

From PID to Active Disturbance Rejection Control

Jingqing Han

Abstract—Active disturbance rejection control (ADRC) can be summarized as follows: it inherits from proportional–integral–derivative (PID) the quality that makes it such a success: the error driven, rather than model-based, control law; it takes from modern control theory its best offering: the state observer; it embraces the power of nonlinear feedback and puts it to full use; it is a useful digital control technology developed out of an experimental platform rooted in computer simulations. ADRC is made possible only when control is taken as an experimental science, instead of a mathematical one. It is motivated by the ever increasing demands from industry that requires the control technology to move beyond PID, which has dominated the practice for over 80 years. Specifically, there are four areas of weakness in PID that we strive to address: 1) the error computation; 2) noise degradation in the derivative control; 3) oversimplification and the loss of performance in the control law in the form of a linear weighted sum; and 4) complications brought by the integral control. Correspondingly, we propose four distinct measures: 1) a simple differential equation as a transient trajectory generator; 2) a noise-tolerant tracking differentiator; 3) the nonlinear control laws; and 4) the concept and method of total disturbance estimation and rejection. Together, they form a new set of tools and a new way of control design. Times and again in experiments and on factory floors, ADRC proves to be a capable replacement of PID with unmistakable advantage in performance and practicality, providing solutions to pressing engineering problems of today. With the new outlook and possibilities that ADRC represents, we further believe that control engineering may very well break the hold of classical PID and enter a new era, an era that brings back the spirit of innovations.

Index Terms—Active disturbance rejection control (ADRC), extended state observer (ESO), nonlinear proportional–integral–derivative (PID), tracking differentiator.

I. INTRODUCTION

THE BIRTH and large-scale deployments of the powerful yet primitive proportional–integral–derivative (PID) control law dates back to the period of the 1920s–1940s in the last century, in response to the pressing demands of industrial automation before, during, and particularly after World War II. Its role in the explosive growth in the postwar manufacturing industry is unmistakable; its dominance is evident even today across various sectors of the entire industry. It is at the same time undeniable that PID is increasingly overwhelmed by the new demands in this era of modern industries where the unending pursuit of efficiency and the lack and cost of skilled labor put a high premium on feedback control technologies. Its merit of simplicity in the analog electronics era has turned

into a liability in the digital control one as it cannot fully take advantage of the new compact and powerful digital processors. It appears that, as any technology, PID will eventually outlive its usefulness, if it has not already done so. The question is, what will replace this hugely successful control mechanism in the 21st century, retaining its basic soundness and, at the same time, shedding its limitations? It is doubtful that such question was even entertained systematically, let alone answered, in the past.

We believe that the answer lies in our understanding of both the characteristics of PID and the challenges it faces. It is such understanding that will lead us to propose further developments in the PID framework and, perhaps, even a drastic innovation toward a new generation of digital control solutions. In this paper, we suggest that there are four fundamental technical limitations in the existing PID framework, and we proceed to propose the corresponding technical and conceptual solutions, including the following: 1) a simple differential equation to be used as a transient profile generator; 2) a noise-tolerant tracking differentiator; 3) the power of nonlinear control feedback; and 4) the total disturbance estimation and rejection. Together, these new tools combine to form the backbone of a new synthesis of digital control law that is not predicated on an accurate and detailed dynamic model of the plant and is extremely tolerant of uncertainties and simple to use. Moreover, we denote this new synthesis active disturbance rejection control or ADRC.

ADRC has been a work in progress for almost two decades [1]–[7], with its ideas and applications appearing in the English literature, amid some questions and confusions, sporadically only in recent years; see, for example, [8]–[14]. In ADRC, we see a paradigmatic change in feedback control that was first systematically introduced in English in 2001 [8]. The conception of active disturbance rejection was further elaborated in [9]. However, even though much success has been achieved in practical applications of ADRC, it appears that this new paradigm has not been well understood and there is a need for a paper that provides a full account of ADRC to the English audience [13]. Such need is unmistakable in the recently proposed terminologies such as equivalent input disturbance [14] and disturbance input decoupling [15], all of which can be seen as a special case of ADRC where only the external disturbance was considered. It is primarily for this reason that this paper is written.

In this paper, we start with, in Section II, classical PID, a dominant technology in industry today, and discuss its characteristics and weaknesses, followed by the proposed remedies in Section III and the resulting ADRC control scheme in Section IV. How this new framework is used to solve various kinds of control problems is shown in Section V. To further help users master this new methodology, some key points in the application of ADRC are presented in Section VI, followed by the concluding remarks in Section VII.

Manuscript received March 20, 2008; revised September 23, 2008. Current version published February 27, 2009.

