Integrator Windup and How to Avoid It

Karl Johan Åström and Lars Rundqwist

Department of Automatic Control
Lund Institute of Technology, S-22100 Lund, Sweden
Phone +46 46 108781, email kja@control.lth.se

Abstract. This paper describes the phenomenon of integrator windup and various ways of avoiding it. It first covers a number of ad hoc schemes. A general procedure to avoid windup which admits a unification of the ideas is given and the results are illustrated on a number of examples.

1. Introduction

Most control systems are designed based on linear theory. This is simple and works very well for regulation around a set point because the process behavior can then be described very well by linear models. To control a system over a wide range it is, however, necessary to consider nonlinear effects. Actuator saturation, which is present in all systems, is particularly important. The effects of saturation are clearly noticeable when the controller is an unstable system. This is always the case for a controller with integral action. Since the feedback loop is broken when the actuator saturates the unstable modes of the regulator may then drift to undesirable values. The consequences are that it may take a long time for the system to reach equilibrium after an upset. The phenomenon which was first noticed in conventional PID control is therefore called *integrator windup*. It is well known to practitioners of automatic control but has not received much attention from theoreticians and it is usually neglected in standard courses on control. This paper reviews the windup phenomenon and different ways to avoid it.

2. Ad hoc methods

Let us first illustrate the phenomenon of windup. Figure 1 shows PI control of an integrator with a saturating actuator. The initial set-point change is so large that the actuator saturates at the high limit. The integral term increases initially because the error is positive and it reaches its largest value at time $t = 10$ when the error goes through zero. The output remains saturated at this point because of the large value of the integral term. It does not leave the saturation limit until the error has been negative for sufficiently long time to let the integral part come down to a small level. The net effect is a large overshoot and a damped oscillation where the control signal flips from one extreme to the other like in relay oscillation. The output finally comes so close to the set point that the actuator does not saturate. The system then behaves linearly and settles.

Incremental Algorithms

In the early phases of feedback control integral action was integrated with the actuator by having a motor drive the valve directly. In this case windup is handled automatically because integration stops when the valve stops. When controllers were implemented by analog techniques and later with computers some manufacturers used a configuration that was a direct translation of the old mechanical design. This led to the so called incremental algorithm. In this algorithm the rate of change of the control signal is first computed and then fed to an integrator. In some cases this integrator is a motor directly connected to the actuator. In other cases the integrator is implemented digitally. With this approach it is easy to handle mode changes and windup. Windup is avoided by stopping the integration whenever the output saturates. If the actuator output is not measured a model which computes the saturated output can be used. The rate of change of the control signal is also easy to limit.

Back-calculation and Tracking

The phenomenon of windup was noticed when the direct connection between the integrator and the actuator was broken. Several tricks were invented to avoid windup. They were described under labels like preloading, batch unit, etc. Although the problem was well understood there were often limits imposed due to the analog implementations. The ideas were often kept as trade secrets and not much spoken about.

The problem of windup was rediscovered when controllers were implemented digitally and several methods to avoid windup were presented in the literature. Back-

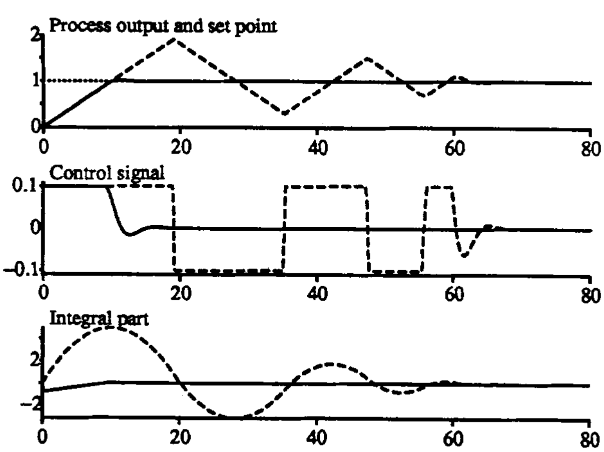

Figure 1. Illustration of integrator windup (dashed line) and control with anti-windup (solid line).

calculation, proposed by Fertik and Ross (1967), works as follows: When the output saturates the integral is recomputed so that its new value gives an output at the saturation limit. It was found advantageous not to reset the integrator in one sampling period but dynamically with a time constant T_t. Figure 2 shows a block diagram of a PID controller with such a back-calculation. The difference e_t between the output of the regulator v and the actuator output u is fed to the input of the integrator through a gain $1/T_t$. The signal e_t is zero when there is no saturation. It will thus not have any effect on the normal operation when the actuator does not saturate. When the actuator saturates the feedback signal prevents the integrator from winding up.

