A Unified Framework for the Study of Anti-Windup Designs

Mayuresh V. Kothare   Peter J. Campo*   Manfred Morari†
Chemical Engineering, 210-41
California Institute of Technology
Pasadena, CA 91125

Carl N. Nett
School of Aerospace Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332

CDS Technical Report No. CIT-CDS 93-011

June 16, 1993

Abstract

We present a unified framework for the study of linear time-invariant (LTI) systems subject to control input nonlinearities. The framework is based on the following two-step design paradigm: "*Design the linear controller ignoring control input nonlinearities and then add anti-windup bumpless transfer (AWBT) compensation to minimize the adverse effects of any control input nonlinearities on closed loop performance*". The resulting AWBT compensation is applicable to multivariable controllers of arbitrary structure and order. All known LTI anti-windup and/or bumpless transfer compensation schemes are shown to be special cases of this framework. It is shown how this framework can handle standard issues such as the analysis of stability and performance with or without uncertainties in the plant model. The actual analysis of stability and performance, and robustness issues are problems in their own right and hence not detailed here. The main result is the unification of existing schemes for AWBT compensation under a general framework.

1 Introduction

All real world control systems must deal with constraints. For example, the control system must avoid unsafe operating regimes. In process control, these constraints typically appear in the form of pressure and temperature limits. In addition, physical limitations impose constraints–pumps and compressors have finite throughput capacity, surge tanks can only hold a certain volume, motors have a limited range of speed. Of special interest and common occurrence are systems with control inputs constraints in an otherwise linear system.

*Current affiliation: GE Corporate Research and Development, P. O. Box 8, Schenectady, NY 12301
†To whom all correspondence should be addressed: phone (818)356-4186, fax (818)568-8743, e-mail: MM@IMC.CALTECH.EDU

All physical systems are subject to actuator saturation. For example, a valve controlling the flow rate of the coolant to a reactor can only operate between fully open and fully closed. We will refer to such a constraint as an *input limitation.* In addition, commonly encountered control schemes must satisfy multiple objectives and hence need to operate in different control modes. Each mode has a linear controller designed to satisfy the performance objective corresponding to that mode. If the operating conditions demand a change of mode, an override or selection scheme chooses the appropriate mode and executes a mode switch. The switch between operating modes is achieved by a selection of the plant input from among the outputs of a number of parallel controllers, each corresponding to a particular mode. We will refer to such a mode switch as a *plant input substitution* since the output of one controller is replaced by that of another.

As a result of substitutions and limitations, the actual plant input will be different from the output of the controller. When this happens, the controller output does not drive the plant and as a result, the states of the controller are wrongly updated. This effect is called *controller windup.* Since the linear controller is designed ignoring actuator nonlinearities, the adverse effect of windup caused by the presence of such nonlinearities is in the form of significant performance deterioration (as compared to the expected linear performance), large overshoots in the output and sometimes even instability [6]. Performance degradation is especially pronounced when the controller is stable with very slow dynamics and gets even worse when the controller is unstable. In addition to windup, when mode switches occur, the difference between the outputs of different controllers results in a bump discontinuity in the plant input. This in turn causes undesirable bumps in the controlled variables. What is required is a smooth transition or *bumpless transfer* between the different operating modes. We will refer to the problem of control system analysis and controller synthesis for the general class of linear time invariant (LTI) systems subject to plant input limitations and substitutions as the *anti-windup bumpless transfer* (AWBT) problem.

Windup problems were originally encountered when using PI/PID controllers for controlling linear systems with control input nonlinearities. One of the earliest attempts to overcome windup in PID controllers was the work by Fertik and Ross [11]. Their strategy has been variously referred to as *anti-reset windup*, *back calculation and tracking* and *integrator resetting.* Experimental evaluation of several digital algorithms for anti-reset windup has been reported by Khanderia and Luyben [16]. Extension of anti-reset windup to a general class of controllers has been reported and is commonly referred to as high gain conventional anti-windup (CAW).

It was recognized later that integrator windup is only a special case of a more general problem. As pointed out by Doyle *et al.* [10], any controller with relatively slow or unstable modes will experience windup problems if there are actuator constraints. Windup is then interpreted as an inconsistency between the controller output and the states of the controller when, for example, the control signal saturates. The "conditioning technique" as an anti-windup and bumpless transfer scheme was originally formulated by Hanus *et al.* [14, 15] as an extension of the back calculation strategy of Fertik and Ross [11] to a general class of controllers. Åström *et al.* [3, 2] proposed that an observer be introduced into the system to estimate the states of the controller and hence restore consistency between the saturated control signal and the controller states. Walgama and Sternby [20] have very clearly exposed this inherent observer property in several anti-windup schemes. Campo and Morari [6] have derived the Hanus conditioned controller as a special case of the observer-based approach. A modified Internal Model Control (IMC) implementation has recently been proposed [21] to improve performance in the face of actuator saturation.

We can summarize the existing approaches to solving the problem of control of LTI systems subject to control input nonlinearities as follows: *Design the linear controller ignoring control input nonlinearities and then add anti-windup bumpless transfer (AWBT) compensation to minimize the* *adverse effects of any control input nonlinearities on closed loop performance.* While many of these schemes have been successful (at least in specific SISO situations), they are by and large intuition based and have little theoretical foundation. Specifically:

- no attempt has been made to formalize these techniques and advance a general AWBT analysis and synthesis theory.
- with the exception of a few [13, 12, 10], no rigorous stability analyses have been reported for anti-windup schemes in a general setting.
- the issue of robustness has been largely ignored (notable exceptions are [6, 7]).
- extension to MIMO cases has not been attempted in its entirety. As pointed out by Doyle *et al.* [10], for MIMO controllers, the saturation may cause a change in the direction of the plant input resulting in disastrous consequences.
- a major void in the existing AWBT literature is a clear exposition of the objectives (and associated engineering trade-offs) which lead to a graceful performance degradation in any reasonably general setting.

The focus of this paper is on setting up a general framework for studying anti-windup and bumpless transfer designs. Although optimal control strategies for saturating systems can be determined using nonlinear optimal control theory, the implementation of such control laws is fairly complicated. Since actuator constraints are relatively simple nonlinearities, we will confine ourselves to the two step design paradigm discussed above. Moreover, we will seek **linear** AWBT compensation for actuator nonlinearities. A comparison of the simple two-step design paradigm discussed above as opposed to a computationally more involved approach which allows the designer to tackle problems of greater complexity has been presented in [17].

This paper is organized as follows: In Section 2, the general AWBT problem is formulated. In order to obtain results with general applicability, we deviate from the existing AWBT literature and work in an abstract framework rather than discussing AWBT methods developed for a particular example. In Section 3, all known LTI anti-windup and/or bumpless transfer schemes are shown to be special cases of the general framework. This unification is the main result of this paper. In Section 4, we briefly discuss how the general framework can handle standard issues such as the analysis of stability and performance, with or without plant uncertainty. Finally, in Section 5, we present concluding remarks.

Notation:: The notation is fairly standard. $\mathcal{R}$ is the field of real numbers. The Lebesgue spaces $\mathcal{L}_2(-\infty, 0]$ and $\mathcal{L}_2[0, \infty)$ consist respectively of square-integrable functions on $(-\infty, 0]$ and $[0, \infty)$. $\mathcal{H}_2$ is the Hardy space of square integrable functions on the imaginary axis with analytic continuation into the right half plane. $\mathcal{H}_\infty$ is the Hardy space of functions bounded on the imaginary axis with analytic continuation into the right half-plane.

For a matrix A, A^* denotes its complex conjugate transpose, A^T denotes its transpose, A^{-1} denotes its inverse (if it exists).

A transfer matrix in terms of state-space data is denoted

$$G(s) = C(sI - A)^{-1}B + D = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

The $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms of a stable transfer matrix $G(s)$ are

$$\|G\|_2 \quad = \quad \left(\frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{trace}[G(j\omega)^* G(j\omega)]\, d\omega\right)^{\frac{1}{2}}$$

$$\|G\|_\infty = \sup_\omega \sigma_{max}[G(j\omega)]$$

where $\sigma_{max}[G(j\omega)]$ is the maximum singular value of $[G(j\omega)]$. The $\mathcal{L}_2$ norm of a vector valued function $u(t) \in \mathcal{L}_2$ on $(-\infty, \infty)$ is

$$\|u\|_2 = \left(\int_{-\infty}^{\infty} u^*(t)u(t)\,dt\right)^{\frac{1}{2}}$$

The same symbol will be used to denote a time-domain signal and its Laplace transform. The distinction should be clear from the context.

2 A General AWBT Framework

In the following sub-sections, the AWBT problem is formulated, the AWBT design criteria are discussed, certain admissibility criteria for AWBT are introduced and a parameterization of all admissible AWBT compensated controllers is presented. Wherever necessary, the assumptions underlying the development are clearly stated.

2.1 Problem Formulation

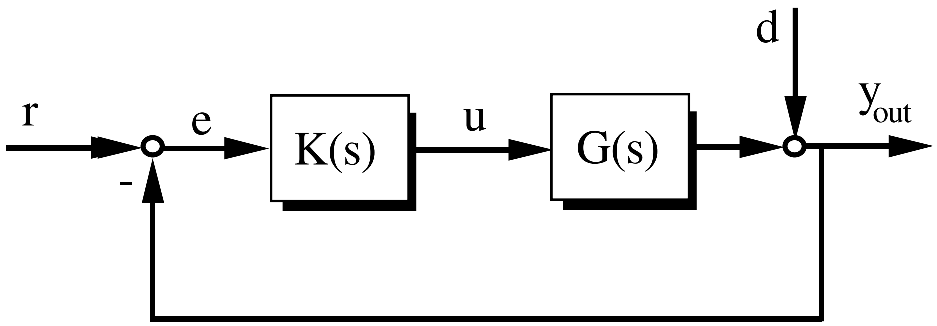

Figure 1: Ideal linear design–error feedback case

The problem considered in this paper can be understood with reference to Fig. 1. In Fig. 1, we have an idealized linear control problem where the linear plant model $G(s)$ is provided. An LTI controller $K(s)$ is designed to meet given performance specifications. These will typically be of the form, "Keep the output tracking error e small despite changes in command r and disturbance d". Due to limitations and/or substitutions, a nonlinearity N is introduced into the interconnection as shown in Fig. 2. As a result, the actual plant input $\hat{u}$ will in general not be equal to the controller output u. This mismatch is the cause for controller windup, controller state initialization errors and a significant transient which must decay after the system returns to the linear regime. This is also the cause for degradation of performance and sometimes instability. The AWBT problem involves the design of $\hat{K}(s)$ shown in Fig. 3. The measured or estimated value of $\hat{u}$ provides information regarding the effect of the generic nonlinearity N and is fed back to the AWBT compensated controller $\hat{K}(s)$. The design criteria to be satisfied by $\hat{K}(s)$ are as follows:

