In this issue of *IEEE Control Systems Magazine* (*CSM*), we are pleased to present reviews of four interesting books that were released recently. These books consider constrained control systems, nonholonomic mechanics, stochastic optimization, and queuing systems. I wish to thank the reviewers for providing in-depth discussions of the contributions of these books. The reviews also include extensive tutorial and informative background material.

IEEE CSM publishes reviews of books that pertain to systems and control. If you are the author (or publisher) of a recent book in these areas, please contact me at the address below. Also, please contact me if you are interested in reviewing a recent or soon-to-be-published book. I can be reached at:

Kirsten Morris
Department of Applied
Mathematics
University of Waterloo
Waterloo, Ontario N2L 3G1 Canada
E-mail: kmorris@uwaterloo.ca

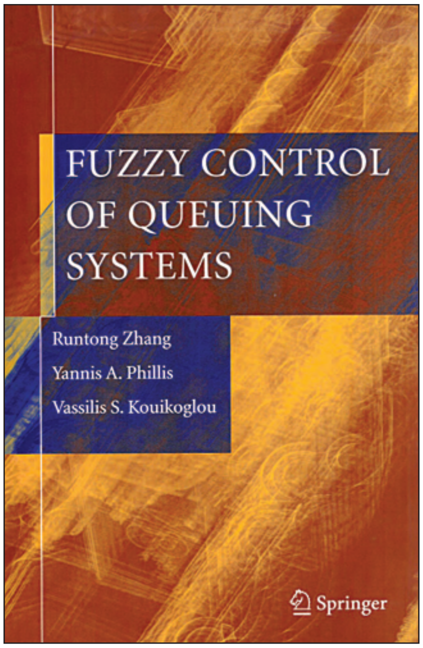
Fuzzy Control of Queuing System by Runtong Zhang, Yannis A. Phillis, and Vassilis S. Kouikoglou, Springer, 2005, 175 pp., ISBN: 1-85233-824-5, US$149.00. *Reviewed by Joseph L. Hellerstein.*

Queuing Theory and Its Applications

As a researcher and software engineer, the appeal of fuzzy control is its potential to simplify greatly the design of complex software systems. However, this book is not about fuzzy control of *software systems*. Rather, this book is about fuzzy control of *queuing systems*. Fortunately, there is a close connection between the two.

To understand this connection, it is useful to review a few aspects of queuing theory. Queuing theory provides techniques for estimating the performance of computing systems and data networks based on assumptions about the interarrival times of requests and their service times. Queuing analysis uses metrics such as utilization (the fraction of time that a server is busy), queue length (the number of requests in the queue), and response time (the time from the arrival of a request until the request is filled). For example, queuing theory has been used to estimate the response time of routers in packet networks such as the Internet. Here the router is a server, work arrives in the form of packets, and service times depend on the speed of the router and the complexity of the packets.

Another area in which queuing theory has been successfully applied is capacity planning, which involves selecting computers and communication links to achieve desired response times at minimal cost. Here, resources such as the central processing unit (CPU) of computers, their memory, network interfaces, and network links are each treated as servers. Work arrives in the form of application requests (often Web requests), which require specified amounts of time for different types of resources. More details on queuing theory can be found in [1]–[3].



Although queuing theory has been used in numerous engineering applications, it is rare for queuing theory to be used by practitioners designing commercial software systems. This gap is due to the mathematical background required to apply queuing theory. It is common, however, for software designers to reason about their systems in terms of the insights from queuing theory. For example, a common insight is that there is a “knee of the utilization curve” at which point response times and queue lengths increase vary rapidly with a small change in utilization. This insight is typically used to construct rules of thumb for resource allocations.

Fuzzy rules are based on the same kind of intuition that software designers use in practice. This book shows how fuzzy rules can be used to control queuing systems. After some background on queuing theory, fuzzy logic, and fuzzy control, the book systematically describes how to apply fuzzy control to individual queuing systems, networks of queuing systems, and data networks.

Contents of the Book

Chapter 1 provides a brief introduction to core concepts used throughout the book. The authors touch on the basics of queuing systems, such as the interarrival and service time processes, the number of servers, the server capacity, the customer population, and the queuing discipline. Also included is a brief summary of several techniques for minimizing convex functions. Although this chapter is a useful refresher for those who have previously encountered the material, the queuing background here is much too brief for someone who is not already familiar with the topic. I would advise supplementing this chapter with one or more of the texts cited earlier. In particular, [1] gives a pragmatic overview that only requires a modest mathematical background, while [2] is the classic reference for queuing theory for those more interested in the

0272-1708/05/$20.00©2005IEEE

mathematical details. Meanwhile, [3] focuses on a particular application (capacity planning) of queuing theory, introducing only the technical details needed for that application.

Chapter 2 gives the basics of fuzzy logic in the form of a summary of key results on fuzzy sets, fuzzy set operations, linguistic variables, and fuzzy reasoning. While the presentation is terse, the information is sufficiently complete for readers who have not had previous in-depth exposure to this material.

Chapter 3 addresses fuzzy control. The material is mostly self contained and written in a narrative style that is helpful to the nonexpert. The architecture considered consists of fuzzification, knowledge base, inference engine, and defuzzification. A number of simple examples add to the readability of this chapter.