The rate at which the regulator output is reset is governed by the time constant T_t, which determines how quickly the integral is reset. Fertik considered a discrete system. In his scheme he used a dead-beat reset. This approach has no exact continuous time equivalent but it can be approximated by making T_t very small. It frequently happens that the actuator output cannot be measured. The anti-windup scheme just described can be applied by estimating the actuator output using a mathematical model of the saturation.

Figure 1 also shows the superior performance of the simple anti-windup scheme. The integral term has a negative value during the initial phase when the actuator is saturated. This behavior is drastically different from that of the conventional PI controller where the integral has a positive value during the initial transient. Some care must be taken when choosing time constant T_t. This is discussed further in Section 4.

A controller with back-calculation can be interpreted as having two modes, the normal *control mode* when it operates like an ordinary controller and a *tracking mode* when the integrator is tracked so that it matches given inputs and outputs. This is discussed further in Åström (1987). Since a controller with tracking can operate in two modes, we may expect that it is necessary to have a logical signal for mode switching. This is not necessary because tracking is automatically inhibited when the tracking signal is equal to the regulator output. This can be used with great advantage when building up complex systems with selectors and cascade control, see Åström (1987) and Wittenmark (1989).

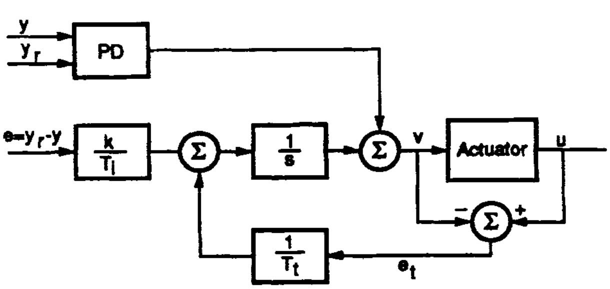

Figure 2. Regulator with anti-windup based on back-calculation.

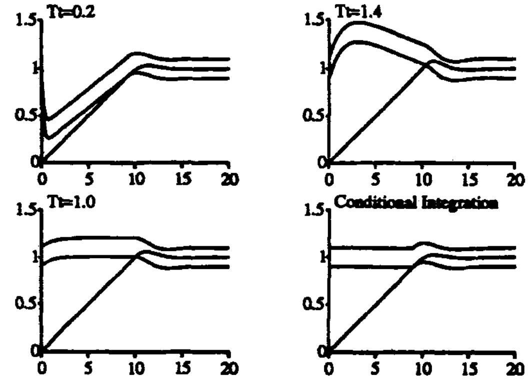

Figure 3. The proportional band for a system with conditional integration and tracking with different tracking time constants.

Conditional Integration

Conditional integration is an alternative to back calculation or tracking. The idea behind this method is to apply integral action only when certain conditions are fulfilled. The conditions are typically that the control error is small and that the actuator does not saturate. This method has been used for a long time. A recent reference is Gallun et al (1985). One possibility to implement conditional integration is to update the integral only when the predicted process output is in the proportional band. When such a condition is imposed it may also be necessary to introduce a hysteresis or some similar condition to avoid chattering.

The *proportional band* is an interval such that the actuator does not saturate if the instantaneous value of the process output or its predicted value is in the interval. For PID control the control signal is given by

$$u = k(by_r - y) + I - kT_d\frac{dy}{dt} \tag{1}$$

Solving for the predicted process output $y_p = y + T_d\frac{dy}{dt}$ we find that the proportional band is (y_l, y_h) where

$$\begin{aligned} y_l &= by_r + \frac{I - u_{max}}{k} \\ y_h &= by_r + \frac{I - u_{min}}{k} \end{aligned} \tag{2}$$

and u_{min}, u_{max} are the values of the control signal such that the actuator saturates.

