A Review of Extremum Control

Sternby, Jan

1979

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Sternby, J. (1979). *A Review of Extremum Control*. (Technical Reports TFRT-7161). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

CODEN: LUTFD2/(TFRT-7161)/1-47/(1979)

A REVIEW OF EXTREMUM CONTROL

J. STERNBY

Department of Automatic Control
Lund Institute of Technology
April 1979

Dokumentutgivare
Lund Institute of Technology
Handläggare Dept of Automatic Control
Jan Sternby
Författare
Jan Sternby

Dokumentnamn
REPORT
Utgivningsdatum
April 1979

Dokumentbeteckning
LUTFD2/(TFRT-7161)/1-47/(1979)
Ärendebeteckning
STU 78-3763

Dokumenttitel och undertitel

A review of extremum control

Referat (sammandrag)

A systematic treatment of different schemes for extremum control is given. It is claimed that the progress in recent years in computing technology and areas like identification, optimization and adaptive control motivates a new look at the problem of extremum control. There are also several possible practical applications. Some of these are discussed at the end of this review.

Referat skrivet av
Author
Förslag till ytterligare nyckelord

Klassifikationssystem och -klass(er)

Indextermer (ange källa)

Omfång
47 pages
Övriga bibliografiska uppgifter
Språk
English
Sekretessuppgifter
ISSN
ISBN

Dokumentet kan erhållas från
Department of Automatic Control
Lund Institute of Technology
P O Box 725, S-220 07 LUND 7, Sweden
Pris
Mottagarens uppgifter

DOKUMENTDATABLAD enligt SIS 62 10 12
SIS-DB 1

Blankett LU 11:25 1976–07

SIGMA·TRYCK
TLTH · Lund

Pervozvanskii, A.A.(1960): Continuous extremum control systems in the presence of random noise. Autom.&Remote Control 21,673.

Price, R.J. and Rippin, D.W.T.(1967): Simultaneous two variable on-line optimization of a water gas shift reactor. IEE Conference publication No.29. "Advances in Computer Control", Second U.K.A.C. Control Conventions, Bristol.

Putsillo, V.P., Petrov, V.V., Kornilov, R.V. and Volodin, E.E.(1960): The principles of construction of one class of extremal control systems for automatized circuits in production processes. Proc. First IFAC World Congress, Moscow.

Rajbman, N.S.(1976): The application of identification methods in the USSR-A survey. Automatica 12,73.

Roberts, J.D.(1965): Extremum or hill-climbing regulation: a statistical theory involving lags, disturbances and noise. Proc. IEE 112,137; (1966): Extremum regulation with self-adjustment of the perturbation amplitude. Proc. Third IFAC World Congress, London.

Simpson, R.J. and Power, H.M.(1972): Correlation techniques for the identification of nonlinear systems. Measurement and Control 5,316.

Smirnova, V.I. and Tay, M.L.(1976): Investigation of a class of algorithms with accumulation in extremal systems. Eng. Cybern. 14, No.5,162.

Tovstukha, T.I.(1960): The effect of random noise on the steady-state operation of a step-type extremal system for an object with a parabolic characteristic. Autom.&Remote Control 21,398.

Tsien, H.S. and Serdengecti, S.(1955): Analysis of peak-holding optimalizing control. J. of the aeronautical sciences 22,561.

Vasu, G.(1957): Experiments with optimalizing controls applied to rapid control of engine pressures with high amplitude noise signals. Transactions of the ASME, April 1957,481.

Xirokostas, D.A. and Henderson, J.G.(1966): Extremum control of dynamic systems in the presence of random disturbances and noise. Proc. Third IFAC World Congress, London.

Zotov, V.V. and Sevryukov, B.N.(1972): The extremal problem of controlling a radiotelescope. Radiophysics&Quantum Electronics 15,1213.

Kisiel, A.J. and Rippin, D.W.T.(1965): Adaptive optimization of a water-gas shift reactor. Proc. IFAC Teddington Conf.

Kotnaur, K.D., Box, G.E.P. and Altpeter, R.J.(1966): A discrete predictor controller applied to sinusoidal perturbation adaptive optimization. ISA Trans. 5,255.

Kushner, H.J. and Clark, D.S.(1978): Stochastic approximation methods for constrained and unconstrained systems. Springer, New York.

Leblanc, M.(1922): Sur l'électrification des chemins de fer au moyen de courants alternatifs de fréquence élevée. Revue Générale de l'Electricité.

Leonov, N.N.(1969a): Difference-type relay extremal systems with transport delays. Autom.&Remote Control 30,1780; (1969b): Tracking process in single-channel extremal systems of the self-oscillating type. Autom.&Remote Control 30,2028.

Moran, F., Berger, C.S. and Xirokostas, D.(1965): Development and application of self-optimising control to coal-fired steam-generating plant. Proc. IEE 115,307.

Morosanov, I.S.(1957): Methods of extremum control. Autom.&Remote Control 18,1077.

Narendra, K.S. and Gallman, P.G.(1966): An iterative method for the identification of nonlinear systems using a Hammerstein model. IEEE Trans. AC-11,546.

Olsen, T.O., Berstad, H. and Danielsen, S.(1976): Automatic control of continuous autogeneous grinding. Proc. IFAC Symp. on Automatic Control in Mining, Minerals and Metals Processing, Johannesburg.

Ostrovskii, Iu.I.(1957): Extremum regulation. Autom.&Remote Control 18,900.

Perelman, I.I.(1961): A statistical investigation of extrapolation systems of extremal control used with an object having a parabolic characteristic. Autom.&Remote Control 22, No.11; (1963): A comparison of the simplest gradient and extrapolation systems of extremal control for the case of an object with real static characteristics. Autom.&Remote Control 24,459.

Perret, R. and Rouxel, R.(1963): Principle and application of an extremal computer. Proc. Second IFAC World Congress, Basel.

TABLE OF CONTENTS

Hammond,P.H. and Rees, N.W.(1968): Self-optimizing and adaptive control systems. In McMillan, R.H.(Ed.): Progress in Control Engineering. Heywood.

Hamza, M.H.(1966): Extremum control of continuous systems. IEEE Trans. AC-11,182.

Heaps, H.S. and Wells, R.V.(1965): The effect of noise on process optimization. Can.J.Chem.Engng 42,319.

Hsia, T.C.(1968): Least square method for nonlinear discrete system identification. Proc. Second Asilomar Conf. Circuits and Systems,423; (1976): A multistage least squares method for identifying Hammerstein model nonlinear systems. Proc. IEEE Conf. Dec.&Control,934.

Jacob, H.G.(1972): Hill-climbing controller for plants with any dynamics and rapid drifts. Int.J.Control 16,17.

Jacobs, O.L.R.(1961): A review of self-adjusting systems in automatic control. J. Electronics Control 10,311.

Jacobs, O.L.R. and Langdon, S.M.(1970): An optimal extremal control system. Automatica 6,297.

Jacobs, O.L.R. and Shering, G.C.(1968): Design of a single-input sinusoidal-perturbation extremum-control system. Proc. IEE 115,212.

