JOURNAL OF GUIDANCE, CONTROL, AND DYNAMICS
Vol. 19, No. 1, January–February 1996

Feedback Design for Unstable Plants with Saturating Nonlinearities: Single-Input, Single-Output

R. A. Hess* and S. A. Snell†
University of California, Davis, Davis, California 95616

Control systems are often required to operate in the presence of saturation nonlinearities, such as actuator displacement and rate limiting. Whereas the occurrence of saturation may possibly be minimized by the choice of closed-loop bandwidth, the nonlinearities themselves are essentially unavoidable, and system performance and even stability can be compromised when the nonlinearities are encountered. To address this problem, a design technique for improving the performance of single-input, single-output feedback systems operating in the presence of saturation nonlinearities is presented. The technique allows application to unstable plants and is an extension of that previously discussed in the literature for stable plants. A simple design example is provided that demonstrates the applicability of the methodology to a case of actuator rate limiting.

Introduction and Background

FEEDBACK design techniques that explicitly consider the possibility of control saturation are receiving increasing attention in the literature. Optimal control strategies for regulators employing inequality constraints on control and state variables are well known.[1] Techniques that attempt to minimize the adverse effects of saturation are also well established, e.g., antiwindup controllers.[2] Feedback systems designed to avoid saturation have been discussed.[3] Nonlinear feedback controllers designed with the goal of improved performance with state and control saturation have also been described.[4] Finally, linear feedback schemes to improve the performance of systems with saturation nonlinearities have been formulated.[5]

The research to be described here falls into the last category. In particular, the treatment of actuator saturation as a nuisance nonlinearity is pursued. This implies that the system in question experiences saturation relatively infrequently and that a nominal control system design was obtained ignoring any nonlinear actuation effects. Such is the case for many aircraft flight control systems in which actuator saturation, either displacement or rate, is a relatively rare occurrence. The rarity of such occurrences is not accidental since the performance requirements of control surface actuators are typically an important part of flight control system specifications. Nonetheless, flight control actuator saturations do occur, typically, with high-performance aircraft.[6–8] When conditionally stable systems are involved (e.g., statically unstable aircraft), the results can be catastrophic.[9,10]

The design technique to be pursued is an extension of that developed by Horowitz[5,11,12] and in Refs. 13 and 14 briefly outlined in the next section. Attention is focused on single-input, single-output (SISO) systems. The design equations are presented next, and the Horowitz technique is extended to unstable plants. The internal stability and input–output stability of the resulting design approach are then discussed, and a simple example is presented which demonstrates the design approach.

Design Technique

Software Limits

The implementation of the design technique to be discussed herein employs what are commonly referred to as software limits. That is, a saturation element is deliberately introduced as part of the compensator in the forward loop that ensures that no commands are sent to the actuator that would result in saturation. This approach effectively moves the nonlinearity from the actuator to the compensator and, in doing so, eliminates any sensing requirements in the design implementation.

Horowitz Design Philosophy

Consider the simple SISO systems shown in Figs. 1 and 2. In these figures, $P(s)$ represents the plant and $A(s)$ the actuator. Both systems employ the software limits alluded to in the preceding paragraph. If $n = 0$, a displacement saturation is created, if $n = 1$, a rate saturation is created, etc. Figure 2 is a modification of Fig. 1 in which an additional feedback loop around the saturation (software) nonlinearity has been created and a different forward loop compensator has been employed. By appropriate choice of n and M, the systems of both Figs. 1 and 2 can eliminate saturation in the actuator. However, Horowitz[11] has shown that one can create independent loop transmissions with respect to the plant and the saturating nonlinearity with the structure of Fig. 2. In doing so, one can significantly improve the performance of the system of Fig. 2 compared to that of Fig. 1 if and when saturations do occur.

Independent Loop Transmissions

Define $-L_n(s)$ as the loop transmission around the nonlinearity N in Fig. 2. That is, open the loop at the plant input $Y(s)$ and determine the transfer function $X(s)/Y(s)$. In addition, define $L(s)$

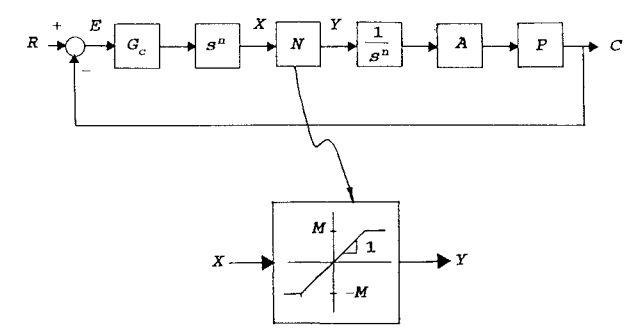

Fig. 1 SISO system with software limiting.

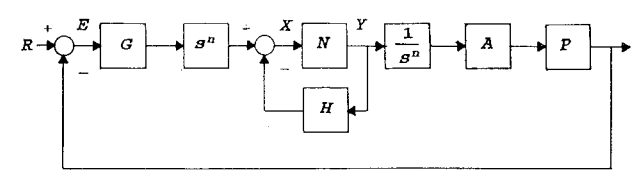

Fig. 2 SISO system with software limiting and independent loop transmissions about P and N.

Received Nov. 30, 1994; presented as Paper 95-0336 at the AIAA 33rd Aerospace Sciences Meeting, Reno, NV, Jan. 9–12, 1995; revision received June 12, 1995; accepted for publication June 27, 1995.

*Professor, Department of Mechanical and Aeronautical Engineering. Associate Fellow AIAA.