1. The nonlinear closed loop system, Fig. 3, must be stable.

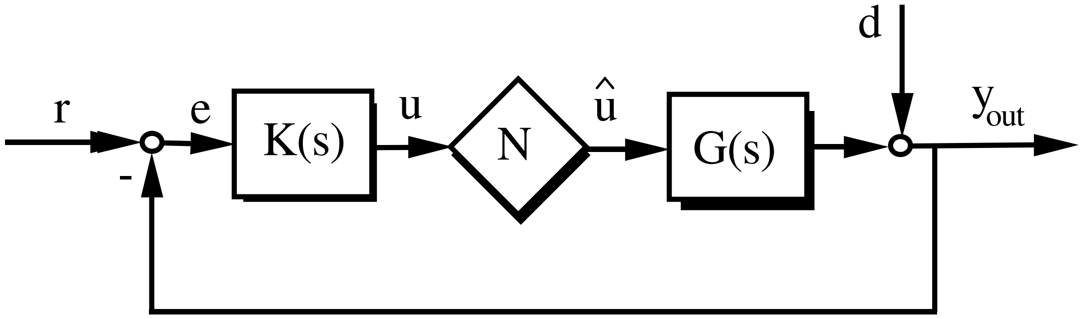

Figure 2: Ideal linear design with nonlinearity N–error feedback case

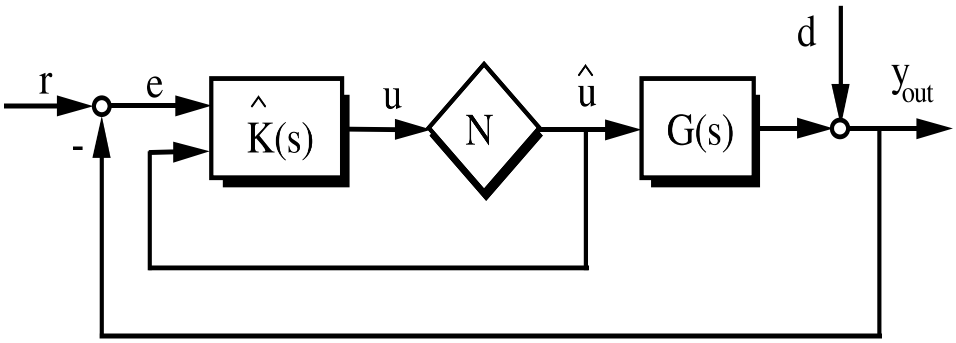

Figure 3: The AWBT problem–error feedback case

2. When there are no limitations or substitutions, ($N \equiv I$), the closed loop performance of the system in Fig. 3 should meet the specifications for linear design in Fig. 1. We call this the *linear performance recovery* requirement.

3. The closed loop performance of the system in Fig. 3 should *degrade gracefully* from the linear performance of Fig. 1 when limitations and/or substitutions occur ($N \not\equiv I$).

In order not to restrict attention to the error feedback case alone, we will consider the linear fractional transformation (LFT) shown in Fig. 4 as the standard interconnection for the idealized linear design. The exogenous input w includes all signals which enter the system from its environment such as commands, disturbances and sensor noise. The input u represents the control effort applied to the plant by the controller $K(s)$. The interconnection outputs z and y_m represent the controlled output which the controller is designed to keep small (e.g. tracking error) and all measurements available to the controller (including commands, measured disturbances, measured plant inputs) respectively. Any feedforward/feedback interconnection of linear system elements can be brought into this general interconnection form. As an example, we consider the error feedback system of Fig 1. The exogenous inputs are the command r, and output disturbance d. Thus, we define $w = \begin{bmatrix} r \\ d \end{bmatrix}$. The controlled output is the tracking error, $e = r - y_{out}$, so we define $z = e$. The information made available to the controller, $K(s)$, is the tracking error, so $y_m = e$. The output of $K(s)$ is the plant input, u. With these definitions, the interconnection $P(s)$ is given by

$$P(s) = \begin{bmatrix} I & -I & -G(s) \\ I & -I & -G(s) \end{bmatrix} \tag{1}$$

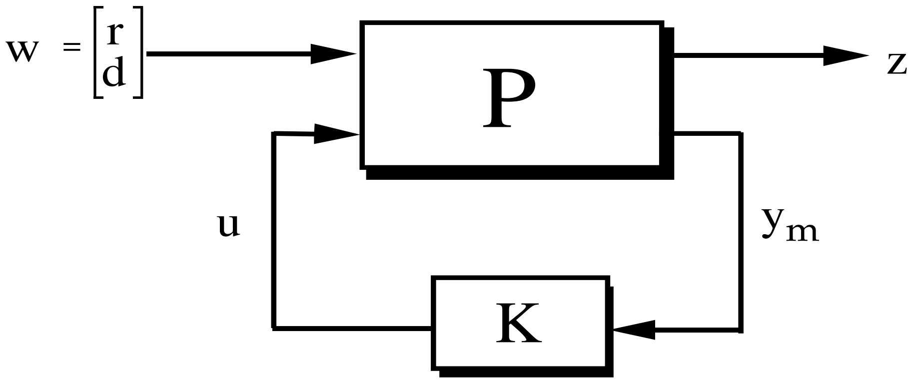

Figure 4: Idealized linear design

With these definitions, the input-output behavior from the exogenous input to the controlled output of the system in Fig. 4 is equivalent to that in Fig. 1.

$P(s)$ and $K(s)$ are assumed to be finite dimensional LTI systems whose state space realizations are assumed to be available. The closed loop transfer function from w to z in Fig. 4 is denoted by $T_{zw}(s)$, and is given by the linear fractional transformation

$$T_{zw}(s) = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21} \tag{2}$$

where

$$P(s) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \tag{3}$$

We assume that performance specifications are provided for the linear design and that the controller $K(s)$ meets these specifications in the absence of limitations and substitutions. For the purpose of this paper we assume that these specifications are of the form

$$\|T_{zw}(s)\| < 1 \tag{4}$$

where the norm, $\|\bullet\|$, is either the $\mathcal{H}_\infty$ norm or the $\mathcal{H}_2$ norm. These frequency domain performance specifications are standard in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ optimal control theory. By including suitable weights in the interconnection $P(s)$, the above performance requirement allows very general specification of the frequency domain characteristics of the closed loop transfer function. The general AWBT problem is based on Fig. 5. The interconnection $\hat{P}$(s) is obtained from $P(s)$ by providing an additional output u_m. Thus,

$$\hat{P}(s) = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \\ P_{31} & P_{32} \end{bmatrix} \tag{5}$$

where

$$u_m = P_{31}w + P_{32}\hat{u} \tag{6}$$

The new signal, u_m, is the measured or estimated value of the actual plant input $\hat{u}$. We allow the general relation (6) to account for measurement noise entering through w (i.e. $P_{31} \not\equiv 0$) and non-trivial measurement dynamics ($P_{32} \not\equiv I$). The situation where a perfect estimate of $\hat{u}$ is available corresponds to $P_{31} \equiv 0$, $P_{32} \equiv I$. As in the error feedback example (Figure 3), the plant input estimate is made available to the controller $\hat{K}(s)$ as a component of the measurement vector y.

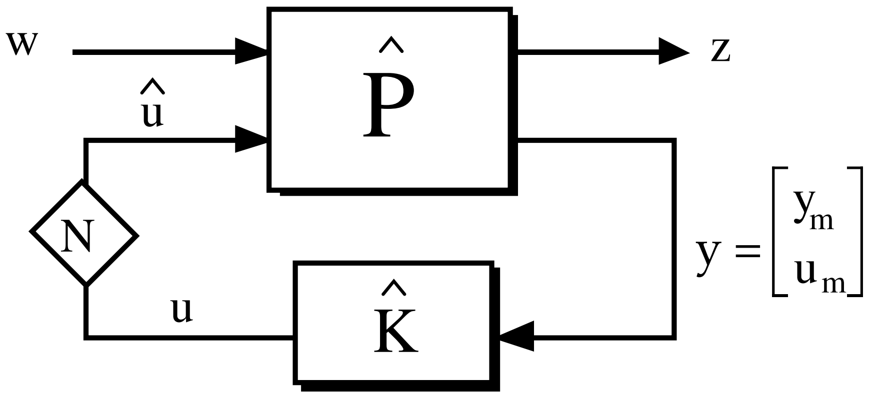

Figure 5: The general AWBT problem

Also included in Fig. 5 is the input limitation/substitution mechanism, represented by the nonlinear block N. N is assumed to be conic sector bounded and of fixed structure (these will be defined in Section 4).

Given this framework, the general AWBT problem amounts to the synthesis of $\hat{K}$(s) which renders the system in Fig. 5 stable, meets our linear performance specifications when $N \equiv I$, and exhibits graceful performance degradation when $N \not\equiv I$.

2.2 Decomposition of $\hat{K}$(s)

Consider Fig. 6 where we express $\hat{K}(s)$ as a feedback interconnection of an LTI block $\hat{\mathcal{K}}(s)$ and an AWBT operator Λ. This linear fractional feedback representation is quite general since at this point we allow Λ to be any, perhaps nonlinear relation. $\hat{\mathcal{K}}(s)$ contains the linear design $K(s)$. The AWBT operator Λ uses information provided to it by $\hat{\mathcal{K}}(s)$ in the form of the input v to generate an AWBT action denoted ξ which is fed back to $\hat{\mathcal{K}}(s)$. In order to maintain complete generality, we provide the AWBT operator, Λ, with full information in $\hat{\mathcal{K}}(s)$, including the state x of $\hat{\mathcal{K}}(s)$ and the input $\begin{bmatrix} y \\ \xi \end{bmatrix}$ to $\hat{\mathcal{K}}(s)$. Partitioning the AWBT action as $\xi = \begin{bmatrix} \xi_1 \\ \xi_2 \end{bmatrix}$ we allow it to act on the state of $\hat{\mathcal{K}}(s)$ via ξ_1 and the output of $\hat{\mathcal{K}}(s)$ via ξ_2. This gives rise to the following realization of $\hat{\mathcal{K}}(s)$

$$\hat{\mathcal{K}}(s) = \left[\begin{array}{c|cccc} A & B & 0 & I & 0 \\ \hline C & D & 0 & 0 & I \\ I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{array}\right] \quad \text{where} \quad K(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \quad \text{and} \quad v = \begin{bmatrix} x \\ y_m \\ u_m \\ \xi_1 \\ \xi_2 \end{bmatrix} \tag{7}$$

Since the state and the input to $\hat{\mathcal{K}}(s)$ fully characterize its output, we say that Λ is provided with full information (FI). Similarly, Λ can drive both the state and the output of $\hat{\mathcal{K}}(s)$ and hence acts with full control (FC). Note that for $\Lambda = 0$, i.e. no corrective AWBT action, we have $\hat{K}(s) = \begin{bmatrix} K(s) & 0 \end{bmatrix}$ which is as expected since in that case, we just have the original linear interconnection of Fig. 4 but with the nonlinearity N between the output of the controller and the input to the plant.