Chapter 4 provides the initial discussion of fuzzy control of queuing systems. The focus is on queuing systems in which the service rate is the actuator. The first problem concerns the case in which the server goes on vacation, such as when the power is turned off. The authors draw on a range of known analytic results for queuing systems to develop a simple decision criterion, namely, servers should be kept on when there are customers present. As with much of the rest of the book, the authors begin by describing the rule base and then the membership function, followed by a numerical example. These results are extended to parallel servers, servers without switching costs, and, finally, servers in series, which are called tandem servers.

In Chapter 5, the authors address control of customer routing. The first case involves parallel servers with different service rates. As in Chapter 4, the authors summarize special cases that have analytical solutions and then develop a more general controller and illustrate it through a numerical example. The second case addresses servers with heterogeneous functions, which means that customer classes can be processed by only a subset of the servers due to memory constraints or the location of data. These cases are then merged to consider the combination of different service rates and server functions. Finally, the authors address parallel servers with uncontrolled arrival streams.

For Chapter 6, the authors consider admission control, which is the manipulation of arriving requests. The cost function includes the effect of rejecting requests. The simplest case is a single-server queuing system, followed by multiple servers with a common queue. This case is then extended to two classes of customers. A further extension is to have two queuing systems in series such that a request can be rejected upon arrival at either queuing system.

Having separately addressed actuation for service rates, routing, and admission control, Chapter 7 considers how to use these tasks in combination. The first case considered is that of two queuing systems in series. The control variables are the service rate used and the customer class selected to enter the server. Next, the authors add considerations for admission control by introducing a second arrival stream, eventually considering both controlled and uncontrolled arrivals.

Finally, Chapter 8 applies the techniques developed in Chapters 4–7 to several Internet control problems. A brief overview of the Internet is given from a control perspective. Then, the authors consider differentiated service, that is, preferential treatment of certain kinds of network traffic such as voice traffic. A technique for enforcing differentiated service is to drop or delay nonpreferred packets. The authors also consider the problem of congestion control. Congestion is a result of competition for routers and communication links as a result of interactions between hosts. Control is exercised in routers by dropping packets when buffers become overly utilized. Considered last is the enforcement of bounded delays through circuit-switched networks. The challenge here is to establish circuits that comply with the quality of service constraints, which is essentially a problem of routing-based control.

Contributions of the Book

In summary, this book makes a much-needed connection between fuzzy control and plants that can readily be modeled as queuing systems, such as computing systems and data networks. The strengths of the book are its systematic approach to structuring control problems and its emphasis on the construction of fuzzy models for various queuing systems. The main weakness of the book is that it includes relatively few real-world examples. Nevertheless, the book is unique in linking queuing theory with fuzzy control, making it a valuable reference.

References

[1] E. Lazowska, J. Zahorjan, G. Graham, and K. Sevcik, *Quantitative System Performance*. Englewood Ciffs, NJ: Prentice Hall, 1984.

[2] L. Kleinrock, *Queuing Systems, Volumes I & II*. New York: Wiley, 1976.

[3] D. Menasce, V. Almeida, and L. Dowdy, *Capacity Planning and Performance Modeling*. Englewood Cliffs, NJ: Prentice Hall, 1994.

Joseph L Hellerstein is a research staff member and manager of the Adaptive Systems Department at the IBM T.J. Watson Research Center in Hawthorne, New York, and an adjunct professor at Columbia University in New York City. He received the Ph.D. in computer science from the University of California at Los Angeles. Since then, his research has addressed various aspects of service quality in computing systems, including predictive detection, automated diagnosis, expert systems, and the application of control theory to computing systems. He has authored or coauthored approximately 80 peer-reviewed articles, a book on expert systems (Addison-Wesley), and *Feedback Control of Computing Systems* (Wiley).

Control Systems with Input and Output Constraints by A.H. Glattfelder and W. Schaufelberger, Springer-Verlag, 2003, 499 pp., US$69.95, ISBN 1-85233-387-1. *Reviewed by A.R. Teel and L. Zaccarian.*

Topic of the Book

Constraints are ubiquitous in practical control design problems. Constraints may be associated with limits on the achievable effort that one can expect from an actuator (*input constraints*) or with operational limits on plant variables (*output constraints*) arising from safety, efficiency, or other concerns. In *Control Systems with Input and Output Constraints*, the authors address control problems for linear plants subject to these constraints. Input constraints are treated as *hard* limits, meaning that the model does not allow the inputs to violate these constraints. These constraints arise due to actuator saturation. Output constraints are treated, in general, as *soft* constraints, meaning that the quantity under consideration may, occasionally and at some cost, exceed the prescribed limits. Saturating sensors are not considered.

Control design for linear plants with constraints has been studied extensively in the academic control community. Constraints are challenging since they cannot be addressed using nonlinear control tools such as feedback linearization or backstepping. An extensive list of references up until 1995 can be found in [1]. The problem receives broad, but less formalized, consideration in industry because of its ubiquity. Glattfelder and Schaufelberger bring their considerable industrial and academic expertise regarding this problem to the research and industrial community through their new book. In this review, we summarize the contributions of the work as we perceive them to be. Alternative summaries are given in [3] and [10].

Control with Constraints

A specific problem within the general class of control problems for linear plants with input constraints is the *antiwindup problem*. This problem arises from the requirement that large signal objectives not affect small signal performance. In the antiwindup problem, a particular linear controller is prespecified and must be used when operation does not cause the input limits to be reached. Subject to this requirement, antiwindup augmentation is designed in an attempt to keep the closed-loop system stable when input saturation occurs and to optimize the performance in this case. While input saturation limits large signal performance, prespecifying the small signal controller and restricting the antiwindup architecture limits the achievable large signal performance. Generally speaking, the required sophistication of antiwindup augmentation increases as the phase lag of the linear system in the feedback loop increases. For this reason, controllers that introduce negative phase, such as programmable integral-differential (PID) controllers, tend to be more susceptible to "windup" and require antiwindup augmentation.