†Assistant Professor, Department of Mechanical and Aeronautical Engineering. Member AIAA.

as the system loop transmission when the nonlinearity is replaced by unity. That is, open the loop at $E(s)$ and calculate the transfer function $C(s)/E(s)$ with $N = 1.0$. Then the following relations can be verified through simple block diagram algebra:

$$C/R = FL/(1 + L) \tag{1a}$$

$$L = GPA/(1 + H) \tag{1b}$$

$$L_n = (H + PAG) \tag{1c}$$

The design actually begins with desired $L(s)$ and $L_n(s)$. Solving Eqs. (1b) and (1c) for $G(s)$ and $H(s)$ yields

$$G = \frac{L(1 + L_n)}{PA(1 + L)} \tag{2a}$$

$$H = (L_n - L)/(1 + L) \tag{2b}$$

Equations (1) allow the designer to independently select L and L_n via $G(s)$ and $H(s)$ from Eqs. (2). The advantage of this independent selection will be explained next.

It is assumed that the $G_c(s)$ in Fig. 1 has been obtained through application of a linear control system design technique that ignores the nonlinearity N, i.e., assumes $N = 1.0$ (nonunity values can be handled by suitable normalization). The result of this linear design is a desired loop transmission $L(s)$ that would, of course, depend on the compensation $G_c(s)$. Now once $L_n(s)$ has been chosen, the system of Fig. 2 could be implemented using Eqs. (2). Concentrating upon Fig. 2, assume that

$$L_n(s) \approx K/s \tag{3}$$

Further, when $N = 1.0$, define

$$y_L(t) = y(t) \tag{4}$$

Let t^i represent the instant of the ith switch from linear to nonlinear behavior [$y(t)$ going into saturation] and let t_i represent the ith switch from nonlinear to linear behavior [$y(t)$ coming out of saturation]. Also assume the quiescent system is always in the linear mode so that $t^1 < t_1$.

Now the following system behavior can be verified.[11] In time intervals in which the signal $y(t)$ is not in saturation, it will approach $y_L(t)$ with dynamics $[1 + L_n(s)]^{-1}$. This means, for $t_i < t \leq t^{i+1}$ and given Eq. (3),

$$y(t) - y_L(t) \propto e^{-K(t - t_i)} \tag{5}$$

Proof of Eq. (5) is not trivial and can be found in Ref. 13.

The rationale behind the choice of the right-hand side of Eq. (3) is now evident. If the designer selects $L_n(s)$ so that the loop around N has transient behavior that is fast as compared to $Y_L(t)$ [i.e., choose K in Eq. (3) large enough], the performance of the system of Fig. 2 can be improved relative to that of the system of Fig. 1, provided the system is stable. In cases where no saturation occurs, the performance of the system of Fig. 2 is identical to that of Fig. 1. Note that this procedure does not eliminate saturation. It alters the dynamic characteristics of the system so that, in coming out of saturation, performance approaches that which would have been in evidence had saturation not occurred.

For reasons to be discussed, the design methodology just described is limited to stable plants. It is obviously desirable to remove this limitation, particularly in aerospace flight control applications. Horowitz and Liao[13] and Liao and Horowitz[14] offer a feedback scheme that can guarantee closed-loop stability for unstable plants with saturation. The design, however, requires the construction of a nonlinear compensation element and does not accommodate disturbance effects. With the exception of the saturation element whose characteristics are known, the system of Fig. 2 is linear and will accommodate disturbance effects.

Extending the Design Technique

Internal Stability

One requirement for a stable plant stems directly from Eq. (2a). In this equation, plant poles will appear in $L(s)$ and, of course, $P(s)$. When the right-hand side of Eq. (2a) is expressed as a ratio of polynomials, the unstable poles will appear as right-half-plane zeros in the numerator of the compensator $G(s)$. These zeros will then cancel the identical unstable poles appearing in the inner closed loop of Fig. 2, and internal instability results. Note, however, that right-half-plane zeros of $P(s)$ will not appear as unstable poles of the compensator $G(s)$, hence, nonminimum-phase plants present no problems.

The approach to be taken here for incorporating unstable plants is to simply modify $L_n(s)$ as follows. Consider expressing the plant transfer function as

$$P(s) = \frac{N(s)}{D_s(s)D_u(s)} \tag{6}$$

where $D_u(s)$ is the polynomial containing the unstable poles of $P(s)$. With nu denoting the number of unstable poles, one can write

$$D_u(s) = \prod_{i=1}^{\text{nu}} (s - p_i) \tag{7}$$

where the p_i may be complex. Now, if one forms

$$L_n(s) = \frac{K}{S} \cdot \left[\prod_{i=1}^{\text{nu}} (s + p_i) \Big/ \prod_{i=1}^{\text{nu}} (s - p_i) \right] \tag{8}$$

Eqs. (2a) and (2b) indicate that the internal stability problem just described disappears. The price one pays for this approach is that $L_n(s)$ will now deviate from the ideal K/s form. This deviation, however, will occur at frequencies well below the crossover frequency of $L_n(s)$, and so the transient response of $[1 + L_n(s)]^{-1}$ will be very similar to that obtained with the original L_n. This is because $L_n(s)$ will exhibit a crossover frequency well above that of $L(s)$ that will, by necessity, exhibit a crossover frequency well above the unstable poles of $P(s)$. The latter is a requirement of any conditionally stable design. Thus, the deviations in the ideal $L_n(s)$ of Eq. (3) will be relegated to very low frequencies relative to its crossover frequency. This will ensure both the stability of $[1 + L_n(s)]^{-1}$ created with the $L_n(s)$ of Eq. (8) and the similarity in transient responses of the $[1 + L_n(s)]^{-1}$ created with either the $L_n(s)$ of Eq. (3) or (8).