We now impose two criteria for the admissibility of the AWBT operator Λ:

1. $\Lambda : v \rightarrow \xi$ is causal, linear, and time invariant.

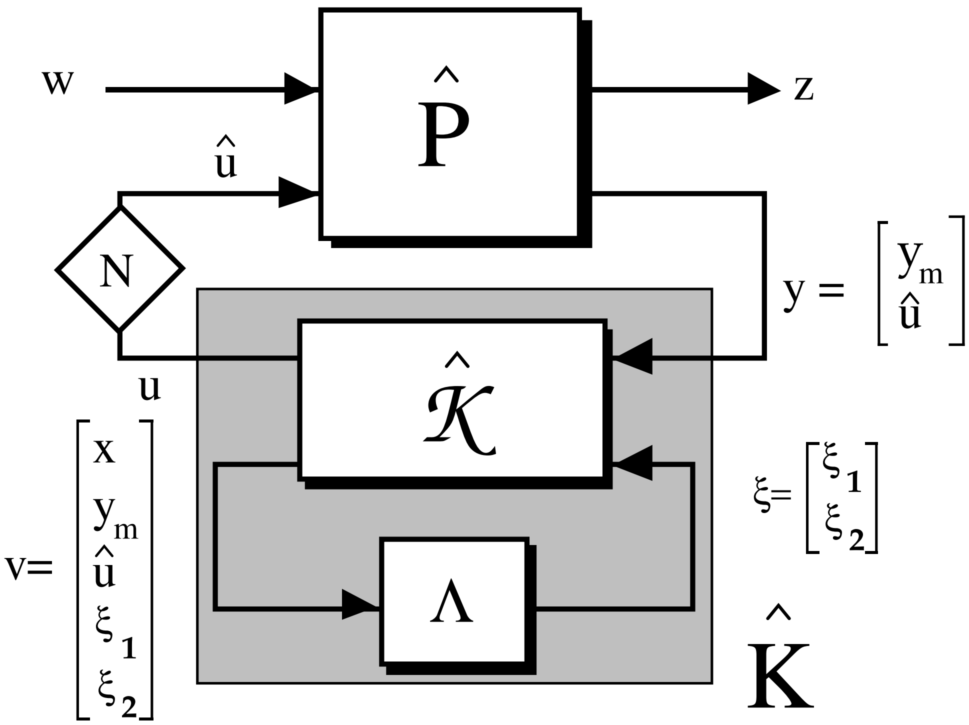

Figure 6: Decomposition of $\hat{K}(s)$

2. $\forall t, u(t) - u_m(t) = 0 \Rightarrow \xi(t) = 0$.

The first condition ensures that the AWBT compensated controller $\hat{K}(s)$ can be realized as an LTI system. As we had stated before, we will be seeking linear AWBT compensation. Moreover, most existing AWBT schemes satisfy this condition. Hence this condition seems reasonable.

The second condition enforces the notion that we do not want the AWBT block Λ to affect the linear closed loop performance achieved by the idealized linear design $K(s)$ when there is no substitution or limitation. Strictly speaking, we would like to have $\xi(t) = 0$ whenever $u(t) - \hat{u}(t) = 0$. In general, since $\hat{u}$ is not available to Λ but only an estimate u_m is available, we cannot enforce the strict linear performance recovery requirement but instead choose to impose a more realistic but weaker condition based on the measurement u_m.

These two criteria imply that any admissible Λ must be memoryless and hence a constant matrix. The two criteria also imply that $\xi(t)$ must be linear in $u_m(t) - u(t)$. Thus,

$$\begin{aligned} \xi &= \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix} (u_m - u) & (8) \\ &= \begin{bmatrix} \Lambda_1 \\ \Lambda_2 \end{bmatrix} \begin{bmatrix} -C & -D & I & 0 & -I \end{bmatrix} v & (9) \\ &= \Lambda v & (10) \end{aligned}$$

Incorporating the AWBT block, Λ, into $\hat{\mathcal{K}}$ we obtain the standard setup of Fig. 5 with

$$\hat{K}(s) = [U(s) \quad I - V(s)] \tag{11}$$

where

$$V(s) = \left[\begin{array}{c|c} A - H_1 C & -H_1 \\ \hline H_2 C & H_2 \end{array}\right] \tag{12}$$

$$U(s) = \left[\begin{array}{c|c} A - H_1 C & B - H_1 D \\ \hline H_2 C & H_2 D \end{array}\right] \tag{13}$$

$$H_1 = \Lambda_1 (I + \Lambda_2)^{-1} \tag{14}$$

$$H_2 = (I + \Lambda_2)^{-1} \tag{15}$$

A necessary condition for well-posedness of the AWBT feedback loop in Figure 6 is that $I + \Lambda_2$ must be nonsingular. Thus, H_2 must be invertible.

The blocks $U(s)$ and $V(s)$ which define the AWBT compensated controller $\hat{K}(s)$ correspond to left coprime factors of $K(s)$. It is easy to verify that

$$K(s) = V(s)^{-1} U(s) \tag{16}$$

for any H_1 and H_2 provided that H_2 is invertible.

If we assume that the realization of $K(s)$ is such that (C, A) is observable, then the eigenvalues of $A - H_1 C$ may be arbitrarily assigned by the selection of H_1. Thus if the eigenvalues of $A - H_1 C$ are chosen to be in the open left half plane, then $U(s)$, $V(s)$ and $\hat{K}(s)$ are stable. Since we will be interested in globally stable systems, we will restrict attention to the case where $\hat{P}(s)$, $U(s)$, $V(s)$ and $\hat{K}(s)$ are stable. This is because global stability of the closed loop with the actuator nonlinearity cannot be guaranteed if either $\hat{K}$ or $\hat{P}$ are unstable. For example, a mode switch from automatic to manual control will leave the loop open. If $\hat{K}$ or $\hat{P}$ are unstable, they will exhibit their unstable characteristics when the system is operating in open loop.

To demonstrate the implementation of the AWBT controller $\hat{K}(s)$, we consider the special case $P_{31} \equiv 0$, $P_{32} \equiv I$ which corresponds to $u_m \equiv \hat{u}$. The input to $\hat{K}$ is $\begin{bmatrix} y_m \\ \hat{u} \end{bmatrix}$. Since $\hat{K}(s) = \begin{bmatrix} U(s) & I - V(s) \end{bmatrix}$, we have

$$u = U(s) y_m + (I - V(s)) \hat{u} \tag{17}$$

This implementation is shown in Fig. 7. Obviously, when $N \equiv I$, we have $\hat{u} \equiv u$ and then, from equation (17), we have

$$\begin{aligned} u &= U(s) y_m + (I - V(s)) u \\ \Rightarrow V(s) u &= U(s) y_m \\ \Rightarrow u &= V(s)^{-1} U(s) y_m \\ &= K(s) y_m \end{aligned}$$

Thus, in this case, when $N \equiv I$, the ideal linear design is recovered exactly.

In general, however, the AWBT implementation is not equivalent to the idealized linear design, even when there are no limitations and substitutions, since $P_{31} \not\equiv 0$ and $P_{32} \not\equiv I$. To see this, we evaluate $T_{zw}(s)$ for the system in Fig. 4 with $N \equiv I$.

$$T_{zw}(s) = P_{11} + P_{12}[I - U P_{22} - (I - V) P_{32}]^{-1} [U P_{21} + (I - V) P_{31}] \tag{18}$$

Thus, performance is clearly different from the idealized linear design for which $T_{zw}(s)$ is given by equation (2)

$$\begin{aligned} T_{zw}(s) &= P_{11} + P_{12} K (I - P_{22} K)^{-1} P_{21} \\ &= P_{11} + P_{12} (V - U P_{22})^{-1} U P_{21} \end{aligned}$$

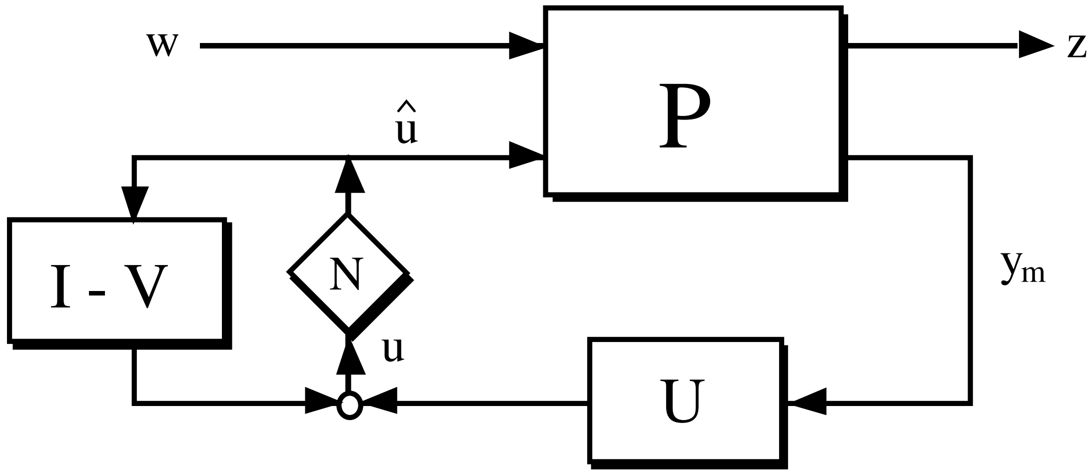

Figure 7: AWBT implementation with perfect measurement of $\hat{u}$

Of course, the two transfer functions are identical if $P_{31} \equiv 0$ and $P_{32} \equiv I$ as can be seen from equation (18) and equation (2).

3 Special cases of the general framework

In the preceding section, a fairly general and abstract framework and AWBT compensation scheme was developed. The AWBT compensated controller $\hat{K}(s)$ was decomposed into an LTI block $\hat{\mathcal{K}}(s)$ and an AWBT operator Λ. Based on certain admissibility criteria for Λ, it was shown that the only allowable Λ are constant matrices. This allowed us to parametrize all admissible AWBT compensated controllers $\hat{K}(s)$ in terms of stable left coprime factors of the initial linear controller $K(s)$. It was shown that the free parameters in the design of $\hat{K}(s)$ are two constant matrices H_1 and H_2, with the restriction that H_2 be invertible.

We will now discuss several known LTI AWBT schemes and show that they are all special cases of the framework and compensation scheme developed in the preceding section. This will enable us to unify all known (somewhat ad-hoc and problem-specific) linear AWBT compensation schemes under a general framework. Some of the schemes discussed here were originally proposed only for taking into account actuator saturation, while some allow consideration of more general actuator nonlinearities. We will use the symbol N for a general actuator nonlinearity, and the saturation block (as shown in Fig. 8) for a saturating actuator, whichever is appropriate in the context.