The notion of proportional band also helps to understand several other schemes for anti-windup. Figure 3 shows the proportional band for the system with tracking for different values of the tracking time constant T_t and a system with conditional integration. The figure shows that in this particular case there is very little difference in performance between conditional integration and tracking. Also notice that the effect of tracking is to move the proportional band closer to the process output. There may be a disadvantage in having it too close because the system may become sensitive to occasional measurement errors.

Multivariable systems

The problem of saturated multivariable systems is treated in Kapasouris et al (1988), where a stabilizing controller is derived. Here direction preservation of the control signal is accomplished by introducing a time varying gain, which is chosen so that saturation is avoided.

Summary

We have thus given two ad hoc methods for avoiding windup, tracking or back-calculation and conditional integration. These schemes can also be combined. In Howes (1986) it is suggested to explicitly manipulate the proportional band for batch control. This is done by introducing so called *cutback points*. The high cutback is above the set point and the low cutback is below. The integrator is clamped when the predicted process output is outside the cutback interval. Integration is performed with a specified tracking time constant when the process output is between the cutback points. The cutback points are considered as controller parameters which are adjusted to influence the response to large set point changes. A similar method is proposed in Dreinhoefer (1988), where conditional integration is combined with back-calculation. In Shinskey (1967) the integrator is given a prescribed value $i = i_0$ during saturation. The value of i_0 is tuned to give satisfactory overshoot at start-up. This approach is also called preloading. Controllers with tracking are also discussed in Glattfelder and Schaufelberger (1983, 1988).

3. General Methods

Some general techniques to avoid integrator windup in a controller will now be given. They are based on ideas from state space theory.

An observer approach

In state space formulation a feedback controller is viewed as a combination of an observer and a state feedback. The dynamics in the controller is given only by the observer, see Figure 4. With such a controller it is easy to understand what goes wrong when the actuator saturates and to devise anti-windup schemes. The following method was originally given in Åström (1983). It is also described in Åström and Wittenmark (1984). Let the controller output be v and the process input be u. When the actuator saturates u is different from v. Since the controller is not aware of the saturation in the actuator it computes the state as if the process input is v. With

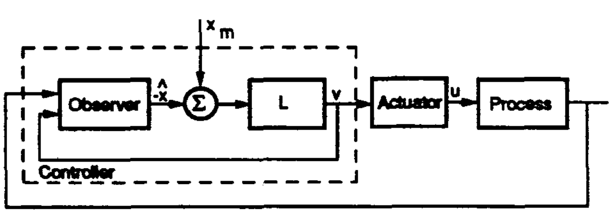

Figure 4. A process with a controller consisting of a state feedback and an observer.

this interpretation it is easy to find a remedy simply by feeding u back to the observer instead of v. If the actuator output is not measured the actuator can be modeled mathematically. The controller can be described by

$$\begin{aligned}\frac{d\hat{x}}{dt} &= A\hat{x} + Bu + K(y - C\hat{x}) \\ u &= \text{sat}\,(u_m + L(x_m - \hat{x}))\end{aligned} \tag{3}$$

where the function sat is defined as

$$\text{sat}(u) = \begin{cases} u_{low} & u \leq u_{low} \\ u & u_{low} < u < u_{high} \\ u_{high} & u \geq u_{high} \end{cases} \tag{4}$$

for a scalar and

$$\text{sat}(u) = \begin{pmatrix} \text{sat}\, u_1 \\ \text{sat}\, u_2 \\ \vdots \\ \text{sat}\, u_n \end{pmatrix} \tag{5}$$

for a vector. The values u_{low} and u_{high} are chosen to correspond to the actuator limitations.

Notice that the dynamics of the controller is given by

$$\det(sI - A + BL + KC) = 0 \tag{6}$$

when the actuator does not saturate and by

$$\det(sI - A + KC) = 0 \tag{7}$$

when the actuator saturates. The dynamics of the controller during saturation is thus given by the observer dynamics.