Final questions regarding the existence of internal stability and stable compensators revolve about determining whether $G(s)$, $H(s)$, and $1/[1 + H(s)]$ are each stable. These questions can be answered in turn, by referring to Eqs. (2) and the following arguments:

Stability of $G(s)$. By referring to Eq. (2a) and recalling that $L(s)$ in the numerator can be written $G_c(s)A(s)P(s)$, one can show that $G(s)$ will be stable under the following conditions.

1) $G_c(s)$ has no right-half plane (RHP) poles. This will be assumed to be the case, as it is assumed that the original linear design is acceptable.

2) $1 + L(s)$ has no RHP zeros. This will be the case since the original linear design is assumed to be stable.

Note that the RHP poles of $1 + L_n(s)$ are canceled by the RHP poles of $1 + L(s)$. This cancellation is exact and can be ensured by the designer by simply omitting these poles from the formulation for $G(s)$.

Stability of $H(s)$. By referring to Eq. (2b), one can show that $H(s)$ will be stable if $1 + L(s)$ has no RHP zeros. This is guaranteed (see condition 2). Note that RHP poles of $L_n(s) - L(s)$ are canceled by RHP poles of $1 + L(s)$. Again, this is an exact cancellation ensured by the designer.

Stability of $1/[1 + H(s)]$. Any RHP poles of $1/[1 + H(s)]$ would also appear as RHP zeros of $G(s)$, exclusive of those RHP zeros in common with $G_c(s)$. This is true since the only RHP poles in $L(s)$ are those from the plant $P(s)$ [again, assuming stable $G_c(s)$]. This leads us to ask if such RHP zeros can exist in $G(s)$. Again, by referring to Eq. (2a), one can show that the only RHP zeros in $G(s)$ will be those in common with $G_c(s)$.

Input–Output Stability

Unfortunately, with unstable plants, the input–output stability of either of the systems of Fig. 1 or 2 is dependent on the duration of saturation encounters. This can be seen by appealing to an argument taken from Ref. 13. Assume in Fig. 1 that $M = 1$, and $y(t)$ goes into saturation. That is,

$$|x(t)| \geq 1 \qquad \text{for} \quad 0 \leq t \leq t^d \tag{9}$$

Now assume that for $t > t^d$, an $x(t)$ of opposite sign and equal magnitude is applied. Express the plant as

$$P(s) = P_s(s) \cdot P_u(s) \tag{10}$$

where P_u contains the unstable poles of $P(s)$. For simplicity, assume $P(s)$ contains only a single unstable pole as $s = p_1$. Then,

$$P(s) = P_s(s) \cdot 1/(s - p_1) \tag{11}$$

and

$$C(s) = \frac{P_s(s)}{(s - p_1)(s^n)} \cdot \left[\frac{1}{s} - \frac{2e^{-t_d s}}{s}\right] \tag{12}$$

The residue associated with the pole of $C(s)$ at $s = p_1$ is

$$\text{residue} = \frac{P_s(p_1)}{p_1^{n+1}}[1 - 2e^{-p_1 t_d}] \tag{13}$$

Thus, for control recovery, the sign of the residue must be negative, or

$$t_d < \frac{\ell_n(2)}{p_1} = \frac{0.693}{p_1} \tag{14}$$

which can be recognized as the time to double amplitude for the mode associated with $s = p_1$. For systems with more than one unstable pole, including possible complex conjugate pairs, Eq. (14) can be generalized to

$$t_d < \frac{\ell_n(2)}{\max_j |\text{Re}(p_j)|} \tag{15}$$

where $\max_j |\text{Re}(p_j)|$ is the largest real part of all unstable poles in $P(s)$. It is important to emphasize that neither of the systems of Fig. 1 nor Fig. 2 will be stable if Eq. (15) is not satisfied.

A parameter that can predict the likely effectiveness of the proposed design approach is

$$\bar{\tau} = \frac{\left[\max_i \left(t_i - t^i\right)\right]}{0.693/\max_j |\text{Re}[p_j]|} \tag{16}$$

where $\bar{\tau}$ is the ratio of the longest saturation interval to the t^d of Eq. (15). If $\bar{\tau}$ exceeds unity, even the maximum control available will not yield stable behavior after a saturation. The term $[\max_i (t_i - t^i)]$ can be estimated from the linear response of the system ($M = \infty$ in Fig. 2) over the input set likely to be encountered in practice, as shown in Fig. 3. Here $y(t)$ with $M = \infty$ for a representative input and/or disturbance is shown along with the saturation limit of the actuator. The actual duration of any saturation with the system of Fig. 2 may be longer than that predicted from the linear response. However, if the system of Fig. 2 is to exhibit performance comparable to that which is in evidence with $M = \infty$ (which is the goal of the proposed design approach), then the saturation durations for Fig. 1 or 2 with $M = \infty$ will, of necessity, be comparable to those of Fig. 2 with finite M.

Summarizing, the greater the value of $\bar{\tau}$ is, the nearer the system comes to violating a necessary condition for closed-loop stability. Thus, one can be confident that the proposed technique will not yield performance comparable to that of a design without saturation for values of $\bar{\tau}$ near unity. As will be seen in the example, however, the technique may still yield performance far superior to that of the system of Fig. 1, with saturation.