Jacobs, O.L.R. and Wonham, W.M.(1961): Extremum control in the presence of noise. J. Electronics Control 11,193.

Jelonek, Z.J., Gardiner, A.B. and Raeside, D.(1965): A theoretical comparison of three types of self-optimizing control systems. Proc.Inst.of Mech. Eng. 179,165.

Kazakevich, V.V.(1960): The process of optimizing control of objects with inertia in the presence of disturbances. Proc. First IFAC World Congress, Moscow; (1966): Algorithms for most rapid search in pulsed extremal control systems. Autom.&Remote Control 27,2068.

Kazakevich, V.V. and Mochalov, I.A.(1977): Sequential algorithms of accelerated hill climbing in inertial plants to be optimized. Autom.&Remote Control 38,32.

Keviczky, L. and Haber, R.(1974): Adaptive dual extremum control by Hammerstein model. Proc. IFAC Conf. on Stochastic Control, Budapest.

Keviczky, L., Udvardi, I. and Bányász, Cs.(1976): Optimization of a cement mill by correlation technique. Simulation results. Proc. of AICA Congress, Delft.

Douce, J.L. and Ng, K.C.(1964): Six-channel adaptive computer. Proc. IEE 111,1757.

Draper, C.S. and Li, Y.T.(1951): Principles of optimalizing control systems and an application to the internal combustion engine. American Society of Mechanical Engineers, New York, Sept. 1951, 160pp.

Eveleigh, V.W.(1963): General stability analysis of sinusoidal perturbation extremum searching adaptive systems. Proc. Second IFAC World Congress, Basel.

Feldbaum, A.A.(1959): Steady-state processes in the simplest discrete extremal system with random noise present. Autom.&Remote Control 20,1024; (1960): Dual control theory I-IV. Autom.&Remote Control 21,874,1033 and 22,1,109.

Florentin, J.J.(1964): An approximately optimal extremal regulator. J. Electronics Control 17,211.

Frait, J.S. and Eckman, D.P.(1962): Optimizing control of single input extremum systems. J. Basic Eng. 84,85.

Frey, A.L., Deem, W.B. and Altpeter, R.J.(1966): Stability and optimal gain in extremum-seeking adaptive control of a gas furnace. Proc. Third IFAC World Congress, London.

Fujii, S. and Kanda, N.(1963): An optimalizing control of boiler efficiency. Proc. Second IFAC World Congress, Basel.

Galkin, L.M.(1976): Questions of the analysis and synthesis of sampled-data extremal systems with independent search. Autom.&Remote Control 37,857.

Gallman, P.G.(1975): An iterative method for the identification of non-linear systems using a Uryson model. IEEE Trans. AC-20,771.

Gallman, P.G. and Narendra, K.S.(1976): Representations of nonlinear systems via the Stone-Weierstrass theorem. Automatica 12,619.

Haber, R. and Keviczky, L.(1976): Identification of nonlinear dynamic systems: A survey. Proc. of 4th IFAC Symp. on System Identification, Tbilisi.

Haist, N.D., Chang, F.H.I. and Luus, R.(1973): Nonlinear identification in the presence of correlated noise using a Hammerstein model. IEEE Trans. AC-18,552.

1. INTRODUCTION

In most control problems, the task of the regulator is to keep some variable at a constant value, or to make it follow a reference signal. In general, the system is then assumed to be linear, and it is possible, in principle, to drive the output to any prescribed value. With such problems, the ordinary PID-regulator can often do a good job. In an extremum control problem on the other hand, the static response curve relating the output to the input has at least one extremum point. It is thus a nonlinear dynamical system. The task of an extremum controller is then to keep the output as close to its extremal value as possible.

There are several examples of practical systems that exhibit this type of behaviour. Control of the air/fuel-ratio for optimal combustion has e.g. been studied on many different plants. Usually, the air flow is then controlled to its optimum setting for the current fuel flow. The optimum may vary e.g. with the fuel quality. Autogeneous ore-grinding is another example, where filling degree in the mill is the input and grinding efficiency is the output. For a water-turbine or a windmill with adjustable blade angles, it is desirable to extract maximum power by a proper setting of the blade angles. This is also an extremal control problem. The paper by Leblanc(1922) shows that such problems have been around for a long time. As a matter of fact, Leblanc uses one of the most wellknown methods, which is based on adding a perturbation signal to the input and observing its effect on the output.

Extremum control problems started to become more popular after the publication of the famous paper by Draper and Li(1951). One reason for this was probably improvements in computing technology that made possible the implementation of more and more complicated controllers.

Towards the end of the 5Ø's a couple of commercial optimalizers became available: the Opcon and Quarie controllers. The interest in extremum control seems to have reached a maximum about then and some years thereafter. The number of published papers was higher than ever since, many of them containing optimistic reports of practical applications.

After a few years, however, it turned out that these controllers did not become as successful as expected. One reason might have been their price, around 25.ØØØ US$ for Opcon in 1959. Since then the publication rate has decreased, especially in the western countries. Nevertheless, some research has continued, and concepts like system identification and adaptive control have been introduced into this area.

In the past decades, computer technology has developed enormously. This is one reason why it might be rewarding to reconsider extremum control problems. It is now possible to implement rather complex control algorithms in low cost microcomputers, as has already been shown with adaptive control. It should then be possible to benefit from inserting more ideas from adaptive control and identification into the extremum control area. Moreover, with today's competition for market shares and increasing system complexity, even small gains in efficiency may be very valuable.

This report is by no means intended to be a complete bibliography for the vaste field of extremum control, but rather an introduction to the area. The selection of source papers has been limited to what was easily available to the author. This excludes especially a lot of work published in Russian and most internal technical reports. Furthermore, the presentation is of course biased by the personal opinions and interests of the author. Nevertheless, it is

6. REFERENCES

Amiyan, L.R., Dovlatyan, R.A. and Kazakevich, V.V.(1972): Algorithm for accelerated hill-climbing and its implementation in a digital optimizer. Autom.&Remote Control 33,1155.

Aseltine, J.A., Mancini, A.R. and Sarture, C.W.(1958): A survey of adaptive control systems. I.R.E. Trans. Autom. Control PGAC-6,1Ø2.

Bamberger, W. and Isermann, R.(1978): Adaptive on-line steady-state optimization of slow dynamic processes. Automatica 14,223.

Berånek, J.(1975): Sulphur removal during fluidized bed combustion. Proc. of 5th Int. Congress CHISA '75, Prague.

Bergholz, G.(1966): Zum Entwurf von Steuerungsalgorithmen für ein Extremalsystem mit hochfrequenten Objektstörungen. Messen, steuern, regeln 9,28Ø.

Billings, S.A. and Fakhouri, S.Y.(1978a): Identification of a class of nonlinear systems using correlation analysis. Proc. IEE 125,691; (1978b): Theory of separable processes with applications to the identification of nonlinear systems. Proc. IEE 125,1Ø51.

Blackman, P.F.(1962): Extremum-seeking regulators. In Westcott(Ed.): An exposition of adaptive control. Pergamon Press.