J. Han, deceased, was with the Institute of Systems Science, Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.

Digital Object Identifier 10.1109/TIE.2008.2011621

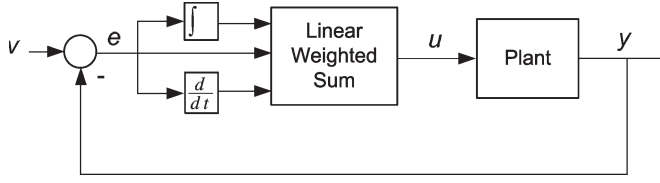


Fig. 1. PID control topology.

II. CLASSICAL PID

Classical PID is a particular rather primitive and simplified implementation of the basic principle in error-based feedback control, as shown in Fig. 1, where the error between the setpoint $v = \text{const}$ and plant y , i.e., $e = v - y$, as well as its differentiation de/dt and integration $\int_0^t e d\tau$ are used in a linear combination to produce the control law

$$u = k_0 \int_0^t e d\tau + k_1 e + k_2 \frac{de}{dt}. \quad (1)$$

This is where PID takes its name: proportional–integral–derivative. It found widespread applications in industry, even when there is little or no information available regarding plant dynamics. To help users understand this, we use the following second-order system equation:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = a_1 x_1 + a_2 x_2 + bu \\ y = x_1 \end{cases} \quad (2)$$

commonly found in practice, such as motion control, to illustrate why is it that a PID can be easily configured and tuned to do its job.

Let $e = v - y = v - x_1 = e_1$, $\dot{e}_1 = -\dot{x}_1 = e_2$, and $\ddot{e} = -\ddot{x}_1$, and the error dynamics can be seen as

$$\begin{cases} \dot{e}_1 = e_2 \\ \dot{e}_2 = a_1 e_1 + a_2 e_2 - a_1 v - bu. \end{cases} \quad (3)$$

Denoting $e_0 = \int_0^t e d\tau$, then $\dot{e}_0 = e = e_1$, and (1) becomes

$$u = k_0 e_0 + k_1 e_1 + k_2 e_2. \quad (4)$$

Together with (3), the error equation can be rewritten as

$$\begin{cases} \dot{e}_0 = e_1 \\ \dot{e}_1 = e_2 \\ \dot{e}_2 = -bk_0 \left(e_0 + \frac{a_1}{bk_0} v \right) + (a_1 - bk_1)e_1 + (a_2 - bk_2)e_2 \end{cases} \quad (5)$$

which is asymptotically stable if

$$\begin{cases} bk_0 > 0, (bk_1 - a_1) > 0, (bk_2 - a_2) > 0 \\ (bk_1 - a_1)(bk_2 - a_2) > bk_0. \end{cases} \quad (6)$$

That is, the design objective $e_1 = e = v - x_1 \rightarrow 0$, or $x_1 \rightarrow v$, is met if the gains k_0 , k_1 , and k_2 are selected to satisfy (6), for the given range of a_1 , a_2 , $b \neq 0$. It is rather obvious that, for most of the plants of the form of (2), a set of PID gains can be easily found, analytically when the model is given or by trial and error when it is not. Although such simplicity and the ease of tuning could very well be behind the popularity of PID, they also mark its fundamental limitations, which are made glaringly

obvious in the presence of ever more demanding control system performance.

To be more specific, we believe that there are four fundamental issues to be addressed in the PID framework. 1) Setpoint is often given as a step function, not appropriate for most dynamics systems because it amounts to asking the output and, therefore, the control signal, to make a sudden jump. 2) PID is often implemented without the D part because of the noise sensitivity. 3) The weight sum of the three terms in (4), while simple, may not be the best control law based on the current and the past of the error and its rate of change. 4) The integral term, while critical to rid of steady-state error, introduces other problems such as saturation and reduced stability margin due to phase lag. It is these fundamental limitations of PID that prompt us to offer the following remedies.

III. INDIVIDUAL REMEDIES

We now offer the following practical solutions to address the aforementioned issues. They are easy to implement and understand, without introducing undue complexity to the existing controller.

A. Setpoint Jump

To avoid setpoint jump, it is necessary to construct a transient profile that the output of the plant can reasonably follow. While this need is mostly ignored in a typical control textbook, engineers have devised different motion profiles in servo systems. A simple easy-to-use solution is offered in this paper.

It is well known that, for a double integral plant

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u \end{cases} \quad (7)$$

with $|u| \leq r$ and v is the desired value for x_1 , the time-optimal solution is

$$u = -r \operatorname{sign} \left(x_1 - v + \frac{x_2 |x_2|}{2r} \right). \quad (8)$$