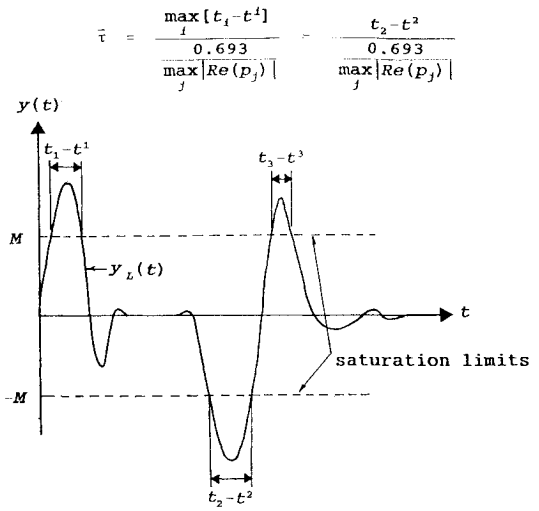

Fig. 3 Estimating $\bar{\tau}$ from the linear system response.

Higher-Order Saturation

In implementing the design of Fig. 2, it is possible that higher order actuator saturations can be induced by the dynamics $[1 + L_n]^{-1}$. For example, consider $n = 1$ in Fig. 2 with $y(t)$ coming out of a saturation and approaching $y_L(t)$ with dynamics $[1+L_n]^{-1}$. This $y(t)$ could cause the actuator to undergo an acceleration saturation. Thus, the selection of the crossover frequency of L_n must involve a tradeoff. That is, it should be large enough to ensure that the dynamics of $[1 + L_n]^{-1}$ are fast as compared to $y_L(t)$ and yet not so large as to induce saturations of higher order in the actuator.

Implementation Considerations

As described in the preceding, the systems of both Figs. 1 and 2 have been created with software rate limits. That is, the purpose of the limiter N is to prevent rate commands above a certain magnitude from being input to the actuator. It is assumed that the actuator output itself is limited and at a value sufficiently higher than that of the software limit and so its limit is never exceeded. Thus, no actuator limitations are included in the analysis, and the requirement for sensing actuator output is removed. Note that nonsymmetric saturation limits are easily accommodated by making the nonlinear elements in Figs. 1 and 2 nonsymmetric. If, as often occurs, the actuator limits are dependent on the actuator operating environment, the limits of the nonlinear elements can be easily scheduled. For example, if one is dealing with the rate limit of a flight control actuator that is dependent upon dynamic pressure, M in Fig. 2 could be scheduled as a function of this variable.

Finally, the nth-order differentiator appearing before the nonlinear element in Fig. 2 will be subsumed into the compensator in implementation of the control law. This can affect the nature of $G(s)$ as will be seen in the example.

Steady-State Errors

The system behavior described by Eq. (5), wherein the output of the nonlinear element exponentially approaches that of the system without saturation, has an obvious intuitive appeal. It is output error $e(t)$, however, that defines closed-loop performance. It is of interest then to investigate $e(t)$ and, in particular, the steady-state behavior of $e(t)$. The steady-state error of the system of Fig. 2 with transient input $r(t)$ can be assessed by examining the equivalent system of Fig. 4. Assume in Fig. 2 that closed-loop stability is in evidence and $y(t)$ has undergone a single, transient saturation from $t^1 \leq t \leq t_1$. The nonlinearity can be removed and replaced by a remnant input $\mathcal{L}^{-1}\{D_n(s)[1 + L_n(s)]\}$, thus producing a quasilinear system that exhibits system time histories identical to those of the system of Fig. 2.[15] The signal $d_n(t)$ is defined as

$$\begin{aligned} d_n(t) &= y(t) - y_L(t) \qquad \text{for} \quad t^1 \leq t \leq t_1 \\ &= 0 \qquad \text{elsewhere} \end{aligned} \tag{17}$$

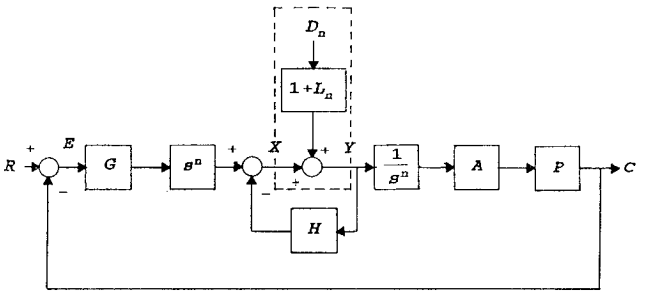

Fig. 4 Linear system equivalent to Fig. 2 for single saturation encounters.

Now employ the standard definition of type as the exponent of any free s appearing in a Laplace transform or transfer function. For notational convenience denote the type of $D_n(s)$ as the integer q^{D_n}, which can be negative (free s in the numerator). Thus, since $d_n(t) = 0$ as $t \to \infty$, the final value theorem guarantees that $D_n(s)$ will be at most type 0, or $q^{D_n} \leq 0$.

Block diagram algebra, the final value theorem, and Fig. 4 can be used to show

$$E(s) = \frac{D_n(s)[1 + L_n(s)]}{G(s)s^n[1 + L(s)]} + \frac{R(s)}{1 + L(s)}$$

$$\begin{aligned} e_{ss}(t) &= \lim_{t \to \infty} e(t) = \lim_{s \to 0} sE(s) \\ &= \frac{D_n(s)[1 + L_n(s)]}{G(s)s^{n-1}[1 + L(s)]} + \frac{R(s)}{s[1 + L(s)]} \end{aligned} \tag{18}$$

Invoking the final value theorem requires that the function in question has no poles in the right-half plane and only real poles on the imaginary axis.[16] This is the case in the first of Eqs. (18) since unstable poles in $1 + L_n(s)$ and $1 + L(s)$ cancel exactly. In addition, if $G(s)$ is nonminimum phase and introduces RHP poles into the first term in $E(s)$, identical zeros will appear in $D_n(s)$ and cancel them.