Boehringer, A.F.(1968): Self-adapting dc converter for solar spacecraft power supply. IEEE Trans. AES-4,1Ø2; (1969): Die selbsttätige Einstellung der Extremwerte von Funktionen der Form $z=g(x^a \cdot y)$ nach einem Kennlinienverfahren. Regelungstechnik 17,544.

Chang, F.H.I. and Luus, R.(1971): A noniterative method for identification using Hammerstein model. IEEE Trans. AC-16,464.

Clarke, D.W. and Godfrey, K.R.(1966): Simultaneous estimation of first and second derivatives of a cost function. Electron.letters 2,338; (1967): Simulation study of a two-derivative hill-climber. Electron. letters 3,261.

Douce, J.L. and Bond, A.D.(1963): The development and performance of a self-optimizing system. Proc. IEE 11Ø,619.

control forward using ideas from these neighbouring areas. Hopefully this survey can help promoting such a progress by presenting the status of extremum control to researchers of these other fields.

hoped that the basic ideas have been accounted for, with brief discussions of different modifications and a reference list of moderate length.

Several survey papers of different kinds have already been published. General surveys of adaptive and self-optimizing control systems that also include extremum control are given by e.g. Aseltine, Mancini and Sarture(1958), Jacobs(1961) and Hammond/Rees(1968). More specialized surveys of extremum control systems are e.g. Morosanov(1957), Ostrovskii(1957) and Blackman(1962). Several basic principles were discussed in detail already by Draper/Li(1951).

The rest of this survey will be organized as follows. In section 2 different models will be discussed. Section 3 is a systematic treatment of proposed schemes for extremum control. A collection of possible practical applications of the theory is discussed in section 4. Most of these have been tried in practise, and the results are described in the existing literature. Finally, a couple of concluding remarks are given in section 5.

In order to just get a feel for existing methods of extremum control it suffices to read parts of this report. Each subsection of section 3 is devoted to one type of method and begins with a short description of the basic principle. Reading these introductory parts and possibly the whole of section 2 gives a quick overview. A complimentary picture can be given by e.g. some of the following basic references.

The survey by Blackman(1962) is a short and well written introduction. The so called perturbation method is used in the application papers by Vasu(1957) and Kisiel/Rippin(1965). The paper by Tsien/Serdengecti(1955) contains a detailed, but interesting analysis of a peakholder. A method for handling dynamics in stepping systems was introduced by Kazakevich(1960). Bamberger/Isermann(1978) give an example of the possibilities of applying modern control theory to extremum control systems.

5. CONCLUDING REMARKS

For some reason most of the research on extremum control has been done in Russia and eastern Europe. It can be mentioned, that out of the papers studied for this report, counting only the ones available in translation, almost 2/3 are from these countries. Most of this work has been published in 'Automation and Remote Control', 'Cybernetics', or the German journal 'Messen, steuern, regeln' with a few papers in the IEEE Transactions on Automatic Control. The early IFAC world conferences are also good sources for further references.

Although quite a few practical applications have been reported, in particular with the perturbation method, most of these have concerned pilot plants or laboratory processes. The field of extremum control still needs further development in order to make the technique easy to apply and well suited for routine use in commercial processes. It is believed that the prerequisites for such a development are now at hand. This has been a main reason for carrying out this survey.

First of all, there has been and is a rapid progress in computer technology with powerful microprocessors now appearing at very low cost. It is even becoming economically feasible to replace ordinary analogue PID-controllers by digital versions implemented in microprocessors. This also adds to the possible flexibility of the controller. The increased computing capacity could then instead be used to implement more complicated control algorithms, such as e.g. extremum controllers.

Secondly, the theory development in disciplines like optimization, identification and adaptive control has been substantial. It should then be possible to bring extremum

Still more applications have been suggested. Examples are control of blade angles in water turbines or wind mills for power generation, and control of distillation columns to yield maximum production. An interesting environmental problem is the removal of sulphur dioxide from the flue gas of a fluidized bed combustor, see Beránek(1975). This can be done by feeding certain additive particles into the bed. To keep down the cost of additive particles, it is then desirable to solve the extremum control problem of controlling the combustion temperature to minimize the contents of sulphur dioxide in the flue gas.

2. MODELS

As already mentioned, extremum control systems have one major characteristic in common. In the absense of disturbances, the steady-state relation between inputs and output should be a function with an extremum point. The object of control is to stay as close to this extremum as possible despite the influence from dynamics, noise or drifts. There is thus just one output being optimized, but several inputs may be used, although most of the discussion will be restricted to single input systems.

The problem of tuning a regulator for a linear system by minimization of a nonlinear criterion will not be considered in this report, even though such problems may have the above characteristic. There are several reasons for this. For one thing, there are many other methods for tuning regulators, like e.g. stochastic adaptive or model reference methods. It would lead too far to cover all these procedures as well in a single report. Furthermore, the extremum control problems treated here will be assumed to have unknown nonlinearities, whereas a nonlinear criterion specified by the designer is of course known to him. This knowledge should then be used in the design. Another special feature of the regulator tuning problem for linear systems is that the basic control loop is linear, but an "artificial" nonlinearity is added in an outer loop. This is in contrast with the extremum systems considered in this report, where the nonlinearity is assumed to be inherent in the system to be controlled.

2.1 Static Systems

A common assumption in the literature is that there is no dynamics in the system. In practise, this condition can be fulfilled by using a sufficiently large sampling interval. But the result may be a slow optimization. In many cases, however, static models may be adequate. One example is the control of machine tools, where one problem is to compromize between production rate and tool wear. This can be achieved by optimizing a complicated but static loss function.

The problem of optimizing a function using noisy measurements can be handled with the stochastic approximation method. A lot of research has been devoted to this subject for a long time. Kushner(1978) gives an account for some of the latest developments in the area together with further references. However, it is not clear how to handle dynamical systems using the method. This is taken as an excuse for limiting the survey by not covering the extensive literature on stochastic approximation.

2.2 Dynamics

More difficult problems arise when dynamics have to be taken into account. It is e.g. not self-evident how to include dynamics in the model. It may happen that the dynamics of the nonlinear system to be controlled is fast compared to actuators and measuring devices. A reasonable model may then be linear dynamics at the output and input of a static nonlinearity.

Gallman/Narendra(1976) consider general nonlinear systems. Based on approximation theory they discuss some series expansion representations of the output, which are valid in a closed interval of time [0,T]. Their presentation

first authors maximized the amount of carbon monoxide converted, and used the steam as control variable. Price/Rippin extended the previous work to include the temperature as a second control variable.

Further applications from different areas are the previously mentioned solar cell optimization by Boehringer(1968) and the adjustment of a radio telescope antenna to maximize the signal received from a moving object, see Zotov/Sevryukov(1972). Bamberger/Isermann(1978) considered the optimization of total power from a steam turbine by controlling the cooling water pump. Table 1 gives an account for what extremum control methods have been used in these applications. As seen from Table 1, no application using self-driving systems has been found. Also, reports of practical work with model-oriented methods are rare, indicating that more research is needed in that area.

