



Contents lists available at ScienceDirect

ISA Transactions

journal homepage: www.elsevier.com/locate/isatrans

On the centrality of disturbance rejection in automatic control

Zhiqiang Gao*

Center for Advanced Control Technologies, Cleveland State University, Cleveland, OH 44115, United States

ARTICLE INFO

Article history:

Received 14 June 2013

Received in revised form

26 August 2013

Accepted 12 September 2013

This paper was recommended for publication by Jeff Pieper

Keywords:

Disturbance

Disturbance rejection

Active disturbance rejection

Copernican revolution

Paradigm shift

Disturbance-free control

ABSTRACT

In this paper, it is shown that the problem of automatic control is, in essence, that of disturbance rejection, with the notion of disturbance generalized to symbolize the uncertainties, both internal and external to the plant. A novel, unifying concept of disturbance rejector is proposed to compliment the traditional notion of controller. The new controller–rejector pair is shown to be a powerful organizing principle in the realm of automatic control, leading to a Copernican moment where the model-centric design philosophy is replaced by the one that is control-centric in the following sense: the controller is designed for a canonical model and is fixed; the difference between the plant and the canonical model is deemed as disturbance and rejected.

© 2013 ISA. Published by Elsevier Ltd. All rights reserved.

1. The problem of disturbance

Disturbance, as defined in the Oxford Dictionary of English, is “the interruption of a settled and peaceful condition”. If engineering is understood as the process of creating an apparatus to serve the human needs, the “settled and peaceful condition” of such an apparatus is the primary concern and the subject of study in the field of automatic control. For example, the apparatus of obtaining mechanical power from steam long preceded the device of automatic control, i.e. the flyball governor that ensures the “peaceful and settled condition” of a constant engine speed, not interrupted by various changes in the operating condition.

There are two distinct areas of study here: (1) the invention of the apparatus to meet the human needs; and (2) the problem of keeping such an apparatus operating in “the settled and peaceful condition”, i.e. the problem of automatic control. This paper is concerned with the latter in a way most general and encompassing. By definition, it appears that the *problem of automatic control* in various domains of applications is reducible, in essence, to the *problem of disturbance*. Understanding automatic control in this manner may give us a refreshingly new outlook and deep insight into the nature of the discipline, beyond its tradition of mathematical rigor. In a way, the strength of a control system is not different from that of a human mind: when disturbed by an adversity, how severely is it interrupted and how fast does it

return to a “peaceful condition”? Similarly in industrial processes, the quality of a control loop is measured by how tough it handles the adversities, i.e. the changes in dynamics, the interferences of unknown forces, etc., collectively known as uncertainties. In this paper, the notion of “*disturbance*” is taken in the most general sense to denote those uncertainties that tend to *disrupt* the working of an apparatus or an organism.

Understood as such, disturbance is what tends to interrupt a nominal course of actions, whether it is a machine or an organism. The problem of disturbance is a problem not just in engineering, but life in general. It is the nature of the interaction between any organism and its environment, and it matters, in terms of survival. It is for this reason that a humanistic approach is taken in this paper, following the trails of ideas and inventions throughout the history of automatic control, the exposition of which will lead to a more comprehensive understanding of the conceptual framework, premises, methodology, and the domains of applications. The end goal is to make the integration seamless between the universal concepts, methods, and solutions of automatic control on one hand, and the idiosyncrasies of particular domains of engineering on the other hand. It is through such integration that practitioners will be able to take advantage of various disturbance rejection techniques freely in solving the pressing problems of today. The need for doing so was articulated some 40 years ago but still rings true today:

“Many scientists are busily at work in laboratories and universities, searching for more advanced control concepts. The principles they are discovering, however, could never realize

* Tel.: +1 2166873528.

E-mail addresses: z.gao@ieee.org, z.gao@csuohio.edu

their full worth if they are not communicated to the people who must apply them. Control problems arise in the plant must be solved in the plant. Until plant engineers and control designers are able to communicate with each other, their mutual problems await solution.” [1]

The questions remain: (1) is there an advance control concept that worth sharing? (2) If so, how can it be communicated to “the people who must apply them”?

Our thesis is therefore quite straightforward: the mutual problem of plant engineers and control designers, which has awaited solution, is the problem of disturbance and it should be the focus of the study in automatic control, the foundation of which begins with the very notion of disturbance, to which we now turn.

2. The evolving notion of disturbance

Like any branch of science, the current generation of researchers in automatic control inherited a particular point of view, an implicit paradigm and an assumed value judgment from the previous generations, the protégés of which continue to shape this discipline from their positions in academia. In the true spirit of science, all points of view, including and especially, the ones that have been dominant, must be continuously reflected upon and judged for relevance and validity. All theories must be scrutinized and verified in practice; cherry picking of evidence to support a particular view, while unavoidable, must be discouraged. “We must be mindful of foundation”, as Prof. Astrom recently advised, or “the building might fall” [2]. Our understanding must be absolutely clear concerning the basic concepts, such as the meaning of *disturbance* or *rejection*, before a solid foundation of automatic control is built.