Now, with the help of Eqs. (2), the type of the first and second terms on the right-hand side of Eq. (18) can be given by

$$\begin{aligned} q^{sE}|_{1\text{st term}} &= [n - 1 + q^{D_n} - q^L + q^p] \\ q^{sE}|_{2\text{nd term}} &= [q^R - q^L - 1] \end{aligned} \tag{19}$$

Thus, if $e_{ss}(t) \to 0$ as $t \to \infty$ then each bracketed term in Eq. (19) must be algebraically less than zero, or

$$\begin{aligned} q^L &> n - 1 + q^{D_n} + q^P \\ q^L &> q^R - 1 \end{aligned} \tag{20}$$

Recalling that $q^{D_n} \leq 0$, a conservative requirement for $e_{ss}(t) \to 0$ is that

$$q^L > n - 1 + q^p \tag{21a}$$

$$q^L > q^R - 1 \tag{21b}$$

For simplicity, the preceding argument assumed a single transient saturation. Of course, multiple saturations can occur, even with a transient input $r(t)$. Equations (21) are still applicable, provided that the saturations are transient, i.e., that $y(t)$ will, at some time, come out of saturation.

Example

The treatment of the preceding sections appears to suggest a very qualitative, ad hoc design procedure. The fundamental approach, however, is essentially that of loop shaping, certainly a familiar and time-honored tool of the control systems engineer. The following simple example will serve to outline the design procedure and indicate the type of performance improvements that are possible with the approach that has been described.

Table 1 Design example transfer functions

$$G_c(s) = \frac{4045.5(0.1)(0.5)}{(0)(20)(50)}\ ^{a,b}$$

$$A(s) = \frac{75^2}{(75)^2}$$

$$P(s) = \frac{0.5}{(0)(-0.5)}$$

$$L_n(s) = \frac{60(0.5)}{(0)(-0.5)}\ ^{c}$$

$$G(s) = \frac{3537.3(0.1)(0.5)(0.509)(75)}{(0.0937)[0.7099, 1.169](14.4)(60.9)}\ ^{b}$$

$$H(s) = \frac{58.23(-0.00349)(0.5)(18.17)}{(0.0937)[0.7099, 1.169](14.4)}$$

[a] $K(z)/[\zeta, \omega_n] = K(s + z)/(s^2 + 2\zeta\omega_n s + \omega_n^2)$.
[b] Compensator shown was augmented by factor $(s + 0.3)/s$ before simulation.
[c] Transfer function prior to increasing type of $G(s)$.

Design

Consider the SISO systems of Figs. 1 and 2 with $n = 1$, i.e., rate limits are being imposed. The plant dynamics, given in Table 1, exhibit an aperiodic divergent mode with a time to double amplitude of approximately 1.4 s. Such modes are often found in modern aircraft with unconventional configurations and concomitant negative static stability margins. Table 1 also lists the assumed actuator dynamics.

Table 1 shows the compensation element $G_c(s)$ for the system of Fig. 1. $G_c(s)$ was the result of a classical loop shaping design. The crossover frequency for the loop transmission $L(s)$ was chosen as 2.0 rad/s and q^L was chosen as 2, satisfying Eqs. (21) for transient $r(t)$. Table 1 also shows the $L_n(s)$ for the system of Fig. 2 as provided by Eq. (7). The K value was chosen by placing the crossover frequency of $L_n(s)$ just beyond the magnitude peak of the $Y(s)/R(s)$ transfer function, thus ensuring that the dynamics of $y(t)$ would be fast as compared to those of $y_L(t)$. Cancellation of approximately colocated poles and zeros in the right-hand sides of Eqs. (2a) and (2b) allowed $G(s)$ and $H(s)$ to be simplified to the forms shown in Table 1.

Using Eq. (2a) resulted in $q^G = 0$. Since the differentiator element $s^n (n = 1)$ in Fig. 2 is to be subsumed into the compensator, however, q^G was increased by one through the inclusion of the factor $(s + 0.3)/s$. This change also affects $L(s)$ and $L_n(s)$ but only at low frequencies, and thus will have a negligible effect on the desired closed-loop characteristics. Finally, to allow a fair comparison between the systems of Figs. 1 and 2 in the simulations to be discussed, the same $(s + 0.3)/s$ factor was included in $G_c(s)$.

Input–Output Stability

As discussed in the preceding, input–output stability of the system of Fig. 2 cannot be guaranteed when the system undergoes saturation. The utility of the technique for any application, however, can be made by estimating $\bar{\tau}$.

Simulation

A series of simulations of the systems of Figs. 1 and 2 were conducted. Two transient test inputs were used:

$$r_1(t) = \mathcal{L}^{-1}[I_1(s)F_I(s)] \tag{22}$$

where

$$i_1(t) = u(t) - 2u(t - 4) + u(t - 8) \tag{23}$$

and

$$r_2(t) = \mathcal{L}^{-1}[I_2(s)F_I(s)] \tag{24}$$

where

$$i_2(t) = \sin(\pi t)[u(t) - u(t - 20)] \tag{25}$$

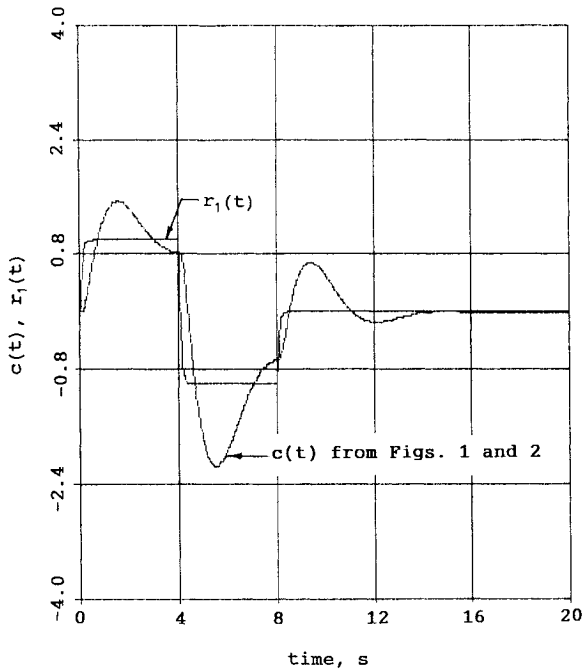

Fig. 5 Doublet responses of the systems of Figs. 1 and 2, without saturation.