Using this principle, a desired transient profile is obtained by solving the following differential equation:

$$\begin{cases} \dot{v}_1 = v_2 \\ \dot{v}_2 = -r \operatorname{sign} \left(v_1 - v + \frac{v_2 |v_2|}{2r} \right) \end{cases} \quad (9)$$

where v_1 is the desired trajectory and v_2 is its derivative. Note that, depending on the physical limitations in each application, the parameter r can be selected accordingly to speed up or slow down the transient profile. In addition, it is well known that this continuous-time time-optimal solution in (8) could introduce considerable numerical errors in a discrete-time implementation. To address this difficulty, a discrete-time solution for a discrete double integral plant

$$\begin{cases} v_1 = v_1 + hv_2 \\ v_2 = v_2 + hu, \quad |u| \leq r \end{cases} \quad (10)$$

was obtained as

$$u = f_{han}(v_1 - v, v_2, r_0, h_0) \quad (11)$$

where h is the sampling period, r_0 and h_0 are controllers parameters, and $fhan(v_1, v_2, r_0, h_0)$ is

$$\begin{aligned} d &= h_0^2 r_0^2, & a_0 &= h_0 v_2, & y &= v_1 + a_0 \\ a_1 &= \sqrt{d(d + 8|y|)} \\ a_2 &= a_0 + \text{sign}(y)(a_1 - d)/2 \\ s_y &= (\text{sign}(y + d) - \text{sign}(y - d))/2 \\ a &= (a_0 + y - a_2)s_y + a_2 \\ s_a &= (\text{sign}(a + d) - \text{sign}(a - d))/2 \\ fhan &= -r_0 \left(\frac{a}{d} - \text{sign}(a) \right) s_a - r_0 \text{sign}(a). \end{aligned}$$

Note that (11) is a time-optimal solution that guarantees the fastest convergence from v_1 to v without any overshoot and r_0 and h_0 are set equal to r and h , respectively. However, when (10) and (11) are used for the purpose of defining a transient profile, r_0 and h_0 can be adjusted individually according to the desired speed and smoothness.

B. Tracking Differentiator

It is common in PID implementation that a differentiation of a signal v is obtained approximately as

$$y = \frac{s}{\tau s + 1} v \quad (12)$$

which can be rewritten as

$$y = \frac{1}{\tau} \left(1 - \frac{1}{\tau s + 1} \right) v$$

or in the time domain as

$$y(t) = \frac{1}{\tau} (v(t) - v(t - \tau)) \approx \dot{v}(t). \quad (13)$$

Such approximation is still quite sensitive to the noise in v because it is amplified by a factor of $1/\tau$. That is, if $v(t)$ contains noise $n(t)$, then $\dot{v}(t)$ contains $n(t)/\tau$ in its first term in (13). We therefore conclude that (12) is not a good way of approximating $\dot{v}(t)$. Instead, we proposed the following approximation:

$$\dot{v}(t) \approx \frac{v(t - \tau_1) - v(t - \tau_2)}{\tau_2 - \tau_1} \quad (14)$$

which can be implemented approximately using the second-order transfer function

$$w_1(s) = \frac{1}{\tau_2 - \tau_1} \left(\frac{1}{\tau_1 s + 1} - \frac{1}{\tau_2 s + 1} \right), \quad \tau_2 > \tau_1 > 0. \quad (15)$$

Moreover, as verified in simulations, this resolves the aforementioned problem of noise amplification. A particular second approximation of a differentiator is $s/(\tau s + 1)^2$, which corresponds to the differential equation with $r = 1/\tau$

$$\ddot{y} = -r^2 (y - v(t)) - 2r\dot{y} \quad (16)$$

where $y(t)$ tracks $v(t)$, $\dot{y}(t)$ tracks $\dot{v}(t)$, and r determines the speed. It is the desire to find a differentiator that yields the

fastest tracking that leads us to the “tracking differentiator” as shown in the following.

Consider $v(t)$ as the input signal to be differentiated; (9), rewritten here as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = -r \text{sign}\left(x_1 - v(t) + \frac{x_2|x_2|}{2r}\right) \end{cases} \quad (17)$$

provides the fastest tracking of $v(t)$ and its derivative subject to the acceleration limit of r . It is in this sense that (17) is denoted as the “tracking differentiator” of $v(t)$. In practical implementations, however, we again advise that its discrete version in (10) and (11) be used to avoid unnecessary oscillations, where r_0 and h_0 are adjusted accordingly as filter coefficients.

C. Nonlinear Feedback Combination

PID, as a control law, employs a linear combination of present, accumulative, and predictive forms of the tracking error and has, for a long time, ignored other possibilities of this combination that are potentially much more effective. As an alternative, we propose the following nonlinear function:

$$\begin{aligned} fal(e, \alpha, \delta) &= \begin{cases} \frac{e}{\delta^{1-\alpha}}, & |e| \leq \delta \\ |e|^\alpha \text{sign}(e), & |e| \geq \delta \end{cases} \\ fhan(x_1, x_2, r, h_0) &= \frac{e}{\delta^{1-\alpha}} + |e|^\alpha \text{sign}(e)(1-s) \end{aligned} \quad (18)$$

$s = \frac{\text{sign}(e + \delta) - \text{sign}(e - \delta)}{2}$

that sometimes provides surprisingly better results in practice. For example, with linear feedback, the tracking error, at best, approaches zero in infinite time; with nonlinear feedback of the form

$$u = |e|^\alpha \text{sign}(e)$$

the error can reach zero much more quickly in finite time, with $\alpha < 1$. Such α can also help reduce steady state error significantly, to the extent that an integral control, together with its downfalls, can be avoided. An extreme case is $\alpha = 0$, i.e., bang-bang control that can bring with it zero steady state error without the I term in PID. It is because of such efficacy and unique characteristics of nonlinear feedback that we propose a systematic and experimental investigation. Such nonlinear feedback functions in the forms of fal and $fhan$ play an important role in the newly proposed control framework, ADRC, as will be presented later in this paper.