Specifically, disturbance rejection is a ubiquitous term used in control theory and it is one of the many design considerations in textbook techniques from loop-shaping to H_∞ . It has essentially become an academic jargon and is used freely without a second thought. In reality, the word “*disturbance*” takes on different meanings as it is used in various contexts, often unknowingly. It often refers to a disruption coming from an external force, as, for example, wind gust acting on an airplane; in some other cases, however, it may not be entirely clear whether the disruption is coming from within or without. People seldom, if ever, ask “what do you mean by disturbance?” or “in what sense a disturbance is said to be *rejected*?”

Going back to the Oxford Dictionary of English again, the word “*disturb*” means “interfere with the normal arrangement or functioning of”, from which “*disturbance*” is inferred to as something abnormal, not part of the original plan, but tends to disrupt it. It is used synonymously with *uncertainty* an automatic control system is designed to deal with. “If there is no uncertainty in the system, the control, or the environment, feedback control is largely unnecessary” [3], said renowned control theorist R. Brockett. If one is only concerned with the type of the uncertainties that tend to “interfere with the normal arrangement or functioning of”, then the problem of uncertainty is no different from the problem of disturbance and will be denoted so in this paper for the sake of simplicity and clarity.

As in any human endeavor, a lot of what is being done stems from habits, consciously or otherwise. It is a habit that the problem associated with the uncertainties internal (external) to a physical system is denoted as the robustness (disturbance) problem. In fact, such robustness problem has become a dominant theme in modern control theory and the topic of endless books and papers; the disturbance problem, on the other hand, is narrowly defined and is treated without much fanfare. The real world, however, does not draw a line to separate internal uncertainty from the external one. In a robot manipulator, for example, what would one

call the problem caused by the coupling among various joints? To each joint the disturbance coming from other joints are external but to the robot as a whole all joints are internal.

Treating the coupling among joints in a robot manipulator as robustness problem presupposes the detailed mathematical model of it, which could be quite nonlinear and complex. This practical hindrance led researchers to a shortcut: to each joint, the coupling force from other joints can be estimated in real time and canceled, resulting in a much simpler and more effective solution called disturbance observer (DOB) [4]. But by habit, the notion of disturbance used in the framework of DOB distinctly refers to something external, even though the solution equally applies to disturbances that are state dependent, thus creating an awkwardness in articulating exactly what kind of uncertainties with which the method deals.

This conceptual ambiguity was resolved once for all by Han in his landmark paper of 1989, which for the first time put forward the idea that, for the purpose of controlling a physical process, linear or nonlinear, having a complete mathematical model is both impractical and unnecessary [5]. Physical system, Han believes, can be controlled without a mathematical model because the information it needs can be extracted from the input–output information. Han went on to open a new front of automatic control and spent the next two decades cultivating it [6–8]. This new area of research is known as active disturbance rejection control (ADRC), with the “disturbance” referring to both internal (state dependent) and external forces that are unknown.

The work of Han demonstrates that many boundaries in control theory are artificial, reflecting not the nature of automatic control but our limitations in comprehending it. Such boundaries include those that divide the systems as linear and nonlinear controls, internal dynamic uncertainty and external disturbance, time varying and time invariance, etc. In fact, all these problems can be seen as one and the same: the problem of disturbance, if the word “disturbance” is allowed to take on the more general meaning described above. Perhaps the most unique contribution from Han’s work is his notion of disturbance, which includes uncertainties both internal and external to the physical process [6–8]. A more detailed account will be given later in this paper.

Taking it one step further, the term *disturbance* can now be used to denote the difference between what the system is and what it is should be, whether this difference belongs to the internal or external uncertainties. But “*what it should be*” seems problem dependent, vague and fleeting; it seems tied to the particular system of interests, not easily describable as a universal concept.

With a penetrating insight, Han gave answer to this puzzle in 1979 by showing that under reasonable conditions, all linear and nonlinear systems can be reduced to the cascade integral form that he termed “canonical form of feedback systems” [9], which is denoted in this paper, for the sake of convenience, as Form Han (FH). In other words, most systems, linear or nonlinear, with state feedback, can be transformed into FH, based on which control design can be standardized. Therefore, FH becomes the point of departure for all design methods because various types of physical systems, once they are reduced to FH, are identical dynamically and can be controlled by a standard, fixed controller.

Note that FH anticipates much of the later, more complete, work on generalized control canonical form for linear and nonlinear systems by M. Fliess and others. See [22] for details and references therein. In particular, Han discussed in [9] the problem of transforming a general nonlinear system to the control canonical form by the use of input-dependent state transformation, to which Fliess gave a rather expanded account a decade later [22].

Han also gives us an alternative answer to the above question of “what an ideal plant should be”. Using FH as the ideal plant,

controllers are readily designed to meet the design specifications. To Han, the problem of automatic control comes down to the problem of what to do with such disturbance, to which we turn next.

3. Disturbance attenuation vs. disturbance rejection: a conceptual divide

To grasp for universal principles with explanation power is a part of human nature. The task has puzzled the mankind from the beginning of civilization, in the quest to transcend the particulars, and it is especially challenging for engineers, who tend to focus on how things are done and not to dwell on what to be done and why. The terms feedback and feedforward, for example, are solutions but the problems they solve are seldom explained clearly. What does feedback do exactly? Prof. Wiener called it “a method of controlling a system by reinserting into it the results of its past performance” [10]. But why? What is the aim and how is it helpful in making the system remember what it has done in the past? Similarly, in the literature we see terms like “two-degree-of-freedom” (2DOF) or “combined system”, begging the questions of “the freedom to do what?” and “what is combined and why?”