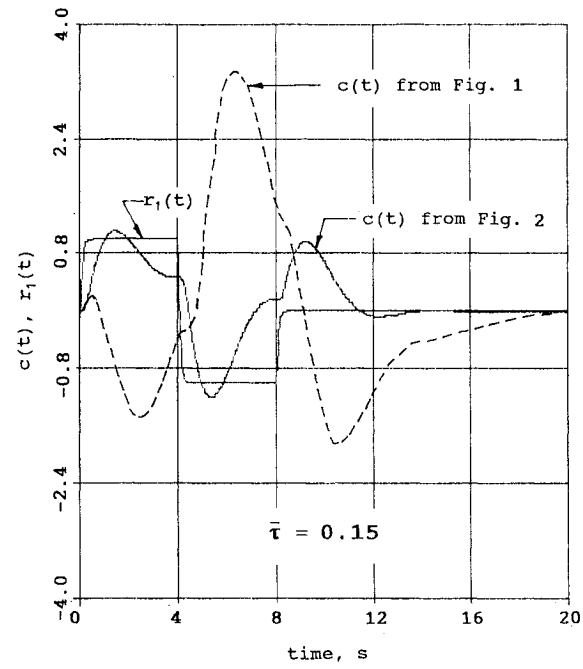

Fig. 6 Doublet responses of the systems of Figs. 1 and 2, with saturation.

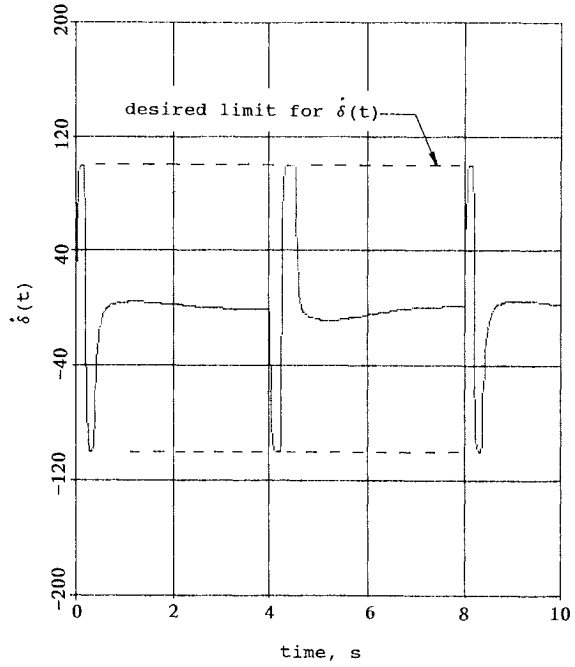

Fig. 7 Actuator output rate limiting in the system of Fig. 2 for doublet input.

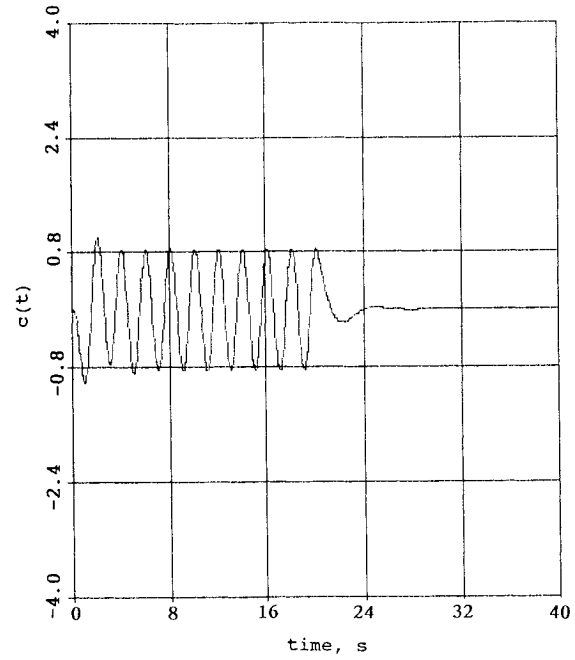

Fig. 8 Transient sinusoid responses of systems of Figs. 1 and 2, without saturation.

In what follows, $r_1(t)$ is referred to as the doublet, and $r_2(t)$ as the transient sinusoid. Finally, the input filter $F_I(s)$ was chosen as

$$F_I(s) = \frac{20^2}{(s+20)^2} \tag{26}$$

In each of the examples to be discussed, rate limits were chosen that produced significant rate saturation with the particular input in question. For the doublet input, M was chosen to allow $\bar{\tau} \ll 1.0$, whereas for the transient sinusoid, M was chosen so that $\bar{\tau}$ was still less than unity but several times larger than the value for the doublet case. Both of these calculations were based on the linear response $y(t)$ as demonstrated in Fig. 3.

Performance with Doublet Input

The rate limit was chosen as 100/s. Figure 5 shows the identical output time histories of both the systems of Figs. 1 and 2 without the saturation element. They are identical since, without saturation, both systems exhibit the same loop transmission about the plant. Figure 6 shows the responses of the two systems when the rate limiter was in place. The $\bar{\tau}$ for this input is also shown in the figure. Note the superiority of the performance of the system of Fig. 2 as compared to that of the system of Fig. 1. Figure 7 shows the actual actuator limiting in $\dot{\delta}(t)$ that is occurring in the system of Fig. 2. Note that the limiter in the nonlinear compensator is successful in preventing the actuator output rate from exceeding the 100/s target value.