3.1 Anti-reset windup

Anti-reset windup [6, 4] has also been referred to as *back-calculation and tracking* [2, 11] and *integrator resetting* [20]. Windup was originally observed in PI and PID controllers designed for SISO control systems with a saturating actuator. Consider the output of a PI controller as shown in Figure 8:

$$K(s) = k(1 + \frac{1}{\tau_I s}) \tag{19}$$

$$= \left[\begin{array}{c|c} 0 & \frac{k}{\tau_I} \\ \hline 1 & k \end{array}\right] \tag{20}$$

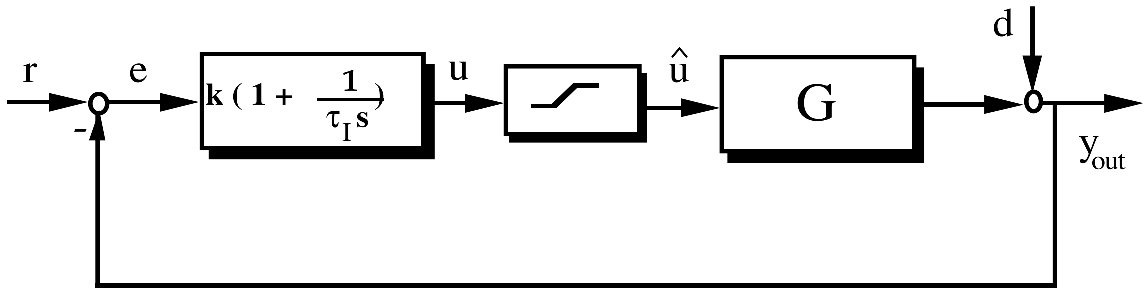

Figure 8: PI control of plant $G(s)$

$$
\begin{aligned}
u &= k(e + \frac{1}{\tau_I}\int_0^t e\ dt) && (21)\\
\hat{u} &= \mathrm{sat}(u) && (22)\\
&= \begin{cases} u_{min} & \text{if } u < u_{min} \\ u & \text{if } u_{min} \le u \le u_{max} \\ u_{max} & \text{if } u > u_{max} \end{cases} && (23)\\
e &= r - y_{out}. && (24)
\end{aligned}
$$

If the error e is positive for a substantial time, the control signal gets saturated at the high limit u_{max}. If the error remains positive for some time subsequent to saturation, the integrator continues to accumulate the error causing the control signal to become "more" saturated. The control signal remains saturated at this point because of the large value of the integral. It does not leave the saturation limit until the error becomes negative and remains negative for a sufficiently long time to allow the integral part to come down to a small value. The adverse effect of this integral windup is in the form of large overshoots in the output y_{out} and sometimes even instability.

To avoid windup, an extra feedback path is provided in the controller by measuring the actuator output $\hat{u}$ and forming an error signal as the difference between the output u of the controller and the actuator output $\hat{u}$. This error signal is fed to the input of the integrator through the gain $\frac{1}{\tau_r}$. The controller equations thus modified are (refer to Figure 9)

$$
\begin{aligned}
u &= k\left[e + \frac{1}{\tau_I}\int_0^t \left(e - \frac{\tau_I}{k\tau_r}(u - \hat{u})\right)\ dt\right] && (25)\\
\hat{u} &= \mathrm{sat}(u) && (26)\\
e &= r - y_{out} && (27)
\end{aligned}
$$

When the actuator saturates, the feedback signal $u - \hat{u}$ attempts to drive the error $u - \hat{u}$ to zero by recomputing the integral such that the controller output is exactly at the saturation limit. This prevents the integrator from winding up.

Rewriting equation (25) in the Laplace domain

$$
\begin{aligned}
u &= \left[ke + \frac{k}{\tau_I s}e - \frac{1}{\tau_r s}(u - \hat{u})\right]\\
\Rightarrow u &= \frac{k\tau_r(1 + \tau_I s)}{\tau_I(1 + \tau_r s)}e + \frac{1}{\tau_r s + 1}\hat{u}
\end{aligned}
$$

In the general framework of Fig. 5,

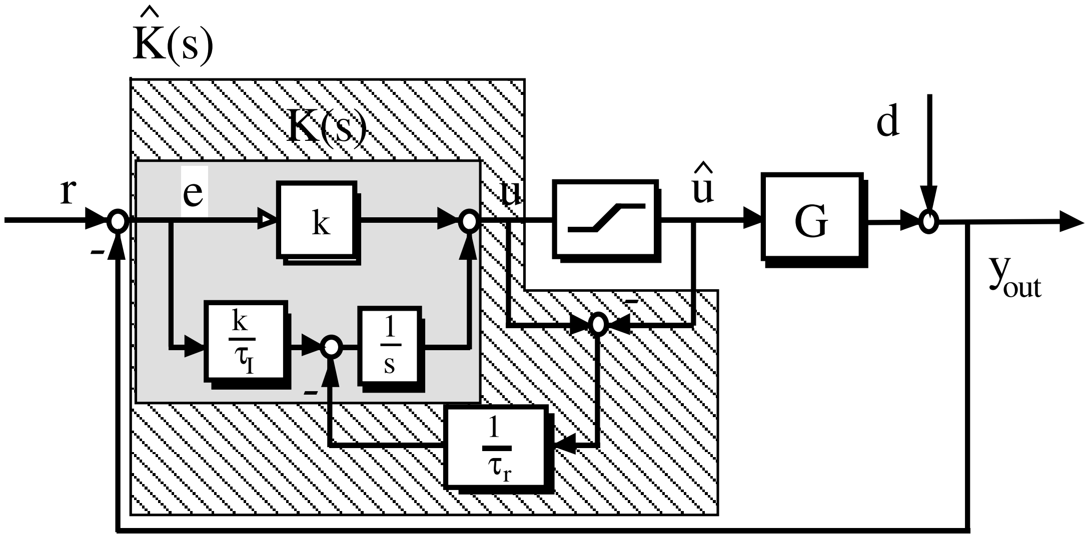

Figure 9: Classical anti-reset windup

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad y = \begin{bmatrix} e \\ \hat{u} \end{bmatrix}, \quad z = e$$

$$\begin{aligned} \hat{K}(s) &= \begin{bmatrix} \frac{k\tau_r(1+\tau_I s)}{\tau_I(1+\tau_r s)} & \frac{1}{\tau_r s+1} \end{bmatrix} \\ \hat{P}(s) &= \begin{bmatrix} I & -I & -G(s) \\ I & -I & -G(s) \\ 0 & 0 & I \end{bmatrix} \end{aligned}$$

A realization of the anti-reset windup compensator $\hat{K}(s)$ is given by

$$\hat{K}(s) = \left[\begin{array}{c|cc} -\frac{1}{\tau_r} & \frac{k}{\tau_I\tau_r}(\tau_r - \tau_I) & \frac{1}{\tau_r} \\ \hline 1 & k & 0 \end{array}\right] \tag{28}$$

Comparing equation (28) with equations (11), (12), (13), we see that the anti-reset windup implementation corresponds to the choices

$$U(s) = \left[\begin{array}{c|c} -\frac{1}{\tau_r} & \frac{k}{\tau_I\tau_r}(\tau_r - \tau_I) \\ \hline 1 & k \end{array}\right] \tag{29}$$

$$V(s) = \left[\begin{array}{c|c} -\frac{1}{\tau_r} & -\frac{1}{\tau_r} \\ \hline 1 & 1 \end{array}\right] \tag{30}$$

$$H_1 = \frac{1}{\tau_r} \tag{31}$$

$$H_2 = 1 \tag{32}$$

in the general framework of Section 2, Fig. 5.

For a PI controller, when integral action is generated as an automatic reset, Åström and Hagglund [1] suggest the implementation shown in Fig. 10 to achieve anti-reset windup compensation. When there is no saturation, it is easy to verify that this implementation results in the standard

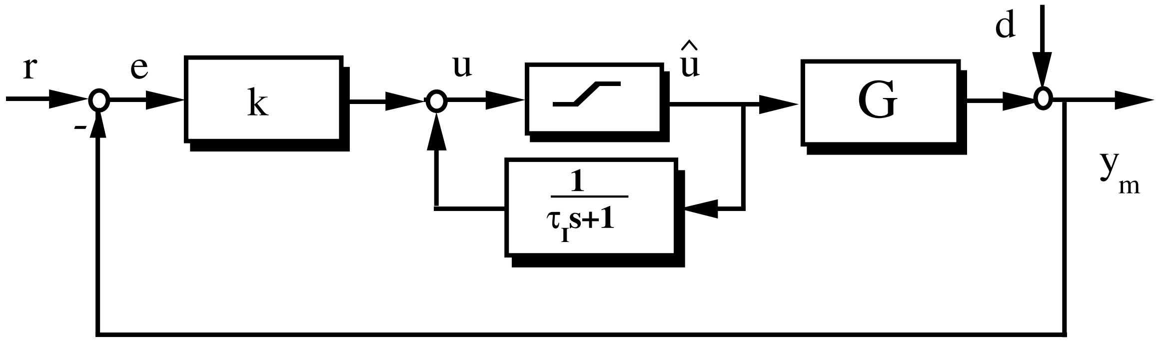

Figure 10: Alternate anti-reset windup implementation

PI controller

$$K(s) \;=\; k(1+\frac{1}{\tau_I s}) = \left[\begin{array}{c|c} 0 & \frac{k}{\tau_I} \\ \hline 1 & k \end{array}\right] \tag{33}$$

In the presence of saturation, the control signal (in Laplace domain) is given by

$$u \;=\; ke + \frac{1}{1+\tau_I s}\hat{u} \tag{34}$$

In the general framework of Section 2, Fig. 5,

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad y = \begin{bmatrix} e \\ \hat{u} \end{bmatrix}, \quad z = e$$

$$\hat{K}(s) \;=\; \begin{bmatrix} k & \frac{1}{1+\tau_I s} \end{bmatrix} = \left[\begin{array}{c|cc} -\frac{1}{\tau_I} & 0 & \frac{1}{\tau_I} \\ \hline 1 & k & 0 \end{array}\right] \tag{35}$$

$$U(s) \;=\; k = \left[\begin{array}{c|c} -\frac{1}{\tau_I} & 0 \\ \hline 1 & k \end{array}\right] \tag{36}$$

$$V(s) \;=\; \frac{\tau_I s}{1+\tau_I s} = \left[\begin{array}{c|c} -\frac{1}{\tau_I} & -\frac{1}{\tau_I} \\ \hline 1 & 1 \end{array}\right] \tag{37}$$

$$H_1 \;=\; \frac{1}{\tau_I} \tag{38}$$

$$H_2 \;=\; 1 \tag{39}$$

It is interesting to note that this anti-reset windup implementation corresponds to $\tau_r = \tau_I$ in the classical anti-reset windup of Fig. 9. Thus, these seemingly different anti-reset windup schemes for PI controllers are identical for the well known heuristic choice of $\tau_r = \tau_I$.

3.2 Conventional Anti-windup (CAW)

High gain conventional conventional anti-windup (CAW) adopts a philosophy similar to that of anti-reset windup. Thus, in some sense CAW can be considered as a direct extension of anti-reset windup to general controllers. The implementation is shown in Fig. 11. The AWBT compensation is provided by feeding the difference $\hat{u} - u$ through a high gain matrix X to the controller input e.