D. Total Disturbance Estimation and Rejection via ESO

In this section, we introduce a new concept: total disturbance and its estimation and rejection. Although such concept is, in general, applicable to most nonlinear multi-input–multioutput (MIMO) time varying systems, we use a second-order single-input–single-output (SISO) example for the sake of simplicity and clarity. Consider

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x_1, x_2, w(t), t) + bu \\ y = x_1 \end{cases} \quad (19)$$

where y is the output, measured and to be controlled, u is the input, and $f(x_1, x_2, w(t), t)$ is a multivariable function of both the states and external disturbances, as well as time.

The objective here is to make y behave as desired using u as the manipulative variable, and for this purpose, unlike any mathematical analysis of (19), $f(x_1, x_2, w(t), t)$ does not need to be expressively known. In fact, in the context of feedback control, $F(t) = f(x_1(t), x_2(t), w(t), t)$ is something to be overcome by the control signal, and it is therefore denoted as the “total disturbance.” At this point, we have transformed a problem that traditionally belongs to system identification to that of disturbance rejection, and its consequence is enormous.

Treating $F(t)$ as an additional state variable, $x_3 = F(t)$, and let $\dot{F}(t) = G(t)$, with $G(t)$ unknown, the original plant in (19) is now described as

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = x_3 + bu \\ \dot{x}_3 = G(t) \\ y = x_1 \end{cases} \quad (20)$$

which is always observable. Now, we construct a state observer, denoted as the extended state observer (ESO), in the form of

$$\begin{cases} e = z_1 - y \\ fe = fal(e, 0.5, \delta), \quad fe_1 = fal(e, 0.25, \delta) \\ \dot{z}_1 = z_2 - \beta_{01}e \\ \dot{z}_2 = z_3 + bu - \beta_{02}fe \\ \dot{z}_3 = -\beta_{03}fe_1 \end{cases} \quad (21)$$

which is implemented in discrete form, with a sampling period of h , as

$$\begin{cases} e = z_1 - y \\ fe = fal(e, 0.5, \delta), \quad fe_1 = fal(e, 0.25, \delta) \\ z_1 = z_1 + h(z_2 - \beta_{01}e) \\ z_2 = z_2 + h(z_3 + bu - \beta_{02}fe) \\ z_3 = z_3 + h(-\beta_{03}fe_1) \end{cases} \quad (22)$$

There are many ways to select the observer gains β_{01} , β_{02} , and β_{03} for a particular problem. For example, the observer gains in (22) can be selected as

$$\beta_{01} = 1 \quad \beta_{02} = \frac{1}{2h^{0.5}} \quad \beta_{03} = \frac{2}{5^2 h^{1.2}} \quad (23)$$

for $F(t) = \gamma \text{sign}(\sin((0.001/h)t))$. Simulation result shows that the observer performs very well for a wide range of h , from 0.0001 to 1000.

For both continuous- and discrete-time forms of ESO in (21) and (22), the observer gains can be made linear for the sake of simplicity in implementation, replacing both fe and fe_1 with e . In this case, the corresponding gains are

$$\beta_{01} = 1 \quad \beta_{02} = \frac{1}{3h} \quad \beta_{03} = \frac{2}{8^2 h^2}. \quad (24)$$

Also, note that the inputs to ESO are the system output y and the control signal u , and the output of the ESO provides the important information $F(t) = f(x_1(t), x_2(t), w(t), t)$. This

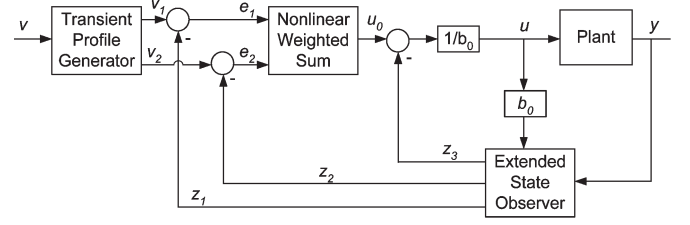


Fig. 2. ADRC topology.

allows the control law $(u_0 - F(t))/b$ to reduce the plant in (19) to a cascade integral form of

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u_0 \\ y = x_1 \end{cases} \quad (25)$$

which can be easily controlled by making u_0 a function of the tracking error and its derivative, i.e., a PD controller. That is, the control problem is transformed to that of estimation and rejection of the total disturbance, and it is greatly simplified. Numerous successful applications of this principle have demonstrated the validity of such an approach.