Performance with Transient Sinusoid Input

The rate limit was chosen as 20/s. Figure 8 shows the identical output time histories of both the systems of Figs. 1 and 2 without the saturation element. Figure 9 shows the responses of the two systems when the rate limiter was in place. The $\bar{\tau}$ value is also given for this

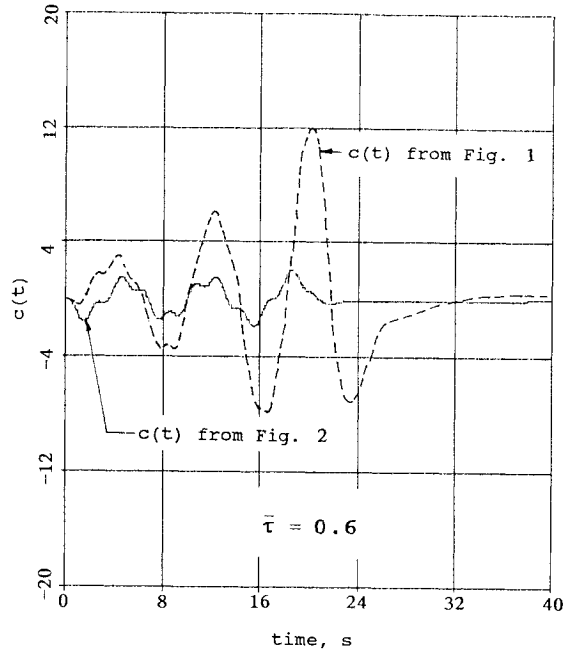

Fig. 9 Transient sinusoid responses of systems of Figs. 1 and 2, with saturation.

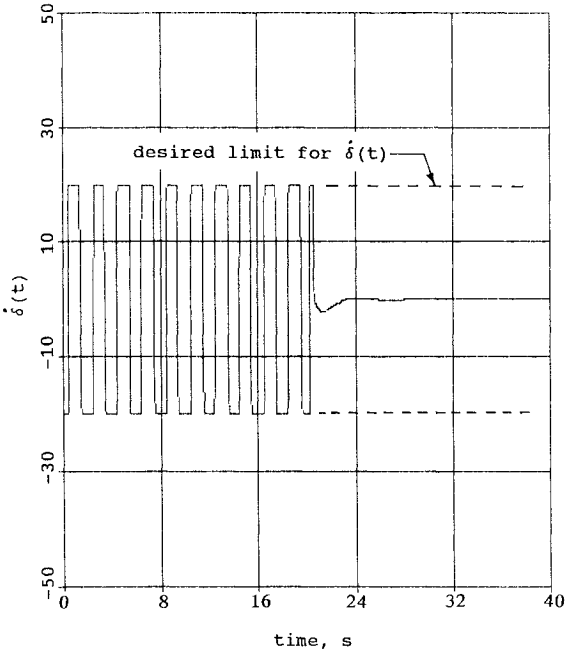

Fig. 10 Actuator output rate limiting in the system of Fig. 2 for transient sinusoid input.

input. Although the system of Fig. 2 did not follow the command input with fidelity, the response was far superior to that for the system of Fig. 1. Note the different ordinate scales in Figs. 8 and 9. Figure 10 shows the actual actuator limiting in $\dot{\delta}(t)$ for the system of Fig. 2 demonstrating both the significant rate limiting that occurred with this system and the fact that the limiter in the compensator was successful in preventing the actuator output rate from exceeding the 20/s target value. The caveat restricting the definition of input–output stability to transient inputs comes to bear in this example. Both time responses of Fig. 9 show an oscillatory divergence (much slower in the case of the system of Fig. 2) for $0 < t < 20$ s, i.e., before the input $r_2(t)$ dies out.

Actuator Acceleration Magnitudes

The system of Fig. 2 achieves its superior performance by forcing $y(t)$, the output of the rate limiter, to approach $y_L(t)$ (the output that would have been in evidence had no limiting occurred) with dynamics $[1 + L_n(s)]^{-1}$. These dynamics are designed to be fast as compared to those of $y_L(t)$. Achieving these rapid dynamics could produce actuator acceleration saturation. If it occurs, this acceleration limiting can severely degrade the performance and stability of this or any feedback system. In the simulation examples considered herein, it is useful to compare the relative magnitudes of the maximum actuator output accelerations for the systems of Figs. 1 and 2. For the doublet input, the maximum acceleration for the system of Fig. 2 was approximately 11% greater than that for the system of Fig. 1. For the transient sinusoid input, the maximum acceleration for the system of Fig. 2 was approximately 122% greater than that for the system of Fig. 1. Recall, however, that estimates of $\bar{\tau}$ indicated that the system of Fig. 2 would not produce performance comparable to a linear system ($M = \infty$) with the transient sinusoid, even without acceleration saturation. Thus, it is unlikely that it would be a candidate for consideration if the 20/s rate limits were in evidence in the actuator. At this juncture, one must either reduce the performance requirements of the system or increase the performance requirements of the actuator.

If predicted actuator acceleration values obtained from simulations such as these are deemed too large but if estimates of $\bar{\tau} \ll 1.0$, then one can reduce the K value in $L_n(s)$.