There is a profound reason for such difficulty. In modern times, the presentation of an idea becomes increasingly more important, to the point of “presentation is everything”. In the field of automatic control, engineering insights were gradually replaced in scholarly transactions by elegant but sometimes empty mathematical symbols and academic jargons. “Everywhere that which is apparent on the surface is reprinted, but nowhere (with very few exceptions) does the investigation go below the surface”, said Prof. Trinks in 1919 [11]. How much progress has been made since then is an open question.

In this section, we strive to continue what Prof. Trinks started in his “book of essentials and principles”: to make clear the fundamental concepts and principles of automatic control. Martin Luther King dreamed that a person is judged not on the color of his skin but on the contents of his character. Likewise, we dream that in the field of automatic control, a scholar and her scholarly work are judged on the contents of her ideas, whoever and wherever the person is.

3.1. The notion of disturbance rejection

The term “disturbance rejection” used in the context of classical and modern control theory refers to how the external disturbance is attenuated as it is propagated through the process, eventually affecting the output. It is measured as the magnitude of the frequency response that defines the disturbance–output relationship. It is specified, usually, in terms of the amount of attenuation and the corresponding frequency range required. Used in such a manner, “rejection” is synonymous to attenuation, or mitigation, or compensation. Shaping system response, in frequency domain, to external disturbances is what it really means by “disturbance rejection” in the current textbooks on control, unfortunately.

The word “reject” comes from Latin, meaning “throw back”, and it has a derivative, rejector, to which we will come back. The word “reject” projects a sense of totality and finality, regarding the object of concern. Putting the two words together “*disturbance rejection*” should, and therefore does from now on, mean literally “no interruptions, whatsoever, of “a peaceful and settled condition”. If, as discussed above, the problem of automatic control is the problem of disturbance, then *disturbance rejection* is *absolutely central*.

To be clear, in an ideal control system, the disturbance, the sum total of the internal dynamics and the external forces, should have absolutely no effect on the operation of the system as designed!

Idealization is a powerful tool in science that helps us distilled pure concepts from messy details of everyday world, as shown in Newtonian physics (the first law of motion) and Einstein’s theory of relativity (a human traveling at the speed of light). Likewise, the concept of disturbance rejection, as defined above, helps us escape the drudgery of academic jargons and engineering idiosyncrasies, to arrive at a universal goal, common to all engineered systems. And, as a side note not to be pursued further, it would not be difficult to see how important disturbance rejection, as redefined, is to us human beings in the happiness and peacefulness of our lives.

The renewed conception of disturbance rejection will help us reestablish the “essentials and principles”, as shown below.

3.2. The principle of flyball governor and disturbance attenuation

Idealization and FH allow us to grasp the principle behind the flyball governor as follows. Let u be the input (driving force), y the output (engine speed), and r the reference (desired speed). In the engine speed control system with the flyball governor, the steam flow that provides the driving force to rotate the engine shaft is made proportional to the tracking error defined as $e=r-y$. That is, the FH of steam engine is

$$J\dot{y} = u \quad (1)$$

where J is the inertia and the flyball governor is described as

$$u = k_p e \quad (2)$$

where k_p is the gain of the governor. The idea behind (2) is that if y is to be governed to follow a given value r , then its derivative needs to be made proportional to the tracking error $e=r-y$. And this is denoted as *the principle of flyball governor* (PFG), which is what behind the invention but has not been articulated as such until now.

From the PFG as defined, it can be seen why people had a hard time swallowing it. “The principle in question has been pronounced defective and faulty, because, to cause the governor to act, it necessitates a change in the quantity to be kept constant”, Trinks wrote in 1919 [11], in the first comprehensive book on the governors and the principles of governing in the English literature. In other words, substantial tracking error must exist, no matter what caused it, before the governor can act on it for the purpose of reducing it to zero. But there is always a cause–effect relationship behind the tracking error; the PFG basically ignores the cause but acts on the effect. It is for this reason we denote the PFG also as the *principle of disturbance attenuation* (PDA) and its explanation powers go far beyond the steam engine. In fact, much of what has been in the realm of existing control theory can be seen as rooted in PDA. It is ubiquitous but unstated in all aspects of automatic control, theory and practice.

Described more generally, there is first a causal relationship between the control variable (input or cause) and the controlled variable (output or effect); second, there is a desired value for the output (setpoint); third, there is an unaccounted force (disturbance) that tends to drive the output away from its desired value; finally the control variable is made a function of the difference between the output and its desired value, thus closing the loop. That is, the control action is driven by the difference between the controlled variable, i.e. the output, and its desired value for the purpose of *attenuating* the effect of the disturbance.

“Faulty” as it appears to be, the PFG, or PDA, survived. It indicates a foundational characteristic an automatic control system possesses: the ability to adjust, by itself, in response to disturbances and to return to the “settled condition” of its original design. This act of “governing” gives the mechanism the name “governor”, from which the name “cybernetics” is derived. In other words, cybernetic

symbolizes a principle of governing that can be described as PDA; it is a goal for which feedback is the means.