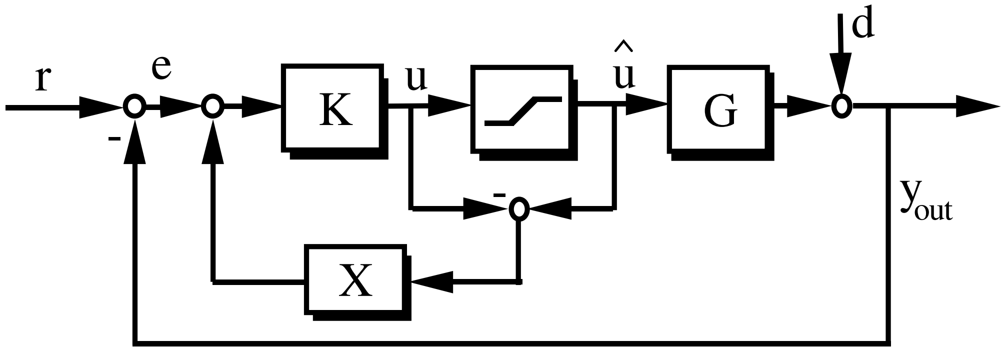

Figure 11: Conventional Anti-windup

Typically, $X = \alpha I$, where $\alpha >> 1$ is a scalar. Given the original linear controller $K(s)$ with state x

$$K(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \tag{40}$$

the modified controller equations based on Fig. 11 are

$$\begin{aligned} \dot{x} &= Ax + B(e + X(\hat{u} - u)) && (41)\\ u &= Cx + D(e + X(\hat{u} - u)) && (42)\\ \Longrightarrow u &= (I + DX)^{-1}Cx + (I + DX)^{-1}De + (I + DX)^{-1}DX\hat{u} && (43) \end{aligned}$$

Substituting equation (43) into equation (41), we get

$$\begin{aligned} \dot{x} &= (A - BX(I + DX)^{-1}C)x + (B - BX(I + DX)^{-1}D)e && (44)\\ &\quad +(BX - BX(I + DX)^{-1}DX)\hat{u} && (45)\\ &= (A - BX(I + DX)^{-1}C)x + B(I + XD)^{-1}e + BX(I + DX)^{-1}\hat{u} && (46) \end{aligned}$$

In the general framework of Section 2, Fig. 5,

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad y = \begin{bmatrix} e \\ \hat{u} \end{bmatrix}, \quad z = e$$

$$\hat{P}(s) = \begin{bmatrix} I & -I & -G(s) \\ I & -I & -G(s) \\ 0 & 0 & I \end{bmatrix} \tag{47}$$

$$\hat{K}(s) = \left[\begin{array}{c|cc} A - BX(I + DX)^{-1}C & B(I + XD)^{-1} & BX(I + DX)^{-1} \\ \hline (I + DX)^{-1}C & (I + DX)^{-1}D & (I + DX)^{-1}DX \end{array}\right] \tag{48}$$

Comparing the realization of $\hat{K}(s)$ with equations (11), (12), (13), we see that CAW corresponds to

$$U(s) = \left[\begin{array}{c|c} A - BX(I + DX)^{-1}C & B(I + XD)^{-1} \\ \hline (I + DX)^{-1}C & (I + DX)^{-1}D \end{array}\right] \tag{49}$$

$$
\begin{aligned}
V(s) &= \left[\begin{array}{c|c} A - BX(I+DX)^{-1}C & -BX(I+DX)^{-1} \\ \hline (I+DX)^{-1}C & (I+DX)^{-1} \end{array}\right] && (50) \\
H_1 &= BX(I+DX)^{-1} && (51) \\
H_2 &= (I+DX)^{-1} && (52)
\end{aligned}
$$

in the general framework of Section 2, Fig. 5.

3.3 Hanus conditioned controller

The conditioning technique was originally formulated by Hanus *et al.* [14, 15] as an extension of the back calculation method proposed by Fertik and Ross [11]. In this technique, windup is interpreted as a lack of consistency between the internal states of the controller and input to the plant when there is a nonlinearity between the controller output and the control input to the plant. Consistency is restored by modifying the inputs to the controller such that if these modified inputs (the so-called "realizable references") had been applied to the controller, its output would not have been different from the control input to the plant.

Consider a simple error feedback controller as shown in Fig. 2, with the nonlinearity N being a saturating actuator.

$$
\begin{aligned}
\dot{x} &= Ax + B(r - y_{out}) && (53) \\
u &= Cx + D(r - y_{out}) && (54) \\
\hat{u} &= \mathrm{sat}(u) && (55)
\end{aligned}
$$

where sat is defined in equation (23).

Following Hanus *et al.* [15], we can apply a realizable reference r^r to the controller such that the output of the controller is $\hat{u}$. Thus,

$$
\begin{aligned}
\dot{x} &= Ax + B(r^r - y_{out}) && (56) \\
\hat{u} &= Cx + D(r^r - y_{out}) && (57)
\end{aligned}
$$

Based on the assumption of "present realizability" (Hanus *et al.* [15]) of the control u, we get

$$u = Cx + D(r - y_{out}) \qquad (58)$$

for the same state x which results from equation (56) after application of the realizable reference r^r. Subtracting equation (58) from equation (57), we get

$$\hat{u} - u = D(r^r - r) \qquad (59)$$

Assuming D is non-singular (i.e., the linear controller $K(s)$ is biproper with $K(\infty)$ invertible) we get

$$r^r = r + D^{-1}(\hat{u} - u) \qquad (60)$$

Combining equations (55), (56), (58), (60), we get

$$
\begin{aligned}
\dot{x} &= (A - BD^{-1}C)x + BD^{-1}\hat{u} && (61) \\
u &= Cx + D(r - y_{out}) && (62) \\
\hat{u} &= sat(u) && (63)
\end{aligned}
$$

This is the AWBT "conditioned controlled".

In the general framework of Section 2, Fig. 5, we have

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad z = r - y_{out}, \quad y = \begin{bmatrix} r - y_{out} \\ \hat{u} \end{bmatrix}$$

A state space realization of the AWBT controller based on the conditioned controller from equations (61), (62), (63) is

$$\hat{K}(s) = \left[\begin{array}{c|cc} A - BD^{-1}C & 0 & BD^{-1} \\ \hline C & D & 0 \end{array}\right] \tag{64}$$

Comparing equation (64) with equations (11),(12),(13), we see that the Hanus conditioned controller corresponds to

$$\begin{aligned} U(s) &= \left[\begin{array}{c|c} A - BD^{-1}C & 0 \\ \hline C & D \end{array}\right] && (65) \\ V(s) &= \left[\begin{array}{c|c} A - BD^{-1}C & -BD^{-1} \\ \hline C & I \end{array}\right] && (66) \\ H_1 &= BD^{-1} && (67) \\ H_2 &= I && (68) \end{aligned}$$

in the general framework of Section 2, Fig. 5.

3.4 Observer based anti-windup

As pointed out before, an interpretation of the windup problem is that the states of the controller do not correspond to the control signal being fed to the plant. This inaccuracy in the state vector of the controller is due to lack of correct estimates of the controller states in the presence of actuator nonlinearities. To obtain correct state estimates and to avoid windup, Åström *et al.* [1, 2] suggest that an observer be introduced into the controller.

Referring to Fig. 4, let the linear controller $K(s)$ be defined by the equations

$$\begin{aligned} \dot{x} &= Ax + By_m && (69) \\ u &= Cx + Dy_m && (70) \end{aligned}$$

Let us assume that there is a nonlinearity N between the controller output and the control input to the plant $P(s)$ so that the input to the plant is given by

$$\hat{u} = N(u) \tag{71}$$

Following Åström *et al.* [1, 2], the nonlinear observer for the controller $K(s)$ (assuming (C,A) detectable) is defined by

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + By_m + L(\hat{u} - C\hat{x} - Dy_m) && (72) \\ u &= C\hat{x} + Dy_m && (73) \\ \hat{u} &= N(u) && (74) \end{aligned}$$

where $\hat{x}$ is an estimate of the controller state and L is the observer gain. Instead of having a separate controller and a separate observer, both are integrated into one scheme to form the AWBT compensator. Thus, the observer comes into the controller structure only in the presence of the actuator nonlinearity ($N \not\equiv I$) and does not affect the linear controller ($N \equiv I$).

In the general framework of Section 2, Fig. 5, a realization for the AWBT compensator described by equations (72), (73) is given by

$$\hat{K}(s) = \left[\begin{array}{c|cc} A - LC & B - LD & L \\ \hline C & D & 0 \end{array}\right] \tag{75}$$

Comparing equation (75) with equations (11), (12), (13), we see that the observer-based AWBT scheme corresponds to

$$U(s) = \left[\begin{array}{c|c} A - LC & B - LD \\ \hline C & D \end{array}\right] \tag{76}$$

$$V(s) = \left[\begin{array}{c|c} A - LC & -L \\ \hline C & I \end{array}\right] \tag{77}$$

$$H_1 = L \tag{78}$$

$$H_2 = I \tag{79}$$

in the general framework of Section 2.

3.5 Internal Model Control

The internal model control (IMC) structure [18, pages 44–45] was never intended to be an anti-windup scheme. Nonetheless, as pointed out in [6, 18, 10, 19], it has potential for application to the anti-windup problem, for the case where the system is open loop stable. The AWBT application of IMC has been studied by Cohen *et al.* [8] and Debelle [9].

Figure 12 shows the IMC structure with an actuator nonlinearity. If the controller is implemented in the IMC configuration, actuator constraints do not cause any stability problems provided the constrained control signal is sent to both the plant and the model. Under the assumption that there is no plant-model mismatch ($G = \tilde{G}$), it is easy to show [6] that the IMC structure remains effectively open loop and stability is guaranteed by the stability of the plant (G) and the IMC controller (Q). Stability of G and Q is in any case imposed by linear design and hence stability of the nonlinear system is assured. Thus the IMC structure offers the opportunity of implement-

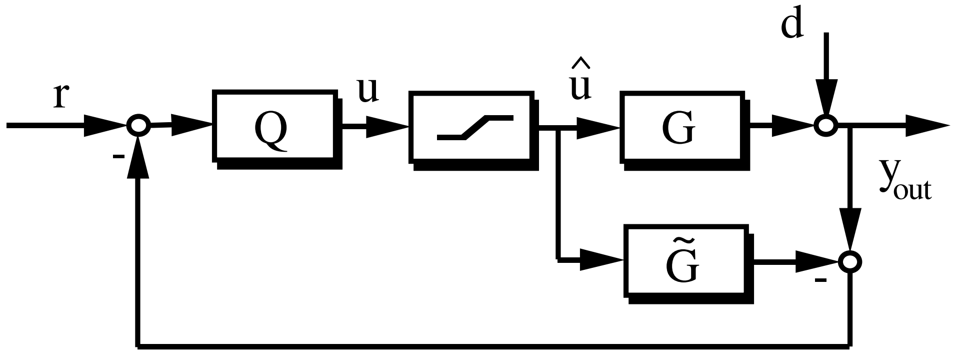

Figure 12: The IMC structure

ing complex (possibly nonlinear) control algorithms without generating complex stability issues, provided there is no plant-model mismatch.