IV. PUTTING IT ALL TOGETHER: ADRC

Combining the transient profile generation, the nonlinear feedback combination, and the total disturbance estimation and rejection, the ADRC takes the form as shown in Fig. 2, with the corresponding control algorithm in (26). Moreover, the observer gains are given in (23), leaving only four tuning parameters; among which, r is the amplification coefficient that corresponds to the limit of acceleration, c is a damping coefficient to be adjusted in the neighborhood of unity, h_1 is the precision coefficient that determines the aggressiveness of the control loop and it is usually a multiple of the sampling period h by a factor of at least four, and b_0 is a rough approximation of the coefficient b in the plant within a $\pm 50\%$ range. Therefore, the main tuning parameter that caters the controller to a particular application is h_1 , with c functioning as a fine-tuning adjustment

$$\begin{cases} fv = fhan(v_1 - v, v_2, r_0, h) \\ v_1 = v_1 + hv_2 \\ v_2 = v_2 + hf v \\ e = z_1 - y \\ fe = fal(e, 0.5, h), \quad fe_1 = fal(e, 0.25, h) \\ z_1 = z_1 + h(z_2 - \beta_{01}e) \\ z_2 = z_2 + h(z_3 + b_0 u - \beta_{02}fe) \\ z_3 = z_3 + h(-\beta_{03}fe_1) \\ e_1 = v_1 - z_1 \quad e_2 = v_2 - z_2 \\ u = -\frac{fhan(e_1, e_2, r, h_1) + z_3}{b_0} \end{cases} \quad (26)$$

The last equation in (26) can be, of course, chosen differently as a control law. Two alternatives are given as

$$\begin{cases} u = \frac{\beta_{01}e_1 + \beta_{02}e_2 - z_3}{b_0} \\ u = \frac{\beta_{01}fal(e_1, \alpha_1, \delta) + \beta_{02}fal(e_2, \alpha_2, \delta) - z_3}{b_0}, \quad 0 < \alpha_1 < 1 < \alpha_2 \end{cases} \quad (27)$$

which amount to a linear and a nonlinear PD controller, respectively. In fact, by using different linear or nonlinear gain combinations in the ESO and the feedback, one can easily find over 100 different controllers in the same ADRC structure. Regardless of which one of these control laws is chosen, we would like to point out that the controller coefficients are not dependent on the mathematical model of the plant, thus making ADRC largely model independent. These coefficients are primarily functions of the “time scale,” i.e., how fast the plant changes. That is, the controller only needs to act as fast as the plant can react, and in the previously given formulation, this is implicitly represented by the choice of the sampling period h .

V. APPLICATION OF ADRC

From the aforementioned illustration, it is apparent that ADRC has a much wider application range than PID. Even though ADRC was presented for a relatively simple second-order plant, the very same idea can be applied to solve problems of much different nature and complexity. In this section, we give a few cases that are particularly challenging to PID and demonstrate the creative process in the applications of ADRC.

A. Time Delay

Plants with time delays, such as

$$y = G(s)e^{-\tau s}u \quad (28)$$

pose difficult challenges to all control design methods. Here, we suggest the following three methods.

1) *Transfer Function Approximation*: There are various methods of approximating the term $e^{-\tau s}$ by using linear transfer functions, for example, the Pade Approximation. However, since ADRC is not predicated on an accurate model of the plant, a simple approximate of $e^{-\tau s} \approx 1/(\tau s + 1)$ may prove to be good enough. That is, (28) is approximated as

$$y = G(s) \frac{1}{\tau s + 1} u. \quad (29)$$

2) *Predictive Output Feedback*: From (28), let

$$y_0 = e^{\tau s} y = G(s)u \quad (30)$$

and design the control law based on the delayless $y_0 - u$ dynamics. Then, implement the control law using a predictive output feedback $e^{\tau s} y$, which is essentially the idea of Smith Predictor.

3) *Predictive Pseudo Input Method*: Let $U = e^{-\tau s} u$ be the pseudo control signal; then, (28) becomes

$$Y = G(s)U. \quad (31)$$

Once U , as a control law, is determined, the real control signal can be found as $u = e^{\tau s} U$. The prediction in both parts 2) and 3) can be readily approximated using the tracking differentiator discussed previously.

B. Multivariable Decoupling Control

Consider the MIMO plant

$$\begin{cases} \dot{x} = F(x, t) + Bu \\ y = Cx \end{cases} \quad (32)$$

where $D = CB$ is a nonsingular $m \times m$ matrix. To formulate the decoupling problem as an ADRC problem, we rewrite (32) as

$$\dot{y} = G + U \quad (33)$$

with the total disturbance $G = CF(x, t)$ and the pseudo control variable $U = Du$. The MIMO plant of (33) is completely decoupled. A SISO ADRC controller is then designed for each channel, each with y_i and $g_i(x, t)$ as the output and the total disturbance, respectively. Once such U is obtained, the real control signal is computed as

$$u = D^{-1}U. \quad (34)$$

In practice, we found that D does not need to be known exactly, as some of its perturbation can also be rejected in ADRC.