Application Area	Authors	Method P	C	S	M
Combustion processes	Draper/Li(1951)		x		
	Vasu(1957)	x			
	Fujii/Kanda(1963)		x		
	Moran et al(1965)	x		x	
	Frey et al(1966)	x			
Chemical proc.	Kisiel/Rippin(1965)	x			
-"-	Price/Rippin(1967)	x			
Solar cell	Boehringer(1968)		x		
Antenna adj.	Zotov/Sevryukov(1972)			x	
Grinding	Keviczky et al(1976)	x			x
-"-	Olsen et al(1976)			x	
Turbine power	Bamberger/Isermann(1978)				x

<u>Table 1</u> - Extremum control methods in the applications.
Methods: P - Perturbation, C - Continuous sweep, S - Stepping, M - Model oriented

4. APPLICATIONS

Quite a few practical applications of extremum control algorithms have been reported in the literature. Combustion processes seem to have been a major concern in earlier work, but later on several other problem areas have been entered. A selection of tested or suggested applications are listed below in order to give a general feel for the wide range of possible applications.

The most common way to optimize a combustion process is to control the air/fuel-ratio through the air flow. Using different measured variables this has been tried by e.g. Draper/Li(1951) for an internal combustion engine, Fujii/Kanda(1963) and Moran et al(1965) for a steam generating plant and Frey et al(1966) in a gas furnace. Draper/Li also varied the ignition timing. The two control variables were alternatively switched to the peakholding regulator. Vasu(1957) varied the fuel flow in a flight propulsion system to maximize a certain pressure indicating performance. Several practical experiments were undertaken to find out the influence on the performance of several design parameters in a perturbation scheme.

In certain grinding mills the grinding efficiency will vary with the filling degree of the mill, which can be controlled through the incoming flow of raw material. The optimal point in maximizing efficiency may depend on the quality and composition of this raw material. This type of application was reported by Keviczky et al(1976) for a cement mill and Olsen et al(1976) for autogeneous ore grinding.

Kisiel/Rippin(1965) and Price/Rippin(1967) considered the water-gas shift reactor, where hydrogene and carbon dioxide is produced from carbon monoxide and steam. The

include the Volterra, Wiener and Uryson series. A Volterra series with separable kernels arises as a nonlinear combination of outputs from different linear systems driven by the common input according to the following equation.

$$y(t) = h_o + \sum_{i=1}^{N_1} \int_0^t h_{1,i}(\tau)\cdot u(t-\tau)d\tau \; +$$

$$+ \sum_{i=1}^{N_2} \sum_{j=1}^{N_3} \int_0^t \int_0^t h_{2,1,i}(\tau_1)\cdot h_{2,2,j}(\tau_2)\cdot u(t-\tau_1)\cdot u(t-\tau_2)d\tau_1 d\tau_2 +$$

$$+ \ldots\ldots \qquad (2.1)$$

Introducing intermediate signals v as outputs from certain linear systems as e.g.

$$v_{21}(t) = \int_0^t \sum_{i=1}^{N_2} h_{2,1,i}(\tau_1)\cdot u(t-\tau_1)d\tau_1 \qquad (2.2)$$

it is possible to rewrite (2.1) as

$$y(t) = h_o + v_1(t) + v_{21}(t)\cdot v_{22}(t) + \ldots.. \qquad (2.3)$$

In the simplest case all the v:s are proportional to $v_1(t)$ and the system can be described as a linear system followed by a polynomial nonlinearity. Such a model is often called a Wiener model.

The Wiener series is related to the Volterra series. Restricting the kernels of the intermediate signals v to be orthonormal Laguerre polynomials, the Wiener series is

$$y(t) = a + \sum_{i=0}^{N_1} \sum_{j=0}^{N_2} a_{ij} H_i[v_j(t)] +$$

$$+ \sum_i \sum_j \sum_k \sum_l a_{ijkl} H_i[v_j(t)] H_k[v_l(t)] + \ldots\ldots \qquad (2.4)$$

where H_i are orthonormal Hermite polynomials.

The Uryson series, finally, is a direct generalization of the Volterra series. It is obtained by placing a polynomial nonlinearity P at the input of each linear subsystem, thus replacing (2.2) by

$$w_{21}(t) = \int_0^t \sum_{i=0}^{N_2} h_{2,1,i}(\tau_1) \cdot P_{2,1,i}[u(t-\tau_1)] d\tau_1 \qquad (2.5)$$

Exchanging v for w, (2.3) is still valid. The so called Hammerstein model is a special case of the Uryson series. It is obtained by truncating this series already after the first order terms. The model is then still nonlinear, since the retained terms contain polynomial nonlinearities. The result is a model with a nonlinearity at the input but not at the output.

The approximation error of the output might be smaller the more terms that are included in the above series expansions. However, for practical reasons, it is necessary to use only a limited number of terms. There may then be essential differences between the representations in the ability to describe the true nonlinear system.

The three expansions have one thing in common: linear dynamics are followed by a nonlinearity, at least if enough terms are included. On the other hand, only the Uryson series include nonlinearities at the input. It may be that the output nonlinearity is in general more important than the input nonlinearity for a good description of a nonlinear

But such comparisons do not give an overall picture. With the large number of existing methods for extremum control it would be expected (and wanted) to find several papers comparing different schemes under shifting circumstances. A few algorithms for static optimization with noisy measurements were compared by Heaps/Wells(1965). Jelonek et al(1965) also used a static, noisy system to evaluate the performance of three extremum-seeking regulators. But no complete comparison of all kinds of methods has been found.

all be made recursive.

Billings/Fakhouri(1978a) consider a more general system with linear dynamics at both input and output of the static nonlinearity. A correlation method is used to find the linear dynamics. The parameters of the nonlinearity can then be found by least squares identification.

An interesting test of structure is developed in Billings/Fakhouri(1978b). It is based on correlating the input u and its square u^2 with the output. Depending on the relation between the correlation functions, it can be determined if the system is linear, Wiener type, Hammerstein type or none of these. Rajbman has developed another test of structure employing the so called dispersion functions. They make it possible to define and check a degree of nonlinearity. This technique is described and further references are given in Rajbman(1976).

In all of the papers mentioned above it is assumed that the input is white Gaussian noise, and this is in some cases of importance for the results to hold. This might be a restriction when using the schemes as part of an extremum controller.

3.5 Comparisons

Many of the papers describing individual methods contain a comparison between the suggested algorithm and some other scheme. In e.g. Moran et al(1965) their improved stepping method is compared to an ordinary perturbation method.

system. For the case of a single nonlinearity connected to a linear system this is intuitively clear.The only possible effect of a known nonlinearity at the input is then to restrict the possible input values for the linear part. The nonlinear control problem can then be transformed to linear control with positive inputs. If the range of the nonlinearity is the whole of the real axis, then a change of control variable will reduce the problem to a linear one.