To rest the entire control theory on the notion of feedback, unfortunately, is to mistake the goal for the means. The mistake is the result of the obsession with the “how” at the cost of ignoring the “what” and the “why”. Feedback control is how we accomplished disturbance attenuation; it is not the goal in itself. A particular governor is designed to govern and is but one embodiment of governing. But to mistake governing itself as a particular governor, no matter how great it is, is unfortunate. Sadly that is exactly what happened, as the well-respected historian, Otto Mayr, testifies that “this field [of automatic control] is essentially based upon a single idea, that of feedback loop” [12]. In other words, the problem of automatic control has been mistaken as that of feedback loop.

3.3. The principle of isochronous governor and disturbance rejection

In the idealize form, the flyball governor of (1) and (2) would never oscillate. But in reality it does, because, as Poncelet [13] points out, there are two sets of motions involved: the motion of steam engine, i.e. the plant, and the motion of the flyball, i.e. the controller. Clerk Maxwell set mathematical control theory on its path of development by describing the control system with a differential equation and determining the condition of oscillation from its characteristic polynomial. Routh continued this line of investigation [14] but came to an erroneous conclusion that the cause of the oscillation is that the governor acts too fast, not realizing it is the lag in the action of the governor that causes the oscillation, as pointed out by Poncelet.

To solve the oscillation problem Poncelet proposed the Isochronous Governor, one that does not have any lag, ideally, in action. Linearizing the steam engine at a particular operating point, a simplified equation is

$$J\dot{y} = -ay + d + u \quad (3)$$

with a as the friction coefficient and d is the external disturbance. The sudden appearance of the disturbance force, d , tends to jolt the engine out of its equilibrium. Poncelet sketched a diagram of a different type of governor, namely the *Isochronous Governor*, that measures d and cancels its effect on the engine with an instantaneous opposing action of steam valve. This is denoted as Poncelet's principle in [11]. A more descriptive definition adopted here is *principle of disturbance rejection* (PDR), which is the opposite of PDA.

If PDA signifies an acceptance, an admission, however reluctant, that the operation of the process will be interrupted or interfered with by disturbance, the principle of disturbance rejection (PDR) is exactly the opposite. True to the root meaning of the word “reject”, PDR signifies the idea that the causes of the output deviation, i.e. disturbance, can be determined and “thrown back”, i.e. canceled. Furthermore, once this cause is removed, the output deviation can be, to a large extent, *avoided*. And this is the main difference between PDA and PDR.

Perhaps the earliest example of PDR is the South-Pointing Chariot (SPC) that, by legend, goes all the way back to the Yellow Emperor of the third millennium BC, the ancestor of all Chinese [15]. This apparatus has a figure mounted on top of a chariot with a finger always pointing to the south, no matter how the chariot moves and turns. It is made possible by an ingenious gear system that measures the rotation of the chariot, from the difference in rotation between the wheels on either side, and counter rotates the pointing figure by the same amount.

Note that in this governor, the output is not measured, the disturbance (rotation of the chariot) is. That is, there is no output feedback, as is defined in feedback control texts. In this sense,

Poncelet's Isochronous Governor is quite similar to SPC, as the steam valve is immediately adjusted by the detection of the load change. More than a century later, Poncelet's idea was revived and formalized as invariance principle (IP) by Shipanov from the then Soviet Union in the form of a set of mathematical conditions under which the output of a process is made “invariant” in the presence of an external disturbance [16]; it led to the two-channel principle [17] where the disturbance is measured and fed to controller for the purpose of achieving the “absolute invariance”, i.e. the ideal disturbance rejection.

Clearly evident in Poncelet and Shipanov's work is the recognition that the problem of control is that of disturbance rejection, despite the various limitations in their solutions. This point view greatly influenced a Chinese graduate student who studied in Moscow in 1960s and later made the next breakthrough in the 1990s, continuing the progress of this idea that simply refuses to go away [5–8,18,19].

Han defines disturbance in a more general sense, to include uncertainties both internal and external. To reject such disturbances, according to Han, requires that their impact on the process be equivalent to an uncertainty in the input channel, which can be treated as an extended state, estimated via a state observer, and canceled by the control action. This new framework of a generalized notion of disturbance, its estimation and cancellation, and the control of the remaining plant by various means were systematically and meticulously developed by Han over a period of two decades, leading to what is known as active disturbance rejection control (ADRC).

Specifically, by including dynamic uncertainties as a part of disturbance internal to the process, a much broader framework is established that could unite the previously separated subjects in control theory: robustness, adaptive control, and disturbance rejection in the narrow sense, etc., reducing the basic problems of automatic control to disturbance rejection. In doing so, the previous principles of Poncelet and Shipanov are made applicable to those processes where the model is mostly unknown and the disturbance is now estimated, not measured, making it a much more practical solution.

4. Disturbance rejector and disturbance rejection control

Close to 200 years had passed from Watt's invention of flyball governor to Wiener's articulation of *feedback* as the underlying concept with enormous explanation power. Consequently, the previous terminologies used by different groups of practitioners from different places at different times were unified under a single concept, the concept of *feedback*. Correspondingly, the term “*controller*”, if not otherwise qualified, has become synonymous with the feedback control mechanism.