C. Cascade Control

Consider the two serial-connected subsystems Σ_1 and Σ_2 with both outputs measured, as described in

$$\begin{cases} \Sigma_1 : \dot{x}_1 = f_1(\cdot) + x_2 \\ \Sigma_2 : \dot{x}_2 = f_2(\cdot) + u \\ y = x_1 \end{cases} \quad (35)$$

where the output of the second subsystem x_2 is the input to the first subsystem. Similar to the backstepping technique, we apply ADRC to Σ_1 using $u_1 = x_2$ as a pseudo control variable. Moreover, once u_1 is determined, it is then used as the setpoint for Σ_2 , where ADRC is also applied. In general, the inner-loop should be designed to be faster than the outer loop, which means that the timescale of the inner loop is made smaller than that of the outer loop.

D. Parallel System Control

In many control problems involving resonant modes, they can be generally represented as

$$\begin{cases} \ddot{x}_1 = -\omega_1^2 x_1 - 2\xi_1 \omega_1 \dot{x}_1 + b_1 u \\ \ddot{x}_2 = -\omega_2^2 x_2 - 2\xi_2 \omega_2 \dot{x}_2 + b_2 u \\ \vdots \\ \ddot{x}_m = -\omega_m^2 x_m - 2\xi_m \omega_m \dot{x}_m + b_m u \\ y = c_1 x_1 + c_2 x_2 + \cdots + c_m x_m \end{cases} \quad (36)$$

which can be rewritten as in the SISO form as

$$\ddot{y} = - \sum_{i=1}^m c_i \omega_i^2 x_i - 2 \sum_{i=1}^m c_i \xi_i \omega_i \dot{x}_i + bu \quad (37)$$

with $b = \sum_{i=1}^m c_i b_i$. This is clearly an ADRC problem and should be treated accordingly if

$$-\left(\sum_{i=1}^m c_i \omega_i^2 x_i + 2 \sum_{i=1}^m c_i \xi_i \omega_i \dot{x}_i\right)$$

is taken as a total disturbance.

VI. SOME KEY POINTS IN HOW THE ADRC IS APPLIED

Even though the ADRC is presented for a second-order plant, it is by no means limited to it. In fact, many complex control systems can be reduced to first or second order, and ADRC makes such simplification much easier by lumping many untrackable terms into “total disturbance.” Still, proper problem formulation and simplification is perhaps the most crucial step in practice, and we offer the following suggestions.

- 1) Identify, among many variables in a physical process, which is the input that can be manipulated and which is the output to be controlled. This may not be very clear, and the choice may not be unique in digital control systems when there are many variables being monitored and many different types of commands can be executed. In other words, the control problem itself sometimes is not well defined at the beginning, and until we clearly identify what the control problem is, including its input and output, no amount of control theory can help us.
- 2) ADRC’s order is chosen according to the relative degree of the plant. In linear systems, this is easily obtained in terms of its transfer function. For others, such as nonlinear and time varying systems, this might not be straightforward. From the diagram of the plant, the order of the system can be determined simply by counting the number of integrators in it. In the same diagram, however, the relative degree can be found as the minimum number of integrators from input to output through various direct paths. For example, the plant shown in Fig. 3 is of fourth order, but its relative degree is two.
- 3) The key in a successful application of ADRC is how well one can reformulate the problem by lumping various known and unknown quantities that affect the system performance into total disturbance. This is a crucial step in transforming a complex control problem into a simple one.
- 4) Another effective method for problem simplification is the intelligent use of the pseudo control variable, as shown previously. ADRC shows an obviously quite different way of going about control design. There is really a paradigm shift. Let us once again use the following second-order system for illumination:

$$\begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = f(x_1, x_2, w(t)) + bu \\ y = x_1 \end{cases} \quad (38)$$

where $w(t)$ is an external disturbance. As an extension of the theory of differential equations, the problem of concern in modern control theory may not necessarily

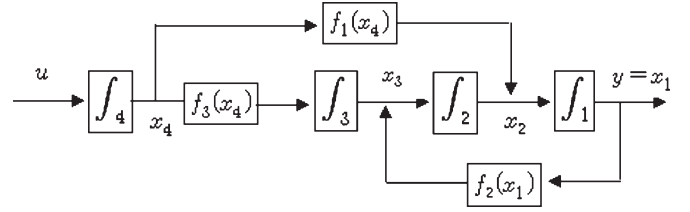


Fig. 3. Relative order determination.

be the same problems faced by practitioners in industry. The key difference may be revealed in how the term $f(x_1, x_2, w(t))$ in (38) is dealt with. Given $u(t)$, the topological structure of the state trajectory in the state space is completely determined by the structure and properties of $f(x_1, x_2, w(t))$, and this problem is known as the open loop analysis problem. Modern control theory is concerned with a different problem. That is, given $f(x_1, x_2, w(t))$, how do we find $u(t)$ as a function of the states so that the topological structure of the state trajectory of (38) has the desired properties. This is the perspective that is dominant in modern control, and this dominance goes back to the 1950s and 1960s of the last century when control theory was seen as a branch of applied mathematics.