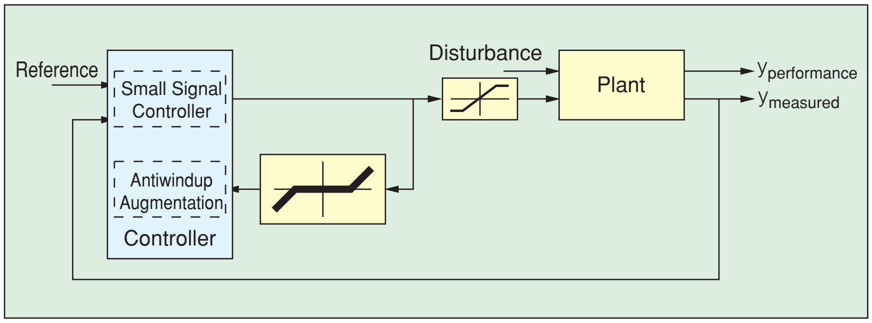


Figure 1. *A simple antiwindup architecture for a control system with input constraints. For small input signals, the output of the deadzone is zero; thus, the antiwindup augmentation is undriven and does not contribute to the control signal. For large signals, the deadzone nonlinearity produces a signal that drives the antiwindup augmentation, which, in turn, affects the control signal.*

Some control problems involving input constraints do not have a small signal controller prespecified. Thus, these input constrained problems are not antiwindup problems as described above. Nevertheless, the ideas behind antiwindup synthesis can still be used to produce an indirect solution to such problems in two steps: 1) design a linear, small signal controller that can be viewed as having been prespecified once this step is completed, and 2) design antiwindup augmentation to achieve large signal performance. Some iteration between the two steps can also be considered. For example, the structure of the small signal controller may be fixed in the form of a PID control, but the gains can be readjusted based on the success or failure of the second step.

The drawback to this indirect approach is that the first step of the design, combined with restrictions on the antiwindup architecture imposed to simplify design, can severely limit achievable large signal performance. However, there are two advantages. First, the initial step is usually not very difficult. Second, when using simple static or low-order antiwindup augmentation within the structure of Figure 1 in the second step, intuitive ideas can often be used to explain reasonable choices for this block. Even with no antiwindup augmentation, the indirect approach guarantees stability and performance when saturation does not occur and also when saturation is not significantly pronounced. Indeed, in the latter case, the values produced by the saturation nonlinearity are not very different from what would be produced by an unconstrained actuator. The task of the antiwindup augmentation block is to extend stability and performance into the region where saturation is severe.

Control problems involving output constraints can also be addressed using a two-step approach either because a small signal controller is prespecified or because taking that viewpoint simplifies the design. Figure 2 shows the associated architecture, which the authors call *override control*, combined with antiwindup control.

Alternatively, direct approaches to feedback control with constraints can be considered. These approaches include model predictive control, nonlinear forwarding techniques (see [5], [6], and [12]) and certain linear matrix inequality (LMI)-based tools that have been developed for direct design of saturating controllers. These approaches do not attempt to replicate a prespecified controller for small signals. Since direct approaches are not encumbered by such requirements, these methods can produce controllers that perform better than indirect approaches for large signals.

Contents of the Book: Part I

In Part I of the book, the authors begin their discussion of control with input constraints by considering designs based on intuitive ideas related to antiwindup synthesis. In Chapter 2, the authors address the antiwindup problem for dominant first-order single-input, single-output (SISO) plants controlled by proportional-integral (PI) controllers. For this problem, several antiwindup constructions are proposed and evaluated in terms of their advantages and limitations for example studies. The book's initial discussion about control with output constraints has a similar flavor. The proposed solutions are based on what the authors call the *control override architecture*. This architecture is similar to the antiwindup architecture of Figure 1 except that the augmentation takes effect when the plant output exceeds the specified output limits. In Figure 2, this type of control override is combined with antiwindup augmentation. As the authors note, control override for output limits has received far less attention in the academic community than the antiwindup problem. In Chapter 3, the authors use a similar approach to that of Chapter 2 by introducing several override techniques for first-order SISO plants under the action of a PI controller and illustrating these techniques through case studies. Subsequently, in Chapter 4, the authors combine the antiwindup and override results introduced in Chapters 2 and 3 and illustrate the unified designs for SISO first-order plants with input and output constraints. Part I of the book ends with Chapter 5, where these approaches are extended to the case of PID controllers with input rate saturation, magnitude saturation, and measurement noise.