Likewise, from Yellow Emperor's SPC, to Poncelet's Isochronous Governor, to Shipanov's invariance principle, and finally to Han's ADRC, these ingenious human inventions share the distinct, if otherwise unstated, trait of *disturbance rejection* not found in feedback control and they also share the unique machinery of a *disturbance rejector*, in parallel to the *controller*. Specifically, “*disturbance rejector*”, or simply “*rejector*”, is defined here as a mechanism that obtains the disturbance information and cancels it out with the control action. Furthermore, the combined controller–rejector pair forms the basic structure of *disturbance rejection control* (DRC), as shown, for example, in Fig. 1.

The basic idea of DRC is to not let the controller interface directly with the messy physical processes, full of nonlinearities, uncertainties, and other nastiness. Instead, the controller regulates the enforced plant, i.e. the transformed plant where the nastiness has been taken out, more or less, by the rejector. The controller's

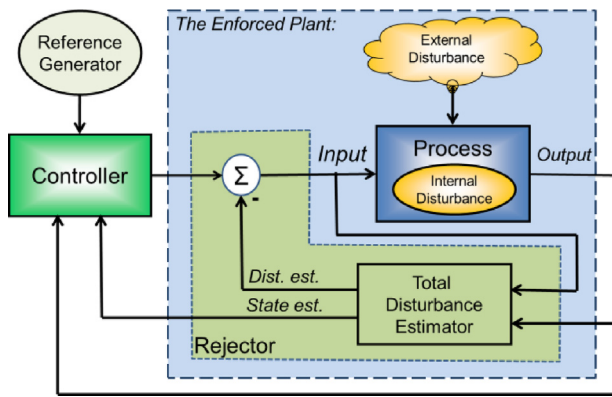


Fig. 1. Disturbance rejection control.

task of meeting the design specifications becomes a lot easier when it deals with the enforced plant, which tends to be much simpler and less uncertain than the actual process. And this is the idea that has been missing in the controller-only paradigm of modern control.

Whether it is robust control or adaptive control, or even the feedback control itself, the goal has always been to deal with uncertainties, and at the same time, meet the performance specifications. But making a single controller do both proves to be unnecessarily complex or even intractable at times, to which DRC provides a genuine alternative.

4.1. Disturbance rejector, the enforced plant and the renewed notion of model

The *rejector*, as defined above and shown in Fig. 1, performs two key tasks: (1) collecting, processing and communicating the disturbance and state information; (2) canceling, or eradiating, the disturbance. Within the framework of DRC, the task of automatic control is shared between the controller and the rejector: the rejector rejects and the controller controls. That is, the rejector performs the main task of rejecting disturbance and controller forces the output to follow the command.

The enforced plant represents the dynamics of the process after the disturbance is removed. With the expanded definition of disturbance, even the very notion of the model is redefined. To put it in another way, it is not that DRC is now “model-free”, because no matter how simple it is, there is always a model that pictures the plant to be controlled, based on which a controller is designed; but rather, the very concept of model is renewed.

In the traditional concept of mathematical model, patterning after the natural sciences, the dynamics of a physical system is described as closely as possible by a set of mathematical equations, known as the model. And this set of equations is the starting point of all analyses and designs. A fundamental, if implicit, premise of the entire modern control theory is that the model is given, more or less. And it is in this sense modern control theory can be seen as the “doctrine of model” suggested by Han [5].

In DRC, the demarcation is broken down between what is known as “internal” dynamics and “external” disturbance. The enforced plant represents the dynamics of the system to be controlled; it could be the mathematical model of the physical process in its traditional sense, or it could be chosen in the form of FH, or somewhere in between. In any case, the rejector treats all that is different from the enforced plant as “disturbance” to be rejected. Taking it this way, control design is never entirely “model-free” or model-independent, even in the case when the enforced plant is chosen in the ideal form of FH. Instead, the controller is designed for the enforced plant, the fidelity of which in relation to process dynamics determines the amount of disturbance

the rejector has to contend with. It is in this sense the task of automatic control is shared between the controller and the rejector.

In modern control, the task is entirely shouldered by the controller, in the absence of the rejector and with the premise that a high fidelity model of the actual process is given. In the case of ADRC proposed by Han, the complexity of the controller is minimized as all dynamics aside from FH is deemed as disturbance and rejected, which makes the quality of control system as good as that of the rejector. The balance between the controller and the rejector in practice will fall in the wide range between these two. The better the disturbance estimator, the less the knowledge of the plant dynamics is needed.

4.2. Various means of obtaining the disturbance information

The notion of rejector can perhaps help reorient the previously not well organized set of tools in a not well-defined field of research. Because of the lack of recognition of the underlying universal principle, even though various forms of disturbance estimation and rejection have been proposed in the solutions of various problems, the researchers are mostly unaware of the field as a whole and each other's work in particular. In early 1970s, the unknown input observer (UIO) was proposed to obtain the disturbance information with a state observer, assuming that the model of both the plant and the disturbance was given; in the late 1980s, another method, the disturbance observer (DOB), was proposed for robotic manipulators where the disturbance is calculated by sending the output of the plant to its inverse transfer function and subtracting the result from the input. DOB was later brought into state space to deal with nonlinear system and the equivalency between UIO and DOB was established for linear time invariant system with external disturbance. In the 1990s, the Extended State Observer (ESO) was proposed specifically to estimate and cancel “total disturbance” in the context of ADRC. A survey of these and other types of state and disturbance observers can be found in [20] and some recent work on DOB based design in the context of robust control can be found, for example, in [21].