The idea based on the observer interpretation can be applied to MIMO as well as SISO systems. For multivariable systems there are several ways to saturate the control signal. For systems having strong coupling it may be useful to saturate the signal so that the direction of the control signal is preserved.

Notice that the controller (3) has high frequency roll-off. The technique can be extended to controllers where the high frequency gain of the controller is constant. Such a controller can be described by

$$\begin{aligned}\frac{dx}{dt} &= Fx + G_r y_r - G_y y \\ u &= Hx + D_r y_r - D_y y\end{aligned} \tag{8}$$

To obtain anti-windup the control algorithm is rewritten so that the control signal appears explicitly, and a feedback from the difference between desired control signal v and the saturated control signal $u = \text{sat}(v)$ is introduced. The following controller is then obtained.

$$\begin{aligned}\frac{dx}{dt} &= Fx + G_r y_r - G_y y + M(u - v) = (F - MH)x \\ &\quad + (G_r - MD_r)y_r - (G_y - MD_y)y + Mu \\ v &= Hx + D_r y_r - D_y y \\ u &= \text{sat}(v)\end{aligned} \tag{9}$$

where $F - MH$, which corresponds to the observer dynamics, has stable eigenvalues.

Conditioning

A technique called conditioning was proposed by Hanus. It is described in Hanus et al (1987). The idea is to compute the reference signal that would just saturate the output. For a controller given in state space form (Eqn (8)) conditioning is equivalent to a special case of the observer approach, with $M = G_r D_r^{-1}$ in Eqn (9). Hence

$$\begin{aligned} \frac{dx}{dt} &= (F - G_r D_r^{-1} H)x - (G_y - G_r D_r^{-1} D_y)y \\ &\quad + G_r D_r^{-1} u \\ u &= \mathrm{sat}(Hx + D_r y_r - D_y y) \end{aligned} \tag{10}$$

The reference signal y_r does not influence the state during saturation. A necessary condition is that the controller has a direct term from the reference signal y_r, i.e. $D_r \neq 0$. The eigenvalues of $F - G_r D_r^{-1} H$ are equal to the transmission zeros of the controller.

Analysis

A system with a saturating actuator can always be reduced to a standard configuration with a linear system having a nonlinear feedback. Approximate analysis of such systems can be done using describing function theory (Atherton, 1975). Sufficient conditions for stability can also be obtained from the circle criterion (Desoer and Vidyasagar, 1975). Global stability is too strict a requirement if the plant is unstable, because an unstable plant can never be stabilised globally when the control signal is bounded.

Consider a system described by

$$\begin{aligned} \frac{dx}{dt} &= Ax + Bu \\ y &= Cx \end{aligned} \tag{11}$$

with the controller

$$\begin{aligned} u &= L(x_m - \hat{x}) \\ \frac{d\hat{x}}{dt} &= A\hat{x} + Bu + K(y - C\hat{x}) \end{aligned} \tag{12}$$

The transfer function of the process is

$$G_p = C(sI - A)^{-1} B \tag{13}$$

and the controller has the transfer function

$$G_c = L(sI - A - BL - KC)^{-1} K \tag{14}$$

Assume that the system has one control variable and one measured and consider the consequences of a saturation of the actuator. The closed loop system is then in the standard configuration with a linear block having the transfer function

$$G(s) = G_p(s) G_c(s) \tag{15}$$

and a nonlinearity $u = \mathrm{sat}(y)$ in the feedback loop. Since the describing function of a saturation is the line segment $(-\infty, -1)$, describing function analysis indicates a limit cycle if the Nyquist curve of $G_p G_c$ intersects this line segment. Furthermore it follows from the circle criterion that the closed loop system is stable provided that

$$\mathrm{Re}\, G_p(i\omega) G_c(i\omega) + 1 > 0 \tag{16}$$

Controller with Anti-Windup

Now consider the consequences of using a controller with anti-windup. It follows from (11) and (12) that

$$\frac{d\tilde{x}}{dt} = (A - KC)\,\tilde{x} \tag{17}$$

where $\tilde{x} = x - \hat{x}$. The estimation error $\tilde{x}$ thus goes to zero exponentially if $A - KC$ is stable. When $\tilde{x} \to 0$, it follows from (17) that the closed loop system is described by