The nice feature of Part I of the book is that it is targeted to an industrial audience and is written in a practically oriented style using terminology that is familiar to industrial engineers. First, the authors rely on block diagrams and examples of simulation studies to illustrate the techniques presented. Next, several control schemes are reported and examples are used to illustrate cases in which one of the schemes leads to the best results. In this manner, an industrial engineer can select and try a number of schemes before deciding which one is most suitable for a specific control problem. This approach is effective when trying to convey a message to a wide audience of control engineers educated on linear control

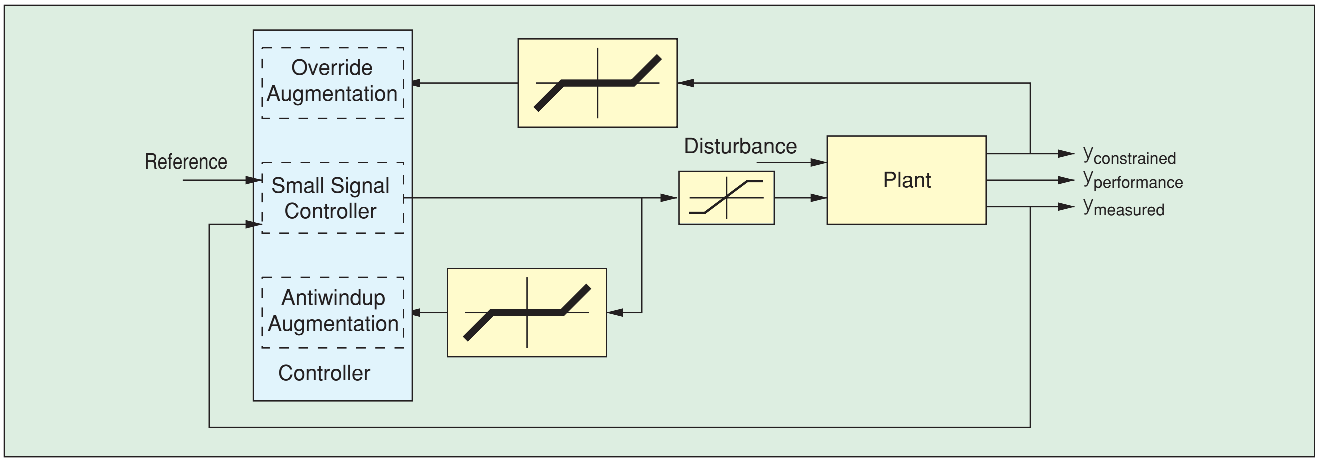


Figure 2. *Antiwindup/override architecture for a control system with input and output constraints. For small input and output signals, the deadzone nonlinearities produce the value zero and thus the augmentation does not contribute to the control signal. For large signals, the deadzone nonlinearities provide signals that drive the augmentation, which in turn affects the control signal.*

loops, Bode diagrams, and Nyquist plots. For analyzing nonlinear systems, the authors provide some exposure to the describing function method and the circle criterion, without subjecting the reader to excessive technical detail.

Contents of the Book: Part II

In Part II of the book, the authors address more general control problems. In Chapter 6, the plant becomes a more complicated (that is, not necessarily dominant first-order) SISO linear system, leading to a wider variety of plant characteristics. Antiwindup-based solutions are developed for the problem of input constraints by generalizing and extending the approaches given in Chapter 2. Direct design approaches, such as model predictive control (MPC), are discussed in this chapter. The structure of Chapter 6 is replicated in Chapter 7 with respect to the override solutions given in Chapter 3 for output-constrained control problems. Also, in Chapter 7, direct design solutions are discussed together with generalized override solutions. Finally, in Chapter 8, the authors provide hints for addressing plants with multiple inputs and outputs.

The motivation for Part II of the text is that the tools described in Part I, which were effective for controlling simple plants with constraints, are often no longer sufficient for characterizing solutions to control problems with more complicated plant dynamics. To preserve the uniformity of the book in terms of style of presentation and target audience, the authors select techniques from the recent literature on advanced direct design and antiwindup design. One technique that is emphasized, and treated quite nicely, is discrete-time model predictive control, which is widely used for sampled-data control of continuous-time plants. On the other hand, some of the other advanced design techniques are treated superficially. For example, the heuristic “nested loops” method, which the authors attribute to [2], bears a striking resemblance to the rigorous nested saturation ideas presented in [9] and [11]. Conversely, the “continued states” approach to antiwindup synthesis, which the authors attribute to [13], bears only a superficial resemblance to the structure used in [13]. In particular, key feedback paths, which greatly reduce the effectiveness of the scheme from [12], are omitted.

There are, of course, many advanced techniques for direct design and antiwindup design that the authors could not include due to space limitations. For other methods, readers can consult the current control literature as well as recent books such as [4] and [8]. We also mention the paper [7], which is written in the same flavor as the book under review but mentions some complementary advanced methods applied to a double integrator with amplitude saturation.

Summary

In summary, the book by Glattfelder and Schaufelberger illustrates designs for input and output saturated systems and antiwindup/override control techniques for typical industrial control systems by adopting a style and notation familiar to the typical industrial engineer. This style has the advantage of targeting a large industrial audience. Its main drawback is that it does not cover in great detail a number of approaches arising from recent research trends that allow more involved control challenges to be addressed. In light of this fact, we believe that the most important contribution of the book is found in Part I, where the adopted tools are adequate for addressing the control problems considered therein. Part II can be seen as an interesting overview of recent advanced techniques for more complicated problem settings. Readers seeking a detailed and comprehensive treatment of approaches to these advanced control problems should, however, refer to papers from the recent literature.

References

[1] D.S. Bernstein and A.N. Michel, “A chronological bibliography on saturating actuators,” *Int. J. Robust Nonlinear Contr.*, vol. 5, pp. 375–380, 1995.

[2] H. Bühler, *Regelkreise mit Begrenzungen*, Fortschritt-Berichte VDI, Nr. 828, Düsseldorf, 2000.

[3] T. Hagglund, “Book review: Control systems with input and output constraints,” *Control Eng. Pract.*, vol. 13, no. 1, p. 131, 2005.