The linear controller $K(s)$ corresponding to the IMC controller $Q(s)$ is given by

$$K(s) = [I - Q(s)\tilde{G}(s)]^{-1}Q(s) \tag{80}$$

Introducing realizations for $\tilde{G}(s)$ and $Q(s)$ (assuming both are stable)

$$Q(s) = \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & D_Q \end{array}\right], \quad \tilde{G}(s) = \left[\begin{array}{c|c} A_G & B_G \\ \hline C_G & D_G \end{array}\right], \tag{81}$$

For simplicity, we will assume that $D_G = 0$, although the case $D_G \neq 0$ can also be considered, but the algebra is messy. A realization for the linear controller $K(s)$ (using equation (80)) is given by

$$K(s) = \left[\begin{array}{cc|c} A_Q & B_Q C_G & B_Q \\ B_G C_Q & A_G + B_G D_Q C_G & B_G D_Q \\ \hline C_Q & D_Q C_G & D_Q \end{array}\right] \tag{82}$$

in the general framework of the idealized linear design of Section 2, Fig. 4, with $w = \begin{bmatrix} r \\ d \end{bmatrix}$, $z = y_m = r - y_{out}$.

Consider the IMC implementation of Fig. 12. Referring to the figure

$$\begin{aligned} y_{out} &= d + G\hat{u} && (83) \\ u &= Q[r - (y_{out} - \tilde{G}\hat{u})] && (84) \\ &= Qe + Q\tilde{G}\hat{u} && (85) \end{aligned}$$

In the general framework of Section 2, Fig. 5, this corresponds to

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad y = \begin{bmatrix} r - y_{out} \\ \hat{u} \end{bmatrix}, \quad z = r - y_{out}$$

$$\begin{aligned} \hat{P}(s) &= \begin{bmatrix} I & -I & -G(s) \\ I & -I & -G(s) \\ 0 & 0 & I \end{bmatrix} && (86) \\ \hat{K}(s) &= \begin{bmatrix} Q & Q\tilde{G} \end{bmatrix} && (87) \end{aligned}$$

Comparing equation (87) with equations (11), (12), (13), and using the realizations for Q and $\tilde{G}$ we get

$$\begin{aligned} U(s) &= Q && (88) \\ &= \left[\begin{array}{c|c} A_Q & B_Q \\ \hline C_Q & D_Q \end{array}\right] && (89) \\ &= \left[\begin{array}{cc|c} A_Q & B_Q C_G & B_Q \\ 0 & A_G & 0 \\ \hline C_Q & D_Q C_G & D_Q \end{array}\right] && (90) \\ V(s) &= I - Q\tilde{G} = \left[\begin{array}{cc|c} A_Q & B_Q C_G & 0 \\ 0 & A_G & -B_G \\ \hline C_Q & D_Q C_G & I \end{array}\right] && (91) \end{aligned}$$

Here we have introduced stable uncontrollable modes of $\tilde{G}(s)$ in the realization of $U(s)$. Comparing equations (90), (91) with equations (12), (13), we see that the IMC implementation of $K(s)$ corresponds to

$$H_1 = \begin{bmatrix} 0 \\ B_G \end{bmatrix} \tag{92}$$

$$H_2 = I \tag{93}$$

in the general framework of Section 2.

For the sake of completeness, we also include the two degree of freedom IMC implementation shown in Fig. 13. The IMC controller in this case is $Q(s) = \begin{bmatrix} Q_1(s) & Q_2(s) \end{bmatrix}$. In our general framework, the idealized linear design of Section 2, Fig. 4, corresponding to this implementation is given by

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad z = r - y_{out}, \quad y_m = \begin{bmatrix} r \\ y_{out} \end{bmatrix}$$

$$P(s) = \begin{bmatrix} I & -I & -G(s) \\ I & 0 & 0 \\ 0 & I & G(s) \end{bmatrix} \tag{94}$$

$$K(s) = \begin{bmatrix} (I - Q_2\tilde{G})^{-1}Q_1 & -(I - Q_2\tilde{G})^{-1}Q_2 \end{bmatrix} \tag{95}$$

The two degree of freedom IMC implementation, considered as an AWBT compensation in our

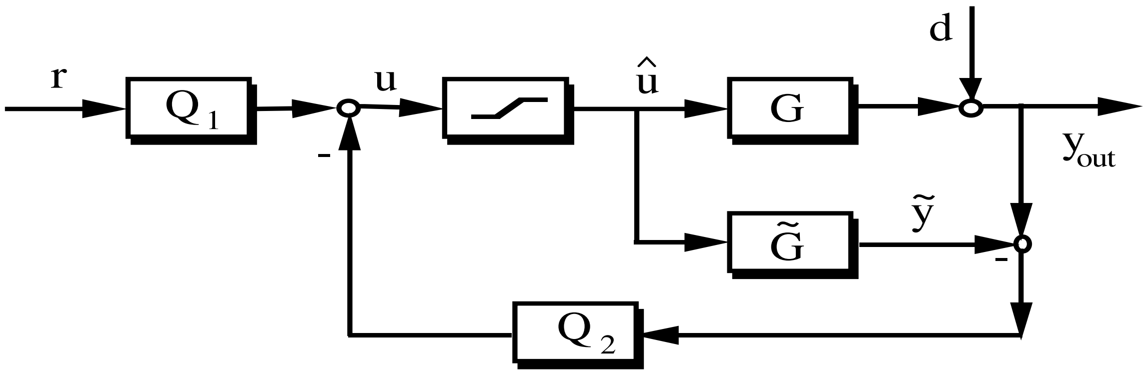

Figure 13: Two degree of freedom IMC structure

general framework of Section 2, Fig. 5 corresponds to

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad z = r - y_{out}, \quad y = \begin{bmatrix} r \\ y_{out} \\ \hat{u} \end{bmatrix}$$

$$\hat{P}(s) = \begin{bmatrix} I & -I & -G(s) \\ I & 0 & 0 \\ 0 & I & G(s) \\ 0 & 0 & I \end{bmatrix} \tag{96}$$

$$\hat{K}(s) = \begin{bmatrix} Q_1 & -Q_2 & Q_2\tilde{G} \end{bmatrix} \tag{97}$$

We use the following state space realizations for $\tilde{G}(s)$ and $Q(s)$

$$Q(s) = \left[\begin{array}{c|cc} A_Q & B_{1Q} & B_{2Q} \\ \hline C_Q & D_{1Q} & D_{2Q} \end{array}\right], \quad \tilde{G}(s) = \left[\begin{array}{c|c} A_G & B_G \\ \hline C_G & D_G \end{array}\right], \tag{98}$$

Again, for simplicity, assuming $D_G = 0$, we obtain realizations for $K(s)$ and $\hat{K}(s)$ from equations (95), (97) as

$$K(s) = \left[\begin{array}{cc|cc} A_G + B_G D_{2Q} C_G & B_G C_Q & B_G D_{1Q} & -B_G D_{2Q} \\ B_{2Q} C_G & A_Q & B_{1Q} & -B_{2Q} \\ \hline D_{2Q} C_G & C_Q & D_{1Q} & -D_{2Q} \end{array}\right] \tag{99}$$

$$\hat{K}(s) = \left[\begin{array}{cc|ccc} A_G & 0 & 0 & 0 & B_G \\ B_{2Q} C_G & A_Q & B_{1Q} & -B_{2Q} & 0 \\ \hline D_{2Q} C_G & C_Q & D_{1Q} & -D_{2Q} & 0 \end{array}\right] \tag{100}$$

Comparing equations (99), (100) with equations (11), (12), (13), we see that the two-degree of freedom IMC implementation of $K(s)$ corresponds to

$$U(s) = \left[\begin{array}{cc|cc} A_G & 0 & 0 & 0 \\ B_{2Q} C_G & A_Q & B_{1Q} & -B_{2Q} \\ \hline D_{2Q} C_G & C_Q & D_{1Q} & -D_{2Q} \end{array}\right] \tag{101}$$

$$V(s) = \left[\begin{array}{cc|c} A_G & 0 & -B_G \\ B_{2Q} C_G & A_Q & 0 \\ \hline D_{2Q} C_G & C_Q & I \end{array}\right] \tag{102}$$

$$H_1 = \begin{bmatrix} 0 \\ B_G \end{bmatrix} \tag{103}$$

$$H_2 = I \tag{104}$$

in the general framework of Section 2.

3.6 Anti-windup design for Internal Model Control (IMC)

In Section 3.5, we discussed certain advantages of the IMC implementation for systems with actuator constraints. In particular, we pointed out that if the controller is implemented in the IMC configuration, actuator constraints do not cause any stability problems provided the constrained control signal is sent to both the plant $G(s)$ and the model $\tilde{G}(s)$. When there is no plant model mismatch, $(G = \tilde{G})$, the IMC structure remains effectively open loop and stability of the system is guaranteed by the stability of the plant $G(s)$ and the IMC controller $Q(s)$.

Unfortunately, the cost to be paid for stability is in the form of somewhat "sluggish" performance. This is because the controller output is independent of the plant output in both the linear and nonlinear regimes. While this does not matter in the linear regime, its implication in the nonlinear regime is that the controller is unaware of the effect of its actions on the output, resulting in some sluggishness. This effect is most pronounced when the IMC controller has fast dynamics which are "chopped off" by the saturation. Moreover, unless the IMC controller is designed to optimize nonlinear performance, it will not give satisfactory performance for the saturating system.