Things look much different from a control engineering perspective. The objective in the practice of control system design is quite clear: Manipulate u so that the $y = x_1(t)$ follows a desired trajectory. In this view, $f(x_1, x_2, w(t))$ is a “disturbance” to be overcome by u in order to drive y where it needs to be. For example, let r be the desired trajectory for y to follow. To achieve $y \approx r$, it is necessary that $\ddot{y} \approx \ddot{r}$, i.e., $f(x_1(t), x_2(t), w(t)) + bu(t) \approx \ddot{r}(t)$ or $bu(t) \approx \ddot{r}(t) - F(t)$, $F(t) = f(x_1(t), x_2(t), w(t))$. In this mind set, what concerns us is not the topological structure of the state or the property of $f(x_1(t), x_2(t), w(t))$, but rather, it is the value of the latter at t , or $F(t)$. Moreover, as shown in the ADRC description, $f(x_1(t), x_2(t), w(t))$ is dealt with as the signal to be estimated based on the input–output data and compensated by the control variable u , and it does not really matter what it is as a function of the states. Moreover, the previously all-important distinction between linear and nonlinear time-invariant and time-varying systems becomes irrelevant.

VII. CONCLUDING REMARKS

As far as practice is concerned, PID has dominated the scene for over 80 years, during which, few fundamental innovations were produced. While much progress has been made in model-based mathematical control science, it has yet to make a significant impact in industry, whose pressures force us to seek alternatives elsewhere. This paper provides readers with such an alternative: a new digital control technology arisen out of the effort to address the shortcomings of PID. This paper gives an account of each component of ADRC as well as its structure and philosophy. This paper also demonstrates the problem solving process and the way of thinking in how to apply ADRC to various engineering problems.

ADRC is the result of years of investigation, largely performed experimentally in computer simulations with the scientific spirit of daring imaginations, painstaking observations, careful generalization and abstraction, and truthful verifications of principles in real-world applications. It may help the newcomers to ADRC greatly if one abandons the initial *How can this be right?* attitude and, instead, run a few simulations of the proposed solutions and observe the results. Perhaps the facts, or data, are more convincing than mere articulation of ideas.

If the PID was born early in the last century, where it was instrumental in the rise of modern industry and if PID is still dominant to this day, conceivably, a major overhaul is long overdue and a new era could be just around the corner. Ushered in by performance and practicality, perhaps ADRC will soon be accepted as a viable alternative to PID, as we wait in anticipation.

ACKNOWLEDGMENT

This paper was translated from Chinese to English by a long-time collaborator of the author, Zhiqiang Gao, at Cleveland State University. He also revised this paper after Prof. Jingqing Han suddenly passed away in April 2008.

REFERENCES

- [1] J. Han, "Control theory: Model approach or control approach," *Syst. Sci. Math.*, vol. 9, no. 4, pp. 328–335, 1989, (in Chinese).
- [2] J. Han and W. Wang, "Nonlinear tracking-differentiator," *Syst. Sci. Math.*, vol. 14, no. 2, pp. 177–183, 1994, (in Chinese).
- [3] J. Han, "Nonlinear PID controller," *J. Autom.*, vol. 20, no. 4, pp. 487–490, 1994, (in Chinese).
- [4] J. Han, "Extended state observer for a class of uncertain plants," *Control Decis.*, vol. 10, no. 1, pp. 85–88, 1995, (in Chinese).
- [5] J. Han, "Auto disturbances rejection controller and its applications," *Control Decis.*, vol. 13, no. 1, pp. 19–23, 1998, (in Chinese).
- [6] J. Han, "From PID to auto disturbances rejection control," *Control Eng.*, vol. 9, no. 3, pp. 13–18, 2002, (in Chinese).
- [7] J. Han, "Active disturbances rejection control technique," *Frontier Sci.*, vol. 1, no. 1, pp. 24–31, 2007, (in Chinese).
- [8] Z. Gao, Y. Huang, and J. Han, "An alternative paradigm for control system design," in *Proc. 40th IEEE Conf. Decis. Control*, 2001, vol. 5, pp. 4578–4585.
- [9] Z. Gao, "Active disturbance rejection control: A paradigm shift in feedback control system design," in *Proc. Amer. Control Conf.*, 2006, pp. 2399–2405.
- [10] B. Sun and Z. Gao, "A DSP-based active disturbance rejection control design for a 1-kW H-bridge DC–DC power converter," *IEEE Trans. Ind. Electron.*, vol. 52, no. 5, pp. 1271–1277, Oct. 2005.
- [11] Y. Su, B. Y. Duan, C. H. Zheng, Y. F. Zhang, G. D. Chen, and J. W. Mi, "Disturbance-rejection high-precision motion control of a Stewart platform," *IEEE Trans. Control Syst. Technol.*, vol. 12, no. 3, pp. 364–374, May 2004.
- [12] Y. Su, C. Zheng, and B. Duan, "Automatic disturbances rejection controller for precise motion control of permanent-magnet synchronous motors," *IEEE Trans. Ind. Electron.*, vol. 52, no. 3, pp. 814–823, Jun. 2005.
- [13] D. Sun, "Comments on active disturbance rejection control," *IEEE Trans. Ind. Electron.*, vol. 54, no. 6, pp. 3428–3429, Dec. 2007.
- [14] J.-H. She, F. Mingxing, Y. Ohyama, H. Hashimoto, and M. Wu, "Improving disturbance-rejection performance based on an equivalent-input-disturbance approach," *IEEE Trans. Ind. Electron.*, vol. 55, no. 1, pp. 380–389, Jan. 2008.
- [15] M. Valenzuela, J. M. Bentley, P. C. Aguilera, and R. D. Lorenz, "Improved coordinated response and disturbance rejection in the critical sections of paper machines," *IEEE Trans. Ind. Appl.*, vol. 43, no. 3, pp. 857–869, May/Jun. 2007.