$$\frac{dx}{dt} = Ax + B\,\mathrm{sat}\, L(x_m - x) \tag{18}$$

This is a linear system with transfer function

$$G_w(s) = L(sI - A)^{-1} B \tag{19}$$

with a saturation in the feedback. Describing function theory then indicates a limit cycle if the Nyquist curve of $G_w(s)$ intersects the segment $(-\infty, -1)$ and it follows from hyperstability theory that the closed loop is stable provided that

$$\mathrm{Re}\, G_w(i\omega) + 1 > 0 \tag{20}$$

We formulate the result as

THEOREM 1
A controller with windup given by (12) gives a stable closed loop system when controlling the process (11) with saturation provided that

$$\mathrm{Re}\, L(sI - A)^{-1} B + 1 > 0 \tag{21}$$

for $s = i\omega$. □

When the control system design is performed it is thus straight forward to investigate if the controller has windup. Notice that the condition (21) does not depend on the observer. Also notice that if L is computed from linear quadratic theory it follows that

$$|1 + G_w(i\omega)| \geq 1 \tag{22}$$

4. Example

Properties of different methods for avoiding windup will now be demonstrated on an example. The process consists of two identical cascaded tanks, used for basic experiments with automatic control, see Åström and Östberg (1986). The control signal is pump speed, which determines the influent flow rate to the upper tank, and the process output is the level of the lower tank. The level

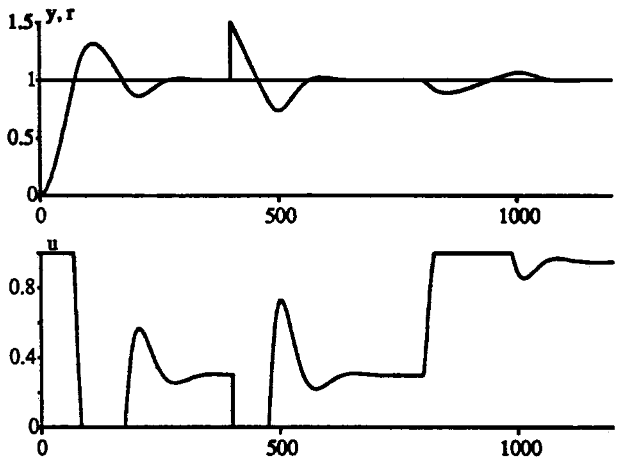

Figure 5. The standard experiment, with delayed disturbances, gives large overshoots due to integrator windup.

control is conducted by a PID controller. The curves given in the paper are results of simulations. The corresponding behavior has however been observed in the laboratory.

A linearized state space model for the double tank is

$$\begin{aligned} \frac{dx}{dt} &= \begin{pmatrix} -\alpha & 0 \\ \alpha & -\alpha \end{pmatrix} x + \begin{pmatrix} \beta \\ 0 \end{pmatrix} u = A_p x + B_p u \\ y &= \begin{pmatrix} 0 & 1 \end{pmatrix} x = C_p x \end{aligned} \tag{23}$$

where x_1 and x_2 are upper and lower tank levels respectively, u is control input and y is measurement, all dimensionless. Parameters $\alpha = 0.015$ (s^{-1}) and $\beta = 0.05$ (s^{-1}). The process input u is restricted to the interval $[0, 1]$. The stationary gain is $\beta/\alpha \approx 3.33$.

The process is controlled by a continuous PID controller with a filtered derivative and a proportional part that only acts on a fraction b of the reference signal y_r (Åström and Hägglund (1988)). The controller parameters are $K = 5$, $T_i = 40$ s, $T_d = 15$ s, $N = 5$ and $b = 0.3$, which gives an overshoot of 10 % and a natural frequency ≈ 0.05 rad/s for the three dominating closed loop poles.

Experiments

A standard experiment is used to test the anti-windup methods. The process and controller starts at stationarity with all signals zero. The experiment is as follows:

1. Start-up: at time $t = 0$ the reference $y_r = 1$.
2. Impulse disturbance: at time $t = 250$ state x_2 is changed to 1.5. This corresponds to quickly pouring a cup of water into the lower tank.
3. Load disturbance: at time $t = 500$ a load on the upper tank is introduced.