In addition, the recent work of Fliess and Join on “model-free control” shows yet another instance of disturbance rejector at work [23], albeit anonymously, the uniqueness of which resides in the real time estimation of disturbances using the algebraic identification method and the proposition that most, if not all, physical systems can be treated as first- or second-order system [24]. Another form of rejector is the embedded model control approach proposed by Canuto [25] where the uncertainty is estimated online and canceled with control action. These are but a few examples, all of which can be well explained by the concept of total disturbance and the structure of DRC, shown in Fig. 1, and understood as a design philosophy of actively estimating and canceling the total disturbance, regardless of the means. In doing so, all methods discussed above share a common objective, understanding, and structure; the only difference is in how the disturbance is estimated and in the premises made.

4.3. Disturbance-free control and a paradigm shift

The unifying concept of disturbance rejector, together with the structure of DRC, marks both a paradigm shift in how controller is designed and a conceptual turning point in how disturbance is dealt with. The schematics of control system took a drastically different form, as shown in Fig. 1. Note that although the internal and external disturbances are very different in nature and they enter into the process throughout its physical layout, they are nonetheless lumped into the “total disturbance” that affects the

process in the same channel where the process input is, and there is a deep insight behind.

Specifically, all disturbances that could cause the output to change is, in a sense, observable; furthermore, since the process input by design is to cause the output to change, the output change due to any disturbances, no matter what the actual disturbance is, can be seen as the result of input abnormality; finally the process input, in turn, can be manipulated to cancel this abnormality and to return the process dynamics to that of the ideal, disturbance-free, form. It is for this reason that the DRC perhaps can be better understood as disturbance-free control (DFC).

This kind of disturbance rejection is deemed active because it doesn't wait for the disturbance to work its way through the physical process and cause significant changes in the state and output. The Total Disturbance Estimator in Fig. 1 denotes the machinery for this purpose and it can take on various forms and names, as discussed above. Regardless of the particular mechanism employed to estimate the disturbance, the role of the Rejector cancels it by subtracting its estimate from the process input. Most, if not all, mechanisms of disturbance rejection described earlier can be reduced to the structure of DRC in Fig. 1, in which the elusive universal principle, the principle of disturbance rejection, is on vivid display.

To describe it yet in another way, Han defines that the mathematical model of all feedback systems to be controlled as that of FH, i.e. in the cascade integral form, on which the controller can be designed using any existing method. The job of disturbance rejector is to enforce this model by estimating and canceling all the discrepancies in the physical system. In this framework, the difficult and complex control problems, such as those associated with nonlinear, uncertain, time-varying process of multi-input and multi-output with strong couplings among variables, become easily solvable after the rejector has done its work. In this sense, the rejector is also an enforcer, enforcing the plant to behave like FH and inducing a completely different mindset on how control system is designed.

This disturbance-centric paradigm shift in control design symbolizes a drastic departure from the so called modern control theory. To be specific, it calls into question the very basic premise: what needs to be understood, at the minimum, regarding the physical process to be controlled. *To assume complete knowledge of it, more or less, is to assume the problem away, more or less.* It merely shifts the burden from control to system identification, often at a significant cost. But the history shows that the temptation of mathematization of automatic control is too strong to resist: *the problem of control design can be turned into a problem of mathematical deduction if both the measure of optimality and the physical process can be reduced to a set of precise mathematical equations.* In reality, however, *neither is the case*; to insist otherwise, to continue to ignore the facts in control practice, is merely delusional.

The practice of control engineering, just like any other human activities, is full of trade-offs, which are hard to put in precise mathematical terms. The benefits are always weighed against the costs, manifested in various forms. All practitioners are concerned with cost and, by extension, the economy of model: high fidelity comes with a high cost and it must be justified by the benefits gained. For example, the use of model predictive control, with costly models, can be justified in certain applications in process industry, even though it is still dominated by PID as a whole. Arguably, the prevalence of PID in all industry sectors speaks a single truth: *the cost of model-centric design, as suggested by modern control theory, does not justify its costs in general.* By the same token, high gain designs, as advocated by some preeminent scholars, often ignored the associated high cost. This failure in recognizing the cost-benefit relationship is at the bottom of the

existing theory-practice divide that is still poorly understood, it appears. On the other hand, the reality of control design with low fidelity models, or with no models at all as some claimed, necessarily shift our attention in control engineering back to the problem of total disturbance. This is at the core of the “paradigm shift” in control science: *shifting from model-centric to disturbance-centric in framework and in mindset.* The result could be profound, as further explained next.

4.4. A Copernican moment

The controller-rejector tandem, shown in Fig. 1, provides a rare Copernican moment where our fundamental conception of automatic control is transformed. Ever since the inception of modern control in the middle of the last century, the mathematical model of the physical process is presumed given and is the point of departure for analysis and synthesis. Unlike the filter design, control design is generally not reusable and is thus performed for each problem, under each set of specifications. Everything revolves around the mathematical model that supposedly captures the plant dynamics globally. But is this absolutely necessary? Is this an overkill that puts unnecessary burden on the part of modeling? Is it realistic to assume complete knowledge of the physical process globally? There were the questions Han contemplated over two decades ago and his answer was a resounding no [5]! Interestingly, Fliess and his collaborators [23] later came to the similar conclusions, symbolized by their *ultra-local* model of dynamic systems. The remarkable parallel between Han's earlier vision and Fliess's later work is striking, even though they were separated geographically, culturally, and in time.