Discussion

Given the simplicity of the example just discussed, it would be inadvisable to draw too strong an analogy with flight control systems and associated problems. It is, however, useful to point out that the simulation involving the transient sinusoid input bears some resemblance to a pilot-induced oscillation (PIO) encounter, both in terms of the frequency of oscillation and the nature of the saturation. The 0.5-Hz frequency is in the approximate range of PIO frequencies reported in the literature for longitudinal PIOs of high-performance aircraft,[17] and the actuator was effectively in nearly constant rate saturation throughout, a phenomenon also frequently associated with severe PIO encounters. The ability of the proposed technique to provide improved performance in this simulation is encouraging.

Directions for Future Research

The design technique proposed herein should be capable of being extended to multi-input, multi-output (MIMO) systems. Indeed, the Horowitz design approach has been extended to MIMO systems with stable plants.[18] The extension to MIMO systems with unstable plants is not trivial, however, as the issue of internal stability can become quite challenging. Research is currently underway in this area.

Conclusions

Based on the research reported herein, the following conclusions can be drawn.

1) A technique for improving the performance of feedback systems subject to actuator saturation has been extended to unstable plants. Implemented as a system to prevent actuator saturation, the design has no sensor requirements.

2) The technique begins with a linear feedback design that ignores any saturation limits. The design produces a system that exhibits performance identical to the original linear design in the absence of saturation and performance superior to the original design in the presence of saturation. This improvement is achieved through the independent selection of loop transmissions around the plant and nonlinearity.

3) Internal stability is guaranteed in the approach but input–output stability in the presence of saturation is not. A method for determining the utility of the design, however, can be offered that is based on the duration of saturation encounters estimated from system time histories with no saturation.

4) A quasilinear representation of the nonlinear system demonstrated the dependence of steady-state error on saturation order (i.e., displacement, rate, etc.) and provided type requirements for the loop transmission around the plant.

5) The selection of the crossover frequency for the loop transmission $L_n(s)$ requires careful consideration to maximize performance

in the presence of saturations explicitly considered in the design and to minimize the occurrence of higher order saturations.

6) A simple example involving a plant with an aperiodically divergent mode demonstrated the utility of the proposed design technique.

7) Future research will focus upon extending these results to MIMO designs.

References

[1]Bryson, A. E., Jr., and Ho, Y. C., *Applied Optimal Control*, Hemisphere, New York, 1989, Chap. 3.

[2]Kothare, M. V., Campo, P. J., Morari, M., and Nett, C. N., "A Unified Framework for the Study of Anti-Windup Designs," *Automatica*, Vol. 30, No. 12, 1994, pp. 1869–1883.

[3]Kapasouris, P., Athans, M., and Stein, G., "Design of Feedback Control Systems for Unstable Plants with Saturating Actuators," *IFAC Symposium on Nonlinear Control Systems Design* (Capri, Italy), edited by A. Isidori, 1989, pp. 302–307.

[4]Mayne, D. Q., and Schroeder, W. R., "Nonlinear Control of Constrained Linear Systems," *International Journal of Control*, Vol. 60, No. 5, 1994, pp. 1035–1043.

[5]Horowitz, I., "A Synthesis Theory for a Class of Saturating Systems," *International Journal of Control*, Vol. 38, No. 1, 1983, pp. 169–187.

[6]Powers, B. G., "Space Shuttle Longitudinal Landing Flying Qualities," *Journal of Guidance, Control, and Dynamics*, Vol. 9, No. 5, 1986, pp. 566–572.

[7]Dornheim, M. A., "Report Pinpoints Factors Leading to YF-22 Crash," *Aviation Week and Space Technology*, Vol. 137, No. 19, 1992, pp. 53, 54.

[8]Anon., "Why the Gripen Crashed," *Aerospace America*, Vol. 32, No. 2, 1994, p. 11.

[9]McRuer, D., and Graham, D., "Retrospective Essay on Nonlinearities in Aircraft Flight Control," *Journal of Guidance, Control, and Dynamics*, Vol. 14, No. 6, 1991, pp. 1089–1099.

[10]Hanson, G. D., and Stengel, R. F., "Effects of Displacement and Rate Saturation on the Control of Statically Unstable Aircraft," *Journal of Guidance, Control, and Dynamics*, Vol. 7, No. 2, 1984, pp. 197–205.

[11]Horowitz, I. M., *Quantitative Feedback Design Theory (QFT)*, QFT Pub., Boulder, CO, 1993, Chap. 11.

[12]Horowitz, I. M., "Feedback Systems with Rate and Amplitude Limiting," *International Journal of Control*, Vol. 40, No. 6, 1984, pp. 1215–1229.

[13]Horowitz, I. M., and Liao, Y. K., "Quantitative Non-Linear Compensation Design for Saturating Unstable Uncertain Plants," *International Journal of Control*, Vol. 44, No. 4, 1986, pp. 1137–1146.

[14]Liao, Y. K., and Horowitz, I. M., "Unstable Uncertain Plants with Rate and Amplitude Saturations," *International Journal of Control*, Vol. 44, No. 4, 1986, pp. 1147–1159.

[15]Graham, D., and McRuer, D., *Analysis of Nonlinear Control Systems*, Dover, New York, 1971, Chap. 5.

[16]Nise, N. S., *Control Systems Engineering*, Benjamin/Cummings, Redwood City, CA, 1992, Chap. 7.

[17]Bjorkman, E. A., Silverthorn, J. T., and Calico, R. A., "Flight Test Evaluation of Techniques to Predict Longitudinal Pilot Induced Oscillations," AIAA Paper 86-2253, Aug. 1986.

[18]Hess, R. A., "Feedback System Design for Stable Plants with Input Saturation," *Journal of Guidance, Control, and Dynamics*, Vol. 18, No. 5, 1995, pp. 1029–1035.