The three parts of the experiment have been designed to saturate the control signal in one or both directions. Figure 5 shows the consequences of windup when the control signal is restricted to the interval $[0, 1]$ without use of anti-windup.

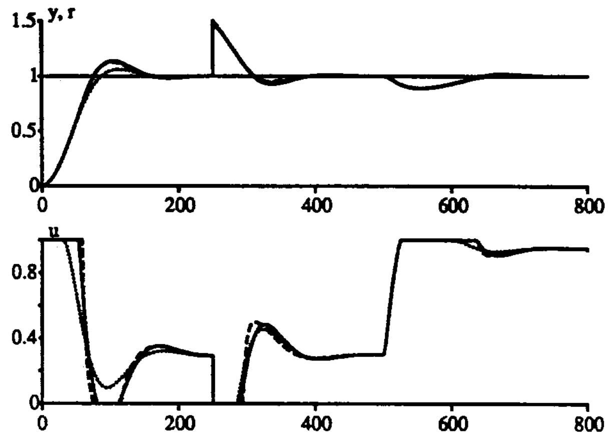

Figure 6. Tracking (solid), the observer approach (dashed) and conditional integration (dotted), with well chosen parameters, are tested on the standard experiment.

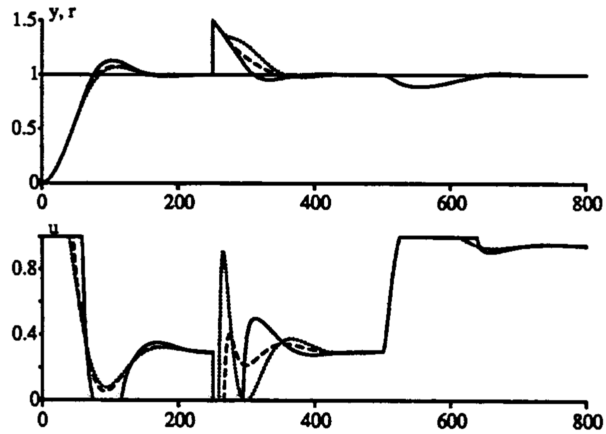

Figure 7. Illustration of the effects of different observer bandwidths ω for anti-windup based on the observer approach. The bandwidths are $\omega = 0.05$ rad/s (solid), $\omega = 0.10$ rad/s (dashed) and $\omega = 0.15$ (dotted).

Anti-windup methods

In this Section some of the described anti-windup methods are compared, namely

- Conditional integration, with integration suspended during saturation.
- Tracking, with $T_t = T_i = 40$ s.
- The observer approach (Eqn (9)), with natural frequency $\omega_0 = 0.05$ rad/s and relative damping $\zeta = 1$.

These methods, see Figure 6, give similar results. Conditional integration gives the same overshoot as tracking with small T_t. The methods are almost identical for load and impulse disturbances.

In Figure 7 the observer approach is compared for three choices of ω, with $\zeta = 1$. $\omega = 0.05$ rad/s is approximately the natural frequency for the three dominating closed loop poles. The two faster observers ($\omega = 0.10$ rad/s and $\omega = 0.15$ rad/s) give less overshoot but the response from the impulse disturbance is very poor. Tracking with $T_t < T_i$ gives similar results.

A number of simulations have been done. Some observations are that the impulse disturbance is the most critical one, that tuning the anti-windup for a good set point response may give a very bad response for an impulse disturbance, and that if the response of an impulse disturbance is good then the other responses are good.

For tracking, $T_t = T_i$ or slightly less ($0.8T_i$), gives good response from the impulse disturbance. Similarly, the natural frequency of the observer should be slightly larger than the natural frequency of the closed loop. Higher natural frequency or shorter T_t deteriorates the response of the impulse disturbance. The results from the simulations agree well with the experimental results.