A modified IMC structure for optimizing performance in the face of saturation has recently been proposed by Zheng *et al.* [21]. In this strategy, the IMC controller $Q(s)$ is factorized as

$$Q(s) = [I + Q_2(s)]^{-1} Q_1(s) \tag{105}$$

and implemented as shown in Fig. 14. Two possible choices of $Q_1(s)$ and $Q_2(s)$ have been discussed in [21] based on two different optimization criteria. Due to space limitations, we will only consider $Q_1(s)$ and $Q_2(s)$ as stable transfer functions without going into details of their specific choices. Let

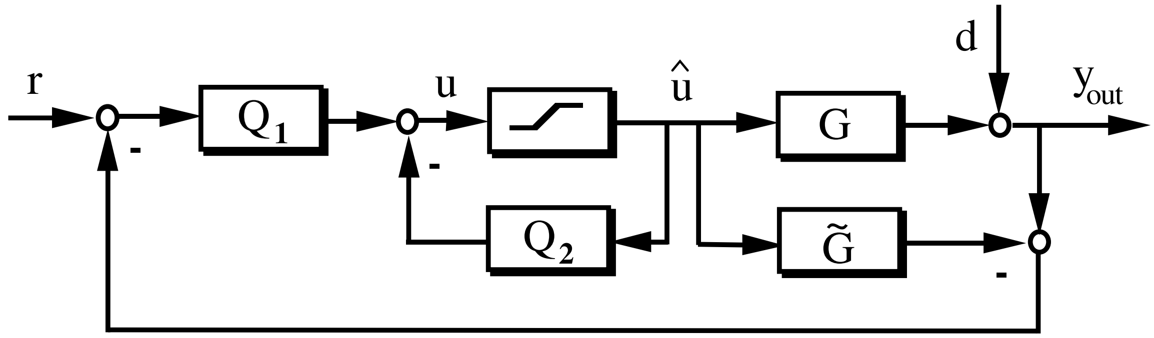

Figure 14: Anti-windup implementation of IMC

us introduce state space realizations for $Q_1(s)$, $Q_2(s)$ and $G(s) = \tilde{G}(s)$ as follows

$$Q_1(s) = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \tag{106}$$

$$Q_2(s) = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] \tag{107}$$

$$\tilde{G}(s) = \left[\begin{array}{c|c} A_G & B_G \\ \hline C_G & D_G \end{array}\right] \tag{108}$$

In order to simplify the algebra, we assume that $Q_2(s)$ and $G(s)$ are strictly proper, i.e., $D_2 = 0$, $D_G = 0$. The more general case where $Q_2(s)$ and $G(s)$ are not strictly proper can also be worked but the algebra is messy and as such does not give any additional insight into the results. Then, in terms of these realizations, the realizations of $Q(s)$ and the linear controller $K(s)$ corresponding to the IMC controller $Q(s)$ are given below

$$Q(s) = [I + Q_2(s)]^{-1} Q_1(s) \tag{109}$$

$$= \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ B_2C_1 & A_2 - B_2C_2 & B_2D_1 \\ \hline C_1 & -C_2 & D_1 \end{array}\right] \tag{110}$$

$$K(s) = [I - Q(s)\tilde{G}(s)]^{-1} Q(s) \tag{111}$$

$$= \left[\begin{array}{ccc|c} A_1 & 0 & B_1C_G & B_1 \\ B_2C_1 & A_2 - B_2C_2 & B_2D_1C_G & B_2D_1 \\ B_GC_1 & -B_GC_2 & A_G + B_GD_1C_G & B_GD_1 \\ \hline C_1 & -C_2 & D_1C_G & D_1 \end{array}\right] \tag{112}$$

Referring to Fig. 14, the control signal u is given by

$$
\begin{aligned}
u &= Q_1[r-(y_{out}-\tilde{G}\hat{u})]-Q_2\hat{u} && (113)\\
&= Q_1(r-y_{out})+(Q_1\tilde{G}-Q_2)\hat{u} && (114)
\end{aligned}
$$

In the general framework of Section 2, Fig. 5, this modified IMC implementation corresponds to

$$
w=\begin{bmatrix} r \\ d \end{bmatrix}, \quad y=\begin{bmatrix} r-y_{out} \\ \hat{u} \end{bmatrix}, \quad z=r-y_{out}
$$

$$
\begin{aligned}
\hat{P}(s) &= \begin{bmatrix} I & -I & -\tilde{G}(s) \\ I & -I & -\tilde{G}(s) \\ 0 & 0 & I \end{bmatrix} && (115)\\
\hat{K}(s) &= \begin{bmatrix} Q_1(s) & Q_1(s)\tilde{G}(s)-Q_2(s) \end{bmatrix} && (116)
\end{aligned}
$$

Comparing equation (116) with equations (11), (12), (13) we see that this implementation of IMC corresponds to

$$
\begin{aligned}
U(s) &= Q_1(s)=\left[\begin{array}{ccc|c} A_1 & 0 & B_1C_G & B_1 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & A_G & 0 \\ \hline C_1 & -C_2 & D_1C_G & D_1 \end{array}\right] && (117)\\
V(s) &= I-Q_1(s)\tilde{G}(s)+Q_2(s) && (118)\\
&= \left[\begin{array}{ccc|c} A_1 & 0 & B_1C_G & 0 \\ 0 & A_2 & 0 & -B_2 \\ 0 & 0 & A_G & -B_G \\ \hline C_1 & -C_2 & D_1C_G & I \end{array}\right] && (119)
\end{aligned}
$$

Note that we have introduced stable uncontrollable modes of $\tilde{G}(s)$ in the realization of $U(s)$. The corresponding values of H_1 and H_2 are given by

$$
\begin{aligned}
H_1 &= \begin{bmatrix} 0 \\ B_2 \\ B_G \end{bmatrix} && (120)\\
H_2 &= I && (121)
\end{aligned}
$$

3.7 The Extended Kalman filter

The last scheme we consider here is an AWBT implementation applicable to observer-based compensators. This implementation is developed to maintain valid state estimates in the observer independent of any nonlinearities between the controller output and the plant input.

Consider the idealized linear design of Fig. 4 with $w=\begin{bmatrix} r \\ d \end{bmatrix}$, $y_m=\begin{bmatrix} r \\ y_1 \end{bmatrix}$. The plant $P(s)$ (Fig. 4) with the state x is described by the state-space equations

$$
\begin{aligned}
\dot{x} &= Ax+B_1r+B_2w+B_3u && (122)\\
z &= C_1x+D_{11}r+D_{12}w+D_{13}u && (123)\\
y_1 &= C_3x+D_{31}r+D_{32}w && (124)
\end{aligned}
$$

i.e.,

$$P(s) = \left[\begin{array}{c|ccc} A & B_1 & B_2 & B_3 \\ \hline C_1 & D_{11} & D_{12} & D_{13} \\ 0 & I & 0 & 0 \\ C_3 & D_{31} & D_{32} & 0 \end{array}\right] \tag{125}$$

Implicit in this realization is the assumption that the command r is available to $K(s)$ without noise.

Let $K(s)$ be the standard observer/state-feedback controller with state $\hat{x}$, whose state-space equations can be described by

$$\dot{\hat{x}} = A\hat{x} + B_1 r + B_3 u + L(y_1 - C_3\hat{x} - D_{31}r) \tag{126}$$

$$u = -F\hat{x} \tag{127}$$

i.e.,

$$K(s) = \left[\begin{array}{c|cc} A - B_3F - LC_3 & B_1 - LD_{31} & L \\ \hline -F & 0 & 0 \end{array}\right] \tag{128}$$

where L is the observer gain and F is the state feedback gain.

In the presence of nonlinearity N between the controller output and the plant input, we have

$$\hat{u} = N(u) \neq u = -F\hat{x}$$

Thus, the observer in (128) will give poor estimates of the true plant state. This is because, equation (128) assumes that $\hat{u} = u = -F\hat{x}$, which is not the case in the presence of the nonlinearity N. Hence, this will result in controller windup.

To provide anti-windup compensation, the observer equations must be modified so that the state estimator is based on the actual input $\hat{u}$ to the plant. Thus, the modified observer/state-feedback compensator is given by

$$\dot{\hat{x}} = A\hat{x} + B_1 r + B_3 \hat{u} + L(y_1 - C_3\hat{x} - D_{31}r) \tag{129}$$

$$= (A - LC_3)\hat{x} + (B_1 - LD_{31})r + B_3\hat{u} + Ly_1 \tag{130}$$

$$u = -F\hat{x} \tag{131}$$

In the general framework of Section 2, Fig. 5, this AWBT compensated controller corresponds to

$$w = \begin{bmatrix} r \\ d \end{bmatrix}, \quad y = \begin{bmatrix} y_m \\ \hat{u} \end{bmatrix},$$

$$\hat{P}(s) = \left[\begin{array}{c|ccc} A & B_1 & B_2 & B_3 \\ \hline C_1 & D_{11} & D_{12} & D_{13} \\ 0 & I & 0 & 0 \\ C_3 & D_{31} & D_{32} & 0 \\ 0 & 0 & 0 & I \end{array}\right] \tag{132}$$

$$\hat{K}(s) = \left[\begin{array}{c|ccc} A - LC_3 & B_1 - LD_{31} & L & B_3 \\ \hline -F & 0 & 0 & 0 \end{array}\right] \tag{133}$$

$$U(s) = \left[\begin{array}{c|cc} A - LC_3 & B_1 - LD_{31} & L \\ \hline -F & 0 & 0 \end{array}\right] \tag{134}$$

$$V(s) = \left[\begin{array}{c|c} A - LC_3 & -B_3 \\ \hline -F & I \end{array}\right] \tag{135}$$

$$H_1 = B_3 \tag{136}$$

$$H_2 = I \tag{137}$$

It is worth mentioning that the classical separation structure property of the observer/state feedback controller is lost with this implementation, in the presence of input nonlinearities. Thus, even though $A - LC_3$ and $A - B_3F$ may have eigenvalues in the open left half plane, the overall closed loop nonlinear system need not be asymptotically stable, and examples can be constructed to demonstrate this.

4 An approach to analysis and synthesis

In the preceding section, all known LTI AWBT schemes were shown to be special cases of the general framework developed in Section 2. We will now briefly discuss how the general framework developed in this paper can handle standard issues such as the analysis of stability and performance, with or without plant uncertainty. We emphasize here that the analysis of LTI systems with actuator constraints is a problem in its own right and a detailed study of the various aspects of analysis is beyond the scope of this paper. The aim of this section is to show that such issues can be handled in the framework we have developed, much in the same fashion as is done in linear control theory.
Definition 1: $\mathcal{L}_{2e}$ is the extended $\mathcal{L}_2$ space of vector valued functions $x(t) \in \mathcal{R}^n$ with the property that

$$\|x(t)\|_T \triangleq \left[\int_0^T x^*(t)x(t)\, dt\right]^{\frac{1}{2}} < \infty$$

for all $T \geq 0$. Note that the extended $\mathcal{L}_2$ space *i.e.,* $\mathcal{L}_{2e}$ allows the consideration of common signals such as persistent disturbances and step inputs which do not belong to $\mathcal{L}_2$.
Definition 2: Any (possibly nonlinear) operator $N : \mathcal{L}_{2e} \rightarrow \mathcal{L}_{2e}$ is said to be inside Cone(C,R) where C and R are LTI operators if

$$\|N(x) - Cx\|_T \leq \|Rx\|_T \tag{138}$$

$\forall T \geq 0$, and $\forall x \in \mathcal{L}_{2e}$.

The operators C and R are referred to as the cone center and radius respectively of the operator N. A conic sector provides an LTI approximation to the input-output behavior of N. The cone center C provides an approximate output Cx for any input x. The cone radius R provides a measure of the error Rx inherent in this approximation.

For example, the SISO saturation nonlinearity (shown in Fig. 15)$N : x \rightarrow \text{sat}(x)$ where

$$\text{sat}(x) = \begin{cases} x(t) & |x(t)| \leq 1 \\ \text{sign}(x(t)) & |x(t)| > 1 \end{cases}$$

lies inside Cone($\frac{1}{2}, \frac{1}{2}$). The operator $C : x(t) \longrightarrow \frac{1}{2}x(t)$ is our linear approximation to N, and $R : x(t) \longrightarrow \frac{1}{2}x(t)$ gives us a measure of the error in this approximation (as much as 100 % in this case).