On the other hand, some restrictions must be imposed on models for extremum control systems. Consider e.g. a system where the output is generated as

$$y = -v_1 v_2 \tag{2.6}$$

where v_1 and v_2 are the outputs of two linear systems with positive static gains. An extremum-seeking regulator might then be used to maximize the average of the output y. But if a frequency exists such that the phase difference is 180^o between the two linear systems, then no maximum will exist. The sign of the product is reversed when a sine wave of this frequency is applied.

To illustrate further the difference between input and output nonlinearities, the following example is given.

Example: Consider the linear system

$$y(t+1) = ay(t) + u(t) + e(t) \tag{2.7}$$

where $e(\cdot)$ is a white noise process, and the static nonlinearity

$$y = u^2 \tag{2.8}$$

The nonlinearity can be placed either at the input or at the output of the linear system. In both cases the object of control is to minimize the expected value of the output. Consider first the case of an input nonlinearity.

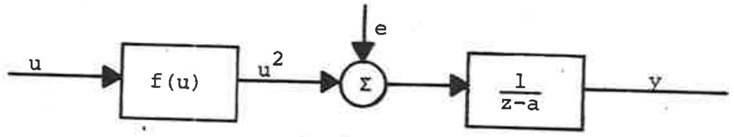

Fig.1 - Example system with input nonlinearity.

The overall system is then

$$y(t+1) = ay(t) + u(t)^2 + e(t) \tag{2.9}$$

Suppose a stationary solution exists ($|a| < 1$). Expected values then are

$$Ey(t) = a\cdot Ey(t-1) + Eu(t-1)^2$$

or

$$Ey = \frac{Eu^2}{1-a} \tag{2.10}$$

The best performance is thus achieved by putting $u(t) = 0$! Furthermore, if $|a|>1$ no stationary solution exists.

Now turn to the other case. The equations are

$$x(t+1) = ax(t) + u(t) + e(t) \tag{2.11}$$

$$y(t) = x(t)^2 \tag{2.12}$$

For a=1 this is the problem considered by Jacobs/Langdon(1970). They show that because of the nonlinear measurement this is a dual control problem in the sense of Feldbaum(1960). The conditional distribution of the state x is discrete, the possible values being $x = \pm|x|$. The conditional mean of x can then be calculated. It is shown that it is not optimal in the long run to have $u(t) = -\hat{x}(t)$. These results would probably not change much if $a = 1-\varepsilon < 1$. Even if a is slightly greater than one, a stationary solution

Bamberger/Isermann(1978) developed a program package employing a gradient method for optimizing a Hammerstein model. The parameters of both linear and nonlinear parts can be identified using either the instrumental variable or correlation methods. In the latter case, the final scheme is closely related to that of Clarke/Godfrey, but with parameters that are independent of the working point. A successful application to power optimization of a turbine is also reported.

Identification

Model identification is an important part of these model oriented methods. An increasing interest in the identification problem for certain nonlinear systems has been noted in recent years. A survey of this area was given by Haber/Keviczky(1976). Some material can also be found in the survey by Rajbman(1976) on identification in Russia. The correlation technique, which seems to be quite useful for nonlinear identification, has been reviewed by e.g. Simpson/Power(1972).

Most of the work has been done for Hammerstein models, starting with Narendra/Gallman(1966). They suggested an iterative method, which alternately updates the nonlinear and linear parts by ordinary least squares and output error least squares respectively. This method is thus not intended for recursive identification. An extended version of this method was used by Gallman(1975) for a more general model with different dynamics for different parts of the nonlinearity.

Some variants of equation error least squares identification of both linear and nonlinear parts of the system have been discussed by Hsia(1968,1976), Chang/Luus(1971) and Haist et al(1973). These methods can

Clarke/Godfrey(1966,1967) estimate the slope and curvature by correlating a 3-level test signal u and its square u^2 with the output. Output dynamics with finite memory will then not influence the result, and for a quadratic nonlinearity the optimum can be reached in one step. It is however necessary to ascertain that the estimate of the second derivative does not become too small. This can be done e.g. using a fixed limit or a first order filter on the estimate.

Roberts(1966) seems to be the first one to suggest a scheme of the second type, where more and more information is gathered about the system. He considers a static system, but includes noise and drift in the model. Several parameters are unknown, including the curvature, position of the optimum, noise level and drift parameters. It is shown that even for known parameters a perturbation signal is needed to follow horizontal drift of the extremum. An optimal perturbation amplitude can be chosen to minimize the mean square deviation from the extremum. When the parameter estimates are correct, a number of signals will have zero mean value. The deviations from zero of these mean values are used to drive the parameter estimates. The input is chosen as the estimated position of the optimum with a superimposed perturbation signal.

Keviczky/Haber(1974) exploited the idea of self-tuning extremum control. They suggested least squares or stochastic approximation identification to find the parameters of a Hammerstein model. The input was then chosen at each step as if the parameter estimates were correct. With this method parameter drift can be handled by a simple modification of the estimation algorithm.

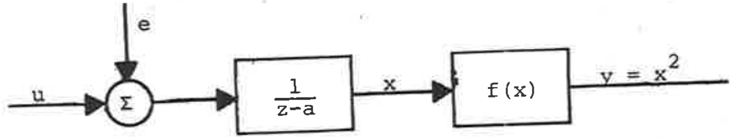

Fig.2 - Example system with output nonlinearity

still seems possible.

There are thus significant differences between the two cases in spite of their identical static response curves. The solution to the second problem includes feedback, and is therefore more attractive although it is more difficult to calculate. #

The second configuration of the example would, of course, be much easier to control if the intermediate signal x were measurable. One possibility would then be to use e.g. minimum variance control of x around the extremum point.

Most of the control algorithms described in the literature have been derived for the static case. Much work has been done to analyse the effect of dynamics on such algorithms. Their behaviour can often be improved by slight modifications of the algorithms to compensate for the dynamics. In an absolute majority of these studies the nonlinearity has been applied at the input, giving a so called Hammerstein model. The linear part is frequently of first order with a known time constant.

Only very few papers discuss what happens when there is an output nonlinearity. In some of those papers it is assumed that the intermediate signal is measured. Others assume that it can be reconstructed because no disturbances enter between the input and the intermediate signal. In such cases the problems with an output nonlinearity are circumvented. But more research is needed to find out how to handle systems where the intermediate signal is not available.

2.3 Noise and Drift

It may be important in practical systems to take noise and drift into account when designing a regulator. Noise is then usually modelled as white and additive, and is applied at the system output as measurement noise. Other possibilities are to apply it in between the linear and nonlinear parts, or at the input.

It is important to note that noise at the input of the nonlinearity is equivalent to a horizontal drift of the nonlinearity. This gives a difficult control problem, which is dual in the sense of Feldbaum. The control signal should thus not be entirely determined to reach the current estimate of the optimum point. It should also assure that future estimates of the optimum are accurate. Jacobs/Langdon(1970) gave a simple example demonstrating this fact. As shown by Florentin(1964) not only control but also estimation is complicated by the presence of horizontal drift. Roberts(1965) showed that a perturbation signal at the input is required to follow the moving optimum.