This Copernican moment, this completely different orientation of the minds, is on vivid display in Fig. 1, where the controller needs to be designed only once and where the rejector forces the physical process of all kinds, linear or nonlinear, time varying or time invariant, deterministic or stochastic, etc. to behave like the predetermined enforced plant based on which the controller is designed. When the enforced plant is chosen to be in the form of FH, for example, the rejector treats all departures from FH as disturbances and cancels it with the control action. Instead of various controller designs revolving around the mathematical model of the plant, this time, in this framework of the controller-rejector tandem, the plants revolve around the controller, trying to fit into the form the controller dictates, with the help of the rejector!

The center of the universe is therefore shifted from the model of the plant to the controller. All imperfections and uncertainties in the process are lumped into what is known as total disturbance to be rejected. In other words, instead of making a controller adjust to the process, like in adaptive or robust control, the process is made “disturbance-free”, by the rejector, to adjust to what the controller is designed for. This, after all, is the authentic doctrine of control, as Han anticipated some 20 years ago, where the problem of automatic control is not assumed away and where the problem of disturbance has returned to its proper place: the center and focus of automatic control. There will be a period of transition, of course, perhaps a generation or more, for this Copernican-like revolution in automatic control to take its course, because it necessarily leads to the devaluation of certain principles and methodologies previously held in high esteem. And it will be a challenging transition, to say the least, for the scholars who built their careers and reputations within the confine of the modern control paradigm. But this ideal of *disturbance-free control* captures the essence of what engineers all aspire to achieve and it will eventually prevail.

5. Conclusions and outlook

In this paper, we first reestablished the concepts of disturbance, disturbance rejection, and disturbance rejection control, the pillars of the on-going investigation. In particular, the expanded notion of disturbance brings under a single umbrella both the state-dependent internal uncertainties and the unknown external forces. This gives us a new paradigm to address the problems of all major branches of modern control theory, such as robust control, nonlinear control, decoupling, adaptive control. In other words, most problems of robust control, nonlinear control, decoupling, adaptive control, etc., can be reformulated as particular forms of disturbance problem and solved in a completely new and exciting way.

Furthermore, two important but otherwise unannounced principles of automatic control are articulated in this paper: the principle of disturbance accommodation and the principle of disturbance rejection, with the distinction between them setting the stage for a new synthesis. The notion of disturbance rejector is proposed by which previously overlapping concepts and methods of disturbance estimation and cancellation are now finally unified. The controller–rejector pair in the framework of disturbance rejection control signifies a sea change in how we understand the business of automatic control: from a model-centric to a control-centric cosmic view; it reminds us of the Copernican revolution in the early days of modern science. It helps us return to the journey that started some thousand years ago: to make our engineered systems *disturbance-free*.

With the renewed focus on disturbance rejection and a fresh new look at automatic control, low hanging fruits abound in the realms of both academic research and technological developments. Led by ADRC, this new form of solutions has found many fields of applications, including, but certainly not limited to, industrial controls such as servo, temperature, web tension; aerospace and aeronautics; high energy physics, to name a few. The readers are referred to [26] for a summary of these applications. Between the early work on flight control [27] and the latest news on the adoption by the industry giant Texas Instruments [28], the mounting evidence testifies for the thesis of this paper: the centrality of disturbance rejection in control engineering practice.

There are enormous opportunities that come with the new paradigm of control science, including three areas of new discoveries that will likely see immense growth in the near future: principles, technologies and applications, as shown in Fig. 2. In addition, theoretical analysis and justification are needed for all three areas, just like the work of Bode and Nyquist which justified and made it intelligible the invention of feedback amplifier.

In particular, under principles, a summarily study of all principles of controller and rejector design, once completed, should give us a fresh new outlook on the state of automatic control and

directions for new discoveries. For example, Fliess and Siramirez's [24] work on algebraic identification provides a drastically different means for disturbance estimation in the future; Fliess and Join's [23] new take on model-free control threatens the conventional definition of the system order and gives a provoking new distinction: the order of the process as it is understood and that as it is controlled. The exposition of such distinction is of great interests but beyond of the scope of this paper.

In terms of technologies, the new principles of disturbance rejection, as applied to address critical needs in every industry sectors, will likely lead to brand new control technologies. The new motion control technology based ADRC is but one example [28].

Finally, in terms of applications, the domain experts, once brought up to speed with the new design concepts, will see the fundamental change in not just how control is designed, but also how systems and solutions are conceived. Shinsky [1] was right in that this advanced concept of controller–rejector needs to be communicated to those at the front line of production; the hundred-fold improvement he observed in the past and the over 50% energy saving obtained recently are testaments of the power of such ideas, as do the applications summarized in [26]. To this end, the centrality of disturbance rejection and the objective of disturbance-free control should be firmly established in all future application researches. To answer the earlier question, these are the ideas and ideals that should be “applied” in future research in automatic control.