[4] T. Hu and Z. Lin, *Control Systems with Actuator Saturation: Analysis and Design*. Cambridge, MA: Birkhauser, 2001.

[5] M. Jankovic, R. Sepulchre, and P.V. Kokotovic, “Constructive Lyapunov stabilization of nonlinear cascade systems,” *IEEE Trans. Automat. Contr.*, vol. 41, no. 12, pp. 1723–1735, 1996.

[6] F. Mazenc and L. Praly, “Adding integrations, saturated controls, and stabilization for feedforward systems,” *IEEE Trans. Automat. Contr.*, vol. 41, no. 11, pp. 1559–1578, 1996.

[7] V.G. Rao and D.S. Bernstein, “Naive control of the double integrator,” *IEEE Contr. Syst. Mag.*, vol. 21, no. 5, pp. 86–97, 2001.

[8] A. Saberi, A.A. Stoorvogel, and P. Sannuti, *Control of Linear Systems with Regulation and Input Constraints*. New York: Springer-Verlag, 2000.

[9] H.J. Sussmann, E.D. Sontag, and Y. Yang, “A general result on the stabilization of linear systems using bounded controls,” *IEEE Trans. Automat. Contr.*, vol. 39, no. 12, pp. 2411–2425, Dec. 1994.

[10] H.G. Tanner, “Book review: Control Systems with Input and Output Constraints,” *Automatica*, vol. 41, no. 4, pp. 731–733, 2005.

[11] A.R. Teel, “Global stabilization and restricted tracking for multiple integrators with bounded controls,” *Syst. Contr. Lett.*, vol. 18, pp. 165–171, 1992.

[12] A.R. Teel, A nonlinear small gain theorem for the analysis of control systems with saturation,” *IEEE Trans. Automat. Contr.*, vol. 41, no. 9, pp. 1256–1270, 1996.

[13] A.R. Teel and N. Kapoor, “The $\mathcal{L}_2$ antiwindup problem: Its definition and solution,” in *Proc. 4th ECC*, Brussels, Belgium, July 1997.

Andrew R. Teel received his A.B. degree in engineering sciences from Dartmouth College in 1987 and his M.S. and Ph.D. degrees in electrical engineering from the University of California, Berkeley, in 1989 and 1992, respectively. After receiving his Ph.D., he was a postdoctoral fellow at the Ecole des Mines de Paris in Fontainebleau, France. In September 1992, he joined the faculty of the Electrical Engineering Department at the University of Minnesota, where he was an assistant professor until September 1997. In 1997, he joined the faculty of the Electrical and Computer Engineering Department at the University of California, Santa Barbara, where he is currently a professor and director of the Center for Control Engineering and Computation. He has received the National Science Foundation Research Initiation and CAREER Awards, the 1998 IEEE Leon K. Kirchmayer Prize Paper Award, and the 1998 George S. Axelby Outstanding Paper Award. He was the recipient of the first SIAM Control and Systems Theory Prize in 1998. He was also the recipient of the 1999 Donald P. Eckman Award and the 2001 O. Hugo Schuck Best Paper Award, both given by the American Automatic Control Council. He is a Fellow of the IEEE.

Luca Zaccarian graduated with a degree in electronic engineering from the University of Roma Tor Vergata in 1995, where he also received his Ph.D. degree in computer science and control engineering in 2000. In 1998–2000, he spent two years as a visiting researcher at the Center for Control Engineering and Computation of the University of California, Santa Barbara. During the summers of 2003 and 2004, he was a visiting professor at the Electrical and Electronic Engineering Department of the University of Melbourne (Australia). Since 2000, he has been an assistant professor ("ricercatore") in control engineering at the University of Roma Tor Vergata. He was the recipient of the 2001 O. Hugo Schuck Best Paper Award given by the American Automatic Control Council.

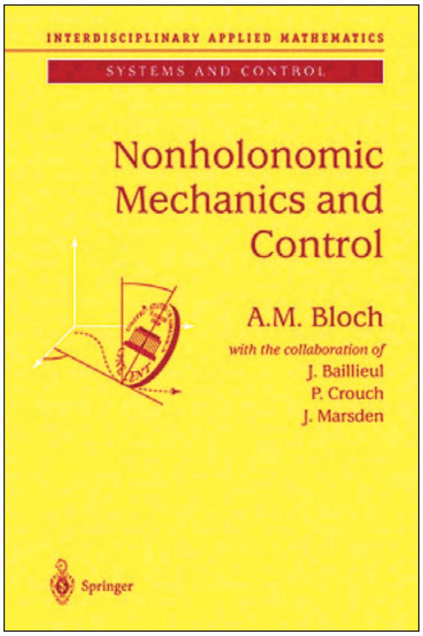
Nonholonomic Mechanics and Control by A.M. Bloch with the collaboration of J. Baillieul, P. Crouch, and J. Marsden. Springer, 2003. ISBN 0-387-95535-6, US$69.95. *Reviewed by A.J. van der Schaft.*

Holonomic and Nonholonomic Constraints

As aptly formulated in its preface, *Nonholonomic Mechanics and Control* links control theory with a geometric view of classical mechanics in both its Lagrangian and Hamiltonian formulations and, in particular, with the theory of mechanical systems subject to kinematic motion constraints. The analysis and control of mechanical systems has been an active research area over the last several decades. The book aims to present some of this material, often scattered throughout the literature, in a cohesive manner.

