.4 and small control noise increases. The computation time for this case (12th-order system with 20 variable parameters) was 4 min on a IBM 360/67 computer.

V. Conclusions

1) The method described in this paper can provide a considerable reduction in the system sensitivity to parameter variations. If the initial system is optimal the output and control rms values will increase due to the sensitivity reduction. However, the nature of the optimality of the nominal system has to be considered carefully since it depends on the assumed values of the process and measurement noise intensities. In the examples that were examined, the sensitivities were reduced by a factor of 2-3, while the output rms increased approximately by a factor of 2.

2) The example given in Sec. IV used state estimate feedback controllers. In this case, the free parameters are the feedback gains and the estimator parameters. The method, however, is by no means limited to such controllers. It is equally applicable to systems with classical compensation networks or other designs that can be represented in state variable form.

3) The computation time for a 12th-order system with 20 variable parameters was 4 min on a 360/67 computer. This computation time is almost insensitive to the number of free parameters. It is therefore recommended to define as such all the parameters that are at the designer's disposition, at least for preliminary runs. If some parameters do not vary in those runs, they may be fixed for subsequent runs.

4) In applying this method, several designs are, in general, recommended with different parameter covariance matrices. The comparison between these designs is conveniently made by means of the stability range and the output and control rms values. If the system has specific performance requirements such as time response envelope or gain and phase margin, these can also be used as comparison criteria. The most satisfactory design, in general, is a matter of subjective designer preference.

References

[1]Kreindler, E., "On Minimization of Trajectory Sensitivity," *International Journal of Control*, Vol. 8, 1968, pp. 89-96.

[2]Porter, B. and Crossley, R., *Modal Control*, Barnes & Noble, New York, 1972.

[3]Bode, H. W., *Network Analysis and Feedback Amplifier Design*, Van Nostrand, New York, 1945.

[4]Cruz, J. B., *System Sensitivity Analysis*, Dowden, Hutchinson & Ross, Inc., Stroudsbourg, Pa., 1973.

[5]Luh, J. Y. S. and Cross, E. R., "Optimal Trajectory Design for Minimum Sensitivity," *Proceedings of the Allerton Conference*, 1967, pp. 152-159.

[6]Cassidy, J. and Lee, I., "On the Optimal Feedback Control of a Large Launch Vehicle to Reduce Trajectory Sensitivity," *Proceedings of the Joint Automatic Control Conference*, 1967, pp. 587-595.

[7]D'Angelo, H., Moe, M. L. and Hendricks, T. C., "Trajectory Sensitivity of an Optimal Control System," *Proceedings of the Allerton Conference*, 1966, pp. 489-498.

[8]Hendricks, T. and D'Angelo, H., "An Optimal Fixed Structure Design with Minimal Sensitivity for a Large Elastic Booster," *Proceedings of the Allerton Conference*, 1967, pp. 142-150.

[9]Rillings, J. H. and Roy, R. J., "Analog Sensitivity Design of Saturn V Launch Vehicle," *IEEE Transactions on Automatic Control*, Vol. AC15-4, Aug. 1970, p. 437.

[10]Rynasky, E. G., Whitbeck, R. F., and Dolbin, B. W., "Sensitivity Considerations of a Flexible Launch Vehicle," Rept. CAL BE-2311-F-1, Cornell Aeronautical Lab., Cornell University, Ithaca, N.Y.

[11]Palsson, T. and Whittaker, H. P., "Parameter Uncertainties in Control System Design," *Proceedings of the Joint Automatic Control Conference*, 1972, pp. 248-254.

[12]Bull, J. S., "Precise Attitude Control of the Stanford Relativity Satellite," Rept. SUDAAR 452, March 1973, Stanford University, Center for System Research, Stanford, Calif.

[13]Hadass, Z., "Design of State Feedback Controllers Including Sensitivity Reduction with Applications to Precision Pointing," SUDAAR No. 482, Aug. 1974, Stanford University, Center for System Research, Stanford, Calif.

[14]Gill, P. E., Murray, W., and Pitfield, R. A., "The Implementation of Two Revised Quasi Newton Algorithms for Unconstrained Optimization," Rept. NAC11, April 1972, National Physics Lab., England.

[15]Hagander, P., "Numerical Solution of $A^TS+SA+Q=0$," *Journal of Information Science*, Vol. 4, 1969, pp. 35-50.

[16]Pace, I. S. and Barnett, S., "Comparison of Numerical Methods for Solving Lyapunov Matrix Equations," *International Journal of Control*, Vol. 15, May 1972.

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

[18]Bingoulac, S. P., "An Alternate Approach to Expanding $PA+A'P=Q$," *IEEE Transactions on Automatic control*, AC-15, Feb. 1970, pp. 135-137.

[19]Forsythe, E. G., Malcolm, M. A., and Moler, C. B., "Computer Methods for Mathematical Computation," Dec. 1972, Stanford University, Computer Science Dept., Stanford, Calif.