Usually, both horizontal and vertical drifts are modelled as first order dynamics driven by white noise. This gives a possibility for tracking the drift, and also seems to be a more realistic model than pure white noise. The nonlinear part may of course contain other parameters than the horizontal and vertical positions. It can e.g. be specified by the three coefficients of a second order polynomial. The linear parts may also contain certain

identifyability of the parameters, it may be necessary to superimpose a perturbation on the control signal in this scheme also.

With a low noise level, the first and second derivatives of a static nonlinearity can be determined approximately using only two search steps. This is the essence of the control law suggested by Perelman(1961,1963)

$$\Delta u = -\frac{1}{2}\cdot\frac{[y(u-a) - y(u+a)]\cdot a}{[y(u+a) + y(u-a) - 2y(u)]} \tag{3.11}$$

where $y(u)$ is the output with input u. Such a scheme was included in the comparison performed by Jelonek et al(1965).

The same idea was elaborated further by Jacob(1972), who included known dynamics before and after the nonlinearity. In the absense of noise, the input and output of the nonlinearity can then be tracked exactly through the dynamics, and a control law similar to (3.11) could be used. Linear or exponential drifts of the extremum can be detected and compensated for. Because of the dynamics, a resting period is introduced, so that each new cycle may start from the steady state.

Higher noise levels can be tolerated if several measurements are made to determine the next control action. Least squares identification is used by Bergholz(1966) to find the parameters of

$$\tilde{y} = \alpha\tilde{u} + \beta\tilde{u}^2 \tag{3.12}$$

where $\tilde{y}$ and $\tilde{u}$ denote deviations from the mean values(within one cycle). The input u must then be perturbed, either deliberately or by noise. The optimum is characterized by $\hat{\alpha}=0$, and is approached by making input changes proportional to $\hat{\alpha}$.

$$\dot{u} = \sqrt{|k \cdot \dot{y}|} \tag{3.10}$$

This modified algorithm was tested on two simulated examples and was found to work well.

3.4 Model Oriented Methods

In the schemes discussed sofar, little information is collected about the system. Only the output, and maybe the slope of the nonlinearity at the current working point are used. Essentially no information is saved for later use. For the methods treated in this section, the control action is calculated from a model obtained by some kind of system identification. The position of the extremum may e.g. be one parameter in the model. The input may then be chosen as the estimated extremum position. In the simplest case the estimation may reduce to the determination of a single parameter from a couple of noise-free measurements.

There are two main groups within this class of methods. For methods in the first group, each control action is preceded by an identification phase. During this phase, the input must be varied deliberately or by noise to produce good parameter estimates. Based on the estimates a control step is then taken, and the cycle is repeated. With this type of scheme, the parameters identified are often allowed to depend on the current working point, as e.g. slope and curvature. Little information, if any, is therefore exchanged between cycles.

For the second group of methods no separate identification phase exists. The parameters are continuously updated, and control steps are taken based on the current estimates. Since more old data are saved in the estimates, this method will be better only if the model parameters do not change very much with changing working points. To ensure

parameters. Drift in these parameters can then be introduced in the same way as above.

Most existing control algorithms are primarily designed for deterministic sytems. System noise is then usually handled by analysing its effect on the closed loop system. Most often this analysis is done by simulation, although some theoretical results have been reported. One way to reduce the effects of noise is of course filtering, which has been found useful and necessary in several schemes.

3. CLASSIFICATION OF ALGORITHMS

Surprisingly few new ideas for extremum control have emerged since the 60's. Most of the work has been concerned with analysing the behaviour of known algorithms or slight modifications. Different difficulties are then considered like e.g. measurement noise, input or output dynamics or drift. This is why the old survey paper by Blackman(1962) can still be recommended as a very good introduction to the field. The classification used in this report will follow Blackman's, even though newer modifications will of course also be reviewed.

The first type of systems to be discussed are perturbation systems. The effect at the output from a known signal added to the input is then used to derive information about the slope of the nonlinearity. In a so called switching system the input is driven at a constant speed until the extremum is passed. The direction of input drift is then reversed according to some fixed rule. Self-driving systems use no preset changes in the input. The measurements are used directly to determine the input.

There is also a fourth class of methods that is not described by Blackman, and seems to have been developed later on. It is based on using a parameterized model in combining parameter identification and extremum control.

Morosanov(1957) has given a separate classification. He also supplies rules of thumb for when to use different methods, and shows how to perform certain design calculations.

The first derivative of the output could then be used to drive the input via an integrator so that

$$u(t) = \int^{t} \dot{y}(t)dt \qquad (3.7)$$

This system would have to be started manually, since $\dot{y}=\dot{u}=0$ is always a stationary point. But if started in the correct direction with $\dot{u}\neq 0$ it will find a point where $f'(u)=0$.

Blackman(1962) discusses several problems with this type of system. As described above, it will e.g. continue in the same direction until $\dot{y}=0$ and then stop. So if started in the wrong direction it will continue. This problem can be handled by measuring $\dot{u}$ as well. Then $f'(u)=\dot{y}/\dot{u}$ can be used in the control law instead of just $\dot{y}$. Dynamics will introduce further problems. As explained by Blackman the system may then stick at other points on the curve $y=f(u)$.

Self-driving systems seem to have been paid very little attention to in the literature. Only the paper by Frait/Eckman(1962) will be mentioned here. They compensate for the dynamics by taking the measured input through a filter to get the signal u^*. This filter should be a good guess of the system dynamics, and a possible control law is then

$$\dot{u} = k\cdot\dot{y}/\dot{u}^* \qquad (3.8)$$

However, to avoid the use of an accurate and therefore expensive divider, Frait/Eckman suggested a modified control law. The sign of $\dot{u}$ was taken from (3.8) according to

$$\text{sign}(\dot{u}) = \text{sign}(\dot{y}\cdot\dot{u}^*) \qquad (3.9)$$

In calculating a proper amplitude of $\dot{u}$, (3.8) was used with $\dot{u}^*=\dot{u}$ to give to show that (3.5) may give significantly better control than (3.3).

Further improvements can be gained by using a variable steplength. Xirokostas/Henderson(1966) use the expression within brackets in (3.5) as an estimate of the slope of the nonlinearity. This estimate is used to choose one out of two or three predetermined steplengths. The argument is that close to the optimum, where the slope will be small, a small steplength should be used to increase the accuracy.

Galkin(1976) analysed the effects of input dynamics in a noise-free system with a constant minimum. The control law used is based essentially on (3.3), but with a threshold k against switching

$$\Delta u_{n+1} = -\Delta u_n \text{sign}(\Delta y_n - k) \tag{3.6}$$

It is examined how the design parameters should be chosen to avoid extra switching due to the dynamics.

3.3 Self-driving Systems

The previously discussed methods employ some form of forced input changes, like a perturbation signal or a predetermined rate of input change. In a self-driving system no such restrictions are imposed on the control signal. Instead, at every instant the available information is used to produce a control signal that will drive the system towards an optimum. Consider once more the static system

$$y = f(u)$$

3.1 Perturbation Methods

Already in 1922 Leblanc suggested an application of a perturbation scheme. This may then be the oldest extremum control method, and has also been quite popular. Several applications have been proposed, see e.g. Vasu(1957), Kisiel/Rippin(1965) or Frey et al(1966).