Kinematic constraints are classically divided into two classes: constraints that can be integrated to yield constraints on the position variables, called *holonomic* constraints, and constraints for which this integration is not possible, called *nonholonomic* constraints. A typical example of a nonholonomic constraint is a wheel rolling vertically *without slipping* on a surface. The constraint on the allowable velocity (the point of contact of the wheel with the surface cannot slip in all directions) cannot be integrated to yield a constraint on the *position* of the wheel. This nonintegrability is intuitively clear, as illustrated by the fact that an automobile can go anywhere it pleases by suitable maneuvering. (This example hints at the intimate relationship between nonholonomic constraints and controllability.) Loosely speaking, mechanical systems with holonomic constraints can be reduced to lower dimensional mechanical systems without constraints; for systems with nonholonomic constraints, this reduction is not possible, and as a result some distinguishing features arise.

Systems with nonholonomic constraints are not only well motivated by applications such as mobile robots but are intellectually interesting as well. From an analytical and geometric point of view, nonholonomic constraints present a class of systems that cannot be described in the standard Lagrangian or Hamiltonian framework; the powerful machinery developed for the standard case must be modified and extended. At the same time, nonholonomic constraints may actually *help* in controlling a system. In particular, for *underactuated* mechanical systems, the presence of nonholonomic constraints in directions that are not directly actuated is helpful for controllability. Interestingly enough, the geometric/analytical and nonlinear control points of view appear to be intimately linked, giving rise to a rich theory that can be illustrated on simple mechanical examples. It is reasonably "easy" to write a difficult and obscure book about this theory, but it is much more difficult to write a book that provides clear access to the topic, which intrinsically has many facets. The author has invested considerable effort in achieving this goal.


INTERDISCIPLINARY APPLIED MATHEMATICS
SYSTEMS AND CONTROL
Nonholonomic Mechanics and Control
A.M. Bloch
with the collaboration of
J. Baillieul
P. Crouch
J. Marsden
Springer

0272-1708/05/$20.00©2005IEEE

Contents

The text begins with four introductory chapters. After introducing Hamilton's principle and the Lagrange- d'Alembert principle, Chapter 1 presents a collection of mechanical systems subject to nonholonomic kinematic constraints. The treatment of these examples illustrates some of the ideas and phenomena discussed in detail in the later chapters. As such, this chapter gives the basic flavor of the material and provides an ideal introduction to the book. Chapters 2, 3, and 4, respectively titled "Mathematical Preliminaries," "Basic Concepts in Geometric Mechanics," and "Introduction to Aspects of Geometric Control Theory," give a brief overview of the mathematical background needed for the rest of the book and introduce key definitions and notation. These chapters can be skipped by readers who are acquainted with this material and can be returned to as needed.

Although, unavoidably, the opening chapters provide only a "crash course" at some points, the material has been written with much care. In fact, in many cases, the clarity of the presentation is unmatched elsewhere in the literature. An example is the treatment of Ehresmann connections in Section 2.9, where the geometric definition and interpretation is linked to the algebraic definition in a very insightful way. Also, the relation with linear (Koszul) connections on tangent bundles is clearly explained. (I would have also linked the definition of a linear connection on a tangent bundle to the geometric interpretation of an Ehresmann connection.) As a result, these introductory chapters provide stimulating reading for a reader well versed in this material as well as for the novice.

A core chapter of the text is Chapter 5, which deals with the geometric formulation of mechanical systems subject to nonholonomic kinematic constraints, both in the Lagrangian and Hamiltonian frameworks. This chapter provides a nice synthesis of the literature, some of which is quite recent. Chapter 6 deals with the control of nonholonomic systems, and the control of mechanical systems subject to nonholonomic constraints. The terminology *nonholonomic (control) systems*, which may be somewhat confusing, originates from the kinematic model of a mechanical system subject to nonholonomic kinematic constraints, for which the feasible space of velocities on the configuration space of the mechanical system is represented as the span of input vector fields. Pure nonholonomicity of the kinematic constraints (in the sense that no part of them can be integrated to yield constraints on the configuration variables) then corresponds to *controllability* of the ensuing control system. Controllability can be characterized by a full-rank condition on the space of vector fields generated by taking (possibly repeated) Lie brackets of the input vector fields. Thus, nonholonomicity is intimately linked to controllability. On the other hand, the linear control paradigm of equivalence between controllability and pole placement breaks down in this case in the sense that there does *not* exist an asymptotically stabilizing continuous state feedback law. The second part of Chapter 6 deals with the control of dynamical models of mechanical systems subject to kinematic constraints and with external input forces.

Chapter 7 describes another interesting connection between nonlinear control theory and nonholonomic constraints, namely, the variational formulation of mechanical systems subject to nonholonomic constraints, as in Hamilton's principle as opposed to the Lagrange-d'Alembert principle. The difference is that, in the first case, the time-integral of the Lagrangian is extremized over all configuration trajectories whose tangent vectors satisfy the kinematic constraints, while in the latter case the conditions for extremality of an arbitrary curve in the configuration space is checked with respect to all tangent vectors (classically called virtual displacements) satisfying the kinematic constraints. This subtlety is not easy to grasp and may seem to be a purely academic discussion (in the bad sense of the word). However, it turns out that the equations of motion obtained from Hamilton's principle are generally different from the "correct" equations of motion obtained by applying the Lagrange-d'Alembert principle; in fact, these equations are the same if and only if the constraints are holonomic!