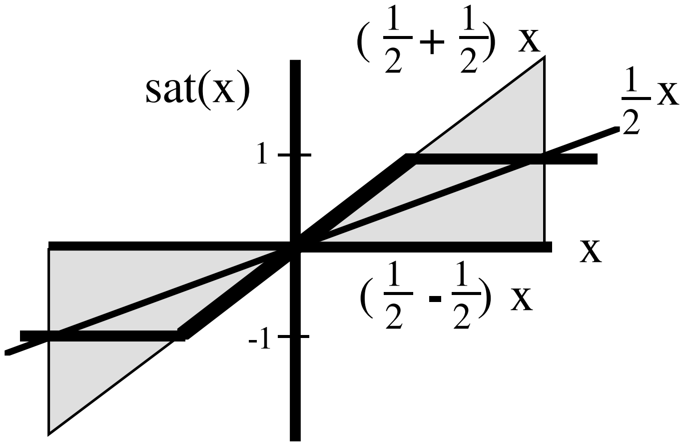

Figure 15: The SISO saturation nonlinearity

The purpose of covering the original nonlinearity by a conic sector is that the conic sector is described in terms of *linear operators*, and stability analysis for sets of nonlinearities bounded by linear operators is much more developed than stability with general nonlinearities. The approach then is to analyze stability for all nonlinearities in the conic sector, giving a sufficient condition for stability of the original nonlinearity.

The set of all nonlinearities in $\mathrm{Cone}(C, R)$ can be replaced by the LTI blocks C, R and the set of all nonlinearities in the $\mathrm{Cone}(0, I)$ because

$$\forall N \in \mathrm{Cone}(C, R) \text{ and } \forall x \in \mathcal{L}_{2e},\ \exists \Delta \in \mathrm{Cone}(0, I) \ni$$

$$N(x) = Cx + R\Delta(x) \tag{139}$$

In this way, we can normalize the nonlinearity N to the nonlinearity Δ which is norm bounded by 1 since it lies in $\mathrm{Cone}(0, I)$. Consider the interconnection shown in Fig. 16 where M is LTI and Δ

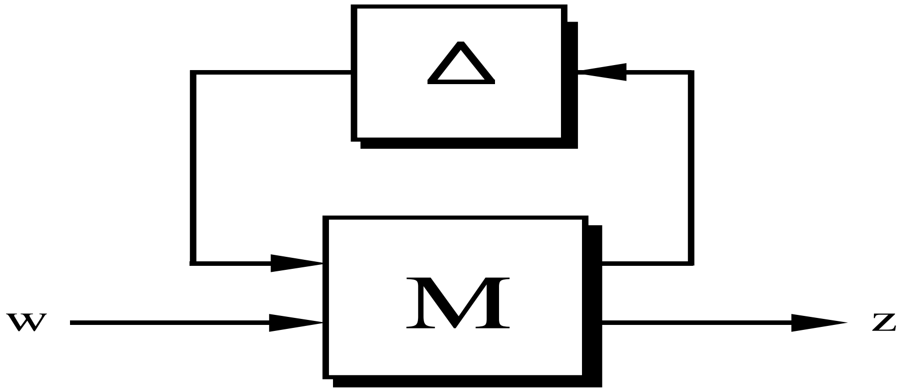

Figure 16: Interconnection for analysis

is a possibly nonlinear operator defined by the set:

$$\mathbf{\Delta} \triangleq \{\Delta \mid \Delta = \mathrm{diag}(\Delta_1, \ldots, \Delta_n) :\ \Delta_i \in \mathrm{Cone}(0, I)\} \tag{140}$$

Note that $\boldsymbol{\Delta}$ has a block diagonal structure. Fig. 16 is obtained from Fig. 5 by expressing N as in (139) and combining $\hat{P}$, $\hat{K}$, C and R into M.

Based on Fig. 16, the analysis of stability of the system can be carried out using small-gain arguments. Moreover, based on the structure of the nonlinearities, input and output scalings can be defined, much in the same way as is done in the analysis of linear systems subject to linear perturbations. In this way, conservatism in the small-gain arguments can be reduced. Linear norm-bounded uncertainties in the plant model can be incorporated by augmenting the Δ block shown above with linear Δ blocks. The corresponding compatible input and output scalings can also be defined.

Similarly, performance can be analyzed by closing the loop from z to w with a performance block. Based on the performance specifications, a controller synthesis problem can be formulated to minimize the worst case gain from w to z. Space limitations preclude a detailed discussion of these issues. These are topics of future research in the analysis of linear time-invariant systems subject to control input nonlinearities.

5 Conclusions

In this paper, we developed a general theoretical framework for studying control systems subject to plant input nonlinearities. The generality of the framework allowed us to consider any control system structure, including feedforward, feedback, multiple degree of freedom, cascade and general non-square controller designs. The theoretical development was based on the following two-step design paradigm "*Design the linear controller ignoring control input nonlinearities and then add AWBT compensation to minimize the adverse effects of any control input nonlinearities on closed loop performance*". This is characteristic of most AWBT schemes reported in the literature. A parameterization of all admissible AWBT compensated controllers was presented in terms of two constant matrices H_1 and H_2. This parameterization allowed us to unify all known LTI anti-windup and/or bumpless transfer schemes under a general framework. The framework developed was shown to hold promise in allowing the consideration of issues such as the analysis of stability and performance, with or without plant model uncertainty, much in the same fashion as is done in linear control theory.

We would like to comment that attempts to unify AWBT schemes have been reported in the past. Notable is the successful attempt by Walgama and Sternby [20] to identify the inherent observer property in a class of anti-windup compensators and to unify several schemes based on this observer property. Thus, unification of these schemes is in terms of a *single* parameter, the observer gain. However, no such observer property can be identified in the conventional anti-windup (CAW) scheme of Section 3.2 when the original linear controller $K(s)$ is not strictly proper. The parameterization we present is in terms of *two* constant matrices, H_1 and H_2. This additional degree of freedom allows us to overcome the shortcoming in the observer-based unification. Thus, as shown in Section 3.2, CAW is a special case of our scheme.

Needless to say, the aim behind our development was primarily to develop a truly general theoretical framework for AWBT controller designs. The resulting AWBT compensation scheme that we have presented and its interpretation are completely different from those reported in the literature. Specifically, the axioms and assumptions leading to our development are novel insofar as the AWBT literature is concerned. Moreover, our framework now allows us to compare and contrast various existing AWBT schemes. Thus, for example, the two seemingly different anti-reset windup strategies discussed in Section 3.1 can now be seen to be identical if $\tau_r = \tau_I$.

In summary, our parameterization of admissible AWBT compensators in terms of H_1 and H_2

allowed us to unify all known LTI AWBT schemes. Thus, rather than employing the older ad-hoc and problem-specific methodologies for AWBT compensation, we can now embark on the development of systematic procedures for choosing H_1 and H_2 for the synthesis of the AWBT compensator $\hat{K}(s)$. For this purpose, quantitative design criteria for AWBT must be defined. An intrinsic part of this step is the complete analysis of systems subject to control input nonlinearities. In Section 4, we briefly outlined how this could be done. Detailed study of the analysis theory and subsequently, AWBT synthesis, are topics of ongoing research.
Acknowledgement: Financial support for this project from the National Science Foundation is gratefully acknowledged.

References

[1] K. J. Åström and T. Hägglund. *Automatic Tuning of PID Controllers.* Instrument Society of America, Research Triangle Park, N.C., 1988.

[2] K. J. Åström and L. Rundqwist. Integrator windup and how to avoid it. In *Proceedings of the 1989 American Control Conference*, pages 1693–1698, 1989.

[3] K. J. Åström and B. Wittenmark. *Computer Controlled Systems Theory and Design.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1984.

[4] P. S. Buckley. Designing override and feedforward controls. *Control Engineering*, 18(8):48–51, 1971.

[5] P. J. Campo. *Studies in robust control of systems subject to constraints.* PhD thesis, California Institute of Technology, 1990.

[6] P. J. Campo and M. Morari. Robust control of processes subject to saturation nonlinearities. *Computers & Chemical Engineering*, 14(4/5):343–358, 1990.

[7] P. J. Campo, M. Morari, and C. N. Nett. Multivariable anti-windup and bumpless transfer: A general theory. In *Proceedings of the 1989 American Control Conference*, pages 1706–1711, 1989.

[8] I. Cohen, R. Hanus, and J. P. Vanbergen. The impact of the control structure on the closed loop efficiency. *Journal A*, 26(1):18–25, 1985.

[9] J. Debelle. A control structure based upon process models. *Journal A*, 20(2):71–81, 1979.

[10] J. C. Doyle, R. S. Smith, and D. F. Enns. Control of plants with input saturation nonlinearities. In *Proceedings of the 1987 American Control Conference*, pages 1034–1039, 1987.

[11] H. A. Fertik and C. W. Ross. Direct digital control algorithm with anti-windup feature. *ISA Transactions*, 6(4):317–328, 1967.

[12] A. H. Glattfelder and W. Schaufelberger. Stability of discrete override and cascade-limiter single-loop control systems. *IEEE Transactions on Automatic Control*, AC-33(6):532–540, June 1988.

[13] A. H. Glattfelder, W. Schaufelberger, and H. P. Fassler. Stability of override control systems. *International Journal of Control*, 37(5):1023–1037, 1983.

[14] R. Hanus and M. Kinnaert. Control of constrained multivariable systems using the conditioning technique. In *Proceedings of the 1989 American Control Conference*, pages 1711–1718, 1989.

[15] R. Hanus, M. Kinnaert, and J. L. Henrotte. Conditioning technique, a general anti-windup and bumpless transfer method. *Automatica*, 23(6):729–739, 1987.

[16] J. Khanderia and W. L. Luyben. Experimental evaluation of digital algorithms for anti-reset windup. *Ind. & Engg. Chem., Proc. Des. and Dev.*, 15:278–285, 1976.

[17] M. Morari. Some Control Problems in the Process Industries. In H L. Trentelman and J. C. Willems, editors, *Essays on Control: Perspectives in the Theory and its Applications*, pages 55–78. Birkhäuser, Boston, 1993.

[18] M. Morari and E. Zafiriou. *Robust Process Control.* Prentice-Hall, Inc., Englewood Cliffs, N.J., 1989.

[19] K. S. Walgama, S. Rönnbäck, and J. Sternby. Generalization of conditioning technique for anti-windup compensators. *IEE Proceedings-D*, 139(2):109–118, March 1992.

[20] K. S. Walgama and J. Sternby. Inherent observer property in a class of anti-windup compensators. *International Journal of Control*, 52(3):705–724, 1990.

[21] A. Zheng, M. V. Kothare, and M. Morari. Anti-windup design for Internal Model Control. *International Journal of Control*, 1994. (in press).