The task of an extremum controller is to keep the gradient of the nonlinearity at zero. The problem is thus reduced to an ordinary control problem if the gradient is measured. This can most often not be done directly. A perturbation method may then provide the necessary information. The basic idea is to add a periodic test signal to the control signal, and observe its effect at the output. This is illustrated in figure 3 for a static nonlinearity. The output and the test signal can e.g. be multiplied and averaged over a number of full periods. The resulting signal is then taken as a substitute for the true gradient, and may e.g. be used in an integral controller as the measured signal that should be kept close to zero.

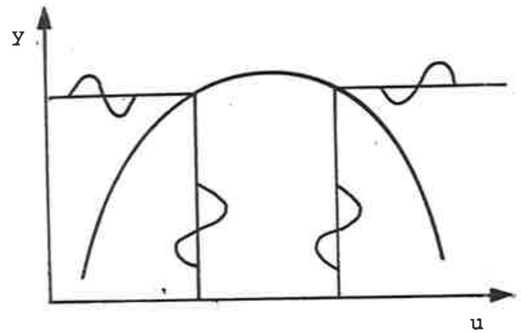

Fig.3 - Effect of an input test signal at the output of a static nonlinearity

Modifications

Dynamics. The basic perturbation method (based on correlating the test signal and the output) may have to be modified if the system contains dynamics. The dynamics will then introduce a phase lag θ in the test signal component of the output. The result of correlation will be multiplied by a factor $\cos\theta$. This gives a sign error in the correlation signal if $\theta > 90^{o}$. The overall system may then become unstable. This situation is avoided if a corresponding phase lag is introduced to the test signal before correlation. Such a feature has been found possible and necessary to include in several of the practical applications reported.

Another way to handle the dynamical effects is to use a perturbation signal of sufficiently low frequency. The phase lag θ will then be small, so that the dynamics can be neglected. This may, however, give a long response time for the overall system.

The control law. In most of the schemes treated in the literature, the input is made proportional to the integral of the correlation signal. A possible improvement would be to use more sophisticated control algorithms based on the same measured signal. One step in this direction was reported by Kotnaur et al(1966). They use a discrete time model with prediction of future disturbances. The correlation signal is taken as the measured error, and minimum variance control is used to keep the process (a gas furnace) at its optimum despite the disturbances.

With the perturbation signal technique, the correlating device must be given a certain amount of time to produce an accurate slope signal. During this time the control signal could be kept constant, so that the total input is varied with the test signal only. The system may then be regarded

Dynamics

For a dynamical system the effect of the last input change on the output may be completely hidden in the responses to earlier input changes. Xirokostas and Henderson(1966) found that no control at all may be better than using (3.3), even in the case of a drifting optimum. This basic algorithm thus needs modifying to handle dynamics.

For the case of known all-pole output dynamics Kazakevich(1960) suggested that a sufficient number of measurements should be made for each new input value, so that the steady state output could be predicted. With no disturbances present n+1 measurements would be enough, where n is the dynamical order. The control law (3.3) can then be used with predicted final output values instead of the current ones. The above approach has been extended in later papers to cover cases with a time delay, measurement noise or drift in the extremum, see e.g. Kazakevich(1966), Amiyan et al(1972) or Kazakevich/Mochalov(1977).

A different method was suggested by Xirokostas/Henderson(1966). They consider an unknown nonlinearity with first order output dynamics and measurement noise. The optimum is assumed to drift around both vertically and horizontally. The dynamics are handled by using a weighted sum of old output differences instead of just the last one in the following way

$$\Delta u_{n+1} = \Delta u_n \cdot \text{sign}[w_o \cdot \Delta y_n + w_1 \cdot \Delta y_{n-1} + w_2 \cdot \Delta y_{n-2} + ..] \quad (3.5)$$

It was shown that the effect of first order dynamics can be completely eliminated using only w_o and w_1 with $w_k = 0$ for $k > 1$. The vertical drift of the optimum is then also well compensated for by much the same choice of w, whereas the measurement noise will impair control. Simulations were used

At first sight it may seem obvious that the stepping period should be kept as small as possible in order to speed up the system. But when dynamics are included in the model this may no longer be true. The easiest way to handle dynamics is to simply wait for the steady state between each input change. But as this may result in too slow a system, several other methods have been proposed, and some of them will be discussed below.

The influence of noise

Measurement noise will introduce a risk of stepping in the wrong direction when using the control law (3.3). This will happen if

$$[\text{sign}(\Delta y) =]\ \text{sign}(\Delta f + \Delta e) \neq \text{sign}(\Delta f) \qquad (3.4)$$

The steady state deviations from the optimum will then be increased. The stochastic distribution of the resulting random walk was analysed for different cases by Feldbaum(1959), Tovstukha(1960) and Jacobs/Wonham(1961). They also examined the influence of the steplength on the resulting loss.

Smirnova/Tay(1976) suggested a modified method to handle noisy systems better. Several measurements of the output are made for each input value. After each measurement a decision is taken either to stay and continue measuring or to move in either direction. The decision is based on the relation between the number of measurements made and the number of favourable ones. No further input changes are thus made until the positive or negative effect of the previous move has been established with a reasonable degree of certainty. The same basic idea has been used by Kazakevich/Mochalov(1977) in a more complicated system that can also handle dynamics.

as a sampled data system where the correlating time is the sampling period.

The test signal. The most commonly used test signal form has been the sinusoid. It is relatively easy to generate using analogue technique, and frequency analysis methods are well suited for examining the effects of such a test signal theoretically. But other test signal forms may also be used, as e.g. a square wave. This is especially easy to generate in a digital computer, and was discussed by e.g. Douce/Bond(1963).

Several inputs. The perturbation method seems to be wellsuited for generalization to more than one input. In order to apply a gradient method in the search for an extremum, the partial derivatives of the static response curve with respect to the different inputs are needed. It is possible to obtain this information by using the correlation method above with perturbation signals of separate frequencies for each input.

Price/Rippin(1967) applied this technique to the optimization of a chemical reactor with two inputs. They used sinusoidal test signals, and found the best frequency relation to be 1:1.5. Douce/Ng(1964) built an analogue six-input extremum-seeking computer with square test signals. They used a frequency separation of 1:1.05 between each channel. The frequency difference should not be made too small, since the correlation time must be increased in order to separate the effects from different test signals.

For the particular case of only two inputs another method is possible. Two test signals of the same frequency, but with a 90° phase difference can be used. A phase lag in the output due to e.g. dynamics will then introduce a cross-coupling in the slope signals, see Blackman(1962).

Analysis

As with most control systems, theoretical analysis is a valuable complement to practical experiments in finding out how perturbation systems work. Such analysis has been carried out to study e.g. stability questions, possible periodic solutions and the influence of different design parameters.

A thorough experimental investigation of a specific system was performed by Vasu(1957). He examined a flight-propulsion system, where the fuel flow into an engine was controlled to maximize a certain pressure indicating engine output. This study covered e.g. steady-state and transient performance, the effect of controller settings, filtering, test signal frequency and shape of the static characteristic.