On the other hand, the framework of Hamilton's principle is linked to nonlinear optimal control theory, while the special case of optimal control of nonholonomic control systems leads to the interesting mathematical topic of sub-Riemannian geometry. Chapter 8 deals with the key problem of stability of mechanical systems with nonholonomic constraints and shows how the delicate stability theory of ordinary Hamiltonian systems can be extended to this case. Finally, Chapter 9 provides a control counterpart by studying the stabilization or setpoint regulation of underactuated mechanical systems. The main method is the theory of controlled Lagrangians. (The inclusion of gyroscopic forces in the Lagrangian description is equivalent to interconnection-damping assignment control of port-Hamiltonian systems.) This method is based on the use of state feedback to shape another *virtual* mechanical system with energy that has a critical point at the desired setpoint. Another method treated is the theory of averaging for underactuated mechanical systems.

Summary

This book is a welcome addition to the existing literature and will certainly become a standard reference. Related books with a different emphasis and partly complementary choice of topics include [1] and the recent [2]. It is to be expected that Bloch's book will be a continuing source of inspiration for further research in this area. Indeed, many aspects remain to be investigated, especially from the systems and control point of view. While important basic understanding has been reached for trajectory planning and classical control problems like setpoint stabilization, the general problem of controlling (by state feedback laws or by interconnection with other dynamical systems) the dynamical behavior of the physical system is still largely open. Important issues here are understanding the robust dynamical behavior of systems with nonholonomic constraints and the exploration of new control paradigms.

In summary, this is a delightful book that will be valuable for both the control community and researchers working on the geometric theory of mechanical systems. With its extensive illustrations and exercises, this book is eminently suited for a graduate course. The author should be congratulated for such an admirable job.

References

[1] J. Cortés Monforte, *Geometric, Control and Numerical Aspects of Nonholonomic Systems* (Lecture Notes in Mathematics 1793). Berlin: Springer-Verlag, 2002.

[2] F. Bullo and A.D. Lewis, *Geometric Control of Mechanical Systems. Modeling, Analysis, and Design for Simple Mechanical Control Systems* (Texts in Applied Mathematics, 49). New York: Springer-Verlag, 2005.

A.J. (Arjan) van der Schaft received his undergraduate and Ph.D. degrees in mathematics from the University of Groningen, The Netherlands, in 1979 and 1983, respectively. In 1982, he joined the Department of Applied Mathematics, University of Twente, Enschede, The Netherlands, where he is presently a full professor in mathematical systems and control theory. His research interests include the mathematical modeling of physical and engineering systems and the control of nonlinear and hybrid systems. He has served as associate editor for *Systems & Control Letters*, *Journal of Nonlinear Science*, *SIAM Journal on Control*, and *IEEE Transactions on Automatic Control*. Currently, he is associate editor for *Systems and Control Letters* and editor-at-large for the *European Journal of Control*. He is coauthor of *System Theoretic Descriptions of Physical Systems* (1984), *Variational and Hamiltonian Control Systems* (1987, with P.E. Crouch), *Nonlinear Dynamical Control Systems* (1990, with H. Nijmeijer), *L_2-Gain and Passivity Techniques in Nonlinear Control* (2000), and *An Introduction to Hybrid Dynamical Systems* (2000, with J.M. Schumacher). He is a Fellow of the IEEE.

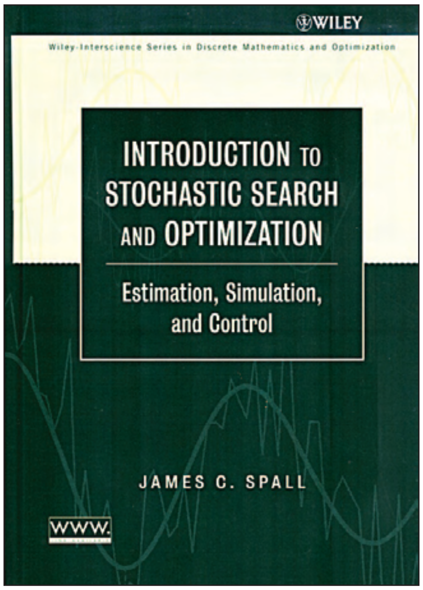
Introduction to Stochastic Search and Optimization by James C. Spall, Wiley, Hoboken, NJ, 2003, ISBN 0-471-33052-3, US$105.00. *Reviewed by G. Yin.*

Search and Optimization Methods

James Spall's book provides a survey of random search (called stochastic search by the author) and optimization methods, including stochastic approximation algorithms, evolutionary computation, Monte Carlo simulation, and related statistical methods. Stochastic approximation algorithms include recursive least squares, adaptive algorithms, and reinforcement learning.

One of the original aims of stochastic approximation is to find roots of a continuous function $f(\cdot)$, where either the precise form of f is not known or f is too complicated to compute and only noisy measurements are available. That is, at iteration n (often referred to as time n) and design point x_n (often referred to as state x_n), it is possible to obtain only a noisy measurement $y_n = f(x_n) + \xi_n$, where ξ_n denotes random noise. In 1951, Robbins and Monro [2] introduced the recursive algorithm $x_{n+1} = x_n + a_n y_n$, where a_n is an appropriately chosen step size. Robbins and Monro coined the name *stochastic approximation* for this procedure. Another one of the original aims of stochastic approximation is to determine the minimizer of a real-valued cost function h using only noisy measurements of $h(x_n)$. To find local minimizers of a smooth function h, one may try to find the roots of the gradient of h. However, the gradient of h is not available when only noisy measurements of h can be used. In 1952, Kiefer and Wolfowitz [3] proposed another recursive algorithm for solving this problem, in which the gradient of h is replaced by its noisy gradient estimate of finite difference type.