References

- [1] Shinsky FG. Process control systems. New York: McGraw-Hill; 1967.
- [2] Karl Johan Astrom. American control conference plenary lecture, Montréal, Canada; June 2012. Available at (<http://www.iececs-oll.org/>).
- [3] Brockett JR. New issues in the mathematics of control. In: Engquist B, Schimidt W, editors. Mathematics unlimited–2001 and beyond. Springer; 2001. p. 189–220.
- [4] Ohishi K, et al. Microprocessor-controlled dc motor for load-insensitive position servo system. IEEE Reactions on Industria Electronics 1987;IE-34:6.
- [5] Han J. Control theory: the doctrine of model or the doctrine of control? System Science and Mathematics 1989;9(4):328–35 [in Chinese].
- [6] Han J. The robustness of control system and Godel's 'incompleteness theorem'. Control Theory and Applications 1999;16(December):149–55 [in Chinese].
- [7] Han J. Active disturbance rejection control technology. The Defense Industry Press; 2008 [in Chinese].
- [8] Han J. From PID to active disturbance rejection control. IEEE Transactions on Industrial Electronics 2009;56(3):900–6.
- [9] Han J. The structure of linear control system and computation in feedback system. Presented at the national meeting on control theory and applications, Xiamen, 1979. Proceedings of the national meeting on control theory and applications, Science Press; 1981 [in Chinese].
- [10] N. Wiener, Cybnetics, MIT Press, 1948.
- [11] Trinks W. Govenors and governing of prime movers. New York: D. Van Nostrand Company; 1919.
- [12] Mayr O. The origin of feedback control. MIT Press; 1970.
- [13] Poncelet J-V. Introduction a la Mecanique Industrielle. In: Kretz M, editor. Physique ou Experimentale. third ed. Paris: Gauthier-Villars; 1870.
- [14] Bennet S. A history of control engineering, 1800–1930. Stevanage: Peter Pergrinus, IEE; 1979.
- [15] Liu Xianzhou. The Chinese history of mechanical inventions. Science Press; 1962 (May).
- [16] Preminger J, Rootenberg J. Some considerations relating to control systems employing the invariance principle.. IEEE Transactions on Automatic Control 1964;9(June (3)):209–15.
- [17] Petrov BN. The invariance principle and the conditions for its application during the calculation of linear and non linear systems. In: Proceedings of the first congress of IFAC, vol. 1, Moscow, USSR, Thornton Butterworth, Ltd., London, England; 1961. p. 117–26.
- [18] Han J. A class of extended state observers for uncertain systems. Control and Decision 1995;10(1):85–8 [in Chinese].
- [19] Han J. Auto-disturbance rejection control and its applications. Control and Decision 1998;13(1) [in Chinese].
- [20] Radke A, Gao Z. A survey of state and disturbance observers for practitioners. In: American control conference; June 2006. p. 6.
- [21] Wei X, Guo L. Composite disturbance-observer-based control and H_∞ control for complex continuous models. International Journal of Robust and Nonlinear Control 2010;20:106–18.
- [22] Fliess Michel. Generalized controller canonical forms for linear and nonlinear dynamics. IEEE Transactions on Automatic Control 1990;35(September (9)).

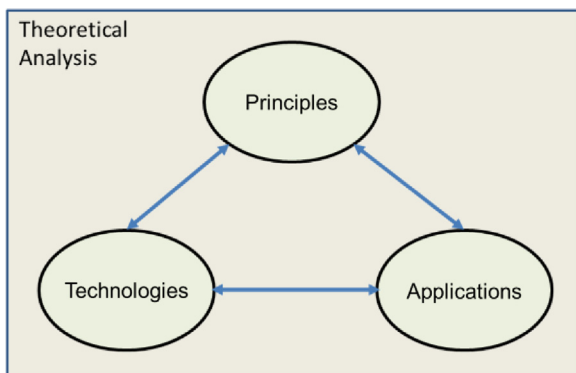


Fig. 2. Future research areas.

- [23] Fliess Michel, Join Cédric. Model-free control. *International Journal of Control* 2013(July). <http://dx.doi.org/10.1080/00207179.2013.810345>.
- [24] Fliess M, Sira-Ramirez H. Closed-loop parametric identification for continuous-time linear systems via new algebraic techniques. In: Garnier H, Wang L, editors, *Identification of continuous-time models from sampled data*. Springer; 2008. p. 362–91. ISBN 978-1-84800-160-2 e-ISBN 978-1-84800-161-9.
- [25] Canuto E, Acuna Bravo W, Molano A, Perez C. Embedded model control calls for disturbance modeling and rejection. *ISA Transactions* 2012;51(5):584–95.
- [26] Qing Zheng, Zhiqiang Gao. On practical applications of active disturbance rejection control. In: *Proceedings of the 2010 Chinese control conference*, Beijing, China; July 29–31, 2010.
- [27] Huang Y, Xu K, Han J, Lam J. Flight control design using extended state observer and non-smooth feedback. In: *Proceedings of the 40th conference on decision and control*; 2001. p. 223–8.
- [28] Achieve improved motion and efficiency for advanced motor control designs in minutes with TI's new InstaSPIN(TM)-MOTION technology. *The Wall Street Journal*, April 18, 2013. Accessed on April 19, 2013 at (<http://online.wsj.com/article/PR-CO-20130418-907338.html>).