The effects of measurement noise and drift of the extremum point were studied by Pervozvanskii(1960). Linearization was used to arrive at an expression for the error variance in tracking the extremum. The total error, including variations due to the test signal, was then minimized to give rules of thumb for choosing the test signal amplitude.

Eveleigh(1963) considered the problem of automatic regulator adjustment for a linear system, but the results apply to more general extremum control systems. The adaptive loop was approximated as a known nonlinearity followed by linear dynamics. The method of describing functions was used to predict the gain limit for the appearance of oscillations and their frequency. The same type of results were also obtained by Frey et al(1966). Moreover, their theoretical

$$
\begin{aligned}
I_\alpha &= k \cdot I_\beta \\
V_\beta &= k \cdot V_\alpha
\end{aligned}
\tag{3.1}
$$

Then $P_\alpha = P_\beta$. This is a special application, but the same technique could be used for other systems where a product of two related measurable factors is to be optimized.

Stepping methods

Consider the static system

$$y = f(u) + e \tag{3.2}$$

where $f(\cdot)$ has a single local maximum but is otherwise arbitrary. To begin with, assume that the disturbance $e=0$. For this system, the input u should be adjusted to give maximal output y. This can be achieved by stepwise changes of u according to the algorithm

$$\Delta u_{n+1} = \Delta u_n \operatorname{sign}(\Delta y_n) \tag{3.3}$$

The closed-loop system will then end up with the input oscillating a few steps around the maximum.

There are two design parameters to choose in such a system, the stepping period and the steplength Δu_n. A large steplength is desired in order to find the maximum quickly, but on the other hand this will imply a large loss in the steady state because of large deviations from the optimum. A variable steplength might then be useful. This will however complicate the algorithm, and it is not selfevident what criterion to use for the changes of steplength.

arbitrary initial conditions better. It is not clear how these systems can cope with higher order dynamics, non-quadratic nonlinearities or time-variations.

When aiming at an extremum point it seems natural to try making the derivative of the output zero. This leads to using the derivative to determine when to reverse the sweeping direction. Leonov(1969b) investigated such a method for two systems with first order dynamics before and after the nonlinearity respectively. A threshold was introduced, so that switching did not occur until the derivative was less than $-\Delta$ after passage of the maximum. For the case of output dynamics it was found best to put $\Delta=0$, but with input dynamics Δ should be a small positive number.

A somewhat different technique was described by Boehringer(1968). His problem was to get maximum power from a solar cell on board a satellite. The current/voltage characteristic will change with the distance to the sun, and may look as in figure 4. The extracted power can then be maximized using a continuous sweep method, where the current and voltage are decreased alternatively so that

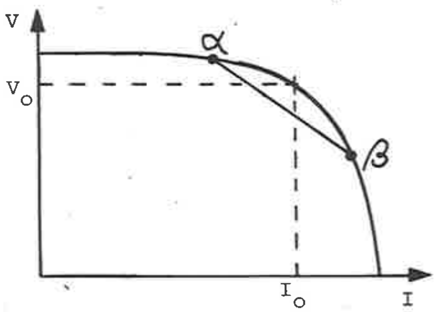

Fig.4 - Current/voltage-characteristic for a solar cell.

results were in good agreement with the results of practical experiments on a gas furnace, maximizing dioxide contents of the flue gas.

Jacobs/Shering(1968) considered a special system with a quadratic static nonlinearity followed by first order linear dynamics. Both input and output were corrupted by noise. The design parameters to determine were the amplitude and frequency for the sinusoidal perturbation signal and the gain in the input driving loop. A loss function was defined as the average deviation of the nonlinearity output from its minimum. The stability properties and the loss were examined for different choices of the design parameters, and a design procedure was given for choosing these parameters.

Although many successful practical applications of the perturbation method have been reported, some potential problems deserve mentioning. As already stated, dynamical effects can be handled by introducing a suitable phase lag to the test signal before correlation with the output. This method works well for systems with a static nonlinearity in series with linear dynamics. But for a general nonlinear system, the phase lag will depend on the point of linearization. The dynamics may then cause more serious trouble.

Another potential problem is that the signal from the correlating device is usually not the exact slope of the nonlinearity. This implies that the optimum will in general not be found exactly if the nonlinearity is not symmetric. However, the deviation will decrease with decreasing perturbation amplitude. This problem can thus be avoided by choosing a sufficiently small perturbation signal at the price of longer correlation times if noise is present.

3.2 Switching Methods

Another basic idea for extremum control is the following. The input is driven at constant speed in the same direction until no further improvement is registered. The drift direction is then reversed. Different algorithms of this type can be described in terms of their specific conditions for altering the direction of input changes. The control law is thus a set of switching conditions. This principle can be mechanized in two ways. The input may be changed continuously or in discrete steps. The second method seems to be quite popular in the Russian literature. Such systems will be called stepping systems.

Continuous sweep

The paper by Tsien/Serdengecti(1955) is a good reference on the continuous sweep method. They consider a static, quadratic nonlinearity with first order dynamics at both input and output. The sweep direction is reversed when the output has decreased from its maximum value by a fixed amount Δ. The design parameters are then the sweep rate and the value of Δ. Tsien/Serdengecti gave design charts and formulae for the input, the output and the so called hunting loss for different values of the design parameters and system time constants. A large portion of the monograph by Draper/Li(1951) is also devoted to an analysis of the continuous sweep method in the presence of dynamics.

If the output is disturbed by noise the above method may give excessive switching unless the value of Δ is sufficiently increased. This higher Δ-value will on the other hand increase the hunting loss. It is thus necessary to compromize in choosing Δ. Filtering is another possibility for reducing the noise sensitivity. The problem is then that more dynamics is introduced into the system,

and the hunting loss will again increase.

Modifications

Unnecessary switching may also be caused by input dynamics. Consider e.g. a maximum-seeking system. After the maximum is passed and the input has been reversed, the input to the nonlinearity will continue to increase for a while due to the input dynamics. The output value at the instant of switching is then taken as the new maximum value, and with large enough dynamic lag this will cause the extra switching. As suggested by Fujii/Kanda(1963) this phenomenon is avoided by waiting for a while before starting to find the new maximum value.

The switching conditions may be chosen in many ways. The output may e.g. be measured only at discrete instants. The difference between successive measurements can then be used as an indicator. This was tried by Putsillo et al(1960) for the control of fuel consumption in a tunnel furnace. Two methods using such differences were analysed by Leonov(1969a). It was found advantageous to keep the input constant for a short while before each reversal of direction.

Several authors have suggested methods relying on diferentiation of the output. Naturally, noise will then be a severe problem that has to be handled by proper filtering. Perret/Rouxel(1963) consider a static quadratic nonlinearity with a time delay followed by first order dynamics. Phase-plane trajectories are calculated for each of the two directions of input drift. From these, switching conditions are derived which employ the second derivative of the output. This algorithm was applied to the maximization of produced reactive power in an alternator. Hamza(1966) describes a very similar method which is claimed to handle