Since stochastic approximation methods were first introduced, there has been significant progress in the development of more sophisticated algorithms. Much of this development has originated from, and has been intertwined with, applications in optimization, control theory, economic systems, signal processing, communication theory, learning, pattern classification, neural networks, and related fields. Emerging applications have been found in wireless communications, repeated stochastic games, and financial engineering. For instance, consider a production planning problem with unreliable machines. Under certain conditions, it can be shown that the optimal control is of the threshold type. For multiple machine problems, it is virtually impossible to find closed-form solutions. However, using stochastic approximation methods, the problem can be recast as a parametric optimization problem that can be solved by



recursive stochastic approximation algorithms. Another example concerns stock-selling decisions, an optimal stopping problem that is normally difficult to solve. Focusing on threshold-dependent selling strategies, one can design stochastic approximation algorithms for this problem. In addition, multiuser detection, time-varying parameter tracking, and adaptive interference suppression problems in CDMA/DS networks can also be solved by stochastic approximation methods. Because of its importance, stochastic approximation has drawn continued attention over the past five decades; see [1] and [4] and the references therein for recent developments.

Contents

The book consists of 17 chapters and five appendices. The chapters, as the author points out, can be divided into two main parts: Part 1 (comprised of chapters 1–12) focuses on stochastic search and optimization, whereas Part 2 (comprised of chapters 13–17) studies modeling, simulation, and estimation. Each chapter ends with a conclusion section that provides some historical perspective. Related problems of interest and appropriate references are also given.

Chapter 1 illustrates randomness in the context of stochastic search and optimization and develops a general problem formulation. Chapter 2 studies random search under noiseless and noisy environments. Recursive estimation for linear models is treated in Chapter 3. Here the discussion is mainly on various least-squares-type algorithms. Kalman filters are also studied.

The next five chapters concentrate on stochastic approximation algorithms. Chapter 4 studies convergence, asymptotic normality, iterate averaging, and time-varying functions of stochastic approximation algorithms. Chapter 5 deals with stochastic gradient algorithms with applications to neural networks, discrete-event systems, and image restoration. Chapter 6 is concerned with optimization of an objective function for which the gradient is not available and measurements are corrupted by noise. Simultaneous perturbation for stochastic approximation (SPSA), a topic of expertise of the author, is treated in Chapter 7. The construction of efficient stochastic approximation algorithms for such purposes has drawn much attention because of practical problems in optimization for high-dimensional systems.

Chapter 8 is devoted to global stochastic approximation algorithms. These algorithms are needed for functions that have multiple local optimizers at which standard gradient search iterates can become trapped. Although global stochastic approximation (or annealing) provides an alternative approach, the convergence rate in annealing algorithms is slow.

Chapters 9–11 focus on learning algorithms. Chapter 9 presents genetic algorithms, and Chapter 10 provides general methods and theory for evolutionary computation. Chapter 11 deals with reinforcement learning using temporal differences, a method that has attracted much attention recently. Chapter 12 is concerned with optimization methods for quantized problems, in which the parameters can take only a finite number of values.

The first chapter in the second part of the book, namely, Chapter 13, is on model selection and statistical information, including the tradeoff between bias and variance, cross validation, and the Fisher information criterion. Chapters 14–16 are concerned with Monte Carlo optimization. Chapters 14 and 15 discuss the interface between simulation and optimization in regard to regenerative systems, finite difference gradient estimates, infinitesimal perturbation analysis methods, and the use of simultaneous perturbation stochastic approximation (SPSA) methods. Chapter 16 studies Markov chain Monte Carlo methods. Finally, Chapter 17 discusses the optimal design of experiments, a process that involves choosing input values according to a given criterion.

The five appendices review selected topics in multivariable calculus, matrix theory, statistical inference, probability theory, random number generation, and Markov processes.

Readership

The prerequisite background for reading this book is modest. The book includes many exercises and selected solutions. It can be used either as a graduate textbook or as a reference for people working in the area of optimization and design of experiments. A one-semester course can cover Chapters 1–8 and Chapters 11 and 12. A second course can include Chapters 9, 10, and 13–17.

Conclusions

One of the main features of Spall's book is the description of the use of recursive stochastic approximation algorithms. This book does a good job of comparing various algorithms in the context of stochastic optimization. It is well written and accessible to a wide audience. Although the book presents theoretical issues in algorithms and methods, it does not include detailed derivations. In summary, the book is a welcome addition to the control and optimization community.

References

[1] H.F. Chen, *Stochastic Approximation and Its Applications*. Dordrecht, The Netherlands: Kluwer, 2002.

[2] H. Robbins and S. Monro, "A stochastic approximation method," *Ann. Math. Statist.*, vol. 22, pp. 400–407, 1951.

[3] J. Kiefer and J. Wolfowitz, "Stochastic estimation of the maximum of a regression function," *Ann. Math. Statist.*, vol. 23, pp. 462–466, 1952.

[4] H.J. Kushner and G. Yin, *Stochastic Approximation and Recursive Algorithms and Applications*, 2nd ed. New York: Springer-Verlag, 2003.

G. George Yin joined Wayne State University in 1987 and achieved the rank of professor in 1996. He was an associate editor of *IEEE Transactions on Automatic Control* from 1994–1998. He is an associate editor of *SIAM Journal on Control and Optimization* and several other journals. He is a Fellow of the IEEE.