Jingqing Han received the B.S. degree in mathematics from Jilin University, Changchun, China, in 1958. From 1963 to 1966, he studied as a Ph.D. student at the Department of Mathematics and Mechanics, Moscow University.

In 1958, he joined the Institute of Mathematics, Chinese Academy of Sciences, Beijing, China, where he was a Professor Emeritus with the Institute of Systems Science, Academy of Mathematics and Systems Science. He published six books and over 200 papers and served in many professional organi-

zations throughout his career. In his pursuit of truth, he was never afraid of putting his reputation on the line and challenging the establishment and the status quo. Trained as a mathematician, he never ceased to seek solutions to the pressing problems of the world in which we live.

Prof. Han was a leading scholar in China for over five decades, receiving numerous prestigious awards, recognitions, and joint appointments at various universities and research institutes.

Message from the Editor-in-Chief

Prof. Jingqing Han was a control theorist and educator who inspired generations of students and colleagues in China. He passed away on April 21, 2008 in Beijing, China, at age 71. He is survived by his wife of 46 years, Pei Renshun, three grandchildren, and his sons, Xuefeng and Xuejun, along with his daughter, Xuehua (Sarah Young), scattered across both sides of the Pacific in China, Japan, and California, working in concert with his followers to keep his legacy alive.

Prof. Han was known to be an independent thinker who made bold moves in his illustrious career several times over, while making important contributions in several distinct areas of study, including optimal control, game theory, guidance and navigation, population growth, computer-aided control system design, and, above all, active disturbance rejection control (ADRC). Perhaps, he made his boldest move in 1989 when, upon reflecting on the several decades of research worldwide and contemplating a seemingly unbridgeable gap between theory and practice, he categorically rejected the basic premise of mathematization of the world and the pure deductive reasoning in research, in a paper titled "Is it a Control Theory or Is it a Model Theory?" In the following two decades, until his death, he devoted his life to finding an alternative, and find it he did, in ADRC, which turns modern control theory on its head. The implication of this change of direction proved to be enormous both in theory and practice, as hundreds of papers and a wide range of applications and patents soon followed.

Born on November 1, 1937 to a poor peasant family of Korean decent in Changbai, Jilin Province, a remote corner of China, Han first rose to prominence in China and beyond through his eye-opening work, in collaboration with Jian Song, on "Analysis and Synthesis of Time Optimal Control System," presented at the 3rd International Federation of Automatic Control World Congress in Moscow in 1963. Interrupted by the Cultural Revolution, he nonetheless returned afterward to lead a nationwide program introducing modern control theory to a generation of scholars and educators in China whose contact with the West had been cut off for over a decade. As a leading researcher, he continued to make original contributions, such as "the constructive method" in linear system theory, which establishes an intimate connection between the state space representation and the polynomial matrix one, unlike anything available in the West.

Prof. Han had a truly active and curious mind that refused to be bound by artificial boundaries; his work transcends the distinctions between linear and nonlinear systems, time-varying and time-invariant dynamics, modeling error and external disturbances, etc., classifications that often box a researcher in for life. He was a true inspiration to all who knew him, particularly his students and long-time collaborators. With his passing, the continuation of the ADRC legacy seems to largely fall on the shoulders of two groups, one headed by Huang Yi in the Chinese Academy of Sciences, the other by Zhiqiang Gao at Cleveland State University.

Prof. Bogdan M. Wilamowski, Editor-in-Chief
IEEE TRANSACTIONS ON INDUSTRIAL ELECTRONICS