5. Conclusions

A number of techniques for avoiding integrator windup have been investigated. Conditional integration is easy to apply to most controllers. The key difficulty is to find appropriate conditions for switching off the integration and to avoid chattering. Tracking is another good method for avoiding windup. This method requires one parameter, the tracking time constant, to be chosen but there is no risk of chattering. Tracking is convenient to use for systems with selectors and cascade control. Selection of the parameters in the anti-windup schemes are important.The tuning is different for disturbances and set point changes. The observer approach is a good general technique. It unifies many approaches and can be extended to many controllers of different types. An advantage of the observer approach is that with a complete linear design no additional parameters have to be chosen. The method also applies directly to multivariable systems.

It should also be pointed out that control design can be formulated as an optimisation problem. LQG is a popular approach which leads to an unconstrained quadratic optimization problem. A natural extension is to include actuator saturation. This leads to a quadratic programming problem. There are very good numerical algorithms available for solving such problems. This is incorporated in quadratic dynamic matrix control (QDMC), see Garcia and Morshedi (1986), which gives a nice solution to the problem of windup at the cost of increased computations.

6. References

Atherton, D. P. (1975): *Nonlinear Control Engineering*, van Nostrand Reinhold, London.

Åström, K. J. (1983): "Practical aspects on digital implementation of control laws," in AGARD Lecture Series No. 128: *Computer-Aided Design and Analysis of Digital Guidance and Control Systems*, AGARD, Neuilly sur Seine, France.

Åström, K. J. (1987): "Advanced control methods – Survey and assessment of possibilities," in H. M. Morris, E. J. Kompass and T. J. Williams (Eds.): *Advanced Control in Computer integrated Manufacturing. Proc. 13th Annual Advanced Control Conference Purdue University.*, Purdue Research Foundation. West Lafayette, Indiana.

Åström, K. J., and T. Hägglund (1988): *Automatic Tuning of PID Regulators*, Instrument Society of America, Triangle Research Park, NC.

Åström, K. J. and B. Wittenmark (1984): *Computer-Controlled Systems Theory and Design*, Prentice-Hall, Englewood Cliffs, NJ.

Åström, K. J. and A.-B. Östberg (1986): "A modern teaching laboratory for process control," *IEEE Contr. Sys. Mag.*, **6**, No. 5, 37–42.

Desoer, C. A. and M. Vidyasagar (1975): *Feedback Systems: Input-Output Properties*, Academic Press, New York.

Dreinhoefer, L. H. (1988): "Controller tuning for a slow nonlinear process," *IEEE Contr. Sys. Mag.*, **8**, No. 2, 56–60.

Fertik, H. A and C. W. Ross (1967): "Direct digital control algorithms with anti-windup feature," *ISA Trans.*, **6**, No. 4, 317–328.

Gallun, S. E., C. W. Matthews, C. P. Senyard and B. Slater (1985): "Windup protection and initialization for advanced digital control," *Hydrocarbon Processing*, June, 63–68.

Garcia, C. E. and A. M. Morshedi (1986): "Quadratic Programming Solution of Dynamic Matrix Control (QDMC)," *Chemical Engineering Communications*, **46**, 73–87.

Glattfelder, A. H. and W. Schaufelberger (1983): "Stability Analysis of Single Loop Systems with Saturation and Antireset-Windup Circuits," *IEEE Trans. Aut. Contr.*, **AC-28**, 1074–1081.

Glattfelder, A. H. and W. Schaufelberger (1988): "Stability of Discrete Override and Cascade-Limiter Single-Loop Control Systems," *IEEE Trans. Aut. Contr.*, **AC-33**, 532–540.

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

Howes, G. (1986): "Control of overshoot in plastics-extruder barrel zones," *EI Technology*, No. 3, Eurotherm International, Brighton, England, 16–17.

Kapasouris, P., M. Athans and G. Stein (1988): "Design of Feedback Control Systems for Stable Plants with Saturating Actuators," *Preprints for 27th IEEE Conf. on Dec. and Contr.*, Austin, Texas.

Shinskey, F. G. (1988): *Process-Control Systems (3:rd ed.)*, McGraw-Hill, New York.

Wittenmark, B. (1989): "Integrators, Nonlinearities and Anti-reset Windup for Different Control Structures," *Proc. American Contr. Conf.*, Pittsburgh, PA.